

On the Optimal Design of Two-Tier Wireless Relay Networks

Vahid Pourahmadi
University of Waterloo,
200 university Ave. W.
Waterloo, ON, Canada
vpourahm@uwaterloo.ca

Shervan Fashandi
University of Waterloo,
200 university Ave. W.
Waterloo, ON, Canada
sfashand@uwaterloo.ca

Aladdin Saleh
Bell Canada,
5099 Creekbank Road,
Mississauga, ON, Canada
aladdin.saleh@bell.ca

Amir K. Khandani
University of Waterloo,
200 university Ave. W.
Waterloo, ON, Canada
khandani@uwaterloo.ca

ABSTRACT

It is known that the achievable data rate per user can be increased when relays are deployed in wireless networks. However, the drawback with this solution is that some of the network resources should be allocated to the relays. In this paper, we consider a two-tier network where all users should send/receive data in two hops (via a relay). Applying vector quantization, we approximately find the location of the relays. These approximate relays' locations are also computed analytically when the number of relays is less than six. Having the relays' locations, the network average transmission rate is evaluated in terms of a set of network parameters. Then, in the multi-dimensional space of these network parameters, we introduce the concept of *neutrality-surface*. The *neutrality-surface* is defined such that the performance of any relay network operating below this surface is inferior to that of a simple no-relay network with the same parameters. Finally, we study the relative and differential relaying gain for different network configurations.

Categories and Subject Descriptors

C.2.1 [Computer-Communication Networks]: Network Architecture and Design—*Wireless Communication* ; C.4 [Performance of Systems]: Design studies, Performance attributes

General Terms

Design, Performance, Theory

Keywords

Resource-limited network, Relay networks, Relay location, Resource assignment

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MSWiM'08, October 27–31, 2008, Vancouver, BC, Canada.
Copyright 2008 ACM 978-1-60558-235-1/08/10 ...\$5.00.

1. INTRODUCTION

Relay networks combine the advantageous features of two radically different design strategies: hierarchical structures and Ad hoc networks. Each of these two opposing architectures has its own pros and cons. For instance, Ad hoc networks show good reliability and scalability. However, they usually require complicated routing algorithms, and are vulnerable to security flaws. Complicated base stations and high signaling overhead are among the disadvantages of hierarchical structures. On the positive side, hierarchical networks are able to work with inexpensive and simple terminals. Relay networks propose a balance between these two extremes.

Relay nodes are used in wireless networks to extend the coverage or enhance the channel quality of the stations. Since relay deployment is a cost-effective solution for many applications, several standardization committees, including IEEE 802.16j, work on adding the relaying functionality to their current standard. For instance, IEEE 802.16j adopts a two-hop network connection, i.e., Source \rightarrow Relay \rightarrow Destination, without a direct link between the source and the destination [7, 14]. IEEE 802.16j relaying schemes are designed such that the relay stations stay transparent to the end users, and let the conventional users connect to the relay network without any modification in their MAC/PHY layers.

Several studies have been conducted to improve the network performance using relays. Cooperative data relaying is studied in [3] and [9]. A multihop relay network is introduced in [8], while [13] discusses intra-cell handover management schemes. Furthermore, many researchers have investigated the efficient scheduling and node assignment techniques in relay networks [4, 14]. In addition, reference [10] evaluates the performance of different feedback schemes in relay networks.

Another interesting direction is to analyze the effect of relaying in a *resource limited networks*. More precisely, as relays are deployed, two opposite effects can be identified. On the one hand, average user transmission rate increases since users transmit to a closer destination and typically experience higher Signal to Noise Ratio (*SNR*) levels. On the other hand, to keep relays in the stable condition (no

buffer overflow), relays must backhaul the aggregated users' traffic to the base-station. This data retransmission takes some of the network resources and degrades the network functionality.

In this work, the network average transmission rate is defined as the metric of the network performance. Relying on this metric, performance of a relay network depends on the following six parameters: the network range (cell size), network scheduler, ratio of the network resource assigned to the users, number of relays, location of the relays, and users transmission power. Given the cell size and the number of relays, we propose a (suboptimum) algorithm based on vector quantization to compute the relays' locations. Having the relays' locations and assuming a fair network scheduler, performance of the network is analyzed with respect to the four remaining parameters. Furthermore, for the cases where the relay count is less than six, this analysis provides a mathematical tool for the network design. Using this tool, the designer can find the proper values of these four parameters based on the design constraints. For instance, these constraints can impose conditions on the cell size, user power, and ratio of the network resources assigned to the users. Study of such conditions are out of the scope of this work.

This paper is organized as follows. First, section 2 describes the network model and the mathematical notations. Then, in section 3, we propose a scheme to estimate the relays' locations. Section 4 presents evaluation of the average transmission rate for both of the relay and no-relay networks. Numerical results and network performance analysis is presented in section 5. Finally, section 6 concludes the paper.

2. NETWORK MODEL

2.1 Network Topology and Channel Model

The network infrastructure is assumed to be a two-tier network [15], where each of the user stations transmits to one of the relays in the relay tier. Then, the relay retransmits the data to the BS. In other words, no direct connection is assumed between the users and the BS. The coverage area of the network is assumed to be bounded within a circle of radius R . User stations are also assumed to be placed uniformly in the network [11]. Thus, using continuous approximation, we can consider a density of ' η ' for the users such that the total number of users in the network is equal to $\eta\pi R^2$.

Depending on the channel condition, each user adapts its modulation scheme. In practice, a discrete set of modulation points is used for adaptive coding and modulation. However, in this work, we approximate the achievable rate of a user with the upper-bound given by the Shannon formula [5]. Thus, if the channel is modeled as a shadowing environment [12], the maximum transmission rate per unit of bandwidth for the link between the user u_i to the relay r_j would be

$$C(d_{u_i r_j}) = \mathbb{E}_{G_{u_i r_j}} \left\{ \log_2 \left(1 + \frac{P G_{u_i r_j}}{N_0} \right) \right\}, \quad (1)$$

where P and N_0 are the users transmission power and the noise power, respectively. The power gain, $G_{u_i r_j}$ is given by $G_{u_i r_j} = S \frac{1}{d_{u_i r_j}^\alpha}$. Here, $d_{u_i r_j}$ represents the distance between the user u_i and its assigned relay, r_j . α is the path

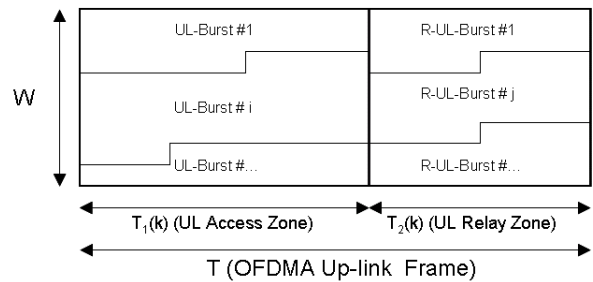


Figure 1: OFDMA Up-link Frame Structure

loss exponent $2 \leq \alpha \leq 5$, and the shadowing factor S has a lognormal distribution with a mean of 0 dB and the variance of $\sigma = 8$ dB [12], i.e., $S = 10^{\frac{Z}{10}}$ where Z is a normal random variable with the mean of zero and standard deviation of 8dB. In the case that there is no relay in the network, equation (1) reduces to

$$C(d_{u_i B}) = \mathbb{E}_S \left\{ \log_2 \left(1 + \frac{P S / d_{u_i B}^\alpha}{N_0} \right) \right\}, \quad (2)$$

where S and $d_{u_i B}$ represent the shadowing factor and the distance between the i^{th} user and the BS, respectively. In the remainder of the paper, we assume that the path loss exponent (α) is equal to 4.

2.2 Frame Structure and Mathematical Representation

In this study, we consider that the network is operating in Time Division Duplex (TDD) mode and Orthogonal Frequency Division Multiple Access (OFDMA) is assumed as the network medium access sharing mechanism [7]. The frame structure is also similar to the frame structure proposed for IEEE 802.16j [7].

Focusing on the Uplink transmission, T denotes the total duration of the uplink frame. Indeed, as Fig. 1 represents, the Uplink frame of each transmission block, say k , is divided into two parts. In the first $T_1(k)$ seconds (**UL Access Zone**), the users transmit to the relay nodes and for the next $T_2(k)$ seconds (**UL Relay Zone**), relays retransmit the aggregated data to the BS. Both of these zones, furthermore, are partitioned into blocks of bandwidth-time which will be assigned to the users or the network relays. For illustration, in Fig. 1, the area labeled as "*UL burst# i*" is a block of bandwidth-time of frame k which is allocated to the user ' i '. We represent this block by $\xi_{u_i}(k)$. Similarly, for the relay zone, we define $\xi_{r_j}(k)$ as the block of bandwidth-time assigned to the relay r_j , in frame k , see "*R-UL burst# j*" in Fig. 1.

In fact, depending on the scheduler decision the value of $T_1(k)$ and $T_2(k)$ change for different frames. However, in each frame, say k , the Access and Relay Zones durations should satisfy $T_1(k) + T_2(k) = T$. Furthermore, assuming no frequency reuse, it is clear that

$$\sum_{i=1}^N \xi_{u_i}(k) = W T_1(k), \text{ and } \sum_{j=1}^{N_r} \xi_{r_j}(k) = W T_2(k), \quad (3)$$

where W , N , and N_r show the available bandwidth, the total number of network users, and the number of relays in the cell, respectively. Note that, in case of no-relay networks,

$T_2(k)$ would be equal to zero and all the uplink duration would be shared by the network users, i.e., $T_1(k) = T$. Thus, in no-relay networks, $\sum_{i=1}^N \xi_{u_i}(k) = W T$.

Given the above notations, let us define $\omega_{\mathcal{T}}^{\mathcal{R}}(k)$ and $\omega_{\mathcal{T}}(k)$ as the average rate that the users transmit in frame k in the cases of relay and no-relay network, respectively. Here, \mathcal{T} denotes the whole area of the network. $\omega_{\mathcal{T}}^{\mathcal{R}}(k)$ and $\omega_{\mathcal{T}}(k)$ can be evaluated as

$$\begin{aligned}\omega_{\mathcal{T}}^{\mathcal{R}}(k) &= \frac{1}{T} \sum_{u_i \in U_{\mathcal{T}}} C(d_{u_i r_j}) \xi_{u_i}(k), \\ \omega_{\mathcal{T}}(k) &= \frac{1}{T} \sum_{u_i \in U_{\mathcal{T}}} C(d_{u_i B}) \xi_{u_i}(k),\end{aligned}\quad (4)$$

where T is the frame duration and

$$U_{\mathcal{T}} = \{u_i: \text{User 'i' is located in the region defined by } \mathcal{T}\}, \quad (5)$$

i.e., all of the network users.

Furthermore, to capture the average performance of the network, we need to average $\omega_{\mathcal{T}}^{\mathcal{R}}(k)$ and $\omega_{\mathcal{T}}(k)$ over different frames. Thus, we define $\Omega_{\mathcal{T}}^{\mathcal{R}}$ and $\Omega_{\mathcal{T}}$ as the average transmission rate over L (tends to infinity) frames for relay and no-relay networks, respectively. Therefore, we have

$$\begin{aligned}\Omega_{\mathcal{T}}^{\mathcal{R}} &= \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{k=1}^L \omega_{\mathcal{T}}^{\mathcal{R}}(k), \\ \Omega_{\mathcal{T}} &= \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{k=1}^L \omega_{\mathcal{T}}(k),\end{aligned}\quad (6)$$

2.3 Network Scheduling

To complete the network model, we require an assumption on the network scheduler. For example, in some applications, the scheduler tries to maximize the network throughput. Therefore, in each frame, it allocates all of the network resources to the user with the best channel condition. In another class of scheduling techniques, the objective is to allocate the network resources to all users in a fair manner. In this work, we assume that a fair scheduling strategy is used in the sense that on the average the network resources are distributed equally among all users (see [2]). In other words, all users have an equal chance to send their information. Based on this assumption, we have

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{k=1}^L \xi_{u_i}(k) = \frac{W \bar{T}_1}{|U_{\mathcal{T}}|}, \quad (7)$$

where \bar{T}_1 is the average duration of the uplink access zone. T is the whole area of the network, and $|U_{\mathcal{T}}|$ shows the cardinality of $U_{\mathcal{T}}$, i.e., the total number of network users.

3. RELAY LOCATION ESTIMATION

Performance of relay networks is strongly affected by the relays' locations. Indeed, there is a coupling between the relays' locations and the amount of resources that the network should assign to the relays. For instance, assume that a relay is placed in an area which is responsible for retransmission of data from a large number of users. Therefore, to keep the relay *stable*, it should receive a larger share of the network resources as compared to a lightly loaded relay. This fact implies that the optimal relays' locations can be found

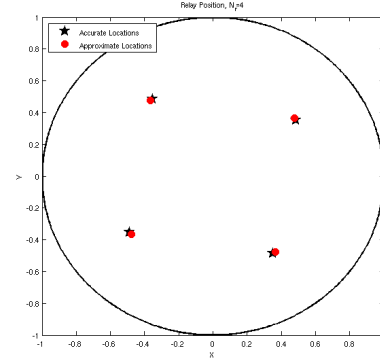


Figure 2: Degradation Imposed by Lower-Bound Maximization

when the resource allocation and relay placement problems are addressed jointly. To reduce the high complexity of this joint problem, in this study, we only propose a suboptimum scheme where the goal is to place the relays such that the users transmission rate, as given in (8), is maximized.

$$\begin{aligned}\Omega_{\mathcal{T}}^{\mathcal{R}} &= \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{u_i \in U_{\mathcal{T}}} \frac{1}{T} \sum_{k=1}^L C(d_{u_i r_j}) \xi_{u_i}(k) \\ &\stackrel{(a)}{=} \frac{W \bar{T}_1}{T |U_{\mathcal{T}}|} \sum_{u_i \in U_{\mathcal{T}}} C(d_{u_i r_j}) \\ &\stackrel{(b)}{\geq} \frac{W \bar{T}_1}{T} \mathbb{E}_S \left\{ \log_2 \left(1 + \frac{P S / N_0}{\left(\frac{1}{|U_{\mathcal{T}}|} \sum_{u_i \in U_{\mathcal{T}}} d_{u_i r_j}^2 \right)^2} \right) \right\}.\end{aligned}\quad (8)$$

where $C(d_{u_i r_j}) = \mathbb{E}_S \{ \log_2 (1 + \frac{P S / d_{u_i r_j}^4}{N_0}) \}$, as in (1). In equation (8), (a) follows from the fair resource allocation assumption (equation (7)) and the lower bound in (b) is based on applying the Jensen's inequality and the fact that $\mathbb{E}_S \{ \log_2 (1 + \frac{P S / d_{u_i r_j}^4}{N_0}) \}$ is a convex function of d^2 . Note that, to find the optimal relays' locations, one should maximize the $\Omega_{\mathcal{T}}^{\mathcal{R}}$, which unfortunately does not have a close form solution. Thus, to have an analytical expression for the relays' locations, in this study, instead of *maximizing* $\Omega_{\mathcal{T}}^{\mathcal{R}}$, we *maximize the lower-bound* in (8). Indeed, this would be a suboptimal solution for the original problem. However, as Fig. 2 depicts, this change of optimization function results in small degradation in the final relays' locations. It is also obvious that *lower-bound maximization* is equivalent to *minimizing the average user to relay distance square* ($D^2 = \frac{1}{|U_{\mathcal{T}}|} \sum_{u_i \in U_{\mathcal{T}}} d_{u_i r_j}^2$).

Given this approximation, we are now able to formulate the relay placement as a *vector quantization problem*. The locations of the users are considered as two-dimensional input points, and the outputs of the vector quantization algorithm would be ' N_r ' two-dimensional vectors representing the relays' locations. For instance, '*Kmeans*' with distance square metric takes users' locations as a series of source samples (\mathbf{x}) and finds ' N_r ' quantization points ($\hat{\mathbf{x}}$) as well as the appropriate mapping between x_i and \hat{x}_i .

Applying this algorithm, for the cases of $N_r \in [2, 6)$, it can be observed that relays are placed on a circle concentric

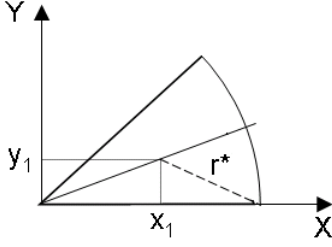


Figure 3: Relaying Region of a Relay

to the main cell. In other words, the relays partition the cell into N_r sectors with the central angle of $\frac{2\pi}{N_r}$ each. Users in the j^{th} sector are assigned to the corresponding relay at (x_j, y_j) placed on the bisector of that sector. This simple topology is no longer valid when $N_r \geq 6$. For such cases, relays are placed on two or more circles. Asymptotically, for large number of relays and uniform distribution of users, ‘Kmeans’ can be approximated, with lower complexity, using lattice quantization. Indeed, as N_r increases, the optimum locations of relays converge to the corners of a parallelogram with the angle of $\frac{\pi}{3}$ [1]. Figure 3 shows the results of the ‘Kmeans’ algorithm for the cases of $N_r = 3, 9, 12$, and 40.

Although ‘Kmeans’ algorithm is able to find the relays’ locations numerically, for the case of $N_r < 6$ (symmetrical structures), we can calculate them analytically as well. For instance, location of the first relay, (x_1, y_1) in Fig. 3, is the point which minimizes the average square distance (D^2) of all users residing inside the relay region of R_1 . For convenience, in the following, the set of users which are associated to the relay R_j is denoted by $U_{\mathcal{R}_j}$. By this notation, D^2 can be written as

$$D^2 = \int_{U_{\mathcal{R}_1}} [(r \cos \theta - x_1)^2 + (r \sin \theta - y_1)^2] r d\theta dr, \quad (9)$$

where $y_1 = \tan(\frac{\pi}{N_r})x_1$. Setting $\frac{\partial D^2}{\partial x_1} = 0$, (x_1, y_1) could be computed as follows

$$(x_1, y_1) = \begin{cases} x_1 = \frac{N_r}{3\pi} \sin(\frac{2\pi}{N_r})R & 2 < N_r < 6 \\ y_1 = \tan(\frac{\pi}{N_r})x_1 & \\ x_1 = 0; y_1 = 0.426R & N_r = 2 \end{cases}. \quad (10)$$

Similarly, location of the j^{th} relay can be found by rotating (x_1, y_1) by $\frac{(j-1)2\pi}{N_r}$ radians.

4. NETWORK AVERAGE TRANSMISSION RATE

According to section 3, the relays’ locations can be computed for different number of relays and any cell size. Having the relays’ locations and assuming a fair network scheduler, the average transmission rate can be computed versus the following four parameters: cell size, ratio of the network resources assigned to the users, number of relays, and users’ transmission power.

The ratio of the network resources which is assigned to the users is identified with the parameters Γ defined as $\Gamma = \frac{\bar{T}_1}{T}$

where \bar{T}_1 denotes the time average of $T_1(k)$ over a large number of frames. $T_1(k)$ is defined in subsection 2.2.

4.1 Average Transmission Rate: No-Relay Network ($\Omega_{\mathcal{T}}$)

For networks with no relay, the average transmission rate ($\Omega_{\mathcal{T}}$) can be computed as follows

$$\begin{aligned} \Omega_{\mathcal{T}} &\stackrel{(a)}{=} \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{k=1}^L \frac{1}{T} \sum_{u_i \in U_{\mathcal{T}}} \xi_{u_i}(k) C(d_{u_i B}) \\ &= \lim_{L \rightarrow \infty} \frac{1}{T} \sum_{u_i \in U_{\mathcal{T}}} C(d_{u_i B}) \frac{1}{L} \sum_{k=1}^L \xi_{u_i}(k) \\ &\stackrel{(b)}{=} \frac{W}{|U_{\mathcal{T}}|} \sum_{u_i \in U_{\mathcal{T}}} C(d_{u_i B}) \\ &\stackrel{(c)}{=} W \int_{r=0}^R C(r) f(r) dr \quad (11) \\ &= \int_{r=0}^R \mathbb{E}_S \left\{ W \log_2 \left(1 + \frac{PS/r^4}{N_0} \right) \right\} \frac{2}{R^2} r dr \\ &= \mathbb{E}_S \left\{ W \log_2 \left(1 + \frac{P S/R^4}{N_0} \right) + \frac{2W}{\ln 2} \frac{\sqrt{\frac{PS}{N_0}}}{R^2} \arctan \left(\frac{R^2}{\sqrt{\frac{PS}{N_0}}} \right) \right\}, \end{aligned}$$

where $U_{\mathcal{T}} = \{\text{all users in the cell}\}$, and $C(d_{u_i B})$ is given in (2). $f(r)$ is defined as the probability density function (pdf) of $d_{u_i B}$, where $d_{u_i B}$ denotes the distance of a station from the BS. For a uniform distribution of users, we have $f(r) = \frac{2\pi r \eta}{\pi R^2 \eta} = \frac{2r}{R^2}$.

In equation (11), (a) follows from equations (4) and (6). (b) results from equation (7), and (c) relies on the continuous approximation.

4.2 Average Transmission Rate: Relay Network ($\Omega_{\mathcal{T}}^{\mathcal{R}}$)

As described in section 2, all stations transmit their data to the BS via one of the relay nodes. Considering a non-cooperative structure, each user is associated with one of the N_r relays. Consequently, the users can be partitioned into N_r non-overlapping groups each of them sending data to a particular relay. Thus, $\Omega_{\mathcal{T}}^{\mathcal{R}}$ can be computed as

$$\begin{aligned} \Omega_{\mathcal{T}}^{\mathcal{R}} &\stackrel{(a)}{=} \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{k=1}^L \sum_{j=1}^{N_r} \frac{1}{T} \sum_{u_i \in U_{\mathcal{R}_j}} \xi_{u_i}(k) C(d_{u_i r_j}) \\ &= \lim_{L \rightarrow \infty} \sum_{j=1}^{N_r} \frac{1}{T} \sum_{u_i \in U_{\mathcal{R}_j}} C(d_{u_i r_j}) \frac{1}{L} \sum_{k=1}^L \xi_{u_i}(k) \quad (12) \\ &\stackrel{(b)}{=} \frac{W}{T} \bar{T}_1 \sum_{j=1}^{N_r} \frac{1}{|U_{\mathcal{T}}|} \sum_{u_i \in U_{\mathcal{R}_j}} C(d_{u_i r_j}) \end{aligned}$$

where $U_{\mathcal{R}_j}$ is the set of users associated with the relay j , and $C(d_{u_i r_j})$ is given in (1). In equation (12), (a) follows from equations (4) and (6), and (b) results from equation (7).

To analyze (12) mathematically, we can simplify it by considering networks which have less than six relays. As section

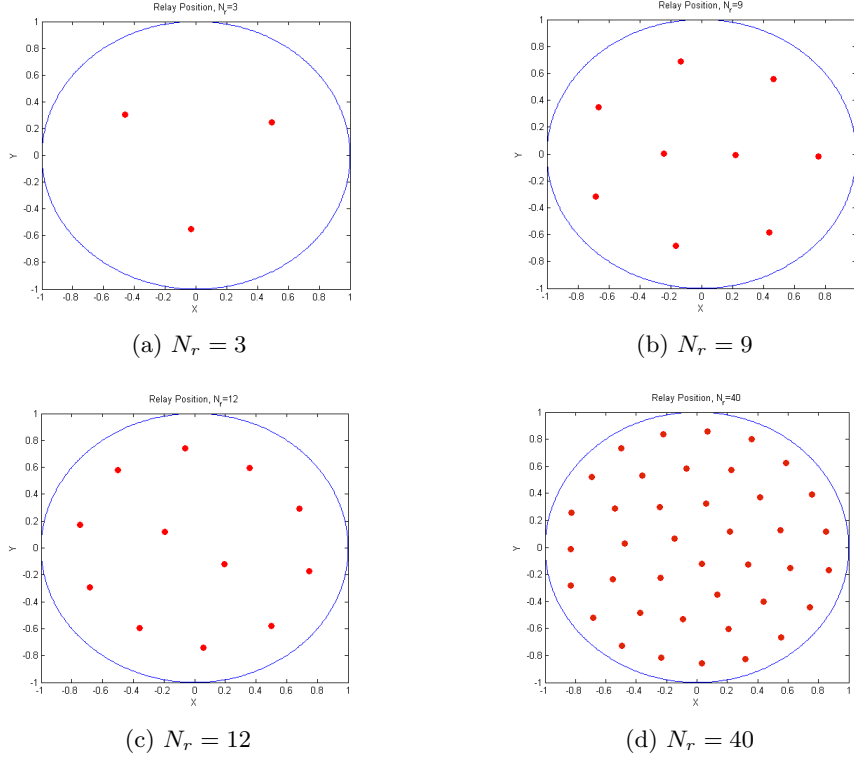


Figure 4: Suboptimum Relay Position

3 shows, in such conditions, the relays should be placed symmetrically in the network. Thus, $U_{\mathcal{R}_j}$ becomes identical for all $j \in \{1, 2, \dots, N_r\}$, and therefore, we can write

$$\begin{aligned}
 \Omega_{\mathcal{T}}^{\mathcal{R}} &= \frac{W \bar{T}_1 N_r}{T |U_{\mathcal{T}}|} \sum_{u_i \in U_{\mathcal{R}_1}} C(d_{u_i r_1}) \\
 &\stackrel{(a)}{=} W \Gamma \frac{1}{|U_{\mathcal{R}_1}|} \sum_{u_i \in U_{\mathcal{R}_1}} C(d_{u_i r_1}) \\
 &\stackrel{(b)}{=} W \Gamma \int_{U_{\mathcal{R}_1}} C(l) \frac{2l}{R^2} d\theta dl \\
 &= W \Gamma \int_{U_{\mathcal{R}_1}} \mathbb{E}_S \left\{ \log_2 \left(1 + \frac{P S / l^4}{N_0} \right) \frac{2l}{R^2} d\theta dl \right\},
 \end{aligned} \tag{13}$$

where $C(l)$ and l are defined as

$$\begin{aligned}
 C(l) &= \mathbb{E}_S \left\{ \log_2 \left(1 + \frac{P S / l^4}{N_0} \right) \right\}, \\
 l^2 &= (r \cos \theta - x_1)^2 + (r \sin \theta - y_1)^2.
 \end{aligned} \tag{14}$$

In equation (13), (a) is based on the assumption of the uniform distribution of users, i.e., $\frac{N_r}{|U_{\mathcal{T}}|} = \frac{1}{|U_{\mathcal{R}_1}|}$ and (b) relies on the continuous approximation, similar to the previous subsection.

5. NUMERICAL RESULTS

Having equations (11) and (13), we are able to compute the network performance for different combinations of “network range (R)”, “number of relay nodes (N_r)”, “ratio of the network resources allocated to the users (Γ)”, and “users

transmission power (P)”. In this section, the performance of the network is evaluated for different values of (R, N_r, Γ) and for the cases that $P = 21$ and 27 dBm and the noise floor is set to $N_0 = -93$ dBm [6].

5.1 Network Neutrality-Surface and Neutrality-Curve

In this subsection, the average transmission rates of relay and no-relay networks are compared. The relay network is determined by the parameters (R, N_r, Γ). To have a fair comparison, it is assumed that both networks have the same cell size. As mentioned in section 4, there is a tradeoff between the performance improvement achieved by relaying and the degradation caused by the loss of resources allocated to the relays. In other words, a relay network with the parameters (R, N_r, Γ) is able to compensate the effect of resource loss *iff*

$$\Omega_{\mathcal{T}}^{\mathcal{R}} \geq \Omega_{\mathcal{T}}, \tag{15}$$

where $\Omega_{\mathcal{T}}^{\mathcal{R}}$ and $\Omega_{\mathcal{T}}$ represent the average transmission rate for the network with and without relays, respectively. In the 3-dimensional space of (R, N_r, Γ), the set of points satisfying (15), with equality, form a *surface*, called the *neutrality-surface*.

In fact, all the triplets below the *neutrality-surface* have lower performance than the corresponding no-relay network. This graph helps the network designer to see how well the designed network performs. To illustrate the *neutrality-surface* visually, we project it over the 2-dimensional space. To this end, the value of Γ is fixed at $\Gamma = \Gamma_1$. Then, the *neutrality-*

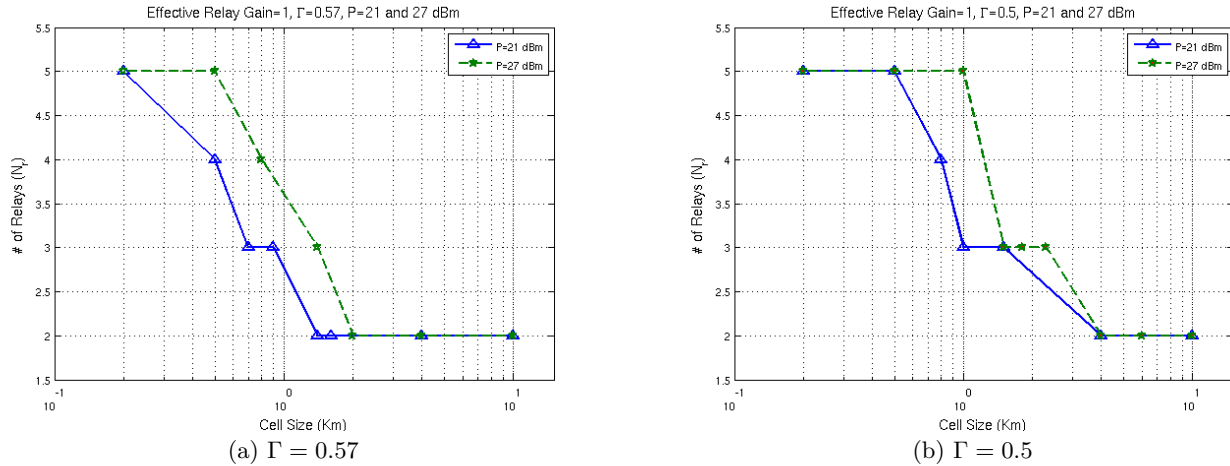


Figure 5: Network Neutrality-Curves

curve is defined as the intersection of the *neutrality-surface* and the *plane* determined by $\Gamma = \Gamma_1$.

Figure 5 depicts the *neutrality-curves* for the two cases where $\bar{T}_1 = \frac{4}{3}\bar{T}_2$ ($\Gamma = 0.57$) and $\bar{T}_1 = \bar{T}_2$ ($\Gamma = 0.5$). In each case, the *neutrality-curve* is plotted for two different values of transmission power.

As Fig. 5 shows, for a fixed Γ , the minimum required number of relays decreases when the network range (cell size) increases. This result can be justified by the fact that for smaller network sizes, the received SNR at the Base Station is high enough for successful decoding. On the contrary, in larger networks, the received SNR reduces because of the path loss. Consequently, even a small number of relays can improve the average transmission rate in larger cells.

Moreover, as Fig. 5 suggests, boosting the transmission power shifts the *neutrality-curve* to the right. In other words, for a fixed cell size, more relays are required to have $\Omega_{\mathcal{T}}^{\mathcal{R}} = \Omega_{\mathcal{T}}$ when $\frac{P}{N_0}$ is higher. Thus, the effect of increasing $\frac{P}{N_0}$ is similar to the effect of reducing the cell size. They both enhance the received SNR and reduce the necessity of relaying.

All of the above observations confirm the intuitive result that relaying improves the performance significantly only in the low SNR region. In the high SNR region, the resource loss imposed by relaying outweighs the gain achieved by increasing the received SNR.

Finally, comparing figures 5(a) and 5(b) shows that for a fixed R , if Γ is reduced, we need more relays to guarantee a gain from relay deployment. In fact, the lower is Γ , the lower will be the fraction of resources allocated to the users. Thus, to keep $\Omega_{\mathcal{T}}^{\mathcal{R}} = \Omega_{\mathcal{T}}$, the user to relay connection has to work in a higher SNR, which implies deploying more relays in the network.

5.2 Relative and Differential Relaying Gain

In this section, two performance metrics are defined to characterize the performance of relay networks. It is shown that these metrics are related to the *neutrality-curve* defined in subsection 5.1.

First, let us define the *relative relaying gain* as the ratio of the average transmission rate with and without relays ($\frac{\Omega_{\mathcal{T}}^{\mathcal{R}}}{\Omega_{\mathcal{T}}}$).

As an example, we have evaluated the relative relaying gain for a typical network where $\Gamma = 0.5$ and $P = 21$ dBm. The results are depicted in Fig. 6.

Figure 6 shows that for fixed values of R and Γ , increasing the number of relays always increases the relative relaying gain. Since $\Omega_{\mathcal{T}}$ is, by definition, independent of N_r , this result is equivalent to $\Omega_{\mathcal{T}}^{\mathcal{R}}$ being an increasing function of N_r .

Fig. 6 also suggests that the rate of growth of the relative relaying gain versus N_r is higher at larger cell sizes. In other words, adding relay nodes improves the network performance more significantly in larger networks. For instance, doubling the relay count from $N_r = 2$ to $N_r = 4$ in a network of cell radius $R = 1$ Km improves the relative relaying gain by about 60%, while this improvement is around 90% when the cell radius is $R = 3$ Km. Clearly, this result is consistent with the intuition that relaying is more advantageous when the network operates in the low SNR region.

It is also interesting to note that the dashed line in Fig. 6 shows the unity gain, i.e., $\Omega_{\mathcal{T}}^{\mathcal{R}} = \Omega_{\mathcal{T}}$, and is another representation of the *neutrality-curve* in Fig. 5.

Other than the relative relaying gain, differential relaying gain can be defined as another important metric for network performance comparison. From the network designer's perspective, the objective is to maximize the profit from a deployed infrastructure. It is also known that users are charged based on their quality of service which itself depends on their data rate. Thus, the no-relay network revenue, \mathbb{B} , can be evaluated as $\mathbb{B} = \mathcal{G} \Omega_{\mathcal{T}}$, where \mathcal{G} stands for the amount of income made per unit of data rate. Similarly, in the case of a relay network, we have $\mathbb{B}^{\mathcal{R}} = \mathcal{G} \Omega_{\mathcal{T}}^{\mathcal{R}}$. Correspondingly, the improvement in network revenue after relay deployment is proportional to the differential relaying gain, $\Omega_{\mathcal{T}}^{\mathcal{R}} - \Omega_{\mathcal{T}}$, as follows

$$\mathbb{B}^{\mathcal{R}} - \mathbb{B} = \mathcal{G}(\Omega_{\mathcal{T}}^{\mathcal{R}} - \Omega_{\mathcal{T}}). \quad (16)$$

This expression shows that other than the value of $\frac{\Omega_{\mathcal{T}}^{\mathcal{R}}}{\Omega_{\mathcal{T}}}$, the difference between the average transmission rate for the relay and no-relay networks ($\Omega_{\mathcal{T}}^{\mathcal{R}} - \Omega_{\mathcal{T}}$) should also be considered in the network design.

As an example, Fig. 7 depicts $\Omega_{\mathcal{T}}^{\mathcal{R}} - \Omega_{\mathcal{T}}$ versus R for

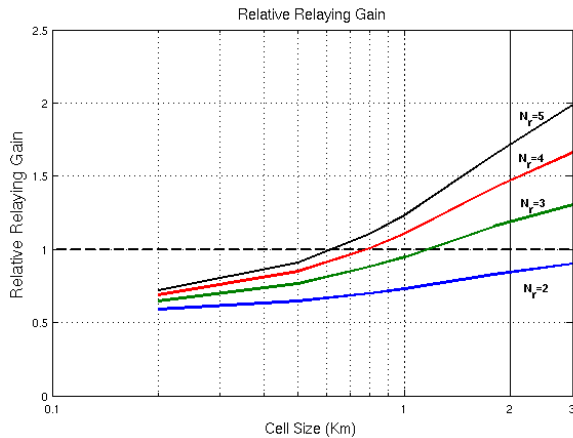


Figure 6: Relative relaying gain as a function of cell size for different values of deployed relays

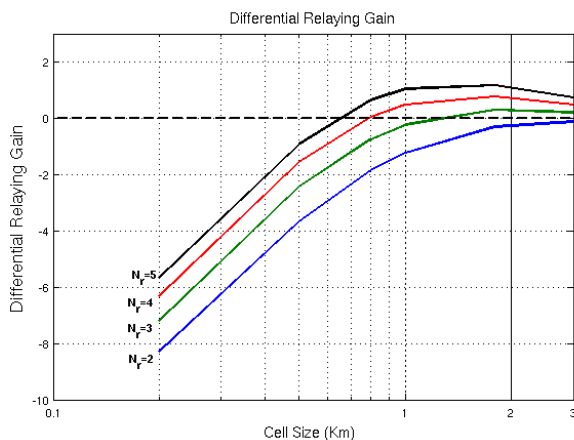


Figure 7: Differential relaying gain as a function of cell size for different values of deployed relays

different values of N_r , while Γ and P are fixed at $\Gamma = 0.57$ and $P = 21\text{dBm}$. At the end, note that the zero crossing points of Fig. 7 characterize the network configurations where $\Omega_{\mathcal{T}}^{\mathcal{R}} = \Omega_{\mathcal{T}}$, and represent another view of the *neutrality-curve*.

6. CONCLUSION

In this paper, a tradeoff associated with relaying in wireless networks is studied. Relaying results in higher SNR values and transmission rates. On the other hand, it potentially degrades the performance by consuming some of the network resources to retransmit the aggregated data to the BS. Considering a two-tier scenario, we discuss how vector quantization can be used to approximately compute the proper relays' locations. Then, in the multi-dimensional space of network parameters, the *neutrality-surface* is introduced. This surface characterizes the points of balance in the tradeoff, i.e., where the two opposite effects of relaying cancel each other. Finally, introducing relative and differential relaying gains as two performance metrics, different network configurations are analyzed.

7. ACKNOWLEDGMENTS

This work is supported by the Nortel Networks, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Ontario Center of Excellence (OCE).

8. REFERENCES

- [1] A. Gersho. Asymptotically Optimum Block Quantization. *IEEE Transactions on Information Theory*, 25:373– 380, 1979.
- [2] B Sadeghi, V Kanodia, A Sabharwal, and E Knightly. Opportunistic Media Access for Multirate Ad Hoc Networks. In *International Conference on Mobile Computing and Networking*, Atlanta, Georgia, USA, 2002.
- [3] G. Kramer, M. Gastpar, and P. Gupta. Cooperative Strategies and Capacity Theorems for Relay Networks. *IEEE Transactions on Information Theory*, 51:3037– 3063, 2005.
- [4] Guoqing Li and Hui Liu. Resource Allocation for OFDMA Relay Networks With Fairness Constraints. *IEEE Journal on Selected Areas in Communications*, 24:2061– 2069, 2006.
- [5] H. Viswanathan and S. Mukherjee. Performance of Cellular Networks with Relays and Centralized Scheduling. *IEEE Transactions on Wireless Communications*, 4:2318– 2328, 2005.
- [6] IEEE 802.16 Task Group. IEEE 802.16m Evaluation Methodology Document. Technical report, [Online] Available: <http://wirelessman.org/tgm/>, 2008.
- [7] IEEE 802.16j Task Group. IEEE 802.16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems: Multihop Relay Specification. *IEEE Standardization Committee*, 2007.
- [8] J. Boyer, D. D. Falconer, and H. Yanikomeroglu. Multihop diversity in wireless relaying channels. *IEEE Transactions on Communication*, 52:1820– 1830, 2004.
- [9] J. N. Laneman and G. W. Wornell. Distributed Space-Time Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks. *IEEE Transactions on Information Theory*, 49:2415– 2525, 2003.
- [10] K. Azarian and H. El Gamal. Cooperation in Outage-limited Multiple-Access Channels. In *IEEE International Zurich Seminar on Communications*, 2006.
- [11] M. Gastpar and M. Vetterli. On the Capacity of Wireless Networks: The Relay Case. In *INFOCOM*, 2002.
- [12] S. Toumpis, A.J. Goldsmith. Capacity Regions for Wireless Ad-Hoc networks. *IEEE Transactions on : Wireless Communications*, 2:736– 748, 2003.
- [13] Shengjie Zhao et. al. Macro Diversity Handover and Fast Access Station Switching for MMR Network. In *IEEE 802 Session 47*, London, UK, January 2007.
- [14] V. Pourahmadi, A. Khandani, W. Tong, P. Zhu. Effective Node Assignment in 2- Hop Fixed Relay Networks. In *IEEE 802 Plenary Session*, Dallas, 12-17 November 2006.
- [15] YT Hou, Y Shi, HD Sherali, and SF Midkiff. On Energy Provisioning and Relay Node Placement for Wireless Sensor Networks. *Wireless Communications, IEEE Transactions on*, 4:2579– 2590, 2005.