In Proc. of MedHocNet 2004, Bodrum, Turkey, June 2004.

Capacity Increase of Multi-hop Cellular WLANs Exploiting Data Rate Adaptation and Frequency Recycling

Seble Mengesha, Holger Karl, Adam Wolisz
Fachgebiet Telekommunikationsnetze, Technische Universität Berlin
mengesha|karl|wolisz@ee.tu-berlin.de

***Abstract*— The use of intermediate terminals for relaying traffic in infrastructure based wireless systems promises improvements in capacity, energy efficiency and coverage. This paper studies the possibility to increase capacity using intermediate mobile terminals as relayers, especially when using a second frequency to perform relaying in addition to a cell's primary frequency. This second frequency is obtained by "recycling" it from neighboring cells. These recycled frequencies are used at a reduced transmission power only in the interior of the cell, close to the access point. We show the trade-off between increased transmission parallelism in one cell and increased interference from other neighboring cells. We present algorithms that solve the two-frequency relaying problem, combining routing, scheduling and data rate adaptation while maintaining fairness in the system. For some combinations of primary and secondary frequency assignments ("macro" and "micro" frequency reuse patterns), we obtained up to 40% gain in capacity.**

I. Introduction

In wireless communication systems, two communicating terminals can be too far apart to reasonably allow direct communication at high data rates. A typical example is a terminal in a wireless cell which is far away from its access point (AP). Reducing this distance by relaying the traffic via intermediate terminals can be beneficial regarding the total capacity, even in cellular-type networks—capacity here understood as the total amount of data transmitted or received by the access points(s) per unit time. The intuition for increased capacity is that over shorter distances, faster data rates (due to adaptive modulation and coding) can be realized. As long as the data rates over the individual links of a two-hop route are more than twice as fast as a direct communication, capacity increases (Section III shows an example). Hence, *rate-adaptive relaying* is one crucial technique for capacity improvement.

Multi-hop communication provides an additional advantage over direct communication in a cellular context: There are potentially more sender-receiver pairs than in the direct case, where the access point always participates and there is only a single pair at a given point in time (disregarding CDMA-based systems for the moment). Hence, relaying offers the possibility to increase capacity even further by *concurrently* transmitting data, in multiple hops, towards the access point. Combined with terminals close to the access point using faster data rates than far-out terminals could use, a pipeline effect should result and increase capacity.

Scheduling such simultaneous transmissions within a cell on a single frequency band (frequency, for short), however, is hardly feasible due to the then increased interference and packet error rate within a cell. But when using a second frequency to schedule concurrent transmissions, they can proceed undisturbed and a considerable gain in capacity per cell can be reached [12].

A second frequency, however, is usually not easily available. A possible caveat would be that doubling the total bandwidth by adding the second frequency should also be done for the direct communication, and only then the performance could be fairly compared. This is a valid concern when looking at a single cell in isolation. However, even in a single cell, two-frequency relaying is a valid option as commodity radio equipment plainly does not support doubling the link bandwidth and direct communication could thus not take advantage of a second frequency if it were available. But more importantly, such frequency (re-)allocations have to be considered not on the level of an individual cell but on a system level.

In typical cellular systems, a total amount of bandwidth is given and it is split into a certain number of frequencies bands that are assigned to individual cells following a frequency reuse pattern. Hence, all available frequencies are (usually) already assigned. An additional frequency for one cell is therefore only available by "borrowing" or "recycling" it from a neighboring cell. An example of such frequency recycling is shown in Figure 1. Such recycling indeed has a viable chance to increase the overall system capacity, when looking at many cells: This second, recycled frequency could be used only in the interior of a cell, far away from the cell it originally belonged to and close to the access point. Because of this closeness, only low transmission power is needed to overcome the short distance, limiting the additional interference in the lending cell. Consequently, there is a trade-off between the increased interference that this recycling will cause on the overall system level and the capacity gains within a cell that can be reached by two-frequency relaying. In fact, this redistribution of frequencies can be regarded as a new option for resource management, enabled by relaying.

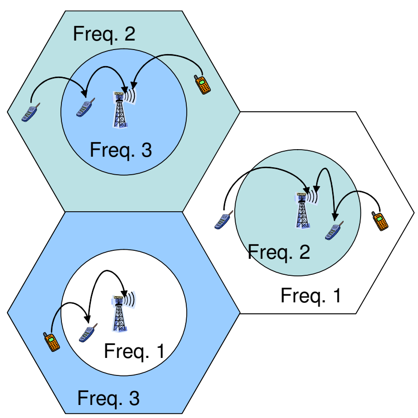


Fig. 1. An example frequency recylcing

Our contribution is to describe such a two-frequency relaying with frequency recycling. We will characterize the performance trade-offs resulting from different ways of (re-)distributing frequencies and discuss frequency reuse patterns appropriate for two-frequency relaying. We will show that relaying admits possibilities for managing frequencies that have a considerable performance gain over direct-only communication systems.

The following Section II outlines our system model. Section III then describes how we combine data rate and transmission power adaptation, scheduling, routing, and frequency recycling to improve capacity in a wireless cell. Some performance results for this scheme are given in Section IV, Section V puts our approach in perspective with related work. Finally, Section VI concludes the paper.

II. System model

To evaluate the combined potential of rate-adaptive relaying with frequency recycling, three aspects of the system are crucial: The physical layer, the MAC structure, and the centralized organization to take decisions.

A. Physical layer

To enable rate-adaptive relaying, a physical layer is required that provides several data rates to choose from, trading off speed against communication distance. As a case study, we use here the physical layer of HiperLAN/2 [8], which provides seven different data rates (modulation plus coding rate). Since this physical layer is very similar to the one employed by IEEE 802.11a, we expect these results to be directly generalizable.

Moreover, we make some simplifying assumptions: We neglect adjacent channel interference and only consider co-channel interference in the evaluations of Section IV; we assume the frequency re-tuning times of a transceiver to be negligible; and we equate channel gains with the distance between sender and receiver. All of these simplifications can be quite easily generalized. In Section IV, the channel gains will also be used to compute signal-to-interference ratios for all data packets, which are converted to uniform packet error rates.

B. MAC structure

The essential choice for a MAC structure is—in the present context—between a contention-based and a scheduled medium access approach. Contention-based systems are currently more popular, but for the question at hand, the particular medium access control *within* a cell should not play the dominating role. A scheduled, TDMA-based access control allows as a finer-grained control over how resources are used and assigned within a cell, admitting deeper insights into the structural properties of our presented concept. In particular, it allows us explicit control over how resources are split among and assigned to terminals within a cell, something that is not immediately possible with a CSMA structure.

Therefore, we decided to use a TDMA-based MAC structure to be able to control resource assignments in detail. Like the physical layer, the details of the MAC correspond to HiperLAN/2—a frame of 2 ms, which is split into slots, assigned to a specific sender–receiver combination (which does not have to include the AP).

C. Centralized organization

Commensurate with the choice of a TDMA-based medium access is our decision to use a per-cell centralized decision making process. The access point only has information about terminals in its own cell—especially, channel gain estimates between terminals—and bases all its decision on this information. In particular, the AP decides in which time slot a given sender and receiver (both belonging to its own cell, including the AP itself) should communicate, which transmission power and which data rate to use. The length of a slot can vary, depending on the chosen data rate.

Channel gain estimates can be obtained by the AP in various ways, e.g., using a concept similar to HiperLAN/2's "radio map" or using approaches like optimistic rate adaptation, e.g. similar to [6] (occasionally switching to a faster rate to check whether communication is still possible and falling back to a slower rate if such an attempt fails).

The advantage of such a centralized organization is a much better control over the system's behavior, giving better insights into the potential capacity contribution of rate-adaptive relaying combined with frequency recycling. How these benefits work out with a distributed organization scheme is objective of current work.

III. Optimizing capacity

A. Example scenario

Consider a simple cellular network with three terminals A, B, and C as shown in Figure 2; assume that B is a potential relayer for A and C is communicating only directly with the AP. Assume also that the medium is shared equally between A, B, and C, resulting in a net data rate of 2, 12, and 9 MBit/s for each terminal, respectively.

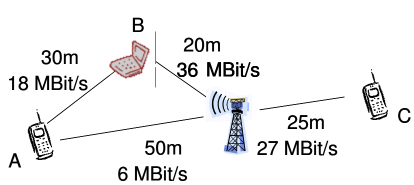


Fig. 2. A simple cellular network scenario

Evidently, A can improve its direct rate from 2 MBit/s to at least 3 (=18/2/3) MBit/s if the time slot for the A to AP communication is split in two equal parts, used for communicating from A to B at 18 MBit/s and from B to AP at 36 MBit/s (the division by two accounts for the need to transmit A's packet twice, the division by three models the sharing of the medium with B's and C's traffic). Splitting the time slot into sub-slots of length 2:1 between the A–B and the B–AP links results in a net data rate for A to the AP of 4 = (18*2/3)/3 = (36*1/3)/3 MBit/s. This *rate-adaptive relaying* has doubled A's net data rate.

Moreover, when using relaying, the AP is idle when A transmits to B. This time can be used for C's transmission to the AP. To schedule these transmissions concurrently, a second frequency is necessary (doing so in the same frequency is sometimes, in very large cells, possible, but in general is not advantageous because of the increased intra-cell interference [2]). The remaining question is then how to organize the frequency recycling pattern.

In summary, we have to solve (a) a transmission power and rate adaptation problem for each individual link; (b) a routing problem to select, for each terminal, a relaying terminal or to decide not to relay at all; (c) a scheduling problem how to assign time slots based on the individual links' data rates and on the option of a second frequency; and lastly (d) we have to find proper frequency recycling patterns. The following sections will provide these solutions. Currently, the algorithm optimizes only the uplink case and considers a single intermediate relay; generalizations are straightforward.

B. Transmission Power and Rate Adaption

The potential benefit of relaying depends on the data rates that can be realized between relayed terminal, relaying terminal, and access point. The effective data rate between any two terminals can be determined based on their channel gain (derived from the distance in the simulation models below) and a target packet error rate which allows to compute, for each data rate, the required transmission power, using the (approximately) known relationship between signal-to-noise and packet error rate [8].

Any data rate that requires more than a maximum allowable power (here, 200 mW) or that does not match minimal required receiver sensitivity is ruled out. As a result, the optimal data rate for this pair of terminals is obtained. The smallest transmission power is used that still meets the target PER for this data rate.[1] An example of such an effective-data-rate selection is shown in Figure 3 for 1% target PER. The shape of this Figure 3 is the primary justification for hoping that relaying will actually improve capacity: *For many distances, the achievable data rate over half the distance is **more than twice** the data rate over the full distance.* This observation holds for other α and target PERs as well.

[1]It is trivial to add a fading margin to this computation to protect against Rayleigh-type fading.

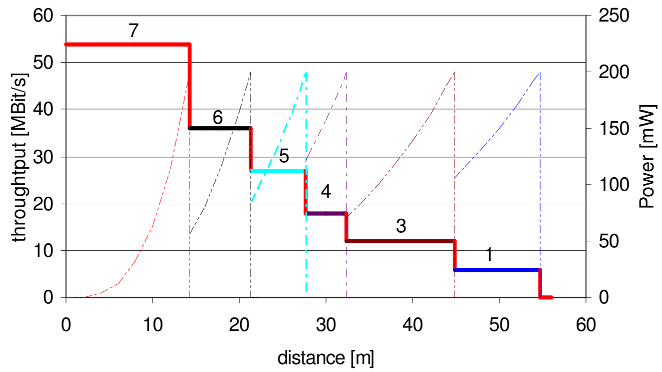


Fig. 3. Effective data rate and transmission power with respect to distance for 1 % PER and $\alpha = 3.2$

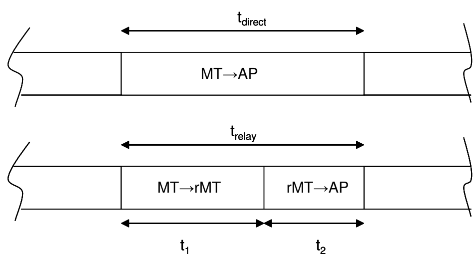


Fig. 4. slot for direct and relay transmission

It should be pointed out that this transmission power/ rate adaptation introduces another tradeoff: By switching to higher rates over shorter distances, it can be necessary to increase the transmission power (we are *not* looking at or trying to optimize energy efficiency in this paper). This increased power in turn increases interference in neighboring cells. Thus, it is not clear whether, in a multi-cell context, this scheme actually provides any benefits—the simulation models used for Section IV do account for these effects.

C. Routing

The routing procedure considered here is relatively simple as we are looking at a cellular context (for each terminal, the destination AP is given), we only consider using at most one intermediate relay (which is reasonable in a cellular-type environment), and we assume a centralized organization. Suppose a far terminal MT has data size y to transmit and this data is directly transmitted to the AP in a time slot of length t_{direct} as shown in Figure 4. Suppose also that there exists a candidate relay mobile terminal r_{MT} for which to decide whether to relay via it or not. Assuming that all terminals fairly share the communication time, in order to use r_{MT} as a relay, we require for a candidate relay terminal:

1) data y can be sent from MT to r_{MT} in time slot t_1
2) relayed data y can be sent from r_{MT} to AP in time slot t_2
3) $t_1 + t_2 = t_{\text{direct}}$
4) $\text{GP}_{\text{relay}} > \text{GP}_{\text{direct}}$

where $\text{GP}_{\text{direct}}$ and GP_{relay} are the effective data rates for the direct and relay transmission, respectively. These data rates are computed as follows.

1) Direct case
The effective data rate for the direct transmission is the correctly transmitted data for a given time, i.e

$$\text{GP}_{\text{direct}} = (1 - \text{PER}_{\text{MT}\to\text{AP}})\text{Rate}_{\text{MT}\to\text{AP}}$$

where $\text{PER}_{\text{MT}\to\text{AP}}$ is the packet error rate from MT to the AP.

2) Relay case
When splitting up the direct time slot t_{direct} into two sub-slots, the question arises how to choose their lengths t_1 and t_2 such that the overall rate is maximized (Section III-A gave an intuition).[2] Leaving aside the error rates on both links for the moment, a simple consideration shows that the data transmitted over both links should be equal to maximize throughput – otherwise, there would be too much or too little data arriving at the relay terminal. Hence:

$$t_1 \text{Rate}_{\text{MT}\to r_{\text{MT}}} = t_2 \text{Rate}_{r_{\text{MT}}\to\text{AP}}$$
$$\Leftrightarrow \frac{t_2}{t_1} = \frac{\text{Rate}_{\text{MT}\to r_{\text{MT}}}}{\text{Rate}_{r_{\text{MT}}\to\text{AP}}} \tag{1}$$

This means that the $t_1 \text{Rate}_{\text{MT}\to r_{\text{MT}}}$ amount of data can be transported from the far terminal to the AP in $t_{\text{direct}} = t_1 + t_2$, resulting in a total throughput of

$$\frac{t_1 \text{Rate}_{\text{MT}\to r_{\text{MT}}}}{t_1 + t_2} = \frac{\text{Rate}_{\text{MT}\to r_{\text{MT}}}}{1 + \frac{t_2}{t_1}} \overset{\text{Eq. (1)}}{=}$$
$$\frac{\text{Rate}_{\text{MT}\to r_{\text{MT}}}}{1 + \frac{\text{Rate}_{\text{MT}\to r_{\text{MT}}}}{\text{Rate}_{r_{\text{MT}}\to\text{AP}}}} = \frac{\text{Rate}_{\text{MT}\to r_{\text{MT}}} \text{Rate}_{r_{\text{MT}}\to\text{AP}}}{\text{Rate}_{\text{MT}\to r_{\text{MT}}} + \text{Rate}_{r_{\text{MT}}\to\text{AP}}}$$

[2] Strictly speaking, this is already a scheduling decision.

Taking the error rates on each link into account, the relay goodput is then:

$$\mathrm{GP_{relay}} = (1 - \mathrm{PER}_{\mathrm{MT}\to r_{\mathrm{MT}}}) \cdot (1 - \mathrm{PER}_{r_{\mathrm{MT}}\to\mathrm{AP}}) \cdot \frac{(\mathrm{Rate}_{\mathrm{MT}\to r_{\mathrm{MT}}})(\mathrm{Rate}_{r_{\mathrm{MT}}\to\mathrm{AP}})}{(\mathrm{Rate}_{\mathrm{MT}\to r_{\mathrm{MT}}}) + (\mathrm{Rate}_{r_{\mathrm{MT}}\to\mathrm{AP}})}$$

Finally, among the candidate relay terminals, the one with maximum $\mathrm{GP_{relay}}$ is selected as a relaying terminal for MT in consideration. Note that there is no issue of "overloading" a relay terminal as the time necessary to do the relaying comes out of the far terminals allocated time slot.

The routing decision takes place independent of interference as the AP has little or no knowledge about ongoing transmissions in other cells. Later on, when terminals sense the presence of interference, they readjust their data rate by first recalculating their actual PER. There can be situations where terminals can no longer transmit their traffic due to excessive interference. Such cases may reduce the overall throughput considerably as they are already scheduled for communication.

D. Scheduling

Before scheduling can be done, the question how to fairly arbitrate resources has to be solved.

1) Fairness schemes: Following the routing decision, terminals are then scheduled in the MAC frame depending on fairness considerations [12]. Fairness among terminals can be maintained by scheduling the communication such that either all terminals obtain an equal share of the total frame time (the "uniform slot size" scheme) or all terminals are allowed to send a uniform amount of data in slots of varying length depending on their modulation and, ultimately, their distance from the access point (the "uniform traffic size" scheme). These two options reflect different perspectives on how to share system resources – the first one might be more provider oriented, the second one rather customer oriented ("I don't care where I am, I want the same effective data rate as everybody else").

The total capacity of a cell which uses a "uniform slot size" scheme is simply the arithmetic mean of the terminal's data rates (whether it is derived by relaying or not is not relevant), times the number of terminals. For the "uniform traffic" scheme, this is more complicated. Looking at the example from Section III-A gives an idea: To transmit, say, 1 Mbit from each terminal takes $1/6$, $1/36$, and $1/27$ seconds, respectively. Hence, 3 MBit are transported in $1/6 + 1/36 + 1/27$ seconds, yielding a throughput of $3/(1/6+1/36+1/27)$. More generally, the uniform traffic scheme corresponds to the *harmonic* mean of the individual data rates. It is worthwhile to point out that the harmonic mean corresponds to the behavior of IEEE 802.11: When every terminal gets the same chance to access the channel to transmit data packets of the same size, traffic fairness ensues. The "performance anomalies" of IEEE 802.11 [5] are, essentially, only the result of the harmonic mean's sensitivity to outliers. Relaying should be particularly beneficial in this fairness scheme as it benefits in particular far terminals, which severely affect the cell capacity under a harmonic mean averaging process.[3]

In the present paper, however, we do not pursue this aspect further. The simplest case, considered here, is that every terminal gets an equal time share of the entire frame.

2) Direct and relay scheduling: The actual scheduling using a single frequency is straightforwadly done by each cell's AP. In direct communication, terminals are simply scheduled sequentially. For one-frequency relaying, the time slot of a relayed terminal is split into two: one sub slot is used for relayed to relaying terminal, the other sub slot for relaying terminal to AP communication.

For the two-frequency relaying case, there can be several alternatives for concurrent transmissions. The scheduler optimally selects the relay to AP communications and schedules them in the second frequency, in a time slot overlapping transmissions in the first frequency, if possible. If there are candidates for two-frequency relaying transmissions but which cannot be scheduled in an overlapping time slot due to lack of independent pairs of entities, then these communications are scheduled normally, using the primary frequency.

E. Frequency recycling

The 3-, 7- and 19-frequency reuse patterns are typical example for frequency distribution on a cell level (partially shown in Figure 5). These assignments represent the "primary" frequency used by each cell; for brevity, we refer to this as the "macro pattern".

This macro frequency reuse pattern assigns every AP a primary frequency for its direct or one-frequency relaying communication. Even in two-frequency relaying,

[3]Consider the example again. For the "uniform slot" fairness scheme, relaying improves the cell capacity from (6+36+27)/3 = 23 MBit/s to (12+36+27)/3 = 25 MBit/s (a gain of about 8.6 %); under a "uniform traffic" scheme, it is improved from 3/ (1/6 + 1/36 + 1/27) = 12.96 MBit/s to 3/ (1/12 + 1/36 + 1/27) = 20.25 MBit/s, a gain of about 56 %!

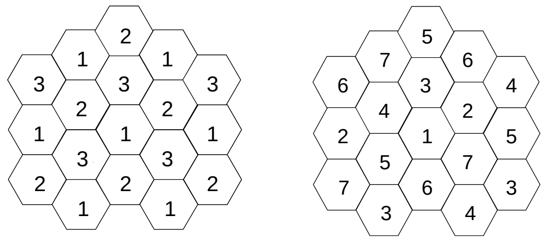
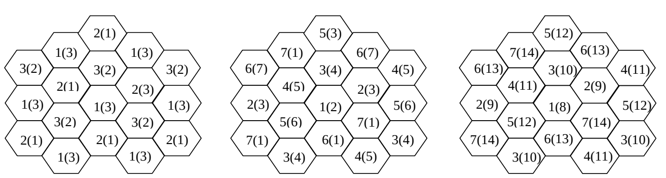


Fig. 5. Macro frequency reuse pattern for 3 and 7 available frequencies

only the relay to AP communication is scheduled using another frequency while the rest continues to always use the primary frequency.

The second frequency is assigned according to a "micro" frequency reuse scheme as shown in Figures 6a and 6b. The frequencies are recycled from a neighbor's primary frequency and are used only close to the center of the cell (according to the stipulations of the scheduler). Since the communication distance is small, the transmission power is in turn small and hence the effect of this recycled freuency in terms of interference on the neighboring cells is not that severe (assuming the data rate is not changed; adapting the data rate for short distances does counteract the effect) . The relayed terminals, which are (typically) located far away from the cell's center and responsible for most of the cell's interference, still use the cell's primary frequency according to the macro frequency reuse pattern.



a. 3- Frequency layout b. 7- Frequency layout c. 7- Frequency layout with 7 additional frequency for relaying

Fig. 6. Macro and micro frequency reuse pattern for 3 and 7 available frequencies; secondary (micro reuse) frequencies are shown in parenthesis.

An alternative approach to frequency reuse is enabled by relaying when a large number of frequencies is available: instead of using, e.g., 19 frequencies to build a very wide-spaced macro frequency pattern, we can use these frequencies to enable two-frequency relaying based on a smaller macro reuse pattern without actual borrowing from neighboring cells. Thus, for the 7 macro frequency pattern, another micro frequency pattern is formed where every cell has a dedicated secondary frequency as in Figure 6c—the frequency recycling happens on a larger scale here.

IV. Results

A. Simulation setup

For the different frequency reuse patterns described above, the capacity (or average throughput) achieved by capacity-oriented schedules for direct communication, one-frequency and two-frequency relaying is evaluated by simulations. The capacity of a cell is obtained by averaging over the goodput of all the terminals within the cell; here we present results for the "uniform slot size" case. These goodput values per terminal are computed in the context of 28 cells in total (56 cells for the 19 frequency case), with the APs located in a regular, hexagonal grid; the cells all have a radius of 35 m. These cells are basically used as a source of interference and as a source of second frequency for relaying purposes; the capacity numbers provided here are only computed from the cell at the center to protect against edge effects.[4] Within each cell, terminals are uniformly, randomly placed. The cell's capacity is then averaged over the capacity results of 55 different placements of terminals.

Note that in all experiments, the *bandwidth per cell* is always constant; different frequency reuse patterns only change the number of different frequency bands in the system at large, but not the allocation per individual cell. Performance differences between the different reuse patterns are thus due to changes in the interference situation.

B. 7-frequency reuse pattern

Figure 7 shows the average throughput of a cell at the AP for a 7 frequency reuse pattern (the confidence interval for all results are computed at a 95% confidence level but they are not shown in the figures as they are very small). The one-frequency relaying shows very little improvement in throughput over the direct communication, mainly because of the interference situation from the macro frequency reuse pattern. Since the reused bandwith is limited to only 7 frequencies, the distance between two cells using the same (primary) frequency is also small, which makes the interference severe. When it comes to the two-frequency relaying case with recycled frequencies, the source of interference is both from the macro and micro frequency reuse pattern. However, since the communication distance is reduced by the relaying

[4]The center cell's capacity stabilizes when about 15 cells are used for 3 and 7 frequency cases; a higher number of cells does not notably change the results.

and since the micro frequency plan is used only close to the AP, the interference is reasonably reduced, increased parallelism outweighs the disadvantages, and hence there is up to 19% improvement in average throughput at the AP.

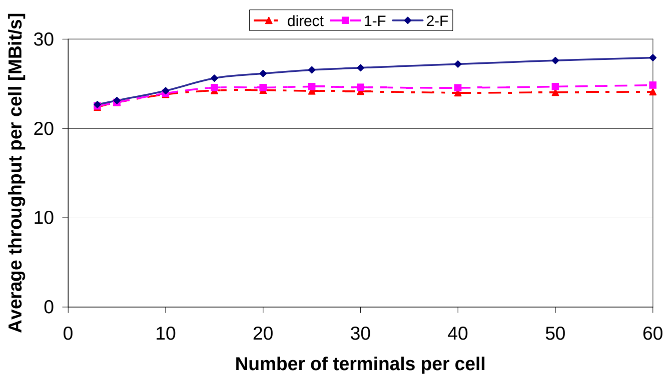


Fig. 7. Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 7 frequency reuse pattern

C. 3-frequency reuse pattern

A similar situation is observed for a macro reuse pattern with 3 frequencies (shown in Figure 8). As the frequency reuse distance is small, every transmission is vulnerable to severe co-channel interference. Though the micro frequency reuse scheme amends the situation to some extent, the overall throughput at the AP is considerably less than with 7 frequencies at hand. Nonetheless, when the available reuse frequencies are very few, relaying in two frequencies inevitably gives a better solution.

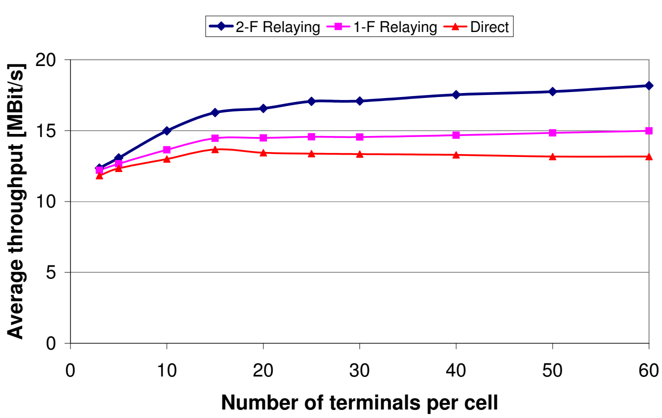


Fig. 8. Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 3 frequency reuse pattern

D. 19-frequency reuse

In the case of a 19 frequency reuse pattern, the interfering distance from co-channel cells is relatively large. As shown in Figure 9 the one-frequency relaying outperforms the direct communication. Up to 24% gain is also obtained when a micro-frequency plan with frequency recycling from neighboring cells is applied.

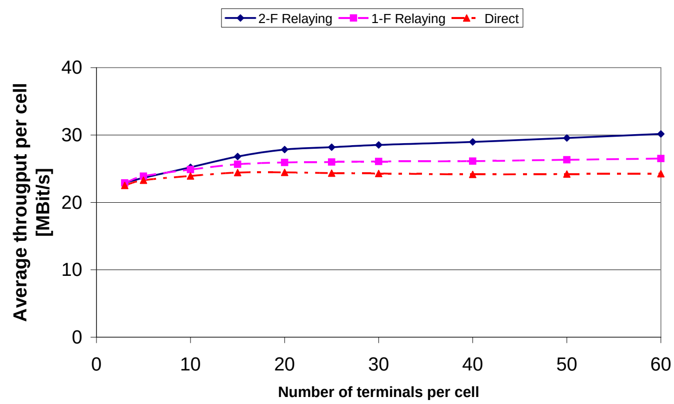


Fig. 9. Average throughput as a function of terminals per cell for 1% PER, $\alpha = 3.2$ and 19 frequency reuse pattern

To complete the description of these results, let us also consider the variation of the path loss coefficient α and the maximum allowed, target packet error rate. Figure 10 shows the average throughput per cell if α increases. Not surprisingly, relaying is not beneficial for small values of α, as all terminals always can communicate directly with the AP; the initial rising slope is due to better "shielding" of inter-cell interference. Relaying does result in performance gains if α exceeds about 2.8.

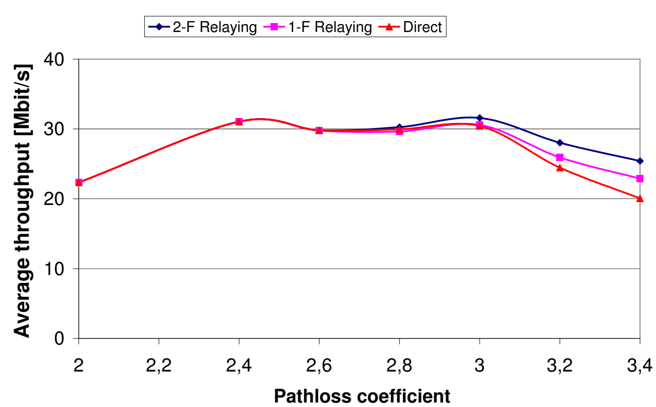


Fig. 10. Average throughput as a function of varying α, 20 number of terminals, 19 frequency reuse pattern

Figure 11 shows the variation of average throughput of two-frequency relaying with respect pathloss coefficient α and 1%, 3%, 5% and 10% maximum target packet error rates. Clearly a higher throughput can be attained when the target packet error rate becomes more tolerant but it remains a trade off between a higher throughput and increased number of retransmissions.

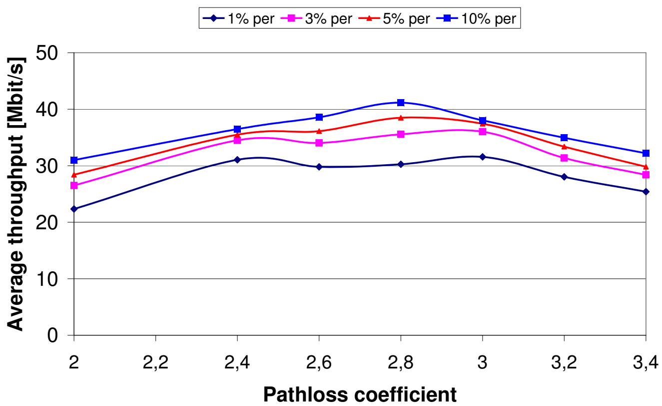


Fig. 11. Average throughput of 2-frequency relaying as a function of varying α, maximum allowed target PER, 20 number of terminals, 19 frequency reuse pattern

E. Separate relaying frequencies

Comparing Figures 7 and 9 shows that the gain from adding 12 frequencies is actually small. Hence, these frequencies might perhaps be better used to support relaying. The interesting situation is when a 7 frequency pattern is used for the macro frequency reuse pattern and additional 7 frequencies (i.e. 14 frequencies in the system as a whole) are available for the micro reuse pattern as described in Section III-E. As the recycled frequencies are not used from the immediate neighboring cells, the reuse distances are increased, which evidently brings about capacity improvement by reducing the level of interference. Figure 12 compares this situation with that of the 19 frequency reuse pattern. Up to 40 % improvement can be obtained by the two-frequency relaying when the new micro frequency reuse pattern is used and yet 5 frequencies can be saved compared to the standard 19 frequency reuse pattern. This represents a good incentive to have a smaller frequency plan with a better figure of merit.

V. Related work

The notion of using relaying to improve capacity in wireless networks has been discussed in several contexts. Gupta and Kumar [3] have studied the capacity of

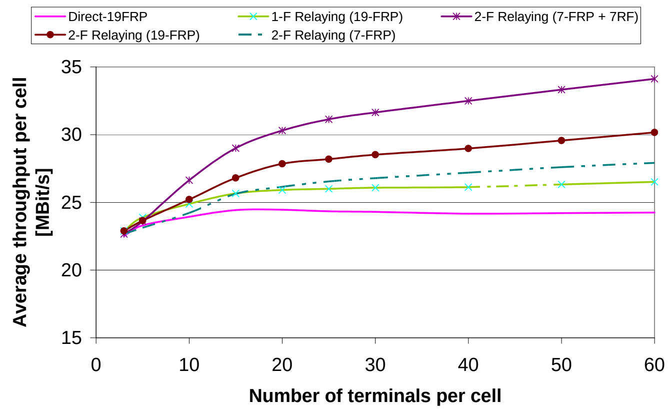


Fig. 12. Comparison of average throughput as a function of terminals per cell for 1% PER and $\alpha = 3.2$

randomly located ad hoc networks. Their result shows that for n identical, randomly located nodes, the maximum achievable throughput under optimal circumstances is $O(\frac{1}{\sqrt{n}})$ and as the number of nodes n per unit area increases, the throughput decreases accordingly. In their purely ad hoc network, they assumed that nodes can transmit using a given common channel and their results do not change if they subdivide the given channel. In our general model where nodes can communicate both with each other and with an AP, we also assumed a given channel at a system level but we subdivided it to reuse for both basic and relay transmissions. As a result, for a network with an increasing number of nodes we obtained a higher capacity upper bound due to frequency recycling than for a network without frequency recycling. Similar theoretical studies are also made in references [1, 11]. Our simulation results stick to these theoretical results in that as the number of terminals increases the per-terminal throughput decreases whereas the overall throughput increases. We do show that our schemes improve upon the total capacity compared to previous schemes.

Relaying in combined cellular and ad hoc systems has also been discussed by various authors. Lin and Hsu [10] compared the end-to-end throughput of the conventional single-hop cellular network with that of a multihop cellular network architecture and by reducing the transmission distance within a cell, they demonstrated how multihop makes simultaneous transmissions possible so as to get improved throughput. But since packets are sent multiple times to arrive at their destinations, the achieved throughput is limited due to bandwidth consumption. Our work is also based on the notion that terminals can communicate with each other for relaying purposes to open possibilities for simultaneous transmissions. Scheduling simultaneous transmissions in the same frequency within a cell without interference requires the cell to be large. Thus, we opted for additional relaying frequency for simultaneous transmissions to minimize interference within the cell. Moreover, we established rate-adaptive multi-hop communication as a crucial optimization technique.

WU et al. [16] addressed the issue of increasing system throughput by introducing ad hoc relay stations in their integrated cellular and ad hoc relay (iCAR) architecture. These ad hoc relays are placed at strategic locations to divert traffic from heavily congested cells to other less congested cells. Such special devices are cost ineffective as they are fully utilized at times of heavy congestion only. But the relaying nodes in our cases are the terminals themselves with their own traffic to send to the AP and hence no additional device is required for the system.

The concept of channel reuse in relation to capacity of a cellular network has been studied by various authors [7, 9, 13, 15].

SRENG et al. [14] applied the reuse of adjacent channels in relaying in TDMA cellular radio network, mainly to enhance coverage. They proposed a relay node selection scheme based on the pathlosses associated with two-hop relay links and relay channel selection schemes based on carrier-to-interference ratio received at the relay nodes from the channels to be reused. Since the proposed schemes are for coverage enhancement, they varied their cell radius from 200 m to 1 km, they also let the maximum transmission power to vary upto 1 W. The radius of the cell in our case is governed by the maximum transmission radius that a terminal can have when communicating directly to the AP with the lowest data rate, i.e., our capacity optimization goal is in the already covered region of the cell and terminals within the cell use minimum transmission power which allows to send with the selected data rate for some target packet error rate. While the idea of relaying and adjacent channel reuse is similar, our approach is based on a joint routing and scheduling mechanism to improve the capacity of the cellular network, while at the same time maintaining (or even explicitly influencing) fairness in the system. We also have macro and micro frequency reuse patterns so as to use adjacent frequencies only close to the AP to minimize co-channel interference. Our routing mechanism is based on optimal data rate with power control which is adaptable in interference-limited environment.

Relaying in CDMA networks is studied in reference [4]. The authors showed how the overall performance of a CDMA relay system depends on the node density and the relative load and suggested that direct transmission eventually becomes favorable with respect to capacity consideration. In fact, our results also suggested that unless channel reuse is used, relaying in single frequency in interference-limited environments does not bring about a significant improvement in capacity (which is also commensurate with theoretical results).

VI. CONCLUSION AND OUTLOOK

Relaying is a viable means to improve the operations of an infrastructure-based wireless communication system as it increases the cell capacity. In this paper, we described rate-adaptive, two-frequency relaying algorithms which improve the system capacity and maintain fairness among terminals. The key technique is to recycle a frequency from a neighboring cell to use it in the interior of a cell to enhance transmission parallelism, hence improving capacity by keeping the access point busy at higher data rates—the disadvantages of higher interference are outweighed by the advantages. Additionally, two-frequency relaying can be used, when many frequencies are available, to improve the overall frequency reuse pattern.

For some macro and micro frequency patterns, we obtained up to 40 % gain in capacity in the network. Moreover, the algorithms are practical as they are not computationally intensive, they can be implemented as iterative online algorithms, and are based on information that can be provided by real systems with acceptable overhead. As no additional infrastructure is necessary and also the requirements on the individual terminals are quite modest, our relaying approach can provide a simple and cheap solution to add capacity to a wireless system, particularly in highly loaded networks. Currently, we are studying the impact of mobility and channel fading and the ensuing errors in channel gain estimates.

With respect to the different fairness schemes, we are currently investigating how relaying influences the capacity in these schemes. In particular, we are interested in the trade-offs between the data rates and the distances that they can cover (at a given error rate) and how this influences capacity. We conjecture that especially the "uniform traffic" case will benefit considerably from a rate-adaptive relaying because of its harmonic mean characteristics.

Acknowledgment

Part of this work has been supported by the Germany Ministry for Education and Research (BMBF) under the HyperNET/IBMS2 project.

References

[1] Michael Gasper and Martin Vertterli. On the Capacity of Wireless Networks: The Relay Case. *IEEE*, 2002.

[2] H.-F. Geerdes and H. Karl. The Potential of Relaying in Cellular Networks. In *Proc. Intl. Network Optimization Conf.(INOC)*, Paris, France, October 2003.

[3] P. Gupta and P. R. Kumar. The Capacity of Wireless Networks. *IEEE Trans. on Information Theory*, November 1998.

[4] Patrick Herhold, Wolfgang Rave, and Gerhard Fettweis. Relaying in CDMA Networks: Pathloss Reduction and Transmit Power Savings. In *VTC2003*, Korea, Spring 2003.

[5] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. Performance Anomaly of 802.11b. In *Proc. INFOCOM*, San Francisco, CA, March 2003.

[6] G. Holland, N. Vaidya, and P. Bahl. A Rate-Adaptive MAC Protocol for Multi-Hop Wireless Networks. In *Proc. 7th Ann. Intl. Conf. on Mobile Computing and Networking*, pages 236–250, Rome, Italy, July 2001. ACM.

[7] I. Katezela and M. Naghshinen. Channel Assignment Schems for Cellular Mobile Telecommunication Systems: a comprehensive survey. In *Proc. IEEE Personal Communication Magazine*, volume 82, pages 1398–1430, 1994.

[8] J. Khun-Jush, G. Malmgren, P. Schramm, and J. Torsner. HIPERLAN type 2 for broadband wireless communication. *Ericsson Review*, 2:108–119, 2000.

[9] Jiandong Li, Zygmunt J. Haas, and Min Sheng. Capacity Evaluation of Multi-Channel Multi-Hop Ad Hoc Networks. In *ICPWC02*, New Delhi, India, December 2002.

[10] Y.-D. Lin and Y.-C. Hsu. Multihop Cellular: A New Architecture for Wireless Communications. In *Proc. INFOCOM*, Tel-Aviv, Israel, March 2000.

[11] B. Liu, Z. Liu, and D. Towsley. On the Capacity of Hybrid Wireless Networks. In *Proc. IEEE INFOCOM*, San Francisco, CA, March 2003.

[12] Seble Mengesha and Holger Karl. Relay Routing and Scheduling for Capacity Improvement in Cellular WLANs. In *Proc. of WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, Sophia-Antipolis, France, March 2003.

[13] Jungmin So and Nitin H. Vaidya. A Multi-channel MAC Protocol for Ad Hoc Wireless Networks. Technical report, University of Illinios at Urbana-Champaign, January 2003.

[14] V. Sreng, H. Yanikomeroglu, and D. Falconer. Coverage Enhancement through Two-hop Relaying in Cellular Radio Systems. In *Proc. IEEE Wireless Communications and Networking Conf.*, Orlando, FL, March 2002.

[15] S.Wu, C.Lin, Y.Tseng, and J.Sheu. A New Multi-channel MAC Protocol with On-Demand Channel Assignment for Multi-Hop Mobile Ad Hoc. In *I-SPAN2000, International Symposium on Parallel Architectures, Algorithms and Networks*, Texas, USA, December 2000.

[16] H. Wu, C. Qiao, and O. Tonguz. Performance Analysis Of iCAR (Integrated Cellular and Ad hoc Relaying System). In *Proc. ICC*, volume 2, pages 450–455, 2001.