

## **UC Santa Cruz**

### **UC Santa Cruz Previously Published Works**

**Title**

Capacity of Social Networks in Wireless Environments

**Permalink**

<https://escholarship.org/uc/item/2fh9x3q4>

**Author**

Garcia-Luna-Aceves, J.J.

**Publication Date**

2011-07-04

Peer reviewed

# Capacity of Social Networks in Wireless Environments

Bitá Azimdoost

Department of Electrical Engineering  
University of California  
Santa Cruz 95064

Email: bazimdoost@soe.ucsc.edu

Hamid R. Sadjadpour

Department of Electrical Engineering  
University of California  
Santa Cruz 95064

Email: hamid@soe.ucsc.edu

J.J. Garcia-Luna-Aceves

Department of Computer Engineering  
University of California  
Santa Cruz 95064

Email: jj@soe.ucsc.edu

**Abstract**—We study capacity of social networks when nodes communicate in a wireless environment. Such hybrid networks that are combination of wireless communication and social networks are defined as composite networks. Each node has at least one local contact in each of four directions of the network area and  $q(n)$  independent long-range contacts, one of which is selected as the destination. We study the throughput capacity for such networks containing  $n$  nodes assuming the same number of social contacts for all nodes. The nodes communicate using multi-hop communications through relaying the packet to one of their local contacts until the packet reaches the destination. The distance between source and its long-range social contacts follows power law distribution with parameter  $\alpha$ . The order capacity is derived and compared for different values of  $\alpha$  and  $q(n)$ .

## I. INTRODUCTION

The throughput capacity of wireless communication networks has been extensively studied in literature. In most available analysis, source-destination pair selection is based on a uniform distribution. Gupta and Kumar [1] computed the achievable throughput capacity in such networks using the routing algorithm which transports information through the shortest path. However, in most practical networks the source-destination association does not follow the uniform and random distribution. Each source belongs to a social group and it only communicates to the members of its social group. Therefore, the social behavior of nodes has some direct effect on the throughput capacity. This hybrid network that has characteristics of both communication and social networks is defined as composite network.

Social networks have been studied extensively for wired networks. For example, the condition to exhibit small-world phenomenon was first discovered by Kleinberg [3]. His work considered a two dimensional grid network that each node has four local contacts and one long-range contact. This paper was based on earlier work by Watts and Strogatz [6] that divided the edges of the network into local and long-range contacts and assumed that there is always an edge between a node and its social contacts. In Kleinberg paper [3], 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  $v$ . Li et al. [2] studied

an extended network's capacity considering almost the same assumptions.

We studied [8] the interaction between communication and social networks in dense networks considering local contacts and a single long-range contact. The source-destination pair selection followed the power law distribution, i.e.,  $Pr(t$  is long-range contact of  $s) = \frac{d^{-\alpha}(s, t)}{\sum_v d^{-\alpha}(s, v)}$ , where  $s$  and  $t$  are any two nodes,  $d(s, v)$  is the Euclidean distance between  $s$  and any other node  $v$ , and  $\alpha \geq 0$  shows how dense the social network is. The results in [8] are limited in scope because in practical systems, each node usually has more than one long-range social contact. In this paper, we study the more general case in which each source has at least one local contact in each perpendicular direction, and  $q$  long-range contacts, one of which is randomly selected as the destination. We will investigate the effect of the density and size of the social groups on the throughput capacity of the network. To the best of our knowledge, this is the first work in literature which studies the interaction between wireless communication and social networks under such a general condition.

The rest of the paper is organized as follows. In section II, we introduce the notations, and some definitions and theorems that we will use throughout the paper. Section III shows that the original power law distribution introduced by Kleinberg [3] cannot be applied when the number of long-range contacts  $q$  is a function of total number of nodes in the network. This limitation was also mentioned by Kleinberg in his paper. Further, a new power law distribution is introduced that is applicable for all values of  $q$ . The main results of our work on the capacity bounds are presented in section IV and derived in details in section V. Section VI discusses the results. The paper is concluded in section VII.

## II. PRELIMINARIES

The network is a dense network in a unit square area with  $n$  uniformly distributed nodes. We use the protocol model [4] for successful communications. Node  $i$  at position  $X_i$  can successfully transmit to node  $j$  at position  $X_j$  if for any node  $k$  at position  $X_k$ ,  $k \neq i$ , that transmits at the same time as  $i$ , then  $|X_i - X_j| \leq r(n)$  and  $|X_k - X_j| \geq (1 + \Delta)r(n)$ ,

where  $X_i, X_j$  and  $X_k$  are the Cartesian positions in the unit square area for these nodes, and  $\Delta > 0$  is the guard zone factor. A common transmission range  $r(n)$  is considered for all the nodes in the network. To guarantee connectivity in this network [5], the transmission range ( $r(n)$ ) is assumed to be  $r(n) = \Theta(\sqrt{\log n/n})$ .

The TDMA medium access control scheme is shown in fig. 1. The network area is divided into square-lets with side-length  $C_1 r(n)$ , ( $C_1 < \frac{1}{4}$ ), and at any given time the cells separated by  $M$  square-lets distance are the only cells allowed to transmit as shown in gray color in fig. 1 where  $M \geq (2 + \Delta)/C_1$ .

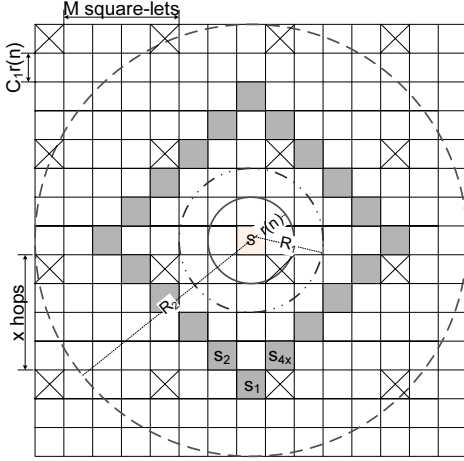


Fig. 1. The solid-line circle shows the transmission range. Dark gray cells ( $s_i$ ) contain the nodes with  $P(X = x)$ .  $R_1$  ( $R_2$ ) are used as the distance of each node in this region instead of their real distances to achieve upper (lower) bounds on  $P(X = x)$ .

The decentralized routing protocol used in this work is very simple. Each node selects one of its local contacts in its four adjacent cells which is the closest one 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 converges.

We use the notation of [7] to denote the elementary symmetric polynomials of the variables  $x = (x_1, \dots, x_n)$  by  $\sigma_{p,n}$ ,  $1 \leq p \leq n$ . In other words,

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

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

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

**Lemma 1-** Let  $x_1, \dots, x_n$  be non-negative real numbers,  $n \geq 2$ . Then for  $1 \leq p \leq n - 1$  we have

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

The proof is described through induction in [7].

In Lemma 2 we will prove that for values of  $p$  that does not grow as fast as  $n$ , this bound is a tight bound.

The notation  $f(n) \equiv g(n)$  is used when  $f(n)$  is in the same order as  $g(n)$ , i.e.,  $f(n)$  and  $g(n)$  have the same asymptotic order ( $f(n) = \Theta(g(n))$ ). Also, the standard notations of  $O$  and  $\Omega$  are used to describe the asymptotic upper and lower bounds.

### III. PROBABILITY FUNCTION OF THE DESTINATION

Kleinberg [3] has studied the wired two dimensional grid with directed edges, in which every node  $s$  has a directed edge to every other node  $v_i$  within lattice distance  $p \geq 1$ , and directed edges to  $q \geq 0$  other nodes using independent random trails. The  $j^{\text{th}}$  directed edge from  $s$  has endpoint  $v_i, i = 1, \dots, n$  with probability proportional to  $d_i^{-\alpha} \hat{=} d^{-\alpha}(s, v_i)$  and normalizing constant  $\sum_{i=1}^n d_i^{-\alpha}$ . Considering the same probability distribution function for long-range contacts, the probability that the long-range contact (LRC) list contains exactly  $q$  independently selected members is

$$\begin{aligned} P(|LRC| = q) &= \sum_{1 \leq i_1 < \dots < i_q \leq n} P(LRC = \{v_{i_1}, \dots, v_{i_q}\}) \\ &= \sum_{1 \leq i_1 < \dots < i_q \leq n} \prod_{j=1}^q P(v_{i_j} \in LRC) \\ &= \sum_{1 \leq i_1 < \dots < i_q \leq n} \frac{d_{i_1}^{-\alpha} \dots d_{i_q}^{-\alpha}}{(\sum_{j=1}^n d_j^{-\alpha})^q}. \end{aligned}$$

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 [3] mentioned that  $q$  is universally constant value and the above derivation proves that the original power law distribution should be modified when  $q$  is a function of  $n$ . In this paper, we assume that each source node has exactly  $q(n)$  long-range contacts selected in independent random trials.

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

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

The normalization factor is obtained using the fact that  $\sum_{1 \leq i_1 < \dots < i_q \leq n} P(LRC = \{v_{i_1}, \dots, v_{i_q}\}) = 1$ .

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

The probability that a particular node  $v_k$  is selected as a long-range contact, i.e. the probability that  $v_k$  is a member of the long-range contact set ( $P(v_k \in LRC)$ ), is given by

$$\begin{aligned} &\sum_{\substack{i_j \neq k \\ 1 \leq i_1 < \dots < i_{q-1} \leq n}} P(LRC = \{v_k, v_{i_1}, \dots, v_{i_{q-1}}\}) \\ &= \frac{\sum_{1 \leq i_1 < \dots < i_{q-1} \leq n, i_j \neq k} d_k^{-\alpha} d_{i_1}^{-\alpha} \dots d_{i_{q-1}}^{-\alpha}}{\sum_{1 \leq i_1 < \dots < i_q \leq n} d_{i_1}^{-\alpha} \dots d_{i_q}^{-\alpha}}. \end{aligned}$$

The above probability function, which shows the probability of node  $v_k$  being in  $LRC$ , is non-decreasing in  $q$ , and also ensures that the described process ends up with a  $q$ -member long-range contact set for each source node.

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

$$P(v_t = v_k) = P(v_t = v_k | v_k \in LRC) \times P(v_k \in LRC) + P(v_t = v_k | v_k \notin LRC) \times P(v_k \notin LRC).$$

Since the destination is only selected from long-range contacts, then  $P(v_k \notin LRC) = 0$ . Further, the selection of destination from long-range contacts has uniform distribution.

$$P(v_t = v_k) = \frac{1}{q} P(v_k \in LRC) = \frac{\sum_{1 \leq i_1 < \dots < i_{q-1} \leq n, i_j \neq k} d_k^{-\alpha} \prod_{j=1}^{q-1} d_{i_j}^{-\alpha}}{q \sum_{1 \leq i_1 < \dots < i_q \leq n} \prod_{j=1}^q d_{i_j}^{-\alpha}}$$

Lets define the notation  $v = (v_1, \dots, v_n)$  for  $(d_1^{-\alpha}, \dots, d_n^{-\alpha})$ , then the above equation can be written as

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

#### IV. MAIN RESULTS

**Theorem 1-** Consider a dense network with nodes associating to social groups communicate in a wireless environment. Each node has at least one local contact in each perpendicular direction, and  $q$  long-range contacts selected independently. Long-range contacts are selected based on power law distribution with parameter  $\alpha$  and one of long-range contacts is the destination for the node. The maximum achievable capacity in this network is

$$\begin{cases} \Theta\left(\frac{1}{\sqrt{n \log n}}\right), & \text{for } q = \Theta(n) \\ \Theta\left(\frac{1}{\sqrt{n \log n}}\right), & \text{for } (q, \frac{q}{n}) \xrightarrow{n \rightarrow \infty} (\infty, 0) \\ \Theta\left(\frac{n-q+1}{n} \frac{1}{\sqrt{n \log n}}\right), & \text{for } q < \infty, 0 \leq \alpha < 2 \\ \Theta\left(\frac{n-q+1}{n^2} \left(\frac{n}{\log n}\right)^{\frac{\alpha-1}{2}}\right), & \text{for } q < \infty, 2 \leq \alpha \leq 3 \\ \Theta\left(\frac{n-q+1}{n} \frac{1}{\log n}\right). & \text{for } q < \infty, 3 < \alpha \end{cases}$$

#### V. THROUGHPUT CAPACITY ANALYSIS

Lets define  $\lambda$  as the data rate for each node and  $X$  as the number of hops traveled by each bit from source to destination. Thus, the total number of concurrent transmissions in such a network would be  $n\lambda E[X]$ , where  $E[X]$  is the average number of hops between any source-destination pair. This value is upper bounded by the total bandwidth  $W$  available divided by the number of non-interfered groups in TDMA scheme as shown in fig. 1 ( $\frac{W}{M^2 C_1^2 r^2(n)}$ ). Therefore, using the minimum transmission range to guarantee connectivity, the maximum data rate in this network is [8]

$$\lambda \leq \lambda_{max} = \Theta\left(\frac{1}{\log n E[X]}\right). \quad (4)$$

As described in [8], the average number of hops can be computed as

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

$P(X = 1)$  is the probability that the packets travel just one hop from source to destination. Thus, it is a positive number less than one, so we can ignore it when deriving the order of expected number of hops.

Since 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  $1 < x < \lceil \frac{1}{C_1} + 1 \rceil$ . And as the maximum number of hops is  $\frac{2}{C_1 r(n)}$ ,  $P(X = x)$  should be calculated for  $x = \lceil \frac{1}{C_1} + 1 \rceil, \dots, \lceil \frac{2}{C_1 r(n)} \rceil$ .

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

The geometric place of the nodes in a distance of  $x$  hops from the source node is a rhombus around it as shown in fig. 1 and explained in [8]. The probability that there exist  $x$  hops between the source and the destination is equal to the probability that the destination is located in one of the cells on the boundaries of this rhombus.

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

Therefore,

$$E[X] \equiv \sum_{x=\lceil \frac{1}{C_1} + 1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \text{ in } s_l} P(v_t = v_k) = \sum_{x=\lceil \frac{1}{C_1} + 1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \text{ in } s_l} \frac{d_k^{-\alpha} \sigma_{q-1, n-1}^k(v)}{q \sigma_{q, n}(v)}. \quad (5)$$

For the rest of the paper, we 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

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

It is easy to show that  $d_k^{-\alpha} \sigma_{n-1, n-1}^k(v) = \sigma_{n, n}(v)$ , therefore

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

There are  $nC_1^2 r^2(n)$  nodes inside each cell  $s_i$ , thus

$$\begin{aligned} E[X] &\equiv \sum_{x=\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} 4x^2 C_1^2 r^2(n) \\ &\equiv r^2(n) \int_{\lceil \frac{1}{C_1}+1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} u^2 du \equiv \frac{1}{r(n)} \equiv \sqrt{\frac{n}{\log n}}. \end{aligned}$$

Now let  $q = \Theta(n)$  but  $q \neq n$ . Let  $Y_i = d_i^{-\alpha}$  be i.i.d. random variables for  $1 \leq i \leq n$  and define the 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  $\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m Z_i = \bar{Z}$  where  $\bar{Z}$  is the expected value of random variable  $Z_i$ .

Thus eq. (3) can be computed as

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

Thus, the value of  $E[X]$  is similar to the case  $q = n$ . Utilizing eq. (4) provides the maximum capacity as

$$\lambda_{max} = \Theta\left(\frac{1}{\sqrt{n \log n}}\right).$$

### B. Case II: $n$ grows much faster than $q$

In most practical social networks such as facebook, the growth rate of the the number of friends (social contacts) does not grow as fast as the network. The expected number of hops between source and destination is derived when  $\lim_{n \rightarrow \infty} \frac{q}{n} = 0$ .

This case includes two different situations,  $\lim_{n \rightarrow \infty} q = \infty$  or  $\lim_{n \rightarrow \infty} q < \infty$ . We will study these cases separately. When  $\lim_{n \rightarrow \infty} q = \infty$ , we can use law of large numbers and similar procedure as before to arrive at the same conclusion as before for both  $E[X]$  and  $\lambda_{max}$ .

Now we consider the case where each node has finite number of contacts ( $\lim_{n \rightarrow \infty} q < \infty$ ). The numerator of  $P(v_t = v_k)$  can be expanded as

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

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

$$P_{lower} \leq P(v_t = v_k) \leq P_{upper}, \quad (6)$$

where  $P_{upper} = \frac{d_k^{-\alpha} \sigma_{q-1, n}(v)}{q \sigma_{q, n}(v)} = P_{lower} + \frac{d_k^{-2\alpha} \sigma_{q-2, n}(v)}{q \sigma_{q, n}(v)}$ .

**Lemma 2-** Let  $X = \{x_1, \dots, x_n\}$  be a set of  $n \geq 2$  non-negative real numbers. Then for finite  $p$ , i.e.  $\lim_{n \rightarrow \infty} p < \infty$ , we have

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

**Proof-** Let random variables  $U_i^p = x_{i_1} \dots x_{i_p}$ ,  $i = 1, \dots, \binom{n}{p}$  where  $1 \leq i_1 < \dots < i_p \leq n$ . Hence due to symmetry, these random variables are identically distributed. Moreover their mean  $\bar{U}_p$  is a function of  $p$ . It can be easily seen that these variables are not independent, as they may have common factors of  $x_{i_j}$ . We partition the set  $X$  into  $p$ -member subsets. Assume that  $T^p$  is the set of all possible such partitionings (each denoted by  $T_i^p$ ) with no common member, i.e.  $T_i^p \cap T_j^p = \phi$ . Thus for finite  $p$ , the number of  $T^p$  members is  $|T^p| \equiv \binom{n}{p} / \binom{n}{p} = \binom{n-1}{p-1}$ .

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

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

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

$$\lim_{n \rightarrow \infty} \sum_{\{x_{i_1} \dots x_{i_p}\} \in T_j^p} x_{i_1} \dots x_{i_p} = \lim_{n \rightarrow \infty} \sum_{\{x_{i_1} \dots x_{i_p}\} \in T_j^p} U_{i, p} = \frac{n}{p} \bar{U}_p$$

Thus,

$$\sigma_{p, n} = \sum_{j=1}^{|T^p|} \frac{n}{p} \bar{U}_p = \binom{n}{p} \bar{U}_p.$$

Similar formulation can be derived for  $\sigma_{p+1, n}$ .

$$\sigma_{p+1, n} = \sum_{j=1}^{|T^{p+1}|} \frac{n}{p+1} \bar{U}_{p+1} = \binom{n}{p+1} \bar{U}_{p+1}$$

Therefore,

$$\frac{\sigma_{1, n} \sigma_{p, n}}{(p+1) \sigma_{p+1, n}} = \frac{\sigma_{1, n} \binom{n}{p} \bar{U}_p}{(p+1) \binom{n}{p+1} \bar{U}_{p+1}}.$$

Note that  $U_i^p$  have identical distribution and  $x_i$  are i.i.d., then the expected value  $\bar{U}_{p+1}$  can be expressed in terms of  $\bar{U}_p$ .

$$\begin{aligned} \bar{U}_{p+1} &= E[U_i^{p+1}] = E[x_{i_1} \dots x_{i_{p+1}}] \\ &= \sum_{x_{i_{p+1}}} E[x_{i_1} \dots x_{i_p} x_{i_{p+1}} | x_{i_{p+1}}] p(x_{i_{p+1}}) \\ &= \bar{U}_p \sum_{x_{i_{p+1}}} x_{i_{p+1}} p(x_{i_{p+1}}) = \bar{U}_p \bar{x} \end{aligned}$$

Further, by utilizing law of large numbers for  $\sigma_{1, n}$  results in  $\sigma_{1, n} \rightarrow n \bar{x}$ . Thus,

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

■

Now returning to the case of finite contacts, we use Lemma 2 (for  $p = q - 1$ ) and inequality (6) to obtain an upper bound for  $E[X]$  in eq. (5).

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

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

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 [8]. 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  $R_2 = A_2 x C_1 r(n)$  ( $A_1, A_2 > 1$ ) for all such nodes to reach upper and lower bounds for this summation (see fig. 1).

$$\sum_{l=1}^{4x} \sum_{v_k \text{ in } s_l} (A_2 x C_1 r(n))^{-\alpha} \leq \sum_{l=1}^{4x} \sum_{v_k \text{ in } s_l} d_k^{-\alpha}$$

$$\leq \sum_{l=1}^{4x} \sum_{v_k \text{ in } s_l} (x C_1 r(n)/A_1)^{-\alpha}$$

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

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

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

(a) is obtained by replacing the sum to integral approximation. After we compute the integral, we arrive at

$$\sum_{\lceil \frac{1}{C_1} + 1 \rceil}^{\lceil \frac{2}{C_1 r(n)} \rceil} x \sum_{l=1}^{4x} \sum_{v_k \text{ in } s_l} d_k^{-\alpha} \equiv \begin{cases} \Theta(n \sqrt{\frac{n}{\log n}}), & 0 \leq \alpha < 3 \\ \Theta(n \sqrt{\frac{n}{\log n}}^{\alpha-2}), & 3 \leq \alpha \end{cases} \quad (9)$$

According to [8],  $\sigma_{1,n}$  can be written as

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

Now we can use these results in inequality (8) and 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.

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

The lower bound capacity follows immediately.

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

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

In order to compute the lower bound for  $E[X]$ , we will study the lower bound of  $P(v_t = v_k)$  in inequality (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 eq. (7).

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

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

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

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

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

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

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

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

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

If we replace the terms in negative part of the above formula with their equivalents from eqs. (9) and (11), it appears that this part will be of the order less than one. Thus, it can be ignored comparing the positive part of this formula and the lower bound for  $E[X]$  will be the same as its upper bound. Therefore, the obtained lower bounds are indeed tight.

$$\lambda_{max} = \begin{cases} \Theta(\frac{n-q+1}{n} \frac{1}{\sqrt{n \log n}}) & , \text{ for } 0 \leq \alpha < 2 \\ \Theta(\frac{n-q+1}{n^2} \sqrt{\frac{n}{\log n}}^{\alpha-1}) & , \text{ for } 2 \leq \alpha \leq 3 \\ \Theta(\frac{n-q+1}{n} \frac{1}{\log n}) & , \text{ for } 3 < \alpha \end{cases}$$

It can be verified that the obtained capacity is achievable for all values of  $q$  and  $\alpha$  and the flow in no node may become the bottleneck. However, due to page constraint, we omit the proof here.

## VI. DISCUSSION

Fig. 2 demonstrates the capacity of the composite network as a function of  $n$  for different values of  $\alpha$  and when the number of long-range contacts is a fixed number, i.e.,  $q = 5$ . It is shown that the capacity order is exponentially decreasing with the increase in the number of nodes. However, the analysis shows that increasing the value of  $\alpha$  will affect the rate of capacity decrease. Small values of  $\alpha$  lead to a rate of decrease in capacity order similar to the results derived by Gupta and Kumar [1]. It is the case where the social groups are highly distributed. In contrast, condensed social groups, i.e. large values of  $\alpha$ , will have the destination within the transmission range with high probability. Thus, the information needs to transport  $\Theta(1)$  hops to reach the destination. Consequently, the maximum throughput capacity is achieved, and the rate of decrease in capacity is much lower than what is seen in the case of small  $\alpha$ .

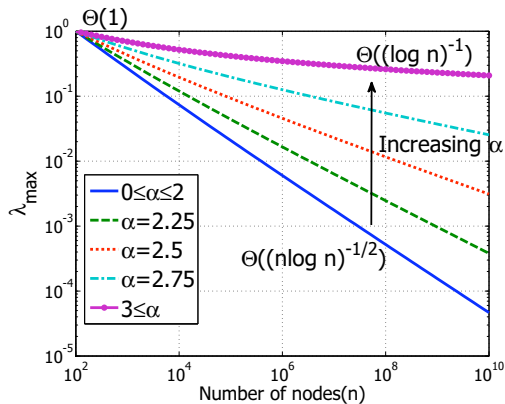


Fig. 2. Throughput capacity vs. the number of nodes when  $q = 5$ .

## VII. CONCLUSION AND FUTURE WORK

This paper presents the first analytical modeling of wireless social networks where each node has local contacts inside its transmission range, and  $q$  long-range contacts. The probability of a node being the long-range contact of the source is inversely proportional to their Euclidean distance with power factor  $\alpha$ . We derived tight bounds on throughput for different values of  $\alpha$  and  $q$ .

Our results show that when the number of long-range contacts does not grow with the network growth, the expected number of hops for each packet to reach the destination is almost independent of the number of long-range contacts. For  $\alpha \leq 2$ , the network behaves like a pure wireless network with no social characteristics. For larger values of  $\alpha$  the rate of capacity increase will be very slow. The maximum capacity is achieved for  $\alpha = 3$  and beyond that, the order throughput capacity does not change.

According to the results, when the number of long-range contacts increases proportional to the number of nodes in the network, the network behavior is similar to a wireless network with no social behavior. Also, the throughput capacity does not

change with parameter  $\alpha$ . This is true when  $q(n) = f(n)$  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 result is rather disappointing since it implies that once each node has a large number of social contacts, then this system does not scale well in wireless environments.

In this work, we assumed that all the nodes have the same number of long-range contacts. As the future work, we can generalize this work when each node has different number of long-range social contacts.

## ACKNOWLEDGMENT

This research was partially sponsored by the U.S. Army Research Laboratory under the Network Science Collaborative Technology Alliance, Agreement Number W911NF-09-0053, by the Army Research Office under agreement number W911NF-05-1-0246, by the National Science Foundation under grant CCF-0729230, and by the Baskin Chair of Computer Engineering. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

## REFERENCES

- [1] P. Gupta, P. R. Kumar, *The capacity of wireless networks*, IEEE Transaction on Information Theory, vol. 46, no. 2, pp. 388-404, 2000.
- [2] J. Li, Ch. Blake, D. S. J. De Couto, H. I. Lee, R. Morris, *Capacity of Ad Hoc wireless networks*, Proceedings of the 7th annual international conference on Mobile computing and networking, p.61-69, July 2001, Rome, Italy.
- [3] J. Kleinberg, *The small-world phenomenon: an algorithm perspective*, Proceedings of the thirty-second annual ACM symposium on Theory of computing, p.163-170, May 21-23, 2000, Portland, Oregon, United States.
- [4] F. Xue, P. R. Kumar, *Scaling laws for ad hoc wireless networks: an information theoretic approach*, Now Publishers Inc., 2006.
- [5] S. Kulkarni, P. Viswanath, *A deterministic approach to throughput scaling in wireless networks*, IEEE Trans. Information Theory, vol. 50, pp. 1041-1049, 2004.
- [6] D. Watts, S. Strogatz, *Collective dynamics of small-world networks*, Nature, pp. 393-440, 1998.
- [7] T. P. Mitev, *New inequalities between elementary symmetric polynomials*, Journal of Inequalities in Pure and Applied Mathematics, volume 4, issue 2, article 48, 2003.
- [8] B. Azimdoost, H. R. Sadjadpour, J.J. Garcia-Luna Aceves, *Capacity of Composite Networks: Combining Social and Wireless Ad Hoc Networks*, IEEE WCNC 2011 conference, Cancun, Mexico, March 28 - 31, 2011.