4th WSEAS International Conference on ELECTRONICS, CONTROL and SIGNAL PROCESSING, Miami, Florida, USA, 17-19 November, 2005 (pp.223-228)

# A Congestion Avoidance Mechanism for Enhancing Performance of Packet Loss Concealment in VoIP Coders

Jaesup Lee[1], Seung Ho Han[2], Won Ryu[1], and Minsoo Hahn[2]
[1] BcN Service Research Group, ETRI
161, Gajeong-dong, Yuseong-gu, Daejeon, 305-350, Korea

[2]School of Engineering, Information and Communications University
119, Munjiro, Yuseong-gu, Daejeon, 305-732, Korea

*Abstract:* - VoIP speech coders have PLC mechanisms to compensate for a packet loss. In this paper, various kinds of packet losses are designed. With these losses the PLC algorithms of speech coders are evaluated and analyzed by PESQ value. Based on these outcomes, reduction of loss rate is crucial, especially for VoIP. Thus, we suggest an enhanced congestion avoidance RED mechanism to enable PLC work better in reliable manner within its perceptual capabilities for VoIP service in the network.

*Key-Words: - VoIP, Internet, Packet Concealment, Congestion Avoidance, Coder, Loss Probability*

## 1 Introduction

Packet loss occurs inherently in an IP network due to network characteristics such as delay, jitter, and network congestion. The packet loss can be occurred due to the missing speech samples for a time interval at receiver. When a packet is lost in VoIP, it brings about the degradation of speech quality. If the packet loss with high loss rate or consecutive packet loss occurs, it is largely harmful to speech quality. Since transmission is real-time process, it is impossible to re-transmit missing data when the data is lost. A concealment technique is required for reconstruction of a missing packet. Speech coders have their own PLC (Packet Loss Concealment) which is also called FEC (Frame Erasure Concealment) mechanisms against the packet loss. At the receiver, new speech signal is generated to reduce the effect of packet loss by the PLC mechanism. Together with this PLC supported coder, we have to provide a reliable VoIP network with the efficient cooperating feature such as RED (Random Early Detection) [1] using an enhanced queue management of the network nodes to reduce the loss rate of packets.

This paper focuses on the performance evaluation and analysis of the PLC algorithms in speech coders associated with a congestion avoidance feature of RED we propose for VoIP network nodes. The various packet losses are generated by EID (Error Insertion Device) [2] and PESQ (Perceptual Evaluation of Speech Quality) [3] is used to evaluate the PLC algorithms. PLC algorithms in G.711 Appendix I, G.723.1, G.729A, and AMR are tested.

## 2 PLC in VoIP Coders and VoIP Network

In the coders, decoding processes including PLC mechanism are done by frame-basis. For simplicity, it is supposed that 1 packet contains 1 speech frame. Since G.711 is waveform coder, the PLC mechanism uses previous signals to generate synthetic signals. PLC algorithm repeatedly inserts pitch period detected from the previous speech in history buffer [4]. Even though G.711 is sample-by-sample coder, it is assumed to be partitioned into 10 ms or 20 ms frames. The PLC algorithm of CELP-based hybrid coders is usually based on extrapolation of the parameters estimated by the last good frame. The PLC algorithm in G.723.1 generates an excitation signal by the different ways according to the classification of voiced/unvoiced [5]. A random number generator generates the excitation signal when the result of classification is unvoiced while pitch estimation is used when the result is voiced. In G.729A, which is complexity reduced version of G.729, the excitation signal is generated by using both adaptive and fixed codebook [6]. The PLC algorithm in AMR has a state machine based on whether the current frame and the previous frames are lost or not [7]. The state stands the number of consecutive frame loss. An excitation signal is generated with controlling fixed codebook gains and adaptive codebook gains under the state.

In VoIP networks, since TCP has congestion control mechanism implemented at the end host, it responds to the packet dropping after a round trip time. Meanwhile, UDP traffic is treated in a worse manner than TCP traffic since their sources are unresponsive to the losses during the period of congestion. Eventually, that the loss rate suffered by UDP/IP voice applications will increase significantly in case of TCP retransmission having a long lengthy of packets.

When a FIFO queue overflows, lots of active TCP sessions are likely to lose an IP packet simultaneously. These sessions will back off and the congestion will disappear. Then all these TCP sessions will increase their window size gradually. These two kinds of events occurred repetitively which can lead to an oscillation of TCP traffic caused by a synchronization of VJ algorithms as specified in RFC 2001 after a queue overflow.

Our research interest is mainly about packet management scheme in routers in terms of an average queue length and an incoming packet length to reduce loss ratio of the real time packet, especially for VoIP service with PLC when congested.

# 3 Proposed RED Mechanism and Protocol

## 3.1 Revised RED Mechanism

To improve the performance of a packet switched network, an active queue management is recommended by Internet Engineering Task Force as a *de-factor* standard for congestion avoidance in the Internet [8]. Therefore, RED that should be used as the default mechanism for managing queues in routers was proposed. However, it has somewhat problems in terms of the throughput and the consecutive drop of packets. Thus, we try to improve voice packet throughput by decreasing a packet drop in a VoIP network which finally resulted in increasing the performance of VoIP packet service as a reliable and scalable environment when a packet loss concealment algorithm is applied.

RED performs two functions. One is the estimation of an average queue size, $s_n$ in Equation (1) below. It is calculated by the instantaneous queue size using an exponentially weighted moving average. The other is the decision based on the average queue length estimated either to drop a corresponding incoming packet or not.

$$\begin{aligned} s_n &= (1-w)s_{n-1} + wq_n, \ \text{if } q_n > 0 \\ s_n &= (1-w)^m s_{n-1}, \qquad \text{otherwise,} \end{aligned} \tag{1}$$

where $w$ is a weight that is applied to the current queue size, and $m$ is $\mu/\lambda$ .

And, decision process is to perform its function mentioned above in accordance with the drop probability, $p_d$ defined by the estimated average queue size as follows:

$$p_d = \begin{cases} 0 & ;\ s_n < k_l \\ (s_n - k_l)\max_p \big/ (k_h - k_l) & ; k_l \le s_n < k_h \\ 1 & ; s_n \ge k_h, \end{cases} \tag{2}$$

where $k_h$ specifies the average queue size above which all packets will be dropped, and $k_l$ specifies the average queue size below which no packet will be dropped. $p_d$ is the drop probability when the average queue size reaches within the designated buffer limits. Between two thresholds the drop probability varies linearly from 0 to the maximum drop probability $\max_p$ .

According to (1), (2) and drop probability as well as queuing delay of finite buffer size (K) based on Markovian model. Overall packet drop in either a random or consecutive occurrence depends on the size of virtual queue length. To improve VoIP QoS even in under high load, we have to reduce packet dropouts by modifying packet drop distribution function into a cubing function so that it may keep drop rate increasing gradually for the capability of the PLC performance at the end user. Thus, we modify $p_d$ as follows:

$$p_d = \begin{cases} 0 & ; s_n < k_l \\ (s_n - k_l)\max_p \big/ (k_h - k_l) & ; k_l \le s_n < k_h \\ \dfrac{(1-\max_p)(s_n - k_h)^3 - \max_p (k_h - K)^3}{(K - k_h)^3} & ; k_h \le s_n < K, \end{cases} \tag{3}$$

where following condition for $\max_p$ should be satisfied:

$$\max_p \le \frac{k_h - k_l}{K - k_h}. \tag{4}$$

From (3), where an average queue size is placed between $k_h$ and K, we might change its value by multiplying some other parameter value as shown in (5). The reason we adopt modification is to consider the behavior of TCP packets that works in a mutually exclusive pattern with UDP packets, but affects each other in terms of bandwidth and throughput of the propagation link channels. The more the queue mechanism accepted the packets until the next drop, the more the drop probability will be increased. Therefore, we might consider that the drop probability for the consecutive packet [13] would be increased in accordance with the number of accepted packets since the last drop.

$$\hat{p}_d = \frac{p_d}{(1-N\cdot p_d)} \cdot \frac{\Phi_i(1 \pm |\rho_{UT}|)}{\left\{ \Phi_i \Theta \left( \frac{(RTT_j - I_i)/RTT_j}{\sqrt{2}} \right)^2 \right\}} \cdot \left( \frac{L}{M} \right)^{m'}, \quad (5)$$

where $0 \le N \cdot p_d < 1$ and $0 < \Phi_i < 1$.

From the equation above, $N$ represents the number of accepted packets since the last drop, and $\Phi_i$ represents the marginal, desirable PLC loss rate chosen by a service provider, as the coder $i$ requires to meet the perceptual performance for VoIP service in the network. $M$ and $L$ represent the maximum transmission unit of packet and the size of an instantaneous packet respectively. $RTT_j$ is the minimum round trip time for the service $j$ required and $I_i$ is an instantaneous mean round trip time for on-going session $i$. Operator $\Theta$ defines the operation as follows:

$$\begin{cases} \text{if } RTT_j \ge I_i, \text{ it performs plus(+) operation,} \\ \text{if } RTT_j < I_i, \text{ it performs minus(-) operation.} \end{cases} \quad (6)$$

But, it performs operation in the other way when the network is equipped with RSVP in order to facilitate UDP path reservation through the network to the VoIP gateway or to the end users.

The reason for which we proposed (5) comes from the TCP goodput estimation formula proposed in [11]. And, $m'$ indicates the size momenting factor, here we select its value as 2 which suggested in [11].

Let $\rho_{UT}$ define the correlation coefficient between two variables, $X_U$ and $X_T$, which represents the loss rate of UDP packet, TCP packet in a gateway, respectively. $\rho_{UT}$ is applied to the TCP or UDP packet in a sequence of +/- operation for the corresponding packet when $\rho_{UT}$ is negative and UDP loss rate at the coder or at the gateway increases during the designated time interval by the receiving-side gateway.

$$\rho_{UT} = \frac{Cov(X_U, X_T)}{\sigma_{X_U} \sigma_{X_T}}. \quad (7)$$

In the RED, the queue maximum threshold $k_h$ is related to the burstiness of incoming traffic. We assume that this burstiness strongly affects PLC coder to have busty loss frames in the VoIP packet for the same session that treated by the uncontrollable UDP packet for voice service. Therefore, initialization of $k_h$ might be set as like $k_h$ =3 $k_l$ , then it will be increased by $\delta_i$ according to the capability up to the maximum which the coder can handle it for error loss concealment considering the number of consecutive loss or error frames in the voice handling time limit. We need the soak time to settle down the desirable value of this parameter in the gateway that does not trespass the capability that is given for coder to handle the burst traffic.

As a number of authors have noted, the loss rate $p$ and throughput T of a TCP connection are approximately related as follows:

$$T = \frac{L \times c}{RTT \times \sqrt{p}}, \quad (8)$$

where T is the observed throughput in octets per second, L is the packet size in octets, RTT is the round-trip time for the session in seconds, p is the packet loss rate and c is a constant in the range from 0.9 to 1.5. Using this relation, one may determine the packet loss rate that would result in a given throughput for a particular session, if a TCP connection was used. The capacity of the shared link is denoted by C, and the round trip propagation delay of connection $i$ is given by $d_i$ and $\gamma$ denotes the total transmission rate of UDP connections for that shared link, which consists of the number of UDP connections $N_U$ and TCP connections $N_T$.

Whilst this relation between packet loss rate and throughput is specific to the TCP congestion control algorithm, it also provides an estimate of the acceptable loss rate for a streaming media application using the same network path, which wishes to coexist fairly with TCP traffic. However, this is not sufficient for fair sharing of a link with TCP traffic, since it does not capture the dynamic behavior of the connection, merely the average behavior, but it does provide one definition of “reasonable” behavior in the absence of real congestion control.

Thus, given $p_n < X_U$ one can compute the queue size at period n+1, $n \geq 1$ as the solution of (10) as follows:

$$\frac{L \cdot c}{\sqrt{p_n \left( d + \frac{q_{n+1} \cdot L}{C} \right)^2}} = \frac{C - \gamma(1 - p_n)}{N - N_u}. \qquad (9)$$

Performance will be presented in the next section. In the case of a fixed $k_h$, we can roughly estimate the result using our RED mechanism from (3) through (9). For varying $k_h$, the result will be better. Any way, by using this mechanism, the average loss rate will be reduced as follows:

$$\Delta = \sum_{s_n = k_h}^{K} \pi(s_n) - \sum_{s_n = k_h}^{K} \pi(s_n) \times \hat{d}, \qquad (10)$$

where $\pi(\cdot)$ is a stationary distribution of the average queue length and $\pi(\aleph) = prob\ (s_n = \aleph)$.

For simplicity and time-scale decomposition approach, Equation (5) can be applied only when the virtual queue length exceeds $k_h$. Because the average queue size of the RED gateway evolves much slower than the TCP dynamics and UDP has no any means in its characteristics against congestion [12]. Thus, we use $\rho_{UT}$ for dealing that matter.

### 3.2 RTP/RTCP Protocol Extension/Usage

Information about the loss rate of UDP at the receiving side can be calculated by using the sequence number, PT, and frame indicating mark bit (M) of RTP header in the voice packet and checking UDP data checksum for data error. These results should be sent to the node or sender as the sender/receiver reports (SR/RR) by the RTCP. RTCP–XR packet format consists of the information in their sub-block such as delay between sender and gateway at the receiving side, which indicates the value of delay since last RR. Also the round trip delay between gateways or gatekeepers might be put in the same manner in the corresponding sub-block periodically.

## 4 Simulation

### 4.1 Packet loss design

A communications channel can be modeled by Gilbert-Elliot model or Bellcore model [9,10]. Gilbert-Elliot model is a binary-state Markov model, and Bellcore model is N-state Markov model. In this paper, STL (Software Tool Library) included in ITU-T Recommendation G.191 is used to model the packet losses occurring in IP network [2]. STL provides EID that generates loss based on Gilbert-Elliot model and Bellcore model. A variety of losses can be generated by modification of parameters in EID.

Burst and random packet losses with 1%, 3%, 5%, 10%, and 15% loss rates are generated. For the efficiency of system, one or multiple speech frames are put together in a packet. It is considered that a packet contains 1 speech frame, 2 speech frames, and 3 speech frames. In this experiment, the frame length is 10 ms for G.711 and G.729A, 30 ms for G.723.1, and 20 ms for AMR.

### 4.2 Performance evaluation of PLC

Measurement of speech quality after PLC processing is required to evaluate the performance of the PLC algorithms. MOS (Mean Opinion Scores) test is generally used for subjective listening test. Since total number of test data is very large, speech quality is measured by the systemic evaluation method. ITU-T recommends the method for objective assessment of quality. It is PESQ included in ITU-Recommendation P.862 [3]. PESQ give us reliable estimation of subjective measurement as MOS-like scale. The output of PESQ provides the score from -0.5 to 4.5 after comparing an original signal with a degraded signal.

Korean dialogue speech from 2 male and 2 female speakers is used as test speech data for each packet loss condition. Total number of speech is 200. Duration of each speech is about 10 seconds.

Figure 1 shows the experimental procedure. First, the speech data are encoded. Second, packet losses are generated by the EID. Third, the encoding files are decoded with the packet losses and without the packet losses. The PLC processes are executed while the encoding files are decoded. Finally, the decoding files with packet losses are evaluated with the reference speech decoded without the packet losses in the speech quality evaluation module.

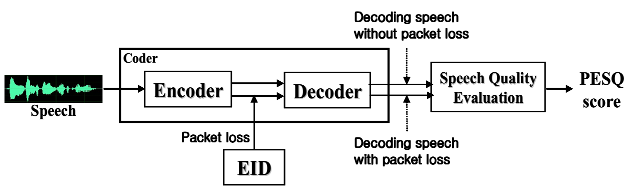


Fig. 1 Experimental procedure

### 4.3 Experimental parameters- RED

We suppose that the arrivals occur randomly according to a Poisson Process with rate $\lambda$ and service times of single server with buffer size K is exponentially distributed at rate $\mu$. We select the value of parameters setting in the simulation as: $\mu$ = 1000 packets/sec., K = 200 packets, $w$ = 0.002, M = 1500 bytes, $\Phi_i$ = 5, RTTj = 150ms, $k_l$ = 10, $k_h$ = 130. In simulation, round trip delays vary between 120ms and 220ms.

## 5 Results

### 5.1 Test result without considering RED

The results of PLC evaluation are shown in Table 1, 2 and 3 as PESQ score. In Tables, 'R' and 'B' imply random packet loss and burst packet loss, respectively.

Table 1. Results of PLC for 1frame per packet

| | G711 | | G723.1 | | G729A | | AMR | |
|---|---|---|---|---|---|---|---|---|
| | R | B | R | B | R | B | R | B |
| 1% | 3.982 | 3.975 | 3.974 | 3.970 | 3.952 | 3.953 | 3.915 | 4.032 |
| 3% | 3.644 | 3.558 | 3.560 | 3.571 | 3.570 | 3.531 | 3.476 | 3.567 |
| 5% | 3.440 | 3.317 | 3.299 | 3.287 | 3.319 | 3.275 | 3.215 | 3.273 |
| 10% | 3.124 | 2.839 | 2.863 | 2.859 | 2.936 | 2.808 | 2.736 | 2.841 |
| 15% | 2.893 | 2.504 | 2.562 | 2.422 | 2.658 | 2.478 | 2.416 | 2.472 |

Table 2. Results of PLC for 2 frames per packet

| | G711 | | G723.1 | | G729A | | AMR | |
|---|---|---|---|---|---|---|---|---|
| | R | B | R | B | R | B | R | B |
| 1% | 3.987 | 3.991 | 4.043 | 4.028 | 3.912 | 4.031 | 3.947 | 4.010 |
| 3% | 3.630 | 3.589 | 3.579 | 3.662 | 3.488 | 3.634 | 3.492 | 3.626 |
| 5% | 3.412 | 3.262 | 3.305 | 3.374 | 3.239 | 3.347 | 3.219 | 3.367 |
| 10% | 3.038 | 2.758 | 2.804 | 2.887 | 2.818 | 2.896 | 2.756 | 2.927 |
| 15% | 2.784 | 2.381 | 2.444 | 2.561 | 2.537 | 2.543 | 2.406 | 2.523 |

Table 3. Results of PLC for 3 frames per packet

| | G711 | | G723.1 | | G729A | | AMR | |
|---|---|---|---|---|---|---|---|---|
| | R | B | R | B | R | B | R | B |
| 1% | 4.018 | 3.983 | 4.019 | 4.015 | 3.939 | 4.019 | 4.023 | 4.058 |
| 3% | 3.621 | 3.600 | 3.542 | 3.689 | 3.500 | 3.698 | 3.527 | 3.683 |
| 5% | 3.390 | 3.305 | 3.238 | 3.453 | 3.234 | 3.406 | 3.256 | 3.393 |
| 10% | 2.965 | 2.756 | 2.732 | 2.953 | 2.776 | 2.936 | 2.778 | 2.983 |
| 15% | 2.667 | 2.4071 | 2.367 | 2.607 | 2.475 | 2.585 | 2.420 | 2.561 |

The results for 1 frame per packet are shown in Table 1. PESQ values of G.711 are greater than any other value of coders at the loss rates for random packet losses while the results of all coders show similar PESQ values for burst packet losses. In G.711 and G.729A, the performance of PLC for burst packet losses is relatively worse than that for random packet losses. In G.723.1, the differences between PESQ values for random packet losses and burst packet losses are relatively small. On the contrary, in AMR, the results for burst packet loss are better than those for random packet loss. Since PLC algorithm of G.723.1 is designed for burst error and AMR has the state machine based on consecutive errors, these results are valid. Thus, G.723.1 and AMR are relatively robust for burst packet loss.

The results for 2 frames and 3 frames per packet are shown in Table 2, and Table 3, respectively. The results show slightly different trend when compared with the result for 1 frame per packet. In AMR, the performance of PLC for random packet losses is relatively improved when compared with the others as the number of frame per packet is increased. The gap of PESQ value with the other coders is decreased. The performance of PLC in G.711 for burst packet losses is relatively largely dropped as the number of frame per packet increases while the others show similar result each other at each loss rate. And, the results of PLC for burst packet loss are better than those for random packet loss in G.723.1, G.729, and AMR.

As the number of frame per packet increases, the performances of PLC for random packet losses are dropped in all coders except AMR. The results of PLC in AMR are similar in 1 frame, 2 frames, and 3 frames per packet. In contrary, the performances show reverse results for burst packet losses. As the number of frame per packet increases, the performances get better in all coders except G.711. The results imply that the performance is worse when short consecutive packet losses are very frequently happened than when long consecutive packet losses occur as the effect of packetization.

These overall test results of PLC coder's algorithms show that at least 5% loss of packets will be allowed at the end user. Thus, on the VoIP gateway, it needs less loss ratio of packets to provide the good quality of voice considering PLC capability at the receiving-side.

### 5.2 Test result with RED

Figure 2 shows that packet handling by ERED on the node is much useful to get the favorable number of consecutive dropouts for fitting PLC performance for an offered load of $\rho$=2. In Figure 3, packet loss ratio per packet length by the drop probability weighted by the square of the ratio of the packet size over the maximum packet size can be a driven force to prevent congestion caused by the long length of packet which is not favorable to CBR UDP VoIP packet service in terms of packet delay in a queue of the network nodes.

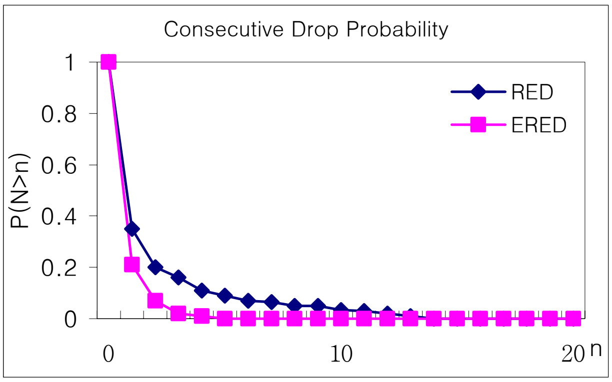


Fig. 2 Consecutive drop probability

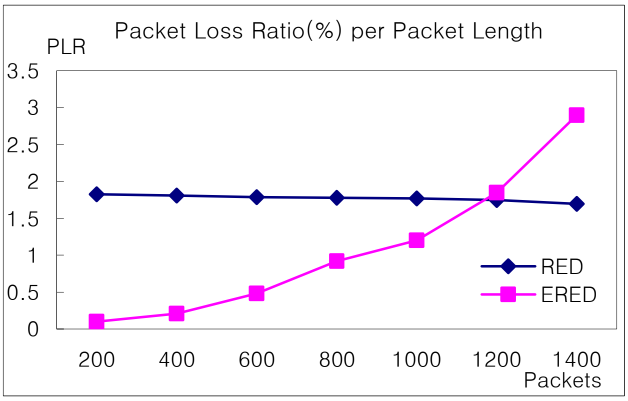


Fig. 3 Packet loss ratio per packet length

## 6 Conclusion

In this paper, firstly, we suggest a new RED mechanism and also analyze the PLC algorithms in the coders by an experimental method to get the marginal loss ratio somewhat roughly enough to handle by the PLC algorithm for every coder to be understood clearly by the end user. Secondly, our mechanism was testified to keep the marginal loss ratio, below 5%. It shows better performance in reducing the ratio of packet drop as more than 30% and makes shorten of the average bursty period as around 15%. In conclusion, considering packet length and RTT parameter as well as the correlation of loss rate between TCP and UDP packets in a VoIP gateway or any necessary nodes, which depends on the network configuration that consists of through the VoIP call connection, a new RED mechanism to reduce the packet loss for enhancing quality level of VoIP calls in the network even though it takes a time for computing, however, is viable and deserves to solve a performance limitation to meet the condition of human perception for VoIP call done by the PLC of the current recommended coders.

*References:*

[1]S. Floyd and V. Jacobson, Random Early Detection gateways for Congestion Avoidance, IEEE/ACM Transactions on Networking, vol.1 n.4, Aug. 1993, pp 397-413.

[2]ITU-T, Software tools for speech and audio coding standardization, ITU-T Recommendation G.191, Nov., 2000.

[3]ITU-T, Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs, ITU-T Recommendation P.862, Feb., 2001

[4]ITU-T, Appendix I: A high quality low-complexity algorithm for packet loss concealment with G.711, ITU-T Recommendation G.711, Sept., 1999.

[5]ITU-T, Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s, ITU-T Recommendation G.723.1, March, 1996.

[6]ITU-T, Annex A: Reduced complexity 8 kbit/s CS-ACEKP speech codec, ITU-T Recommendation G.729 Annex A, Nov., 1996.

[7]ETSI, Universal mobile telecommunications system (UMTS) AMR speech codec; Error concealment of lost frames, ETSI TS 126 091, June, 2002.

[8]B. Braden et al., Recommendations on Queue Management and Congestion Avoidance in the Internet," RFC 2309, Apr. 1998.

[9]Mushkin, M, Bar-David, I, Capacity and coding for the Gilbert-Elliot channels, IEEE Transactions on Information Theory, Vol. 35, Nov. 1989, pp 1277-1290.

[10]Vijay K. Varma, Testing Speech Coders for Usage in Wireless Communications Systems, Proceedings of IEEE workshop on Speech Coding for Telecommunication, Oct. 1993, pp 94-94.

[11]M. Mathis, et. al., The macroscopic behavior of the TCP congestion avoidance algorithm, Computer Communication Review, vol. 27 n. 3, pp 67-82, July 1997.

[12]J. Postel, User Datagram Protocol, RFC0793.

[13]T.Bonald, et. al., Analytic Evaluation of RED Performance, IEEE INFOCOM2000, pp. 26-30, Mar. 2000.