Modeling Dynamic Behavior in Large Evolving Graphs

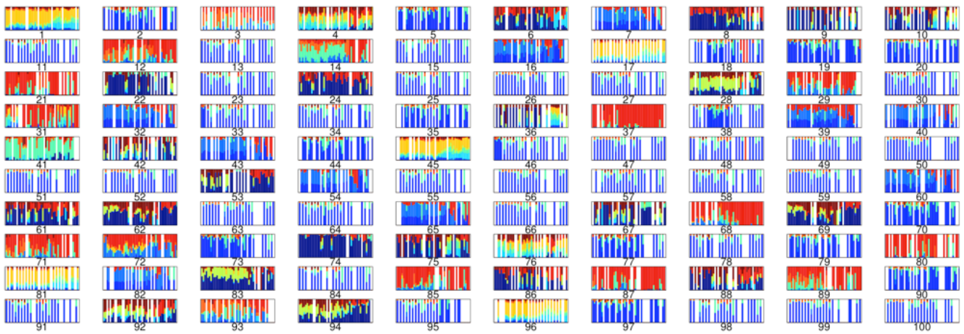
Ryan A. Rossi, Brian Gallagher,
Jennifer Neville, Keith Henderson

Purdue University

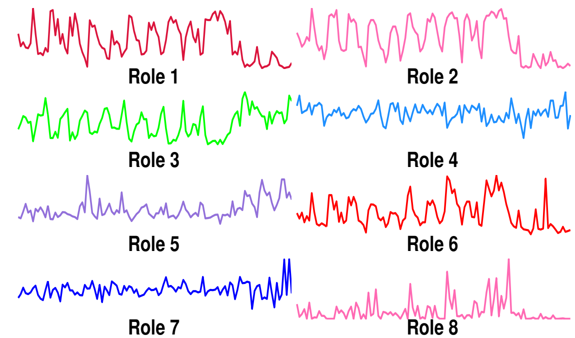
Lawrence Livermore National Laboratory

Modeling Dynamic Graphs

1. **Identify** dynamic patterns in node behavior

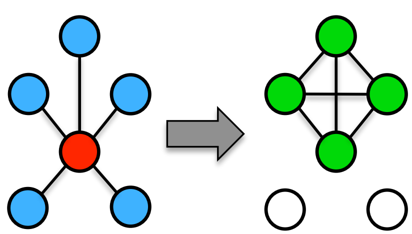


Evolving mixed-role memberships



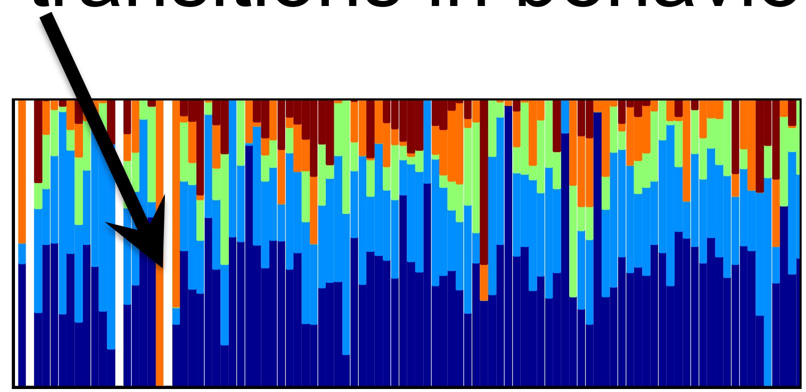
Role contributions

2. **Predict** future structural changes

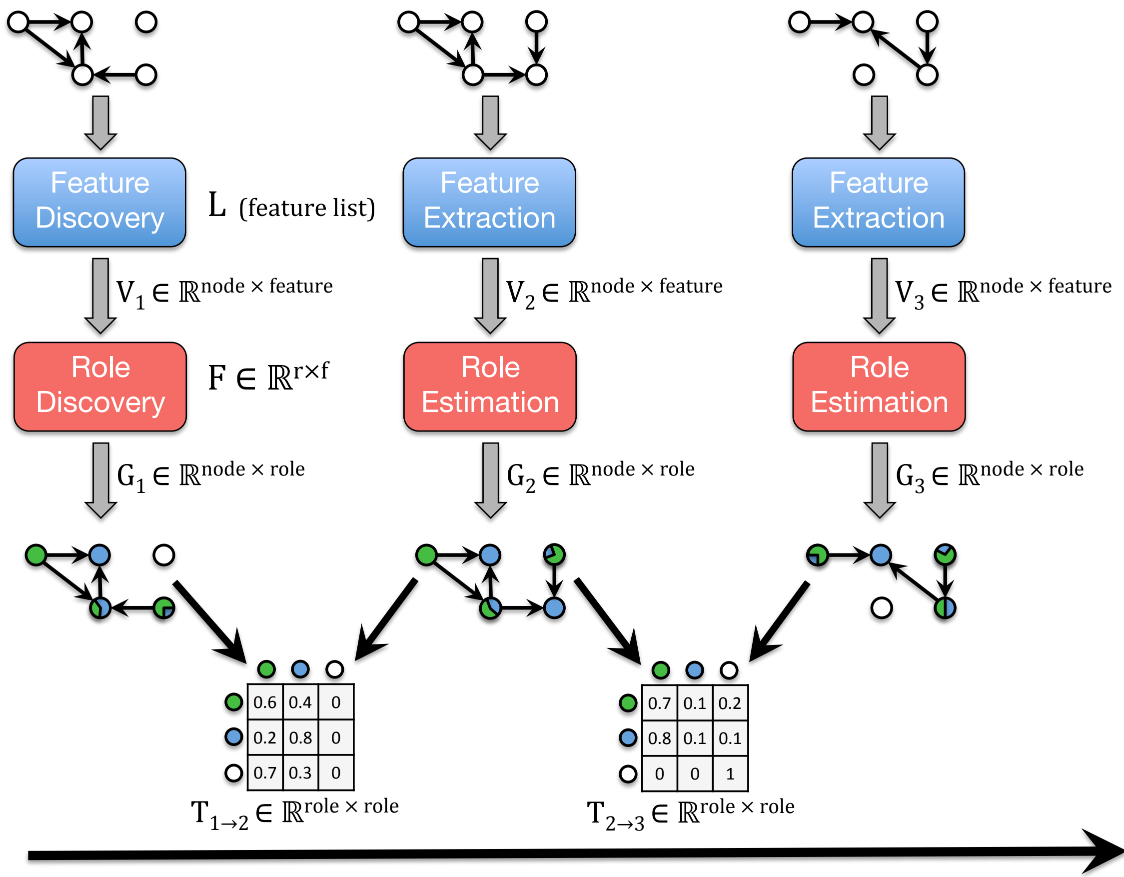


Transition from star to clique

3. **Detect** unusual transitions in behavior



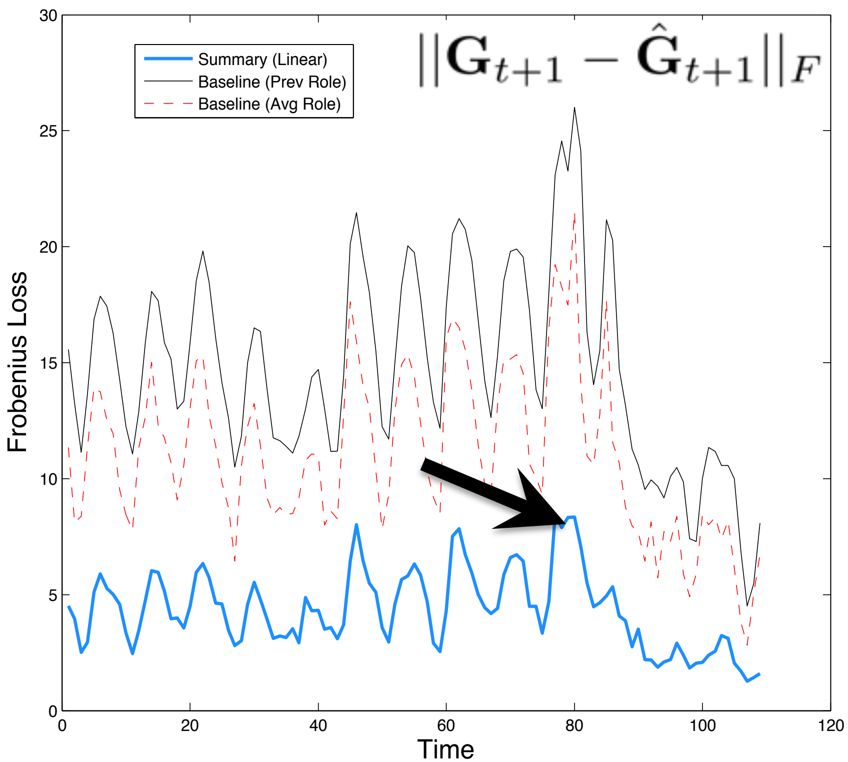
Dynamic Behavioral Mixed-Membership (DBMM) Model



1. Compute set of features
2. Estimate the features on each snapshot graph
3. Learn roles from features using NMF, number of roles selected via MDL
4. Extract roles from each feature matrix over time
5. Use NMF to estimate transition model

The DBMM model is: (1) Scalable for **BIG** graphs (2) Easily parallelizable (3) Non-parametric & data-driven (4) Flexible and interpretable

Predicting Structural Behavior



Twitter

Given G_{t-1} and G_t find a transition model T that minimizes the functional:

$$f(\mathbf{G}_t, \mathbf{G}_{t-1}) = \frac{1}{2}||\mathbf{G}_t - \mathbf{G}_{t-1}\mathbf{T}||_F^2$$

All models predict G_{t+1} using G_t as

$$\mathbf{G}'_{\mathbf{t}+1} = \mathbf{G}_\mathbf{t}\mathbf{T}$$

<u>Summary model</u>: Weight training examples from k previous timesteps

Baseline models: Predict future role based on

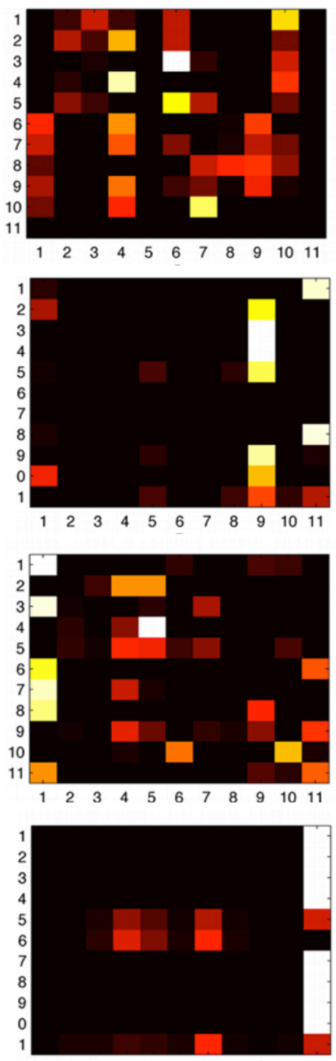
(1) previous role

(2) average role distribution

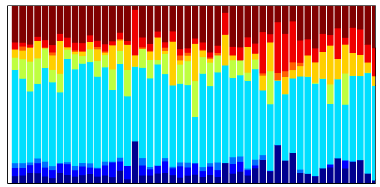
DBMM is more accurate at predicting future behavior than baselines

Dynamic Network Analysis with Roles

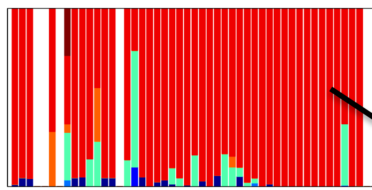
Role transition matrices



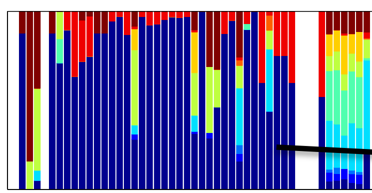
Role proportions over time



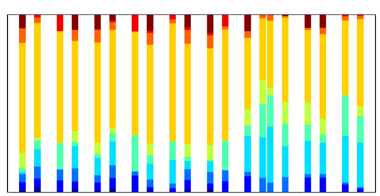
Stationary roles



Homogeneous roles



Abrupt transition



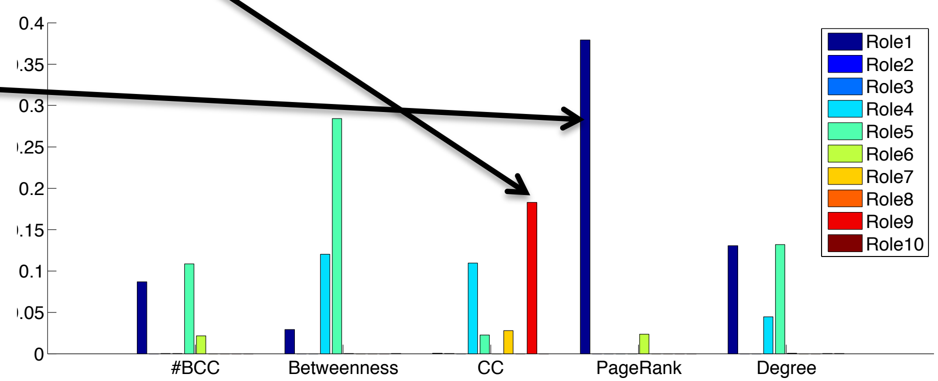
Periodic, similar roles

Roles exhibit many of the traditional time-series patterns

Roles are interpretable

Fit role-model to matrix of network statistics:

$$\mathbf{G}_t\mathbf{E}_t \approx \mathbf{M}_t$$

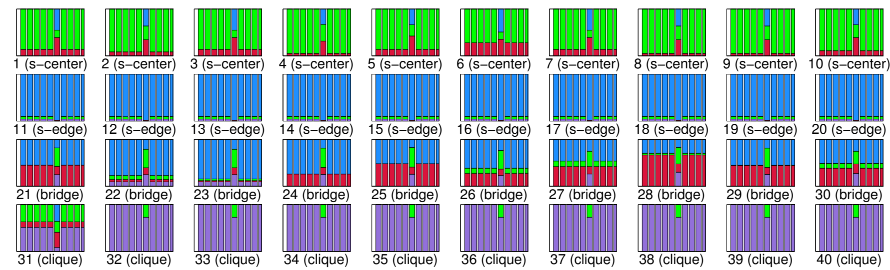


Anomalous Structural Transitions

Problem: detect nodes with unusual structural transitions

Anomaly score:

1. Estimate transition model T for v
2. Use it to predict v's memberships
3. Take the difference from actual



Inject anomalies into synthetic data:
Detected 88.5% over 200 repeated trials

DBMM model finds nodes that are anomalous for only short time-periods