Exploiting Spatial Diversity in Rate Adaptive WLANs with Relay Infrastructure

Aaron So and Ben Liang
Department of Electrical and Computer Engineering
University of Toronto, Toronto, Ontario, Canada
{aaronso, liang}@comm.utoronto.ca

Abstract— **The throughput capacity of a Wireless Local Area Network (WLAN) can be improved synergically by 1) the multi-rate capability of modern WLAN equipments and 2) the spatial diversity provided by its relay infrastructure. In this work, we investigate the effect of multi-path fading and opportunistic utilization of a fixed number of immobile relay nodes on the throughput capacity of a rate adaptive WLAN. We develop an analytical framework that computes the throughput capacity of an IEEE 802.11 WLAN with relay infrastructure in the Rayleigh fading environment. We compare the performance of MAC-layer and network-layer relaying. Our results show that up to 200% performance gain can be achieved by an optimal relay infrastructure over a network with no relay. Furthermore, for a wide range of system parameters, optimally placed relay nodes can significantly increase the network throughput capacity.**

I. Introduction

Wireless Local Area Networks (WLANs), which provide low-cost wireless broadband data access for mobile Internet users, are expected to create a plethora of business opportunities. Currently, the most commonly implemented WLANs in North America are based on the IEEE 802.11b standard, which is capable of supporting bit rates up to 11Mbps in the 2.4GHz spectrum. As the number of hotspot users proliferates and the demand from wireless Internet users increases, new strategies have to be employed to increase the throughput of future WLANs. The IEEE 802.11a standard [1], which is capable of supporting bit rates up to 54Mbps in the 5GHz spectrum, is a promising technology to improve hotspot throughput. However, due to the higher frequency, signal attenuates more severely than that of lower frequency systems. As a result, relay nodes can play an important role in large 802.11a networks. In general, as future wireless communications systems migrate to higher frequencies for more spectrum, relay infrastructure will become even more crucial.

The multi-rate capability of modern WLAN equipments and spatial diversity provided by relay infrastructures can synergically improve the throughput capacity of a WLAN. In a WLAN with relay infrastructure, the source can either transmit its data to the destination directly, or relay its data via the relay node. Since WLANs are usually located in a multi-path fading environment, the qualities of wireless links vary with time. Depending on the instantaneous channel condition, if a circuitous route can result in a higher bit rate than the direct route, the source should use the relay node to relay its data. In this work, an analytical model has been developed to understand the benefits of MAC-layer relaying in a multi-rate WLAN with relay infrastructure. Moreover, the optimal locations of the relay nodes which maximize the throughput capacity of a network with respect to different system parameters are numerically derived. Furthermore, the effects of network-layer relaying can be studied by using a slightly modified version of the MAC-layer model. Our results show that up to 200% performance gain can be achieved by an optimal relay infrastructure over a network with no relays.

This paper is organized as follows. In Section II we review recent related work in multihop networks. In Section III we explain our system model and discuss the objective of our design. In Section IV, the channel model and its relationship with transmit bit rate is discussed. In Section V, we explain the relay strategy and its relation with packet transmission time. In Section VI, the optimal placement of relay nodes in a two dimensional single AP scenarios is discussed. In Section VII, we discuss the numerical results obtained from our calculations. Finally, in Section VIII, we give conclusions.

II. Related Work

Inspired by recent advances in ad hoc networking [2], the concept of using peer mobile hosts to relay data has been explored in the context of cellular networks [3]. A relay routing protocol which emphasizes on capacity maximization has been proposed in [4]. In a more recent work, the problem of joint routing, link scheduling and power control in such multihop networks has been investigated [5]. Moreover, issues about frequency assignment and frequency recycling in such multihop networks have been addressed in [6].

The concept of using immobile relay nodes to relay traffic, e.g., [7], has received less attention. Immobile relay nodes have several advantages when compare with mobile relay nodes. First, because of their sedentariness, it is reasonable to assume that they have access to power supply. Consequently, energy is not a constraint. Second, the fixed relay nodes can be optimally configured to maximize their beneficial effects.

In [8], the expected throughput capacity of a wireless network with relay infrastructure and the optimal placement of the relay nodes were investigated based on the wireless relay channel capacity derived in [9] and [10]. In [11], an efficient extension point placement algorithm aiming at improving the throughput of a rectilineal network through *network-layer* relaying has been proposed. In this work, we focus on understanding the effect of opportunistic *MAC-layer* relaying and relay infrastructure configuration on a rate adaptive WLAN in a multipath fading environment. In particular, we explore the spatial diversity provided by the relay infrastructure in a discrete multi-rate WLAN in the Rayleigh fading environment, and numerically derive the optimal placement of relay nodes in such WLANs.

III. System Model and Design Objective

The system that we investigate consists an access point which is connected to the wired network and provides wireless coverage to an area. Mobile Hosts (MHs) locating in this coverage area wirelessly communicate with the wired network via this access point. In this study, we assume the transmission schedule for all transmitters is completed perfectly by the AP, and communications are performed in a poll-and-response fashion. At any given time, only one transmitter is allowed to transmit, so that no interference is experienced at a receiver, and no collision can occur.

The time axis is divided into time varying cycles. The AP communicates with each MH in its coverage area in a round robin fashion. In each cycle, the AP transmits a downlink packet to the chosen MH, and the MH transmits a uplink packet to the AP. For both uplink and downlink communications, a relay node (RN) can be used to forward the packets if taking such circuitous route can curtail the overall packet transaction time. In this study, we assume the lengths of the uplink and downlink packets are fixed, and the AP always has a packet to send to each MH and vice versa.

Let *x* be the total number of bits of an uplink and a downlink packet combine. Let T_i represents the packet transaction time of an AP-MH pair in the i^{th} cycle. Thus, T_i is a sequence of independent and identically distributed random variables. By the Law of Large Numbers, the throughput capacity of the network is

$$C = \lim_{n\to\infty} \frac{nx}{\sum_{i=1}^{n} T_i} = \frac{x}{E[T_i]} \quad . \tag{1}$$

In order to maximize the throughput capacity of the network, $E[T_i]$ has to be minimized. Thus, the design objective of our system is to minimize the expected time that an AP-MH pair completes a single downlink and uplink packet exchange, which we called expected packet transaction time in this paper.

IV. Channel Model and Transmit Bit Rates

The packet transaction time is inversely proportional to the communication rate between a pair of transmitter and receiver. Since different random network parameters such as distance between transmitter and receiver and channel fading affect the reliable communication rate between a transmitter and a receiver, the instantaneous bit rate supported by a transmitter receiver pair, and thus the packet transaction time, is random. In this paper, the signal strength of a transmitted signal is attenuated according to a propagation path loss model [12] and the received signal is subject to Rayleigh fading. For a particular received signal strength, the appropriate bit rate is determined by the IEEE 802.11a standard [1].

A. Propagation Path Loss Model

In this work, the following propagation path loss model is used [12].

$$P_b = \frac{P_a}{d_b^{\alpha}} \quad . \tag{2}$$

where P_a and P_b are the signal power measured at unit distance and d_b meters away from the transmitter respectively, and α is a positive constant representing the path loss roll off factor. The *reference power* P_a can be obtained via field measurement at distance one meter away from the transmitter or calculated using the following free space path loss formula [12].

$$P_a = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 L}, \tag{3}$$

where P_t is the transmitted power, G_t is the transmitter antenna gain, G_r is the receiver antenna gain, λ is the wavelength of the transmitted signal, and L ($L \geq 1$) is the loss factor not related to propagation. In this work, the parameters in (3) are obtained from the CISCO Aironet 1200 series access point specifications [13], and L is set to 1. In the next subsection, we describe how fading affects the reliable communication rate between a transmitter receiver pair.

B. IEEE 802.11a under Rayleigh Fading Channel

Suppose M data rates, which are denoted as $r_1, r_2, ..., r_M$, are supported by the physical layer. Reliable communication by using rate r_m can be realized only if the signal strength at the receiver is above a certain threshold, say η_m. Consequently, for the set of M data rates, there is a set of M thresholds, $\eta_1, ..., \eta_M$. We further define $\eta_0 = 0$ and $\eta_{M+1} = \infty$. For IEEE 802.11a, there are 8 supported bit rates, and the corresponding thresholds are specified in the standard [1].

If the transmitter reference power is P, while the distance between the transmitter and the receiver is l, where $l > 1$, the average received signal power, P_r, is given by (2). Then, according to the Rayleigh fading channel model, the probability of receiving a signal with instantaneous power, γ, is exponentially distributed with pdf $p(\gamma) = \frac{1}{P_r} e^{\frac{-\gamma}{P_r}}$.

Since the network can only support a finite number of bit rates, the transmitter bit rate is a discrete random variable, which is denoted by $R(P, l)$ in this paper. The probability that the transmitter can transmit at rate r_m, is

$$P(R(P, l) = r_m) = \int_{\eta_m}^{\eta_{m+1}} \frac{l^{\alpha}}{P} e^{\frac{-\gamma l^{\alpha}}{P}} d\gamma. \tag{4}$$

where $m = 0, 1, 2, ..., M$ and $r_0 = 0$. $P(R(P, l) = r_0)$ is the probability that the receiver is located in a deep-fade area. In the next section, we use the above channel model to derive the packet transaction time of a transmitter receiver pair.

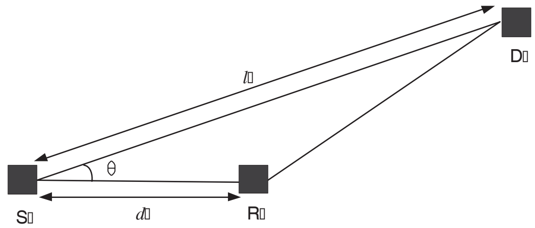

Fig. 1. Relative locations of the source(S), relay node(R) and destination(D).

V. Relay Strategy and Packet Transaction Time

Consider the single user scenario in Fig. 1. The source, S, can either transmit its packet directly to the destination, D, or relay it via the relay node, R. If the relay node is used, the packet has to be transmitted twice. Thus, the relay node may or may not be beneficial and the effect of the relay node varies depend on the channel condition. At each time instance, each link can support reliable communication at a particular rate which is described probabilistically in (4). The relay node will be employed to forward the packet if the combined S-R R-D link can achieve a higher bit rate than that of the direct S-D link. In this work, we assume the channel coherence time exceeds the packet transmission time; thus, the channel condition is essentially invariant during each packet transmission. Moreover, we also assume the distance between any two transmitters is larger than the coherence distance; thus, each link is faded independently. Before the source transmits, a small channel probing time, T_p, is required for the source to determine the channel conditions. Moreover, we assume that the source of a packet can always compute the optimal decision, and the packet can always be transmitted successfully if the channel allows. In addition to the channel probing time, a node needs a small inter-frame time, T_{if}, to change from transmit state to receive state and vice versa.

Let $T(P,l,x)$ be a discrete random variable which represents the transmission time of an x-bit packet by a transmitter with reference power P and located l meters away from the receiver excluding the probing and inter-frame time. The probability distribution of $T(P,l,x)$ is

$$P\left(T(P,l,x)=\frac{x}{r_m}\right)=P(R(P,l)=r_m). \quad (5)$$

Thus, the probability density function of this single-hop raw packet transmission time is

$$f_{T(P,l,x)}(z)=\sum_{i=0}^{M}P(R(P,l)=r_i)\delta\left(z-\frac{x}{r_i}\right). \quad (6)$$

Consider the direct link from the source to the destination. Let $T_d(P_1,l,x)$ represents the packet transmission time from S to D together with the probing and inter-frame time, where P_1 is the reference power of the source's transmitter. The pdf of this random variable is

$$f_{T_d(P_1,l,x)}(z)=f_{T(P_1,l,x)}(z-T_p-T_{if}). \quad (7)$$

Now, let $T_{SRD}(P_1,P_2,l,d,\theta,x)$ be a discrete random variable which represents the transmission time of an x-bit packet, excluding the probing and inter-frame time, transmitted by the source and forwarded to the destination via the relay node, where P_1 and P_2 are the reference powers of the source and the relay node respectively. Since the packet is transmitted twice, this two-hop raw packet transmission time is a sum of two independent random variables. I.e.,

$$T_{SRD}(P_1,P_2,l,d,\theta,x)=T(P_1,d,x)+T(P_2,e,x). \quad (8)$$

where $e=(l^2+d^2-2ld\cos\theta)^{\frac{1}{2}}$. Thus, the pdf of this two-hop raw packet transmission time is the convolution of two single-hop raw packet transmission time pdf's. I.e.,

$$f_{T_{SRD}(P_1,P_2,l,d,\theta,x)}(z)=f_{T(P_1,d,x)}(z)*f_{T(P_2,e,x)}(z) \ . \quad (9)$$

Similar to the direct case, let $T_r(P_1,P_2,l,d,\theta,x)$ be the discrete random variable represents the two-hop packet transmission time from S to D through R, together with the probing and inter-frame time. The pdf of this random variable is

$$f_{T_r(P_1,P_2,l,d,\theta,x)}(z)=f_{T_{SRD}(P_1,P_2,l,d,\theta,x)}(z-T_p-2T_{if}). \quad (10)$$

To simplify notations, we denote $T_d(P_1,l,x)$ and $T_r(P_1,P_2,l,d,\theta,x)$ by T_d and T_r respectively.

The source always takes the route (direct or circuitous) which results in the minimal packet transmission time. The resulting packet transmission time is the minimum of two random variables, i.e., $T_m=\min(T_d,T_r)$. The pdf of this random variable is

$$f_{T_m}(z)=f_{T_d}(z)[1-F_{T_r}(z)]+f_{T_r}(z)[1-F_{T_d}(z)], \quad (11)$$

where $F_{T_d}(z)$ and $F_{T_r}(z)$ are the cdf's of T_d and T_r respectively.

When the destination is located in a deep-fade area (the direct link fails and either one of the links in the two-hop route fails), transmission is not possible. We assume the source can detect such condition after channel probing, and the source will postpone its transmission. However, every time such undesired condition happens, one channel probing time is wasted and the wasted channel probing time(s) add(s) to the total packet transmission time.

Let the events of "unsatisfactory" and "acceptable" channel conditions be F and S respectively. Then the probability of F is $P[T_m=\infty]$. Let $T_t(P_1,P_2,l,d,\theta,x)$ represents the total packet transmission time taking consideration of different channel states. Again for notation simplification, let's denote $T_t(P_1,P_2,l,d,\theta,x)$ by T_t. The expected total packet transmission time can be calculated as follows.

$$\begin{aligned}E[T_t]&=E[T_t|F]\times P[T_m=\infty]+E[T_t|S]\times[1-P[T_m=\infty]]\\&=[E[T_t]+T_p]\times P[T_m=\infty]+\\&\quad\left[\frac{\int_0^{T_{max}}zf_{T_m}(z)dz}{1-P[T_m=\infty]}\right]\times[1-P[T_m=\infty]] \ .\end{aligned} \quad (12)$$

After some rearrangement,

$$E[T_t]=\frac{P[T_m=\infty]T_p+\int_0^{T_{max}}zf_{T_m}(z)dz}{1-P[T_m=\infty]} \ , \quad (13)$$

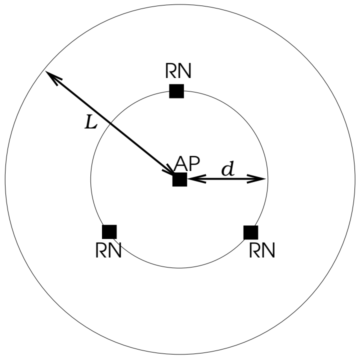

Fig. 2. Network Configuration.

where T_{max} is the largest successful packet transmission time. For comparison purposes, we have to calculate the expected total packet transmission time with the absence of the relay node. Let $T_a(P_1, l, x)$ represents the total packet transmission time when the relay node in Fig. 1 is removed. Following similar analysis,

$$E[T_a(P_1, l, x)] = \frac{P[T_d = \infty]T_p + \int_0^{T_{max}} z f_{T_d}(z)dz}{1 - P[T_d = \infty]} \ . \quad (14)$$

In the next section, we will study the scenario, where MHs are randomly distributed in a given area. Moreover, given a set of system parameters, the optimal placement of the RNs in a two dimensional network will be discussed.

VI. Optimal Placement of Relay Nodes

In this section, we consider a model as shown in Fig. 2, where a number of RNs, N, (in this example, $N = 3$) is given, and they are placed uniformly around and d meters away from the AP. The MHs are uniformly distributed within the circle with radius L centered at the AP.

The following system parameters and notations are used:

- x_d = downlink packet length (bits).
- x_u = uplink packet length (bits).
- $\beta = \frac{x_d}{x_d + x_u}$ =proportion of downlink data.
- α = path loss roll off factor.
- P_{ar} = reference power of AP or RN.
- P_m = reference power of MH. ($P_m \leq P_{ar}$)
- d = location of RN.
- l = location of a MH.
- $\theta = \angle$ MH-AP-RN
- $\phi = \angle$ AP-HM-RN = $\arctan\left(\frac{d \sin\theta}{l - d\cos\theta}\right)$
- $e = (l^2 + d^2 - 2ld\cos\theta)^{\frac{1}{2}}$

As described in Section III, the goal of the system is to minimize the expected packet transaction time of the network. To evaluate the effect of relay infrastructure in a network, we first evaluate the expected packet transaction time of the network without RNs. By using (14), the expected packet transaction time of the network is

$$\overline{T_{direct}} = \frac{1}{\pi L^2} \int_{0+\epsilon}^{L} \int_0^{2\pi} l\{E[T_a(P_{ar}, l, x_d)] + E[T_a(P_m, l, x_u)]\} dl d\theta, \quad (15)$$

where $\epsilon > 0$ is small.

We then evaluate the network expected packet transaction time with the help of relay nodes. As described in the previous section, the MH may directly communicate with the AP or use the *closest* RN to help relay its packet to and from the AP. By using (13) the network expected packet transaction time is

$$\overline{T_{relay}}(d) = \frac{N}{\pi L^2} \int_{0+\epsilon}^{L} \int_{-\frac{\pi}{N}}^{\frac{\pi}{N}} l\{E[T_t(P_{ar}, P_{ar}, l, d, \theta, x_d)] + E[T_t(P_m, P_{ar}, l, e, \phi, x_u)]\} dl d\theta. \quad (16)$$

For each set of system parameters, there exists an optimal $d^* \in (0, L)$ such that $\overline{T_{relay}}(d^*)$ is minimal. Since $\overline{T_{relay}}(d)$ is a continuous function of d, all the critical points of $\overline{T_{relay}}(d)$ can be found when we set $\overline{T_{relay}}\,'(d) = 0$. This can be solved numerically. In the next section, we discuss some observations obtained from the numerical results.

VII. Numerical Analysis

In this section, we discuss the benefit of the optimally placed RNs with respect to different system parameters. The effects of both MAC-layer and network-layer relaying were studied. When network-layer relaying is performed, the capacity of the network can be computed by the above model except a MH always uses a RN if it can curtail the *expected* packet transaction time, i.e., the decision of whether to use a RN is computed based on the MH location instead of the instantaneous channel condition. For each set of parameters, there is an optimal placement of RNs and two capacities: C_{direct} and C_{relay}. C_{relay} and C_{direct} represent the throughput capacity of the network with optimally placed relay nodes and without relay nodes respectively. The throughput capacity of the network is defined in (1). Thus, $C_{direct} = \frac{x}{\overline{T_{direct}}}$ and $C_{relay} = \frac{x}{\overline{T_{relay}(d^*)}}$, where d^* represents the optimal location for the RNs. We define the performance gain of the network as follows.

$$\frac{C_{relay} - C_{direct}}{C_{direct}} \ . \quad (17)$$

In the numerical analysis, all the parameters for the hardware are obtained from the CISCO Aironet 802.11a 1200 Series specifications [13]. These hardware parameters and the fixed system parameters we used are summarized in Table I.

Parameter	Value	Parameter	Value
AP's transmitter power	16 dBm	G_t	6 dBi
RN's transmitter power	16 dBm	G_r	6 dBi
MH's transmitter power	13 dBm	T_{if}	0.03 ms
λ	0.06 m	T_p	0.2 ms

TABLE I

Parameters for AP, RN, and MH

A. Roll-off Factor and Number of Relay Nodes

In Fig. 3, we study the effect of the roll off factor, α and the number of RNs in the network. When α is small, RNs are only used to help the distanced MHs; thus, they are placed further away from the AP. However, when α is large, the optimal RN placement depends on the marginal benefit provide by the relay

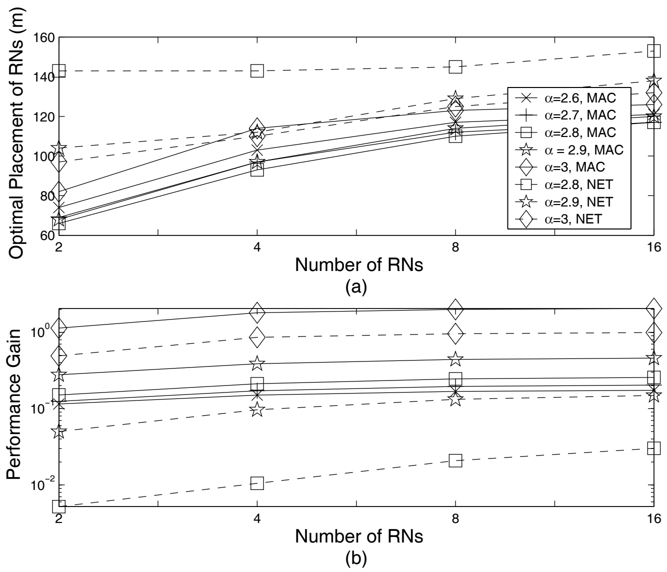

Fig. 3. Effect of roll off factor (α) and the number of RNs for $x_d = 2kbyte$, $x_u = 1kbyte$, $L = 300m$.

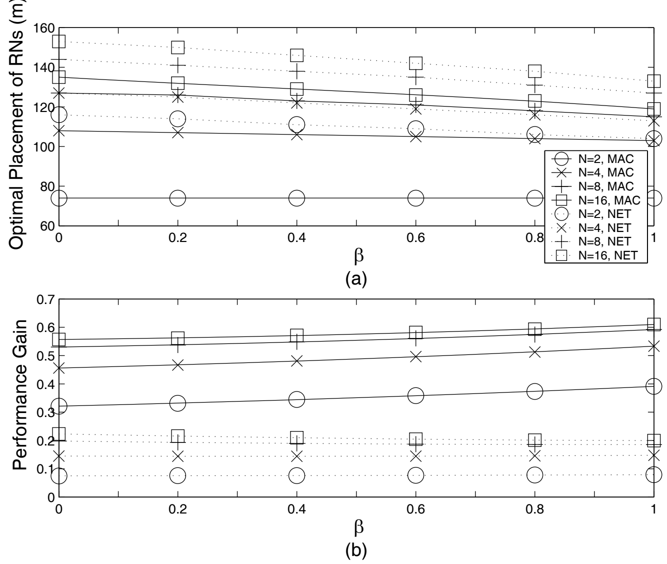

Fig. 4. Effect of proportion of downlink data and the number of RNs for $x_d + x_u = 2kbyte$, $\alpha = 2.9$, $L = 300m$.

strategy to MHs at different locations. Thus, as shown in Fig. 3a, the optimal locations of RNs computed by MAC-layer and network-layer relaying show different trends. From Fig. 3b, the performance gain increases as the roll off factor increases. This is because as the signal decades faster, the beneficial effect of a relay node that boost the signal power becomes more and more significant. When the channel attenuation is high ($\alpha = 3$), by using a small number of RNs ($N = 8$), the performance gain can be more than 200%.

B. Proportion of Downlink Data and Number of Relay Nodes

In Fig. 4, we study the effect of the proportion of downlink data and the number of RNs in the network. From Fig. 4a, the optimal RN position moves away from the AP as the proportion of downlink data, β, decreases. From Fig. 4b, when MAC-layer relaying is used, the performance gain increases as the proportion of downlink data, increases. The contrary is observed when network-layer relaying is used. This observation suggests that the benefit of opportunistic transmission is more prominent when β is large. It may be because with spatial diversity, the beneficial effects of the RN on downlink data is actually more significant than that on uplink data. Thus, when MAC-layer relaying is used, higher performance gains were observed when β is high.

VIII. Conclusion

In this work, we investigate the integration of immobile relay node to a WLAN infrastructure. We studied the benefits of spatial diversity provided by relay infrastructures on multi-rate WLANs in the Rayleigh fading environment. We have showed that in most cases, by using a few optimally placed relay nodes, the throughput capacity of the network can be improved significantly. In particular, MAC-layer relaying can significantly outperform network-layer relaying. As relay infrastructure is expected to play a more important role in future communications system, it is important to understand their impact on wireless networks.

References

[1] "IEEE standard for information technology, local and metropoliton networks - specific requirements, part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications - high-speed physical layer in the 5 GHz band," 1999.

[2] Z. Haas, J. Deng, B. Liang, P. Papadimitratos, and S. Sajama, *Wiley Encyclopedia of Telecommunications*, John Wiley & Sons, 2002.

[3] Y. Lin and Y. Hsu, "Multihop cellular: A new architecture for wireless communications," in *Proc. of IEEE INFOCOM*, Tel-Aviv, Israel, Mar. 2000, vol. 3, pp. 1273 – 1282.

[4] S. Mengesha and H. Karl, "Relay routing and scheduling for capacity improvement in cellular WLANs," in *Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt'03)*, Sophia-Antipolis, France, Mar. 2003.

[5] R. Cruz and A. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," in *Proc. of IEEE INFOCOM*, San Francisco, California, USA, Apr. 2003, vol. 1, pp. 702–711.

[6] S. Mengesha, H. Karl, and A.Wolisz, "Capacity increase of multi-hop cellular WLANs exploiting data rate adapatation and frequency recycling," Tech. Rep., Technical University Berlin Telecommunication Networks Group, 2003.

[7] H. Wu, C. Qiao, S. De, and O. Tonguz, "Integrated cellular and ad hoc relaying systems: iCAR," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 5, pp. 2105–2215, October 2001.

[8] A. So and B. Liang, "Effect of relaying on capacity improvement in wireless local area networks," in *in Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2005.

[9] M. Khojastepour, A. Sabharwal, and B. Aazhang, "On capacity of gaussian 'cheap' relay channel," in *Proc. of IEEE Global Telecommunications Conference GLOECOM'03*, San Francisco, California, USA, Dec. 2003, vol. 3, pp. 1776 – 1780.

[10] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation in wireless relay channel," *IEEE Transactions on Information Theory*, Jun. 2005.

[11] A. So and B. Liang, "An efficient algorithm for the optimal placement of wireless extension points in rectilineal wireless local area networks," in *Proc. of The Second International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QShine)*, Orlando, FL, U.S.A., Aug. 2005.

[12] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, 2 edition, 2001.

[13] "CISCO Aironet 1200 series access point: Data sheet," Tech. Rep., CISCO Systems, 2003.