Dictionary Based Reconstruction and Classification of Randomly Sampled Sensor Network Data

Grigorios Tsagkatakis*, and Panagiotis Tsakalides*†

*Institute of Computer Science (ICS) - Foundation for Research & Technology - Hellas (FORTH), Crete, Greece
†Department of Computer Science, University of Crete, Heraklion, Crete, Greece
e-mail: {greg, tsakalid}@ics.forth.gr

Abstract—In this paper, we propose a method for recovering and classifying WSN data while minimizing the number of samples that need to be acquired, processed, and transmitted. The problem is formulated according to the recently proposed framework of Matrix Completion (MC), which asserts that a low rank matrix can be recovered from a small number of randomly sampled entries. The application of MC in WSN data is motivated by the assumption that sensory data exhibit intra-sensor correlations and that these data can be represented using known examples. We formulate the problem as that of recovering the low rank measurement matrix by encoding the contributions of known examples, the dictionary elements, for reconstructing and classifying the data. Experimental results using artificial data suggest that the proposed scheme is able to accurately reconstruct and classify the sensory data from a small number of measurements.

I. INTRODUCTION

Wireless Sensor Networks (WSNs) serve as a novel framework for data acquisition, where a large number of small and inexpensive sensors record a set of environmental variables and communicate them to a base station for processing. WSNs have been used in numerous applications with very promising results owing to their low cost, their ability to cover large areas and flexible deployment. Despite their many applications, WSNs are characterized by three major limitations, namely low processing capabilities, communication bandwidth constraints, and limited power. Given these limitations, the objective of algorithms developed for such systems is to necessitate moderate processing capabilities and to operate at low transmission rates.

To address these issues, the recently proposed framework of Compressed Sensing (CS) [1], [2] has been investigated and different applications have been proposed. According to the CS framework, perfect reconstruction of a signal can be obtained by recording a compressed version of the data, provided the data can be sparsely represented in a dictionary of atoms. More importantly this compression can be achieved by applying a linear transform via a matrix consisting of random values, drawn from an appropriate distribution, a property termed universal dimensionality reduction. Decoding and reconstruction of the original signal is achieved by solving an l_1 minimization problem or by greedy approaches.

Even though the theory of CS has only been recently proposed, it has be utilized in various applications in the field of WSN. This can be attributed to the key characteristics of CS, such as asymmetric complexity, universal encoding, and robustness to noise, among others. The majority of papers employing CS in WSN has focused on reconstructing the randomly projected data at the base station.

Although reconstructing the measurements in the base station is important, many applications are focused on the identification of specific patterns. For example, consider a WSN monitoring an environmental variable such as humidity over a large area. For applications such as farming and water quality monitoring, identifying patterns, recognizing events, and classifying measurements are equally important to signal reconstruction. In these scenarios, classification schemes could be used to generate complex trigger events that request the sensors to sample at higher rates or actuators to take appropriate actions.

In this work, we propose a new scheme based on the recently developed framework of Matrix Completion (MC) for reconstruction and classification of WSN data. To achieve this goal, we propose the exploitation of the inherent correlations between the sensed data and a dictionary of known examples, in order to reconstruct the original measurements or perform classification of the sensed data from a much smaller number of random measurements.

The rest of the paper is organized as follows: Section II discusses the motivation of the proposed scheme while Section III presents a short overview of previous work. Sections IV and V present the formulation of MC and the proposed scheme for the application in WSN data processing. Experimental results are provided in Section VI and the paper concludes in Section VII.

II. MOTIVATION

To motivate this work, we consider the scenario where a number of sensors monitor the environment and report their measurements to a base station where the reconstruction and the classification of the data takes place. To address the limited processing power, the communication constraints, and the energy limitations, we propose to perform random sampling of the environment where each sensor collects measurements at random intervals or where only a random subset of sensors take measurements of environment at every instant.

While random sensing has numerous benefits such as reduction in communication cost, which plays a central role in the life-time of the network [10], recovering such data under

the MC framework is only possible if correlation between the sensed data exists. To illustrate the low rank nature of a WSN measurement matrix, Figure 1 shows an example of data gathered from a simulated WSN of 20 sensors, each one collecting 50 measurements. In this example, the measurements of each sensor where modeled as the summation of a baseline autoregressive signal with Gaussian noise resulting in an average SNR=14dB.
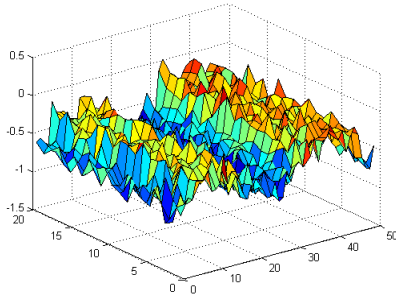

Figure 1.   Recorded signals from the sensor network.

Although the underlying trend in the data is evident, the added noise masks this trend. This effect can be seen in Figure 2 where the corresponding singular values of the measurement matrix are shown. We observe that the gathered measurements matrix can still be characterized by a low rank while the added noise is responsible for additional singular values that exhibit much lower energy.
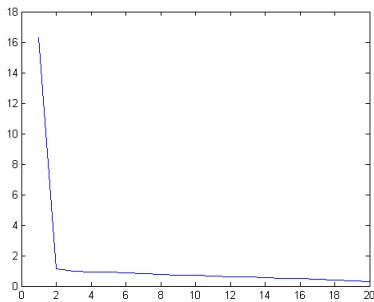

Figure 2.   Singular values of measurement matrix.

III. Previous Work

Communicating sensor data over a wireless network is challenging. On the one hand, data rates should be kept to a minimum to reduce the energy consumption of the communication, while on the other hand, the limited processing power of the nodes and the distributed nature of the system, prevent the application of sophisticated compression schemes. Recent techniques try to achieve a balance between these constraints by performing simultaneous sensing and compression via the Compressed Sensing framework.

Methods such as [3], [4], [5] and [6] exploit the correlation between the sensor data in an effort to achieve reconstruction with reduced transmission overhead. Although CS based methods offer high quality reconstruction, they require the application of Random Projections (RP) as an encoding scheme. According to the RPs sampling scheme, each sensor needs to collect the measurements, usually in a vector form, and then multiply them by a random matrix that performs universal dimensionality reduction.

More recently, the CS framework was also utilized for classifying WSN data. In [7], an action recognition system that operated on randomly projected data from body sensors was developed. In order to recognize the activity, the authors proposed the Distributed Sparsity-based Classification (DSC) framework, where the sensor measurements were first projected into a low dimensional space via the application of Random Projections and then classified locally by an ℓ_1 minimization classification scheme on a dictionary of examples.

The above scheme differs significantly from the proposed one where Random Sampling (RS) of raw data measurements takes place. The benefits of RS over RP in the case of WSN is that in addition to reducing communication requirements, RS also reduces sampling requirements. The frequency of sampling may play a critical part in the lifetime of the WSN system, since specific sensor types, such as biosensors, exhibit sensing quality degradation. We discuss the theory of Matrix Completion and RS next.

IV. Low Rank Matrix Recovery

Let $\mathbf{X} = [X_1, X_2, ..., X_s] \in \mathbb{R}^{n \times s}$ be a measurement matrix consisting of s measurements from n different sources. In general, one cannot recover the ns entries of the matrix $\mathbf{X}$ from a smaller number of k entries, where $k << ns$. However, it has been recently shown that such a recovery is possible when the rank of the matrix is small compared to the dimensions of the matrix. More interestingly, one can recover the matrix $\mathbf{X}$ from $k \geq Cn^{6/5}r log(n)$ random measurements, where $n > s$ and $rank(\mathbf{X}) = r$ [13]. To recover the unknown matrix X from the measurements M, the following minimization problem needs to be solved

$$\min\{\ rank(\mathbf{M})\ :\ \ \mathcal{A}(\mathbf{X}) = \mathcal{A}(\mathbf{M})\} \qquad (1)$$

where $\mathcal{A}$ is a linear map from $\mathbb{R}^{ns} \to \mathbb{R}^k$. In the case of MC, $\mathcal{A}$ is defined as a random sampling operator that records a small number of entries from the matrix. In general, to solve the MC problem, the linear map $\mathcal{A}$ must satisfy a modified Restricted Isometry Property, which is the case for uniform Random Sampling in both rows and columns [8].

Although solving the above problem will generate a low rank matrix consistent with the observations, rank minimization is an NP-hard problem and therefore cannot be used in practice. Recently, a relaxation of the above problem [13] was shown to produce arbitrary accurate results, by replacing the rank constraint by the tractable nuclear norm minimization

$$\min\{\ \|\mathbf{M}\|_*\ :\ \ \mathcal{A}(\mathbf{X}) = \mathcal{A}(\mathbf{M})\} \qquad (2)$$

where the nuclear norm is defined as $\|\mathbf{L}\|_* = \sum \|\lambda_i\|_1$ *i.e.* the ℓ_1 norm of the singular values. For the noisy case, an approximate version can be solved [12]

$$\min\{\ \|\mathbf{M}\|_*\ :\ \ \|\mathcal{A}(\mathbf{X}) - \mathcal{A}(\mathbf{M})\|_F^2 \leq \epsilon\} \qquad (3)$$

where $\|\mathbf{X}\|_F^2 = \sum \lambda_i^2$ denotes the Frobenius norm and ϵ is the acceptable approximation error.

V. PROPOSED METHOD

The MC formulation is able to recover the matrix by making no assumption about the process that generates the matrix, except that it is of low rank. In this work, we consider the scenario where the sensors measure variations of a small number of baseline signals and our goal is to identify these baseline signals from a small number of measurements. According to our model, the measurement matrix $\mathbf{X}$ can be described as a linear combination of known examples *i.e.* $\mathbf{X} = \mathbf{DL}$, where $\mathbf{D}$ is a dictionary of examples and $\mathbf{L}$ is a low rank matrix that contains the representation coefficients. Recovery of the representation coefficients is possible by solving

$$\min\{ \ \|\mathbf{L}\|_* \ : \ \|\mathbf{X} - \mathbf{DL}\|_F^2 < \epsilon \} \qquad (4)$$

The above formulation, termed Low Rank Representation (LLR) [16], has been recently applied in subspace segmentation by setting the dictionary as the data matrix *i.e.* $\mathbf{D} = \mathbf{X}$ and seeking the low rank matrix that encodes the contributions.

However, other models for the dictionary can also be used. In our model, we consider a combination of LLR and MC by seeking the low rank presentation coefficients matrix $\mathbf{L}$ from a small number of measurements $\mathcal{A}(\mathbf{X})$. The updated minimization is given by

$$\min\{ \ \|\mathbf{L}\|_* \ : \ \mathbf{Z} = \mathbf{DL}, \ \|\mathcal{A}(\mathbf{X}) - \mathcal{A}(\mathbf{Z})\|_F^2 < \epsilon \} \quad (5)$$

The problem in (5) can be transformed to a semidefinite programming problem and solved using interior point methods [11]. In our experimental section we utilized the CVX [14],[15] packet for solving (5). The proposed formulation bares many analogies to dictionary based CS reconstruction methods where vectors are replaced by matrices and the assumption of sparse representation on a dictionary of examples is reframed as a low rank representation on the dictionary. This model is also related to high dimensional statistics where the problem is similar to a multivariate regression with rank constraints [17].

To classify the data, we follow an approach similar to CS classification where data are classified by selecting the class whose examples introduce the minimum error defined as

$$class(\mathbf{X}) = \min_{\mathbf{j}} \|\mathcal{A}(\mathbf{X}) - \mathcal{A}(\mathbf{Z_j})\|_\mathbf{F}^\mathbf{2} \qquad (6)$$

where $\mathbf{Z_j} \equiv \mathbf{D_j L_j}$ selects only the examples corresponding to class j.

VI. EXPERIMENTAL RESULTS

In this section, we provide two sets of experimental results that illustrate the characteristics of the proposed scheme for reconstruction and classification, respectively. We present simulation results on a network of 10 and 20 sensors, where each sensor must communicate 50 measurements. To generate the sensor data, we modeled a baseline measurement signal according to an autoregressive model, initialized with a random seed, i.e. $\mathbf{Y_{i+1}} = a\mathbf{Y_i} + n$, where a and n are i.i.d from $\mathcal{N}(0, \sigma)$ The simulated measurements were modelled as noisy versions of the baseline signal with various amounts of white

Gaussian noise, simulating different levels of inter-sensor correlations. This approach was employed for generating a dictionary of 70 atoms. In this dictionary, one atom was a noisy version of the baseline signal while the other atoms where drawn according to different models. Each data point in the results is averaged over 20 independent trials.

First, we present the results on measurement matrix reconstruction where we compare the performance of traditional MC with the dictionary driven MC scheme. Figures 3, 4 and 5 present the reconstruction error as a function of sampling rate for SNRs 65dB, 40dB and 14dB, respectively. We observe in the graphs that the dictionary guided MC is able to achieve lower reconstruction error compared to the traditional MC, especially in low sampling rates. Furthermore, the reduced error rates are more pronounced in low SNRs. This observation implies that the dictionary based approach is more robust to noise compared to the traditional MC, due to the additional information that is encoded in the dictionary.
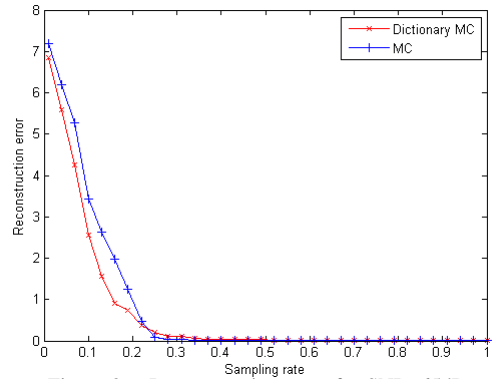

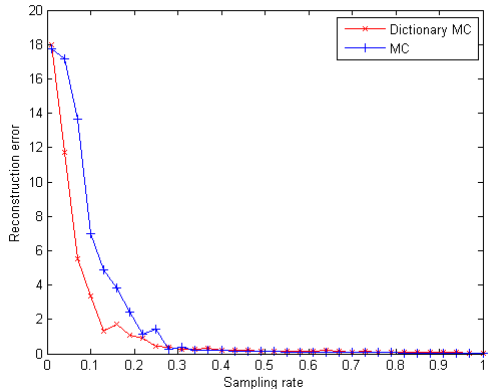Figure 3.   Reconstruction error for SNR=65dB.

Figure 4.   Reconstruction error for SNR=45dB.

In the second set of experiments, we investigated the application of the proposed scheme for the classification of the WSN data. Tables I and II present the classification accuracy for different amounts of noise on networks of 10 and 20 nodes respectively. By classification accuracy we measured the probability of correcting classifying the testing signals. The objective of this experiment was to evaluate the performance of our scheme as a function of both network size and data correlation levels. From these results we can make two basic
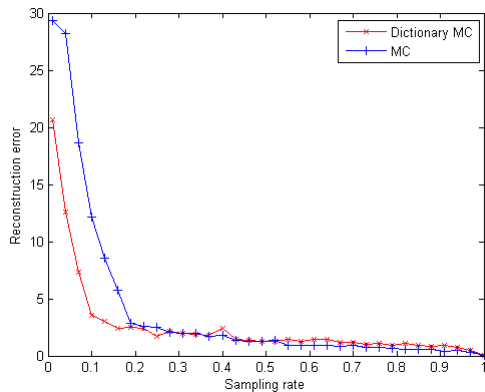
Figure 5. Reconstruction error for SNR=14dB.

observations. By considering the effects of correlation, we observe the the proposed scheme is able to achieve increased accuracy in scenarios where the signals are affected by a small amount of noise and therefore are more correlated, since the noise component is independent of the signals. Regarding network size, we observe that larger networks are more robust to noise which implies that these networks are able to capture data correlation more efficiently. For example very high classification performance can be achieved half the measurements, a reduction in sampling rate that can have significant advantages for resource constrained WSN.

Table I
CLASSIFICATION ACCURACY FOR A NETWORK OF 10 SENSORS

Classification Accuracy			
sampling rate	65dB	40dB	14dB
0.1	0.50	0.05	0.00
0.2	0.60	0.20	0.15
0.3	0.80	0.35	0.25
0.4	0.95	0.80	0.45
0.5	1.00	0.95	0.40
0.6	1.00	1.00	0.55
0.7	1.00	1.00	0.60
0.8	1.00	1.00	0.60
0.9	1.00	1.00	0.70
1	1.00	1.00	0.70

Table II
CLASSIFICATION ACCURACY FOR A NETWORK OF 20 SENSORS

Classification Accuracy			
sampling rate	65dB	40dB	14dB
0.1	0.65	0.05	0.00
0.2	0.90	0.35	0.35
0.3	0.95	0.95	0.65
0.4	0.95	0.95	0.75
0.5	1.00	0.95	0.70
0.6	1.00	0.95	0.80
0.7	1.00	1.00	0.70
0.8	1.00	1.00	0.75
0.9	1.00	1.00	0.70
1	1.00	1.00	0.70

VII. CONCLUSIONS

We presented a sampling scheme for WSN that can significantly reduce both acquisition and communication require-ments by exploiting the correlation between the sensory data. The proposed scheme is based on a variation of the recently proposed framework of Matrix Completion where a set of known examples are used for generating a dictionary and utilizing it for data recovery. Experimental results in simulated data suggest that the proposed scheme can achieve superior performance for the case of reconstruction compared to the traditional MC. Furthermore, results suggest that significant reduction in sampling rate can be obtained with limited effects on classification performance. In future work, validation of the proposed scheme will be applied to real WSN data.

REFERENCES

[1] D.L. Donoho, "Compressed sensing", in *IEEE Trans. on Information Theory*, vol. 52, No. 4, pp. 1289-1306, 2006.
[2] E.J. Candes, and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?", in *IEEE Trans. on Information Theory*, vol. 52, No. 12, pp. 5406-5425, 2006.
[3] M.F. Duarte, M.B. Wakin, D. Baron, and R.G. Baraniuk, , "Universal distributed sensing via random projections", in *ACM Int. Conf. on Information Processing in Sensor Networks*, pp. 177-185, 2006.
[4] L. Chong, W. Feng, S. Jun, and C. Chang Wen "Compressive data gathering for large-scale wireless sensor networks, in *ACM Int. Conf. on Mobile Computing and Networking (MobiCom)*, pp. 145-156, 2009.
[5] G. Tzagkarakis, J.L. Starck, and P. Tsakalides, "Joint Sparse Signal Ensemble Reconstruction in a WSN Using Decentralized Bayesian Matching Pursuit", in *19th European Signal Processing Conf.*, 2011.
[6] J. Haupt, W.U. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data", in *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 92-101, 2008.
[7] A.Y. Yang, R. Jafari, S.S. Sastry, and R. Bajcsy, "Distributed recognition of human actions using wearable motion sensor networks", *Journal of Ambient Intelligence and Smart Environments*, Vol. 1, No. 2, pp. 103–115, IOS Press, 2009.
[8] B. Recht, M. Fazel, and P.A. Parrilo, "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization", *SIAM Review*, vol. 52, No. 471, 2010.
[9] M. Fazel, H. Hindi, and S. Boyd, "A rank minimization heuristic with application to minimum order system approximation", in *Proceedings of the American Control Conference*, Vol. 6, pp. 4734-4739, 2001.
[10] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey", in *Ad Hoc Networks*, vol. 7, No. 3, pp. 537-568, 2009.
[11] Z. Liu, and L. Vandenberghe, "Interior-point method for nuclear norm approximation with application to system identification" in *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3,pp. 1235-1256, 2009.
[12] R. Mazumder, T. Hastie, and R. Tibshirani, "Spectral regularization algorithms for learning large incomplete matrices" in *The Journal of Machine Learning Research*, vol. 11, pp. 2287-2322, 2010.
[13] E.J. Candès, and, B. Recht, "Exact matrix completion via convex optimization" in *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717-772, 2009.
[14] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, version 1.21", 4/2011.
[15] M. Grant and S. Boyd "Graph implementations for nonsmooth convex programs", in *Recent Advances in Learning and Control*, pp. 95-110, 2008.
[16] G. Liu and Z. Lin and Y. Yu "Robust Subspace Segmentation by Low-Rank Representation", in *Int. Conf. Machine Learning*, 2010.
[17] A. Tsybakov, and A. Rohde, "Estimation of High-Dimensional Low Rank Matrices", in *The Annals of Statistics*, vol. 39, No. 2, pp.887-930, 2011.