

## Research Article

# ODMBP: Behavior Forwarding for Multiple Property Destinations in Mobile Social Networks

Jia Xu,<sup>1,2,3,4</sup> Jin Xin Xiang,<sup>1</sup> Xiang Chen,<sup>1</sup> Fang Bin Liu,<sup>4</sup> and Jing Jie Yu<sup>4</sup>

<sup>1</sup>School of Computer Science & Technology, Nanjing University of Posts and Telecommunications, Jiangsu, Nanjing 210003, China

<sup>2</sup>Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Jiangsu, Nanjing 210003, China

<sup>3</sup>Institute of Computer Technology, Nanjing University of Posts and Telecommunications, Jiangsu, Nanjing 210003, China

<sup>4</sup>Department of Information Technology, Nanjing General Hospital of Nanjing Military Command, Jiangsu, Nanjing 210002, China

Correspondence should be addressed to Jia Xu; xujia@njupt.edu.cn

Received 2 November 2015; Accepted 3 January 2016

Academic Editor: Tingting Chen

Copyright © 2016 Jia Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The smartphones are widely available in recent years. Wireless networks and personalized mobile devices are deeply integrated and embedded in our lives. The behavior based forwarding has become a new transmission paradigm for supporting many novel applications. However, the commodities, services, and individuals usually have multiple properties of their interests and behaviors. In this paper, we profile these multiple properties and propose an Opportunistic Dissemination Protocol based on Multiple Behavior Profile, ODMBP, in mobile social networks. We first map the interest space to the behavior space and extract the multiple behavior profiles from the behavior space. Then, we propose the correlation computing model based on the principle of BM25 to calculate the correlation metric of multiple behavior profiles. The correlation metric is used to forward the message to the users who are more similar to the target in our protocol. ODMBP consists of three stages: user initialization, gradient ascent, and group spread. Through extensive simulations, we demonstrate that the proposed multiple behavior profile and correlation computing model are correct and efficient. Compared to other classical routing protocols, ODMBP can significantly improve the performance in the aspect of delivery ratio, delay, and overhead ratio.

## 1. Introduction

In recent years, the smartphones have increased rapidly. According to the data from the International Data Corporation (IDC) Worldwide Quarterly Mobile Phone Tracker, the vendors shipped a total of 334.4 million smartphones worldwide in the first quarter of 2015 [1]. Wireless mobile networks are evolving and integrating with many aspects of our lives since we can read news, watch videos, listen to music, communicate with others, send and receive emails, browse and search the web, share contents to Internet and trade online, and so forth, through the smartphones conveniently. The wide spread smartphones promote the combination of the online social network and the mobile smart terminals, accelerating the development of the mobile social network (MSN) [2]. MSN involves the interactions between participants with similar interests and objectives through their mobile devices within virtual social networks.

Due to the dynamic and volatile nature of MSN, opportunistic networks operate under a completely new networking paradigm where traditional routing protocols cannot be applied [2]. Opportunistic networks are wireless mobile self-organizing networks in which the topology is extremely dynamic and unstable. Thus, in most cases, there might not exist the complete link from the source to the destination simultaneously. There have been many research efforts on opportunistic forwarding. However, most of them deliver the message based on IP or device address, which is not effective in many interest-aware or behavior-aware MSN applications.

The unprecedented tight coupling between mobile devices and their users provides new approaches to infer users' behavior and interest from mobile devices. The mobile devices can now act as distributed behavioral sensors of users to capture their interests and enable implicit interest profiling [3].

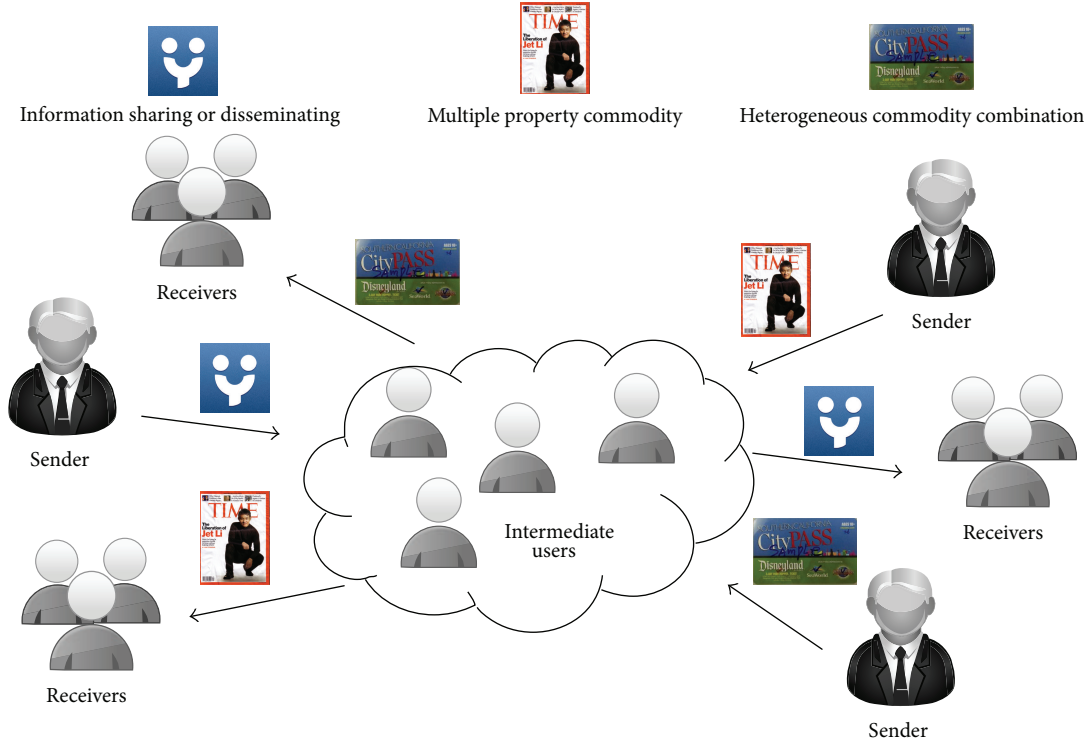


FIGURE 1: Multiple interests oriented opportunistic dissemination.

There are many popular location based applications in MSN. For instance, the location based services can help mobile users to find friends who are currently in their vicinity. Another example is Contact Recommendation Mechanism [4], which can efficiently select contacts in order to address them as a social group, so as to ease the initialization of group interactions.

The basic idea of these applications is extracting the interest profiles or relationship profiles from the social behavior. However, few research efforts consider the multiple behavior properties of MSN users. In fact, many MSN applications deal with the objects with multiple properties. Moreover, the human being has multiple interests naturally. There are some types of typical scenarios:

- (i) sharing or disseminating the information to the people with similar multiple interest profiles: as an example, Bob is a student, and he wants to find a roommate who is in the same university. He also hopes that the roommate likes swimming, just like himself. Now, Bob wants to push the message to the persons who have the great possibility to be the roommate;
- (ii) recommendation of commodities or services with multiple properties: for example, the editorial office wants to recommend a new magazine, which includes multiple topics, such as pop music, clothing, and bodybuilding. The editorial staff need to disseminate the advertisement to the potential readers who are interested in most of the topics;

- (iii) recommendation of the combination of heterogeneous commodities or services: for example, the merchant wants to publicize a discount combo of films and snacks and to send the information to the people who like both of them.

In the aforementioned scenarios, the message sender has a set of target interests, which can represent the commodities, services, or individuals with multiple properties. As shown in Figure 1, the message sender wants to send the message to the receivers who have the same or similar interests to the target interests.

In this paper, we focus on extracting multiple behavior properties from the daily traces of users and exploring the Opportunistic Dissemination Protocol based on the Multiple Behavior Profile in mobile social network. The key contributions of our work are summarized as follows.

- (i) We aim to deal with the data dissemination in a class of ubiquitous application scenarios, where the multiple properties of objects or the multiple interests of people need to be considered.
- (ii) We map the multiple properties or the multiple interests to the behavior space and profile the multiple behavior properties. Moreover, we propose the correlation computing model based on the principle of BM25 [5] for multiple behavior profiles.
- (iii) We design an Opportunistic Dissemination Protocol based on Multiple Behavior Profile (ODMBP) in mobile social networks.

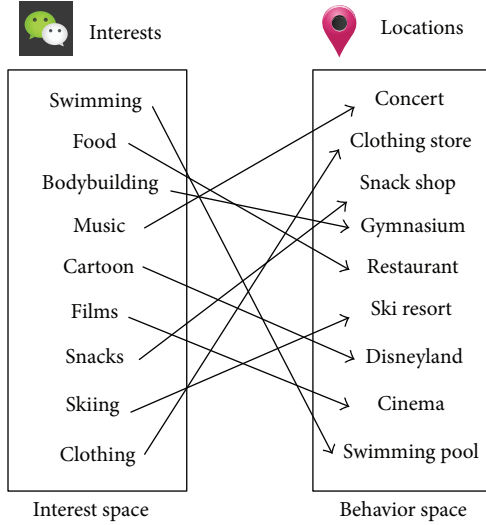


FIGURE 2: Mapping the interest space to the behavior space.

- (iv) The extensive simulations show that the proposed multiple behavior profiles and correlation computing model are correct and efficient. Compared to other classical routing protocols, ODMBP achieves high delivery ratio and low delay in the scenarios of multiple property data dissemination.

The remainder of this paper is organized as follows. Section 2 presents the challenges and the rationale of designed protocol. Section 3 introduces the multiple behavior profile and the correlation computing model. We present our opportunistic dissemination protocol in Section 4. The performance evaluation is presented in Section 5. We review the related work in Section 6. We conclude the paper in Section 7.

## 2. Challenges and Rationale

There are some challenges to design the opportunistic dissemination protocol for multiple property objects. First, we need to give a computable expression of interests, from which the multiple properties can be obtained. Second, for multiple properties, we need to consider the number of matched properties; that is, the designed protocol should try its best to match more properties in appointed properties. Thus, it might be inefficient to summate the similarity value of each property straightforwardly. Moreover, due to the energy limited devices and the intermittent link of opportunistic network, the designed protocol should meet the desired properties of distributed design, low computation complexity, low overhead, and high expandability.

As can be seen in Figure 2, many interests are closely related to the individual's daily trace and can be represented by the specific locations in the trace. It is shown in [6] that social relationships can explain about 10% to 30% of all human movements based on an analysis of different kinds of location datasets. On the other side, a large body of research has demonstrated that people show striking persistence in

their mobility profiles. For example, in [7], the authors state that the similarity of the mobility profile of a given user to its future profile is high, above 0.75 for eight days, and remains above 0.6 for five weeks. The observations demonstrate that the mobility profile is indeed an intrinsic property and a valid representation of the user, even if only a short history of mobility profile is used. Therefore, in this work, we assume that the locations can represent the user's interests; moreover, the longer the time in one location is, the stronger the corresponding interest is.

The basic idea of ODMBP is mapping the interest space to the behavior space and extracting the similarity between users' multiple behavior profiles and the target profiles. ODMBP uses the locations and the corresponding time spent at the locations to reflect users' preferences. The multiple behavior profile of each user is extracted from the quantized behavior space. Further, a reasonable correlation computing model should be applied to calculate the correlation metric. The designed opportunistic dissemination protocol then takes the correlation metric based forwarding strategy as the basic principle.

## 3. Multiple Behavior Profile and Correlation Computing Model

In this section, we will introduce the multiple behavior profile and the correlation computing model. The behavior profile should reflect multiple behavior properties, respectively, according to multiple interests. The correlation computing model should quantize the correlation for each user and should match as many behaviors as possible in the set of appointed behaviors.

**3.1. Multiple Behavior Profile.** Assume that there are a set  $U = \{1, 2, \dots, n\}$  of users and a set  $L = \{1, 2, \dots, m\}$  of locations, where  $n \geq 2$  and  $m \geq 2$ . Each location  $j$  in behavior space represents the corresponding interest. Each user  $i \in U$  has a user multiple behavior profile  $UMBP_i = (t_{i1}, t_{i2}, \dots, t_{im})$ , where  $t_{ij}$ ,  $j \in L$ , is the total time that user  $i$  spent at location  $j$ .

Note that  $t_{ij}$  is a cumulative time based on current trace of user  $i$ , and the value would be changed when time goes on. The time that user spent at the specific location can be measured through different ways. A widely used method is sensing the location information continuously through GPS sensors, which are integrated universally in the smartphones. Alternatively, the connection log of WiFi or switches in specific location can also help to obtain the time that the user spent. This work does not involve the specific persistent sensing, and the energy consumption can be very low.

The user multiple behavior profiles can be expressed as  $UMBP = (UMBP_1, UMBP_2, \dots, UMBP_n)$ . It can be viewed as  $n \times m$  behavior matrix with element  $t_{ij}$ . An example of UMBP is given in Figure 3. In most cases, the behavior matrix is a sparse matrix since most users only stay at a small fraction of all  $m$  locations. Thus, some specific data structures such as triple table can be used to reduce the space and time

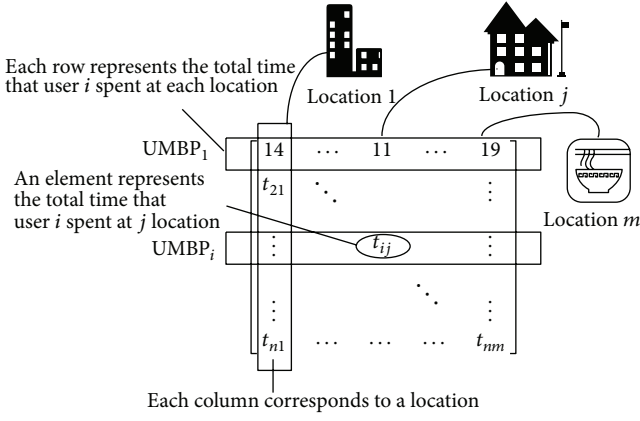


FIGURE 3: Illustration of user multiple behavior profiles.

complexity. Each element  $t_{ij}$  in UMBP is associated with a behavior indicator  $x_{ij}$ , where

$$x_{ij} = \begin{cases} 1, & \text{if } t_{ij} \neq 0 \\ 0, & \text{if } t_{ij} = 0. \end{cases} \quad (1)$$

There is a target multiple behavior profile  $\text{TMBP} = (tx_1, tx_2, \dots, tx_m)$  for each specific data dissemination application, where  $tx_j = 1$ ,  $1 \leq j \leq m$ , if the message sender hopes that the receivers have the behavior property associated with location  $j$ ; else,  $tx_j = 0$ . We assume that TMBP, which can be obtained through mapping the interests to the corresponding locations, is known in advance. We further denote the number of behavior properties in TMBP as  $r = \sum_{j=1}^m tx_j$ .

**3.2. Correlation Computing Model.** To find the potential receivers, we need a computing model to calculate the correlation between the target multiple behavior profile and user  $i$ 's multiple behavior profile,  $1 \leq i \leq n$ . The correlation is quantized by metric  $\text{Score}_i(\text{TMBP}, \text{UMBP})$ . We use the principle of the ranking function, named BM25 [8], to calculate this metric. BM25 uses the ideas of Robertson-Sparck-Jones (RSJ) probability model [9] and is a ranking function used by search engines to rank matching documents according to their relevance to a given search query. So far, it is the most successful model for calculating the correlation [10–12].

We first define the behavior factor of user  $i$  to location  $j$  in TMBP as

$$\text{BF}_{ij} = \frac{t_{ij}(K+1)}{t_{ij} + K}, \quad (2)$$

$$\text{s.t. } 1 \leq i \leq n, 1 \leq j \leq m, j = \arg \max_{1 \leq k \leq m} tx_k = 1,$$

where  $K$  is an empirical parameter, which represents the importance of behavior factor in  $\text{Score}_i(\text{TMBP}, \text{UMBP})$ .

The behavior factor  $\text{BF}_{ij}$  measures the total time user  $i$  spent at location  $j$  in TMBP.  $\text{BF}_{ij}$  provides a basic correlation evaluation. However, it might not meet the requirement of

matching as many behaviors as possible in TMBP since the behavior factor does not consider the distinction of user distribution at different locations. Actually, there might be some locations where few people stay in general. Thus, the behavior factor, which only considers the time the user spent at the location, might lose sight of these sparsely populated locations. To balance the sparsely populated locations, we introduce  $w_j$ , the weight of location  $j$ ,

$$w_j = \log \frac{n - q_j + 0.5}{q_j + 0.5}, \quad (3)$$

where  $q_j = \sum_{i=1}^n x_{ij}$  is the number of users at location  $j$ .

Note that the greater the value of  $q_j$  is, the smaller the value of  $w_j$  is. The weight of location  $j$  reflects the distinction of user distribution at different locations and can promote the importance of sparsely populated locations in TMBP.

Now we can give the ultimate formula to calculate the correlation metric:

$$\text{Score}_i(\text{TMBP}, \text{UMBP}) = \sum_{j=1}^m w_j \times \text{BF}_{ij}, \quad (4)$$

$$\text{s.t. } 1 \leq i \leq n, 1 \leq j \leq m, j = \arg \max_{1 \leq k \leq m} tx_k = 1.$$

Note that the Cosine similarity [13, 14] is another method to compute this metric as well. The Cosine similarity is widely used for computing the similarity of the text. It is not difficult to use Cosine similarity based on our user multiple behavior profiles. However, the Cosine similarity does not consider the distinction of user distribution at different locations. Further analyses and evaluation will be given in Section 5.

## 4. Opportunistic Dissemination Protocol Design

In this section, we attempt to design an opportunistic dissemination protocol for the services with multiple property objects. According to the principle of small world [15], people have high clustering property, and the users with similar behavior property have high probability of encounter. ODMBP disseminates the messages based on users' multiple behavior profiles and corresponding correlation metric.

The UMBP of each user will change with elapsed time, and it should be updated in distributed way. In ODMBP, each user  $i$  stores UMBP in his mobile device.  $\text{UMBP}_i$  can be updated by itself through position sensor or network connection log, while  $\text{UMBP}_j$ ,  $j \neq i$ , will be updated when user  $i$  encounters user  $j$ .

As shown in Algorithm 1, ODMBP consists of three stages: user initialization, gradient ascent, and group spread. In the user initialization stage, for each encountered user  $j$ , the message sender  $i$  matches the  $\text{UMBP}_j$  with the unmatched target multiple behavior profile  $\text{TMBP}'$ . If there is at least one matched location  $k$ , that is,  $t_{jk} = tx'_k = 1$ , the message sender  $i$  sends the message to user  $j$ . Once all locations in target multiple behavior profile are matched, that is,  $\text{TMBP}' = 0$ , the message sender  $i$  deletes the message. By this way, the user

```

(1)  $TMBP' \leftarrow TMBP$ ;
(2) foreach  $j$  encountered do
(3)   update UMBP for  $i$  and  $j$ ;
(4)   if  $i$  is a message sender then
(5)     if  $TMBP' \neq 0$  then // Stage 1: User Initialization
(6)       foreach  $k \in \{1, 2, \dots, m\}$  do
(7)         if  $t_{jk} \neq 0$  and  $tx'_k \neq 0$  then
(8)           send message to  $j$ ;
(9)            $tx'_k \leftarrow 0$ ;
(10)          break;
(11)     else
(12)       delete message in  $i$ ;
(13)     else if  $\delta > \text{Score}_j(TMBP, UMBP) > \text{Score}_i(TMBP, UMBP)$  then // Stage 2: Gradient Ascent
(14)       send message to  $j$ ;
(15)       delete message in  $i$ ;
(16)     else if  $\text{Score}_j(TMBP, UMBP) > \text{Score}_i(TMBP, UMBP) \geq \delta$  then // Stage 3: Group Spread
(17)       send message to  $j$ ;

```

ALGORITHM 1: ODMBP.

initialization stage can parallelize the dissemination process and decrease the delay efficiently.

Then, in the gradient ascent stage, the message holder forwards the message to the users with higher correlation score. The gradient ascent stage is derived from the fact that the multiple behavior profiles of the users with higher correlation score are more similar to the target receivers.

In the group spread stage, if the correlation score of the message holder is higher than threshold  $\delta$ , where  $\delta$  is a parameter of ODMBP, the message holder copies the message to the users with higher correlation score. This means ODMBP considers all user  $i$  satisfying  $\text{Score}_i(TMBP, UMBP) \geq \delta$  as the receivers.

## 5. Performance Evaluation

**5.1. Methodology and Settings.** In this section, we conduct thorough simulations to investigate the performance of ODMBP. We use the real trace dataset StudentLife [16], which contains the sensor data, EMA data, survey responses, and educational data. For our simulations, we adopted a part of this dataset, named Wifi-Location, which contains the data of 49 volunteers moved around 92 buildings in Dartmouth College within a month. The Wifi-Location, which contains nearly 0.192 million mobility records, acquires WiFi AP deployment information from Dartmouth Network Services and records participants' on-campus rough locations and unix time stamp. As an example, the record (1364359102, in (Kemeny)) indicates a volunteer moved in the building called Kemeny at the unix time 1364359102. The buildings can be seen as the locations in UMBPs and TMBP. We removed the interference items in the real movement trace such as the duplicate data and the invalid users. Figure 4 describes the number of locations of each user of the processed data and this number mostly falls in the interval [5, 100].

All the simulations were run on ONE simulator [17]; it is an opportunistic network environment simulator which provides a powerful tool for generating mobility traces, running

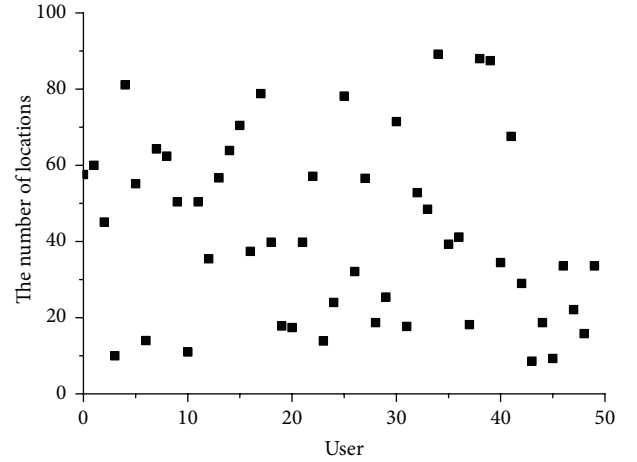


FIGURE 4: Distribution of location number.

DTN messaging simulations with different routing protocols. All the results are averaged over 1000 runs. The settings of the ONE simulator have been listed in Table 1. We first integrate the continuous records with the same location into a new record in order to compute their time difference. We also need to remove some interference items such as duplicate data and invalid user. We take this final output results as the external events connection data for the simulator. The number of hosts and the number of locations are 49 and 92, respectively, which are equal to the number of volunteers and buildings in Wifi-Location trace.

We use the time-location pairs to structure  $UMBP_i$  of any user  $i$ . As the simulator time goes on, the time spent in specific location can be obtained through calculating the elapsed time from the time stamp of user  $i$ 's current mobility record. By this way, we can obtain  $UMBP_i$  through accumulating all such elapsed time for each location. The UMBP is privacy information for each user and is calculated and updated dynamically with the simulator time. Moreover, the



TABLE 1: Simulation parameter setting.

Parameter	Value
Scenario name	ODMBP/ODMBP-Cos/EP/S & W
Number of hosts	49
Message time to live (Message TTL)	500 s
Simulation connection	False
Buffer size of nodes	1 M
Movement model warm-up time	1000 s
Simulation duration	960000 s
Number of locations	92
External events connection	True

users can connect with the users who are in the same location simultaneously. So, we can structure the external event connection data dynamically for the ONE simulator based on the above processing method. In our simulations, the behavior locations in TMBP are selected randomly among 92 buildings.

In our simulations, we first reveal the impacts of the key parameters on delivery ratio and delay of ODMBP. Moreover, we evaluate ODMBP further by comparing it with other protocols: Epidemic routing [18], Spray and Wait [19], and ODMBP-Cos.

To explore the differences between Cosine and BM25, we apply the Cosine similarity to our system model, and the correlation score function based on Cosine similarity is

$$\text{CosSim}_i(\text{TMBP}, \text{UMBP}) = \frac{x_{ij} \cdot tx_k}{\|x_{ij}\| \times \|tx_k\|}, \quad (5)$$

$$\text{s.t. } 1 \leq i \leq n, 1 \leq j \leq m, j = \arg_{1 \leq k \leq m} tx_k = 1,$$

where  $\cdot$  is the vector product and  $\|x\|$  is the Euclidean norm of  $x$ ; that is,  $\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$ .

We substitute the correlation score function with CosSim (TMBP, UMBP) in stage two and stage three of Algorithm 1, respectively. We call the protocol using Cosine similarity as ODMBP-Cos.

**5.2. Revealing the Impacts of the Key Parameters.** There are three key parameters: the empirical parameter  $K$ , the number of behavior properties in target multiple behavior profile  $r$ , and the threshold  $\delta$ . We will vary them for exploring the impacts of these parameters, respectively.

**5.2.1. Impact of  $K$ .** Based on our correlation computing model,  $K$  represents the importance of behavior factor in ultimate score metric. When using BM25 model in searching,  $K$  usually gets the value of 1.2 based on past experience. However, this setting might not be applicable in our multiple behavior dissemination scenarios. For the purpose of revealing the impact of  $K$  on ODMBP, we measure the delivery ratio and delay of ODMBP with different value of  $K$  when setting

$r = 5$ . As shown in Figure 5, ODMBP gets the best delivery ratio and delay when  $K = 1.75$ . Based on the observation, we fix  $K = 1.75$  in the following simulations. However, the setting of  $K$  may be closely related to the real dataset adopted.

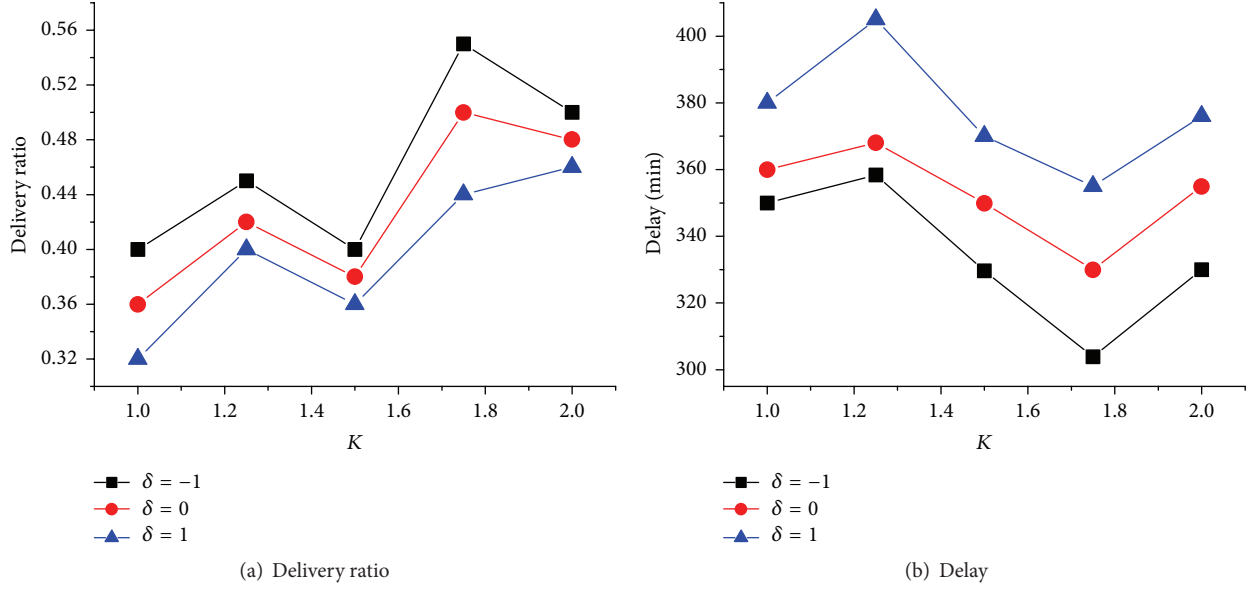
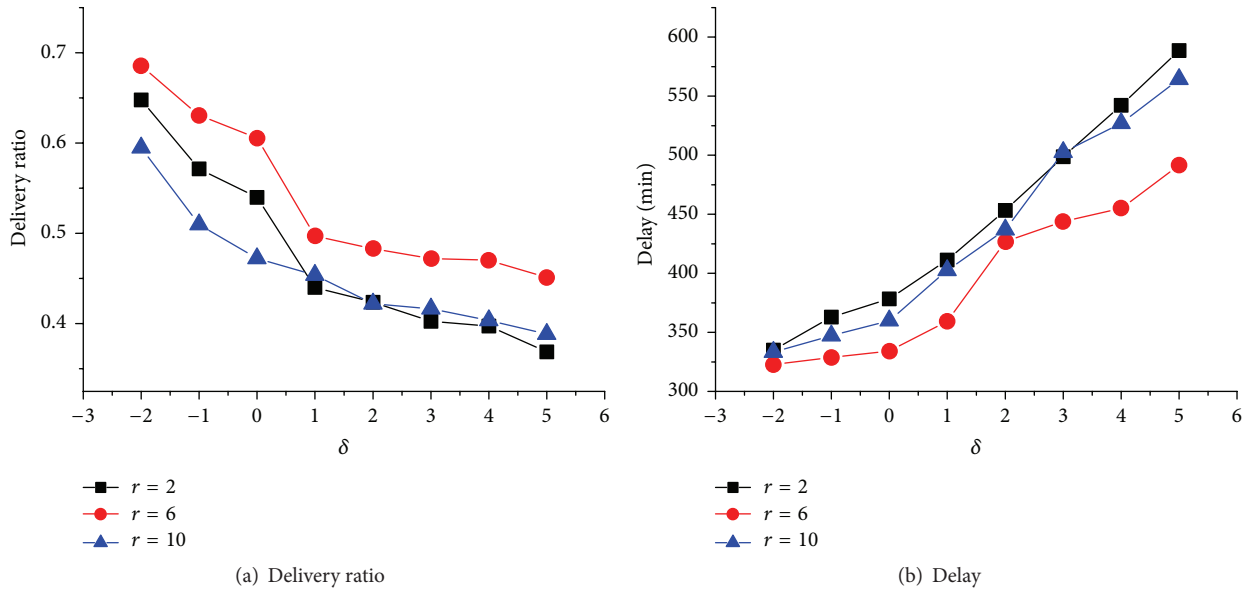
**5.2.2. Impact of  $\delta$ .** Threshold  $\delta$  is a criterion to judge whether the user is a receiver. It is also the trigger of ODMBP to enter the group spread stage. We measure the performance of ODMBP with different  $\delta$ . Figure 6 shows three groups of results corresponding to  $r = 2$ ,  $r = 6$ , and  $r = 10$ , respectively. When the value of  $\delta$  goes on, the delivery ratio decreases drastically for all settings of  $r$ . This is because the number of receivers reduces when the threshold increases. Accordingly, it takes more time to find the receivers, and the delay increases.

**5.2.3. Impact of  $r$ .** The number of behavior properties in target multiple behavior profile  $r$ , which is provided in advance, indicates the comprehensiveness of commodities/services or people's versatility. We cannot adjust the value of  $r$  to improve the performance of ODMBP; however, we can evaluate the scalability of designed protocol through the observation of the impact of  $r$  on ODMBP. We can see from Figure 7 that the curves of delivery ratio are not monotonous. Based on formula (4),  $\text{Score}_i(\text{TMBP}, \text{UMBP}) = \sum_{j=1}^m w_j \times \text{BF}_{ij}$ ; thus, the score is a summation value for all behavior locations in TMBP. Note that  $w_j = \log((n - q_j + 0.5)/(q_j + 0.5))$ , and the value of  $w_j$  will be negative if  $q_j > n/2$ . Thus, the value of score might decrease with great value of  $r$ . As a result, the number of receivers would reduce. As can be seen from Figure 6, ODMBP achieves the best performance in the aspect of delivery ratio and delay when  $r = 6$  among all measured  $r$  in our simulations.

**5.3. Compare with Other Protocols.** We compare ODMBP with other classical routing protocols, Epidemic routing and Spray and Wait in opportunistic network. In Epidemic routing protocol, the message is delivered to each encountered node that does not have the same message. The Spray and Wait routing protocol consists of two phases: Spray and Wait. The message copies are forwarded to  $l$  different nodes in the spray phase, and then the direct transmission is performed in the wait phase. We set  $l = 6$  and apply binary mode in the spray phase of Spray and Wait routing protocol. We set  $r = 6$ ,  $K = 1.75$ , and  $\delta = 0$  for ODMBP. We also compare the performance of ODMBP and ODMBP-Cos. We set  $r = 6$ ,  $\delta = 0.7$  in order to obtain the best performance of ODMBP-Cos. Such settings are based on the similar measures in Section 5.2.

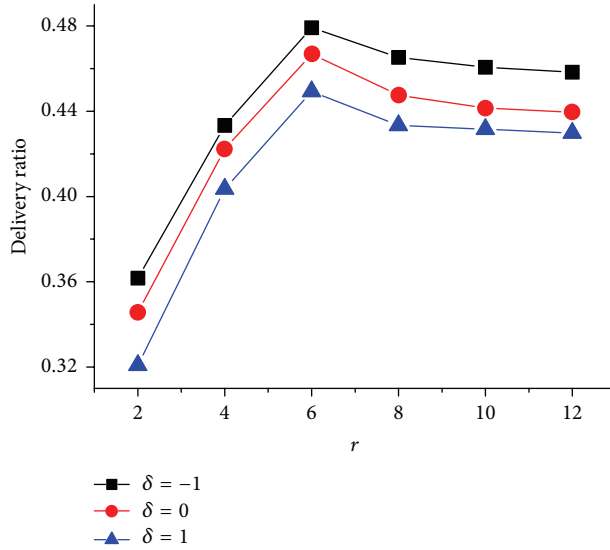
As shown in Figure 8, ODMBP has higher delivery ratio compared with ODMBP-Cos. This is because there are some locations where few people stay in general. Thus, the correlation function based on Cosine similarity, which only considers the time spent at the location, might lose sight of these sparsely populated locations. However, ODMBP can balance it well. On the other hand, the delay performance of two protocols is close.

The delivery ratio increases with increasing message TTL for all four protocols. This is because there is more time to

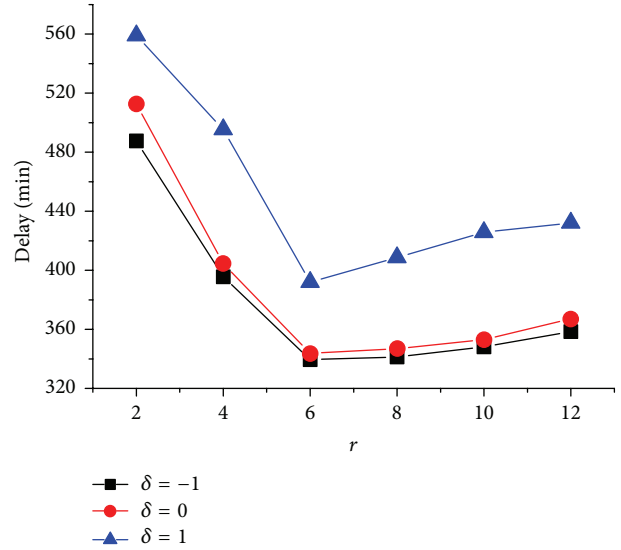
FIGURE 5: Impact of  $K$ .FIGURE 6: Impact of  $\delta$ .

deliver the message to the receivers before dropping it in the forwarding queue. However, the delay increases when the message TTL increases. Epidemic routing achieves the best performance among four protocols; however, it will suffer high overhead and is not efficient in our mobile social network applications. This is because Epidemic routing does not provide filtering scheme in the dissemination. The ONE simulator defines the parameter overhead ratio (number of

relayed messages – number of delivered messages)/number of delivered messages, while ODMBP has the threshold to filtrate the user with different correlations. Thus, ODMBP can reduce the amount of relayed messages. As shown in Figure 8, ODMBP has lower overhead ratio than Epidemic routing. In most cases, the performance of ODMBP is better than Spray and Wait, and ODMBP improves 11.6% and 12.5% in the aspect of delivery ratio and delay, respectively, on average.



(a) Delivery ratio



(b) Delay

FIGURE 7: Impact of  $r$ .

This is because ODMBP can forward the message to the users who are more similar to the target, while Spray and Wait does not consider the correlation metric.

## 6. Related Work

At present, there are many studies on exploring the behavior attributes of users in mobile social networks. In [7], Hsu et al. established a user behavior oriented communication model through the analysis of the participants' mobile data in university, demonstrating that the user has high stability in his mobile attribute. On this basis, they presented a protocol for the profile-cast service with high transmissibility and low delay, named CSI. However, only single behavior attribute is considered in their designed protocol. InterestCast [20], a novel communication protocol, also considers users' interests, solving the problem for a wide range of social scenarios and applying to an opportunistic network where nodes are the personal devices of moving individuals, possibly interacting with fixed road-side devices. In [21], Zhao et al. study a new coverage problem, opportunistic coverage, to characterize the sensing quality of such people-centric sensing systems. Compared with the traditional static coverage and dynamic coverage in sensor networks, opportunistic coverage has some unique characteristics caused by the requirements of urban sensing applications and human mobility features. The interest-aware implicit multicast (iCast) [22] is a new casting paradigm and works based on the inferred interest profiles. In this paradigm, messages are sent to a behavioral interest profile (not to an IP or device address). It combines user's interest and behavior for multicast communication. In [23], Elshierief et al. explore the notion of mobile users' similarity as a key enabler of innovative applications hinging

on opportunistic mobile encounters. SANE [24] combines the advantages of both social-aware and stateless approaches. It is based on the observation that individuals with similar interests tend to meet more often. In [25], Matsuo et al. propose an efficient boundary detection method in dense mobile wireless sensor networks. Each node preliminarily recognizes locations of itself and all its neighboring nodes. The authors determine the node forwarding direction by comparing the similarity score with the encounter nodes. Cheng et al. present iZone [26], a mobile social networking system based on the analysis of general requirements of MSN and location based services (LBS). The ultimate goal is developing and establishing an integrated framework for providing social network based healthcare information services targeting patient safety, empowerment, and guidance. Besides, [27] explains the interaction relationship of social network users and mutual influence and social network privacy behavior characteristics and motivation, including the prediction of user behavior as well.

## 7. Conclusion

We have extracted the multiple behavior profiles from the users' daily trace through mapping the multiple properties in the interest space to the behavior space. The BM25 based correlation computing model was proposed to calculate the correlation metric of multiple behavior profiles. Moreover, we have proposed an Opportunistic Dissemination Protocol based on Multiple Behavior Profile termed ODMBP, in mobile social networks. It consists of three stages: user initialization, gradient ascent, and group spread. Through extensive simulations, we have demonstrated that the proposed multiple behavior profiles and correlation computing model are



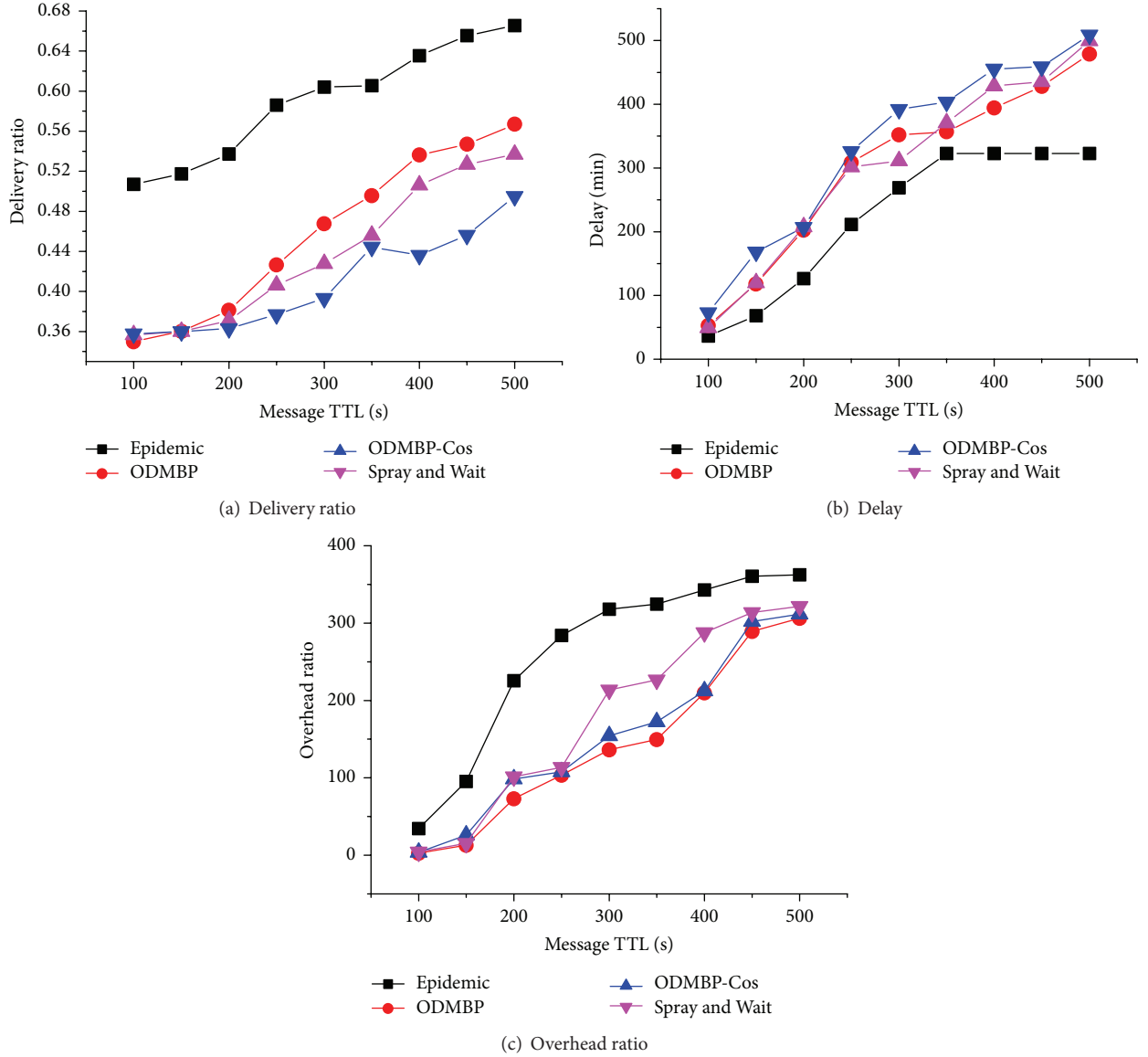


FIGURE 8: Compare ODMBP with other protocols.

efficient. Compared to the other classical routing protocols, ODMBP can significantly improve the performance in the aspect of delivery ratio and delay.

In the future work, we will consider more complex scenarios. For example, the behavior locations in the target multiple behavior profile can be associated with specific weights, which indicate the importance of the behavior locations.

## Notations

$U, n$ : Set of users and the number of users  
 $L, m$ : Set of locations and the number of locations

TMBP,  $r$ : Target multiple behavior profile and the number of behavior properties in target multiple behavior profile  
 UMBP,  $UMB P_i$ : User multiple behavior profiles and user  $i$ 's user multiple behavior profile  
 $x_{ij}$ : Behavior indicator of user  $i$  to location  $j$   
 $tx_j$ : Target behavior indicator of location  $j$   
 $w_j$ : Weight of location  $j$   
 $BF_{ij}$ : Behavior factor of user  $i$  to location  $j$

$t_{ij}$ :	Total time that user $i$ spent at location $j$
$q_j$ :	Total number of users at location $j$
$K$ :	A parameter used for computing the behavior factor
$\delta$ :	A threshold used for control of entering the group spread stage
$\text{Score}_i(\text{TMBP}, \text{UMBP})$ :	The correlation metric between the target multiple behavior profile and user $i$ 's multiple behavior profile.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work is sponsored in part by NSFC Grants (nos. 61472193, 61472192, and 61373139), The Natural Science Foundation of Jiangsu Province (nos. BK20141429, BK20130852), Scientific and Technological Support Project (Society) of Jiangsu Province (no. BE2013666), CCF-Tencent Open Research Fund (no. CCF-Tencent RAGR20150107), China Postdoctoral Science Foundation (no. 2014M562662, 2013T60553), Jiangsu Postdoctoral Science Foundation (no. 1402223C), Independent Research Project of Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks (no. WSNLBZY201524), NUPTSF (Grant no. NY215098), and the "1311" Talent Project of NJUPT.

## References

- [1] The worldwide smartphone market compared to the previous quarter's unprecedented results, 2015, <http://www.idc.com/prodserv/smartphone-market-share.jsp>.
- [2] X. Hu, T. H. S. Chu, V. C. M. Leung, E. C.-H. Ngai, P. Kruchten, and H. C. B. Chan, "A survey on mobile social networks: applications, platforms, system architectures, and future research directions," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1557–1581, 2014.
- [3] S. Moghaddam and A. Helmy, "Interest-aware implicit multicast (iCast): opportunistic mobile data dissemination without per-group management," in *Proceedings of the 9th ACM MobiCom Workshop on Mobility in the Evolving Internet Architecture (MobiArch '14)*, pp. 55–60, ACM, Maui, Hawaii, USA, September 2014.
- [4] R. Grob, M. Kuhn, R. Wattenhofer, and M. Wirz, "Cluestr: mobile social networking for enhanced group communication," in *Proceedings of ACM International Conference on Supporting Group Work (GROUP '09)*, pp. 81–90, Sanibel Island, Fla, USA, May 2009.
- [5] R. Blanco and P. Boldi, "Extending BM25 with multiple query operators," in *Proceedings of the 35th Annual ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '12)*, pp. 921–930, August 2012.
- [6] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '11)*, pp. 1082–1090, San Diego, Calif, USA, August 2011.
- [7] W.-J. Hsu, D. Dutta, and A. Helmy, "CSI: a paradigm for behavior-oriented profile-cast services in mobile networks," *Ad Hoc Networks*, vol. 10, no. 8, pp. 1586–1602, 2012.
- [8] S. E. Robertson and H. Zaragoza, "The probabilistic relevance framework: BM25 and beyond," *Foundations and Trends in Information Retrieval*, vol. 3, no. 4, pp. 333–389, 2009.
- [9] S. E. Robertson and K. S. Jones, "Relevance weighting of search terms," *Journal of the American Society for Information Science*, vol. 27, no. 3, pp. 129–146, 1976.
- [10] M. Murata, H. Nagano, R. Mukai, K. Kashino, and S. Satoh, "BM25 with exponential IDF for instance search," *IEEE Transactions on Multimedia*, vol. 16, no. 6, pp. 1690–1699, 2014.
- [11] S. E. Robertson, S. Walker, M. Hancock-Beaulieu, A. Gull, and M. Lau, "Okapi at TREC4," in *Proceedings of the 4th Text REtrieval Conference (TREC '96)*, pp. 21–30, Vienna, Va, USA, May 1996.
- [12] S. Robertson, "On the history of evaluation in IR," *Journal of Information Science*, vol. 34, no. 4, pp. 439–456, 2008.
- [13] H. Yildirim and M. S. Krishnamoorthy, "A random walk method for alleviating the sparsity problem in collaborative filtering," in *Proceedings of the ACM Conference on Recommender Systems (RecSys '08)*, pp. 131–138, ACM, Lausanne, Switzerland, October 2008.
- [14] C. Cattuto, D. Benz, A. Hotho, and G. Stumme, "Semantic analysis of tag similarity measures in collaborative tagging systems," in *Proceedings of the 3rd Workshop on Ontology Learning and Population (OLP3 '08)*, pp. 39–43, Patras, Greece, July 2002.
- [15] A. Banerjee, R. Agarwal, V. Gauthier, C. K. Yeo, H. Afifi, and F. B.-S. Lee, "A self-organization framework for wireless ad hoc networks as small worlds," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 6, pp. 2659–2673, 2012.
- [16] R. Wang, F. L. Chen, Z. Y. Chen et al., "StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proceedings of the ACM Conference on Ubiquitous Computing (UbiComp '14)*, pp. 3–14, Seattle, Wash, USA, September 2014.
- [17] F. Ekman, A. Keränen, J. Karvo, and J. Ott, "Working day movement model," in *Proceedings of the 1st ACM SIGMOBILE Workshop on Mobility Models*, pp. 33–40, ACM, Hong Kong, May 2008.
- [18] A. Vahdat and D. Becker, *Epidemic routing for partially-connected ad hoc networks [M.S. thesis]*, 2000.
- [19] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *Proceedings of the ACM SIGCOMM Workshop on Delay-Tolerant Networking (WDTN '05)*, pp. 252–259, ACM, August 2005.
- [20] E. Pagani, L. Valerio, and G. P. Rossi, "Weak social ties improve content delivery in behavior-aware opportunistic networks," *Ad Hoc Networks*, vol. 25, pp. 314–329, 2015.
- [21] D. Zhao, H. D. Ma, and L. Liu, "Energy-efficient opportunistic coverage for people-centric urban sensing," *Wireless Networks*, vol. 20, no. 6, pp. 1461–1476, 2014.
- [22] S. Moghaddam and A. Helmy, "Interest-aware implicit multicast (iCast): opportunistic mobile data dissemination without per-group management," in *Proceedings of the 9th ACM Workshop*

- on Mobility in the Evolving Internet Architecture (MobiArch '14)*, pp. 55–60, ACM, Maui, Hawaii, USA, September 2014.
- [23] M. Elshierief, T. Elbatt, A. Zahran, and A. Helmy, “The quest for user similarity in mobile societies,” in *Proceedings of the IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS '14)*, pp. 569–574, Budapest, Hungary, March 2014.
  - [24] A. Mei, G. Morabito, P. Santi, and J. Stefa, “Social-aware stateless routing in pocket switched networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 1, pp. 252–261, 2014.
  - [25] K. Matsuo, K. Goto, A. Kanzaki, T. Hara, and S. Nishio, “Overhearing-based efficient boundary detection in dense mobile wireless sensor networks,” in *Proceedings of the 15th IEEE International Conference on Mobile Data Management (IEEE MDM '14)*, pp. 225–234, IEEE, Queensland, Australia, July 2014.
  - [26] R. Cheng, Z. Yang, and F. Xia, “IZone: a location-based mobile social networking system,” in *Proceedings of the 3rd International Symposium on Parallel Architectures, Algorithms and Programming (PAAP '10)*, pp. 33–38, Dalian, China, December 2010.
  - [27] A. Acquisti, L. Brandimarte, and G. Loewenstein, “Privacy and human behavior in the age of information,” *Science*, vol. 347, no. 6221, pp. 509–514, 2015.

