Map-based Compressive Sensing Model for Wireless Sensor Network Architecture, A Starting Point

Mohammadreza Mahmudimanesh, Abdelmajid Khelil*, Nasser Yazdani

University of Tehran, Technical University of Darmstadt, University of Tehran
m.mahmoudi@ece.ut.ac.ir, khelil@informatik.tu-darmstadt.de, yazdani@ut.ac.ir

Abstract. Sub-Nyquist sampling techniques for Wireless Sensor Networks (WSN) are gaining increasing attention as an alternative method to capture natural events with desired quality while minimizing the number of active sensor nodes. Among those techniques, Compressive Sensing (CS) approaches are of special interest, because of their mathematically concrete foundations and efficient implementations. We describe how the geometrical representation of the sampling problem can influence the effectiveness and efficiency of CS algorithms. In this paper we introduce a Map-based model which exploits redundancy attributes of signals recorded from natural events to achieve an optimal representation of the signal.

Keywords: Wireless Sensor Networks, Compressive Sensing, Map-based WSN, WSN Architecture.

1 Introduction

A Wireless Sensor Network (WSN) is an instance of a distributed sensing network with a field of sensor nodes. But a very important constraint, differentiates a WSN from other similar configurations of sensing networks. WSN nodes usually have a very limited source of power. Every node must try to conserve as much energy as it can to extend the whole lifetime of the WSN. This energy is mostly consumed during sensing and multihop transport of the sensed data to a base station called the sink [2]. During the operation of a WSN usually a low quality of information (QoI) [11] is requested by the WSN user to make a decision about the occurrence of a specific event. Therefore, it is wiser to keep down as many nodes as possible to conserve more energy while satisfying user's needs [3]. The main responsibility of a WSN is to monitor the physical parameters of a natural operational environment (such as a desert and jungle). One attribute of signal representation of natural events (such as environment temperature and humidity) is that they mostly have very smooth changes and gradients over a plane. Compressive Sensing (CS) [4-7] is a novel sampling approach which tries to exploit compressibility of signals in order to reduce the

* Research supported in part by DFG GRK 1362 (TUD GKmM)

minimum samples required to reconstruct the whole signal. CS offers a mathematically concrete method to capture only *m* samples of all *n* available samples of a signal (where $m \ll n$) and then exactly recover the original signal with an overwhelming probability. It exploits the compressibility attribute of a signal to wisely select only those samples which are important in signal reconstruction. We argue that based on the nature of environment and the geometrical distribution of the specific signal, we can find an arrangement that maximizes CS recovery performance and hence improves its quality.

The remainder of this paper is organized as follows. Section 2 gives a general view of our system model that we use throughout this paper. in Section 3 we summarize the mathematical basics of CS in general with special attention to its applications in WSN. In Section 4, based on our formulations in Sections 3, we discuss how selecting a suitable geometrical framework can improve the performance of compressive wireless sampling recovery. Section 4 presents an evaluation of the model, and Section 5 refers to some related works. Finally, we conclude with a summary and a brief description of future works.

2 System Model

Our WSN scenario consists of a network that is composed of *n* stationary resource-constraint sensor nodes (SN) and a static resource-rich sink. Commonly, WSNs are built utilizing hundreds to thousands of cheap SNs. For a WSN consisting of *n* SNs, there are $n!$ possible vectors of size *n* which can describe the environment as reported by sensors, at any instance of time. Here we focus on the spatial sampling of the environment using the SNs at a fixed instance of time. The main idea of our proposed model is to find a mapping between SNs and sensed vector elements, that results the most compressible signal vector. We will formally define *compressibility* in the next section, but informally one can think of a compressible vector, as a vector that may have redundant data. The *n* SNs sample the physical signal *f* of interest, which we assume to be compressible. A data sample is characterized by the ID and the location of the SN as well as the sensor reading and its timestamp. We assume that SNs know their own geographic position. The clocks of SNs are synchronized, e.g., via GPS or any efficient synchronization protocol [12]. Positioning information is required to find the optimal spatial sampling method.

3 Compressive Sensing: Mathematical Basics and applicability to WSN

Suppose a discrete signal *f* of *n* samples, which can be mathematically represented as a discrete-time *n*-dimensional vector signal $f \in \mathbb{R}^n$. CS is interested in *low-rate sampling*, in which the number *m* of available measurements is smaller than the dimension *n* of signal *f* . The vector *f* is in fact a discretely sampled version of a continuous signal of real environment. The sampling rate is either dictated by Nyquist sampling rate, or desired reconstruction resolution or quality. Therefore, when $m<n$, we are indeed faced to a sub-Nyquist sampling situation. Reconstructing the actual signal $f \in \mathbb{R}^n$ from its sub-sampled version $y \in \mathbb{R}^m$ may seem infeasible. CS promises that if the signal satisfies some preconditions, it can be accurately or even exactly recovered from fewer *compressed samples*.

Definition 1: Sensing basis Φ is the basis in which signal is discretely represented. The classical case is when the sampling basis is delta Dirac waveforms, $\phi_k(t) = \delta(t-k)$ that results in signal vector *y* consisting of samples $y_k = \langle f, \phi_k \rangle$, $k = 1, \ldots, n$.

Definition 2: Sensing matrix is an $m \times n$ matrix *A* which is used to select an under-sampled edition of signal $f \in \mathbb{R}^n$, namely $y \in \mathbb{R}^m$, by the matrix multiplication: $y = Af$.

3.1 Sparse and Compressible Signals

Suppose we have a vector $f \in \mathbb{R}^n$ which we expand it in an orthonormal basis (such as a wavelet or Fourier basis) as follows:

$$f(t) = \sum_{k=1}^{n} x_k \psi_k(t) \tag{1}$$

where *x* is the coefficients vector under Ψ-transform of f, $x_k = \langle f, \psi_k \rangle$. More conveniently we can express *f* as Ψ*x,* where Ψ is the $n \times n$ transformation matrix with $[\psi_1, \psi_2, \ldots, \psi_n]$ as columns. For example if we take Fourier basis as Ψ-basis, then $\psi_k(t) = n^{-1/2} e^{i2k\pi t/n}$ and *x* is the Fourier transform of *f.* Note that because we deal with time-discrete signals, *t* is limited only to integral values between 1 and *n*, $t \in \{1, \ldots, n\}$. Therefore, $\psi_k(t)$ determines the *n* items of the *k*-th column of matrix Ψ, by substituting *t* by 1,2, ... , *n*.

Definition 3: Representation basis Ψ is the basis in which the signal is transformed and represented for final storage or communication purposes. It can also be referred as Ψ-domain. The final purpose, is demanded by the specific application of signal sampling, compression, storage and recovery. For example wavelet-domain is a suitable Ψ-domain for image compression.

Definition 4: *S*-Sparse vector is a vector with only *S* nonzero items. We also call a signal *f*, *S*-Sparse, if its representation (in Ψ) is a *S*-Sparse vector.

Definition 5: Compressible Signal is a signal whose representation vector has many small (near zero) items and only a few relatively large and meaningful items.

In general for a vector v of size n, we define v_S as a S-sparse version of v by setting $n-S$ items of v to zero. By selecting a well-chosen S, the S-sparse vector f_S can be extracted from compressible signal f while maintaining the reconstruction error bellow a certain level. Hereafter, we may use "sparsity" and "compressibility" interchangeably. In the ongoing sections, sometimes by "sparsity" we mean that many elements of vector representation of signal are so small (near zero compared to other elements of vector), that we can easily neglect them.

3.2 Incoherent Sparse Sampling and Recovery

Suppose we are given a pair (Φ,Ψ) of orthonormal bases for vectors in $\mathbb{R}^n$. The first basis Φ is used for sampling the signal f in time or space domain, and the second is used to represent f in frequency domain. We define coherence between these two bases as follows:

Definition 6: Coherence. The coherence between the sensing basis Φ and the representation (frequency) basis Ψ is defined as:

$$\mu(\Phi,\Psi) = \sqrt{n}\max(\langle \phi_k, \psi_j\rangle) \quad \text{over all} \quad 1 \le k, j \le n \tag{2}$$

in other words the coherence represents the largest correlation value between any two elements of Φ and Ψ.

CS is mainly concerned with low coherence pairs of sampling and representation bases. If Φ is the canonical or spike basis of delta Dirac functions ($\phi_k(t) = \delta(t-k)$), and Ψ is the Fourier basis ($\psi_k(t) = n^{-1/2}e^{i2k\pi t/n}$), then it can be shown that $\mu(\Phi,\Psi)=1$ and we have maximal incoherence. The interesting part of CS theory is that if we even select an orthobasis Φ uniformly at random, then with high probability, the coherence between Φ and Ψ is about $\sqrt{2\log n}$. Random waveforms with independent identically distributed (i.i.d.) entries, also have a very low coherence with any fixed orthonormal representation bases Ψ.

Subsampling refers to sampling less than all available measurement. In WSN, this means that among all available n sensor nodes, we only observe a subset of all nodes and collect the data $y_k = \langle f, \phi_k\rangle$, $k \in M$ where $M \subset \{1,\ldots, n\}$ is a subset of cardinality $m < n$. Note that our reduced sensing basis has fewer dimensions, i.e. the signal is projected over fewer basis vectors.

Recovering original signal from these incomplete set of samples is performed by l_1-norm minimization [10]; the proposed reconstruction f^* is given by $f^* = \Psi x^*$, where x^* is the solution to the following convex optimization program:

$$\text{minimize} \, ||\tilde{x}|| \quad \text{subject to} \quad y_k = \langle \phi_k, \Psi\tilde{x}\rangle, \quad \forall k \in M \tag{3}$$

where $\tilde{x} \in \mathbb{R}^n$ and l_1-norm is defined as $||x||_{l_1} := \Sigma_i|x_i|$. In simple words, among all vectors in Ψ-domain which are consistent with the collected data, select the x^* whose

l_1 norm is globally minimum. Then the recovered signal f^* can be calculated from Ψx^*.

Fundamental theorem of CS: Suppose that the Ψ transform x of f in the Ψ-domain is S-sparse. If we select m measurements in the Φ-domain uniformly at random so that

$$m \geq C \cdot \mu^2 (\Phi,\Psi) \cdot S \cdot \log n \qquad (4)$$

for some positive constant C, then the solution to the optimization problem (3) can accurately or even exactly recover the original signal f.

3.3 CS Advantages and its Application in WSN

The following results from above mathematical summary leads us to think of CS as a very interesting and useful tool for sub-Nyquist sensor sampling, specially for WSN sensor reporting:

1) The smaller the coherence, the fewer samples are needed, hence CS emphasizes on low coherence systems. The measurement matrix can be even random or noise-like, because any randomly generated orthonormal basis has low coherence with Ψ transformation matrices such as Fourier or wavelet.
2) The resulting undersampled signal suffers almost no information loss if only about any random set of m coefficients are captured. The number of captured samples m may be far less than the signal dimensionality, if the coherence between sampling and representation bases is a small bounded value.

The CS theorem suggests a concrete and extremely efficient acquisition protocol: first sample the signal in an incoherent domain. Incoherence is the only perquisite; one can simply use random sampling. The sampling matrix needs not to be adaptive to the signal, i.e. this random subsampling technique can apply to any network topology and natural event recording. This makes CS to stand out as one of the best proposed sub-Nyquist sampling techniques for WSN. But as we see in the next section, we can even better utilize CS by considering geometrical attributes of the WSN environment.

4 CS-oriented Map-based (CSM) architecture

Mapping is the core of the CSM model that reorders the vector signal elements in a way that maximizes CS recovery performance. In other words, mapping is the preparatory phase prior to CS sampling and recovery that prepares an alternative representation of sampling and recovery problem. We define mapping in an abstract way that assigns each element of signal vector f one and only one sensed value of the set of SNs readings. In comparison to other Map-based models for WSNs, our model is more abstract and general [9].

Definition 7: Mapping function *M* is a one-to-one function from $\{1,2,\ldots,n\}$ to $\{1,2,\ldots,n\}$ that defines a mapping between sensor readings and signal vector elements :

$$\forall\, k\in\{1,\ldots,n\}\ ,\quad f_k = s_{M(k)} \tag{5}$$

in which *s* is the vector of sensor readings arranged by SN identifier (ID) indexes.

As an example, suppose that *s* is the sensor readings vector ordered by sensor IDs of 3 sensors: (sensor 1, sensed value 13), (sensor 2, sensed value 11), (sensor 3, sensed value 10). Then $s=(13,11,10)$. Now we define the mapping function *M* as: $M(1)=2,\ M(2)=1,\ M(3)=3$. The resulting signal vector *f* under mapping *M* is $f=(s_2, s_1, s_3)=(11,13,10)$.

Mapping gives us the flexibility to choose a more efficient view of the environment. To show that why this flexibility is required, first we need to analyze the fundamental theorem of CS. For a fixed setup of a WSN in CSM model with *N* sensor nodes that uses Φ and Ψ as its sampling and representation bases, we find that $\mu^2(\Phi,\Psi)$ and $\log n$ are constants. Then (4) is reduced to the following linear equation:

$$m \geq C_0 S \tag{6}$$

where $C_0 = C \cdot \mu^2(\Phi,\Psi) \cdot \log n$. Then when all other parameters are fixed (which is of course so, after WSN is deployed), the minimum number of required random samples to recover the vector *f* depends only on the sparsity of the Ψ-transform of *f*. Note that the final recovery is done after applying the inverse mapping M^{-1} to vector *f* and assigning the recovered values to corresponding SNs.

In this model, a centralized or distributed processing entity is responsible for finding the best representation of the world with the most benefit. By representation we mean the mapping *M*, and benefit is in fact the sparsity (or compressibility) of vector *f* under mapping *M*. Here we don't try to present an algorithm to find the most beneficial mapping. Proposing an efficient algorithm can be another challenging topic, and can be even specialized for specific operational environments. In this paper we only try to show that WSNs that use CS approaches and exploit the geometrical properties of their operational environment can reach better CS recovery performance. This emphasizes the importance of considering Map-based WSN models in compressive wireless sensing (CWS) [1]. Next section presents a simple evaluation that runs a trivial exhaustive "efficient mapping" search algorithm to find the optimal mapping.

5 Evaluation of CSM with an Exhaustive Algorithm

Figures 1.a and 2.a show the actual environment in which our test sensor network will operate. In our tests, the environment is the same, but in each test we have randomly distributed the SNs over the operational area. The WSN consists of SNs that capture the temperature at different nodes of a natural area – such as a forest. In practice, such an image doesn't exist prior to sampling. In fact this figure is the ideal target of our WSN. We are going to apply CS sampling and recovery algorithms to this sensor network.

Our CS sampling method, selects randomly m of n sensor readings. Therefore we can even imagine of other $n-m$ sensors, being inactive. This is the situation that may happen in the case of sleep-scheduling energy preservation [2]. In fact the measurement matrix used in our test, was a simple $m \times n$ 0-1 matrix, that was constructed by randomly selecting m rows of $I_{n \times n}$. The Ψ-domain is the classical Fourier-domain, i.e. matrix Ψ is the inverse discrete Fourier transform matrix (IDFT). We have measured the quality of recovered vector, by computing the mean square error (MSE) between recovered and original vector. Because of the random nature of the tests, we ran each test for every WSN consisting of randomly distributed SNs, for one hundred times, and then averaged MSE of each run over all 100 results.

We evaluated and compared two sampling approaches:

a) In the first method we sample the nodes in a raster-like fashion from top to bottom and from left to right.

b) In the second method, we use the knowledge about the current actual temperature map. We examine SNs reported values and try to arrange them in a vector that is most compressible. This requires that for any possible mapping M, we reorder the sensor readings according to M, and compute the sparsity of its Ψ transform. The optimal mapping M^* is the one that generates the most compressible representation of vector f.

We don't try to invent an efficient algorithm that finds the optimal mapping in reasonable amount of time. We used a trivial exhaustive algorithm that searches among all $n!$ possible mappings and chooses the most optimal one. Because of the exhaustive search algorithm inefficiency, we tried the evaluation for a sensor network of only 10 SNs – that is n = 10.

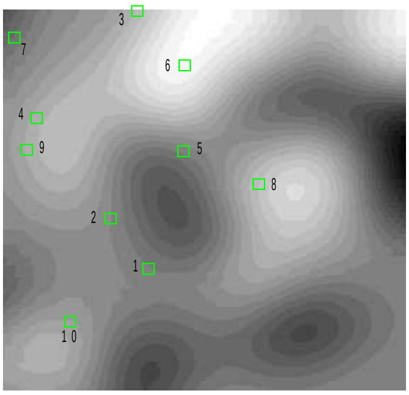

Fig. 1 (a) Randomly distributed SNs of the first test.

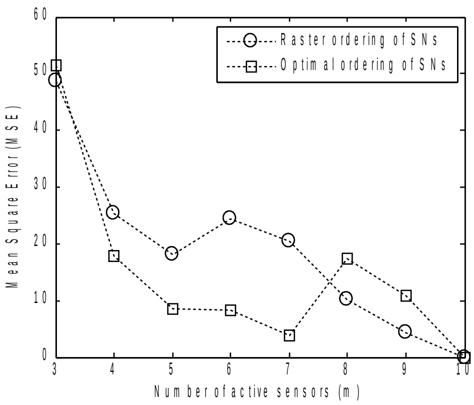

Fig. 1 (b) Average MSE for different cardinalities of randomly selected active sensor nodes**..**

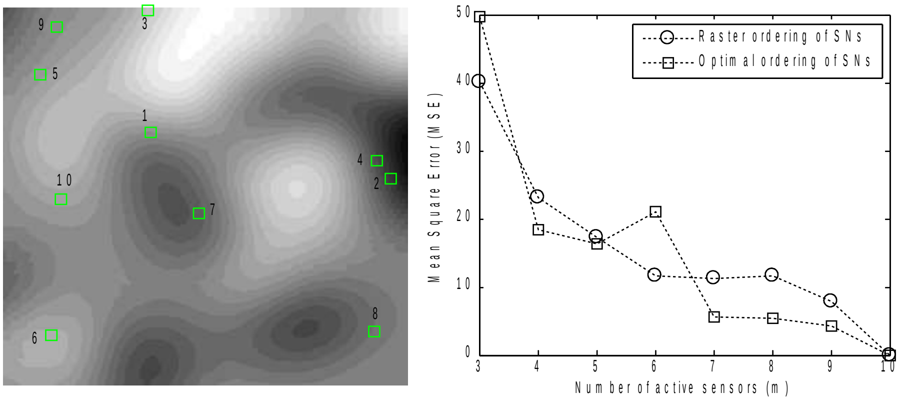

Fig. 2 (a) Randomly distributed SNs of the second test**..**

Fig. 2 (b) Average MSE for different cardinalities of randomly selected active sensor nodes**..**

Figures 1 and 2 depict two runs of our tests for two different positioning of SNs. In each of the figure pairs, Figure a shows the specific positioning of SNs in the test run, and Figure b shows a chart with the corresponding test results. In each of these two sample charts, the dashed line with circle markings determines the average MSE over all 100 CS recoveries of randomly sampled data using first sampling approach (a). Because of random nature of CS sampling, we perform the recovery procedure, one hundred times for each *m* in every test, and then take the average MSE as the actual MSE of that test. The dashed line with square markings shows averaged MSE using second sampling approach (b). For each simulated WSN, we run CS sampling and recovery for different number of active SNs. Because for this case, $m < 3$ can not give us a meaningful result, we started *m* from 3 to *n*=10. As expected, by increasing the number of active SNs, MSE decreases and we have more accurate recovered signal vectors.

As we were expecting, the sampling strategy that tries to achieve the most compressible vector representation of environment signal, leads to lower MSE. However there are some anomalies for some *m*'s, but in most cases, MSE of the tests based on sampling method (b) is lower than that of (a). The anomalies may occur because of randomness nature of problem, or small number of SNs. More satisfying tests can be performed using more SNs. But as we tested for some other testing WSNs, a similar behavior has been observed.

6 Related Work

There is a growing set of literature about specific uses of CS in WSNs, which try to model the WSN to fit the CS framework. Among them Compressive Wireless Sensing (CWS) [1] can be assumed to be pioneering work. Some other early papers have adapted CS for WSN and proposed concrete and efficient CS models for WSNs [13,14]. CWS methods are also tightly related to data compression techniques in

WSN [15-18]. But CWS has more practical results, as it promises efficient data acquisition protocol for distributed sensor networks with high cost of sampling.

Map-based World Model (MWM) for WSN refers to a generalized framework of WSN modelling, that views the world model of a WSN beyond a simple distributed sensing system [8]. As an improvement to previously proposed WSN design models, it also considers the topology of WSN nodes and specific geometrical distribution of the desired signal. Map-based models can be well adopted to exploit this attribute of natural event recording. By defining a suitable mapping that gives us a vector of signal samples with sparsest frequency-domain representation, we can improve the performance of CS signal reconstruction. The mapping procedure, that we have proposed in this work, is rather abstract and more general than MWM, but follows similar ideas.

There are also another set of proposed data acquisition techniques which specially try to exploit the spatial correlation of SNs reportings [19-21]. One can put all series of work under a more general field of study, that deals with distributed sampling in a relatively error-prone and resource-constrained network architecture. In such situations, because of some information loss, we try to reconstruct the signal from a fewer number of available samples [22-24].

7 Conclusion and Future Work

In this paper, we have introduced a new approach to signal acquisition in WSN. We have stated that selecting a suitable geometric framework can utilize better the smoothness attribute of natural signals. By exploiting the spacial attributes of the operational environment, a more compressible vector view of environment signal can be represented. This affects an essential factor of Compressive Sensing (CS) techniques performance, called the *sparsity* of signal in *representation basis*. The sparser or more compressible the representation of a signal, the fewer samples needed to be captured for signal reconstruction using CS algorithms.

However, we didn't try to derive an efficient algorithm to find the optimal mapping, but presented a test on a small scale WSN, which shows that the approach that follows our CS-oriented Map-based (CSM) model, can achieve better results. The lack of an efficient optimal map finding algorithm, limits us to try the model on a WSN consisting of more SNs. Deriving such an efficient algorithm can be a future challenge. But for this small-scale testing WSN, number of tries are high enough to fade away the effect of random sampling.

References

1. W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, “Compressive wireless sensing”, Int. Conf. on Information Processing in Sensor Networks (IPSN), 2006.
2. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, “Wireless sensor networks: a survey”, Comput. Networks 38, 2002.

3. Raghavendra, Cauligi S.; Sivalingam, Krishna M.; Znati, Taieb (Eds.) “Wireless Sensor Networks”, 2nd printing, 2005, ISBN: 978-1-4020-7883-5.
4. E. Candès and M. Wakin, “An introduction to compressive sampling”, IEEE Signal Processing Magazine, 25(2), 2008.
5. E. Candès, “Compressive sampling”, Int. Congress of Mathematics, Madrid, Spain, 2006.
6. R. Baraniuk, “Compressive sensing”, IEEE Signal Processing Magazine, 24(4), 2007.
7. D. Donoho, “Compressed sensing”, IEEE Trans. on Information Theory, 52(4), 2006.
8. A. Khelil, F. K. Shaikh, B. Ayari and N. Suri "MWM: A Map-based World Model for Event-driven Wireless Sensor Networks", In Proc. of The 2nd ACM International Conference on Autonomic Computing and Communication Systems (AUTONOMICS), 2008.
9. A. Khelil, F. K. Shaikh, A. Ali and N. Suri "gMAP: An Efficient Construction of Global Maps for Mobility- Assisted Wireless Sensor Networks", In the Sixth Annual Conference on Wireless On demand Network Systems and Services (WONS), 2009.
10. E. Candès and J. Romberg “Sparsity and incoherence in compressive sampling”, Inverse Problems, 23 (3) ISSN 0266-5611, 2007.
11. S. Zahedi and C. Bisdikian, “A framework for QoI-inspired analysis for sensor network deployment planning”, In 2nd Int’l Workshop on Performance Control in Wireless Sensor Networks (PWSN), 2007.
12. B. Sundararaman et al. Clock Synchronization for Wireless Sensor Networks: A Survey. Ad-Hoc Networks, 3(3), May 2005.
13. V. Cevher, A. C. Gurbuz, J. H. McClellan, and R. Chellappa, “Compressive wireless arrays for bearing estimation of sparse sources in angle domain,” ICASSP 2008.
14. B. Hern, “Robustness of Compressed Sensing in Sensor Networks”, Bachelore thesis, 2008.
15. N. Kimura and S. Latifi, “A Survey on Data Compression in Wireless Sensor Networks.” In Proceedings of the international Conference on information Technology: Coding and Computing (ITCC) - Volume II, 8-13. IEEE Computer Society, 2005.
16. K. Barr, and K. Asanovi, “Energy aware lossless data compression.” In Proceedings of the 1st international Conference on Mobile Systems, Applications and Services (MobiSys) 231-244. ACM Press, New York, NY, 2003.
17. J. Kusuma, L. Doherty and K. Ramchandran, "Distributed compression for sensor networks," in Proc. International Conf. Image Processing (ICIP), vol. 1, pp. 82-85, Oct. 2001.
18. T. Arici, B. Gedik, Y. Altunbasak, and L. Liu, “PINCO: a Pipelined In-Network Compression Scheme for Data Collection in Wireless Sensor Networks,” In Proceedings of 12th International Conference on Computer Communications and Networks, October 2003.
19. M. Adler. “Collecting correlated information from a sensor network.” In Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA), 2005.
20. D. Chu, A. Deshpande, J. Hellerstein and W. Hong. “Approximate data collection in sensor networks using probabilistic models.” In Proceedings of the International Conference on Data Engineering (ICDE), 2006.
21. S. Pattem, B. Krishnamachari, and R. Govindan. “The impact of spatial correlation on routing with compression in wireless sensor networks.” In Proceedings of the International Conference on Information Processing in Sensor Networks (IPSN), 2004.
22. S. Pradhan and K. Ramchandran. “Distributed source coding using syndromes (DISCUS): Design and construction.” IEEE Transactions on Information Theory, 49(3), 2003.
23. A. Silberstein, G. Puggioni, A. Gelfand, K. Munagala and J. Yang. “Making Sense of Suppressions and Failures in Sensor Data: A Bayesian Approach.” In Proceedings of the International Conference on Very Large Data Bases (VLDB), 2007.
24. D. Slepian and J. Wolf. “Noiseless coding of correlated information sources.” IEEE Transactions on Information Theory, 19(4), 1973.