Reliability and Efficiency Analysis of Distributed Source Coding in Wireless Sensor Networks

C. Fischione, S. Tennina, F. Santucci, F. Graziosi

Abstract—We propose a comprehensive theoretical framework to evaluate reliability and energy consumption of distributed source coding (DSC) in wireless sensor networks (WSNs) applications. Energy efficiency and the amount of measurements that can be successfully decoded in tree-based WSNs employing DSC in the presence of different coding topologies and packet aggregation schemes (PA) are accurately characterized. The system model includes a realistic network architecture with multi-hop communication, automatic repeat request protocol (ARQ), packet losses due to channel impairments and collisions, and correlation properties of the sensed phenomena. Four DSC topologies and three alternatives of PA are considered. The analysis is carried out by evaluating the expressions of reliability of DSC in terms of probability of measurements that cannot be decoded (loss factor), and the efficiency in terms of average energy consumption of the network. Numerical results show that the best choice of DSC topology and packet aggregation depends highly on the network parameters and source characteristics. Therefore, the analysis developed in this paper can be used as an effective mean to optimize network operations.

Keywords: Wireless Sensor Networks (WSNs), Distributed Source Coding (DSC), Packet Aggregation (PA).

I. INTRODUCTION

Distributed Source Coding (DSC) exploits the richness of information usually provided by spatial correlation of measurements taken by the nodes of a Wireless Sensor Network (WSN), to compress data without loss of information. Although DSC has recently found in WSNs a relevant application domain, the theoretical foundations date back to the pioneering work by Slepian and Wolf [1]. A practical code construction has been proposed in [2]. In [3] an algorithm to achieve DSC over WSNs using a single codebook with variable compression rate has been proposed. In [4] it has been pointed out that performance of DSC is largely influenced by the DSC topology, i.e., the nodes’ position with respect to the spatial correlation profile of the sensed phenomena. The problem of optimizing the DSC topology has been investigated in [5], [6] for tree-like networks. The interaction between DSC and routing has been studied in [7]–[9]. Routing with packet aggregation allow for decreasing protocol overhead, with reduction of energy consumption, as it was noted [7]. The problem of finding the optimal packet aggregation alternative for WSNs with DSC has been dealt in [10].

These contributions assume that DSC acts above the typical protocol stack of a sensor network platform. However, many other mechanisms for WSNs (e.g., routing and MAC protocols) are targeted to the same goal of achieving better energy usage. As it was observed in [11], DSC interactions with lower layers of the protocol stack (network, data link, and physical layer) deserve particular attention to fully exploit the claimed benefits. For instance, header overheads that appear in any protocol data unit (PDU) format should be taken into account. A node may aggregate packets coming from other nodes with its own data in order to reduce the impact of packet overhead in multi-hop routing. It can be argued that DSC has a relevant impact on network, data link, and physical layers, and vice versa. Although the above mentioned relevant contributions address fundamental topics of DSC, they neglect some important aspects that relate to network design and operations in realistic scenarios, namely: packet losses, packet aggregation and fragmentation, overhead reduction, and cross-layer interaction. We believe that a crucial concern is represented by a proper (joint) combination of DSC topologies and packet aggregation (PA) in realistic multi-hop communication scenarios.

In this paper we investigate the interplay of main communication system parameters that affect DSC performance. We develop a complete framework for a joint analysis of the reliability (loss factor) and overall energy consumption. Our contribution is related to [4] because we include the four alternatives of coding topology therein proposed. It is also related to [7], because we consider three PA techniques. However, we simultaneously consider DSC in a system scenario including an accurate model of the physical layer, a data link layer, packet aggregation, multi-hop routing, and the correlation pattern of the sensed phenomena.

II. SYSTEM MODEL

Consider an area wherein $N$ nodes are deployed: each of them takes some measurements and attempts to transmit reports towards a sink. A tree-like network topology is considered. Each node belongs to a generic level $i$ (with $i = 1, \ldots, L$), $i$ being the distance of a node from the sink in terms of number of hops (see Fig. 1). Let denote by $i = 0$ the sink node, which is placed at the level 0. Tree-based architecture is a relevant network topology, as pointed out in numerous contributions (see, for instance, [4] [6], [12]), and for the IEEE 802.15.4 standard [13, pag. 15–16]. Denote with $l_i$ the number of nodes belonging to level $i$. We set $l_0 = 1$, which is the number of nodes of level 0 (only the sink). The equality $\sum_{i=1}^{L} l_i = N$ holds. Each node is connected to a parent node and takes $m$ measurements of $m$ different physical phenomena per sensing interval. Assume that the measure-
ments of different phenomena are continuous i.i.d. random variables, and that the measurements of a given phenomenon by different sensors are spatially correlated [3], [6]. Let \( Y_{i,j} \), for \( i = 1, \ldots, L \), and for \( j = 1, \ldots, L_i \), be the quantized measurement taken from the generic node \( j \) of level \( i \) at a given time instant for one of the \( n \) phenomena. Measurements are quantized by an analog-to-digital converter having the same resolution in each sensor. A set of variables \( Y_{i,j} \)'s at a given time instant are generically spatially correlated, the correlation pattern being constant in time under the assumption of stationarity [4].

Each node performs the following tasks: it encodes the measurements according to the DSC algorithm proposed in [3], receives PDUs from the child nodes, aggregates them, and then transmits them towards its parent node. Let \( R(Y_{i,j}) \) be the coding rate of the DSC technique [3] of the considered source.

At the physical layer we consider a protocol data unit (PDU) that we name packet. As a consequence of operations at the various layers of the protocol stack, it is composed as follows: a frame length indicator plus a preamble and a synchronization symbol of \( P \) bits; a data link header of \( O \) bits, a network header of \( Q \) bits; a payload of \( s_i \) bits incorporating coded measurements taken by nodes and \( ID \) bits for node identification; \( CRC \) bits for the forward error detection [14]. The payload may have variable size as a consequence of the DSC topology and aggregation mechanism, as it will be clear in the following subsections. The packet structure is adherent to specifications by the IEEE 802.15.4 communication standard [13] for the data link layer and physical layer of the protocol stack, while we refer to Tmote Sky sensors [15] for the network layer.

A. DSC Topology and Packet Aggregation

Three prominent DSC coding schemes are here studied and their performance compared to the benchmark case denoted by no DSC (NODSC): Sequential DSC (SEQ), Clustered DSC (CL), and DSC Master Slave (MS). These schemes were first proposed in [4]. In the following, we briefly summarize their characteristics.

In the NODSC topology, each sensor encodes measurements independently from other nodes. No distributed source coding is adopted.

In the SEQ scheme (see Fig. 2), node \( (1,1) \) performs the encoding of \( Y_{1,1} = R(Y_{1,1}) \) bits; node \( (1,2) \) encodes \( Y_{1,2} \) with \( R_{1,2} = R(Y_{1,2}|Y_{1,1}) \) bits, provided that the decoder knows \( Y_{1,1} \); in general, node \( (i,j) \) encodes \( Y_{i,j} \) using the following number of bits \( R_{i,j} = R(Y_{i,j}|Y_{1,1}, \ldots, Y_{i-1,1}, Y_{1,2}, \ldots, Y_{i-1,2}, \ldots, Y_{1,L}, \ldots, Y_{i-1,L}) \), provided that the decoder knows \( Y_{1,1}, \ldots, Y_{1,l}, Y_{2,1}, \ldots, Y_{2,2}, \ldots, Y_{i-1,1}, \ldots, Y_{i-1,L} \).

In the CL topology (see Fig. 3), nodes are grouped in \( K = L_1 \) clusters, where each cluster consists of a sub-tree having a node of level 1 as a root. Each cluster includes a number of nodes, which is denoted by \( N_k \). For each node of a cluster, a SEQ coding topology is adopted independently from other clusters and starting from the root node \( (1,k) \), for \( k = 1, \ldots, L_1 \), of cluster \( k \).

In the MS topology, nodes are grouped in clusters as in the case of CL, however each root \( (1,k) \), for \( k = 1, \ldots, L_1 \), of a sub-tree is elected as master node. Furthermore, in contrast to CL, for each node of a cluster \( k \), the DSC is performed only with respect to the master \((1,k)\) of the cluster the node belongs to.

We consider three alternatives for aggregation: the classical multi-hop (CMH), where nodes relay packets without any aggregation; the aggregated multi-hop (AMH), where nodes collect received packets, aggregate them at the MAC layer into a single frame and relay it; the Fragmented Aggregated Multi-hop (FAMH), which differs from AMH as it allows us to do packet fragmentation: when an aggregated packet reaches a size that exceeds a maximum threshold, it is fragmented in multiple packets.

B. Performance Indexes

Two performance indexes are considered in this paper, namely: the loss factor and the energy consumption of the network, which we denote with \( \Lambda \) and \( E_N \), respectively. The loss factor is defined as the fraction of measurements generated by the network that cannot be reconstructed at the sink. The energy consumption of the network is instead expressed as the average number of bits (per reporting interval) transmitted.\(^1\)

The characterization of both \( \Lambda \) and \( E_N \), which is the core contribution of this paper, is founded on the packet loss probability, as we see next.

III. PACKET LOSS PROBABILITY

Consider a wireless channel which exhibits non-selective fading behavior both in frequency and in time, and a QPSK modulation.\(^1\)

\(^1\)Note that the performance indexes do not include the packet delay caused by the coding topology. By the assumptions that the correlation properties of the sensed phenomena are stationary, that the number of measurements per sampling rate is the typical one provided by sensors off-the-shelf [15] (i.e. less than ten), and that the number of levels is not large, then the delays are negligible.
probability is communication standard [13]. Therefore, the average bit error loss probability of a packet generated at level \( i \) is given by (III.3) single-hop packet loss probability. When an average packet loss probability at level \( i \) is defined as the average value of the Signal-to-Noise Ratio (SNR) computed at distance \( d \) from the transmitting node. The packet loss probability over a multi-hop path from node \( i \) to the sink can be computed with (III.4) and (III.7).

\[
\Psi(s_i) = \frac{1}{d_{max} - d_{min}} \int_{d_{min}}^{d_{max}} \Psi(\rho, s_i) d\rho ,
\]

where \( d_{min} \) and \( d_{max} \) are, respectively, the minimum and maximum distances between any pair of child-parent nodes. We call (III.3) single-hop packet loss probability. When an ARQ protocol is implemented at level \( j \), with \( M_j \) maximum number of retransmissions, we can easily express the packet loss probability of a packet generated at level \( i \) across one hop as follows:

\[
\tilde{\Psi}(s_i, M_j) = \Psi(s_i)^{M_j} .
\]

The packet loss probability over a multi-hop path from node \( i \) up to the sink can be computed with

\[
\tilde{\Psi}(s_i) = 1 - \prod_{j=1}^{i} [1 - \tilde{\Psi}(s_i, M_j)] .
\]

**Remark 3.1:** Observe that (III.4) depends on the coding topology and aggregation scheme through the payload size \( s_i \), which will be characterized in the next subsections for the four alternatives of coding topology and the three schemes of aggregation.

For the sake of the performance analysis, we define the average coding rate of the level \( i \) as \( R_i = \frac{1}{n} \sum_{l=1}^{n} R_{i,l} \). We define the average connectivity of a node belonging to level \( i \) as the average number of children of that node, namely \( C_i = l_{i+1}/l_i, \ i = 0, \ldots, L - 1 \), where we impose \( C_L = 0 \).

### A. CMH

The contribution to the average payload size for the CMH scheme at level \( i \), as given from measurements generated by a node of level \( j \), is

\[
s_i = ID + m \cdot R_j , \quad i = 1, \ldots, L \text{ and } j \geq i ,
\]

Therefore, the single-hop packet loss probability is given by (III.4), where \( s_i \) is given by (III.6).

### B. AMH

Consider the AMH scheme. Then, the average payload size at the \( i \)th level is [16]

\[
s_i = \begin{cases} 
1 + \sum_{j=1}^{L-i} \prod_{k=1}^{j} n_{i+k} & i = 1, \ldots, L - 1 , \\
ID + m \cdot R_L & i = L .
\end{cases}
\]

The packet loss probability over a single-hop is

\[
\tilde{\Psi}(s_i) = 1 - [1 - \tilde{\Psi}(s_i, M_i)]^{C_{i-1}} , \text{ for } i = 2, \ldots, L ,
\]

Finally, expression (III.9) can be used in (III.4) to obtain the single-hop packet loss probability.

### IV. Loss Factor

Here we build on the packet loss probability analysis of previous section to derive the loss factor. First, let us introduce some definitions. Denote with \( D_i \) the average number of descendants of the generic node of level \( i \) whose measurements are successfully received at that node. Denote with \( D_{i,j} \) the average number of descendants at level \( j \) whose measurements are successfully received at the generic node of level \( i \). We also impose \( D_{1,i} = 1 \) and \( D_{i,j} = 0 \) if \( i > j \). Then, at level \( i \) of the network, the average number of descendants is [16]

\[
D_i = \sum_{j=i+1}^{L} D_{i,j} , \quad i = 0, \ldots, (L - 1) .
\]
where
\[ D_{i,j} = \prod_{k=i+1}^{j} [1 - \tilde{\Psi}(s_j, M_k)] \cdot C_{k-1}. \] (IV.2)

In the next subsections, we characterize the loss factor for each case of coding topology and aggregation mechanism.

A. NODSC

In the CMH scenario, the probability that a packet generated at level \( i \) does not reach the sink is given by (III.5). Hence, after averaging over all levels and number of nodes, the loss factor is
\[ \Lambda_{\text{NODSC}}^{\text{CMH}} = \sum_{i=1}^{L} \left\{ \prod_{j=1}^{i} [1 - \tilde{\Psi}(s_i, M_j)] \right\} \cdot \frac{l_i}{N}. \] (IV.3)

When aggregation procedures are adopted, derivation of the loss factor has to take into account that loosing a packet generated from a node of level \( i \) determines the avalanche effect of losing measurements successfully received by that node and coming from lower levels of the network. A packet loss at level \( i \) causes an average loss of \( 1 + D_i \) measurements. Therefore, the loss factor for the AMH (b) and FAMH (c) cases is
\[ \Lambda_{\text{NODSC}}^{(b,c)} = \sum_{i=1}^{L} \left\{ \prod_{j=1}^{i} [1 - \tilde{\Psi}(s_i, M_j)] \right\} \cdot (1 + D_i) \cdot \frac{l_i}{N}, \] (IV.4)

where the computation of \( \tilde{\Psi}(s_i, M_j) \), which appears also in the \( D_i \), is performed by using (III.3), (III.4), and (III.7) for AMH (b), and using (III.4), (III.8), and (III.9) for FAMH (c), respectively.

B. SEQ

Let us first consider the CMH case. The loss factor of the SEQ topology in the CMH aggregation is [16]
\[ \Lambda_{\text{SEQ}}^{\text{CMH}} = \frac{1}{N} \sum_{i=1}^{L} A_{\text{SEQ},i} \cdot \sum_{j=1}^{l_i} B_{\text{SEQ},i,j} \cdot C_{\text{SEQ},i,j}, \] (IV.5)

where
\[ A_{\text{SEQ},i} = \left\{ 1 - \prod_{m=1}^{i} [1 - \tilde{\Psi}(s_i, M_m)] \right\}, \]
\[ B_{\text{SEQ},i,j} = N + 1 - j - \sum_{m=1}^{i-1} l_m, \]
\[ C_{\text{SEQ},i,j} = \prod_{k=1}^{i-1} \prod_{m=1}^{k} \left\{ 1 - \tilde{\Psi}(s_k, M_m) \right\}^{l_k}, \]
\[ \cdot \prod_{m=1}^{i} \left[ 1 - \tilde{\Psi}(s_i, M_m) \right]^{j-1}. \]

and where \( \tilde{\Psi}(s_i, M_k) \) is computed with (III.4).

In the cases of AMH and FAMH, expressions for the loss factors are the same of (IV.5). In fact, the loss of a packet coming from node \((i, j)\) induces the loss of all measurements aggregated from lower layers of the network until \((i, j)\): its average value is given by \( D_i \) in (IV.1). However, this number of losses is already included in the number of measurements that, being coded with respect to \( Y_{i,j} \), cannot be reconstructed. From Fig. 1, one sees that this metric is \( N + 1 - j - \sum_{k=1}^{i-1} l_k \). Therefore, we can readily express the loss factor for the AMH and FAMH cases using (IV.5).

where the probability \( \tilde{\Psi}(s_j, M_j) \) has to be computed using (III.3), (III.4), and (III.7) for the AMH and using (III.4), (III.8), and (III.9) for the FAMH.

C. CL

When the network is partitioned into clusters, for each cluster the loss factor can be computed as in the SEQ case. Hence, the loss factor for CMH (a), AMH (b), and FAMH (c) is
\[ \Lambda_{\text{CL}}^{(a,b,c)} = \frac{1}{K} \sum_{k=1}^{K} A_{\text{SEQ},k}, \] (IV.6)

where the subscript \( k \) denotes the loss factor of cluster number \( k \), and it is computed using (IV.5). When computing \( \Lambda_{\text{SEQ},k}^{(a,b,c)} \), \( N \) and \( l_i \) must be replaced, whenever they appear in (IV.5), with \( N_k \) and \( l_i^{(k)} \) respectively, where \( l_i^{(k)} \) is defined as the number of nodes of level \( i \) of cluster \( k \). Notice that \( l_0^{(k)} = 1 \) is the number of nodes of level 0 (only the master node) of cluster \( k \).

D. MS

The analysis is similar to the CL, the only difference being in the fact that there is only one node with respect to which the coding is performed (see [16] for details).

V. AVERAGE ENERGY CONSUMPTION

Expressing the average number of bits transmitted by the network requires characterization of the average number of transmitted packets at the generic level \( i \). Recall that such a number depends on the coding topology by the payload size. In the following, we derive the energy consumption for the CMH, AMH and FAMH schemes.

The number of retransmissions of a packet sent from a node of level \( i \) over the hop between level \( i \) and \( i - 1 \) is denoted as \( \Theta_i(s_i) \). It is expressed as [14]
\[ \Theta_i(s_i) = 1 + M_i \cdot \tilde{\Psi}(s_i) M_{i+1} - (M_i + 1) \cdot \tilde{\Psi}(s_i) M_i. \]

Therefore, the average number of transmitted bits from the generic node of level \( i \), when an ARQ protocol is employed, is
\[ B_i = \sum_{j=1}^{L} (P + O + Q + s_j + CRC) \cdot D_{i,j} \cdot \Theta_j(s_j), \]
where \( D_{i,j} \) has been defined in (IV.2).

Consider the CMH case. The average number of transmitted bits is
\[ E_N^{\text{CMH}} = \sum_{i=1}^{L} B_i \cdot l_i, \] (V.1)

where \( \Theta_i(s_i) \) is computed using (III.3).

Let us consider now the AMH and FAMH cases. Each node of level \( i \) transmits only one packet containing both the measurements taken from local sensing and the measurements coming from lower levels of the network and encapsulated therein. Therefore, \( E_N \) is readily expressed as follows:
\[ E_N^{\text{AMH}} = \sum_{i=1}^{L} (P + O + Q + s_i + CRC) \cdot \Theta_i(s_i) \cdot l_i, \] (V.2)
where $\Theta_i(s_i)$ is computed using (III.3) and (III.7). Indeed, in the FAMH case, the average number of transmitted bits is

$$E_N^{FAMH} = \sum_{i=1}^{L} (P + O + \tilde{s}_i + CRC) \cdot F_i \cdot \Theta_i(s_i) \cdot l_i, \quad (V.3)$$

where $\Theta_i(s_i)$ is computed using the payload size $\tilde{s}_i$ for the FAMH case, with (III.8) and (III.9).

VI. Numerical Examples

In this section we report numerical results obtained by a Matlab implementation and simulation of the analytical framework developed in the previous sections.

A. Simulation Parameters

The parameters setting adopted for simulations is representative of the Tmote Sky sensors [15] and the communication standard IEEE 802.15.4 [13], as introduced in the sequel. We consider a network deployed in a square area having facet of 11m with the sink located in the middle and $N = 64$ nodes distributed in $L = 4$ levels, where each node is randomly located within a sub-square of 1.2m. The number of nodes for each level and average connectivity are $[l_0,\ldots,l_L] = [1,4,12,20,28]$ and $[C_0,\ldots,C_L] = [4,3,20/12,28/20,0]$, respectively. Each node senses $m = 8$ measurements per sensing event. The packet frame format is as follows: $P = 48$, $O = 184$, $Q = 56$, $ID = 32$, $CRC = 16$, and $s_{\text{max}} = 760$, where the unit is intended in bit. Transmission power has been set to $-10$ dBm. The noise floor is $P_n = -120$ dBm [15]. The collision probability has been set to $\phi_i = 10^{-5}$ (this corresponds to a packet collision probability of about 0.005, which is quite large if related to the sampling rate, the packet duration, and the number of nodes). Smaller values for the collision probability basically induce similar effects. Without loss of generality, and coherently with [13], we set the maximum retransmission iterations of the ARQ protocol for levels 1, $\ldots$, $L = 4$, to 3, 3, 1, 1, respectively. This choice is motivated by the fact that nodes closer to the sink are subject to larger packet losses as a consequence of larger traffic load.

Measurements are characterized with a $N$ dimensional multi-variate normal distribution having average $\mu$ and co-variance matrix $K = [K_{kl}]$, where $K_{kk} = \sigma^2$ and $K_{kl} = \sigma^2 e^{-r d_{kl}}$ for $k \neq l$, where $r$ is the spatial correlation decay parameter, and $d_{kl}$ is the distance between nodes $k$ and $l$. The correlation decay parameter is defined such that it tends to 0 for highly correlated sources, whereas tends to 1 for low correlated sources [6]. We used the upper bound for the DSC coding rate $R(\cdot)$ we derived in [16].

B. Simulation Results

We consider two representative scenarios: (i) large spatial correlations among measurements ($r = 10^{-3}$) and (ii) low correlation patterns ($r = 10^{-1}$).

In Fig. 4, the loss factor is plotted for various cases of the coding topology and aggregation scheme as obtained in Section IV, for a highly correlated measurements scenario. We observe that ARQ decreases significantly the amount of measurements that cannot be decoded at the sink node. Among DSC topologies, the SEQ case shows largest values of the loss factor. The same effect appears also in the CL topology, but clustering reduces the amount of performance loss. A relevant conclusion that can be drawn from Fig. 4 is that MS seems to be the most appealing coding topology in terms of reliability, and that aggregation mechanisms exhibit similar performance within the same coding topology.

In Fig. 5, values of the energy consumption $E_N$ for the various DSC topologies and aggregation schemes are plotted as obtained for highly correlated sources, as presented in Section V. Although ARQ introduces a larger number of packets per node to be relayed, we can clearly observe that it determines a small energy rise. In terms of coding schemes, the NODSC topology exhibits the worst performance, since it uses no compression at all, whereas SEQ, CL and MS topologies do not show large differences. This is mainly due to the large spatial correlation, so that measurements from a few nodes are enough to achieve good compression rates.

For a better understanding of energy consumption, define the following coding efficiency metric:

$$\eta_{CT} = 1 - \frac{E_{N,CT}}{E_{N,NODSC}}, \quad (VI.1)$$

where the subscript CT denotes one of the coding topologies SEQ, CL, or MS. Large values of $\eta_{CT}$ mean that the corresponding coding topology exhibits reduced energy consumptions when compared to the NODSC case.

To compare both efficiency and reliability, in Fig. 6 the coding efficiency is reported as a function of the loss factor.

2Recall that losing a measurement determines the avalanche effect of losing all other nodes related measurements.
for the case of highly correlated sources. It turns out that the SEQ alternative often offers the largest coding efficiency, but also induces larger loss factors. On the contrary, the MS case exhibits poor energy savings, but good loss factor. The CL topology presents fair joint performance. Finally, it is confirmed that aggregation mechanisms significantly impact the energy consumption.

In Fig. 7, we reported the coding efficiency as a function of the loss factor, for low correlated measurements. Although the maximum achievable efficiency is smaller than that obtained for highly correlated measurements, the remarks done for Fig. 6 still hold. However, a drop of efficiency of the AMH and FAMH mechanisms is evident, particularly in the MS case when compared to the CL CMH.

VII. CONCLUSIONS AND FUTURE PERSPECTIVES

A comprehensive theoretical framework to evaluate the loss factor and energy efficiency of distributed source coding in the presence of four alternatives of coding topology (no DSC, sequential, clustered and master slave) and three alternatives of packet aggregation mechanisms (classic multi-hop, aggregated multi-hop and aggregated and fragmented multi-hop) was proposed.

Given a network topology and correlation pattern of the sensed phenomena, and considering the application requirements (reliability and energy consumption), our analysis can be used for an efficient network deployment. Since the analytical results capture the trade-off between the network parameters, they can be employed effectively by network designers to provide a global optimization of the DSC operations. Therefore, our analysis is a major tool to guarantee a long lifetime of the network under reliable communications.

We plan to extend the analysis developed in this paper to a more general topology than tree-like network. The goal is the joint optimization of routing, aggregation scheme, and source coding with respect to the correlation pattern of the distributed source. We believe that non-linear mixed integer-real optimization theory will be useful for the solution of this challenging problem.

REFERENCES