ROLES IN NETWORKS

Apr 30, 2020
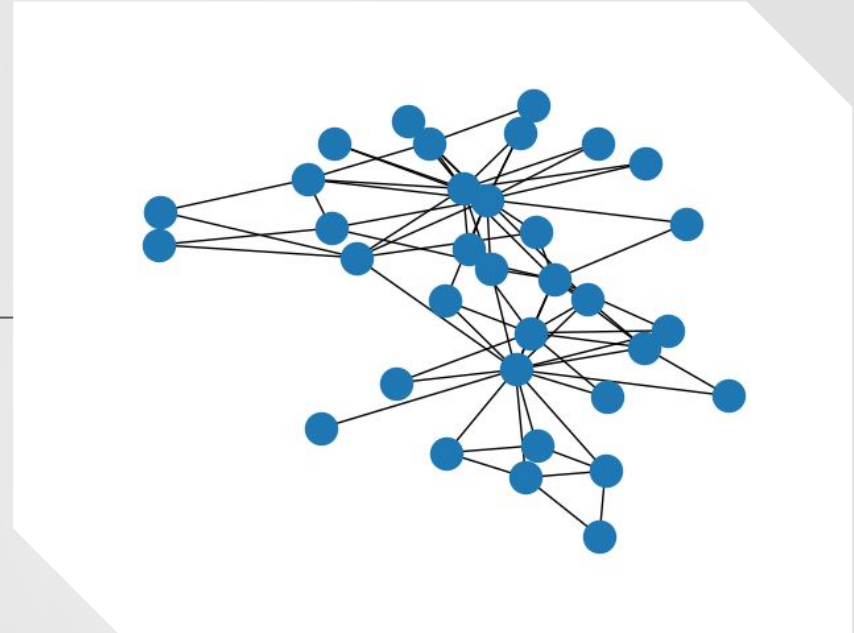
CONTENTS

1
INTRODUCTION

INTRODUCTION

NETWORKS

Networks are, by definition, sets of *nodes* with some pairs connected by *edges* [1].

INTRODUCTION

There are two sides of the same coin — *communities* and *roles*. They are different but complementary, and thus equally important. While communities capture the local structure of the network, roles are on a higher level, and they do even transfer across networks.

Both communities are roles have many applications, but some tasks are easy to resolve with communities but not with roles, and the other way around.
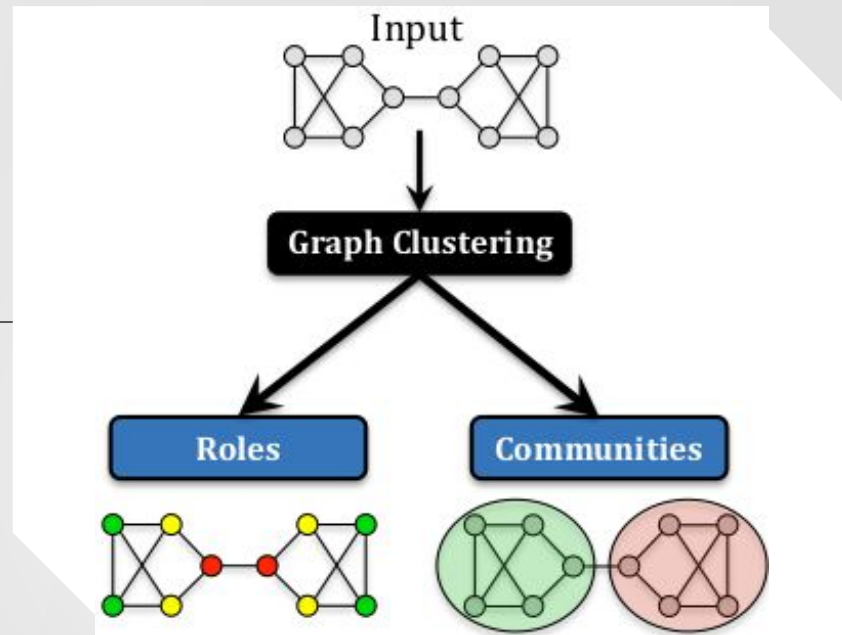
INTRODUCTION

GRAPH CLUSTERING

It is possible to group the nodes by either community or role. Rather than graph clustering, we say *community detection* in the former and *role discovery* in the latter case.

https://arxiv.org/abs/1405.7134

2
EXAMPLE

Some say that an example is worth a thousand words.

EXAMPLE

Once upon a time, in a land far away, there was a company.

It is possible to group employees by either department — community — or title — role.

Graph.
- *Nodes.* We add one node per employee.
- *Edges.* We add one edge between Employee A and Employee B if they are likely to give or take advice to or from one another, for example.
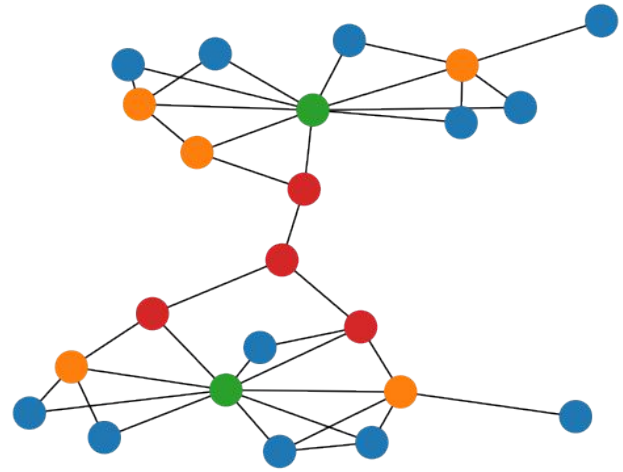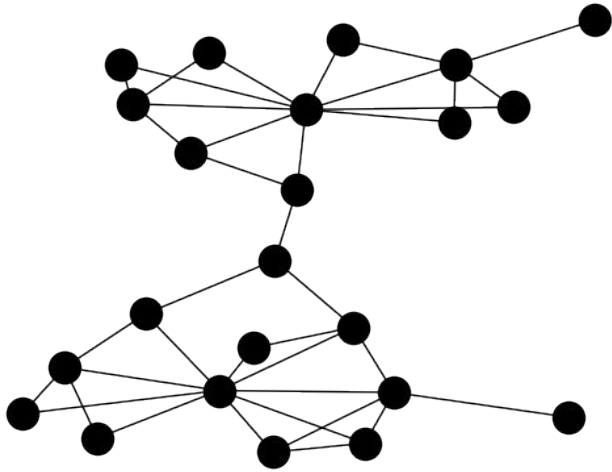
COMMUNITY DETECTION

https://youtu.be/S4QZiUPJkRI

ROLE DISCOVERY

https://youtu.be/S4QZiUPJkRI

3
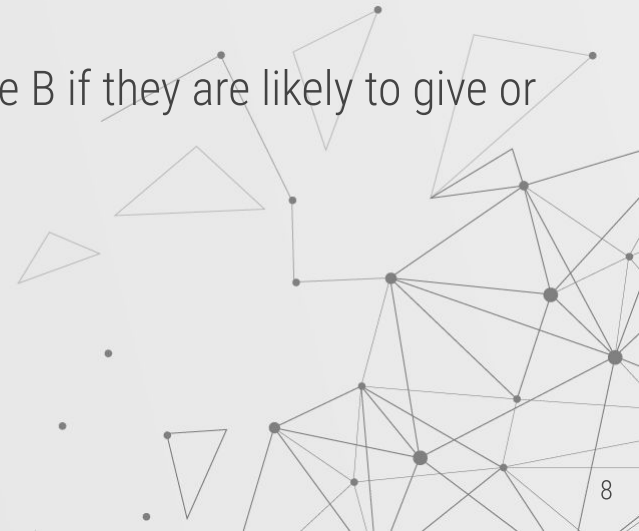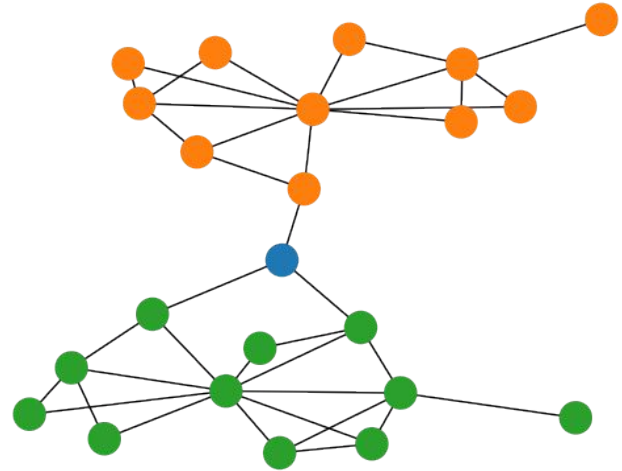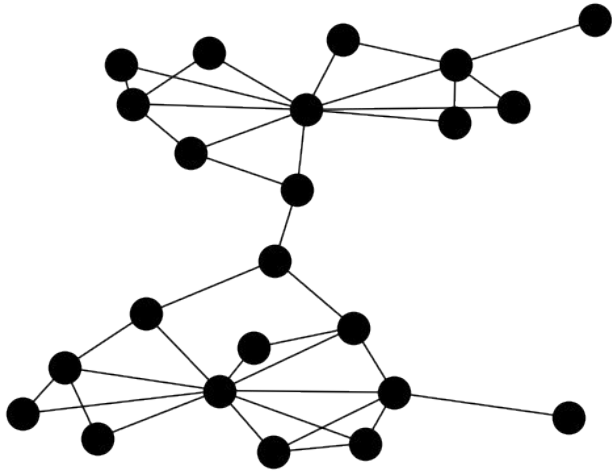
COMMUNITIES

Communities and Community Detection

COMMUNITIES

Communities are, by definition, sets of nodes that are likely to be connected to the nodes within and unlikely to be connected to the nodes without, that is, sets of nodes that are dense and separated from one another [3].
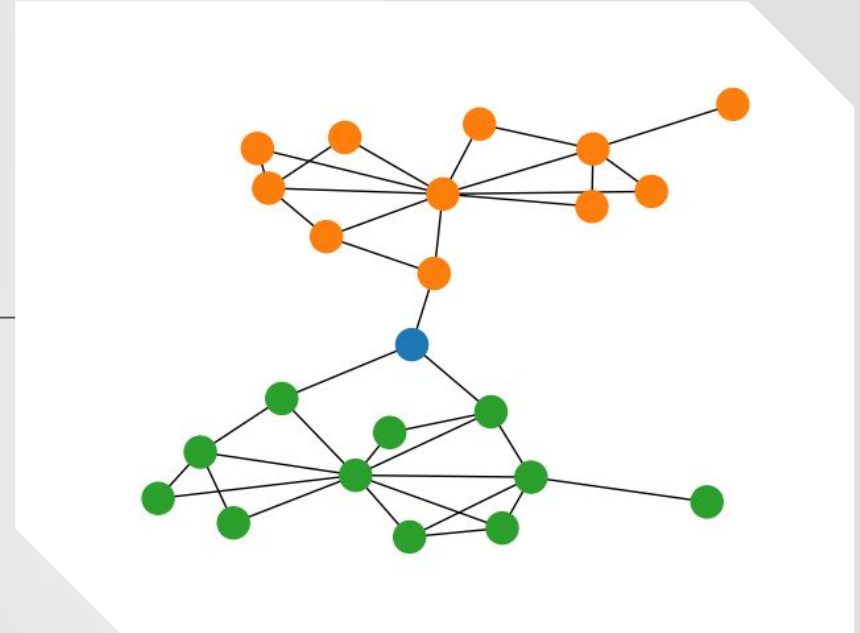
COMMUNITIES

COMMUNITY DETECTION

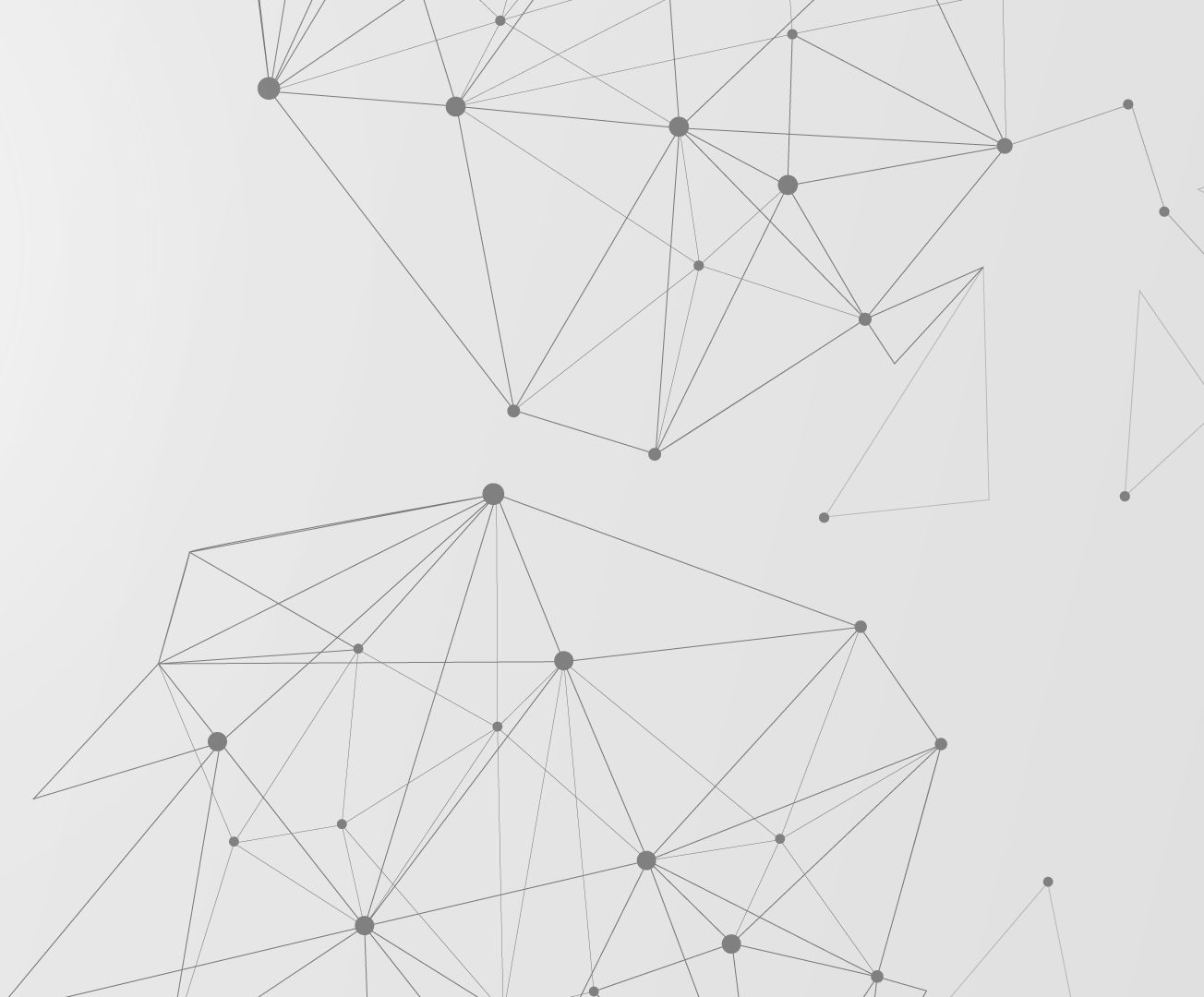Community detection is the task of dividing nodes into communities.

It is impossible not to be spoilt for choice of community detection algorithms, which maximize internal and minimize external connectivity.

https://youtu.be/S4QZiUPJkRI

4

ROLES

Roles and Role Discovery

ROLES

Roles are, by definition, sets of nodes that are structurally similar [6, 7].

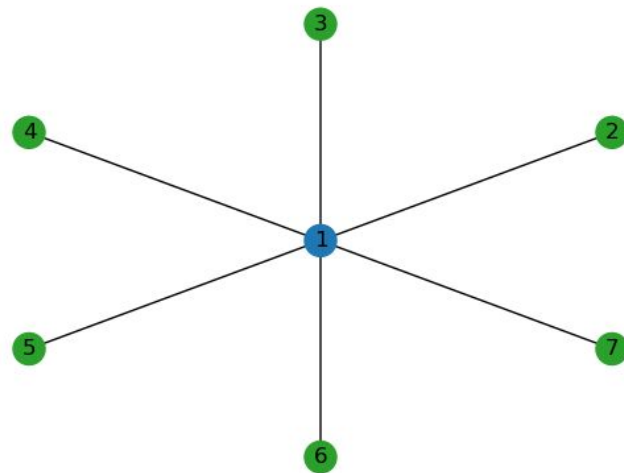They are originally from Social Network Analysis.

ROLES

STRUCTURAL EQUIVALENCE

Nodes are *structurally equivalent* if they are connected to the same neighbors [6].

Example. Nodes 2–7 are structurally equivalent.

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0103733

ROLES

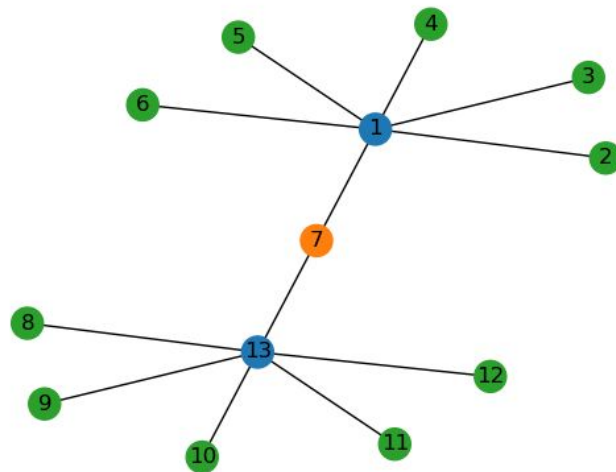## REGULAR EQUIVALENCE

Nodes are *regularly equivalent* if they are connected to role-equivalent neighbors [6].

Example.
- Nodes 1 and 13 are regularly equivalent.
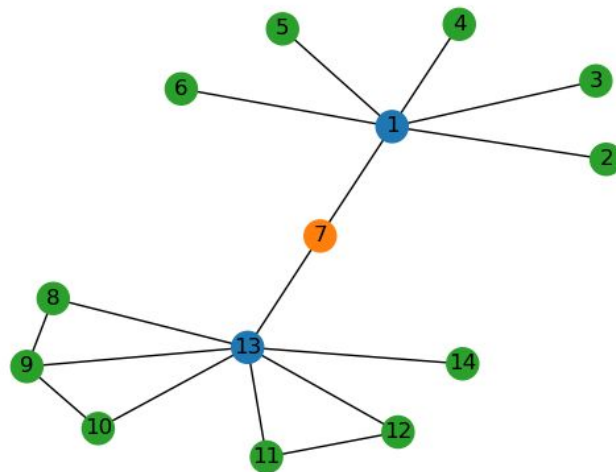- Nodes 2–6 and 8–12 are regularly equivalent.

ROLES

STRUCTURAL SIMILARITY

Nodes are *structurally similar* if they are connected to role-similar neighbors [7].

Example.
- Nodes 1 and 13 are structurally similar.
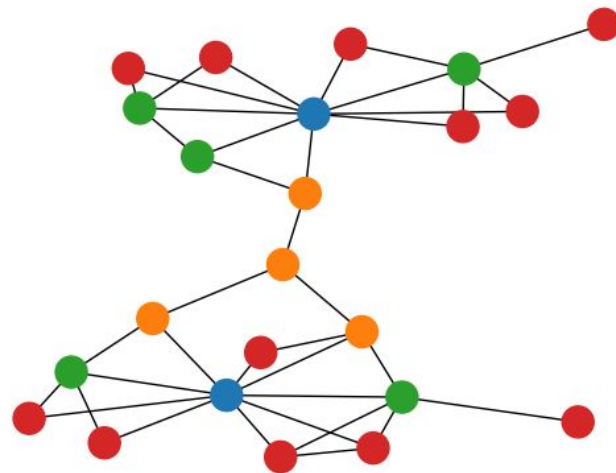- Nodes 2–6 and 8–14 are structurally similar.

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0103733

ROLES

ROLE DISCOVERY

Role discovery is the task of dividing nodes by roles.
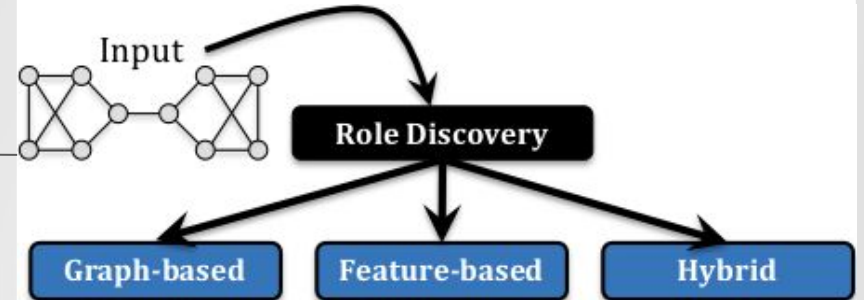
https://youtu.be/S4QZiUPJkRI

ROLE DISCOVERY

CATEGORIES

According to [6], there are three categories of role discovery — *graph-based*, *feature-based*, and *hybrid* methods. While *graph-based* methods extract roles from networks directly, *feature-based* methods extract roles from networks indirectly.


https://arxiv.org/abs/1405.7134

BLOCK MODELS

Based on permutations, block models are able to divide sets of nodes into subsets of *structurally* or *stochastically equivalent* nodes [6].

If *u* and *v* are similarly connected to the same blocks, then we will assign *u* and *v* to the same block, or role [2].

With (stochastic) block models, it is possible to summarize/compress blocks in order that relationships between roles may be highlighted [6, 7].

BLOCK MODELS
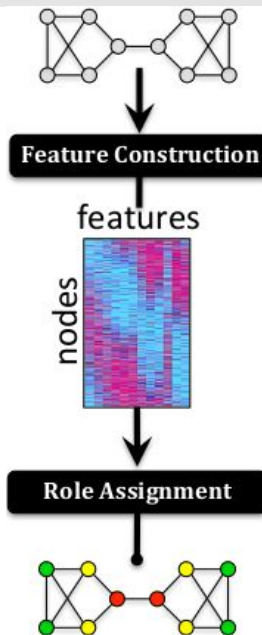

https://arxiv.org/abs/0912.5410
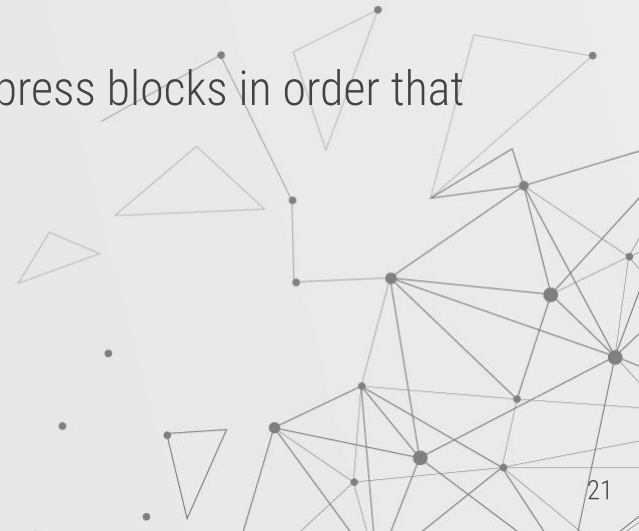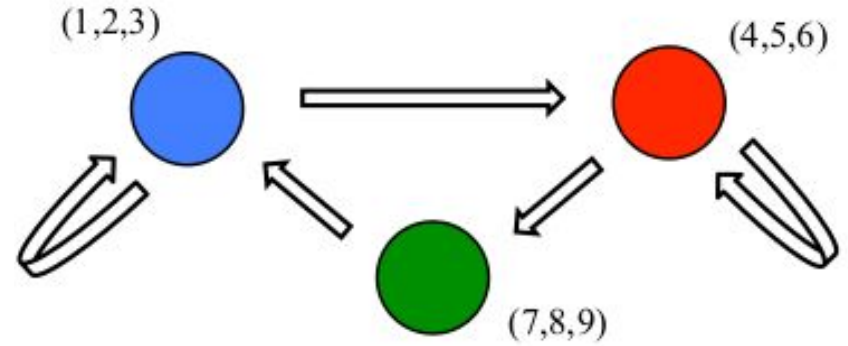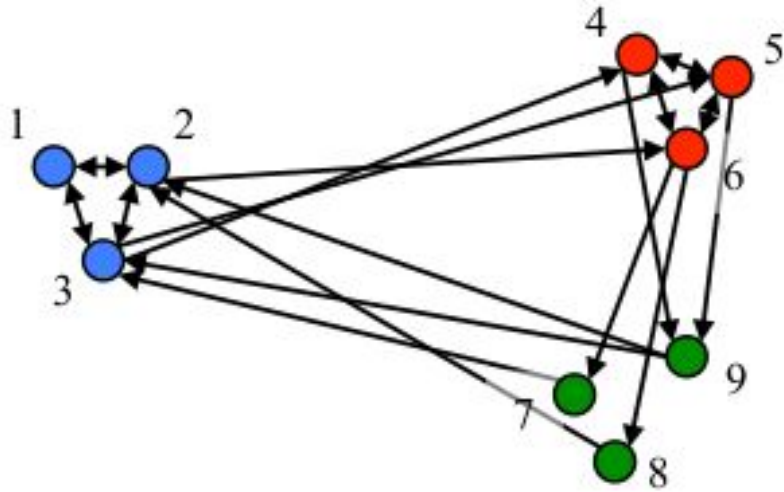
FEATURE-BASED

According to [6], there are two steps.

- *Feature Construction.* First, it is necessary to construct a nodes-features matrix.
- *Role Assignment.* Second, it is necessary to assign roles to nodes by *role clustering* or *low-rank approximation* methods.

Example. RolX [4]

https://arxiv.org/abs/1405.7134
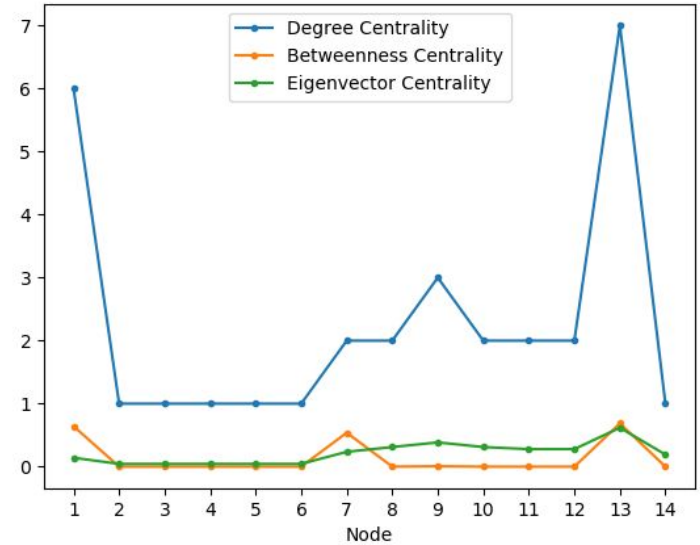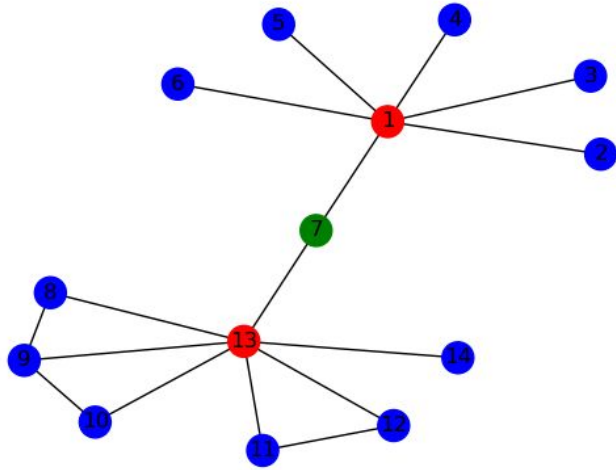
FEATURES

It is important to select the right features. Not all those (internal or external) features are *structural features*.

It is necessary to make a feature selection because there is a dependency between the quality of the features and the quality of the roles.

FEATURES

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0103733

5

APPLICATIONS

APPLICATIONS

NODE CLASSIFICATION

LINK PREDICTION

ANOMALY DETECTION

SUMMARIZATION

COMPRESSION

VISUALIZATION

○ Icons made by Freepik from www.flaticon.com

NODE CLASSIFICATION

- *Strong Homophily.* Communities are better for node classification if it is likely that the nodes within the clusters are similar to one another.
- *Weak Homophily.* If the homophily is weak, then we will prefer roles.

It is worth remembering that *roles generalize across networks* [7].

LINK PREDICTION

According to [7], both communities and roles are good for link prediction.

- *Short Range.* If the hypothesis that there is a missing link between nodes that do share a common neighbor (or more) is true, then we will prefer communities.
- *Long Range.* If the hypothesis that there is a missing link between nodes that do not share a common neighbor is true, then we will prefer roles.
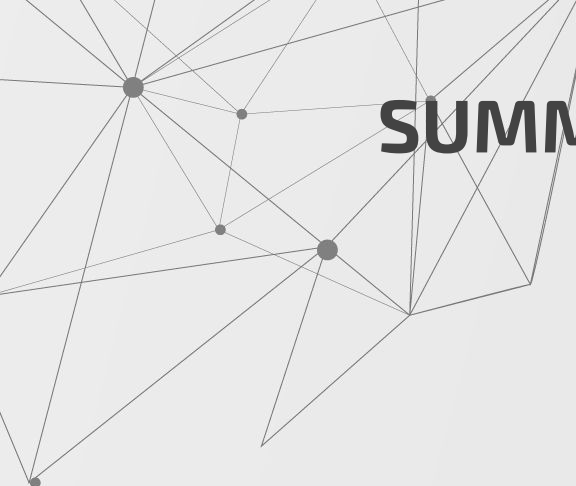
ANOMALY DETECTION
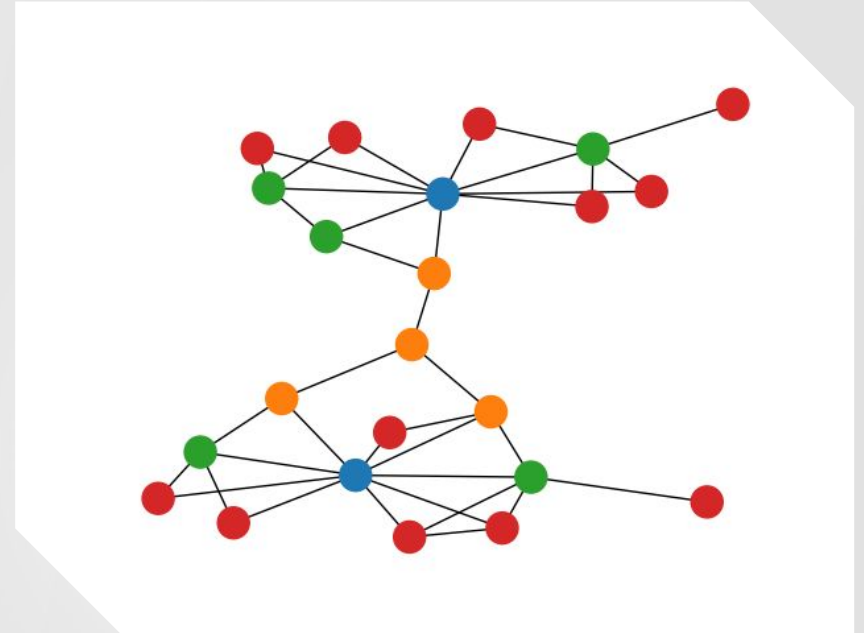
Roles are better for anomaly detection [7].

- *Static.* Anomalies are nodes whose role is anomalous.
- *Dynamic.* Anomalies are nodes that move from role to role in an anomalous manner.
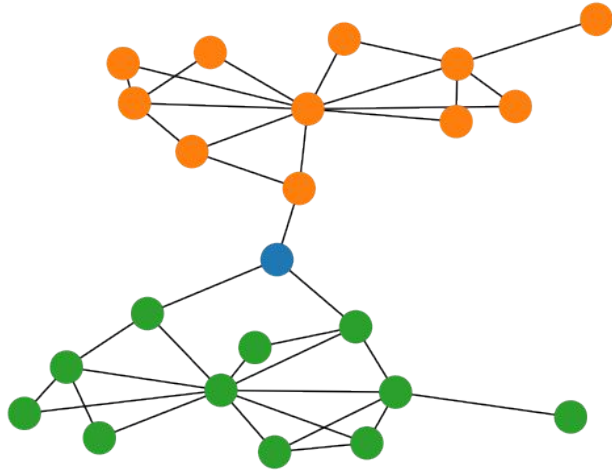
SUMMARIZATION/COMPRESSION

Both communities and roles are able to highlight relationships between sets of nodes.
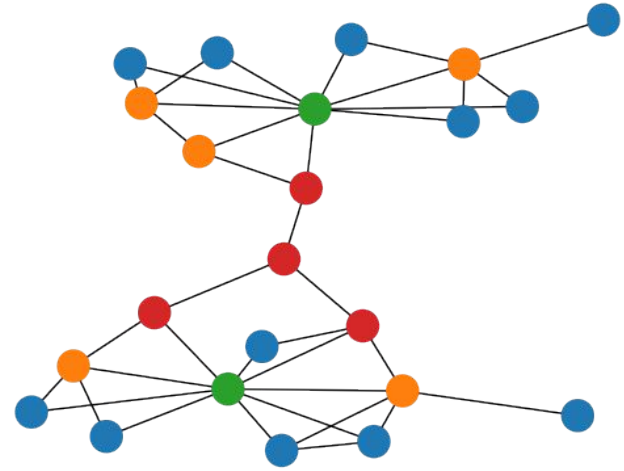
https://youtu.be/S4QZiUPJkRI

VISUALIZATION

Communities

Roles

https://youtu.be/S4QZiUPJkRI

6
CONCLUSION

CONCLUSION

Both *communities* and *roles* are helpful. It is important not to mix communities and roles, since they are not interchangeable.

While *community detection* has received keen interest, *role discovery* is still in an early stage, despite its many applications, such as network alignment and data anonymization.

REFERENCES

[1] David Easley and Jon Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World.* 2010.

[2] Anna Goldenberg et al. *A Survey of Statistical Network Models.* 2009. arXiv: 0912.5410 [stat.ME].

[3] Santo Fortunato. *Community Detection in Graphs.* 2010. Physics Reports.

[4] Keith Henderson et al. *RolX: Structural Role Extraction & Mining in Large Graphs.* Proceedings of the 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2012.

[5] Shaobin Huang et al. *Identifying Node Role in Social Network Based on Multiple Indicators.* PLoS ONE. 2014.

[6] Ryan A. Rossi and Nesreen K. Ahmed. *Role Discovery in Networks.* IEEE Transactions on Knowledge and Data Engineering (TKDE). 2015.

[7] Ryan A. Rossi et al. *From Community to Role-based Graph Embeddings.* 2019. arXiv: 1908.08572 [cs.SI].

QUESTIONS?

Susanna Pozzoli
spozzoli@kth.se