QoS-Aware Resource Allocation for Slowly Time-Varying Channels

Guerino Giancola, Luca De Nardis, Maria-Gabriella Di Benedetto
Info-Com Department
University of Rome La Sapienza
Via Eudossiana, 18 – 00184, Rome, Italy
giancola@infocom.uniroma1.it

Abstract—**In this paper the problem of optimizing transmission efficiency at the Medium Access Control (MAC) level when considering applications requiring Quality of Service (QoS) constraints is addressed. The analysis focuses on the design of error protection mechanisms, and an analytical approach for selecting and dimensioning error protection for both real-time and non-real-time services is proposed. This approach is formalized as an optimization problem, and is based on the assumption of a slowly time-varying channel, as typical in Wireless Local Area Networks (WLANs). The result of the analysis is an iterative algorithm which is capable of minimizing required capacity according on both the state of the channel and the requested QoS.**

Keywords: MAC, Error Protection, QoS, WLAN.

I. INTRODUCTION

Two basic problems in the design of Medium Access Control (MAC) protocols with Quality of Service (QoS) are the efficient management of the resource, and the need for fulfilling QoS requirements despite the unpredictable behaviour of the channel. If the MAC is capable of adjusting error protection to the time-varying nature of the channel, then the above problems can be strongly mitigated.

In the ideal scenario where all the transmitted data units correctly reach destination, the MAC module is able to fulfil QoS constraints by simply reserving a portion of the available capacity for each source which is admitted in the system. This portion can be easily evaluated analytically, and in general the MAC module needs only to identify the set of traffic descriptors and QoS parameters for the source under examination. Since each transmitted data unit is received without any corruption, the only event which causes loss of information is buffer overflow. Therefore, if the MAC module is able to avoid buffer filling, QoS fulfilment can be guaranteed at admission time for the whole duration of the connection.

When considering a channel which introduces errors on the transmitted units, mechanisms of either retransmission (i.e. Automatic Repeat on reQuest, ARQ) or error correction (i.e. Forward Error Correction, FEC), or both, may be needed in the MAC in order to improve system efficiency [1]-[3]. ARQ mechanisms are based on the repetition of corrupted MAC Protocol Data Units (MACPDUs) [4]. In this case, error detecting codes, such as a Cyclic Redundancy Check (CRC), are embedded in each transmitted data unit, and are used by the receiver to detect errors on incoming data packets. FEC schemes introduces redundancy in each MACPDU in order to provide the receiver with the capability to correct a certain number of errors [5]. FEC schemes may be also combined within ARQ mechanisms, giving rise to the so-called Hybrid ARQ [6]. The solution based on the ARQ introduces delays which might be incompatible with real-time features. On the other hand, the soultion based on the FEC has the drawback of requiring overhead transmission, and therefore efficiency loss. As a consequence, transmission efficiency can be guaranteed only when the MAC protocol is capable of adjusting error protection according to both the time-varying nature of the channel and the characteristics of those applications which are admitted to the system.

In this paper, we propose an analytical approach for selecting and designing error protection schemes for traffic sources requiring QoS. The resulting algorithm works with both real-time and non-real-time traffic sources, and it maximizes transmission efficiency by selecting and dimensioning an error protection mechanism which takes into account both channel status and QoS constraints. The paper is organized as follows. Section 2 illustrates the reference scenario. Section 3 describes the proposed algorithm for error protection design and resource allocation. Section 4 evaluates performance by introducing a time-varying channel model. Finally, Section 5 contains the conclusions.

II. REFERENCE SCENARIO

Each traffic source is characterized by two sets of parameters. The first, denoted *Tspecs*, collects parameters describing source traffic activity. It can consists for example of the Dual Leaky Bucket (DLB) parameters described in [7], i.e. the peak rate of the flow p (bit/s), the average rate of the flow r (bit/s), the token buffer dimension b (bits), and the maximum source packet size M (bits). The second set of parameters, denoted *Qspecs*, defines two QoS requirements: the maximum tolerable end-to-end delay D_{MAX} (s), and the minimum percentage of packets F (0÷100) required at destination within D_{MAX}. Note that the same parameters are used for both real-time and non-real-time services, with no explicit need for defining classes of traffic.

The proposed MAC protocol uses fixed-size MACPDUs of L_{PDU} bits, composed of a fixed-size *header* of L_H bits and a fixed-size *payload* of L_P bits. The *header* contains the information used by the MAC for managing the transmission of a MACPDU. It may contain error detection codes such as CRC, but no FEC. The *payload* conveys bits originating from source packets segmentation, and redundancy bits eventually introduced by the FEC. In other words, we assume that the introduction of corrective overhead is realized by removing the corresponding bits from the MACPDU payload. The payload is therefore composed of two parts: a FEC field of L_{FEC} bits, and an effective payload for user data of $L_{EFF}=L_P-L_{FEC}$ bits. Note that while L_{FEC} and L_{EFF} may vary in different MACPDUs, L_{PDU}, L_H, and L_P are fixed. With reference to the ARQ, a Selective Repeat (SR) strategy is implemented in order to avoid unnecessary re-transmissions which could affect simulation results.

Resource allocation is based on the definition of a MAC frame of D_F secs. D_{sys} is the maximum system delay introduced by the MAC for the transmission of a single MACPDU.

We assume a slowly time-varying channel characterized by a Bit Error Rate (BER) indicated by p_b. We also assume that the transmitter knows the exact value of p_b by estimation of the reverse channel. We discard incorrectable errors in the *header* field, and restrict the present analysis to error protection on the *payload.*

III. QoS-Aware Resource Allocation

A. Error Protection design

The proposed algorithm for optimizing resource allocation is parametrical on the number of retransmissions N_R which are allowed by the ARQ. For any N_R between 0 (i.e. no ARQ) and a maximum value denoted $N_R^{(\max)}$, the algorithm evaluates the amount of capacity which is required by the MAC for fulfilling QoS. Transmission efficiency is therefore guaranteed by selecting the N_R value leading to the lower amount of capacity.

The procedure starts by expressing in analytical terms the trade-off which exists between the amount of reserved capacity C (bit/s) and the delay D (s) which is experienced by the transmitted MACPDUs. This task is exploited by introducing two basic functions: the Capacity function *X(Tspecs,D)* and the Delay function *Δ(Tspecs,C)*. The Capacity function *X(Tspecs,D)* is the capacity in bit/s which is necessary for guaranteeing a maximum delay D to a generic source described by *Tspecs*. The Delay function *Δ(Tspecs,C)* evaluates the end-to-end delay when the MAC reserves the capacity C. If for example *Tspecs* are expressed in terms of DLB parameters, both the Capacity function and the Delay function can be derived based on [7], and write as follows:

$$\mathrm{X}(Tspecs, D) = \frac{p\cdot b - r\cdot M}{(D-D_{sys})\cdot(p-r)+b-M} \tag{1}$$

$$\Delta(Tspecs, C) = \left[\frac{p-C}{p-r}(b-M)+M\right]\cdot\frac{1}{C}+D_{sys} \tag{2}$$

Tspecs are used by the MAC for predicting the number of bits emitted by the source in a generic time interval, and based of this data for evaluating the minimum bit rate $R_{\min}$ which is necessary in order to avoid overflow of the source buffer. In the DLB case, $R_{\min}$ is equal to the token buffer rate r, i.e. to the average flow rate. The MAC operates by first assigning $R_{\min}$, which in turn leads to the following delay:

$$D_0 = \Delta(Tspecs, R_{\min}) \tag{3}$$

Qspecs however impose that D_0 be lower than D_{MAX} . In case $D_0>D_{MAX}$, the MAC understands that R_{min} is too low. The general rule for required capacity is therefore:

$$C = \mathrm{X}(Tspecs, \min\{D_{MAX}, \Delta(Tspecs, R_{\min})\}) \tag{4}$$

If however errors are present during transmission, equation (4) must be modified in order to take into account the presence of ARQ, i.e:

$$C = \mathrm{X}(Tspecs, \min\{D_{MAX} - N_R\cdot RTT, \Delta(Tspecs, R_{\min})\}) \tag{5}$$

where *RTT* is the estimated round-trip-time, i.e. the time necessary for the retransmission request plus the time needed for the retransmission of a MACPDU. According to (5), the maximum value for N_R is given by:

$$N_R \le N_R^{(\max)} = \left\lfloor \frac{D_{MAX}-D_{sys}}{RTT}\right\rfloor \tag{6}$$

where $\lfloor x \rfloor$ is the inferior integer part of x.

Because of the presence of both the MACPDU header and the FEC field, the capacity C defined in (5) leads to the following effective capacity C_{eff}:

$$C_{eff} = C\cdot\frac{L_{PDU}}{M}\cdot\left\lceil\frac{M}{L_{PDU}-L_{FEC}}\right\rceil \tag{7}$$

where M is the maximum source packet size and where $\lceil x \rceil$ is the superior integer part of x. The value of M can be eventually found in *Tspec*s (as in the DLB case), but can also be dynamically estimated by the MAC in correspondance to each admitted source. For a given source (i.e. given *Tspecs* and *Qspecs*) the effective capacity derived in (7) depends on both the value of N_R and the size of the FEC field. The following step of the proposed procedure therefore consists in evaluating the size of L_{FEC} as a function of N_R.

In order to proceed with the FEC design, the MAC must operate the conversion of F, i.e. the minimum percentage of source packets which should be correctly delivered within D_{MAX}, into a MAC parameter. If we denote with M the maximum source packet size (in bits), we can translate the constraint given by F into the following maximum tolerable MACPDU loss probability P_L:

$$P_L \le 1-\left(\frac{F}{100}\right)^{\frac{1}{\lceil M/L_{EFF}\rceil}} \tag{8}$$

If both FEC and ARQ are implemented, then larger P_L values can be tolerated. In particular, since each MACPDU is lost when (N_R+1) different transmissions are discarded at the receiver, P_L becomes:

$$P_L \le \sqrt[1+N_R]{1-\left(\frac{F}{100}\right)^{\frac{1}{\lceil M/L_{EFF}\rceil}}} \triangleq \Pi(F, L_{EFF}, N_R) \tag{9}$$

In (9) we assume independent and identically distributed errors on retransmissions.

The function $\Pi(F,L_{EFF},N_R)$ is the desired MAC parameter. The MAC can evaluate the necessity for the presence of an error protection mechanism on the basis of $\Pi(F,L_{EFF},N_R)$. Consider in fact a channel with a bit error rate p_b. In case of unprotected transmission each MACPDU would suffer the following packet loss probability:

$$P_E^{(MAX)}(p_b) = 1-(1-p_b)^{L_P} \tag{10}$$

The MAC establishes whether an error protection mechanism is needed by comparing (10) with the maximum tolerable MACPDU loss rate when the MAC has neither FEC nor ARQ as expressed by:

$$E_0(F) = \Pi(F, L_P, 0) = 1-\left(\frac{F}{100}\right)^{\frac{1}{\lceil M/L_P\rceil}} \tag{11}$$

If $P_E^{(MAX)} \le E_0(F)$, meaning that the channel produces an average packet loss rate which is acceptable in terms of QoS fulfilment, all MACPDUs can be transmitted without any protection mechanism

Once the MAC has established the necessity for an error protection, it must decide the FEC size as a function of the number of retransmissions N_R. In order to select a FEC size, the MAC first evaluates the maximum tolerable MACPDU packet loss rate, with no FEC (i.e. $L_{EFF}=L_P$), and N_R allowed retransmissions, expressed by:

$$E_1(F, N_R) = \Pi(F, L_P, N_R) = \sqrt[N_R+1]{1-\left(\frac{F}{100}\right)^{\frac{1}{\lceil M/L_P\rceil}}} \tag{12}$$

The MAC then compares $E_1(F,N_R)$ against $P_E^{(MAX)}$. The case $E_1(F,N_R) \ge P_E^{(MAX)}$ indicates that N_R retransmissions are sufficient for fulfilling the requirements with no need for FEC. Oppositely, when $E_1(F,N_R) < P_E^{(MAX)}$, a FEC scheme becomes necessary in order to increase MACPDU robustness against errors. For given channel condition (i.e. for a given p_b) and given N_R, the FEC must compensate the gap between $P_E^{(MAX)}$ and $E_1(F,N_R)$. The larger the gap, the larger the FEC size required by each MACPDU payload. But the modification of the FEC size aimed at filling up the gap between $P_E^{(MAX)}$ and $E_1(F,N_R)$, also has the effect of altering the required level of protection for each MACPDU, as expressed by (9). This problem can only be solved by means of an iterative algorithm.

If k indicates the maximum number of errors accepted by the FEC, and under the hypothesis of statistical independent errors, the MACPDU loss rate P_L is given by:

$$P_L = 1-\sum_{j=0}^{k}\binom{L_P}{j}\cdot p_b^{\,j}\cdot(1-p_b)^{L_P-j} \tag{13}$$

Equation (13) must be overturned in order to express the corrective capability k as a function of both bit error rate p_b and a target packet loss probability P_L. As shown in [8], this can be obtained as follows:

$$k = \mathrm{K}(p_b, P_L) = \left\lceil L_P\cdot p_b + erfc^{-1}(2\cdot P_L)\cdot\sqrt{2\cdot L_P\cdot p_b\cdot(1-p_b)}\right\rceil \tag{14}$$

where $erfc(x)$ is the complementary error function. Once the corrective capability k is fixed, the MAC must evaluate the number of redundancy bits which must be introduced in the payload in order to guarantee such a level of protection. Since the FEC field length however depends on the particular FEC scheme, it is convenient to introduce the following unspecified function:

$$L_{FEC} = \Lambda(k, L_P) \tag{15}$$

For a given packet size L_P, and for a requested corrective capability k, $\Lambda(k,L_P)$ determines the corresponding FEC field size L_{FEC} . Suppose for example a FEC scheme based on a Reed-Solomon code using a word length equal to 8 bits [5]. One has:

$$\Lambda(k, L_P) = \min\left\{8\cdot\left\lceil\frac{L_P}{8}\right\rceil, 16\cdot k\right\} \tag{16}$$

According to Eqs.(14) and (16), in order to guarantee a packet loss probability P_L with a bit error rate p_b, the FEC size must be:

$$L_{FEC} = \Lambda[\mathrm{K}(p_b, P_L), L_P] \tag{17}$$

As observed above, we cannot proceed however by simply introducing $P_L^{(MAX)} = \Pi(F,L_{EFF},N_R)$ as expressed by (9) into (17). We propose thus an iterative algorithm based on successive approximations. This algorithm is parametrical in N_R and is initialized by estimating the current bit error rate p_b, which allows evaluation of $P_E^{(MAX)}$ as expressed in (10). If $P_E^{(MAX)} > E_0$ (see (11)), the MAC module proceeds with the selection of the error protection mechanism, else it transmits all MACPDUs without protection on the payload field. The first step for selecting the FEC size is comparing $P_E^{(MAX)}$ with $E_1(F,N_R)$ (see Eq.(12)). If $P_E^{(MAX)} \le E_1(F,N_R)$, no FEC is required (i.e. $L_{EFF}=L_P$). Otherwise, i.e. $P_E^{(MAX)} > E_1(F,N_R)$, the following iterative steps are applied:

- step 1) The *new* packet loss probability nP_L is set equal to $E_1(F,N_R)$;
- step 2) The *old* packet loss probability oP_L is set equal to nP_L;
- step 3) The FEC size L_{FEC} is determined through Eq.(17) by considering oP_L as the target packet loss probability P_L;

step 4) The effective payload length L_{EFF} is evaluated ($L_{EFF}=L_P-L_{FEC}$);

step 5) According to Eq.(9), the new packet loss probability nP_L is given by $\Pi(F,L_{EFF},N_R)$;

step 6) nP_L and oP_L are compared:

- $nP_L \neq oP_L$ indicates that the maximum tolerable MACPDU loss rate $P_L^{(MAX)}=\Pi(F,L_{EFF},N_R)$ has been affected by the new FEC size. The procedure must be repeated from step 2) to step 6). The procedure converges thanks to the non-linear dependence on L_{EFF} of the function $\Pi(F,L_{EFF},N_R)$.
- $nP_L = oP_L$ indicates that the FEC has been correctly designed.

Once the FEC size L_{FEC} is fixed, the MAC can evaluate the capacity required for the source as given in (7). This capacity must then be expressed in terms of the number N_{PDU} of MACPDU per frame which should be reserved for the source. This value can be expressed as follows:

$$N_{PDU}=\left\lceil\left(\frac{C_{eff}\cdot D_F}{L_{PDU}}\right)\cdot\left(1+\Delta N_{ARQ}\right)\right\rceil \tag{18}$$

where the term ΔN_{ARQ} takes into account the average number of MACPDUs per frame which will be retransmitted by the ARQ mechanism. This term can be expressed as follows:

$$\Delta N_{ARQ}=\sum_{j=1}^{N_R}\left[\Pi(F,L_{EFF},N_R)\right]^j \tag{19}$$

B. Resource Allocation

For a given number of allowed retransmission N_R, equation (18) expresses the required number of MACPDUs per frame for a source with *Qspecs* which generates traffic according to *Tspecs*. In order to perform resource allocation, the MAC must simply select the value of N_R between 0 and $N_R^{(MAX)}$ leading to the lower N_{PDU} value. The resulting amount of capacity maximizes transmission efficency, but it depends on the actual p_b value; The procedure should therefore be repeated at each observed bit error rate variation.

Performance of the proposed algorithm is presented in Figs.1 and 2 for the ideal case of a channel with fixed BER. For each BER value, we represent both the value of N_{PDU} resulting by the application of the proposed algorithm, and the corresponding number of retransmissions N_R. Fig.1 refers to a typical real-time application (*source A*), while Fig.2 refers to a typical high bit rate and non-real-time application (*source B*). *Tspecs* and *Qspecs* parameters for these traffic sources are listed in Table 1. In both cases, the following system parameter are adopted: L_P = 512 bits, L_H = 88 bits (i.e. L_{PDU} =600 bits), D_{sys} = 0.02 secs., D_F = 0.01 secs., RTT = 0.05 secs. Reed-Solomon codes are used for the FEC.

For both *source A* and *source B*, we observe that when the channel is worsening, it becomes necessary to increase the number of retransmissions in order to fulfil QoS constraints.

	Parameter	Symbol	Source A	Source B
Tspecs	peak rate	p	64 Kb/s	1 Mb/s
	average rate	r	32 Kb/s	768 Kb/s
	token buffer dimension	b	11520 bits	409600 bits
	maximum packet size	M	576 bits	4096 bits
Qspecs	maximum tolerable end-to-end delay	D_{MAX}	0.1 s	1 s
	minimum percentage of packet required at destination	F	99.9	99.999999

Table 1. *Tspecs* and *Qspecs* parameters for *source A* and *source B*

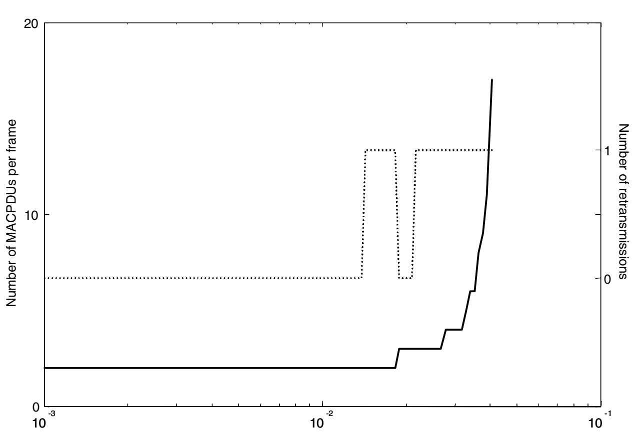

Figure 1. Number of MACPDU per frame vs. BER (solid line) and optimum number of retransmissions (dotted line) for *source A*.

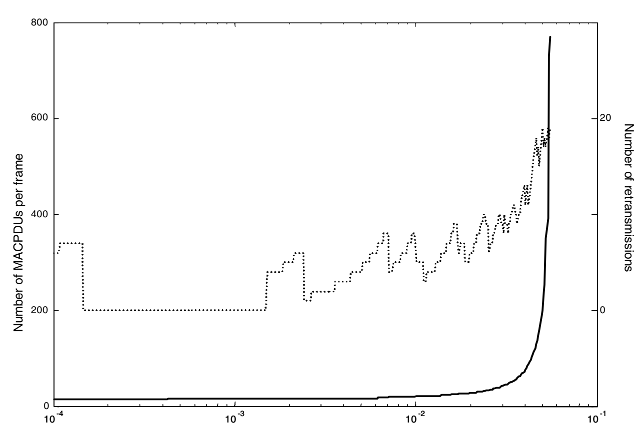

Figure 2. Number of MACPDU per frame vs. BER (solid line) and optimum number of retransmissions (dotted line) for *source B*.

IV. Performance Analysis in Presence of a Slowly Time-Varying Channel

In order to verify the performance of the proposed algorithm in the case of a slowly time-varying channel, we simulated the mobility of the receiver in a multipath environment. The reference scenario consists of a fixed transmitter and a mobile receiver characterized by a constant speed v m/s. A Multi-Carrier Code Division Multiple Access (MC-CDMA) scheme [9] with 64 sub-carriers was considered for transmission, and the Jakes channel model [10] was used for characterizing the multipath propagation. In particular, a channel impulse response composed by three main paths was considered. The capability of the algorithm to provide the required QoS for various values of the constant speed v was analyzed.

Fig. 3 shows the results of the simulation in the case of two real-time sources characterized by the same QoS parameter F=99%. The first source (*source 1*) has $D_{MAX}^{(1)}=60ms$, while the second source (*source 2*) has $D_{MAX}^{(2)}=110ms$. Three different v values are considered. In the case of v=0 (i.e. no mobility), the algorithm guarantees the fulfillment of the QoS for both sources. Problems occurr when increasing the receiver speed to v=0.35 m/s, leading to a channel coherence time equal to T_c=80 ms. In this case, we observe that the percentage of source packets which are correctly delivered to destination is below 99% threshold for both *source 1* (F=97.8%) and *source 2* (F=98.3%). When the speed is further increased to v=0.50 m/s, leading to a channel coherence time equal to $T_c=50ms$, system performance decreases only for *source 1* (F=93.4%) while the MAC is capable of guaranteeing the QoS for *source 2* (F=99.15%). According to these results, we derive that the proposed protocol is robust to the variation of BER only when the channel coherence time is not comparable with packet lifetime. If $T_c \gg D_{MAX}$, only a small percentage of MACPDUs experiences a change of BER during their lifetime. On the contrary, several changes of BER value are experienced by most of the transmitted MACPDUs when considering $T_c \ll D_{MAX}$. Also in this case, however, channel performance can be assumed constant in average terms. This conclusion was confirmed by considering several other sources with higher D_{MAX} values. Even in the presence of very severe constraints regarding the F value, the algorithm provides the required QoS for all sources with D_{MAX} values sufficiently higher than the channel coherence time.

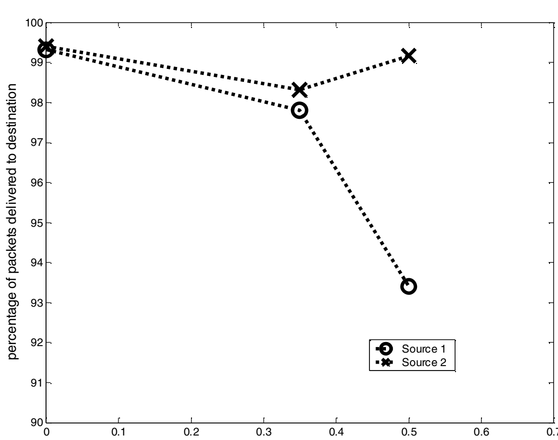

Figure 3. Percentage of source packets delivered to destination as a function of the receiver speed for *source 1* (circles) and *source 2* (crosses).

V. Conclusions

An analytical approach for selecting and designing error protection schemes at the MAC layer for traffic sources requiring QoS was proposed. The resulting algorithm maximizes transmission efficiency by selecting and dimensioning an error protection mechanism which takes into account both channel status and QoS constraints. The algorithm operates with both real-time and non-real-time sources, and it was demonstrated to be capable of optimizing resource allocation by adapting error protection to channel performance.

In the case of propagation over a slowly time-varying channel, as typical in WLANs, performance degradation is observed when the channel coherence time is comparable to the maximum tolerable end-to-end delay required by the sources. In a scenario with low mobility, the proposed algorithm can thus support the QoS for both real-time and non-real-time-applications. In a scenario with high mobility, QoS cannot be guaranteed for real-time applications only.

References

[1] D. A. Eckhardt, P. Steenkiste, "Improving wireless LAN performance via adaptive local error control", in Proceedings of the Sixth International Conference on Network Protocols, 13-16 Oct 1998, Page(s): 327-338.

[2] Jeong-Hwan Seo, Hyun-Taek Oh, Seung-Woo Seo, "A new MAC and DLC scheme considering channel conditions in wireless ATM", Vehicular Technology Conference Proceedings, 2000, VTC 2000, Tokyo. 2000 IEEE 51st, Volume: 2, 2000, Page(s): 987-991 vol.2

[3] Peng Ge, P. K. McKinley, "Comparisons of error control techniques for wireless video multicasting", in Proceedings of the 21st IEEE International Conference on Performance, Computing, and Communications, 2002, Page(s): 93-102.

[4] R. Benice, A. Frey Jr., "An Analysis of Retransmission Systems", Communications, IEEE Transactions on, Volume: 12, Issue: 4, Dec 1964, Page(s): 135-145.

[5] M. Y. Rhee, "Error-correcting coding theory", McGraw-Hill, 1989, Page(s). 319-325.

[6] S. B. Wicker, "An error control technique for high data rate communication networks", INFOCOM '89. Proceedings of the Eighth Annual Joint Conference of the IEEE Computer and Communications Societies., 23-27 Apr 1989, Page(s): 594-601 vol.2.

[7] P. P. White, *"RSVP and integrated services in the Internet: a tutorial", IEEE Communications Magazine, vol. 35, May 1997, Page(s). 100-106.*

[8] G. Giancola, S. Falco, M. G. Di Benedetto, "A Novel Approach to Error Protection in Medium Access Control Design", in Proceedings of the 4th IEEE Workshop on Mobile and Wireless Communication Networks, MWCN2002, Sept. 2002, Page(s): 337-341.

[9] R. van Nee, R. Prasad, "OFDM for Wireless Communications", Artech House, 2000.

[10] W. Jakes, "Microwave Mobile Communications", IEEE Press Classic Reissue Series, 1994.