

A Survey of Social-Based Routing in Delay Tolerant Networks: Positive and Negative Social Effects

Ying Zhu, Bin Xu, *Member, IEEE*, Xinghua Shi, and Yu Wang, *Senior Member, IEEE*

Abstract—*Delay tolerant networks (DTNs) may lack continuous network connectivity. Routing in DTNs is thus challenging since it must handle network partitioning, long delays, and dynamic topology in such networks. In recent years, social-based approaches, which attempt to exploit social behaviors of DTN nodes to make better routing decision, have drawn tremendous interests in DTN routing design. In this article, we summarize the social properties in DTNs, and provide a survey of recent social-based DTN routing approaches. To improve routing performance, these methods either take advantages of positive social characteristics such as community and friendship to assist packet forwarding or consider negative social characteristics such as selfishness. We conclude by discussing some open issues and challenges in social-based approaches regarding the design of DTN routing protocols.*

Index Terms—DTN routing; Social-based approaches; Social graphs; Social network analysis; Delay tolerant networks.

I. INTRODUCTION

DELAY or disruption tolerant networks (DTNs) [1]–[3] have recently drawn much attention from networking researchers due to the wide applications of these networks in challenging environments, such as space communications, military operations, and mobile sensor networks. Intermittent connectivity in DTNs results in the lack of instantaneous end-to-end paths, large transmission delay and unstable network topology. These characteristics make the classical ad hoc routing protocols [4]–[6] not being applicable for DTNs, since these protocols rely on establishment of a complete end-to-end route from the source to the destination.

Many routing schemes [7]–[20] have been proposed for DTNs. Most of these DTN routing protocols belong to three categories: *message-ferry-based*, *opportunity-based* and *prediction-based*. In message-ferry-based methods [8]–[11], systems usually employ extra mobile nodes as ferries for message delivery. The trajectory of these ferries is controlled to improve delivery performance with store-and-carry. However, controlling these nodes leads to extra cost and overhead. In opportunity-based schemes [3], [12], [13], nodes forward

messages randomly hop by hop with the expectation of eventual delivery, but with no guarantees. Generally, messages are exchanged only when two nodes meet at the same place, and multiple copies of the same message are flooded in the network to increase the chance of delivery. Some DTN routing protocols [14]–[19] make relay selection by estimating metrics relative to successful delivery, such as delivery probability or expected delay based on a history of observations. Most of these protocols focus on whether two nodes will have a contact and when such contact happens if they do contact. Liu and Wu [20] also proposed a forwarding method based on a probabilistic forwarding metric, which is derived by modeling each forwarding as an optimal stopping rule problem.

All of the current DTN routing methods share a similar paradigm, the “store and forward” fashion. If there is no connection available at a particular time, a DTN node can store and carry the data until it encounters other nodes. When the node has such a forwarding opportunity, all encountered nodes could be the candidates to relay the data. Thus, relaying selection and forwarding decision need to be made by the current node based on certain routing strategy. Various DTN routing approaches adopt different strategies based on different metrics. Example of such metrics include estimated delivery probability to the destination node, network resources available (including bandwidth, storage, and energy), estimated delay, and current network congestion level. However, the unpredictable mobility and restricted resource in DTNs significantly obstruct us from designing an ideal forwarding mechanism.

Lately, the consideration of social characteristics provides a new angle of view in the design of DTN routing protocols. In most of the DTN applications (e.g. vehicular networks [21], [22], mobile social networks [23]–[27], disease epidemic spread monitoring and pocket switched networks (PSNs) [28]), a multitude of mobile devices are used and carried by people, whose behaviors are better described by social models. This opens new possibilities of social-based DTN routing, in which the knowledge of social characteristics are used to make better forwarding decision in DTN routing. Notice that social relations and behaviors among mobile users are usually long-term characteristics and less volatile than node mobility. Based on this observation and taking the recent advances in social network analysis, several social-based DTN routing methods [29]–[34] have been proposed recently to exploit various social characteristics in DTNs (such as community and centrality)

Manuscript received 4 January 2011; revised 10 June 2011 and 5 December 2011.

Y. Zhu and Y. Wang are with Department of Computer Science, the University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223, USA (e-mail: {yzhu17,yu.wang}@unc.edu).

B. Xu is with the Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China.

X. Shi is with Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts 02215, USA.

Digital Object Identifier 10.1109/SURV.2012.032612.00004

to assist the relay selections. Hereafter, we call these social characteristics *positive social characteristics* since correctly using them can improve the DTN routing performance.

Meanwhile, there is the other side of the story. Since mobile nodes in DTNs are owned and controlled by rational entities, such as people or organizations, these nodes usually behave selfishly, i.e., attempt to maximize their own utilities and conserve their limited resources during the operation. Carrying and forwarding packets for others in DTN routing consumes their own energy and other resources, and thus selfish nodes may be reluctant to serve as relays for others. Since routing is an inherently cooperative activity, routing operation can be critically impaired by selfish behaviors. Motivated by this observation, several social-based DTN routing protocols [35]–[40] have been proposed to take users' selfish behaviors into consideration by either punishing misbehaving nodes or stimulating cooperations among nodes. We call such a social characteristic of selfishness *negative social characteristic*.

In this article, we survey the utilizations of both positive and negative social characteristics in the design of social-based DTN routing protocols. The rest of the paper is organized as follows. Section II reviews recent advances in social network analysis, and Section III introduces several social properties related to DTNs routing. Detailed social-based DTN routing protocols, considering positive and negative social characteristics, are reviewed and analyzed in Section IV and Section V respectively. Section VI concludes this article with a summary of all methods aforementioned and a discussion of potential future work.

II. SOCIAL NETWORK ANALYSIS

Social network analysis (SNA) [41], [42] has attracted a significant attention in many research areas such as anthropology, biology, communication studies, economics, information science, computer science and engineering. SNA mainly focuses on studying relationships among social entities and the patterns and implications of these relationships. With the increasing popularity of online social networks and new information technologies (such as mobile computing, E-commerce, distributed systems, and smart sensing), SNA becomes a more powerful tool to study the relationships and ties among users, and thus may guide the design of new policies, protocols, or applications for different information systems. In this section, we briefly review some recent advances in SNA and its applications in information systems, with a focus on four major aspects: *community detection*, *information propagation*, *recommendation system*, and *security and privacy*. Some of these advances will become the basis of social-based approaches for DTN routing.

Community detection and the study of social structures are among the most fundamental problems in social network analysis. In recent years, it has been well studied regarding how to effectively and accurately discovering communities and clusters in social networks. Here, we review some of the proposed methods for these problems. Girvan and Newman [43] introduced a community detection method which uses centrality indices to find community boundaries. By removing the edges with high edge betweenness (a type of centrality

measures), it separates the communities from one another. In [44], Newman proposed a modularity-based community detecting algorithm which uses the metric of quality function “modularity”. Modularity is defined as a benefit function that measures the quality of a particular division of a network into communities. The community structure for a network can be found by searching for network divisions that have positive and preferably large values of modularity. Blondel *et al.* [45] proposed a greedy community detection algorithm, which consists of two phases. The first phase looks for small communities by optimizing modularity locally and the second phase aggregates nodes of the same community and builds a new network whose nodes are these identified communities. The two steps are repeated iteratively until a maximum of modularity is attained. Other types of methods of finding community structures include minimum-cut method, hierarchical clustering, clique-based methods, and so on. More detailed review on these methods could be found in [46]. With the increasing availability of large social network data (especially online social network data), much such recent research focuses on how to detect communities more efficiently in large-scale social networks. Satuluri *et al.* [47] proposed a method to sparsify a graph to enable faster graph clustering. This method ranks the edges using a simple similarity-based heuristic and selects the top few edges to be retained in the sparsified graph so that the edges that are likely to be part of the same cluster are retained. Xu *et al.* [48] introduced several methods to identify online communities with similar sentiments in online social networks, which aim to help companies with market segmentation and design of marketing strategies. In [49], Liu and Murata studied community detection in heterogeneous systems, and their method is competent for not only one-to-one correspondence but also many-to-many correspondences. Nguyen *et al.* [50] studied how to detect the overlapping community structures in a dynamic network, with a focus on the adaptive update of community structures for mobile applications. There are also several recent studies on how to evaluate the communities. Leskovec *et al.* [51] compared several community detection methods by treating community quality as a function of size. Mitzlaff *et al.* [52] presented an approach for the evaluation of communities using implicit information. Almeida *et al.* [53] studied five major quality metrics for graph clustering and proved that all of them have strong biases.

Information propagation in social networks is another key topic in social network analysis. Information propagation has been used in epidemiology to help understand how patterns of human contact aid or inhibit the spread of diseases such as HIV in a population. Such a study is very relevant to data dissemination or data forwarding in communication networks (such as routing packets in delay tolerant networks). Since we will cover details on social-based DTN routing in later sections, here we just sample a few recent studies of information propagation in general (online) social networks. Lind *et al.* [54] studied a simple model of information propagation in social networks, by introducing the concepts of spread factor (the average maximal fraction of neighbors of a given node that interchange information among each other) and spreading time (i.e. the time needed for the information to

reach such a fraction of nodes). They applied this model to real empirical networks and compared spreading dynamics with different types of networks. They found that the number of neighboring connections strongly influences the probability of being gossiped. Yildiz *et al.* [55] considered the problem of asymmetric information diffusion with gossiping protocols in both static and dynamic networks. They derived conditions under which the network converges to the desired result within limit, and provided policies that offers a trade-off between accuracy and increased mixing speed for the dynamic asymmetric diffusion problem. In [56], [57], Tang *et al.* proposed new temporal distance metrics to quantify and compare the speed of information diffusion with the consideration of the evolution of a network from a local and global view. Lee *et al.* [58] proposed a method to find influentials by considering link structure and the temporal order of information adoption in Twitter. In [59], Zhao *et al.* used communication motifs and maximum-flow communication motifs as the tools to characterize the patterns of information propagation in two real-life social networks (networks from cellular call record and Facebook wall-post history). They concluded that the patterns of information propagation within both social networks are stable overtime, but these patterns are different and sensitive to the cost of communication in synchronous and asynchronous social networks. The speed and the amount of information propagated through a network are correlated and dependent on individual profiles. In [60], Bakshy *et al.* studied the content propagation via user-to-user content transfer history in a time-evolving social network (Second Life). They found that the social network plays a significant role in the propagation of content. Additionally, adoption rate increases as the number of adopting friends increases, but this effect varies with the connectivity of a particular user. They also found that sharing among friends occurs faster than sharing among strangers and some individuals play a more active role in distributing content than others. Kuhlman *et al.* [61] studied the problem of inhibiting diffusion of complex contagions such as rumors, undesirable fads and mob behavior in social networks by removing a small number of critical nodes from the network. They showed that finding minimum number of such nodes is NP-hard, and proposed efficient heuristics for such tasks. All of the studies above confirm that social structures and properties indeed strongly influence information propagation. These observations inspire the development of social-aware routing protocols for different communication networks as well.

Online social networks are emerging as useful resources which provide enriched social related information, such as social relationships, patterns, and interests of users, for different applications. Via social network analysis, various social properties can be extracted from online social networks and applied to diverse applications. Such examples include online recommendation systems [62]–[66] (which recommend contents or products based on user’s social property and relationship) and search engines [67], [68] (which match online content with user’s social interests). In [64], Tyler *et al.* studied user behaviors of an online review website for market analysis. By using user behavior information, such information as the category of a new product as well as which users will

follow the produce can be more accurately predicted. Tian *et al.* [66] introduced a graph predication-based recommendation strategy, *link revival*, that suggests users re-connect with their old friends. Their method selects proper candidates based on the prediction of their future behaviors and the resulted connections improve social network connectivity. In [62], [63], Phelan *et al.* proposed a recommendation system *Buzzer*, which uses a content-based approach to rank RSS news by mining trending terms from public or private Twitter timelines. In [65], Lopes *et al.* presented an approach to recommend collaborations on the context of academic social networks. Zhao *et al.* [67] studied ranking algorithms for microblogs (namely Twitter) and their search algorithm takes six social network properties into consideration. These social properties measure users’ social influences in Twitter. Choudhury *et al.* [68] also studied topic-based content exploration in Twitter. They proposed a greedy iterative clustering algorithm to select a set of items on a given topic that matches a specified level of diversity. In order to identify the ‘best’ or ‘right’ set of items, they align the characteristics of information space including specific content or social attributes and the information diversity of the results set, with measurements of human information processing.

Finally, security and privacy issues are essential for social network applications. Recently, there have been many new research studies in this area, such as anti-spam [69], [70] in social media and access control [71] and privacy preservation [72]–[74] in (online) social networks. Here we list a few examples. Grier *et al.* [69] studied the characterization of spams on Twitter and found that Twitter is a highly successful platform for coercing users to visit spam pages. Without spam filtering, they found that blacklists are too slow at identifying new threats in Twitter. Vanetti *et al.* [70] then proposed a content-based message filtering for online social networks, which allows a user to customize the filtering criteria with a machine learning based soft classifier automatically labeling messages. In [71], Hu and Ahn proposed a multiparty authorization control model which enables collaborative management of shared data in online social networks. Liu and Terzi [74] considered the privacy issue in online social networks from an individual user’s viewpoint. They calculated the privacy score of a user, which indicates the user’s privacy risk, based on his or her participation in the network. Fang and Lefevre [73] proposed a template for the design of a social networking privacy wizard. The wizard is built based on an active learning paradigm which iteratively asks the user to assign privacy ‘label’ to selected friends, and uses this input to construct a classifier. The classifier can be used to automatically assign privileges to the rest of the user’s friends. Narayanan and Shmatikov [72] presented a framework for analyzing privacy and anonymity in social networks and developed a re-identification algorithm for anonymized social networks.

III. SOCIAL PROPERTIES IN DTNS

In this section, we introduce some social properties related to DTNs routing. They can be categorized into *positive properties* which benefit the relay selection (e.g. community, centrality, similarity and friendship) and *negative properties* which hurt the network performance (e.g. selfishness). Many

of these social properties have been recently studied in social network analysis.

A. Social Graph and Contact Graph

The most popular way, to study the social relations among people and extract their social properties, is building a *social graph* (also called social network). A social graph is a global mapping of everybody and how they are related. Such a graph is an abstract graph where vertices represent individual people and edges describe social ties between individual people. Social ties can be expressed in many forms. For example, different types of social ties may describe different social relationships among people such as friends, family members, and co-workers. Social graphs have been widely used in many applications, such as analysis of online social networks [75] or terrorist networks [76]. With a social graph, a variety of social metrics (e.g., communality, centrality, and similarity) can be easily calculated or estimated, and these metrics can be then used by social-based approaches. Therefore, it is crucial to obtain social graphs for social-based approaches.

A social graph is an intuitive source for many social metrics such as community and friendship. Unfortunately it is not always available (due to either privacy or security reasons) or hard to be obtained via disclosed social data. However, with new networking technology, we can study relationships among people by observing their interactions and interests over wireless networks. Building a *contact graph* is a common way to study the interactions among people in a network and thus analyze their relationships and estimate the social metrics among them. In DTNs, each possible packet forwarding happens when two mobile nodes are in contact (i.e., within transmission range of each other). By recording contacts seen in the past, a contact graph can be generated where each vertex denotes a mobile node (device or person who carries the device) and each edge represents one or more past meetings between two nodes. An edge in this contact graph conveys the information that two nodes encountered each other in the past. Thus the existence of an edge intends to have predictive capacity for future contacts. A contact graph can be constructed separately for each single time slot in the past, or it can be constructed to record the encounters in a specific period of time by assigning a set of parameters to each edge to record the time, the frequency and the duration of these encounters. From the observation that people with close relationships such as friends, family members, etc. tend to meet more often, more regular and with longer duration, we can extract DTN nodes' relationships from the recorded contact graph, estimate their social metrics, and use such information to choose relays with higher probabilities of successful forwarding.

How to detect people's relationships and create the relative social graph from the recorded contact graph may affect estimation accuracy and the efficiency of social-based approaches. Most of the current social-based DTN routing algorithms [28], [30], [31] directly treat the aggregated contact graph (merging the contact graphs of several time slots into one graph) as the social graph of all entities in the network, and uses this graph to generate social metrics for forwarding selection. This strategy is based on the observation that although the contact

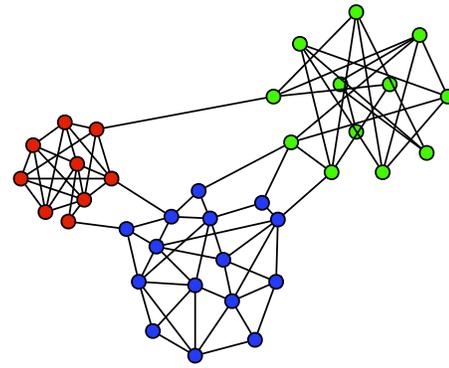


Fig. 1. Illustration of community structures in a contact/social graph. Three colors represent three community structures.

graph reflects the encounter history while the social graph reflects the social relations among people, the aggregated contact graph (the sum of contact graph over time) and the social graph are statistically similar. However, Hossmann *et al.* [77], [78] showed that the performance of these algorithms heavily depends on the way the graph is constructed out of observed contacts (i.e., contact aggregation) and proposed a method to select an appropriate aggregation period for contact aggregation. After building the aggregated contact graph, different social metrics can be obtained. For example, Hui, *et al.* [79]–[82] proposed several community detection approaches (simple, k -clique, modularity, etc.) with great potential to detect both static and temporal communities. Bulut *et al.* [34] introduced a method of detecting the quality of friendship by calculating the social pressure metric (SPM) from contact graphs.

B. Community

Community is an important concept in ecology and sociology [83]–[85]. In ecology, a community is an assemblage of two or more populations of different species occupying the same geographical area. In sociology, community is usually defined as a group of interacting people living in a common location. Community ecologists and sociologists study the interactions between species/people in communities at many spatial and temporal scales [44], [83]–[86]. It has been shown that a member of a given community is more likely to interact with another member of the same community than with a randomly chosen member of the population [86]. Therefore, communities naturally reflect social relationship among people.

Since wireless devices are usually carried by people, it is natural to extend the concept of social community into DTNs to explore interactions among wireless devices. It is believed that devices within the same community have higher chances to encounter each other. Figure 1 illustrates an example of three community structures in a social/contact graph. Therefore, the knowledge of community structures could help a routing protocol to choose better forwarding relays for particular destinations, and hence improve the chance of delivery. Many proposed community detection algorithms [50], [79]–[82] are available for identifying social communities from the contact graph of DTNs.

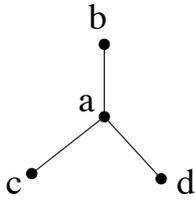


Fig. 2. Illustration of three centrality measures over a simple graph.

C. Centrality

In graph theory and network analysis, *centrality* is a quantitative measure of the topological importance of a vertex within the graph. A central node, typically, has a stronger capability of connecting other nodes in the graph. In a social graph, the centrality of a node describes the social importance of its represented person in the social network. In DTNs, the sociological centrality metrics [87] can also be used for relay selections (nodes with high centralities are always good candidates of relay nodes).

There are several ways to define centrality in a graph. Three common centrality measures are *degree centrality*, *betweenness centrality*, and *closeness centrality* [88]–[90]. Degree centrality is the simplest centrality measure which is defined as the number of links (i.e., direct contacts) incident upon a given node. For example, in Figure 2, the degree centrality of node *a* is 3 while those of the other nodes are 1. A node with a high degree centrality is a popular node with a large number of possible contacts, and thus it is a good candidate of a message forwarder for others (i.e., a hub for information exchange among its neighborhood). Betweenness centrality measures the number of shortest paths passing via certain given node. For example, the betweenness centrality of node *a* in Figure 2 is 6, since every shortest path passes through it. But for the other nodes, their betweenness centralities are 3. Nodes that occur on many shortest paths between other nodes have higher betweenness than those that do not. A node with high betweenness centrality can control or facilitate many connections between other nodes, thus it is ideal for a bridge node during message exchange. The closeness centrality of a node is defined as the inverse of its average shortest distance to all other nodes in the graph. If a node is near to the centre of the graph, it has higher closeness centrality and is good for quickly spreading messages over the network. For the example in Figure 2, the closeness centrality of node *a* is 1 since its average shortest distance to all others is 1. For any of the other nodes, the closeness centrality is 0.6 since their average shortest distance is $\frac{5}{3}$.

D. Similarity

Similarity [30] is a measurement of the degree of separation. It can be measured by the number of common neighbors between individuals in social networks. Sociologists have long known that there is a higher probability of two people being acquainted if they have one or more other acquaintances in common. In a network, the probability of two nodes being connected by a link is higher when they have a common neighbor. When the neighbors of nodes are unlikely to be in contact with each other, diffusion can be expected to take

longer than when the similarity is high (with more common neighbors). In addition, there are other ways to define the similarity beyond common neighbors, such as similarity on user interests [91] and similarity on user locations [92].

E. Friendship

Friendship is another concept in sociology which describes close personal relationships. In DTNs, friendship can be defined between a pair of nodes. On the one hand, to be considered as friends of each other, two nodes need to have long-lasting and regular contacts. On the other hand, friends usually share more common interests as in real world. In sociology, it has been shown that individuals often befriend others who have similar interests, perform similar actions and frequently meet with each other [93]. This observation is called *homophily phenomenon*. Therefore, the friendship in DTNs can be roughly determined by using either contact history between two nodes [34] or common interests/contents claimed by two nodes [33].

F. Selfishness

Selfishness has been well-studied in sociology and economics [94], and has recently been considered in design of computer networks [95]–[103]. In DTNs, selfishness can describe the selfish behaviors of DTN nodes controlled by rational entities. Selfish nodes can behave selfishly at individual level and aim to only maximize their own utilities without considering system-wide criteria. They can also behave selfishly in a social sense and are willing to forward packets for nodes with whom they have social ties but not the others. A selfish DTN node may drop others' messages and excessively replicate its own messages to increase its own delivery rate while significantly degrading other users' performance or even cause starvation [104].

IV. TAKING ADVANTAGE OF POSITIVE SOCIAL CHARACTERISTICS IN DTN ROUTING

In this section, we review several social-based DTN routing methods that take advantage of positive social characteristics in DTN networks.

A. Label Routing

Hui and Crowcroft [29] introduced a routing method (called as label routing hereafter) based on community labels in Pocket Switched Networks (PSNs). A PSN [28] is a type of DTN where mobile devices are carried by people and communicate with each other when people meet. To reduce the amount of traffic created by forwarding messages in PSNs, the proposed routing method uses a labeling strategy to select forwarding relay. Since people in the same community are likely to meet regularly, they are appropriate forwarders for messages destined to the members of their community. In their solution, Hui and Crowcroft assumed that each node has a small label telling others about its affiliation/group (i.e., its social community), just like name badges used in a conference. Based on the labels, label routing chooses to forward messages

to destinations directly or to next-hop nodes which belong to the same group (label) with the destinations.

Label routing takes the advantage of the knowledge of social community. It assumes that people from the same community tends to meet more often than people from different communities and hence can be good forwarders to relay messages destined to the other members in the same community (with the same label). Label routing requires very little information about each individual (only its group/affiliation). This is easy to implement in PSN applications, by tapping a mobile device and writing down the affiliation of the owner. In other words, the community (or group) information relies on user inputs in label routing. However, user-defined communities may not always reflect the position/contact relationship among nodes. For example, two DTN nodes in the same community may be physically far away and could never meet with each other. In this scenario, using one node to be the forwarder for the other may not be a good choice. In addition, in label routing, the message forwarding from the source to the destination is purely via the members within the same community of the destination. This may significantly increase the delay or even fail to deliver the message. For instance, message delivery will fail when the source does not meet any member from the destination's community, even though there are possible relay nodes from other communities.

B. SimBet Routing

Daly and Haahr [30] proposed a social-based routing protocol (called SimBet routing hereafter) which uses *betweenness centrality* and *similarity* metrics to identify some "bridge" nodes (with high values of these metrics) in networks. To avoid exchanging information of the entire network topology, they only estimated the betweenness centrality Bet_n for each node n in its local neighborhood. For similarity metric, they considered the similarity $Sim_n(d)$, the number of common neighbors, of the current node n with the destination node d . Both of the social metrics are maintained and updated dynamically in DTNs. Therefore, the proposed SimBet routing makes forwarding decision by considering not only the pre-estimated betweenness centrality metric but also the locally determined social similarity. Nodes with high betweenness centralities are those nodes who can act as bridges in their neighborhood, while nodes with high similarities with the destination are more likely to find a common neighbor with the destination which can act as the forwarder.

In SimBet routing, when a DTN node n meets another DTN node m and holds a message with destination d , n calculates its relative betweenness utility and similarity utility to node m :

$$SimUtil_n = \frac{Sim_n(d)}{Sim_n(d) + Sim_m(d)}$$

$$BetUtil_n = \frac{Bet_n}{Bet_n + Bet_m}$$

Then node n can compute its SimBet utility, which is a weighted combination of betweenness utility and similarity utility:

$$SimBetUtil_n(d) = \alpha SimUtil_n(d) + (1 - \alpha) BetUtil_n.$$

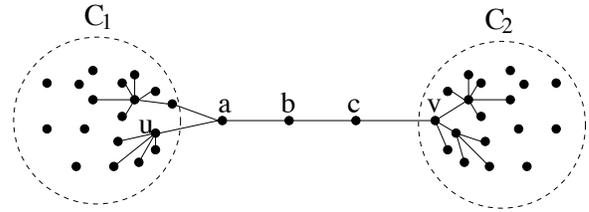


Fig. 3. Illustration of problems of the SimBet routing.

Here, α is a tunable parameter which can adjust the relative importance of the two utilities. For the message with d as its destination, if $SimBetUtil_m(d) > SimBetUtil_n(d)$, node n forwards the message to node m . Otherwise, it continues to hold the message. Via possible multi-hop relays, the message may eventually reach d .

In summary, SimBet routing uses two social metrics (centrality and similarity) to estimate or predict the probability that potential relay nodes may meet the destination. It is obvious that both metrics are effective at identifying suitable relays in different scenarios respectively. Take an example graph, as shown in Figure 3, where a few low-degree bridges (i.e., a , b and c) connect two well-connected components C_1 and C_2 . Assume that node u wants to send a message to node v . When node u encounters node a , it compares its SimBet utility with that of node a 's. Both u and a have zero similarity to v , but u 's global betweenness centrality is less than a 's since a sits on more of the shortest paths. Thus, u will transfer the message to a based on SimBet routing. In this case, centrality metric helps to pick the better relay node. On the contrary, if node a wants to send a message to v and it encounters node b , similarity metric will play a role since the global betweenness centralities of a and b are the same. Therefore, a has a smaller similarity (zero common neighbor) to v than b has (one common neighbor with v). Therefore, combining multiple social metrics may make the social-based protocol more effective in broad situations. However, due to the uncertainty of future encounters and underlying social graph, it is still possible that the node with high SimBet utility fails to delivery the message to the destination.

To avoid global information exchanges, SimBet routing provides a distributed method to calculate social metrics locally, which is desirable in a DTN environment. However, estimating centrality based solely on local information may lead to inaccurate "bridge" identification. For instance, in the example shown in Figure 3, it is assumed that u wants to send a message to v . When u encounters node a , based on the two-hop information, u 's local betweenness Bet_u is much larger than a 's Bet_a . Since both u and a have zero similarities to v , the overall $SimBetUtil_u(v) > SimBetUtil_a(v)$. Then, node u will not pass this message to node a , and thus miss the opportunity to delivery the message. Nonetheless, considering global betweenness, each of the nodes of a , b and c has highest betweenness in the entire network (since they form the only path connecting components C_1 and C_2), and can then be correctly identified. A possible way to increase the chance of correct "bridge" identification is using larger neighborhood information, although this may increase communication cost. Similarly, to increase the chance of delivery, multiple relay

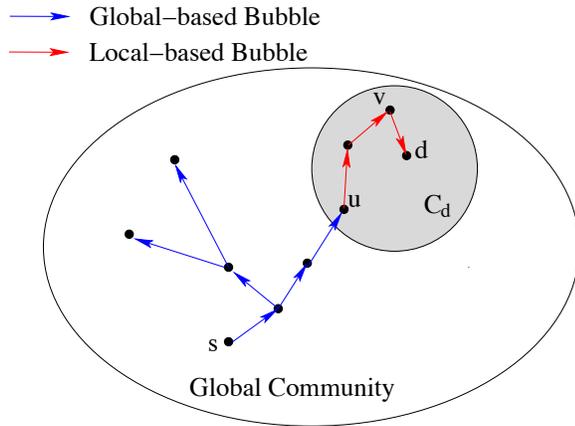


Fig. 4. An illustration of the Bubble Rap Forwarding from source s to destination d : blue and red arrows show the bubble-up operations based on global centrality in global community and local centrality in d 's community C_d , respectively.

nodes could be used. The trade-off is always between delivery performance and communication cost.

C. Bubble Rap Forwarding

The forwarding strategy, *Bubble Rap Forwarding*, proposed by Hui *et al.* [31] also relied on two social characteristics (community and centrality). They assumed that each node belongs to at least one community and its node centrality (either betweenness or degree centrality) in the community describes the popularity of the node within this community. Each node has a global centrality across the whole network (or called global community), and a local centrality within its local community. A node may also belong to multiple communities and hence have multiple local centralities. Taking advantages of these social characteristics, Bubble Rap Forwarding basically includes two phases: a bubble-up phase based on global centrality and a bubble-up phase based on local centrality. In both phases, the bubble-up forwarding strategy is utilized to forward messages to nodes which are more popular than the current node (i.e., with higher centrality). When a node s has a message with destination of d , it first bubbles the message up based on the global centrality, until the message reaches a node which is in the same local community C_d as the destination d . This procedure is shown as blue arrows in Figure 4. After the message reaches d 's community at node u , Bubble Rap Forwarding switches to the second phase which uses members of C_d as relays. This forwarding strategy continues to bubble up the message through the local community based on local centrality until the destination is reached. This later procedure is shown as red arrows in Figure 4. In order to reduce cost, it is also required that whenever a message is delivered to the community, the original carrier delete this message from its buffer to prevent further dissemination.

Bubble Rap Forwarding uses the concept of community in addition to node centrality to help with the forwarding decision. The introduction of local centrality inside a community is more beneficial than local centrality around local neighborhood (i.e., k -hop) [30]. The bubble-up operations allow fast transfer of a message towards the destination or its community.

However, such a strategy may fail when the destination belongs only to the communities whose members are all with low global centrality values. In this case, the bubble-up process in the first phase of Bubble Rap Forwarding cannot find the relay node which is in the same local community as the destination node. A possible solution for this problem to have a timeout timer for bubble-up process and exchange to other backup strategy for data delivery after timeout. In [31], the authors used a flat community (not hierarchical) to demonstrate the efficiency of Bubble Rap Forwarding. However, they did not provide details about how to handle hierarchical communities where the destination d may belong to multiple overlapping communities. In that scenario, they may face problems in the second phase of Bubble Rap Forwarding. For example, if the current encountering node u shares multiple communities with d , a problem arises regarding which one of d 's local communities should be chosen to bubble-up. A simple solution to this problem is picking the local community with which d have highest centrality. This solution also matches the spirit of Bubble Rap Forwarding which keeps looking for nodes with high centralities.

D. Social Based Multicasting

All of the social based routing methods discussed above are unicast routing protocols for DTNs. Social based approaches can be applied to multicast routing protocols for DTNs as well. Recently, Gao *et al.* have [32] proposed a set of multicast routing methods which use both centrality metric and community for relay selection.

Instead of using traditional centrality metrics such as betweenness, Gao *et al.* introduced a new metric (called *cumulative contact probability*) based on the Poisson modeling of social networks. Betweenness is purely defined based on the topology of the contact graph, and may not be sufficient to represent the probabilities for a node to contact others. Thus, a weighted social network model is used to differentiate the contact frequencies of different node pairs. In such model, the contact process of each node pair (i, j) is formulated as a Poisson process with an average contact rate of $\lambda_{i,j}$. Then the cumulative contact probability of node i can be defined as

$$C_i = 1 - \frac{1}{N-1} \sum_{j=1, j \neq i}^N e^{-\lambda_{ij}T}.$$

Here, N is the total number of nodes in the network, and T is the total time period. In other words, C_i indicates the average probability that node i meets a random node within time T . This centrality metric (or its variation) is used in the proposed multicast methods to select a relay node with higher centrality.

In [32], two multicast problems are considered: single-data multicast and multiple-data multicast, whose goals are to deliver a data item or a set of data items to a set of destinations within the time constraint T . The additional optimization objective is to minimize the number of relays used to achieve the average delivery ratio p .

For single-data multicast, the authors assumed the destinations are uniformly distributed, and thus they tried to ensure that all nodes are contacted by the data source or its selected relays within T . Based on cumulative contact

probabilities (i.e., centrality) of nodes, a minimal number of relay nodes are selected among the contacted neighbors of data source to guarantee the average delivery ratio is larger than p . This selection problem can be solved as a unified knapsack problem. The centrality metric is also refined for the case where the data source does not meet its contacted neighbors.

For multi-data multicast, the authors proposed a community-based approach which only requires nodes to maintain the probabilities of forwarding each data item to other nodes in the same community. When the destinations are in other communities, data forwarding is conducted through some gateway nodes which belong to multiple communities. The data source selects relays among its contacted neighbors based on the local centrality metric, and places appropriate data items on each relay. They used a two-stages relay selection schemes, where both data item selection and relay selection are modeled as knapsack problems. Their method can ensure that the average delivery probability is larger than p .

In summary, the solutions proposed in [32] provide new multicast strategies for DTNs based on the new centrality metric and community concept. The cumulative contact probability based on the Poisson modeling of social networks represents more accurate measurement of the probabilities for a node to contact other nodes. But this new model increases computational complexity as well. The authors addressed the optimization problem to minimize the number of relay nodes while satisfying the required delivery ratio and time constraint.

E. Homophily Based Data Diffusion

Zhang *et al.* [33] proposed a data diffusion scheme based on the “homophily” phenomenon in social networks. Here, data diffusion aims to deliver data to all nodes in DTNs. In DTNs, data may not be completely delivered from one node to another during a contact between them, since the contact time is too short to transmit the data or the buffer available at the receiving node is insufficient to hold the data. Therefore, in the design of data diffusion protocol, not only the contact probability between nodes but also the data propagation orders (which data should be propagated first) affect the diffusion speed and data access delay.

To choose an appropriate relay node to diffuse and an appropriate data item to buffer, Zhang *et al.* introduced a method using the *friendship* among nodes and the “homophily” phenomenon. The “homophily” phenomenon describes the trend in real word that friends usually share more common interests than strangers. By applying the same idea from “homophily” phenomenon, their proposed data diffusion strategy diffuses the most similar data items between friends, and diffuses the most different data items between strangers. If a node meets a new contact who is a friend, it first diffuses the most similar data items of their common interests to its friend first until the contact time is over. If the new contact is a stranger, it starts from the data item most different from their common interest. By theoretical analysis, Zhang *et al.* showed that this data diffusion scheme achieves better diffusion speed and data access delay than the other three possible schemes (including

diffusing the most similar data to any encounter, diffusing the most different data to any encounter, and diffusing the most different data between friends and the most similar data between strangers).

This proposed method provides a new angle to social-based approaches. It considers the need of managing data propagation orders which is an important aspect of design issues in DTN routing. With the same amount of communication opportunity and duration, more useful information can be transmitted under this proposed method. In addition, the proposed method is not conflicted with other DTN routing protocols. It can be used together with other DTN routing protocols to make better relay decisions with efficient data propagation orders. In the proposed method, social friendship is the only metric used to predict the encounter’s needs of information and friendship is defined by users. However, user defined friendship is not always available in DTNs. Therefore, it is another challenging direction need to be further explored regarding how to efficiently detect the friendship in dynamic DTNs .

F. Friendship Based Routing

Bulut *et al.* [34] also used friendship to aid the delivery of packets in DTNs. They introduced a new metric, social pressures metric (SPM), to accurately detect the quality of friendship. Different from [33], where friendship is defined by users based on their social relationships, this approach considered friends as nodes which contact to each other frequently and have long-lasting and regular contacts. Therefore, the social pressures metric between nodes i and j can be estimated from the encounter histories of these nodes (recorded by the nodes) as: $SPM_{i,j} = \frac{\int_{t=0}^T f(t)dt}{T}$, where $f(t)$ denotes the remaining time to the first encounter of these nodes after time t and T is the total time period. SPM describes the average forwarding delay if node i has a message destined to j at each time unit. Then, the link quality $w_{i,j}$ between each pair of nodes, (i, j) , is defined as $w_{i,j} = \frac{1}{SPM_{i,j}}$. The authors assumed that the bigger value of $w_{i,j}$ represents the closer friendship between i and j . Using the value of $w_{i,j}$, each node can construct its friendship community for each period T as a set of nodes whose link quality with itself is larger than a threshold. When a node i , having a message destined to d , meets with node j , it forwards the message to j if and only if (1) j and d are in the same friendship community (in the current period) and (2) j is a stronger friend of d than i .

In summary, this friendship based routing method uses the node contact information in each period to calculate the friendship metric (i.e., SPM), and constructs the friendship community. These social metrics can indeed help with making smarter forwarding decisions. However, the calculation of these metrics needs the whole contact information during each period, which may not be realistic in most DTNs. To obtain $f(t)$ in the current period, node i needs to know the time of its first encounter to node j after time t in this period, which is an event in future. Therefore, either the values in contact history from previous periods are used for this calculation at the current period or the estimated future contacts in this period are available for this calculation. This

is clearly a drawback of this proposed method. In addition, this friendship based routing uses a similar forwarding scheme to label routing [29], which may lead to the same problem. If the source node fails to meet with any node in the same friendship community with the destination node, the delivery fails. Therefore, more felicitous forwarding strategies should be studied for this friendship based routing.

Although the friendship based method [34] and homophily based method [33] both use friendship metrics for delivery data in DTNs, they are designed for different purposes. In [33] the friendship measurement is used to select which data items to diffuse, while in [34] the friendship metric is used to detect communities and select which relay nodes to forward. Therefore, different social metrics or various calculation methods need to be designed for specific design purposes. There is no universal solution for all applications.

G. Other Social-based Routings

Besides the social-based DTN routing strategies reviewed above, there are also a few recent social-based approaches which define their own social-related metrics to improve either the scalability or accuracy of routing. We briefly review them in this subsection.

In [91] Mei *et al.* took advantage of the observation, that people with similar interests tend to meet more often, to propose a social-aware and stateless routing (SANE) for packet switched networks. This routing strategy represents the interest profile of an individual u as an k -dimensional vector I_u . To express the interest similarity between two individuals u and v , the *cosine similarity* is defined as, $\Theta(I_u, I_v) = \cos(\angle I_u I_v) = \frac{I_u I_v}{\|I_u\| \|I_v\|}$. In SANE, a message should be forwarded to individuals whose interest profiles closely resemble that of the destination. They assume that the interest profile of a message m is the interest profile of its destination. Thus, a message m will be relayed to a node u only if the *cosine similarity* of the interest profile between message m and node u is higher than a given threshold ρ . One of the advantages of this method is that each node only needs to maintain the interest profile without extra storage. The cost of maintaining and updating this social metric is also relevantly easy. These advantages improve the scalability of this routing method.

Gao and Cao [105] proposed a user-centric data dissemination approach which considers both social centrality and user interests simultaneously. Different from the concept of centrality used in [30]–[32], this approach creates its own concept of centrality, which indicates the expected number of interesters (nodes interested in the data item held by i) that node i can encounter during the remaining time $T_k - t$ of data dissemination. Here, T_k is the time constraint of the data item and t is the current time. Then their relay selection makes sure that a new relay always has better capability of disseminating data to interesters than the existing relays based on this newly defined time-varying centrality. They consider both local centrality (centrality defined over one-hop neighborhood) and multi-hop centrality (which takes multi-hop opportunistic connection into consideration). With multi-hop centrality, more forwarding chances are considered, and

this strategy may thus lead to more accurate estimation of forwarding probability.

In [106], Fabbri and Verdone proposed a sociability-based DTN routing, which is based on the idea that nodes with high degrees of sociability (frequently encountering many different nodes) are good forwarding candidates. They defined the *sociability indicator* metric to evaluate the forwarding ability of a node. This metric quantifies the social behavior of a node by counting its encounters with all the other nodes in the network over a period T and is therefore a time-varying parameter. The routing strategy forwards packets to the most sociable nodes only. It is worth to notice that the strategy also considers both *first hop-based sociability* and *kth hop-based sociability*. For *kth* hop-based sociability, the highly sociable neighbors are considered during the calculation of a user's sociability.

From these new social-aware approaches, we can see that the design of social-based DTN routing tends to be more sophisticated. It not only directly uses social concepts from social networks but also considers its own reality in a DTN environment.

V. DEALING WITH NEGATIVE SOCIAL CHARACTERISTICS IN DTN ROUTING

In DTNs, most nodes are controlled by rational users or organizations, the selfish behaviors of these entities, which attempt to maximize their utilities and conserve their resources can significantly impact the network performance. Upendra *et al.* [35] demonstrated that network performance incurs serious degradation when selfish behaviors exist in DTNs. They compared the performance of fully cooperative DTNs with selfish DTNs via simulations with synthetic and real mobility traces, and their simulation results showed that the presence of selfish users can degrade the total delivered traffic to less than 20 percentage of what can be delivered under full cooperation. Motivated by this observation, the selfishness of DTN nodes have been considered in some of recent DTN routing solutions [35]–[40]. Most of these solutions focus on developing appropriate incentive mechanisms to stimulate individually selfish nodes to forward messages for all other nodes.

Incentive mechanisms [95]–[103] have been well-studied in traditional wireless ad hoc networks. However, limited resources and unreliable communications make it more challenging to design incentive mechanisms for DTNs than those for traditional wireless networks. In DTNs, the common assumption of an end-to-end path between the source and the destination no longer holds, and the delivery paths used by DTN routing cannot be predetermined. Thus, selfish actions are extremely difficult to detect. Even worse, large transmission delay makes the feedback information become useless. For these reasons, traditional incentive mechanisms do not work in DTNs naturally.

Existing incentive mechanisms for DTN routing can be categorized into three categories: *reputation based*, *tit-for-tat based* and *credit (virtual currency) based*. In reputation based schemes, forwarding services are provided to nodes depending on their reputation records. When a node provides

services for other nodes, it gains good reputation. Nodes with good reputations can receive services from other nodes. On the contrary, misbehaving nodes get bad reputations and will be denied participation in the network. The fear of detection and punishment motivates nodes to cooperate. In tit-for-tat based schemes, every node forwards as many messages for a neighbor as the neighbor forwards for it. In this way, a node autonomously lowers services to a neighbor if it detects that the neighbor is misbehaving. The credit based schemes introduce some form of credit or virtual currency to regulate the packet-forwarding relationships among different nodes. Nodes earn virtual currency by forwarding packets for others. These credits can be used to obtain forwarding service from any node in the network. For every forwarding request, the virtual bank charges the sender an extra amount of virtual currency, and the intermediate nodes redeem their rewards at the bank after successful delivery. In the remaining of this section, we review some incentive-based DTN routing protocols to handle selfish DTN nodes.

A. Give2Get

Mei *et al.* [36] proposed a reputation-based incentive scheme (Give2Get) for DTN routing, which can detect misbehaving nodes and remove them from DTN routing. They proved that their proposed scheme achieves a Nash equilibria, that is, no rational node has any incentive to deviate, i.e., selfish node cannot find a better choice other than truthfully following the protocol. They provided two versions of their proposed scheme for epidemic forwarding (where messages are forwarded to firstly encountered nodes) or delegation forwarding (where messages are forwarded to nodes with higher forwarding qualities) respectively. The basic ideas of both versions are the same: (1) hiding the content of the message (including its source and destination) to the candidate relay before the relay agrees to serve; and (2) requiring *proof of relay* after the selected relay agrees to serve and receive the encrypted message. Here, a proof of relay is just an encrypted message sent from the relay to the sender. The sender can show this message to the source node later to prove it has forwarded the message.

Here, we only use epidemic forwarding as the example to explain their proposed Give2Get scheme. The nodes generate the messages hide sender information to every possible relay except for the destination (by using destination's public key). Once a message m is generated, the sender s tries to find two other nodes and relay the message to them. When s encounters a node u , it first negotiates a symmetric key with u , then every communication after this point is encrypted with this key. Node s asks node u if it has already handled a message m by sending u the hash $H(m)$ of m . Node u would not lie because it does not know the content of m (including its destination), i.e., whether the destination is itself. If u lies, it may reject a message destined to itself. If u accepts the relay request, s encrypts the message with a random key k and send it to u . Then u sends a *proof of relay* to s . Finally, s sends the key k to u . After u gets message m , u follows the same steps as s did. It is required to collect two *proofs of relay* before a timeout Δ_1 and hold the message until Δ_2 if it cannot collect

two proofs within $[0, \Delta_1]$. During $[\Delta_1, \Delta_2]$, node s can test node u , if it meets u again, by requesting u to show either two *proofs of relay* or a proof of still having the message in its memory. Otherwise, s will broadcast a *proof of misbehavior* to the whole network that, in turn, will remove node u from the network unless u can prove that s is wrong. Node u can stop looking for relays only when two proofs are collected or Δ_1 expires, and discard the message after Δ_2 . The timeout Δ_1 and Δ_2 should be carefully chosen to achieve high success forwarding rate and positive probability of that node s will meet node u again before Δ_2 expires.

The Give2Get algorithm uses encryption techniques to avoid that selfish nodes lie to elude their duties or forge *proof of relay* to obtain good reputations. The Nash equilibria of the proposed scheme makes sure that no rational node has any incentive to deviate. It is worth to notice that the choices of the timeouts Δ_1 and Δ_2 could significantly affect the protocol's performance. The longer Δ_1 is, the higher success rate could be achieved. However, as Δ_1 and $\Delta_2 - \Delta_1$ increase, the buffer of nodes will be occupied for a longer time and the packet loss ratio will increase as well. Therefore, the choice of timeout could significantly depend on application requirements and the underlying network mobility. In [36], the authors provide discussions on the choice of timeout over their datasets.

B. Tit-For-Tat

Shevade *et al.* in [35] proposed the use of pair-wise tit-for-tat (TFT) as a simple and robust incentive mechanism for DTN routing. This incentive-aware routing protocol allows selfish nodes to maximize their own performances without any significant degradation of system-wide performance. Rather than attempting to detect misbehavior, it focuses on detecting good behaviors by using TFT.

Differing from traditional TFT, the method in [35] incorporates *generosity* and *contrition*. Generosity enables bootstrapping and absorbs transient asymmetries, while contrition prevents mistakes from causing endless retaliation. Bootstrapping happens when two nodes meet for the first time. Since no packets have ever been successfully relayed by both nodes, the basic TFT prevents the start of any relay. Generosity allows a node to send ε number of packets more than it has earned. This can absorb up to ε amount of traffic imbalance and stimulate the beginning of cooperative. However, this also allows selfish nodes to do ε less work than others. While generosity absorbs small amount of traffic imbalance, any imbalance exceeding ε could lead to lengthy retaliation between two neighbors. Therefore, contrition is introduced in their solution to prevent mistakes from causing endless retaliation and provide a way to return to stability after perturbation, by refraining a node from reacting to a valid retaliation to its own mistake.

The proposed TFT-based algorithm is easy to be implemented in DTNs, since the verification of relaying only exchanges between neighboring nodes. However, the generosity in this new TFT-based method may be exploited by selfish nodes. These nodes may reject to provide generosity to their neighbors while enjoy others' generosity. Therefore, this TFT-based algorithm cannot fully avoid selfish behaviors.

C. SMART and MobiCent

In this subsection, we review two recent credit-based incentive schemes in DTNs: SMART [37], [38] and MobiCent [39]. Notice that credit-based incentive schemes have been well-studied in traditional wireless networks [100]–[103]. Some of them could be extended to DTNs, but limited resources and the lack of instantaneous end-to-end paths make these schemes not as efficient as in traditional wireless networks. Therefore, here we only focus on reviewing two DTN schemes. SMART (Secure multilayer credit-based incentive) scheme [37], [38] uses credits to provide incentives to selfish nodes. This scheme allows the credits to be transferred/distributed by the current intermediate nodes without the involvement of the sender. Such scheme adopts a novel layer concatenation technique to withstand cheating actions of selfish nodes. MobiCent [39] assumes that each mobile device is capable of operating in two modes: a long-range low-bandwidth radio (e.g., cellular interface) to maintain an always-on connection, and a short-range high-bandwidth link (e.g., Wi-Fi) to opportunistically exchange a large amount of data with peers in its vicinity. MobiCent provides a credit-based algorithmic mechanism design approach to address selfish attacks.

In SMART [37], [38], a virtual electronic currency, a *layered coin*, is defined to stimulate the cooperation among nodes. Such a coin is composed of multiple layers, and each layer is generated by the source/destination or an intermediate node. The first layer (i.e., the base layer) is generated by the source to indicate the payment rate (credit value), remuneration conditions, the class-of-service (CoS) requirement, and other reward policies. It implies that the forwarding node agrees to provide forwarding service under the predefined CoS requirement and will be rewarded according to the reward policy in the future. During the subsequent propagation process, each intermediate node generates a new layer (i.e., the endorsed layer) based on the previous layers by appending a non-forgeable digital signature. With these endorsed layers, it is easy to track the propagation path and determine each intermediate node by checking the signature of each endorsed layer. In the rewarding and charging phase, if the provided forwarding service satisfies remuneration conditions defined in the predefined reward policy, each forwarding node along one or multiple path(s) will share the credit defined in this coin depending on different data-forwarding algorithms (single-copy/multi-copy forwarding) and the actual forwarding results (bundle delivered along one or multiple paths). A layer concatenation technique is presented in [37], [38] to ensure the security of layered coins. This technique concatenates different layers with each other by injecting the generator information of the next layer into the previous layer. However, the computation and transmission overhead of SMART may be unaffordable in some DTN applications. Even two optimization techniques are proposed to improve the overall efficiency, their overheads are still considerable.

MobiCent [39] is another credit-based incentive DTN system with its focus on truthful mechanism design for selfish DTNs. This mechanism not only encourages cooperation among nodes but also addresses two types of attack, edge insertion attack and edge hiding attack, from selfish nodes.

In edge insertion attacks, a node v can create a sybil v' such that it can be rewarded twice. For edge insertion attack, MobiCent designs a new incentive-compatible payment algorithm (multiplicative decreasing reward, MDR) where the relay's reward decreases with the path length and the client's payment increases with the path length. Under the MDR payment algorithm, both relays and client have no incentive to launch edge insertion attack. In edge hiding attacks, node v can hold the package instead of forwarding it to other relays. For edge hiding attack, MobiCent designs new payment set selection algorithms to determine an incentive-compatible relay set by examining a sufficient subset of paths ever revealed before deadline. With the proposed selection algorithms, relays and client have no incentives to launch edge insertion attack or edge hiding attack. In summary, MobiCent carefully designs the payment and rewarding algorithms to set the client's payment and the relays' rewards so that whatever the nodes do (edge insertion or hiding) they cannot gain more credits than they act truthfully. Therefore, nodes will always forward packets without adding phantom links, and never waste contact opportunities unless the reward is inadequate or it is the decision of the underlying routing protocol. One weakness of MobiCent is its requirement of always-on long-range connection. In many DTN applications, such centralized connection does not exist.

D. SSAR

In all of the methods above, a selfish node is considered as having the same level of selfishness to every other node. This selfishness can be called *individual selfishness*. However, in a more realistic scenario, a selfish node may have different levels of selfishness to different groups of nodes. In [104], Hui *et al.* found that in PSNs each node tends to help with forwarding more for people inside their community and less for people in other communities. This kind of selfish behaviors is called *social selfishness* to distinguish with individual selfishness. In [40], Li *et al.* considered social selfishness in DTNs by allowing a user to define different willingness (i.e., level of selfishness) to other users. The authors proposed a Social Selfishness Aware Routing (SSAR), which enforces users' social selfishness in routing meanwhile maintains acceptable routing performance. With the consideration of social selfishness, SSAR allocates resources such as buffers and bandwidths based on packet priority which is related to the social relationship among nodes (i.e., willingness defined by nodes themselves). To maintain routing performance, SSAR quantifies a relay's willingness to evaluate its forwarding capability and thus reduces packet dropping rate. Furthermore, SSAR formulates the forwarding process as a multiple knapsack problem with assignment restrictions and uses a greedy algorithm to select the forwarding packet set.

VI. CONCLUSION

In this article, we discuss different social characteristics of DTN nodes and survey current social-based routing protocols which use these social characteristics to assist packet forwarding in DTNs. Table I summarizes the social characteristics used by these routing protocols. To improve routing

TABLE I
COMPARISONS OF SOCIAL-BASED DTN ROUTING PROTOCOLS

Routing Protocols / Social Characteristics	Community	Centrality	Similarity	Friendship	Interest	Selfishness	
						Individual	Social
Label [29]	✓						
SimBet [30]		✓	✓				
Bubble Rap [31]	✓	✓					
Social Based Multicasting [32]	✓	✓					
Homophily Based Data Diffusion [33]				✓	✓		
Friendship Based Routing [34]	✓			✓			
SANE [91]			✓		✓		
Gao and Cao [105]		✓			✓		
Fabbri and Verdone [106]		✓					
Give2Get [36]						✓	
TFT [35]						✓	
SMART [37], [38]						✓	
MobiCent [39]						✓	
SSAR [40]							✓

performance in DTNs, these methods either take advantages of positive social characteristics (such as community and friendship) to make a smarter selection of relay node or use incentive mechanisms to stimulate individual nodes with negative social characteristics (such as selfishness).

From the analysis and comparison of these methods, we can conclude that social-based approaches are more promising than pure opportunity-based routing protocols for DTNs since these social-based approaches take advantages of relatively stable characteristics (social properties) efficiently to predict and deal with the dynamics of DTNs. Nonetheless, none of these approaches can guarantee the perfect routing performance (such as delivery guarantee or minimum delay). The prediction always has certain probability to fail in some cases. To provide more accurate prediction, multiple social metrics could be applied together (e.g., SimBet [30] combines centrality and similarity as their social-based routing metric). One step further, combining these social-based metrics with traditional opportunity-based metrics [107], [108] or other factors (such as location profile [92], [109] or mobility pattern [110], [111]) may also improve the performance in different settings. Our experiences show that combining multiple metrics may provide opportunities to improve the overall performance. However, such improvement is not guaranteed. On the contrary, inappropriate combination of multiple routing metrics may hurt the performance under certain circumstances. For example, in a very sparse DTN, it may significantly reduce the number of relay candidates and consequently lead to poor performance by enforcing the advance in multiple metrics. Therefore, It is an interesting research challenge regarding how to smartly take the advantage from multiple metrics. In addition, multiple forwarding strategies (e.g., combining social-based and opportunity-based approaches as in Give2Get [36]) may further improve the performance in broader situations. However, combining too many forwarding strategies may make the routing procedure too complex. Meanwhile, there is always a trade-off between performance and complexity in the protocol design. For example, SimBet [30] considers estimating centrality only based on local information, but it may miss identifying certain bridge nodes. Several methods [105], [106] also consider estimating their social metrics based on multi-hop relays to improve their predictions. However,

increasing the number of hops will cause larger overhead. Last, the performance of a routing protocol in DTNs depends heavily on the mobility model, environment, node density, social structure, and many other facts. Therefore, a universal routing solution for all DTN application scenarios is extremely hard. In other words, for certain DTN applications, we need to design specific routing protocols and mobility/social models to fulfill the requirements.

Even though social-based routing in DTNs has lately received much attention in the wireless network community as a relatively new area, there are still quite a few challenges left. First, most of the current social-based approaches use only simple definitions of one or two social characteristics (such as k-clique for community, node degree for centrality, contact frequency for friendship). Thus, it is interesting to see whether there are other more realistic and accurate social characteristics which can be used to further improve performance of DTN routing (even if it would be more complex). Second, social-based approaches significantly rely on accurate modelling of the social characteristics used by them. However, due to the lack of continuous connectivity and time-varying topology, it is hard to accurately estimate certain social characteristics without global or future information even with simple definitions. Therefore, it is another challenging task as to how to model and extract accurate social characteristics in dynamic DTNs. Last, there are not many social-based solutions which consider both positive and negative social characteristics together. It is a potential direction worthy for investigation in the future. For example, it could be more harmful for a node with higher social importance (with higher centrality or similarity) than a node with lower importance to be selfish. Considering positive social metrics together with negative ones, different approaches can be applied to avoid selfish behaviors from nodes with different social roles.

In this article, we mainly focus on social-based DTN routing protocols. But there are also other interesting directions within DTNs in which social studies may be used for protocol design, such as privacy protection [112], cooperative caching [113], [114], and content based sharing [115]. We believe that social-based approaches will be applied to wider research topics in communication networks far beyond DTN routing protocols.

VII. ACKNOWLEDGMENTS

This work is supported in part by the US National Science Foundation (NSF) under Grant No. CNS-0915331 and CNS-1050398, by the National Science Foundation of China (NSFC) under Grant NO. 61170212, and by Tsinghua National Laboratory for Information Science and Technology (TNList). The authors are grateful for a variety of valuable comments from the reviewers and the editor.

REFERENCES

- [1] Z. Zhang, "Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges," *IEEE Commun. Surveys Tutorials*, vol. 8, no. 1, pp. 24–37, 2006.
- [2] Y. Wang, H. Dang, and H. Wu, "A survey on analytic studies of delay-tolerant mobile sensor networks" *Wirel. Commun. Mob. Comput.*, vol. 7, no. 10, pp. 1197–1208, 2007.
- [3] P. Juang, H. Oki, Y. Wang, M. Martonosi, L. S. Peh, and D. Rubenstein, "Energy-efficient computing for wildlife tracking: design tradeoffs and early experiences with zebnet," *SIGOPS Oper. Syst. Rev.*, vol. 36, no. 5, pp. 96–107, 2002.
- [4] C. Perkins and E. Royer, "Ad-hoc on-demand distance vector routing," in *Proc. 2nd IEEE Workshop on Mobile Computing Systems and Applications*, February 1999.
- [5] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, T. Imielinski and H. F. Korth (Eds.), Norwood, MA: Kluwer Academic Publishers, p.153–181, 1996.
- [6] E. Royer and C. Toh, "A review of current routing protocols for ad-hoc mobile wireless networks," *IEEE Pers. Commun.*, vol. 6, no. 2, pp. 46–55, Apr. 1999.
- [7] Q. Yuan, I. Cardei, and J. Wu, "Predict and relay: an efficient routing in disruption-tolerant networks," in *MobiHoc '09: Proc. 10th ACM international symposium on Mobile ad hoc networking and computing*, 2009, pp. 95–104.
- [8] R. C. Shah, S. Roy, S. Jain, and W. Brunette, "Data mules: modeling a three-tier architecture for sparse sensor networks," in *Proc. 1st IEEE International Workshop on Sensor Network Protocols and Applications*, pp. 30–41, 2003.
- [9] B. Burns, O. Brock, and B. N. Levine, "MV routing and capacity building in disruption tolerant networks," in *Proc. 24th IEEE International Conference on Computer Communications (INFOCOM)*, pp. 398–408, 2005.
- [10] W. Zhao, M. Ammar, and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," in *MobiHoc '04: Proc. 5th ACM international symposium on Mobile ad hoc networking and computing*, pp. 187–198, 2004.
- [11] W. Zhao, M. Ammar, and E. Zegura, "Controlling the mobility of multiple data transport ferries in a delay-tolerant network," in *Proc. 24th IEEE International Conference on Computer Communications (INFOCOM)*, 2005.
- [12] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Duke University Technical Report CS-200006, April 2000.
- [13] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *WDTN '05: Proc. 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, 2005, pp. 252–259.
- [14] A. Lindgren, A. Doria, and O. Schelén, "Probabilistic routing in intermittently connected networks," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 7, no. 3, pp. 19–20, 2003.
- [15] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "Performance modeling of epidemic routing," *Comput. Netw.*, vol. 51, no. 10, pp. 2867–2891, 2007.
- [16] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine, "Maxprop: Routing for vehicle-based disruption-tolerant networks," in *Proc. 25th IEEE International Conference on Computer Communications (INFOCOM)*, 2006.
- [17] J. Leguay, T. Friedman, and V. Conan, "Evaluating mobility pattern space routing for DTNs," in *Proc. 25th IEEE International Conference on Computer Communications (INFOCOM)*, Barcelona, Spain, April 2006.
- [18] C. Liu and J. Wu, "Routing in a cyclic mobispace," in *MobiHoc '08: Proc. 9th ACM international symposium on Mobile ad hoc networking and computing*, pp. 351–360, 2008.
- [19] C. Liu and J. Wu, "Scalable routing in delay tolerant networks," in *MobiHoc '07: Proc. 8th ACM international symposium on Mobile ad hoc networking and computing*, pp. 51–60, 2007.
- [20] C. Liu and J. Wu, "An optimal probabilistic forwarding protocol in delay tolerant networks," in *MobiHoc '09: Proc. 10th ACM international symposium on Mobile ad hoc networking and computing*, pp. 105–114, 2009.
- [21] X. Zhang, J. Kurose, B. N. Levine, D. Towsley, and H. Zhang, "Study of a bus-based disruption-tolerant network: mobility modeling and impact on routing," in *MobiCom '07: Proc. 13th annual ACM international conference on Mobile computing and networking*, pp. 195–206, 2007.
- [22] M. Piorkowski, N. Sarafijanovic-Djukic, and M. Grossglauser, "A parsimonious model of mobile partitioned networks with clustering," in *Proc. 1st International Conference on Communication Systems and NETWORKS (COMSNETS)*, pp. 1–10, 2009.
- [23] A. Chaintreau, P. Fraigniaud, and E. Lebarh, "Opportunistic spatial gossip over mobile social networks," in *WOSN '08: Proc. first ACM Workshop on Online Social Networks*, 2008, pp. 73–78.
- [24] F. Chierichetti, S. Lattanzi, and A. Panconesi, "Gossiping (via mobile?) in social networks," in *DIAL M-POMC '08: Proc. 5th ACM international workshop on Foundations of mobile computing*, pp. 27–28, 2008.
- [25] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell, "Sensing meets mobile social networks: the design, implementation and evaluation of the CenceMe application," in *SenSys '08: Proc. 6th ACM conference on Embedded network sensor systems*, pp. 337–350, 2008.
- [26] A. Beach, M. Gartrell, S. Akkala, J. Elston, J. Kelley, K. Nishimoto, B. Ray, S. Razgulin, K. Sundaresan, B. Surendar, M. Terada, and R. Han, "WhozThat? evolving an ecosystem for context-aware mobile social networks," *IEEE Network*, vol. 22, no. 4, pp. 50–55, 2008.
- [27] Z.-B. Dong, G.-J. Song, K.-Q. Xie, and J.-Y. Wang, "An experimental study of large-scale mobile social network," in *Proc. 18th International World Wide Web Conference (WWW2009)*, April 2009.
- [28] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Pocket switched networks and the consequences of human mobility in conference environments," in *WDTN '05: Proc. 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, 2005.
- [29] P. Hui and J. Crowcroft, "How small labels create big improvements," in *International Workshop on Intermittently Connected Mobile Ad hoc Networks in conjunction with IEEE PerCom 2007*, March 19-23, 2007.
- [30] E. M. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant networks," in *MobiHoc '07 Proc. 8th ACM international symposium on Mobile ad hoc networking and computing*, 2007.
- [31] P. Hui, J. Crowcroft, and E. Yonek, "Bubble rap: Social-based forwarding in delay tolerant networks," in *Proc. 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, May, 2008.
- [32] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective networks," in *MobiHoc '09: Proc. 10th ACM international symposium on Mobile ad hoc networking and computing*, 2009.
- [33] Y. Zhang and J. Zhao, "Social network analysis on data diffusion in delay tolerant networks," in *MobiHoc '09: Proc. tenth ACM international symposium on Mobile ad hoc networking and computing*, 2009.
- [34] E. Bulut and B. K. Szymanski, "Friendship based routing in delay tolerant mobile social networks," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, Dec, 2010.
- [35] U. B. Shevade, H. H. Song, L. Qiu, and Y. Zhang, "Incentive-aware routing in DTNs," in *Proc. 16th IEEE International Conference on Network Protocols (ICNP 2008)*, Orlando, FL, USA, October 2008.
- [36] A. Mei and J. Stefa, "Give2get: Forwarding in social mobile wireless networks of selfish individuals," in *ICDCS '10 Proc. 2010 IEEE 30th International Conference on Distributed Computing Systems*, 2010.
- [37] H. Zhu, X. Lin, R. Lu, and X. Shen, "A secure incentive scheme for delay tolerant networks," in *Proc. 3rd International Conference on Communications and Networking in China (ChinaCom)*, Aug. 2008.
- [38] H. Zhu, X. Lin, R. Lu, Y. Fan, and X. Shen, "Smart: A secure multilayer credit-based incentive scheme for delay-tolerant networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4628–4639, Oct. 2009.
- [39] C. BinBin and C. M. Choon, "Mobicent: a credit-based incentive system for disruption tolerant network," in *INFOCOM'10: Proc. 29th IEEE International conference on Computer Communications*, Mar. 2010.
- [40] Q. Li, S. Zhu, and G. Cao, "Routing in socially selfish delay tolerant networks," in *INFOCOM'10: Proc. 29th IEEE International conference on Computer Communications*, Mar. 2010.

- [41] J. Scott, *Social Network Analysis: A Handbook*. Sage Publications, 2000.
- [42] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.
- [43] M. Girvan and M.E.J. Newman, "Community structure in social and biological networks," in *Proc. National Academy of Sciences*, vol. 99, no. 12, pp. 7821-7826, 2002.
- [44] M.E.J. Newman, "Modularity and community structure in networks," *Proc. National Academy of Sciences*, vol. 103, no. 23, pp. 8577, 2006.
- [45] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of community hierarchies in large networks," *J. Stat. Mech.*, vol. 2008, no. 10, page: P10008, 2008.
- [46] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, no. 3-5, pp. 75174, 2010.
- [47] V. Satuluri, S. Parthasarathy and Y. Ruan, "Local graph sparsification for scalable clustering," in *Proc. 2011 international conference on Management of data*, 2011.
- [48] K. Xu, J. Li, S.S. Liao, "Sentiment community detection in social networks," in *Proc. 2011 iConference*, pp. 804-805, 2011.
- [49] X. Liu and T. Murata, "Extracting the mesoscopic structure from heterogeneous systems," in *Proc. 22nd ACM conference on Hypertext and hypermedia*, 2011.
- [50] N.P. Nguyen, T.N. Dinh, S. Tokala, and M.T. Thai, "Overlapping communities in dynamic networks: their detection and mobile applications," in *Proc. 17th ACM annual international conference on Mobile computing and networking (MobiCom '11)*, 2011.
- [51] J. Leskovec, K.J. Lang and M. Mahoney, "Empirical comparison of algorithms for network community detection," in *Proc. 19th international conference on World wide web*, 2010.
- [52] F. Mitzlaff, M. Atzmueller, D. Benz, A. Hotho and G. Stume, "Community assessment using evidence networks," in *Proc. 2010 international conference on Analysis of social media and ubiquitous data*, pp. 79-98, 2010.
- [53] H.Almeida, D. Guedes, W. Meira and M.J. Zaki, "Is there a best quality metric for graph clusters?" in *Proc. 2011 European conference on Machine learning and knowledge discovery in databases*, 2011.
- [54] P.G. Lind, L.R. da Silva, J.S. Andrade, and H.J. Herrmann, "Spreading gossip in social networks," *Physical Review E*, vol. 76, no. 3, 2007.
- [55] M.E. Yildiz, A. Scaglione, and A. Ozdaglar, "Asymmetric information diffusion via gossiping on static and dynamic networks," in *Proc. 49th IEEE Conference on Decision and Control (CDC 2010)*, 2010.
- [56] J. Tang, M. Musolesi, C. Mascolo, V. Latora, "Temporal distance metrics for social network analysis," in *Proc. 2nd ACM workshop on Online social networks*, 2009.
- [57] J. Tang, M.Musolesi, C. Mascolo, V. Latora, "Characterising temporal distance and reachability in mobile and online social networks," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 1, 2010.
- [58] C. Lee, H. Kwak, H. Park, S. Moon, "Finding influentials based on the temporal order of information adoption in Twitter," in *Proc. 19th international conference on World wide web*, 2010.
- [59] Q. Zhao, Y. Tuan, Q. He, N. Oliver, R. Jin, W. Lee, "Communication motifs: a tool to characterize social communications," in *Proc. 19th ACM international conference on Information and knowledge management*, 2010.
- [60] E. Bakshy, B. Karrer, L. Adamic, "Social influence and the diffusion of user-created content," in *Proc. tenth ACM conference on Electronic commerce*, 2009.
- [61] C.J. Kuhlman, V.S.A. Kumar, M.V. Marathe, S.S. Ravi, D.J. Rosenkrantz, "Finding critical nodes for inhibiting diffusion of complex contagions in social networks," in *Proc. 2010 European conference on Machine learning and knowledge discovery in databases: Part II*, 2010.
- [62] O. Phelan, K. McCarthy, M. Bennett, B. Smyth, "On using the real-time web for news recommendation & discovery," in *Proc. 20th international conference companion on World wide web*, 2011.
- [63] O. Phelan, K. McCarthy, M. Bennett, B. Smyth, "Term of a feather:content-based news recommendation and discovery using twitter," in *Proc. 33rd European conference on Advances in information retrieval*, 2011.
- [64] S.K. Tyler, S. Zhu, Y. Chi, Y. Zhang, "Ordering innovators and laggards for product categorization and recommendation," in *Proc. third ACM conference on Recommender systems*, 2009.
- [65] G.R. Lopes, M.M. Moro, L.K. Wives, J.P.M.D Oliveira, "Collaboration recommendation on academic social networks," in *Proc. 2010 international conference on Advances in conceptual modeling: applications and challenges*, 2010.
- [66] Y. Tian, Q. He, Q. Zhao, X. Liu, W. Lee, "Boosting social network connectivity with link revival," in *Proc. 19th ACM international conference on Information and Knowledge management*, 2010.
- [67] L. Zhao, Y. Zeng, N. Zhong, "A weighted multi-factor algorithm for microblog search," in *Proc. 7th international conference on Active media technology*, 2011.
- [68] M.D. Choudhury, S. Counts, M. Czerwinski, "Identifying relevant social media content: leveraging information diversity and user cognition," in *Proc. 22nd ACM conference on Hypertext and hypermedia*, 2011.
- [69] C. Grier, K. Thomas, V. Paxson, M. Zhang, "@spam:the underground on 140 characters or less," in *Proc. 17th ACM conference on Computer and communications security*, 2010.
- [70] M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, E. Ferrari, "Content-based filtering in online social networks," in *Proc. international ECML/PKDD conference on Privacy and security issues in data mining and machine learning*, pp. 127-140, 2010.
- [71] H. Hu, G. Ahn, "Multiparty authorization framework for data sharing in online social networks," in *Proc. 25th annual IFIP WG 11.3 conference on Data and applications security and privacy*, 2011.
- [72] A. Narayanan and V. Shmatikov, "De-anonymizing social networks," in *Proc. 30th IEEE Symposium on Security and Privacy*, 2009.
- [73] L. Fang and K. Lefevre, "Privacy wizards for social networking sites," in *Proc. 19th international conference on World wide web*, 2010.
- [74] K. Liu and E. Terzi, "A framework for computing the privacy scores of users in online social networks," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2010.
- [75] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in *Proc. 7th ACM SIGCOMM conference on Internet measurement (IMC '07)*, 2007.
- [76] V. E. Krebs, "Mapping networks of terrorist cells," *Connections*, vol. 24, no. 3, pp. 43-52, 2002.
- [77] T. Hossmann, F. Legendre, and T. Spyropoulos, "From contacts to graphs: pitfalls in using complex network analysis for DTN routing," in *INFOCOM'09: Proc. 28th IEEE International conference on Computer Communications Workshops*, pp. 260-265, 2009.
- [78] T. Hossmann, T. Spyropoulos, and F. Legendre, "Know thy neighbor: Towards optimal mapping of contacts to social graphs for DTN routing," in *INFOCOM'10: Proc. 29th IEEE International conference on Computer Communications*, 2010.
- [79] P. Hui, E. Yoneki, S.Y. Chan, and J. Crowcroft, "Distributed community detection in delay tolerant networks," in *Proc. ACM SIGCOMM Workshop, MobiArch'07*, 2007.
- [80] E. Yoneki, P. Hui, and J. Crowcroft, "Visualizing community detection in opportunistic networks," in *Proc. ACM MobiCom Workshop on Challenged Networks (CHANTS)*, September, 2007.
- [81] P. Hui, E. Yoneki, and J. Crowcroft, "Identifying social communities in complex communications for network efficiency," in *Proc. 1st International Conference on Complex Sciences: Theory and Applications (Complex 2009)*, February, 2009.
- [82] S. Y. Chan, P. Hui, and K. Xu, "Community detection of time-varying mobile social networks," in *Proc. 1st International Conference on Complex Sciences: Theory and Applications (Complex 2009)*, February, 2009.
- [83] G.A. Hillery, "Definitions of community: areas of agreement," *Rural Sociology*, vol. 20, no. 4, pp. 111, 1955.
- [84] D.W. McMillan and D.M. Chavis, "Sense of community: A definition and theory," *Journal of Community Psychology*, vol. 14, no. 1, pp. 6-23, 1986
- [85] D.D. Perkins, P. Florin, R.C. Rich, A. Wandersman, and D.M. Chavis, "Participation and the social and physical environment of residential blocks: Crime and community context," *American Journal of Community Psychology*, vol. 18, pp. 83-115, 1990.
- [86] S. Okasha, "Altruism, group selection and correlated interaction," *British Journal for the Philosophy of Science*, vol. 56, no.4, pp. 730-725, 2005.
- [87] P. V. Marsden, "Egocentric and sociocentric measures of network centrality," *Social Networks*, vol. 24, no. 4, pp. 407-422, October 2002.
- [88] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40(1), pp. 35-41, 1977.
- [89] L. C. Freeman, "Centrality in social networks: conceptual clarification," *Social networks*, vol. 1(3), pp. 215-239, 1979.
- [90] M. J. Newman, "A measure of betweenness centrality based on random walks," *Socail networks*, vol. 27(1), pp. 39-54, 2005.
- [91] A. Mei, G. Morabito, P. Santi and J. Stefa, "Social-aware stateless forwarding in pocket switched networks," in *Proc. 30th IEEE Conference on Computer Communications(INFOCOM) mini-conference*, 2011.
- [92] J. Leguay, T. Friedman, and V. Conan, "DTN routing in a mobility pattern space," in *Proc. ACM WDTN*, 2005.

- [93] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, vol. 27, pp. 415–444, 2001.
- [94] R. Dawkins, *The Selfish Gene*. Oxford University Press, 1990.
- [95] L. Buttyan and J. Hubaux, "Stimulating cooperation in self-organizing mobile ad hoc networks," *ACM/Kluwer Mobile Networks and Applications*, vol. 5, no. 8, October 2003.
- [96] M. Jakobsson, J.-P. Hubaux, and L. Buttyan, "A micro-payment scheme encouraging collaboration in multi-hop cellular networks," in *Proc. Financial Cryptography*, 2003.
- [97] L. Blazevic, L. Buttyan, S. Capkun, S. Giordano, J. P. Hubaux, and J. Y. L. Boudec, "Self-organization in mobile ad-hoc networks: the approach of terminodes," *IEEE Commun. Mag.*, vol. 39, no. 6, June 2001.
- [98] L. Buttyan and J. P. Hubaux, "Enforcing service availability in mobile ad-hoc WANS," in *Proc. 1st ACM international symposium on Mobile ad hoc networking & computing*, Boston, Massachusetts, 2000, pp. 87–96.
- [99] V. Srinivasan, P. Nuggehalli, C. F. Chiasserini, and R. R. Rao, "Cooperation in wireless ad hoc wireless networks," in *Proc. 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom)*, 2003.
- [100] L. Anderegg and S. Eidenbenz, "Ad hoc-VCG: a truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents," in *Proc. 9th annual international conference on Mobile computing and networking*, pp. 245–259, 2003.
- [101] W. Wang, S. Eidenbenz, Y. Wang, and X.-Y. Li, "OURS: Optimal unicast routing system in non-cooperative wireless networks," in *Proc. 12th ACM Annual International Conference on Mobile Computing and Networking (MobiCom 2006)*, 2006.
- [102] W. Wang, X.-Y. Li, and Y. Wang, "Truthful multicast in selfish wireless networks," in *Proc. 10th ACM Annual International Conference on Mobile Computing and Networking (MobiCom 2004)*, 2004.
- [103] S. Zhong, L. Li, Y. G. Liu, and Y. R. Yang, "On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks," in *Proc. 11th ACM Annual International Conference on Mobile Computing and Networking (MobiCom 2005)*, 2005.
- [104] P. Hui, K. Xu, V. O. Li, J. Crowcroft, V. Latora, and P. Lio, "Selfishness, altruism and message spreading in mobile social networks," in *Proc. 1st IEEE International Workshop on Network Science for Communication Networks (NetSciCom09)*, April 24, 2009.
- [105] W. Gao and G. Cao, "User-centric data dissemination in disruption tolerant networks," in *Proc. 30th IEEE Conference on Computer Communications (INFOCOM)*, 2011.
- [106] F. Fabbri and R. Verdone, "A sociability-based routing scheme for delay-tolerant networks," in *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, January, 2011.
- [107] W. Gao and G. Cao, "On exploiting transient contact patterns for data forwarding in delay tolerant networks," in *Proc. 18th IEEE Int'l conference on Network Protocols (ICNP)*, 2010.
- [108] E. Bulut, S. Geyik and B. Szymanski, "Efficient routing in delay tolerant networks with correlated node mobility," in *Proc. 7th IEEE International Conference on Mobile Ad-hoc and Sensor Systems(MASS)*, Nov, 2010.
- [109] J. Fan, Y. Du, W. Gao, J. Chen and Y. Sun, "Geography-aware active data dissemination in mobile social networks," in *Proc. Mobile Ad Hoc and Sensor Systems (MASS)*, Nov. 2010.
- [110] W. Gao and G. Cao, "Fine-grained mobility characterization: Steady and transient state behaviors," in *Proc. 11th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2010.
- [111] L. Vu, Q. Do and K. Nahrstedt, "3R fine-grained encounter-based routing in delay tolerant networks," in *Proc. 12th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, WoWMoM*, 2011.
- [112] R. Lu, X. Lin, and X. Shen, "SPRING: A social-based privacy-preserving packet forwarding protocol for vehicular delay tolerant networks," in *Proc. 29th IEEE International Conference on Computer Communications (INFOCOM)*, 2010.
- [113] X. Zhuo, Q. Li, G. Cao, Y. Dai, B. Szymanski, and T. La Porta, "Social-based cooperative caching in DTNs: a contact duration aware approach," in *International Conference on Mobile Ad-hoc and Sensor Systems (IEEE MASS)*, 2011.
- [114] W. Gao, G. Cao, A. Iyengar, and M. Srivatsa, "Supporting cooperative caching in disruption tolerant networks," in *Proc. IEEE International Conference on Distributed Computing Systems (ICDCS)*, 2011.
- [115] G. Karbaschi and A.C. Viana, "A content-based network coding to match social interest similarities in delay tolerant networks," in *Proc. 1st Extreme Workshop on Communication (ExtremeCom)*, 2009.



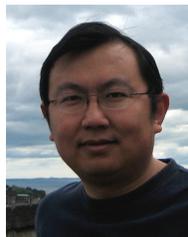
Ying Zhu received her B.Eng. degree in automatic control and M.Eng. degree in pattern recognition from University of Electronic Science and Technology of China in 2002 and 2005, respectively. She is currently a Ph.D. student in the University of North Carolina at Charlotte, majoring in computer science. Her current research focuses on wireless networks, ad hoc and sensor networks, delay tolerant networks, and algorithm design.



Bin Xu received his PhD degree from the Department of Computer Science and Technology at Tsinghua University, China, in 2006. He is currently an associate professor of computer science in Tsinghua University. His research interests mainly include sensor networks, mobile computing, and service composition. He is a member of the IEEE and the ACM.



Xinghua Shi received her Ph.D.(2008) and M.Sc. (2003) in computer science from the University of Chicago and her M.Eng (2001) and B. Eng (1998) in computer science from Beijing Institute of Technology, China. Currently, she is a research fellow at Brigham and Women's Hospital, Harvard Medical School, and a visiting postdoctoral fellow at Broad Institute of Harvard and MIT. Her research interests span computational and systems biology, bioinformatics, genetics and epigenetics, scientific computing, and networks.



Yu Wang is an associate professor of computer science at the University of North Carolina at Charlotte. He received his Ph.D. degree (2004) in computer science from Illinois Institute of Technology, his B.Eng. degree (1998) and M.Eng. degree (2000) in computer science from Tsinghua University, China. His research interest includes wireless networks, ad hoc and sensor networks, delay tolerant networks, mobile computing, complex network, and algorithm design. He has published more than 100 papers in peer-reviewed journals and conferences. He has served as program chair, publicity chair, and program committee member for several international conferences (such as IEEE INFOCOM, IEEE IPCCC, IEEE GLOBECOM, IEEE ICC, IEEE MASS, etc.). He is a recipient of Ralph E. Powe Junior Faculty Enhancement Awards from Oak Ridge Associated Universities in 2006 and a recipient of Outstanding Faculty Research Award from College of Computing and Informatics at UNC Charlotte in 2008. He is a senior member of the ACM, IEEE, and IEEE Communications Society.