Overcoming Intrinsic Losses with a Physical-transport Cross-layer Control System for Low-SNR Links

Claudio Estevez, *Member, IEEE,* Cesar Azurdia-Meza, *Member, IEEE,* Sandra Céspedes, *Member, IEEE*

***Abstract*—Communication over low-SNR environments faces various challenges and data detection designing can be arduous. An aggregative sampling technique with TCP feedback is proposed to transmit in low-SNR channels. The proposed scheme overcomes intrinsic losses by having a physical-transport cross-layer interaction. Samples are aggregated to make a single bit decision. As the quantity of aggregated samples is increased, the bit error rate (BER) is reduced. The transport-layer loss information is fed back to the physical layer to dynamically control the amount of redundancy, therefore reducing intrinsic loss. Results show that with an SNR of 0 dB the system is able to reach over 10 Mbps with a BER of near 10^{-9}. It is demonstrated that by implementing the proposed technique it is feasible to reduce the BER by a factor of 10^9 by reducing the effective throughput by a factor of 3. For SNR environments of over -10 dB a BER of 10^{-8} is achieved. Performance improvement of 11 dB or more is obtained compared to the analyzed techniques.**

***Index Terms*—cross-layer, intrinsic loss, low-SNR, noise, physical-transport.**

I. Introduction

Low signal-to-noise ratio (SNR) systems are relatively pervasive; consequently, the challenge of successfully detecting the transmitted signal in noisy environments has an axiomatic urgency. Various techniques have been proposed to address low-SNR detection, including coding [1], [2], [3], improved phase detection [4], stochastic resonator [5], energy-based [6], pre-filtering [7], [8], and spread spectrum [9] solutions. These systems report poor BER performance (greater than 10^{-4}) for SNR values much lower than 0 dB. For standard-size packets (∼12,000 bits) with no error correction (only detection, such as checksum) the probability of packet loss in these systems is near 100%. Even recent work [5] in low-SNR link detection struggle to obtain low BERs, demonstrating that this is an important on-going topic. In this work, the proposed system is compared to a coding and a spread spectrum case. Using a discrete integrator technique and a physical-transport cross layer control design, it is possible to achieve high reliability detection under very low-SNR schemes, without relinquishing considerable throughput. The transmission control protocol (TCP) is designed to manage data flow based on congestion assuming losses are extrinsic, nevertheless, it can be configured to provide a dynamic control over physical layer parameters with the objective of reducing intrinsic loss. This technique is best suited for lossy-channel systems with low number of simultaneous TCP transmissions and large round-trip time (RTT). Because the physical layer is unique to this system, it is unlikely it can communicate with standard networking equipment, therefore it is necessary to surround the low-SNR link with TCP proxies to prevent out-of-order segments, as shown in Figure 1. An additional advantage of implementing TCP proxies is the reduction of the perceived RTT, hence increasing the end-to-end throughput. Employing TCP proxies surrounding low-performance links has been previously employed in satellite networks [10], but not under a cross-layer physical-transport system.

Authors are with the Department of Electrical Engineering, Universidad de Chile, Santiago, Chile. e-mail: cestevez@ing.uchile.cl
Work partially funded by project FONDECYT 11121655 and 11140045

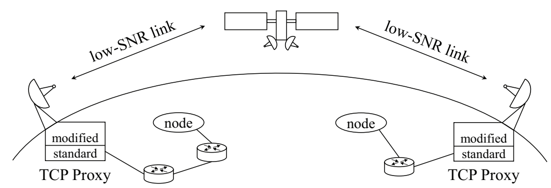

Fig. 1. Architecture of proposed cross-layer solution showing the placement of the TCP proxies.

II. Cross-layer Control System Overview

The proposed system is aimed at improving the detection of digital signals by implementing a control system that incorporates elements from the physical and the transport layer and have them interact. The versatility of this mechanism supports any physical layer solution in which the redundancy can be dynamically varied and any transport protocol with a predictable congestion window size ($cwnd$). Here, an aggregative sampling technique alongside traditional TCP is presented.

A. Cumulative Sampling Technique

The proposed physical counterpart of this strategy consists of a discrete technique referred here as the aggregative sampling technique. The strategy consists on representing a single bit by one or more pulses. Upon arrival at the receiver each pulse is sampled and each sample is added to the previous sample for a predetermined amount of redundancy. The system is based on a binary antipodal pulse. The decoding operation is the sum of all the samples that correspond to a single bit. Notice that employing a square signal with unit gain is equivalent to employing repetition coding. This is a low-level software modification, and it will not affect the hardware, hence, it does not require a forklift upgrade.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LCOMM.2015.2490672

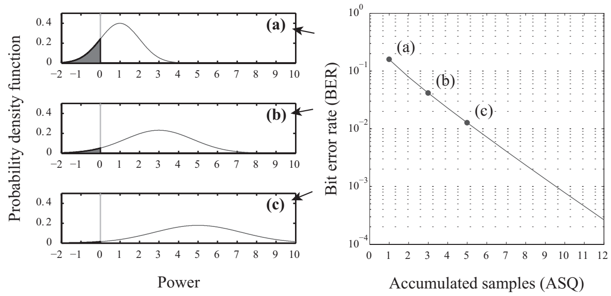

Fig. 2. Portraying the BER decrease with ASQ increase for 0-dB SNR.

It is essential to study the effect of sample aggregation on the BER. The probability density function (pdf) of the sample amplitude, assuming augmented white Gaussian noise (AWGN), for the positive single-sample case is described by the expression $\phi_X(x) = (\sqrt{2\pi\sigma^2})^{-1} e^{-\frac{(x-A_b)^2}{2\sigma^2}}$ or simply $X \backsim \mathcal{N}(A_b, \sigma^2)$, where σ^2 is the noise variance. Expanding this to the multiple-sample case, where each bit is represented by S pulses, has $X_S = \sum_{i=1}^{S} X$, which yields $X_S \backsim \mathcal{N}(SA_b, S\sigma^2)$ assuming all X are independent of each other. S is the amount of modulation redundancy and referred to as the Aggregated Sampling Quantity (ASQ). Figure 2 shows the cases for (a) single sample ($S=1$), (b) $S=3$, and (c) $S=5$. The BER P_b for the multiple-sample case is given by the following expression:

$$\begin{aligned} P_b = {} & P(b{=}1)P(x_S \le \mathcal{T}\,|\,b{=}1) + \\ & P(b{=}0)P(x_S \ge \mathcal{T}\,|\,b{=}0) = \\ & \frac{P(b{=}1)}{\sigma\sqrt{2\pi S}} \int_{-\infty}^{\mathcal{T}} e^{-\frac{(x-SA_b)^2}{2S\sigma^2}} dx + \frac{P(b{=}0)}{\sigma\sqrt{2\pi S}} \int_{\mathcal{T}}^{\infty} e^{-\frac{(x+SA_b)^2}{2S\sigma^2}} dx, \end{aligned} \quad (1)$$

where $\mathcal{T}$ is the decision threshold, b is the received bit, and A_b is the pulse amplitude. Assuming that $P(b=1) = P(b=0) = \frac{1}{2}$ and the threshold $\mathcal{T} = 0$ then (1) is simplified to:

$$P_b = \int_{-\infty}^{0} \frac{e^{-\frac{(x-SA_b)^2}{2S\sigma^2}}}{\sigma\sqrt{2\pi S}} dx = \Phi\left(\frac{-A_b}{\sigma}\sqrt{S}\right) \quad (2)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, i.e., $\mathcal{N}(0,1)$, hence:

$$P_b = \frac{1}{2} + \frac{1}{2} erf\left(-\frac{A_b}{\sigma}\sqrt{\frac{S}{2}}\right) \quad (3)$$

The shaded area of Figure 2(a), 2(b), and 2(c) represents the P_b. Figure 2 is included to show that using the aggregative sampling technique the BER decreases with increasing ASQ, but the BERs that are of interest to our work are significantly lower and therefore the aggregated samples are greater.

B. Physical-Transport Cross-layer Interaction

The main objective of the cross-layer interaction is to achieve good and stable performance under low-SNR environments. To achieve this, a control system is proposed where the feedback capabilities of TCP are combined with the aggregative sampling technique described earlier. The proposed mechanism is designed to maintain segment synchronization, where in case of loss the system will recover in (at most) one RTT, as does conventional TCP. The throughput is dependent on the $cwnd$. The initial $cwnd$ in traditional TCP is set to 2, 3, or 4 [11]. The transport layer interacts in two ways with the physical layer: an indirect path for the data, and a direct path for the control information. The indirect path has to undergo the network (L3) and data-link (L2) layers for appropriate processing and encapsulation. The direct path is an interface to share TCP information, such as RTT, $cwnd$, segment loss detection, and any other useful information that the design requires. In this work, only the $cwnd$ and RTT information are used, the loss information is conveyed indirectly through the $cwnd$. The ASQ S varies as a function of $cwnd$, as described by the expression $S = \left\lfloor \frac{f_p RTT}{cwnd\, WPS} \right\rfloor$, where f_p [pulses/second] is the pulse transmission rate, the whole packet size (WPS) is defined here as $WPS = MSS$ + overhead [bits/packet], MSS is the maximum segment size, and the S [pulses per bit] is the ASQ. RTT and $cwnd$ have units of seconds and packets, resp. Because of the high RTT and low $cwnd$ values the resulting ASQ values can be enormous, this could result in overkill. An S_{MAX} value can be established such that the S value always resides within the 1 to S_{MAX} range. Both end-terminals must have exactly the same $cwnd$ algorithm than the opposite end-terminal so that the ASQs are always synchronized. If for any reason there was a mismatch, the receiver will be unable to interpret the data. The mechanisms are shown in Figure 3.

The $cwnd$ control is the same as in traditional TCP. To maintain synchronization, the receiver knows the algorithm used by the transmitter. The connection is initialized using S_{MAX} to have the highest probability of success during TCP's three-way handshake. It continues the transmission using the defined relation $S(cwnd)$. If the checksum is successful, the segment was received with no errors and the receiver sends an acknowledgement and increases its $cwnd$ value according to the selected TCP version, in this case traditional TCP. This continues on until the checksum indicates that an error has occurred. In this case, the receiver sends a duplicate acknowledgement while artificially inflating the $cwnd$ for a period of RTT, after RTT the transmitter would have received notice of segment loss and adjusted its $cwnd$, and hence its S value. If both systems fail to find synchronization after RTT, the system will timeout (standard TCP behavior) and will resume at the defined initial $cwnd$, therefore both systems will resume at an ASQ of S_{MAX}. So even in unfavorable conditions the system should recover in at most $2\,RTT$ seconds. It is important to mention that the receiver only keeps track of $cwnd$, but does not perform any congestion control. It only sends feedback to the transmitter. The feedback or acknowledgement segments are always sent using S_{MAX} to ensure the best chance of successful reception. Even if one acknowledgement arrives with errors, as long as the following segment acknowledges a higher sequence number the system will continue unaffected. Notice from the required architecture (Section I, Figure 1) that there cannot be any missing segments or out-of-order segments, only segments with errors. This trait is necessary for successful synchronization of $cwnd$ and therefore S values.

Using this cross-layer solution it is possible to configure a

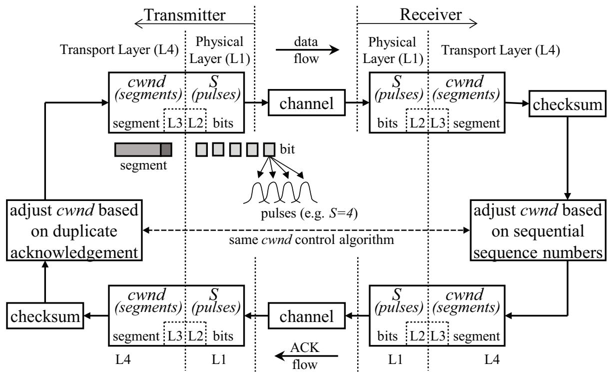

Fig. 3. Physical-transport cross-layer control system.

bit error tolerance. Given the highest tolerable BER $P_{b_{TOL}}$ of a system, then (3) can be solved for the ASQ value. This will yield the lowest allowable ASQ value, denominated S_{MIN}, that will maintain the BER below the $P_{b_{TOL}}$ threshold, which is:

$$S_{MIN} = \left\lceil 2\left(\frac{\sigma}{A_b} erf^{-1}(2P_{b_{TOL}} - 1)\right)^2 \right\rceil \tag{4}$$

This expression will be analyzed in the following section.

III. Results and Discussion

To demonstrate the advantages of the cross-layer control system, two tests are performed. The first shows what S_{MIN} values correspond to the respective $P_{b_{TOL}}$ tolerances. The second shows the throughputs obtained using the proposed system and compares it to other techniques. Because, to the best of our knowledge, there are no physical-transport cross-layer solutions, we have chosen some physical layer work combined with traditional TCP with no cross-layer interaction.

In the first test, the S_{MIN} is obtained analytically for different BER tolerances, see Table 1. The SNR for this test is kept constant at -10 dB. It is observed that a logarithmic decrease of the tolerance causes an approximately linear increase of ASQ. This is very advantageous, to put this into perspective lets observe a particular case. The BER of 10^{-6} requires 226 samples and 10^{-15} requires 631 samples. Since the ASQ is inversely proportional to the throughput, this implies that reducing the throughput by a factor of three will approximately reduce the BER by a factor of 10^9. Note that establishing an ASQ lower bound results in a throughput upper bound. By having a minimum ASQ value (S_{MIN}) a Packet Loss Rate (PLR) guarantee is attained, which is commonly found in Quality of Service (QoS) requirements. The minimum tolerable PLR is $PLR_{TOL} = 1-(1-P_{b_{TOL}})^{WPS}$, where WPS must be in unit of bits.

TABLE I
S_{MIN} FOR DIFFERENT TOLERANCE VALUES

Tolerance	10^{-6}	10^{-9}	10^{-12}	10^{-15}
S_{MIN}	226	360	495	631

In the second test, a simulation is performed to examine the throughput performance of the proposed system under different SNR conditions. The cross-layer solution is compared to: 1) a blind estimation on direct-sequence spread spectrum (DSSS) technique [9] and 2) a union bound for energy allocation optimization [1] technique using, in both cases, traditional TCP with no cross-layer interaction. Both of these works use a binary modulation scheme under an additive white Gaussian noise (AWGN) channel. The simulation parameters are specified in Table II. An $S_{MAX} = 360$ is chosen to limit the value of S (can reach up to 21,114 with the chosen simulation parameters), and $S = 360$ corresponds to a $BER = 10^{-9}$, which is near the resolution of the simulation (lower BERs are perceived as lossless transmissions). An $S_{MIN} = 0$ shows the bitrate can reach the native pulse transmission rate f_S. f_S depends highly on the type of network, $f_S = 10^9$ is a pervasive wired bitrate and reasonable assumption for next-generation wireless networks. The file size determines the BER resolution, 100 MB is chosen to have $BER \approx 10^{-9}$. The remaining parameters are standard (IEEE 802.3, RFC791, RFC793). To compare this physical layer work to the proposed work it is necessary to compute the PLR, based on the BER, i.e., $PLR = 1-(1-BER)^{WPS}$ (WPS in unit of bits). Using the PLR, the steady-state throughput can be computed using the well-known expression from [12], $T = \frac{MSS}{RTT\sqrt{2p/3}+T_0 min(1,3\sqrt{3p/8})p(1+32p^2)}$, where T_0 is the initial timeout time, T is the steady-state throughput, and p is the PLR. The steady-state throughput is defined here as the average throughput after TCP has exited the initial slow start state. The data for the cases where $PLR > 0.99$ is not considered as at these rates there is a very low probability of completing the transmission. The S_{MAX} limit of 360 samples will cause an elbow effect in the throughput performance near the -10 dB SNR region. This is beneficial if the actual link SNR is above the -10 dB region because at a segment loss event, TCP resets the $cwnd$ to S_{MAX} and if it where too large it will take more time to reach the desired region of operation. Increasing the value of S_{MAX} will improve further the performance. From Figure 4 it can be seen that the proposed solution improved the traditional method by 25 dB. When compared to the DSSS work, an improvement of ∼11 dB is observed (at $BER = 10^{-6}$). For the coding technique the improvement is of ∼14 dB (at $BER = 10^{-6}$). At SNR values of -8 dB the ASQ steady-state values are near 100, which means that a throughput of 1/100 the native pulse rate can be achieved, in this case 10 Mbps. This is a

TABLE II
SIMULATION PARAMETERS

RTT	500 milliseconds
S_{MAX}	360 samples
S_{MIN}	no restriction (1 sample)
f_p	10^9 pulses/second
MSS	1460 bytes
header size	20 bytes
T_0	1 second
consecutive retransmissions before quitting	5 attempts
initial $cwnd$	2 segments
file size	100 megabytes

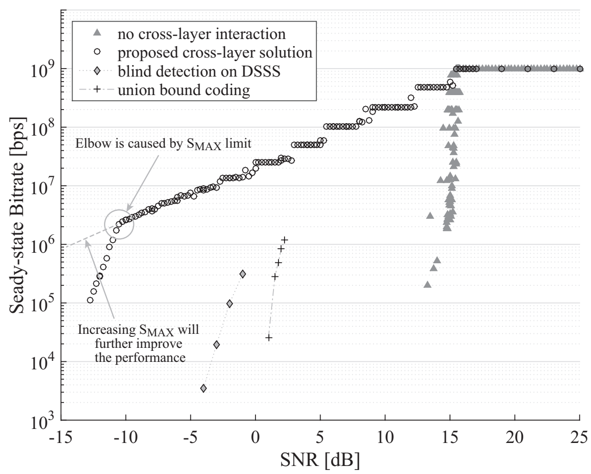

Fig. 4. Effect of the SNR on the effective throughput.

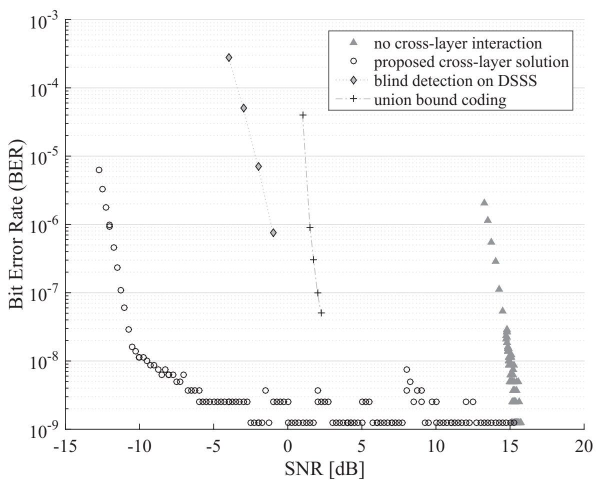

Fig. 5. Effect of the SNR on the BER.

reasonable trade-off for enabling transmissions at low-SNR environments. From Figure 5, it can be observed that values of $SNR > -10$ dB yield $BER < 10^{-8}$. This can be increased or decreased by changing the $S(cwnd)$ behavior described in Section II-B. Synchronization of pulses is a difficult task under noisy channels but having high redundancy repetition coding and the use of integrators/adders at the receiver, this problem can be alleviated.

Using the aggregative sampling technique the cost of having a system with low-error tolerance is fairly low. It is convenient to design systems with error tolerances in the order of 10^{-15} or lower since it will have a small impact on the effective throughput. Note that as the effective throughput approaches the value of f_S the granularity decreases, this is a result of having less samples per bit.

IV. Conclusion

A physical-transport cross-layer solution is proposed for increased performance in low-SNR environments. Results of the proposed aggregative sampling technique with a TCP-based feedback control have shown high tolerance to noisy conditions. For example, systems with SNR levels of 0 dB can potentially achieve over 10 Mbps. Since the system aggregates sample information, the required modification to implement the proposed scheme is software-based, hence a forklift upgrade is not necessary. The proposed technique allows an easy configuration of the BER tolerance by setting a minimum amount of redundancy (S_{MIN}). This enables the configuration of a PLR tolerance guarantee, typically found in QoS requirements. An important contribution of this work is that this technique can decrease the BER significantly with a relatively small impact on the effective throughput, specifically, results have shown a factor of 10^9 decrease in BER with a factor of 3 cost on the effective throughput. In terms of throughput performance, the cross-layer solution can obtain the throughput of the no-cross-layer-interaction method with SNR values that are approximately 25 dB smaller, which is a very significant difference. When compared to other physical layer work, but without cross-layer interaction, significant improvement is obtained. Compared to the blind estimation on DSSS and the union bound coding work an $\sim$11 dB and $\sim$14 dB SNR improvement is observed, respectively. For this scenario the BER was lower than 10^{-8} for SNR values greater than -10 dB.

References

[1] S. Xia, Z. Weidang, and L. Ping, "Improving the bit error rate performance by optimising energy allocation based on union bound," *IET Communications*, vol. 8, no. 18, pp. 3366–3371, 2014.

[2] M. Thakur and M. Médard, "On optimizing low SNR wireless networks using network coding," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*. IEEE, 2010, pp. 1–5.

[3] R. Koetter and M. Médard, "An algebraic approach to network coding," *Networking, IEEE/ACM Transactions on*, vol. 11, no. 5, pp. 782–795, 2003.

[4] W. Jiang and Y. Cui, "Performance analysis of MPSK phase detectors for carrier synchronization PLLs at low SNRs," *Communications Letters, IEEE*, vol. 18, no. 12, pp. 2133–2136, Dec 2014.

[5] J. Liu, Z. Li, L. Guan, and L. Pan, "A novel parameter-tuned stochastic resonator for binary PAM signal processing at low SNR," *Communications Letters, IEEE*, vol. 18, no. 3, pp. 427–430, March 2014.

[6] P. Cheraghi, Y. Ma, Z. Lu, and R. Tafazolli, "A novel low complexity differential energy detection for sensing OFDM sources in low SNR environment," in *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*. IEEE, 2011, pp. 378–382.

[7] F. Ferdousi, A. Tahseen, M. Sharmin, M. Murshed, I. Kabir, N. Jahan, M. Khan, and M. K. Hassan, "Parameter estimation of AR systems at a very low SNR using prefiltering in the autocorrelation domain," in *TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region*, vol. 4. IEEE, 2003, pp. 1377–1380.

[8] J. Kim and J. Cioffi, "Transmission optimization with a space-time filter at low SNR wireless environment," in *Global Telecommunications Conference, 1999. GLOBECOM'99*, vol. 1. IEEE, 1999, pp. 889–893.

[9] D. Li, X. Gu, and Q. Guo, "Estimation-based blind detection in low SNR on direct-sequence spread spectrum signal," in *Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on*. IEEE, 2010, pp. 1–4.

[10] N. Itoh and M. Yamamoto, "Proxy-based TCP with adaptive rate control and intentional flow control in ad hoc networks," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*. IEEE, 2008, pp. 1–6.

[11] M. Allman, V. Paxson, and W. Stevens, "RFC 2581: TCP congestion control," 1999.

[12] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP throughput: A simple model and its empirical validation," in *ACM SIGCOMM Computer Communication Review*, vol. 28, no. 4. ACM, 1998, pp. 303–314.