Abstract—In this paper, we study how the cooperative relaying can improve both capacity and fairness in cellular network. The capacity and fairness have a trade-off relationship, so increasing cell throughput deteriorates fairness and vice versa. First, we show that the achievable average throughput region can be enlarged by using the cooperative relaying. This enlarged region means that capacity and fairness can be improved at the same time with an adequate scheduling algorithm. Thus, secondly we propose a generalized scheduling algorithm for cooperative relaying. The proposed scheduling algorithm can improve both capacity and fairness at the expense of cooperation among users. From simulations, we show that the trade-off relationship can be surpassed and the unfairness problem in the heterogeneous channel condition can be solved by the opportunistic relaying.

I. INTRODUCTION

Since multi-user diversity concept was proposed in [1], the fading of the wireless channel has been considered as a source of performance improvement. Using the channel state feedback, the scheduler at base station can assign the time slot to a user with the best channel state and this scheme can maximize the cell capacity. However, this makes a user who has a great channel gain monopolizes the whole time slot, so may cause a serious unfairness problem. To compromise these two goals, i.e., capacity and fairness, many scheduling algorithms have been proposed.

Even though the simultaneous increase of both capacity and fairness has been tried in [2]–[4], these works just show that capacity and fairness are in trade-off relationship. To maximize the cell capacity, the fairness cannot be guaranteed, i.e., increasing the cell capacity decreases the fairness, and vice versa. In the modern wireless system which operates at the high frequency, the radio propagations are significantly different between users with the line-of-sight from the base station and those with non-line-of-sight. The more heterogeneous channels make the severer trade-off. Hence, one of two goals should be more sacrificed to maximize the other.

This paper starts from the following question: How can we alleviate the trade-off relationship between capacity and fairness. Moreover, how can we improve capacity and fairness at the same time? The trade-off comes from the fact that more resources should be assigned to the users with bad channel conditions for fairness, thereby decreasing the resources allocated to the users with good channel conditions. This leads to the decreases of the system capacity. We exploit this good but unused channel for other users with bad channels, i.e., cooperative relaying. In this way, we can improve capacity and fairness simultaneously, which cannot be achieved in the traditional scheduling.

Some previous works have shown that the relaying method can improve the performance between the source and destination. In this paper, we consider the cooperative relaying among various relaying methods. By applying the cooperative relaying with an adequate scheduling algorithm, we show that both capacity and fairness can be improved. Even though the link capacity of cooperative relaying has been analyze through several works, to the best of our knowledge, this is the first work to show how this cooperative relaying can be applied to the system level.

We organize the remainder of paper as follows. Section II introduces the cooperative diversity and how relaying can be exploited for opportunistic scheduling. Section III and IV show that the opportunistic relaying scheme can improve the performance in terms of cell capacity and fairness. Section V shows the simulation results verifying the opportunistic relaying algorithm followed by conclusions in section VI.

II. RELATED WORK

A. Cooperative Diversity

The capacity in cooperative network was comprehensively studied in [5], [6]. When source transmits data to the destination, a neighbor node with the better channel gain from the source can relay packet to the destination. While a source and a destination are linked as a direct line in a conventional network, cooperative relaying makes the link as a triangle which is drawn in Fig. 1. In triangular link model, we use X and \( \tilde{X} \) to denote the signals sent from the source and relay, respectively, \( \tilde{Y} \) to denote the signals received at the relay, \( Y_1 \) and \( Y_2 \) to denote the signals received at the destination from orthogonal channels. The relationships between the input and output symbols in the orthogonal channels at each symbol time instant can be written as

\[
Y_1 = \sqrt{g_{RD}} X + N, \\
Y_2 = \sqrt{g_{RD}} \tilde{X} + N, \\
\tilde{Y} = \sqrt{g_{R}} X + N.
\]
where $g_D$, $g_{RD}$ and $g_R$ represent the link gain of direct link to destination, relay link and direct link to relay node respectively. The $N$ and $\hat{N}$ are additive Gaussian noise at destination and relay node respectively. If $Y_1$ and $Y_2$ are from different frequency band, the signals can be decoded without interfering each other.

[6] shows that the lower bound can be maximized by allocating the channel fraction optimally which is assigned for relaying channel. The lower bound on the capacity of cooperative networking is given by

$$C_{\text{coop}} \geq C_{\text{low}} = \max_{0 \leq \alpha \leq 1} \min \left\{ \alpha \log_2 \left( 1 + \frac{g_D P_S}{N \alpha} \right), \alpha \log_2 \left( 1 + \frac{g_{RD} P_R}{\hat{N} (1 - \alpha)} \right) \right\},$$

(2)

where $\alpha$ is the channel resource allocation ratio to the broadcasting channel of source, $P_S$ and $P_R$ are transmit power of source and relaying node respectively. The $(1 - \alpha)$ portion of the resource is allocated to the relaying channel. This cooperative relaying can achieve a rate that is at least equal to the capacity of the direct link i.e. $\log_2 (1 + g_D P_S/N)$. Furthermore, the lower bound approaches to the data rate by the classical relay channel i.e. $\log_2 (1 + g_R P_R/\hat{N})$ as the channel gain between relay node and destination goes to infinity.

When this cooperative relaying is applied to the cellular network system, cell throughput may increase highly. However, the improvement should come along with an adequate scheduling scheme. When a scheduler decides to send data through triangular link, it should utilize the good channel of relay node. Then, the relay node which has a good channel state should lend its opportunity to the other node for the capacity increase. However, when there are packets to send to the relaying node, this opportunity should be used for the relaying node as well. Hence, the scheduler should be proposed to utilize the channel opportunity for the adequate user.

B. Opportunistic Scheduling

Then, how can the relaying channel be used for the cell-wise performance improvement? The good relaying channel should be used for both capacity and fairness improvement. Thus far, proposed opportunistic schedulers [2]–[4] assume

that the achievable data rate per each user is given according to its own channel gain. By exploiting cooperative relaying, each user can have 1 direct link and $(K - 1)$ triangular links, if $K$ users are in a cell and the users are totally connected one another. This means that the base station can utilize the more opportunities by increasing the set to schedule. From this point of view, we show that the multi-user diversity gain can be increased by user cooperation.

III. MODEL DESCRIPTION

A. System Model

We consider the downlink in a cellular network with cooperative relaying system. A single broadband channel is shared by all users in TDMA (time division multiple access) manner, that is, the base station transmits to only one destination at any given time slot using one relaying node in case. The link gain is varying independently among users and is fixed during a time slot. The base station communicates with $K$ mobile users and uses the feedback of channel conditions to measure the exact achievable data rate for each user. The base station maintains a packet queue for each user, and there is always enough data in the queue, i.e., infinite backlog in the queue.

At the beginning of time slot, the base station gets a direct channel gain $g_i^D$ and relaying channel gain vector $g_{ik}^R = [g_{i1}^R, \ldots, g_{iK}^R]$ from each user, where $g_{ik}^R$ is the relaying channel gain between user $i$ (destination) and user $k$ (relaying node).

If the base station transmits the packets to the user $i$ through the direct link, the achievable data rate is given by

$$R_i^D = \log_2 \left( 1 + \frac{g_i^D P_{BS}}{N} \right),$$

(3)

where $P_{BS}$ is the transmit power of base station, $N$ is the additive white Gaussian noise and $g_i^D$ is direct channel gain. If base station transmits the packets to the user $i$ through the triangular link using user $k$ as a cooperative relaying node, the lower bound of achievable data rate is given by

$$R_{ik}^T = \max_{0 \leq \alpha_{ik} \leq 1} \min \left\{ \alpha_{ik} \log_2 \left( 1 + \frac{g_{ik}^R P_{BS}}{N \alpha_{ik}} \right), \alpha_{ik} \log_2 \left( 1 + \frac{g_k^D P_{BS}}{N \alpha_{ik}} \right) \right\},$$

(4)

where $\alpha_{ik}$ is a frequency allocation fraction in case that user $i$ is the destination and user $k$ roles as a cooperative relaying node. $\hat{N}$ is additive Gaussian noise and $P_R$ is the transmit power of relaying node. We assume that the scheduler uses the triangular link only when the channel gain to the relaying node is better than the channel gain to the destination i.e. $g_k^D > g_i^D$. Under this assumption, the optimal frequency allocation $\alpha_{ik}^*$ can be found as shown in [6]. This achievable data rate is lower bound, but we consider this as an achievable capacity for the triangular link and show the performance improvement. Actually, the capacity improvement might be greater than the lower bound.

After the base station calculates the available data rate through direct link and triangular link for each user from (3)
and (4) respectively, the base station determines the best link for each user \( i \) as follows:

\[
R^E_i = \max_{k \in K'} \left\{ R^D_i, R^\Delta_i \right\}, \quad \forall i, \tag{5}
\]

where \( K' \) is the set of \( \{ k : g^D_k > g^D_i \} \). We call \( R^E_i \) as an effective data rate which means the highest data rate that the user \( i \) can achieve through the direct link or triangular link. In the case that the triangular link is selected, the base station keep information about the relaying node \( k \) and the frequency allocation fraction \( \alpha^{\Delta}_{ik} \) and assign a time slot following the relaying node and frequency allocation fraction.

**B. Characterization of the Average Throughput Region**

In this section, we will obtain an explicit characterization of the average throughput region. An average throughput vector \( \mathbf{T} \) is defined as

\[
\mathbf{T} = (T_1, T_2, \ldots, T_K) \in \mathbb{R}^K, \tag{6}
\]

where \( K \) is the number of users in a cell and \( T_i \) is the average data rate of user \( i \). Because only one user is served during a time slot by the scheduling policy, the average data rate is updated as

\[
T_i(n) = \frac{\sum_{l=1}^{n} R_i(l) 1_{\{i = i^\ast(l)\}}}{n}, \tag{7}
\]

where \( R_i(l) \) is available data rate for user \( i \) at time slot \( l \). Thus, \( R_i(l) \) is \( R^D_i(l) \) or \( R^E_i(l) \) when the system uses direct link only or triangular link respectively. Besides, \( 1_{\{i = i^\ast(l)\}} \) is the indicator function of event that user \( i \) is selected at time slot \( l \) by scheduling policy. From the TDMA assumption that at most only one user is selected as a destination, we have an allocation constraint: \( \sum_{i=1}^{K} 1_{\{i = i^\ast(l)\}} = 1 \) or 0 at every time slot \( l \). We show how the long term average throughput region is changed by using the triangular path other than the direct links only.

Let \( C^D_\Pi \) denote the long term average throughput region under the allocation constraint. \( C^D_\Pi \) consists of the average data rate vectors obtained by all possible stationary resource allocation schemes. For the only direct link case and with triangular link case, the data rate region can be expressed as below respectively

\[
C^D_\Pi = \left\{ (T^D_1, \ldots, T^D_K) \in \mathbb{R}^K \right\}
\]

\[
: T^D_i = \frac{1}{n} \sum_{l=1}^{n} R^D_i(l) 1_{\{i = i^\ast(l)\}}, \quad i = 1, \ldots, K, \pi \in \Pi \right\}
\]

\[
C^\Delta_\Pi = \left\{ (T^\Delta_1, \ldots, T^\Delta_K) \in \mathbb{R}^K \right\}
\]

\[
: T^\Delta_i = \frac{1}{n} \sum_{l=1}^{n} R^\Delta_i(l) 1_{\{i = i^\ast(l)\}}, \quad i = 1, \ldots, K, \pi \in \Pi \right\}
\]

for all \( n \) and \( \Pi \) is the set of all possible scheduling policies which do not violate the allocation constraint.

The achievable data rate region \( C^D_\Pi \) is proved to be a convex set in [2]. Following the proof, it is easy to prove that \( C^\Delta_\Pi \) is also a convex set. Thus, we skip the proof in this paper. Besides the convexity, this achievable data rate region has the following property that shows the relation between the achievable average throughput regions by only the direct link and triangular link.

**Proposition 1.** The achievable data rate region \( C^\Delta_\Pi \) is a superset of \( C^D_\Pi \) at a given channel condition.

**Proof.** Assume that \( \pi^D \) is the scheduling policy which achieves one of the (upper-right) boundary point in direct link case. The scheduled user at each time slot \( l \) by policy \( \pi^D \) can be expressed as \( i^\ast(l) \). If we allocate the same time slot \( n \) to the same user \( i^\ast(l) \) using effective data rate other than direct link, the achieved data rate vector will be increased, because we have \( R^D_i(l) \leq R^E_i(l) \). Thus, \( C^D_\Pi \subseteq C^\Delta_\Pi \) holds.

As long as there are some time slots where the triangular link gain is larger than the direct link gain, the achievable data rate region is proper superset of one using only direct link. Additionally, there are some characteristics of the achievable data rate region using the triangular links.

1. As the relaying channel states get better, the achievable data rate region stretches outward. Fig. 2 illustrates this.

2. One of the boundary points in \( C^D_\Pi \) cannot be stretched outward by using triangular links. This point \( A \) in Fig. 2 is the rate vector \( T^{D^\ast} \) which is achieved by the scheduling policy \( i^\ast = \arg \max \{ R^D_i \} \) at every time slot, i.e., max C/I scheduler. No matter how much the relaying channel state get better, the achieved rate vector using triangular link \( T^{\Delta^\ast} \) by scheduling policy \( i^\ast = \arg \max \{ R^\Delta_i \} \) is equal to \( T^{D^\ast} \). This is because \( \max \{ R^D_i(l) \} = \max \{ R^E_i(l) \} \) at every time slot \( l \).

3. As the relaying channel states get better and eventually \( G^R_{ik} \) goes to infinity, the rate region stretch outward until every intercept with the each users’ axis are \( \sum_{l=1}^{n} \max \{ R^E_i(l) \} / n \).

The achievable average throughput region is enlarged by using
the triangular link compared to the one by using only direct link. The enlarged data rate region means that the achieved data rate vector using triangular link can surpass the capacity and fairness trade off relationship of direct link case.

To attain data rate vector which reside in such an enlarged region, an adequate scheduling policy is required. Thus, we propose scheduling policies which achieve those data rate vectors in the next section.

IV. SCHEDULER DESIGN

A. General Scheduler Using Cooperative Relaying

Kushner et al. [7] shows that the achieved throughput vector by Proportional Fairness (PF) scheduler resides on the boundary of achievable throughput region $C^D_{\Pi}$ in (8). Then, how can we make the achieved throughput vector reside on the boundary point of $C^D_{\Pi}$? To achieve this goal, we need to design a noble scheduler that jointly decides:

- **User selection**: Which user will be selected?
- ** Relay node selection**: Whether the selected user will utilize a neighbor relay node or not? If yes, which relay node will be utilized?

For example, a scheduling policy which allocates the same time slot to the same user with the direct link can achieve the better capacity and fairness, if the base station uses the triangular link instead of direct link. Even though the performance can be improved by allocating the time slots in this way, it is not difficult to see that such a scheduler is not optimal in that there can exist a better choice for given channel condition.

We propose an optimal scheduling algorithm, provided that the channel gain is ergodic and stationary. This assumption is necessary for the convergence and optimality of gradient-based scheduler [7], which we will apply to develop our scheduling policy. The following proposition presents the proposed scheduling algorithm and its properties. The proof is straightforward following the one in [7].

**Proposition 2.** The achieved average throughput vector $\mathbf{T}$ is the solution of problem $\mathbf{P}$ if and only if the following assignment strategy is used at every time slot

$$i^* = \arg \max \frac{R^E_i}{(T^i)^\alpha} \quad \text{where} \quad R^E_i = \max \{ R^D_{i_k}, R^{\Delta}_{i_k} \}, \quad \alpha \neq 1,$$

$$\alpha \neq 1,$$

$$\mathbf{P}: \max_{\pi \in \Pi} U(\mathbf{T}^\pi) \tag{10}$$

where

$$U(\mathbf{T}^\pi) = \begin{cases} \sum_i \frac{1}{1 - \alpha} (T^\pi_i)^{1-\alpha}, & \alpha \neq 1, \\ \sum_i \log(T^\pi_i), & \alpha = 1. \end{cases} \tag{11}$$

Besides, the achieved data rate vector is on the boundary of the data rate region $C^\Delta_{\Pi}$.

This proposed scheduler is easy to adjust capacity and fairness trade-off relationship by varying $\alpha$, while the achieved point is on the boundary of capacity region and satisfy the alpha fairness in [8].

B. PF Schedululer Using Cooperative Relaying

We call the scheduling algorithm with $\alpha = 1$ in (9) as opportunistic relaying scheduler (OR). In this section, we examine the characteristics of achieved rate vector by OR and compare it with one by PF scheduler. The next proposition explains the properties of capacity and fairness when the triangular link is used, that is, when users cooperate each other opportunistic.

**Proposition 3.** The logarithm sum of achieved average throughput by OR is always larger than one using only the direct links by PF.

*Proof.* From Proposition 2, the achieved point is on the boundary and maximizes the log sum of achieved average throughput. Because the utility function is strictly concave and $C^\Delta_{\Pi} \supseteq C^D_{\Pi}$, it is straightforward that $U(\mathbf{T}^{OR}) \geq U(\mathbf{T}^{PF})$, where $\mathbf{T}^{OR}$ and $\mathbf{T}^{PF}$ are achieved average throughput vectors by OR and PF scheduler respectively. From the previous section, $C^\Delta_{\Pi}$ is the proper superset of $C^D_{\Pi}$ when there are some triangular link with the better gain. Thus, $U(\mathbf{T}^{OR}) > U(\mathbf{T}^{PF})$, unless the optimal point is on the point where the boundary of $C^D_{\Pi}$ and $C^\Delta_{\Pi}$ converge. ■

The average data throughput that is achieved by OR scheme is different from the conventional PF scheduler. In the sense of original channel, this achieved average rate by OR does not mean proportional fair allocation any more. However, from the effective channel point of view, the achieved average data throughput satisfies the proportional fairness.

This opportunistic relaying scheduler can improve the performance by handing over the good opportunity to the mobile which has comparably weak channel for a long time. From this property of OR, we can show how OR improves the performance at the extreme situation.

**Proposition 4.** When the channel gains between the mobile nodes go to infinity and the mobiles are fully connected each other, the achieved data rate by OR scheduling scheme satisfies the max-min fairness. At the same time, the total throughput is same as the max CI scheduler.

*Proof.* Because the relaying channel gain is infinite, the triangular link gain for a user $i$ is determined by the best channel gain between base station and the $i$’s neighbor. Furthermore, this value is same as the highest channel gain given in a cell, because we assume that the mobiles are fully connected each other. At every time slot, the effective channel gain for all users is same with the maximum value of users’ direct channel gains because every user can use it as effective channel gain of themselves through cooperative relaying.

Consequently, the scheduling priority metric is changed into

$$i^* = \arg \max_{\pi \in \Pi} \max \{ R^D_{i_k} \} / T^i \tag{12}$$

Thus, at every time slot, scheduler chooses user $i$ who has got the least amount of data so far. This implies the max-min fairness scheduling. Furthermore, the scheduler adds to the cell capacity with the maximum value which the system can support at every single time slot. Consequently, the total
V. SIMULATION

Through various simulations, we try to show how much cooperative relaying can improve the system performance such as capacity and fairness. We use the fairness index which was used in [9]. This index is defined as

$$\frac{\left(\sum_i T^*_i T_i\right)^2}{\left(\sum_i T_i^2\right)\left(\sum_i T_i^2\right)}$$

where $T^*$ is the max-min fair throughput allocation and $T_i$ is the achieved data rate for user $i$.

Fig. 3, Fig. 4 shows the capacity and fairness are in trade-off relationship in cellular networks. When alpha is 0, the system capacity is maximized, but it has the worst fairness performance. In Fig. 4, we cannot illustrate the fairness index when alpha is 0 because of the log scale. The value of fairness index is 0.364. By increasing the value of alpha in (9), the achieved average throughput is decreased, while the fairness index is increased. At the same value of alpha, both capacity and fairness index using the cooperative relaying is higher than the case without the cooperative relaying. From the excellence in capacity and fairness, we can conclude that the system exploiting the cooperative relaying can alleviate the trade-off relationship, and can find the improved region in terms of capacity and fairness.

To investigate the performance per user closely, we plot the average throughput per user with respect to Round-Robin (RR), PF and OR in Fig. 5. We note that under RR, the throughput achieved by a user is proportional to its average channel gain, because RR does not consider the channel state and allocate the same fraction of time to all users. First, we can see that each user gets higher throughput under PF than under RR. The fraction among users follows the fraction of RR because the PF scheduler guarantees achievable data rate proportional to their average channel gain. In OR, the user who gets the relatively good channel state such as 4, 6 and 8 users get little smaller amount of capacity than in PF, while the user who gets the relatively weak channel state such as 2, 9 and 10 users get higher capacity than in PF. This result implies that the opportunity of users with strong channel gain can be
Var the average channel gain of whole users, i.e., shows how widely we scatter the users in a cell. The value of distance from the base station. In Fig. 6, the horizontal axis get different channel statistics by differentiating the mobiles’ distance attenuation and Rayleigh fading, we can get a lot of burden to the system. The reduction of the best link to assign a time slot. The channel gain feedback can cause a lot of burden to the system. The reduction of the amount of feedback could be an interesting future study.

Because the improvement of OR is based on borrowing and lending of good channel, the gain is increased as the channel becomes more heterogeneous. We vary the variance of average channel gain among users to change heterogeneity in the system. First we make users in the same distance from the base station so that they have homogeneous channel statistics. All users have identical Rayleigh fading statistics and distance attenuation. After then, we spread them over the cell so that they have the heterogeneous channel states by varying the distance attenuation while we maintain the statistics of users’ Rayleigh fading. Because we assume that the channel consists of distance related attenuation and Rayleigh fading, we can get different channel statistics by differentiating the mobiles’ distance from the base station. In Fig. 6, 7, the horizontal axis shows how widely we scatter the users in a cell. The value indicates the ratio of the variance of average channel gain to the average channel gain of whole users, i.e., $\frac{\text{Var}(g^2)}{E[g^2]}$.

Fig. 6 shows the tendency that the capacity decreases as the heterogeneity increases. However, the decrement of OR is marginal compared to those of other schedulers. As the channel gets more heterogeneous, there are more possibilities to use the cooperative relaying to overcome the weak channel state. Even though PF scheduler tries to serve the user having relatively good channel state, the user having (even severely) bad channel conditions also grab the chance to be served because the PF scheduler metric is inversely proportional to the empirical throughput. In the same situations, OR can borrow the good channel to overcome the weak channel problem without harming the user with good channel state. Unless the weak channel user uses the good channel, it is not used for the system at that time slot at all. This good channel which is not used in the traditional scheduling can be used for cooperation. This can increase the fairness too as shown in Fig. 7. While the fairness in RR and PF is degraded as the channels get more heterogeneous, OR can maintain fairness due to the cooperative relaying.

VI. CONCLUSION AND DISCUSSION

In this paper, we propose an opportunistic relaying scheme aiming to resolve two challenging issues in wireless network, i.e., capacity and fairness. Based on channel gain information of relay paths, our noble scheduler construct the effective channel gain and utilizes this gain vector opportunistically for downlink scheduling. This cooperative relaying can enlarge the achievable capacity region and our scheduling algorithm achieves the boundary point of enlarged capacity region. Simulation results show that the unfairness problem can be solved by handing over the strong channel opportunity to the user with poor channel gain.

Throughout the paper, we assume that the scheduler at the base station knows every channel gain, so it can choose the best link to assign a time slot. The channel gain feedback can cause a lot of burden to the system. The reduction of the amount of feedback could be an interesting future study.

REFERENCES