Social-Tie-Based Information Dissemination in Mobile Opportunistic Social Networks

Yunsheng Wang and Jie Wu
Department of Computer and Information Sciences
Temple University, Philadelphia, PA 19122

***Abstract*—A mobile opportunistic social network (MOSN) is a new type of delay tolerant network (DTN), in which the mobile users contact each other opportunistically. Information dissemination is a challenging problem in MOSNs, due to uncertainty and intermittent connectivity. In this paper, we propose a distributed social tie strength calculation mechanism to identify the relationship between each set of pairwise mobile nodes. Following arguments originally proposed by Mark Granovetter's seminal 1973 paper, *The Strength of Weak Ties*, the majority of the novel information dissemination is generated by weak ties. We first evaluate the strength of weak ties in *MIT reality mining data*. Then, a social-tie-based information dissemination protocol is presented, which is a token-based information dissemination scheme, including two phases: *weak tie-driven forwarding* and *strong tie-driven forwarding*. In the weak tie-driven forwarding phase, the susceptible nodes with more weak ties will receive more tokens for future forwarding. The number of forwarding tokens is related to the number of weak ties of two encountered nodes. After a while, the information will have been spread to multiple communities. Our scheme switches to a strong tie driven forwarding phase, in which the influential nodes are more important. The number of forwarding tokens is proportional to the number of strong ties of two encountered nodes. Extensive simulations are conducted in comparison to several approaches in real world mobile traces.**

***Index Terms*—Influential, information dissemination, local bridge, mobile opportunistic social networks (MOSNs), social tie strength, susceptible.**

I. Introduction

Social influence is empirically elusive in the social sciences. Scholars from different fields as diverse as business, computer science, physics, and sociology are interested in who influences whom, how to efficiently disseminate information, and how to prevent virus contagion. The answers to these questions, which are critical to policies, depend on the robustness of estimations of the degree to which contagion is at work during the social information influence [1].

Delay tolerant networks (DTNs) [2] are characterized by intermittent connectivity and limited network capacity, in which most of the time there does not exist an end-to-end path between some or all of the nodes in the network. With the popularization of smart phones, mobile opportunistic social networks (MOSNs), a new type of DTN, have recently become popular. In MOSNs, the individuals carrying smart phones walk around and communicate with each other via Bluetooth or WiFi, when they are in each other's transmission range. Because of the short contact duration and intermittent connectivity, designing an efficient information dissemination policy in such a network environment becomes a challenging problem.

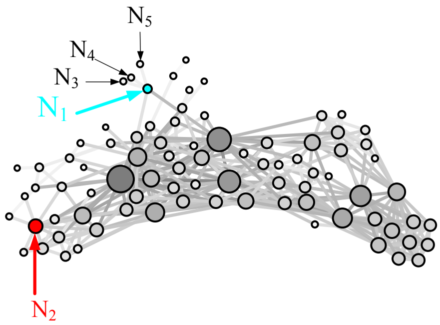

Fig. 1. An illustration of the Facebook friends network: the nodes are the Facebook users. The size of the nodes show the degree of the nodes. Larger nodes have larger degrees. The thickness of the links show the social tie strength between pairwise nodes. Thicker links have larger tie strength.

One of the powerful roles that networks play is to bridge the local and global, which guide the information flows through a social network. The strength of the weak ties hypothesis from sociology [3], illustrates the importance of weak ties in information dissemination. The tie strength has been modeled in many online social network researches [4–6]. Fig. 1 illustrates a Facebook friends network. N_2 has a larger degree than N_1, while N_1 has more weak ties than N_2. Without N_1, three individuals N_3, N_4, and N_5 cannot receive the information, while the neighbors of N_2 do not have this problem. Therefore, N_1 is considered more important than N_2 in information dissemination. Weak ties play an important role in information dissemination. However, these studies need the global information of whole social networks, which is not suitable in MOSNs. In this paper, we propose a distributed tie strength measurement mechanism. Every mobile node maintains a *tie strength table*, which records the social tie relationship with its encountered nodes.

One particularly controversial argument is the "influentials" hypothesis that influential individuals (with many strong ties) catalyze the information dissemination in society [7]. Despite this popular argument, a variety of researches suggest that susceptibility (with many weak ties) is the key trait that drives the diffusion of novel social information [8–10]. In this paper, we make use of the social tie table to identify the *influential* and *susceptible* members in MOSNs, as to enhance the efficiency of information dissemination.

In [11], Easley and Kleinberg claimed that *If a node A in*

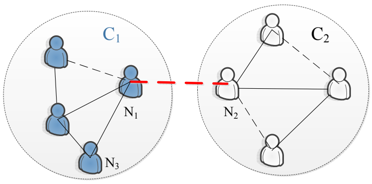

Fig. 2. An illustration of the local bridge linking two communities: *solid* lines are the *strong* ties and dashed lines are the weak ties.

a network ratifies the Strong Triadic Closure Property[1] *and is involved in at least two strong ties, then any local bridge*[2] *it is involved in must be a weak tie*. As shown in Fig. 2, if the link between N_1 and N_2 is a strong tie, following the Strong Triadic Closure, there must be a link between N_2 and N_3; however, the definition of the local bridge says it cannot. If the local bridge is a strong tie, then all other links connected to the endpoints of the local bridge must be weak ties based on the Strong Triadic Closure Property. Hence, in a large network, the local bridges are more likely to be weak ties. If a node has many weak ties, the likelihood of one being a local bridge is relatively high. Therefore, to search for a local bridge, which can connect multiple communities, the focus should be on searching the susceptible nodes with many weak ties.

In this paper, we consider the information dissemination problem as a token-based broadcast routing problem. In order to reduce the cost, only the mobile nodes with the tokens can forward the message to the encountered nodes. The routing process we proposed has two phases: *weak tie-driven forwarding* and *strong tie-driven forwarding*.

In the weak tie-driven forwarding phase, we propose to spread the message to more susceptible nodes, which is like an *inter-community* information spread. Hence, forwarding more tokens to the nodes with more weak ties can increase the spread speed. Therefore, at the beginning of information dissemination, the number of tokens assigned to the relay node is proportional to the number of weak ties of the two encountered nodes.

After a while, the information has been propagated to multiple communities. Hence, the strong ties will play a more important role, which means influential individuals (with many strong ties) can disseminate the information to many individuals in a short period. Therefore, the next stage is a strong-tie driven forwarding, which is like an *intra-community* information dissemination.

More specifically, the key contributions of our work can be summarized as follows:

- We illustrate a novel distributed social tie strength calculation mechanism, which can simply calculate the social tie strength based on the contact and social feature information in a cost-effective way.
- We propose a two-phase social-tie-based information dissemination algorithm, which is a token-based message forwarding scheme, according to the tie strength of the nodes with its neighbors.
- We present the design, implementation, and evaluation for quantitatively measuring the performance of the social-tie-based information dissemination scheme in real world mobile traces. The effectiveness of our approach is verified through extensive simulations.

[1]If a node A has links to nodes B and C, then the $B-C$ link is especially likely to form if A's links to B and C are both strong ties [11].

[2]A link joining two nodes in a graph is a *local bridge*, if its endpoints have no friends in common [11].

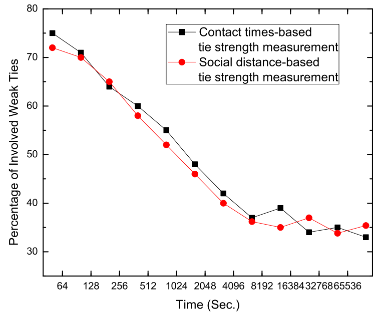

Fig. 3. The percentage of involved weak ties in MIT reality mining data.

The remainder of this paper is organized as follows: we review the related work in Section II. We then present the details of our approach in Section III. The simulation and evaluation are shown in Section IV. Finally, we conclude our work in Section V.

II. Related Work

Mobile opportunistic social networks (MOSNs), a new type of DTNs, become more and more interesting, according to the widespread use of smart phones. Researchers study MOSNs from a social networking point of view [12–16]. Most of them are considering routing problems in MOSNs. They use the mobile users' social features [13], their community properties [14, 15], as well as the mobility of the mobile users [16] to enhance the routing protocols. In this paper, we use the contact information and social features of each individual to measure the strength of their social ties.

There has been some work on data dissemination in DTNs [17–20]. In [17], Ning et al. proposed a credit-based incentive-aware data dissemination scheme in DTN, which effectively tracks the value of a message, which highly depends on its probability to be delivered by an intermediate node. Gao and Cao proposed a user-centric data dissemination in [20]. Their approach was based on a social centrality metric, which considers the social contact patterns and interests of mobile users simultaneously, and thus ensures effective relay selection. In this paper, we utilize the strength of weak ties for data dissemination.

III. Social-Tie-Based Information Dissemination

In this section, we first introduce the motivation of our work. The datasets we used will be presented next. Then, we discuss our proposed tie strength calculation mechanism and two phase token-based information dissemination protocol.

A. Motivation

Mark Granovetter's seminal 1973 paper [3], *The Strength of Weak Ties*, demonstrated that weak ties play a key role in the novel information dissemination. We verify the strength of

TABLE I
SOCIAL FEATURES IN DATASETS.

Infocom Feature	MIT Feature
Affiliation	*Neighborhood*
City	*Daily commute*
Nationality	*Hangouts*
Language	*Working hour*
Country	*Affiliation*
Position	*Research group*

weak ties in real mobile datasets. Fig. 3 shows the percentage of involved weak ties in MIT reality mining data by using message flooding. We use the number of contacts or social feature distance to identify the weak and strong ties, which we will discuss in detail in Section III. We find that, at the early stage, more weak ties are involved in the information dissemination than the strong ties. Later on, strong ties become the dominant factor.

B. Datasets

For our work, we exploit two datasets – MIT reality mining data [21] and Infocom 2006 conference trace [22]. These two datasets include *activity-based* and *survey-based* data. The activity-based data includes the contact information between pairwise nodes. We built an activity-based network whereby participants act as mobile nodes, and the number of contacts between two nodes act as contact information for tie strength calculation. In the MIT reality mining data, the contact information is recorded by the *call logs* by phone, and the *proximity data* by Bluetooth. The Infocom 2006 trace only includes the proximity data. The data from a survey provides self-reported personality, which we consider to be the social features of the participants.

C. Tie Strength Calculation

In the MOSNs, there are many factors that affect the tie strength, such as the number of contacts, contact frequencies, contact durations, last contact time, social distance, and so on, between two encountered nodes.

Definition: *The strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie* [3].

In this paper, we use two factors to measure the tie strength: *contact information*, which is from the activity-based data; and *social information*, which is from survey-based data. For the contact information, we use the number of contacts between pairwise nodes as the measurement metric. For the social information, we extract the social features from two real datasets: MIT reality mining data and Infocom2006 trace, as shown in Table I. Each node has a social feature vector, which indicates its characteristic in the social features. We use a 2-dimensional social feature vector as an example. Dimension 1 corresponds to *city* with three distinct values: New York (0), London (1), and Paris (2); and dimension 2 corresponds to *gender*, with two distinct values: male (0) and female (1). A user with social feature vector $[0, 1]$ represents a female working in New York. If two nodes have exactly the same value in one dimension, we assume their distance

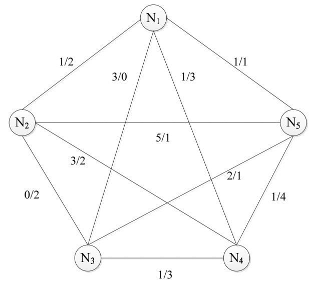

Fig. 4. An illustration of a five-nodes social network. The value (x/y) on the links represent the number of contacts/the social feature distance of linked vertices.

in this dimension is 0. Otherwise, their distance is 1. The social feature distance is the summation of the distances in all dimensions [13]. For example, a female working in New York with social feature vector $[0, 1]$, has social feature distance 1 to a female working in Paris with social feature vector $[2, 1]$, and social feature distance 2 to a male working in London with social feature vector $[1, 0]$. We use the social feature distance as another metric to measure the tie strength.

In this paper, we model tie strength as a linear combination of the number of contacts and social feature distance:

$$w_{ij} = \alpha C_{ij} + \beta \frac{1}{1 + D_{ij}}, \tag{1}$$

where w_{ij} represents the tie strength of nodes N_i and N_j. C_{ij} ($C_{ij} \in [0, 1]$) represents the normalized number of contacts between these two nodes, which includes the call logs and proximity data. D_{ij} ($D_{ij} = 0, 1, 2, 3, ...$) represents the social feature distance between these two nodes, which is used to measure the closeness between two nodes. α and β represent the impact of contact information and social feature information, respectively. At the same time, $\alpha + \beta = 1$.

In the learning process, each node calculates the tie strength with the encountered nodes, and creates a weighted adjacency matrix. Then, the definition of strong and weak ties was established as follows: following [23, 24], from the weighted adjacency matrix, we used as a threshold the 59th percentile of the link weights cumulative distribution; then, links weighted higher than or equal to the threshold were considered as strong ties, while links with a weight less than the threshold were marked as weak ties.

For example, Fig. 4 illustrates a five-nodes social network, $G(V, E)$, where each link shows the condition of the linked pairwise nodes. As shown in Fig. 4, x/y labeled on each link represents *the number of contacts/the social feature distance*. In this example, each node has 5 social features. Here, we assume α and β in Eq. 1 are the same, with value $\frac{1}{2}$. Therefore, we can create a 5×5 weighted symmetric adjacency matrix according to Eq. 1:

$$\begin{bmatrix} 0 & \frac{4}{15} & \frac{4}{5} & \frac{9}{40} & \frac{7}{20} \\ \frac{4}{15} & 0 & \frac{1}{6} & \frac{7}{15} & \frac{3}{4} \\ \frac{4}{5} & \frac{1}{6} & 0 & \frac{9}{40} & \frac{9}{20} \\ \frac{9}{20} & \frac{7}{15} & \frac{9}{40} & 0 & \frac{1}{5} \\ \frac{7}{20} & \frac{3}{4} & \frac{9}{20} & \frac{1}{5} & 0 \end{bmatrix} \tag{2}$$

Algorithm 1 Weak Tie-Driven Forwarding

/$*$ When a message holder N_i with c tokens meets node N_j without the message. $*$/
/$*$ Forward the message to N_j. $*$/
if $W_j > W_i$ **then**
Give $\left\lceil \frac{W_j}{W_i+W_j} c \right\rceil$ number of tokens to N_j

Based on this adjacency matrix, we can calculate the threshold ($(\frac{4}{5} - \frac{1}{6}) \times 59\% \approx 0.374$). Finally, we can distinguish the strong and weak ties. In this example, the links (N_1, N_3), (N_2, N_4), (N_2, N_5), and (N_3, N_5) are strong ties, while the links (N_1, N_2), (N_1, N_4), (N_1, N_5), (N_2, N_3), (N_3, N_4), and (N_4, N_5) are weak ties.

Then, each node can maintain its social tie table, recording the relationship to its encountered nodes. In each contact, the encountered nodes will exchange their social tie tables.

D. Two-Phase Token-based Message Forwarding

1) Influential and susceptible nodes: Local bridges linking multiple communities can disseminate the information among these communities. In a large social network, the local bridges are more likely to be the weak ties [11]. The *susceptible* nodes with many weak ties have high probability, located on the local bridges. A variety of literatures have claimed that susceptible nodes play a key role in novel information dissemination. In this paper, we propose a token-based message forwarding with two phases: *weak tie-driven forwarding* and *strong tie-driven forwarding*. The mobile nodes with the message tokens can forward the message to the encountered nodes. In the early stage, susceptible nodes will receive more tokens, which we call 'weak tie-driven forwarding'. After a while, the message has been spread to multiple communities. It is more important that *influential* nodes, with more strong ties, deliver the message to the group members. Therefore, we change to a strong tie-driven forwarding scheme, which forwards more tokens to influential nodes.

2) Weak tie-driven forwarding: The individuals with many weak ties are considered to be the susceptible members in the MOSNs, who do not cluster in the network, while the influential individuals do [1]. A node with many weak ties has a relatively high probability of locating on a local bridge, which links different communities; therefore, susceptible members play a key role in novel information dissemination. Here, we propose a weak tie-driven forwarding algorithm in the early stage of information dissemination. When a message holder N_i, with c number of tokens, encounters a node N_j without that message, it forwards the message to N_j. If N_j has more weak ties than N_i, N_i will give $\left\lceil \frac{W_j}{W_i+W_j} c \right\rceil$ number of tokens to N_j, where W_i is the number of weak ties of N_i. Algorithm 1 shows the whole process of weak tie-driven forwarding.

3) Strong tie-driven forwarding: The influential individuals, with many strong ties, cluster in the network and can influence the other infected nodes with a high probability in the local communities. When the information has been spread to multiple communities, our forwarding strategy will turn to a strong tie-driven forwarding.

TABLE II
THE PERFORMANCE OF DIFFERENT TIE STRENGTH CALCULATION SCHEMES IN MIT REALITY MINING TRACE

	$\alpha = 0.5$ $\beta = 0.5$	$\alpha = 0.75$ $\beta = 0.25$	$\alpha = 0.25$ $\beta = 0.75$	$\alpha = 1$ $\beta = 0$	$\alpha = 0$ $\beta = 1$
Delivery ratio	81.3%	83%	80.8%	79.3%	76.5%
Latency (Sec.)	328k	310k	323k	327k	343k

TABLE III
THE PERFORMANCE OF DIFFERENT TIE STRENGTH CALCULATION SCHEMES IN INFOCOM2006 TRACE

	$\alpha = 0.5$ $\beta = 0.5$	$\alpha = 0.75$ $\beta = 0.25$	$\alpha = 0.25$ $\beta = 0.75$	$\alpha = 1$ $\beta = 0$	$\alpha = 0$ $\beta = 1$
Delivery ratio	89.5%	89.1%	91%	87.8%	88.3%
Latency (Sec.)	40.3k	41.2k	40k	44.3k	42.5k

There are many factors that affect the two phases' switch time. However, considering the properties of MOSNs, every node does not have the global information, such as how many nodes in the network have received the new information. Therefore, the token holders can only use the number of tokens they have to estimate the information dissemination situation.

When the number of tokens held by the mobile nodes is below a predefined threshold, our information dissemination process switches to the second phase: strong tie-driven forwarding. The strong tie-driven forwarding phase is also a token-based broadcast process. When a message holder N_i, with c number of tokens, encounters a node N_j without a message, it forwards the message to N_j. The number of tokens assigned to N_j is $\left\lceil \frac{S_j}{S_i+S_j} c \right\rceil$, if N_j has more strong ties than N_i. Here, S_i is the number of strong ties of N_i.

IV. Simulation and Evaluation

We evaluate the performance of the proposed two-phase token-based message forwarding scheme (**TTF**) through trace-driven simulations.

A. Simulation Setting and Comparison Scheme

In all experiments, the first half of the trace is used for the learning process, which is for the accumulation of the network information, the process of collecting the contact information, and calculating the tie strength combined with the social feature information. After the first half learning process, we can create an adjacency matrix, then distinguish the weak ties and strong ties following the method we discussed in Section III-C. The new information generation and information dissemination happens during the second half of both traces.

In the simulation, we compare our proposed scheme with the following information dissemination schemes in MOSNs:

1) **Flooding (F)**, in which the message holder will forward half of the token to the encountered node without the message.

2) **Weak tie-driven forwarding (WF)**, which is the same as the first phase of our proposed two-phase token-based scheme.

3) **Strong tie-driven forwarding (SF)**, which is the same as the second phase of our proposed two-phase token-based scheme.

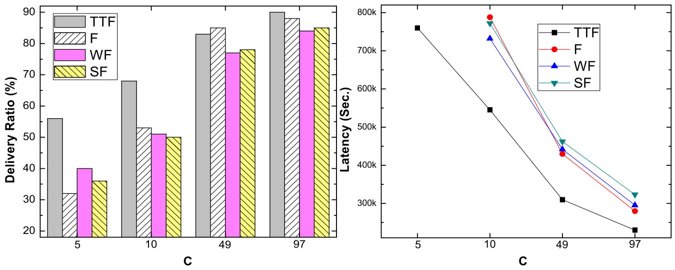

Fig. 5. Comparing the performance of different schemes in different initial numbers of tokens in the MIT reality mining trace: (L): delivery ratio; (R): latency.

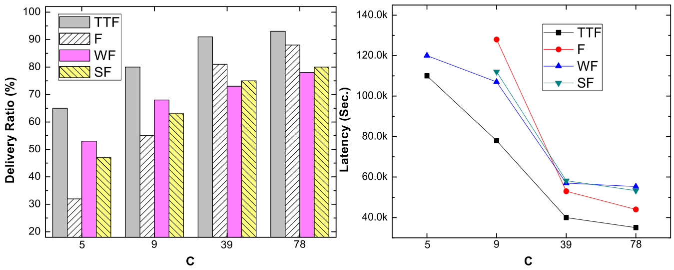

Fig. 6. Comparing the performance of different schemes in different initial numbers of tokens in the Infocom2006 trace: (L): delivery ratio; (R): latency.

B. Different Schemes for Tie Strength Calculation

First, we compare the performance of different tie strength calculation schemes. As we discussed in Section III-C, we will compare the performance in different values of α and β in Eq. 1. We set the initial number of tokens (C) created by the source to $\frac{n}{2}$, where n is the number of of nodes in the whole network. When the number of tokens held by the message holder is below a threshold $\lceil \log_2 C \rceil$, the token forwarding strategy will switch from the weak tie-driven forwarding phase to a strong tie-driven forwarding phase.

As shown in Tables II and III, we can see that when one set of the information is excluded from learning process ($\alpha = 1,\ \beta = 0$ or $\alpha = 0,\ \beta = 1$), the delivery ratio will reduce, and latency will increase. In the MIT trace, contact information can predict the tie strength more accurately, especially when $\alpha = 0.75$ and $\beta = 0.25$, in Table II. Therefore, in the rest of the simulation, we set α to 0.75 and β to 0.25 in the MIT trace. In the Infocom2006 trace, the social feature information is more important than the contact information. When $\alpha = 0.25,\ \beta = 0.75$, the performance is best among all schemes, as shown in Table III. Therefore, in the rest of the simulation, we set α to 0.25 and β to 0.75 in the Infocom2006 trace.

C. Performance in Limited Initial Tokens

Then, we evaluate different protocols in different initial numbers of tokens (C). The initial number of tokens created by the source is $\left\lceil \frac{\sqrt{n}}{2} \right\rceil$, $\lceil \sqrt{n} \rceil$, $\frac{n}{2}$, and n. When the number of tokens held by the message holder is below a threshold $\lceil \log_2 C \rceil$, in **TTF**, the token forwarding strategy will switch from the weak tie-driven forwarding phase to the strong tie-driven forwarding phase. Here, we compare the delivery ratio in different settings, and latency when the delivery ratio reaches 50%.

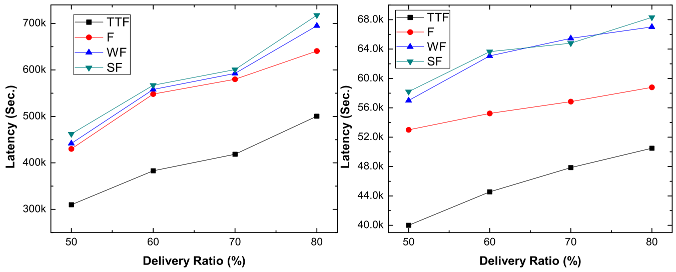

Fig. 7. Comparing the performance of different schemes in different delivery ratios: (L): MIT; (R): Infocom.

From Figs. 5 and 6, we can see that our proposed two-phase token-based information dissemination scheme has a much higher delivery ratio compared with the other three forwarding schemes, especially when the initial number of tokens is smaller. The simulation results indicate the robustness of our proposed scheme, and it performs much better than other schemes in limited resource conditions (here, we consider the initial number of tokens as the resource).

Since some schemes can not achieve 50% delivery ratio, when the initial number of tokens is $\left\lceil \frac{\sqrt{n}}{2} \right\rceil$ in both MIT reality mining and Infocom2006 traces, we consider that the latency of these schemes in this condition is infinite. In Figs. 5 and 6, we find that our scheme can dramatically reduce the latency in all conditions.

D. Performance in Different Delivery Ratios

Here, we compare the latency of different protocols in different delivery ratios. We set the delivery ratio to 50%, 60%, 70%, and 80%. The initial number of tokens created by the source is $\frac{n}{2}$ in this part. The two phases' switch threshold is also $\lceil \log_2 C \rceil = \lceil \log_2 n/2 \rceil$ in **TTF**.

Fig. 7 shows the latency comparison of different schemes in different delivery ratio constraints. Compared to weak tie-driven forwarding and strong tie-driven forwarding, our proposed two-phase token-based message forwarding scheme reduces the latency by about 15% in the MIT reality mining trace, and 21% in the Infocom2006 trace in Fig. 7. Our scheme performance is much better in the lower delivery ratio condition, compared to the flooding scheme. This means that our scheme can spread the new information quickly at the beginning of information dissemination, while flooding scheme may waste the contacts among the friend nodes, who contact each other in high frequency.

E. Impact of The Two Phases' Switch Threshold

We also compare the performance of our proposed two-phase token-based message forwarding in different switch thresholds. The comparison switch thresholds are set as $\left\lceil \frac{\log_2 C}{2} \right\rceil$, $\lceil \log_2 C \rceil$, and $\lceil 2 \times \log_2 C \rceil$, respectively. The initial number of tokens created by the source is set to $\frac{n}{2}$. Here, we compare the delivery ratio in different settings, and latency when the delivery ratio reaches 50%.

From Tables IV and V, we can see that in both MIT and Infocom2006 traces, the performance varies with the change of the two phases' switch threshold. When the threshold is

TABLE IV
THE PERFORMANCE OF OUR SCHEME IN DIFFERENT TWO PHASES' SWITCH THRESHOLDS IN MIT REALITY MINING TRACE

Threshold	$\lceil \frac{\log_2 C}{2} \rceil$	$\lceil \log_2 C \rceil$	$\lceil 2 \times \log_2 C \rceil$
Delivery ratio	69%	83%	75%
Latency (Sec.)	363k	310k	345k

TABLE V
THE PERFORMANCE OF OUR SCHEME IN DIFFERENT TWO PHASES' SWITCH THRESHOLDS IN INFOCOM2006 TRACE

Threshold	$\lceil \frac{\log_2 C}{2} \rceil$	$\lceil \log_2 C \rceil$	$\lceil 2 \times \log_2 C \rceil$
Delivery ratio	85%	91%	82%
Latency (Sec.)	42.7k	40k	44.9k

too small, our scheme will work as weak tie-driven forwarding the majority of the time. When the threshold is too large, our scheme will switch to strong tie-driven forwarding quickly. In both situations, the delivery ratio will decrease, while the latency increases. Therefore, choosing an accurate switch threshold is very important for the performance of our proposed two-phase token-based message forwarding scheme.

F. Summary of Simulation

In this section, we evaluate our proposed two-phase token-based information dissemination scheme in two real world mobile traces: the MIT reality mining campus trace, and the Infocom2006 conference trace. The simulation results show that our proposed scheme has better performance in both delivery ratio and latency in different network environments, compared to other schemes. When the resources are limited, which means the source node generates a small amount of tokens, our approach performs even better, increasing the delivery ratio and decreasing latency more dramatically than other approaches. The results in various delivery ratio scenarios indicate that our scheme can spread the novel information to certain fractions of users in the network quickly. By changing the value of the two phases' switch threshold, we find that our scheme performs best when the switch threshold is equal to $\lceil \log_2 C \rceil$, where C is the initial number of tokens generated by the source node.

V. CONCLUSION

In this paper, we present a social-tie-based information dissemination scheme in MOSNs. We leverage the strength of weak ties and susceptible nodes in novel information dissemination for token split guidance. We design a tie strength calculation mechanism to distinguish the weak and strong ties, which considers both the contact and social feature information. Then, a two-phase token-based message forwarding algorithm is introduced and evaluated in different network environments in real world mobile traces. The simulation results verify the effectiveness of our proposed approach. Our future work will include more experiments on different social network traces to validate the effectiveness of our approach. We also plan to develop a mobile phone application to exploit the social network properties in large scale mobile social networks.

REFERENCES

[1] S. Aral and D. Walker, "Identifying social influence in networks using randomized experiments," *IEEE Intelligent Systems*, vol. 26, no. 5, pp. 91–96, Sep. 2011.

[2] K. Fall, "A delay-tolerant network architecture for challenged internets," in *Proc. of the ACM conference on Applications, technologies, architectures, and protocols for computer communications(SIGCOMM)*, 2003, pp. 27–34.

[3] M. Granovetter, "The Strength of Weak Ties," *The American Journal of Sociology*, vol. 78, pp. 1360–1380, 1973.

[4] L. A. Adamic and E. Adar, "Friends and neighbors on the web," *Social Networks*, vol. 25, pp. 211–230, 2003.

[5] E. Sun, I. Rosenn, C. Marlow, and T. M. Lento, "Gesundheit! modeling contagion through facebook news feed," in *Proc. of the 3rd International Conference on Weblogs and Social Media (ICWSM)*, 2009.

[6] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic, "The role of social networks in information diffusion," in *Proc. of the ACM 21st international conference on World Wide Web(WWW)*, 2012, pp. 519–528.

[7] T. W. Valente, *Network models of the diffusion of innovations*. Hampton Press, Cresskill, N.J., 1995.

[8] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proc. of the 9th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2003, pp. 137–146.

[9] D. J. Watts and P. S. Dodds., "Influentials, networks, and public opinion formation," *Journal of Consumer Research*, vol. 34, pp. 441–458, 2007.

[10] D. Centola and M. W. Macy, "Complex contagions and the weakness of long ties," *American Journal of Sociology*, vol. 113, pp. 702–734, 2007.

[11] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.

[12] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective," in *Proc. of ACM MobiHoc*, 2009, pp. 299–308.

[13] J. Wu and Y. Wang, "Social feature-based multi-path routing in delay tolerant networks," in *Proc. of IEEE INFOCOM*, 2012, pp. 1368 –1376.

[14] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: Social-based forwarding in delay-tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 11, pp. 1576 –1589, 2011.

[15] A.-K. Pietilänen and C. Diot, "Dissemination in opportunistic social networks: the role of temporal communities," in *Proc. of ACM MobiHoc*, 2012, pp. 165–174.

[16] A. Ribeiro, R. Sofia, and A. Zuquete, "Improving mobile networks based on social mobility modeling," in *Proc. of IEEE ICNP*, 2011, pp. 289 –291.

[17] T. Ning, Z. Yang, X. Xie, and H. Wu, "Incentive-aware data dissemination in delay-tolerant mobile networks," in *Proc. of 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, 2011, pp. 539 –547.

[18] G. Karlsson, V. Lenders, and M. May, "Delay-tolerant broadcasting," in *Proc. of the ACM SIGCOMM workshop on Challenged networks (CHANTS)*, 2006, pp. 197–204.

[19] Y. Wang, Y. Guo, and J. Wu, "Making many people happy: Greedy solutions for content distribution," in *Proc. of IEEE International Conference on Parallel Processing (ICPP)*, 2011, pp. 693 –702.

[20] W. Gao and G. Cao, "User-centric data dissemination in disruption tolerant networks," in *Proc. IEEE INFOCOM*, 2011, pp. 3119 –3127.

[21] A. P. N. Eagle and D. Lazer, "Inferring social network structure using mobile phone data," in *Proc. of the National Academy of Sciences*, vol. 106(36), 2009, pp. 15,274–15,278.

[22] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD trace cambridge/haggle/imote/infocom2006 (v. 2009-05-29)," 2009.

[23] Y. Kalish and G. Robins, "Psychological predispositions and network structure: The relationship between individual predispositions, structural holes and network closure," *Social Networks*, vol. 28, no. 1, pp. 56 –84, 2006.

[24] J. Staiano, B. Lepri, N. Aharony, F. Pianesi, N. Sebe, and A. Pentland, "Friends don't lie: inferring personality traits from social network structure," in *Proc. of the 2012 ACM Conference on Ubiquitous Computing (UbiComp)*, 2012, pp. 321–330.