Capacity of Scale Free Wireless Networks

Bita 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

Abstract—We study the impact of social connectivity on the capacity of wireless networks by considering different values of concentration factor and degree dispersion in scale-free networks. The result shows that a capacity similar to Gupta and Kumar [1] is achieved. Further investigation reveals that traditional transport capacity definition provides misleading conclusions for such network models. We show that nodes with different social status impact the capacity differently. By separating nodes with different social status in frequency and allocating separate bandwidth to them, it is shown that majority of nodes scale in this network. The results imply that in a network with social and communication characteristics, social behavior of the nodes has significant influence on the performance of such networks.

I. INTRODUCTION

Wireless networks and their performance have been widely studied during the past decade. Gupta and Kumar's work [1] showed that the capacity order of wireless communication networks with n nodes, each participating in unicast transmissions, is $\Theta(\frac{1}{\sqrt{n \log n}})$. They assume that each source selects a destination with no priority, so any node in the network can be selected as destination with uniform probability.

On the other hand, an early work by Milgram on the small-world phenomenon [2] has evoked a considerable attention to the modeling of social networks which include a large family of networks. Studies have shown that the Web [3], [4], scientific collaboration on research papers [5], and general social networks [6] have small-world properties. Several models have been proposed and analyzed in this category. Watts and Strogatz [7] divided the edges of a network into local and long-range contacts and assumed that there is always an edge between a node and any of its local or long-range social contacts. Dietzfelbinger et al. [8] studied a ring-based network where each node is connected to its left and right neighbors and possibly to some further nodes, and the long-range contacts may be selected through any distribution.

Liben-Nowell et al. [9] found a strong correlation between friendship and geographic location in social networks by using data from Live Journal, and Backstrom et. al. [10] observed that in practical networks like Facebook the geography and social relationships are inseparable; the nodes that interact with each other are more likely to be geographically close.

Kleinberg [11] proposed a model to explain the small-world phenomenon. His model consists 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 α shows the density of the social network. Li et al. [12] computed upper bounds on the capacity of a wireless network in which source-destination pairs followed a power-law distribution as in Kleinberg's model. Azimdoost et al. [13] also studied the interaction between communication and social networks by considering one single long-range contact per node, selected through a power-law distribution and found out that the maximum capacity, λ_{max}, in dense social networks is much larger than in Gupta and Kumar's result. More recently, the same authors extended their earlier work [14] by considering nodes having any number of long-range social contacts and for all possible number of long-range contacts. Their results show that for highly concentrated networks ($\alpha > 3$), the maximum possible throughput of $\Theta(\frac{1}{\log n})$ in a point-to-point wireless communication can be achieved if each node has limited number of contacts, while for the networks where the social connections of each node increases with the network growth, the maximum capacity of order $\Theta(\frac{1}{\sqrt{n \log n}})$ is achieved which is similar to that of Gupta and Kumar [1].

Earlier studies [13], [14] focused on the homogeneous social networks where all nodes have the same number of social group members. However, study of social networks reveals that the number of social group members for each node is different. More specifically, a small portion of the nodes have a very large social group size while majority of nodes have social contact with a few other nodes in the network. Therefore, different nodes in the network have different social status, i.e., few nodes are highly popular while most of the nodes are less popular in the network. Such networks have been named scale-free networks [15], [16]. The present paper thus addresses the following questions:

1) In a scale-free wireless network, what is the order of the maximum achievable information rate?
2) How can this information rate be improved?
3) Is traditional transport capacity definition [1] appropriate for such networks? Is there any better approach to demonstrate the performance of such networks in terms of throughput capacity?

Theorem 1 is proved in section II and implies that these networks are not scalable and similar throughput capacity

order of Gupta and Kumar [1] is achieved.

Theorem 1. *Consider a social wireless network consisting of n nodes with the following properties.*

- *Each node has $q = 1, 2, .., $ or $n - 1$ social contacts and the number of nodes with q social contacts is inversely proportional to q^γ, where γ is the social degree distribution exponent.*
- *The probability that two nodes in Euclidean distance d away from each other are socially connected is inversely proportional to d^α, where α is the social group concentration factor.*
- *Each source communicates with one of its social contacts randomly with no preference.*

Under these conditions and assuming the minimum wireless transmission range of $r(n) = \Theta(\sqrt{\frac{\log n}{n}})$ to guarantee connectivity ([17]), the throughput capacity will be $\lambda_{max} = \Theta(\frac{1}{\sqrt{n \log n}})$ for sufficiently large n.

Theorem 1 shows that these networks are not scalable. However, Theorem 2 demonstrates that by separating few popular nodes (nodes with large number of social contacts) from the rest of nodes that have few social contacts using different sections of available bandwidth, the majority of nodes can scale.

Theorem 2. *Consider the social network characterized in theorem 1 and assume that social connectivity between nodes is highly concentrated ($\alpha > 2$) with large social degree distribution exponent ($2 < \gamma$). Let's divide the total bandwidth (W) into two distinct parts, $W/2$ each; one part to be used to transfer the information generated from the highly connected source nodes ($G_{>q_0}$) and the other part to be used for communication by the source nodes with small social group size ($G_{\leq q_0}$) where q_0 is a constant value independent of n.*

The maximum data rate for the first group ($G_{>q_0}$) is

$$\lambda_{max} = \Theta(\frac{1}{\sqrt{n \log n}}), \ for \ 2 < \alpha. \quad (1)$$

The maximum data rate for the second group ($G_{\leq q_0}$) is

$$\lambda_{max} = \begin{cases} \Theta(\frac{1}{\sqrt{n^{3-\alpha} \log^{\alpha-1} n}}) & for \ 2 < \alpha < 3 \\ \Theta(\frac{1}{\log n}) & for \ 3 < \alpha \end{cases} \quad (2)$$

Section IV will discuss the results.

II. THROUGHPUT OF SCALE FREE NETWORKS

Studies show that a large number of networks like WWW [3] can be characterized as scale-free networks. According to [18], the number of nodes with q contacts in these networks is inversely proportional to q^γ, where the exponent γ illustrates the clustering property of the network. It has been shown [18] that for each social network, this exponent is a constant number which does not change over time. Thus, the number of social contacts of a node, or its degree, is a random variable (Q) which can take the values $q = 1, .., n - 1$ with the probability distribution

$$Pr(Q = q) = \frac{q^{-\gamma}}{N_{q,\gamma,n}}, \quad (3)$$

where the normalization factor for this probability ($N_{q,\gamma,n}$) is

$$N_{q,\gamma,n} = \sum_{q=1}^{n-1} q^{-\gamma}. \quad (4)$$

Let's assume that each node selects its destination in random from its social contacts, so that the average number of hops (X) passed by the information from source v_i to its destination is

$$E[X|Source = v_i] = \sum_{q=1}^{n-1} Pr(Q = q)E[X|Source = v_i, Q = q].$$

Replacing the value of $Pr(Q = q)$ from (3) and $E[X|Source = v_i, Q = q]$ from [14] in the above equation results in

$$
\begin{aligned}
&E[X|Source = v_i] \\
&= \sum_{1}^{\frac{1}{r(n)}} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} \sum_{q=1}^{n-1} \frac{q^{-\gamma}}{\sum_{b=1}^{n-1} b^{-\gamma}} \frac{d_k^{-\alpha} \sigma_{q-1,n-1}^{\overline{k}}(v)}{q \sigma_{q,n}(v)},
\end{aligned}
\quad (5)
$$

where d_k is the distance between source and any other node v_k in the network. s_l represents a square cell at distance of x hops from the source node v_i. $\sigma_{q,n}(v)$ is the polynomial symmetric function described in [19] and is equal to $\sum_{1 \leq i_1 < .. < i_q \leq n} \prod_{j=1}^{q} d_{i_j}^{-\alpha}$. $\sigma_{q-1,n-1}^{\overline{k}}(v)$ is defined in [14] as $\sum_{1 \leq i_1 < .. < i_{q-1} \leq n, i_h \neq k} \prod_{j=1}^{q-1} d_{i_j}^{-\alpha}$. Expanding the elementary symmetric polynomials, we have

$$\frac{d_k^{-\alpha} \sigma_{q-1,n-1}^{\overline{k}}(v)}{\sigma_{q,n}(v)} = \frac{\sum_{1 \leq i_1 < .. < i_q \leq n, \exists h:i_h=k} \prod_{j=1}^{q} d_{i_j}^{-\alpha}}{\sum_{1 \leq i_1 < .. < i_q \leq n} \prod_{j=1}^{q} d_{i_j}^{-\alpha}}. \quad (6)$$

Since each d_{i_j} is an independent sample of a random variable (distance between the source and any other random node), we can define i.i.d. random variables $Y_{i_j} = d_{i_j}^{-\alpha}$ for $1 \leq i_j \leq n$ and the random variable sequence $Z_{i_j} = \log Y_{i_j}$ for all values of i_j, which will obviously be i.i.d. as well.

$$
\begin{aligned}
\frac{d_k^{-\alpha} \sigma_{q-1,n-1}^{\overline{k}}(v)}{\sigma_{q,n}(v)} &= \frac{\sum_{1 \leq i_1 < .. < i_q \leq n, \exists h:i_h=k} \prod_{j=1}^{q} Y_{i_j}}{\sum_{1 \leq i_1 < .. < i_q \leq n} \prod_{j=1}^{q} Y_{i_j}}, \\
&= \frac{\sum_{1 \leq i_1 < .. < i_q \leq n, \exists h:i_h=k} \exp(\sum_{j=1}^{q} Z_{i_j})}{\sum_{1 \leq i_1 < .. < i_q \leq n} \exp(\sum_{j=1}^{q} Z_{i_j})},
\end{aligned}
$$

For sufficiently large value of q_0 that is independent of n, we can apply the Law of Large Numbers (LLN) for $q > q_0$ and q random variables of type Z_i. Thus for $q > q_0$, for any small $\epsilon > 0$ we can find small $\delta(\epsilon)$ such that $\lim_{Large \ q} \frac{1}{q} \sum_{i=1}^{q} Z_i = \overline{Z} + \epsilon$, with probability $1 - \delta(\epsilon) \rightarrow 1$, where $\overline{Z}$ is the expected value of random variable Z_i. Therefore the order of the above equation equals to[1]

---

[1] We use the notation $\equiv$ to show the order equality.

$$\frac{\sum_{1 \le i_1 < .. < i_q \le n, \exists h: i_h = k} \exp(q(\overline{Z}+\epsilon))}{\sum_{1 \le i_1 < .. < i_q \le n} \exp(q(\overline{Z}+\epsilon))},$$

$$= \frac{\binom{n-1}{q-1}}{\binom{n}{q}} = \frac{q}{n}. \qquad (7)$$

Let's define E_1 as

$$E_1 =$$
$$\sum_{x=1}^{\frac{1}{r(n)}} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} \sum_{q=q_0+1}^{n-1} \frac{q^{-\gamma-1}}{\sum_{b=1}^{n-1} b^{-\gamma}} \frac{d_k^{-\alpha} \sigma_{q-1,n-1}^{\overline{k}}(v)}{\sigma_{q,n}(v)}, \qquad (8)$$

which is equal to

$$E_1 \equiv \sum_{x=1}^{\frac{1}{r(n)}} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} \sum_{q=q_0+1}^{n-1} \frac{q^{-\gamma-1}}{\sum_{b=1}^{n-1} b^{-\gamma}} \frac{q}{n} \qquad (9)$$

The term $\sum_{q=q_0+1}^{n-1} \frac{q^{-\gamma-1}}{\sum_{b=1}^{n-1} b^{-\gamma}} \frac{q}{n}$ is not a function of k or l, so it can be taken out of the summation, and the number of terms of the two summations over k and l is in the order of $x(nr^2(n))$. Thus,

$$E_1 \equiv \frac{r^2(n)}{\sum_{b=1}^{n-1} b^{-\gamma}} \sum_{x=1}^{\frac{1}{r(n)}} x^2 \sum_{q=q_0+1}^{n-1} q^{-\gamma}$$

$$\equiv \frac{1}{r(n) \sum_{b=1}^{n-1} b^{-\gamma}} \sum_{q=q_0+1}^{n-1} q^{-\gamma}. \qquad (10)$$

The last equality is computed by approximating the sum as integral, i.e., $\sum_1^{\frac{1}{r(n)}} x^2 \equiv \frac{1}{r^3(n)}$.

Now define

$$E_2 =$$
$$\sum_{x=1}^{\frac{1}{r(n)}} x \sum_{l=1}^{4x} \sum_{v_k \ in \ s_l} \sum_{q=1}^{q_0} \frac{q^{-\gamma-1}}{\sum_{b=1}^{n-1} b^{-\gamma}} \frac{d_k^{-\alpha} \sigma_{q-1,n-1}^{\overline{k}}(v)}{\sigma_{q,n}(v)}. \qquad (11)$$

Note that $E[X] = E_1 + E_2$. We will compute E_1 and E_2 for different values of α and γ and investigate which one of them will be the dominant factor in computation of $E[X]$.

Lemma 1. E_1 has higher order value than E_2 for all values of α and γ.

Proof: We observe that E_1 is not a function of α.

$$E_1 \equiv \frac{1}{r(n) \sum_{q=1}^{n-1} q^{-\gamma}} \sum_{q=q_0+1}^{n-1} q^{-\gamma}$$

If $0 \le \gamma \le 1$, it can be easily shown that the order of $\sum_{q=q_0+1}^{n-1} q^{-\gamma}$ and $\sum_{q=1}^{n-1} q^{-\gamma}$ are both equal to $\frac{n^{1-\gamma}}{1-\gamma}$. Therefore, $E_1 \equiv \frac{1}{r(n)}$, and as this is the maximum number

of hops that $E[X]$ can have in a unit square area, therefore E_2 does not have any effect on the order of $E[X]$.

If $1 \le \gamma$, $\sum_{q=1}^{n-1} q^{-\gamma} \equiv \sum_{q=q_0+1}^{n-1} q^{-\gamma} \equiv 1$. Therefore $E_1 \equiv \frac{1}{r(n)}$ and again E_1 will be the dominant factor in computation of $E[X]$. $\blacksquare$

Lemma 1 implies that regardless of the density and clustering degree of the social network, each piece of information needs to travel $\Theta(\sqrt{\frac{n}{\log n}})$ hops on average to reach the destination. Finally combining this result and by utilizing the minimum transmission range to assure connectivity in the network, the maximum data rate is equal to $\lambda \le \lambda_{max} = \Theta(\frac{1}{E[X] \log n})$. Then, Theorem 1 is readily proved.

$$\lambda_{max} = \Theta(\frac{W}{E[X] \log n}) \qquad (12)$$

$$= \Theta(\frac{1}{\sqrt{n} \log n}) \qquad (13)$$

W is the available bandwidth in the network.

III. IMPROVING THE CAPACITY OF SCALE FREE NETWORKS

We observe from [14] and computation of E_1 and E_2 that when nodes have large social contact size, they require significant network resources to transport unicast data to destinations. On the other hand, when nodes have small social contact size, they require much less network resources in order to transport packets from sources to destinations. Such significant disparity in capacity behavior among nodes suggests that the conventional definition of transport capacity for wireless communication networks is not appropriate for scale free wireless social networks. In these networks, a more accurate analysis should be based on the fact that nodes with different social status in terms of popularity (i.e., number of social contacts) should be grouped separately. For this reason, we divide the bandwidth W into two equal parts and allow communication for each group of nodes within its allocated bandwidth. Note that as we are talking about the order of the throughput capacity, any bandwidth allocation which does not depend on the number of nodes will not change the order of throughput capacity result. Clearly, in order to preserve the connectivity in the network, we still allow nodes in different social status to relay messages for the other group of nodes. However, this condition requires that each node is equipped with two radios, each one operating in different frequency. Further discussion on the details of this approach is beyond the scope of this paper.

Lemma 2. Let q_0 be a large constant number. For small social degree distribution exponent ($0 < \gamma < 1$), the number of nodes with more than q_0 social contacts ($N_{>q_0}$) is $\Theta(n)$ and the number of nodes with less than q_0 social contacts ($N_{\le q_0}$) is $\Theta(n^\gamma)$. Further, for large social degree distribution exponent ($2 < \gamma$), the ratio of the number of nodes with more than q_0 social contacts to the number of nodes with less than q_0 social contacts ($\frac{N_{>q_0}}{N_{\le q_0}}$) is $\Theta(q_0^{1-\gamma})$ which is a very small number for sufficiently large q_0.

Proof: According to [18], the number of nodes with q connections is proportional to $q^{-\gamma}$ and based on (3), the number of nodes with q connections is on average equal to $n\frac{q^{-\gamma}}{\sum_{q=1}^{n-1} q^{-\gamma}}$. Therefore the number of nodes with more than q_0 social connections ($N_{>q_0}$)is

$$N_{>q_0} = n\frac{\sum_{q=q_0+1}^{n-1} q^{-\gamma}}{\sum_{q=1}^{n-1} q^{-\gamma}}, \tag{14}$$

and the number of nodes with less than $q_0 + 1$ social connections ($N_{\leq q_0}$)is

$$N_{\leq q_0} = n\frac{\sum_{q=1}^{q_0} q^{-\gamma}}{\sum_{q=1}^{n-1} q^{-\gamma}}. \tag{15}$$

These summations can be approximated by integrals.

$$1 + \int_2^n \frac{dq}{q^\gamma} \leq \sum_{q=1}^{n-1} q^{-\gamma} \leq 1 + \int_2^n \frac{dq}{(q-1)^\gamma}$$

$$1 + \int_2^{q_0+1} \frac{dq}{q^\gamma} \leq \sum_{q=1}^{q_0} q^{-\gamma} \leq 1 + \int_2^{q_0+1} \frac{dq}{(q-1)^\gamma}$$

$$\int_{q_0+1}^n \frac{dq}{q^\gamma} \leq \sum_{q=q_0+1}^{n-1} q^{-\gamma} \leq \int_{q_0+1}^n \frac{dq}{(q-1)^\gamma}$$

Therefore the upper and lower bounds for $N_{>q_0}$ and $N_{\leq q_0}$ are

$$n\frac{\frac{1}{1-\gamma}(n^{1-\gamma} - (q_0+1)^{1-\gamma})}{1 + \frac{1}{1-\gamma}((n-1)^{1-\gamma} - 1)} \leq N_{>q_0}$$

$$\leq n\frac{\frac{1}{1-\gamma}((n-1)^{1-\gamma} - q_0^{1-\gamma})}{1 + \frac{1}{1-\gamma}(n^{1-\gamma} - 2^{1-\gamma})}, \tag{16}$$

and

$$n\frac{1 + \frac{1}{1-\gamma}((q_0+1)^{1-\gamma} - 2^{1-\gamma})}{1 + \frac{1}{1-\gamma}((n-1)^{1-\gamma} - 1)} \leq N_{\leq q_0}$$

$$\leq n\frac{1 + \frac{1}{1-\gamma}(q_0^{1-\gamma} - 1)}{1 + \frac{1}{1-\gamma}(n^{1-\gamma} - 2^{1-\gamma})}. \tag{17}$$

Based on these inequalities, it can be easily seen that for small γ (less than 1), the number of nodes with large number of social connections is a tight bound, i.e., $N_{>q_0} = \Theta(n)$. For larger γ, the number of such nodes decreases significantly and will be negligible compared to the number of nodes with very small number of social contacts. When γ is larger than 2, both $N_{>q_0}$ and $N_{\leq q_0}$ are $\Theta(n)$ but their ratio is inversely proportional to $q_0^{\gamma-1}$. Figure 1(a) shows the ratio of the number of nodes with more than q_0 social contacts to the number of nodes with less social connections for $\gamma = 2.3$ and $n = 10^7$, and Figure 1(b) illustrates the same ratio for similar γ and $q_0 = 1000$. It can be seen that the network size (n) does not considerably affect this ratio as long as it is much more than q_0, and the value of q_0 changes the ratio exponentially. ∎

In other words, for large γ, the number of nodes involving in E_1 is much less than the nodes which generate the E_2 part of

the total average number of hops. E_2 is calculated in appendix and it is shown that for large values of γ and α this term is much smaller than E_1. The following Lemma describes the size of $E1$ and E_2 for large values of α and γ.

Lemma 3. In highly concentrated social networks (large α) with large social degree distribution exponent (large γ) , a very small group of nodes ($N_{>q_0}$) use the majority of the resources (due to the large average number of hops traveled by each packet to reach the destination), while a large group of nodes ($N_{\leq q_0}$) use a small portion of the resources.

This Lemma implies that conventional definition of transport capacity may not be appropriate for scale-free networks. In these networks, transportation of a single packet requires different amount of network resources in terms of relaying and average number of hops to reach destination. Based on this observation, it makes sense that we divide the nodes into two categories. One group of nodes are less popular and their social group size is small, i.e., $N_{\leq q_0}$ and the other group of nodes are those nodes that are more popular with higher social status with many social contacts, i.e., $N_{>q_0}$. We divide the available bandwidth W into two equal parts and allow communication for each group inside their own bandwidth. By doing so, there is more fairness in each group in terms of utilizing the network resources for transmission of packets to destinations which will ultimately allow us to better understand the performance of the network.

For example if $q_0 = 100$ and $\gamma = 2.5$, then it is easy to show that 99.9% of nodes can scale while only 0.1% of nodes with larger than 100 social contacts will not scale.

We have computed E_1 before and by utilizing (12), the maximum data rate for sources with $N_{>q_0}$ is given by

$$\lambda_{max_{>q_0}} = \Theta(\frac{W/2}{E_1 \log n}) = \Theta(\frac{1}{\sqrt{n \log n}}). \tag{18}$$

We use the results of appendix and particularly equation (31) for sources in the second category, i.e., $N_{\leq q_0}$, to compute the throughput capacity.

$$\lambda_{max_{\leq q_0}} = \Theta(\frac{W/2}{E_2 \log n}) \tag{19}$$

$$= \begin{cases} \Theta(\frac{1}{\sqrt{n \log n}}) & for\ 0 < \alpha < 2 \\ \Theta(\frac{1}{\sqrt{n^{3-\alpha} \log^{\alpha-1} n}}) & for\ 2 < \alpha < 3 \\ \Theta(\frac{1}{\log n}) & for\ 3 < \alpha \end{cases} \tag{20}$$

These two capacity results prove Theorem 2.

IV. DISCUSSION AND FUTURE WORK

This work utilizes the basic framework introduced in [14] for wireless social networks and incorporates the power-law degree distribution which is an important feature observed in social networks. The model characterizes a wireless network of n nodes each having social contact with a random number of nodes. The probability that a source has q social contacts is proportional to $q^{-\gamma}$. Selection of the social group members is also a random selection and the probability of a node being
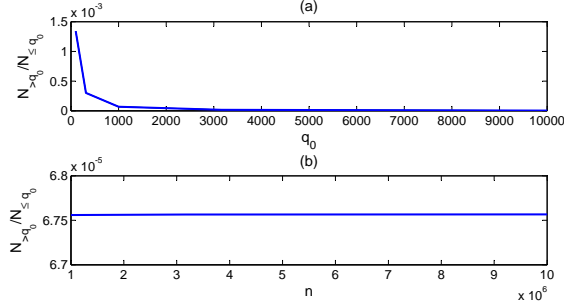
Fig. 1. The ratio between the number of very popular nodes and less popular ones ($\frac{N_{>q_0}}{N_{\leq q_0}}$) for large degree distribution exponent ($\gamma = 2.3$) and (a) diffrent values of social group size threshold q_0 for a fixed number of nodes ($n = 10^7$), (b) different network sizes and a fixed value for q_0 (10^3).

a social contact of a source is inversely proportional to their Euclidean distance with power factor α. The multiple access interference model is considered according to the protocol model.

The order of the throughput capacity in such networks is derived and proved to be $\Theta(\frac{1}{\sqrt{n \log n}})$. Further investigation reveals that nodes with different social status, i.e., different number of social contacts, have different effect on throughput capacity. Therefore, traditional transport capacity concept for wireless networks is not appropriate for these types of networks. However, if we divide the nodes into two groups based on their social status and assign to each group half of the available bandwidth, then nodes with small number of social group members can easily scale. On the other hand, the limiting factor in scaling the capacity is the existence of few nodes with high social status that consume majority of the network resources in terms of relaying requirements. More specifically, it was shown that the nodes that limit the capacity consist of a small portion of the network under the condition that the social groups are geographically highly concentrated ($\alpha > 2$) and the degree distribution exponent is large ($\gamma > 2$). Figures 2(a) and (b) demonstrate data rates for these two groups of nodes, when $\gamma > 2$.

There exist many other features of social groups that we can add in our future work and study the throughput capacity performance of such networks. However, these preliminary results show that the results of Gupta and Kumar and many other papers followed that work are overly pessimistic and social connection among nodes may actually help in scaling wireless networks.

APPENDIX

Here we calculate the upper bound for E_2. In (11), we can write $d_k^{-\alpha}\sigma_{q-1,n-1}^{\overline{k}}(v)$ as
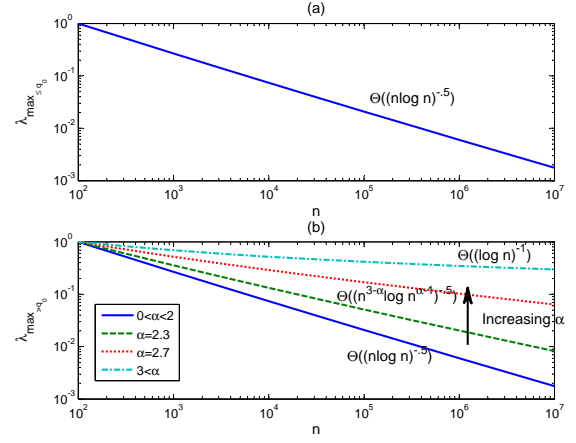

Fig. 2. Maximum achievable data rate order for (a) highly connected source nodes (Group $G_{>q_0}$), (b) nodes with small social group (Group $G_{\leq q_0}$), with large degree distribution exponent ($\gamma > 2$).

$$
d_k^{-\alpha}\sigma_{q-1,n-1}^{\overline{k}}(v),
$$
$$
= d_k^{-\alpha}(\sigma_{q-1,n}(v) - d_k^{-\alpha}\sigma_{q-2,n-1}^{\overline{k}}(v)),
$$
$$
= d_k^{-\alpha}(\sigma_{q-1,n}(v) - d_k^{-\alpha}(\sigma_{q-2,n}(v) - d_k^{-\alpha}\sigma_{q-3,n-1}^{\overline{k}}(v))).
$$
$$
\tag{21}
$$

Since $d_k^{-\alpha}$ and $\sigma_{q-2,n-1}$ are positive values, $d_k^{-\alpha}\sigma_{q-1,n}(v)$ provides an upper bound for $d_k^{-\alpha}\sigma_{q-1,n-1}^{\overline{k}}(v)$.

Lemma 4. Let $\Psi = \{\psi_1,...,\psi_n\}$ be a set of $n \geq 2$ non-negative real numbers. Then for a finite p that is not a function of n, we have

$$
\frac{\sigma_{1,n}(\Psi)\sigma_{p,n}(\Psi)}{(p+1)\sigma_{p+1,n}(\Psi)} = \Theta\left(\frac{n}{n-p}\right). \tag{22}
$$

Lemma 4 was proved in [14] and according to this Lemma, p requires to be very small such that $\frac{n}{p}$ is sufficiently large and we use Law of Large Numbers. Therefore, we select q_0 such that when n goes to infinity and for $q \leq q_0$, we have

$$
\frac{\sigma_{q-1,n}(v)}{\sigma_{q,n}(v)} \equiv \frac{nqd_k^{-\alpha}}{(n-q+1)\sigma_{1,n}(v)}. \tag{23}
$$

By incorporating the upper bound of (21) into (11), we arrive at

$$
E_2 <
$$
$$
\sum_{x=1}^{\frac{1}{r(n)}} x \sum_{l=1}^{4x} \sum_{v_k \; in \; s_l} \sum_{q=1}^{q_0} \frac{q^{-\gamma-1}}{\sum_{b=1}^{n-1} b^{-\gamma}} \frac{nqd_k^{-\alpha}}{(n-q+1)\sigma_{1,n}(v)}.
$$
$$
\tag{24}
$$

It can be easily seen that the order of the distances between the source and all the nodes inside s_l for all $l = 1,..,4x$ equals to $xr(n)$ and the number of nodes with a distance of x equals to $\Theta(xnr^2(n))$, thus

$$E_2 <$$

$$\frac{n^2 r^{2-\alpha}(n)}{\sigma_{1,n}(v) \sum_{b=1}^{n-1} b^{-\gamma}} \sum_{x=1}^{\frac{1}{r(n)}} x^{2-\alpha} \sum_{q=1}^{q_0} \frac{q^{-\gamma}}{n-q+1}. \tag{25}$$

Since q_0 is not growing with n, then $n - q + 1 = \Theta(n)$ for $q \le q_0$. Therefore,

$$E_2 < \frac{n r^{2-\alpha}(n)}{\sigma_{1,n}(v) \sum_{b=1}^{n-1} b^{-\gamma}} \sum_{x=1}^{\frac{1}{r(n)}} x^{2-\alpha} \sum_{q=1}^{q_0} q^{-\gamma}. \tag{26}$$

By approximating the summations in the above equation with integration for sufficiently large networks, these summations are given by

$$\sum_{x=1}^{\frac{1}{r(n)}} x^{2-\alpha} \equiv \begin{cases} \frac{1}{(3-\alpha)r^{3-\alpha}(n)} & for \ 0 \le \alpha \le 3 \\ \frac{1}{\alpha-3} & for \ 3 \le \alpha \end{cases}, \tag{27}$$

and

$$\sigma_{1,n} \equiv \begin{cases} \Theta(n) & for \ 0 \le \alpha \le 2 \\ \Theta(n(\sqrt{\frac{n}{\log n}})^{\alpha-2}) & for \ 2 \le \alpha \end{cases}. \tag{28}$$

Therefore the order of E_2, the average number of hops for sources with low number of social connections is

$$E_2 < \frac{1}{\sum_{b=1}^{n-1} b^{-\gamma}} \sum_{q=1}^{q_0} q^{-\gamma} \begin{cases} \frac{1}{r(n)} & for \ 0 \le \alpha \le 2 \\ \frac{1}{r^{3-\alpha}(n)} & for \ 2 \le \alpha \le 3 \\ 1 & for \ 3 \le \alpha \end{cases}. \tag{29}$$

By utilizing (21) and the results from [14], we arrive at

$$\frac{\sigma_{q-2,n}(v)}{q\sigma_{q,n}(v)} = \Theta\left(\frac{n^2}{\sigma_{1,n}^2(v)} \frac{q-1}{(n-q+1)(n-q+2)}\right). \tag{30}$$

We can use similar calculations to prove that this upper bound is actually a tight bound. Therefore for $\gamma > 1$.

$$E_2 = \begin{cases} \Theta(\frac{1}{r(n)}) & for \ 0 \le \alpha \le 2 \\ \Theta(\frac{1}{r^{3-\alpha}(n)}) & for \ 2 \le \alpha \le 3 \\ \Theta(1) & for \ 3 \le \alpha \end{cases} \tag{31}$$

It is obvious that for $\alpha > 2$, the value of E_2 is much less than E_1. Under this condition, a small number of nodes ($N_{>q_0}$) are using a large portion of the resources and are limiting the total throughput of the network.

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] S. Milgram, "The small world problem," Psychology Today, 2(60), 1967.

[3] R. Albert, H. Jeong, and A. L. Barabasi, "The diameter of the world wide web," Nature, 401:130, 1999.

[4] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, "Graph structure in the web: Experiments and models," Proc. of the 9th International World Wide Web Conference (WWW'00), Amsterdam, May 2000.

[5] M. E. J. Newman, "The structure of scientific collaboration networks," Proc. of the National Academy of Sciences (PNAS), 98:409-415, 2001.

[6] L. A. Adamic, O. Buyukkokten, and E. Adar, "A social network caught in the Web," First Monday, 8(6), 2003.

[7] D. Watts, S. Strogatz, "Collective dynamics of small-world networks," Nature, 393:440-442, 1998.

[8] M. Dietzfelbinger, P. Woelfel, "Tight Lower Bounds for Greedy Routing in Uniform Small World Rings," Proc. ACM STOC 09, pp. 591-600, 2009.

[9] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins, "Geographic routing in social networks," Proc. of the National Academy of Sciences (PNAS), 102(33):11623-11628, 2005.

[10] L. Backstrom, E. Sun, C. Marlow, "Find me if you can: Improving geographical prediction with social and spatial proximity," Proc. WWW'10, Raleigh, NC, 2010.

[11] J. Kleinberg, "The small-world phenomenon: an algorithm perspective," Proc. 32nd Annual ACM Symposium on Theory of Computing, May 21-23, 2000, Portland, Oregon, USA.

[12] J. Li, Ch. Blake, D. S. J. De Couto, H. I. Lee, R. Morris, "Capacity of ad hoc wireless networks," Proc. 7th Annual Int'l Conf. on Mobile Computing and Networking, July 2001, Rome, Italy.

[13] B. Azimdoost, H. R. Sadjadpour, J.J. Garcia-Luna Aceves, "Capacity of composite networks: Combining social and wireless ad hoc networks," Proc. IEEE WCNC 2011, Cancun, Mexico, March 28-31, 2011.

[14] B. Azimdoost, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, " The impact of social groups on the capacity of wireless networks," IEEE International Workshop on Network Science (NSW), 2011.

[15] A. L. Barabasi, R. Albert, "Emergence of scaling in random networks," Science 286, 1999.

[16] R. Albert, A. L. Barabasi, "Statistical mechanics of complex networks," Reviews of Modern Physics, Vol. 74, Jan. 2002.

[17] S. Kulkarn, P. Viswanath, "A deterministic approach to throughput scaling in wireless networks," IEEE Trans. Information Theory, Vol. 50, pp. 1041-1049, 2004.

[18] S. Boccalettia, V. Latorab, Y. Morenod, M. Chavez, D. U. Hwang, "Complex networks: Structure and dynamics," Physics Reports 424-Elsevier, pp. 175-308, 2006.

[19] T. P. Mitev, "New inequalities between elementary symmetric polynomials," Journal of Inequalities in Pure and Applied Mathematics, vol. 4, issue 2, 2003.