Improving Goodput by Relaying in Transmission-Power-Limited Wireless Systems

Seble Mengesha, Holger Karl, Adam Wolisz
Telecommunication Networks Group
Institute of Telecommunication Systems,
Faculty of Electrical Engineering and Computer Science,
Technische Universität Berlin
Einsteinufer 25, 10587 Berlin, Germany,
{mengesha,karl,wolisz}@ee.tu-berlin.de

Abstract

In wireless communication systems, the capacity of a cell (the amount of correctly delivered traffic in unit time) is a precious resource that cannot be arbitrarily increased. This paper looks at the consequences limited transmission power has for the cell capacity and how these limitations can be overcome by using relaying between terminals. The challenge faced here is that we require the relaying terminal not only to transmit relayed traffic but also traffic originated locally. The key technique to handle this challenge is to dynamically adapt transmission power and in particular modulation and hence error susceptibility depending on the actual configuration. We show that it is possible to increase the capacity in certain configurations considerably; critical parameters are the location of terminals, the path-loss coefficient, and the noise. Under certain fairness constraints (all terminals must obtain a minimum share of total goodput), relaying can provide improvements of up to 350%.

Keywords: Wireless LAN, relaying, ad hoc, capacity

1. Introduction

Improving the capacity of successful transmission of user traffic is of practical importance when radio spectrum becomes a scarce and valuable resource. Many different approaches such as coding, modulation and appropriate medium access techniques have been used to improve the capacity; and recently the use of relaying has been considered not only to extend the coverage of a wireless ad-hoc network but also for capacity improvements.

When terminals are far away from their destination, they will transmit at high transmission power levels in order to reach their destination. This transmission power is large enough to cause interference to other transmitting terminals and hence causes high packet error rates. Moreover, even if there are no neighboring terminals that would suffer from interference, the transmission power cannot be arbitrarily increased because of regulatory limitations on effective isotropic radiated power (EIRP), and thus high packet errors are inevitable when mobile terminals are far away from their access points or far away from each other.

In such a situation, relaying packets as opposed to direct transmission has prominent advantages: the interference problem is mitigated and terminals close to the maximum range can now transmit their traffic to the access point via terminals closer to the access point (relaying terminals) with reduced transmission power. For terminals that are already using maximum transmission power, at least the packet error rate can be reduced. These effects should promise an increase in the total amount of traffic which can be supported by a set of terminals in a given amount of time. The typical scenario we are considering in this paper is a cellular setup where there is an access point managing a certain cell, containing mobile terminals.

The downside of relaying is the overhead for the relaying terminal. Since the relaying terminals not only transmit the relayed traffic but also their own traffic, a faster modulation with a higher bit rate than the other terminals should be used in order to handle the additional traffic load in a given amount of time. The use of higher bit rates, however, has a penalty of higher error rates as compared to the lower ones at the same level of noise and interference. Hence there is a tradeoff between the use of shorter distances (improving PER) and the need for higher modulations (increasing...
This tradeoff is investigated in the present paper, taking into account realistic channel models and PER models as well as limitations on transmission power. Both direct and relaying transmission in a wireless cellular network are considered and their goodput (the total amount of correctly delivered bits per unit time) is compared. A simple cell model will be used with one access point and four transmitting terminals. The radio technology under consideration is HiperLAN/2 (as this paper was developed in the context of the IBMS² project, which uses HiperLAN/2 as a case study), and both the concrete error models as well as the maximum allowed transmission power of 200 mW steam from this assumption. However, the results should easily carry over to other similar radio technologies as well (e.g., IEEE 802.11 PCF).

This paper is the first paper to investigate this particular combination of impacts of distance, limited transmission power, and, in particular, modulation when comparing direct and relaying communication. The main result is that for certain combinations of distances between access point, relay terminal, and normal terminal, relaying does outperform direct communication and thus increases the capacity of wireless systems. In a related paper [10], we consider instead of the limited transmission power the consequences of interference reduction between cells and derive similar results (this paper does not consider limited transmission power but also uses choosing among different modulations as a key technique).

The remainder of this paper is organized as follows. The following Section 3 considers related work, Section 4 describes the system model, which in turn is analyzed in Section 5. Numerical evaluations of the analysis are shown in Section 6 and finally, Section 7 contains conclusions and a discussion of possible future work.

2. Related Work

Relaying has been considered as a means to many different ends. In the context of ad-hoc networks, the routing question has been extensively addressed (an overview can be found e.g. in [11]). In the present context, particularly the question of capacity improvements by relaying are of interest.

GUPTA and KUMAR [3] have studied the capacity of ad-hoc networks with randomly located nodes having random communication patterns. Their result shows that for \( n \) nodes, the maximum achievable throughput per node under optimal circumstances is \( \frac{1}{\sqrt{n}} \) times the transmission rate. Hence, as the number of nodes increases, the throughput per node drops to zero. However, the main point here is that communication patterns are not local: as the number of nodes increases, the average distance also increases, requiring more and more spectrum and space. In our model, communication distances do not arbitrarily increase as we are mostly interested in improving the capacity of a fixed cell (i.e. supporting more users at a minimum bandwidth), but even for mobile terminals that are relatively far away from their access points.

GROSSGLAUSER and TSE [4] used a similar situation as in [3] but introduced mobility into the model and showed how the capacity of this wireless ad-hoc network can be increased. Their result is, in fact, in sharp contrast to that of [3] since the average long-term throughput they obtained can be kept constant even as the number of nodes \( n \) per unit area is increased. They also showed that a single relay is sufficient to use the entire throughput capacity of the network within the limits of their interference model. However, the increase of the capacity is at the cost of considerable delay.

In the context of IEEE 802.11 [8], [11] have performed a simulation study of the relaying behavior of 802.11’s distributed coordination function (DCF). They considered different scenarios (regular and random placement of nodes) and attempted to communicate packets over multiple hops between different nodes. Their main result is that, while IEEE 802.11 DCF does find reasonably efficient communication schedules even in complicated networks, only local communication patterns can be efficiently supported (in agreement with the theoretical results derived by [3]).

Recently, BRONZEL et al. [11] have shown for CDMA systems (also in the context of the IBMS² project) how the capacity of a single-relay network can be increased by reducing the interference in the network. They investigated the achievable average transmit power reduction and capacity gain that results from using mobiles as relay stations for other terminals. Their result shows that relaying in a mobile network generally reduces the total transmit power. Furthermore, relaying provides an increase in achievable capacity gain of around 20% for uplink transmission in a CDMA system. An interesting option would be to use the ideas of the present paper, namely adapting the modulation rate, for a CDMA system as well, e.g. in the form of variable spreading gains.

As the IBMS² project is targeting HiperLAN/2 as a demonstration platform, we are particularly interested in using interference-to-error-rate mappings that describe the actual behavior of a HiperLAN/2 system. The only numbers available are those published by KHUN-JUSH et al. [8]. Their model assumes large office environments with non-line-of-sight propagation, modeling indoor wireless communication systems. The mapping from SNR to PER we are using here [11] is based on their simulation results. They also showed the ideal achievable link throughput with respect to SNR or C/I for the different physical layer modes.
A similar result is described by [HOLLAND et al. [8]] who show how to choose the modulation rate in order to obtain optimal link throughput in an IEEE 802.11 environment. The main contribution of their work is a heuristic scheme on how to choose modulation from a receiver-oriented point of view and how the 802.11 standard would need to be extended so as to accommodate their signaling scheme. In contrast to the work presented here, they do not consider the effects of employing different modulations in relaying yet only look at a single link. In [7], we also consider how lower interference levels due to relaying affect the choice of modulations/transmission power levels. In a sense, the work described in [3] could serve as a building block in relaying systems — this is subject to further investigations.

3. Model

To investigate the capacity improvement of relaying transmission over the direct transmission case, the cell model shown in Figure 1 is used. The cell contains an access point at the center of the cell, a relaying terminal at a distance \( d \) close to the access point and three other transmitting stations \( d' \) away from the relaying terminal. For simplification, we assume that the three “far” terminals MT\(_2\) to MT\(_4\) are located very close to each other, so that the differences in distance to MT\(_1\) or AP are negligible.

![Figure 1: System scenario](image)

A transmission power \( P_T \) is used to transmit the traffic load from the terminals either to the access point or to the relaying terminal. A simple path-loss model which depends only on the distance \( x \) and the path-loss coefficient \( \alpha \) is used to relate the transmission power \( P_T \) to the received power \( P_R \):

\[
P_R = \frac{P_T}{x^\alpha}
\]

As the intention of this paper is to investigate the effect of limited transmission power in a wireless system, the model is assumed to have negligible interference from other cells or adjacent channels (an assumption that is dropped in [8]). Moreover, TDMA is applied and no two terminals transmit at the same time using the same frequency band. Hence, there is no interference within the cell and the only sources of packet errors are noise and limited transmission power. The signal to noise ratio is expressed simply as \( \text{SNR} = \frac{P_R}{N} \) or, if SNR is expressed in dB, \( \text{SNR} = 10 \log_{10} \frac{P_R}{N} \).

Our analysis is carried out based on the interference-to-error-rate mapping data for HiperLAN/2 [8]. An exponential curve fitting [11] is used to mathematically describe the mapping and the PER is:

\[
\text{PER}(s, i) = 10^{a_i x^2 + b_i x + c_i}
\]

where \( s = 10 \log_{10} \text{SNR}, \) the SNR in dB, \( i \) is the index of the modulation (HiperLAN/2 offers seven different modulations) and \( a_i, b_i, \) and \( c_i \) are as defined in Table 1, which also shows name and nominal bit rate (NBR) for these modulations. An appropriate cutoff is chosen to ensure that the PER remains between 0 and 1. While these numbers describe the indoor model quite accurately, we are also applying them to scenarios for which they were not intended (e.g., different path-loss coefficients). Nevertheless, we believe that this approximation should be close enough to provide an initial understanding of the relaying capacity problem.

Throughout the analysis, uplink phase is considered and the terminals are assumed to have an unlimited of data they want to transmit with maximum possible data rate. For simplicity, only the nominal capacity of a given transmission is considered (protocol overhead is not taken into account, which anyway would be almost identical for both direct and relaying). As a perfect transmission over a wireless link is not possible, terminals try to obtain a packet error rate of at most 0.1% if feasible.

This traffic assumption is actually conservative as it puts considerable requirements on the relaying node, which has not only to transmit its own traffic but also the other nodes’ traffic. This results in the need to use faster modulations between inner terminal and access point (otherwise, queues would overflow in the relay node) with inferior error behavior. Under lower load, this could be avoided, which should in turn improve the performance of relaying.

The performance metric evaluated here is the total amount of traffic, coming from all mobile terminals, that the access point can successfully receive in unit time.

4. Analysis

In this section, for both direct and relaying communication, the goodput at the access point in the uplink direction is derived. This analysis is based on SNR values given a certain amount of transmission power and distance and on
the mapping from SNR values to packet error rates. A predefined communication schedule is needed to see the effect of direct and relay communication in this TDMA based analysis.

### 4.1. Direct Communication

The schedule is straightforward for the direct case: A four-time-slot schedule is shown in Table 2. Each terminal sends its traffic load directly to the access point, selecting one of the seven possible modulations.

<table>
<thead>
<tr>
<th>Slot</th>
<th>Communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$MT_1 \rightarrow AP$</td>
</tr>
<tr>
<td>2</td>
<td>$MT_2 \rightarrow AP$</td>
</tr>
<tr>
<td>3</td>
<td>$MT_3 \rightarrow AP$</td>
</tr>
<tr>
<td>4</td>
<td>$MT_4 \rightarrow AP$</td>
</tr>
</tbody>
</table>

**Table 2: Direct communication schedule (uplink)**

In addition to the modulation, each terminal also has to choose a transmission power; any choice between 0 and 200 mW is acceptable. Depending on the distance the terminals are away from the destination of their packets, the terminals can adjust their transmission power to keep their packet error rate PER to less than some specified value. As the PER is the model considered here, they try to maintain a minimal SNR value. With increasing distance, transmission power has to be increased, but can not be raised beyond the maximum of 200 mW. For distances that would require a larger transmission power, the PER can not be maintained and starts to increase; the goodput will also decrease accordingly. This situation is shown in Figure 2 for two different combinations of $\alpha$ and distance between terminals $d$.

Each mobile terminal adjusts its transmission power according to the following relation in order to maintain the required SNR level:\footnote{As this assumes perfect power control, it is unrealistic to hope for an actual system to achieve such performance. It only serves as an upper bound. How to construct relaying systems that actually approach the upper bound computed here is the topic of current research.}

\[
\text{SNR}(x) = \frac{P_t(x)}{N} \Leftrightarrow \\
\text{SNR}(x) = \frac{P_t(x)}{N x^{\alpha}} \Leftrightarrow \\
P_t(x) = N x^{\alpha} \text{SNR}(x)
\]
Transmitting from $MT_1$ to $AP$:

$$\text{SNR}_{MT_1 \rightarrow AP} = \frac{P_i(d)}{d^{\alpha} N} \quad (2)$$

Transmitting from $MT_2$ to $AP$:

$$\text{SNR}_{MT_2 \rightarrow AP} = \frac{P_i(d + d')}{(d + d')^{\alpha} N}$$

The outer terminals ($MT_2$, $MT_3$, and $MT_4$) are assumed to have identical SNR. Once the SNR values of each transmission is computed, the PERs can easily be determined from Equation (1) for the employed modulation type $j$. The goodput $GP_{\text{direct}}$ for a particular terminal can, then, be approximated as $GP_{\text{direct}} = (1 - \text{PER})(\text{modulation’s nominal bit rate})$. This assumes a very simple link layer which does not provide any ARQ protocols. However, since the terminals have to share the channel with each other, the nominal data rate needs to be divided by four (without considering a downlink phase, which would further scale down the goodput). Thus, the goodput is (using the PER Equation (1)):

$$GP_{\text{direct}}(i, j_i) = (1 - \text{PER}(10 \log_{10}(\text{SNR}_{MT_i \rightarrow AP}), j_i)) \frac{\text{NBR}_{j_i}}{4}$$

where $i$ is the terminal under consideration and $j_i$ is the modulation $j$ it uses to send the traffic load to the access point. The total goodput at the access point is finally obtained by adding up the individual terminals’ goodput in the cell.

In order to obtain the optimal possible throughput $GP_{\text{optimal}}^{\text{direct}}$, the best combination of modulations is used. Hence:

$$GP_{\text{optimal}}^{\text{direct}} = \max_{i \in \{1, \ldots, 4\}, j_i \in \{1, \ldots, 4\}} \sum_{i=1}^{4} GP_{\text{direct}}(i, j_i)$$

where $i$ is the terminal under consideration and $j_i$ is the modulation it uses to send the traffic load to the access point. The total goodput at the access point is finally obtained by adding the individual terminal goodput in the cell.

### 4.2. Relay Communication

The relay communication schedule, shown in Table 3, uses the same number of slots as in the direct case. But the outer terminals use the inner terminal as a relay to send their traffic load to the access point. The inner terminal should now transmit both its own and the relayed traffic. Here, the inner terminal should employ a faster modulation and higher data rate than the overall nominal bit rate of the outer terminals.

This schedule is actually a pessimistic approximation. The relay terminal is always using the faster, more error-prone modulation, whether or not the outer terminals have sent any data that is to be relayed to the access point. In the particular traffic model assumed here, the outer terminals always have data to send, which is unlikely to be the case in realistic scenarios. Hence, adding a mechanism that uses faster modulation between relay and access point only when necessary should again improve the error behavior of this link. Such a mechanism, however, could imply additional signaling overhead and delay— we leave this problem for further study.

Similar to the direct case, the transmission power has to be chosen based on the PER requirement. The PER requirement for the two relay hops is slightly different from the overall PER, which remains the same as that of the direct case. To guarantee this, the target PER of each individual hop in the relaying case should be chosen such that two consecutive transmission have the same target error rate as in the direct case. Thus,

$$(1 - \text{PER}_{\text{rel}})^2 = 1 - \text{PER}_{\text{direct}} \Leftrightarrow \text{PER}_{\text{rel}} = 1 - \sqrt{1 - \text{PER}_{\text{direct}}}$$

Again based on the schedule, the outer terminals will have SNRs of:

### Transmitting from $MT_i$ to $AP$:

$$\text{SNR}_{MT_i \rightarrow AP} = \frac{P_i(d)}{(d')^{\alpha} N}$$

where $d' > 1$ is an outer terminal. The inner terminal has the same SNR as in Equation (3). Note that the outer terminals are now sending their traffic to a closer relaying terminal and they can maintain the assumed PER level to a considerable distance with the limited transmission power, even though now the target PER for an individual hop is larger than in the direct case.

The determination of goodput is different from that of the direct case as we have to deal with different modulations over the different hops. Consider first the goodput...
for terminal $MT_1$, which uses modulation $i$ with nominal bit rate $NBR_i$ to transmit to the access point. Since $MT_1$ is transmitting only for a quarter of the time, $NBR_i$ has to be divided by four. Moreover, it is only part of the data that belongs to $MT_1$ in the given time slot — some of the time has to be devoted to the relaying of data received from other terminals. Thus, the actual goodput of $MT_1$ for transmitting its own data is

$$GP_{relay}(1, j_1) = (1 - \text{PER}(10\log_{10}(\text{SINR}_{MT_1 \to AP}, j_1))) \cdot \frac{NBR_1 - \sum_{i=2}^{4} NBR_i}{4}$$

where $NBR_i$ is the nominal bit rate of the outer terminal $i$. The goodput of the outer terminals depend on PERs of both hops (from the terminals to relay and from relay to access point) and is computed as:

$$GP_{relay}(i, j_i) = (1 - \text{PER}(10\log_{10}(\text{SINR}_{MT_1 \to AP}, j_1))) \cdot \frac{NBR_i}{4}$$

where $i$ is the index of the outer terminal and $j_1$ and $j_i$ are the modulation indexes used by the relaying and outer terminals, respectively. The total goodput at the access point is the overall goodput of each terminal in the cell, the optimal total goodput $GP_{relay}^{\text{optimal}}$ is again the goodput achieved with the best combination of modulations, just as in the direct case.

$$GP_{relay}^{\text{optimal}} = \max_{j_i(1, ..., 4)} \sum_{i=1}^{4} GP_{relay}(i, j_i)$$

5. Numerical Evaluation

Following the equations obtained in Section 4, we have examined the total goodput of the direct and relay communication for the wireless network model under consideration. Setting the PER requirement to 0.1%, we varied the terminal separations $d$ and $d'$ for a given path-loss coefficient $\alpha$ and noise $N$. The transmission power $P_t$ adjusts itself accordingly for each terminal until it reaches the maximum limit. However, in this section we introduce an additional requirement: all terminals should get a certain minimum amount of goodput; otherwise, we declare the total goodput to be zero. This constraint is introduced to avoid very unbalanced communication scenarios, e.g., a direct communication case were the near terminal is transmitting at full speed while the outer terminals do not achieve any goodput at all (a case which, were it accepted, would put relaying at an unfair disadvantage).

The following graphs (Figure 3 to Figure 5) give an overview of the ratio of the optimal total goodput in a cell between direct and relaying communication for various values of $\alpha$; i.e., they plot $GP_{relay}^{\text{optimal}}/GP_{direct}^{\text{optimal}}$. Ratios with values larger than 1 indicate that relaying performs better than direct communication. One representative example graph for the optimal total direct goodput is also shown in Figure 6.

![Figure 3: Ratio of total goodput for $\alpha = 3$ at $N = -70dBm$](image)

![Figure 4: Ratio of total goodput for $\alpha = 4$ at $N = -70dBm$](image)

As expected, relaying outperforms direct communication when cells are large and the outer terminals have almost lost connection with the access point. The result of
our analysis indicates that there is an improvement of communication by relaying when the path-loss coefficient $\alpha$ is larger. And even better results are obtained when the communication environment has a larger noise figure. In the particular example shown in Figure 5, relaying has a better performance well above 50% for certain configurations. For smaller $\alpha$ values, an improvement is observed only when the terminals are very far apart. It is evident that when the terminals are not very far from the access point, they can operate within the limited range of transmission power, maintaining the required $\text{SNR}$ level, and hence, they can have a better performance without using relaying. However, it might be difficult to draw conclusions since the interference-to-error-rate mapping data we are relaying is obtained for an indoor environment, which typically is around $\alpha = 4$.

There are cases where a single terminal gets the maximum share of the total goodput and the rest performs with very small bit rate. Although such cases can increase the total goodput and the capacity ratio considerably, we avoided them (as mentioned above) not to have unnecessarily unbalanced goodput distribution in the system. The sharp edges shown in the figures are the results of such cases. The actual limits when an unbalanced communication scenario is rejected are very low (on the order of kilobits per second are required as minimum goodput for each terminal); for realistic requirements, even more direct communication scenarios would have to be rejected and the case for relaying would even be stronger.

As Figure 6 shows, there are certain configurations in which direct communication does not achieve any throughput at all, while relaying still performs well (not shown). In a sense, this means that relaying is able to extend the coverage of a cell — a not surprising result, which is however not the primary goal of this paper.

6. Conclusion and Future Work

This paper has compared direct and relaying communication in wireless communication systems under the perspective of cell capacity. The main limiting factors for capacity in a direct system considered in this paper are noise and limited transmission power. We have shown that relaying does indeed have the potential to improve upon the direct system when terminals are far away from their access points. The degree of improvement depends on the location (relaying performs better over large distances), the noise (relaying performs better in environments with a lot of noise), and the path-loss coefficient (relaying performs better when attenuation is large). This combination should make relaying particularly attractive in semi-static, noisy indoor environments, were a lot of traffic should be handled with few access points; examples for such environ-
ments are exhibition halls or temporary, ad-hoc installations with only few connections to a wired backbone.

Even though the model considered here is fairly simple, it allows some generalizations. For example, only a few numbers of terminals were considered, nevertheless, the results direct scale to large numbers as the relative share of time used by the terminals is the same in both approaches. As analytical treatment of larger configurations is difficult, work is currently under way to build a simulation environment mirroring the relevant aspects of HiperLAN/2 (along with a proper channel model). Based on this simulation system, larger configurations as well as different scenarios, in particular, multi-cell scenarios, will be considered.

Other problems to overcome are the fact that the optimization here can only use a discrete amount of modulation schemes (seven in the case of HiperLAN/2) whereas a continuous optimization might be both easier and provide better results. The problem here is to correctly model “mixed” modulations in their error characteristics as well as how to actually make them feasible (a simple approach would be to switch between different modulations with a certain ratio in order to obtain an intermediary error characteristic). This is currently under investigation.

The long-term goal of this research is to devise mechanisms that automatically decide whether or not to use relaying in a particular situation, in order to optimize cell capacity (this is closely akin to solving routing problems in ad-hoc networks). The main question to answer is who will be relaying terminal for which other terminal, and how to configure transmission power and modulation. Ideally, this should be solvable with information gathered from the terminals themselves.

References