CWI-ADE2016 Dataset

Sensing nightclubs through 40 million BLE packets

Sergio Cabrero
Centrum Wiskunde & Informatica
s.cabrero@cwi.nl

Jack Jansen
Centrum Wiskunde & Informatica
j.jansen@cwi.nl

Thomas Röggla
Centrum Wiskunde & Informatica
t.roggla@cwi.nl

John Alexis Guerra-Gomez
Los Andes University
john.guerra@gmail.com

David A. Shamma
Centrum Wiskunde & Informatica
aymans@acm.org

Pablo Cesar
Centrum Wiskunde & Informatica
Delft University of Technology
p.s.cesar@cwi.nl

ABSTRACT

The CWI-ADE2016 Dataset is a collection of more than 40 million Bluetooth Low Energy (BLE) packets and of 14 million accelerometer and temperature samples generated by wristbands that people wore in a nightclub. The data was gathered during Amsterdam Dance Event 2016 in an exclusive club experience curated around human senses, which leveraged technology as a bridge between the club and the guests. Each guest was handed a custom-made wristband with a Bluetooth-enabled device that broadcasted movement, temperature and other sensor readings. A network of Raspberry Pi receivers deployed for the occasion captured BLE broadcast packets both from wristbands and from any other device in the environment. This data provides a full picture of the performance of the real life deployment of a sensing infrastructure and gives insights to designing sensing platforms, analysing network performance, understanding crowds behaviour or studying opportunistic sensing. In this paper, we describe the method used to collect the data and retain privacy. We also provide analysis of various aspects of the dataset and illustrate its use measuring network performance and crowd movement to fuel the multimedia and sensing research community. Interested readers can access the dataset, related code and other assets in https://github.com/cwi-dis/CWI-ADE2016-Dataset.

CCS CONCEPTS

•Networks →Network performance analysis; •Applied computing →Performing arts;

KEYWORDS

Dataset, Crowd, Sensing, BLE, IoT, nightclubs, accelerometer, wearables, activity, location

ACM Reference format:
Sergio Cabrero, Jack Jansen, Thomas Röggla, John Alexis Guerra-Gomez, David A. Shamma, and Pablo Cesar. 2017. CWI-ADE2016 Dataset. In *Proceedings of ACM Multimedia Systems Conference, Taiwan, Taipei (MMSYS'2017),* 7 pages.
DOI: 10.475/123_4

1 INTRODUCTION

Club culture is about getting together and enjoying multisensory experiences with other people. These experiences are curated by the event organizers [1] and each individual average club goer typically has little impact on the experience as a whole. But what if the club could actually react to the level of excitement of the crowd? What if the people could actively influence the overall experience by their activity? Or more generally: what would the club of the future look like? We asked ourselves these questions for a two-day event in which a sensing platform was specifically developed to enhance the experience of over 900 party-goers, held in the context of the *Amsterdam Dance Event*[1] in October 2016. The central component in our approach are custom-made wristbands by the Dutch fashion designer ByBorre[2] that collect a variety of sensor readings from the wearer. This data is broadcast using Bluetooth Low Energy (BLE) Advertisement packets and collected with a network of Raspberry Pis (RPIs) deployed in an empty building transformed into an ad-hoc nightclub. The data was forwarded to a central server where it was stored and processed for different purposes, such as activity recognition, localization of guests, driving a real time data visualization or affecting light and sound of a room within the environment [10].

Datasets gathered during real-life experiments are highly valuable for researchers of different fields. However, larger scale data from sensing human activity in real-life scenarios is not abundant. Some datasets [5] exist, but they mainly represent experiments related to academic activities. The CWI-ADE2016 Dataset aims to improve this situation with the data extracted from over 40 million BLE broadcast packets gathered during two nights from our own and other devices present in the club. Additionally, almost 14 million of these packets include accelerometer and temperature readings from guests' wristbands, which are also published in this release.

This paper describes the data collected, the platform used, the context in which it was developed and the lessons we learned in the process. We describe the two-day event and the infrastructure developed for the occasion. We also provide an overview of the CWI-ADE2016 Dataset from different perspectives and examples of usage. The first example describes some insights that this data provides to understand network performance, either to model it better or to tackle realistic application design in this environment.

[1] https://www.amsterdam-dance-event.nl/
[2] http://www.byborre.com

The second example hypothesises over the use of this data to understand people's movement during the event. The goal is to encourage and inspire other researchers to use this dataset, to enquire us for complementary data that they find necessary and to aid them when tackling similar challenges. The following section describes some knowledge about the event production and how it influenced the design of the data collection system: §3 describes how information is organised in the dataset. §4 and §5 will analyse the contents of the dataset and provide some examples of usage to inspire other researchers. We conclude the paper with §6.

2 DATA COLLECTION

For the data collection, we start by describing the context that defined the requirements. Next, we document the devices used as data sources. Finally, we elaborate on the infrastructure used to receive and process the data.

2.1 Context: two club nights for 900 guests

The system came into existence as part of a collaboration on wearable technology with ByBorre. For a two-day club event with around 900 guests within the context of the annual *Amsterdam Dance Event* held in October 2016 in Amsterdam, we wanted to explore what the club of the future might look like. The selected venue was the first floor of the emblematic Het Bungehuis[3] building in Amsterdam's city center. The core idea was to find ways to learn about the guests' behaviour and try to communicate with the environment with the goal to bring people together and design an experience which would stimulate all the senses at once: Specially created dinner menus, drinks and perfumes, an adaptive sound system and light show with technology playing the role of connecting all the senses into an all-encompassing experience. For this, we evaluated a series of candidate sensor technologies and ways to make a club experience more participatory. Ideally we wanted something compact and unobtrusive, which could be integrated into textiles for people to wear without impacting their experience. We opted for specially designed wristbands with embedded sensors.

Two types of guests where invited to the event: VIPs who enjoyed the party, and FoRB guests who also enjoyed dinner and a welcoming tour. Around 18:00, Amsterdam time, each day, FoRBs were welcomed and given a tour of the space, which included a special cocktail in the Housewarming Bar. At around 18:30, dinner started. VIPs started arriving around 19:00 and the party ended after midnight. All guests collected their wristbands at the reception as they entered the building, but they carried it with them when they left. Figure 1 shows the space, the rooms, the approximate tour itinerary, and the location of our infrastructure that we will describe later in this section.

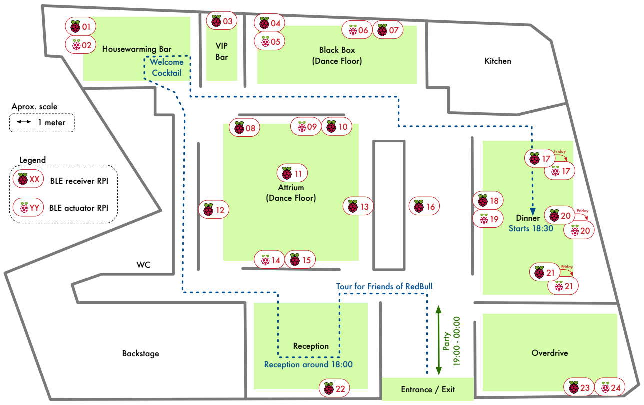

Figure 1: Location of RPIs in the improvised night club.

2.2 Data sources: custom-made wristbands

We fitted guests' wristbands with off-the-shelf, BLE-enabled circuit boards. These boards needed to meet the requirements of being small and having long battery life. We decided to create two different types of bands, as the special programme for FoRBs required some of the bands to be able to provide direct feedback to the wearer in the form of LED lights. Out of the total 900 wristbands that we produced, 800 were fitted with *Estimote Sticker* boards for VIP guests. These coin-sized boards broadcast a UUID, 3-axis accelerometer values and temperature readings using a protocol similar to Apple's *iBeacon* over BLE, i.e. they embed sensor readings in Manufacturer Data BLE advertisements [3]. The second type of wristband for FoRB guests, of which only 100 were made, uses a *SensorTag CC2650* board from *Texas Instruments (TI)*. It is slightly larger than the Estimote board, but it is a more general-purpose board for IoT applications, has more sensors built-in and is fully programmable. We mounted a small strip of RGB LEDs on them. The idea behind this is that the sensor could be *actuated*, i.e. it could flash LEDs in different colours, should some specified event occur. This was used to signal some of the guests that the next part of their special programme was about to begin.

Both devices use BLE to broadcast their sensor readings periodically. We selected BLE because it is present in most commercial devices and it is a mature technology. We also chose broadcast-based communication over other possibilities, such as pairing each wristband with guests' phones, because it requires zero configuration and the guests' anonymity is easy to preserve. In our system, data from the wristbands is received when they are in range of a receiver—and no collisions or interference prevent it—without any further action. This simplicity comes at the cost of potential disadvantages, such as unreliable data delivery or potential interception of data by BLE receivers external to our system.

BLE broadcasting of advertisement packets uses three frequency channels; each packet is sent over the three channels consecutively, unless the device is configured otherwise. Then, the device must wait before advertising again. This advertising frequency can be freely chosen by developers and devices, but must be inside the ranges defined by BLE. In our case, *Estimote Stickers* broadcast two types of packets, one regular *iBeacon* packet approximately every 5 seconds and one *Nearable* packet every 1.25 seconds. This time is doubled, i.e. 2.5 seconds, when the sensor is not in motion. *Nearable* packets contain the sensor readings we are interested in. So we designed our applications around their sampling frequencies and, consequently, we programmed *TI SensorTags* to also broadcast their data every 1.25 seconds. To extend their battery life, we implemented a sleep/awake mechanism. So when asleep, *TI SensorTags*

[3] https://nl.wikipedia.org/wiki/Bungehuis

broadcast one packet every 10 seconds. We woke them up just a couple of hours before they were needed. Both platforms are highly configurable in broadcast frequencies, transmission powers and several other parameters. Unless specifically stated, we used the default factory values.

2.3 Infrastructure: a network of Raspberry Pis

Tracking 450 devices per night inside an improvised club space of about 500 m^2 is challenging because of the amount of devices, the expected density in areas such as the dance floors and the limitations of BLE [2, 4]. The 2.4 GHz frequency band used by BLE is shared with WiFi. Although we avoided setting up networks in this band, there were other teams involved in the production that used it. So our data collection system, and the whole data pipeline supporting it, needed to factor in possible interferences and packet losses. The system also needed to be easy to install and remove, as well as cost effective. For these reasons, we opted for a network of Raspberry Pis (RPIs) that listened to BLE advertisement channels. They were then connected via Ethernet—to diminish interferences—to a central server, which stored the data, in a MongoDB database, and forwarded it to the rest of the systems. This system not only captured and stored BLE packets emitted by wristbands, but also by any BLE devices broadcasting in range of our receivers. Since these packets occupy BLE resources anyway, we captured them to understand network performance and to explore potential correlations between wristbands and other devices.

The approximate location of the RPIs in the space is shown in Figure 1. For completeness, the map also shows the location of the RPIs used as actuators. The role of these devices was to connect to *TI SensorTags* and activate their LED lights. Digging more into the role of, and the system behind, these RPIs is out of the scope of this paper. However, we must mention that four RPIs in the dining room changed their role from Thursday to Friday. Thus, we used 17 receivers on Thursday and 13 on Friday. RPIs were carefully placed to be able to cover the whole space and, whenever possible, to be able to tell sensor location—at least at room level—by knowing which RPI or RPIs received its packets. However, we must admit that calibration was not possible due to the short time available for testing on location and the need of deploying Ethernet wires well in advance.

3 DATASET

There are two distinct files in the dataset: one containing information about BLE packets (*blepackets* file), and one containing information about accelerometer and temperature sensors carried in some of these packets (*sensordata* file), namely those sent by in *Estimote Nearable* boards and *TI SensorTag* boards. Some of the data is completely revealed, e.g. accelerometer or temperature readings, but we chose to not publish or hash part of it, only to the end of avoiding the leaking of personal data. Nevertheless, we see this dataset as a living entity, and we will do further work to expand it and encourage other researchers to share their benchmarks, results, and visualisations.

The *blepackets* file contains more than 40 million records, and *sensordata* contains almost 14 million. We have structured the files using a field as a pointer between them, i.e. *Packet Id.*. In other words, a record in *sensordata* has the same *Packet Id.* as the BLE packet in *blepackets* that carried that sensor data. Table 1 contains a brief description of the fields[4]. For further details into sensor readings, we refer you to the Estimote Sticker[5] and TI Sensortag CC2650[6] documentations.

4 ANALYSIS

Of the 40 million BLE packets in the CWI-ADE2016 dataset, almost 14 million come from 812 *Estimote Stickers* and 109 *TI SensorTag* devices. The rest come from an unknown, but potentially high, number of other BLE chips. The number of devices is slightly higher than the number of wristbands because some of them were used for testing. The amount of wristbands used and packets collected each day is similar, but slightly higher on Friday than on Thursday. We provide analysis of three different aspects of the dataset. First, we present the number of records generated in our database and an estimation of the number of BLE packets that produced them. This is important to understand how records in the dataset were generated, and why they are not equivalent to BLE packets. Second, we present the different types of packets that were captured by the system. This shows the high BLE noise levels encountered during the event and the heterogeneous set of devices detected. Finally, we comment on the number of packets received by each RPI and the number of wristbands seen per minute. This provides interesting insights on the performance of our infrastructure and its design. All these analyses consider data captured only between 16:00 and 00:00 (8 hours) each day, which shows the system during the event and in the immediate preparations before. The dataset contains data of two complete days.

4.1 Packets and copies

Because we have deployed several receivers close to each other, every single BLE packet from a wristband can potentially create several records in the dataset, one for each RPI that received it. Distinguishing between an 'original' packet and its 'copies' is not trivial. Packets are timestamped at the moment when they are processed by our scanning software. So even accurately synchronised receivers can give slightly different timestamps to the same packet if delays are produced in the Bluetooth stack. Identifying copies by having similar timestamps and the same payload is also not completely accurate. BLE can produce three packets in every advertisement interval, which will be very close together in time—in the range of 20 milliseconds—and they are likely to have the same payload, although *Estimote Nearable* packets change it slightly. So packets transmitted over different channels can be easily mistaken as different packets. Being aware of these difficulties and just to help us give an overview of the dataset, we will establish the following criteria for our analysis. We consider that packets from the same source, i.e. the same source MAC Address, that are received with less than 100 milliseconds difference between them, are the same packet, and that the one received first is the original and the rest are copies. Although we are aware that this is not completely

[4] More details in https://github.com/cwi-dis/CWI-ADE2016-Dataset
[5] http://developer.estimote.com/nearables/
[6] http://processors.wiki.ti.com/index.php/SensorTag2015

Table 1: Fields in dataset files

Name	*Type*	*In File*	*Description*
Packet Id.	Integer	blepackets, sensordata	A unique identifier of a packet in the dataset.
Timestamp	Float	blepackets, sensordata	Unix time in seconds of the moment in which the packet was received by our BLE scanning software in the RPI. Add 2 hours to obtain Amsterdam local time.
Type	Integer	blepackets, sensordata	Type of device that transmitted the packet. For wristbands three different packet formats: ‘`estimote-nearable`’, ‘`estimote-iBeacon`’ and ‘`sensortag`’. For packets that contain Manufacturer Data: their Company Identifier in hexadecimal, e.g. ‘`0×004C`’ for Apple devices[7]. For others: ‘`unknown`’.
Raspberry Pi	String	blepackets	RPI that received the packet. Figure 1 shows the name and the approximate location of our receivers.
RSSI	Integer	blepackets	The Received Signal Strength Indicator in dB.
Address Hash	Hex	blepackets	A salted hash of the MAC address of the source of the BLE broadcast packet.
Sensor Id.	Hex	blepackets	An identifier of the device that sent the packet. The default is identical to *Address Hash*. For *Estimote Stickers* is a hash of their UUID to overcome MAC Address randomising. *TI Sensortags* maintain their MAC Address constant.
Payload Hash	Hex	blepackets	A hash of the payload bytes.
Payload Length	Integer	blepackets	Number of packet bytes after the MAC Address.
Temperature	Float	sensordata	The ambient temperature in °C measured by wristbands.
Accelerometer (X,Y,Z)	Vector Float	sensordata	The readings of the accelerometer in the wristbands.
Is moving?	Boolean	sensordata	*(Estimotes only)* Indicates if the sticker is in movement.
Previous Motion State	Integer	sensordata	*(Estimotes only)* Seconds the sticker was in its previous state, either moving or static.
Current Motion State	Integer	sensordata	*(Estimotes only)* Seconds the sticker has been in its current state, either static or moving.

accurate, this technique is effective in grouping the packets produced by our wristbands, including those in different frequency channels, which contain the same sensor readings.

Using these criteria, there are between 2 and 3 copies of each packet. Figure 2 illustrates the number of records generated in the database per minute and the corresponding BLE packets that created them. The amount of packet copies is relatively low at the beginning of the evening and then increases drastically at around 19:00. This reflects the fact that wristbands were stored at the reception, where just receiver pi22 was close, and as the party started, people carried them to areas with a higher density of receivers, such as the dinning room or the dance floors. Having several copies of each packet and receiving data from the same device in different RPIs can be useful to apply this data for localisation purposes.

4.2 BLE noise

BLE is an extremely popular protocol, so we expected interferences from devices unknown to our system, such as smartphones carried by guests. The reality was even more extreme: most of the packets we captured are from unknown devices and, even during the party, the amount of unknown packets received was almost equal to the number of packets from the wristbands. This means that in real life scenarios, when a protocol in the 2.4 GHz band is used, one must expect a highly congested spectrum. We were curious about what kind of BLE broadcast packets were received, so we parsed the advertisements in these packets and identified manufacturers when possible. Figure 3 shows the result. The x-axis shows the hexadecimal Company Id., unless it is a packet from the wristbands, or a packet we could not identify. The y-axis is the number of packets and copies received of that type in a logarithmic scale. We observe a great diversity of devices, from Apple (`0×004C`) to Samsung (`0×0075`). Although we have used the same mechanism to parse all Company Ids from the BLE packets, we suspect that some packets declare it as little endian and some others as big endian. Thus, it is possible that, both `0×0075` and `0×7500` belong to Samsung devices.

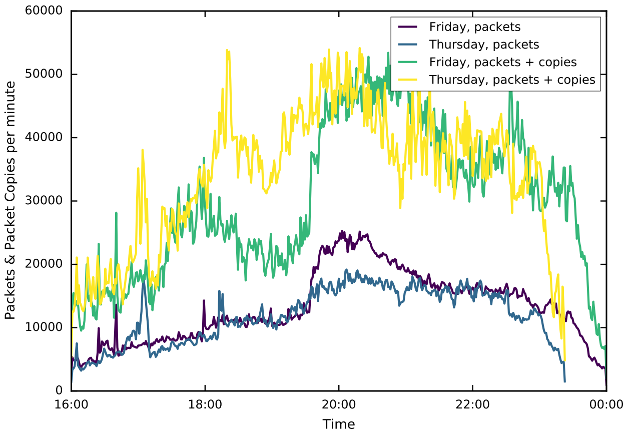

Figure 2: Messages received vs. records stored.

4.3 Infrastructure performance

One of our uncertainties was the performance of BLE in a dense environment with many known and unknown devices, and people affecting the propagation of signals. We deployed 17 receivers located in strategic positions to receive as many packets as possible and provide us with the desired functionality. As expected, the RPI

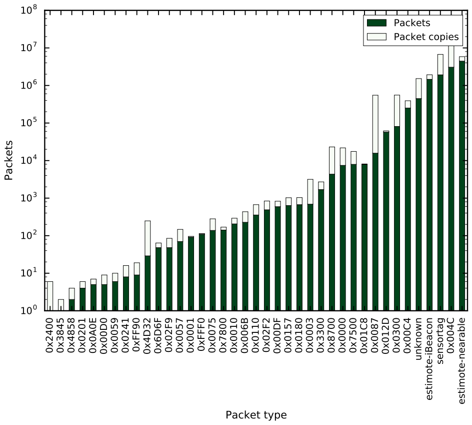

Figure 3: Number of packets received by type (log **scale).**

in the reception was the one that collected most packets, because all wristbands were stored there for a few hours. It is noticeable how all RPIs contributed to the number of received packets. Even those that were only used on Thursday (pi17, pi20 and pi21) were useful to increase the packet count then. Looking at the most active receivers at each moment provides insights about the space and the behaviour of people during the event. For example, pi10, which was located in between the two dance floors, was not the primary receiver for many wristbands, but it captured a lot of data. Whereas pi11 was in the middle of the main dance floor, where people gathered most. Thanks to the relatively high density of RPIs, it should be possible to obtain a coarse-grained location of people, by discriminating between the receivers detecting each device at each moment.

We would like to provide some insights on how our sensing infrastructure performed in real-time. It received in the order of 13,000 packets per minute including packets both from wristbands and other devices. If we consider 450 wristbands, each of them broadcasting one packet every 1.25 seconds, it adds up to 21,600 packets per minute. This number is much higher than the number of packets received. Even if half the advertisement rate is considered, i.e. 2.5 seconds, the number of packets is just slightly inferior to the number of packets registered in the system. These numbers suggest that this dataset was collected in a congested BLE network.

We do not know the total number of BLE devices in the environment, because of the frequently used MAC Address randomising mechanism, but we know that on average 400 wristbands per minute were detected. Figure 4 shows the number of *Estimote Stickers* and *TI SensorTag* devices seen per minute in the club. For *TI SensorTag*, variations are very slow. On Friday right after 16:00 we can hint the activation of a few wristbands. Then, around 23:00 we see people leaving the space. For *Estimote Stickers* before 18:00, there is a lot of instability. This is because on Thursday at some point all of the wristbands were at the venue, probably congesting the receivers close to them. Then, on Friday, some extra wristbands were brought around to accommodate more guests. Then, after 18:00 the situation is more stable, and we see an steady decreasing trend that likely indicates people leaving. Note that although the number of wristbands seen is overall pretty stable, it has variations from one minute to the next. This implies that there are minutes in which we did not receive any data from a few unlucky guests.

5 EXAMPLES OF USAGE

There are several example use cases and illustrations for the CWI-ADE2016 dataset. Our examples focus on two research communities. On one hand, we target those interested in network performance, network modelling and designing applications for congested networks. On the other hand, we offer some initial insights on how to use the dataset to analyse different aspects of crowd movement.

5.1 BLE dense network performance

Networks with a high number of devices sharing the same medium are difficult to model and predict. When the number of devices increases—specially with heterogeneous devices—it is difficult to formulate theoretical models, which often consider ideal signal propagation conditions. This situation is aggravated in indoor spaces crowded with people. For these reasons, we believe that this dataset can be used as a tool to approximate network behaviour in similar conditions. Naturally, every real life event is different, but this data can be used to complement and extend the conclusions of known experiments, models and new techniques [2, 4, 6]. Our observations show consistent patterns when analysing the RSSI levels of the packets received by the RPIs. We believe that this data can be leveraged by the modelling community to validate and enhance their models of BLE broadcasting behaviour.

Designing applications to support congested networks is always a challenge. Being aware of this, our applications were designed to cope with data losses. Although they benefit from receiving as much data as possible, they do not require it. If applications neglect this issue, they are bound to underperform. This dataset can be used to determine network reliability in challenging BLE environments and to design applications accordingly. For example, Figure 5 represents the time in seconds between consecutive packets of the same sensor as a Cumulative Distribution Function. According to the specifications, our wristbands should broadcast every 1.25 seconds—or 2.5 if it is a static *Estimote Sticker*. However, once the number of devices sharing the air increases and we add people that move freely, i.e. we add real conditions, the expectations are not fulfilled. During the event, 50% of the *Estimote Nearable* packets coming from the same sensor had an interval longer than 2.6 seconds between them, and for 10% it was longer than 10 seconds. The impact was smaller for *TI Sensortag* devices—potentially because of their more powerful hardware—but also existed. This data, combined with other such as sensor location, can help researchers in designing better applications or better BLE beaconing schemes.

5.2 Crowd Movement

Our second example reasons about people's behaviour during the event: can we use this type of sensors to tell how they moved around the space? Can we find groups of people? Indoor location using BLE or other technologies has had a lot of attention lately [7, 9]. However, getting accurate results is extremely difficult in the presence of crowds that unpredictably attenuate signals. Fortunately,

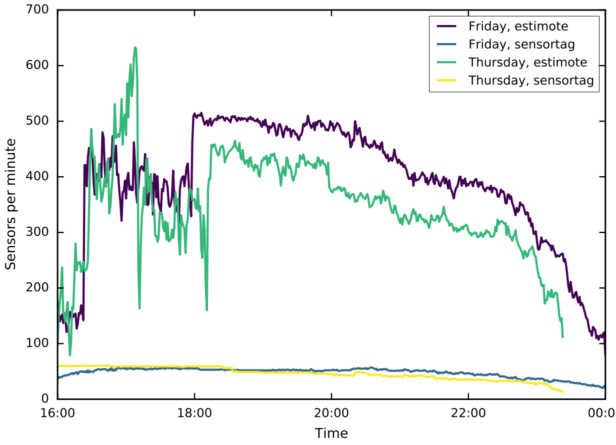

Figure 4: Number of wristbands seen per minute.

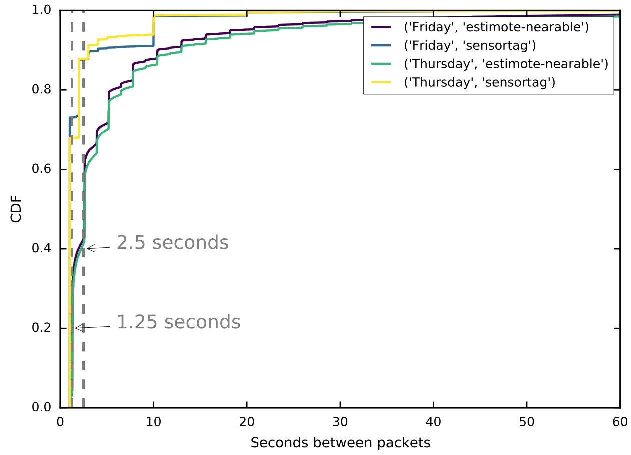

Figure 5: Cumulative Distribution Function of the time between packets from the same device.

to understand crowd behaviour, precise location is not strictly necessary and other metrics can be used, such as proximity [8] among people or to defined points.

Following this approach, we have created an interactive visualisation[8] that shows the location of the wristbands at room level. A wristband is in the room where its last packet was received, excluding copies with the 100 milliseconds criterion explained in Section 4. If a packet is not received in 10 minutes, we consider that the wristband left the space. We find that this interactive visualisation shows the movement of FoRB during dinner, an event that can help as an approximation for ground truth. This kind of analysis is not accurate to study individual behaviour, as errors in locating an individual are easy to make. However, it is a powerful tool to show the crowd as a whole where errors are relatively low.

Another interesting question is to see if we can cluster people from our data. For example, if we can spot groups of friends that enjoyed the party together. We have made an initial attempt, looking into the data of the FoRB wristbands. This group is easy to differentiate from the other guests, because we know that they were shown around and spent some time together having dinner, so it is possible to establish some ground truth. We have looked into the sequence of rooms that each wristband visited during the night, then we carried out a pairwise comparison looking for matches, i.e. two wristbands in the same room at the same time. We counted the matches and constructed an index that we fed into a hierarchical/agglomerative clustering algorithm. In the dataset website[9], the reader can find a report containing the figures of this analysis —and all the others. We have not included them here due to space limitations. However, we have observed that it shows clustering of some FoRB, indicating that there were at least two groups of people that apparently spent a lot of time together. Thus, we believe further analysis is worth it.

6 CONCLUSIONS

This paper has scratched the surface of the data we gathered by handing BLE enabled wristbands with embedded sensors to guests of a two-day club event. We find the experiment and data valuable as it offers researchers millions of BLE sensor packets in a relatively small space concentrated during an event. However, not only the data is valuable, but also the lessons learned in the process of collecting it. We have shown different aspects of our design, with its advantages and disadvantages, and we are willing to share our experience and the software used, both during the event and for the analysis. We would like to encourage other researchers to use this data in different ways and to enquire us about information that is not yet in the dataset and they believe to be useful.

REFERENCES

[1] A. Ahmed, S. Benford, and A. Crabtree. 2012. Digging in the Crates: An Ethnographic Study of DJS' Work. In *SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 1805–1814.

[2] M. Aiello, R. de Jong, and J. de Nes. 2009. Bluetooth broadcasting: How far can we go? An experimental study. In *2009 Joint Conferences on Pervasive Computing (JCPC)*. 471–476. DOI:http://dx.doi.org/10.1109/JCPC.2009.5420140

[3] SIG Bluetooth. 2010. *Bluetooth core specification version 4.0.* Specification of the Bluetooth System.

[4] Keuchul Cho, Gisu Park, Wooseong Cho, Jihun Seo, and Kijun Han. 2016. Performance analysis of device discovery of Bluetooth Low Energy (BLE) networks. *Computer Communications* 81 (2016), 72 – 85. DOI:http://dx.doi.org/10.1016/j.comcom.2015.10.008

[5] Nathan Eagle and Alex (Sandy) Pentland. 2006. Reality Mining: Sensing Complex Social Systems. *Personal Ubiquitous Comput.* 10, 4 (March 2006), 255–268. DOI:http://dx.doi.org/10.1007/s00779-005-0046-3

[6] Robin Kravets, Albert F Harris, III, and Roy Want. 2016. Beacon Trains: Blazing a Trail Through Dense BLE Environments. In *Proceedings of the Eleventh ACM Workshop on Challenged Networks (CHANTS '16)*. ACM, New York, NY, USA, 69–74. DOI:http://dx.doi.org/10.1145/2979683.2979687

[7] Dimitrios Lymberopoulos, Jie Liu, Xue Yang, Romit Roy Choudhury, Vlado Handziski, and Souvik Sen. 2015. A Realistic Evaluation and Comparison of Indoor Location Technologies: Experiences and Lessons Learned. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks (IPSN '15)*. ACM, New York, NY, USA, 178–189. DOI:http://dx.doi.org/10.1145/2737095.2737726

[8] Claudio Martella, Armando Miraglia, Jeana Frost, Marco Cattani, and Maarten van Steen. 2016. Visualizing, clustering, and predicting the behavior of museum visitors. *Pervasive and Mobile Computing* (2016), –. DOI:http://dx.doi.org/10.1016/j.pmcj.2016.08.011

[9] R. Harle R. Faragher. 2014. An Analysis of the Accuracy of Bluetooth Low Energy for Indoor Positioning Applications. In *Proceedings of the 27th International*

[8] http://johnguerra.co/viz/crowdMovement/

[9] https://github.com/cwi-dis/CWI-ADE2016-Dataset/blob/master/analysis/README.md

Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014). 201–210.

[10] Thomas Röggla, Sergio Cabrero, Demosthenis Katsouris, Zhiyuan Zheng, Amritpal Singh Gill, Jack Jansen, Judith A. Redi, Pablo Cesar, and David A. Shamma. 2017. The Club of The Future: Participatory Clubbing Experiences. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3027063.3052967 (To Appear.).