TKN Telecommunication Networks Group

Technical University Berlin

Telecommunication Networks Group

Selecting Vertical Handover Candidates in WLAN Hotspots

Sven Wiethoelter, Adam Wolisz

wiethoel@tkn.tu-berlin.de

Berlin, August 2008

TKN Technical Report TKN-08-009

This work has been supported by the German Research Foundation (DFG) in the context of the project ”POL4G”

TKN Technical Reports Series

Editor: Prof. Dr.-Ing. Adam Wolisz

Abstract

Future wireless networks will consist of multiple heterogeneous access technologies such as UMTS, WLAN, and WiMax. These technologies differ greatly regarding medium access scheme, network capacity, QoS support, choice of data rates, and other various parameters such as power consumption and AAA aspects. In the literature, different network access selection and handover strategies have been discussed in order to maximize the utilization for such a heterogeneous network. Thereby, it is still an open issue how to select ”inefficient” mobiles as vertical handover candidates. This work presents a novel scheme for the selection of handover candidates in WLAN hotspots. After discussing the design rationale of the decision metric, simulation studies show the impact of single and multiple handovers on remaining users in the cell.

Contents

Chapter 1

Introduction

Today's access networks differ greatly regarding coverage area, supported degree of mobility and user data rates. WLANs, for example, can offer high data rates in small coverage areas but only limited to no mobility support. Contrary, UMTS supports high mobility with large coverage but rather low data rates. As a result, combining different access technologies is a very promising approach to deal with various conditions such as user mobility, different traffic patterns and QoS requirements.

The classical selection approach for such a heterogeneous network follows the basic rationale of each access technology where cellular networks are primarily applied for voice services while WLAN hotspots are used to serve web-traffic. It is reality today that a certain fraction of users have terminals with network interface cards (NICs) for several technologies while other users have only access to one technology. For a cellular network provider, it may therefore be beneficial even to release a voice user with a multi-NIC terminal to a WLAN hotspot such that the capacity can be re-allocated to other users being bound to cellular networks. Although the transport of voice traffic in WLAN implies a high overhead, this approach becomes more and more promising due to the increasing density of WLAN hotspots.

Lots of research effort has been placed in the area of network selection and handover strategies for heterogeneous networks, recently. All approaches target to identify costs and revenues for each access network by combining certain input parameters to cost functions. These input parameters are nothing else than metrics reflecting criteria for a performance comparison of one specific aspect (such as load or QoS conditions).

The design of a cost function consists of two parts. Firstly, the input parameters have to be selected and secondly, they are concatenated to a cost function. Lots of work was published considering not only different parameters but also various approaches for the concatenation of cost functions. In the literature, the considered parameters range from users' preferences and QoS conditions, the load of access networks, the power consumption of network interface cards, and monetary costs—just to name only the most common ones. Other work focusses on the concatenation of these parameters to cost functions or other comparable decision models. Stevens-Navarro et al. [13] provide a short introduction and comparison of four common approaches.

Historically, Wang et al. [16] were the first who used a cost function for handover decisions in a heterogeneous network. This early work applied a linear combination of offered bandwidth, power consumption, and costs, whereby the natural logarithm was used as normalization function. Among other aspects, McNair et al. [10] extended Wang's cost function by a factor that reflects the ability of a network to fulfill certain requirements such as minimum bandwidth or maximum latency.

Above approaches define the available bandwidth of access networks as well the allocated resources to a user simply in terms of throughput. This however does not reflect the actual load in each technology that is required to achieve this ”net” throughput. Recent work targets at minimizing the load of the heterogeneous system while still meeting QoS constraints of different user types. Gazis et al. [3] applied a generic utility model and formulated a general concept as a Knapsack problem whereby the authors assume—without specifying them—that there exist appropriate mappings of QoS-to-resource, quality-to-utility, and resource-to-cost.

Fodor et al. [2] considered access selection in a coupled GSM/EDGE and WCDMA network. Despite service-based assignment, they propose an algorithm that considers the measured radio resource consumption of every single user. Although this outperforms the service-based algorithm in case of error-free measurement, it is shown that the performance greatly depends on the accuracy of the measured resource consumption.

Yilmaz et al. [18] investigated several access selection principles ranging from simple ”WLAN-if-coverage” and load-based SNR thresholds to schemes that base on the achievable bitrates and residual capacity in each RAN. According to their results, ”WLAN-if-coverage” performs well for low to moderate density of WLAN networks, while schemes based on load-based SNR thresholds and achievable bitrates perform best in case of high probability for WLAN coverage. The latter only outperforms others if information regarding signal quality and system load is sufficiently accurate.

This work contributes to access selection and handover strategies for heterogeneous networks as follows. We present a handover candidate selection scheme which targets to increase the utilization of allocated resources. Therefore, we derive a novel WLAN decision metric that extends the idea of using the ”real” radio resource consumption from previous work. The metric bases on the efficiency of the occupied airtime for transmissions on the wireless channel. This work shows that the metric allows for the identification of users with low utilization and studies the effect on users remaining in the cell in case of multiple handovers according to our scheme.

The paper is structured as follows. Firstly, Section 2 describes the system model and discusses the proposed cost function for WLAN hotspots. We derive the candidate selection approach with the decision metric for WLANs in Section 3. Section 4 presents the results of the simulation studies. Finally, conclusions are drawn in Section 5.

Chapter 2

System Model

The system under study consists of a single WLAN cell and another heterogeneous access technology (denoted as "AT_2" in the following). Thereby, the WLAN cell is completely within the coverage area of AT_2. We assume that all mobile terminals are equipped with a network interface card for each access technology such that they can perform vertical handovers.

Secondly, we assume that each access network has an own cost function reflecting the effort and the revenue serving a specific user who belongs to a certain traffic class. These cost functions are involved in two processes: the initial access selection and the (vertical) handover decision.

The access selection consists of two main steps: For terminals in the coverage area of WLAN and AT_2, the access selection takes places by computing and comparing cost function values of each access network. Accordingly, the choice will be made for the network, in which the mobile gains a better cost function value.

For the handover decision process, the involved access networks are divided into two conceptual parts, namely the originator and the recipient network. In principal, there exist three general concepts regarding the placement of the handover decision. This can be realized within the originator network, the recipient network, or by a separate arbitration entity.

In the following, we discuss the outstanding tasks for a handover if the decision is made in the originator network. In this approach, the originator network

1. identifies potential handover candidate(s),
2. estimates the gain due to the potential handovers,
3. requests cost function value estimates from the recipient network via appropriate means of signaling,
4. compares candidates' cost function values within originator and recipient network, and
5. finally decides for or against a vertical handover for each candidate.

Contrary, the recipient network

1. estimates the cost function value for each potential handover candidate currently served by the originator network, and
2. assesses the impact of a handover on other users.

As described above, we assume to have a cost function for each access network. For WLAN, the proposed cost function reflects

1. the load in each access network,
2. the utilization of allocated resources, and
3. the QoS level for every user.

Here, we consider a cost function that linearly combines parameters for these three parts:

$$\begin{gathered} c_{WLAN}(i) = \omega_1 \frac{t_a(i)}{\Delta t} + \omega_2 D(i) + \omega_3 QoS(i) \\ \text{with} \sum_{k=1}^{3} \omega_k = 1 \end{gathered} \tag{2.1}$$

While $t_a/\Delta t$ is the occupied airtime on the channel in relation to measurement interval Δt, D represents the (normalized) decision metric evaluating the resource utilization on behalf of each traffic stream. The QoS parameter separates into several parts dependent on the requested service. Here, we distinct dependent on different QoS classes. For VoIP, it would consist of the end-to-end delay, jitter, and packet loss (normalized by their maximum tolerable values).

A handover for user i from WLAN to AT_2 takes place, iff

$$c_{WLAN}(i) > c_{WLAN}(j) \;\; \forall \text{ users } j \neq i \tag{2.2}$$

$$c_{WLAN}(i) > c_{AT_2}(i) \tag{2.3}$$

Eq. 2.2 represents the identification of potential handover candidates within the WLAN cell, i.e., the selection of the user with the highest cost function value. Eq. 2.3 describes the comparison of candidate's cost function value in WLAN and AT_2. Only in case that his value is significantly better in AT_2, a handover will be triggered. This part is indispensable, since serving the user with the highest costs in WLAN may still be cheaper than putting him into AT_2.

Chapter 3

Inefficiency Metric

The goal of the proposed metric is to reflect the total utilization of allocated resources. This goal leads to the key question how to identify the parts (e.g., certain users or traffic flows) that contribute to the load in the access network drastically but benefit only marginally from these expenditures. Such behavior is denoted as ”inefficiency” in the following.

In radio technologies such as WLAN, the load evoked by a user depends on various factors such as path loss, fading, and interference. In order to maximize the total number of users in the system without violating their QoS constraints, we follow the approach that the most ”inefficient” users are selected as handover candidates.

Assuming ideal conditions (i.e., no path loss, collisions, packet errors, etc.) results in transmissions evoking the lowest possible load on the channel. Contrary, all means for error control and adoption to channel conditions (e.g., power control, rate adaptation, retransmissions) lead to an increase of the load on the wireless channel in real systems.

The design goal for the decision metric is to find a measure for the expenditures required to deal with these realistic conditions. It consists of two parts: the surcharge ζ and the overhead factor ϱ. While the surcharge is a measure for additional expenditures required for error control and correction, the overhead factor allows for an evaluation of different data packet sizes regarding their suitability in WLANs. In the following, both parts as well as their composition to the final inefficiency metric are discussed in detail.

3.1 Surcharge

This part is derived from the very basic definition of efficiency: In engineering, efficiency is usually defined as relation of system's output ϑ to the overall effort ψ one has to insert:

$$\eta = \frac{output}{effort} = \frac{\vartheta}{\psi} \tag{3.1}$$

Efficiency η can range in the interval $[0, 1]$, whereby effort values much larger than output values ($\psi >> \vartheta$) lead asymptotically towards efficiency values of zero.

The design rationale behind this part is to identify terminals with smallest efficiency values as handover candidates. However, it may be difficult to distinguish between two very small efficiency values near zero although the corresponding difference of effort values may be significant. Hence, not the efficiency itself but its reciprocal is applied to enable comparability.

$$\textit{surcharge}\ \zeta = \eta^{-1} = \frac{\psi}{\vartheta} \tag{3.2}$$

In WLANs, the effort ψ for a single transmission of an MPDU depends on the state of the wireless channel, the choice of a modulation scheme, the collision level as well as the number of retransmissions. All these parts have impact on the effort for a transmission in a way that they affect its duration. Thus, it is straightforward to consider the duration for a complete transmission sequence in order to determine its effort (Eq. 3.3). There, the number of trials represents the (re)transmissions that have been required to ensure the delivery of the MSDU.

$$\psi = t_a = \sum_{i=0}^{\#trials} \Delta t_i \tag{3.3}$$

For each trial, Δt_i represents the amount of time that the wireless medium is occupied (or reserved, in case of inter-frame spaces and NAV settings*):

$$\Delta t_i = t_{IFS} + t_d(Rate_i) + t_{ack} \tag{3.4}$$

This includes the whole transmission sequence consisting of the inter-frame spaces DIFS or AIFS and SIFS (t_{IFS}), the duration t_d of the complete data frame ”on air”, where the data part is encoded with a certain modulation scheme $Rate_i$ and the acknowledgment t_{ack}.

Secondly, we define system's output at MAC level as the smallest possible duration for the whole transmission that would be required in case of an ideal error free channel (Eq. 3.5).

$$\vartheta = \Delta t_{opt} = t_{IFS} + t_d(max\, Rate) + t_{ack} \tag{3.5}$$

Note that the output definition includes the duration of the whole data frame when the data part is encoded with the highest modulation $max\, Rate$ and only the single transmission attempt. Thus it serves as a reference case that implies the smallest possible effort.

3.2 Overhead Factor

While the surcharge is a measure for the efficiency regarding the transmission of MSDUs, it does not tell anything about the suitability of WLANs to transport these MSDUs with their specific size. IEEE 802.11 introduces a fixed amount of overhead (PHY framing, inter-frame spaces and immediate ACK) for one transmission regardless of the MSDU size. Thus, the smaller the MSDU size, the less becomes 802.11 optimally utilized. To accommodate this behavior, we introduce the overhead factor as

$$\alpha = \frac{\Delta t_{opt}}{\Delta t_{MSDU_{opt}}} \tag{3.6}$$

While Δt_{opt} is again the smallest possible duration for a frame exchange (Eq. 3.5), Δt_{MSDU} represents the duration of the bare MSDU assuming the highest modulation scheme.

*The backoff duration does not apply here, since only the occupation of the channel is of interest.

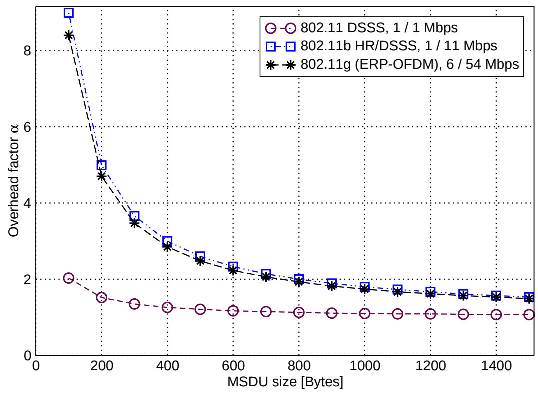


Figure 3.1: Overhead factors of 802.11/b/g PHYs

Figure 3.1 displays the overhead factors for different MSDU sizes and three different 802.11 PHYs (figure's legend box specifies PHYs' basic / data rate, other parameters according to [4]). 802.11 and 802.11b curves clearly show that the higher the data rates, the higher is the overhead especially for small MSDU sizes such as VoIP (e.g., 200 Bytes in case of G.711-coded speech and a packetization of 20 ms). The overhead values of 802.11g ERP OFDM PHYs are slightly lower than the 802.11b curve. This results from the fact that 802.11g ERP OFDM comes up with smaller slot times, shorter PLCP preamble and header as well as a higher basic rate of 6 Mbps.

3.3 Metric Composition

In order to allow the handover candidate selection among users with heterogeneous traffic patterns, overhead factor α and surcharge ζ are combined to the inefficiency metric:

$$D = \alpha\zeta \tag{3.7}$$

Although the metric design has been discussed for a transmission of a single MPDU only, it is simple to extend it over multiple transmissions / larger time scales. Just by calculating resp. measuring output ϑ and effort ψ over fixed-size intervals, one is able to compute the surcharge value afterwards. A detailed discussion about the interval size for WLAN is included in [17]. Within this work, $100\,ms$ were applied.

Chapter 4

Impact of Inefficiency Metric

4.1 Goals of Investigation

Firstly, the simulative proof-of-concept study shows that the inefficiency metric allows to identify candidates in a WLAN network for a (heterogeneous) handover. Secondly, we identify the impact of the selection scheme on users remaining in the WLAN cell, if a single candidate performs a handover from WLAN to AT_2. Thirdly, we are interested in the impact of multiple handovers according to our approach: There, we handover several candidates from WLAN to AT_2, while the same number of users (with the same service type) are put from AT_2 to WLAN.

4.2 Set of Experiments

For the above goals, a set of three experiments has been performed:

1. Max. #nodes,
2. Reduced set, and
3. Replaced set(s)

The first experiment determines the maximum number of nodes such that the WLAN network is loaded (but not saturated) in a way that the QoS constraints of at least one node are violated.

Secondly, we show the impact of a single handover from WLAN to AT_2 when choosing the most ”inefficient” WLAN user. This experiment is called ”reduced set” since the total number of WLAN users decreases. In comparison to the maximum number of nodes, this experiment gives an idea about the approximate range of improvements due to the single handover of the most ”inefficient” user.

Thirdly, we study the impact of multiple handovers according to our strategy. There, we conduct a replacement of nodes based on the results of the ”max. #node” experiment. Under replacement, we understand here that a node with a high metric value is triggered to perform a handover from WLAN to AT_2, while the WLAN network accommodates another node (either due to a handover from AT_2 or a new, arriving user). Here, it is assumed that this new node is present near the AP with a distance of $10\ m$ and represents the same user type as the one put from WLAN to AT_2. This third experiment is conducted with one to three replacements in total.

4.3 Simulation Scenario

The scenario under study consists of a WLAN network interconnected to a heterogeneous access technology AT_2. It is assumed that AT_2 provides full coverage in the region of interest and is capable to serve all arriving users. Contrary, the WLAN part consists of a single hotspot that covers only a certain part of the area, e.g., like a departure lounge in an air-port.

The considered IEEE 802.11g AP is 11e-capable by providing EDCA functionality. We assume to have VoIP users only, which are equally distributed over the area of interest. Users are stationary and equipped with AT_2 as well as with WLAN devices. The latter applies 802.11g Extended Rate Physicals (ERPs) with OFDM modulation—from 6 up to 54 Mbps. The 802.11e/g parameters were chosen according to the default values of [4]. *

In the large-scale in-house environment described above, radio signals are not only affected due to path loss but also due to multipath propagation. Path loss of radio signals is modelled by the TwoRayground model of ns-2. For multi-path propagation, the Ricean fading model of Punnose et al. [11] was applied, whereby the slow movements of the environment have been set to $2\,km/h$. The Ricean K-Factor, which specifies the ratio between the amount of signal power received on line of sight and the variance of the multipath [12], was set according to the measurements of [15] to $3\,dB$. A wireless channel with such characteristics requires an appropriate rate selection algorithm. For this we implemented Adaptive Automatic Rate Fallback (AARF) of Lacage et al. [9].†

As discussed by Kochut et al. [7], ns-2's wireless channel model for WLANs does not accurately model capture effects. With Ricean fading, capture effects may also occur in our investigations. Therefore, ns-2's Wireless PHY model has been extended by an SINR (Signal to Interference plus Noise Ratio) part, where each receiver keeps track of the power level present on the wireless channel. For each single packet arriving at the receiver, this allows firstly to decide whether

- the PLCP preamble can be decoded correctly,
- the SINR is large enough to decode the PSDU (which may be transmitted at higher data rates).

Secondly, in case of multiple arriving packets at receiver's PHY, the novel SINR add-on allows to decide whether

- a frame captures others,
- an arriving frame is too weak to harm an already receiving one,
- a collision occurs between multiple frames.

4.4 Node placement and Traffic Model

In the simulation scenario, WLAN VoIP nodes are distributed equally over the area of interest, whereby the AP is located at the corner of the considered environment, such that no hidden nodes appear (Figure 4.1).

*The TXOPLimits were set to zero such that a single transmission per medium access attempt is performed.

†An overview of rate-adaptation mechanisms as well as a short explanation of AARF together with the results of a single-terminal test simulation is provided in [17].

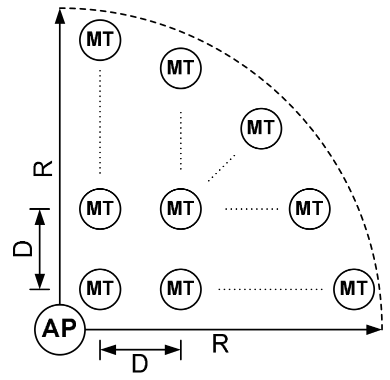


Figure 4.1: WLAN Scenario

All wireless stations have a VoIP call with a wired node outside the WLAN. The delay between AP and the wired nodes was set to $100\,ms$. All stations use an exponential ON/OFF model to generate voice traffic. According to ITU-T Recommendation P.59 [5], mean ON and OFF durations of $1.004\ s$ and $1.587\ s$ have been applied. During the ON periods, voice packets are generated according to the ITU-T codec G.711 with a packetization of $20\,ms$, i.e., each voice flow has a $64\ kbps$ peak rate with $160\,Byte$ audio packets.

4.5 Metrics and QoS constraints

4.5.1 Surcharge

Each transmitting station determines its surcharge value over an interval of $100\ ms$ as described in Section 3. Therefore, all stations measure their injected traffic in the uplink, while the AP observes the traffic to each STA in the downlink. The surcharge values are calculated only if there were any transmission attempts during the interval.‡

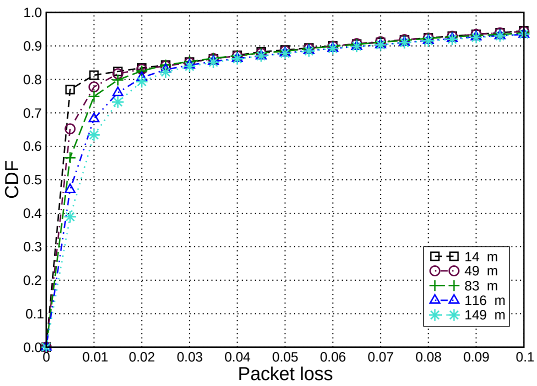
4.5.2 Application-level losses

In order to assess the quality of the VoIP calls, we measure the loss of audio packets on application level over certain intervals. A loss can either occur due to lost or late packets. A packet is considered to be late if it arrives after a maximum network delay of $150\ ms$ (similar to [8]) at the receiver such that it cannot be played out anymore.

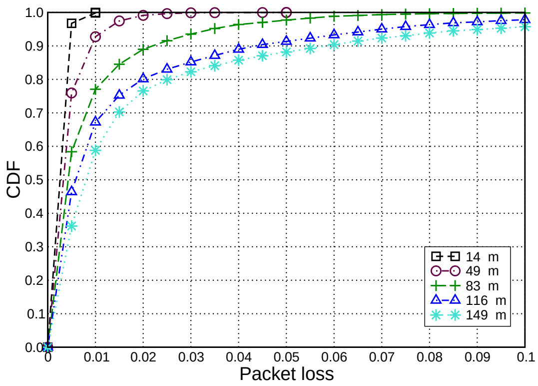
4.5.3 QoS constraints

For every single VoIP call, the quality should stay on an acceptable level. ”Acceptable” thereby means that a certain boundary for application level losses—consisting of packet losses and late packets—is not violated. In the following, we discuss the choice of this QoS boundary.

‡Since this work considers the same VoIP traffic pattern for all nodes, the overhead factor is just a constant value. Thus we focus on the surcharge part of the inefficiency metric in these simulation studies.

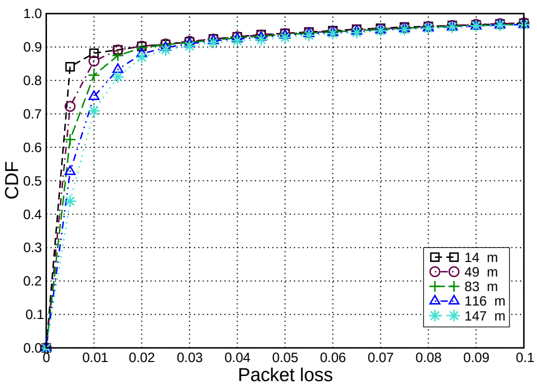


(a) Downlink

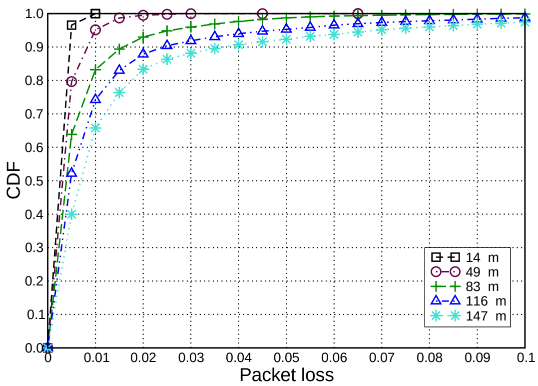


(b) Uplink

Figure 4.2: Max. #nodes (48 nodes), application-level losses of voice packets



(a) Downlink

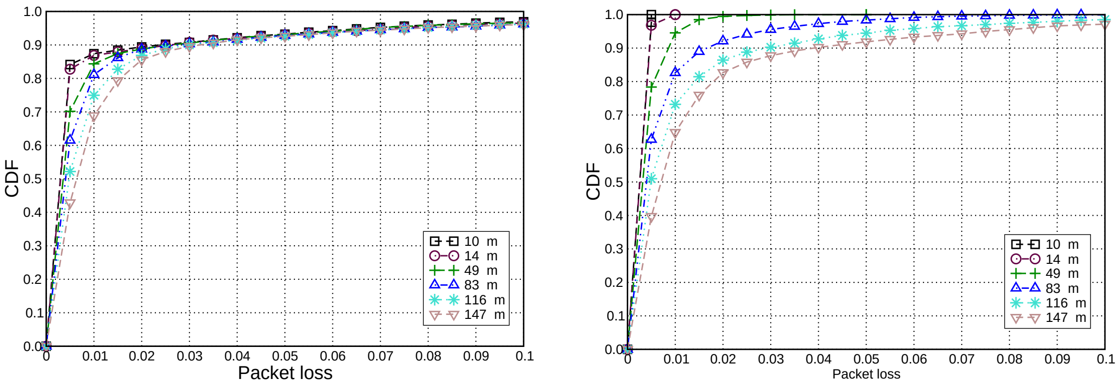


(b) Uplink

Figure 4.3: Reduced set (47 nodes), application-level losses of voice packets

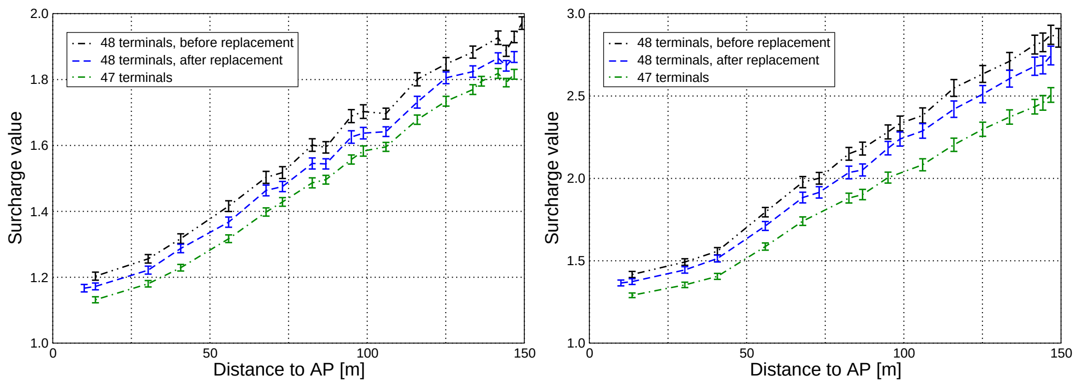
With Packet Loss Concealment (PLC) schemes and one-way delays up to 200 ms, random losses of up to 5 percent for G.711 are acceptable [14, p. 38, Fig. 29]. If five or more percent of the VoIP packets are lost, i.e., they have been dropped or they arrive with a network delay larger than 150 ms, the perceived quality is assumed to be temporary lousy. The interval over which this criterion is evaluated has been set to 4 seconds. §

In our work, the QoS boundary is defined as follows: If the perceived quality is temporary lousy in 10 or more percent of the overall number of intervals, the quality degradation of the complete call is defined as not acceptable anymore.



(a) Downlink, after 2nd replacement (b) Uplink, after 2nd replacement

Figure 4.4: Application-level Losses of voice packets



(a) Downlink (b) Uplink

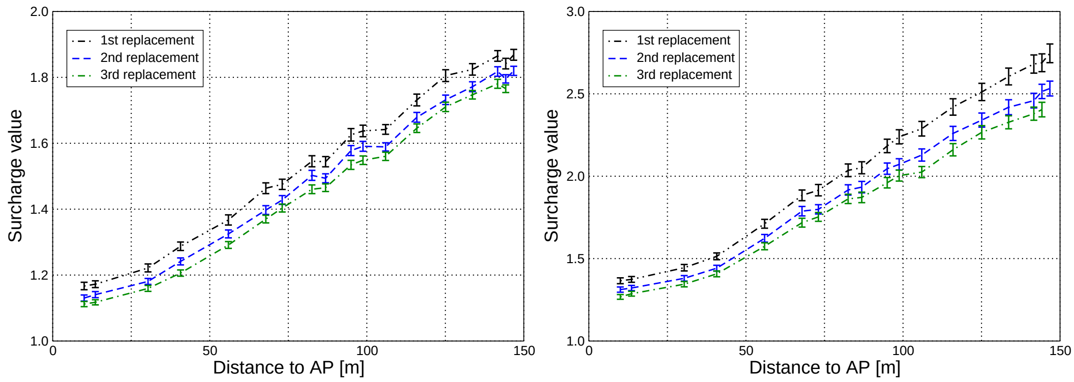
Figure 4.5: Comparison of surcharge values from all three experiments

4.6 Results

All surcharge results have been evaluated by batch means analysis, whereby the batch sizes were chosen such that the relation of the autocovariance between successive batch means to the variance [6] stayed below 4 percent. All mean values are plotted with their 95 percent confidence interval.

In the first experiment, the maximum number of nodes has been determined such that the QoS constraints of at least a single node are violated. This is achieved with 48 VoIP nodes in total. Figure 4.2(a) and 4.2(b) show the cumulative distribution function (CDF) of application-level losses for down- and uplink direction. While the QoS-boundary is violated for all nodes in the downlink, the losses depend greatly on the distance between AP and STAs for the uplink, where boundaries are crossed for far nodes, only. This effect results from the asymmetric traffic distribution between AP

§It was chosen in the order of seconds so that one is able to get an impression about the incidence of periods with frequent (non-random) losses.



(a) Downlink (b) Uplink

Figure 4.6: Surcharge values after one, two, and three replacements

and STAs and is discussed further below.

After identifying the operational point of the network where QoS constraints of several clients are being violated, the second set of experiments shows the impact of a single handover. There, the handover candidate was selected according to the novel strategy of selecting the most ”inefficient” user. After this single handover, i.e., 47 active VoIP nodes in total, the packet loss is below 3 and 4 percent in 90 percent of the evaluation intervals for downlink and uplink direction, respectively (Figure 4.3). Thus, QoS constraints as defined in Section 4.5.3 are met for all 47 nodes due to a single handover following the efficiency-aware selection approach.

Now, let’s consider the effect of a single replacement, i.e., the most inefficient node is triggered to perform a handover from WLAN to the AT_2 network, while the WLAN network accommodates another VoIP node (with a distance of 10 meters to the AP). Figure 4.5 shows the surcharge values in up- as well as downlink direction for all three experiments. There, the surcharge values increase with larger distances between nodes and AP. This result was expected since the probability for lower data rates and higher number of retransmissions increases with the distance. All these impacts are now unified in the single surcharge metric. Not surprisingly, the ”max. #nodes” experiment results in highest surcharge values for all nodes, while the single replacement experiment leads to a significant reduction: the surcharge values drop by around 2.3 to 3.9 percent (downlink) and 2.9 to 5.9 percent (uplink). Lowest surcharge values are gained with the ”reduced set” experiment, where the most inefficient node was selected as handover candidate. There, the surcharge values of all other remaining nodes drop by 3.6 to 7.9 percent in the downlink and 3.5 to 12 percent in the uplink compared to ”max. #nodes” results.¶

It attracts attention that surcharge values are higher for the up- than for the downlink direction. This stems from the asymmetric traffic conditions: the AP has to serve 48 VoIP streams in the downlink, i.e., 48 times more traffic than each single VoIP node in the uplink. This asymmetric traffic distribution leads to a lower a collision probability for the AP. Beside other aspects, Cai et al. investigated this effect already analytically in their work [1] . The discrepancy between up- and downlink amplifies here, since the collision level has also impact on the rate adaptation scheme.

¶Although being relatively close to each other, confidence intervals at each single distance do not overlap.

Table 4.1: Uplink: Quantiles at 5-percent packet loss

	Distance to AP [m]						
	14	49	83	116	134	144	149
max. #nodes	1.0	1.0	0.98	0.91	0.89	0.88	0.88
reduced set	1.0	1.0	0.99	0.95	0.93	0.93	—
1st replacement	1.0	1.0	0.98	0.93	0.90	0.90	—
2nd replacement	1.0	1.0	0.98	0.94	0.93	0.92	—
3rd replacement	1.0	1.0	0.99	0.95	0.94	0.94	—

Table 4.2: Downlink: Quantiles at 5-percent packet loss

	Distance to AP [m]						
	14	49	83	116	134	144	149
max. #nodes	0.89	0.89	0.89	0.88	0.88	0.88	0.88
reduced set	0.94	0.94	0.94	0.93	0.93	0.93	—
1st replacement	0.91	0.91	0.91	0.90	0.90	0.90	—
2nd replacement	0.93	0.93	0.93	0.93	0.93	0.93	—
3rd replacement	0.95	0.94	0.94	0.94	0.95	0.94	—

From the surcharge curves, one can see that there exist certain plateaus at distances from 68 to 73, 83 to 87, and 99 to 106 meters. This results from the fact that nodes within a certain region are likely to have similar SINR values on average, thus being able to meet SNR thresholds for similar modulation schemes.

The positive impact of further replacements is displayed in Fig 4.6, again for up- as well as downlink. While the second replacement leads again to a relatively large decrease, no significant improvement was gained with the third replacement (i.e., confidence intervals of 2nd and 3rd replacements overlap at several distances).

Lastly, we consider the impact of replacements on users' QoS. Tables 4.1 and 4.2 show the cumulative probability of having five or less percent of application losses for all experiments in up- and downlink. While the first replacement does not improve the application losses greatly for up -and downlink, it is the second replacement that avoids a violation of QoS constraints. From Fig. 4.4(a) and 4.4(b), we can observe that less than 4 percent losses occur in 90 percent of the intervals for up- as well as downlink direction. Now, the third replacement brings users' QoS up to level of the reduced-set experiment, which means that we gain comparable quality although there are 48 instead of 47 nodes. Interestingly, there are only small differences in QoS values between 2nd and 3rd replacement. This is directly in line with the surcharge results, where confidence intervals overlap such that there's no significant difference at certain points anymore. From the replacement study we observe the interesting aspect that a non-significant impact of a replacement on the surcharge also implies only marginal differences in users' QoS in case of VoIP traffic.

Chapter 5

Conclusions and Future Work

This work extends approaches for handover decisions being based on the measured load for every user resp. each traffic flow. The novel metric selects most ”inefficient” users as handover candidates. ”Inefficient” thereby refers to parts contributing to the cell load greatly but benefiting only marginally from this effort. The two parts of the decision metric are discussed in detail by focussing on today’s IEE 802.11 networks. Firstly, proof-of-concept simulations are used to document that the novel metric is suitable to select most ”inefficient” users. Secondly, simulation results show the improvements for users remaining in the WLAN cell, after performing a handover of the most inefficient candidate. Finally, we study the impact of our scheme in case of multiple handovers, where ”inefficient” WLAN users are replaced by suitable candidates from other heterogeneous access networks.

As future work, we do not only consider the investigation of our approach with a mixture of elastic and inelastic traffic but also a study on the impact of user mobility on our selection scheme.

Acknowledgments

The authors thank Marc Emmelmann for the numerous discussions regarding the presentation of the approach and the results. During the final phase, his valuable review comments improved this work a lot.

Bibliography

[1] L.X. Cai, Xuemin Shen, J.W. Mark, L. Cai, and Yang Xiao. Voice capacity analysis of WLAN with unbalanced traffic. *Vehicular Technology, IEEE Transactions on*, 55(3):752–761, May 2006.

[2] G. Fodor, A. Furuskar, and J. Lundsjo. On access selection techniques in always best connected networks. In *ITC Specialist Seminar on Performance Evaluation of Wireless and Mobile Systems*, August 2004.

[3] V. Gazis, N. Alonistioti, and L. Merakos. Toward a generic ”always best connected” capability in integrated WLAN/UMTS cellular mobile networks (and beyond). *Wireless Communications, IEEE*, 12(3):20–29, June 2005.

[4] IEEE 802.11. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Std 802.11-2007, June 2007.

[5] ITU-T Recommendation P.59. Artifical Conversational Speech, March 1993.

[6] Raj Jain. *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling.* John Wiley & Sons Inc., May 1991.

[7] A. Kochut, A. Vasan, A. U. Shankar, and A. Agrawala. Sniffing out the correct physical layer capture model in 802.11b. In *Network Protocols, 2004. ICNP 2004. Proceedings of the 12th IEEE International Conference on*, pages 252–261, IEEE, 2004.

[8] A. Koepsel and A. Wolisz. Voice transmission in an IEEE 802.11 WLAN based access network. In *In Proc. of WoWMoM 2001*, pages 24–33, July 2001.

[9] Mathieu Lacage, Mohammad Hossein Manshaei, and Thierry Turletti. IEEE 802.11 rate adaptation: a practical approach. In *Proc. of MSWiM 2004*, pages 126–134, New York, NY, USA, 2004. ACM Press.

[10] Janise McNair and Fang Zhu. Vertical handoffs in fourth-generation multinetwork environments. *Wireless Communications, IEEE*, 11(3):8–15, June 2004.

[11] Ratish J. Punnoose, Pavel V. Nikitin, and Daniel D. Stancil. Efficient Simulation of Ricean Fading within a Packet Simulator. In *Vehicular Technology Conference*, volume 2, pages 764–767. IEEE VTS-Fall VTC 2000. 52nd, 24–28 September 2000.

[12] Theodor S. Rappaport. *Wireless Communications—Principles and Practice*. Prentice Hall, 2002.

[13] Enrique Stevens-Navarro and Vincent W.S. Wong. Comparison between vertical handoff decision algorithms for heterogeneous wireless networks. In *Proc of 63rd Vehicular Technology Conference, VTC 2006-Spring.*, volume 2, pages 947–951. IEEE, 2006.

[14] Telecommunications Industry Association (TIA). TSB-116-A: Voice Quality Recommendations for IP Telephony, March 2006.

[15] E. Walker, H.J. Zepernick, and T.Wysocki. Fading Measurements at 2.4 GHz for the Indoor Radio Propagation Channel. In *Proc. of IZS 1998*, pages 171–176, February 17–19, 1998.

[16] H. J. Wang, R. H. Katz, and J. Giese. Policy-enabled handoffs across heterogeneous wireless networks. In *WMCSA '99: Proceedings of the Second IEEE Workshop on Mobile Computer Systems and Applications*, page 51, Washington, DC, USA, 1999. IEEE Computer Society.

[17] S. Wiethölter. Virtual Utilization and VoIP Capacity of WLANs Supporting a Mix of Data Rates. Technical Report TKN-05-004, Telecommunication Networks Group, Technische Universität Berlin, 2005.

[18] O. Yilmaz, A. Furuskar, J. Pettersson, and A. Simonsson. Access selection in WCDMA and WLAN multi-access networks. In *in Proc. of Vehicular Technology Conference, VTC Spring, IEEE*, volume 4, pages 2220–2224, 2005.