We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

3

Distributed Compressed Sensing of Sensor Data

Vasanth Iyer[1] and Dhananjay Singh[2]
[1]*International Institute of Information Technology, Hyderabad A.P. India. 500 032*
[2]*National Institute of Mathematical Sciences (NIMS), Daejeon, 305-340, South Korea*

1. Introduction

Intelligent Information processing in distributed wireless sensor networks has many different optimizations by which redundancies in data can be eliminated, and at the same time the original source signal can be retrieved without loss. The data-centric nature of sensor network is modeled, which allows environmental applications to measure correlated data by periodic data aggregation. In the distributed framework, we explore how Compressed Sensing could be used to represent the measured signals in its sparse form, and model the framework to reproduce the individual signals from the ensembles in its sparse form expressed in equations(1,3). The processed signals are then represented with their common component; which is represented by its significant coefficients, and the variation components, which is also sparse are projected onto scaling and wavelet functions of the correlated component. The overall representation of the basis preserves the temporal (intra-signal) and spatial (inter-signal) characteristics. All of these scenarios correspond to measuring properties of physical processes that change smoothly in time, and in space, and thus are highly correlated. We show by simulation that the framework using cross-layer protocols can be extended using sensor fusion, and data-centric aggregation, to scale to a large number of nodes.

1.1 Cross Layer Sensor Nodes

Sensor network due to its constrained resources such as energy, memory, and range uses a cross layer model for efficient communications. The cross layer model uses *pre-processing*, *post-processing* and *routing*, to accomplish sensor measurements and communications with sensor nodes. Cross layer based routing protocols use different OSI layers to do multi-hop communications. Due to high deployment node densities and short bursts of wireless transmission, not all layers are connected, and can only be coordinated and scheduled by a higher level network function which keeps track of the node states. Due to this limited connectivity between layers, one needs to efficiently schedule the sensor nodes, and its states from the lower-level physical layers, to the higher routing and application layers. The energy spent at each layer needs to be carefully profiled so that any redundancy due to network scalability can further deteriorate the power-aware routing algorithm. One common motivation is to use the least number of bits to represent the data, as the transmission cost per bit increases non-linearly (Power Law) with distance (S. B. Lowen and M. C. Teich (1970)). The other relevant factors, which influence the accuracy of the measured sensor values, versus the total number of sensors deployed, can be divided into pre- and post processing of sensing application parameters. The lower-layer

pre-processing involves, (a) the number of measurement needed so that the measured values can be represented without loss by using intra-sensor redundancy, (b) as the sensor measurements show temporal correlation with inter sensor data, the signal is further divided into many blocks which represent constant variance. In terms of the OSI layer, the pre-processing is done at the physical layer, in our case it is wireless channel with multi-sensor intervals. The network layer data aggregation is based on variable length pre-fix coding, which minimizes the number of bits before transmitting it to a sink. In terms of the OSI layers, data aggregation is done at the data-link layer periodically buffering, before the packets are routed through the upper network layer.

1.2 Computation Model

The sensor network model is based on network scalability the total number of sensors N, which can be very large upto many thousand nodes. Due to this fact an application needs to find the computation power in terms of the combined energy it has, and also the minimum accuracy of the data it can track and measure. The computation steps can be described in terms of the cross-layer protocol messages in the network model. The pre-processing needs to accomplish the minimal number of measurements needed, given by $x = \sum \vartheta(n)\Psi_n = \sum \vartheta(n_k)$, where Ψ_n^k is the best basis. The local coefficients can be represented by 2^j different levels, the search for best basis can be accomplished, using a binary search in $O(\lg m)$ steps. The post processing step involves efficient coding of the measured values, if there are m coefficients, the space required to store the computation can be accomplished in $O(\lg_2 m)$ bits. The routing of data using the sensor network needs to be power-aware, so these uses a distributed algorithm using cluster head rotation, which enhances the total lifetime of the sensor network. The computation complexity of routing in terms of the total number of nodes can be shown as $OC(\lg N)$, where C is the number of cluster heads and N total number of nodes. The computational bounds are derived for pre- and post processing algorithms for large data-sets, and is bounds are derived for a large node size in Section, Theoretical bounds.

1.3 Multi-sensor Data Fusion

Using the cross-layer protocol approach, we like to reduce the communication cost, and derive bounds for the number of measurements necessary for signal recovery under a given sparsity ensemble model, similar to Slepian-Wolf rate (Slepian (D. Wolf)) for correlated sources. At the same time, using the collaborative sensor node computation model, the number of measurements required for each sensor must account for the minimal features unique to that sensor, while at the same time features that appear among multiple sensors must be amortized over the group.

1.4 Chapter organization

Section 2 overviews the categorization of cross-layer pre-processing, CS theories and provides a new result on CS signal recovery. Section 3 introduces routing and data aggregation for our distributed framework and proposes two examples for routing. The performance analysis of cluster and MAC level results are discussed. We provide our detailed analysis for the DCS design criteria of the framework, and the need for pre-processing. In Section 4, we compare the results of the framework with a correlated data-set. The shortcomings of the upper layers which are primarily routing centric are contrasted with data centric routing using DHT, for the same family of protocols. In Section 5, we close the chapter with a discussion and conclusions. In appendices several proofs contain bounds for scalability of resources. For prerequisites and programming information using sensor applications you may refer to the book

by (S. S. Iyengar and Nandan Parameshwaran (2010)) Fundamentals of Sensor Programming, Application and Technology.For Applications on emerging domains and implementations, please refer to (Dhananjay Singh and Hoon-Jae Lee, 2010a;b).

2. Pre-Processing

As different sensors are connected to each node, the nodes have to periodically measure the values for the given parameters which are correlated. The inexpensive sensors may not be calibrated, and need processing of correlated data, according to intra and inter sensor variations. The pre-processing algorithms allow to accomplish two functions, one to use minimal number of measurement at each sensor, and the other to represent the signal in its loss-less sparse representation.

2.1 Compressive Sensing (CS)

The signal measured if it can be represented at a sparse Dror Baron (Marco F. Duarte) representation, then this technique is called the sparse basis as shown in equation (1), of the measured signal. The technique of finding a representation with a small number of significant coefficients is often referred to as Sparse Coding. When sensing locally many techniques have been implemented such as the Nyquist rate (Dror Baron (Marco F. Duarte)), which defines the minimum number of measurements needed to faithfully reproduce the original signal. Using CS it is further possible to reduce the number of measurement for a set of sensors with correlated measurements (Bhaskar Krishnamachari (Member)).

$$x = \sum \vartheta(n)\Psi_n = \sum \vartheta(n_k)\Psi_{n_k}, \tag{1}$$

Consider a real-valued signal $x \in R^N$ indexed as x(n), $n \in 1, 2, ..., N$. Suppose that the basis $\Psi = [\Psi_1, ..., \Psi_N]$ provides a K-sparse representation of x; that is, where x is a linear combination of K vectors chosen from, Ψ, n_k are the indices of those vectors, and $\vartheta(n)$ are the coefficients; the concept is extendable to tight frames (Dror Baron (Marco F. Duarte)). Alternatively, we can write in matrix notation $x = \Psi\vartheta$, where x is an $N \times 1$ column vector, the sparse basis matrix is $N \times N$ with the basis vectors Ψ_n as columns, and $\vartheta(n)$ is an $N \times 1$ column vector with K nonzero elements. Using $\| \, . \, \|_p$Ąto denote the ℓ_p norm, we can write that $\| \, \vartheta \, \|_p = K$; we can also write the set of nonzero indices $\Omega 1, ..., N$, with $|\Omega| = K$. Various expansions, including wavelets (Dror Baron (Marco F. Duarte)), Gabor bases (Dror Baron (Marco F. Duarte)), curvelets (Dror Baron (Marco F. Duarte)), are widely used for representation and compression of natural signals, images, and other data.

2.2 Sparse representation

A single measured signal of finite length, which can be represented in its sparse representation, by transforming into all its possible basis representations. The number of basis for the for each level j can be calculated from the equation as

$$A_{j+1} = A_j^2 + 1 \tag{2}$$

So staring at $j = 0$, $A_0 = 1$ and similarly, $A_1 = 1^2 + 1 = 2$, $A_2 = 2^2 + 1 = 5$ and $A_3 = 5^2 + 1 = 26$ different basis representations.
Let us define a framework to quantify the sparsity of ensembles of correlated signals $x_1, x2, ..., xj$ and to quantify the measurement requirements. These correlated signals can be represented

by its basis from equation (2). The collection of all possible basis representation is called the sparsity model.

$$x = P\theta \tag{3}$$

Where P is the sparsity model of K vectors ($K << N$) and θ is the non zero coefficients of the sparse representation of the signal. The sparsity of a signal is defined by this model P, as there are many factored possibilities of $x = P\theta$. Among the factorization the unique representation of the smallest dimensionality of θ is the sparsity level of the signal x under this model, or ϵ which is the smallest interval among the sensor readings distinguished after cross-layer aggregation.

2.3 Distributed Compressive Sensing (DCS)

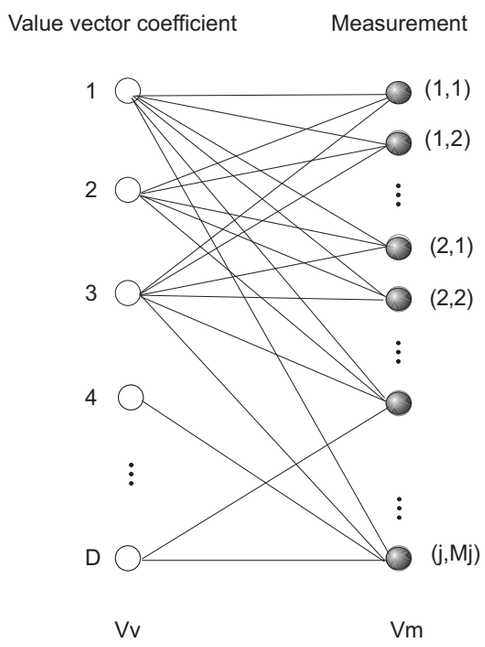


Fig. 1. Bipartite graphs for distributed compressed sensing.

DCS allows to enable distributed coding algorithms to exploit both intra-and inter-signal correlation structures. In a sensor network deployment, a number of sensors measure signals that are each individually sparse in the some basis and also correlated from sensor to sensor. If the separate sparse basis are projected onto the scaling and wavelet functions of the correlated sensors(common coefficients), then all the information is already stored to individually recover each of the signal at the joint decoder. This does not require any pre-initialization between sensor nodes.

2.3.1 Joint Sparsity representation

For a given ensemble X, we let $P_F(X) \subseteq P$ denote the set of feasible location matrices $P \in P$ for which a factorization $X = P\Theta$ exits. We define the joint sparsity levels of the signal ensemble as follows. The joint sparsity level D of the signal ensemble X is the number of columns of the smallest matrix $P \in P$. In these models each signal x_j is generated as a combination

of two components: (i) a common component z_C, which is present in all signals, and (ii) an innovation component z_j, which is unique to each signal. These combine additively, giving

$$x_j = z_C + z_j, j \in \forall \tag{4}$$

$$X = P\Theta \tag{5}$$

We now introduce a bipartite graph $G = (V_V, V_M, E)$, as shown in Figure 1, that represents the relationships between the entries of the value vector and its measurements. The common and innovation components K_C and K_j, $(1 < j < J)$, as well as the joint sparsity $D = K_C + \sum K_J$. The set of edges E is defined as follows:

- The edge E is connected for all K_c if the coefficients are not in common with K_j.
- The edge E is connected for all K_j if the coefficients are in common with K_j.

A further optimization can be performed to reduce the number of measurement made by each sensor, the number of measurement is now proportional to the maximal overlap of the inter sensor ranges and not a constant as shown in equation (1). This is calculated by the common coefficients K_c and K_j, if there are common coefficients in K_j then one of the K_c coefficient is removed and the common Z_c is added, these change does not effecting the reconstruction of the original measurement signal x.

3. Post-Processing and Routing

The computation of this layer primarily deals with compression algorithms and distributed routing, which allows efficient packaging of data with minimal number of bits. Once the data are fused and compressed it uses a network protocol to periodically route the packets using multi-hoping. The routing in sensor network uses two categories of power-aware routing protocols, one uses distributed data aggregation at the network layer forming clusters, and the other uses MAC layer protocols to schedule the radio for best effort delivery of the multi-hop packets from source to destination. Once the data is snap-shotted, it is further aggregated into sinks by using Distributed Hash based routing (DHT) which keeps the number of hops for a query path length constant in a distributed manner using graph embedding James Newsome and Dawn Song (2003).

3.1 Cross-Layer Data Aggregation

Clustering algorithms periodically selects cluster heads (CH), which divides the network into k clusters which are in the CHs Radio range. As the resources at each node is limited the energy dissipation is evenly distributed by the distributed CH selection algorithm. The basic energy consumption for scalable sensor network is derived as below.
Sensor node energy dissipation due to transmission over a given range and density follows Power law, which states that energy consumes is proportional to the square of the distance in m^2 transmitted.

$$PowerLaw = 1^2 + 2^2 + 3^2 + 4^2 + ... + (d-1)^2 + d^2 \tag{6}$$

To sum up the total energy consumption we can write it in the form of Power Law equation [7]

$$PowerLaw = f(x) = ax^2 + o(x)^2 \tag{7}$$

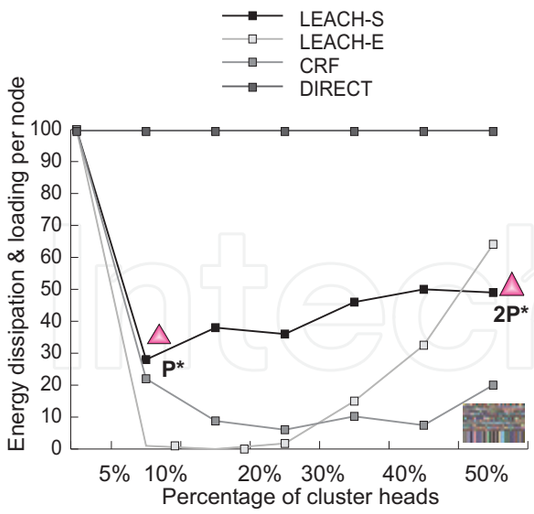


Fig. 2. Cost function for managing residual energy using LEACH routing.

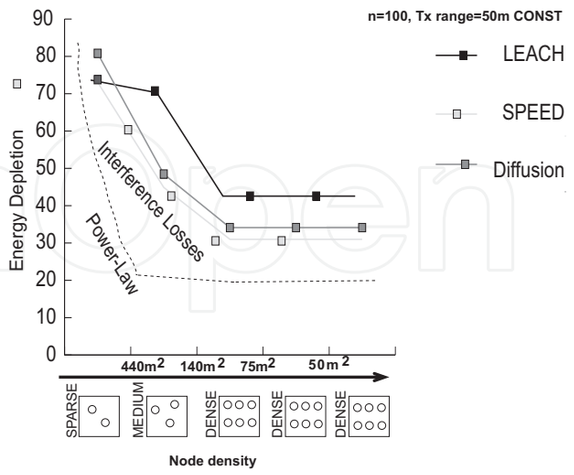


Fig. 3. Power-aware MAC using multi-hop routing.

Substituting d-distance for x and k number of bits transmitted, we equate as in equation (7).

$$PowerLaw = f(d) = kd^2 + o(d)^2 \tag{8}$$

Taking Log both sides of equation (8),

$$\log(f(d)) = 2\log d + \log k \tag{9}$$

Notice that the expression in equation (10) has the form of a linear relationship with slope k, and scaling the argument induces a linear shift of the function, and leaves both the form and slope k unchanged. Plotting to the log scale as shown in Figure 3, we get a long tail showing a few nodes dominate the transmission power compared to the majority, similar to the Power Law (S. B. Lowen and M. C. Teich (1970)).
Properties of power laws - Scale invariance: The main property of power laws that makes them interesting is their scale invariance. Given a relation $f(x) = ax^k$ or, any homogeneous polynomial, scaling the argument x by a constant factor causes only a proportionate scaling of the function itself. From the equation (10), we can infer that the property is scale invariant even with clustering c nodes in a given radius k.

$$f(cd) = k(cd^2) = c^k f(d) \alpha f(d) \tag{10}$$

This is validated from the simulation results (Vasanth Iyer (G. Rama Murthy)) obtained in Figure (2), which show optimal results, minimum loading per node (Vasanth Iyer (S.S. Iyengar)), when clustering is $\leq 20\%$ as expected from the above derivation.

3.2 MAC Layer Routing

The IEEE 802.15.4 (Joseph Polastre (Jason Hill)) is a standard for sensor network MAC interoperability, it defines a standard for the radios present at each node to reliably communicate

Sensors	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
Value	4.7 $\pm$ 2.0	1.6 $\pm$ 1.6	3.0 $\pm$ 1.5	1.8 $\pm$ 1.0	4.7 $\pm$ 1.0	1.6 $\pm$ 0.8	3.0 $\pm$ 0.75	1.8 $\pm$ 0.5
Group	-	-	-	-	-	-	-	-

Table 1. A typical random measurements from sensors showing non-linearity in ranges

with each other. As the radios consume lots of power the MAC protocol for best performance uses *Idle*, *Sleep* and *Listen* modes to conserve battery. The radios are scheduled to periodically listen to the channel for any activity and receive any packets, otherwise it goes to idle, or sleep mode. The MAC protocol also needs to take care of collision as the primary means of communication is using broadcast mode. The standard carrier sense multiple access (CSMA) protocol is used to share the channel for simultaneous communications. Sensor network variants of CSMA such as B-MAC and S-MAC Joseph Polastre (Jason Hill) have evolved, which allows to better handle passive listening, and used low-power listening(LPL). The performance characteristic of MAC based protocols for varying density (small, medium and high) deployed are shown in Figure 3. As it is seen it uses best effort routing (least cross-layer overhead), and maintains a constant throughput, the depletion curve for the MAC also follows the Power Law depletion curve, and has a higher bound when power-aware scheduling such LPL and Sleep states are further used for idle optimization.

3.2.1 DHT KEY Lookup

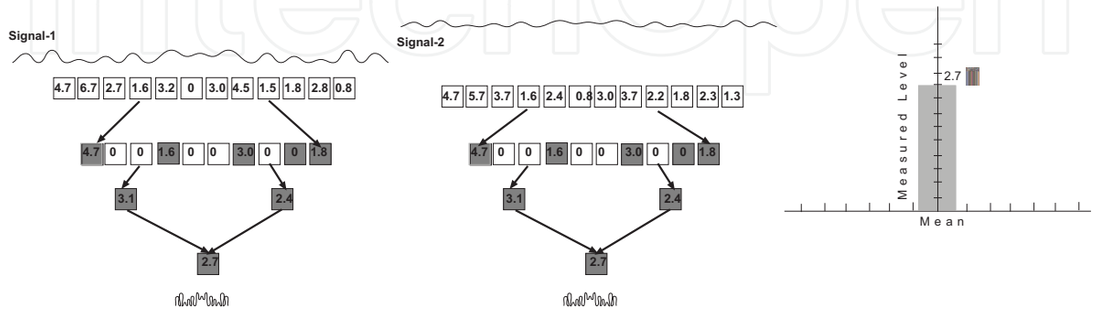
Topology of the overlay network uses an addressing which is generated by consistent hashing of the node-id, so that the addressing is evenly distributed across all nodes. The new data is stored with its $< KEY >$ which is also generated the same way as the node address range. If the specific node is not in the range the next node in the clockwise direction is assigned the data for that $< KEY >$. From theorem:4, we have that the average number of hops to retrieve the value for the $< KEY, VALUE >$ is only $O(\lg n)$ hops. The routing table can be tagged with application specific items, which are further used by upper layer during query retrieval. In domain specific application such as medical integration, the performance measure can be found in the related work of (Dhananjay Singh and Daeyeoul Kim, 2010a;b).

4. Comparison of DCS and Data Aggregation

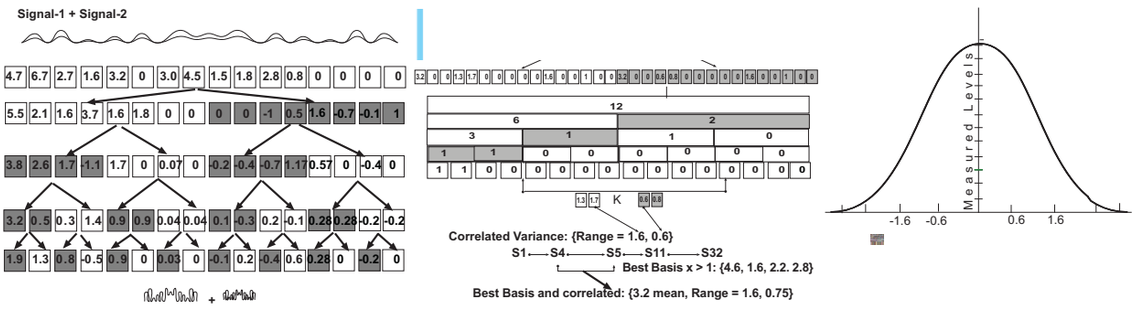
In Section 4 and 5, we have seen various data processing algorithms, in terms of communication cost they are comparable. In this Section, we will look into two design factors of the distributed framework:

1. Assumption1: How well the individual sensor signal sparsity can be represented.
2. Assumption2: What would be the minimum measurement possible by using joint sparsity model from equation (5).
3. Assumption3: The maximum possible basis representations for the joint ensemble coefficients.
4. Assumption4: A cost function search which allows to represent the best basis without overlapping coefficients.
5. Assumption5: Result validation using regression analysis, such package R (Owen Jones (Robert Maillardet)).

The design framework allows to pre-process individual sensor sparse measurement, and uses a computationally efficient algorithm to perform in-network data fusion.
To use an example data-set, we will use four random measurements obtained by multiple sensors, this is shown in Table 1. It has two groups of four sensors each, as shown the mean value are the same for both the groups and the variance due to random sensor measurements vary with time. The buffer is created according to the design criteria (1), which preserves the sparsity of the individual sensor readings, this takes three values for each sensor to be represented as shown in Figure (4).



(a) Post-Processing and Data Aggregation



(b) Pre-Processing and Sensor Data Fusion

Fig. 4. Sensor Value Estimation with Aggregation and Sensor Fusion

In the case of post-processing algorithms, which optimizes on the space and the number of bits needed to represent multi-sensor readings, the fusing sensor calculates the average or the mean from the values to be aggregated into a single value. From our example data, we see that for both the data-sets gives the same end result, in this case $\mu = 2.7$ as shown in the output plot of Figure 4(a). Using the design criteria (1), which specifies the sparse representation is not used by post-processing algorithms. Due to this dynamic features are lost during data aggregation step.
The pre-processing step uses Discrete Wavelet Transform (DWT) (Arne Jensen and Anders la Cour-Harbo (2001)) on the signal, and may have to recursively apply the decomposition to arrive at a sparse representation, this pre-process is shown in Figure 4(b). This step uses the design criteria (1), which specifies the small number of significant coefficients needed to represent the given signal measured. As seen in Figure 4(b), each level of decomposition reduces the size of the coefficients. As memory is constrained, we use up to four levels of

decomposition with a possible of 26 different representations, as computed by equation (2). These uses the design criteria (3) for lossless reconstruction of the original signal.
The next step of pre-processing is to find the best basis, we let a vector Basis of the same length as cost values representing the basis, this method uses Algorithm 1. The indexing of the two vector is the same and are enumerated in Figure of 4(b). In Figure 4(b), we have marked a basis with shaded boxes. This basis is then represented by the vector. The basis search, which is part of design criteria (4), allows to represent the best coefficients for inter and intra sensor features. It can be noticed that the values are not averages or means of the signal representation, it preserves the actual sensor outputs. As an important design criteria (2), which calibrates the minimum possible sensitivity of the sensor. The output in figure 4(b), shows the constant estimate of S_3, S_7 which is $Z_C = 2.7$ from equation (4).
To represent the variance in four sensors, a basis search is performed which finds coefficients of sensors which matches the same columns. In this example, we find $Z_j = 1.6, 0.75$ from equation (4), which are the innovation component.

$$Basis = [0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ \ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$$

$$Correlated\ range = [0\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ \ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$$

4.1 Lower Bound Validation using Covariance

The Figure 4(b) shows lower bound of the overlapped sensor i.i.d. of $S_1 - S_8$, as shown it is seen that the lower bound is unique to the temporal variations of S_2. In our analysis we will use a general model which allows to detect sensor faults. The binary model can result from placing a threshold on the real-valued readings of sensors. Let m_n be the mean normal reading and m_f the mean event reading for a sensor. A reasonable threshold for distinguishing between the two possibilities would be $0.5(\frac{m_n+m_f}{2})$. If the errors due to sensor faults and the fluctuations in the environment can be modeled by Gaussian distributions with mean 0 and a standard deviation σ, the fault probability p would indeed be symmetric. It can be evaluated using the tail probability of a Gaussian Bhaskar Krishnamachari (Member), the Q-function, as follows:

$$p = Q\frac{\left((0.5(\frac{m_n+m_f}{2}) - m_n)\right)}{\sigma} = Q\left(\frac{m_f - m_n}{2\sigma}\right) \tag{11}$$

From the measured i.i.d. value sets we need to determine if they have any faulty sensors. This can be shown from equation (11) that if the correlated sets can be distinguished from the mean values then it has a low probability of error due to sensor faults, as sensor faults are not correlated. Using the statistical analysis package R Owen Jones (Robert Maillardet), we determine the correlated matrix of the sparse sensor outputs as shown This can be written in a compact matrix form if we observe that for this case the covariance matrix is diagonal, this is,

$$\Sigma = \begin{pmatrix} \rho_1 & 0 & .. & 0 \\ 0 & \rho_2 & .. & 0 \\ : & : & \searrow & : \\ 0 & 0 & .. & \rho_d \end{pmatrix} \tag{12}$$

The correlated co-efficient are shown matrix (13) the corresponding diagonal elements are highlighted. Due to overlapping reading we see the resulting matrix shows that S_1 and S_2 have higher index. The result sets is within the desired bounds of the previous analysis using DWT. Here we not only prove that the sensor are not faulty but also report a lower bound of

Sensors	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
$i.i.d._1$	2.7	0	1.5	0.8	3.7	0.8	2.25	1.3
$i.i.d._2$	4.7	1.6	3	1.8	4.7	1.6	3	1.8
$i.i.d._3$	6.7	3.2	4.5	2.8	5.7	2.4	3.75	2.3

Table 2. Sparse representation of sensor values from Table:1

the optimal correlated result sets, that is we use S_2 as it is the lower bound of the overlapping ranges.

$$\Sigma = \begin{pmatrix} \overrightarrow{4.0} & 3.20 & 3.00 & 2.00 & 2.00 & 1.60 & 1.5 & 1.0 \\ 3.2 & \overrightarrow{2.56} & 2.40 & 1.60 & 1.60 & 1.28 & 1.20 & 0.80 \\ 3.0 & 2.40 & \overrightarrow{2.250} & 1.50 & 1.50 & 1.20 & 1.125 & 0.75 \\ 2.0 & 1.60 & 1.50 & \overrightarrow{1.00} & 1.00 & 0.80 & 0.75 & 0.5 \\ 2.0 & 1.60 & 1.50 & 1.00 & \overrightarrow{1.00} & 0.80 & 0.75 & 0.5 \\ 1.6 & 1.28 & 1.20 & 0.80 & 0.80 & \overrightarrow{0.64} & 0.60 & 0.4 \\ 1.5 & 1.20 & 1.125 & 0.75 & 0.75 & 0.60 & \overrightarrow{0.5625} & 0.375 \\ 1.0 & 0.80 & 0.750 & 0.50 & 0.50 & 0.40 & 0.375 & \overrightarrow{0.250} \end{pmatrix} \quad (13)$$

5. Conclusion

In this topic, we have discussed a distributed framework for correlated multi-sensor measurements and data-centric routing. The framework, uses compressed sensing to reduce the number of required measurements. The joint sparsity model, further allows to define the system accuracy in terms of the lowest range, which can be measured by a group of sensors. The sensor fusion algorithms allows to estimate the physical parameter, which is being measured without any inter sensor communications. The reliability of the pre-processing and sensor faults are discussed by comparing DWT and Covariance methods.
The complexity model is developed which allows to describe the encoding and decoding of the data. The model tends to be easy for encoding, and builds more complexity at the joint decoding level, which are nodes with have more resources as being the decoders.
Post processing and data aggregation are discussed with cross-layer protocols at the network and the MAC layer, its implication to data-centric routing using DHT is discussed, and compared with the DCS model. Even though these routing algorithms are power-aware, the model does not scale in terms of accurately estimating the physical parameters at the sensor level, making sensor driven processing more reliable for such applications.

6. Theoretical Bounds

The computational complexities and its theoretical bounds are derived for categories of sensor pre-, post processing and routing algorithms.

6.1 Pre-Processing

Algorithm 1 DWT: Using a cost function for searching the best sparse representation of a signal.

1: Mark all the elements on the bottom level
2: Let $j = J$
3: Let $k = 0$
4: Compare the cost v_1 of the element k on level $(j-1)$ (counting from the left on that level) to the sum v_2 of the cost values of the element $2k$ and the $2k+1$ on the level j.
5: if $v_1 \leq v_2$, all marks below element k on level $j-1$ are deleted, and element k is marked.
6: if $v_1 > v_2$, the cost value v_1 of element k is replaced with v_2 $k = k+1$. If there are more elements on level j (if $k < 2^{j-1} - 1$)), go to step 4.
7: $j = j-1$. If $j > 1$, go to step 3.
8: The marked sparse representation has the lowest possible cost value, having no overlaps.

Theorem 1. *The Slepian-Wolf rate as referenced in the region for two arbitrarily correlated sources x and y is bounded by the following inequalities, this theorem can be adapted using equation*

$$R_x \geq H\left(\frac{x}{y}\right), R_y \geq H\left(\frac{y}{x}\right) \text{ and } R_x + R_y \geq H(x,y) \tag{14}$$

Theorem 2. *minimal spanning tree (MST) computational and time complexity for correlated dendrogram. First considering the computational complexity let us assume n patterns in d-dimensional space. To make c clusters using $d_{min}(D_i, Dj)$ a distance measure of similarity. We need once for all, need to calculate $n(n-1)$ interpoint distance table. The space complexity is n^2, we reduce it to $\lg(n)$ entries. Finding the minimum distance pair (for the first merging) requires that we step through the complete list, keeping the index of the smallest distance. Thus, for the first step, the complexity is $O(n(n-1))(d^2+1) = O(n^2d^2)$. For clusters c the number of steps is $n(n-1) - c$ unused distances. The full-time complexity is $O(n(n-1)-c)$ or $O(cn^2d^2)$.*

6.2 Post-processing

Theorem 3. *Properties of Pre-fix coding: For any compression algorithm which assigns prefix codes and to uniquely be decodable. Let us define the **kraft Number** and is a measure of **the size** of L. We see that if L is 1, 2^{-L} is .5. We know that we cannot have more than two L's of .5. If there are more that two L's of .5, then $K > 1$. Similarly, we know L can be as large as we want. Thus, 2^{-L} can be as small as we want, so K can be as small as we want. Thus we can intuitively see that there must be a strict upper bound on K, and no lower bound. It turns out that a prefix-code only exists for the codes IF AND ONLY IF:*

$$K \leq 1 \tag{15}$$

The above equation is the Kraft inequality. The success of transmission can be further calculated by using the equation For a minimum pre-fix code $a = 0.5$ as $2^{-L} \leq 1$ for a unique decodability.
Iteration $a = 0.5$
In order to extend this scenario with distributed source coding, we consider the case of separate encoders for each source, x_n and y_n. Each encoder operates without access to the other source.
Iteration $a \geq 0.5 \leq 1.0$
As in the previous case it uses correlated values as a dependency and constructs the code-book. The

compression rate or efficiency is further enhanced by increasing the correlated CDF higher than $a > 0.5$. *This produces very efficient code-book and the design is independent of any decoder reference information. Due to this a success threshold is also predictable, if* $a = 0.5$ *and the cost between* $L = 1.0$ *and* 2.0 *the success* $= 50\%$ *and for* $a = 0.9$ *and* $L = 1.1$, *the success* $= 71\%$.

6.3 Distributed Routing

Theorem 4. *The Cayley Graph* (S, E) *of a group: Vertices corresponding to the underlying set S. Edges corresponding to the actions of the generators. (Complete) Chord is a Cayley graph for* $(Zn, +)$. *The routing nodes can be distributed using* $S = Z \bmod n\,(n = 2^m)$ *very similar to our simulation results of LEACH (Vasanth Iyer (G. Rama Murthy)). Generators for one-way hashing can use these fixed length hash* $1, 2, 4, , 2^m - 1$. *Most complete Distributed Hash Table (DHTs) are Cayley graphs. Data-centric algorithm Complexity: where Z is the original ID and the key is its hash between* $0 - 2^m$, $ID + key$ *are* uniformly *distributed in the chord (Vasanth Iyer (S. S. Iyengar)).*

7. References

S. Lowen and M. Teich. (1970). Power-Law Shot Noise, IEEE Trans Inform volume 36, pages 1302-1318, 1970.

Slepian, D. Wolf, J. (1973). Noiseless coding of correlated information sources. Information Theory, IEEE Transactions on In Information Theory, IEEE Transactions on, Vol. 19, No. 4. (06 January 2003), pp. 471-480.

Bhaskar Krishnamachari, S.S. Iyengar. (2004). Distributed Bayesian Algorithms for Fault-Tolerant Event Region Detection in Wireless Sensor Networks, In: *IEEE TRANSACTIONS ON COMPUTERS*,VOL. 53, NO. 3, MARCH 2004.

Dror Baron, Marco F. Duarte, Michael B. Wakin, Shriram Sarvotham, and Richard G. Baraniuk. (2005). Distributed Compressive Sensing. In Proc: *Pre-print*, Rice University, Texas, USA, 2005.

Vasanth Iyer, G. Rama Murthy, and M.B. Srinivas. (2008). Min Loading Max Reusability Fusion Classifiers for Sensor Data Model. In Proc: *Second international Conference on Sensor Technologies and Applications*, Volume 00 (August 25 - 31, SENSORCOMM 2008).

Vasanth Iyer, S.S. Iyengar, N. Balakrishnan, Vir. Phoha, M.B. Srinivas. (2009). FARMS: Fusionable Ambient Renewable MACS, In: *SAS-2009*, IEEE 9781-4244-2787, 17th-19th Feb, New Orleans, USA.

Vasanth Iyer, S. S. Iyengar, Rama Murthy, N. Balakrishnan, and V. Phoha. (2009). Distributed source coding for sensor data model and estimation of cluster head errors using bayesian and k-near neighborhood classifiers in deployment of dense wireless sensor networks, In Proc: *Third International Conference on Sensor Technologies and Applications SENSORCOMM*, 17-21 June. 2009.

Vasanth Iyer, S.S. Iyengar, G. Rama Murthy, Kannan Srinathan, Vir Phoha, and M.B. Srinivas. INSPIRE-DB: Intelligent Networks Sensor Processing of Information using Resilient Encoded-Hash DataBase. In Proc. Fourth International Conference on Sensor Technologies and Applications, IARIA-SENSORCOMM, July, 18th-25th, 2010, Venice, Italy (archived in the Computer Science Digital Library).

Vasanth Iyer, S.S. Iyengar, N. Balakrishnan, Vir. Phoha, M.B. Srinivas. (2009). FARMS: Fusionable Ambient Renewable MACS, In: *SAS-2009*,IEEE 9781-4244-2787, 17th-19th Feb, New Orleans, USA.

GEM: Graph EMbedding for Routing and DataCentric Storage in Sensor Networks Without Geographic Information. Proceedings of the First ACM Conference on Embedded Networked Sensor Systems (SenSys). November 5-7, Redwood, CA.

Owen Jones, Robert Maillardet, and Andrew Robinson. Introduction to Scientific Programming and Simulation Using R. Chapman & Hall/CRC, Boca Raton, FL, 2009. ISBN 978-1-4200-6872-6.

Arne Jensen and Anders la Cour-Harbo. Ripples in Mathematics, Springer Verlag 2001. 246 pp. Softcover ISBN 3-540-41662-5.

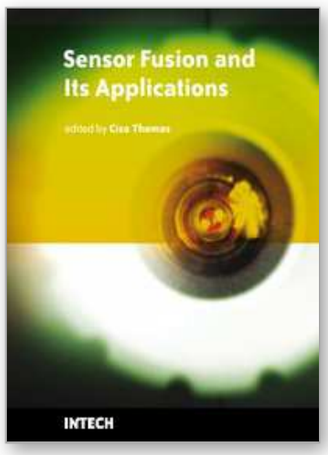
S. S. Iyengar, Nandan Parameshwaran, Vir V. Phoha, N. Balakrishnan, and Chuka D Okoye, Fundamentals of Sensor Network Programming: Applications and Technology. ISBN: 978-0-470-87614-5 Hardcover 350 pages December 2010, Wiley-IEEE Press.

Dhananjay Singh and Hoon-Jae Lee, Design and Performance Evaluation of a Proactive Micro Mobility Protocol for Mobile Networks, in Book: Handheld Computing for Mobile Commerce: Applications, Concepts and Technologies, IGI Global USA, Feb. 2010, edited by DrWen-Chen Hu and Dr Yanjun Zuo.

Dhananjay Singh and Hoon-Jae Lee, Global Healthcare Monitoring Applications: IPv6 Based Wireless Sensor Networks, in Book: E-Healthcare Systems And Wireless Communications: Current And Future Challenges, IGI Global USA, Feb. 2010.

Dhananjay Singh and Daeyeoul Kim, MMSP: Design a Novel Micro Mobility Sensor Protocol for Ubiquitous Communication, UBICOMM2010, Florence, Italy, October 25 30, 2010.

Dhananjay Singh and Daeyeoul Kim, Performance Analysis of Gateway Discovery Techniques: IPv6 based Wireless Sensor Networks, INTERNET2010, Valencia, Spain, September 20 25, 2010.



Sensor Fusion and its Applications
Edited by Ciza Thomas

ISBN 978-953-307-101-5
Hard cover, 488 pages
Publisher Sciyo
Published online 16, August, 2010
Published in print edition August, 2010

This book aims to explore the latest practices and research works in the area of sensor fusion. The book intends to provide a collection of novel ideas, theories, and solutions related to the research areas in the field of sensor fusion. This book is a unique, comprehensive, and up-to-date resource for sensor fusion systems designers. This book is appropriate for use as an upper division undergraduate or graduate level text book. It should also be of interest to researchers, who need to process and interpret the sensor data in most scientific and engineering fields. The initial chapters in this book provide a general overview of sensor fusion. The later chapters focus mostly on the applications of sensor fusion. Much of this work has been published in refereed journals and conference proceedings and these papers have been modified and edited for content and style. With contributions from the world's leading fusion researchers and academicians, this book has 22 chapters covering the fundamental theory and cutting-edge developments that are driving this field.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen