Retransmission Strategy Performance Evaluation of Relay network based on IEEE 802.11 WLAN

Hossein Bobarshad

Faculty of New Sciences and Technologies University of Tehran, Iran

ABSTRACT

Combining adaptive transmission rate and ARQ, the performance of the end-to-end packet loss rate, delay and spectral efficiency of the systems with multiple sources is evaluated. The evaluation is performed for two different scenarios: 1) Server retransmission only and 2) Source retransmission only. It is shown that the first scenario has the better performance regarding to the end-to-end delay, packet loss and TCP throughput in low SNRs, however the second scenario demonstrates better performance in high SNRs. Two scenarios fairly show similar spectral efficiency over wireless link.

Key words: Retransmission, ARQ, Performance Evaluation, WLAN.

Introduction

Efficient use of radio resources is becoming a very challenging task for future wireless communications. Limited radio resources, time and frequency-varying channel conditions, and different quality of service (QoS) requirements of the users necessitate the optimum use of spectrum, power, and time as the three main resources in communication systems. As the wireless channel is inherently a shared medium, efficient resource sharing mechanisms depend strongly on both the stochastic nature of user activity as well as the selection of physical layer coding and modulation schemes. Resource allocation based only on channel state information may cause a certain user’s queue to become unacceptably large, resulting in long queuing delay as well as packet loss. Such an information theoretic approach to resource allocation, which ignores the randomness in packet arrivals and queuing, cannot guarantee the users’ QoS and stability of queuing systems. On the other hand, QoS requirements affect the overall system efficiency, the reason being the lack of information from upper layers congestion control or link layer automatic repeat request (ARQ). Therefore to achieve a better performance it is necessary to consider the effect of the different layers on each other. Through such a cross-layer design, additional gains can be achieved for wireless systems and results in the efficient utilization of the radio resources and system performance improvement (Yi Shi, 2010; Maharshi, A., 2003; Wang, X. and J.K. Tugnait, 2003; Joint, 2004; El Gamal, H., 2006; Liu, Q., 2004; Liu, Q., 2005; Wang, X., 2007).

Authors in (Wang, X., 2007) prepared a joint design of truncated ARQ and adaptive modulation based on instantaneous signal to noise ratio (SNR). They analyzed the queuing process which induced by both ARQ and adaptive modulation. With their derived analytical expressions of pertinent performance metrics, they jointly specified the retry limit for the truncated ARQ protocol as well as the prescribed packet error rate (PER) for adaptive modulation to optimize the system throughput for QoS-guaranteed traffic, for a single source destination scenario. In (Liu, Q., 2004) the performance of the TCP protocol is studied in a system with adaptive modulation at physical layer and a finite length queue at the data link layer.

Up until now, the cross layer studies have been focused on the single source-destination pair scenario, however in a practical communication network, multiple sources and destinations may occur, such as in sensor and computer networks. As all the sources share the same spectrum resources, there is a need to investigate the medium access control (MAC) that regulates the transmission and receiving of sources/destinations.

This paper fills the gap between (Wang, X., 2007) and (Liu, Q., 2004) by designing jointly the truncated ARQ protocol and adaptive modulation scheme based on instantaneous SNR in a multiple sources system. The queuing process induced by ARQ, adaptive modulation and multiple sources is analyzed and TCP performance in terms of average end-to-end delay and throughput is investigated.

In section II the system model is described and two different scenarios in terms of adaptive retransmissions are introduced. Queuing analysis of combined adaptive modulation and ARQ is investigated for multiple sources system in section III, and packet loss rate expressions correspond to two different scenarios are derived. Performance of TCP layer in terms of end-to-end delay and throughput and performance of physical layer spectral efficiency are studied respectively in sections IV and V in the aforementioned system. Numerical results and discussion draw in section VI.
Fig. 1: Wireless system with multiple sources.

System Model:

As illustrated in Fig. 1 multiple sources transmit their information packets, simultaneously and through the wired lines to a common server in order to transmit them through the wireless link. All packets from different sources have the same length of equal to $L$ bits. All of the arrived packet streams from different sources lined up in a common finite length queue at the server which operates in first input first output (FIFO) mode. The server transmits its packets, frame by frame through the wireless link where each frame contains a fixed number of symbols $N_s$. Given a fixed symbol rate (symbols/second), each frame may contain one or more packets. After modulation with rate $R(\text{bits/symbol})$, each packet mapped to $L/R$ symbol-block and multiple of these blocks together make one frame to be transmitted through the wireless link. At the network layer each datagram has fixed length of bytes which is contained in only one packet at the data link layer, and at the transport layer each segment contains a fixed number of bytes which is transported by one datagram through the network. The segment, packet and frame structures are depicted in Fig. 2.

The wired links between the sources and the server are assumed to be error free, whereas the wireless channel is assumed to be frequency flat and remains constant during one frame duration but is allowed to vary from frame to frame. As a consequence the server adjusts the adaptive modulation frame by frame.

We assume that the received SNR over the wireless link, $\gamma$, is Nakagami-m distributed

$$f_\gamma(\gamma) = \frac{m^m \gamma^{m-1}}{\bar{\gamma}^m \Gamma(m)} \exp\left(-\frac{m\gamma}{\bar{\gamma}}\right)$$

where $\bar{\gamma}$ is the average received SNR and $\Gamma(m) = \int_0^\infty t^{m-1} \exp(-t) dt$. Let $N$ denote the total number of available modes. For these $N+1$ choices available of the AMC selector, the entire SNR range is partitioned into $N+1$ non overlapping consecutive intervals with boundary points denoted as $\{\gamma_n\}_{n=0}^N$ such that $n$ is chosen when $\gamma \in [\gamma_n, \gamma_{n+1})$. In the presence of additive white Gaussian noise, the PER can be approximated as

$$\text{PER}(\gamma) = \begin{cases} 
1, & 0 < \gamma < \gamma_{pm} \\
\alpha_n \exp(-g_n \gamma) & \gamma > \gamma_{pm}
\end{cases}$$

where $n$ is the mode index, and $\alpha_n$, $g_n$ and $\gamma_{pm}$ are mode dependent parameters, which are listed in Table-I for packet length $L = 1080$ bits.

| Table 1: Transmission Modes With Convolutionally Coded Modulation. |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Modulation | Mode 1 | Mode 2 | Mode 3 | Mode 4 | Mode 5 |
| Modulation | BPSK | QPSK | QPSK | 16-QAM | 64-QAM |
| Code rate | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{3}{4}$ | $\frac{3}{4}$ | $\frac{3}{4}$ |
| $R_s$ (bits/sym.) | 0.5 | 1 | 1.5 | 3 | 4.5 |
| $\alpha_n$ | 274.723 | 90.251 | 67.618 | 53.399 | 35351 |
| $g_n$ | 7.993 | 3.5 | 1.688 | 0.376 | 0.09 |
| $\gamma_{pm}$ | -1.533 | 1.094 | 3.972 | 10.249 | 15.978 |
These parameters are obtained by fitting (2) to the exact PER, as explained in [6, App.]. Given the mode selection scheme and the pdf of $\gamma$ in (1), then the average PER corresponding to mode $n$, known as $\overline{PER}_n$, is given by (Wang, X., 2007)

$$\overline{PER}_n = \frac{1}{\Pr(n)} \int \overline{PER}(\gamma) f_\gamma(\gamma) \, d\gamma = \frac{a_n m^n (\Gamma(m,b,\gamma_n) - \Gamma(m,b,\gamma_{n+1}))}{\Pr(n) \Gamma(m) \overline{\gamma}_n ^n b_n}$$

(3)

where $\Pr(n) = \int_{\gamma_n}^{\gamma_{n+1}} f_\gamma(\gamma) \, d\gamma$ and $b_n = m/\overline{\gamma} + g_n, n \in [1,N]$. Average packet error rate $\overline{PER}$ then can be obtained from

$$\overline{PER} = \frac{\sum_{n=1}^{N} R_n \Pr(n) \overline{PER}_n}{\sum_{n=1}^{N} R_n \Pr(n)}.$$  

(4)

If a packet is received incorrectly at the receiver, the receiver will send a retransmission request (RRQ) to the server; the server may retransmit the packet immediately or may broadcast the RRQ to all sources. The sources retransmit their packets upon receiving the RRQ from the server.

Fig. 2: Segment, packet and frame structure.

Depend on which one is responsible of packet retransmission, the server or the sources; we introduce two scenarios as follows:

I-scenario- Upon receiving the RRQ the server immediately retransmits the packet and there is no retransmission between the sources and the server.

II-scenario- RRQ is broadcasted to the sources and the sources are responsible of retransmissions; the server acts as an intermediate node in this case.

In the following sections we compare these two scenarios with respect to end-to-end delay, packet loss rate and spectral efficiency performance. In both scenarios, due to the system constraints, each packet is allowed to be retransmitted up to $A_{\text{max}}$ times either from the source or from the server, i.e. the packet is dropped if it does not received correctly after $A_{\text{max}} + 1$ times of transmission.

If the packet error rate at physical layer is guaranteed to be always less than $P_0$, (i.e. $\overline{PER}_n \leq P_0$), which naturally leads to $\overline{PER} \leq P_0$) the algorithm searching for the thresholds $\{\gamma_n\}_{n=0}^{N+1}$ to achieve the prescribed $P_0$, per mode operates as follows [6]: 1) Set $n = N$, and $\gamma_{n+1} = \infty$. 2) For each $n$, search for the unique $\gamma_n \in [\gamma_{n+1}, \gamma_{n+1}]$ that satisfies $\overline{PER}_n = P_0$, or if there is no such $\gamma_n$, pick $\gamma_n = \gamma_{n+1}$. 3) If $n > 1$, set $n = n - 1$, and go to Step 2; otherwise, set $\gamma_0 = 0$ and stop. Given a prescribed $P_0$, this algorithm guarantees that $\overline{PER}_n < P_0$ for all $n \in [1,N]$. 


Packet Loss Rate and Queuing:

The end-to-end packet loss rate $P_L$, can be due to packet dropping from the finite length queue at server with probability $p_d$, or due to corruption through the channel $PER(\gamma)$. Packet dropping rate is defined as the ratio of the number of packets which are dropped from the queue due to overflow, to the total number of packets arrived to the queue.

Here it is considered that all the arrived packet streams, from $K$ different sources, are merged and lined up in a common finite length queue at the server. The inter arrival time distribution of each stream is $f_i(x), (i = 1, ..., K)$ with average arrival rates of $\lambda_i$ and $\text{SqV}$ of $C_i$. Squared coefficient of variation or $\text{SqV}$, is defined as the ratio of the variance to the square of the mean.

Inter arrival time distribution and its characteristic parameters directly affect on the packet dropping rate and queuing performance. As a consequence the first step to queuing performance evaluation is to determine the characteristics of the arrival packet stream distribution into the queue, which is, as aforementioned, the superposition of the $K$ different streams. In the following subsection we derive the characteristic parameters of the inter arrival time distribution of the superposition (or merged) stream.

A. Merged Process Approximation:

There are several methods of approximation which can be used to determine the distribution and the first two moments of the merged process. Let $F_i(x)$ note the Cumulative Distribution Function (CDF) of the $i^{th}$ arrival stream and $F_i^c(x) = 1 - F_i(x)$; then the CDF, and the average arrival rate of the merged process, $F(x)$ and $\lambda$ respectively, can be obtained as follows (Kouvatsos, D.D., 1989)

$$F(x) \approx 1 - \frac{1}{\lambda} \sum_{i=1}^{K} \lambda_i F_i^c(x) \prod_{m \neq i} \lambda_m \int_{u=\lambda}^{\infty} F_m^c(u) du, \quad \lambda = \sum_{i=1}^{K} \lambda_i.$$  

(5)

It is shown in that the above approximated solution approaches the correct solution as $K$ approaches to infinity. As it will be discussed in numerical results section, using the above equation, it is straightforward to show that the superposition of two or more streams drown from general exponential (GE) distribution can be approximated by a GE distribution.

Defining the traffic intensity as $\ell = \lambda / \mu$, $0 < \ell < 1$, according to (Whitt, W., 1995), the $\text{SqV}$ of the merged stream $C(\ell)$ can be quite accurately approximated as

$$C(\ell) \approx \sqrt{w(\ell) \sum_{i=1}^{K} \left(\frac{\lambda_i}{\lambda}\right)^2 C_i^2 + 1 - w(\ell)}.$$  

(6)

where $w(\ell) = \left[1 + 4(1-\ell)^2(1-v)\right]^{-1}$ and $v = \left[\sum_{i=1}^{K} \left(\frac{\lambda_i}{\lambda}\right)^2\right]^{-1}$.

Hence the approximation of the inter arrival time distribution of the merged stream, and its mean and variance can be obtained (4) and (5).

B. Packet Dropping Rate Calculation:

Let denote $q_0$ as the number of packets in the finite length queue of length $Q$, packet dropping probability $p_d$ is defined as

$$p_d = \Pr(q_0 = Q)$$  

(7)

It means that $p_d$ is equal to the probability of having $Q$ packets in the finite length queue of length $Q$. If the processing time distribution is modeled by exponential distribution with mean $1/\mu$ seconds, according to (Kouvatsos, D.D., 1989; Nagarajan, R., 1991; Kaur, H.T., 2003), $p_d$ can be approximated by

$$p_d = \frac{(1-\omega)\omega^Q}{1-\omega^{Q+1}}.$$  

(8)
where $\omega$ is found by solving the following equation

$$\omega = \mathcal{L}\{f(x)\}_{[\omega=\mu]} = \frac{1}{\omega} \int_0^\infty f(x) e^{-(1-\omega)\mu} \, dx$$

(9)

The right hand side of the above equation is Laplace transform of the probability density function (pdf) of the packet inter-arrival time of the merged process, $f(x)$, at $s=(1-\omega)\mu$. Average packet dropping rate is obtained from averaging $p_d$ over all channel conditions, $\overline{p}_d$.

The most common choice for telecommunication network design is based on the exponential assumption of holding times. Usual choice is the Poisson arrival of the packets, however today’s networks and applications generate a traffic that is self-similar and bursty over a wide range of time scale. Using Poisson processes to model packet arrivals in a computer network will certainly lead to a failure simply since there is far too much correlation among packet arrivals to have an assumption of independent arrivals (Willinger, W., 1997).

Assuming squared coefficient of variation, $\text{SqV}$ as the ratio of the variance to the square of the mean, bursty distributions are defined as the distributions whose $\text{SqV}$ are greater than one. Some example distributions that are burstier than Poisson are Generalized Exponential (GE), Weibull, Pareto, and lognormal distributions. The GE distribution, which allows for modeling of a wide range of bursty and self-similar distributions, is used in this work to model the packet arrivals. It can be shown that all the aforementioned distributions are special form of Levy-skew-$\alpha$-stable distribution (Willinger, W., 1997).

C. Packet Loss Rate:

As mentioned before, each packet can be lost either due to dropping from the finite length queue at the server or due to corruption through the channel, which leads to end-to-end packet error rate of $P_l$.

If the packet error rate at physical layer is guaranteed to be always less than $P_0$, in the first scenario after $A_{\text{max}}$ times of packet retransmission from the server, the probability of packet corruption over the wireless channel would be guaranteed to be always less than $P_0^{A_{\text{max}}+1}$ and consequently the end-to-end average packet loss rate correspond $p_l$, would be guaranteed to be

$$p_l \leq \overline{p}_d + (1-\overline{p}_d)P_0^{A_{\text{max}}+1}$$

(10)

In the II-scenario, when each packet could undergo up to $A_{\text{max}}$ times of retransmission from the source, if the packet error rate at physical layer is guaranteed to be always less than $P_0$, then $p_l$ would be guaranteed to be

$$p_l \leq (\overline{p}_d + (1-\overline{p}_d)P_0)^{A_{\text{max}}+1}$$

(11)

Therefore, in each scenario the number of transmissions that a given packet undergoes is a geometric random variable with mean $\overline{A}_I$ and $\overline{A}_{II}$ respectively in the first and the second scenario as

$$\overline{A}_I = 1 + P_0 + \cdots + P_0^{A_{\text{max}}} = \frac{1-P_0^{A_{\text{max}}+1}}{1-P_0},$$

(12)

$$\overline{A}_{II} = 1 + p_l + \cdots + p_l^{A_{\text{max}}} = \frac{1-p_l^{A_{\text{max}}+1}}{1-p_l}.$$  

(13)

In the first scenario, packets arrive to the queue with the average rate of $\lambda$, (4) and each packet is transmitted $\overline{A}_I$ times in average from the server over the wireless link. Equivalently the average processing time of each packet multiplies by $\overline{A}_I$ and the packet processing rate can be modeled by $\mu = R\mu_0/\overline{A}_I$, where $\mu_0$ is the processing rate corresponds to $R=1$ and no retransmission.

In the second scenario, as each packet is transmitted $\overline{A}_{II}$ times in average from the source toward the queue, the packet inter arrival rate into the queue can be modeled by $\overline{A}_{II}\lambda$. The packets leave the queue with the rate of equal to $\mu = R\mu_0$ where $\mu_0$ is the service rate corresponding to $R=1$.

It is important to note that, in the second scenario, the $P_l$ is defined in terms of $\overline{A}_{II}$, this implies that equation (12) is a fixed point equation and it is necessary to justify that it has a unique solution. Start from $g(x) = 0$ where
observe that \( g(1) \leq 0 \), and \( g(A_{\text{max}}) \geq 0 \), and as \( \partial g(x)/\partial x > 0 \) it is straightforward to verify that \( \partial g/\partial x > 0 \). It follows that there is a unique solution for the above equation in \([1, A_{\text{max}}]\).

### TCP Performance:

In the system we are modeling here, the round trip time of TCP (RTT) of each source includes the propagation delay from the sources to the queue \( T_s \), the random transmission delay in wireless link \( T_w \), and the random queuing delay \( T_q \). The transmission time for acknowledgment over the wireless link, \( T_{w_f} \) is assumed to be negligible and \( T_s \) is equal for all the sources. Hence from each source we have

\[
RTT = 2T_s + T_w + T_q
\]  

where \( T_w \) is equal to the average service time upon the queue, \( 1/\mu \), and \( T_s \) is assumed to be constant during the connection.

To obtain the average waiting time per packet in the queue, \( T_q \), we use the Pollaczek-Khinchin formula (Bertsekas, D. and R. Gallager, 1992)

\[
T_q = \frac{\lambda \times \text{2nd moment of service time}}{2(1 - \lambda \times \text{average service time})}
\]  

where \( \lambda \) is the average number of arrival bits per second from the source.

### Spectral Efficiency:

The spectral efficiency is defined as the number of bits transmitted per second per unit of the bandwidth. Taking into account the ARQ and the queuing affects the spectral efficiency (SE) is expressed as

\[
\text{SE} = Pr(q_0 \geq \mu) \frac{R}{\bar{A}} + Pr(q_0 < \mu) \frac{R}{\mu}
\]  

where \( \bar{A} \) is the average packet transmission number over the wireless link. From (7) it is also straightforward to show

\[
Pr(q_0 \geq \mu) = 1 - Pr(q_0 < \mu) = 1 - \left[ \frac{(1 - \omega\mu)}{1 - \omega} + 2 \frac{(1 - \omega\mu)^2}{1 - \omega^2} + \cdots + (\mu - 1) \frac{(1 - \omega\mu)^{\mu-1}}{1 - \omega^{\mu-1}} \right].
\]

Fig. 3: End to end connection.
In order to highlight the difference between $\overline{A}$ and $\overline{A}_I$, it is important to note that in the first scenario $\overline{A} = \overline{A}_I$, but in the second scenario, as packets may be dropped due to overflow, the average number of transmission that any particular packet undergoes from the source $\overline{A}_I$, is not necessarily equal to the average number transmission that, that particular packet experiences over the wireless link, i.e $\overline{A} \neq \overline{A}_I$. In this case $\overline{A}$ can be obtained from

$$\overline{A} = 1 + P_0(1 - \overline{p}_d) + \ldots + P_0^{d_{\max}}(1 - \overline{p}_d)^{d_{\max}} = \frac{1 - (P_0(1 - \overline{p}_d))^{d_{\max}}}{1 - P_0(1 - \overline{p}_d)}$$

(20)

The average spectral efficiency over fading channels would be obtained by averaging (17) over the all channel conditions.

**Numerical Results and Discussion:**

We assume that there are 10 different sources $K = 10$ transmit their packets of length 1000 bits to the server. The packet inter arrival time distributions are assumed to be generalized exponential where their average rates $\lambda_i$ are assumed to take on arbitrary values between 0.9 and 1.1. SqV of the inter arrival time distributions are set to be 2 in order to guarantee the bursty arrival packets. $\mu_0 = 30$ and the queue length is $Q = 20$ and it is assumed that the rate can take any integer value between 0 to eight, $r = 8$.

The probability density function of GE distributions, $f_i(x)$, is given by (Nagarajan, R., 1991)

$$f_i(x) = (1 - a_i)\delta(x) + a_i^2 \lambda_i e^{-a_i x}, \quad 0 < a_i \leq 1$$

(20)

where $\delta(x)$ is the delta function and $a_i = 2/(C_i + 1)$. Bearing in mind that the superposition of two or more streams drawn from GE, can be approximated by a GE distribution, the mean inter arrival rate rate of the merged process, $f(x) = (1 - a)\delta(x) + a^2 \lambda e^{-ax}, 0 < a \leq 1$, can be obtained from $\lambda = \sum_{i=1}^{K}\lambda_i$. Also $a = 2/(C(\ell) + 1)$, where $C(\ell)$ can be found from (5). Taking Laplace transform of $f(x)$ for $s = (1 - \omega)\mu$, it is shown that $\omega$ is obtained solving

$$\omega = (1 - a) + \frac{a^2 \lambda}{(1 - \omega)\mu + a \lambda},$$

(21)

where $\omega$ is derived as $\omega = 1 - a + a\lambda/\mu$. Using (13), the packet dropping probability is obtained from

$$p_d = (1 - \omega)\omega^{\ell}/(1 - \omega^{\ell + 1}).$$

Fig. 3 depicts the end-to-end packet loss rate when the packet error rate at wireless link is guaranteed to be always less than $P_0 = 10^{-3}$. It can be seen that the first scenario has better packet loss performance in low SNRs than the second scenario. In low SNRs the transmission rate over the wireless channel is small and as the service rate is proportional to $R$, the packet loss is mostly due to packet dropping from the queue rather than corruption in wireless channel. Consequently packet retransmission from the source to the queue increases the packet dropping rate and descends the packet loss performance. However in high SNRs as the packet dropping rate is very small as compare to $P_0$, from (10) it can be seen that retransmission from the sources improves the packet loss performance more than retransmission from the server.

End-to-End delay is plotted versus SNR for two scenarios in Fig. 4. The first scenario has less end to end delay over all channel conditions than the second one.

The average TCP throughput of $K$ sources, $\xi = \sum_{i=1}^{K} \xi_i/K$ is shown in Fig. 5. The average TCP throughput of the first scenario is higher in low SNRs than the II-scenario. The proof is straightforward from (16) and definition of $p_t$ in correspond to the first and second scenario. The spectral efficiency is fairly the same for two scenarios except in low SNRs and high $P_0$. Fig. 6 compares the SE of two scenarios.

From above one can decide to use the first scenario in low SNRs, and the second scenario in high SNRs in order to get better performance in terms of the packet loss, delay and TCP throughput.
Fig. 3: Packet loss rate versus channel SNR, $A_{\text{max}} = 1$ and $P_0 = 10^{-3}$.

Fig. 4: Delay versus channel SNR, $A_{\text{max}} = 1$ and $P_0 = 10^{-3}$

Fig. 5: Average TCP throughput versus SNR, $A_{\text{max}} = 1$ and $P_0 = 10^{-3}$
Fig. 6: SE versus SNR for $A_{max} = 1$ and $P_0 = 10^{-4}, 10^{-1}$.

References


