Social-Aware Forwarding Improves Routing Performance in Pocket Switched Networks*

Josep Díaz[1], Alberto Marchetti-Spaccamela[2], Dieter Mitsche[1], Paolo Santi[3] and Julinda Stefa[2]

[1] Univ. Politècnica de Catalunya, Barcelona, Spain
[2] Sapienza Università di Roma, Italy
[3] Istituto di Informatica e Telematica del CNR, Pisa, Italy

Abstract. We study and characterize social-aware forwarding protocols in opportunistic networks and we derive bounds on the expected message delivery time for two different routing protocols, which are representatives of social-oblivious and social-aware forwarding. In particular, we consider a recently introduced stateless, social-aware forwarding protocol using interest similarity between individuals, and the well-known BinarySW protocol, which is optimal within a certain class of stateless, social-oblivious forwarding protocols. We compare both from the theoretical and experimental point of view the asymptotic performance of Interest-Based (IB) forwarding and BinarySW under two mobility scenarios, modeling situations in which pairwise meeting rates between nodes are either *independent of* or *correlated to* the similarity of their interests.

1 Introduction

Opportunistic networks, in which occasional communication opportunities between pairs or small groups of nodes are exploited to circulate messages, are expected to play a major role in next generation short range wireless networks [17,18,19]. In particular, pocket-switched networks (PSNs) [9], in which network nodes are individuals carrying around smart devices with direct wireless communication links, are expected to become widespread in a few years. Message exchange in opportunistic networks is ruled by the *store-carry*-and-*forward* mechanism typical of delay-tolerant networks [6]: a node (either the sender, or a relay node) stores the message in its buffer and carries it around, until a communication opportunity with another node arises, upon which the message can be forwarded to another node (the destination, or another relay node).

Given this basic forwarding mechanism, a great deal of attention has been devoted in past years to optimize the forwarding policy of routing protocols. Recently, several authors have proposed optimizing forwarding strategies for PSNs

* The three first authors were partially supported by the EU through project FRONTS. The work of P. Santi was partially supported by MIUR, program PRIN, Project COGENT.

C. Demetrescu and M.M. Halldórsson (Eds.): ESA 2011, LNCS 6942, pp. 723–735, 2011.

based on the observation that, being these networks composed of *individuals* characterized by a collection of social relationships, these social relationships can actually be reflected in the meeting patterns between network nodes. Thus, knowledge of the social structure underlying the collection of individuals forming a PSN can be exploited to optimize the routing strategy, e.g., favoring message forwarding towards "socially well connected" nodes. Significant performance improvement of social-aware approaches over social-oblivious approaches has been experimentally demonstrated [3,8,11].

Most existing social-aware forwarding approaches hinge on the ability of *storing* state information that can be used to keep trace of history of past encounters and/or to attempt to predict future meeting opportunities [1,2,3,8,10,11]. On the other hand, socially-oblivious routing protocols such as epidemic [20], two-hops [7] and the class of Spray-and-Wait protocols [18], do not require storing additional information in the node buffers, which are then exclusively used to store the messages circulating in the network. Thus, comparing performance of social-aware vs. social-oblivious forwarding approaches would require modeling node buffers, which renders the resulting network model very complex. If storage capacity on the nodes is not accounted for in the analysis, unfair advantage would be given to social-aware approaches, which extensively use state information.

In [14], a *stateless* approach has been presented; this approach is motivated by the observation that individuals with similar interests meet relatively more often than individuals with diverse interests. The definition of this Interest-Based forwarding approach (IB forwarding in the following) allows a *fair* comparison – i.e., under the same conditions for what concerns usage of storage resources – between social-aware and social-oblivious forwarding approaches in PSNs.

Our contributions. The main goal of this paper is to compare IB and BinarySW forwarding. BinarySW [18] is chosen as a representative element of the class of social-oblivious forwarding protocols, since it is shown in [18] to be optimal within the class of Spray-and-Wait forwarding protocols, and given the extensive simulation-based evidences of its superiority within the class of stateless, social-oblivious approaches. To the best of our knowledge, the notion of interest-based mobility – although empirically verified in [14,16] – has never been formalized in the literature.

Namely, main contributions of this paper are:

1. An asymptotic analysis of IB and BinarySW forwarding performance in two different scenarios: *interest-based* mobility and *social-oblivious* mobility. The first scenario models the situation where node mobility is highly correlated to similarity of individual interests, while the second one models the opposite situation in which node mobility is independent of individual interests. We consider the case when only one relay node can be used to speed up message delivery and we prove, under reasonable probabilistic assumptions, that *IB forwarding provides asymptotic performance benefits compared to BinarySW*: IB forwarding yields *bounded* expected message delivery time and BinarySW yields *unbounded* expected message delivery time. The result that IB forwarding is better than BinarySW forwarding when nodes meet according to an interest based model that favors encounters among similar people might not be surprising. However we remark that such result was not formally proved before; we also observe that while one has constant expected time the other one is unbounded.

2. We confirm the analysis of **1.** through simulations based both on a real-world data trace and a synthetic human mobility model recently introduced in [13].

3. We extend the analysis of **1.** in several ways. First, we consider the case when many relay nodes, more copies of the message, and more hops can be used to speed up message delivery. We show that the expected delivery time of BinarySW is asymptotically the same. We also consider a version of the forwarding algorithm in which the sender knows the ID of the destination, but it does not know its *interest profile* (see next section for a formal definition of interest profile). We show that expected message delivery time with IB forwarding remains bounded even in this more challenging networking scenario if we allow a limited number of relay nodes.

A byproduct of the above analysis is the definition of a simple model of pairwise contact frequency correlating similarity of individual interests with their meeting rate. We believe this model might be useful in studying other social-related properties of PSNs, and we deem such model a contribution in itself.

Due to length constraints, proofs are not reported, and are presented in the full version of the paper [5].

2 The Network Model

We consider a network of $n+2$ nodes, which we denote $\mathcal{N} = \{S, D, R_1, \ldots, R_n\}$: a *source node* S, a *destination node* D, and n potential *relay nodes* $R_1, \ldots, R_n$. Following the model presented in [14], we model each of the $n+2$ nodes as a point in an m-dimensional *interest space* $[0,1]^m$, where m is the total number of interests and $m \ll n$. We assume $m = \Theta(1)$. The m-dimensional vector associated with a node defines its *interest profile*, i.e., its degrees of interest in the various dimensions of the interest space. Each node $A \in \mathcal{N}$ is thus assigned an m-dimensional vector $A[a_1, \ldots, a_m]$ in the interest space. As in [14], we use the well-known *cosine similarity* metric [4], which measures similarity between two nodes A and B as $\cos(\angle(AB))$, the cosine of the angle formed by A the origin and B. Since the cosine similarity metric implies that the norm of the vectors is not relevant, we can consider all vectors normalized to have unit norm.

We assume S and D to have orthogonal interests, namely $S[1, 0, \ldots, 0]$, and $D[0, 1, \ldots, 0]$. We call this scenario the *worst-case delivery scenario* since it corresponds to the worst case situation under the interest-based mobility model. Furthermore, in the analysis below, we assume the following concerning the distribution of interest profiles in the interest space: first, the angle α_i between the i-th interest profile and S's interest profile is chosen uniformly at random in $[0, \pi/2]$; then, from all unit vectors in the intersection of the positive orthant of the m-dimensional sphere with that $(m-1)$-dimensional subspace, one vector is chosen uniformly at random.

It is important to observe that, while nodes are assumed to move around according to some mobility model $\mathcal{M}$, node coordinates in the interest space *do not change over time*. This is coherent with what happens in real world, where individual interests change at a much larger time scale (months/years) than needed to exchange messages within the network. Thus, when focusing on a single message delivery session, it is reasonable to assume that node interest profiles correspond to fixed points in the interest space.

In particular, we assume that the mobility metric relevant to our purposes is the *expected meeting time*, which is formally defined as follows:

Definition 1. *Let A and B be nodes in the network, moving in a bounded region R according to a mobility model $\mathcal{M}$. Assume that at time $t = 0$ both A and B are independently distributed in R according to the stationary distribution of $\mathcal{M}$, and that A and B have a fixed transmission range. The* first meeting time *T between A and B is the random variable (r.v.) corresponding to the time interval elapsing between $t = 0$ and the instant of time where A and B first come into each other transmission range. The* expected meeting time *is the expected value of r.v. T.*

Following the literature [17,18,19], we assume that $T_{AB} \sim \exp(\lambda_{AB})$, therefore $\mathbb{E}[T_{AB}] = \frac{1}{\lambda_{AB}}$ (i.e. the meeting time between any pair of nodes A and B is described by a Poisson point process of intensity λ_{AB}).

In the sequel we consider two mobility models and two routing algorithms. The mobility models *social oblivious* and *interest based* are defined as follows:

– *social oblivious mobility*: for any $A, B \in \mathcal{N}$, the meeting rate $\lambda_{AB} = \lambda$ for some $\lambda > 0$ independent on A and B. This corresponds to the situation in which node mobility is not influenced by the social relationships between A and B, and it is the standard model used in opportunistic network analysis [17,18,19].

– *interest-based mobility*: the rate λ_{AB} is defined as $\lambda_{AB} = k \cdot \cos(\alpha_{AB}) + \delta(n)$. Note that the cos term implies higher correlation between nodes with more similar interests while the $\delta(n) > 0$ term accounts for the fact that occasional meetings can occur also between perfect strangers; we are interested in the case $\delta(n) \to 0$ as $n \to \infty$, which corresponds to the fact that as n grows, the probability of a meeting by chance decreases. Finally, $k > 0$ is a parameter modeling the intensity of the interest-based mobility component.

We characterize the performance of *routing algorithms*, i.e. the dynamics related to delivery of a message M from S, to D. We use S, D, or R_i to denote both a node, and its coordinates in the interest space. The dynamics of message delivery is governed by a routing protocol, which determines how many copies of M shall circulate in the network, and the forwarding rules. Namely, the following two routing algorithms are considered:

– *FirstMeeting* (FM) [18]: S generate two copies of M; S always keeps a copy of M for itself. Let R_j be the first node met by S amongst nodes $\{R_1, \ldots, R_n\}$. If R_j is met before node D, the second copy of M is delivered to node R_j. From this point on, no new copy of the message can be created nor transferred to other nodes, and M is delivered to D when the first node among S and R_j gets in touch with D. If node D is met by S before any of the R_i's, M is delivered directly. This protocol is equivalent to Binary SW as defined in [18].

– *InterestBased* [14]: IB(γ) routing is similar to FM, the only difference being that the second copy of M is delivered by S to the first node $R_k \in \{R_1, \ldots, R_n\}$ met by S such that $cos(R_k, D) \geq \gamma$, where $\gamma \in [0, 1]$ is a tunable parameter. Note that IB(0) is equivalent to FM routing. If it happens that after time n still no node in $\{R_1, \ldots, R_n\}$ satisfying the forwarding condition is encountered, then the first relay node meeting S after time n is given the copy of M independently of similarity between interest profiles.

We remark that IB routing is a stateless approach: interest profiles of encountered nodes are stored only for the time needed to locally compute the similarity metrics, and discarded afterwards. Although stateless, IB routing requires storing a limited amount of extra information in the node's memory besides messages: the node's interest profile, and the interest profile of the destination for each stored message. However, note that interest profiles can be compactly represented using a number of bits which is independent of the number n of network nodes, the additional amount of storage requested on the nodes is $O(1)$. On the other hand, stateful approaches such as [1,2,3,8,10,11] require storing an amount of information which is at least proportional to the number of nodes in the network, i.e., it is $O(n)$ (in some cases it is even $O(n^2)$). Thus, comparing IB routing with a socially-oblivious routing protocol can be considered fair (in an asymptotic sense) from the viewpoint of storage capacity. Based on this observation, in the following we will make the standard assumption that node buffers have unlimited capacity [17,18,19], which contributes to simplifying the analysis.

We denote by T_X^{μ} the random variable corresponding to the time at which M is first delivered to D, assuming a routing protocol $X \in \{\text{FM,IB}(\gamma)\}$, where *so* and *ib* represent social-oblivious and interest-based mobility, respectively, under mobility model $\mu \in \{so, ib\}$.

For both algorithms and both mobility models we consider the following random variables: T_1 is the r.v. counting the time it takes for S to meet the first node in the set $\mathcal{R} = \mathcal{N} \setminus \{S\}$; T_2 is 0 if D is the first node in $\mathcal{R}$ met by S; otherwise, if R_j is the relay node, then, T_2 is the r.v. counting the time, starting at T_1, until the first out of S and R_j meets D.

3 Bounds on the Expected Delivery Time

The following proposition states that in the social oblivious mobility scenario $\mathbb{E}[T_{FM}^{so}]$ and $\mathbb{E}[T_{IB(\gamma)}^{so}]$ are asympotically equal. We state the result of IB routing assuming $\gamma := \frac{0.29}{m-1}$ (the extension to other values of $\gamma \in (0, 1)$ is omitted).

Proposition 1. $\mathbb{E}[T_{FM}^{so}] = \mathbb{E}[T_{IB(\gamma)}^{so}] = \frac{1}{2\lambda}(1 + o(1))$ *where* $\gamma := \frac{0.29}{m-1}$.

We now consider the interest base mobility model. Consider first the case when FM routing is used in presence of interest-based mobility. The difficulty in performing the analysis stems from the fact that, under interest-based mobility, the rate parameters of the exponential r.v. representing the first meeting time between S and the nodes in $\mathcal{R}$ are r.v. themselves.

Denote by α_i the r.v. representing $\angle(S, R_i)$ in the interest space, and let $\lambda_i = k\cos\alpha_i + \delta$ be the r.v. corresponding to the meeting rate between S and R_i. Notice that the probability density function for any α_i to have value $x \in [0, \pi/2]$ is $2/\pi$. To compare results for the two mobility cases, we first compute $\mathbb{E}[\lambda_i]$, and set k in such a way that $\mathbb{E}[\lambda_i] = \lambda$. We have

$$\mathbb{E}[\lambda_i] = \int_0^{\pi/2} \frac{2}{\pi}(k\cos(\alpha) + \delta)d\alpha = \frac{2k}{\pi} + \delta, \tag{1}$$

and thus $k = \frac{\pi}{2}(\lambda - \delta)$. To compute $\mathbb{E}[T_1]$ exactly, we have to consider an n-fold integral taking into account all possible positions of the nodes $R_1, \ldots, R_n$ in the interest space. We will see that T_1 is asymptotically negligible compared with T_2, and hence we can use the trivial upper bound $\mathbb{E}[T_1] \leq \frac{1}{n\delta}$. Computing $\mathbb{E}[T_2]$ exactly also seems difficult. The following theorem gives a lower bound.

Theorem 1. $\mathbb{E}[T^{ib}_{FM}] \geq \min\{\Omega(n/\log n), \Omega(\log(1/\delta))\}$.

The theorem implies that if $\delta = \delta(n) = o(1)$ then $\mathbb{E}[T^{ib}_{FM}] \to \infty$. Theorem 2 analyses IB($\gamma$) routing in presence of interest-based mobility.

Theorem 2. *For some constant* $c > 0$ *and any* $0 < \gamma < 1$, $\mathbb{E}[T^{ib}_{IB(\gamma)}] \leq m\gamma/c$.

Theorems 1 and 2 establishes asymptotic superiority of IB(γ) over FM routing in case of interest-based mobility.

3.1 Extensions

More copies and more hops. We now discuss how to extend the analysis for interest based mobility to $\ell > 2$ hops and $q > 2$ copies. We consider a variation of the FM routing protocol for interest-based mobility, which we call FM*: we assume that the message M is forwarded from node A to node B only if the interest profile of node B is *more similar* to the destination node than that of node A. If a node has already forwarded M to a set of nodes, then it will forward M only to nodes which are closer to the destination than all the previous ones. We have that, $T^{so}_{FM^*} \leq T^{so}_{FM}$, since the first one at least partially accounts for similarity of interest profiles when forwarding messages. Note that the difference between FM* and IB routing is that, while in IB a minimum similarity threshold with D must be satisfied to forward M, in FM* even a tiny improvement of similarity is enough to forward M.

Observe that upper bounds on the asymptotic performance provided by IB routing remain valid also for $\ell, q > 2$. We now show that, even allowing more copies and/or hops and the smarter FM* forwarding strategy, $\mathbb{E}[T^{ib}_{FM^*}]$ do not improve asymptotically, with respect to $\ell = 2$ and $q = 2$.

Consider the case of FM* routing with $\ell \geq 2$ ($\ell = \Theta(1)$) hops and exactly 2 copies of M. Let T_1 the r.v. of the first meeting time between S and the first relay node in $\mathcal{N} - \{S\}$, and let T_i be the r.v. of the meeting time between the

$(i-1)$-st relay node and the i-th relay node. Let T_ℓ be the r.v. counting the time it takes for the first relay node among $\mathcal{N} \setminus \{D\}$ to meet D.

The following theorem gives a lower bound on FM* routing with two copies and a constant number of hops.

Theorem 3. $\mathbb{E}[T^{ib}_{FM^*}] = \Omega(\log(1/\delta))$.

Let us consider the situation when we have $q > 2$ copies of M and ℓ hops. Assume wlog $q = 2^w$ for $w \in \mathbb{N}$, and that the copies of M are forwarded at each hop as follows: whenever a node contains 2^s, $s \geq 1$ copies of a message and meets a node different from D, in the independent mobility model it always gives to that node 2^{s-1} copies M – this is the Binary SW strategy of [18]. In the interest-based mobility model it gives to the node 2^{s-1} copies only if the new node is closer to D than the previous hops containing some copies of M. Assume also that all relay nodes keep the last copy for itself and deliver it only if they meet D. Therefore the number of hops is at most $\log_2 q$.

Theorem 4. *For any constant number of copies* $\mathbb{E}[T^{ib}_{FM^*}] = \Omega(\log(1/\delta))$.

Unknown destination. A major limitation of the interest based routing previously considered is that the sender must know the interest profile of the destination i.e., the coordinates $D[a_1, a_2, \ldots, a_m]$ in the interest space. We now relax this assumption assuming that S knows the identity of node D (so delivery of M to D is possible), but not its interest profile and we show that a modified version of the $IB(\gamma)$ routing that uses more than one copy of the message also provides asymptotically the same upper bound.

The idea is that the routing chooses $m-1$ relay nodes with the characteristic that each one the $m-1$ relay nodes will be "almost orthogonal" to the others and to S, and S will pass a copy to each one of them, and keep one. Namely, let $\hat{R}_j$ denote the j-th relay chosen node, $j = 1, 2, \ldots, m-1$. We consider the following routing algorithm $\texttt{Mod-}IB(\gamma)$ to choose relay nodes:

If S meets a node with coordinates $R_i[r_1, r_2, \ldots, r_m]$, the node becomes the j-th relay node $\hat{R}_j$, $j = 1, 2, \ldots, q-1$, if the following conditions are met: $0.05 \leq R_i[1] \leq 0.1$; $\exists k$, $2 \leq k \leq m$ s. t. $0.8 \leq R_i[k] \leq 0.85$; $\forall s$, $1, \leq s \leq j-1$, $\hat{R}_s[k] < 0.8$

Theorem 5. *For a constant* $c > 0$ *and* $\gamma = \frac{0.29}{m-1}$ *we have* $\mathbb{E}[T^{ib}_{\texttt{Mod}-IB(\gamma)}] \leq m\gamma/c$.

4 Simulations

We have qualitatively verified our asymptotic analysis through simulations, based on both a real world trace collected at the Infocom 2006 conference – the trace used in [14,16] –, and the SWIM mobility model of [13], which is shown to closely resemble fundamental features of human mobility.

4.1 Real-World Trace Based Evaluation

A major difficulty in using real-world traces to validate our theoretical results is that no information about user interests is available, for the vast majority of available traces, making it impossible to realize IB routing. One exception is the Infocom 06 trace [8], which has been collected during the Infocom 2006 conference. This data trace contains, together with contact logs, a set of user profiles containing information such as nationality, residence, affiliation, spoken languages etc. Details on the data trace are summarized in Table 1.

Table 1. Detailed information on the Infocom 06 trace

Experimental data set	Infocom 06
Device	iMote
Network type	Bluetooth
Duration (days)	3
Granularity (sec)	120
Participants with profile	61
Internal contacts number	191,336
Average Contacts/pair/day	6.7

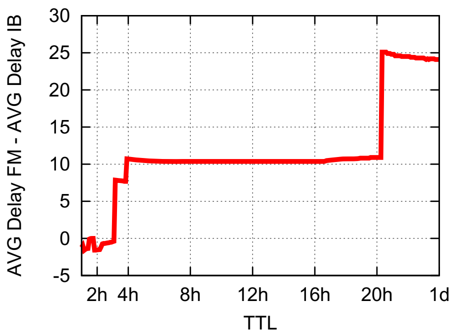


Fig. 1. Difference between average packet delivery delay with FM and IB routing with the Infocom 06 trace as a function of the message TTL

Similarly to [14], we have generated 0/1 interest profiles for each user based on the corresponding user profile. Considering that data have been collected in a conference site, we have removed very short contacts (less than $5min$) from the trace, in order to filter out occasional contacts – which are likely to be several orders of magnitude more frequent than what we can expect in a non-conference scenario. Note that, according to [14], the correlation between meeting frequency of a node pair and similarity of the respective interest profiles in the resulting data trace (containing 53 nodes overall) is 0.57. Thus, the Infocom 06 trace, once properly filtered, can be considered as an instance of interest-based mobility, where we expect IB routing to be superior to FM routing.

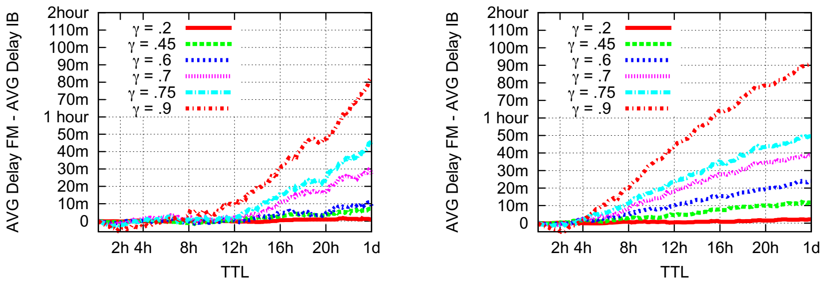
In order to validate this claim, we have implemented both FM and IB routing. We recall that in case of FM routing, the source delivers the second copy of its message to the first encountered node, while with IB routing the second copy of the message is delivered by the source to the first node whose interest similarity with respect to the destination node is at least γ. The value of γ has been set to $0.29/(m-1)$ as suggested in the analysis, corresponding to 0.0019 in the Infocom 06 trace. Although this value of the forwarding threshold is very low, it is nevertheless sufficient to ensure a better performance of IB vs. FM routing.

The results obtained simulating sending 5000 messages between randomly chosen source/destination pairs are reported in Figure 1. For each pair, the message is sent with both FM and IB routing, and the corresponding packet delivery time are recorded. Experiments have been repeated using different TTL (TimeToLive) values of the generated message. Figure 1 reports the difference between the average delivery time with FM and IB routing, and shows that a lower average delivery time is consistently observed with IB routing, thus qualitatively confirming the theoretical results derived in the previous section.

4.2 Synthetic Data Simulation

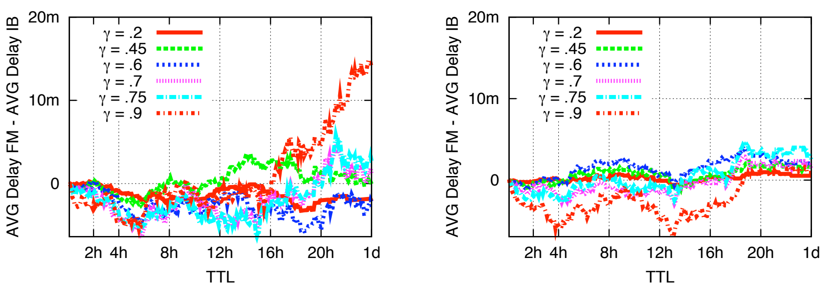
The real-world trace based evaluation presented in the previous section is based on a limited number of nodes (53), and thus it cannot be used to validate FM and IB scaling behavior. For this purpose, we have performed simulations using the SWIM mobility model [13], which has been shown to be able to generate synthetic contact traces whose features very well match those observed in real-world traces. Similarly to [14], the mobility model has been modified to account for different degrees of correlation between meeting rates and interest-similarity. We recall that the SWIM model is based on a notion of "home location" assigned to each node, where node movements are designed so as to resemble a "distance from home" vs. "location popularity" tradeoff. Basically, the idea is that nodes tend to move more often towards nearby locations, unless a far off location is very popular. The "distance from home" vs. "location popularity" tradeoff is tuned in SWIM through a parameter, called α, which essentially gives different weights to the distance and popularity metric when computing the probability distribution used to choose the next destination of a movement. It has been observed in [13] that giving preference to the "distance from home" component of the movement results in highly realistic traces, indicating that users in reality tend to move close to their "home location". This observation can be used to extend SWIM in such a way that different degrees of interest-based mobility can be simulated. In particular, if the mapping between nodes and their home location is random (as in the standard SWIM model), we expect to observe a low correlation between similarity of user interests and their meeting rates, corresponding to a social-oblivious mobility model. On the other hand, if the mapping between nodes and home location is done based on their interests, we expect to observe a high correlation between similarity of user interests and their meeting rates, corresponding to an interest-based mobility model.

Interest profiles have been generated considering four possible interests ($m = 4$), with values chosen uniformly at random in $[0, 1]$. In case of interest-based mobility, the mapping between a node interest profile and its "home location" has been realized by taking as coordinates of the "home location" the first two coordinates of the interest profile. In the following we present simulation results



(a) IM network of 1000 nodes. (b) IM network of 2000 nodes.

Fig. 2. Difference between average packet delivery delay with FM and IB routing with SWIM mobility in the Interest-based mobility (IM) scenario, as a function of the message TTL



(a) NIM network of 1000 nodes. (b) NIM network of 2000 nodes.

Fig. 3. Difference between average packet delivery delay with FM and IB routing with SWIM mobility in the Non Interest-based mobility (NIM) scenario, as a function of the message TTL

referring to scenarios where correlation between meeting rate and similarity of interest profiles is -0.009 (denoted Non-Interest based Mobility – NIM – in the following), and 0.61 (denoted Interest-based Mobility – IM – in the following), respectively. We have considered networks of size 1000 and 2000 nodes in both scenarios, and sent 10^5 messages between random source/destination pairs. The results are averaged over the successfully delivered messages. In the discussion below we focus only on average delay. However, we want to stress that in both IM and NIM scenarios, the IB routing slightly outperforms FM in terms of delivery rate (number of messages actually relayed): The difference of delivery rates is about 0.015% in favor of IB.

Figure 2 depicts the performance of the protocols for various values of γ on IM mobility. As can be noticed by the figure, the larger the relay threshold γ, the more IB outperforms FM. Moreover, as predicted by the analysis, the performance improvement of IB over FM routing becomes larger for larger networks. Indeed, for $\gamma = .9$ and $TTL = 24h$, message delivery with IB is respectively

$80m$ and $90m$ faster on the network of respectively 1000 nodes (see Figure 2(a)) and 2000 nodes (see Figure 2(a)). This means that, on IM mobility, IB routing delivers more messages with respect to FM, and more quickly.

Notice that the results reported in Figure 2 apparently are in contradiction with Theorem 2, which states an upper bound on the expected delivery time which is directly proportional to γ – i.e., higher values of γ implies a looser upper bound. Instead, results reported in Figure 2 show an increasingly better performance of IB vs. FM routing as γ increases. However, we notice that the bound reported in Theorem 2 is a bound on the *absolute* performance of IB routing, while those reported in Figure 2 are results referring to the *relative* performance of IB vs. FM routing.

The performance of the protocols on NIM mobility is depicted in Figure 3. In this case, the two protocols interchangeably perform better or worse in terms of delay. The negative values in the figure are due to the few more messages that IB delivers to destination whereas FM does not. Some of these messages reach the destination slightly before message TTL, by thus increasing the average delay. However, independently of γ, the values are close to zero. Indeed, note the difference of the $y-$axis between Figures 2 and 3. This indicates that, if mobility is not correlated to interest similarity, as far as the average delay is concerned the selection of the relay node is not important: A node meeting the forwarding criteria in IB routing is encountered on average soon after the first node met by the source.

4.3 Discussion

The Infocom 06 trace is characterized by a moderate correlation between meeting frequency and similarity of interest profiles – the Pearson correlation index is 0.57. However, it is composed of only 53 nodes. Despite the small network size, our simulations have shown that IB routing indeed provides a shorter average message delivery time with respect to FM routing, although the relative improvement is almost negligible (in the order of 0.06%).

To investigate relative FM and IB performance for larger networks, we used SWIM, and simulated both social-oblivious and interest-based mobility scenarios. Once again, the trend of the results qualitatively confirmed the asymptotic analysis: in case of social-oblivious mobility (correlation index is -0.009), the performance of FM and IB routing is virtually indistinguishable for all network sizes; on the other hand, with interest-based mobility (correlation index is 0.61), IB routing provides better performance than FM. It is interesting to observe the trend of performance improvement with increasing network size: performance is improved of about 5.5% with 1000 nodes, and of about 6.25% with 2000 nodes. Although percentage improvements over FM routing are modest, the trend of improvement is clearly increasing with network size, thus confirming the asymptotic analysis. Also, IB forwarding performance improvement over FM forwarding becomes more and more noticeable as the value of γ, which determines selectivity in forwarding the message, becomes higher: with $\gamma = 0.2$ and 2000 nodes, IB improves delivery delay w.r.t. FM forwarding of about 0.1%; with $\gamma = 0.6$ improvement becomes 1.7%, and it raises up to 6.25% when $\gamma = 0.9$.

5 Conclusion

We have formally analyzed and experimentally validated the delivery time under mobility and forwarding scenarios accounting for social relationships between network nodes. The main contribution of this paper is proving that, under fair conditions for what concerns storage resources, social-aware forwarding is asymptotically superior to social-oblivious forwarding in presence of interest-based mobility: its performance is never below, while it is asymptotically superior under some circumstances – orthogonal interests between sender and destination.

We believe several avenues for further research are disclosed by our initial results, such as considering scenarios in which individual interests evolve in a short time scale, or scenarios in which forwarding of messages is probabilistic instead of deterministic.

References

1. Boldrini, C., Conti, M., Passarella, A.: ContentPlace: Social-Aware Data Dissemination in Opportunistic Networks. In: Proc. ACM MSWiM, pp. 203–210 (2008)
2. Costa, P., Mascolo, C., Musolesi, M., Picco, G.P.: Socially-Aware Routing for Publish-Subscribe in Delay-Tolerant Mobile Ad Hoc Networks. IEEE Journal on Selected Areas in Communications 26(5), 748–760 (2008)
3. Daly, E., Haahr, M.: Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs. In: Proc. ACM MobiHoc, pp. 32–40 (2007)
4. Deza, M.M., Deza, E.: Encyclopedia of Distances. Springer, Berlin (2009)
5. Diaz, J., Marchetti-Spaccamela, A., Mitsche, D., Santi, P., Stefa, J.: Social-Aware Forwarding Improves Routing Performance in Pocket Switched Networks. Tech. Rep. IIT-11-2010, Istituto di Informatica e Telematica del CNR, Pisa (2010)
6. Fall, K.: A Delay-Tolerant Architecture for Challenged Internets. In: Proc. ACM Sigcomm, pp. 27–34 (2003)
7. Grossglauser, M., Tse, D.N.C.: Mobility Increases the Capacity of Ad-Hoc Wireless Networks. In: Proc. IEEE Infocom, pp. 1360–1369 (2001)
8. Hui, P., Crowcroft, J., Yoneki, E.: BUBBLE Rap: Social-Based Forwarding in Delay Tolerant Networks. In: Proc. ACM MobiHoc, pp. 241–250 (2008)
9. Hui, P., Chaintreau, A., Scott, J., Gass, R., Crowcroft, J., Diot, C.: Pocket-Switched Networks and Human Mobility in Conference Environments. In: Proc. ACM Workshop on Delay-Tolerant Networks (WDTN), pp. 244–251 (2005)
10. Ioannidis, S., Chaintreau, A., Massoulie, L.: Optimal and Scalable Distribution of Content Updates over a Mobile Social Networks. In: Proc. IEEE Infocom, pp. 1422–1430 (2009)
11. Li, F., Wu, J.: LocalCom: A Community-Based Epidemic Forwarding Scheme in Disruption-tolerant Networks. In: Proc. IEEE Secon (2009)
12. McPherson, M.: Birds of a feather: Homophily in Social Networks. Annual Review of Sociology 27(1), 415–444 (2001)
13. Mei, A., Stefa, J.: SWIM: A Simple Model to Generate Small Mobile Worlds. In: Proc. IEEE Infocom (2009)
14. Mei, A., Morabito, G., Santi, P., Stefa, J.: Social-Aware Stateless Forwarding in Pocket Switched Networks. In: Proc. IEEE Infocom, MiniConference (2011)
15. Mitzenmacher, M., Upfal, E.: Probability and Computing. Cambridge U.P. Cambridge (2005)

16. Noulas, A., Musolesi, M., Pontil, M., Mascolo, C.: Inferring Interests from Mobility and Social Interactions. In: Proc. ANLG Workshop (2009)
17. Resta, G., Santi, P.: The Effects of Node Cooperation Level on Routing Performance in Delay Tolerant Networks. In: Proc. IEEE Secon (2009)
18. Spyropoulos, T., Psounis, K., Raghavendra, C.S.: Efficient Routing in Intermittently Connected Mobile Networks: The Multi-copy Case. IEEE Trans. on Networking 16(1), 77–90 (2008)
19. Spyropoulos, T., Psounis, K., Raghavendra, C.S.: Efficient Routing in Intermittently Connected Mobile Networks: The Single-copy Case. IEEE Trans. on Networking 16(1), 63–76 (2008)
20. Vahdat, A., Becker, D.: Epidemic Routing for Partially Connected Ad Hoc Networks. Tech. Rep. CS-200006, Duke University (April 2000)