Delay and Jitter Closed-form Expressions for Cross-Layer Hybrid ARQ Schemes

Aude Le Duc*, Philippe Ciblat†, and Christophe J. Le Martret*

*Thales Communications, Colombes, France
†ENST, Paris, France

Abstract—New ARQ or HARQ schemes taking into account the link between the MAC and IP layers have been recently introduced. In this paper, we analyze these new schemes in terms of delay and jitter by deriving these metrics in closed-form expressions at the IP level. As the framework developed for obtaining such terms in closed-form is generic, we show that the proposed expressions hold for any standard ARQ or HARQ scheme (ARQ, Incremental Redundancy HARQ, Chase Combining HARQ without cross-layer optimization) at any level (MAC or IP).

I. INTRODUCTION

One of the most popular coding mechanisms in wireless systems (e.g. HSPA, Wimax) is the Hybrid-ARQ (HARQ) which is a combination of ARQ based packet retransmission principle and forward error coding. Various HARQ schemes have been introduced such as Chase-Combining (CC) HARQ ([1]) and Incremental Redundancy (IR) HARQ ([2], [3]). Theoretical analysis of HARQ schemes through Packet Error Rate (PER), efficiency, delay and jitter is a crucial task since it enables to provide practical insights about HARQ schemes design.

Usually the theoretical or practical comparisons of different HARQ schemes are done at the MAC level, i.e., the MAC packets are considered independent of each other. In such a context, there are papers focusing on HARQ theoretical performance analysis. Regarding the basic ARQ schemes, [4] has established closed-form expressions of PER and efficiency at the MAC level. Extensions of these expressions to HARQ schemes can be found in [5], [6] for Packet Error Rate and in [7] for efficiency.

Recently the authors in [8] propose to analyze and to optimize ARQ schemes at the IP level by taking into account the fact that the MAC packets belong to the same packet coming from the network layer. Indeed the network layer packets (also called IP packets) are fragmented in order to have the length of the packets authorized by the data link layer (MAC packets). As an optimization, the authors suggested to transfer the number of retransmissions per MAC packet into a global transmission credit associated with the IP packet in the context of ARQ scheme. Such cross-layer strategy has been extended to HARQ scheme in [9]. Theoretical expressions for Packet Error Rate and Delay are given in [8] for ARQ schemes at the IP level with and without taking into account the suggested optimization.

In this paper, we evaluate the performance of any HARQ scheme in terms of delay (defined as mean number of transmitted packets per successful information packet transmission), and jitter (defined as delay variance) at the IP level. Our contribution is to provide closed-form expressions for these two metrics in the cross-layer and conventional strategies at the IP level. Notice that PER and efficiency for any HARQ schemes at the IP level have been analytically expressed in the following companion paper [9].

The paper is organized as follows: in Section II, we describe the communication scheme. In Section III, we derive the closed-form expressions of the considered metrics. Section IV is devoted to numerical illustrations. Conclusion is drawn in Section V.

II. COMMUNICATION SCHEME DESCRIPTION

A. Layer Model

In our system, we consider only the three first ISO layers : Physical, Data Link (also called MAC) and Network (also called IP) layers. We use the conventional naming of Service Data Unit (SDU) and Packet Data Unit (PDU) which is specified according to the layer [10]. For instance, we use DSU and DPDU at the MAC layer. For sake of simplicity, we focus on single user case.

At the transmitter, the MAC layer receives a DSDU packet (also called IP packet) of length $L_{IP}$ from the IP layer. The DSDU is fragmented into $N$ fragments (also called MAC packets) of length $L_{MAC} = L_{IP}/N$. From these MAC packets, the transmitter generates DPDU packets of length $L > L_{MAC}$ which are then sent through the propagation channel. The generation of DPDU packets depends on the retransmission scheme and will be detailed later on. After propagating through the physical channel, the received packet is demodulated (and decoded if necessary) and the DPDU packet is sent to the MAC layer which decides whether the transmission is successful or not. The MAC layer sends an ACKnowledgment (ACK) or a Negative ACKnowledgment (NACK) back to the transmitter accordingly. Then the delivered fragmented MAC packets obtained from DPDU packets enable the MAC layer to reconstitute the DSDU packets that are sent to IP layer.

---

*This work was prepared through collaborative participation in the URC: Urban Planning for Radio Communications project, sponsored by the Systematic Paris-Région Cluster.
B. Retransmission Schemes

For distinguishing different ARQ/HARQ schemes from each others, we describe i) the retransmission mechanism (RM) operating at the transmitter side, and then ii) the receiver processing (RP) performing at the receiver side. The global retransmission scheme is the combination of both.

1) Retransmission Mechanisms: The RM explains the transmitter behavior when continuously receiving NACKs. For all the RMs, an header is systemically added to the incoming fragment, followed by a CRC encoding in order to check the packet integrity at the receiver side.

RM1: ARQ. The DPDU packet is equal to the concatenation of MAC packet, header and CRC. The transmitter may retransmit the same DPDU at most \( P = P + 1 \) times where \( P \) is the so-called persistence.

RM2: HARQ of Type I. The DPDU packet is constituted by MAC packet with header and CRC that is then encoded by a Channel Coding of rate \( R = 1/r \). The transmitter retransmits the same coded DPDU packet at most \( P = P + 1 \) times.

RM3: HARQ of Type II (with Incremental Redundancy). MAC packet with header and CRC is encoded by a Channel Coding of rate \( R_0 = 1/r_0 \) (known as mother code). The redundancy bits are then broken up into \((r_0-1)\) subblocks of same length \( L \) and transformed into a set of DPDUs numbered as \( \{DPDU(i)\}_{i=1}^{r_0} \). DPDU(1) corresponds to the information bits whereas other DPDU(i) (for \( i = 2, \cdots, r_0 \)) correspond to redundancy. The transmitter starts to transmit DPDU(1), then DPDU(2), till DPDU\((r_0)\). As the persistence is given per MAC packet, there are at most \( P = r_0(P + 1) \) successive packet transmissions.

2) Receiver Processing: The RP depicts the way the incoming DPDUs are processed in order to decide if the associated MAC packet is corrupted or not. In the sequel, a packet is "received" when no error occurs.

RP1: Packet by packet. The receiver treats DPDU packets one by one without memory.

RP2: Flush memory. It consists in processing sequentially the incoming packets: checking CRC for DPDU(1), then after receiving DPDU(2) decoding concatenation of DPDU(1) and DPDU(2), and checking CRC, and so on till reception of DPDU\((r_0)\). If the MAC packet is not received after the DPDU\((r_0)\) decoding and the persistence is not reached, the received packet memory is flushed (put to zero) and the process starts again.

RP3: Chase Combining. The receiver combines the received packets using the Chase Combining algorithm [1].

C. Cross-layer Strategies

The conventional retransmission schemes are usually applied at the MAC layer. Recently, based on the fact that if one DPDU at the MAC layer is missing at the receiver side, the corresponding DSDU at the IP layer is dropped, the authors in [8] proposed to enhance the ARQ scheme by granting a global retransmission credit, noted \( C \), to the set of MAC packets belonging to the same IP packet. Thus, rather than allowing each of the \( N \) fragments (belonging to the same DSDU) to be transmitted \( P \) times, the new scheme allocates \( C \) transmissions to the set of \( N \) fragments. This strategy at the MAC layer actually takes into account the fact that the IP layer is expecting complete packets and constitutes a cross-layer optimization of the link control. Extending this strategy to HARQ (RM2, RM3 or any other HARQ) is straightforward and has been done in the companion paper [9].

In the later we will refer the conventional one to as PDU-Based Strategy (PBS) and the cross-layer one to as SDU-Based Strategy (SBS).

III. PERFORMANCE CLOSED-FORM EXPRESSIONS

Let \( \pi_j \) be the PER corresponding to the \((j + 1)\)-th DPDU transmission when the \( j \) previous DPDU transmissions have failed. Those PER are computed only one time (by simulation) and will serve to calculate all the performance using the analytical expressions provided later. Our objective is to derive the considered metrics with respect to \( \pi_j \) and all the design parameters \((P, N, C)\).

We derive the metric closed-form expressions in two steps. First we derive expressions using the probability \( p_n^x(k) \) which is defined as the probability that \( n \) successive MAC packets are successfully received in exactly \( k \) DPDU transmissions for xBS (with either \( 'x'='S' \) or \( 'x'='P' \)). Note that according to the definition, we have \( p_n(0) = 0 \), for \( k < n \). The obtained expressions derived for both PBS and SBS are valid whatever the RM and RP. The second step consists in expressing \( p_n^x(k) \) as a function of the elementary probability \( p_1(k) \) that is independent of the cross-layer strategy. According to the definition, \( p_1(k) \) is the probability that one MAC packet is received correctly after \( k \) DPDU transmissions. It can be thus written as

\[
p_1(k) = (1 - \pi_{k-1}) \cdot \prod_{i=0}^{k-2} \pi_i. \tag{1}
\]

For the PBS case, all the metrics can be considered at the MAC and IP levels, whereas for the SBS case, only the IP level makes sense. We also assume without loss of generality that the ACK/NACK transmissions are delay-free.

A. Metric Expressions vs. \( p_n^x(k) \)

1) Packet Error Rate: We remind the expressions of PER since they are needed for deriving those of Delay and Jitter. These expressions are proven and available in [9]. At the MAC layer, the PER writes in the PBS case as follows

\[
\Pi_{MAC}^P = 1 - \sum_{k=1}^{P} p_1(k), \tag{2}
\]

where \( P_x \) is the maximum number of transmitted DPDUs that takes different expressions depending on the RM (see Section II-B). At the IP level, we have in the PBS case

\[
\Pi_{IP}^P = 1 - (1 - \Pi_{MAC}^P)^N. \tag{3}
\]

In the SBS context, we obtain at the IP level

\[
\Pi_{IP}^S = 1 - \sum_{k=N}^{C} p_n^S(k), \tag{4}
\]
where \( C \) is the global transmission credit for the \( N \) DPDU packets per IP packet.

2) Delay: At the MAC level, the mean number of transmitted DPDUs per successful MAC packet transmission is given for PBS by

\[
\bar{n}_{\text{MAC}}^P = \frac{P_s}{\sum_{k=1}^{P_s} k \cdot \Pr\{\text{MAC packet received in \( k \) DPDUs} \mid \text{MAC packet received}\}
\]

(5)

By applying the Bayesian rule

\[
\Pr\{\text{MAC packet rec. in \( k \) DPDUs} \mid \text{MAC packet rec.}\} = \frac{\Pr\{\text{MAC packet rec. in \( k \) DPDUs}\}}{\Pr\{\text{MAC packet rec.}\}}
\]

in (5), we finally get

\[
\bar{n}_{\text{MAC}}^P = \frac{\sum_{k=1}^{P_s} k \cdot P_t}{1 - \Pi_{\text{MAC}}} = \frac{\sum_{k=1}^{P_s} k P_t}{\sum_{k=1}^{P_s} P_t}.
\]

(6)

At the IP level for the PBS case, the result is straightforward since the MAC packet transmissions are assumed independent. Thus we have \( \bar{n}_{\text{IP}}^P = N \cdot \bar{n}_{\text{MAC}}^P \), which leads to

\[
\bar{n}_{\text{IP}}^P = \frac{N \cdot \sum_{k=1}^{P_s} k P_t}{\sum_{k=1}^{P_s} P_t}.
\]

(7)

As for the SBS case, the reasoning is the same as for \( \bar{n}_{\text{IP}}^P \) but at the IP level. We get

\[
\bar{n}_{\text{IP}}^S = \frac{\sum_{k=1}^{P_s} k P_t}{1 - \Pi_{\text{IP}}} = \frac{\sum_{k=1}^{P_s} k P_t}{\sum_{k=1}^{P_s} P_t}. \]

(8)

3) Jitter: The jitter represents the variation of the instantaneous delay, denoted by \( \bar{n}_{\text{X}}^P \) (for XBS context at the X level) around its mean value given by the delay \( \bar{n}_{\text{X}}^P \) previously calculated. The jitter, denoted by \( \sigma_{\text{X}}^2 \) (for XBS context at the X level) is actually the variance of \( \bar{n}_{\text{X}}^P \) and is given by

\[
\sigma_{\text{X}}^2 = \mathbb{E}[(\bar{n}_{\text{X}}^P)^2] - (\mathbb{E}[\bar{n}_{\text{X}}^P])^2,
\]

(9)

where \( (\mathbb{E}[\bar{n}_{\text{X}}^P])^2 \) is simply given by \( (\bar{n}_{\text{X}}^P)^2 \).

At the MAC level for the PBS case, we get

\[
\sigma_{\text{MAC}}^2 = \mathbb{E}[(n_{\text{MAC}}^P)^2] - (\bar{n}_{\text{MAC}}^P)^2.
\]

(10)

We then have to identify the first term in Eq. (10), which is the second order moment of the delay. It is expressed as follows \( \mathbb{E}[(n_{\text{MAC}}^P)^2] = (1 - \Pi_{\text{MAC}})^{-1} \sum_{k=1}^{P_s} k^2 P_t(k) \). From this result and by inserting Eq. (6) in Eq. (10), we thus obtain

\[
\sigma_{\text{MAC}}^2 = \frac{\sum_{k=1}^{P_s} k^2 P_t(k)}{1 - \Pi_{\text{MAC}}} - \left(\frac{\sum_{k=1}^{P_s} k P_t(k)}{1 - \Pi_{\text{MAC}}}\right)^2.
\]

(11)

At the IP level, thanks to the MAC packets independence in the PBS case, one can prove that \( \sigma_{\text{IP}}^2 = N \sigma_{\text{MAC}}^2 \). Proof is omitted due to lack of space. The final expression is thus

\[
\sigma_{\text{IP}}^2 = N \left(\frac{\sum_{k=1}^{P_s} k^2 P_t(k)}{1 - \Pi_{\text{MAC}}} - \left(\frac{\sum_{k=1}^{P_s} k P_t(k)}{1 - \Pi_{\text{MAC}}}\right)^2\right).
\]

(12)

where \( \Pi_{\text{MAC}}^P \) is given by Eq. (2).

For the SBS case, the jitter expression is only derived at the IP level. Following the previous reasoning, we can easily obtain

\[
\sigma_{\text{IP}}^2 = \frac{\sum_{k=1}^{P_s} k^2 P_t(k)}{1 - \Pi_{\text{IP}}} - \left(\frac{\sum_{k=1}^{P_s} k P_t(k)}{1 - \Pi_{\text{IP}}}\right)^2.
\]

(13)

where \( \Pi_{\text{IP}} \) is given by Eq. (4).

B. \( p_{n_k}^P \) vs. \( p_1(k) \)

In the previous section, we have expressed all the metrics as functions of the probabilities \( p_{n_k}^P(k) \). One can note that all the metrics related to the PBS case involve only \( p_1(k) \) which can be easily deduced from \( \pi_j \) by Eq. (1). Therefore the metric derivation in the PBS case, i.e., in conventional ARQ/HARQ schemes without cross-layer consideration at the MAC and IP levels, is completed.

For the SBS case, the equations involve general terms \( p_{n_k}^S(k) \) which cannot be expressed in a simple form as a function of \( \pi_j \). We find however that \( p_{n_k}^S(k) \) can be expressed as a function of the \( p_1(k) \) probabilities which makes the link with \( \pi_j \). The idea is to remark that the different events related to \( p_{n_k}^S(k) \) are constituted by \( n \) successive independent successful transmissions of one DPDU, and thus can be written as:

\[
p_{n_k}^S(k) = \sum_{\mathbf{q} \in Q_{k,n}^S} p_1(q_1)p_1(q_2)\cdots p_1(q_n).
\]

(14)

where \( \mathbf{q} := (q_1, q_2, \ldots, q_n) \) and

\[
Q_{k,n}^S = \left\{ (q_1, q_2, \ldots, q_n) \mid \sum_{i=1}^{n} q_i = k, \text{ and } q_i > 0 \right\}.
\]

The set \( Q_{k,n}^S \) takes into account the fact that the \( n \) packets are received in exactly \( k \) transmissions.

As remarked in [8] (with a simpler set \( Q \) than ours), the Buzen’s algorithm [13] enables us to calculate recursively Eq. (14) as follows

\[
p_{n_k}^S(k) = \sum_{k'=1}^{k-n+1} p_1(k')p_{n-k'+1}^S(1 - k').
\]

(15)

First of all, using the previous delay and jitter generic expressions, we are now able to evaluate the performance of any HARQ scheme at any level without simulating it. The computational burden is given by the evaluation of the \( \pi_j \) as a function of the Signal to Noise Ratio (SNR), which is done only once for a given channel coding scheme.

Secondly, notice that, to our best knowledge, one of the novelty of this paper is to provide exact closed-form expressions for delay (as defined in the introduction) and jitter even in standard PBS case for any HARQ schemes at the MAC and IP levels. In the literature, only expression for delay in ARQ scheme is available for the PBS and SBS cases at the MAC and IP levels in [8].
C. Simplified Performance Closed-Form Expressions for ARQ

In the previous section, we derived the delay and the jitter as a function of \( p_1(k) \). Those expressions cannot be simplified further, except for the ARQ scheme (RM1-RP1). In such a case, \( \forall j, \pi_j = \pi_0 \) and real closed-form expressions for the delay and the jitter can be found for both PBS and SBS cases at the MAC and IP levels.

In ARQ scheme, Eq. (1) takes the simple form

\[
p_1(k) = (1 - \pi_0)\pi_0^{(k-1)}.
\]

(15)

For the PBS case, putting Eq. (15) into Eq. (7) leads to

\[
\bar{n}_P^P = N \left(\frac{P + \pi_0 - 2}{\pi_0 - 1} + \frac{P + 1}{\pi_0^{(P+1)}} - 1\right).
\]

(16)

Notice that \( \bar{n}_P^P \) is obtained by putting \( N = 1 \) into Eq. (16).

For the PBS case, putting Eq. (15) into Eq. (18). Eqs (16), (17), (18), and (19) are more compact with \( \pi_0 \) than those obtained in the previous section and enable their numerical evaluations with very low computational effort.

IV. Numerical Illustrations

Due to lack of space, we only inspect IR-HARQ scheme (i.e., RM3-RP2) at the IP level. The considered communication system satisfies the following assumptions:

- The IR-HARQ is implemented with the Rate Compatible Punctured Convolutional (RCPC) codes with a mother code of rate \( R_0 = 1/4 \) [3]. The modulation is a Binary Phase Shift Keying (BPSK) modulation.
- The channel is an Additive White Gaussian Noise (AWGN) channel.
- The ACK/NACK feedback is error-free. The CRC is assumed to be ideal and without overhead.

In order to evaluate numerically the proposed theoretical expressions for delay and jitter, we only need to estimate the empirical values of \( \pi_j \) (with \( j \in \{0, \ldots, k-1\} \)). These empirical values are then put into Eq. (1). Besides, empirical delay and jitter are obtained by sending one thousand DSDUs.

In Fig. 1, we plot theoretical and empirical delay at the IP level versus SNR for PBS and SBS. The number of fragments per DSDU is fixed to \( N = 2 \). The global retransmission credit is equal to 16, which means that \( C = 16 \) and \( P = 1 \) respectively in the PBS and SBS cases. We observe a perfect matching between theoretical curves and empirical ones.

In Fig. 2, we display theoretical and empirical jitter at IP level versus SNR in the PBS and SBS contexts. The simulation setup is the same as in Fig 1. Once again, a perfect agreement between theoretical curves and empirical ones occurs. Consequently, only theoretical curves for delay and jitter enable us to analyse properly system latency with respect to different parameters or approaches (influence of \( N \), PBS versus SBS, etc).

Fig. 3 and Fig. 4 represent theoretical PER versus theoretical delay and theoretical jitter respectively. Different SNRs and different global transmission credits per IP packet are considered. We remind that the global transmission credit per IP
packet is equal to \( N(P + 1)/R_0 \) in the PBS case and to \( C \) in the SBS one. The number of fragments is \( N = 3 \) and the persistence for the PBS case is \( P = 1 \). At "high" SNR, we observe that the SBS approach offers better performance than the PBS one since smaller PER is obtained for given common delay, jitter and SNR. In constrast, at "low" SNR, a trade-off has to be made between PER and latency (given by delay and jitter). Indeed, at \( SNR = -5\, \text{dB} \), the PBS approach leads to a slight loss in PER but a significant gain in latency. We finally notice that the SBS and PBS approaches can offer similar performance when the global transmission credit is lower for the SBS case than for the PBS one. This is a great advantage for the SBS case.

V. CONCLUSION

Closed-form expressions for delay and jitter in any HARQ scheme at the IP level are proposed in this paper. This analytical framework now enables to evaluate quickly the performance of any HARQ scheme in terms of various design parameters (such as the global transmission credit, number of fragments per DSDU, cross-layer or not optimization, etc).

REFERENCES