Towards Optimum Cost in Multi-Hop Networks with Arbitrary Network Demands

Kumar Viswanatha, Emrah Akyol and Kenneth Rose
{kumar, eakyol, rose}@ece.ucsb.edu
Department of Electrical and Computer Engineering
University of California at Santa Barbara, CA-93106.

***Abstract*—This paper considers the problem of minimizing the communication cost for a general multi-hop network with correlated sources and multiple sinks. For the single sink scenario, it has been shown that this problem can be decoupled, without loss of optimality, into two separate subproblems of distributed source coding and finding the optimal routing (transmission structure). It has further been established that, under certain assumptions, such decoupling also applies in the general case of multiple sinks and arbitrary network demands. We show that these assumptions are significantly restrictive, and further provide examples to substantiate the loss, including settings where removing the assumptions yields unbounded performance gains. Finally, an approach to solving the unconstrained problem, where routing and coding cannot be decoupled, is derived based on Han and Kobayashi's achievability region for multi-terminal coding.**

***Index Terms*—Multi-hop sensor networks, distributed source coding, multi-terminal information theory**

I. Introduction

Compression of sources in conjunction with communication over a network has been an important research area, notably with the recent advancements in distributed compression of correlated sources and network (routing) design, coupled with the deployment of various sensor networks. Encoding correlated sources in a network, such as a sensor network with multiple nodes and sinks, has conventionally been approached from two different directions. The first approach is routing the information from different sources in such a way as to efficiently recompress the data at intermediate nodes without recourse to distributed coding methods (we refer to this approach as joint coding via "explicit communication"). The second approach is exploiting the correlation between sources while compressing at the nodes even if the sources do not explicitly communicate with each other, i.e., distributed source coding (DSC), and routing the information accordingly. Relevant background on DSC and route selection in a network is given in the next section.

For a single sink network, it has been shown that optimal DSC followed by an optimal shortest path routing mechanism is unbeatable with respect to minimizing the total communication cost when the cost is a convex function of the link data rate [1], [2]. Also, it has been shown that for large networks, DSC followed by optimum routing can have unbounded gains compared to explicit communication [1]. But for the more general network consisting of multiple sinks with arbitrary network demands, such a mechanism has been shown to be optimal only under an important assumption. In this paper we primarily look at the extent of sub-optimality caused due to this assumption. Motivated by the potential unbounded gains by removing the assumption, we design a solution to the problem using prior results from multi-terminal information theory.

The rest of the paper is organized as follows. In section II, we discuss the related work. In section III we formulate the problem and discuss the prior results to substantiate the assumption. We present examples of settings to show unbounded gains in section IV and propose a solution to the unconstrained problem in section V.

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

II. Background and Related Work

A. Distributed Source Coding

The field of DSC began in the seventies with the seminal work of Slepian and Wolf [3]. 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 [4] derived a lossy coding extension that bounds the rate-distortion performance in the presence of decoder side information. Han and Kobayashi [5] extended the Slepian-Wolf result to general multi-terminal source coding scenarios (see also [6]). We adopt their results in this work.

B. Compression in Multi-hop Networks

A survey of routing techniques in sensor networks is given in [7]. Most of these approaches only exploit inter-source correlations via "explicit communication", i.e., joint compression is performed at intermediate nodes where all information is available, without appeal to distributed coding principles. However, such approaches tend to be wasteful at all but the last hops of the communication path. 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. Cristascu et.al [1] considered joint optimization of Slepian-Wolf coding and routing, and provided a solution to this problem whose optimality depends on constraining assumptions. The scenario of multi-sink was considered in [8], where a practical suboptimal distributed scheme was proposed. [9]

provided a proof of the optimum communication route for 'correlated data gathering' is an NP complete problem, along with low complexity suboptimal algorithms. [10] compared different joint compression-routing schemes for a correlated sensor grid and also proposed an approximate, practical, static source clustering scheme to achieve compression efficiency. Note that there has been considerable amount of work in the field of network coding related to compression in networks with capacity constraints [11], [12]. In this paper we restrict ourselves to the conventional routing schemes and do not consider the possibility of network coding at intermediate nodes.

III. Problem setup

A. Problem Formulation

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. Source node i has access to source random variable X_i. The joint probability distribution of $(X_1 \ldots X_N)$ is known at all the nodes. The sinks are denoted $S_1, S_2 \ldots, S_M$. 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 $S^i \subseteq \{S_1, S_2 \ldots, S_M\}$. For any subset, $\mathcal{B}$, of encoders, we denote by $X_{\mathcal{B}} = \{X_i : i \in \mathcal{B}\}$.

Define traffic matrix (or "request" matrix) T, for network graph G as the $N \times M$ binary matrix that specifies which sources must be 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 $S^i = \{S_j : T_{ij} = 1\}$. We denote by E^i, the set of all paths from source i to its destination sinks S^i.

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 minimize the overall network cost (calculated given the set of link rates and edge weights) for lossless reconstruction of source information at appropriate sinks.

B. Prior Results

Highly relevant to this work are recent results from [1] on "networked Slepian-Wolf".

i) *Single sink requesting all sources*: It is key to observe here that regardless of the rate allocation at the nodes, to minimize the communication cost, source bits must traverse the network through the minimum cost path to reach the destination sink. This observation offers considerable simplification, as it implies that optimization of the transmission structure can be done independently of rate allocation. Then, given the optimal transmission structure, the optimal rate allocation can be found. In fact, the optimal transmission structure can be easily shown to be the shortest path tree (SPT) rooted at the sink. Define

$$d_i^* = \min_{\mathcal{P} \in E^i} \sum_{e \in \mathcal{P}} w_e$$

as the minimum total weight a bit encounters on its path from source i to the sink. The remaining task is to find a rate allocation at the sources so that the total communication cost is minimized while ensuring lossless reconstruction. This implies that the optimal rates allocated have to lie in the Slepian-Wolf achievable rate region, which boils down to solving the linear programming problem for the rates :

$$\min_{\{R_i\}_i^N} \sum_{i=1}^{N} R_i d_i^*$$

subject to the Slepian-Wolf rate constraints.

$$\sum_{i \in Y} R_i \geq H(X_S | X_{S^c}) \quad \forall S \subseteq \{X_1, X_2 \ldots X_N\} \tag{1}$$

Remarkably, Liu et. al. [2] claim broader optimality of the single sink networks of [1] (Slepian Wolf encoding and shortest path routing): It is optimal over all possible joint coding-routing schemes, including any processing at intermediate nodes, i.e., network coding cannot improve the performance when the link communication cost is a convex function of the link data rate [1].

ii) *Multiple Sinks* : Here we no longer assume that sinks request all sources, hence we have a non-trivial request matrix T, which specifies which sources need to be reproduced at each sink. We further say that source i communicates with sink S_j if the sink receives the bits transmitted by the source, and denote this by $i \to S_j$. We define a $N \times M$ *communication matrix* C as:

$$C_{ij} = \begin{cases} 1 & \text{if } i \to S_j \\ 0 & \text{else,} \end{cases}$$

Note that C and T need not be the same. For the multiple-sink scenario, the above simple scheme of shortest path routing and Slepian-Wolf encoding is shown to be optimal under the following assumption [1]:

Assumption: Only requested sources, $i \in V^j$, communicate with sink S_j, i.e., $C = T$.

This assumption is clearly valid in the case of independent sources, but questionable otherwise. An unrequested but correlated source may provide less expensive information on requested sources. We will show how this assumption leads to substantial suboptimality. This assumption being the primary motivation for this work, in the next section, we first illustrate the sub-optimality using a simple example. Then we show settings where the performance gains achievable by removing the assumption is unbounded.

[1]Note that a capacity constraint makes the cost function non-convex

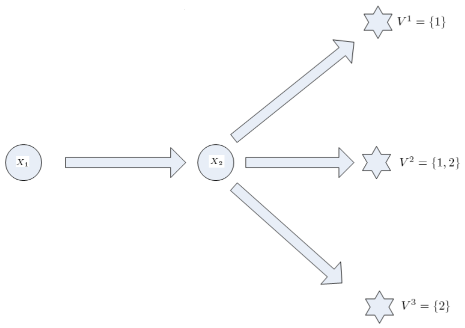

Figure 1: A simple example to illustrate the suboptimality due to the assumption

IV. Suboptimality due to the assumption

A. Example

The simple network of figure 1 connects two correlated identically distributed sources (X_1 and X_2) and three sinks (S_1,S_2,S_3). Let the marginal source entropies be $H(X_1) = H(X_2) = H$, and the conditional entropies be $H(X_1|X_2) = H(X_2|X_1) = h$, where $(h < H)$. Sinks S_1 and S_3 request for source X_1 and X_2, respectively, whereas sink S_2 requests for both X_1 and X_2. The available communication links are depicted in the figure 1. For simplicity, assume that $w_e = 1, \forall e \in E$.

The request matrix T is given by:

$$T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Clearly, communication matrix $C_1 = T$, allows for lossless reconstruction of the requested sources at each sink. Also, there are 3 other communication mtarices which allow for lossless reconstruction. They are given by:

$$C_2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} C_3 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, C_4 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

It is easy to see that the optimal communication cost for each of the above cases is given by:

$$\begin{aligned} W_1 = &\quad 3 \times H(X_1) + 2 \times H(X_2) = & 5H \\ W_2 = &\quad 3 \times H(X_2) + 4 \times H(X_1|X_2) = & 3H + 4h \\ W_3 = &\quad 4 \times H(X_1) + 2 \times H(X_2|X_1) = & 4H + 2h \\ W_4 = &\quad 3 \times H(X_2) + 3 \times H(X_1|X_2) = & 3H + 3h \end{aligned} \quad (2)$$

Clearly, $W_4 < W_2, W_3$. It follows directly that if $h/H < 2/3$, i.e., X_1 and X_2 are sufficiently dependent, then C_4 yields a lower communication cost than $C = T$. This simple example establishes the suboptimality of the assumption due to inter-source dependencies.

B. Asymptotic gains

It is of interest whether the gain due to eliminating the above limiting assumption is bounded. We analyze the asymptotic gain as the number/density of source nodes is increased within a given spatial region. Consider a sensor network example

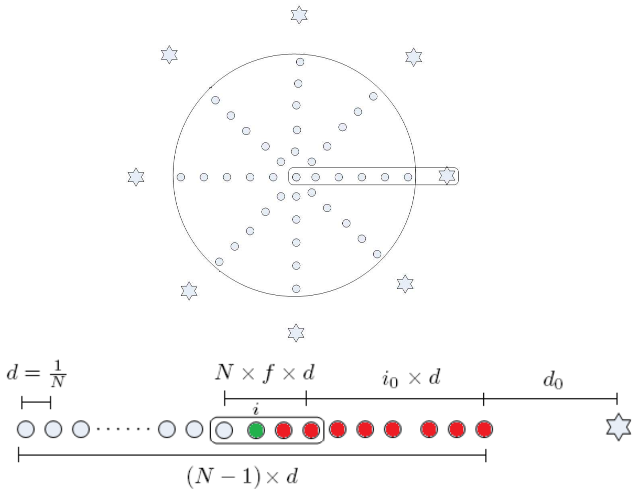

Figure 2: An example of a large network to illustrate asymptotic gains when all sources are allowed to communicate

shown in Fig. 2. The sources are distributed along several radii of the unit circle and the sinks are located at a distance d_0 from the periphery. For the sake of simplicity, we assume that there are N uniformly spaced sources on any radius and there are M sinks (M radii) in the network. It is assumed that a sink requests for all sources within a sub-interval of length $f \in [0, 1]$ on the radius connecting it with the center. Therefore each sink requests Nf sources. The communication cost of transmitting 1 bit over a distance d is assumed to be d^l for a given l.

Let $f_0 \in [0, 1-f]$ be the distance of the requested interval from the periphery. If we enumerate the source nodes on the radius by $0, 1, \ldots, N-1$, where node i produces source X_i, then the nearest requested source is $i_0 = Nf_0$. The sink requests for the subset of sources numbered $(i_0, i_0 + 1, \ldots i_0 + Nf)$. Under the assumption, only these sources may communicate their information to the sink. We will evaluate how much can be gained by allowing other (nearer) sources to send their information as well.

We define gain, γ, as the ratio of the minimum communication cost achievable under the assumption ($C = T$) to that of an unconstrained competitor ($C = C_0$) to be specified,

$$\gamma = \frac{Cost(C = T)}{Cost(C = C_0)} \quad (3)$$

We further introduce the following notation to denote conditional entropies:

$$a^i_{(j_1,j_2)} = H(X_i | X_{j_1}, X_{j_1 - 1} \ldots X_{j_2}) \ , \ j_2 < j_1$$

It follows that the minimum cost as a function of i_0 and f when $C = T$ is given by:

$$Cost(C = T) = \left(i_0 d^l + d_0^l\right) H(X_{i_0}) + \sum_{i=i_0+1}^{i_0+fN} i d^l a^i_{(i-1,i_0)} \quad (4)$$

Beyond the requested sources, the chosen competitor, C_0, also allows communicating information from all sources between the sink and the requested interval. The minimum cost when $C = C_0$ is given by:

$$Cost(C = C_0) = d_0^l H(X_{i_0}) + \sum_{i=2}^{i_0+fN} i d^l a^i_{(i-1,1)} \qquad (5)$$

Note that the maximum gain achievable by discarding the assumption is lower bounded by γ since C_0 may conceivably still be suboptimal. Also note that γ could be less than 1 for cases when C_0 has a higher communication cost compared to $C = T$. For those cases, the gain achievable by removing the assumption is lower bounded by 1. Remember that the maximum gain equals 1 if the sources are independent. To calculate the costs, we next assume that the sources $\{X_i\}$ form a stationary Markov chain.

1) Stationary Markov Chain: Denote the source entropies and conditional entropies by H and h respectively $(H > h)$, i.e. $H(X_i) = H$ and $H(X_i|X_{i-1}) = h \ \forall i$. Equations (4) and (5) simplify substantially and the costs as functions of i_0 and f are given by:

$$Cost(C = T) = \left(N f_0 d^l + d_0^l\right) H + \sum_{i=Nf_0+1}^{(f_0+f)N} i d^l h \qquad (6)$$

$$Cost(C = C_0) = d_0^l H + \sum_{i=2}^{(f_0+f)N} i d^l h \qquad (7)$$

For $l = 1$, letting $N \to \infty$ and expressing γ as a function of f_0 and f gives:

$$\gamma(f_0, f) = 1 + \left[\frac{2f_0 - cf_0^2}{2d_0 + c\,(f_0 + f)^2}\right] \qquad (8)$$

where $c = \lim_{N\to\infty} Nh(N)/H$. Note that the sensor density increases with N and the correlation increases correspondingly, hence $h(N) \to 0$. In the special case where the sources are sampled from a continuous-space Gauss-Markov process, $c(N) \to c$ where c directly determines the exponential decay of the correlation with distance. Pattem et.al [10] derived an empirical expression for joint entropy function for a real sensor network scenario using a data-set pertaining to rainfall. Their expression satisfies the above Markov property validating the use of such a source model to obtain gains.

Next we assume that the sinks are allowed to request for any Nf contiguous sources with equal probability. This implies that the distribution for the closest source follows a uniform density, i.e. $i_0 \sim \mathcal{U}(0, N - Nf + 1)$. Also, as the network size grows, $N, M \to \infty$, the distance between the sink and the requested interval approaches a continuum and therefore f_0 follows a continuous uniform density, i.e. $f_0 \sim \mathcal{U}(0, 1-f)$. The expressions for average costs as a function of f are given by:

$$Cost(C = T) = \frac{1}{1-f}\int_0^{1-f} \left[\left(N f_0 d^l + d_0^l\right) H\right] df_0 + \frac{1}{1-f}\int_0^{1-f}\left[\sum_{i=Nf_0+1}^{(f_0+f)N} i d^l h\right] df_0$$

$$Cost(C = C_0) = \frac{1}{1-f}\int_0^{1-f}\left[d_0^l H + \sum_{i=2}^{(f_0+f)N} i d^l h\right] df_0$$

For $L = 1$, evaluating the integrals and letting $N \to \infty$, we get γ as a function of f as:

$$\gamma(f) = \frac{3\left[(1-f) + cf + 2d_0\right]}{6d_0 + c(1 + f + f^2)} \qquad (9)$$

Cases of interest:

- $c(N) \to 0 \Rightarrow \gamma(f) \to 1 + ({(1-f)}/{2d_0})$. As the sources become highly dependent, $\gamma > 1$ for $f < 1$. In the Gauss-Markov setting this is the case that the correlation coefficient approaches 1. Also, as the sinks approach the periphery of the circle, i.e. $d_0 \to 0$, the gain grows unboundedly regardless of the size of the requested interval.
- $f \to 0 \Rightarrow \gamma = {(6d_0+3)}/{(6d_0+c)}$: There are a large number of sources and each sink requests for a very small subset. Clearly, for small values of c and d_0, which is typically the case in sensor networks, $\gamma(f) \gg 1$.

The bottom line is that there are considerable potential gains if one removes the assumption.

V. Optimization when assumption is removed

We next consider the problem of finding the optimum communication cost when the assumption is removed. This involves finding optimum communication matrix, C^*, the rate allocation for each source and the route through the network which minimize the total communication cost. Note that, in this paper we only consider finding the optimum cost and do not attempt to solve the issue with the complexity of the optimization.

Let C_A be the set of all communication matrices that allow for loss-less reconstruction of the requested sources at each sink. Note that for any $C \in C_A$, at least the requested sources must communicate with each sink. We first pick some $C \in C_A$. It determines the subset of sinks that a source has to transmit its information to. We denote by $E^i(C)$ the set of all paths from source i to the corresponding subset of sinks. From arguments similar to that in [1], the optimum route from the source to these sinks is determined by a spanning tree optimization (minimum Steiner tree). 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 that node i communicates with (determined by C). Mathematically:

$$d_i^*(C) = \min_{\mathcal{P}\in E^i(C)} \sum_{e\in\mathcal{P}} w_e \qquad (10)$$

It is known that the Steiner tree optimization is NP - Complete and hence requires approximate algorithms to solve in practice. Again it is key to note that, immaterial of the rate allocation at the nodes, for a fixed communication matrix, the total cost a bit from source i encounters is $d_i^*(C)$. Therefore the remaining task of finding a rate allocation for each source can be solved independent of the Steiner tree optimization.

For a fixed request matrix T, the entire achievable rate region for loss-less reconstruction of the requested sources depends on the communication matrix. We denote this region by $\mathcal{R}_T(C)$. The optimum rate tuple that minimizes the communication cost, $\{R_1(C), R_1(C) \ldots R_N(C)\}$, must therefore lie in $R_T(C)$. Hence, having found the optimum weights, $d_i^*(C)$, the optimum rate tuple can be determined as follows:

$$\min_{\{R_i(C)\}_i^N} \sum_{i=1}^{N} R_i(C) d_i^*(C)$$

subject to,

$$\{R_1(C), R_1(C) \ldots R_N(C)\} \in R_T(C) \tag{11}$$

and we denote the optimum cost by $D^*(C)$.

A single letter characterization of the entire region, $R_T(C)$, is still unknown. Han and Kobayashi provide a partial achievable region for the problem and we use their results to characterize $R_T(C)$. We denote the subset of encoders that communicate with sink j by $\Sigma_j(C) = \{i : C_{ij} = 1\}$. Let $U_1, U_2 \ldots U_N$ be auxiliary random variables on finite sets satisfying the following Markovian properties:

- The random variables $U_{\Sigma_j(C)}$ are conditionally independent given $X_{\Sigma_j(C)}$ $\forall j$.
- For each $i \in \Sigma_j(C)$, the conditional distribution of U_i given $X_{\Sigma_j(C)}$ depends on X_i alone.

For each such set of axillary random variables $U = \{U_1, U_2 \ldots U_N\}$, define $R_T(C, U)$ as the set of all the rate tuples, $\{R_1, R_2 \ldots R_N\}$, satisfying the following constraints. For each $j \in \{1 \ldots N\}$ and all $S \subseteq \Sigma_j(C)$:

$$\sum_{i \in S} R_i \geq I(X_S, U_S | U_{\Sigma_j(C) \setminus S}) + \psi_S(C) \tag{12}$$

where

$$\psi_S(C) = \sum_{i \in S} \max_{k : T_{ik} = 1} H(X_i | U_{\Sigma_k(C)})$$

Then $R_T(C)$ is given by the closure of $R_T(C, U)$ over all auxiliary random variables U satisfying the above properties:

$$R_T(C) = \cup_U R_T(C, U) \tag{13}$$

Note that due to its dependence on auxiliary random variables an explicit characterization of $R_T(C)$ is hard, except for some particular communication matrices. If we choose the communication matrix to be equal to T and the auxiliary random variables $U_i = X_i \ \forall i$, equation (12) becomes the usual Slepian - Wolf constraints given in [1]. i.e. for each $j \in \{1 \ldots N\}$ and all $S \subseteq V^j$:

$$\sum_{i \in S} R_i \geq H(X_S | X_{V^j \setminus S}) \tag{14}$$

The optimum communication cost, D^*, is the minimum over all $C \in C_A$:

$$D^* = \min_{C \in C_A} D^*(C) \tag{15}$$

and the corresponding communication matrix that minimizes is C^*. Note that the number of admissible communication matrices (i.e. the cardinality of the set C_A), usually grows of the order of $2^{N \times M}$. Hence, $2^{N \times M}$ $D^*(C)$ are to be computed to find the optimum. It is clear from the above result that, if the sources are allowed to communicate with any sink, finding the rate allocation for each source and the transmission structure through the network cannot, in general, be decoupled.

VI. Conclusion and Future work

In this paper we addressed the problem of optimizing the communication cost for a general network with multiple sinks and arbitrary network demands. We showed that the assumption under which this problem was solved earlier is highly limiting. We also gave examples of scenarios where the gain by removing the assumption is unbounded. We proposed a theoretical solution to the problem using results from multi-terminal information theory. Future research directions include design of low complexity algorithms to find close to optimal solutions. Also, this work takes a step closer towards finding the best joint coding-routing scheme for general networks.

References

[1] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, "Networked slepian-wolf: Theory, algorithms and scaling laws," *IEEE Trans. on Information Theory*, vol. 51, no. 12, pp. 4057–4073, 2005.

[2] J. Liu, M. Adler, D. Towsley, and C. Zhang, "On optimal communication cost for gathering correlated data through wireless sensor networks," in *in Proceedings of the 12th annual international conference on Mobile computing and networking. ACM*, 2006.

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

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

[5] T. S. Han and K. Kobayashi, "A unified achievable rate region for a general class of multiterminal source coding systems," *IEEE Trans. on Information Theory*, vol. IT-26, pp. 277–288, May 1980.

[6] I. Csiszar and J. Korner, "Towards a general theory of source networks," *IEEE Trans. on Information Theory*, vol. IT-26, pp. 155–165, Mar 1980.

[7] H. Luo, Y. Liu, and S. K. Das, "Routing correlated data in wireless sensor networks: A survey," *IEEE Network*, vol. 21, no. 6, pp. 40–47, 2007.

[8] B. L. K. Yuen and B. Liang, "Distributed data gathering in multi-sink sensor networks with correlated sources," *Lecture Notes in Computer Science*, vol. 3976, p. 868, 2006.

[9] R. Cristescu, B. Beferull-lozano, M. Vetterli, and R. Wattenhofer, "Network correlated data gathering with explicit communication: Np-completeness and algorithms," *IEEE/ACM Transactions on Networking*, vol. 14, pp. 41–54, 2006.

[10] S. Pattem, B. Krishnamachari, and G. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," *ACM Trans. on Sensor Networks*, vol. 4, no. 4, 2008.

[11] A. Ramamoorthy, K. Jain, P. Chou, and M. Effros, "Separating distributed source coding from network coding," *IEEE Trans. on Information Theory*, vol. 52, no. 6, pp. 2785–2795, 2006.

[12] T. Ho, M. Medard, M. Effros, and R. Koetter, "Network coding for correlated sources," in *in Proceeding of CISS*, 2004.