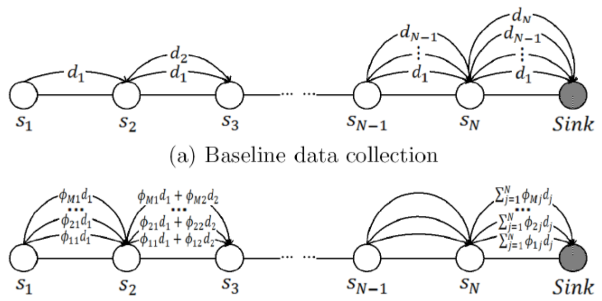
CS-541 Wireless Sensor Networks

Homework #1

Due date 26/3/2018

Part 1: Compressive data gathering

In the first part of the HW, you will explore the potential of Compressed Sensing (CS) in data gathering in WSN. We will assume a network of n sensors, where only k observe a non-zero value (1), while the rest do not (0). In non-CS setting, each sensor will have to transmit its measurements to nodes in its path, which will forward its measurements to the sink. The total number of packets is thus #sensors x #nodes_in_each_path. In CS, a random node is selected which will transmit its measurements to the sink through its path. Instead of simply forwarding the packet, each node (source and intermediate) will multiple its own measurement with a random number and add it to the packet before transiting it to the next node in the path. The following figure showcases the process and is adapted from Luo, Chong, Feng Wu, Jun Sun, and Chang Wen Chen. "Compressive data gathering for large-scale wireless sensor networks." In ACM Proceedings of the 15th annual international conference on Mobile computing and networking, pp. 145-156., 2009.



(a) Baseline data collection

(b) Compressive data gathering

Formally, assume the following path $s_4 \rightarrow s_9 \rightarrow s_2 \rightarrow s_1$.

If s_2 is selected, the $p_1 = \varphi_{2,1} m_2$ where m_2 is the measured value and $\varphi_{2,1}$ a random number from the normal distribution $\mathcal{N}(0,1)$ which is used by node 2 during packet 1 transmission.

If s_9 is selected next, the $p_2 = \varphi_{2,2} m_2 + \varphi_{9,2} m_9$ where m_9 is the measured value, while $c_{2,2}$ and $c_{9,2}$ are random numbers from the normal distribution $\mathcal{N}(0,1)$ for nodes 2 and 9 during transmission of packet 2. This can be written as $p_2 = \boldsymbol{\varphi}_2^T \boldsymbol{m}$ where $\boldsymbol{m} = [m_1, m_2, \dots, m_n]$ and $\boldsymbol{\varphi}_2 = [\varphi_{1,2}, \varphi_{2,2}, \dots, \varphi_{9,2}, \dots, \varphi_{n,2}]$

If s_4 is selected for packet t, then the acquired packet will be the linear combination of all measurements along the path $s_4 \rightarrow s_9 \rightarrow s_2 \rightarrow s_1$, multiplied with random values in $\boldsymbol{\varphi}_t$ i.e. $p_t = \boldsymbol{\varphi}_t^T \boldsymbol{m}$

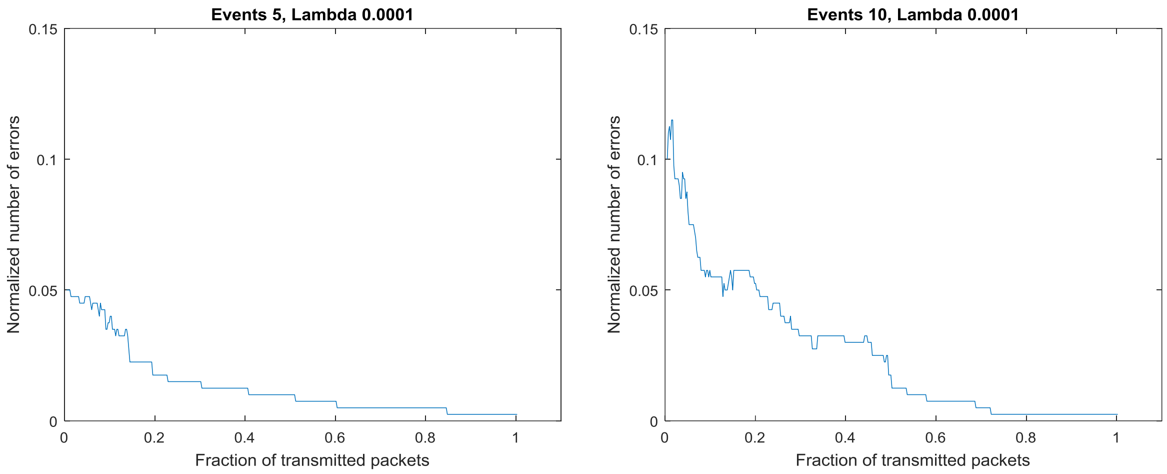
Once the sink receives a number of packets, its solve the sparse coding problem given by $\|\boldsymbol{p} - \boldsymbol{\Phi}^T \boldsymbol{m}\|_2 + \lambda |\boldsymbol{m}|_1$ where $\boldsymbol{p}$ is a vector will all the acquired packets, $\boldsymbol{m}$ is a vector with the measurements and $\boldsymbol{\Phi} \in \mathbb{R}^{t \times n}$ is the matrix containing the random coefficients for packets $1 \dots t$.

For this homework you must consider the following steps

1. Given the connectivity matrix, execute the shortest-path algorithm on each node in order to estimate the path to the sink node (assume this is node 1). (useful functions: graph, shortestpath)
2. Count the number of non-CS packets that have to be transmitted (for each node count the number of relay packets to the sink)

3. Realize the CS data gathering using the matlab lasso function with arguments, the random matrix $\boldsymbol{\Phi}$ and the received packets $\boldsymbol{p}$ in order to estimate the measurements $\boldsymbol{m}$ (set lambda to 0.0001)

Once original measurements are reconstructed, the error will be measured by the number of different measurements between the original measurements and the reconstructed ones over the total number of measurements. The results you will acquire should look like the following figures:



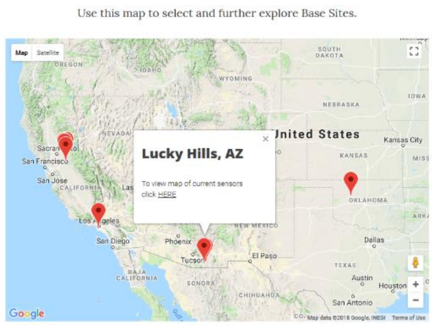
Normlized number of detection errors for 5 and 10 events as a fucntion of tranmitted packets

What is the impact of the number of compressed measurements for a given number of events versus the transmission of the full set of observations?

What is the impact of large number of events (non-zero elements) on the detection rate (test 5 and 10 events)?

More info lecture 5b and Luo, Chong, Feng Wu, Jun Sun, and Chang Wen Chen. "Compressive data gathering for large-scale wireless sensor networks." In ACM Proceedings of the 15th annual international conference on Mobile computing and networking, pp. 145-156., 2009.

Part 2: WSN Data analysis



In the second part of this HW, we will explore the use of low rank matrix and tensor modeling can be employed for the recovery of WSN data with missing measurements. To achieve this goal, we will consider real observations from a wireless soil moisture sensing WSN, the SoilSCAPE[1], deployed in the continental USA.

To recover the missing measurements, two approaches will be explored. For tensor completion, you can utilize the Parallel matrix factorization for low-rank tensor completion[2] method while for matrix completion, the method based on the augmented Lagrange multiplier method [3]

[1] http://soilscape.usc.edu/bootstrap/
[2] Xu, Yangyang, Ruru Hao, Wotao Yin, and Zhixun Su. "Parallel matrix factorization for low-rank tensor completion." arXiv preprint arXiv:1312.1254 (2013).
[3] Lin, Zhouchen, Minming Chen, and Yi Ma. "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices." arXiv preprint arXiv:1009.5055 (2010).

This exercise will involve the following steps

- Collect measurements: The full resolution should be: 3 measurements at 5, 15 and 30 cm per hour, for a period of 7 days, from the Lucky Hills, AZ WSN (Nodes #1500-#1500).
- Form the corresponding tensor and the associated matrices.
- Subsample the data such that each sensor takes 1 measurements every 1, 10, and 24 hours.
- Evaluate the performance of matrix and tensor completion with respect to number of collected measurements.

The reconstruction quality can be measured in terms of any of the mean-squared-error (MSE), root-mean-squared-error, (RMSE).

Deliverables

1. Report 2-3 pages (incl. diagrams & results)
2. Code (matlab or similar) with some comments