A Novel Compressive Sensing Based Data Aggregation Scheme for Wireless Sensor Networks

Cheng Zhao *†‡§, Wuxiong Zhang *†‡, Xiumen Yang*†‡, Yang Yang*†‡, and Ye-Qiong Song¶
*Shanghai Institute of Microsystem and Information Technology (SIMIT), CAS
†Shanghai Research Center for Wireless Communications (WiCO)
‡Key Laboratory of Wireless Sensor Network & Communication, SIMIT, CAS
§University of Chinese Academy of Sciences
¶LORIA - University of Lorraine, France

***Abstract*—The random distribution of sensors and the irregularity of routing paths lead to unordered sensory data which are difficult to deal with in Wireless Sensor Networks (WSNs). However, for simplicity, most existing researches ignore those characteristics in the designs of Compressive Sensing based Data Aggregation Schemes (CSDAS). Since conventional sparsification bases (e.g., DCT, Wavelets) are inefficient to deal with unordered data, performances of CSDAS with conventional bases are inevitably constrained. In this work, a novel CSDAS which adopts Treelet transform as a sparse transformation tool is proposed. Our CSDAS is capable to exploit both spatial relevance and temporal smoothness of sensory data. Moreover, our CSDAS contains a novel correlation based clustering strategy which is realized with the localized correlation structure of sensory data returned by Treelets and facilitates energy saving of CSDAS in WSNs. Comparative results show the reconstruction error rate with adopting Treelet transform in CSDAS is about 18% lower than that of conventional ones when the normalized energy consumption is 0.3. Even larger performance gain will be obtained at higher energy consumption level. Meanwhile, simulations results further show that our novel correlation based clustering strategy is of great potential. Specially, there is a gain of roughly 35% for total energy savings with our proposed clustering strategy.**

I. Introduction

In recent years, how to prolong lifetime of WSNs has been a hot research topic. Generally, a WSN consists of one(or few) sink node(s) and a large amount of randomly sowed sensor nodes. The sink node is connected to a wired network to gather sensory data from sensor nodes, which usually have constrained resources. Massive unordered data are generated in WSN applications (e.g., environment monitoring), and thus proper data transmission/processing schemes have to be designed to facilitate applications which are sensitive to network lifetime.

Aiming at prolonging WSNs lifetime, many in-network data aggregation technologies [1] have been proposed via reducing energy consumption or easing computational burden of sensor nodes. Compressive Sensing (CS) [2][3] offers us a novel perspective and is another promising method to prolong WSNs lifetime. CS achieves a precise recovery with far fewer measurements than the dimension of original signal. Meanwhile, CS owes an inherent characteristic that consisting of a low complexity encoder at the cost of a high complexity decoder. This characteristic facilitates the reduction of energy consumption and computational burdens of sensor nodes in WSNs.

Many prior explorations have laid solid foundations for studying CS in WSNs. Early work [4] focuses on exploiting Distributed Compressive Sensing [5] into single-hop sensor networks to reduce energy consumption of each sensor node. However, multi-hop network and in-network data processing are not considered. Work in [6] introduces CS into multi-hop WSNs. Its key contribution is to introduce a practical implementation of compressing records from all sensor nodes through random linear projections. In that work, projections of records are delivered through a simple gossip algorithm based communication scheme which requires large amount of in-network communications. Work in [10] proposes an energy-efficient scheme: each node generates a sparse random vector (sensory record multiplies local sparse random coefficient vector). The vector is added up with that from previous node(s), and then the result is forwarded to its next hop node. This scheme reduces the amount of communications by generating Sparse Random Projections (SRP). However, SRP leads to frequently changed routing paths, which are difficult to achieve in WSNs. Work in [8] designs a Compressive Data Gathering (CDG) scheme by collecting Dense Random Projections (DRP) of signals. This scheme is decoupled with route and topology of nodes. However, work in [9] shows that plain DRP based data aggregation schemes may not be able to reduce the whole energy consumption in tree-based topologies.

All of the aforementioned and other researches contribute a lot to implementations of CS in WSNs, however they take few considerations of how to make full use of the sparsity, which is a prerequisite for CS, of unordered data Work in [11] shows that standard transformations widely used in 2-D image processing are not sufficient for real world sensory data, which are unordered due to the random distribution of sensors and irregularity of routing paths (detailed in [7]). To address this problem, work in [12] exploits Principal Component Analysis (PCA) to discover transformations from training data set. Compared with conventional method and DCT-based CS technology, this scheme is preferable in terms of reconstruction error with energy consumption as a trade-off.

In this work, considering spatial relevance and temporal smoothness of sensory data, we exploit Treelet transform [13]

978-1-4799-2003-7/14/$31.00 ©2014 IEEE

to derive proper sparsification bases (transformations). Compared with traditional transformations (e.g., DCT, Wavelets) capable for smooth data processing, Treelet transform is data-driven which extends localized multi-scale analysis to unordered data. Compared with PCA which is also data-driven, Treelet transform focuses more on localized correlation structures while PCA aims at constructing a global optimal linear representation of noisy observations. Treelet transform achieves better performance when the number of variables p is much larger than that of observations n, since the principle components discovered by PCA may be masked by noise in that regime. This merit of Treelet transform is significant, since generating training set with precise observations is very expensive in terms of energy consumptions in WSNs. Besides, Treelet transform has a merit of flexibility in introducing external information (geographic information in this work) for specific applications due to its customizable similarity metric. Aiming at the design of a practical energy-efficient CSDAS in multi-hop WSNs, for the first time, we propose to exploit the localized correlation structure obtained by Treelet transform for realizing proper clustering strategy for nodes, which facilitates energy saving of CSDAS by reducing the number of nodes in one cluster without impacting the localized correlation among nodes. Simulation results with real world signals show it is effective to introduce Treelets into CS. Compared with PCA, the reconstruction error rate with Treelets is about 18% lower when the number of measurements accounts for 30% of the dimension of original signals and even larger performance gain is obtained at higher energy consumption levels. Meanwhile, simulations results further show that our novel correlation based clustering strategy is of great potential. More specifically, there is a gain of roughly 35% for total energy savings with our proposed clustering strategy.

The rest of this paper is organized as follows. In Section II, a brief introduction of Treelets and CS is presented. Our CSDAS is detailed in Section III. Simulation results and discussions are provided in Section IV. We conclude this work and point out future work in Section V.

II. Preliminary

To facilitate readers' understanding of this work, brief introductions to the basic ideas of Treelets and CS are presented at first.

A. Treelets [13]

Treelet transform is a novel algorithm aiming at the construction of multi-scale bases, which extends multi-scale analysis to unordered data. Treelet transform is constructed on a hierarchical cluster tree. At each level of the tree, the most correlated variables measured by a customizable similarity metric are transformed and replaced by a coarse-grained Sum Variable and a residual Difference Variable. It returns a hierarchical tree and an orthogonal basis which both reflect the internal localized correlation structure of the data, and are stable to noise.

The treelet algorithm is described in Algorithm 1.

Algorithm 1 Treelet transform

Require: The original sample data set, $\mathbf{x}$; The maximum level, L;

Ensure: The transformation matrix, $B^{(L)}$;

1: At level $l = 0$, compute the sample covariance and correlation matrix $\hat{\Sigma}^{(0)}$ and $\hat{M}^{(0)}$ from original data $\mathbf{x}$; Initialize $\mathbf{x}^{(0)} = \mathbf{x}$, transformation matrix $B^{(0)}$ ($p \times p$ identity matrix);

2: **Repeat** 3-5, **for** $l = 1, \ldots, L$:

3: Find the two most similar variables according to the similarity matrix $\hat{M}^{(l-1)}$. Assume α and β are the indexes of two most similar variables;

4: Perform a local Jacobi rotation on this pair. Find a rotation matrix $J(\alpha, \beta, \theta_l)$, $\theta_l \leq \pi/4$ and $\hat{\Sigma}^{(l)}_{\alpha\beta} = \hat{\Sigma}^{(l)}_{\beta\alpha} = 0$.

5: Update the following matrices $\hat{\Sigma}^{(l)} = J^{\mathrm{T}} \hat{\Sigma}^{(l-1)} J$, $B^{(l)} = B^{(l-1)} J$, $\mathbf{x}^{(l)} = J^{\mathrm{T}} \mathbf{x}^{(l-1)}$, and update $\hat{M}^{(l)}$ accordingly;

6: **return** $B^{(L)}$;

In above table, $\mathbf{x} = B^{(l)} \mathbf{x}^{(l)}$ at level l. $\mathbf{x}^{(l)}$ owns a beneficial property that there are only a small number of non-zero/significant elements when other elements are zero/close to zero. When adopted in WSNs, Treelet transform can derive sparsification bases from training sets, and these sparsification bases can be exploited to predict correlation structures of sensory data following the training set for a period of time.

B. Compressive Sensing [2][3]

CS finds a sparse solution of an ill-posed inverse problem when the signal of interest is known to be sparse and compressible. Denote $\mathbf{x} = B\mathbf{s}$, where $\mathbf{x} \in \mathbb{R}^N$ is the signal of interest, and $B \in \mathbb{R}^{N \times N}$ is a transformation space. Signal $\mathbf{x}$ is K-sparse if $\mathbf{s}$ has only K significant elements when other elements are zero or close to zero. In CS, "compressed" measurements instead of periodic signal samples are directly acquired. Each measurement y_i is the inner product of $\mathbf{x}$ and the measurement vector $\phi_i \in \mathbb{R}^N$, i.e., $y_i = \langle \phi_i, \mathbf{x} \rangle$. Define $\Phi = [\phi_1, ..., \phi_M]^{\mathrm{T}}$, we have

$$\mathbf{y} = \Phi \mathbf{x} = \Phi B \mathbf{s}, \tag{1}$$

where $\Phi \in \mathbb{R}^{M \times N}$, and $M \ll N$.

In general, Equation (1) has infinite solutions. However, if ΦB satisfies the so-called Restricted Isometry Property (RIP) [3], the sparsest solution will be obtained by CS. Intriguingly, a large class of random matrices whose entries are *i.i.d.* Gaussian, Bernoulli (± 1), etc, have the RIP with high probability (see more details in [3]). The sparsest solution is obtained by solving

$$\min_{\mathbf{s}} \|\mathbf{s}\|_0 \quad \text{s.t.} \quad \Phi B \mathbf{s} = \mathbf{y}. \tag{2}$$

Prior researches in CS mainly focus on reducing the number of measurements M, increasing robustness, and reducing computational complexity of recovery algorithms. So far, robust recovery of K-sparse signals with $M = \mathcal{O}(K \log(N/K))$ noisy measurements can be realized by convex optimization solvers or greedy algorithms. In our scheme, we imbibe the formers' research production about CS theory itself and focus on making full use of it for data aggregation schemes in multi-hop networks.

III. Our Novel CS-based Data Aggregation Scheme

In this section, details of our CSDAS for unordered data in WSNs are provided. At first, a brief description of the system model is presented.

Consider that a WSN consists of only one sink node and a large amount of randomly sowed sensor nodes. Readings of all sensor nodes (denoted as $\mathbf{x}(k) = [x_1(k), x_2(k), \ldots, x_n(k)]$ at time k as follows, where $x_i(k)$ is the reading of node i) need to be gathered by the sink node periodically. There are spatio-temporal correlations among readings from all sensor nodes. We aim at designing a CS-based technology to collect information energy-efficiently from the monitoring area by making full use of spatio-temporal correlations.

A. Treelets as Sparse Transformation Tool to Derive Sparsification Bases

For a typical CSDAS, an immediate task is to accomplish the collection of measurements $\mathbf{y}(k)$ of sensory data $\mathbf{x}(k)$. In this work, the basic idea of compressing records from all sensor nodes through dense random linear projection (DRP) [6][8] is adopted. For sensor node i, its own observation $x_i(k)$ is multiplied with a local random coefficient vector ϕ_i', and then the sum $\sum_{j=1}^{i-1} \phi_j' x_j(k) + \phi_i' x_i(k)$ is sent to its next hop node, where $\sum_{j=1}^{i-1} \phi_j' x_j(k)$ is transmitted to node i by its previous node.[1] The measure procedure is formulated as

$$\mathbf{y}(k) = \Phi \mathbf{x}(k). \tag{3}$$

The columns of Φ contain random coefficient vectors ϕ', and *i.i.d* Gaussian matrix is adopted as measurement matrix Φ in our simulations due to its proven performance [2].

Another task we confront is to discover the correlation structure (sparsification basis) of unordered sensory data. In this work, we propose to adopt Treelet transform to solve this problem. Different from conventional methods, Treelet transform discovers the localized correlation structure from training set $T(k)$ at time k, which consists of K sequential observations, i.e., $T(k) = \{\mathbf{x}(k-K), \mathbf{x}(k-K+1), \cdots, \mathbf{x}(k-1)\}$. By applying Treelet transform, an orthogonal basis $B(k)$ is obtained, then it is adopted to sparsify sensory data after the training set for a period of time.[2] However, the performance of the basis is impacted by the similarity metric. Thus in this work, not only the common correlation coefficient but also geographic information is introduced into the similarity metric to avoid misjudgement caused by noise and improve the reliability of the transformation. The similarity is defined as

$$M_{ij} = |C_{ij}| - \lambda |d_{ij}|, \tag{4}$$

where C_{ij} is the common correlation coefficient, d_{ij} is the Euclidian distance between node i and j normalized by the maximum Euclidian distance between all pairs of nodes and λ is a non-negative constant. By this similarity metric, nodes which are far apart showing a strong correlation in terms of readings are eliminated from the class of similar node pairs. That is because this kind of correlation may be caused by noise or some accident situation and it can not be used to predict the correlation structure of sensory data in the following period of time.

Once the sparsification basis $B(k)$ is obtained, $\mathbf{x}(k)$ can be sparsified as

$$\mathbf{x}(k) - \bar{T}(k) = B(k)\mathbf{s}(k), \tag{5}$$

where $\bar{T}(k) = \frac{1}{K}\sum_{n=1}^{K} \mathbf{x}(k-n)$, and $\mathbf{s}(k)$ has only a few significant elements. Due to the orthogonality of $B(k)$, (5) can also be written as

$$\mathbf{s}(k) = B^{\mathrm{T}}(k)\left(\mathbf{x}(k) - \bar{T}(k)\right). \tag{6}$$

With (6), Equation (3) can be reorganized as

$$\mathbf{y}(k) - \Phi\bar{T}(k) = \Phi B(k)\mathbf{s}(k). \tag{7}$$

The aforementioned procedure accomplish the collection of measurements, then the signal of interest can be recovered by solving the following problem at the sink side

$$\min_{\mathbf{s}(k)} \|\mathbf{s}(k)\|_0 \ \text{s.t.}\ \mathbf{y}(k) - \Phi\bar{T}(k) = \Phi B(k)\mathbf{s}(k). \tag{8}$$

Various algorithms are available for solving problem (8). Generally speaking, these algorithms can be classified into two categories: convex optimization based algorithms and greedy algorithms. Convex optimization based algorithms have better performance at the cost of more computational complexity. In this work, considering the tradeoff between performance and complexity, a typical greedy algorithm named Compressive Sampling Matching Pursuit (CoSaMP) [14] is adopted.

B. A Localized Correlation based Clustering Strategy

Since the number of measurements M grows large as $M \sim \mathcal{O}(K \log(N/K))$ [3],[3] DRP based techniques perform poor in terms of total energy consumption [9], though the energy consumption of each node is balanced. Aiming at ameliorating this problem, we propose a novel clustering strategy, which is realized from the localized correlation structure returned by Treelet transform.

As mentioned above, Treelet transform is constructed on a hierarchical cluster tree, thus nodes are clustered into several clusters naturally. Besides, nodes in the same cluster are selected according to their similarity. That is, these nodes have strong correlation which means their readings can be well

[1] ϕ' is the column vector of Φ and M determines the length of a wireless frame that each node transmits.

[2] In the clustering strategy which will be introduced in next subsection, Treelet transform returns a group of orthogonal bases, since each cluster needs a respective one.

[3] Larger N means larger energy consumption of each node.

sparsified. Obviously, at different levels of the hierarchical cluster tree, the clustering strategy is different. Hence, we can customize the clustering strategy by stopping clustering nodes at a certain level to reserve a corresponding number of clusters.[4] Under our clustering model, a typical CS-based data aggregation is performed in each cluster separately, and the sink recovers the original signal from these multi-cluster measurements individually. Since the localized correlation based clustering strategy not only decreases the number of nodes in one cluster but also reserves strong correlations among nodes in the same cluster, the required number of measurements M of each node is reduced. In other words, the communication load of each node is reduced. The localized correlation based clustering can be regarded as an enhanced technique if there is strong localized correlation among sensory data.

C. The Work Process of Our Proposed Scheme

The work process of our CS-based scheme is present in this subsection. The most significant difference between our scheme and the conventional ones is that our scheme need training rounds to collect training set. Hence, the main problem in the work process is how to arrange two alternate procedures, i.e., training and monitoring rounds. Considering the valid time of derived information from training sets, a plain idea is to collect training sets in training rounds, and adopt the information discovered from training set for the next μ monitoring rounds before collecting the next training set. Two enhancements are made in the work. Firstly, during monitoring rounds, $\bar{T}$ is updated/re-calculated. For instance, if the collection of the last training set is completed at time k, then $\bar{T}(k+1) = \frac{1}{K}\{\sum_{n=1}^{K-1} \mathbf{x}(k-n) + \hat{\mathbf{x}}(k)\}$, where $\hat{\mathbf{x}}(k+K-1)$ is the recovered readings at time k. Secondly, a feedback control is designed to adjust μ adaptively according to an original error estimation method. μ is related to the degree of temporal correlation of signals and impacts the performance of the system significantly. The original error estimation method is formulated as

$$\zeta = \frac{\mathrm{norm}\left(\hat{\mathbf{x}}(k-K) - \mathbf{x}(k-K)\right)}{\mathrm{norm}\left(\mathbf{x}(k-K)\right)}, \tag{9}$$

where the operator $\mathrm{norm}(\cdot)$ is 2-norm, $\mathbf{x}(k-K)$ is the first precise observation of newly collected training set $T(k)$. Usually, we only have the knowledge of $\hat{\mathbf{x}}$ while $\mathbf{x}$ is not available. Fortunately, we note that training sets contain some precise readings. For the first observation of training round $T(k)$, we adopt the material, e.g., sparsification basis, measurement matrix and $\bar{T}(k-K-1)$, used for the last monitoring round (at time $k-K-1$) and obtain $\hat{\mathbf{x}}(k-K)$ from $\mathbf{x}(k-K)$. Then error estimation method (9) can be carried out. Actually, though only the reconstruction error of the last monitoring round before each training round is estimated, it is sufficient to give a judge standard that whether μ is reasonably set. The benefit of adaptively adjusting μ according to ζ is not to achieve performance gain but to ensure that the performance for continuous working is within a reasonable range to avoid unnecessary energy wasting and unacceptable reconstruction distortion. In addition, if strong localized correlation among sensory data is found from T, the localized correlation based clustering strategy is performed and each cluster derives its parameters/information during monitoring rounds and collects measurements individually. Then the sink recovers observations for each cluster. In summary with a flowchart, the work process of our scheme is shown in Fig. 1.

[4]Other methods are also investigated, e.g., selecting a level at which the Difference Variable, which reflects the similarity between two subtrees, is larger than a predefined value.

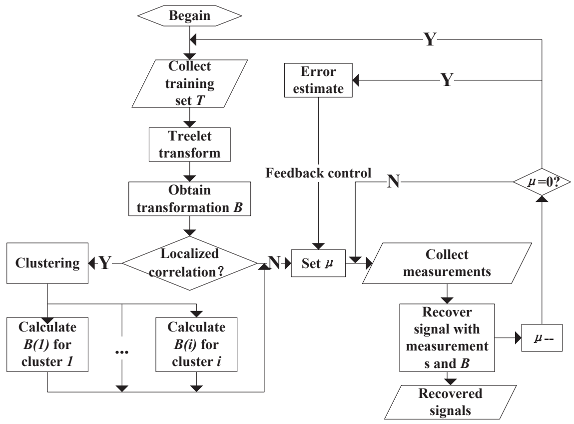

Fig. 1: The work process of our scheme.

IV. Numerical Results

In this section, we compare the performance of our CS-based scheme, labeled as Treelet-CS, with PCA-CS [12] through extensive simulations, since PCA-CS has the most similar structure with ours and it has been proven to be an energy efficient data aggregation scheme [12][16]. Besides, the efficiency of our localized correlation based clustering strategy is presented.

1) Simulation environment configurations: The WSN testbed at the *Ecole Polytechnique Fédérale de Lausanne* [15] which consists of 97 sensors deployed across the EPFL campus is adopted for our simulations. Different observations of environmental conditions at EPFL (e.g., temperature, humidity, voltage, luminosity) are collected every five minutes. Thanks to the work of [16], observations of invalid nodes are eliminated, and thus we are able to exploit the database of actual temperature for our simulations conveniently. In addition, readings of all nodes at the same time k form a signal of interest $\mathbf{x}(k)$. To validate the effectiveness of our proposed scheme, we conduct simulations with the same configurations with the aforementioned WSN testbed.

2) Parameter definitions: Similar with previous sections, $\mathbf{x}(k)$ is signal of interest at time k, N is the number of nodes, M is the number of measurements, K is the number of observations in a training set, μ is the number of monitoring rounds, and $\xi = K/\mu$ is the ratio of observations used for training and monitoring.

3) Metrics of energy consumption and reconstruction error: For our localized correlation based clustering strategy, M of each cluster is different. Hence we define

$$\eta = \frac{\Sigma_i M_i N_i}{N^2} \tag{10}$$

as normalized energy consumption metric, where i is the index of each cluster. $\eta = \frac{M}{N}$ when the clustering strategy is not adopted. The reconstruction error is measured as

$$\varepsilon = \sum_{k=1}^{F} \frac{\mathrm{norm}\left(\mathbf{x}\left(k\right) - \hat{\mathbf{x}}\left(k\right)\right)}{\mathrm{norm}\left(\mathbf{x}\left(k\right)\right)} \Bigg/ F, \tag{11}$$

where F is the number of total frames adopted in simulations.

4) Simulation results: The impact of K to reconstruction effect is presented in Fig. 2. The performance curves with different K are provided. For each curve, the reconstruction error declines as normalized energy consumption increases. Besides, it is roughly seen that the shorter training set length K is, the better performance is. This result seams to go against conventional thought that longer training set will lead to more accurate prediction. However, in this set of simulations, ξ is set to be constant to guarantee that simulations for different K are performed with the same proportion of two different collections.[5] At the same ξ, larger K means larger μ. Importantly, the validity of training sets decreases over μ due to the variation of spatial correlation. Thus, we doubt that the reconstruction error are mainly caused by the variation of spatial correlation over time not by inaccuracy of shorter K. What calls for special attention is that the performance of $K = 2$ is similar with that of $K = 3$. That is because when $K = 2$, the covariance matrix of training set has only two possible values: "1" and "-1". The two-value covariance is less stable to noise, and cannot distinguish the strength of correlations. Thus it has partly performance loss.

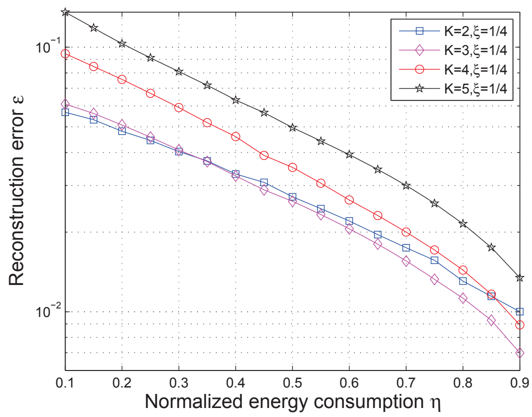

Fig. 2: Impact of different length of training set.

The impact of μ to reconstruction effect is presented in Fig. 3. Spots (in blue) are the average result of 20 times simulation for the same piece of real data set, and to make the impact clearer to readers, we plot a variation trend with the red curve. The figure shows a obvious conclusion that the bigger μ the bigger reconstruction error, i.e., the reconstruction error are mainly caused by the variation of spatial correlation over time. Hence, μ impact the performance of whole scheme directly, which means dynamically adjust μ according to error estimation will keep the reconstruction error in a expected range.

[5]The observation collection methods during training and monitoring rounds are different.

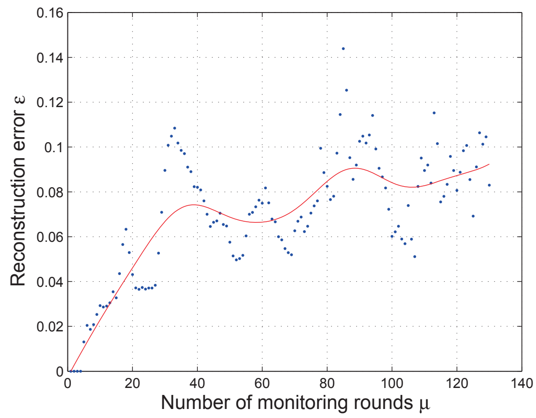

Fig. 3: Performance of adjusting μ.

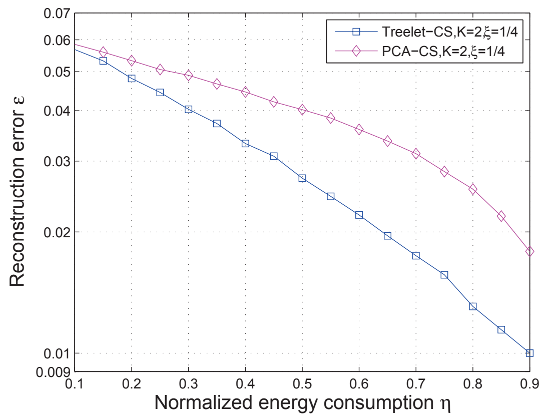

Fig. 4: Performance comparisons between PCA-CS and Treelet-CS.

Performance comparisons between PCA-CS and Treelet-CS are presented in Fig. 4. The performance of classical methods are not included due to the proven performance of PCA-CS compared to classical ones [12]. It is seen that for real world signals, compared to PCA-CS the performance gain with Treelets is about 18% when the number of measurements accounts for 30% (normalized energy consumption metric, η) of the dimension of original signals and even more performance gain could be obtained when η is higher. This result proves the effectiveness of exploiting Treelet transform and agrees with our analysis aforementioned that when the number of observations is fewer than that of variables (nodes in this work), the principle components discovered by PCA are masked by noise.

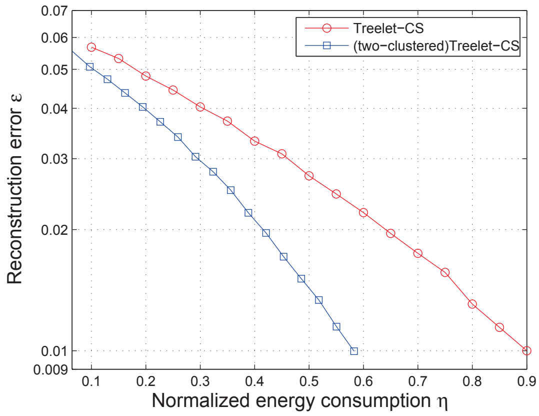

Fig. 5: Performance comparison between clustering and non-clustering for Treelet-CS.

Performance comparison between clustering strategy and plain method (non-clustering) for Treelet-CS is presented in Fig. 5. For this set of simulations, K is fixed to two. Hence, as aforementioned, correlations between nodes are only distinguished as positive or negative (“1” or “-1”). Hence, nodes are clustered into two clusters at most. Certainly, for other value of K, the strength of correlations can be distinguished and can result in more sophisticated and complicated clustering strategy. For the sake of simplicity, we set $K = 2$. Result in Fig. 5 shows that the localized correlation based clustering strategy is of great potential, as a gain of roughly 35% for total energy savings is seen under different construction error levels.

V. Conclusion

In this work, we have investigated the feasibility of the applications of both Treelet transform and novel clustering strategy for CSDAS in WSNs. Specially, Treelet transform is adopted as sparse transformation tool to derive transformations from training data. In contrast with prior global methods (PCA in this work), Treelet transform is a well-suited localized method in solving data aggregation problem since it has better performance when only a few precise training observations are available. Moveover, we have also proposed a novel clustering strategy based on the localized correlation structure obtained by Treelet transform. Simulation results with real world signals show that it is efficient to introduce Treelets into CSDAS, as compared with PCA. Meanwhile, simulations results further show that our novel correlation based clustering strategy is of great potential. In the future, we will investigate CSDAS, taking the importance of different observations into account, since abnormal events usually contain more information. Detailed routing protocol for our novel CSDAS will be investigated. Moreover, we will extend Treelets-CS into other pilot-based coding and decoding technologies.

Acknowledgment

This research is partially supported by the National Science and Technology Major Project of China under grant 2014ZX03005001, the NSFC under grant 61231009, International Science & Technology Cooperation Project of Qinghai Province(2013-H-811), and the STCSM under grant 11JC1412300.

References

[1] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi,“In-network aggregation techniques for wireless sensor networks: a survey,” IEEE Wireless Communications Magazine, vol. 14, no. 2, pp. 1526-1284, April. 2007

[2] E. Candès and M. Wakin, “An introduction to compressive sampling,” IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 21-30, Mar. 2008.

[3] D. Donoho, “Compressed sensing,” IEEE Transactions on Information Theory, vol.52, no. 4, pp. 1289-1306, 2006.

[4] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, “Compressive wireless sensing,” in Proceeding of 2006 International Conference on Information Processing in Sensor Networks.

[5] D. Baron, M. Wakin, M. Duarte, S. Sarvotham, and R. Baraniuk, “Distributed compressed sensing,” Technical Report ECE-0612, Electrical and Computer Engineering Department, Rice University, Dec. 2006.

[6] M. Rabbat, J. Haupt, A. Singh, and R. Nowak, “Decentralized compression and predistribution via randomized gossiping,” in Proceeding of 2006 International Symposium on Information Processing in Sensor Networks, IPSN, pp. 51-59, 2006.

[7] G. Shen, S. Y. Lee, S. Lee, and et al, “Novel distributed wavelet transforms and routing algorithms for efficient data gathering in sensor webs,” in proceeding of NASA Earth Science Technology Conference (ESTC2008), University of Maryland, MD, USA, Jun. 2008.

[8] C. Luo, F. Wu, J. Sun, and C. Chen, “Compressive data gathering for large-scale wireless sensor networks,” in Proceeding of the 15th annual international conference on Mobile computing and networking. ACM, pp. 145-156, 2009.

[9] L. Xiang, J. Luo, A. Vasilakos, “Compressed data aggregation for energy efficient wireless sensor networks,” in Proceeding of international conference on Sensing Communication and Networking, Salt Lake City, pp. 27-30, June 2011.

[10] W. Wang, M. Garofalakis, and K. Ramchandran, “Distributed sparse random projections for refinable approximation,” in Proceeding of 2007 International Symposium on Information Processing in Sensor Networks, IPSN, pp. 331-339, 2007.

[11] G. Quer, R. Masiero, D. Munaretto, M. Rossi, J. Widmer, and M. Zorzi, “On the Interplay Between Routing and Signal Representation for Compressive Sensing in Wireless Sensor Networks,” in ITA Workshop, San Diego, CA, US, Feb. 2009.

[12] R. Masiero, G. Quer, D. Munaretto, and et al, “Data Acquisition through joint Compressive Sensing and Principal Component Analysis,” in Proceeding of IEEE Globecom 2009, Honolulu, Hawaii, USA, Nov.-Dec. 2009.

[13] A. B. Lee, B. Nadler, and L. Wasserman, “Treelets–an adaptive multi-scale basis for sparse unordered data,” Annals of Applied Statistics, vol. 2, no. 2, 2008.

[14] D. Needell and J. Tropp, “CoSaMP: iterative signal recovery from incomplete and inaccurate samples,” Applied and Computational Harmonic Analysis, vol. 26, no. 3, pp. 301-321, 2009.

[15] “EPFL LUCE SensorScope WSN,” (Access Date: June 2012) [Online]. Available: http://sensorscope.epfl.ch/.

[16] R. Masiero, G. Quer, G. Pillonetto, M. Rossi, and M. Zorzi, “Sensing, Compression and Recovery for Wireless Sensor Networks: Sparse Signal Modelling and Monitoring Framework Design,” IEEE Transactions on Wireless Communications, vol.11, no. 10, pp. 3447-3461, 2012.

[17] X. Chu, D. Lopez-Perez, Y. Yang, and F. Gunnarsson, “Heterogeneous Cellular Networks: Theory, Simulation and Deployment,” Cambridge University Press, 2013.