Cross-layer modeling of wireless channels for data-link and IP layer performance evaluation

Dmitri Moltchanov*, Yevgeni Koucheryavy, Jarmo Harju

Networks and Protocols Group, Institute of Communications Engineering, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland

Received 1 August 2005; accepted 1 August 2005

Abstract

To provide a tool for performance evaluation of IP-based delay- and loss-sensitive applications running over wireless channels we propose a novel cross-layer wireless channel modeling approach. We firstly develop simple and computationally efficient wireless channel modeling algorithm. For this purpose we adopt the special solution of the inverse eigenvalue problem and show that its complexity significantly decreases when the time-series is covariance stationary two-valued in nature. Our model explicitly takes into account autocorrelation and distributional properties of empirical data. Then, we extend this model to IP layer using the cross-layer mappings. The resulting model is represented by the IP packet error process and reflects memory properties of initial bit error process. We show that our approach allows to get accurate estimators of IP packet error probabilities in presence of FEC at the data-link layer eliminating the need for computationally expensive time-consuming bit level simulations. It also provides a way to choose the required correction capabilities of FEC codes resulting in best possible performance at the data-link and IP layers.

© 2005 Published by Elsevier B.V.

Keywords: Cross-layer wireless channel modeling; Performance evaluation

1. Introduction

While next generation (NG) mobile systems are not completely defined, there is a common agreement that they will rely on IP protocol as a consistent end-to-end transport technology. The motivation is to introduce a unified service platform for future 'mobile Internet' known as 'NG All-IP' mobile systems.

In addition to wireless access to the Internet, NG All-IP networks must satisfy requirements of QoS-aware applications. This is an inherent problem for many service types even in fixed IP networks. Wireless and mobility issues add their own problems on top of this inherent IP flaw. Time-varying nature of wireless channels, teletraffic and mobility issues must be addressed before wireless multimedia services will be successfully deployed.

Abstract

To provide a tool for performance evaluation of IP-based delay- and loss-sensitive applications running over wireless channels we propose a novel cross-layer wireless channel modeling approach. We firstly develop simple and computationally efficient wireless channel modeling algorithm. For this purpose we adopt the special solution of the inverse eigenvalue problem and show that its complexity significantly decreases when the time-series is covariance stationary two-valued in nature. Our model explicitly takes into account autocorrelation and distributional properties of empirical data. Then, we extend this model to IP layer using the cross-layer mappings. The resulting model is represented by the IP packet error process and reflects memory properties of initial bit error process. We show that our approach allows to get accurate estimators of IP packet error probabilities in presence of FEC at the data-link layer eliminating the need for computationally expensive time-consuming bit level simulations. It also provides a way to choose the required correction capabilities of FEC codes resulting in best possible performance at the data-link and IP layers.

© 2005 Published by Elsevier B.V.

Keywords: Cross-layer wireless channel modeling; Performance evaluation

1. Introduction

While next generation (NG) mobile systems are not completely defined, there is a common agreement that they will rely on IP protocol as a consistent end-to-end transport technology. The motivation is to introduce a unified service platform for future 'mobile Internet' known as 'NG All-IP' mobile systems.

In addition to wireless access to the Internet, NG All-IP networks must satisfy requirements of QoS-aware applications. This is an inherent problem for many service types even in fixed IP networks. Wireless and mobility issues add their own problems on top of this inherent IP flaw. Time-varying nature of wireless channels, teletraffic and mobility issues must be addressed before wireless multimedia services will be successfully deployed.

To evaluate performance of applications in fixed networks it is usually sufficient to estimate performance degradation caused by packet forwarding procedures. Dealing with wireless networks we also must take into account performance degradation caused by bit errors of wireless channels. It has completely different nature compared to what we dealt in fixed networks and may contribute a lot in end-to-end performance degradation.

Data-link layer techniques such as forward error correction (FEC), automatic repeat request (ARQ), and combination of them (hybrid ARQ/FEC) may allow to recover from losses caused by bit errors. ARQ techniques eliminate the influence of errors allowing to retransmit incorrectly received frames. To notify the sender about the requested frame, ARQ protocols require a feedback channel. When the channel conditions are relatively 'bad' ARQ may introduce significant delay and jitter that may not always be tolerable for multimedia applications such as streaming video or real-time two-way voice communications. On the other hand, FEC procedures try to eliminate influence of errors in advance introducing error correction redundancy. This redundancy is efficiently exploited at the receiver to recover from bit errors. The major advantage of FEC
techniques for delay-constrained applications is that they do not introduce long delays allowing some information to be lost. Depending on a particular implementation, FEC capabilities can be used at both physical and data-link layers. In this paper, we concentrate our attention on delay-sensitive applications for which only FEC techniques are feasible.

To predict performance of information transmission over wireless channels, models representing the channel state in terms of signal-to-noise ratio (SNR) or bit errors are often used (see [1–5] among others and [6,7] for review). However, those models cannot be directly used in performance evaluation studies and must be properly extended to the data-link and IP layers at which QoS parameters are usually defined. For such an extension to be accurate, we must take into account specific peculiarities of underlying layers including statistical properties of involved stochastic processes, error concealment techniques at physical and data-link layers, segmentation procedures between different layers, etc. An adequate performance evaluation model of the wireless channel at the layer of interest is a complex cross-layer function of underlying layers and propagation characteristics. These models along with traffic models have to be further applied to predict QoS expectations experienced by applications running over the wireless channels.

In this paper, we propose a novel cross-layer performance evaluation framework. Firstly, based on statistical analysis we show that the autocorrelation function (ACF) of the bit error process exhibits nearly geometrical behavior and approximate such a behavior using a single geometrically distributed component. To find suitable Markovian model approximating a given ACF, we formulate and solve the inverse eigenvalue problem, i.e. we find a Markov process such that its transition probability matrix posses a predefined eigenvalue approximating the behavior of empirical ACF. We also show that there is unique Markov modulated process approximating distributional and autocorrelation properties of the bit error trace. The associated fitting algorithm is simple, fast, and computationally efficient. Then, we extend the bit error model to IP layer using the cross-layer mappings. The resulting model is represented by the IP packet error process, preserves memory properties of the bit error process and captures specific peculiarities of protocols at layers below IP. We show that our approach allows to get accurate estimators of IP packet error probabilities and efficiently evaluate the required rate of FEC codes to provide best possible QoS parameters at the data-link and IP layers.

Our paper is organized as follows. In Section 2 we overview propagation characteristics of wireless channels and models used to capture them. In Section 3 we propose the wireless channel modeling algorithm at the physical layer. Cross-layer extension to IP layer is introduced in Section 4, and associated performance evaluation framework is presented in Section 5. In Section 6 numerical examples are considered. Conclusions are drawn in the last section.

2. Propagation properties of wireless channels

The propagation path between the transmitter and a receiver may vary from line-of-sight (LOS) to very complex ones due to diffraction, reflection and scattering. To predict performance of wireless channels, propagation models are often used. The received signal strength predicted by propagation models is then related to SNR considering interference and thermal noise [8]. We usually distinguish between the large-scale and small-scale propagation models [9].

2.1. Large-scale propagation models

When a mobile user moves away from the transmitter over large distances the received local average signal strength (RLASS) gradually decreases. This signal strength can be predicted using large-scale propagation models. There are a number of large-scale propagation models developed to date (see [10–12] among others). Such models do not take into account rapid fluctuations of the received signal strength. As a result, they cannot be effectively used in performance evaluation studies. Indeed, the instantaneous received signal strength may vary quickly resulting in fast changes in the performance parameters perceived at higher layers.

2.2. Small-scale propagation models

When a mobile user moves over short distances the instantaneous received signal strength may vary. The reason is that the received signal is a sum of many components coming from different directions due to reflection, diffraction and scattering. Since phases, amplitudes and arriving times of components are random, the resulting signal may vary quickly and significantly. To capture small-scale propagation characteristics we distinguish between two cases: there is a LOS between the transmitter and a receiver and there is no LOS. In presence of non-fading component the small-scale propagation envelop distribution is Rician. As the dominant component fades away due to shadowing by obstacles the small-scale propagation envelop distribution degenerates to Rayleigh one.

The small-scale propagation models capture characteristics of wireless channel on a finer granularity than large-scale ones [1,2,6,13]. Additionally, these models implicitly take into account movement of users over short travel distances [9].

2.3. Markovian small-scale propagation model

Assume a discrete-time environment, i.e. time axis is slotted, the slot duration is constant and given by \( \Delta t = (t_{i+1} - t_i), i = 0,1, \ldots \). We choose \( \Delta t \) such that it equals to the time to transmit a single symbol at the wireless channel.
Small-scale propagation characteristics are often represented by the doubly-stochastic process \( L(n) \) modulated by the discrete-time Markov chain \( S_L(n) \), each state of which is associated with conditional probability distribution function of the received signal strength \([1,2,6]\). The underlying modulation allows to take into account autocorrelation properties of the signal strength process. Since, it is allowed for the Markov process \( S_L(n), n=0,1,… \) to change its state in every time slot, two successive bits may experience completely different propagation characteristics.

An illustration of such a model is shown in Fig. 1, where states are associated with conditional distribution functions \( F_L(k\Delta f | i) \). In what follows, we will drop dependence on \( \Delta f \) as it is implicitly assumed. Let \( D_L \) and \( \pi_L \) be the one-step transition probability matrix and the stationary probability vector of \( \{ S_L(n), n=0,1,… \} \), respectively. Parameters \( M \), \( D_L \), and \( F_L(k\Delta f | i) \), are estimated from statistical data \([6]\).

### 2.4. SNR process

For a model to be useful, in addition to the received signal strength we have to take into account the level of noise and consider the signal-to-noise ratio (SNR), instead. In the uplink direction, the interference at the receiver is given by \( I = (W + U) \), where \( W \) the thermal noise, \( U = \sum U(1) \) is received signal strength from \( L \) interfering mobile terminals from neighboring interfering cells. \( W \) is usually assumed to be constant. \( U \) is stochastic, depends on the distance and propagation path between the receiver and transmitters in interfering cells. In what follows, we assume that a given cellular network is well-provisioned, i.e. the best possible frequency plan is used. Due to these assumptions it is fair to assume that interference, \( I \), is small and constant. Using the value of interference the best possible frequency plan is used. Due to these assumptions it is fair to assume that interference, \( I \), is small and constant. Using the value of interference the best possible frequency plan is used. Due to these assumptions it is fair to assume that interference, \( I \), is small and constant.

### 2.5. Bit error process

If the wireless channel model is defined in terms of the SNR process, it should be further extended to bit error process. Consider an arbitrary state \( i \) of the Markov chain \( \{ S_L(n), n=0,1,… \} \) associated with the conditional probability distribution function \( F_L(k\Delta f | i) \), \( k=1,2,…,N \), of the SNR. Assume that the probability of a single bit error is the deterministic function of the SNR value. According to this assumption, all values of \( F_L(k\Delta f | i) \) that are less or equal to a computed value of the so-called bit error threshold \( B_T \) lead to incorrect bit reception. Those values which are greater than \( B_T \) do not cause bit error. Therefore, each state \( i, i=1,2,…,M \) of the Markov process \( \{ S_L(n), n=0,1,… \} \) could be associated with the following bit error probability \( P_{E,i} \):

\[
P_{E,i} = \Pr(E(n) = 1 | S_L(n) = i) = \sum_{k=1}^{B_T} \Pr(L(n) = k|S_L(n) = i),
\]

where \( \{ E(n), n=0,1,… \} \) is the bit error process for which ‘1’ denotes an incorrectly received bit, ‘0’ refers to a correctly received bit. \( \{ S_L(n), n=0,1,… \} \) is the underlying Markov chain of \( \{ E(n), n=0,1,… \} \) is the underlying Markov chain of \( \{ E(n), n=0,1,… \} \). Note, that \( \{ S_L(n), n=0,1,… \} \) and \( \{ S_E(n), n=0,1,… \} \) are defined on the same state-space, and \( \pi_E = \pi_L, D_E = D_L \), where \( D_E \) and \( \pi_E \) are transition probability matrix and stationary distribution vector of \( \{ S_E(n), n=0,1,… \} \), respectively.

In general, \( B_T \) must be estimated based on a modulation scheme, type of the receiver and other specific features utilized at a given wireless channel \([9]\). Unfortunately, analytical expressions for (1) are available for simple modulation schemes only \([6,14]\). Therefore, models of the SNR process and the received signal strength process are not convenient for performance evaluation purposes requiring extensive field measurements of empirical relationship between the SNR value and the bit error probability.

### 3. Bit error model of the wireless channel

It was shown that the bit error process of the wireless channels can be modeled by the Markov modulated process \([15–17]\). Such a model provides a useful trade-off between complexity of the model and accuracy of fitting to statistical data. However, parameters matching algorithms developed for this model do not take into account second-order properties of statistical data in terms of their ACF. It may lead to incorrect representation of the memory of bit error process.

In what follows, we propose a model for bit error traces. In this study we use IEEE 802.11b bit error traces available...
from [18]. Setup and background information related to collection of traces are given in [19,20]. We however, do not restrict our attention to wireless local area networks. These traces are used for illustrative purposes. We found our traces to be partially covariance stationary and explicitly capture statistical characteristics of covariance stationary segments including bit error rate and empirical ACF. The proposed approach is not only applicable for bit error traces but can be used to model frame error traces of wireless channels [21] as long as empirical ACFs can be sufficiently well approximated by a single geometrically distributed component.

3.1. Model of bit errors

A bit error trace is a sequence of successive events of correct and incorrect bit receptions. To use theory of stochastic processes, we define a bit error trace as a sequence of RVs. We assume that ‘1’ represents an incorrectly received bit and ‘0’ denotes correctly received bit. Successive realizations of this RV compose the bit error trace \( W_G(n), n = 0, 1, \ldots \), \( W_G(n) = \{0, 1\} \). A number of studies suggested that this process can be treated as, at least, partially covariance stationary one [19,20,22]. Using these properties, one can compute all important statistics using a single (a part of) bit error trace. We use \( f_g(1) \) to denote probability of bit error as seen by time-averages. Correspondingly, \((1 − f_g(1))\) denotes probability of correct bit reception.

Due to a binary nature of covariance stationary bit error traces, distributional properties are completely determined by the mean of the process and ACF. We suggest that this process is characterized by nearly geometrical distribution of ACF for small lags. As an example, empirical ACFs of two bit error traces observed on IEEE 802.11b wireless channel are shown in Fig. 2.

To model a bit error trace we propose to use the discrete-time Markov modulated process with two states of the modulating Markov chain. This process is a special case of D-MAP, known as discrete-time Markovian arrival process (D-MAP). According to D-MAP, value of the process is modulated by the discrete-time Markov chain \( \{S(n), n = 0, 1, \ldots\\}, S(n) = \{1, 2\} \). Let \( D \) be its transition probability matrix. We define D-MAP using two matrices, \( D(0) \) and \( D(1) \), containing transition probabilities from state to state associated with 0 or 1, respectively. Thus, for each pair of states \( D_{ij}(k) = \lim_{n→∞} \Pr [W(n) = k, S(n) = j | S(n − 1) = i] \), \( k = 0, 1 \) are conditional probability distribution functions. Without loss of generality we do not allow those arrivals when the underlying Markov chain changes its state from \( i \) to \( j, i ≠ j \).

Let \( \{W(n), n = 0, 1, \ldots\}, W(n) = \{0, 1\} \), be the D-MAP process modeling the bit error trace \( W_G(n), n = 0, 1, \ldots \), \( W_G(n) = \{0, 1\} \), and let \( \{G(n), n = 0, 1, \ldots\} \) be its mean process. Mean process of D-MAP with two-states of the modulating Markov chain is characterized by ACF distributed according to geometrical law with a single component determining its decay: \( R_G(l) = φ \lambda^l \), where \( φ \) is the variance of the process, \( λ \) is the non-unit eigenvalue. Such a behavior may produce fair approximation of the statistical ACF.

3.2. The matching algorithm

To approximate empirical ACF, we propose to minimize the error of approximation \( γ \) by varying the value of coefficient \( λ \) according to the following expression

\[
g = \frac{1}{T} \sum_{i=1}^{T} \left( \frac{K_E(i) − \lambda^i}{K_E(i)} \right)^2, \quad i = 1, 2, \ldots, \tag{2}
\]

where \( i_0 \) is the lag up to which the ACF is statistically significant, and \( K_E(i) \) is the value of empirical ACF for lag \( i \). Note that \( λ = K_E(1) \) and \( γ = 0 \) for \( i_0 = 1 \).

The construction of the Markov modulated process involves the inverse eigenvalue problem. The general solution of this problem does not exist. However, it is possible to solve it when some limitations on form of eigenvalues are set. Our limitation is that the non-unit eigenvalue, \( λ \), must be located in \((0,1)\) fraction of 0X axis. Since, all eigenvalues of transition probability matrix of irreducible aperiodic Markov chain are located in \([-1,1]\).
fraction of 0X axis, \( \lambda = [-1, 1] \). Finally, \( \lambda = (0, 1] \) must be fulfilled by the solution of the inverse eigenvalue problem.

Let us introduce the mean process of \( \{W(n), n=0,1,\ldots\} \), \( \{G(n), n=0,1,\ldots\} \). Stochastic properties of \( \{G(n), n=0,1,\ldots\} \) are completely characterized by a triplet \((E[W], \phi, \lambda)\). Values of \((E[W], \phi, \lambda)\) are related to parameters of \( \{G(n), n=0,1,\ldots\} \) using the following equations

\[
\begin{align*}
E[W] &= \alpha G_2 + \beta G_1 \\
\lambda &= 1 - \alpha - \beta \\
\phi &= \alpha \beta \left( \frac{G_1 - G_2}{\alpha + \beta} \right)^2,
\end{align*}
\]

where \( G_1 \) and \( G_2 \) are means in states 1 and 2 respectively, \( \alpha \) and \( \beta \) are probabilities of transition from state 1 to state 2 and from state 2 to state 1, respectively.

Observing (3) one may note that in order to completely parameterize the mean process of \( \{W(n), n=0,1,\ldots\} \), we must provide four parameters \((G_1, G_2, \alpha, \beta)\). If one choose \( G_1 \) as a free variable with constraint \( G_1 \leq E[W] \) to satisfy \( \lambda = (0, 1] \), one can determine \( G_2, \alpha, \) and \( \beta \) from the next set of equations

\[
\begin{align*}
G_2 &= \frac{\phi E}{E[W]} - G_1 \\
\alpha &= \frac{(1 - \lambda)(E[W] - G_1)}{G_2 - G_1} \\
\beta &= \frac{(1 - \lambda)(G_1 - E[W])]}{G_2 - G_1}
\end{align*}
\]

Let us now identify a distinctive feature of the proposed matching algorithm that uniquely identifies the process we are looking for. Observing the first equation in (4) one may conclude that there are an infinite number of processes matching \((E[W], \phi, \lambda)\). However, there is an additional restriction on the value of \( G_1 \), Consider the first equation in (4), and rewrite it using a well known result for covariance stationary observations \( \phi E = E[W] = (E[W])^2 \)

\[
G_2 = \frac{E[W] - E[W] G_1}{E[W] - G_1},
\]

To represent the bit error trace, \( \{W(n), n=0,1,\ldots\} \) must be defined on the binary state space \( W(n) = \{0,1\} \). Thus, the value of \( G_2 \) must be equal to or less than one for any state of \( S(n), n=0,1,\ldots \). To identify what values of \( G_1 \) must be chosen to satisfy \( 0 < G_1 < 1, i=1,2 \), one have to consider two extreme cases

\[
\lim_{G_1 \rightarrow E[W]} G_2 = \infty, \quad \lim_{G_1 \rightarrow 0} G_2 = 1.
\]

Observing (6) one may note that \( G_1 = 0, G_2 = 1 \), gives the only process matching \((E[W], \phi, \lambda)\). Hence, given \( G_1 = 0, G_2 = 1 \), the only parameters we have to determine to match the mean process of the bit error trace are \( \alpha \) and \( \beta \)

\[
\begin{align*}
\alpha &= (1 - \lambda)E[W] \\
\beta &= (1 - \lambda)(1 - E[W])
\end{align*}
\]

4. Cross-layer wireless channel modeling

The bit error model defined in the previous section cannot be directly used in performance evaluation studies and must be properly extended to the data-link and IP layers at which QoS parameters are usually defined. To do so, we have to take into account specific peculiarities of underlying layers including statistical properties of involved stochastic processes, error concealment techniques at physical and data-link layers, segmentation procedures between different layers etc.

Cross-layer mobility-dependent design of the wireless channel model at the IP layer is illustrated in Fig. 3, where each layer is associated with appropriate stochastic process modeling the reception of protocol data units (PDU). Stochastic process at each layer is a cross-layer function of specific peculiarities of underlying layers.

In what follows, we denote the developed bit error model by \( \{E(n), n=0,1,\ldots\} \), \( E(n) = [0,1] \) modulated by the underlying Markov chain \( \{S(n), n=0,1,\ldots\} \), \( S(n) = \{1, 2\} \). We also denote by \( d_{E,n}(k) = \Pr\{E(n) = k | S(n) = j, S(n-1) = i\} \), \( k = 0,1 \), the transition probability of state \( i \) to state \( j \) with correct \( (k=0) \) and incorrect \( (k=1) \) bit reception, respectively. These probabilities can be represented in a compact form using matrices \( D_E(1) \) and \( D_E(0) \) such that \( D_E(1) + D_E(0) = D_E \). Note that in our case the state from which a transition occurs completely determines the bit error probability. The state to which a transition occurs is used for convenience of matrix notation.

4.1. FEC at physical or data-link layer

Assume that the length of frames is constant and equals to \( m \) bits. The sequence of consecutively transmitted bits, denoted by gray rectangles, is shown in Fig. 4, where \( (l-1), l, (l+1) \) denote time intervals whose length equals to the time to transmit a single frame; \( i, j, k \), denote the state of the Markov chain \( \{S(n), n=0,1,\ldots\} \) in the beginning of these intervals.

Assume that a kind of FEC redundancy is added to the original stream of bits at the physical layer. To proceed we

\[
\begin{align*}
\text{Network layer} & \quad \{P(n) = P(m), n=0,1,\ldots\} \\
\text{Data-link layer} & \quad \{F(n) = F(E(n), \ldots, E(n+m)) \}, n=0,1,\ldots\} \\
\text{Physical layer} & \quad \{E(n), n=0,1,\ldots\}
\end{align*}
\]

Fig. 3. Cross-layer modeling concept.
also assume that the stream of bits (with redundancy bits added to them) are divided into frames of \( m \) bits in length, among which up to \( l \) bits can be corrected. Block codes are directly parameterized by \((m,l)\). To determine \((m,l)\) for convolutional codes, approximations should be used. The proposed interpretation is valid for both physical and data-link layer.

### 4.2. Frame error process without FEC

Consider the stochastic process \( \{N(l), l=0,1,\ldots\} \), \( N(l) = \{0,1,\ldots,m\} \), describing the number of incorrectly received bits in consecutive bit patterns of length \( m \). This process is doubly stochastic, modulated by the underlying Markov chain \( \{S_{\ell}(l), l=0,1,\ldots\} \). Process \( \{N(l), l=0,1,\ldots\} \) can be completely parameterized via parameters of the bit error process.

To completely parameterize \( \{N(l), l=0,1,\ldots\} \) we have to determine transition probabilities between states \( i \) and \( j \) of the modulating Markov chain \( \{S_{\ell}(n), n=0,1,\ldots\} \) with exactly \( k, k=0,1,\ldots,m \), incorrectly received bits. Let us denote the probability of going from the state \( i \) to state \( j \) for the Markov chain \( \{S_{\ell}(l), l=0,1,\ldots\} \) with exactly \( k, k=0,1,\ldots,m \), incorrectly received bits in a bit pattern of length \( m \) by \( d_{N,j}(k) = \Pr(N(l)=j|S_{\ell}(l) = i, l = 1,2,\ldots,m) \). Let us denote the number of bit errors in consecutive frames can be related to the frame error process \( \{F(l), l=0,1,\ldots\} \), \( F(l) = \{0,1\} \), as follows:

\[
D_F(0) = \sum_{k=0}^{F_T-1} D_N(k), \\
D_F(1) = \sum_{k=F_T}^{m} D_N(k),
\]  

(9)

where \( F_T \) is the frame error threshold. Expressions (9) are interpreted as follows: if the number of incorrectly received bits in a frame is greater or equal to a computed value of the frame error threshold \( k > F_T - 1 \), frame is incorrectly received and \( F(l) = 1 \). Otherwise \( k < F_T \), it is correctly received and \( F(l) = 0 \).

Assume that FEC is not used at the data-link layer. It means that every time a frame contains at least one bit error, it is received incorrectly. Thus, the transition probability matrices (9) of the frame error process take the following form:

\[
D_F(0) = D_N(0), \\
D_F(1) = \sum_{k=1}^{m} D_N(k).
\]  

(10)

One may note that is not required to compute \( D_N(k), k=1,2,\ldots,m \) from (8). Alternatively to (8), \( D_F(1) \) can be found using the following relation

\[
D_F(1) = \left( \sum_{k=0}^{m} D_N(k) \right)^m - D_N(0) = D_E^m - D_E^0.
\]  

(11)

### 4.3. Frame error process with FEC

The frame error threshold \( F_T \) depends on FEC capabilities. Assume that the number of bit errors that can be corrected by a FEC code is \( l \). Then, \( F_T = (l+1) \) and the frame is incorrectly received when \( k > l \). Otherwise, it is correctly received. Thus, the transition probability matrices (9) take the following form:

\[
D_F(0) = \sum_{k=0}^{l} D_N(k), \\
D_F(1) = \sum_{k=l+1}^{m} D_N(k).
\]  

(12)
In (12) we only need expressions for \( D_M(k), k=0,1,\ldots,l \). Using the same reasoning as in (11), the expression for \( D_F(1) \) is simplified as below

\[
D_F(1) = D_E^m - \left( \sum_{k=0}^{l} D_N(k) \right)^m.
\]

(13)

The slot durations of \{N(t), l=0,1,\ldots\} and \{F(t), l=0,1,\ldots\} are the same, \( \Delta t' \), and related to the slot duration of the bit error process \{E(n), n=0,1,\ldots\} as \( \Delta t' = m\Delta t \).

4.4. IP packet error process

Assume that IP packet is segmented into \( z \) frames of equal size at the data-link layer. The sequence of consecutively transmitted frames, denoted by gray rectangles, is shown in Fig. 5, where \((h-1), h, (h+1)\) denote time intervals whose length equals to the time to transmit a single packet; \(k, i, j\), denote the state of the Markov chain \{\( S_p(n), n=0,1,\ldots \)\} in the beginning of intervals. Note that assumption of constant packet size does not restrict the generality of results as long as the only one traffic source is allowed to be active at any instant of time.

Consider stochastic process \{\( M(h), h=1,2,\ldots \), \( M(h)=0,1,\ldots,z \)\}, describing the number of incorrectly received frames in a consecutive frame patterns of length \( z \). This process is doubly-stochastic, modulated by the underlying Markovian chain \{\( S_m(h), h=0,1,\ldots \)\}, and can be completely defined via parameters of the frame error process.

To completely parameterize \{\( M(h), h=0,1,\ldots \)\} we have to determine transition probabilities between states \( i \) and \( j \) of the modulating Markov chain \{\( S_p(l), l=0,1,\ldots \)\} with exactly \( k, k=0,1,\ldots,z \), incorrectly received frames. Let us denote the probability of going from state \( i \) to state \( j \) for the Markov chain \{\( S_p(h), h=0,1,\ldots \)\} with exactly \( k, k=0,1,\ldots,z \), incorrectly received frames in a frame pattern of length \( z \) by \( d_{M,}\{i,j,k\}=Pr\{M(h)=k|S_p(h)=i\} \). Let the set of matrices \( D_M(k) \), \( k=0,1,\ldots,z \), contain these transition probabilities. Matrices \( D_M(k) \), \( k=0,1,\ldots,z \), can be found using \( D_F(k), k=0,1 \), of the frame error process (similarly to (8)) as shown below

\[
\begin{align*}
D_M(0) &= D_F(0) \\
D_M(1) &= \sum_{k=1}^{z} D_{F,k-1}D_F(1)D_M(k) \\
\vdots \\
D_M(z) &= D_F(1)
\end{align*}
\]

(14)

Fig. 5. Sequence of consecutively transmitted frames at the wireless channel.

Let us now introduce the packet error process \{\( P(h), h=0,1,\ldots \), \( P(h)=0,1 \)\}, where \( '0' \) indicates the correct reception of the packet, \( '1' \) denotes incorrect packet reception. Process \{\( P(h), h=0,1,\ldots \)\} is modulated by the underlying Markov chain \{\( S_p(h), h=0,1,\ldots \)\}. Note that the state space of \{\( S_p(h), h=0,1,\ldots \)\} and \{\( S_m(h), h=0,1,\ldots \)\} is the same. Let us denote the probability of going from state \( i \) to state \( j \) for the modulating Markov chain \{\( S_p(h), h=0,1,\ldots \)\} with exactly \( k, k=0,1,\ldots,z \), incorrectly received packets by \( d_{P,}\{i,j,k\}=Pr\{P(h)=k|S_p(h)=i\} \). These probabilities are then combined in matrices \( D_P(0) \) and \( D_P(1) \). Process \{\( M(h), h=0,1,\ldots \), \( M(h)=0,1,\ldots,z \)\}, describing the number of incorrectly received frames in consecutively transmitted packets can be related to the packet error process \{\( P(h), h=0,1,\ldots \)\} as follows

\[
D_P(0) = \sum_{k=0}^{P_T-1} D_M(k), \quad D_P(1) = \sum_{k=P_T}^{z} D_M(k).
\]

(15)

where \( P_T \) is the packet error threshold.

Expressions (15) are interpreted as follows: if the number of incorrectly received frames in a packet is greater or equal to a computed value of the packet error threshold \( (k > P_T-1) \), packet is incorrectly received and \( P(h)=1 \). Otherwise, it is correctly received and \( P(h)=0 \). Since no error correction procedures are defined for IP layer, \( P_T=1 \). That is, every time a packet contains at least one incorrectly received frame, the whole packet is received incorrectly

\[
D_P(0) = D_M(0), \quad D_P(1) = \sum_{k=1}^{z} D_M(k).
\]

(16)

One should note that in order to determine (16) only \( D_M(0) \) must be computed in (14). Expression for \( D_P(1) \) is given by

\[
D_P(1) = \left( \sum_{k=1}^{m} D_M(k) \right)^m - D_M(0) = (D_F(0) + D_F(1))^m - D_F^m(0).
\]

(17)

The slot durations of \{\( P(h), h=0,1,\ldots \)\} and \{\( M(h), h=0,1,\ldots \)\} are the same, \( \Delta t'' \), and related to the slot duration of the bit error process \{\( E(n), n=0,1,\ldots \)\} as \( \Delta t'' = mz\Delta t \).

4.5. Illustration of the proposed extension

An illustration of proposed cross-layer mapping is shown in Fig. 6, where time diagrams of \{\( E(n), n=0,1,\ldots \)\}, \{\( N(l), l=0,1,\ldots \)\}, \{\( F(l), l=0,1,\ldots \)\}, \{\( M(h), h=0,1,\ldots \)\}, \{\( P(h), h=0,1,\ldots \)\} are shown. Error thresholds \( F_T \) and \( P_T \) must estimated as outlined previously and then used to compute transition probability matrices of error processes at different layers. The resulting IP packet error process is classified as D-MAP.

One should note that the proposed cross-layer mapping limits memory of the bit error process at lag \( m \), where \( m \) is the length of the frame in bits. Additionally, memory of
the frame error process is also limited at lag $z$, where $z$ is the number of frames to which an IP packet is segmented. However, in practice, $m$ and $z$ are sufficiently large, allowing to accurately capture memory of the initial bit error process at layers, above physical. The threshold, $m_{\text{min}}$, at which the frame error model becomes valid can be computed using Fig. 7, where the function $y(m) = \lambda^m$ is plotted. Recall that the ACF of the bit error model is distributed as $R_{\text{e}}(i) = \phi^i$, where $\phi$ is the variance of the process, $\lambda$ is the non-unit eigenvalue. The same reasoning applies for IP packet error process.

5. Performance evaluation

To define models of incorrect reception of PDUs at different layers we implicitly assumed that appropriate PDUs are consecutively transmitted at corresponding layers. Hence, the IP packet error process is conditioned on the event of consecutive transmission of packets. Such a conditioning may not provide sufficient details regarding how the application actually performs over the wireless channel. Real packet arrival processes can be autocorrelated, leading to unexpected performance parameters at the IP layer even when the wireless channel conditions are the same.

5.1. Queuing model

The straightforward way to represent the packet transmission process over the wireless channel is to use the special case of $G_\alpha/G_\beta/1/K$ queuing system, where $G_\alpha$ is the packet arrival process, $G_\beta$ is the packet service process of the wireless channel, $K$ is the capacity of the system. Here, the service process is defined as times required to successfully transmit successive packets over the wireless channel.

It is known that both interarrival times of packets from the traffic sources and transmission times of packets can be autocorrelated. These features significantly complicate analysis of $G_\alpha/G_\beta/1/K$ queuing system even when both arrival and service processes can be accurately modeled by Markov modulated processes. Indeed, the theoretical background of queuing systems with autocorrelated arrival and service processes is not well-studied.

Consider now a class of preemptive-repeat priority systems with two Markovian arrival processes. We allow both processes to have arbitrary autocorrelation structure of homogenous Markovian type. Assume that the first arrival process represents the packet arrival process from the traffic source. To provide an adequate representation of unreliable transmission medium, we assume that the second arrival process is the packet error process of the wireless channel. Making this process to be high priority one, and allowing its arrivals to interrupt ongoing service of low priority arrivals, we assure that when an arrival occurs from this process, it immediately seizes the server for service. Since no error concealment techniques are defined for IP layer, a packet whose service is interrupted immediately leaves the system. Such a behavior is interpreted as an incorrect reception of
a packet from the traffic source, and the priority discipline is referred to as preemptive-repeat.

Analysis of queuing systems with priority discipline is still a challenging task. Among others, preemptive-repeat is probably most complicated priority discipline. However, a number of assumptions can be introduced to make the queuing model less complicated. In what follows, we limit our model to discrete-time environment and require each arrival from any arrival process to have a service time of one slot in duration. Since there can be at most one arrival from the arrival process representing the packet error process of the wireless channel, these arrivals do not wait for service, enter the service in the beginning of nearest slots, and, if observed in the system, are being served. We also have to ensure that all arrivals from the packet error process are accommodated by the system. Following such assumptions, it is no longer needed to require preemptive-repeat priority discipline. Since all arrivals occur simultaneously in batches, it is sufficient for such a queuing-system to have non-preemptive priority discipline.

We assume that the packet arrival process from the traffic source is modeled by D-MAP process. Since both processes are D-MAP in nature, the queuing system reduces to non-preemptive priority discipline. Since all arrivals occur simultaneously in the beginning of nearest slots, and, if observed in the system, are being served. We also have to ensure that all arrivals from the packet error process are accommodated by the system. Following such assumptions, it is no longer needed to require preemptive-repeat priority discipline. Since all arrivals occur simultaneously in batches, it is sufficient for such a queuing-system to have non-preemptive priority discipline.

When a constant bit rate channel (CBR) is assigned to a mobile terminal during the whole duration of a call, to estimate performance parameters of packet transmission process we can directly apply D-MAP + D-MAP/D/1/1 queuing system. Such a model provides an adequate representation of time-division multiple access (TDMA) organization of the wireless access. According to TDMA, each source has an unrestricted access to a circuit-switched channel, on which packet transmission is organized. Thus, there is no IP layer concurrent traffic competing for resources, and the only performance degradation stems from unreliable nature of the wireless channel.

The sample path of the model is shown in Fig. 8, where transmissions of packet and error arrivals are marked by grey and black rectangles, respectively. Numbers are used to identify packets. One may note that the ‘transmission of error packets’ mimics behavior of the lossy wireless channel. To study the packet transmission process over the wireless channel, we have to derive performance parameters of the packet arrival process in D-MAP + D-MAP/D/1/1 queuing system.

Note that \( \sum \text{D-MAP}/D/1/1 \) queuing model can be extended to capture effects of contention access over a CBR channel. This can be done by allowing additional arrival process to the queuing system representing the concurrent traffic. This extension is however out of the scope of this paper.

5.2. Loss performance

Due to a simple structure of the queuing system, loss analysis can be accomplished by taking into account the nature of arrival processes and using their superposition.

Let the RV \( Z \) be its PDF. Since only a single arrival is allowed from the packet arrival process to the queuing system representing the concurrent traffic. This event occurs when and only when arrivals from both processes occur simultaneously. Contrarily, an arrival from the packet arrival process is not lost when no arrivals occur from the packet error process. We have

\[
f_L(1) = \lim_{n \to \infty} \Pr\{W_A(n) = 1, W_P(n) = 1\}. \tag{18}
\]

Consider superposition of the packet arrival and packet error processes. Transition probability matrices of the underlying Markovian chain of such a superposition are given by

\[
\begin{align*}
D_S(2) &= D_A(1) \otimes D_P(1) \\
D_S(1) &= D_A(1) \otimes D_P(0) + D_A(0) \otimes D_P(1) \tag{19} \\
D_S(0) &= D_A(0) \otimes D_P(0)
\end{align*}
\]

The state-space of the modulating Markov chain, \( \{S_A(n), n = 0,1,\ldots\} \), of the new model is given by

\[
S_A(n) = (1, 1, (1, 2), \ldots, (1, M_A), (2, 1), (2, 2), \ldots, (2, M_A)), \tag{20}
\]

where \( M_A \) is the number of states of the modulating Markov chain of the packet arrival process. Let us denote the steady-state vector of \( \{S_A(n), n = 0,1,\ldots\} \) by \( \pi_S = (\pi_{S_A(1)}, \pi_{S_A(2)}, \ldots, \pi_{S_A(1)}, \pi_{S_A(2)}, \pi_{S_A(3)}, \ldots, \pi_{S_A(2M_A)}) \). Using \( \pi_S \) we have the following expressions for \( f_L(1) \)
where $e$ is the vector of ones of appropriate size.

6. Numerical examples and performance evaluation

6.1. Bit error model

We applied our algorithm to available bit error traces and found that it matches both bit error rate and empirical ACF. Values of $\lambda$, used to approximate empirical ACFs, are shown in Table 1. One may note that given a certain $i_0$, $\gamma$ does not provide information regarding relative accuracy of the approximation. For comparison purposes coefficient $\phi$ is introduced. It is the error of ACF approximation computed at the maximum chosen lag, max($i_0$). In this study we choose max($i_0$) = 3 due to the following reasons. Firstly, the approximation capability of geometrically decaying function is rather limited. Indeed, increase of the coefficient $\lambda$ may result in overestimation of the lag-1 autocorrelation. One may expect that the value of lag-1 autocorrelation produces more impact on the performance of the model as compared to greater lags [23]. Thus, lag-1 autocorrelation should not be significantly overestimated or underestimated. Secondly, although increase of max($i_0$) may potentially decrease the error of ACF approximation, it also results in increase of complexity of the model. Keeping these reasons in mind we chose the model with $i_0 = 2$. It provides accuracy of approximation and simplicity of estimation.

The models obtained for our traces are as follows

\[ P_1(1) = 0, P_2(1) = 1 \]

\[ D_{E,1} = \begin{pmatrix} 0.99982 & 0.00018 \\ 0.56282 & 0.43718 \end{pmatrix}, \]

\[ D_{E,2} = \begin{pmatrix} 0.99859 & 0.00141 \\ 0.69059 & 0.30941 \end{pmatrix}. \]

Comparison of means of empirical data, model and generated traces is shown in Table 2. Recall, that the mean value and ACF provide sufficient information regarding distributional properties of the covariance stationary binary stochastic process. The mean of the model exactly matches mean of the empirical data. Mean of the generated trace deviates from mean of the model. This is due to statistical fluctuations and inaccuracies caused by simulation of small probabilities in (22).

ACFs of empirical traces and ACFs of obtained models are shown in Fig. 9. Depending on the trace, good approximation of the empirical data takes place up to lags 2–5.

6.2. Performance at the data-link layer

Let us consider performance of the data-link layer. Assume that the length of the frame is $m$ bits and FEC is not used at the data-link layer. The conditional mean of correctly received frames (this parameter can be interpreted as the mean number of correctly received frames given that they are generated according to Bernoulli process with probability of a single arrival set to 1) is given by the following expression

\[ E[N_F] = \pi_F D_F(0)e. \]
The estimated values of $E[N_F]$ for two models with different values of $m$ are shown in Fig. 10. Despite of completely different bit error rates observed for two traces, the difference at the data-link layer is not significant even for small frame sizes. This difference becomes negligible for $m > 100$.

Consider now the effect of FEC. Assume that a FEC code may correct up to $l$, $l = 0, 1, 2$, bit errors ($F_T = l + 1$). Then, the conditional mean of the correctly received frames is given by (23) with redefined $D_J(0)$. Results for different traces and values of $(m, l)$ are shown in Fig. 11. For illustrative purposes non-realistic values of $(m, l)$ are also included. Comparing obtained results one may note that the performance at the data-link layer can be significantly improved using FEC codes. Even usage of simplest FEC codes may drastically improve performance at the data-link layer. Again, there is no significant difference between two traces.

Let us now consider how parameters of the wireless channel model influence parameters at the data-link layer. Indeed, we found that first-order characteristics of the wireless channel at the physical layer may not always be important from the point of view of performance parameters at the data-link layer. So it is natural to ask: is it actually needed to look for precise models introducing additional complicatedness?

In addition to models given by (22) we also consider the following models. We keep the mean number of bit errors for both traces intact and change the autocorrelation introduced in models. It is accomplished by varying parameter $\lambda$. For both models we consider two additional values of $\lambda$: $\lambda = 0.0$ and $\lambda = 0.9$. The models we get are as follows ($f_1(1) = 0$, $f_2(1) = 1$)

$$
D_1^{z=0} = \begin{pmatrix} 0.99968 & 0.00032 \\ 0.99968 & 0.00032 \end{pmatrix},
$$

$$
D_1^{z=0.9} = \begin{pmatrix} 0.99997 & 0.00003 \\ 0.99997 & 0.90003 \end{pmatrix},
$$

$$
D_2^{z=0} = \begin{pmatrix} 0.99796 & 0.00204 \\ 0.99796 & 0.00204 \end{pmatrix},
$$

$$
D_2^{z=0.9} = \begin{pmatrix} 0.99980 & 0.00020 \\ 0.99980 & 0.90020 \end{pmatrix}.
$$

Estimated values of $E[N_F]$ for different values of $m$ when no FEC procedures are used at the data-link layer are shown in Fig. 12. For both models decrease in linear dependence between successive values causes significant performance
degradation. Moreover, increase of the bit error rate causes more deviation. We may interpret these results as follows: when more noise is added to the model and FEC is not used, the channel quality as seen at the data-link layer becomes worse.

Consider what happens when FEC procedures are added to the data-link layer. Computed values of $E[N_F]$ for both traces and different values of $(m,l)$ are shown in Fig. 13. One may see that, in general, the channel quality as seen at the data-link layer is much better when the channel is noisier and FEC procedures are used. However, it may not always hold in practice. Indeed, as shown in Fig. 13(c) when weak FEC codes are used noisy channels may lead to worse performance for large $m$.

Consider now the effect of strong channel memory. Estimated values of $E[N_F]$ for different values of $m$ when no FEC procedures are used at the data-link layer are shown in Fig. 14. Depending on the bit error rate of the wireless channel, strong memory may lead to either worse or better performance at the data-link layer.

Computed values of $E[N_F]$ for both traces and different values of $(m,l)$ are shown in Fig. 15. An interesting observation is that strong channel memory does not affect performance at the data-link layer significantly when the bit error rate is low and FEC code with strong correction capabilities is used.

Depending on our observations we conclude that both distributional and autocorrelation properties of the wireless
channel must be accurately taken into account. Changes in either bit error rate or memory of the channel may lead to undesirable effects causing degradation of performance parameters at the data-link layer. Additionally, strong channel memory may not always lead to degradation of the performance parameters, especially when FEC capabilities are weak. At the same time, noisiness of the wireless channel does not always guarantee better performance, compared to channels with stronger memory.

6.3. Performance at the IP layer

Consider how packetization between the data-link and IP layers influence performance at the IP layer. In what follows, we use models of wireless traces defined in (22), and assume that IP packets are generated according to MMBP (a special case of D-MAP, widely used for voice source modeling) with the following parameters

\[
D_A(0) = \begin{pmatrix} 0.12 & 0.24 \\ 0.36 & 0.12 \end{pmatrix},
\]

\[
D_A(1) = \begin{pmatrix} 0.09 & 0.56 \\ 0.24 & 0.28 \end{pmatrix},
\]

(25)

We also assume that the length of the frame is 20 bits, FEC code is used at the data-link layer and may correct up to 2 bits \(P_T = 3\). The mean number of correctly received packets, \(E[N_F]\), can be obtained from (21) and given by

![Fig. 14. Probabilities of correct frame reception without FEC; (a) trace 1, (b) trace 2.](image1)

![Fig. 15. Probabilities of correct frame reception with FEC; (a) trace 1, (b) trace 1, (c) trace 2, (d) trace 2.](image2)
Fig. 16. Probabilities of correct packet reception with FEC; (a) trace 1: \( z = 1,2,\ldots,10 \), (b) trace 2: \( z = 1,2,\ldots,10 \), (c) trace 1: \( z = 10,20,\ldots,100 \), (d) trace 2: \( z = 10,20,\ldots,100 \).

the following expression

\[
E[N_p] = \pi_S(D_4(1) \otimes D_R(0))e.
\]  

(26)

Computed values of \( E[N_p] \) for both traces, different values of \((m,l), z = 1,2,\ldots,10, \) and \( z = 10,20,\ldots,100 \) are shown in Fig. 16. As was expected FEC procedures significantly improve the channel quality as seen at the IP layer. However, with increasing of the number of frames to which a packet is segmented, the probability of correct packet reception may decrease significantly even when FEC is used at the data-link layer. Comparing Fig. 16(b) and (c) one may observe the effect of bit error propagation to IP layer. Indeed, roughly the same quality is seen at the IP layer with Trace 1, \( z = 100 \) and Trace 2, \( z = 10 \).

7. Conclusions

In this work we proposed analytical performance evaluation framework suitable for delay-sensitive multimedia applications for which only data-link or physical layer FEC techniques are feasible. We firstly proposed simple and computationally efficient wireless channel modeling algorithm. The bit error model was then extended to IP layer using the cross-layer mappings. Our model is represented by the IP packet error process reflecting memory properties of the initial bit error process.

We showed that our approach allows to analytically derive estimators of IP packet error probabilities in presence of FEC at the data-link layer eliminating the need for computationally expensive time-consuming bit level simulations. It also provides a way to analytically evaluate the required correction capabilities of FEC codes to provide best possible performance at IP layer. The submitted approach can be used when the size of frames at the data-link layer and the size of packets at the IP layer are constant. It is also suitable for data-link layer performance evaluation when the size of IP frames is variable.

We also showed that inaccurate matching of first- and second-order properties of the bit error process may lead to either overestimation or underestimation of the performance at higher layers. We revealed that strong channel memory may not always lead to degradation of the performance parameter, especially when FEC capabilities are weak. However, noisiness of the wireless channel does not always guarantee better performance, compared to channels with stronger memory.

Acknowledgements

The financial support of Graduate School of Technology and Automation (GETA), and Nokia Foundation is gratefully acknowledged. This work has also been partially supported by E-NEXT project.

References


