STATISTICAL MULTIPLEXING FOR MITIGATING LATENCY IN ADAPTIVE MODEMS

J.M. Torrance, L. Hanzo

Dept. of Electr. and Comp. Sc.,Univ. of Southampton, SO17 1BJ, UK.
Tel: +44-703-593 125, Fax: +44-703-594 508
Email: lh@ecs.soton.ac.uk http://www-mobile.ecs.soton.ac.uk

ABSTRACT

The latency of instantaneously adaptive modulation is mitigated using statistical multiplexing. The proposed scheme exhibits typically higher integrity than fixed modems, as seen in Figure 5 or requires an approximately 2 dB lower channel SNR, as suggested by Figure 4.

1. INTRODUCTION

Adaptive modulation has been proposed to exploit the time-variant Signal-to-noise ratio (SNR) of slowly-fading pedestrian indoors channels [1, 2], while its latency aspects have been studied in [3]. In this treatise a statistical multiplexing scheme is proposed in order to mitigate the associated delay, which arises due to disabling transmissions during the low instantaneous channel SNR intervals of slowly-fading channels. In our previous studies a lower-delay, 10^{-2} bit error rate (BER) speech system and a 10^{-4} BER, higher-integrity, higher latency data system were characterised [3]. For each schemes two sets of switching levels have been proposed that achieve the above-mentioned mean- and peak-BERs, which are shown in Table 2. Specifically, l_1, l_2, l_3, l_4 represent the levels, at which the modem switches from 'No transmission' to BPSK, QPSK, 16QAM and 64QAM, respectively. Here we only considered the performance of the 30ms-latency speech system using both the peak-and mean-switching levels, since the computer data scheme is less sensitive to system latency.

2. STATISTICAL MULTIPLEXING

In our proposed statistical multiplexing scheme the number of slots allocated to a user is adaptively controlled in order to compensate for the variation in terms of the number bits per symbol used by the transmission scheme. Table 1 summarises, how many slots will be used for each transmission scheme. As the channel conditions deteriorate, the modulation order reduces and, therefore, the number of slots required per frame, in order to maintain a certain maximum latency, increases. The channel-quality dependent demand for slots in a transmission frame is independent for all users and, therefore, with a sufficient number of users the variation in demand should be averaged. When the channel becomes so poor that transmissions will result in an unacceptable BER, that is the modem is in the 'no transmission' mode, no slots need to be occupied for transmission, however, 1 up- and 1 down-link slot is reserved for channel estimation.

Modulation Scheme	Slots Occupied
No Transmission	2
BPSK	8
QPSK	6
Square 16 QAM	4
Square 64 QAM	2

Table 1: Total number of up- and down-link slots in a frame for balanced-duplex TDD with statistical multiplexing.

Scheme	l_1	l_2	l_3	l_4
Peak-Speech	4.3	7.3	13.9	19.4
Mean-Speech	3.31	6.48	11.61	17.64
Peak-Computer Data	8.3	11.3	18.2	24.2
Mean-Computer Data	7.98	10.42	16.76	26.33

Table 2: Modulation switching levels for speech and computer data systems through a Rayleigh channel, expressed in terms of instantaneous channel SNR (dB) required for achieving Mean and Peak BERs of $1x10^{-2}$ and $1x10^{-4}$

The variation in throughput exhibits a bursty nature. Therefore, similar techniques may be used to statistically multiplex the transmission of adaptive modulation bursts to those that have been employed for speech multiplexing using Voice Activity Detectors (VAD). However, rather than a VAD determining, whether a slot will be occupied or not, the number of slots used by each mobile will vary depending upon the channel conditions. Figure 1 shows a TDD/TDMA frame, which is similar to that proposed by the Pan-European FRAMES Consortium operated under the auspices of the European ACTS Programme. The number of slots that are occupied by each user in the frame will vary. If a user's channel improves, such that a more efficient modulation scheme may be employed, the mobile will not transmit in certain slots. This will de-allocate slots that were reserved for that user. Other users may contend for slots by transmitting in time slots that are not occupied. All mobiles are aware of which slots are allocated to which user on the basis of the broadcast information contained in the control slots. A slot that would have only partially be filled with information bits is not transmitted and the bits are queued for the next frame.

Figure 2 illustrates the operation of the base-station for em-

PIMRC'97, Sept. 1997, Helsinki, Finland

0-7803-3871-5/97/$10.00 © 1997 IEEE

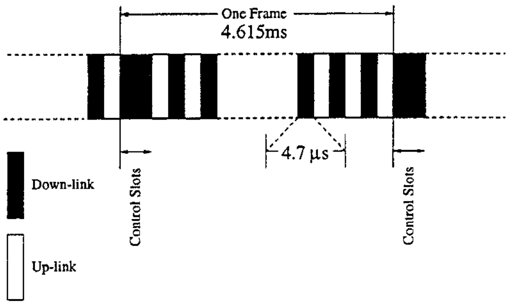


Figure 1: TDD/TDMA frame for employment of statistical multiplexing evaluation of latency in FRAMES TDD mode.

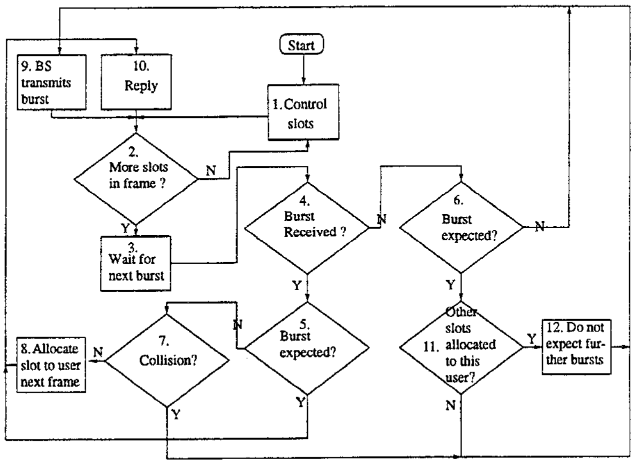


Figure 2: The base-station implementation of statistical multiplexing algorithm

ploying statistical multiplexing, which is not detailed here due to lack of space. Similarly, Figure 3 illustrates the operation of the mobile-station for using statistical multiplexing, again without considering the specific details of it.

Below it will be demonstrated that statistical multiplexing will result in reduced latency. This will allow the system throughput to be close to the mean theoretical BPS performance of the adaptive modem, without introducing unacceptable delay [3]. The statistical multiplexing scheme was evaluated for 5, 10, 15, 20, 25 and 30 users and 20, 30, 40 and 50 dB average channel SNRs. All users were assumed to be experiencing the same average channel SNR and results were obtained for normalised Doppler frequencies of 0.0042, 0.025, 0.054, 0.079 and 0.133, which correspond to mobile velocities of 0.136, 0.812, 1.75, 2.57 and 4.32 ms^{-1}. The permission probability for the statistical multiplexing assumed values of $0.1, 0.2, \ldots, 0.9$ and we found that the optimum value was 0.3. Let us now study the performance of the proposed scheme in comparison to conventional fixed modulation.

3. COMPARISON WITH FIXED MODULATION, USING BLOCK CODING

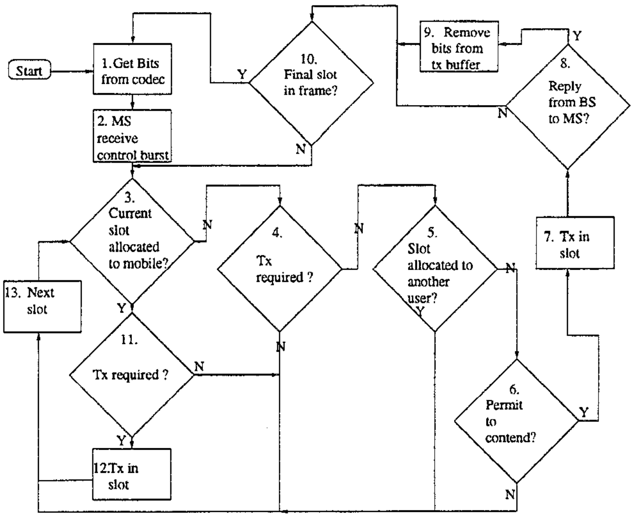


Figure 3: The mobile-station's statistical multiplexing algorithm

The adaptive and fixed modems are compared in terms of how many duplex users they can support on a single frequency at pratical indoors velocities. Block coding is employed and and indoor propagation environment is assumed. The transmission frames are 4.615 ms in duration, they contain 64 slosts easch with a payload of 64 data symbols and because of ramp, guard and training symbols, the gross symbol rate is 1.3 MBd. A continuous duplex-balanced service is assumed with a data throughput of 192 bits / 4.615 ms = 41.6 Kbits/s. The intended application for such a transmission scheme is a mixed video and speech service, for example using the 8 kbps ITU G.729 speech codec and lip-synchronised video encoded using the ITU H.263 video codec. For such a service the transmission Frames Error Rate (FER) should not exceed 1%. The latency that can be afforded by the transmission is 30 ms. This is based upon a maximum of 60 ms system delay (GSM has a speech delay of 57.5ms), where for example 10 ms are required for encoding and decoding of the source framas, and a further 10 ms are reserved for the so-called processing delay. Interference is not considered at this stage.

3.1. Fixed Scheme

The performance of BPSK, QPSK, Square 16 and 64 QAM was evaluated seperately over Rayleigh fading channels. The throuput was 41.6 Kbit/s for all schemes and, therefore they will use six, three, two and one slots per user, respectively. For Square 64 QAM, QPSK and BPSK this results in 384 bits/frame, and for Square 16 QAM 512 bits/frame. The 384-bit frame of Square 64 QAM and BPSK consists of three Bose-Chaudhuri-Hocquenghem (BCH) (127,64,10) block codes, where three transmitted bits will be 'wasted', since $3 \times 127 = 381$. The corresponding number of primary information bits is $3 \times 64 = 192$. The Square 16 QAM will use four BCH (127,50,13) block codes and waste four bits. The coding rates are close to 0.5 and half rate codes are widely used in wireless communications. All of these values are summarised in Table 3.

For each of the fixed schemes the data was interleaved over four frames. If this is achieved in the most simple manner the total latency is 8×4.615 ms. However, it is possible to interleave over four frames and maintain a transmission

Fixed Scheme	Bits Per Slot	Slots Per Frame	Bits Per Frame	Block Code	Codewords Per Frame	Info Bits Per Frame
BPSK	64	6	384	127,64,10	3	192
QPSK	128	3	384	127,64,10	3	192
QAM 16	256	2	512	127,50,13	4	200
QAM 64	384	1	384	127,64,10	3	192

Table 3: Summary of burst configurations of fixed modulation schemes to be used for comparison with adaptive modulation.

	Slow		Fast	
Mod. Scheme	Not Interl'd	Interl'd	Not Interl'd	Interl'd
BPSK	20.0	20.0	20.3	17.7
QPSK	23.0	23.0	23.1	20.7
QAM 16	27.7	27.7	28.0	23.7
QAM 64	33.9	33.9	34.2	30.5

Table 4: Average Channel SNR (dB) required to obtain 1% FER when using the fixed modulation schemes described in Table 3 under Slow (0.136 m/s speed) and Fast (4.32 m/s speed) Rayleigh fading conditions.

Current Tx. Mode	Codewords Per Slot	Information Bits Per Slot	Desired Multiplexed Slots	Info Per Frame
No Tx.	0	0	1	0
BPSK	1	39	5	195
QPSK	2	78	2.5	195
QAM 16	4	156	1.25	195
QAM 64	6	234	$0.8\dot{3}$	195

Table 5: Transmission rate for statistically multiplexed adaptive modulation scheme showing the desired number of slots to achieve a throughput comparable with the fixed scheme in Table 3

delay below 30 ms. In this case, for the Square 64 QAM, QPSK and BPSK 12 BCH codewords will be interleaved over four TDD/TDMA frames and in the case of Square 16 QAM, 16 BCH codewords will be interleaved over four TDD/TDMA frames.

Full bit-level simulation of fixed modulation with block coding was undesirable because simulation times become prohibitive for accurate results. However, full mathematical analysis is complicated and therefore the following quasi bit level simulation was performed. The fading was assumed to be constant over the length of a slot because of the propagation environment and mobile velocities concerned. This assumption holds also for the adaptive modulation analysis. The power spectral density (PSD) of the noise was assumed to be constant for the duration of the experiment. A correlated Rayleigh fading profile was invoked to yield a received power profile. The quotient of this faded received power and the noise power gave an instantaneous SNR, which, depending upon the modulation scheme, was substituted into analytical solutions, in order to obtain the BER for the slot. The BER was then used to determine the probability that the BCH code used was not overloaded and the average frame error rate (FER) was evaluated over a minimum of 10000 frames.

The FER was evaluated for the four fixed schemes shown in Table 3 over the Rayleigh fading channels of Section 2. The FERs were evaluated with and without interleaving and the SNR, where the 1% FER was achieved, was recorded. The SNRs required for the lowest and highest normalised Doppler frequencies are shown in Table 4. Under slow fading conditions the interleaving had no effect upon the required average SNR, however, for the fastest fading the required SNR was 2.6 to 4.3 dB higher.

3.2. Adaptive Scheme

In adaptive modulation a frame error can occur for two reasons, either as a result of the noise corrupting sufficient bits so that the BCH codeword is overloaded, or, alternatively the frame experiences too high a delay and is therefore not transmitted. The frames lost due to corruption yield the FER, while the frames lost due to delay result in a Delay Error Rate (DER) and the total rate of corrupted frames is the Total Error Rate (TER). For the fixed schemes the DER was zero and therefore TER equals FER. An ideal adaptive modulation scheme balances the DER and FER in order to minimise the TER. The switching levels to be used for the adaptive modulation scheme were the speech system's optimised mean BER levels and the speech system's peak BER levels from Table 2. The block code BCH(63,39,4) was chosen, since it was the most powerful code from the $n = 63$ long family that allowed us to maintain the 41.6 Kbit/s target rate stated in Section 2. Depending upon which mode the adaptive modem was in, 0, 1, 2, 4 or 6 codewords could be transmitted per frame as exemplified by Table 5.

The performance of the adaptive scheme was evaluated using the statistical multiplexing algorithm from Section 2. Here, the adaptive modulation and statistical multiplexing support 195 bits/frame for every user and therefore, irrespective of the average channel SNR, each user adds five BCH(63,39,4) codewords to its queue, delivering a 'payload' of $5 \times 39 = 195$ bits. This results in a throughput of 42.25 Kbit/s, which is in excess of the target value of 41.6 Kbit/s. Frames that were in the transmission buffer longer than 30 ms were removed from the buffer and were not transmitted. The relative frequency of this event gave the DER estimate. Perfectly coherent detection and perfect channel estimation was assumed, as it had been for the fixed modulation, and the SNR was assumed to be constant over a 72μs slot. The FER was evaluated using the same technique as was used for fixed modulation and the results were obtained for the same Slow and Fast fading rates, corresponding to pedestrian speeds of 0.136 m/s and 4.32 m/s, respectively. All users transmitting in the frame had the same throughput requirements, average channel SNR and fading statistics. Table 6 shows the DER, FER and TER for both sets of switching levels, in both the Slow and Fast Rayleigh fading channels, for various numbers of users at average channel SNRs, when all the users are generally achieving the desired performance.

4. RESULTS AND CONCLUSIONS

Several observations may be made from Table 6:

- As expected, the DER with mean BER switching levels is greater than with peak BER switching levels, since the mean BER switching levels are higher and, therefore, the throughput is on average lower. Hence more frames are dropped from the transmission buffer due to delayed transmission.
- The FER with mean BER switching levels is less than the FER with peak BER switching levels. This is the reverse situation compared to above, since the higher switching levels result in higher transmission integrity.
- The FER is the same for the equivalent Fast and Slow scenarios, however the DER reduces with increased mobile speed. The FER is constant as a function of mobile speed, because the modulation scheme that is used depends upon the instantaneous SNR only. However, the lower speeds incur higher DERs, since the reduced channel correlation prevents the transmission scheme from changing frequently.
- When the average channel SNR increases for the same number of users, the DER reduces. This is because all users will on average require fewer slots or experience fewer 'no transmission' instances.

From Table 6 it can be observed that combined adaptive modulation and statistical multiplexing can support 5, 10, 15, 20, 25 or 30 TDD users at 20, 20, 22, 26, 28 and 30 dB average channel SNRs, respectively, at either velocity or either set of switching levels. The number of users supported for the fixed modulation schemes can be calculated by taking the 32 duplex slot pairs and dividing them by the number of slots required per frame, as given in Table 3. The results are shown in Figure 4 for the Fast and Slow environments, with and without interleaving. The adaptive modem's performance is also shown. However, it should be noted that the line given is for the Slow fading channel and better performance is achieved by the Fast fading channel, hence the label 'Worst Case'.

Figure 4 shows that the adaptive scheme supports at least as many, and often 25 % more multi-media users than the BPSK, QPSK and Square 16 fixed modulation schemes over the Slow fading channel. Over the Fast fading channel, with the exception of 5 users, the adaptive modulation matchs the performance all the fixed schemes, but at lower average channel SNRs. There are two points that further improve the relative performance of adaptive modulation:

- The overall effect of a given DER and FER upon the quality of a wireless service is not the same. For fixed schemes the DER is zero and this has the advantage that since systematic block codes are used, the decoder can output the received primary information bits, rather than erroneously decoding them, when the code is overloaded. Alternatively, the overloaded BCH codeword may be dropped at the receiver. Either way, without the use of Automatic Repeat Request techniques - which exhibit an associated delay, redundancy and additional complexity - the transmitting encoder does not know the condition

		Normalised Doppler Frequency 0.0042 (velocity 0.136 ms^{-1})						Normalised Doppler Frequency 0.133 (velocity 4.32 ms^{-1})					
		Mean BER switching levels			Peak BER switching levels			Mean BER switching levels			Peak BER switching levels		
Users	SNR (dB)	FER	DER	TER	FER	DER	TER	FER	DER	TER	FER	DER	TER
5	18	3.0×10^{-3}	9.5×10^{-3}	1.3×10^{-2}	4.4×10^{-5}	1.4×10^{-2}	1.4×10^{-2}	3.0×10^{-3}	1.8×10^{-3}	4.8×10^{-3}	4.4×10^{-5}	2.9×10^{-3}	2.9×10^{-3}
5	20	2.3×10^{-3}	5.8×10^{-3}	8.1×10^{-3}	4.1×10^{-5}	7.2×10^{-3}	7.2×10^{-3}	2.3×10^{-3}	1.0×10^{-3}	3.3×10^{-3}	4.1×10^{-5}	1.4×10^{-3}	1.4×10^{-3}
10	20	2.3×10^{-3}	6.1×10^{-3}	8.4×10^{-3}	4.1×10^{-5}	7.6×10^{-3}	7.6×10^{-3}	2.3×10^{-3}	1.1×10^{-3}	3.4×10^{-3}	4.1×10^{-5}	3.4×10^{-3}	3.4×10^{-3}
10	22	1.7×10^{-3}	2.0×10^{-3}	3.7×10^{-3}	3.3×10^{-5}	3.4×10^{-3}	3.4×10^{-3}	1.7×10^{-3}	2.2×10^{-4}	1.9×10^{-3}	3.3×10^{-5}	3.7×10^{-4}	4.0×10^{-4}
15	22	1.7×10^{-3}	3.0×10^{-3}	4.7×10^{-3}	3.3×10^{-5}	5.5×10^{-3}	5.5×10^{-3}	1.7×10^{-3}	2.9×10^{-4}	2.0×10^{-3}	3.4×10^{-5}	4.1×10^{-4}	4.4×10^{-4}
15	24	1.2×10^{-3}	9.3×10^{-4}	2.1×10^{-3}	2.4×10^{-5}	1.5×10^{-3}	1.5×10^{-3}	1.2×10^{-3}	6.7×10^{-5}	1.3×10^{-3}	2.4×10^{-5}	1.0×10^{-4}	1.2×10^{-4}
20	26	8.7×10^{-4}	1.1×10^{-3}	2.0×10^{-3}	1.7×10^{-5}	1.7×10^{-3}	1.7×10^{-3}	9.0×10^{-4}	7.3×10^{-5}	9.7×10^{-4}	1.7×10^{-5}	1.0×10^{-4}	1.2×10^{-4}
20	28	4.6×10^{-4}	6.6×10^{-4}	1.1×10^{-3}	1.2×10^{-5}	1.0×10^{-3}	1.0×10^{-3}	4.8×10^{-4}	3.5×10^{-5}	5.2×10^{-4}	1.2×10^{-5}	7.4×10^{-5}	8.6×10^{-5}
25	28	4.6×10^{-4}	8.6×10^{-4}	1.3×10^{-3}	1.1×10^{-5}	8.4×10^{-3}	8.4×10^{-3}	4.7×10^{-4}	2.5×10^{-4}	7.2×10^{-4}	1.2×10^{-5}	3.7×10^{-4}	3.8×10^{-4}
25	30	3.8×10^{-4}	5.4×10^{-4}	9.2×10^{-4}	7.0×10^{-6}	6.2×10^{-3}	6.2×10^{-3}	3.8×10^{-4}	2.1×10^{-5}	4.0×10^{-4}	7.0×10^{-6}	3.6×10^{-5}	4.3×10^{-5}
30	30	4.0×10^{-4}	2.0×10^{-3}	2.4×10^{-3}	7.0×10^{-6}	1.2×10^{-2}	1.2×10^{-2}	4.1×10^{-4}	1.8×10^{-3}	2.2×10^{-3}	8.0×10^{-6}	4.4×10^{-3}	4.4×10^{-3}
30	32	2.4×10^{-4}	6.4×10^{-4}	8.8×10^{-4}	6.0×10^{-6}	3.1×10^{-3}	3.1×10^{-3}	2.5×10^{-4}	4.1×10^{-5}	2.9×10^{-4}	6.0×10^{-6}	4.8×10^{-5}	5.4×10^{-5}

Table 6: The Frame Error Rate (FER) due to noise, Dropped Error Rate (DER) and Total Error Rate (TER) for statistically multiplexed, block coded adaptive modulation scheme using the peak and mean BER switching values at two velocities

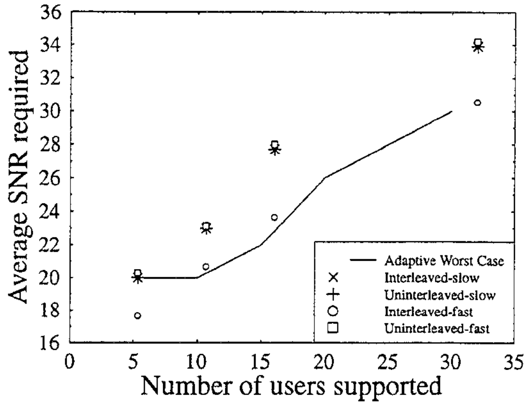
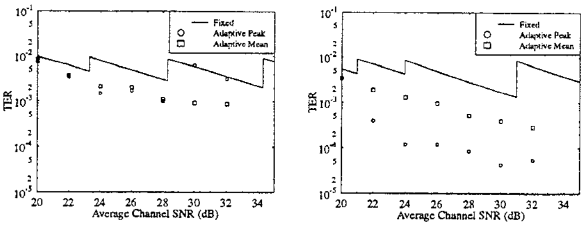


Figure 4: The average channel SNR required to support a given number of users with fixed and adaptive modulation on the basis of the comparison case study. Adaptive performance is for the 'Slow' channel and is hence labelled 'Worst Case'.



(a) Slow Channel

(b) Fast Channel

Figure 5: Total Error Rate for adaptive and interleaved fixed modulation schemes at normalised Doppler frequencies of 0.004170 (Slow) and 0.133427 (Fast). The Adaptive schemes shown with both mean and peak BER switching levels and the number of users that are supported can be inferred from Figure 4.

of the received BCH codeword. With adaptive modulation and the peak switching levels the FER can become very low, as seen Table 6, and the DER ≈ TER. Therefore, it is possible for the source encoder to have near perfect knowledge of what is received at the decoder. This is extremely useful in mitigating the propagation of errors resulting from the encoder ignorantly assuming that its local decoder's state is the same as the remote decoder's state.

- The results for the fixed modems were evaluated using BPSK, QPSK, square 16 and 64 QAM. They are based around a particular operating average channel SNR. In the fixed modems the number of slots allocated to a user is not reassigned, as the average channel SNR varies. Without this reassignment every fixed modulation user will potentially require a margin of SNR to protect it against degrading average channel SNRs. Furthermore, as the average channel SNR increases, the lack of ability to reassign slots results in fewer users being supported than possible. Slot reassignment is possible for the fixed modulation schemes, although it incurs an additional transceiver reconfiguration capability beyond the scope of this work.

Figure 4 shows that adaptive modulation increases the number of users that can be supported at a given average channel SNR over a Rayleigh fading channel, when compared to fixed modulation combined with perfect channel reassignment. Although all of the points in Figure 4 achieve the TER of 1% the illustration does not show by what margin this target is achieved. To rectify this, Figure 5 considers the TER for the adaptive and fixed modulation schemes from Figure 4. The number of users is, therefore, a variable in Figure 5, however, the number of users supported by the adaptive schemes is always equal or greater than the number of users supported by the fixed schemes, as was shown in Figure 4.

Considering Figure 5(a), it can be seen for the Slow fading channel that with the exception of one value the adaptive scheme results in a better channel quality than the fixed scheme. Furthermore, for the slow fading channel there is no clear advantage in using either the peak- or mean- BER adaptive switching levels. At the higher speed portrayed in Figure 5(b), both sets of switching levels show even higher improvements over the fixed schemes and there is a marked advantage with the peak BER switching levels. The mean and peak switching levels, result respectively, in up to a factor of three or 30 reduction in TER, respectively. The peak BER switching levels are much more attractive than the mean BER switching levels, since the DER becomes negligible in both cases, as the channel becomes less correlated. Therefore, the TER tends to the FER, which is much lower with the peak BER switching levels. The conclusion of Figures 4 and 5 is that adaptive modulation, improves the quality of the link expressed in terms of FER, while accommodating more users.

5. ACKNOWLEDGEMENT

The financial support of the following organisations is gratefully acknowledged: Motorola ECID, Swindon, UK; European Community, Brussels, Belgium; Engineering and Physical Sciences Research Council, Swindon, UK; Mobile Virtual Centre of Excellence, UK.

6. REFERENCES

[1] Toyoki Ue, Seiichi Sampei and Norihiko Morinaga,"Symbol Rate and Modulation Level Controlled Adaptive Modulation/TDMA/TDD for Personal Communication Systems," in *Proceedings of* 45^{th} *Vehicular Technology Conference*, pp.306-310, IEEE, 1995.

[2] Soon-Ghee Chua and Andrea Goldsmith, "Variable-Rate Variable-Power MQAM for Fading Channels," in *Proceedings of* 46^{th} *Vehicular Technology Conference*, pp. 815-819, IEEE, 1995.

[3] J. Torrance, L. Hanzo: Latency considerations for adaptive modulation in slow Rayleigh fading, Proc. of IEEE VTC'97, Phoenix, USA, 1997, pp 1204-1209