UNSUPERVISED GROUPING OF MOVING OBJECTS BASED ON AGGLOMERATIVE HIERARCHICAL CLUSTERING

Kaori Fujinami
Graduate School of Engineering
Department of Computer and Information Sciences
Tokyo University of Agriculture and Technology
2-24-16 Naka-cho, Koganei, Tokyo, Japan
Emails: fujinami@cc.tuat.ac.jp

Submitted: Aug. 17, 2016 ***Accepted: Oct. 30, 2016*** ***Published: Dec. 1, 2016***

Abstract- In this article, we present a method to identify a grouping of sensor nodes that show similar movement patterns in an ad-hoc manner. The motivation behind the ad-hoc grouping is to allow a system to monitor complex and concrete situations of people and/or devices such as "who is/are utilizing what object(s)" and "what objects are carried together" without any supervision of human before and at the time of interaction. An agglomerative hierarchical clustering algorithm was applied to a data stream to find the group members as a set of clusters within a certain height. A threshold was also determined in an unsupervised way based on simple statistics obtained from the previous clustering results. An off-line analysis was conducted on data collected in realistic situations. Although grouping two of the same but unrelated activities proved to be difficult, the proposed algorithm performed well in other relaxed cases such as walking with a bag vs. pushing a platform hand truck. Furthermore, we confirmed the effectiveness of clustering-based grouping in comparison with simple distance-based grouping.

Index terms*:* Smart Objects, Agglomerative Hierarchical Clustering, Grouping, Accelerometer.

I. INTRODUCTION

Our daily lives and environments are full of computational devices and large amounts of information due to technological advances. We can obtain information anytime at any place. However, at the same time, it takes effort to acquire useful information. Offering appropriate information in particular situations is a major research topic in ubiquitous/pervasive computing environments, so that users can make correct decisions in a user-friendlier system. This is known as *context-awareness* [7]. Any information that is relevant to an application can be a *context*. A context can be obtained not only from a single entity such as person, place, and object [1][2][10][21], but also from an attribute of a group, and grouping can complement a missing piece of information in each entity [11]. For example, suppose that a camera and a train ticket are in a bag. It is likely that the bag belongs to *someone* who is traveling by train, although each object just provides a partial piece of information. The carrier of the bag can further be identified if both the bag and the person's wearable objects, e.g., clothes and accessories, are grouped. In this case, the bag can provide a personalized service: if the bag senses that an unauthorized person is carrying it, it can issue an alert to its owner. Additionally, the person can be suggested to take a camera with her if she does not have one in the bag based on the custom that she usually takes camera with her when she travels. These examples are realized by observing the *coexistence* of objects. For example, a radio frequency identification (RFID) receiver worn by a person or embedded into a bag can be used to detect RFID-tagged objects [5, 19]. However, the similarity of the movement pattern would provide more reliable information than mere coexistence.

In this article, we present a method to identify a group of sensor nodes that move together in an ad-hoc manner. An agglomerative hierarchical clustering [8] is applied to the data stream in an unsupervised manner to specify the candidates for group members. Then, in the *dendrogram* of the clustering result, a branch that contains a key "leaf" is cut off at a particular level of threshold. Here, the threshold is determined based on the statistics obtained from the previous clustering results. Thus, no parameter needs to be specified for a certain object or a usage scenario. Such ad-hoc and real-time grouping is important so that the user can focus on his/her main task. The rest

of this article is organized as follows. In Section II, related work is examined. Section III and Section IV present the proposed algorithm and the evaluation, respectively. Finally, we conclude the article in Section V.

II. RELATED WORK

In the last decade, due to the advancement of microelectromechanical systems (MEMS) and wireless communication technologies, accelerometer-based grouping has become popular [3][4][12][13][15][17][18][22][24]. A correlation coefficient, which is a metric of the proximity of signal patterns in the time domain, is often used in associating (or grouping) one or more sensor nodes [3][4][12][17][22]. Although the correlation coefficient is lightweight, it is sensitive to the delay between signals. If the delay is due to a communication delay, the timestamps by synchronized local clocks on the node sides allow correction. However, as pointed out in [12] and [17], the placement of sensor nodes on structurally loose frames leads to another delay. For example, a delay could occur between a sensor-augmented toothbrush and a wrist-worn sensor because of the wrist joint. Also, boxes stacked on a platform on a hand truck could have propagation delays of the driving force. So, an approach based on the correlation coefficient seems more suitable for explicit shaking in one's hand [3] and utilization under the condition where objects are put on the same rigid frame.

A *coherence function*, the measurement of the extent to which two signals are linearly related at each frequency, has also been used to determine if two devices are carried by the same person [15] and to realize device-to-device authentication [18] based on the shared movement pattern. Since the coherence function deals with the frequency domain, the delay can be ignored. However, it initially requires a periodic movement of objects like those on a person's body who is walking or objects shaken in one's hand.

RFID technology can also make a group of coexistent objects. A person might wear a glove with a receiver such as that in [19]. This is very simple but there is a tradeoff in the range of tag detection. The short-range detection realized by a passive tag reduces the chance of detection of an incorrect object that is not used but just exists in the range. But it also has the advantages of low cost and maintenance-free tags. However, the long detection range realized by an active tag system makes the grouping complex. If there are more than two people wearing tag readers in

close vicinity, it is difficult to identify who is the true carrier since all the readers might detect the same tags. Similar to this approach, radio signal strength can measure the proximity of two devices [5]. The signal strength is easily affected by the orientation of the body and the placement of devices. Efstratiou et al. used packet loss to approximate the distance between two entities [9]. In their application, an accelerometer is attached to a drill and the user is identified as a person who is within 2 meters of the drill. They successfully leveraged the business rule that no person except the user comes close to the drill. However, the assumption might need to be reconsidered when the rule changes.

The approach of Perianu et al. [16][17] enables the formation of a group of sensor nodes based on similar movement patterns by a static threshold in a distributed manner. Here, the similarity is represented by a correlation coefficient. In [16], an algorithm to determine the threshold was introduced. In this algorithm, two distributions of correlation coefficients are required: one for correct pairs and the other for incorrect pairs. The distributions depend on the definitions of the pairs; however, we consider "incorrect pair" to be difficult to define. For example, suppose that a situation is defined as "carried by a person"; then, the complement set is "not carried by the person." This can be "carried by another person who is following the person" (very similar movement), "carried on a platform hand truck," or "merely left on a desk" (no correlation at all). Application requirements should be carefully analyzed to avoid oversights. Rocha et al. proposes a framework called "semantic clustering", in which a cluster of sensor nodes is formed based on semantic correlation [20]. A fuzzy inference system calculates the similarity of a situation observed sensor nodes, in which a domain expert is assumed to supply a set of fuzzy IF-THEN rules. The challenge in our approach is to form a group without such definitions. A group needs to be formed in a self-organized manner. The similarity of features between a reference node and the other neighboring nodes are evaluated based on a threshold that reflects the characteristics of the group in recent time steps.

III. DESIGNING THE AD-HOC GROUPING ALGOTITHM

a. Assumptions

The assumptions in designing the algorithm are as follows.

AS1. Wireless accelerometer nodes are attached on objects and/or people.

AS2. The sensor nodes communicate with each other using a broadcast channel whose range is approximately 1 meter. Therefore, communication delay, packet loss and multi-hopping are not taken into account.

AS3. The number of members in a group is unknown.

AS4. A node called the Group Leader (GL) is a special node that is attached to an object. It initiates the grouping process for nodes that reside within the communication range (neighboring nodes). The GL is also capable of detecting a change of its state, such as "started moving."

AS5. A person carries at least one node on his/her body whose ID is not known by the GL, though.

AS6. Nodes that reside in the vicinity of the GL do not always belong to the same group as the GL, but coincidentally reside near the GL.

A typical scene is illustrated in Figure 1 (a), where three nodes form a group (indicated as the ellipse); the GL (shown in red) is attached to the bag, a normal node embedded on a mobile phone is in a chest pocket, and the other normal node embedded on an umbrella is in the bag. The fifth assumption (AS5) indicates that the GL and a node on a person's body should move together to some extent, since someone must be engaged in the activity when an object is utilized or carried. For the sixth assumption (AS6), a person who is not carrying the abovementioned bag may be walking close to the carrier (within the transmission range of the GL on the bag) (Figure 1 (b)). This means that a set of neighboring nodes of the GL consists of either of the following: 1) only the group members, or 2) a combination of members and non-members. Note that we exclude a node carried by the nearby person from the group formed by the GL on the bag because the node on the second person has no direct relationship with the GL. Moreover, we handle the situation where more than two people carry one container (Figure 1 (c)). In this case, a direct relationship exists between the box and the two people.

Figure 1. Typical Scenes

b. Requirements

Based on the assumptions, the requirements for the algorithm are as follows.

RQ1. The similarity measurement should consider that the movements of the nodes are not always highly correlated.

RQ2. The grouping decision should be made in an unsupervised manner based on the local data.

The first requirement comes from our observations and the cases assumed in AS1 and AS5. The three nodes that appear in Figure 1 (a) might have propagation delays of the driving force. So, the signals might also have a delay or lose their "micro-similarity." Therefore, a measurement that can capture a macroscopic characteristic is needed. In a complex and realistic application setting, it is difficult to collect a dataset and find the parameters that completely cover the characteristics of the *group* and the *non-group*. So, the grouping decision should be made in an unsupervised manner using datasets obtained for a specific number of time steps, which is the second requirement.

c. Grouping Based on an Agglomerative Hierarchical Clustering

Figure 2 shows the flow of a grouping that consists of windowing, feature calculation, clustering, branch cutting, and threshold updating. As can be seen in the top right side of the figure, grouping is conducted for each processing window. Then, a feature vector that meets the first requirement is calculated. Here, we specify two features that characterize the similarity of the movement of nodes for a period of time: *standard deviation* and *mean-crossing number*. The

mean-crossing number is obtained by counting the number of crossings over a mean value in a time window, which represents the frequency-domain characteristic in a lightweight manner. A correlation coefficient has not been selected since, as described in Section II, it is sensitive to the delay between signals and the deformation of a signal. We also tested other features in the time domain (e.g., mean and inter-quartile range) as well as the frequency domain (e.g., frequency range power and frequency entropy [14]); however, the combination of the standard deviation and mean-crossing number performed best.

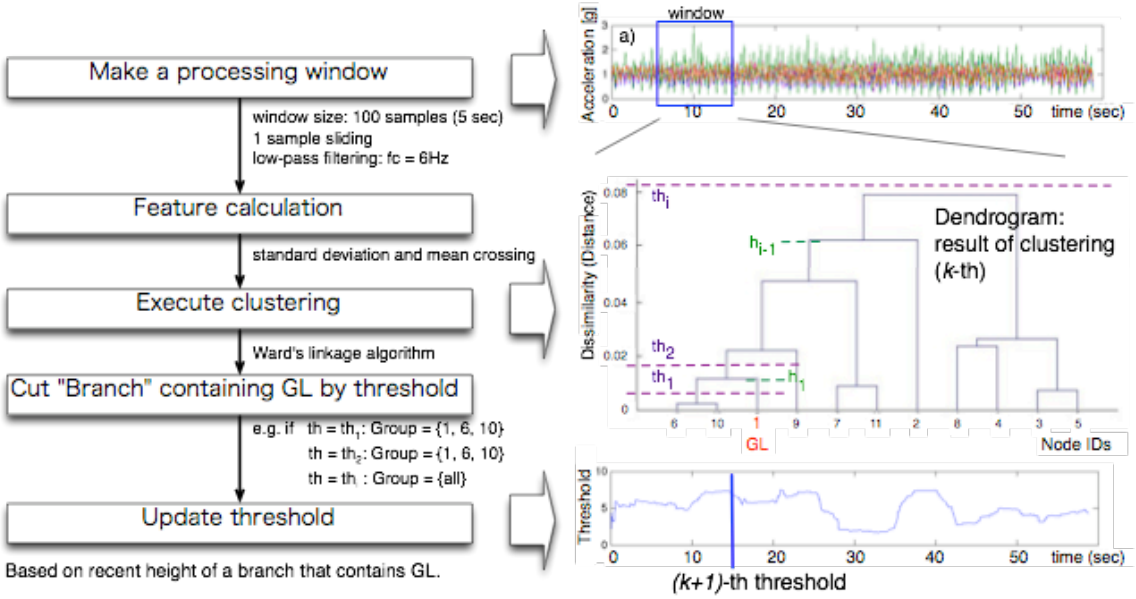

Figure 2. Flow of a grouping

For the second requirement, we applied an unsupervised clustering technique, in which the group of interest is represented as a cluster. In addition to the group of interest, another group may exist that is formed by non-member nodes. The number of such groups is not known in advance. So, we cannot simply apply a method that assumes the number of classes is known, such as discriminant analysis and *k*-means clustering (refer to [8] for an example). We decided to apply Agglomerative Hierarchical Clustering (AHC) (e.g., [8]) to the data stream. As its name suggests, AHC generates a hierarchical cluster tree. The leaf nodes in the cluster hierarchy are the sensor nodes. They are singleton clusters from which all higher clusters are built. The Euclidean distance is used as a dissimilarity measurement for clustering. A particular linkage algorithm generates the tree, and a dissimilarity matrix is given as input. We selected Ward's method [25]

from the various types of linkage algorithms. The method chooses the successive clustering steps to minimize the increase in the *error sum of squares* after fusing two clusters into a single cluster. The distributions of the features of all the cluster members are taken into account. We consider this algorithm to be robust in the case that a feature vector is not actually a member of the group of interest but happens to be near the GL, and the feature vectors of the true members are gathered at a different area in the feature space. If only the distances between the GL and each node were considered, the grouping would fail to reject the non-member node.

The result of AHC is usually represented by a *dendrogram*. The dendrogram expression provides a hierarchical view of the relationship of the nodes. In the middle-right side of Figure 2, an example of a dendrogram is shown, in which 10 sensor nodes exist as the neighboring nodes of id 1 (GL), and id 6 and 10 are the closest pair since the dissimilarity is the smallest (0.0023). To identify the group members in this case, the tree is "cut" with a specific threshold. Multiple branches can be cut, among which the one containing the GL is finally selected as the group of interest, since our goal is to form a group that contains the GL, not all the groups that are found in the neighboring nodes. In Figure 2, the set of nodes {1,6,10} is a group defined by th_2, and {1,6,9,10} forms another group defined by th_3. In the first case, the other neighboring nodes are classified into the *non-member* group.

Note that if there is no branch that contains the GL below the threshold, the GL is combined with the nearest *sub-branch* so that the new branch can be regarded as a group. Suppose that the tree is cut with th_1, the sub-branch {1,6,10} is obtained by {1} and {6,10}. This is explained by AS5, which states that at least one node (on the person's body) moves together with the GL. So, it is natural to combine the separated nodes. However, if th_4 is applied, all the neighboring nodes are classified into the same group.

The threshold is determined based on the statistics obtained from the previous clustering results to meet the second requirement (RQ2). Here, the threshold at the *i*-th time step is defined as follows:

$$th_i = average(H_{i-1}) + \alpha \cdot sd(H_{i-1}) \quad (\alpha \geq 0) \quad (1)$$

$$H_{i-1} = \{h_{i-L}, h_{i-L+1}, h_{i-L} \dots h_{i-1}\} \quad (2)$$

$$h_k = height(GL) \quad (3)$$

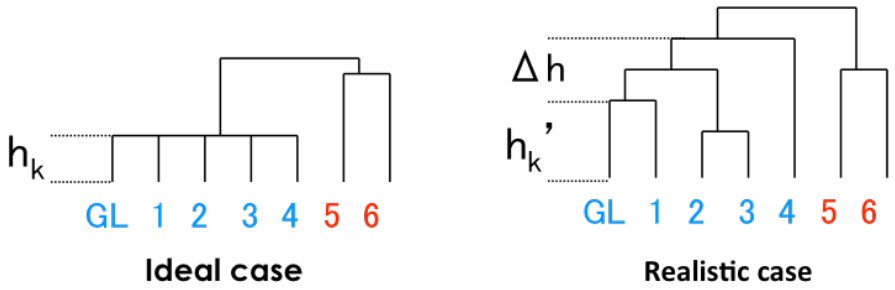

Figure 3. The deviation of the dissimilarity measurement (height), assuming that nodes 1 to 4 belong to the same group as the GL. In an ideal case, all members of a group have the same dissimilarity measurements, whereas, in a realistic case, the measurements vary (Δh) due to the delay and deformation of the waveform.

The first and second terms on the right side of Equation (1) are the average and the standard deviation of the recent L samples of the dissimilarity measurement, respectively. The measurement is represented as the height of the branch that contains the GL as a leaf (Equation (3)). We call the addition to the standard deviation a *margin-scaling factor*, where $\alpha > 0$. An ideal group has no or very small deviation in the height of the branch (Figure 3 left), but in a realistic environment, variance occurs due to the delay and deformation of the waveform (Figure 3 right). Thus, the deviation of the dissimilarity measurement is added to relax the cutting threshold, which means going up the dendrogram and accepting more distant nodes. If th_2 is used as a threshold, then h_1 is obtained as the height of the GL. Once the height is obtained by Equation (2), Equation (1) updates the threshold for the next time step. Thus, the grouping is realized in an ad-hoc and unsupervised manner.

IV. EXPERIMENTS

In this Section, we show a preliminary evaluation of the *grouping performance*, the capability that the algorithm finds the correct grouped nodes among others and eliminates non-related nodes.

a. Experimental Situations of Correct and Confused Grouping

We specified a situation of grouping for the evaluation, where a person is carrying a bag with three items: a mobile phone, a key, and a purse. Here, the GL is attached to the bag. Two wireless accelerometer nodes are attached on his right wrist and belt. The three items are also augmented with the same type of sensor nodes (see Figure 4 (a)). These six nodes form a group. The status

of the grouping prompts an application that notifies an alert to the person with the wrist-worn sensor when the bag detects the absence of a required object, such as a key.

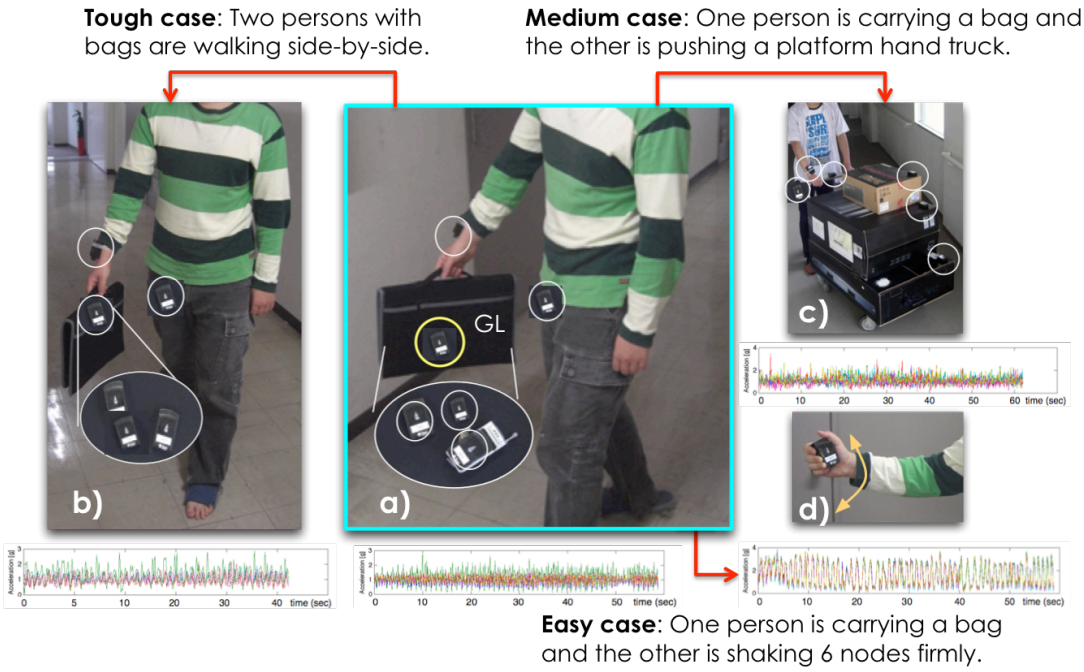

Figure 4. Experimental situations of correct and confused grouping

Three situations were considered as confusing ones for the "correct" group defined above. These cases occur if the nodes are just within the range of communication of the GL (see the description of AS6 in Section III.a). The three cases shown in Figure 4 are as follows: (b) six nodes are in a person's bag, (c) a person is pushing a platform hand truck on which two boxes are stacked, and (d) six nodes are held firmly and shaken. The combinations (a) and (b) are confusing since the activities are the same and thus the features might be very similar. The combinations (a) and (c) are more relaxed, and finally (d) is selected to see the performance of the algorithm between completely different activities.

b. Data Collection and Processing

We programmed a SunSPOT [23] node so that it can acquire three-axis accelerometer readings from an onboard sensor, ST Microsystems LIS3L02AQ, and then we calculated and transmitted the two features at 20 Hz to a data collection PC via an IEEE 802.15.4 compliant radio transceiver. The features are calculated against the magnitude of the acceleration vector

(Equation (4)), because 1) the amount of data transmitted is reduced to 1/3, and 2) the system's interest is not in the orientation of the force. All the experiments in this section ran in MATLAB using real sensor data collected by the nodes fed to the algorithm in an off-line manner. Data were collected for 40 to 60 seconds (800 to 1200 samples) for each situation.

$$A_i = \sqrt{a_x^2 + a_y^2 + a_z^2}$$

The data packets were sent asynchronously from multiple nodes, and timestamps were added on the receiver PC side. After storing all the data in a database, the range of data for all coexistent nodes was extracted based on the timestamp. Ideally, the maximum time difference in two nodes was less than one sampling interval, 50 [msec], since we assumed no communication delay and no packet loss. The size of the window for calculating the features was 200 (=10 sec) based on a preliminary observation. This means the delay was 0.5% of the window size. We assumed that this caused no major problems. To simulate a confused situation where other nodes exist near the truly grouped nodes, one of the other three datasets, either (b), (c), or (d) in Figure 4, was combined with dataset (a). The proposed algorithm attempts to identify the five nodes that belong to the same group as the GL of case (a). For example, the case of (a) with (b) assumed that two people were walking side by side, as shown in Figure 1 (b).

We utilized a sliding window with the overlapping of 199 samples **(**due to sliding windows for every sample.). Before calculating the features, a second-order Butterworth low-pass digital filter [6] was applied to eliminate noise, in which the cut-off frequency was set to 6 [Hz] based on prior observations in the frequency domain. Furthermore, the margin-scaling factor was heuristically set to 5 in the experiment to generate a large difference between the accuracy and the aFPR (defined in the next section). The effect of the margin scale factor is examined in Section IV.d.ii.

c. Grouping Performance Metrics and Visualization

We calculated three association performance metrics: *accuracy*, *average true positive ratio* (aTPR), and *average false positive ratio* (aFPR). The accuracy, defined in Equation (5), represents the capability of identifying entire nodes. If the accuracy is 1.0, the algorithm always identifies all the members of the group. However, even if the accuracy is low, e.g., 0.1, this result does not directly imply that the algorithm is not working as expected. The aTPR also needs to be

considered, since it is the capability of the algorithm to correctly identify on average the number of member nodes (Equation (6)). The performance is considered good if the aTPR is close to 1.0. The aFPR, which indicates the average number of false positive nodes among all the windows (Equation (7)), needs to be close to zero. Here, *false positive* indicates the situation in which a node that is not a member of the group of interest is classified as a member. The basic characteristics of the ratios (aTPR and aFPR) and the grouping capabilities are summarized in Table 1.

$$Accuracy = \frac{the\ number\ of\ correctly\ grouped\ nodes}{total\ number\ of\ data\ windows\ (= N_w)} \tag{5}$$

$$aTPR = \frac{1}{N_w}\sum_{k=1}^{N_w}\frac{the\ number\ of\ member\ nodes\ classified\ as\ member\ (= N_{tp,k})}{the\ number\ of\ the\ member\ nodes\ (= N_m)} \tag{6}$$

$$aFPR = \frac{1}{N_w}\sum_{k=1}^{N_w}\frac{the\ number\ of\ non\ member\ nodes\ classified\ as\ member\ (= N_{fp,k})}{the\ number\ of\ the\ non\ member\ nodes\ (= N_{nm})} \tag{7}$$

Table 1. The basic characteristics of the ratios and the grouping capabilities.

aTPR	**aFPR**	**Characteristics**
Low	High	Clustering failed.
High	Low	Both clustering and thresholding performed well.
High	High	The threshold was too high to exclude unrelated cones
Low	Low	The threshold was too low to include necessary ones.

The result is visualized using 2D images for each combined situation (Figure 5), in which the accuracy, the aTPR, and the aFPR are represented. The color level represents the ratio, which varies smoothly from black (0.0) through shades of red, orange, and yellow, to white (1.0). The numbers of the member and non-member nodes varied in the off-line analysis. The vertical and horizontal sides of the images indicate the number of group members (N_m in Equation (6)) and that of non-members (N_{nm} in Equation (7)), respectively. By multiplying the number of the vertical side of an aTPR image minus 1, the expected number of true positives is obtained. The reason for the subtraction of 1 is to exclude the GL itself. Meanwhile, the expected number of false positives was calculated in the same manner by multiplying the horizontal number in an

aFPR image. The number zero on the horizontal sides of the accuracy and the aTPR images assume that the neighboring nodes fully consist of the member nodes and there is no confusing node. This is the extreme condition that can be realized by the optimized communication range between the GL and the member nodes. Furthermore, since we assumed that there are at least two nodes in a group (see AS5), the vertical sides of the images begin with 2. For example, a pixel of the accuracy image referred to as (3, 3) represents the ratio for the case of "3-out-of-6", in which the GL node is included among the three member nodes. These images allow us to view all the variations of the combined situations, including the number of true cases and the number of false cases.

d. Discussion

d.i Overall Performance of Grouping

As shown in the top row of Figure 5, the combined situation with (b) is difficult to discriminate. The almost black area in the accuracy image indicates that it does not identify all the members. Actually, the accuracy ranges from 0 to 0.005 in this area. Meanwhile, the aTPRs are not as bad as imagined from the accuracy. In Figure 5, the expected number of true positives at the pixel marked as "A" is 2.4 (=$(6\text{-}1)\times 0.48$). This indicates that 2.4 out of 5 true member nodes excluding the GL itself are correctly identified when a total of 10 nodes exist near the GL. This is not bad result considering the similarity of the situations of (a) and (b). However, the aFPRs are higher (i.e., brighter) than the other cases. This indicates that a non-member node is likely to be confused as a true member. In the case of position "B" in Figure 5, the value is 4.0 ($\approx 6\times 0.66$), which means four nodes were mistakenly identified as group members on average from six non-member nodes. From the results in Table 1, we conclude that the clustering did not perform well. This is because the activities of (a) and (b) are the same, and therefore the features are similar. To meet the second requirement RQ2, we utilized common features, i.e., *standard deviation* and *mean-crossing number*, for any object grouping as described in Section III-c. However, there should be a certain set of features that can express slight difference between nodes attached to different persons engaged in the same activity. So, if RQ2 can be relaxed so that the features can be tuned to an object to which the GL is attached, the clustering performance is improved. At the same time, we can consider a hierarchy of groups, where the large group of "walking side by side"

contains sub-groups of "objects carried by a person." The algorithm of such a high-level grouping is future work.

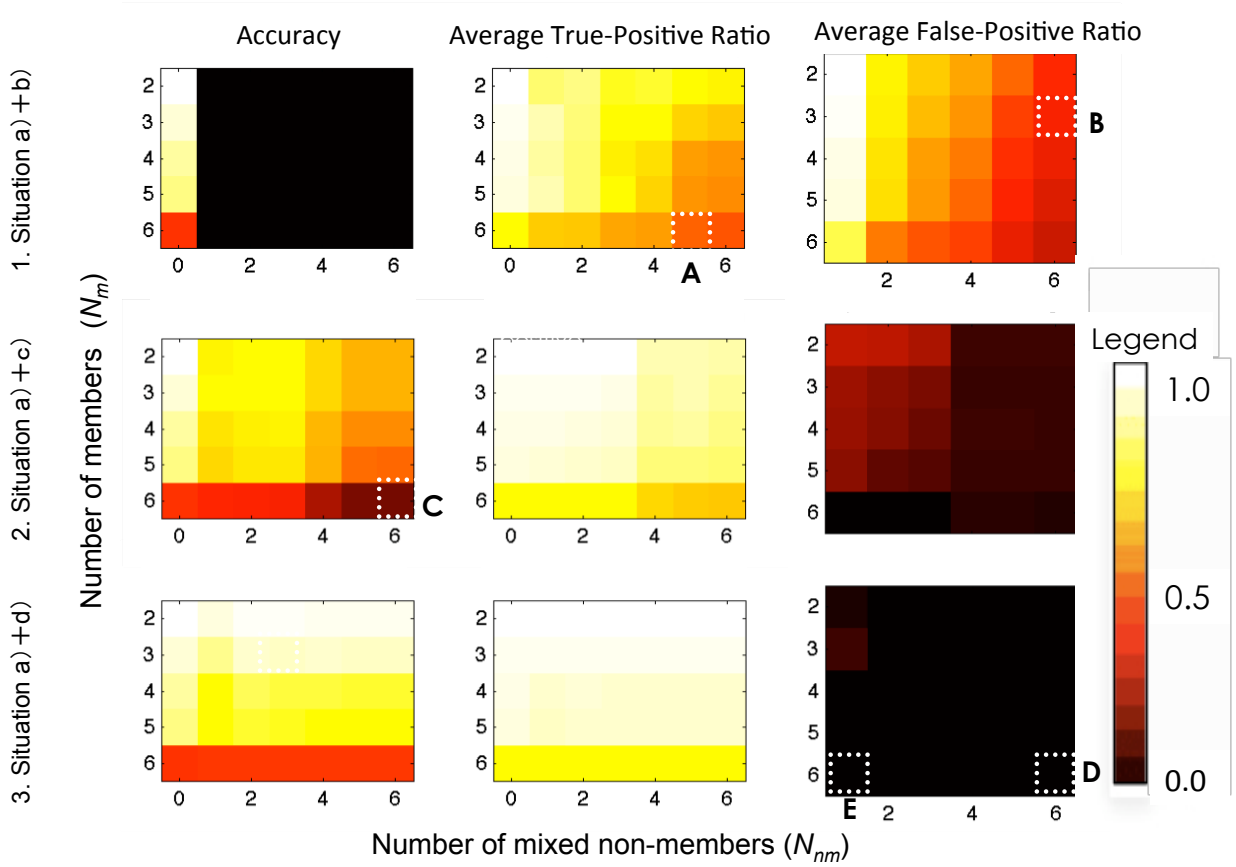

Figure 5. Resultant Images: Accuracy, aTPR and aFPR are represented with 2D images for each combined situation. The vertical and the horizontal sides of an image indicate the number of group members and that of non-members, respectively. The darker the color or level a pixel becomes, the closer the ratio is to zero. Here, the margin-scaling factor (α) is 5.0.

The combined situation with (c) is at an intermediate level. The worst accuracy is 0.14 in the 6-out-of-12 case (marked "C" in Figure 5). However, it is much better than the probability ($_{(12-1)}C_{(6-1)}$ = 0.004) when the true five members (excluding the GL) are randomly selected from 11 (= 5 + 6) nodes. Figure 6 shows the details of the two situations. The first and second graphs are the waveforms of situations (a) and (c), respectively. The third graph indicates the temporal change of aTPR and aFPR, and the bottom graph is the variation of the threshold with time. The aFPR is very small (i.e., almost 0) and is stable. However, the aTPR looks unstable. As shown in Table 1, we consider the cluster of the group was relatively well formed and the non-member nodes are

far from the cluster, but the threshold is not high enough to contain all the member nodes in the cluster. If the threshold is set higher, such missing nodes will appear in the group of interest. Finally, the images of the case with (d) in Figure 5 show the best performance in the three situations: high accuracies, high aTPRs, and low aFPRs. Especially, the aFPRs are remarkably low. The aFPRs range from 0.0011 (six non-member nodes marked "E") to 0.0765 (one non-member node marked "D"), and thus a non-related node is almost not classified as a member of the group of interest. This is because the two activities are quite different from each other.

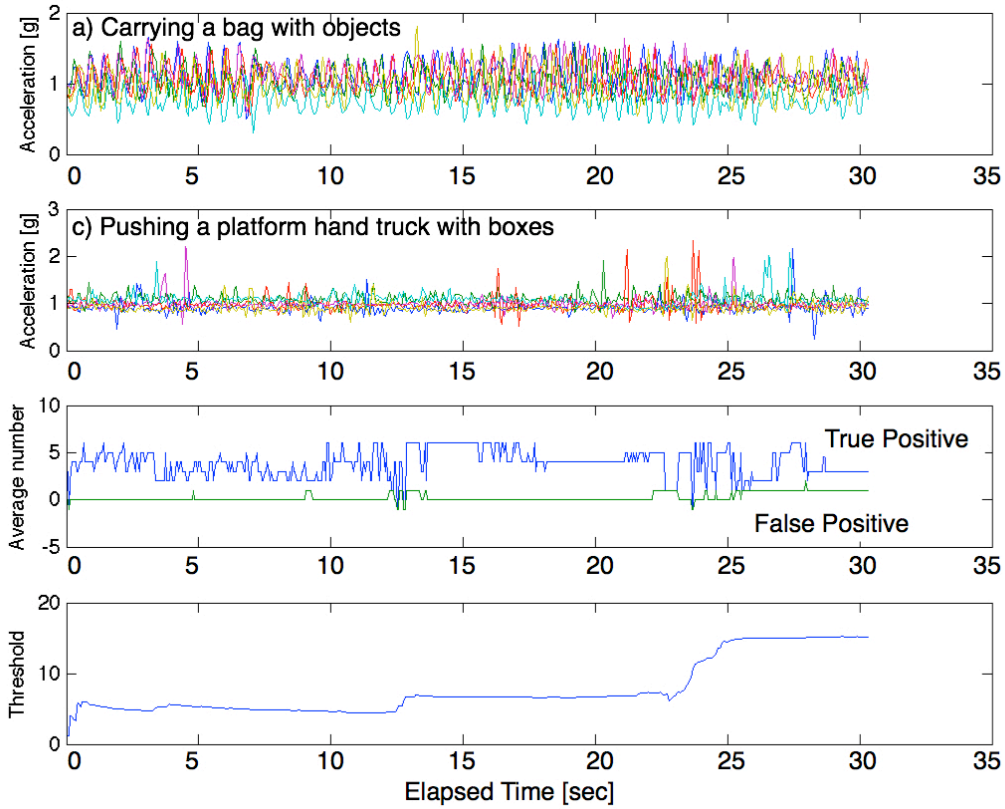

Figure 6. Waveforms of situations (a) and (c), and the temporal changes of aTPR, aFPR, and the thresholds

d.ii Margin and Performance

Figure 7 shows the characteristics of the ratios, including the accuracy, the aTPRs, and the aFPRs, against various margin-scaling factors. The horizontal position represents the margin-scaling factor α. The values were obtained by averaging all the combinations of the number of the true cases and those of the false cases, which means one value is obtained from one image in Figure 5. The left graph of Figure 7 looks different from the other two: the accuracy is saturated at 0.14

without any peak, aTPR is relatively low, and aFPR is much higher. This means that the overlapping of situations (a) and (b) was too large to separate by only increasing the threshold. However, the other two cases show similar patterns; that is, accuracy peaks occur. A peak is considered as an upper bound of the accuracy for a particular combination of situations. The accuracy drastically increases up to the peak as the threshold (or the margin-scaling factor) becomes high, and the number of the non-members remains low. Once the threshold exceeds the peak, the non-member nodes are likely to be contained, and the chance of identifying the entire member is decreased; that is, the accuracy is low. Note that the zero value of α on the horizontal axis indicates that no margin is adopted. By comparing the ratios at α = 0 with α = *peak* (or the saturated value in case 1, (a) + (b)), we can confirm the effectiveness of the margin that contributes to the increase of the accuracy. In future work, we need to investigate an optimal thresholding algorithm that finds the peak accuracy.

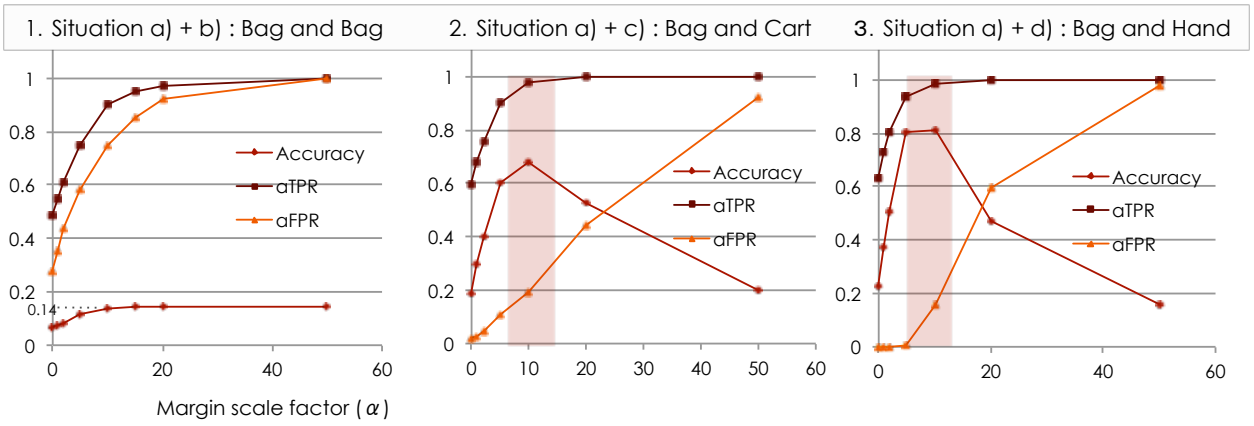

Figure 7. The characteristics of the ratios against various margin-scaling factors for the three combinations of situations.

d.iii Effectiveness of AHC-based Grouping

We applied AHC to obtain the overall distance information among neighboring nodes before branch cutting with a certain threshold. We also consider another approach that seems more straightforward. Here, the distances between the GL and all the other neighboring nodes are calculated, and a threshold is applied to identify the nodes within a specific range. Figure 8 illustrates this approach. Node 3 exists within certain distance (th_i) to the GL, and therefore it is identified as a member. Such a threshold is defined in a similar manner to that in Equation (1);

however, one exception is that the *height* in Equation (3) is replaced with the minimum distance to the GL at each time step (Equation (8)), since according to the fifth assumption (AS5), at least one node should belong to the same group as the GL. We also tested another linkage algorithm for AHC, called *single linkage* [8]. The single linkage (the nearest neighbor) algorithm uses the smallest distance between objects in the two clusters, and so it is simpler than Ward's method.

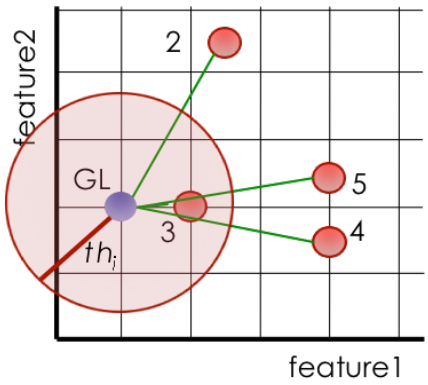

Figure 8. Grouping based on the distance between the GL and neighboring nodes.

$$h'_k = \min_{i \in neighboring\ nodes} (distance\ to\ GL_i) \qquad (8)$$

Figure 9 compares the performance between the AHC and the distance-based grouping in situation (a) + (c). The graphs for the aTPRs show almost the same shapes, but the AHC-based grouping provides better accuracy and better aFPR than the distance-based one. The peaks of accuracy obtained from the AHC-based approaches are higher than the distance-based one, that is, D < AS < AW. Regarding the difference between the peak of the accuracy and the peak of the corresponding aFPR (marked AW, AS, D and AW', AS', D'), an approach is considered superior if the difference, AS − AS', is large, so that the accuracy is high and the aFPR is low. The performance metrics in Figure 9 are 0.48, 0.41, and 0.33 for AHC-Ward, AHC-single linkage and distance-based grouping, respectively. This suggests that the AHC-based linkage algorithms work better than the simple distance-based approaches. We consider that this is because the distributions of features of all the cluster members are taken into account in the AHC.

Moreover, we consider that Ward's method is superior in terms of robustness against changes of the margin-scaling factor. In Figure 9, the accuracy of Ward's method decreases moderately while the aFPR increases gradually after passing the peaks. However, the changes for the single

linkage method appear drastically with the sudden dropping of accuracy and the increasing aFPR. This indicates that the Ward-based approach does not terribly degrade the grouping performance even if the thresholding algorithm fails to determine an optimal value.

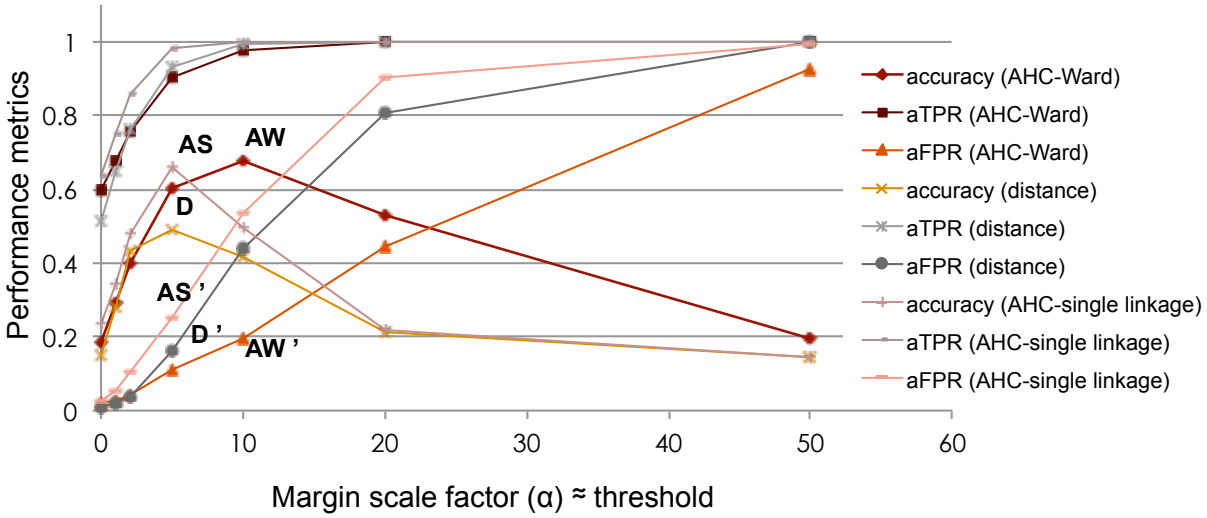

Figure 9. Comparison of AHC and distance-based grouping with various thresholds in situation (a) + (c). Note that AHC-based grouping is further divided into two methods: Ward's method and the single linkage method. The annotations with apostrophes indicate the peaks for the aFPRs, and those without apostrophes correspond to the peaks for the accuracy.

V. CONCLUSION

In this article, we presented a method to identify multiple sensor nodes that have similar movement patterns in an ad-hoc manner. Ward's method and Agglomerative Hierarchical Clustering (AHC) were applied at every time step. The utilization of these two methods was intended to find the true group members while taking into account the spatial distribution of the feature vectors. This is considered to be robust in the case in which a feature vector of a non-member happens to be near the GL, and the feature vectors of the true members are gathering is at a different area in the feature space. With clustering, the unrelated node is likely to be in a different cluster that is far from the clusters of the group of interest. Moreover, the utilization of a standard set of two time-series features, namely, the standard deviation and the mean-crossing

number, makes the grouping of sensor nodes that move in a loosely correlated manner robust compared with a correlation coefficient-based grouping.

We conducted an off-line analysis to evaluate the basic performance of the proposed grouping. In this analysis, a person who has an attached node on the right wrist is carrying a bag of items with attached sensor nodes. Three other situations were combined as false cases existing within the communication range of the GL. Although the grouping in a combined situation of the same but unrelated activities proved to be difficult, the algorithm performed well in the other relaxed cases. We also confirmed the effectiveness of the clustering-based approach, especially Ward's method-based clustering, in comparison with a simple distance-based grouping and the shortest linkage clustering. As future work, the algorithm will be tested in a more dynamic and heterogeneous environment to observe the generality.

REFERENCES

[1] M.R.Addlesee, S. Hodges, J. Newman, P. Steggles, and A. Ward, "Implementing a Sentient Computing System", *IEEE Computer*, *34* (8), (2001), 50-56.

[2] M. Beigl, H.W. Gellersen, and A.Schmidt, "MediaCups: Experience with Design and Use of Computer-Augmented Everyday Objects", Computer Networks*, 35* (4), (2001), 401-409.

[3] D.Bichler, G.Stromberg, M.Huemer, and M.Löw, "Key generation based on acceleration data of shaking processes", Proceedings of the 9th International Conference on Ubiquitous Computing (Ubicomp2007), pp. 304-317, 2007.

[4] S.Bosch, R.Marin-Merianu, P.Havinga, A.Horst, M.Marin-Perianu, and A.Vasilescu, "A study on automatic recognition of object use exploiting motion correlation of wireless sensors", Personal and Ubiquitous Computing, 16 (7), (2012), pp. 875-895.

[5] W.Brunette, C.Hartung, B.Nordstrom, and G.Borriello. "Proximity Interactions between Wireless Sensors and their Application", Proceedings of the Second ACM International Workshop on Wireless Sensor Networks and Applications (WSNA 2003), pp. 30-37, 2003.

[6] J.T. Bryant, H.W. Wevers, and P.J. Lowe, "Methods of data smoothing for instantaneous centre of rotation measurements", Medical and Biological Engineering and Computing, 22(6), (1984), pp.597-602.

[7] A. Dey, “Providing Architectural Support for Building Context-Aware Applications”, PhD Thesis. Georgia Institute of Technology, 2000.

[8] R.O.Duda, P.E.Hart and D.G.Stork, “Pattern Classification - Second Edition –“, John Wiley & Sons, 2001.

[9] C.Efstratiou, N.Davies, G.Kortuem, J.Finney, R.Hooper, and M.Lowton, “Experiences of designing and deploying intelligent sensor nodes to monitor hand-arm vibrations in the field”, Proceedings of the 5th International Conference on Mobile Systems, Applications and Services (MobiSys’07), pp. 127-138, 2007.

[10] K. Fujinami and T. Nakajima, “Sentient Artefact: Acquiring User’s Context Through Daily Objects”, Proceedings of the 2nd International Symposium on Ubiquitous Intelligence and Smart Worlds (UISW2005), pp. 335-344, 2005.

[11] K.Fujinami and S. Pirttikangas, “Kuka: An Architecture for Associating an Augmented Artefact with its User using Wearable Sensors”, Proceedings of IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC2008), pp.154-161, 2008.

[12] K.Fujinami and S.Pirttikangas, “A Study on a Correlation Coefficient to associate an Object with its User”, Proceedings of the 3rd IET International Conference on Intelligent Environment (IE07), pp. 288-295, 2007.

[13] L.E. Holmquist, F.Mattern, B.Schiele, P.Alahuhta, M.Beigl, and H.-W. Gellersen, “Smart-Its Friends: A Technique for Users to Easily Establish Connections between Smart Artefacts”, Lecture Notes in Computer Science 2201 (Ubicomp2001), pp. 116-22, 2001.

[14] K.Kunze, P.Lukowicz, H.Junker, and G.Tröster, “Where am I: Recognizing On-body Positions of Wearable Sensors”, Proceedings of International Workshop on Location and Context-Awareness (LoCA 2005), LNCS 3479, pp. 264-2275, 2005.

[15] J.Lester, B.Hannaford, and G.Borriello, “Are You With Me? - Using Accelerometers to Determine if Two Devices are Carried by the Same Person”, Proceedings of International Conference on Pervasive Computing (Pervasive 2004), pp. 33-50, 2004.

[16] R.Marin-Perianu, C.Lombriser, P.J.M. Havinga, H.Scholten, and G.Tröster, “Tandem: A context-aware method for spontaneous clustering of dynamic wireless sensor nodes”, Proceedings of the 1st International Conference on the Internet of Things (IOT2008), pp. 341-359, 2008.

[17] R.Marin-Perianu, M.Marin-Perianu, P.J.M. Havinga, and H.Scholten, "Movement-based group awareness with wireless sensor networks", Proceedings of the 5th International Conference on Pervasive Computing (Pervasive2007), pp. 298-315, 2007.

[18] R.Mayrhofer and H.Gellersen, "Shake well before use: Intuitive and secure pairing of mobile devices", IEEE Transactions on Mobile Computing, 8(6), (2009), pp.792-806.

[19] M.Philipose, K.P. Fishkin, M.Perkowitz, D.J. Patterson, D.Fox, H.Kautz, and D.Hähnel, "Inferring Activities from Interactions with Objects", IEEE Pervasive Computing, 3:50-57, 2004.

[20] A.R.Rocha, L.Pirmez, F.C.Delicato, E.Lemos, I.Santos, D.G.Gomes, and J.N.Souza, "WSNs clustering based on semantic neighborhood relationships", Computer Networks, 56(5), 2012, pp. 1627-1645.

[21] D.Roggen, N.B. Bharatula, M.Stäger, P.Lukowicz, and G.Tröster, "From Sensors to Miniature Networked SensorButtons*",* Proceedings of the 3rd International Conference on Networked Sensing Systems (INSS'06), pp. 119-122, 2006.

[22] H. Scholten and P. Bakker, "Opportunistic sensing in train safety systems", International Journal on Advances in Networks and Services, Vol. 4, No. 3-4, (2011), pp. 353-362.

[23] Sun SPOT World, `http://sunspotdev.org/` (accessed: May 2016)

[24] H.Yüzugüzel, J. Niemi, S. Kiranyaz, and M. Gabbouj, "ShakeMe: Key Generation from Shared Motion", Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM), pp. 2130-2133, 2015.

[25] J.H. Ward, "Hierarchical groupings to optimize an objective function", Journal of the American Statistical Association, 58(301), pp. 236-244, 1963.