BaG: Behavior-aware Group Detection in Crowded Urban Spaces using WiFi Probes

Jiaxing Shen
The Hong Kong Polytechnic University
csjshen@comp.polyu.edu.hk

Jiannong Cao
The Hong Kong Polytechnic University
csjcao@comp.polyu.edu.hk

Xuefeng Liu
Huazhong University of Science and Technology
csxfliu@gmail.com

ABSTRACT

Group detection is gaining popularity as it enables various applications ranging from marketing to urban planning. The group information is an important social context which could facilitate a more comprehensive behavior analysis. An example is for retailers to determine the right incentive for potential customers. Existing methods use received signal strength indicator (RSSI) to detect co-located people as groups. However, this approach might have difficulties in crowded urban spaces since many strangers with similar mobility patterns could be identified as groups. Moreover, RSSI is vulnerable to many factors like the human body attenuation and thus is unreliable in crowded scenarios. In this work, we propose a behavior-aware group detection system (BaG). BaG fuses people's mobility information and smartphone usage behaviors. We observe that people in a group tend to have similar phone usage patterns. Those patterns could be effectively captured by the proposed feature: number of bursts (NoB). Unlike RSSI, NoB is more resilient to environmental changes as it only cares about receiving packets or not. Besides, both mobility and usage patterns correspond to the same underlying grouping information. The latent associations between them cannot be fully utilized in conventional detection methods like graph clustering. We propose a detection method based on collective matrix factorization to reveal the hidden associations by factorizing mobility information and usage patterns simultaneously. Experimental results indicate BaG outperforms baseline approaches by 3.97% ∼ 15.79% in F-score. The proposed system could also achieve robust and reliable performance in scenarios with different levels of crowdedness.

CCS CONCEPTS

• **Information systems → Mobile information processing systems**; **Clustering**.

KEYWORDS

group detection; WiFi; probe request; collective matrix factorization

ACM Reference Format:
Jiaxing Shen, Jiannong Cao, and Xuefeng Liu. 2019. BaG: Behavior-aware Group Detection in Crowded Urban Spaces using WiFi Probes. In *Proceedings of the 2019 World Wide Web Conference (WWW '19), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3308558.3313590

WWW '19, May 13–17, 2019, San Francisco, CA, USA

ACM ISBN 978-1-4503-6674-8/19/05.
https://doi.org/10.1145/3308558.3313590

1 INTRODUCTION

Group detection plays an important role in many applications including marketing [29], healthcare [22, 35], and urban planning [3, 39]. The grouping information is an essential contextual feature in multi-target tracking [26] and behavior analysis [12]. This social context facilitates better interpretation of observed events and detection of abnormal behaviors [13]. A typical scenario is for retailers to determine the right incentive for potential customers. For example, "Buy X get Y free" promotions would be more attractive to group customers than individual customers [29]. It is commonly assumed a group could be discerned by people's physical locations and behaviors [6, 23, 29, 34]. Therefore, *group detection* is defined as a task of clustering a set of people into disjoint subsets based on their locations and behaviors.

Existing works exploit *probe requests*[1] to detect co-located people as a group [12, 19]. Co-location is achieved through received signal strength indicator (RSSI) contained in probe requests. Besides, those probes could be collected without difficulty since they are broadcast automatically to seek information about nearby access points (APs). Compared with other approaches, this method requires neither high deployment cost (e.g., deploy cameras [9, 34]) nor user intervention (e.g., carry wearable devices or install apps [18, 24, 29]). The probe-based methods are thus capable of detecting large-scale spontaneous groups in naturalistic environments.

However, current probe-based methods might have some difficulties in crowded environments. First, many strangers are close to each other in crowded spaces [29]. For example, two strangers might walk closely along an aisle. This indicates detecting co-located people as a group in crowded areas is error-prone. Second, RSSI is vulnerable to many factors like device diversity, multipath fading, and body attenuation [30]. It is unreliable to use RSSI to detect groups in dynamic environments crowded with people moving around.

We ask the following question: *can we reliably detect groups with WiFi probes in crowded urban spaces?* In this paper, we provide an affirmative answer by proposing a Behavior-aware Group detection (BaG) system that integrates both mobility information and phone usage behaviors. We have a key observation that people in the same group tend to have similar phone usage patterns. The observation is closely related to the concept of "phubbing" that is rife throughout the world. "Phubbing" is defined as the act of snubbing others in social interactions and instead focusing on one's smartphone [11].

[1] *Probe request*, *WiFi probe*, and *probe* are used interchangeably

According to a survey of 276 participants, the correlation of "phubbing" and "being phubbed" is 0.6 [4], which indicates individuals in a group would have similar phone usage patterns. The authors [4] explained this phenomenon occurs might due to false-consensus effects. Individuals assume that others think and do the same as themselves. Besides, we contribute a new feature (*number of bursts, NoB*) extracted from WiFi probes that could effectively capture phone usage patterns. A *burst* is a set of probes sent over a very short period of time (less than 1 second) [8]. The higher frequency a phone is used, the more NoB is generated. Unlike RSSI, NoB is resilient to crowded environments since it only cares about whether there are probes or not.

The vision of BaG, however, entails sparsity and fusing challenges when applied to real conditions.

1) How to handle spatiotemporal sparsity of WiFi data? First, an area is usually covered by a limited number of APs (spatial sparsity). Second, the probing frequency of a smartphone is affected by many factors like operating systems and users' usage patterns [8] (temporal sparsity). The data in common of two users could be much sparser. Measuring user similarity with such data could lead to highly biased and unreliable results.

2) How to cluster users fusing mobility and behaviors? The integration of mobility and behaviors remains an open issue. Mobility and behaviors have latent associations since both of them are different perspectives of the real grouping information. Simply combing them like calculating their arithmetic mean might limit the potential of hidden associations and thus derive unsatisfactory results.

For the sparsity challenge, we first represent mobility and behavior information into two matrices respectively. Then we apply matrix factorization (MF) to handle data sparsity by decomposing an input matrix into the product of several factor matrices. For the fusing challenge, we consider sparsity-constrained collective nonnegative MF (SCNMF). The advantages are two-fold. First, collective MF (CMF) takes the correlation of both input matrices into consideration by factorizing them simultaneously. Second, the sparsity constraint of CMF makes it an alternative for clustering so that grouping results could be derived directly without extra clustering processes.

According to our experimental evaluation with 82 volunteer groups in a large shopping mall, BaG achieves robust and reliable performance in scenarios with different levels of crowdedness. Compared with baseline approaches, BaG improves F-score of detection by 3.97% $\sim$ 15.79% in labeled data and 6.67% $\sim$ 20.69% in synthetically labeled data, respectively.

Our main contributions are summarized as follows.

- We introduce a new improvement of group detection in crowded environments: phone usage behaviors.
- A new feature (NoB) is extracted from WiFi probes that could effectively capture phone usage behaviors.
- We propose a new group detection method (SCNMF) that fuses mobility and behaviors and derives the grouping results without extra clustering processes.

The remainder of this paper is organized as follows. We present preliminaries in Section 2. Section 3 introduces design details of BaG, followed by experimental evaluation of the system in Section 4. Related work is introduced in Section 5. We conclude this paper and introduce the future work in the last section.

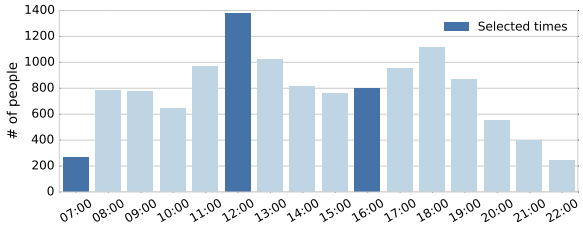


Figure 1: Average number of people in a day in the mall.

2 PRELIMINARIES

First of all, we use a two-day pilot study to illustrate relationships of the following variables in crowded environments: 1) RSS difference vs. groups 2) Phone usage behaviors vs. groups ; 3) Phone usage behaviors vs. NoB.

Experiments are conducted in a large shopping mall where we have deployed 20 APs to sniff WiFi probes. More detailed settings could be found in Section 4. Although it is difficult to define crowdedness directly, different numbers of onsite people could reflect different levels of crowdedness. Figure 1 highlights selected times of the experiment and the average number of people in the mall. The order of crowdedness is 12:00 > 16:00 > 7:00.

In each experiment, we recruit around 10 volunteer groups (each group consists of 2 $\sim$ 5 people) and record their grouping information and MAC addresses. We install an app on their smartphones to record screen on-off states as phone usage events. To ensure the authenticity, volunteers are only told to keep WiFi enabled and behave as normal without knowing the purposes of the experiment.

2.1 RSS Difference vs. Groups

RSSI methods calculate RSS (received signal strength) difference of two users. To this end, we first process users' WiFi data with an interval of 60 seconds, as probes are sent in a granularity of minutes [12]. Then user i's WiFi data in the t-th minute is represented as a vector:

$$\mathbf{v}_i^t = [rss_1, rss_2, \cdots, rss_N], \tag{1}$$

where rss_n represents received signal strength of the n-th AP, N is the number of total APs. If no RSS is available for AP n, then rss_n is set to 0. According to [37], RSS difference of users i and j at time t could be calculated with Eq. 2 and 3:

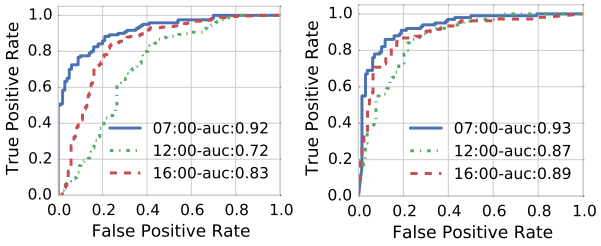
$$d(\mathbf{v}_i^t, \mathbf{v}_j^t) = \frac{1}{|A_i \cap A_j|} \cdot \sum_{a \in A_i \cap A_j} \left| rss_a^i - rss_a^j \right| \tag{2}$$

$$A_i = \left\{ k \middle| rss_k^i \neq 0, k \in [1, N] \right\} \tag{3}$$

The averaged RSS difference $D(i, j)$ over all common time slots is calculated with Eq. 4, where T is a set of common time slots of user i and j.

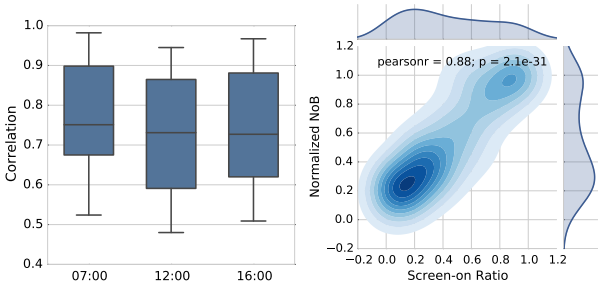
$$D(i, j) = \frac{1}{|T|} \sum_{t \in T} d(\mathbf{v}_i^t, \mathbf{v}_j^t), \tag{4}$$

Then we derive a distribution $\mathbf{X}_g(k)$ of $D(i, j)$ for all group pairs appear in time slot $k \in \{7{:}\,00, 12{:}\,00, 16{:}\,00\}$. Formal definition



(a) RSS difference (b) Behavior difference

Figure 2: ROC curves of different features at selected times. (a) With RSS difference; (b) With behavior difference.



(a) (b)

Figure 3: (a) Boxplot of the correlation for all users at selected times; (b) An example of correlation analysis of *screen-on ratio* and *normalized NoB* for a single user.

of $\mathbf{X}_g(k)$ is shown in Eq. 5, where boolean function $g(i,j)$ returns 'True' when user i and j are from the same group and U_k is a set of users appear in time slot k. Similarly, for all non-group pairs, we have another distribution $\mathbf{X}_{\tilde{g}}(k)$.

$$\begin{aligned}\mathbf{X}_g(k) &= \left\{D(i,j)\middle|g(i,j) = True; i,j \in U_k\right\}\\ \mathbf{X}_{\tilde{g}}(k) &= \left\{D(i,j)\middle|g(i,j) = False; i,j \in U_k\right\}\end{aligned} \tag{5}$$

Usually, $\mathbf{X}_g(k)$ and $\mathbf{X}_{\tilde{g}}(k)$ overlap to some extent, so we cannot easily find a cut-off point that completely separates both distributions. However, the smaller the overlap, the better the separation performance. To compare the effectiveness of using RSS difference more objectively, we exploit Receiver Operating Characteristic (ROC) curve, which is usually used to illustrate the diagnostic ability of a binary classifier system as its discrimination threshold varies. The curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings.

Figure 2a shows the comparison of ROC curves using RSS difference at selected times. It is clear that the effectiveness of different times varies significantly, indicating RSS difference is not a consistent feature in crowded environments. Besides, the order of effectiveness is 7:00 > 16:00 > 12:00, which is exactly the opposite order of the crowdedness. This implies crowded environment might have a significant negative impact on RSSI methods.

2.2 Smartphone Usage Behaviors vs. Groups

To represent phone usage behaviors, we use a straightforward way. If a smartphone is in use, the screen must be on. Otherwise, the screen is off. To represent user i's usage behaviors, we use a behavior vector $\mathbf{b}_i$:

$$\mathbf{b}_i = \left[\frac{s_1}{60}, \frac{s_2}{60}, \cdots, \frac{s_t}{60}, \cdots\right], \tag{6}$$

where s_t is a total number of seconds when the screen is on in the t-th minute. Behavior difference $D'(i,j)$ of user i and j is measured with Euclidean distance in Eq. 7.

$$D'(i,j) = \frac{1}{60}\sqrt{\sum_{t=1}(s_t^i - s_t^j)^2} \tag{7}$$

The behavior difference distributions $\mathbf{Y}_g(k)$ and $\mathbf{Y}_{\tilde{g}}(k)$ at time k are defined as:

$$\begin{aligned}\mathbf{Y}_g(k) &= \left\{D'(i,j)\middle|g(i,j) = True; i,j \in U_k\right\}\\ \mathbf{Y}_{\tilde{g}}(k) &= \left\{D'(i,j)\middle|g(i,j) = False; i,j \in U_k\right\}\end{aligned} \tag{8}$$

Similarly, we draw ROC curves for $\mathbf{Y}_g(k)$ and $\mathbf{Y}_{\tilde{g}}(k)$. As shown in Figure 2b, the ROC curves at different times are close to each other, indicating behavior difference is stable in different crowdedness settings. Besides, the performance (AUC, area under curve, are all over 0.85) proves behavior difference is effective for group detection.

2.3 Smartphone Usage Behaviors vs. NoB

Although different smartphones have different probing patterns, it is reported that a smartphone will send more probes when the screen is on [8]. Therefore, the more frequently a smartphone is used, the more NoB is generated. A recent research work [15] uses the number of wireless packets to predict the screen on-off states and achieves an accuracy ranging from 93% to 100%. NoB is different from the work in two aspects. First, only WiFi probes are used here. Second, we use the number of bursts instead of the number of packets.

We represent user i's NoB feature with a NoB vector:

$$\mathbf{n}_i = [n_1, n_2, \cdots, n_t, \cdots], \tag{9}$$

where n_t is the number of bursts in the t-th minute. To handle the impact of device diversity, we further calculate a normalized NoB vector with Eq. 10, where $max(.)$ find the largest element of an input vector.

$$\hat{\mathbf{n}}_i = \mathbf{n}_i / max(\mathbf{n}_i) \tag{10}$$

To analyze the correlation between smartphone usage behaviors and NoB, we calculate the Pearson Correlation Coefficient of the normalized NoB vector ($\hat{\mathbf{n}}_i$) and the behavior vector ($\mathbf{b}_i$) for users in all experiments. Figure 3a shows the boxplot of the correlation from all users. It is clear that smartphone usage behaviors and NoB are closely correlated in different crowdedness settings. Furthermore, an example of correlation analysis for a single user is illustrated in Figure 3b. The correlation coefficient and the p-value are 0.88 and $2.1e^{-31}$, respectively, which shows it is significant that screen-on ratio is positively correlated with NoB for the user. Furthermore, we could find that the correlation is higher when both variables are small. This means if the screen is on, it is more likely that the smartphone will send probes.

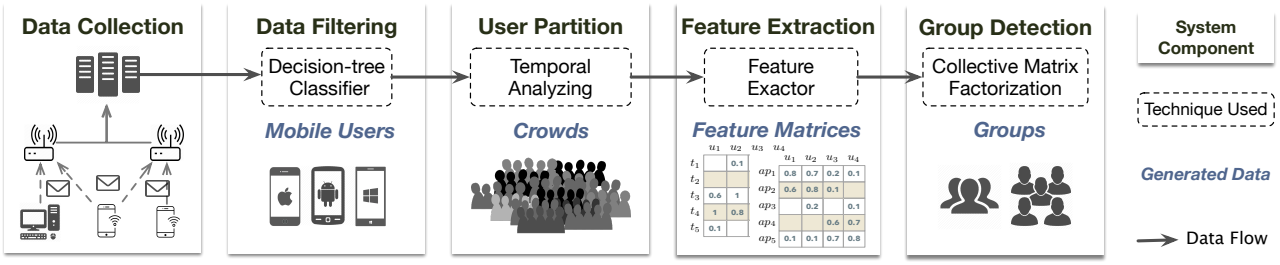


Figure 4: System overview of BaG.

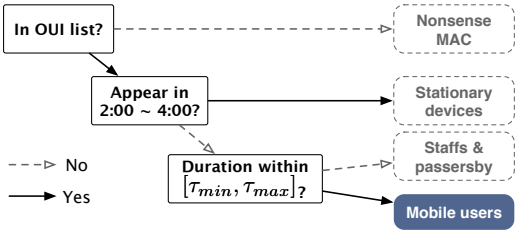


Figure 5: Filtering out WiFi data from non-mobile users.

3 SYSTEM DESIGN

In this section, we elaborate on design details of BaG. The overview of the proposed system is illustrated in Figure 4. First, we passively sniff WiFi probes from different devices (Data Collection). Then we filter out data from non-mobile users (Data filtering) and separate the filtered data into several crowds (User Partition). Besides, mobility and behavior related features are extracted and represented in matrices (Feature Extraction). Last, we use collective matrix factorization to fuse two matrices and derive grouping results (Group Detection) without extra clustering processes.

3.1 Data Collection

We exploit off-the-shelf WiFi APs to sniff WiFi probes from nearby devices and upload the data to a server. The AP configurations are as follows: AR9341 (WLAN chip), 64M (RAM), and 8M (Flash Memory). We use those APs for collecting probe requests only since they have limited computational and storage capacity.

Each AP works under OpenWrt[2] with a virtual network interface enabled in monitor mode. We use Tcpdump (a utility for capturing network traffic) to sniff WiFi probes and upload the collected data to the server at 2:00 AM when the bandwidth is fully available. Once the upload process is complete, local WiFi data will be deleted to make space for the new coming data.

3.2 Data Filtering

Data filtering services as a preprocessing of the raw WiFi data including three steps: combat with MAC randomization, filter out data from non-mobile users, and extract useful data fields.

To prevent third parties from tracking devices with MAC (Media Access Control) address, several vendors have implemented MAC address randomization which implies that probe requests no longer use the real MAC address of the device. However, this mechanism may not work as expected in reality since it is reported defeated in several recent works [8, 21, 38]. Besides, this randomization mechanism is not activated in the majority of old devices, especially for Android systems. We follow the practice described in [21] to defeat MAC randomization.

As the collected data may come from stationary devices (like desktops and IP cameras) and mobile devices, we remove the data of non-mobile users devices with a simple decision-tree classifier as illustrated in Figure 5. If a MAC is not in OUI (Organizationally Unique Identifier) list, it might be a forged MAC or from unregistered companies which makes no sense to us. Then we use the routine of human activities to filter out stationary devices as during 2:00 ~ 4:00, most public places are closed, and most of the people are in sleep. Last, with the help of two scenario-dependent parameters τ_{min} and τ_{max} which represent the minimum and maximum dwell time, we could filter out data from staffs and passersby. In our system, we empirically set $\tau_{min} = 10$ and $\tau_{max} = 240$ which means people stay for at least 10 minutes and up to 4 hours. This setting is based on our observation and is intended to identify customer groups. If we want to detect staff groups, both parameters should be relatively larger. More details could be found in Reference [32].

Lastly, we extract desired fields from WiFi probes. For each WiFi probe, we can get the following data entry:

$$< \text{Timestamp, Device\_MAC, AP\_MAC} >,$$

where *Timestamp* indicates the time of receiving a probe, *MAC* is MAC address which uniquely represents a smartphone or an AP.

3.3 User Partition

Instead of detecting groups out of all mobile users in a whole day, we first utilize temporal constraint of groups to separate users into different *crowds* and then identify groups out of each crowd. The main advantage of user partition is that the efficiency of group detection can be greatly improved since it breaks down the original problem into many subproblems which can be solved independently.

The idea of the temporal constraint is very simple. The time gap of the appearance of groups members would never be too large. An illustration is shown in Figure 6. For an extreme example, group members would not appear separately in the morning and in the night. Formally, the temporal constraint is defined as

$$max(|ta_i - ta_j|) \leq \psi, \tag{11}$$

[2]OpenWrt is a highly extensible GNU/Linux distribution for embedded devices (typically wireless routers).

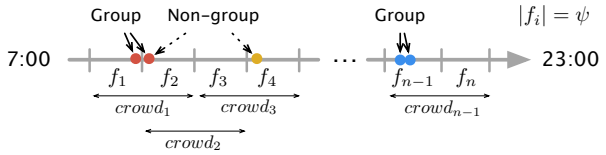


Figure 6: Partitioning users of a day into crowds.

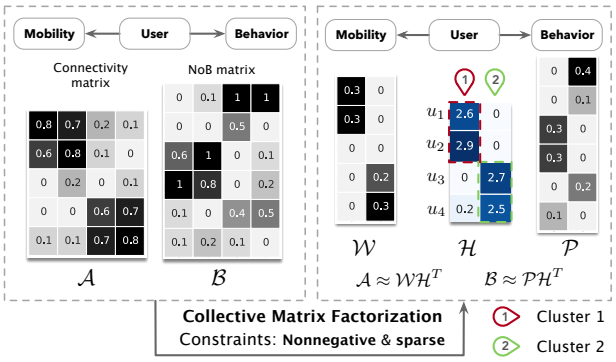


Figure 7: Illustration of using CMF to cluster users.

where ta_i means the appearance time of user i and ψ is a threshold. The physical meaning of parameter ψ is the dwell time of the most people. The selection of ψ is discussed in Section 4.

Then we equally partition the daytime into non-overlapping fragments using ψ. The partition process is depicted in Figure 6. For any two sequential fragments f_i and f_{i+1}, we take out WiFi data entries of users whose start time falls in both fragments. Those users are regarded as a *crowd* and their WiFi data entries are denoted as $crowd_i$. If a person appears in f_i, his/her group mate may appear in the same fragment or the next fragment, but it is barely possible to appear in f_{i+2} since ψ restricts the dwell time of most people.

3.4 Feature Extraction

Given data entries of a crowd, we represent two features as two feature matrices. The first feature is NoB that indicates a user's phone usage behaviors. The other is *AP connectivity* which indicates a user's mobility information.

As introduced in Section 1, NoB refers to how many times a device has sent WiFi probes. AP connectivity means how long a user has been in the coverage area of an AP. It reflects where a user has been to and how long they have stayed. Compared to RSSI, although AP connectivity is more coarse-grained, it is also more robust and reliable. Both AP connectivity and NoB are insensitive to crowded environments as they only care about whether a smartphone sends WiFi probes or not which cannot be readily affected.

We first group WiFi data entries by users' MAC addresses. For each user, we raise the granularity of the data from seconds to minutes, as this practice reduces the impact of sparse WiFi probes. To eliminate the negative impact of device diversity, we normalize connectivity vector and NoB vector for each user. Lastly, we combine all users' connectivity and NoB vectors and generate corresponding matrices.

3.5 Group Detection

Group detection is essentially a *hard clustering* problem which means each user can only belong to a cluster or not. Previous works [12, 29] need to measure pairwise user similarity and then construct a user graph with nodes representing users and the edge between nodes indicating their similarity. After that, an explicit graph clustering process like Markov Cluster Algorithm is applied to the graph to detect groups.

However, it would be inefficient and cumbersome to construct the graph by measuring pairwise user similarity. Instead, we use nonnegative matrix factorization (NMF) to derive the group results directly without extra clustering processes. We choose NMF rather than other matrix factorization like singular value decomposition (SVD) as entries in both original matrices have physical meaning (number of bursts and duration time). Positive factors facilitate direct physical connections. Besides, collective nonnegative matrix factorization (CNMF) is used to jointly factorize connectivity and NoB matrices to get better results as both matrices share the same latent feature: the grouping information. An illustration of the whole process is shown in Figure 7. We have two relation matrices $\mathcal{A}$ (user-mobility matrix) and $\mathcal{B}$ (user-behavior matrix) as input. After collective matrix factorization, there are three factors corresponding to three entities. The user entity matrix $\mathcal{H}$ contains the group results. Each row of $\mathcal{H}$ indicates a potential group with $\mathcal{H}_{ij}$ representing the strength that user i belongs to group j. From $\mathcal{H}$ we can see clearly that the group results are $(u_1, u2)$ and (u_3, u_4).

In the following subsections, we first briefly review the rationale of using NMF for clustering. Then we focus on solving the group detection problem using CNMF.

3.5.1 ***NMF and k-means.*** Given an input matrix $\mathcal{A} \in \mathbb{R}^{m \times n}$ and an integer $k < \min(m, n)$, NMF aims to find two nonnegative factors $\mathcal{W} \in \mathbb{R}^{m \times k}$, $\mathcal{H} \in \mathbb{R}^{n \times k}$ such that $\mathcal{A} \approx \mathcal{W}\mathcal{H}^T$. $\mathcal{W}$ and $\mathcal{H}$ can be found by solving the optimization problem:

$$\min_{\mathcal{W},\mathcal{H}} F_k(\mathcal{W},\mathcal{H}) = \frac{1}{2}\left\|\mathcal{A} - \mathcal{W}\mathcal{H}^T\right\|_F^2 s.t. \quad \mathcal{W},\mathcal{H} \geq 0, \tag{12}$$

where $F_k(\cdot)$ is the loss function and k is the reduced dimension. $\|\mathcal{A}\|_F$ means Frobenius Norm of matrix $\mathcal{A}$. The reasons of using Frobenius Norm are that it has a Guassian noise interpretation and the objective function can be easily transformed to a matrix trace version.

Existing works [17] provide evidence that sparsity constrained NMF is a viable alternative as a clustering method. The objective of k-means is to minimize the sum of squared distances from each data point to its centroid. With $\mathcal{A} = [a_1, ..., a_n] \in \mathbb{R}^{m \times n}$, the objective function J_k with given integer k can be written as:

$$J_k = \sum_{j=1}^{k} \sum_{a_i \in C_j} \left\|a_i - c_j\right\|^2 = \left\|\mathcal{A} - C\mathcal{B}^T\right\|_F^2, \tag{13}$$

where $\mathcal{B} \in \mathbb{R}^{n \times k}$ denotes the clustering assignment. If the i-th observation is assigned to the j-th cluster $\mathcal{B}_{ij} = 1$, otherwise $\mathcal{B}_{ij} = 0$. $C = [c_1, ..., c_k] \in \mathbb{R}^{n \times k}$ is the centroid matrix where c_j is the cluster centroid of cluster C_j. C can also be written as $C = \mathcal{A}\mathcal{B}\mathcal{D}^{-1}$

Algorithm 1 Alternating nonnegative least squares

Require: Initialize $\mathcal{H}$ and $\mathcal{P}$ with random nonnegative values.

1: **while** the convergence criterion is not satisfied **do**

2: $\mathcal{W} \leftarrow \arg\min_{\mathcal{W}} \left\| \begin{pmatrix} \sqrt{\alpha}\mathcal{H} \\ \sqrt{\eta}\mathbf{I}_k \end{pmatrix} \mathcal{W}^T - \begin{pmatrix} \sqrt{\alpha}\mathcal{A}^T \\ \mathbf{0}_{k\times m} \end{pmatrix} \right\|_F^2$

3: $\mathcal{P} \leftarrow \arg\min_{\mathcal{P}} \left\| \begin{pmatrix} \sqrt{1-\alpha}\mathcal{H} \\ \sqrt{\eta}\mathbf{I}_k \end{pmatrix} \mathcal{P}^T - \begin{pmatrix} \sqrt{1-\alpha}\mathcal{A}^T \\ \mathbf{0}_{k\times l} \end{pmatrix} \right\|_F^2$

4: $\mathcal{H} \leftarrow \arg\min_{\mathcal{H}} \left\| \begin{pmatrix} \sqrt{\alpha}\mathcal{W} \\ \sqrt{1-\alpha}\mathcal{P} \\ \sqrt{\beta}\mathbf{e}_{1\times k} \end{pmatrix} \mathcal{H}^T - \begin{pmatrix} \sqrt{\alpha}\mathcal{A} \\ \sqrt{1-\alpha}\mathcal{B} \\ \mathbf{0}_{1\times n} \end{pmatrix} \right\|_F^2$

5: **end while**

where $\mathcal{D}^{-1}$ is a diagonal matrix defined in:

$$\mathcal{D}^{-1} = diag(\frac{1}{|C_1|}, \frac{1}{|C_2|}, ..., \frac{1}{|C_k|}) \in \mathbb{R}^{k\times k} \tag{14}$$

$|C_j|$ denotes the number of data points in cluster j. Since a diagonal matrix can be factored as any two diagonal matrices $\mathcal{D}^{-1} = \mathcal{D}_1\mathcal{D}_2$. Defining $\mathcal{F} = \mathcal{B}\mathcal{D}_1$ and $\mathcal{H} = \mathcal{B}\mathcal{D}_2$, the objective function J_k can be rephrased as:

$$J_k = \left\|\mathcal{A} - \mathcal{A}\mathcal{B}\mathcal{D}_1\mathcal{D}_2\mathcal{B}^T\right\|_F^2 = \left\|\mathcal{A} - \mathcal{A}\mathcal{F}\mathcal{H}^T\right\|_F^2 \tag{15}$$

where $\mathcal{F}$ and $\mathcal{H}$ have exactly one positive element in each row. Factor $\mathcal{H}$ has exactly one nonzero element for each row which represents a hard clustering result of the corresponding data point. If we let $\mathcal{W} = \mathcal{A}\mathcal{F}$, J_k is similar to objective function F_k of NMF. The difference is that NMF formulation does not have the constraints that $\mathcal{H}^T$ is a sparse matrix. To this end, sparse NMF (SNMF) is proposed. The idea is to use ℓ_1-norm regularization to achieve sparsity of the factorization [36].

3.5.2 ***SCNMF for Group Detection***. Now we provide our formulation and solution to the group detection problem. Given two input matrices: connectivity matrix $\mathcal{A} \in \mathbb{R}^{m\times n}$ and NoB matrix $\mathcal{B} \in \mathbb{R}^{l\times n}$, and an integer $k < \min\{m, n, l\}$, sparsity-constrained collective nonnegative matrix factorization (SCNMF) aims to find three nonnegative factors $\mathcal{W} \in \mathbb{R}^{m\times k}$, $\mathcal{H} \in \mathbb{R}^{n\times k}$ (sparse matrix), and $\mathcal{P} \in \mathbb{R}^{l\times k}$ such that:

$$\begin{cases} \mathcal{A} \approx \mathcal{W}\mathcal{H}^T \\ \mathcal{B} \approx \mathcal{P}\mathcal{H}^T \end{cases} \tag{16}$$

The solutions $\mathcal{W}, \mathcal{H}, \mathcal{P}$ can be found by solving the optimization problem:

$$\begin{aligned} \min_{\mathcal{W},\mathcal{H},\mathcal{P}} \frac{1}{2}\Big[\alpha\left\|\mathcal{A} - \mathcal{W}\mathcal{H}^T\right\|_F^2 + (1-\alpha)\left\|\mathcal{B} - \mathcal{P}\mathcal{H}^T\right\|_F^2 + \\ \beta\sum_{j=1}^{n}\left\|\mathcal{H}_j\right\|_1^2 + \eta(\|\mathcal{W}\|_F^2 + \|\mathcal{P}\|_F^2)\Big] \text{ s.t. } \mathcal{W},\mathcal{H},\mathcal{P} \geq 0, \end{aligned} \tag{17}$$

where $\mathcal{H}_j$ is the j-th row vector of $\mathcal{H}$. Parameter $\alpha \in [0, 1]$ is scenario-dependent which weights the relative importance of two input matrices. Parameter $\beta > 0$ balances the trade-off between the accuracy of approximation and the sparseness of $\mathcal{H}$. Parameter $\eta > 0$ controls the size of the elements of $\mathcal{W}$ and $\mathcal{P}$ respectively. It is usually determined by the largest element of input matrices [17].

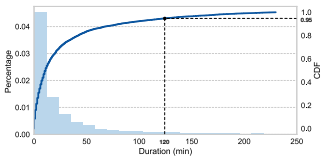


Figure 8: Distribution and CDF of dwell time in the mall.

Although solving Eq. 17 is a non-convex problem [10], it is convex separately in each factor, i.e., finding the optimal factor $\mathcal{W}$ corresponding to fixed factors $\mathcal{H}$ and $\mathcal{P}$ reduce to a convex optimization problem. Algorithms based alternating nonnegative least squares (ANLS) are often used for sparse NMF. Then we use Algorithm 1 to iterate the ANLS until a convergence criterion, which is set to iterate for 200 times [16], is satisfied. In the algorithm, $\mathbf{I}_k$ is an identity matrix of size $k \times k$, $\mathbf{0}_{k\times m}$ is a zero matrix of size $k \times m$, and $\mathbf{e}_{1\times k}$ is a row vector with all components equal to one. More details can be found in [16].

3.5.3 ***Determine the correct number of*** k. Parameter k is an input of the system which corresponds to the potential number of groups. In [17], consistency of clustering algorithms is used to determine the correct number of k from data. Consistency means the consistency clustering results under random initializations. To measure consistency, we construct a matrix $C_k \in \mathbb{R}^{n\times n}$ where n is the number of all users and k is the number of clusters. For pairwise users i and j, $C_k(i, j) = 1$ if they are assigned to the same cluster, otherwise $C_k(i, j) = 0$. Then we calculate the averaged $\hat{C}_k$ over trails. $\hat{C}_k(i, j)$ represents the possibility user i and user j being assigned to the same cluster. If the clustering results were consistent throughout all trails, each element in $\hat{C}_k$ would be close to either 0 or 1. A general quality of consistency is proposed by dispersion coefficient:

$$\rho_k = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} 4\Big[\hat{C}_k(i, j) - \frac{1}{2}\Big]^2 \tag{18}$$

where $\rho_k \in [0, 1]$ and $\rho_k = 1$ indicates the perfect consistent clustering results. After obtaining the values of ρ_k for various k, the number of clusters could be determined by the point k' where $\rho_{k'}$ drops. More detailed information can be found in [17].

4 EXPERIMENTAL EVALUATION

4.1 Settings

4.1.1 ***Setup***. We conduct experiments in a large shopping mall with 4 floors covering an area of $4890m^2$. The experiments here and the experiments conducted in Section 2 are different in terms of purposes and scales.

There are originally 20 APs installed in the mall for customers to access the Internet. We use those APs to collect the WiFi data. Within one week, we recruit 82 volunteer groups including 294 volunteers during 28 experiments at different times of a day. For each experiment, we recruit $2 \sim 4$ volunteer groups with each group containing $2 \sim 5$ people and record their MAC addresses and grouping information. The majority of experiments last from 0.5 hour to 1.5 hours. To ensure authenticity, volunteers are only

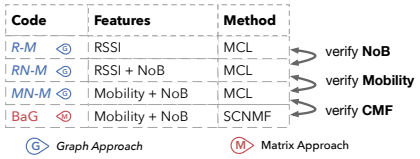


Figure 9: Illustration of baseline approaches.

told to keep their WiFi enabled without knowing the purpose of experiments.

4.1.2 ***Collected data.*** During the week, 4, 184, 778 WiFi probes are received from 59, 282 devices a day on average. According to our data filtering (Section 3.2), only 6.66% (3, 951) of them are from customers' mobile devices.

Figure 8 depicts the distribution of customers' dwell time which obeys the power law distribution. We also draw cumulative distribution function (CDF) of dwell time in Figure 8. It shows that 95% customers stay in the mall for less than 2 hours. This finding is useful because it helps to determine the value of parameter ψ (in Section 3.3).

4.1.3 ***Ground truth.*** BaG takes the collected data of each day as input and output the detected groups. We have no ways of knowing the grouping information for everyone in the collected data. However, we could know the relationships of a subset of those people.

Based on the information of our volunteers, we have prepared two evaluation datasets. The difference between two datasets is the number of non-group pairs. We call the first dataset as *labeled dataset* as it is entirely contributed by volunteers. In other words, we know the relation (group and non-group) of every pair of users. This dataset consists of 438 group pairs (positive samples) and 1, 135 non-group pairs (negative samples).

The second dataset is *synthetically labeled dataset* which includes some non-volunteers. When recruiting volunteer groups, we engage all group members in the experiment. Therefore, user pairs consisting of volunteer and non-volunteers must be non-group. Eventually, we get a dataset of 438 positive samples, and 2, 311 negative samples.

Even though we could find significantly more negative samples, we carefully control such negative samples. One reason is that such synthetic negative sample should also obey the temporal constraint as mentioned in Section 3.3, otherwise it could be too easy to detect. Another reason is that the difference of numbers between positive and negative samples should not be too large, otherwise the evaluation metrics might be affected.

4.1.4 ***Metrics.*** There is no consensus on which metrics should be used to evaluate groups detection [34]. Here we use *Precision*, *Recall* and *F-score* to measure the performance of group detection, which are defined as follows.

$$\begin{cases} \text{Precision(P)} = \dfrac{tp}{tp + fp} \\ \text{Recall(R)} = \dfrac{tp}{tp + fn} \\ \text{F-score} = 2 \cdot \dfrac{\text{P} \cdot \text{R}}{\text{P} + \text{R}} \end{cases}$$

Truth / Detection	g	$\tilde{g}$
g	tp	fp
$\tilde{g}$	fn	tn

g : Group
$\tilde{g}$: Non-group

Table 1: Detailed results of different α when $\beta = 0.2$.

α	precision	recall	F-score
.1	.861	.887	.874
.3	.897	.898	.897
.5	.822	.914	.866
.7	.743	.884	.807
.9	.690	.847	.760

Table 2: Detailed results of different β when $\alpha = 0.5$.

β	precision	recall	F-score
.1	.815	.912	.861
.2	.822	.918	.867
.3	.829	.915	.870
.4	.813	.905	.857
.5	.817	.896	.855

4.1.5 ***Baseline approaches.*** As mentioned in Introduction, two of our contributions are an effective feature (NoB) and a clustering method (CMF) for detecting groups. To verify the effectiveness of both NoB and CMF, we set three baselines R-M, RN-M, and MN-M. A baseline has two important components: feature(s) and method. Figure 9 shows detailed configurations of all approaches. According to the clustering methods used, those approaches can be further classified as graph approach (using Markov Cluster algorithm, MCL) and matrix approach (using SCNMF). Steps of graph approaches: 1) Measure user similarity based on some features. 2) Construct a user graph with users as nodes and their similarities as edge weights. 3) Run graph clustering algorithms to derive clusters or groups.

For different graph approaches, different feature combinations are used to measure user similarity. For multiple features, we combine them with arithmetic mean to measure similarity. Finally, MCL is applied to the constructed graph to detect groups. MCL is a popular graph clustering algorithm used by many existing works [12, 29]. It works well when the cluster size is small and it does not require the number of clusters as an input.

For matrix approach, BaG is a typical one. Therefore, it is unnecessary to repeat the steps again here. The precondition of this approach is to have matrices as input. That is why we do not have a matrix approach with RSSI and NoB since it is non-trivial to find an appropriate way of representing RSSI in the matrix format.

Table 3: Performance in labeled dataset.

Codes	7:00			16:00			12:00		
	P[1]	R[2]	F[3]	P	R	F	P	R	F
R-M	.874	.886	.880	.755	.788	.771	.700	.720	.710
RN-M	.920	.935	.927	.848	.876	.862	.815	.799	.807
MN-M	.905	.922	.914	.856	.886	.871	.817	.849	.832
BaG	.934	.954	.944	.887	.915	.901	.859	.892	.875

[1] Precision [2] Recall [3] F-score

Table 4: Performance in synthetically labeled dataset.

Codes	7:00			16:00			12:00		
	P	R	F	P	R	F	P	R	F
R-M	.751	.806	.778	.652	.715	.682	.618	.655	.636
RN-M	.821	.886	.852	.757	.832	.793	.729	.756	.742
MN-M	.806	.850	.827	.758	.810	.783	.727	.779	.752
BaG	.849	.905	.876	.804	.868	.835	.777	.842	.808

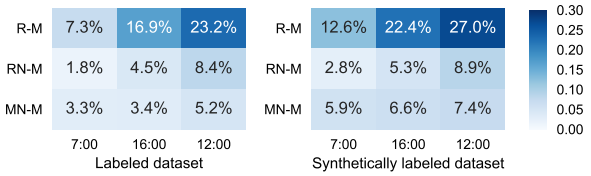


Figure 10: Performance gain ratios of BaG.

4.1.6 ***Parameter selection.*** In this system, there are 3 parameters to determine: ψ for user partition (Section 3.3) and α, β for group detection (Section 3.5).

In user partition, ψ represents the maximum duration time of onsite customers. As discussed in Section 3.3, the physical meaning of ψ is the dwell time of customers. While Figure 8 shows that 95% customers stay in the mall for less than 2 hours. Therefore, we set ψ to 120 minutes in our system.

For group detection, parameter α controls the relative importance of connectivity matrix over NoB matrix. According to the results shown in Table 1, we set α to 0.3. Generally, small α is preferred since connectivity matrix contains coarse-grained mobility information that has limited capacity to detect groups in crowded environments. Parameter β balances the trade-off between accuracy of approximation and sparseness. The performance of different β is illustrated in Table 2. Even though the performance is not that sensitive to β, too big β is undesirable since that might lead to worse approximation [17], so β is set to 0.2.

4.2 Evaluation

As explained in Section 2.1, we use the number of onsite people to represent different crowdedness settings. To have a comprehensive understanding of the performance of BaG, the evaluation is further explained with two datasets, four approaches, three feature combinations, and two clustering methods.

4.2.1 ***Two datasets.*** According to Figure 1, the order crowdedness for selected hours is 12:00 > 16:00 > 7:00. Detailed performance of all approaches on labeled and synthetically labeled datasets are shown in Table 3 and 4, respectively.

Compared with labeled dataset, all approaches have worse performance on the synthetically labeled dataset. This is reasonable since changes of the number of people to be detected might alter the clustering results. Besides, the probability that strangers are close to each other or have similar smartphones usage patterns are likely to increases with the number of people.

Interestingly, the average change of precision (−11.1%) is more significant than that of recall (−0.7%) over all approaches and times. One possible reason is the boost of negative samples increases false detections especially false positive detections (non-groups are detected as groups) and leads to a significant decline in precision.

4.2.2 ***Four approaches.*** Averaging the performance over all times, BaG improves F-score by 3.97% ∼ 15.79% and 6.67% ∼ 20.69% in labeled and synthetically labeled datasets, respectively.

We further demonstrate the performance of BaG by showing the *F-score gain ratio* over baselines. The ratio is defined in Eq. 19 where bl represents a baseline approach.

$$F\text{-score gain ratio} = \frac{F_{\text{BaG}} - F_{\text{bl}}}{F_{\text{bl}}}, \tag{19}$$

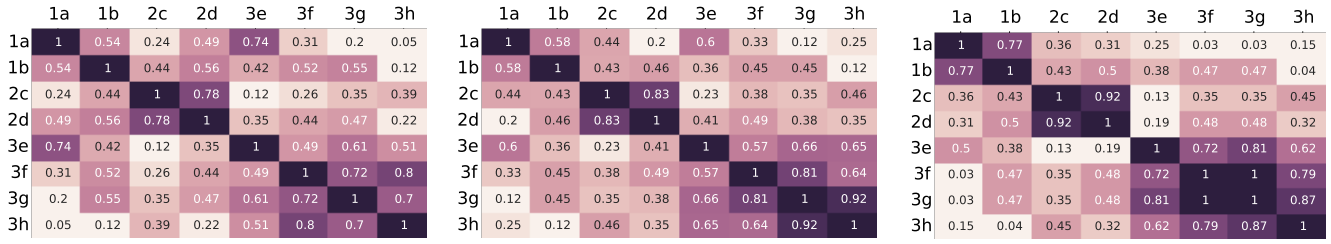
As shown in Figure 10, the ratio increases with the crowdedness indicating BaG is more accurate and reliable in crowded environments. Besides, the improvement is slightly higher in synthetically labeled data.

4.2.3 ***Three feature combinations.*** R-M, RN-M, and MN-M are all based on the same clustering method, so we could match them to compare the performance of different feature combinations.

From the performance tables, we have identified two issues on both datasets. First, the performance of feature combinations with NoB (RN-M and MN-M) are consistently higher than the other approach (R-M) over all times. The reasons might be two-fold. On one hand, the additional feature NoB has been used. On the other, it also proves NoB has the capability to capture the underlying grouping information in different crowdedness settings.

The second issue is that in the least crowded setting RN-M slightly outperforms MN-M, but the situation is just the opposite in the most crowded setting. This phenomenon results from different recall declining rates. When the crowdedness increases, both methods become worse, but recall of RN-M declines faster than that of MN-M. Potential reasons for different declining rates might be as follows. RSSI is more fine-grained than mobility but it is also more unreliable when the environment gets crowded since people frequently move around and hold their smartphones in different ways. In crowded environments, group users might have low similarity based on RSSI distance measurement and thus generate more false negative detections (groups are detected as non-groups) that degrade recall.

To support this, in Figure 11 we show an example of similarity matrices from a single experiment in 12:00 with 3 groups. As we can see in Figure 11a, some group users (like e and f) have low similarity



(a) RSSI **(b) RSSI+NoB** **(c) Mobility+NoB**

Figure 11: User similarity matrix from a single experiment in 12:00 using different feature combinations. Alphabets represent people, numbers reveal their grouping information. For example, $1a$ means user a is in group 1.

using RSSI. Although NoB has a certain effect, the improvement is quite limited as illustrated in Figure 11b. As a contrast in Figure 11c, the combination of mobility and NoB measures user similarity in a better way as mobility is more coarse-grained and thus insensitive to environmental changes.

4.2.4 ***Two clustering methods.*** Two clustering methods refer to graph approach and matrix approach. We could compare them by matching the performance of BaG and MN-M from performance tables and Figure 10. Overall, BaG slightly outperforms MN-M in all situations ranging from 3.3% $\sim$ 7.4%.

The main advantage of the matrix approach relies on the fusion of mobility and NoB. Mobility and NoB are two different perspectives of the real grouping information and thus have some hidden associations. With collective matrix factorization, we could simultaneously factorize them to reveal their hidden associations. While in the graph approach mobility and NoB are simply combined using arithmetic mean which might have limited the potential of the latent association.

5 RELATED WORK

Group detection can be classified as vision-based approaches, sensor-based approaches, and WiFi-based approaches according to different means.

Vision-based approaches regard group detection as a task of clustering set of users' trajectories into disjoint subsets [9, 34]. However, this kind of methods have some apparent limitations. First of all, the biggest issue is privacy erosion. Besides, video surveillance suffers from environmental issues such as non-line-of-sight, and low brightness.

Sensor-based approaches use wearable devices or install apps on smartphones to collect users' behavioral data. Groups are detected through correlation analysis of multiple sensor data. For instance, MIT researchers use specially designed wearable devices called "Sociometric Badges" [24, 25, 33] to measure group behavior through face-to-face interaction and physical proximity. Some research works [18, 20, 29] combine several sensor modalities (WiFi, accelerometer, compass, etc) to measures users similarity. However, these methods might be difficult to collect data on a scale, as they require user intervention which would be cumbersome in some scenarios. Besides, engaging multiple sensors drains smartphone battery more quickly.

WiFi-based approaches utilize the information contained in probe request to detect groups. Compared to other approaches, WiFi-based approaches do not require high deployment cost or user intervention. The probe contains significant information like time stamp, smartphone MAC address, RSSI, and SSID, which enables a wide range of applications like passive tracking [7], crowd counting[28, 40], and facility utilization analysis [27]. SSID and RSSI are two frequently used information to detect groups. Cunche et. al. [1, 2, 5] link different smartphones through SSID similarity. However, 80% of the devices reply with empty SSID list [14], approaches that rely on SSID may not work well anymore. Then researchers' focus transfer to RSSI which indicate users' mobility. Kjærgaard et. al. [19] extract spatial features, signal-strength features, and pseudo-spatial features from signal strength to detect social groups which they call pedestrian flocks. It is found that the performance of spatial features is unreliable, since mapping RSSI into locations is not accurate enough. Besides, the mapping process itself is usually time-consuming and labor-intensive. To avoid the cumbersome mapping process, directly measure the similarity of RSSI fingerprints to detect co-located mobile users. These methods get rid of absolute locations, thus eliminate labor-intensive calibration and protect users' privacy. SocialProbe [12] considers the hardware diversity and uses the normalized RSSI vector to achieve co-location detection. However, hardware difference is not the only factor, other factors like body attenuation, environmental disturbance, and multipath fading can also significantly affect RSSI. Therefore, RSSI based approach may not achieve reliable performance in some scenarios, especially in crowded urban spaces like shopping malls. SNOW [31] focuses on a certain type of group in shopping malls using WiFi data from associated devices rather than WiFi probes.

6 CONCLUSION

In this paper, we present a group detection system (BaG) that is verified reliable and accurate in different crowdedness settings. One of our main contributions is a new feature (NoB) extracted from WiFi probes. NoB could effectively capture phone usage patterns which is a new perspective for detecting groups. The second contribution is a new detection method (SCNMF) that fuses mobility information and usage behaviors. SCNMF could reveal the hidden associations between mobility and behaviors by decomposing them simultaneously. The experimental evaluation in a large shopping

mall demonstrates the effectiveness and robustness of BaG in scenarios with different levels of crowdedness. Compare to baseline approaches, BaG shows a significant improvement by increasing F-score of detection by 3.97% $\sim$ 15.79% in labeled dataset and 6.67% $\sim$ 20.69% in synthetically labeled dataset respectively.

ACKNOWLEDGMENTS

We sincerely appreciate the efforts of the anonymous reviewers and their insightful comments for improving this manuscript. The presented work was supported by National Key R&D Program of China (2018 YFB1004801). It was also partially supported by Shenzhen Basic Research Funding Scheme (JCYJ20170818104222072), and NSFC with project No. 61332004 and 61572218.

REFERENCES

[1] Marco V Barbera, Alessandro Epasto, Alessandro Mei, Vasile C Perta, and Julinda Stefa. [n. d.]. Signals from the crowd: uncovering social relationships through smartphone probes. In *Proceedings of ACM IMC, 2013.*

[2] Ningning Cheng, Prasant Mohapatra, Mathieu Cunche, Mohamed Ali Kaafar, Roksana Boreli, and Srikanth Krishnamurthy. [n. d.]. Inferring user relationship from hidden information in wlans. In *Proceedings of MILCOM, 2012.*

[3] Yohan Chon, Suyeon Kim, Seungwoo Lee, Dongwon Kim, Yungeun Kim, and Hojung Cha. [n. d.]. Sensing WiFi packets in the air: practicality and implications in urban mobility monitoring. In *Proceedings of ACM Ubicomp, 2014.*

[4] Varoth Chotpitayasunondh and Karen M Douglas. 2016. How "phubbing" becomes the norm: The antecedents and consequences of snubbing via smartphone. *Computers in Human Behavior* (2016).

[5] Mathieu Cunche, Mohamed Ali Kaafar, and Roksana Boreli. [n. d.]. I know who you will meet this evening! linking wireless devices using wi-fi probe requests. In *Proceedings of WoWMoM, 2012.*

[6] Marzieh Dashti, Mohd Amiruddin Abd Rahman, Hamed Mahmoudi, and Holger Claussen. [n. d.]. Detecting co-located mobile users. In *Proceedings of IEEE ICC, 2015.*

[7] Adriano Di Luzio, Alessandro Mei, and Julinda Stefa. [n. d.]. Mind your probes: De-anonymization of large crowds through smartphone WiFi probe requests. In *Proceedings of IEEE INFOCOM, 2016.*

[8] Julien Freudiger. [n. d.]. How talkative is your mobile device?: an experimental study of Wi-Fi probe requests. In *Proceedings of ACM WiSec, 2015.*

[9] Weina Ge, Robert T Collins, and R Barry Ruback. 2012. Vision-based analysis of small groups in pedestrian crowds. *IEEE TPAMI* (2012), 1003–1016.

[10] Nicolas Gillis and François Glineur. 2012. A multilevel approach for nonnegative matrix factorization. *J. Comput. Appl. Math.* (2012), 1708–1723.

[11] A Haigh. 2015. Stop phubbing.

[12] Hande Hong, Chengwen Luo, and Mun Choon Chan. [n. d.]. SocialProbe: Understanding Social Interaction Through Passive WiFi Monitoring. In *Proceedings of ACM MobiQuitous, 2016.*

[13] Timothy M Hospedales, Jian Li, Shaogang Gong, and Tao Xiang. 2011. Identifying rare and subtle behaviors: A weakly supervised joint topic model. *IEEE TPAMI* (2011).

[14] Xueheng Hu, Lixing Song, Dirk Van Bruggen, and Aaron Striegel. [n. d.]. Is there wifi yet?: How aggressive probe requests deteriorate energy and throughput. In *Proceedings of ACM IMC, 2015.*

[15] Shuja Jamil, Sohaib Khan, Anas Basalamah, and Ahmed Lbath. [n. d.]. Classifying smartphone screen ON/OFF state based on wifi probe patterns. In *Proceedings of ACM UbiComp, 2016.*

[16] Hyunsoo Kim and Haesun Park. 2008. Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method. *SIAM journal on matrix analysis and applications* (2008), 713–730.

[17] Jingu Kim and Haesun Park. 2008. *Sparse nonnegative matrix factorization for clustering.* Technical Report. Georgia Institute of Technology.

[18] Mikkel Baun Kjærgaard, Martin Wirz, Daniel Roggen, and Gerhard Tröster. [n. d.]. Detecting pedestrian flocks by fusion of multi-modal sensors in mobile phones. In *Proceedings of ACM UbiComp, 2012.*

[19] Mikkel Baun Kjærgaard, Martin Wirz, Daniel Roggen, and Gerhard Tröster. [n. d.]. Mobile sensing of pedestrian flocks in indoor environments using wifi signals. In *Proceedings of IEEE PerCom, 2012.*

[20] Youngki Lee, Chulhong Min, Chanyou Hwang, Jaeung Lee, Inseok Hwang, Younghyun Ju, Chungkuk Yoo, Miri Moon, Uichin Lee, and Junehwa Song. [n. d.]. Sociophone: Everyday face-to-face interaction monitoring platform using multi-phone sensor fusion. In *Proceedings of ACM MobiSys, 2013.*

[21] Jeremy Martin, Travis Mayberry, Collin Donahue, Lucas Foppe, Lamont Brown, Chadwick Riggins, Erik C Rye, and Dane Brown. 2017. A study of MAC address randomization in mobile devices and when it fails. *Proceedings on Privacy Enhancing Technologies* (2017).

[22] S Anne Moorhead, Diane E Hazlett, Laura Harrison, Jennifer K Carroll, Anthea Irwin, and Ciska Hoving. 2013. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of medical Internet research* (2013).

[23] Kazuaki Nakamura, Tsukasa Ono, and Noboru Babaguchi. [n. d.]. Detection of groups in crowd considering their activity state. In *Proceedings of IEEE ICPR, 2016.*

[24] Daniel Olgum Olgun, Peter A Gloor, and Alex Sandy Pentland. [n. d.]. Capturing individual and group behavior with wearable sensors. In *Proceedings of AAAI, 2009.*

[25] Daniel Olguín Olguín, Benjamin N Waber, Taemie Kim, Akshay Mohan, Koji Ara, and Alex Pentland. 2009. Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* (2009).

[26] Stefano Pellegrini, Andreas Ess, Konrad Schindler, and Luc Van Gool. 2009. You'll never walk alone: Modeling social behavior for multi-target tracking. In *Computer Vision, 2009 IEEE 12th International Conference on.* IEEE.

[27] Thor S Prentow, Antonio J Ruiz-Ruiz, Henrik Blunck, Allan Stisen, and Mikkel B Kjærgaard. 2015. Spatio-temporal facility utilization analysis from exhaustive wifi monitoring. *Pervasive and Mobile Computing* (2015).

[28] Lorenz Schauer, Martin Werner, and Philipp Marcus. [n. d.]. Estimating crowd densities and pedestrian flows using wi-fi and bluetooth. In *Proceedings of EAI MOBIQUITOUS, 2014.*

[29] Rijurekha Sen, Youngki Lee, Kasthuri Jayarajah, Archan Misra, and Rajesh Krishna Balan. [n. d.]. Grumon: Fast and accurate group monitoring for heterogeneous urban spaces. In *Proceedings of ACM SenSys, 2014.*

[30] Souvik Sen, Romit Roy Choudhury, and Srihari Nelakuditi. [n. d.]. SpinLoc: Spin once to know your location. In *Proceedings of Workshop on ACM MobiSys, 2012.*

[31] Jiaxing Shen, Jiannong Cao, Xuefeng Liu, and Shaojie Tang. 2018. SNOW: Detecting Shopping Groups Using WiFi. *IEEE Internet of Things Journal* (2018).

[32] Jiaxing Shen, Jiannong Cao, Xuefeng Liu, and Chisheng Zhang. 2017. DMAD: Data-Driven Measuring of Wi-Fi Access Point Deployment in Urban Spaces. *ACM Transactions on Intelligent Systems and Technology (TIST)* 9, 1 (2017), 11.

[33] Jiaxing Shen, Oren Lederman, Jiannong Cao, Florian Berg, Shaojie Tang, and Alex Pentland. 2018. GINA: Group Gender Identification Using Privacy-Sensitive Audio Data. In *2018 IEEE International Conference on Data Mining (ICDM).* IEEE, 457–466.

[34] Francesco Solera, Simone Calderara, and Rita Cucchiara. 2016. Socially constrained structural learning for groups detection in crowd. *IEEE TPAMI* (2016).

[35] Sandra Soto, Elva M Arredondo, Miguel T Villodas, John P Elder, Elena Quintanar, and Hala Madanat. 2016. Depression and chronic health conditions among latinos: The role of social networks. *Journal of immigrant and minority health* (2016), 1292–1300.

[36] Robert Tibshirani. 2011. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* (2011), 273–282.

[37] Geert Vanderhulst, Afra Mashhadi, Marzieh Dashti, and Fahim Kawsar. [n. d.]. Detecting human encounters from wifi radio signals. In *Proceedings of ACM MUM, 2015.*

[38] Mathy Vanhoef, Célestin Matte, Mathieu Cunche, Leonardo S Cardoso, and Frank Piessens. 2016. Why MAC address randomization is not enough: An analysis of Wi-Fi network discovery mechanisms. In *Proceedings of ACM ASIACCS, 2016.* ACM.

[39] T Wiesenthal, G Leduc, P Cazzola, W Schade, and J Köhler. 2011. Mapping innovation in the European transport sector. *An assessment of R&D efforts and priorities, institutional capacities, drivers and barriers to innovation. JRC Scientific and Technical Report* (2011).

[40] Wei Xi, Jizhong Zhao, Xiang-Yang Li, Kun Zhao, Shaojie Tang, Xue Liu, and Zhiping Jiang. [n. d.]. Electronic frog eye: Counting crowd using wifi. In *Proceedings of IEEE INFOCOM, 2014.*