DISTRIBUTED SOURCE CODING AND DISPERSIVE INFORMATION ROUTING : AN INTEGRATED APPROACH WITH NETWORKING AND DATABASE APPLICATIONS

Kumar Viswanatha, Emrah Akyol, Sharadh Ramaswamy and Kenneth Rose

{kumar, eakyol, rsharadh, rose}@ece.ucsb.edu
Department of Electrical and Computer Engineering
University of California at Santa Barbara, CA-93106.

ABSTRACT

This paper considers the design of optimal joint compression-routing schemes for networks with correlated sources and multiple sinks. Such a setting is typically encountered in sensor networks. It is sometimes tempting to assume that an optimal distributed compression scheme followed by standard routing (as designed for independent sources) would be nearly optimal. We instead propose a joint design approach that integrates distributed source coding with a mechanism called dispersive information routing. Unlike network coding, dispersive information routing can be realized using conventional routers without recourse to recoding at intermediate nodes. We also point out the direct connections between dispersive information routing and a related problem in sensor network databases, namely, fusion coding for selective retrieval. We propose an efficient practical design strategy, variants of which are adopted for each of the two problems. Simulation results provide evidence for substantial gains over conventional schemes.

Index terms - Compression, distributed source coding, routing, sensor networks, database retrieval

1. INTRODUCTION

The field of distributed source coding (DSC) began in the seventies with the seminal work of Slepian and Wolf [13]. They showed, in the context of lossless coding, that side-information available only at the decoder can nevertheless be fully exploited as if it were available to the encoder, in the sense that there is no asymptotic performance loss. Later, Wyner and Ziv [14] derived a lossy coding extension. A number of theoretical publications followed, primarily aimed at solving the general multi-terminal source coding problem. It was not until the late nineties, when practical DSC schemes adopting principles from channel coding were being designed, where notably influential was the "DISCUS" approach [6]. An alternative approach to DSC sprung directly from principles of source coding. Algorithms for distributed vector quantizer design have been proposed in [3, 11]. Algorithms to optimize the fundamental tradeoffs in the (practically unavoidable) case of sources with memory have recently appeared in [12]. Similar design approaches have shown to be efficient to handle the exponential growth in codebook size in the context of distributed coding for large number of sources [10]. The source coding approaches will be most relevant to us here and will be discussed briefly in Section 3.2.

Compression in multi-hop networks has gained significant importance in recent years, mainly due to its relevance in sensor network applications. Review paper [4] describes the various joint compression-routing schemes that have been developed so far. Encoding correlated sources in a network with multiple sources and sinks has conventionally been looked at from two different directions. The first approach performs compression at intermediate nodes [4], where all the information is available, without appealing to DSC principles. However, such approaches tend to be wasteful at all but the last hops of the communication path. The second approach uses distributed source coding to exploit correlation at the source nodes followed by simple routing at intermediate nodes. Well designed DSC could provide considerable performance improvement and/or complexity/energy savings. Various aspects of DSC for routing have been considered in a number of publications, and notably in [2], where the authors consider joint optimization of Slepian - Wolf coding and conventional routing.

Optimal routing schemes, designed for independent sources (conventional routing), have been studied extensively [1], primarily due to the growth of the Internet. It may be tempting to assume that an optimal distributed source code, which tries to eliminate the dependencies between sources, followed by a conventional routing mechanism would achieve optimality for correlated sources. In this paper, we introduce the concept of "dispersive information routing" (DIR) coupled with joint (compression-routing) optimization approach, which offer substantial performance gains for correlated sources. We first show using a simple network example the sub-optimality of conventional methods and motivate the approach. To demonstrate its potential we then derive a gradient based method for joint DSC - DIR design and apply it to a sensor grid with multiple sources and sinks. Unlike network coding [7], DIR can be realized using conventional routers without recourse to expensive coders at intermediate nodes.

A different problem that is highly relevant to correlated sources, and that perhaps surprisingly reveals underlying conceptual similarity, is that of storage and retrieval of correlated sources from a database at a fusion center - called fusion coding. This problem has been studied recently both from the information theoretic [5] and the design perspectives [8]. In this paper, we review the motivations, ideas and the methodologies involved. Our primary outlook is to point out the close connection between fusion coding and dispersive information routing and thereby illustrate the general applicability of the approaches herein.

2. NOTATIONS AND PROBLEM SETUP

Let a network be represented by an undirected graph $G = (V,E)$. Each edge $e \in E$ is a network link whose communication cost depends on the edge weight w_e. The nodes V consist of N source nodes, M sinks, and $|V| - N - M$ intermediate nodes[1]. Source node i has access to source random variable X_i. The joint probability distribution of $(X_1 \dots X_N)$ is known at all the nodes. The sinks are denoted $S_1, S_2 \dots, S_M$. Each sink requests the information of a subset of sources. Let the subset of nodes requested by sink S_j be $V^j \subseteq V$. Conversely, source i has to be reconstructed at a subset of sinks denoted $S^i \subseteq \{S_1, S_2 \dots, S_M\}$.

Define traffic matrix (or "request" matrix) T, for network graph G as the $N \times M$ binary matrix that specifies which sources must be

The work was supported by the NSF under grant CCF-0728986

[1] For any set B, we use $|B|$ to denote the set cardinality.

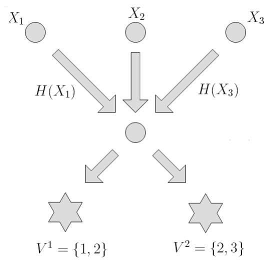


Figure 1: Example to motivate DIR: S_1 requests for X_1 and X_2 while S_2 requests for X_2 and X_3.

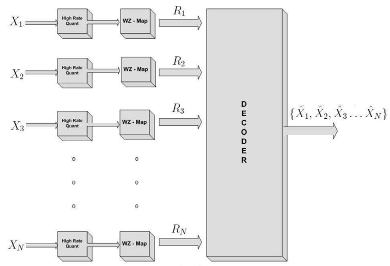


Figure 2: Conventional DSC design

reproduced at which sinks:

$$T_{ij} = \begin{cases} 1 & \text{if } i \in V^j \\ 0 & \text{else,} \end{cases}$$

i.e., $V^j = \{i : T_{ij} = 1\}$ and $S^i = \{S_j : T_{ij} = 1\}$. Without loss of generality we assume that every source is requested by at least 1 sink.

The cost of communication through a link is a function of the bit rate flowing through it and the edge weight, which we will assume for simplicity to be a simple product $f(r, w_e) = rw_e$, noting that the approach is directly extendible to more complex cost functions. The objective is to design encoders, routers and decoders to minimize the overall network cost (calculated given the set of link rates and edge weights) at a prescribed level of average distortion.

We denote by E_B^i, the set of all paths from source i to the subset of sinks $B \subseteq \{S_1 \ldots S_M\}$. The optimum route from the source to these sinks is determined by a spanning tree optimization (minimum Steiner tree) [1]. More specifically, for each source node i, the optimum route is obtained by minimizing the cost over all trees rooted at node i which span all sinks $S_j \in B$. The minimum cost of transmitting a single bit from source i to the subset of sinks B, denoted by $d_i^*(B)$, is given by:

$$d_i^*(B) = \min_{P \in E_B^i} \sum_{e \in P} w_e \qquad (1)$$

3. DISTRIBUTED SOURCE CODING AND DISPERSIVE INFORMATION ROUTING

3.1 Information Theoretic Motivation

Figure 1 depicts a simple network with 3 sources (X_1, X_2 and X_3) and two sinks (S_1 and S_2). S_1 requests for X_1 and X_2 while S_2 requests for X_2 and X_3. There is one intermediate node, c (which we call the collector), which serves as a router. The sources can communicate with the sinks only through the collector. This is a toy example or simplification of a large sensor network with all intermediate nodes collapsed to a single collector node. Note that, although the motivation is provided for clarity within the lossless setting, the practical design approach will be generally applicable to the lossy coding setting.

The collector has to transmit enough information to S_1 so that it can decode both X_1 and X_2, and enough information to S_2 to decode X_2 and X_3. Hence the rates on the edges (c, S_1) and (c, S_2) are at least $H(X_1, X_2)$ and $H(X_2, X_3)$, respectively. Let the edge weights into the collector be much smaller than out of the collector: $w_{1,c}, w_{2,c}, w_{3,c} \ll w_{c,S_1}, w_{c,S_2}$. This would force source X_1 and X_3 to transmit data at rates $H(X_1)$ and $H(X_3)$, respectively. As source X_2 has to transmit enough data for both the sinks to decode it losslessly:

$$R_2 \geq \max(H(X_2|X_1), H(X_2|X_3)) \qquad (2)$$

Conventional routing methods (designed for independent sources) do not "split" a packet at an intermediate node and hence would forward all the bits from X_2 to both the sinks. This would mean sub-optimality on either one of the branches (c, S_1) or (c, S_2) if $H(X_2|X_1) \neq H(X_2|X_3)$.

But instead, let us now relax this restriction and allow the collector to "split the packet" and route only a subset of bits on each edge. We could equivalently think of source X_2 transmitting 3 smaller packets to the collector; first packet has rate $R_{2,\{1,2\}}$ bits and is destined to both the sinks. Two other packets have rates $R_{2,1}$ and $R_{2,2}$ bits and are addressed to sinks S_1 and S_2, respectively. Technically, in this case, the collector would just have to route the received packets. Perhaps surprisingly, it can be shown using random product binning arguments similar to [13] that the rate tuple $(R_{2,\{1,2\}}, R_{2,1}, R_{2,2}) = (H(X_2|X_1, X_3), I(X_2, X_3|X_1), I(X_2, X_1|X_3))$ is indeed achievable, and the rates on edges $(c, S1)$ and $(c, S2)$ achieve their respective lower bounds, i.e., this routing approach is implemented at no asymptotic rate loss, and hence outperforms conventional routing.

We term such a routing mechanism, where each intermediate node in a multi-hop network can *route any subset* of the received bits on each of the forward paths as "dispersive information routing" (DIR). Note the clear difference from network coding. DIR does not require expensive coders at intermediate nodes, but rather can always be realized using conventional routers with each source transmitting multiple packets into the network intended to different subsets of sinks. Also note that such a routing mechanism is inessential when the sources are independent.

3.2 Conventional distributed source coder

We return to general lossy coding and begin with a description of the conventional DSC system with a single sink. Consider N correlated sources, $\{X_i, i = 1...N\}$ transmitting information at rate R_i, respectively, to the sink. The encoding consists of 2 stages. The first stage is the discretization of the source-space by a high-rate quantizer (a practical engineering necessity, see e.g., [11]), which partitions the input source space into N_i regions, i.e.,

$$\mathscr{H}_i : \mathscr{R} \rightarrow Q_i = \{1 \ldots N_i\} \qquad (3)$$

Note that the high rate quantizers are designed independent of the rest of the modules. The second stage, which we call the 'Wyner-Ziv map' (WZ-map), is a module that relabels the N_i quantizer regions with a smaller number, 2^{R_i}, of labels, which if properly designed exploits the correlation between the sources and aids in distributed compression. Mathematically, for each source i, the WZ map is the function,

$$\mathscr{W}_i : Q_i \rightarrow I_i = \{1 \ldots 2^{R_i}\} \qquad (4)$$

and the encoding operation can be expressed as:

$$\mathscr{E}_i(x_i) = \mathscr{W}_i(\mathscr{H}_i(x_i)) \quad \forall i \qquad (5)$$

We use $I = I_1 \times I_2 \ldots I_N$ to denote the set of all possible received symbols at the decoder. The decoder $\mathscr{D}_i$ for each source is given by:

$$\mathscr{D}_i : I \rightarrow \hat{X}_i \subset \mathscr{R} \qquad (6)$$

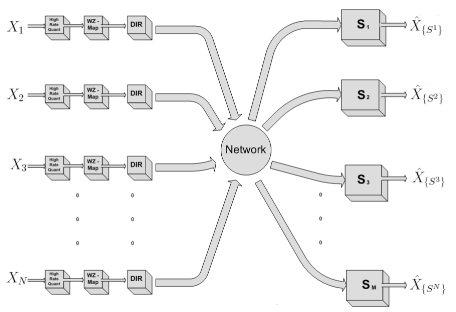


Figure 3: Integrated DSC and DIR

We next consider a network as formulated in section 2 with each sink S_j requesting for a subset $V^j \subseteq V$ of sources. We first describe two routing options used in practice and point to their drawbacks.

1) Broadcasting : All the bits from each source are routed to all the sinks (Broadcasted). Then the decoder $\mathscr{D}_{ij}$ for source X_i at sink S_j can be expressed as:

$$\mathscr{D}_{ij} : I \rightarrow \hat{X}_{ij} \subset \mathscr{R} \quad \forall j, i \in V^j \tag{7}$$

Such a routing mechanism is obviously highly inefficient in terms of communication costs, as unused information is delivered to various sinks.

2) Conventional routing : All the bits transmitted by source i are routed to its destination sinks $S_j \in S^j$. If we use $I_{S_j} = \prod_{i \in S^j} I_i$, to denote the set of all possible received symbols at sink S_j, then the decoder $\mathscr{D}_{ij}$ can be expressed as:

$$\mathscr{D}_{ij} : I_{S_j} \rightarrow \hat{X}_{ij} \subset \mathscr{R} \quad \forall j, i \in V^j \tag{8}$$

Such a routing technique is questionable in the presence of inter-source dependencies. An unrequested but correlated source may be nearer and provide "less costly" information on the sources requested. Also, only a subset of the bits transmitted by a source may be sufficient to reconstruct it at certain decoders (for eg. section 3.1).

3.3 Integrated DSC and dispersive routers

We now describe our approach to dispersive information routing. *We allow each bit to be routed to any subset of the sinks.* We introduce a new module at each encoder which decides the route for each bit generated at that encoder. Note that if each bit follows the route prescribed by the encoders, every intermediate node would just be forwarding a subset of the received bits on each of the forward paths. We call this module the "dispersive information router" to indicate the routing mechanism it induces in the network. We denote by $S = \{S_1, S_2 \ldots S_M\}$ the set of all sinks and by 2^S the power set (set of all subsets) of S. Formally, the router at the i^{th} encoder is given by:

$$\mathscr{C}_i : \{1 \ldots R_i\} \rightarrow 2^S \tag{9}$$

and denote by $\mathscr{C} = \mathscr{C}_1 \times \mathscr{C}_2 \ldots \mathscr{C}_N$. The routers uniquely determine the set of all the received bits at each sink. The decoder at each sink is now modified to be the mapping:

$$\mathscr{D}_{ij} : I \times \mathscr{C} \rightarrow \hat{X}_{ij} \subset \mathscr{R} \quad \forall j, i \in V^j \tag{10}$$

The total communication cost W of the system is given by:

$$W = \sum_{i=1}^{N} \sum_{j=1}^{R_i} d_i^* \left(\mathscr{C}_i(j)\right) \tag{11}$$

and the average reconstruction distortion is measured as:

$$D = E\left[\sum_{j=1}^{M} \sum_{i=1}^{|V^j|} \gamma_{ij} d_{ij}(X_i, \hat{X}_{ij})\right] \tag{12}$$

where $d_{ij} : R \times R \rightarrow [0, 1)$ is a well-defined distortion measure and γ_{ij} are used to weigh the relative importance of each source at each sink to the total distortion.

Hereafter, we will specialize to the squared error distortion. In practice, we only have access to training sets and not the actual source distributions. Assuming ergodicity, expectation is approximated by simple averaging over the training set. If the training set is denoted as $\mathscr{T}$, we measure distortion as,

$$D = \frac{1}{|\mathscr{T}|}\left[\sum_{\mathbf{x} \in \mathscr{T}} \sum_{j=1}^{M} \sum_{i=1}^{|V^j|} \gamma_{ij}(x_i - \hat{x}_{ij})^2\right] \tag{13}$$

where $\mathbf{x} = \{x_1 \ldots x_N\}$. The trade off between the distortion and the communication cost is controlled using a Lagrange parameter $\lambda \geq 0$ to optimize the weighted sum of the two quantities. From (13) and (11), the Lagrangian cost to be minimized is:

$$L = \left[\sum_{\mathbf{x} \in \mathscr{T}} \sum_{j=1}^{M} \sum_{i=1}^{|V^j|} \frac{\gamma_{ij}}{|\mathscr{T}|}(x_i - \mathscr{D}_{ij}(I, \mathscr{C}))^2\right] + \lambda \sum_{i=1}^{N} \sum_{j=1}^{R_i} d_i^* \left(\mathscr{C}_i(j)\right) \tag{14}$$

where $\mathbf{I} = [\mathscr{E}_1(x_1), \mathscr{E}_2(x_2) \ldots \mathscr{E}_N(x_N)]^T$ denotes the set of all bits being routed in the network. The objective is to find $\mathscr{E}_i's$, $\mathscr{C}_i's$ and $\mathscr{D}_{ij}'s$ that minimize L for a given λ.

3.4 Necessary conditions for optimality

In the following two sections, we derive the necessary conditions for optimality of all modules and propose an iterative gradient based method for optimal design.

1) Optimal encoders : Let $\mathscr{T}_{i,j} = \{\mathbf{x} \in \mathscr{T} : \mathscr{H}_i(x_i) = j\}$. Note that the second term in the Lagrangian does not depend on the WZ maps and hence, from (14), the optimum WZ-map for fixed routers and decoder codebook is given by:

$$\mathscr{W}_i(j) = k^* = \arg\min_{k \in I_i} \sum_{\mathbf{x} \in \mathscr{T}_{i,j}} \sum_{m=1}^{M} \sum_{l=1}^{|V^j|} \gamma_{lm}(x_l - \mathscr{D}_{lm}(\mathbf{I}_{i,k}, \mathscr{C}))^2 \tag{15}$$

where

$$\mathbf{I}_{i,k} = [\mathscr{E}_1(x_1), \ldots, \mathscr{E}_{i-1}(x_{i-1}), k, \mathscr{E}_{i+1}(x_{i+1}), \ldots, \mathscr{E}_N(x_N)]^T$$

2) Optimal routers : For fixed encoders and decoders the optimum dispersive information router for bit j transmitted by source i is given by:

$$\mathscr{C}_i(j) = e^* = \arg\min_{e \in 2^S} \sum_{\mathbf{x} \in T} \sum_{m=1}^{M} \sum_{l=1}^{|V^j|} \gamma_{lm}(x_l - \mathscr{D}_{lm}(\mathbf{I}, \mathbf{C}_{i,j}^e))^2 + \lambda d_i^*(e) \tag{16}$$

where

$$\mathbf{C}_{i,j}^e = [C_1(1), C_1(2) \ldots C_i(j-1), e, C_i(j+1), \ldots C_N(R_N)]^T$$

3) Optimal decoders : If $\mathbf{I} = [\mathscr{E}_1(x_1), \mathscr{E}_2(x_2) \ldots \mathscr{E}_N(x_N)]^T$ represents the bits transmitted from all the sources, and if e represents the positions of bits received by a decoder, then we use $\mathbf{I}_e$ to denote the index represented by the bits in $\mathbf{I}$ at the positions indicated by e. Setting to zero the gradient of L (14) with respect to the reconstruction values, yields the optimal decoder:

$$\hat{x}_{ij}(I, e) = \mathscr{D}_{ij}(\mathbf{I}, e) = \frac{1}{|\mathscr{F}|} \sum_{\mathbf{x} \in \mathscr{F}} x_i \tag{17}$$

where $\mathscr{F} = \{\mathbf{x} \in \mathscr{T} : ([\mathscr{E}_1(x_1), \mathscr{E}_2(x_2) \ldots \mathscr{E}_N(x_N)]^T)_e = \mathbf{I}_e\}$.

3.5 Algorithm for joint DSC - DIR design

A natural design approach, given the necessary conditions for optimality, is to use gradient descent. For each value of λ, all the modules are initialized randomly and iteratively optimized till convergence is reached. When each module is optimized, the Lagrangian cost reduces. Since there are only a finite number of source partitions possible, convergence to a locally optimal solution is (in principle) guaranteed. An operational cost-distortion curve is obtained by varying λ, thereby trading off total communication cost to average distortion. To mitigate the issue of local minima, the system is optimized over multiple random initializations. Global optimization techniques such as DA [11], can be used to avoid poor local minima, but are beyond the scope of this paper. Also note that the proposed design approach is centralized in the sense that the optimization is done offline, at a central location, before the network operates. In this paper, we aim to establish potential gains from using DIR in a practical setting, which in turn promotes future research for developing efficient decentralized design strategies.

4. ANALOGIES TO DATABASE APPLICATIONS - FUSION CODING

Consider a fusion center, which receives information from multiple correlated sources and has to store the data for later retrieval and usage, such as in a sensor network monitoring temperatures. Periodically, users request information from *subsets of these sources* ("user queries" in database lingo), where the interesting subsets are a priori unknown to the fusion center. Now, an interesting trade-off between the (compression) storage rate and the retrieval time (rate), very similar to the one discussed previously in the context of networks, arises in such a setting. On the one hand, joint coding of all the sources reduces the overall storage requirement. But on the other hand, such joint coding may necessitate retrieval of all the stored bits to decode any subset of the sources. Hence, joint coding could be highly wasteful with respect to the amount of information retrieved when a future query may select only a few of the sources for retrieval.

Both the information theoretic (asymptotic) characterization [5] and the practical fusion coder design [8, 9] have been considered earlier. We refer the reader to these papers for a more comprehensive overview of the design approaches and trade-offs involved. Here we briefly describe the motivation and the methodologies and point out the similarities to dispersive information routing.

4.1 Information theoretic motivation (lossless setting)

Consider N correlated sources $(X_1, X_2 \dots X_N)$ to be stored at a fusion center. Define a query q $(q \in Q)$ as the subset of sources that need to be retrieved, where Q is the set of all subsets of sources. We assume that the user queries q with a known prior probability of $P(q)$ $(\sum_{q\in Q} P(q) = 1)$. The total storage rate is denoted by R_s and the average retrieval rate is denoted by R_r. The number of bits retrieved for query q is denoted by R_q.

First consider joint storage of all the sources. This requires $R_s = H(X_1, X_2 \dots X_N)$ storage bits. But all the stored bits have to be retrieved for any query implying an average retrieval rate of:

$$R_r = H(X_1, X_2 \dots X_N) \tag{18}$$

The minimum retrieval rate possible is given by,

$$R_{r,min} = \sum_{q\in Q} P(q) H(X_q) \tag{19}$$

which requires to be separately compressed and stored each subset of sources that could be queried. Important intuition to be gained from the above argument is that better retrieval rates can be obtained by storing more bits than required and retrieving a subset of them for each query. Information theoretic achievable bounds on the minimum R_r for a fixed R_s were obtained in [5].

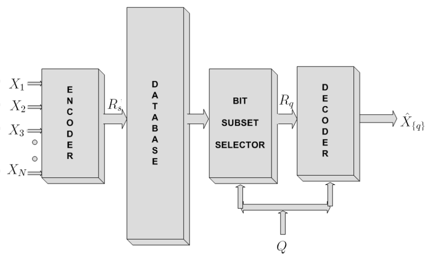


Figure 4: Fusion Coder

Note the similarities between dispersive information routing and fusion coding. In the former, each sink uses a subset of the transmitted bits to reconstruct the requested sources. Whereas in the fusion coding setup, each query retrieves a subset of the stored bits to reconstruct the queried sources. Similarly, note the analogy between joint compression in the context of databases and broadcasting in networks. Hence we use very similar design approaches in both the contexts.

4.2 Fusion coder design

The Fusion Coder (FC) (Figure 4) is composed of three modules - an encoder, a decoder and a 'bit-subset selector'. The last module forms the central component of FC as it retrieves a subset of the encoded bits for each query (subset of queried sources). The encoder is defined by the function:

$$\mathscr{E} : R^N \rightarrow I = \{0,1\}^{R_s} \tag{20}$$

which compresses the N dimensional input vector to R_s bits. Formally, the bit-subset selector is the mapping:

$$\mathscr{S} : Q \rightarrow \mathscr{B} = 2^{\{1\dots R_s\}} \tag{21}$$

where $\mathscr{B}$ is the set of all subsets of the stored bits. Observe that the role played by the bit-subset selector in the context of databases resembles closely the function of the dispersive information routers for networks. The decoder can be expressed as:

$$\mathscr{D} : I \times B \rightarrow \hat{X} \subset R^N \tag{22}$$

The average distortion and the average retrieval rate are given by:

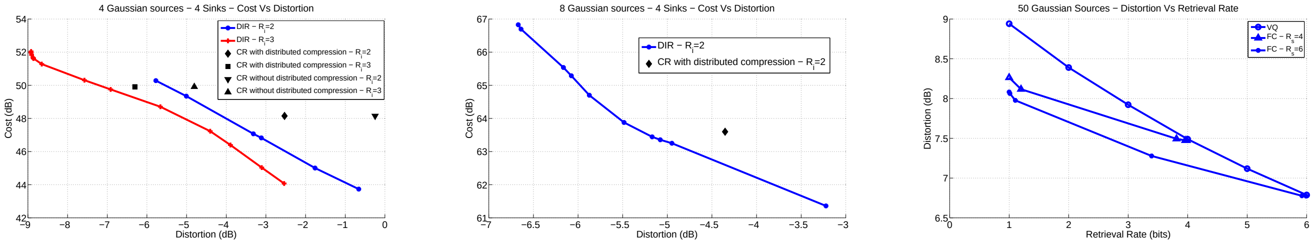
$$D = E\left[d_q(X, \mathscr{D}(\mathscr{E}(X), \mathscr{S}(q)))\right] \qquad R_r = \sum_{q\in Q} P(q)|\mathscr{S}(q)| \tag{23}$$

A Lagrangian similar to (14) is set up to trade off distortion with the retrieval rate. We refer the reader to [8] for further details on the necessary conditions for optimality and the algorithm for system design, which go hand in hand with those given in Section 3.4.

5. RESULTS

5.1 Dispersive information routing

We simulated a sensor network with 4 sinks located at the 4 verticess of a square grid (100×100). The sources and the intermediate nodes were placed randomly inside the grid, to mimic a real world scenario with inaccessible regions. Each element in the request matrix, T, was generated using uniform binary $(0,1)$ random variable. We considered synthetic Gaussian sources, $N(0,1)$, with the correlation dropping exponentially with the distance. Specifically, the correlation between two sources at a distance d was assumed to be $\rho^{\frac{d}{d_0}}$. We used the square and cube of the distances as the edge weights (w_e) for simulations 1 and 2 respectively. The Steiner tree optimized costs $(d_i^* s)$ were computed before the design



(a) DIR-DSC - 4 Source example (b) DIR-DSC - 8 Source example (c) Fusion Coding - 50 Sources

Figure 5: Simulation results

of the modules, using an optimal Steiner tree algorithm [2]. In the figures we use DIR to indicate our approach and CR to indicate conventional routing.

For both our experiments, we chose $\rho = 0.95$, $d_o = 100$ and $\gamma_{ij} = 1 \forall i, j$. The system was trained using a training sequence and tested on a test set, each of size $200,000$. For a source transmitting at a rate R_i, the high rate quantizer partitioned the source space into 2^{R_i+3} regions, for eg. if the source rate was 2 bits, $N_i = 32$. For all methods, we report the best performance over several random initializations (limited to 25).

4 Sources : Figure 5a shows the results obtained for 4 sources with 4 other intermediate nodes. The simulations were conducted at two different source rates, $R_i = 2$ and 3. Our approach gains about 1.5 dB and 1 dB in distortion over conventional routing for the two rates respectively. The reduction in gain for $R_i = 3$ is probably due to more local minima in the cost function at higher source rates. For comparison, we have also indicated the distortion obtained with conventional routing without using distributed compression.

8 Sources : We simulated the sensor grid with 8 sources at a source rate of $R_i = 2$. Figure 5b shows a comparison between the two approaches. Our approach gains about 1 dB in distortion over conventional routing, or conversely, we gain about 1 dB in cost for a fixed distortion.

5.2 Fusion coding

Similar plots were obtained to show distortion versus average retrieval rate for a database with 50 synthetic Gaussian sources with a correlation coefficient of 0.8. An *exponential query* distribution was used to generate the queries. Further details on the implementation are available in a coming paper [8]. As can be seen from Figure 5c, our design approach reduces the retrieval rate by a factor of $3X$ (and hence speeds up retrieval by about $3X$), at a distortion of 8 dB.

6. CONCLUSION

We introduced dispersive information routing (DIR) which allows for subsets of a source information to be directed towards different sinks by intermediate nodes in the network. The approach was motivated by consideration of a simple network and information theoretic derivation in the lossless setting which clearly demonstrate the means by which DIR gains over conventional approaches. We proposed a joint optimization of distributed coding and dispersive router, and for practical demonstration implemented it via gradient descent. We also pointed out the similarities to a related problem of joint compression and selective retrieval of correlated sources, which was studied recently in the context of sensor network databases. Simulation results show considerable gains of our approach over conventional methods. Future work includes obtaining information theoretic bounds on the minimum achievable cost using DIR and developing distributed approaches for efficient design.

[2] It is known that the optimum Steiner tree optimization is NP - Complete and hence requires approximate algorithms to solve in practice for large networks [2].

REFERENCES

[1] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to algorithms, second edition*. McGraw-Hill Science/Engineering/Math, July 2001.

[2] R. Cristescu, B. Beferull-Lozano, and M. Vetterli. Networked Slepian-Wolf: Theory, algorithms and scaling laws. *IEEE Trans. on Information Theory*, 51(12):4057–4073, 2005.

[3] M. Fleming, Q. Zhao, and M. Effros. Network vector quantization. *IEEE Trans. on Information Theory*, 50:1584–1604, Aug 2004.

[4] H. Luo, Y. Liu, and S. K. Das. Routing correlated data in wireless sensor networks: A survey. *IEEE Network*, 21(6):40–47, 2007.

[5] J. Nayak, S. Ramaswamy, and K. Rose. Correlated source coding for fusion storage and selective retrieval. In *Proceedings of IEEE International Symposium on Information Theory*, pages 92–96, Sep 2005.

[6] S. S. Pradhan and K. Ramchandran. Distributed source coding using syndromes (DISCUS): Design and construction. *IEEE Trans. on Information Theory*, 49:626–643, 1999.

[7] A. Ramamoorthy, K. Jain, P. A. Chou, and M. Effros. Separating distributed source coding from network coding. *IEEE Trans. on Information Theory*, 52(6):2785–2795, 2006.

[8] S. Ramaswamy, J. Nayak, and K. Rose. Fusion coding of correlated sources for storage and selective retrieval. *IEEE Trans. on Signal Processing*, 58, Mar 2010.

[9] S. Ramaswamy and K. Rose. Shared descriptions fusion coding of correlated sources for storage and selective retrieval. In *Proceedings of IEEE Data Compression Conference*, 2008.

[10] S. Ramaswamy, K. Viswanatha, A. Saxena, and K. Rose. Towards large scale distributed coding. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Mar 2010.

[11] A. Saxena, J. Nayak, and K. Rose. Robust distributed source coder design by deterministic annealing. *IEEE Tran. on Signal Processing*, 58:859 – 868, Feb 2010.

[12] A. Saxena and K. Rose. Distributed predictive coding for spatio-temporally correlated sources. *IEEE Tran. on Signal Processing*, pages 4066–4075, Oct 2009.

[13] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE. Trans. on Information Theory*, 19:471–480, Jul 1973.

[14] A. D. Wyner and J. Ziv. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. on Information Theory*, 22:1–10, Jan 1976.