ICASSP 2016
at Shanghai, China

# Effecient Sensor Position Selection Using Graph Signal Sampling Theory


Akie Sakiyama[1], Yuichi Tanaka[1], Toshihisa Tanaka[1] and Antonio Ortega[1,2]

[1]Tokyo University of Agriculture and Technology, [2]University of Southern California

Email: sakiyama@msp-lab.org, ytnk@cc.tuat.ac.jp, tanakat@cc.tuat.ac.jp, ortega@sipi.usc.edu


## I. INTRODUCTION

### Sensor Placement Problem

- Selects $F$ sensor locations from $N$ locations to optimize performance
- Often has to handle *large number* of sensors *distributed nonuniformly*
- Many conventional researches [1, 2]
  - Assume that the spatial phenomena are modeled as a *Gaussian process*
  - Place the sensors at the most informative locations

### Graph Signal Processing

- Can efficiently analyze *complex, irregular, and high-dimensional* data
- Relationship between data points is represented as edges in a graph

### This Paper

- ***Proposes a sensor selection approach based on the sampling theory for graph signals [4, 5, 6]***
- ***Shows that the conventional approaches can be viewed as graph vertex domain operations***

## II. PRELIMINARIES

### Graph Signals

- Graph: $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ where $\mathcal{V}$ and $\mathcal{E}$ are sets of nodes and links
- Adjacency Matrix: $A(m,n)$ = weight of link between node $m$ and $n$
- Diagonal Degree Matrix: $D(m,m) = \sum_n A(m,n)$
- Normalized Graph Laplacian Matrix (GLM): $\boldsymbol{\mathcal{L}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$

  → eigenvalues: $\{\lambda_i\}_{i=0\dots N-1} = \sigma(\boldsymbol{\mathcal{L}})$ = *spectrum of graph*

  eigenvectors: $\{\boldsymbol{u}_{\lambda_i}\}_{i=0\dots N-1}$
- $\mathbf{X}_{\mathcal{AB}}$ : matrix extracting $\mathcal{A}$ rows and $\mathcal{B}$ columns from $\mathbf{X}$

  $\boldsymbol{x}_{\mathcal{A}}$ : restriction of $\boldsymbol{x}$ to its components indexed by $\mathcal{A}$

## III. SENSOR PLACEMENT PROBLEM

### Problem Setting

- Selecting $|\mathcal{S}| = F$ sensors from $|\mathcal{V}| = N$ possible locations
- Signal $\boldsymbol{f}$ has the Gaussian joint zero-mean distribution:

$$p(\boldsymbol{f}) = \frac{1}{(2\pi)^{\frac{N}{2}}|\mathbf{K}|}\exp\left(-\frac{1}{2}\boldsymbol{f}^T\mathbf{K}^{-1}\boldsymbol{f}\right) \quad \mathbf{K}: \text{covariance matrix}$$

### Entropy Criterion [1]

★***Sensors are selected so that the uncertainty of a measurement with respect to previous measurements is maximized***

$$\mathcal{S}^* = \mathop{\arg\min}_{\mathcal{S}\subset\mathcal{V}:|\mathcal{S}|=k} H(\boldsymbol{f}_{\mathcal{S}^c}|\boldsymbol{f}_{\mathcal{S}}) = \mathop{\arg\max}_{\mathcal{S}\subset\mathcal{V}:|\mathcal{S}|=k} H(\boldsymbol{f}_{\mathcal{S}}) \qquad \mathcal{S}^c = \mathcal{V}\setminus\mathcal{S}$$

- Conditional entropy: $H(f(y)|\boldsymbol{f}_{\mathcal{S}}) = \frac{1}{2}\log(2\pi e(\mathcal{K}(y,y) - \mathbf{K}_{y\mathcal{S}}\mathbf{K}_{\mathcal{SS}}^{-1}\mathbf{K}_{\mathcal{S}y}))$

### Mutual Information Criterion [2]

★***Selects the locations that more significantly reduce the uncertainty of the rest of the space***

$$\mathcal{S}^* = \mathop{\arg\max}_{\mathcal{S}\subset\mathcal{V}:|\mathcal{S}|=k} H(\boldsymbol{f}_{\mathcal{S}^c}) - H(\boldsymbol{f}_{\mathcal{S}^c}|\boldsymbol{f}_{\mathcal{S}}) := \mathrm{MI}(\mathcal{S}) \qquad \bar{\mathcal{S}} = \mathcal{V}\setminus(\mathcal{S}\cup y)$$

## IV. GRAPH SIGNAL FOR SENSOR SELECTION

- Sensors and observed signals are viewed as nodes and graph signals

### How to determine the connection of nodes

- Assumption: the random signals have following distributions:

$$p(\boldsymbol{f}) \propto \exp\left(-\sum_i\sum_j A(i,j)(f(i)-f(j))^2 - \delta\sum_i f(i)^2\right)$$

$$= \exp\left(-\boldsymbol{f}^T(\mathbf{L}+\delta\mathbf{I})\boldsymbol{f}\right) \qquad \delta: \text{parameter}$$

★***Graph Laplacian matrix is obtained from the covariance matrix***

$$\mathbf{L} = \mathbf{K}^{-1} - \delta\mathbf{I}$$

## REFERENCES


[1] M. C. Shewry and H. P. Wynn, "Maximum entropy sampling," *Journal of Applied Statistics*, vol. 14, no. 2, pp. 165–170, 1987.

[2] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *The Journal of Machine Learning Research*, vol. 9, pp. 235–284, 2008.

[3] A. Gadde, A. Anis, and A. Ortega, "Active semi-supervised learning using sampling theory for graph signals," in *Proc. KDD'14*, 2014, pp. 492–501.

[4] S. Chen, R. Varma, A. Sandryhaila, and J. Kovacevic , "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, 2015.

[5] I. Shomorony and A. S. Avestimehr, "Sampling large data on graphs," in *Proc. GlobalSIP'14*, 2014, pp. 933–936.

[6] A.Gadde and A.Ortega, "A probabilistic interpretation of sampling theory of graph signals," in *Proc. ICASSP'15*, 2015, pp. 3257–3261.


## V. SELECTION USING GRAPH SAMPLING THEORY

★**Chooses sensors so as to *maximize the cut-off frequency in the graph spectral domain:*** $\mathcal{S}^* = \arg\max_{\mathcal{S}} \omega_c(\mathcal{S})$

Cut-off frequncy $\omega_c(\mathcal{S})$ : the maximum frequency of a signal that can be perfectly recoverd from the samples on the subset $\mathcal{S}$

- For optimization, we use three methods based on; eigenvalue [3] (EV), singlar value [4] (SVD), and standard basis [5] (SV)

### Relationships with Existing Methods

- Representing the entropy and MI criteria by using graph Laplacian matrix
  - Using the relationship [6]: $\mathbf{L}_{\mathcal{S}^c} + \delta\mathbf{I} = (\mathbf{K}_{\mathcal{S}^c} - \mathbf{K}_{\mathcal{S}^c\mathcal{S}}(\mathbf{K}_{\mathcal{S}})^{-1}\mathbf{K}_{\mathcal{S}^c\mathcal{S}}^T)^{-1}$

*Entropy criterion*

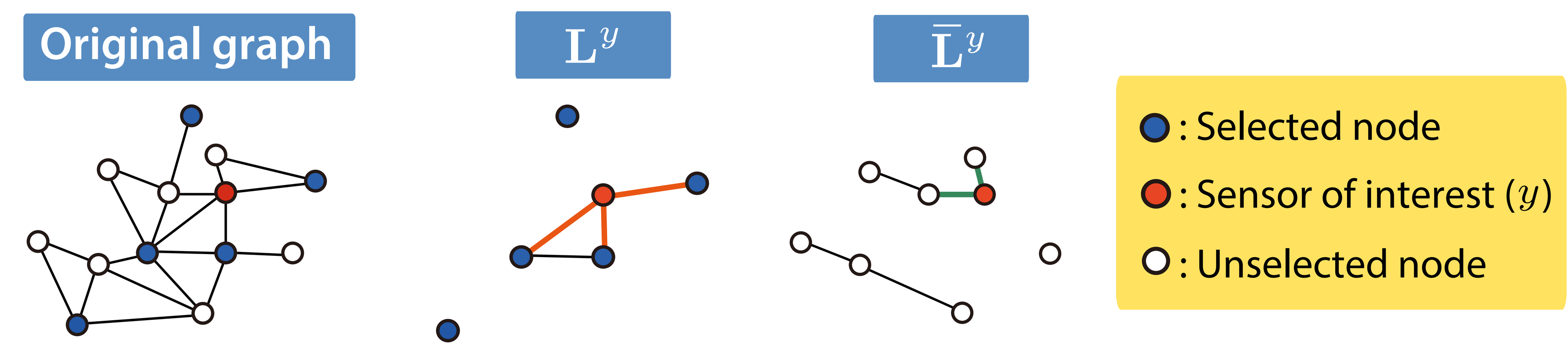
$$y^* \leftarrow \mathop{\arg\max}_{y\in\mathcal{S}^c} \frac{1}{L^y(y,y)+\delta^y}$$

*Choosing node having the weaknest connection with selected sensors*

*Mutual information criterion*

$$y^* \leftarrow \mathop{\arg\max}_{y\in\mathcal{S}^c} \frac{\bar{L}^y(y,y)+\bar{\delta}^y}{L^y(y,y)+\delta^y}$$

*Choosing node having the weaknest connection with selected sensors and the strongest connection with unobserved locations*




$L^y(y,y)$ = degree of $y$th node in graph having nodes $\mathcal{S}\cup y$ and edges in $\mathcal{S}\cup y$

$\bar{L}^y(y,y)$ = degree of $y$th node in graph having unselected nodes $\mathcal{S}^c$ and edges in $\mathcal{S}^c$

***Conventional approach can be viewed as graph vertex domain operations, whereas our approach can be viewed as graph spectral domain approach***

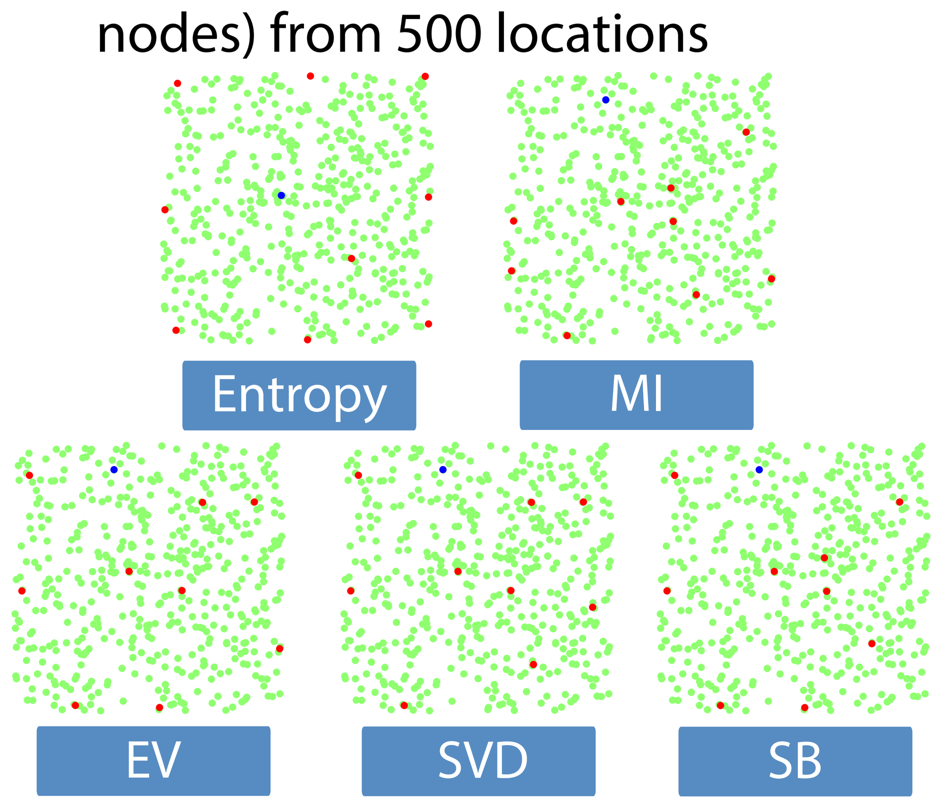
## VI. EXPERIMENTAL RESULTS

### 1. Execution Time

- $|\mathcal{S}| = |\mathcal{V}|/10$ sensors are selected with various number of $\mathcal{V}$


### 2. Choosing Sensor Positions

- 10 selected locations (red and blue nodes) from 500 locations




### 3. Predicting Values on Unobserved Locations

- Test signals are randomly generated according to the GP model and are corrupted by the additive white Gaussian noise
- Original signals are reconstruct only from the signals on $\mathcal{S}$
- Estimated signal is obtained by method in [3]

Performance Comparison (Average of 500 Tested Signals): SNR [dB]

| $\lvert\mathcal{S}\rvert$ | 10 | 20 | 40 | 60 | 80 | 100 |
|---|---|---|---|---|---|---|
| Entropy | -2.77 | -2.17 | 0.28 | 2.02 | 2.13 | 5.75 |
| MI | 3.44 | 5.88 | 7.57 | 8.75 | 9.38 | 10.25 |
| EV | 3.52 | **6.46** | 7.74 | 8.53 | **9.44** | **10.37** |
| SVD | **3.83** | 5.98 | 7.55 | **8.88** | 9.14 | 10.29 |
| SB | 3.61 | 6.34 | **7.86** | 8.21 | 8.67 | 10.31 |

## VII. CONCLUSION

- Optimal sensor selection method based on the graph sampling theorem has been proposed
- All the proposed methods achieved better performance
- As a future work, we will further investigate the theoretical issues and better reconstruction algorithms