Bluetooth Based Collaborative Crowd Density Estimation with Mobile Phones

Jens Weppner, Paul Lukowicz
Embedded Intelligence
University of Kaiserslautern, German Research Center for Artificial Intelligence (DFKI)
Kaiserslautern, Germany
{*jens.weppner, paul.lukowicz*}@*dfki.de*

***Abstract*—We present a technique for estimating crowd density by using a mobile phone to scan the environment for Bluetooth devices. The paper builds on previous work directed to use Bluetooth scans to analyze social context and extends it with more advanced features, leveraging collaboration between close by devices, and the use of relative features that do not directly depend on the absolute number of devices in the environment. The method is evaluated on a data set from an experiment at the public viewing event in Kaiserslautern during the European soccer championship showing over 75% recognition accuracy on seven discrete classes.**

***Keywords*-Bluetooth scan based sensing; crowd sensing; collaborative sensing; participatory sensing**

I. Introduction

Knowing the density of a crowd can be relevant for a number of applications. Examples range from crowd control and emergency services through urban planning to consumer applications recommending where to go out (based where many other people have also gone to). While in some applications dedicated infrastructure such as access control gates or CCTV cameras [1] may be used, in others it would be desirable to be able to estimate crowd density without pre-installed infrastructure. One possibility is to recruit enough users to be able to estimate the density from the number of devices which report from being in the relevant area. The obvious disadvantage of this method is that a significant number of users must be recruited, which is not always possible. In this paper we present an alternative method that requires only few users moving through the environment with their mobiles scanning for discoverable Bluetooth devices.

II. Related Work and Paper Contributions

A. Related Work

The work most similar to ours is by Nicolai et al. [2] where the discovery time of Bluetooth devices as well as the relation between the number of people and the number of discoverable Bluetooth devices was investigated. As opposed to our approach the work relied on static Bluetooth sensing locations and only the absolute number of discovered Bluetooth devices was used. Along the same lines Morrison et al. [3] investigated crowd density estimation in stadium-based sporting events. However they did not attempt rigorous automatic classification and focused on a visualization tool for Bluetooth logs. Another use case of Bluetooth scanning is described in [4] by Kostakos. They recorded passenger journeys in public transportation by analyzing Bluetooth fingerprints. O'Neill et al. [5] presented initial findings in Bluetooth presence and Bluetooth naming practices. Table I shows an overview about different existing Bluetooth scanning approaches.

Slightly distant from our work, Eagle et al. and Hermersdorf et al. showed [6], [7] how to recognize social patterns in daily user activity, infer relationships and identify socially significant locations from using Bluetooth scans. Versichele et al. [8] performed an experiment at a mass event where they covered an area with static Bluetooth scanning devices to extract statistics and visitor profiles. BLIP Systems [9] exploited a stationary Bluetooth based people tracking system. Based on multiple Bluetooth zones scenarios like queue length at airports or travel times by car are indicated.

Campbell et al. [10] and Burke et al. [11] introduced the general concept of people-centric sensing and participatory sensing. Wirz et al. [12] demonstrated the specific need for detecting potentially critical crowd situations at an early stage during city-wide mass gatherings. They collected GPS traces to create a crowd condition visualization which was monitored by the city police.

B. Contributions

In this paper we present a Bluetooth scan based methodology which can detect different discrete crowd densities. The main contributions beyond the related work above are as follows:

1) We rely not just on the number of devices seen by a Bluetooth scan, but also take information about the link structure between actively scanning Bluetooth devices, ratio of discovered devices in the current scan window to previous scan windows, team-wise diversity of discovered devices, number of semi-continuous device visibility periods, and device visibility durations into account.

Table I: **Bluetooth crowd sensing strategies in related work**

Bluetooth Measurement Setup	Description
Gate flow measurements (cumulative people counting)	A Bluetooth scanning device is positioned stationary near a narrow entry/gate/turnstile/security check point etc. This is mostly accomplished with fixed Bluetooth scanning hardware. The number of discovered Bluetooth devices is summed up over time to make a statement about the current crowd density in a bounded area. Instantaneous crowd density evaluations are not possible. [5]
Queue measurements (waiting time)	A Bluetooth scanning device is positioned near a queue or waiting area (i.e. supermarket check-out/airport check-in/public transport etc.) to measure the time between the appearance and disappearance of unique Bluetooth devices to estimate the current waiting time. This is useful for estimating a relatively small group of locally bounded stationary people. This approach does not consider dynamic people who are not bound to a single location. [9], [13]
Checkpoint measurements (people flow detection)	Two or more specific Bluetooth scanning devices are fixed at two or more separate places with a well known distance to each other. Time is measured between a discovery of a unique device at place 1 and place 2. This approach is similar to queue length estimation but works in a more widespread area. Requirement of this approach is that people are walking from place 1 to place 2 and that the crowd density has a direct relationship with the time needed between both places. [8]
Solitary stationary measurements (instantaneous people counting)	One Bluetooth device is located at a specific location (i.e. shop etc.) with limited dimensions. The instantaneous number of discoverable Bluetooth devices in the covered area is mapped to a number of people by assuming a fixed proportion between the number of discoverable Bluetooth devices and the number of people. [2]
Collaborative and in combination of stationary and dynamic measurements (crowd density)	This novel approach is considered in this paper. With multiple people equipped with ubiquitous smartphones we present new features to evaluate the crowd-density instantaneously. The collaboration allows a coverage of a larger area and a crowd-density estimation with features which are even independent of the proportion between discoverable Bluetooth devices and the number of people.

2) We propose a method to combine the sensor information from several mobile phones carried by different stationary and dynamic close by users (only 0.2% of all people are equipped with a Bluetooth scanning mobile phone) to determine the crowd density in an area of $2500m^2$.
3) We propose "relative" features based on the ratio between values observed by different devices, rather than on the absolute number of Bluetooth IDs seen during a scan. This makes the system more robust against variations in the number of discoverable devices that may result from the background of the people in the crowd rather than the crowd density.

We evaluate the method on a data set recorded during 3 days at the European soccer championship public viewing event in Kaiserslautern which is attended by thousands of visitors. Looking at seven discrete densities that cover the range from a nearly empty space (around 0.01 people per m^2) to dense crowd (above 2.0 people per m^2) we demonstrate recognition rates of over 75% using both relative and absolute features. This is over 30% better than the simple approach from previous work that relies on the number of devices found solely.

III. Approach

A. Background

The foundation of our Bluetooth based crowd density sensing technique is based on the general observation that many people have the Bluetooth transceivers of their mobile phone in the discoverable mode as default setting. This is illustrated in Table II and Figure 1 on data sets from 5 different locations and venues across Europe: (1) several soccer games from the German first and second division, collected in and around the stadium, (2) the world-famous Munich Oktoberfest beer festival, (3) the England-France soccer game at Wembley Stadium in November 2010, (4) a music festival in Malta, and (5) the 3-day European championship public viewing event in Kaiserslautern.

From the above only the Munich Oktoberfest beer festival and the public viewing event in Kaiserslautern data was collected explicitly for crowd density estimation and thus contains crowd density ground truth that is used for the quantitative evaluation later in the paper. During the Munich Oktoberfest experiment we only had a small number of people walking synchronously back and forth on the event's main street. Regarding the Octoberfest experiment we only can utilize a subset of the features presented in this paper because of the lack of information of the bi-directional link structure between actively scanning Bluetooth devices. The Kaiserslautern public viewing experiment gives us a complex data set with asynchronously walking or standing people and all feature calculation requirements to demonstrate our approach.

The other data sets were collected for different purposes, such as inertial navigation and activity recognition. However all data sets include regular Bluetooth scans collected over periods of days by several volunteers walking through the area of the specific event during times of different crowd density. It can be seen that the median of the number of devices discovered per scan is between 8 and 13 with thousands of distinct devices having been recognized over the course of each experiment. Figure 1 shows that only less than 10% of the scans returned no discoverable devices and up to 50 devices were seen when in dense crowd.

We observed that most *discoverable* Bluetooth devices are smartphones and cell phones mostly manufactured by Samsung, Nokia and Sony Ericsson. See table III and table IV for complete listings.

Table II: Statistics about performed Bluetooth crowd-density experiments. This publication covers the public viewing event in Kaiserslautern during the European soccer championship.

Event	Duration	Participants	Number of Bluetooth scans	Average devices per scan	Median devices per scan	Discriminative devices
Kaiserslautern public viewing event (DE)	3 days	10	4100	5.84	6	410
Munich Oktoberfest (DE)	3 days	3	2775	13.35	13	4454
Malta open-air festival (MT)	3 days	12	5500	8.70	8	1088
Wembley Stadium (UK)	1 day	6	4958	15.44	10	2509
Allianz Arena Soccer (DE)	4 day	10	14087	10.87	8	3944

Table III: Types of discovered Bluetooth devices

Bluetooth major device class	Percentage
Smartphone	72%
Mobile phone	28%
Laptop	0.2%
Cordless phone	0.02%
Audio/Video headset	0.01%
Other	0.04%

Table IV: Bluetooth device manufacturers

Bluetooth device manufacturer	Percentage
Samsung	29%
Nokia	32%
Sony Ericsson	12%
RIM	7%
LG	7%
Texas Instruments	3%
HTC	1%
Other	8%

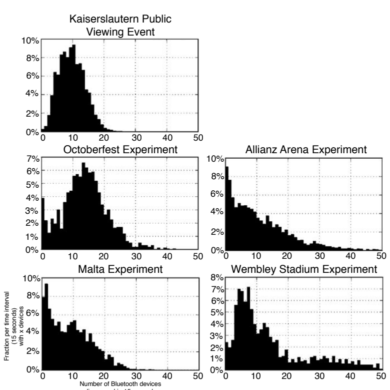

Figure 1: Distribution showing the fraction of the number of Bluetooth devices discovered in a 15 second time window at multiple experiment venues.

B. General Considerations

An obvious way to estimate crowd density is to perform a scan for discoverable devices and assume that the number of devices it returns is an indication of the number of people in the vicinity defined by Bluetooth range (typically around 10m). Unfortunately, this simple approach contains a number of problems.

Firstly, there is the issue of sufficient statistics. With the scan limited to a radius of about 10m (approximately a circle with $300m^2$ area) anything between a few and a few hundred people can be within range. While in a dense crowd with a few hundred people we may get a representative sample, in less crowded areas we are likely to see very strong variations between samples. Assuming the probability of any single user having a discoverable Bluetooth device to be 10% the probability that no device is seen when 20 people are within range is $0.9^{20} = 0.12$. Thus we may sometimes be in a group of people who do not even have activated mobile phones while at other times we may be surrounded by a group where everyone has an active Bluetooth device.

Secondly, there is the question of signal attenuation. At 2.4GHz (which is the transmission frequency of Bluetooth) the human body has a high absorption coefficient. This means that in a dense crowd (where we would expect to have good statistics) the effective scan range is reduced and therefore "falsifying" the results.

Finally, we have to consider cultural factors. This means that the average number of people carrying a discoverable Bluetooth device may significantly vary depending on who the persons in the crowd are. For the same crowd density at a student party of a technical university a different number of devices may be present than at a fifth division soccer game in a poor rural area.

To mitigate the influences above our method does not rely solely on the absolute number of discovered devices. Instead we also use the average signal strength and signal strength variations. In addition, we look at collaborative estimation from several (up to around 10) devices. In doing so, we focus on differential features that are not directly dependent on the absolute number of discoverable devices in the environment or the absolute signal strength. As shown in the evaluation section IV the above measures lead to over 30% improvement in recognition rate over a method based on the absolute number of discovered devices.

C. Features

We are calculating our features based on multiple partially distributed sensors because we want to achieve a statement of the crowd density of the whole event area as we assume all sensors together are covering a large portion of the area during movement in the area.

1) Feature: Averaged sum of distinct devices discovered by all sensors in scan window: This simple feature describes the current number of discovered distinct devices for every

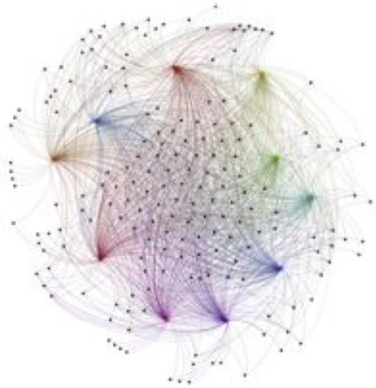

Figure 2: Bluetooth link structure graph showing all 10 actively scanning smartphones (colored) and discovered devices (black nodes) during the experiment.

snap-shot (a *Bluetooth scan window* is hereafter also referred to as a snap-shot) of the experiment. For each snap-shot each of the sensors delivers a set of unique devices identified by the unique Bluetooth MAC address.

Calculating the union of all discovered devices (by all sensors) divided by the number of sensors results in this feature. Bluetooth devices discovered by multiple sensors at the same time are not influencing this feature.

This feature relies directly on the level of distribution of the sensors over the event area. Since the Bluetooth range is very limited a sensor distribution over a larger area obviously leads to a larger number of distinct devices and the other way around.

The downside of this feature is its direct relation to the ratio of discoverable Bluetooth devices to the number of people to be detected.

See figure 3 for a visualization of the feature during one experiment.

2) Feature: Ratio bi-directional link structure of sensors to average pairwise distance of sensors multiplied with average sensor speed: This composite feature characterizes the context of the snap-shot of the collaborative sensor data more explicit.

The feature takes the bi-directional Bluetooth link structure between the sensors (actively scanning *Bluetooth devices* are hereafter also referred to as sensors) into account. A directional link (hereafter also referred to as sensor discovery) between a pair of sensor 'a' and 'b' is defined as established when sensor 'a' discovers sensor 'b' in the current snap-shot. Another directional link is defined as established when sensor 'b' discovers sensor 'a'. This implies a pair of sensors might both link to each other or one sensor links to the other or none of both sensors links to the other. All combinations of pairs between sensors are monitored. Maximum established links between ten sensors would be 90 links, the minimum number of established links would be zero. The bi-directional link structure is defined as the sum of all links between all sensors.

The average pairwise distance calculation is based on the GPS sensor data information. A snap-shot contains multiple GPS locations per sensor (GPS location is sampled at 1Hz). Only locations with a GPS accuracy better than 15m are taken into account. Based on the filtered locations we calculate the average since most promising location of the sensor. For each pair of sensors we calculate the distance between them. There are $n!/2/(n-2)!$ distances calculated per snap-shot where n is the number of sensors. The distance between all sensor pairs is then averaged. The average speed is calculated for every snap-shot and each individual sensor based on averaged GPS information per snap-shot each with an accuracy of better than 15m. Afterwards the average speed is calculated for all sensors. Finally, the feature is calculated by the number of bi-directional links divided by the average pairwise distance of sensors multiplied with the average sensor speed. It is important to mention that this feature is completely independent of external (others than the used Bluetooth sensors) discoverable Bluetooth devices. It uses the relationship between the number of links to the distance between sensors, based on the assumption that a more dense crowd shields the sensor links heavier than in a low dense crowd with the same underlying distance.

See figure 4 on page 5 for a visualization of the feature during one experiment.

3) Feature: Ratio of discovered devices in current snap-shot to discovered devices in last x minutes: This feature characterizes the crowd movement during the snap-shot of the collaborative sensor data more explicit.

Newly detected Bluetooth devices in a snap-shot are defined as a set of all unique devices discovered during the snap-shot by all sensors. Calculating the union of unique discovered devices by all sensors in a snap-shot leads to the collaborative measurement. The second part of the ratio is the size of the set of unique Bluetooth devices discovered during previous 15 snap-shots (depending on the size of the snap-shot, this signifies a monitoring of the previous 1 to 10 minutes). Finally, the feature is calculated by the size of the collaborative set of discovered devices at the snap-shot divided by the size of the collaborative set of Bluetooth devices discovered before. This implies that the value is smaller in a less moving crowd than in a more likely moving crowd. This is caused by the fact that the number of different devices seen during x snap-shots is smaller if there is less movement (less devices are rushing by) than for strong crowd movement (high crowd flow).

See figure 5 for a visualization of the feature during one experiment.

4) Feature: Average team-wise *diversity of discovered devices per scan window (ratio not concurrent devices to concurrent devices):* We define a team by two persons staying in close adjacency while each person is carrying a sensor. A team can either move dynamically or be stationary, but continuously stays together.

This feature takes into account the *team-wise* diversity of discovered Bluetooth devices for each snap-shot. In this context we define diversity as the ratio between the number of Bluetooth devices not concurrently discovered and concurrently discovered devices. Not concurrently discovered devices are defined in set theory as the symmetric difference. Either sensor 'a' or sensor 'b' but not both sensors discover the same device in a snapshot. Concurrent devices appear both in the current snap-shot of sensor 'a' and sensor 'b'. The ratio is averaged for all teams in each snap-shot for a collaborative measurement.

This feature calculates the diversity of discovered devices between two sensors which are close to each other. This gives us a value depending on the crowd between and around the team as well as the unambiguousness of the two sensor measures.

See figure 6 on page 5 for a visualization of the feature during one experiment.

5) Feature: Average number of semi-continuous unique device visibility periods (finite state machine approach): We define a semi-continuous device visibility period per sensor as the number of consecutive snapshots, whereas in each snapshot a unique device is discovered with the exception of very short vanishings during the period. A *short vanishing* is defined as a single snap-shot without a discovery among other snap-shots including the presence of a specific device. Multiple short vanishings may appear during a semi-continuous device visibility period. The period ends when a device vanishes at least for two consecutive snap-shots. The same unique device then may reappear again or vanish for a longer time or forever.

The data is further processed by calculating the sum of present semi-continuous unique device visibility periods during a snap-shot. By definition, the sum of unique devices might include a device which is not seen in the current snap-shot. We implemented the calculation of this feature by a finite state machine for each unique device (d) and for each sensor (s). Resulting in $d * s$ finite state machines. Finally, the collaborative overall average value is calculated per snap-shot over all sensors. The feature value can be high for a small number of discoverable devices which are in range for a longer time. The value can be low for a high number of discoverable devices which are in range for a shorter time.

See figure 7 on page 6 for a visualization of the feature during one experiment.

6) Feature: Average durations of semi-continuous unique device visibility periods (finite state machine approach): This feature is based on semi-continuous device visibility similar to to feature III-C5 but calculates the duration. Therefore the pre-processing is similar to feature III-C5.

The duration of a semi-continuous visibility of a unique device is defined as the number of sequential snap-shots where a specific device is seen. This duration factors into all snap-shots that the semi-continuous visibility is covering.

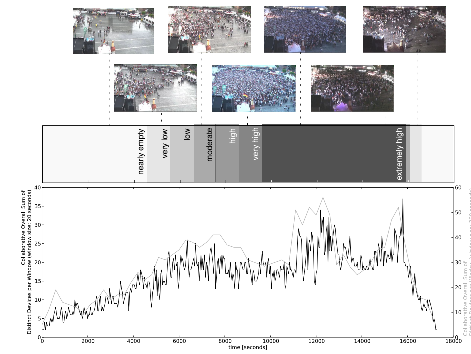

Figure 3: Feature: 'Size of device set of all distinct discovered devices by all sensors in time frame'. Overview of crowd density levels shown by different background grey levels.

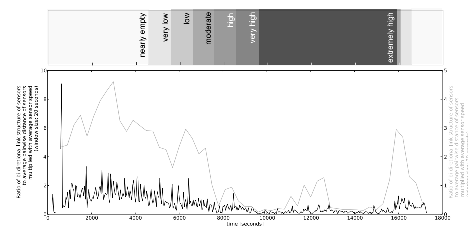

Figure 4: Feature: 'Ratio bi-directional link structure of sensors to average pairwise distance of sensors multiplied with average sensor speed'

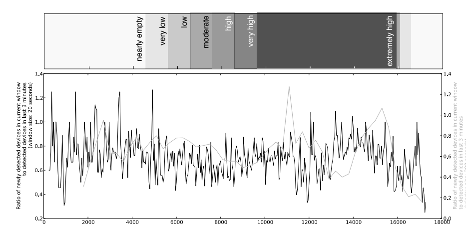

Figure 5: Feature visualization of 'Ratio of discovered devices in current snap-shot to discovered devices in last x minutes'

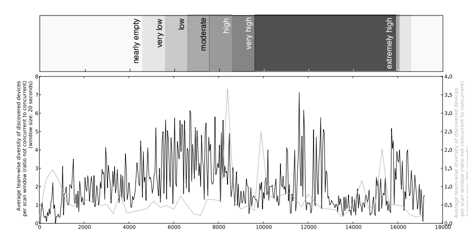

Figure 6: Feature visualization of 'Average team-wise diversity of discovered devices per scan window (ratio *not* concurrent to concurrent)'

Averaging the duration for one sensor of all device visibility durations at one snap-shot is the value per sensor. Averaging this value of all sensors per snap-shot results in the value of this feature.

See figure 8 for a visualization of the feature during one experiment.

(a) nearly empty 0.01-0.05 $people/m^2$ (b) very low 0.05-0.2 $people/m^2$ (c) low 0.2-0.3 $people/m^2$ (d) moderate 0.3-0.4 $people/m^2$ (e) high 0.4-1.0 $people/m^2$ (f) very high 1.0-2.0 $people/m^2$ (g) extremely high 2.0++ $people/m^2$

Figure 10: Crowd density classes ranging from *nearly empty* to *extremelyhigh*. Excerpts of the HD ground truth video.

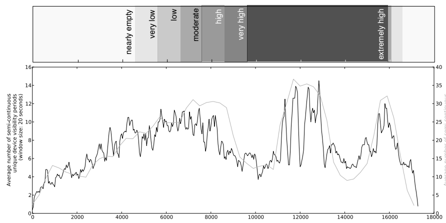

Figure 7: Feature visualization of 'Average number of semi-continuous unique device visibility periods'

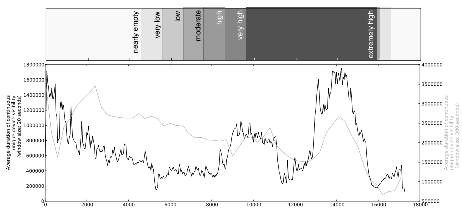

Figure 8: Feature visualization of 'Average durations of semi-continuous unique device visibility periods (finite state machine approach)'

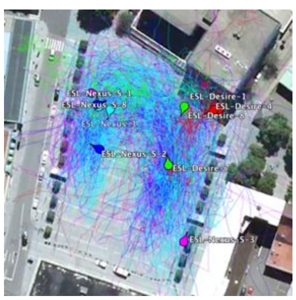

Figure 9: Satellite image of the event area with GPS traces of the Bluetooth discoverers.

IV. Evaluation

A. Experimental Environment

To evaluate our features we set up three experiments on three days during the European soccer championship 2012 at the official public viewing event at the town-center (marketplace called 'Stiftsplatz') of Kaiserslautern (Germany). The evaluated experiment area has a dimension of 48.5 to 48.5 meters allowing up to 5200 people to enter the fenced area.

Each experiment had a duration of about 4 hours consisting of 2 hours before the soccer championship kick-off began, 45 minutes during the first half of the soccer match, 15 minutes during the half time break, 45 minutes during the second half of the soccer match, and 20 minutes after the game.

We started our experiment early before spectators began entering the event area. During two hours the area was then filled up to a level where no more people where allowed to enter the area by the event organization for safety reasons.

We gathered sensor data of different crowd densities including levels *nearly empty* $(0.01-0.05people/m^2)$, *very low* $(0.05-0.2people/m^2)$, *low* $(0.2-0.3people/m^2)$, *moderate* $(0.3-0.4people/m^2)$, *high* $(0.4-1.0people/m^2)$, *very high* $(1.0-2.0people/m^2)$, *extremely high* $(2.0+people/m^2)$. See figure 11 for the complete course of the crowd density levels during the experiment and figure 10 with excerpts from the ground truth video for each crowd density class.

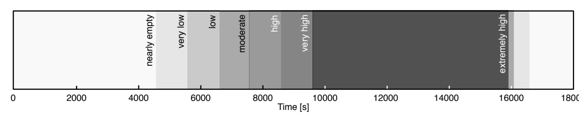

Figure 11: Crowd density ground truth information

The crowd flow was moderate during the filling phase of the event since attendees went slowly from the event entry towards the big screen in the opposite corner. There was no crowd flow during the first and second half of the soccer match. During the break the crowd flow range was between moderate and high. Beyond the end of the soccer match the crowd flow was very high since the attendees wanted to leave the event area through multiple exits as fast as possible because the German soccer team had lost the match. See figure 12 for the complete course of the crowd flow. Our presented crowd density measurement technique is robust to differing crowd flow levels as they are not considered in the labeling procedure of the feature vectors and are not correlated to the crowd density levels.

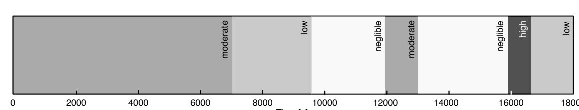

Figure 12: Crowd flow ground truth information

B. Experimental Setup

For each experiment we recruited 10 students. We divided the students into 5 teams with 2 students each. Team members always stayed in close contact (up to 1 meter distance) to each other. Teams were instructed to be either stationary (2 teams, 4 students) or dynamic (3 teams, 6 students).

Stationary is defined as continuously standing on the spot. We placed stationary teams near the entrance of the area. One team near the left side and the other team on the right side of the entry.

Dynamic is defined as continuously walking around on the event area. Teams were told to walk on a curved path covering 3 sides of the event area and mostly covering the edge regions (see below for exceptions) of the crowd since those regions were common to walk on because nearby food and beverage stands.

The idea behind the stationary and dynamic scripted setup is to represent a natural behavior of people during such events. Some people are standing still watching the performance. Other people are walking around to food/beverage stands, meet friends, change to a better viewing spot etc.

Multiple dynamic teams are allowed to walk asynchronously. The walking speed of dynamic teams is not scripted, allowing to choose the personal optimum walking speed (we can determine the walking speed by evaluating our GPS log information). We allowed teams to move to a self determined place in the middle of the crowd during the first and second half of the soccer match excluding the break.

In a real-life scenario people do not have to be categorized to behave stationary or dynamically continuously. Smartphone sensor information allows to dynamically detect the type of behavior. Because of this random natural behavior we do not manually apply any information to our algorithms about stationary or dynamic behavior.

Each student was equipped with one Android smartphone which is placed in the trousers pocket.

We deployed Android smartphones of different types including HTC Desire, Google Nexus and Samsung Nexus S each based on the most recent version of the Android operating system. On all devices we were running our custom Android application called ContextRobot which records multiple sensor data streams onto the microSD card for later off-line analysis. Our Android applications continuously scanned for discoverable Bluetooth devices (a Bluetooth scan is defined as a time interval which emits a set of unique Bluetooth device with the restriction of unrepeated occurrences of a unique device). A single log entry of a device discovery during a Bluetooth scan is associated with a timestamp, a serial Bluetooth scan interval number, personal Bluetooth device name, unique Bluetooth id (Bluetooth MAC address), and the Bluetooth signal strength as a RSSI (received signal strength indication) value.

An exact temporal begin and end of a Bluetooth scanning interval cannot be specified during the recording of Bluetooth sensor data since the operating systems restricts to certain length of scan intervals depending on internal thresholds. The average duration of a Bluetooth scan interval is about ten seconds (with little variations). Our application triggers a new Bluetooth scan when the previous scan has ended. Multiple collaborative Android devices recording Bluetooth scans intervals are synchronized in an off-line manner. At a given time window of a length of 20 seconds we determine one scan interval which fits this window entirely.

In addition to Bluetooth sensor information we record location information by the GPS sensor at a frequency of $1Hz$. Our Android application continuously records timestamp, latitude, longitude and accuracy onto the microSD cards. Figure 9 visualizes the walking traces of the dynamic team members. Location information is required for some feature computations which rely on distances between multiple students and their walking speed.

For obtaining ground truth data about the crowd density we set up a HD video camera on top of a neighbored hotel building with view of the whole event area. Figure 10 shows excerpts of the video footage for different crowd density classes. The ground truth labels are based on the video footage which we labeled every 10-15 minutes with a crowd density class ranging from nearly empty to extremely high.

C. Results

We analyzed the collaborative features for crowd density classification with a granularity of 40 seconds to achieve a statement of the overall crowd density at the event area. Nevertheless, some feature calculations are inspecting a few preceding sensor values, but we do not apply any additional filtering by applying a sliding window to the data.

We trained a decision tree classifier with all features and evaluated it with 10-fold cross-validation. We achieve an accuracy of 75.3% for estimating the correct crowd density class on seven discrete crowd density classes. Figure 13 shows the confusion matrix of the classification results. The majority of the predictions are distributed along the diagonal. False classifications are prevailingly located near the diagonal implying a classification in an adjacent crowd density class.

Even without features relying on the absolute number (without feature III-C1 and III-C5) of discoverable Bluetooth devices we achieve a classification accuracy of 67.8%. See figure 14 for the confusion matrix of the classification results.

V. Conclusion and Future Work

We have shown how Bluetooth scan data from just a few users equipped with standard mobile phones can be used to estimate crowd density. The core of the method is the comparison and fusion of data from different devices which leads to over 30% improvement in accuracy over a simple single device approach. The just over 75% accuracy

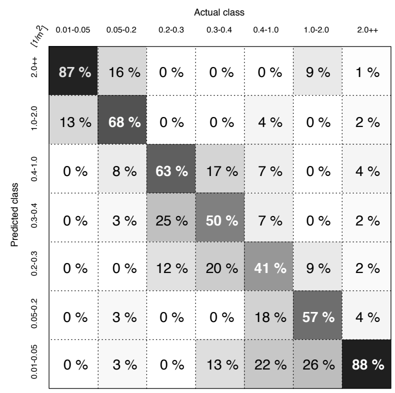

Figure 13: Confusion matrix with six features. *With* features depending on number of discoverable Bluetooth devices.

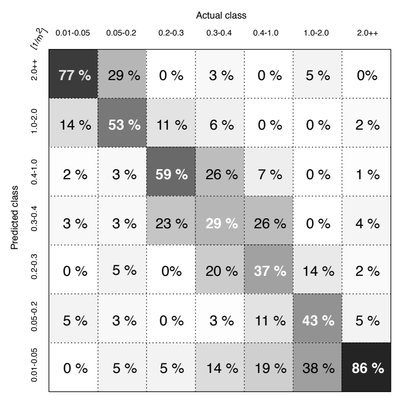

Figure 14: Confusion matrix with four features. *Without* features depending on absolute number of discoverable Bluetooth devices.

on seven classes must be seen in the context of noisy ground truth resulting from arbitrary class definition, extrapolation between the ground-truth based crowd density extraction every 10 to 15 minutes and inaccuracies in the counting process. In addition, confusions occur nearly exclusively between neighboring classes (see figure 13 and figure 14). Note that the experimental data did not include the ”totally empty space” class which can be trivially recognized from the near absence of Bluetooth devices and could be easily integrated into the system.

In summary, we believe that the method presented in this paper is potentially useful for many applications. To further improve this method future work will focus on better understanding and modeling the relative features. We will also collect data from other events and countries to verify the hypothesis that the relative features are robust against culture related differences in the percentage of people carrying discoverable Bluetooth devices.

ACKNOWLEDGMENT

This work is supported by the FP7 ICT Future Enabling Technologies programme of the European Commission under the grant agreement No. 231288 (SOCIONICAL).

REFERENCES

[1] H. Su, H. Yang, and S. Zheng, “The large-scale crowd density estimation based on effective region feature extraction method,” in *Proceedings of the 10th Asian conference on Computer vision - Volume Part III*, ser. ACCV’10. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 302–313.

[2] T. Nicolai and H. Kenn, “About the relationship between people and discoverable bluetooth devices in urban environments,” in *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology*. ACM, 2007, pp. 72–78.

[3] A. Morrison, M. Bell, and M. Chalmers, “Visualisation of spectator activity at stadium events,” in *2009 13th International Conference Information Visualisation*. IEEE, 2009, pp. 219–226.

[4] V. Kostakos, “Using bluetooth to capture passenger trips on public transport buses,” *arXiv*, vol. 806, 2008.

[5] E. O’Neill, V. Kostakos, T. Kindberg, A. Schiek, A. Penn, D. Fraser, and T. Jones, “Instrumenting the city: Developing methods for observing and understanding the digital cityscape,” *UbiComp 2006: Ubiquitous Computing*, pp. 315–332, 2006.

[6] N. Eagle and A. (Sandy) Pentland, “Reality mining: sensing complex social systems,” *Personal Ubiquitous Comput.*, vol. 10, pp. 255–268, March 2006.

[7] M. Hermersdorf, H. Nyholm, J. Salminen, H. Tirri, J. Perkiö, and V. Tuulos, “Sensing in rich bluetooth environments,” in *In Proceedings of WSW’06 at SenSys’06. SenSys’06*, 2006.

[8] M. Versichele, T. Neutens, M. Delafontaine, and N. V. de Weghe, “The use of bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the ghent festivities,” *Applied Geography*, vol. 32, no. 2, pp. 208 – 220, 2012.

[9] (2013, 10) Blip systems a/s. [Online]. Available: http://www.blipsystems.com/

[10] A. Campbell, S. Eisenman, N. Lane, E. Miluzzo, R. Peterson, H. Lu, X. Zheng, M. Musolesi, K. Fodor, and G.-S. Ahn, “The rise of people-centric sensing,” *Internet Computing, IEEE*, vol. 12, no. 4, pp. 12 –21, july-aug. 2008.

[11] J. A. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, “Participatory sensing,” 2006.

[12] M. Wirz, T. Franke, D. Roggen, E. Mitleton-Kelly, P. Lukowicz, and G. Troster, “Inferring crowd conditions from pedestrians’ location traces for real-time crowd monitoring during city-scale mass gatherings,” in *Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2012 IEEE 21st International Workshop on*, june 2012, pp. 367 –372.

[13] V. Kostakos, T. Camacho, and C. Mantero, “Wireless detection of end-to-end passenger trips on public transport buses,” in *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, sept. 2010, pp. 1795 –1800.