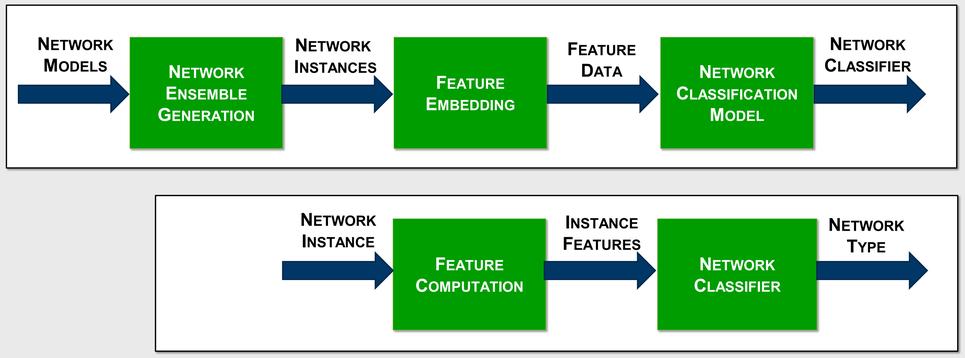
Identifying the Breaking Point: Phase Shift Behavior in Classifying Graph Models

Problem

- Given an observed graph and a collection of generative models, what is the closest generative model?
- Are topological features robust for discriminating graph models?

Approach

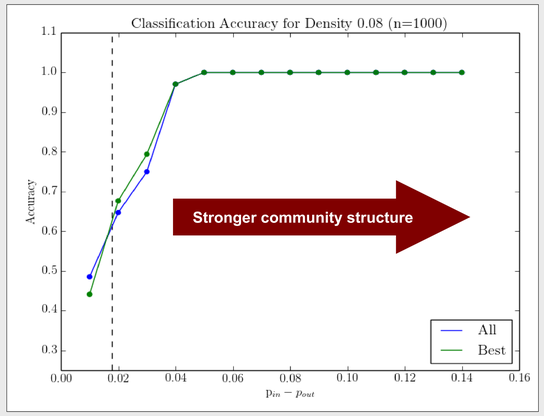
- Given a model, construct an ensemble of graph instances that captures its variation
- Given different models and mappings of their graph ensembles to feature spaces, test if models can be discriminated



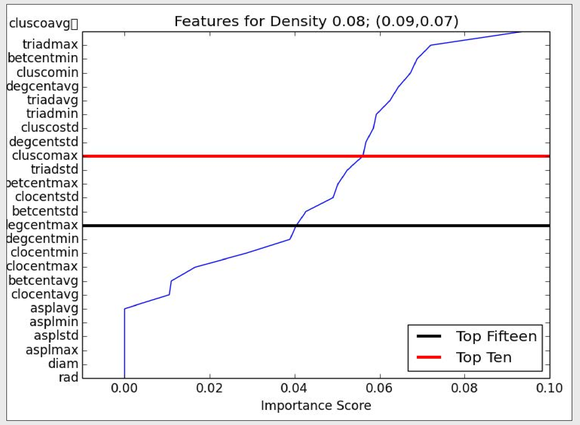
- Main Goals:
  - Full characterization of model space
  - Graphs with realistic properties
  - Complex feature representation
  - Computationally feasible framework

Experimental Setup/Results

- Models: Erdős-Rényi, Stochastic Block Model
- Parameters: Sparse Regime
- Features: Topological
- Classifier: Random Forest (RF)



RF Performance



RF Feature Selection

- RF is able to separate a sparse block from a sparse Erdős-Rényi
- Accuracy of our method approaches the theoretical limit
- Only a small subset of features needed to discriminate
- Discriminating features change for different parameter settings

Future Work

- Analyze RF performance as a function of graph size
- Consider additional graph realistic models
- Consider alternative feature representations of graphs

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

This work is sponsored by Assistant Secretary of Defense for Research and Engineering under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the United States Government. Approved for public release; distribution is unlimited.

Leah Weiner, Rajmonda S. Caceres, Matthew C. Schmidt, Benjamin A. Miller
MIT Lincoln Laboratory