University of Würzburg
Institute of Computer Science
Research Report Series

UMTS Load Metrics for a Policy-Based Vertical Handover Framework

Rastin Pries, Andreas Mäder, Dirk Staehle

Report No. 395 July 2008

University of Würzburg
Department of Computer Science
Am Hubland, D-97074 Würzburg, Germany
{pries,maeder,staehle}@informatik.uni-wuerzburg.de

UMTS Load Metrics for a Policy-Based Vertical Handover Framework

Rastin Pries, Andreas Mäder, Dirk Staehle
University of Würzburg
Department of Computer Science
Am Hubland, D-97074 Würzburg, Germany
{pries,maeder,staehle}@informatik.
uni-wuerzburg.de

Abstract

In most European countries, UMTS networks have been set up and completely cover urban areas. At the beginning of UMTS, there were only a few users but with a special pricing policy, the number of users is increasing drastically. A few providers already offer flat rates for UMTS and their networks get more and more crowded. Therefore, an efficient radio resource management (RRM) is one of the key aspects. In UMTS, the capacity is limited by the allowed NodeB transmit power on the downlink, the interference on the uplink, and the number of used channelization codes. In this technical report, we will show how to examine both loads and furthermore, compute the code utilization.

1 Introduction

Todays mobile phones are equipped with several wireless devices like UMTS, GPRS, and WLAN. Which network device to use is normally up to the end user who can select the preferred technology. This is annoying for the user, who does not care about the network as long as the costs and the provided services are similar. With such devices, where you can choose a preferred technology, a lot of battery power is wasted because every wireless interface is turned on and searching for attachment points. Considering the network providers, they are interested in a near optimal load distribution in order to serve more customers.

In order to solve these problems, Wang et al. [1] introduced a policy-based handover mechanism across heterogeneous wireless networks. For the handover decision, they use three parameters, the power consumption of using a network access device, the costs for the network and the available bandwidth. Though, how to determine the available bandwidth is not explained. In this document, we evaluate the cell load for the UMTS network and the load each user produces. The user load is dominated by its position and the radio access bearer (RAB).

This document is organized as follows. Section 2 gives a detailed overview of how to perform radio resource management (RRM) in UMTS. This is followed by Section 3 validating the analytical models for the load calculation using simple simulation scenarios. Finally, Section 4 concludes this report.

2 UMTS Load Metrics

The UMTS cell (or sector) capacity is limited by several factors which we use as an indicator for the load. We consider the following air interface radio resource metrics: uplink load, which depends on the received interference at the NodeB, downlink load, which is defined by the NodeB transmit power, and code utilization. In the literature, many models are available for the characterization for the uplink and downlink air interface of UMTS, some examples are [2, 3, 4, 5, 6]. We follow the general insight that the capacity is on the downlink limited by the maximum transmit power and on the uplink by the interference pole equation. The uplink and downlink loads are therefore "soft" measures, which are influenced by several stochastic variables like user positions, speed, and fading. In contrast, the code utilization is a "hard" measure, since it depends only on the number of active downlink connections.

We assume that the system is able to measure the most relevant parameters for the calculation of the load metrics either at the NodeB or at the UE. The measurements are defined in [7]. Measurements which are performed in the UEs are reported to the RNC with measurement reports, which are either triggered by the radio resource management or sent periodically. A special case is the distance between NodeBs and UEs, for which separate procedures exist. In [8], four positioning methods are described from which OTDOA (observed time difference of arrival), U-TDOA (uplink – time difference of arrival) and Assisted GPS provide sufficient accuracy for vertical handovers.

2.1 Uplink

The uplink load is defined as the ratio between the multiple access interference (MAI) and the MAI plus the thermal noise:

$$\eta_{ul} = \frac{\hat{I}_{or} + \hat{I}_{oc}}{W\hat{N}_0 + \hat{I}_{or} + \hat{I}_{oc}}, \tag{1}$$

with $\hat{I}_{or}$ as own-cell interference stemming from the users power controlled by the considered NodeB, $\hat{I}_{oc}$ as other-cell interference generated by mobiles in the surrounding sectors, $W = 3.84\,\text{Mcps}$ as system chip rate and $N_0 = -174\,\text{dBm/Hz}$ as one-sided thermal noise power density. The individual load one user contributes to the total load depends on its transmit power and location. If the UE is power controlled by the considered NodeB, the UE will contribute to the own-cell load, otherwise to the other-cell load. The own-cell interference is therefore the sum of the received transmit powers of all power controlled UEs:

$$\hat{I}_{or} = \sum_{k \in x} \hat{S}_{k,x}, \tag{2}$$

with $\hat{S}_{k,x}$ as received power for mobile k at NodeB x. Correspondingly, the other-cell interference is the sum of all received powers of all mobiles which are not power controlled by NodeB x:

$$\hat{I}_{oc} = \sum_{y \in \mathcal{L} \setminus x} \sum_{k \in y} \hat{S}_{k,y} \tag{3}$$

This means that every mobile in the network contributes to the own-cell interference at the own NodeB and to the other-cell interference at surrounding NodeBs. The extent of this contribution depends for the own-cell interference mainly on the bit rate, for the other-cell interference mainly on the position of the mobile in the cell. If the mobile is more on the cell edge, it is closer to the non-serving NodeBs and generates more interference. Due to the cell coupling effect, which describes the mutual dependence between own-and other-cell interference this relation is not exact, since the generated other-cell interference in a way "comes back" to the own cell because a higher other-cell interference also requires higher transmit powers to meet the target-E_b/N_0-values at the serving NodeB. However, since the UTRAN measures the SIR and reports the transmit powers of the UEs, the reported values already include this effects.

We redefine the uplink load in order to reflect the influence of the own- and other-cell interferences as following:

$$\eta_{ul} = \eta_{ul}^{or} + \eta_{ul}^{oc} = \frac{\hat{I}_{or}}{W\hat{N}_0 + \hat{I}_0} + \frac{\hat{I}_{oc}}{W\hat{N}_0 + \hat{I}_0}, \tag{4}$$

where $\hat{I}_0 = \hat{I}_{or} + \hat{I}_{oc}$. We propose to use the following definition for the load of a single user in the network:

$$\tilde{\omega}_k = \frac{\hat{S}_{k,x}}{W\hat{N}_0 + \hat{I}_{x,0}} + \sum_{y \in \mathcal{L} \setminus x} \frac{\hat{S}_{k,y}}{W\hat{N}_0 + \hat{I}_{y,0}} \tag{5}$$

Note that the second term denotes the interference the mobile k generates at the non-serving NodeBs. The total received interference in the denominators is measured at the NodeB as the *total received wideband power*. The received power at x can be calculated from the measured SIR. The received powers at $y \neq x$ is not immediately available, however it can be estimated from the position of the mobile as following:

$$\hat{S}_{k,y} = \hat{S}_{k,x} \cdot \hat{\Delta}_{k,x} \quad \text{where} \quad \hat{\Delta}_{k,x} = \frac{\hat{d}_{k,y}}{\hat{d}_{k,x}}, \tag{6}$$

where $\hat{d}_{k,x}$ is the estimated path loss between UE k and NodeB x. Note that the maximum value of $\hat{\Delta}_{k,y}$ is 1, since otherwise we assume that power control is transferred to NodeB x.

The own-cell uplink load is additive (meaning that the single user own-cell load for one user stays constant independent of the number of users in the system) due to the fact that the received power for a single mobile can be expressed by the following equation:

$$\hat{S}_{k,x} = \phi_k^{ul} \cdot (W\hat{N}_0 + \hat{I}_x), \tag{7}$$

where ϕ_k^{ul} is the service load factor (an effective bandwidth measure) which is defined as

$$\phi_k^{ul} = \frac{\hat{\varepsilon} R_k}{\hat{\varepsilon} R_k + W}. \tag{8}$$

Substituting the received powers in Eq. (4) with (8) shows that the own-cell load can also be expressed as

$$\eta_{ul,x}^{or} = \sum_{k \in x} \phi_k^{ul}. \tag{9}$$

2.2 Downlink

Although the WCDMA downlink is in principle interference limited, the main limiting factor is the maximum allowed transmit power of the NodeB, T_{max}, which is usually 10 W or 20 W. The reason is that in the downlink direction orthogonal codes are used, which reduce the intra-cell interference theoretically to zero. However, due to multi-path propagation, perfect orthogonality is destroyed such that a certain fraction α of the transmit power is seen as interference. The orthogonality factor α depends on the propagation situation and is usually between 0.2 and 0.6. The interference between NodeBs is not reduced, since the orthogonal codes are only used within a cell, such that the transmit powers of other cells are received with an orthogonality factor of 1. The downlink E_b/N_0-equation is

$$\hat{\varepsilon} = \frac{W}{R_k} \cdot \frac{\hat{T}_{x,k}\hat{d}_{k,x}}{W\hat{N}_0 + \sum_{y \in \mathcal{L} \setminus x} \hat{T}_y \hat{d}_{k,y} + \hat{d}_{k,x}\alpha_k(\hat{T}_{x,d} + \hat{T}_{x,c} - \hat{T}_{x,k})}. \tag{10}$$

Here, $\hat{T}_{x,k}$ is the transmit power at NodeB x for mobile k, R_k is the bit rate, $\hat{T}_{x,c}$ is the constant power for the pilot and shared channels and $\hat{T}_{x,d}$ is the combined transmit power for all dedicated channel users. Solving for the transmit power yields

$$\hat{T}_{x,k} = \phi_k^{dl} \hat{Q}_k, \tag{11}$$

with

$$\phi_k^{dl} = \frac{\hat{\varepsilon}_k R_k}{W + \alpha_k \hat{\varepsilon}_k R_k} \tag{12}$$

and

$$\hat{Q}_k = \frac{W \hat{N}_0}{\hat{d}_{x,k}} + \sum_{y \in \mathcal{L} \setminus x} \hat{T}_y \hat{\Delta}_{y,k} + \alpha_k (\hat{T}_{x,d} + \hat{T}_{x,c}). \tag{13}$$

The first factor ϕ_k^{dl} only depends on the bit rate and E_b/N_0, thus it is independent of the position of the users, while the second factor $\hat{Q}_k$ can be interpreted as positional load factor. The first term in (13) is nearly irrelevant due to the very low values for the path loss $\hat{\delta}_{k,x}$ in most cases. The second term describes the influence of the other-cell interference on the transmit power. It becomes relevant if the mobile is near the edge of the cell, such that the path gain ratio $\hat{\Delta} >> 0$. The third term describes the influence of the own-cell interference, which is nearly independent of the location of the user if we assume only small variations for the orthogonality factor α_k. In contrast to the uplink, where the load in a single cell environment is additive over the single user loads regardless of the user positions, the downlink load is not additive due to the positional load factor. We therefore propose a load definition which is additive in the single cell case. For this reason, we add a correctional factor ξ to the transmit power which leads to the definition of $\hat{T}^*_{x,k}$, which denotes the transmit power for a single user under the assumption that the NodeB sends with the maximum transmit power $\hat{T}_{x,max}$:

$$\hat{T}^*_{x,k} = \hat{T}_{x,k} + \xi_k \quad \text{with} \quad \xi_k = \alpha_k \phi_k^{dl} \cdot (\hat{T}_{x,max} - \hat{T}_{x,c} - \hat{T}_{x,d}), \tag{14}$$

and

$$\hat{T}^*_{x,d} = \sum_{k \in x} \hat{T}^*_{x,k}. \tag{15}$$

The downlink cell load is then defined as

$$\eta_x^{dl} = \frac{\hat{T}^*_{x,d}}{\hat{T}_{x,\max} - \hat{T}_{x,c}}. \tag{16}$$

The single user load is correspondingly defined as

$$\omega_k^{dl} = \frac{\hat{T}^*_{x,k}}{\hat{T}_{x,max} - \hat{T}_{x,c}}. \tag{17}$$

The term in the denominator in both equations denotes the maximum allowed transmit power for the dedicated channels. In contrast to the uplink, the downlink load does not distinguish between own- and other-cell load, since the cell coupling effects are already included with the positional load factor. With the exception of the orthogonality factor, all relevant variables are available on the NodeB. The E_b/N_0 values can be approximated with the current target-E_b/N_0 value of the radio bearer.

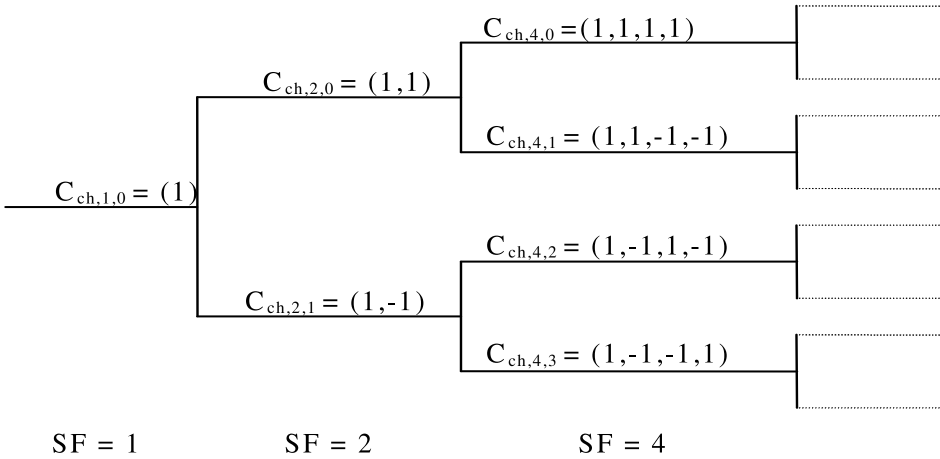


Figure 1: Orthogonal variable spreading factor code tree

2.3 Code Utilization

A further limiting factor are the number of codes available on the downlink. The orthogonal variable spreading factor (OSVF) codes are taken from the OSVF code tree, which is shown in Figure 1. As it can be seen in the figure, the number of available codes depends on the spreading factor. If a code is in use, all codes in the subtree of this code cannot be used anymore. This makes the OSVF codes a quite scarce resource, which is often the most dominant limiting factor on the UMTS downlink, see e.g. [9]. For example, the lowest spreading factor used in UMTS FDD mode is 8 for the 384 kbps bearer. Since one code (i.e. subtree) is reserved for the common and signaling channels, the number of 384 kbps bearers in one cell is limited to 7.

The highest possible spreading factor is $\text{SF}_{max} = 512$. All other spreading factors can be expressed as a fraction of the maximum spreading factor, such that

$$\text{SF}_k = \frac{\text{SF}_{max}}{n_k}, \tag{18}$$

where n_k is the number of maximal spreading codes which are blocked by SF_k. For example, the spreading code SF_8 occupies an equivalent number of $n_k = \frac{512}{8} = 64$ SF_{max} spreading codes.

For the *downlink code utilization* we must further consider the number of minimal spreading codes which are reserved for signaling and common channels, n_R, which we assume as $n_R = 32$. The number of available maximum spreading codes is therefore $n_A = 512 - n_R = 480$. The total code utilization at a NodeB x then the number of occupied maximum spreading codes divided by the number of available codes:

$$\eta_{x,c} = \sum_{k \in x} \frac{n_k}{n_{x,A}}. \tag{19}$$

If a mobile is in soft handover, a radio bearer is established to all NodeBs in the active set of the mobile. Consequently, the mobiles occupies a spreading code for each connection, which leads to the following definition of the *single user code utilization*:

$$\omega_k^c = \sum_{x \in AS} \frac{n_k}{n_{x,A}}. \tag{20}$$

3 Evaluation

First of all, we want to evaluate the importance of the RAB assignment. All UMTS service providers assign a fixed RAB to the users, no matter what bearer is really needed. To show the influence on the load, we simulate a low-rate Voice over IP (VoIP) scenario with several RAB. The ITU-T G.729 [10] voice codec is used with a data rate of 8 kbps and a frame size of 10 ms. If we add the RTP, UDP, IPv4 overhead, the total data rate is 40 kbps. We place four users in the simulation scenario and each of these users has a distance of 750 m to the NodeB. The users start their voice conversation with an offset of 200 seconds. The RAB is increased from 64 kbps up to 192 kbps on both, the uplink and the downlink. The results for the downlink load can be seen in Figure 2 and for the uplink in Figure 3. It is obvious that the produce load on both, the uplink and the downlink, increases with increased

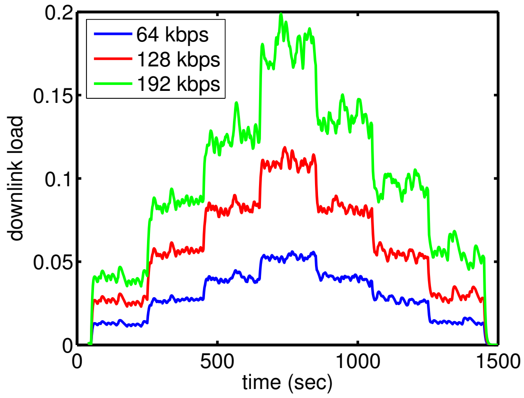


Figure 2: Downlink load for different RAB

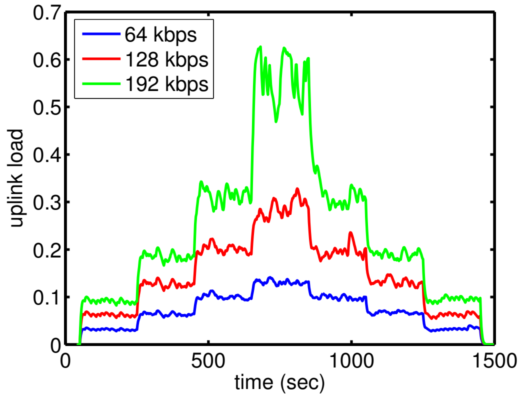
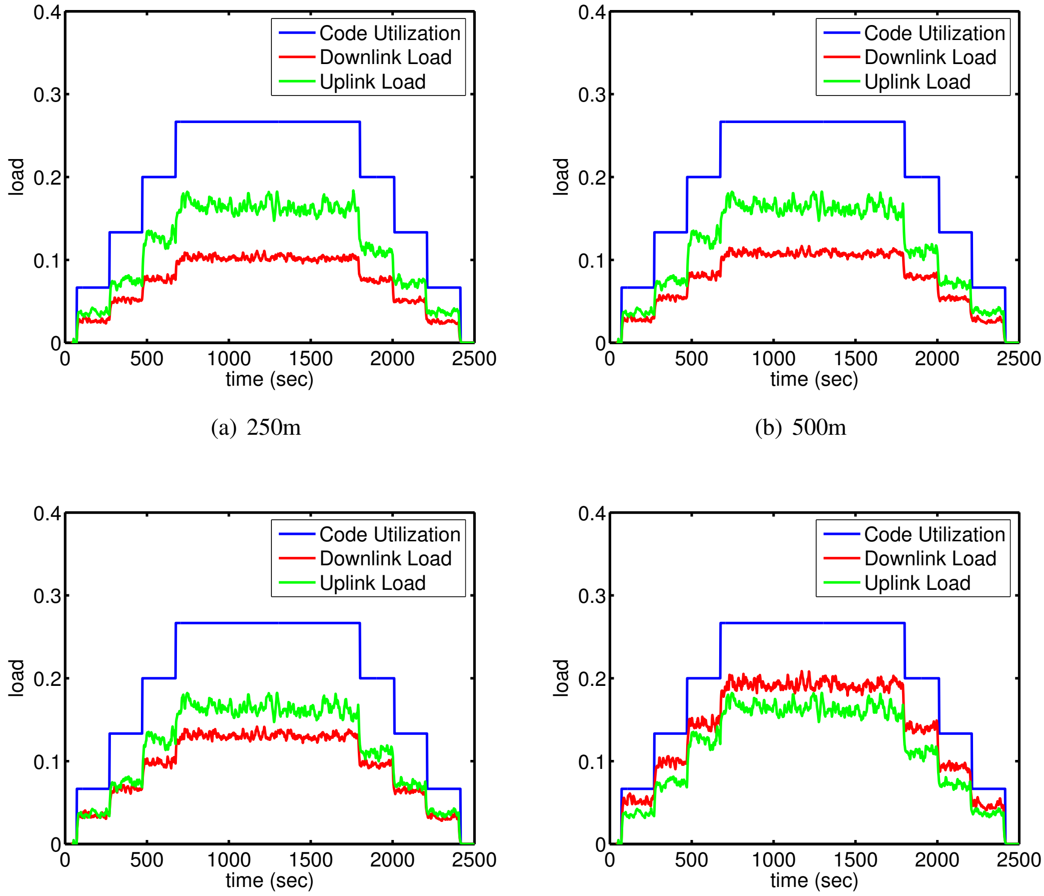


Figure 3: Uplink load for different RAB

RABs. If we compare the load on the uplink when every of the four users have started their voice conversation, the total uplink load is 5.5 times higher when assigning a 192 kbps RAB instead of a sufficient 64 kbps RAB. The RAB has to be assigned in communication with the used application. This can be really complex but using the IP Multimedia Subsystem (IMS) might be one possibility.
Now, we take a look on the three different loads for different distances to the NodeB. We still use four users, but they are downloading an FTP file instead of performing a voice conversation. The FTP file size is 20 MBytes and the users start the download with an offset of again 200 seconds. We set the RAB to 128 kbps for the uplink and downlink direction. Four different distances are used, 250 m, 500 m, 750 m, and 1000 m. Figure 4 shows the results.

The increasing steps at the code utilization curves show the time instants when the users start their FTP download. The end of the download is shown with the decreasing code use after around 1700 seconds. The code utilization and the uplink load stays the same, no matter how far the users are away from the NodeB. The only load which increases with the distance is the downlink load. This was already described in Section 2. Here, the load increases from around 10 percent for 250 and 500 meters to 20 percent for 1000 meters distance to the NodeB. The inner circle with around 500 meters or less to the NodeB is therefore not at interesting as the circle between 500 and 1000 meters distance.



(a) 250m

(b) 500m

(c) 750m

(d) 1000m

Figure 4: Uplink Load, Downlink Load, and Code Utilization for the following distances to the NodeB: (a) 250m,(b) 500m,(c) 750m, and(d) 1000m.

For the first simulation runs, we have only considered single cell scenarios without any interference from neighboring cells. Let us now also take these neighboring cells into account. We place 6 neighboring cells around our considered NodeB. Each of these cells has a NodeB with an isotropic antenna and is connected to the same RNC as our first NodeB. For this simulation setup, we increase the number of users in the neighboring cells. For our first NodeB, the number of users is stepwise increased to four. For the neighboring cells, we start simulation runs with one user in each cell and increase the users to five resulting in 34 users in total. The users in the neighboring cell have a distance of 500 meters to "their" NodeB and a distance of 1250 meters to the first NodeB. Our four users of the center cell have a distance of 750 meter to the NodeB.

One, two, or three users in the neighboring cells do not have any influence on the load in our considered cell. Only when we increase the number of users in the neighboring cells to more than three, the uplink load is influenced. The results are shown in Figure 5.

The code utilization and the downlink load are just illustrated once because they do not change with increasing other cell interference. The uplink load curve for three interferers in each neighboring cell

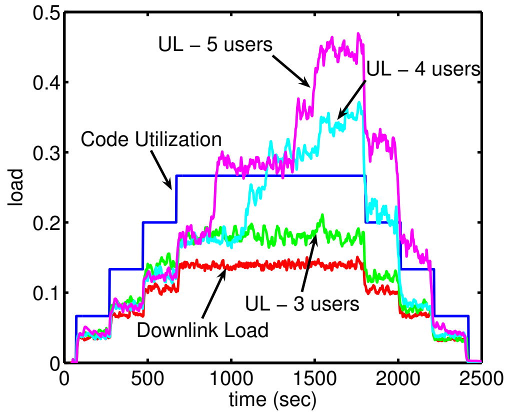


Figure 5: Increasing number of other cell users

looks similar to the one from Fig. 4(c). If we further increase the number of interferers to four or five, the uplink load even exceed the code utilization. Both curves, for four and five interferers still show a stepwise behavior when already every user is downloading the FTP file. The reason for this increase is the implemented outer loop power control. The own cell interference and the other cell interference result in a lot of packet loss. With each lost packet, the UE power is increase by one dB, whereas the power is decreased only by one dB multiplied with the BLER ($\frac{1}{500}$) in case of a successful transmitted packet. The load is decreased when some users have finished their download and the interference decreases.

Finally, we want to evaluate the load in a mobile environment. Therefore, we are using a four cell scenario which is shown in Figure 6. The NodeBs are placed with a distance of 750m around the center of the scenario. The user is moving from position one at the left side of the scenario to the right, marked with a three and back. When the user is at the center of the cell, marked with a two, she will have all four NodeB's in the active set. At this point, we expect a peak in the load.

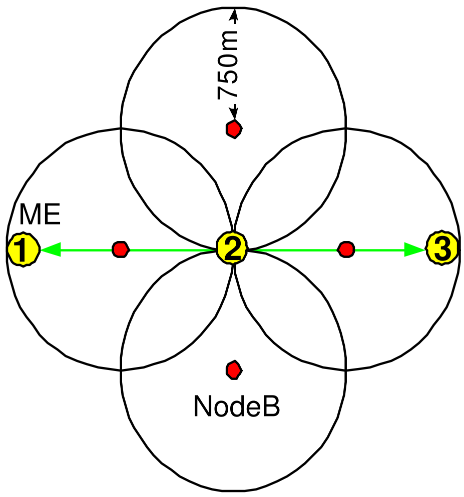


Figure 6: Mobile user scenario

This effect is shown in Figure 7 and you can see two peaks, one at around 30 seconds and the other peak at around 200 seconds. This is the time when the user is at position two. At this position, she needs a spreading code from every NodeB, illustrated with the blue curve. Furthermore, the uplink load triples due to the produced interference which reaches a maximum because of the same distance to all NodeB's. The downlink load stays almost the same and only the serving NodeB changes.

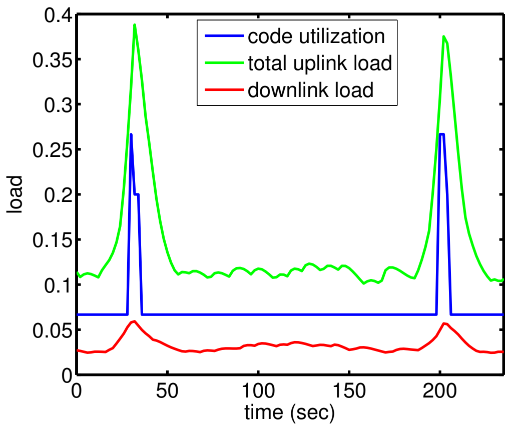


Figure 7: Mobile user

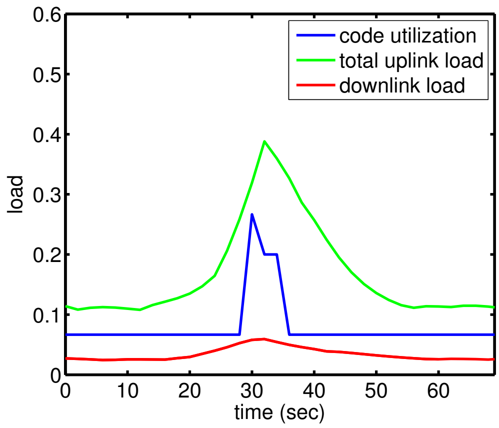


Figure 8: Mobile user in the overlapping area

A closer look into the scenario at the peak, shown in Figure 8, clarifies the effect. Looking at the blue curve for the code utilization shows that at the beginning, the mobile user only has one NodeB in the active set. At around 29 seconds all three remaining NodeB's are added to the active set and the code utilization increases to 25 percent. After another three seconds, the NodeB from the left side of the scenario is dropped from the active set which can be seen in the code utilization decrease from 25 down to 20 percent. When the user reaches the right side of the scenario, also the two NodeB's from the center of the scenario are dropped from the active set.

This mobile scenario shows that it is not sufficient to compute the load of a user when the user starts the conversation. The load has to be computed continuously.

4 Conclusion

In this report, we calculated the cell capacity in a UMTS network. The UMTS cell (or sector) capacity is limited by several factors which we used as an indicator for the load. We considered the following air interface radio resource metrics: Uplink load, which depends on the received interference at the NodeB, downlink load, which is defined by the NodeB transmit power, and the code utilization. The uplink and downlink loads are "soft" measures, which are influenced by several stochastic variables like user position, speed, and fading. In contrast, the code utilization is a hard measure, since it depends only on the number of active downlink connections.

The simulation results have shown that the code utilization has to be taken into account in order to calculate the cell capacity. In all scenarios, this was the main limiting factor. If we want to perform a policy-based vertical handover with the network load as one policy, the code utilization has to be computed.

References

[1] H. J. Wang, R. H. Katz, and J. Giese, "Policy-Enabled Handoffs Across Heterogeneous Wireless Networks," in *WMCSA*, (New Orleans, LA, USA), February 1999.

[2] K. Sipilä, Z.-C. Honkasalo, J. Laiho-Steffens, and A. Wacker, "Estimation of capacity and required transmission power of WCDMA downlink based on a downlink pole equation," in *Proc. of VTC spring '00*, vol. 2, (Tokyo, Japan), pp. 1002–1005, 2000.

[3] H. Holma and A. T. (Eds.), *WCDMA for UMTS.* John Wiley & Sons, Ltd., Feb 2001.

[4] J. Pérez-Romero, O. Sallent, R. Agusti, and M. A. Díaz-Guerra, *Radio Resource Management Strategies in UMTS.* John Wiley & Sons, Ltd, 2005.

[5] D. Staehle and A. Mäder, "An Analytic Model for Deriving the Node-B Transmit Power in Heterogeneous UMTS Networks," in *IEEE VTC Spring*, (Milano, Italy), May 2004.

[6] A. Mäder and D. Staehle, "Analytic Modelling of the WCDMA Downlink Capacity in Multi-Service Environments," in *16th ITC Specialist Seminar*, (Antwerp, Belgium), pp. 229–238, Aug 2004.

[7] 3GPP, "3GPP TS 25.215 V3.13.0 Physical layer - Measurements (FDD)," tech. rep., 3GPP, Mar 2005.

[8] 3GPP, "3GPP TS 25.305 V7.1.0 Stage 2 functional specification of User Equipment (UE) positioning in UTRAN," tech. rep., 3GPP, Sep 2005.

[9] D. Staehle, "On the code and soft capacity of the UMTS FDD downlink and the capacity increase by using a secondary scrambling code," in *IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, (Berlin, Germany), Sep 2005.

[10] International Telecommunication Union, "Coding of Speech at 8 kbit/s using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)." Recommendation G.729, Telecommunication Standardization Sector of ITU, March 1996.