Multimodal Co-Presence Detection with Varying Spatio-Temporal Granularity

Haus, Michael; Ding, Aaron Yi; Ott, Jorg

DOI
10.1109/PerComWorkshops48775.2020.9156105

Publication date
2020

Document Version
Accepted author manuscript

Published in
Proceedings - 2020 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2020

Citation (APA)
Haus, M., Ding, A. Y., & Ott, J. (2020). Multimodal Co-Presence Detection with Varying Spatio-Temporal Granularity. In *Proceedings - 2020 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2020* Article 9156105 (2020 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2020). IEEE. https://doi.org/10.1109/PerComWorkshops48775.2020.9156105

Multimodal Co-Presence Detection with Varying Spatio-Temporal Granularity

Michael Haus†, Aaron Yi Ding*, Jörg Ott†

†Department of Computer Science, Technical University of Munich, Germany
*Department of Engineering Systems and Services, Delft University of Technology, Netherlands

***Abstract*—Pervasive computing environments are characterized by a plethora of sensing and communication-enabled devices that diffuse themselves among different users. Built-in sensors and telecommunication infrastructure allow co-presence detection. In turn, co-presence detection enables context-aware applications, like those for social networking among close-by users, and for modeling human behavior. We aim to support developers building better context-aware applications by a deepened understanding of which set of context information is appropriate for co-presence detection. We have gathered a multimodal dataset for proximity sensing, including several proximity verification sets, like Bluetooth, Wi-Fi, and GSM encounters, to be able to associate sensor's data with a spatial granularity. We show that sensor modalities are suitable to recognize the spatial adjacency of users with different spatio-temporal granularity. We find that individual user mobility has only a minor, negligible effect on co-presence detection. In contrast, the heterogeneity of device's sensor hardware has a major negative impact on co-presence detection. To reveal energy pitfalls with respect to usability, we perform an energy analysis pertaining to the usage stemming from different sensors for co-presence detection.**

***Index Terms*—Co-presence detection, Multimodal sensor dataset, User mobility, Device heterogeneity, Sensor energy use**

I. Introduction

Portable devices accompany mobile users almost everywhere surrounded by a pervasive wireless infrastructure that is typically composed of Cellular, Wi-Fi, and Bluetooth. Besides providing connectivity, these infrastructures offer an unprecedented opportunity for co-presence detection which is further supported by sensors typically included in mobile devices, such as smartphones and tablets. A quantitative measure of co-presence defines two individuals as "close" when their similarity of context information is large [1]. Context describes any information that can be used to characterise the situation of a person, place, or object that is considered relevant to the interaction between a user and an application [2]. The ability to identify semantically close entities enables context-aware applications, such as data offloading [3], distributed ad hoc networking [4], romantic matchmaking, and social networking [5]. Mobile users of social networks mainly rely on their virtual online communities, which lack the "physical" and contextual interactions among users. We can augment social networks with local interactions by using proximity as a metric to determine who is discoverable on a network of spontaneously and opportunistically connected nodes.

The proximity of mobile devices is typically inferred in two ways: 1) by calculating distances between entities using coordinates from a positioning system, and 2) by computing similarities of context information that is sensitive to the user's activity or location. In this work, we focus on the latter case, i.e., similarity of sensor or wireless data for user's proximity. We address the question, whether multimodal sensor data are suitable to achieve a co-presence detection with varying spatio-temporal granularity? The main problem is a practical dataset, including several proximity verification sets and sensor data of mobile devices where we ensure the physical co-presence of user devices. Therefore, we conducted a study with 126 subjects as part of the lecture on "social computing", over three months, that resulted in a multimodal dataset for co-presence detection (details in Section III). Our study identifies how effective sensor modalities, e.g., barometer, are to detect physical proximity of users with different spatio-temporal granularities. Moreover, we highlight the impact of device heterogeneity due to different sensor hardware and user's mobility involving movement patterns and variability on co-presence detection of spontaneous groups.

We summarize our contributions as follows:

- We gathered a multimodal dataset for co-presence detection, including multiple proximity verification sets, to be able to associate sensor's data with spatial granularity. Our dataset is publicly available [6] as an anonymized subset.
- We show that sensor modalities are suitable for co-presence detection with a signal distance ratio of 5.6x among nearby and remote users. We quantify the impact of device heterogeneity where the co-presence accuracy decreases by 47 %, while the user's mobility has a negligible effect on the co-presence detection.
- We perform an energy analysis on mobile devices to assess the energy demand of different sensors for co-presence detection. The system idle dominates the total energy consumption of the smartphone with 98 %, compared to the phone sensors with only 2 %.

II. Related Work

Our work can be positioned within the field of proximity detection or co-presence detection (used interchangeably in this paper). A number of approaches have been proposed for co-presence detection using the similarity of Bluetooth signals [7], Wi-Fi signals [8], ambient sound [9], images [10], and accelerometer data [11]. Some works concentrate on the estimation of face-to-face interaction among users up to 1.5 m using Bluetooth signals [12], proximity sensors [13], or comparing magnetometer readings to link devices in close proximity of a few centimeters [14]. Other solutions for co-presence detection require infrastructure support such as beacons for Bluetooth low energy (BLE) [15] or ultrasound [16] to emit messages to recognize user's co-presence. In contrast to our work, most of existing studies only analyze device-to-device proximity and not the proximity of device groups, without evaluating the impact of other factors, such as user mobility and hardware heterogeneity on proximity reasoning. The system in [17] is similar to our analysis considering a multitude of sensor data for group detection. However, their work lacks a direct comparison of different sensor data regarding proximity accuracy, e.g., whether the similarity of accelerometer data is higher compared to barometer pressure. We do not aim to mitigate context-manipulation attacks [18] which prevent co-presence detection or trick the proximity reasoning to include remote users into a group of nearby users.

III. Proximity Data Collection

In accordance with ethical requirements, we have undergone a standardized process via the data protection officer of the institute covering appropriate data protection and respecting user's privacy; we conveyed explicitly the purpose of data collection, the type of data gathered from different people, how and where the data is stored, and who would process and use the data. Finally, our data collection is built on participants who approved our privacy agreement to join the study.

A. Sensing Framework

We used the AWARE framework [19] (version 4.0.708) in a client and server setting to gather sensor data from mobile devices. The AWARE server runs on a Linux server located at the department with Apache, MySQL database, PHP, and a Mosquitto MQTT broker, which enables TLS transmissions between student's mobile devices and the AWARE server. On the client side, the AWARE app is available for Android and iOS devices. Table I shows the configured study data to be collected from the student's mobile devices. For the sensor sampling rates, we consider the trade-off between energy and memory consumption, and whether the sampled sensor data being usable for co-presence detection. Our proximity verification sets include GPS and network locations, and Bluetooth, Wi-Fi, and GSM encounters. To obtain a sufficiently fine-grained Bluetooth verification of user's proximity, we distributed 50 BLE beacons over the campus to cover main entrances, lecture halls, library, and cafeteria.

B. Dataset Preparation for Proximity Analysis

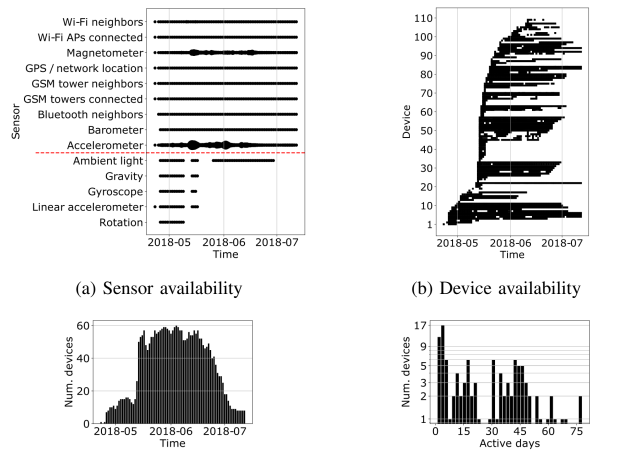
Our proximity data collection contains sensor data from 126 devices. We do not consider the following sensor data for co-presence detection due to limited (few users) sampling points: ambient light, ambient temperature, gravity, gyroscope, linear accelerometer, and rotation. This results in a cleaned proximity data collection including 110 mobile devices with 69 % Android devices and 31 % iOS devices. To reduce the demanded resources of I/O and CPU when using the gathered data, we convert the raw data stored in MySQL dumps (122 GiB) to Apache Parquet (23 GiB).

C. Overview of Proximity Data Collection

As our co-presence detection is dependent on sensor data sensed from multiple devices at the same time and place, Fig. 1(a) proofs that most of the desired sensor data are evenly distributed over time being able to infer user's proximity. Additionally, Figs. 1(b) and 1(c) confirm that we have enough active users who contributed sensor data over a longer period. With respect to the possible size of device groups, Fig. 1(d) presents the number of active devices broken down by days of collected data. We select the following wireless and sensor data targeted for proximity reasoning: Wi-Fi access points (APs) connected, Wi-Fi neighbors, Bluetooth neighbors, GSM towers connected, GSM tower neighbors, GPS and network location, accelerometer, barometer, and magnetometer.

TABLE I
Overview of study dataset

Data characteristic	Sampling rate	Study data
User activity	5 s	(Linear) accelerometer
User position	5 min	GPS, network
User environment	5 min	Barometer, magnetometer, temperature, light, gravity, gyroscope, rotation, GSM towers, Bluetooth and Wi-Fi devices



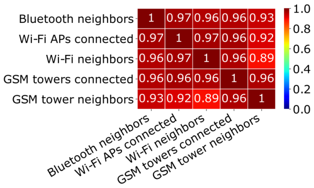
(a) Sensor availability
(b) Device availability
(c) Number of active devices
(d) Histogram of active days

Fig. 1. Overview of proximity data collection

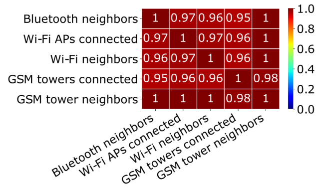
IV. Verification Sets for Proximity Reasoning

A. Performance Comparison of Proximity Verification Sets

We use several proximity verification sets to check the spatial adjacency of device groups inferred by sensor data,



(a) Close-by devices



(b) Remote devices

Fig. 2. Pearson correlation among different proximity verification sets for performance comparison of co-presence detection

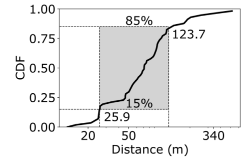
such as magnetometer. To be able to compare the verified co-presence results deduced from sensor data, we analyze that each proximity verification set achieves a similar performance in detecting device groups. Hence, we compute the correlation of verified device groups based on Bluetooth neighbors, Wi-Fi APs connected, Wi-Fi neighbors, GSM towers connected, and GSM tower neighbors as shown in Fig. 2. By using a moving time window of two hours over the proximity verification sets, we limit the runtime and identify close-by devices for each proximity verification set via encountered Bluetooth devices, Wi-Fi APs, or GSM cell towers. Afterwards, we remove duplicate device groups and determine remote devices. We compute the mean Pearson correlation for both device groups: close-by and remote devices, if we recognized device groups from at least two different proximity verification sets for each time window, e.g., Wi-Fi and GSM neighbors.

For devices in proximity, Fig. 2(a) shows the mean Pearson correlation between different proximity verification sets. The correlation is evenly distributed over all verification sets and ranges between 0.89 and 0.97. Fig. 2(b) presents a similar result with a slightly increased correlation among proximity verification sets for remote devices ranging between 0.95 and 1. Additionally, the mean group size of nearby devices amounts to four devices compared to distant device groups with 11 devices.

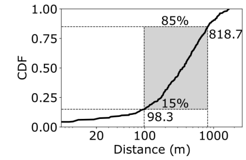
B. Spatial Granularity of Proximity Verification Sets

To be able to associate sensor's data with a spatial granularity for co-presence detection, we compute the geographic expansion of each proximity verification set by using the user's daily moving distance based on encountered location-tagged wireless devices. Therefore, we manually link our self-distributed BLE beacons with latitude and longitude coordinates. For the positions of the Wi-Fi APs, we use a list from our IT department with the MAC address of each Wi-Fi access point and the nearest room number; our institution's room finder provides latitude and longitude for each room on campus. For the positions of the GSM cell towers, we use a publicly available dataset[1]. Based on this, we group the data of each proximity verification set after user devices joining the positions of encountered BLE beacons, Wi-Fi APs, or GSM cell towers. To achieve a more accurate user's daily moving distance, we resample the scans of surrounding Bluetooth devices, Wi-Fi APs, and GSM cell towers by using the median scan period and we take the mean of user positions for each scan. Finally, to get a user's daily moving distance, we sum over the geodesic distances for a series of user positions. Fig. 3 presents the ascending spatial granularity of different proximity verification sets, among 15 % and 85 % of all users.

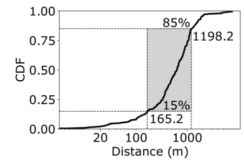
[1] https://www.opencellid.org



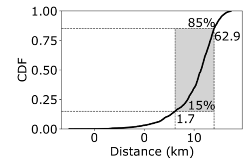
(a) Bluetooth neighbors



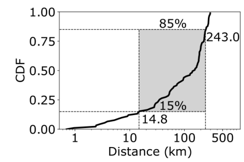
(b) Wi-Fi APs connected



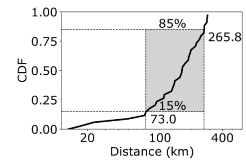
(c) Wi-Fi neighbors



(d) GPS and network locations



(e) GSM towers connected



(f) GSM tower neighbors

Fig. 3. Spatial granularity of proximity verification sets based on users' encounters of location-tagged wireless devices

V. Evaluation of Multimodal Group Proximity

Our aim is to analyze whether sensor modalities are effective for co-presence detection by using proximity verification sets with a variety of spatial granularity, such as Bluetooth, Wi-Fi, and GSM tower neighbors. Our sensor data (barometer, magnetometer, and accelerometer) is dependent on the user's location or activity with the assumption that people share the same context similarity. We define the verification of users' proximity in the following way: at least two or more user devices sense the same wireless device, e.g., Bluetooth device, Wi-Fi access point, and GSM cell tower, within a limited time period like 10 min. To identify remote devices which are not close to each other, we take the difference among all devices and nearby devices within the proximity period. To ensure timely aligned sensor's data across user devices for co-presence detection, our sensing framework performs the clock drift correction of user devices during the daily data upload to the data collection server.

A. Time Periods of Device Encounters for Proximity Detection

As prerequisite for a meaningful co-presence detection, we identify the best encounter times in terms of most devices in proximity, a sufficient number of remote devices, and the largest set with same sensed sensor data across user devices. To compute the aforementioned proximity statistics to select

TABLE II
FOR EACH VERIFICATION SET WE IDENTIFY THE MOST EFFECTIVE SENSOR MODALITY TO DETECT CO-PRESENCE FULFILLING DIFFERENT SPATIO-TEMPORAL GRANULARITY

Verification set	Sensor data	Spatial granularity	Proximity period	Proximity signal distance $\bar{\delta}_p$	Non-proximity signal distance $\bar{\delta}_{np}$	Signal distance ratio $\bar{\delta}_{np}/\bar{\delta}_p$
Bluetooth neighbors	Accelerometer	26–124 m	25 min	992.9	1888.5	1.9
Wi-Fi APs connected	Magnetometer	98–819 m	20 min	652.9	1455.8	2.2
Wi-Fi neighbors	Barometer	165 m–1.2 km	30 min	0.7	13.2	18.3
GSM towers connected	Magnetometer	15–243 km	15 min	772.5	2181.8	2.8
GSM tower neighbors	Magnetometer	73–266 km	30 min	2906.3	9353.8	3.2

the best encounter times for co-presence detection, we use a moving non-overlapping time window of two hours for each combination of proximity verification set, sensor data, user groups with different mobility, and device groups with varying sensor hardware. Given the dataset diversity and the nature of proximity detection over short periods, we choose an empirical two hour time window to strike a balance between granularity and fidelity, comparing with the proximity time window of 5–30 minutes.

B. Results of Co-Presence Detection

We perform our co-presence detection for the best encounter times of 55 different parameter sets, including proximity verification sets, sensor data, user groups, and device groups. We use multiple proximity periods $\in$ [5, 10, 15, 20, 25, 30] min to evaluate the time granularity of user's co-presence. To verify the user's proximity, two user devices have to encounter the same wireless device, e.g., access point or BLE beacon, within the proximity time window. We aim at recent proximity encounters and hence set the time range to 5–30 min.

We use the dynamic time warping distance named δ to compute the similarity of sensor data across different user devices. For comparison, we calculate the signal similarity within each group of devices in proximity, defined by the proximity verification set and between each device in proximity and all distant devices. We take the mean of signal distances among close-by devices $\bar{\delta}_p$ as well as between remote devices $\bar{\delta}_{np}$. Our assumption is that the signal similarity among close-by devices is higher compared to that of remote devices. We use the raw sensor signal to evaluate the basic performance of our co-presence detection.

As a result, based on the maximum signal distance ratio $\bar{\delta}_{np}/\bar{\delta}_p$ between the proximity $\bar{\delta}_p$ and non-proximity signal distance $\bar{\delta}_{np}$, Table II presents the most effective sensor modality and proximity period for each proximity verification set with a varying spatial granularity. A signal distance ratio of one means no difference in the signal similarity among nearby and remote devices, hence co-presence detection is not possible. The larger the signal distance ratio the better for proximity reasoning. The proximity period shows how much time elapses before we can infer the most effective co-presence detection. In addition, we are able to associate sensor's data

TABLE III
SPATIAL GRANULARITY AND PROXIMITY PERIOD FOR EACH SENSOR MODALITY

Sensor data	Spatial granularity	Proximity period	Proximity signal distance $\bar{\delta}_p$	Non-proximity signal distance $\bar{\delta}_{np}$	Signal distance ratio $\bar{\delta}_{np}/\bar{\delta}_p$
Accelerometer	26–124 m	25 min	992.9	1888.5	1.9
Barometer	165 m–1.2 km	30 min	0.7	13.2	18.3
Magnetometer	165 m–1.2 km	30 min	4320.1	28337.9	6.6

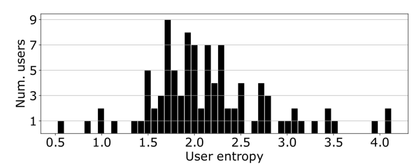


Fig. 4. Distribution of user entropy based on positions from encountered and connected Wi-Fi access points

with a spatial granularity and proximity period as shown in Table III, given by the maximum signal distance ratio.

To sum up, we see a clear distinction of signal similarity among nearby and distant devices, allowing for co-presence detection with a varying spatio-temporal granularity. Via the identified spatial granularity and proximity period, depending on the use case, developers of context-aware applications are able to choose the appropriate wireless or sensor data for co-presence detection. For instance, the magnetometer data offers the most diverse spatial granularity of user's co-presence from a few hundred meters up to kilometers, compared to the accelerometer data with a working range between 30–100 m.

C. Impact of User Mobility on Co-Presence Detection

To evaluate the impact of user mobility on co-presence detection, we compute the mean entropy per user based on positions of encountered wireless devices (e.g., BLE beacons and Wi-Fi APs) or directly via sensed GPS and network positions. We use the entropy to define the user mobility via random device encounters at different locations covering both movement patterns and variability; this is better than the user's

TABLE IV
IMPACT OF USER MOBILITY ON CO-PRESENCE DETECTION

Sensor data	User entropy	Mean proximity period	Mean signal distance ratio $\bar{\delta}_{np}/\bar{\delta}_p$
Accelerometer	1.9	19.2 min	1.3
	2.2	21.2 min	1.6
	3.1	17.5 min	1.3
Barometer	1.9	17.5 min	9.3
	2.2	17.5 min	9.5
	3.1	–	–
Magnetometer	1.9	18.8 min	3
	2.2	18.1 min	1.9
	3.1	21.9 min	2.1

daily moving distance, which neglects the randomness of user behavior. For instance, a user moves several hundred meters each day but only between two positions resulting in a higher moving distance but low entropy. The randomness of user's mobility is more crucial for co-presence detection.

We compute the user entropy for each proximity verification set based on users' encounters of location-tagged wireless devices and apply x-means clustering to find user groups with different mobility behavior. We select user groups inferred by the connected Wi-Fi APs because we can cover 88.2 % of all users, i.e., 97 of 110 users. Fig. 4 shows the distribution of user entropy including two user groups: 34 users with a mean entropy of 3.05 meaning high mobility and 63 users with a mean entropy of 1.86 meaning low mobility. For comparison, we treat all users as a third binned user group with a mean entropy of 2.16 meaning medium mobility.

For each sensor, Table IV shows the impact of user mobility on co-presence detection using three user groups with different mobility behavior. For the barometer sensor, the user group with the highest mobility entropy was too sparse and no proximity encounters could be found. The user entropy, reflecting the users' mobility, has only a minor impact on the mean proximity period and mean signal distance ratio of each user group. For example, there is no trend wherein users with a higher mobility have more or less encounters with other users compared to less randomly moving users.

D. Impact of Device Heterogeneity on Co-Presence Detection

We present a two-fold analysis of device heterogeneity, including sensor hardware statistics and quantifying the effect of different sensing ranges and sensitivities of device sensors on our co-presence detection using sensor's signal similarity.

Regarding the diversity of mobile device's sensor hardware, Fig. 5(a) illustrates that 70 % of all device sensors are produced by only three vendors and Fig. 5(b) shows per sensor that only 30 % of all device sensors are unique and only 10 % of all device sensors are from different vendors. In more detail, Fig. 5(c) presents sensor components and vendors for each sensor, on average, we have 17 unique sensors from six vendors. Many mobile devices from different manufacturers are using the same sensor hardware. This leads to a reduced impact of device heterogeneity on our co-presence detection.

TABLE V
IMPACT OF DEVICE DIVERSITY ON CO-PRESENCE DETECTION

Sensor data	Sensor hardware	Mean proximity period	Mean signal distance ratio $\bar{\delta}_{np}/\bar{\delta}_p$
Accelerometer	mixed	16.9 min	1.2
	same	22.2 min	1.5
Barometer	mixed	17.5 min	2.9
	same	17.5 min	15.9
Magnetometer	mixed	21.9 min	1.7
	same	17.9 min	2.8

Besides that, we quantify the impact of device heterogeneity on our co-presence detection. To this end, we enrich nearby devices defined by our proximity verification set with sensor names or device models, in case of missing hardware information. Different device models like iPhone 6, iPhone 6s, and iPhone 6s Plus are handled as one device model because they use the same sensor hardware[2]. We split the nearby devices according to their sensor hardware or device model to achieve a potentially higher signal similarity among devices in proximity using only the same sensor hardware. We treat remote devices as one device group regardless of their sensor hardware. Table V presents the impact of device diversity on co-presence detection. Per device group, the mean proximity period remains the same across different sensor hardware. In contrast, the mean signal distance ratio increases over device groups separating close-by and distant devices, as expected, if we are only using devices with the same sensor hardware for each proximity group. The co-presence detection with accelerometer data slightly improves by 1.25x, similarly to magnetometer data with 1.6x. The barometer data achieves the greatest improvement with a 5.5x greater signal distance ratio, compared to mixed sensor hardware for co-presence detection.

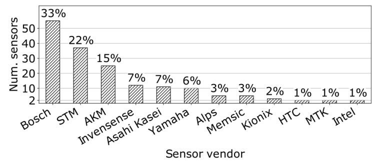
VI. ENERGY ANALYSIS FOR CO-PRESENCE DETECTION

With respect to usability, we analyze the impact of the energy consumption of different device sensors for proximity sensing on the limited battery capacity of mobile devices.

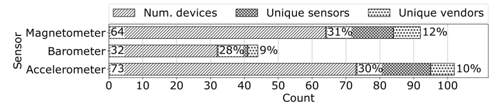
A. Testbed for Sensor Energy Measurements

Our proximity dataset contains five Samsung Galaxy S5 devices (model: SM-G900F). As a sampling device for our sensor energy measurements we use the Samsung Galaxy S5 with Android 6.0.1, in which we replaced the detachable battery with a Monsoon high-voltage power monitor. The Monsoon device directly powers the smartphone with 3.85 V and we take the energy measurements, e.g., time, voltage, and current, via the Python library of the Monsoon power monitor.

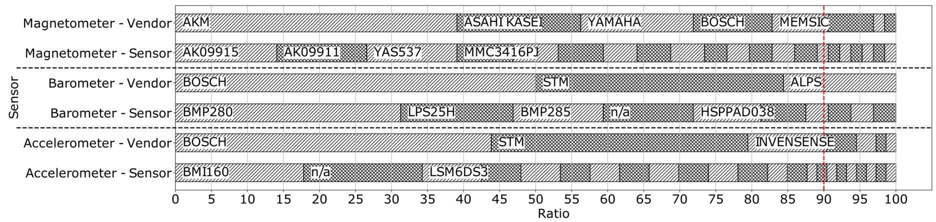
[2]https://www.ifixit.com/Teardown



(a) Ratio of sensor vendors over all devices



(b) Ratio of unique sensor components and vendors per sensor over all devices



(c) Sensor components and vendors per sensor over all devices

Fig. 5. Sensor hardware statistics for device heterogeneity

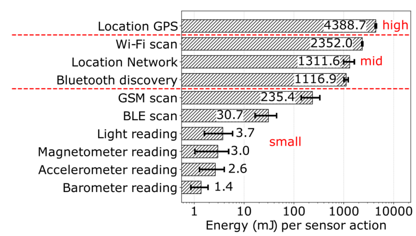
To compute the energy consumption of each device sensor, our Android test application performs different sensor actions, e.g., magnetometer reading and Wi-Fi scan, during the energy measurements lasting one minute in each of our ten evaluation rounds. To purely compute a sensor's energy, we measure the energy used by the smartphone's idle (with disabled wireless connections, including GPS, Wi-Fi, Bluetooth, and GSM) and for a specific sensor's energy we only activate the corresponding wireless interface. For instance, we only activate the Bluetooth interface for Bluetooth discovery or BLE scan. We take as many sensor readings (n_{sensor}) as possible and count them to normalize the consumed energy, resulting in a time-independent energy scale in mJ for each sensor action. The default sampling rate is 10 s or as fast as the sensor is able to provide the information, e.g., GPS location every 30 s.

B. Impact of Sensor Energy Consumption for User's Proximity

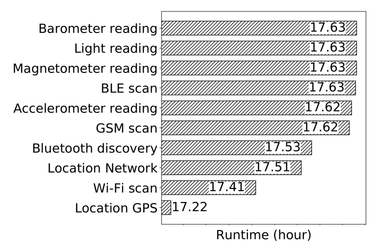
For our energy measurements we record the time, voltage, and current to compute the energy consumption $E(mJ) = U(V) \cdot I(mA) \cdot t(s) = \sum_{i=1}^{n} U_{t_i} \cdot I_{t_i} \cdot (t_{i+1} - t_i)$. We show the energy consumption for each sensor action E_{sa} in Fig. 6(a), defined by $E_{sa} = {}^{E_{\text{sensor}} - E_{\text{system idle}}}/_{n_{\text{sensor}}}$. We apply k-means with three clusters on the median energy consumption for each sensor action, to classify the different sensor actions into three energy consumption levels: small with 34.2 mJ, medium with 1558.5 mJ, and high with 4341.9 mJ as shown in Fig. 6(a). For location requests the network provider determines the device location based on cell towers and Wi-Fi APs, whereas the GPS provider determines the device location using satellites.

We aim to highlight the effect of the sensor's standalone energy consumption on the limited battery capacity of mobile devices. Therefore, we compute the smartphone's runtime t_s using different device sensors for proximity detection as in

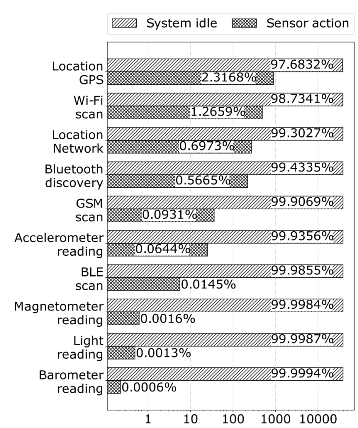
$$\begin{aligned} E_{\text{battery}} &= E_{\text{system idle}} + E_{\text{sensor}} \\ &= U \cdot I \cdot t_s + E_{sa} \cdot {}^{t_s}/_{s_r} \\ \rightarrow t_s &= \frac{E_{\text{battery}} \cdot s_r}{U \cdot I \cdot s_r + E_{sa}} \end{aligned}$$



(a) Standalone sensor's energy consumption with energy levels*



(b) Smartphone runtime using different sensors



(c) Energy ratio between smartphone's system idle and sensor action*

Fig. 6. Energy analysis for co-presence detection (*logarithmic scale)

where E_{sa} is the energy consumption for each sensor action from Fig. 6(a) and s_r is the sampling rate of each device sensor for our proximity data collection in Table I. We use the battery capacity of the Samsung Galaxy S5 with $E_{\text{battery}} = 10.78\,Wh$, defined by $E(Wh) = {}^{Q(mAh)\cdot U(V)}/_{1000}$, electric charge defined as $Q = 2800\,mAh$. For voltage U and current I we take the mean voltage and current from our energy measurements. The battery life of the smartphone in idle state without sensor actions and disabled wireless interfaces is 17.63 hours. Fig. 6(b) shows that the different sensor actions have only a minor effect on the smartphone's runtime t_s. For instance, the standalone sensor's energy consumption to receive GPS locations is 3193x higher compared to barometer readings, whereas the smartphone's runtime only decreases by 25 minutes. Hence, we analyze the ratio of energy consumption between the system idle and sensor actions. Fig. 6(c) highlights that the system idle dominates the total energy consumption of the smartphone with $\approx 98\,\%$. We cannot recognize an effect on the energy relation between system idle and sensor actions if the smartphone is moving or not.

VII. Conclusion

Our study on co-presence detection focuses on using wireless and sensor data from mobile devices. For proximity sensing, we collected a multimodal dataset from 126 participants over three months. We associate the collected data with an effective spatial and time granularity and we identify which sensor data from mobile devices is appropriate for proximity detection. Furthermore, we show that user mobility has only a minor impact on the proximity reasoning and that the device heterogeneity with diverse sensor hardware heavily affects the co-presence detection. Finally, we have conducted an energy analysis of different device sensors for proximity detection. The idle system consumes the most battery capacity of the mobile device while the effect of sensor reading is negligible.

For future work, we plan to further analyze the timely performance of co-presence detection throughout the day, e.g., morning, noon, and evening, and how the user activity, e.g., standing, sitting, and moving, affects the proximity reasoning. Moreover, we plan to enrich our co-presence detection by estimating the social relation among users. For the ground truth of social relationships, we have conducted a survey among the participants of our data collection including the type of relationship, presented as ranked categories, e.g., friend, classmate, and stranger. Our aim is to better understand the social dynamics of a group of people related to proximity, e.g., how the strength of social ties correlates with the spatial adjacency.

Acknowledgment

The authors would like to thank Georg Groh and the study participants for their support in making this work possible. In addition, we thank the reviewers and Leonardo Tonetto for their meaningful feedback.

References

[1] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, "Measurement of Proximity," in *Cluster Analysis*, ser. Wiley Series in Probability and Statistics. Wiley, 2011, pp. 43–69.

[2] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a Better Understanding of Context and Context-Awareness," in *Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing (HUC)*, 1999, pp. 304–307.

[3] A. Pyattaev, K. Johnsson, S. Andreev, and Y. Koucheryavy, "Proximity-Based Data Offloading via Network Assisted Device-to-Device Communications," in *Proceedings of the IEEE 77th Vehicular Technology Conference (VTC Spring)*, 2013, pp. 1–5.

[4] J. Li, J. Jannotti, De Couto, Douglas S. J., D. R. Karger, and R. Morris, "A Scalable Location Service for Geographic Ad Hoc Routing," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2000, pp. 120–130.

[5] Y. Wang, A. V. Vasilakos, Q. Jin, and J. Ma, "Survey on Mobile Social Networking in Proximity (MSNP): Approaches, Challenges and Architecture," *Wireless Networks*, vol. 20, no. 6, pp. 1295–1311, 2014.

[6] M. Haus, A. Y. Ding, and J. Ott. (2020) Proximityness. [Online]. Available: http://proximity.cm.in.tum.de

[7] A. Ghose, C. Bhaumik, and T. Chakravarty, "BlueEye: A System for Proximity Detection using Bluetooth on Mobile Phones," in *Proceedings of the ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp Adjunct)*, 2013, pp. 1135–1142.

[8] P. Sapiezynski, A. Stopczynski, D. Kofoed Wind, J. Leskovec, and S. Lehmann, "Inferring Person-to-person Proximity Using WiFi Signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 2, pp. 1–20, 2017.

[9] W.-T. Tan, M. Baker, B. Lee, and R. Samadani, "The Sound of Silence," in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2013, pp. 1–14.

[10] M. Maier, C. Marouane, M. Klette, F. Dorfmeister, P. Marcus, and C. Linnhoff-Popien, "SURFtogether: Towards Context Proximity Detection Using Visual Features," in *Proceedings of the 3rd International Conference on Context-Aware Systems and Applications (ICCASA)*, 2014, pp. 86–91.

[11] K. A. Nguyen, R. N. Akram, K. Markantonakis, Z. Luo, and C. Watkins, "Location Tracking Using Smartphone Accelerometer and Magnetometer Traces," in *Proceedings of the 14th International Conference on Availability, Reliability and Security (ARES)*, 2019, pp. 1–9.

[12] S. Liu, Y. Jiang, and A. Striegel, "Face-to-Face Proximity Estimation Using Bluetooth On Smartphones," *IEEE Transactions on Mobile Computing*, vol. 13, no. 3, pp. 811–823, 2014.

[13] C. Martella, A. Miraglia, M. Cattani, and M. van Steen, "Leveraging Proximity Sensing to Mine the Behavior of Museum Visitors," in *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2016, pp. 1–9.

[14] R. Jin, L. Shi, K. Zeng, A. Pande, and P. Mohapatra, "MagPairing: Pairing Smartphones in Close Proximity Using Magnetometers," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1306–1320, 2016.

[15] P. C. Ng, J. She, and S. Park, "High Resolution Beacon-Based Proximity Detection for Dense Deployment," *IEEE Transactions on Mobile Computing*, vol. 17, no. 6, pp. 1369–1382, 2018.

[16] B. Thiel, K. Kloch, and P. Lukowicz, "Sound-based Proximity Detection with Mobile Phones," in *Proceedings of the 3rd International Workshop on Sensing Applications on Mobile Phones (PhoneSense)*, 2012, pp. 1–4.

[17] A. A. de Freitas and A. K. Dey, "Using Multiple Contexts to Detect and Form Opportunistic Groups," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW)*, 2015, pp. 1612–1621.

[18] H. T. T. Truong, J. Toivonen, T. D. Nguyen, C. Soriente, S. Tarkoma, and N. Asokan, "DoubleEcho: Mitigating Context-Manipulation Attacks in Copresence Verification," in *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2019.

[19] D. Ferreira, V. Kostakos, and A. K. Dey, "AWARE: Mobile Context Instrumentation Framework," *Frontiers in ICT*, vol. 2, pp. 1–9, 2015.