www.ietdl.org

Published in IET Communications
Received on 27th February 2009
Revised on 13th November 2009
doi: 10.1049/iet-com.2009.0140

ISSN 1751-8628

Cross-layer optimisation of network performance over multiple-input multiple-output wireless mobile channels

M. Luccini   A. Shami   S. Primak

Bell Centre for Information Engineering, Department of Electrical and Computer Engineering, The University of Western Ontario, London, Ontario N6A 5B9, Canada
E-mail: ashami@eng.uwo.ca

Abstract: In this study a wireless multiple-input multiple-output (MIMO) communication system operating over a fading channel is considered. Data packets are stored in a finite size buffer before being released into the time-varying MIMO wireless channel. The main objective of this work is to satisfy a specific quality of service (QoS) requirement, i.e. the probability of data loss because of both erroneous wireless transmission and buffer overflow, as well as to maximise the system throughput. The theoretical limit of ergodic capacity in MIMO time-variant channels can be achieved by adapting the transmission rate to the capacity evolving process. In this study, the channel capacity evolving process has been described by a suitable autoregressive model based on the capacity time correlation and a finite state Markov chain (FSMC) has been derived. The joint effect of channel outage at the physical layer and the buffer overflow at the medium access control layer has been considered to describe the probability of data loss in the system. The optimal transmission strategy must minimise that probability of data loss and has been derived analytically through the Markov decision process (MDP) theory. Analytical results show the significant improvements of the proposed optimal transmission strategy in terms of both system throughput and probability of data loss.

1 Introduction

Multiple-input multiple-output (MIMO) wireless systems gained significant interest because of their high-spectral efficiency and increased channel capacity. According to the Shannon's theory, the channel capacity upper bounds the maximum amount of information that can be reliably transmitted over the communication channel, with an asymptotically small probability of error. The ability of transmitting close to the capacity limit relies on an accurate channel state information (CSI), which must be provided at the receiver for effective decoding. Moreover, pilots-assisted channel estimation can provide an accurate CSI for flat-fading channel even for moderate/high user mobility [1, 2]. Furthermore, if the CSI is available not only at the receiver but also at the transmitter, a high multiplexing gain can be achieved in multi-antennas transmissions [3], further increasing the spectral efficiency.

The knowledge of the channel state at the transmitter can be obtained by a feedback channel as described in [4]. Feeding the predicted channel capacity to the transmitter would allow to choose the right modulation, coding and power to match the capacity of the wireless channel, with a reasonable amount of feedback information.

Shannon's channel capacity is achieved by utilising long codes. Thus, the channel capacity plays a primary role in the modelling of the communication system, when there are no constraints on energy and coding/decoding complexity at the transmitter/receiver side. In this case it is possible to focus on the MIMO channel capacity, which represents the maximum amount of information supported by the MIMO channel. The main objective of this study is to investigate the optimal approach to the variable signalling rate control for a MIMO wireless time-variant channel.

In this study, a discrete autoregressive (DAR) prediction model is used to predict the capacity variation at specific instants of time. The DAR model is based on the capacity correlation and on the CSI at a previous moment in time. Thanks to the prediction model, the transmitter can adapt its signalling rate to the predicted capacity value. The DAR-1 model assigns a Markov nature to the instantaneous capacity process. Dividing the continuous capacity process into a finite number of discrete states, a finite state Markov chain model (FSMC) for the instantaneous MIMO capacity is obtained. The principle of FSMC is to discretise a continuous process into a finite number of states, over which the process itself can be qualified separately. The FSMC is a well accepted block fading channel model for slow-varying flat fading channels, where the channel is assumed to stay in the same state within one block period. In wireless communication, the Markovian assumption is widely used to model the channel fading and the signal to noise ratio (SNR) variation [5–7]. Moreover, the FSMC model and adaptive modulation and coding (AMC) strategies are commonly accepted as the fundamental techniques for developing effective cross-layer protocols and algorithms.

If a transmission system with a single finite buffer is considered, higher layer data packets are enqueued into a finite size buffer space before being released onto the time-variant wireless channel. The choice of the optimal signalling rate must take into account the buffer state: over the physical (PHY) layer, the buffer at the MAC layer is itself a FSMC dependent on the arrival process. The joint consideration of the capacity state and buffer state leads to a two-dimensional optimisation process, where the optimal signalling rate should be chosen according to the state of the instantaneous capacity and how many packets are present in the buffer. The choice of the signalling rate at each time frame must maximise the performance at each state: the probability of successful transmission (throughput) at that specific signalling rate. According to the outage definition provided in [8], the objective of this analysis is to minimise the average joint packet loss rate because of both outage and buffer overflow. In this case there are two design objectives (outage and buffer overflow) that jointly define the optimisation target: to minimise the end-to-end packet loss. The optimal transmission policy design is investigated analytically with the application of dynamic programming (DP) and Markov decision process (MDP) theories.

The rest of this paper is organised as follows. In Section 2, the motivations and contributions of this work are discussed. In Section 3, we give the system model used for this study. Capacity distribution for mobile MIMO channel is presented in Section 4. In Section 5, we briefly review some related research work. In Sections 6 and 7 FSMC for MIMO channel capacity and the proposed two-dimensional cross layer optimisation approach are discussed, respectively. Section 8 concludes this work.

2 Motivations and contributions

This study is motivated by the following considerations. First, the channel variation because of user mobility results into an instantaneous channel capacity variation. The instantaneous capacity can be modelled as a Gaussian random process, whose mean value is defined as the channel ergodic capacity. The ergodic capacity represents the theoretical limit of the amount of information, which can be reliably transmitted over the channel in a long-time observation. From a system point of view, the theoretical limit denoted by the channel ergodic capacity can be achieved in transmission by adapting the signalling rate to the evolution of the instantaneous capacity [8]; the signalling rate represents the amount of information transmitted on a specific timescale and it is a percentage of the channel capacity. By considering a channel evolution based on discrete time steps (i.e. the frame duration T_f), we propose to adapt the signalling rate to the capacity variation to reach the ergodic capacity. Second consideration is that the information to be transmitted is not a continuous process but rather arrives from the upper layer (MAC) and can be subject to loss due the finite buffer size. Therefore in order to optimise the system transmission, the signalling rate must be chosen by considering both the actual channel capacity and the buffer state. In other words, the maximum amount of information can be transmitted by adapting the signalling rate to the channel capacity and buffer state: the aim is to transmit at the maximum reliable rate and to minimise the probability of buffer overflow (i.e. losing packets at the buffer with no chance to recover them).

The major contributions of this work are summarised as follows:

1. this study proposes an analytical definition of FSMC for the MIMO channel capacity;

2. this study develops an analytical model of the proposed algorithm for evaluating quality of service (QoS) performance metrics at the MAC layer, such as the system throughput and probability of packet loss;

3. combined with previous research, this paper proposes an efficient method for cross-layer performance optimisation.

3 System model

Fig. 1 illustrates the general structure of the system model used for this study. Specifically, the PHY layer is characterised by multiple antennas at both sides of the communication link, which results in a significant channel capacity increase. At the PHY layer, the amount of information sent over the MIMO channel per unit time/bandwidth is defined as the signalling rate of the transmitter. The MAC layer of the system is characterised by a single finite buffer, in which data arriving from the higher application layers are enqueued. The dependence

IET Commun., 2010, Vol. 4, Iss. 6, pp. 683–696
doi: 10.1049/iet-com.2009.0140

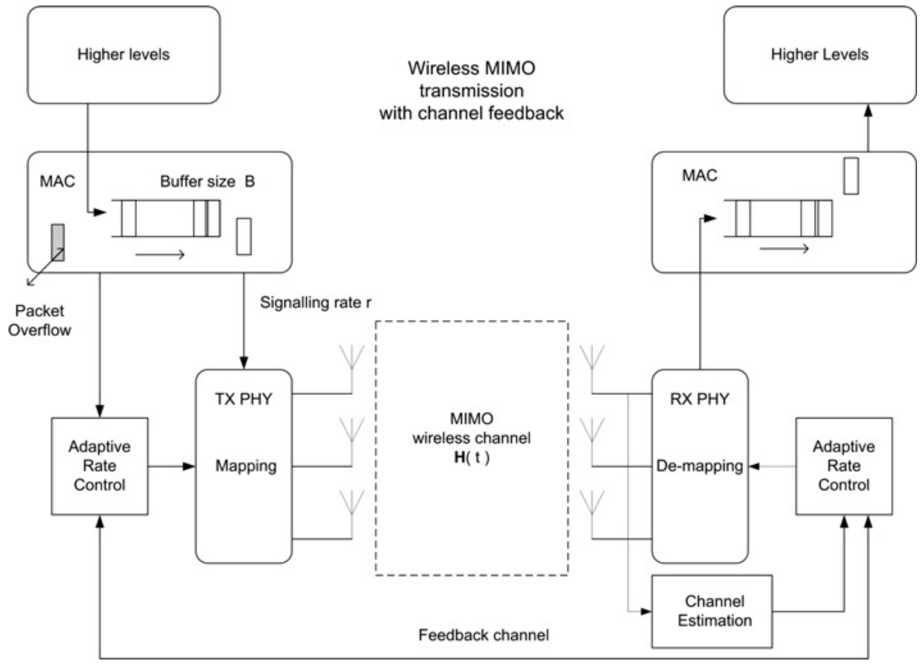

Figure 1 *Communication system scheme with feedback channel*

between the PHY and MAC layers lies in the PHY signalling rate, which represents the number of packets which are de-queued from the buffer and transmitted over the MIMO channel per unit time/bandwidth. This definition states the equivalence between the signalling rate at the PHY layer and the buffer service rate at the MAC layer. At the PHY layer, AMC is applied to achieve adaptive multirate transmission over the time-variant MIMO channel.

The instantaneous channel capacity can be well approximated by a Gaussian process [3, 9]. According to the definition of Shannon capacity [8], the channel capacity can be interpreted as the maximum amount of information that can be reliably transmitted on the MIMO channel, with a suitable code and modulation scheme at the PHY layer of the system. With no constraints on energy availability or delay in transmission, the instantaneous capacity represents the limit of information that can be transmitted by the system. The channel spectral efficiency (bits/s/Hz) is related to the signalling rate (packets/s) by the bandwidth required by the system [10] and the packet size. Data packets of size N_b bits are collected in a finite buffer of B packets. The system works with time frames of fixed duration T_f. The frame duration also represents the time step at which the signalling rate can be adjusted based on the channel capacity and buffer state.

4 Capacity distribution for mobile MIMO channel

The characteristics of the channel capacity in MIMO systems allow to develop a useful model for the instantaneous capacity variation, which will be used later to obtain the FSMC for the capacity process. In the literature, many studies deal with the statistical properties of the MIMO wireless capacity. Some of those statistical properties are the probability density function (PDF), the cumulative distribution function (CDF), the level crossing rate (LCR) and the average fade duration (AFD) of the channel capacity. However, deriving exact analytical expressions for all those statistical properties appears to be a complicated task. Referring to the principle of maximum entropy [11], we assume that elements of the channel matrix are independent and identically distributed (i.i.d.) zero mean Gaussian processes. If CSI is available at the receiver but not at the transmitter side, MIMO capacity is a random process, approximately Gaussian for large number of transmitting and receiving antennas N_T, N_R [9]. Defining $\beta = N_T/N_R$ and γ the average channel SNR, the mean and the variance of the capacity process are given by [9]

$$\frac{\bar{C}}{N_R} = \beta \ln\left[1 + \frac{\gamma}{\beta} - \frac{1}{4}F\left(\frac{\gamma}{\beta}, \beta\right)\right] + \ln\left[1 + \gamma - \frac{1}{4}F\left(\frac{\gamma}{\beta}, \beta\right)\right] - \frac{\beta}{4\gamma}F\left(\frac{\gamma}{\beta}, \beta\right) \quad (1)$$

$$\sigma_C^2 = -\ln\left(1 - \beta\left[\frac{1}{4\gamma}F\left(\frac{\gamma}{\beta}, \beta\right)\right]^2\right) \quad (2)$$

where

$$F(x, z) = \left[\sqrt{x(1+\sqrt{z})^2 + 1} - \sqrt{x(1-\sqrt{z})^2 + 1}\right]^2 \quad (3)$$

For small values of SNR γ these equations could be simplified to produce

$$\frac{\bar{C}}{N_R} = \gamma - \frac{1}{2}\frac{\beta+1}{\beta}\gamma^2 \tag{4}$$

$$\sigma_C^2 = \gamma - \frac{1}{2}\frac{\beta+2}{\beta}\gamma^2 \tag{5}$$

while for large SNR γ the following approximations are valid for $\beta = 1$ [12]

$$\frac{\bar{C}}{N_R} \simeq \ln\frac{\gamma}{e} + \frac{2}{\sqrt{\gamma}} - \frac{1}{\gamma} \tag{6}$$

$$\sigma_C^2 \simeq \frac{1}{2}\left(\ln\frac{\gamma}{4} + \frac{2}{\sqrt{\gamma}} - \frac{1}{4\gamma}\right) - 0.193 \tag{7}$$

Interestingly, Hochwald *et al.* [3] considered the case of a reasonable number of antennas N_T, N_R, proving that the Gaussian distribution holds in the case of small numbers of antennas and i.i.d. channel coefficients. High accuracy in the capacity distribution is obtained even for 2×2 MIMO systems. The Gaussian approximation becomes more and more accurate increasing N_T and N_R.

Referring to the problem under analysis, we are interested in a prediction model for the capacity variation. The definition of capacity for MIMO channels [9, 13] suggests that the instantaneous capacity evolves in time according to the fading process. This consideration has been deeply investigated in [14]. Specifically, referring to the block fading model, the time variation of the channel is described by the normalised Doppler frequency $f_m = f_d T_f$, where f_d is the Doppler frequency in Hz and T_f is the system frame duration. The value of f_m does not affect the statistical channel description, once a proper observation interval is set to derive the statistical characteristics. If the process exhibits a long autocorrelation function, the need of independent samples to achieve accurate statistics leads to a long process observation time. The time variation of the instantaneous capacity could be described in terms of DAR models, as presented in [15, 16]. In the simplest case of DAR-1 model, the capacity evolution equation is given by

$$\Delta C(n+1) = \rho \Delta C(n) + \sqrt{1-\rho^2}\,\xi(n+1) \tag{8}$$

where $\Delta C(n+1) = (C(n+1) - \bar{C})/\sigma_C$ is the normalised deviation of capacity $C(n+1)$ at the discrete time $n+1$ from its time average $\bar{C}$; $\xi(n+1)$ is a sequence of i.i.d. zero mean Gaussian variables with unit variance. The analytical investigation of the correlation coefficient ρ is still an open issue. It can be estimated from experimental data [15] or its value could be deduced based on some theoretical development. Although a complete analytical description of the MIMO channel capacity is still under research, the study in [14] suggests that the instantaneous capacity evolves with the same statistical characteristics of the fading channel. The results presented in that study support the conclusion that the channel autocorrelation can be a valid model also for the capacity correlation coefficient ρ. Starting from the Gaussian wide sense stationary (WSS) uncorrelated scattering model, the autocorrelation sequence for ideally generated in-phase and quadrature Gaussian processes at discrete time is given by

$$R[n] = J_0(2\pi f_m |n|) \tag{9}$$

where J_0 is the zero-order Bessel function of the first kind. Equation (9) allows to obtain the correlation information of the capacity process once the discrete prediction interval nT_f is set. At the transmitter, the DAR-1 prediction model can be set up with the knowledge of the statistical moments of the capacity process to obtain the instantaneous capacity value for the next time frame. Moreover, the DAR-1 model assigns Markov nature to $C(n)$.

4.1 Outage probability

Capacity specifies how much information the channel can support, which turns into a maximum amount of information the system can reliably transmit. The probability of outage describes the frequency at which capacity falls under a specific signalling rate: in that case no reliable transmissions are possible. A certain maximum probability of outage P_{out} is required for the constant data rate transmission over the MIMO channel. The outage appears when the signalling information rate r exceeds the instantaneous capacity C_n at time n, therefore [12]

$$P_{out} = \mathrm{Prob}(C_n < r) = \int_{-\infty}^{R} p_C(x)\,dx = Q\left(\frac{\bar{C}-r}{\sigma_C}\right) \le \frac{1}{2}\exp\left[-\frac{1}{2}\left(\frac{\bar{C}-r}{\sigma_C}\right)^2\right] \tag{10}$$

where $Q(x)$ is the well-known Q-function [10] and $\bar{C}$ and σ_C can be calculated using (1) and (2). Solving (10) for r in terms of desired P_{out} allows one to set a proper information rate on the channel. If $P_{out} \ll 1$, which is usually a case, one can obtain a simple approximation from the upper bound in (10)

$$r \simeq \max\left\{\bar{C} - \sqrt{-2\sigma_C^2 \ln 2P_{out}},\ 0\right\} / \ln(2) \quad \text{(bits/s/Hz)} \tag{11}$$

where r is the maximum signalling rate the channel can support given a desired probability of outage, $\bar{C}$ is the average channel capacity per antenna specified in (1), P_{out} is the desired outage probability and ln is the natural logarithm.

IET Commun., 2010, Vol. 4, Iss. 6, pp. 683–696
doi: 10.1049/iet-com.2009.0140

5 Review of previous works on finite state Markov model for radio communication channels

The FSMC for random processes is a useful model for studying the process behaviour in time. Of primary importance for the FSMC model are the state probabilities and the transition probabilities among different states. We present here the main model for communication channel and an efficient way of deriving the transition probabilities among all the possible states based on an eigenvalues analysis. The literature reveals many attempts to model the fading envelope of a time variant communication channel and the resulting error flow using finite Markov chains. In this section, the meaning and importance of FSMC as a model of radio communication channels are discussed in order to introduce the problem of the FSMC model for the MIMO channel capacity.

The study of the FSMC to model a communication channel emerges from the early work of Gilbert [17] and Elliott [18] who studied a two-states Markov channel known as the Gilbert–Elliott channel. Their model is composed of a good state G and a bad or burst state B, the transition probabilities are made in order to simulate the burst error conditions. More complex Markov models followed this first one, in the attempt of modelling more accurately the communication error statistical behaviour [19]. When the channel quality varies significantly, the two-state Gilbert–Elliott model is not adequate. A straightforward solution is to increase the possible states of the model to a finite set $S = \{s_0, s_1, \ldots, s_{K-1}\}$; if the process under analysis is ergodic, the resulting Markov process $\{S(n)\}$, for $n = 1, 2, \ldots$ will be a constant Markov process with the related properties. Looking at wireless communications in fading environments, many works suggested the idea of creating a FSMC for the channel state [6, 7, 20], obtained by partitioning the SNR at the receiver side in K possible states. The transition probabilities between two channel states depend on the statistical description of the random fading process. Wang and Moayeri [20] proposed a well-known analytical model to calculate the transition probabilities by looking at the fading process and SNR statistics. The core of this model is the evaluation of the LCR for the SNR process. Given a random process x, the LCR of a specific value $\hat{x}$ in a time interval Δt is defined as

$$\int_t^{t+\Delta t} p\{x > \hat{x}\}\,\mathrm{d}t \tag{12}$$

and it is dependent on the duration of the observation interval Δt. The LCR is defined as the number of crossing per second of a given threshold. Defining $\Delta x = \dot{x}\Delta t$, a possible way to obtain the LCR is to derive the statistical distribution of the first derivative $\dot{x}$, which represents the velocity of the process, and then evaluate the probability of x being around the threshold $\hat{x}$ with all the possible velocities

$$\mathrm{LCR}(\hat{x}) = \int_0^{\infty} \dot{x} p(\hat{x}, \dot{x})\,\mathrm{d}\dot{x} \tag{13}$$

where $p(\hat{x}, \dot{x})$ is the joint density of the process $\hat{x}$ and its first derivative $\dot{x}$. In the case of Gaussian process this integral can be solved in a closed form which involves the second derivative of the correlation coefficient $\rho_{xx}(\Delta t)$ of the process x under analysis. Knowing the statistics of the received SNR, the probability inside the integral can be expressed in a closed form. The study of the LCR allows to approximate the transition probabilities to the adjacent channel states in the simple form

$$p_{i,i+1} \simeq \frac{\mathrm{LCR}_{i+1}}{R_t^{(i)}}, \quad i = 0, 1, \ldots, K-2 \tag{14}$$

where LCR_{i+1} is the number of up-crossing rate per second and $R_{\mathrm{t}}^{(i)} = R_{\mathrm{t}} \times p_i$ is the symbol transmission rate weighted by the state probability p_i. The LCR can be derived analytically from the received SNR distribution [20]

$$\mathrm{LCR}(a) = \sqrt{\frac{2\pi a}{\bar{\gamma}}} f_{\mathrm{d}} \exp\left\{-\frac{a}{\bar{\gamma}}\right\} \tag{15}$$

where $\gamma = E\{A\}$ denotes the mean value of the received SNR and $f_{\mathrm{d}} = v/\lambda$ is the Doppler frequency of the mobile user normalised to the carrier wavelength λ. Equation (14) is referred as the Wang–Moayeri model for adjacent states transition probabilities, which can be determined as

$$\begin{aligned} p_{k,k+1} &= \frac{N_{k+1}T_{\mathrm{f}}}{p_k}, \quad k = 0, \ldots, K-1 \\ p_{k,k-1} &= \frac{N_k T_{\mathrm{f}}}{p_k}, \quad k = 1, \ldots, K \end{aligned} \tag{16}$$

in which N_k is the cross rate for state k, either upward or downward and p_k is the kth state probability. By the assumption in the model [20], the probability of remaining in the same state k is defined as

$$p_{k,k} = \begin{cases} 1 - p_{k,k+1} - p_{k,k-1}, & \text{if } 0 < k < K \\ 1 - p_{0,1}, & \text{if } k = 0 \\ 1 - p_{K,K-1}, & \text{if } k = K \end{cases} \tag{17}$$

The Wang–Moayeri method (for the evaluation of transition probabilities of a random process) has been used in [5] to generate a FSMC of the channel state based on the received SNR. In [5], the modulation and coding at the transmitter sides are adapted to the variation of the SNR of the communication link, modelled as a FSMC. The proposed channel partition method CPM maintains a certain level of average packet error rate (PER) over the time-variant channel when corresponding AMC mode is applied for each channel state. When the target PER is not fixed a priori, the overall packet loss at PHY and MAC layers can be minimised through cross-layer analysis. The use of LCR for MIMO capacity appears in [15], in which the estimation of LCR is a good parameter to obtain the transition probabilities in case of slow fading environments.

One limit of the above framework is that it can only provide a simple way to obtain the transition probabilities between adjacent states. Consequently, it can be inferred that this model works well for slow varying processes which evolve only to adjacent states in the time observation interval Δt. Defining f_{d} as the Doppler frequency in Hz and T_{f} as the frame duration in s, for high normalised Doppler frequency $f_{\mathrm{m}} = f_{\mathrm{d}} T_{\mathrm{f}}$ the process could jump to non-adjacent states and the crossing probability between the adjacent states may turn in very low values. In order to overcome this limitation, in case of moderate/high normalised Doppler frequency a different approach to the problem must be taken.

The main disadvantage of evaluating the transition probabilities by the Wang–Moayeri model [20] or by the analytical framework in [15] is the fact that, while the elements of the transition matrices can be analytically calculated, it is especially difficult to obtain analytical expressions for eigenvalues and eigenvectors of the matrix of transitional probabilities. The model proposed in [21] allows to determine the transition probability matrix for a K-state Markov chain with the knowledge of the process correlation interval d, which is the second largest eigenvalue of the generated process [22]. The exact knowledge of eigenvalues greatly reduces the complexity and accumulation of numerical error. Let $\boldsymbol{P} = [p_{ij}]$ be a matrix of transitional probabilities of the DAR-1 K-states Markov chain. It is shown in [21] that $\boldsymbol{P}$ is defined as

$$\boldsymbol{P} = \boldsymbol{Q} + d \times (\boldsymbol{I} - \boldsymbol{Q}), \quad 0 \le d < 1 \tag{18}$$

where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{Q}$ is composed by the steady state probabilities as shown

$$\boldsymbol{Q} = \begin{bmatrix} p_0 & p_0 & \cdots & p_0 \\ p_1 & p_1 & \cdots & p_1 \\ \cdot & \cdot & \cdots & \cdot \\ p_{K-1} & p_{K-1} & \cdots & p_{K-1} \end{bmatrix} \tag{19}$$

where $\{p_k\}_{k=0,\ldots,K-1}$ are the stationary probabilities. The resulting Markov process ι at discrete time $\kappa = 0, 1, \ldots$ has an exponential autocorrelation function

$$R_u(\kappa) = R_u(0) d^{\kappa} \tag{20}$$

which matches the desired autocorrelation of an AR-1 generated process. The value d corresponds to the autocorrelation of the process at time $\kappa = 1, 2, \ldots$; if the information on the autocorrelation is available, it is possible to obtain the transition probabilities by the first two matrix eigenvalues. In the case of a mobile wireless communication system, the channel autocorrelation between two frames of duration T_{f} is $\rho = J_0(2\pi f_{\mathrm{d}} T_{\mathrm{f}})$ [10], where f_{d} is the Doppler frequency of the mobile user and T_{f} is the frame duration.

The method proposed in [21] provides the transition probabilities among all the possible states K and not only between adjacent states. For this reason, especially for significant normalised Doppler frequency f_{m} (moderate/high fading), the transition probabilities obtained by this model are to be preferred to the ones based on LCR in (16) and (17).

The eigenvalues method will be used in the next sections to evaluate the transition probabilities for the FSMC of the instantaneous channel capacity.

6 FSMC for MIMO channel capacity

As in [11, 12], the instantaneous capacity is well described by a random Gaussian process with mean $\boldsymbol{C}$ and variance $\sigma_C^2 : C \sim N(\boldsymbol{C}, \sigma_C^2)$. Referring to the analysis in Section 4, the DAR-1 evolution model of the channel capacity process in (8) assigns a Markov nature to the process, in the specific case of the first order. In this model, the information on the next state is gained only from the current state, assuming that information corresponding to previous states is negligible. Following the idea proposed in Section 5, a FSMC can be obtained by partitioning the instantaneous capacity process into a finite number of intervals or states $S = \{s_0, s_1, \ldots, s_{K-1}\}$, with corresponding thresholds $\{c_k\}_{k=0}^{K}$. Capacity is said to be in state s_k, $k = 0, 1, \ldots, K-1$ if the value C of the process is in the interval $[c_k, c_{k+1})$, that is

$$c_k \le C < c_{k+1} \tag{21}$$

As discussed in [21], partitioning must be performed such that the highest state probability is assigned to the state s_k^*, which contains the average value of the capacity process, by selecting boundaries $[c_k^*, c_{k+1}^*)$ such that

$$\bar{C} = \frac{c_k^* + c_{k+1}^*}{2} \tag{22}$$

The states surrounding s_k^* on the left and right of the process PDF must have decreasing probabilities with respect to s_k^*. If the partitioned process is ergodic, the Markov process S_k is a stationary process, with the property that transition probabilities are time invariant [23]. The evolution of the capacity process is related to a fixed timescale defined by the frame duration T_{f} of the system. According to the general block fading model, the channel capacity is assumed to remain constant within one block period, with a block length equal to T_{f}. The instantaneous capacity evolution model has been presented in (8). Given the value of capacity at time index nT_{f} with $n = 0, 1, \ldots$, the capacity process at time $n + 1$ is a Gaussian random

doi: 10.1049/iet-com.2009.0140

variable related to the original process by the correlation coefficient $\rho(T_{\mathrm{f}})$ (from now on the parameter T_{f} will be omitted). Specifically, the Gaussian process at one step prediction is characterised by the following first two moments

$$\begin{aligned}\bar{C}_{n+1} &= (1-\rho)\bar{C}_n + \rho C(n) \\ \sigma^2_{n+1} &= (1-|\rho|^2)\sigma^2_n\end{aligned} \tag{23}$$

For a FSMC model the two probabilities of main interest are the steady state probability, which describes the asymptotic probability of being in a given state s_k and the transition probabilities, which drive the transitions among different states.

1. *Steady State Probabilities:* the probability p^o_k that the instantaneous capacity is in state s_k is defined as

$$p^o_k = \int_{c_k}^{c_{k+1}} p(C)\,\mathrm{d}C \tag{24}$$

for $k = 0, 1, \ldots, K-1$. For Gaussian random variables, the steady state probability is easily expressed by $Q(c_k) - Q(c_{k+1})$, where $Q(x)$ represents the well-known Q-function [10]. The steady state probabilities can be arranged in the following vector form

$$\boldsymbol{\pi} = \lim_{n\to\infty} p(s(n) = s_k) \tag{25}$$

for time index $n = 0, 1, 2 \ldots$ and for $k = 0, 1, \ldots, K-1$.

2. *Transition Probabilities:* the probability of transitions between two states is a conditional probability which can be obtained by the joint PDF of the state distribution, according to the Bayes' rule

$$p(b|a) = \frac{p(b, a)}{p(a)} \tag{26}$$

The two variables under attention are the capacity values C at time n and $n+1$. According to (8), the capacity evolution $C(n)$ is still a Gaussian variable. For the particular case of two Gaussian random variables, the joint probability is given by [24]

$$p\{x_1, x_2\} = \int_{c_i}^{c_{i+1}}\int_{c_j}^{c_{j+1}} \frac{1}{\sqrt{2\sigma_1^2\sigma_2^2(1-\rho^2)}}\exp\left[-\frac{1}{2(1-\rho^2)} \times \left(\frac{\delta x_1^2}{\sigma_1^2} + \frac{2\rho\delta x_1\delta x_2}{\sigma_1\sigma_2} + \frac{\delta x_2^2}{\sigma_2^2}\right)\right]\mathrm{d}x_1\,\mathrm{d}x_2 \tag{27}$$

where δx is defined as $x - \mu_x$ and ρ is the correlation coefficient between x_1 and x_2 at the sampling time T_{f}. The transition probability is then the joint PDF weighted by the marginal density of one of the two variables. Transition probabilities are arranged in a matrix $\boldsymbol{P}$.

The steady state and transition probabilities can be evaluated numerically once a set of thresholds $\{c_k\}_{k=0}^K$ has been defined. The steady state probabilities of the capacity process are evaluated by (24) as the probability of being a specified state

$$p\{C \in s_k\} = \int_{c_k}^{c_{k+1}} p_C(x)\,\mathrm{d}x \tag{28}$$

where $p_C(x)$ is the Gaussian PDF of the instantaneous capacity.

Referring to the theory in [23], a probability distribution $\{p_k | k \geq 0\}$ is said to be a stationary distribution for the Markov chain if

$$p_k = \sum_{i=0}^{K-1} p_i p_{i,k}, \quad k = 0, 1, \ldots, K-1 \tag{29}$$

Equation (29) can be extended to all the states in the vector equation

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P}, \quad \sum_{k\in S}\pi_k = 1 \tag{30}$$

Equation (30) shows that $\boldsymbol{\pi}$ is left eigenvector of the matrix of transition probabilities $\boldsymbol{P}$ corresponding to the eigenvalue 1. Thus the eigenvalues method proposed in [21] seems to be appropriate for evaluating the elements of the transition probability matrix. We considered (30) to validate the state and transition probabilities. Numerical results showed that the method based on the joint PDF in (27) is easily subject to numerical error because of the precision of the integral evaluation, which can only be solved numerically. As previously discussed, the LCR method shows an unrealistic high probability remaining in the same capacity state even for high normalised Doppler f_{m}. The eigenvalue framework proposed in [21] has shown the best behaviour, matching with high accuracy the Markov property in (30).

In conclusion, once the transition probabilities are computed, the FSMC for the channel capacity is modelled as a $(K+1)\times(K+1)$ transition probability matrix, with the form

$$\begin{bmatrix} P_{0,0} & P_{0,1} & \ldots & P_{0,K} \\ P_{1,0} & P_{1,1} & \ldots & \vdots \\ P_{2,0} & \ddots & \ddots & P_{2,K} \\ \vdots & P_{K-1,K-2} & P_{K-1,K-1} & P_{K-1,K} \\ \ldots & \ldots & P_{K,K-1} & P_{K,K} \end{bmatrix} \tag{31}$$

7 MDP theory: policy domain performance optimisation

The knowledge of the variation of the instantaneous capacity process leads to describe the evolution of the maximum amount of information that can be sent over the channel. Once a set of possible signalling rates is available at the PHY layer, the question of which rate must be chosen according to the capacity state becomes of primary importance. This kind of problem is referred in the literature as a two-dimensional optimisation problem, which can be investigated by the application of DP and MDP theories for the optimal transmission policy design.

The question of which is the best signalling rate can be addressed by introducing the definition of reward for each possible choice [25]. Generally speaking the definition of 'reward' refers to a specific performance metric. The idea beyond the optimisation procedure is that the total expected reward must be maximised either after a finite number of iterations or in the perspective of an indefinite working horizon.

In this study the theory of MDP is applied to the choice of the best signalling rate according the system conditions. Specifically, the value of 'reward' is defined as the number of packets correctly received at the destination MAC layer. Consider a completely ergodic K-states Markov process: the process states are described by the steady-state probabilities and the transition probability matrix $\boldsymbol{P}$, the reward matrix $\boldsymbol{G}$ associates a specific reward with each state and transition. The process is allowed to evolve for a very long time and the parameter of interest is the total expected reward of the process. Since the process is completely ergodic, the limiting steady state probabilities $\pi_{i,j}$ are independent of the staring state, the gain of the process is defined as [25]

$$g = \sum_{i=0}^{K-1} \pi_i q_i \tag{32}$$

where the quantity q_i is the immediate reward in state i defined as [25]

$$q_i = \sum_{j=0}^{K-1} p_{i,j} r_{i,j} \tag{33}$$

It is important to note that every ergodic Markov process with reward will have a gain defined as in (32). The same equation can be arranged in a vector form considering all the states $i = 0, \ldots, K-1$

$$g = \boldsymbol{\pi} \times \boldsymbol{q}^{\mathrm{T}} \tag{34}$$

If different actions can be taken at any state, the transition matrix and the reward structure will change according to the chosen action. At each state, a specific action has effects on the probability distribution of transitions and on the rewards, since each action reasonably introduces a specific cost and result. In the case of study, the set of available signalling rates constitutes the action space $A(s)$ for each capacity state $s \in S$. A policy is defined when one specific action $a \in A$ has been defined for each state of the model. Bringing into the model the system buffer, a convenient signalling rate can be chosen according both to the capacity state and queue buffer state. Suppose that the system has a single buffer of B packets. The buffer state defines how many packets are waiting in the buffer and it is itself a FSMC dependent on the arrival process. The higher levels arrival process is described by a Poisson model [23]

$$\mathrm{Prob}\{A(\tau) = m\} = \frac{\mathrm{e}^{-\lambda\tau}(\lambda\tau)^m}{m!}, \quad m = 0, 1, \ldots \tag{35}$$

where $A(\tau) = m$ denotes the event of m arrivals in the time interval τ. λ is the average arrival rate in packets/s. The capacity transitions and the arrival process are considered two independent processes because of their different nature. Considering both the capacity and buffer states leads to a two-dimensional problem, in which the system is completely characterised by the state pair $s(k, q)$, where k is the capacity state index and q is the MAC layer queue length.

In order to refer to the MDPs theory, we define the state transition probability matrix and the reward matrix as follows.

1. *State transition probability matrix* $\boldsymbol{T}$*:* It is a three-dimensional matrix which orders the possible actual states, the possible succeeding states and the possible actions (or alternatives), which can be chosen in each state. Each 'slice' ζ_a corresponds to the set of all possible transitions of the pair state capacity-buffer: $(k, q)(k', q')$ for $k = 0, 1, \ldots, K-1$, $q = 0, 1, \ldots, B$ (see (36))

The transition probability $p^a_{(k,q),(k',q')}$ is function of the parameters (a, k, q, k', q') and represents the probability of passing from the state (k, q) to the new state configuration (k', q') in terms of capacity and buffer size. The value a specifies the action to take at time T_{f}, in this case the

$$\zeta_a = \begin{bmatrix} p^a_{(0,0);(0,0)} & \cdots & p^a_{(0,0);(1,0)} & \cdots & p^a_{(0,0);(K-1,B-1)} \\ p^a_{(0,1);(0,0)} & \cdots & \cdots & \cdots & \vdots \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ p^a_{(K-1,B-1);(0,0)} & \cdots & p^a_{(K-1,B-1);(1,0)} & \cdots & p^a_{(K-1,B-1);(K-1,B-1)} \end{bmatrix} \tag{36}$$

IET Commun., 2010, Vol. 4, Iss. 6, pp. 683–696
doi: 10.1049/iet-com.2009.0140

choice of the signalling rate for the current block. The value of $p^a_{(k,q),(k',q')}$ in $\boldsymbol{T}$ is defined for any possible action in the current $a \in \mathcal{A}(s_{(k,q)})$, as

- if $(q - \min(q, \varphi^a_{\max}) \le q' < B)$, then

$$p^a_{(k,q),(k',q')} = p\{A(T_f) = q' - [q - \min(q, \varphi^a_{\max})]\} \times p_{k;k'} \tag{37}$$

- if $(q - \min(q, \varphi^a_{\max}) \le q' = B)$, then

$$p^a_{(k,q),(k',q')} = 1 - \sum_{q'=0}^{B-1} p\{A(T_f) = q' - [q - \min(q, \varphi^a_{\max})]\} \times p_{k;k'} \tag{38}$$

where $\varphi^a_{\max}$ is the maximum number of packets that can be served in the time interval T_f when a specific action a is taken. $A(T_f)$ denotes the expected number of packets that will arrive in the next interval T_f according to the arrival process. The probability of arrival are modelled according to the Poisson arrival process, the capacity transition probabilities are derived for every state transition according to the FSMC of channel capacity. The optimal choice for the action a to take is the optimisation problem which requires the knowledge of the rewards for each action.

2. *State transition reward matrix* $\boldsymbol{G}$: Each element specifies the reward associated with a state transition (k, q), (k', q') for a given action a. With this definition, elements of $\boldsymbol{G}$ are function of the set $\{a, k, q, k', q'\}$ and defined as

$$g^a_{(k,q);(k'q')} = \begin{cases} \min(q, \varphi^a_{\max}) \times [1 - P^a_{\text{out}}(k)], & p^a_{(k,q);(k',q')} \neq 0 \\ 0, & \text{otherwise} \end{cases} \tag{39}$$

where $P^a_{\text{out}}(k)$ denotes the probability of outage on the channel, which is the probability of no reliable transmission as seen in Section 4.1.

The optimisation problem works on the variable a, which denotes the possible actions to take at every decision epoch T_f. In an optimal model, at each step T_f the transmitter predicts the instantaneous capacity and obtain a set of possible signalling rates associated with outage probability. The action a is the decision of a specific rate to use in the next time interval T_f. A decision on the signalling rate will have consequences on the buffer queue (how many packets the system will serve) and on the outage probability in transmission. When the optimal action is decided for every possible state, a policy μ is obtained. From the matrices $\boldsymbol{T}$ and $\boldsymbol{G}$, the optimal policy for signalling rate selection can be solved by the policy iteration method discussed in [25]. The resulting policy μ_0 is a vector $\boldsymbol{d}$, which specifies the best action a to take at each possible state of the process.

Since any policy also defines a specific state probability $p^a_{(k,q),(k',q')}$ and reward $g^a_{(k,q),(k'q')}$ for each process state, each policy μ thus specifies a unique Markov process with rewards. Given the specific definition of reward, the transmission policy resulting from the MDP analysis is optimal in the sense that it maximises the total reward of the process.

8 Two-dimensional cross layer optimisation

In this section the improvement in the system performance brought by the two-dimensional optimisation is analysed. The possibility of adapting the signalling rate to the channel capacity state and the buffer state reduces the total probability of losing packets in transmission.

Referring to the analysis in [26], the probability of data loss depends on both capacity outage and buffer overflow. Moreover, Bai and Shami [27, 28] emphasise that the optimal choice of the signalling rate must take into consideration not only the channel state evolution but also the system buffer state. In terms of throughput at the MAC layer, the optimal transmission strategy depends also on the amount of information available at the buffer: the optimal signalling rate allows transmitting all the data in the buffer without incurring in the channel outage probability. This fact leads to a deep cross-layer analysis, whose purpose is to transmit at optimal signalling rate given the conditions of the buffer and communication channel.

To avoid a deep analysis of the modulation and code scheme at the PHY layer, this work proceeds on the assumption that the information level described by the channel capacity can be achieved by a proper code and modulation scheme. This hypothesis holds where there are no constrains on transmission delay and energy consumption. The optimisation target turns into the MAC layer throughput for the optimal transmission policy design.

Consider the FSMC model of the capacity evolution described in Section 6. The frame duration T_f represents the time interval in which the system transmits with a fixed information rate, established at the beginning of the time frame. Specifically, let us consider the knowledge of the instantaneous capacity $C(n)$ at a specific time n. $C(n+1)$ is the predicted instantaneous capacity at the next time step according to the specific DAR-1 evolution model in (8), Section 4. The probability distribution of the predicted instantaneous capacity is given by $p_C(C(n+1)|C(n))$: it is conditioned to the previous value $C(n)$ and it is still Gaussian. The mean value and variance of prediction are obtained by taking the expectation of the AR model, resulting in (23). The transmitter predicts the instantaneous state of the channel capacity $C(n+1)$ based on the knowledge of the previous state $C(n)$ and establish a possible information rate $r(n+1) \leq C(n+1)$. The corresponding probability of outage is given by [12]

$$P_{\text{out}} = \text{Prob}(C(n+1) < r(n+1)) = \int_{-\infty}^{R} p_C(x)\text{d}x$$
$$= Q\left(\frac{C(n+1) - r(n+1)}{\sigma_{n+1}}\right) \tag{40}$$

For any given rate $r(n+1)$ the corresponding outage probability can be calculated with the use of the mean and variance of the DAR-1 predicted capacity. As from (23), the moments of the distribution depends on the correlation coefficient ρ of the capacity process. For a long prediction interval the correlation coefficient decays to zero resulting into the mean and variance of prediction equal to the mean and variance of the original process: no additional information is available for the prediction. If a specific outage is required as a quality metric in the transmission, the corresponding maximum signalling rate can be derived according to [12]

$$r(n+1) \simeq \max\left\{\bar{C}_{n+1} - \sqrt{-2\sigma^2_{n+1}\ln(2P_{\text{out}})}\right\} \tag{41}$$

where P_{out} is the desired outage probability and $\boldsymbol{C}_{n+1}$, σ^2_{n+1} are the mean and variance of the DAR-1 prediction.

From a MAC layer point of view, deciding for a specific rate $r(n+1)$ has two effects: on one side a value of $\varphi(n+1)$ packets will be transmitted in the next T_{f} if present in the buffer; on the other side a price will be paid in terms of probability of loosing packets because of outage. The FSMC of the instantaneous capacity defines the transition probability matrix of the capacity process. The capacity transition in the next time step is driven by the values of the transition matrix, with each arriving state associated with a probability value. For each one of the capacity states we define a proper set of permitted signalling rates $R = \{r_k\}_{k=0}^{K-1}$, where

$$r_k \leq c_k, \quad r_k \in R, \quad 0 \leq k \leq K-1 \tag{42}$$

For each one of the possible rates, the outage probability would be the probability that capacity at the predicted state is under the rate value. When outage occurs no reliable transmission are possible since the channel do not support and carry the information that is being sent. For each capacity state, the possible signalling rates are a subset of the set R, according to the outage requirement. If the signalling rates r_k are expressed in packets/s, the packet error rate (PER) is given by $r(n+1) \times P_{\text{out}}$ while a successful data transmission is obtained for $r(n+1) \times (1 - P_{\text{out}})$.

8.1 Numerical simulations

We compared the system behaviour in case of the MDP adaptive rate transmission policy and the constant rate transmission policy. In the model specified in Section 7, the reward structure of the problem is defined as the number of packets which are correctly received on the other side of the communication link, as specified in (39). For this reason the process reward g can be related directly to the throughput η of the system through the time frame duration

$$\eta = g/T_{\text{f}} \tag{43}$$

A second QoS parameter considered is the packet loss rate at the source. Disregarding the origin of the loss, the total packet loss rate is the complementary part of both the total packet arrived at the source and the packet correctly received. Therefore the source packet loss rate is defined as

$$\varepsilon = 1 - \frac{g}{\lambda T_{\text{f}}} \tag{44}$$

According to (39), the reward matrix for the constant rate transmission is a matrix of equal elements since the signalling rate is constant for all the states. Once the signalling rate is specified, the corresponding outage probability limits the process reward.

The simulations parameters are presented Table 1. We considered a MIMO 4 × 4 system with AWGN and SNR equal to 10 dB. The number of antennas equal to 4 is a reasonable choice as it represents a good trade-off between performance improvements and complexity. Referring to the 802.16e standard, the transmission band is equal to 1.25 MHz. We considered a carrier of 2 GHz and a user speed of 10 m/s, which is a reasonable average speed in urban environments. The buffer size is set to $B = 20$ packets and the capacity process is divided into $K = 7$ states. The signalling rates are chosen according to (41) for a target P_{out} of the system. This leads to a state transition

Table 1 System parameters

Description	Parameter	Value
SNR	SNR	10 dB
band	W	1.25 MHz
Tx, Rx antennas	N_{T}, N_{R}	4
bits per packet	N_{b}	1080
carrier frequency	f_{c}	2 GHz
user speed	v	10 m/s
frame duration	T_{f}	0.002 s
buffer size (packets)	B	20
capacity states	K	7

IET Commun., 2010, Vol. 4, Iss. 6, pp. 683–696
doi: 10.1049/iet-com.2009.0140

matrix of size $B \times K \times A$ where A is the number of suitable signalling rates in the set R specified by (42).

The two sources of data loss are the probability of outage P_{out} at the PHY layer and the probability of buffer overflow P_{overflow} at the MAC layer. In order to describe the whole system behaviour, the total probability of failure is defined as the sum of those two components: $P_{\mathrm{FAIL}} = P_{\mathrm{out}} + P_{\mathrm{overflow}}$. Note that the sum of the two probabilities is the lower bound of the successful transmission probability, defined as

$$\nu = (1 - P_{\mathrm{out}})(1 - P_{\mathrm{overflow}}) \tag{45}$$

The buffer overflow and the outage events are independent; the first one is related to the arrival process and buffer size, the second one depends on the channel capacity random process. Fig. 2 shows the average system throughput for a constant rate transmission with two different strategies. In the first case (dashed circle line), the signalling rate is chosen to avoid channel outage, in the second case a more aggressive strategy is implemented and the system works at a signalling rate equal to the channel ergodic capacity (mean value of the channel capacity distribution). In the first case, the system works with a low signalling rate which becomes insufficient to avoid the buffer overflow when the packet arrival rate increases: in (2) the system throughput is limited by the buffer overflow (P_{overflow}). In the second case, the system works with a higher signalling rate: the average system throughput increases for higher arrival rates, even though it is continuously limited by the outage probability (P_{out}).

Fig. 2 shows that the average system throughput of a constant rate transmission based on either the buffer overflow probability or the channel outage probability is significantly limited. This consideration leads to improve the system performance by adapting the signalling rate and optimising the transmission strategy.

The following figures show the difference of process gains and data loss probability for the constant rate transmission and the adaptive rate transmission obtained by the MDP analysis. The optimal policy obtained by the MDP analysis adapts the signalling rate to the state of buffer and capacity at every time frame T_{f}, forcing the system to work at high values of signalling rate only when the buffer is full.

Fig. 3 shows the benefits in terms of system throughput obtained by adapting the signalling rate to both the buffer occupancy and the channel capacity state. We assume that a certain maximum outage probability P_{out} needs to be met for MIMO transmission. Dashed lines show the system throughput for constant rate transmission using the highest signalling rate allowed by the P_{out} requirement. As expected, the system throughput decays when the packet arrival rate is allowed to increase. The MDP optimisation ensures that, at each possible state, the best signalling rate is chosen maximising the system gain: the average system throughput reaches a stable value also for high arrival rates. The average system throughput obtained with the MDP transmission strategy does not show the strong decline in gain which affects the constant rate transmission in Fig. 2 for high packet arrival rates.

The study of the packet loss rate confirms the better performance obtained by adapting the transmission rate based on the channel/buffer state. Fig. 4 shows the average packet loss rate resulting from the MDP transmission strategy. The dashed lines represent the packet loss rate for constant transmission rates based on two different outage requirements. If no outage is allowed, the system is penalised at high arrival rates when the low transmission rate cannot support the system limited buffer. On the other

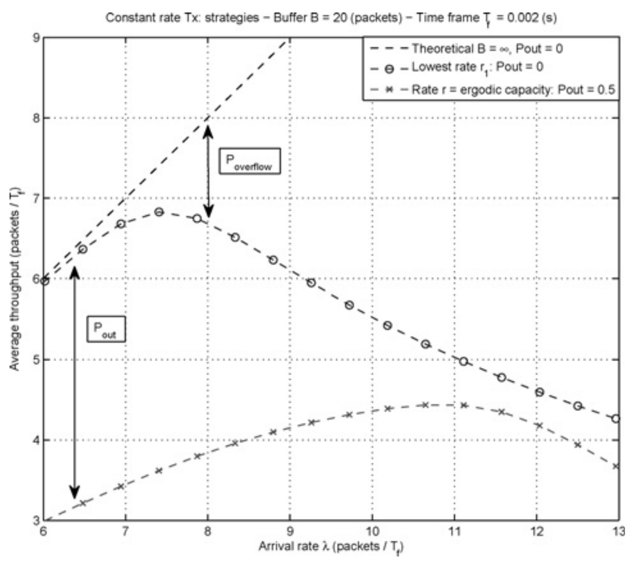

Figure 2 *Average system throughput for constant rate transmission*

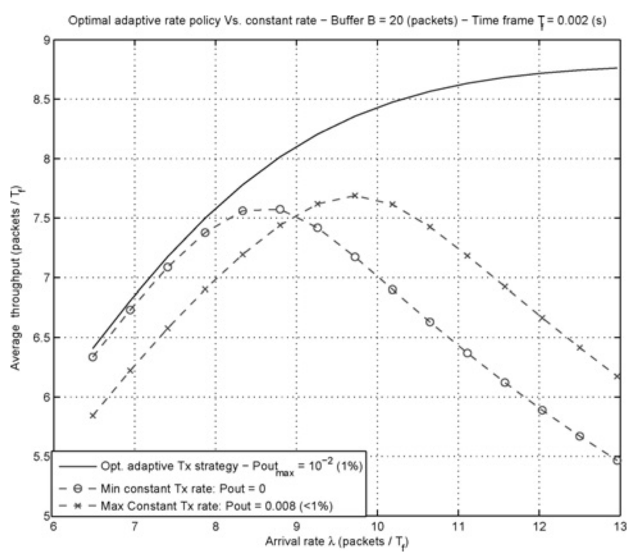

Figure 3 *Average system throughput for different arrival rates*

doi: 10.1049/iet-com.2009.0140

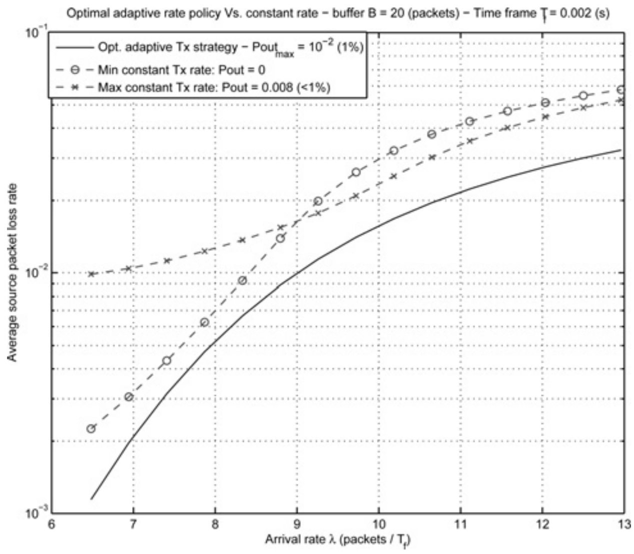

Figure 4 *Average packet loss rate for different arrival rates*

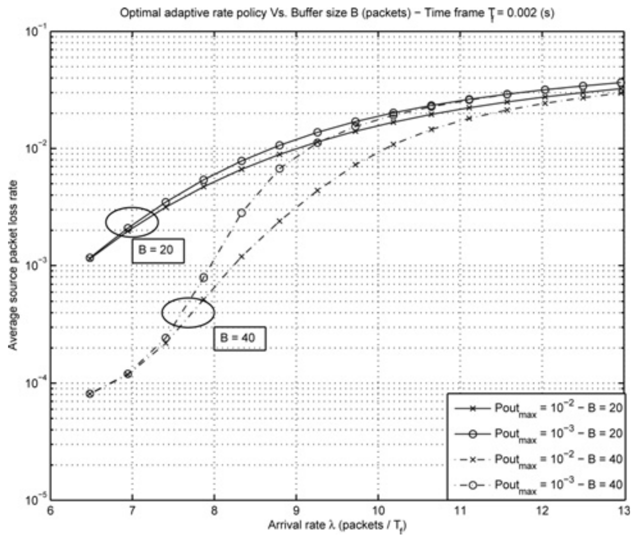

Figure 6 *Average packet loss rate*

hand, allowing a little outage on the channel helps in reducing the packet loss for high arrival rates. The joint consideration of the two behaviours leads to a better transmission strategy, which shows a significantly lower average packet loss rate. The strategy obtained by the MDP results to be the optimal one in terms of average packet loss rate: the adaptive rate transmission, limiting the allowed signalling rates to match a specific outage requirement, works with a lower rate but results in overall better performance described by the controlled increase of the packet loss rate.

A target outage probability P_{out} can be specified as a system requirement. A target P_{out} will lead to the choice of a suitable set of signalling rate to match the requirement. Fig. 5 shows the average system throughput with a target $P_{out} = 10^{-3}$. A

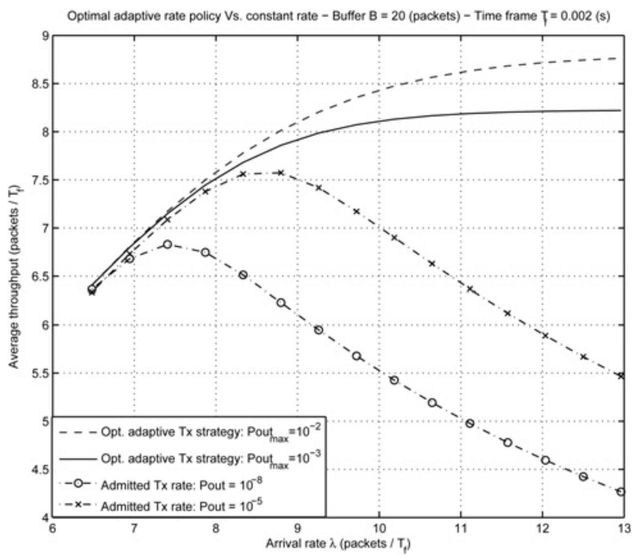

Figure 5 *Average system throughput with target $P_{out} = 10^{-3}$*

more strict requirement in terms of outage imposes lower signalling rates which turns into a lower gain compared with the optimal transmission strategy with $P_{out} = 10^{-2}$ (dashed line). The number of packets per time frame correctly received is lower and an increase in the buffer overflow probability is expected. The adaptive strategy gain is always higher than the constant rate transmission with the same suitable rates, confirming the efficiency of the MDP optimisation.

Finally, when imposing a target P_{out}, the system buffer can be designed to match a specific packet loss rate. Fig. 6 shows the average packet loss rate for different arrival rates, P_{out} and buffer sizes B. The system buffer can be specified to achieve a specific requirement on the packet loss rate. Moreover, according to the system condition (i.e. packet arrival rate), the system performance can be enhanced by allowing a higher P_{out} instead of changing the buffer size. The analysis presented helps to discuss the best system design strategy.

9 Conclusions

In this study, we have proposed a two-dimensional performance optimisation technique for data transmission over a MIMO wireless channel. The proposed scheme considers the evolving MIMO wireless channel capacity as well as the buffer occupancy in order to achieve the optimum wireless system throughput.

The MIMO channel capacity was modelled as a FSMC and the transition probabilities were analytically derived for adjacent and also non-adjacent states. The problem of the optimal transmission signalling rate was resolved referring to the MDP theory. The optimal signalling rate results from both the predicted capacity in the future time frame and the number of data packets in the buffer, in order to minimise the probability of data loss both from the buffer

IET Commun., 2010, Vol. 4, Iss. 6, pp. 683–696
doi: 10.1049/iet-com.2009.0140

(overflow) and the channel (outage). Assuming a frame duration T_f, the signalling rate is adapted at every frame according to both the buffer state and the capacity transition behaviour. The MDP optimisation derives the optimal transmission rate for every possible state of buffer and capacity, obtaining the optimal transmission strategy.

The joint consideration of P_{out} and $P_{overflow}$ is necessary to evaluate the system performance and the MDP analysis was an efficient method to maximise the system throughput. Numerical analysis indicates that using the derived optimal strategy increases significantly the system performance, both in throughput per frame and in packet loss rate, especially when the system experiences high data loads.

10 Acknowledgment

The authors thank Tricia Willink and Kareem Baddour, CRC, Canada, for the helpful discussions.

11 References

[1] ZEMEN T., MECKLENBRAUKER C.: 'Time-variant channel estimation using discrete prolate spheroidal sequences', *IEEE Trans. Signal Process.*, 2005, **53**, (9), pp. 3597–3607

[2] SEJDIĆ E., LUCCINI M., PRIMAK S., BADDOUR K., WILLINK T.: 'Channel estimation using DPSS based frames'. Proc. IEEE ICASSP., March–April 2008

[3] HOCHWALD B.M., MARZETTA T.L., TAROKH V.: 'Multiple-antenna channel hardening and its implications for rate feedback and scheduling', *IEEE Trans. Inf. Theory*, 2004, **50**, (9), pp. 1893–1909

[4] XIAODONG C., GIANNAKIS G.B.: 'Adaptive PSAM accounting for channel estimation and prediction errors', *IEEE Trans. Wirel. Commun.*, 2005, **4**, (1), pp. 246–256

[5] LIU Q., ZHOU S., GIANNAKIS G.B.: 'Queuing with adaptive modulation and coding over wireless links: cross-layer analysis and design', *IEEE Trans. Wirel. Commun.*, 2005, **4**, (3), pp. 1142–1153

[6] ZORZI M., RAO R., MILSTEIN L.: 'Error statistics in data transmission over fading channels', *IEEE Trans. Commun.*, 1998, **46**, (11), pp. 1468–1477

[7] TAN C.C., BEAULIEU N.C.: 'On first-order markov modeling for the rayleigh fading channel', *IEEE Trans. Commun.*, 2000, **48**, (12), pp. 2032–2040

[8] BIGLIERI E., CALDERBANK R., CONSTANTINIDES A., GOLDSMITH A., PAULRAJ A., POOR H.: 'MIMO wireless communications' (Cambridge University Press, Cambridge, 2007)

[9] TULINO A., VERDÚ S.: 'Random matrix theory and wireless communications', *Found. Trends Commun. Inf. Theory*, 2004, **1**, pp. 1–182

[10] PROAKIS J.: 'Digital communications' (McGraw-Hill, New York, 2001, 4th edn.)

[11] DEBBAH M., MULLER R.R.: 'MIMO channel modeling and the principle of maximum entropy', *IEEE Trans. on Information Theory*, 2005, **51**, (5), pp. 1667–1690

[12] LOYKA S., LEVIN G.: 'Diversity-multiplexing tradeoff via asymptotic analysis of large MIMO systems'. Proc. IEEE ISIT, Nice, France, June 2007

[13] TELATAR I.E.: 'Capacity on multi-antenna Gaussian channels', *Eur. Trans. Telecommun.*, 1999, **10**, (6), pp. 585–595

[14] HOGSTAD B.O.: 'Mulitple-input mulitple-output fading channel models and their capacity'. PhD dissertation, Department of Communication Technology, Faculty of Engineering and Science, Aalborg University, Aalborg, Denmark, 2008

[15] SMITH P.J., ROY S., SHAFI M.: 'Capacity of MIMO systems with semicorrelated flat fading', *IEEE Trans. Inf. Theory*, 2003, **49**, (10), pp. 2781–2788

[16] VAIHUNTHAN S., HAYKIN S., SELLATHURAI M.: 'MIMO channel capacity modeling using Markov models'. Proc. IEEE VTC Spring, 2005

[17] GILBERT E.: 'Capacity of a burst noise channel', *Bell Syst. Technol. J.*, 1960, **39**, pp. 1253–1265

[18] ELLIOTT E.O.: 'Estimates of error rates for codes on a burst-noise channel', *Bell Syst. Technol. J.*, 1963, **42**, pp. 1977–1997

[19] KANAL L.N., SASTRY A.R.K.: 'Models for channels with memory and their applications to error control', *Proc. IEEE*, 1978, **66**, (7), pp. 724–744

[20] WANG H.S., MOAYERI N.: 'Finite-state Markov channel – a useful model for radio communication channels', *IEEE Trans. Veh. Technol.*, 1995, **44**, (1), pp. 163–171

[21] SUBOTIC V.: 'Error performance evaluation of wireless communication systems in the presence of non-rayleigh channel fading'. PhD dissertation, Department of Electrical and Computer Engineering, Faculty of Engineering, The University of Western Ontario, London, Ontario, Canada, 2005

[22] PRIMAK S., KONTOROVICH V., LYANDRES V.: 'Stochastic methods and their applications to communications: SDE approach' (John Wiley and Sons, Chichester, 2004)

[23] BERTSEKAS D., GALLAGER R.: 'Data networks' (Prentice-Hall, Saddle River, NJ, 1987)

doi: 10.1049/iet-com.2009.0140

[24] PAPOULIS A.: 'Probability, random variables and stochastic processes' (McGraw-Hill, Boston, MA, 1991, 3rd edn.)

[25] HOWARD R.A.: 'Dynamic programming and Markov processes' (The MIT Press, Cambridge, MA, 1960)

[26] LUCCINI M., PRIMAK S., SHAMI A.: 'On the capacity of mimo channels and its effect on network performance'. Proc. IEEE ICTTA, May 2008

[27] BAI X., SHAMI A.: 'Two dimensional cross-layer optimization for packet transmission over fading channel', *IEEE Trans. Wirel. Commun.*, 2008, **7**, (10), pp. 3813–3822

[28] BAI X., SHAMI A.: 'A new perspective of cross-layer optimization for wireless communication over fading channel'. Proc. IEEE ICCCN, 13–16 August 2007, pp. 1–6

doi: 10.1049/iet-com.2009.0140