SCIENCE CHINA
Information Sciences

• **Supplementary File** •

The Capacity of k-connectivity d-dimensional Wireless Networks with Node Failure

Wei LI[1], Junyu LIU[1], Min SHENG[1] & Jiandong LI[1,2]

[1]*State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an,* 710071*, China;*
[2]*Department of Broadband Communication, Peng Cheng Laboratory, Shenzhen,* 518055*, China*

Appendix A Detailed description

Appendix A.1 Details of the Network Protocol

The data packets are transmitted from the source node to the destination node by using multihop strategy. Specifically, the transport process of data packets involves cell partitioning and building/rebuilding source-destination (S-D) lines. To better understand the transport process of data packets in the presence of node failure, we will use the data packet transport process for $d = 2$ as an example to illustrate. For the case of $d > 2$, using network protocol is akin to the $d = 2$, except for the network dimension and the critical transmission radius $r_d(n, k) = \left(\frac{\log n+(k-1)\log\log n-\log\Gamma(k)+\mu_n}{n}\right)^{1/d}$, where $\Gamma(k) = (k-1)!$ [3].

Step 1: To ensure k-connectivity of the network structure, first, we tessellate the unit square with small squares of side length $a = \frac{r_2(n,k)}{2}, k \geqslant 2$, as shown in Figure A1(a). $a = \frac{r_2(n,k)}{2}$ is set to ensure that each cell contains at least k nodes.

Step 2: For any S-D pair in which neither the source node nor the destination node is within any failure regions, the data path can be rebuilt based on whether the original data path is blocked by the failure regions. Specifically, when the data path for a source-destination pair is not blocked, nodes within the cells traversed by the S-D line can relay data packets to the corresponding destination node. This is illustrated in Figure A1(b). If the data path is blocked, however, the original S-D line must be rebuilt. In this case, the data packets are delivered by the nodes in the cells passed by the rebuilt S-D line, as shown in Figure A1(c).

Note that the network protocol model presented here is only intended to serve as a basis for analyzing network capacity. In practical network design, specific details need to be considered. In other words, the network protocol used here captures the inherent features of a multi-hop strategy that are common to all practical data transport protocols about data packet delivery [8–12].

Appendix A.2 Interpretation of Condition 1

In wireless networks, simultaneous transmission of nodes inevitably interferes with another nodes, which is also the main internal factor limiting network capacity. Therefore, a protocol model is established from a geometric perspective, which is widely used in the field of network capacity analysis [1, 4, 5]. For any receive node, specifically, it can successfully receive a data packet only when no other simultaneously receive nodes are within the circles centered at the current receive nodes with radii of $\frac{\triangle}{2}$ hop distance. Otherwise, it cannot successfully receive the data packet due to interference from other nodes. For the former, we refer to it as simultaneous transmission is feasible, and for the latter, simultaneous transmission is infeasible, as shown in Figure A2.

Next, we will rigorously prove the above statements. Without loss of generality, we assume that there are two pairs of nodes, X_i and X_k, simultaneously transmitting in the network, with corresponding receiving nodes X_j and X_l, respectively. Using the triangle inequality and Condition 1, we have

$$\begin{aligned}\rho(X_j, X_l) &\geqslant \rho(X_j, X_k) - \rho(X_l, X_k)\\ &\geqslant (1+\triangle)\rho(X_i, X_j) - \rho(X_l, X_k).\end{aligned} \tag{A1}$$

Similar to (A1)

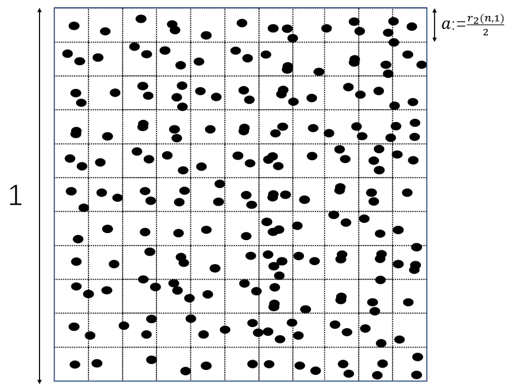
$$\begin{aligned}\rho(X_l, X_j) &\geqslant \rho(X_l, X_i) - \rho(X_i, X_j)\\ &\geqslant (1+\triangle)\rho(X_k, X_l) - \rho(X_i, X_j).\end{aligned} \tag{A2}$$

Combining with (A1) and (A2), we obtain

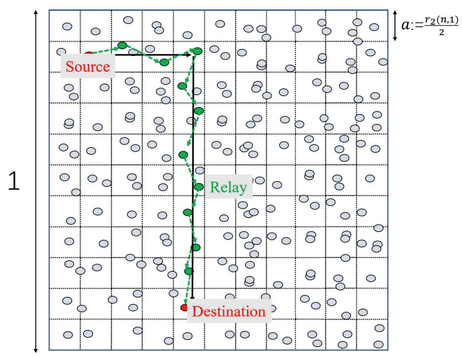
$$\rho(X_l, X_j) \geqslant \frac{\triangle}{2}\left[\rho(X_k, X_l) + \rho(X_i, X_j)\right]. \tag{A3}$$

According to (A3), the simultaneous successful reception of data packets is feasible only when the circles centered at the receive nodes X_j and X_l, with radii of $\frac{\triangle}{2}$ hop distance, do not overlap, as shown in Figure A2(a). This also holds for simultaneously transmission node pairs greater than 2.

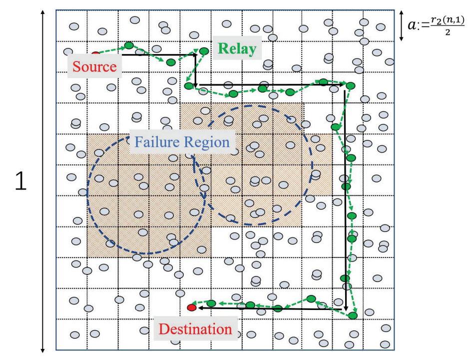
* Corresponding author (email:)



(a) Cell partition.



(b) Original data path of data packets without any node failures.



(c) Circumvented data path of data packets with node failures.

Figure A1 Data path construction strategies under different scenarios when $d = 2$.

Appendix A.3 Interpretation of Definition 1

The network capacity is defined from the data path level, which captures the impact of node failures on the network capacity. Specifically, $\varepsilon_d(n, m, k)$ is utilized to quantify the fraction of unserved S-D pairs, integrating the connectivity parameter k, the intensity of the initial failure nodes m, and the node intensity n.

Two typical cases of the network capacity are given in the following.

- When $m = 0$ and $k > 1$, network capacity can degenerate into the feasible throughput of wireless networks without node failure, which is denoted by $T_d(n, k)$ (bps/Hz). In this case, all S-D pairs can be served.
- When $m = 0$, $d = 2$ and $k = 1$, network capacity is consistent with the definition of feasible throughput $T(n)$ (bps/Hz) in [1]. This definition has been used to analyze the network capacity in the absence of node failures.

Note that the above definitions are only used to capture the relationship between network capacity with the number of nodes, connectivity parameters and number of failure nodes. In actual network design, the fraction of unserved S-D pairs, which is a key indicator to measure the robustness of the network structure, can be used to measure the ability of the wireless network to combat node failures. Moreover, this definition is not dependent on specific network protocols and methods, so conclusions based on this definition are applicable to all multi-hop wireless networks.

Appendix B Proofs

Lemma 1. Let X be a Poisson random variable with mean λ, then

$$\mathbb{P}\{X \geqslant x\} \leqslant \frac{e^{-\lambda}(e\lambda)^x}{x^x},\ x > \lambda. \tag{B1}$$

Proof. The proof has been given in Appendix II of [2].

Appendix B.1 Proof of Proposition 1

Proof. An S-D pair, denoted by i, cannot be served if either its source or destination node is contained in a failure region or a closed region surrounded by failure regions (see Figure B1). There are two situations in which S-D pairs are not served:

1) When at least one node of the S-D pair is located in a failure region, as shown in Figure B1(a);

2) When at least one node of the S-D pair is located within a closed region surrounded by failure regions, as shown in Figure B1(b).

Therefore, to analyze the fraction of unserved S-D pairs, it is crucial to calculate the failure volume, which consists of the failure regions and closed regions. In particular, the failure volume is given by

$$V_d(m) = V_{d,1}(m) + V_{d,2}(m), \tag{B2}$$

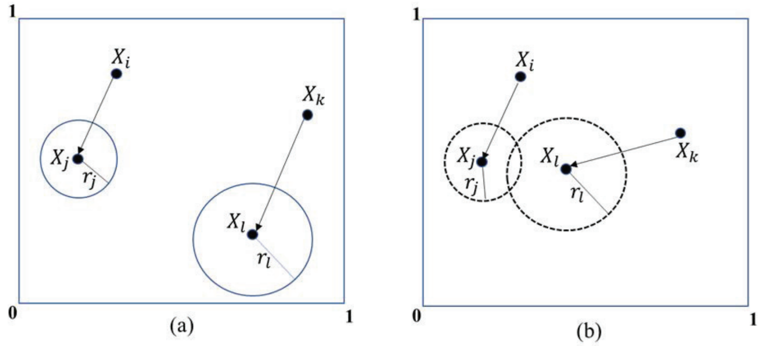


Figure A2 An example of two transmission simultaneously nodes ($d = 2$, $\triangle = 0.05$, $r_j = \frac{\rho(X_i, X_j)\triangle}{2}$ and $r_l = \frac{\rho(X_k, X_l)\triangle}{2}$). (a) X_i and X_k simultaneouly transmission is feasible. (b) X_i and X_k simultaneouly transmission is infeasible.

where $V_{d,1}(m)$ is the volume of failure regions and $V_{d,2}(m)$ is the volume of the closed regions.

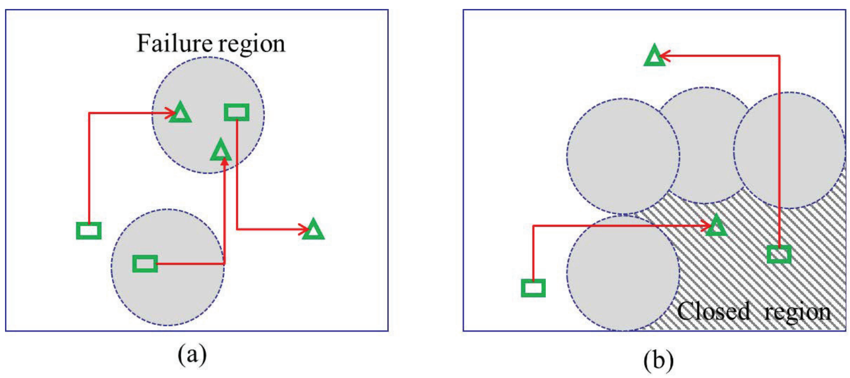


Figure B1 An example of unserved S-D pairs in the plane ($d = 2$). (Rectangle: source node. Triangle: destination)

Next, we analyze the upper bounds of $V_{d,1}(m)$ and $V_{d,2}(m)$.

The upper bound of $V_{d,1}(m)$**.** In the d-dimensional Euclidean space, each failure region contains at most 4^d cells and the volume of each cell is

$$a^d = \left(\frac{r_d(n,k)}{2}\right)^d, \ d \geqslant 2. \tag{B3}$$

The number of initial failure nodes is at most $(1+\sigma)m$ ($\sigma > 0$) since the initial failure nodes are distributed uniformly [7]. Therefore, the upper bound of $V_{d,1}(m)$ is given by

$$V_{d,1}(m) \leqslant (1+\sigma)m4^d a^d = (1+\sigma)m2^d r_d^d(n,k). \tag{B4}$$

The upper bound of $V_{d,2}(m)$**.** Since the initial failure nodes are distributed uniformly, each closed region contains at most 2ζ the failure regions, where $\zeta = \log m + (k-1)\log\log m - \log\Gamma(k)$ (See Lemma 2 in [7]). In the worst case, all the closed regions are surrounded by at most 2ζ the failure regions. In other words, the number of closed regions is at most $\frac{(1+\sigma)m}{2\zeta}$. The largest closed region is formed at the corner of in the hypercube $[0,1]^d$. The size of surface that each failure region contributes to the largest closed region is given by $\frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}(\sqrt{d-1}a)^{d-1}$. By calculating $\frac{1}{2^d}\frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}R_{max,d}^{d-1} = 2\zeta\frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}(\sqrt{d-1}a)^{d-1}$, the radius $R_{max,d}$ of the largest closed region is obtained. Hence, we have

$$\begin{aligned} V_{d,2}(m) &\leqslant \frac{(1+\sigma)m}{2\zeta}\frac{1}{2^d}c_d R_{max,d}^d \\ &= \frac{(1+\sigma)mc_d}{2^{d+1}\zeta}\left(2^{\frac{2}{d-1}}\zeta^{\frac{1}{d-1}}\sqrt{d-1}r_d(n,k)\right)^d, \end{aligned} \tag{B5}$$

where $c_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$ is the volume of the unit sphere in the d-dimensional Euclidean space.

Combining (B4) and (B5), we obtain

$$V_d(m) = V_{d,1}(m) + V_{d,2}(m) \leqslant f(d,m,k)(1+\sigma)mr_d^d(n,k), \tag{B6}$$

where $f(d,m,k) = 2^d + 2^{\frac{-d^2+2d+1}{d-1}}c_d(d-1)^{\frac{d}{2}}\zeta^{\frac{1}{d-1}}$.

Denote n_* as the number of nodes in $V_d(m)$ that follows Poisson random variable with mean $\lambda = nV_d(m)$. Using Lemma 1, we have

$$\mathbb{P}\{n_* \geqslant 2\lambda\}=\mathbb{P}\{n_* \geqslant 2nV_d(m)\} \leqslant \left(\frac{e}{4}\right)^{nV_d(m)} \to 0,\ n \to +\infty. \tag{B7}$$

Thus, it is given by

$$n_* \leqslant 2nV_d(m) \leqslant \varepsilon_{d,0}(n,m,k)(1-\sigma)\frac{n}{2}, \tag{B8}$$

where $(1-\sigma)\frac{n}{2}$ is the lower bound of the total number of S-D pairs $N(n)$ w.h.p (See. Lemma 1a in [6]) and

$$\varepsilon_{d,0}(n,m,k) = 2\frac{1+\sigma}{1-\sigma}mf(d,m,k)r_d^d(n,k) = O\left(\frac{(k\log n)^{\frac{d}{d-1}}}{n^{1-\beta^{-1}}}\right).$$

Appendix B.2 Proof of Proposition 2

Proof. The line segment i of each the S-D line is extended in the i-th dimension, which is called the extend line-i with length 1. Assuming that all the cells through which the extend line-i passes should deliver the corresponding data packets of the line segment i, the order of network capacity cannot be affected [5]. Next, the maximum number of data paths served by the regular cell and loaded cell is obtained by using the above method, as shown in Figure B2.

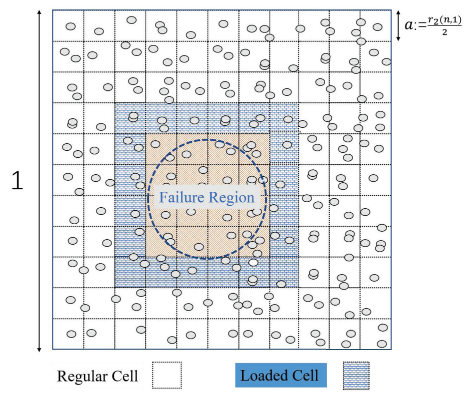


Figure B2 Regular cells and loaded cells.

Let $n_{i,0}$ and $n_{i,1}$ denote the numbers of the non-rebuild extend line-i and rebuild extend line-i carried by a cell in the i-th dimesion, respectively.

Regular cell. For a regular cell, it only carries the data packets generated by the non-rebuiled lines. The extend line-i of all the nodes located in the volume $1 \times a^{d-1}$ should be handled by the regular cell. Hence, the number of the non-rebuild extend line-i is a Poisson random variable with mean na^{d-1}. Using Lemma 1, we obtain

$$\mathbb{P}\left\{n_{i,0} \geqslant 2na^{d-1}\right\} \leqslant \left(\frac{e}{4}\right)^{na^{d-1}},\ i = 1, \cdots, d. \tag{B9}$$

Using (B9), we have

$$\mathbb{P}\left\{\sum_{i=1}^{d} n_{i,0} \geqslant 2nda^{d-1}\right\} \leqslant \sum_{i=1}^{d}\mathbb{P}\left\{n_{i,0} \geqslant 2na^{d-1}\right\} \leqslant d\left(\frac{e}{4}\right)^{na^{d-1}} \to 0,\ n \to +\infty. \tag{B10}$$

Hence, each regular cell can carry data paths at most

$$\nabla = \sum_{i=1}^{d} n_{i,0} = 2nda^{d-1}. \tag{B11}$$

Loaded cell. For a loaded cell, it also needs to carry the rebuild extended lines. Since each of the closed regions contains 2ζ failure regions, the length L of a closed region projection on the i-th dimension is at most $2\zeta \times 4a$. Similarly to (B9), we have

$$\mathbb{P}\left\{n_{i,1} \geqslant 16\zeta na^{d-1}\right\} \leqslant \left(\frac{e}{4}\right)^{8\zeta a^{d-1}},\ i = 1, \cdots, d. \tag{B12}$$

Accordingly, we obtain

$$\mathbb{P}\left\{\sum_{i=1}^{d} n_{i,1} \geqslant 16nd\zeta a^{d-1}\right\} \leqslant \sum_{i=1}^{d}\mathbb{P}\left\{n_{i,1} \geqslant 16n\zeta a^{d-1}\right\} \leqslant d\left(\frac{e}{4}\right)^{8\zeta a^{d-1}} \to 0,\ n \to +\infty. \tag{B13}$$

From the above analysis, each loaded cell can carry data paths at most

$$\sum_{i=1}^{d}(n_{i,0}+n_{i,1}) = (8\zeta+1)\nabla. \tag{B14}$$

Appendix B.3 Proof of Theory 1

Proof. **Case I ($\beta > 2$).** The number of extended data paths carried by a loaded cells is denoted as

$$n_0 = \sum_{i=1}^{d} n_i, \tag{B15}$$

where n_i is the number of extended data paths of the i-th dimesion.

In this case, most of the cells are the regular cells. This means that the distance between any two failure regions, which is projected on the i-th dimesion, is greater than or equal to $2a$. The extended data paths in the i-th dimesion, which pass through the loaded cells, are generated by all the nodes located in the volume of $1 \times (1+\sigma)m(6a)^{d-1}$ because each failure region at most contains 4^d cells. Note that $(1+\sigma)m$ is the upper bound of the initial failure nodes and $\sigma > 0$ is arbitrarily small real number. Accordingly, the upper bound of n_i is a Poisson random variable with mean $\lambda_i = (1+\sigma)mn(6a)^{d-1}$, $i = 1, \cdots, d$. Using Lemma 1, we have

$$\mathbb{P}\left\{n_i \geqslant 2(1+\sigma)mn(6a)^{d-1}\right\} \leqslant \left(\frac{e}{4}\right)^{(1+\sigma)mn(6a)^{d-1}}, \; i = 1, \cdots, d. \tag{B16}$$

Using (B16), we obtained

$$\begin{aligned} &\mathbb{P}\left\{n_0 = \sum_{i=1}^{d} n_i \geqslant 2d(1+\sigma)mn(6a)^{d-1}\right\} \\ &\leqslant \sum_{i=1}^{d} \mathbb{P}\left\{n_i \geqslant 2(1+\sigma)mn(6a)^{d-1}\right\} \\ &\leqslant d\left(\frac{e}{4}\right)^{(1+\sigma)mn(6a)^{d-1}} \to 0, \; n \to +\infty. \end{aligned} \tag{B17}$$

Combining (B16) and (B17), the upper bound of n_0 is given by

$$n_0 \leqslant 2d(1+\sigma)mn(6a)^{d-1} \leqslant \varepsilon_{d,1}(n,m,k)\frac{1-\sigma}{2}n, \tag{B18}$$

where

$$\begin{aligned} \varepsilon_{d,1}(n,m,k) &= \frac{1+\sigma}{1-\sigma}6^{d-1}2^{3-d}dm\left[r_d(n,k)\right]^{d-1} \\ &\leqslant \frac{1+\sigma}{1-\sigma}6^{d-1}2^{3-d}dm\left(\frac{k\log n}{n}\right)^{\frac{d-1}{d}} \\ &= O\left(\frac{(k\log n)^{\frac{d-1}{d}}}{n^{\frac{d-1}{d}-\beta^{-1}}}\right). \end{aligned}$$

and $(1-\sigma)\frac{n}{2}$ is the lower bound of the number of S-D pairs $N(n)$, which is given by Lemma 1a in [6]. The fraction of data paths served by the loaded cells is almost zero since $\lim_{n\to+\infty} \varepsilon_{d,1}(n,m,k) = 0$.

Then, we define the sum of the S-D pairs through the loaded cells and the unserved S-D pairs as the fraction of unserved S-D pairs $\varepsilon_d(n,m,k)$. Based on Proposition 1, $\varepsilon_d(n,m,k)$ is given by

$$\begin{aligned} \varepsilon_d(n,m,k) &\leqslant \varepsilon_{d,0}(n,m,k) + \varepsilon_{d,1}(n,m,k) \\ &= O\left(\frac{(k\log n)^{\frac{d}{d-1}}}{n^{1-\beta^{-1}}}\right) + O\left(\frac{(k\log n)^{\frac{d-1}{d}}}{n^{\frac{d-1}{d}-\beta^{-1}}}\right) \\ &= O\left(\frac{(k\log n)^{\frac{d-1}{d}}}{n^{\frac{d-1}{d}-\beta^{-1}}}\right). \end{aligned} \tag{B19}$$

Next, the feasible throughput $T_d(n,m,k)$, which is determined by the total number of data paths through the regular cell (See. Proposition 2) is given. It is ensured that the cell out of in the failure region obtains one time slot to transport data packets in every τ time slots. Hence, $T_d(n,m,k)$ can be given by

$$T_d(n,m,k) \geqslant \frac{1}{\tau\nabla} \geqslant \frac{1}{2\tau nda^{d-1}} \geqslant \frac{1}{c_0(k\log n)^{\frac{d-1}{d}} n^{\frac{1}{d}}}, \tag{B20}$$

where $c_0 = 2\tau d2^{1-d}$ and $a = \frac{r_d(n,k)}{2}$.

Combining Definition 1, (B19) and (B20), network capacity is given by

$$\begin{aligned} S_d(n,m,k) &= (1-\varepsilon_d(n,m,k))\frac{(1-\sigma)n}{2}T_d(n,m,k) \\ &\geqslant c_1(1-\varepsilon_{d,1}(n,m,k))\left(\frac{n}{k\log n}\right)^{\frac{d-1}{d}}, \end{aligned} \tag{B21}$$

where $c_1 = \frac{1-\sigma}{2c_0}$ and $\varepsilon_{d,1}(n,m,k) = O\left(\frac{(k\log n)^{\frac{d-1}{d}}}{n^{\frac{d-1}{d}-\beta^{-1}}}\right)$.

Case II ($1 < \beta \leqslant 2$). Compared with **Case I**, the ratio of loaded cells becomes larger. This means that most of the data paths can pass the loaded cells w.h.p. Therefore, by Proposition 1, the fraction of unserved S-D pairs $\varepsilon_d(n,m,k)$ is upper bounded by

$$\varepsilon_d(n,m,k) \leqslant \varepsilon_{d,0}(n,m,k) = O\left(\frac{(k\log n)^{\frac{d}{d-1}}}{n^{1-\beta^{-1}}}\right). \tag{B22}$$

In this case, the feasible throughput $T_d(n,m,k)$ is determined by the number of S-D pairs passing through the loaded cells. From Proposition 2, $T_d(n,m,k)$ is given by

$$T_d(n,m,k) \geqslant \frac{1}{(8\zeta+1)\nabla\tau} \geqslant c_2 \frac{1}{k\log n}\left(\frac{n}{k\log n}\right)^{\frac{d-1}{d}}, \tag{B23}$$

where $c_2 = \frac{1}{\tau}$.

Similarly to (B21), combining (B22) and (B23), network capacity is given by

$$S_d(n,m,k) \geqslant c_3(1-\varepsilon_{d,0}(n,m,k))\frac{1}{k\log n}\left(\frac{n}{k\log n}\right)^{\frac{d-1}{d}}, \tag{B24}$$

where $c_3 = \frac{1-\sigma}{2\tau}$.

References

1 P. Gupta and P. Kumar. The capacity of wireless networks. IEEE Trans. Inf. Theory, 2000, 46: 388-404
2 M. Franceschetti, O. Dousse, D. N. C. Tse and P. Thiran. Closing the gap in the capacity of wireless networks via percolation theory. IEEE Trans. Inf. Theory, 2007, 53: 1009-1018
3 M. D. Penrose. On k-connectivity for a geometric random graph. Random Structures and Algorithms, 1999, 15: 145-164
4 N. Lu and X. S. Shen. Scaling laws for throughput capacity and delay in wireless networks - a survey. IEEE Commun. Surveys Tuts., 2014, 16: 642-657
5 A. El Gamal, J. Mammen, B. Prabhakar, et al. Optimal throughput-delay scaling in wireless networks - Part I: the fluid model. IEEE Trans. Inf. Theory, 2006, 52: 2568-2592
6 S. Jeon, N. Devroye, M. Vu, S. Chung, et al. Cognitive networks achieve throughput scaling of a homogeneous network. IEEE Trans. Inf. Theory, 2011, 57: 5103-5115
7 K. -H. Cho, S. -H. Lee and V. Y. F. Tan. Throughput scaling of covert communication over wireless Adhoc metworks. IEEE Trans. Inf. Theory, 2020, 66: 7684-7701
8 M. Sheng, X. Li, X. Wang, and C. Xu. Topology control with successive interference cancellation in cognitive radio networks. IEEE Trans. Commun., 2017, 65: 37-48
9 Y. Shi, H. Sun, M. Sheng, et al. Constructing a robust topology for reliable communications in multi-channel cognitive radio Ad Hoc networks. IEEE Commun. Mag, 2018, 56: 172-179
10 A. Hawbani et al. Zone probabilistic routing for wireless sensor networks. IEEE Trans. Mob. Comput., 2019, 18: 728-741
11 L. Dai. A theoretical framework for random access: stability regions and transmission control. IEEE/ACM Trans. Netw., 2022, 30: 2173-2200
12 S. Rezaei, M. Gharib and A. Movaghar. Throughput analysis of IEEE 802.11 multi-hop wireless networks with routing consideration: a general framework. IEEE Trans. Commun., 2018, 66: 5430-5443