Progressive Joint Coding, Estimation and Transmission Censoring in Energy-Centric Wireless Data Gathering Networks

Sheryl L. Howard and Paul Flikkema
EE Department
Northern Arizona University
Flagstaff, AZ 86001
sheryl.howard,paul.flikkema@nau.edu

Abstract

Energy-constrained wireless sensor networks often are designed to measure a spatio-temporal process that is correlated in space, time, or both. The goal of these data-gathering networks is a description of the process that provides the required fidelity with a minimum expenditure of energy. Our approach combines (1) channel coding and estimation/decision processing of coded messages for in-network data censoring with (2) estimation of the encoded and censored information at a fusion center where energy is plentiful. Nodes examine their own messages together with messages from preceding nodes, and compute the fidelity of the estimate at the next node as a function of data censoring proposals. Our algorithm exploits redundancy of two types: intra-message redundancy from channel coding, and inter-message redundancy due to spatio-temporal correlation of the samples. This redundancy is used to alleviate overall energy consumption and message congestion near the fusion center by allowing relaying nodes to censor messages that might otherwise be forwarded, if those messages can be inferred from other messages, given the correlation model. The effect of censoring on fidelity and energy consumption is characterized, and our censoring algorithm shown to provide significant energy savings.

1 Introduction

Many wireless sensor network applications involve the estimation of a time-varying vector field at a destination node. Sensor samples are forwarded, often via multihop networking, to the destination or fusion center where the samples are processed into an estimate of the field. The estimate is formed under severe resource constraints, the primary constraint being the energy available to each node. This, along with economic pressure to limit node costs, reduces communication reliability. The goal is an estimate with the best precision possible given limited energy and unreliable communication.

1-4244-2575-4/08/$20.00

This paper views the problem from a distributed signal processing perspective, and integrates tools normally separated into distinct layers of the network protocol stack.

Source coding, and, for networks, distributed source coding (see the review [1]) is a presentation- or application-layer technique that approaches the problem as one of minimizing uncoded communication rates. Channel coding, an essential link layer technique, plays a role as well due to the poor link quality in practical sensor networks [2]. It is well-known that the separation theorem justifying independent design of source and channel codes breaks down (i) for the small block sizes often found in sensor networks and (ii) when multiple access (not just a collection of point-to-point links) exists. Clearly, source-channel coding should play a role. This is true in both distributed fusion problems (e.g., the consensus problem, where all nodes form an estimate of the entire field) and the centralized fusion problem—where samples are funneled to the destination—considered here. This latter problem is not a special case of the distributed fusion problem, since the distributed fusion problem can be solved with gossiping algorithms [3] leading to uniform traffic loads across the network. In centralized fusion without source coding, traffic, and hence congestion, rapidly increases as the samples merge on their way to the fusion center. The problem is not limited to the highest and lowest layers, but directly challenges medium-access design as well. A recent paper [4] treats congestion at the MAC layer in isolation; our cross-layer approach tackles the root cause of congestion and energy efficiency.

This paper is organized as follows. Section 2 discusses our philosophy of reducing energy consumption in the network. Section 3 presents related work, and Section 4 presents the source model used for this paper. Section 5 discusses combined source-channel (SC) decoding/estimation at the fusion center, and Section 6 proposes censoring within the network. Section 7 presents the simulation model used in this paper. Section 8 discusses a relay network, and shows results for combined SC decoding/estimation without transmission censoring in a relay network. Section 9 presents results for combined SC decoding/estimation with censoring. Section 10 concludes this paper.

2 Approach

We assume that a prior model for the measured field exists in the form of a multivariate distribution over a defined set of samples in space and/or time. This model represents knowledge about the field that can be used in a Bayesian inference model, often implemented as a simple MAP estimator. This assumption is completely general in that the prior model may be uninformative. More generally, bounds can usually be established with high confidence in the absence of a completely specified distribution. It was shown in [5] how the channel-coded samples received at the destination form a SC product codeword (formed from stacking the received samples) that combines deterministic parity constraints of the channel codes with probabilistic constraints encoding the correlation of the prior model for the measured field.

This paper extends [5] to every node in the network serving as a relay. Relaying nodes can exploit the redundancy in the information they receive (and collect themselves) to make informed decisions on whether to forward samples, or *censor* a portion of them. Our algorithm spans layers since it employs SC decoding to evaluate proposed censoring decisions. Censoring is a form of distributed adaptive signal processing, as the decision is a function of the model, the data, and the targeted fidelity. From a coding perspective, the algorithm progressively punctures a dynamic SC product codeword as it is shuttled to the destination. The resulting benefits are two-fold. First, the amount of data transmitted, and hence energy consumption, is reduced. Congestion is also relieved: as the samples travel to the destination, more redundant information may be available, so transmission rates are naturally throttled where most needed.

3 Related Work

Several researchers have addressed the question of what tools to use in getting the required information from a sensor network with the minimum expenditure of energy. The problem is complex because it touches almost every traditional layer of wireless networking. Our adaptive SC censoring algorithm shares a goal of estimate and forward (EF) techniques (see [6] and the references therein) studied in cooperative communication: maximizing the amount of information recoved at the destination. However, our approach is different: we take into account the limited capabilities of low-cost radios for which sophisticated physical-layer strategies (such as maximal-ratio combining detection) are not possible, and our goal is energy efficiency, not capacity. In our strategy a node does not relay in the sense of cooperative communication; it instead makes a decision on what information to send—if any—based in part on received multicasts from neighboring nodes. Application-layer Bayesian inference techniques have been explored in [7, 8] and [9]; [10] describes a technique that combines channel coding and opportunistic listening. Our approach also couples the application and logical link layers in that we exploit prior knowledge of the sensed environment jointly with channel coding.

An error recovery approach in wireless sensor networks was proposed in [11] that progressively increases the code rate as data nears the fusion center, reflecting the fact that the overall source-destination channel is a composition of single-hop channels. Their criterion is throughput instead of energy efficiency, and the proposed algorithm treats only channel coding. Related work includes correlated data gathering [12] and censoring in detection [13].

4 Model

Many wireless sensor network (WSN) applications, particularly environmental monitoring, involve multi-node deployments that measure multiple environmental processes correlated in space and time. The goal is a faithful representation of the data with respect to a model at minimum energy cost. In some applications (e.g., camera networks), bandwidth may also be limited. However, in this paper we assume the low data rates found in many environmental sensing applications, and neglect bandwidth efficiency.

Let x_i be a sample computed from some physical process measured by a transducer of a sensor node. Here i is a general index subsuming sensor node, transducer, and time indices. The measurement m_i of x_i is corrupted by transducer noise and then quantized into the information word u_i, which is encoded into a codeword v_i. The codeword is forwarded to the destination, which receives $y_i = v_i + z_i$ where z_i is a binary error vector with addition mod 2.

The entire collection of data samples is x, and similarly for m, u, v, and y. From Bayes' theorem, we have

$$\begin{aligned} p(x,m,u,v|y) &\propto p(y|x,m,u,v)p(x,m,u,v); \quad (1) \\ &\propto p(y|v)p(v|u)p(u|m)p(m|x)p(x), \end{aligned}$$

where $p(x)$ is the prior information included in the source model as $p(v,x)$ or $p(u)$.

But the mapping of the per-measurement channel code from m to v is deterministic, so we can use

$$p(x,m,u,v|y) \propto p(y|v)p(v|x)p(x). \quad (2)$$

The model can be simplified depending on the sensing scenario. For example, in most cases of interest transducer noise and channel error processes are independent.

Due to space limitations, this paper emphasizes the effects of channel errors and hence focuses on inference of the collection of quantized samples u. We thus re-define u to incorporate the properties of the source statistics and measurement noises in a symbol- or bit-level probabilistic model. Then the MAP estimate of the transmitted SC product codeword is

$$\hat{v} = \arg\max_{v} p(v|y) = \arg\max_{v} p(y|v)p(v). \tag{3}$$

There is a direct mapping from $\hat{v}$ to $\hat{u}$; the most likely measurements $\hat{m}$ are then found directly from $\hat{u}$.

5 Processing at the Fusion Center

The fusion center receives y and estimates the samples v. In the most general, computationally expensive case, it computes the posterior probability law according to equation (3), or an approximation thereof. For reduced complexity, an iterative inference algorithm coupling source and channel estimator/decoders [5] can be used to compute a point estimate. Iterative source-channel decoders were first explored in [14], and later for a single first-order correlated source over an AWGN (additive white Gaussian noise) channel in [15].

Both the full Bayesian search and the iterative algorithm incorporate prior knowledge of the source statistics, channel code structure, network reliability, as well as missing (censored) samples. Note that the symbols of censored codewords are treated as completely ambigious.

The iterative decoder is comprised of component decoders for both the channel codes and the source model. Probabilities are exchanged between soft-decision channel decoders and source decoder, similar to turbo decoding [16], allowing for multiple iterations of decoding to improve performance. The source model decoder takes in the *a posteriori* channel decoder output, and converts bit probabilities $p(v_{ij}|y_{ij})$ for bit j of codeword v_i to symbol probabilities on m_i for its *a priori* input. These *a priori* symbol probabilities $p(m_i)$ are further combined to form product symbol probabilities $p(\mathcal{M}_{/\mathrm{k}})$, where $\mathcal{M}_{/\mathrm{k}}$ is the combined message word for all samples, excluding k. These product symbol probabilities are assumed symmetric for all k, i.e., invariant with respect to measurement ordering. Finally, the source decoder combines these product symbol probabilities $p(\mathcal{M}_{/\mathrm{k}})$ with conditional probabilities $p(m_i = M|\mathcal{M}_{/\mathrm{k}})$ from the source model as

$$p(m_i = M) = \sum_{\mathcal{M}_{/\mathrm{k}}} p(m_i = M|\mathcal{M}_{/\mathrm{k}})p(\mathcal{M}_{/\mathrm{k}}). \tag{4}$$

The source decoder then outputs these (hopefully) improved symbol probabilities $p(m_i)$, which are then converted back to bit probabilities $p(v_{ij})$ for *a priori* input to each channel decoder i. A reduced set of conditional bitwise probabilities could be used instead of symbol probabilities, but would introduce further suboptimality. This simplification was used in [5], which did not include results for a source decoder using symbol probabilities.

In contrast, the full Bayesian search algorithm is non-iterative. This decoder chooses the most-likely product codeword $\hat{v}$ with the MAP decision of (3), and is termed the MAP decoder. The *a priori* product codeword probabilities $p(v)$ may be found from the source model. The MAP decoder is used as a performance benchmark, as it provides optimal decoding results. The iterative decoder is slightly suboptimal, as it is composed of separate decoders for channel codes and source model. The ability to iteratively improve its performance allows the iterative decoder to approach the optimal performance of the MAP decoder.

6 In-Network Processing

The potential for recovering censored samples motivates the question: What information should be sent to the fusion center? More specifically, given a limited energy supply in each node of the network, what samples should any node forward toward the fusion center, and for what samples should transmission be censored?

We propose an adaptive censoring algorithm: it weighs the value of the data in terms of its "informativeness" in the context of the prior model and transmission cost. Large excursions from the model are clearly more useful in characterizing the process than samples near expected values, and their transmission is a better use of energy.

At a receiving node, the censoring decision of a sender is viewed as a puncturing of the SC product codeword. Individual samples are encoded using an (n, k, d) channel code. If a node has N potential samples to send but L are sent (and thus $N - L$ are censored), then the formal source-product code rate is $\frac{kN}{nL}$, assuming the decoder infers all N samples. Note that the code rate is not fully descriptive since some samples are more valuable. For example, communication of two product codewords of the same rate corresponding to two censoring decisions have the same energy cost, but may not be equally informative. For this reason the censoring decision must rely on the data itself in addition to the rate.

The algorithm begins when a node receives a noisy product codeword y. The censoring decisions made by the transmitters are known, since it is assumed that every measurement is associated with a unique label (for example, the node/transducer ID and time). In practice, the label (suitably defined, compressed, and code-protected) can apply to all the information in the transmitted product codeword.

The node performs the inference computation of section 5 on the received product codeword to generate an estimate of the samples. Once computed, this *a posteriori* distribution becomes the prior distribution for the node's censoring algorithm. While the complete APP distribution $p(v|y, c)$ could be used, in this paper computational energy cost is limited by using the singleton MAP decision $v_0 = \arg\max_v p(v|y, c)$ as an approximation.

The censoring algorithm has two steps: (1) computing an *estimate* of the fidelity and energy cost of a set of censoring proposals, and (2) choosing one proposal as the censoring *decision* based on a *censoring policy*.

How should the estimate be computed? Here is where the algorithm's model of the network becomes important. The simplest approach is to use the data as-is; this implicitly assumes that the data communicated as a result of the estimation/decision algorithm will arrive without error. In our algorithm, the estimate is *predictive* in the sense that the node accounts for the unreliability of channel to the next node, and beyond that, to the fusion center. In this paper, we assume a model of local (one-hop) collaboration: the transmitting node requires only knowledge of the quality of the hop to the next node in the data gathering tree, and assumes that the next node will take appropriate action based on its local knowledge.

The overall goal of the censoring policy is to decide which codewords in the product codeword to censor. Each censoring proposal c_l, $\forall l = 1, \ldots, 2^N$, censors a unique pattern of codewords in v_0. The posterior distribution

$$p(v|y, c_l) = \frac{p(y|v_0, c_l)p(v)}{p(y, c_l)} \tag{5}$$

is computed by the node for each proposal c_l, where $p(v)$ is the prior distribution over the information represented by the product codeword.

The subset of evaluated proposals is determined as part of the censoring policy. Another aspect of the policy is the criterion for ranking the censoring proposals with respect to a summary measure of the distribution. One of many possible measures is $\mathrm{Cov}(V_l, V) = \mathrm{E}(V_l - V)^2$, where V is the random variable associated with the distribution $p(v|y, c)$ and V_l is the random variable associated with the posterior conditional distribution (5). Alternatively, and more simply, we can consider the variance of the posterior about the MAP estimate, $\mathrm{Var}\ V_l = \mathrm{E}(V_l - v_0)^2$. In this paper, we consider the simplest measure, computed as the squared difference $(\hat{v}_l - v_0)^2$ between the MAP estimate of the node's in-hand data v_0 and the point estimate $\hat{v}_l = \arg\max p(v|y, c_l)$ from (5) for each proposal c_l.

The chosen proposal is the decision; the non-censored samples are encoded and transmitted to the next node. The receiving node repeats the processing, computing $p(v|y, c)$ (or an approximation) for its received product codeword and running the censoring algorithm.

7 Simulation Experiments

A symmetric multivariate Gaussian source model is used for the sensed process x. The *a priori* information uncertainty σ_x^2 is assumed uniform; the covariance matrix for the source model has diagonal elements $\sigma_{x_i}^2 = \Sigma_{ii} = \sigma_x^2$ for all i. The cross-correlation r is found as $r_{ij} = \Sigma_{ij}/\Sigma_{ii}$ for all $i \neq j$. The information mean μ_x is known *a priori*. In all simulations, $\sigma_x^2 = 0.2$, $\mu_x = 10$, and $r_{ij} = 0.9$.

The transducer noise t is assumed Gaussian; thus $p(m_i|x_i)$ is also Gaussian with mean x_i and variance σ_t^2. The noisy measurement $m_i = x_i + t_i$ is then quantized to a data sequence u_i; R_{U_i} is the range of the analog-to-digital converter (ADC) input. Encoding u_i results in the codeword v_i. Due to the deterministic mapping from m to v, the conditional codeword probability $p(v_i|x_i)$ is given by $p(v_i|x_i) = p(v_i = V_i|x_i) = \int_{R_{U_i}} f_m(s)ds$, where $f_m(\cdot) \sim N(x_i, \sigma_t^2)$. In all simulation results, σ_t^2=0 to clearly show the effect of channel noise.

The channel code used for encoding each sensor measurement is a $(n, k, d_{\min})$ = (7,4,3) Hamming code. This low-complexity, short block-length code was used to minimize sensing node computational load and assess usefulness of the approach.

All simulations model a single hop as a binary symmetric channel (BSC) with crossover probability ρ. The BSC models the hard-decision output of most available single-chip radio transceivers for sensor networks.

The fusion center decodes the received noisy product codeword y. Both an iterative SC decoder, described in [5], and a MAP decoder, also described in [5], are examined in the simulations. The individual channel decoders used in the iterative decoder are trellis decoders implementing the BCJR [17] algorithm on the (7,4,3) Hamming parity-check trellis. Our figure of merit is the mean-squared error (MSE) $E[(m_i - \hat{m}_i)^2]$ between the measurements and the estimated measurements at the hub. Each MSE value is calculated using 50,000 Monte Carlo simulations.

8 Relay Network

In a relay or line network, the sensor nodes form a relay chain or line that forwards messages to the fusion center. Each node takes its own measurements and transmits them along with messages it receives from the preceding node.

8.1 Relay Network without Transmission Censoring

A relay network with N=3 forwarding nodes and a fusion center is considered. Each node takes M=2 measurements of its own, and forwards those as well as any received messages on to the next node. The fusion center receives a correlated product codeword of size 6 x 7, where the channel code length is 7. This section examines performance when each forwarding node simply forwards all measurements to the next node, with no transmission censoring, and only the destination node performs any decoding. This provides us with a baseline performance for maximal energy cost.

8.2 Results

Each relay node encodes its own data with a (7,4,3) Hamming code. MSE results for MAP decoding of the SC product code, the iterative SC decoder after 5 decoding iterations, and channel decoding alone (with ML decoding of each received codeword) are compared to an uncoded system with hard-decision decoding. Two different iterative decoders are examined, one using conditional symbol probabilities, and a suboptimal version using bitwise-only conditional probabilities.

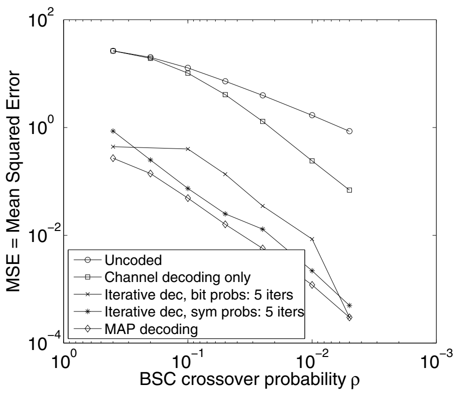

Figure 1. MSE vs ρ for 3-node relay network without transmission censoring.

The performance improvement possible with either the iterative or MAP SC decoders on the relay network is significant, as shown in Figure 1. The bottom curve (-◇-) displays the performance of the optimum MAP SC decoder. The iterative SC decoder with conditional symbol probabilities (-∗-) shows some loss due to the suboptimality of separate decoders for source and channel. A lower complexity approach [5] uses bitwise-only conditional probabilities in the iterative source decoder, eliminating the need for word assembly and disassembly at each iteration. This approach (-x-) shows noticeable performance loss compared to the iterative decoder with conditional symbol probabilities. All SC decoders perform much better than channel decoding alone (-□-) or the uncoded system (-o-).

9 Relay Network: Transmission Censoring

In this section we assess the performance of inference-based censoring in the relay network. With an eye toward minimum energy cost for computations, we consider a highly constrained censoring policy with a simple decision criterion. Each relay node uses a self-only censoring policy, where it considers proposals that censor only subsets of its own samples; all data received from other nodes is forwarded. Proposals are considered in order of energy cost from least to most, so the proposal that censors all of the node's own samples is considered first. Each proposal c_l is evaluated by computing the predicted codeword $\hat{v}$ generated by the receiving node's decoder/estimator when the proposal's associated product codeword v_l is transmitted. If $\hat{v} = v^*$ (where v^* is the node's true, in-hand product codeword), the search is terminated and the v_l is sent. To minimize complexity, the full MAP prediction is not computed; instead, the node uses one iteration of the source decoder. Since the effect of channel errors is considered, but the gain of the channel code is not, our results are pessimistic but simplicity is gained.

9.1 Results

A 3-node relay network is examined; each relay node encodes its own data with a (7,4,3) Hamming code, and each hop modeled as a BSC with crossover probability ρ.

Figure 2 compares MSE performance with and without censoring for two inference algorithms at the fusion center: the MAP SC decoder, and the iterative SC decoder using symbol probabilities. As expected, MAP decoding without censoring (-◇-) provides an upper performance bound, but MAP decoding *with* censoring (-□-) is surprisingly close. At most channel reliabilities, iterative decoding without censoring (-+-) is better than with censoring (-◁-), as expected. However, when using the iterative decoder in very poor channels (ρ=0.4), it is better to censor. It is well-known that most channel codes perform worse than uncoded communication in poor channels. This is exactly what is occurring here: the bit LLRs of censored samples input to the channel decoder are set to zero. The resulting channel decoder output LLRs are also zeros in the first iteration; in effect, no decoding is performed.

The effect of censoring on both MSE performance and energy efficiency is shown in Figure 3, where the MSE is plotted against the fraction of messages censored for different decoders at $\rho = .05$. These decoders include the MAP SC decoder, the iterative SC decoder using bitwise conditional probabilities,and the iterative SC decoder using conditional symbol probabilities. The uncoded case is also shown for comparison. MSE results without censoring are also shown; these fall on the y-axis, as the fraction of censored messages is zero for no censoring.

For the optimum MAP SC decoder at the fusion center, the MSE performance loss due to censoring is negligible, with a huge gain in energy efficiency—80% of the messages are censored. As expected, the iterative SC decoder has poorer fidelity than the MAP decoder; but its energy efficiency is similar despite lower complexity. Note that the MAP SC decoder is likely to be feasible in terms of computational load and energy consumption.

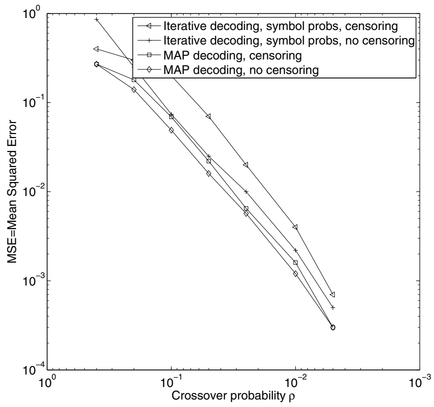

Figure 2. MSE vs ρ for 3-node relay network with transmission censoring.

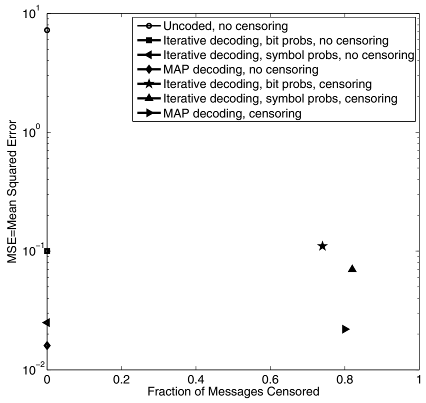

Figure 3. MSE vs fraction of messages censored for relay network.

10 Conclusion

In a range of important applications, wireless sensor networks are tasked to gather correlated data with minimum energy expenditure. We have proposed an approach that integrates channel coding of samples and an adaptive strategy of data sample censoring. The in-network progressive prediction of the fidelity and energy cost of censoring proposals exploits decoding/estimation using intra-codeword redundancy of channel coding and inter-codeword redundancy of the information, and allows the explicit trade-off of energy cost and fidelity that can be chosen in the context of the data gathering application. A related benefit of censoring is the management of congestion as data flows toward the fusion center. The significant performance improvement and energy-efficiency of this approach was demonstrated in simulations.

Acknowledgment

The work of the first author (Howard) was supported by an NAU Intramural Grant. The work of the second author (Flikkema) was supported in part by NSF grant CNS-0540414 and an NAU ERDENE grant.

References

[1] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks", *IEEE Signal Proc. Mag.*, pp. 80-94, Sep. 2004.

[2] Y. Zhao and R. Govindan, "Understanding packet delivery performance in dense wireless sensor networks", in *Proc. Sensys '03*, 2003.

[3] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus", in *Proc. IPSN '05*, 2005.

[4] G.-S. Ahn *et al.*, "Funneling-MAC: A localized, sink-oriented MAC for boosting fidelity in sensor networks", in *Proc. Sensys '06*, 2006.

[5] S. Howard and P. Flikkema, "Integrated source-channel decoding for correlated data-gathering sensor networks", *Proc. IEEE WCNC '08*, 2008.

[6] A. Chakrabarti, E. Erkip, A. Sabharwal, and B. Aazhang, "Code designs for cooperative communication", *IEEE Signal Proc. Mag.*, pp. 16-26, Sep. 2007.

[7] S. Mukhopadhyay, D. Panigrahi, and S. Dey, "Model based error correction for wireless sensor networks", in *Proc. IEEE SECON 2004*, 2004, pp. 575-584.

[8] G. Hartl and B. Li, "infer: A Bayesian inference approach towards energy efficient data collection in dense sensor networks, *Proc. IEEE ICDCS 2005*, 2005, pp. 371-380.

[9] A. Silberstein *et al.*, "Suppression and failures in sensor networks: A Bayesian approach", in *Proc. VLDB 2007*, 2007.

[10] H. Dubois-Ferrire, D. Estrin, and M. Vetterli, "Packet combining in sensor networks", in *SenSys 2005*, 2005.

[11] S. Qaisar and H. Radha, "OPERA: An optimal progressive error recovery algorithm for wireless sensor networks", in *Proc. IEEE SECON '07*, 2007.

[12] R. Cristescu and B. Beferull-Lozano, "Lossy network correlated data gathering with high-resolution coding", in *Proc. IPSN '05*, 2005.

[13] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks", *IEEE Trans. Wireless Comm.*, vol. 4, no. 6, pp. 2668-2673, Nov. 2005.

[14] N. Görtz, "On the iterative approximation of optimal joint source-channel decoding", *IEEE Journal Sel. Areas in Comm.*, vol. 14, pp. 1662-1670, Sep. 2001.

[15] J. Hagenauer and N. Görtz, "The turbo principle in joint source-channel coding", in *Proc. IEEE Inform. Theory Workshop (ITW) '03*, 2003, pp. 275-278.

[16] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near optimum error correcting and decoding: Turbo codes", *IEEE Trans. Comm.*, vol. 44, pp. 1261-1271, Oct. 1996.

[17] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284-287, Mar. 1974.