

Detecting Walking in Synchrony through Smartphone Accelerometer and Wi-Fi traces

Enrique Garcia-Ceja^{2,1}, Venet Osmani¹, Alban Maxhuni¹, and Oscar Mayora¹

¹ CREATE-NET, Via alla Cascata 56D, Trento, Italia

² Tecnológico de Monterrey, Campus Monterrey, Av. Eugenio Garza Sada 2501 Sur, Monterrey, N.L., México

A00927248@itesm.mx, {venet.osmani, alban.maxhuni, oscar.mayora}@create-net.org

Abstract. Social interactions play an important role in the overall well-being. Current practice of monitoring social interactions through questionnaires and surveys is inadequate due to recall bias, memory dependence and high end-user effort. However, sensing capabilities of smartphones can play a significant role in automatic detection of social interactions. In this paper, we describe our method of detecting interactions between people, specifically focusing on interactions that occur in synchrony, such as walking. Walking together between subjects is an important aspect of social activity and thus can be used to provide a better insight into social interaction patterns. For this work, we rely on sampling smartphone accelerometer and Wi-Fi sensors only. We analyse Wi-Fi and accelerometer data separately and combine them to detect walking in synchrony. The results show that from seven days of monitoring using seven subjects in real-life setting, we achieve 99% accuracy, 77.2% precision and 90.2% recall detection rates when combining both modalities.

Keywords: social interactions – accelerometer – wi-fi – ambient intelligence – health and wellbeing

1 Introduction

Interaction between people is one of the basic human activities and it impacts a number of life aspects, including wellbeing. The association between wellbeing and human interaction, specifically social interaction has been well established. Subjects with low quantity of social relationships are typically less healthy, psychologically and physically, manifesting higher risks for a wide range of conditions – from psychiatric disorders to accidents and even mortality [1], while individuals who maintain a certain level of social engagements are shown to be more successful in coping with stress [2]. Several studies have demonstrated that the amount of social activity is negatively correlated with depressive behavior while socialization can help in alleviating depressive symptoms [3, 4].

The current practice of using surveys, diaries and self reporting methods to record social interactions has several drawbacks such as recall bias, memory

dependence, and high end-user effort for continuous long term monitoring [5]. Also, many serious mental illness (SMI) require lifelong management and individuals need to constantly look for signs that might indicate relapse [6], which is infeasible with manual methods. The use of technology can alleviate many of these problems, specifically, wearable devices with embedded sensors, which can continuously record data in an unobtrusive manner.

Miniaturization of sensors makes it possible to collect different types of data. It is common to have wearable devices with several embedded sensors like accelerometers, gyroscopes, Bluetooth and Wi-Fi, to name a few. In recent years, several research works have taken advantage of those sensors to infer different aspects of the user’s context, including physical activities, location and proximity to other users. For example, the accelerometer sensor has been used to recognize physical activities, including walking, running and sleeping [7, 8]. It has been shown that it is also possible to recognize social actions like drinking, laughing, speaking, with a single accelerometer [9]. Another important aspect for context aware systems is *location*. The Bluetooth and Magnetometer sensors have been used to locate a user in indoor environments where GPS usually fails [10, 11, 12, 13]. The Wi-Fi sensor has also been used to classify proximity between mobile devices [14], which is important to identify social structures and relationships [15, 16, 17].

However, the majority of these works have focused on detecting activities that pertain to individual users; that is, establishing statistical models of activities and using sensor data to classify these activities. One issue with this approach is that it is difficult to generalize models across diverse population, since it is challenging to obtain vast set of diverse training data. In contrast, our work focuses on understanding synchrony of activities between people, and we are particularly interested in detecting synchrony of activities when people are in proximity. For this work, we focus on detecting walking activity synchrony using accelerometer data and Wi-Fi traces collected using smartphones worn by subjects in an everyday setting. Our main motivation in detecting walking in synchrony between subjects, is that walking together (going to lunch for example) is an important aspect of the overall social interaction, which has been studied in our previous work [18, 19, 20].

In this work we first used the accelerometer and Wi-Fi data independently to detect walking in synchrony. Then we evaluated performance of our method when combining both modalities. Using both sources of information helped to increase the overall accuracy of detecting joint walking activity.

This paper is organized as follows: Section 2 will present related work in this area. In Section 3 we will describe the methodology which includes: data collection details, pre-processing, accelerometer data analysis and Wi-Fi data analysis. Section 4 presents the experiments and results and finally, in Section 5 we draw the main conclusions.

2 Related Work

The use of sensors to infer user context and interactions has gained a lot of interest in recent years. For example, in [21], the authors used Wi-Fi signal strengths to infer the location of a client and to classify whether or not the client is in motion. To perform the inference, they used Hidden Markov Models (HMMs). They achieved an accuracy of 87% for the motion classification and a median error of 1.5 meters for the location inference. Welbourne et al. [22] also used Wi-Fi signals to infer the user’s place but they also added more sensors like accelerometers and barometric pressure sensors in order to highlight significant places based on the activity that occurs there. The aforementioned works perform the inferences in a user-centric manner, i.e., they do not take into account interactions between users. In [23] they started to explore the use of sensor data from mobile phones to discover the relationships between users. The information they collected includes call logs, Bluetooth proximity devices, cell tower ids, application usage and phone status. In their study, 100 subjects participated in the data collection process over a period ranging from 2 to 7 months and they were able to infer friendship relationships with over 90% accuracy.

Another work that takes into account relations between users is the one of Vu et al. [24]. They used Wi-Fi and Bluetooth data to predict what persons a user will meet, predict future location and stay duration at that location. Sekara & Lehmann [25] demonstrated that weak links obtained from Bluetooth signal strength data have a lower probability of being observed at later times and thus, a lower probability of sharing an online friendship. Carreras et al. [26] used Wi-Fi RSSI to detect proximity between users without the need of any external infrastructure since they introduced a duty-cycle method. This method consists of allowing the devices to alternatively act as Portable Hot Spot (PHS) and client mode.

Social Network Analysis (SNA) visualizations and methods have been used to extract friendship relationships from sensor data. Oloritun et al. [15], predicted friendship ties based on duration of interactions from smartphones data by modeling them as undirected weighted graphs where the weight represents the duration of the interaction. Their results suggest that there is a significant correlation between close friendship and interactions in weekend nights. Another method from SNA is Link Prediction, which consists of predicting future or missing links in a network. Yang et al. [27] used Link Prediction with human mobility information in order to predict future proximity topologies.

In this work we focus on the problem of inferring walking interactions between users. Several of the mentioned works relay on the assumption that two persons interact if they are close to each other, however this may not be the case. It is more likely that any two given persons are interacting if they are walking together than if they are just close to each other so our results may benefit other works that use interaction information to infer higher order social relationships. To the best of our knowledge there are no works that attempt to classify joint walking interactions *between users* by using just accelerometer data or by combining acceleration and Wi-Fi sensors. One of the advantages of using these type of

sensors is that they are non-visual and non-auditory thus, they mitigate privacy concerns and do not interfere with the individual’s daily routines [18].

Table 1 shows a summary of related work classified by the different data sources used and by type. User-centric type means that the work does not take into account the interaction between users but is oriented towards single user i.e., it focuses on context details rather than higher-level interactions. Multi-user means that the work does take into account several users and their possible interactions but does not infer social relationships. This work lies between user-centric and social inference (i.e., multi-user) because it does take into account the interaction between users but it does not try to infer social structures and higher-level relations.

Table 1. Classification of different related works according to data source and type

Work	Data sources	Type
Krumm & Horvitz [21]	Wi-Fi	user-centric
Welbourne et al. [22]	Wi-Fi, accelerometers, pressure sensors	user-centric
Carreras et al. [26]	Wi-Fi	multi-user
This work	Accelerometer, Wi-Fi	multi-user
Eagle & Pentland [23]	Bluetooth, call logs, cell towers, application usage, phone status	social inference
Eagle et al. [28]	call logs, cell towers, Bluetooth	social inference
Vu et al. [29]	Bluetooth, Wi-Fi	social inference
Vu et al. [24]	Bluetooth, Wi-Fi	social inference
Sekara & Lehmann [25]	Bluetooth	social inference
Lerman et al. [17]	Digg, Twitter	social inference
Carrasco et al. [30]	Simulated Wi-Fi	social inference
Oloritun et al. [15]	Bluetooth	social inference
Yang et al. [27]	different datasets	social inference
Yu et al. [16]	Bluetooth, location, voice calls, text messages	social inference
Stopczynski et al. [31]	Questionnaires, Facebook, Phone sensors	social inference

3 Methodology

In this section we describe the details of the data collection process, pre-processing steps and finally the accelerometer and Wi-Fi data analysis.

3.1 Data Collection

The data was collected continuously using two smartphones running Android OS. An application was developed to collect accelerometer and Wi-Fi data from the smartphones' sensors. The sampling rate for the accelerometer sensor was set at 5Hz. This sensor returns the acceleration for each of the x, y and z axes along with a timestamp. The application continually scans Wi-Fi Access Points and stores their Basic Service Set Identifier (BSSID), which is a unique address that identifies each Access Point through MAC address.

The data was collected with 7 subjects in 7 different days during working hours (approx. between 9 a.m and 6 p.m). The experiment was carried out in an office environment where most of the time the users are working on a computer. For each day, one of the smartphones was carried by a *control user* and the other phone was given to a *test user*. There was 1 *control user* and 6 *test users* (one test user collected data for two different days). Three of the *test users* were unknown to the *control user* and the other three *test users* work in the same office as the *control user*. The *control user* took note of the start and end time of each of the walking interactions, i.e., when both users walked together while going for coffee, lunch and other joint walking activities, such as going home. The users were free to use their phone in any way they wanted; that is, there were no restrictions put on the users to carry the phone in a specific manner or at a specific position of the body.

3.2 Pre-processing

An initial examination of raw accelerometer data revealed substantial noise, as expected. Hence, the raw accelerometer data was smoothed with an average filter of length 3:

$$v_s(t) = \frac{1}{n} \sum_{i=t-n}^{t-1} v(i) \quad (1)$$

where v is the original vector, v_s is the smoothed vector and n is the window length. This filter was applied to each of the 3 acceleration axes. Then, in order to take into account the overall movement of the smartphone the magnitude of the acceleration was computed as follows:

$$M(t) = \sqrt{a_x(t)^2 + a_y(t)^2 + a_z(t)^2}, \quad (2)$$

where $a_x(t)$, $a_y(t)$ and $a_z(t)$ are the accelerations at time t for each of the axes. Since each user has its own way of moving (some users tend to move more than others), the magnitude M was standardized ($\mu = 0$ and $\sigma = 1$) relative to each user.

3.3 Acceleration Analysis

The next step was to synchronize in time the data of both users because the application on each phone was not started exactly at the same time and we also needed to take into account the small variations of the internal clocks of the phones. To accomplish this, we divided the time into slots of ten seconds each and computed the mean magnitude for each of the slots such that the data for both users lied in a common time scale. Figure 1 shows the magnitudes for the control user and one of the test users.

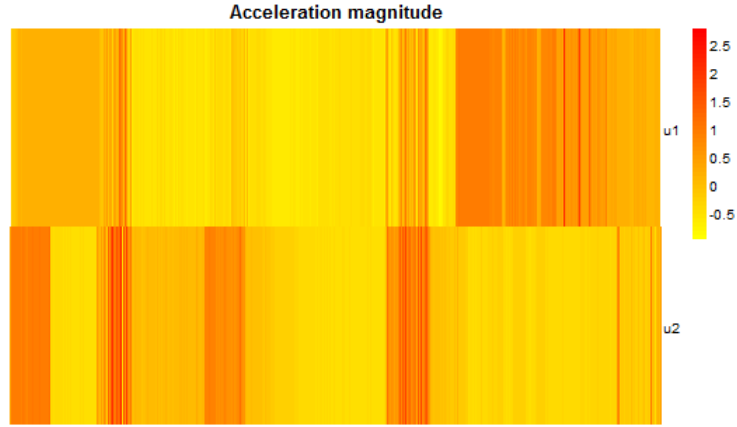


Fig. 1. Heat map of the acceleration magnitudes of the control user and one of the test users. u1 is the control user and u2 represents the test user. The x axis represents time.

Since we are interested in significant changes of acceleration between both users that occur almost at the same time, we now compute the magnitude difference for each user as follows:

$$D(t) = |M(t) - M(t - 1)| \quad (3)$$

Figure 2 shows the result of applying this filter. Now we can see that the magnitude changes are emphasised. We can also see that the red lines are fragmented but we want to arrange them into blocks so it is easier to detect when both users are moving at the same time. To 'group' the fragmented lines into blocks we applied the following filter:

$$V(t) = \Theta \left(\sum_{i=t-(n-1)}^{t-1} |D(i) - D(i + 1)| - \varepsilon \right) \quad (4)$$

where Θ is the Heaviside function, ε is a threshold (set to 1) and n is the window size which was set to 10. Figure 3 shows the result after applying this

filter. Now we can see more clearly, when both users were moving at the same time by looking at the intersection of the blocks.

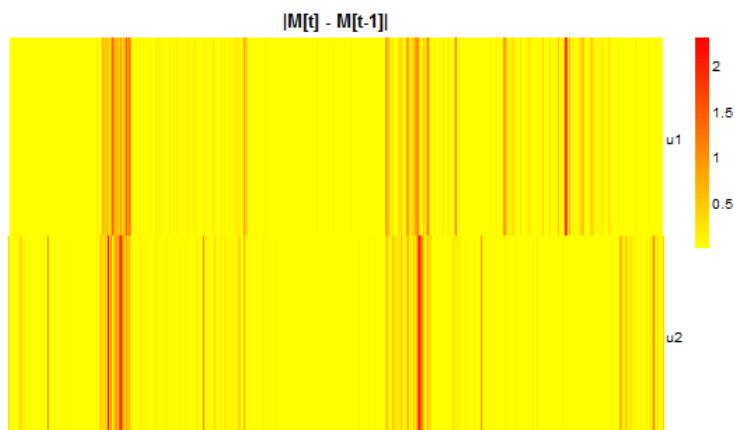


Fig. 2. Heat map of the magnitude differences for each user. u_1 is the control user and u_2 represents the test user. The x axis represents time.

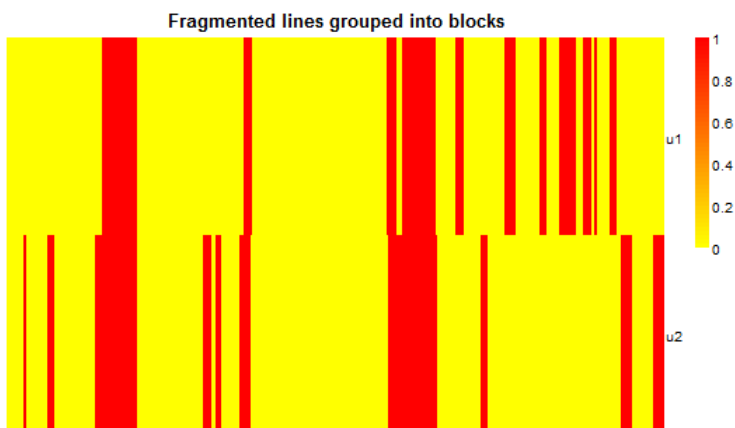


Fig. 3. Resulting blocks after grouping fragmented lines. u_1 is the control user and u_2 represents the test user. The x axis represents time.

Once we have found the resulting blocks we predict a positive walking interaction if the time slot is within a block and $V(t)_{u_1} == 1$ and $V(t)_{u_2} == 1$ where $V(t)_{u_1}$ and $V(t)_{u_2}$ are the resulting vectors after applying Eq.(4) for the *control user* and the *test user*, respectively.

3.4 Wi-Fi Analysis

If a user is walking from one place to another, the difference between the in-range Access Points at time t and time $t-1$ should be greater than that of user staying at the same place. In the latter case, when the user is staying at the same place the variability between Access Points at time t and $t-1$ should be low since the smartphone detects almost the same BSSIDs during scanning. Intuitively, if the variability of Access Points is high for two users at the same time, while sharing a number of common Access Points, it is an indication that they may be walking together. As variability measure we used the Jaccard distance [32]:

$$D_j(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|} \quad (5)$$

where A and B are the sets of BSSIDs at time t and $t-1$, respectively. In the extreme case, when the smartphone detects the same Access Points over time, i.e., the user is not moving, the variability will be 0. On the other hand, if the Access Points at time t are completely different from those at time $t-1$ the variability will be 1. Figure 4 shows the Wi-Fi BSSIDs variance for both users over time. We can see that the variance tends to increase within the red blocks (when the user is moving based on accelerometer data). The green horizontal lines represent the self reported walking interactions. To predict a positive walking interaction we added the variances of both users and if this sum is greater than a given threshold, then it is set as positive. The threshold selection was done using a ROC curve [33] which will be presented in Section 4.

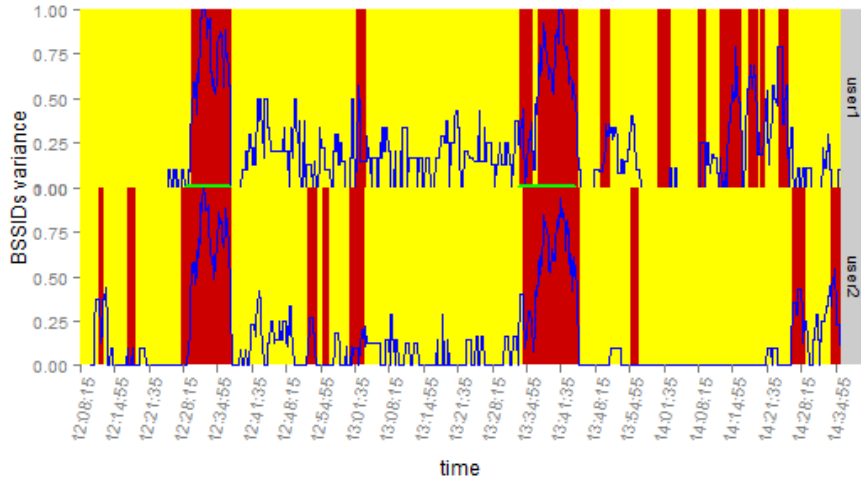


Fig. 4. BSSIDs variance over time. user1 is the control user and user2 represents the test user. The green lines show the self reported walking interactions.



Fig. 5. Predictions when using just accelerometer data for all subjects. Red lines represent negative predictions and Green lines represent positive predictions (walking together).

4 Experiments and Results

In this Section we present the results of the walking interactions detection when using: 1) Accelerometer data; 2) Wi-Fi data; and 3) Accelerometer data combined with Wi-Fi data. We performed the recognition between the control user and each test user (7 trials in total).

Figure 5 shows the resulting predictions when using just accelerometer for 7 different days (quantified confusion matrix is shown in Figure 7). Red lines represent negative predictions and Green lines represent positive predictions (walking together). It can be seen that the classes are unbalanced, i.e., most of the time slots are negatives (97.2%). The overall accuracy of the predictions was 0.984; but given the unbalanced nature of data we also present specificity, sensitivity (also known as *recall*), positive predictive value (PPV) (also known as *precision*) and negative predictive value (NPV). These measures are summarized in Table 2. All metrics are greater than 0.9, except the PPV which was 0.657. The PPV was low because there were many false positives (See Figure 5). The false positives are due to the fact that users were moving or walking at the same time even though they were not in proximity to each other.

Table 2. Results with accelerometer data.

Accuracy	Sensitivity	Specificity	PPV	NPV
0.984	0.902	0.986	0.657	0.997

Now we evaluate the predictions by using just Wi-Fi data. Figure 6 shows the resulting ROC curve ³ for different threshold values of the BSSIDs variance computed with Eq.(5). This plot represents the trade off between sensitivity and specificity for different threshold values. We can see that a good threshold value is around 0.424 since it produces a combination of sensitivity and 1 - specificity that is close to the upper left corner. The diagonal represents a random classification so, a value below the diagonal means that the classifier for that specific threshold performed worse than a random guess. The resulting Area Under the Curve was 0.95 and it represents the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. Based on the ROC

³ The ROC curves were produced with the ROCR package [34]

curve it seems that using just acceleration is better than using just Wi-Fi data. This is because there is no threshold value in the ROC curve that produces a combination of both sensitivity and specificity that are greater or equal to the ones obtained when using accelerometer data.

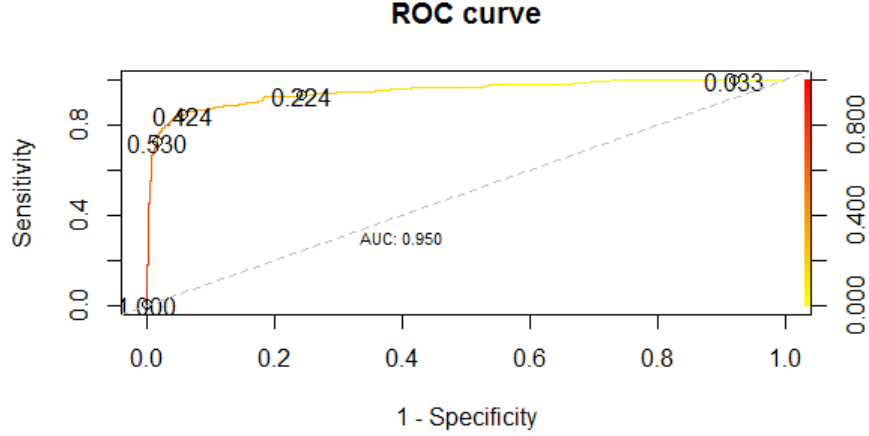


Fig. 6. Resulting ROC curve for different threshold values of the BSSIDs variance

Finally, we performed the prediction by combining both, accelerometer and Wi-Fi data. Many of the false positives are there because the users may be walking at the same time even though they are in different places. We used Eq.(5) to compute the distance between the control user and the test user at any given time slot. If the distance between the users is 1 it means that they are in different places so the prediction can be corrected. Table 3 shows the results when combining both sources of information.

Table 3. Results with accelerometer and Wi-Fi data.

Accuracy	Sensitivity	Specificity	PPV	NPV
0.990	0.902	0.992	0.772	0.997

We can see that the accuracy and specificity increased; the sensitivity and NPV remained the same and the PPV had a significant improvement (from 0.657 to 0.772). Figure 7 shows the resulting confusion matrices when using just accelerometer data and when using accelerometer with Wi-Fi data. We can see that the second column of the matrices remained the same. The first columns of the matrices had a significant change because the false positives were reduced considerably by ruling out detected interactions from accelerometer data that were not real because the users were in different places.

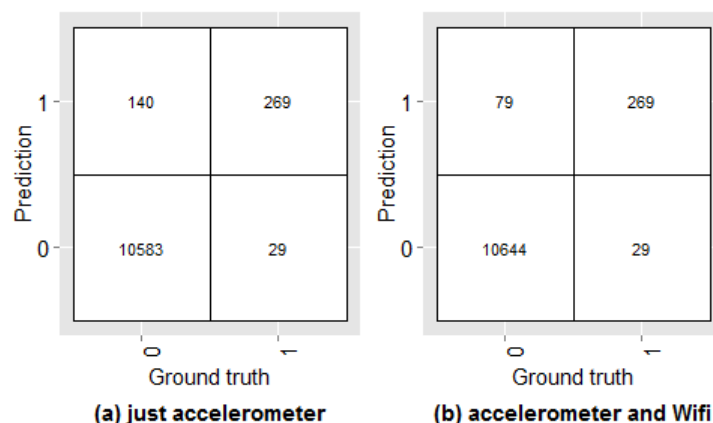


Fig. 7. Confusion matrices when a) using just accelerometer data and b) combining accelerometer and Wi-Fi. 1 means a walking interaction and 0 no interaction.

5 Conclusions and Future Work

In this work we showed how accelerometer and Wi-Fi data collected from smartphones can be used to classify walking in synchrony between users. First, we used accelerometer data and Wi-Fi data independently. Our results suggest that using just an accelerometer gives better results than using just Wi-Fi. Then, we combined both sensors and the overall accuracy increased because when adding Wi-Fi it was possible to reduce the false positives due to the fact that the persons may be walking at the same time but in different places. Our experiments were tested in an office environment and these results are indicative of performance of our method. We plan to extend this work to a larger sample of subjects and investigate other types of interactions in synchrony, including commuting interactions such as being in the same car, bus or train.

Acknowledgements Enrique Garcia-Ceja would like to thank Consejo Nacional de Ciencia y Tecnología (CONACYT) and the AAAMI research group at Tecnológico de Monterrey for the financial support in his PhD. studies. This work was supported by EIT ICT Labs, Social Gym project.

References

1. James S House, Karl R Landis, and Debra Umberson. Social relationships and health. *Science*, 241(4865):540–545, 1988.
2. Sheldon Cohen and Thomas A. Wills. Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, 98(2):310 – 357, 1985.
3. Vivian Isaac, Robert Stewart, Sylvaine Artero, Marie-Laure Ancelin, and Karen Ritchie. Social activity and improvement in depressive symptoms in

- older people: A prospective community cohort study. *The American Journal of Geriatric Psychiatry*, 17(8):688 – 696, 2009.
4. Hayden B. Bosworth, Judith C. Hays, Linda K. George, and David C. Stephens. Psychosocial and clinical predictors of unipolar depression outcome in older adults. *International Journal of Geriatric Psychiatry*, 17(3):238–246, 2002.
 5. Nathan Norfleet Eagle. *Machine perception and learning of complex social systems*. PhD thesis, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 2005.
 6. M. Matthews, S. Abdullah, G. Gay, and T. Choudhury. Tracking mental well-being: Balancing rich sensing and patient needs. *Computer*, 47(4):36–43, April 2014.
 7. Edmond Mitchell, David Monaghan, and Noel E O’Connor. Classification of sporting activities using smartphone accelerometers. *Sensors*, 13(4):5317–5337, 2013.
 8. O.D. Lara and M.A. Labrador. A survey on human activity recognition using wearable sensors. *Communications Surveys Tutorials, IEEE*, 15(3):1192–1209, 2013.
 9. Hayley Hung, Gwenn Englebienne, and Jeroen Kools. Classifying social actions with a single accelerometer. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp ’13*, page 207–210, New York, NY, USA, 2013. ACM.
 10. S. Aparicio, J. Perez, AM. Bernardos, and J.R. Casar. A fusion method based on bluetooth and wlan technologies for indoor location. In *Multisensor Fusion and Integration for Intelligent Systems, 2008. MFI 2008. IEEE International Conference on*, pages 487–491, Aug 2008.
 11. Carlos E. Galván-Tejada, José C. Carrasco-Jiménez, and Ramon F. Brena. Bluetooth-wifi based combined positioning algorithm, implementation and experimental evaluation. *Procedia Technology*, 7(0):37 – 45, 2013. 3rd Iberoamerican Conference on Electronics Engineering and Computer Science, CIECC 2013.
 12. Jaewoo Chung, Matt Donahoe, Chris Schmandt, Ig-Jae Kim, Pedram Razavai, and Micaela Wiseman. Indoor location sensing using geomagnetism. In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services, MobiSys ’11*, pages 141–154, New York, NY, USA, 2011. ACM.
 13. Carlos E. Galván-Tejada, José C. Carrasco-Jimenez, and Ramon Brena. Location identification using a magnetic-field-based FFT signature. *Procedia Computer Science*, 19(0):533 – 539, 2013. The 4th International Conference on Ambient Systems, Networks and Technologies (ANT 2013), the 3rd International Conference on Sustainable Energy Information Technology (SEIT-2013).
 14. Alessandro Carlotto, Matteo Parodi, Carlo Bonamico, Fabio Lavagetto, and Massimo Valla. Proximity classification for mobile devices using wi-fi environment similarity. In *Proceedings of the first ACM international workshop*

- on Mobile entity localization and tracking in GPS-less environments, pages 43–48. ACM, 2008.
15. Rahman O. Oloritun, Anmol Madan, Alex Pentland, and Inas Khayal. Identifying close friendships in a sensed social network. *Procedia - Social and Behavioral Sciences*, 79(0):18–26, 2013. 9th Conference on Applications of Social Network Analysis (ASNA).
 16. Zhiwen Yu, Kingshe Zhou, Daqing Zhang, Gregor Schiele, and Christian Becker. Understanding social relationship evolution by using real-world sensing data. *World Wide Web*, 16(5-6):749–762, 2013.
 17. Kristina Lerman, Suradej Intagorn, Jeon-Hyung Kang, and Rumi Ghosh. Using proximity to predict activity in social networks. In *Proceedings of the 21st International Conference Companion on World Wide Web, WWW '12 Companion*, page 555–556, New York, NY, USA, 2012. ACM.
 18. A Matic, V. Osmani, and O. Mayora. Trade-offs in monitoring social interactions. *Communications Magazine, IEEE*, 51(7):114–121, July 2013.
 19. A Matic, V. Osmani, and O. Mayora. Speech activity detection using accelerometer. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 2112–2115, Aug 2012.
 20. Aleksandar Matic, Venet Osmani, and Oscar Mayora-Ibarra. Analysis of social interactions through mobile phones. *Mobile Networks and Applications*, 17(6):808–819, 2012.
 21. J. Krumm and E. Horvitz. LOCADIO: inferring motion and location from wi-fi signal strengths. In *Mobile and Ubiquitous Systems: Networking and Services, 2004. MOBIQUITOUS 2004. The First Annual International Conference on*, pages 4–13, August 2004.
 22. Evan Welbourne, Jonathan Lester, Anthony LaMarca, and Gaetano Borriello. Mobile context inference using low-cost sensors. In *Proceedings of the First International Conference on Location- and Context-Awareness, LoCA'05*, page 254–263, Berlin, Heidelberg, 2005. Springer-Verlag.
 23. Nathan Eagle and Alex (Sandy) Pentland. Reality mining: Sensing complex social systems. *Personal Ubiquitous Comput.*, 10(4):255–268, March 2006.
 24. Long Vu, Quang Do, and K. Nahrstedt. Jyotish: A novel framework for constructing predictive model of people movement from joint wifi/bluetooth trace. pages 54–62, March 2011.
 25. Vedran Sekara and Sune Lehmann. The strength of friendship ties in proximity sensor data. *PLoS ONE*, 9(7):1–8, 2014.
 26. I Carreras, A Matic, P. Saar, and V. Osmani. Comm2sense: Detecting proximity through smartphones. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pages 253–258, March 2012.
 27. Yang Yang, N.V. Chawla, P. Basu, B. Prabhala, and T. La Porta. Link prediction in human mobility networks. In *Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference on*, pages 380–387, August 2013.
 28. Nathan Eagle, Alex (Sandy) Pentland, and David Lazer. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, 106(36):15274–15278, 2009.

29. Long Vu, Klara Nahrstedt, Samuel Retika, and Indranil Gupta. Joint bluetooth/wifi scanning framework for characterizing and leveraging people movement in university campus. In *Proceedings of the 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, MSWIM '10, page 257–265, New York, NY, USA, 2010. ACM.
30. José C. Carrasco-Jiménez, José M. Celaya-Padilla, Gilberto Montes, Ramon F. Brena, and Sigfrido Iglesias. Social interaction discovery: A simulated multiagent approach. In Jesús Ariel Carrasco-Ochoa, José Francisco Martínez-Trinidad, Joaquín Salas Rodríguez, and Gabriella Sanniti di Baja, editors, *Pattern Recognition*, volume 7914 of *Lecture Notes in Computer Science*, pages 294–303. Springer Berlin Heidelberg, 2013.
31. Arkadiusz Stopczynski, Vedran Sekara, Piotr Sapiezynski, Andrea Cuttone, Mette My Madsen, Jakob Eg Larsen, and Sune Lehmann. Measuring large-scale social networks with high resolution. *PLoS ONE*, 9(4):1–24, 2014.
32. P. Jaccard. *Nouvelles recherches sur la distribution florale*. Bull. Soc. Vaud. Sci. Nat., 1908.
33. Charles E. Metz. Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, 8(4):283 – 298, 1978.
34. Tobias Sing, Oliver Sander, Niko Beerenwinkel, and Thomas Lengauer. ROCR: visualizing classifier performance in R. *Bioinformatics*, 21(20):3940–3941, 2005.