Integer Programming Models for Optimization-Based Clustering and Analysis of Networks

Samin Aref (University of Toronto)

Joint work with Mark Wilson, Andrew Mason, Zachary Neal, Ly Dinh, Shadi Rezapour, Jana Diesner, Mahdi Mostajabdave, Hriday Chheda, and Boris Ng

Frameworks, research and applications in complex Networks with signed edges
Satellite at NetSci'25, Maastricht, 2025-06-03

Mechanical & Industrial Engineering
UNIVERSITY OF TORONTO

MAX-PLANCK-INSTITUT FÜR DEMOGRAFISCHE FORSCHUNG
MAX PLANCK INSTITUTE FOR DEMOGRAPHIC RESEARCH

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

Outline

Background

1. Balance in signed networks

Methodology

2. How can we measure balance?

Use Case

3. Hidden coalitions in the US Congress

Optimization of unsigned networks

4. Network clustering (Bayan algorithm and Troika algorithm)

UNIVERSITY OF TORONTO

Part 1

Background

1. Balance in signed networks

Methodology

2. How can we measure balance?

Use Case

3. Hidden coalitions in the US Congress

Optimization of unsigned networks

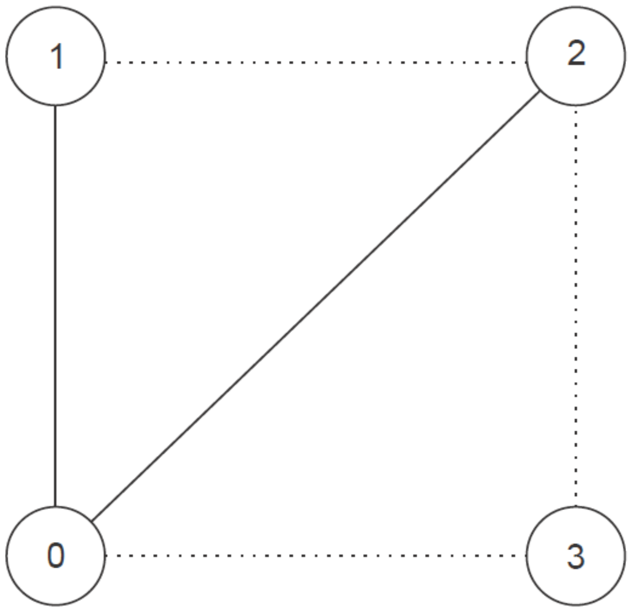
4. Network clustering (Bayan algorithm and Troika algorithm)

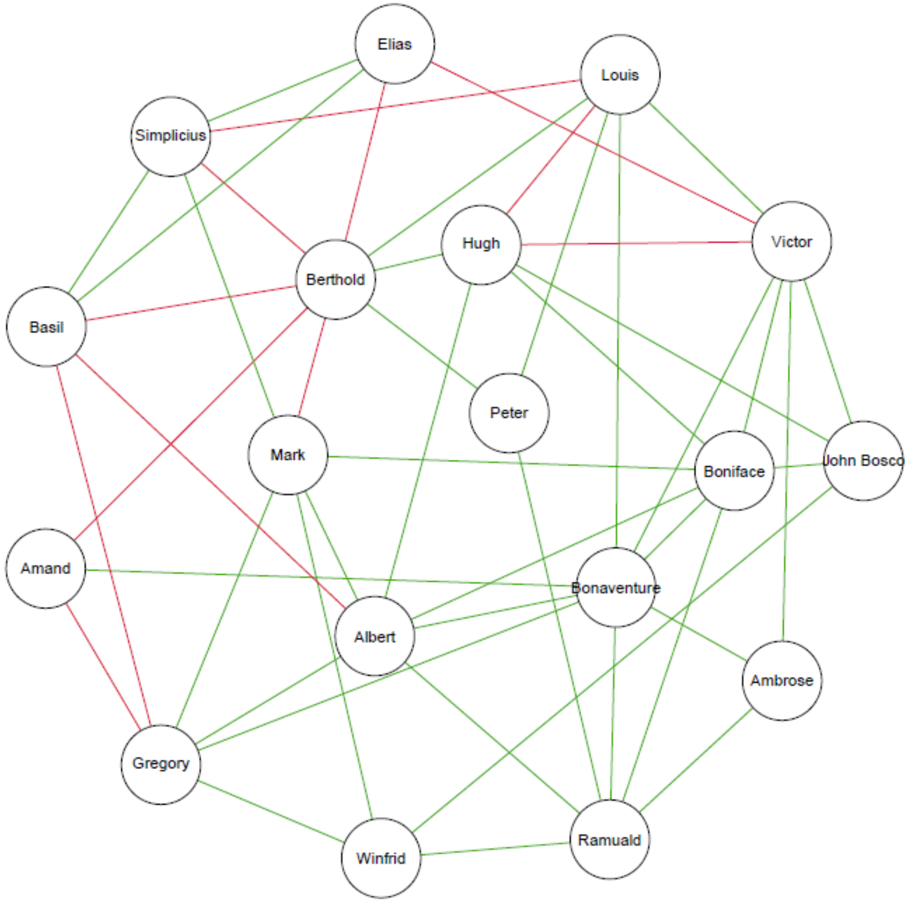
Signed network (signed graph)

Two types of edges:

-positive edges (solid line or green colour)

-negative edges (dotted line or red colour)





UNIVERSITY OF TORONTO

Theory of Structural Balance

(Heider 1944) (Cartwright and Harary 1956)

In *balanced* signed networks:

Enemy of an enemy = friend
Friend of a friend = friend
Enemy of a friend = enemy
Friend of an enemy = enemy

UNIVERSITY OF TORONTO

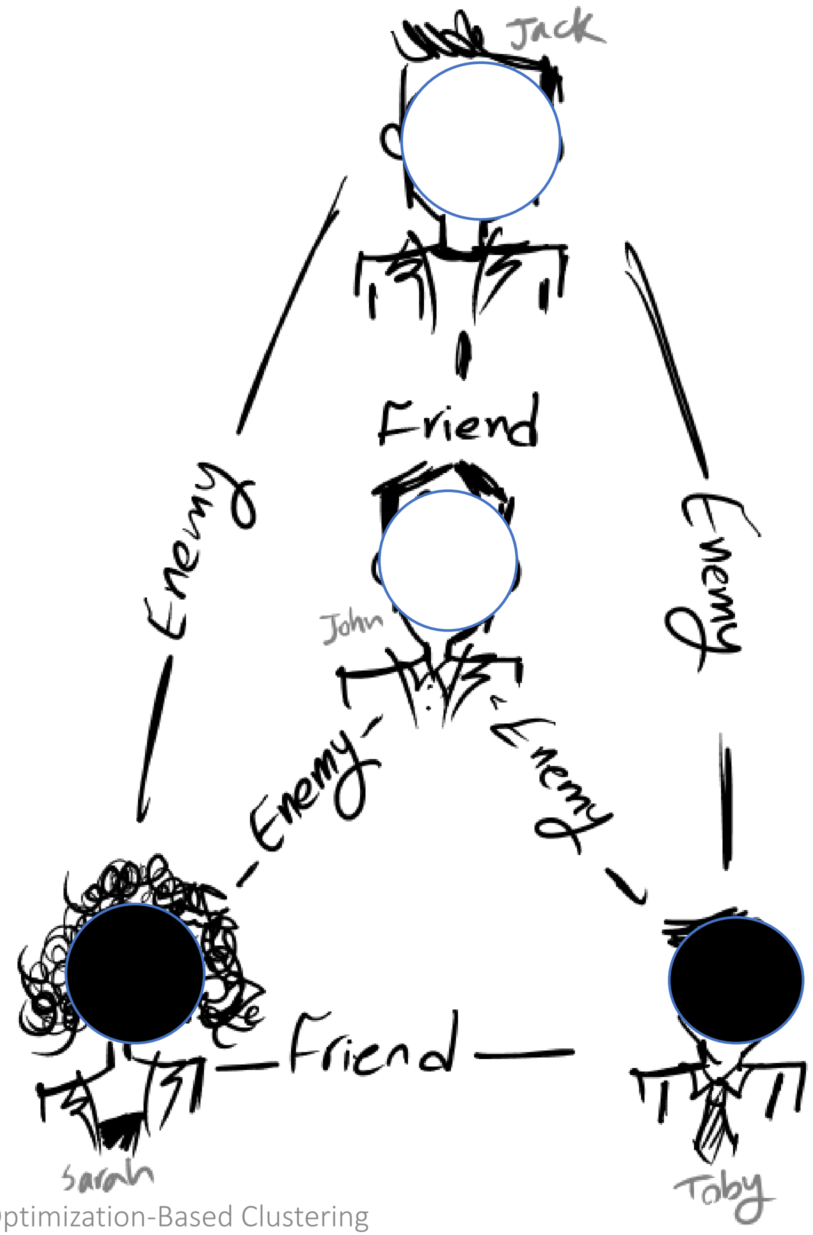
Theory of Structural Balance

(Heider 1944) (Cartwright and Harary 1956)

In *balanced* signed networks:

Enemy of an enemy = friend ✓
Friend of a friend = friend ✓
Enemy of a friend = enemy ✓
Friend of an enemy = enemy ✓

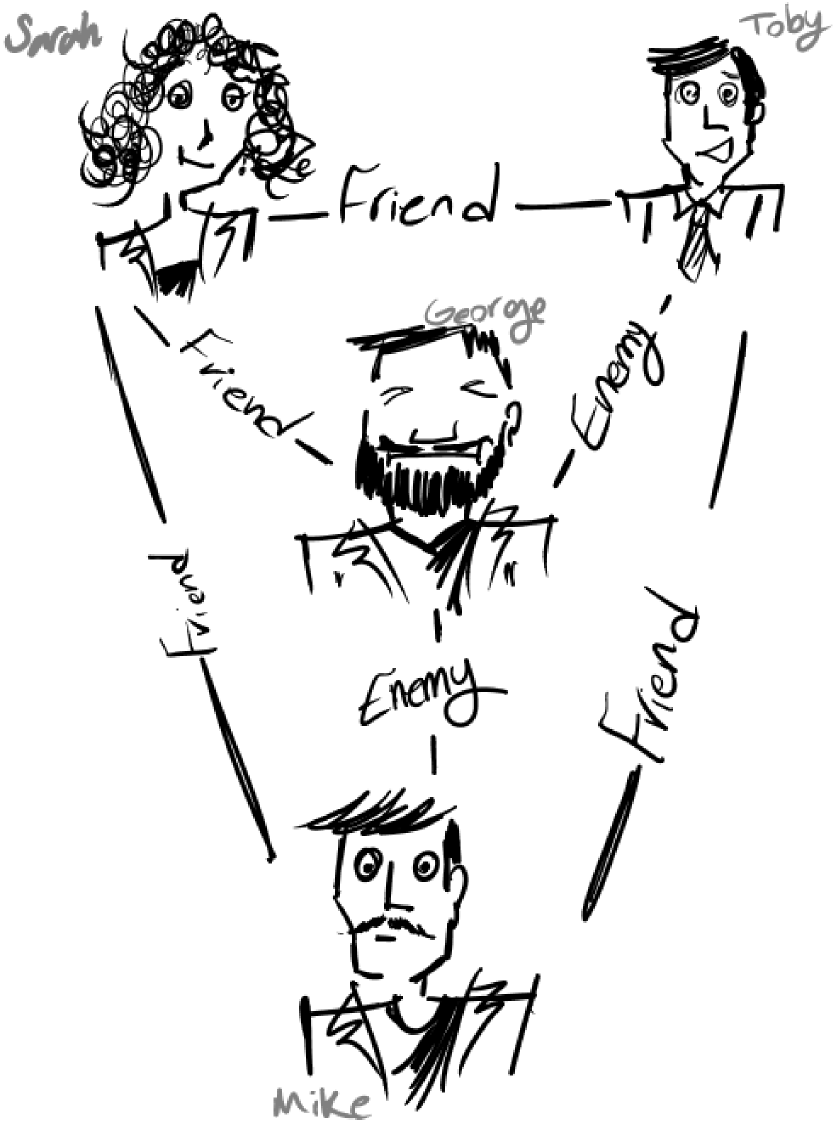
Balanced



UNIVERSITY OF TORONTO

Unbalanced

Enemy of an enemy = friend ✓
Friend of a friend = friend ✗
Enemy of a friend = enemy ✗
Friend of an enemy = enemy ✗



UNIVERSITY OF TORONTO

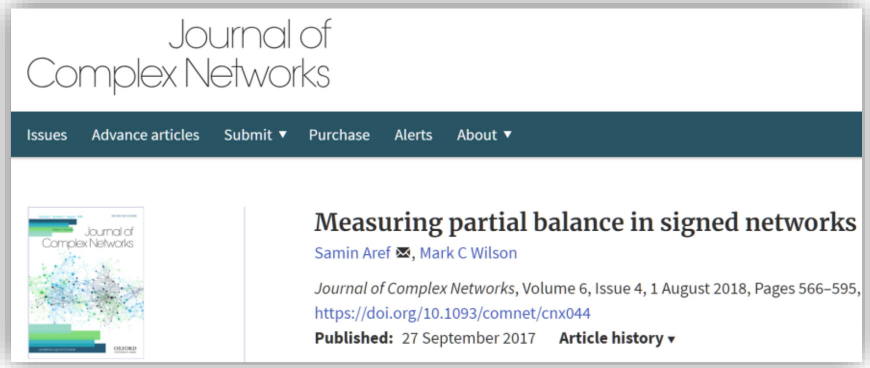
Balanced subgraph

It is 1 edge away from balance

Frustration index of a signed graph:

Minimum number of edges whose removal makes the resulting subgraph balanced.

- The *frustration index* is 1


Journal of Complex Networks

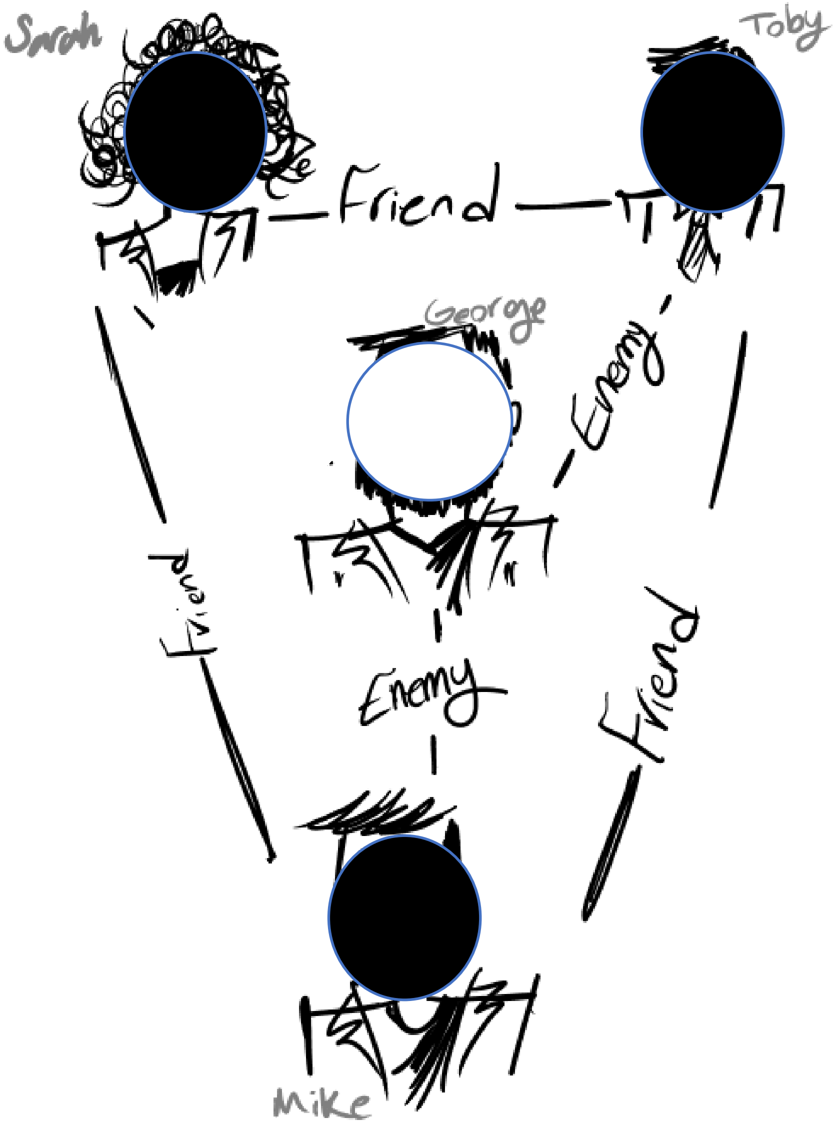
Issues | Advance articles | Submit ▾ | Purchase | Alerts | About ▾

Measuring partial balance in signed networks

Samin Aref ✉, Mark C Wilson

Journal of Complex Networks, Volume 6, Issue 4, 1 August 2018, Pages 566–595, https://doi.org/10.1093/comnet/cnx044

Published: 27 September 2017 **Article history** ▾



UNIVERSITY OF TORONTO

Part 2

Background

1. Balance in signed networks

Methodology

2. How can we measure balance?

Use Case

3. Hidden coalitions in the US Congress

Optimization of unsigned networks

4. Network clustering (Bayan algorithm and Troika algorithm)

Undirected signed graph $G = (V, E, \sigma)$

$$\sigma : E \rightarrow \{-1, +1\}$$

$$|V| = n$$

$$|E| = m = m^+ + m^-$$

$$|E^+| = m^+$$

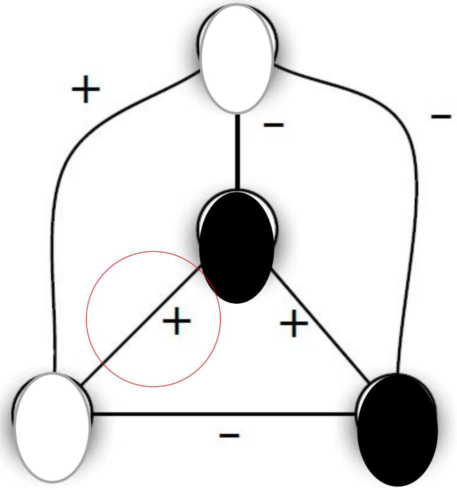
$$|E^-| = m^-$$

Defining frustration based on the clusters of nodes

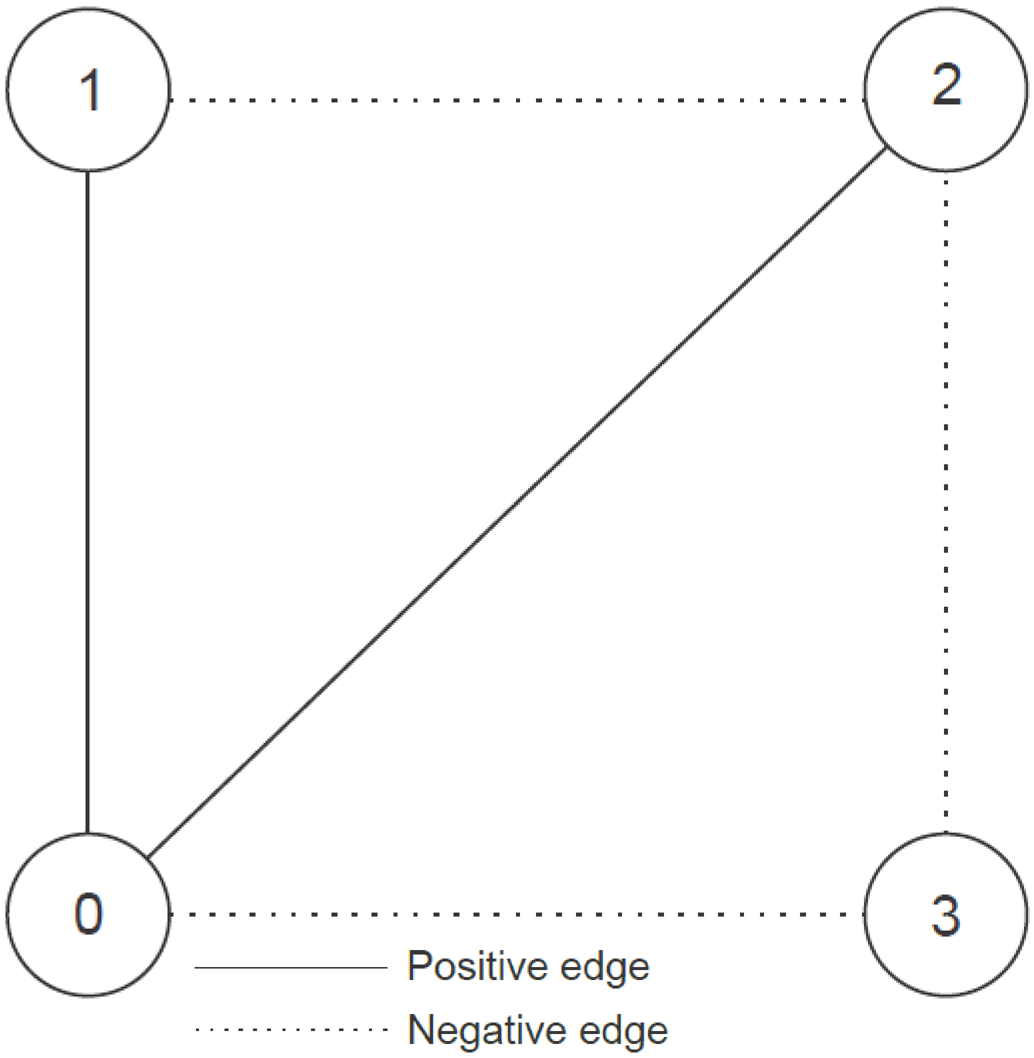
Frustrated edge:

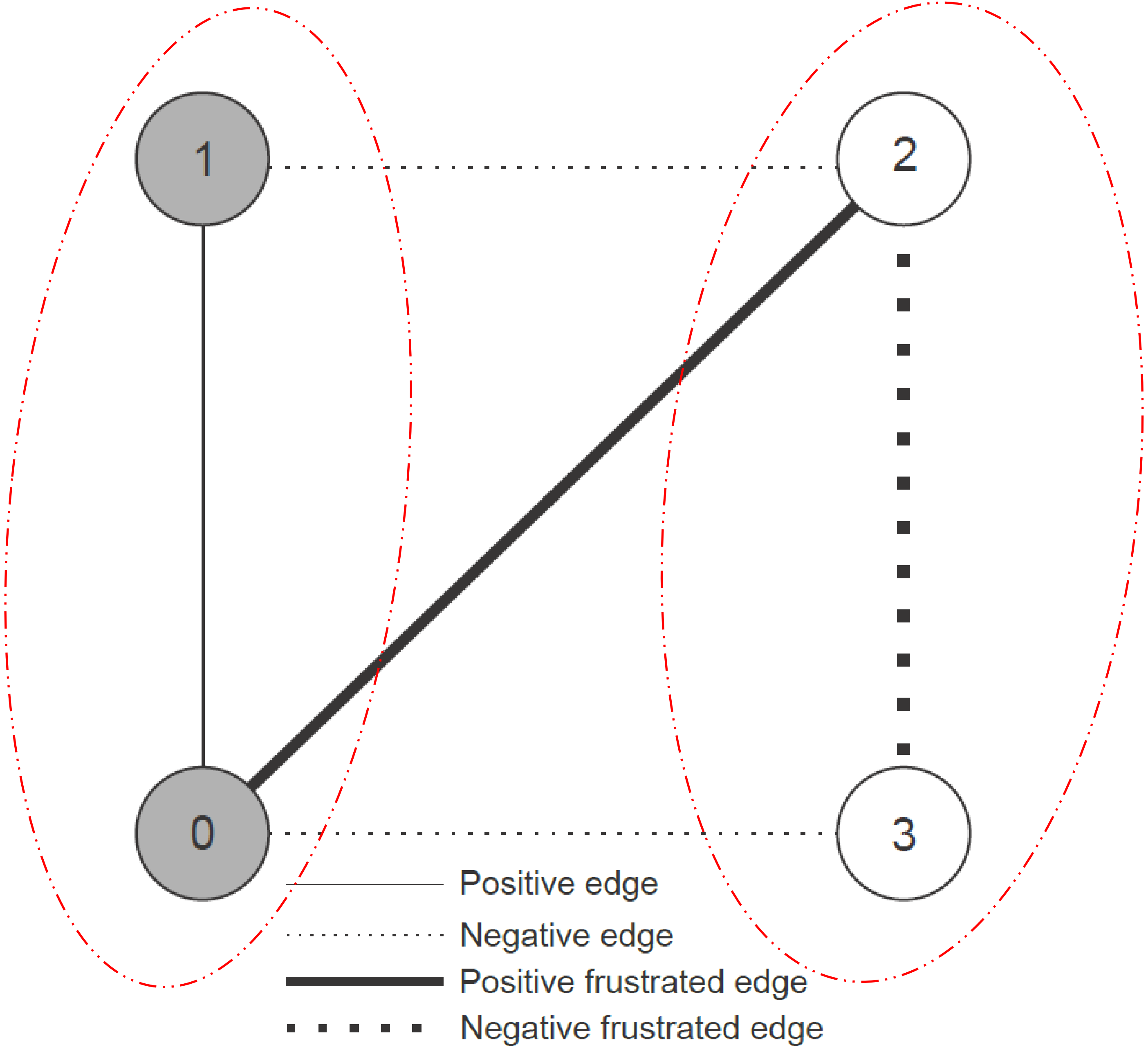
-positive edge with endpoints in different clusters

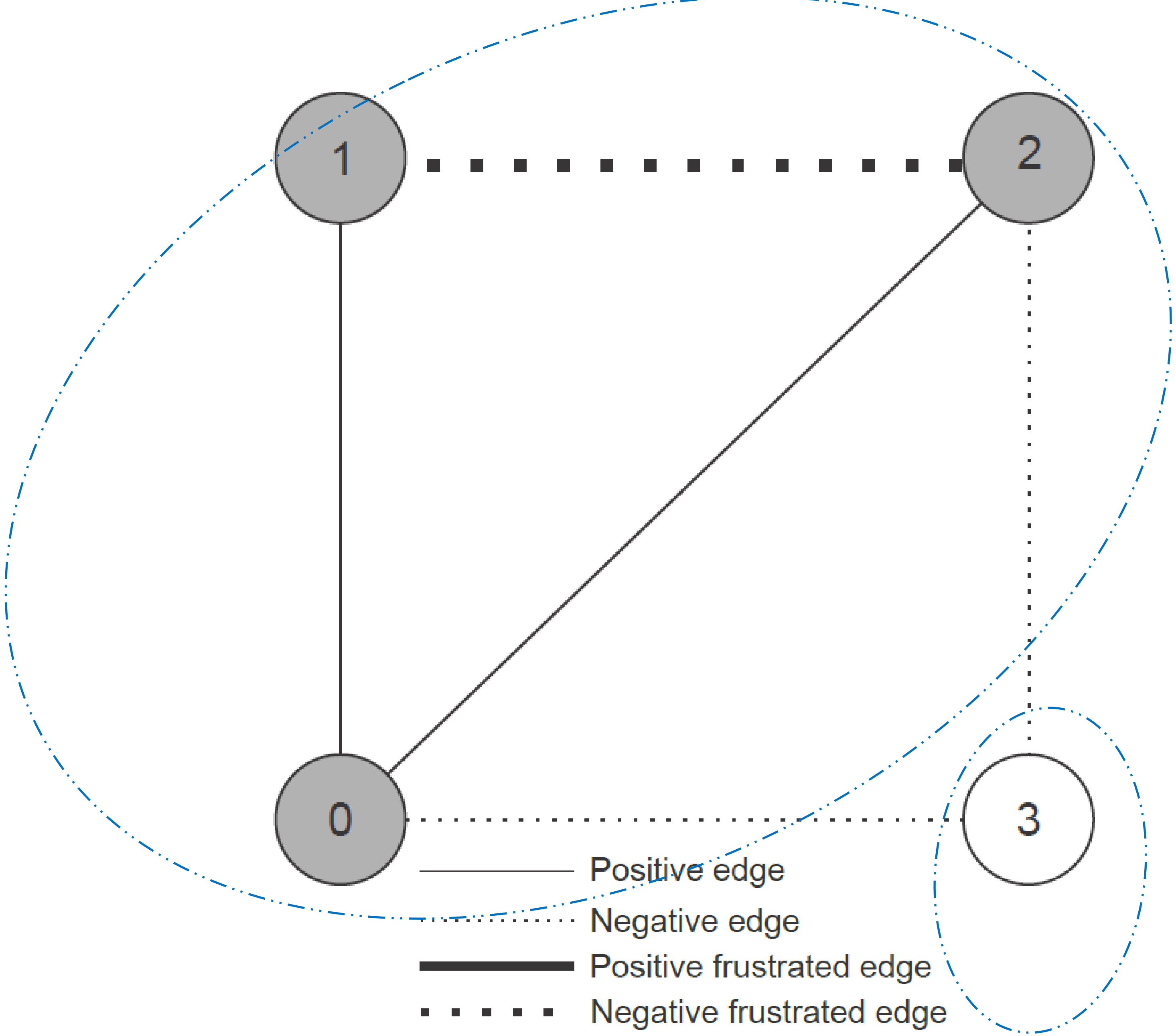
-negative edge with endpoints in the same cluster



UNIVERSITY OF TORONTO


1
2
0
3
Positive edge
Negative edge


1
2
0
3
Positive edge
Negative edge
Positive frustrated edge
Negative frustrated edge

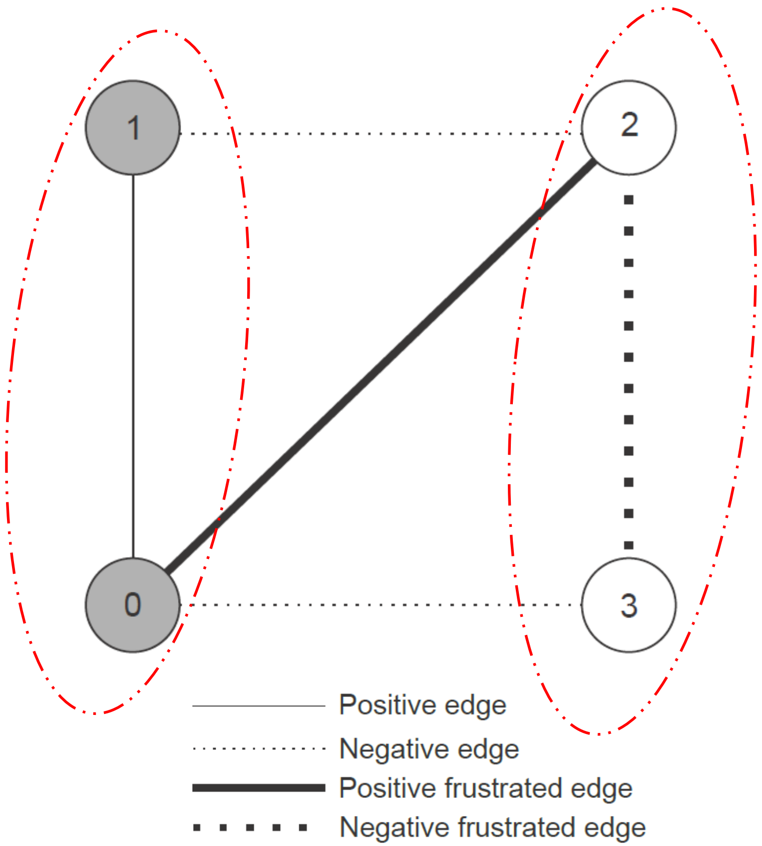

1
2
0
3
Positive edge
Negative edge
Positive frustrated edge
Negative frustrated edge

Frustration state of edge (i,j) under partition P=(X, V\X)

x_i : binary variable representing cluster of node i

$$x_i = 1 \text{ if } i \in X$$
$$x_i = 0 \text{ if } i \in V \setminus X$$

$$f_{ij}(X) = \begin{cases} 0, & \text{if } x_i = x_j \text{ and } (i,j) \in E^+ \\ 1, & \text{if } x_i = x_j \text{ and } (i,j) \in E^- \\ 0, & \text{if } x_i \neq x_j \text{ and } (i,j) \in E^- \\ 1, & \text{if } x_i \neq x_j \text{ and } (i,j) \in E^+ \end{cases}$$



Frustration count (of a graph G under partition P=(X, V\X))

$$f_G(X) := \sum_{(i,j)\in E} f_{ij}(X)$$

Frustration index (of graph G)

$$L(G) = \min_{X \subseteq V} f_G(X)$$

Minimizing the number of frustrated edges

x_i : binary variable representing cluster of node i

f_{ij} : binary variable representing the frustration of the signed edge (i, j) based on endpoint clusters

$$f_{ij} = x_i + x_j - 2x_i x_j \;\forall (i,j) \in E^+$$

$$f_{ij} = 1 - (x_i + x_j - 2x_i x_j) \;\forall (i,j) \in E^-$$

Model UBQP

$$\min_{x_i : i \in V} Z = \sum_{(i,j) \in E^+} x_i + x_j - 2x_i x_j + \sum_{(i,j) \in E^-} 1 - (x_i + x_j - 2x_i x_j)$$

$$\text{s.t.} \quad x_i \in \{0,1\} \quad \forall i \in V$$

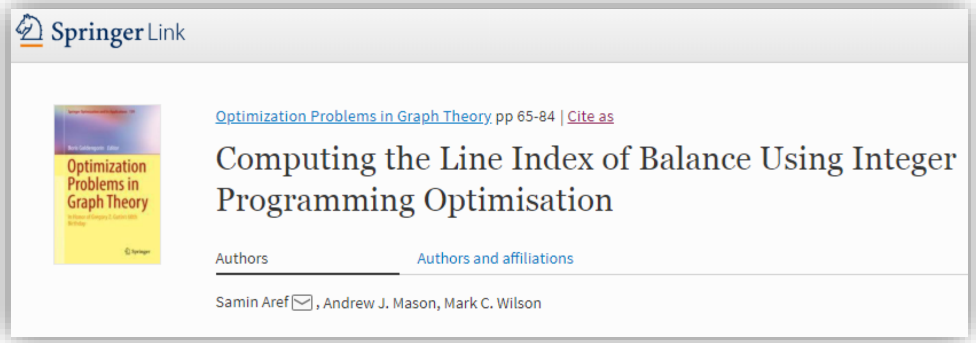
UNIVERSITY OF TORONTO

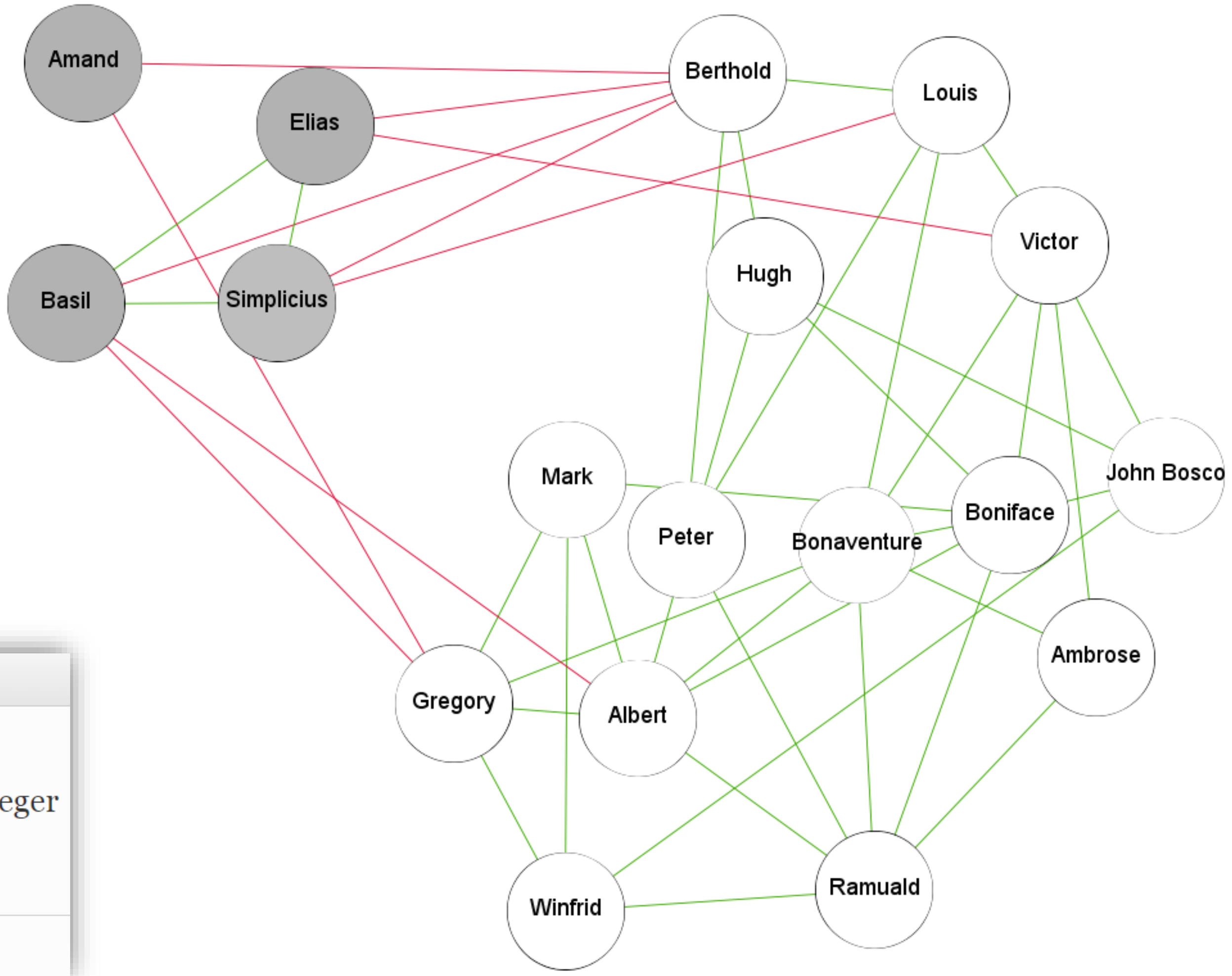
Sampson’s monastery interactions

Social context

Nodes: people (monks)
Edges: top choices for like and dislike

Negative edges are shown in red

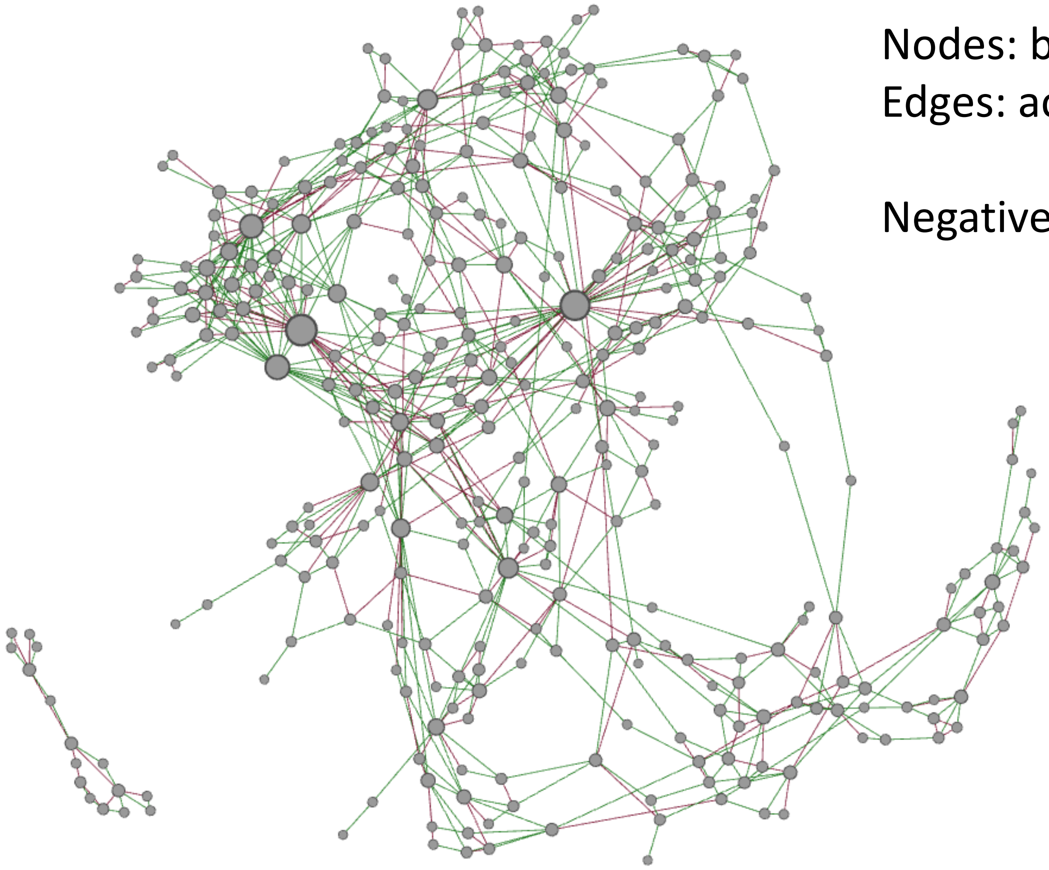




UNIVERSITY OF TORONTO

Signed network of a protein (EGFR)

Biological context



Nodes: biological molecules
Edges: activation or inhibition relations

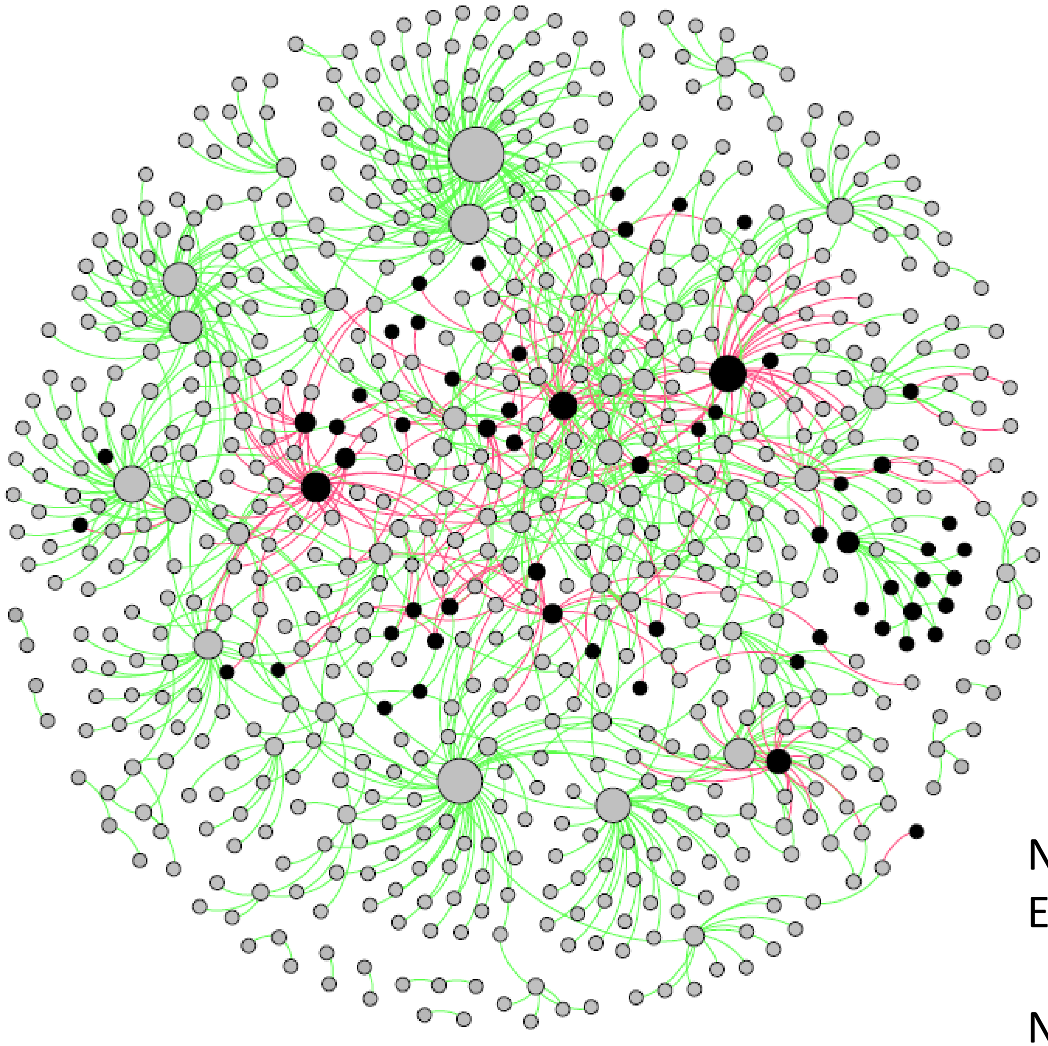
Negative edges are shown in red

n	m	m^-
329	779	264

UNIVERSITY OF TORONTO

Baker's yeast

Biological context



n	m	m^-
690	1080	220

NETWORKS
An International Journal

RESEARCH ARTICLE | Full Access

A modeling and computational study of the frustration index in signed networks

Samin Aref, Andrew J. Mason, Mark C. Wilson

First published: 03 October 2019 | https://doi.org/10.1002/net.21907

Nodes: biological molecules
Edges: activation or inhibition relations

Negative edges are shown in red

UNIVERSITY OF TORONTO

(-1) x (-1) = (+1)

Transitivity of signed ties in networks

Balance is not restricted to social and biological networks though.

We observe it in certain financial networks, networks of international relations, some molecular networks, and in polarized political networks.


Journal of Complex Networks

Issues | Advance articles | Submit ▾ | Purchase | Alerts | About ▾

Article Contents

Abstract

1. Introduction

Balance and frustration in signed networks

Samin Aref, Mark C Wilson

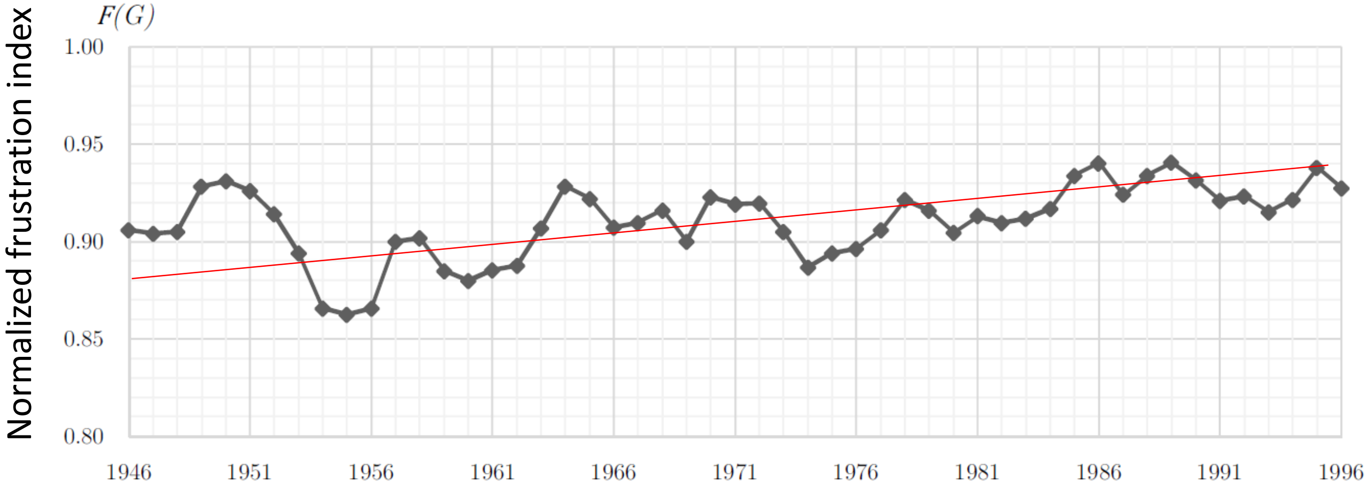
Journal of Complex Networks, cny015, https://doi.org/10.1093/comnet/cny015

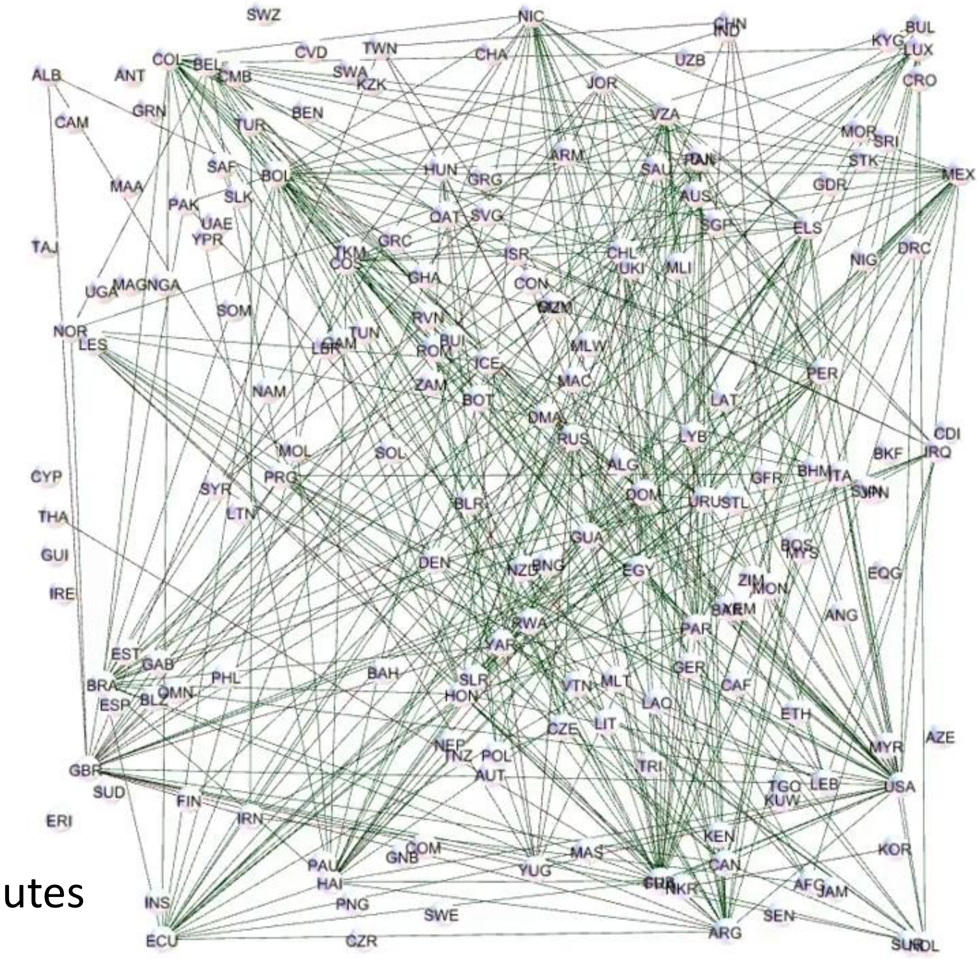
Published: 14 August 2018 **Article history** ▾

International relations

Signed international relations
Correlates of War dataset (1946-1996)

YouTube link: TinyURL.com/CoWDynamics





Nodes: Countries

Positive edges: Formal military alliances

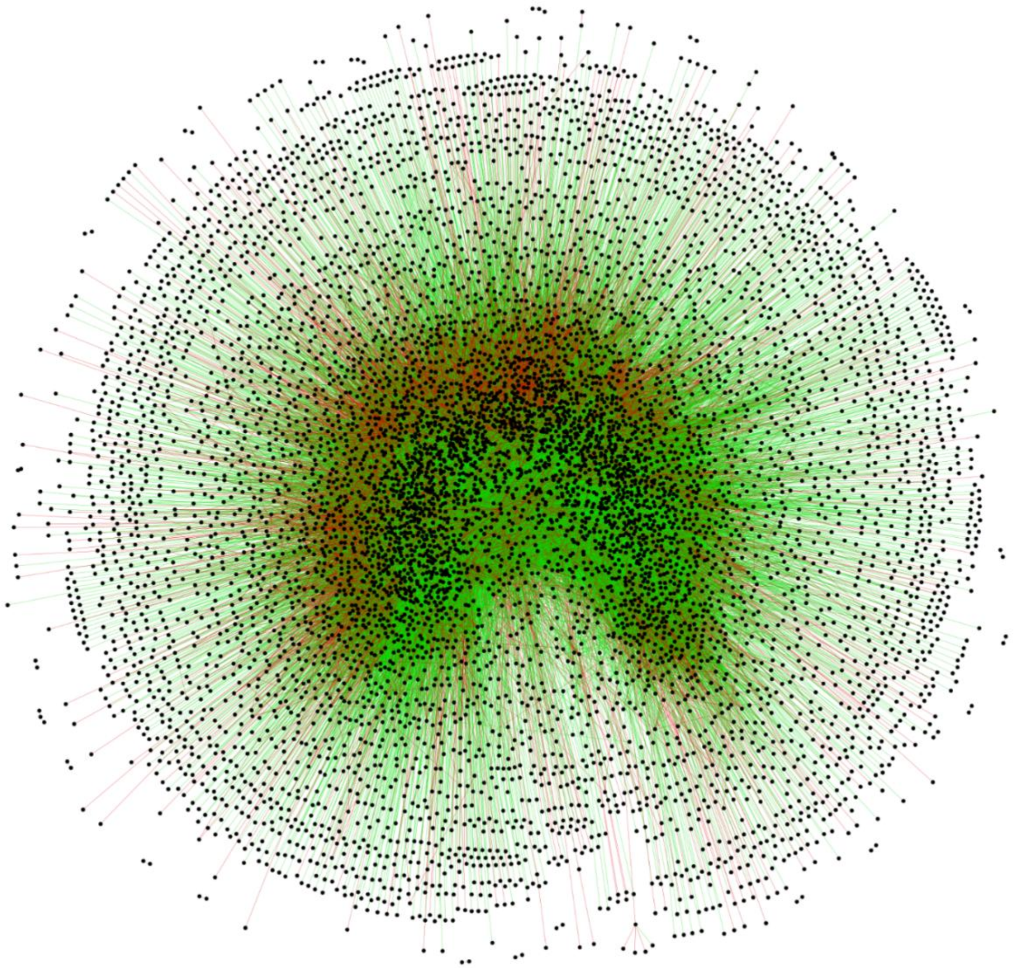
Negative edges: Military conflicts, ideological conflicts, or border disputes

UNIVERSITY OF TORONTO

Extends to directed signed networks

Social context

n	m	m^-
7112	99917	21837




scientific reports

Explore our content ⌄ Journal information ⌄

nature › scientific reports › articles › article

Article | Open Access | Published: 17 September 2020

Multilevel structural evaluation of signed directed social networks based on balance theory

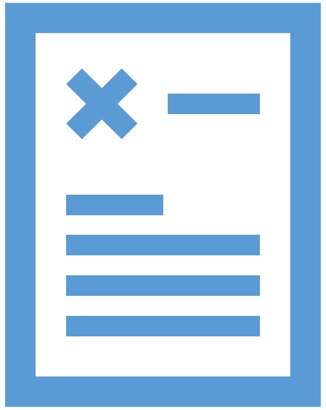
Samin Aref, Ly Dinh, Rezvaneh Rezapour & Jana Diesner

Wikipedia Elections
Nodes: Wikipedia authors
Edges: approval or disapproval for promotion to administrator

UNIVERSITY OF TORONTO

Part 3

Networks of US Congress legislators
20 signed networks: one for each session (2-year period) of the Congress



Bill co-sponsorship data

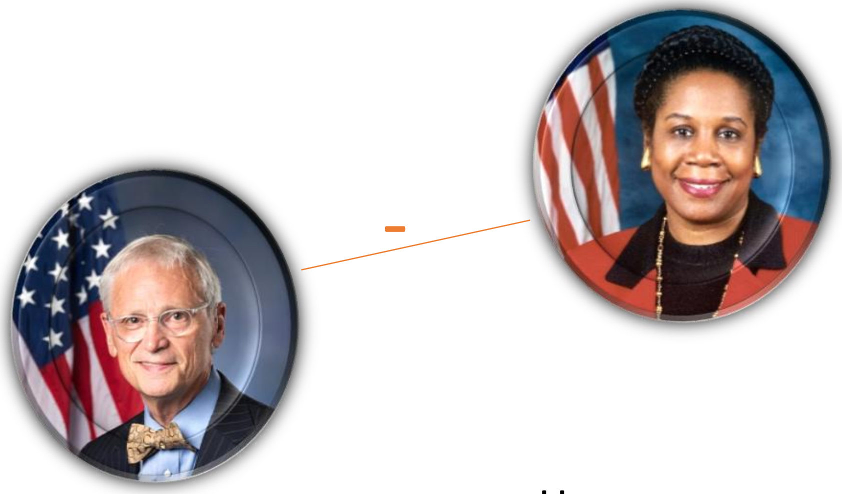
who has co-sponsored which bill

20 sessions: 1979-2017

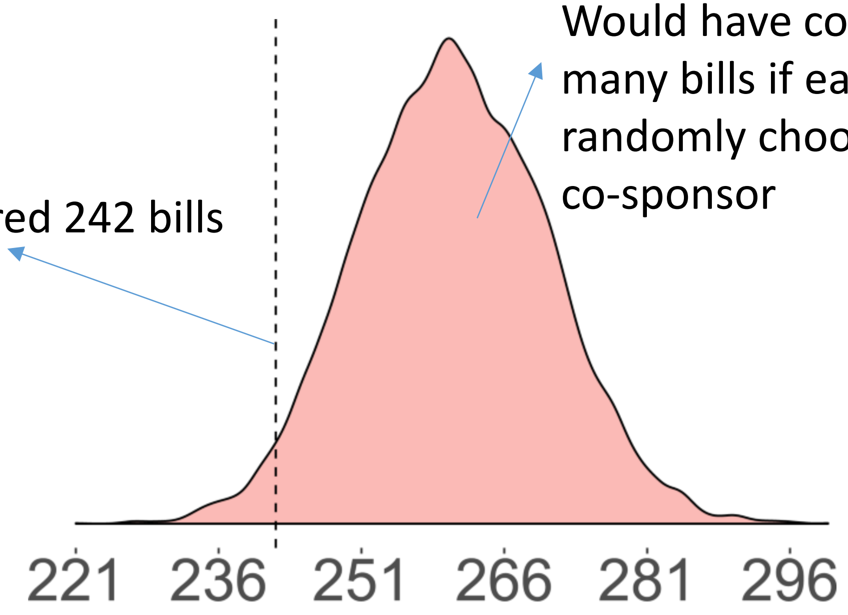
~435 US Representatives ✕ ~6000 bills

UNIVERSITY OF TORONTO

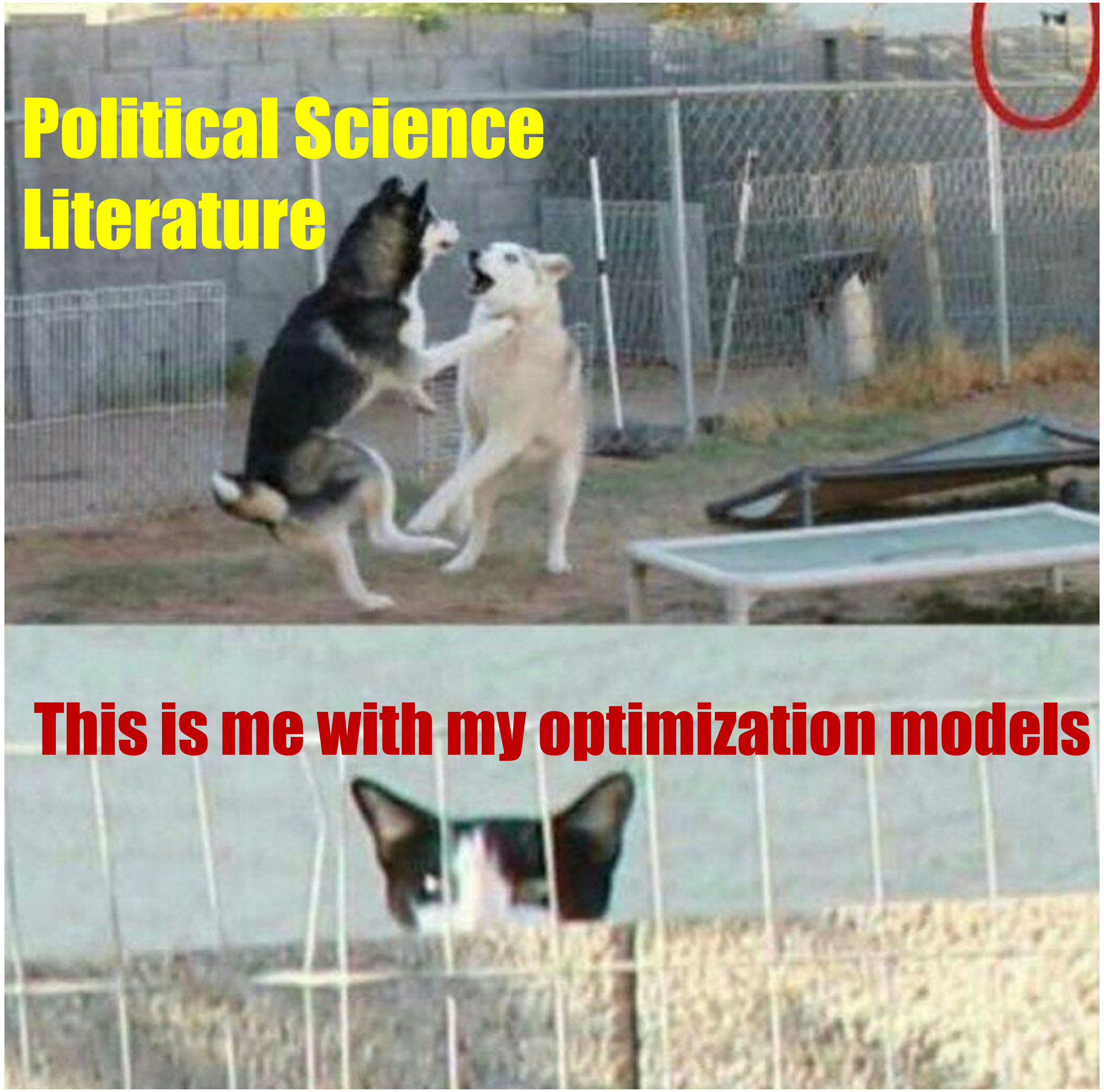
Inferring signed networks of US legislators



Have co-sponsored 242 bills



Would have co-sponsored these many bills if each legislator randomly chooses which bills to co-sponsor


Political Science
Literature
This is me with my optimization models

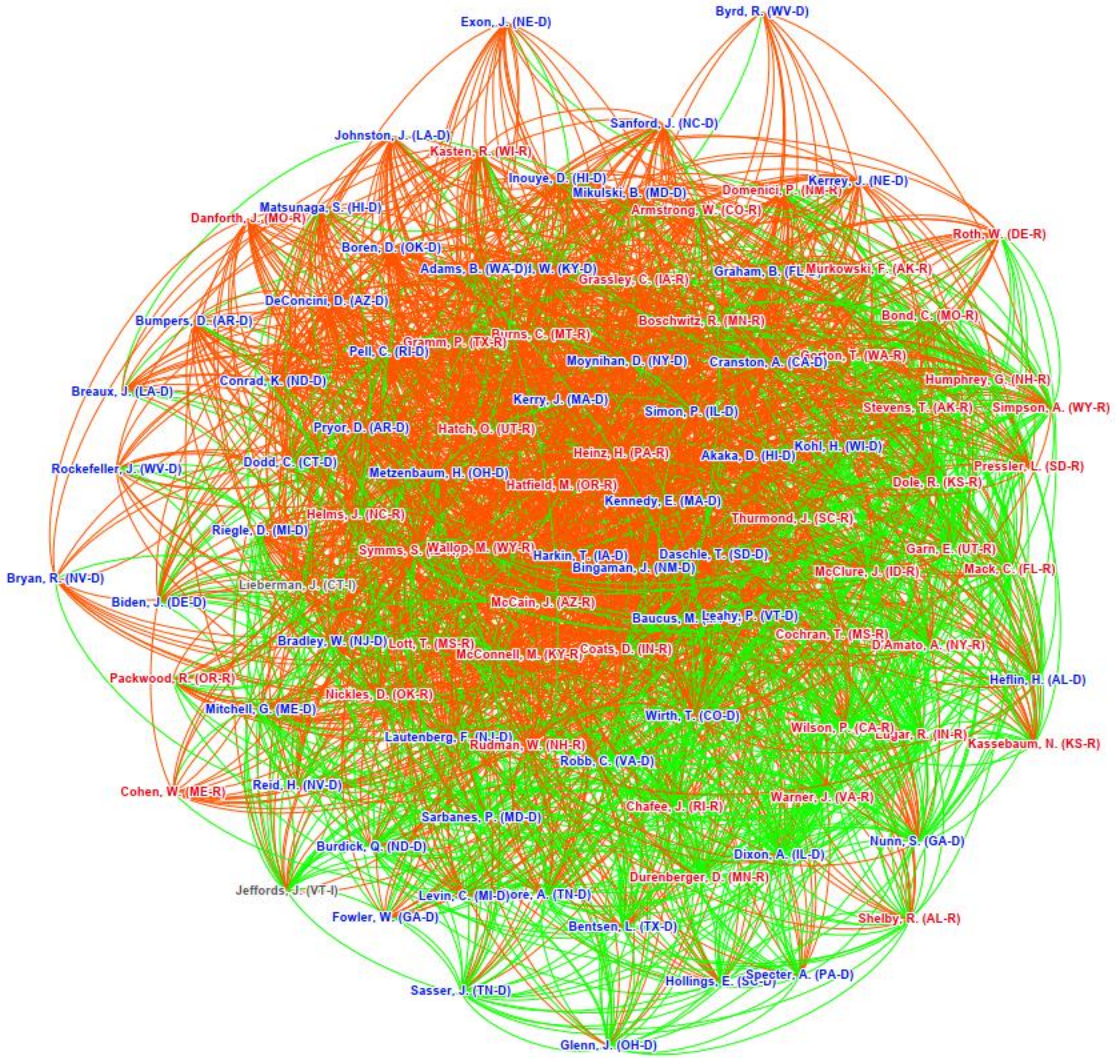
US Senate network 1989-1991 (101st session)

Node colors=party affiliation:

- Republican
- Democrat

Edge colors:

- Significant collaboration
- Significant lack of collaboration



UNIVERSITY OF TORONTO

1979

Node colors=party affiliation:

Republican

Democrat

Edge colors:

Significant collaboration

Significant lack of collaboration


scientific reports

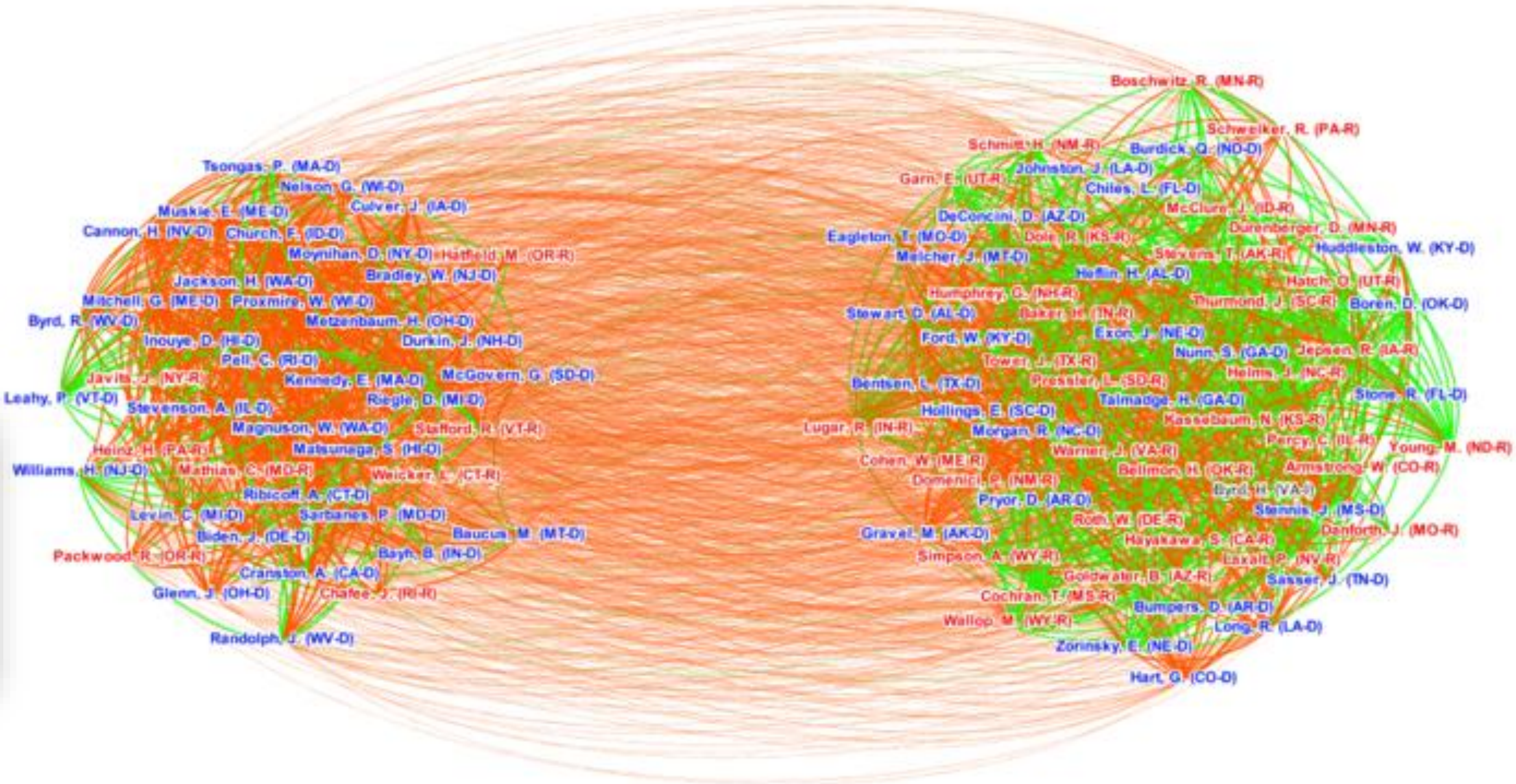
Explore our content ⌄ Journal information ⌄

nature › scientific reports › articles › article

Article | Open Access | Published: 30 January 2020

Detecting coalitions by optimally partitioning signed networks of political collaboration

Samin Aref & Zachary Neal



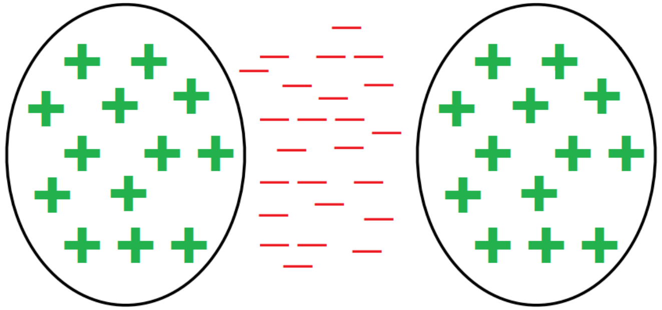
UNIVERSITY OF TORONTO

Classic balance vs. generalized balance

Structurally balanced if and only if:

The network can be partitioned into 2 clusters with positive edges being within clusters and negative edges being between clusters

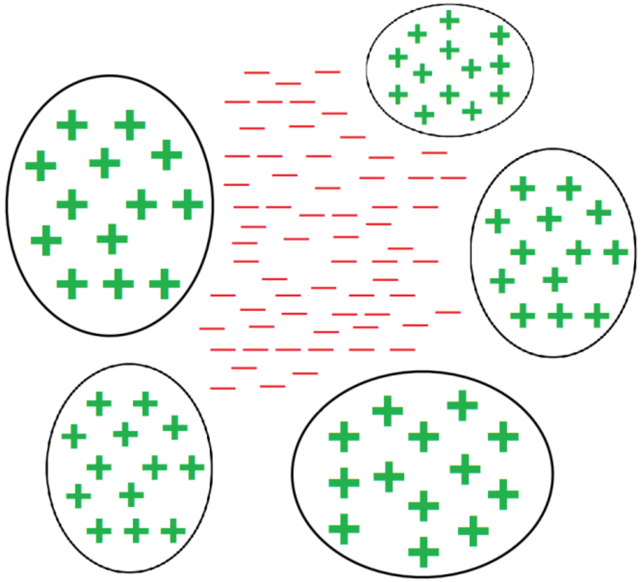
(Cartwright-Harary 1956)



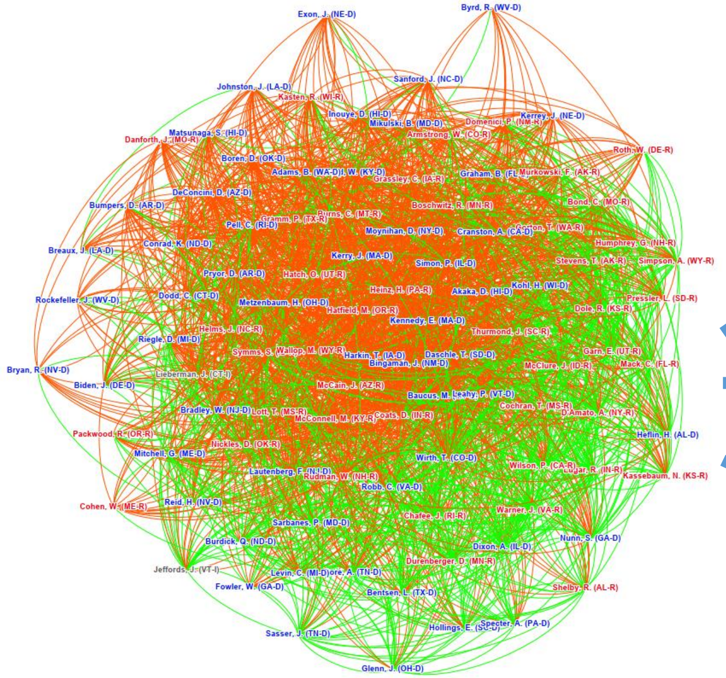
Weakly balanced if and only if:

The network can be partitioned into k clusters with positive edges being within clusters and negative edges being between clusters

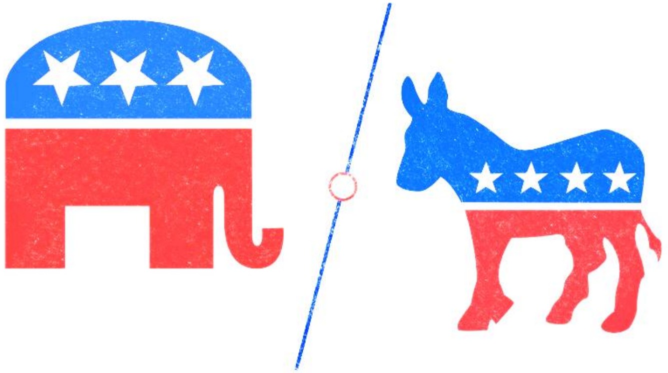
(Davis 1967)



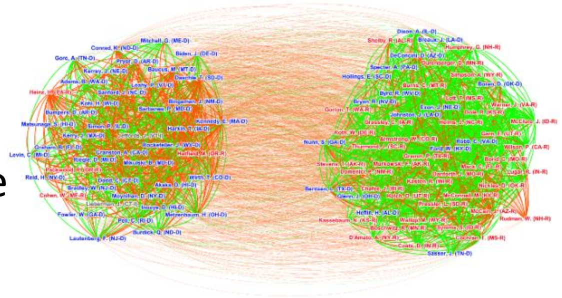
UNIVERSITY OF TORONTO



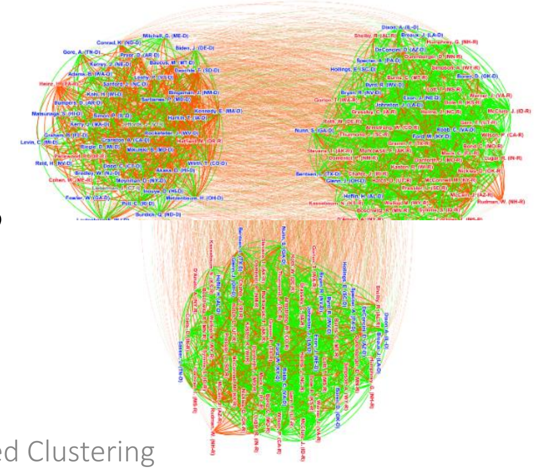
Partition legislators by party



Partition legislators by classic balance

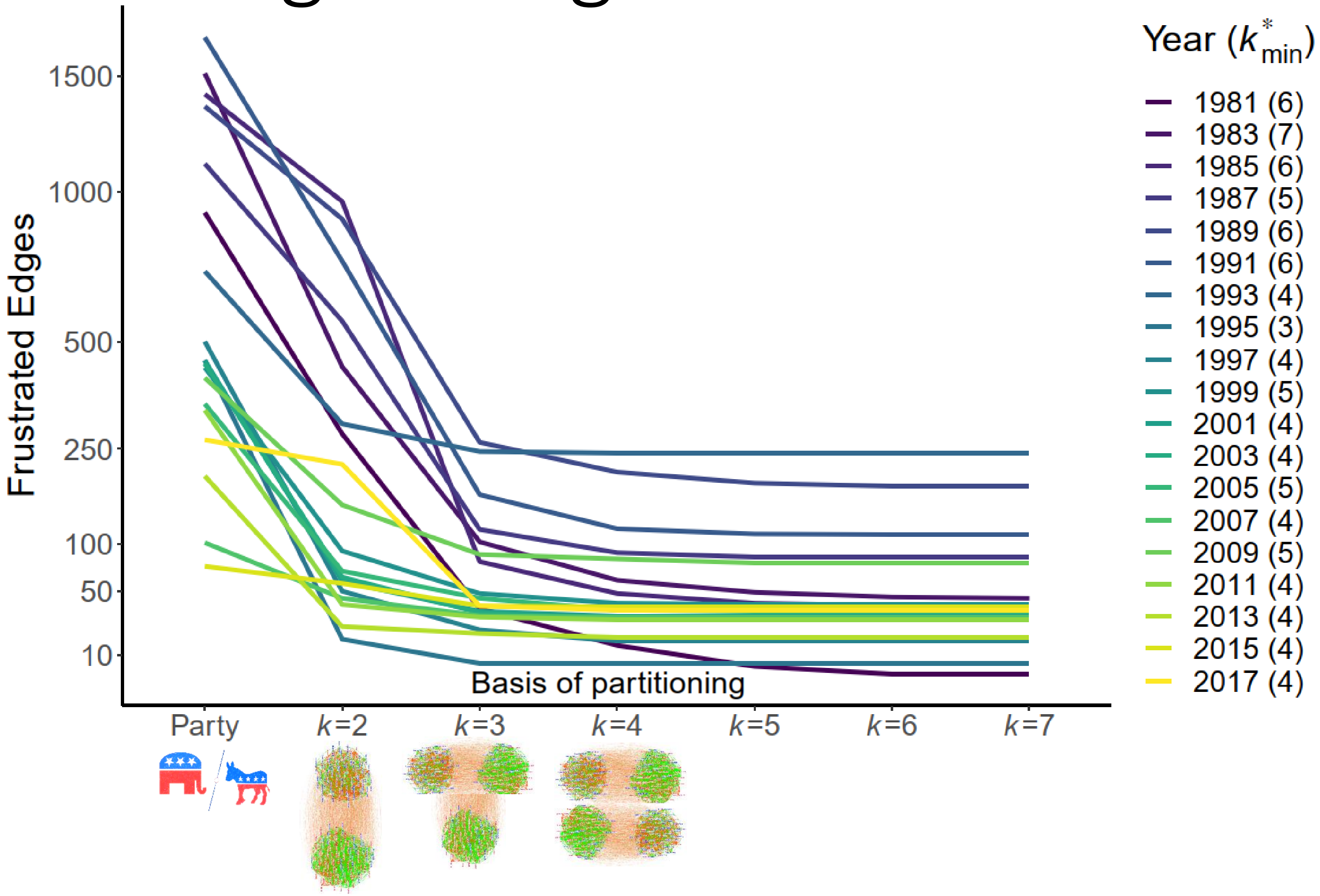


Partition legislators by generalized balance

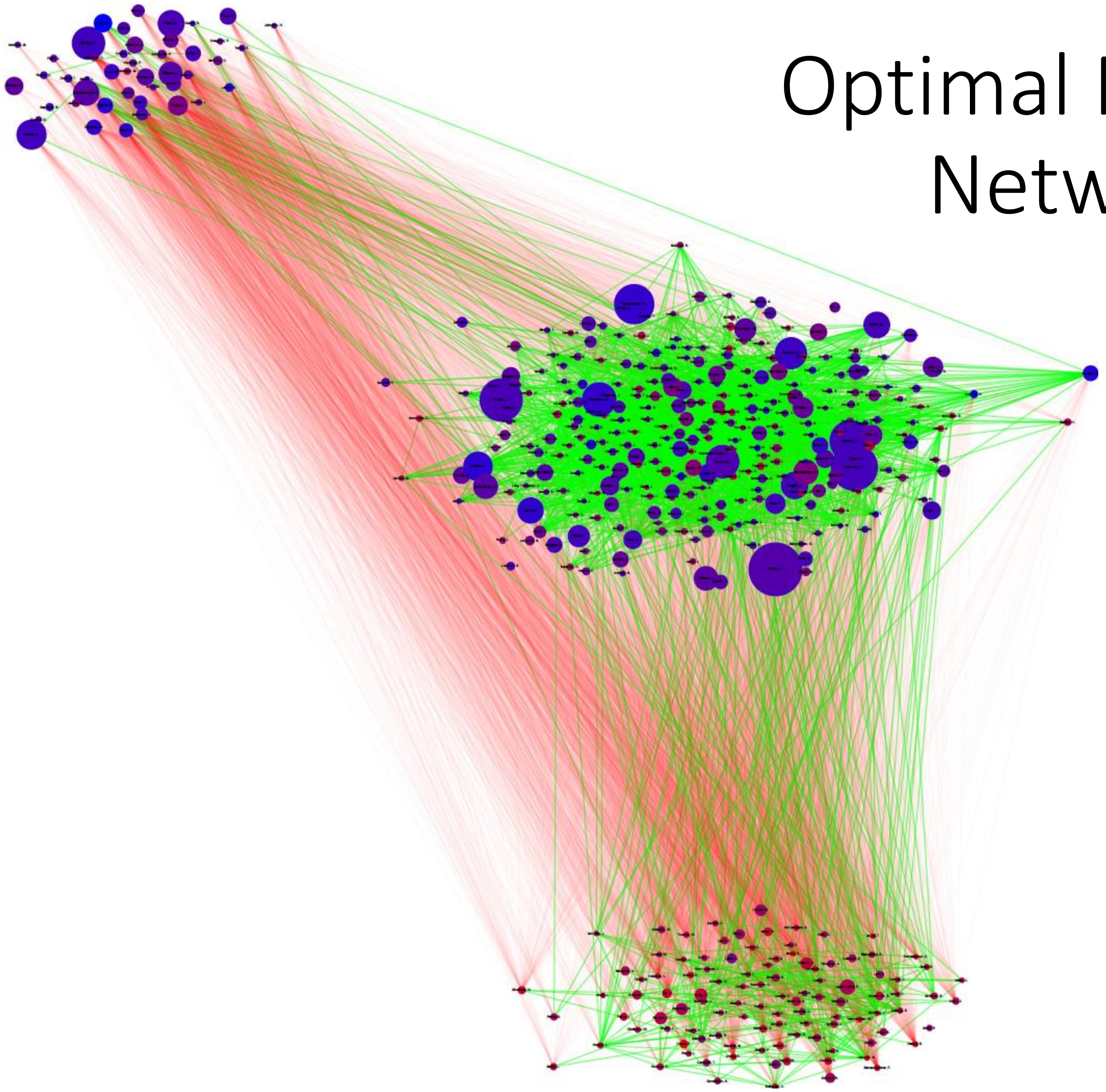


UNIVERSITY OF TORONTO

Partitioning the legislators into clusters



Optimal Partition of US House Network 101 (1989-1991)



scientific reports

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › scientific reports › articles › article

Article | Open Access | Published: 07 October 2021

Identifying hidden coalitions in the US House of Representatives by optimally partitioning signed networks based on generalized balance

Samin Aref & Zachary P. Neal

UNIVERSITY OF TORONTO

Part 4

Background

1. Balance in signed networks

Methodology

2. How can we measure balance?

Use Case

3. Hidden coalitions in the US Congress

Optimization of unsigned networks

4. Network clustering (Bayan algorithm and Troika algorithm)

- There are varying degrees to the success of different algorithmic attempts at optimizing any function.

- Over the past 15-20 years, the term **modularity maximization (MM)** has been used for referring to **attempts at maximizing modularity** (whose success rates have remained unknown).

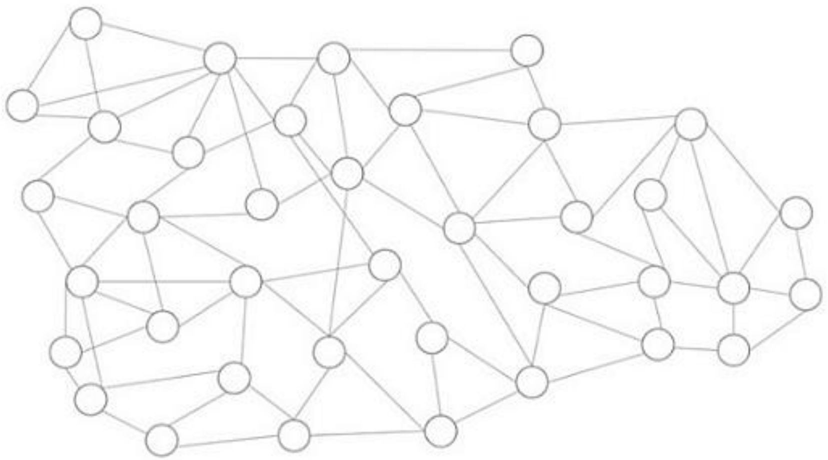
- If those MM attempts fail more than they succeed, we actually do not know much about maximum modularity partitions (for better or worse).

UNIVERSITY OF TORONTO

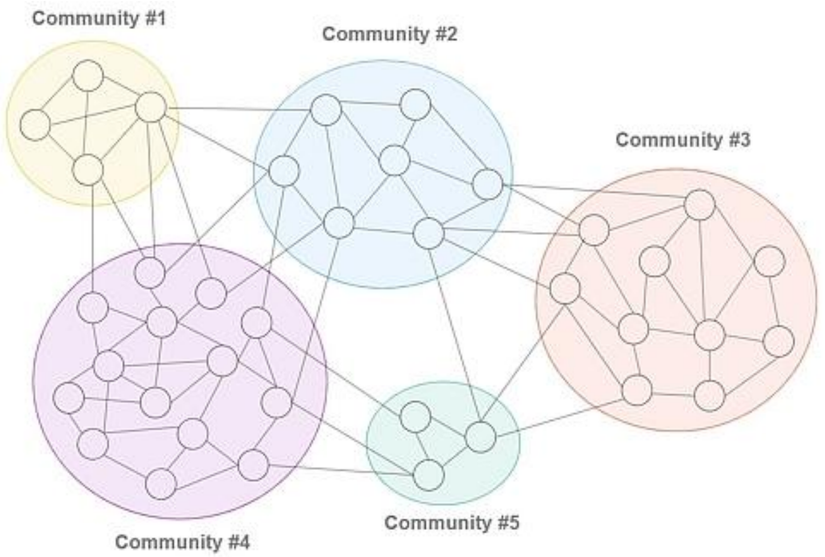
Problem definition

- Input: graph $G(V,E)$ with n nodes and m edges
- Method: a clustering (descriptive community detection) algorithm
- Output: a partition of the nodes into communities (node colours)
- Goal: grouping tightly connected nodes into communities (clusters) which can be gently connected to the rest of the network (Schaub et al. 2017).

Input



Output



Figures from timbr.ai

Schaub et al. (2017) The many facets of community detection in complex networks

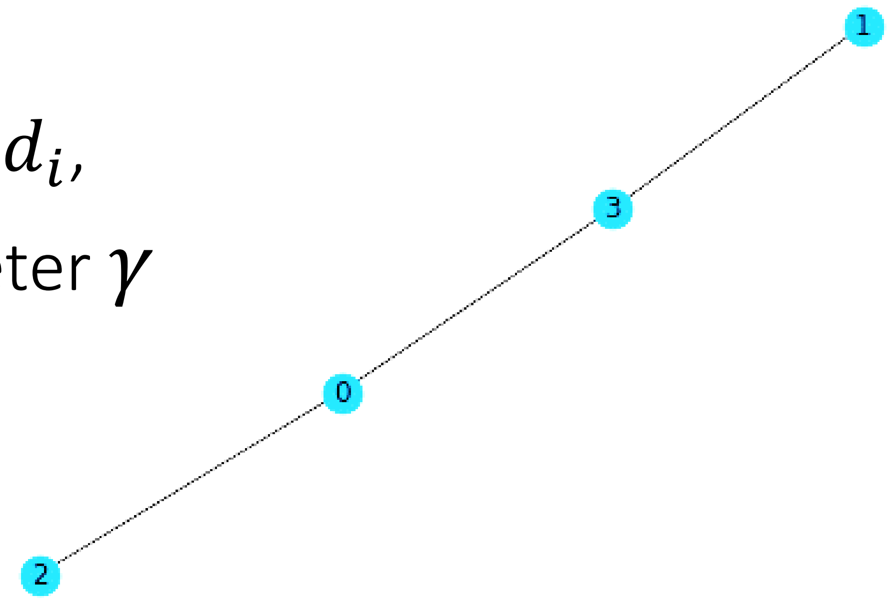
Modularity

Given a network (graph) *G(V,E),* find a partition of *V* into X={V_1, V_2, …, V_c} such that the modularity *Q(G,X)* is maximized.

Modularity entry b_{ij} : a function of degrees d_i,

connections a_{ij}, and the resolution parameter γ

$$b_{ij} = a_{ij} - \gamma \frac{d_i d_j}{2m}$$



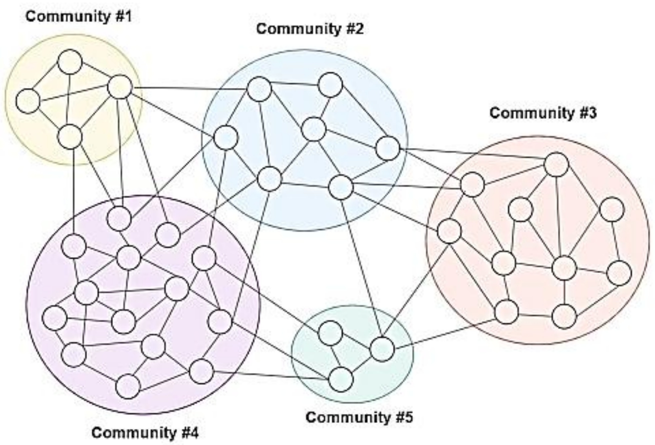
Detecting communities via Modularity Maximization (MM)

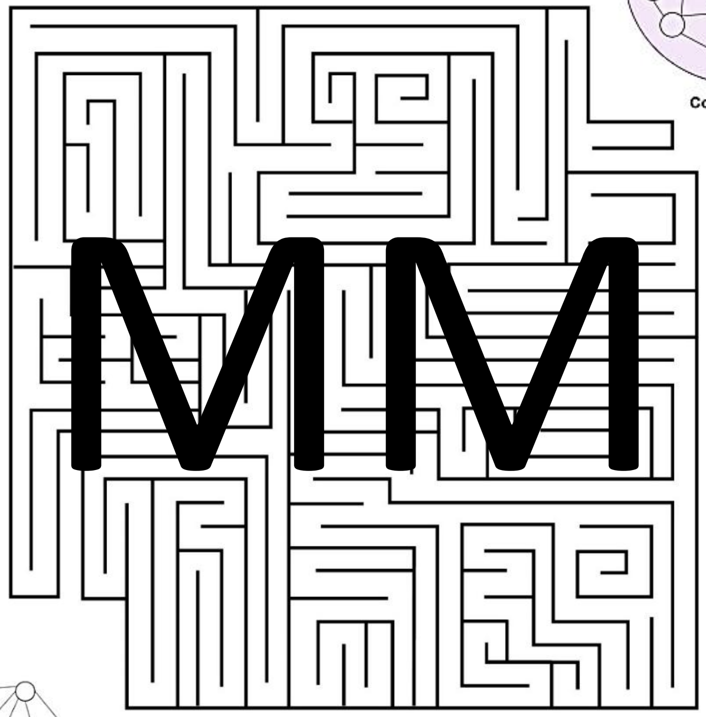
Given graph *G(V,E)* with modularity entries b_{ij}, find a partition of the node set *V* into $X=\{V_1, V_2, \ldots, V_k\}$ to maximizes the objective function *Q(G,X)*.

$$Q_{(G,X)} = \frac{1}{2m} \sum_{(i,j) \in V^2} b_{ij}$$

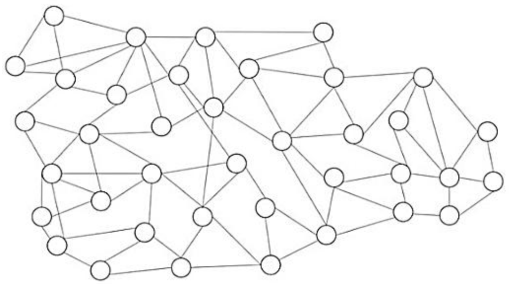
i and j are together in partition X

Output: communities





Input: a simple graph



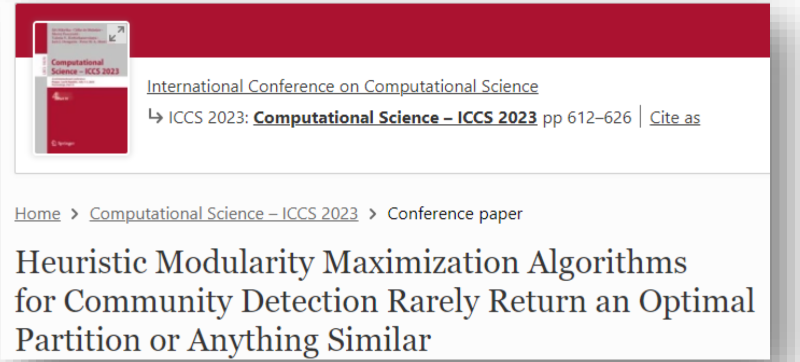
Figures from timbr.ai

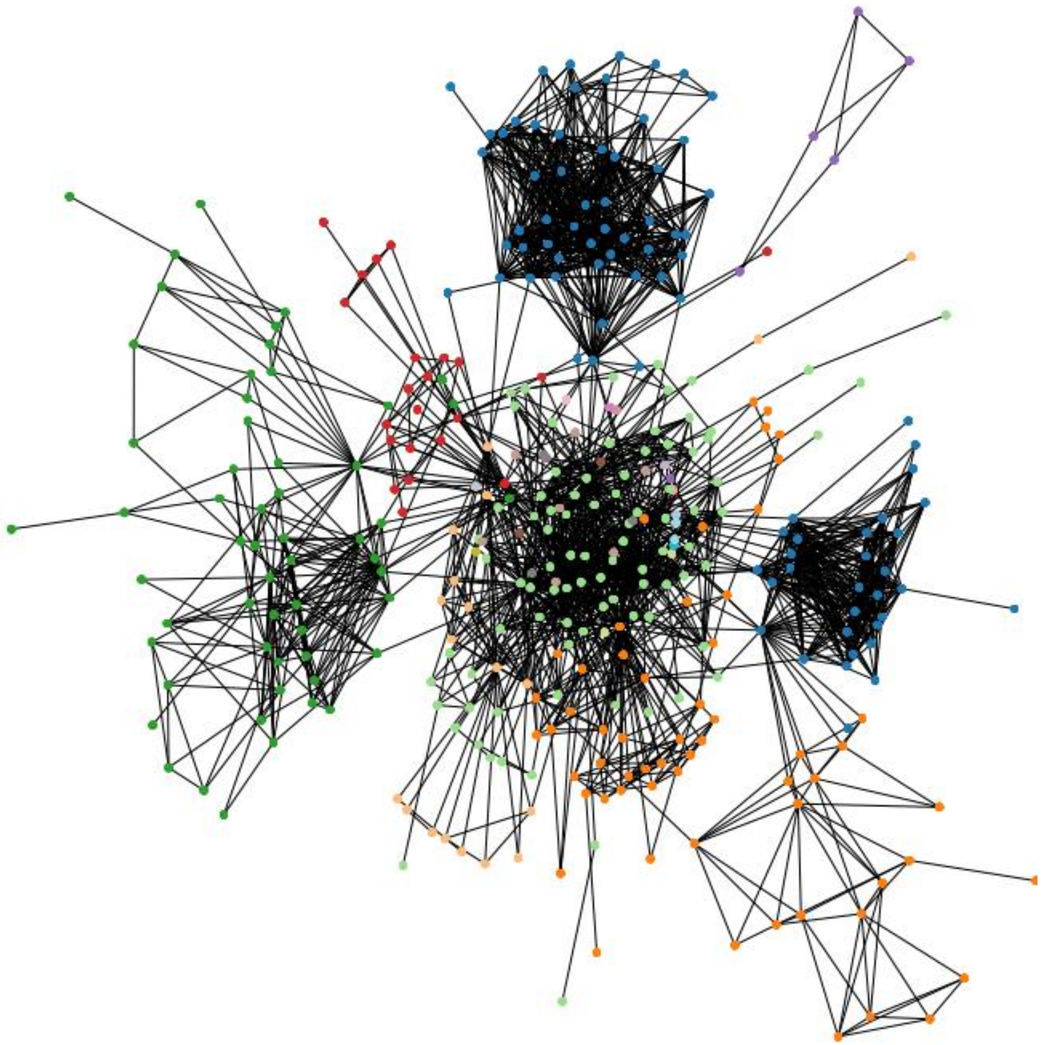
UNIVERSITY OF TORONTO

Communities depend on the algorithm even if the objective function is the same.

Dataset:
facebook_friends
m=1988

Q: modularity for a partition
Q*: maximum modularity of the graph
k: number of communities
AMI: adjusted mutual information
(similarity to an optimal partition)



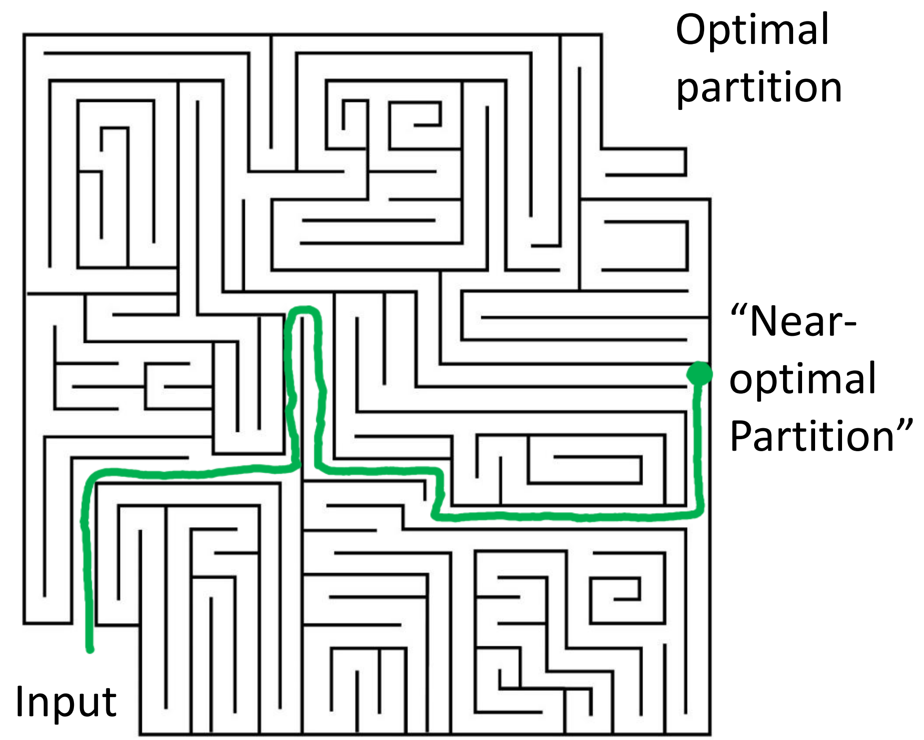


(a) Bayan, $Q^* = 0.7157714$, $k = 28$, $\mathrm{AMI} = 1$

UNIVERSITY OF TORONTO

Approach 1: Modularity Maximization Heuristics

1. Edge Motif (EdMot) (Li et al. 2019)
2. Leiden (Traag et al. 2019)
3. Paris (Bonald et al. 2018)
4. Belief (Zhang & Moore 2014)
5. Combo (Sobolevsky et al. 2014)
6. Leicht-Newman (LN) (Leicht & Newman 2008)
7. Louvain (Blondel et al. 2008)
8. Greedy (CNM) (Clauset et al. 2004)
9. Graph Neural Network (GNN) (Sobolevsky et al. 2022)



UNIVERSITY OF TORONTO

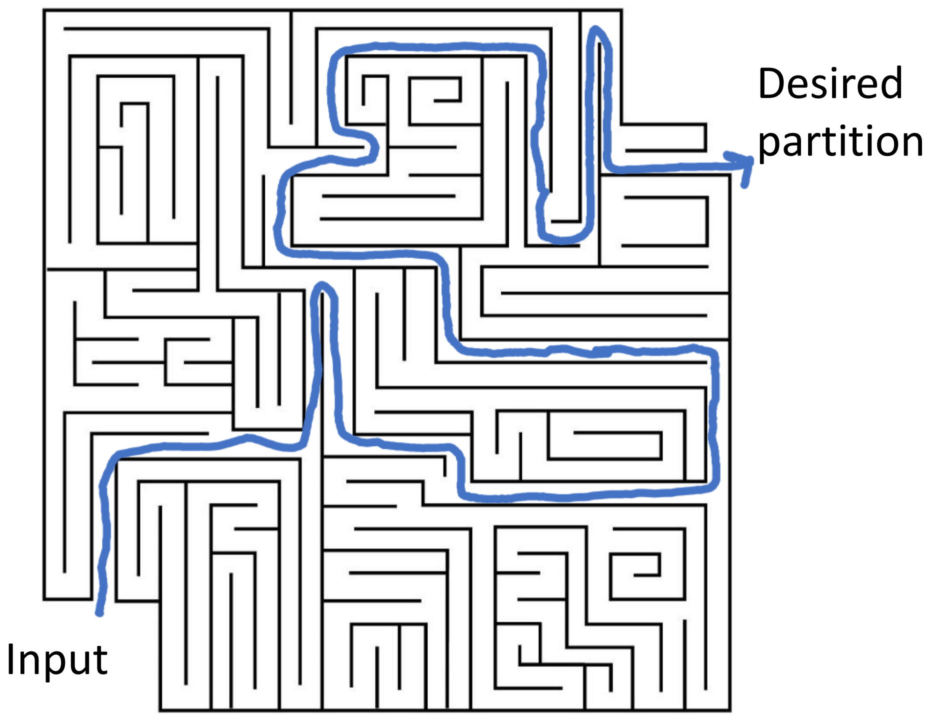
Approach 2: Exact/Approximate Modularity Maximization

- Integer Programming - IP (Brandes et al. 2007)
- IP and LP rounding (Agarwal & Kempe 2008)
- Column generation (Aloise et al. 2010)
- Sparse IP and LP rounding (Dinh & Thai 2015)
- Approximation (Kawase et al. 2021)
- Bayan (Aref et al. 2022) for graphs with m<3000

$$\max_{x_{ij}} Q = \frac{1}{2m}\left(\sum_{(i,j)\in V^2, i<j} 2b_{ij}(1-x_{ij}) + \sum_{i\in V} b_{ii}\right)$$

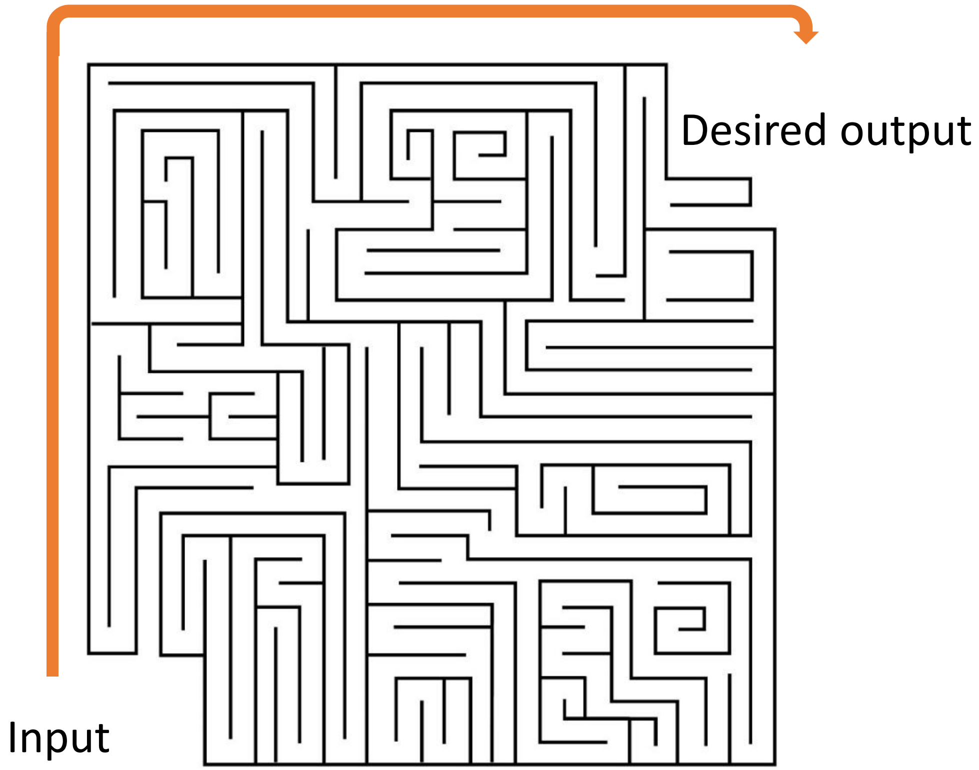
$$\text{s.t.} \quad x_{ik} + x_{jk} \geq x_{ij} \quad \forall (i,j) \in V^2, i<j, k \in K(i,j)$$

$$x_{ij} \in \{0,1\} \quad \forall (i,j) \in V^2, i<j$$

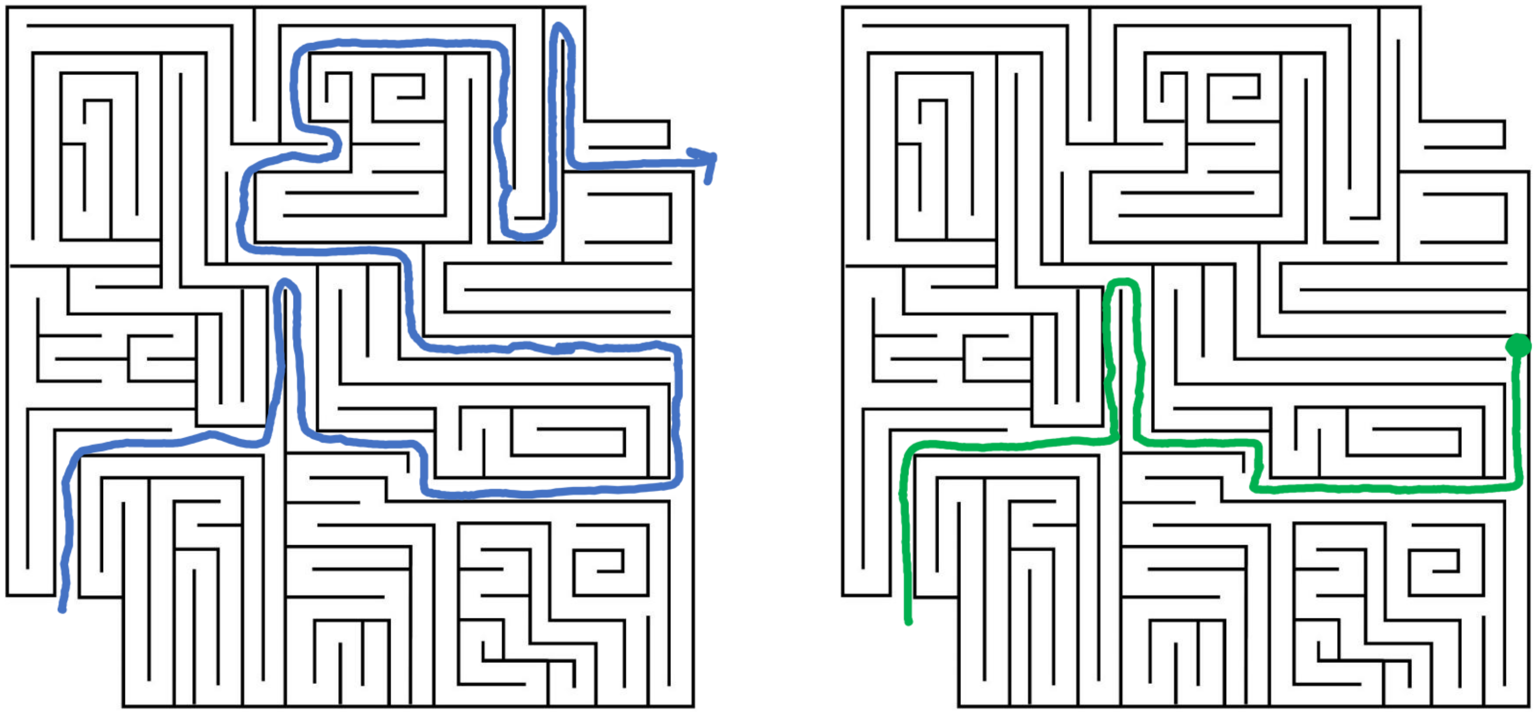


Approach 3: Other Methods

1. Kernighan-Lin bisection (Kernighan and Lin 1970)
2. RB Potts model with Erdős–Rényi as null (Reichardt & Bornholdt 2006)
3. Chinese whispers (Biemann et al. 2006)
4. Walktrap (Pons & Latapy 2006)
5. k-cut (Ruan & Zhang 2007)
6. Asynchronous label propagation (Raghavan et al. 2007)
7. Infomap (Rosvall & Bergstrom 2008)
8. Genetic Algorithm (Pizzuti 2008)
9. Semi-synchronous Label propagation (Cordasco & Gargano 2010)
10. Constant Potts Model (CPM) (Traag et al. 2011)
11. Significant scales (Traag et al. 2013)
12. Stochastic Block Model (SBM) (Peixoto 2014a)
13. SBM with Monte Carlo Markov Chain (MCMC) (Peixoto 2014b)
14. WCC (Prat-Pérez et al. 2014)
15. Surprise (Traag et al. 2015)
16. Diffusion Entropy Reducer (DER) (Kozdoba and Mannor 2015)
17. GemSec (Rozemberczki et al. 2019)
18. Bayesian Planted Partition (BPP) (Zhang and Peixoto, 2020)
19. Markov Stability (PyGenStability) (Arnaudon et al., 2023)

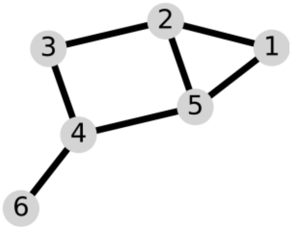


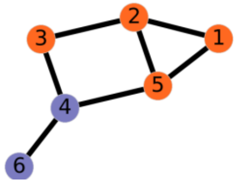
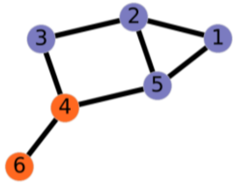
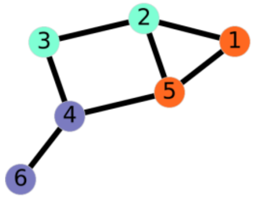
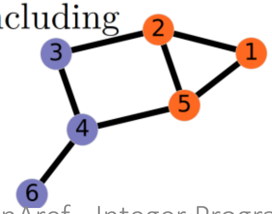
Assessing the optimality of partitions among ten modularity-based algorithms



UNIVERSITY OF TORONTO

Evaluating ten modularity-based algorithms



Ten algorithms (and their variations)	Partition (obtained by an MM algorithm)	Modularity	Optimality percentage (GOP)	Maximum similarity to any optimal partition - AMI (or RMI or ECS)
Method 1		Q_1=0.122	Q_1/Q^*=60%	36%
Method 2		Q_2=0.122	Q_2/Q^*=60%	36%
...	...	...	...	...
Method 19		Q_{19}=0.092	Q_{19}/Q^*=45%	30%
Exact method (Integer Programming)	All optimal partitions including 	Q^*=0.204	100%	100%
			Y-axis	X-axis

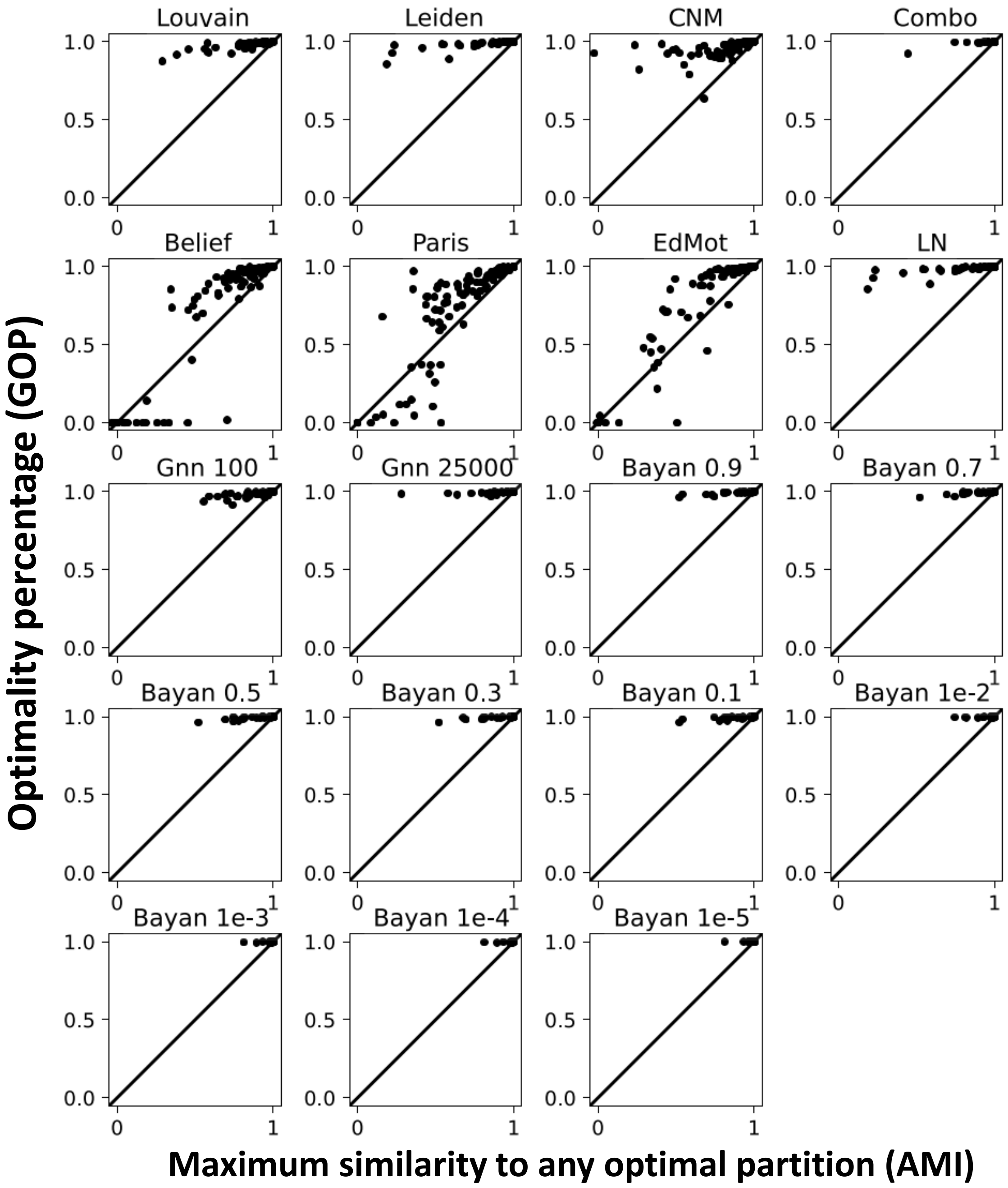
UNIVERSITY OF TORONTO

Test cases are 104 graphs with modular structure:

- 54 real networks
- 20 LFR random graphs
- 30 ABCD random graphs

Each datapoint represents the performance of one algorithm on one test case.

1. Y-values: Most partitions obtained from heuristics are sub-optimal
2. X-values: Many partitions are dissimilar to any optimal partition
3. 45°-line: near-optimal partitions are not similar to any optimal partitions



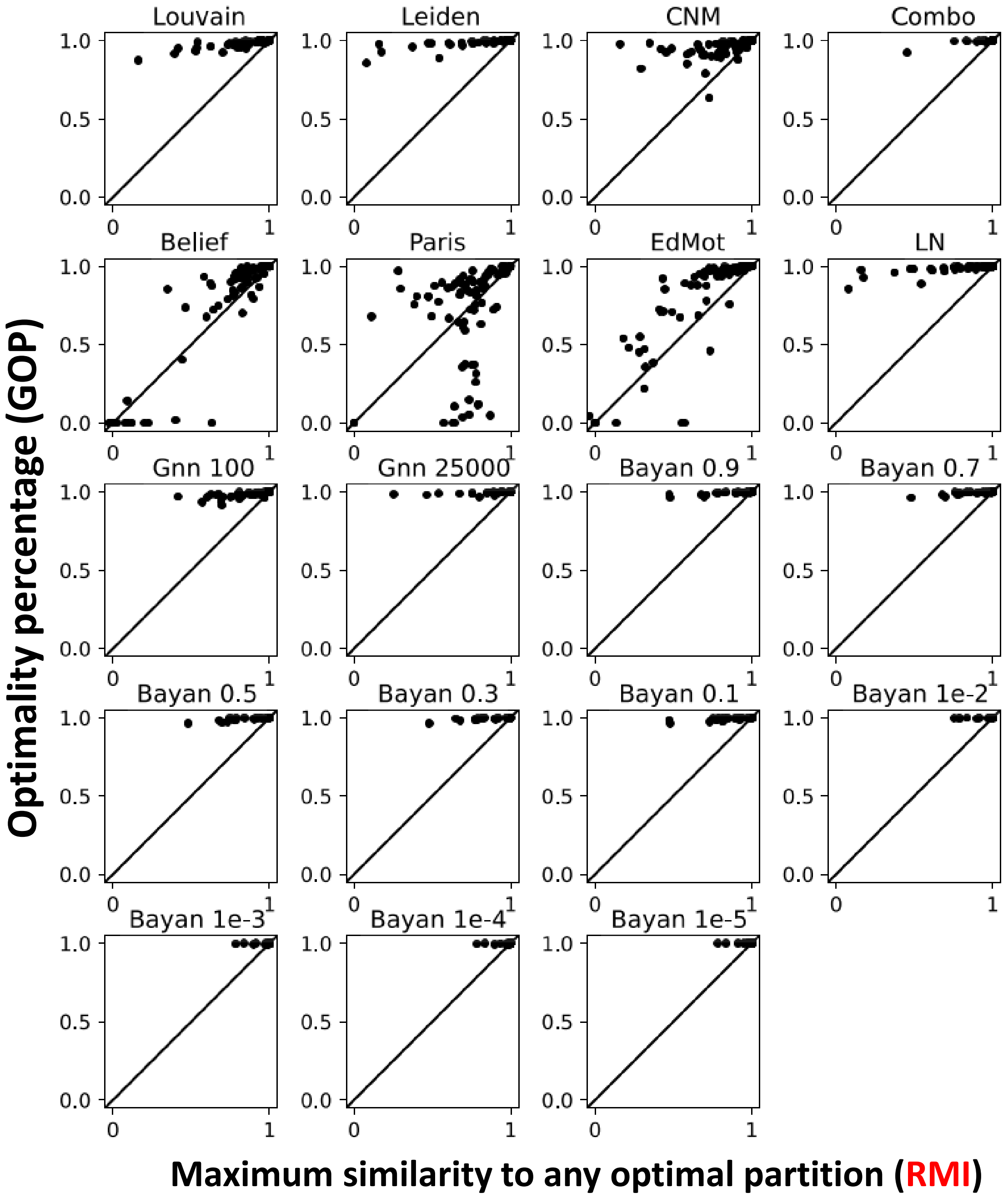
UNIVERSITY OF TORONTO

Test cases are 104 graphs with modular structure:

- 54 real networks
- 20 LFR random graphs
- 30 ABCD random graphs

Each datapoint represents the performance of one algorithm on one test case.

1. Y-values: Most partitions obtained from heuristics are sub-optimal
2. X-values: Many partitions are dissimilar to any optimal partition
3. 45°-line: near-optimal partitions are not similar to any optimal partitions



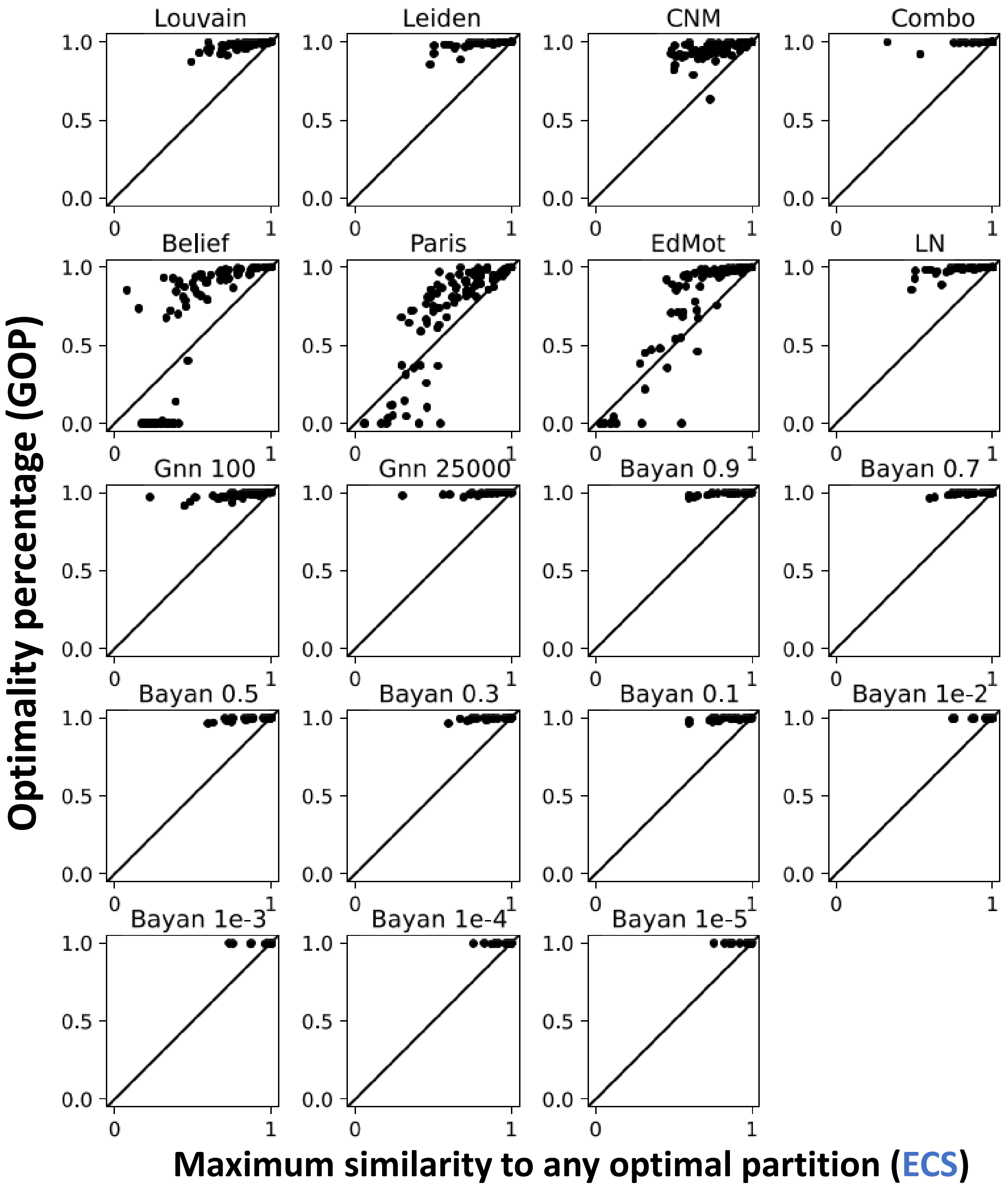
UNIVERSITY OF TORONTO

Test cases are 104 graphs with modular structure:

- 54 real networks
- 20 LFR random graphs
- 30 ABCD random graphs

Each datapoint represents the performance of one algorithm on one test case.

1. Y-values: Most partitions obtained from heuristics are sub-optimal
2. X-values: Many partitions are dissimilar to any optimal partition
3. 45°-line: near-optimal partitions are not similar to any optimal partitions

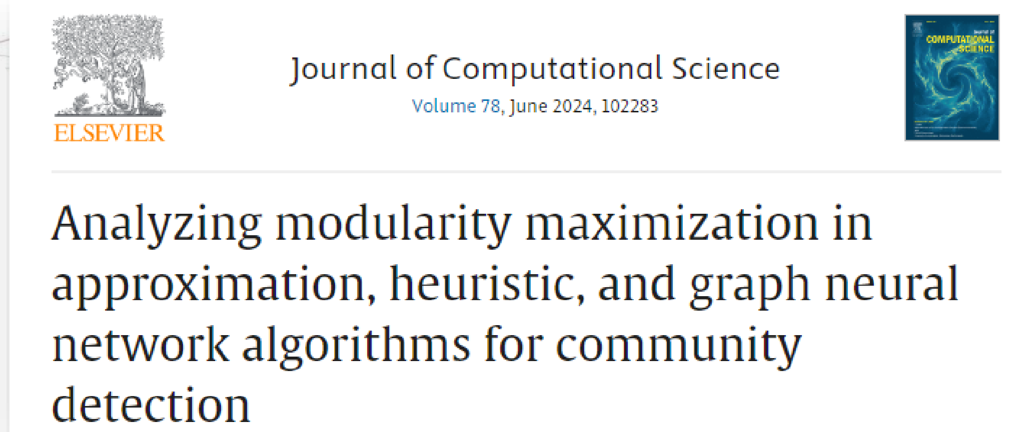


UNIVERSITY OF TORONTO

1

Heuristic modularity maximization algorithms rarely* maximize modularity.

*Only 43.9% of the times according to our experiments on 104 synthetic and real networks.


ELSEVIER

Journal of Computational Science

Volume 78, June 2024, 102283

Analyzing modularity maximization in approximation, heuristic, and graph neural network algorithms for community detection

UNIVERSITY OF TORONTO

2

Suboptimal partitions of heuristic algorithms are disproportionately dissimilar to any optimal partition.

An x% suboptimality is often associated with a dissimilarity much larger than x% from any optimal partition.

UNIVERSITY OF TORONTO

Can we globally maximize modularity?

- YES, for some networks we can!
- Bayan algorithm maximizes modularity in networks with fewer than 3000 edges and approximates maximum modularity in slightly larger networks on ordinary computers.

```
%pip install bayanpy
```

```
import networkx as nx
import bayanpy
```

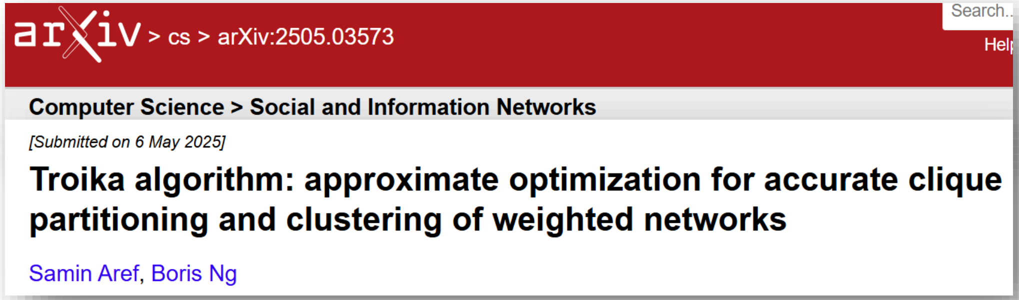
```
G = nx.barbell_graph(5,2)
```

```
bayanpy.bayan(G)
```

UNIVERSITY OF TORONTO

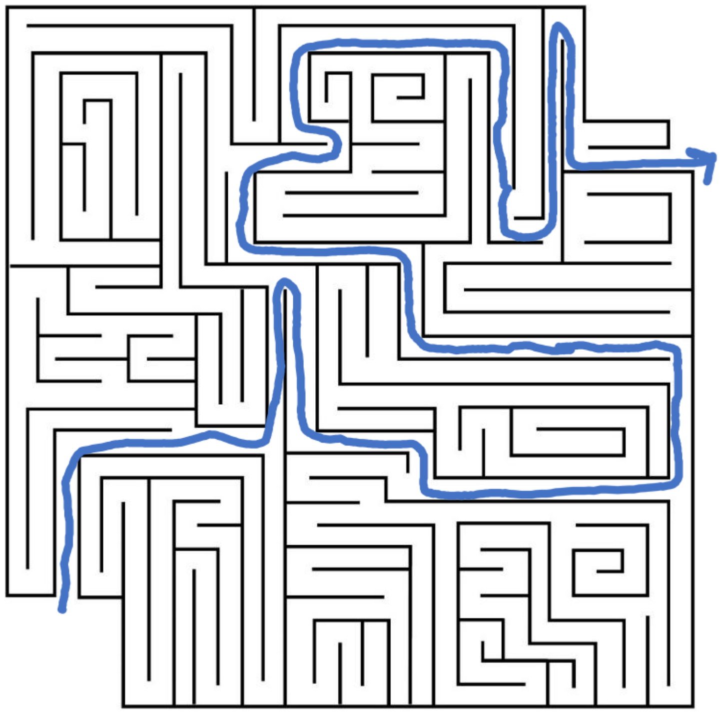
Can we globally optimize other functions?

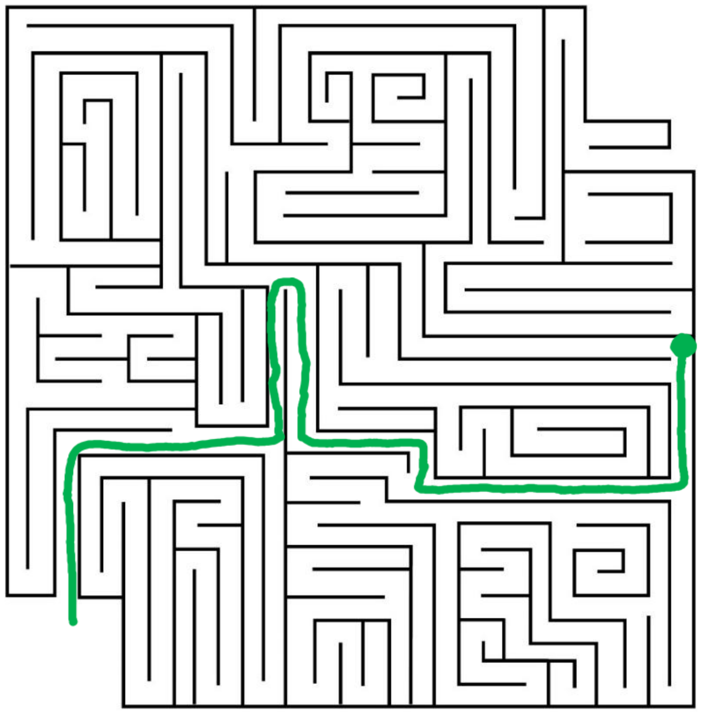
- YES, for some networks we can!
- Bayan algorithm maximizes modularity in networks with fewer than 3000 edges and approximates maximum modularity in slightly larger networks on ordinary computers.
- Modularity maximization on an unsigned graph reduces to *clique partitioning* on a weighted signed graph.
- Troika algorithm solves clique partitioning (optimizes total within-cluster weight).



UNIVERSITY OF TORONTO

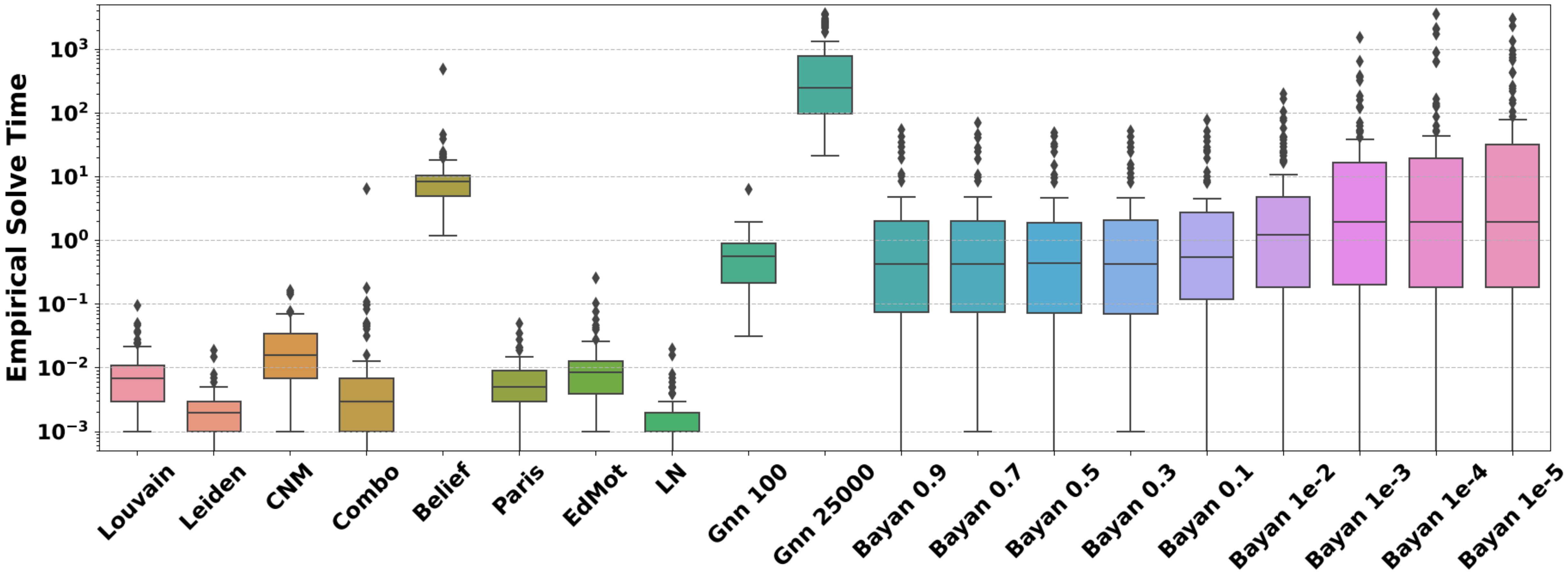
Assessing the run time of ten modularity-based methods (and their variations)





UNIVERSITY OF TORONTO

Solve time of ten modularity-based methods

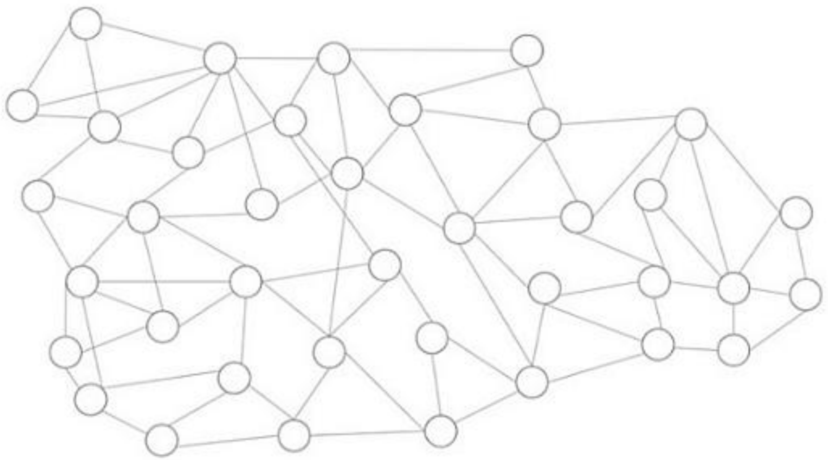


Wait! But clustering is not about modularity maximization!

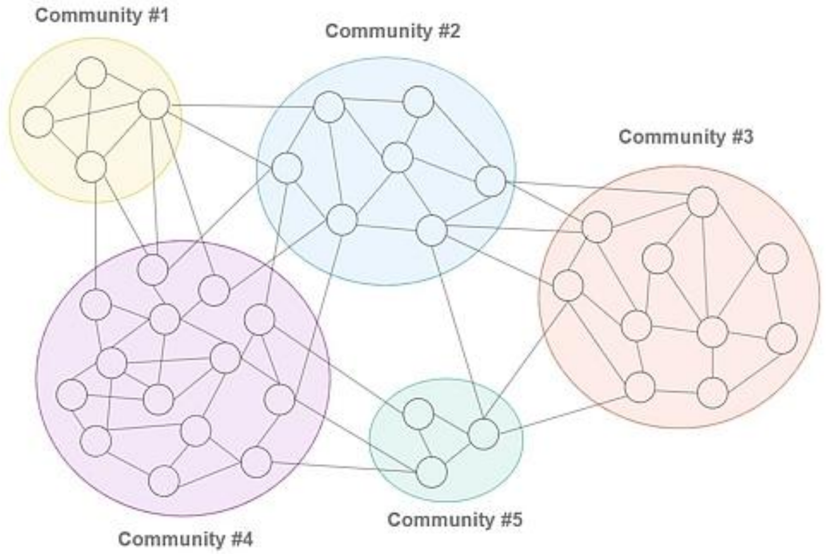
<u>A supervised comparison of 30 CD algorithms</u>

- Input: graph
- Method: a network clustering algorithm
- Output: a partition
- Test: **retrieval of planted partitions** in synthetic benchmark networks (e.g. LFR)

Input

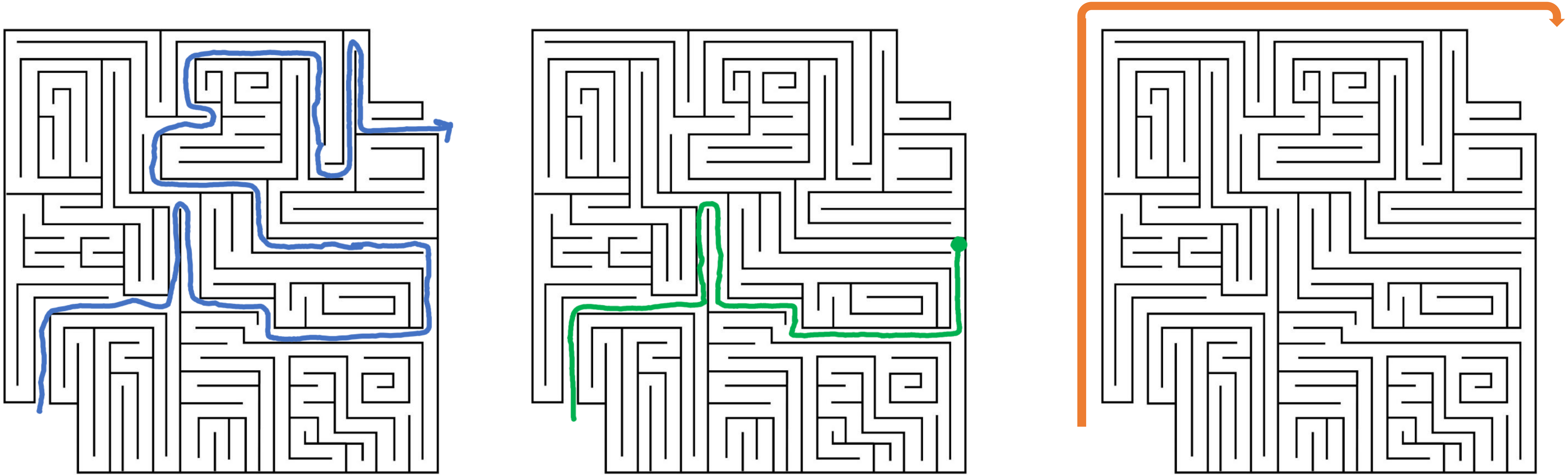


Output



Figures from timbr.ai

Comparing 30 algorithms based on retrieving planted partitions



Planted partition retrieval assessment

Standard benchmark graphs:

- LFR benchmarks (**L**ancichinetti-**F**ortunato-**R**adicchi benchmarks)
- ABCD benchmarks (**A**rtificial **B**enchmark for **C**ommunity **D**etection)

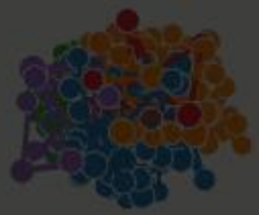
Retrieval performance indicator:

- Similarity with the planted partitions

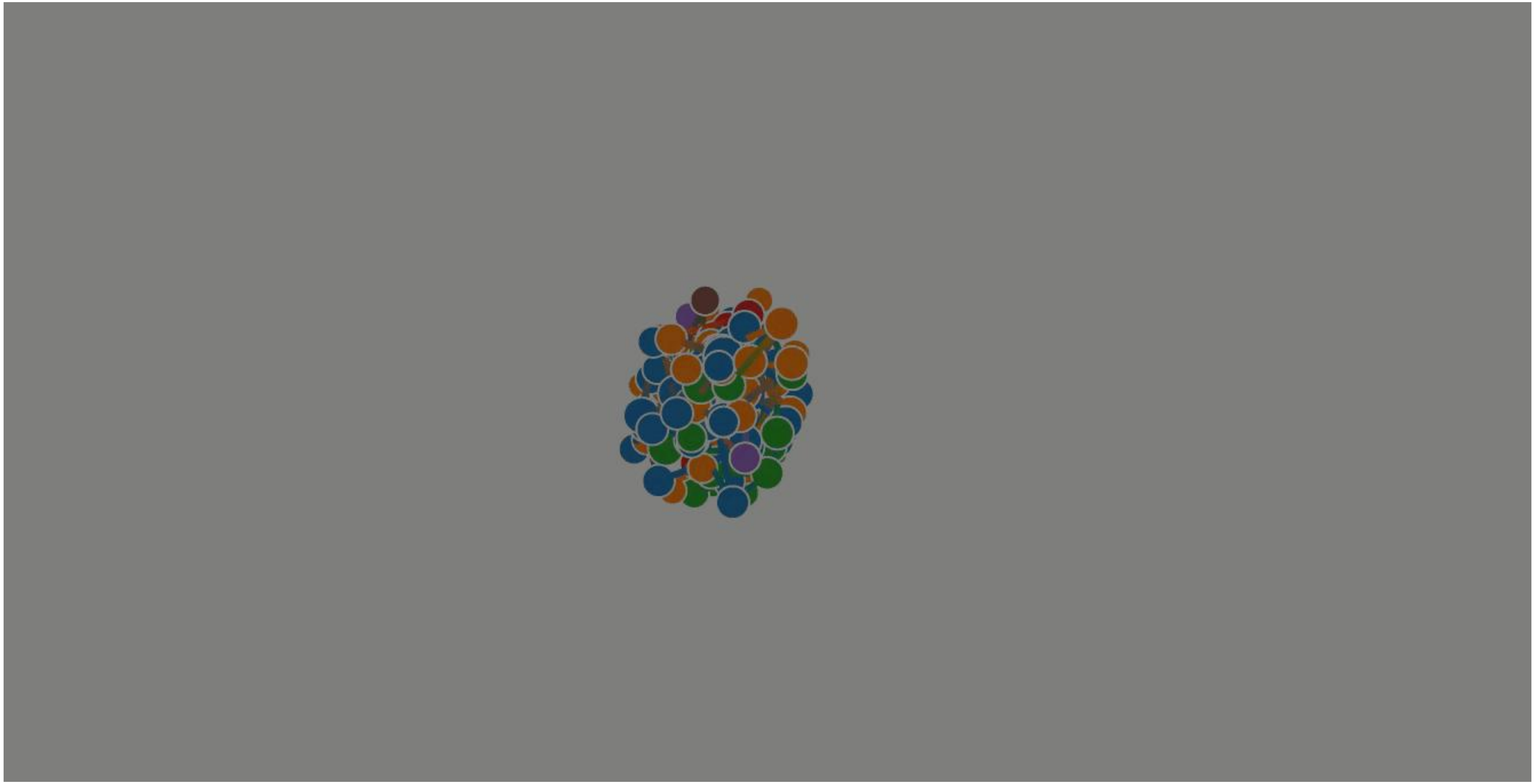
adjusted mutual information (or RMI[1] or ECS[2]) averaged over 100 graphs

1. RMI: Reduced Mutual Information (Newman et al. 2020, Jerdee and Newman 2023)
2. ECS: Element-Centric Similarity (Gates et al. 2019)

LFR benchmark with **low** noise

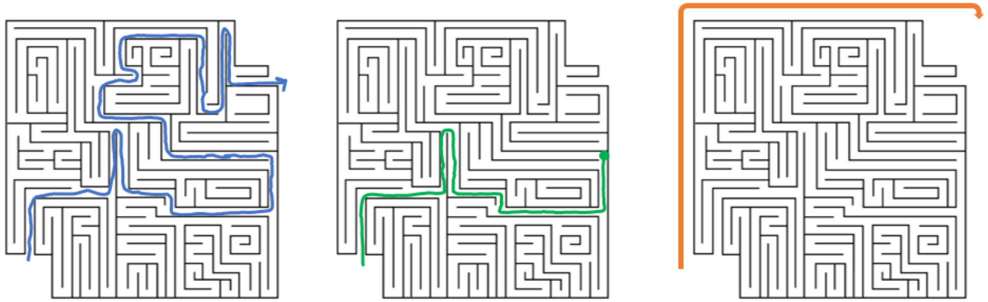


LFR benchmark with **high** noise



UNIVERSITY OF TORONTO

Comparing 30 algorithms



Test cases:

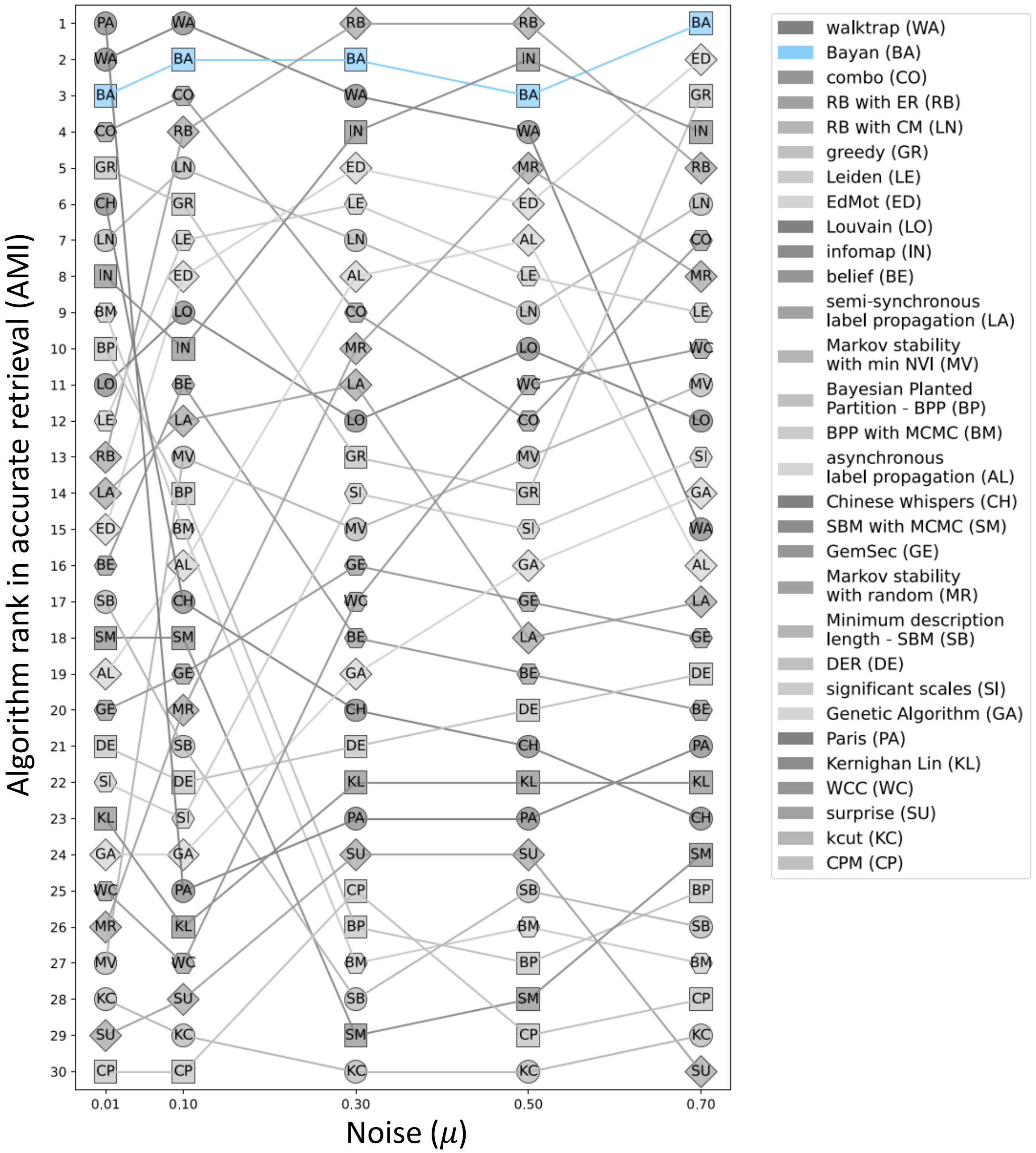
- 500 LFR benchmark graphs with up to 1000 edges (generated with planted partitions)
- Noise μ (fraction of inter-community edges)

$$\mu \in \{1\%, 10\%, 30\%, 50\%, 70\%\}$$

UNIVERSITY OF TORONTO

Comparing 30 algorithms

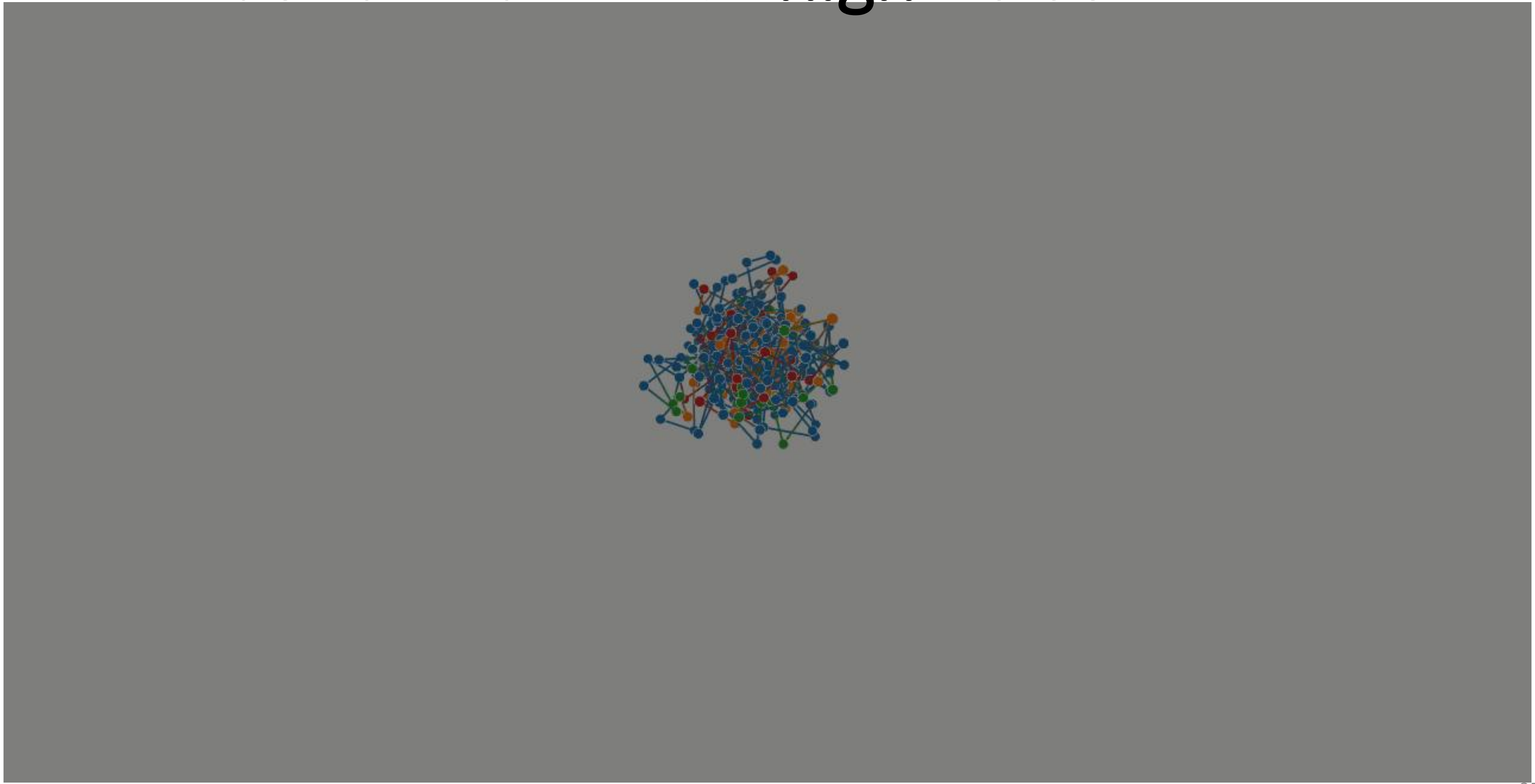
Based on retrieving the planted partitions of LFR benchmarks



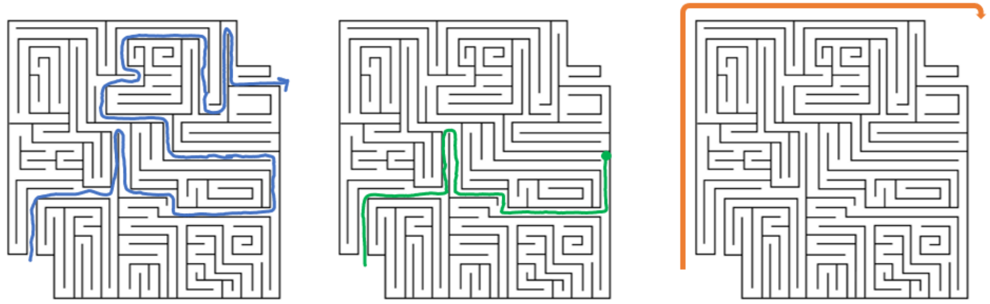
UNIVERSITY OF TORONTO

ABCD benchmark with **low** noise

ABCD benchmark with **high** noise



Comparing 30 algorithms



Test cases:

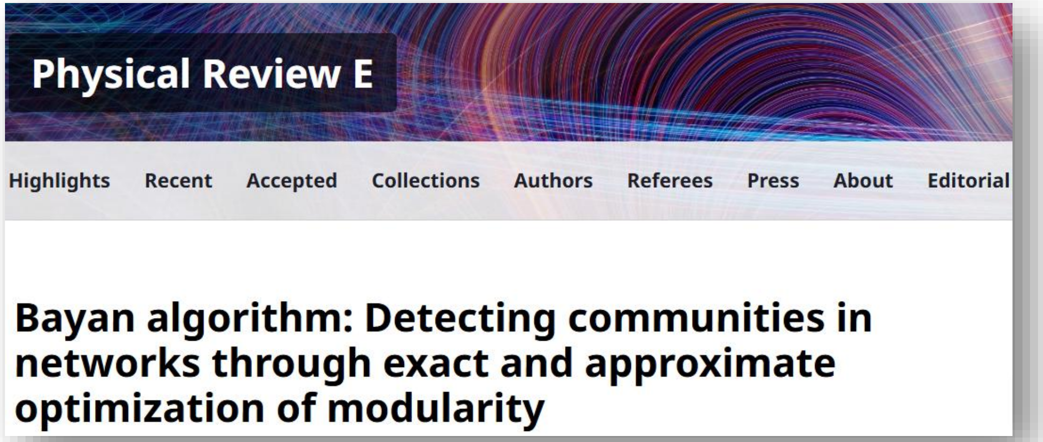
- 500 ABCD benchmark graphs with up to 1000 edges (generated with planted partitions)
- Noise ξ (mixing parameter)

$$\xi \in \{10\%, 30\%, 50\%, 70\%, 90\%\}$$

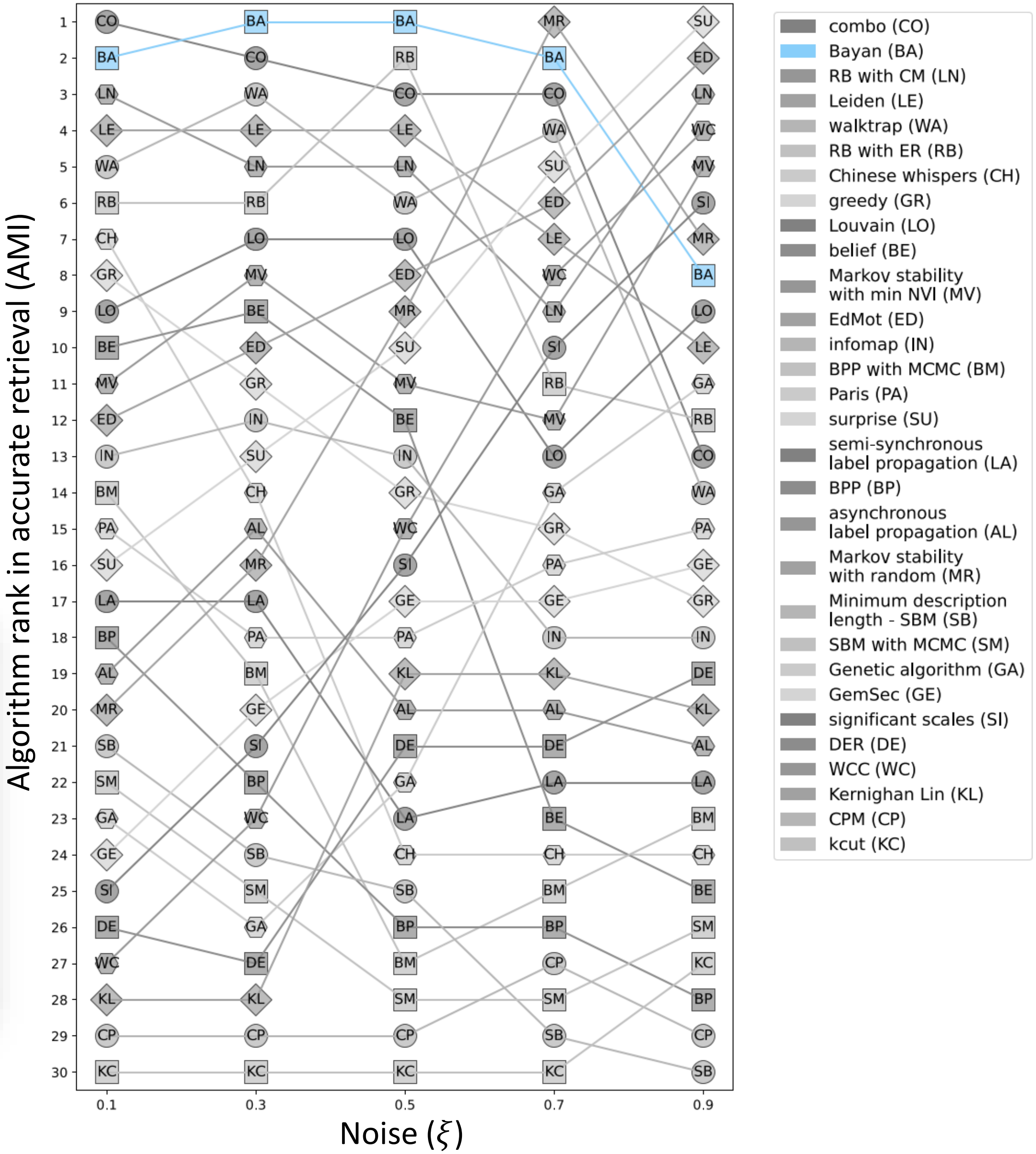
UNIVERSITY OF TORONTO

Comparing 30 algorithms

Based on retrieving the planted partition of ABCD benchmark

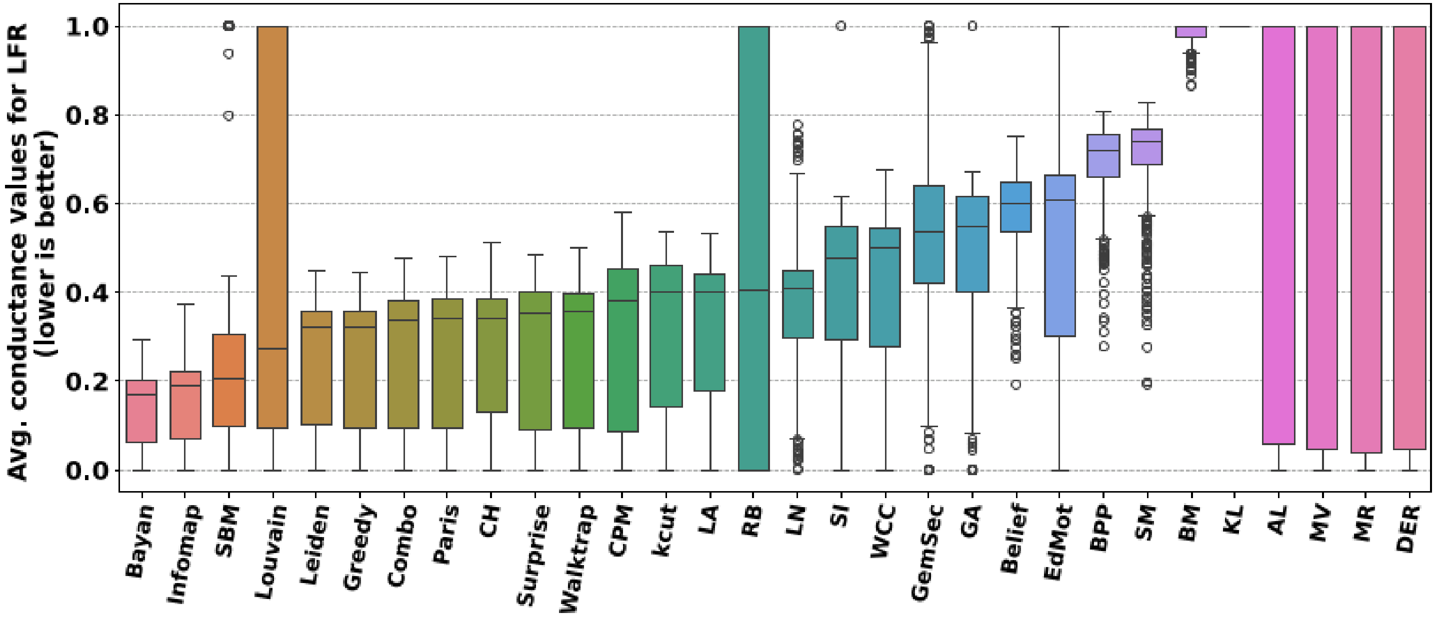

Physical Review E
Highlights Recent Accepted Collections Authors Referees Press About Editorial

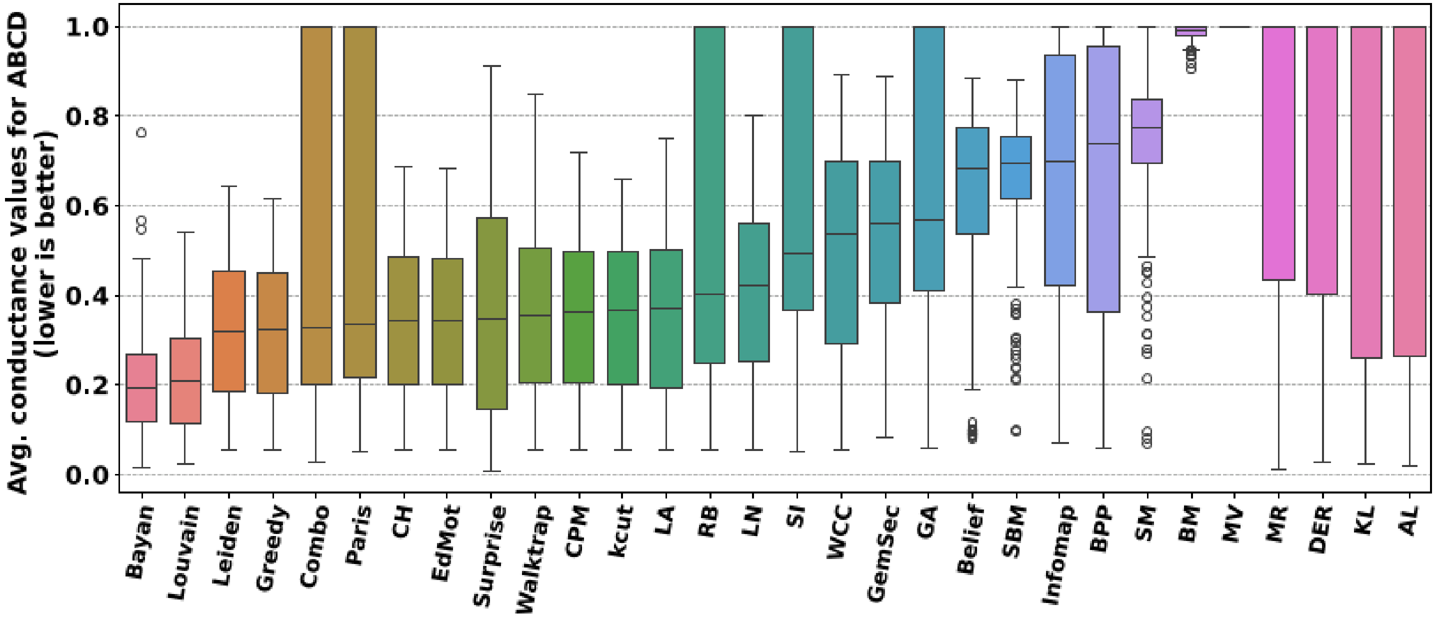
Bayan algorithm: Detecting communities in networks through exact and approximate optimization of modularity



UNIVERSITY OF TORONTO

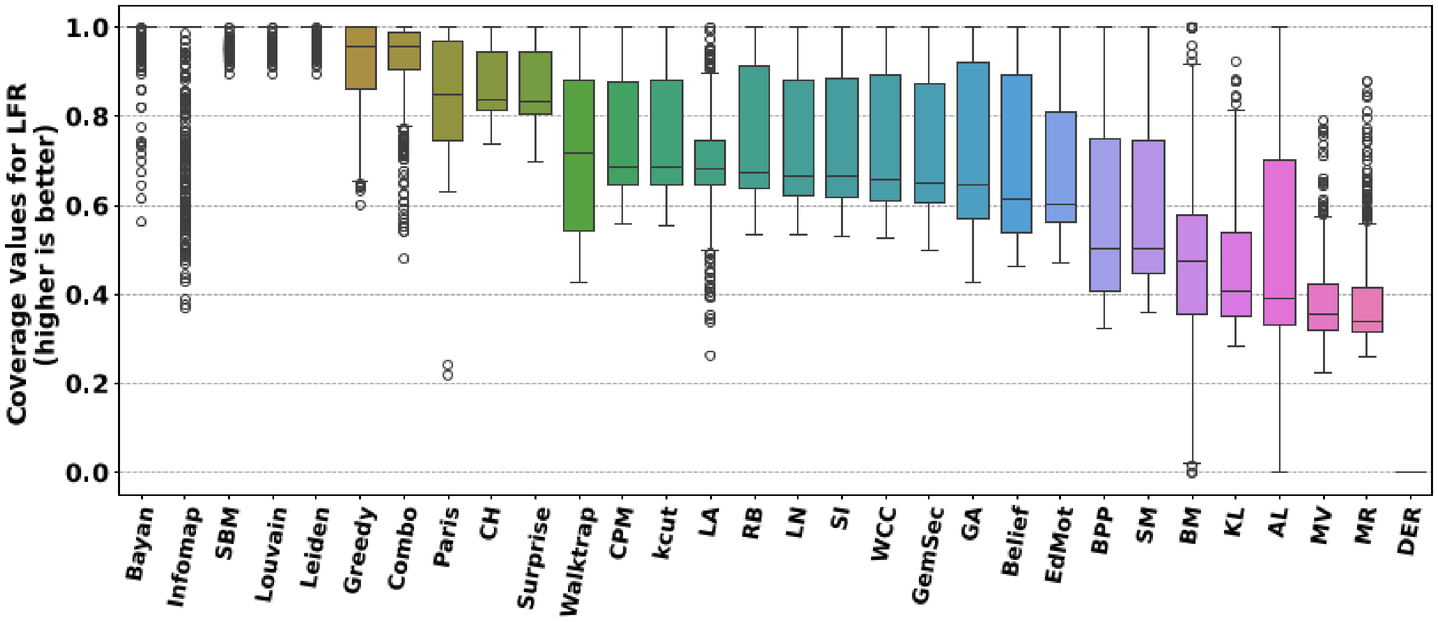
Unsupervised assessment: average conductance

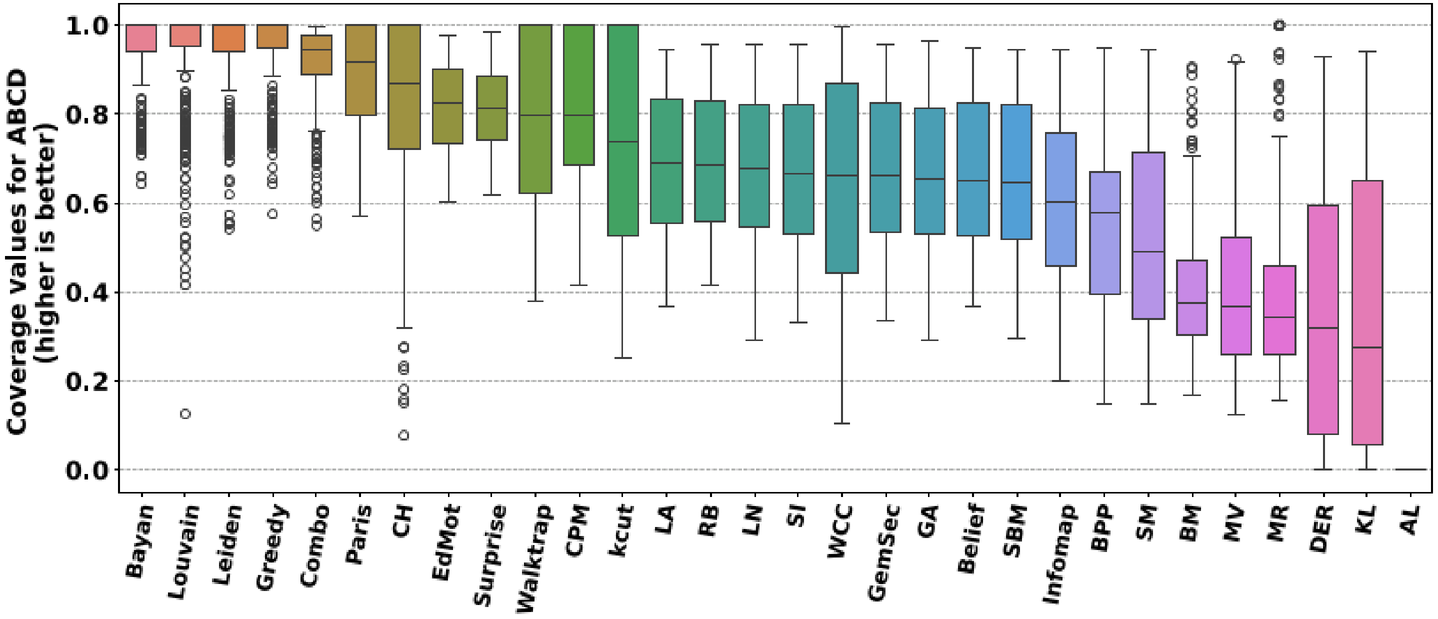




UNIVERSITY OF TORONTO

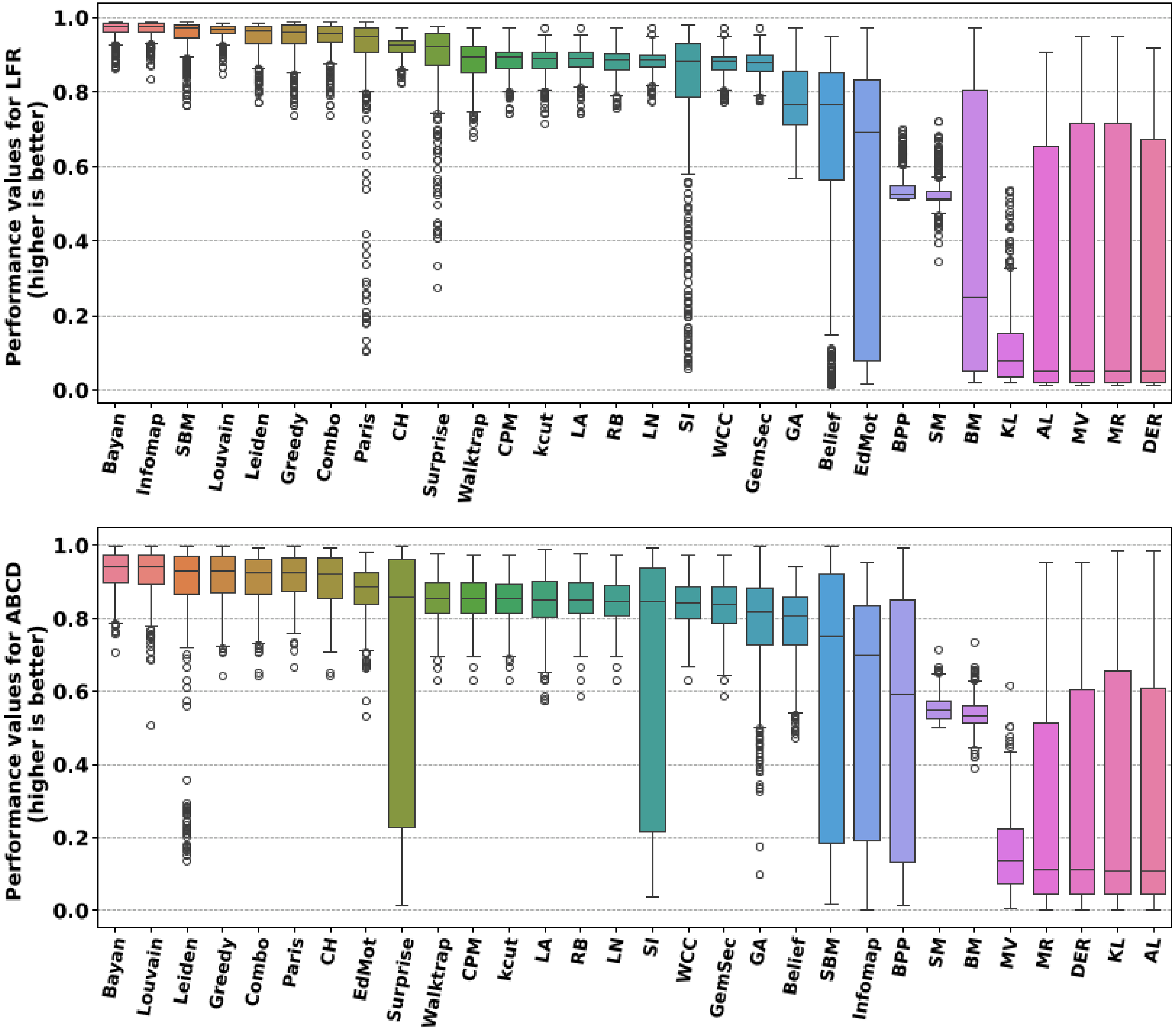
Unsupervised assessment: partition coverage





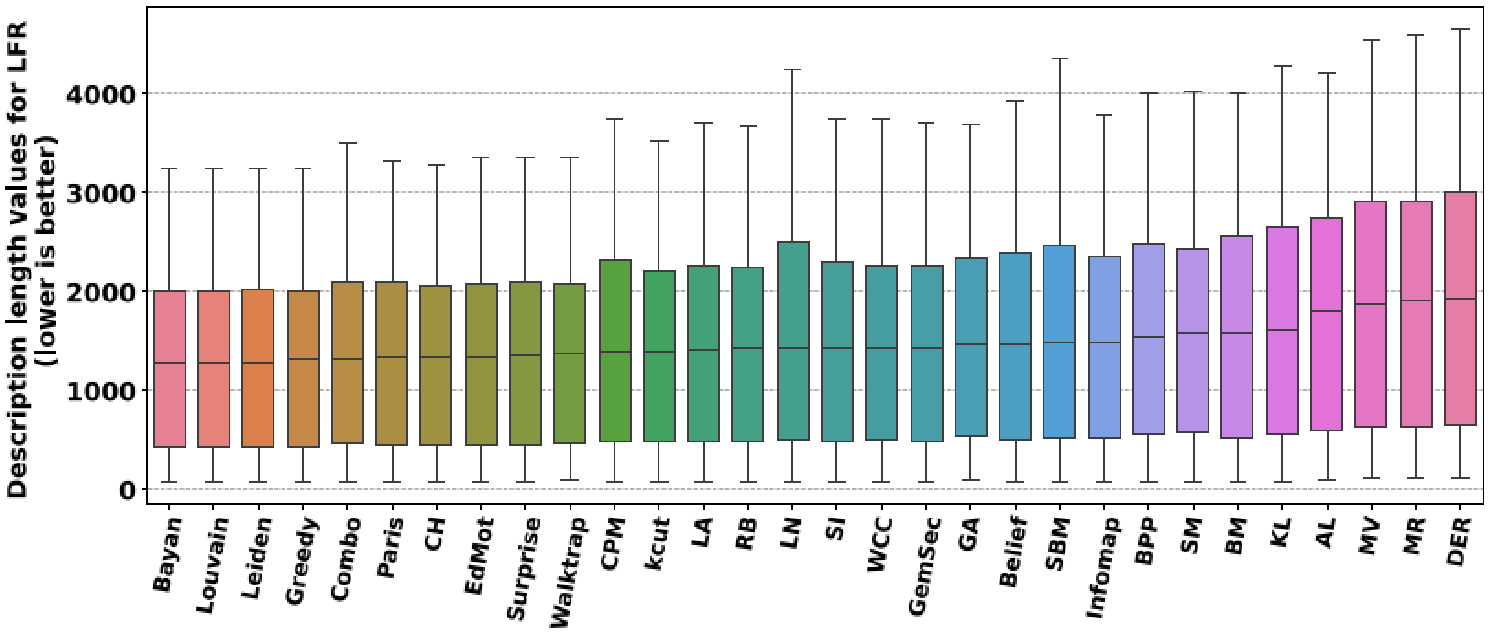
UNIVERSITY OF TORONTO

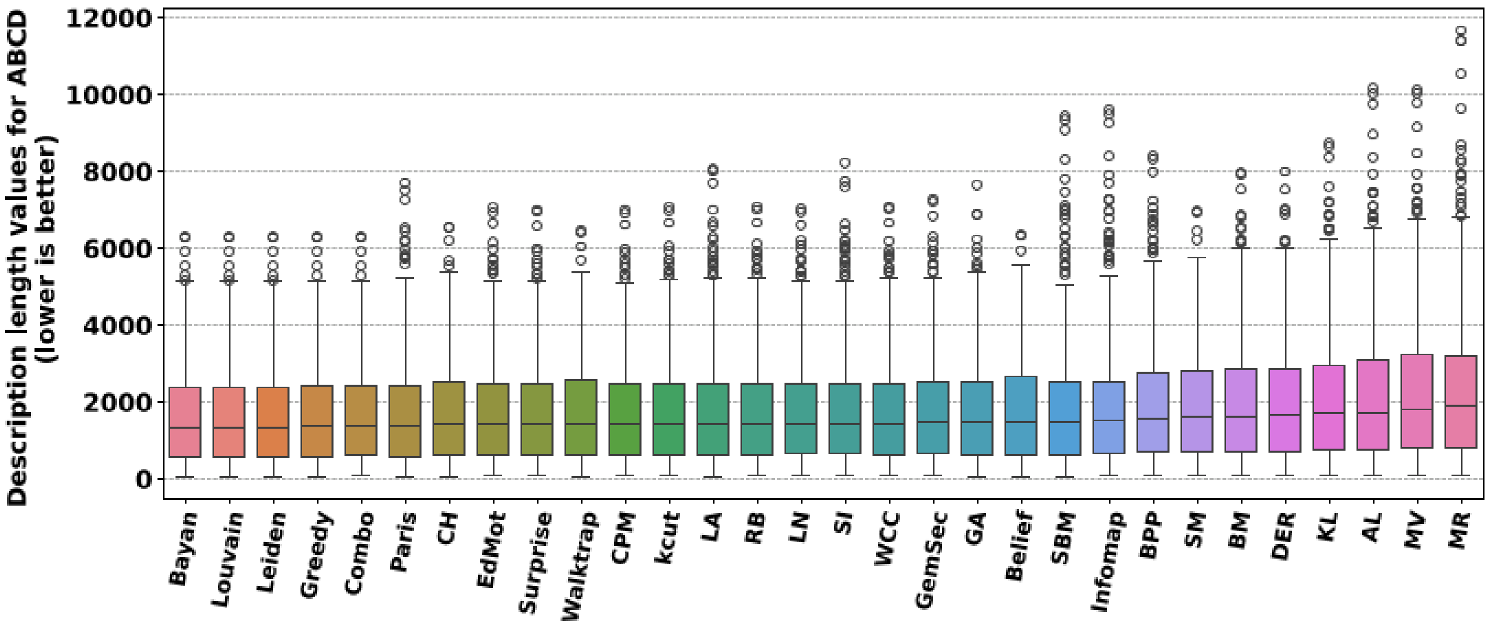
Unsupervised assessment: partition performance



UNIVERSITY OF TORONTO

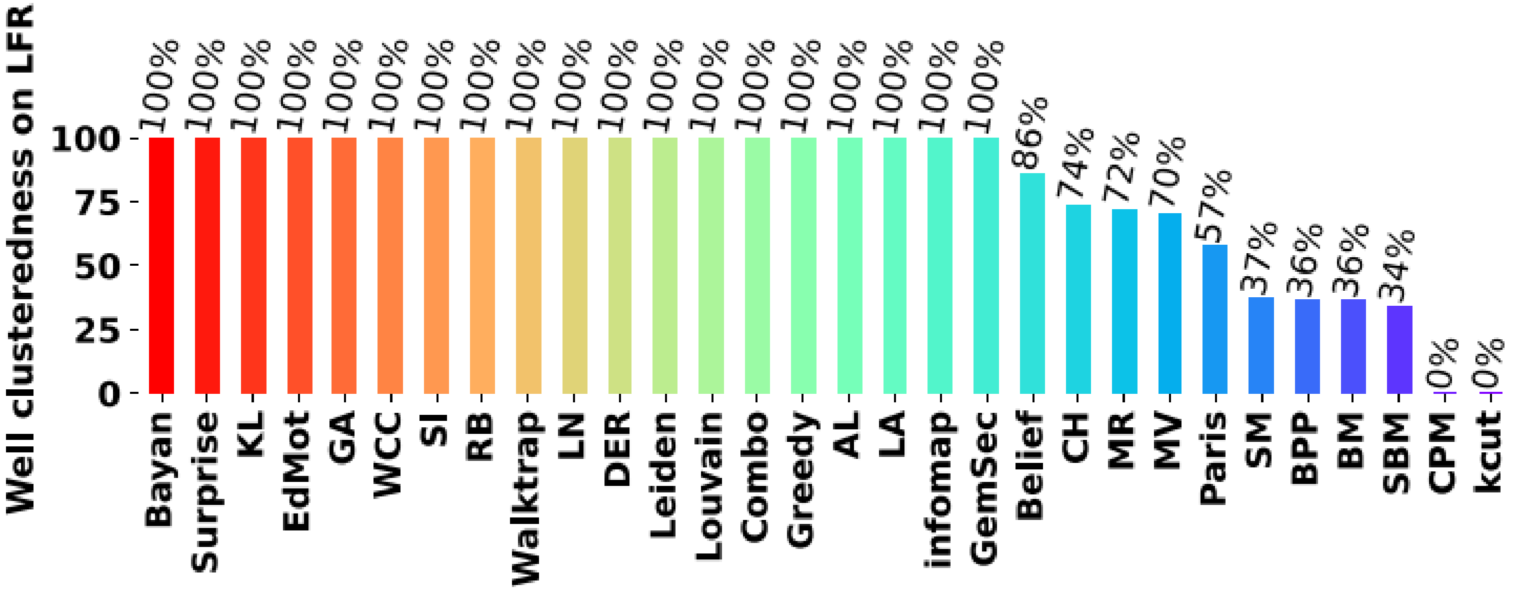
Unsupervised assessment: description length

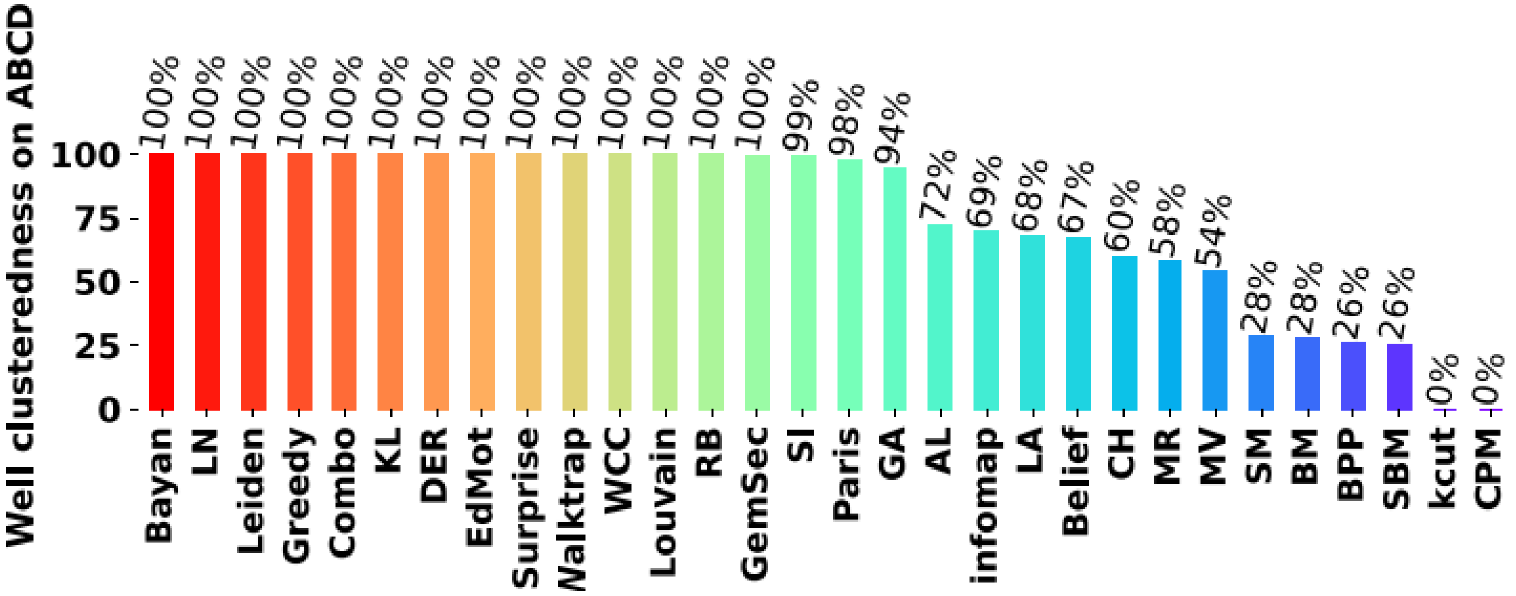




UNIVERSITY OF TORONTO

Unsupervised assessment: well-clusteredness





UNIVERSITY OF TORONTO

Resources on the Bayan algorithm

```
%pip install bayanpy
```

```
import networkx as nx
import bayanpy
```

```
G = nx.barbell_graph(5,2)
```

```
bayanpy.bayan(G)
```

Paper: doi.org/10.1103/PhysRevE.110.044315
GitHub Repo: github.com/saref/bayan
Project website: bayanproject.github.io

Google Colab examples:
tinyurl.com/bayancolab

Peer-reviewed papers → ArXiv.org
Models and code → github.com/saref
Network data → FigShare and OSF saref.github.io

UNIVERSITY OF TORONTO

Papers presented

Measuring balance arxiv.org/abs/1509.04037

Computing frustration index arxiv.org/abs/1710.09876

Frustration index of large networks arxiv.org/abs/1611.09030

Applications of frustration and partitioning arxiv.org/abs/1712.04628

Multilevel evaluation of balance arxiv.org/abs/2005.09925

Partitioning dense signed networks of the US Congress arxiv.org/abs/1906.01696

Partitioning signed networks based on generalized balance arxiv.org/abs/2105.01913

Modularity-based heuristics arxiv.org/abs/2302.14698

Analyzing modularity maximization arxiv.org/abs/2310.10898

Bayan algorithm arxiv.org/abs/2209.04562

Troika algorithm arxiv.org/abs/2505.03573

UNIVERSITY OF TORONTO

Special thanks to my collaborators:

Mark C. Wilson
Information and Computer Sciences
University of Massachusetts

Andrew J. Mason
Engineering Science
University of Auckland

Zachary Neal
Department of Psychology
Michigan State University

Hriday Chheda
Mechanical and Industrial Engineering
University of Toronto

Boris Ng
Mechanical and Industrial Engineering
University of Toronto

Ly Dinh
School of Information
University of South Florida

Rezvaneh (Shadi) Rezapour
Information Science
Drexel University

Jana Diesner
School of Social Science and Technology
Technical University of Munich

Mahdi Mostajabdave
Mathematical and Industrial Engineering
Polytechnique Montreal

UNIVERSITY OF
TORONTO

Abelson RP, Rosenberg MJ (1958) Symbolic psycho-logic: A model of attitudinal cognition. Behavioral Science 3(1):1–13.

Aref S, Neal ZP (2020) Detecting coalitions by optimally partitioning signed networks of political collaboration. Scientific Reports, 10(1):1–10.

Aref S, Mason AJ, Wilson MC (2018) Computing the line index of balance using integer programming optimisation in Optimization Problems in Graph Theory, ed. Goldengorin B. (Springer), pp. 65–84.

Aref S, Mason AJ, Wilson MC (2020) A Modeling and Computational Study of the Frustration Index in Signed Networks. Networks, 75(1):95-110.

Aref S, Wilson MC (2016) Measuring partial balance in signed networks. Journal of Complex Networks 6(4):566–595.

Aref S, Wilson MC (2019) Balance and frustration in signed networks. Journal of Complex Networks 7(2):163–189.

Aref S, Dinh, L., Rezapour, R., & Diesner, J. (2020). Multilevel structural evaluation of signed directed social networks based on balance theory. Scientific reports, 10(1), 1-12.

Aref S, Neal ZP (2021). Identifying Hidden coalitionss in the US House of Representatives by optimally partitioning signed networks based on generalized balance. Scientific reports, 11(1), 1-9.

Cartwright D, Harary F (1956) Structural balance: a generalization of Heider's theory. Psychological Review 63(5):277–293

Cartwright D, Harary F (1976) Balance and clusterability: An overview. In P. W. Holland and S. Leinhardt, editors, Social networks: Surveys, advances, and commentaries, pages 25–50. Academic Press.

Davis, JA (1967) Clustering and structural balance in graphs. Human Relations 20 (2): 181–187.

Demaine, E. D., Emanuel, D., Fiat, A. & Immorlica, N. Correlation clustering in general weighted graphs. Theoret. Comput. Sci. 361, 172–187 (2006).

Doreian P, Mrvar A (1996) A partitioning approach to structural balance. Social networks. 1;18(2):149-68.

Harary F (1959) On the measurement of structural balance. Behavioral Science 4(4):316–323.

Heider F (1944) Social perception and phenomenal causality. Psychological Review 51(6):358–378.

Neal ZP (2014) The backbone of bipartite projections: Inferring relationships from coauthorship,co-sponsorship, co-attendance and other co-behaviors. Social Networks 39:84–97.

Neal ZP (2020). A sign of the times? Weak and strong polarization in the US Congress, 1973–2016. Social Networks, 60, 103-112.

Terzi E, Winkler M (2011) A spectral algorithm for computing social balance in Proceedings of International Workshop on Algorithms and Models for the Web-Graph, WAW 2011, eds. Frieze A, Horn P, Prałat P. (Springer), pp. 1–13.

Zaslavsky T (1987) Balanced decompositions of a signed graph. Journal of Combinatorial Theory, Series B 43(1):1–13.

Additional references

Dinh, T.N., Thai, M.T.: Toward optimal community detection: From trees to general weighted networks. Internet Mathematics 11(3), 181–200 (2015)

Clauset, A., Newman, M.E., Moore, C.: Finding community structure in very large networks. Physical review E 70(6), 066111 (2004)

Biemann, C.: Chinese whispers: an efficient graph clustering algorithm and its application to natural language processing problems. In: Proceedings of the First Workshop on Graph Based Methods for Natural Language Processing. TextGraphs-1, pp. 73–80. (2006)

Reichardt, J., Bornholdt, S.: Statistical mechanics of community detection. Physical Review E 74(1), 016110 (2006).

Pons, P., Latapy, M.: Computing communities in large networks using random walks. J. Graph Algorithms Appl. 10(2), 191–218 (2006)

Ruan, J., Zhang, W.: An efficient spectral algorithm for network community discovery and its applications to biological and social networks. In: Seventh IEEE International Conference on Data Mining (ICDM 2007), pp. 643–648. IEEE, (2007)

Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment 2008(10), 10008 (2008).

Rosvall, M., Bergstrom, C.T.: Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences 105(4), 1118–1123 (2008).

Leicht, E.A., Newman, M.E.J.: Community structure in directed networks. Physical Review Letters 100(11), 118703 (2008).

Pizzuti, C.: GA-Net: A Genetic Algorithm for Community Detection in Social Networks. In: Proceedings of the 10th International Conference on Parallel Problem Solving from Nature — PPSN X - Volume 5199, pp. 1081–1090. Springer, Berlin, Heidelberg (2008)

Cordasco, G., Gargano, L.: Community detection via semi-synchronous label propagation algorithms. In: 2010 IEEE International Workshop On: Business Applications of Social Network Analysis (BASNA), pp. 1–8. IEEE, (2010)

Traag, V.A., Van Dooren, P., Nesterov, Y.: Narrow scope for resolution-limit-free community detection. Physical Review E 84(1), 016114 (2011)

Traag, V.A., Krings, G., Van Dooren, P.: Significant scales in community structure. Scientific reports 3(1), 1–10 (2013)

Peixoto, T.P.: Efficient monte carlo and greedy heuristic for the inference of stochastic block models. Physical Review E 89(1), 012804 (2014)

Prat-P´erez, A., Dominguez-Sal, D., Larriba-Pey, J.-L.: High quality, scalable and parallel community detection for large real graphs. In: Proceedings of the 23rd International Conference on World Wide Web, pp. 225–236 (2014)

Sobolevsky, S., Campari, R., Belyi, A., Ratti, C.: General optimization technique for high-quality community detection in complex networks. Physical Review E 90(1), 012811 (2014)

Zhang, P., Moore, C.: Scalable detection of statistically significant communities and hierarchies, using message passing for modularity. Proceedings of the National Academy of Sciences 111(51), 18144–18149 (2014)

Traag, V.A., Aldecoa, R., Delvenne, J.-C.: Detecting communities using asymptotical surprise. Physical Review E 92(2), 022816 (2015)

Bonald, T., Charpentier, B., Galland, A., Hollocou, A.: Hierarchical graph clustering using node pair sampling. In: MLG 2018 - 14th International Workshop on Mining and Learning with Graphs, London, UK (2018)

Traag, V.A., Waltman, L., van Eck, N.J.: From Louvain to Leiden: guaranteeing well-connected communities. Scientific Reports 9(1) (2019).

Li, P.-Z., Huang, L., Wang, C.-D., Lai, J.-H.: EdMot: An edge enhancement approach for motif-aware community detection. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 479–487 (2019)

Rozemberczki, B., Davies, R., Sarkar, R., Sutton, C.: Gemsec: Graph embedding with self clustering. In: Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 65–72 (2019)

Questions?
Thank you!

s.aref@utoronto.ca
@SaminAref
saref.github.io

UNIVERSITY OF TORONTO

Slides are available at https://saref.github.io/presentation/aref-netsci2025-talk.pdf