RANDOM MULTIRESOLUTION REPRESENTATIONS FOR ARBITRARY SENSOR NETWORK GRAPHS

Wei Wang and Kannan Ramchandran

University of California, Berkeley
Department of Electrical Engineering and Computer Sciences
Email: `wangwei, kannanr@eecs.berkeley.edu`

ABSTRACT

We propose a distributed multiresolution representation of sensor network data so that large-scale summaries are readily available by querying a small fraction of sensor nodes, anywhere in the network, and small-scale details are available by querying a larger number of sensors, locally in the region of interest. A global querier (such as a mobile collector or unmanned aerial vehicle) can obtain a lossy to lossless representation of the network data, according to the desired resolution. A local querier (such as a sensor node) can also obtain either large-scale trends or local details, by querying its immediate neighborhood. We want the encoding to be robust to arbitrary, even time-varying, wireless communication connectivity graphs. Thus we want to avoid cluster heads or deterministic hierarchies that are not robust to single points of failure. We propose a randomized encoding which enables both robustness, and distributed computation that does not require long distance coordination or awareness of network connectivity at individual sensors. Our distributed encoding algorithm operates on local neighborhoods of the communication graph.

1. INTRODUCTION

Sensor networks may have a significant impact on sciences that previously had limited or no data, but scientists must be able to deal with the huge quantities of data measured by sensor networks. Signal processing will play an important role in analyzing and compressing such large quantities of data. Information of significant interest to the querier, such as summaries, averages, or anomalies, should be readily available from anywhere in the network. The querier should also be able to obtain all the data if needed, but in a multiresolution refinable way. Since the functions of interest may depend on the end application, or vary for different types of queriers, the sensor network should be able to serve diverse queries.

As technology advances make wireless sensor motes smaller and cheaper, communication continues to be the dominant cost in power consumption. Furthermore, in wireless communications, connectivity is highly time-varying and spatially nonuniform. Thus signal processing algorithms for sensor networks should be robust and distributed, without requiring rigid coordination, synchronization, and knowledge.

In this paper, we propose a distributed multiresolution representation of sensor network data that encodes on local neighborhoods and is robust to arbitrary, time-varying communication links. The main contributions of our paper are:

- Multiresolution querying. Our algorithm guarantees, with high probability, that large-scale averages and differences are available by querying a small number of sensors, anywhere in the network. Small-scale averages and differences are available, with high probability, by querying a larger number of sensors, locally in the region of interest. The original data can be reconstructed without error by querying one symbol from every node.
- Distributed robust encoding. Each sensor independently and randomly encodes on local neighborhoods of the communication graph. Thus sensors do not need to coordinate with or be aware of other sensors' states, and no routing over relays is needed. Furthermore our algorithm is robust to arbitrary, time-varying communication graphs in which links can fail. Sensors do not need to know the network connectivity.

1.1. Past Work

Distributed multiresolution and wavelet algorithms have been proposed [1, 2, 3, 4, 5, 6]. These past works typically have sink models where a root node collects all the information. Layers of intermediate cluster heads collect data from their subregions of the network and compute the wavelet coefficients. The wavelet transform decomposition tree is thus overlayed on the network communication graph. These schemes are not robust to single node or link failures. They require rigid coordination and synchronization between nodes, as well as knowledge of the network connectivity at individual nodes. These schemes have restrictive models on the network connectivity graph such as grids or trees.

Consensus algorithms have been proposed [7, 8, 9, 10] to compute the average of sensor network data. Our proposed scheme has a similar goal of using only local computation that does not require knowledge or coordination, and formulating the global effect in terms of adjacency matrices of the communication graph. However, consensus algorithms are iterative schemes whose convergence relies on properties of the 1-hop adjacency matrix of the graph. Our proposed algorithm is not iterative and does not require very large numbers of communications to mix the data. Our scheme further utilizes properties of more general adjacency relationships of the graph.

Network coding algorithms [11] encode independent source data in a distributed manner and guarantees decodability of all of the original data. It is possible that such schemes may be extended to allow decoding of subsets of the original data. These schemes do not allow, however, encoding of functions of the original data, such as averages or transforms, unless clusterheads first collect and compute the functions.

2. MULTIRESOLUTION GRAPH-DEPENDENT BASES

A sensor network can be represented by a finite undirected graph $G = (V, E)$, where sensors are located on the vertices V and communication links are represented by edges E. Each sensor $v \in V$ makes a measurement $x(v) \in \mathbb{R}$. Let the number of sensor nodes be $n = |V|$, then the aggregate sensor data is $\mathbf{x} = [x(v_1), \ldots, x(v_n)]^T \in \mathbb{R}^n$. The vertex set V of a graph forms a complete metric space, with the shortest path distance metric (in hops) $d(v_i, v_j), v_i, v_j \in V$. The h-hop neighborhood of a node $v_i \in V$ is defined as $\mathcal{N}_h(v_i) = \{v_j \in V : d(v_i, v_j) \leq h\}$.

We can thus define functions $f(v)$, that are mappings from the vertex set V into the reals $\mathbb{R}$, and have localized support on V. In other words, a graph-dependent function takes the weighted sum of the sensor values in an h-hop neighborhood, and is zero outside the neighborhood.

$$f_{i,h_i}(v_j) = \begin{cases} w_{ij} & \text{for } v_j \in \mathcal{N}_{h_i}(v_i), \\ 0 & \text{for } v_j \notin \mathcal{N}_{h_i}(v_i) \end{cases}$$

for $i \in \{1, \ldots, |V|\}$ and $h_i \in \mathbb{N}$. These functions can be scaled to different resolutions, corresponding to different neighborhood sizes h_i, and shifted to center over each node v_i in the graph. We want to construct sets of graph-dependent bases, $\{\mathbf{f}_{i,h_i} = [f_{i,h_i}(v_1), \ldots, f_{i,h_i}(v_n)]\}$, which span $\mathbb{R}^n$, so that every possible data vector measured by the sensor network can be uniquely represented as a linear combination of the basis functions, $\mathbf{x} = \sum_i c_i \mathbf{f}_{i,h_i}$.

Motivated by simple Haar scaling and wavelet functions consisting of averages and differences, we define two graph-dependent bases.

Weighted Average Basis Function: computes the weighted average on the h_i-hop neighborhood of node v_i, placing more weight on the data measured at v_i.

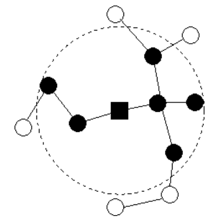

Fig. 1. A locally-supported graph-dependent basis function.

$$f_{i,h_i}(v_j) = \begin{cases} \frac{a}{d_{i,h_i}} & \text{if } v_j \in \mathcal{N}_{h_i}(v_i) \setminus v_i, \\ (1-a) + \frac{a}{d_{i,h_i}} & \text{if } i = j, \\ 0 & \text{otherwise} \end{cases}$$

where $0 < a < \frac{1}{2}$ is a constant and $d_{i,h_i} = |\mathcal{N}_{h_i}(v_i)| \geq 1$ is the degree of the h_i-hop neighborhood. Note that all the weights are nonnegative.

Weighted Difference Basis Function: computes the weighted difference of the value of node v_i to the values of its h_i-hop neighbors.

$$g_{i,h_i}(v_j) = \begin{cases} -\frac{b}{d_{i,h_i}} & \text{if } v_j \in \mathcal{N}_{h_i}(v_i) \setminus v_i, \\ (1+b) - \frac{b}{d_{i,h_i}} & \text{if } i = j, \\ 0 & \text{otherwise} \end{cases}$$

where $b > 0$ is a constant and $d_{i,h_i} = |\mathcal{N}_{h_i}(v_i)| \geq 1$. Note that the weight on node v_i's value is positive ($1+b-b/d_{i,h_i} > 0$) and the weights on all other h_i-hop neighbors of v_i are negative ($-b/d_{i,h_i} < 0$).

It is worth noting that an unweighted average or difference would not generate a basis for arbitrary graphs. We can easily construct counterexamples where subgraphs of the network connectivity are fully connected, resulting in the same unweighted average centered at different sensors. But if each node stores both the weighted average and weighted difference basis coefficients, at no extra communication cost, then the querier can obtain the unweighted average (and difference) of the h-hop neighborhood of the queried node as $\frac{1+a}{2a}\mathbf{f}_{i,h_i} - \frac{1-a}{2a}\mathbf{g}_{i,h_i}$ (and $\left(\frac{d}{2a} + \frac{d}{2} + 1\right)\mathbf{f}_{i,h_i} - \left(\frac{d}{2a} - \frac{d}{2} - 1\right)\mathbf{g}_{i,h_i}$), setting $a = b$ in the functions above.

A set of graph-dependent functions can be equivalently expressed in matrix form.

$$\mathbf{W} = \begin{bmatrix} f_{1,h_1}(v_1) & \ldots & f_{1,h_1}(v_n) \\ \vdots & \ddots & \vdots \\ f_{n,h_n}(v_1) & \ldots & f_{n,h_n}(v_n) \end{bmatrix}$$

The set of functions $\{\mathbf{f}_{1,h_1}, \ldots, \mathbf{f}_{n,h_n}\}$ forms a basis iff the matrix $\mathbf{W}$ is invertible. The expansion and reconstruction can now be written as $\mathbf{y} = \mathbf{W}\mathbf{x}$ and $\mathbf{x} = \mathbf{W}^{-1}\mathbf{y}$. Notice that $\mathbf{W}$ has the sparsity structure of the h-hop connectivity matrix $\mathbf{H}$

of the graph, where $\mathbf{H}_{ij} = 1(v_j \in \mathcal{N}_{h_i}(v_i))$ and $\mathbf{W}_{ij} = 0$ iff $\mathbf{H}_{ij} = 0$.

Fact 1. *For any nonnegative integers $h_1, \ldots, h_n \in \mathbb{N}$, the set of Weighted Average (or Weighted Difference) graph-dependent functions $\{\mathbf{f}_{1,h_1}, \ldots, \mathbf{f}_{n,h_n}\}$ (or $\{\mathbf{g}_{1,h_1}, \ldots, \mathbf{g}_{n,h_n}\}$) defined above, forms a basis for $\mathbb{R}^n$ over any finite undirected graph $G = (V, E)$, $|V| = n$,*

Proof. For any undirected graph $G = (V, E)$, with $|V| = n$, define a stochastic nonnegative matrix $\mathbf{A}$ which computes the average on local neighborhoods: $\mathbf{A_{ij}} = \frac{1}{|\mathcal{N}_h(v_i)|}$ if $v_j \in \mathcal{N}_h(v_i)$; 0 otherwise. Since $\mathbf{A1} = \mathbf{1}$, 1 is an eigenvalue of $\mathbf{A}$ with corresponding positive eigenvector $\mathbf{1}$. By the Perron-Frobenius theorem for nonnegative matrices [12], the spectral radius $\rho(\mathbf{A}) = 1$. Since $\mathbf{A}$ has strictly positive diagonal entries, define $\hat{\mathbf{F}} = \mathbf{A} + c\mathbf{I}$ and $\hat{\mathbf{G}} = \mathbf{A} - c\mathbf{I}$ for some constant $c > 1$. Thus 0 is not an eigenvalue of either $\hat{\mathbf{F}}$ or $\hat{\mathbf{G}}$, and hence both $\hat{\mathbf{F}}$ and $\hat{\mathbf{G}}$ are invertible. Multiplying by a scalar does not change the rank of a matrix, and thus the stochastic matrices $\mathbf{F} = \frac{1}{1+c}\hat{\mathbf{F}}$ and $\mathbf{G} = \frac{1}{1-c}\hat{\mathbf{G}}$ are invertible. □

Thus we can easily construct locally supported graph dependent bases. We defined two simple bases that allow the querier to obtain averages or detect anomalies, at different resolutions corresponding to the neighborhood hop-size. Our framework guarantees a basis set for any communication connectivity graph, and for any choice of neighborhood hop-size parameters $\{h_1, \ldots, h_n\}$. In fact each sensor node can choose the neighborhood size parameter of its encoding basis function randomly and independently from other nodes. Thus our framework allows us to now design a robust distributed encoding algorithm which allows sensors to operate independently without global coordination or knowledge. The degree distribution according to which each node chooses its neighborhood size then determines the probabilistic model for multiresolution querying, as well as the encoding communication cost.

3. RANDOM DISTRIBUTED ENCODING ALGORITHM

Distributed Encoding Algorithm. Each node in the sensor network computes and stores its encoding coefficient according to the following algorithm:

- Encoding node v_i randomly and independently chooses a neighborhood size H_i according to an encoding degree distribution $p(h)$.
- Encoding node v_i sends a request for data to all data nodes v_j in its H_i-hop neighborhood, down the minimum spanning tree.
- Data nodes v_j compute and send the aggregated sum of their data back up the tree to v_i.
- Encoding node v_i then stores the coefficient computed using the graph-dependent basis function $\mathbf{f}_{i,H_i}$ (or $\mathbf{g}_{i,H_i}$).

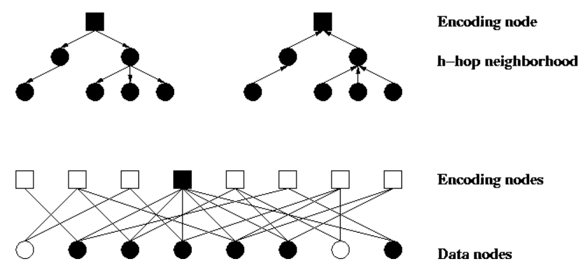

Fig. 2. The encoding algorithm randomly generates a bipartite graph between encoding nodes and data nodes.

We now describe each of the encoding steps in detail to show that this encoding algorithm can be implemented in a distributed and robust way. The first step requires the design of an appropriate degree distribution $p(h)$, which is key in controlling the tradeoff between the communication cost and the querying availability. Each encoding node v_i sends to its 1-hop neighbors, a request for data (v_i, H_i, hop-count), where v_i serves as a request ID, H_i is the hop parameter chosen iid $\sim p(h)$, and hop-count $= 0$. If the 1-hop neighbors have not received the data request ID previously, and hop-count $\leq H_i$, they augment the hop-count and send the data request to their 1-hop neighbors. The process repeats until hop-count $= H_i$. This process thus transmits a single data request down the minimum spanning tree between the encoding node v_i and all the data nodes in its H_i-hop neighborhood. The communication cost per encoding node of this data request is thus $O(|\mathcal{N}_{H_i}(v_i)|)$, the number of nodes in the H_i-hop neighborhood of v_i.

When the data request hop-count reaches H_i, the leaf data nodes of the tree send their data back up the tree to their parents. All intermediate nodes in the tree compute the aggregate sum of their data with their children's data, and sends the aggregate to their parents. The process repeats until the root node v_i is reached. Finally sensor node v_i can compute and store its coefficients $\mathbf{f}_{i,H_i}$ and $\mathbf{g}_{i,H_i}$. This process thus transmits a single data symbol on each link of the spanning tree, again at a cost of $O(|\mathcal{N}_{H_i}(v_i)|)$.

Note that the minimum spanning tree can be maintained in a distributed manner by each node remembering only its parent in the tree. Each sensor is the root of one tree, and a child in many trees. The number of trees it is a member of, in other words the source node degree distribution, is induced by the encoding node degree distribution. The average degree can thus be made small enough. Furthermore, the decoder needs to know the encoding graph in order to decode, by inverting the matrix $\mathbf{W}$. It is sufficient for each encoding node to encode a header containing the indices of the data nodes it encoded. This adds an overhead cost of $O(E(p(h)))$, which should be small relative to the data packet size.

Since our encoding functions form a basis on *any* communication graph, this algorithm does not require a deterministic tree structure around each node. Thus our encoding algorithm is robust to link or node failures. This encoding algorithm can be represented by a bipartite graph between the data nodes and the encoding nodes, where each sensor is both a data node and an encoding node, by flattening the h-hop neighborhood trees. The communication cost is proportional to the number of edges in the bipartite graph.

The multiresolution querying can now be modeled probabilistically, as a function of the encoding degree distribution $p(h)$. Suppose a querier wants to obtain a scale-h coefficient, corresponding to an h-hop neighborhood. Each queried node has a scale-h coefficient with probability $p(h)$. Let Y denote the number of nodes queried until the decoder receives a scale-h coefficient. Thus Y is a geometrically distributed random variable with mean $\mu = \frac{1}{p(h)}$ and variance $\sigma^2 = \frac{1-p(h)}{(p(h))^2}$. Since $P(Y \leq y) = 1 - (1 - p(h))^y$, the probability of not receiving the desired scale coefficient exponentially decays. The querier receives a scale-h coefficient with probability $1 - (1 - p(h))^y$ by querying y nodes.

The communication cost depends on the encoding degree distribution $p(h)$ and the communication graph. The communication cost per encoding node is $O(|\mathcal{N}_{H_i}(v_i)|)$, and thus the overall communication cost of our encoding algorithm is $O(\sum_{i=1}^{n} |\mathcal{N}_{H_i}(v_i)|)$.

Since the size of the h-hop neighborhood depends on the communication graph, we compute the average communication cost per node for randomly generated graphs. $E[|\mathcal{N}_{H_i}(v_i)|] = \sum_{h=0}^{R} E[|\mathcal{N}_{H_i}(v_i)| \mid H_i = h]P(H_i = h)$. Let $F(h) = E[|\mathcal{N}_{H_i}(v_i)| \mid H_i = h]$. If the network communication graph is connected, as h goes to the maximum degree R, $F(h)$ goes to n. When $h = 0$, $F(h) = 1$. $F(h)$ for $0 < h < R$ is a monotonically increasing function, and the shape depends on the communication graph. To achieve an encoding communication cost per node of $c(n)$, a sufficient condition on the degree distribution $p(h)$ is for $E(p(h)) \leq F^{-1}(c(n))$.

Our encoding algorithm uses the following encoding degree distribution $p(h)$, which allows us to traverse the tradeoff between communication cost and querying availability. For $h = 0, \ldots, R$, $p(h) = \frac{1}{c}(\alpha_1 e^{-\lambda_1 h} + \alpha_2 e^{-\lambda_2 (R-h)})$, where α_1 and α_2 control the height of the peaks of the exponentials, λ_1 and λ_2 control the decay rates, R is the largest degree, and c is the normalizing constant.

4. DISCUSSION AND FUTURE WORK

Our multiresolution querying model is based on a distribution on the neighborhood hop-sizes of the local bases. However since we have made no assumptions on the communication connectivity graph, a large hop parameter may not correspond to a large number of nodes in the network, depending on the connectedness of the graph. If the network connectivity is changing in time, we may be able to obtain more consistent resolution scaling by using larger time windows.

As future work, we want to define significance querying models that are data driven, so that the most significant encoding coefficients are readily available everywhere in the network.

5. REFERENCES

[1] J.M. Hellerstein, W. Hong, S. Madden, and K. Stanek, "Beyond average: Towards sophisticated sensing with queries," in *Proc of IPSN*, 2003.

[2] D. Ganesan, B. Greenstein, D. Perelyubskiy, D. Estrin, and J. Heidemann, "An evaluation of multi-resolution storage for sensor networks," in *Proc of the ACM SenSys Conf*, 2003.

[3] D. Ganesan, D. Estrin, and J. Heidemann, "Dimensions: Why do we need a new data handling architecture for sensor networks?," *ACM SIGCOMM Computer Communications Review*, vol. vol. 33, no. 1, pp. 143–148, 2003.

[4] J.M. Hellerstein and W. Wang, "Optimization of in-network data reduction," in *Proc of DMSN*, 2004.

[5] R. Wagner, S. Sarvotham, and R. Baraniuk, "A multiscale data representation for distributed sensor networks," in *Proc of IEEE ICASSP*, 2005.

[6] R. Wagner, H. Choi, R. Baraniuk, and V. Delouille, "Distributed wavelet transform for irregular sensor network grids," in *Proc of IEEE Workshop on Stat Signal Processing*, 2005.

[7] M. Alanyali, S. Venkatesh, O. Savas, and S. Aeron, "Distributed bayesian hypothesis testing in sensor networks," in *Proc of American Control Conf*, 2004.

[8] D.S. Scherber and H.C. Papadopoulos, "Locally constructed algorithms for distributed computations in ad-hoc networks," in *Proc of IPSN*, 2004.

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

[10] V. Saligrama, M. Alanyali, and O.Savas, "Asynchronous distributed detection in sensor networks," *submitted to IEEE Trans on Signal Processing*, 2005.

[11] A.G. Dimakis, V. Prabhakaran, and K. Ramchandran, "Ubiquitous access to distributed data in large-scale sensor networks through decentralized erasure codes," in *Proc of IPSN*, 2005.

[12] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, NY, 1985.