Localization of Wireless Sensors using Compressive Sensing for Manifold Learning

Chen Feng[1,2], Shahrokh Valaee[1], Zhenhui Tan[2]
[1] Department of Electrical and Computer Engineering, University of Toronto
[2] State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University
Email: {chenfeng, valaee}@comm.utoronto.ca, zhhtan@center.njtu.edu.cn

***Abstract*— In this paper, a novel compressive sensing for manifold learning protocol (CSML) is proposed for localization in wireless sensor networks (WSNs). Inter-sensor communication costs are reduced significantly by applying the theory of compressive sensing, which indicates that sparse signals can be recovered from far fewer samples than that needed by the Nyquist sampling theorem. We represent the pair-wise distance measurement as a sparse matrix. Instead of sending full pair-wise measurement data to a central node, each sensor transmits only a small number of compressive measurements. And the full pair-wise distance matrix can be well reconstructed from these noisy compressive measurements in the central node, only through an ℓ_1-minimization algorithm. The proposed method reduces the overall communication bandwidth requirement per sensor such that it increases logarithmically with the number of sensors and linearly with the number of neighbors, while achieves high localization accuracy. CSML is especially suitable for manifold learning based localization algorithms. Simulation results demonstrate the performance of the proposed protocol on both the localization accuracy and the communication cost reduction.**

Keywords- Compressive Sensing, Manifold Learning, CSML, localization, Wireless Sensor Networks

I. Introduction

Accurate and low cost sensor localization is one of the fundamental and crucial challenges in Wireless Sensor Networks (WSNs) [1]. In various applications, including environmental monitoring, vehicle tracking and emergency response, it is often necessary to know each sensor's location in advance for its data to be meaningful [2]. In addition, location-awareness routing protocols can save energy significantly by eliminating route discovery. In designing protocols for localization, low communication costs with a high level of accuracy should be especially considered.

Manifold learning (ML)-based algorithms (e.g, multidimensional scaling (MDS) [3][4], and isometric mapping (Isomap) [5]) formulate the localization problem from pair-wise measurements as a dimensionality reduction problem on a Riemann manifold. Compared with other measurement-based algorithms (e.g, time-of-arrival (TOA) or angle-of-arrival (AOA) measurements of ultra-wideband (UWB) [6]), ML-based algorithms avoid expensive devices, since the only requirement for learning in a central node is pair-wise measurements, which could be any of the physical readings that indicate distance information among sensors, such as the received signal strength (RSS) or the hop-count. However, with the increasing number of sensor nodes, the scale of pair-wise measurements becomes very large. Communication cost between each sensor node and the central node is a bottleneck in these cases. Patwari *et al.* [7] showed the accuracy and robustness of ML algorithms, but neglected the large communication costs when obtaining the pair-wise measurements by assuming them to be known.

In [8], it has been realized that, with the help of choosing landmark sensor nodes, a computationally efficient approximation to MDS-based algorithms can be achieved. The fundamental insight was that, MDS was only applied on a few landmark points. Then, a relative position map was obtained with high accuracy, by formulating the localization problem of rest of the sensor nodes as a triangulation problem. However, due to the constraints of getting connectivity between each sensor node to these landmark nodes, as well as getting large pair-wise measurements among them, this method is still energy exhausted to be applied in large scale networks.

Shang *et. al.* [9] use an MDS-MAP(P) strategy as a variant of MDS to solve the communication cost problem by using patches of relative maps. The main idea is to build a local map at each node of the immediate vicinity and then merge these maps together to form a global map. However, since each node computes a local MDS algorithm individually, it leads to high error accumulation, large energy costs for measuring and computing at each sensor side, as well as large delay from the root node to the last leaf node.

In this paper, with the goal of addressing this open but important problem, our original contributions are as follows. First, we propose compressive sensing for manifold learning *(CSML)*, a new approach to reduce

This work was supported by the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, under project No.863 (2007AA01Z277).

the communication cost. In this algorithm, each sensor transmits only a small number of compressive measurements to a central node by a local random projection. The central node reconstructs a sparse pair-wise distance matrix, through an ℓ_1-minimization algorithm. The overall communication bandwidth requirement per sensor is reduced logarithmically. Second, we derive the least sparsity level that could be achieved in ML-based algorithms, under different large network densities. Simulation results demonstrate the performance of CSML on both the localization accuracy and the communication cost reduction. For simplicity, our methods are illustrated with 2-D sensor networks and they can easily be extended to 3-D cases under the same methodology.

The remainder of this paper is organized as follows. Problem statements and mathematical models are described in Section II. In Section III, we describe the CSML algorithm, presenting and analyzing the idea of reducing communication cost based on the CS theory, including the measurement stage at each sensor node and the learning stage at the central node. The effectiveness of CSML is demonstrated through simulations in Section IV. Finally, Section V concludes the paper.

II. Problem Statement

We consider a scenario that sensors are randomly deployed on an area, taking distance measurements (by RSS readings) from their neighbors, and passing the measured data to a central node through multi-hops, as Fig. 1 illustrates. Based on the RSS measurements, the goal is to determine locations of all sensor nodes simultaneously in the central node with a high level of accuracy, while reducing the total communication costs within the network.

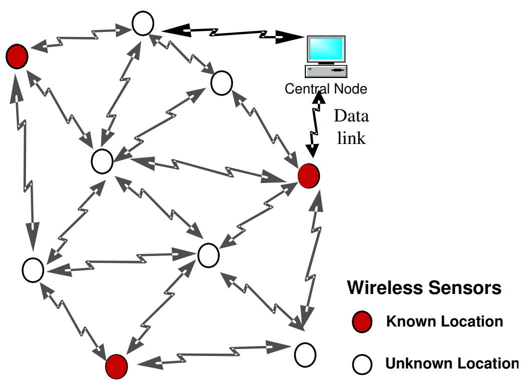

Fig. 1. The scenario of localization in WSNs.

We assume that the number of unknown location sensors is p, while the number of anchor nodes with known locations is q, with $q \ll p$, and $p + q = n$. The known locations of anchor nodes are assumed to be: $[(x_{p+1}, y_{p+1}), (x_{p+2}, y_{p+2}), ..., (x_{p+q}, y_{p+q})]$. The pair-wise distance measurement matrix is represented by: $D = [D_1, ..., D_i, ..., D_n], i = 1, 2, ..., n$, where $[D_i]_{n\times 1} = [d_{i1}, d_{i2}, ..., d_{ij}, ..., d_{in}]^T, j = 1, 2, ..., n$, where A^T returns the transpose of matrix A. The random projection operator $\Phi_{m\times n}$ is a matrix with *i.i.d* Gaussian random entries, with $m \ll n$.

The objective is to determine the physical positions of the n sensor nodes simultaneously in a central node, *i.e.*, $[(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)]$, with a high level of accuracy, while reducing the communication cost between each sensor node and the central node significantly.

III. CSML: Protocol Design

Manifold learning algorithms (*e.g*, MDS, Isomap) are based on a full pair-wise measurement matrix, which indicates the distance information among sensor nodes. However, they ignore the large communication cost for obtaining this matrix in large scale networks. In this paper, we introduce a new protocol, referred to as CSML, to solve this problem. Based on the CS theory, it uses an ℓ_1-minimization algorithm to recover a sparse pair-wise distance matrix.

The block diagram of CSML is shown in Fig. 2 with two stages. On the measurement stage, random projections are conducted locally at each sensor node, and only a small number of noisy compressive measurements are transmitted to the central node. In the learning stage, a sparse pair-wise distance matrix, indicating neighboring connectivity, is reconstructed at the central node through an ℓ_1-minimization algorithm, and Isomap is applied on this reconstructed matrix to learn the locations of all the sensor nodes.

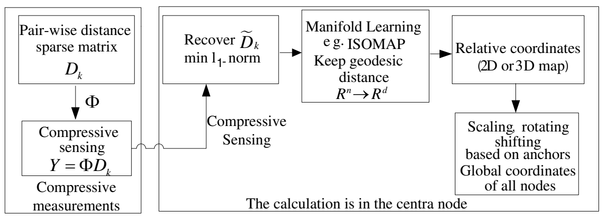

Fig. 2. The framework of CSML protocol.

A. Measurement stage

Let $D_k \in R^n$ be a sparse representation of the full pair-wise distance matrix D. $D_k = [D_{k1}, D_{k2}, ..., D_{ki}, ..., D_{kn}]$, where each D_{ki} is a k-sparse vector, namely, $||D_{ki}||_0 \le k$, $i \in \{1, 2, ..., n\}$. Thus, for each sensor node i, D_{ki} represents the pair-wise distance measurement from its k nearest neighbors, and leaves the other entries to be zeros. It is noted that the positions of the k non-zero entries in the sparse vector are unknown.

According to the CS theory, noisy compressive measurements in an m-dimensional space are needed at each sensor node, with $m \ll n$. The CS theory indicates that, rather than designing a sensor to measure a signal $[D_i] \in R^n$, it often suffices to measure a much smaller vector $Y_i = \Phi D_{ki}$ [10] to [11]. If Φ is properly designed, from only m noisy compressive measurements, it is possible to recover D with the quality comparable to its proximity to the nearest k-sparse signal (D_k). If D itself is k-sparse, then the reconstruction is almost exact. The quality of the recovered signal is as good as if the position of the k largest values of each D_i is known ahead of time and can be measured directly. Based on the CS theory, the number of compressive measurements m should be larger than $c_0 k \log(n/k))$, where c_0 is a constant. Our simulations show that with $c_0 = 1$, we can estimate the locations of sensors with a very small error.

Compressive measurements are obtained by taking incoherent projections on the original signal.

$$Y_i = \Phi D_{ki} + \varepsilon_i, Y_i \in R^m, \forall i = 1, 2, ..., n \tag{1}$$

where $\Phi_{m \times n}$ is the measurement matrix, which is properly designed such that it is incoherent with the signal basis, and ε_i is the measurement noise. It has been shown that incoherence holds with very high probability between an arbitrary basis and a random matrix with *i.i.d.* Gaussian distributed entries [10]. Therefore, we choose Φ as an $m \times n$ random matrix with *i.i.d.* Gaussian distributed entries for applying the CS theory.

By multiplying a random projection $\Phi_{m \times n}$ locally on the measurement vector at each sensor node, a noisy compressive vector Y_i is measured and sent to the central node. Compared to sending a vector of $[D_i]_{n \times 1}$, or sending k measurements from k nearest neighbors with large overheads to indicate the nodes, only an $m \times 1$ vector Y_i is transmitted, where $m = O(k \log(n/k))$. Note that $m \ll n$.

B. Learning stage

By applying the CS theory, a sparse pair-wise distance matrix could be recovered at the cental node based on the random seeds $\Phi_{m \times n}$ through an ℓ_1-minimization algorithm. Thus, the traditional ill-posed recovery problem through far fewer samples is solved here only by using the following linear algorithm, which is effectively solved in polynomial time.

$$\begin{aligned} \tilde{D_k} = arg \min_{\tilde{D_k} \in R^n} \parallel \tilde{D_k} \parallel_1, \\ s.t.\ \tilde{Y} = \Phi D_k + \epsilon. \end{aligned} \tag{2}$$

Minimization (2) is the Basis Pursuit (BP) algorithm [12], which formulates the problem with equality constraints, and solves the problem by a primal-dual interior point method. Furthermore, the recovery error of signal $\tilde{D_k}$ compared to the full pair-wise signal D is proved to be bounded as [10]:

$$\parallel D - \tilde{D_k} \parallel_2 \leq c_1 k^{-1/2} \parallel D - D_k \parallel_1 + c_2 \varepsilon \tag{3}$$

where c_1 and c_2 are constants, and ε is the measurement noise.

Next, the ML-based algorithm (Isomap) can be applied on this well recovered $\tilde{D_k}$ matrix, and this process is undertaken in the central node. Basically, three steps of Isomap are conducted.

1) A geodesic distance matrix D'_k is computed based on the recovered sparse pair-wise distance matrix $\tilde{D_k}$, by using either Dijkstra's or Floyd's shortest path algorithm. The near-zero entries in $\tilde{D_k}$ are set to a large value which implies large distances.
2) Classical MDS algorithm is applied on the geodesic distance matrix D'_k;
   a) We square and double centeralize matrix D'_k, and obtain the Gram matrix B by: $B = -\frac{1}{2} J D'^2_k J$, where $J = I_{n \times n} - \frac{1}{n} e \times e^T$, and $e = (1, 1, ..., 1)^T$;
   b) We eigen-decompose matrix B, *i.e.*, $B = UVU^T$, and keep the largest d positive eigenvalues (V_i) and the first d columns of eigenvectors $(U_i), i = \{1, 2, ..., d\}$;
   c) Finally, we learn the relative low dimensional coordinates by: $(x_i, y_i) = U_i V_i^{1/2}, i = \{1, 2, ..., d\}$.
3) The output of step 2 is a 2-D or 3-D relative position map, thus global position of all the sensor nodes is obtained by a position alignment, *i.e.*, mapping the relative coordinates into global coordinates through scaling, rotating and shifting based on the prior location knowledge of the q anchor nodes.

IV. Simulations

The effectiveness and properties of CSML are studied through simulations. Sensors are randomly placed in a unit square region. The number of sensors is 50, 100, and 200 respectively. Sensors only hear from their neighbors, while the data transmission between sensors and the central node can take place through multi-hops.

The ML error with respect to the number of nearest neighbors is computed to determine the necessary number of neighbors (parameter k), under different network densities. The ML error is defined by the residual variance, which is the sum of residual eigenvalues in step 2 of Sec. III-B. It represents how well the Isomap is applied for localization.

$$R = \sum_{i=d+1}^{rank(B)} V_i. \tag{4}$$

The recovery rate of CS is defined as the ratio of the number of correctly recovered non-zero entries in $\tilde{D_k}$

to the overall number of non-zero entries in D_k. A non-zero element in $\tilde{D}_k$ is assumed to be correctly recovered if $|d_{ij} - \tilde{d_{ij}}| < \delta$, where δ is a design parameter, (in our simulation, $\delta = 1e-3$).

$$r = c/C \tag{5}$$

$$c = \begin{cases} 1 & \text{if } |d_{ij} - \tilde{d_{ij}}| < \delta \\ 0 & \text{else} \end{cases}.$$

Meanwhile, the accuracy of the overall CSML localization algorithm is measured by averaging the Euclidean distances between the estimated locations and their true locations over the whole area:

$$P_e = \frac{1}{n} \sum_{i=1}^{n} \sqrt{(x'_i - x_i)^2 + (y'_i - y_i)^2}. \tag{6}$$

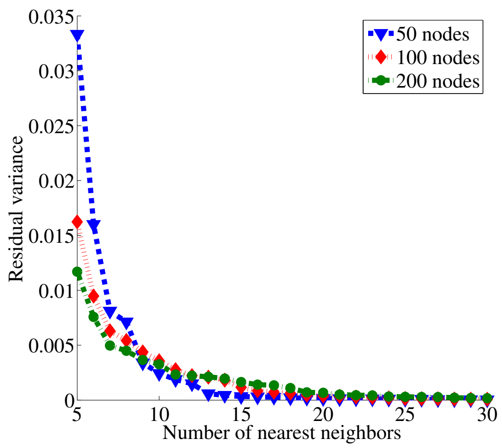

Fig. 3. Residual variances with respect to the number of nearest neighbors.

Figure 3 shows residual variances with respect to the number of nearest neighbors under different network densities, from 50 to 200 nodes. Under these situations, residual variances converge to a small number only when the number of neighbors is larger than a certain number. This turning point could be considered as the least sparsity level (k) achieved with which ML works well. Thus, when designing random seeds $\Phi_{m\times n}$, where $m = O(k\log(n/k))$, the smallest k is constrained by ML algorithms. We will set $k = 10$ to achieve a good performance by ML in the following simulations.

Figure 4 illustrates the recovery rates of the CS algorithm with respect to the number of compressive measurements (m) under k nearest neighbors ($k = 10$ according to Fig. 3) under different scales of networks. Results indicate that using far fewer number of measurements, it is possible to recover an approximate full pair-wise measurements matrix with extremely high recovery rate. To recover at least 90% of the pair-wise matrix, the communication costs reduced by CSML reaches

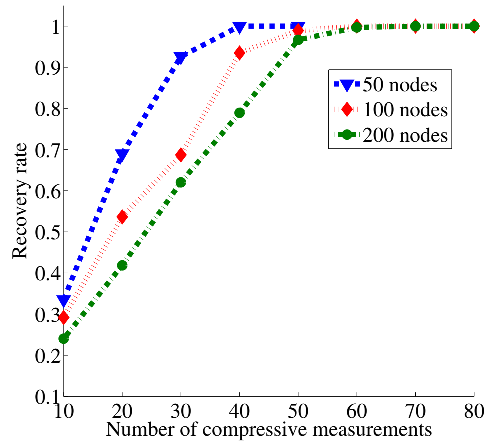

Fig. 4. Recovery rates with respect to the number of compressive measurements.

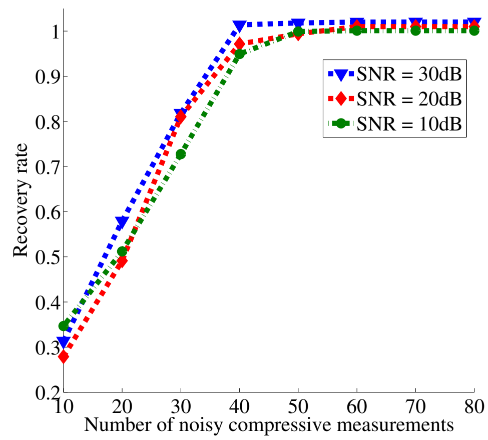

Fig. 5. Recovery rates under noisy compressive measurements.

40%, 60% and 75% respectively, under 50, 100 and 200 sensor nodes. The larger scale of the network, the more percentage of reduction of the communication cost. The result also holds in noisy environment. Figure 5 illustrates the corresponding recovery rate under noisy measurements in the scenario of 100 sensor nodes with 10 neighboring connectivities. Signal-to-noise (SNR) varies from 10dB to 30dB, which is defined as the ratio of the transmit signal power to the noise power at the receiver. Noise is assumed to follow Gaussian distribution in the simulation.

Figure 6 shows the intuitive effectiveness of CSML algorithm by locating 100 randomly deployed nodes in a unit square area. The result is an average output of 100 experiments, each of which is under 50% compressive measurements from 10 nearest neighbors. Green stars represent the true locations of these 100 nodes, while red circles represent the output of the whole CSML algorithm.

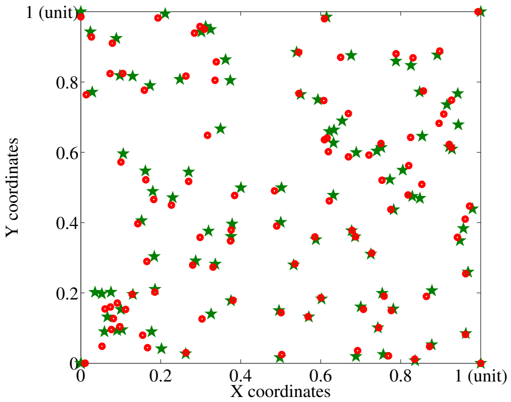

Fig. 6. CSML localization result under 100 random nodes.

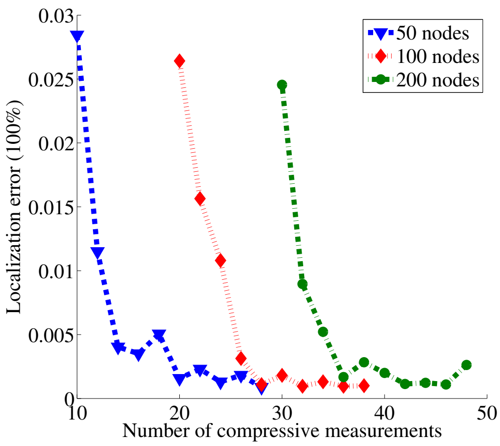

Fig. 7. The localization error with respect to the number of compressive measurements.

Figure 7 shows the statistical results of the CSML localization accuracy with respect to the number of compressive measurements that each sensor node takes. With sufficient number of measurements ($m \ll n$), localization error can be less than 1% (defined in (6), which is well acceptable for corresponding location-based services.

In classical Isomap approach, each sensor transmits $[D_i]_{n\times 1}$ measurements to the central node, which brings large communication cost problem, especially in large scale networks; However, by applying the CS theory, compared to the ambient dimension n, CSML reduces the overall communication bandwidth requirement per sensor, such that it increases logarithmically with the number of sensors and linearly with the number of neighbors. Meanwhile, with sufficient number of compressive measurements, CSML achieves an extremely high localization accuracy by applying Isomap directly on the reconstructed sparse matrix.

V. Conclusion

In this paper, we first show that the CS theory could be applied to solve the large communication cost problem in the ML-based localization algorithms. The CSML protocol is proposed to formulate the pair-wise measurements among neighbor sensors as a sparse matrix, and an approximate full pair-wise matrix from noisy compressive measurements is reconstructed at the central node, only through an ℓ_1-minimization algorithm. It is shown that CSML significantly reduces the transmission costs, while maintaining a high level of localization accuracy. CSML is especially suitable for ML-based localization algorithms in wireless sensor networks.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communication Magazine*, vol. 40, no. 8, pp. 102–114, August 2002.

[2] N. Patwari, J. N. Ash, and S. Kyperountas, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, pp. 54–69, July 2005.

[3] J. A. Costa, N. Patwari, and A. O. H. III, "Distributed weighted-multidimensional scaling for node localization in sensor networks," *ACM Transactions on Sensor Networks*, vol. 2, pp. 39–64, February 2006.

[4] P. E. Green, F. J. Caromone, and S. M. Smith, "Multidimensional scaling: Concepts and applications," *Newton*, 1989.

[5] J. B. Tenenbaum, V. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, pp. 2319–2323, 2000.

[6] A. Catovic and Z. Sahinoglu, "The cramer-rao bounds of hybrid toa/rss and tdoa/rss location estimation schemes," *Technical Report, Mitsubishi Electric Research Lab*, Janurary 2004.

[7] N. Patwari and A. O. H. III, "Manifold learning algorithms for localization in wireless sensor networks," *IEEE Signal Processing Magazine*, pp. 54–69, July 2005.

[8] V. Silva and J. B. Tenenbaum, *Sparse Multidimensional Scaling using Landmark Points*. Technical Report, Stanford, 2004.

[9] X. Ji and H. Zha, "Sensor positioning in wireless ad-hoc sensor networks with multidimensional scaling," *IEEE INFOCOM*, pp. 2652–2661, 2004.

[10] J. C. Emmanuel and B. W. Michael, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, pp. 21–30, March 2008.

[11] R. G. Baraniuk, M. Davenport, R. Devore, and M. B. Wakin, "A simple proof of the restricted isomatry property for random matrices," *Constructive Approximation*, 2008.

[12] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, pp. 33–61, 1998.