

# Detecting Human Encounters from WiFi Radio Signals

Geert Vanderhulst, Afra Mashhadi, Marzieh Dashti, Fahim Kawsar

Bell Laboratories

{firstname.lastname}@bell-labs.com

## ABSTRACT

We present the design, implementation and evaluation of a novel human encounter detection framework for measuring and analysing human behaviour in social settings. We propose the use of WiFi *probes*, management frames of WiFi, that periodically radiate from mobile devices (as proxies for humans), and existing WiFi access points to automatically capture radio signals and detect human copresence. Based on the spatio-temporal properties of this copresence and their interplay we defined a model, borrowing theories from sociology, to detect *human encounters* – short-lived, spontaneous human interactions. We evaluated our framework using controlled and in-the-wild experiments yielding a detection performance of 96% and 86% respectively. As such, our framework opens up interesting opportunities for designing proxemic and group applications, as well as conducting large-scale studies in the areas of computational social sciences.

## CCS Concepts

• **Human-centered computing** → *Ubiquitous and mobile computing systems and tools*;

## Author Keywords

Spontaneous Encounters, WiFi Signals, Hybrid Sensing

## INTRODUCTION

Understanding human social interaction is an important facet of modelling individual and group behaviour, with the potential to transform our understanding of our lives, organisations, and societies. Indeed, in the digital world, thanks to the unprecedented traces from users of social media [19, 26], emails [9, 15], and call data records [7, 23], the quantification of human social interactions has been studied extensively to compile comprehensive pictures of the ways we learn, act, think, and behave as social human beings.

However, the emergence of such a data-driven “computational social science” based on the physical world has been much slower. Only in the recent years, modern technologies, such as video surveillance [5], wearable smart badges [2, 22], and multimodal sensor equipped smartphones [10, 24] are used to explore the human social interactions in the physical

world. Although accurate, most of these technologies come at the expense of high deployment and management costs, and sometimes at the risk of privacy erosion. Besides, the intrinsic requirement of carrying special purpose hardware limits the potential of these approaches beyond short-term studies. Being able to detect human interactions with minimum obstruction to their daily life routines and settings could enable new avenues of research in the social psychology domain where encounters and social engagement of individuals, such as those with behavioural disorders (e.g., depression), could be monitored for medical diagnostics.

A possible solution is thus to rely on the devices that are already embedded in our lives such as smartphones to act as proxies for humans, granting the assumption that people mostly carry their phones with them [16]. For instance, previous work has shown that people carry their mobile phone at the workplace, 48% of the time within arm-reach and 82% of time within 5 meters [8], making the phones’ location a good approximation for users’ location in the offices. However, active sensing approaches on mobile phones have been shown to be inefficient in terms of energy consumption [3, 14].

Motivated by these issues, in this paper we aim to build a novel framework for detecting human *spontaneous encounters*. In particular we are interested in detecting those social interactions that are short-lived in nature and are spontaneous between a small set of individuals. Our framework leverages existing WiFi infrastructure and WiFi signals radiating from mobile devices to detect human copresence. Borrowing theories from sociology, we propose a novel model for detecting human encounters based on the spatio-temporal properties of this copresence and their interplay. As such, the contributions of our work are twofold:

- A model, grounded on sociology theory, for detecting spontaneous human encounters in the physical world from copresence data.
- The design, implementation and evaluation of an end-to-end framework embodying this model that actively leverages existing WiFi infrastructure with minimal but wilful participation from the users.

Our encounter detection framework essentially transforms quantified network noise into social signals. In doing so, it opens up opportunities for novel proxemic applications and large-scale human behaviour analytics at low cost. After outlining the terminology used throughout the text, we elaborate on the design challenges and implementation of our framework, followed by an evaluation (controlled and in-the-wild experiments) and concluding remarks.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

MUM '15, November 30–December 02, 2015, Linz, Austria

© 2015 ACM. ISBN 978-1-4503-3605-5/15/11...\$15.00

DOI: <http://dx.doi.org/10.1145/2836041.2836050>

## COPRESENCE AND HUMAN ENCOUNTER

In this section we first define the terminology that we use throughout the paper. We then describe the scope and scenarios of human encounters that we want to capture.

### Terminology

In order to define what we mean by an *encounter* and its characteristics, we visit the definitions from the social science literature where different aspects of human groups and interactions have been studied for many years by sociologists. In the study of human groups [13], a main distinction between *primary* and *secondary* groups is made. Primary groups are those in which everyone knows everyone else on face-to-face basis and thus can be treated as a small social system. In contrast, the secondary groups are characterised by impersonal relations and are goal-oriented. However, in this work we are not interested in the *social relationships* between the individuals, rather in their interactions. Thus, to make our taxonomy independent of this relational aspect, we avoid using the term *group* and instead borrow two other major terms from sociology: *copresence* and *encounter*.

**Copresence.** This term refers to the spatio-temporal conditions under which humans can interact with each other. Defined by Goffman [12], *copresence* is the condition when people “sense that they are close enough to be perceived in whatever they are doing, including their experiencing of others, and close enough to be perceived in this sensing of being perceived”. These conditions are expected to persist throughout a contained space like a room, and to apply to anyone in that space. For example, in the context of a workspace, employees who share an office can be referred to as being copresent to each other. Whereas in public spaces, such as streets, the area for which the defined mutual perceiving prevails is hard to define. To quantify the physical distance over which one person can perceive another with the naked sense, we borrow from [20] where a distance limit of up to 4 meters is proposed. Furthermore we are interested in *corporeal* presence, situated in semi-private spaces (e.g., workplace), which refers to bodily presence and excludes remote interactions such as telephony or online social interaction [29].

While the term *copresence* captures the condition of being close to each other from the human sense, it does not capture interactions amongst the copresent individuals. For example, a group of individuals waiting at a bus stop could be referred to as copresent despite no human interaction amongst them is exhibited.

**Encounter.** Defined by Goffman [12] the term *encounter* comprises “all those instances of two (or more) participants in a situation joining each other openly in maintaining a single focus of cognitive and visual attention (single mutual activity), entailing preferential communication rights”. A simple example of an encounter is when persons meet each other in the workplace corridor and engage in a conversation. The term encounter implies two main properties by definition: Firstly, it is not only constrained to verbal conversations, but can also include gestures, gaze and non-verbal interactions amongst persons. However, regardless

of the communication channel, the participants are bound to be mutually involved. Secondly, the term is independent of social ranking of human interactions – persons with a different social status or background can be engaged in an encounter – and covers occasions which bring two persons into close access to each other. Finally, to conclude the definition and properties of *encounter*, we borrow from Zhao [29] to constrain our definition to real-time, reciprocal encounters which exclude short non-intentional interactions, such as greeting a postman passing by.

In this work we are interested in human encounters which are intrinsically short-lived, which we refer to as *spontaneous encounters*.

### Scope Definition

In order to define the scopes and scenarios of the spontaneous encounters that we are interested in this work, we consider two variables that are exhibited in every human encounter: *duration* and *size*.

**Duration ( $\delta t_g$ ).** The duration represents the amount of time that an encounter existed, from the time of formation to its decomposition. Duration is a simple metric but a good indicator for differentiating various types of encounters. For example, a high  $\delta t_g$  – in the order of hours – can be reflective of a cohabiting group such as a family where the members spend long periods of time together at home. Whereas a smaller value of  $\delta t_g$  – for example less than 10 minutes – could be more representative of spontaneous encounters such as coffee breaks at the workplace and casual chats.

Furthermore, as the encounters are defined based on the engagement of the individuals, the termination of the encounter (and thus  $\delta t_g$ ) is highly affected by the entrance and departure of a participant. We will discuss this property and how it is used in our encounter detection algorithm later in this paper.

**Size ( $N$ ).** The size represents the number of distinct individuals in the encounter. In the domain of social psychology, size has been shown to play an important role in determining the nature of groups. J.C.Turner, a social psychologist, proposed a theory of size and depersonalisation [25]. He states that as the size of the human group grows, the group becomes more depersonalised.

Based on these two variables and their interplay, we now empirically define a model for human encounters.

- $N \geq 2$ . By definition an encounter should have at least two members.
- $\delta t_g > 1min$ . For an encounter to be considered a meaningful reciprocal engagement (as opposed to random short-lived copresence), its duration should be at least 1 minute. This means any encounter that is less than 1 minute in duration is excluded from our definition of encounter, as illustrated in Figure 1a.
- As the size of the persons in the encounter  $N$  increases, the duration of the encounter also increases as to allow the individuals to mutually involve in the communication and

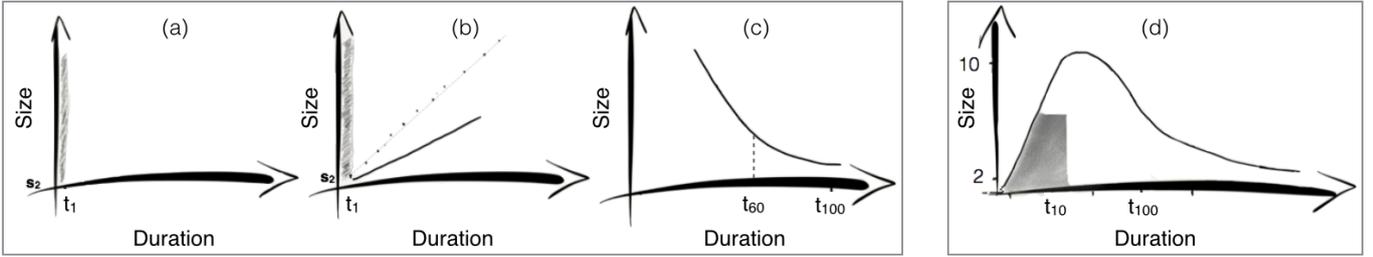


Figure 1: Scope and boundaries of a human encounter: (a) the exclusion of random temporary copresence; (b) the sub-linear growth of copresent users; (c) the decrease in size of an encounter for longer durations; (d) an empirically defined model from three previous social observations.

cognitive/visual attention [12]. We model this property as a sub-linear growth (Figure 1b). For example, while a two minute encounter for 2 persons could be long enough for a social chat, a 5 persons encounter might require more than 5 minutes. We claim this requirement based on two properties from sociology. Firstly, as the number of persons increases, the longer it takes for the *ceremonial rituals*<sup>1</sup> of encounters formation and decomposition [11]. Secondly, it is more likely for the focus of attention (e.g., topic of the conversation) to be shifted when more individuals are involved in a verbal encounter. This property is illustrated in Figure 1b by a sub-linear growth of size.

- Finally, as the duration  $\delta t_g$  increases, the likelihood of the encounter between a large number of individuals decreases. This is because it is less conventional for a large number of people to actively and mutually interact with each other for a long period of time (e.g., 10 people interacting for over 1 hour), unless the copresence is a planned gathering such as a meeting which falls outside the scope of this work. This property is illustrated in Figure 1c by a descending line.

Figure 1d brings together the above properties, for an empirically defined boundary of size and duration. In this paper, we are interested in those encounters that fall in the shaded area which correspond to intrinsically short-lived spontaneous encounters. These spontaneous encounters could be of three different types: (i) spontaneous encounter of two or more individuals where coming together does not seem to have a previously defined rationale, such as small talk in the corridor; (ii) spontaneous encounter of two or more individuals with a clear rationale for their gathering, for example people going for a coffee break; (iii) spontaneous encounter of two or more individuals with additional persons joining in a later time. This encounter could be related to a specific mutual instrumental activity. For example, a group of friends having a debate and asking a third party to join their conversation.

In order to confirm that the proposed model is viable, we looked at the measured interactions of 23 employees over 30 days as collected by [22]. This dataset monitored the interaction amongst employees of a data server configuration firm for one month through infrared badges. Figure 2 presents the

<sup>1</sup>The ceremonial ritual refers to the conversations or actions that are necessary to keep participants in line and give a closure to the mutual activity sustained in the encounter.

duration and size of these interactions. As illustrated, the majority of the interactions fall within short-lived spontaneous ones with a few participants. The figure also demonstrates that as the number of participants increases so does the duration of the interactions.

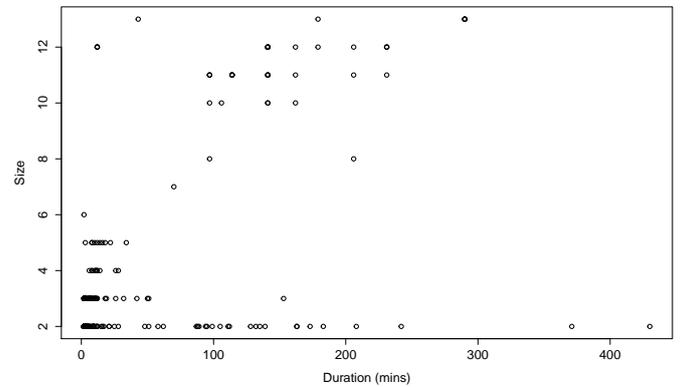


Figure 2: The size and duration of interactions amongst employees in a real enterprise.

## DESIGN CHALLENGES

In the previous section, we described the foundation of our spontaneous human encounter detection framework. Essentially, to detect these encounters automatically we need sensory signals that can accurately and reliably track the copresence of humans. This copresence is then modeled to detect human encounters. There are multiple alternatives to acquire the sensory signals with varying degree of characteristics in terms of system requirements, measurement complexity, and privacy concerns. For instance, past works have explored capturing copresence information using dedicated multimodal sensing infrastructure [5], dedicated sensor badges [2, 22] or using mobile sensing techniques [10, 24]. Although accurate, these approaches often come at the expense of high deployment costs and privacy erosion. In this paper, we divert from these traditional approaches and apply a network-centric approach: we rely on existing WiFi infrastructure with minimal but willful user participation to detect human copresence from WiFi signals. Our rationale behind using WiFi signals is grounded upon the findings from past research suggesting people carry their phones with them most of the time [16] having WiFi switched on [3]. We argue that

this is a reasonable assumption for most of the spaces we target: workspaces, conferences, museum, networking events, etc. Besides, by using existing WiFi Infrastructure together with pervasive mobile phones, we aim to reduce the deployment and management costs substantially without compromising the quality of detection. However, to realise our proposed approach, we need to overcome three major challenges as discussed below:

**Zero Mobile Sensing** We aim to detect human encounters without the need of sensory information from the user’s personal device, such as audio/video signals, GPS coordinates, accelerometer data or WiFi scans. Moreover, we want to step away from discovery protocols (e.g., Bluetooth) due to the risk of detecting peripherals that do not uniquely identify a user (e.g., audio headset, wearables, printer, etc), and the fact that Bluetooth is usually turned off by users [3]. Instead, we aim to leverage day to day WiFi radio signals exchanged between personal devices and local network infrastructure.

**Detection Accuracy** We rely on common network infrastructure such as WiFi APs to capture WiFi signals originating from the users’ devices. However, as pointed out in [21], no passive WiFi tracking system can guarantee accurate tracking performance, due to a number of issues: (i) temporal sparsity of detections, (ii) spatial sparsity of detections across network nodes and (iii) unpredictable path loss, e.g., caused by walls, furniture, people, etc. Since many human encounters are short-lived, it is particularly challenging to guarantee device detections in sub minute intervals, whilst devices might or might not be connected to a WiFi network.

**Detection Reliability** When opportunistically leveraging radio signals to detect human encounters, a valid concern is the rate of erroneous detections. This error rate is challenged by several phenomena beyond our control, such as signal fluctuations within an environment and device heterogeneity. For instance, different phone designs (e.g., used WiFi chip, antenna position, casing, ...) might affect the strength of a signal leaving a phone. Besides, a phone sitting in a user’s pocket might behave differently from a phone laying on a desk.

While consistent transmission of WiFi signals from mobile devices can mitigate these challenges, it requires mobile devices to constantly send data packets, which is often energy expensive. However, according to the 802.11 standard, a device can periodically radiate a WiFi signal to actively search for available networks as long as the WiFi chipset is on. This WiFi signal, a *Probe Request* – further referred to as ‘probe’ – is a WiFi management frame that can be either directed to a specific network, by indicating its SSID, or broadcasted to any network within range. Figure 3 shows an example of a directed probe. Mobile devices periodically send out these probes with a frequency that is vendor-specific and also depends on the power state of the device. Upon receiving such probe, an addressed AP replies with a Probe Response enabling the device to initiate a connection, or associate with an AP with a stronger signal. Since probes are not encrypted,

they can easily be intercepted using commodity hardware, i.e., a WiFi chip set to monitor mode. In our framework we use these probes as the primary means to capture radio signals transmitted from users’ devices (as proxies of humans).

DA	SA	BSSID	SSID
ff:ff:ff:ff:ff:ff	90:68:c3:be:34:9f	ff:ff:ff:ff:ff:ff	DIRECT-

Figure 3: A (partial) Probe Request originating from a device with MAC address “90:68:c3:be:34:9f” that is looking for an AP with SSID “DIRECT”.

It is known that the frequency by which probes are sent greatly varies between vendors and device models [21]. Indeed, in a small experiment we conducted at our office space, we noticed the number of probes sent by individual devices ranging from less than 20 up to more than 100 in a period of 10 minutes. The majority of devices (laptops and mobile phones) being used at our office space have an active WiFi connection which causes mobile phones (e.g., Moto G phones) not to send probes at all within our time window. These observations demand for a strategy to increase the number of probes being sent by mobile phones. We consider two options to guarantee predictable probe transmissions:

**Network-centric** Prompting devices for additional WiFi transmissions from within the network. For instance, in [21] Musa et al. have shown that by emulating the SSID of popular or previously visited WiFi access points (AP), devices can be tricked into (re-)associating with an AP and generating extra probes. The advantage of this approach is that it is fully passive, albeit it exploits the chattiness of WiFi chips in a privacy-invasive way and interrupts active WiFi connections.

**Device-centric** Synchronise the transmission of probes at the device-end. While this approach requires a small device-centric component, it guarantees a high detection rate. Most importantly this allows users to opt in and out to our human encounters detection framework.

Using either of these approaches, we argue that the *Detection Accuracy* challenge can be realistically addressed. While we can exercise control over the timing or probes, we cannot mitigate the potential effects of device heterogeneity on signal fluctuations. As a result, some inaccuracies are likely to be introduced when calculating human copresence. By feeding aggregated copresence data (as sensed by multiple network nodes) through a semantic model capturing human limits (e.g., duration and size of human groups), we aim to increase the *Detection Reliability* in our framework.

#### Energy Consumption of WiFi Probes

In what follows, we assess the energy gains that can be expected from our approach at the device-end. We compared the energy consumption of a WiFi scan and a WiFi probe (i.e., listening to radio signals from nearby APs versus transmitting a radio signal). To gain control over probes without rooting a phone and modifying its WiFi drivers, we leverage the WiFi Direct standard. WiFi Direct is designed to enable devices to easily connect with each other without requiring a wireless

AP. An important step in this process is the device discovery phase, when devices send out probes to advertise their presence. These probes can be distinguished from other probes by their SSID value (“DIRECT-”) and are broadcasted on channels 1, 6 and 11 in the 2.4GHz band. To measure the energy consumption of sending out a controlled WiFi Direct probe, we initiate a WiFi Direct device discovery and immediately abort it as soon as one probe has been sent out. Performing a WiFi scan is supported natively on Android phones, albeit we witness other manufacturers (i.e., Apple) to shield this functionality from developers.

We disassembled an Android Moto G phone and measured its battery’s discharge current for either task using a Tektronix Power Analyser. The results of our energy measurements are listed in Figure 4. While the peak current drawn by both tasks is very similar, the WiFi scan consumes more energy because it takes longer. During a WiFi scan, the chip needs to tune to each WiFi channel for at least 100ms and listen for incoming (beacon) frames that are typically broadcasted every 100ms by APs. During a WiFi direct discovery process, devices also listen to each other’s probe responses which introduces some overhead – a trade off between leveraging existing APIs and manipulating a phone’s firmware.

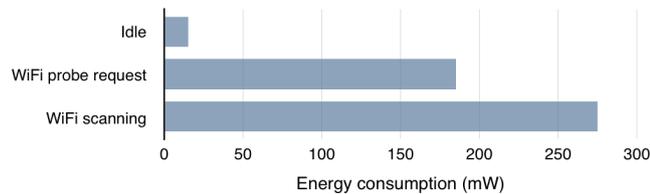


Figure 4: Comparison of the energy consumption for sending a WiFi probe and performing a WiFi scan on a Moto G phone.

As the figure shows, the actual power gain of a WiFi Direct probe over a WiFi scan is approximately 30%. However, the real gain comes from the fact that as soon as the probe leaves the phone, it is in the network. In contrast, collecting signal strength (RSS) measurements on a phone requires active network connection to communicate its scan results to a server in the network for further analysis. In the next section, we explain the detail technical approach of our framework that actively leverages the decisions informed in this section.

## DESIGN AND IMPLEMENTATION OF THE FRAMEWORK

The framework proposed in this paper, consists of three main components as illustrated in Figure 5.

### Radio Signal Capturing Engine

Based on insights from our feasibility study, we opted for a hybrid sensing approach consisting of:

- A small device-centric component that is responsible for synchronising probe transmissions at the device-end. Hence we can aim to minimise the effects of temporal changes in the environment with probes being broadcasted simultaneously.
- A network-centric component that runs on local network nodes and which is responsible for collecting probes and feeding them to our algorithms.

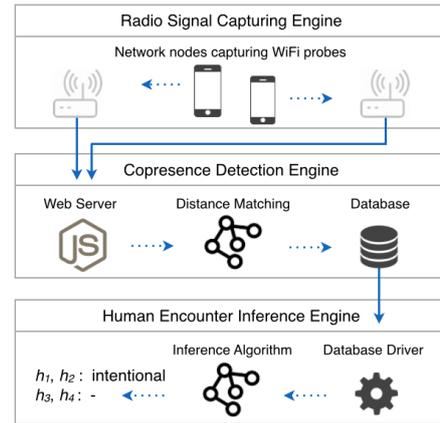


Figure 5: Architecture pipeline.

As such, our architecture is designed with the ubiquitous availability of small wireless cells (i.e., WiFi APs) with on-board computing capabilities in mind. As a proxy of a small cell, we opted for Meshlium devices from Libelium<sup>2</sup> which have all the properties of a ‘real’ WiFi AP including the ability to easily program them. A Meshlium features a 500Mhz x86 processor, two WiFi interfaces (one of which acts as an AP by default) and runs an embedded Debian Linux operating system. However, any device with a WiFi chip that can be put in monitoring mode can be leveraged for collecting probes and reporting signal measurements to our copresence engine. We involve the user in the loop as we only capture probes with SSID “DIRECT-” which are intentionally broadcasted by an app running on a user’s personal device. In the next sections, we explain how we time probe transmissions and collections.

### Heartbeat Synchronisation

Clocks embedded in electronic devices typically consist of cheap quartz oscillators which are prone to clock drift, such that they count time at slightly different rates. To correct for these errors, devices should regularly synchronise their clocks with an accurate, external time source such as GPS or an NTP server. Recent smartphones (e.g., Motorola Moto G, Google Nexus 5, iPhone 6) automatically synchronise their clocks and we found them to be on par with an accuracy up to 1 second which suits our purpose. Besides, we manually configured the Meshliums to synchronise their clocks every hour with a NTP server, over a LAN connection. To lock on to the heartbeat, a smartphone or network node initially sleeps for  $(60 - S) \bmod I$  seconds with  $S$  being the number of seconds past the minute of the current time and  $I$  being the interval between pulses (with  $I$  chosen less than 60 seconds).

### Scheduling Probe Requests

We orchestrate phones to send out a probe request in concert at each pulse of the heartbeat. We generate these probes in the same way as we did to measure the energy consumption, e.g., by leveraging the WiFi Direct standard. As such, the only requirement for a phone to be used with our framework

<sup>2</sup><http://www.libelium.com>

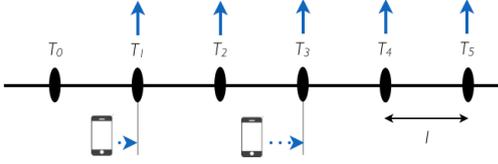


Figure 6: Phones locking on to a shared heartbeat. Once synchronised, phones send out a probe at each  $t_i$ .

is to have WiFi switched on – the phone may or may not be connected to an AP.

#### Scheduling Probe Request Collection

On the Meshliums, we put one of its WiFi interfaces in monitor mode and continuously listen for “DIRECT-” probe requests on channel 6 (albeit channel 1 or 11 could be used as well). Thus, we only capture those probe requests which we assume that are intentionally sent by a phone to the Meshliums – all other probe requests, e.g., originating from devices trying to connect to a WiFi AP, are discarded. Each Meshlium expects to pick up a probe request from each phone in its vicinity at each heartbeat pulse  $t_i$  and stores the MAC address of the phone along with the probe’s perceived signal strength as reported by the WiFi hardware. Next, at each  $t_i+2$  (i.e., two seconds after each pulse), all MAC addresses and RSSI values are bundled and reported to the copresence engine. The collective reports received from all Meshliums on the server are then passed as input to the copresence detection algorithm.

#### Copresence Detection Engine

This component consists of a NodeJS Web server to which network nodes report their RSS measurements and a MongoDB database where the copresence results are stored. The idea behind our copresence technique is that the devices of copresent users will exhibit similar radio signal characteristics at a shared location. To this end, we compare the RSS values from different detected devices with each other according to the algorithm outlined in [6]. The aggregated set of RSS values obtained from different network nodes is a vector that acts as a fingerprint of the user’s current location. For instance, if network nodes  $M_1$ ,  $M_2$  and  $M_3$  reported the detection of a device with MAC address “90:68:c3:be:34:9f” at time  $T_0$ , its corresponding fingerprint vector entry resembles:

$$t_0, 90:68:c3:be:34:9f, [M_1 : -48, M_2 : -64, M_3 : -58]$$

We calculate the distance (in the signal domain) between fingerprints from every  $j$ -th and  $k$ -th device,  $d_{j,k}$ . To this end, we used the Manhattan distance metric, in which the sum of the absolute differences of signal strengths is computed. Assume that  $r_j^{M_i}$  and  $r_k^{M_i}$  denote RSS values from the  $j$ -th and  $k$ -th device observed by network node  $M_i$  with  $N$  corresponding to the number of common network node detections, their Manhattan distance can be calculated as:

$$d_{j,k} = \sum_{i=1}^N |r_j^{M_i} - r_k^{M_i}| \quad (1)$$

Note that two fingerprint vectors may include an unequal list of different observed devices. Before applying the distance metric for a pair of devices, we first extract common detections. To have a fair comparison, we then normalise the calculated distance by dividing it by  $N$ .

Users whose device fingerprints differ less than a dissimilarity threshold  $\delta$  are considered to be copresent. This threshold  $\delta$  should be chosen to approximate a distance of up to 3 meters between devices which has been identified as the boundary of a potential human encounter. We refer to our Sensitivity Analysis where we identify optimal configuration parameters such as the threshold value and preferred heartbeat interval. To increase the confidence level that detected copresent users were actually together, we aggregate copresence samples over a 1 minute window as explained in the same section.

#### Human Encounter Inference Engine

Using the two variables described earlier, we now build a feature which reflects the *stability* of the encounter’s gathering by taking into account the entrance and departure of the participants. We define this stability feature by reasoning that an encounter exhibits mutual commitment of the participants. This means the members all contributed equally to the encounter by staying the same amount of time together. We formulate this as a function  $f(G)$  for a given group of copresent users  $G$  where  $G$  contains  $N$  distinct individuals:

$$f(G) = \frac{\sum_{h_i \in G}^N \delta t_{h_i}}{N \times \delta t_G} \quad (2)$$

Here  $h_i$  denotes the  $i$ -th human in the copresence group and  $\delta t_{h_i}$  denotes the time the participant stayed in the encounter. For  $G$  to be a spontaneous encounter, we expect  $f(G) \rightarrow 1$ . As  $f(G) \rightarrow 0$  we observe volatility in the group which can correspond to copresence scenarios such as people queueing at a coffee machine.

Based on this feature we built an *online* inference model layered on top of the copresence engine. The input to this model is a list of copresent humans, fed to our model every minute as shown in the example below.

$$t_0, \{h1, h2, h3\}, \{h4, h5\}$$

$$t_1, \{h2, h3\}, \{h1, h4, h5\}$$

In order to infer spontaneous encounters from this input stream, at every timestamp (i.e., every minute), we infer to which of the past events the current copresence list corresponds to. We do this by taking into account the membership of the copresent individuals, and deciding at each timestamp whether they already fit a previously identified copresence group or if we need to create a new copresence group. For example, based on the above example, at time  $t_0$ , two copresence groups co-exist ( $\{h1, h2, h3\}, \{h4, h5\}$ ). At time  $t_1$ , our model receives  $\{h2, h3\}$  and  $\{h1, h4, h5\}$  as input. Given that  $h2$  and  $h3$  were previously seen at time  $t_0$ , we add  $\{h2, h3\}$  as the follow-up time sequence to the

$\{h1, h2, h3\}$  copresence group. Similarly  $\{h1, h4, h5\}$  is added to both  $\{h1, h2, h3\}$  and  $\{h4, h5\}$  given the participation of  $h1$  and  $h4$  in those copresence groups at time  $t_0$ . Therefore at time  $t_1$  we extract the following possible copresence groups:

$$CG1 \left( \begin{array}{l} t_0 : \{h1, h2, h3\} \\ t_1 : \{h2, h3\} \end{array} \right) CG2 \left( \begin{array}{l} t_0 : \{h1, h2, h3\} \\ t_1 : \{h1, h4, h5\} \end{array} \right)$$

$$CG3 \left( \begin{array}{l} t_0 : \{h4, h5\} \\ t_1 : \{h1, h4, h5\} \end{array} \right)$$

As the number of copresence groups at the same timestamp grows, the number of *potential* combinations of the copresence groups grows accordingly. For each potential copresence group after time  $t_0$  – recall that our encounter definition imposes a minimum of at least one minute copresence – we calculate the defined stability feature  $f(G)$ . However, as our feature calculation happens *online*,  $f(G)$  needs to be calculated gradually over time. We adjust our feature calculation as follows:

$$f(W) = \frac{\sum_{h_i \in W} \delta t_{h_i}}{N_w \times \delta t_w} \quad (3)$$

where  $W$  denotes a sliding window,  $N_w$  is the number of distinct individuals in the window timeframe and  $\delta t_w$  is the duration of the window (i.e., the window size). For instance, in the above example  $f(W)$  for a window size of 2 minutes ( $\delta t_w = 2$ ) at time  $t_1$ , will result in 0.83, 0.6 and 0.83 for CG1, CG2 and CG3 respectively. As time passes (per minute), this feature is recalculated through the sliding window  $W$ . Finally, at each timestamp  $t$  we decide if an encounter has formed if  $f(W) \approx 1$ . In the above example none of the copresence groups are detected as an encounter. To account

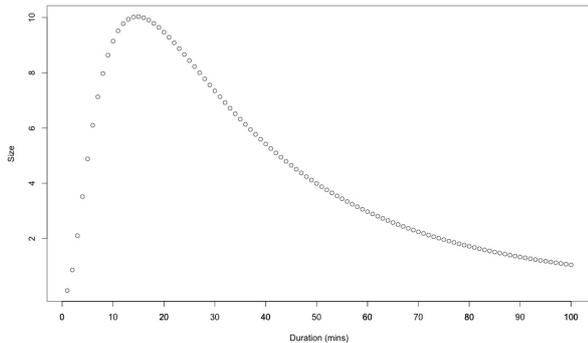


Figure 7: Log normal distribution with  $\mu = 3.5$  and  $\sigma = 0.9$ .

for the properties and limits of spontaneous encounters that we defined earlier, we set the window size  $\delta t_w$  in our feature calculation dynamically by reasoning on Figure 1. At every timestamp, we reason on the number of copresent people and follow the distribution in Figure 1 to decide how long those individuals must have been together to be considered as an encounter. To this end, we model the empirically drawn curve by experimenting with a log normal distribution. We set the

mean and standard deviation of this distribution as  $\mu = 3.5$  and  $\sigma = 0.9$  respectively through experimental parameter fitting, so to best correspond to our empirical distribution. The final distribution is presented in Figure 7. We use this distribution to detect the minimum duration for a given size of copresent people, and use this duration to set the window size dynamically. We refer to this approach as *Adaptive Window*.

Finally, in order to detect the group *decomposition*, we calculate an exponential moving average as follows:

$$S_t = (1 - \alpha) \times S_{t-1} + \alpha \times f_t(W) \quad (4)$$

where  $f_t(W)$  denotes the result of  $f(W)$  at time  $t$ . This approach allows us (by setting  $\alpha$ ) to decide on whether the present or the past should have more influence on detecting a copresence as an encounter. It also allows us to incorporate tolerance into our model as to be able to account for temporal volatility in the encounter, e.g., caused by a person passing by an encounter. We detect an encounter to be terminated if at time  $t$  the moving average  $S_t$  is less than a threshold denoted by  $th_{exit}$ .

## EXPERIMENTAL SETUP

In our setup, we use 4 WiFi monitoring nodes (Meshliums) which we deployed in our office, where about 50 people share an open office space. In placing the monitoring nodes, we followed the configuration of the existing corporate WiFi access points, assuming that future small cells would also be placed by the facility manager based on the coverage they provide. As the result of this placement strategy, these four nodes were located 15 to 20 meters apart, as illustrated in Figure 8. From the device-end, we used 9 Moto G phones (3rd generation) on which we installed a prototype application for sending out synchronised WiFi probes, as discussed in our architecture. Each of these phones was attributed a unique color, based on its MAC address. To validate the correct operation of our setup, we implemented a feedback loop in our prototype showing on each device its detected copresent peers.

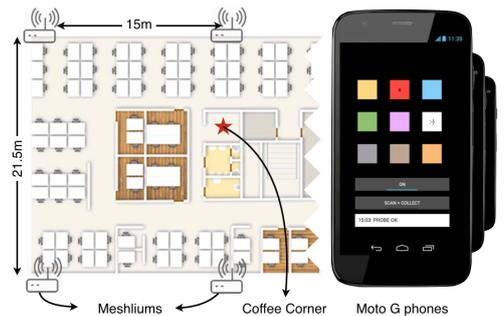


Figure 8: Floorplan of the testing area and screenshot of our Android application prototype.

## Sensitivity Analysis

Prior to evaluating the performance of our framework, we conducted a number of experiments to optimise the configuration parameters of the copresence detection engine. To this end, we collected offline training data consisting of sets of RSS measurements of pairs of copresent (less than 3 meters apart, i.e., the limit we set for a human encounter) and

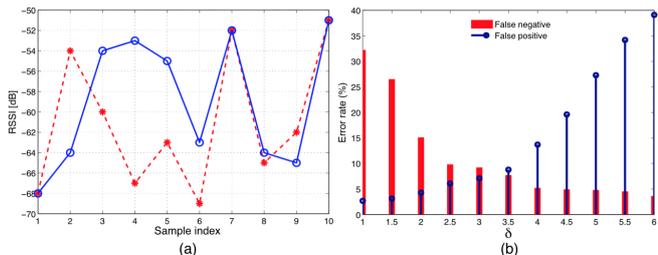


Figure 9: RSS fluctuations of copresent devices (a) and determining the optimal threshold  $\delta$  for minimising copresence detection errors.

dispersed devices at different locations in our coverage area. Figure 9a depicts the RSS fluctuations from two copresent devices at the coffee corner, as perceived by one of the Meshlium over 70 seconds. While we experience a similar variation trend, we witness sporadic RSS jumps which are consistent throughout our training data samples. Furthermore, not all probes are always captured by a Meshlium, even at close signal distance. To compensate for sudden RSS fluctuations between copresent devices as well as probe losses, we aggregate multiple copresence detection samples over a 1 minute window. For instance, for a probe frequency of 5 seconds, we require two users to be detected at the same location at least 3 times in our time window (empirically defined) to regard them as being copresent in that minute. Note that we opted for a 1 minute window to maintain sufficient granularity for detecting human encounters from copresent data.

In order to find the optimal dissimilarity threshold  $\delta$ , we construct a connectivity graph based on fingerprint similarities. If the distance  $d_{j,k}$  between a  $j$ -th and a  $k$ -th device is less than  $\delta$ , their corresponding nodes are connected by an edge in the graph, i.e.,  $C_{j,k} = 1$ . Otherwise, their nodes are disconnected, i.e.,  $C_{j,k} = 0$ . We then evaluated the algorithm’s performance for different values of  $\delta$  using a connectivity error metric. If two users and their devices  $j$  and  $k$  were in reality co-present, we define  $C'_{j,k} = 1$  indicating a true link between these devices, otherwise  $C'_{j,k} = 0$ . We calculate  $e_{j,k} = C_{j,k} - C'_{j,k}$  for every two pairs of nodes. The connectivity error  $e_{j,k} = -1$  indicates a false negative error and  $e_{j,k} = 1$  indicates a false positive error. From our training data, we analysed the connectivity errors for different values of  $\delta$ . Figure 9b suggests that the optimal threshold should be chosen between 3 and 3.5 – we used 3.25.

## CONTROLLED EVALUATION

### Dataset

In order to evaluate our framework, we scripted a set of use cases which were imitated using the previously described smartphones and the application. The scripted use cases corresponded to the scope and definition of the spontaneous encounters and fall within the three categories of scenarios described earlier.

For each of these scenarios we collected multiple variations of spontaneous encounters ranging from 2 up to 6 persons, for different durations. In total we collected nine different instances of these scenarios. Due to the space constraints, we

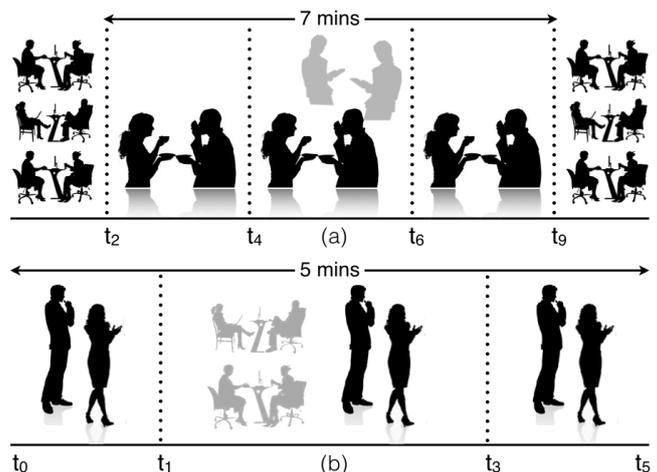


Figure 10: Two example scenarios of the controlled data collection

Scenario	Window size	False Positive
Queue with 4	2-5	0
Queue with 6	2-5	0
Queue with 8	2-5	0
Workspace	2	4
Workspace	3-5	2
Workspace	6+	0

Table 1: The number of false positives for the instances of scenarios which did not include any encounters.

will only describe two specific cases in detail assisted by Figure 10. The timeline indicates the composition and decomposition of encounters across individuals, whereas grayed out individuals are either copresent or passing by but do not join the encounters. In addition to these cases we also collected instances of scenarios of copresent individuals which do not correspond to encounters (e.g., people sharing a desk).

## Results

In this section we present the results of our encounter detection engine for the data collected through the controlled setting. We first analyse the validity of our algorithm by examining the cases which did not include any encounter (i.e., queue and workspace scenario). In so doing, we measure the number of false positives for each of these cases by varying the window size (Table 1). The results indicate the correctness of our approach in *not* detecting any encounters for the queue scenarios. However, in the case of the workspace scenario, we observe encounters detected for smaller window sizes. This is because in our online processing, we cannot distinguish between individuals joining (or departing) the workspace or an encounter, unless the copresence is longer than what we consider to be the boundary of an encounter (outside the shaded area of the Figure 1d). We discuss and address these cases in the evaluation of our live deployment in the In-the-Wild section, where the duration of individuals’ copresence in the open office resonates with the actual working day. Next we evaluated our inference engine accuracy

in detecting spontaneous encounters based on the rest of the collected scenarios described earlier. We evaluate our algorithm in terms of two aspects: (i) the detection accuracy, i.e., whether a ground-truth spontaneous encounter is detected; (ii) decomposition delay, i.e., how accurately our algorithm detected the duration of the spontaneous encounter.

**Detection Accuracy.** We evaluate the accuracy of our algorithm in detecting spontaneous encounters, by comparing the result of our inference engine against the ground-truth for all the scenarios. To this end, we set the parameter  $\alpha = 0.5$ , and  $th_{exit} = 0.75$  and measure precision and recall for a variable window size ( $\delta t_w$ ). Furthermore, we set the condition for a spontaneous encounter to  $f(W) = 1$ . Figure 11a illustrates the F-Score results for a varying window size versus an adaptive window size where the window size was determined based on our human interaction model. As shown in the figure, in case of the variable window size, the wider the window size the lower the F-Score. As  $\delta t_w$  gets bigger, the algorithm fails to detect short-lived spontaneous encounters which may happen between a few people. This is because the bigger  $\delta t_w$ , the longer duration each member is required to stay together for the spontaneous encounters to be detected ( $f(W) = 1$ ). This causes the short-lived spontaneous encounters to go undetected, thus lowering the recall. Smaller window sizes, however, also suffer from the opposite problem of low precision. This is because they also capture people who are accidentally passing by as participants of the encounters.

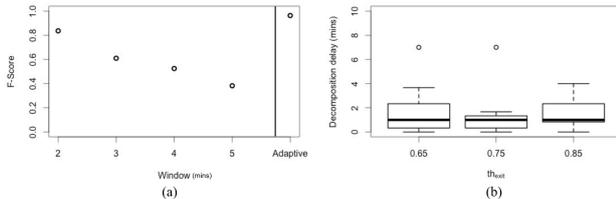


Figure 11: Framework performance: (a) F-Score for varying window size versus adaptive window size and (b) the decomposition delay.

Moving to the comparison with the adaptive window, our algorithm performs with an F-Score accuracy of 0.96 ( $precision = 1$ ,  $recall = 0.92$ ). As the number of copresent people increases, our window size is adaptively widened to reflect back on the longer membership duration that is needed for the given size to be considered as a spontaneous encounter. These are encouraging results, as we are able to detect the spontaneous encounters from a diverse set of scenarios with a very high precision and recall. However, the F-Score by itself cannot capture the validity of our model since it is equally important to detect an encounter’s decomposition (next to its formation).

**Decomposition Accuracy** In our experiments we first fix  $\alpha = 0.5$  and examined the effect of  $th_{exit}$ . By definition, the higher  $th_{exit}$  the more conservative our algorithm would be, that is the encounter decomposition would happen earlier. We found on average our approach had delta

decompositions of less than 2 minutes. The smallest decomposition delay was observed for  $th_{exit} = 0.75$  as illustrated in Figure 11b. As the parameters  $\alpha$  and  $th_{exit}$  exhibit an interaction, we also experimented with varying the  $\alpha$  parameter. The results were consistent and intuitive, i.e., as  $\alpha \rightarrow 1$  and more weight is given to the present, the sooner the encounters end resulting in negative delta decomposition time. As  $\alpha \rightarrow 0$  and more weight is given to the past, the longer it takes for  $S_t$  to decay causing the encounters to be detected for a longer duration they actually existed.

## IN-THE-WILD EVALUATION

In this section we report on the live deployment of our framework for 7 hours. The in-the-wild study started at 10am and involved 8 employees who share an open office space. In order to eliminate the challenges that are raised by different device manufacturers in this study, we asked the participants to carry the same devices as before. We then ran a diary-based study, where we asked the participants to record their encounters during the day. A major advantage of diary studies is that they bring the task of data collection into the respondents’ everyday world, however, a common disadvantage is that they rely on human recall and memory.

The participants were told to perform their daily work routines, carrying the phones we provided them. The participants were also asked to take their coffee break and their social chats near our coverage area to ensure they would be in the range of our network nodes. In our deployment setting, four of the participants shared a proximate desk area with each other, leading to a potential continuous copresence amongst them. In order to account for this situation, we tuned our encounter detection algorithm to learn the continuity of these copresences and ignore them as encounters. To this end, we added a parameter which corresponds to this continuity, and set its limit to 20 minutes. That is, if the same set of individuals have been copresent for longer than 20 minutes, we conclude they share a common sitting area. We set this limit to 20 minutes as it corresponds to the maximum duration of spontaneous encounters we are interested in – recall the shaded area of Figure 1d. For the evaluation, we also used the adaptive window and set  $\alpha = 0.5$  and  $th_{exit} = 0.75$ . The rest of the settings were consistent with those in the previous section.

We report and discuss the final results of our evaluation in terms of precision and recall of the overall framework, and do not dwell into the accuracy of separate components. We measure the recall in terms of the number of the *correctly* identified encounters out of the self-reported ones, and the precision as the *total* number of detected encounters out of the self-reported ones. Our ground-truth data based on the diaries corresponded to 17 distinct self-reported encounters between 8 participants. Given that these diaries suffer from imprecise timekeeping caused by human error, we allow a +/- 10 minutes offset from the actual time of the encounter to the reported time. That is, we consider a spontaneous encounter to be detected even if its composition time is shifted by 10 minutes compared to the reported ground-truth. The results of the in-the-wild study show that our framework was able to

detect 10 out of 17 encounters leading to 58% recall, however, our framework also detected additional 5 false encounters (15 in total) making the precision of our technique to 86%.

Looking at these results closely and comparing them to the ground-truth, we can see the success of our algorithm is in detecting the false positives that result in sharing a desk area, leading to high precision. However, the same factor causes the recall to suffer. That is, our framework fails to detect encounters that happen between those individuals who share a desk space, even if they go for a coffee together. While this is a limitation of our current approach, this can be mitigated through applying a filter in the network node, e.g., in the co-presence detection engine. The filter should detect a significant and almost identical fluctuation of RSSIs across a pair (or a set) of co-located devices (as proxies of humans) indicating that these devices are continuously preserving the same distance. This indication would be enough for the inference engine to adapt accordingly to deal with the phenomenon described above. Figure 12 depicts the duration of encounters in comparison to the self-reported ground-truth. Here, the x-axis presents the time of the day (in minutes) while the y-axis shows the duration of the encounter, both for detected and ground-truth encounters. Two observations can be made based on this figure. First, the horizontal distance between the detected and the ground-truth data shows that all the detected groups fall within 10 minutes offset from their ground-truth counterpart. Secondly, for 9 of these detected encounters the detected duration error was on average 1.5 minutes (median=1) – with an exception of an outlier at time 230 where the duration delta is 12 minutes.

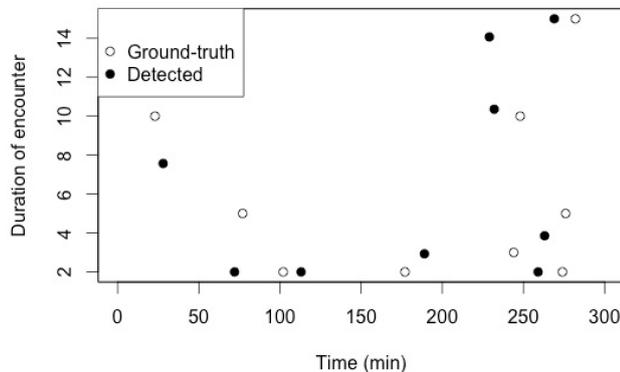


Figure 12: Delta delay in detection and decomposition time.

## RELATED WORK

A large body of work exploring WiFi-based localisation and co-localisation algorithms exists in literature [1, 28]. In our framework, we rely on a distance-based co-localisation algorithm borrowed from [6], whilst more advanced algorithms could be plugged in as well. Our contribution lays in the combination of such algorithm with a social model for inferring human encounters. Other works have also relied on Bluetooth to e.g., discover crowds [27], but studies indicate that few devices have Bluetooth turned on by default as compared to WiFi [3]. Also, due to Bluetooth frequency hopping, timely detections of mobile devices cannot be guaranteed for inferring short-lived human encounters. In general,

discovery-based tracking solutions result in each device having its own perception of a copresent group, lacking consensus. Passive tracking solutions such as [3, 21] leverage the network to determine users’ trajectories. These systems also struggle with deterministic device detections, albeit Musa et al. [21] have shown that tricks can be applied to increase the detection rate. However, deploying sniffing hardware in public areas remains a sensitive topic w.r.t. privacy given that it tracks users without their consent and therefore is forbidden by law in several countries. We advocate for solution in which users can opt in and opt out anytime.

Several ways of detecting human groups have also been introduced by computational social sciences [4, 17]. Olguín et al. proposed the sociometric badge, a wearable device that is used to measure face-to-face interactions in the workplace [22]. Brown et al. [2] tracked face-to-face interactions by using light-weight RFID tags which exploit short-range wireless sensing technologies to infer an interaction taking place. Other approaches have relied on quantifying interaction through cameras [5] and signals captured through mobile phones [10, 24] such as audio signals indicating the meta-linguistic contexts of conversation between individuals (i.e., turn-takings and pace) [18]. Indeed, to detect face-to-face interaction, high precision technologies such as infrared light with a transmission range between 2 to 4 meters and a narrow cone of 20 degrees are typically required [20].

## CONCLUDING REMARKS

In this paper, we presented a framework for detecting spontaneous human encounters. We measure human copresence from ubiquitous WiFi radio signals and feed this data to a model grounded upon social theory to infer face-to-face interactions. This framework was evaluated through controlled experiments accompanied by live deployment. While our initial results are encouraging, we acknowledge that our evaluation was limited in terms of number of participants and device variety. In particular, we expect congestion to happen in the case of denser deployments (with hundreds of individuals) where a large number of probes would share the same channel simultaneously. Indeed alternative robust techniques such as those that incorporate randomisation could be developed as part of our framework. We also plan to assess the impact of different device models and antenna designs.

Furthermore, while no modifications to a phone are needed, a present limitation of our approach is that it depends on a small application running on a user’s personal device to guarantee the detection of the device in a limited time window. We envision that this component could be part of future mobile operating systems, as to give developers control over WiFi probes and offering users with the option to indicate what they may be used for (i.e., opt in/out to particular services). The major advantage of our system lays in the fact that it leverages existing network infrastructure with minimal energy expenditure (zero sensing) on mobile devices. We have shown that even without highly accurate copresence data (e.g., sub meter), we can detect human encounters with high accuracy thanks to our inference model taking social patterns into account.

## REFERENCES

1. Wain Sy Anthea Au, Chen Feng, Shahrokh Valaee, Sophia Reyes, Sameh Sorour, Samuel N. Markowitz, Deborah Gold, Keith Gordon, and Moshe Eizenman. Indoor tracking and navigation using received signal strength and compressive sensing on a mobile device. *IEEE Trans. Mob. Comput.*, 12:2050–2062, 2013.
2. Chloë Brown, Christos Efstratiou, Ilias Leontiadis, Daniele Quercia, Cecilia Mascolo, James Scott, and Peter Key. The architecture of innovation: Tracking face-to-face interactions with ubicomp technologies. In *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing*, pages 811–822, 2014.
3. Yohan Chon, Suyeon Kim, Seungwoo Lee, Dongwon Kim, Yungeun Kim, and Hojung Cha. Sensing wifi packets in the air: practicality and implications in urban mobility monitoring. In *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing*, pages 189–200, 2014.
4. Tanzeem Choudhury and Alex Pentland. Sensing and modeling human networks using the sociometer. In *Proceedings of the International Symposium on Wearable Computers*, pages 216–216, 2003.
5. Roy D, Patel R, DeCamp P, Kubat R, Fleischman M, Roy B, Mavridis N, Tellex S, Salata A, and Guinness J. The human speechome project. In *28th Annual Meeting of the Cognitive Science Society*, 2006.
6. Marzieh Dashti, Mohd Amiruddin Abd Rahman, Hamed Mahmoudi, and Holger Claussen. Detecting co-located mobile users. In *Proceedings of the International Conference on Communications*, 2015.
7. Manlio De Domenico, Antonio Lima, and Mirco Musolesi. Interdependence and predictability of human mobility and social interactions. *Pervasive and Mobile Computing*, 9(6):798–807, 2013.
8. Anind K. Dey, Katarzyna Wac, Denzil Ferreira, Kevin Tassini, Jin-Hyuk Hong, and Julian Ramos. Getting closer: An empirical investigation of the proximity of user to their smart phones. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, pages 163–172, 2011.
9. Jean-Pierre Eckmann, Elisha Moses, and Danilo Sergi. Entropy of dialogues creates coherent structures in e-mail traffic. *Proceedings of the National Academy of Sciences of the United States of America*, 101(40):14333–14337, 2004.
10. Christos Efstratiou, Ilias Leontiadis, Marco Picone, Kiran K Rachuri, Cecilia Mascolo, and Jon Crowcroft. Sense and sensibility in a pervasive world. In *Pervasive Computing*, pages 406–424. 2012.
11. Erving Goffman. *Interaction ritual: Essays in face to face behavior*. AldineTransaction, 2005.
12. Erving Goffman. *Behavior in public places*. Simon and Schuster, 2008.
13. George C Homans. *The human group*, volume 7. Routledge, 2013.
14. Mikkel Baun Kjærgaard, Sourav Bhattacharya, Henrik Blunck, and Petteri Nurmi. Energy-efficient trajectory tracking for mobile devices. In *Proceedings of the 9th International Conference on Mobile Systems, Applications and Services*, pages 307–320, 2011.
15. Gueorgi Kossinets and Duncan J Watts. Empirical analysis of an evolving social network. *Science*, 311(5757):88–90, 2006.
16. Nicholas D Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T Campbell. A survey of mobile phone sensing. *Communications Magazine, IEEE*, 48(9):140–150, 2010.
17. David Lazer, Alex Sandy Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, et al. Life in the network: the coming age of computational social science. *Science*, 323(5915):721, 2009.
18. Youngki Lee, Chulhong Min, Chanyou Hwang, Jaeung Lee, Inseok Hwang, Younghyun Ju, Chungkuk Yoo, Miri Moon, Uichin Lee, and Junehwa Song. Sociophone: Everyday face-to-face interaction monitoring platform using multi-phone sensor fusion. In *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*, pages 375–388, 2013.
19. Adriana M Manago, Tamara Taylor, and Patricia M Greenfield. Me and my 400 friends: the anatomy of college students’ facebook networks, their communication patterns, and well-being. *Developmental psychology*, 48(2):369, 2012.
20. Claudio Martella, MV Steen, AV Halteren, Claudine Conrado, and Jie Li. Crowd textures as proximity graphs. *Communications Magazine, IEEE*, 52(1):114–121, 2014.
21. A. B. M. Musa and Jakob Eriksson. Tracking unmodified smartphones using wi-fi monitors. In *Proceedings of the 10th Conference on Embedded Network Sensor Systems*, pages 281–294, New York, NY, USA, 2012.
22. Daniel Olguín, Benjamin N Waber, Taemie Kim, Akshay Mohan, Koji Ara, and Alex Pentland. Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 39(1):43–55, 2009.
23. J-P Onnela, Jari Saramäki, Jorkki Hyvönen, György Szabó, David Lazer, Kimmo Kaski, János Kertész, and A-L Barabási. Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18):7332–7336, 2007.

24. Daniel Roggen, Martin Wirz, Gerhard Tröster, and Dirk Helbing. Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods. *arXiv:1109.1664*, 2011.
25. John C Turner. Social categorization and the self-concept: A social cognitive theory of group behavior. *Advances in Group Processes*, 2:77–122, 1985.
26. Bimal Viswanath, Alan Mislove, Meeyoung Cha, and Krishna P Gummadi. On the evolution of user interaction in facebook. In *Proceedings of the Workshop on Online Social Networks*, pages 37–42, 2009.
27. Jens Weppner and Paul Lukowicz. Bluetooth based collaborative crowd density estimation with mobile phones. In *Proceedings of the 11th International Conference on Pervasive Computing and Communications*, pages 193–200, 2013.
28. Zheng Yang, Chenshu Wu, and Yunhao Liu. Locating in fingerprint space: Wireless indoor localization with little human intervention. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, pages 269–280, New York, NY, USA, 2012.
29. Shanyang Zhao. Toward a taxonomy of copresence. *Presence: Teleoperators and Virtual Environments*, 12(5):445–455, 2003.