

# The Effect of Information on Scheduling Performance in Multi-hop Wireless Networks

Jun Hong and Victor O. K. Li

Department of Electrical and Electronic Engineering  
The University of Hong Kong, Pokfulam Road, Hong Kong, China  
Email: {hongjun,vli}@eee.hku.hk

**Abstract**—Previous research has estimated the capacity of wireless networks by assuming that each node in the network can obtain precise network information. However, in reality, available network information is mostly imprecise and incomplete. In this paper, we study the relationship between the information obtained by each node and the capacity of the wireless network. We also consider the communication overhead of collecting network information, and analyze the tradeoff between the network capacity improvement and the network information collection overhead. The analysis in this paper can be a valuable tool on determining information collection parameters in wireless networks.

## I. INTRODUCTION

Wireless networks have been a hot research topic and the capacity of wireless networks has drawn significant attention. Gupta and Kumar determine the capacity of wireless networks under some idealized assumptions [1]. Other researchers analyze the impact of interference on the multi-hop wireless network performance [2], [3], [4]. Garetto et al. use a Markovian model to estimate the scheduling effects on the throughput of CSMA channels statistically [5]. Kolar and Abu-Ghazaleh evaluate the scheduling interactions among a given set of active links and analyze the scheduling effects on network capacity [6]. We observe that the wireless network capacity depends not only on the scheduling algorithm, but also on the network information. However, previous research on the capacity of wireless networks only provides an upper bound on the achievable capacity under ideal situation. Network information, which has a large impact on the capacity of wireless networks, is not taken into account.

Network information is an essential and important factor in wireless network scheduling. Obviously, if every node gets more network information, the scheduling will be more efficient and the network capacity can be improved. However, due to the limitations of wireless networks, collecting and disseminating such information is not an easy task, and is consumed valuable bandwidth resource. In this paper, we develop a model to analyze the relationship between the information obtained by each node and the capacity of the wireless network. Some approximations are made in the analysis and these analytic results are compared with simulation results, and we show that the estimation error on the wireless network capacity is small. We also consider the communication overhead of collecting network information, and analyze its impact on the network performance. To the best of our knowledge, this is

the first paper considering the effects of network information on wireless network capacity.

The rest of the paper is organized as follows. Section II describes the network model used throughout the paper and a conflict graph coloring model is introduced. Section III analyzes the relationship between available network information and network capacity, and the overhead of obtaining network information is also presented. The experimental evaluation is presented in Section IV. Section V concludes the paper.

## II. MODEL AND DEFINITIONS

### A. Network Model

Suppose that  $N$  nodes are located in a region of area  $S$   $m^2$ . The nodes of the networks are distributed as a two-dimensional Poisson point process with density  $\lambda$  [7], i.e.

$$\Pr(i \text{ nodes in } C) = (\lambda C)^i \exp(-\lambda C) / i!, \quad i = 0, 1, \dots$$

where  $C$  is the size of the area of interest.

We assume that all the transmissions employ the same power and communication parameter. The communication range is  $R_C$ , i.e. each node can transmit with a maximum radius  $R_C$ , and the circle with a radius of  $R_C$  is called the node's communication area, denoted by  $S_C$ , where  $S_C = \pi R_C^2$ . The interference is  $R_I$ , i.e. a transmitter may interfere with the receivers which are within a range of  $R_I$ .  $n_i$  denotes the  $i$ -th node, and  $d_{ij}$  denotes the distance between nodes  $n_i$  and  $n_j$ . Node  $n_i$  and node  $n_j$  are said to be each other's neighbor if  $d_{ij} \leq R_C$ . It is assumed that all nodes always have packets waiting to send (heavy traffic condition), and each node chooses one of its neighbors (if it has any) randomly to send a packet.

Time is divided into time slots of the same length, and grouped into frames. A frame consists of  $L$  time slots. In each frame, every node is assigned one of the slots, and starts transmission of its packet at the beginning of the slot assigned. If a conflict occurs<sup>1</sup>, the transmission will be assigned a new time slot in the next frame. Note that the frame size, which depends on the current traffic, is adjustable.

We use the protocol interference model [1] to define the conditions for a successful wireless transmission.

<sup>1</sup>The assignments will be made to minimize the conflicts, but due to incomplete and imprecise network information, conflicts may still occur.

**Protocol Interference Model:** the transmission from node  $n_i$  to  $n_j$  is successful if 1)  $d_{ij} \leq R_C$ , and 2) any node  $n_k$ , such that  $d_{kj} \leq R_I$ , is not transmitting.

The information collection range is  $R_D$ , which means that each node can obtain the information of the nodes within a range of  $R_D$ . Since each node can communicate with its neighbors directly, we assume that  $R_D \geq R_C$ , i.e. every node can obtain the information of its neighbors. According to the protocol interference model, nodes out of the interference range can not directly affect the transmission, so the information collection range  $R_D$  can be less than  $R_I$ . Hence, we only consider the situation that  $R_C \leq R_D \leq R_I$ .

A coloring algorithm based on the collected information is used to obtain an optimal scheduling of the transmissions, and  $L$  is the number of required time slots in each frame by applying this optimal coloring algorithm. The details of the coloring algorithm will be introduced in Section II-B.

### B. Conflict Graph Coloring Model

A wireless network is represented as a bi-directional physical graph  $G^p$ . The vertices of  $G^p$  correspond to the wireless nodes and the edges correspond to the wireless links between the nodes. However, wireless interference, which is a key issue impacting network performance, cannot be modeled accurately by  $G^p$ . The effects of interferences can be modeled as a conflict graph [2], [8], [9], whose vertices correspond to the links in graph  $G^p$ , and an edge between two vertices indicates that the corresponding links cannot be active simultaneously, i.e. there is an edge between vertex  $l_{ij}$  and  $l_{mn}$  if 1)  $d_{in} \leq R_I$ , or 2)  $d_{mj} \leq R_I$ .

A vertex coloring of a graph is an assignment of colors to the vertices such that no two adjacent vertices are assigned the same color. The graph coloring problem is then to find a vertex coloring for a graph using the minimum number of colors possible. We can see that the principle of graph coloring is the same as the scheduling rule of conflict graph, and the transmission scheduling in a wireless network is equivalent to the coloring of a conflict graph [10], [11].

An example of a physical graph is shown in Figure 1, where the distance between the adjacent nodes is just  $R_C$ . Since the ratio of  $R_I$  and  $R_C$  is typically between 2 and 3 [12], here  $R_I$  is set to  $2.5 \times R_C$ . The corresponding conflict graph of Figure 1 is shown in Figure 2(a), which is colored using three colors. So, the number of successful packets per time slot is  $4/3$ . However, the true conflict graph in Figure 2(a) is obtained based on the complete network information. If each node can get only part of the network information, the conflict graph may be incorrectly estimated, which will lead to inappropriate scheduling and collisions. Suppose  $R_D = R_C$ , then Figure 2(b) shows the corresponding conflict graph. Since the edge between  $l_{12}$  and  $l_{43}$ , as well as the edge between  $l_{23}$  and  $l_{53}$ , are not detected in Figure 2(b), a packet collision happens, and the number of successful packets per time slot is 0. It is obvious incomplete information results in degraded performance.



Fig. 1. A Chain-topology Wireless Network

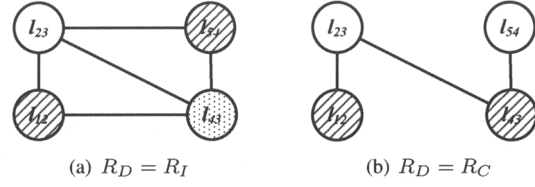


Fig. 2. Conflict Graph

## III. ANALYSIS

### A. Throughput Analysis

In this section, we will analyze how the wireless network capacity increases with the information collection range.

As shown in Figure 3, a node  $n_i$  can communicate with the nodes within the communication range  $R_C$  directly, and it can get the information of the nodes within a range of  $R_D$ . If  $n_i$  receives a packet, the transmission may be affected by the transmitting nodes within the interference range  $R_I$ . The optimal scheduling algorithm can guarantee that the transmission is assigned to a time slot which is different from that occupied by the transmitters in the information collection area  $S_D$ , so no conflicts will happen among  $n_i$  and  $V_D$ , where  $V_D$  is a set of the nodes in  $S_D$ . For  $n_i$  as a receiver, the transmission may fail only if at least one node in the shaded region  $A$  is transmitting using the same time slot, where  $A$  is the region for which information is not collected by node  $n_i$ .  $A$  also refers to the size of the region, and  $A = \pi(R_I^2 - R_D^2)$ .

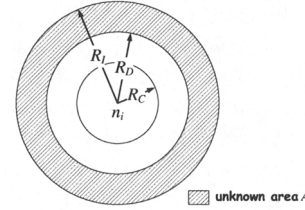


Fig. 3. Relationship between  $R_C$ ,  $R_D$  and  $R_I$

For node  $n_i$ , define

$$x_i = \begin{cases} 1 & \text{if } n_i \text{ sends a packet successfully,} \\ 0 & \text{otherwise.} \end{cases}$$

The network capacity  $G$  is defined as the number of successful transmissions per time slot. Hence,

$$E(G) = \frac{E\left(\sum_{i=1}^N x_i\right)}{L} = \frac{E\left(E\left(\sum_{i=1}^N x_i \mid N = n\right)\right)}{L} \quad (1)$$

$$= \frac{\sum_{n=0}^{\infty} n \Pr(x_i = 1 \mid N = n) \Pr(N = n)}{L}.$$

The node  $n_i$  sends a packet successfully if both of the following conditions are satisfied,

- 1)  $n_i$  has one or more neighbors,
- 2) no node in  $A$  is transmitting.

Let  $B$  be the event that  $n_i$  has one or more neighbors and there are  $M$  nodes in  $A$ . Then applying The Law of Total Probability, the conditional probability of the event  $\{x_i = 1\}$ , given  $N$ , is,

$$\Pr(x_i = 1 | N) = \sum_B \Pr(x_i = 1 | B, N) \Pr(B | N). \quad (2)$$

Given that  $N = n$ , the locations of the nodes are  $n$  independent variates with identical distribution that is uniform in the region  $S$  [13]. Let the function  $N(C)$  represent the number of nodes in region  $C$ . So, the conditional probability of event  $B$ , given  $N$ , is

$$\begin{aligned} \Pr(B | N = n) &= \Pr(N(S_C) \geq 1 \& N(A) = M | N = n) \\ &= C_{n-1}^M \left(\frac{A}{S}\right)^M \left(1 - \frac{A}{S}\right)^{n-M-1} \\ &\quad - C_{n-1}^M \left(\frac{A}{S}\right)^M \left(1 - \frac{A}{S} - \frac{S_C}{S}\right)^{n-M-1}. \end{aligned} \quad (3)$$

According to the conflict graph coloring model, in each frame, each link is randomly assigned to one of the  $L$  time slots. The coloring algorithm can ensure that the assigned time slot is different from those occupied by  $V_D$ . So, when the number of nodes in  $A$  is  $M$ , the probability that a link has no conflict with the nodes in  $A$  is  $(1 - 1/L)^M$ , i.e.

$$\Pr(x_i = 1 | B, N) = (1 - 1/L)^M. \quad (4)$$

Now substituting (3) and (4) in (2) gives,

$$\Pr(x_i = 1 | N) = \left(1 - \frac{A}{SL}\right)^{n-1} - \left(1 - \frac{A}{SL} - \frac{S_C}{S}\right)^{n-1}. \quad (5)$$

Combining (1) and (5) yields the result,

$$E(G) = \frac{\lambda S}{L} \exp\left(-\frac{\lambda A}{L}\right) (1 - \exp(-\lambda S_C)). \quad (6)$$

### B. Overhead Analysis

The intuition behind formula (6) is that the network throughput improves when each node gets more information. However, collecting information consumes bandwidth resource, which may affect the network capacity. In this section, we will analyze the tradeoff between the network capacity improvement and the information collection overhead.

Nodes in the network distribute information by information packets which can be piggybacked onto data packets. Each information packet is marked with the location of the node which sends the information originally. According to the constraint of the information collection range and the location of the packet sender, the nodes receiving the information packet can decide whether the information should be forwarded.

Suppose each node can transmit at  $W$  bits per second, and the state of a node can be described using  $I$  bits. If a node

sends its own information to the nodes within its information collection range, the number of information packets that should be transmitted is equal to the number of nodes in its information collection area  $S_D$ . So, the average length of the information packets that should be transmitted by all the nodes in the network is  $NN(S_D)I$ . Since there are  $G$  successful transmissions per time slot, the total time consumed to transmit the information packets is  $NN(S_D)I/WG$ .

Suppose the length of each time slot is  $\tau$  seconds, then in each frame, the successfully transmitted effective data is  $G\tau WL$  bits and the time spent in transmitting effective data is  $L\tau$  seconds. Hence, it takes  $L\tau + N\frac{N(S_D)I}{WG}$  seconds in total to complete the transmissions in a frame. Therefore, the network can achieve an effective data rate of

$$G' = \frac{G\tau WL}{L\tau + N\frac{N(S_D)I}{WG}}.$$

Let  $P = \frac{I}{W\tau}$ , i.e.  $P$  is the ratio of the information packet length to the payload. Then  $G'$  can be rewritten as

$$G' = \frac{GL}{GL + NN(S_D)P} GW, \quad (7)$$

where  $\frac{GL}{GL + NN(S_D)P}$  is the proportion of effective data. It can be seen from (7) that both  $G$  and  $N(S_D)$  increase with the information collection range  $R_D$ , and a properly chosen  $R_D$  can maximize the effective data rate  $G'$ .

## IV. EXPERIMENTAL VALIDATION

We simulate three quantities, namely, throughput, communication overhead, and effective data rate.

### A. Simulation Setup

Nodes are randomly located in a  $50 \times 50$  square area according to a Poisson distribution. Four different network densities are chosen for the simulation:  $\lambda = 1, 2, 3$ , and  $4$ .

The communication range is 1 and the interference range is 2.5. The information collection range varies from 1 to 2.5. We only observe the data of the nodes in the central region, of size  $40 \times 40$ , so that the edge effects are isolated from other phenomena. The simulations are implemented in C++ and the simulation results are averaged over 900 runs.

### B. Validation of Throughput Analysis

Figure 4 plots the throughput calculated by the formulation and the average throughput observed by simulation against the information collection range  $R_D$ . It is obvious that as  $R_D$  increases, the throughput rises, as it should be. From the simulation result, we can see that in all the scenarios, the error between the analytical results and the simulation results decreases as  $R_D$  increases. The reason of the phenomenon is as follow. As shown in Figure 5, node  $n_i$  is sending a packet to node  $n_j$ . We can observe that when  $R_D$  is less than  $2R_C$ , there is an overlap (the shaded area) between  $n_j$ 's unknown area  $A$  and  $n_i$ 's communication area  $S_C$ , which means that the number of nodes in  $A$  and that in  $S_C$  are not independent. However, we make an approximation in equation (3) that the

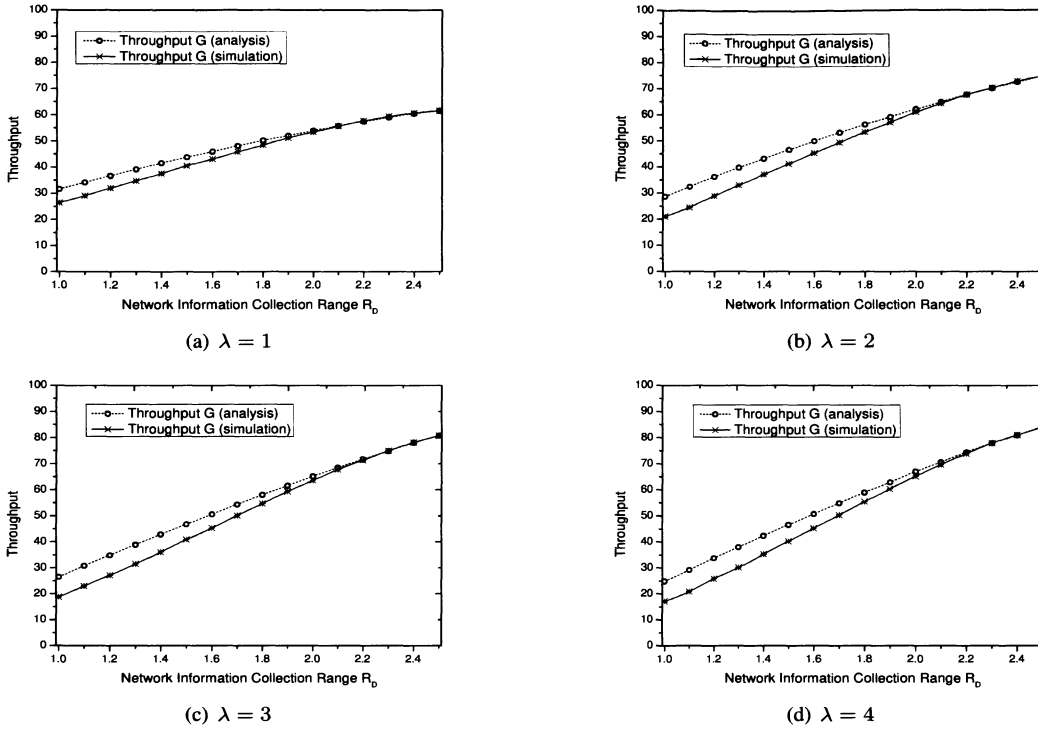


Fig. 4. Throughput Comparison between Analysis and Simulation

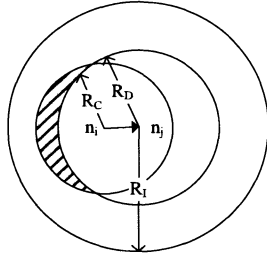


Fig. 5. Overlap between  $A$  and  $S_C$

number of nodes in  $A$  and that in  $S_C$  are independent, which leads to the difference between the analytical results and the simulation results. The area of the overlap declines as the information collection range increases, and finally the overlap disappears when  $R_D \geq 2R_C$ . Therefore, the two results match well under that condition.

### C. Communication Overhead

Communication overhead is incurred when nodes send information packets to distribute their own information. So, we use the number of information packets to evaluate the communication overhead. In Figure 6, we plot the average number of information packets that are required to get certain information versus the information collection range, for the four scenarios, respectively. It is obvious that the number of information packets increases with the information collection range and the network density.

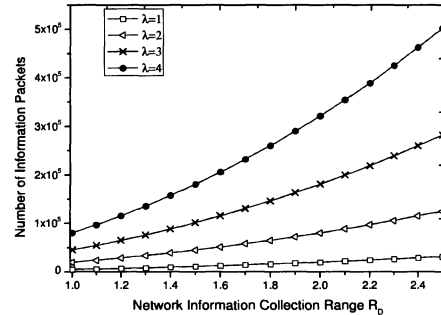


Fig. 6. Overhead of Information Collection

### D. Effective Data Rate

Figure 7 plots the effective data rate  $G'$  under the four scenarios. The ratio of the information packet length to the payload is set to 1%, 5% and 10%, respectively. We can see that even when the overhead is considered, the effective data rate increases with the information collection range. This is because the number of information packets is linearly proportional to  $S_D$ , while  $G$  increases exponentially with  $S_D$ . So  $G^2$  increases much faster than  $N(S_D)$  as  $R_D$  increases. Therefore,  $G'$  increases with  $R_D$  and achieves its maximum value when  $R_D = R_I$ .

## V. CONCLUSIONS AND FUTURE WORK

In this paper, we first observe the importance of network information, which was ignored before, in determining the

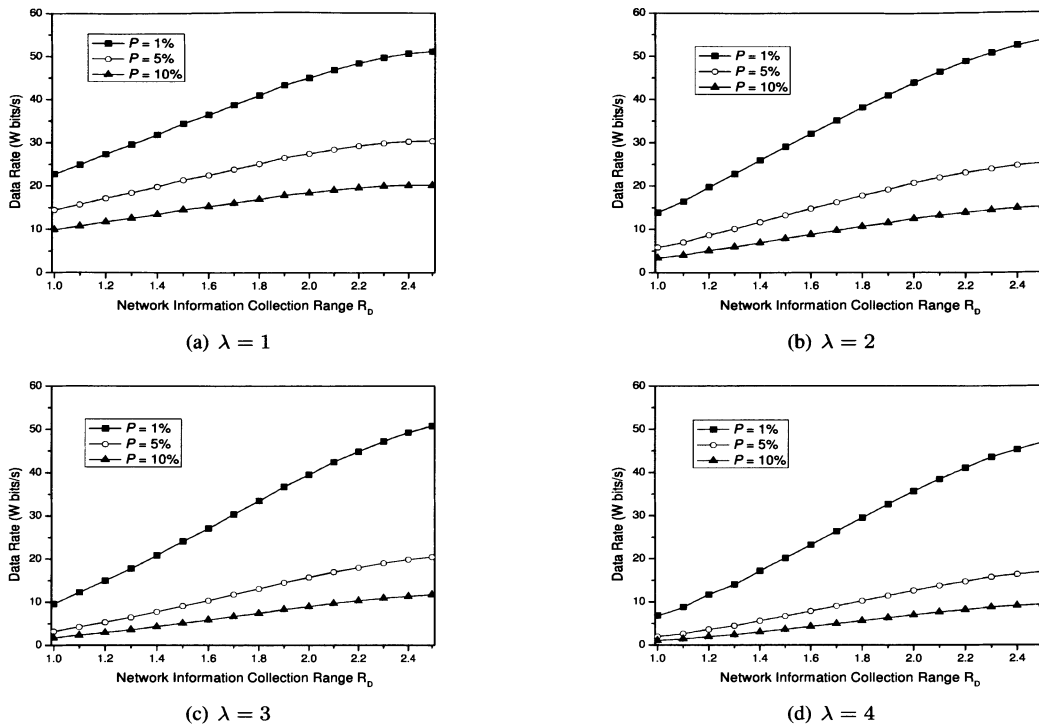


Fig. 7. Effective Data Rate

capacity of wireless networks. We use a conflict graph model to describe the wireless interference and an optimal graph coloring algorithm is employed to maximize the network capacity with limited information. The analytical result on the relationship between network capacity and the information got by each node is then derived. Since distributing information requires communication overhead, which is also an important metric to evaluate the network performance, we analyze the overhead, and reconsider the network performance by calculating the effective data rate.

We notice that the network capacity is determined not only by obtained information, but also by many other factors, such as the scheduling scheme, the mobility of wireless networks and so on. There is an opportunity for further research in considering the effects of scheduling and network mobility together with network information.

#### ACKNOWLEDGMENT

This research is supported in part by the Research Grants Council of Hong Kong, under Grant No. HKU 7148/06E.

#### REFERENCES

- [1] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, 2000.
- [2] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," in *the 9th annual international conference on Mobile computing and networking (MobiCom '03)*. New York, NY, USA: ACM Press, 2003, pp. 66–80.
- [3] M. Kodialam and T. Nandagopal, "The effect of interference on the capacity of multi-hop wireless networks," in *IEEE Symposium on Information Theory*, 2004.

- [4] V. Kolar and N. Abu-Ghazaleh, "A multi-commodity flow approach for globally aware routing in multi-hop wireless networks," in *the Fourth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom 2006)*, Mar. 2006.
- [5] M. Garetto, T. Salonidis, and E. W. Knightly, "Modeling per-flow throughput and capturing starvation in CSMA multi-hop wireless networks," in *the 25th IEEE International Conference on Computer Communications (INFOCOM 2006)*, 2006, pp. 1–13.
- [6] V. Kolar and N. Abu-Ghazaleh, "The effect of scheduling on link capacity in multi-hop wireless networks," *Arxiv preprint cs.NI/0608077*, 2006.
- [7] T. Hou and V. O. K. Li, "Transmission range control in multihop packet radio networks," *IEEE Transactions on Communications*, vol. 34, no. 1, pp. 38–44, 1986.
- [8] X. Yang and N. Vaidya, "Priority scheduling in wireless ad hoc networks," *Wireless Networks (Springer)*, vol. 12, no. 3, pp. 273–286, 2006.
- [9] A. Puri, "Optimizing traffic flow in fixed wireless networks," in *IEEE Wireless Communications and Networking Conference (WCNC 2002)*, vol. 2, 2002, pp. 904–907.
- [10] M. V. Marathe, A. Panconesi, and L. D. Risinger, "An experimental study of a simple, distributed edge coloring algorithm," in *the twelfth annual ACM symposium on Parallel algorithms and architectures (SPAA '00)*. New York, NY, USA: ACM Press, 2000, pp. 166–175.
- [11] X. Zhang, J. Hong, L. Zhang, X. Shan, and V. O. K. Li, "CC-TDMA: Coloring- and Coding-based multi-channel TDMA scheduling for wireless Ad Hoc networks," in *IEEE Wireless Communications and Networking Conference (WCNC 2007)*, March 2007, pp. 133–137.
- [12] A. Raniwala and T.-c. Chiueh, "Architecture and algorithms for an IEEE 802.11-based multi-channel wireless mesh network," in *the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005)*, vol. 3, 2005, pp. 2223–2234.
- [13] R. G. Gallager, *Discrete Stochastic Processes*. Springer, 1996.