Georgia Tech College of Engineering
School of Electrical and Computer Engineering

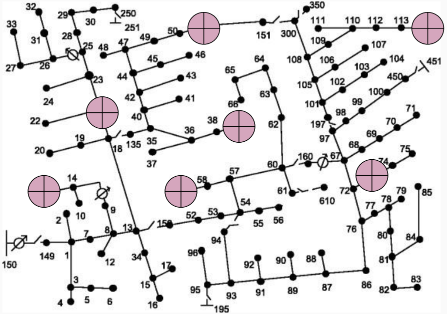
Strategic distribution network sensing

Samuel Talkington, Rahul Gupta, Richard Asiamah, Paprapee Buason, Daniel K Molzahn

December 2024

Traditional sensor placement: offline, static

With sensors



Without sensors

Select *locations* (a subset of nodes) to install sensors and measure continuously

Motivation

Dave Rieken,
Vice President of
Research

HUBBELL

Sensor *placement* in power systems is **wrong**!

We have sensors at ***every* node:** The problem is ***how* we sense!**

We make smart meters, folks!

70% of North American households have smart meters, EIA, 2020.

Here's an algorithm that selects sensor locations

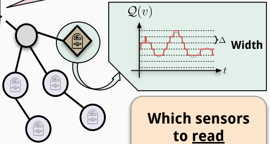
Samuel Talkington

Sensor placement–or sampling?

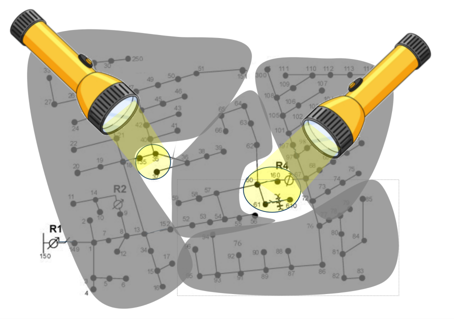
Dave Rieken,
Vice President of
Research

HUBBELL

The key limitation is a
communication
bandwidth limit



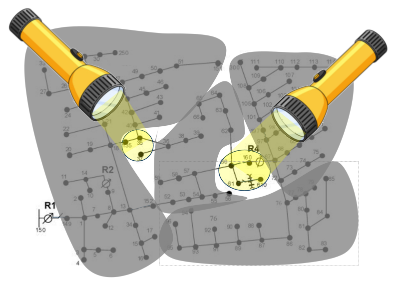
Which sensors
to read



70% of North American households have smart meters, EIA, 2020.

The problem

- Power distribution networks have high levels of sensors already, but with...
- *Limited communication* ***bandwidth***.
- **How do we dynamically monitor these sensor networks efficiently?**
- **i.e., how to move these flashlights around?**



Power distribution systems

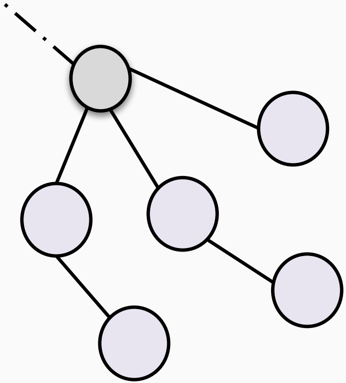


Figure 1: A distribution network can be modeled as a **tree network**, $|\mathcal{N}| = n$, and $|\mathcal{E}| = n - 1$.

Select only a few sensors

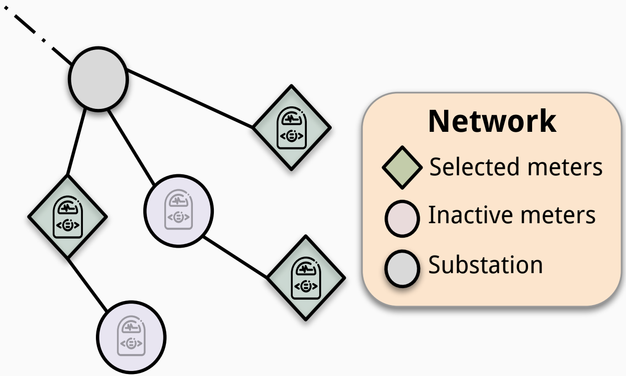


Figure 2: Key idea: we can only select a few sensors

Select $\mathcal{S}$, find worst case in $\mathcal{S}$

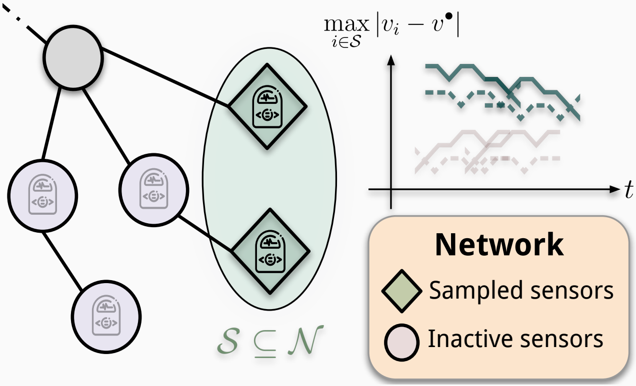


Figure 3: From $\mathcal{S}$, what's the **worst case** voltage?

Grid model

Power flow equations: Recap

- A grid is a graph: $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, with $n = |\mathcal{N}|$ nodes.
- Nodal voltages: $\boldsymbol{u} = \boldsymbol{v} \circ \exp(j\boldsymbol{\theta}) \in \mathbb{C}^n$

  $\boldsymbol{v} \in \mathbb{R}^n$ voltage magnitudes

  $\boldsymbol{\theta} \in (-\pi, \pi]^n$ voltage phase angles
- Nodal power injections: $\boldsymbol{s} = \boldsymbol{p} + j\boldsymbol{q} \in \mathbb{C}^n$

  $\boldsymbol{p} \in \mathbb{R}^n$, "active" power

  $\boldsymbol{q} \in \mathbb{R}^n$, "reactive" power
- $\boldsymbol{Y} \in \mathbb{C}^{n \times n}$ nodal admittance matrix (generalized, complex-valued graph Laplacian)

Power flow equations $\boldsymbol{s} : \mathbb{C}^n \to \mathbb{C}^n$

$$\boldsymbol{s} = \text{diag}(\boldsymbol{u}) \underline{\boldsymbol{Y} \boldsymbol{u}}$$

Linear power flow model

A simple power flow model is formed by inverting the power flow Jacobian at the flat start condition:

$$\begin{bmatrix} p \\ q \end{bmatrix} \approx \begin{bmatrix} G & -B \\ -B & -G \end{bmatrix} \begin{bmatrix} v - 1 \\ \theta \end{bmatrix} \iff \begin{bmatrix} v - 1 \\ \theta \end{bmatrix} \approx \begin{bmatrix} R & X \\ X & -R \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}, \qquad (1)$$

where $G, B \succeq 0$ are the real and imaginary components of the $n \times n$ reduced admittance matrix $Y = G + \mathrm{j}B$, and $R, X \succeq 0$ are the *resistance and reactance* matrices.

Linear power flow model

For *distribution (tree) networks*, the voltage magnitudes $\boldsymbol{v} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ can be approximated as a linear system:

$$\boldsymbol{v} \approx \mathbf{1} + \boldsymbol{R}\boldsymbol{p} + \boldsymbol{X}\boldsymbol{q}.$$

Linear distribution network model

Denoting $\boldsymbol{\epsilon} := \boldsymbol{v} - \mathbf{1}$ as the *voltage magnitude perturbations*, we will analyze:

$$\boxed{\boldsymbol{\epsilon} = \boldsymbol{R}\boldsymbol{p} + \boldsymbol{X}\boldsymbol{q}.}$$

Uncertain linear power flow model

Uncertainty in p, controllable q

Assumptions

Introduce generic uncertainty with the following assumptions:

- The reactive power injections $\boldsymbol{q}$ are set by a linear controller with a gain (ratio of reactive to active injections): $\kappa = q_i / p_i$, that is known for all nodes.
- The active power injections $\boldsymbol{p}$ are **random** with an **unknown distribution** with bounds $p_i \in \left[\underline{p}, \overline{p}\right]$ computed from historical data.

Key points about randomness in p:

- The uncertainty assumptions for $\boldsymbol{p}$ that are **neither Gaussian, independent, nor identically distributed.**
- Only requires **bounds**, which can be arise in engineering contexts such as:
  - Hosting capacity values.
  - Global horizontal irradiance (GHI) clear sky model data.
  - Device manufacturer limits.
  - Optimal power flow or other engineering constraints.

Main result

Theorem (Concentration of Voltages Under Uncertain Power Injections)

Let $\mathbf{p}$ be an n-dimensional vector of random active power injections that are bounded between $\overline{p}$ and $\underline{p}$, and let $\Delta := \overline{p} - \underline{p}$ denote the bound width. Let $\boldsymbol{K}$ be a fixed $n \times n$ control matrix such that $\mathbf{q} = \boldsymbol{K}\mathbf{p}$. Then $\mathbf{v} = \mathbf{1} + (\boldsymbol{R} + \boldsymbol{X}\boldsymbol{K})\,\mathbf{p}$, and perturbations in nodal voltages satisfy

$$\mathrm{E}\left[||\mathbf{v} - \mathbf{1}||_\infty\right] \leq \frac{1}{2}\Delta\,||\boldsymbol{R} + \boldsymbol{X}\boldsymbol{K}||_\infty \sqrt{2\log(2n)}; \tag{2}$$

moreover, for any $t > 0$,

$$\Pr\left[||\mathbf{v} - \mathbf{1}||_\infty > t\right] \leq 2n\exp\left\{\frac{-2t^2}{\Delta^2\,||\boldsymbol{R} + \boldsymbol{X}\boldsymbol{K}||_\infty^2}\right\}. \tag{3}$$

Graph Fourier transform

From the fixed power factor assumption, there is an orthonormal $\boldsymbol{W} \in \mathbb{R}^{n \times n}$, specifically, a **graph Fourier basis**, such that $\psi := \boldsymbol{W}^\top \epsilon$ is the *graph Fourier transform* of the voltage magnitudes. In summary,

$$\epsilon = \underbrace{(\boldsymbol{R} + \boldsymbol{X}\boldsymbol{K})}_{=\boldsymbol{L}^{-1}} \boldsymbol{p} = \boldsymbol{W}\boldsymbol{\Lambda}^{-1}\boldsymbol{W}^\top \boldsymbol{p} = \boldsymbol{W}\psi \tag{4}$$

Benefit: There exist efficient algorithms for sampling sensors with this special structure (more on this later).

Spectral bandit algorithm outline

Strategy

At each time t: the learner picks b nodes to **check the security.**

The set of all strategies is the **set of all subsets of b nodes.**

$$\mathcal{A} = \left\{ \mathcal{S} \in 2^{\mathcal{N}} \ : \ |\mathcal{S}| \le b \right\}, \tag{5}$$

so there are $|\mathcal{A}| = \binom{n}{b}$ possible strategies… challenging in general!

Reward

When the learner has selected sensors $\mathcal{S}_t \in \mathcal{A}$ to *ping*, she observes a *reward* $f : \mathcal{A} \to \mathbb{R}$ that looks like

$$\boxed{f(\mathcal{S}) = \text{Worst case voltage in } \mathcal{S}.} \tag{6}$$

In symbols:

$$f(\mathcal{S}) = \max_{i \in \mathcal{S}_t} |\epsilon_i| = \max_{i \in \mathcal{S}_t} |v_i - 1| = \max_{i \in \mathcal{S}_t} |\langle \mathbf{w}_i, \psi \rangle| \,. \tag{7}$$

This reward is the **maximum voltage magnitude observed in the sampling strategy.**

How to catch a bandit

To pick the best sampling strategy, minimize the **regret**:

$$\boxed{\text{Regret} = \text{E}\left[\text{Best voltage sampling strategy} - \text{Your voltage sampling strategy}\right]}$$

If at first you don't succeed... try again!

Spectral bandit algorithm

Solution approach: At each timestep t, recursively compute an estimate of the *Fourier coefficients* ψ for the voltage magnitudes $\boldsymbol{v}$:

$$\hat{\psi}_t = \underset{\boldsymbol{\psi} \in \mathbb{R}^n}{\arg\min} \sum_{s=1}^{t-1} \left(v_s - \langle \boldsymbol{w}_s, \boldsymbol{\psi} \rangle \right)^2 + \beta \, ||\boldsymbol{\psi}||_{\boldsymbol{\Lambda}}^2 , \tag{8}$$

where $\beta > 0$ is a regularization parameter that you choose. The indices $s = 1, \ldots t-1$ are the **sampled nodes!**

Spectral regularization

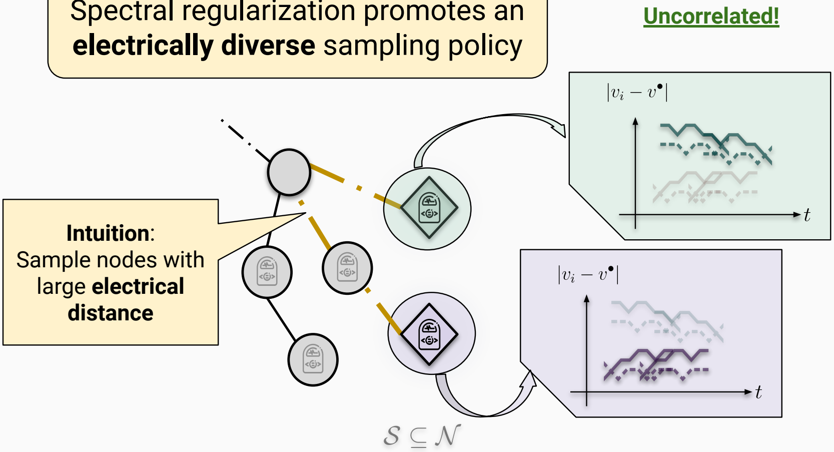
The regularization term, $||\psi||_{\mathbf{\Lambda}}$, promotes predictions of the voltages that are **electrically diverse**:

$$||\psi||_{\mathbf{\Lambda}} := \sqrt{\psi^{\top}\boldsymbol{L}\psi} = \sqrt{\sum_{(i,j)\in\mathcal{E}} y_{ij}(\psi_i - \psi_j)^2}. \tag{9}$$

This is also known as the **Dirichlet energy** of the graph.

Relates to *effective resistance*...check out the paper for more information

Spectral regularization promotes an **electrically diverse** sampling policy



Bandit algorithm solution

The regression problem has a closed form solution at each timestep t:

$$\hat{\psi}_t = \left(\sum_{s=1}^{t-1} \mathbf{w}_s \mathbf{w}_s^\top + \beta \mathbf{\Lambda} \right)^{-1} \left(\sum_{s=1}^{t-1} \mathbf{w}_s v_s \right) := \mathbf{V}_t^{-1} \left(\sum_{s=1}^{t-1} \mathbf{w}_s v_s \right).$$

Where $s = 1, \ldots, t-1 \in \mathcal{N}$ are the **sampled nodes!** The voltage at one node is often similar to its neighbor.

Q: How do we pick those samples?

Answer: Need to **bridge the gap** between the signal processing technique (spectral bandits) and the structural concentration results.

How do we pick those sampled nodes?

Selecting the sample s for each time step:

- Given a sampling budget b, pick the top b nodes ranked by upper confidence bounds on the voltages
- Estimate $\hat{\psi}_t$
- Update **upper confidence bounds** (UCBs) for all nodes:

$$\text{UCB} = \underbrace{\left| \boldsymbol{w}_i^{\top} \psi - 1 \right|}_{\text{exploitation}} + \underbrace{c \, ||\boldsymbol{w}_i||_{V_t^{-1}}}_{\text{exploration}}$$

  The exploration term is determined by our concentration result (see the paper).

- Select the top b nodes greedily
- Continue on...

Extension to sampling strategies

Theorem (Concentration of voltage within sampling strategies)

Let $\mathcal{S} \subseteq \mathcal{N}$ be a sampling of b nodes. Suppose that $\Delta_t := \Delta$ for all t, and suppose that LinDistFlow accurately represents the network model. If the assumptions hold, we have

$$\mathrm{E}\left[\max_{i\in\mathcal{S}} |v_i - 1|\right] \lesssim \frac{1}{2}\Delta \max_{i\in\mathcal{S}} \left|\left|\boldsymbol{\Lambda}^{-1}\boldsymbol{w}_i\right|\right|_2^2 \sqrt{2\log(b)}; \qquad (10)$$

moreover, for all $\epsilon > 0$

$$\Pr\left[\max_{i\in\mathcal{S}} |v_i - 1| > \epsilon\right] \leq 2b\exp\left\{\frac{-2\epsilon^2}{\Delta^2 \max_{i\in\mathcal{S}} \left|\left|\boldsymbol{\Lambda}^{-1}\boldsymbol{w}_i\right|\right|_2^2}\right\}. \qquad (11)$$

Guaranteed performance

The regret of the sampler over m periods is bounded as

$$R_m \leq \tilde{\mathcal{O}}(d\sqrt{m}), \tag{12}$$

where d is the **effective dimension** of the graph Laplacian:

$$d := \max_{i \in \mathcal{N}} \; i \quad \text{s.t.} \quad (i-1)\lambda_i \leq \frac{m}{\log(1 + m/\lambda_1)}, \tag{13}$$

where λ_1 is the smallest eigenvalue of L.

The optimal hyperparameter β depends on the effective dimension, the spectrum of the Laplacian. See our paper or[a] for more.

[a] T. Kocák, et al., "Spectral Bandits", Journal of Machine Learning Research, 21 (1), Jan. 2020.

Key take-away

Question: Why is this an improvement?

Answer: The worst case regret with standard least-squares is

$$R_m \leq \tilde{O}(n\sqrt{m}),$$

where n is the number of nodes. Our result, by **incorporating the graphical structure**[1] of the power flow equations,

$$R_m \leq \tilde{O}(d\sqrt{m}),$$

reduces the scaling factor to the *intrinsic dimension*, $d < n$, of the graph Laplacian.

(This is a huge improvement, as we will see empirically.)

[1] T. Kocák, et al., ”Spectral Bandits”, Journal of Machine Learning Research, 21 (1), Jan. 2020.

New metric: AC regret

Limitations of traditional regret metric

- The traditional regret metric uses the **linear power flow approximation** as the "ground truth" for the best voltage sampling strategy
- Robust theoretical guarantees (more on this later), but not a good empirical metric due to lack of physical realism.
- The **AC power flow (ACPF)** provides a much more realistic model of the power flow equations (non-linear).

In the power system setting we can define the (empirical) metric we term **AC regret:**

$$\text{AC regret} = \text{E}\left[\text{Clairvoyant ACPF voltage sampling strategy} - \text{Your strategy}\right]$$

Note: Involves solving a non-linear estimation problem... no guarantees

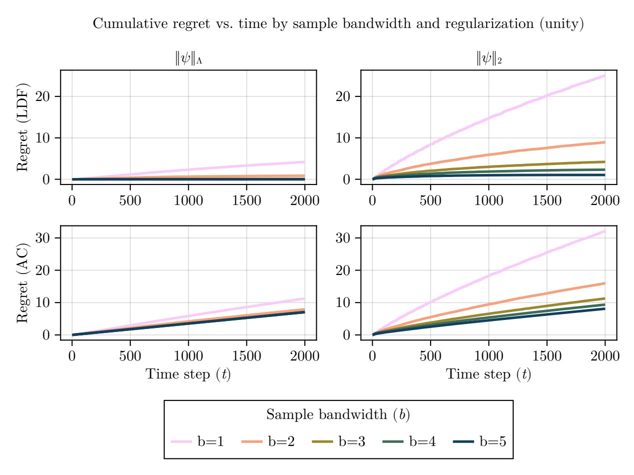


Figure 4: Fixed power factor: Regret of the bandwidth-constrained maximal voltage risk sampler vs. time with spectral (left) and ℓ_2 (right) regularization.

Additional empirical results for randomized control

We can relax the assumption on $\boldsymbol{q} = \boldsymbol{K}\boldsymbol{p}$, and let the entries of $\kappa_i := K_{ii}$ be random, e.g.,

$$\kappa_i \sim \text{Uniform}(\underline{\kappa}_i, \overline{\kappa}_i) \quad i = 1, \dots, n.$$

The following numerical results demonstrate that this works empirically, future work will generalize this.

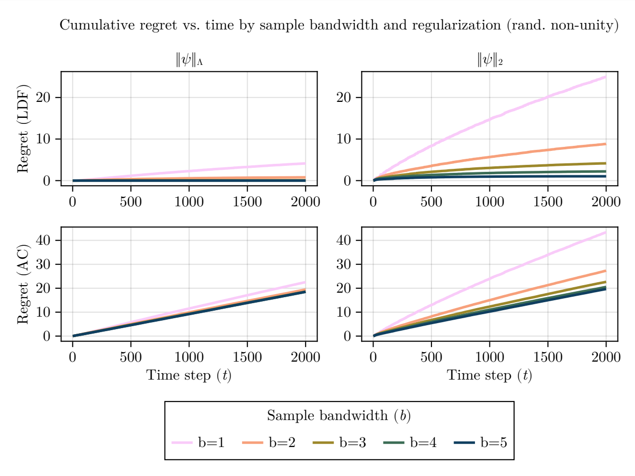


Figure 5: Non-fixed power factor: Regret of the bandwidth-constrained maximal voltage risk sampler vs. time with spectral (left) and ℓ_2 (right) regularization.

Thanks! Keep in touch: talkington@gatech.edu

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-2039655. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation.