

# Measuring Interaction Proxemics with Wearable Light Tags

ALESSANDRO MONTANARI\*, Computer Laboratory, University of Cambridge

ZHAO TIAN\*, Department of Computer Science, Dartmouth College

ELENA FRANCU, School of Business and Economics, Maastricht University

BENJAMIN LUCAS, Nottingham University Business School, University of Nottingham

BRIAN JONES, Computer Laboratory, University of Cambridge

XIA ZHOU, Department of Computer Science, Dartmouth College

CECILIA MASCOLO, Computer Laboratory, University of Cambridge

The proxemics of social interactions (e.g., body distance, relative orientation) influences many aspects of our everyday life: from patients' reactions to interaction with physicians, successes in job interviews, to effective teamwork. Traditionally, interaction proxemics has been studied via questionnaires and participant observations, imposing high burden on users, low scalability and precision, and often biases.

In this paper we present *Protractor*, a novel wearable technology for measuring interaction proxemics as part of non-verbal behavior cues with fine granularity. Protractor employs near-infrared light to monitor *both the distance and relative body orientation* of interacting users. We leverage the characteristics of near-infrared light (i.e., line-of-sight propagation) to accurately and reliably identify interactions; a pair of collocated photodiodes aid the inference of relative interaction angle and distance. We achieve robustness against temporary blockage of the light channel (e.g., by the user's hand or clothes) by designing sensor fusion algorithms that exploit inertial sensors to obviate the absence of light tracking results.

We fabricated Protractor tags and conducted real-world experiments. Results show its accuracy in tracking body distances and relative angles. The framework achieves less than 6° error 95% of the time for measuring relative body orientation and 2.3-cm – 4.9-cm mean error in estimating interaction distance. We deployed Protractor tags to track user's non-verbal behaviors when conducting collaborative group tasks. Results with 64 participants show that distance and angle data from Protractor tags can help assess individual's task role with 84.9% accuracy, and identify task timeline with 93.2% accuracy.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**; • **Computer systems organization** → **Embedded systems**;

Additional Key Words and Phrases: Face-to-face interactions, non-verbal behaviors, light sensing.

\*Both authors contributed equally to this work.

Authors' addresses: Alessandro Montanari, Computer Laboratory, University of Cambridge, [alessandro.montanari@cl.cam.ac.uk](mailto:alessandro.montanari@cl.cam.ac.uk); Zhao Tian, Department of Computer Science, Dartmouth College, [tianzhao@cs.dartmouth.edu](mailto:tianzhao@cs.dartmouth.edu); Elena Francu, School of Business and Economics, Maastricht University, [e.francu@maastrichtuniversity.nl](mailto:e.francu@maastrichtuniversity.nl); Benjamin Lucas, Nottingham University Business School, University of Nottingham, [benjamin.lucas@nottingham.ac.uk](mailto:benjamin.lucas@nottingham.ac.uk); Brian Jones, Computer Laboratory, University of Cambridge, [bdj23@cl.cam.ac.uk](mailto:bdj23@cl.cam.ac.uk); Xia Zhou, Department of Computer Science, Dartmouth College, [xia@cs.dartmouth.edu](mailto:xia@cs.dartmouth.edu); Cecilia Mascolo, Computer Laboratory, University of Cambridge, [cecilia.mascolo@cl.cam.ac.uk](mailto:cecilia.mascolo@cl.cam.ac.uk).

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).

© 2018 Association for Computing Machinery.

2474-9567/2018/3-ART25 \$15.00

<https://doi.org/0000001.0000001>

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 2, No. 1, Article 25. Publication date: March 2018.



Fig. 1. Example of using Protractor to track team interaction when conducting a creative task (the Marshmallow challenge).

#### ACM Reference Format:

Alessandro Montanari, Zhao Tian, Elena Francu, Benjamin Lucas, Brian Jones, Xia Zhou, and Cecilia Mascolo. 2018. Measuring Interaction Proxemics with Wearable Light Tags. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 25 (March 2018), 30 pages. <https://doi.org/00000001.00000001>

## 1 INTRODUCTION

Face-to-face social interactions play a significant role in different contexts. In the workplace, serendipitous interactions between employees have been demonstrated to be key in team coordination, cohesiveness, and productivity [24, 60, 68]. Architects have studied how to increase unplanned interactions by changing the layout and design of physical spaces [22, 38, 69, 106, 113].

Traditional sociology has placed high importance in observing the non-verbal aspects of social interactions such as interaction proxemics (e.g., interaction distance and relative body orientation). Non-verbal cues on interaction proxemics reveal user attitudes and emotions [14, 99]. They also shed light on the role of culture [42, 48, 105] in influencing people's interaction behaviors. Observing these cues can facilitate many important applications. We list four specific examples:

- *team collaboration*: interaction details such as body distances and relative angles are important cues to study team collaboration (e.g., task timeline, individual roles) on creative tasks and assess team's creativity [95];
- *job interviews*: non-verbal skills such as eye contact, energy level, and affect (expressed with hand gestures and body movement) can be the subject of training to improve the interview outcome [32, 85];
- *doctor-patient interactions*: patient satisfaction is affected by the physician's expressiveness that includes non-verbal behaviors like more forward leaning, nodding, gestures and gazing [83, 110];
- *marketing and sales*: customer's engagement with the sales representative depends on his engagement abilities, which are therefore also important in sales training [79].

For all these examples, an accurate monitoring of body distance and relative orientation is crucial. The interaction distance between people has been estimated to be in approximately 7-cm intervals with a temporal granularity of 7 seconds in social interactions [98]. Angles of interactions are significant to study communicator's attitude towards his interlocutor and should be estimated to the nearest  $10^\circ$  based on prior study [86].

To monitor interaction proxemics continuously, conventional approaches in behavioral sciences have relied on questionnaires, participant observations, or the use of non-human objects (e.g., life-sized photographs, miniature dolls or silhouettes) [64]. Based on self-reporting, these approaches not only impose high burden on users, imply various biases, but also fail to provide behavioral information during a contact. Technology has progressed substantially in capturing fine-grained face-to-face interactions [28, 29, 52], however existing work still falls short: some either infer *only* user proximity [28, 29, 88] or body distances [52, 119], or analyze speech-related non-verbal signals with no information on interaction distance and relative orientation [89]. Others focus on very specific

contexts reproduced in the lab (e.g. job interviews, public speeches) and require cameras that bring privacy concerns and entail heavy environmental instrumentation, limiting the flexibility of the system [16, 17, 33].

The goal of this work, thereby, is to seek a more scalable and accurate approach to continuously measuring interaction proxemics as part of non-verbal behaviors during social interactions. To eliminate the need of infrastructure support, we consider a lightweight wearable tag resembling an access badge worn with a lanyard or clip (Figure 1). We leverage such tags to track *both the actual interaction distance and relative body orientation of users involved in a social interaction*. Specifically, each tag emits wireless beacons encoded with its tag ID and listens to beacons from other nearby tags. Based on the received beacons, the tag then identifies other tags/users within its sensing range, and estimates the relative angle and distance to each of these tags/users. These angle and distance numbers are used to identify participants and recorded as their interaction proxemics during an interaction. At first sight, the problem appears to be a standard problem of relative device positioning. However, the context of tracking interaction proxemics presents three new challenges. We next overview each challenge and our solution.

First, accurately identifying the participants in an interaction is challenging. Two users in close distance may not be in an interaction, as they may stand with other people in between or are not facing each other (see examples in Figure 2). Thus, it is key to recognize both the line-of-sight proximity and user's *relative* body orientation. To this end, prior methods using radio frequency (RF) signals (e.g., Bluetooth, Wi-Fi) [22, 75, 88] or microphones [76, 107, 119] are all prone to false positives, since RF signals and sound penetrate human bodies. Also, relative body orientation cannot be simply obtained by compass sensors, which measure only the absolute orientation of the user/tag itself, rather than how it relates to other tags, as shown in § 4.1.

To reduce such false positives and enable accurate tracking without the need for expensive and cumbersome infrastructure, we choose near-infrared (NIR) light as the wireless medium for tags to transmit beacons. With wavelengths in nanometers, NIR light is imperceptible, directional, and cannot penetrate opaque macroscopic objects (e.g., human body). Thus, it is the ideal medium for measuring line-of-sight proximity in our context. Furthermore, to infer relative angles and distances to other tags, we leverage two collocated infrared photodiodes each pre-configured with a different orientation (§ 4.1). By analyzing the difference of light intensity sensed by the photodiodes, we can compute the incident angle and distance to each sensed tag.

The second challenge lies in enabling reliable tracking using infrared light beacons. Light beacons can be accidentally blocked by user's hands, clothes, another user passing by, or other objects (e.g., book, paper) introduced during the interaction; the motion of user body can cause tags suddenly moving beyond each other's sensing range. In all these cases, the tracking results using NIR light can either become unavailable or have low fidelity. To deal with these artifacts and realize reliable tracking, we augment light-based tracking with inertial sensors (i.e., accelerometer, gyroscope). Although inertial sensors measure only the motion status (e.g., velocity, orientation) of the tag itself, we design a data fusion algorithm (§ 4.2) that leverages inertial sensor data to extrapolate missing relative angles and distances upon losses of light beacons.

The third challenge is to ensure that tags operate with low power to avoid frequent charging and to ease tag distribution for various studies. Certain components (e.g., NIR LED) consume relatively higher power than others, and directly detecting short (e.g., 1.8  $\mu$ s) NIR light pulses imposes energy burden of high analog-to-digital (ADC) sampling (e.g., 500 KHz). To improve system energy efficiency, we design strategies (§ 4.3) for a tag to adapt its operation mode to the current context (e.g., presence of nearby tags, user's motion status). It selectively switches off more energy-demanding modules to save energy without much sacrificing sensing temporal granularity. We also judiciously design the NIR sensing circuit to eliminate the need of high ADC sampling (§ 4.1).

We have implemented our designs and fabricated Protractor tags using off-the-shelf, low-cost hardware. Each tag is measured  $74 \times 54 \times 15$  mm in size and 40 g in weight. We have evaluated the efficacy of our tags in ranging and angle detection using controlled experiments. To examine its practical implications, we have further deployed our tags to track user interaction proxemics when users collaborate on the Marshmallow challenge as the creative

task. We conduct the deployment with 16 groups of 4 users each and seek to examine the feasibility of using fine-grained angle and distance data to understand complex processes comprising interrelated sets of human actions such as team creativity in an organizational environment. Our main findings are as follows:

- Protractor achieves  $2.2^\circ$  mean angular error and  $5.2^\circ$  95th percentile in estimating interaction angles and 2.3-cm – 4.9-cm mean error in ranging;
- Protractor is robust in diverse settings (e.g., tag height offsets, indoor lighting variations, reflections from nearby objects) and effectively mitigates occasional missing or unreliable NIR tracking results with data fusion;
- Protractor is capable of running continuously for 5 days with a single charge by switching into low power modes based on contextual information;
- Protractor is capable of supporting organizational science studies by providing objective data that can be used to predict subjective assessments (e.g., task role, task timeline) with an accuracy from 84.9% to 93.19%, which outperforms similar prior studies [117].

We see the potential of Protractor not only in the support of social research but also for practical applications (e.g., providing real-time behavioral feedback during interactions, novel human-computer interaction interfaces). In comparison to approaches using cameras, Protractor serves as a more lightweight and scalable alternative. Its unobtrusive nature and the wearable form factor could ease privacy concerns and potentially reduce biases for accurate behavioral monitoring.

## 2 RELATED WORK

**Profiling User Interactions.** We overview prior works on profiling user interactions based on their technology.

1) *RF technologies*: Bluetooth Classic has been widely used for detecting user proximity [9, 10, 35, 53]. However, its high power leads to low sensing granularity (e.g., sampled every few minutes to avoid draining the battery quickly [9, 35, 53]). Recent works [20, 56, 57, 88] use Bluetooth Low Energy (BLE) to collect data on user proximity and mobility. Customized devices have also been used to monitor interactions. Forsys et al. used wireless sensors based on the IEEE 802.15.4 standard to collect human contacts for epidemiology research [40]. Cattuto et al. used RFID to sense interactions by exchanging low-power radio packets, which are shielded by the human body and cannot be received if people are not facing each other [26].

2) *Microphones*: Another source of data is the microphone on unmodified phones. Lee et al. monitored conversation by matching the volume signature with a topography database built during a learning phase [76]. Zhang et al. exploited the Doppler effect to detect the trajectories of approaching people and adopted voice profiling to confirm the occurrence of a conversation [119]. By monitoring the actual conversation, however, these approaches are sensitive to false positives if other nearby users are in a different conversation. Also, the use of microphones might raise ethical and privacy issues, preventing the wide adoption of the system. Tan et al. proposed the use of audio silence pattern to detect collocation in a privacy-preserving manner [107]. The resulting power consumption, however, still prevents the system from running continuously.

3) *Sensor Combinations*: Other approaches have adopted a combination of technologies. Choudhury and Pentland combined infrared transceivers and microphone for measuring face-to-face interactions [28]. The system has been deployed in many contexts [29, 74, 89]. Ultrasound and RF signals have also been combined to capture interactions, using the time-difference-of-arrival technique. As examples, the iBadge [102] applies this principle to capture interactions between kids, teachers and objects in a kindergarten classroom; Opo [52] further boosts the ranging accuracy (5-cm accuracy) with a temporal fidelity of 2 seconds.

4) *Cameras*: Video cameras or more sophisticated versions, such as Microsoft Kinect, have been used for the automatic recognition of non-verbal behaviors and analysis of group interactions. As examples, in [117], Zancanaro et al. detect participant's functional role in group interactions by setting up nine cameras and several microphones in a room and analyzing video with machine learning algorithms. Jayagopi et al. study conversational

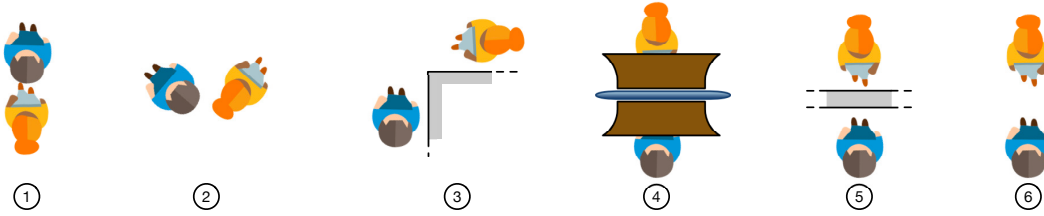


Fig. 2. The first five are example scenarios with users in proximity but not in an interaction, where two users are 1): in line with one's back facing the other; 2) in opposite directions; 3) on either side of a corner, without line of sight; 4) occupying opposite desks with a partition in the middle; 5) on either side of a wall when working in neighboring offices. Scenario 6 is an actual interaction with users facing each other at a distance of 1 m. We use the last scenario as a reference for comparison.

group dynamics (e.g., conversational topics, leadership styles) [58], and group dominance [59], using nonverbal cues extracted from an existing dataset with 100-hr meeting recordings [25]. In [54], group cohesion is studied using hours of audio-visual group meeting data. [100] uses prosodic and turn-taking behaviors to identify participant's speaking role. [55] estimates group formations in crowded environments using a graph clustering algorithm. The analysis is based on video footage of over 50 people presenting scientific work in a poster session. Similarly, [31] estimates F-formations of human interactions using video footage of coffee breaks in a social event. [101] infers emergent leaders using nonverbal cues extracted from audio and video channels.

Although camera-based approaches capture rich data of social contacts, they face several limitations. First, they require considerable effort in instrumenting buildings [25, 54, 117] to enable data collection. An infrastructure-free solution is superior from this point of view since it allows to collect data even in areas that cannot be instrumented, like public spaces or during large events. As a result, so far cameras have been used in very specific and static contexts reproduced in the lab (e.g. job interviews or public speeches) [16, 17, 33]. Second, in general, the analysis of non-verbal behaviors through video recording requires a considerable amount of storage space and processing power, limiting the usability of the system. Finally, cameras raise privacy concerns from the user and from people being recorded without their consent. This is particularly true for wearable cameras that are always with the user, even during private or intimate moments [51].

*Overall, prior works either commonly focus only on user proximity or interaction distance, or rely on invasive cameras. Our work differs in that we also continuously and unobtrusively capture relative body orientation as an important part of non-verbal interaction cues.*

**Relative Device Positioning.** Another related line of work is on sensing the relative position and orientation among devices. Particularly for short-range positioning, existing work has explored the use of ultrasound (18 – 20 kHz, or 40 kHz) and infrared. Ultrasound methods measure time of flight of acoustic signals to position devices by multilateration [49, 81, 87, 96] and estimate device orientation by measured phase offset [97] or positions of multiple devices [50]. These systems, however, require either additional RF radio [49, 87, 96], or the aid of multiple nodes (pre-deployed anchor nodes with known locations [81, 97] or multiple peer nodes [50]). Infrared-based systems have commonly been used in robotics, which measure the reflected infrared light to detect surrounding obstacles and distances [19], or use static stereo-cameras to track moving objects that carry active tags emitting infrared signals [11, 13]. Our scenario differs in that we enable mobile tags to track each other without any infrastructure support. In [41], Frantal et al. measured infrared incident angle using 12 photodiodes each facing a different direction. Its limited angular resolution and the resulting form factor make it not suitable in our context.

### 3 A CASE FOR LIGHT-BASED TAGS

Our design of the wearable tag starts with seeking the suitable wireless medium to transmit beacons, which are exploited to infer incident angles and distances to other tags/users in an interaction. The ideal medium should best



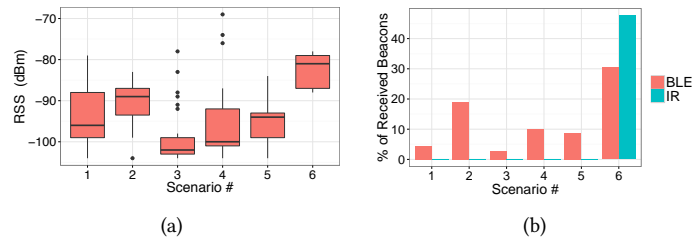


Fig. 3. (a): Received signal strength of BLE packets for the six scenarios described in Figure 2. (b): Percentage of received beacons by BLE and Infrared for the six scenarios described in Figure 2.

facilitate the measure of line-of-sight distance and incident angle, so that we can correctly identify participants in a contact. We now discuss three candidates: RF signals (e.g., Wi-Fi, Bluetooth), ultrasound, and light.

**Radio Frequency.** Prior studies have utilized RF signals on wearable devices or smartphones to monitor social interactions [9, 26, 88]. These systems examine the received signal strength (RSS) to infer if users carrying or wearing these devices are engaged in an interaction. However, RF signals are omni-directional, penetrate human bodies and objects, and are susceptible to multi-path effects. All these characteristics can make the identification of close encounters and relative orientation difficult and prone to false positives.

To verify this problem, we take Bluetooth Low Energy (BLE) as an example, and devise simple experiments that recreate realistic scenarios that involve two people in an indoor environment (Figure 2). These scenarios represent different combinations of people and objects between the transmitting devices. In each scenario, users wear on the chest a BLE device (Nordic nrf51822 SoC), transmitting advertisement beacons at 10-Hz rate with  $-20$  dBm TX power<sup>1</sup> and scanning for beacons every 100 ms with each scan lasting 20 ms. We collect RSS traces from each BLE device for 60 seconds in each scenario. The experiments are conducted outside office hours to avoid the presence of moving people in the vicinity. However, the environment presents various surfaces that could reflect radio signals (e.g., walls, the floor, the ceiling) and there are also 5 Wi-Fi access points active (2.4 GHz).

Figure 3(a) shows box plots of RSS values in dBm in all scenarios, where a higher value indicates a higher received signal strength. Figure 3(b) shows the percentage of received BLE beacons. We make two main observations. First, as expected, in all scenarios (1 – 5) where users are not in a social contact, BLE packets can still be received even when two devices are not in line of sight. The reception ratio of BLE packets is below 30% because the device does not scan continuously but performs a 20ms scan every 100ms and thus misses advertisement beacons. Second, although users are stationary, RSS values vary significantly in a single setting and across different settings. This is because BLE uses three channels (separated by 2 MHz) to transmit advertisement beacons, resulting in fades at different spatial positions for different channels, even when transmitter and receiver are static [37]. Wi-Fi signals present similar characteristics. We conclude that RF is not the proper choice in our context.

**Ultrasound.** Next, we examine ultrasound for transmitting beacons. With wavelengths in millimeters, ultrasound has been shown to have line-of-sight propagation and be unable to penetrate objects. This has been exploited by earlier studies to sense interaction distances [52] or to position devices [87, 96, 97]. In our experiment, we modify the HC-SR04 [2] ultrasonic transducer (4.5-cm in diameter) with 40 kHz center frequency, commonly used by prior studies [52, 87, 96, 97]. It sends carrier bursts for 8 cycles periodically (1 transmission every 2.5 seconds in our experiment). Because of the difficulty of modulating ultrasound pulses<sup>2</sup>, these bursts are treated as pure pulses at the receiving end without any decoding, and we use an oscilloscope to inspect the signal and

<sup>1</sup>The same power level has been used on the same BLE chip by the Openbeacon.org project to detect proximity encounters [4].

<sup>2</sup>As shown in [96], ultrasound suffers from severe multi-path effects and is hard to modulate. Thus, it is commonly used in combination with RF for ranging.

its amplitude. We repeat the experiment in the same scenarios in Figure 2. Our results confirm that ultrasound cannot penetrate objects in scenario 1, 4, and 5, whereas in scenario 2 and 3, we occasionally observe weak pulses, possibly due to reflection. Such pulses can trigger incorrect detection of social contacts if the appearances of pulses are used for ranging [52].

**Light.** We now move on to examining light as the final candidate. Specifically we consider NIR light rather than visible light<sup>3</sup>, because NIR is imperceptible to human eyes and keeps the wearable tag sensing unobtrusive. We repeat the same scenarios in Figure 2, where users wear on their chest an NIR transceiver transmitting one NIR beacon per second. For each scenario we log the received and decoded beacons for 60 seconds and compute the percentage of received beacons (Figure 3(b)). We observe that the NIR transceiver does not receive any beacons in scenarios (1 – 5) where the devices are not in line of sight. The beacon losses in scenario 6 are due to errors during the decoding at the receiver end, which prevent the identification of the correct beacon<sup>4</sup>. NIR light propagates as a directional beam in a cone shape, thus it serves as a good medium to detect and monitor relative angle and distance of interacting people. Additionally, typical NIR emitters and receivers have a very small form factor (e.g., 5×5×7-mm), which is desirable for building a wearable device to be worn all day.

Based on all above experiments, we decide to choose NIR light as the wireless medium for sensing non-verbal cues in social contacts.

## 4 PROTRACTOR DESIGN

The core of Protractor is to measure relative angles and distances of interacting users in an accurate and a reliable manner. Protractor achieves accuracy by exploiting the propagation characteristics of NIR light for precise angle detection and ranging. It ensures the tracking reliability by fusing inertial sensors and NIR sensors to compensate for the occasional loss (e.g., light being blocked) of light tracking results. Above all, as a wearable tag, Protractor is designed to operate with low power. Next, we elaborate on each design component.

### 4.1 Angle Detection and Ranging

A face-to-face interaction can occur in various forms. Two important non-verbal interaction cues are the distance between any two involved users and their relative body orientation [47]. We define the latter as the *interaction angle*, which is the angle between the body normal and the line connecting the two users (Figure 4(a)).

At the first look, interaction angles seemingly can be obtained using the magnetometer/compass sensor, which measures user's absolute orientation. Then by exchanging the information with nearby users, one can estimate relative angles to others. However, knowing absolute orientation alone is inadequate to infer interaction angles. Figure 4(b) shows a simple example, where even if both user A and B's absolute orientations are known, their interaction angle still cannot be determined. Because B can be at location B' with the same orientation, which yet results into a different interaction angle  $\theta_{B'}$ . Adding the knowledge of A and B's distance does not help either (B' and B are at an equal distance to A). Such angle ambiguity can be resolved with A and B's absolute locations, obtained by

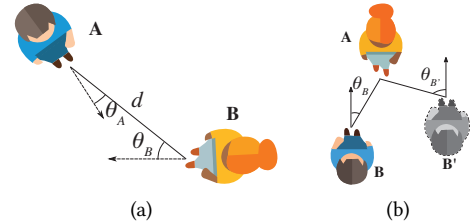


Fig. 4. (a): Interaction distance  $d$ , and relative body orientation, i.e., interaction angles ( $\theta_A, \theta_B$ ), in a social contact. (b): The same absolute orientation of B can lead to different orientation relative to A.

<sup>3</sup>A recent study [108] uses ultra-short visible light pulses to enable imperceptible communication. It can also be a candidate.

<sup>4</sup>Note that as a simple proof of concept, this experiment is to compare different media, rather than an extensive analysis of general success rates in decoding NIR beacons. Our tags achieve much higher success rates in decoding by regulating beacon transmissions and adding random delays. We will discuss our tag design in § 5 and detailed experiments on its decoding robustness with multiple tags in § 6.3.

existing user-centric indoor localization methods [72, 80, 118]. But still, user's 2D location coordinates indicate little on the actual occurrence of face-to-face contacts. As shown in earlier examples (Figure 2), nearby users can be separated by other indoor objects (e.g., a wall, desk partitioner) and thus not in a social contact.

Protractor overcomes the above problem by directly measuring the line-of-sight channel between two chest-worn Protractor tags using NIR light. Its key design elements are the NIR light beacons emitted by each tag, the detection of incident angle, and the estimation of line-of-sight distance.

**NIR Light Beacons.** A Protractor tag periodically (1 beacon every 5 s in our implementation) emits NIR light beacons, each of which encodes the user ID. We choose the NIR wavelength of 940 – 950 nm for the beacon transmission. It is commonly used in consumer wireless infrared communication such as TV remote control. To encode data, an NIR emitter (i.e., LED) flashes at a carrier frequency (38 kHz) in bursts. Among various IR modulation/coding schemes, Sony IR coding [111] is chosen in our implementation for its popularity. As illustrated in Figure 5, bit 1 is encoded as 1200  $\mu$ s carrier frequency burst followed by an off duration (600  $\mu$ s), while bit 0 is 600  $\mu$ s carrier frequency burst followed by an off duration (600  $\mu$ s). To reduce the power consumption, we decrease LED's duty cycle of the carrier to 7%. To decode light beacons, we use an infrared receiver module [8], which outputs logic LOW continuously for carrier frequency (mark) and logic HIGH for off duration (space). The micro-controller polls the receiver's output every 50  $\mu$ s to detect the duration of each mark and decode bits.

In addition to conveying the user/tag ID, the received signal strength (RSS) of a light beacon is utilized later for deriving interaction angles and distances. Here a light beacon's RSS equals the peak amplitude of the light pulse minus the ambient light baseline (Figure 5). Measuring the RSS is challenging at a low-power wearable device, because the common IR carrier frequency is 38 KHz, meaning that the light pulse can be as short as 1.8  $\mu$ s (7% duty cycle). Detecting such short light pulses requires a sampling rate higher than 500 KHz, imposing a high energy overhead to the tag. To address this problem, Protractor leverages an envelope detector (Figure 5 and 9(b)) that holds the signal at its peak until the end of a beacon. It allows the micro-controller to sample the peak amplitude with much lower rates (1 kHz in our implementation).

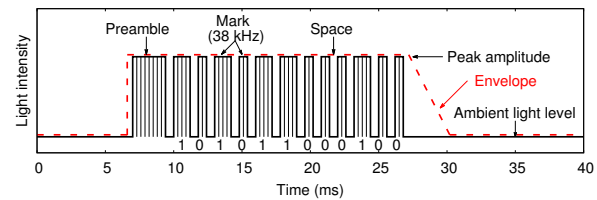


Fig. 5. Time series of NIR light pulses of an example beacon.

**Deriving Interaction Angle.** Protractor reuses light beacons to derive the interaction angle to the user/tag that each received beacon corresponds to. In the RF literature, estimating the signal's angle of arrival commonly relies on multiple antennas placed with known intervals to measure phase offset [43, 66, 104, 116] or mechanically rotating antennas [70, 71]. These methods are not applicable in our context, because of the tag's small form factor. Also, since LED is incoherent light source, there is no phase information as in RF technologies.

Instead, Protractor leverages the fact that an NIR photodiode responds to incoming light with different sensitivity depending on the light's incident angle, which is referred to as photodiode's *angular response*. Thus, if two collocated NIR photodiodes face different directions, incoming light with a given incident angle can result into different signal strength perceived by each photodiode. If we can obtain the one-to-one mapping between the light incident angle and the resulting signal strength pattern at photodiodes, we can then derive incoming light's incident angle based on measured RSS values at photodiodes.

Before diving into the detail of the above method, we first describe the optical channel model characterizing the propagation of NIR light. For a pair of LED and photodiode with distance  $d$ , assume that LED's light ray with irradiance angle  $\phi$  hits the photodiode with incident angle  $\theta$  (Figure 6(a)), and  $I$  denotes the RSS at the



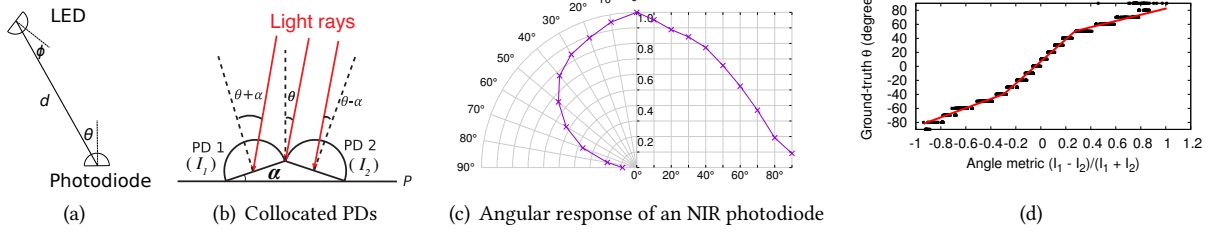


Fig. 6. Estimating the interaction/incident angle  $\theta$  using two collocated photodiodes (PD). (a) shows the optical channel between an LED and a photodiode, with irradiance angle  $\phi$  at the LED and incident angle  $\theta$  at the photodiode. (b) shows two collocated photodiodes facing different directions. Because of the photodiode's angular response (c), two PDs perceive different signal strength  $I_1, I_2$ . The incident angle  $\theta$  and the angle metric in Eq. (2) has a linear relation, which can be used to estimate  $\theta$  on the fly.

photodiode.  $I$  can then be calculated as [65, 80]:

$$I = A \frac{F(\phi)G(\theta)}{d^2}, \quad (1)$$

where  $A$  is a constant determined by the transmit power and receiver's gain,  $F(\phi)$  is LED's irradiation pattern at irradiance angle  $\phi$ , and  $G(\theta)$  is photodiode's angular response at incident angle  $\theta$ .

Now consider two collocated photodiodes that are rotated clockwise and counter-clockwise respectively, by a pre-defined angle  $\alpha$  with respect to the reference plane  $P$  (Figure 6(b)). Suppose  $\theta$  is the interaction angle, i.e., the angle between the incoming light and the normal of  $P$ . Then for the first and second photodiode, the light's incident angle is  $\theta + \alpha$ , and  $\theta - \alpha$  respectively, causing different RSS at each photodiode. Using the optical channel model (Eq. (1)), we can compute the RSS at each photodiode as:

$$I_1 = I_0 G(\theta + \alpha), \quad I_2 = I_0 G(\theta - \alpha)$$

where  $I_0 = AF(\phi)/d^2$ . We consider the same  $I_0$  for both photodiodes because  $\phi$  and  $d$  are the same for both photodiodes, given that the distance from the LED to the photodiodes (e.g., 30 cm to 2 m for normal social contacts) is much larger than the photodiode size (5 mm in diameter).  $\alpha$  is a known parameter, so the question now is how to derive  $\theta$  after measuring  $I_1$  and  $I_2$ . A straightforward method is to exhaustively measure photodiode's angular response at different incident angles and to seek the best-fit  $G(\cdot)$  function. Then  $\theta$  can be computed by solving the equation  $I_1/I_2 = G(\theta + \alpha)/G(\theta - \alpha)$ . This method, however, is ineffective. Because  $G(\cdot)$  can be complicated (e.g.,  $\cos^m(\theta)$ ) or even without analytical form, there is no closed-form solution. Numerical method such as the Newton's method are too computationally intensive.

To circumvent the need to solve the complicated equation, we seek a metric that is computed based on  $I_1, I_2$  and has a simple 1-1 mapping with  $\theta$ . To this end, we define an *angle metric*  $i$  as

$$i \equiv \frac{I_1 - I_2}{I_1 + I_2} = \frac{G(\theta + \alpha) - G(\theta - \alpha)}{G(\theta + \alpha) + G(\theta - \alpha)}. \quad (2)$$

Since the angular response of NIR photodiodes are typically symmetric (i.e.,  $G(\cdot)$  is an even function), the relation between  $i$  and  $\theta$  has the following properties: first,  $i$  is zero when  $\theta = 0$ , as  $G(\alpha) = G(-\alpha)$ ; second, the relation between  $\theta$  and  $i$  is approximately linear, even when  $G(\cdot)$  is non-linear, such as  $\cos^m(\theta)$ , based on our simulation, indicating that we can always apply linear regression to seek the relation between  $\theta$  and  $i$ .

To verify the relation between  $i$  and  $\theta$ , we conduct a benchmark experiment using two NIR photodiodes (OSRAM SFH 205 F [5]) with the measured angular response in Figure 6(c). We arrange the two photodiodes with  $\alpha = 22.5^\circ$  (Figure 6(b)) on a table and move the IR transmitter to emulate different interaction angles ( $-90^\circ$  to

90°) and different distances (50 cm to 200 cm) (Figure 10(a)). At each location, the transmitter sends beacons for 30 seconds. We measure  $I_1$  and  $I_2$  at two photodiodes and compute the metric  $i$  (Eq. (2)). We then plot all  $i$  values along with  $\theta$  in Figure 6(d). We observe that  $\theta$  is piecewise linearly<sup>5</sup> related to  $i$ . With the linear relation obtained offline through sample measurements, we can derive  $\theta$  on the fly after computing  $i$  based on measured  $I_1$  and  $I_2$ .

**Estimating Interaction Distance.** Protractor estimates the interaction distance by leveraging the optical channel model (Eq. (1)) and derived interaction angles. Specifically, for a pair of tags  $m$  and  $n$ , each tag first detects its interaction angle to the other tag, i.e.,  $\theta_m, \theta_n$ . Since the interaction/incident angle of a tag is also the irradiance angle of the other tag, we can compute the distance  $d_{mn}$  between  $m$  and  $n$  as  $d_{mn} = \sqrt{A \frac{F(\theta_n)G(\theta_m)}{I_m}}$ , where  $I_m$  is the RSS of light beacons from tag  $n$  measured at tag  $m$ .

Directly computing the above formula requires knowing the value of  $A$ . Instead, we define a *distance metric*  $l$  as  $l = F(\theta_n)G(\theta_m)/I$  and rewrite  $d_{mn}$  as

$$\ln(d_{mn}) = a \ln(l) + b. \quad (3)$$

We compute the logarithm in the above equation because the exponent of the distance  $d$  is not exactly 2, as shown in our measurements. We calibrate parameter  $a$  and  $b$  using benchmark experiments, where we collect  $l$  values<sup>6</sup> along with the ground-truth distance  $d_{mn}$ , and ground-truth interaction angles  $\theta_m, \theta_n$ . We then perform a linear regression to determine  $a$  and  $b$ . Figure 7 shows our benchmark experiment results and the linear model. With the trained linear model (Eq. (3)), we then compute interaction distances based on the derived interaction angles.

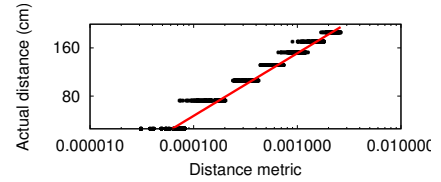


Fig. 7. Distance measurement model. Both axes are in log scale.

## 4.2 Sensor Data Fusion

While providing precision, NIR light tracking alone is not reliable for a number of reasons: light can be easily blocked by other objects (e.g., a waving hand, a book, a piece of paper) introduced in an interaction; the chest-worn tags can occasionally move beyond each other's sensing range, due to user's body movement during a contact. To enhance the tracking reliability, Protractor leverages inertial sensors (i.e., accelerometer, gyroscope) to compensate for the low fidelity of light tracking results in those occasions. We choose inertial sensors because they are small in size (2.5×3 mm) and consume low power (e.g., 2.8 mW). They can be easily fit in the wearable tag and operate continuously in the background with negligible energy overhead.

The challenges of using inertial sensors lie in sensory measurement noise. Such noise is particularly troublesome when measuring small displacement (e.g., centimeter-level distance change)<sup>7</sup>. In a social interaction, users tend to remain static at their 2D locations while changing body orientation by a greater extent. Thus, we consider fusing only the gyroscope data and the estimated interaction angles, while using accelerometer to sense large location displacement for determining the start/end of a new sensor fusion process.

To fuse the NIR angle detection results and gyroscope readings, we adopted the Kalman filter algorithm [36, 45, 61] for its simplicity and efficiency. Specifically, we model the interaction angle as a discrete-time hidden

<sup>5</sup>We run a linear regression at different intervals ( $[-90^\circ, -30^\circ]$ ,  $[-30^\circ, 30^\circ]$ , and  $[30^\circ, 90^\circ]$ ) to obtain the linear relation. For photodiodes with single-slope linear angular response, the relation will also be single-slope linear.

<sup>6</sup>We estimate  $F(\cdot)$  and  $G(\cdot)$  based on sampled measurements.

<sup>7</sup>Our experiments with inertial measurement unit Bosh BMI160 show non-zero sensor readings (e.g.,  $0.03 \text{ m/s}^2$  at x-axis) in the stationary mode even after removing the constant offset. It translates into 1.5-m location drift after only 10 seconds.

Markov model (HMM):

$$\begin{aligned}\theta_t &= \theta_{t-1} + \Delta\theta_t + w_t, & w_t &\sim \mathcal{N}(0, \sigma_{w,t}^2) \\ \tilde{\theta}_t &= \theta_t + v_t, & v_t &\sim \mathcal{N}(0, \sigma_{v,t}^2)\end{aligned}$$

where  $\theta_t$  is the hidden state (i.e., the actual interaction angle) at time  $t$ ,  $\tilde{\theta}_t$  is the observation (i.e., the estimated interaction angle using NIR measurements),  $\Delta\theta_t$  is the orientation change measured by the gyroscope sensor,  $v_t$  denotes the Gaussian observation noise (i.e., the angle detection errors using NIR light), and  $w_t$  is the Gaussian noise of gyroscope readings. Given that it is a linear Gaussian Bayesian model, Kalman filter has been proven to seek the optimal solution recursively [61].

Our data fusion based on Kalman filter recursively conducts two steps: prediction and updating. The prediction step produces the estimated mean and variance of the interaction angle at  $t$ , before the arrival of new NIR measurements at  $t$ . It predicts the interaction angle by:

$$\begin{aligned}\hat{\theta}_{t|t-1} &= \hat{\theta}_{t-1|t-1} + \Delta\theta_t \\ \sigma_{\theta,t|t-1}^2 &= \sigma_{\theta,t-1|t-1}^2 + \sigma_{w,t}^2.\end{aligned}$$

Upon the arrival of new NIR measurements and thus newly derived interaction angle  $\tilde{\theta}_t$ , the updating step then incorporates the new observation into the prior estimate and obtains improved posteriori estimates. It updates estimates as follows:

$$\begin{aligned}\hat{\theta}_{t|t} &= \hat{\theta}_{t|t-1} + k_t(\tilde{\theta}_t - \hat{\theta}_{t|t-1}) \\ \sigma_{\theta,t|t}^2 &= \sigma_{\theta,t|t-1}^2 - k_t\sigma_{\theta,t|t-1}^2\end{aligned}$$

where  $k_t = \sigma_{\theta,t|t-1}^2 / (\sigma_{\theta,t|t-1}^2 + \sigma_{v,t}^2)$ .

The update step can mitigate large accidental errors in NIR measurements, such as incorrect pulse amplitude detection due to ADC malfunction. The data fusion addresses the problem of occasional losses of NIR measurements, as its prediction step produces estimated interaction angle without new NIR measurements.

We start the fusion with an NIR measurement:  $\hat{\theta}_{0|0} = \tilde{\theta}_0$ ,  $\sigma_{\theta,0|0}^2 = \sigma_{v,0}^2$ . We model the variance of noise  $v_t$  and  $w_t$  based on our experimental observations. Specifically, our experiments show that NIR angle detection errors tend to have a small variance when both photodiodes have large pulse amplitude readings. Thus, we model the variance  $\sigma_{v,t}^2$  of the observation noise  $v_t$  as  $\sigma_{v,t}^2 \propto 1/(I_{t,1} + I_{t,2})$ . We model the gyroscope noise variance as  $\sigma_{w,t}^2 \propto \Delta t$ , because of the drifting problem of gyroscope sensor. The orientation change is an integration of the gyroscope readings and thus its error accumulates over time. We terminate the data fusion process when large location displacement is discovered from accelerometer readings [73], e.g., users walk away from their previous locations.

### 4.3 Adaptive Sampling

Given our goal of continuously tracking social contacts, Protractor's battery life is a critical aspect of our design. To ensure efficient use of available power while keeping the tag operational, we apply context-aware duty cycling. Succinctly, when no interactions are detected for a period or, when the tag is not being used, the more energy-demanding modules are switched off or reduced in their capability to save energy. The more energy-demanding modules are the angle detection module (mainly the transimpedance amplifier, § 6.4) and the NIR LED, while the inertial measurement unit (IMU) and NIR receiver consume low energy. We thus use the IMU and NIR receiver to infer the current context (i.e., presence of other devices nearby and user's motion status) and adapt Protractor's operation accordingly.

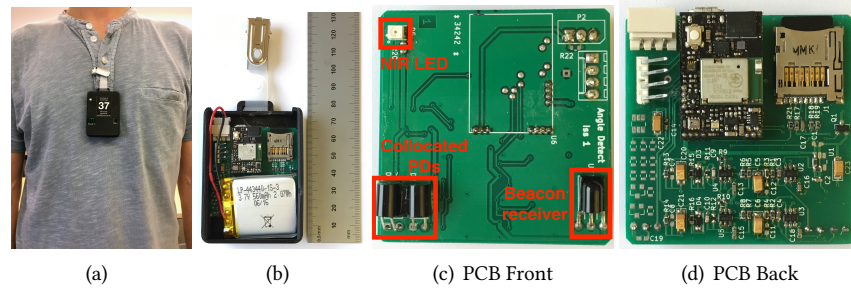


Fig. 8. Protractor prototype. (a) and (b) show the assembled tag and its internal components. (c) and (d) is the PCB we design and fabricate to host the NIR transceiver and angle detection module.

We define three states that a Protractor tag can be in at any given time: 1) *High Power*: all modules are powered on, NIR beacons are transmitted every 5 s and the two photodiodes in the angle detection module are sampled at 1kHz; 2) *Low Power*: the angle detection module is powered off and NIR beacons are transmitted every 20 s; 3) *System Off*: the angle detection module and the NIR transceiver are powered off.

In High Power and Low Power states, the IMU and NIR receiver are powered on and functional because they are used to trigger the state change. In System Off state only the IMU is powered on. The NIR transmission rate (one beacon every 5 s) is selected to reduce the probability of collisions in presence of multiple tags. By contrast, we adopt a transmission period of 20 s in the Low Power state to save energy but be reactive in case of interaction and do not severely sacrifice the temporal granularity of collected data. Additionally, even if NIR transmission has a relatively high power consumption, its duration is short (i.e., few tens of milliseconds) and thus its impact on the overall energy consumption is limited (see power profiles of individual components in § 6.4).

We define two rules for the state transition. *Rule 1*: no interaction has been detected in the last 20 minutes; *Rule 2*: no movement has been detected in the last 20 minutes. Rule 1 is to detect scenarios where people are not in interactions for long (e.g., when completing individual work). Thus there is no need to monitor angle/distance with high granularity and the tag switches to the Low Power state. Rule 2 is to infer when the tag is not in use and triggers the transition to System Off state. Once any above condition is not met, the tag reverts to High Power state. The 20-min window from the last interaction is chosen to avoid missing short contacts with short intervals. Similarly, the 20-min window for body movements prevents the transition to System Off state when the user is stationary for a while with the tag still in use. In § 6.4, we will examine the tag's energy consumption and the benefit of adaptive sampling.

## 5 PROTRACTOR PROTOTYPE

We fabricate 6 Protractor tags using off-the-shelf hardware contained in a 3D-printed case. The final assembled tag (Figure 8(a)) resembles an access badge that can be worn using a lanyard or a clip using the loop in the upper part. It measures  $74 \times 54 \times 15$  mm in size and 40 g in weight (with a 560 mAh battery). Figure 9(a) shows its main internal components, including the NIR sensing module, the battery, the IMU, and the micro-controller. We next describe three key components (NIR sensing module, IMU, and micro-controller) in detail.

**NIR Sensing.** The NIR sensing components are the NIR transceiver and angle detection module, which are hosted by a customized printed circuit board (PCB) we design and fabricate (Figure 8(c) and 8(d)). For the NIR beacon transmitter, we choose OSRAM SFH 4240 [1] as the NIR LED, because it provides  $\pm 60^\circ$  3dB beam angle that enables a wide sensing range. Its wavelength peaks at 950 nm. We use an NPN transistor to driver the LED. We choose Vishay TSOP38238 as the NIR receiver, which includes both the photodetector and pre-amplifier. The receiver outputs low when it senses the carrier frequency 38 kHz. Its output signal is connected to the

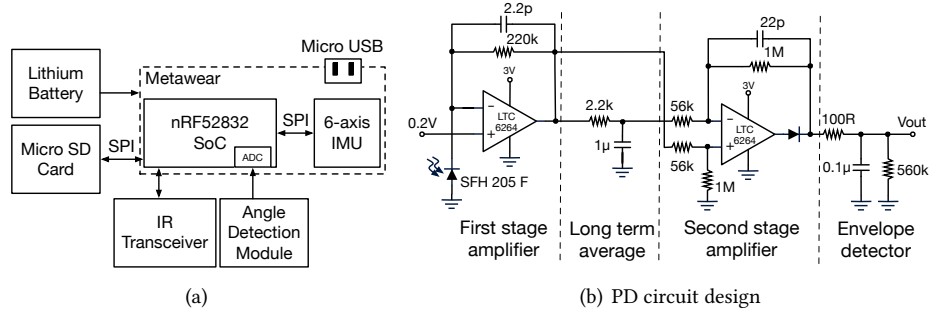


Fig. 9. Protractor design. (a) is the block diagram of the components. (b) is the circuit design of the angle detection module, including a two-stage amplifier, a long-term average to remove the influence of ambient light, and an envelope detector.

micro-controller for decoding. We use the Sony Serial Infra-Red Control (SIRC) protocol (12-bit) to transmit the tag ID every 5 seconds. We select 5s as transmission period to balance power consumption and resolution of the collected data. To prevent collisions in case two or more devices have their transmissions synchronized, we perturb each transmission by adding a random delay (4 – 1020 ms). A collision of multiple NIR beacons makes beacon decoding impossible. In this case the beacon is discarded and it is not used to infer angle and distance, hence not affecting the inference accuracy.

The angle detection module has two NIR photodiodes (OSRAM SFH 205 F [5]) with spectral range of sensitivity from 800 nm to 1100 nm. They are arranged on a 3D-printed base and their orientations form a  $45^\circ$  angle.

Figure 9(b) shows this module's circuit design including a two-stage amplifier and an envelope detector. We adopted a two-stage amplifier in order to detect the light beacons even in environments with high light levels. The first stage is a transimpedance amplifier with a relatively low gain to avoid saturation in bright conditions. The second stage is a differential amplifier which measures the difference between the average light level (RC network between the two stages) and the instant light level and amplifies the signal further with a gain of 17.8. This configuration allows to remove the ambient light level which is added to beacon signal and might cause the amplifier to saturate preventing a correct measure of the amplitude of the signal.

**Inertial Measurement Unit.** We use the Bosch BMI160 6-axis IMU that embeds an accelerometer and gyroscope in the same package. The IMU operates with low power (around  $950\mu A$  with accelerometer and gyroscope in full operation mode) and contains an on-board FIFO buffer where sensor readings can be accumulated without CPU intervention. This allows the micro-controller to sleep for longer periods, leading to a longer battery life. Accelerometer and gyroscope are sampled at 25Hz.

**Micro-Controller.** All components are controlled by a Nordic's nRF52832 SoC that includes a 32-bit ARM-M4F CPU and a 2.4GHz radio transceiver. We use a nRF52832 developer board from Mbienlab Inc. that contains the main SoC, the Bosh IMU and associated circuitry. We attach a micro SD card socket to the SoC using the Serial Peripheral Interface (SPI). The entire device is powered by a 560mAh 3.7V lithium battery that can be recharged via a micro-USB interface.

The micro-controller samples the output of the two photodiodes (after the amplifier and envelope detection) every 1ms (1kHz) using the on-board 14-bit ADC and logs the data on the SD card. The sampling is stopped during the transmission of NIR beacons to avoid the detection of false pulses from the same device.



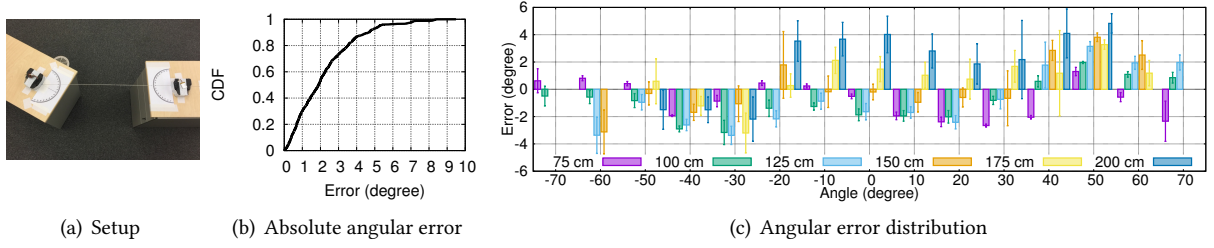


Fig. 10. Protractor's accuracy in estimating interaction angles.

## 6 SYSTEM EVALUATION

We evaluate the systems performance of Protractor prototypes, aiming to examine Protractor's accuracy in determining interaction angle and distance, the impact of practical factors (e.g., differences in user body heights, reflections, ambient light), its scalability with multiple tags, and its energy consumption. We will also examine the efficacy of data fusion in enhancing the tracking reliability.

### 6.1 Accuracy

**Experimental Setup.** We conduct controlled experiments with two static tags to examine Protractor's tracking accuracy using only NIR light. In particular, we place each tag on a different table and support each tag via a piece of foam to emulate the actual usage scenario where tags face each other (Figure 10(a)). Two tags are at the same height and we vary their distance and relative orientation. To obtain the ground truth on the distance  $d$ , we connect tags with a string and measure the string length. To obtain the ground truth on interaction angles  $\theta$ , we place a printed angle meter under each tag to measure their relative orientation. For estimated angle  $\hat{\theta}$  and distance  $\hat{d}$ , we compute the angular and distance error as  $(\hat{\theta} - \theta)$  and  $(\hat{d} - d)$ , respectively. All experiments are indoor with normal lighting (300–400 lux, fluorescent lights).

**Angle.** We start with examining Protractor's accuracy in angle detection. We rotate the table of a tag (tag 1) and keep the other table/tag (tag 2) fixed and facing tag 1. As a result, the interaction angle of tag 1 varies while the interaction angle of tag 2 remains  $0^\circ$ . We vary the interaction angle of tag 1 from  $-90^\circ$  to  $90^\circ$  with  $10^\circ$  interval and the distance from 75 cm to 2 m with a 25-cm step. For each distance/angle combination, we let tags transmit light beacons for one minute. We then compute the interaction angle of tag 1 using the method in § 4.1.

We plot the absolute angular errors (Figure 10(b)) and the error distribution under different angle/distance combinations, where error bars show the standard deviation (Figure 10(c)). Because we rotate each photodiode by  $22.5^\circ$ , one tag will not detect the other's light beacons once the interaction angle exceeds  $67.5^\circ$ . Therefore, the tag's angular sensing range spans approximately from  $-70^\circ$  to  $70^\circ$ . We observe that within the sensing range, the mean error is  $2.2^\circ$  and the 95th percentile is  $5.2^\circ$ , expected to be sufficient for detecting interpersonal contacts. As we scrutinize the error for each angle/distance combination, we observe that large errors occur at long distances and large angles (e.g., 2 m and  $60^\circ$ ) with weak signal strengths. Since ADC's resolution is fixed, the ADC error ratio (error/pulse amplitude) is larger under weaker signals, leading to less precise RSS and larger angular errors.

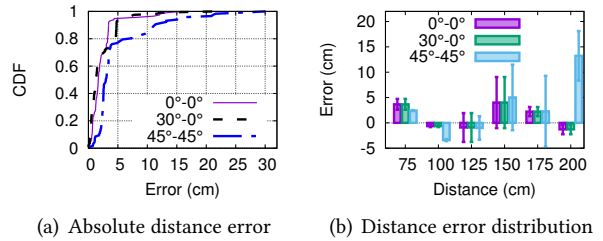


Fig. 11. Accuracy of distance measurement.

**Distance.** We next examine Protractor’s accuracy in ranging. Instead of exhaustively testing all possible combinations of distance and relative angles ( $\approx 1\text{K}$  test cases), we select three representative interaction scenarios with different configurations on the two tags’ interaction angles: (1) face-to-face interaction ( $0^\circ - 0^\circ$ ), (2) one people talking to many others ( $30^\circ - 0^\circ$ ), and (3) two users discussing in front of a white-board ( $45^\circ - 45^\circ$ ). In each scenario, we vary the tag distance from 75 cm to 2 m with 25-cm interval. We then measure the interaction angle at each tag and derive the interaction distance. We plot the CDF of absolute distance errors in Figure 11(a). We observe that the three scenarios have similar mean errors (2.3 cm, 2.4 cm, and 4.9 cm respectively), while scenario (3) has a longer tail, with 11.4 cm as the 90th percentile compared to 3.4 cm and 4.7 cm in the other two scenarios. As we further examine the error distribution across distances for each scenario (Figure 11(b)), we find that the longer tail in scenario (3) is due to the error jump (10 cm) under 2-m distance. The error jumps in this case because the distance (2 m) approaches the sensing limit, and the interaction angle ( $45^\circ$ ) at each tag approaches the half 3dB viewing angle ( $60^\circ$ ) of our photodiode or LED. It results into weak RSS, increasing ADC error ratios and ranging errors.

## 6.2 Robustness

As a chest-worn tag, Protractor can be affected by various practical factors, such as height differences among tags, reflection of NIR light caused by nearby objects (e.g., walls), and ambient light. We now examine the impact of these factors on Protractor’s accuracy, using controlled experiments with the same setup as Figure 10(a).

**Height Offset.** We first examine Protractor’s performance when tags are at different heights. Such height offset can be caused by user’s body height difference, or the way users are wearing tags or interacting with each other (e.g., a sitting user talking to a standing user). For this purpose, we test three settings of tags’ interaction angles ( $0^\circ - 0^\circ$ ,  $30^\circ - 0^\circ$ , and  $60^\circ - 0^\circ$ ) and two distances (75 cm and 125 cm). For each combination, we increase a tag’s height by raising its supporter and vary the height offset from 0 cm to 50 cm, which is proximately the height difference between a sitting user and a standing one. Figure 12(a) shows the angular errors in different combinations of interaction distance and angle. Our main observation is that angular errors do not exceed  $10^\circ$  even under 50-cm height offset, which demonstrates that Protractor’s angle detection is robust against tag height offset. The reason is that without any pitch rotation of the body, the vertical incident angle is the same for both photodiodes and thus has been canceled out (similarly to the  $I_0$  term) in our angular metric (Eq. (2)). The height offset, however, does affect ranging. As shown in Figure 12(b), Protractor increasingly overestimates the distance as the height offset increases. This is because we currently detect only horizontal interaction angles. Thus, height offset leads to a larger vertical angle and higher signal attenuation. Without knowing vertical angles, our method attributes the increase in attenuation to a longer distance. Overall, the maximum distance error caused by height offset is 20 cm. To diminish this error, we can add a pair of photodiodes to detect vertical angles, with the cost of a slightly bigger form factor and higher energy consumption. In this case, even if ambient light (e.g., office lighting) might affect the upper and lower photodiodes unevenly, it will not affect the angle detection because we deduct the background ambient light when extracting the beacon amplitude. We leave this extension to future work.

**Reflection.** Next, we examine how Protractor’s performance is affected by NIR light reflection from nearby objects. In this experiment, we set two tags 1-m away. We then arrange another object in parallel to the line

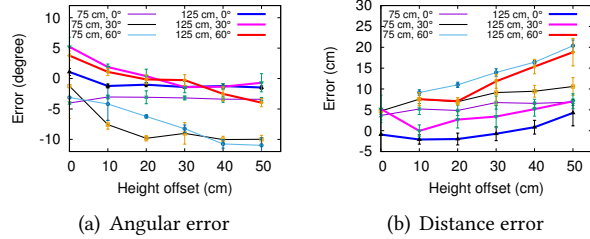


Fig. 12. Influence of height offset.

connecting two tags with a 50-cm perpendicular distance. We test two interaction angles ( $0^\circ$  and  $30^\circ$ ) for tag 1<sup>8</sup> while keeping tag 2's interaction angle as  $0^\circ$ . We test three types of reflection objects: human bodies, screens, and walls. We conduct the experiment in a large office for the former two and in a corridor (1.8-m width) for walls.

In Figure 13, we plot angular and distance errors for tag 1 under different types of reflection objects, where error bars show the standard deviation. As a reference, we also include the result when no reflecting objects are nearby. We make three main observations. *First*, reflection consistently causes underestimates of tag 1's interaction angles. This is because reflection strengthens the RSS perceived by the photodiode closer to the reflection object, which biases the incident light towards the reflection objects. *Second*, among different reflection objects, walls better reflect NIR lights and thus cause larger angular/distance errors, while reflections by human bodies and screens cause absolute errors no more than  $4^\circ$  and 5 cm. *Third*, as for distance errors, wall reflection consistently causes underestimates, because it strengthens RSS and triggers our method to infer shorter distances. On the other hand, reflections by human bodies and screens are weaker and do not necessarily strengthen RSS, leading to possible overestimates. In summary, we observe that only strong reflections by nearby walls present a challenge for Protractor, while smaller objects such as human bodies and screens introduce marginal effects.

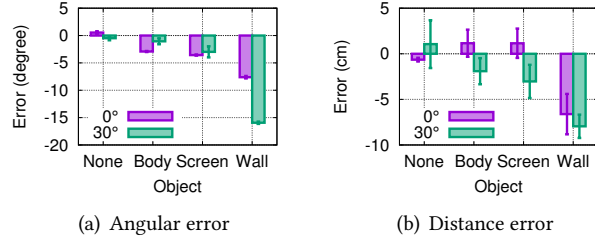


Fig. 13. Influence of reflection.

**Ambient Light.** We also examine the impact of ambient light on Protractor. From our experiments under different levels of indoor lighting, we observe that changes in indoor lighting do not affect Protractor's accuracy in angle detection and ranging. The reason is twofold. First, indoor artificial lights (e.g., fluorescent lights) emit mainly visible light, whereas our NIR sensor [5] is not sensitive to visible light (390 – 700 nm), as its spectral sensitivity range is 800 – 1100 nm. Second, the measured RSS at each photodiode is the amplitude after subtracting the sensed ambient light (§ 4.1). Thus ambient light changes do not affect estimated angles/distances, as long as photodiodes are not saturated.

Table 1. Closest working distance in various ambient lighting.

Ambient light (lux)	250	550	1220
Closest working distance (cm)	10	13	20

However, the saturation problem can occur under high ambient NIR light (e.g., bright sunlight through the window), which affects the closest working distance of our tags. Table 1 lists the closest working distance under different ambient lighting. The result manifests that our system sustains for common social interaction distance (longer than 20 cm) even in bright indoor environment (higher than 1000 lux). We also observe that tags cannot detect light beacons any more when its perceived illuminance exceeds 2500 lux. This level is well below the typical indoor illuminance that ranges between 300 and 500 lux [6, 7]. For comparison in full daylight (not directed towards the sun) there is an illuminance between 10k and 25k lux [103]. We are able to achieve this robustness against variation in ambient light levels thanks to our two-stage amplifier which removes most of the ambient light from the beacon signal.

**Occasional Low Fidelity in NIR Tracking.** We examine the efficacy of data fusion (§ 4.2) in compensating for occasional low-fidelity NIR tracking results. Using the setup in Figure 10(a), we set two tags 1.25-m away

<sup>8</sup>For the  $30^\circ$  angle, we rotate tag 1 towards the reflection object.

facing each other with  $0^\circ$  relative angle. We emulate two cases: occasional blockage of the light channel, and tags temporarily moving outside each other's sensing range.

We test three blockage scenarios by considering whether any tag changes its orientation during the blockage. Figure 14(a) shows three blockage periods: (1) 50"–90": we place a cardboard between tags and then remove it; (2) 130"–190": we place a cardboard, rotate a tag by  $40^\circ$ , and then remove the cardboard; (3) 270"–370": we place a cardboard, rotate a tag by  $40^\circ$ , rotate it back, and then remove the cardboard. The tag orientation remains in period (1), while it changes once and twice in period (2) and (3), respectively. We observe that although NIR angular results are absent during the blockage periods, our data fusion can immediately and accurately extrapolate missing angles using the prediction step<sup>9</sup>. Scenario (2) and (3) also demonstrate the necessity of data fusion, which is capable of tracking the orientation change during the blockage. In comparison, methods such as using the most recent NIR angular result would completely miss the orientation change, which can be important non-verbal cues in a social contact.

We next test the scenario when tags move outside the sensing range. Our prior experiment (§ 6.1) shows that the maximum half sensing angle is  $70^\circ$  for our current prototype. Thus, we start with two tags directly facing each other, rotate a tag by  $90^\circ$ , and later rotate it back. Figure 14(b) plots the estimated angle with and without data fusion. We see that estimated angles using NIR sensors alone are around  $67^\circ$ , translating into  $-23^\circ$  error. With data fusion, the estimated angle is  $85^\circ$  with only a  $-5^\circ$  error. Overall, our results validate data fusion's efficacy in augmenting NIR tracking when NIR tracking is not available or reliable.

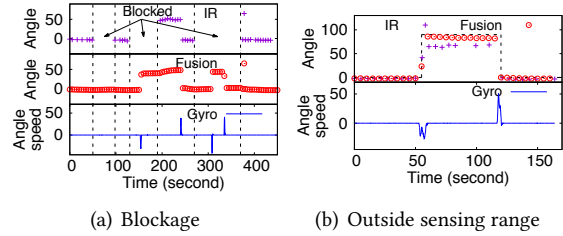


Fig. 14. Data fusion of NIR and inertial sensors.

### 6.3 Scalability

After extensive experiments with two static tags, we now examine the scalability of our design with more than two tags. The presence of more tags can increase the likelihood of NIR beacon collisions, during which signals from multiple NIR beacons add up, potentially causing errors in signal measurements and the decoding of NIR beacons. However, since our system discards collided beacons, beacon collisions do not affect the accuracy of ranging and angle estimation (our prior accuracy results with two tags hold); rather, they affect only the temporal granularity of the data.

To examine the efficacy of our system design – low transmission rate of beacons and random transmission delay (§ 5) – in reducing beacon collisions, we perform a test with six tags. We set up the tags on a table in two rows, where the front row is 80 cm away from the second row. Tags send and receive beacons with the configured transmission rate (0.2 Hz) for 21 hours. For each pair of devices (30 pairs in total), we compute the percentage of received beacons that are successfully decoded. Overall, we observe that the average success rate is 79.5% with 78.3% as the minimum and 80.8% as the maximum. The average duration between received beacons is 6.3 seconds. We conduct similar experiments with four tags and the average success rate in beacon decoding is 84.7%. Our results shows that our system gracefully scales to larger number of tags by recording sufficient number of beacons and thus providing satisfactory temporal granularity. To support applications that require denser deployment of tags and finer temporal granularity, we will examine more sophisticated beacon designs that allow beacons to be extracted under collisions. We plan it for future work.

<sup>9</sup>We smooth gyroscope data using a sliding window of length 25.

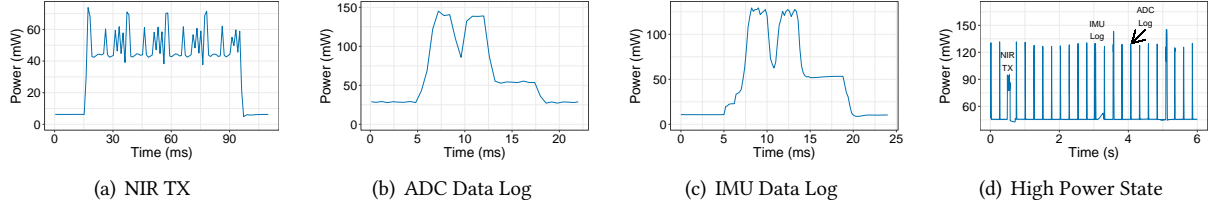


Fig. 15. Power profiles of Protractor's main operations and activities in the High Power state.

## 6.4 Energy Consumption

Finally, we report on the energy consumption of our prototype. We first analyze the power profile of each component using a Monsoon power monitor. Figure 15(a) shows the power trace of NIR beacon transmissions. For each transmission we repeat the same code 4 times to increase the chances of a successful decoding and to have enough data to infer distance and angle. A longer burst (i.e., > 4 beacons) would provide more data for the angle and distance estimation but also increase the power consumption substantially.

Figures 15(b) and 15(c) show the power profiles of the ADC and IMU data logged on the micro SD card. The power consumed by the ADC during a conversion is low ( $\approx 700\mu\text{W}$ ). The constant high power in Figure 15(b) is due to the transimpedance amplifier used to amplify photodiode signals in the angle detection module. This is the most power-demanding component in our prototype. To save energy, we buffer ADC and IMU readings (512-byte and 1024-byte respectively) and then log on the SD card only when the buffers are full. The power consumed by the NIR receiver is negligible in comparison, as it only entails the digital reading of a GPIO pin every  $50\mu\text{s}$ .

We also measure the average power consumed in each of the three power states (§ 4.3) and we obtain: 51.75mW for High Power, 9.42mW in Low Power and 7.96mW for System Off. The tag is powered by a 560-mAh (2.07Wh) battery, however, the battery life of the tag depends on its usage pattern. To estimate the battery life, we compute the average energy consumed per hour as:

$$P_{\text{hour}} = \frac{P_{\text{high}}t_{\text{high}} + P_{\text{low}}t_{\text{low}} + P_{\text{off}}t_{\text{off}}}{24}, \quad (4)$$

where  $t_{\text{high}}$ ,  $t_{\text{low}}$  and  $t_{\text{off}}$  are the number of hours spent respectively in High Power, Low Power and System Off state while  $P_{\text{high}}$ ,  $P_{\text{low}}$  and  $P_{\text{off}}$  are the respective power levels in each state. Assuming on a normal working day a user spends 5 hrs interacting with people<sup>10</sup> (i.e., tag in High Power state), 4 hours on individual work (Low Power state), and does not interact for the rest of the day (System Off state), we can compute the battery life by dividing the battery capacity (560mAh) by  $\frac{P_{\text{hour}}}{3.7\text{V}}$  and obtain an estimated lifetime of about 120 hrs (i.e., 5 days with a single charge). If the device was configured to stay in High Power state (9 hrs per day), without adaptive sampling, the battery would last 85 hrs.

## 7 REAL-WORLD DEPLOYMENT

In this section we explore the potential of Protractor for the use in the context of organizational science and human resource management research. We conduct experiments using a team creativity task.

### 7.1 Task Motivation

Social dynamics happening in small groups of people are usually complex and might affect the performance of teams in corporate settings. Regardless of the impact of the context on social contacts, the common aspect is that *people do not interact with one another as anonymous beings. They come together in the context of specific environments and with specific purposes. Their interactions involve behaviors associated with defined statuses and*

<sup>10</sup>Previous work found that university students spend on average 4.5 hrs per day in face-to-face conversation[76].



*particular roles. These statuses and roles help to pattern our social interactions and provide predictability*[109]. Given the importance of roles in the understanding and of social interactions, in this section we propose to employ Protractor for the automatic recognition of roles people take while working on a defined task. We conducted experiments in a controlled setting, where we assigned participants an creative problem-solving task widely used for assessing teams' creative potential [115], to simulate a team working together within an organizational environment (e.g. in new product development).

Although human networks and social structures have been featured prominently in the fields of organizational behavior and human resource [84, 94], recent research also highlights the importance of taking actions and tasks as analytical focuses in understanding people working within organizations [46, 92, 93]. This research though has tended to focus on higher-level perspectives such as organizational routines [39, 92], and not yet on leveraging the capacity of sensor technologies to examine micro-space and proxemic behavior as a basis for studying actions [47], which has been the case with other research on interactions and dynamics [90, 91]. The goal of this deployment is to provide an initial exploration of the possibilities offered by Protractor in the understanding of complex, and often abstract processes, comprising multiple, interrelated sets of human actions such as creativity in an organizational environment. In particular we explore the possibility of predicting, using only proxemics information (i.e., angle and distance between pairs of participants), two aspects of team dynamics: (1) **task role**: the verbal role assumed by each participant, and (2) **task timeline**: the different building phases of the creative task.

## 7.2 Experimental Setup

We employed an existing creativity task “The Marshmallow Challenge” [3], which was designed to help teams experience fundamental collaboration dynamics in creative problem solving. We recruited participants from the Computer Laboratory at the University of Cambridge (U.K.), and the Department of Computer Science at Dartmouth College (U.S.). We formed 16 teams of four participants ( $n = 64$ )<sup>11</sup>. 90% of the participants were aged 18 to 29 years old and 79% of our participants were men. The participants have been compensated by entering a raffle for an Amazon voucher (6 vouchers available valued £50 or \$50 each).

The teams were welcomed in the experiment room and then given the instructions and rules for the building process upon receiving and wearing the Protractor tags. The participants were given eighteen minutes to build the tallest freestanding structure using 20 sticks of spaghetti, one yard of tape, one yard of string, and one marshmallow which, most importantly, had to be supported by the freestanding spaghetti construction. All participants wore the Protractor and were video recorded during the entire building process. At the end of the eighteen minutes, the building process was stopped; each participant then filled out a survey assessing the creative exchanges of ideas between herself and every other team member on a five-level Likert scale, and demographics.

To summarize, two main sources of data were collected for this study: (1) angle and distance data recorded by Protractor for every two participants (i.e., for every dyad) at approximately five-second resolution; (2) team members' verbal interactions (i.e., their individual task role described by their verbal exchange) and the timeline of the teams' building process (i.e., the timeline of the building task of each team), which we coded from the video recordings using Atlas.ti<sup>12</sup>.

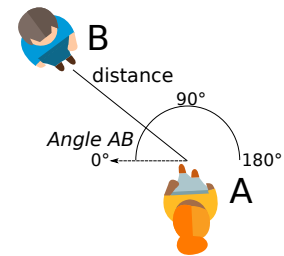


Fig. 16. Illustration of angle and distance in the classification.

<sup>11</sup>Ethical approvals have been obtained from both local institutions before the study.

<sup>12</sup><http://atlasti.com/>.

**Angle and Distance.** The raw angle values detected by Protractor range from  $-90^\circ$  to  $90^\circ$  for every A-B dyad approximately at five-second intervals, where  $0^\circ$  represents participants A and B facing each other, the negative interval indicates B to the left of A, and the positive interval indicates B to the right of A. We rescaled the raw data to the  $[0^\circ, 180^\circ]$  interval to make the results more interpretable whilst preserving the left-right dichotomy. Thus, as shown in Figure 16, in the rescaled dataset,  $90^\circ$  represents participants A and B facing each other,  $0^\circ$  to  $90^\circ$  interval indicates B to the left of A, and  $90^\circ$  to  $180^\circ$  interval indicates B to the right of A. Distance is measured in centimeters and captures the distance between dyads of participants. Just as the angle data, distance is measured approximately at five-second intervals (Figure 16).

**Task Roles.** We coded the team members' verbal behaviors during the building process by using the Advanced Interaction Analysis (*act4teams*) video coding scheme from Lehmann-Willenbrock et al. [77]. This coding scheme has been employed to code verbal behaviors in video recordings of team interactions. The scheme covers four main categories<sup>13</sup> of statements, namely:

- Problem-focused (labeled *ProblFcs*): it identifies communication directly related to the issues subject of the meeting. Problem-focused communication includes discussions about the problems, formulation of ideas, solutions and their analysis. This category includes the following types of statements: identifying a problem, connections with problems, defining the objective, identifying a solution, describing a solution, problem with a solution, arguing for a solution, organizational knowledge, and more.
- Procedural: this kind of communication describes statements related to the structure and organization of the discussion. It is possible to distinguish between Positive and Negative statements. Positive statements are the ones that are beneficial for the organization of the discussion while the Negatives have a negative influence and lead to a loss of structure and loss of thought. In our dataset we label the two as *ProcedPos* and *ProcedNeg* respectively.
- Socio-emotional: it captures the social relationships within a team. Also in this case it is possible to have Positive (labeled *SocEmPos*) and Negative behaviors (labeled *SocEmNeg*). The Positive category includes statements used to show solidarity and support, release tension or agreement. On the other hand, Negative behaviors comprise self-promotion, criticizing, offending or interrupting others and having side conversations that demonstrate disengagement.
- Action-oriented: it describes statements aimed at improving the team's work by showing willingness to take action. Positive statements (labeled *ActOrtPos*) show proactive behavior, willingness of taking responsibility or planning of concrete actions. By contrast, negative statements (labeled *ActOrtNeg*) manifest no interest in change, complaining, lack of initiative, seeking someone to blame or denying responsibility.

These categories describe solution-oriented behavior [63] and have been shown to help teams grow aware of their dynamics in meetings and affect team and organizational success [12, 62, 78]. We use these labels (7 in total) to code our participants' individual verbal statements in the building process at five-second increments. We nominated starting points for each verbal code and assigned these codes to all subsequent time increments; as a new verbal behavior occurred, the new code replaced the previous code in subsequent time increments and so on.

**Task Timeline.** In the original design of the challenge[3], the building phases described are orient, plan, build, and ta-da or oh-no. Empirically, we adapted the phases to collect a more fine-grained taxonomy of the teams' building processes and we labeled our data with:

- (1) *intro* for the introduction time before the actual discussion;
- (2) *materials and logistics* for the discussions covering the tools at hand, planning the building, and starting to put together pieces of structure or to check their strength and stability;

<sup>13</sup>For a more detailed description of the four categories refer to the works from Lehmann-Willenbrock et al. [77] and Kauffeld et al. [62].

- (3) *building levels one, two, three, and four* for assembling the materials and stacking them into the final structure;
- (4) *consolidating level one* for reinforcing the base of the structure to ensure the structure is freestanding;
- (5) *marshmallow on top* for the attempts to place the marshmallow on top of the structure to test the strength of the construction or to finalize it;
- (6) *outro* for the time they finished building to the end of the allocated eighteen minutes.

Once we identified the start and end points of each phase we annotated the angle and distance data within a phase interval with the relative label.

### 7.3 Classification of Video Coded Labels

We classified the video coded task roles and task timeline labels. We treat the data as not formally sequence or temporally dependent, but rather by using all five second increments across all groups as separate instances for classification. We do this to examine the informational value of the angle and distance data collected by Protractor at the most basic level. Classifiers were selected from the WEKA machine learning library [82] and applied as multi-class classifiers using a one-versus-all approach. We initially compared a selection of classifiers (e.g., SVM, MLP, Random Tree, REP Tree and Tree Ensemble classifiers), before selecting the Random Forest classifier (run with 50 and 100 trees, referred to as RF50 and RF100 respectively henceforth) for its simplicity, performance and ability to minimize overfitting [21]. Analysis comprised the following main steps:

- (1) The input features were the angles of each dyad A-B (Angle AB and Angle BA) as well as the distance between A and B collected by the protractor. The data were then normalized to [0, 1] interval.
- (2) Next, we partitioned data into 70/30 splits for classifier training and testing. We applied the SMOTE oversampling procedure [27] to oversample the minority classes. Model performance was assessed with reference to (a) sound precision, recall and F-measure scores across classification targets, (b) reasonable balance between these scores across targets, and (c) good overall model accuracy.
- (3) We further assessed model performance using 10-fold cross-validation with stratified sampling (without any oversampling).

To gain more insights into the role of the features, we also ran a simple forward feature selection loop using the same multi-class Random Forest Classifier with 50 trees (RF50 henceforth). Table 2 lists the results. For the Task Role classes, “distance.b.a” contributes the most to the overall model accuracy (39.45%), followed by varied contributions from the angle features (the highest being “angle.a.d” with 10.62%, and “angle.a.c” with 11.32%), before peaking at eight of nine features (79.46%). For the Task Timeline classes, “angle.b.d” contributes 28.04% to the overall accuracy result, followed by gains of 15.88% (“angle.c.d”), 16.60% (“angle.a.d”), and 11.44% (“distance.b.c”). Thereafter, gains are comparatively modest, peaking at fifteen of eighteen features (91.44%). The intention here is not to show a generalizable pattern of feature contributions, but rather to shed light on the role of the distance and angle features together as markers of focal task related behaviors and interactions.

**7.3.1 Results.** We summarize the overall model accuracy results for RF50 and RF100 in Table 3. Overall accuracy results for cross-fold validation with RF100 are summarized in Table 4. Classes are ordered by F-measure. Next, we discuss the results for each classification task.

**1) Task Role.** The Task Role classes represent the nature of the verbal communication that was taking place among participants during the activities. To predict the instant role of a participant, we use as features her angle towards each other group member (3 features), the angle of each other member toward her (3 features), and the distance between her and each other member (3 features). We concatenate these pairwise dyadic features in a

Table 2. Forward features selection results run with multi-class Random Forest classifier with 50 trees (RF50). The procedure involves adding features iteratively to the model with the goal of maximizing overall accuracy. Overall accuracy at any given feature shows the total model accuracy achieved up until that feature. Change shows the increase / decrease in model accuracy from one feature to the next.

Target	Order	Added Feature	Accuracy	Change	Target	Order	Added Feature	Accuracy	Change
Task Role	9	distance.a.d	79.40%	-0.06%	Task Timeline	18	distance.a.c	88.97%	-1.44%
	8	distance.a.c	79.46%	1.80%		17	distance.c.d	90.41%	-0.41%
	7	angle.d.a	77.65%	2.71%		16	distance.b.d	90.82%	-0.62%
	6	angle.a.b	74.95%	3.72%		15	angle.d.a	91.44%	1.86%
	5	angle.c.a	71.22%	5.62%		14	distance.a.d	89.59%	0.00%
	4	angle.a.c	65.61%	11.32%		13	angle.d.c	89.59%	-0.82%
	3	angle.a.d	54.29%	10.62%		12	angle.b.c	90.41%	1.24%
	2	angle.b.a	43.67%	4.22%		11	angle.a.c	89.18%	0.62%
	1	distance.a.b	39.45%	39.45%		10	angle.c.a	88.56%	1.13%
						9	angle.d.b	87.42%	2.89%
						8	angle.b.a	84.54%	1.24%
						7	distance.a.b	83.30%	2.27%
						6	angle.c.b	81.03%	5.15%
						5	angle.a.b	75.88%	3.92%
						4	distance.b.c	71.96%	11.44%
						3	angle.a.d	60.52%	16.60%
						2	angle.c.d	43.92%	15.88%
						1	angle.b.d	28.04%	28.04%

Table 3. Overall accuracy of predicting participant's instant task role and group's task timeline. We show precision, recall, and F-measure scores, as well as the overall accuracy when using Random Forest with 50 trees (RF50) and 100 trees (RF100) respectively.

RF50 Multi-class Classifier					RF100 Multi-class Classifier				
Task Role	Recall	Precision	F-measure	Accuracy	Task Role	Recall	Precision	F-measure	Accuracy
ProcedPos	0.84	0.84	0.84	79.3%	ProcedPos	0.85	0.84	0.85	80.7%
SocEmPos	0.72	0.87	0.79		SocEmPos	0.74	0.87	0.80	
ProblFcs	0.81	0.72	0.76		ProblFcs	0.82	0.75	0.78	
ActOrtPos	0.69	0.78	0.73		ActOrtPos	0.70	0.81	0.75	
ProcedNeg	0.65	0.65	0.65		ActOrtNeg	0.64	0.75	0.69	
SocEmNeg	0.61	0.67	0.64		ProcedNeg	0.61	0.67	0.64	
ActOrtNeg	0.43	0.67	0.52		SocEmNeg	0.58	0.68	0.62	
Task Timeline	Recall	Precision	F-measure	Accuracy	Task Timeline	Recall	Precision	F-measure	Accuracy
Intro	0.778	1.000	0.875	91.06%	Intro	0.778	1.000	0.875	91.88%
Materialsandlogistics	0.957	0.935	0.946		Materialsandlogistics	0.962	0.944	0.953	
Buildinglevelone	0.934	0.901	0.917		Buildinglevelone	0.941	0.905	0.923	
Buildingleveltwo	0.883	0.929	0.905		Buildingleveltwo	0.883	0.929	0.905	
Marshontop	0.836	0.868	0.852		Marshontop	0.836	0.885	0.860	
Consolidatinglevelone	0.896	0.883	0.889		Consolidatinglevelone	0.929	0.903	0.916	
Outro	0.750	0.857	0.800		Outro	0.625	0.833	0.714	
Buildinglevelthree	0.906	1.000	0.951		Buildinglevelthree	0.875	1.000	0.933	
Buildinglevelfour	0.750	1.000	0.857		Buildinglevelfour	1.000	1.000	1.000	

Table 4. Model accuracy (recall, precision, and F-measure) of predicting participant's instant task role and group's timeline using Random Forest with 100 trees (RF100) and 10-fold cross validation.

Task Role	Recall	Precision	F-measure	Accuracy	Task Timeline	Recall	Precision	F-measure	Accuracy
ProcedPos	0.96	0.82	0.88	84.9%	Intro	0.75	0.91	0.82	93.2%
ProblFcs	0.80	0.87	0.83		Materialsandlogistics	0.96	0.94	0.95	
SocEmPos	0.73	0.93	0.82		Buildinglevelone	0.96	0.93	0.95	
ActOrtPos	0.72	0.89	0.79		Buildingleveltwo	0.92	0.94	0.93	
ProcedNeg	0.70	0.91	0.79		Marshontop	0.72	0.89	0.80	
SocEmNeg	0.38	0.98	0.55		Consolidatinglevelone	0.96	0.92	0.94	
ActOrtNeg	0.27	0.92	0.42		Outro	0.87	0.96	0.91	
					Buildinglevelthree	0.88	0.97	0.92	
					Buildinglevelfour	0.92	1.00	0.96	

feature vector to perform task-role classification at group level. In total, it leads to 11454 data instances<sup>14</sup>. We list the class distribution of these instances in Table 5.

Table 5. Class distribution of all (11454) instances in the task of classifying individual's instant task role.

Label	ProcedPos	ProblFcs	ActOrtPos	SocEmPos	SocEmNeg	ProcedNeg	ActOrtNeg
Count	5351	3144	1592	1138	110	75	44

As shown in Table 3, the overall accuracy of classifying task role was 79.3% (RF50) and 80.7% (RF100) respectively. Cross-fold validation with RF100 achieved overall accuracy of 84.9% (Table 4). Recall and precision are strong among all classes with ProcedPos scoring the highest. SocEmNeg, ActOrtNeg, and ProcedNeg score lower because they are minority classes with fewer instances, where SocEmNeg has 110 instances, ActOrtNeg has 44, and ProcedNeg has 75 instances (Table 5). However, our use of SMOTE [27] to oversample minority classes has helped to improve the recall of minority classes, in comparison to the result without any class balancing. This follows our objective of reaching a balance between the scores (precision, recall and F-measure) across classes.

As we further analyze the confusion matrix in Table 6 using RF100, we can see that despite our oversampling of the minority classes, the classifier is still slightly biased towards the majority classes (*ProblFcs* and the *positive* ones), resulting in more predictions of these classes. However, we did not want to oversample the minority classes further, due to the limited number of instances available. At this stage we considered satisfiable a precision of around 70% (or more) for all classes with a loss in recall for the less frequent *negative* classes (Table 3).

Overall, our results show that task-role classes are quite distinguishable, meaning that the angle and distance data parallel the verbal behavior. This adds to the qualitative seminal work on verbal and non-verbal behaviors (i.e., the spatial orientations) that pull people together or push them apart. Most notably, this literature stream that emerged discretely in the realms of anthropology and mental-health in the mid-sixties [30, 48] is now making its way to social psychology, where the interconnectedness of behavioral channels is a new research endeavor [18].

**2) Task Timeline.** The Task Timeline classes represent stages in the building process followed by the participants. To predict the stage in the building process, we examine the configurations (i.e., relative orientations and distances) of all participants in the group, based on the rationale that these configurations vary across different stages of the building process. As examples, in the *intro* phase, participants might have longer distances from each other since they are not yet actively working; in the *materials and logistics* phase, they might come closer to one another and form sub-groups (pairs of people with short distance and angle close to 90°) while they get familiar with the materials or prototype a basic structure. Thus, we concatenate pairwise distances and angles

<sup>14</sup>The total number of instances theoretically is # of people per group × # of groups × # of 5-second intervals in 18 minutes =  $4 \times 16 \times 216 = 13824$ . We obtain a lower number of instances because some groups finished the challenges before the 18-minute mark and participants stopped working on the structure and interacting.

Table 6. Confusion matrix of classifying instant task role of each participant using RF100 while training on 70% of the data (model results in Table 3). We concatenate a participant's instant angle and distance data to all other group members as the feature vector and predict her instant role. The matrix contains 3440 testing instances (30% of the total data).

	ProblFcs	ProcedPos	SocEmPos	ActOrtPos	ActOrtNeg	SocEmNeg	ProcedNeg
ProblFcs	776	123	8	31	1	2	3
ProcedPos	167	1371	23	36	1	7	1
SocEmPos	40	36	252	12	0	0	2
ActOrtPos	45	90	7	334	1	0	1
ActOrtNeg	1	3	0	1	9	0	0
SocEmNeg	2	11	0	1	0	19	0
ProcedNeg	7	2	0	0	0	0	14



of all participants in the group as the feature vector to predict the Task Timeline. It results into 3231 instances, given the five-second resolution of Protractor data<sup>15</sup>. Table 7 lists the class distribution of these instances.

Table 7. Class distribution of all (3231) instances in the task of classifying task timeline.

Label	Buildinglevelone	Consolidatinglevelone	Materialsandlogistics	Buildingleveltwo	Marshontop	Buildinglevelthree	Outro	Intro	Buildinglevelfour
Count	908	701	700	539	183	106	53	28	13

We summarize the overall accuracy of classifying task timeline in Table 3. We observe that the majority of tasks in the timeline can be distinguished well, achieving 91.06% (RF50) and 91.88% (RF100) overall (Table 3). Cross-fold validation with RF100 achieves 93.19% accuracy (Table 4). Recall and precision are strong for most classes, with *Intro* and *Marshontop* having a slightly lower recall (Table 4).

From the confusion matrix in Table 8, we observe that the classifier exhibit inter-class misclassification mainly between *building level one*, *building level two*, *materials and logistics*, *marshmallow on top* and *consolidating level one* (central section of the matrix). One possible reason for this is that the labels we selected might be too fine-grained and represent the same underlining action (e.g. *working on the structure*). A different coding scheme might account for these similarities and aggregate some of the labels we employed for this work.

**7.3.2 Result Comparison to Existing Works.** We also have compared our results to prior similar studies on social dynamics of small groups using different technologies. These prior studies commonly collected much richer data (e.g., speaking turn and prosodic cues [100, 101], head and body activity [101]) in both visuals and audios, while our tags collect only body distances and orientation. The most relevant work is [117], where Zancanaro et al. used cameras and microphones to analyze the roles played by team members in relation to the tasks the group has to face (“Task Area”) and in relation to the functioning of the group (“Socio-Emotional Area”). Using the behavioral traits of speech activity (presence or absence) and fidgeting (e.g. tapping on table, playing with glasses) the authors were able to predict the manually coded “Task Area” and “Socio-Emotional Area” roles of 10 participants with accuracy between 65% and 68% (F-score between 0.52 and 0.55). The problem setting is similar to our classification of Task Role, where we achieve better results with an accuracy of 84.9% and F-score ranging from 0.42 to 0.88 (Table 4).

<sup>15</sup>The total number of instances theoretically is # of groups  $\times$  # of 5-second intervals in 18 minutes =  $16 \times 216 = 3456$ . We obtain a lower number of instances because some groups finished the challenges before the 18-minute mark and participants stopped working on the structure and interacting.

Table 8. Confusion matrix of classifying timeline states, using RF100 while training on 70% of the data (model results in Table 3). We aggregate the dyadic angles and distances of all members in a group as group-level features to predict the current stage of the building process. The matrix contains 973 testing instances (30% of the total data).

	Intro	Materials and logistics	Building level one	Building level two	Marsh on top	Consolidating level one	Outro	Building level three	Building level four
Intro	7	2	0	0	0	0	0	0	0
Materials and logistics	0	203	6	2	0	0	0	0	0
Building level one	0	9	257	3	1	3	0	0	0
Building level two	0	0	7	143	0	12	0	0	0
Marsh on top	0	1	3	3	46	2	0	0	0
Consolidating level one	0	0	6	2	5	196	2	0	0
Outro	0	0	1	1	0	4	10	0	0
Building level three	0	0	4	0	0	0	0	28	0
Building level four	0	0	0	0	0	0	0	0	4

## 8 DISCUSSIONS

### 8.1 Implications for Organizational Science

Our deployment study makes three theoretical contributions to the team dynamics and proxemics literature. First, we identify and validate the conceptual link between fluid spatial arrangements (described by the variation in angle and distance between team members) and communication content. Second, we extend Hall's spatial orientation with distinct dyadic data where left-right orientations demonstrate how proxemics match verbal behavior. Third, we demonstrate the boundaries of "problem solving" as an only-positive verbal team role.

More broadly, we also address recent calls for research in micro-meso level behavioral processes that have overwhelmingly been researched using retrospective self-reports [34, 67]; known for being bias-prone and inaccurate, these tools have been called by social scientists to be supplemented by unobtrusive, data-dense, and continuous measurement systems [44, 114]. By employing the Protractor we were able to accurately observe nuances in proxemic behavioral changes. Practitioners such as trainers, supervisors, or team facilitators can use our findings to better understand their teams' dynamics and intervene (with procedural interpolations) during intense problem-focused discussions to achieve higher creativity.

Finally, Protractor also allows us to study the impact of culture on proxemics behaviors. Hofstede et al. have shown that cultural backgrounds can impact the way people think, feel, and act while working with others [42]. Cultural differences and personal preferences could alter the way people approach others in terms of interpersonal distance and relative orientation. Some cultures for example tend to have closer distances when interacting with strangers than other cultures [48, 105]. These differences could have a significant effect in today's highly international workplaces and would need to be factored in when studying non-verbal cues. We leave these aspects for future work.

### 8.2 Other Applications

We see the potential of such system not only in the support of organizational science research but also for other practical applications. Protractor tags could be used during job interviews or sales training sessions to collect data that can be later analyzed by the trainee in order to assess her behavior and to improve it over time. We also envision the possibility of using this system to provide behavioral real-time feedback during social interactions, similarly to what has been done in the past with cameras in controlled environments [33]. We believe the unobtrusive wearable form factor could ease privacy concerns and potentially reduce biases.

Another area where Protractor could show its strengths is the design of novel human-computer interfaces based on people orientation and movements, also called *Proxemic Interactions*. In this field, expensive motion tracking systems (e.g., Vicon<sup>16</sup>) are used to create prototypes, but obviously they are not deployable in real applications at scale [15, 23, 112]. Protractor represents instead a viable less expensive option to gather continuous user's orientation and motion without relying on invasive cameras. Protractor devices can be integrated into tangible objects and in the environment in order to gather accurate information about people orientations in space and in relation to objects. The data generated by Protractor could then be used to offer innovative interaction paradigms for smart and connected objects. Our current prototype can also rely on the availability of Bluetooth Low Energy which could be used to locate users in the environment.

In our evaluation we have shown the maximum accuracy that we were able to achieve with our prototype. Given that different applications entail different requirements on accuracy and power consumption, our approach can be adapted to various requirements. As an example, by tuning the beacon transmission frequency and ADC sampling rate it is possible to trade temporal granularity and accuracy for battery life. This is beneficial for applications that do not need continuous accurate angle and distance measurements but would prefer a longer

<sup>16</sup><http://www.vicon.com/>

operation period with a single charge. In this situation the beacons can be transmitted less frequently and the transimpedance amplifier can be switched off for longer periods of time.

## 9 CONCLUSION AND FUTURE WORK

We have introduced Protractor, a wearable technology to accurately monitor non-verbal cues in social interactions. The novelty of our approach lies in its ability of detecting relative body orientation via judicious use of near-infrared light and sensor fusion algorithm exploiting inertial sensors. We demonstrated Protractor efficacy with prototype experiments and a user study with 64 participants using a team collaboration task.

Moving forward, we plan for the following future work. *First*, we will conduct larger-scale deployment of Protractor tags over a longer term, aiming to facilitate studies on human social interactions and deepen understandings on the role of non-verbal body languages in social contacts. While we have examined team collaboration with Protractor tags, other applications that we are interested in investigating include sales representative training sessions and meetings between healthcare practitioners and patients. *Second*, on the technology front, we will investigate performance enhancements of Protractor tags, including lower-range solutions to improve accuracy in very bright settings, alternative NIR beacon designs to allow beacon decoding upon collisions, and adaptive duty cycling based on sliding-window average. We will also consider outdoor scenarios where strong NIR light contained in sunlight can saturate light sensors in our tags. We will explore adaptation of light sensor gains to mitigate the saturation problem and broaden the application scenario of Protractor tags.

## ACKNOWLEDGMENT

We sincerely thank the reviewers for their insightful comments that helped improve the paper. This work is in part supported by the Faculty International Travel Grant at Dartmouth College, National Science Foundation under CNS-1552924, and the Alfred P. Sloan fellowship. Alessandro Montanari acknowledges the support of the EPSRC and Qualcomm Inc. through a Research Studentship in Computing. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the funding agencies or others.

## REFERENCES

- [1] [n. d.]. High Power Infrared Emitter (940 nm). [http://www.osram-os.com/osram\\_os/en/products/product-catalog/infrared-emitters,-detectors-and-sensors/infrared-emitters/power-emitter-gt40mw/emitter-with-940-nm/sfh-4240/index.jsp](http://www.osram-os.com/osram_os/en/products/product-catalog/infrared-emitters,-detectors-and-sensors/infrared-emitters/power-emitter-gt40mw/emitter-with-940-nm/sfh-4240/index.jsp). ([n. d.]). Last accessed 6 March 2018.
- [2] [n. d.]. Making a better HC-SR04 Echo Locator. [http://uglyduck.ath.cx/ep/archive/2014/01/Making\\_a\\_better\\_HC\\_SR04\\_Echo\\_Locator.html](http://uglyduck.ath.cx/ep/archive/2014/01/Making_a_better_HC_SR04_Echo_Locator.html). ([n. d.]). Last accessed 6 March 2018.
- [3] [n. d.]. The Marshmallow Challenge. <https://www.tomwujec.com/design-projects/marshmallow-challenge/>. ([n. d.]). Last accessed 6 March 2018.
- [4] [n. d.]. Physical Web on OpenBeacon. <http://get.openbeacon.org/>. ([n. d.]). Last accessed 6 March 2018.
- [5] [n. d.]. Silicon PIN Photodiode with Daylight Blocking Filter. [http://www.osram-os.com/osram\\_os/en/products/product-catalog/infrared-emitters,-detectors-and-sensors/silicon-photodetectors/photodiodes/pin-photodiodes-in-through-hole-package/sfh-205-f/index.jsp](http://www.osram-os.com/osram_os/en/products/product-catalog/infrared-emitters,-detectors-and-sensors/silicon-photodetectors/photodiodes/pin-photodiodes-in-through-hole-package/sfh-205-f/index.jsp). ([n. d.]). Last accessed 6 March 2018.
- [6] 2010. Code of Federal Regulations 29 CFR standard 1926.56. (2010).
- [7] 2011. Standard UNI EN 12464-1, Lighting of workplaces, UNI. (2011).
- [8] 2016. IR receiver modules for remote control systems. (2016). Rev. 1.5.
- [9] Nadav Aharoni, Wei Pan, Cory Ip, Inas Khayal, and Alex Pentland. 2011. Social fMRI: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing* 7, 6 (2011), 643–659.
- [10] Allison E. Aiello, Amanda M. Simanek, Marisa C. Eisenberg, Alison R. Walsh, Brian Davis, Erik Volz, Caroline Cheng, Jeanette J. Rainey, Amra Uzicanin, Hongjiang Gao, Nathaniel Osgood, Dylan Knowles, Kevin Stanley, Kara Tarter, and Arnold S. Monto. 2016. Design and methods of a social network isolation study for reducing respiratory infection transmission: The eX-FLU cluster randomized trial. *Epidemics* 15 (2016), 38 – 55.

- [11] Erwin Aitenbichler and Max Mühlhäuser. 2003. An IR Local Positioning System for Smart Items and Devices. In *Proc. of the 23rd International Conference on Distributed Computing Systems*.
- [12] Joseph A Allen, Nale Lehmann-Willenbrock, and Steven G Rogelberg. 2015. *The Cambridge handbook of meeting science*. Cambridge University Press.
- [13] Y. Arai and M. Sekiai. 2003. Absolute position measurement system for mobile robot based on incident angle detection of infrared light. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 1. 986–991.
- [14] Michael Argyle, Veronica Salter, Hilary Nicholson, Marilyn Williams, and Philip Burgess. 1970. The Communication of Inferior and Superior Attitudes by Verbal and Non-verbal Signals. *British journal of social and clinical psychology* 9, 3 (1970), 222–231.
- [15] Till Ballendat, Nicolai Marquardt, and Saul Greenberg. 2010. Proxemic interaction: designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces*. ACM.
- [16] Ligia Batrinca, Giota Stratou, Ari Shapiro, Louis-Philippe Morency, and Stefan Scherer. 2013. Cicero-towards a multimodal virtual audience platform for public speaking training. In *Intelligent Virtual Agents*. Springer, 116–128.
- [17] Tobias Baur, Ionut Damian, Florian Lingenfelser, Johannes Wagner, and Elisabeth André. 2013. Nova: Automated analysis of nonverbal signals in social interactions. In *Human Behavior Understanding*. Springer, 160–171.
- [18] Geoffrey Beattie and Andrew W Ellis. 2017. *The psychology of language and communication*. Routledge.
- [19] G. Benet, F. Blanes, J.E. Simões, and P. Pálrez. 2002. Using infrared sensors for distance measurement in mobile robots. *Robotics and Autonomous Systems* 40, 4 (2002), 255 – 266.
- [20] Tjeerd W Boonstra, Mark E Larsen, and Helen Christensen. 2015. Mapping dynamic social networks in real life using participants' own smartphones. *Heliyon* 1, 3 (2015), e00037.
- [21] Leo Breiman. 2001. Random Forests. *Machine Learning* 45, 1 (Oct. 2001), 5–32.
- [22] Chloë Brown, Christos Efstratiou, Ilias Leontiadis, Daniele Quercia, Cecilia Mascolo, James Scott, and Peter Key. 2014. The Architecture of Innovation: Tracking Face-to-face Interactions with Ubicomp Technologies. In *Proc. of UbiComp*.
- [23] Frederik Brudy, David Ledo, Saul Greenberg, and Andreas Butz. 2014. Is Anyone Looking? Mitigating Shoulder Surfing on Public Displays through Awareness and Protection. In *Proceedings of The International Symposium on Pervasive Displays*. ACM, 1.
- [24] Ronald S Burt. 2004. Structural holes and good ideas. *American journal of sociology* 110, 2 (2004), 349–399.
- [25] Jean Carletta, Simone Ashby, Sebastien Bourban, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, Vasilis Karaiskos, Wessel Kraaij, Melissa Kronenthal, Guillaume Lathoud, Mike Lincoln, Agnes Lisowska, Iain McCowan, Wilfried Post, Dennis Reidsma, and Pierre Wellner. 2005. The AMI Meeting Corpus: A Pre-announcement. In *Proc. of the Second International Conference on Machine Learning for Multimodal Interaction*.
- [26] Ciro Cattuto, Wouter Van den Broeck, Alain Barrat, Vittoria Colizza, Jean-François Pinton, and Alessandro Vespignani. 2010. Dynamics of person-to-person interactions from distributed RFID Sensor Networks. *PLoS ONE* 7 (07 2010), 1–9.
- [27] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. 2002. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research* 16 (2002), 321–357.
- [28] Tanzeem Choudhury and Alex Pentland. 2002. The sociometer: A wearable device for understanding human networks. In *Proc. of the CSCW'02 Workshop: Ad hoc Communications and Collaboration in Ubiquitous Computing Environments*.
- [29] Tanzeem Choudhury and Alex Pentland. 2003. Sensing and Modeling Human Networks Using the Sociometer. In *Proc. of ISWC*.
- [30] William S Condon and William D Ogston. 1966. Sound film analysis of normal and pathological behavior patterns. *The Journal of nervous and mental disease* 143, 4 (1966), 338–347.
- [31] Marco Cristani, Loris Bazzani, Giulia Paggetti, Andrea Fossati, Diego Tosato, Alessio Del Bue, Gloria Menegaz, and Vittorio Murino. 2011. Social interaction discovery by statistical analysis of F-formations. In *Proceedings of the British Machine Vision Conference*.
- [32] Amy JC Cuddy, Caroline A Wilmuth, Andy J Yap, and Dana R Carney. 2015. Preparatory power posing affects nonverbal presence and job interview performance. *Journal of Applied Psychology* 100, 4 (2015), 1286.
- [33] Ionut Damian, Chiew Seng (Sean) Tan, Tobias Baur, Johannes Schöning, Kris Luyten, and Elisabeth André. 2015. Augmenting Social Interactions: Realtime Behavioural Feedback Using Social Signal Processing Techniques.
- [34] Richard P DeShon. 2012. Multivariate dynamics in organizational science. *The Oxford handbook of organizational psychology* 1 (2012), 117–142.
- [35] Nathan Eagle and Alex (Sandy) Pentland. 2006. Reality Mining: Sensing Complex Social Systems. *Personal Ubiquitous Comput.* 10, 4 (March 2006).
- [36] Ramsey Faragher et al. 2012. Understanding the basis of the Kalman filter via a simple and intuitive derivation. *IEEE Signal Processing Magazine* 29, 5 (2012), 128–132.
- [37] Ramsey Faragher and R Harle. 2014. An analysis of the accuracy of bluetooth low energy for indoor positioning applications. In *Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014)*, Tampa, FL, USA, Vol. 812.
- [38] Anne-Laure Fayard and John Weeks. 2007. Photocopiers and water-coolers: The affordances of informal interaction. *Organization studies* 28, 5 (2007), 605–634.

- [39] Martha S Feldman, Brian T Pentland, Luciana DăĂŽAdderio, and Nathalie Lazaric. 2016. Beyond routines as things: Introduction to the special issue on routine dynamics. *Organization Science* (2016).
- [40] Andrzej Forys, Kyeong Min, Thomas Schmid, Warren Pettey, Damon Toth, and Molly Leecaster. 2013. WRENMining: Large-Scale Data Collection for Human Contact Network Research. In *Proc. of First International Workshop on Sensing and Big Data Mining*.
- [41] D. Frantal and D. Arbula. 2015. Mobile node localization using infrared angle of arrival sensor. In *International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*.
- [42] Hofstede Geert, Hofstede Gert Jan, and Michael Minkov. 1991. *Cultures And Organizations: Software of The Mind*. McGrawHill USA.
- [43] Jon Gjengset, Jie Xiong, Graeme McPhillips, and Kyle Jamieson. 2014. Phaser: Enabling Phased Array Signal Processing on Commodity WiFi Access Points. In *Proc. of MobiCom*.
- [44] James A Grand, Michael T Braun, Goran Kuljanin, Steve WJ Kozlowski, and Georgia T Chao. 2016. The dynamics of team cognition: A process-oriented theory of knowledge emergence in teams. *Journal of Applied Psychology* 101, 10 (2016), 1353.
- [45] Paul D Groves. 2013. *Principles of GNSS, inertial, and multisensor integrated navigation systems*. Artech House.
- [46] Thorvald Hærem, Brian T Pentland, and Kent D Miller. 2015. Task complexity: Extending a core concept. *Academy of Management Review* 40, 3 (2015), 446–460.
- [47] Edward T. Hall. 1963. A System for the Notation of Proxemic Behavior. *American Anthropologist* 65, 5 (1963), 1003–1026.
- [48] Edward Twitchell Hall. 1966. The hidden dimension. (1966).
- [49] Andy Harter, Andy Hopper, Pete Steggles, Andy Ward, and Paul Webster. 1999. The anatomy of a context-aware application. In *Proc. of MobiCom*.
- [50] Mike Hazas, Christian Kray, Hans Gellersen, Henoc Agbota, Gerd Kortuem, and Albert Krohn. 2005. A relative positioning system for co-located mobile devices. In *Proc. of MobiSys*.
- [51] Roberto Hoyle, Robert Templeman, Steven Armes, Denise Anthony, David Crandall, and Apu Kapadia. 2014. Privacy behaviors of lifeloggers using wearable cameras. In *Proc. of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*.
- [52] William Huang, Ye-Sheng Kuo, Pat Pannuto, and Prabal Dutta. 2014. Opo: A Wearable Sensor for Capturing High-fidelity Face-to-face Interactions. In *Proc. of SenSys*.
- [53] Pan Hui, Augustin Chaintreau, James Scott, Richard Gass, Jon Crowcroft, and Christophe Diot. 2005. Pocket Switched Networks and Human Mobility in Conference Environments. In *Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-tolerant Networking (WDTN '05)*. ACM, New York, NY, USA, 244–251. <https://doi.org/10.1145/1080139.1080142>
- [54] H. Hung and D. Gatica-Perez. 2010. Estimating Cohesion in Small Groups Using Audio-Visual Nonverbal Behavior. *IEEE Transactions on Multimedia* 12, 6 (Oct 2010), 563–575.
- [55] Hayley Hung and Ben Kröse. 2011. Detecting F-formations As Dominant Sets. In *Proc. of ICMI*.
- [56] Shuja Jamil, Anas Basalamah, and Ahmed Lbath. 2014. Crowdsensing traces using bluetooth low energy (BLE) proximity tags. In *Proc. of the ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*.
- [57] Shuja Jamil, Anas Basalamah, Ahmed Lbath, and Moustafa Youssef. 2015. Hybrid participatory sensing for analyzing group dynamics in the largest annual religious gathering. In *Proc. of UbiComp*.
- [58] Dinesh Babu Jayagopi and Daniel Gatica-Perez. 2009. Discovering Group Nonverbal Conversational Patterns with Topics. In *Proc. of International Conference on Multimodal Interfaces (ICMI-MLMI)*.
- [59] D. B. Jayagopi, H. Hung, C. Yeo, and D. Gatica-Perez. 2009. Modeling Dominance in Group Conversations Using Nonverbal Activity Cues. *IEEE Transactions on Audio, Speech, and Language Processing* 17, 3 (March 2009), 501–513.
- [60] Phillip Jeffrey and Andrew McGrath. 2000. Sharing Serendipity in the Workplace. In *Proceedings of the Third International Conference on Collaborative Virtual Environments (CVE '00)*. ACM, New York, NY, USA, 173–179. <https://doi.org/10.1145/351006.351037>
- [61] Rudolph Emil Kalman. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82, 1 (1960), 35–45.
- [62] Simone Kauffeld and Nale Lehmann-Willenbrock. 2012. Meetings matter: Effects of team meetings on team and organizational success. *Small Group Research* 43, 2 (2012), 130–158.
- [63] Simone Kauffeld and Renee A Meyers. 2009. Complaint and solution-oriented circles: Interaction patterns in work group discussions. *European Journal of Work and Organizational Psychology* 18, 3 (2009), 267–294.
- [64] Mark L Knapp, Judith A Hall, and Terrence G Horgan. 2013. *Nonverbal communication in human interaction*. Cengage Learning.
- [65] Toshihiko Komine and Masao Nakagawa. 2004. Fundamental analysis for visible-light communication system using LED lights. *IEEE transactions on Consumer Electronics* 50, 1 (2004), 100–107.
- [66] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. SpotFi: Decimeter Level Localization Using WiFi. In *Proc. of SIGCOMM*.
- [67] Steve WJ Kozlowski. 2015. Advancing research on team process dynamics Theoretical, methodological, and measurement considerations. *Organizational Psychology Review* 5, 4 (2015), 270–299.
- [68] David Krackhardt and Jeffrey Hanson. 1993. Informal Networks: The Company Behind the Chart. *Harvard Business Review* 71, 4 (jul/aug 1993), 104–111.



- [69] Robert E Kraut, Robert S Fish, Robert W Root, and Barbara L Chalfonte. 1990. Informal communication in organizations: Form, function, and technology. In *Human reactions to technology: Claremont symposium on applied social psychology*. Citeseer, 145–199.
- [70] Swarun Kumar, Stephanie Gil, Dina Katabi, and Daniela Rus. 2014. Accurate Indoor Localization with Zero Start-up Cost. In *Proc. of MobiCom*.
- [71] Swarun Kumar, Ezzeldin Hamed, Dina Katabi, and Li Erran Li. 2014. LTE Radio Analytics Made Easy and Accessible. In *Proc. of SIGCOMM*.
- [72] Ye-Sheng Kuo, Pat Pannuto, Ko-Jen Hsiao, and Prabal Dutta. 2014. Luxapose: Indoor positioning with mobile phones and visible light. In *Proc. of MobiCom*. ACM, 447–458.
- [73] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. 2011. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter* 12, 2 (2011), 74–82.
- [74] Mathew Laibowitz, Jonathan Gips, Ryan Aylward, Alex Pentland, and Joseph A Paradiso. 2006. A sensor network for social dynamics. In *Proceedings of the 5th international conference on Information processing in Sensor Networks*. ACM.
- [75] Youngki Lee, Younghyun Ju, Chulhong Min, Seungwoo Kang, Inseok Hwang, and June-hwa Song. 2012. CoMon: Cooperative Ambience Monitoring Platform with Continuity and Benefit Awareness. In *Proc. of MobiSys*.
- [76] Youngki Lee, Chulhong Min, Chanyou Hwang, Jaeung Lee, Inseok Hwang, Younghyun Ju, Chungkuk Yoo, Miri Moon, Uichin Lee, and June-hwa Song. 2013. Sociophone: Everyday face-to-face interaction monitoring platform using multi-phone sensor fusion. In *Proc. of MobiSys*.
- [77] Nale Lehmann-Willenbrock, Joseph A Allen, and Simone Kauffeld. 2013. A sequential analysis of procedural meeting communication: How teams facilitate their meetings. *Journal of Applied Communication Research* 41, 4 (2013), 365–388.
- [78] Nale Lehmann-Willenbrock, Anna Grohmann, and Simone Kauffeld. 2011. Task and relationship conflict at work. *European Journal of Psychological Assessment* (2011).
- [79] Thomas W Leigh and John O Summers. 2002. An initial evaluation of industrial buyers' impressions of salespersons' nonverbal cues. *Journal of Personal Selling & Sales Management* 22, 1 (2002).
- [80] Liqun Li, Pan Hu, Chunyi Peng, Guobin Shen, and Feng Zhao. 2014. Epsilon: A visible light based positioning system. In *Proc. of NSDI*. 331–343.
- [81] Kaikai Liu, Xinxin Liu, and Xiaolin Li. 2013. Guoguo: Enabling fine-grained indoor localization via smartphone. In *Proc. of MobiSys*.
- [82] Machine Learning Group at the University of Waikato. 1997. Weka 3: Data Mining Software in Java. (1997). Retrieved 09/12/2015 from <http://www.cs.waikato.ac.nz/ml/weka/>
- [83] Marianne Schmid Mast. 2007. On the importance of nonverbal communication in the physician–patient interaction. *Patient education and counseling* 67, 3 (2007), 315–318.
- [84] John E Mathieu, John R Hollenbeck, Daan van Knippenberg, and Daniel R Ilgen. 2017. A century of work teams in the Journal of Applied Psychology. *Journal of Applied Psychology* 102, 3 (2017), 452.
- [85] Thomas V McGovern and Howard EA Tinsley. 1978. Interviewer evaluations of interviewee nonverbal behavior. *Journal of Vocational Behavior* 13, 2 (1978), 163–171.
- [86] Albert Mehrabian. 1968. Some referents and measures of nonverbal behavior. *Behavior Research Methods* 1, 6 (1968).
- [87] Masateru Minami, Yasuhiro Fukuju, Kazuki Hirasawa, Shigeaki Yokoyama, Moriyuki Mizumachi, Hiroyuki Morikawa, and Tomonori Aoyama. 2004. *DOLPHIN: A practical approach for implementing a fully distributed indoor ultrasonic positioning system*. Springer Berlin Heidelberg, 347–365.
- [88] Alessandro Montanari, Sarfraz Nawaz, Cecilia Mascolo, and Kerstin Sailer. 2017. A Study of Bluetooth Low Energy Performance for Human Proximity Detection in the Workplace. In *Pervasive Computing and Communications (PerCom), 2017 IEEE International Conference on*. IEEE.
- [89] Daniel Olguin Olguin and Alex Sandy Pentland. 2008. Social sensors for automatic data collection. *Proceedings of the 14th Americas Conference on Information Systems (AMCIS)* (2008).
- [90] Jukka-Pekka Onnela, Benjamin N Waber, Alex Pentland, Sebastian Schnorf, and David Lazer. 2014. Using sociometers to quantify social interaction patterns. *Scientific reports* 4 (2014), 5604.
- [91] Alex Pentland, Tanzeem Choudhury, Nathan Eagle, and Push Singh. 2005. Human dynamics: computation for organizations. *Pattern Recognition Letters* 26, 4 (2005), 503–511.
- [92] Brian T Pentland, Thorvald Hærem, and Derek Hillison. 2010. Comparing organizational routines as recurrent patterns of action. *Organization Studies* 31, 7 (2010), 917–940.
- [93] Brian T Pentland, Alex P Pentland, and Roger J Calantone. 2017. Bracketing off the actors: Towards an action-centric research agenda. *Information and Organization* 27, 3 (2017), 137–143.
- [94] Jill E Perry-Smith and Pier Vittorio Mannucci. 2017. From creativity to innovation: The social network drivers of the four phases of the idea journey. *Academy of Management Review* 42, 1 (2017), 53–79.
- [95] Jill E Perry-Smith and Christina E Shalley. 2003. The social side of creativity: A static and dynamic social network perspective. *Academy of management review* 28, 1 (2003), 89–106.

- [96] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. 2000. The Cricket location-support system. In *Proc. of MobiCom*.
- [97] Nissanka B. Priyantha, Allen KL Miu, Hari Balakrishnan, and Seth Teller. 2001. The Cricket compass for context-aware mobile applications. In *Proc. of MobiCom*.
- [98] Martin S Remland, Tricia S Jones, and Heidi Brinkman. 1995. Interpersonal distance, body orientation, and touch: Effects of culture, gender, and age. *The Journal of social psychology* 135, 3 (1995).
- [99] Irene Rossberg-Gempton and Gary D Poole. 1993. The effect of open and closed postures on pleasant and unpleasant emotions. *The Arts in psychotherapy* 20, 1 (1993), 75–82.
- [100] Hugues Salamin, Alessandro Vinciarelli, Khiat Truong, and Gelareh Mohammadi. 2010. Automatic Role Recognition Based on Conversational and Prosodic Behaviour. In *Proceedings of the ACM International Conference on Multimedia*.
- [101] D. Sanchez-Cortes, O. Aran, M. S. Mast, and D. Gatica-Perez. 2012. A Nonverbal Behavior Approach to Identify Emergent Leaders in Small Groups. *IEEE Transactions on Multimedia* 14, 3 (June 2012), 816–832.
- [102] A. Savvides. 2002. Design of a Wearable Sensor Badge for Smart Kindergarten. In *Proc. of ISWC*.
- [103] Paul Schlyter. 2009. Radiometry and photometry in astronomy. <http://stjarnhimlen.se/comp/radfaq.html>. (2009).
- [104] R. Schmidt. 1986. Multiple emitter location and signal parameter estimation. *Antennas and Propagation, IEEE Transactions on* 34, 3 (March 1986), 276–280.
- [105] Agnieszka Sorokowska, Piotr Sorokowski, Peter Hilpert, Katarzyna Cantarero, Tomasz Frackowiak, Khodabakhsh Ahmadi, Ahmad M Alghraibeh, Richmond Aryeetey, Anna Bertoni, Karim Bettache, et al. 2017. Preferred Interpersonal Distances: A Global Comparison. *Journal of Cross-Cultural Psychology* 48, 4 (2017), 577–592.
- [106] James B Stryker and Michael D Santoro. 2012. Facilitating face-to-face communication in high-tech teams. *Research-Technology Management* 55, 1 (2012), 51–56.
- [107] Wai-Tian Tan, Mary Baker, Bowon Lee, and Ramin Samadani. 2013. The sound of silence. In *Proc. of the 11th ACM Conference on Embedded Networked Sensor Systems*.
- [108] Zhao Tian, Kevin Wright, and Xia Zhou. 2016. The Darklight Rises: Visible Light Communication in the Dark. In *Proc. of MobiCom*.
- [109] Henry Tischler. 1990. *Introduction to Sociology*. Harcourt Brace College Publishers.
- [110] John M Travaline, Robert Ruchinskas, and Gilbert E D’Alonzo. 2005. Patient-physician communication: why and how. *J Am Osteopath Assoc* 105, 1 (2005), 13–8.
- [111] Vishay Semiconductors 2013. *Data formats for IR remote control*. Vishay Semiconductors. Rev. 2.1.
- [112] Daniel Vogel and Ravin Balakrishnan. 2004. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM.
- [113] Benjamin N Waber, Daniel Olguin Olguin, Taemie Kim, and Alex Pentland. 2010. Productivity through coffee breaks: Changing social networks by changing break structure. *Available at SSRN 1586375* (2010).
- [114] Theodore A Walls and Joseph L Schafer. 2006. *Models for intensive longitudinal data*. Oxford University Press.
- [115] Tom Wujec. 2011. *Imagine, Design, Create: How Designers, Engineers And Architects Are Changing Our World*. (2011).
- [116] Jie Xiong and Kyle Jamieson. 2013. ArrayTrack: A Fine-grained Indoor Location System. In *Proc. of NSDI*.
- [117] Massimo Zancanaro, Bruno Lepri, and Fabio Pianesi. 2006. Automatic Detection of Group Functional Roles in Face to Face Interactions. In *Proc. of ICMI*.
- [118] Chi Zhang and Xinyu Zhang. 2016. LiTell: Robust Indoor Localization Using Unmodified Light Fixtures. In *Proc. of MobiCom*.
- [119] Huanle Zhang, Wan Du, Pengfei Zhou, Mo Li, and Prasant Mohapatra. 2016. DopEnc: Acoustic-based Encounter Profiling Using Smartphones. In *Proc. of MobiCom*.

Received August 2017; revised November 2017; accepted January 2018