

EE-ARQ: a Green ARQ-Based Algorithm for the Transmission of Video Streams on Noise Wireless Channels

Alfio Lombardo¹, Carla Panarello², Giovanni Schembra¹

¹DIEEI - University of Catania, Italy - {lombardo,schembra}@dieei.unict.it

²CNIT - Research Unit of University of Catania, Italy - cpana@dieei.unict.it

Received: January 30, 2013

Accepted: March 15, 2013

Published: March 31, 2013

DOI: 10.5296/npa.v5i1.3180

URL: <http://dx.doi.org/10.5296/npa.v5i1.3180>

Abstract

The challenge in the next future is to increase the energy efficiency of telecommunications networks, and specifically wireless devices that present the highest energy consumption coefficient per bit transmitted among all the networking devices. However, making a network device green can cause performance deterioration. The target of this paper is to propose a new algorithm for the transmission of multiplexed rate-controlled multimedia streams over wireless channels. The algorithm is an energy-efficient variant of ARQ to exploit the correlation of the wireless channel behavior. In addition, in order to compensate transmission bandwidth reduction due to the energy saving policies, a cross-layer approach is applied introducing a source Rate Controller working to modify the video quality according to the state of the transmission buffer. An analytical model of the whole system is presented in order to evaluate performance and provide some guidelines to design the configuration parameters of the proposed algorithm.

Keywords: Wireless communications, Energy efficient algorithm, Cross-layer approach, ARQ, Video transmission, Performance Evaluation.

1. Introduction

Appealing services of telecommunications networks today are given by video transmissions on wireless networks [28]. They are supported by the enormous diffusion of portable computers and new generation cellular phones on the one hand, on the other have favored the realization of very complex widespread wireless networks.

At the same time, in the last few years power consumption has become very relevant in our life, and there is EU-wide incentive to reduce carbon dioxide emissions by 20% before year 2020. More in deep, 3% is expected to come from the ICT sector and a major role in "greening" telecommunications will be played by wireless networking technologies. For this reason, if up to now the main goal of research and industrial work in telecommunications has been to maximize performance, or reduce energy consumption in mobile devices to lengthen their battery life, the challenge in the next future will be to realize green telecommunications networks, and specifically wireless devices that present the highest energy consumption coefficient per transmitted bit, among all the networking devices [20].

In this perspective a significant amount of works have been done in recent years to make both wired [19, 40, 34, 24, 41, 22, 36, 35]) and wireless transmissions energy efficient [37, 29]. More specifically, in wireless networks the most part of them have been devoted to save energy in low-power battery devices, with the aim of increasing their lifetime [13, 4]. Since the launch of 3G access, mobile networks are a major consumer of electricity, and with LTE mobile operators we have to prepare for even further increases in power consumption per base station. The energy bill of current deployed wireless networks is already more than significant, surpassing the 20% of the operating costs for some of them [3]. Thus, means to lower the energy consumption of wireless networks are very valuable. Unfortunately, the power amount necessary for efficient and reliable transmissions makes wireless network devices, like wireless routers and network interfaces, the most critical devices to be optimized. Therefore, greening a wireless network device can cause performance deterioration. It is widely accepted that a good way to improve performance from the network level to the application level is to use a cross-layer approach [33, 10, 9]. However, it is challenging to maintain perceived Quality of Service (QoS) acceptable when energy saving strategies are implemented.

Energy consumed for efficient wireless transmissions is strongly related to techniques for maintaining reliable communications over noisy channels, such as forward error correction (FEC) and automatic repeat request (ARQ). However, the channel-state unaware behavior makes both FEC and traditional ARQ techniques energy inefficient. For this reason, some channel-adaptive link layer protocol ideas, such as GBN-ARQ and SR-ARQ based on channel probing [44, 45], and ARQ based on stochastic learning automaton [39], have been proposed earlier. Now a challenging task for a successful deployment of mobile video services is to focus at the same time both power consumption at the transmission level and quality of service (QoS) at the application level, with the aim of providing system designers with a tool for achieving a tradeoff between these two above opposite targets.

With this in mind, in this paper we propose a new algorithm for the transmission of

multiplexed rate-controlled multimedia streams over wireless channels. A cross-layer approach is coupled with the algorithm to avoid quality deterioration due to the application of the energy saving policy. The overall system consists of both a channel-adaptive ARQ-based protocol and an adaptive video transmission system. The system is therefore adaptive in both video source coding and ARQ transmission. More specifically, we propose a new version of the SW-ARQ algorithm, in the following referred to as Energy-Efficient ARQ (EE-ARQ), in order to exploit the correlation of the wireless channel behavior, so minimizing transmission when the channel state is bad. In addition, in order to compensate transmission bandwidth reduction due to the energy saving policies, a Rate Controller is introduced to follow a feedback law to control the encoding rate of the sources. Finally, using an analytical model of the system, some guidelines are provided to design algorithm parameters.

The paper is structured as follows. Section 2 describes the proposed Green Adaptive Video Wireless Transmission system we consider in the rest of the paper. Section 3 introduces the proposed algorithm for energy-efficient transmission on wireless channels. Section 4 introduces the Markov model of the system and Section 5 defines the main performance parameters regarding both video encoding quality, queue, ARQ transmission and consumed power. In Section 6 the model is then applied to a case study to evaluate the performance of a real case. Finally, Section 8 concludes the paper.

2. System Description

In this section we will describe the *Green Adaptive Video Wireless Transmission* system we consider in the rest of the paper. It is constituted by a *Video Multiplexer* loaded by V *Adaptive Video Sources* connected to it through high-speed low-delay links. The *Video Multiplexer Queue* is served by a wireless channel with time-variant bit error rate (BER) behavior; channel losses are managed with the Automatic Repeat reQuest (ARQ) protocol. The wireless output link constitutes the system bottleneck. When channel conditions get worse, more retransmissions are needed and the Video Multiplexer queue length increases. In order to avoid congestion, video sources are adaptive, that is, their emission bit rate is modified by a *Rate Controller* located in the Video Multiplexer, according to the state of the Video Multiplexer Queue, with a mechanism described later in this section.

Adaptive Video Sources we are considering are any video sources that, according to a given feedback, can modify their encoding rate run time in order to change their emission rate.

Let $\underline{\Psi}$ be the encoding rate array, containing all the available emission bit rates of the video sources, and $\underline{\Omega}$ the quality array, containing the quality levels associated to the available emission bit rates, expressed in terms of peak signal to noise ratio (PSNR); let G be the number of encoding levels, that is, the cardinality of the sets $\underline{\Psi}$ and $\underline{\Omega}$.

Packets coming from video sources are subdivided in *ARQ blocks* of H bits to be managed by the ARQ protocol. These ARQ blocks are buffered in the Video Multiplexer

Queue whose dimension, defined as the maximum number of blocks that can be accommodated in the queue and in the server facility, is K . Let C be the transmission rate on the wireless link, expressed in bits per second, as seen by the Video Multiplexer Queue.

Thus the time needed to transmit one ARQ block is $\Delta_{ARQ} = H/C$.

In the considered scenario the most appropriate version of ARQ is the stop-and-wait ARQ (SW-ARQ) because delays introduced by it are not too high, given that link propagation delays are negligible as compared to the ARQ block transmission time; on the other hand it ensures that packets are received at destination in the same order as they were sent by the transmitter.

The system architecture is synthesized in Fig. 1. As we can note, the approach we propose is cross-layer. In fact, as shown in the figure, we can identify three distinct layers: the Adaptive Video Source layer, constituted by the video sources that encode video and produce IP packets that are sent to the underlying Video Multiplexer layer; the Video Multiplexer layer that fills ARQ blocks with the flows coming from the video sources, and buffers them in the Video Multiplexer Queue; the Wireless Transmitter layer that transmits ARQ block on the wireless channel. The above three layers, although remaining distinct, change their behavior according to the feedback received by the other layers. In fact, as already described so far, according to the state of the queue, the Video Multiplexer sends "Rate-increase" or "Rate-decrease" messages to the Adaptive Video Sources. On the other hand, according to the result of the wireless transmission, the Video Multiplexer with the EE-ARQ inside, decides whether a new, or the same, ARQ block has to be sent to the Wireless Transmitter, for a new attempt of transmission. The design is cross-layer as well, since system parameters can be chosen taking into account at the same time perceived quality at the Video source layer, delay and loss probability at the Video Multiplexer layer, and energy consumption at the Wireless transmitter layer.

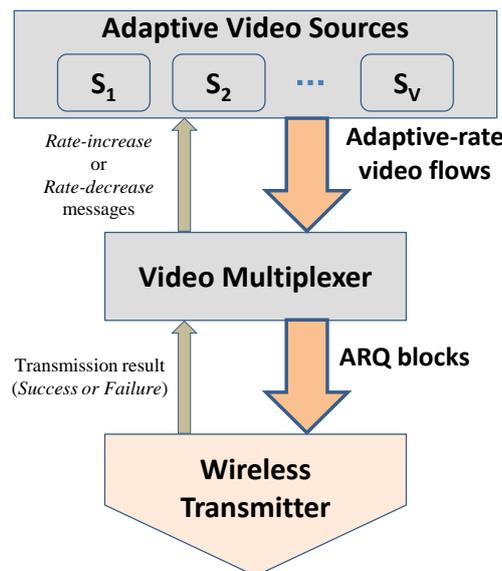


Figure 1: Cross-layer view of the system architecture

In the next section we will describe the algorithm to be applied to ARQ to achieve energy saving transmissions on wireless channels.

3. EE-ARQ Algorithm

In this paper we propose a modified version of the SW-ARQ protocol, in order to make it energy efficient. We will refer to this new algorithm as EE-ARQ (Energy-Efficient ARQ). The motivation at the base of it is that the quality behavior of the underlying wireless transmission channel is a strongly-correlated stochastic process, that is, the same signal-to-noise ratio (SNR) level is maintained for a period that is very long as compared to the transmission duration of a single ARQ block. For this reason, if a transmission has failed, it is highly likely that an immediately successive attempt will follow the same sort. Starting from this consideration, we propose to use a retransmission algorithm where the transmission is attempted with a probability depending on the number of previous attempts. In such a way, the sender deduces the state of the channel and transmits more rarely when the channel is considered bad. More specifically, as in the classical SW-ARQ protocol, let ρ be the counter of retransmission attempts for the same ARQ block ($\rho = 0$ when the block is transmitted for the first time). The counter ρ is incremented by one at each retransmission attempt, and reset to zero when a block is removed from the service facility because successfully transmitted or discarded because the maximum number of retransmissions, $s_{MAX}^{(R)}$, has been reached.

According to the new EE-ARQ protocol, in a generic instant when the sender should transmit a block according to the classical SW-ARQ protocol, the transmission is attempted with a probability $\wp^{(Tx)}(\rho)$ depending on the number of previous attempts, ρ . It is defined as:

$$\wp^{(Tx)}(\rho) = \frac{1}{Tx_{law}(\rho)} \quad (1)$$

where $Tx_{law}(\rho)$ is the transmission law associated to the retransmission policy. In this paper we consider and analyze a retransmission policy similar to the exponential backoff adopted by IEEE 802.3 CSMA/CD protocol. When this policy is adopted, the transmission law, i.e. the denominator of the probability to attempt a transmission, increases exponentially with the number of previous attempts of retransmission. It is defined as:

$$Tx_{law}(\rho) = \gamma^\rho, \quad \text{with } \gamma \geq 1 \quad \text{and} \quad \rho \in \{0, \dots, s_{MAX}^{(R)}\} \quad (2)$$

Using this law allows the sender to deduce the state of the channel to transmit more rarely when the channel is bad. As we will see in Section 6, the choice of γ has a strong impact on the overall system performance in terms of both application quality and energy consumption.

As already observed so far, the task of the *Rate Controller* is to control the emission bit rate of the video sources with the target of maintaining the queue length as much constant as possible, avoiding situations in which the buffer empties or overflows due to some channel condition variations. To this purpose the Rate Controller periodically monitors the state s_Q of the Video Multiplexer Queue, defined as the number of ARQ blocks which are present in the queue and in the service facility; based on it, the Rate Controller implements a feedback law that determines at each time slot whether sending “rate-increase” or “rate-decrease” feedback messages to the video sources. In order to tune the system reaction time, the Rate Controller decides the number of sources that have to change their rates. More specifically, at each time slot, the Rate Controller first decides the kind of message to send to the video sources according to the state of the queue: when s_Q is less than a given threshold q_L , it sends “rate-increase” messages, while it sends “rate-decrease” messages when s_Q is greater than another threshold, q_H ; no messages are sent when $s_Q \in [q_L, q_H]$. Then the Rate Controller decides the number of sources that have to receive the above messages. This number is calculated as $\tilde{v}(s_Q) = \phi \cdot f_{law}(s_Q)$, where $f_{law}(s_Q)$ is the per-source feedback law mask defined as follows:

$$f_{law}(s_Q) = \begin{cases} -\frac{s_Q}{q_L} + 1 & 0 \leq s_Q < q_L \\ 0 & q_L \leq s_Q \leq q_H \\ \frac{s_Q - q_H}{K - q_H} & q_H < s_Q \leq K \end{cases} \quad (3)$$

where K is the Video Multiplexer Queue dimension, as defined at the beginning of Section II. The function f_{law} is shown in Fig. 2. The coefficient $\phi \in \{1, \dots, V\}$ allows the Rate Controller to decide the maximum number of sources that can be contacted simultaneously.

Since the number $\tilde{v}(s_Q)$ may not be an integer, it is rounded to one of the closest integer values with probabilities proportional to its distance from them. So the final number of sources that have to change their rate is given by:

$$v(s_Q) = \begin{cases} \lfloor \tilde{v}(s_Q) \rfloor & \text{with probability : } 1 - \wp_V(s_Q) \\ \lfloor \tilde{v}(s_Q) \rfloor + 1 & \text{with probability : } \wp_V(s_Q) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where $\wp_V(s_Q) = \tilde{v}(s_Q) - \lfloor \tilde{v}(s_Q) \rfloor$, and $\lfloor x \rfloor$ indicates the maximum integer lower than or equal to x . The $v(s_Q)$ sources to be contacted are randomly chosen among the V sources loading the buffer, starting from the ones encoding at the highest bit rate if $s_Q \in [q_H + 1, K]$, or from the ones encoding at the lowest bit rate, if $s_Q \in [0, q_L - 1]$.

The parameters q_L and q_H can be used to tune the reaction time of the system against the stability of the quality of service perceived at the user level: the higher the distance between q_L and q_H , the less reactive the system, but the more stable the encoding quality.

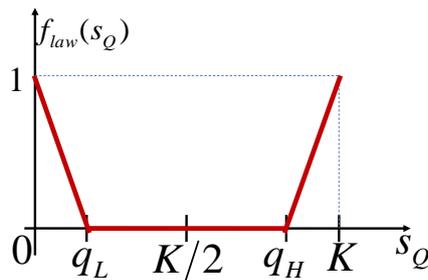


Figure 2: Feedback law

4. System Model

In this section we will model the Green Adaptive Wireless Transmission system described in the previous section, and in the following indicated as Σ . As already said, it is a queueing system loaded by an aggregate of V video adaptive sources, and served by a wireless channel with the energy efficient EE-ARQ mechanism described so far. The model will be defined by using the most general Markov-modulated process in the discrete-time domain, the Switched Batch Bernoulli process (SBBP) [25].

4.1 Markov Chain State Definition

As usual, the first step to define a discrete-time system model is to choose the time slot. As mentioned before, ARQ blocks are the queueing units. Therefore, in the following we will consider, as the model time slot, the so-called ARQ reaction period, indicated as Δ , and defined as the time needed to transmit an ARQ block and receive the relative ack on the wireless link.

Let us indicate the system constituted by the queue and the service facility as the Multiplexer Queue. Let us define the following processes:

- $N(p)$ be the queue drain process of this system, representing the number of ARQ blocks removed from the system in the P -th slot;
- $W(p)$ be the emission process of the Adaptive Source aggregate, representing the number of ARQ blocks sent to the queue from the source aggregate during the P -th slot.

A complete description of the system Σ at the p -th slot requires a two-dimensional Markov chain, whose state in the generic p -th slot is defined as $\underline{S}^{(\Sigma)}(p) = (\underline{S}^{(W)}(p), \underline{S}^{(S)}(p))$, where:

- $\underline{S}^{(W)}(p)$ is the state of the underlying Markov chain of the adaptive source aggregate emission process $W(p)$. It is a vector of G elements; the g -th element, for $g \in \{1, \dots, G\}$, represents the number of sources using the g -th encoding level for the frame to be encoded in the p -th slot;
- $\underline{S}^{(S)}(p)$ is the state of the Markov chain of the queueing system process; it is a two-dimensional Markov chain defined as $\underline{S}^{(S)}(p) = (S^{(Q)}(p), \underline{S}^{(N)}(p))$ where:
 - $S^{(Q)}(p) \in \{0, \dots, K\}$ is the queue state, i.e. the number of ARQ blocks in the queue and in the server facility at the p -th slot;
 - $\underline{S}^{(N)}(p)$ is the state of the Markov chain characterizing the queue drain process $N(p)$; it is defined as $\underline{S}^{(N)}(p) = (S^{(R)}(p), S^{(C)}(p))$ where:
 - $S^{(R)}(p) \in \{w', w_0', 0, w_1', 1, \dots, w_{s_{MAX}}^{(R)}, s_{MAX}^{(R)}\}$ is the ARQ retransmission state. More specifically, $S^{(R)}(p) = \rho$, for $\rho \in \{0, \dots, s_{MAX}^{(R)}\}$, if a block is transmitted at the p -th slot and this transmission is the ρ -th retransmission for that block (note that $S^{(R)}(p) = 0$ for the first transmission of each block);

$S^{(R)}(p) = w_\rho$, for $\rho \in \{0, \dots, s_{MAX}^{(R)}\}$, if the block in the service is not transmitted at the p -th slot, waiting for the ρ -th retransmission in a later slot; $S^{(R)}(p) = w$ (queue waiting for a packet) if there are no blocks in the queue to be transmitted (i.e. $S^{(Q)}(p) = 0$).

- $S^{(C)}(p)$ is the channel state during the transmission at the p -th slot.

4.2 System Time Evolution Model

Since we are applying a discrete-time model, the first step is to define the sequence of events in each slot. To this end let us consider the values of the system state $\underline{S}^{(\Sigma)}(p)$ in the generic $(p-1)$ -th and p -th slots as $\underline{s}_{\Sigma 1} = (\underline{s}_{W1}, \underline{s}_{S1})$ and $\underline{s}_{\Sigma 2} = (\underline{s}_{W2}, \underline{s}_{S2})$, respectively, where $\underline{s}_{S1} = (s_{Q1}, \underline{s}_{N1}) = (s_{Q1}, (s_{R1}, s_{C1}))$ and $\underline{s}_{S2} = (s_{Q2}, \underline{s}_{N2}) = (s_{Q2}, (s_{R2}, s_{C2}))$, respectively.

Let us now define the structure of the generic system state transition from the generic slot $p-1$ to the slot p . The system state is monitored at the end of each slot, in a time instant referred to as the *Observation Point*, containing the starting values for the system variables of the next slot. Each transition is structured according to the following steps:

1. at the beginning of the p -th slot, the *transmission result* (i.e. failure or success) of the attempt of transmission done in the $(p-1)$ -th slot, is known: it determines the occurrence that in the p -th slot a retransmission for the same block is needed (in case of failure), or a new block in the queue, if any, enters the service substituting the successfully transmitted one (in case of success); therefore, the value $d \in \{0, 1\}$ of the process $N(p)$, representing the number of ARQ blocks removed from the service, is determined;
2. at the same time, the feedback messages (if any) generated by the Rate Controller on the basis of the queue state value s_{Q1} in the $(p-1)$ -th slot, determine the source aggregate state transition $\underline{s}_{W1} \rightarrow \underline{s}_{W2}$ and, consequently, the number of blocks, $W(p) = r_w$, emitted by the source aggregate during the p -th slot;

3. the number of blocks removed from the service, together with the number of emitted blocks by the source aggregate, deterministically determine the queue state transition

$$s_{Q1} \rightarrow s_{Q2};$$

4. afterwards, based on the number of packets removed from the service and on the new value s_{Q2} of the queue state in the p -th slot, and according to the retransmission

probability $\wp^{(Tx)}(\rho)$ in (1), the transition of the retransmission counter state,

$$s_{R1} \rightarrow s_{R2}, \text{ is derived;}$$

5. finally, the channel state transition, $s_{C1} \rightarrow s_{C2}$, occurs.

Once the system state transition is completed, if there is a block in the service in the p -th slot, an attempt of transmission is done according to the retransmission probability $\wp^{(Tx)}(\rho)$. The *transmission result*, based on the channel state s_{C2} in the p -th slot, will be known at the Observation Point at the end of the p -th slot. This result will determine the system state transition in the next slot.

4.3 Model of the Whole Green Adaptive Wireless Transmission System

The target of this section is to derive the transition probability matrix $P^{(\Sigma)}$ of the Markov process $\underline{S}^{(\Sigma)}(p)$, representing the state of the Transmission system as a whole. The generic element of the transition probability matrix $P^{(\Sigma)}$ is defined as:

$$P_{[s_{\Sigma 1}, s_{\Sigma 2}]}^{(\Sigma)} = Prob \left\{ \begin{array}{l} \underline{S}^{(W)}(p) = \underline{s}_{W2} \mid \underline{S}^{(W)}(p-1) = \underline{s}_{W1} \\ \underline{S}^{(S)}(p) = \underline{s}_{S2} \mid \underline{S}^{(S)}(p-1) = \underline{s}_{S1} \end{array} \right\} \quad (5)$$

Now, applying the total probability theorem for each possible value of the number of arrivals, r , and the number of blocks drained from the queue, d , after some mathematical elaborations accounting the memoryless property of Markov chains, we obtain:

$$\begin{aligned}
 P_{[\underline{s}_{\Sigma 1}, \underline{s}_{\Sigma 2}]}^{(\Sigma)} &= \sum_{r=0}^{r_{MAX}^{(W)}} \sum_{d=0}^1 \text{Prob} \left\{ N(p) = d \begin{array}{l} S^{(Q)}(p-1) = s_{Q1} \\ S^{(R)}(p-1) = s_{R1} \\ S^{(C)}(p-1) = s_{C1} \end{array} \right\} \cdot \text{Prob} \left\{ \begin{array}{l} \underline{s}^{(W)}(p) = \underline{s}_{W2} \\ W(p) = r \end{array} \middle| \begin{array}{l} \underline{s}^{(W)}(p-1) = \underline{s}_{W1} \\ S^{(Q)}(p-1) = s_{Q1} \end{array} \right\} \\
 &\cdot \text{Prob} \left\{ \begin{array}{l} S^{(Q)}(p) = s_{Q2} \\ W(p) = r \\ N(p) = d \end{array} \right\} \cdot \text{Prob} \left\{ \begin{array}{l} S^{(R)}(p) = s_{R2} \\ S^{(C)}(p-1) = s_{C1} \\ S^{(Q)}(p) = s_{Q2} \\ N(p) = d \end{array} \right\} \\
 &\cdot \text{Prob} \left\{ S^{(C)}(p) = s_{C2} \middle| S^{(C)}(p-1) = s_{C1} \right\}
 \end{aligned} \tag{6}$$

The above equation can be rewritten as follows:

$$P_{[\underline{s}_{\Sigma 1}, \underline{s}_{\Sigma 2}]}^{(\Sigma)} = B_{[(s_{Q1}, (s_{R1}, s_{C1}), d)]}^{(N)} \cdot C_{[\underline{s}_{W1}, \underline{s}_{W2}]}^{(W)}(r, s_{Q1}) \cdot P_{[s_{Q1}, s_{Q2}]}^{(Q)}(r, d) \cdot P_{[s_{R1}, s_{R2}]}^{(R)}(s_{C1}, s_{Q2}, d, l) \cdot P_{[s_{C1}, s_{C2}]}^{(C)} \tag{7}$$

where:

- $B_{[(s_{Q1}, (s_{R1}, s_{C1}), d)]}^{(N)}$ is the first probability in (6), and represents the probability that d ARQ blocks are drained from the queue in the p -th slot;
- $C_{[\underline{s}_{W1}, \underline{s}_{W2}]}^{(W)}(r, s_{Q1})$ is the second probability in (6); it is the probability to move from the state \underline{s}_{W1} to the state \underline{s}_{W2} , including the probability that r ARQ blocks are emitted by the source aggregate when its Markov chain is in the state \underline{s}_{W2} ;
- $P_{[s_{Q1}, s_{Q2}]}^{(Q)}(r, d)$, $P_{[s_{R1}, s_{R2}]}^{(R)}(s_{C1}, s_{Q2}, d, l)$ and $P_{[s_{C1}, s_{C2}]}^{(C)}$ are the third, fourth and fifth probabilities in (6), and are elements of the transition probability matrices of the Queue, the ARQ retransmission process and the wireless channel state, respectively.

The matrix $C^{(W)}(r, s_{Q1})$ will be derived in Section 4.4. The other matrices will be calculated in Section 4.5.

Once the matrix $P^{(\Sigma)}$ is known, we can calculate the steady-state probability array of the system Σ as the solution of the following linear system:

$$\begin{cases} \underline{\pi}^{(\Sigma)} \cdot P^{(\Sigma)} = \underline{\pi}^{(\Sigma)} \\ \underline{\pi}^{(\Sigma)} \cdot \underline{1}^T = 1 \end{cases} \tag{8}$$

where $\underline{1}^T$ is a column array whose elements are all equal to 1. It may be difficult to solve

(8) directly since the number of states grows explosively as the queue size increases. Nevertheless, many algorithms [38] enable us to calculate the array $\underline{\pi}^{(\Sigma)}$ maintaining a linear dependency on K .

4.4 Adaptive Source Aggregate Model

In this section we derive the model of the emission process of the Adaptive Source Aggregate. As said in Section 2, the emission rate of each source of the aggregate is driven by the Rate Controller which implements a feedback law that, according to the state s_Q of the Video Multiplexer Queue, determines the number of sources a feedback message has to be sent to.

Now let us define the underlying Markov chain state of the Adaptive Source Aggregate emission process $W(p)$. This chain has a G -dimensional state, i.e. $\underline{s}^{(W)}(p) = (S_{[1]}(p), S_{[2]}(p), \dots, S_{[G]}(p))$, where G is the number of encoding levels, and the value of $S_{[g]}(p)$, for each $g \in \{1, \dots, G\}$, represents the number of sources encoding with the g -th encoding level in the p -th slot.

At the beginning of the p -th slot, the Rate Controller randomly chooses a set of sources among the V sources loading the buffer, starting from the ones encoding at the highest bit rate, if $s_Q \in [q_H + 1, K]$, or from the ones encoding at the lowest bit rate, if $s_Q \in [0, q_L - 1]$. Then it sends them a "rate-decrease" or "rate-increase" message according to the result of the comparison of s_Q with the two thresholds, q_L and q_H , as described in Section 2.

Therefore the generic state transition $\underline{s}_{W1} \rightarrow \underline{s}_{W2}$ depends, through the feedback law, on the queue state value s_{Q1} in the $(p-1)$ -th slot. For this reason, let us indicate the transition probability matrix of the underlying Markov chain of $W(p)$ as $P^{(W)}(s_{Q1})$. Its generic element can be defined as follows

$$P_{\lfloor \underline{s}_{W1}, \underline{s}_{W2} \rfloor}^{(W)}(s_{Q1}) = \begin{cases} 1 - \wp_V(s_{Q1}) & \text{if } s_{Q1} < q_L & \underline{s}_{W1}, \underline{s}_{W2} \in (\mathcal{A} \text{ or } \mathcal{B}) \\ \wp_V(s_{Q1}) & \text{if } s_{Q1} < q_L & \underline{s}_{W1}, \underline{s}_{W2} \in (\mathcal{C} \text{ or } \mathcal{D}) \\ 1 - \wp_V(s_{Q1}) & \text{if } s_{Q1} > q_H & \underline{s}_{W1}, \underline{s}_{W2} \in (\mathcal{E} \text{ or } \mathcal{F}) \\ \wp_V(s_{Q1}) & \text{if } s_{Q1} > q_H & \underline{s}_{W1}, \underline{s}_{W2} \in (\mathcal{G} \text{ or } \mathcal{H}) \\ 1 & \text{if } s_{Q1} < q_L & \underline{s}_{W1}, \underline{s}_{W2} \in \mathcal{I} \\ 1 & \text{if } s_{Q1} > q_H & \underline{s}_{W1}, \underline{s}_{W2} \in \mathcal{J} \\ 1 & \text{if } q_L \leq s_{Q1} \leq q_H & \underline{s}_{W1}, \underline{s}_{W2} \in \mathcal{K} \\ 0 & \text{otherwise} & \end{cases} \quad (9)$$

where the conditions in the last column of (9) are detailed in Table 1. More specifically:

- **conditions \mathcal{A} and \mathcal{B}** identify the state transitions happening when the queue length is below the threshold q_L and the Rate Controller sends a "rate-increase" message to $\lfloor \tilde{v}(s_{Q1}) \rfloor$ sources randomly chosen among the sources encoding with the minimum bit rate;
- **conditions \mathcal{C} and \mathcal{D}** identify the state transitions happening when the queue length is below the threshold q_L and the Rate Controller sends an "rate-increase" message to $(\lfloor \tilde{v}(s_{Q1}) \rfloor + 1)$ sources;
- **conditions \mathcal{E} and \mathcal{F}** identify the state transitions happening when the queue length is over the threshold q_H and the Rate Controller sends a "rate-decrease" message to $\lfloor \tilde{v}(s_{Q1}) \rfloor$ sources;
- **conditions \mathcal{G} and \mathcal{H}** identify the state transitions happening when the queue length is over the threshold q_H and the Rate Controller sends a "rate-decrease" message to $(\lfloor \tilde{v}(s_{Q1}) \rfloor + 1)$ sources;
- **condition \mathcal{I}** identifies the state transitions happening when the queue length is below the threshold q_L and the Rate Controller should send a "rate-increase" message to $\lfloor \tilde{v}(s_{Q1}) \rfloor$ sources, but in the state \underline{s}_{W1} there are less than $\lfloor \tilde{v}(s_{Q1}) \rfloor$ sources that can increase their encoding level;
- **condition \mathcal{J}** , opposite to the condition \mathcal{I} , identifies the state transitions happening when

the queue length is over the threshold q_H and the Rate Controller should send a "rate-decrease" message to $\lfloor \tilde{v}(s_{Q1}) \rfloor$ sources; however, in the state \underline{s}_{W1} there are less than $\lfloor \tilde{v}(s_{Q1}) \rfloor$ sources that can decrease their encoding level;

- **condition \mathcal{K}** identifies the cases where the queue length is within the two thresholds q_L and q_H and no feedback message is sent by the Rate Controller.

\mathcal{A} :	$\exists g \in \{1, \dots, G-1\} : \underline{s}_{W1[g]} \geq \lfloor \tilde{v}(s_{Q1}) \rfloor, \underline{s}_{W2[g]} = \underline{s}_{W1[g]} - \lfloor \tilde{v}(s_{Q1}) \rfloor,$ $\underline{s}_{W2[g+1]} = \underline{s}_{W1[g+1]} + \lfloor \tilde{v}(s_{Q1}) \rfloor, \underline{s}_{W2[G]} \neq V$
\mathcal{B} :	$\exists g \in \{1, \dots, G-2\} : \underline{s}_{W1[g]} < \lfloor \tilde{v}(s_{Q1}) \rfloor, \underline{s}_{W2[g]} = 0,$ $\underline{s}_{W2[g+1]} = V - (\lfloor \tilde{v}(s_{Q1}) \rfloor - \underline{s}_{W1[g]}), \underline{s}_{W2[g+2]} = \lfloor \tilde{v}(s_{Q1}) \rfloor - \underline{s}_{W1[g]}$
\mathcal{C} :	$\exists g \in \{1, \dots, G-1\} : \underline{s}_{W1[g]} \geq \lfloor \tilde{v}(s_{Q1}) \rfloor + 1,$ $\underline{s}_{W2[g]} = \underline{s}_{W1[g]} - (\lfloor \tilde{v}(s_{Q1}) \rfloor + 1), \underline{s}_{W2[g+1]} = \underline{s}_{W1[g+1]} + \lfloor \tilde{v}(s_{Q1}) \rfloor + 1$
\mathcal{D} :	$\exists g \in \{1, \dots, G-2\} : \underline{s}_{W1[g]} < \lfloor \tilde{v}(s_{Q1}) \rfloor + 1, \underline{s}_{W2[g]} = 0,$ $\underline{s}_{W2[g+1]} = V - \lfloor \tilde{v}(s_{Q1}) \rfloor - 1 + \underline{s}_{W1[g]}, \underline{s}_{W2[g+2]} = \lfloor \tilde{v}(s_{Q1}) \rfloor + 1 - \underline{s}_{W1[g]}$
\mathcal{E} :	$\exists g \in \{2, \dots, G\} : \underline{s}_{W1[g]} \geq \lfloor \tilde{v}(s_{Q1}) \rfloor, \underline{s}_{W2[g]} = \underline{s}_{W1[g]} - \lfloor \tilde{v}(s_{Q1}) \rfloor,$ $\underline{s}_{W2[g-1]} = \underline{s}_{W1[g-1]} + \lfloor \tilde{v}(s_{Q1}) \rfloor, \underline{s}_{W2[1]} \neq V$
\mathcal{F} :	$\exists g \in \{3, \dots, G\} : \underline{s}_{W1[g]} < \lfloor \tilde{v}(s_{Q1}) \rfloor, \underline{s}_{W2[g]} = 0,$ $\underline{s}_{W2[g-1]} = V - (\lfloor \tilde{v}(s_{Q1}) \rfloor - \underline{s}_{W1[g]}), \underline{s}_{W2[g-2]} = \lfloor \tilde{v}(s_{Q1}) \rfloor - \underline{s}_{W1[g]}$
\mathcal{G} :	$\exists g \in \{2, \dots, G\} : \underline{s}_{W1[g]} \geq \lfloor \tilde{v}(s_{Q1}) \rfloor + 1, \underline{s}_{W2[g]} = \underline{s}_{W1[g]} - (\lfloor \tilde{v}(s_{Q1}) \rfloor + 1),$ $\underline{s}_{W2[g-1]} = \underline{s}_{W1[g-1]} + \lfloor \tilde{v}(s_{Q1}) \rfloor + 1$
\mathcal{H} :	$\exists g \in \{3, \dots, G\} : \underline{s}_{W1[g]} < \lfloor \tilde{v}(s_{Q1}) \rfloor + 1, \underline{s}_{W2[g]} = 0,$ $\underline{s}_{W2[g-1]} = V - (\lfloor \tilde{v}(s_{Q1}) \rfloor + 1 - \underline{s}_{W1[g]}), \underline{s}_{W2[g-2]} = \lfloor \tilde{v}(s_{Q1}) \rfloor + 1 - \underline{s}_{W1[g]}$
\mathcal{I} :	$\underline{s}_{W1[G-1]} \leq \lfloor \tilde{v}(s_{Q1}) \rfloor, \underline{s}_{W2[G]} = V$
\mathcal{J} :	$\underline{s}_{W1[2]} \leq \lfloor \tilde{v}(s_{Q1}) \rfloor, \underline{s}_{W2[1]} = V$
\mathcal{K} :	$\underline{s}_{W1} = \underline{s}_{W2}$

Table 1: State sets in Eq. (9)

Now, let us derive the emission probability matrix $B^{(W)}$, whose generic element is defined as:

$$B_{[s_W, r]}^{(W)} = Prob\{W(p) = r | \underline{S}^{(W)}(p)\} \quad (10)$$

For the evaluation of (10), we need to determine the number r of ARQ blocks emitted by the source aggregate during the generic p -th time slot when the state of the underlying Markov chain is $\underline{S}^{(W)}(p)$. Previously, we have defined the encoding rate array $\underline{\Psi}$ as an array containing the emission bit rate of a single source for each encoding level. Therefore, let us define $\underline{\Psi}^{(W)}$ as the encoding rate array expressed in terms of ARQ blocks, each of H bits, emitted by the source aggregate during Δ seconds.

Note that for the conversion from bits/s to ARQ blocks/slot, we have to take into account that the H bits of an ARQ block include the *overhead* given by the sum of all the layer headers. The generic element of $\underline{\Psi}^{(W)}$, representing the number $r \in \{0, \dots, r_{MAX}^{(W)}\}$ of ARQ blocks emitted by the source aggregate during a time slot when its state is s_W , can be calculated as:

$$\Psi_{[s_W]}^{(W)} = \frac{\Delta}{H} \cdot \left(\sum_{g=1}^G s_{W[g]} \Psi_{[g]}^* \right) \quad \forall g \in \{1, \dots, G\} \quad (11)$$

where $\underline{\Psi}^*$ is the encoding rate array derived from $\underline{\Psi}$ taking into account the overhead introduced by the higher-layer protocol headers.

Note that the maximum number $r_{MAX}^{(W)}$ of ARQ blocks that can be emitted by the source aggregate in each slot is obtained from (11) by considering the state value s_W where all the sources emit with the maximum bit rate, (i.e. $s_{W[g]} = 0, \forall g \in \{1, \dots, G-1\}$ and $s_{W[G]} = V$), that is:

$$r_{MAX}^{(W)} = V \cdot \left\lceil \frac{\Delta}{H} \cdot \Psi_{[G]}^* \right\rceil \quad (12)$$

where $\lceil x \rceil$ indicates the minimum integer greater than or equal to x . Since $\underline{\Psi}_{[s_W]}^{(W)}$ in

(11) can be a non-integer number, we assume that, in each slot, the number of emitted ARQ blocks by the aggregate $W(p)$ is:

$$W(p) = \begin{cases} \lfloor \underline{\Psi}_{[s_W]}^{(W)} \rfloor & \text{with prob. } 1 - \wp'_{[s_W]} \\ \lfloor \underline{\Psi}_{[s_W]}^{(W)} \rfloor + 1 & \text{with prob. } \wp'_{[s_W]} \end{cases} \quad (13)$$

where $\wp'_{[s_W]} = \underline{\Psi}_{[s_W]}^{(W)} - \lfloor \underline{\Psi}_{[s_W]}^{(W)} \rfloor$;

In summary, the generic element of the emission probability matrix $B^{(W)}$ of the SBBP modeling the source aggregate emission process $W(p)$, is:

$$B_{[s_W, r]}^{(W)} = \begin{cases} 1 - \wp'_{[s_W]} & \text{if } r = \lfloor \underline{\Psi}_{[s_W]}^{(W)} \rfloor \\ \wp'_{[s_W]} & \text{if } r = \lfloor \underline{\Psi}_{[s_W]}^{(W)} \rfloor + 1 \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

So far, we have derived the matrices $P^{(W)}(s_{Q1})$ and $B^{(W)}$ completely characterizing the SBBP $W(p)$. Now, according to [25], we can equivalently represent $W(p)$ with the matrix $C^{(W)}(r, s_{Q1})$, defined as the transition probability matrix with the probability of emitting a given number r of ARQ blocks. Its generic element is defined as follows:

$$C_{[s_{W1}, s_{W2}]}^{(W)}(r, s_{Q1}) = Prob \left\{ \begin{array}{l|l} W(p+1) = r & S^{(Q)}(p) = s_{Q1} \\ \underline{S}^{(W)}(p+1) = s_{W2} & \underline{S}^{(W)}(p) = s_{W1} \end{array} \right\} \quad (15)$$

The matrix $C^{(W)}(r, s_{Q1})$ can be derived from $P^{(W)}(s_{Q1})$ and $B^{(W)}$ as follows:

$$C^{(W)}(r, s_{Q1}) = P^{(W)}(s_{Q1}) \cdot \text{diag}(B_{[r, r]}^{(W)}) \quad (16)$$

where $\text{diag}(B_{[r, r]}^{(W)})$ is the diagonal matrix containing the r -th column of the emission probability matrix $B^{(W)}$.

4.5 Multiplexer System Model

The target of this section is to define the Markov model of the Multiplexer System. According to the system state transition steps defined in Section 4.2, we will first consider the queue drain process $N(p)$. Then we will derive the transition probability matrices of the queue and the ARQ retransmission counter state. Finally, we will consider the wireless channel state transition.

1. Queue Drain Process

Let us describe the queue drain process $N(p)$. It has been defined as the number of ARQ blocks removed from the service in the p -th slot. According to the choice of the model time slot Δ , the number of ARQ blocks removed from the service cannot be greater than 1. More specifically, $N(p) = 1$ if the block transmitted in the $(p-1)$ -th slot has arrived correctly, or the maximum number of retransmissions has been reached in the same $(p-1)$ -th slot; otherwise $N(p) = 0$. Of course, the fact that $N(p) = 1$ means that the block which is in the service at the $(p-1)$ -th slot is removed, and a new block, if there is any, enters the service facility at the p -th slot.

The SBBP processes $N(p)$ is modulated by the underlying Markov chain $\underline{S}^{(N)}(p) = (S^{(R)}(p), S^{(C)}(p))$. More specifically, its value in the p -th slot depends on $\underline{S}^{(N)}(p-1)$, i.e. the state of the same underlying Markov chain in the $p-1$ -th slot. In addition it depends on the queue state $S^{(Q)}(p-1)$, as we will see below.

Let us indicate the values of the retransmission counter state $S^{(R)}(p-1)$, the channel state $S^{(C)}(p-1)$, and the queue state $S^{(Q)}(p-1)$ as s_{R1} , s_{C1} and s_{Q1} , respectively. Consequently, we have:

$$\begin{aligned}
 B_{[(s_{Q1}, s_{R1}, s_{C1}), d]}^{(N)} &= \text{Prob} \left\{ N(p) = d \left| \begin{array}{l} S^{(Q)}(p-1) = s_{Q1} \\ S^{(R)}(p-1) = s_{R1} \\ S^{(C)}(p-1) = s_{C1} \end{array} \right. \right\} = \\
 &= \begin{cases} 1 & s_{R1} = 'w' & s_{Q1} = 0 & d = 0 \\ 1 & s_{R1} \in \{ 'w_0', \dots, 'w_{s_{MAX}^{(R)}}' \} & s_{Q1} > 0 & d = 0 \\ \wp_{[s_{C1}]}^{(ABE)} & s_{R1} \in \{ 0, \dots, s_{MAX}^{(R)} - 1 \} & s_{Q1} > 0 & d = 0 \\ 1 - \wp_{[s_{C1}]}^{(ABE)} & s_{R1} \in \{ 0, \dots, s_{MAX}^{(R)} - 1 \} & s_{Q1} > 0 & d = 1 \\ 1 & s_{R1} = s_{MAX}^{(R)} & s_{Q1} > 0 & d = 1 \\ 0 & \text{otherwise} & & \end{cases} \quad (17)
 \end{aligned}$$

where $\wp_{[s_{C1}]}^{(ABE)}$ represents the ARQ block error probability in the channel state s_{C1} , and will be calculated later as in (21).

2. Queue Model

Let us now consider the queue state transition. To this end let us consider two generic values of the queue state, s_{Q1} and s_{Q2} , in the $(p-1)$ -th and the p -th slots, respectively. As discussed so far, the ARQ blocks are the queuing units. Therefore, at the generic p -th slot, the queue is decremented by d blocks and incremented by r blocks. Thus the value of the queue state at the p -th slot can be obtained through the following equation:

$$s_{Q2} = \min \{ \max \{ s_{Q1} - d, 0 \} + r, K \} \quad (18)$$

where K is the Multiplexer queue size. Thus, given s_{Q1} , r and d , and taking into account that $d = 0$ if $s_{Q1} = 0$, the generic element of the queue transition matrix is defined as follows:

$$P_{[s_{Q1}, s_{Q2}]}^{(Q)}(r, d) = \begin{cases} 1 & \text{if } \min \{ s_{Q1} - d + r, K \} = s_{Q2} \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

3. EE-ARQ Retransmission Counter Model

Now we model the EE-ARQ mechanism. To this end, let us describe how the

retransmission counter state is updated. Note that the retransmission counter state transition depends on the evolution of the service processes $N(p)$ and the queue state $S^{(Q)}(p) = s_{Q2}$ in the p -th slot. Thus the generic element of the transition probability matrix $P_{[s_{R1}, s_{R2}]}^{(R)}(s_{Q2}, d)$ corresponding to the transition from the state $S^{(R)}(p-1) = s_{R1}$ to the state $S^{(R)}(p) = s_{R2}$, is based on the values s_{Q2} and d . Summarizing, we have:

$$P_{[s_{R1}, s_{R2}]}^{(R)}(s_{Q2}, d) = \begin{cases} 1 & \text{if } s_{R1} = 'w' & s_{R2} = 'w' & s_{Q2} = 0 & d = 0 \\ 1 & \text{if } s_{R1} = \rho & s_{R2} = 'w' & s_{Q2} = 0 & d = 1 & \forall \rho \in \mathbf{R} \\ \wp^{(Tx)}(0) & \text{if } s_{R1} = 'w' & s_{R2} = 0 & s_{Q2} > 0 & d = 0 \\ 1 - \wp^{(Tx)}(0) & \text{if } s_{R1} = 'w' & s_{R2} = 'w_0' & s_{Q2} > 0 & d = 0 \\ \wp^{(Tx)}(\rho + 1) & \text{if } s_{R1} = \rho & s_{R2} = \rho + 1 & s_{Q2} > 0 & d = 0 & \forall \rho \in \mathbf{R}^- \\ 1 - \wp^{(Tx)}(\rho + 1) & \text{if } s_{R1} = \rho & s_{R2} = 'w_{\rho+1}' & s_{Q2} > 0 & d = 0 & \forall \rho \in \mathbf{R}^- \\ \wp^{(Tx)}(\rho) & \text{if } s_{R1} = 'w_{\rho}' & s_{R2} = \rho & s_{Q2} > 0 & d = 0 & \forall \rho \in \mathbf{R} \\ 1 - \wp^{(Tx)}(\rho) & \text{if } s_{R1} = 'w_{\rho}' & s_{R2} = 'w_{\rho}' & s_{Q2} > 0 & d = 0 & \forall \rho \in \mathbf{R} \\ \wp^{(Tx)}(0) & \text{if } s_{R1} = \rho & s_{R2} = 0 & s_{Q2} > 0 & d = 1 & \forall \rho \in \mathbf{R} \\ 1 - \wp^{(Tx)}(0) & \text{if } s_{R1} = \rho & s_{R2} = 'w_0' & s_{Q2} > 0 & d = 1 & \forall \rho \in \mathbf{R} \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

where we have indicated as \mathbf{R} the set of values $\{0, \dots, s_{MAX}^{(R)}\}$ and \mathbf{R}^- the set of values $\{0, \dots, s_{MAX}^{(R)} - 1\}$.

4. Wireless Channel Model

Finally, let us describe the wireless channel behavior. As usual (e.g. see [32]), it is described by means of an M -state Markov process. In this way, the channel statistical behavior can be characterized by the $M \times M$ transition probability matrix of the Markov chain, $P^{(C)}$, and the bit error rate, $BER_{[s_C]}$, associated to each state s_C . Assuming that losses of bits are in the channel uniformly distributed during the time slot, the ARQ block error probability when the channel state is s_C can be calculated as follows:

$$\wp_{[s_C]}^{(ABE)} = 1 - (1 - BER_{[s_C]})^H \quad (21)$$

where $(1 - BER_{[s_C]})^H$ is the probability that the block is transmitted correctly.

5. Performance Evaluation

In this section we introduce the main performance parameters for the proposed system, and analytically derive them. First, from the system steady-state probability array let us calculate the probability arrays regarding the queue length and the queue drain processes:

$$\pi_{[s_Q]}^{(Q)} = Prob\{S^{(Q)}(p) = s_Q\} = \sum_{\underline{s}_W} \sum_{\underline{s}_N} \pi_{[\underline{s}_W, s_Q, \underline{s}_N]}^{(\Sigma)} \quad (22)$$

$$\pi_{[\underline{s}_N]}^{(N)} = Prob\{S^{(N)}(p) = \underline{s}_N\} = \sum_{\underline{s}_W} \sum_{s_Q} \pi_{[\underline{s}_W, s_Q, \underline{s}_N]}^{(\Sigma)} \quad (23)$$

Now, using the above distributions, we can calculate the ARQ block loss probability defined as the probability that the ARQ protocol discards a block since unsuccessfully retransmitted for $s_{MAX}^{(R)}$ times:

$$\wp_{Loss_{ARQ}} = \sum_{s_C} \wp_{[s_C]}^{(ABE)} \pi_{[s_{MAX}^{(R)}, s_C]}^{(N)} \quad (24)$$

where we have considered that $\underline{s}_N = (s_R, s_C)$. Therefore $\wp_{Loss_{ARQ}}$ is given by the channel error probability $\wp_{[s_C]}^{(ABE)}$ of each channel state s_C , multiplied by the probability that the channel state is s_C and the retransmission counter state is $s_{MAX}^{(R)}$.

Another important parameter that can be derived from the above distributions is the probability that the buffer empties or not empties:

$$\begin{aligned}\wp_{QueueEmpty} &= Prob\{S^{(Q)}(p) = 0\} = \pi_{[0]}^{(Q)} \\ \wp_{QueueNotEmpty} &= 1 - \wp_{QueueEmpty} = \sum_{s_Q=1}^K \pi_{[s_Q]}^{(Q)}\end{aligned}\quad (25)$$

Assuming that video sources are enough adaptive, the two above metrics, $\wp_{Loss_{ARQ}}$ and $\wp_{QueueNotEmpty}$, can be used to quantify the capacity of the system to maximize link utilization when the channel gets better quality, and avoid losses when the channel get worse. For example, the loss probability can be used to find the minimum acceptable channel quality that avoids buffer saturation, given a minimal video bitrate.

Let us now derive the total power consumption as follows:

$$\bar{\mathcal{P}} = \bar{\mathcal{P}}_{Tx_{SUCCESS}} + \bar{\mathcal{P}}_{Tx_{FAILURE}} + \bar{\mathcal{P}}_{IDLE} \quad (26)$$

where $\bar{\mathcal{P}}_{Tx_{SUCCESS}} = \mathcal{P}_{Tx} \cdot \wp_{Tx_{SUCCESS}}$, $\bar{\mathcal{P}}_{Tx_{FAILURE}} = \mathcal{P}_{Tx} \cdot \wp_{Tx_{FAILURE}}$, and

$\bar{\mathcal{P}}_{IDLE} = \mathcal{P}_{IDLE} \cdot \wp_{IDLE}$ represent the power consumed during a successful transmission, a failed transmission and an ARQ transmitter IDLE state (corresponding to a state when the transmitter is idle because the queue is empty, or the EE-ARQ algorithm imposes the transmitter to not transmit in order to save power), respectively.

The terms \mathcal{P}_{Tx} and \mathcal{P}_{IDLE} are input parameters, representing the power consumed in one slot to transmit an ARQ block or to be in the IDLE state, respectively. The terms $\wp_{Tx_{SUCCESS}}$, $\wp_{Tx_{FAILURE}}$ and \wp_{IDLE} are the probabilities of a successful transmission, a failed transmission and a transmitter IDLE state, respectively. They can be calculated as follows:

$$\wp_{Tx_{SUCCESS}} = \sum_{s_C} \sum_{s_R} (1 - \wp_{[s_C]}^{(ABE)}) \cdot \pi_{[s_R, s_C]}^{(N)} \quad (27)$$

$$\wp_{Tx_{FAILURE}} = \sum_{s_C} \sum_{s_R} \wp_{[s_C]}^{(ABE)} \cdot \pi_{[s_R, s_C]}^{(N)} \quad (28)$$

$$\wp_{IDLE} = 1 - \wp_{Tx} = 1 - (\wp_{Tx_{SUCCESS}} + \wp_{Tx_{FAILURE}}) \quad (29)$$

6. Numerical results

In this section we apply the proposed model to numerically evaluate the performance of the whole Green Adaptive Video Wireless Transmission system when it applied the proposed

EE-ARQ algorithm. The evaluation will be carried out in terms of both quality of service and power saving, with the purpose of showing how the combination of the EE-ARQ protocol with the source Rate Controller reduces waste of energy while assuring good QoS. Moreover, we will show how the analytical framework proposed in this paper allows to deduce some design guidelines on the choice of the retransmission policy parameter γ , according to the given requirements specified in terms of QoS and power saving. In the following we consider a video sequence in CIF format (352 x 288 pixels), related to a documentary captured from the BBC International Television, and encoded using the vcodec mpeg2video with the ffmpeg encoder. We have applied $G = 5$ different encoding levels ($\underline{\Psi} = \{80, 120, 160, 200, 240\}$ kbit/s). The correspondent PSNR values are: $\underline{\Omega} = \{39.5, 44.0, 46.0, 47.5, 49.0\}$ dB. We have set a number of sources of $V = 10$ and we have first considered the case where the traffic load is produced by sources encoding at constant bitrate, and then the case where the Rate Controller is applied. We consider a Rate Controller with a V-shaped feedback mask, i.e. $q_L = q_H = K/2$, which represents the most reactive feedback mask, as demonstrated in [43], and we set the parameter $\phi = 1/V$ in such a way that only one source at time is requested to change its encoding rate. The buffer size is set to $K = 25$ ARQ blocks, each ARQ block is constituted by $H = 512$ bits and the maximum number of retransmissions is $s_{MAX}^{(R)} = 7$. As far as the wireless transmission side is concerned, we set C to 5 Mbit/s; therefore the slot duration is $\Delta_{ARQ} = H/C = 0.1$ ms. According to [7], we assumed that the wireless interface consumes a power of $P_{Tx} = 1350$ mW during transmission, and $P_{IDLE} = 66$ mW when it is idle. Although the analytical model is able to capture both first- and second-order statistics of the channel behavior, here we present an analysis for a constant BER, equal to $1 \cdot 10^{-3}$; in this way our results are independent on the channel time correlation.

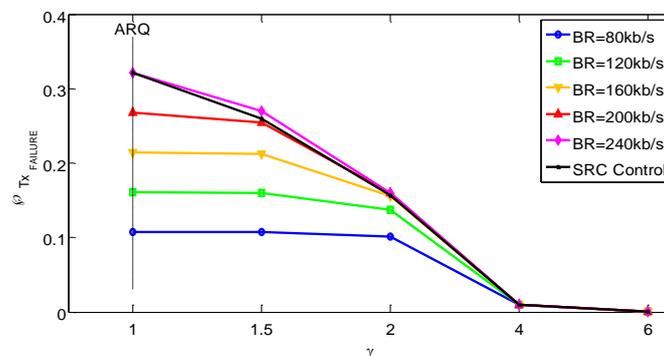


Fig. 3 Probability of Transmission Failure

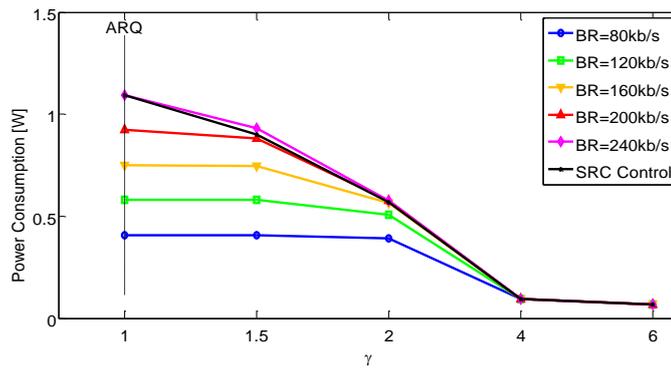


Fig. 4 Power consumption

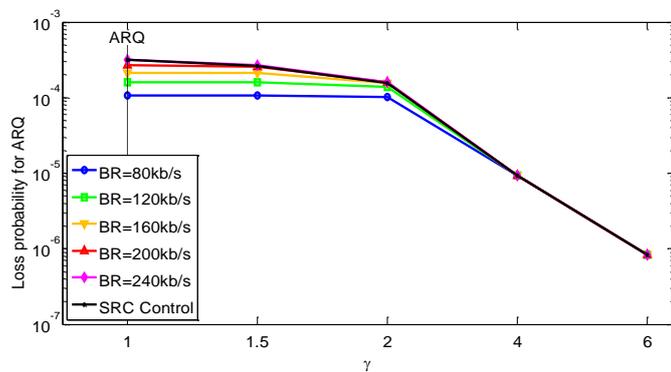


Fig. 5 Loss probability

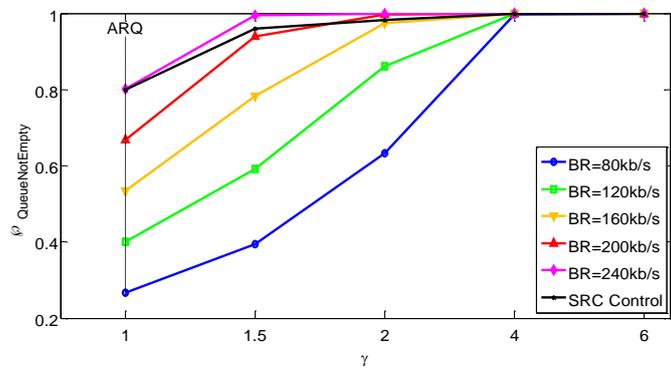


Fig. 6 Probability of Queue not empty

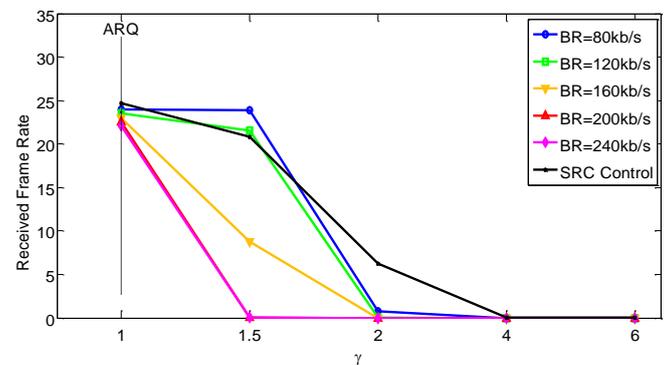


Fig. 7 Received Frame Rate

In order to deeply analyze cross-layer effects of the proposed algorithm, we compare results with different cases of constant bit rate encoding, when the encoder uses a target bit rate ranging between 80 kbit/s (the worst encoding quality but the lowest output bit rate) and 240 kbit/s (the best encoding quality but the highest output bit rate).

First we evaluate the capability of saving energy by considering the number of transmission failures. Fig. 3 and Fig. 4 show the probability of transmission failure and the total power consumption, for both constant-bit-rate and rate-controlled cases, against the γ parameter value. Note that for $\gamma=1$, the retransmission policy assures that packets are always transmitted with probability 1, so this case coincides with the classical ARQ. Results in Fig. 3 and Fig. 4 show that the proposed retransmission policy allows to reduce the probability of failure so reducing the power consumption with respect to the case of the classical ARQ. Moreover the higher the γ parameter value, the higher the energy saved.

In order to quantify the capacity of the system to both avoid ARQ loss when the channel get worse, and maximize link utilization when the channel gets better quality, Fig. 5 and Fig. 6 present the loss probability due to ARQ block discarding because the maximum number of retransmissions are exceeded, and the probability that the queue is not empty, respectively. As we can notice, the proposed mechanism improves the classical ARQ performance, and the improvement increases with the γ value. More specifically, the ARQ loss probability decreases and the link utilization increases for increasing values of γ . This occurs for any kind of video encoding, both constant-rate and adaptive-rate encoding. This is due to the property of our mechanism to learn the state of the channel, and therefore transmit when the channel state is better.

Of course, the counterpart of this approach is that the service rate of the buffer decreases with γ , therefore causing some losses. In order to evaluate the overall performance perceived at destination, we derived the received frame rate. It is displayed in Fig. 7. Analyzing this figure we can notice that EE-ARQ presents a destination frame rate greater than high encoding bitrates. If compared with low rate encoding sources it has a lower received frame rate, but presents a better encoding quality. Finally, as compared with the classical ARQ, the lower received frame rate is compensated with the power save shown in Fig. 4.

A peculiarity of our approach is that the designer can choose the value of γ to find the desired tradeoff between received quality and consumed power. In addition, another important characteristic that is worth attention is that EE-ARQ is able to learn the channel state. Therefore, when the channel is bad, as the case studied so far, EE-ARQ saves power while maintaining perceived performance acceptable; instead, when the channel state is good, it works like classical ARQ, but not wasting power because retransmissions are not needed.

In this context the Rate Controller plays an important role, because it adapts video encoding to decrease encoding quality to minimize losses, and so maximize frame rate perceived at destination. Therefore the source Rate Controller, when coupled with EE-ARQ, allows to tune the encoding bit rate of the source aggregate in order to provide the best video quality with the maximum received frame rate, while reducing power consumption with respect to the classical ARQ.

7. Related Work

Although energy efficiency of wireless devices has been one of the principal research topic in the literature of telecommunications networks in at least the last decade, a main focus was on how to save energy in low-power battery devices, with the aim of increasing their lifetime [13, 4]. Most of the pioneering research in the area of energy-saving wireless communications has focused on transmission schemes to minimize the transmission energy per bit. In [42], optimal strategies are illustrated to minimize the energy per bit required for reliable transmission in the wideband regime. The work in [11] proposes an optimal scheduling algorithm to minimize transmission energy by maximizing the transmission time for buffered packets. Other scheduling techniques are proposed in [26] and [23] to minimize the transmission energy. Energy saving techniques for short distance wireless communications are studied in [5, 6, 12]. The work in [30] analyzes the energy efficiency of three classes of Hybrid-ARQ (HARQ) protocols, considering the packet transmission time as a degree of freedom that can be optimized in order to maximize energy-efficiency.

However, most of the above papers assume constant channel conditions, and none of them accounts the impact of the considered techniques on the performance perceived at the application level. On the contrary, one of the key features of this paper is the cross-layer energy-efficient approach for the transmission of video traffic over wireless channels. Many previous works considered this issue (see for example [27, 21, 8, 10, 9]), but at the best of our knowledge this is the first work that proposes a system to control energy efficiency of the wireless channel, and performance at both the transmission and the application levels simultaneously, making adaptive both the wireless transmitter and the video sources; this goal is achieved thanks to the introduction of a source Rate Controller at the application level and a transmission law at the link level. In addition this paper provides some guidelines to support system design. The same guidelines can be used to modify system feedback laws and parameters run-time according to the time-varying channel behavior.

Another important topic covered by the paper is the model of the whole system. Many analytical models of ARQ transmission systems [15, 16, 2, 1], data multiplexer [31, 17] and multimedia sources [14, 18] are present in the past literature but, at the best of our knowledge, this is the first model that is able to capture the cross-layer behavior of a green transmission system constituted by a multiplexer of video sources served by an energy-efficient ARQ-based wireless link.

8. Conclusions and Future Works

In this paper we have proposed a cross-layer approach for the transmission of multiplexed rate-controlled multimedia streams over wireless channels. The proposed approach is adaptive in both the video sources and the wireless transmitter. In order to compensate transmission bandwidth reduction due to the energy saving policies, a source Rate Controller is introduced. An analytical model of the system has been defined and a numerical analysis has been used as an example on how to derive some guidelines to choose the design parameters of the proposed transmission law. However, the analytical framework proposed in this paper, is more general and can be used to evaluate and design different energy-saving retransmission policies and different feedback masks for the source Rate-Controller. As said in Section 5, evaluation of some minimum bound for the channel quality that avoids buffer saturation, and maximum bound for the video encoding bitrate that avoids link bandwidth waste can be an interesting application of the proposed model in future research.

Acknowledgement

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n. 257740 (Network of Excellence "TREND").

References

- [1] K. Ausavapattanakun and A. Nosratinia, "Analysis of selective-repeat arq via matrix signal-flow graphs," *Communications, IEEE Transactions on*, vol. 55, no. 1, pp. 198–204, 2007. <http://dx.doi.org/10.1109/TCOMM.2006.885092>
- [2] L. Badia, M. Rossi, and M. Zorzi, "Queueing and delivery analysis of sr arq on markov channels with non-instantaneous feedback," in *Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE*, vol. 6, 2005, pp. 5 pp.–3721. <http://dx.doi.org/10.1109/GLOCOM.2005.1578465>
- [3] A. Berl, E. Gelenbe, M. Di Girolamo, G. Giuliani, H. De Meer, M. Q. Dang, and K. Pentikousis, "Energy-efficient cloud computing," *The Computer Journal*, vol. 53, no. 7, pp. 1045–1051, 2010. <http://dx.doi.org/10.1093/comjnl/bxp080>
- [4] S. Chen, P. Sinha, N. Shroff, and C. Joo, "A simple asymptotically optimal energy allocation and routing scheme in rechargeable sensor networks," in *INFOCOM, 2012 Proceedings IEEE*, march 2012, pp. 379–387. <http://dx.doi.org/10.1109/INFCOM.2012.6195775>
- [5] S. Cui, A. Goldsmith, and A. Bahai, "Energy-efficiency of mimo and cooperative mimo techniques in sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 6, pp. 1089–1098, 2004. <http://dx.doi.org/10.1109/JSAC.2004.830916>
- [6] Shuguang Cui, A. J. Goldsmith, A. Bahai, , "Energy-constrained modulation optimization," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 5, pp. 2349–2360, 2005. <http://dx.doi.org/10.1109/TWC.2005.853882>

- [7] L. M. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad hoc networking environment," in *In IEEE Infocom, 2001*, pp. 1548–1557.
- [8] P. Frossard, C. Chen, C. Sreenan, K. Subbalakshmi, D. Wu, and Q. Zhang, "Guest editorial cross-layer optimized wireless multimedia communications," *Selected Areas in Communications, IEEE Journal on*, vol. 25, no. 4, pp. 641–644, 2007. <http://dx.doi.org/10.1109/JSAC.2007.070501>
- [9] L. Galluccio, F. Licandro, G. Morabito, and G. Schembra, "An analytical framework for the design of intelligent algorithms for adaptive-rate mpeg video encoding in next generation time-varying wireless networks," *IEEE Journal on Selected Areas of Communications*, vol. 23, pp. 369–384, February 2005.
- [10] L. Galluccio, G. Morabito, and G. Schembra, "Transmission of adaptive mpeg video over time-varying wireless channels: Modeling and performance evaluation," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 2777–2788, November 2005.
- [11] A. E. Gamal, R. Nair, B. Prabhakar, E. Uysal-biyikoglu, and S. Zahedi, "Energy-efficient scheduling of packet transmissions over wireless networks," in *in Proc. INFOCOM Conf*, pp. 1773–1783.
- [12] A. Goldsmith and S. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *Wireless Communications, IEEE*, vol. 9, no. 4, pp. 8–27, 2002. <http://dx.doi.org/10.1109/MWC.2002.1028874>
- [13] C. Guo, R. V. Prasad, P. Pawelczak, and R. Hekmat, "Designing energy efficient automatic repeat request protocol in wireless sensor networks," in *Proceedings of the 4th ACM workshop on Challenged networks, ser. CHANTS '09. 1em plus 0.5em minus 0.4em* New York, NY, USA: ACM, 2009, pp. 35–42. <http://dx.doi.org/10.1145/1614222.1614229>
- [14] H. Hefes and D. Lucantoni, "A markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance," *Selected Areas in Communications, IEEE Journal on*, vol. 4, no. 6, pp. 856–868, 1986. <http://dx.doi.org/10.1109/JSAC.1986.1146393>
- [15] T. Issariyakul, E. Hossain, and A. Alfa, "Markov-based analysis of end-to-end batch transmission in a multi-hop wireless network," in *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, vol. 5, 2005, pp. 3505–3509 Vol. 5. <http://dx.doi.org/10.1109/ICC.2005.1495071>
- [16] T. Issariyakul, E. Hossain, A.S. Alfa, "Analysis of latency for reliable end-to-end batch transmission in multi-rate multi-hop wireless networks," in *IEEE International Conference on Communications 2005. ICC 2005. 2005*, vol. 5, 2005, pp. 3494–3498 Vol. 5. <http://dx.doi.org/10.1109/ICC.2005.1495069>
- [17] A. Lombardo, G. Morabito, and G. Schembra, "An accurate and treatable markov model of mpeg-video traffic," in *INFOCOM '98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 1, 1998, pp. 217–224 vol.1. <http://dx.doi.org/10.1109/INFCOM.1998.659657>
- [18] A. Lombardo, G. Morabito, G. Schembra, "Modeling intramedia and intermedia relationships in multimedia network analysis through multiple timescale statistics," *IEEE Transactions on Multimedia*, vol. 6, no. 1, pp. 142–157, 2004.

<http://dx.doi.org/10.1109/TMM.2003.819750>

[19]C. Panarello, A. Lombardo, G. Schembra, L. Chiaraviglio, and M. Mellia, “Energy saving and network performance: a trade-off approach,” in Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking, ser. e-Energy '10. New York, NY, USA: ACM, 2010, pp. 41–50. <http://dx.doi.org/10.1145/1791314.1791321>

[20]P. Mahadevan, P. Sharma, S. Banerjee, and P. Ranganathan, “A power benchmarking framework for network devices,” in Proceedings of the 8th International IFIP-TC 6 Networking Conference, ser. NETWORKING '09. 1em plus 0.5em minus 0.4em Berlin, Heidelberg: Springer-Verlag, 2009, pp. 795–808. http://dx.doi.org/10.1007/978-3-642-01399-7_62

[21]G. Miao, N. Himayat, Y. G. Li, and A. Swami, “Cross-layer optimization for energy-efficient wireless communications: a survey,” *Wirel. Commun. Mob. Comput.*, vol. 9, no. 4, pp. 529–542, Apr. 2009. <http://dx.doi.org/10.1002/wcm.v9:4>

[22]S. Nedeveschi, L. Popa, G. Iannaccone, S. Ratnasamy, and D. Wetherall, “Reducing network energy consumption via sleeping and rate-adaptation,” in Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation, ser. NSDI'08. 1em plus 0.5em minus 0.4em Berkeley, CA, USA: USENIX Association, 2008, pp. 323–336.

[23]P. Nuggehalli, V. Srinivasan, and R. Rao, “Delay constrained energy efficient transmission strategies for wireless devices,” in INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 3, 2002, pp. 1765–1772 vol.3. <http://dx.doi.org/10.1109/INFCOM.2002.1019430>

[24]A. Lombardo, C. Panarello, and G. Schembra, “Achieving energy savings and qos in internet access routers,” *SIGMETRICS Perform. Eval. Rev.*, vol. 38, no. 3, pp. 76–80, Jan. 2011. <http://dx.doi.org/10.1145/1925019.1925035>

[25]A. Lombardo and G. Schembra, “Performance evaluation of an adaptive-rate mpeg encoder matching intserv traffic constraints,” *IEEE/ACM Trans. Netw.*, vol. 11, no. 1, pp. 47–65, Feb. 2003. <http://dx.doi.org/10.1109/TNET.2002.804830>

[26]C. Schurgers and M. Srivastava, “Energy efficient wireless scheduling: adaptive loading in time,” in Wireless Communications and Networking Conference, 2002. WCNC2002. 2002 IEEE, vol. 2, 2002, pp. 706–711 vol.2. <http://dx.doi.org/10.1109/WCNC.2002.993354>

[27]G. Shah, W. Liang, and X. Shen, “Cross-layer design for qos support in wireless multimedia sensor networks,” in Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE, 2010, pp. 1–5. <http://dx.doi.org/10.1109/GLOCOM.2010.5683279>

[28]S. Sharangi, R. Krishnamurti, and M. Hefeeda, “Energy-efficient multicasting of scalable video streams over wimax networks,” *IEEE Transactions on Multimedia*, vol. 13, pp. 102–115, February 2011. <http://dx.doi.org/10.1109/TMM.2010.2076799>

[29]K. Son and B. Krishnamachari, “Speedbalance: Speed-scaling-aware optimal load balancing for green cellular networks,” in INFOCOM, 2012 Proceedings IEEE, march 2012, pp. 2816–2820. <http://dx.doi.org/10.1109/INFCOM.2012.6195707>

[30]I. Stanojev, O. Simeone, Y. Bar-Ness, and D. Kim, “On the energy efficiency of hybrid-arq protocols in fading channels,” in Communications, 2007. ICC '07. IEEE International Conference on, june 2007, pp. 3173–3177. <http://dx.doi.org/10.1109/ICC.2007.526>

- [31] D. Tse, R. Gallager, and J. Tsitsiklis, "Statistical multiplexing of multiple time-scale markov streams," *Selected Areas in Communications, IEEE Journal on*, vol. 13, no. 6, pp. 1028–1038, 1995. <http://dx.doi.org/10.1109/49.400658>
- [32] B. Vucetic, "An adaptive coding scheme for time-varying channels," *Communications, IEEE Transactions on*, vol. 39, no. 5, pp. 653–663, may 1991. <http://dx.doi.org/10.1109/26.87156>
- [33] J. Xu, X. Shen, J. Mark, and J. Cai, "Adaptive transmission of multi-layered video over wireless fading channels," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2305–2314, June 2007. <http://dx.doi.org/10.1109/TWC.2007.05838>
- [34] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti, "Energy efficiency in the future internet: A survey of existing approaches and trends in energy-aware fixed network infrastructures," *Communications Surveys Tutorials, IEEE*, vol. 13, no. 2, pp. 223–244, 2011. <http://dx.doi.org/10.1109/SURV.2011.071410.00073>
- [35] R. Bruschi, A. Lombardo, C. Panarello, F. Podda, G. E. Santagati, and G. Schembra, "Active window management: Reducing energy consumption of TCP congestion control," in *IEEE ICC 2013 - Selected Areas in Communications Symposium (ICC'13 SAC)*, Budapest, Hungary, jun 2013, pp. 2747–2751, udR Catania, TREND.
- [36] C. Panarello, M. Mellia, M. Marsan, M. Meo, A. Lombardo, and G. Schembra, "On the intertwining between capacity scaling and tcp congestion control," in *Future Energy Systems: Where Energy, Computing and Communication Meet (e-Energy)*, 2012 Third International Conference on, 2012, pp. 1–4.
- [37] C. E. Jones, K. M. Sivalingam, P. Agrawal, and J.-C. Chen, "A survey of energy efficient network protocols for wireless networks," *Wireless Networks*, vol. 7, pp. 343–358, 2001.
- [38] A. Kamal, "Efficient solution of multiple server queues with application to the modeling of atm concentrators," in *Proceedings of IEEE INFOCOM 1996*, vol. 1, mar 1996, pp. 248–254 vol.1. <http://dx.doi.org/10.1109/INFCOM.1996.497900>
- [39] K. S. Kumar, R. Chandramouli, and K. P. Subbalakshmi, "On stochastic learning in predictive wireless arq," *Wirel. Commun. Mob. Comput.*, vol. 8, pp. 871–883, September 2008. <http://dx.doi.org/10.1002/wcm.v8:7>
- [40] A. Lombardo, C. Panarello, D. Reforgiato, and G. Schembra, "Measuring and modeling energy consumption to design a green netfpga giga-router," in *Global Communications Conference (GLOBECOM)*, 2012 IEEE, 2012, pp. 3062–3067. <http://dx.doi.org/10.1109/GLOCOM.2012.6503584>
- [41] A. Lombardo, D. Reforgiato, V. Riccobene, and G. Schembra, "Modeling temperature and dissipation behavior of an open multi-frequency green router," in *In Proc. of IEEE GreenCom 2012, The Second IEEE Online Conference on Green Communications*, September 2012.
- [42] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theor.*, vol. 48, no. 6, pp. 1319–1343, Sep. 2006. <http://dx.doi.org/10.1109/TIT.2002.1003824> =Opt
- [43] F. Licandro, C. Panarello, and G. Schembra, "A rate-controlled voip system based on wireless mesh network infrastructure: Design issues and performance analysis," in *Wireless Communications 2007 CNIT Tyrrhenian Symposium*, ser. Signals and Communication Technology, S. Pupolin, Ed. Springer US, 2007, pp. 235–243.

http://dx.doi.org/10.1007/978-0-387-73825-3_19

[44]M. Zorzi, R. R. Rao, and S. Member, “Error control and energy consumption in communications for nomadic computing,” IEEE Transactions on Computers, vol. 46, pp. 279–289, 1997.

[45]M. Zorzi and R. R. Rao, “Energy constrained error control for wireless channels,” IEEE Personal Communications, vol. 4, pp. 27–33, 1997.

Copyright Disclaimer

Copyright reserved by the author(s).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).