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

Jun Hong, *Student Member, IEEE,* and Victor O.K. Li, *Fellow, IEEE*

***Abstract*—Previous research has estimated the performance of wireless networks by assuming that nodes 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 wireless network performance and available network information. It is assumed that each node in the network can obtain the information about other nodes within its information collection range, and a distributed graph coloring algorithm is employed to perform scheduling with the available information. The analytical result on the quantitative relationship between the information collection range and the network throughput is derived. We also consider the communication overhead of collecting information, and analyze the tradeoff between network capacity improvement and information collection overhead. Based on the derived result, an optimal information collection range which maximizes the net data rate can be found. Since wireless networks are typically mobile, and the collected information may be inaccurate due to the dynamics of the networks, we analyze the effect of information for mobile wireless networks by considering the information updating rate, and the result can be used to determine the information collection range as well as the information updating period.**

***Index Terms*—Network state information, scheduling, multi-hop wireless networks.**

I. INTRODUCTION

NETWORK information, such as network topology, channel state, and traffic 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 may consume valuable bandwidth resource in wireless networks. In this paper, we use a conflict graph to model the interference relationship between wireless links, and an algorithm based on graph coloring is applied to perform distributed scheduling with limited information. Then we analyze the relationship between available network information and the achievable network performance. Some approximations are made in the analysis. Simulation results show that the relative error due to such approximations of the estimate 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. Since node movements in wireless networks may change the network states, we establish a model for mobile wireless networks and analyze how information changes with the mobility rate. The analysis in this paper may be used to determine the information collection parameters in wireless networks.

The rest of the paper is organized as follows. Section II presents the state of the art. Section III describes the network model used throughout the paper and a conflict graph coloring model is introduced. Section IV analyzes the relationship between available network information and wireless network performance, and determines the tradeoff between information collection overhead and network performance improvement. The experimental evaluation is presented in Section V. Section VI concludes the paper.

II. RELATED WORK

Previous research has estimated the performance of wireless networks. Gupta and Kumar [1] first determine the capacity of wireless networks. Franceschetti *et al.* [2] close the gap between the capacity upper and lower bounds in Gupta and Kumar's original results. Zhang and Hou [3] extend this work to networks with unlimited bandwidth. Some researchers analyze the impact of interference on multi-hop wireless network performance. Jain *et al.* [4] use a conflict graph to model the wireless interference and compute upper and lower bounds on the network throughput. Kodialam and Nandagopal [5] analyze the effect of interference on the achievable data rate in wireless networks. Kolar and Abu-Ghazaleh [6] study the performance of globally aware routing which is cognizant of the wireless interference. The scheduling effects on wireless network performance have also been studied. Garetto *et al.* [7] use a Markovian model to estimate the effects of scheduling on the throughput of CSMA channels. Kolar and Abu-Ghazaleh [8] evaluate the scheduling interactions among several given links and analyze the scheduling effects on network capacity. We have observed that wireless network performance not only depends on the scheduling algorithm, but also on the available network information. However, most existing work on the capacity of wireless networks only provides an upper bound on the achievable capacity under the assumption that precise network information is available.

By noting that global information collection may be infeasible and may incur a large overhead in wireless networks, many distributed scheduling algorithms which only require local information are proposed [9], [10], [11], and the bounds on the performance of localized scheduling are given[10]. However, in addition to performance bound of localized scheduling, we are also interested in the quantitative relationship between available network information and scheduling performance.

Manuscript received December 8, 2009; revised April 22, 2010 and June 29, 2010; accepted August 19, 2010. The associate editor coordinating the review of this paper and approving it for publication was F. A. Cruz-Perez.

J. Hong and V. O.K. Li are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, China (e-mail: {hongjun, vli}@eee.hku.hk).

Digital Object Identifier 10.1109/TWC.2010.090210.091764

The first work to study the required network information theoretically can be found in [12], where rate distortion theory is used to find limits on the information required to indicate the start time and length of messages. Several recent papers also use rate distortion theory to derive the lower bound on the routing overhead required to restrict the error of routing information within a given threshold [13], [14], [15]. While [13], [14], [15] focus on routing, [16] analyzes the effect of information on scheduling performance by using an information-theoretic framework based on rate distortion theory. However, we found that as the number of nodes in the network increases, it is difficult to derive the network performance degradation measure, which is required as an input parameter for the rate distortion function.

In our previous work [17], we have analyzed the effect of information on wireless network performance, but the dynamics of the networks are ignored. We observe that the collected information may be inaccurate due to the dynamics of the networks. In this work, we analyze the effect of information for mobile wireless networks by considering the information updating rate as well as the information collection range.

III. Model and Definitions

A. Network Model

Suppose that N nodes are located in a region of area $S\ m^2$, and S is also used to refer to the region itself. The nodes of the networks are distributed as a two-dimensional Poisson point process with density λ[18], i.e., $\Pr(i \text{ nodes in } C) = (\lambda C)^i \exp(-\lambda C) \big/ i!, i = 0, 1, \cdots$. We assume that the nodes are homogeneous and all data transmissions employ the same power and communication parameters. 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 range is R_I, i.e., a transmitter may interfere with the receivers which are within a range of R_I, and $R_I = \rho R_C$, where ρ is a constant typically between 2 and 3 [19]. n_i denotes the i-th node, and d_{ij} denotes the distance between nodes n_i and n_j. Nodes n_i and n_j are said to be each other's neighbor if $d_{ij} \le R_C$. 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 with the same length, and grouped into frames. It is assumed that the frame size is fixed, and each frame consists of L time slots, where L is set to $\lfloor \lambda \pi R_I^2 \rfloor$. In each frame, every node is assigned one of the slots, and starts transmission at the beginning of the time slot assigned. $T(i)$ denotes the time slot assigned to n_i.

We use the protocol interference model [1] to define the conditions for successful transmissions.

Protocol Interference Model: a transmission between two nodes n_i and n_j is successful if

1) the two nodes are within communication range of each other, i.e., $d_{ij} \le R_C$,
2) no nodes within a receiving node's interference range is transmitting using the same time slot.

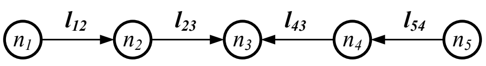
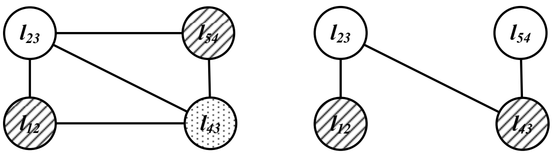


Fig. 1. A chain-topology wireless network.



(a) $R_D = R_I$ (b) $R_D = R_C$

Fig. 2. Conflict graphs.

The information collection range is R_D, which means that each node can obtain the information (including location, assigned time slot, traffic information, etc.) of the nodes within a range of R_D. According to the protocol interference model, nodes out of the interference range can not affect the transmission, and since the objective of collecting information is to avoid conflicts, the information collection range R_D is less than R_I. Hence, we only consider the situation when $0 \le R_D \le R_I$, and $R_D = 0$ means that the nodes in the network do not collect any information which indicates the time slots occupied by other transmissions. A coloring algorithm based on the collected information is used to perform the scheduling of the transmissions. The details of the coloring algorithm will be introduced in Section III-C.

B. Conflict Graph Coloring Model

A wireless network can be represented as a bi-directional graph G^p. The vertices of G^p correspond to the wireless nodes and the edges correspond to the wireless links between the nodes. There is a directed link $l_{ij} (i \ne j)$ from n_i to n_j if 1) there is a packet from n_i to n_j, and 2) $d_{ij} \leqslant R_C$. However, wireless interference, which is a key issue impacting network performance, cannot be modeled accurately in G^p. The effects of interference in such a network can be modeled as a conflict graph[4],[20],[21], 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 $d_{in} \leqslant R_I$, or $d_{mj} \leqslant R_I$.

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. In other words, the adjacent vertices with different colors correspond to two links that cannot be activated simultaneously. Thus, the transmission scheduling in a wireless network is equivalent to the coloring of a conflict graph[22],[23].

An example of a chain-topology wireless network is shown in Figure 1, where the distance between the adjacent nodes is just R_C. Since the ratio of the interference range and the communication range is typically between 2 and 3 [19], here ρ is set to 2.5, and $R_I = 2.5 \times R_C$. The corresponding conflict graph of Figure 1 is shown in Figure 2(a). The four vertices in the conflict graph correspond to the four directional links in Figure 1 and the edges between the vertices represent their conflict relationships. As shown in Figure 2(a), the conflict graph is colored using three colors, where a color corresponds to a time slot in TDMA scheduling. 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 complete network information, i.e., precise topology information and traffic information. If each node collects only part of the network information, the conflict graph may be incorrectly estimated, leading to inappropriate scheduling and collisions. Suppose each node collects information within a range of R_C, i.e., $R_D = R_C$, then Figure 2(b) shows the corresponding conflict graph, the edges of which is a subset of that in the true conflict graph. Since the edge between l_{12} and l_{43}, as well as the edge between l_{23} and l_{53}, exist in reality but is not detected and ignored in Figure 2(b), a packet collision happens. As a result, although the conflict graph is colored with two colors, the number of successful packets per time slot is 0. It is obvious that incomplete information results in degraded performance.

C. Algorithm Specification

We specify the details of the implementation of the coloring algorithm. This algorithm is based on the M coloring algorithm introduced in [22]. Some changes are made to adapt to our network scenarios. Suppose each node has two palettes, i.e., sender-palette and receiver-palette. The sender-palette indicates the available colors for the node as a sender, while the receiver-palette indicates the available colors for the node as a receiver. In an M coloring algorithm, the size of the palettes, which depends on the degree of the network, is flexible, so global information exchange is required for synchronization. However, in our network model, each node only exchanges information with nodes within a range of R_D, so flexible palette size is infeasible. Two problems may arise when the palette size is fixed, 1) the palette size is too small so that some of the links may not have available colors, and 2) the palette size is too large so that some of the colors are wasted. So, the palette size must be chosen carefully. According to the experimental results in [22], when the initial palette size is larger than 1.05Δ, the difference between the initial palette size and the final palette size (which indicates the number of actually used colors) is less than 3%, where Δ is the maximum degree of the graph. So, in our algorithm, the size of palettes L is set to $\lfloor \lambda\pi R_I^2 \rfloor$. The colors of all the palettes are the same, namely, color 1, color 2, $\cdots$, and color L, which correspond to time slot 1, time slot 2, $\cdots$, and time slot L.

Each link executes the following seven steps to get a color,

1) The sender chooses one of its neighbors randomly as receiver, and sends a message to the receiver to inform it of the transmission.
2) The receiver sends a list of the unavailable colors to the sender.
3) Based on the received list, the sender marks the unavailable colors in its sender-palette, and chooses a color from the available colors at random.
4) The sender sends the chosen color to the receiver.
5) If there is no available colors, the sender will inform the receiver, and cancel its transmission.
6) Both the sender and the receiver send their chosen colors to the nodes which are within their respective information collection ranges.
7) The nodes which receive the information from the sender mark the received colors as unavailable in its receiver-palette, and the nodes which receive the information from the receivers mark the received colors as unavailable in its sender-palette.

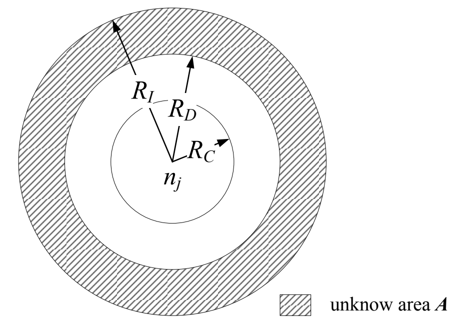


Fig. 3. Relationship between R_C, R_D and R_I.

IV. Analysis

A. Performance Analysis

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

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,

$$\begin{aligned} E(G) &= \frac{E\left(\sum_{i=1}^{N} x_i\right)}{L} = \frac{E\left(E\left(\sum_{i=1}^{n} x_i \,|N=n\right)\right)}{L} \\ &= \frac{\sum_{n=0}^{\infty} E\left(nx_i \,|N=n\right)\Pr\left(N=n\right)}{L} \\ &= \frac{\sum_{n=0}^{\infty} n\Pr\left(x_i=1\,|N=n\right)\Pr\left(N=n\right)}{L}. \end{aligned} \tag{1}$$

where L is the number of time slots in each frame, and N is the number of nodes in the network.

As shown in Figure 3, a node $n_j\,(j=1,2,\cdots,N)$ can obtain the information of the nodes within a range of R_D. If n_j receives a packet, the transmission may be affected by the transmitting nodes within the interference range R_I. The proposed 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_j and V_D, where V_D is a set of the nodes in S_D. For a receiver n_j, 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_j. A also refers to the size of the region, and $A = \pi(R_I^2 - R_D^2)$.

As a result, node n_i sends a packet successfully if both of the following conditions are satisfied,

1) n_i has one or more neighbors which can receive the packet,
2) none of the nodes in the receiver's unknown area A is transmitting using the same time slot.

Hence,

$$\Pr(x_i = 1 \,|N = n) = \Pr(\text{node } n_i \text{ has neighbors} \;\&\; \text{no nodes in } A \text{ use the same slot} \,|N = n). \quad (2)$$

Note that Equation (2) ignores the probability that there is no available color for node n_i.

Given that $N = n$, the locations of the nodes are n independent variates with identical distribution that is uniform in the region S[24]. So, the conditional probability that a given node n_i has neighbors is

$$\Pr(n_i \text{ has neighbors } |N = n) = 1 - (1 - \tfrac{S_C}{S})^{n-1}, \quad n \geq 1$$

According to the conflict graph coloring model, in each frame, each link is randomly assigned to one of the L time slots. So, the probability that a transmitter has conflict with the receiver is $\frac{A}{SL}$. Since there are $n-2$ (when $n \geq 2$) nodes in the network except the transmitter and the receiver, and we assume that the nodes affect the transmission independently, then the probability that none of the transmitters will affect the transmission is given by

$$\Pr(x_i = 1 \,|n_i \text{ has neighbors} \;\&\; N = n) = (1 - \frac{A}{SL})^{n-2}, n \geq 2 \quad (3)$$

Then we have,

$$\Pr(x_i = 1 \,|N = n) = \left(1 - (1 - \frac{S_C}{S})^{n-1}\right)\left(1 - \frac{A}{SL}\right)^{n-2}, n \geq 1 \quad (4)$$

Since it is a Poisson point process, the probability distribution function of N is,

$$\Pr(N = n) = \frac{(\lambda S)^n \exp(-\lambda S)}{n!}, n = 0, 1, 2, \cdots \quad (5)$$

Now substituting (4) and (5) in (1) gives,

$$E\left(\sum_{i=0}^{N} x_i\right) = \lambda S\left(\frac{\exp\left(-\frac{\lambda A}{L}\right)}{1 - \frac{A}{SL}}\right)\left(1 - \exp\left(-\lambda S_C(1 - \frac{A}{SL})\right)\right) \quad (6)$$

Combining (1) and (6) yields the result,

$$E(G) = \frac{\lambda S}{L}\left(\frac{\exp\left(-\frac{\lambda A}{L}\right)}{1 - \frac{A}{SL}}\right)\left(1 - \exp\left(-\lambda S_C(1 - \frac{A}{SL})\right)\right) \quad (7)$$

In Equation (7), the first term $\frac{\lambda S}{L}$ can be interpreted as the average number of transmissions allocated in each time slot. The second term $\frac{\exp\left(-\frac{\lambda A}{L}\right)}{1-\frac{A}{SL}}$ is identified as the effect of unavailable information on network throughput. The third term $\left(1 - \exp\left(-\lambda S_C(1 - \frac{A}{SL})\right)\right)$ can be viewed as an indication of the network connectivity. In particular, when the unknown area $A = 0$, which means that each node gets complete information, the second term is equal to one, i.e., the performance degradation due to incomplete information is eliminated, and the corresponding network throughput is $\frac{\lambda S}{L}(1 - \exp(-\lambda S_C))$.

B. Net Data Rate

Equation (7) says that the network performance improves when the information collection range increases, thus getting more information. However, collecting information consumes bandwidth resource, which may adversely affect the network performance. More specifically, transmitting network information diverts valuable bandwidth resource which may be used for data transmissions, thereby reducing network capacity. In this section, we will analyze the tradeoff between the network performance improvement and the network information collection overhead.

Suppose each node employs the same fixed power to broadcast coloring information to the nodes within the information collection range, i.e., when transmitting coloring information, the communication range R_C is set to R_D. Note that the power can be different from that used to transmit data. Since the transmission of coloring information is different from that of data in terms of transmission mode (broadcast and unicast) and applied power, when transmitting coloring information, the corresponding network capacity, denoted by G_I, is different from that derived by Equation (7). According to the protocol interference model specified in Section III-A, for each receiver, the coloring information is received successfully if none of the nodes within its interference range, which is equal to ρR_D, is transmitting simultaneously. Since there is no information to assist in the transmissions of coloring information, the unknown area for each information receiver is $A_I = \pi(\rho R_D)^2$. Then G_I can be calculated by using the same procedure introduced in Section IV-A, and we have

$$G_I = \frac{\lambda S}{L_I}\left(\frac{\exp\left(-\frac{\lambda\pi(\rho R_D)^2}{L_I}\right)}{1 - \frac{A_I}{S}}\right) \quad (8)$$

where L_I is set to $\left\lfloor \lambda\pi(\rho R_D)^2 \right\rfloor$. Then Equation (8) becomes

$$G_I \approx \frac{\lambda S}{L_I}\left(\frac{\exp(-1)}{1 - \frac{A_I}{S}}\right) \quad (9)$$

We can see that as R_D increases, which means that each node broadcasts the coloring information with a larger power, the network capacity G_I decreases.

Suppose each node can transmit at W bits per second, the length of each time slot is μ seconds, and the coloring information transmitted by each node is I bits. Since there is an average of λS links in the network, and the sender and receiver of each link should broadcast their coloring information once, the total time consumed to transmit the coloring information is $\frac{2\lambda SI}{WG_I}$. Suppose the update period is L_u time slots, which means that the nodes collect network information every $L_u\mu$ seconds. We assume that the movement of the nodes is quasi-static, i.e., the locations of the nodes remain unchanged during each updating period. So, in each L_u time slot, the number of bits of data transmitted successfully is $GWL_u\mu$, and the time for transmitting the data is $L_u\mu$. Therefore, the net data rate, which is defined as the data transmitted per second, is given by

$$G' = \frac{GWL_u\mu}{L_u\mu + \frac{2\lambda SI}{WG_I}}.$$

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

$$G' = \frac{L_u}{L_u + \frac{2\lambda SP}{G_I}} GW, \quad (10)$$

where $\frac{L_u}{L_u+\frac{2\lambda SP}{G_I}}$ is the proportion of data. It can be seen from (10) that as R_D increases, G increases and G_I decreases, and a properly chosen information collection range can maximize the net data rate G'.

C. Mobility Analysis

The nodes of wireless networks are typically mobile. The obtained network information may be inaccurate due to variations of the network topology. In Section IV-B, we just assume that the network state remains unchanged during each information updating period, and the collected information is accurate within each information updating period. However, this may not be true in practical systems. Similarly, a shorter information updating period results in a better network performance, because the available information is more accurate, but it also incurs more communication overhead, which may degrade the network performance. So, in this section, the effect of information on wireless network performance is reevaluated, and we consider the combined effect of information collection range and information updating period on network performance. Based on this analysis, an optimal information collection range, as well as an optimal information updating period, can be found.

The network model introduced in Section III is used. It is assumed that the sender-receiver pair does not change with time despite the mobility of the nodes, i.e., once a node chooses one of its neighbors to send a packet, it will not send packets to other nodes until the transmission is fulfilled or the information is updated. The speed and direction of the movements of nodes are time- and location-independent. All the nodes follow the same mobility model, and the movements of different nodes are independent and identically distributed (i.i.d). Then obviously, at any given time, the nodes of the networks are distributed as a two-dimensional Poisson distribution.

Consider a transmission from n_i to n_j. As the topology changes, the receiver n_j which was within the communication range when information was last collected may be out of the communication range when n_i sends a packet; some nodes which interfered with n_i may be out of the interference range, while other nodes which did not interfere with n_i may move into n_i's interference range. For each pair of nodes, four states can be used to indicate their location relationship, and the dynamics of the nodes can be described by the changes of the states. For a certain node n_i, the state of Node n_j at the t-th time unit, denoted by $S^{(t)}(j)$, is set according to the distance between n_i and n_j. Specifically, when $R_D \geq R_C$,

$$S^{(t)}(j) = \begin{cases} 1 & \text{if } d_{ij}^{(t)} < R_C \\ 2 & \text{if } R_C \leq d_{ij}^{(t)} < R_D \\ 3 & \text{if } R_D \leq d_{ij}^{(t)} < R_I \\ 4 & \text{if } d_{ij}^{(t)} \geq R_I \end{cases} \quad (11)$$

and when $R_D < R_C$,

$$S^{(t)}(j) = \begin{cases} 1 & \text{if } d_{ij}^{(t)} < R_D \\ 2 & \text{if } R_D \leq d_{ij}^{(t)} < R_C \\ 3 & \text{if } R_C \leq d_{ij}^{(t)} < R_I \\ 4 & \text{if } d_{ij}^{(t)} \geq R_I \end{cases} \quad (12)$$

where $d_{ij}^{(t)}$ is the distance between n_i and n_j at the t-th time unit.

Without loss of generality, we just focus on the first situation, i.e., $R_D \geq R_C$, and the second situation can be analyzed by simply mapping the States 1, 3, and 4 in Equation (11) to the States $1 \cup 2$, $2 \cup 3$, and 4 in Equation (12), respectively.

Let $\alpha_{kl}^{(\tau)}(j)$ denote the transition probability that the state of n_j changes from k to l during τ time units, i.e., $\alpha_{kl}^{(\tau)}(j) = \Pr\left(S^{(t+\tau)}(j) = l \,\middle|\, S^{(t)}(j) = k\right)$, and $\sum_{l=1}^{4} \alpha_{kl}^{(\tau)}(j) = 1$. Since the movements of the nodes are i.i.d, the transition probabilities are the same for all the nodes, i.e., $\alpha_{kl}^{(\tau)}(1) = \alpha_{kl}^{(\tau)}(2) = ... = \alpha_{kl}^{(\tau)}$.

The probability that the state of n_j is k at the t-th time unit is denoted by $P_k^{(t)}(j)$, i.e., $P_k^{(t)}(j) = \Pr\left\{S^{(t)}(j) = k\right\}$,and $\sum_{k=1}^{4} P_k^{(t)}(j) = 1$. The state probability distributions of all the nodes are identical, i.e., $P_k^{(t)}(1) = P_k^{(t)}(2) = ... = P_k^{(t)}$.

It is assumed that the length of each time slot is less than or equal to the length of each time unit, so the transmission will not be interrupted. Suppose the network information is collected at the t-th time unit, and a certain transmission is performed at the $(t+\tau)$-th time unit ($\tau = 0, 1, 2 \cdots$) based on the collected information. Then a node n_i sends a packet successfully when the following conditions are satisfied

1) at t, n_i has one or more neighbors,
2) the selected receiver, say n_j, is within n_i's communication range at $t+\tau$ (i.e., $S^{(t+\tau)}(j) = 1$),
3) any node, say n_k, which is out of n_j's information collection range at t (i.e., $S^{(t)}(k) > 2$), and within n_j's interference range at $t+\tau$ (i.e., $S^{(t+\tau)}(k) < 4$) does not send packets in the same time slot as n_i.

Given that $N = n$, for a certain node n_i, the probability that there is at least one node within its communication range at the t-th time unit is,

$$\begin{aligned} P_{\bar{0}}^{(t)} &= \Pr(\text{one or more neighbors at } t | N = n) \\ &= 1 - (1 - P_1^{(t)})^{n-1} \end{aligned}$$

The probability that the selected receiver is still within n_i's communication range at the $(t+\tau)$-th time unit is

$$\Pr\left(S^{(t+\tau)} = 1 \,\middle|\, S^{(t)} = 1\right) = \alpha_{11}^{(\tau)}$$

For a certain node n_k, the probability that it has a conflict with n_j is

$$P_C^{(\tau)} = \Pr\left(d_{kl}^{(t)} > R_D \,\&\, d_{kj}^{(t+\tau)} < R_I \,\&\, T(i) = T(k) | N = n\right)$$

where $T(i)$ is the time slot assigned to n_i.

According to the analysis in Section IV-A,

$$\Pr\left(T(j) = T(k) \left| d_{kj}^{(t)} > R_D, N = n\right.\right) = \frac{1}{L}$$

And

$$\begin{aligned} &\Pr\left(d_{kj}^{(t)} > R_D \,\&\, d_{kj}^{(t+\tau)} < R_I\right) \\ &= \Pr\left(d_{kj}^{(t)} > R_D\right) \Pr\left(d_{kj}^{(t+\tau)} < R_I \left| d_{kj}^{(t)} > R_D\right.\right) \\ &= P_3^{(t)}(1-\alpha_{34}^{(\tau)}) + P_4^{(t)}(1-\alpha_{44}^{(\tau)}) \end{aligned} \quad (13)$$

So,

$$P_C^{(\tau)} = \frac{P_3^{(t)}(1-\alpha_{34}^{(\tau)}) + P_4^{(t)}(1-\alpha_{44}^{(\tau)})}{L} \quad (14)$$

Hence, when there are n nodes in the network, the probability that no nodes has a conflict with n_j is

$$P_{\bar{c}}^{(\tau)} = \left(1 - P_C^{(\tau)}\right)^{n-2}$$

So,

$$\Pr\left(x_i = 1 \left| T(i) = t+\tau, N = n\right.\right) = P_{\bar{c}}^{(\tau)} P_{\bar{0}}^{(\tau)} \alpha_{11}^{(\tau)}$$

Hence, the network throughput is given by

$$\begin{aligned} &E(G_m) \\ &= \frac{1}{L_u} \sum_{n=0}^{\infty} \sum_{\tau=0}^{L_u-1} n \Pr\left(x_i = 1 \& T(i) = t+\tau \& N = n\right) \\ &= \frac{1}{LL_u} \sum_{n=0}^{\infty} \sum_{\tau=0}^{L_u-1} n P_{\bar{c}}^{(\tau)} P_{\bar{0}}^{(t)} \alpha_{11}^{(\tau)} \frac{(\lambda S)^n e^{-\lambda S}}{n!} \\ &= \frac{\lambda S}{LL_u} \sum_{\tau=0}^{L_u-1} \alpha_{11}^{(\tau)} \frac{e^{-\lambda S P_C^{(\tau)}}}{1 - P_C^{(\tau)}} \left(1 - e^{-\lambda S_C \left(1 - P_C^{(\tau)}\right)}\right) \end{aligned} \quad (15)$$

Therefore, $E(G_m)$ is a function of $P_C^{(\tau)}$ and $\alpha_{11}^{(\tau)}$, which in practice, can be obtained by measurement or from historical data. From (15), we can see that the network throughput depends on the mobility rate, and the updating period L_u, which represents the delay between information collection and transmission. As the mobility rate and L_u increase, the collected information becomes less precise, resulting in performance degradation.

When the network is immobile,

$$\alpha_{kl}^{(\tau)} = \begin{cases} 0 & \text{if } k \neq l \\ 1 & \text{if } k = l \end{cases},$$

and according to Equation (14),

$$\forall \tau \in \{0, 1, \cdots, L_u - 1\}, P_C^{(\tau)} = \frac{P_3^{(t)}}{L} = \frac{A}{SL}.$$

In such a case, the network throughput becomes $\frac{\lambda S}{L}\left(\frac{\exp\left(-\frac{\lambda A}{L}\right)}{1-\frac{A}{SL}}\right)\left(1 - \exp\left(-\lambda S_C(1 - \frac{A}{SL})\right)\right)$, which we can see, is exactly the same as Equation (7).

According to the analysis in Section IV-B, we can also derive the net data rate when the impact of information updating period is considered, i.e.,

$$G'_m = \frac{L_u}{L_u + \frac{2\lambda SP}{G_I}} G_m W, \quad (16)$$

V. Experimental Validation

In this section, we have the following goals:

1) We verify the analysis of network performance in Section IV by comparing the analytical results with the simulation results.
2) We evaluate the network performance when the communication overhead is taken into account.
3) We evaluate the network performance when the dynamics of the networks and the information updating period are considered.

A. Simulation Setup

Nodes are randomly located in a 40×40 square area according to a Poisson distribution. The network density is set to two different values, i.e., λ is set to 1 and 4, respectively. We only observe the data of the nodes in the central region, of size 30×30, so that the edge effects are isolated from other phenomena. The nodes move according to a 2-D Random Walk Mobility Model[25]. That is, in each time unit, each node chooses a direction from $[0, 2\pi]$ randomly, and moves with the chosen direction and a constant speed v. The speed v is set to 0.1 and 0.5 per time unit. The length of each time unit is set to the length of the time slot. The information updating period L_u varies from 2 to 20 time slots.

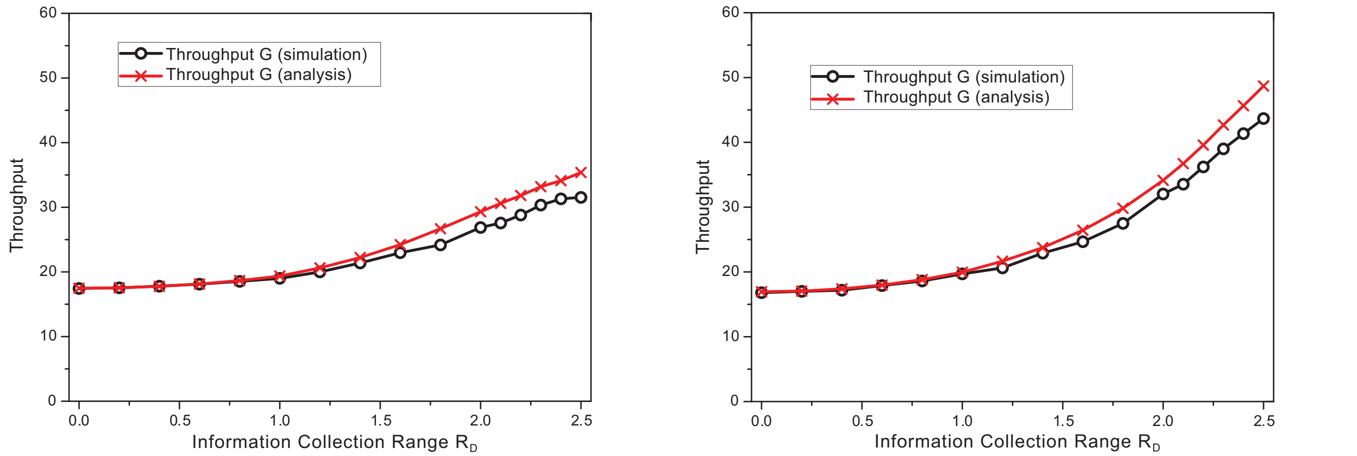
The communication range of nodes is 1 and the interference range is 2.5. The information collection range varies from 0 to 2.5. The simulations for all scenarios are implemented in C++ and each simulation result is averaged over 900 runs.

We evaluate two quantities, namely, throughput and net data rate, which determine the performance of wireless networks. The definitions of the two quantities are as follow:

1) Throughput: the average number of successful packets in each time slot.
2) Net data rate: the number of bits of data transmitted per second.

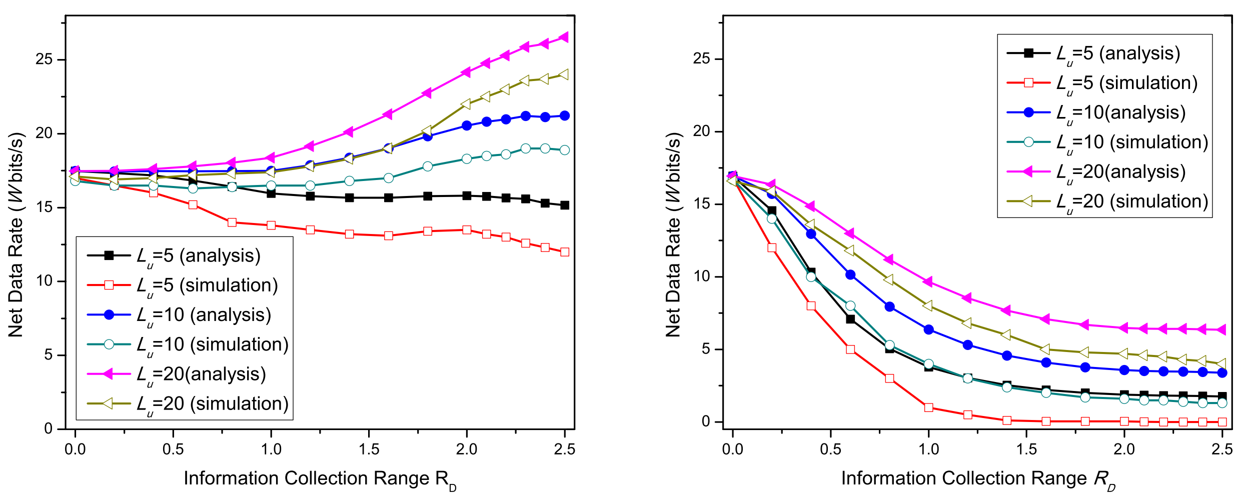
B. Validation of Network Throughput

Figure 4 compares the analytical throughput and the throughput by simulation against the information collection range R_D. It is obvious that as the information collection range R_D increases, the throughput increases, as it should. As network density increases, the network throughput increases slightly, because there are more links in the network. The difference between the analytical results and the simulation results are mainly due to 1) the assumption of independence among nodes (Equation (3)), 2) due to the fixed frame size, some nodes may not have colors to choose from, and 3) the failure of the coloring process due to the lack of information and coordination. From the simulation results, we can see that as the information collection range increases, the difference between the analytical results and the simulation results



(a) $\lambda = 1$ (b) $\lambda = 4$

Fig. 4. Comparison of throughputs obtained by analysis and simulation.



(a) $\lambda = 1, P = 1\%$ (b) $\lambda = 4, P = 5\%$

Fig. 5. Net data rate as a function of the information collection range.

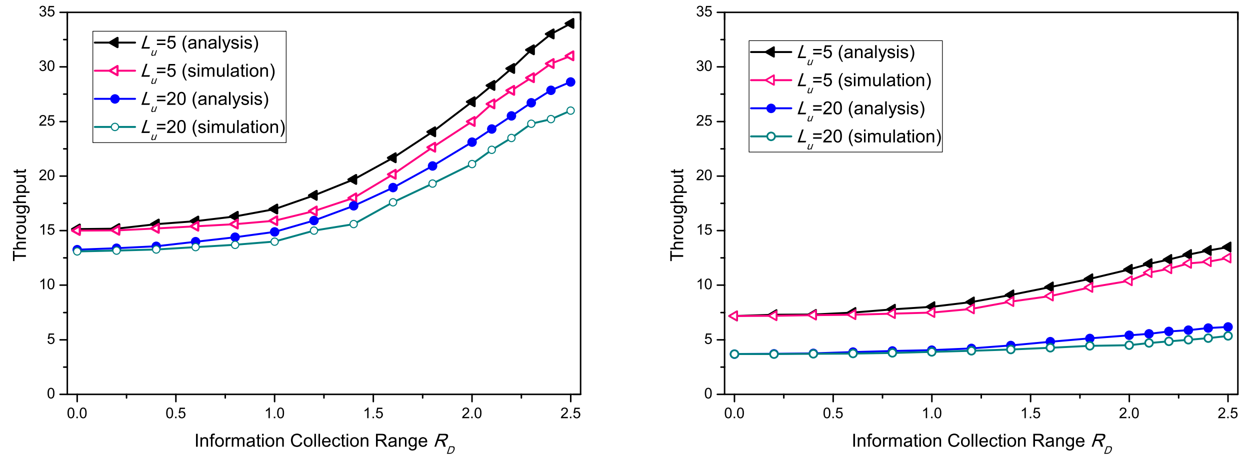
becomes larger. This is because as the information collection range increases, the dependency among nodes becomes stronger, and the probability that a node may not have any colors to choose from becomes larger.

C. Validation of Net Data Rate

Figure 5 plots the net data rate G'. The ratio of the information packet length to the data packet length is set to 1% and 5%, respectively. The information updating period L_u is set to 5, 10, and 20 time slots. Note that in this scenario, the information updating period is a parameter which indicates the dynamics of the networks, and we assume that during each information updating period, the network state remains unchanged. We will measure the effect of information updating period in Section V-D. We can see that when $P = 1\%$ and $L_u = 20$, which means that the information packet is short, and the network topology changes slowly, larger information collection range brings larger net data rate. As the information packet length and the dynamics of the network increase, the extra communication overhead incurred by enlarging the information collection range is considerable, and the net data rate decreases as the information collection range increases.

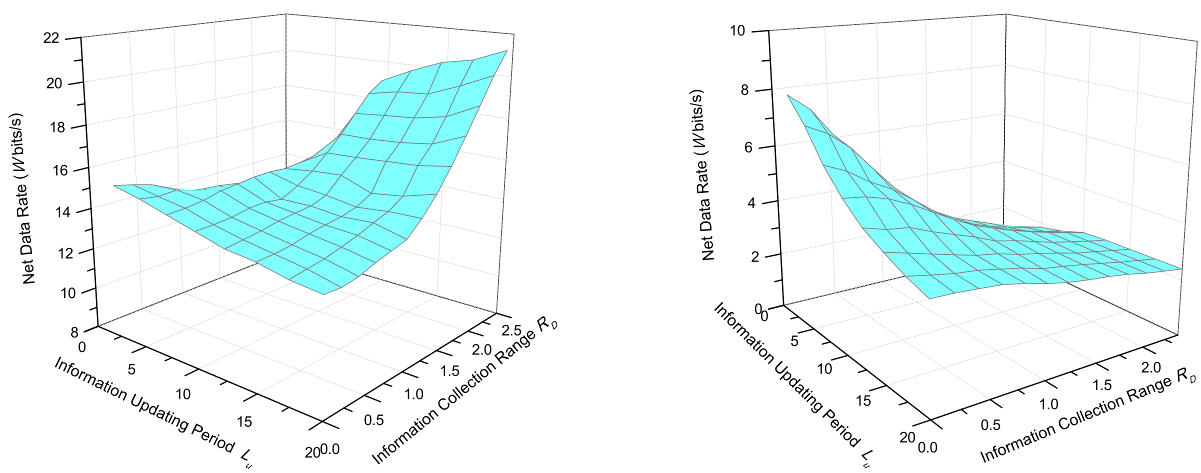
D. Mobile Networks

Figure 6 compares the analytical throughput and the throughput by simulation against the information collection range R_D for two different network mobility rates. As shown in the figure, network throughput decreases with the speed of the nodes. This is because with the same information collection range and information updating period, the collected information is more likely to be inaccurate when the nodes move faster. When the mobility rate is small ($v = 0.1$), we can see that the network throughput increases as the information collection range increases, but the benefit obtained with a shorter information updating period is relatively small. Specifically, when L_u decreases from 20 time slots to 5 time slots, the network throughput increases by less than 20%. As the mobility rate increases, we can see that the effect of the information updating period becomes larger, and when $v = 0.5$, by reducing the updating period from 20 time slots to



(a) $v = 0.1$ (b) $v = 0.5$

Fig. 6. Comparison of throughputs of mobile networks obtained by analysis and simulation.



(a) $v = 0.1, P = 1\%$ (b) $v = 0.5, P = 5\%$

Fig. 7. Net data rate of mobile networks as a function of information collection range and information updating period.

5 time slots, the network throughput is improved by more than 100%. However, the network throughput is hardly improved by the information collection range, because the collected network information becomes imprecise and useless due to the rapid variation of the network topology. Note that when the information collection range R_D is 0, which means that the nodes do not disseminate coloring information to any other nodes, the network throughput with shorter updating periods is larger. This is because we assume that at the beginning of each updating period, each transmitter will update the information of its receiver (even though the receiver is out of its information collection range) so as to establish a link.

Figure 7 plots the net data rate when considering the effect of the information updating period. When $v = 0.1$ and $P = 1\%$, i.e., the mobility rate and the information packet are small, a longer information updating period and a larger information collection range are preferred, and the optimal net data rate is achieved when $L_u = 20$ and $R_D = 2.5$. As the mobility rate and the information packet length increase, a shorter information updating period and a smaller information collection range result in larger net data rate, and the net data rate is maximized when $L_u = 2$ and $R_D = 0$, which means that when the network is highly dynamic, to maximize the net data rate, each transmitter should update its receiver information frequently, and should not exchange the coloring information with any other nodes.

VI. Conclusions and Future Work

Most existing results on wireless network capacity have been derived based on the implicit assumption of perfect network state information and negligible overhead in obtaining network state information. In this work, we investigate the quantitative relationship between the available network information and the scheduling performance in wireless networks. We use a conflict graph model to describe the wireless interference and an graph coloring algorithm is employed to perform scheduling with limited information. The analytical result on the relationship between network performance and the information got by each node is then derived. Since collecting information requires communication overhead which degrades the network performance, we analyze the overhead, and reconsider the network performance by calculating the net data rate. We notice that the network performance is determined not only by the collected information, but also by the mobility of wireless networks, so we evaluate the effect of network information when the nodes in the networks are mobile.

In fact, the scheduling algorithm itself also influences the network performance. In the future, we intend to consider the impact of scheduling algorithms on the network performance. In addition, we intend to consider the impact of information on the performance of a complete MAC protocol.

REFERENCES

[1] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, 2000.

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

[3] H. Zhang and J. Hou, "Capacity of wireless ad-hoc networks under ultra wide band with power constraint," in *Proc. 24th IEEE Conference on Computer Communications (INFOCOM 2005)*, vol. 1, 2005, pp. 455–465.

[4] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," in *Proc. 9th Annual International Conference on Mobile Computing and Networking (MobiCom '03)*, New York, pp. 66–80.

[5] M. Kodialam and T. Nandagopal, "The effect of interference on the capacity of multi-hop wireless networks," *IEEE Symposium on Information Theory*, 2004.

[6] V. Kolar and N. Abu-Ghazaleh, "A multi-commodity flow approach for globally aware routing in multi-hop wireless networks," in *Proc. Fourth Annual IEEE International Conference on Pervasive Computing and Communication*, Mar. 2006.

[7] M. Garetto, T. Salonidis, and E. W. Knightly, "Modeling per-flow throughput and capturing starvation in CSMA multi-hop wireless networks," in *Proc. 25th IEEE International Conference on Computer Communications*, 2006, pp. 1–13.

[8] V. Kolar and N. Abu-Ghazaleh, "Scheduling aware network flow models for multi-hop wireless networks," in *Proc. IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2008*, pp. 1 –8.

[9] A. Ephremides and T. Truong, "Scheduling broadcasts in multihop radio networks," *IEEE Trans. Commun.*, vol. 38, no. 4, pp. 456–460, 1990.

[10] Y. Xu and W. Wang, "Scheduling partition for order optimal capacity in large-scale wireless networks," in *Proc. 15th annual International Conference on Mobile Computing and Networking*, 2009, pp. 109–120.

[11] J.-H. Ju and V. O. K. Li, "An optimal topology-transparent scheduling method in multihop packet radio networks," *IEEE/ACM Trans. Networking*, vol. 6, no. 3, pp. 298–306, 1998.

[12] R. G. Gallager, "Basic limits on protocol information in data communication networks," *IEEE Trans. Inf. Theory*, vol. 22, no. 4, pp. 385–398, 1976.

[13] G. Cheng and N. Ansari, "Rate-distortion based link state update," *Computer Networks*, vol. 50, no. 17, pp. 3300–3314, 2006.

[14] N. Bisnik and A. Abouzeid, "Capacity deficit in mobile wireless ad hoc networks due to geographic routing overheads," in *Proc. 26th IEEE International Conference on Computer Communications*, 2007, pp. 517–525.

[15] D. Wang and A. Abouzeid, "Link state routing overhead in mobile ad hoc networks: a rate-distortion formulation," in *Proc. 27th IEEE International Conference on Computer Communications*, 2008, pp. 1337–1345.

[16] J. Hong and V. O. K. Li, "Impact of information on network performance—an information-theoretic perspective," *IEEE Global Telecommunications Conference 2009*.

[17] ——, "The effect of information on scheduling performance in multi-hop wireless networks," in *Proc. 19th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2008.

[18] T. Hou and V. O. K. Li, "Transmission range control in multihop packet radio networks," *IEEE Trans. Commun.*, vol. 34, no. 1, pp. 38–44, 1986.

[19] A. Raniwala and T.-c. Chiueh, "Architecture and algorithms for an IEEE 802.11-based multi-channel wireless mesh network," in *Proc. 24th IEEE Conference on Computer Communications*, vol. 3, 2005, pp. 2223–2234.

[20] X. Yang and N. Vaidya, "Priority scheduling in wireless ad hoc networks," *Wireless Networks*, vol. 12, no. 3, pp. 273–286, 2006.

[21] A. Puri, "Optimizing traffic flow in fixed wireless networks," in *Proc. IEEE Wireless Communications and Networking Conference*, vol. 2, 2002, pp. 904–907.

[22] M. V. Marathe, A. Panconesi, and L. D. Risinger, "An experimental study of a simple, distributed edge coloring algorithm," in *Proc. 12th Annual ACM Symposium on Parallel Algorithms and Architectures*, New York, 2000, pp. 166–175.

[23] 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 *Proc. IEEE Wireless Communications and Networking Conference*, Mar. 2007, pp. 133–137.

[24] R. G. Gallager, *Discrete Stochastic Processes*. Springer, 1996.

[25] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Commun. and Mobile Computing*, vol. 2, no. 5, pp. 483–502, 2002.

Jun Hong received the B.S. and M.S. degrees in Electronic Engineering from Beijing University of Posts and Telecommunications in 2004 and Tsinghua University in 2007, respectively. She is currently a PhD candidate in the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong. Her research interest is in the general area of wireless networks. Specifically, she has been working on the effect of network state information on wireless network performance.

Victor O.K. Li (S'80 - M'81 - F'92) received SB, SM, EE and ScD degrees in Electrical Engineering and Computer Science from MIT in 1977, 1979, 1980, and 1981, respectively. He is Associate Dean of Engineering and Chair Professor of Information Engineering at the University of Hong Kong (HKU), and Guest Chair Professor of Wireless Communication and Networking at Tsinghua University, Beijing, China. He also served as Managing Director of Versitech Ltd., the technology transfer and commercial arm of HKU, and on the boards of Sunevision Holdings Ltd. and China.com Ltd. Previously, he was Professor of Electrical Engineering at the University of Southern California (USC), Los Angeles, California, USA, and Director of the USC Communication Sciences Institute. Sought by government, industry, and academic organizations, he has lectured and consulted extensively around the world. He has received numerous awards, including the PRC Ministry of Education Changjiang Chair Professorship at Tsinghua University, the UK Royal Academy of Engineering Senior Visiting Fellowship in Communications, the Croucher Foundation Senior Research Fellowship, and the Order of the Bronze Bauhinia Star, Government of the Hong Kong Special Administrative Region, China. He is a Registered Professional Engineer and a Fellow of the IAE, and the HKIE.