

Persistent Roles in Online Social Networks

Carlotta Domeniconi

Joint work with: Matt Reville and Aditya Johri

George Mason University

ECML PKDD 2016



Motivation

- ▶ Node roles...
 - ▶ permit compact representations of networks
 - ▶ can be useful for understanding temporal patterns of node behavior
- ▶ Do different networks have the same roles?
- ▶ How do people's roles change over time?

What is a Social Network?

- ▶ Nodes are people
- ▶ What about edges?
 - ▶ Relationships (family, friends, etc.)
 - ▶ This is really an ontology about people
 - ▶ Interactions
 - ▶ One-time (“Alice became friends with Bob”)
 - ▶ Recurring (Alice and Bob message multiple times a week)
- ▶ Recurring interactions require allocation of an inelastic resource (e.g., time and attention)
- ▶ Social Interaction Networks
 - ▶ Nodes are people
 - ▶ Edges are dynamic and based on ongoing, pair-wise interaction events

Temporal Network Snapshots

- ▶ Calculate the average time between interactions $\langle \delta t_{ij} \rangle$ for all pairs
- ▶ Calculate an observation window threshold Ω from all $\langle \delta t_{ij} \rangle$
- ▶ Full methodology described in:
 - ▶ Miritello, G., Lara, R., Cebrian, M., Moro, E.: Limited communication capacity unveils strategies for human interaction. *Scientific reports* 3 (2013)
 - ▶ Revelle, M., Domeniconi, C., Johri, A.: Evidence of temporal artifacts in social networks. In: *Proceedings of the 6th International Workshop on Mining Ubiquitous and Social Environments (MUSE)*, pp. 35–42 (2015)

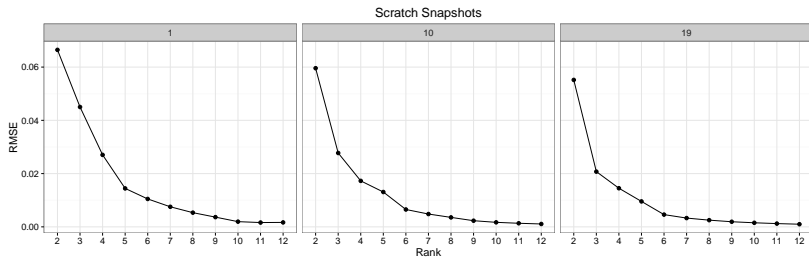
Node Features

	Name	Description
1	<u>In-degree</u>	Count of incoming edges
2	<u>Out-degree</u>	Count of outgoing edges
3	<u>Weighted in-degree</u>	Count of incoming interactions
4	<u>Weighted out-degree</u>	Count of outgoing interactions
5	<u>Reciprocity</u>	Ratio of reciprocated edges over all outgoing edges
6	<u>New activity count</u>	Count of new outgoing edges
7	<u>Social strategy</u>	Ratio of new outgoing edges over all outgoing edges
8	<u>Betweenness centrality</u>	Number of all shortest paths which pass through the node
9	<u>PageRank</u>	PageRank measure of centrality
10	<u>Weighted PageRank</u>	Weighted variant of PageRank
11	<u>Transitivity</u>	Probability any two neighbor nodes are connected (local clustering coefficient)
12	<u>Weighted transitivity</u>	Weighted variant of transitivity

Role Discovery and Membership

- ▶ Each network snapshot is processed independently
- ▶ Decompose node feature matrix $X \in \mathbb{R}^{D \times N}$
- ▶ Two outputs:
 - ▶ Basis matrix (roles) $U \in \mathbb{R}^{D \times L}$
 - ▶ Coefficient matrix (membership weight) $V \in \mathbb{R}^{L \times N}$
 - ▶ Where L is the factorization rank
- ▶ Goal is to find an approximation with low error:
 - ▶ $X \approx UV$

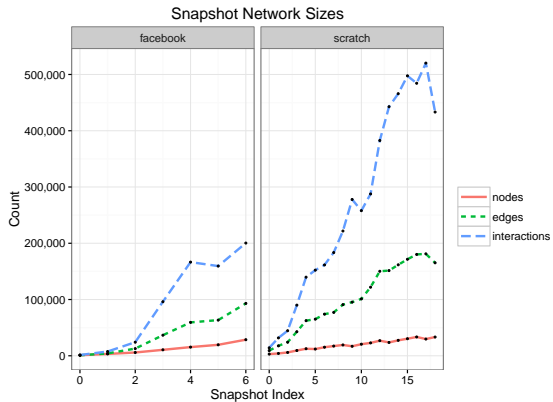
Model Selection



- ▶ Error curves for the first, middle, and final network snapshots in Scratch
- ▶ Select rank of 6 for all snapshots

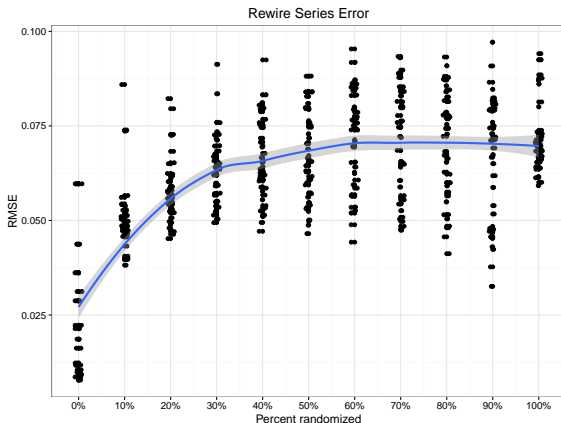
Datasets

- ▶ Facebook Wall Posts
- ▶ Scratch Project Posts

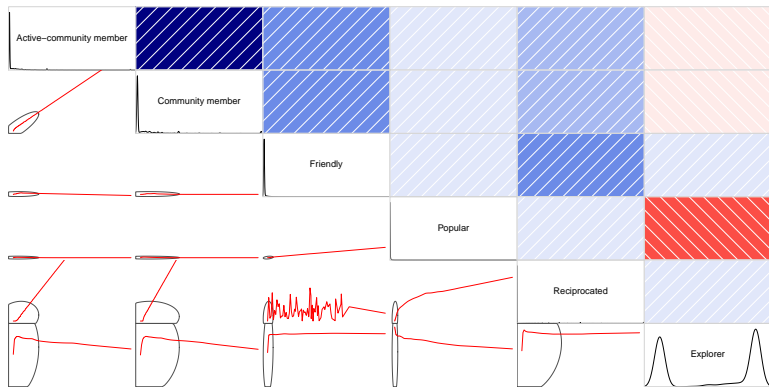


Role Dependence on Network Structure

- ▶ Are the roles describing the networks?
- ▶ Generate approximations using discovered roles
- ▶ Errors plotted for 50 series of randomly rewired networks

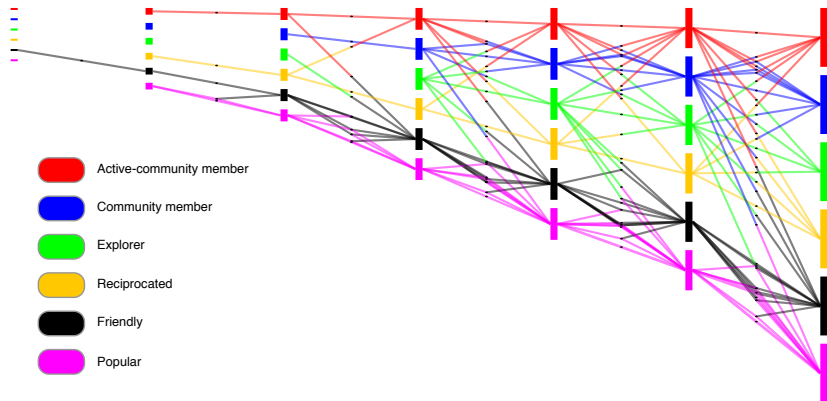


Role Membership



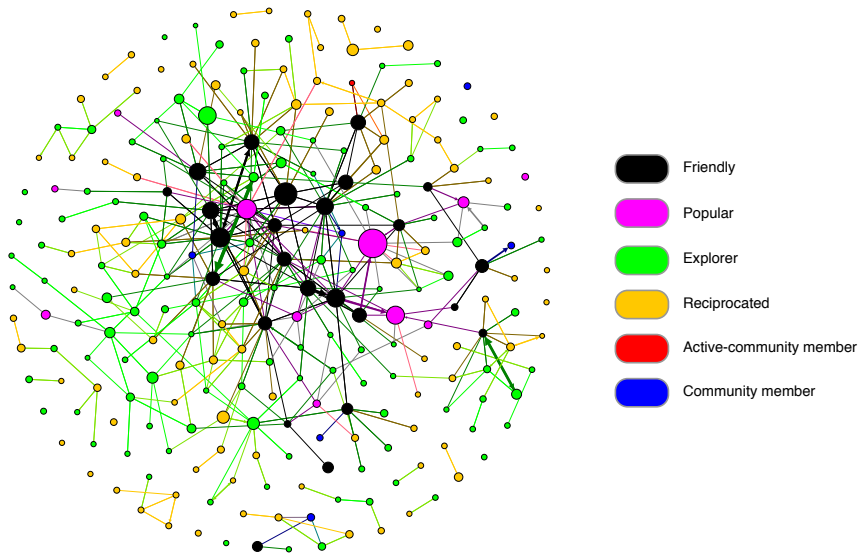
- ▶ Upper diagonal: positive correlation (blue) and negative correlation (red)
- ▶ Diagonal: distribution of role membership weights
- ▶ Role membership correlations tend to be similar across all snapshots

Role Transitions



- The role transitions for top-5% users in each role over all Facebook snapshots

Role Affinity



Conclusion

- ▶ Methodology for discovery of persistent roles
- ▶ Persistent roles enable new comparative analysis of networks
- ▶ Basis for modeling role membership and interaction dynamics

Thank You

