APPENDICES to "Capacity Scaling of Wireless Social Networks"

Cheng Wang, Lu Shao, Zhong Li, Lei Yang, Xiang-Yang Li, *Senior Member, IEEE*, and Changjun Jiang

APPENDIX A
USEFUL KNOWN RESULTS

Lemma A.1 (Minimal Spanning Tree, [1]): Let X_i, $1 \leq i < \infty$, denote independent random variables with values in $\mathbb{R}^d$, $d \geq 2$, and let M_n denote the cost of a minimal spanning tree of a complete graph with vertex set $\{X_i\}_{i=1}^n$, where the cost of an edge (X_i, X_j) is given by $\Psi(|X_i - X_j|)$. Here, $|X_i - X_j|$ denotes the Euclidean distance between X_i and X_j and Ψ is a monotone function. For bounded random variables and $0 < \sigma < d$, it holds that as $n \to \infty$, with probability 1, one has $M_n \sim c_1(\sigma, d) \cdot n^{\frac{d-\sigma}{d}} \cdot \int_{\mathbb{R}^d} f(X)^{\frac{d-\sigma}{d}} dX$, provided $\Psi(x) \sim x^\sigma$, where $f(X)$ is the density of the absolutely continuous part of the distribution of the $\{X_i\}$.

Lemma A.2 (Kolmogorov's Strong LLN, [2]): Let $\{X_n\}$ be an i.i.d. sequence of random variables having finite mean: for $\forall n$, $\mathbf{E}[X_n] < \infty$. Then, a strong law of large numbers (LLN) applies to the sample mean: $\bar{X}_n \xrightarrow{a.s.} \mathbf{E}[X_n]$, where $\xrightarrow{a.s.}$ denotes *almost sure convergence*.

APPENDIX B
PROOFS FOR SOME THEOREMS AND LEMMAS

B.1 Proof of Lemma 6

We firstly give a basic lemma for the final proof.

Lemma B.1: For all social-broadcast sessions $\mathcal{S}_k$ ($k = 1, 2, \cdots, n$) under the social model $\mathbb{P}(\delta = 0, \gamma, \beta)$, with high probability, the lower bounds on $\sum_{k=1}^n |\operatorname{EMST}(\mathcal{P}_k)|$ hold as described in Table.B.1.

Proof: Let N_l denote the number of sessions with l destinations. First of all, to simplify the proof, we let

$$N_l = n \cdot \Pr(q_k = l) = n \cdot \left(\sum_{j=1}^{n-1} j^{-\gamma}\right)^{-1} \cdot l^{-\gamma},$$

which has no impact on the analysis in order sense according to laws of larger numbers. Based on all $\mathcal{S}_k$, define two sets $\mathcal{K}^1 := \{k | q_k = \Theta(1)\}$ and $\mathcal{K}^\infty := \{k | q_k = \omega(1)\}$. Then,

$$\sum_{k=1}^n |\operatorname{EMST}(\mathcal{P}_k)| = \underline{\Sigma}^1 + \underline{\Sigma}^\infty, \tag{B.1}$$

where $\underline{\Sigma}^1 = \sum_{k \in \mathcal{K}^1} |\operatorname{EMST}(\mathcal{P}_k)|$, $\underline{\Sigma}^\infty = \sum_{k \in \mathcal{K}^\infty} |\operatorname{EMST}(\mathcal{P}_k)|$.

First, we consider $\underline{\Sigma}^1$. Since for $q_k = \Theta(1)$, it holds that $|\operatorname{EMST}(\mathcal{P}_k)| = \Theta(|X - v_k|)$, then we have $\underline{\Sigma}^1 = \sum_{k \in \mathcal{K}^1} |X - v_k|$. For $k \in \mathcal{K}^1$, define a sequence of random variables $\xi_k^1 := |X - v_k| / \sqrt{n}$ having finite mean: $\mathbf{E}[\xi_k^1] = \mathbf{E}[|X - v_k|] / \sqrt{n}$, where $\mathbf{E}[|X - v_k|]$ is presented in Lemma 2. Then, $\underline{\Sigma}^1 = \Theta(\sqrt{n} \cdot \sum_{k \in \mathcal{K}^1} \xi_k^1)$. Therefore, by Lemma A.2, with probability 1,

$$\sum_{k \in \mathcal{K}^1} \xi_k^1 = \Theta(|\mathcal{K}^1| \cdot \mathbf{E}[|X - v_k| / \sqrt{n}]),$$

where $|\mathcal{K}^1|$ denotes the cardinality of $\mathcal{K}^1$. Thus, we get that with probability 1,

$$\underline{\Sigma}^1 = \Theta(|\mathcal{K}^1| \cdot \mathbf{E}[|X - v_k|]). \tag{B.2}$$

Next, we consider $\underline{\Sigma}^\infty$. For $k \in \mathcal{K}^\infty$, all random variables $|\operatorname{EMST}(\mathcal{P}_k)|$ are *independent*; moreover, from Lemma 3, with probability 1, $|\operatorname{EMST}(\mathcal{P}_k)| = \Omega(L_{\mathcal{P}}(\beta, q_k))$, where $L_{\mathcal{P}}(\beta, q_k)$ is defined in Eq.(10). Thus, with probability 1,

$$\underline{\Sigma}^\infty \geq \sum_{l:(1,n]} n \cdot (\sum_{j=1}^{n-1} j^{-\gamma})^{-1} \cdot l^{-\gamma} \cdot L_{\mathcal{P}}(\beta, l). \tag{B.3}$$

Finally, combining Eqs.(B.1), (B.2) and (B.3), we complete the proof. □

Then, we begin to prove Lemma 6. Since $\sum_{k=1}^n q_k = \Theta(\sum_{l=1}^{n-1} n \cdot \Pr(q_k = l) \cdot l)$, we get $\sum_{k=1}^n q_k = Q(\gamma)$, where

$$Q(\gamma) = \begin{cases} \Theta(n), & \gamma > 2; \\ \Theta(n \log n), & \gamma = 2; \\ \Theta(n^{3-\gamma}), & 1 < \gamma < 2; \\ \Theta(n^2 / \log n), & \gamma = 1; \\ \Theta(n^2), & 0 \leq \gamma < 1. \end{cases} \tag{B.4}$$

Moreover, for all $v_k \in \mathcal{V}$, $\mathbf{E}[|v_{k_i} - p_{k_i}|] = \Theta(\int_0^{\sqrt{n}} x \cdot e^{-\pi \cdot x^2} dx)$, that is, $\mathbf{E}[|v_{k_i} - p_{k_i}|] = \Theta(1)$. Thus, according to Lemma A.2, with high probability,

$$\sum_{k=1}^n \sum_{i=1}^{q_k} |v_{k_i} - p_{k_i}| = \Theta(\sum_{k=1}^n q_k). \tag{B.5}$$

Finally, combining with Lemma B.1, we prove the lemma.

B.2 Proof of Lemma 7

For each routing tree $\mathcal{T}_k$, denote the number of cells in $\mathbb{V}(\sqrt{n}, \sqrt{2})$ as $N(\mathcal{T}_k, \sqrt{n}, \sqrt{2})$. From Lemma 4, $\sum_{k=1}^n N(\mathcal{T}_k, \sqrt{n}, \sqrt{2}) = \Omega\left(\sum_{k=1}^n |\operatorname{EMST}(\mathcal{S}_k)|\right)$. Combining with Lemma 6, we get that

$$\sum_{k=1}^{n_s} N(\mathcal{T}_k, \sqrt{n}, \sqrt{2}) = \Omega(H(\gamma, \beta)),$$

where $H(\gamma, \beta)$ is presented in Table.4. By pigeonhole principle, there is at least one cell that will be used by at least $\Omega(H(\gamma, \beta)/n)$ sessions. Since the total throughput capacity of any cell in $\mathbb{V}(\sqrt{n}, \sqrt{2})$ is of order $O(1)$, [3], it holds that under any strategy the social-broadcast throughput is at most of order $O(n / H(\gamma, \beta))$ due to the congestion in some cells.

TABLE B.1
Lower Bounds on $\sum_{k=1}^{n} |\mathrm{EMST}(\mathcal{P}_k)|$

β \ γ	$\gamma > 3/2$	$\gamma = 3/2$	$1 < \gamma < 3/2$	$\gamma = 1$	$0 \le \gamma < 1$
$\beta > 2$	$\Omega(n)$	$\Omega(n \log n)$	$\Omega(n^{\frac{5}{2}-\gamma})$	$\Omega(n^{3/2}/\log n)$	$\Omega(n^{3/2})$
$\beta = 2$	$\Omega(n \cdot \log n)$	$\Omega(n \cdot (\log n)^2)$	$\Omega(\log n \cdot n^{\frac{5}{2}-\gamma})$	$\Omega(n^{3/2})$	$\Omega(n^{3/2} \cdot \log n)$
$1 < \beta < 2$	$\Omega(n^{2-\frac{\beta}{2}})$	$\Omega(n^{2-\frac{\beta}{2}} \cdot \log n)$	$\Omega(n^{\frac{7}{2}-\gamma-\frac{\beta}{2}})$	$\Omega(n^{(5-\beta)/2}/\log n)$	$\Omega(n^{(5-\beta)/2})$
$\beta = 1$	$\Omega(n^{3/2}/\sqrt{\log n})$	$\Omega(n^{3/2} \cdot \sqrt{\log n})$	$\Omega(n^{3-\gamma}/\sqrt{\log n})$	$\Omega(n^2/(\log n)^{3/2})$	$\Omega(n^2/\sqrt{\log n})$
$0 \le \beta < 1$	$\Omega(n^{3/2})$	$\Omega(n^{3/2} \cdot \log n)$	$\Omega(n^{3-\gamma})$	$\Omega(n^2/\log n)$	$\Omega(n^2)$

B.3 Proof of Lemma 8

Denote an island in $\mathbb{V}(\sqrt{n}, \frac{1}{3}\sqrt{\log n})$ by $\mathcal{I}$. For a link, say $u \to v$, where the receiver v is located in $\mathcal{I}$, its length is $|uv| = \Omega(\sqrt{\log n})$, then the capacity of this link is not larger than $B \log_2(1 + \frac{P \cdot |uv|^{-\alpha}}{N_0}) = O((\log n)^{-\frac{\alpha}{2}})$. Consider the initial transmission load of $\mathcal{I}$, by the *tails of Binomial distribution* and *union bounds*, we have that with uniformly high probability, the initial transmission load of each cell in $\mathbb{V}(\sqrt{n}, \frac{1}{3}\sqrt{\log n})$, denoted by $\mathrm{Load_I}$, is of order

$$\mathrm{Load_I} = \begin{cases} \Omega(\log n), & \gamma > 2; \\ \Omega((\log n)^2), & \gamma = 2; \\ \Omega(\log n \cdot n^{2-\gamma}), & 1 < \gamma < 2; \\ \Omega(n), & \gamma = 1; \\ \Omega(n \cdot \log n), & 0 \le \gamma < 1. \end{cases}$$

In addition, there are at most $\Theta(\log n)$ simultaneous links terminating (or initiating) in $\mathcal{I}$ since it contains $\Theta(\log n)$ nodes inside. By the *pigeonhole principle*, there exists a link whose load is of $\Omega(\mathrm{Load_I}/\log n)$. Then, the lemma follows from $R_{u,v} = O((\log n)^{-\frac{\alpha}{2}})$.

B.4 Proof of Lemma 9

By the proposed construction of EST, we have that

$$\sum\nolimits_{k=1}^{n} \mathrm{EST}(\mathcal{S}_k) = O(\Sigma^{\mathcal{A}} + \Sigma^{vp} + \Sigma^{r}) \qquad \text{(B.6)}$$

where

$$\Sigma^{\mathcal{A}} = \sum\nolimits_{k=1}^{n} |\mathrm{EMST}(\mathcal{A}_k)|,\ \Sigma^{vp} = \sum\nolimits_{k=1}^{n} \sum\nolimits_{i=1}^{q_k} |v_{k_i} - p_{k_i}|$$

and $\Sigma^{r} = \sum_{k=1}^{n} \min_i\{|v_k - v_{k_i}|\}$. Let $\Sigma^{\mathcal{A}} = \overline{\Sigma}^1 + \overline{\Sigma}^\infty$, where

$$\overline{\Sigma}^1 = \sum_{k \in \mathcal{K}^1} |\mathrm{EMST}(\mathcal{A}_k)|,\quad \overline{\Sigma}^\infty = \sum_{k \in \mathcal{K}^\infty} |\mathrm{EMST}(\mathcal{A}_k)|.$$

By similar procedures to Eqs.(B.2) and (B.3), we obtain that

$$\begin{aligned} \Sigma^r &= O(n \cdot \mathbf{E}[|X - v_k|]), && \text{(B.7)} \\ \overline{\Sigma}^1 &= O(|\mathcal{K}^1| \cdot \mathbf{E}[|X - v_k|]), && \text{(B.8)} \\ \overline{\Sigma}^\infty &= \sum\nolimits_{l:(1,n]} (\sum\nolimits_{j=1}^{n-1} \frac{1}{j^\gamma})^{-1} \cdot \frac{n}{l^\gamma} \cdot L_{\mathcal{P}}(\beta, l), && \text{(B.9)} \\ \Sigma^{vp} &= \Theta(\sum\nolimits_{k=1}^{n} q_k). && \text{(B.10)} \end{aligned}$$

Combining Eq.(B.6) and Eqs.(B.7-B.10), we get the lemma.

APPENDIX C
SUPPLEMENTARY CONTENT FOR LOWER BOUNDS

C.1 Routing Backbone System

For the sake of succinctness, we first introduce a notion called *scheme lattice* by modifying the *lattice* in [3].

Definition C.1 (Scheme Lattice, [3]): Divide the deployment region/torus $\mathcal{O}(n) = [0, \sqrt{n}]^2$ into a lattice consisting of square cells of side length b, we call the lattice *scheme lattice* and denote it by $\mathbb{L}(\sqrt{n}, b, \theta)$, where $\theta \in [0, \pi/4]$ is the minimum angle between the sides of the deployment region and produced cells.

Our social-broadcast schemes are based on two types of routing backbones, i.e., *highways* and *parallel arterial roads*.

C.1.1 Highway System

The highways are built based on scheme lattice $\mathbb{L}(\sqrt{n}, \sqrt{c^2}, \pi/4)$, as illustrated in Fig.C.1. Then, there are m^2 cells, where $m = \lceil \sqrt{n}/\sqrt{2}c \rceil^2$. A cell is non-empty (open) with the probability of $p \to 1 - \exp(-c^2)$, as $n \to \infty$, independently from each other. Based on $\mathbb{L}(\sqrt{n}, \sqrt{c^2}, \pi/4)$, draw a horizontal edge across half of the squares, and a vertical edge across the others, to obtain a new lattice as described in Fig.C.1. An edge $\hbar$ in the new lattice is *open* if the cell crossed by $\hbar$ is *open*, and call a path comprised of edges in the new lattice (Fig.C.1) *open* if it contains only open edges. Based on an open path penetrating the deployment region, as illustrated in Fig.C.1, we choose a node from each cell in $\mathbb{L}(\sqrt{n}, \sqrt{c^2}, \pi/4)$ corresponding to the edge of open path, call this node *highway-station*, connect a pair of highway-stations in two adjacent cells, and finally obtain a crossing path, and call it *highway*, as in Fig.C.1.

For a given constant $\kappa > 0$, partition the scheme lattice $\mathbb{L}(\sqrt{n}, \sqrt{c^2}, \pi/4)$ into horizontal (or vertical) rectangle slabs of size $m \times \kappa \log m$ (or $\kappa \log m \times m$), denoted by $\mathcal{R}_i^{\mathrm{H}}$ (or $\mathcal{R}_i^{\mathrm{V}}$), where $m = \frac{\sqrt{n}}{\sqrt{2}c}$. Denote the number of disjoint horizontal (or vertical) highways within $\mathcal{R}_i^{\mathrm{H}}$ (or $\mathcal{R}_i^{\mathrm{V}}$) by N_i^{H} (or N_i^{V}). The next lemma follows.

Lemma C.1: ([4]) For any κ and $p \in (5/6, 1)$ satisfying $2 + \kappa \log(6(1-p)) < 0$, there exists a $\eta = \eta(\kappa, p)$ such that

$$\lim_{m \to \infty} \Pr[N^h \ge \eta \log m] = 1, \lim_{m \to \infty} \Pr[N^v \ge \eta \log m] = 1,$$

where $N^{\mathrm{H}} = \min_i N_i^{\mathrm{H}}$ and $N^{\mathrm{V}} = \min_i N_i^{\mathrm{V}}$.

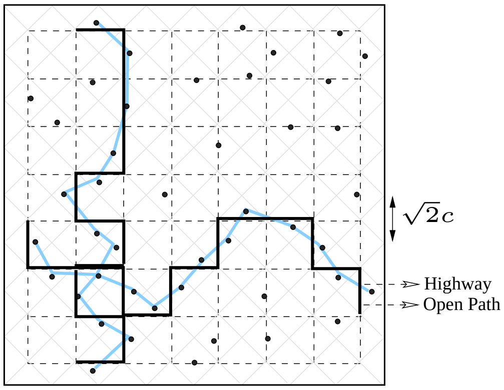


Fig. C.1. Building Horizontal and Vertical Highways.

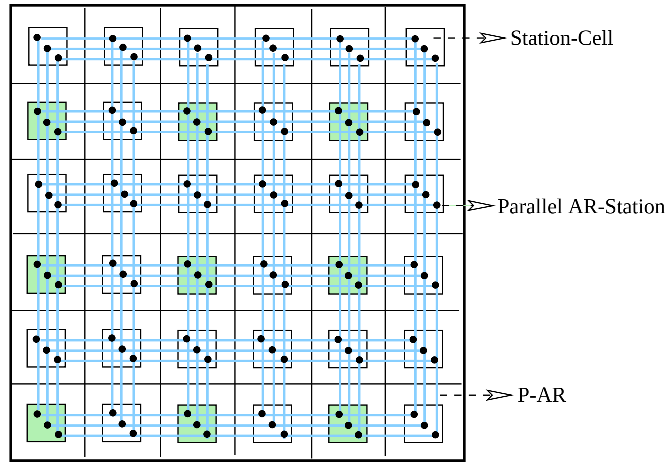


Fig. C.2. Parallel Arterial Roads. The shaded station-cells can be scheduled simultaneously. In any time slot, there are $2\log n$ concurrent links initiated from every activated station-cell.

The highways can be scheduled by a 9-TDMA scheme based on scheme lattice $\mathbb{L}(\sqrt{n}, \sqrt{c^2}, \pi/4)$, [4]. By a similar procedure to Theorem 3 in [4], it can be proven that all highways can sustain w.h.p. the rate of order $\Omega(1)$.

C.1.2 Parallel Arterial Road System (P-AR system)

The P-AR system is constructed based on the scheme lattice $\mathbb{L}(\sqrt{n}, 3\sqrt{\log n}, 0)$. Then, there are $\frac{n}{9\log n}$ cells in $\mathbb{L}(\sqrt{n}, 3\sqrt{\log n}, 0)$, called *AR-cells*. Denote each row (or column) by $\tilde{\mathcal{R}}_i^h$ (or $\tilde{\mathcal{R}}_i^v$), where $i \in [1, \frac{\sqrt{n}}{3\sqrt{\log n}}]$. Then, by Chernoff bounds, for all $\frac{n}{9\log n}$ AR-cells, the number of nodes is *w.h.p.*within $[\frac{9}{2}\log n, 18\log n]$.

In the center of each AR-cell, we set a smaller square of side length $2\sqrt{\log n}$, as illustrated in Fig.C.2, call it *station-cell*. Then, it holds that for all station-cells, there are, *w.h.p.*, at least $2\log n$ nodes.

The *horizontal* arterial roads in $\tilde{\mathcal{R}}_i^h$ is constructed by using the following operations: Firstly, for all $\frac{\sqrt{n}}{3\sqrt{\log n}}$ station-cells in $\tilde{\mathcal{R}}_i^h$, choose $2\log n$ nodes from each station-cell, called *parallel AR-stations*. Secondly, connect those parallel AR-stations in the station-cells contained in the edge-adjacent AR-cells in a one-to-one pattern, as illustrated in Fig.C.2. In a similar way, we can construct the *vertical* arterial roads. We say that two arterial roads are *disjoint* if no station is shared by them. According to the procedure of construction above, there are $2\log n$ disjoint horizontal (or vertical) arterial roads in every row (or column) of $\mathbb{L}(\sqrt{n}, 3\sqrt{\log n}, 0)$.

A 4-TDMA scheme, as depicted in Fig. C.2, is adopted to schedule arterial roads. The main technique called *parallel transmission scheduling* is: Instead of scheduling only one link in each activated station-cell (or cell) in each time slot, we consider scheduling $2\log n$ links initiating from the same station-cell (or cell) together. It can be proven that this modification increases the total throughput for each cell by order of $\Theta(\log n)$, compared with only scheduling one link in each cell. To be specific, each P-AR can sustain a rate of $\lambda^{\mathrm{AR}} = \Theta((\log n)^{-\frac{\alpha}{2}})$.

C.1.3 Access Paths to P-AR System (P-APs)

We call those links, along which the nodes outside drain the packets to P-AR system or the stations in P-AR system deliver the packets to the nodes outside, *parallel access paths* (P-APs).

For every node outside parallel arterial roads, say v, where $v \in \tilde{\mathcal{R}}_j^v$ and $v \in \tilde{\mathcal{R}}_i^h$, it drains the data packets into a parallel AR-station located in the adjacent AR-cell in $\tilde{\mathcal{R}}_j^v$, denoted by $\mathbf{S}_{\mathrm{p}}(v)$, by a single hop called *parallel draining path* (Please see the illustration in Fig.C.3(a)); and receives the packets from the station, located in the adjacent AR-cell in $\tilde{\mathcal{R}}_i^h$, of a specific arterial road by a single hop called *parallel delivering path* (Please see the illustration in Fig.C.3(b)). Specifically, each AR-cell is further divided into $2\log n$ subsquares, called *parallel assignment cell* (PA-cell), of area $\frac{9\log n/\lambda}{2\log n} = \frac{9}{2\lambda}$. Connect all nodes in the same PA-cell with the same P-AR station in the adjacent AR-cell to build the P-APs.

A 2-TDMA scheme is capable to schedule the draining paths (delivering paths, resp.) except those initiating from (terminating to, resp.) nodes in $\tilde{\mathcal{R}}_\delta^h$ ($\tilde{\mathcal{R}}_\delta^v$, resp.), where $\delta = \frac{\sqrt{n}}{3\sqrt{\log n}}$, and use an additional 1-TDMA scheme to schedule other draining paths (delivering paths, resp.). Please see the illustrations in Fig.C.3(a) and Fig.C.3(b). Then, it follows that the rate of each parallel access path, including parallel draining and parallel delivering paths, can also be sustained of $\lambda^{\mathrm{AR}} = \Theta((\log n)^{-\frac{\alpha}{2}})$.

C.2 Social-Broadcast Schemes

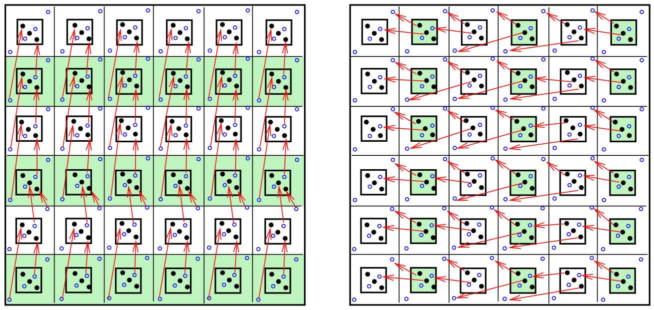
For any social-broadcast session $\mathcal{S}_k$, denote the set of all edges in the Euclidean spanning tree $\mathrm{EST}(\mathcal{S}_k)$ by $\mathcal{E}_k$.

C.2.1 Assignment of Backbones

Now, we determine which backbones, including highway and AR, can be used by a specific communication-pair, *i.e.*, a link $u \to v \in \mathcal{E}_k$.

Assignment of Arterial Roads: Denote the vertical P-AR passing through the parallel AR-station $\mathbf{S}_{\mathrm{p}}(u)$ by $\mathbf{AR}_{\mathrm{p}}^{\mathrm{V}}(u)$; and denote the horizontal P-AR passing through the parallel AR-station $\mathbf{S}_{\mathrm{p}}(v)$ by $\mathbf{AR}_{\mathrm{p}}^{\mathrm{H}}(v)$.

Assignment of Highways: Recall from Lemma C.1 that in each horizontal (or vertical) rectangle slab $\mathcal{R}_i^{\mathrm{H}}$ (or $\mathcal{R}_i^{\mathrm{V}}$) of



(a) Parallel Draining Paths (b) Parallel Delivering Paths

Fig. C.3. (a) The shaded cells can be scheduled simultaneously. All draining paths except those initiating from nodes in $\mathcal{R}_\delta^h$, where $\delta = \frac{\sqrt{n}}{3\sqrt{\log n}}$, can be scheduled once in $2 \times \frac{16 \log n}{2 \log n} = 16$ time slots. In each slot, $2 \log n$ links can be scheduled simultaneously. Here, $16 \log n$ is the maximum number of nodes in each cell, and $2 \log n$ is the number of stations in each cell. In addition, the nodes in $\mathcal{R}_\delta^h$ drain packets to the stations in $\mathcal{R}_{\delta-1}^h$, and those access paths can be scheduled by additional $\frac{16 \log n}{2 \log n} = 8$ time slots. (b) The shaded station-cells can be scheduled simultaneously. All delivering paths except those terminating to nodes in $\mathcal{R}_\delta^v$, can be scheduled once in $2 \times \frac{16 \log n}{2 \log n} = 16$ time slots. In each slot, $2 \log n$ links can be scheduled simultaneously. In addition, the nodes in $\mathcal{R}_\delta^v$ receive packets from the stations in $\mathcal{R}_{\delta-1}^v$, and those access paths can be scheduled by additional $\frac{16 \log n}{2 \log n} = 8$ time slots.

area $\sqrt{n} \times \kappa\sqrt{2c} \cdot \log \frac{\sqrt{n}}{\sqrt{2c}}$ (or $\kappa\sqrt{2c} \cdot \log \frac{\sqrt{n}}{\sqrt{2c}} \times \sqrt{n}$), there are at least $\eta \cdot \log \frac{\sqrt{n}}{\sqrt{2c}}$ horizontal (or vertical) highways. Divide further each horizontal (or vertical) slab into horizontal (or vertical) slice of area $\sqrt{n} \times \frac{\kappa\sqrt{2c}}{\eta}$ (or $\frac{\kappa\sqrt{2c}}{\eta} \times \sqrt{n}$). Choose any $\eta \cdot \log \frac{\sqrt{n}}{\sqrt{2c}}$ highways from each slab, and define an arbitrary bijection from those highways to the slices. For any node u located in a horizontal slice $\mathbf{Slice}_j^{\mathrm{H}}$ (or vertical slice $\mathbf{Slice}_j^{\mathrm{V}}$), the packets initiating from u and terminating to v are assigned to the horizontal highway $\mathbf{H}^{\mathrm{H}}(u)$ and vertical highway $\mathbf{H}^{\mathrm{V}}(v)$ that are mapped to the slices $\mathbf{Slice}_j^{\mathrm{H}}$ and $\mathbf{Slice}_j^{\mathrm{V}}$, respectively.

C.2.2 Social-Broadcast Routing Schemes

For each social-broadcast session $\mathcal{S}_k$ with a Euclidean spanning tree $\mathrm{EST}(\mathcal{S}_k)$, we build two types of social-broadcast routing trees by two corresponding schemes, denoted by $\mathbb{B}_{\mathrm{p\&h}}$ and $\mathbb{B}_{\mathrm{p}}$.

For each edge $u \to v \in \mathcal{E}_k$:

Under $\mathbb{B}_{\mathrm{p\&h}}$, u drains the packets into the parallel AR-station $\mathbf{S}_{\mathrm{p}}(u)$ along a specific P-AP; the packets are transported along the vertical parallel AR $\mathbf{AR}_{\mathrm{p}}^{\mathrm{V}}(u)$ to the assigned horizontal highway $\mathbf{H}^{\mathrm{H}}(u)$; the packets are carried along $\mathbf{H}^{\mathrm{H}}(u)$ and then the vertical highway $\mathbf{H}^{\mathrm{V}}(v)$; the packets are transported along $\mathbf{AR}_{\mathrm{p}}^{\mathrm{H}}(v)$ to the parallel AR-station $\mathbf{S}_{\mathrm{p}}(v)$; and this station delivers the packets to v.

Under $\mathbb{B}_{\mathrm{p}}$, u drains the packets into the assigned parallel AR-station $\mathbf{S}_{\mathrm{p}}(u)$ along a specific P-AP; the packets are transported along the parallel ARs (first parallel vertical AR $\mathbf{AR}_{\mathrm{p}}^{\mathrm{V}}(u)$ then horizontal one $\mathbf{AR}_{\mathrm{p}}^{\mathrm{H}}(v)$) by a Manhattan routing pattern to the parallel AR-station $\mathbf{S}_{\mathrm{p}}(v)$; and this station delivers the packets to v.

When all links in $\mathcal{E}_k$ are checked, merge the same edges (hops) and remove the circles that cannot break the connectivity of $\mathrm{EST}(\mathcal{S}_k)$. Finally, we obtain the corresponding social-broadcast routing trees.

C.3 Proof of Theorem 6

Define an event $\mathrm{Event}_k(\mathbb{B}, v)$ as: Session $\mathcal{S}_k$ is routed through v under the scheme $\mathbb{B}$ in the corresponding phase.

C.3.1 Under Scheme $\mathbb{B}_{\mathrm{p\&h}}$

For a highway-station [5], say v^{H}, we can get that $\Pr(\mathrm{Event}_k(\mathbb{B}_{\mathrm{p\&h}}, v)) = O(\frac{1}{n} \cdot (|\,\mathrm{EST}(\mathcal{S}_k)| + q_k \cdot \log n))$. By laws of larger numbers, the load of highway-station v^{H} in the *highway phase*, denoted by $\mathrm{Load}(v^{\mathrm{H}})$, holds that $\mathrm{Load}(v^{\mathrm{H}}) = O(\frac{1}{n} \cdot (\sum_{k=1}^{n} |\,\mathrm{EST}(\mathcal{S}_k)| + \min\{\log n \cdot \sum_{k=1}^{n} q_k, n^2\}))$. Combining the fact that v^{H} can sustain the rate of order $\Theta(1)$ in highway phase, we get the throughput during highway phase as follows:

$$\underline{\Lambda}^{\mathrm{H}} = \Omega(n/(H(\gamma,\beta) + \min\{\log n \cdot Q(\gamma), n^2\})), \tag{C.1}$$

where $H(\gamma,\beta)$ and $Q(\gamma)$ are defined in Table.4 and Eq.(B.4), respectively. Similarly, we get the throughput in *AR phase*,

$$\underline{\Lambda}^{\mathrm{AR}} = \Omega(\lambda^{\mathrm{AR}} \cdot n/\min\{\sqrt{\log n} \cdot Q(\gamma), n^2\}), \tag{C.2}$$

where $\lambda^{\mathrm{AR}} = \Omega((\log n)^{-\frac{\alpha}{2}})$ is the rate of parallel AR-station in AR phase.

According to the principle of bottleneck, combining Eqs.(C.1) and (C.2), we obtain the achievable throughput $\underline{\Lambda}^{\mathbb{B}_{\mathrm{p\&h}}}$.

C.3.2 Under Scheme $\mathbb{B}_{\mathrm{p}}$

For a parallel AR station, say v^{pa}, it can sustain the rate of order $\Omega((\log n)^{-\frac{\alpha}{2}})$. By a similar analysis method, we get the achievable throughput $\underline{\Lambda}^{\mathbb{B}_{\mathrm{p}}}$ by

$$\underline{\Lambda}^{\mathbb{B}_{\mathrm{p}}} = \Omega((\log n)^{\frac{1-\alpha}{2}} \cdot n/(H(\gamma,\beta) + \sqrt{\log n} \cdot Q(\gamma))). \tag{C.3}$$

Appendix D
A Validation of Population-Based Social Model

In this section, we provide the validation of our model via Google+ users' dataset, [6], [7].

D.1 Google+ Dataset

Google+ was opened to everyone 18 years of age or older on September 20, 2011. It provides each user with an incoming friend list (i.e., "have you in circles"), an outgoing friend list (i.e., "in your circles") and a profile page. In other words, "u is a friend of v" means that "u is in the circles of v".

We adopted the daily snapshots of public Google+ social network structures and user profiles on October 11, 2011. The used dataset consists of 13090 users distributed on 724 locations in USA. Please see the illustration in Fig.D.1.

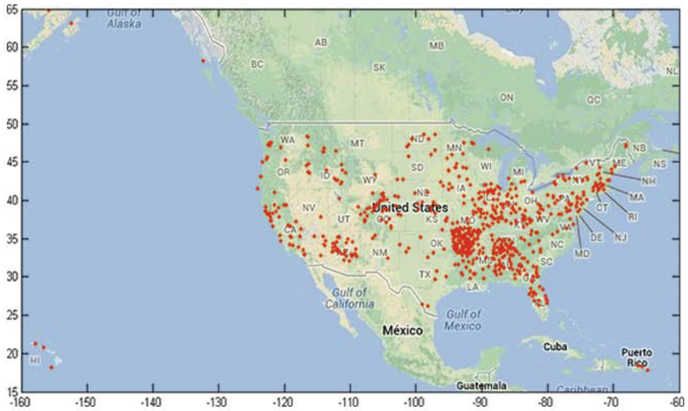


Fig. D.1. Geographical Distribution of Google+ Users.

D.2 Geographical Distribution of Google+ Users

As illustrated in Fig.D.1, users are distributed in an inhomogeneous pattern. Particularly, it shows that there are only several "center points", which matches the assumption of our center-clustering random model (CCRM) that the whole distribution of users generates based on a constant number of center points. Note that for the capacity scaling issue, our CCRM with exact one center point is equivalent to the extended CCRM with a constant number of center points.

Then, the geographical distribution of Google+ users can be regarded as a real-world approximation of our center-clustering random model (CCRM).

D.3 Degree Distribution of Google+ Users

Recall that we assume that the number of friends of a particular node $v_k \in \mathcal{V}$, denoted by q_k, follows a Zipf's distribution [8], i.e., $\Pr(q_k = l) = \left(\sum_{j=1}^{n-1} j^{-\gamma}\right)^{-1} \cdot l^{-\gamma}$. We validate the Zipf's degree distribution of social relations by investigating the *negative linear correlation* between Y and X, where $X = \log K_{out}$ with K_{out} representing an outgoing degree, and $Y = \log N_{out}(K_{out})$.

In dataset of Google+, the relationships between Y and X is described as Fig.D.2. It shows that the relationship tendency is approximated to a line segment with negative slope, which basically matches our proposed model.

D.4 Validation of Population-Based Model

Let $d(u, v)$ denote the distance between user u and user v; let $\mathcal{D}(u, v)$ denote the disk centered at u with a radius $d(u, v)$; and let $N(u, v)$ denote the number of nodes in the disk $\mathcal{D}(u, v)$. Furthermore, we define a variable

$$\mathbf{I}(u, v) = \mathbf{1} \cdot \{v \text{ is a friend of } u\}.$$

We validate the power-law degree distribution of social formation by investigating the *negative linear correlation* between Y and X, where $X = \lg N$ with N denoting a number of nodes, and

$$Y = \lg \left(\frac{\sum_{<u,v> \in \mathcal{E}} \mathbf{1} \cdot \{N(u, v) = N\}}{\sum_{u,v \in \mathcal{V}} \mathbf{1} \cdot \{N(u, v) = N\}} \right),$$

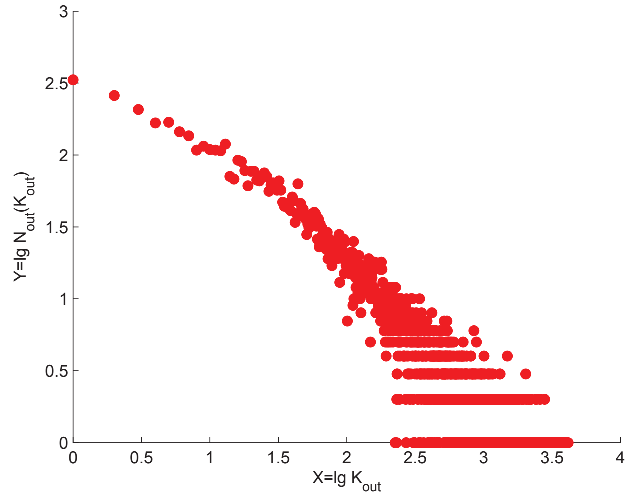


Fig. D.2. Social Degree Distribution of Google+ Users.

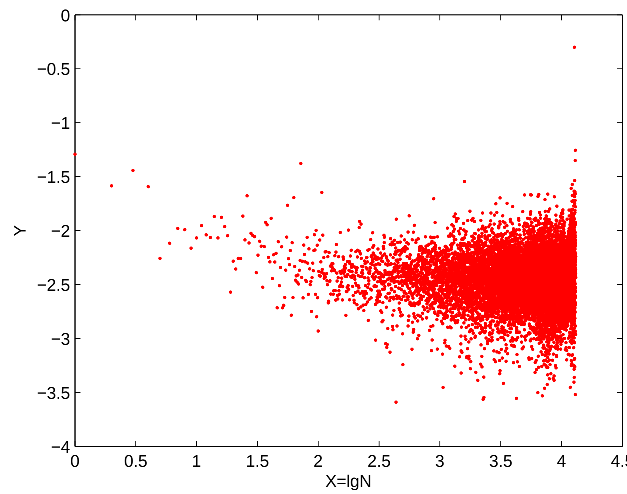


Fig. D.3. Population-based Social Probability Distribution of Google+ Users.

with $\mathcal{V}$ and $\mathcal{E}$ denoting the set of all users and the set of all social links, respectively.

In dataset of Google+, the relationships between Y and X is described as Fig.D.3. It shows that the relationship tendency is approximated very coarsely to a line segment with negative slope. The experimental result also basically validates our proposed model, although it does not perfectly match. The main reason of mismatch lies in the fact that: (1) The locations of users in the dataset are coarse-grained, and are indeed estimated in the experiments. This reduces the accuracy of experiments. (2) Based on this dataset, more than 90% results fall within the part with $X > 3$. The accumulation of experimental errors here leads to a "bloated" tail in the validation graph.

References

[1] J. Steele, "Growth rates of euclidean minimal spanning trees with power weighted edges," *The Annals of Probability*, pp. 1767–1787, 1988.

[2] D. Williams, *Probability with martingales*. Cambridge university press, 1991.

[3] C. Wang, C. Jiang, X.-Y. Li, S. Tang, Y. He, X. Mao, and Y. Liu, "Scaling laws of multicast capacity for power-constrained wireless networks under gaussian channel model," *IEEE Transactions on Computers*, vol. 61, no. 5, pp. 713–725, 2012.

[4] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran, "Closing the gap in the capacity of wireless networks via percolation theory," *IEEE Transactions on Information Theory*, vol. 53, no. 3, pp. 1009–1018, 2007.

[5] C. Wang, C. Jiang, X.-Y. Li, S. Tang, and P. Yang, "General capacity scaling of wireless networks," in *Proc. IEEE INFOCOM 2011*.

[6] N. Z. Gong, W. Xu, and D. Song, "Reciprocity in social networks: Measurements, predictions, and implications," *arXiv preprint arXiv:1302.6309*, 2013.

[7] N. Z. Gong, W. Xu, L. Huang, P. Mittal, E. Stefanov, V. Sekar, and D. Song, "Evolution of social-attribute networks: measurements, modeling, and implications using google+," in *Proc. ACM IMC 2012*.

[8] C. Manning and H. Schütze, *Foundations of statistical natural language processing*. MIT press, 1999.