Social-Aware Stateless Routing in Pocket Switched Networks

Alessandro Mei, *Member, IEEE*, Giacomo Morabito,
Paolo Santi, *Member, IEEE Computer Society*, and Julinda Stefa

Abstract—Existing social-aware routing protocols for pocket switched networks make use of the information about the social structure of the network deduced by state information of nodes (e.g., history of past encounters) to optimize routing. Although these approaches are shown to have superior performance to social-oblivious, stateless routing protocols (BinarySW, Epidemic), the improvement comes at the cost of considerable storage overhead required on the nodes. In this paper we present SANE, the first routing mechanism that combines the advantages of both *social-aware* and *stateless* approaches. SANE is based on the observation—that we validate on a real-world trace—that individuals with similar interests tend to meet more often. In SANE, individuals (network members) are characterized by their *interest profile*, a compact representation of their interests. By implementing a simple routing rule based on interest profile similarity, SANE is free of network state information, thus overcoming the storage capacity problem with existing social-aware approaches. Through thorough experiments, we show the superiority of SANE over existing approaches, both stateful, social-aware and stateless, social-oblivious. We discuss the statelessness of our approach in the supplementary file, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPDS.2014.2307857, of this manuscript. Our interest-based approach easily enables innovative networking services, such as interest-casting. An interest-casting protocol is also introduced in this paper, and evaluated through experiments based on both real-world and synthetic mobility traces.

Index Terms—Delay-tolerant networks, opportunistic networks, mobile social networks, community structure, routing

✦

1 INTRODUCTION

THE vision of a near future in which a multitude of handheld devices establish direct wireless communication links in an opportunistic fashion has attracted the attention of the research community. In particular, the benefits of complementing Internet connectivity with opportunistic communications have been recently demonstrated in [20]. Pocket switched networks (PSNs) are specific types of delay tolerant networks (DTNs) where nodes are powerful devices communicating through short-range technology (e.g., BlueTooth) carried by *individuals* [14]. Nodes mobility, coupled with a *store-and-forward* mechanism is the fundamental means of communication in DTNs and PSNs.

Existing routing approaches for PSNs have both advantages and disadvantages. Social-aware routing protocols [7], [11], [12], [19] are shown to have superior performance to social-oblivious schemes (e.g., BinarySW [27]). But, the improvement comes at the cost of processing a significant amount of state information (e.g., information on past encounters among nodes, current structure in communities,

etc.) and storing it at the local memory of the nodes. A common feature of existing social-aware routing protocols is that they are *stateful*, i.e., nodes need to collect, compute, and store information about previous interactions to be able to make correct routing decisions about future messages. So, a node that leaves the network for, e.g., one week, when it returns it is not ready to take part in routing decisions. Conversely, a stateless approach like, e.g., Epidemic, frees nodes from these restrictions: Nodes can join and leave in a very simple way and can always take part in the routing mechanism. All they need to know is the destination's ID.

We present Social-Aware NEtworking (SANE), a *social-aware*, *stateless* approach, that combines the advantages of social-aware routing with the benefits of stateless protocols. SANE exploits a sociological observation [21]: People with similar interests tend to meet more often. The authors of [25] use this observation independently and alternatively to ours, showing that mobility patterns can be used to accurately predict individual interests. A first significant contribution of this work (also included in its preliminary conference version [40]) is a *quantitative* validation of the same observation, based on the largest real-world mobility trace enriched with user profiles information at the time of our writing [12], [13]. Later on, the same observation has been exploited in other works that span from community detection to collection of network data for research purposes [35], [36], [37].

In SANE, each network individual is characterized with an *interest profile* (IP)—a compact representation of user's interest—belonging to the network's *interest space*. Such information is then used by SANE routing: When individual A carrying a message M destined to individual D (whose interest-profile is contained in M) meets another individual B, she compares D and B's interest profiles. Depending on

---

- *A. Mei and J. Stefa are with the Department of Computer Science, Sapienza University of Rome, via Salaria 113, 00198 Rome, Italy. E-mail: {mei, stefa}@di.uniroma1.it.*
- *G. Morabito is with the Department of Computer Science and Telecommunications Engineering, University of Catania, 95125 Catania, Italy. E-mail: gmorabi@diit.unict.it.*
- *P. Santi is with the Department of Computational Mathematics, Istituto di Informatica e Telematica, Area della Ricerca via G., Moruzzi 1, 56124 Pisa, Italy. E-mail: paolo.santi@iit.cnr.it.*

the outcome of this comparison, she decides whether to forward M to B. Besides from routing, the SANE approach seems naturally built for *innovative* networking services like interest-casting, where a message is meant to reach a certain group of users in the network: Those to which this message is of interests. So, we characterize a message M circulating in the network by a *message relevance profile*. The relevance profile of the message is also represented as a point in the interest space, and the goal is to deliver a copy of M to all potentially interested users, i.e., individuals whose interest profile is "close enough" (according to a certain similarity metric) to M's relevance profile. An interest-casting protocol is also introduced in this paper. Our experiments with both real-world and synthetic traces show the superiority of our proposed social-aware, stateless routing and interest-casting approaches over existing stateful, social-aware as well as stateless, social-oblivious routing approaches.

The paper is organized as follows: Section 2 discusses related work; Section 3 introduces the notions of interest space and interest profile on which our approach is based. In Section 4, we introduce Social Aware NEtworking. Section 5 presents the evaluation of SANE's performance with both real-world and synthetic mobility traces. Finally, Section 6 concludes the paper.

2 RELATED WORK

The works that exploit social information between nodes in PSNs can be divided into three main categories. The first includes protocols that exploit network structure, node centralities measures, or node meeting rates for routing [2], [7], [11], [12], [15], [19], [39]. These existing social-aware approaches heavily build upon the ability of storing and handling large amounts of state information by nodes [2], [7], [11], [19], [39], or require computation of complex metrics upon the whole network like community membership as in LABEL [15]. Often, this information only does not suffice. Indeed, LABEL is outperformed by its successors requiring more information than just community membership like BUBBLE [12], which also requires ranking of nodes within all of their communities, and within the whole network. In a word, all these protocols are *stateful*. It is worthy discussing that in [39] the authors exploit a centrality measure of nodes in disseminating data in a user-centric way. They take into account users' interests. Differently from SANE, interests are considered unrelated to their capability to contact others interested in the same data.

Publish-subscribe mechanisms exploiting social-ties [1], [4], [18], [34] compose the second category. In [18] a service provider (e.g., a cellphone operator) selectively sends users dynamic content updates that can be shared with others when a communication opportunity arises. The performance improves when the provider considers the social-ties strength in the network. In [1] data is shared with the goal of optimizing content availability through careful, social-aware data placement. The HiBOp [34] approach learns and represents through context information the users social relations. Nodes store (and update) information on their identity and meeting history, and use it in a routing mechanism that resembles Milgram's experiment. As the authors point out, the information that nodes need to store and update on

the other network nodes can potentially become very large. In [4], the authors present SocialCast, a routing mechanism that exploits predictions based on metrics of social interactions. SocialCast remains a publish-subscribe scheme, and again, it requires nodes to store a considerable amount of state information. Similarly to the routing protocols previously discussed, all of these publish-subscribe schemes are stateful—they rely on complex metrics computed either on the whole network or by the nodes, based on information observed (and stored) during past interactions.

The third category is that exploiting mobility-profiles for routing [16], [17], [31], [32], [33]. Mobility-profiles represent the likelihood of users to visit geographic locations within the network (e.g., bar, home, a given bus station). They can either be built during a training-phase, as the nodes visit the places and get feedback from, e.g., Wi-Fi APs (as in [31], [32]), and then be used to predict future visits, or, be built during a training phase and then dynamically updated from the nodes' mobility pattern [16], [17]. Besides capturing similarities between "friends" frequenting similar geographic locations, this approach also captures communication opportunities between "familiar strangers": e.g., strangers that take the same bus.

Note that the mobility-profiles are different from the interest-profiles that SANE is based upon. As a matter of fact, they only give information on the probabilities/time spent by the nodes in a certain network area. Even though they capture communication opportunities between familiar strangers, approaches based on mobility-profiles miss other types of casual communication opportunities: those that arise between people with similar interests, yet having typically different mobility patterns (e.g., people that live in different areas of the city and never cross each other's path, yet sharing the same interest of going to the movies). We describe more in details how SANE identifies these more subtle, yet important, contact opportunities, in the supplementary file, available online, to this paper. In addition, as we will also see from the experimental results, this aspect of the interest-profiles makes so that SANE out-performs not only social-oblivious competitors, but also complex social-aware ones like BUBBLE, and mobility-profile based ones like MobySpace [31], [32]. It is worth observing that routing approaches based on a notion of user profile inspired by our model of interest space have recently been proposed in the literature [10], [29], [30].

Lastly, the work in [38] investigates features of similarity in people's rates on films and their relation to the way users determine trust in online recommendation systems.

3 INTEREST SPACE AND PROFILES

3.1 The Model

We assume each individual in the network can be represented through a unique identifier (her *address*) and her *interest profile*, i.e., a compact representation of her interests within the *interest space*. The interests, in the sense of our work are defined as follows: An attribute of an individual that potentially induces mobility patterns that bring him physically close to other people. So, for example, liking lonely walks through the woods is not an interest in the sense of our work. Conversely, liking Shakespeare, is. It
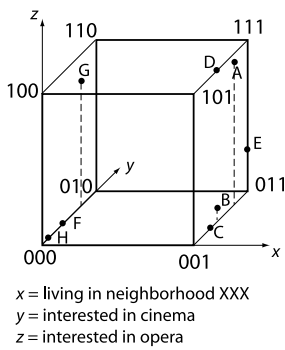
Fig. 1. An example of network with three interest dimensions and eight individuals. Individuals $A, \ldots, E$ live in neighborhood XXX, and have different degrees of interest in cinema and opera, while individuals F, G, and H live outside neighborhood XXX.

potentially pushes an individual to go to a theater showing a Shakespeare's play, or to the bookshop/library to get Shakespeare's books, and thus meet other people.

We represent the interest space as an m-dimensional unit cube $\mathcal{C} = [0, 1]^m$, where m is the total number of interests in the network under consideration. For a given $i \in \{1, \ldots, m\}$, the value of the corresponding dimension in the interest space (the ith interest dimension) is either a 0/1 value or an arbitrary real value in $[0, 1]$. This enables 0/1 interests like "membership to a certain community", and arbitrary "degree of interest" in a given topic.

The notion of interest profile introduced above is very general, and can be used to capture not only similarity between interests, but also, to some extent, similarity between mobility patterns. For instance, if two individuals live in the same neighborhood—thus, have a value of 1 in the corresponding interest dimension—it is likely that their mobility patterns display some similarity. In the supplementary file, available online, of this manuscript we discuss in details how to efficiently acquire user interest-profiles and how to handle their dynamics.

Given the above definition of interest space, it is quite natural to represent the interest profile of an individual A with an m-dimensional vector reporting, for each possible interest dimension, A's degree of interest in the particular topic/community (either a real number or a binary value). Thus, we can think of individual interest profiles as points in the m-dimensional interest space. For example, Fig. 1 represents a set of eight individuals, denoted as $A, B, \ldots, H$, in a network with three interest dimensions. The x dimension corresponds to "living in neighborhood X" and allows only binary membership values. The y and z dimensions, respectively "interested in cinema" and "interested in opera" have continuous membership values. To express similarity between individual interests, and thus quantitatively measure "homophily" (degree of interest similarity [21]), we use the well-known *cosine similarity* metric [26]:

Definition 1. *Given two m-dimensional vectors A and B, the cosine similarity metric, denoted $\Theta(A, B)$, is defined as:*

$$\Theta(A, B) = \cos(\angle AB) = \frac{A \cdot B}{\parallel A \parallel \parallel B \parallel},$$

where $\parallel X \parallel$ represents the length of vector X.

TABLE 1
Detailed Information on the Infocom 06 Trace

Experimental data set	Infocom 06
Device	iMote (Bluetooth)
Duration (days)	3
Granularity (sec)	120
Participants number	78
Participants with profile	61

From the definition of interest space in our model we have that $0 \leq \Theta(A, B) \leq 1$, with higher values of $\Theta(A, B)$ corresponding to a higher "homophily". Lastly, we observe that the cosine similarity metric defined above is very simple to compute. Thus, it should not induce noticeable computation overhead to powerful nowadays smartphones.

3.2 Validation

As mentioned, our stateless protocols are based on a simple and natural observation from everyday life: Our movements are guided in a large part by our interests. To validate this intuition quantitatively we use traces collected during an experiment done with real Bluetooth communicating devices distributed to part of the participants of the Infocom 2006 conference [12], [13]. The devices were configured to perform a Bluetooth baseband layer "inquiry" discovering the MAC addresses of other Bluetooth nodes in range of communication. The results of the inquiry were written to flash RAM, recording contact periods between devices, in the form of {MAC, start time, end time}. This data trace contains not only contact logs, but it also reports information on participants' nationality, residence, languages spoken, affiliation, scientific interests, etc. From this information we can easily generate an interests profile vector of 0/1 coordinates: We count all the possible nationalities, countries and cities of residence, languages spoken, affiliations, possible scientific interest topics, declared by the participants. These are the interests. Then, we build, for each participant, a profile vector that has as many coordinates as interests. A 1 in the ith coordinate of a given participant's profile vector corresponds to the fact that that participant is either interested in the scientific topic, or speaks that particular language, or comes from that particular country (depending on what interest dimension i represents). In the process, we discard participants that have not declared any of the above interests, so that the validation is not biased by incomplete profiles. The number of the participants in the trace after this selection reduces to 61. Although there are other data-traces available on line describing contact among participants in different experimental settings [9], [12], [13], [14], only few of them contains some information on participants' profiles, coupled with fine-grained tracing of pair-wise meetings. To the best of our knowledge, Infocom 06 is the largest available data trace to date that displays both features, thus in this paper we focus on this data trace. More details about the data trace can be found in Table 1.

To support our intuition, we first calculate the cosine similarity between the interest profiles for every pair of participants. Then, we compute the Pearson correlation index among this value and the total meeting duration/meeting frequency among every couple. These values result to be 0.28 and 0.08, respectively. The second correlation

TABLE 2
Correlation of Interest Profiles and the Meetings of Participants

AVG meet time	C_d	C_f	Nodes
> 0 (min)	.28	.08	61
> 5 (min)	.55	.57	53
> 10 (min)	.67	.67	26

C_d and C_f indicate the Pearson correlation coefficient between pairwise similarity of participants' profiles and, respectively, total meeting durations and meeting rates.

coefficient is small: This is more than reasonable, being this trace the result of the movement in a big conference, where there is a high "mixing" of people and thus a high number of short, mostly random meetings. For example, almost all the attendees meet during the coffee break. Yet, the first correlation coefficient (the one related to the duration of the contacts between people) shows that even in the presence of a high number of casual meetings, people with similar profile tend to meet for longer times. To confirm this observation, we then compute the correlation coefficients among profile similarities and meeting duration/meeting frequency, only for pairs of individuals who spend, on the average, more than a certain amount of time together. This way the effect of the random short meetings is attenuated. The results are presented in Table 2. As can be seen, when we focus on longer meetings, the correlation of meeting frequency and similarity of interest profiles is considerably higher, reaching 0.67. As discussed more in depth in the supplementary file, available online, of this manuscript, these results support the conclusion that our intuition is sound and that it can be used as the basic mechanism of social-aware, stateless routing protocols.

4 SOCIAL AWARE NETWORKING

Here we introduce *Social Aware NEtworking*, a protocol suite that enables the efficient delivery of information to relevant destinations in PSNs. Besides the traditional unicast (called UN-SANE in the rest of the paper), SANE supports also *interest-cast*, a novel communication service where messages are destined to a set of nodes: those to which the message is of interest, according to their profile.

We assume that each node can be a relay and the messages are carried out in a *store-and-forward* fashion. Messages are equipped with a header containing the target interest-profile (*message relevance profile* throughout the paper), an integer N_{replicas} representing the number of replicas of the message that the node is allowed to forward to other relays, and a time-to-live (TTL) value. In UN-SANE (the unicast version), the header contains also the destination's node ID, and the *message relevance profile* coincides with the destination's interest-profile. In the interest-cast version, the header, along with the *message relevance profile*, contains also a threshold value α that determines the set of relevant destination nodes as explained later on in this section.

In PSNs nodes can exchange information as communication opportunity turns up. Accordingly, SANE procedures are triggered each time a node (say A) enters within the radio coverage of another node (say B). Initially, nodes exchange their interest profiles as they will

be used to take the most appropriate routing decisions; then, each node starts scanning its buffer of the messages to relay. The treatment of each message depends on its type, (i.e., unicast or interest-cast), and will be described respectively in Sections 4.1 and 4.2. After all messages in the buffer have been processed, the node updates the buffer. This is achieved by

- *removing messages that are obsolete.* To this aim a deadline instant, t_{dead}, is assigned to each message in the buffer.
- *handling the messages relayed by the other node.* If the node is a destination then the message is delivered to the application; if the node is a relay then it inserts the message in the buffer.

4.1 Unicast

In the unicast case we aim at the best tradeoff between communication overhead, delivery success (i.e., the probability that the packet reaches the destination before it elapses), and delivery delay. According to our interest-based approach, a message M should preferably be forwarded to individuals whose interest profile closely resembles the one of the destination. As in [27], we assume that in order to keep the communication overhead under control, the same message can be relayed at most for N_{replicas}^* times. The source is responsible for initializing the values of N_{replicas}, which must be a power of 2 and represents the maximum values of replicas of the message in the network, and the value of TTL, which represents the maximum delay acceptable for the delivery of the message. We denote the initial values of N_{replicas} and TTL as N_{replicas}^* and TTL^*, respectively. The message relevance profile is set equal to the interest profile of the destination node.

More specifically, the forwarding rule is as follows: Message M is relayed to node B if and only if both the following conditions hold:

- the current value of N_{replicas} is higher than 1;
- the cosine similarity metric between the relevance of message M, denoted as $R(M)$, and the interest-profile of B, denoted $IP(B)$, is higher than a given threshold ρ that we call *relaying threshold* as below:

$$\Theta(R(M), IP(B)) \geq \rho. \tag{1}$$

Note that if N_{replicas} equals 1, the message is not relayed to node B, even though the cosine similarity between $IP(B)$ and $R(M)$ is higher than the cosine similarity between $IP(A)$ and $R(M)$. Even though such design choice may seem counter-intuitive it guarantees an upper bound on the communication overhead.

Before passing M to B the value of N_{replicas} contained in $M's$ header is halved, whereas the value of TTL is set equal to the difference between the deadline instant and the current time. The so modified copy of M is sent to B. This is equivalent to handing node B half of the copies of M currently in node A's buffer, as done in BinarySW [27]. When B is the destination of the message this latter is transmitted to B regardless of the value of N_{replicas}. In this case, M is removed from $A's$ buffer after transmission.

As the threshold ρ decreases, the routing strategy becomes more aggressive—more and more nodes result possible relays. This results in faster delivery and an increase of both success delivery and cost. This latter, denoted with $c(M)$, is proportional to the number of copies of M spread in the network. A few extreme cases can be considered:

- $N^*_{\text{replicas}} = \infty$: in this case there is no bound on the number of copies of the message circulating in the network. We call the resulting version of our protocol suite *epidemic* SANE, and we denote it with UN-SANE EP. The SANE version corresponding to the case $N^*_{\text{replicas}} < \infty$ is instead called *spray & wait* SANE and denoted UN-SANE SW.
- $\rho = 0$: The relay threshold is not used, and the proposed routing strategy becomes the same as BinarySW [27]. Furthermore, if N^*_{replicas} is set to ∞ then our protocol behaves like Epidemic [28], which is the fastest but also the more costly forwarding strategy.
- $\rho = 1$: Only direct message delivery from source to destination is possible. Message delivery cost is minimized, but message delivery delay can be very high.

We include a thorough experimental study of the impact of the thresholds α and ρ on the performance of the routing strategy in the supplementary file, available online, of this manuscript.

4.2 Interest-Cast

PSNs can create innovative services realized within the PSN itself, without the need of resorting to pre-existing communication facilities. *Interest-cast* is an example of such services in which a user wants to communicate a certain information (for instance, announcing a movie at a local theater about opera composer Puccini) to the maximum possible number of interested users, within a certain time (e.g., the time of the last movie show). Interested users might have an interest in opera, or cinema, or both, and should be located in the "neighborhood" of the theater, so to be able to reach the theater if interested. This type of communication paradigm matches very well with the localized nature of PSN communications: the information is spread relatively fast in the neighborhood of the sender, while it takes longer to propagate to remote areas (which are typically less interested in the information, though).

Assume individual C wants to send a message M to all or the largest possible number of potentially interested individuals within the network. First, C must set the message relevance profile of M, which can be done assigning for each of the m interest dimensions a "relevance" value in the $[0, 1]$ interval. Such m-dimensional vector associated with a message is used (coupled with the individuals' interest profiles) to drive information propagation within the PSN. Recall that the notion of message relevance profile $(R(M))$ allows to represent message M—similarly to individuals—as a point in the interest space. The set of *relevant destinations* for M, denoted $RD(M)$, is the set of network individuals for which message M is deemed relevant; i.e., nodes to which message M should be delivered within an upper bound on the delivery time denoted TTL^*. Whether a message M is relevant for a certain individual B is determined using a certain *relevance metric*. In this paper we use the well-known cosine similarity metric [26] to determine whether message M is relevant for individual B.

Both individuals' interests and message relevance profiles take values in the same m-dimensional interest space. Thus, for any individual B and message M, the angle between $IP(B)$ and $R(M)$ is in $[0, \pi/2]$, implying that $\Theta(B, M)$ is indeed in $[0, 1]$. The relevance of a message M to individual B is decided through the following rule: M is relevant to B if and only if $\Theta(IP(B), R(M)) \geq \alpha$, where α is a suitably chosen *relevance threshold*.

The routing mechanism of interest-cast is similar to the unicast case. In fact, if the two conditions given in Section 4.1 for the unicast case hold then the message is relayed to B in the same way. If the above two conditions are not met but B is a relevant destination, then the message is transmitted with N_{replicas} set to one and TTL evaluated as explained in Section 4.1. Note that the above transmission does not have impact on the communication overhead.

We want to stress the difference between the notion of interest-casting defined herein and more traditional communication paradigms and services such as multi-casting and publish-subscribe. In interest-casting, the only action taken by a "content provider" (an individual generating a message) is determining the message relevance profile. After that, the message is injected in the network, and information propagation is driven by the notions of relevance and interest profile. As we shall see, these notions are used not only to dynamically determine the set of relevant destinations, but also to govern the routing process. Thus, in interest-casting the content-provider is not aware of the set of destinations the content should be delivered to. This is in sharp contrast with the traditional notion of multi-casting where multi-cast groups are explicitly defined and typically known to the content provider. Furthermore, interest-cast destinations must not explicitly subscribe to a specific "topic". All individuals do in our network specify what they are interested in. This is enough for them to get relevant messages of topics of their interests, as we will see from the experimental results. This is also in sharp contrast with publish-subscribe mechanisms. These typically require users to explicitly subscribe to one or several specific "topics" to be able to receive the corresponding information.

5 EXPERIMENTAL SETUP AND RESULTS

Here we present experimental results on the performance of both the unicast and interest-cast version of SANE compared to that of well known opportunistic routing protocols. For the evaluation we use both real-world traces (Infocom 06) and synthetic ones obtained with the SWIM mobility model [22]. SWIM allows us to evaluate the performance under different correlations among profiles and meeting-rates. This is not possible with the Infocom 06 trace.

5.1 Evaluation Using Infocom 06 Traces

To validate the protocols on the Infocom 06 trace we generate 5,000 messages with a uniform traffic pattern (source-destination chosen uniformly at random), and we set messages' relevance profile to be equal to the destination's interest profile. Then, we let the messages be forwarded in the network according to the different routing schemes, and

average the results. As already discussed in Section 3.2, the correlation between node interest profiles and their meeting frequencies is low (see first row of Table 2) without filtering out short meetings; on the other hand, filtering out short meetings to increase correlation would considerably reduce the size of the data set, making simulation results scarcely significant. In view of this, we have decided to keep the user population as large as possible (61 users, with a 0.08 meeting frequency correlation). The relevance threshold in all cases is chosen in such a way that (1) the relevant destinations of most of messages (more than 90 percent) are more than 1 to allow us to evaluate the multicast version of the protocols, and (2) the success rate of the Epidemic protocol is higher than 70 percent for $TTL = 1h$. In this trace the second condition is met for relevance thresholds lower than $\alpha = 0.45$. The relay threshold was then tuned to $\rho = 0.25$ as it was the lowest value that allowed for a differentiation in terms of performance between Epidemic and SANE-EP. Such values are used for both unicast and interest-cast.

5.1.1 Unicast

We compare the unicast version of SANE (UN-SANE) to the well known stateless routing protocols BinarySW [27] and Epidemic [28], and to a state-of-the-art of social-aware routing protocol, BUBBLE [12]. In implementing BUBBLE, we took care of putting the protocol in the best possible conditions, i.e., complete knowledge of the social graph and of the local/global ranking metrics. We consider both the spray and wait (UN-SANE SW) and the epidemic (UN-SANE) versions in our experiments. The parameter N^*_{replicas} (number of message copies) of BinarySW and UN-SANE SW is set to 4—this maximizes the performance of BinarySW according to the authors [27]. We then measure the *avg delay* (avg delivery time for successfully delivered messages), the *cost* (avg number of message copies in the network per generated message), and success rate in dependence of message TTL. For every metric we compute the confidence intervals as $[avg - \sigma; avg + \sigma]$, σ denoting the standard deviation (see Fig. 2).

As can be seen, both versions of UN-SANE provide significantly higher success rate than that of competing protocols (excluding, of course, Epidemic); also, the delay provided by the two versions of UN-SANE is better than that of both BinarySW and BUBBLE. The two versions of UN-SANE provide different performance/cost tradeoffs, with the SW version providing lower success rate among the two (around 60 percent instead of about 68 percent), but with a much lower cost (four times less cost). Note also that the cost of UN-SANE SW is about the same as that of BinarySW, and only slightly higher than that of BUBBLE for small TTLs.

5.1.2 Interest-Cast

Here, we show results related to the two interest-cast versions of our protocol: SANE SW, and SANE EP. Since there is no immediate way of extending BUBBLE into an interest-cast protocol, we compare SANE protocols only to Epidemic and BinarySW, whose interest-cast versions are straightforward (simply delivers a copy of the message to all relevant destinations, computed offline). The way we generate messages and the input tuning parameters of BinarySW and

(a) Success Rate
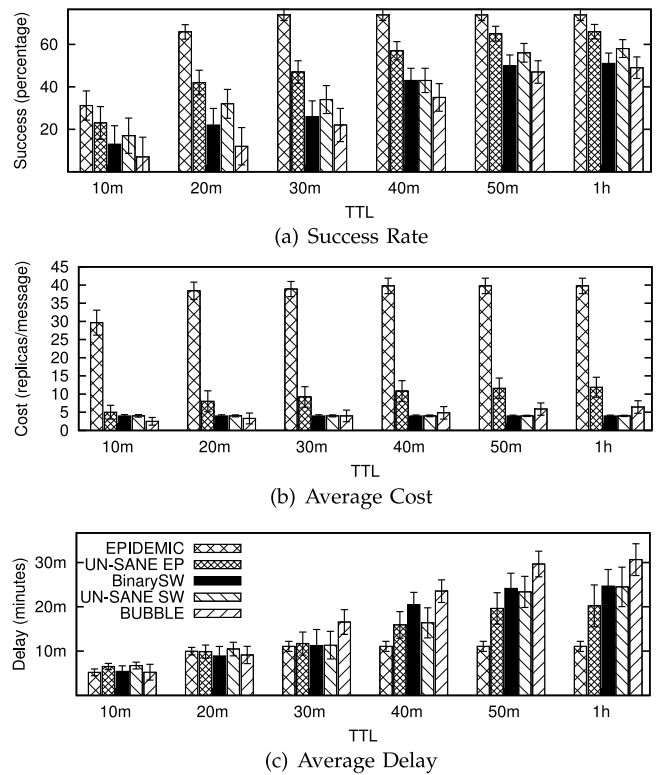
(b) Average Cost

(c) Average Delay

Fig. 2. Unicast on Infocom 06. Unlimited buffer.

SANE SW are the same as in the previous section. Again, for every metric we also compute the confidence intervals as $[avg - \sigma; avg + \sigma]$, σ denoting the standard deviation. The results are shown in Fig. 3. In this case, *coverage* refers to the percentage of relevant destinations holding a copy of the message when the TTL expires. Recall that relevant destination for a given message M are all the individuals such that the cosine similarity between their interest profile and the relevance of message M, $R(M)$, is larger than the relevance threshold α. As seen from the figures, SANE protocols perform very well, providing comparable coverage of relevant destinations to that of Epidemic (for TTLs > 30 min), but with a much reduced cost (as much as 10-fold cost reduction with respect to Epidemic, in case of SANE SW). The benefits of our proposed social-aware routing approach are evident comparing the relative performance of BinarySW and SANE SW: with a comparable cost, SANE SW provides higher coverage and lower delay as compared to BinarySW.

5.2 Evaluation Using SWIM Traces

The synthetic traces we use for evaluation have been obtained from the SWIM mobility model [22], [23], [24]. In SWIM, nodes are assigned a home point in the network area, assumed to be a square. Each time a node chooses its next destination, it trades off distance from its home point and popularity of the possible destinations. Thus, nodes with relatively close home points (*neighbors*) tend to go to the same locations and get in contact more often.

5.2.1 Experimental Setup

To run SANE on SWIM's traces, we first generate a network of 200 nodes and of an area of $1,000 \times 1,000$. Each node is
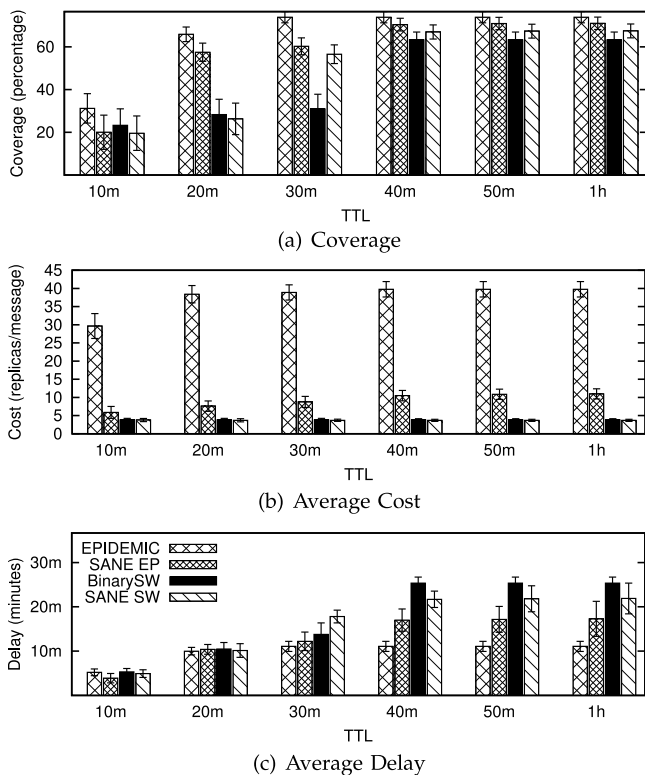
(a) Coverage

(b) Average Cost

(c) Average Delay

Fig. 3. Multicast on Infocom 06. Unlimited buffer.

(a) Success Rate

(b) Average Cost

(c) Average Delay

Fig. 4. Unicast on SWIM. Unlimited buffer.

equipped with a randomly generated four-dimensional interest-profile vector—the entries are chosen independently and uniformly at random in $[0, 1]$. Each profile vector is then normalized to 1—this way, we make sure that no node has very low interests or no interests at all.

In SWIM, neighbors tend to have a higher meeting rate. The amount of correlation between vicinity of home points and meeting rate in SWIM is controlled by a parameter η: The higher this parameter, the higher this correlation will be. Thus, tuning a relatively high meeting rate between nodes with similar profiles is easy: First we derive, for every node, its home point from the interest profile through a linear mapping, in such a way that nodes with similar profiles happen to be neighbors. This is done by using the first two coordinates of the profile as home-point coordinates. The correlation between profile similarity and home-point distances results very high (in our case it is —0.9). Then, we generate SWIM mobility traces, controlling the resulting correlation between node profile similarity and their meeting frequency by tuning SWIM's η parameter. The resulting correlation between interest profile similarity and pairwise meeting rates with these settings is about 0.7, allowing a wider range of variation for the relevance and relay threshold parameters of the SANE protocols.

SWIM allows us to implement a mobility-profile based protocol with which to compare our interest-profile based mechanisms.[1] For the comparison we chose MobySpace [31], [32]. Following the suggestion of the authors in [31],

[32], we build a mobile-profile for each node in SWIM containing the probabilities of visits in the different locations (SWIM's cells) of the network area during the training period (100 days long in our simulation). They are exploited by MobySpace rule [31] and MobySpace-Flooding rule [32] for respectively unicast and multicast to select relays with a mobility-profile that is closer to the destination's (according to the euclidean distance).

5.2.2 Unicast

We first compare the performance of UN-SANE protocols versus Epidemic, BinarySW, MobySpace (denoted with MS), and BUBBLE. Again, as in the previous scenario, the α parameter is chosen in such a way that the (1) relevant destinations of most of messages (more than 90 percent) are more than 1 to allow us to evaluate the multicast version of the protocols, and (2) the success rate of the Epidemic protocol is higher than 70 percent for TTL = 1h. Such conditions are met for $\alpha = 0.95$. Then, $\rho = 0.7$ is the lowest value to allow a differentiation between SANE-EP and Epidemic in terms of performance. Thus we chose this parameter. As can be noticed from Fig. 4, the advantages of UN-SANE protocols over competitors become even more evident than in Infocom 06 simulations: in particular, UN-SANE EP provides the *same* delivery performance as Epidemic starting from moderate TTL values, with a four-fold reduction in cost, and only a moderate increase in delivery delay. UN-SANE SW reduces the cost even more substantially than UN-SANE EP (more than 20-fold reduction over Epidemic), while only marginally sacrificing delivery performance (success probability as high as 95 percent). As compared to BinarySW, UN-SANE SW has

---

1. This is not possible with the Infocom 06 trace—it lacks information on the geographic locations of the user movements.

(a) Coverage
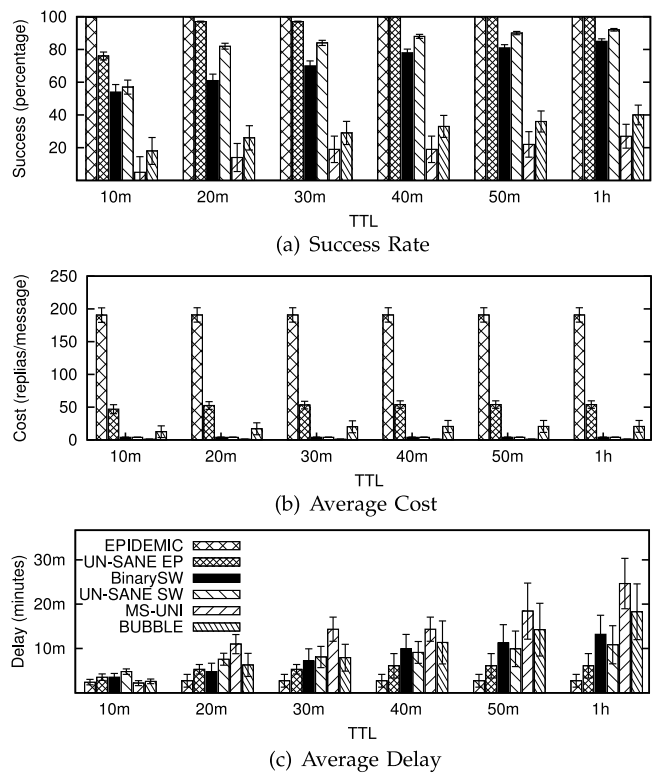

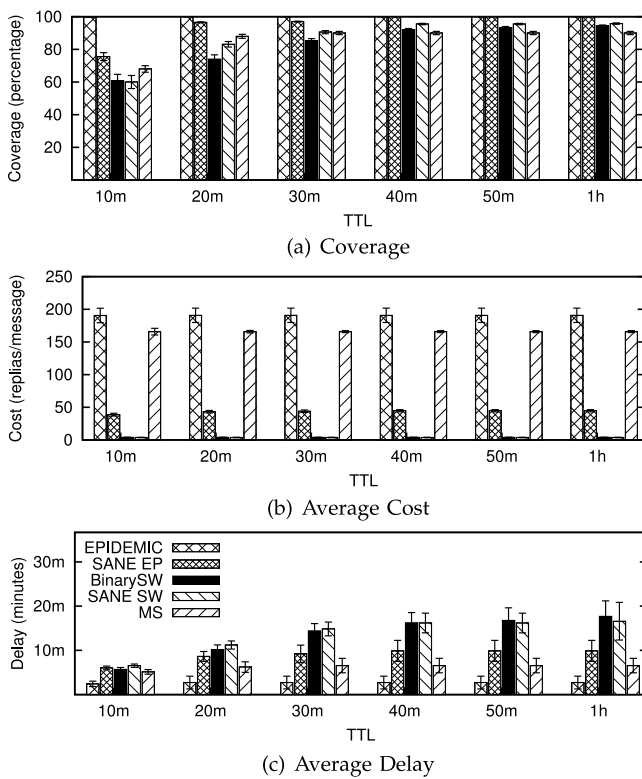(b) Average Cost

(c) Average Delay

Fig. 5. Multicast on SWIM. Unlimited buffer.

approximately the same cost and comparable delay, but substantially increases the success probability, thus confirming the effectiveness of our proposed social-aware routing approach. Finally, both UN-SANE and UN-SANE SW also over-perform their mobility-profile contestant MS in both terms of success percentage (above four times higher success rates for TTLs of 10 m, and above three times higher success rate for bigger TTLs) and delay (above two times faster for all TTLs). We believe this is due to the particularity of contact opportunities that SANE exploits, and that are missed by MobySpace (as discussed in the Related Works section). The cost of MS is contained as only one copy per message travels the network, however, it is similar to that of SANE-SW.

5.2.3 Interest-Cast

The results for the interest-cast versions of SANE, with similar settings as in the previous subsection, are reported in Fig. 5. Again, the advantages of SANE over competitors are evident: SANE EP provides nearly the same coverage as EP (*complete coverage* for TTL values $\geq$40 min), with an eightfold cost reduction w.r.t. Epidemic, at the price of a moderate increase in delay; SANE SW reduces cost even more (about 40-fold reduction w.r.t. Epidemic), at the price of a certain coverage decrease (which becomes negligible for TTL values $\geq$ 40 min), and a more substantial increase in delay. As in the case of unicast, the advantages of SANE SW over BinarySW are tangible, proving the validity of our social-aware, stateless routing approach. When it comes to MS in its controlled flooding version, we notice that it slightly out-performs SANE SW in terms of coverage for TTLs up to 20 min, then SANE SW takes

over for TTLs larger than 30 min. This is expected: the flooding flavor of MS does not have the restrictions of SW protocols in terms of number of copies. Thus, they reach their best performance sooner (see Fig. 5c), at the price of very high costs, comparable to that of Epidemic (see Fig. 5b). In any case, SANE EP always out-performs MS in terms of coverage, of at least 10 percent (see Fig. 5a) reducing to less than 1/3 the cost (see Fig. 5b).

Finally, recall that for both BUBBLE and MS we needed to pre-process the traces so to compute the community membership and node rankings (BUBBLE), and to gather the mobility-profile information on nodes during the training-period (for MS). Conversely, no pre-computation on the traces needed to be done for the stateless SANE.

5.3 Discussion

When collectively considered, the experimental results presented in this section clearly show the superiority of SANE protocols over both social oblivious, stateless and social-aware, stateful approaches. However, it is worth mentioning the following: For very low TTLs, it is very difficult for all protocols to reach the destinations in time, especially in the multicast routing with limited buffer size (as shown in the experimental results presented in the supplementary file, available online, of this manuscript). As long as the TTL increases, all protocols perform better and better, approaching an upper bound: The maximum success rate achievable, by that protocol, on the trace. The goodness of a given protocol is determined by how fast (in terms of TTLs) it approximates this upper bound. According to the experimental results, SANE is very efficient in quickly approaching its upper bound.

SANE protocols provide significantly higher average delay than Epidemic (the benchmark), though lower delays than other competitors; however, each packet is considered successfully received only if delivered within its TTL, a time which is deemed as acceptable for message delivery by the user sending the message. Thus, although the SANE average delivery is higher compared to Epidemic, it is considered as acceptable by the users, as packets are still delivered to destination(s) within the TTL.

Comparing the performance of (UN)SANE SW and BinarySW gives an empirical proof that the superiority of SANE is due to its stateless, social-aware routing mechanism: Both protocols have the same upper bound on the number of replicas circulating in the network, so, an almost identical message overhead, but a different routing mechanism. This difference makes (UN)SANE SW consistently performs better than BinarySW in terms of success rate and delay.

Finally, we deem important to discuss the limitations of our approach in terms of efficiency, overhead, and practicability. Both efficiency (success rate and delay) and networking overhead (transmissions) are affected, and thus limited, by the parameter ρ, which regulates the selectiveness of forwarding in SANE: As shown in the experiments presented in the supplementary file, available online, of this manuscript, the lower this parameter the less selective is the routing approach, but the higher the efficiency (lower delay and higher success rate), and vice versa. This parameter has to be carefully tuned to the highest possible value (lowest

possible overhead) that provides the required efficiency. In practical scenarios, one might start off with a very selective relay strategy (high ρ), and eventually switch to a less selective one (lower the ρ) as the message TTL approaches, so to guarantee delivery within the requested time. Another possibility in real scenarios is to exploit different ρ values for different messages, in dependence of their importance. As an alternative, one can always rely on SANE-SW, which has a limited transmission overhead (bounded by the number of replicas), but yields good results in terms of efficiency.

Regarding memory overhead, we note that SANE relies on interest-profiles (similar to IP addresses). So, we believe that the memory overhead to store the routing information can be of a few hundreds of bytes in practical contexts—not a severe issue on today's smartphones that feature several GBs of memory. Another aspect of practicability is related to the interest-profiles: Probably, in campus-like scenarios, the most adequate interest-profile is one including courses information; in a city-like scenario however, one might deem more important exploiting interest-profiles including information on the neighborhood, work, hobbies, and so on. In addition, consider that in many practical cases the interest-profile may come for free without user intervention (in a university campus it may suffice to use the list of courses the student has registered to along with the list of clubs and campus societies she belongs to as her interest-profile).

6   CONCLUSIONS

In this paper, we have first validated the intuition that individuals with similar interests tend to meet more often than individuals with diverse interests, and then used this intuition to design the first social-aware, stateless routing mechanism for opportunistic networks, called SANE. A nice feature of the SANE approach is that it can be used not only for traditional unicast communication, but also for realizing innovative networking services for PSNs, such as interest-casting. The results of extensive simulations based on both real-world and synthetic mobility traces have shown a clear superiority of our SANE approach over existing competitors. In particular, the comparison with BinarySW and MobySpace clearly shows the benefits of social-aware routing. Finally, in the supplementary file, available online, that accompanies this manuscript we include further experimental results that study the performance of SANE in networks with limited buffer size, in dependence of the α and ρ parameters, and in networks with low correlation among profile-similarity and meeting rates.

REFERENCES

[1]   C. Boldrini, M. Conti, and A. Passarella, "ContentPlace: Social-Aware Data Dissemination in Opportunistic Networks," *Proc. 11th Int'l Symp. Modeling, Analysis and Simulation of Wireless and Mobile System (MSWiM '08)*, 2008.

[2]   A. Lindgren, A. Doria, and O. Schelén, "Probabilistic Routing in Intermittently Connected Networks," *ACM SIGMOBILE Mobile Computing and Comm. Rev.*, vol. 7, pp. 19-20, 2003.

[3]   M. Claypool, D. Brown, P. Le, and M. Waseda, "Inferring User Interest," *IEEE Internet Computing*, vol. 5, no. 6, pp. 32-39, Nov./Dec. 2001.

[4]   P. Costa, C. Mascolo, M. Musolesi, and G.P. Picco, "Socially-Aware Routing for Publish-Subscribe in Delay-Tolerant Mobile Ad Hoc Networks," *IEEE J. Selected Areas in Comm.*, vol. 26, no. 5, pp. 748-760, June 2008.

[5]   G. Costantino, F. Martinelli, and P. Santi, "Privacy-Preserving Interest-Casting in Opportunistic Networks," *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC)*, 2012.

[6]   G. Costantino, F. Martinelli, P. Santi, and D. Amoruso, "An Implementation of Secure Two-Party Computation for Smartphones with Application to Privacy-Preserving Interest-Cast," *Proc. Mobicom*, 2012.

[7]   E. Daly and M. Haahr, "Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs," *Proc. ACM MobiHoc*, 2007.

[8]   J. Diaz, A. Marchetti-Spaccamela, D. Mitsche, P. Santi, and J. Stefa, "Social-Aware Forwarding Improves Routing Performance in Pocket Switched Networks," *Proc. 19th European Conf. Algorithms (ESA '11)*, 2011.

[9]   N. Eagle and A. Pentland, "Reality Mining: Sensing Complex Social Systems," *Personal and Ubiquitous Computing*, vol. 10, no. 4, pp. 255-268, 2006.

[10]   D. Eppstein, M.T. Goodrich, M. Loffler, D. Strash, and L. Trott, "Category-Based Routing in Social Networks: Membership Dimension and the Small-World Phenomenon," *Proc. Int'l Conf. Computational Aspects of Social Networks (CASoN)*, 2011.

[11]   W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in Delay Tolerant Networks: A Social Network Perspective," *Proc. ACM MobiHoc*, 2009.

[12]   P. Hui, J. Crowcroft, and E. Yoneki, "BUBBLE Rap: Social-based Forwarding in Delay Tolerant Networks," *Proc. ACM MobiHoc*, 2008.

[13]   P. Hui, E. Yoneki, S.Y. Chan, and J. Crowcroft, "Distributed Community Detection in Delay Tolerant Networks," *Proc. ACM Mobility in the Evolving Internet Architecture (MobiArch)*, 2007.

[14]   P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Pocket-Switched Networks and Human Mobility in Conference Environments," *Proc. ACM SIGCOMM Workshop Delay-Tolerant Networking (WDTN '05)*, 2005.

[15]   P. Hui and J. Crowcroft, "How Small Labels Create Big Improvements," *Proc. Fifth IEEE Int'l Conf. Pervasive Computing and Comm. (PERCOMW '07)*, Mar. 2007.

[16]   W.J. Hsu, D. Dutta, and A. Helmy, "Profile-Cast: Behavior-Aware Mobile Networking," *ACM SIGMOBILE Mobile Computing and Comm. Rev.*, vol. 12, no. 1, pp. 52-54, 2008.

[17]   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, pp. 1586-1602, 2012.

[18]   S. Ioannidis, A. Chaintreau, and L. Massoulie, "Optimal and Scalable Distribution of Content Updates over a Mobile Social Net," *Proc. IEEE INFOCOM*, Apr. 2009.

[19]   F. Li and J. Wu, "LocalCom: A Community-Based Epidemic Forwarding Scheme in Disruption-Tolerant Networks," *Proc. Sixth Ann. IEEE Comm. Soc. Conf. Sensor, Mesh and Ad Hoc Comm. and Networks (SECON '09)*, 2009.

[20]   F. Malandrino, M. Kurant, A. Markopoulou, C. Westphal, and U. Kozat, "Proactive Seeding for Information Cascades in Cellular Networks," *Proc. IEEE INFOCOM*, 2012.

[21]   M. McPherson, "Birds of a Feather: Homophily in Social Networks," *Ann. Rev. Sociology*, vol. 27, no. 1, pp. 415-444, 2001.

[22]   A. Mei and J. Stefa, "SWIM: A Simple Model to Generate Small Mobile Worlds," *Proc. IEEE INFOCOM*, 2009.

[23]   S. Kosta, A. Mei, and J. Stefa, "Small World in Motion (SWIM): Modeling Communities in Ad-Hoc Mobile Networking," *Proc. Seventh Ann. IEEE Comm. Soc. Conf. Sensor, Mesh and Ad Hoc Comm. and Networks (SECON '10)*, 2010.

[24]   S. Kosta, A. Mei, and J. Stefa, "Large-Scale Synthetic Social Mobile Networks with SWIM," *IEEE Trans. Mobile Computing*, vol. 13, no. 1, pp. 116-129, Jan. 2014.

[25]   A. Noulas, M. Musolesi, M. Pontil, and C. Mascolo, "Inferring Interests from Mobility and Social Interactions," *Proc. Workshop Analyzing Networks and Learning with Graphs (Colocated with NIPS '09)*, 2009.

[26]   G. Salton, A. Wong, and C.S. Yang, "A Vector Space Model for Automatic Indexing," *Comm. ACM*, vol. 18, no. 11, pp. 613-620, 1975.

[27]   T. Spyropoulos, K. Psounis, and C.S. Raghavendra, "Efficient Routing in Intermittently Connected Mobile Networks: The Multi-copy Case," *IEEE Trans. Networking*, vol. 16, no. 1, pp. 77-90, Feb. 2008.

[28]   A. Vahdat and D. Becker, "Epidemic Routing for Partially Connected Ad Hoc Networks," technical report, Duke Univ., 2000.

[29] J. Wu and Y. Wang, "Social Feature-Based Multi-Path Routing in Delay Tolerant Networks," *Proc. IEEE INFOCOM*, 2012.

[30] J. Wu and Y. Wang, "Hypercube-Based Multi-Path Social Feature Routing in Human Contact Networks," *IEEE Trans. Computers*, vol. 63, no. 2, pp. 383-396, Feb. 2014.

[31] J. Leguay, T. Friedman, and V. Conan, "Evaluating Mobility Pattern Space Routing for DTNs," *Proc. IEEE INFOCOM*, 2006.

[32] J. Leguay, T. Friedman, and V. Conan, "Evaluating MobySpace-Based Routing Strategies in Delay-Tolerant Networks," *Wireless Comm. and Mobile Computing—Wireless Ad Hoc and Sensor Networks*, vol. 7, pp. 1171-1182, 2007.

[33] H.A. Nguyen, S. Giordano, and A. Puiatti, "Probabilistic Routing Protocol for Intermittently Connected Mobile Ad Hoc Network," *Proc. IEEE Int'l Symp. World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2007.

[34] C. Boldrini, M. Conti, and A. Passarella, "Exploiting Users Social Relations to Forward Data in Opportunistic Networks: The HiBOp Solution," *Pervasive and Mobile Computing*, vol. 4, no. 5, pp. 633-657, 2008.

[35] E. Jaho, M. Karaliopoulos, and I. Stavrakakis, "ISCoDe: A Framework for Interest Similarity-Based Community Detection in Social Networks," *Proc. IEEE Third Int'l Workshop Network Science for Comm. Networks (NetSciCom)*, 2011.

[36] T. Hossmann, G. Nomikos, T. Spyropoulos, and F. Legendre, "Collection and Analysis of Multi-Dimensional Network Data for Opportunistic Networking Research," *Computer Comm.*, vol. 35, pp. 1613-1625, 2012.

[37] S. Scellato, A. Noulas, and C. Mascolo, "Exploiting Place Features in Link Prediction on Location-Based Social Networks," *Proc. 17th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD)*, 2011.

[38] J. Golbeck, "Trust and Nuanced Profile Similarity in Online Social Networks," *ACM Trans. Web*, vol. 3, no. 4, article 12, 2009.

[39] W. Gao and G. Cao, "User-Centric Data Dissemination in Disruption Tolerant Networks," *Proc. IEEE INFOCOM*, 2011.

[40] A. Mei, G. Morabito, P. Santi, and J. Stefa, "Social-Aware Stateless Forwarding in Pocket Switched Networks," *Proc. IEEE INFOCOM*, 2011.

Alessandro Mei received the laurea degree in computer science, summa cum laude, from the University of Pisa, Italy, in 1994. He continued his studies as a PhD student at the Department of Mathematics of the University of Trento, Italy, and as a visiting scholar at the Department of EE-Systems of the University of Southern California during 1998 and part of 1999. He received the PhD degree in mathematics from the University of Trento in 1999. After a postdoctoral position at the University of Trento, in 2001 he joined the faculty of the Computer Science Department at the Sapienza University of Rome, Italy. His main research interests include computer system security and parallel, distributed, and networked systems. He was presented with the Best Paper Award of the 16th IEEE International Parallel and Distributed Processing Symposium in 2002, the EE-Systems Outstanding Research Paper Award of the University of Southern California for 2000, and the Outstanding Paper Award of the Fifth IEEE/ACM International Conference on High Performance Computing in 1998. He is a past associate editor of the *IEEE Transactions on Computers* (2005-2009), and the general chair of IEEE IPDPS 2009, Rome, Italy. He is a former Marie Curie Fellow and a 2013 Google Faculty Research Grant Awardee. He is a member of the ACM and the IEEE.

Giacomo Morabito received the laurea degree in electrical engineering and the PhD degree in electrical, computer and telecommunications engineering from the Istituto di Informatica e Telecomunicazioni, University of Catania, Italy, in 1996 and 2000, respectively. From November 1999 to April 2001, he was with the Broadband and Wireless Networking Laboratory of the Georgia Institute of Technology as a research engineer. Since April 2001, he has been with the Dipartimento di Ingegneria Elettrica, Elettronica e Informatica of the University of Catania where he is currently an associate professor. His current research interests include analysis and protocols for wireless networks and microfluidic networks.

Paolo Santi received the Laurea degree and PhD degree in computer science from the University of Pisa in 1994 and 2000, respectively. He is a senior researcher at the Istituto di Informatica e Telematica del CNR in Pisa, Italy. He visited the Georgia Institute of Technology in 2001, and Carnegie Mellon University in 2003. His research interests include fault-tolerant computing in multiprocessor systems (during PhD studies), and the investigation of fundamental properties of wireless multihop networks such as connectivity, topology control, lifetime, capacity, mobility modeling, and cooperation issues. He has contributed more than 70 papers and a book in highly reputed conferences and journals in the field of short range wireless networking and mobile computing. He is the recipient of the 2004 ITG Outstanding Paper Award. He was the general co-chair of ACM VANET 2007 and 2008, technical program co-chair of IEEE WiMesh 2009 He is involved in the organizational and technical program committee of several conferences in the field. He has been a guest editor of the *Proceedings of the IEEE*, and he is an associate editor for the *IEEE Transactions on Mobile Computing and IEEE Transactions on Parallel and Distributed Systems*. He is a member of the IEEE Computer Society, and a senior member of ACM and SIGMOBILE.

Julinda Stefa received the laurea degree in computer science, summa cum laude, and the PhD in computer science from the Sapienza University of Rome, Italy, in July 2006 and February 2010, respectively. She is a post-doc at the Computer Science Department of the Sapienza University of Rome. In 2005, she joined Google Zurich for three months as an engineering intern. She was a visiting scholar at the CS Department of UNC-Chapel Hill from November 2008 to April 2009, and a research intern at Microsoft Research, Cambridge, United Kingdom, from January to April in 2011. Her research interests include computer systems and network security, parallel and distributed systems, and analysis and modeling of social mobile wireless networks. She received a 2011 research grant from Working Capital PNI and offered by Telecom Italia (30 winners out of 2138), and the IEEE INFOCOM 2013 and IEEE SECON 2013 Best Demo Award.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.