Capacity Optimization for Adaptive Multi-Rate in UMTS Systems

Gerhard Helmer(*) and Andrea Garavaglia(*)
WCDMA System Design
Ericsson Eurolab Nürnberg
Nuremberg, Germany

Gunther Knopp(*) and Wolfgang Koch
Ericsson-chair for Mobile Communications
University of Erlangen-Nuremberg
91058 Erlangen, Germany
koch@nt.e-technik.uni-erlangen.de

Jürgen Bolik(*)
Cognitas GmbH
Ottobrunn, Germany

Alfred Wassermann
University of Bayreuth
Bayreuth, Germany

Abstract— **Adaptive Multi-Rate (AMR) will be used for voice applications in UMTS systems where high spectral efficiency and system stability are required while guaranteeing a good speech quality. In this paper we investigate the trade-off between the system resource utilization and the provided quality of voice services, by proposing a method that considers both aspects of quality and capacity for the optimal assignment of speech data rates in 3rd generation cellular systems. Preliminary results, obtained with a simple model, indicate that a substantial capacity increase can be achieved in comparison to non-optimal cases. Moreover, the simplicity of the approach allows a straightforward application in real-time assignment strategies.**

Keywords – UMTS; WCDMA; AMR; Rate Assignment; Radio Resource Management; Discrete Optimization;

I. INTRODUCTION

The evolution of personal wireless communication systems aims to support larger flexibility in applications and radio resource usage with the introduction of different types of packet and circuit switching connections. In this regard, the Adaptive Multi-Rate (AMR) is the code standardized by 3GPP [1-3] that will be used for speech calls in the UMTS system. UMTS is the WCDMA member of the IMT2000 family of third generation systems. The specified coding scheme is able to support up to eight different data rates, from 4.75 kb/s to 12.2 kb/s, offering a wide range of settings to match the available rates in worldwide already deployed systems and the desired Quality of Service (QoS). Each of the available rates provides a proper speech quality, by means of appropriate information redundancy, and consumes a different amount of radio resources. Generally speaking, rates are switched down to increase the system capacity and are switched up to increase the quality of the voice service. Therefore, from a network point of view, a clear trade-off between the capacity of the system and the overall quality provided to the end users can be identified for rate assignment of speech calls.

In this paper, we introduce a mathematical model and present a network optimization method, which considers together quality and capacity in order to find out optimal rate assignments for 3rd generation wireless communication systems. The assumptions made about the base station maximum power, the radio network scenario and the supported data rates are rather general, thus allowing full flexibility to adapt the method to different systems and network configurations. We apply the method to UMTS systems.

The paper first presents a simple, yet comprehensive model for the quality of voice applications and the corresponding resource usage. Based on that, the problem of optimal rate assignment is introduced together with the proposed solution approach. System capacity performance is then evaluated by means of numerical simulations, and finally, conclusions are drawn.

II. QUALITY AND CAPACITY IN UMTS SYSTEMS

In order to define the optimization problem, we first identify suitable quantities to summarize the quality and capacity aspects of UMTS systems.

In case of speech calls, the quality of service depends on the target Block Error Ratio (BLER) and on the data rate for the standardized AMR mode. These perceived speech quality is anyway not directly proportional to the particular data rate, being the AMR modes a collection of eight different speech codecs rather than a single codec with eight different data rates. Therefore an ad hoc metric shall be defined to characterize the quality of the different AMR modes.

For realistic rate assignment algorithms, a suitable approach for the quality of service measure is directly taken from [4], by introducing a quantity called User Satisfaction, which is measured experimentally in terms of averaged voice-call perception. Good perceived quality will lead to a high user satisfaction, while a poor perceived quality or a blocked or dropped will lead to a low or negative user satisfaction. By indicating with $us_i(t)$ the instantaneous user satisfaction of user i at time t, the user satisfaction of a entire speech call of duration t_i can be defined as

$$\overline{us_i} = \frac{1}{t_i}\int_{t=0}^{t_i} us_i(t)\,dt \qquad (1)$$

Reasonably assuming that a user is annoyed for a certain amount of time after experiencing an event of either blocked or dropped call, a proper duration can also be assigned to those events. Therefore, the call duration t_i can be divided in H time intervals of duration $t_{i,h}$, characterized by constant AMR mode and BLER, and the integral in equation (1) can be substituted by a sum over the H time intervals:

$$\overline{us_i} = \frac{1}{t_i}\sum_{h=1}^{H} us_{i,h}\cdot t_{i,h} \qquad (2)$$

where $us_{i,h}$ is the user satisfaction of user i in the time interval h.

Expanding the sum and collecting together the terms concerning the generic AMR mode k, the user satisfaction for the user i switching rate among m modes can be rewritten as

$$\overline{us_i} = \frac{1}{t_i}\sum_{k=0}^{m} u_{i,k,BLER}\cdot t_{i,k} \qquad (3)$$

where $u_{i,k,\mathrm{BLER}} = f(k,\mathrm{BLER})$ is the user satisfaction perceived by the user i when using the mode k, for a certain BLER. Assuming that an ideal power control is deployed, the BLER can be considered constant during the call (i.e. the connection quality is locked to the set target) and thus removed from the function above. Therefore, for the AMR mode k at given BLER, the user satisfaction can be written as u_k. The values for the u_k for the different AMR modes at BLER of 1% are listed in Table I [4]. Note that besides the normal eight AMR modes, the two extra modes $k = 0$ and $k = -1$, respectively for blocked and dropped calls (thus having a negative user satisfaction) are also included.

According to the definition above, two new significant parameters, namely the average user satisfaction US and the system performance SP can be introduced:

$$US = \frac{\sum_{i=1}^{N_u}\overline{us_i}\cdot t_i}{\sum_{i=1}^{N_u} t_i} \qquad (4)$$

$$SP = \frac{US\cdot\sum_{i=1}^{N_u} t_i}{T\cdot N_c} = \frac{\sum_{i=1}^{N_u}\overline{us_i}\cdot t_i}{T\cdot N_c} \qquad (5)$$

TABLE I. USER SATISFACTION FOR AMR MODES AT 1% BLER [4]

AMR mode k:	**0** (blocked)	**1**	**2**	**3**	**4**
Bit rate r_k [kbit/s]	0.00	4.75	5.15	5.90	6.70
User satisfaction u_k	-5.00	0.446	0.649	0.710	0.758
AMR mode k:	**-1** (dropped)	**5**	**6**	**7**	**8**
Bit rate r_k [kbit/s]	0.00	7.40	7.95	10.20	12.20
User satisfaction u_k	-1.00	0.746	0.855	0.937	0.892

They represent the quality of a particular service over a group of N_u users and the observed quality during time T of a complete system consisting of N_c cells, respectively.

Another aspect of interest for the optimization of rate assignment is the system capacity. Considering the case of a UMTS FDD (Frequency Division Duplex) system, it is possible to separate the capacity constraints for uplink and downlink, thus simplifying the analysis. In uplink, the capacity is limited mainly by interference from other mobiles and from the base stations, due to the non-orthogonality of the spreading codes used in the uplink. In downlink instead, the capacity is limited by the total available power at the base stations and by the hierarchical structure of a limited number of spreading codes.

In highly loaded WCDMA networks, the downlink power becomes the limiting factor of radio resources for voice services. Therefore, the capacity of a cell can be assumed limited by the maximum downlink power P_0 available for speech connections, which takes into account the maximum base station carrier power, the power dedicated to pilot and synchronization channels, and a possible margin to cope with measurement errors. Based on this fact, the other factors of code limitations or uplink resource budget will not be considered further in this work for sake of simplicity. In case needed, they may still be added as additional constraints to the optimization problem formulated in the next section.

III. THE OPTIMIZATION PROBLEM

The problem we want to solve consists in maximizing the overall user satisfaction (i.e. the total satisfaction accumulated over all users), under the constraint that the sum of all individual downlink transmit powers p_{ik} in a cell shall not exceed the total downlink power P_0 available for speech services. The solution will be represented by the rate assignment for each of the N_u users present in the system. The quantity that shall be maximized is an indication of the quality of service for the whole system of users and benefits from a high amount of users, which sums up to overall user satisfaction. On the other hand, the values of individual user satisfactions are limited by the cumulated user satisfaction of all users, and therefore an optimum for the trade-off between quality and capacity shall be found.

Considering a system of N_u users switching among m AMR modes, we introduce the decision variable $x_{ik} \in \{0,1\}$ to represent the solution of the optimization problem above:

$$x_{ik} = \begin{cases} 1 & \text{if the user } i \text{ selects mode } k \\ 0 & \text{otherwise} \end{cases}$$

where $i \in \{1,\ldots, N_u\}$ and $k \in \{1,\ldots, m\}$. The problem of optimizing the total user satisfaction in the system, while satisfying the constraint on limited capacity, can be then expressed as

$$\textbf{Maximize} \quad \sum_{i=1}^{N_u}\sum_{k=1}^{m} u_k\cdot x_{ik} \quad \text{such that}$$

$$\sum_{i=1}^{N_u}\sum_{k=1}^{m} p_{ik}\cdot x_{ik} \le P_0 \qquad (6)$$

$$\sum_{k=1}^{m} x_{ik} = 1 \qquad (7)$$

$$\sum_{k=1}^{m} r_{ik}\cdot x_{ik} = r_i \qquad (8)$$

$$x_{ik} \in \{0,1\} \qquad (9)$$

The inequality (6) imposes the downlink power constraint; while equations (7) and (8) ensure that each user in the system selects a mode and transmits with the corresponding data rate, respectively. Finally, equation (9) imposes the integrality of the solution.

To solve the problem, the individual consumed powers p_{ik} shall be derived from the system parameters and the AMR mode characteristics. The details of this calculation are reported in Appendix A. It is important to notice that the values of the individual powers depend on the selected solution, i.e. $p_{ik} = p_{ik}(x_{ik})$, thus introducing a non-linearity in the optimization problem.

From a computational complexity point of view, the AMR capacity and quality optimization problem is a very difficult problem (NP-hard problem), and even simple cases demand a huge amount of time to search for the exact solution. The difficulty of solving this problem arises from the fact that the value of p_{ik} is related to the level of background noise N and total interference I at the mobile receiver to the bit rate r_k, and to the path loss L_i of a user i at a given time t. Since the level of interference that the users experience depends on the power transmitted from the base stations, which in turn depends on the various p_{ik} belonging to the chosen solution, the discrete optimization problem increases exponentially in size, and ready available common solvers cannot be directly used for a real case. Therefore, we decide to investigate the possibility of modifying the optimization problem, by reducing it in dimension.

One promising approach is given by a simplification with the help of a user grouping method, where users are classified and grouped according to their power consumption. This is based on the idea of characterizing each user as belonging to a certain group, within which every user has the same common properties in terms of path loss or of transmitted power. The grouping method provides also a good possibility to base a real time rate assignment algorithm on, since users are stepwise classified and rate changes are limited, being only needed when a user changes group in the system to improve the overall quality.

In [5] different grouping techniques have been tested and compared in terms of fairness of the user distribution, calculation time, and quality of the final solution (e.g. how close is the solution found with grouping to the bounds of the solution without grouping). The use of a grouping according to power consumption with a logarithmic distribution has been selected as a good alternative to obtain an appropriate equivalent reduced-size problem.

Considering a number N_C of user groups and denoting with n_i the number of links in group i, the optimization problem can be rewritten as:

$$\textbf{Maximize} \quad \sum_{i=1}^{N_C}\sum_{k=1}^{m} u_k\cdot x_{ik} \quad \text{such that}$$

$$\sum_{i=1}^{N_C}\sum_{k=1}^{m} p_{ik}\cdot x_{ik} \le P_0 \qquad (10)$$

$$\sum_{k=1}^{m} x_{ik} = n_i \qquad (11)$$

$$\sum_{k=1}^{m} r_{ik}\cdot x_{ik} = r_i \qquad (12)$$

$$x_{ik} \in \{0,\ldots,n_i\} \qquad (13)$$

$$\sum_{i=1}^{N_C} n_i = N_u \qquad (14)$$

where the equations (10)-(13) replace the equations (6)-(9) and the additional equation (14) ensures that all the users are taken into account exactly once.

IV. NUMERICAL RESULTS

To solve the described optimization problem, a numerical approach has been used, based on snapshot simulations. The scenario we considered assumes uniform cell deployment with omni-directional antennas at the Nodes B, and speech-only services with uniform traffic distribution. Three different environments have been modeled using the Okamura-Hata attenuation model, according to the characteristics reported in Table II.

For the AMR modes, the values reported in Table I are used. Because of snapshots simulations, no distinction is possible between blocking and dropping events. Therefore we decided to use only the user satisfaction u_0 = -1 for both.

After choosing the desired scenario, the corresponding cumulative distribution of the path loss is evaluated and values are statistically assigned to the users, for different level of traffic. To complete the problem initialization, we assume all the users having the best rate and apply the equation (A3) in the appendix to initialize the power coefficients. To avoid unrealistic power assignments, a link power limitation of 2W is introduced, and then users are grouped according to their power. For N_C the values $N_C = 10$ and $N_C = 20$ have been selected for this purpose. Finally, several examples have then been solved with *LPSolve* [6].

TABLE II. CHARACTERISTICS OF THE SIMULATED ENVIRONMENTS

Environment	Urban	Suburban	Rural
Attenuation Constant	32.5 dB	25.32 dB	4.14 dB
Attenuation exponent	3.52	3.52	3.44
Shadowing (σ)	8 dB	8 dB	8 dB
Correlation Distance	40 m	30 m	100 m
Radio Channel	Pedestrian A	Vehicular A	Vehicular A
Cell Radius	500 m	1000 m	3400 m
Mobile Distribution	Uniform	Uniform	Uniform
Average mobile speed	3 km/h	50 km/h	80 km/h
Power Control	Ideal	Ideal	Ideal
Considered Bit Rate	r_k(AMR) + 3.4 kbit/s	r_k(AMR) + 3.4 kbit/s	r_k(AMR) + 3.4 kbit/s
Handover	No SHO	No SHO	No SHO
DL Available Power P_0	14 W	14 W	14 W

As comparison for the optimized solution, we have selected for this paper the simple reference case of assigning a single rate for all users, which can be actually regarded as a good example of non-optimal rate assignment. In figure 1 the results obtained for the urban scenario are represented in terms of overall user satisfaction *US* and System Performance *SP*. The optimization solution outperforms significantly the reference case for the complete range of traffic. Also for the cases of Suburban (figure 2) and Rural (figure 3) scenarios, the optimal solution behaves better than the single rate, though the improvements are not as consistent as in Urban scenario. This is due to the fact that the use of AMR is getting more advantageous when more diversity is available in the system, by adapting the different rates to the different power levels of the mobile terminals.

V. CONCLUSIONS

To conclude, we have presented a mathematical model and a simplified approach to optimize rate assignment of AMR modes in a cellular system. We have applied this method to AMR rate assignment in the downlink of UMTS.

The method allows the evaluation of an approximated solution that provides a significant increase in capacity in comparison with non-optimal cases. The fast calculation time and the reduced complexity approach given by the grouping technique enable its use in real time applications for AMR rate assignment strategies.

The results have shown that the difference between the optimal solution when the complete set of AMR modes is considered, and the case when only a single mode is used throughout the whole network, is significant. This confirms the expectation of a quality increase for a given capacity (or vice versa, a capacity increase for a desired quality).

The solution of the optimization problem provides upper bounds for the cumulated user satisfaction, under power consumption constraints, and can be used as benchmarks in order to evaluate present ad-hoc algorithms.

APPENDIX A

For the generic user *i*, an effective individual downlink power pi is transmitted from the network to the mobile phone, depending on the selected AMR mode, on the target BLER, and on channel conditions and cell interference. Under the assumption of ideal power control, the target error rate is always reached, and the Signal to Noise and Interference Ratio (SINR) of a generic user keeps constant for the selected bit rate. For higher data rates, the SINR has to increase proportionally to the effective bit rate in order to keep the error rate on target. Therefore, following the analysis reported in [5], the following relation can be written for the user *i* in linear power domain:

$$SINR_i = \frac{p_i}{L_i \cdot (N + I_i)} = \beta \cdot (r_i + r_s) \tag{A1}$$

where L_i is the path loss, N is the receiver noise, I_i is the total interference, r_i and r_s are the user and signaling data rates, respectively, and β a coefficient relating the SINR to the bit rate.

The total interference is the sum of the intracell interference (i.e. the interference generated by the cell to which the user is connected) and of the intercell interference (i.e. the interference generated by all the other cells). The intracell interference is caused by non-orthogonal downlink power contributions of the base station:

$$I_{\text{intra}} = \frac{1}{L_i}\left[\alpha_i \left(P_{CPICH} + P_{DPCH} - p_i\right) + P_{SCH}\right] \tag{A2}$$

where P_{CPICH} is the power of the Common Pilot Channel, P_{DPCH} is the total power of Dedicated Physical Channels, P_{SCH} is the power of the Synchronization Channel and α_i is the orthogonality factor, accounting the lost of orthogonality due to multipath propagation. The intercell interference is caused by the downlink transmitted power of neighboring base stations and can be regarded as a sum of many contributions coming from all directions. In the considered scenario with symmetrical deployment and uniform traffic distribution, the intercell interference can be approximated by a constant term I_{inter} without loss of generality.

By substituting the total interference $I_i = I_{\text{intra}} + I_{\text{inter}}$ in equation (A1) and solving, with the help of the relation (A2), for p_i we obtain:

$$p_i = \gamma_i \cdot \left[L_i\left(N + I_{inter}\right) + \alpha_i\left(P_{CPICH} + P_{DPCH}\right) + P_{SCH}\right] \tag{A3}$$

where

$$\gamma_i = \frac{\beta \cdot (r_i + r_s)}{1 + \alpha_i \beta \cdot (r_i + r_s)} \tag{A4}$$

In equation (A3), the term P_{DPCH} has still to be evaluated as a function of the other system parameters, before using it directly in the optimization problem. For a cell serving N_u users, the total power used for the dedicated physical channels is

$$P_{DPCH} = \sum_{i=1}^{N_u} p_i \tag{A5}$$

By substituting equation (A3) in (A5) and solving for P_{DPCH} we obtain

$$P_{DPCH} = \frac{\left(N + I_{inter}\right) \cdot \sum_{i=1}^{N_u} \gamma_i L_i + P_{SCH} \sum_{i=1}^{N_u} \gamma_i + P_{CPICH} \sum_{i=1}^{N_u} \alpha_i \gamma_i}{1 - \sum_{i=1}^{N_u} \alpha_i \gamma_i} . \tag{A6}$$

Its combination with (A3) and (A4) gives the needed solution. Note that the parameter α_i depends on the selected rate for the user *i*, which in turn depends on the solution according to equation (8), thus yielding to non-linearity.

REFERENCES

[1] 3GPP, "AMR Speech Codec; General Description", TS 26.071, (www.3gpp.org)

[2] 3GPP, "AMR Speech Codes; Transcoding Function", TS 26.090, (www.3gpp.org)

[3] 3GPP, "Performance Characterization of the AMR Speech Codec", TR 26.975, (www.3gpp.org)

[4] M. Karlsson et al., "Joint Capacity and Quality Evaluation for AMR Telephony Speech in WCDMA Systems", in Proc. of IEEE VTC 2002 Fall

[5] G. Knopp: "Analysis and Optimisation of Rate Assignment Algorithms for Adaptive Multi-Rate in 3G-Networks", Master Thesis, University of Erlangen-Nürnberg, Nov. 2002

[6] See for example ftp://ftp.ics.ele.tue.nl/pub/lp_solve/

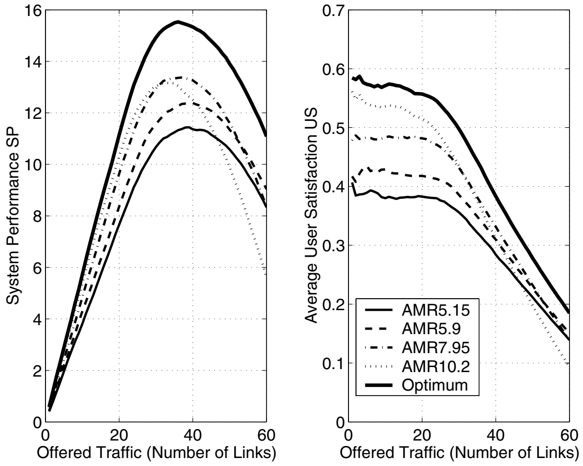

Figure 1. SIMULATION RESULTS FOR URBAN SCENARIO

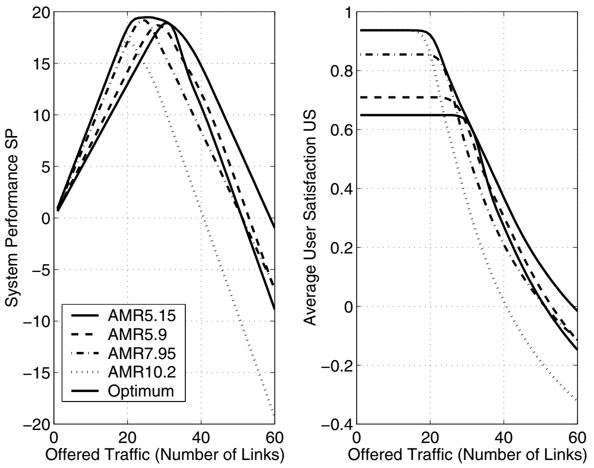

Figure 2. SIMULATION RESULTS FOR SUBURBAN SCENARIO

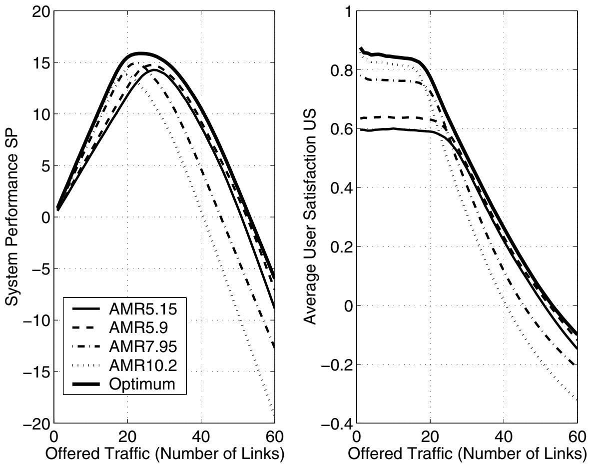

Figure 3. SIMULATION RESULTS FOR RURAL SCENARIO

(*) Gerhard Helmer, Jürgen Bolik and Andrea Garavaglia were with Ericsson until June 2003. Gunther Knopp was with University Erlangen-Nuremberg until April 2003.