Group Walking Recognition Based on Smartphone Sensors

Qimeng Li[1], Raffaele Gravina[1](✉), Sen Qiu[2], Zhelong Wang[2], Weilin Zang[3], and Ye Li[3]

[1] Department of Informatics, Modeling, Electronics and Systems, University of Calabria, 87036 Rende, Italy
qimeng.li@unical.com, r.gravina@dimes.unical.it

[2] School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China
{qiu,wangzl}@dlut.edu.cn

[3] Key Laboratory for Health Informatics, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
{wl.zang,ye.li}@siat.ac.cn

Abstract. Human group activity represents a potentially valuable contextually relevant source of information, which can be analyzed to support diverse human-centric applications. In recent year, more and more sensors are being pervasively spread in daily living environments, so giving excellent opportunities for using ubiquitous sensing to recognize group activities. In this paper, we used smartphone-based data and edge computing technologies to address group activity recognition, with particular focus on group walking. The data is provided by two groups of participants using a smartphone with embedded 9-DoF inertial sensors; several features are generated to identify group membership of each subject. Our results showed that the accelerometer rarely can be used alone to identify the group motion; in most situations, multiple sensor sources are required to determine group membership. Moreover, the use of 9-DoF sensors to identify group affiliation is still challenging, because, in a multi-user scenario, individual behaviors often have mutual contingency; therefore, the concept of proximity is also introduced to improve the classification algorithm.

Keywords: Multi-user activity · Group membership · Walking group clustering · Wearable sensors

1 Introduction

In recent years, thanks to ubiquitous sensing and Internet of things (IoT), more and more human-centric data has been captured, analyzed, and used to support or improve human's life. Human activity recognition (HAR) [1–3], in this

L. Mucchi et al. (Eds.): BODYNETS 2019, LNICST 297, pp. 91–102, 2019.
https://doi.org/10.1007/978-3-030-34833-5_8

situation, becomes a relevant building block to address several problems in diversified application areas such as public safety, transportation, healthcare, wellness, manufacturing.

HAR, which is distinguished by the number of people involved in a given task, can be roughly divided in individual activity and group (multi-user) activity. From decades, most of studies were focused on recognizing individual activities, while research on group activity detection is much less developed. With the advent of technologies such as cloud computing, many researchers began to use computer-vision [4,5] to identify group activities because still images and video streams are very information-rich signals. Even though the recognition performance can be excellent, computer-vision have non trivial drawbacks and limitation, such as the raising privacy concern, deployment location selection of a fixed camera or battery capacity and energy consumption of a mobile camera. The recent continuous development of microelectronics and Internet of Things (IoT) technology, which have capability to provide massive amount of data from the more and more sensors being pervasively spread in daily living environments, has given an excellent opportunity for using ubiquitous sensing to recognize group activities.

In this paper, we propose a preliminary research on group membership identification in different walking groups. We used a smartphone to collect the 9-DoF (degree of freedom) data and generate features to recognize individual activity; then, all features with recognition results will be sent to the Edge or the Cloud layer to determine the walking group membership.

The remainder of the paper is organized as follows. Section 2 discusses some the state-of-art works on group activity recognition. Section 3 describes the proposed group activity detection architecture. Section 4 reveals the building blocks of the programming model for group activity detection. Finally, Sect. 5 concludes the paper and outlines planned future work.

2 Related Work

There is an established literature on group activities, involving group identification, group membership affiliation selection, group activity recognition. Most studies focuses specifically on the recognition of specific group activities. In this particular case, most researches focused on detecting individual activities/behaviors from each group member, and used such meta-knowledge to recognize the group activities.

Computer-vision approaches have been often applied. Ibrahim et al. [3] adopted a long short-term memory (LSTM) algorithm to build up a deep model for recognize group activity recognition. In their model, 2-stage LSTM (i.e., person dynamics and group dynamics) was used and experimental results demonstrate that the model has a good performance for group activity recognition. Deng et al. [4] proposed a framework which combined graphical models with deep neural networks, and their results showed the model could handle highly structured learning tasks on group activity recognition.

Another recent emerging approach, called *channel state information*, proposes the analysis of Wi-Fi signals to recognize human activities. Feng et al. [5]proposed a scheme termed Multiple Activity Identification System (MAIS) to identify multiple activities of different subjects in a group. However, all of the approaches above could not properly address requirements such as privacy, low portability, low power consumption, and mobility.

Thanks to the technical improvements of microelectronics, the wearable sensors are becoming more lightweight, less energy demanding, and often embedded in daily life wearable accessories (glasses, watches, rings) and garments (t-shirts, shoes, gloves). Abkenar et al. [6] proposed a modeling language (i.e., GroupSense-L) and a distributed middleware (i.e., GroupSense) for mobile group activity recognition (GAR). Experimental analysis showed excellent results in recognizing group activities such as playing table tennis, eating together, and picking cherry in an orchard. Gordon et al. [7] presented a method using mobile phone sensor data to recognize group affiliations in multi-group environments. The experimental results showed the approach can correctly detect 93% of group affiliations. Bourbia et al. [8] presented a generic framework, which is based on the concept of *patterns* for mining interval-based relationships between users' temporally overlapped actions, for group activity recognition using simple non-obtrusive sensors. Yu et al. [9] also proposed a framework with a two-stage process (i.e., sensing modality selection and multimodal clustering) to identify subgroups in a homogeneous group activity using smartphone embedded sensors.

In contrast with previous literature, with the aim of accurately recognizing group activities and group membership, we propose a framework that uses smartphone sensors to identify walking group membership and the group activities which generated by different groups. In this framework, we take into account additional information (i.e. proximity) to support the group activity recognition step.

3 Methodology

In this paper, we propose a framework (see in Fig. 1) to identify walking groups. It is composed of three different layers:

- *Sensing layer* - to collect and analyze all sensor data for further recognition;
- *Edge layer* - if the area is covered with Edge devices, this layer will receive the generated features and classified individual activities to recognize the group activity and identify group membership;
- *Cloud layer* - if the area is not covered by edge devices, the smartphone will connect to Cloud backend which will support the group activity recognition.

Figure 2 depicts the processing workflow of the proposed method. It is composed of three main tasks:

- *Data collection* - to acquire and store raw sensor data;
- *Pre-processing* - the pre-processing step reduces noise from the raw signals;

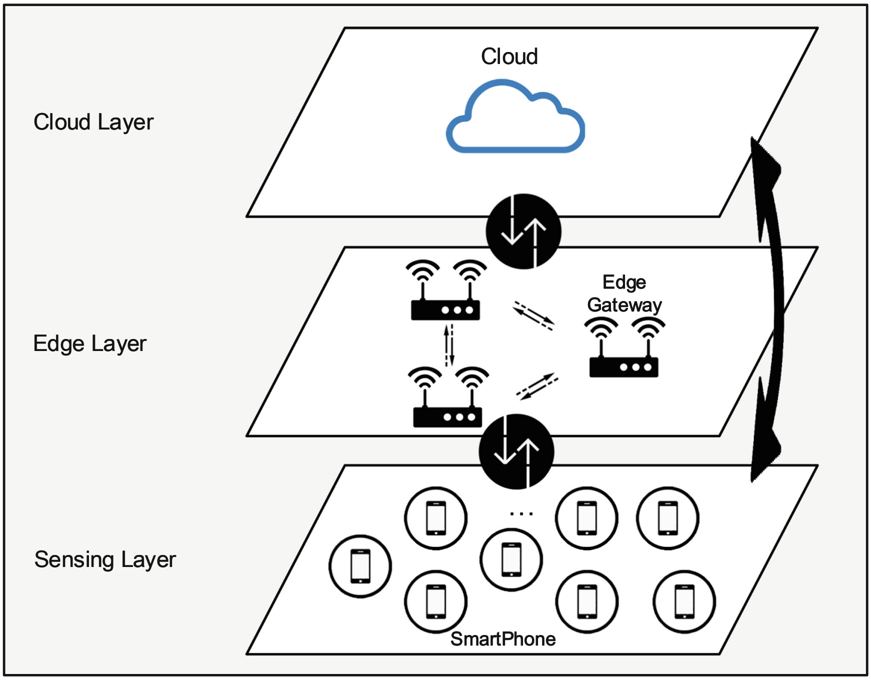

Fig. 1. Framework of the proposed method.

- *Data segmentation and feature generation* - data is segmented into fixed length windows; the segmented data will generate the feature sets;
- *Individual activity classification* - feature sets are sent to a local classifier running on the smartphone and results will be sent to the Edge (or Cloud) layer for group activity recognition;
- *Walking group recognition* - the Edge (or Cloud) layer executes a group recognition classifier that considers only feature sets from potential groups (i.e. cluster of users that are in mutual proximity).

3.1 Data Collection

Data is directly generated by embedded 9-DoF inertial sensor. Collected data is stored in a file which is structured into CSV format. The Edge and Cloud layers will calculate the distance from each user's smartphone.

3.2 Pre-processing

To reduce the noise present in the raw signals, a Finite Impulse Response (FIR) filter is adopted to inertial sensor streams to smooth the data. After de-noising, the raw smartphone sensor (accelerometer, gyroscope, magnetometer) data is normalized in the range $[0, 1]$.

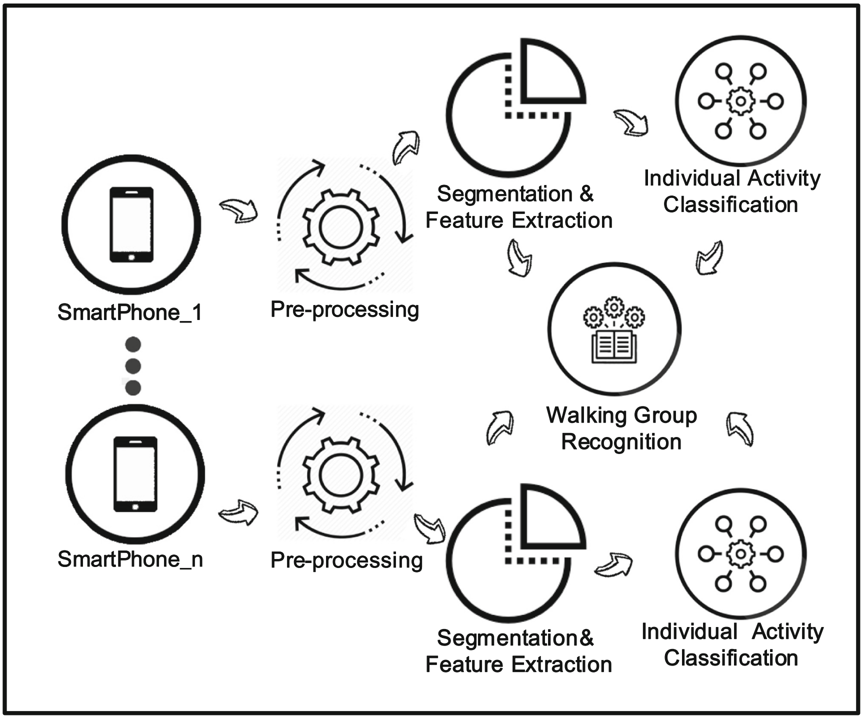

Fig. 2. Processing workflow of the proposed method.

3.3 Data Segmentation and Feature Generation

For the accelerometer data, we fixed the time length for the sliding windows as $l = 1\,\mathrm{s}$. Similarly, the gyroscope and magnetometer data can be divided into windows corresponding to the segmentation of acceleration data.

Feature extraction and selection are crucial in a machine learning process; in this paper, we extracted the following features: direction, signal magnitude vector (SMV), mean, root mean square (RMS), standard derivation (STD), and proximity. The most important features are defined and described in the next subsections.

Direction. Magnetometer and acceleration signals can be fused to estimate the angle between Magnetic North and the user heading. Equation 1 and Eq. 2 calculate respectively the pitch angle θ and the roll angle γ of the user. The g, g_x, g_y, g_z represent the components on three axes of the gravity acceleration vector.

$$\theta = \arcsin\left(-\frac{g_x}{g}\right) \tag{1}$$

$$\gamma = \arctan\left(\frac{g_y}{g_z}\right) \tag{2}$$

Then, the heading angle ψ (with respect to magnetic north) can be calculate using Eq. 3, which $m_x, m_y, and m_z$ represent the three axes elements of the magnetic field.

$$\psi = \arctan\left(-\frac{m_y \cos\gamma - m_z \sin\gamma)}{m_x \cos\theta + m_y \sin\theta \sin\gamma + m_z \sin\theta \cos\gamma}\right) \tag{3}$$

Signal Magnitude Vector. Signal magnitude vector provides a measure of the degree of movement intensity; it can be calculated from the tri-axial acceleration values using Eq. 4.

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{4}$$

Proximity. In our scenario, the concept of proximity represents a feature to measure the distance among different users and between the user and the edge device. In the framework, we consider two different situations:

- *in-door* - if the area is covered with edge devices, the Edge layer will calculate the distance from each surrounding user's smartphone; this allows to construct and keep updated a (user-to-user) proximity map;
- *out-door* - if the area is not covered by edge devices, the GPS module is activated and all the data (feature sets and individual activity class) transmitted to the Cloud layer will be geo-localized. In this way the Cloud layer is able to construct and keep updated a (user-to-user) proximity map similar to the one constructed by the Edge layer.

3.4 Individual Activity Classification

In this step, a previously proposed HMM-based algorithm [10] is used to classify individual user activity. In this work, we specifically focus on two activities: walking and standing. In our HMM, a continuous activity can be represented by a finite number of states. Each state consists of transition probabilities to other states as well as observation probabilities of all the discrete symbols from every state. In this model, each kind of activity represents a trained HMM. Thus, when we obtain a discrete observations sequence $\{O_1, O_2, \ldots, O_n\}$, the appropriate HMM can be found through the maximum likelihood and the recognition had been carried out by the Eq. 5.

$$c = \arg\max_{\mu} P(O_{1:n}, S_{1:n} \mid \lambda_\mu) P(\lambda_\mu) \tag{5}$$

3.5 Walking Group Recognition

When individual activities are classified, the classification results and previous generated features will be sent to an Edge node or Cloud (see Fig. 1) to recognize the group activities. First of all, Algorithm 1 will calculate the size of the group and make a list of possible group members.

Algorithm 1. Member list for classification.

Input: Proximity (P_n) The sets of previous classification result for current batch, $S = \{S_1, S_2, \ldots, S_n\}$; Previous group member list, L_{t-1}; Previous potential member list, R_{t-1};
Output: Group member list, L_t; Removed member list, R_t;
initial list: Calculate proximity and make member list;
function UPDATE GROUP MEMBER LIST(L_t, R_t)
    Calculate the number of the no walking state occurrences (N) for group member list, in period t
    if $N > Threshold$ **then**
        Remove the User for the group member list, add to potential member list;
    else
        Keep;
    end if
end function
function UPDATE POTENTIAL MEMBER LIST(L_t, R_t)
    Calculate the number of the no walking state occurrences (N) for potential member list, in period t
    if $N < Threshold$ **then**
        Remove the User for the potential member list, add to group member list;
    else
        Keep;
    end if
end function
return Group member list, L_t; potential member list, R_t;

In this step, we used a k-NN algorithm to determine the group. Assuming that the sample set $D = \{x_1, x_2, \ldots, x_n\}$ contains n unlabeled samples, each of $x_i = \{x_{i1}, x_{i2}, \ldots, x_{im}\}$ is a m-dimension feature vector (including direction, signal magnitude vector (SMV), mean, etc.); the clustering algorithm divides the sample set D into k disjoint clusters $\{C_l | l = 1, 2, \ldots, k\}$. The output recognition result can be described as in Eq. 6.

$$\lambda = \lambda_1, \lambda 2, \ldots, \lambda_n \tag{6}$$

4 Experiment and Results

4.1 Experiment Description

The data has been collected during an experiment carried out in a previous research [11]. In this experiment, ten volunteers used the same smartphone model (i.e., Samsung Galaxy sIII) to collect data in a real-world environment. Each participant held the mobile phone in the hand during walking sessions. Along other information, accelerometer, gyroscope, magnetometer data were also collected although not used nor further analyzed in that work. The ten participants were

Algorithm 2. Walking group classification.

Input: The sets of features and previous classification result for current batch, $D = \{x_1, x_2, \ldots, x_n\}$; Cluster number, k;
Output: Cluster division on the current batch, C;

initial vector: randomly select k samples from set D: $\{\mu_1, \mu_2, \ldots, \mu_k\}$;
repeat
    Let $C_i = \emptyset (1 \leq i \leq k)$;
    for $j = 1, 2, \ldots, n$ **do**
        Calculate distance: $d_{ji} = \|x_j - \mu_i\|_2$;
        Determine cluster tag: $C_{\lambda_j} = C_{\lambda_j} \cup \{x_j\}, \lambda_j \in \{1, 2, \ldots, k\}$;
    end for
    for $i = 1, 2, \ldots, k$ **do**
        Calculate new mean vector: $\mu_i^{'} = \frac{1}{|C_i|} \sum_{x \in C_i} x$;
        if $\mu_i^{'} \neq \mu_i$ **then**
            Update the current mean vector μ_i to $\mu_i^{'}$;
        else
            Keep the current mean vector μ_i;
        end if
    end for
until No updates;
return Cluster division, C;

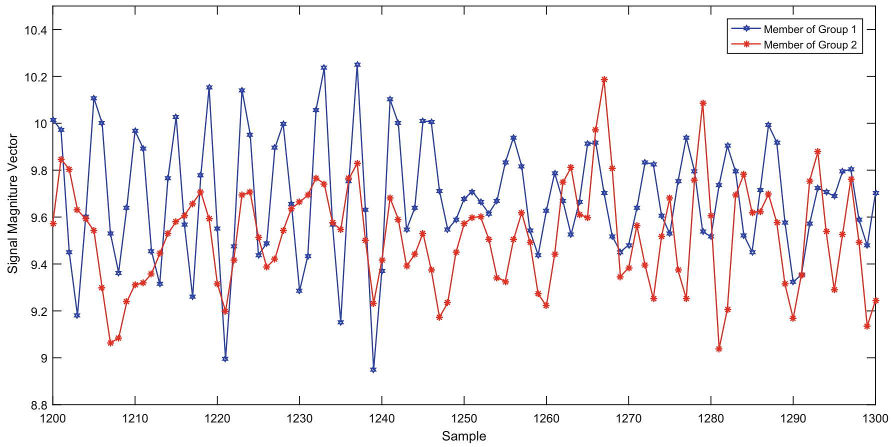

Fig. 3. Filtered acceleration data from different groups.

equally divided into two groups. The protocol of the experiment consisted of the two groups crossing each other, walking in opposite but parallel directions. The two groups were staggered for a short period, then they re-aggregated and continue walking forward altogether.

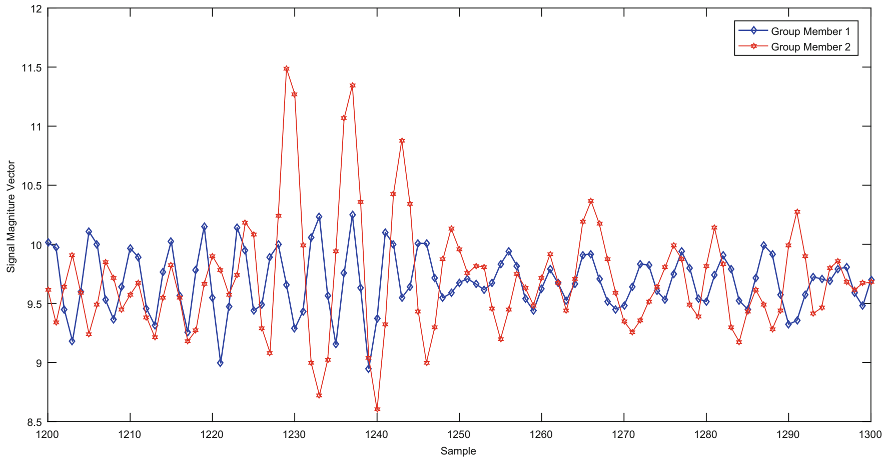

Fig. 4. Filtered acceleration data from same group.

4.2 Heading Direction Detection

In Fig. 5, the red and black lines indicate that there are two participants standing back to back, and the angle between their smartphones is almost 180°. The blue line represents the walking participant direction, and it is nearly the same direction of the standing participant which is marked in red line. This means that the walking participant is facing magnetic north. Obviously, when the participant changes direction, the indication of the magnetic field will change accordingly.

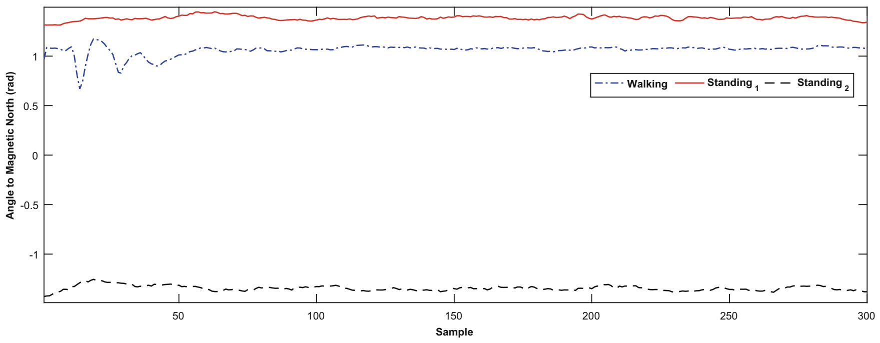

Fig. 5. Group member direction detection. (Color figure online)

4.3 Acceleration Data Analysis

As shown in Fig. 3, users from the same group often accelerated or decelerated in the same pattern; and in Fig. 4, the pattern in some areas is different, due to the

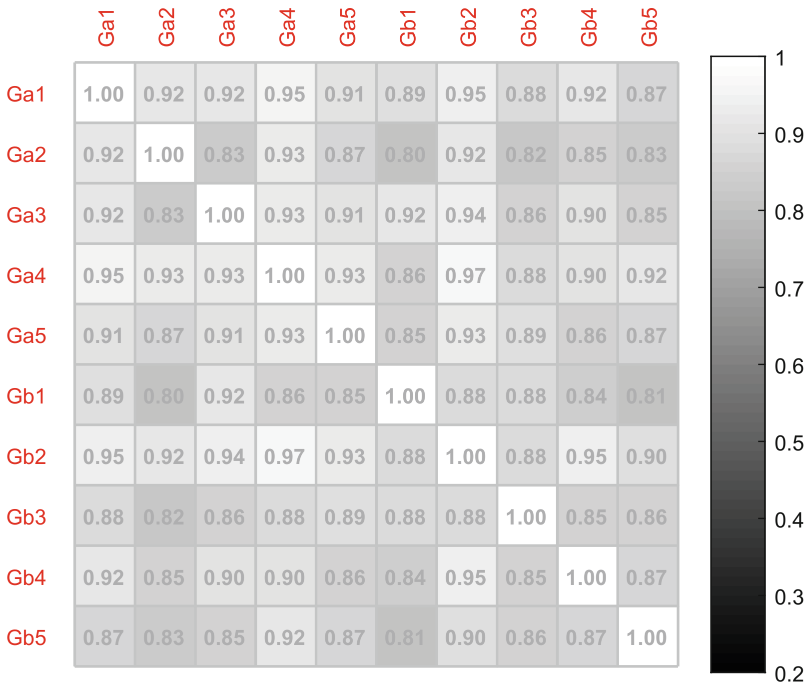

Fig. 6. Cosine similarity between each member.

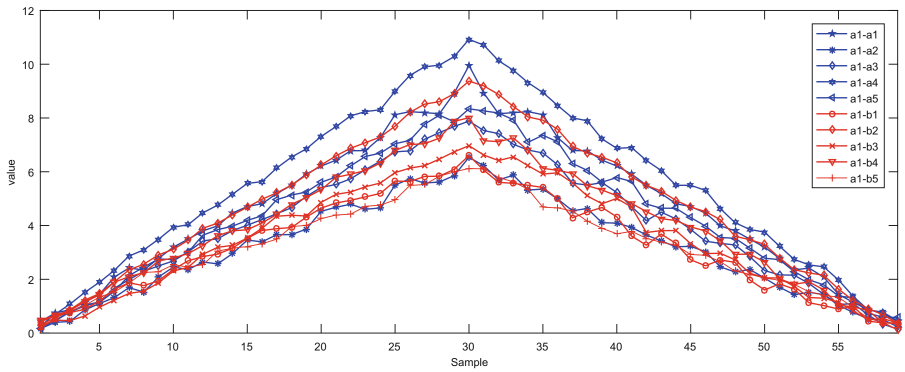

Fig. 7. Cross-correlation of each member (30 samples).

participants are in a different accelerative situation. Still, all of the above can not indicate that users have a similar pattern just from the same group; their motion pattern also could have a similarity between different groups. Human motion patterns are usually random, and different groups of people may have the same motion pattern at the same time. Besides, human activities are usually very subjective, and individual consciousness may lead to inconsistent actions at the same time, and it will result in different motion patterns for people in the same group.

Figures 6 and 7 show the cosine similarity and the cross-correlation between each member; these indicators show that there is no very strong relation by the acceleration changes between each member, even if the participants are in similar patten. Thus, the acceleration data only can be used in recognize the motion state (e.g., waling, standing) of the participant.

5 Conclusion

In this paper, we proposed a framework for the recolonization of group activity and walking group membership. The work is based on mobile, edge, and cloud computing which are selected to support the recognition in different scenarios. Data is generated by several smartphones from each participant, and it is used to identify individual activities. Then, the classified data and features are send to a Raspberry PI edge node or cloud to mining the group activity and the membership.

Results showed that the accelerometer can be used to recognize individual activities such as walking and standing but it has important limitations in mining the pattern of group activities. Therefore, we are planning to apply information fusion methods to combine accelerometer data with heading direction and proximity indicators, so to support a more accurately recognition of the group membership.

References

1. Gravina, R., Li, Q.: Emotion-relevant activity recognition based on smart cushion using multi-sensor fusion. Inform. Fusion **48**, 1–10 (2018)
2. Lara, O.D., Labrador, M.A.: A survey on human activity recognition using wearable sensors. IEEE Commun. Surv. Tutor. **15**(3), 1192–1209 (2013)
3. Ibrahim, M.S., Muralidharan, S., Deng, Z., Vahdat, A., Mori, G.: A hierarchical deep temporal model for group activity recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1971–1980 (2016)
4. Deng, Z., Vahdat, A., Hu, H., Mori, G.: Structure inference machines: Recurrent neural networks for analyzing relations in group activity recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4772–4781 (2016)
5. Feng, C., Arshad, S., Liu, Y.: MAIS: multiple activity identification system using channel state information of WiFi signals. In: Ma, L., Khreishah, A., Zhang, Y., Yan, M. (eds.) WASA 2017. LNCS, vol. 10251, pp. 419–432. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60033-8_37
6. Abkenar, A.B., Loke, S.W., Zaslavsky, A., Rahayu, W.: GroupSense: recognizing and understanding group physical activities using multi-device embedded sensing. ACM Trans. Embed. Comput. Syst. (TECS) **17**(6), 98 (2019)
7. Gordon, D., Wirz, M., Roggen, D., Tröster, G., Beigl, M.: Group affiliation detection using model divergence for wearable devices. In: Proceedings of the 2014 ACM International Symposium on Wearable Computers, pp. 19–26, September 2014

8. Bourbia, A.L., Son, H., Shin, B., Kim, T., Lee, D., Hyun, S.J.: Temporal dependency rule learning based group activity recognition in smart spaces. In: 2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC), Vol. 1, pp. 658–663, June 2016
9. Yu, N., Zhao, Y., Han, Q., Zhu, W., Wu, H.: Identification of partitions in a homogeneous activity group using mobile devices. Mob. Inform. Syst. **2016**, Article ID 3545327, 14 p. (2016)
10. Ma, C., Li, Q., Li, W., Gravina, R., Zhang, Y., Fortino, G.: Activity recognition of wheelchair users based on sequence feature in time-series. In: Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 3659–3664, October 2017
11. Álvarez Lacasia, J., Leppänen, T., Iwai, M., Kobayashi, H., Sezaki, K.: A method for grouping smartphone users based on Wi-Fi signal strength. In: Forum on Information Technology, vol. 32, pp. 450–452 (2013)