Data Driven Modeling for Critical State Estimation in Power Grids
Sanmukh R Kuppannagari, Advisor: Dr. Viktor Prasanna

Introduction
- Power system state estimation critical to ensure smooth grid operations
- State estimation: \( z = Hx + e \)
- Data Injection Attacks: spoof meter readings with \( a = Hc \), so that error is undetected
- LMP: Locational Marginal Price
- Problem Statement: Given \( N \) buses, identify \( k << N \) buses which can determine the system state (as represented by LMPs) accurately

Lasso Regression
\[
\min \left( \frac{1}{2m} \sum_{i=1}^{m} (y_i - \beta_0 - x_i^T \beta)^2 + \lambda \sum_{j=1}^{n} \beta_j \right)
\]
- Inputs: \( x \): load vector, \( y \): price
- Outputs: \( \beta \) weights for each bus
- Results:
  - Lasso with 10 fold validation
  - Just 4-5 buses ensure <5% deviation from 14 bus error rate
- Limitations: Low accuracy

Exploiting Power Flow Structure
- System Pattern Region: Range of load-vectors for which same \( N \) equations become tight
- Theorem: Load pattern space can be partitioned into disjoint convex sets each of which correspond to a unique system pattern region
- Approach (linear cost function):
  - Preprocess historical data and identify \( l \) most frequent system patterns (using LMPs)
  - Generate \( \frac{l(l-1)}{2} \) one-to-one SVM models
  - Use PCA on the \( R^\frac{l(l-1)}{2} \times n \) weight matrix to identify the minimum number of states which account for maximum variance (to meet threshold criteria)

Future Work
- Generalize for quadratic cost functions
- Challenge: LMP not unique for a region

Email: {kuppanna, prasanna}@usc.edu
Website: http://dslab.usc.edu/