Analysis of Social Interactions Through Mobile Phones

**Aleksandar Matic · Venet Osmani ·
Oscar Mayora-Ibarra**

Abstract Equipment of mobile phones with various kinds of sensors is transforming these devices from mere capabilities of voice and internet access to devices capable of sensing a number of phenomena pertaining to their users. In this paper we make use of these capabilities of phones to detect social interactions between people and analyze social context by using embedded sensors found in typical smart phones. Work carried out in this area has typically used dedicated hardware to establish social interactions, and we contend on the suitability of mobile phone, since additional devices that user is not familiar with influence natural user behavior and thus their social interaction patterns. Our work shows that two parameters that can be detected through mobile phone sensing, namely interpersonal distance and relative body orientation, provide a solid basis for inferring social interactions. We describe how these factors are acquired using smart phones and describe our analysis. The experiments demonstrate that we can recognize not only whether a social interaction is taking place, but also the type of social interaction, distinguishing between formal and informal social settings.

Keywords social interaction analysis · nonverbal behavior · small group interactions · proximity detection · informal and formal communication

1 Introduction

Smart phones are becoming smarter, not just from the point of view of the increasingly complex services they can offer,

A. Matic (✉) · V. Osmani · O. Mayora-Ibarra
Center for Research and Telecommunication Experimentation for
Networked Communities, CREATE-NET,
Via alla Cascata 56/D,
38123 Povo, Trento, Italy
e-mail: aleksandar.matic@create-net.org

but also from the aspects of information they provide about their users. Recent off-the-shelf phones come equipped with an array of sensors that have enabled an "awareness" of our daily routines. With respect to social interactions, mobile phones can track precise logs of our phone calls, short massages, emails and activities in virtual networks. However, when it comes to face-to-face social interactions, mobile phones still have a limited awareness.

The work carried out in detecting social interactions typically has relied on having a dedicated device, equipped with sensing modalities designed to allow monitoring of user behavior. However, having dedicated hardware and asking subjects to wear it, introduces its own set of issues including subject stigmatization, resulting in behavior change. This occurs because wearing unfamiliar and visible sensing hardware increases the awareness of being monitored, thus affecting the natural behavior of the subjects. On the other hand, mobile phones are already ubiquitous devices that have been adopted faster than any technology in human history [1]. Therefore, the monitoring behavior process through mobile phones fades into the background, having a minimal effect on the users' behavior and consequently their social interaction patterns. However, the challenge is how to address monitoring of specific activities relying on existing sensing technologies that are embedded in mobile phones, which is the issue not encountered when using purpose-manufactured devices which already have dedicated sensors incorporated.

Monitoring of social interactions using mobile phones is typically based on sensing proximity or on detecting speech activity. A frequently applied approach for inferring social activity through the detection of proximity relies on the use of Bluetooth [2], [3]. Since the Bluetooth communications range is in the order of ten meters, this approach provides only a coarse spatial granularity in recognizing interpersonal distances; therefore, the knowledge about proximity

between individuals is used to model the dynamics of social interactions at large scale rather than detecting each single social encounter which takes place at small spatio-temporal scale. As an alternative, Wyatt et. al [4] proposed the method of extracting audio data features using microphones from a pair of co-located mobile phones, in order to detect who was speaking and when thus detecting face-to-face interactions. The algorithm does not capture raw audio data but a set of features which does not contain verbal information. However, the microphone-based approaches are sensitive to false positives as nearby conversations can be unintentionally picked up. In addition, activating microphone typically entails compromising privacy and ethical issues – in a number of situations (for example, in public spaces) audio data cannot be obtained due to legal or ethical norms [5].

This paper provides a solution that uses non-auditory sensors embedded in the current smart phones to detect the occurrence of social interactions which occur on a small spatio-temporal scale. Our work shows that two parameters that can be inferred through mobile phone sensing, namely interpersonal distance and relative body orientation, provide a solid basis for monitoring social interactions. Furthermore, we demonstrate high predictive power of spatial parameters to detect social context of face-to-face interactions, that is formal or informal as perceived by the subjects.Therefore, the system proposed in this paper has the potential to generate rich and large-scale data considering the ubiquitous nature of mobile phones while not relying on sensitive data. The fact that people habitually carry the mobile phone makes this device an ideal tool for unobtrusive and continuous monitoring of individuals' behavior. The goal is to provide a tool for acquiring a better insight into social activity of subjects and the contexts of individual social interactions thus to potentially support the research in social networks analysis and the investigation of formal/informal structures.

The paper is organized as follows. Section 2 provides a review of the current work on mobile phone-based sensing of social interactions. In Section 3 we present our system that can be widely deployed as a mobile phone application in order to infer interpersonal spatial settings. Then, in Section 4 and Section 5 we demonstrate how the system can be used to detect the occurrence of social interactions and the social context. Finally, we provide a summary in Section 6.

2 Related work

Smart-phones have been proposed as an alternative to using dedicated hardware or external infrastructure for gathering social interaction data. Monitoring of social interactions through smart phones typically rely on detecting proximity and on audio analysis.

Using Bluetooth as a proximity sensor to reconstruct social dynamics at large scale has been extensively investigated under the umbrella of reality mining initiative [6, 7, 1]. MIT Media Lab's Reality Mining project launched in 2004 with the goal of sensing complex social systems which included inferring patterns in daily user activity, relationships, socially meaningful locations, and organizational structures [6]. Along the same line, Raento et al. [8] were one of the first who proposed mobile phone data collection for large-scale context sensing. More recent algorithm for identifying social groups and inferring frequency/duration of meetings within each group was proposed by Mardenfeld et al. [9] who tested their approach on the Reality Mining dataset. In addition to modeling the patterns of person-to-person interactions, Do and Gatica-Perez [10] showed that it is possible to infer different interaction types using a probabilistic model applied on longitudinal Bluetooth data. However, Bluetooth scans indicate the presence of nearby devices in a radius of 10 m, which does not provide sufficient information to detect an ongoing social interaction which takes place on a small spatio-temporal scale; rather, such an approach is used to model the longitudinal dynamics of social interactions.

In order to address the limitation of Bluetooth scan to detect actual face-to-face proximity between subjects, the Virtual Compass project [11] estimates interpersonal distances using RSSI analysis of Bluetooth and Wi-Fi signals. By applying empirical propagation models, the approach achieves the median accuracy between 0.9 m and 1.9 m while also detecting position of subjects in 2D plane. However, the lack of subjects' orientation information might not be sufficient for modeling the occurrence of face-to-face social interactions.

The fact that co-location of subjects does not always imply their interaction [4] raises the requirement for alternative methods to detect voice conversation using microphone embedded in mobile phones. A number of approaches associated the level of detected ambient noise to social interactions [12], while more advanced methods included analysis of voice segments for each pair of microphones thus identifying conversations among two or more subjects [4]. However, the limitations of microphone-based approaches include: 1) sensitivity to false positives since the conversations occurring in close proximity of the monitored subjects in which they are not involved, can be incorrectly classified, 2) activating microphone can face ethical issues and can negatively affect the perception of privacy in subject.

In contrast to previous studies, we propose the mobile phone-based solution for monitoring social interactions which occur on small spatio-temporal scale, without relying

on sensitive data. Our work demonstrates that by using sensing capabilities available in mobile phone it is possible to detect not only whether a social interaction is taking place, but also the type of social interaction, distinguishing between formal and informal social settings.

3 Estimating interpersonal distance and relative body orientation

3.1 Distance estimation

3.1.1 Overviewof our approach

Our concept for estimating distance between two mobile phones is based on the RSSI analysis which has been already shown to be a promising solution. In contrast to the Virtual Compass [11] (which is, to our knowledge, the only similar approach), our method of building a generic empirical model to map RSSI values to distance regardless of the phone used, relies on supervised learning thus trading-off the user effort in signal fingerprint collection for the accuracy in distance estimation. The reason for using a more costly method in terms of the end user effort is the fact that one of the pre-dominant factors affecting RSSI patterns is the receiver's characteristics [13] whose capturing can lead to a higher system's accuracy. This hypothesis was tested in the experiments that follow, demonstrating that environmental factors have less prevailing impact on RSSI patterns than receiver's characteristics due to relatively short distances and no obstacles between receiver and transmitter. Unlike time-consuming measurements typically required for fingerprinting methods, the user effort will be decreased to only a couple of minutes to calibrate the phone signal, while achieving a comparable accuracy to full fingerprinting method. The concept for estimating distance is tested using Wi-Fi signals. Nevertheless, other radio transmitting/receiving mechanisms with accessible RSSI values (such as FM or Bluetooth) available on mobile phones could be used for the same purpose or in combination with Wi-Fi.

3.1.2 Estimating distance between two mobile phones using Wi-Fi RSSI

Similar to indoor positioning systems that use fingerprinting technique, our method for distance estimation is based on analyzing RSSI values, observed on an unknown distance from the phone which transmits Wi-Fi signal (colloquially known as Wi-Fi Hotspot or Personal Hotspot). The distance is estimated by applying the model built using a database that matches RSSI values (fingerprints) with the actual distances.

To acquire the training set we used two smart phones, one in transmitting (tethering) mode, the other in receiving (client) mode, to carry out Wi-Fi signal measurements. We used different distances following a grid of 0.5 m thus collecting patterns of RSSI in the database that was used as training set. The transmitting power of 0 dBm provided the smoothest and the most monotone RSSI to distance dependency thus proving to be the best fit for short distance estimation (Fig. 1). However, it can be seen that the RSSI demonstrates the instability and fluctuations of the Wi-Fi signal, typically due to environmental factors. Therefore, the distance estimation approach based on a simple RSSI threshold analysis (assigning ranges of RSSI values to corresponding distances) did not suffice which led to applying machine-learning techniques for distance estimation

Our testbed consisted of six smart phones (with Android operating system) including three different models, namely *HTC Desire*, *Samsung Nexus S* and *HTC Nexus One* that used a modified firmware to allow adjustment of transmitting power. Different phone units were distinguished by MAC addresses. Measurements were taken in three offices with dimensions of *12×8 m*, *6×5 m* and *6×3 m*, a balcony of *12×2.5 m* and a meeting room of *10×8 m*. For testing the system's accuracy we used a pair of phones – one in transmitting and the other one in receiving mode.

Figure 2 shows the system's accuracy when using the same phone (i.e. same model) acting as a receiver in both training and test phase. The results correspond to Naïve Bayes classifier with Kernel Density Estimation (KDE) and Gaussian Process (GP) regression, however several other techniques were tested performing similarly (including linear classification and SVM). The median estimation error (50th percentile) of approximately 0.5 m was achieved using
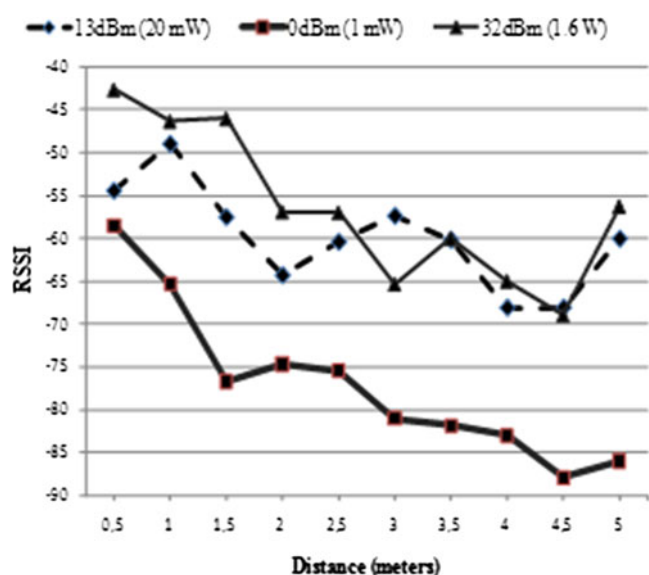

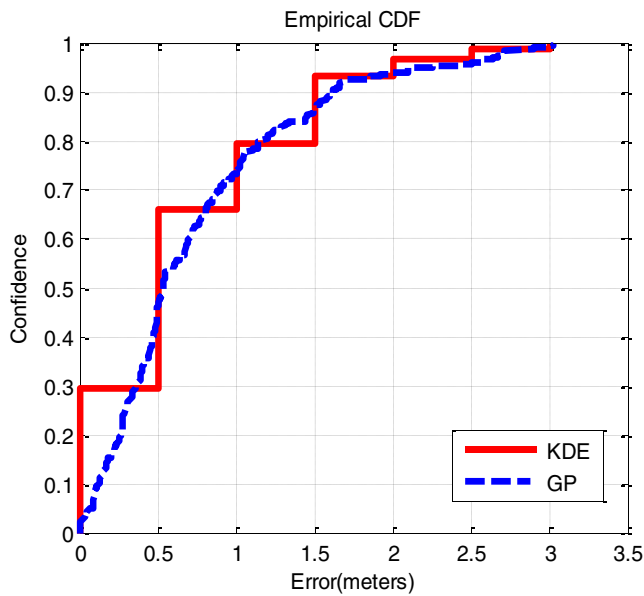**Fig. 1** Dependence RSSI on the distance (different power levels)

Fig. 2 Distance estimation accuracy (same receiving phone for training and test phases)

both machine learning techniques. Naive Bayes with KDE showed a slightly better overall performance, providing distance estimation with a 50 % percentile error of 0.5 m and 95th percentile error of approximately 2 m.

3.1.3 Fast calibration method

When *different* phone models were used for training and test phase, the system's accuracy significantly degraded (Fig. 3). This is due to the fact that RSSI patterns highly depend on
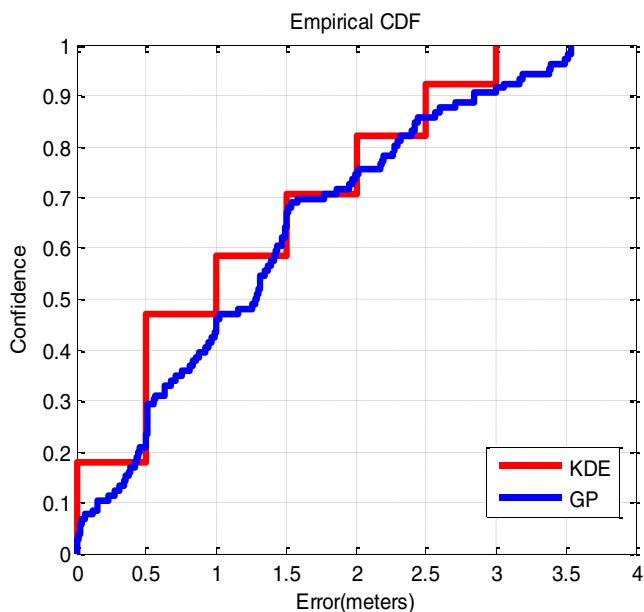
the receiver characteristics [13] which are likely to be different across different phone models. In order to tackle this problem while avoiding the repetition of RSSI measurements which would be laborious and time-consuming, we "calibrated" only one point by measuring RSSI for a couple of minutes on a fixed distance of, for instance, 1 m. Once the RSSI is captured, the training set is estimated applying the following propagation model [14]:

$$P(d)[dBm] = P(d_o)[dBm] - 10n \log\left(\frac{d}{d_0}\right)[dBm] - X[dBm]$$

$$(1)$$

where n is the path loss exponent, $P(d0)$ is the signal power at the reference distance $d0$ from the transmitter phone (in our case 1 m) and d is the distance in which RSSI is estimated by applying the model. X is a component that reflects the sum of losses induced by each wall between the transmitter and receiver. We have found empirically that the best suited value for the coefficient n is 1.5, while for X is zero (there are no walls or other obstacles between points). Figure 4 shows a cumulative distribution function of distance estimation errors (for each of the phone model that we tested applying cross-validation method across different environments), showing again 50th percentile error of 0.5 m and 95th percentile error between 2 m and 2.5 m (for each of the phone model that we tested applying cross-validation method across different environments). Gaussian processes regression achieved also median error of 0.5 m while 95th percentile of 2.5 m but for brevity reasons we presented only classification method (Bayesian with KDE). Therefore, by investing a minimal effort of performing the
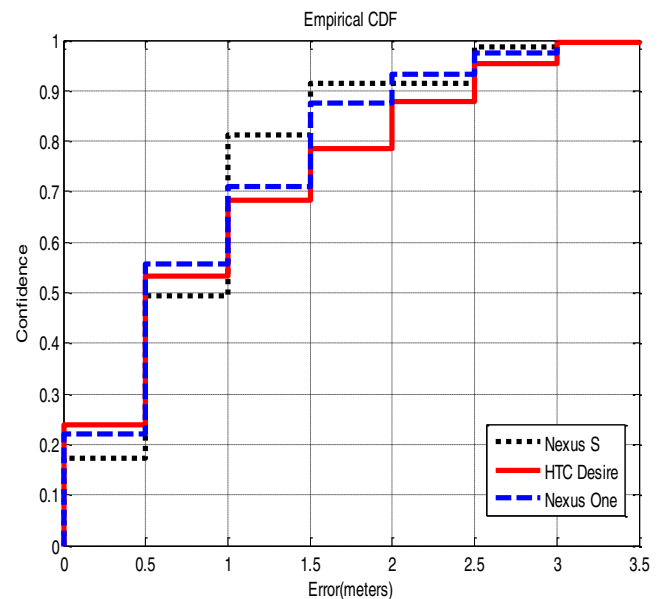

Fig. 3 Distance estimation accuracy (different receiving phone for training and test phases)

Fig. 4 Distance estimation accuracy using training set generated by applying propagation method

calibration for a few minutes at a single distance, we demonstrated that it is possible to achieve similar performance as in the case of acquiring a full training set (Figs. 2 and 4).

Calibrating the phone and testing in the same environment provided similar accuracy as in the case of performing calibration and testing in different environments (which was evidenced across all six environments). This may be indicative that of the pre-dominant factor that influences RSSI pattern lies in receiver's characteristics. Less prevailing impact of environmental conditions may be explained by relatively short distances and no obstacles between receiver and transmitter which could affect the signal propagation. This was further evidenced in the experiments of real-life settings conducted in a wide array of environments and phone models which will be presented in Section 4 and Section 5.

To recap, in comparison to the existing solutions based on mobile phone sensing, our system provides a higher accuracy in estimating distance between phones, does not require communication between devices and broadcasting the distance to each of peers, while the training phase is facilitated with a fast calibration method which makes the approach adaptive to different applications, environments and phone models.

3.2 Estimating relative body orientation

Relative body orientation refers to the angle between the orientations of torsos [15] considering two subjects that are facing each other. To recognize the relative body orientation of subjects carrying mobile phones, we use the embedded orientation sensor that provides the following values (expressed in degrees): Azimuth – the angle between the magnetic north direction and the y-axis, around the z-axis (0° to 359°); Pitch – the rotation around x-axis (−180° to 180°) with positive values when the z-axis moves towards the y-axis; and Roll – the rotation around y-axis (−90° to 90°) with positive values when the x-axis moves towards the z-axis. Knowing the relative position between the body and the phone orientation is a fundamental condition in order to recognize the individual's body orientation and the relative body orientation between subjects. Once this relationship is determined, calculating relative body orientation would require relative processing of *azimuth*, *pitch* and *roll* values. In our experiments, we were always aware of the exact position where participants carried the phone. However, in their recent study, Shi et al. [16] demonstrated that it is possible to automatically detect on-body position of the mobile phone by utilizing the fusion of accelerometer and gyroscope.

In this section, we evaluate the proposed concept for estimating the distance between mobile phones and described the use of compass for extracting relative body orientation between subjects. Distance estimation accuracy

results were consistent for three tested phone models across six environments; evaluating performance for other phone models is out of the scope of this paper, however we do not expect large disparities based on the phone models already tested. In the sections that follow we assess whether the extracted parameters are sufficiently accurate for detecting social encounters.

4 Detecting social interaction occurences

Interpersonal distances, typically held in social interactions, are investigated through the study of proxemics [17] According to this study, interpersonal distances include the following metrics: *intimate distance* (close: 0–0.15 m, far: 0.15–0.45 m), *personal distance* (close: 0.45–0.76 m, far: 0.76–1.2 m), *social distance* (close: 1.2–2.1 m, far: 2.1–3.6 m) and *public distance* (close: 3.6–7.6 m, far: 7.6 and more). These four categories of interpersonal distance are typically used for the following activities: *intimate distance* for embracing, touching or whispering; *personal distance* for interactions among good friends or family members; *social distance* for interactions among acquaintances; and, *public distance* used for public speaking. When it comes to the relative body orientation, two people may hold any of the relative orientations between 0° and 180° during social interaction. The relative body orientation is often used in studies to describe the immediacy of interaction, subject's attitude or similar phenomena in social interactions rather than to recognize whether the interaction exists or not [15]. However, Groh et al. [15] demonstrated that interpersonal distance and relative body orientation together provide sufficient evidence to infer the occurrence of social interactions. The authors recognized the two parameters relying on a highly precise commercial camera system (with the accuracy of <1 mm and <1°) installed in a room of 3×3 m, and the reported accuracy in detecting the occurrence of social interactions was approximately 80 %. In this work we analyze the potentials of mobile phone sensing to recognize social interactions based on the two parameters, interpersonal distance (denoted with d) and the relative body orientation (denoted with α), detected in an unobtrusive way while not confining the experimental settings to one room.

Detection of social interactions that will be presented in this section is based on analyzing spatial parameters between a pair of subjects that carry mobile phones. If more than two subjects are involved in the same conversation, our method recognizes other participants by examining information for each pair of individuals involved in the social interaction. On the other hand, as the number of participants increases, interpersonal distances expand and the angles become wider, thus putting constraints on developing a single model of social interactions, regardless of the number

of participants. However, these effects (such as changes in angles) are typically neglected in the literature since practical experience suggests that when there are more than four or five individuals, they frequently split up into sub-groups [15] [18]. Therefore, the experiments that follow were conducted under the assumption that in real-life setting, the number of individuals which actively participate social interaction is limited to four of five (usually referred to as a "small-group" interaction) [15] [18].

We chose time frame of 10 s to process data as suggested by [19] in order to capture dynamic changes in social interactions while at the same time to discriminate between existing and non-existing social interactions. Therefore, interpersonal distances were estimated using a sequence of Wi-Fi RSSI values for every 10-second frame while body orientation is averaged for every 10 s (i.e. 10 samples). Relative body orientation of subjects was considered only if the standard deviation of the samples was less than or equal to 10° for each subject (regarding the 10-second time frame), otherwise the current frame of samples was left out. This was done in order to analyze situations in which subjects held stable relative orientation, such that random body movements are removed as a source of orientation uncertainty. The threshold of 10° was confirmed to be a trade-off between decreasing the standard deviation of the estimated relative body orientation (proportional to decreasing threshold) and decreasing the amount of discarded data (proportional to increasing threshold). Overall, approximately 20 %–25 % of unstable orientation data was discarded. We installed the application in five phones, two *HTC Desire*, two *HTC Desire S* and one *Samsung Galaxy S* with synchronized clocks to ensure correct data aligning which was important, considering a short time frame of 10 s for data analysis. Focusing on small-group of co-located face-to-face social interactions, we performed the experiments in four types of scenarios:

4.1 Experiments

4.1.1 Controlled experiments

Participants, that partially knew each other, were asked to communicate for an amount of time of their choice, while carrying the mobile at a known place. The first trial involved 6 participants (4 males, 2 females, age: 31±4 years) that were talking to each other, maximum four at a time, at 14 randomly selected locations, including 12 indoor and 2 outdoor environments. The duration of these interactions was 5.6±3.8 min. The second trial was conducted in a meeting room and it consisted of two 15-minute sessions in each involving 4 people (6 males, 2 females, age: 29±4 years) who were let to communicate freely as they wanted. This experimental trial resulted in 1300 pairs of relative body orientation and

interpersonal distance (α, d) with a time frame window of 10 s for processing and averaging data.

4.1.2 Break room settings

The break room is the place where employees in our research center typically socialize. This created the opportunity to monitor social interaction in a natural setting. When people were coming to the break room, we asked them to place the phone in a case attached on the right hip and to continue their interaction. Overall, we recorded 15 interactions of duration 6.2±3.5 min that included 24 different people. This experimental trial resulted in 1300 (α, d) pairs.

4.1.3 Continous monitoring

Aiming to analyze social interactions in continuous settings, the third trial of experiments was performed during working time for one week i.e. 5 working days, involving 5 colleagues that share the same office. They were asked to provide a label whenever social interactions occurred outside of the office, through a button press on the phone. Overall, during one week of measurements there were 9 standing social interactions labeled which involved either all of 5 participants or their subset. The locations were random in the building with duration of 6.4±8.1 min, resulting in 900 (α, d) pairs. However, the fact that one week of measurements resulted in 9 labeled standing conversations was questioning and it turned out that the participants did not label several social encounters which they clarified at the end of experiments.

The results shown in Fig. 5 indicate that social interactions between two subjects were centralized around 180° – the relative body orientation that corresponds to a perfect face-to-face position. Wider range of relative orientations was perceived in the cases of break room and continuous monitoring settings (both with SD of 45°) in comparison to the controlled experiments (with SD of 25°). This may reflect the fact that participants were more relaxed and held less steady orientation when they were participating in break room social interactions, in comparison to social interactions where participants were instructed to communicate. It can be seen that participants were mostly having shorter interpersonal distances in the break room and continuous monitoring, which were both natural setting.

4.1.4 Non-existing social interaction

In order to assess the potential of using spatial (α, d) parameters to distinguish existing and non-existing social interactions, it was necessary to create also a solid corpus of the pairs that do not correspond to social interactions. Four subjects that attended a fair called "Researchers Night" were monitored
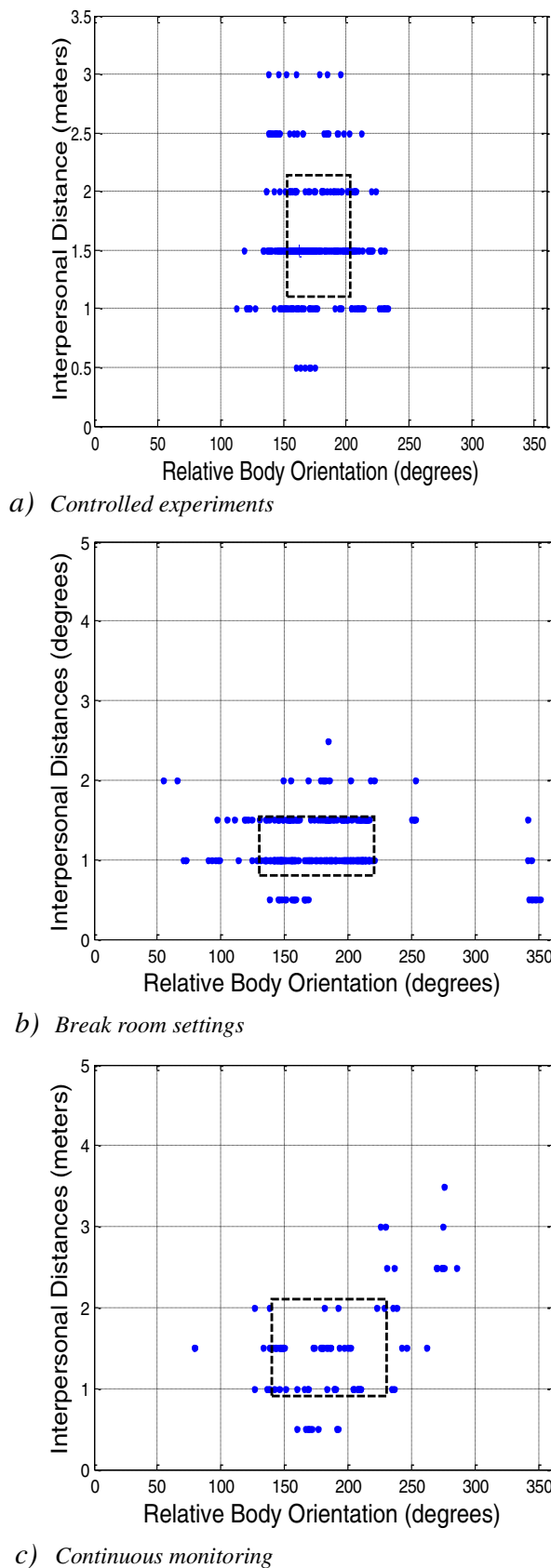
a) *Controlled experiments*

b) *Break room settings*

c) *Continuous monitoring*

Fig. 5 Analyzing social interactions through the relative body orientation (α) and the interpersonal distance (d)

while being asked to report any social encounter among them. Measurements from one-hour period in which they reported no social interactions was extracted as a suitable data set containing overall 1400 (α, d) pairs for creating non-existing social interaction corpus; being at the stand implied their constant proximity and random relative body orientations (while sitting/standing/moving) – Fig. 6. In addition, there were added measurements from previously described controlled experiments which included subjects that were in concurrent social interactions and in a close proximity (all social encounters occurred within 5×5m space).

4.2 Results

Table 1 presents the results of distinguishing occurrence of social interactions (denoted as SI) and situations when no social interaction occurs (denoted as NonSI) represented with a feature-vector (α, d) by applying Linear Classification and Naïve Bayes with KDE techniques. The classification performance was evaluated using 10-fold cross validation.

The results demonstrate the accuracy of 74 % in detecting social interactions based on interpersonal distance and relative body orientation. Naïve Bayes with KDE performed slightly better in identifying social interaction pairs while Linear Classifier provided lower rate of false positives. A contributing factor to this performance is also a simple method of taking out of the consideration (α, d) pairs corresponding to the situations in which subjects did not hold a stable relative body orientation, thus eliminating the source of uncertainty created in most cases by random body movements.
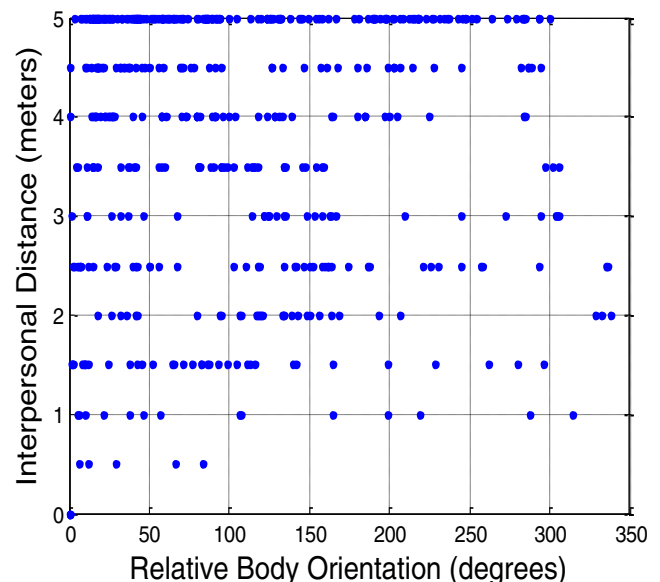

Fig. 6 (α, d) pairs corresponding to situations without social interactions taking place

Table 1 Classification results (orientation, distance)

	Naïve Bayes (KDE)		Linear Classifier	
	SI/ NonSI		SI/ NonSI	
SI	74 %	26 %	72 %	28 %
NonSI	24 %	76 %	22 %	78 %

However, instead of using the standard deviation (SD) of relative body orientation for identifying "unstable" (α, d) pairs, we attempted to use it also as a classification feature that can be considered as an index of holding stable relative position of participants in a social interaction. SD of relative body orientation (denoted with σ) was also calculated for each 10-second frame (i.e. for 10 samples) and combined with distance d and averaged relative body orientation α, constituting 2-feature vector (σ, d) and 3-feature vector (σ, α, d). Table 2 shows the results of 10-fold cross validation.

The combination of interpersonal distance and SD of relative body orientation provided higher accuracy in comparison to the previous case of using relative body orientation angle (Table 2). This may be due to the fact that feature-vector (σ, d) does not discriminate classes based on the absolute angle between body orientations in social interactions thus allowing for more situations to be included in the model in comparison to feature-vector (α, d). As expected, this resulted in a higher rate of false positives that occurred mostly when subjects were in a close proximity, having a stable body orientations but not interacting (for instance, sitting or being in concurrent social interactions). The highest accuracy was achieved using 3-feature vector (σ, α, d) that resulted in 89 % of successfully classified vectors corresponding to social interactions and 26 % of false positives.

The results demonstrate that the accuracy of estimating interpersonal distances and relative body orientations achieved with mobile phone sensing was sufficiently discriminative to identify social interactions. Note that the position of the phone does not affect the standard deviation (SD) of relative body orientation, thus the model based on 2-feature vector (σ, d) does not require users to carry the phone on a pre-defined/known position on the body.

5 Analysis of type of social interactions

5.1 Background

Monitoring of social interactions has a particular application in workplace, where socialization patterns can be used to influence workplace policies towards a more efficient work environment, since high complexity information is mostly exchanged through face to face interaction [20]. Various studies investigated methods of improving communication channels to enable more efficient knowledge transfer between employees. While most of the outcomes suggested the promotion of informal type of communications [21, 22], several investigations claim the opposite, arguing that formal interactions are an efficient knowledge transfer strategy [23]. However, there is a general consensus that improving communication channels used by knowledge workers requires a deeper understanding of both formal and informal types of interactions [21, 23, 24]. The difficulty in monitoring and measuring informal/formal networks was identified to be a key challenge towards making substantial steps in the efficient information transfer and consequently for increasing productivity in knowledge-driven communities [24]. Therefore, this section evaluates the potential of using our system to indicate the type of social interaction once its occurrence is already detected on small spatio-temporal scale, as elaborated in the previous section. Although the term *type* of social interaction may include various connotations (such as competitive, cooperative, decision making, and other types of conversation), it is used here to denote formal or informal context. Interpersonal distances and relative body orientations that can be extracted using the proposed system are evaluated regarding the predictive power in classifying between formal and informal type of interaction.

The main postulates of the proxemics study [17] suggest that people unconsciously organize the space around them, corresponding to different degrees of intimacy. It is even intuitively known that having a chat with a close friend, talking to the boss or talking to the queen differ in spatial settings conventions i.e. interpersonal distances are affected by level of formality in social interaction. Furthermore, according to social psychology, the formality is bound by roles and hierarchies among participants [21] which is further mirrored in spatial arrangements. The matching

Table 2 Classification results (2-feature and 3 feature vectors)

	(σ, d)				(σ, α, d)			
	Naïve Bayes (KDE)		Linear Classifier		Naïve Bayes (KDE)		Linear Classifier	
	SI/ NonSI		SI/ NonSI		SI/ NonSI		SI/ NonSI	
SI	79 %	21 %	74 %	26 %	89 %	11 %	87 %	13 %
NonSI	31 %	69 %	28 %	72 %	26 %	74 %	21 %	79 %

between social relations and the spatial formations in social interactions was recently investigated using computer vision system for estimating distances among subjects, confirming strong positive correlation [17]. Therefore, we opted for interpersonal distance in the attempt to classify between formal and informal social context.

Regarding spatial settings detection, the proposed system allows measures of relative body orientation and its standard deviation (as an index of stable relative body position between participants) that demonstrated high predictive power of detecting social encounter occurrence. Social psychology literature does not directly associate body orientations and the degree of formality in conversations. However, the relative body orientation is often used in studies to describe the immediacy of interaction, subject's attitude or similar phenomena in social interactions [31]. Therefore, it is hypothesized that the body orientation related cues (namely relative body orientation and its standard deviation) might also correlate with the level of formality thus being selected as suitable parameters aiming to formal versus informal interaction classification.

5.2 Description of experiments

Experiments in analysis of type of social interactions were conducted in a number of locations, including three meeting rooms, three offices, three coffee rooms, two balconies and an entrance hall with dimensions that did not physically confine subjects (the dimensions of the smallest room were 5×4 m) thus not affecting interpersonal distances. We interrupted face to face communications that were about to occur or were already in progress and for subjects that accepted the participation in experiments, we provided smart phones that were broadcasting/receiving Wi-Fi signal (to estimate interpersonal distance) and sampling orientation. Subjects were given a case to carry the phone and in this manner we were aware of the position of the phone with respect to the body, in order to calculate relative body orientation. Once the social interactions ended (during which we were not present) participants were asked to fill out a short check-box questionnaire that requested a description of the interaction. In order to infer whether the conversation was formal or informal we used a questionnaire similar to the one used in the study of function of informal interactions in companies [21]. Overall, we collected 33 face-to-face communications, 21 informal (duration of 9 ± 5 min) and 12 formal (duration 21 ± 10 min), which included participation of 50 subjects (33 males/17 females, with an age of 32.7 ± 6.6 years) resulting in approx. 12 h of sensor data. Wi-Fi and orientation were sampled with 1 Hz and we estimated
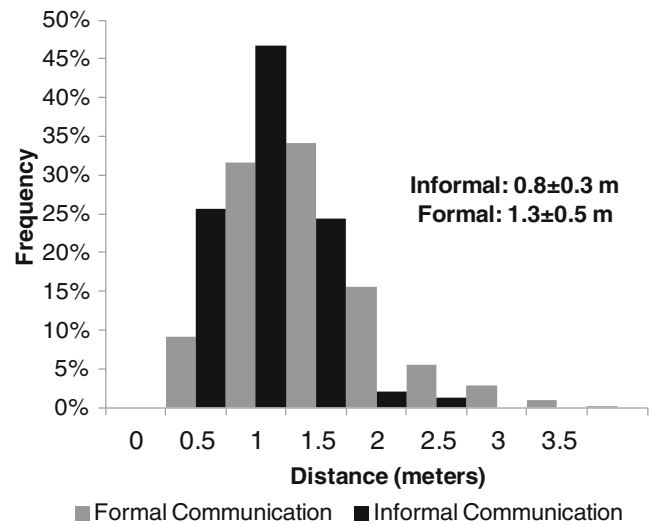

Fig. 7 Interpersonal distances in formal/informal social interactions

interpersonal distance and the relative body orientation for each time frame of 10 s.

5.3 Differences between formal and informal social context across monitored parameters

5.3.1 Interpersonal distances

Figure 7 shows the histogram of interpersonal distances, plotted for each time frame of 10 s recorded during formal and informal communications. According to the study of proxemics [15], interpersonal distances detected in informal communications mostly belong to the Personal Space having the mean value of 0.8 m. In formal communications, the results show distances that correspond to both Personal and Social Space, with the mean value on the border of these two zones, at 1.3 m. These absolute measures should be taken only illustratively considering the distance estimation precision (provided in Section 3). However, both distributions in Fig. 7 were acquired using the same system thus embedding the same median error in estimated distances. Therefore, whereas the absolute measures cannot be reliably claimed with a precision less than 50 cm due to the system's accuracy, the relative difference between interpersonal distances may provide more reliable estimate of the actual phenomenon. Furthermore, the results demonstrate that the distinction between formal and informal social interactions is reflected in interpersonal distances that are estimated using mobile phone sensing despite the median accuracy of 50 cm. Due to the considerable intersection of values related to interpersonal distances in formal and informal social interactions, relying solely on this temporal cue to distinguish the two types of social interactions would not suffice. Rather, it will be investigated if interpersonal distance can be combined with other parameters to distinguish different social contexts.
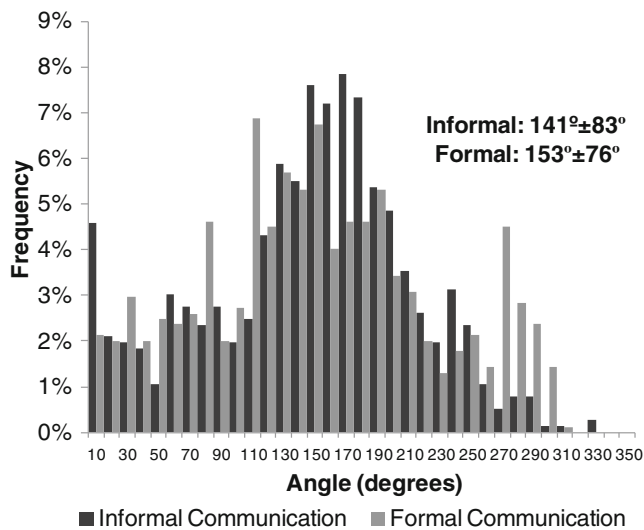
Fig. 8 Relative body orientationsin formal/informal socialinteractions

5.3.2 Relative body orientation

Relative body orientations in formal and informal social interactions were analyzed after discarding all 10-seconds frames with the standard deviation greater than 10° as previously described. The results are presented in Fig. 8.

In both formal and informal communications the mean value of relative body orientation was between 140° and 150° (180° corresponds to a direct face-to-face orientation) thus not demonstrating major differences between the two types of communication. It cannot be concluded if such results pertain to the phenomenon of formal/informal communications or it was due to the limited accuracy in estimating relative body orientation using the compass sensor embedded in phones and approximating the angle between the body and the phone orientation. However, when recognized with mobile phone, relative body orientation did not
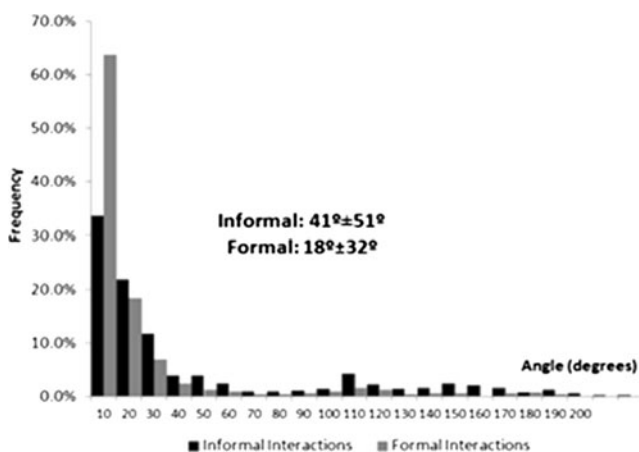

Fig. 9 Standard deviation of relative body orientations (calculated each 10 seconds)

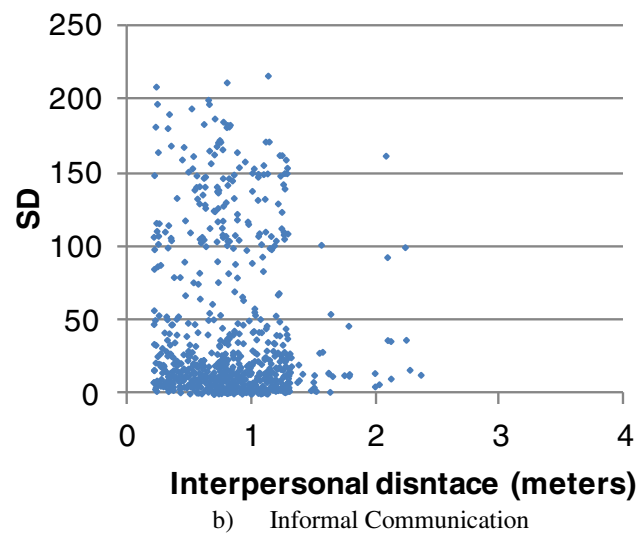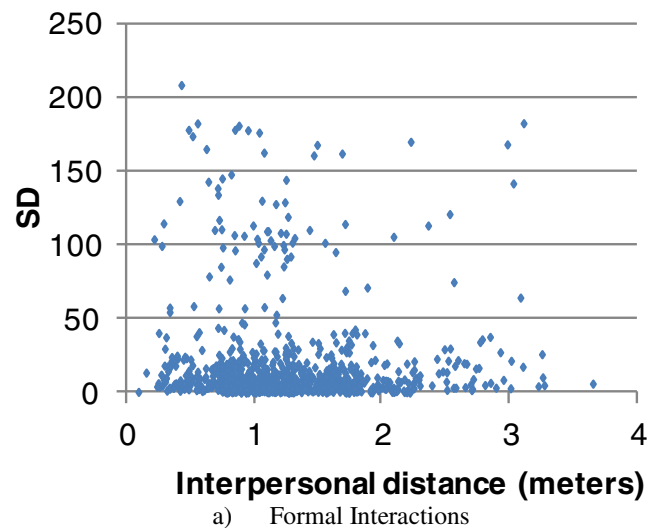a) Formal Interactions

b) Informal Communication

Fig. 10 Interpersonal distances and standard deviation of relative body orientation plotted pair-wise

mirror the difference between formal and informal conversational context.

5.3.3 Standard deviation of relative body orientation

Figure 9 shows the histograms of standard deviation for each 10-second frame during formal/informal social interactions. The results demonstrate that subjects were more flexible in holding their relative body orientation during

Table 3 Classification results (formal vs. informal)

Feature vector	Naïve Bayes (KDE)	SVM
(d, α)	64 %	66 %
(d, α, σ)	75 %	75 %
(d, σ)	77 %	81 %

informal communications than in the case of formal interactions. In formal interactions relative body orientation of subjects had a tendency to remain stable for longer periods (in contrast to informal social context), which may be due to maintaining eye contact for example, or some other external factor such as a video beam or a monitor that focused subjects' attention. Therefore, we attempted to classify formal versus informal communications on the basis of two parameters, namely interpersonal distance and the standard deviation of relative body orientation.

5.4 Classification of formal / informal interactions

The pairs of interpersonal distance and standard deviation of relative body orientation calculated for each 10-second time frame are plotted in Fig. 10 separately for formal and informal social context. The visualization of the data shows the differences between formal and informal interactions, which further prompted us to investigate the classification between these two types of social interactions. Please note that interpersonal distances were estimated applying GP regression for a more precise illustration (unlike in Section 4) of the differences between the two types of social interactions.

Table 3 presents the results for Naïve Bayes classification with Kernel Density Estimation and SVM. The performance was evaluated using 10-fold cross validation. These results demonstrate that interpersonal distance and standard deviation of relative body orientation are well suited features to discover the type of face-to-face communication. Furthermore, the accuracy of detecting these two parameters achieved with mobile phone sensing sufficed for this purpose, which may be the substantial basis for a number of context-aware mobile computing applications. Computing both parameters does not require the phone to be at a known place on the body thus affording an unobtrusive monitoring of subjects that habitually carry mobile phone.

6 Conclusions

Understanding social interactions is important for a number of disciplines, including social psychology, epidemiology, medicine, economics and anthropology. The solutions for continuous (mobile) monitoring of social interactions are typically based on the use of dedicated devices which introduces a set of issues including subjects' stigmatization and consequently their behavioral change. This occurs because wearing unfamiliar and visible sensing hardware increases the awareness of being monitored. One way to address this issue is to utilize the sensing capabilities available in one of the most widely adopted devices – mobile phone. However, current research on mobile phone sensing to monitor social interactions remains limited: it either includes audio analysis

which often raises privacy concerns and ethical issues or it focuses on quantifying dynamics of social activity over time while being limited in analyzing social interactions which occur on spatial scales of meters and time scales of minutes.

This paper presented the design and evaluation of the system able to infer social interactions which take place on small spatio-temporal scale relying on non-auditory sensing modalities found in typical mobile phones. We have described the challenges faced when using only mobile phone sensing as a substrate for analysis of social interactions and inference of the type of social interactions. Our results show that we can accurately detect interpersonal distance with median error of 0.5 m, even when using different mobile phone models, which combined with relative body orientation provides a reliable inference of social interactions. We have also found that stability of relative body orientation is a suitable parameter to detect occurrence of social interactions since it does not constrain subjects to wear the phone on a known place on body. Furthermore, the combination of interpersonal distance and standard deviation of relative body orientation showed a high predictive power in classifying the type of social interaction - formal or informal. We envision that our system for social interaction analysis can provide an instrument for gathering rich and large-scale social interaction data thus supporting the research in social networks and the investigation of formal/informal structures.

References

1. Eagle NN (2005) "Machine perception and learning of complex social systems," Massachusetts Institute of Technology
2. Eagle N, (Sandy) Pentland A (2005) Reality mining: sensing complex social systems. Pers Ubiquit Comput 10(4):255–268
3. Do TMT, Gatica-Perez D "Contextual grouping: discovering real-life interaction types from longitudinal Bluetooth data," *infoscience.epfl.ch*
4. Wyatt D, Choudhury T, Bilmes J (2011) "Inferring colocation and conversation networks from privacy-sensitive audio with implications for computational social science." ACM Transactions on Intelligent Systems and Technology (TIST) 2(1)
5. Cristani M, Pesarin A, Vinciarelli A (2011) "Look at who's talking: voice activity detection by automated gesture analysis," in *Workshop on Interactive Human Behavior Analysis in Open or Public Spaces*
6. Eagle N, (Sandy) Pentland A (2005) Reality mining: sensing complex social systems. Pers Ubiquit Comput 10(4):255–268
7. Hidalgo CA, Rodriguez-Sickert C (2007) The dynamics of a mobile phone network. Physica A: Stat Mech Appl 387(12):3017–3024
8. Raento M, Oulasvirta A, Petit R, Toivonen H (2005) ContextPhone: a prototyping platform for context-aware mobile applications. IEEE Pervasive Comput 4(2):51–59
9. Mardenfeld S, Boston D, Pan SJ, Jones Q, Iamntichi A, Borcea C (2010) "GDC: Group Discovery using Co-location traces," *2010*

IEEE Second International Conference on Social Computing, pp. 641–648

10. Do TMT, Gatica-Perez D (2011) "GroupUs: Smartphone proximity data and human interaction type mining," in *5th annual International Symposium on Wearable Computers*, no. 2

11. Banerjee N, Agarwal S, Bahl P, Chandra R, Wolman A, Corner M (2010) "Virtual compass: relative positioning to sense mobile social interactions," *Pervasive Computing*, pp. 1–21

12. Lane ND et al (2010) "BeWell: A smartphone application to monitor, model and promote wellbeing," in *5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth2011)*

13. Bhagwat P, Raman B, Sanghi D (2004) Turning 802.11 inside-out. ACM SIGCOMM Comput Commun Rev 34(1):33

14. Matic A, Papliatseyeu A, Osmani V, Mayora-Ibarra O (2010) "Tuning to your position: FM radio based indoor localization with spontaneous recalibration," *2010 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 153–161, Mar

15. Groh G, Lehmann A, Reimers J, Frieß MR, Schwarz L (2010) "Detecting social situations from interaction geometry," in *IEEE International Conference on Social Computing/IEEE International Conference on Privacy, Security, Risk and Trust*

16. Shi Y, Shi Y, Liu J (2011) "A rotation based method for detecting on-body positions of mobile devices," in *Proceedings of the 13th International Conference on Ubiquitous Computing*

17. Hall E (1966) The hidden dimnesion. Double Day Anchor Books, New York

18. Gatica-Perez D (2009) Automatic nonverbal analysis of social interaction in small groups: a review. Image Vis Comput 27 (12):1775–1787

19. Groh G, Lehmann A, Souza MD (2011) "Mobile Detection of Social Situations with Turn Taking Patterns," in *WAC2011*

20. Allen TJ (2000) "Architecture and communication among product development engineers," *Proceedings of the 2000 IEEE Engineering Management Society*, vol. 49, no. 2. UNIVERSITY OF CALIFORNIA, p. 153

21. (1990) "Informal communication in organizations: Form, function, and technology," in *Human reactions to technology: Claremont symposium on applied social psychology*

22. Whittaker S, Frohlich D, Daly-Jones O (1994) "Informal workplace communication: What is it like and how might we support it?," in *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, pp. 131–137

23. Aalbers R, Koppius O, Dolfsma W (2006) "On and off the beaten path: Transferring knowledge through formal and informal networks," *CIRCLE Electronic Working Paper Series*

24. Fischbach K, Gloor PA, Schoder D (2008) Analysis of informal communication networks – a case study. Bus Inf Syst Eng 1 (2):140–149