A Group Cooperative Coding Model for Dense Wireless Networks

El Miloud Ar-Reyouchi1, Ahmed Lichiou2
Dept. of Telecommunication and Computer Science
Abdelmalek Essaadi University and SNRT
Tetouan, Rabat, Morocco

Salma Ratoual3
EEA&TI Laboratory Faculty of Sciences and Techniques
Hassan 2nd University
Mohammedia, Casablanca, Morocco

Abstract—Generally, node groups in dense wireless networks (WNs) often pose the problem of communication between the central node and the rest of the nodes in a group. Adaptive Network Coded Cooperation (ANCC) for wireless central node networks adapts precisely to an extensive and dense WN, at this level the random linear network coding (RLNC) scheme with Low-Density Parity-Check (LDPC) are used as the essential ANCC evolution. This paper suggests two phases effective technique then studies the influence of the randomly chosen number of coded symbols which are correctly received in the second phase on bit error rate (BER). The proposed technique also focuses on the role of dispersion impact related to the LDPC code generating matrix. The results of the simulation allow selecting the best compromise for the best BER.

Keywords—Dense Wireless networks; Bit Error Rate (BER); Low-Density Parity-Check (LDPC) codes

I. INTRODUCTION

The present paper analyzes a network coding scheme for dense wireless networks as in the case of the wireless sensor networks (WSNs) especially as regards the energy efficiency a cording to [1], [2], [3] or improvement planning at Bit-error rate (BER) as in [4]. This type of network consists of a large set of nodes that cover a relatively small area and whose nodes collect information from the environment and send it periodically to one or more common destinations. In this type of networks it cannot be guaranteed that all nodes have a direct connection with the destination node, so it is very common for the rest of the nodes to collaborate correctly so that the original information reaches its destination by means of retransmissions as in [5]. The use of isotropic antennas allows the nodes to establish neighborhoods, since they are in the coverage radius of other nodes.

The Network Coding (NC) scheme, which has been used as base of the development, is the ANCC (Adaptive Network Coded Cooperation) [6], [7], that adapts precisely to the extensive and the dense WSN and also to the internet of things (IoT) networks [8].

Network Coding [9] is a special case of Cooperative Communication where the topology of a network is exploited to increase its performance [10] and according to [11] it can also improve end to end delays. The network's own nodes are those that encode the transmission, allowing parts of the message that is transmitted to its destination to go in different ways, thus increasing the capacity of the system in multicast networks [9], [12]. Another possibility is that the different ways to reduce the number of errors in transmission by sending several copies of the same messages by different paths to their destination [13].

Employing wireless NC [14], [15] in user cooperation, (ANCC) [16], [17] completely exploits time division techniques (TDMA) for channel code in distributed destinations, which is composed of two phases: sending phase and resending phase. In the sending phase, each node sends its own original data while the others keep silent and decode the packets which have been arrived with the appropriate level. In the resending phase, the central node, randomly chooses several packets to form the resending packets.

In this scheme, it is proposed to combine the network in the form of a graph (network on graph), describing the instant topology of a network, with the codes based on graphs (codes on graph) and which can be adapted to the ad-hoc nature of a wireless network, such as LDPC codes.

The LDPC codes are a category of error correction codes proposed in 1963 [18] by Robert Gallager .These codes are characterized by having a little dense generating matrix G, that is, a binary generating matrix containing a high number of zeros and therefore a low number of ones. Gallager defined the LDPC codes as a CLDPC code (n, j, k) where n is the length of the code words (in English codeword), j the number of ones in per column and k the number of ones per row.

ANCC flexibly produces a selection of LDPC codes [19], [20] in a distributed aspect at the destination. In these codes, all the code words that can be formed from the same generating matrix they find, like in [21], each other at the same distance from each other.

Initially, the LDPC codes were not very successful and fell into oblivion because they required more computing power for their decoding and the capacity of the computers of the time was quite limited. Only authors in [22] such as Tanner have attempted to obtain a valid implementation of Gallager codes. The appearance of turbo-codes in 1993 [23] and advances in computability led the scientific community to rediscover these codes in the mid-1990s [24], [25]. The RNLC and LDPC are used in [26] to more accurately measure the standing wave ratio of an antenna system in a dense WN.

The most typically exploited algorithm for decoding LDPC codes is the Sum-Product algorithm (SPA) [25], or belief
propagation algorithm. SPA is the most basic form to implement the iterative decoding of LDPC codes. This algorithm is an iterative algorithm that estimates the posterior probability of the symbols from the parity check matrix, the received symbols and the likelihoods of the symbols passing through the channel. According to [10], it should be mentioned here that the increase of iterations number has no influence on the system performance.

The main contribution in this paper is the development of a technique using LDPC codes and RLNC in ANCC for reducing the BER value in dense WN, it is distinguishable to analyze how the value of parameter $\beta$ influences the results and what values of this parameter are valid for the nodes group in dense network

The remainder of this paper is structured as follows. The Section II summarizes the mathematical principles of LDPC codes. Section III presents a system model and the description of the proposed technique. Section IV presents simulation results. Section V closes this paper with a conclusion by determining the subject of the future work.

II. MATHEMATICAL PRINCIPLES OF LDPC

The coding of the LDPC codes does not differ from the coding of any block code. Below we summarize the vectors and matrices that intervene in the coding [27] of our approach

1) Vector of information $u$: It is the vector composed of the $N$ bits of information, each one of them comes from each of the nodes that transmit the information:

$$u = [u_1, u_2, ..., u_N]$$

(1)

2) Generator matrix $G$: is a matrix of size $N \times 2N$ that is used to generate the coded words. This matrix is defined by the identity matrix $I_N$ and the matrix $P$, which is responsible for generating the parity bits presented in the coded words. The way to obtain the matrix $P$ depends on the parity equations that we want to use.

$$G = [I_N \ P]$$

(2)

3) Vector of coded information $(c)$ is the vector that results from coding the information vector $u$ with the generating matrix $G$. It has length $2N$ so that the first $N$ bits correspond to the original information bits and the last $N$ bits correspond to linear combinations (defined by the matrix $P$) of the information bits.

$$c = \mu G = [c_1, c_2, ..., c_m, c_{m+1}, c_{m+2}, ..., c_{2m}]$$

(3)

where

$$[c_1, c_2, ..., c_m] = [\mu_1, \mu_2, ..., \mu_m]$$

(4)

and

$$[c_{m+1}, c_{m+2}, ..., c_{2m}] = [\mu_1, \mu_2, ..., \mu_m]P$$

(5)

4) Parity check matrix $H$ is matrix designed for the systematic code generated by $G$ that will have dimensions $N \times 2N$. The $H$ matrix which is used for simulation is constructed according to orthogonality of the matrix $G$ and the matrix $H$ row vectors multiplication. So that:

$$H = [P^T I_N]$$

(6)

$$G.H^T = 0$$

(7)

Where $P^T$ and $H^T$ are the transposed matrices of $P$ and $H$ respectively.

III. SYSTEM MODEL AND THE DESCRIPTION OF THE PROPOSED TECHNIQUE

A. System Model

Assume a monitoring system model where nodes are connected either by wireless as illustrated in Fig. 1. The purpose of this monitoring system model is to provide data, which accurately reflects the critical conditions of renewable energy.

Consider a network consists of $m$ nodes (monitoring labeled $n_1, n_2, n_3, ..., n_m$, and a central node as topology shown in Fig. 1. We assume that each node transmit one native symbol in this phase and we model the network by a directed graph $(N, L)$, where $N$ is the set of nodes $\{u_1, u_2, u_3, ..., u_m\} \in N$ and $L$ is the set of the wireless directed links. The central node, at the center of the figure, wishes exchange symbols denoted $s_1, s_2, s_3, ..., s_m$ coming from all nodes of the network. Suppose now that the nodes are in the communication range.

Fig. 1 represent a directed graph consisted of set $m$ nodes, each node is represented by a circle, with the name of the node inside. The starting node is named by $n_1$.
B. Scenario Description

1) Sending phase: Fig. 1 shows the transmission of nodes $n_1, n_2, n_3, \ldots, n_m$. The node $n_1$ transmits its symbol and both the central node and the rest of the nodes remain silent and listen to the symbol that node $n_1$ has transmitted. The rest of the nodes analyze the received symbol and verify if it is considered correctly received or not. If the node considers that it has been properly received, it proceeds to store the symbol sent by node $n_1$. Otherwise, the symbols must be discarded. In addition, the central node can also stores the symbol $s_1$ sent by node $n_1$.

After the transmission of the node $n_1$, the node $n_2$ begins the transmission of a symbol $s_2$. As explained for node $n_1$, the node $n_2$ transmits its symbol, while the rest of the nodes of the network listen to the channel and store the symbol $s_2$ if they estimate that it has been received correctly.

In this case, the nodes $n_1, n_2, n_3, \ldots, n_m$ (except node $n_2$) store the symbol $s_2$. The central node also stores the symbol $s_2$. These steps are repeated for the transmission of the symbols of the rest of the network nodes. Finally, the central node has received all the symbols $s_1, s_2, s_3, \ldots, s_m$. Consequently, the information vector is composed of the symbols that have transmitted all the nodes $u = (s_1, s_2, s_3, \ldots, s_m)$.

Once all the nodes in the network have transmitted their symbols, the resending phase begins which starts with choosing to combine the symbols received correctly. The idea of embedding the coefficients used in the linear combination, in the packet header, was introduced in [28] by the work of Chou et al. As well as the idea of local opportunistic coding to present a practical implementation of a network coded system was used in [29] for unicast.

Therefore, each node chooses the $\beta$ symbols that have received perfectly in the previous phase. When one of the nodes has fewer than $\beta$ nodes, it would transmit only the nodes that have correctly received, by producing a generator matrix $G$ with a number below the $\beta$ ones in the corresponding column.

The behavior of the resending phase is quite similar to that of the nodes in the sending phase. Each node emits the elaborated symbol (under another form: encoded symbol) while the rest of the symbols remain silent.

The main difference between the two phases is that in this second phase, it is no longer necessary for the nodes to listen to the symbols transmitted by the rest of the nodes. In this second phase only the central node must listen and store the symbols.

2) Resending phase: The process of the second phase is as follows:

First, all the nodes in the network prepare the symbol they are going to send in the resending phase. Subsequently, node $n_1$ sends the corresponding symbol $\sum_{i=1}^{\beta} \oplus s_i$, $\beta \leq m$ together with the headers needed to decode it in the central node where $\oplus$ represents the binary sum, with what the expression $\sum_{i=1}^{\beta} \oplus s_i$ represents the binary sum of the symbols $s_1, s_2, \ldots, s_\beta = s^c$.

The central node receives and stores the received information. These steps are repeated for nodes $n_1, n_2, n_3, \ldots, n_m$, which transmit their corresponding symbols $s_1, s_2, s_3, \ldots, s_m$, respectively.

Finally, the central node has received all the data corresponding to both phases. Thus, after these two phases, and from (4) and (5) we have that the encoded information vector $c$ is:

$$c = [s_1, s_2, s_3, \ldots, s_m, m \times \sum_{i=1}^{\beta} \oplus s_i]$$

From the previous data the central node can build the generator matrix $G$. This requires the matrix that expresses the equations parity $P$. This matrix $P$ is built from the information of the resending phase. The element $s^c_{i,j}$ (coded symbol) has a value 1, if the node corresponding to row $i$ participates in the binary sum of the response what issues the node that corresponds to column $i$, it is 0 if not involved. So the resulting matrix $P$ is:

$$P = \begin{bmatrix}
 n_1 & n_2 & \cdots & n_m \\
 s^c_{1,1} & s^c_{1,2} & \cdots & s^c_{1,m} \\
 \vdots & \vdots & \ddots & \vdots \\
 s^c_{m,1} & s^c_{m,2} & \cdots & s^c_{m,m}
\end{bmatrix}$$

where on the top, it has been tagged the node which correspond to each column and the on right side, the node corresponds to each row.

According to (2) and the matrix $P$, we can build the matrix $G$ responsible for generating this code, which is as follows:
which leads to the next matrix:

\[
G = \begin{bmatrix}
1 & 0 & \ldots & 0 & s_{1,1}^c & s_{1,2}^c & \ldots & s_{1,m}^c \\
0 & 1 & \ldots & 0 & s_{2,1}^c & s_{2,2}^c & \ldots & s_{2,m}^c \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1 & s_{m,1}^c & s_{m,2}^c & \ldots & s_{m,m}^c
\end{bmatrix}
\]

Equation (2) defines that the generating matrix \( G \) can be divided into two halves. The first half corresponds to an \( I_m \) identity matrix of size \( m \), with ones on the principal diagonal and zeros elsewhere. This identity matrix represents the sending phase, in which each node of the network transmits its data, and this is represented by a 1 for each element \( s_{i,i} \) where \( i \in \{1, m\} \). The second half of the matrix \( G \) corresponds to the matrix \( P \), which defines the parity equations.

In this way, it is appreciated that in the generating matrix \( G \), each node has two columns, the column \( i \) for the transmission in the sending phase and the column \( m+1 \) for the transmission in the resending phase, which corresponds to a distributed coding.

As shown above, where it is labeled that the first and the \( n_{m+1} \)-th column correspond to the two times transmitted by node by node \( n_1 \), one in the sending phase and another in the resending phase. In addition, from the matrix \( P \) the central node can construct the parity check matrix \( H \), as indicated in (6):

\[
H = \begin{bmatrix}
s_{1,1}^c & s_{1,2}^c & \ldots & s_{1,m}^c & 1 & 0 & \ldots & 0 \\
s_{2,1}^c & s_{2,2}^c & \ldots & s_{2,m}^c & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
s_{m,1}^c & s_{m,2}^c & \ldots & s_{m,m}^c & 0 & 0 & \ldots & 1
\end{bmatrix}
\]

The bipartite graph regarding a regular parity check matrix \( H \) will be used by the SPA to acquire the transmitted information vector.

### C. Influence of Parameter \( \beta \)

In order to obtain the best BER, it is necessary to analyze how the value of the parameter \( \beta \) influences the results of the proposed technique and what values of this parameter are valid for this study.

Recall that parameter \( \beta \) corresponds to the number of symbols from the nodes of the group that intervene in the response of each of the nodes during the resending phase. A higher value of this parameter implies that the transmitted symbol includes information about a greater number of nodes, but a very high number may imply that the generator matrix stop being very scattered, increasing the probability of short cycles appearing. This causes the LDPC codes to stop having good properties.

We define the degree of dispersion \( D \) as the relationship between the symbols are combined in the resending phase and the number of nodes that can participate in the combination. It refers to the dispersion of the generator matrix \( G \) of the LDPC code.

\[
D = \frac{\beta}{m - 1}
\]

Before giving the results, we can present the useful parameters values in an attempt to improve our results. The parameters used in our simulation are shown in Table I.

### TABLE I. SIMULATION PARAMETERS OF THE FIRST SENDING AND RESENDING PHASES

<table>
<thead>
<tr>
<th>Designation</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node total number</td>
<td>50/100/200/300/400/500</td>
</tr>
<tr>
<td>Parameter ( \beta )</td>
<td>4/6/8/10/12/14/16</td>
</tr>
<tr>
<td>Iteration number</td>
<td>10</td>
</tr>
<tr>
<td>Power</td>
<td>50 mw</td>
</tr>
<tr>
<td>Node Placement</td>
<td>Uniform</td>
</tr>
</tbody>
</table>

### IV. RESULTS AND DISCUSSIONS

The main objective of this section is to obtain the parameter values \( \beta \) for the better behavior of the proposed method.

#### A. Simulation Results

Fig. 2 represents the different curves obtained for a group of 50, 100, 200, 300, 400 or 500 nodes, varying the value of the parameter \( \beta \) and using 10 iterations of the Sum-Product decoding algorithm. The dispersion percentage \([10]\) refers to the dispersion having the generator matrix of the LDPC code.

According to Fig. 2, we note that the increase of the nodes as well as the increase of the parameter causes an increase of the percentage of dispersion. The dispersion percentage presented by the curve of 50 nodes is too high and could be...
associated more with dispersed matrices than with very dispersed matrices, so it does not seem a good option to use groupings of this size. We are looking for a compromise between the numbers of the nodes and the value of the parameter for a good value of BER.

Fig. 3 represents the different curves obtained for a different value of parameter $\beta$ which are 4, 6, 8, 10, 12, 14 or 16, varying the value of nodes using 10 iterations of the Sum-Product decoding algorithm.

As shown in Fig. 2 and Fig. 3, parameters $\beta$ have a crucial role for the dispersion of the elements of the matrix generator.

Fig. 4 shows the different curves obtained for a network containing 100 nodes varying the value of the parameter $\beta$.

By analyzing the graphs using 100 nodes, we can see that the curves, with greater compromise, are given for the parameter values = 6, 7 and 8, whose matrices offer the dispersion percentage that are displayed in Fig. 2 and Fig. 3.

Fig. 5 shows the curves obtained in simulations of 500 nodes, using a maximum of 10 iterations of the decoding algorithm Sum-Product. The parameter $\beta$ is varied between 8 and 13 and the BER simulation results versus SNR are represented in Fig. 5 below.

With 500 nodes, the generator matrix of the LDPC code has a size of 500x1000. In addition, the parameter $\beta$ is varied between 8 and 13 and the BER simulation results versus SNR are shown in the Fig. 5. By observing the resulting graphics to these groups, it can be seen that the curves, that offer a better compromise, are given for the parameter values $\beta$ comprised between 9 and 12.

For the nodes 200, 300 and 400 nodes we have followed the same procedure as for the nodes 100 and 500 node as shown in Fig. 2 and Fig. 3.

The curves represented in Fig. 6 are the curves that offer better compromise for studied system. By analyzing the curves obtained in Fig. 6, we see that the reliability parameter increases when increasing the nodes number.
The curves shown in Fig. 6 are the curves that offer the best compromise for each group. In this way, a value of parameter $\beta = 6$ has been selected for a group of 100 nodes. The parameter $\beta = 8$ has been chosen for group of 200 nodes. For group of 450 nodes a value of the parameter is $\beta = 10$ and for group of 500 nodes a value of the parameter is $\beta = 12$.

B. Discussion, Limitations and Future Work

This paper has presented bases of the ANCC scheme using Network Coding with extensive LDPC codes in dense wireless networks based on an efficient technique. The chosen values of the parameter $\beta$, obtained for different scenarios with the proposed technique, are a very good compromise to obtain an optimal BER. The number of the selected nodes to a group has an inseparable relationship with the parameter $\beta$ when applying network coding. The proposed technique results improve the BER performance while maintaining error-free communication with negligible chance of interference.

All the nodes that appeared in the group reached to the central node in a single hop. This scheme facilitated and simplified enough the analysis. But, despite efforts to increase BER in dense wireless networks, this study has a notable limitation crossing several nodes. Consequently, this scheme is difficult to be represented in a real environment because several hops can appear to reach the central node. This depends, among other factors, of the number of nodes, the distance between them and the used technology. The future work is to develop congestion avoidance algorithm improving the BER in central node through several nodes for dense wireless network.

V. Conclusions

This study investigates the improvement of BER by applying an efficient technique of communication between a central node and a node groups using network coding with LDPC. It focuses most notably on dispersion percentage and the number of the coded symbols correctly received in the resending phase. The authors have been examined and performed different scenarios by selecting the best BER curves versus SNR, with a varying number of nodes and parameters $\beta$ in the network. The results clearly concludes that the choice of parameters $\beta$ with their corresponding BER nodes can improve considerably and significantly the BER versus SNR.

REFERENCES