Node sampling by partitioning on graphs via convex optimization

Cristian Rusu and John Thompson

The University of Edinburgh

[dstl]

EPSRC
Engineering and Physical Sciences Research Council

The setup: A network composed of m sensors is operating. Each sensor can measure some environmental parameter, for example temperature, and send to a central node.

The problems:

1. Sensors are low size, weight, power and cost (SWaP-C) devices and therefore they cannot perform the measurements continuously for long periods of time
2. Sensors are heterogeneous and have different (but known) power profiles
3. We need to run the network over a period of time

The solution: We propose an algorithm, that takes into account the structure of the network, to schedule the activation of each sensor over time such that sensors are not over/under used and do not deplete their energy.

Characteristics of the solution

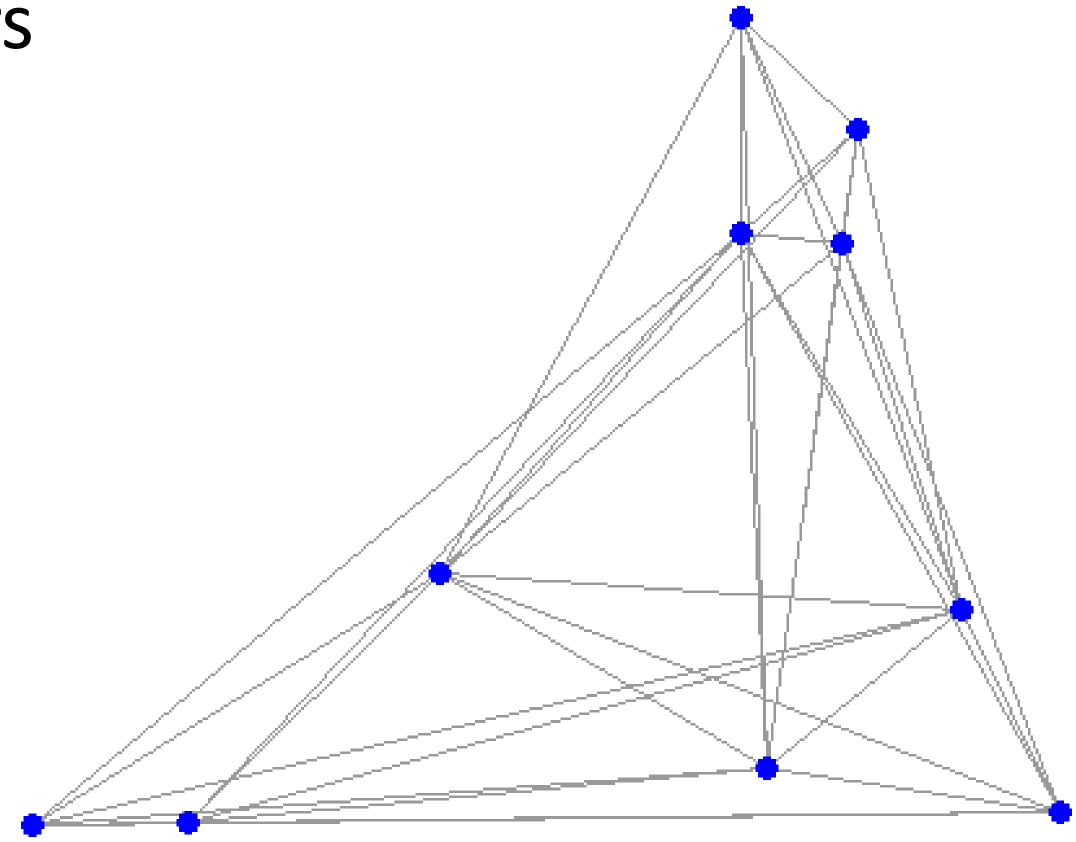
- **We know** something about the sensor network: **how the nodes are correlated**

- **We select** at each time instance only **a subset of the sensors** in the network to perform a measurement **and send** it to the central node

- The goal is **to rebuild** from this subset the **approximate values** for **all the network nodes**

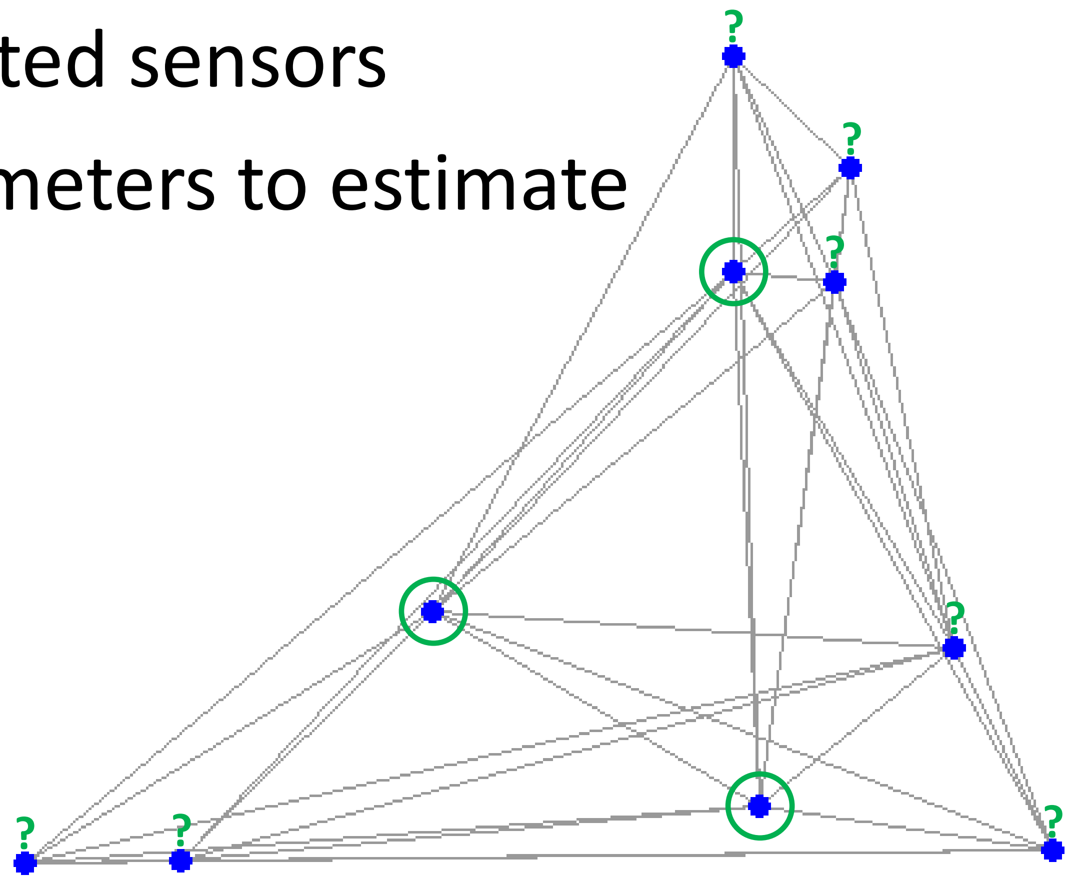
A typical (small) sensor network

10 sensors

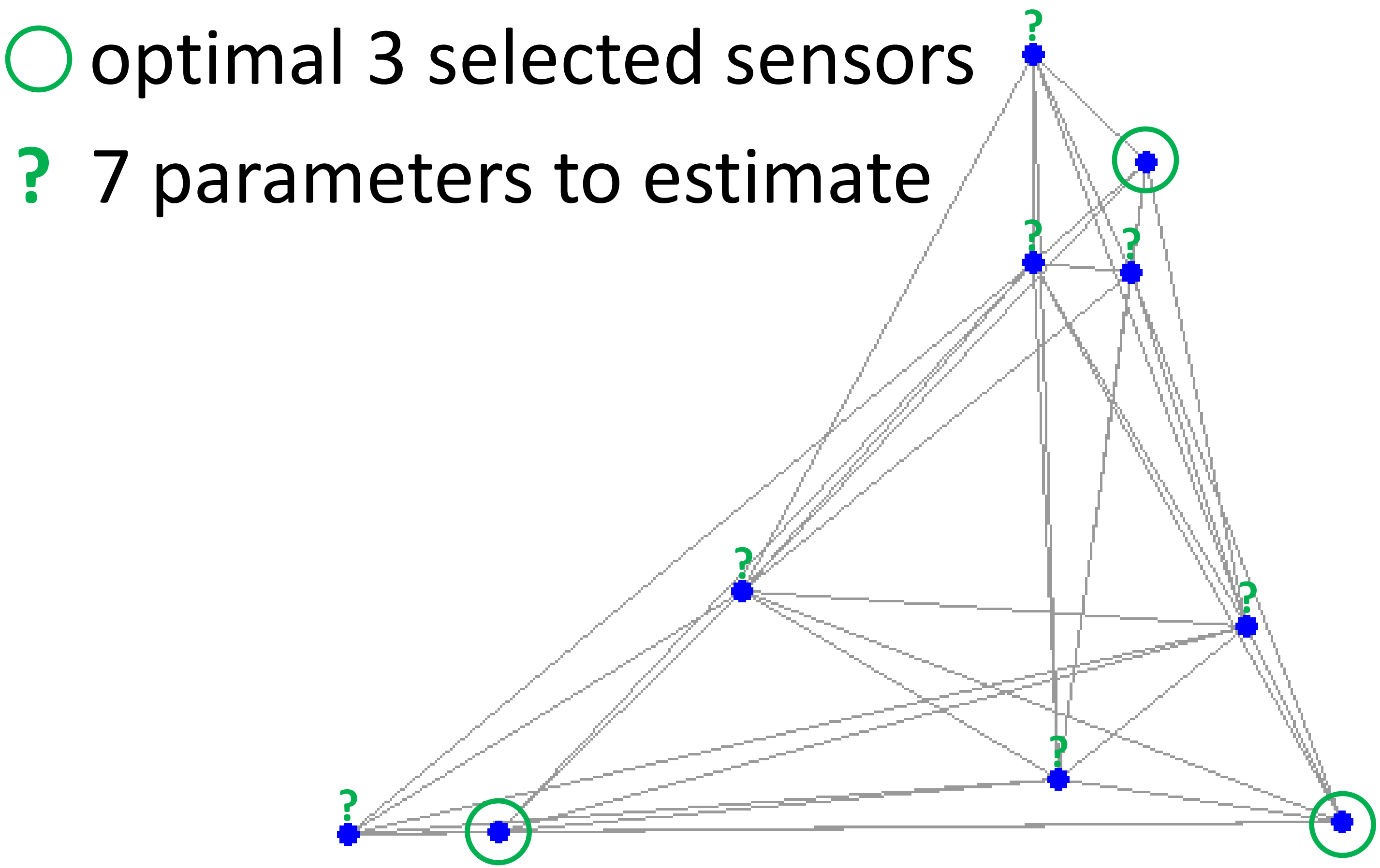


A typical (small) sensor network

- ◯ 3 selected sensors
- ? 7 parameters to estimate



A typical (small) sensor network



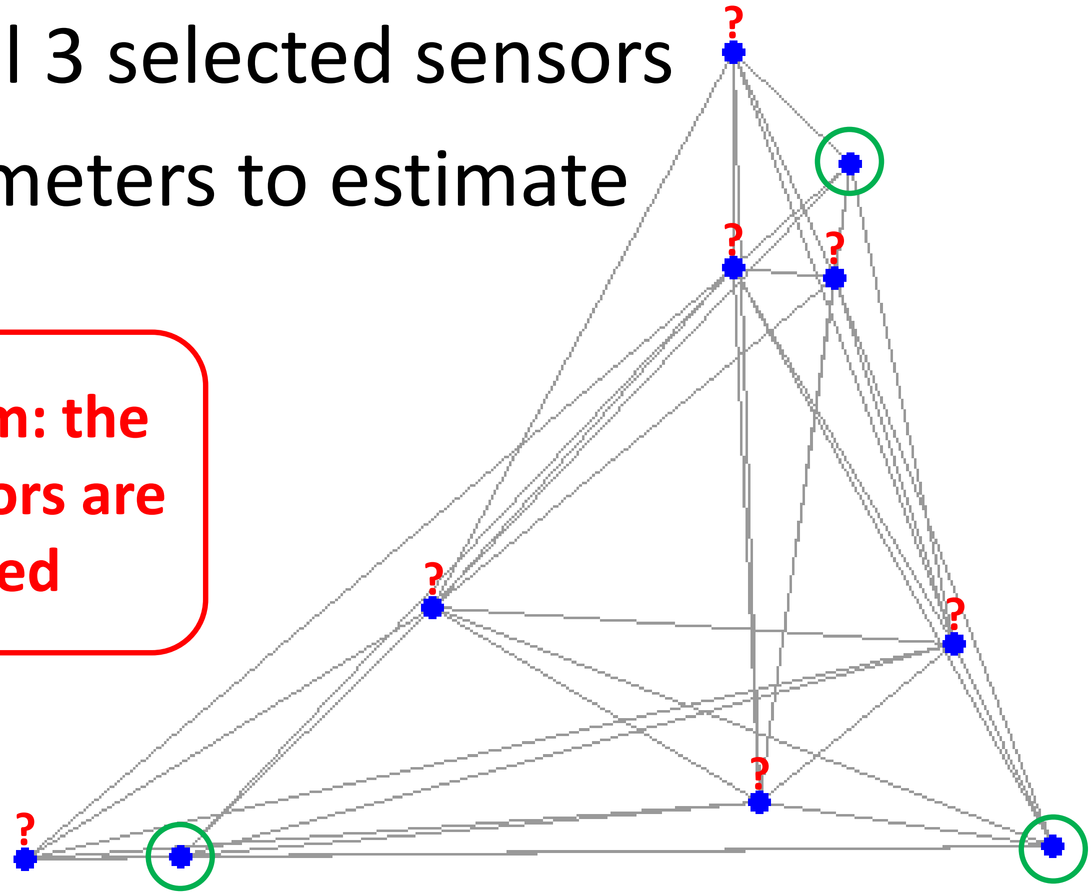
Information theory tells us which is the optimal selection.

A typical (small) sensor network

○ optimal 3 selected sensors

? 7 parameters to estimate

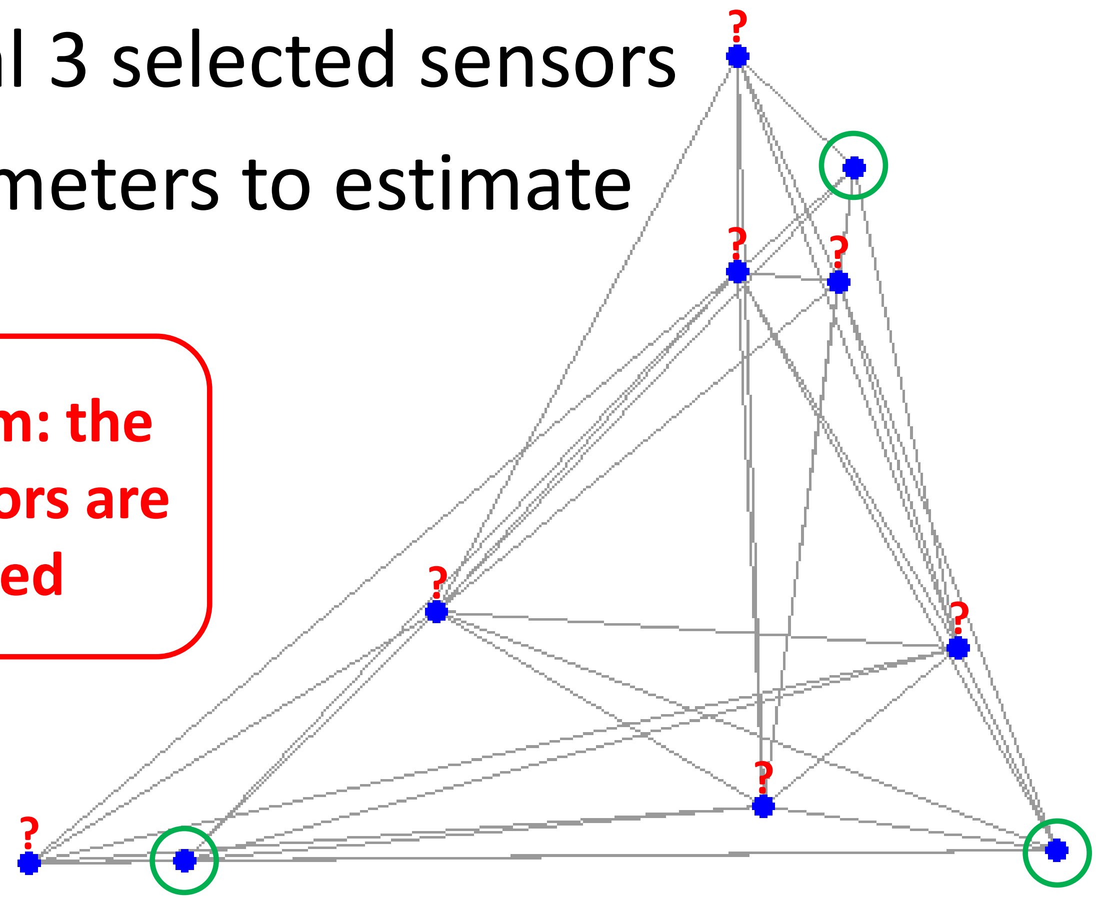
One problem: the same 7 sensors are never used



A typical (small) sensor network

○ optimal 3 selected sensors

? 7 parameters to estimate



One problem: the same 7 sensors are never used

Balance activation of sensors such that all are used <u>equally</u>

Scheduling Table for the Network

- The **Scheduling table Z** defines which sensor is activated when (N sensors × P time slots)
- $z_{ij} = 1$: activate sensor i at time instance j

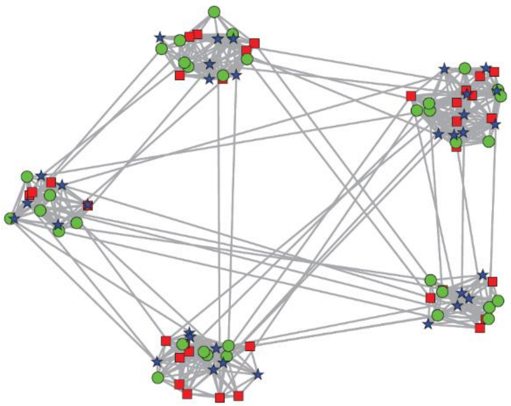
Time →

$$\mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & \dots & 0 \\ 1 & 0 & 1 & \ddots & & 1 \\ \vdots & \vdots & \vdots & & \ddots & . \\ 0 & 1 & 0 & \dots & \dots & 0 \end{bmatrix}$$

↓ **Sensor index**

- *Aim:* the sum of each row should be equal to balance sensor activations

A primer on graph signal processing

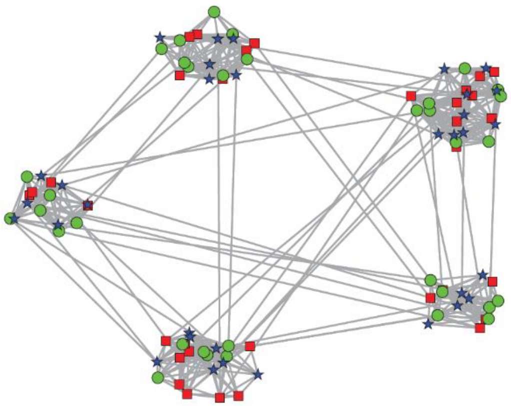


- Graph signal samples $\mathbf{f} \in \mathbb{R}^m$
- Adjacency matrix **W** shows connections between nodes:

$$\mathbf{W} = \begin{bmatrix} 0 & 0.2 & 0.5 & \dots & 0.3 \\ 0.2 & 0 & 0 & \dots & 0 \\ 0.5 & 0 & 0 & & 0.2 \\ \vdots & \vdots & \vdots & \ddots & . \\ 0.3 & 0 & 0.2 & \dots & 0 \end{bmatrix}$$

- Entry w_{ij} shows strength of connections between nodes, e.g. based on distance

A primer on graph signal processing



- Graph signal $\mathbf{f} \in \mathbb{R}^m$
- Square Degree matrix **D:**

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & & 0 \\ \vdots & \vdots & \vdots & \ddots & . \\ 0 & 0 & 0 & \dots & d_m \end{bmatrix}$$

- The entry d_j is just the sum of the jth row of the adjacency matrix **W**

A primer on graph signal processing

- Graph signal $\mathbf{f} \in \mathbb{R}^m$
- Degree matrix **D**, adjacency matrix **W**
- *m*x*m* positive definite Laplacian matrix

$$\mathbf{L} := \mathbf{D} - \mathbf{W} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$$

U^T is called the graph Fourier transform

Λ contains the frequency information about the graph

Bandlimited graph signal processing

- $\mathbf{L} := \mathbf{D} - \mathbf{W} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$
- The graph Fourier transform of signal f

$$\tilde{\mathbf{f}} = \mathbf{U}^T\mathbf{f}$$

- ω-bandlimited signal $\tilde{f}_i = 0$ when $\lambda_i > \omega$

$$\mathbf{f} = \sum_{i=1}^{r} \tilde{f}_i \mathbf{u}_i = \mathbf{U}_{\mathcal{V}\mathcal{R}}\tilde{\mathbf{f}}_{\mathcal{R}}$$

Bandlimited graph signal processing

- $\mathbf{L} := \mathbf{D} - \mathbf{W} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$
- The graph Fourier transform of signal f

$$\tilde{\mathbf{f}} = \mathbf{U}^T\mathbf{f}$$

- ω-bandlimited signal $\tilde{f}_i = 0$ when $\lambda_i > \omega$

$$\mathbf{f} = \sum_{i=1}^{r} \tilde{f}_i \mathbf{u}_i = \mathbf{U}_{\mathcal{VR}}\tilde{\mathbf{f}}_{\mathcal{R}}$$

The idea: we keep only the slow varying (low frequencies) features of the graph, we ignore local and rapid varying behavior (high frequencies) to capture some general characteristics of the graph signals

Exploit bandlimited signals

- $\mathbf{U}_{\mathcal{VR}}$ encodes information about the low frequency components of the graph
- Graph signal reconstruction involves selecting only a subset of nodes to estimate **f**
- How do we go about choosing disjoint groups of nodes to estimate **f** at different times?

The idea: we reconstruct the estimated complete graph signal by using the observed measurements and knowledge about the low frequency components of the graph

Exploit bandlimited signals

- How well can we recover the complete signal?

Selecting more sensors always helps

The properties of the GFT matrix **U** key

Split Sensors into **Disjoint Sets** to average out network energy cost

Aim to equalize **Mean Squared Error** of each set of sensors

The convex optimization problem

$$\underset{y,\ \mathbf{Z}\in[0,1]^{m\times P}}{\text{minimize}} \quad y + \lambda \sum_{p=1}^{P} \mathbf{w}_p^T \mathbf{z}_t$$

relaxation

The objective does two things: tries to minimize error (y) while also pushing the elements in **Z** to be binary (the regularization)

$$\text{subject to} \quad \sum_{p=1}^{P} \mathbf{z}_p = \mathbf{1}_{m\times 1}$$

This constraint guarantees partitioning

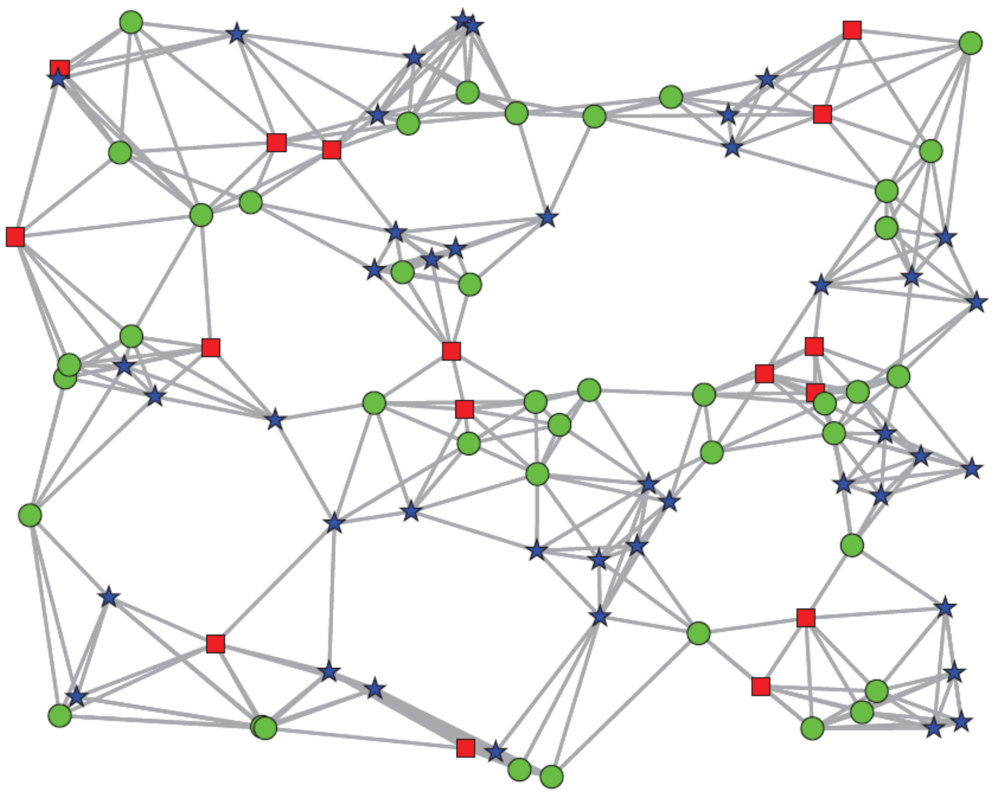
$$\mathbf{1}_{1\times m}\mathbf{z}_p \geq |\mathcal{R}|, p = 1, \ldots, P$$

Make sure each partition can invert the problem

$$\text{tr}((\mathbf{U}_{\mathcal{R}}^T \text{diag}(\mathbf{z}_p)\mathbf{U}_{\mathcal{R}})^{-1}) \leq y, p = 1, \ldots, P$$

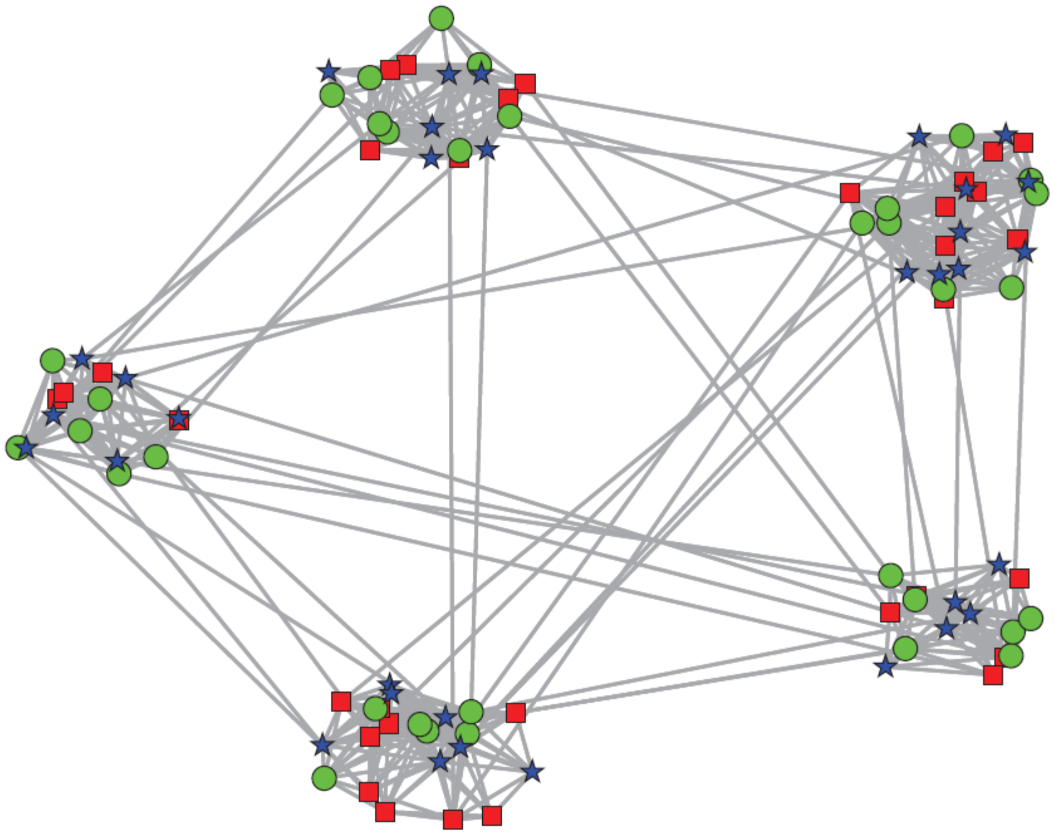
These constraints impose low MSE

Results: sensor network



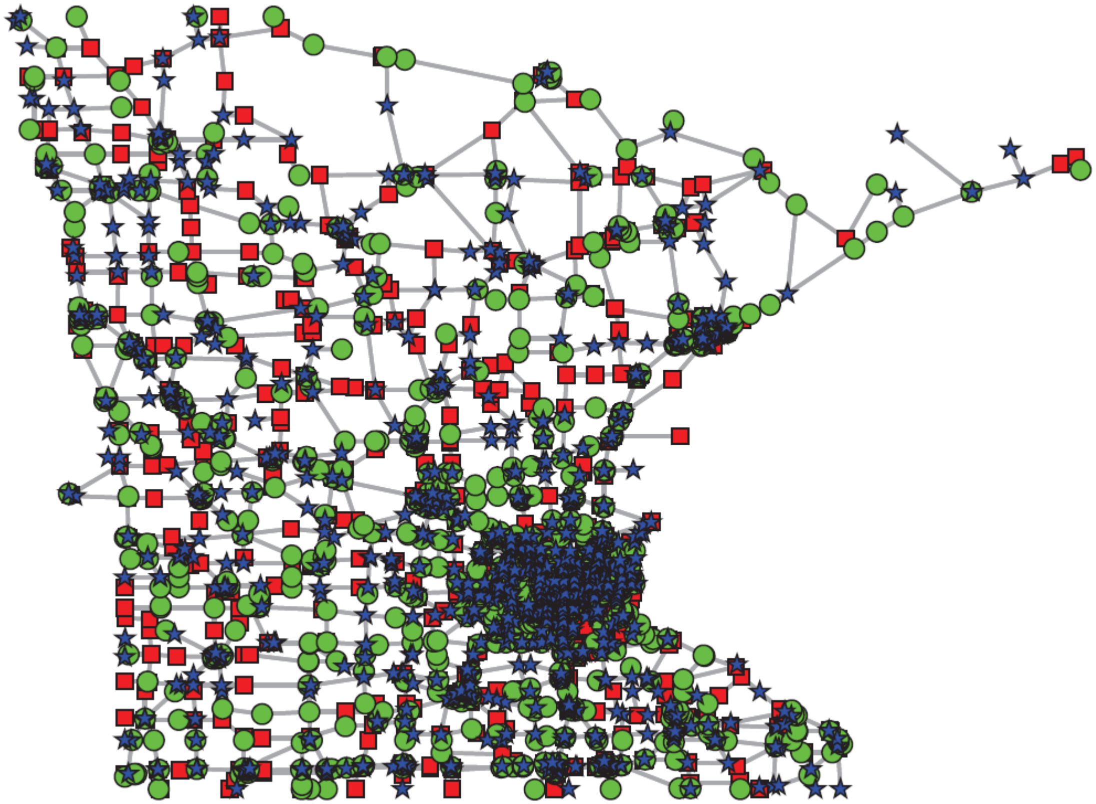
100 nodes, 3 partitions

Results: community network



100 nodes with 5 communities separated into 3 partitions

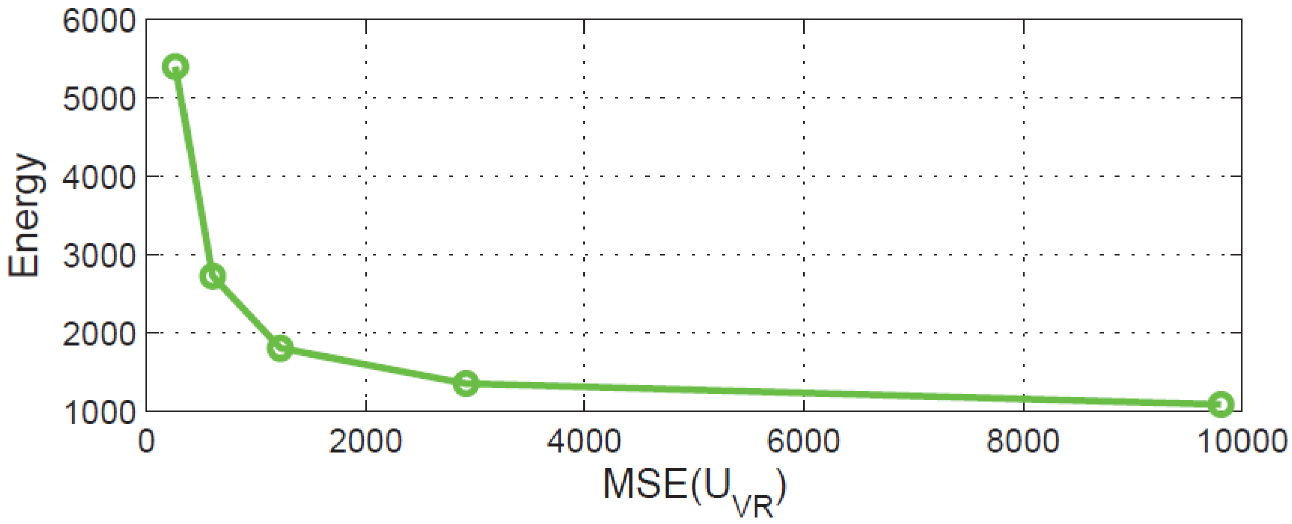
Results: Minnesota network



2642 nodes, 3 partitions

Results: Minnesota network

- Consider there is a **central node**
- The **cost to transmit** to it is $s_i = O(\|\mathbf{c} - \mathbf{v}_i\|_2)$



Trade-off between MSE and energy consumption

Conclusions

- Networks can be **scheduled in time efficiently**
- Exploit the concept of **bandlimited graph signals**
- Scenarios (**SWaP-C devices**) where energy is limited are particular interesting
- Get a **trade-off** between energy consumption and accuracy of operating the sensor network
- Allowing **less accuracy** in parameter estimation, we can get **large gains in energy** consumption

References

- ***Sensor scheduling with time, energy and communication constraints*** (https://arxiv.org/abs/1702.04927)
- ***Balanced sensor management across multiple time instances via l-1/l-infinity norm minimization*** (in Proc. ICASSP 2017)