Node Metrics in Complex Networks: Foundations, Perspectives and Applications

Sergey Shvydun

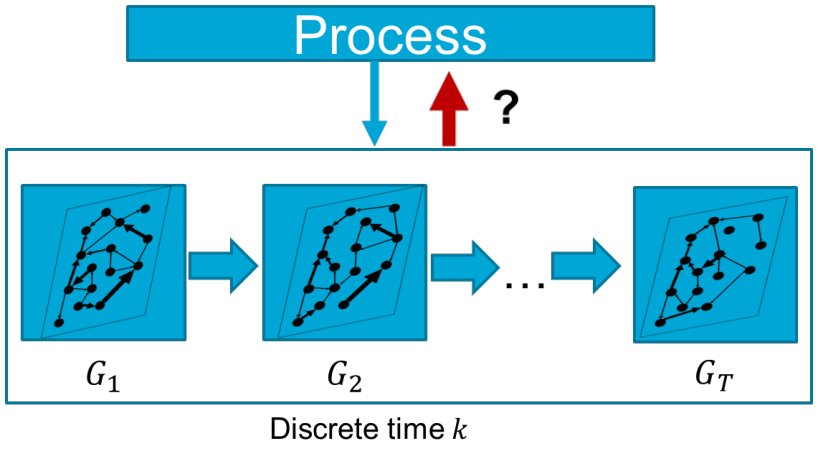
TU Delft NAS

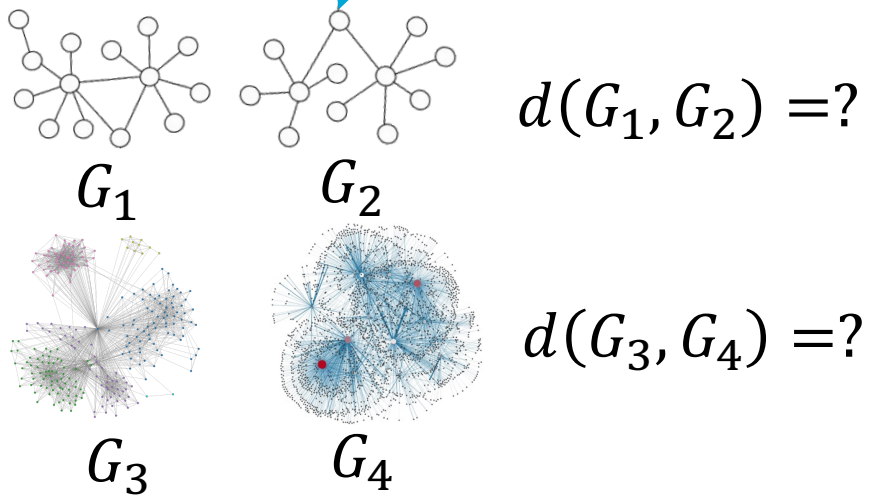
Graphs&Data@TUDelft seminar
06/11/2025

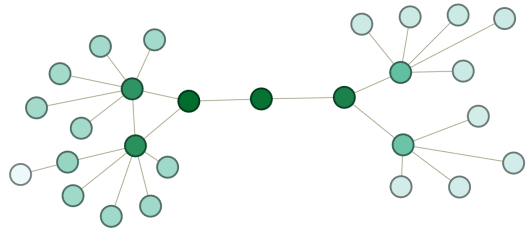
About me

- Postdoc at the Network Architectures and Services (NAS) group, TU Delft

Research interests: machine learning, **network science**, social choice theory.







1) Temporal networks

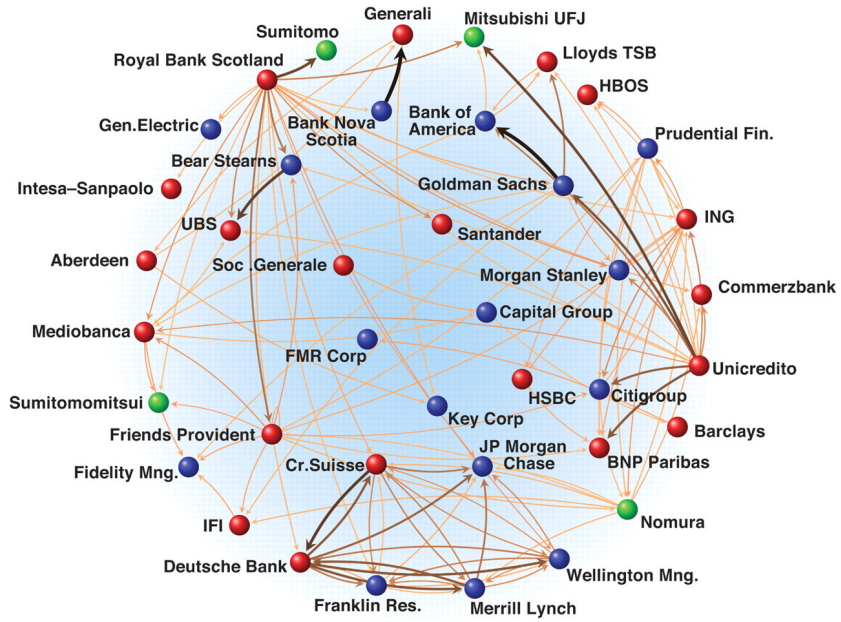
2) Graph similarity

3) Node Metrics

Research Area

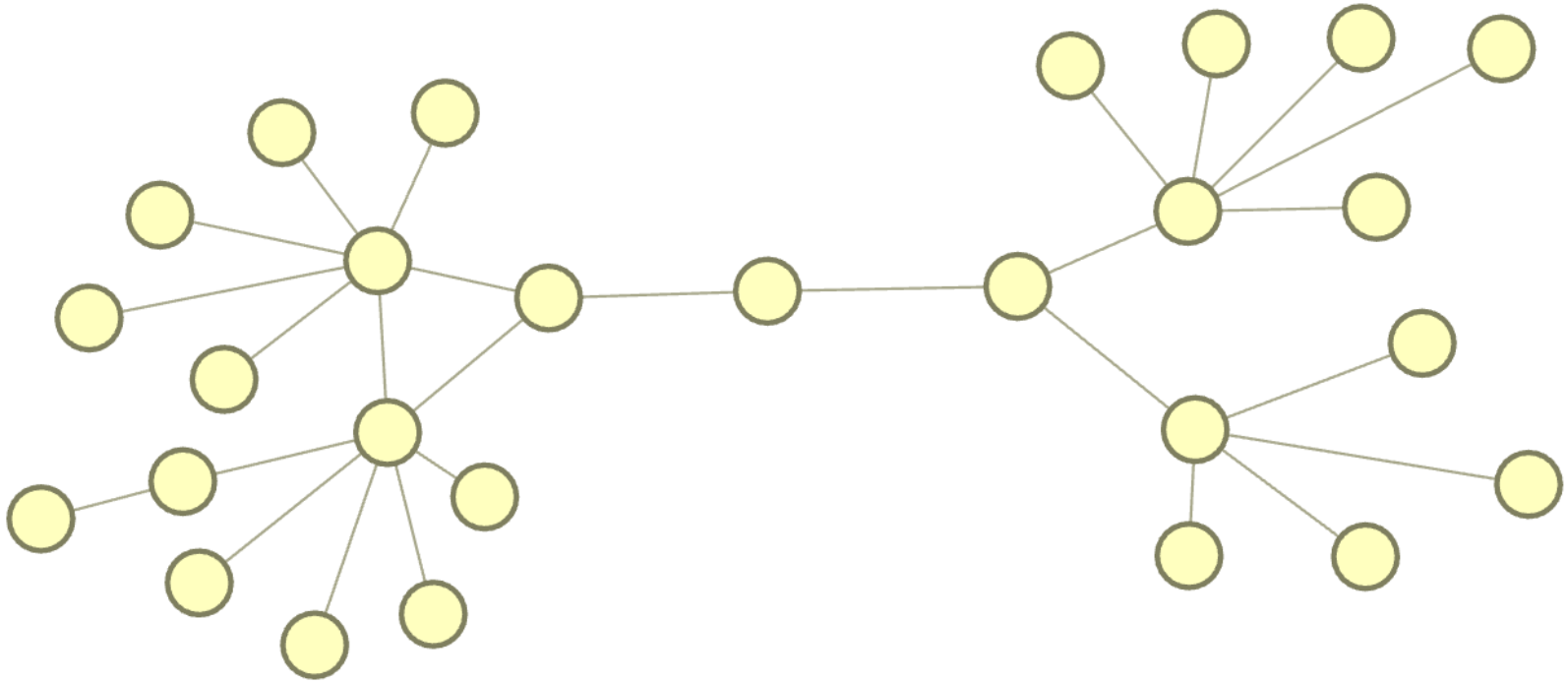
Consider a graph $G = G(\mathcal{N}, \mathcal{L})$ with N nodes and L links:

- $\mathcal{N}$ is a set of nodes (people, financial institutes, countries, cities, computers, etc.);
- $\mathcal{L}$ is a set of edges (relationships: roads, contacts, flows, calls, etc.).



Problem Statement

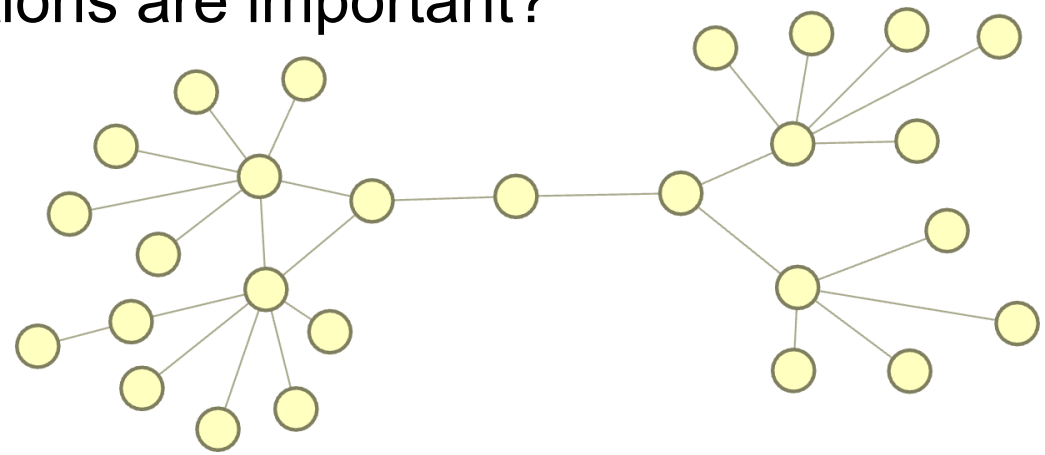
Consider a graph $G = G(N, L)$ with N nodes and L links.



- How can we count and classify node positions in different graph topologies?
- Which nodes are the most important in the network?

Node Metrics: Applications

Which these questions are important?

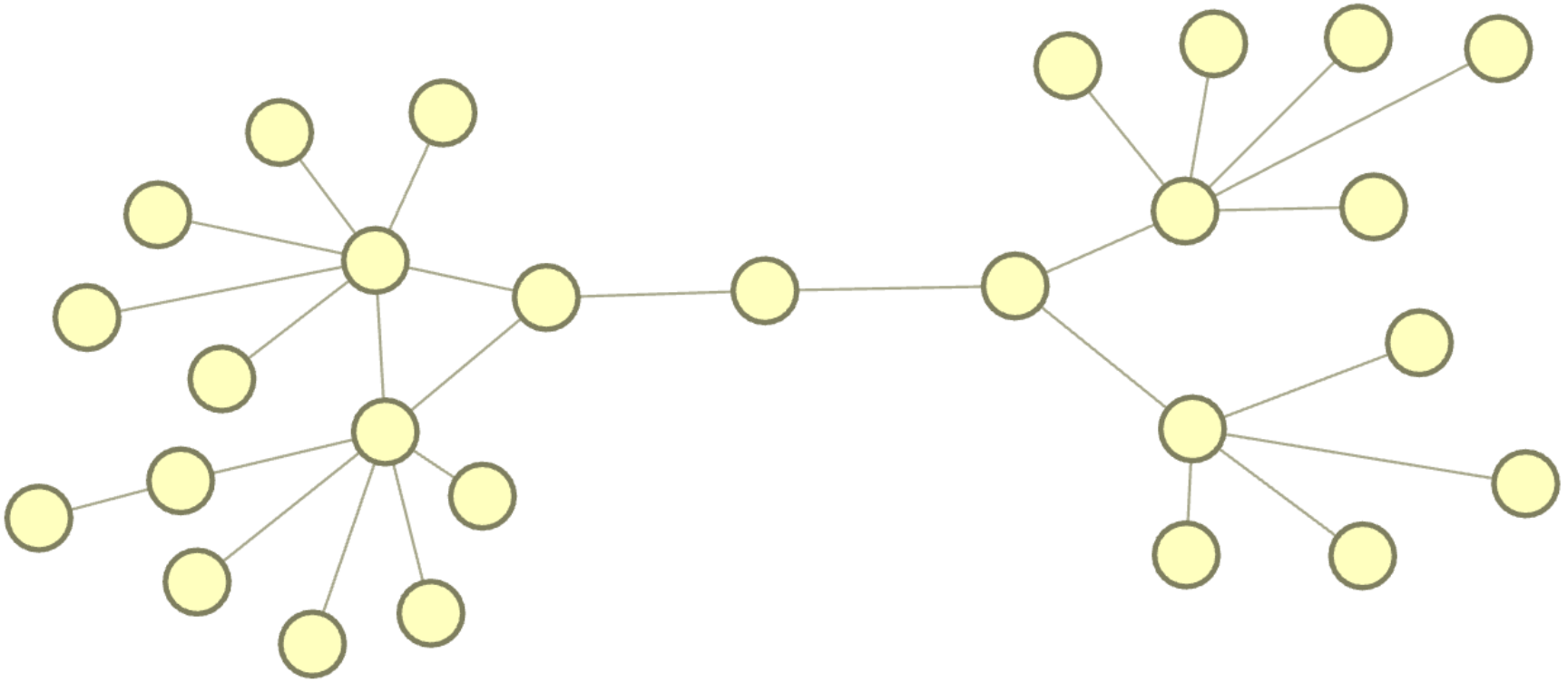


Applications:

- Social networks: communication patterns, leadership roles and career success;
- Marketing: influencers and opinion leaders;
- Network robustness and resilience: critical infrastructure;
- Transportation: optimizing traffic flow;
- Epidemiology: “super-spreaders“;
- Economics: systemically important financial institutions;
- Science of science: influential papers/researchers/research groups;
- Graph Machine learning: link & node prediction.

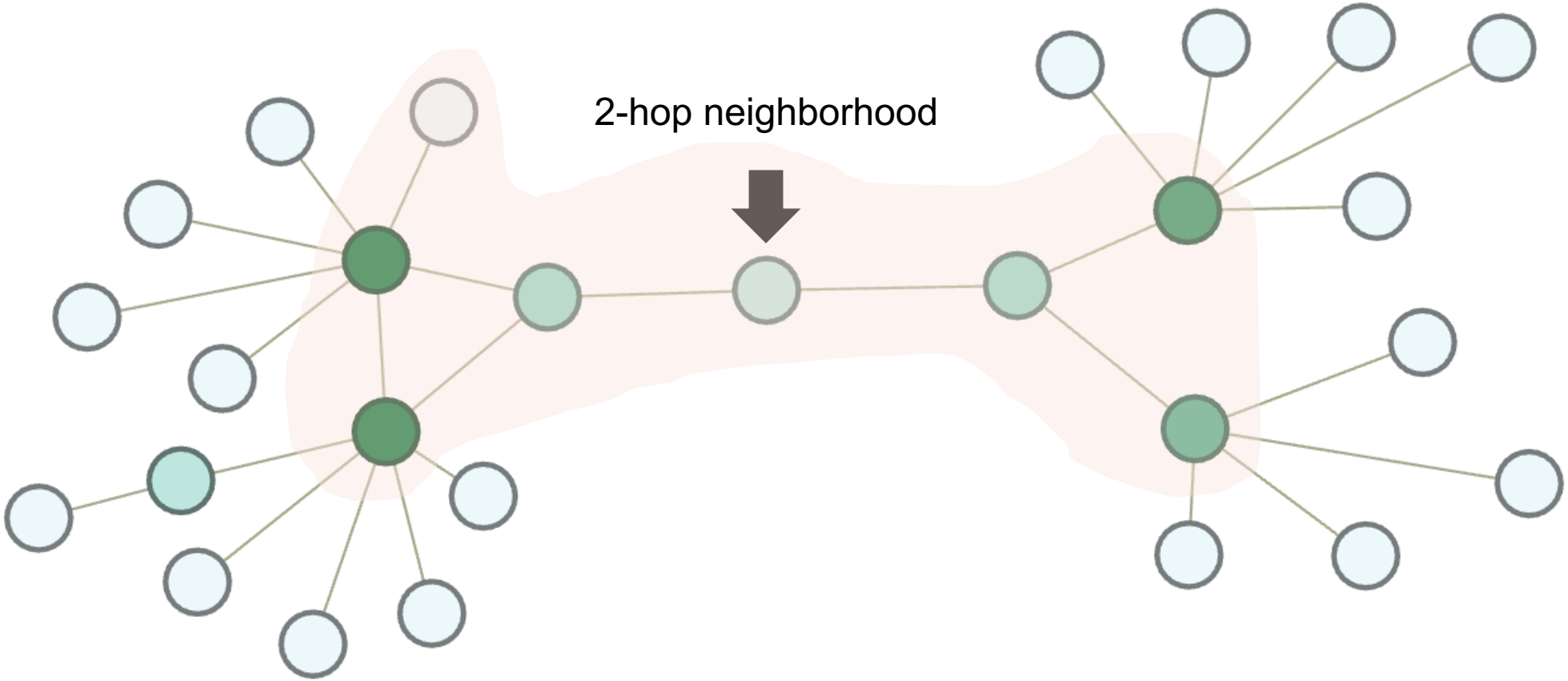
Node Metrics: Applications

Which nodes are the most important in the network?



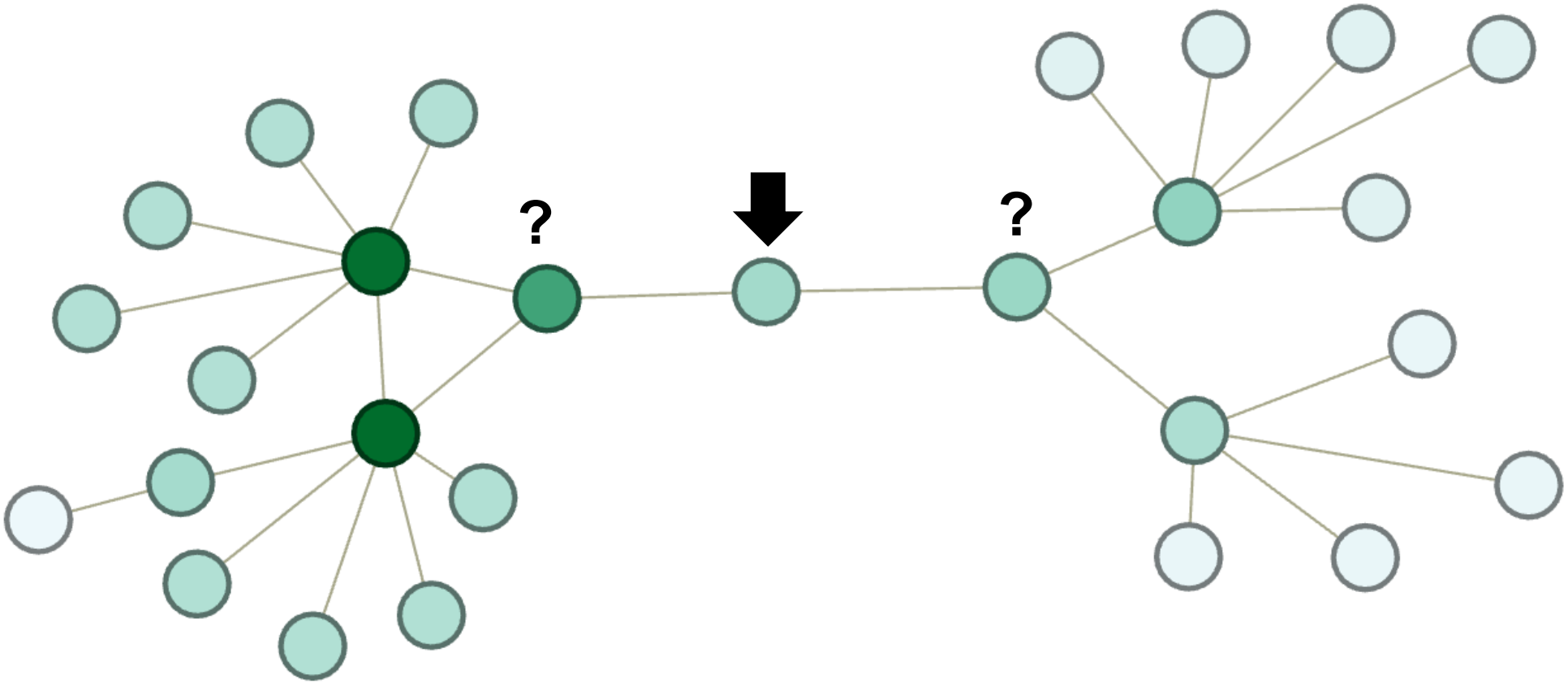
No unique solutions (it depends on the type of network and the process)!

Examples: Local/Semi-local measures



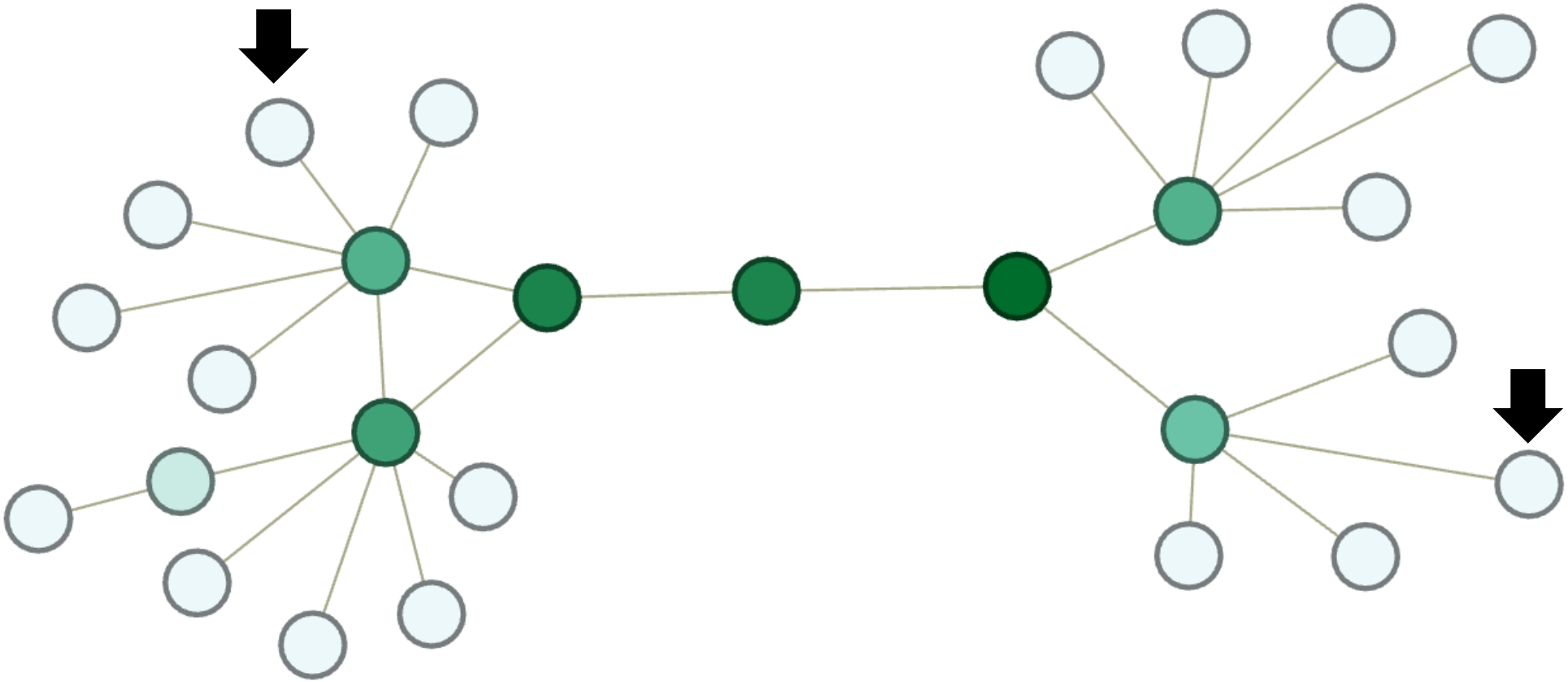
- **Local measures:** how many neighbors do I have?
  Examples: node degree, 2-hop degree.

Examples: Spectral measures



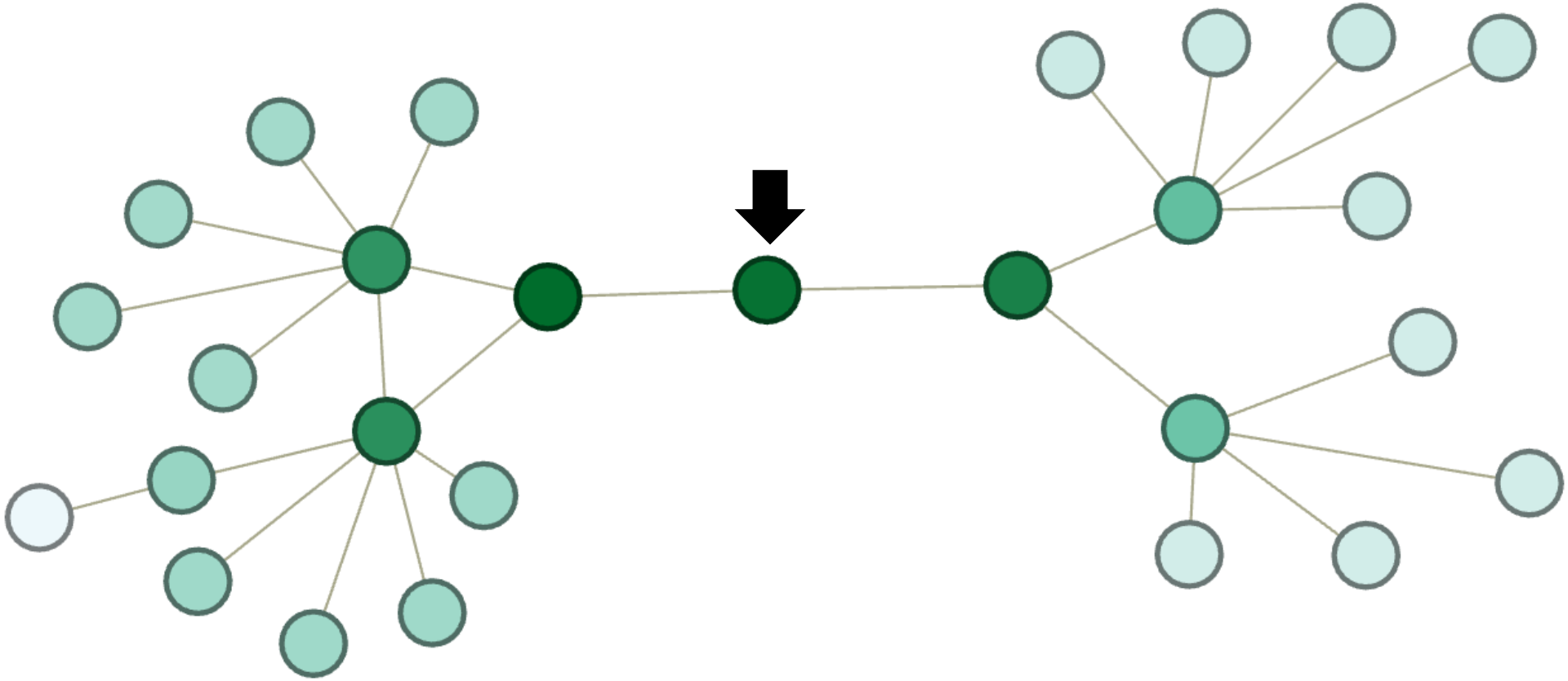
- **Spectral measures:** my importance depends on the importance of my neighbors. *Examples*: eigenvector, Katz and PageRank centralities.

Examples: betweenness-based measures



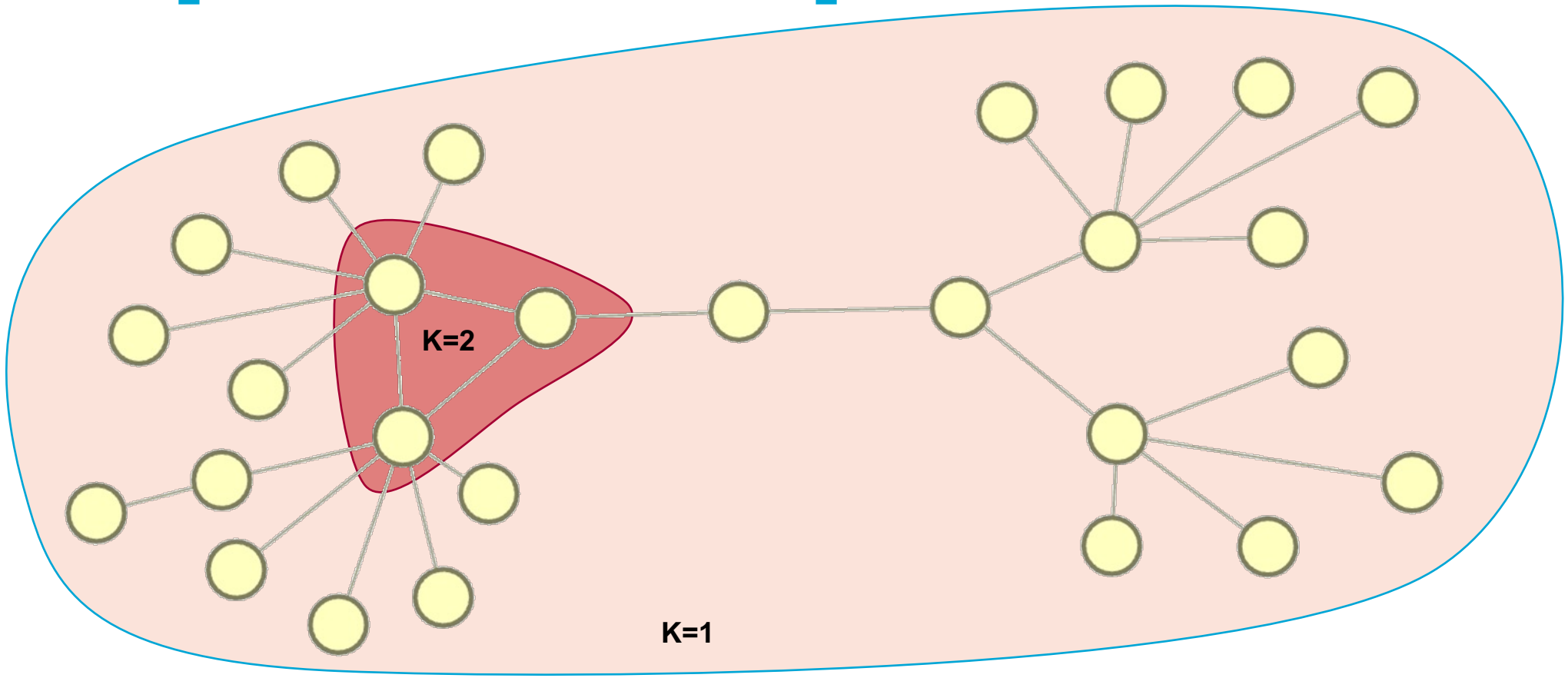
- **Betweenness-based measures:** how many paths go through me?
  Examples: stress, betweenness centrality.

Examples: closeness-based measures



- **Closeness-based measures:** how close am I to all other nodes? Examples: closeness centrality, harmonic centrality.

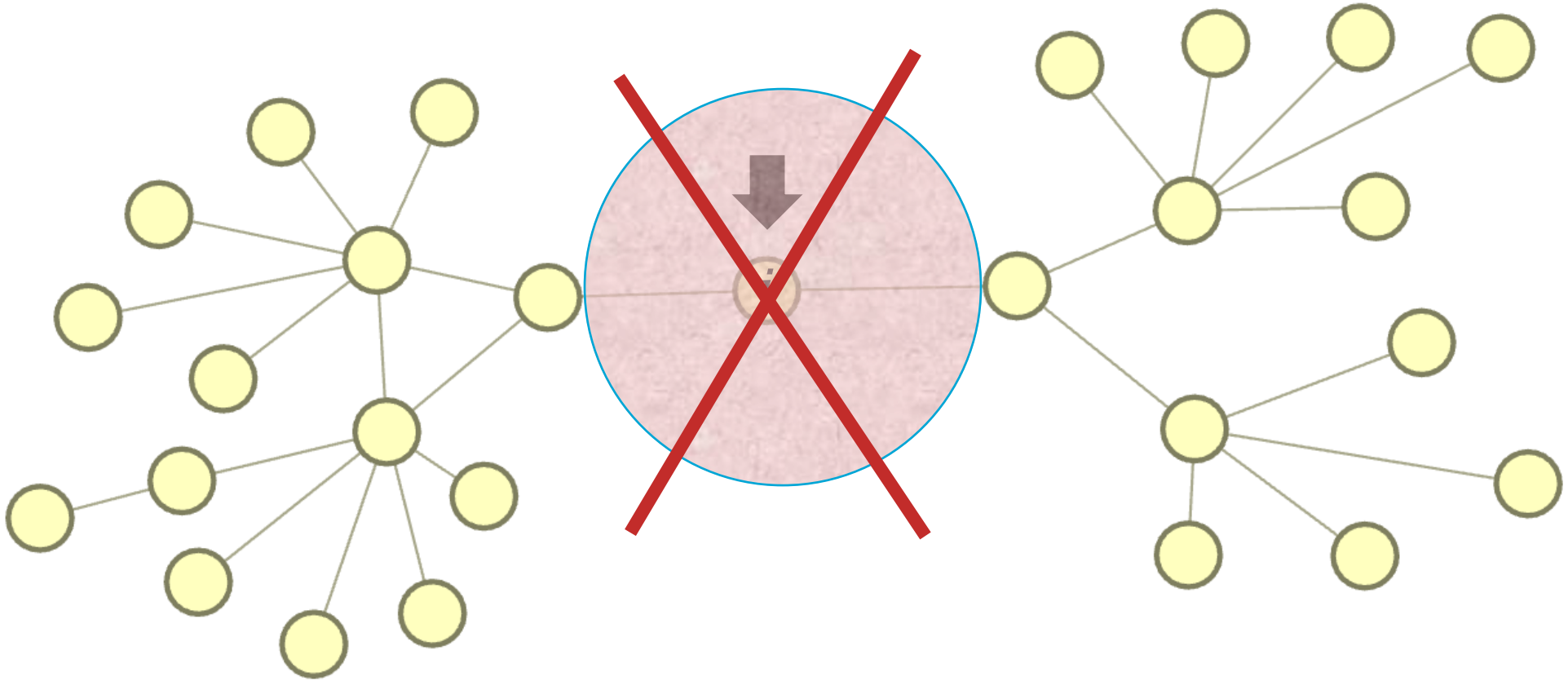
Examples: core decomposition measures



- **K-shell centralities:** network decomposition

  Examples: k-shell centrality, weighted k-shell decomposition.

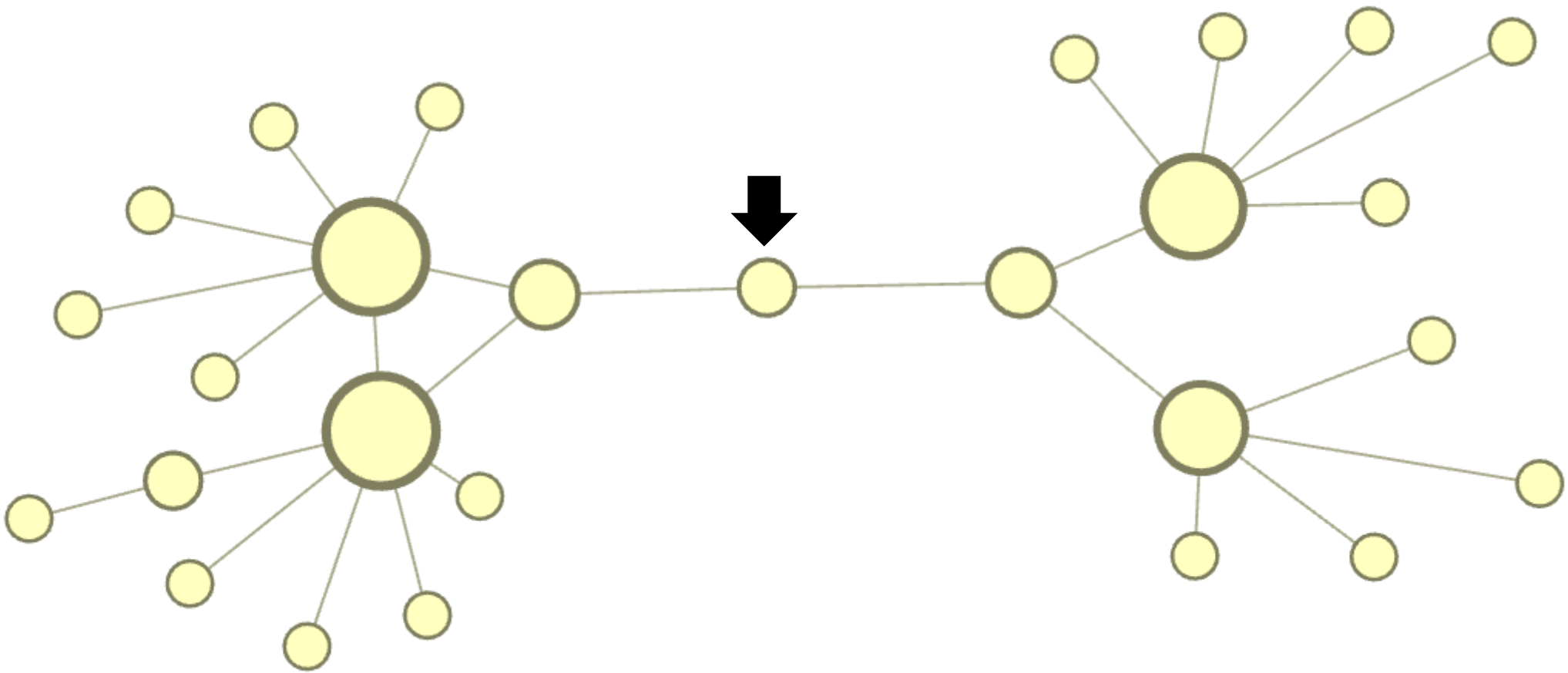
Examples: vitality-based measures



- **Vitality-based measures:**

$$c(i) = f(G) - f(G_i)$$

Examples: gravity-based measures



- **Gravity models:** $c(i) = \sum_{j=1}^{N} \frac{m_i m_j}{d_{ij}^2}$ (m_i is the mass of node i)

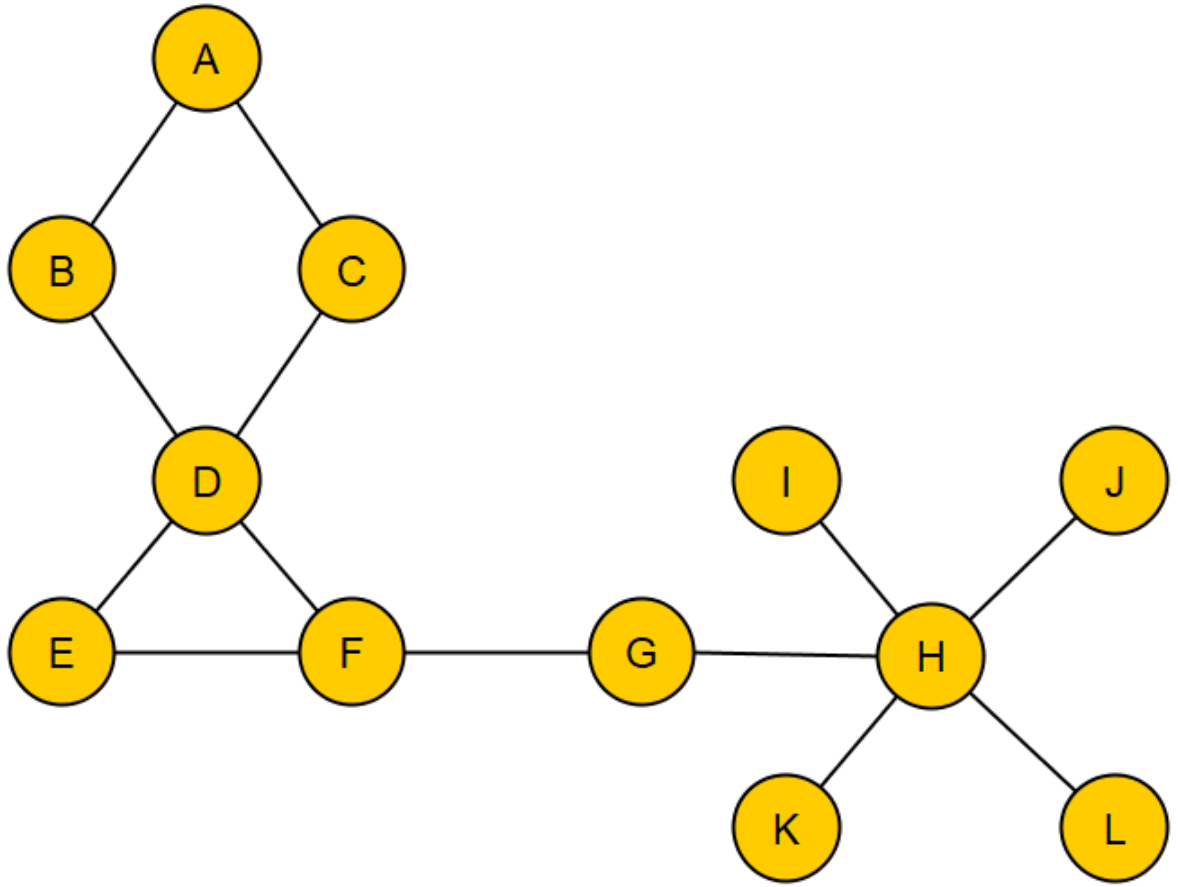
Classical measures for nodes

1. Degree (node H)
   (How many neighbors do I have?)

2. Eigenvector (node D)
   (My importance depends on the importance of my neighbors, etc.)

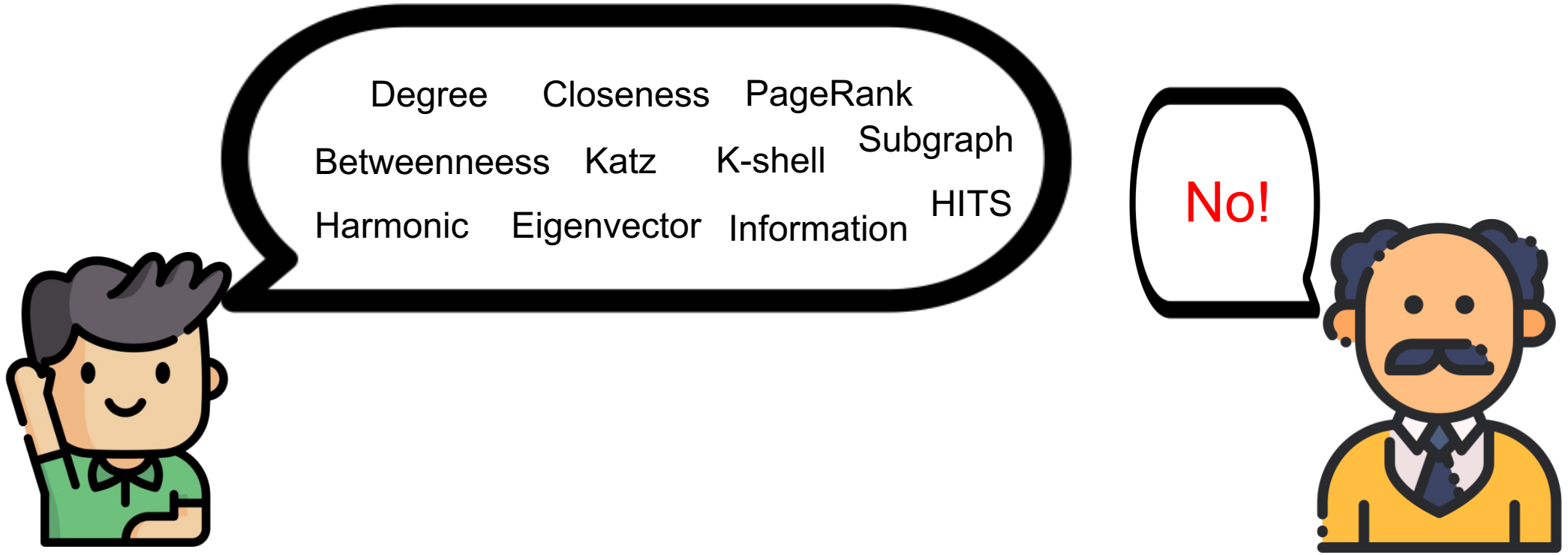
3. Betweenness (node H)
   (How many shortest paths go through me?)

4. Closeness (nodes F, G)
   (How close am I to all other nodes?)



Node metrics: Problems and Challenges

The number of proposed metrics over the graph history is **overwhelming!**



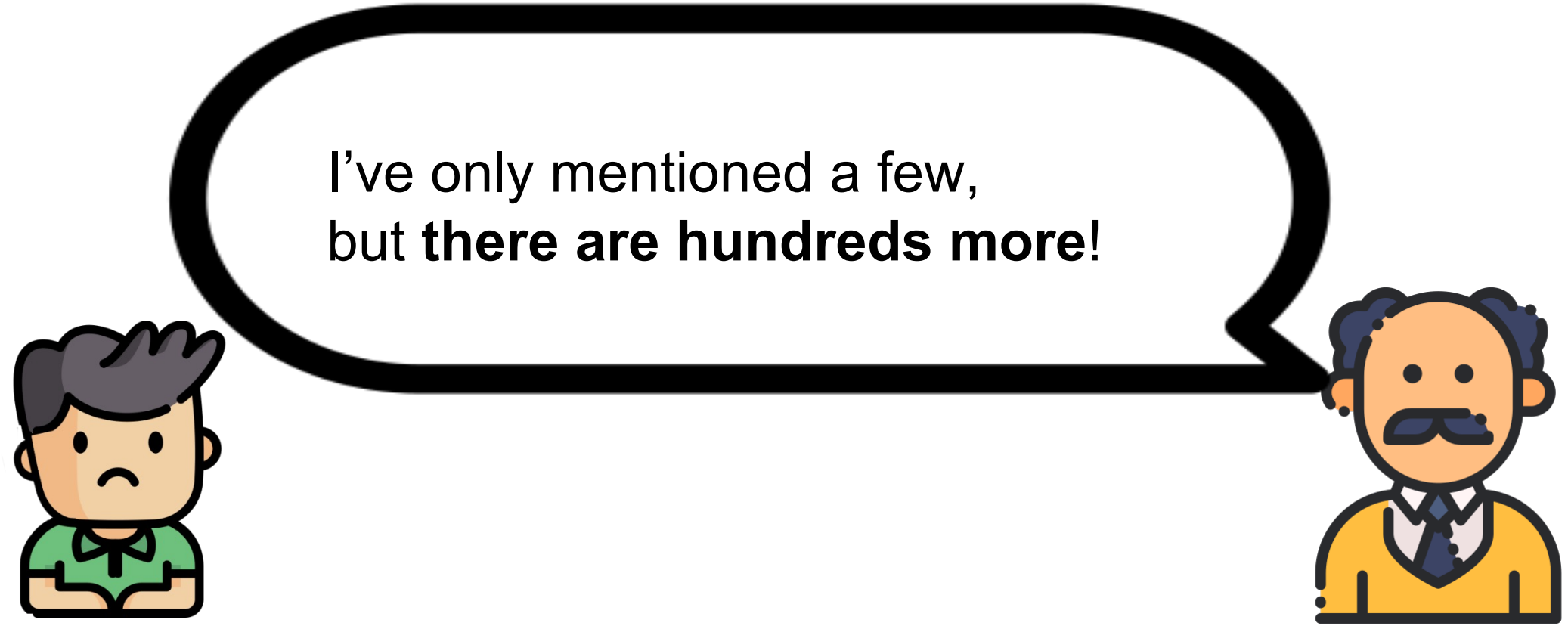
Node metrics: Problems and Challenges

The number of proposed metrics over the graph history is **overwhelming!**

Degree Eigenvector Katz Hubbel Betweenness Closeness Harmonic Residual closeness Decay d-closeness Current-flow closeness Flow betweenness Current-flow betweenness d-betweenness Subgraph PageRank LeaderRank SRIC LRIC LRIC-sim Interdependence Laplacian Percolation k-shell Collective Influence BottleNeck Centroid Closeness vitality ClusterRank Communicability betweenness Cross-Clique Connectivity Diffusion Degree MNC DMNC Load m-reach k-path Geodesic k-path Local reaching Second order Trophic level VoteRank Entropy Degree mass EPC Leverage Lin's index Lobby index Markov Pairwise disconnectivity Radiality SALSA LocalRank Topological Neighborhood connectivity Local H-index CON Score Eccentricity Stress Dynamical influence Edge-disjoint k-path Vertex-disjoint k-path k-betweenness Distance-weighted fragmentation Eigentrust AIC Nieminen's closeness e-betweenness WKPaths LineRank Seeley's index BG-index MCC Burt's constraint Community centrality Bridging Coefficient Bridging Centrality Egocentric betweenness Localized bridging centrality (LBC) Extended LBC Effective size Shapley value Local clustering coefficient Redundancy

Node metrics: Problems and Challenges

The number of proposed metrics over the graph history is **overwhelming!**



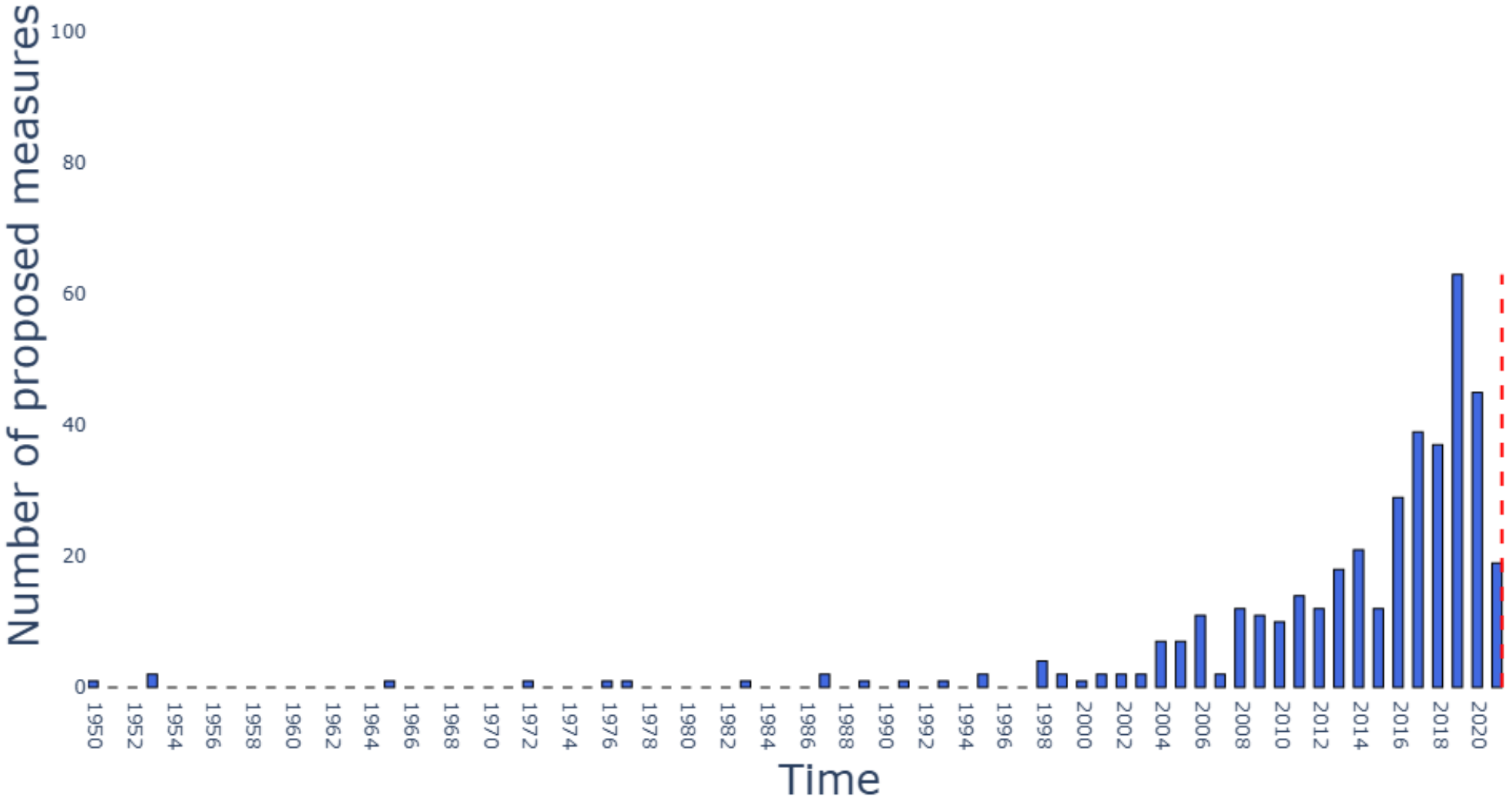
Node metrics: Problems and Challenges

The number of proposed metrics over the graph history is **overwhelming!**

1. Centiserver (https://www.centiserver.ir/): 403 measures

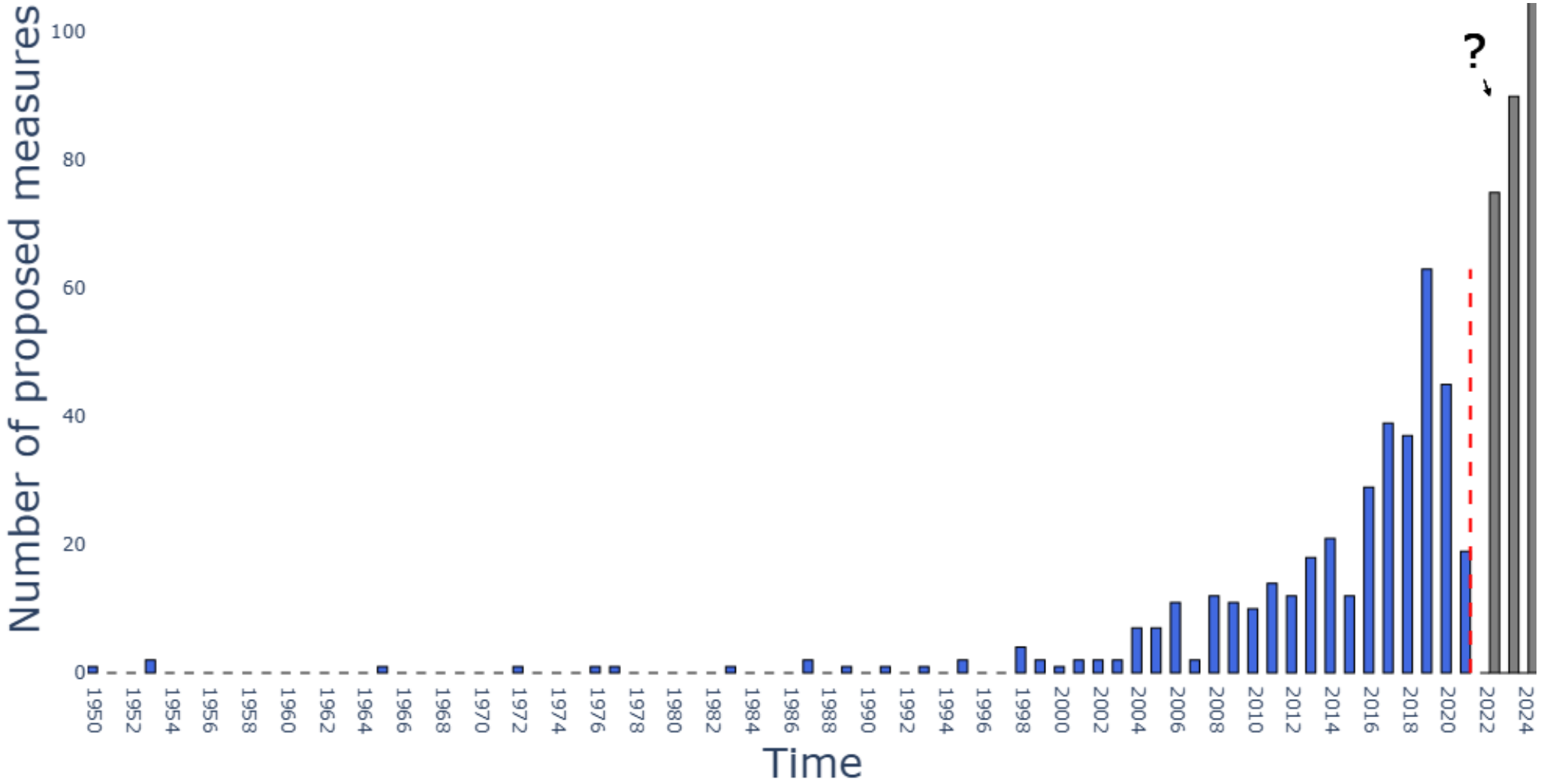
- Data reflects information available up to 2021.
- The website (list of measures) contains inaccuracies.

Evolution of Node Metrics



Source: Centiserver (site not updated after 2021)

Evolution of Node Metrics



Source: Centiserver (site not updated after 2021)

Node Metrics: Problems

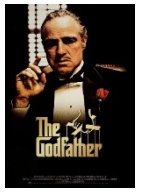
The number of proposed metrics is overwhelming.

Problem 1. Most existing measures remain unknown.

- Hard to find these measures;
- Most extensive literature reviews: ≈ 50-80 measures.

There are promising measures based on ideas from

- Information theory (entropy-based measures),
- Cooperative game theory (game centrality),
- Voting theory (VoteRank),
- Dempster-Shafer evidence theory (new evidential centrality),
- Multi-criteria decision-making (TOPSIS),
- Signal processing (graph Fourier transform),
- Physics (e.g., gravity centrality),
- Biology (e.g., physarum centrality),
- Geometry (e.g., curvature index)
- **and many other fields**
- Godfather Index (by Matthew O. Jackson)



Node Metrics: Problems

Problem 2. Access to many centrality models is limited.

Existing packages:

- Igraph
- NetworkX
- CentiLib
- CentiBiN
- graph-tool
- Neo4j
- Pajek
- Sna
- CytoNCA
- Gephi
- tidygraph
- ...

The most comprehensive libraries contain no more than 40 measures.

Node Metrics: Problems

Problem 3. Duplication: many models are being reinvented.

Example 1:

Harmonic centrality is also known in the literature as

- Latora closeness centrality
- Nodal efficiency
- Efficiency centrality

Example 2:

- Redundancy (1997)
- Local average connectivity (2011)

Example 3:

- Information centrality (1989)
- Current-flow closeness (2005)

Example 4:

- Node displacement (2010)
- Topological centrality (2013)
- Zeta vector centrality(2017)

Example 5:

- Generalized gravity centrality (2021)
- Clustering gravity model (2022)

and many others…

Node Metrics: Problems

Problem 4. Naming conflict: many new measures share the same name.

Examples:

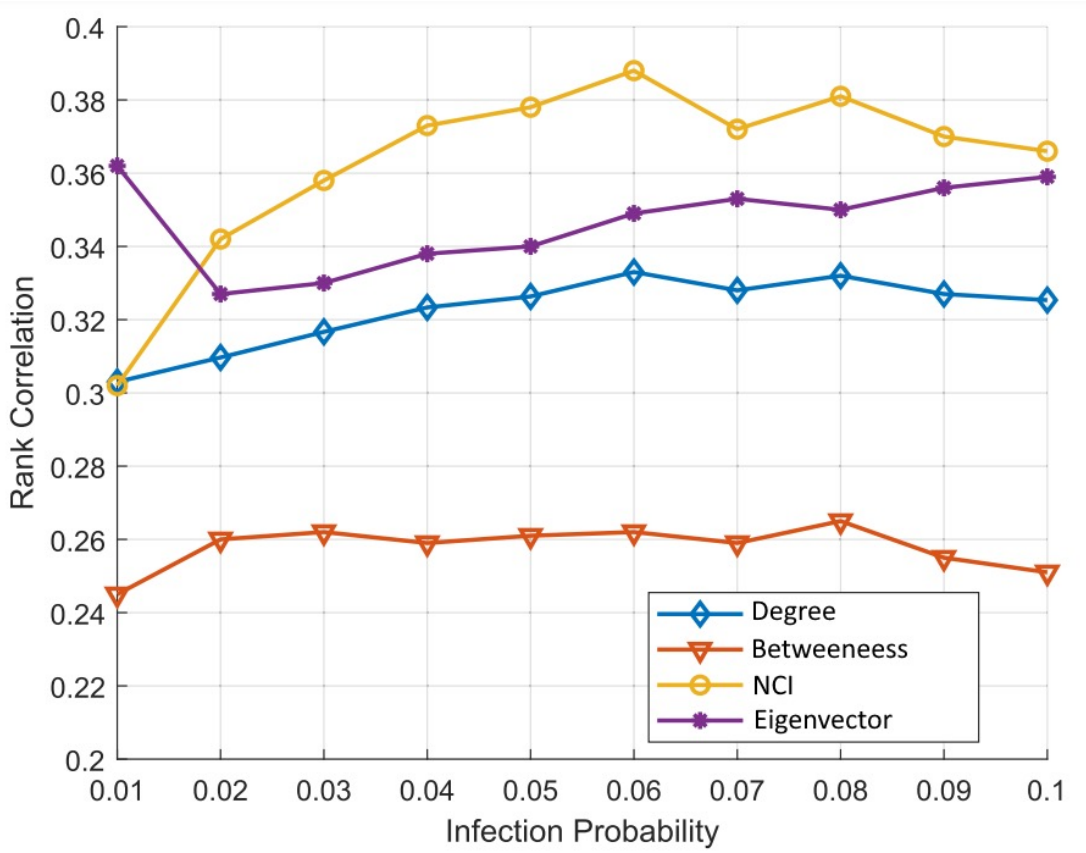
- Neighborhood centrality;
- Local centrality
- Improved <name> centrality.
- Hybrid centrality

Node Metrics: Problems

Problem 5. Poor validation of new models.

Example:

Comparison usually limited to 3–5 (typically classic) models, not the most advanced.



Node Metrics: Problems

Problem 6. Both researchers and reviewers face significant challenges when comparing new models with existing ones.

- Reviewers may not be familiar with all centrality measures, making it difficult to assess the significance of a new model.

- Authors may be unaware that their proposed measure already exists in the literature or that an alternative measure performs better.

Node Metrics: Challenges

Challenges:

- How to select the most appropriate measure?
- How to compare these models?
- How to validate new models?

Node Metrics

How to compare/validate centrality measures?

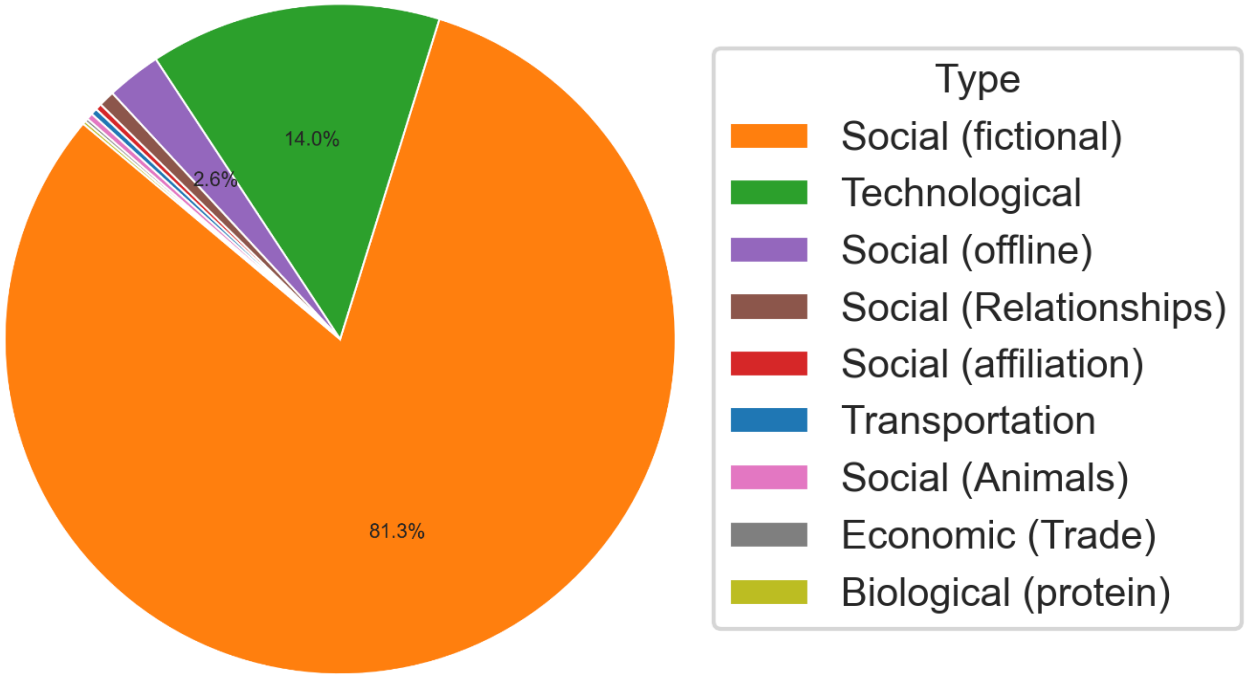
- Comparison to the ground-truth data (if available);
- Model comparison in downstream tasks: Identifying Influential Nodes (IIN) & Network Dismantling;
- Check the axioms for centrality;
- Evaluate the sensitivity of centrality measures to incomplete data.

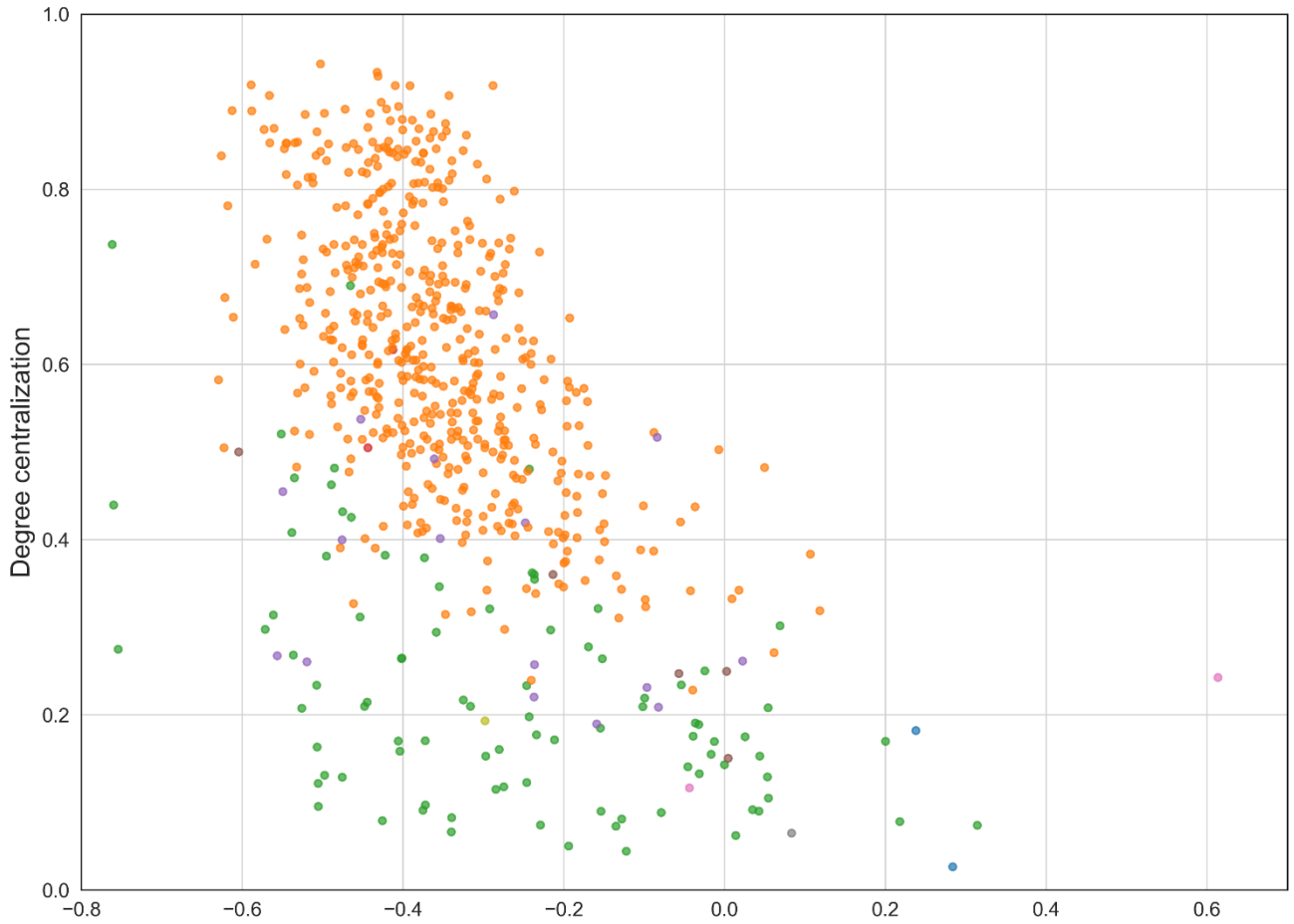
Comparison of the models: correlation analysis

Centrality Measures: Comparison

Correlation analysis

Data: 648 empirical networks, Index of Complex Networks (ICON).

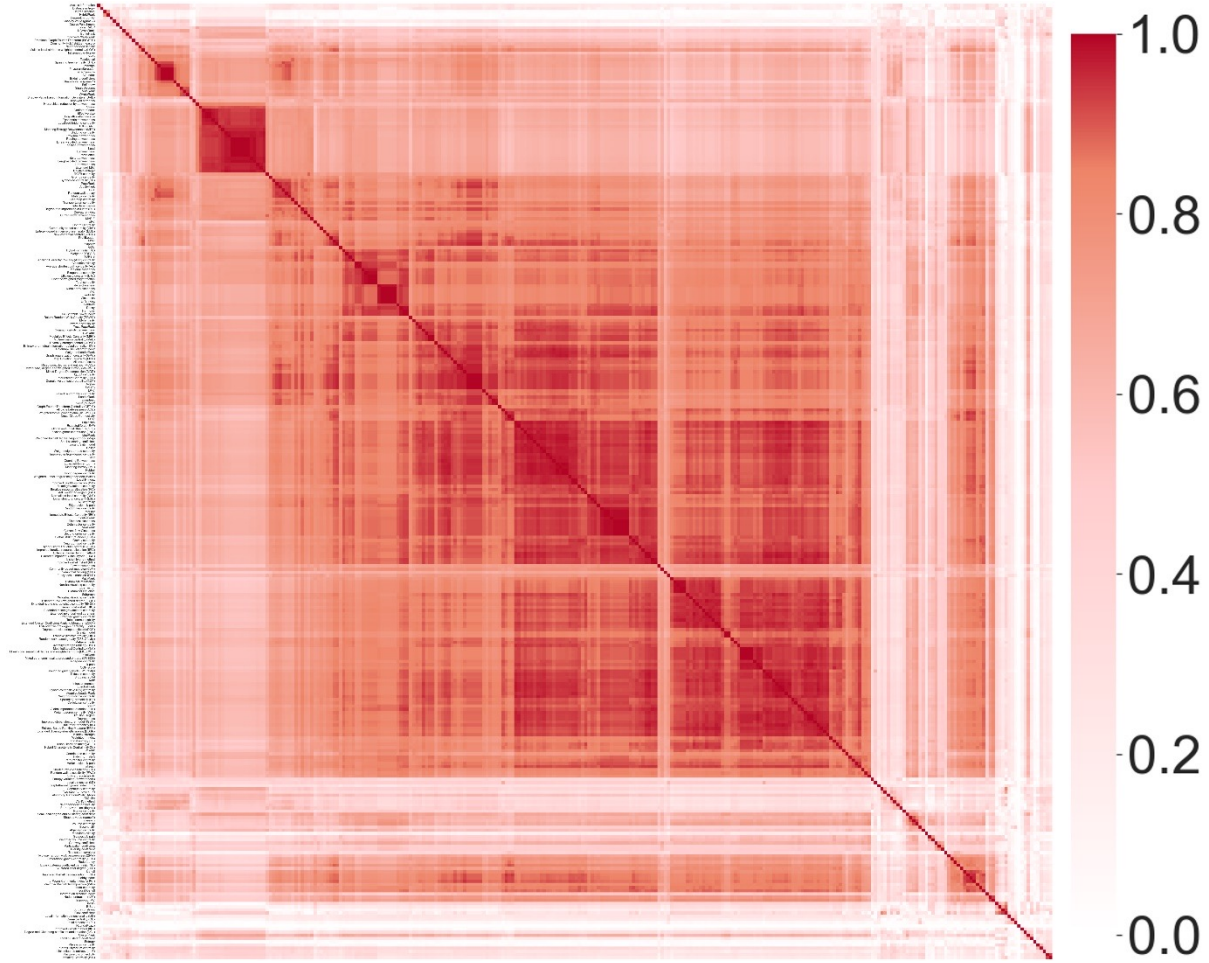




Comparison of the models: correlation analysis

Do we need all these measures?

Average correlation coefficient between >300 centrality measures on empirical networks



Many measures are correlated!

Comparison of the models: axiomatic approach

Comparison of the models: axiomatic approach

What properties do node metrics satisfy?

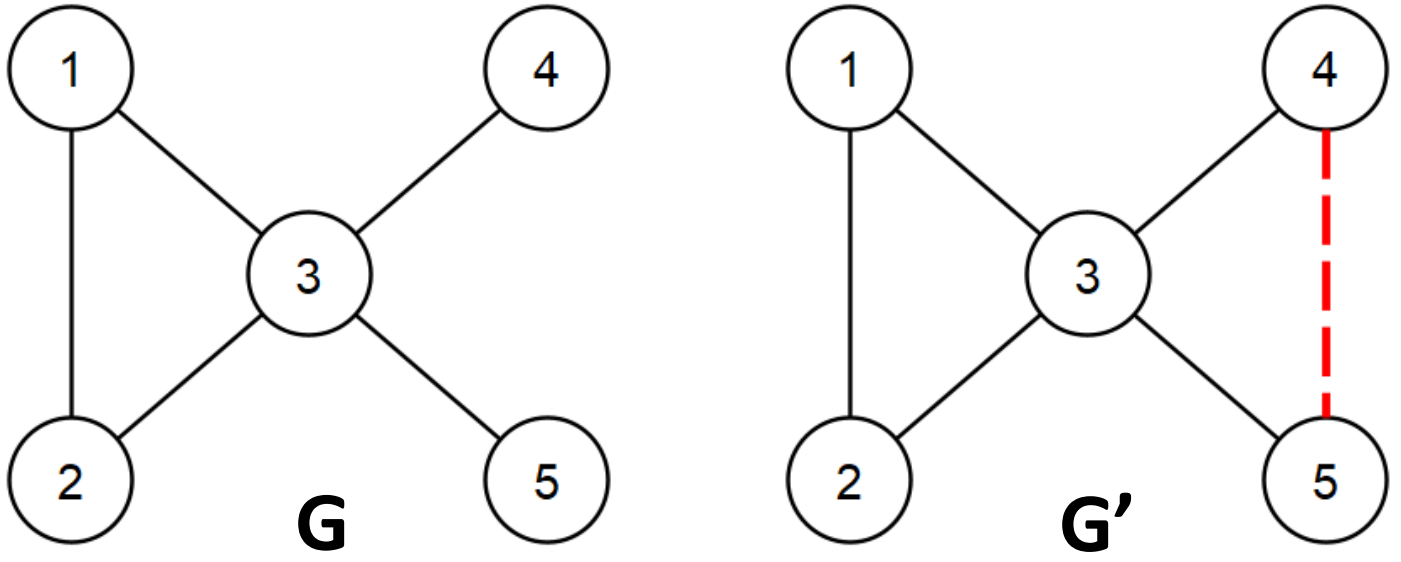
Axioms:

1) Anonymity;
2) Endpoint Increase;
3) Monotonicity;
4) Top Node;
5) Fairness;
6) Balanced Contributions;
7) …

References:
1) Sabidussi, G., The Centrality Index of a Graph, Psychometrika, 1966, vol. 31, no. 4, pp. 581–603.
2) Boldi, P. and Vigna, S., Axioms for Centrality, Internet Math., 2014, vol. 10, pp. 222–262.
3) Skibski, O., Michalak, T.P., and Rahwan, T., Axiomatic Characterization of Game-Theoretic Centrality, J. Artif. Int. Res., 2018, vol. 62, no. 1, pp. 33–68.

Axioms: Examples

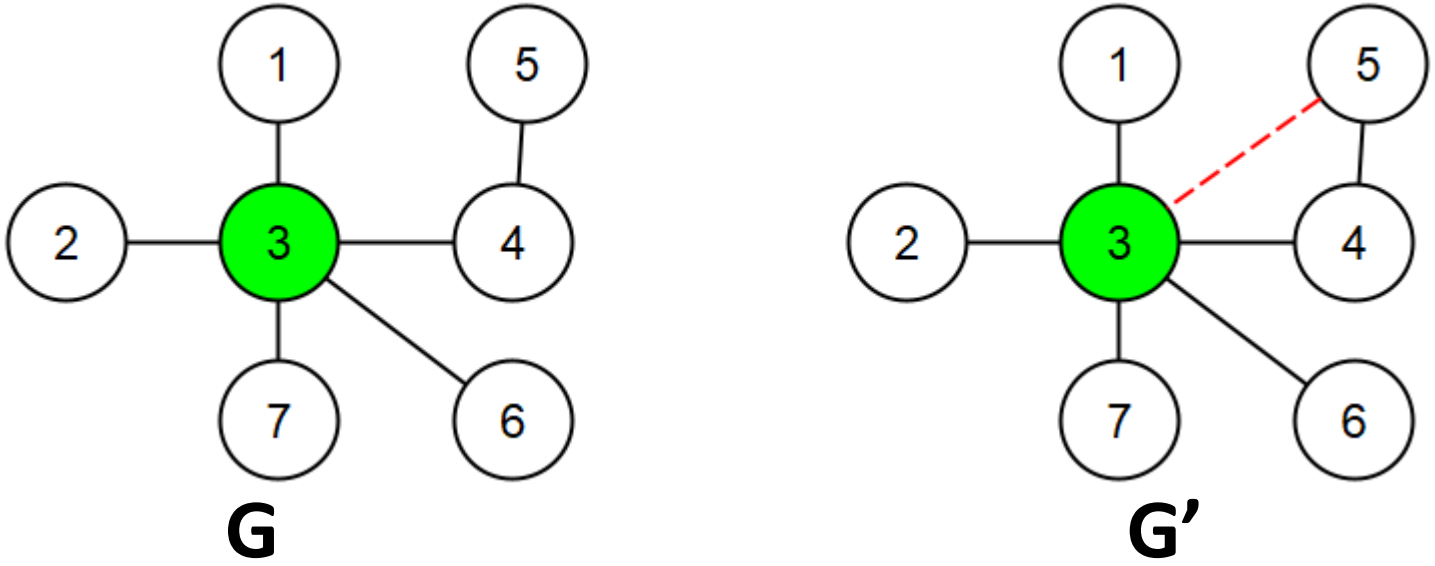
1. Endpoint Increase: “adding an edge increases the centrality of both endpoints”.



$$\begin{cases} C_4(G') > C_4(G) \\ C_5(G') > C_5(G) \end{cases}$$

Axioms: Examples

2. Top Node: *"if a node has the highest centrality, then adding an incident edge will not change that".*



$$\forall u, v \in V\colon\ C_u(G) \geq C_w(G) \Rightarrow C_u(G') \geq C_w(G')$$

Existing Axiomatic Results for Centrality Measures

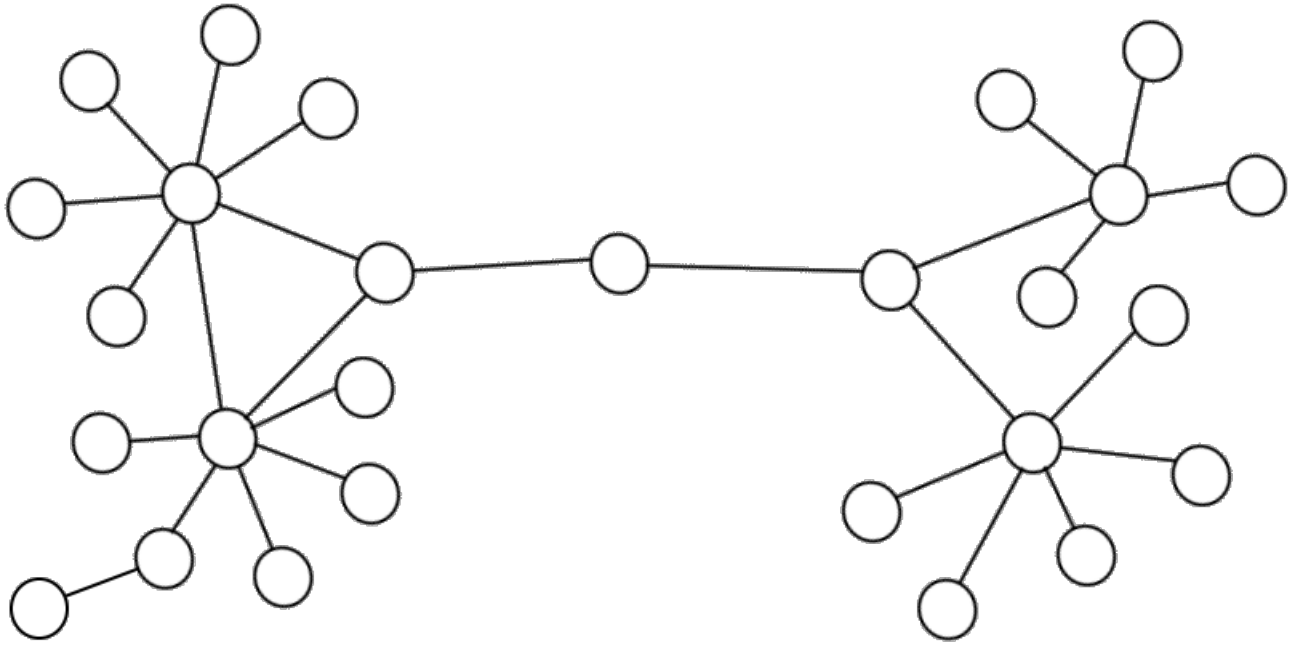
Centrality Measures	Axioms										
	Anonymity	Endpoint Increase	Monotonicity	Top Node	Fairness	Balanced Contributions	Add Edge Distance	Bridge	Cut-Vertex Additivity	Density	Size
Degree centralities	+	+	+	+	+	+	+	-	+	+	-
Eigenvector centrality	+	-	-	-	-	-	-	-	-	+	-
Katz centrality	+	+	+	?	-	-	-	-	-	+	-
PageRank	+	?	-	?	-	-	-	-	-	+	-
HITS	+	?	?	?	?	?	?	?	?	?	?
Closeness centrality	+	+	+	-	-	-	+	+	-	+	-
Betweenness centrality	+	-	-	-	-	-	-	-	-	-	-

Comparison of the models: sensitivity analysis

Comparison of the models: sensitivity analysis

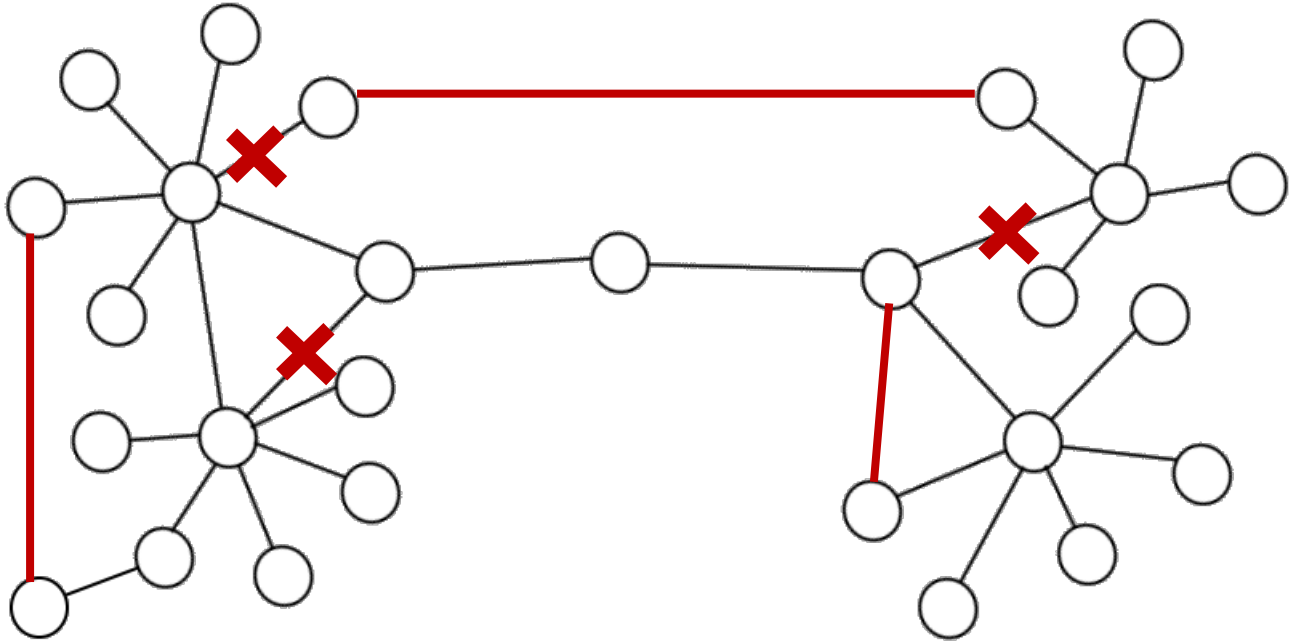
Many centralities are misused in real-world applications, leading to incorrect interpretations.

Example: Consider a real network:



Nodes	Centrality
1	0.8
2	**1**
3	0.3
4	0.2
…	…
27	0.05

Comparison of the models: sensitivity analysis



Nodes	Centrality
1	0.8
2	**1**
3	0.3
4	0.2
…	…
27	0.05

Nodes	Centrality
1	0.6
2	0.4
3	**0.8**
4	0.3
…	…
27	0.1

Many real networks are partially observed!

- Missing links
- Incorrect links

Node Metrics: Problems

The number of proposed metrics is overwhelming.

Problem 1. Most existing measures remain unknown.

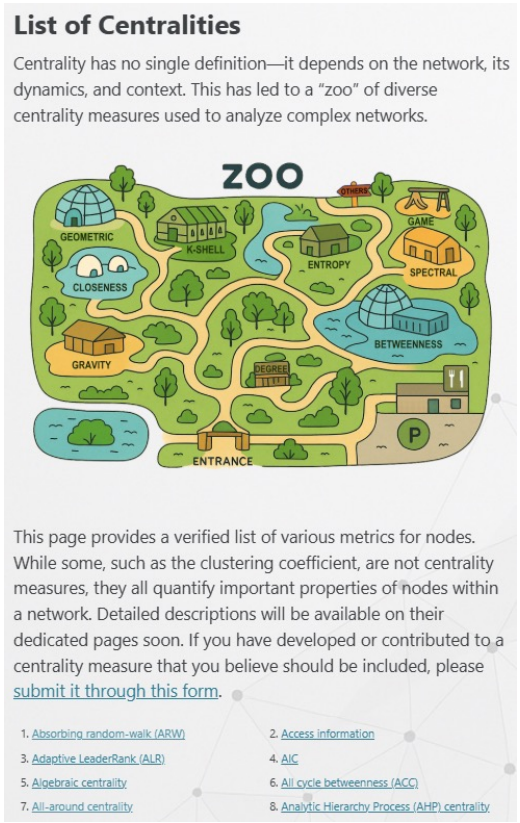
Problem 2. Access to many centrality models is limited.

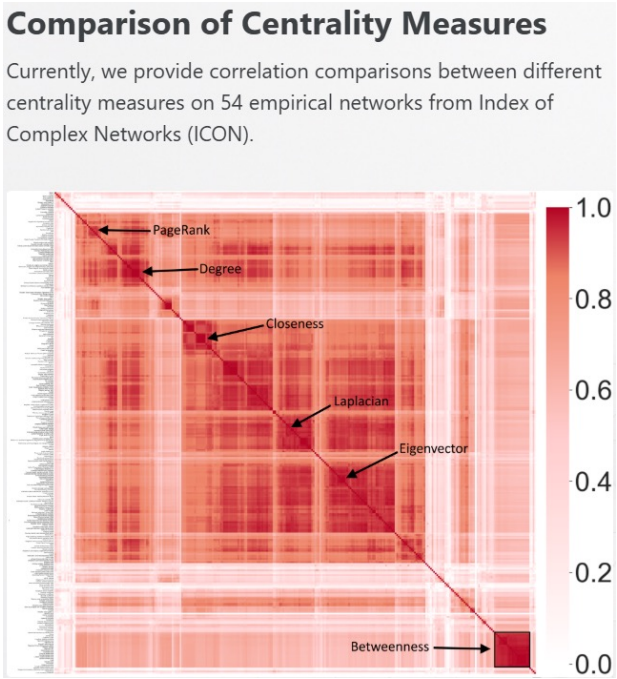
Problem 3. Duplication: many models are being reinvented.

Problem 4. Naming conflict: many new measures share the same name.

Zoo of Centralities: Website

Visit: https://centralityzoo.github.io/





List of measures and their description

Comparison of models (TBA)

+ Code (Python, TBA)

Conclusion

- The field of node metrics is messy and needs better organization.
- **https://centralityzoo.github.io/**: visit the website for updates.
- Metric missing? Visit the website to fill the form.
- Open for collaboration.

Thank you!