# Impact of Information on Network Performance – An Information-Theoretic Perspective

Jun Hong, 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*—**Available network information is an important factor in determining network performance. In this paper, we study the basic limits on the amount of network information that should be transmitted in the network to achieve a given level of network performance. From the perspective of information theory, network information is an information source, and the lower bound on network information is the minimum code letters required to encode the source. We propose a general information-theoretic framework, which can be applied to any network, to study the effect of network information on the performance of any network protocol. We also analyze the tradeoff between network performance improvement and network information collection overhead. To illustrate our approach, we use the framework to determine the lower bound on the traffic information for a simple scheduling protocol in wireless networks. The results in this paper may be used to analyze and evaluate network protocols and guide future designs.**

## I. INTRODUCTION

It is intuitively obvious that network information, such as network topology, channel state and traffic information, is an essential and important factor in network performance. If more network information is obtained, the network protocol will be more efficient and the network performance can be improved. For example, consider the hidden terminal problem [1] shown in Figure 1. Both nodes $n_1$ and $n_3$ want to send a packet to node $n_2$. Obviously, there will be a conflict if $n_1$ and $n_3$ transmit simultaneously. But if the two nodes get the scheduling information of each other, the conflict can be avoided and the network performance will be improved. However, collecting and disseminating network information may consume valuable bandwidth resource. So, the objective of our research is to investigate the relationship between network performance and network information. There are two fundamental questions,

1) How much information is required for effective decision making?
2) What is the tradeoff between network performance improvement and the communication overhead of transmitting network information?

Information theory provides a way of measuring information and the ultimate limits of communication, and may help us investigate the relationship between information and network performance. We find that the major purpose of rate distortion theory is to find the lower bound on the binary digits required to encode a source so that it can be reconstructed to achieve a given fidelity criterion [2]. The basic idea of our research is that the distortion in rate distortion theory can be extended to network performance measures, such as network capacity and packet loss rate, and the minimum quantity of information required to achieve a given level of network performance can be found by applying rate distortion theory. The major contribution of our work is an information-theoretic framework, which is generally applicable to any network (and to any resource allocation system), to formulate the quantitative relationship between network information and network performance. We demonstrate this framework using a wireless network example, to obtain the lower bound on the amount of information required to achieve certain network performance. We also consider the communication overhead of collecting network information, and analyze its impact on the network performance. Although in this example we only focus on the scheduling performance in wireless networks, our approach can also be applied to wired networks, and to measuring the effect of information on the performance of other network protocols.

The rest of the paper is organized as follows. Section II presents the state of the art. Section III defines the problem and presents a framework based on rate distortion theory, with which the relationship between information and network performance can be analyzed. Section IV illustrates the framework by studying the effect of information on wireless network performance. Section V computes the net network data rate. Section VI concludes the paper.

## II. RELATED WORK

The first work to study the required network information theoretically can be found in [3], where rate distortion theory is used to find limits on 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 overhead required to restrict the information error within a given threshold [4]–[6]. Our work is fundamentally different in that, through the development of a performance degradation function, we have developed an information-theoretic framework which is generally applicable to any network, and to any network protocols. By choosing a proper performance degradation function, our framework can be used to study the limits on protocol information and overhead in [4]–[6].
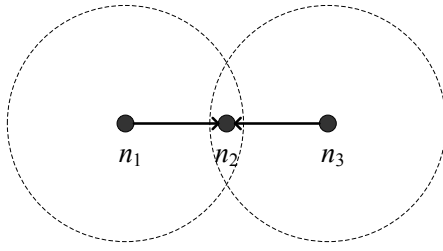
Fig. 1. Hidden Terminal Problem

There is also a body of work on "network information theory" as exemplified by [7]–[9], but it focuses on the capacity region of the network and the distributed source coding strategies, and is not directly related to our work.

Another body of work focuses on estimating the performance of wireless networks. Gupta and Kumar first determine the capacity of wireless networks [10]. Franceschetti *et al.* close the gap between the capacity upper and lower bounds in Gupta and Kumar's original results [11]. Zhang and Hou extend this work to networks with unlimited bandwidth [12]. Some researchers analyze the impact of interference on multi-hop wireless network performance. Jain *et al.* use a conflict graph to model the wireless interference and compute upper and lower bounds on the network throughput [13]. Kodialam and Nandagopal analyze the effect of interference on the achievable data rate in wireless networks [14]. Kolar and Abu-Ghazaleh study the performance of globally aware routing which is cognizant of the wireless interference [15]. The scheduling effects on wireless network performance have also been studied. Garetto *et al.* use a Markovian model to estimate the effects of scheduling on the throughput of CSMA channels [16]. Kolar and Abu-Ghazaleh evaluate the scheduling interactions among several given links and analyze the scheduling effects on network capacity [17].

We observe that the wireless network capacity depends not only on the wireless interference or 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 conditions, and ignores available network information. In our previous research, we have analyzed the effect of information on the performance of a particular scheduling algorithm [18]. In this paper, instead of focusing on a specific application, we develop a general information-theoretic framework, which can be applied to any network, to study the effect of information on the performance of any network protocol.

## III. INFORMATION-THEORETIC FRAMEWORK

In this section, we build an information-theoretic framework, with which the quantitative relationship between the network information and network performance can be formulated. We extend the concept of distortion in rate distortion theory to network performance measures, and take the network information as an information source. Then the minimum information required to achieve a given network performance

can be found by minimizing the code letters required to encode the source. Table I compares the original terms used in rate distortion theory and the terms used in our approach.

Suppose the complete network information is $X$, and the estimated network information is $\hat{X}$. If $\hat{X} = X$, i.e. the complete network information is obtained, naturally, the decisions can be correct and the network performance is optimal. If $\hat{X} \neq X$, i.e. the network information is incorrectly estimated, the decisions based on the inaccurate information may be inappropriate, which will lead to degradation of the network performance. In reality, available network information is mostly imprecise and incomplete. In such cases, we may demand that the degradation of the network performance is within a given threshold $D$.

We define a network performance degradation measure $d$ as a function of $X$ and $\hat{X}$, to describe the impact, due to the difference between the complete network information $X$ and its estimation $\hat{X}$, on the network performance. The value of $d_{X\hat{X}}(x, \hat{x})$ is a measure of the cost of estimating $x$ as $\hat{x}$, where $x$ and $\hat{x}$ are instances of $X$ and $\hat{X}$, respectively. Let the function $O(\hat{x})$ represent the network performance when the collected network information $\hat{X}$ is equal to $\hat{x}$, then

$$d_{X\hat{X}}(x, \hat{x}) = O(x) - O(\hat{x})$$

Suppose $P_{X\hat{X}}(x, \hat{x})$ is a joint probability on $X$ and $\hat{X}$, then the average performance degradation $\bar{D}$ is given by,

$$\bar{D} = \sum_{\hat{X}} \sum_{X} P_{X\hat{X}}(x, \hat{x}) d_{X\hat{X}}(x, \hat{x})$$

The degradation threshold $D$ is specified as a maximum tolerable value for the average degradation $\bar{D}$. So, there is a constraint on $P_{X\hat{X}}(x, \hat{x})$,

$$\sum_{\hat{X}} \sum_{X} P_{X\hat{X}}(x, \hat{x}) d_{X\hat{X}}(x, \hat{x}) \leq D$$

Depending on the performance of interest, the degradation measure $d_{X\hat{X}}(x, \hat{x})$ can be negative. For example, the degradation measure will be negative when we are interested in network delay or packet loss rate. In such cases, the degradation threshold $D$, which should be non-positive, is specified as the minimum value for $\bar{D}$, and the constraint on $P_{X\hat{X}}(x, \hat{x})$ is,

$$\sum_{\hat{X}} \sum_{X} P_{X\hat{X}}(x, \hat{x}) d_{X\hat{X}}(x, \hat{x}) \geq D^1$$

The mutual information between $X$ and $\hat{X}$ is,

$$I\left(X; \hat{X}\right) = \sum_{\hat{X}} \sum_{X} P_{X\hat{X}}(x, \hat{x}) \log_2 \frac{P_{X\hat{X}}(x, \hat{x})}{P_X(x) P_{\hat{X}}(\hat{x})}$$

Hence, we can get the minimum amount of required information by minimizing $I\left(X; \hat{X}\right)$ over all joint probabilities subject to the constraint that $\bar{D} \leq D$,

[1]Without loss of generality, we will assume the degradation measure is non-negative and use the constraint $\bar{D} \leq D$ in the rest of the paper.

TABLE I
RATE DISTORTION THEORY VS. OUR APPROACH

| Rate Distortion Theory | Our Approach |
|---|---|
| Information source | Complete network information |
| Distortion | Network performance degradation |
| Minimum required code letters | Minimum required network information |

$$R(D) = \min_{P_{X\hat{X}}(x,\hat{x}):\bar{D}\leq D} I\left(X;\hat{X}\right)$$

$R(D)$, known as the rate distortion function, is a lower bound on the information required to guarantee that the network performance degradation is less than or equal to $D$. $R(D)$ can be obtained by solving the optimization problem. Suppose the network performance is $O$ when the complete network information is obtained, then when collecting network information of $R(D)$ bits, the network performance can achieve $O - D$. In the following sections, we will illustrate this framework with wireless networks.

## IV. ANALYSIS OF WIRELESS NETWORKS

In this section, we will illustrate the framework by analyzing the effect of traffic information on wireless network performance. Suppose there is a central controller which maintains the topology of the network. The channel access protocol is assumed to be Time Division Multiple Access (TDMA). The sender of each link informs the controller whether there is a packet waiting to be sent on this link in each time slot. Then the controller performs scheduling based on the obtained traffic information, and distributes the scheduling results to the nodes. In this part, we focus on the effect of *traffic information* on wireless network performance. It does not mean that traffic information is the only required information. Here we assume that other control information is available, and the aim of this assumption is to separate the effect of traffic information from other phenomena.

Let $N$ be the number of links in the network, $\mathbb{Y}^{(N)} = (Y_1, Y_2, \cdots, Y_N)$ be the complete traffic information, and $\hat{\mathbb{Y}}^{(N)} = \left(\hat{Y}_1, \hat{Y}_2, \cdots, \hat{Y}_N\right)$ be the estimated traffic information, where $Y_i$ indicates whether there is a packet waiting to be sent on the $i$-th link,

$$Y_i = \begin{cases} 1 & \text{link } i \text{ has a packet to send;} \\ 0 & \text{otherwise.} \end{cases}$$

and $\hat{Y}_i$ is the traffic information of the $i$-th link obtained by the controller.

Based on the collected information $\hat{\mathbb{Y}}^{(N)}$, the central controller performs scheduling according to the network topology and a given scheduling algorithm. For any given network topology and scheduling algorithm, there exists a function $Z$ which describes the relationship between the scheduling result and the obtained network information. Let

$\mathbf{y} = (y_1, y_2, \cdots, y_N)$ and $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_N)$ be instances of $\mathbb{Y}^{(N)}$ and $\hat{\mathbb{Y}}^{(N)}$, respectively, and $\mathbb{Z}^{(N)}(\hat{\mathbf{y}}) = (Z_1(\hat{\mathbf{y}}), Z_2(\hat{\mathbf{y}}), \cdots, Z_N(\hat{\mathbf{y}}))$ represent the scheduling result given that the estimated information $\hat{\mathbb{Y}}^{(N)}$ is equal to $\hat{\mathbf{y}}$, where $Z_i(\hat{\mathbf{y}})$ denotes the scheduling decision for the $i$-th link, i.e.,

$$Z_i(\hat{\mathbf{y}}) = \begin{cases} 1 & \text{link } i \text{ transmits;} \\ 0 & \text{link } i \text{ does not transmit.} \end{cases}$$

Note that we can measure performance degradation using different metrics, such as network throughput, transmission delay, packet loss rate, and so on. Here, we take the network throughput $G$, which is defined as the average number of successful transmissions per time slot, as an example. Therefore, the performance degradation measure $d_{\mathbb{Y}^{(N)}\hat{\mathbb{Y}}^{(N)}}(\mathbf{y},\hat{\mathbf{y}})$ is defined as,

$$d_{\mathbb{Y}^{(N)}\hat{\mathbb{Y}}^{(N)}}(\mathbf{y},\hat{\mathbf{y}}) = \sum_{i=1}^{N} Z_i(\mathbf{y}) y_i - \sum_{i=1}^{N} Z_i(\hat{\mathbf{y}}) y_i \quad (1)$$

The first term on the right side of equation (1) is the number of successful transmissions when the network information $\mathbf{y}$ is estimated correctly, and the second term is the achievable network throughput when the information $\mathbf{y}$ is estimated as $\hat{\mathbf{y}}$.

Suppose $P_{\mathbb{Y}\hat{\mathbb{Y}}}^{(N)}(\mathbf{y},\hat{\mathbf{y}})$ is a joint probability measure on $\mathbb{Y}^{(N)}$ and $\hat{\mathbb{Y}}^{(N)}$, then the constraint on $P_{\mathbb{Y}\hat{\mathbb{Y}}}^{(N)}(\mathbf{y},\hat{\mathbf{y}})$ is,

$$\sum_{\hat{\mathbb{Y}}^{(N)}} \sum_{\mathbb{Y}^{(N)}} P_{\mathbb{Y}\hat{\mathbb{Y}}}^{(N)}(\mathbf{y},\hat{\mathbf{y}}) d_{\mathbb{Y}^{(N)}\hat{\mathbb{Y}}^{(N)}}(\mathbf{y},\hat{\mathbf{y}}) \leq D$$

and the minimum amount of required information per time slot is

$$R(D) = \min_{P_{\mathbb{Y}\hat{\mathbb{Y}}}^{(N)}(\mathbf{y},\hat{\mathbf{y}}):\bar{D}\leq D} I\left(\mathbb{Y}^{(N)};\hat{\mathbb{Y}}^{(N)}\right)$$

Note that this result is based on the assumption that other control information is completely obtained, and the "network information" mentioned here is the traffic information.

Now we will study how the network throughput increases with the traffic information by analyzing a simple "network" consisting of two links. Although the network is simple and perhaps unrealistic, it provides insight into the problem, and results of larger and more realistic networks can be derived by applying the same procedure.

It is assumed that the two links $L_1$ and $L_2$ interfere with each other, i.e. only one link can be active in each time slot. Otherwise, there will be collisions. Suppose $Y_1$ and $Y_2$ are independent identically distributed variables which follow the
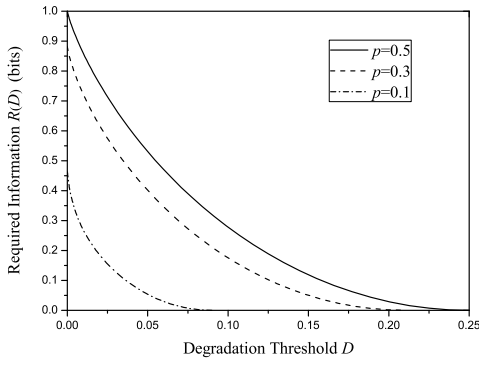
Fig. 2.  Minimum Required Traffic Information as Functions of $D$ and $p$

Bernoulli $(p)$ distribution, i.e. $\Pr(Y_1 = 1) = \Pr(Y_2 = 1) = p$, and $\Pr(Y_1 = 0) = \Pr(Y_2 = 0) = q$. At each time slot, the controller makes decision according to the following scheduling rules:

1) If $\hat{Y}_1 = 0$, and $\hat{Y}_2 = 0$, the time slot is assigned to $L_1$.
2) If $\hat{Y}_1 = 0$, and $\hat{Y}_2 = 1$, the time slot is assigned to $L_2$.
3) If $\hat{Y}_1 = 1$, and $\hat{Y}_2 = 0$, the time slot is assigned to $L_1$.
4) If $\hat{Y}_1 = 1$, and $\hat{Y}_2 = 1$, the time slot is assigned to $L_2$.

Therefore, the function $\mathbb{Z}^{(2)}$, which produces the scheduling result, is defined as

$$\mathbb{Z}^{(2)}\left((0,0)\right) = \mathbb{Z}^{(2)}\left((1,0)\right) = (1,0)$$
$$\mathbb{Z}^{(2)}\left((0,1)\right) = \mathbb{Z}^{(2)}\left((1,1)\right) = (0,1) \quad (2)$$

When $\mathbb{Y}^{(2)} \neq (0,0)$, if the traffic information is correctly estimated, there will be one successful transmission in each time slot; otherwise, the time slot may be assigned to a link which does not have any packets to send, thereby wasting a time slot and leading to performance degradation. Then the degradation measure $d_{\mathbb{Y}^{(2)}\hat{\mathbb{Y}}^{(2)}}(\mathbf{y}, \hat{\mathbf{y}})$ is given by,

$$
\left[d_{\mathbb{Y}^{(2)}\hat{\mathbb{Y}}^{(2)}}(\mathbf{y}, \hat{\mathbf{y}})\right] =
\begin{array}{c}
\\
00 \\ 01 \\ 10 \\ 11
\end{array}
\begin{array}{cccc}
00 & 01 & 10 & 11 \quad \leftarrow \hat{\mathbf{y}} \\
\begin{bmatrix}
0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
$$
$$\begin{array}{c}\uparrow \\ \mathbf{y}\end{array}$$

So, the average performance degradation is

$$\bar{D} = \sum_{\mathbf{y}=(0,0)}^{(1,1)} \sum_{\hat{\mathbf{y}}=(0,0)}^{(1,1)} P_{\mathbb{Y}\hat{\mathbb{Y}}}^{(2)}(\mathbf{y}, \hat{\mathbf{y}}) d_{\mathbb{Y}^{(2)}\hat{\mathbb{Y}}^{(2)}}(\mathbf{y}, \hat{\mathbf{y}})$$

Since $Y_1$ and $Y_2$ are independent, and the messages are sent to the scheduler independently, we can get

$$P_{\mathbb{Y}\hat{\mathbb{Y}}}^{(2)}(\mathbf{y}, \hat{\mathbf{y}}) = P_{Y_1}(y_1) P_{\hat{Y}_1|Y_1}(\hat{y}_1|y_1) P_{Y_2}(y_2) P_{\hat{Y}_2|Y_2}(\hat{y}_2|y_2)$$

Suppose the transition probabilities $P_{\hat{Y}_i|Y_i}(0|0) = \alpha_i$ and $P_{\hat{Y}_i|Y_i}(1|1) = \beta_i$, then

$$\bar{D} = pq(2 - \alpha_2 - \beta_2) \leq D \quad (3)$$

Note that $\bar{D}$ depends only on $P_{\hat{Y}_2|Y_2}$. This is because according to equation (2), the scheduling result is based on $\hat{Y}_2$, and the information of $Y_1$ will not affect the network performance. So. the minimum quantity of required information is

$$R(D) = \min_{\alpha_2,\beta_2:\, \bar{D}\leq D} I\left(Y_2; \hat{Y}_2\right) \quad (4)$$

Figure 2 shows the numerical results of the rate distortion function $R(D)$. It is obvious that $R(D)$ is nonincreasing with the degradation threshold $D$. When $p$, which indicates the probability that there is a packet on the link, is no more than 0.5, $R(D)$ increases with $p$. This is because the entropy of the information increases as $p$ increases. Note that this result is based on the given scheduling algorithm. If the scheduling algorithm is changed, the performance degradation measure will be different, and the tradeoff between network protocol information and network performance may become different.

## V. Analysis of Net Data Rate

We have analyzed how network throughput varies with available network information. It is obvious that the network throughput improves when the scheduler collects more information. However, we have not considered the overhead of collecting network information. In fact, collecting information consumes bandwidth resource, which may affect the network data rate. Since the network throughput, defined in Section IV, is the rate of transmitting all types of data (including the effective data and the network information), we need to find the net network data rate, i.e., the effective data that can be carried over the network per time slot.

Suppose the number of bits in each time slot is $\mu$. In each time slot, $R(D)$ bits is consumed to achieve a network data rate of $\mu(G_o - D)$ bits per time slot. Therefore, the effective data transmitted in each time slot (i.e. the net data rate) is

$$G_N = \mu(G_o - D) - R(D) \quad \text{bits/time slot} \quad (5)$$

It can be seen from (5) that when $R(D)$ becomes larger, i.e. the scheduler gets more information, $\mu(G_o - D)$ increases, but it is unclear whether the net network data rate $G_N$ will increase or not. A properly chosen information quantity can maximize the net network data rate. Figure 3 plots the net data rate $G_N$ against the network information $R(D)$, where $p$ is set to 0.5, and $\mu$ is set to 2, 5, 10 and 20, respectively. We can see that when $\mu$ is small ($\mu = 2, 5$), collecting information occupies a large part of the bandwidth resource, and more information does not mean better performance. When $\mu$ becomes larger ($\mu = 20$), the net network data rate increases with the network information steadily. This is because transmitting network information only takes a small part of the network capacity.

## VI. Conclusions and Future Work

In this paper, we have developed an information-theoretic framework to model the relationship between network information and network performance, and the minimum quantity of information required for a given network performance is derived. Our proposed framework is generally applicable to

(a) $\mu = 2$ bits



(b) $\mu = 5$ bits



(c) $\mu = 10$ bits
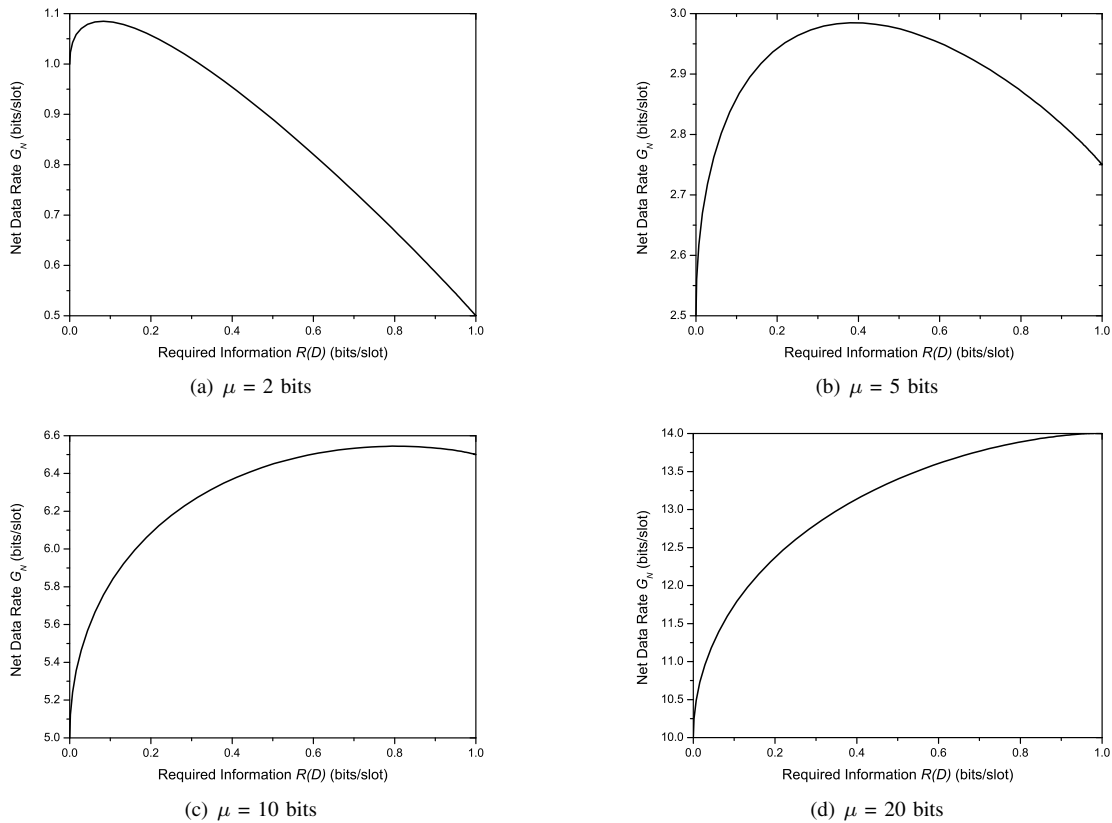


(d) $\mu = 20$ bits

Fig. 3.   Net Data Rate

any network, to study the effect of network information on the performance of any network protocol. As part of our future work, we intend to consider the details of implementation, and to develop specific strategies that approximate the performance limits derived in this work.

## ACKNOWLEDGMENT

## REFERENCES

[1] F. A. Tobagi and L. Kleinrock, "Packet switching in radio channels: Part ii - the hidden terminal problem in carrier sense multiple-acces and the busy-tone solution," *IEEE Transactions on Communications*, vol. 23, no. 12, pp. 1417–1433, 1975.

[2] R. G. Gallager, *Information Theory and Reliable Communication*. New York, Wiley, 1968.

[3] ——, "Basic limits on protocol information in data communication networks," *IEEE Transactions on Information Theory*, vol. 22, no. 4, pp. 385–398, 1976.

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

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

[6] D. Wang and A. Abouzeid, "Link State Routing Overhead in Mobile Ad Hoc Networks: A Rate-Distortion Formulation," in *the 27th IEEE International Conference on Computer Communications (INFOCOM 2008)*, 2008, pp. 1337–1345.

[7] T. Cover and J. Thomas, *Elements of Information Theory*. New York, Wiley, 2006.

[8] A. El Gamal and T. M. Cover, "Multiple user information theory," in *Proceedings of the IEEE*, vol. 68, no. 12, 1980, pp. 1466–1483.

[9] L. L. Xie and P. R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation," *Information Theory, IEEE Transactions on*, vol. 50, no. 5, pp. 748–767, 2004.

[10] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[11] 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.

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

[13] 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)*, 2003, pp. 66–80.

[14] M. Kodialam and T. Nandagopal, "The effect of interference on the capacity of multi-hop wireless networks," in *International Symposium on Information Theory (ISIT 2004)*, 2004, p. 472.

[15] V. Kolar and N. Abu-Ghazaleh, "A multi-commodity flow approach for globally aware routing in multi-hop wireless networks," in *the 4th Annual IEEE International Conference on Pervasive Computing and Communications (PerCom 2006)*, 2006.

[16] 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.

[17] V. Kolar and N. Abu-Ghazaleh, "The effect of scheduling on link capacity in multi-hop wireless networks," *Arxiv preprint cs.NI/0608077*, 2006.

[18] J. Hong and V. O. K. Li, "The effect of information on scheduling performance in multi-hop wireless networks," in *the 19th IEEE International*

*Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2008)*, 2008.