Integrating Social Links into Wireless Networks: Modeling, Routing, Analysis and Evaluation

Zhuo Lu, *Member, IEEE,* Yalin E. Sagduyu, *Senior Member, IEEE,* and Yi Shi, *Senior Member, IEEE,*

Abstract—Social connections among network nodes have been well investigated as an additional opportunity in network design (e.g., in routing strategies and trusted networking). This paper presents a paradigm shift that explores the design and performance analysis of combining social links jointly with communication links for message delivery in wireless networks. In a combined multi-layer social and communication network, communication links are based on conventional wireless technologies (e.g., WiFi, Bluetooth) and social links are overlaid over a communication infrastructure (e.g., cellular network) that provides an alternative way for data transmission. The goal is to characterize the performance analytically when routing is designed by combining social and communication links. A distance discretization technique is applied to model the reliability and delay of message delivery. The analytical foundation is developed to analyze the end-to-end delay and success probability under various effects of persistent transmission, potential error in distance estimation, and mobility. Systematic routing strategies that employ network inference are then designed to improve the performance in different aspects, such as delivery delay, delivery success probability, and energy-saving. A network emulation testbed is implemented with actual radios and real-world social network datasets to measure the performance of a heterogeneous network with social and communication links. The results in this paper show that the integration of social links in wireless network routing as a multi-layer design leads to substantial performance improvement for delay and reliability of message delivery.

Index Terms—Wireless networks; network science; multi-layer networks; social links; delay; reliability.

1 Introduction

Leveraging social relationships to improve the network performance has been recently investigated in network routing protocol design, such as social-aware routing in delay-tolerant network (DTN) [1]–[6]. In these protocols, social ties are typically used as abstract or conceptual links for a node's decision making in routing.

A social network can be considered not only as a logical topology that represents social connections used for decision making, but also as an overlay network over a physical infrastructure for information delivery. For example, when people make phone calls to their friends, they can communicate with each other because they have social links and the cellular network infrastructure serves as the underlying communication medium for such social links. In this regard, cellular networks feature a heterogeneous architecture, in which social links on top of the cellular (such as LTE) infrastructure constitute an overlay network, and at the same time WiFi or Bluetooth links of smart phones form a conventional wireless network with peer-to-peer communication. Alternatively, social links can be also considered as other communication means (such as using airborne relays, satellite communications, high power or directional transmissions) to enable social connections. Current routing and data delivery processes in cellular networks, by default, operate over one network interface (e.g., web surfing in smart phones through either the WiFi or cellular network interface). If network design jointly takes into account the overlaid social links and the conventional communication links, the network performance can be potentially improved with more reliable end-to-end delivery, higher throughput, or smaller message delay.

Although a social network infrastructure may support transmissions to non-friends (strangers), a stranger may not want to forward data towards a destination. For example, a person may send a message over a cellular network to anyone and ask for forwarding this message. A social friend is likely to forward but a stranger may not. As a result, we assume that a node only sends data to friends in a social network. Therefore, social links as a whole become vital routing information at the network layer. There exists a design space with performance benefits for integrating social links as another type of data links into a wireless network to form a combined social and communication network.

Some recent studies have already considered a combined network with both social and communication links [7]–[10]. The typical assumption for analytical foundations in these works is that there exist an infinite number of users in a finite-area network such that greedy routing [11]–[14] can always reliably find a next hop neighbor to move a message closer to the destination. Nonetheless, such an assumption does not always hold due to two main reasons: (i) there are only a finite number of nodes in a network therefore reliable routing is not always guaranteed; and (ii) message delivery

- *Zhuo Lu is with the Department of Electrical Engineering, University of South Florida, Tampa FL 33620. E-mail: zhuolu@usf.edu.*
- *Yalin E. Sagduyu and Yi Shi are with Intelligent Automation Inc, Rockville MD 20855. E-mails: {ysagduyu, yshi}@i-a-i.com.*
- *A preliminary version appears in the 2016 IEEE International Workshop on Network Science for Communication Networks.*
- *Part of the material presented in this paper is based upon work supported by the Air Force Office of Scientific Research under Contract FA9550-12-C-0037. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Air Force Office of Scientific Research.*

may also fail due to communication or social link failures (as data delivery on social links still rely on an underlying communication infrastructure).

In this paper, we aim to model the performance of a combined social and communication network with a finite number of nodes and link unreliability. We derive multi–layer network routing strategies to optimize the performance according to this model. Our approach to analyze the performance of data delivery is based on a novel *distance discretization* technique, which gradually aggregates the delay or success probability of a message that reaches a discretized distance to the destination. Based on this approach, we model and analyze the delivery delay and success probability of a combined social and communication network with respect to a variety of network conditions, such as node density, mobility, and link failure pattern. We also develop a proof-of-concept testbed to perform real-time high-fidelity evaluation of combined social and communication network design. The paper can be considered as an exploratory work to model, analyze and evaluate a new type of wireless networks with social links integrated as data links to show the potential benefits of such a new network type. Our contributions can be summarized as follows.

1) We developed a novel network discretization technique to model the message delivery of combined social and communication networks.
2) We analyzed end-to-end delay and success probability under effects of persistent transmission, potential error in distance estimation, and mobility.
3) We designed and implemented systematic routing strategies that employ network inference to improve the performance in different aspects, such as delivery delay, delivery success probability, and energy-saving.
4) We implemented a testbed with actual radios and real-world social network datasets to measure the performance of combined social and communication networks.
5) Our results motivate the integration of social links into wireless network design to improve the message delivery performance.

The remainder of this paper is organized as follows. In Section 2, we introduce the preliminaries and models. In Section 3, we present the performance analysis and simulations of combined social and communication networks. In Section 4, we show how to improve the performance by changing routing strategies. In Section 5, we describe the network emulation testbed and experimental results. Section 6 reviews related work. Finally, we conclude this paper in Section 7.

2 Preliminaries and Models

In this section, we present assumptions and models, and then formulate the problem of message delivery in combined social and communication networks.

2.1 Wireless and Social Network Models

We consider a multi-layer social and communication network with a finite number of nodes, in which communication links are based on short-range wireless connections (e.g., WiFi, Bluetooth) and social links are overlaid over long-range communication infrastructures (e.g., satellite and cellular networks). Both links can be used for data transmission. Hence, the combined social and communication network model differs from existing models [15]–[18] that leverage social links only for decision making in routing.

To provide an analytical framework to model such a scenario, we integrate social links into the traditional approach of wireless ad hoc network modeling. We consider the commonly adopted multi-hop wireless network model (e.g., [19]–[21]) where N nodes are uniformly and independently distributed on a disk area with radius R. Nodes can communicate using communication links if they are within each other's communication range $r_c \ll R$, indicating that the maximum number of hops in the network can be considerably large. The uniform node distribution and the common communication range are assumed to achieve a balance between mathematical traceability and practical indication in our modeling. As many mobility models (e.g., random walk and random direction model [22]) statistically lead to the uniform (or closely uniform) distribution of nodes over the network given a snapshot of the network topology at a particular time, our modeling and analysis apply to static networks as well as model networks with node mobility resulting in the stationary uniform node distribution.

Besides communication links, nodes can also communicate via social links. In this paper, the social links are high-level abstract links built upon a realistic data communication infrastructure (e.g., cellular or satellite networks) that spans over the network area. Such social links are not used as logical links in traditional networks, but can be considered as super links used to deliver data between two nodes that are socially connected. Hence, throughout this paper, social links are assumed to also have the capability of data delivery. As a result, how nodes are connected via social links is essential for the performance of the combined social and communication network. It is well known that social networks exhibit the small-world phenomenon, i.e., social actors are linked by short chains of acquaintances [23]–[26]. We adopt the Octopus model [27], [28] to capture such a phenomenon together with additional degree distribution characterization such as scale-free network properties.

In the Octopus model, there are two types of social links: short-range (SR) and long-range (LR) connection links to account for close and far social connections, respectively. In practice, social links may be coupled with communication links between nodes. For example, two socially close friends may be also geographically close to each other (e.g., friends living in the same apartment); then, they have both social and communication links. To accommodate such correlations, we use probabilities γ_{CS} and γ_{CL} to denote the probabilities that an SR or LR social link, respectively, correspond to a communication link. Similarly, we denote by γ_{NCS} or γ_{NCL} the probabilities, respectively, that an SR or LR social link exists between two nodes without a communication link.

Fig. 1 shows an example of how the combined social and communication network is modeled: nodes B and C are within node A's wireless transmission range. Therefore, there exist two communication links for node A: A $\leftrightarrow$ B, A $\leftrightarrow$ C. In addition, the fact that A and B or A and C are physically close may indicate that they are also socially

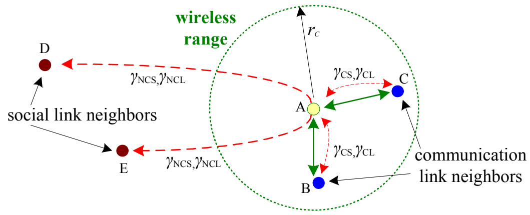

Fig. 1. Social and communication links of a node in a combined network.

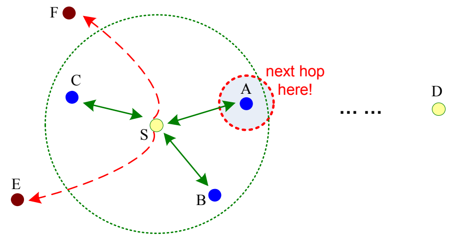

Fig. 2. Under greedy routing, source S wants to deliver a message to destination D.

close (i.e., there is a chance for them to be socially close). Therefore, we assume that there exists an SR social link with probability γ_{CS} between nodes A and B (or C). On the other hand, they may not be socially close but still know each other (i.e., there is an LR link between them according to the Octopus model). We assume that there is an LR social link with probability γ_{CL} between nodes A and B (or C) if there is no SR link. Consider nodes D and E in Fig. 1; because they are not within node A's wireless transmission range, there is no communication link from node A to D (or E). However, there still exists an SR social link with probability γ_{NCS} between nodes A and D (or E). If the SR link does not exist, there is an LR social link with probability γ_{NCL} between nodes A and D (or E).

2.2 Greedy Routing Procedure

In a combined social and communication network, a message can be transmitted over social or communication links along an end-to-end delivery path. We adopt the greedy routing mechanism [7], [11], [27] for message delivery, as a viable solution over a multi-hop path using only local information at each node. In particular, each node in the network maintains a table about their social neighbors as well as their one-hop communication network neighbors. The mechanism makes routing adaptive to a dynamically-changing network topology, where there exists no stable end-to-end path between two nodes. There are several assumptions associated with the greedy routing: (i) there is no centralized routing coordination in the network; (ii) the source only uses its local next-hop information to make the routing decision; (iii) the routing source or a forwarding node has knowledge about where the destination is (or its distance to the destination); (iv) greedy routing always forwards the data closer to destination (i.e., the routing decision of the source or a forwarding node is always to make sure the next hop has a smaller physical distance than the source or the forwarding node to the destination). Assumptions (i) and (ii) are to ensure that greedy routing can be used under a dynamically-changing network topology. But when only local information is used, greedy routing still needs some level of global information to route data to the correct destination; therefore, the knowledge of its distance to the destination becomes essential, which is assumption (iii). This knowledge does not need to be exactly accurate. The impact of approximate knowledge with distance estimation errors is analyzed in Section 3.3.2. Assumption (iv) is the heuristic nature of greedy routing over geographical areas.

A forwarding node under greedy routing always attempts to find the next-hop node in all of its social link and communication link neighbors, whose distance[1] to the destination is the shortest, and at the same time smaller than the forwarding node's distance to the destination. An example is shown in Fig. 2, where source node S wants to transmit a message to destination node D. Node S first checks its neighbors via both social links (nodes E and F) and communication links (nodes A, B, and C). Among all neighbors, it chooses the one that is closest to destination D, namely, node A in this case. Then, the message will be transmitted to node A, who will follow the same greedy routing procedure to find the next hop node. If nodes A and B do not exist in Fig. 2, node C would become the neighbor closest to destination D. However, its distance to D is larger than source S's distance to D. This means that the message would be delivered farther away from the destination, which is, however, not allowed by greedy routing. Therefore, source S would simply drop the message and claim delivery failure.

Note that social information is considered in routing decisions. Social and wireless links are combined in a joint graph, where the delay is minimized with local decisions. Implicitly, social information (such as the social links and the corresponding link delays and success probabilities) is used in routing decisions.

2.3 Problem Formulation

With network model and routing protocol defined, we evaluate the success probability and delay of message delivery in a combined social and communication network to analyze the benefit of a joint (multi-layer) network design for the routing protocol.

To facilitate performance evaluation, we define the hop distance k between two nodes in the network as $k = \lceil d/r_c \rceil$, where d is their distance and r_c is the communication range. In addition, we denote by β_s and β_c the social and communication link success ratios, respectively. All the notations used throughout this paper are shown in Table 1.

In this paper, we aim to evaluate the performance of a combined social and communication network in terms of the success probability of message delivery S_k and average delay of message delivery T_k both as functions of hop distance k.

3 Analysis and Evaluation

In this section, we derive the success probability and delay of end-to-end message delivery analytically under greedy

1. In this paper, distance is referred to as the geographical or physical distance, not the social distance, unless specified otherwise.

TABLE 1
Notations used in the paper.

N:	number of nodes in the network.
R:	network size.
λ:	node density, defined as $\lambda = N/(\pi R^2)$.
r_c:	transmission range of a communication link.
β_c:	success ratio of a communication link.
β_s:	success ratio of a social link.
D_c:	communication link delay.
D_s:	social link delay.
γ_{CS}:	probability that there exists an SR social link given the presence of a communication link.
γ_{CL}:	probability that there exists an LR social link given the presence of a communication link.
γ_{NCS}:	probability that there exists an SR social link given the non-presence of a communication link.
γ_{NCL}:	probability that there exists an LR social link given the non-presence of a communication link.
ρ:	probability that two nodes are socially connected.
S_k:	delivery success probability at hop distance k.
T_k:	average delivery delay at hop distance k.
S'_m:	delivery success probability at mini-hop distance m.
T'_m:	average delivery delay at mini-hop distance m.

routing, and then use simulations to validate the analysis.

3.1 Performance under Greedy Routing

To facilitate tractable analysis, our methodology is to propose a new approximation technique called *distance discretization*: First, rings are drawn with radii r_c, $2r_c$, $\cdots$, centered around the destination as shown in Fig. 3. If the source has a hop distance of k to the destination, it will fall between the $(k-1)$-th and k-th rings. Second, $n-1$ mini-rings are drawn with equal space $r'_c = r_c/n$ between adjacent rings. The mini-hop distance m between two nodes is defined as $m = \lceil d/r'_c \rceil$. Thus, if the source-destination hop distance is k, its mini-hop distance satisfies $(k-1)n + 1 \leq m \leq nk$. The number n adjusts the level of refinement of discretization (higher n means closer approximation to continuous distances).

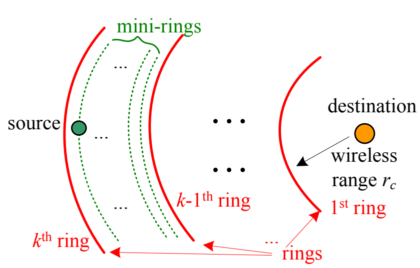

Fig. 3. The source-destination distance divided by rings and mini-rings.

The success probability S'_m is defined as the probability that message delivery is successful for a source-destination path with mini-hop distance of $m \geq 1$, and the delivery delay T'_m is defined as the delivery delay for a source-destination path with mini-hop distance of $m \geq 1$.

Given the complexity of combined social and communication networks, the direct derivation of S'_m (or T'_m) is mathematically intractable. Our approach is to derive a recursive solution to S'_m (or T'_m) that only includes the set of $\{S'_j\}_{1\leq j\leq m-1}$ (or $\{T'_j\}_{1\leq j\leq m-1}$) such that a numerical value of S'_m (or T'_m) can be computed given any network setup.

3.1.1 Delivery Success Probability

We first compute S'_m. Suppose that a source has a message to send to its destination with mini-hop distance m. If $1 \leq m \leq n$ (i.e., the source is within one hop to the destination), the source can always send the message directly to the destination using the communication link. Thus, the delivery success probability is the link success ratio of the communication link, i.e., $S'_m = \beta_c$ for $1 \leq m \leq n$.

Now consider the case that $m > n$. Under greedy routing, the source tries to find among its neighbors the next-hop node with the smallest mini-hop distance (to the destination) to forward the message. The next-hop node must have a mini-hop distance smaller than m. Let $Z_{m,j}$ denote the event that there exists a next hop node (via either communication or social link) that reduces the mini-hop distance by j, where $1 \leq j \leq m$, i.e., the message will be forwarded to a next-hop node with mini-hop distance $m-j$. The next-hop node will then use the same greedy routing strategy to forward the message. Thus, the delivery success probability from the next-hop node is S_{m-j}. Then, we can write S'_m recursively as

$$S'_m = \begin{cases} \beta_c \sum_{j=1}^{n} \mathbb{P}(Z_{m,j}) S'_{m-j} + \\ \quad \beta_s \sum_{j=n+1}^{m} \mathbb{P}(Z_{m,j}) S'_{m-j} & m \geq n+1 \\ \beta_c & 1 \leq m \leq n, \end{cases} \tag{1}$$

where $S'_0 = 1$ and $\mathbb{P}(e)$ is the probability of event e.

Next, we solve for $\mathbb{P}(Z_{m,j})$ in (1). We denote $E_{m,x}$ as the event that the forwarding node can find a node via either communication or social link that reduces the mini-hop distance by x, where $1 \leq x \leq m$. Recall that $Z_{m,j}$ is the event that the next hop node reduces the mini-hop distance by j, where $1 \leq j \leq m$. Thus, event $Z_{m,j}$ is equivalent to the event that $E_{m,j}$ happens but at the same time $E_{m,j+1}$, $E_{m,j+2}$, $\cdots$, and $E_{m,m}$ do not happen. Therefore, we can express $Z_{m,j}$ as

$$Z_{m,j} = E_{m,j} \bigcap \left(\bigcap_{x=j+1}^{m} E^c_{m,x} \right), \tag{2}$$

where $E^c_{m,x}$ denotes the complementary of event $E_{m,x}$, i.e., the event that $E_{m,x}$ does not happen. Because of the independent node distribution in the network, we have from (2) that

$$\begin{aligned} \mathbb{P}(Z_{m,j}) &= \mathbb{P}\left(E_{m,j} \bigcap \left(\bigcap_{x=j+1}^{m} E^c_{m,x} \right) \right) \\ &\approx \mathbb{P}(E_{m,j}) \prod_{x=j+1}^{m} \mathbb{P}\left(E^c_{m,x}\right) \\ &= p_{m,j} \prod_{x=j+1}^{m} (1 - p_{m,x}), \end{aligned} \tag{3}$$

where $p_{m,x} = \mathbb{P}(E_{m,x})$. Note that the approximation in (3) is to assume that $\{E_{m,x}\}_{x\in[j,m]}$ are independent of each other. From our comprehensive simulations, the approximation does not substantially affect the results in terms of success probability and delay of message delivery.

Computing $p_{m,x}$ (i.e., the probability that the forwarding node with mini-hop distance m can find a node to reduce the mini-distance by x) consists of two parts in terms of hop distance: $1 \leq x \leq n$ and $n+1 \leq x \leq m$.

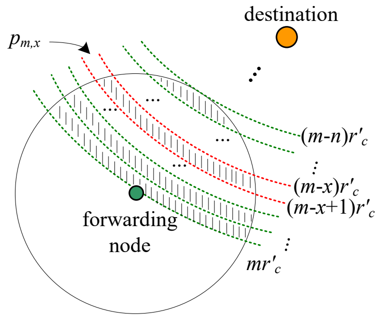

Fig. 4. Areas between mini-ring pairs inside the communication range.

1) $1 \leq x \leq n$: $p_{m,x}$ is the probability that there exists a node in the shaded area between adjacent mini-rings $(m-x)r'_c$ and $(m-x+1)r'_c$, as shown in Fig. 4. We denote by $E^{\text{IN}}_{m,x}$ and $E^{\text{OUT}}_{m,x}$ the events that the forwarding node can find a node between adjacent mini-rings $(m-x)r'_c$ and $(m-x+1)r'_c$ in Fig. 4 that reduces the mini-hop distance by x inside and outside the transmission range, respectively. If event $E^{\text{IN}}_{m,x}$ happens, the node can be reached via communication link. To compute $E^{\text{IN}}_{m,x}$, we denote $p_1 = \mathbb{P}((E^{\text{IN}}_{m,x})^c)$ as the probability that there exists no node on the area between mini-rings $(m-x)r'_c$ and $(m-x+1)r'_c$, which can be computed via Poisson point process approximation [29] as

$$p_1 = \exp(-\lambda(A((m-x+1)r'_c, nr'_c, mr'_c) \\ -A((m-x)r'_c, nr'_c, mr'_c))). \tag{4}$$

where $\lambda = N/(\pi R)^2$ is the node density on the network area, and $A((m-x+1)r'_c, nr'_c, mr'_c) - A((m-x)r'_c, nr'_c, mr'_c)$ is the area between mini-rings $(m-x)r'_c$ and $(m-x+1)r'_c$ inside the communication range, in which $A(r_a, r_b, d)$ is a function to compute the intersection area of two circles [30] with distance d that have radii r_a and r_b, respectively, satisfying

$$A(r_a, r_b, d) = r_a^2 \cos^{-1} \frac{r_a^2 + d^2 - r_b^2}{2dr_a} + \\ r_b^2 \cos^{-1} \frac{r_b^2 + d^2 - r_a^2}{2dr_b} - \\ \frac{\sqrt{(-d+r_a+r_b)(d-r_a+r_b)(d+r_a-r_b)(d+r_a+r_b)}}{2}.$$

If event $E^{\text{OUT}}_{m,x}$ happens, the node can be only reached via a social link. Fig. 5 shows an example of such an event: the potential next-hop node is outside the communication range of the forwarding node, but is socially connected to it. Thus, it can reach the next-hop via the social link to reduce the mini-hop distance by x.

In this case, as two nodes are socially connected with some probability ρ, this probability can be computed based

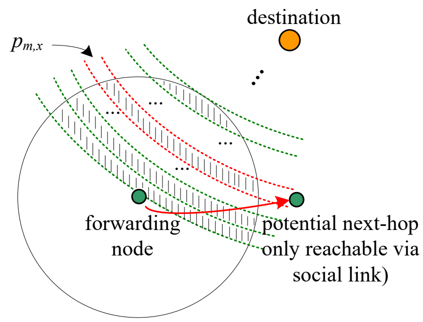

Fig. 5. An example of the scenario that event $E^{\text{OUT}}_{m,x}$ happens when $1 \leq x \leq n$.

on the social-communication link correlation model in Section 2, i.e.,

$$\rho = \gamma_{\text{NCS}} + (1 - \gamma_{\text{NCS}})\gamma_{\text{NCL}}. \tag{5}$$

We denote by $p_2 = \mathbb{P}((E^{\text{OUT}}_{m,x})^c)$ the probability that there exists no node on the area between mini-rings $(m-x)r'_c$ and $(m-x+1)r'_c$ outside the communication range. Using the Poisson point process approximation and the thinning theorem [29], we can get p_2 as

$$p_2 = \exp(-\rho\lambda((2(m-x)+1)\pi r'^2_c - A((m-x+1)r'_c, \\ nr'_c, mr'_c) + A((m-x)r'_c, nr'_c, mr'_c))). \tag{6}$$

It follows from (4) and (6) that

$$p_{m,x} = 1 - \mathbb{P}((E^{\text{IN}}_{m,x})^c)\mathbb{P}((E^{\text{OUT}}_{m,x})^c) = 1 - p_1 p_2. \tag{7}$$

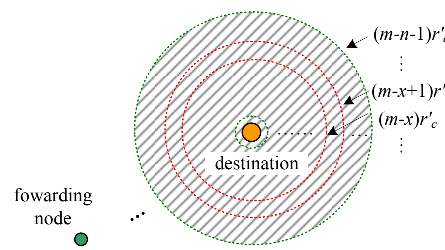

Fig. 6. Areas between mini-ring pairs outside the communication range.

2) $n+1 \leq x \leq m$: We consider two cases that $x = m$ and $n+1 \leq x < m$. If $x = m$, $p_{m,x}$ is the probability that the forwarding node is socially connected to the destination. Then, we have $p_{m,x} = \rho$ given in (5). If $n+1 \leq x < m$, $p_{m,x}$ is the probability that there exists a node in the shaded area between mini-rings $(m-x)r'_c$ and $(m-x+1)r'_c$, as shown in Fig. 6. For a Poisson point process with density $\rho\lambda$ on the area, we obtain

$$p_{m,x} = 1 - e^{\rho\lambda\pi(2(m-x)+1)r'^2_c}. \tag{8}$$

In summary, we have the delivery success probability at mini-hop distance of m as

$$S'_m = \begin{cases} \beta_c \sum_{j=1}^{n} \mathbb{P}(Z_{m,j}) S'_{m-j} + \\ \quad \beta_s \sum_{j=n+1}^{m} \mathbb{P}(Z_{m,j}) S'_{m-j} & m \geq n+1 \\ \beta_c & 1 \leq m \leq n, \end{cases} \tag{9}$$

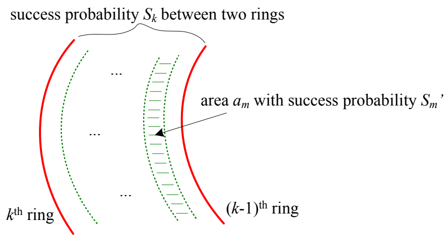

Fig. 7. Compute S_k from S'_m.

where the initial condition is $S'_0 = 1$, $\mathbb{P}(Z_{m,j}) = p_{m,j}\prod_{x=j+1}^{m}(1-p_{m,x})$,

$$p_{m,x} = \begin{cases} 1-p_1p_2 & 1 \le x \le n \\ 1-e^{\rho\lambda\pi(2(m-x)+1)r_c'^2} & n+1 \le x < m \\ \rho & x = m, \end{cases}$$

$\rho = \gamma_{\text{NCS}} + (1-\gamma_{\text{NCS}})\gamma_{\text{NCS}}$, and $\lambda = N/(\pi R)^2$. Accordingly, as shown in Fig. 7, the success probability S_k at hop distance k can be computed from mini-hop distance S'_m as

$$S_k = \frac{\sum_{m=(k-1)n+1}^{kn} a_m S'_m}{\pi(2k-1)r_c^2},$$

where a_m is the area between the $(m-1)$-th and m-th mini-rings, satisfying $a_m = (2m-1)r_c'^2$.

3.1.2 *Delivery Delay*

Next, we proceed to derive the average delivery delay T'_m. We denote by A_m the event that message delivery at mini-hop distance m is successful. Then, $\mathbb{P}(A_m)$ is the delivery success probability and it is given in (9). Conditioned on event A_m, the average delay at mini-hop distance m is expressed in a recursive way as

$$T'_m = \sum_{j=1}^{m} \mathbb{P}(Z_{m,j}|A_m)(D_j + T'_{m-j}), \tag{10}$$

where the initial condition is $T'_0 = 0$, D_j is the delay over j hops, satisfying

$$D_j = \begin{cases} D_c(\text{communication link delay}) & j \le n \\ D_s(\text{social link delay}) & j > n, \end{cases} \tag{11}$$

and

$$\begin{aligned} &\mathbb{P}(Z_{m,j}|A_m) = \mathbb{P}(Z_{m,j} \cap A_m)/\mathbb{P}(A_m) \\ &= \mathbb{P}(Z_{m,j} \cap ((B_{m,j} \cap Z_{m,j} \cap A_{m-j}) \cup \\ &\quad (\cap_{k=1,k\neq j}^{m}(B_{m,k} \cap Z_{m,k} \cap A_{m-k}))))/\mathbb{P}(A_m) \\ &= \mathbb{P}(B_{m,j})P(Z_{m,j})P(A_{m-j})/\mathbb{P}(A_m). \end{aligned} \tag{12}$$

In (12), $B_{m,j}$ is the event that the link from the node with mini-hop distance m to the node with mini-hop distance $m-j$ does not fail. It follows from (10) and (12) that

$$\begin{aligned} T'_m \;=\; & \beta_c \sum_{j=1}^{n} \frac{(\mathbb{P}(Z_{m,j})S'_{m-j}}{S'_m}(D_c + T'_{m-j}) \\ & +\beta_s \sum_{j=n+1}^{m} \frac{\mathbb{P}(Z_{m,j})S'_{m-j}}{S'_m}(D_s + T'_{m-j}). \end{aligned} \tag{13}$$

Then, the average delivery delay T_k at hop distance k can be computed from mini-hop distance delay T'_m as

$$T_k = \frac{\sum_{m=(k-1)n+1}^{kn} a_m T'_m}{\pi(2k-1)r_c^2},$$

where a_m is the area between the $(m-1)$-th and m-th mini-rings, satisfying $a_m = (2m-1)r_c'^2$.

Consequently, the delivery success probability and delay can be computed by (9) and (13), respectively. Note that (9) and (13) are based on the distance discretization technique, which separates the distance between two nodes into mini-rings with space r_c/n, as shown in Fig. 3. It is expected that the best approximation is achieved when $n \to \infty$ (i.e., the number of mini-rings used to discretize the continuous distance goes to infinity).

3.2 Simulation Validation

Simulations are used to validate the theoretical analysis and measure the performance of combined social-communication networks. We set up a 100-node network on a disk with radius 700 meters (m). The communication range of each node is $r_c = 65$m. We map the social network on the communication network with the following parameters: $\gamma_{\text{CS}} = 0.3$, $\gamma_{\text{CL}} = 0.01$, $\gamma_{\text{NCS}} = 0.01$, and $\gamma_{\text{NCL}} = 0.077$, which represent geographically correlated social relationships between nodes. These numbers are computed based on the Reality Mining dataset [31] that includes self-report relationship/friendship data and cell tower proximity information used in our experiments in Section 5. We use $n = 20$ to compute the theoretical results based on distance discretization in all simulations and experiments.

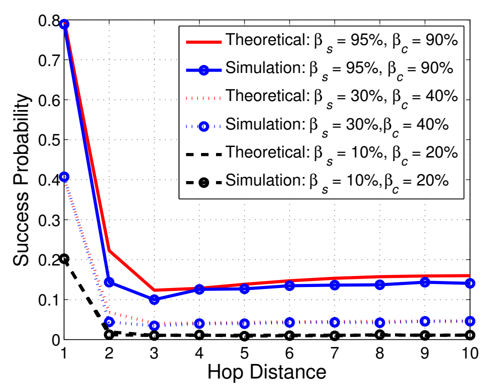

Fig. 8. Success probability with different communication and social link failures.

We measure the success probability in Fig. 8, which shows that as the hop distance increases, the success probability first sharply decreases then stabilizes and converges to some constant. In addition, more communication or social link failure leads to smaller success probability. This means the successful message delivery relies heavily on the reliability of communication and social links. Note that there are 100 nodes over the disk region with radius 700m and the wireless communication range $r_c = 65$m. This means that the average coverage is approximately $100 \times (65^2\pi)/(700^2\pi) \approx 0.86$ in the network. As a result, it is a relatively sparse random network scenario and a

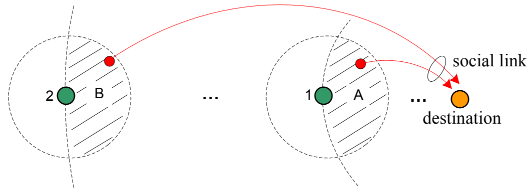

Fig. 9. Distant node has more chance to find through the communication link a next hop that has a direct social link to the destination under greedy routing.

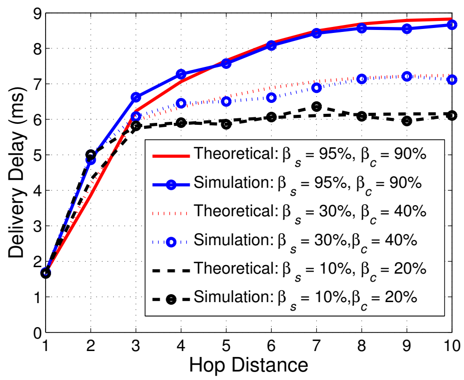

Fig. 10. Average delivery delay with different communication and social link failures.

node is likely not to find a next-hop node to forward data. Therefore, the success probability (when the hop distance is larger than 1) appears to be relative low in Fig. 8. Moreover, if we consider the network as a traditional multi-hop wireless ad hoc network without any social link, the success probability decays exponentially to zero as the hop distance increases. In contrast, by integrating social links into wireless networks, we can see that from Fig. 8 that the exponential decay to zero vanishes and the success probability remains positive even when the hop distance is large. This in fact shows that the joint greedy routing over the combined network brings benefits to the network performance.

It is worth noting that as Fig. 8 shows, after the initial sharp decrease, the success probability slightly increases when the hop distance increases. This indicates that larger distance between two nodes does not always lead to less reliable delivery. The reason behind this apparently counter-intuitive result is that greedy routing always attempts to move a message closer to the destination. However, a closer node in fact has less chance to find a next hop with a social link to the destination. An example is shown in Fig. 9, where node 1 is closer than node 2 to the destination. Under greedy routing, they will examine if there exist neighbors (as potential next hop nodes) in areas A and B, respectively. However, area B is larger than area A. This means that node 2 will have more chance to find a node through the communication link as the next hop that has a social link directly to the destination. As a result, Fig. 8 demonstrates that after the initial sharp decrease, the success probability then slightly increases and remains approximately constant as the hop distance increases.

Fig. 10 shows the average delivery delay as a function of hop distance. It is observed from Fig. 10 that the delay does not increase linearly as the hop distance increases, but starts to converge when the hop distance is large. The reason is that a node can always have a chance to find a social link that reduces the hop distance larger than 1. Moreover, an interesting phenomenon in Fig. 10 is that the delivery delay decreases as the link failure increases. The reason is that it is very hard to deliver a message over multiple hops with high failure probabilities. Therefore, when a message delivery is successful, it is very likely to be delivered over one or few hops. Thus, the delay is in fact smaller with higher failure probabilities.

Figs. 8 and 10 demonstrate that there is a close match between the theoretical analysis and simulation result (e.g., the maximum deviation for success probability is 6.7% at hop distance $d = 10$), thereby validating the distance discretization model to analyze the delay and success probability.

3.3 Further Analysis and Simulations

3.3.1 *Improve Reliability using Persistent Transmission*

In practice, either communication or social link could fail to deliver a message. In some applications with high reliability requirements, we can significantly improve the reliability of hop-by-hop message delivery by retransmitting a message until the transmission is successful at the cost of increasing the delay. This means that the one-hop reliability via social or communication link is guaranteed to be 1 while the expected delay is increased to D_s/β_s and D_c/β_c for social and communication links, respectively.

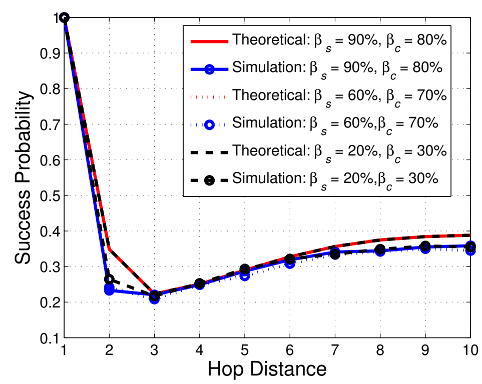

Fig. 11. Success probability under persistent transmission with different communication and social link failures.

Fig. 11 shows the success probability under persistent transmission. The delivery success probability always remains the same with different link failure probabilities. Therefore, persistent transmission is able to substantially improve the delivery success in the combined social and communication network.

The cost or penalty of persistent transmission is the degradation in delay performance. Fig. 12 demonstrates the delivery delay under persistent transmission. It is clear in Fig. 12 that higher failure probability increases the delivery delay. For example, when the network and social link success probabilities change from 80% and 90% to 30% and 20%, respectively, the delivery delay at the hop distance

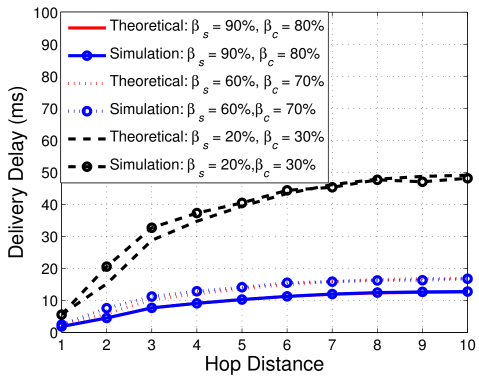

Fig. 12. Average delivery delay under persistent transmission with different communication and social link failures.

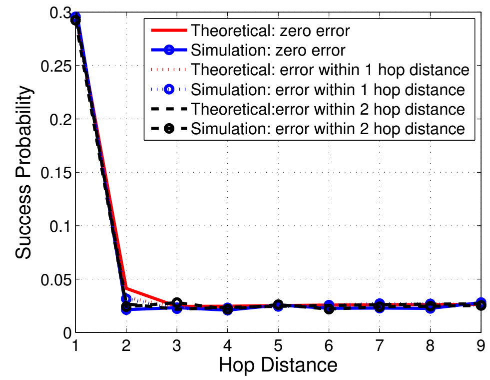

Fig. 13. Success probability with distance estimation errors.

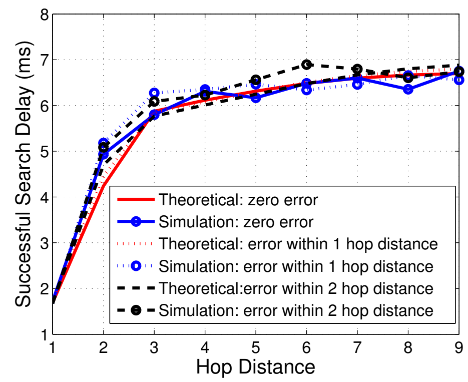

Fig. 14. Delivery delay with distance estimation errors.

of 10 increases from 12.7 ms to 48.12 ms under persistent transmission, as shown in Fig. 12. In practice, we should design the persistent transmission scheme with a reasonable trade-off between the success probability improvement and the delay degradation. For example, we should limit the number of persistent transmissions for delay-sensitive traffic, and choose a fairly large number of retransmissions for best-effort traffic.

3.3.2 Impact of Distance Estimation Error

If every node has only incomplete local information, it may estimate the distance values of its neighbors with errors. For example, estimation errors in signal based distance estimation methods are unavoidable [32]–[34]. We model such uncertainty in location information as the error probability $e_{a,b}$, which is the probability that a node with actual mini-hop distance a to a destination will be estimated as b.

To accommodate distance errors in analysis, we recompute the probability that the next forwarding node reduces the hop distance by j, i.e., $\mathbb{P}(\tilde{Z}_{m,j})$. First, we denote $\tilde{E}_{m,x}$ as the event that the forwarding node can find a node (not necessarily the next hop node because greedy routing always chooses the node nearest to the destination as next hop) via either communication or social link that reduces the mini-hop distance by x, where $1 \le x \le m$. This means that event $\tilde{Z}_{m,j}$ is equivalent to the event that $\tilde{E}_{m,j}$ happens but at the same time $\tilde{E}_{m,j+1}$, $\tilde{E}_{m,j+2}$, $\cdots$, and $\tilde{E}_{m,m}$ do not happen, or they happen but the errors make the forwarding node believe they have longer distances.

Generally speaking, it is difficult to precisely analyze the success probability and delay with arbitrary distance errors. However, if the distance error is relatively small compared with the mini-hop distance m, we can neglect the effect of neighbors with original distances longer than the distance of the forwarding node, and obtain approximately

$$\begin{aligned}\tilde{Z}_{m,j} \;\approx\;& \tilde{E}_{m,j} \cap (\cup_{j^*=1}^{m-1}(U_{m-j,m-j^*} \cap (\cap_{x\neq j^*}(\\ & \tilde{E}^c_{m,x} \cup (\tilde{E}_{m,x} \cap (\cup_{e=1}^{j^*-1} U_{m-x,m-e})))))), \quad (14)\end{aligned}$$

where $U_{a,b}$ denotes the event that the distance a is estimated as b, and $\mathbb{P}(U_{a,b}) = e_{a,b}$. Accordingly we have

$$\begin{aligned}\mathbb{P}(\tilde{Z}_{m,j}) \;\approx\;& \tilde{p}_{m,j} \sum_{j^*=1}^{m-1} (e_{m-j,m-j^*} \Pi_{x=1,x\neq j^*}^{m} \\ & ((1-\tilde{p}_{m,x}) + \tilde{p}_{m,x} \sum_{e=1}^{j^*} e_{m-x,m-e})), \quad (15)\end{aligned}$$

where $\tilde{p}_{m,x} = \tilde{\mathbb{P}}(E_{m,x})$. Then, the success probability follows immediately.

Figs. 13 and 14 show the success probability and delivery delay comparisons, respectively, between theoretical analysis and simulation results under different distance estimation errors. The simulation setups are the same as those in Figs. 8 and 10 except for the distance estimation error: when each node estimates the distance, the estimation error is uniformly distributed within 1 or 2 hop distance.

It is seen from Figs. 13 and 14 that the simulation results closely match the theoretical analysis, and also the distance error does not have a significant impact on the routing performance.

3.3.3 Impact of Mobility on Performance

Next, we analyze how mobility can affect the message delivery. We consider the worst case for a multi-hop wireless network: when a node transmits a message to the next hop, the next hop always moves out of the transmitting node's wireless range and also farther away to the destination. In other words, the next hop distance always increases after the delivery. In this setup, messages cannot be delivered by communication links, because a node will always move out of the communication range during the message transmission.

For each transmission, we denote by w the maximum increase of the mini-hop distance of the next hop. That is, when a node makes a routing decision to transmit to the next hop with mini-hop distance m, after the transmission, the mini-hop distance of the next hop becomes $m+w$ in the worst case. In this regard, a message can never be delivered in wireless networks with communication links. However, it can still have a chance to be delivered when social links in a combined social and communication network are used.

We denote by S_m^{worst} the success probability of mini-hop distance m under such a worst-case scenario. Because communication links always fail due to the worst-case mobility, S_m^{worst} becomes a function of social link success probability β_s, and we can write S_m^{worst} in a similar way as

$$S_m^{\text{worst}} = \begin{cases} \beta_s \mathbb{P}(Z_{m,m}) + & \\ \ \beta_s \sum_{j=n+1}^{m-1} \mathbb{P}(Z_{m,j}) S_{m-j+w}^{\text{worst}} & m \geq n+1 \\ \beta_s & 1 \leq m \leq n. \end{cases} \tag{16}$$

For a practical network, it is feasible to assume that the mini-hop distance is always bounded above; i.e., $m \leq M_{\max}$, where $M_{\max} = \lceil 2Rn/r_c \rceil$ is the maximum mini-hop distance between two nodes at the network boundaries. Then, we can formulate and solve (16) as a linear system for all S_m^{worst}, $1 \leq m \leq M_{\max}$. Table 2 shows a numerical example of the success probability due to the worst-case mobility when (i) routing is performed with communication links only, (ii) combined social and communication links are used.

TABLE 2
Success probability vs. hop distance.

Hop distance:	1	3	5	7	9
Success probability (i)	0%	0%	0%	0%	0%
Success probability (ii)	6.1%	6.2%	6.5%	6.9%	7.3%

Table 2 shows that if we use communication links only, the success probability is always 0 because the next hop will always move farther away from the destination. However, if the routing design uses both social and communication links, there always exists a positive probability that a message can be delivered via social links even under the worst-case mobility for multi-hop wireless networks.

4 Improved Routing Protocols

In this section, we present how routing with only local information can be adapted to improve either the delay or the success probability by using analytical results and network inference.

4.1 Improve the Delay Performance

We first show how to improve the delay performance. As shown in Fig. 15, suppose that node A knows social link delay D_s, communication link delay D_c, and the average mini-hop distance advancement denoted by h_m at mini-hop distance m by using the communication link. When node A finds a social link to forward a message to node B that can reduce the mini-hop distance by d, it knows that the delivery delay is D_s, if it uses the social link, and it also knows that the average delivery delay to node B will be dD_c/h_m, if it can find a path to node B with communication links only. This means that if node A wants to reduce the delivery delay, it should take the risk to find a path with communication links only to node B if $dD_c/h_m < D_s$.

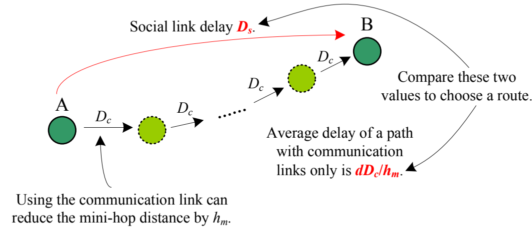

Fig. 15. Illustration of how a node improves the delay.

Thus, greedy routing for delay improvement chooses a social link compared to a communication link if the mini-hop distance d between two nodes satisfies

$$d \geq (h_m D_s)/D_c, \tag{17}$$

where h_m is given by

$$h_m = \frac{\sum_{j=1}^{n} j\mathbb{P}(Z_{m,j})}{n \sum_{j=1}^{n} \mathbb{P}(Z_{m,j})}. \tag{18}$$

Then, the success probability becomes

$$S_m^{\text{opt1}} = \begin{cases} \beta_c \sum_{j=1}^{n} \mathbb{P}(Z_{m,j}) S'_{m-j} + & \\ \ \beta_s \sum_{j=y(h_m)}^{m} \mathbb{P}(Z_{m,j}) S'_{m-j} & m \geq n+1 \\ \beta_c & 1 \leq m \leq n, \end{cases} \tag{19}$$

where $y(h_m) = n+1+\lfloor \frac{n h_m D_s}{D_c} \rfloor$, and $S_0^{\text{opt1}} = 1$. In a similar way, we obtain the delay as

$$\begin{aligned} T'_m &= \beta_c \sum_{j=1}^{n} \frac{\mathbb{P}(Z_{m,j}) S'_{m-j}}{S'_m} (D_c + T'_{m-j}) \\ &\quad + \beta_s \sum_{j=y(h_m)}^{m} \frac{\mathbb{P}(Z_{m,j}) S'_{m-j}}{S'_m} (D_s + T'_{m-j}). \end{aligned} \tag{20}$$

4.2 Improve the Delivery Reliability

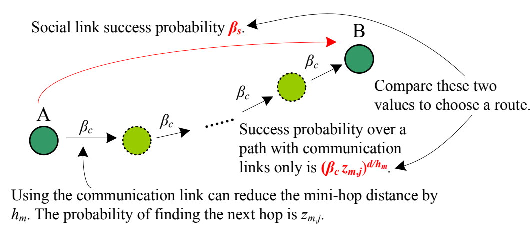

Fig. 16. Illustration of how a node improves the success probability.

Similar to the greedy routing for delay improvement, as shown Fig. 16, the greedy routing for success probability improvement chooses a social link compared to a communication link in the combined network, if the mini-hop distance satisfies

$$d \geq \frac{h_m \log \beta_s}{\log(z_{m,j}\beta_c)}, \tag{21}$$

where h_m is the average hop advancement for the communication link, and $z_{m,j} = \sum_{j=1}^{n} \mathbb{P}(Z_{m,j})$ denotes the probability that node A can find a next-hop node through the communication link. Then, the success probability and delay have the same expressions as (19) and (20), respectively, where $y(h_m)$ is different and follows from

$$y(h_m) = n + 1 + \left\lfloor \frac{n h_m \log \beta_s}{\log(\sum_{j=1}^{n} \mathbb{P}(Z_{m,j})\beta_c)} \right\rfloor. \tag{22}$$

4.3 Improved Routing with Practical Implementation

Note that improved routing must work with the knowledge of D_s, D_c, β_s, β_c, h_m, and $z_{m,j}$. The first four terms are point-to-point parameters that can be measured locally between a node and its neighbors. The last two terms are non-local network statistics, which can be both computed theoretically from node density λ and the probability ρ that two nodes are socially connected.

If ρ and λ are unknown, a node can reverse-engineer them out of observations (i.e., using network inference to obtain ρ and λ). In particular, a node can compute the success probability $\hat{S}_m$ from its messages received at different mini-hop distance m (e.g., the node can compute the probability by comparing the sequence numbers of received messages). Then, based on the observation $\hat{S}_m$ and the theoretical modeling S'_m, the following grid search is performed to estimate the network parameters ρ and λ:

$$(\hat{\rho}, \hat{\lambda}) = \arg\min_{\rho,\ \lambda} \sum_m |S'_m(\rho, r_c) - \hat{S}_m|. \tag{23}$$

The estimates $(\hat{\rho}, \hat{\lambda})$ are sent to the routing process to improve either the delivery delay or the success probability. Note that developing an efficient solution to (23) is outside the scope of this paper. We simply use exhaustive grid search for our experiments. In a practical system, a destination node can measure $\hat{S}_m$ by checking the sequence numbers of packets. For example, if one packet with sequence number 1005 is followed by another packet with sequence number 1010, the destination can deduce that 4 packets are lost, from which it can measure $\hat{S}_m$.

It is worth noting that the delay measurements can also be used as a way to infer ρ and λ, i.e.,

$$(\hat{\rho}, \hat{\lambda}) = \arg\min_{\rho,\ \lambda} \sum_m |T'_m(\rho, r_c) - \hat{T}_m|. \tag{24}$$

An issue in (24) is that the destination must have a synchronized clock with the source in order to precisely measure the time delay, which is usually not available. Thus, we choose to use (23) as a network inference method to improve the routing performance.

4.4 Energy-Saving Routing

In addition to the two types of improved routing protocols, another routing strategy can be built upon network inference to save energy consumption at each node. Note that after performing network inference, a node can intelligently foresee message delivery outcomes with inferred information. One direct application of such capability is to save energy in the network by avoiding transmitting delay-sensitive messages that have a high chance to miss the deadline. Specifically, a node can infer an approximate delay distribution for each hop distance. When it has a message to transmit/forward, it can first compute the probability that the message could miss the deadline. If the probability is larger than a threshold, the node can simply discard the message to save energy due to the high chance that the message cannot reach the destination on time. We will evaluate the performance of such a protocol in experiments.

5 Testbed and Experiments

In this section, we set up a proof-of-concept testbed to measure the performance of combined social and communication networks under greedy routing and its improved versions. The testbed consists of programmable WiFi radios, RouterStation Pros [35], that represent network nodes, an Ethernet switch, and a high-fidelity multi-hop wireless channel emulator, called RFnest$^{\text{TM}}$ [36].

5.1 Network Setup

Each node has one WiFi interface as the wireless communication link. All nodes are connected via (radio frequency) RF cables to network channel emulator RFnest$^{\text{TM}}$ that can attenuate realistic RF signals according to any specific network topology. They are also connected via Ethernet cables to an Ethernet switch to emulate the social links. A social network server is also connected with the switch. Any message transmitted over a social link will go to the server first, then be forwarded to the next hop. Thus, the social link delay and failures are emulated at the social server to accommodate various social network and link conditions. During our experiments, the total number of nodes is set to be 21 due to current hardware configurations.

5.2 Social Dataset

We use the Reality Mining dataset [31] to generate social connections between nodes in the testbed. The Reality Mining project was conducted from 2004-2005 at the MIT Media Laboratory. There are measurements of 94 subjects using mobile phones pre-installed with software that recorded and sent the data about call logs, Bluetooth devices in proximity of approximately five meters, cell tower IDs, application usage, and phone status. The dataset also collected self-report relational data from each individual, where subjects were asked about friendship with others. We use the data of 21 individuals with the largest number of friends. We also use the network connectivity data in Reality Mining to set up the wireless network topology in RFnest$^{\text{TM}}$.

5.3 Performance of Greedy Routing

We first measure the success probability and delivery delay in Figs. 17 and 18, respectively. It is observed from Fig. 17 that as the hop distance increases, the success probability

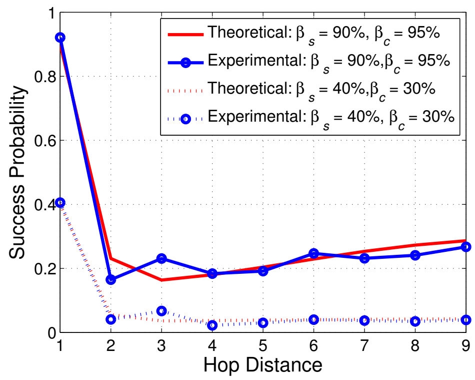

Fig. 17. Measurements of success probability with different communication and social link failures.

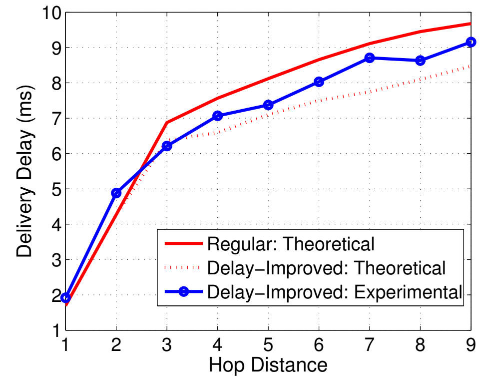

Fig. 19. Delivery delay under delay-improved greedy routing (14.5% inference error, 20% communication and 40% social link failure).

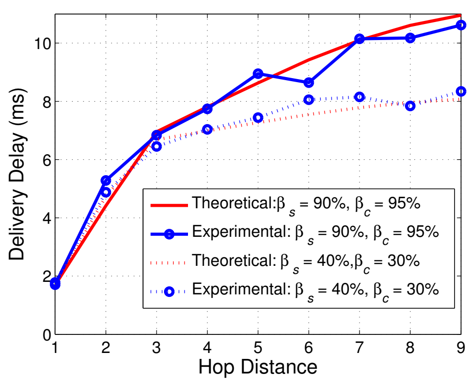

Fig. 18. Measurements of delivery delay with different communication and social link failures.

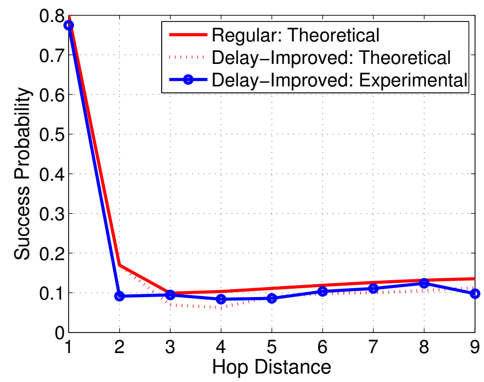

Fig. 20. Success probability under delay-improved greedy routing (14.5% inference error, 20% communication and 40% social link failure).

first sharply decreases then converges. In addition, if the social link is more reliable with low link failure probability, the end-to-end success probability will also be improved. Note that it is also found in experimental results that when the hop distance increases, the success probability also slightly increases. This means that under greedy routing, farther nodes may have larger success probability than closer nodes for message delivery to the destination.

Fig. 18 shows the delivery delay as a function of hop distance. The delay does not increase linearly as the hop distance increases, but starts to converge when the hop distance is large. This is because a node can always have a chance to find a social link that reduces the hop distance larger than 1.

Figs. 17 and 18 also show that there is a reasonable match between the theoretical analysis and experimental results. For example, the maximum derivation for success probability is 20.1% at hop distance $d = 3$ and the average deviation is 9.3%. Therefore, our modeling does provide a good prediction for the performance of the testbed system.

5.4 Performance of Improved Greedy Routing

Next, we consider improved greedy routing to improve delay and success probability. Note that in practice, not every node has knowledge of network parameters. Thus, nodes have to obtain network parameters via network inference. During our experiments, the probability ρ that two nodes are socially connected and node density λ are both unknown to every node. The improved greedy routing always uses $\hat{\rho}$ and $\hat{\lambda}$ obtained from network inference (23) to make routing decisions. We found that even with sufficiently large number of samples of received packets, there always exists an error floor of 10% - 20% errors to deduce network parameters to improve greedy routing. The error floor is due to two causes: (i) the sampled measurements are subject to hardware randomness and errors; and (ii) the theoretical analysis is based on an approximation technique, which will lead to small errors in practice.

We first consider the delay-improved greedy routing. Fig. 19 shows the delivery delay under delay-improved greedy routing. We can see that although there is a performance gap for the improved greedy routing between ideal and inferred network parameters, the inference results can still evidently help reduce the delay for the improved greedy routing compared with the regular greedy routing.

Fig. 20 illustrates the comparison of success probability between delay-improved routing and regular greedy routing. We can observe from Fig. 20 that there exists a slight penalty to improve the delay performance as the success probability is reduced under the delay-improved routing.

We next evaluate the performance of success probability-improved greedy routing. Table 3 shows the comparison of experimental results at hop distances 2 and 3 between regular greedy routing and improved greedy routing for

TABLE 3
Performance of Success Probability-Improved Routing.

Hop distance:	Regular	Improved
2	1.326e-1	1.387e-1
3	7.034e-2	8.984e-2

Parameters: 14.5% inference error, 10% communication and 99% social link failures.

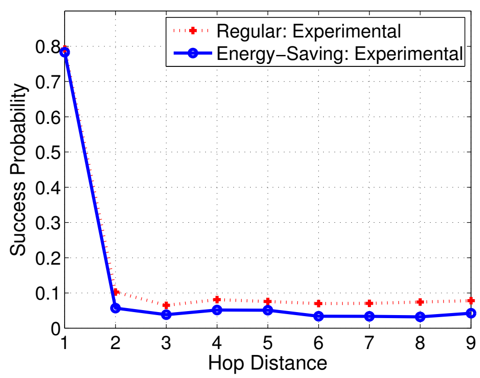

Fig. 21. Comparison of success probability under regular greedy routing and energy-saving greedy routing (20% communication and 40% social link failure).

success probability. Table 3 shows that the success probabilities are very small with slight improvements. This is because if a social link is so unreliable that nodes have to choose a communication link, it is still very hard to find a next hop towards the destination via the communication link as there are only a finite number of nodes in the network.

5.5 Energy-Saving Routing

An energy-saving routing is also implemented to measure how energy can be saved by estimating message delivery outcomes based on network inference. In the testbed, we generate messages with a 5-ms deadline. Every node is implemented to discard a message if it infers that the message cannot be delivered with a probability larger than 80%. Fig. 21 shows that there is a slight performance loss in terms of success probability for such an energy-saving scheme. However, Fig. 22 shows that the average delivery delay is significantly reduced to values within the 5-ms deadline.

Fig. 23 shows the ratio of the number of saved transmissions with the energy-saving scheme over that without energy-saving, which essentially represents how much energy is saved. It is noted from Fig. 23 that when the hop distance is large, around 80% energy can be saved to avoid unnecessary transmission of delay-sensitive messages.

5.6 Discussions and Limitations

We designed and implemented three routing protocols on top of the testbed to decrease the delay of message delivery, improve reliability of message delivery, and save transmission energy, respectively. Our studies showed that adequately combining social links with wireless network design can substantially benefit wireless networks in many

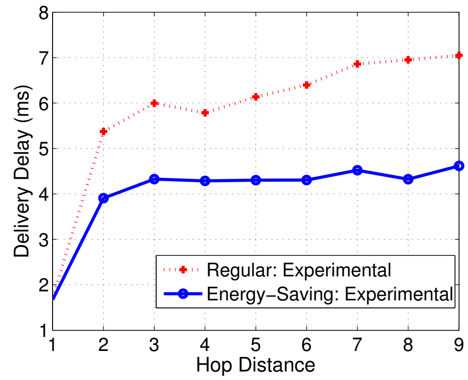

Fig. 22. Comparison of delivery delay under regular greedy routing and energy-saving greedy routing (20% communication and 40% social link failure).

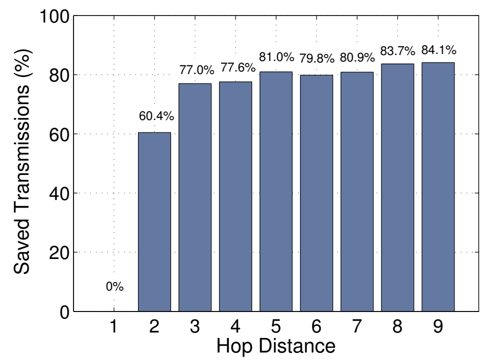

Fig. 23. Percentage of saved transmissions in energy-saving greedy routing (20% communication and 40% social link failure).

perspectives such as delay, reliability and energy saving. Today's network infrastructures provide an underlying architecture that overlays social networks over wireless communication medium. Therefore, combined social and communication design for data delivery is a promising technique for network performance optimization.

5.6.1 *Complexity of Analytical Framework*

In this paper, the analytical framework to solve the probability of delivery success (e.g., (9)) and delay (e.g., (13)) requires recursive computations with computational complexity proportional to the parameter n in the distance discretization approach. This means that the computation time becomes longer if we increase n to improve the computation accuracy (and the time will grow to ∞ when $n \to \infty$). From our simulations, we find that a range of $n \in [15, 25]$ is sufficient to obtain an approximate match between the theoretical and simulated results, while reducing the computation time in theoretical analysis. For example, in Figs. 8 and 10, we choose $n = 20$ and observe the approximate matches under various conditions.

5.6.2 *Limitations*

In this paper, we present an exploratory work to analyze the network performance of a new type of wireless networks with social links integrated as additional data links into the wireless communication links. Although today's network systems have not yet adopt such a design paradigm, their underlying infrastructures as well as people's social conductivities do provide all necessary conditions for a combined social and communication network design. Such a combined network design shows potential benefits to improve the network performance, such as delivery reliability and delay. Our proof-of-concept testbed evaluation was conducted at a relatively small scale. Hence, thorough investigations are still needed to understand how a combined network design can perform in a practical large-scale network environment. The work in this paper can be served as the first step to a comprehensive understanding on how to combine social and communication network design.

6 Related Work

Social-Aware Routing: Many social-aware routing protocols have been developed in the literature for DTNs. These designs leverage social relationships to improve the network performance [1]–[6]. In these protocols, social ties are typically used as abstract or conceptual links for a node's decision making in routing, and are not used as links for data delivery. In particular, DTNs aim to tolerate the delay.

There have been few studies on data delivery and routing (with local information only) in purely social networks [28], where the delay is the main objective without considering the potential link failures and there is also no communication network or overlaid/hybrid network architecture considered.

Combined Social and Communication Network Analysis: Some recent studies have already considered a combined network with both social and communication links [7]–[10] under the geometry-based greedy routing. The typical assumption for analytical foundations in these works is that there exist an infinite number of users in a finite-area network such that the greedy routing [11]–[14] can be applied by assuming a next hop neighbor closer to the destination can be always found. For example, [7] and [10] analyzed the performance of a combined network with an infinite number of users, and [9] considers a mobile scenario, in which there are an infinite number of nodes moving around in the network.

Nonetheless, the assumption of infinite nodes does not always hold. There are two under-explored issues in modeling and evaluating a combined social and communication network with a finite number of nodes:

1) It is not clear how to analyze the performance of a practical network with a finite number of nodes, where message delivery may fail not only due to social or communication link failures, but also due to the non-existence of a next-hop node closer to the destination.
2) Existing design and schemes (either social aware routing heuristics or combined network analysis with an infinite number of nodes) were validated in simulations only, but it is still unknown how routing will perform in system-level experiments with actual radios used for communication links and a realistic social network dataset used for social links.

Our work addresses the issues in existing studies by modeling the performance of message delivery under greedy routing in a combined social-mobile network with a finite number of nodes and link unreliability. We further improve the greedy routing strategy used in existing studies based on our modeling and analysis, and evaluate the improved strategies in comprehensive simulations and real-world implementations.

Note that this paper extends the preliminary models proposed in our conference version [37] in three aspects: (i) we integrate more practical settings, including distance estimation error and mobility, into the analytical framework; (ii) we propose different improved routing strategies based on the modeling and analysis; and (iii) we perform comprehensive simulations and experiments to validate the theoretical analysis. The framework of [37] for finite node analysis has been applied in [38] to learning in mobile ad hoc networks under the assumption of reliable links.

7 Conclusion

In this paper, we systematically studied the performance of combined social and communication networks with a finite number of nodes. We proposed a distance discretization technique to derive the success probability and delay of message delivery. The model incorporates the effects of persistent transmission, potential error in distance estimation, and mobility. We used the analytical results and the inferred network properties to improve routing protocol design in terms of success probability, delay, and energy saving. We built a testbed to implement routing strategies to improve the delivery reliability, energy, and delay performance. We also conducted a variety of experiments on the testbed to measure the performance of social and communication networks. Our results motivate the use of social links in a wireless network to substantially improve the performance of end-to-end message delivery. The future work includes adapting the modeling into realistic real-world protocols and finding the best application domains for social-model network integration.

References

[1] F. Li, H. Jiang, H. Li, Y. Cheng, and Y. Wang, "SEBAR: Social energy based routing for mobile social delay tolerant networks," *IEEE Trans. Vehicular Technology*, 2017.

[2] Q. Li, S. Zhu, and G. Cao, "Routing in socially selfish delay tolerant networks," in *Proc. of IEEE INFOCOM*, Mar. 2010.

[3] K. Wei, X. Liang, and K. Xu, "A survey of social-aware routing protocols in delay tolerant networks: Applications, taxonomy and design-related issues," *IEEE Communications Surveys & Tutorials*, vol. 16, 2014.

[4] J. Hom, L. Good, and S. Yang, "A survey of social-based routing protocols in delay tolerant networks," in *Proc. of ICNC*, Jan 2017, pp. 788–792.

[5] K. Chen and H. Shen, "SMART: Lightweight distributed social map based routing in delay tolerant networks," in *Proc. of IEEE ICNP*, 2012.

[6] E. Bulut and B. Szymanski, "Exploiting friendship relations for efficient routing in delay tolerant mobile social networks," *IEEE Trans. Parallel and Distributed Systems*, 2012.

[7] K. Neema, Y. E. Sagduyu, and Y. Shi, "Search delay and success in combined social and communication networks," in *Proc. of IEEE Globecom*, 2013.

[8] M. W. Bigrigg, K. M. Carley, K. Manousakis, and A. McAuley, "Routing through an integrated communication and social network," in *Proc. of MILCOM*, 2009.

[9] Y. E. Sagduyu and Y. Shi, "Navigating a mobile social network," *IEEE Wireless Communications Magazine*, Oct. 2015.
[10] Y. E. Sagduyu, Y. Shi, and K. Neema, "Search in combined social and wireless communication networks: Delay and success analysis," *IEEE Trans. Wireless Communications*, Sept. 2015.
[11] J. M. Kleinberg, "Navigation in the small world," *Science*, 2000.
[12] R. West and J. Leskovec, "Human wayfinding in information networks," in *In Proc. of WWW*, 2012.
[13] O. Simsek and D. Jensen, "Navigating networks by using homophily and degree," *PNAS*, 2008.
[14] O. Yagan, D. Qian, J. Zhang, and D. Cochran, "Information diffusion in overlaying social-physical networks," in *In Proc. of CISS*, 2012.
[15] A. Mei, G. Morabito, P. Santi, and J. Stefa, "Social-aware stateless forwarding in pocket switched networks," in *Proc. of IEEE INFOCOM*, 2011.
[16] J. Wu and Y. Wang, "Social feature-based multi-path routing in delay tolerant networks," in *Proc. of IEEE INFOCOM*, 2012.
[17] X. Lin, R. Lu, X. Liang, and X. Shen, "STAP: A social-tier-assisted packet forwarding protocol for achieving receiver-location privacy preservation in VANETs," in *Proc. of IEEE INFOCOM*, 2011.
[18] P. Costa, C. Mascolo, M. Musolesi, and G. P. Picco, "Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 748–760, 2008.
[19] M. Ji, G. Caire, and A. F. Molisch, "Fundamental limits of caching in wireless d2d networks," *IEEE Trans. Information Theory*, vol. 62, no. 2, pp. 849–869, Feb 2016.
[20] B. Liu, Z. Liu, and D. Towsley, "On the capacity of hybrid wireless networks," in *Proc. of IEEE INFOCOM*, vol. 2, 2003, pp. 1543–1552.
[21] U. C. Kozat and L. Tassiulas, "Throughput capacity of random ad hoc networks with infrastructure support," in *Proc. of ACM MobiCom*, 2003, pp. 55–65.
[22] M. Zhao and W. Wang, "A unified mobility model for analysis and simulation of mobile wireless networks," *ACM-Springer Wireless Networks*, vol. 15, no. 3, pp. 365–389, April 2009.
[23] A. L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.
[24] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, pp. 440–442, 1998.
[25] S. Schnettler, "A small world on feet of clay? A comparison of empirical small-world studies against best-practice criteria," *Social Networks*, vol. 31, pp. 179–189, 2009.
[26] ——, "A structured overview of 50 years of small-world research," *Social Networks*, vol. 31, pp. 165–178, 2009.
[27] M. C. H. Inaltekin and H. V. Poor, "Average message delivery time for small-world networks in the continuum limit," *IEEE Trans. Information Theory*, vol. 56, pp. 4447–4470, 2010.
[28] H. Inaltekin, M. Chiang, and H. Poor, "Delay of social search on small-world random geometric graphs," *Journal of Mathematical Sociology*, 2012.
[29] M. Penrose, *Random Geometric Graphs*. Oxford University Press, 2003.
[30] Circle-Circle Intersection, http://mathworld.wolfram.com/Circle-CircleIntersection.html.
[31] RealityMining, http://realitycommons.media.mit.edu/realitymining.html.
[32] C.-H. Chang and W. Liao, "Revisiting relative location estimation in wireless sensor networks," in *Proc. of IEEE ICC*, 2009.
[33] T. K. Le and N. Ono, "Closed-form and near closed-form solutions for TDOA-based joint source and sensor localization," *IEEE Trans. Signal Processing*, vol. 65, no. 5, March 2017.
[34] A. Akhlaq, R. McKilliam, R. Subramanian, and A. Pollok, "Selecting wavelengths for least squares range estimation," *IEEE Trans. Signal Processing*, vol. 64, no. 20, pp. 5205–5216, Oct 2016.
[35] Router Station Pro - Ubiquiti networks, http://www.ubnt.com/routerstation.
[36] J. Yackoski, B. Azimi-Sadjadi, A. Namazi, J. H. Li, Y. E. Sagduyu, and R. Levy, "RFnest: Radio frequency network emulator simulator tool," in *Proc. of MILCOM*, 2011.
[37] Z. Lu, Y. E. Sagduyu, and Y. Shi, "Friendships in the air: Integrating social links into wireless network modeling, routing, and analysis," in *Proc. of IEEE INFOCOM-Workshop on Network Science for Communication Networks (INFOCOM-NetSciCom)*, Apr. 2016.
[38] Y. E. Sagduyu, Y. Shi, S. J. Mackey, D. H. Cansever, M. P. Patel, and B. F. Panettieri, "Analytical framework for MANET learning and routing," in *Proc. of MILCOM*, 2017.

Zhuo Lu is an Assistant Professor in the Department of Electrical Engineering, University of South Florida. He received his B.S. and M.S. degrees from Xidian University, Xi'an China, in 2002 and 2005, respectively, and his Ph.D. degree in computer engineering from North Carolina State University, Raleigh NC, in 2013. His research interests include network science, cyber security, data analytics, cyber-physical systems, mobile computing and wireless networking. He is a member of IEEE, ACM and USENIX.

Yalin E. Sagduyu (S'02–M'08–SM'15) received the B.S. degree in Electrical and Electronics Engineering from Bogazici University, Turkey, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Maryland at College Park. He is an Associate Director at Intelligent Automation Inc., Rockville, MD, USA. His research interests are in the areas of design and optimization of wireless networks, cognitive radio networks, network coding, information theory, security, game theory, social networks, machine learning, and big data analytics. He co-chaired several IEEE and ACM Workshops on cognitive radio networks. He also served as the MAC and Cross-Layer Design Track Co-Chair at IEEE PIMRC, 2014, and in the organizing committee of IEEE GLOBECOM 2016.

(S'02–M'08–SM'13) received his Ph.D. degree in Computer Engineering from Virginia Tech in 2007. He is an Adjunct Assistant Professor in the Bradley Department of Electrical and Computer Engineering, Virginia Tech and a Senior Research Scientist in Intelligent Automation Inc. Dr. Shi's research focuses on algorithm design, optimization, and machine learning for cognitive radio networks, MIMO and cooperative communication networks, sensor networks, ad hoc networks, satellite networks, and social networks. Dr. Shi was IEEE Communications Surveys and Tutorials Exemplary Editor in 2014. He was a recipient of Best Paper Award at IEEE INFOCOM 2008, a recipient of the only Best Paper Award Runner-Up at IEEE INFOCOM 2011, and a recipient of Best Student Paper Award at ACM WUWNet 2014. Dr. Shi is an Editor for IEEE Communications Surveys and Tutorials.