International Conference on Automation, Control Engineering and Computer Science (ACECS)
Proceedings of Engineering and Technology – PET
Vol.20 pp.21 - 24

Optimization of a modular ad hoc land wireless system via joint Source-Network Coding for correlated sensors

Dina Chaal
ENSA-Tangier
LTI Laboratory, Morocco
Email: dina.chaal@gmail.com

Abdelouahid Lyhyaoui
ENSA-Tangier
LTI Laboratory, Morocco
Email: lyhyaoui@gmail.com

Frédéric Lehmann
Télécom SudParis
SAMOVAR Laboratory, France
Email: frederic.lehmann@telecom-sudparis.eu

***Abstract*—We address the problem of collecting data in a wireless sensor network, where network coding is used for data transmission. The temporal correlation between the measurements is exploited to recover the data at the receiver. For this purpose, network encoding operations are considered in conjunction with lossy source compression, performed by an LDGM code set generated during transmission. The decoding is carried out using the iterative doping algorithm based on the sum product algorithm, on a graph which represents the LDGM encoding equations. Consequently, we first study the design tradeoffs of LDGM-based lossy source coding for significant parameters, such as packet size reduction and distorsion. In addition, we aim to show the prospective impact of the network coding on a wireless sensors network. This involves identifying how linear codes improve key performance criteria for networks such as rate, delay, and convergence time. Our work was applied to a real case study to highlight consistency and efficiency of our proposed system.**

***Index Terms*—Wireless Sensor Networks, Adhoc Networks, Lossy Compression, Source and Network Coding, Low-density generator matrix codes (LDGM), Random Linear Network Coding (RLNC), Differential pulse code modulation system (DPCM).**

I. Introduction

In recent years the widespread availability of wireless communication and handheld devices has stimulated research on self-organizing networks that do not require a pre-established infrastructure. These ad hoc networks, as they are commonly called, consist of autonomous nodes that collaborate in order to transport information. Due to their limited power supply, energy consumption is a key issue in the design of protocols and algorithms for WSNs. The energy consumption of radio communication is directly proportional to the number of bits of data, that is, data traffic, transmitted within the network [1]. Therefore, using compression to reduce the number of bits to be transmitted has the potential to drastically reduce communication energy costs and so increase network lifetime. Two categories of compression can be distinguished [2] according to their uses, some compression algorithms are designed to support exact reconstruction of the original data after decompression (lossless)[3], in other cases, the reconstructed data is only an approximation of the original (lossy), and contains degradation. Our approach to lossy source coding is based on the low density generator matrix (LDGM) codes.

The trend of communication systems is to promote mobility, cooperation between users, and diffusion. Moreover, the delays in routing data frames in communication networks are often constrained. To address this problem, as well as others (flow, reliability and energy consumption), researchers devised a new approach to information transmission in communication networks, their solution is based on linear combinations of messages from different streams to approximate the theoretical limits of bandwidth usage. Since then, this Network Coding theory [4] has become a discipline in its own right.

II. The Proposed System

As shown in the figure 1, the proposed system is composed of three principle blocks described in section II-A-II-C.

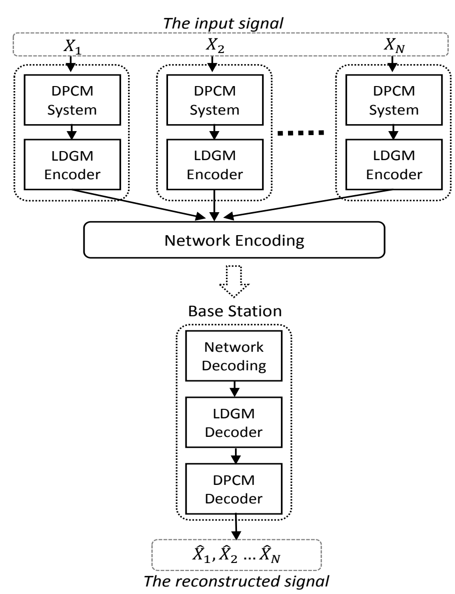

Fig. 1: Structure of the proposed framework

ISSN 2356-5608

A. The Differential pulse code modulation (DPCM) system [2]

Since Linear source codes are known to achieve the entropy rate of memoryless sources [5]. We used the DPCM system. It is mainly a procedure of converting analog to digital signal in which analog signal is sampled and then difference between actual sample value and its predicted value (predicted value is based on previous sample or samples) is quantized and then encoded forming digital value. Thus DPCM code words represent differences between samples. Figure 2 shows the block diagrams for DPCM transmitter.

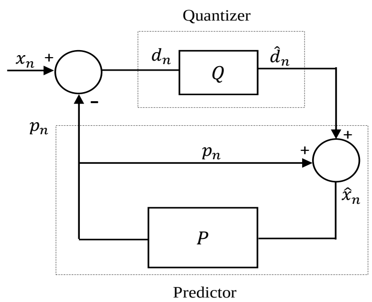

Fig. 2: Differential Predictive Encoder

Where the used symbols have the following meaning:

- x_n : Sampled values of input signal.
- d_n : Prediction error, difference between actual and predicted value $d_n = x_n - p_n$.
- $\hat{d}_n$: Quantized prediction error, $\hat{d}_n = Q[d_n]$
- p_n : predicted value.
- $\hat{x}_n$: Reconstructed value of sampled signal

$$\hat{x}_n = \hat{d}_n + p_n$$

The predicted value is formed using prediction factors and previous samples, usually linear prediction is used, so predicted value can be given as a weighed linear combination of N previous samples using a_i weighting factors :

$$p_n = \sum_{i=1}^{N} a_i \hat{x}_{n-i}$$

.

We choose weighting factors in order to minimize the prediction error, this leads us to minimization of quantization noise (better signal-to-noise ratio).

B. The LDGM Encoder [6]

Given a $Ber\left(\frac{1}{2}\right)$ source, any particular $i.i.d.$ realization $s \in \{0,1\}^n$ is referred to as a source sequence. The goal is to compress source sequences s by mapping them to shorter binary vectors $z(s) \in \{0,1\}^m$, with length m where the quantity $R := \frac{m}{n}$ is the compression ratio. The source decoder then maps the compressed sequence z to a reconstructed source sequence $\hat{s}$. For a given pair $(s, \hat{s})$, the reconstruction fidelity is measured by the Hamming distortion

$$d_H(s, \hat{s}) := \frac{1}{n} \sum_{i=1}^{n} \mid s_i - \hat{s}_i \mid$$

The overall quality of our encoder-decoder pair is measured by the average Hamming distortion $D := E[d_H(s, \hat{s})]$. As illustrated in figure 3, the n source bits are lined up at the left of the graph, and each check, in turn, is connected to (some subset of) the m information bits at the right part of the graph. Note that there is a one-to-one correspondence between source bits and checks. We use $s_1, s_2, ..., s_n$ to refer to elements of C, corresponding to a source bit. Conversely, we use $z_1, z_2, ..., z_m$ to refer to information bits in the set V.

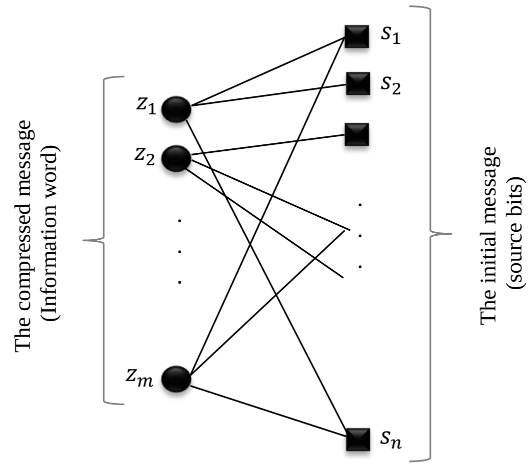

Fig. 3: Graphical representation of an LDGM Code

C. The Network coding processor [7]

The last block component allows recipient nodes to exploit overlapping paths in a multicast tree in order to increase the use of network bandwidth. It lets the intermediate nodes responsible for routing the messages to perform linear operations, so that this optimal throughput is attainable. We model a network consisting of two point-to-point links in tandem as shown in figure 4 below.

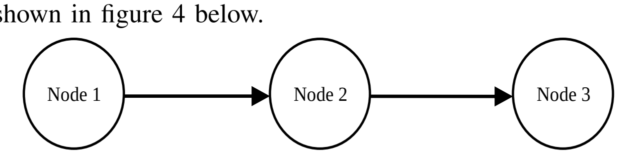

Fig. 4: A two link tandem network

We wish to establish a connection of rate arbitrary close to R packets per unit time from node 1 to node 3. We suppose further that random linear network coding is run for a total time Δ, and that, in this time, a total of N packets is received by node 2. we call these packets $v_1, v_2, ..., v_N$. Any packet u received by a node is a linear combination of $v_1, v_2, ..., v_N$, so we can write $u = \sum_{n=1}^{N} (\beta_n v_n)$ And since v_n is formed by a random linear combination of the message packets $\omega_1, \omega_2, ..., \omega_K$, we have :

$$v_n = \sum_{k=1}^{K} (\alpha_{nk} \omega_k)$$

For $n = 1, 2, , N$.

ISSN 2356-5608

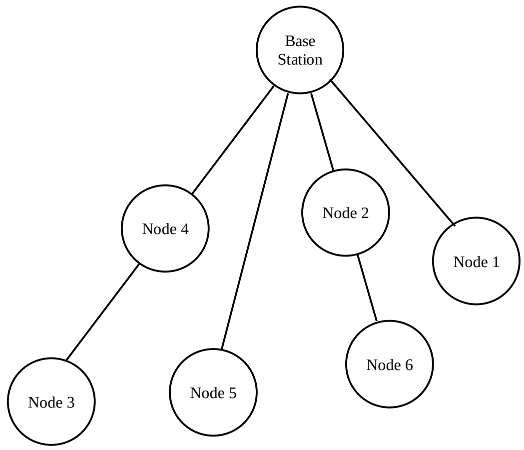

Fig. 5: Model of the structure

The coefficients are generated randomly at the time of sending.

$$u = \sum_{k=1}^{K} \left(\sum_{n=1}^{N} (\beta_n \alpha_{nk}) \right) \omega_k$$

- K: Number of initial packets.
- N: The number of packets present in the node at the transmission time.
- β: Coefficients of the linear combination established in this node.
- α: Global encoding coefficients transmitted by the predecessor node.
- u: The message to be sent.

Concerning the decoding phase, we applied the basic decoding algorithm, described in [8]. This algorithm is a modified version of the Gauss algorithm.

D. Description of the study case

The discussed system is applied to enhance the work of [9], who presented a practical alternative for gathering temperature data using an adhoc wireless sensors network, by minimizing the bandwidth allocated for data transmission. [9] tested their solution on different regions in North Morocco, the Ksar Sghir example has been token subject to our application since it contains the highest number of deployed sensors and the biggest database. Figure 5 describes the network layout, each disc represents a deployed node and the top disc represents the base station. Each sensor collects 6 measurements of temperature each minute, the Packet routing is done using the Ad-hoc On-demand Distance Vector (AODV) protocol [10] and the nodes are interconnected, due to overlapping transmission range.

III. Simulation Results

The database is formed by measurement of 6 sensors, each sensor collects a temperature value every 10 seconds, the measured values of all the sensors vary between $15°C$ and $45°C$, this is converted to a uniformly quantized interval of $-5°C$ to $5°C$ after being processed by the DPCM, which significantly reduces the number of bits required to encode the information. We used an order one predictor at the DPCM system, since the mean squared error of its predictive values remains constant. A bit stream of 900 bits containing 15 minutes of data is then established and subjected to LDGM compression. In order to determine the compression ratio to be applied for each sensor, the entropy of the different nodes is calculated. Figure 6 presents the entropy for each node of our network.

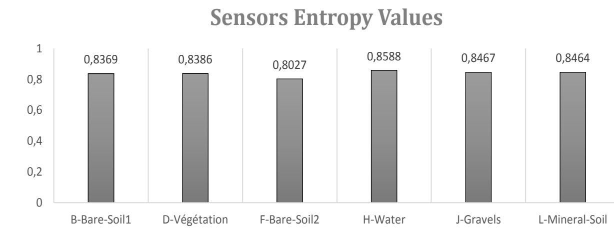

Fig. 6: Entropy value for each sensor

Based on the above results we simulated data transfer using 0.8 and 0.85 as compression ratio. We performed lossy compression using LDGM codes, more precisely a systematic sparse generator matrix, we reconstructed the data using the doping algorithm described on Caire paper [11]. This latter is based on a belief propagation algorithm, it combines a message passing approach with iterative doping of the most significant bits. In our system we used a 50% compression ratio in addition to 30% or 35% doped bits. The mean error for all sensors varies between $0.3°C$ and $0.5°C$. Figures 7-9 compares the reconstructed data and the input data for different samples.

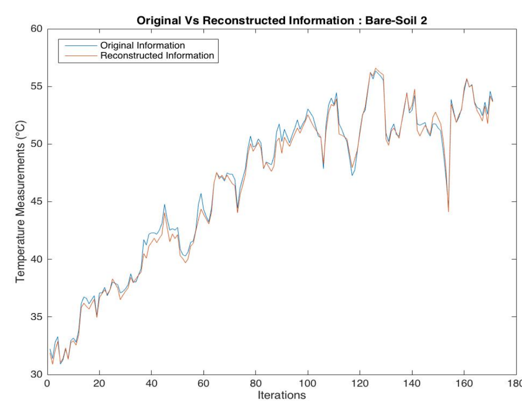

Fig. 7: Input data Vs Reconstructed data for sensor Bare-Soil2, Mean Error=$0.49°C$

All the sensors divide their data into 20 packets then inject them into the network at the same instant of time, each node transmits a linear combination of the received packets along with its own packets, and the decoding is established when there is enough data to reconstruct the initial packets. The decoding curve for all sensors is presented in figure 10.

The system´ s initial performances reveal a delay of 0.6 seconds, and a 33% rate loss of packets transfer. The im-

ISSN 2356-5608

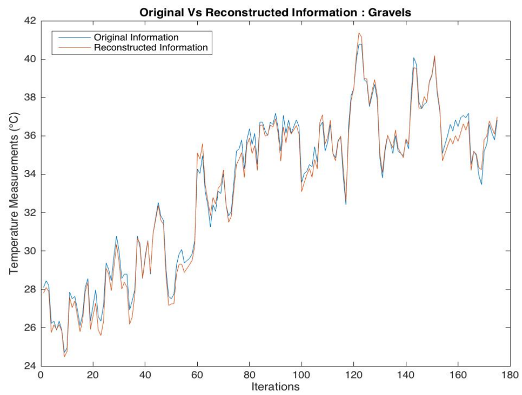

Fig. 8: Input data Vs Reconstructed data for sensor Gravels, Mean Error=$0.33°C$

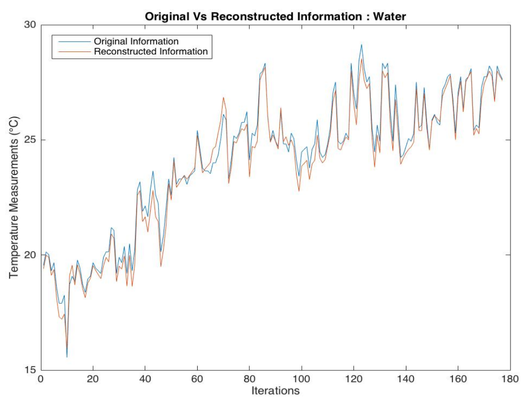

Fig. 9: Input data Vs Reconstructed data for sensor Water, Mean Error=$0.30°C$

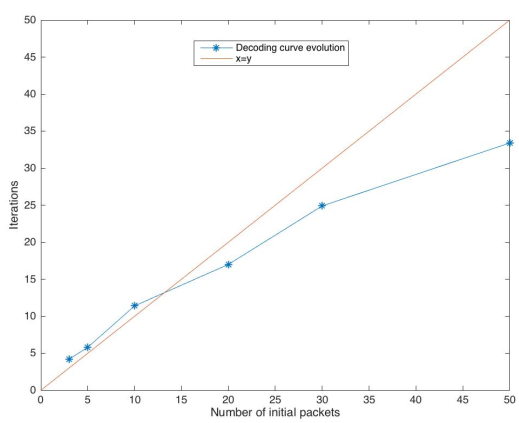

Fig. 10: The decoding curve evolution

plementation of the compression and network coding insured a smaller delay comparing to the classic multi-hop packet routing and clearly a bandwidth alleviation. The obtained results showed significant performance in comparison with some existing compression algorithms [12] and [13], however more enhancement are to be done in order to achieve results which draws near to Shannon bound.

IV. Conclusion

At the source, data compression using predictive encoding and LDGM codes allowed us to optimize the number of bits to be sent on the channel, thus optimizing the bandwidth. The results show that with a compression ratio of 68% over the global framework structure, it is possible to reconstruct the initial data with an average distortion of 0.16. This latter can be translated to $[0.3° \ 0.49°]C$ in the temperature measurement.

While the network coding allowed us to optimize the necessary processing time to route packets from all the nodes to the sink, along with preventing errors in data transfers. Extensions to more developed scenarios are possible in the context of the distributed source coding and advanced relay techniques applied to the system. A channel coding study will be conducted as well in order to enhance the network coding contribution in minimizing the packet loss rate.

Acknowledgment

The authors would like to thank Pr. Naoufal Raissouni and Pr. Asaad Chahboun, for giving us access to their collected database, for their collaboration, and helpful discussions.

References

[1] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, pages 10 pp. vol.2–, January 2000.

[2] Khalid Sayood. Introduction to data compression, December 2005.

[3] Su Chen. *Data Compression in Dynamic Systems*. ProQuest, 2008. Google-Books-ID: 24O8r8bggiIC.

[4] R. Bassoli, H. Marques, J. Rodriguez, K. W. Shum, and R. Tafazolli. Network Coding Theory: A Survey. *IEEE Communications Surveys Tutorials*, 15(4):1950–1978, 2013.

[5] Imre Csiszár and János Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, 1997. Google-Books-ID: c_JEIQAACAAJ.

[6] W. Zhong, H. Lou, and J. Garcia-Frias. LDGM codes for joint source-channel coding of correlated sources. In *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, volume 1, pages I–593–6 vol.1, September 2003.

[7] Tracey Ho and Desmond Lun. *Network Coding: An Introduction*. Cambridge University Press, Cambridge ; New York, 1 edition edition, April 2008.

[8] Janus Heide, Morten V. Pedersen, and Frank H. P. Fitzek. Decoding Algorithms for Random Linear Network Codes. In *NETWORKING 2011 Workshops*, pages 129–136. Springer, Berlin, Heidelberg, May 2011. DOI: 10.1007/978-3-642-23041-7_13.

[9] N. Raissouni, J. A. Sobrino, A. Chahboun, N. Ben Achhab, M. Lahraoua, and A. Azyat. First results towards building up a reliable in situ measurements database for LST algorithm validations using modular WSN: Northern Morocco campaigns case study. *International Journal of Remote Sensing*, 34(9-10):3153–3163, May 2013.

[10] V. Kumar, A. S. Baghel, and P. Mishra. Performance evaluation of DSDV, AODV and LSGR protocol in ad-hoc networks. In *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, pages 4261–4266, March 2016.

[11] Giuseppe Caire, Shlomo Shitz Shamai, and Sergio Verdu. A new data compression algorithm for sources with memory based on error correcting codes. 2003.

[12] P. A. Regalia. A modified belief propagation algorithm for code word quantization. *IEEE Transactions on Communications*, 57(12):3513–3517, December 2009.

[13] Rethnakaran Pulikkoonattu. A source coding scheme using sparse graphs: Modern Coding Theory Course Exam 2008. 2008.

ISSN 2356-5608