Proceedings of the 32nd Chinese Control Conference
July 26-28, 2013, Xi'an, China

Redundancy Control in Large Scale Sensor Networks via Compressive Sensing

Liwen Xu[1], Yongcai Wang[1], Changjian Hu[2]
[1]Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, P. R. China
[2]NEC Labs, Beijing, P. R. China

Abstract: In wireless sensor networks for smart city or smart planet applications, massive volumes of real-time sensory data are being generated in every second, which pose great challenges to the power-limited sensor nodes, bandwidth-limited transmission links, and require high data storage and management costs. To deal with these challenges, compressive sensing (CS) converts the the spatially and temporally correlated information to sparse signals in some transformed domains (Such as DCT and FFT), and conducts cost-efficient, low-rank sensing. This paper presents a cost-centric comparison between recent compressive sensing solutions, i.e., Compressive Data Gathering (CDG) and Compressive Sparse Function (CSF), with traditional sensing technologies, in the means of sensing, transmission, storage and computation costs. It shows by a city temperature collection example that CDG performs similarly to CSF, both of which can prolong the network lifetime for almost one magnitude than traditional multi-hop sensing, while providing enough information for recovering the temperature distributions.

Key Words: Compressive Sensing, Sensor Networks, Energy Efficiency, Data Gathering, Redundancy Control

1 Introduction & background

With the growing demands of “smart planet” and “smart city” applications, large-scale wireless sensor networks (LWSN) are fast spreading. Such sensor networks contain thousands or ten thousands of wireless sensor nodes, which collect massive, real-time information for environment, traffic or resident information monitoring applications.

The sensor nodes are generally limited in power and communication range. They take samples following application-defined duty-cycle rules for the purpose of energy saving. Because the huge number of the sensor nodes, they jointly produce large amount of real-time, sensory data. These big volume of real-time data generally needs to be forwarded in multiple hop manner towards the sink because of the limited communication range of each node.

A most recent example of LWSN for city environment monitoring is CitySee project [7], which is developed in Wuxi City, China. The authors deployed 100 *sensor nodes* and 1096 *relay nodes* to monitor the urban CO2 in a 5000(m)*4000(m) city region. Fig.1 is a snapshot of a part of the CitySee sensor network. Since each sensor in the network has only 100(m) communication range, the longest path in the CitySee network is up to 20 hops. Therefore, each byte of generated data by the *sensor nodes* will be forwarded up to 20 times before reaching the sink.

1.1 Critical Problems in LWSN

Due to the requirement of multi-hop forwarding to transmit data through long distances, in LWSN, it is critically important to allow only small number of *sensor nodes* to generate only small volume of sensing data. This is important for not only reducing the energy consumption of the *sensor nodes*, but also can remarkably reduce the transmission costs of the *relay nodes*. In CitySee, for the purpose of energy saving, only 100 *sensor nodes* can generate data in the 20 square kilometers region with duty cycle 0.04. The other 1096 sensors only relay data without generating observations.

However, for monitoring real-time information of large areas, an easily raised question for LWSNs is that if only small volume of information is collected for the purpose of energy efficiency, how can the LWSNs guarantees comprehensive and accurate information monitoring? This question is actually concerning the tradeoff between:

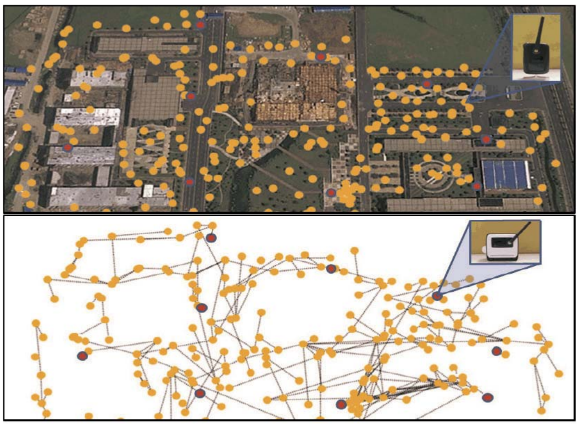


Fig. 1: A part of large scale sensor network in Citysee[7]

1) monitoring comprehensiveness and accuracy and,
2) sensing and data transmission costs

in LWSNs.

To address these challenging problems, an important factor that can be exploited is that the physical information, such as city environment information are highly correlated in temporal and spatial domains, which are highly redundant if they are collected densely or continuously. Such kinds of information can be transformed in to sparse signals by Fast Fourier Transformation (FFT) or Discrete Cousin Transformation (DCT) etc.

By exploiting such information sparsity features in redundant data, the revolutionary compressive sensing (CS) theory was proposed in the literature to enable energy efficient information sensing and transmission in LWSNs. These CS-based information sensing schemes not only guarantees information monitoring comprehensiveness and accuracy, but also preserves energy efficiency.

1.2 Contribution of This Paper

This paper presents a cost-centric comparison between traditional sensing technology and two most recent compressive sensing technologies in LWSN, i.e., Compressive Data Gathering (CDG)[4] and compressive sparse function (CSF)[6]. It compares the sensing, transmission, storage and computation costs and interprets how these CS-based methods can achieve energy efficient data collection while preserving information monitoring accuracy.

It firstly introduces the basic idea of compressive sensing to interpret how CS can exploit the information redundancy to reduce the sensing cost. Then the CDG and CSF schemes are overviewed, with detailed explanation and comparison of how they reduce the sensing and data communication costs while preserving the information monitoring accuracy.

At last, a simulated city temperature monitoring example is used to to numerically show the efficiency of CDG and CSF. It shows that CSF can prolong the network lifetime for almost one magnitude than that of traditional sensing, while providing enough information for recovering the temperature distributions.

2 Sensing Cost Reduction in CS Theory

In this section, we firstly introduce Compressive Sensing (CS) theory[1, 2]. CS theories consider the problem of how much information are required to be collected for comprehensively and accurately monitoring information from an area-of-interest. They focus solely on the fundamental sensing cost reduction problem while guaranteeing information monitoring accuracy, without considering the detailed implementation in the sensor networks and the transmission cost issues. These issues in sensor networks were addressed by recent advantages of Compressive Data Gathering and Compressive Sparse Function schemes, which will be introduced in Section 3.

Compressive Sensing is a revolutionary technique to address sensing cost reduction problem by exploiting the information sparsity among the redundant data. Generally speaking, CS can successfully recover high dimensional data from its projections in a lower dimensional space, when the monitored data has a small sparsity in some transformed domains.

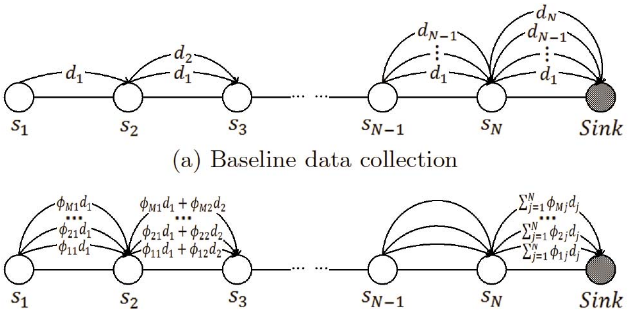
The physical information, such as city environment information are highly correlated in temporal and spatial domains. Let's suppose the information to be measured in the monitored area forms a length-n vector, which is denoted by $\mathbf{x} \in \mathbb{R}^n$. The vector $\mathbf{x}$ can be transformed to a sparse vector $\mathbf{s} = \Psi\mathbf{x}$, where $\Psi \in \mathbb{R}^{n*n}$ is a transformation basis such as Fast Fourier Transformation (FFT) or Discrete Cousin Transformation (DCT). After transformation, the vector $\mathbf{s} \in \mathbb{R}^n$ becomes k-sparse which contains at most $k \ll n$ non-zero values and all the other items are zeros.

By transforming $\mathbf{x}$ to a sparse vector $\mathbf{s}$, the vector $\mathbf{s}$ can be sampled by a low rank matrix, i.e,:

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

where $\Phi \in \mathbb{R}^{m*N}$ is a low-rank sampling matrix, and $\mathbf{y} = \mathbb{R}^m$ is a low-rank observation vector, in which, $m < N$.

Because $m < n$ makes Φ underdetermined, Eqn.(1) is hard to resolve. However since $\mathbf{s}$ is a k-sparse vector, we can solve this system by finding the sparsest version of $\mathbf{s}$ as an approximation, which is to solve a l_0-norm minimization



d_1
d_2
d_1
d_{N-1}
d_1
d_N
d_{N-1}
d_1
s_1
s_2
s_3
s_{N-1}
s_N
$Sink$
$\phi_{M1}d_1$
$\phi_{21}d_1$
$\phi_{11}d_1$
$\phi_{M1}d_1 + \phi_{M2}d_2$
$\phi_{21}d_1 + \phi_{22}d_2$
$\phi_{11}d_1 + \phi_{12}d_2$
$\sum_{j=1}^{N}\phi_{Mj}d_j$
$\sum_{j=1}^{N}\phi_{2j}d_j$
$\sum_{j=1}^{N}\phi_{1j}d_j$

(a) Baseline data collection

(b) Compressive data gathering

Fig. 2: Compressive Data Gathering in chain topology

$$\underset{\hat{\mathbf{s}} \in \mathbb{R}^n}{argmin} \|\hat{\mathbf{s}}\|_{l_0}, \text{ such that } \Phi\hat{\mathbf{s}} = \mathbf{y}, \tag{2}$$

where $\|\cdot\|_{l_0}$ counts the non-zero elements in a vector. This problem is proved to be NP-hard, however CS amazingly uses l_1-norm minimization

$$\underset{\hat{\mathbf{s}} \in \mathbb{R}^n}{argmin} \|\hat{\mathbf{s}}\|_{l_1}, \text{ such that } \Phi\hat{\mathbf{s}} = \mathbf{y}, \tag{3}$$

to take the place of l_0-norm minimization and the solution to Eq(3) exactly solves Eq(2) if Φ satisfies Restricted Isometry Property (RIP) and

$$m \geq ck \log n, \tag{4}$$

where c is a positive constant. Practically, it is sufficient that $1 \leq c \leq 4$.

The isometry constant of a $m \times n$ matrix Φ is defined as δ which holds

$$1 - \delta \leq \frac{\|\mathbf{\Phi x}\|_{l_2}}{\|\mathbf{x}\|_{l_2}} \leq 1 + \delta \tag{5}$$

for all the k-sparse vectors $\mathbf{x}$. We can simply consider that Φ holds for RIP if δ is small enough, e.g. $\delta < 1$. RIP describes the property that Φ behaves isometrically which means a vector can nearly keep its l_2-norm under the transformation of Φ.

By the low-rank sampling technique in CS, we only take $m < n$ samples, which needs much less sensing cost than traditional sensing methods.

However, such classical CS theory has only considered the efficient sensing problem, which has not considered the multi-hop data transmission costs in LWSNs.

3 Sensing and Transmission Cost Reduction in CDG and CSF

Because the multi-hop data communication consumes more energy than the sensing operations, how to reduce the data communication cost in LWSN while preserving the information monitoring accuracy is a critical challenging problem. On this issue, several CS-based strategies have been proposed to devote their effort. Among these methods, Compressive Data Gathering (CDG) and Compressed Sparse Functions (CSF) are two representatives of state-of-the-art efficient data collection strategies in sensor networks.

CDG uses random coding at each intermediate node in a distributed fashion and decodes at data center. CSF works in an even simpler way that it randomly collects data from a part of nodes and reconstructs the complete version at data center.

They both reduce the sensing cost and the transmission cost in LWSNs while preserving the information monitoring accuracy. We are going to give a more detailed introduction about these two methods.

3.1 CDG: Compressive Data Gathering

The key point of the high efficiency of CS is the use of a smaller matrix to project the original data vector into a lower dimensional space. CDG brings this process inside the network in a distributed fashion via *in-network encoding*. Instead of passing its original data to the next node, each node firstly encode its own data and the data obtained from the previous node and then pass the encoded version to the next hop.

More specifically, Fig.(2) illustrates a simple example which compares the traditional data gathering scheme and the CDG scheme in a multi-hop chain-style sensor network.

- In traditional data gathering, s_1 sends message d_1 to s_2; s_2 sends d_1 and d_2 to s_3. The other sensors repeat the multi-hop data relay scheme until all the messages are reported to the sink. The overall number of messages transmitted by sensors is $(n-1)n/2$, where n is the number of sensors. Therefore the transmission cost is in the magnitude of $O(n^2)$.
- In CDG, s_1 passes $\phi_{11}d_1$ to s_2 where ϕ_{11} is a random number, then s_2 passes $\phi_{11}d_1 + \phi_{12}d_2$ to s_3 and so on until the last encoded value is transmitted to the sink node. Clearly, the sink node gets a linear combination of all the original data $y_1 = \sum_{i=1}^{n} \phi_{1i}d_i$. The similar collection process is carried out for m times, thus the sink node have a series of encoded, i.e. linear combinations of, data which can be written as $\mathbf{y} = \mathbf{\Phi d}$. Each ϕ_{ij} is a random number generated by a global seed thus the sink node can reconstruct the whole random matrix $\mathbf{\Phi}$. The overall number of messages transmitted by sensors is mn, i.e., the overall transmission cost is in the magnitude of $O(knlog(n))$, because m is in the magnitude of $O(klog(n))$. Therefore, CDG much reduces the transmission cost.

In case the vector $\mathbf{d}$ is not sparse, CDG finds a sparse representation of $\mathbf{d}$ by $\mathbf{d} = \Psi\mathbf{x}$. Since the sink node can know Ψ and reconstruct the whole random matrix $\mathbf{\Phi}$ by the global random seed, the sink node can recover the vector $\mathbf{x}$ by l_1 norm minimization algorithms to solve the problem in (3). So that, CDG can successfully compress the n-dimensional data into a lower m-dimensional space through in-network-encoding-based transmission. It can reduce the transmission cost efficiently.

3.2 CSF: Compressed Sparse Functions

In the following part, we are going to briefly introduce the basic of Compressive Sensing technique and CSF strategy and see it makes sensor networks more energy-efficient.

Compressed Sparse Functions view the data in the area-of-interest as values generated by a function F, which is called *satisfying function* of the monitoring area. For information monitoring in 2-D space by sensor network, we assume $\mathbf{d} = (d_1, d_2, \cdots, d_n)$ are n sensing points, where a sensor is deployed at each sensing point to capture data and we are trying to collect all the data from these sensors. The value sensed by sensor i is denoted by $d_i = F(i)$, where $i = 1, 2, \cdots, n$ denote n sensors at the n sensing points.

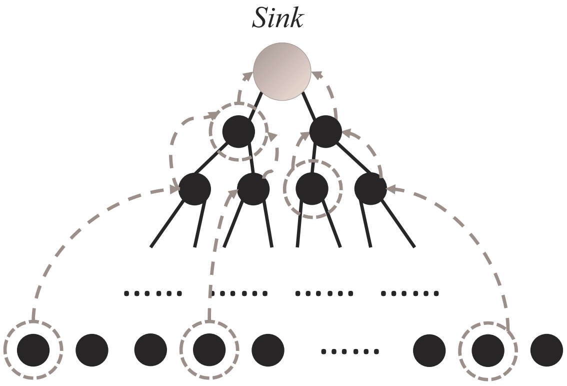


Fig. 3: Compressed Sparse Functions in Binary-Tree-Structured network

Different from transitional CS and CDG, which collect data by low rank linear encoding, CSF tries to identify the satisfying function and to calculate the data at the sensing points by the estimated satisfying function. The basic idea is that the satisfying function can generally be determined by only a small set of function parameters. These small set of parameters can be determined by using even less measurements than CDG.

In order to identify this satisfying function efficiently, CSF defines a series of function called *function base*, under which the satisfying function can be sparsely represented. Letting $\mathbf{P} = \{p_1, p_2, \cdots, p_n\}$ denote the function base, F can be represented as $a_1p_1 + a_2p_2 + \cdots + a_np_n$ where $\mathbf{a} = (a_1, a_2, \cdots, a_n)$ is a k-sparse vector, which indicates the coefficients of the function basis. This basic idea is an extension to function fitting, with utilization of sparseness of coefficients.

With above basic idea, CSF selects only $m = ck\log n$ sensors out from n sensors to report their sampling values. Assuming $\{i_1, i_2, \cdots, \imath_m\}$ are the indices of the selected nodes, thus the values of these nodes can be written as

$$\begin{pmatrix} d_{i_1} \\ d_{i_2} \\ \vdots \\ d_{i_m} \end{pmatrix} = \begin{pmatrix} p_1(i_1) & p_2(i_1) & \cdots & p_n(i_1) \\ p_1(i_2) & p_2(i_2) & \cdots & p_n(i_2) \\ \vdots & \vdots & \ddots & \vdots \\ p_1(i_m) & p_2(i_m) & \cdots & p_n(i_m) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

According to (4) in CS theory, the k-sparse vector can be recovered with knowing m samples and the function base P. In most cases of natural information monitoring system, the data can usually be sparsely represented in the Discrete Cosine Transform (DCT) domain, so DCT is a good choice to form a function base which works well enough to provide accurate and robust recoveries.

Then we look into how much improvement in energy-efficiency CSF can bring in a data gathering sensor net-

	Order of # of packets		
	Chain	Grid	Binary Tree
Baseline	n^2	$n\sqrt{n}$	$n \log n$
CDG	$kn \log n$	$kn \log n$	$kn \log n$
CSF	$kn \log n$	$k\sqrt{n} \log n$	$k \log^2 n$

Table 1: Efficiency of three methods when collecting data from n nodes in different topologies

work. Fig.(3) demonstrates CSF in a binary-tree-structrued network. In such a network of n nodes, the depth of the tree is $\log n$, therefore collecting m random nodes' data needs only $O(k \log^2 n)$ transmissions which is amazingly a sub-linear number of the total scale of the network. In a network consisting of over 1000 nodes, more than 60% energy than the basic strategy.

3.3 Comparison of CDG and CSF

Although both CDG and CSF utilize CS technique for compression, CDG uses local encoding at each node and CSF explores the inner correlation between data to calculate the satisfying function.

Table 1 intuitively compares the number of packets that have to be transmitted in the network in order to collect n pieces of data via three different methods, where k is a parameter which measures the inner correlation among the data that is collected.

From the table we can generally conclude that CDG and CSF both provide efficient data gathering in chain topology and some 2D topologies, but in some topologies, such as binary-tree-structured networks, CDG even costs more than the basic strategy while CSF keeps high efficiency no matter in which cases.

4 Simulation and example

In this section, we use an simple example to demonstrate CDG and CSF, as well as provide comparison about their performance on power consumption, recovery accuracy, storage and computational cost.

4.1 Simulation Settings

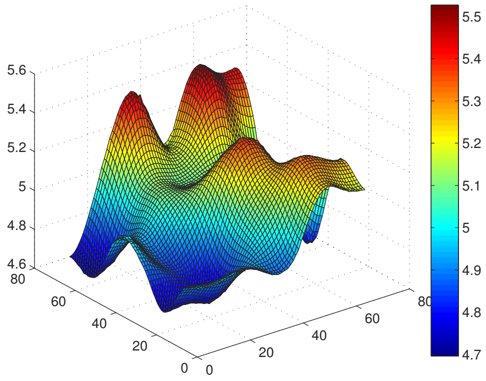


Fig. 4: Data in sensing area

Within a sensing area, $n = 4096$ sensor nodes are uniformly distributed in a 64×64 grid. The data in the sensing area that is to collected is assumed mixture Gaussian distributed. The system collects data every minute, and Fig.(4) shows a snapshot of the data for one moment.

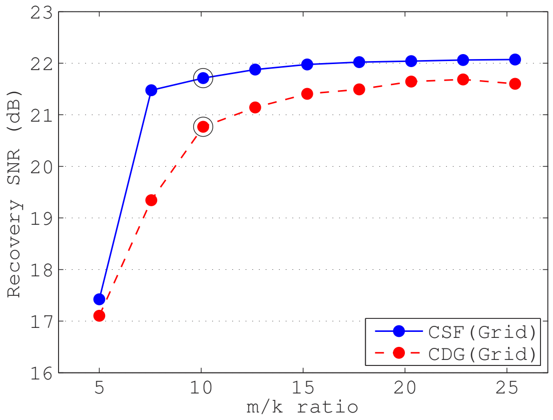


Fig. 5: Recovery accuracy according to different m/k-ratio

	Disk occupation	Normalized ratio	Computational cost
Baseline	~8.4GB	100%	0
CDG	~4.0GB	48%	$O(n^3)$
CSF	~4.0GB	48%	$O(n^3)$

Table 2: Storage and computational cost comparison

Each piece of data is assumed to be a real number occupying 4B, thus the size of one snapshot is 16KB. The data is, naturally, highly correlated in temporal and spacial domains. Each sensor node can communicate with its adjacent neighbors.

For simplicity, the sparsity parameter k is set to be 4% of n. It is reasonable because k-biggest coefficients in DCT domain of the original data contain more than 95% of the total energy in this case.

In the network simulation, CSMA is employed as the MAC protocol, and a 100s period of network-organizing is firstly performed before collecting data.

4.2 Simulation results

In this section, we list some results comparing CDG, CSF and non-CS method in power consumption, collection accuracy, storage cost and computational cost based on previously mentioned settings.

The baseline method honestly collects every single piece of data through the whole network leading to exactly accurate collection of data, so we only plot the accuracy comparison between CDG and CSF in Fig.(5) according to different m/k-ratio. The Signal-to-Noise-Ratio (SNR) is used to measure the quality of recoveries. Though CS-based methods cannot collect exact data, we have confidence of highly accurate recoveries. In addition, CSF recovers better than CDG provided the same m/k-ratio. The little circles in the figure indicate an obvious threshold of an necessary m/k-ratio.

Fig.(6) shows power consumption along the time axis. From the figure we can see after the first 100s, which is a network-organizing period, data gathering process begins, and CSF consumes least power among three methods, and CDG also performs much efficiently than the baseline.

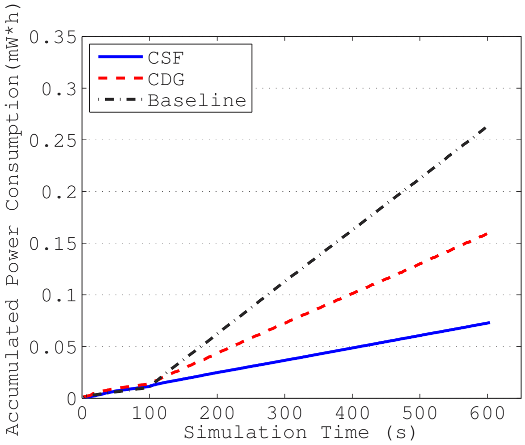


Fig. 6: Recovery accuracy according to different m/k-ratio

The storage and computational comparison between three methods is listed in Table 2. The first column counts the total disk occupation of the data collected for one year. Not surprisingly, CS-based methods requires less storage space because they collect fewer samples. The second column shows the normalized comparison according the baseline case. However, CS-base methods need more computational resource since the full version of data has to be extracted from the collected, but compressed data. Commonly used CS recovery algorithms have the time complexity of $O(n^3)$, but CDG needs to separately reconstruct each random matrix before recovering, so CDG cost a bit more computational resource than CSF.

5 Conclusion

In this paper, we introduce two state-of-the-art methods for efficient data gathering in sensor networks which both take the advantage of Compressive Sensing techniques: Compressive Data Gathering and Compressed Sparse Functions. Through description of mechanism and an example, we illustrate the high performance of CS-based methods no matter in total power consumption or storage efficiency.

Though accuracy loss exists in these methods, both CDG and CSF can guarantee confident recovery by collecting enough samples. CS-based methods need more computational resources than trivial methods, however, we believe, it is valuable to have a highly-energy-efficient network via sacrificing a little, but cheap computational resources.

References

[1] E. Candès, T. Tao. *Decoding by linear programming*. IEEE Trans. Inf. Theory, Vol. 51.

[2] D. Donoho. *Compressed Sensing*. IEEE Trans. Inf. Theory, volume 52, 2006.

[3] Emmanuel J. Candes, Michael B. Wakin. *An Introduction to Compressive Sampling*. Signal Processing Magazine, IEEE 25.2 (2008): 21-30.

[4] C. Luo, F. Wu, J. Sun, *Compressive Data Gathering for Large-scale Wireless Sensor Networks*, in Proc. ACM Mobicom'09.

[5] Liwen Xu, Yuexuan Wang, Yongcai Wang. *Major Coefficients Recovery: a Compressed Data Gathering Scheme for Wireless Sensor Network*. 2011 IEEE Global Communications Conference.

[6] Liwen Xu, Xiao Qi, Yuexuan Wang, Thomas Moscibroda. *Efficient Data Gathering using Compressed Sparse Functions*. IEEE INFOCOM 2013, April 14-19, 2013.

[7] Xufei Mao, Xin Miao, Yuan He, Tong Zhu, Jiliang Wang, Wei Dong, XiangYang Li, Yunhao Liu. *CitySee: Urban CO_2 Monitoring with Sensors*, (PDF), IEEE INFOCOM 2012, Orlando, Florida, USA, March 25-30, 2012.

[8] Wu, Xiaopei, and Mingyan Liu. *In-situ soil moisture sensing: measurement scheduling and estimation using compressive sensing.* Proceedings of the 11th international conference on Information Processing in Sensor Networks. ACM, 2012.

[9] J. Luo, L. Xiang, and C. Rosenberg. *Does Compressed Sensing Improve the Throughput of Wireless Sensor Networks?*. ICC09, pp.1-6, 2009.

[10] Ahmed, Nasir, T. Natarajan, and Kamisetty R. Rao. *Discrete cosine transform.* Computers, IEEE Transactions on 100.1 (1974): 90-93.