194.050 Social Network Analysis

Topic 04: Equivalence-based role mining

Intermediate presentation

Babiy, Ivan
Chan, Yat Hin
Fan, Wing Yan
Muriasova, Karina
Olsiakova, Terezia

Content

Motivation

Theoretical background

- Definition of role equivalences
- Techniques for finding or quantifying equivalences

Proposed approach

- Which edge lists to use
- Methods for role mining
- Expected results

Motivation and Task Description

Task: Grouping users into equivalence classes that define roles

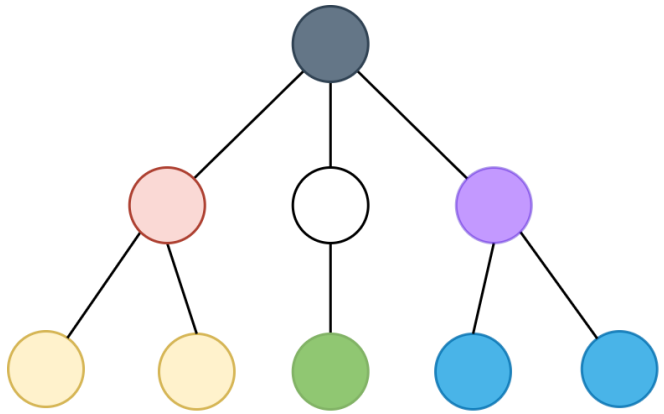
No prior knowledge of user roles

Equivalences are designed in a way that the results may carry a sociological interpretation [1]

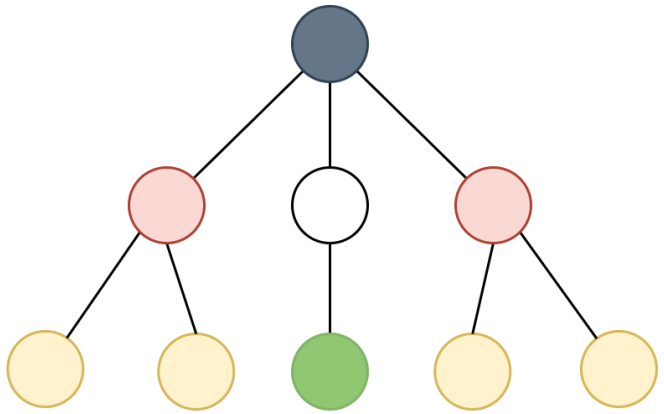
Theoretical background – Equivalence definitions

Comparing Positions: how nodes are embedded in its ego-network [2]

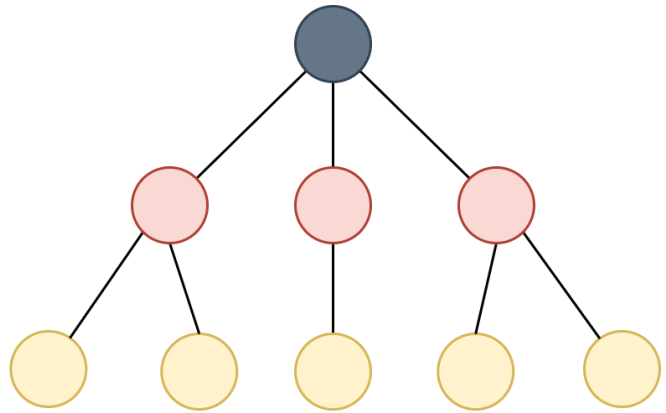
Structural equivalence



Automorphic equivalence



Regular equivalence



Most restrictive

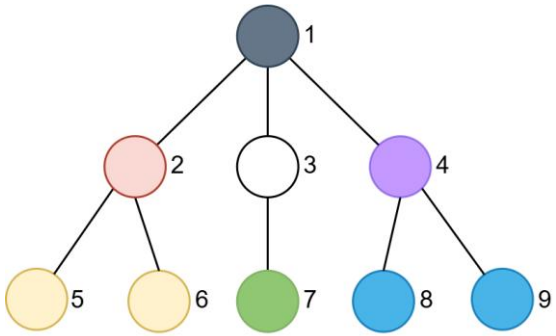
Least restrictive

Theoretical background – Structural equivalence

Exactly identical composition of neighbourhood

e.g. {5, 6}

{8, 9}



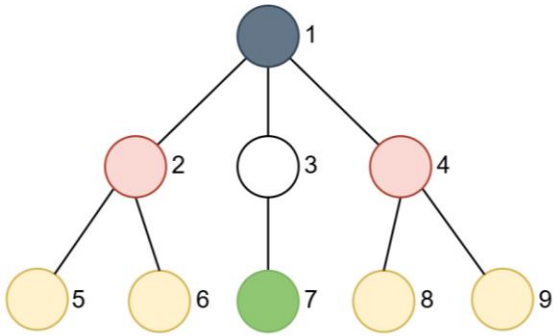
Approximated by K-means with compatible similarity measure
(corrected Euclidian or Manhattan distance function)

Theoretical background – Automorphic equivalence

Identically-shaped neighbourhoods with same degree

The connected neighbours can be different entities

e.g. {5, 6, 8, 9}



RoleSim [3]

Compute RoleSim score matrix of maximal weighted matching of the two vertices' neighbourhood

Iterate until convergence of the maximal weighted matching value

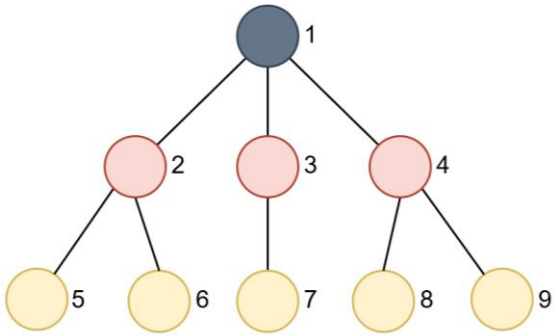
Output: a similarity matrix for performing clustering

Theoretical background – Regular equivalence

Existence of connections with neighbour(s) of the same class

The connected neighbours can be different entities
The degree can be different

e.g. {5, 6, 7, 8, 9}



REGE [4]

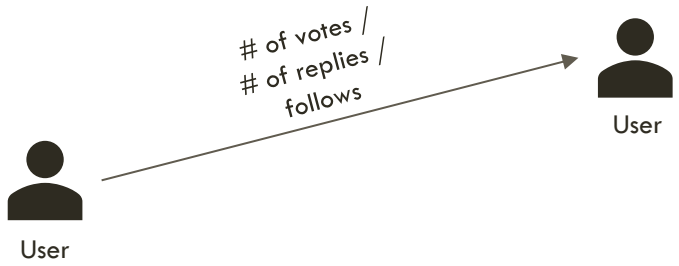
Compute similarity score matrix that quantifies matching edge type (incoming/outgoing/bidirectional) exists in the two vertices' neighbourhood

Iteratively update matrix with the previous matching score as a weight for fixed iteration

Output: a similarity matrix for performing clustering

Proposed approach – Graphs to create

We will create 3 directed graphs with users as nodes as follows:



Network of Users Who Vote on Postings

- Edge list: `df_edge_list_directed_users_votes_to_postings_net.parquet`
- Edge weights: Number of votes

Network of Users Who Reply to Postings

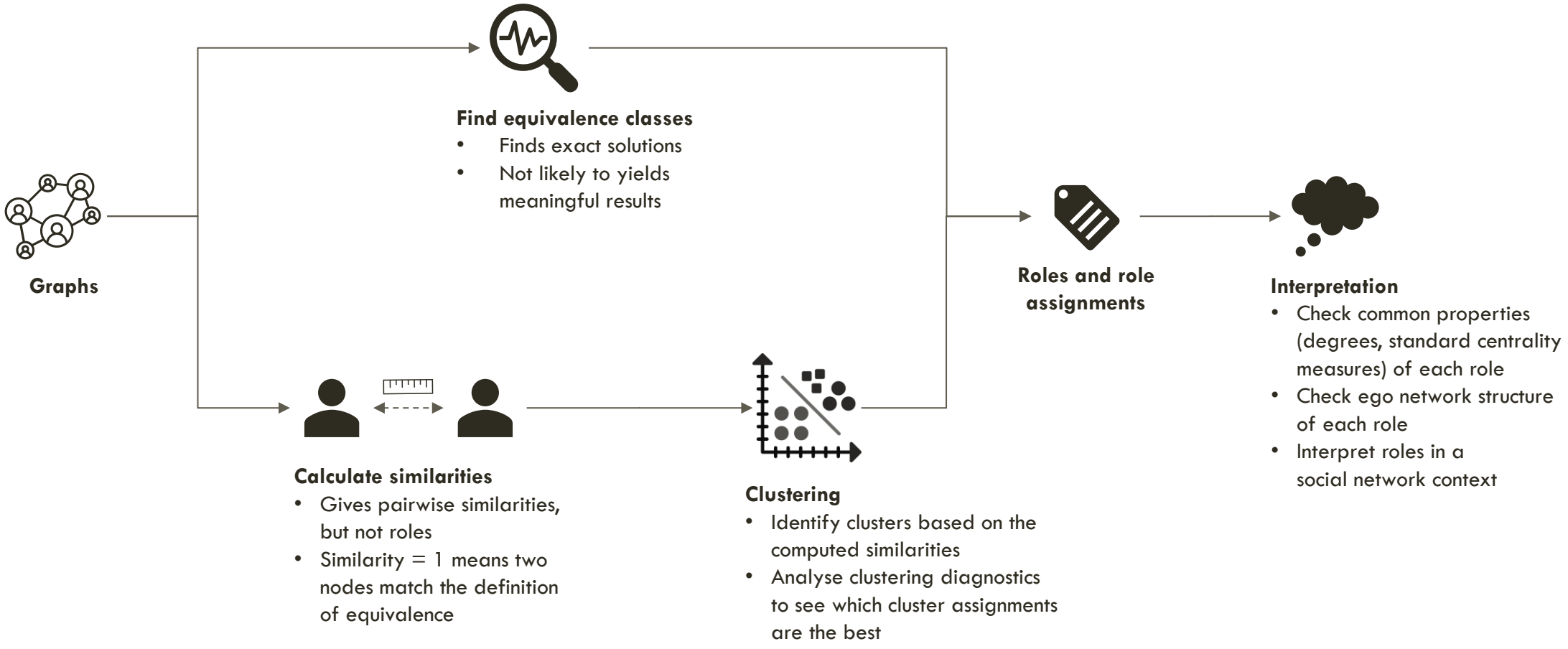
- Edge list: `df_edge_list_directed_users_postings_replies.parquet`
- Edge weights: Number of replies

Network of Users Who Follow Other Users

- Edge list: `df_edge_list_directed_users_combined_postings_replies_and_votes_to_postings_net_and_follow_connections.parquet`
- Edge weights: 1 (Unweighted)

We do not plan to use the `weight_total` column in the edge lists since it hides the nature of user interactions

Proposed approach – Role mining



Expected Results

Role Identification

✓Use equivalence clustering metrics to determine if distinct roles emerge

Role Interpretation

✓**Activity Patterns:** Voting, replying, following

✓**Ego-Network Structure:** Neighborhood composition and characteristics

Role Characteristics

✓**Degree Distributions:** In-degree, out-degree

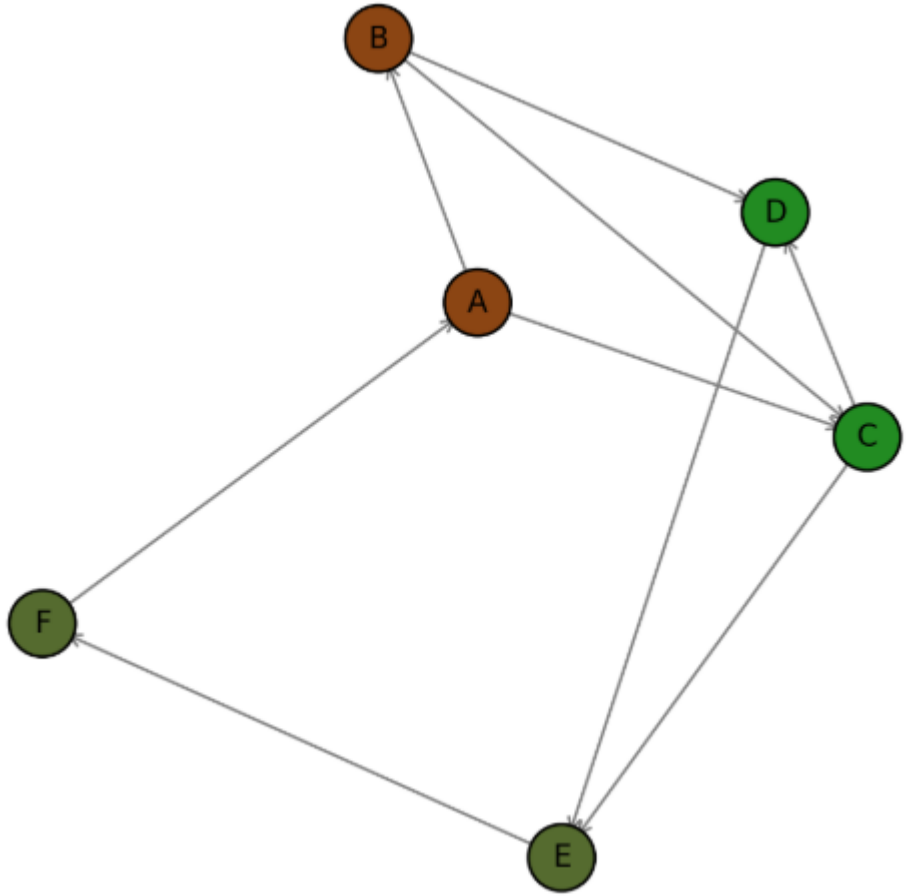
✓**Network Metrics:** Centrality, connectivity

✓**Equivalence Properties:** Structural, automorphic, regular

Common Features

✓**Neighborhood Structure:** Connections and configurations

✓**Behavioral Patterns:** Interactional or functional similarities



References

[1] Doran, D. (2017). *Network Role Mining and Analysis*. Springer International Publishing, p. 31.

[2] Doran, D. (2017). Equivalence-Based Role Mining. *Network Role Mining and Analysis*. Springer International Publishing.

[3] Jin, R., Lee, V. E., & Hong, H. (2011). Axiomatic ranking of network role similarity. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 922–930. ACM.

[4] Borgatti, S. P., & Everett, M. G. (1993). Two algorithms for computing regular equivalence. *Social Networks*, *15*(4), 361–376.