Resource Allocation and Relay Selection for Collaborative Communications

Saeed Akhavan Astaneh, Saeed Gazor

***Abstract*—We investigate the relay selection problem in a network where users are able to collaborate with each other; decode and forward the messages of each other along with their own messages to the destination. We study the performance obtained from collaboration in terms of 1) increasing the achievable rate, 2) saving the transmit energy, and 3) reducing the resource (time-bandwidth) requirement. To ensure fairness, we assume that the transmit energy to the rate ratio is fixed for all users. We allocate resource optimally for the proposed collaborative protocol (CP) and compare the result with the non-collaborative protocol (NCP) where users transmits their messages directly to the destination. The collaboration gain allows us 1) to decide whether to collaborate or not and 2) to select one relay among the possible relay users. We show that a considerable gain can be obtained if the direct source-destination channel gain is significantly smaller than those of alternative links. We demonstrate that a rate and energy improvement of up to $\left(1+\sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta}$ can be obtained, where η is the environment path loss exponent and k is the ratio of the rates of involved users. We also show that the collaboration is only beneficial for the middle range rate ratio.**

***Index Terms*—Collaboration, relay selection, resource allocation, rate improvement, energy saving, resource efficiency.**

I. Introduction

In wireless networks, the main interrelated quantities are achievable rate, transmit or receive energy and efficiency of resource. Many recent results, e.g., [1]–[3], show that collaboration among users in wireless networks, depending on channel condition and available energy, may increase the rate, save on the energy or reduce the resource requirement. Here, we ask the questions: Depending on channel condition and available energy, when collaboration is beneficial?, what are the involved gain or loss from possible collaboration?, and how to select one relay among the possible candidates? In order to answer these questions, we consider a network of users intending to send independent information to a their corresponding destinations (see Figure 1, for different scenarios). We propose that users assist each other only if in a fair way, the collaboration offers benefit in terms of rate, energy or resource. Here, the notion of fairness means that the achievable rates of different users would be proportional to their energy levels. We evaluate the effect of proposed collaboration protocol on system performance and then, based on the achieved gain or loss, present our protocol and relay selection.

Most of the existing CPs, e.g., [1]–[3], implicitly assume that a relay is already chosen, although, selective schemes have been investigated recently and several interesting methods have been proposed to choose the best relay among the potential relay users using different optimization criteria, for example the error rate in [4], [5], energy consumption and network lifetime in [6]–[8], diversity gain and outage probability in [9]–[11], the pricing technique in [12] and convex optimization in [13]. In all these references [1]–[13], it is assumed that the relay node provides free service to the source which is obviously beneficial to the source. Following [14], we study the problem of relay and protocol selection using three criteria; rate, energy and resource. In addition, to capture fairness among users, we assume that users will assist each other in relaying only if they gain from such a collaboration, thus those users having no data to transmit will not engage in such a collaboration. In contrast to [14], as it is important to take into consideration the different rate demands of various users, we introduce a new priority parameter. This parameter is imposed by an upper layer in order to determine the ratio of rate demands of involved users in the network. The motivation of this approach is to provide differentiated/prioritized services (see [15]). In this paper assuming that a rate ratio is provided by the upper layer, we either maximize the achieved rate, minimize the energy consumption or the resource utilization.

The remainder of the paper is organized as follows. Next we present the system model and present the protocols in Section II. In Section III we study single relay networks and investigate the rate, energy and resource improvement from possible collaboration. We then provide conditions on the location of the relay user for collaboration to be beneficial. In Section IV, we present our relay selection protocols. Extensions to the general network with multiple source and relay topology are discussed in Section V. Finally, in Section VI we give our concluding remarks.

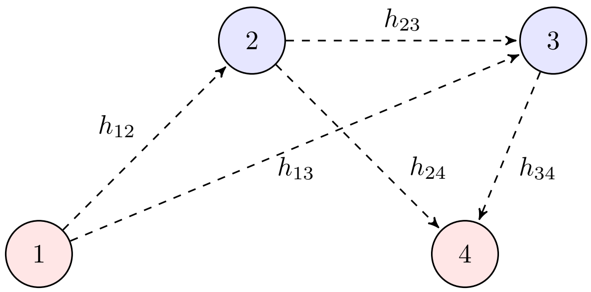


Fig. 1. A collaborative network, the channel energy gain between $i^{\rm th}$ and $j^{\rm th}$ user is denoted by h_{ij}. Consider three scenarios: 1) the $1^{\rm st}$ and the $2^{\rm nd}$ users transmit to the $3^{\rm rd}$ user, 2) the $1^{\rm st}$ user transmit to the $3^{\rm rd}$ user and the $2^{\rm nd}$ user broadcasts to the $3^{\rm rd}$ and the 4th users, 3) the $1^{\rm st}$ to the $3^{\rm rd}$, the $2^{\rm nd}$ and the $3^{\rm rd}$ to the 4th.

Manuscript received October 23, 2008; revised May 8, 2009; accepted August 25, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr Young-Chai Ko.

Authors are with the Department of Computer Engineering, Queen's University, Kingston, Ontario K7L 3N6, Canada, s.gazor, astaneh@queensu.ca.

II. System Model and Protocols

Consider the first scenario in Figure 1, where we assume that the 1^{st} and 2^{nd} users wish to transmit independent messages respectively with rates R_1 and R_2 to the 3^{rd} user over an additive white Gaussian noise channel and the 2^{nd} user may also assist the 1^{st} user to transmit its messages to the 3^{rd} user. Let denote the channel gain of the communication link between the i^{th} and j^{th} user by h_{ij}. We assume that the gain of all the channel links are perfectly known to all the receivers and transmitters. Thus, users need to acquire their channel gains via efficient channel estimation algorithms (see for example [16]) and make it available for all other users. To this end, we assume that collaborators could initially exchange messages to establish a collaboration protocol before main data streams are transmitted. Apparently, such exchanges consume a fraction of the available energy and resource. In this paper, we ignore the extra cost induced by this communication overhead. In addition, upper layers, such as MAC layer, must be intelligent (see [17], [18]) which is beyond the scope of this paper.

When the users collaborate, the network is a multi-hopping network where one user receives the messages of another user and forwards the decoded messages to the intended receiver as well as its own messages. Otherwise, they form a multiple access channel, i.e., they transmit directly to the receiver via a resource sharing method.

We assume that users transmit via a resource division protocol where the i^{th} user can transmit over a portion β_i of available resource (by resource we mean the product of used time and bandwidth, i.e., $B \times T$.) Using time/frequency division requires perfect time/frequency synchronization. In this paper, we ignore the required overhead to achieve this synchronization and assume perfect synchronization. The received energy to noise ratio within the resource slot $\beta_i BT$ can be expressed as $\frac{h_{ij}E_i}{N\beta_i BT}$, where E_i denotes the transmit energy of the i^{th} user and N denotes the received noise power. Unless otherwise stated, we consider the case where the available resource BT is unit, i.e., $BT = 1$. Let define the ratio of transmit energy to received noise power (TERN) as $\epsilon_i = \frac{E_i}{N}$. Thus, the achievable rate for the user i is given by

$$R_i = \beta_i \log\left(1 + \frac{h_{ij}\epsilon_i}{\beta_i}\right). \tag{1}$$

To pose the fairness issue in a multi-user communication network, we first need to define a fairness constraint. Most data applications are sensitive to error but tolerant to delay. It is clear that a higher signal-to-interference- plus-noise (SINR) ratio level at the output of the receiver will generally result in a lower bit-error rate, and hence higher throughput. However, achieving a high SINR level requires the user terminals to transmit at a high power, which in turn results in low battery life. However, similar to [19], [20], we impose the constraint $\frac{\epsilon_i}{R_i} = \mathrm{cte}$ in order to maintain the fairness in the network for all users. This constraint ensures fairness among users as the energy spent by users is proportional to their demand for rate. This constraint captures the tradeoff between the throughput and energy consumption and is particularly suitable for applications in which energy efficiency is critical [19]. Fixing the transmit-energy-to-rate ratio among all users (which is way to guarantee a fairness in network [19], [20]) does penalize the nodes with higher channel gain versus other nodes. However, this approach can be deemed as fair since from a consumer perspective, users can control their rate demand versus the required transmit energy and have no control over their channel gains. The fairness constraint also can be recast as

$$\frac{R_2}{R_1} = \frac{\epsilon_2}{\epsilon_1} \stackrel{\mathrm{def}}{=} k. \tag{2}$$

which hereafter we denote k as the rate ratio and as a design parameter imposed by upper layers of the network such as MAC layer.

We consider a half-duplex communication network where each user can either transmit or receive (but not both) at any time and any frequency band. Throughout this paper, we consider two following communication protocols:

- Non Collaborative protocol where users transmit directly to the destination via a resource (time and frequency) division method.
- Collaborative protocol where over the first resource slot, the 1^{st} user transmits its message and the 2^{nd} user decodes the message of the 1^{st} user. Then, over the 2^{nd} resource slot, the 2^{nd} user re-encodes the decoded message of the 1^{st} user in conjunction with its own message, the 2^{nd} message, and broadcasts the encoded message.

In the CP, the 2^{nd} user consumes a portion of the available energy to relay for the 1^{st} user. In return, the 2^{nd} user may be compensated by receiving more resource. The more resource implicitly allow users to save on energy. Therefore, an incentive for engaging in a collaboration could be access to excess of resource.

III. Collaboration in Single Relay Networks

In the following we study some properties of proposed protocols and investigate upper and lower bounds for achievable rates. We use the Shannon capacity as the performance metric. In order to justify this assumption, we assume that user employ a capacity-achieving channel coding.

A. Non-Collaborative Protocol (NCP)

In this protocol, during 1^{st} portion of resource slot, i.e., β_1, the 1^{st} user transmits its message. The receiver, the 3^{rd} user, decodes this message correctly for a maximum rate of $R_1 = \beta_1 \log\left(1 + \frac{h_{13}\epsilon_1}{\beta_1}\right)$. In a similar manner, the maximum rate of the 2^{nd} user which could be decoded reliably at the 3^{rd} user is $R_2 = \beta_2 \log\left(1 + \frac{h_{23}\epsilon_2}{\beta_2}\right)$. Since, we assume that one unit of resource is available, i.e., $\beta_1 + \beta_2 = 1$, hereafter, we denote $\epsilon_1 \stackrel{\mathrm{def}}{=} \epsilon$, $\epsilon_2 = k\epsilon$, $\beta_1 \stackrel{\mathrm{def}}{=} \beta$ and $\beta_2 = 1 - \beta$. Hence, we get the following optimization problem for NCP:

$$\begin{aligned} R_{\mathrm{NCP}} &= \max_{\beta} \left(R_1(\beta) + R_2(1-\beta)\right) \\ \text{s.t.} \quad & \frac{R_2}{R_1} = k \end{aligned} \tag{3}$$

where R_{NCP} is the achievable sum rate of users and $R_1(\beta) = \beta \log\left(1 + \frac{h_{13}\epsilon}{\beta}\right)$ and $R_2(1-\beta) = (1-\beta)\log\left(1 + \frac{h_{23}k\epsilon}{1-\beta}\right)$. Since $R_1(\beta)$ and $R_2(1-\beta)$ are increasing and decreasing function of β, respectively, the solution of the above optimization is the unique solution of the following

$$\begin{aligned} R_{\mathrm{NCP}} &= (k+1)\beta \log\left(1 + \frac{h_{13}\epsilon}{\beta}\right) \\ &= \frac{k+1}{k}(1-\beta)\log\left(1 + \frac{h_{23}k\epsilon}{1-\beta}\right). \end{aligned} \tag{4}$$

The optimal resource β_i as a solution of (4) is a function of h_{ij} and ϵ. However for ease of notation and abbreviation, we denote the optimal resource only by β_i.

B. Collaborative protocol (CP)

In this protocol, over the 1^{st} portion of the resource slot, i.e., β, the 1^{st} user transmits its messages at rate R_1. During this time, The 3^{rd} user is switched off and thus ignores the received signal from the 1^{st} user. The 2^{nd} user attempts to decode the messages of the 1^{st} user. Hence, the maximum achievable rate for the 1^{st} user is expressed as $R_1 = \beta \log\left(1 + \frac{h_{12}\epsilon}{\beta}\right)$. Over the remaining portion of resource slot, i.e., $1-\beta$, the 2^{nd} user re-encodes the decoded messages of the 1^{st} user and transmits the messages of the 1^{st} user as well as its own messages to the intended destination. In fact, during this time, the 2^{nd} user must transmit at rate of $\frac{k+1}{k}R_2$ to accommodate both data. The maximum achievable rate which may be decoded reliably at the 3^{rd} user is $R_2 = \frac{k(1-\beta)}{k+1}\log\left(1 + \frac{h_{23}k\epsilon}{1-\beta}\right)$. This yields the following max-min resource allocation problem:

$$\begin{aligned} R_{\mathrm{CP}} &= \max_{\beta}\left(R_1(\beta) + R_2(1-\beta)\right) \\ \text{s.t.} \quad & \frac{R_2}{R_1} = k \end{aligned} \tag{5}$$

where R_{CP} is the achievable sum rate of users which will be compared with R_{NCP}. In a similar way, the optimal solution is the unique solution of the following equation with respect to β:

$$\begin{aligned} R_{\mathrm{CP}} &= (k+1)\beta \log\left(1 + \frac{h_{12}\epsilon}{\beta}\right) \\ &= (1-\beta)\log\left(1 + \frac{h_{23}k\epsilon}{1-\beta}\right). \end{aligned} \tag{6}$$

Similar to the NCP, for ease of notation, we merely denote the optimal resource by β_i.

C. Rate Improvement for Given Resource and Energy

In this section, we define the collaboration gain as the ratio of achievable sum rate of the CP to that of the NCP, i.e., $\frac{R_{\mathrm{CP}}}{R_{\mathrm{NCP}}}$. This ratio represents the achievable sum rate improvement of these protocols. We derive tight upper and lower bounds and study the asymptotic behavior of the collaboration gain at low and high TERN and rate ratio.

Since $R_1(\beta)$ and $R_2(1-\beta)$ are increasing and decreasing convex and continuous functions of β, respectively, the maximization (4) is guaranteed to have a unique solution. Unfortunately, this solution has no closed form expression. In Appendix A, we derive the upper and lower bounds in (7b) and (7a) for these achievable rates.

$$R_{\mathrm{NCP}} > \frac{\frac{1}{k}\log(1+kh_{23}\epsilon)\log(1+h_{13}\epsilon)}{\frac{1}{k}\log(1+kh_{23}\epsilon) + \log(1+h_{13}\epsilon)}. \tag{7a}$$

These bound are tight for high TERN $\epsilon \to \infty$; this is the case where the noise power is negligible compared with the received signal powers. In high TERN regime, the available resource is allocated to the users receive in proportion with their rate demands, i.e., $\lim_{\epsilon\to\infty} \beta = \frac{1}{k+1}$. The lower bound in (7a) is obtained the intersection point of the two lines connecting end points of the rate curves

Using the same approach, we can find the bounds in (8a) and (8b) for the achievable sum-rate of the CP

$$R_{\mathrm{CP}} > \frac{\frac{1}{k+1}\log(1+kh_{23}\epsilon)\log(1+h_{12}\epsilon)}{\frac{1}{k+1}\log(1+kh_{23}\epsilon) + \log(1+h_{12}\epsilon)} \tag{8a}$$

which are tight in high TERN regime.

From (7b) and (8a), it is easy to see that $\lim_{\epsilon\to\infty}\frac{R_{\mathrm{CP}}}{R_{\mathrm{NCP}}} \geq \frac{k+1}{k+2}$. In addition, from (8b) and (7a), we can see that $\lim_{\epsilon\to\infty}\frac{R_{\mathrm{CP}}}{R_{\mathrm{NCP}}} \leq \frac{k+1}{k+2}$. Thus $\lim_{\epsilon\to\infty}\frac{R_{\mathrm{CP}}}{R_{\mathrm{NCP}}} = \frac{k+1}{k+2}$. Thus the sum rate gain $\frac{k+1}{k+2}$ is smaller than one in high TERN regime; this means that where large amount of received energy to noise ratio is available the collaborative schemes are not attractive.

In Appendix A, we also derive the tight bounds in (9a) and (9b) for low TERN regime (small values of ϵ)

$$\begin{aligned} R_{\mathrm{NCP}} &< \min\left\{\log(1+h_{13}\epsilon), \frac{1}{k}\log(1+kh_{23}\epsilon)\right\} \\ &\leq \epsilon \min\{h_{13}, h_{23}\}. \end{aligned} \tag{9b}$$

In addition, the achievable rate is also lower bounded by two end points of the curves. This upper bound is tight in low TERN regime, i.e., where the received signal is dominated by noise power. From the above, we conclude that

$$\lim_{\epsilon\to 0^+}\frac{R_{\mathrm{NCP}}}{\epsilon} = \min\{h_{13}, h_{23}\}. \tag{10}$$

Similar to the non-collaborative case, we derive the following upper and lower bounds for CP:

Thus, we conclude that

$$\lim_{\epsilon\to 0^+}\frac{R_{\mathrm{CP}}}{\epsilon} = \min\{h_{12}, \frac{k}{k+1}h_{23}\}. \tag{12}$$

By combining (10) and (12), we get the following result

$$\lim_{\epsilon\to 0^+}\frac{R_{\mathrm{CP}}}{R_{\mathrm{NCP}}} = \frac{\min\left\{h_{12}, \frac{k}{k+1}h_{23}\right\}}{\min\{h_{13}, h_{23}\}}. \tag{13}$$

In addition, It is easy to show that $\frac{R_{\mathrm{CP}}}{R_{\mathrm{NCP}}}$ is always smaller than $\frac{\min\left\{h_{12}, \frac{k}{k+1}h_{23}\right\}}{\min\{h_{13}, h_{23}\}}$, i.e., $\frac{R_{\mathrm{CP}}}{R_{\mathrm{CP}}} \leq \frac{\min\left\{h_{12}, \frac{k}{k+1}h_{23}\right\}}{\min\{h_{13}, h_{23}\}}$. This means that the rate gain can be greater than unity only if $h_{13} \leq \min\{h_{12}, h_{23}\frac{k}{k+1}\}$. In this case, the maximum rate gain $(\min\{\frac{h_{12}}{h_{13}}, \frac{h_{23}}{\frac{k+1}{k}h_{13}}\})$ is only achievable in low TERN regime.

Now, we examine the collaborative gain when the rate ratio is large. It is easy to see that for large k, the optimal β, which is either the solution of (4) or (6), tends to zero, i.e., $\beta \to 0$. This implies that more resource should be allocated to the higher demanding user. Hence, it is easy to show that $\lim_{k\to\infty}\frac{\log(k)}{k}R_{\mathrm{NCP}} = \lim_{k\to\infty}\frac{\log(k)}{k}R_{\mathrm{CP}} = 1$. Then, it follows

$$R_{\rm NCP} < \frac{\frac{\log(1+h_{13}(k+1)\epsilon)}{\left(1-\frac{1}{1+h_{13}(k+1)\epsilon}-\log(1+h_{13}(k+1)\epsilon)\right)}+\frac{k\log(1+h_{23}(k+1)\epsilon)}{\left(1-\frac{1}{1+h_{23}(k+1)\epsilon}-\log(1+h_{23}(k+1)\epsilon)\right)}}{\frac{(k+1)}{\left(1-\frac{1}{1+h_{13}(k+1)\epsilon}-\log(1+h_{13}(k+1)\epsilon)\right)}+\frac{k(k+1)}{\left(1-\frac{1}{1+h_{23}(k+1)\epsilon}-\log(1+h_{23}(k+1)\epsilon)\right)}} \tag{7b}$$

$$R_{\rm CP} < \frac{\frac{\log(1+h_{12}(k+1)\epsilon)}{\left(1-\frac{1}{1+h_{12}(k+1)\epsilon}-\log(1+h_{12}(k+1)\epsilon)\right)}+\frac{(k+1)\log\left(1+h_{23}\frac{k(k+2)}{k+1}\epsilon\right)}{\left(1-\frac{1}{1+h_{23}(k+1)\epsilon}-\log\left(1+h_{23}\frac{k(k+2)}{k+1}\epsilon\right)\right)}}{\frac{k+2}{\left(1-\frac{1}{1+h_{12}(k+1)\epsilon}-\log(1+h_{12}(k+1)\epsilon)\right)}+\frac{(k+1)(k+2)}{\left(1-\frac{1}{1+h_{23}\frac{k(k+2)}{k+1}\epsilon}-\log\left(1+h_{23}\frac{k(k+2)}{k+1}\epsilon\right)\right)}} \tag{8b}$$

$$R_{\rm NCP} > \epsilon\frac{2h_{23}+2h_{13}-h_{13}^2\epsilon-kh_{23}^2\epsilon-\sqrt{4(h_{23}-h_{13})^2+\epsilon^2\left(h_{13}^2+kh_{23}^2\right)^2+4\epsilon(h_{23}-h_{13})\left(h_{13}^2-kh_{23}^2\right)}}{4}. \tag{9a}$$

$$R_{\rm CP} < \min\left\{\log(1+h_{12}\epsilon), \tfrac{1}{k+1}\log(1+kh_{23}\epsilon)\right\} \le \epsilon\min\left\{h_{12}, \tfrac{k}{k+1}h_{23}\right\} \tag{11a}$$

$$R_{\rm CP} > \epsilon\frac{\frac{2kh_{23}}{k+1}+2h_{12}-h_{12}^2\epsilon-\frac{h_{23}^2k^2\epsilon}{k+1}-\sqrt{4\left(\frac{kh_{23}}{k+1}-h_{12}\right)^2+\epsilon^2\left(h_{12}^2+\left(\frac{kh_{23}}{k+1}\right)^2\right)^2+4\epsilon\left(\frac{kh_{23}}{k+1}-h_{12}\right)\left(h_{12}^2-(\frac{kh_{23}}{k+1})^2\right)}}{4} \tag{11b}$$

that

$$\lim_{k\to\infty}\frac{R_{\rm CP}}{R_{\rm NCP}} = 1 \tag{14}$$

On the other hand, if k tends to zero (where the rates of the $1^{\rm st}$ user is larger than the rate of the $2^{\rm nd}$ user), the optimal β for NCP tends to unity, while for CP tends to zero. Thus, the collaborative gain for small values of k, i.e., $k \to 0$, is

$$\lim_{k\to 0^+}\frac{1}{k}\frac{R_{\rm CP}}{R_{\rm NCP}} = \frac{h_{23}\epsilon}{\log\left(1+h_{13}\epsilon\right)}. \tag{15}$$

It follows that for small enough rate ratio the achievable rate of NCP is strictly greater than that of CP, i.e, $R_{\rm NCP} > R_{\rm CP}$.

D. Energy Saving for Given Capacity and Resource

In the following, we are interested in quantifying the advantage of the collaboration in terms of energy saving. This is in contrast to the previous section where the rate is maximized provided a fixed amount of available energy. Here, we assume that each user requires some specified rate R_i and has to allocate TERN proportional to R_i. In order to meet these rate requirements, users may collaborate (or not) to use available resource efficiently. Given a unit of shared resource, we minimize the TERN as follows

$$\text{CP}: \quad \begin{cases} \min \epsilon_{\rm CP}, \\ \quad \text{s.t. } R = \beta\log\left(1+\frac{h_{12}\epsilon_{\rm CP}}{\beta}\right) \\ \quad\quad = \frac{1-\beta}{k}\log\left(1+\frac{h_{23}k\epsilon_{\rm CP}}{1-\beta}\right) \end{cases} \tag{16a}$$

$$\text{NCP}: \quad \begin{cases} \min \epsilon_{\rm NCP}, \\ \quad \text{s.t. } R = \beta\log\left(1+\frac{h_{13}\epsilon_{\rm CP}}{\beta}\right) \\ \quad\quad = \frac{1-\beta}{k+1}\log\left(1+\frac{h_{23}k\epsilon_{\rm CP}}{1-\beta}\right) \end{cases} \tag{16b}$$

Since the rates in (3), (5) are monotonically increasing functions of TERN, thus, it is easy to show that optimization problem (16) is the dual of (3) and (5). This means that under similar channel gains, the TERN collaboration gain (i.e., the ratio of TERN in NCP to that of collaborative one $\frac{\epsilon_{\rm NCP}}{\epsilon_{\rm CP}}$) obtained from (16) is the same as the rate collaboration gain from (3) and (5). More specifically from this duality, we conclude that

$$\frac{\epsilon_{\rm NCP}}{\epsilon_{\rm CP}} \le \frac{\min\left\{h_{12}, \frac{k}{k+1}h_{23}\right\}}{\min\left\{h_{13}, h_{23}\right\}}. \tag{17}$$

Similarly, the maximum gain is obtained when the rate demand is small, i.e., as $R \to 0$.

E. Resource Efficiency for Given Capacity and Energy

Let $\beta_{i,\rm NCP}$ ($\beta_{i,\rm CP}$), denote the required resource for the $i^{\rm th}$, $i = 1, 2$, user to transmit its own information, R and kR (R and $(k+1)R$), under TERN constraints of ϵ and $k\epsilon$, respectively, in NCP (CP). We also define resource efficiency as $\frac{\beta_{\rm NCP}}{\beta_{\rm CP}}$, where $\beta_{\rm NCP}$ and $\beta_{\rm CP}$ are solution of the following equations:

$$\begin{cases} R = \beta_{1,\rm NCP}\log\left(1+\frac{h_{13}\epsilon}{\beta_{1,\rm NCP}}\right) = \frac{\beta_{2,\rm NCP}}{k}\log\left(1+\frac{h_{23}\epsilon}{\beta_{2,\rm NCP}}\right), \\ \beta_{\rm NCP} = \beta_{1,\rm NCP}+\beta_{2,\rm NCP}, \end{cases} \tag{18a}$$

$$\begin{cases} R = \beta_{1,\rm CP}\log\left(1+\frac{h_{12}\epsilon}{\beta_{1,\rm CP}}\right) = \frac{\beta_{2,\rm CP}}{k+1}\log\left(1+\frac{h_{23}\epsilon}{\beta_{2,\rm CP}}\right), \\ \beta_{\rm CP} = \beta_{1,\rm CP}+\beta_{2,\rm CP}. \end{cases} \tag{18b}$$

Note that we have feasible solution only if $R \le \epsilon\min\{h_{13}, h_{23}\}$ for the NCP and $R \le \epsilon\min\{h_{12}, h_{23}\frac{k}{k+1}\}$ for the CP. As the required rates approach these upper bounds the resource usage tends to infinity. In both protocols, due to the fairness constraint, the user with the worst channel obtains a larger amount of resource.

F. Effect of Network Geometry

In the following, we investigate the impact of the location of the relay user on the collaboration gain. In particular, we assume that the signal attenuation is governed by geometry of users as $h_{ij} = \frac{1}{{d_{ij}}^{\eta}}$ on two dimensional plane, where d_{ij} denotes the distance between the i^{th} and j^{th} users. In order to understand the impact of users relative locations on the collaboration gain, we investigate the region where transmission via collaboration provides more gain. We assume that in the two dimensional plane, the source, relay and destination are located on $(-\frac{1}{2}, 0)$, (x, y) and $(\frac{1}{2}, 0)$, respectively. Plugging the channel gains as $\frac{1}{d^{\eta}}$ and $\frac{1}{(1-d)^{\eta}}$ into the equations (4) and (6), we obtain the rate improvement of both protocols as a function of geometry of relay user. Figure 2 depicts the region where collaboration provides more benefit, i.e., the rate of CP is more than that of the NCP. This figure also depicts the contours of rate gain, where the ratio of achievable rate of protocols is fixed numbers (we plotted for the rate gains of 1, 2 and 4). We observe that as the rate ratio k increases the collaboration contours enlarge. Further increasing the rate ratio, the gain contours reduce. It implies that if the users with middle rate demand have incentive to collaborate with other users.

Since the channel gains are symmetric in two dimensional space, it is clear that the optimal relay user lies on the line connecting the source to the destination. We observe that the gain contours are approximately the intersections of two arcs with the radii $(gc)^{1/\eta}$ and $\left(\frac{k+1}{k}gc\right)^{1/\eta}$ with gc being $gc = \frac{R_{\text{CP}}}{R_{\text{NCP}}}$. In order to find the optimal placement of the relay user we examine the equation (13). It is easy to see that the optimal location is

$$d = \frac{1}{1 + \left(\frac{k}{k+1}\right)^{1/\eta}} \tag{19}$$

where at that point the maximum rate gain of

$$\frac{R_{\text{CP}}}{R_{\text{NCP}}} \leq \left(1 + \sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta} \tag{20}$$

is achievable.

Clearly the optimal location lies on the line connecting nodes 1 and 3. We now assume that all three nodes are on a one dimensional line and are located at 0, d and 1, respectively, i.e. $h_{12} = \frac{1}{d^{\eta}}$, $h_{13} = 1$ and $h_{23} = \frac{1}{(1-d)^{\eta}}$. Figure 3 shows the resulting ratio of maximum achievable rates using the CP and the NCP versus the location of the relay node $d \in [0, 1]$ for $\eta = 3$. As intuited from the upper bound in (20), for higher rate ratio k, more gain is expected (see in Figure 3(a)), however, as depicted in Figure 3(b), for large enough available energy of users ϵ, the collaboration gain degrades as rate ratio k increases. Figure 4 presents the rate improvement from CP and NCP protocols versus the rate ratio of users k. We observe that for small rate ratio, the rate improvement is zero and for large values of k, the rate improvement tends to unity which also confirms that for high rate ratio k, the collaboration is not beneficial. Figure 5 depicts the resource gain of the CP compared with NCP, i.e., $\frac{\beta_{\text{NCP}}}{\beta_{\text{CP}}}$ (18), for a required rate of $0.5h_{13}\epsilon$ versus location of the relay node. We observe that for a given required rate, depending on the relay channel condition, the resource gain is greater than unity. We have noticed that for small rate ratio k, CP provides more gain in terms of resource usage. In addition, for small rate ratio, the best location for relay user is almost in the middle of the source and destination users. Figure 6 shows the energy gain of the CP compared with the NCP, i.e., $\frac{\epsilon_{\text{NCP}}}{\epsilon_{\text{CP}}}$ (16), for a given required rate of $R = 0.09h_{13}$ versus the location of the relay node. Employing the CP, we obtain significant energy savings even for $\eta = 3$, provided that the relay is located appropriately. In addition, we observe that for higher rate ratio (see Figure 5), users benefit less in terms of resource efficiency. We deduce that only users which are interested in resource efficiency, with less rate requirement, can gain from possible collaboration. It is interesting to note that the CP provides rate/energy gain even for $\eta = 2$, by contrast, for such a small η there is no gain in rate/energy if the relay has no information to transmit (traditional multi-hopping) [3].

IV. Collaboration in Multiple Relay Networks

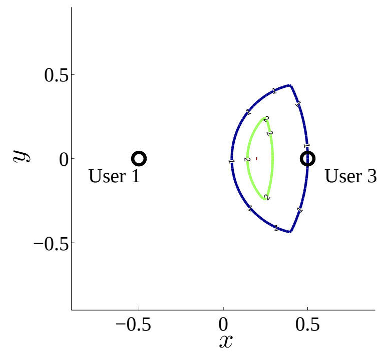
In the following, we propose our relay selection protocols based on the collaboration gain which is introduced in previous section. We use the channel gains to select one relay among the available relay users to participate in collaboration. We note that if the NCP outperforms the collaborative one, we fall back on the NCP, i.e., no relay user would be selected and the source sends its information to the destination directly. Otherwise, the source employs one relay in forwarding its information to the destination. The main objective of the proposed protocols are to achieve higher collaboration gain, higher rate improvement, energy saving or resource efficiency while guaranteeing fairness for all users. However, in large networks, the cost induced by communication overhead must be considered in future works.

A. Relay Selection: Rate Improvement and Energy Saving

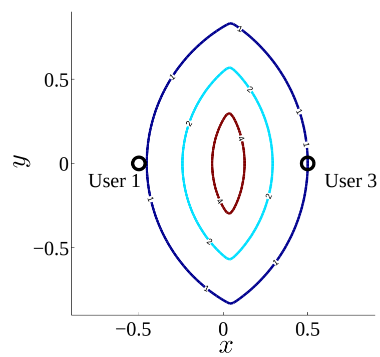
First, we consider the rate improvement as a criterion to select the best relay. As shown in previous section, the energy minimization problem is dual of the rate maximization problem, hence the relay selection protocol holds for the energy saving as well.

The result in (13) is very intuitive and suggests a strategy in deciding to use collaboration and to choose a relay user among the potential candidates. Given the full CSI, collaboration protocol is preferred if $\epsilon \ll 1$ and $h_{13} \ll \min\{h_{12}, h_{23}\frac{k}{k+1}\}$. In order to maximize the rate gain, the best relay user is the one that maximizes the $\frac{\min\{h_{12}, h_{23}\frac{k}{k+1}\}}{h_{13}}$.

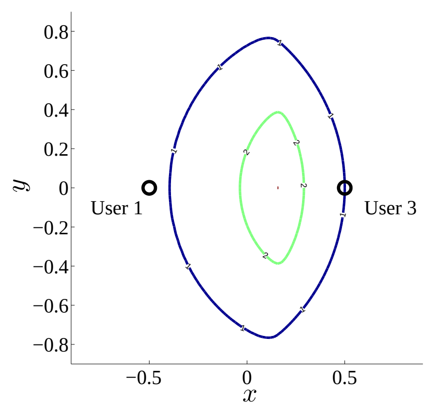
The results in (14) and (15) also provide an attractive guideline that for low and high rate ratio, NCP is preferred. The equation (19) implies that the best relay user, in order to maximize the rate gain, is located almost in the middle of the source and destination users. We observe that under severe path loss, users benefit more from the proposed collaboration relative to direct transmission. Ochiai, Mitran and Tarokh [2] showed the same result in the context of diversity gain which is not in the scope of this paper. This result also appears



(a)



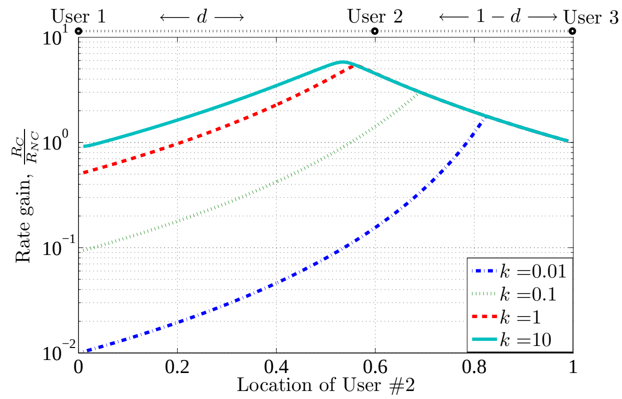
(b)



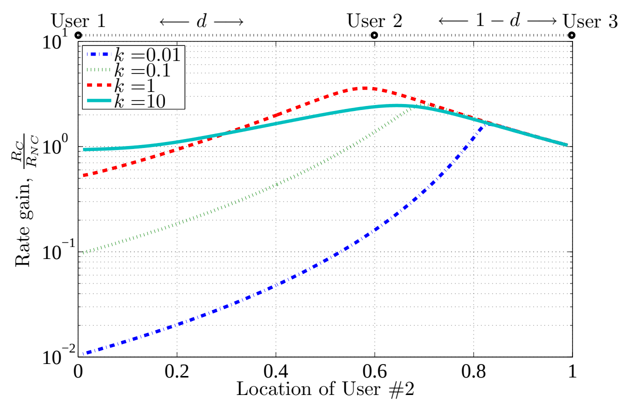
(c)

Fig. 2. Contours of the rate gain $\frac{R_{\text{CP}}}{R_{\text{NCP}}}$ (4), (6) versus relay (2[nd] user) location (x, y) for $\epsilon = 0.01$, $h_{ij} = \frac{1}{d_{ij}^{\eta}}$ and $\eta = 3$, (a) $k = 0.1$, (b) $k = 10$, (c) $k = 100$

very attractive that, in contrast to traditional multi-hopping, appropriately designed collaboration can provide a significant rate gain (see e.g., [3]). Figures 3(a) and 3(b) confirm the above results. This indicates that the best location for the relay user is in the vicinity of the midpoint between the transmitter and the receiver pair. This means that by appropriately selecting the relay user, we efficiently take advantage of the geometrical distribution of users. A relay with optimal location almost achieves (13), which serves for relay selection. Note that by selecting one relay, the multiple relay network becomes



(a)



(b)

Fig. 3. Effect of the relay location d on rate improvement $\frac{R_{\text{CP}}}{R_{\text{NCP}}}$ (4), (6) for $h_{12} = \frac{1}{d^{\eta}}$, $h_{13} = 1$, $h_{23} = \frac{1}{(1-d)^{\eta}}$, for $\eta = 3$, $k = 0.01, 0.1, 1$ and 10, respectively, and different TERN values (a) $\epsilon = 0.01$, and (b) $\epsilon = 0.1$.

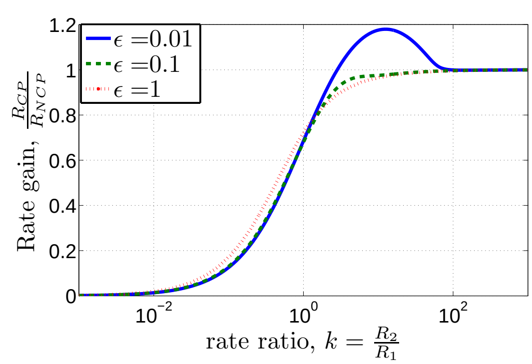


Fig. 4. Effect of rate ratio k on rate improvement, $\frac{R_{\text{CP}}}{R_{\text{NCP}}}$, (4), (6), for $h_{12} = \frac{1}{d^{\eta}}$, $h_{13} = 1$, $h_{23} = \frac{1}{(1-d)^{\eta}}$ for a fixed relay location $d = 0.5$ for $\eta = 3$ and different TERN values $\epsilon = 0.01, 0.1$ and 1.

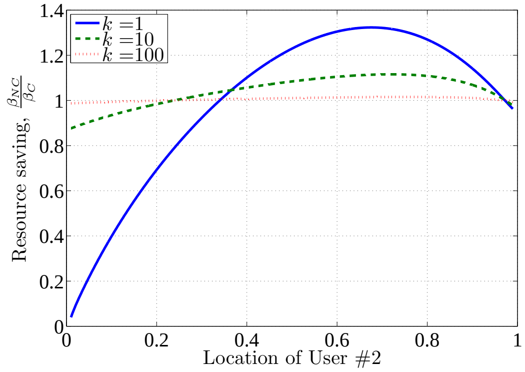


Fig. 5. Ratio of resource usage in CP and NCP $\frac{\beta_{\text{NCP}}}{\beta_{\text{CP}}}$ (18) for $h_{12} = \frac{1}{d^{\eta}}$, $h_{13} = 1$ and $h_{23} = \frac{1}{(1-d)^{\eta}}$ versus relay location d for a required rate of $R = 0.5h_{13}\epsilon$, $\eta = 3$ and $h_{13}\epsilon = 0.01$, and $k = 1, 10$ and 100.

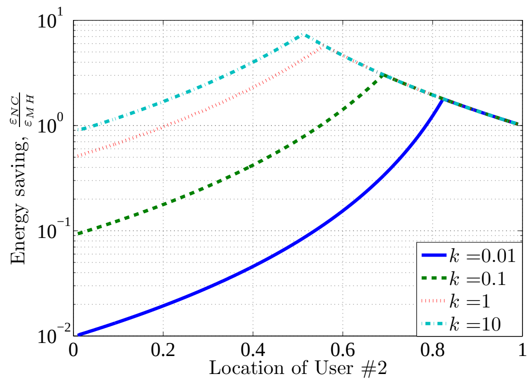


Fig. 6. Ratio of energy usage in CP and NCP $\frac{\epsilon_{\text{NCP}}}{\epsilon_{\text{CP}}}$ (16) for $h_{12} = \frac{1}{d^{\eta}}$, $h_{13} = 1$, $h_{23} = \frac{1}{(1-d)^{\eta}}$ and $\eta = 3$ versus relay location d for unit resource and a given required rate of $R = h_{13}/100$, (a) $k = 0.01, 0.1, 1$ and 10.

a single relay network. Thus, the exact rate improvement or energy saving can be examined as in (6), (4) and (16).

B. Relay Selection: Resource Efficiency

Now, we address resource efficiency and the objective is to select a relay user among the potential candidates and to decide whether to collaborate or not. We propose the following procedure:

- Feasibility check: We compare R with $\epsilon \min\{h_{13}, h_{23}\}$ for the NCP and with $\epsilon \min\{h_{12}, h_{23}\frac{k}{k+1}\}$ for the CP. Then, we ignore the protocol which is not feasible.
- Resource usage: If both are feasible, we must choose the protocols with the least resource usage. The resource usages β_{NCP} and β_{CP} are the solutions of (18). It is worth noting that this criterion is different from the rate and energy criteria since here we are willing to minimize resource usage for a given amount of energy and rate which does not necessarily yields the same result as maximizing rate or energy ratios (compare Figures 5 with Figure 3 and Figure 6 for difference between obtained rate, energy and resource gains). However, in order to maximize the resource efficiency, the simulation result (Figures 5) shows that the best location for the relay user is almost in the middle of the source and the destination user.
- Collaborator selection: Similarly, we can use the resource usages for the criterion to select the collaborator among multiple feasible candidates.

V. Collaboration in General Networks

We can extend the proposed protocols to the general networks where more than one (relay) user are available to relay the messages of multiple users (as source users) toward the different destinations. As we have shown here, we focus on one relay system and look for the best user to serve as relay to maximize the achievable rate, minimize the energy consumption or utilize the available resource more efficiently. To this end, we provide a rough guideline that if direct link channel gain is smaller than the other links, often the CP outperforms the NCP. Otherwise, if a fixed rate is required, the feasibility of different scenarios must be verified. Among feasible solutions, we must choose the protocol and relays which provide maximum rate, or maximize savings on resource (18) or on energy (16). For CP, a relay among possible candidates must be selected which maximizes $\min\{h_{23}k/(k+1), h_{12}\} \gg h_{13}$.

For example, suppose that in Figure 1 the 1^{st} user wishes to send data to the 3^{rd} user, while the 2^{nd} user wishes to broadcast independent messages to the 3^{rd} and 4^{th} users. Using this guideline, the 2^{nd} user can collaborate with the 1^{st} user via acting as relay (the more information, the more incentive to collaborate). In this example the 3^{rd} user has no data to send and thus, ironically, has no incentive to collaborate. So the 2^{nd} user should send his data directly to the 4th user.

So far, we have assumed the same destination for both transmissions. We might relax this constraint easily. For example in Figure 1, suppose that the 1^{st} user wishes to send messages to the 3^{rd} user and the 2^{nd} and 3^{rd} users wish to send messages to the 4th user. Using the CP, the 2^{nd} user can act as the relay between the 1^{st} and 3^{rd} users and the 3^{rd} user acts as the relay between the 2^{nd} and 4th users.

We have shown that collaboration have the potential to increase the rate gain of the users by a factor of at most $\left(1 + \sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta}$. This result shows that appropriately choosing the relay user and collaboration protocol considerably save the transmit energy, and also reduce interference amongst the users. Our proposed protocols not only improves rate, energy or resource utilization of the involved users, but also have the potential to decrease the overall interference of the network. We have shown that collaboration can mitigate the effects of path loss, thus, users can save transmit energy. This saving reduces interference among users which allows to increase density of users in the network through resource reusing. Minimizing the consumed transmit energy will lead to increased node lifetimes in terms of battery power as well. A positive side effect of this is that smaller transmit powers will also reduce the overall interference in the network.

$$\beta=\frac{1}{k+1}+\frac{\frac{1}{k+1}\log\left(\frac{1+(k+1)h_{23}\epsilon}{1+(k+1)h_{13}\epsilon}\right)}{\log\left(\left(1+\left(k+1\right)h_{23}\epsilon\right)\left(1+\left(k+1\right)h_{13}\epsilon\right)\right)-\frac{(k+1)h_{13}\epsilon}{1+(k+1)h_{13}\epsilon}-\frac{(k+1)h_{23}\epsilon}{1+(k+1)h_{23}\epsilon}} \tag{21}$$

VI. Conclusion

We used rate, energy and resource usage as criteria for collaboration and relay user selection. We found the conditions under which the collaboration is preferred for all users. Interestingly, the gain of the users from collaboration in various terms (increase their achievable rate, reduce their transmit energy or use resources more efficiently) can be more significant at low TERN, where the background noise is strong. Clearly, if the background noise is very weak, the collaboration is less attractive. The relative geometrical location of users (i.e., channel responses) must be considered in the relay selection. Very simple criteria are proposed for relay selection. If the relay is in the vicinity of the midpoint between the transmitter and the receiver pair, collaboration can offer good performance. A maximum rate gain (as well as energy saving gain) of up to $\left(1+\sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta}$ can be obtained provided that a collaboration is established with an appropriately located relay, where η is the environment path loss exponent. Furthermore, we present several protocols on how to select the best relay among the possible candidates to maximize the cooperation gain.

Appendix

Proof of (7b): We use the first-order Taylor series approximation at point $\frac{1}{k+1}$ for $R_1(\beta)$ and $R_2(1-\beta)$ which is accurate for high TERN regime. The intersection point of the approximate lines gives an upper bound for achievable capacity for the NCP. The coordinates of this intersection point are given by (21) and (7b) which as noted before is tight for high TERN regime. ■

Proof of (9a): To find a lower bound, we can approximate functions in (4) by their second order Taylor series versus ϵ and obtain $R_{\rm NC} \geq \max\{h_{13}\epsilon - \frac{h_{13}^2\epsilon^2}{2\beta}, h_{23}\epsilon - \frac{kh_{23}^2\epsilon^2}{2(1-\beta)}\}$. To find a tight bound we solve $\left(h_{23}-h_{13}\right)\beta^2+\left(\frac{h_{13}^2\epsilon}{2}+\frac{kh_{23}^2\epsilon}{2}+h_{13}-h_{23}\right)\beta-\frac{h_{13}^2\epsilon}{2}=0$. This quadratic equation has only one feasible solution in the interval $[0,1]$. This bound is described by (9a) and (22)

$$\beta = \frac{(h_{23}-h_{13}-\frac{\epsilon k h_{23}^2}{2}-\frac{\epsilon h_{13}^2}{2})}{2(h_{23}-h_{13})} + \frac{\sqrt{\left(h_{23}-h_{13}-\frac{\epsilon k h_{23}^2}{2}-\frac{\epsilon h_{13}^2}{2}\right)^2+2(h_{23}-h_{13})\epsilon h_{13}^2}}{2(h_{23}-h_{13})} \tag{22}$$

which as noted before is tight for low TERN regime ■

References

[1] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part I. system description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, 2003.

[2] H. Ochiai, P. Mitran, and V. Tarokh, "Variable-rate two-phase collaborative communication protocols for wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4299–4313, 2006.

[3] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, 2005.

[4] Y. Zhao, R. Adve, and T. Lim, "Improving amplify-and-forward relay networks: optimal power allocation versus selection," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3114–3123, 2007.

[5] A. Ibrahim, A. Sadek, W. Su, and K. Liu, "Cooperative communications with relay-selection: When to cooperate and whom to cooperate with?" *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, 2008.

[6] M. Chen, S. Serbetli, and A. Yener, "Distributed power allocation strategies for parallel relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 552–561, 2008.

[7] Z. Zhou, S. Zhou, J.-H. Cui, and S. Cui, "Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3066–3078, August 2008.

[8] W. Huang, Y. Hong, and C. Kuo, "Lifetime maximization for amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5 Part 2, pp. 1800–1805, 2008.

[9] A. Bletsas, H. Shin, and M. Win, "Cooperative communications with outage-optimal opportunistic relaying," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3450–3460, 2007.

[10] E. Beres and R. Adve, "Selection cooperation in multi-source cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 118–127, 2008.

[11] A. Nosratinia and T. Hunter, "Grouping and partner selection in cooperative wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, p. 369, 2007.

[12] T. C.-Y. Ng and W. Yu, "Joint optimization of relay strategies and resource allocations in cooperative cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 328–339, February 2007.

[13] L. Le and E. Hossain, "Cross-layer optimization frameworks for multi-hop wireless networks using cooperative diversity," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, 2008.

[14] S. A. Astaneh and S. Gazor, "Joint protocol and relay node selection in collaborative networks," *2008 24th Biennial Symposium on Communications*, pp. 162–165, 2008.

[15] C. Dovrolis, D. Stiliadis, and P. Ramanathan, "Proportional differentiated services: Delay differentiation and packet scheduling," in *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*. ACM New York, NY, USA, 1999, pp. 109–120.

[16] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity inmobile wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 461–471, 1999.

[17] H. Zhu and G. Cao, "rDCF: A Relay-Enabled Medium Access Control Protocol for Wireless Ad Hoc Networks," *IEEE Trans. Mobile Comput.*, vol. 5, no. 9, pp. 1201–1214, Sept. 2006.

[18] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. Panwar, "CoopMAC: A cooperative MAC for wireless LANs," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, p. 340, 2007.

[19] F. Meshkati, D. Guo, H. Poor, and S. Schwart, "A unified approach to power control in large energy-constrained cdms systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1208–1216, 2008.

[20] D. Goodman and N. Mandayam, "Power control for wireless data," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 48–54, 2000.

Saeed Akhavan Astaneh(S ' 08) Saeed A.Astaneh received the B.S. and M.S. degrees in Electrical Engineering from Sharif University of Technology, Tehran, Iran, in 2003 and 2005, respectively. From 2007 to 2008, he was a researcher at Queen's University, ON, Canada. He is now a Ph.D. student at Queens University. His research interests are in the areas of game theory, cognitive radio, and cooperative communications.

Saeed Gazor (S ' 94 M ' 95 SM ' 98) received the B.Sc. degree in electronics engineering in 1987 and the M.Sc. degree in communication systems engineering in 1989, from Isfahan University of Technology, Isfahan, Iran, both summa cum laude, the highest honors. He received the Ph.D. degree in signal and image processing from the Dpartement Signal (cole Nationale Suprieure des Tlcommunications/ ENST), Telecom Paris, France, in 1994. Since 1999, he has been on the Faculty at Queens University, Kingston, ON, Canada, and currently holds the position of Professor of the Department of Electrical and Computer Engineering. He is also cross-appointed to the Department of Mathematics and Statistics at Queens University. Prior joining Queens University, he was an Assistant Professor with Isfahan University of Technology, Department of Electrical and Computer Engineering, from 1995 to 1998. He was a Research Associate with the University of Toronto from January 1999 to July 1999. His main research interests are statistical and adaptive signal processing, cognitive radio and signal processing, array signal processing, speech processing, image processing, detection and estimation theory, MIMO communication systems, collaborative networks, channel modeling, and information theory. Dr. Gazor has received a number of awards including a Provincial Premiers Research Excellence Award, a Canadian Foundation of Innovation Award, and an Ontario Innovation Trust Award. He is a member of the Professional Engineers Ontario and is currently serving as Associate Editor for the IEEE SIGNAL PROCESSING LETTERS.