Capacity of Wireless Networks with Social Behavior

Bita Azimdoost, Hamid R. Sadjadpour, *Senior Member, IEEE,*
and J. J. Garcia-Luna-Aceves, *Fellow, IEEE*

Abstract—The capacity of a wireless network is studied when nodes communicate with one another in the context of social groups. All the nodes are assumed to have the same number of independent long-range social contacts, one of which each selects randomly as its destination. The Euclidean distance between a source and its social group members follows a power-law distribution and communication between any two nodes takes place only within the physical transmission range resulting in communication over multi-hop paths. The capacity order of such a composite network is derived as a function of the number of nodes, the social-group concentration, and the size of social groups. Our results demonstrate that when each node has constant number of contacts which does not increase with network size growth, and are geographically concentrated, then the network behaves similar to social networks and communication network does not have any effect on the throughput capacity. On the other hand, when the social contact population grows in time, or social connectivity among nodes is highly distributed, then the communication network is the dominant factor and the composite network behaves similar to wireless networks, i.e., the capacity is the same as Gupta and Kumar results. When neither social connectivity nor communication network is dominant, then the throughput capacity results are between these two extreme cases.

Index Terms—Social networks, wireless networks, throughput capacity.

I. INTRODUCTION

STARTING with the work by Gupta and Kumar [1], the order throughput capacity of wireless communication networks has been studied extensively in the recent past, and all this prior work has assumed that sources select their destinations according to a uniform distribution. However, some studies [2] show that the social influence, which is the probability of having interaction, is heavily determined by the distance. In other words, if the network nodes are social entities or their relations are defined in a social context, the probability of communicating is proportional to the inverse of the Euclidean distance between two nodes with some power

exponent. It has also been observed [3], [4] that there are some patterns common in many different types of social networks ; the examples include small-world [4], and power-laws [3] to name a few.

Backstorm *et al.* [3] observed that the likelihood of having contact with a person decreases with Euclidean distance and follows a power-law distribution. Although prior results have made significant contribution to the understanding of the behavior of wireless networks, but these results on the capacity order of wireless networks did not consider the social characteristics of these networks.

On the other hand, as a result of early work by Milgram on the small-world phenomenon [4], the modeling of social networks have received considerable attention. For example, Watts and Strogatz [5] divided the edges of a network into local and long-range contacts and assume that there is always an edge between a node and any of its local or long-range social contacts. Dietzfelbinger *et al.* [6] calculated the average number of steps between any source and target along a ring-based network in which each node is connected to its left and right neighbors and possibly to some further vertices, and the long-range contacts may be selected through any distribution. Fraigniaud *et al.* [7] assumed that the probability of a node being the long-range contact of a source is proportional to the rank of their distance among the distances from the source to all the other nodes and derived the upper bound for the expected number of steps for any source-target pair. Kleinberg [8] introduced a model for the characterization of the small-world phenomenon consisting of a two-dimensional extended grid with point-to-point links in which each node has four local contacts and one long-range contact. The source node s selects any other node v as its long-range contact with a probability proportional to $d^{-\alpha}(s, v)$, where $d(s, v)$ is the lattice distance between s and v, and $\alpha \geq 0$ shows the density of the social network. Based on these assumptions, some upper or lower bounds were derived for different social network densities. In summary, the analysis of social networks did not consider the effect of multihop communication on the capacity order.

What is needed to understand the true performance of wireless networks is a model that captures the restrictions imposed by the communication infrastructure, together with the distribution of flows rendered by social groups. Therefore, a model for composite networks that has both social and communication characteristics is needed. In this regard, Li *et al.* [9] studied the capacity of a wireless network in which source-destination pairs follow a power-law distribution as in

Kleinberg's model; however, they provide only upper bounds that need not be tight and provide no insight on the impact of social-group sizes. More recently, Azimdoost *et al.* [10] studied the impact of the interaction between communication and social networks on the wireless network capacity, by considering four local contacts and a single long-range contact per node, with the source knowing the location of its four local contacts and the destination. The source-destination pair selection follows a power-law distribution that is a function of the Euclidean distance between the source and the rest of the nodes. While this model is a marked improvement over models that assume a uniform distribution for source-destination pairs, its results are limited in scope because a node usually has more than one long-range social contact in its social group.

In this paper, we study the case of a wireless network in which nodes communicate with others in the context of social groups. Section II introduces the notation and some definitions and results used throughout the paper.

Section III shows that the original power-law distribution introduced by Kleinberg [8] cannot be used when the number of long-range contacts q is a function of the total number of nodes n in the network. In fact, this limitation was also mentioned by Kleinberg [8] . In addition, a modified power-law distribution is introduced that is applicable for all values of $q(n)$.

The main contribution of this paper is stated in the following theorem, which considers what we call a *wireless social or composite network* . In such a network, each of the n nodes has a social group consisting of $q(n)$ long-range contacts selected independently. Long-range social contacts are selected based on the power-law distribution with parameter α identified in Section III, and one of those long-range contacts is the destination of the node's flow. Communication between any two nodes can take place only if they are within transmission range ($r(n)$) and such communication succeeds according to the protocol model of multiple access interference [1].

Theorem 1.1: Consider a wireless network consisting of n nodes with social behavior modeled by the following properties.

- Any two nodes in distance d away from each other are socially connected with a probability inversely proportional to d^α, where α is the social group density.
- All the nodes have exactly q independent social contacts where $q = 1, .., n - 1$.
- Each source selects one of its social contacts as its destination randomly with no preference.

Under these conditions and assuming the wireless transmission range $r(n) = \Theta(\sqrt{\frac{\log n}{n}})$ to guarantee connectivity [11], the maximum capacity order in this wireless social network is

$$\begin{cases} \Theta(\frac{1}{nr(n)}) & for \ q = \Theta(n) \\ \Theta(\frac{1}{nr(n)}) & for \ (q, \frac{q}{n}) \stackrel{n \to \infty}{\to} (\infty, 0) \\ \Theta(\frac{n-q+1}{n^2 r(n)}) & for \ q < \infty, 0 \le \alpha < 2 \\ \Theta(\frac{n-q+1}{n^2 r^{\alpha-1}(n)}) & for \ q < \infty, 2 \le \alpha \le 3 \\ \Theta(\frac{n-q+1}{n^2 r^2(n)}) & for \ q < \infty, 3 < \alpha \end{cases}$$

Section IV presents the proof of Theorem 1 by deriving the upper and lower bounds of the throughput capacity that

coincide for the various values of $q(n)$. This result shows that the scaling properties of a wireless network are a function of the density (α) and size of the social groups q. If the size of social groups is proportional to the network size, then the order network capacity is the same as if no social groups existed, which is the same result by Gupta and Kumar [1]. Interestingly, this is the case even when the size of social groups is an insignificant fraction of the number of nodes n as n goes to infinity, and is intuitive by noticing that in such a case source-destination pairs must consume communication resources along large multi-hop paths linking sources with destinations. By contrast, when social group sizes do not grow as fast as the number of nodes in the network n, and their spread is localized ($\alpha \ge 2$), then such network has capacity order much higher than a pure wireless network. Section V concludes the paper by discussing the implications of our results.

To the best of our knowledge, this is the first work that considers the interaction between social groups and the underlying wireless communication infrastructure in an analytical framework of the capacity order of wireless networks. However, our model still lacks many characteristics of social networks that should be added in the future works in order to better understand the behavior of composite networks. For this reason, the results presented in this paper are also limited in scope.

II. PRELIMINARIES

The network is a dense network in a unit square area with n uniformly distributed nodes. We use the protocol model [12] to determine the success of communication in the presence of multiple access interference (MAI). In particular, if χ_i, χ_j and χ_k denote the Cartesian positions in the unit square area for nodes v_i, v_j and v_k, assuming that node $v_k \ne v_i$ transmits on the same sub-channel at the same time as v_i, and $r(n)$ is the common transmission range of all the nodes in the network, then node v_i can successfully transmit to node v_j if $|\chi_i - \chi_j| \le r(n)$ and $|\chi_k - \chi_j| \ge (1 + \Delta)r(n)$, where $\Delta > 0$ is the guard zone factor. To guarantee connectivity in this network [11], the transmission range ($r(n)$) is assumed to be $r(n) = \Omega(\sqrt{\log n/n})$.[1]

As Figure 1 illustrates, a TDMA medium access control scheme is assumed to avoid MAI. The network area is divided into square cells with side-length $C_1 r(n), (C_1 < \frac{1}{4})$, and at any given time the cells separated by M-cell distance are the only cells allowed to transmit as shown with a cross sign inside the cells in figure 1 where $M \ge (2 + \Delta)/C_1$.

The routing of information is very simple. Each node is assumed to know the locations of its intended destination and its immediate neighbors, and selects as its next hop to the destination that local contact that is closest to the destination. The local contacts are within the radio range since they are the one hop physical neighbors of the node. Assuming that there is at least one local contact in each of the four adjacent cells of the source guarantees that this simple routing protocol

---

[1]For n points placed uniformly at random on the unit square, the probability that there is no node in the $r(n)$ vicinity of any selected node tends to zero if $r(n)$ is at least $\Theta(\sqrt{\frac{\log n}{n}})$.
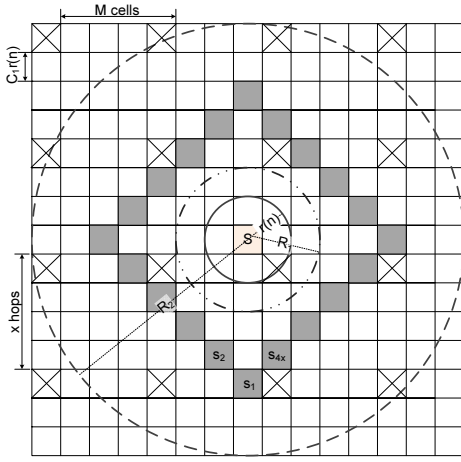
Fig. 1.   The solid-line circle shows the transmission range. Dark gray cells (s_i) contain the nodes with $X = x$. R_1 (R_2) are used as the distance of each node in this region instead of their real distances to achieve upper (lower) bounds on $P(X = x)$.

converges. If each node has more than four local contacts, i.e., all nodes within transmission range are local contacts, then the order throughput capacity computation does not change and the same results can be derived. The four local contacts assumption was first considered in [8] for grid networks.

We use the notation of [13] to denote the elementary symmetric polynomials of the variables $x = (x_1, ..., x_n)$ by $\sigma_{p,n}, 1 \le p \le n$. In other words,

$$\sigma_{p,n}(x) = \sigma_{p,n}(x_1, ..., x_n) = \sum_{1 \le i_1 < i_2 < .. < i_p \le n} x_{i_1}...x_{i_p}.$$

Moreover, we define the elementary symmetric polynomials of the same set of variables except one, x_k, as

$$\sigma_{p,n-1}^{\overline{k}}(x_1, ..., x_n) = \sigma_{p,n-1}(x_1, ..., x_{k-1}, x_{k+1}, ..., x_n).$$

Lemma 2.1: [13] Let $x_1, ..., x_n$ be non-negative real numbers, $n \ge 2$. Then for $1 \le p \le n - 1$, we have

$$\sigma_{1,n}\sigma_{p,n} \ge \frac{n(p+1)}{n-p}\sigma_{p+1,n}.$$

In Section IV (Lemma 4.1), we prove that this is a tight bound for values of p that do not grow as fast as n.

The standard notations of O and Ω are used to describe the asymptotic upper and lower bounds respectively. When $f(n) = \Theta(g(n))$, then it is denoted by $f(n) \equiv g(n)$.

III. A POWER-LAW DISTRIBUTION FOR SOCIAL GROUPS

In Kleinberg's model [8], every node s has a directed edge to every other node v_i within lattice distance $p \ge 1$, and directed edges to $q \ge 0$ other nodes using independent random trials. Each directed edge from s has endpoint $v_i, i = 1, .., n$ with probability proportional to $d_i^{-\alpha} \triangleq d^{-\alpha}(s, v_i)$ and normalizing factor $\sum_{i=1}^{n} d_i^{-\alpha}$. Considering the same probability distribution function for *long-range social contacts* (LSC), the probability that the LSC list contains exactly q independently selected members is the summation of all possible q-member

subsets of nodes probabilities.

$$P(|LSC| = q) = \sum_{1 \le i_1 < ... < i_q \le n} P(LSC = \{v_{i_1}, ..., v_{i_q}\})$$

$$= \sum_{1 \le i_1 < ... < i_q \le n} \prod_{j=1}^{q} P(v_{i_j} \in LSC)$$

$$= \sum_{1 \le i_1 < ... < i_q \le n} \frac{d_{i_1}^{-\alpha}...d_{i_q}^{-\alpha}}{(\sum_{j=1}^{n} d_j^{-\alpha})^q}.$$

where v_{i_j} is the i_j^{th} node in the network for $j = 1, ..., q$ and $i_j = 1, ..., n$. As can be seen, this probability is close to one for $q = \Theta(1)$, decreases by increasing q, and approaches zero when $q = \Theta(n)$. Kleinberg [8] assumed that q is a universally constant value and the above derivation proves that the original power-law distribution used in his paper should be modified to consider those cases when q is a function of n. We assume that each source node has the same number of LSCs $q(n)$ selected in independent random trials.

The long-range contacts are selected independently, while closer nodes to the source have a better chance of being selected as a LSC, thus, the probability that a particular q-member set is the LSC set is proportional to the product of the inverse of the distances of its members from the source. This probability can be written as

$$P(LSC = \{v_{i_1}, ..., v_{i_q}\}) = \frac{d_{i_1}^{-\alpha}...d_{i_q}^{-\alpha}}{N_{\alpha,q}}. \quad (1)$$

The normalization factor $N_{\alpha,q}$ is obtained using the fact that $\sum_{1 \le i_1 < ... < i_q \le n} P(LSC = \{v_{i_1}, ..., v_{i_q}\}) = 1$.

$$N_{\alpha,q} = \sum_{1 \le i_1 < ... < i_q \le n} d_{i_1}^{-\alpha}...d_{i_q}^{-\alpha} \quad (2)$$

The probability that a particular node v_k is selected as a LSC (i.e., the probability that v_k is a member of the LSC set) is given by[2]

$$P(v_k \in LSC)$$

$$= \sum_{1 \le i_1 < ... < i_{q-1} \le n, i_j \ne k} P(LSC = \{v_k, v_{i_1}, ..., v_{i_{q-1}}\}),$$

$$= \frac{\sum_{1 \le i_1 < ... < i_{q-1} \le n, i_j \ne k} d_k^{-\alpha} d_{i_1}^{-\alpha}...d_{i_{q-1}}^{-\alpha}}{\sum_{1 \le i_1 < ... < i_q \le n} d_{i_1}^{-\alpha}...d_{i_q}^{-\alpha}}.$$

The above probability function denotes the probability of node v_k being in LSC, and is non-decreasing in q. It also guarantees that the described process ends up with a q-member LSC set for each source node.

Let ϑ_t be a random variable denoting the destination node. Then, for each particular $v_k \in V$ (the set of nodes except source), we have

$$P(\vartheta_t = v_k) = P(\vartheta_t = v_k \mid v_k \in LSC) \times P(v_k \in LSC)$$
$$+ P(\vartheta_t = v_k \mid v_k \notin LSC) \times P(v_k \notin LSC).$$

Given that the destination is only selected from LSCs, $P(\vartheta_t = v_k \mid v_k \notin LSC) = 0$. Furthermore, the selection

---

[2]Note that we assume that |LSC| is equal to q for all sources.

of destination from LSCs has a uniform distribution.

$$P(\vartheta_t = v_k) = \frac{1}{q}P(v_k \in LSC)$$

$$= \frac{\sum_{1 \leq i_1 < ... < i_{q-1} \leq n, i_j \neq k} d_k^{-\alpha} \prod_{j=1}^{q-1} d_{i_j}^{-\alpha}}{q \sum_{1 \leq i_1 < ... < i_q \leq n} \prod_{j=1}^{q} d_{i_j}^{-\alpha}}$$

Let $v = (v_1, ..., v_n)$ denote $(d_1^{-\alpha}, ..., d_n^{-\alpha})$, then the above equation can be written as

$$P(\vartheta_t = v_k) = \frac{d_k^{-\alpha} \sigma_{q-1, n-1}^{\overline{k}}(v)}{q \sigma_{q,n}(v)}. \qquad (3)$$

IV. THROUGHPUT CAPACITY ANALYSIS

Let λ denote the data rate for each node and X be the number of hops traveled by each bit from source to destination. The total number of concurrent transmissions per second in such a network is then $n\lambda E[X]$, where $E[X]$ is the average number of hops in a route for any given source-destination pair. This value is upper bounded by the total bandwidth W available, divided by the number of non-interfered groups in the TDMA scheme as shown in Figure 1 (*i.e.*, $\frac{W}{M^2 C_1^2 r^2(n)}$). Therefore, the maximum data rate for each node is [10]

$$\lambda \leq \lambda_{max} = \Theta(\frac{1}{nr^2(n)E[X]}). \qquad (4)$$

The average number of hops can be computed as

$$E[X] = \sum_{x=1}^{x_{max}} xP(X=x) = P(X=1) + \sum_{x=2}^{x_{max}} xP(X=x).$$

$P(X = 1)$ is the probability that the packets travel just one hop from source to destination, and its value resides between 0 and 1. Since each packet needs to travel at least one hop from the source to reach the destination, the average number of hops between the sources and destinations cannot be less than 1. Therefore, $P(x = 1)$ does not change the order of the $E[X]$ and can be ignored when deriving the order of expected number of hops.

To compute $P(X = x)$ for $x > 1$, we need to consider the long-range contacts outside the circle with radius $r(n)$ centered at the source node. Given that all the nodes inside the transmission range of a source receive the data transmitted from it in just one hop, $P(X = x) = 0$ for $1 < x < \lceil \frac{1}{C_1} + 1 \rceil$. The information between source and destination located on two opposite corners of the network area passes through the maximum number of hops which is $\lceil \frac{2}{C_1 r(n)} \rceil$. Thus, $E[X]$ can be calculated as

$$E[X] \equiv \sum_{\lceil \frac{1}{C_1} + 1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} xP(X=x).$$

To compute $P(X = x)$ for $x = \lceil \frac{1}{C_1} + 1 \rceil, ..., \lceil \frac{2}{C_1 r(n)} \rceil$, we need to compute the number of nodes at a distance of x hops from the source and their corresponding Euclidean distances from the source. The geometric place of such nodes is a rhombus around the source node as shown in Figure 1 and explained in [10]. The probability that the number of hops between source and destination is x hops equals the probability that the destination is located in one of the cells on the boundaries of this rhombus. Hence,

$$P(X = x) = \sum_{l=1}^{4x} P(destination\ is\ inside\ s_l)$$

$$= \sum_{l=1}^{4x} \sum_{v_k\ in\ s_l} P(\vartheta_t = v_k).$$

Therefore,

$$E[X] \equiv \sum_{\lceil \frac{1}{C_1} + 1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k\ in\ s_l} P(\vartheta_t = v_k)$$

$$\equiv \sum_{\lceil \frac{1}{C_1} + 1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k\ in\ s_l} \frac{d_k^{-\alpha} \sigma_{q-1,n-1}^{\overline{k}}(v)}{q \sigma_{q,n}(v)} \qquad (5)$$

We now compute the average number of hops based on different values of q as a function of n.

A. Case I: q grows with n

If $q = n$, then $E[X]$ can be rewritten as

$$E[X] \equiv \sum_{x=\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k\ in\ s_l} \frac{d_k^{-\alpha} \sigma_{n-1,n-1}^{\overline{k}}(v)}{n \sigma_{n,n}(v)}.$$

Since

$$d_k^{-\alpha} \sigma_{n-1,n-1}^{\overline{k}}(v) = d_k^{-\alpha} \prod_{i=1, i \neq k}^{n} d_i^{-\alpha}$$

$$= \prod_{i=1}^{n} d_i^{-\alpha} = \sigma_{n,n}(v),$$

then $E[X] \equiv \sum_{x=\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k\ in\ s_l} \frac{1}{n}$.

Because nodes are uniformly distributed over the network area, there are $nC_1^2 r^2(n)$ nodes inside each cell s_l with high probability. Thus[3]

$$E[X] \equiv \sum_{x=\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} 4x^2 C_1^2 r^2(n)$$

$$\equiv r^2(n) \int_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} u^2 du \equiv \frac{1}{r(n)}.$$

Hence, the per-node throughput capacity is $\frac{1}{nr(n)}$, which will lead to the same result obtained by Gupta and Kumar ($\frac{1}{\sqrt{n \log n}}$) [1], if we use the minimum transmission range necessary to guarantee connectivity. This result is consistent, because the number of social contacts is equal to the total number of nodes in the network, and one of these nodes is selected randomly and uniformly as the destination, which is a similar assumption to that of the original work by Gupta and Kumar [1].

The second case is when $q = \Theta(n)$ but $q \neq n$. Define i.i.d. random variables $Y_i = d_i^{-\alpha}$ for $1 \leq i \leq n$ and define the

---

[3]Note that we are computing the order of $E[X]$ dropping constant factors.

sequence $Z_i = \log Y_i$ for all values of i. It is obvious that Z_i are i.i.d. as well. Utilizing the law of large numbers, we have $\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} Z_i = \overline{Z}$ where $\overline{Z}$ is the expected value of random variable Z_i. Thus equation (3) can be computed as

$$
\begin{aligned}
P(\vartheta_t = v_k) &\equiv \frac{\sum_{1 \le i_1 < .. < i_q \le n, \exists h : i_h = k} \prod_{j=1}^{q} Y_{i_j}}{q \sum_{1 \le i_1 < .. < i_q \le n} \prod_{j=1}^{q} Y_{i_j}} \\
&\equiv \frac{\sum_{1 \le i_1 < .. < i_q \le n, \exists h : i_h = k} \exp \sum_{j=1}^{q} Z_{i_j}}{q \sum_{1 \le i_1 < .. < i_q \le n} \exp \sum_{j=1}^{q} Z_{i_j}}, \\
&\equiv \frac{\sum_{1 \le i_1 < .. < i_q \le n, \exists h : i_h = k} \exp q \overline{Z}}{q \sum_{1 \le i_1 < .. < i_q \le n} \exp q \overline{Z}} \\
&\equiv \frac{\binom{n-1}{q-1}}{q \binom{n}{q}} = \frac{1}{n}.
\end{aligned}
$$

Therefore, the value of $E[X]$ is similar to the case $q = n$.

$$
E[X] \equiv \sum_{x = \lceil \frac{1}{C_1} + 1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} \frac{1}{n} \equiv \frac{1}{r(n)}.
$$

Using equation (4) provides the capacity as $\lambda_{max} = \Theta(\frac{1}{nr(n)})$.

B. Case II: n grows much faster than q

In this case, the expected number of hops between source and destination is obtained when $\lim_{n \to \infty} \frac{q}{n} = 0$, and two mutually exclusive situations must be considered, namely: $lim_{n \to \infty} q = \infty$ and $\lim_{n \to \infty} q < \infty$.

When $\lim_{n \to \infty} q = \infty$, we can use law of large numbers and a similar procedure as before to arrive at

$$
E[X] = \Theta(\frac{1}{r(n)}), \lambda_{max} = \Theta(\frac{1}{nr(n)}).
$$

When each node has finite number of contacts ($\lim_{n \to \infty} q < \infty$), the numerator of $P(\vartheta_t = v_k)$ can be expanded as

$$
\begin{aligned}
&d_k^{-\alpha} \sigma_{q-1,n-1}^{\overline{k}}(v) = d_k^{-\alpha} \sigma_{q-1,n}(v) - d_k^{-2\alpha} \sigma_{q-2,n-1}^{\overline{k}}(v) \\
&= d_k^{-\alpha} \sigma_{q-1,n}(v) - d_k^{-2\alpha} \sigma_{q-2,n}(v) + d_k^{-3\alpha} \sigma_{q-3,n-1}^{\overline{k}}(v).
\end{aligned}
$$

Note that $d_k^{-\alpha}$ and $\sigma_{q-i,n-j}$ are positive values; therefore, the upper and lower bounds for $P(\vartheta_t = v_k)$ are obtained as

$$
\begin{aligned}
d_k^{-\alpha} \frac{\sigma_{q-1,n}(v) - d_k^{-\alpha} \sigma_{q-2,n}(v)}{q \sigma_{q,n}(v)} &\le P(\vartheta_t = v_k), \\
P(\vartheta_t = v_k) &\le \frac{d_k^{-\alpha} \sigma_{q-1,n}(v)}{q \sigma_{q,n}(v)}. \quad (6)
\end{aligned}
$$

Lemma 4.1: Let $\Psi = \{\psi_1, ..., \psi_n\}$ be a set of $n \ge 2$ non-negative real numbers. Then for a finite p, i.e., $\lim_{n \to \infty} p < \infty$, we have

$$
\frac{\sigma_{1,n}(\Psi) \sigma_{p,n}(\Psi)}{(p+1) \sigma_{p+1,n}(\Psi)} = \Theta \left(\frac{n}{n-p} \right). \quad (7)
$$

Proof: Define random variables $U_i^p = \psi_{i_1}...\psi_{i_p}$ for $i = 1, .., \binom{n}{p}$ where $1 \le i_1 < .. < i_p \le n$. Due to symmetry, these random variables are identically distributed. Moreover, their mean $\overline{U_p}$ is a function of p. It can be easily seen that these random variables are not independent, as they may

have common factors of ψ_{i_j}. We partition the set Ψ into p-member subsets. Assume that T^p is the set of all possible such partitioning (each denoted by T_i^p) with no common member, i.e., $T_i^p \cap T_j^p = \phi$. For a finite p, the number of T^p members is $|T^p| \equiv \binom{n}{p} / (\frac{n}{p}) = \binom{n-1}{p-1}$.

Now we can expand $\sigma_{p,n}(\Psi)$ to separate summations over different partitions described above. Thus,

$$
\sigma_{p,n} = \sum_{1 \le i_1 < .. < i_p \le n} \psi_{i_1}..\psi_{i_p} = \sum_{j=1}^{|T^p|} \sum_{\{\psi_{i_1}..\psi_{i_p}\} \in T_j^p} \psi_{i_1}..\psi_{i_p}.
$$

Because each inner summation is applied over one possible partitioning of Ψ, it is performed over $\frac{n}{p}$ of independent U_i as described before. The law of large numbers can be applied here.

$$
\lim_{n \to \infty} \sum_{\{\psi_{i_1}..\psi_{i_p}\} \in T_j^p} \psi_{i_1}..\psi_{i_p} = \lim_{n \to \infty} \sum_{\{\psi_{i_1}..\psi_{i_p}\} \in T_j^p} U_i^p = \frac{n}{p} \overline{U_p}
$$

Thus,

$$
\sigma_{p,n} = \sum_{j=1}^{|T^p|} \frac{n}{p} \overline{U_p} = \binom{n}{p} \overline{U_p}.
$$

A similar formulation can be derived for $\sigma_{p+1,n}(\Psi)$.

$$
\sigma_{p+1,n} = \sum_{j=1}^{|T^{p+1}|} \frac{n}{p+1} \overline{U_{p+1}} = \binom{n}{p+1} \overline{U_{p+1}}
$$

Therefore,

$$
\frac{\sigma_{1,n} \sigma_{p,n}}{(p+1) \sigma_{p+1,n}} = \frac{\sigma_{1,n} \binom{n}{p} \overline{U_p}}{(p+1) \binom{n}{p+1} \overline{U_{p+1}}}.
$$

Note that U_i^p have identical distribution and ψ_i are i.i.d.. Therefore, the expected value $\overline{U_{p+1}}$ can be expressed in terms of $\overline{U_p}$

$$
\begin{aligned}
\overline{U_{p+1}} &= E[U_i^{p+1}] = E[\psi_{i_1}...\psi_{i_{p+1}}] \\
&= \sum_{\psi_{i_{p+1}}} E[\psi_{i_1}...\psi_{i_p} \psi_{i_{p+1}} | \psi_{i_{p+1}}] p(\psi_{i_{p+1}}), \\
&= \sum_{\psi_{i_{p+1}}} \psi_{i_{p+1}} E[\psi_{i_1}...\psi_{i_p}] p(\psi_{i_{p+1}}) \\
&= \overline{U_p} \sum_{\psi_{i_{p+1}}} \psi_{i_{p+1}} p(\psi_{i_{p+1}}) \\
&= \overline{U_p} . \overline{\psi_{p+1}} = \overline{U_p} . \overline{\psi}.
\end{aligned}
$$

Furthermore, by utilizing law of large numbers for $\sigma_{1,n}$ results in $\sigma_{1,n}(\Psi) \to n \overline{\psi}$. Thus

$$
\frac{\sigma_{1,n}(\Psi) \sigma_{p,n}(\Psi)}{(p+1) \sigma_{p+1,n}(\Psi)} \equiv \frac{n \binom{n}{p}}{(p+1) \binom{n}{p+1}} = \frac{n}{n-p}. \quad \blacksquare
$$

Returning to the case of finite contacts, we use Lemma 4.1 (for $p = q - 1$) and inequality (6) to obtain an upper bound

for $E[X]$ in equation (5).

$$E[X] \leq \sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} \frac{d_k^{-\alpha} \sigma_{q-1,n}(v)}{q\sigma_{q,n}(v)}$$

$$\equiv \frac{n}{n-q+1} \sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} \frac{d_k^{-\alpha}}{\sigma_{1,n}} \quad (8)$$

Referring to the results presented in [10], it can be observed that the average number of hops in this case is $\frac{n}{n-q+1}$ times more than the case when there is only one long-range contact for each source. To calculate the above summation, we need to compute the distance between each node in s_i and the source. To simplify the problem, we use distances $R_1 = xC_1 r(n)/A_1$ and $R_2 = A_2 xC_1 r(n)$ $(A_1, A_2 > 1)$ for all such nodes to reach upper and lower bounds for this summation (see figure 1).

$$\sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} (A_2 xC_1 r(n))^{-\alpha} \leq \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} d_k^{-\alpha}$$

$$\leq \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} (xC_1 r(n)/A_1)^{-\alpha}$$

By replacing the number of nodes in each cell by $nC_1^2 r^2(n)$ and ignoring the constant values in the above inequality, we can see that the order of both upper and lower bounds are the same.

$$\sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} d_k^{-\alpha}$$

$$\equiv nr^{2-\alpha}(n) \sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x^{2-\alpha}$$

$$\overset{a}{\equiv} nr^{2-\alpha}(n) \int_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil+1} u^{2-\alpha} du$$

The last equality (a) is obtained by replacing the sum by its integral approximation. After computing that integral for a sufficiently large value of n which leads to sufficiently small transmission range, we arrive at

$$\sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} d_k^{-\alpha}$$

$$\equiv \begin{cases} \Theta(\frac{n}{r(n)}) & , for \ 0 \leq \alpha \leq 3 \\ \Theta(\frac{n}{r^{\alpha-2}(n)}) & , for \ 3 \leq \alpha \end{cases} \quad (9)$$

Moreover, $\sigma_{1,n}$ can be written as

$$\sigma_{1,n} = \sum_{v_k} d_k^{-\alpha} \equiv \int_{r(n)}^{\gamma d_{max}} nu^{1-\alpha} du,$$

where d_{max} is the maximum distance between any two nodes in the network, and $\gamma \leq 1$. Calculating the integral for a

sufficiently large value of n leads to

$$\sigma_{1,n} \equiv \begin{cases} \Theta(n) & for \ 0 \leq \alpha \leq 2 \\ \Theta(\frac{n}{r^{\alpha-2}(n)}) & for \ 2 \leq \alpha \end{cases} \quad (10)$$

The derivations of equations (9) and (10) are described in the Appendix.

Now we can use these results in equation (8) to obtain the following upper bound for $E[X]$. Note that $E[X] \geq 1$; therefore, if the computation ends up with $E[X] < 1$, we replace it with 1.

$$E[X] = \begin{cases} O(\frac{n}{n-q+1} \frac{1}{r(n)}) & for \ 0 \leq \alpha < 2 \\ O(\frac{n}{n-q+1} \frac{1}{r^{3-\alpha}(n)}) & for \ 2 \leq \alpha \leq 3 \\ O(\frac{n}{n-q+1}) & for \ 3 < \alpha \end{cases}$$

The lower bound capacity follows immediately.

$$\lambda_{max} = \begin{cases} \Omega(\frac{n-q+1}{n^2 r(n)}) & for \ 0 \leq \alpha < 2 \\ \Omega(\frac{n-q+1}{n^2 r^{\alpha-1}(n)}) & for \ 2 \leq \alpha \leq 3 \\ \Omega(\frac{n-q+1}{n^2 r^2(n)}) & for \ 3 < \alpha \end{cases}$$

Thus, these are the upper bounds of $E[X]$ and the lower bounds on the capacity if the number of long-range contacts is a finite number greater than one.

To compute the lower bound for $E[X]$, we will study the lower bound of $P(\vartheta_t = v_k)$ in equation (6). First, we calculate the order of $\frac{\sigma_{q-2,n}(v)}{q\sigma_{q,n}(v)}$. This value is obtained by replacing $p = q - 1$ and $p = q - 2$ in equation (7).

$$\frac{\sigma_{1,n} \sigma_{q-1,n}}{q\sigma_{q,n}} = \Theta\left(\frac{n}{n-q+1}\right)$$

$$\frac{\sigma_{1,n} \sigma_{q-2,n}}{(q-1)\sigma_{q-1,n}} = \Theta\left(\frac{n}{n-q+2}\right)$$

By multiplying these two equations and combining with equation (10), we arrive at

$$\frac{\sigma_{q-2,n}}{q\sigma_{q,n}} = \Theta(\frac{(q-1)n^2}{(n-q+1)(n-q+2)\sigma_{1,n}^2})$$

$$= \begin{cases} \Theta(\frac{(q-1)}{(n-q+1)(n-q+2)}) & for \ 0 \leq \alpha \leq 2 \\ \Theta(\frac{(q-1)r^{2\alpha-4}(n)}{(n-q+1)(n-q+2)}) & for \ 2 \leq \alpha \end{cases} \quad (11)$$

The lower bound for $E[X]$ is derived by combining equations (5) and (6).

$$E[X] \geq$$

$$\sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} \frac{d_k^{-\alpha} \sigma_{q-1,n}(v) - d_k^{-2\alpha} \sigma_{q-2,n}(v)}{q\sigma_{q,n}(v)}$$

$$= \frac{\sigma_{q-1,n}(v)}{q\sigma_{q,n}(v)} \sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} d_k^{-\alpha}$$

$$- \frac{\sigma_{q-2,n}(v)}{q\sigma_{q,n}(v)} \sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} d_k^{-2\alpha}. \quad (12)$$

Following similar steps for deriving equation (9), we have

$$\sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} d_k^{-2\alpha}$$

$$\equiv \begin{cases} \Theta(\frac{n}{r(n)}) & , for \ 0 \leq \alpha \leq 3/2 \\ \Theta(\frac{n}{r^{2\alpha-2}(n)}) & , for \ 3/2 \leq \alpha \end{cases} \quad (13)$$

If the terms in the negative part of equation (12) are replaced with their equivalents from equations (11) and (13), it is easy to show that for connected networks (minimum transmission range $r(n) = \Theta(\sqrt{\frac{\log n}{n}})$), these negative parts will be of an order less than one.

$$\frac{\sigma_{q-2,n}(v)}{q\sigma_{q,n}(v)} \sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} d_k^{-2\alpha}$$

$$= \begin{cases} \Theta(\frac{1}{nr(n)}) & , for \ 0 \leq \alpha \leq 3/2 \\ \Theta(\frac{1}{r^{2\alpha-2}(n)}) & , for \ 3/2 \leq \alpha \leq 2 \\ \Theta(\frac{1}{r^2(n)}) & , for \ 2 \leq \alpha \end{cases}$$

$$= \begin{cases} \Theta(\frac{1}{\sqrt{n \log n}}) & , for \ 0 \leq \alpha \leq 3/2 \\ \Theta(\frac{1}{n^{2-\alpha} \log^{\alpha-1} n}) & , for \ 3/2 \leq \alpha \leq 2 \quad = o(1) \\ \Theta(\frac{1}{\log n}) & , for \ 2 \leq \alpha \end{cases}$$

Thus, these terms can be ignored compared to the positive part of $E[X]$ and the lower bound for $E[X]$ is the same as its upper bound. Therefore, the obtained bounds on capacity are indeed tight bounds.

However, it is important to compute the traffic carried in each cell and find out if this throughput capacity can be supported for each cell.

It can be verified that the obtained capacity order is achievable for all values of q and α and the flow in no node may become the bottleneck. In order to do that, we just need to compute the total traffic load that a cell is required to accommodate and check if it is not greater than the maximum rate a cell can support.

The traffic load of a node may appear in different situations of being source, relay or destination, the maximum of this value multiplied by the number of nodes in a cell (traffic load of a cell) should not exceed the maximum rate that each cell can support which is $\Theta(1)$.

- Traffic load of a source node

Each source is assumed to transmit data at rate λ, so the maximum load created by each source will be $\Theta(\lambda_{max})$.

- Traffic load of a relay node

We need to compute the maximum number of paths passing through each relay node. To compute this value, we calculate the maximum number of source-destination paths passing through each cell which is ([14], [15])

$$E[X]Pr(Path_i \ intersects \ cell_j) = O(E[X]r^2(n)) < O(nr^2(n)).$$

As they are $\Theta(nr^2(n))$ nodes in each cell, using a routing protocol that randomly and uniformly selects one node in a cell

to forward the packets will result in the maximum traffic load of a relay node to be $\Theta(\lambda_{max}nr^2(n)/nr^2(n)) = \Theta(\lambda_{max})$.

- Traffic load of a destination node

The power-law distribution of the social contacts leads to a non-uniform distribution for destinations. However, we prove that for large n this distribution is asymptotically uniform. The probability that a node v_k is selected as destination is

$$Pr(v_k \ is \ destination)$$

$$= \sum_{v_i} Pr(v_k \ is \ destination|v_i \ is \ source)Pr(v_i \ is \ source),$$

and as the source nodes are uniformly distributed, this probability is equal to

$$\frac{1}{n}\sum_{v_i} Pr(v_k \ is \ destination|v_i \ is \ source).$$

As we have shown in equation (3), the probability inside the summation is equal to $\frac{d_{k_i}^{-\alpha} \sigma_{q-1,n-1,i}^{\overline{v_k}}(v)}{q\sigma_{q,n,i}(v)}$, where index i shows that all the distances in this equation are measured toward source v_i. Replacing this value, which has been shown by $P(\vartheta_t = v_k)$ throughout this paper, with the equivalent values obtained for different values of q, it can be easily seen that the probability that v_k is destination will be $\Theta(\frac{1}{n})$.

If q goes to infinity for sufficiently large n,

$$Pr(v_k \ is \ destination|v_i \ is \ source) = \frac{1}{n}.$$

Thus,

$$Pr(v_k \ is \ destination) = \frac{1}{n}\sum_{v_i}\frac{1}{n} = \frac{1}{n}.$$

If q does not grow with n,

$$Pr(v_k \ is \ destination|v_i \ is \ source) = \frac{n}{n-q+1}\frac{d_{k_i}^{-\alpha}}{\sigma_{1,n,i}}.$$

Thus,

$$Pr(v_k \ is \ destination) = \frac{1}{n-q+1}\sum_{v_i}\frac{d_{k_i}^{-\alpha}}{\sigma_{1,n,i}}.$$

Since $\sigma_{1,n,i}$ has the same order for all i, and by definition is equal to $\sum_{v_i} d_{k_i}^{-\alpha}$, the above equation is equivalent to

$$\frac{1}{n-q+1}\frac{\sum_{v_i} d_{k_i}^{-\alpha}}{\sigma_{1,n}} = \frac{1}{n}.$$

Now since the maximum number of routes passing through each cell is $\Theta(nr^2(n))$, it can be concluded that the maximum number of paths destined to each cell has the same order. Thus, similar to the relay traffic load, each destination traffic load will be $\Theta(\lambda_{max})$.

The total traffic load of a node is $\lambda_{max}(\Theta(1)+\Theta(1)+\Theta(1))$, which results in a total traffic of $\lambda_{max}\Theta(nr^2(n))$ for each cell. If the transmission range is greater than $\Theta(\sqrt{\frac{\log n}{n}})$, the traffic load will be $\lambda_{max}\Theta(\log n)$ which is less than $\Theta(1)$ for all values of α and q.

Therefore, the maximum throughput capacity is upper bounded by the inverse of this traffic [16], i.e., $\lambda_{max} \leq$
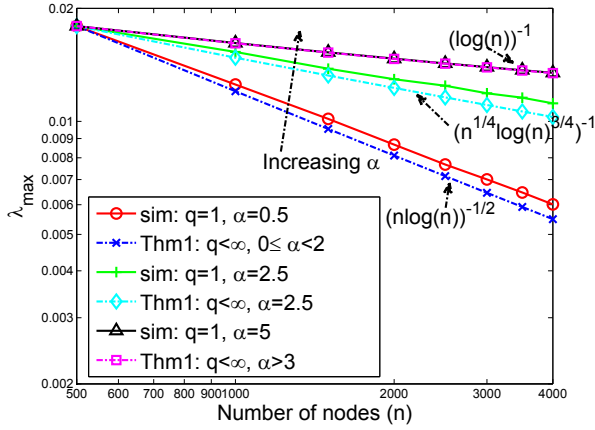
Fig. 2. Throughput capacity vs. the number of nodes for different social network densities α, when each source has $q = 1$ long-range contact based on Theorem 1 results (dash-dot curves), and the simulation results (solid curves).

Fig. 3. Throughput capacity is constant (or changes) with respect to social network density α when each source has finite (or infinite $q = f(n)$) number of social contacts.

$\Theta(\frac{1}{\log n})$, which does not violate the throughput capacity bounds we derived earlier.

V. DISCUSSION AND FUTURE WORK

This paper presents a modeling framework for the capacity of a wireless network in which nodes communicate in the context of social groups and successful transmissions can occur only between nodes within transmission range of each other. The model characterizes a wireless network of n nodes each with a social group size that is a function of the number of nodes n, the probability of a node being a long-range social contact of a source that is inversely proportional to their Euclidean distance with power factor α, and MAI is modeled according to the protocol model.

Figure 2 illustrates the results of Theorem 1 by plotting the network capacity as a function of n for different values of α (shown in dash-dot lines) when the transmission range is the minimum value and the number of long-range contacts is a fixed number, i.e., $q(n) = 1$. The solid lines show similar results obtained through simulations which follow closely the theoretical results. It can be observed that the capacity order decreases exponentially as the number of nodes increases. However, increasing the value of α affects the rate of this capacity decrease. Small values of α correspond to the case in which the social groups are highly distributed in the wireless network, and lead to a rate of order-capacity decrease similar to the results derived by Gupta and Kumar [1], in which no social groups exist.

In contrast, for large values of α, social groups are localized, the paths from sources to destinations involve only $\Theta(1)$ hops, and the maximum throughput capacity is achieved. Furthermore, rate of order-capacity decrease is much smaller than with small values of α.

Figure 3 shows the throughput capacity versus the power law exponent (α) for two values of $q(n)$. In one case, $q(n)$ is a function of n, i.e., $q(n) = f(n)$, where $f(n)$ is an increasing function of n, and in the second case $q(n)$ is a constant value, i.e., $q(n) = 100$. It can be concluded that if the number of
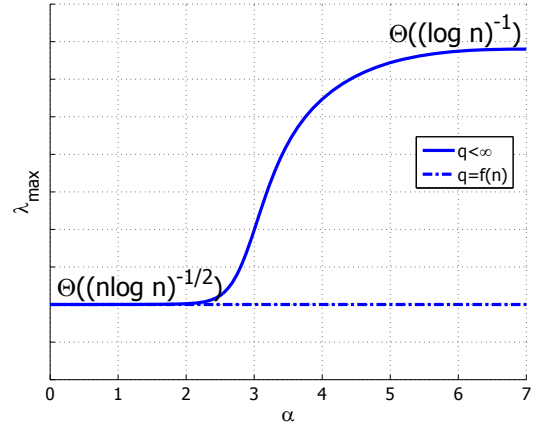
long-range contacts is not a function of the number of nodes, the resulting capacity changes with the parameter α. If α assumes small values ($\alpha \leq 2$), the network behaves as if there were no social groups. For medium values of α ($2 < \alpha < 3$), an exponential growth is observed in the throughput capacity from $\Theta(1/\sqrt{n \log n})$ to $\Theta(1/\log n)$. For large values of α ($\alpha \geq 3$), each source selects its destination along a path involving only $\Theta(1)$ hops w.h.p. and the resulting capacity is the maximum capacity that can be obtained. We also observe that the rate of capacity increase is very slow for $\alpha > 4$.

However, if the number of long-range social contacts $q(n)$ grows proportional to the number of nodes n, the network behaves as if the network had no social groups, independently of the rate of growth for $q(n)$, and each node selects its destination randomly from all the other network nodes. In this case, the throughput capacity does not change with parameter α, and this is true even if $q(n)$ is much smaller than n, i.e., $q(n) = \log \log(n)$, which is a small number even when n is a very large number.

This phenomenon can be described considering the probability of the source-destination distance (d_{st}) order being $\Theta(1)$. When the number of social contacts of each node is a finite number, this probability is very small, even if that finite number is very large. While in the latter case, if the number of social contacts grows with the network, it can be proved that with high probability the source-destination distance is $\Theta(1)$.

$$Pr(d_{st} = \Theta(1)) = Pr(D_1 < d_{st} < D_2)$$

$$= Pr(destination\ is\ inside\ the\ Ring(source, D_1, D_2))$$

$$\equiv \int_{D_1}^{D_2} \frac{n x^{1-\alpha} \sigma_{q-1,n-1}^{\overline{x}}}{q \sigma_{q,n}} dx \qquad (14)$$

where $D_1, D_2 < \infty$ are real finite numbers, and Ring (source, D_1, D_2) is a ring with the inner radius of D_1 and outer radius of D_2 centered on the source. Using the approximations and techniques used in the paper, the following
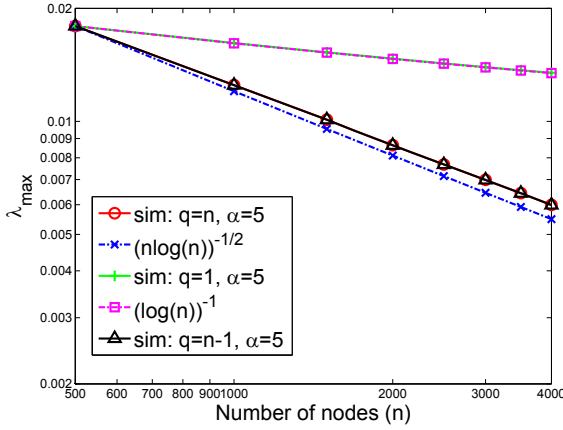
Fig. 4. Simulation results for network size $n = 500 - 4000$ and social group size $q = 1, n-1, n$ for social network density $\alpha = 5$. The dash-dot and solid lines correspond to theoretical and simulation results, respectively.

probability is proved in Appendix.

$$Pr(d_{st} = \Theta(1)) = \begin{cases} 0 & for\ q < \infty, 2 \le \alpha \\ 1 & for\ q \to \infty\ or\ 0 \le \alpha \le 2 \end{cases} \quad (15)$$

Figure 4 illustrates the simulation results for a fixed large social density $\alpha = 5$ for three social group sizes; $q = 1, n-1, n$. It can be seen that the results are very close to the analytical results even when n is not a very large value, i.e., $500 < n < 4000$.

In this work we have made many assumptions to simplify our analytical framework. For example, we have assumed that all the nodes have the same social group size and dispersion, that each source unicasts with a single destination in its social group, that the protocol model is used to model MAI, and that all radios are similar. In addition, we have not addressed the role of content popularity or common interest in content within social groups. We hope to relax these assumptions in our future work, and that this paper will inspire more modeling work on the impact of social groups in wireless networks.

APPENDIX

DETAILED DERIVATION OF EQUATION (9)

$$\sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k\ in\ s_l} d_k^{-\alpha}$$

$$\equiv nr^{2-\alpha}(n) \int_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil+1} u^{2-\alpha} du$$

$$= \frac{nr^{2-\alpha}(n)}{3-\alpha} ((\lceil \frac{2}{C_1 r(n)} \rceil + 1)^{3-\alpha} - (\lceil \frac{1}{C_1} + 1 \rceil)^{3-\alpha})$$

If the transmission range decreases with increasing n, then for sufficiently large n, we have

$$\left(\lceil \frac{2}{C_1 r(n)} \rceil + 1 \right)^{3-\alpha} = \Theta(\frac{1}{r^{3-\alpha}(n)}).$$

If $\alpha < 3$,[4]

$$(\lceil \frac{2}{C_1 r(n)} \rceil + 1)^{3-\alpha} - (\lceil \frac{1}{C_1} + 1 \rceil)^{3-\alpha}$$

$$\equiv \Theta(\frac{1}{r^{3-\alpha}(n)}) - \Theta(1)$$

$$= \Theta(\frac{1}{r^{3-\alpha}(n)})$$

Therefore,

$$\sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k\ in\ s_l} d_k^{-\alpha}$$

$$\equiv \frac{nr^{2-\alpha}(n)}{3-\alpha} \Theta(\frac{1}{r^{3-\alpha}(n)}) \equiv \Theta(\frac{n}{r(n)})$$

For dense social networks in which $\alpha > 3$, we have

$$\sum_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k\ in\ s_l} d_k^{-\alpha}$$

$$\equiv \frac{nr^{2-\alpha}(n)}{\alpha - 3} ((\frac{1}{\lceil \frac{1}{C_1} + 1 \rceil})^{\alpha-3} - (\frac{1}{\lceil \frac{2}{C_1 r(n)} \rceil + 1})^{\alpha-3}),$$

and for large n

$$(\frac{1}{\lceil \frac{1}{C_1} + 1 \rceil})^{\alpha-3} - (\frac{1}{\lceil \frac{2}{C_1 r(n)} \rceil + 1})^{\alpha-3}$$

$$\equiv \Theta(1) - \Theta(r^{\alpha-3}(n)) \equiv \Theta(1)$$

Thus, the above summation is equivalent to

$$\frac{nr^{2-\alpha}(n)}{\alpha - 3} \Theta(1) \equiv \Theta(\frac{n}{r^{\alpha-2}(n)}).$$

DETAILED DERIVATION OF EQUATION (10)

For large n with minimum transmission range, we have

$$\sigma_{1,n} \equiv \int_{r(n)}^{\gamma d_{max}} nu^{1-\alpha} du$$

$$= \frac{n}{2-\alpha} ((\gamma d_{max})^{2-\alpha} - r^{2-\alpha}(n))$$

For $\alpha < 2$ and small $r(n)$, we arrive at $\sigma_{1,n} \equiv \frac{n}{2-\alpha}(\gamma d_{max})^{2-\alpha} \equiv \Theta(n)$.
And for $\alpha > 2$, $\sigma_{1,n}$ is

$$\sigma_{1,n} \equiv \frac{1}{r^{2-\alpha}(n)} - (\frac{1}{\gamma d_{max}})^{\alpha-2})$$

$$\equiv \frac{n}{(\alpha - 2)r^{2-\alpha}(n)} \equiv \Theta(\frac{n}{r^{\alpha-2}(n)})$$

DETAILED DERIVATION OF EQUATION (15)

Based on the computations in section IV and if $\lim_{n \to \infty} q = \infty$, the probability of a node v_k at distance $d_k = x$ away from the source being the destination (equation (3)) will be

$$\frac{x^{-\alpha} \sigma_{q-1,n-1}^{\bar{x}}}{q \sigma_{q,n}} \equiv \frac{1}{n}.$$

---

[4]Note that for $\alpha = 3$, both Cases I and II give the same result.

Thus equation (14) equals to

$$Pr(d_{st} = \Theta(1)) \quad \equiv \quad \int_{D_1}^{D_2} x\,dx = \Theta(1).$$

For networks with finite number of contacts per node, $\lim_{n\to\infty} q < \infty$, the probability that a node at distance x is selected as the destination is $\frac{nx^{-\alpha}}{(n-q+1)\sigma_{1,n}}$ (equation (8)). By replacing this value in equation (14) we have

$$Pr(d_{st} = \Theta(1)) \equiv \frac{n}{n-q+1} \int_{D_1}^{D_2} \frac{nx^{1-\alpha}}{\sigma_{1,n}}\,dx$$

$$\equiv \Theta\left(\frac{n}{\sigma_{1,n}}\right) \equiv \begin{cases} 1 & for\ 0 \le \alpha \le 2 \\ \sqrt{\frac{\log n}{n}}^{\alpha-2} & for\ 2 \le \alpha \end{cases}$$

As can be seen, for both finite and infinite values of q when $0 \le \alpha \le 2$, with high probability the destinations are at distance of $\Theta(1)$ from the sources, while for concentrated social networks($2 \le \alpha$) with finite q, this probability is asymptotically negligible.

REFERENCES

[1] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[2] B. Latané, J. H. Liu, A. Nowak, M. Bonevento, and L. Zheng, "Distance matters: physical space and social impact," *Personality and Social Psychology Bulletin*, vol. 21, no. 8, pp. 795–805, Aug. 1995.

[3] L. Backstrom, E. Sun, and C. Marlow, "Find me if you can: improving geographical prediction with social and spatial proximity," in *Proc. 2010 International Conference on World Wide Web*, pp. 61–70.

[4] S. Milgram, "The small world problem," *Psychology Today*, vol. 2, pp. 60–67, 1967.

[5] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, June 1998.

[6] M. Dietzfelbinger and P. Woelfel, "Tight lower bounds for greedy routing in uniform small world rings," in *Proc. 2009 ACM Symposium on Theory of Computing*, pp. 591–600.

[7] P. Fraigniaud and G. Giakkoupis, "On the searchability of small-world networks with arbitrary underlying structure," in *Proc. 2010 ACM Symposium on Theory of Computing*, pp. 389–398.

[8] J. Kleinberg, "The small-world phenomenon: an algorithm perspective," in *Proc. 2000 ACM Symposium on Theory of Computing*, pp. 163–170.

[9] J. Li, C. Blake, D. S. J. De Couto, H. I. Lee, and R. Morris, "Capacity of ad hoc wireless networks," in *Proc. 2001 International Conference on Mobile Computing and Networking*, pp. 61–69.

[10] B. Azimdoost, H. R. Sadjadpour, and J. J. Garcia Luna-Aceves, "Capacity of composite networks: combining social and wireless ad hoc networks," in *Proc. 2011 IEEE Wireless Communications and Networking Conference*, pp. 464–468.

[11] M. D. Penrose, "The longest edge of the random minimal spanning tree," *The Annals of Applied Probability*, vol. 7, no. 2, pp. 340–361, 1997.

[12] F. Xue and P. R. Kumar, "Scaling laws for ad hoc wireless networks: an information theoretic approach," in *Foundations and Trends in Networking*, 2006, pp. 16–47.

[13] T. P. Mitev, "New inequalities between elementary symmetric polynomials," *J. Inequalities in Pure and Applied Mathematics*, vol. 4, no. 2, 2003.

[14] Z. Wang, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, "Fundamental limits of information dissemination in wireless ad hoc networks—part II: multi-packet reception," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 803–813, Mar. 2011.

[15] P. Gupta and P. R. Kumar, "Internets in the sky: the capacity of three dimensional wireless networks," in *Commun. in Inf. and Syst.*, 2001, pp. 33–50.

[16] B. Liu, D. Towsley, and A. Swami, "Data gathering capacity of large scale multihop wireless networks," in *Proc. 2008 IEEE International Conference on Mobile Ad Hoc and Sensor Systems*, pp. 124–132.

Bita Azimdoost received her B.S. and M.S. degrees from Sharif University of Technology in 1998 and 2000, respectively. Since 2010, she has been pursuing Ph.D. degree with the Department of Electrical Engineering at University of California, Santa Cruz, USA. Her research interests span over wireless ad hoc networks, social networking, and information-centric architectures.

Hamid R. Sadjadpour is a Professor at University of California, Santa Cruz. More detailed information can be found at March 2011 IEEE TRANSACTION ON WIRELESS COMMUNICATIONS.

J. J. Garcia-Luna-Aceves is a Professor at University of California, Santa Cruz. More detailed information can be found at March 2011 IEEE TRANSACTION ON WIRELESS COMMUNICATIONS.