Karlsruher Institut für Technologie

Karlsruhe Reports in Informatics 2013,9
Edited by Karlsruhe Institute of Technology,
Faculty of Informatics
ISSN 2190-4782

Peer-to-Peer Group Affiliation Detection using Mobile Phones

(Technical Report, September 19th 2013)

Dawud Gordon, Michael Beigl, Martin Wirz, Gerhard Tröster, Daniel Roggen

2013

KIT – University of the State of Baden-Wuerttemberg and National Research Center of the Helmholtz Association

Fakultät für **Informatik**

Peer-to-Peer Group Affiliation Detection using Mobile Phones

(Technical Report, September 19$^{\text{th}}$ 2013)

Dawud Gordon, Michael Beigl
Karlsruhe Institute of Technology
Karlsruhe, Germany
{Firstname.Lastname}@kit.edu

Martin Wirz, Gerhard Tröster
ETH Zürich
Zürich, Switzerland
{wirz,troester}@ife.ee.ethz.ch

Daniel Roggen
Newcastle University
Newcastle, England
d.roggen@ncl.ac.uk

***Abstract*—As social creatures we spend most of our time in groups, which are more productive than the individuals in them alone. Methods for recognizing group boundaries and affiliations have been presented using centralized instances to aggregate and evaluated data from mobile devices. However in situations which centralized instances are not available due to network restrictions and scalability, novel methods for detecting affiliation must be found. We present a method for distributed, peer-to-peer (P2P) recognition of group affiliations in multi-group environments, using the divergence of sensor data distributions as an indicator of similarity. The method combines novel approaches for assessing pairwise similarity between individuals, and then interpreting that information in a distributed manner. An experiment was conducted with 10 individuals in different group configurations, and the P2P as well as contemporary centralized approaches are evaluated. The results show that although the output of the proposed method fluctuates, we can still correctly detect 93% of group affiliations by applying a filter. At the same time our method only requires local communication with P2P neighbors, reducing response time by up to 7% or total energy consumed to recognition by up to 43%. We foresee applications in mobile social networking, life logging, smart environments, crowd situations and possibly crowd emergencies.**

I. Introduction and Motivation

Around 70% of the time we spend in public areas is done together with other people [16]. In general we are social creatures and spend a great deal of our time in groups of one form or another [9]. This social behavior is also useful, as it has been shown that groups are better than individuals at accomplishing tasks, which is often why they are formed in the first place [9]. Understanding group behavior and context is then crucial for systems which are trying to assist these groups in some fashion. Before an understanding of the group's context can be reached, the boundaries of groups and individual affiliations must be identified through the precess of group affiliation detection (**GAD**). Often times several groups can occupy the same space at once [16], making it important to detect non-affiliation as well as affiliation.

We humans have an innate ability to visually recognize these groups quickly [16], using unconscious processes which can be described using the Gestalt Laws [9]. Our minds automatically observe and group objects together based on proximity, similarity and interaction: objects which are similar or near each other belong together. It is this perception process of detecting groups and affiliations which GAD to imitate [17]. Since we are trying to imitate human perception, we are therefore bound to that perception as it defines correct and incorrect affiliation decisions. The problem is then to differentiate inter-group similarity from intra-group similarity.

Individuals who are within the same group will behave similarly to each other, e.g. walking in the same direction, at the same speed, or conducting activities at the same time [17]. Through their social interaction they also behave similarly due to norms within the group [9], as well as the "Chameleon Effect" [7], which states that they will subconsciously mimic each other's social cues. Since their physical behavior is similar, the sensors on their mobile devices which monitor physical behavior will also produce similar signals. It therefore follows that by monitoring wearable sensor signals of multiple individuals, it is possible to estimate group affiliation [17], and recognize affiliative behavior.

Such methods aggregate data from distributed sensors and sensing modalities [14] and then analyze the emergent result [17]. However, there are conditions where centralized aggregation cannot be achieved [11]. This can occur in non-instrumented environments where a centralized instance is not reachable, or when the bandwidth is too expensive either in terms cost to the user device or to the infrastructure provider. Finally, under some circumstances access is simply not available, such as during emergencies where infrastructure is usually the first casualty [4]. For these conditions, new methods for evaluating group affiliation using P2P analysis systems must be explored.

By changing the angle of approach from centralized to P2P, the definition of the problem also changes. The point of view which we wish to imitate changes from that of the observer of the emergent behavior to the point of view of the individual in the P2P network. The problem now becomes assessing individual-to-individual affiliation across neighboring individuals and nodes. Complexity moves from the method for clustering groups, to the method for evaluation similarity in a distributed fashion.

We present a novel method for assessing P2P affinity by modeling the data as a distribution and then calculating the disparity (or similarity) as the Jeffrey's divergence between models from different individuals. We call this method divergence-based affiliation detection (**DBAD**). We compare DBAD with centralized and distributed approaches using signal correlation which is the basis for previous approaches [17]. Such approaches require sensor data exchanges between nodes in order to perform time-series analyses. The DBAD approach is sensor-independent, requiring only a sensor which measures personal characteristics which are in some way indicative of inter and intra-group similarity and can be modeled as distributions, e.g. BlueTooth fingerprints, GPS locations, proximity sensors, etc. Furthermore DBAD has the potential to use multiple modalities for a single pair-wise affinity analysis which would solve several existing issues [14], although we do not present this here for brevity. We present 2 methods for accomplishing GAD, one where nodes exchange Gaussian probability density functions (DBAD-P) of sensor data, and another where they exchange histograms of observations (DBAD-H).

We evaluate these methods with an experiment involving 10 individuals with varying group numbers, sizes and affiliations conducting a homogenous activity: a scenario with high difficulty. The resulting data set is also published as part of the contribution of this work (see Sec. IV). We evaluate the methods using two different types of sensor data each with different types of distributions; accelerometers and magenetometers, modeled as normal and Von Mises distributions respectively. The DBAD methods perform significantly worse in terms of identifying inter and intra-group similarities at any given instant with a maximum of 63% compared to a 74% for centralized correlation. However, filtering similarity values over time improves recognition to 93%. The reduced P2P communication range limits the number of inter-group neighbors increasing accuracy to 80% even without filtering.

A centralized correlation approach requires little memory (under 2kB) and energy (13.6 mJ), and the total response time for each node is low as processing is offloaded. Distributed applications of correlation algorithms are however are not viable due to the time and energy required for communicating sensor data. The DBAD-H requires marginally greater memory than the centralized approach (around 2.1kB), and decreases response time and totally energy consumed by 8% and 24% respectively compared to the centralized approach. DBAD-P has a higher response time due to local processing and requires double the memory, but the total energy expenditure is less than both centralized correlation and DBAD-H by 43% and 24% respectively due to reduced communication.

II. Related Work

The behavior of groups or crowds is emergent behavior resulting from individual members' actions, their interactions with each other and the environment, as well as their initial states and predispositions [9]. In animals, individual behavior, interactions and states can be quite simple yet still generate complex group behavior [18], allowing straightforward modeling. For humans, modeling such systems is a very difficult problem due to the complexity and cardinality of variables.

GAD differs from behavioral modeling in that we are not interested in understanding the behavior, but rather in assessing if the behaviors are similar, regardless of the form of the behaviors themselves. Marin-Perianu et al. [15] proposed detecting groups of smart goods in the supply chain using the degree of correlation between the sensor signals. This approach was later applied to human beings, where the correlation of acceleration signal variance was used to identify group affiliation [17]. This was conducted with heterogeneous behavior over groups (e.g. walking, climbing stairs, etc.) and homogeneous behavior between individuals within the group.

This work was expanded to describe the approach for centralized detection of group affiliation [17]. First behavior-relevant information is extracted from sensor signals (signal features) and a correlation analysis is conducted to generate a pair-wise disparity (or similarity) matrix. The relational graph represented by the disparity matrix can then be clustered in order to obtain a fairly accurate group affiliation label for all individuals [17]. For multi-modal sensing systems, the clusters generated from each sensing modality can be fused in order to combine information from both multiple modalities [14]. Here the focus is now on trying to create methods which achieve the same goals, but without requiring centralized instances.

Brdiczka et al. [5] recognized changes in group configurations by calculating the Jeffrey's divergence over histograms of multi-modal sensor data. There, divergences have been shown to indicate changes in the group dynamic based on the emergent image of sensor data. Here we we investigate if these methods can also be an indicator of one to one group affinity. Since probability density functions over human trajectories characterize them well [6], it follows that these models could be useful for detecting similarities in that behavior.

BlueTooth has also been used as a sensing modality to recognize device proximity [8], as have microphone sensors [19]. In both cases fingerprinting methods were used to compare individuals to each other. However the principle is the same. At any given time, similarity metrics can give us an indication of "proximity" between individuals, often corresponding to physical proximity. However it is the similarity in these proximities over time which indicate group affiliation, requiring exchanging time-lines of measurements or features. We present a method to avoid exchanging observations, using model parameters instead, and a method for filtering these similarities over time to create an indicator for group affiliation.

III. Divergence-based Affiliation Detection

The previous work on centralized approaches [17] describes that approach to group affiliation detection as following. Sensor data streams from devices monitoring potential group members are analyzed and behavior-relevant information is extracted, e.g. acceleration variance, as indicators of individual activity cues [17]. A cross-correlation ρ analysis of a given time window of these extracted signals is conducted in a pair-wise fashion, resulting in a disparity matrix $\overline{\mathcal{M}}$ in which index i, j indicates the strength of the correlation between the observational data $\mathcal{D}$ of subject i and subject j over a period of time t.

$$\overline{\mathcal{M}}^t_{ij} = \rho(\mathcal{D}^t_i, \mathcal{D}^t_j) = \frac{\gamma(\mathcal{D}^t_i, \mathcal{D}^t_j)}{\sigma(\mathcal{D}^t_i)\sigma(\mathcal{D}^t_j)} \tag{1}$$

where $\mathcal{D}_i$ represents sensor data from subject i over the time period t, γ is the covariance and σ the variance over the windows. The multi-dimensional similarity graph represented by $\overline{\mathcal{M}}$ can then be clustered using semi-supervised clustering algorithms, resulting in an assignment of group affiliation.

The problem with this approach is that in order to evaluate the similarity between two subjects the data streams from

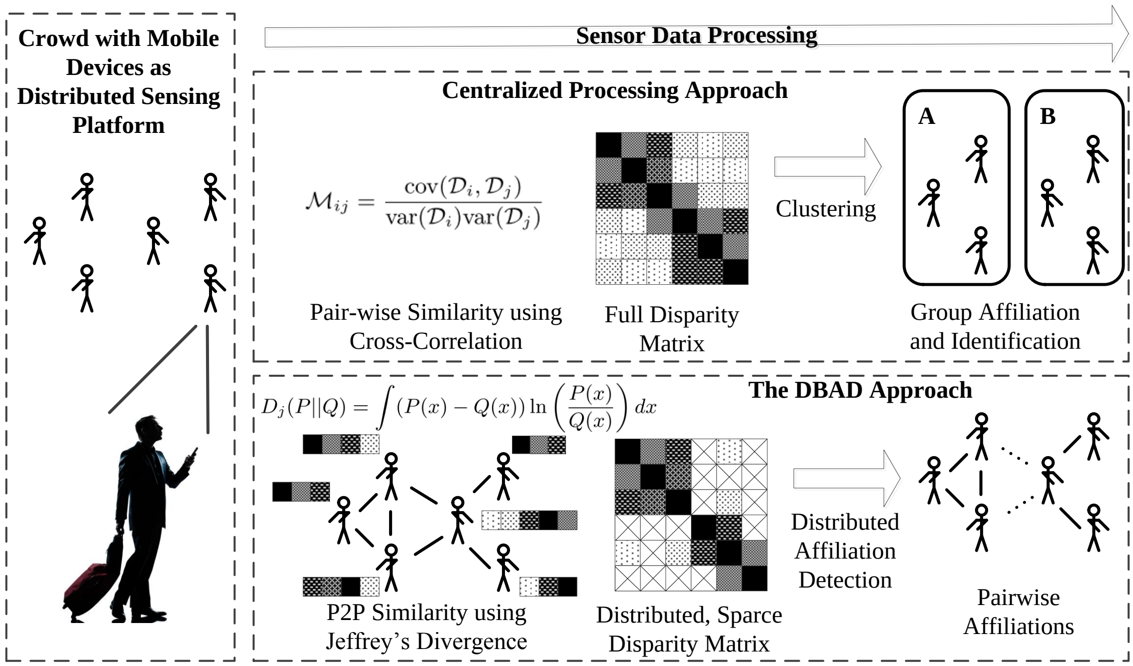

Fig. 1. The centralized and the novel distributed approach to group affiliation detection

both subjects are required. This is due to the cross-correlation algorithm γ in the numerator of Eq. (1), which requires calculation of a function of the point-wise multiplication of both signals. In order to avoid communicating raw sensor data, new methods of analysis which do not rely on time-based signal analysis are required.

We present a model-based approach to this problem called divergence-based affiliation detection (DBAD). The approach works for any sensing modality which delivers similar values for similar inter-individual behavior and can be expressed as a histogram over a window. Theoretically it can also be used to combine several modalities into one similarity measurement although we have not yet evaluated this aspect. Each device computes a model of local data based on the sensor signals it has collected over a specified time window. Here the specific modeling approach taken is to use probability density functions (PDF) for modeling windows of local data. Once these models have been fitted, devices exchange these models with their neighbors. Each device calculates similarity to its neighbors using its own data and the models from its neighbors by calculating the divergence of the PDFs using an extension of the Kullback-Leibler divergence. Based on this information, neighboring devices then collaboratively decide if they are affiliated with each other or not using a distributed clustering algorithm. Depending on the model, this should reduce communication volume and hopefully maintain affiliation detection accuracy with respect to related work. As we will show in Sec. V, this is indeed the case with a few caveats.

The volume of raw data communication is dependent on the sensor sample and re-sample rate and is therefore not fair to directly compare these rates. One approach to reducing this communication load without sacrificing recognition would be to compress the data losslessly before transmission, thereby reducing the amount of information which must be transmitted. We therefore evaluated how lossless compression would affect communication volumes by compressing the data using a two-step differential encoding followed by the DEFLATE[1] algorithm. Initially lossless audio compression algorithms where tried as these outperform DEFLATE for audio data which is of a somewhat similar nature to the sensor data used here, but these performed poorly.

A. Distributed Modeling

Histograms of sensor observations have been shown to be indicative of group behavior, albeit for different purposes [5]. They are also fairly straight-forward to model as probability density functions (PDF) and characterize motion and trajectory data well [6]. Therefore, we chose to use these as a basis for modeling the data locally.

DBAD is then as follows. For each sample window, nodes extract relevant activity cues from the sensors. In this work, we monitored the acceleration and orientation of the subjects, using accelerometers and magnetometers respectively. For the acceleration we used the variance of signal magnitude, indicative of walking speed [17]. For orientation the circular mean of the azimuth, indicative of the direction of walking heading. The circular mean of a vector of angles $\bar{\theta}$ consisting of N angles θ is given by [2]:

$$\mu(\bar{\theta}) = \mathrm{atan2}\left(\frac{\mathrm{imag}(\bar{r})}{\mathrm{real}(\bar{r})}\right), \text{ where } \bar{r} = \frac{1}{N}\sum_{j}^{N} e^{i\theta_j} \qquad (2)$$

These signals build the basis for the comparative analysis of individuals. Based on these signals each node fits a mixture model of distributions to the window, where the type of distribution used is based on the type of sensor used. For the

[1] https://tools.ietf.org/html/rfc1951

acceleration signal, this is then modeled using a mixture of Gaussians. For the orientation sensor, the data is modeled using a mixture of von Mises distributions [3] due to the circular nature of the data [6], given by:

$$\text{vonMises}(\theta|\mu, m) = \frac{1}{2\pi I_0(m)} e^{m\cos(\theta-\mu)} \tag{3}$$

where the circular variance σ is given by $\sigma(\bar{\theta}) = 1 - \bar{r}$ and $I_0(m)$ is a normalization coefficient, given the zeroth-order modified Bessel function of the first kind:

$$I_0(m) = \frac{1}{2\pi} \int_0^{2\pi} e^{m\cos(\theta)} d\theta \tag{4}$$

For both models, the number of components is identified using subtractive clustering, with expectation maximization for parameter fitting [3], [6]. The results is a mixture model consisting of K Gaussian components:

$$P(x) = \sum_{k=1}^{K} \pi_k \mathit{Distr}_k(\mathcal{D}) \tag{5}$$

where the type of distribution $\mathit{Distr}_k(\mathcal{D})$ used depends on the data being modeled, using standard Gaussians $\mathcal{N}(x|\mu_k, \sigma_k)$ for acceleration, or $\text{vonMises}(\theta|\mu_k, m_k)$ for orientation data.

B. Distributed Affinity Analysis

Once these mixture models haven been built, nodes (belonging to individuals) exchange these models with their single-hop neighbors. Each node n_i in the set of all nodes with dimension N can now calculate their disparity to neighboring nodes based on the Jeffrey's divergence of these two nodes. The Jeffrey's divergence is an extension of the Kullback-Leibler divergence, selected because it is numerically stable and symmetric. The Jeffrey's divergence D_j between two distributions P and Q is given by:

$$D_J(P||Q) = \int (P(x) - Q(x)) \ln\left(\frac{P(x)}{Q(x)}\right) dx \tag{6}$$

Each node calculates its pairwise disparity to all other nodes within its single-hop communication neighborhood $\mathcal{V}^t$ at time t. Which nodes are in this neighborhood is dependent on the range of communication ψ (complexities in wireless communication are ignored here), and the physical Euclidean distance between two nodes at the time:

$$\mathcal{V}^t = \{[n_i, n_j]\}|\text{dist}^t(n_i, n_j) \leq \psi \tag{7}$$

The behavioral distance between neighboring nodes can then be acquired as the value of the Jefferey's divergence between distributions of the sensor data of the two nodes.

$$\forall_{[n_i,n_j]\in\mathcal{V}^t} |\overline{\mathcal{M}}_{ij}^t = D_J(\mathit{Dist}(\mathcal{D}_i^t)||\mathit{Dist}(\mathcal{D}_j^t)) \tag{8}$$

In this way the D_J is commutative and both nodes will conclude the same similarity based on the same models. In the centralized approach, the results of the complete pairwise metrics are centrally calculated, yielding a complete similarity matrix for all nodes as shown Fig. 1. Clustering this matrix to find affinity is a relatively straight-forward task, requiring only parameter fitting for clustering thresholds [17]. In a distributed approach this is not the case.

Each mobile device can only communicate with other nodes within reach of local p2p communication, which has 2 important repercussions. First, the similarity matrix is distributed across the complete set of user devices and is not available to any single device. Since the assumption is that global communication is either unavailable, intermittently unavailable, or cannot be used for cost reasons (i.e. Bandwidth), it also implies that this distributed data entity cannot be fully queried by any single device. Second, its distributed nature also means that the disparity matrix is incomplete or sparse, as disparity is not measured between devices which are not within communication range. This presents a challenge of evaluating a distributed, sparse disparity matrix across multiple devices. The result is a distributed, sparse disparity matrix as shown in the bottom portion of Fig. 1, where each row of the disparity matrix is located on a different device, and several positions contain no data (when $[n_i, n_j] \notin \mathcal{V}$). Since individuals are mobile over time, the vacancy of a position in the disparity matrix $\mathcal{M}_{ij}^t$ at time t also varies over time as well.

The output of the distributed similarity analysis may fluctuate from window to window, therefore a moving average of the disparity matrices is used as a low-pass filter to smooth the output. A fifo buffer of length b is taken, where the length of the buffer represents how many disparity matrices where used in the average process. This buffer forms the basis for the low-pass moving average filter, where at any point in time t the smoothed disparity matrix $\widetilde{\mathcal{M}}^t$ is given by point-wise average of the disparity matrices in the buffer:

$$\widetilde{\mathcal{M}}_{ij}^t = \frac{1}{b} \sum_{\tau=0}^{b-1} \overline{\mathcal{M}}_{ij}^{t-\tau} \tag{9}$$

For example, a buffer length of one is the same as using no buffer, where a buffer of length 3 means that two previous matrices as well as the current one are averaged together before clustering.

1) Distributed Threshold-based Clustering: Once distributed disparity has been assessed, nodes must then convert this into affiliation information. A threshold-based approach was followed where each device makes a decision based on locally observed disparity values and a predefined threshold ϕ. For each node n_i, the clustering is conducted using only the information in $\overline{\mathcal{M}}_{ij}^t$ where $[n_i, n_j] \in \mathcal{V}^t$, or the information local to the node at time t. The result is a subset $\mathcal{V}_{\text{affil}_i}^t \in \mathcal{V}$ of nodes which are affiliated withe node n_i at time t, based on their disparity:

$$\mathcal{V}_{\text{affil}_i}^t := [n_i, n_j] \in \mathcal{V}^t | (\overline{\mathcal{M}}_{ij}^t \leq \phi) \tag{10}$$

Where the converse is true for local non-affiliation decision:

$$\mathcal{V}_{\text{non-affil}_i}^t := [n_i, n_j] \in \mathcal{V}^t | (\overline{\mathcal{M}}_{ij}^t > \phi) \tag{11}$$

From a global point of view one can ignore the pairwise neighborhood memberships and cluster the full disparity matrix using the threshold value ϕ, ignoring $\mathcal{V}^t$. However, in real situations, this information is not available, therefore making local decisions based on local information necessary. The optimal value used for ϕ is dependent on the physical activity of the subject, as well as the sensors used to monitor that behavior. For practical purposes, the threshold can be experimentally obtained by maximizing the accuracy.

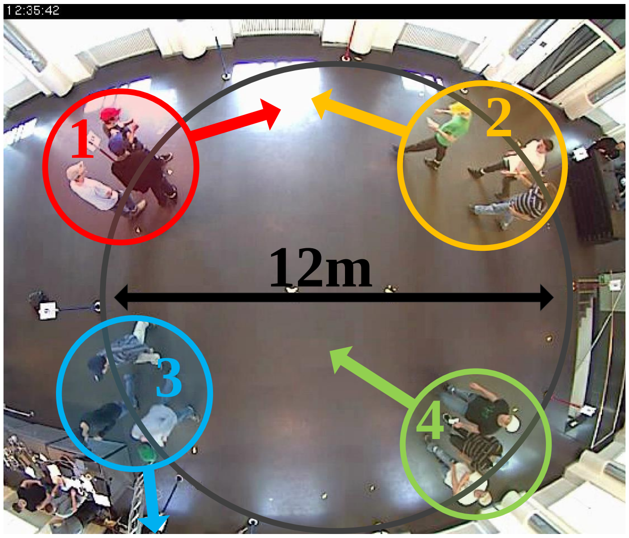

Fig. 2. A still image from the experiment video showing a four group configuration, with annotated group affiliating and heading

Fig. 3. The Android devices used for subject monitoring (left) and the on-body position of the devices (right)

IV. Group Behavior Experiment

Previous experiments with centralized group behavior detection [17] were conducted with groups performing various heterogeneous activities and acceleration sensors. During emergency situations, inter-group behavioral differences may not occur in this fashion. Here, it is quite possible that the activity performed by all participants is homogeneous during crowd and potential emergency situations, such as walking, queuing or fleeing [13]. To evaluate performance under these more difficult conditions, an experiment and data set was created using homogeneous activity behavior, namely walking, of several individuals in different group configurations.

The experiment was conducted in a large open room in a university setting. 12 subjects walked through the room in various group configurations while being monitored by wearable android mobile sensing devices attached to the hip of each subject as shown in Fig. 3. The devices monitored a single subject each using 3D accelerometers and magnetic field sensors, as well as ambient audio. For each subject, the data set contains 51 minutes of data, although 2 devices contained faulty motion sensors, leaving 10 usable subjects.

The experiment was recorded using a wide-angle lens on the ceiling of the room, and each subject was given head-gear of a different color to enable offline individual identification as shown in Fig. 2. 12 labeled posts where set up in a circle with a diameter of 12 meters inside of a large room, where each post displayed a unique number clearly on a sign in clockwise order. A single member from each group was given a list of numbers, and each group then followed that member from post to post in the randomly assigned order on the list. Between experiments, group affiliations where reassigned and the experiment was repeated in the following configurations: one group (all together), 2 groups, 3 groups, 4 groups, no groups (each subject was given a separate list)[2]. Before each group experiment, subjects hopped in unison 3 times which was used to synchronize data by aligning the periods of free-fall (zero acceleration) across subjects.

[2]Data set available at: http://www.teco.edu/~gordon/GAD/data_set.zip

Location data for each subject was annotated after the fact using a mixture of manual and automated color tracking software. For this purpose the video of the experiment was taken and the pixel coordinates of the subject's hat was tracked throughout the experiments. The location is given in pixel coordinates from the top left of the video. We converted these coordinates into meters using the diameter of the circle (12 meters = 430 pixels) as a reference. These coordinates contain the elliptical distortion of the wide-angle lens, but can theoretically be transposed into spacial coordinates using the known dimensions of the room and the location of the camera. We argue that for the purpose of this research, this approximation suffices.

The performance of both centralized and the DBAD algorithms was implemented in MATLAB and then simulated using this data set. The simulation was performed on both the accelerometer and orientation data respectively to evaluate new and previous methods for the emergency situation scenario. For this purpose, the data from the experiments was cut up into windows whose length was varied. The variance of the acceleration data was calculated over a 15 second moving window, as this was shown to be effective for centralized forms of group affiliation detection in other scenarios [17]. For the orientation data, the azimuth was taken around the vertical axis of the subject, and a moving average of one second was used as an indicator of walking direction.

GAD was then performed using the centralized approach based on the signal cross-correlation, as well as the DBAD algorithms. Numerical integration of a PDF is carried out by estimating a histogram of the PDF. In order to evaluate the effect of modeling error on performance, the same process was also conducting using histograms of the individual sample windows constructed using the data windows directly as well. The resulting sparse, distributed similarity matrices where then clustered for both the PDF-based and histogram-based data, and the results where evaluated in terms of correct and incorrect pairwise affiliation detections.

Pairwise affiliations are binary in nature, either indicating affiliation or non-affiliation of two subjects. However for a given group configuration, the distribution of affiliation and non-affiliation is not independent and identical. Given N subjects divided into M groups, with size(m_i) subjects in each group $m_i \in M$, the total number of pairwise comparisons is:

$$X_{\text{total}} = \frac{(N)(N-1)}{2} \quad (12)$$

The number of subject affiliations is given by:

$$X_{\text{affil}} = \sum_{i}^{M} \frac{\text{size}(m_i)(\text{size}(m_i)-1)}{2} \quad (13)$$

The number of non-affiliations is given by:

$$X_{\text{non-affil}} = \prod_{i}^{M} \text{size}(m_i) = X_{\text{total}} - X_{\text{affil}} \quad (14)$$

The accuracy for recognition is then defined as the fraction of pairwise affiliations correctly estimated by the system. A true positive is an affiliation which is judged as an affiliation, as true negative is for non-affiliations. False positive is a non-affiliation judged to be an affiliation, and false negative when affiliated subjects are judged non affiliated. In experiments where there is either only one group, or experiments where everyone acts independently, a row of the confusion matrix is then empty or zero. It is important to note that standard evaluation metrics such as precision, recall and f-measure are then undefined in these cases due to division by zero.

V. Evaluation

Before evaluating communication range, all algorithms where evaluated using a sliding window whose length was varied between 1 to 60s. The results of this simulation are shown in Fig. 4. For all algorithms, results using the acceleration sensor are shown to be independent of window length and remain constant at around 63.5%. This value represents the noise level, meaning the accuracy achieved using random guesses. The reason this is above 50% is the imbalance between the number of affiliative and non-affiliative links as was mention in Sec. IV. The centralized approach performed best of the three algorithms and improves monotonically with the length of the window, achieving just under 74% for a window length of 60 seconds. Using DBAD-H on the histograms and clustering the resulting complete similarity matrices yields an optimum of around 66.2% at 5 seconds. Further increasing the sample window reduces the accuracy of the algorithm, as it asymptotically approaches the noise level at 60 seconds. DBAD-P performs only slightly worse than using a histogram, behaving similarly with an optimum of 66.0% at a window length of 5 seconds which then drops off into noise.

While not necessarily a negative result, the recognition rates achieved are not high enough to be useful in such situations. On further inspection of performance, we identified that individual affiliation values extracted from sample windows where noisy from frame to frame. A simulation was conducted using a low pass filter, in this case a sliding window moving average, in order to assess the informational content provided by the distributed methods. For this purpose the optimum window length for the distributed algorithms of 5 seconds was used, as well as the filter in Eq. (9).

The results of this evaluation are shown in Fig. 5. Here the accuracy using the acceleration sensor remains almost constant showing only slight increases with filtering, indicating that noisy data is not causing the low values. Using orientation data however, it can be clearly seen that the centralized approach as well as both distributed approaches benefit from filtering, eventually all converging at values of around 93.3%. Optimum values are reached after about 250 seconds of monitoring, or a filter of length 50 classifications over 5 second windows.

One of the goals of the proposed methods is to reduce communication volumes, thereby alleviating stress on the network and reducing battery life of the individual devices. We monitored the rate of exchange of data during the course of the simulations for the different algorithms, the results of which can be seen in Fig. 6. The centralized approach requires each node to exchange the entire sample window's worth of sensor data, in this case sampled at 50 Hz. Regardless of window length, 50 Hz of sensor data must be transferred per second, requiring 4 bytes of data per measurement, or 200 B/s.

For smaller window sizes, the compression overhead reduces the advantages of compression (orientation) or even makes it counter-productive (acceleration), where as window size increases the savings become more pronounced, at about 175 B/s for acceleration and 150 B/s for orientation, being able to save around 12.5% and 25% respectively. The distributed algorithms however greatly outperform the centralized approaches. At their optimal window length of 5 seconds, communicating histograms between nodes (in this case 20 buckets) requires only 8 Bytes/second of communication, and communicating models a factor of 10 less. Concretely these are either π, μ and σ values for acceleration data, π, θ and m values for orientation data respectively), as shown in Eq. (5). This is 94.7% and 99.5% reduction when compared even to the centralized approach with lossless data compression for the histogram and model-based distributed methods respectively.

One major difference between the distributed approaches and the centralized approach is the use of P2P communication which has a limited communication range. We evaluated the effect of this by varying the effective communication range of individual nodes using the location information annotated from the video. For a given range, nodes are able to only communicate with other nodes which are within a circle with radius equal to the range.

Fig. 7 shows the accuracy results when the communication range of the devices is limited in simulation. At maximum range all nodes can communicate with each other across all experiments. As the range is decreased, the accuracy of the all methods increases to an optimum at 4.5m of 83.1% for the centralized approach, 79.6% for the histogram-based approach, and 81.2% for the approach using model divergence. Decreasing the communication range further incurs a sharp drop, with accuracy eventually dropping off to noise as the distance approaches 0. The optimum of 4.5m is there length where affiliated links are maximized and non affiliated links are minimized within the neighborhood of each node.

The results are demonstrated in Fig. 8, where similarity

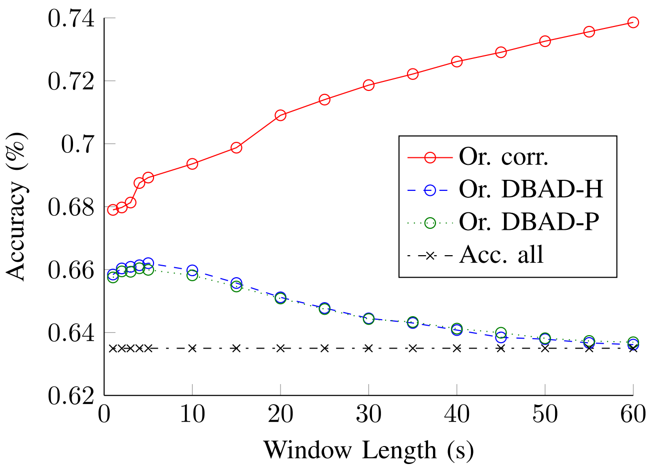

Fig. 4. Performance for cross-correlation (corr.) and DBAD over orientation (or.) and acceleration (acc.) signals

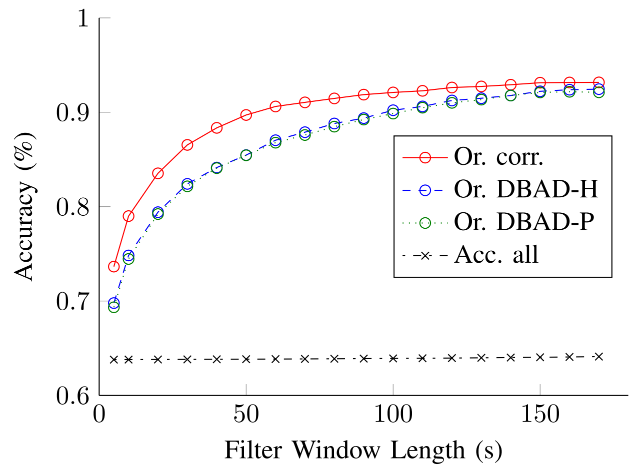

Fig. 5. Performance for cross-correlation (corr.) and DBAD over orientation (or.) and acceleration (acc.) signals filtering over a 5s window

matrices are displayed instead of disparity for visibility reasons. Each row and column are subjects from 1 to 10, and index i, j is the similarity between subject i and j. In Fig. 8a) a typical clustering of a 5 second window by DBAD-P algorithm is shown for two groups. The difference in the similarity between subjects can be seen, but two groups can be identified, one in the upper left and one on the lower right. This also leads to noise in the identification of group affiliation in the same column of Tab. I. In Fig. 8b) both are in different locations but the heading is similar, as is the case with groups 2 and 4 in Fig. 2. This leads to a drop in precision in Tab. I for that window. In Fig. 8c), a communication range of 5m greatly increases precision as most inter-group links are removed, but recall lags, as intra-group similarity fails to correlate group affinity. Filtering over the entire experiment Fig. 8d) improves all values, but errors are still caused by intra-group similarity values. The problem with intra-group similarities is demonstrated by Fig. 2(1), where the heading of the individuals in the group differs dramatically. Note that here we use precision and recall for demonstration purposes, but for experiments where with 1 or no groups, F-score and either one or the other of these metrics is undefined.

Finally, we ran a simulation to compare the local resource footprints of the various approaches. The values presented in Tab.II are simplified approximations, calculated from the bitrate and power consumption of different communication technologies [1][3]. This is modeled on an Android Nexus 4 device where processing occurs on a single core which has a consumption of 0.5W. Detecting affiliation using distributed cross-correlation is impractical due to the high response time and total energy cost of classification. The costs are due to the high communication volumes and consumptions caused by communicating raw sensor data over P2P channels. The centralized approach however has an expensive communicator, but the high bandwidth means low communication times. Processing is also offloaded, therefore processing time is low, and total energy is low as well.

DBAD-H has low processing time because model fitting is avoided, and P2P communication reduces the cost of communication even with the reduced bitrate. The total cost of energy of DBAD-H is therefore 24% lower than for centralized cross-correlation. DBAD-P has a more processing for model fitting and analysis than DBAD-P, and therefore increased response time as well, but the total energy required drops due to reduced communication. Nonetheless, DBAD-P reduces total energy consumption with respect to DBAD-H by a further 24% or by 43% compared to centralized cross-correlation.

[3]http://www.csr.com/sites/default/files/white-papers/comparisons_between_low_power_wireless_technologies.pdf

VI. Discussion

Due to the nature of the problem, subjects who are in the same group generate similar sensor patterns for reasons discussed in Sec. I. However, subjects in different groups may appear to be similar for periods of time, e.g. when both groups walk in the same direction, as is the case with groups 2 and 4 in Fig. 2. By observing subjects for a long enough period (extending window size), the centralized approach can make these temporary phenomena irrelevant as demonstrated in Sec. V. For the distribution-based approaches however, extending the window size reduces effectiveness as the characteristics of the signal disappear into a flat distribution after enough directional changes. This effect is also compounded by a weakness in the distributed methods themselves, as PDFs and histograms both ignore the time component of the signals.

Take for example two individuals who walk in opposite directions for a period of time, then turn around 180 degrees and walk back the way they came for the same period. In this scenario, although the individuals exhibited very different behavior, the distribution over orientation for that period would appear identical. For this same reason, intra-group affiliations are difficult to correctly recognize, as heading varies over members depending on their location. This is a sensor issue, which indicates that the heading feature is not a perfect fit for intra-group affiliation. However, correlation does not use the absolute value of the signal but rather analyzes covariance over time [12]. The distributed method is therefore slightly worse, even with filtering, compared to the correlation approach which

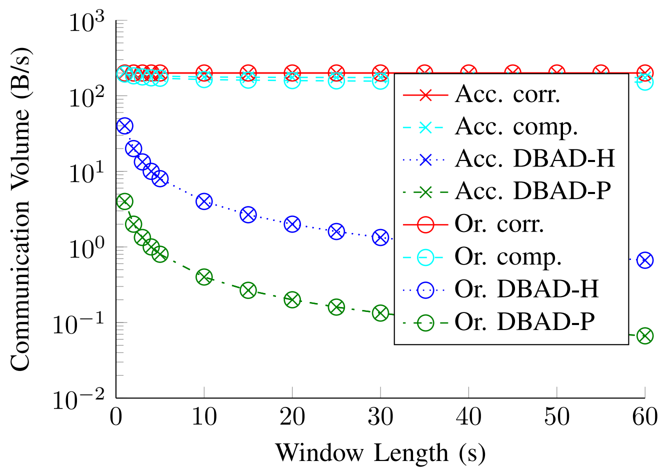

Fig. 6. Communication volumes for cross-correlation (corr.), cross-correlation using compressed values (comp.) and DBAD over orientation (or.) and acceleration (acc.) signals

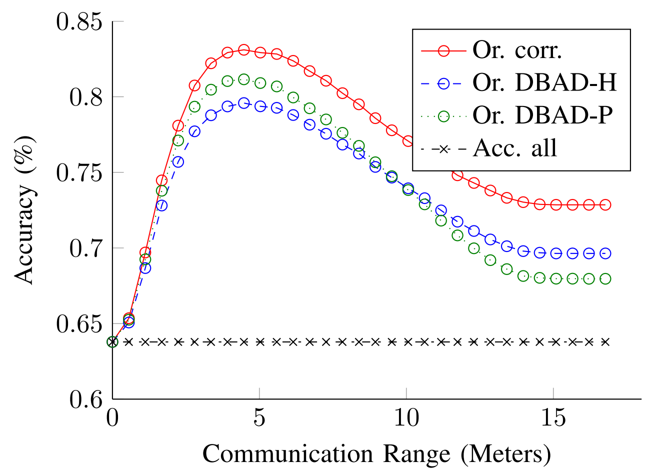

Fig. 7. Performance over communication range for cross-correlation (corr.) and DBAD over orientation (or.) and acceleration (acc.) signals for a 5s window

TABLE I. CONFUSION MATRICES OF AFFILIATION (AF) AND NON-AFFILIATION (NAF) WITH GROUND TRUTH (GT) AND CLASSIFICATION (CL) AND RESULTING METRICS ACCURACY (ACC.), PRECISION (PREC.), RECALL (REC.) AND F-SCORE, CORRESPONDING TO THE DISPARITY MATRICES IN FIG 8.

↓ GT, CL →	AF	NAF
AF	38	12
NAF	6	44

Acc.	Prec.	Rec.	F-score
0.82	0.86	0.76	0.6

↓ GT, CL →	AF	NAF
AF	40	10
NAF	44	6

Acc.	Prec.	Rec.	F-score
0.46	0.48	0.8	0.81

↓ GT, CL →	AF	NAF
AF	38	12
NAF	8	42

Acc.	Prec.	Rec.	F-score
0.8	0.83	0.76	0.79

↓ GT, CL →	AF	NAF
AF	42	8
NAF	0	50

Acc.	Prec.	Rec.	F-score
0.92	1	0.84	0.91

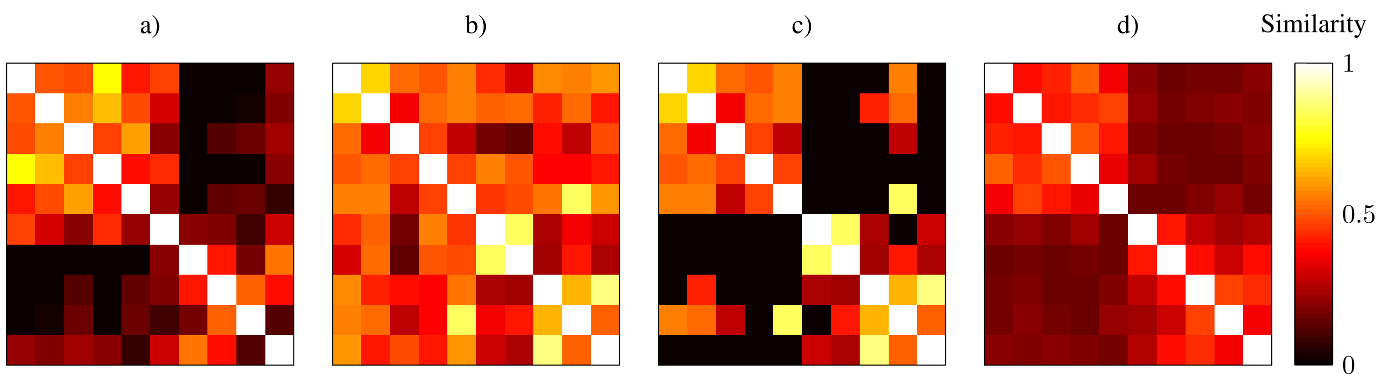

Fig. 8. Similarity between subjects in a two-group experiment for a window size of 5s using orientation and the DBAD-P method. **a)** under normal conditions, **b)** when both groups have similar headings, **c)** when the communication range is 5m and **d)** when averaged over the whole experiment.

is more robust in this respect. The indication is that the P2P DBAD methods are weak against variance from sensors with respect to intra-group affiliation.

The fact that the distribution-based approaches bring with them this inherent weakness also explains why the low-pass filter is so effective. The filter allows the p2p methods to deal with short-term similarity between non-affiliated subjects by extending the observation range for any given affiliation decision. Reducing communication range however can remove these ambiguities entirely, as the members of different two groups are often not compared with each other if they are outside the communication range ψ (again observe groups 2 and 4 in Fig. 2).

One application is for support of social network applications by allowing automated sharing or tag recommendation based on user affiliations. Other applications include life-logging systems which could document who we spent time with. The DBAD approach can also be used to support P2P group activity recognition [10] by allowing group constituents to be identified. The novel algorithms presented here have not been evaluated in large groups or crowds, however the evaluation gives some insight into the uses there. The P2P methods only use neighboring nodes, meaning that the effort required by each device is dependent on the density of the crowd and not the size as a whole. In emergency situations, prevention or management systems must be aware of group affiliations in order to manage groups of individuals as whole. Contradictory instructions to different individuals of the same social group will cause confusion and may be partially or

TABLE II. RESOURCE CONSUMPTION ANALYSIS FOR A CENTRALIZED CORRELATION APPROACH, DISTRIBUTED CORRELATION APPROACHES AND NOVEL DBAD-BASED METHODS FOR 1 CLASSIFICATION OF 5 SECONDS

Approach	Memory Used (B)	Comm. Tech.	Comm. Per Classification (B)	Comm. Time (ms)	Comm. Energy (mJ)	Proc. Time (ms)	Proc. Energy (mJ)	Total Time (ms)	Total Energy (mJ)
Cent. Corr.	2000	3G	1500	29.8	13.53	61.95	0.03	91.75	13.56
Acc. Corr. Comp.	2000	BT 4.0	18000	417.66	191.85	110.4	0.06	528.05	191.91
Or. Corr. Comp.	2000	BT 4.0	21000	487.27	223.83	110.4	0.06	597.66	223.89
DBAD-Hist	2160	BT 4.0	960	22.28	10.23	62.98	0.03	85.26	10.26
DBAD-PDF	4280	BT 4.0	720	16.71	7.67	168.47	0.08	185.17	7.76

fully ignored or disobeyed. Using these methods, management systems could disseminate messages to different individuals, and then allow these messages to disseminate along P2P links classified as intra-group affiliation. Furthermore a combination of in-network similarity assessment and server-side clustering approaches would alleviate bandwidth consumption caused by GAD in crowds while enabling a full emergent group analysis.

VII. CONCLUSION

Humans are social creatures and often build groups for social reasons, and because groups can be better at reaching goals than the individuals separately [9]. However, often several different groups have different goals and occupy the same space, and must therefore be differentiated. Current differentiation methods consist of centrally aggregating sensor information and then clustering the emergent sensor image. However this approach is not feasible when network communication is too expensive, either due to the scale or the environment.

We present a method for distributed, P2P recognition of group affiliations using the divergence of sensor data distributions as an indicator of similarity (DBAD). When addressing the problem from a P2P standpoint, the challenge changes slightly from recognizing group boundaries from the observers point of view, to recognizing subjective affiliations to local neighbors from the point of view of each group member. Divergences can either be calculated using models of individual behavior (DBAD-P) or using histograms of sensor data (DBAD-H). The results show that the output of the proposed method is noisy, with instantaneous recognition rates only slightly over random. However group affiliations can still be detected 93% of the time by applying a low-pass filter to that output signal.

We show that only having a limited range of communication actually improves system performance, by allowing the devices to implicitly use location information without requiring a further sensor. This results in non-affiliative connections being removed from the equation, making the decision or classification easier. Analysis of resource consumption indicates that time-series analysis approaches in the network are infeasible due to time and energy required for communication. However the novel methods reduce the energy footprint of classification by up to 43% and the response time by up to 7%, although memory required by each device increases, without sacrificing affiliation recognition accuracy. However, the P2P algorithms are not as helpful as centralized approaches in some situations, as they do not attempt to recognize group boundaries, but rather local affiliations.

REFERENCES

[1] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani. Energy consumption in mobile phones: a measurement study and implications for network applications. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, IMC '09, pages 280–293, New York, NY, USA, 2009. ACM.

[2] P. Berens. Circstat: A matlab toolbox for circular statistics. *Journal of Statistical Software*, 31(10):1–21, 9 2009.

[3] C. M. Bishop. *Pattern Recognition and Machine Learning*, volume 4 of *Information science and statistics*. Springer, 2006.

[4] A. Boin and A. McConnell. Preparing for Critical Infrastructure Breakdowns: The Limits of Crisis Management and the Need for Resilience. *Journal of Contingencies and Crisis Management*, 15(1):50–59, Mar. 2007.

[5] O. Brdiczka, J. Maisonnasse, P. Reignier, and J. L. Crowley. Detecting small group activities from multimodal observations. *Applied Intelligence*, 30(1):47–57, July 2007.

[6] S. Calderara, A. Prati, and R. Cucchiara. Mixtures of von mises distributions for people trajectory shape analysis. *Circuits and Systems for Video Technology, IEEE Transactions on*, 21(4):457–471, 2011.

[7] T. L. Chartrand and J. a. Bargh. The chameleon effect: the perception-behavior link and social interaction. *Journal of personality and social psychology*, 76(6):893–910, June 1999.

[8] T. M. T. Do and D. Gatica-Perez. Groupus: Smartphone proximity data and human interaction type mining. In *Wearable Computers (ISWC), 2011 15th Annual International Symposium on*, pages 21–28, 2011.

[9] D. Forsyth. *Group Dynamics*. International student edition. Thomson/Wadsworth, 4th edition, 2006.

[10] D. Gordon, J.-H. Hanne, M. Berchtold, A. Shirehjini, and M. Beigl. Towards collaborative group activity recognition using mobile devices. *Mobile Networks and Applications*, pages 1–15, 2012.

[11] D. Gordon, M. Scholz, Y. Ding, and M. Beigl. Global peer-to-peer classification in mobile ad-hoc networks: a requirements analysis. In *Proceedings of the 7th conference on Modeling and using context (CONTEXT)*, pages 108–114. Springer, 2011.

[12] G. R. Gordon and J. Surkis. A general simulation methodology for optical bar code evaluation. *Journal of Statistical Computation and Simulation*, 8(2):85–101, 1978.

[13] D. Helbing and P. Mukerji. Crowd Disasters as Systemic Failures: Analysis of the Love Parade Disaster. *Disasters*, 2012.

[14] M. Kjærgaard and M. Wirz. Detecting pedestrian flocks by fusion of multi-modal sensors in mobile phones. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012.

[15] R. Marin-Perianu, M. Marin-Perianu, P. Havinga, and H. Scholten. Movement-based group awareness with wireless sensor networks. In *Proceedings of the 5th international conference on Pervasive computing*, PERVASIVE'07, pages 298–315, Berlin, Heidelberg, 2007. Springer-Verlag.

[16] M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz. The Walking Behaviour of Pedestrian Social Groups and Its Impact on Crowd Dynamics. *PLoS ONE*, 5(4):7, 2010.

[17] D. Roggen, D. Helbing, G. Tröster, and M. Wirz. Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods. *Networks and Heterogeneous Media*, 6(3):521–544, Aug. 2011.

[18] D. Sumpter. The principles of collective animal behavior. *Philosophical Trans. of the Royal Society*, 2006.

[19] M. Wirz, D. Roggen, and G. Troster. A wearable, ambient sound-based approach for infrastructureless fuzzy proximity estimation. In *Wearable Computers (ISWC), 2010 International Symposium on*, pages 1–4, 2010.