GOODPUT OPTIMISATION OF LONG-LIVED TCP CONNECTIONS IN A RAIN-FADED SATELLITE CHANNEL[1]

Nedo Celandroni, *Member, IEEE*, Erina Ferro, Francesco Potortì

***Abstract*— The optimization of the end-to-end throughput of a TCP connection over geostationary satellite links is a challenging research topic because the high delay-bandwidth product, together with a non-negligible random loss of packets, are conditions which differ considerably from the original environment for which TCP was originally designed. As a result, TCP performance is significantly impaired by the channel bit error rate. In this paper we investigate the application of different FEC (forward error correction) types/rates and different bit rates, for the optimization of TCP goodput, in transmissions over a rain-faded geostationary satellite channel, provided that the end-to-end protocols are left unaltered. We compare physical-level FEC techniques, such as convolutional encoding/Viterbi decoding and Reed Solomon, link-level erasure codes and their combinations, over a wide field of signal-to-noise conditions of the satellite channel. The case of multiple connections per link is also analyzed, in addition to that of a single connection per link. . In order to evaluate the throughput of TCP long-lived connections we used a fluid simulator ad-hoc developed.**

***Index Terms*— Long-lived TCP connections, rain fading, satellite channel, FEC techniques.**

I. INTRODUCTION

IN order to improve the efficiency of TCP over geostationary satellite channels, where improving the error rate is generally expensive, in our study we operated at the physical or link OSI levels by trading the bandwidth of the satellite link for packet loss rate. This technique does not interfere in any way with the normal behaviour of the TCP stack, but simply entails the choice of transmission techniques over the satellite link and the appropriate tune-up of relevant parameters. Generally speaking, for any given satellite transmission equipment, a number of parameters are chosen to obtain a target performance in terms of bit error rate (BER) and information bit rate (IBR). For a given available radio spectrum, antenna size and maximum transmission power, the selection of a modulation scheme and various FEC (forward error correction) types usually allows a wide range of choices. Commonly used satellite systems make such choices in a static fashion, by possibly permitting the user to change some of them manually, or in some cases by dynamically switching between a limited number of preset parameters [1]. The reason for switching from one set of parameters to another is that most satellite links have highly variable physical characteristics: for example, they are subject to a variable atmospheric attenuation of the signal. In addition, the bandwidth itself on a given link may vary with time, due to dynamic bandwidth allocation among different users of the same channel. In this paper we argue that, given an available radio spectrum, antenna size and transmission power, the selection of a modulation scheme and a FEC type allows choosing the link's BER and IBR that maximize the throughput of a TCP connection (also called goodput), i.e. the end-to-end transfer rate. This optimization can be made for different channel quality conditions whose variability is due, for instance, to the variable atmospheric attenuation of the signal. In order to make the proposed method usable in practice, optimal transmission parameters, for each channel condition, can be reported in look-up tables and dynamically applied in an adaptive fashion. A similar approach was followed by Barakat and Altman in [2], where the FEC is operated at link-level by using a block erasure code and the packet losses are assumed to be both independent and correlated [3]. In [4] we presented a method based on the point-by-point optimization of the transmission parameters by operating with a modem that had the ability to switch between BPSK (binary phase shift keying) and Q (quadrature) PSK and with a convolutional/Viterbi encoder/decoder. In this paper we assumed, as in [4], to face with additive white Gaussian noise (AWGN) so that we can consider independent packet losses. This assumption is reasonable when operating with geostationary satellite links and fixed user antennas. Here we extended the results obtained in [4], by introducing Reed Solomon and block erasure codes other than mixed techniques (concatenated codes). A wide scenario of the various techniques, including the one used in [2], was produced for different channel quality conditions, and the relevant comparison was made possible. We also considered the case in which multiple TCP connections share the same satellite link.

I. MINIMUM BER VERSUS AVAILABLE INFORMATION BIT RATES

Let us suppose that, for a given satellite transmission system, several of the parameters are dynamically tunable, so that the channel BER can be traded for the IBR. One

[1] This work is supported by the Ministero dell'Istruzione, Università e Ricerca of Italy under the "TANGO" FIRB project.

possibility is changing the modulation scheme; another, which is much more flexible and usually cheaper, is to change the type of FEC used. Most modern transmission systems provide for variable information bit rates by changing the FEC redundancy; some do so for each individual packet. FEC can be applied at different levels: using rate-compatible punctured convolutional codes [5] it is possible to seamlessly change the FEC while maintaining bit time synchronization of the data stream. Reed-Solomon codes can be applied at the phisical level, typically as outer codes with respect to an inner convolutional code. A similar effect can be obtained at the link level by splitting the TCP packets into smaller units, to which an erasure code is then applied prior to transmission [2], [6].

For a given transmission equipment, it is possible to define a family of graphs of minimum BER versus IBR for each channel condition. These indicate the gain in the BER obtained at expenses of the IBR, i.e. by increasing the redundancy of transmitted data by using one or more of the aforementioned methods. Each member of such a family of graphs is a non-decreasing function of the BER versus IBR.

Let us define $\boldsymbol{C}$ as the set of possible channel conditions $\boldsymbol{c}$ for some given communication equipment. For each element $\boldsymbol{c}$ of $\boldsymbol{C}$, a BER vs. IBR graph can be drawn representing the minimum BER that can be achieved by varying the different transmission parameters. In order to describe the available trade-offs for a given communication system, the minimum BER vs. IBR graph should be computed or measured for each element $\boldsymbol{c}$ of the set $\boldsymbol{C}$. In practical situations, $\boldsymbol{C}$ can be reduced to a manageable set of easily measurable channel conditions. For example, in a geostationary satellite communication system, $\boldsymbol{C}$ may be reduced to the span of atmospheric attenuations in the system operating range.

Since the set of available channel parameters is generally discrete, the graphs will be staircases. Let us define $\boldsymbol{P}$ as the set of possible parameter settings $\boldsymbol{p}$ of the equipment considered. For a given $\boldsymbol{c}$, each $\boldsymbol{p}$ will correspond to a point in the BER-IBR plane. The staircase that connects the minimum points is the graph relative to $\boldsymbol{c}$, and the corresponding $\boldsymbol{p}$ points are the parameters that define the operating points relative to the graph. Figure 1 shows an example of such graphs, where the parameter settings useful for obtaining the minimum BER are circled.

Armed with the graphs for the channel conditions $\boldsymbol{c} \in \boldsymbol{C}$, we want to find the optimum operating point for each $\boldsymbol{c}$, from the viewpoint of a TCP connection. Since the set of graphs is known *a priori* and only depends on the characteristics of the equipment, the search for the optimum operating point is performed only when tuning some given equipment, so its computational complexity is not an issue. The result of the tuning operation is a lookup table that is used during normal operations to get the transmission parameters $\boldsymbol{p}$ given the channel condition $\boldsymbol{c}$. In the case of dynamic bandwidth allocation, such a lookup table should be pre-computed for each possible bandwidth share. With reference to Fig. 1, we want to discover, for each channel condition $\boldsymbol{c}$, which is the best-circled point (and, consequently, the optimum transmission parameters $\boldsymbol{p}$) on the IBR-BER plane, from the viewpoint of maximizing the goodput of a TCP connection.

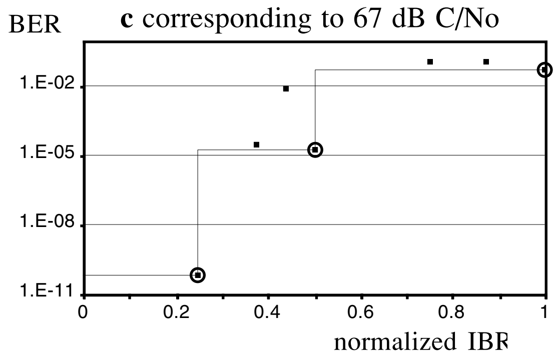

Fig. 1. Minimum BER vs. IBR graph for the C/N_0 condition of 67 dB.

Since we are referring to a geostationary satellite, by far the most important parameter for describing the channel conditions is the signal-to-noise ratio, so we only consider this parameter, expressed in decibels, as the carrier power to the one-sided noise spectral density ratio (C/N_0 [dB]). Let us further suppose that our modem is capable of working in a 12 dB range for C/N_0, from 66 to 78 dB, and that the granularity for the measurement of the attenuation is 1 dB. Thus, the set $\boldsymbol{C}$ consists of 13 possible different channel conditions $\boldsymbol{c}$. For each $\boldsymbol{c}$ we should build the minimum BER vs. IBR graph, as shown in Fig. 1.

I. TCP GOODPUT EVALUATION

We consider a Reno TCP implementation without the SACK [9] and with the Window Scale [10] options. Details on the TCP Reno congestion control mechanisms can be found in [4], [7], and [8].

Similarly to what assumed in [7], in our system a buffer of capacity B is associated with the bottleneck link, whose transmission rate is μ segments/s. Denoting by T the complete round-trip delay, consisting of the link latency τ plus the segment service time, we have $T = \tau + 1/\mu$. We also define the *normalized buffer size* $\beta = B/(\mu T)$.

The TCP fluid simulator used, called TGEP (TCP goodput estimation program), has been written in *Mathematica*™, and it is available in [11]. It is suitable for long-lived connections and high delay-bandwidth products. Given the number of connections N sharing a link of capacity μ [segments/s], the segment loss (assumed geometrically distributed) rate q, the size of the buffer β, the round trip delay T and the number b of segments acknowledged by each ACK, TGEP allows the estimation of the goodput within a chosen confidence interval at a chosen confidence level. The limit of TGEP is that it does not consider TCP timeouts and it does not model multiple losses within any interval T, but in cases in which the

TABLE I
GOODPUT OF 5 CONNECTIONS- NS2 AND TGEP COMPARISON

TCP Reno (no delayed ACKs) goodput of 5 connections sharing a link with a bottleneck rate μ=455 segments/s, τ=0.5 s, β=0.8. Confidence intervals are at 99% level.

			ns2 simulations				
q	Connect. #1	Connect. #2	Connect. #3	Connect. #4	Connect. #5	Total	Relation (1) T_o=2s
10^{-4}	0.198±3.1%	0.199±2.5%	0.199±0.5%	0.196±2.2%	0.202±2.3%	0.994±0.03%	2.656
10^{-3}	0.165±2.4%	0.169±2.6%	0.164±4.3%	0.169±5.3%	0.169±1.3%	0.836±0.67%	0.8345
10^{-2}	0.0469±0.64%	0.0471±2.3%	0.470±1.5%	0.0466±3.1%	0.0468±1.6%	0.234±0.4%	0.2478
10^{-1}	0.0090±3%	0.0090±0.8%	0.0090±0.7%	0.0090±0.6%	0.0092±2%	0.0452±4%	0.0429
			TGEP output				
q	Connect. #1	Connect. #2	Connect. #3	Connect. #4	Connect. #5	Total	Relation (1) T_o=1s
10^{-4}	0.199±3.5%	0.198±0.6%	0.199±5.8%	0.199±3.7%	0.199±1.89%	0.994±0.01%	2.657
10^{-3}	0.169±5.1%	0.168±7.7%	0.166±8.2%	0.171±7.5%	0.171±7.2%	0.845±1%	0.8375
10^{-2}	0.0498±5.1%	0.0509±13.8%	0.0515±8.4%	0.0495±6.7%	0.0494±5.8%	0.248±1%	0.2565
10^{-1}	0.0156±1.3%	0.0157±1.7%	0.0156±1.6%	0.0155±1.5%	0.0157±1.9%	0.0781±1%	0.0568

random loss rate q is << 1, which are significant cases when the delay-bandwidth product is high, TGEP performs satisfactorily.

Unless otherwise specified all results of TGEP in the rest of the paper have been obtained with a confidence interval of ±1% at 99% level. In order to obviate to the non-consideration of TCP timeouts, we can also consider the expression (1) of the normalized TCP throughput [12], valid for unlimited bandwidth links:

$$T_g = \sqrt{\frac{3}{2}}\,\frac{1}{\mu\left(\sqrt{bq}T + T_o \min\left[1,\ \sqrt{3bq/8}\right]q(1+32q^2)\right)},\quad (1)$$

where T_o is the value of the timeout, which can be assumed as max{1, 4 RTT} [13], where RTT is the minimum round trip time experienced by the connection. In Table I we report the comparison of TGEP, relation (1) and ns2 simulation for 5 connections. We can observe a good agreement between ns2 and relation (1) for low values, and between ns2 and TGEP for high values of the goodput. This follows our expectations, because relation (1) considers TCP timeouts and unlimited bandwidth, while TGEP considers the bottleneck rate of the link and does not keep timeouts into account. A suitable threshold to use when selecting one or the other procedure can be assumed to be equal to 50%, even if TGEP agrees with relation (1) (and with ns2 simulation) also for much lower values of the goodput.

IV TCP GOODPUT WITH DIFFERENT FECS AND MODULATION SCHEMES

Let us now return to our original problem, which was to find, for each channel condition $\boldsymbol{c}$, the point on the relative graph where the TCP goodput is maximum. With reference to Fig. 1, the optimum operating point is chosen from the circled points on the graph by making use of TGEP or relation (1).

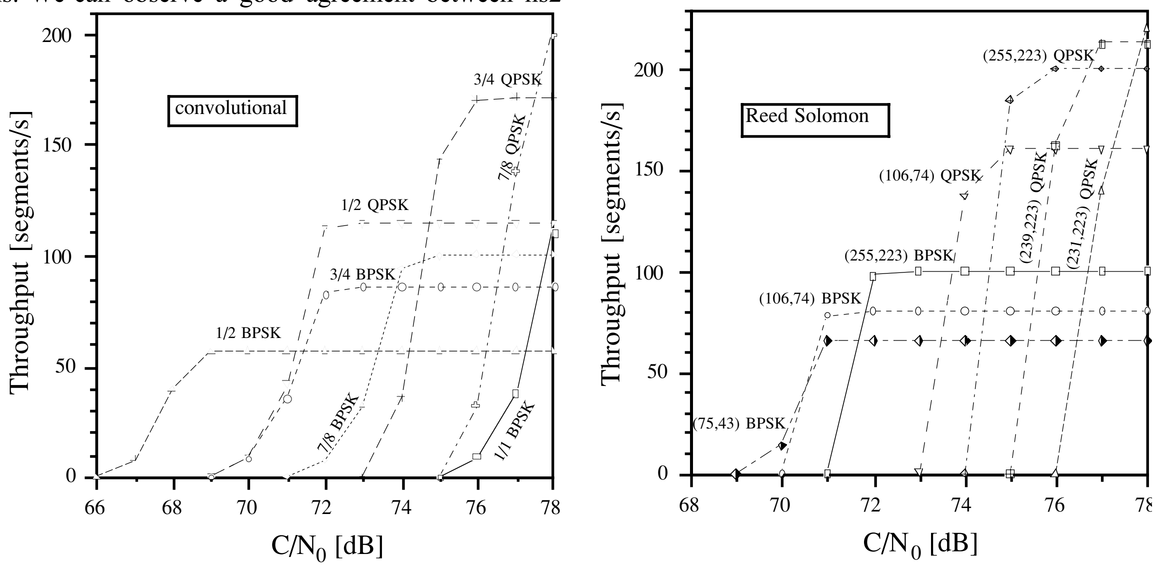

Fig. 2. TCP goodput versus C/N_0 for convolutional (CV) and Reed-Solomon (RS) codes.

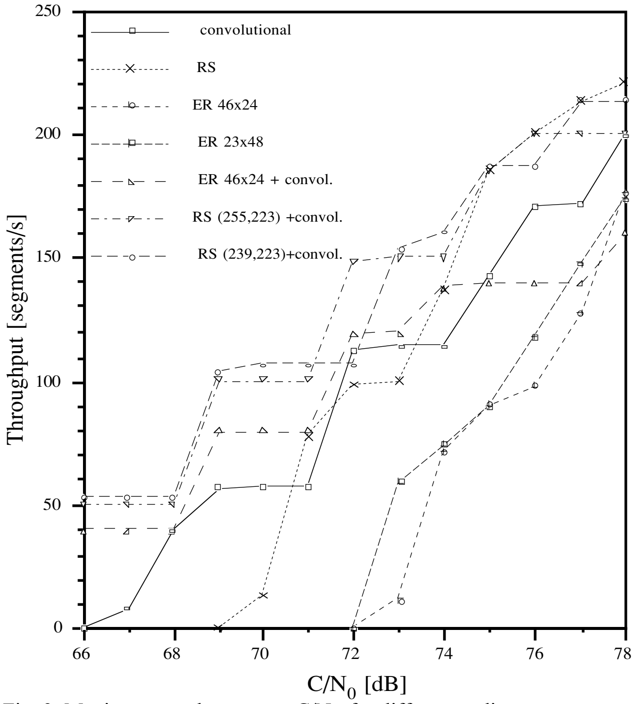

Fig. 3. Maximum goodput versus C/N_0, for different coding types.

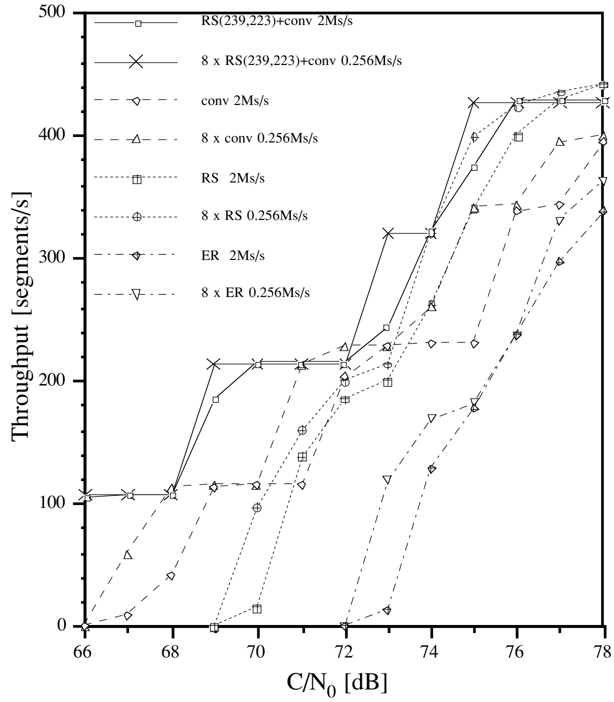

Fig. 4. Goodput envelopments of one TCP connection and the aggregated goodput of 8 TCP connections that share the same amount of bandwidth, versus C/N_0, for different coding types.

We estimated the goodput of both a single TCP connection and of multiple TCP connections, for different conditions of the channel quality. We assume having a modem, working at 2 Msymbols/s, which is able to switch between BPSK and QPSK modulation schemes. Thus the resulting channel bit rate is 2 and 4Mb/s, respectively. This choice was made both in order to represent a commonly used piece of equipment and because the spectrum occupancy is the same in both cases. We also assume that the channel bandwidth is shared in TDMA (time division multiple access) mode among the users and that the considered TCP connection received an assignment of one-half of the carrier capacity. The C/N_0 ratio available at the demodulator input spans between 66 and 78 dB. C/N_0 depends on the signal attenuation, e.g. due to atmospheric conditions (rain fading). The bit rate of uncoded data is thus 2 or 1Mb/s when the modem works in QPSK or BPSK, respectively. This situation was chosen in order to facilitate the comparison with the case of multiple connections. The decision concerning the selection of the transmission parameters is made at the earth station on the basis of the link quality, which is supposed to be known.

The FEC codes we considered are a convolutional coder/Viterbi decoder (CV) with puncturing [5, 14] over BPSK/QPSK modulated symbols, Reed Solomon (RS) codes [15, 16], and erasure codes (ER) [6]. The modulation scheme selection and CV coding/decoding is performed at the physical level; RS coding/decoding is preferably performed at the physical level, but it could be made by software at link level, while ER coding/decoding is performed at link level [6]. In SCPC (single channel per carrier) mode, to obtain the same situations as in TDMA, we must consider a modulation rate of 1Msymbols/s. When a carrier of 2Msymbols/s is entirely assigned to the TCP connection, the results are still valid if we decrement by 3 dB the range of C/N_0 ratio (63-75).

Figure 2 reports the results relative to the CV and RS codes employed in the single TCP connection per link case. All points lying on the envelopment curves represent the maximum goodput obtainable with the available hardware for each given channel quality condition. In [4] it is shown that the adoption of a different criterion, such as the commonly used one which imposes the BER below a certain threshold, produces results that may sensibly differ from optimality, even if the threshold is made variable with the link capacity. In the RS case we used shortened codes which derive from the RS(255,223). The erasure code relative to 23-byte blocks performs better than the 46-byte one if the code is used alone. When the code is concatenated with the CV one, the 46-byte code performs slightly better than the 23-byte one.

In order to make comparisons easier, in Fig. 3 we reported the envelopes of the cases shown in Fig. 2 plus two ER codes, an ER+CV and an RS+CV code. We note that, as expected, the best performance is achieved by using concatenated codes, in particular by using RS plus CV ones. RS code outperforms the CV one only at high C/N_0 values, while its performance is worse at low C/N_0 values. ER codes exhibit the worst performance; their most significant peculiarity is that a number of redundant blocks in a packet may not only be corrupted but even be lost, to allow a successful segment delivery, and an AWGN channel does not take sufficient advantage of this.

By using the same point-by-point optimization method, we made a comparison between the goodput achieved by a single connection with the aggregated goodput of 8 TCP connections each of which has its own buffer and a bandwidth share equal to 1/8 of the entire bandwidth. The comparison of the envelopment curves, which express the maximum values of the goodput versus C/N_0 for different coding types, is reported in Fig. 4.

The whole channel bandwidth is 4 or 2Mb/s for QPSK or BPSK modulation schemes, respectively.

When the bandwidth is divided among 8 connections and it is assigned individually to the requesting users, the transmission parameter optimization is made for TCP bottleneck rates of 512 (256) kb/s in the QPSK (BPSK) case.

In the other case, the optimization is made for the full rate and a single TCP connection, because all the bandwidth is assigned to a single user. Figure 4 shows that the total goodput sensibly differs in the two cases (single and multiple connections per link, respectively), for many of the C/N_0 values considered and for all coding types. The difference is null when the selection of the transmission parameters only allows excessively low values of the SER.

When multiple TCP connections share the same link, it was empirically observed in [7] by making use of simulation, that the bandwidth is shared evenly among connections if they have the same latency. As the latency introduced by a geostationary satellite is quite high (half a second) it is reasonable to assume that the additional latency introduced by terrestrial hops in the entire link paths is negligible with respect to the satellite one. We thus assume that all connections obtain an equal share of the link bandwidth. This is strongly supported by our simulations as well (see Table I). In addition, we simulated the case in which 8 TCP connections share the same buffer and the same bottleneck link. The results obtained for the aggregated goodput of all connections do not sensibly differ from the one obtained by considering 8 separated links and buffers with 8 connections (reported in Fig 4). Furthermore, the difference of the two results decreases as the size of the buffer increases. This is a useful issue because it allows reducing the number of the look-up tables necessary to take the optimum transmission parameters. In fact, provided that the buffer employed is sufficiently large, different tables for a different number of TCP connections are not required. It is sufficient to divide the allocated bandwidth by the number of connections and use the resulting value to enter the table that gives the optimum transmission parameters as a function of the bandwidth in the current channel condition (C/N_0).

VI. CONCLUSIONS

By using a rapid simulation tool (TGEP) ad-hoc developed, we depicted a scenario that shows the maximum achievable TCP goodput in a realistic situation in which a bandwidth portion of a geostationary satellite transponder is assigned to a user in TDMA mode, for single or multiple TCP connections. Several coding techniques, used for two different modulation schemes (B/QPSK), have been applied to TCP connections, in a wide range of signal-to-noise ratios of the AWGN channel considered. The case studied closely emulates a situation in which the received signal level is affected by a variable attenuation due to different atmospheric conditions. The comparison of the various coding techniques shows the supremacy of the CV case, especially when it is concatenated with RS codes, over the ER codes. This is the result in an AWGN channel. However, we think that the strong peculiarity of the ER codes, which allow the recovery of a packet even with some lost blocks, can be better exploited when different types of impairment affect the channel.

REFERENCES

[1] C. B. Cox and T. A. Coney, "Advanced Communications Technology Satellite (ACTS) fade compensation protocol impact on Very Small Aperture Terminal bit error rate performance", *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 2, February 1999, pp. 173-179.

[2] C. Barakat and E. Altman, "Bandwidth tradeoff between TCP and link-level FEC", *Computer Networks*, vol. 39, no. 2, June 2002, pp. 133-150.

[3] E. N. Gilbert, "Capacity of a Burst noise Channel" *Bell Systems Technical Journal*, vol. 39, (1960), pp. 1253-1266.

[4] N. Celandroni, F. Potorti', "Maximizing single connection TCP goodput by trading bandwidth for BER", *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 63-79, 2003.

[5] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications", *IEEE Transactions on Communications*, vol. 36, no. 4, April 1988, pp. 389-400.

[6] L. Rizzo, "Effective erasure codes for reliable computer communication protocols", *Computer Communication Review*, vol. 27, no. 2, April 1997, pp. 24-36.

[7] T. V. Lakshman and U. Madhow, "The performance of TCP/IP for networks with high bandwidth-delay products and random loss", *IEEE/ACM Transactions on Networking*, vol. 5, no. 3, June 1997, pp. 336-350.

[8] M. Allman, V. Paxson and W. S. Stevens, "TCP congestion control", *RFC 2581*, April 1999.

[9] M. Mathis, J. Mahdavi, S. Floyd and A. Romanow, V. Jacobson and R. Braden, "TCP selective acknowledgement options", *IETF, RFC 2018*, October 1996.

[10] V. Jacobson, R. Braden and D. Borman, "TCP extensions for high performance", *IETF, RFC 1323*, May 1992.

[11] http://wnet.isti.cnr.it/software/tgep.html

[12] J. Padhye, V. Firoiu, D.F. Towsley and J.F. Kurose, "Modeling TCP Reno performance: a simple model and its empirical validation", *IEEE/ACM Trans. Networking*, vol. 8, pp. 133-145, 2000.

[13] S. Shenker and L. Zhang, "Some Observations on the Dynamics of a Congestion Control Algorithm", *ACM SIGCOMM-CCR*, vol. 20, no. 5, October 1990, pp. 30-39.

[14] S. Lin and D.J. Costello, "Error Control Coding: Fundamental and Applications": Prentice-Hall, 1983.

[15] M. Riley and I. Richardson, "Reed_Solomon Codes", online at http://www.4i2i.com/reed_solomon_codes.htm, November 2002.

[16] M. K. Simon, J.K. Omura, R. A. Scholtz and B. K. Levitt, "Spread Spectrum Communications Handbook", revised edition: McGraw-Hill, 1994.