Towards a 4G IP-based Wireless System Proposal

Tony Ottosson[1], Mikael Sternad[2], Anders Ahlén[2] Arne Svensson[1] and Anna Brunström[3]

[1] Dept. Signals and Systems, Chalmers University of Technology, SE-41296 Göteborg, Sweden
[2]Signals and Systems, Uppsala University, PO Box 528,SE-751 20 Uppsala, Sweden.
[3]Dept. Computer Science, Karlstad University, SE-65188 Karlstad, Sweden

Abstract

A packet switched wireless cellular system with wide area coverage, high throughput and high spectral efficiency is proposed. Smart antennas at both base stations and mobiles improve the antenna gain and improve the signal to interference ratio. The small-scale fading is predicted and a slotted OFDM radio interface is used, in which time-frequency bins are allocated adaptively to different mobile users, based on their predicted channel quality. This enables efficient scheduling among sectors and users, as well as fast adaptive modulation and power control. We here estimate the spectral efficiency of the suggested downlink. The resulting channel capacity grows with the number of simultaneous users and with the number of antenna elements in terminals. A high efficiency, around 4 bits/s/Hz at 16dB SNR is attained already for a moderate numbers of users and terminal antennas. An outline is given of research pursued within the PCC Wireless IP Project to improve and investigate this type of system.

1 Introduction

With the current trend towards integration of systems and services, wireless data services for mobile vehicular users will require throughputs and Quality of Service (QoS) levels that approach those obtainable in present WLAN's. For economic reasons, this challenge cannot be met solely by an expansion of the bandwidth. A much higher spectral efficiency will be a key feature of any acceptable future 4G solution.

The Wireless IP project [1] within the Swedish research program for Personal Computing and Communications (PCC) studies problems that are important in the evolution of UMTS toward higher data rates, as well as in future 4G technologies for mobile systems. Our goal is to improve the spectral efficiency for packet data, in particular IP traffic, with sufficient quality of service for various traffic classes.

We here outline a radio interface for a spectrally efficient and highly flexible packet switched system that provides high throughput in both the downlink and the uplink. This concept aims at a much higher spectral efficiency than obtained by 3G systems and their enhancements, such as the HSDPA mode of WCDMA, release 5 [2]. Our ultimate goal is a capacity of $50 - 100$ Mbit/s per sector or beam of a base station, with a somewhat lower capacity in the uplink. This would provide wireless and mobile users with a capacity that corresponds to the present wireline Ethernet.

Our work so far, summarized here, has resulted in methods that can be expected to provide a downlink spectral efficiency of around 4 bits/s/Hz. This corresponds to 20Mbit/s per base station sector/beam for 5MHz channels.

What are the difficulties involved in obtaining a significantly improved spectral efficiency? The transmission of packet data to/from mobile vehicular users encounters several complications:

- A significant fraction of the users will be on the move, and encounter channel properties that vary on all time-scales.
- Uplink transmissions will not be precisely synchronized[1].
- A constrained transmit power in particular in the uplink will limit the transmission range and/or the data rate.

To overcome these difficulties, we develop and combine tools and methods that are outlined below in Section 2. The properties of the proposed downlink are specified in Section 3 and its capacity is analyzed in Section 4. Some open issues under present study are discussed in Section 5. For instance, our effort to improve the efficiency of lower layers are complemented by efforts to improve the transport layer for wireless links.

2 Elements of a 4G Radio Interface

To obtain high data rates in wireless packet data transmission is, in itself, not particularly difficult, but to do this with wide area coverage, without a very dense and expensive infrastructure, is more challenging. To provide high data rate services also to vehicular users with reasonable quality of service is more challenging still, due to their rapidly varying channels. To attain a high spectral efficiency for vehicular users is the toughest challenge of all.

[1]Exact synchronization among terminals and base stations is a very difficult task, that would impose tremendous requirements on the hardware and on the algorithms of the system.

Many types of solutions could be considered, but very few are able to meet all these requirements simultaneously. Unsynchronized packet data systems, perhaps with a collision avoidance scheme, would encounter problems in heavily loaded situations. The same is probably true for ad-hoc radio networks in unregulated frequency bands, at least when wide area coverage, heavy usage and adequate quality of service are to be attained simultaneously.

We have therefore focused on slotted cellular systems in (new) regulated frequency bands, systems that partly build on the 3G philosophy, but that in important aspects depart radically from it.

UMTS and earlier cellular systems are based on averaging, which by necessity limits its spectral efficiency. We instead propose an approach that utilizes feedback control and fast optimization of the resources at hand. Coding overhead is to be minimized. Significant steps in this direction are today taken within 3GPP with the aim of improving the downlink data rate of UMTS [2], and one of our intents is to investigate how far this road might lead.

One of the many difficulties is to avoid ending up with a system with an overwhelming demand for feedback control bandwidth. There is also the risk of developing advanced estimation algorithms that estimate parameters that are in the end of little relevance for the system performance. Such aspects force us to consider the system design as a whole, rather than investigating different algorithms in isolation.

The main purpose of this paper is to outline our provisional target system proposal, so that it can be submitted to discussion, criticism and improvement.

We propose a cellular slotted FDD system and the use of multiple access via OFDM in both uplink and downlink. Uplink transmissions from different mobile users are assumed to be only roughly synchronized and slotted FDD avoids the interference problems encountered in asynchronous TDD systems. Bursty but delay-insensitive data, and streaming audio/video, are expected to comprise major parts of the traffic. Delay-sensitive traffic will also be present. The use of slotted OFDM enables a more fine-grained resource optimization, in both time and frequency, as compared to the use of either TDMA or CDMA. Within such systems, we are studying the application and co-optimization of several methods, introduced below, for improving the spectral efficiency.

2.1 Channel Quality Feedback, Scheduling and Link Adaptation

The variation of the received power with time and frequency is illustrated by Figure 1 for one particular user and fading pattern. Channels to and from different users will fade independently. With many potential users present, the channel could ideally almost always be allocated on a time-frequency slot basis to users who encounter favorable conditions within these particular slots. This effect is sometimes denoted multiuser diversity [3]. The exploitation of this effect requires *channel predictors* (in the mobiles for the downlink and at the base station for the uplink), that predict the short-term fading for each user. It also requires the feedback of the resulting slot quality estimates to a *scheduler*, which then allocates time-frequency slots among the users of each base station sector. The allocation is based on SNR estimates for each user, taking goals and constraints such as throughput, user satisfaction, delay and QoS into account [4, 5].

In the downlink, the time-frequency bins to be allocated are proposed to have a duration of 0.667ms and a width of 200kHz. Their duration corresponds to the slot length of UMTS, but a bandwidth of 5 MHz will be partitioned into 25 bins. The channel will be approximately time and frequency invariant within these time-frequency bins.

The uplink allocation will be complicated by the need for significant guard-bands, due to lack of exact clock synchronization between mobile transmitters. We would here share each 0.667ms time-slot among fewer users, possibly by the use of adaptive bandwidth allocation, and use a few narrow-band channels for short packets [6].

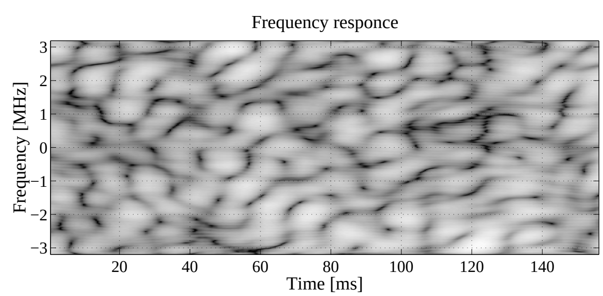

Figure 1: Time-frequency representation of an estimated channel obtained from real measurement data on a 6.4MHz channel. White color denotes high power whereas dark color denotes low power. The dynamic range and the speed of the mobile is approximately 40 dB and 50 km/h, respectively. The coherence bandwidth is 0.6 MHz.

The channel quality feedback enables *fast link adaptation* (adaptive modulation, adaptive coding and power control). Robustness against the use of too high modulation formats due to channel estimation errors can be obtained, at little coding overhead. One tool is Hybrid type-II ARQ in link level retransmission of delay-insensitive traffic [4]. The SNR limits for using different modulation formats can furthermore be optimized with the prediction accuracy taken into account [7]. The system efficiency might be improved further by using power control by waterfilling [8] among the users in different frequency bins within a time slot, but this remains to be investigated.

2.2 Prediction and Predictive Scheduling

The attainable throughput, and in particular also the fulfillment of QoS constraints, can be improved if the scheduling can be performed over a time horizon comprising several slots [5]. Predictors and predictive schedulers can be implemented at reasonable computational complexity [5].

From extensive investigations on broadband measurement data, we have found that the received signal power can be predicted with reasonable accuracy for time intervals corresponding to a movement of up to half a carrier wavelength [9, 10]. Half a wavelength at 1900MHz would at vehicular speeds of about 50 km/h correspond to 5ms, or about 7 of our proposed 0.667ms slots.

2.3 Slot Scheduling among Sectors

Some pairs or groups of sectors, within a cell or in neighboring cells, may create a large amount of mutual interference. In order to reduce this co-channel interference, time-frequency bins are allocated exclusively within these groups of sectors, according to the traffic load in each sector. Information on the traffic load is exchanged between base stations via an inquiry procedure.

This interference reduction could enable the use of aggressive power control schemes in combination with the adaptive modulation, if this is deemed advantageous. It also improves the channel estimation accuracy and reduces the problem of predicting time-varying interference levels.

2.4 Use of Multiple Antenna Elements

To improve the received power, reduce interference and attain sufficient range of the broadband transmission, we propose smart antennas to be used at the base station and preferably also in the terminals. In this first outline, we assume only straightforward robust techniques of low computational complexity, but the use of more sophisticated schemes for array signal processing and spatial multiplexing will be investigated as the project continues.

In the downlink, base stations are assumed to use multiple antennas to form S sectors (beams), that are more narrow than standard 120^o sectors. The terminals may use L antennas/diversity branches for Maximum Ratio Combining (MRC).

In the uplink, base station receivers are furthermore assumed to use MRC of several diversity branches within each sector. The mobiles may there use a few antenna elements to form crude lobes.

3 The Downlink

We assume each time-frequency bin of 0.667ms $\times$200kHz to be partitioned into 6 symbols on 20 subcarriers. Each symbol is 100μs long, with a cyclical prefix of 11μs. This is adequate to cover the delay spread in most urban environments. Of the 120 symbols per bin, five are known pilots that are used for training.[2]

A 5 MHz system will thus use 500 subcarriers (or in practice 512, with 12 idle) and the symbol rate on each subcarrier is about 9000 symbols/s.

The design target vehicular speed is 100 km/h. The system should work efficiently for populations of users with this and lower speeds.

Each mobile user predicts the quality of all bins an appropriate number of time-slots ahead. All mobiles then signal their predicted quality estimates on an uplink control channel, by transmitting the suggested appropriate modulation levels to be used in the different frequency bins of the predicted time slot. A scheduler located close to the base station then allocates time-frequency bins to the different users, based on the predicted channel capacities of all active users, their data queue status, QoS requirements and priorities. The scheduler broadcasts its allocation decisions on a downlink control channel. The downlink transmission for the allocated time-slots can then begin. In each bin, the modulation level used is that which was suggested by the appointed user.

4 Potential of Proposed Downlink

We here estimate the spectral efficiency for transmission to K active users within one base station sector, under the following simplifying assumptions:

- Flat fading AWGN channels that are time-invariant within each bin of 200kHz $\times 0.667$ms and independent Rayleigh fading between bins,
- all users are assigned equal average channel power by a slow power control scheme and the channels to different users fade independently,
- accurate channel predictions,
- the target service is reliable packet data transmission, while fairness between users with different rates, QoS requirements, and delay constraints are neglected,
- all users do always have data to transmit, and the payload symbols within the allocated slots are fully utilized by the designated user.

The assumed scheduler thus works as a selection diversity scheme, where the user with the best predicted SNR in each time-frequency bin out of all K users will transmit in that bin. In the receiver we assume MRC with L antennas.

The resulting pdf of the received SNR (γ) after MRC and multiuser selection diversity can then be shown to be

[2] Algorithms for channel estimation that are located in the mobile terminals will use also the received payload symbols in their channel estimators, to obtain channel estimates with high time-frequency resolution which are then used as input data to the channel predictors.

(see e.g. pp. 186–194 in [11] for guidance in performing this calculation)

$$p(\gamma) = \frac{K e^{-\gamma/\bar{\gamma}} (\frac{\gamma}{\bar{\gamma}})^L}{\gamma \Gamma^K(L)} \left(\Gamma(L) - \Gamma(L, \tfrac{\gamma}{\bar{\gamma}}) \right)^{K-1}, \quad (1)$$

where $\bar{\gamma}$ is the average SNR per receiver antenna, and

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt \quad (2)$$

$$\Gamma(a, z) = \int_z^\infty t^{a-1} e^{-t}\, dt \quad (3)$$

are the gamma and incomplete gamma functions, respectively.

We utilize an adaptive modulation system that uses $N = 8$ modulation formats: BPSK, 4-QAM, 8-QAM, 16-QAM, 32-QAM 64-QAM, 128-QAM, and 256-QAM. Equal transmitter power (equal energy per symbol) is used in all bins. The target symbol error rate that determines the choice of modulation level must now be determined.

Since the target service is reliable packet transmission we need an ARQ scheme. For simplicity we assume simple ARQ without coding. For this ARQ to be useful, we select a target packet error rate of $P_E = 0.1$.

A packet error rate of 0.1 for the link level packets containing 115 payload symbols corresponds to a symbol error rate of 9.16×10^{-4}.

The raw spectral efficiency must for our proposed system be multiplied by a factor $1 - P_E = 0.9$ due to packet retransmissions, by a factor $G_c = 100/111$ due to cyclical prefixes and a factor $G_p = 115/120$ since five of the 120 symbols in a bin are used as pilots. The spectral efficiency using AMS and simple ARQ will then be obtained by a weighted sum over the modulation levels, weighted by the probabilities that those particular levels will be utilized

$$\eta \geq (1 - P_E) G_c G_p \sum_{i=0}^{N-1} k_i \int_{\gamma_i}^{\gamma_{i+1}} p(\gamma)\, d\gamma \ \text{[bits/s/Hz]} \quad (4)$$

Here, $k_i = \log_2(M)$ is the number of bits transmitted using uncoded M-ary QAM modulation. The values γ_i, $i = 0, \ldots, 7$, define the symbol SNR regions in which to use the different modulation levels: 2^{k_i}-ary QAM is used in the SNR interval $[\gamma_i, \gamma_{i+1})$, with $\gamma_N = \infty$. No transmission is performed below SNR $= \gamma_0$. These γ_i values are given in Table 1 for the target symbol error rate 9.16×10^{-4}. The inequality in (4) comes from the fact that once the SNR is above the γ_i level, the symbol error rate (and thus the packet error rate) is lower than the target.

The spectral efficiency in (4) is evaluated numerically for a symbol energy to noise ratio of $\bar{\gamma} = 16$ dB per receiver antenna. The result is presented in Figure 2, which shows a dramatic improvement with an increasing multiuser selection diversity and with the number of receiving antennas. Already for two antennas and five users per sector, the spectral efficiency is considerable.

i	Modulation	k_i	γ_i (dB)
0	BPSK	1	6.86
1	4-QAM	2	10.41
2	8-QAM	3	14.57
3	16-QAM	4	17.69
4	32-QAM	5	21.03
5	64-QAM	6	24.02
6	128-QAM	7	27.16
7	256-QAM	8	30.14

Table 1: The limits γ_i for the SNR per receiver antenna in dB that determine the use of different modulation levels. These limits are calculated as the SNR required to achieve the target symbol error rate 9.16×10^{-4} using $M = 2^{k_i}$ QAM. Here k_i is the number of bits per symbol.

The spectral efficiency saturates for a high number of users, where most bins are occupied by users who can utilize a high modulation degree in that bin. The addition of more receiver diversity branches tends to decrease the multiuser diversity effect.

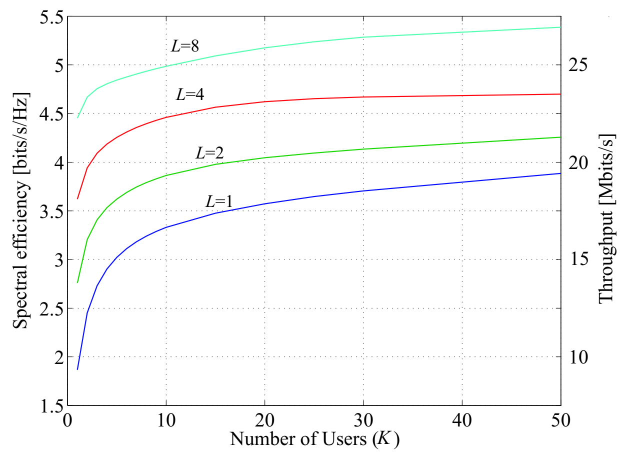

Figure 2: Spectral efficiency using AMS and simple ARQ with Lth order MRC diversity in the mobile and Kth order of selection diversity between the users. The symbol energy to noise ratio per receiver antenna is $\bar{\gamma} = 16$ dB. Right-hand scale: The corresponding capacity within one sector in a 5MHz downlink.

In theory, the spectral efficiency within one sector, η, should be multiplied by the number of sectors, S, used per cell. However, even if the interference is reduced by using many sectors, we cannot use all time-frequency bins in all sectors. The spectral efficiency of each cell would therefore be larger than is indicated by Fig. 2, but smaller than ηS. Interference between sectors has to be studied carefully for a more precise estimate.

By multiplying the spectral efficiency in bits/s/Hz by 5, it is converted into an equivalent capacity in Mbit/s of a base station sector in a system with 5MHz bandwidth. This is illustrated by the right-hand scale of Figure 2. Although the test situation is not exactly comparable, it can be noted that the maximum capacity of the HSDPA mode of WCDMA is claimed to be 8-10Mbit/s.

4.1 Discussion

The above estimate is conservative, in that the packet error rate 10% is attained only at the limits γ_i of the SNR regions. Inside the regions, the error rate is less, with a corresponding lower need for retransmissions. On the other hand, we have neglected prediction errors, which will tend to increase the error rate and decrease the attained throughput.[3] The impact of these two effects is a topic for our near-term research.

Another important aspect is the requirement for control bandwidth by the channel quality feedback and the transmission scheduling. The channel quality feedback in the uplink is in the form of a suggested modulation format. With 8 allowed formats, this requires 3 bits for each of the 25 frequency bins in a 0.667ms time slot. With K active users, the required uplink control bandwidth is $B_{uc} = K \times 112$ kbits/s. For $K = 10$ active users per sector, this corresponds to 1.1Mbit/s. This is considerable (75 times more than for power control in WCDMA) but not overwhelming; The total uplink capacity is under the assumptions of this section expected to be on the order of 15Mbit/s per base station sector [6]. Furthermore, there are many conceivable ways of modifying the signaling to reduce the bandwidth demand.

The downlink control information for broadcasting the scheduling decisions consists of a user number per slot. For 2^n simultaneous users and a code rate of R used for the control information, the required downlink control bandwidth is $B_{dc} = (n/R) \times 37.5$ kbits/s. It is small as compared to the capacity per sector.

5 Concluding Remarks

Packet data systems that use feedback information may be constructed in a way that promises to attain extreme spectral efficiency. The type of system outlined in this paper is promising and it also raises many fascinating research questions. Some of them concern the interplay between channel prediction and adaptive modulation [7], others the interplay between Hybrid ARQ retransmission and predictive scheduling [4]. The efficient use of smart antennas to improve the resource allocation is yet another topic where many questions await answers.

The interplay with higher layers is also crucial for obtaining a system with overall good performance. The flow control of TCP is here of particular interest. At present we investigate the performance of different variants of TCP over wireless links with the optimized link layers outlined above. We also investigate how the use of information from lower layers may improve the performance of higher layers. We are making a major effort to create an appropriate simulation environment for studying such issues [12]. Using real TCP and UDP stacks, it enables us to generate packets that propagate in real time over realistically simulated wireless links, so that important aspects and timescales of lower layers, and inter-layer interactions can be studied.

[3] For predictive scheduling combined with link level retransmission, our preliminary results indicate that the use of bad power prediction (standard deviation 5dB) reduces the throughput by about 20% as compared to the use of perfect prediction. The attainable Quality of Service parameters such as delays do suffer much worse degradation [4, 5].

In this conference, [7] provides an estimate of the spectral efficiency when the levels γ_i are adjusted by taking the prediction error variance into account. For flat Rayleigh fading channels, a normalized prediction error variance level of 0.1 there results in only a negligible reduction of the spectral efficiency as compared to perfect prediction. This prediction error variance would be reasonable in our proposed downlink [10].

References

[1] The Wireless IP Project within the Swedish SSF PCC (Personal Computing and Communication) program. See M.Sternad, "The Wireless IP Project," RVK02, June 2002. www.signal.uu.se/Research/PCCwirelessIP.html

[2] 3rd Generation Partnership Project (3GPP), Technical Specification Group Radio Access Network: "UTRA High Speed Downlink Packet Access (Release 4)," *3GPP TR 25.950 V2.0.0*.

[3] R. Knopp and P.A. Humblet, "Information capacity and power control in single-cell multiuser communications," *IEEE ICC*, Seattle, 1995.

[4] N. C. Ericsson, S. Falahati, A. Ahlén and A. Svensson, "Hybrid type-II ARQ/AMS supported by channel predictive scheduling in a multi-user scenario," *IEEE VTC2000-Fall*, Boston, MA, USA, Sept. 24-28, 2000, pp. 1804-1811.

[5] N. C. Ericsson, *On Scheduling and Adaptive Modulation in Wireless Communications*. Licentiate Thesis, Signals and Systems, Uppsala University, Sweden, June 2001.

[6] T. Ottosson, A. Ahlén, A. Brunström, M. Sternad and A. Svensson, "Toward 4G IP-based wireless systems: A proposal for the uplink." *5th Wireless World Research Forum Meeting*, Temple, AR, March 2002.

[7] S. Falahati, A. Svensson, T. Ekman and M. Sternad, "Effect of prediction errors on adaptive modulation systems for wireless channels," *RVK02*, Stockholm, June 2002.

[8] D.N.Tse, "Optimal power allocation over parallel Gaussian channels," *ISIT*, Ulm, Germany, June 1997.

[9] M. Sternad, T. Ekman and A. Ahlén, "Power prediction on broadband channels," *IEEE Vehicular Technology Conference*, Spring 2001, Rhodes, Greece, May 6-9 2001.

[10] T. Ekman, M. Sternad and A. Ahlén, "Unbiased power prediction on broadband channels," *RVK02*, Stockholm, June 2002.

[11] M. D. Yacoub, *Foundations of Mobile Radio Engineering*, CRC Press, 1993.

[12] A. Ewerlid and N.C. Ericsson "Real time simulations of wireless IP communications," *RVK02*, Stockholm, June 2002.