Northeastern University
Network Science Institute

So Many Choices, So Few Guidelines: Sifting through Functions on Networks

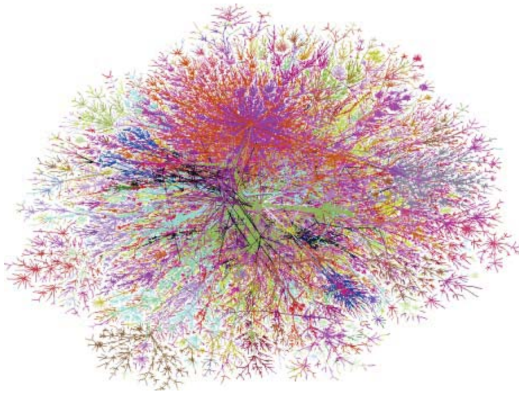
Tina Eliassi-Rad

tina@eliassi.org

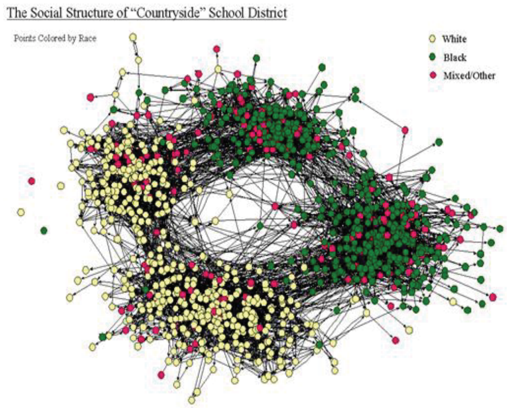
Supported by NSF, DTRA, DARPA, IARPA, DOE/LLNL & WaPo Labs

Complex Networks are Ubiquitous

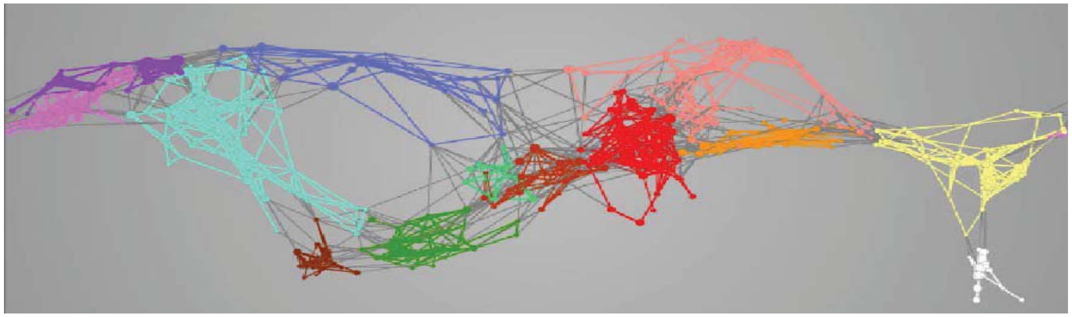
Technological Networks
(e.g., Internet)



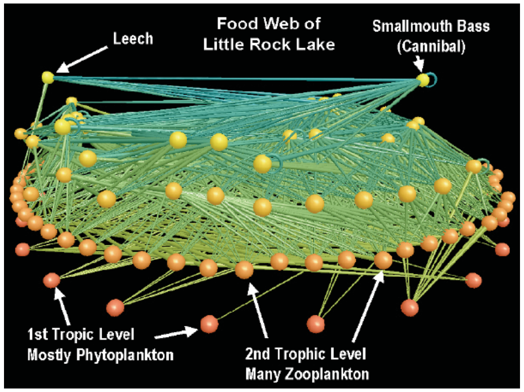
Social Networks
(e.g., Friendship Network)



Information Networks
(e.g., Map of Science)



Biological networks
(e.g., Food Web)



WE ♥ OUR NETWORKS

So, What's the Problem?

- The literature is cluttered with various measures of tie-strength, centrality, robustness influence, community, discovery, network similarity, …

- Deciding which measure to choose is nontrivial because little exists in terms of theoretical work

Roadmap

- A theoretical guide to tie-strength measures
- An empirical guide to network-similarity methods

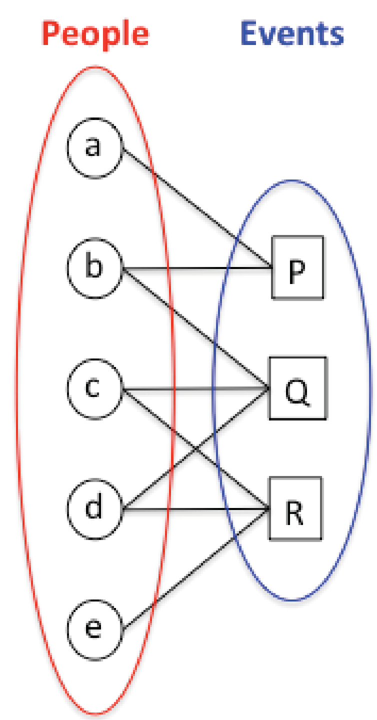
Roadmap

- A theoretical guide to tie-strength measures
  - Mangesh Gupte and Tina Eliassi-Rad: Measuring tie strength in implicit social networks. In *ACM WebSci* 2012.
    - http://eliassi.org/papers/gupte-websci12.pdf
- An empirical guide to network-similarity methods

Problem Definition

- Given a bipartite graph with people as one set of vertices and events as the other set, measure *tie strength* between each pair of individuals

- Assumption
  - Attendance at mutual events implies an implicit weighted social network between people



Motivation

- Most real-world networks are bipartite and are converted to unipartite (e.g., via AA^T)
- Explicitly declared friendship links can suffer from a low signal-to-noise ratio (e.g., Facebook friends)

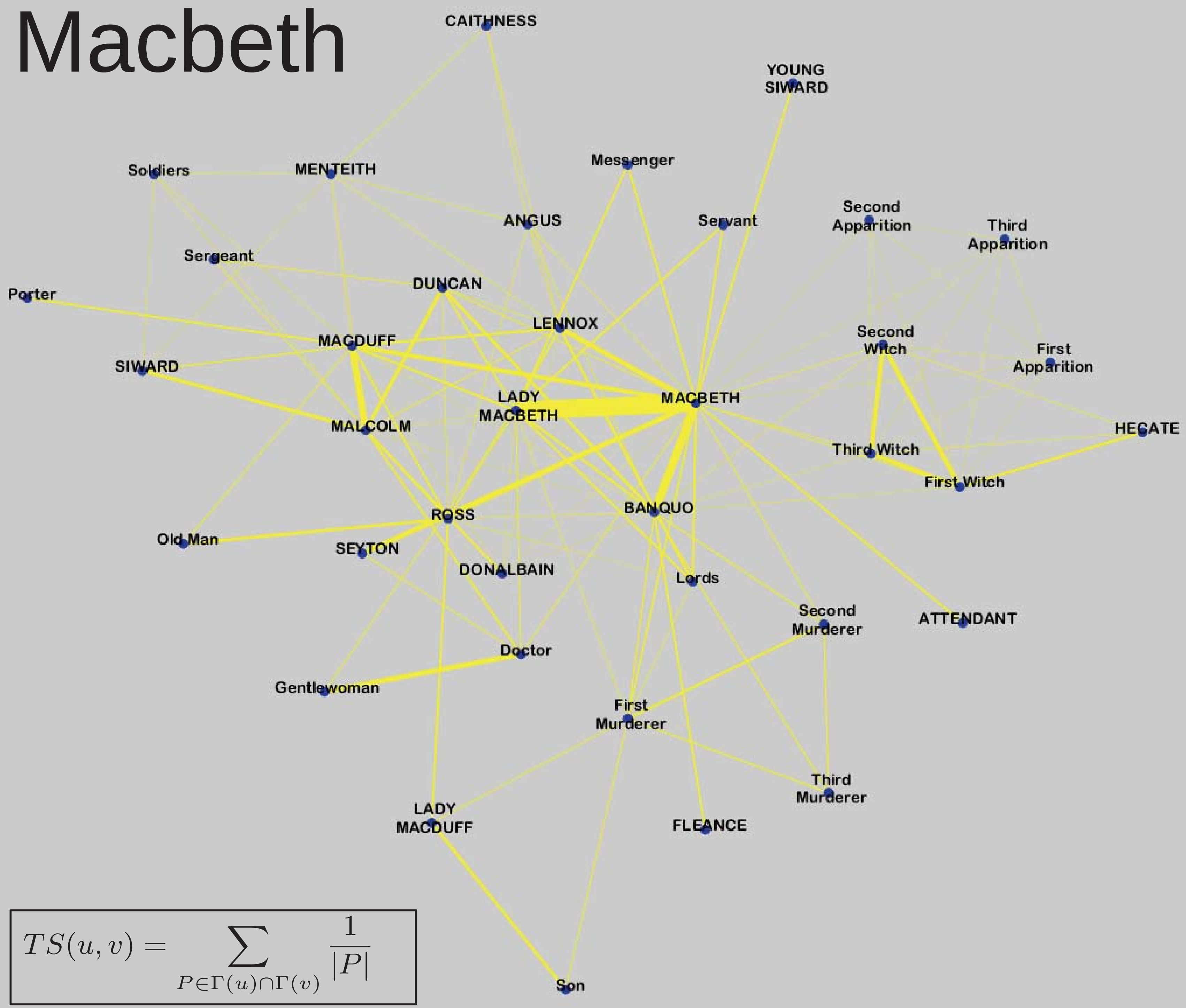
- **Challenge:** Detect which of links in the unipartite graph are important
- **Goal:** Infer the implicit weighted social network from people's participation in mutual events

Tie Strength

- A measure of tie strength induces
  - a ranking on all the edges, and
  - a ranking on the set of neighbors for every person

- Example of a simple tie-strength measure
  - Common neighbor measures the total number of common events to a pair of individuals

Macbeth



$$TS(u,v) = \sum_{P \in \Gamma(u) \cap \Gamma(v)} \frac{1}{|P|}$$

Decisions, Decisions

- There are many different measures of tie-strength
  1. Common neighbor
  2. Jaccard index
  3. Max
  4. Linear
  5. Delta
  6. Adamic and Adar
  7. Preferential attachment
  8. Katz measure
  9. Random walk with restarts
  10. Simrank
  11. Proportional
  12. …

Which one should you choose?

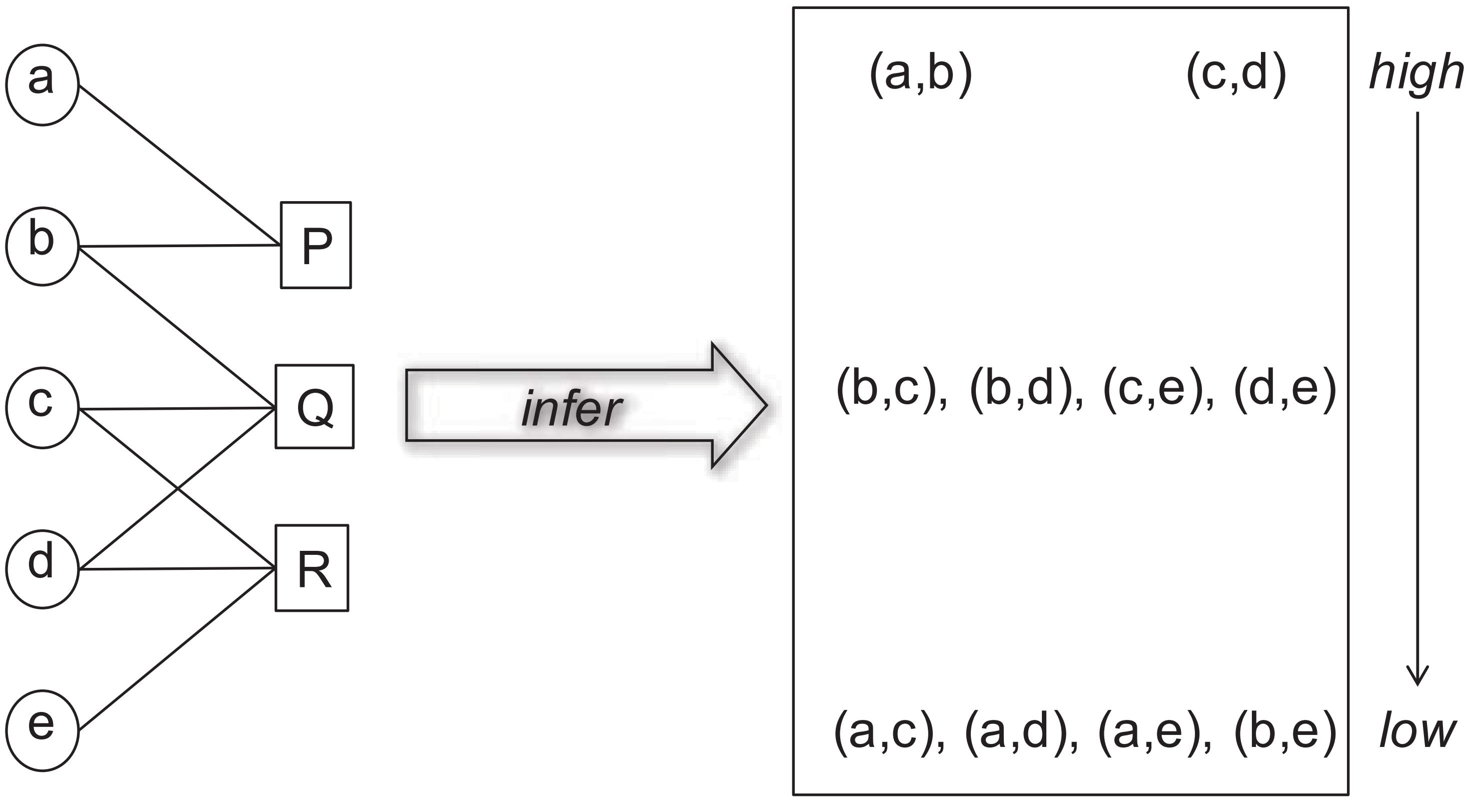
Outline for Tie-Strength

- An axiomatic approach to the problem of inferring implicit social networks by measuring tie strength
- A characterization of functions that satisfy all our axioms
- Classification of prior measures according to the axioms that they satisfy
- Experiments
- Wrap-up

Running Example

Input
People × Event Bipartite Graph

Output
Partial Order of Tie Strength among People

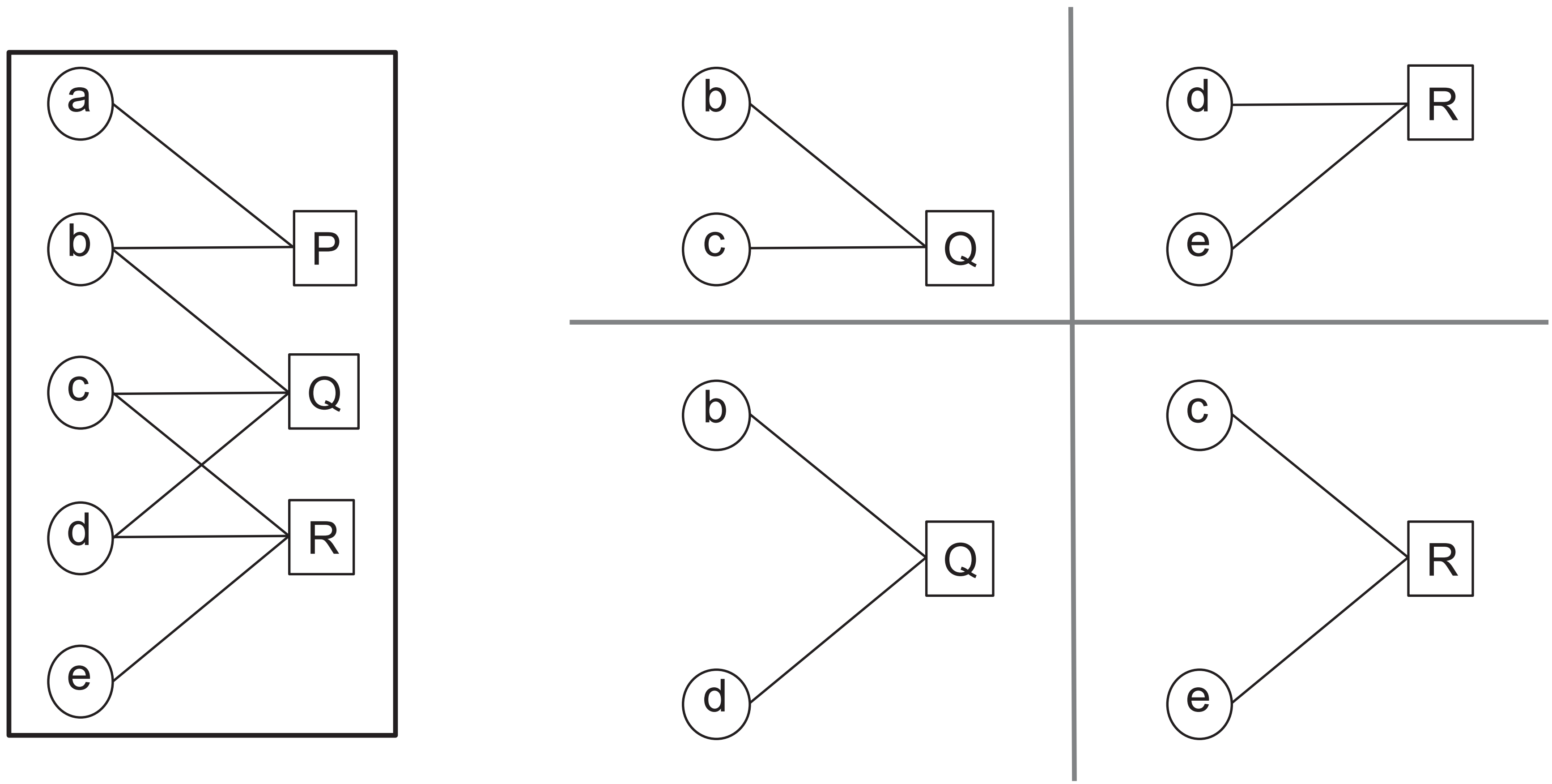


Axioms

- Axiom 1: Isomorphism
- Axiom 2: Baseline
- Axiom 3: Frequency
- Axiom 4: Intimacy
- Axiom 5: Popularity
- Axiom 6: Conditional Independence of People
- Axiom 7: Conditional Independence of Events
- Axiom 8: Submodularity

Axiom 1: Isomorphism

- Tie strength between *u* and *v* is independent of the labels of *u* and *v*



Axiom 2: Baseline

- If there are no events, then tie strength between each pair *u* and *v* is 0

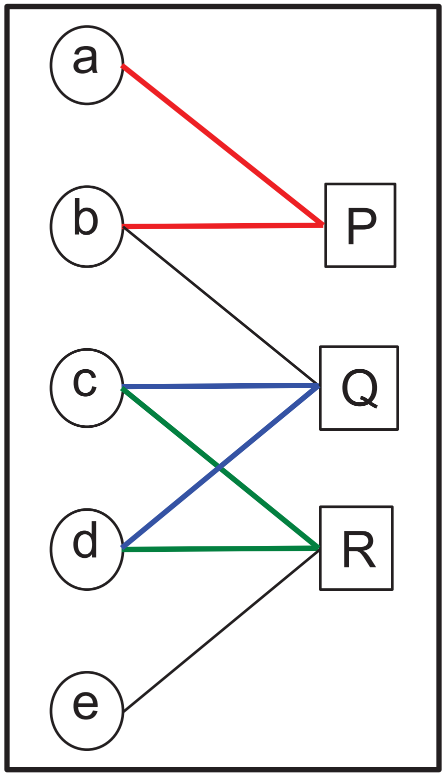
$$TS_{\varnothing}(u, v) = 0$$

- If there are only two people *u* and *v* and a single event *P* that they attend, then their tie strength is at most 1

$$TS_P(u, v) \leq 1$$

  - Defines an upper-bound for how much tie strength can be generated from a single event between two people

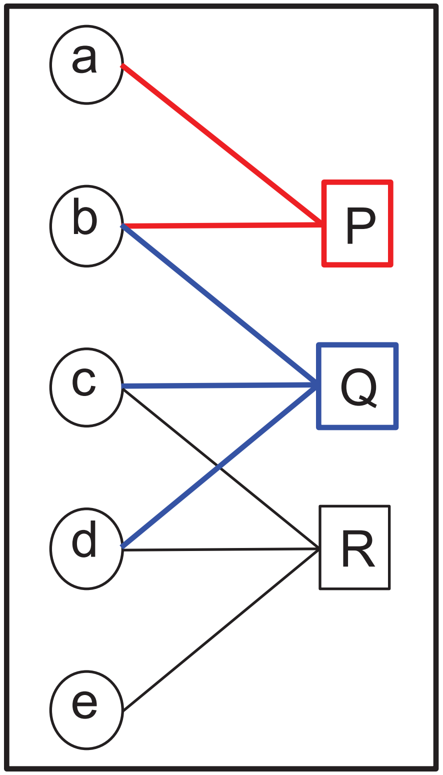
Axiom 3: Frequency & Axiom 4: Intimacy



- Axiom 3 (**Frequency**)
  - More events create stronger ties
  - All other things being equal, the more events common to *u* and *v*, the stronger their tie-strength
- Axiom 4 (**Intimacy**)
  - Smaller events create stronger ties
  - All other things being equal, the fewer invitees there are to any particular event attended by *u* and *v*, the stronger their tie-strength

Axiom 5: Popularity

- Larger events create more ties
- Consider two events *P* and *Q*
- If $|Q| > |P|$, then the total tie strength created by *Q* is more than that created by *P*



Axioms 6 & 7: Conditional Independence of People and of Events

- Axiom 6: **Conditional Independence of People**
  - A node *u*'s tie strength to other people does **not** depend on events that *u* does **not** attend
- Axiom 7: **Conditional Independence of Events**
  - The increase in tie strength between *u* and *v* due to an event *P* does **not** depend on other events, just on the existing tie strength between *u* and *v*
  - $TS_{(G+P)}(u, v) = g(TS_G(u, v), TS_P(u, v))$
    - where *g* is some monotonically increasing function

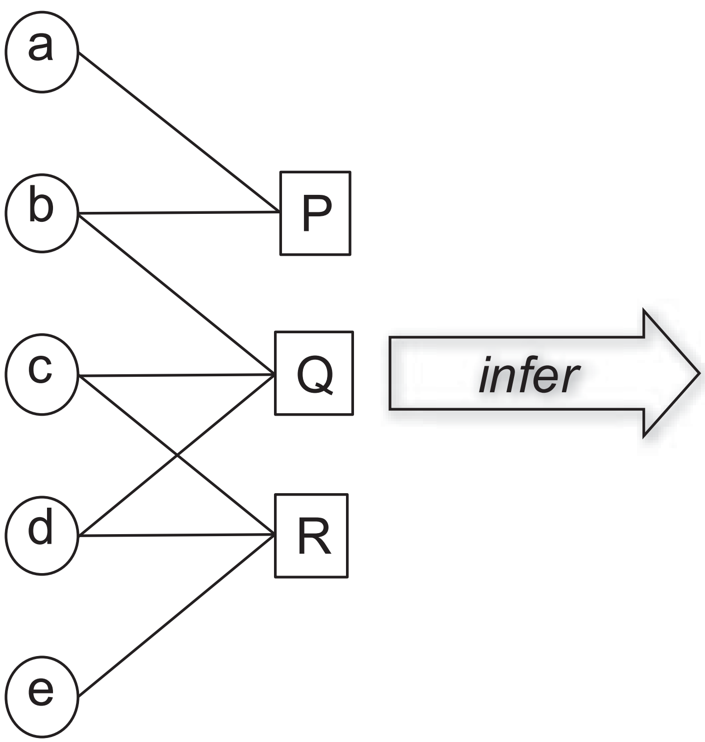
Axiom 8: Submodularity

- The marginal increase in tie strength of u and v due to an event Q is at most the tie strength between u and v if Q was their only event
- If G is a graph and Q is a single event, then $TS_{(G+Q)}(u, v) - TS_G(u, v) \leq TS_Q(u, v)$

Example – Mapping to Axioms

Input
People × Event Bipartite Graph

Output
Partial order of Tie Strength



(a,b) (c,d)	*high*	*Axiom 4 (Intimacy) & Axiom 3 (Freq)*
(b,c), (b,d), (c,e), (d,e)		*Axiom 1 (Isomorphism)*
(a,c), (a,d), (a,e), (b,e)	*low*	*Axiom 2 (Baseline) & Axiom 6 (Cond. Indep. of Vertices) & Axiom 7 (Cond. Indep. of Events)*

Observations on the Axioms

- Our axioms are fairly intuitive

A1: Isomorphism	A2: Baseline	A3: Frequency	A4: Intimacy
A5: Popularity	A6: Cond. Indep. of people	A7: Cond. indep. of events	A8: Submodularity

- But, several previous measures in the literature break some of these axioms
- Satisfying all the axioms is not sufficient to uniquely identify a measure of tie strength
  - One reason: inherent tension between Axiom 3 (Frequency) and Axiom 4 (Intimacy)

Inherent Tension Between Frequency & Intimacy

- Scenario #1 (intimate)
  - Mary and Susan go to 2 parties, where they are the only people there.
- Scenario #2 (frequent)
  - Mary, Susan, and Jane go to 3 parties, where they are the only people there.
- In which scenario is Mary's tie to Susan stronger?

Observations on the Axioms (cont.)

A1: Isomorphism	A2: Baseline	A3: Frequency	A4: Intimacy
A5: Popularity	A6: Cond. Indep. of people	A7: Cond. indep. of events	A8: Submodularity

- Axioms are equivalent to a natural partial order on the strength of ties
  - Pertinent to ranking application
- Choosing a particular tie-strength function is equivalent to choosing a particular linear extension of this partial order
  - Non-obvious decision
  - Details in WebSci 2012 paper

Preamble to the Characterization Theorem

- Let $f(n)$ = total tie strength generated in a <u>single</u> event with n people
- If there is a <u>single</u> party with n people, the tie strength of each tie is

$$\frac{f(n)}{\binom{n}{2}}$$

  - Based on Axiom 1 (Isomorphism)

- The total tie strength created at an event P with n people is a monotone function $f(n)$ that is bounded by

$$1 \leq f(n) \leq \binom{n}{2}$$

  - Based on Axiom 2 (Baseline) and Axiom 4 (Intimacy) and Axiom 5 (Popularity)

Characterizing Tie Strength

A way to explore the space of valid functions for representing tie strength and find which work given particular applications

Theorem. Given a graph $G = (L \cup R, E)$ *and two vertices u and v*, *if* the tie-strength function *TS* follows Axioms (1-8), then the function has to be of the form

$$TS_G(u, v) = g(h(|P_1|), h(|P_2|), \ldots, h(|P_k|))$$

- $\{P_i\}_{1 \le i \le k}$ are the events common to both *u* and *v*
- h is a monotonically decreasing function bounded by $1 \geq h(n) \geq \frac{1}{\binom{n}{2}}, n \geq 2; h(1) = 1; h(0) = 0.$
- g is a monotonically increasing submodular function

Many Measures of Tie Strength

1. Common neighbor
2. Jaccard index
3. Max
4. Linear
5. Delta
6. Adamic and Adar
7. Preferential attachment
8. Katz measure
9. Random walk with restarts
10. Simrank
11. Proportional

$$TS(u,v) = |\Gamma(u) \cap \Gamma(v)|$$

$$TS(u,v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|}$$

$$TS(u,v) = \max_{P \in \Gamma(u) \cap \Gamma(v)} \frac{1}{|P|}$$

$$TS(u,v) = \sum_{P \in \Gamma(u) \cap \Gamma(v)} \frac{1}{|P|}$$

$$TS(u,v) = \sum_{P \in \Gamma(u) \cap \Gamma(v)} \frac{1}{\binom{|P|}{2}}$$

$$TS(u,v) = \sum_{P \in \Gamma(u) \cap \Gamma(v)} \frac{1}{\log |P|}$$

$$TS(u,v) = |\Gamma(u)| \cdot |\Gamma(v)|$$

$$TS(u,v) = \sum_{q \in \text{ path between } u,v} \gamma^{-|q|}$$

$$TS(u,v) = \begin{cases} 1 & \text{if u = v} \\ \gamma \cdot \frac{\sum_{a \in \Gamma(u)} \sum_{b \in \Gamma(v)} TS(a,b)}{|\Gamma(u)| \cdot |\Gamma(v)|} & \text{otherwise} \end{cases}$$

$$TS(u,v) = \sum_{P \in \Gamma(u) \cap \Gamma(v)} \frac{\epsilon}{|P|} + (1-\epsilon) \frac{TS(u,v)}{\sum_{w \in \Gamma(u)} TS(u,w)}$$

Non Self-Referential Tie Strength Measures

- Common neighbor
  - The total # of common events that both u and v attended
- Jaccard Index
  - Similar to common neighbor
  - Normalizes for how “social” *u* and *v* are
- Adamic and Adar [2003], Delta, and Linear
  - Tie strength increases with the number of events
  - Tie strength is 1 over a simple function of event size
- Max
  - Tie strength does not increase with the number of events
  - Tie strength is the maximum tie strength from all common events

Self-Referential Tie-Strength Measures

- Katz measure [Katz,1953]
  - Tie strength is the number of paths between u and v, where each path is discounted exponentially by the length of the path
- Random walk with restarts
  - A non-symmetric measure of tie strength
  - Tie strength is the stationary probability of a Markov chain process
  - With probability α, jump to a node u; and with probability 1-α, jump to a neighbor of a current node.
- Simrank [Jeh & Widom, 2002]
  - Tie strength is captured by recursively computing the tie strength of neighbors
- Proportional
  - Tie strength increases with # of events
  - People spend time proportional to their tie-strength at a party

Measures of Tie-Strength that Satisfy All the Axioms

A1: Isomorphism	A2: Baseline	A3: Frequency	A4: Intimacy
A5: Popularity	A6: Cond. indep. of P	A7: Cond. indep. of E	A8: Submodularity

	A1	A2	A3	A4	A5	A6	A7	A8	$g(a_1, \ldots, a_k)$ $h(\lvert P_i \rvert) = a_i$
Common Neighbors	✓	✓	✓	✓	✓	✓	✓	✓	$g(a_1, \ldots, a_k) = \Sigma a_i$ $h(n) = 1$
Delta	✓	✓	✓	✓	✓	✓	✓	✓	$g(a_1, \ldots, a_k) = \Sigma a_i$ $h(n) = 2(n(n-1))^{-1}$
Adamic & Adar	✓	✓	✓	✓	✓	✓	✓	✓	$g(a_1, \ldots, a_k) = \Sigma a_i$ $h(n) = (\log(n))^{-1}$
Linear	✓	✓	✓	✓	✓	✓	✓	✓	$g(a_1, \ldots, a_k) = \Sigma a_i$ $h(n) = n^{-1}$
Max	✓	✓	✓	✓	✓	✓	✓	✓	$g(a_1, \ldots, a_k) = \max\{a_i\}$ $h(n) = n^{-1}$

Measures of Tie-Strength that Do **Not** Satisfy All the Axioms

A1: Isomorphism	A2: Baseline	A3: Frequency	A4: Intimacy
A5: Popularity	A6: Cond. indep. of P	A7: Cond. indep. of E	A8: Submodularity

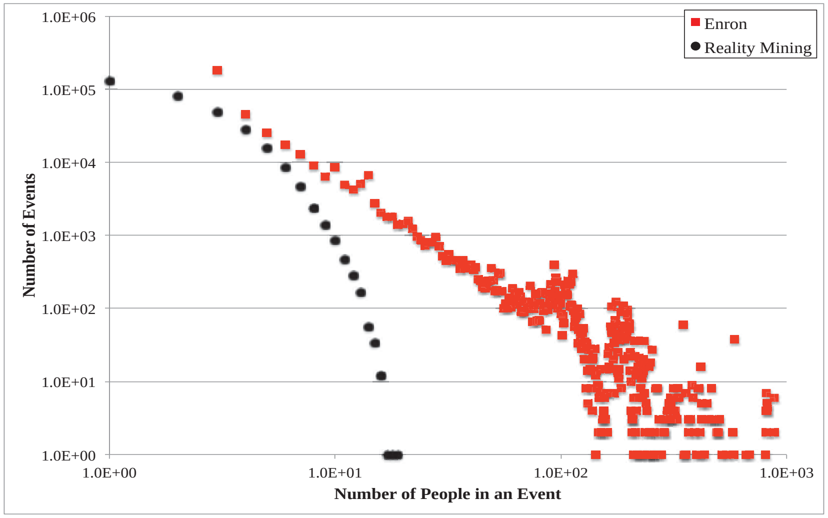
	A1	A2	A3	A4	A5	A6	A7	A8	$g(a_1, \ldots, a_k)$ $h(\lvert P_i \rvert) = a_i$
Jaccard Index	✓	✓	✓	✓	✓	✗	✗	✗	✗
Katz Measure	✓	✗	✓	✓	✓	✓	✗	✗	✗
Preferential Attachment	✓	✓	✗	✓	✓	✓	✗	✗	✗
RWR	✓	✗	✗	✗	✓	✓	✗	✗	✗
Simrank	✓	✗	✗	✗	✗	✗	✗	✗	✗
Proportional	✓	✗	✗	✓	✗	✓	✗	✗	✗

Data Sets

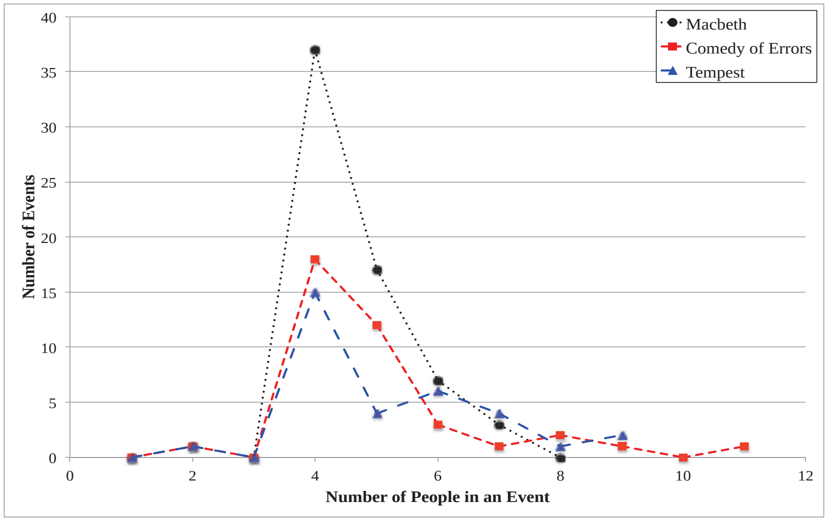
Graphs	# of People	# of Events
Southern Women	18	14
The Tempest	19	34
A Comedy of Errors	19	40
Macbeth	38	67
Reality Mining Bluetooth	104	326,248
Enron Emails	32,471	371,321

Degree Distributions

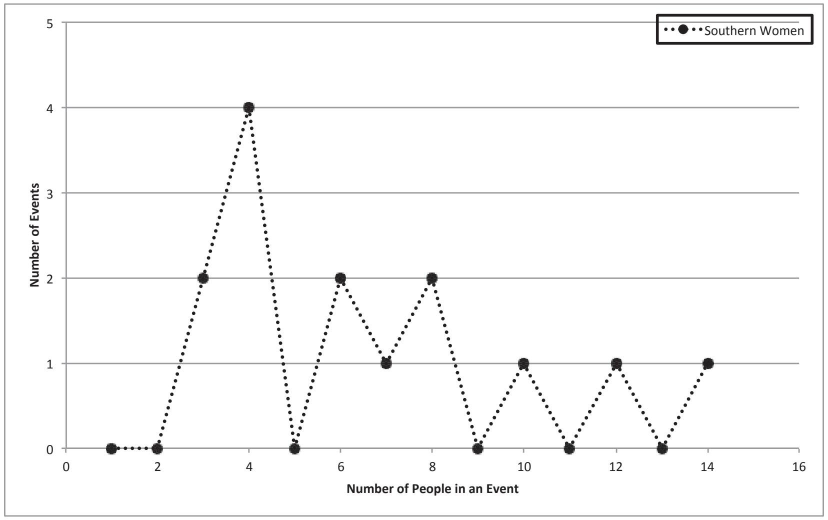
Enron & Reality Mining



Shakespeare's Plays



Southern Women



Completeness of Axioms 1-8

(Number of Ties **Not** Resolved by the Partial Order)

Dataset	Tie Pairs	Incomparable Pairs (%)
Southern Women	11,628	683 (5.87)
The Tempest	14,535	275 (1.89)
A Comedy of Errors	14,535	726 (4.99)
Macbeth	246,753	584 (0.23)
Reality Mining	13,794,378	1,764,546 (12.79)

- % of tie-pairs where different tie-strength functions can differ
  - Smaller is better
  - Generally, percentages are small
  - Large real-world networks have more unresolved ties

$$\text{\# of tie pairs} = \binom{\binom{n}{2}}{2}$$

Take-away Point #1

% of tie pairs on which different tie-strength functions can differ is small.*

* Disclaimer: For ranking applications <u>and</u> tie-strength functions that satisfy the axioms

Two Tie-Strength Functions that Do **Not** Satisfy the Axioms

- Jaccard Index
  - Normalizes for how "social" *u* and *v* are

$$TS(u,v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|}$$

- Temporal Proportional
  - Increases with number of events
  - People spend time proportional to their tie-strength in a party
  - Events are ordered by time

$$TS(u,v,t) = \begin{cases} TS(u,v,t-1) & \text{if } u \text{ and } v \text{ do not attend } P_t \\ \epsilon \frac{1}{|P_t|} + (1-\epsilon)\frac{TS(u,v,t-1)}{\sum_{w \in P_t} TS(u,w,t-1)} & \text{otherwise} \end{cases}$$

Soundness of Axioms 1-8

(Number of Conflicts Between the Partial Order and Tie-Strength Functions **Not** Satisfying the Axioms)

Dataset	Tie Pairs	Jaccard (%)	Temporal (%)
Southern Women	11,628	1,441 (12.39)	665 (5.72)
The Tempest	14,535	488 (3.35)	261 (1.79)
A Comedy of Errors	14,535	1,114 (7.76)	381 (2.62)
Macbeth	246,753	2,638 (1.06)	978 (0.39)
Reality Mining	13,794,378	290,934 (0.02)	112,546 (0.01)

- % of tie-pairs in conflict with the partial order
  - Smaller is better
  - Generally, percentages are small
  - They decrease as the dataset increases

More on Soundness

- **Question 1:**

  Are the number of conflicts, between the partial order and tie-strength functions not satisfying the axioms, small because most of the tie-strengths are zeros (sparsity of real graph)?

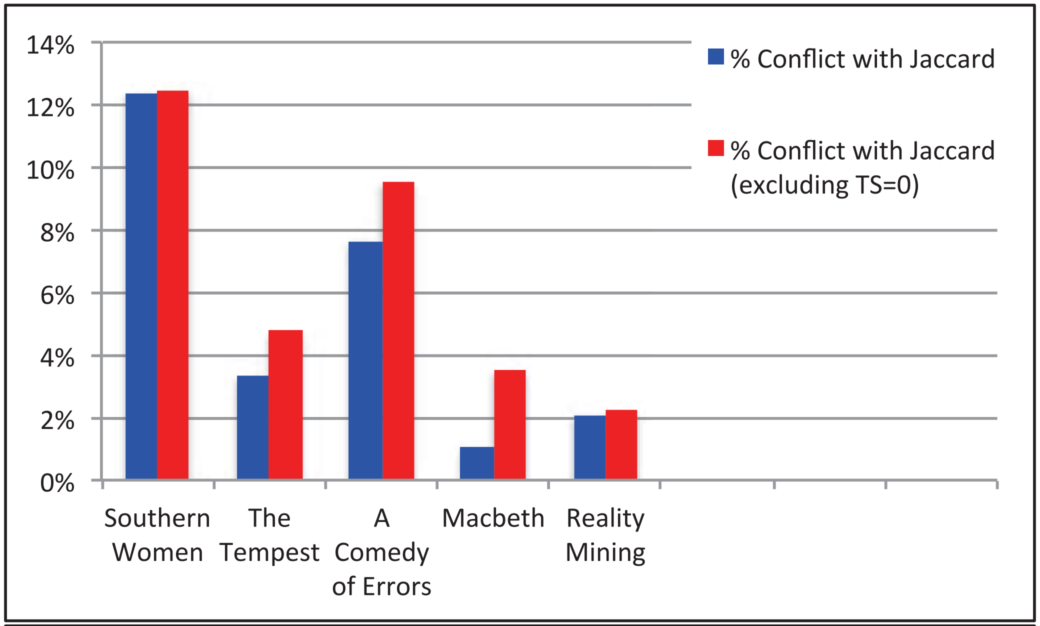
- **Answer:**
  - This is partially true.
  - For some pairs, the tie-strength being set to zero is caused by the axioms.
  - It may or may not be true that all these pairs have tie-strength zero in the actual function used.
    - For example, this won't be true for some self-referential functions like Simrank, Random Walk with Restart, etc.

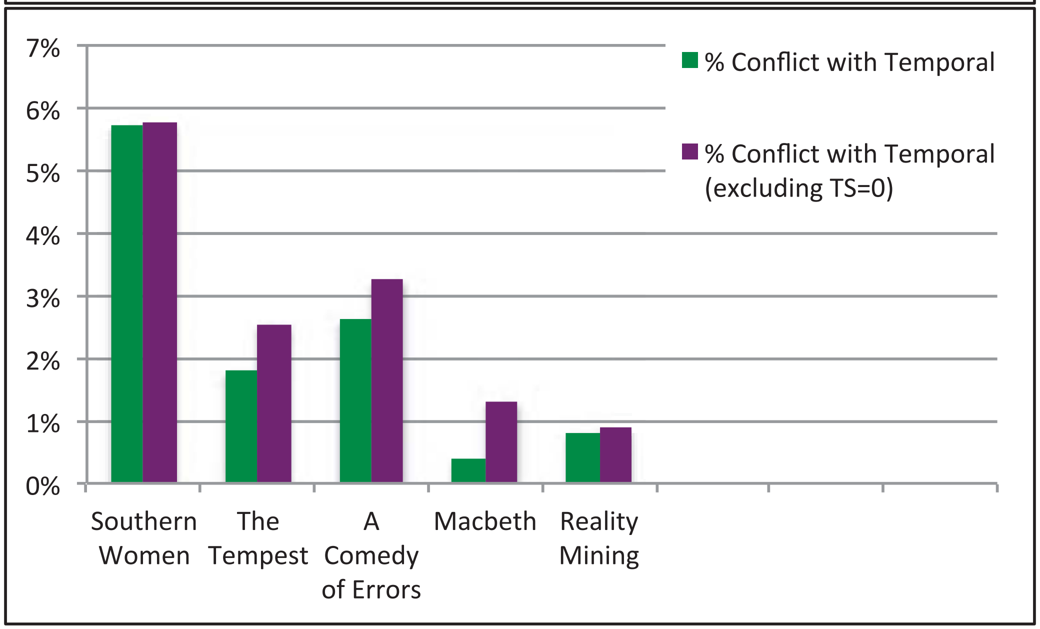
Even More on Soundness

- **Question 2:** How do the conflict numbers change if we only looked at tie pairs that have nonzero tie-strengths?
- **Answer:** The percentages go up but not by much.

Dataset	Tie Pairs	Tie Pairs (excluding TS=0)	Jaccard	Temporal
Southern Women	11,628	11,537	1,441	665
The Tempest	14,535	10,257	488	261
A Comedy of Errors	14,535	11,685	1,114	381
Macbeth	246,753	74,175	2,638	978
Reality Mining	13,794,378	12,819,272	290,934	112,546

Even More on Soundness





Take-away Point #2

% of conflicts between our axioms and tie-strength functions **not** satisfying our axioms is small.*

* Disclaimer: For ranking applications

Putting Take-away Points #1 & #2 Together

1. % of tie pairs on which different tie-strength functions can differ is small

2. % of conflicts between our axioms and tie-strength functions <u>not</u> satisfying our axioms is small

3. If your application is ranking, just pick the most computationally efficient tie-strength measure (e.g. common neighbor).

Disclaimer: For ranking applications

Scalability Issue

- # of tie pairs = $\binom{\binom{n}{2}}{2}$
- Enron has 32,471
- # of tie pairs in Enron ≈ 138 quadrillion

$$\binom{\binom{32471}{2}}{2} = 138,952,356,623,361,270$$

- Ignore zero tie-strengths

A Real-world Application

- **Input:** Data from an online friendship network and its social reader*
- **Q1:** How can we effectively capture the similarities between the reading behaviors of a user and her friends over time?
- **Q2:** How can we effectively summarize such similarities across users?

Details in our *NewsKDD* 2014 paper.
http://eliassi.org/papers/le-newskdd14.pdf

* A reading application deployed on a social network

Related Work

- Strength of ties
  - Spread of information in social networks [Granovetter, 1973]
  - Use external information to learn strength of tie
    - [Gilbert & Karahalios, 2009], [Kahanda & Neville, 2009]
- Very few axiomatic work approaches to graph measures
  - PageRank axiomatization [Altman & Tennenholtz, 2005]
  - Information theoretic measure of similarity [Lin, 1998]
    - Assumes probability distribution over events
- Link prediction
  - [Adamic & Adar, 2003]
  - [Liben-Nowell & Kleinberg, 2003]
  - [Sarkar, Chakrabarti, Moore, 2010 & 2011]

Wrap-up of Tie Strength

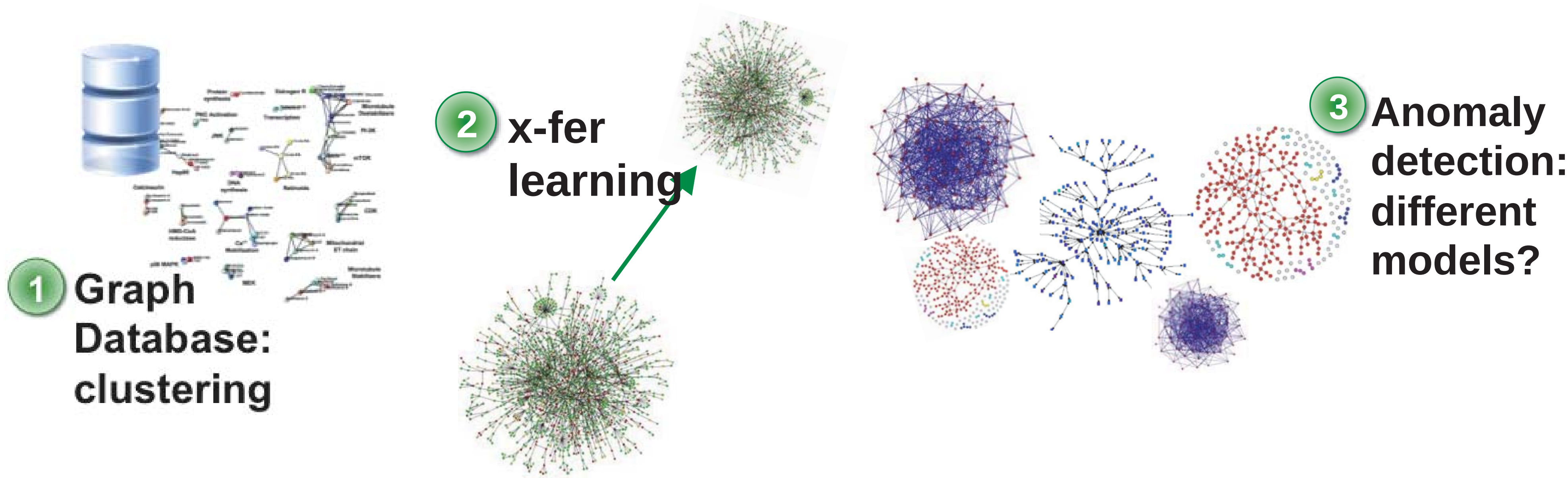
- Presented an axiomatic approach to the problem of inferring implicit social networks by measuring tie strength
- Characterized functions that satisfy all the axioms
- Classified prior measures according to the axioms that they satisfy
- Demonstrated coverage of and conflict with axioms
    - See paper for experiments on correlations among tie-strength measures
- In ranking applications, the axioms are equivalent to a natural partial order
- Mentioned application to a real-world problem
- Future work: Axiomatization of network similarity

Roadmap

- A theoretical guide to tie-strength measures
- An empirical guide to network-similarity methods
  - Sucheta Soundarajan, Tina Eliassi-Rad, and Brian Gallagher: A Guide to Selecting a Network Similarity Method. In *SIAM SDM* 2014.
    - http://eliassi.org/papers/soundarajan-sdm14.pdf

Network Similarity

- An important task with many applications



- How should we measure similarity between two or more networks (without known node-correspondence)?

Many Methods Exist for Comparing Networks

Density

Community Structure

Transitivity

NetSimile

DEGREE

Leadership-Bonding-Diversity

Eigenvalues

RWR

...and many more!

How do we pick one?

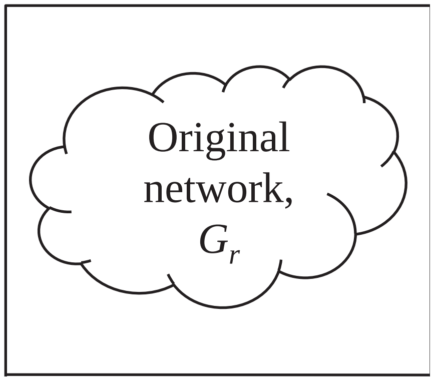
Approaches

- **Theoretical:** Use an axiomatic approach to sort through the different methods
  - Very hard
  - What does similarity of 0 mean?
- **Empirical:** Compare a large variety of network similarity methods and investigate how different are the results [SDM'14]

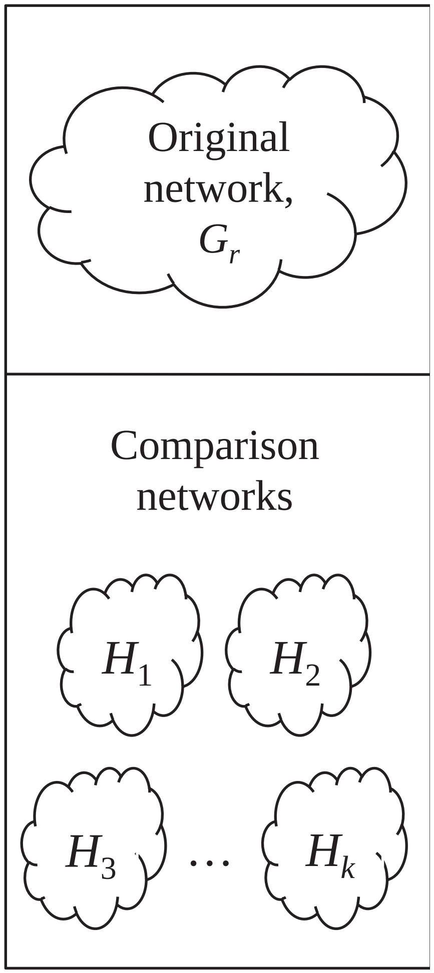
Questions

1. How correlated are different network-similarity methods to each other?
2. How can one automatically find groups of methods that behave comparably?
3. How can one select a single consensus method from a group of network-similarity methods?

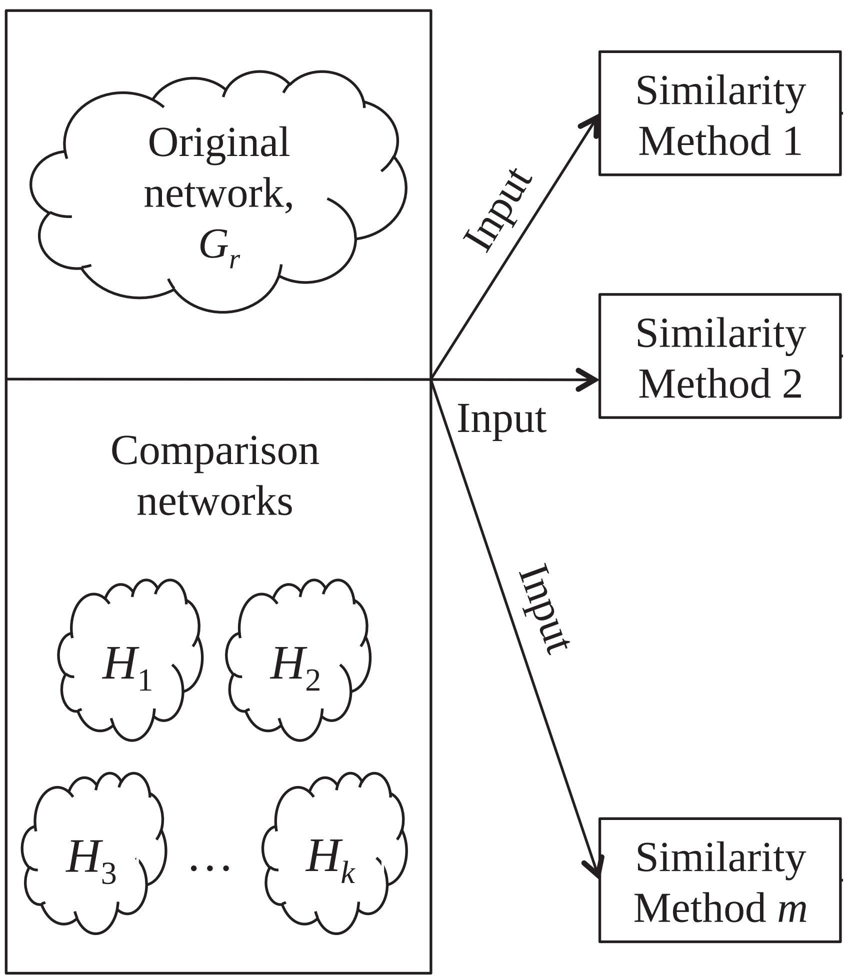
Our Approach



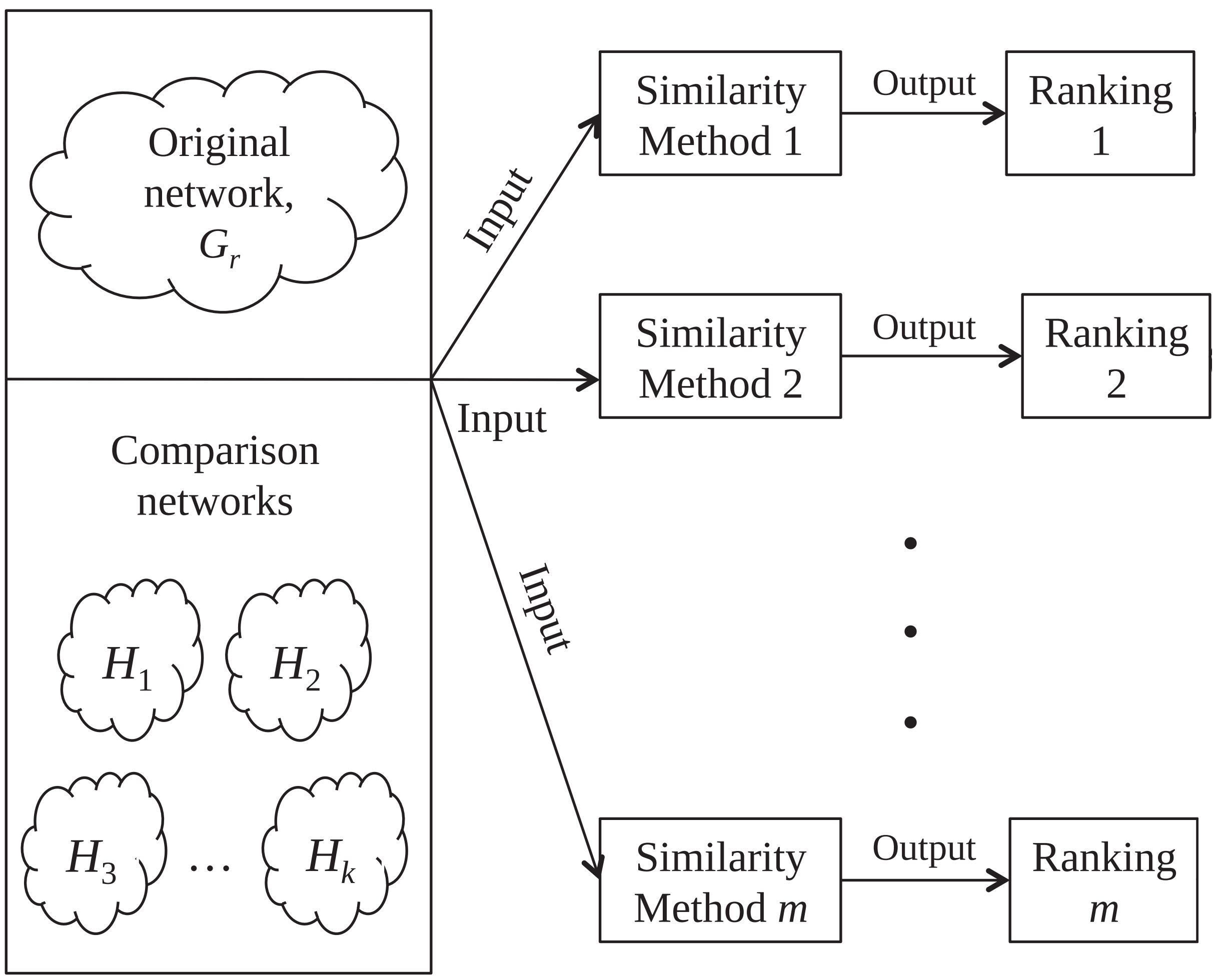
Our Approach



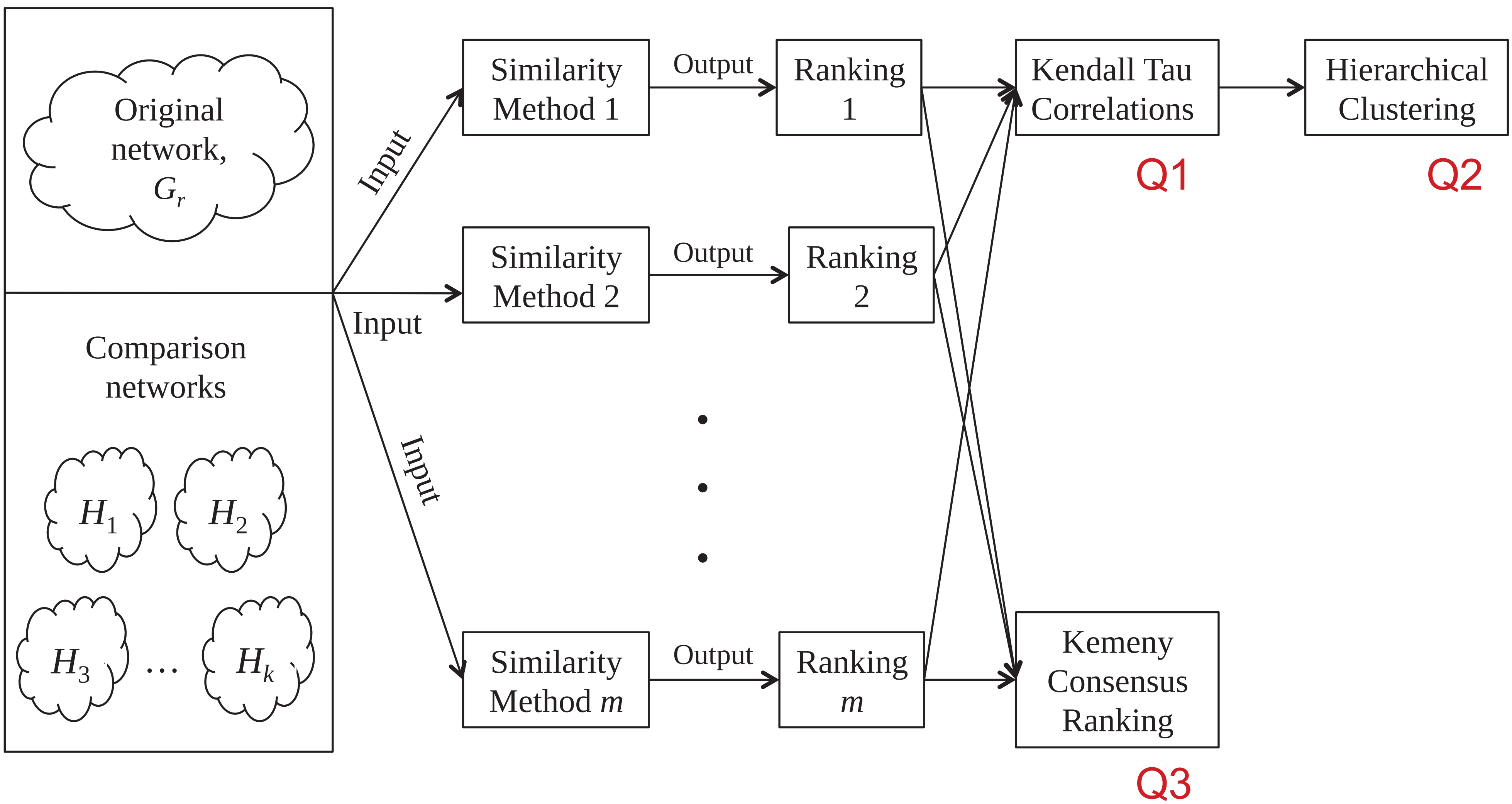
Our Approach



Our Approach



Our Approach

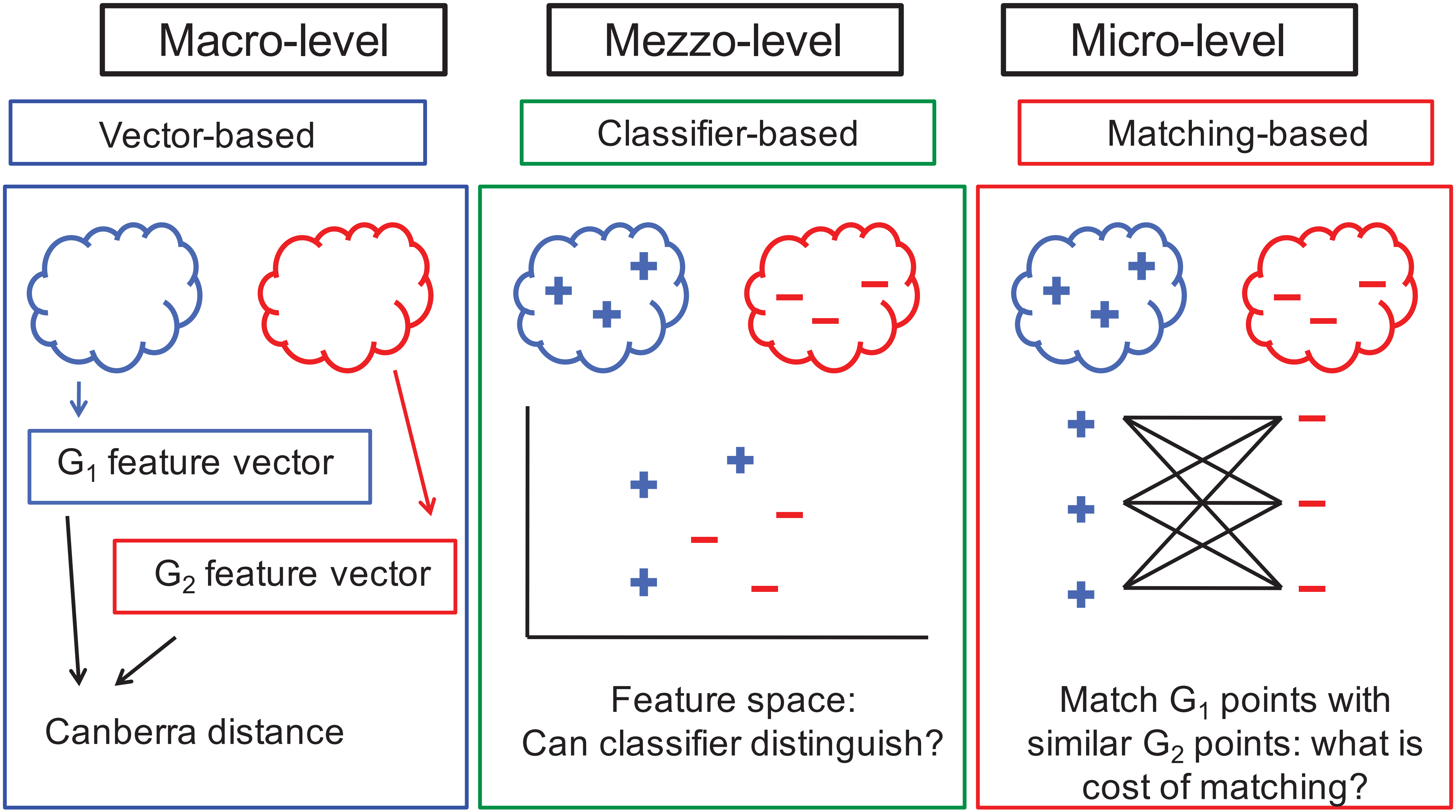


Network Similarity Methods Categorization

1. At what level of the network does it operate?
   - Micro
   - Mezzo
   - Macro

2. What type of comparison does it use?
   - Vector-based
   - Classifier-based
   - Matching-based

Network Similarity Methods



Network Similarity Methods

	Micro-level	Mezzo-level	Macro-level
Vector-based	1. NetSimile	4. Random Walk Distances 5. InfoMap-In 6. InfoMap-Known 7. InfoMap-In&Known	16. Degree 17. Density 18. Transitivity 19. Eigenvalues 20. LBD
Classifier-based	2. NetSimileSVM	8. AB 9. BFS 10. Random Walk (RW) 11. Random Walk with Restarts (RWR)	--
Matching-based	3. NetSimile-Match	12. AB-Match 13. BFS-Match 14. RW-Match 15. RWR-Match	--

Data

Dataset	# Nodes	# Edges	Avg. Degree	Density	Transitivity	# Comps
Grad	500	3000	13	0.03	0.43	2
Undergrad	1200	43K	71	0.06	0.24	1
DBLP	740K	2.5M	7	0.00001	0.21	3,840
Amazon	270K	740K	5	0.00002	0.23	36,444
LJ1	500K	10M	43	0.0001	0.04	1
LJ2	500K	10M	43	0.0001	0.08	1
Enron	37K	180K	10	0.0003	0.09	1,065

Experimental Setup

- Use various methods to rank the similarity of a set of networks, {DBLP, Amazon, …}, to a given network, Grad
- Include two "baseline" versions of each graph

Similarity Ranking to Grad

BFS	NetSimile	…
GradReWr	GradDel	…
GradDel	GradReWr	…
DBLP	Amazon	…
Amazon	DBLP	…
Email	Email	…
LJ2	Undergrad	…
Undergrad	LJ2	…
LJ1	LJ1	…

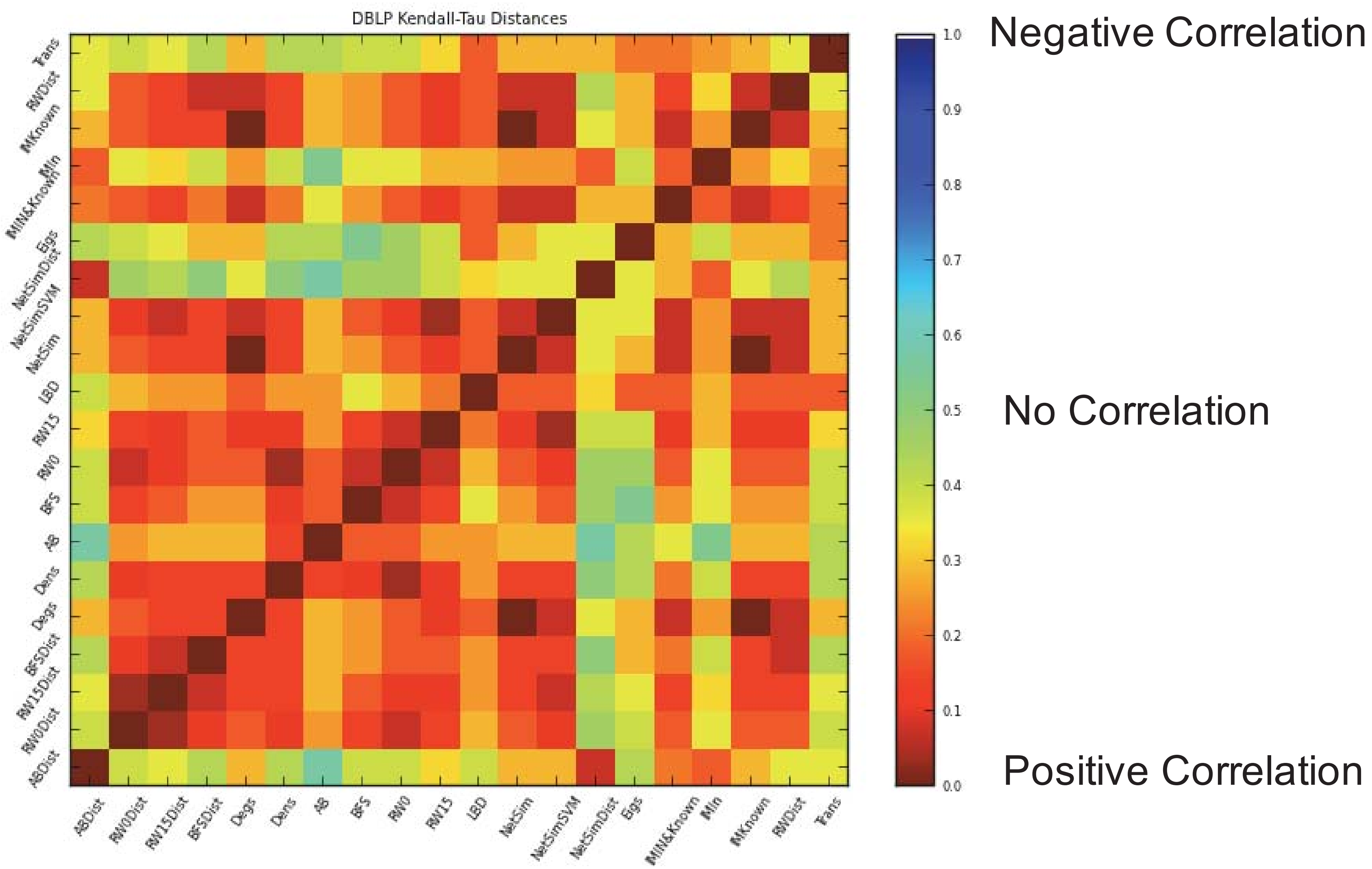
Q1. How correlated are different network-similarity methods to each other

- **Solution:** Kendall-Tau Rank Distances
- For each pair of items:
  - If they are in same order in both lists, 0 points

  - If they are in opposite order, 1 point

- Take average over all pairs
- Distance of 0 means same ordering
- Distance of 1 means inverse ordering

Example: DBLP Kendall-Tau Distances



Q2. How to automatically find groups of methods that behave comparably

- **Solution:** Agglomerative hierarchical clustering (complete-linkage)
- Which clusters appear in many datasets over 1000 runs?

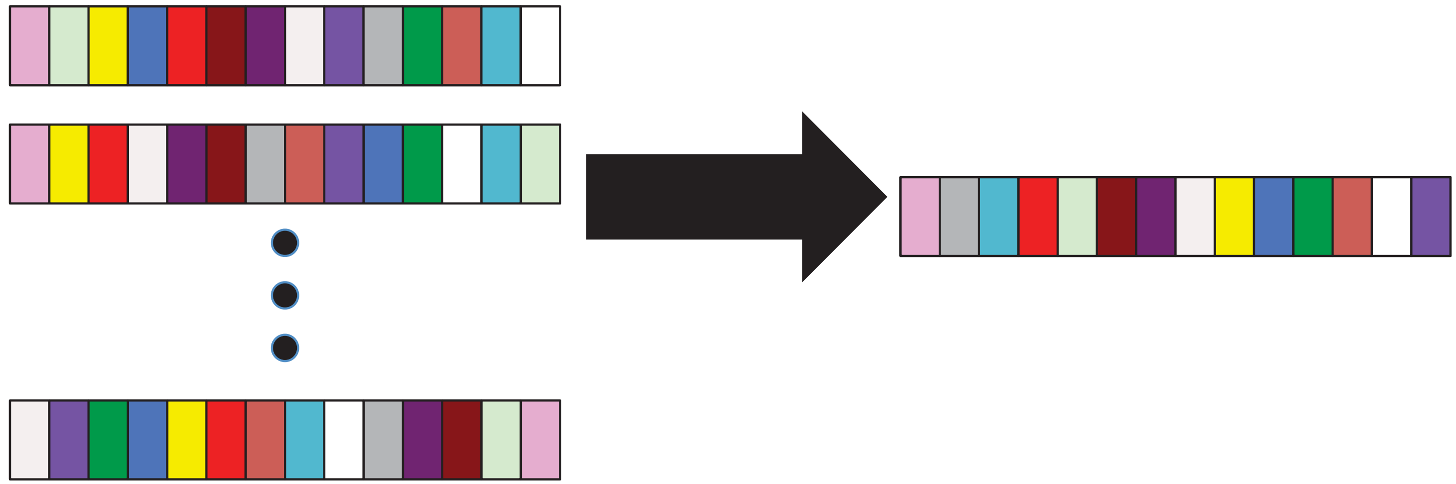
InfoMap-In&Known, InfoMap-Known
RW-Dist, RWR-Dist
RW, RWR, BFS, NetSimileSVM
LBD, Transitivity
NetSimileDist, InfoMap-In
RW-Dist, RWR-Dist, BFS-Dist, Density

Legend

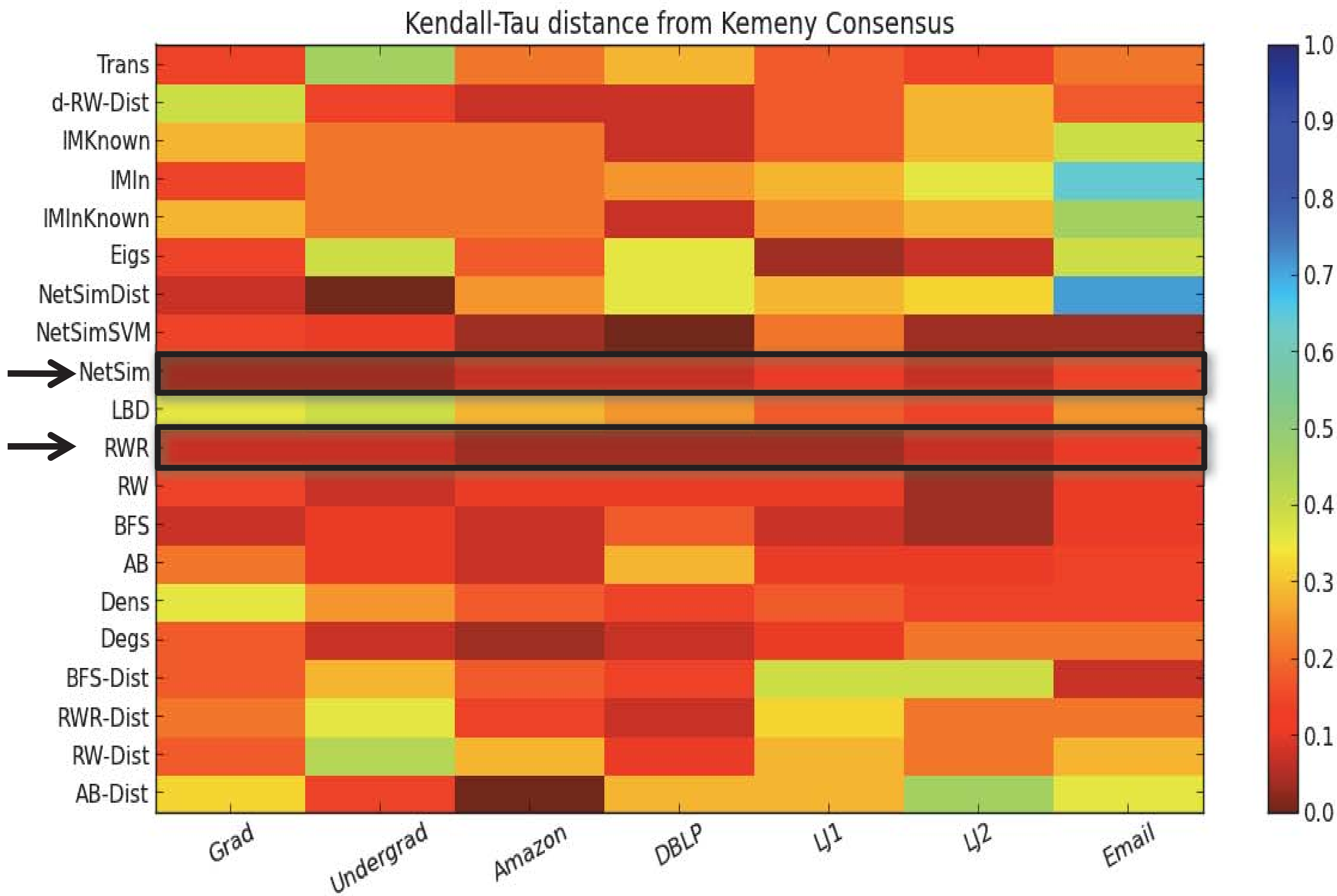
Vector-based
Classifier-based
Matching-based

Q3. How to select a consensus method from many rankings?

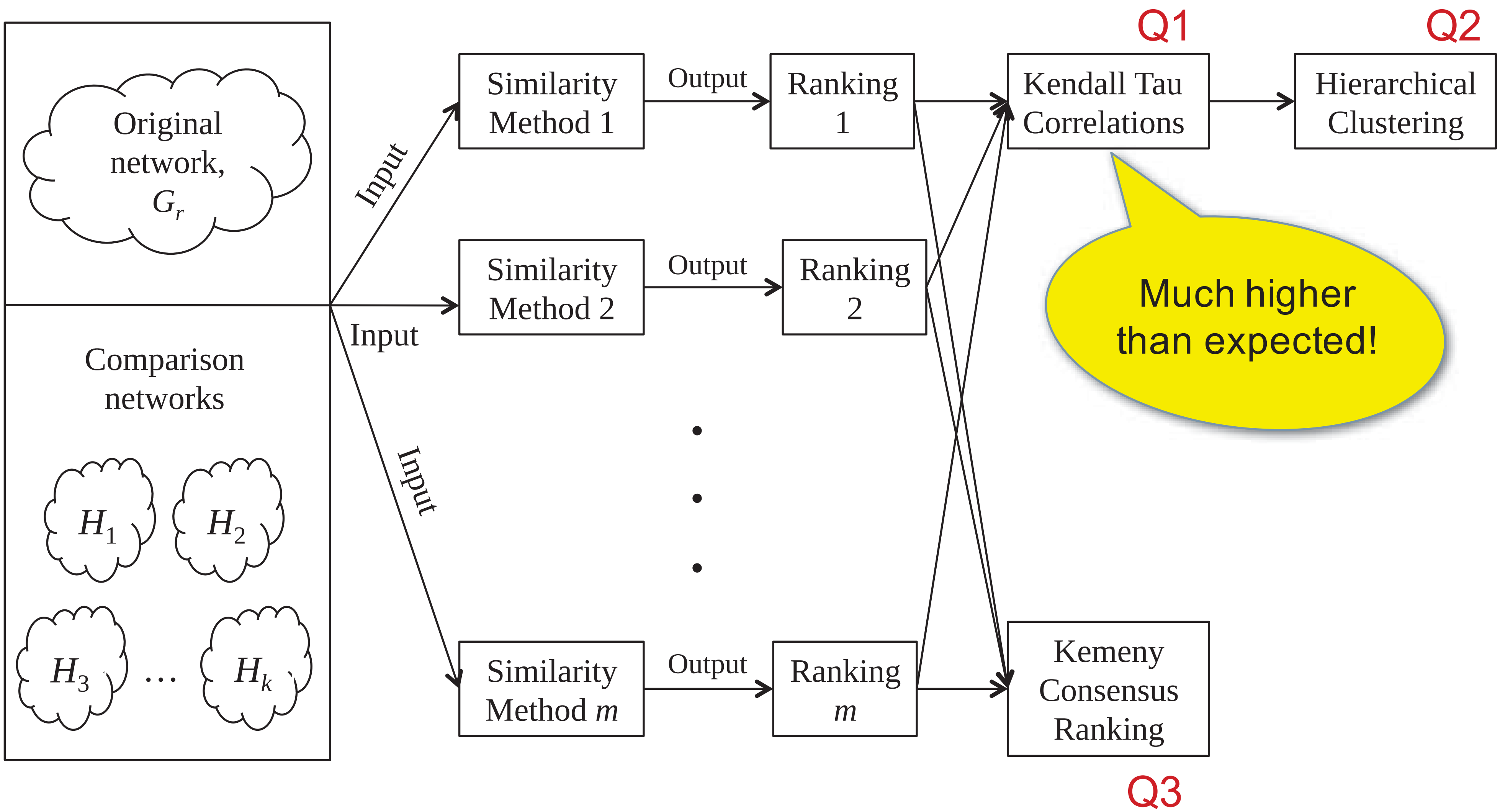
- **Solution:** Kemeny-Young method for ranking aggregation
- Assume rankings are noisy estimates of “true” ranking
- Consensus is the maximum likelihood estimator for “true” ranking
- Can be interpreted as “median” vote



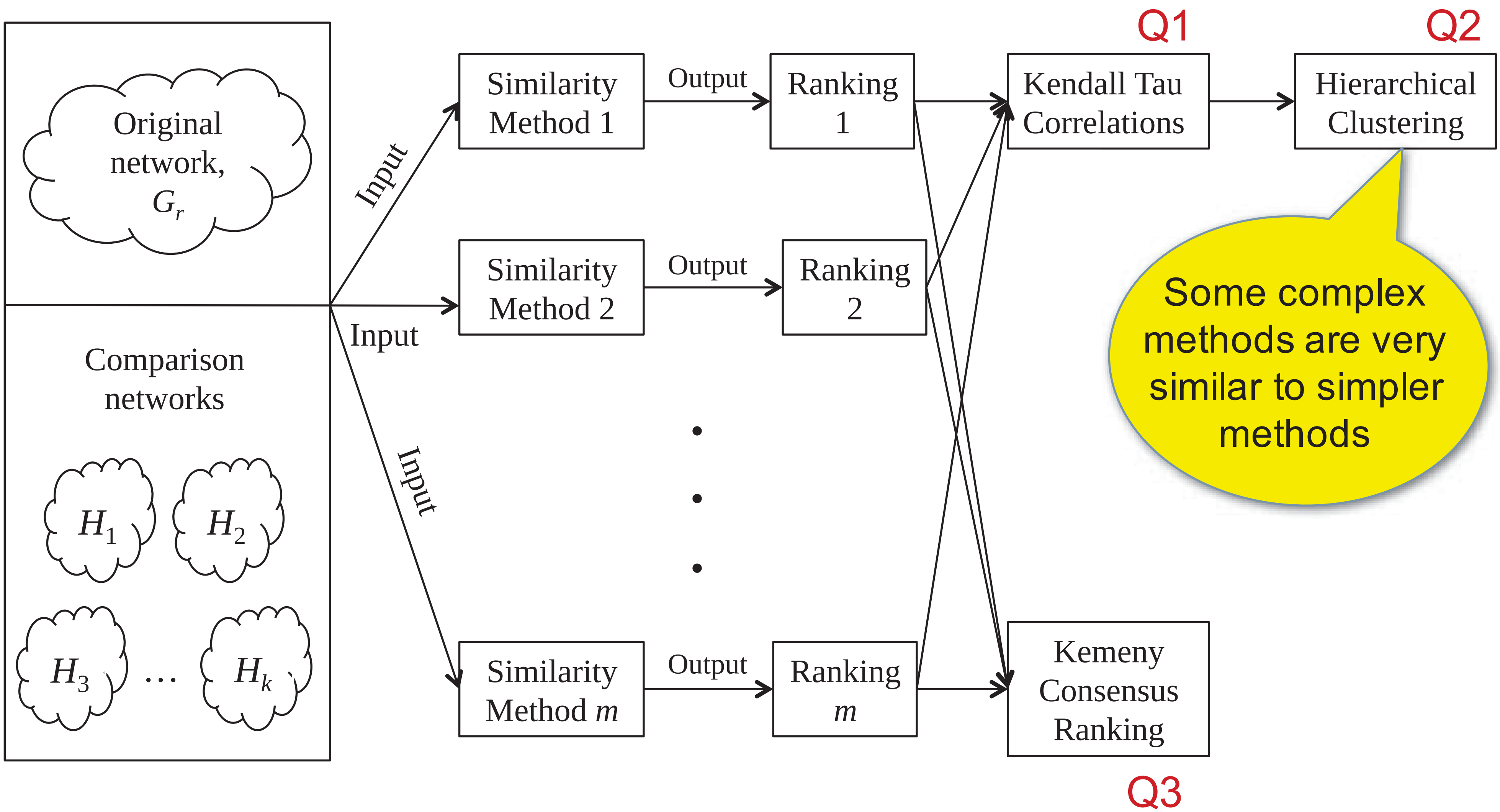
Which Methods are Closest to Consensus?



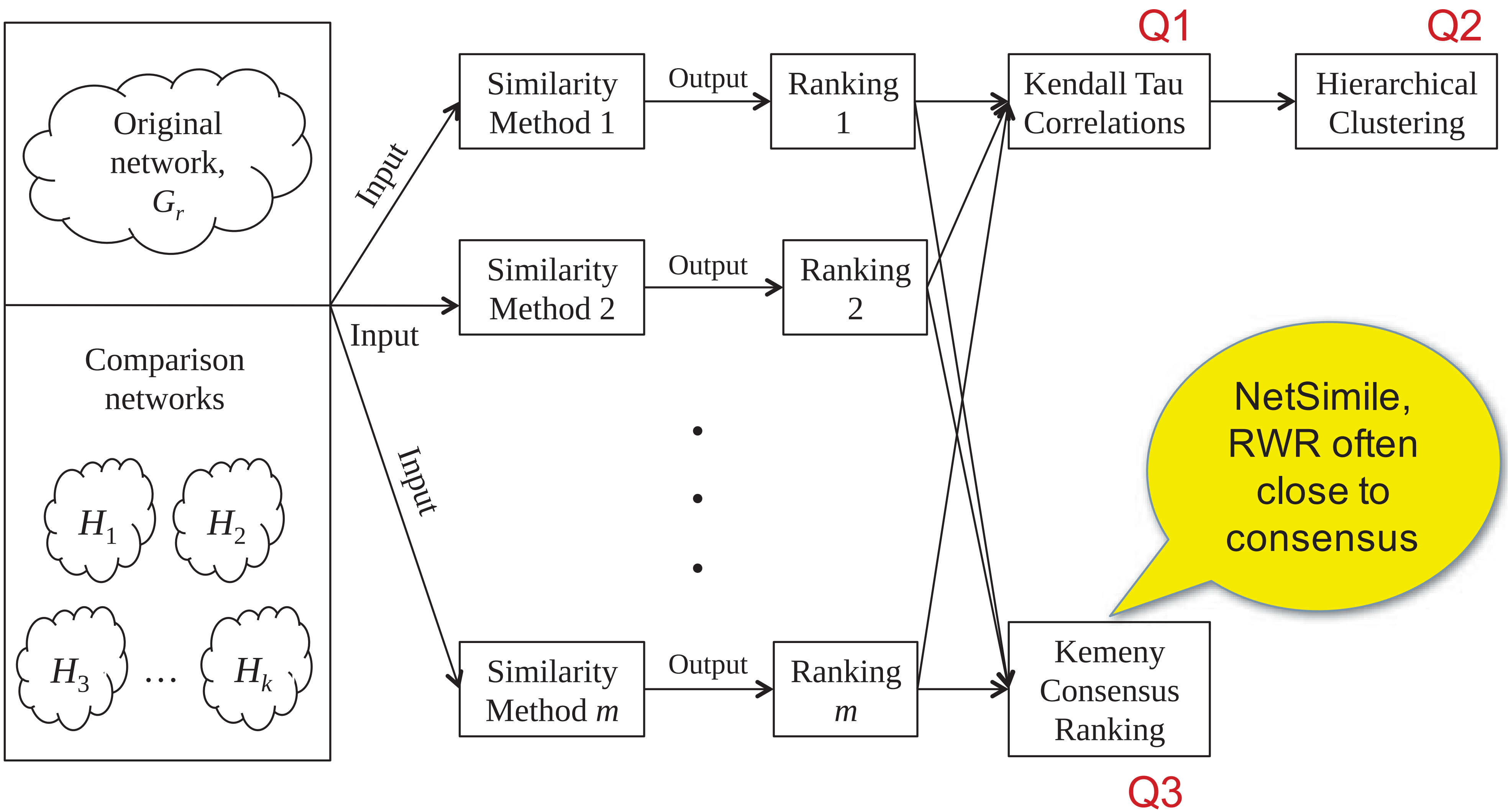
Wrap-up of Network Similarity



Wrap-up of Network Similarity

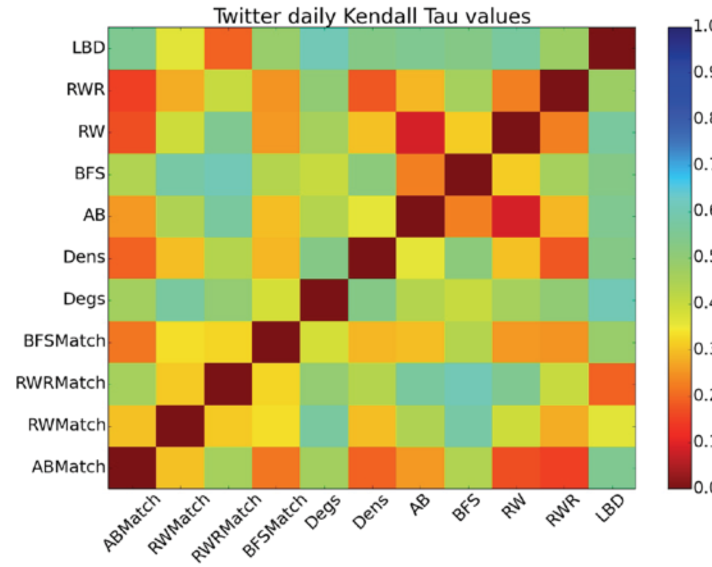


Wrap-up of Network Similarity

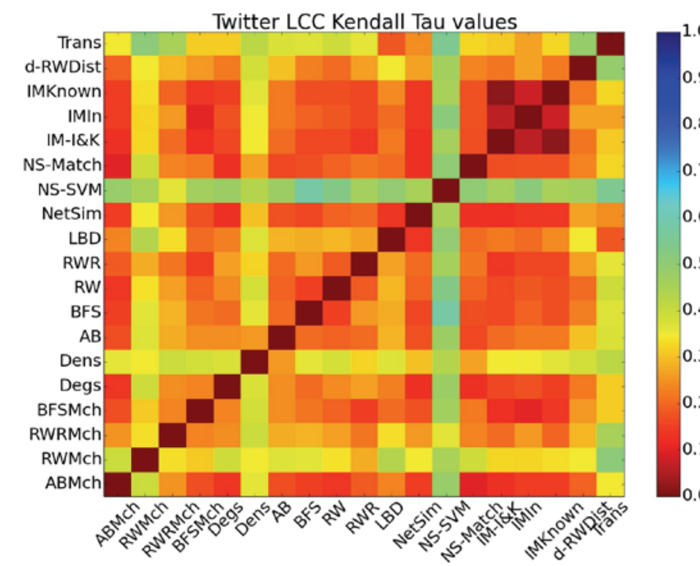


Network Similarity in Longitudinal Networks

- Did not see as much correlations in rankings
- Why?
  - Naïve aggregation of time-stamped edges into networks produces sets of dyads and not “structurally mature” networks
- ADAGE: a systematic framework for generating graph snapshots from streaming edge data
  - Details in our *WWW* 2016 poster
  - http://eliassi.org/papers/soundarajan-www16poster.pdf



(a)



(b)

Recent Tutorial on …

Node and Graph Similarity: Theory and Applications

- At *IEEE ICDM* 2014, with Danai Koutra and Christos Faloutsos
- Tutorial abstract and slides at http://web.eecs.umich.edu/~dkoutra/tut/icdm14.html

Our work inspired others to sift through functions on networks

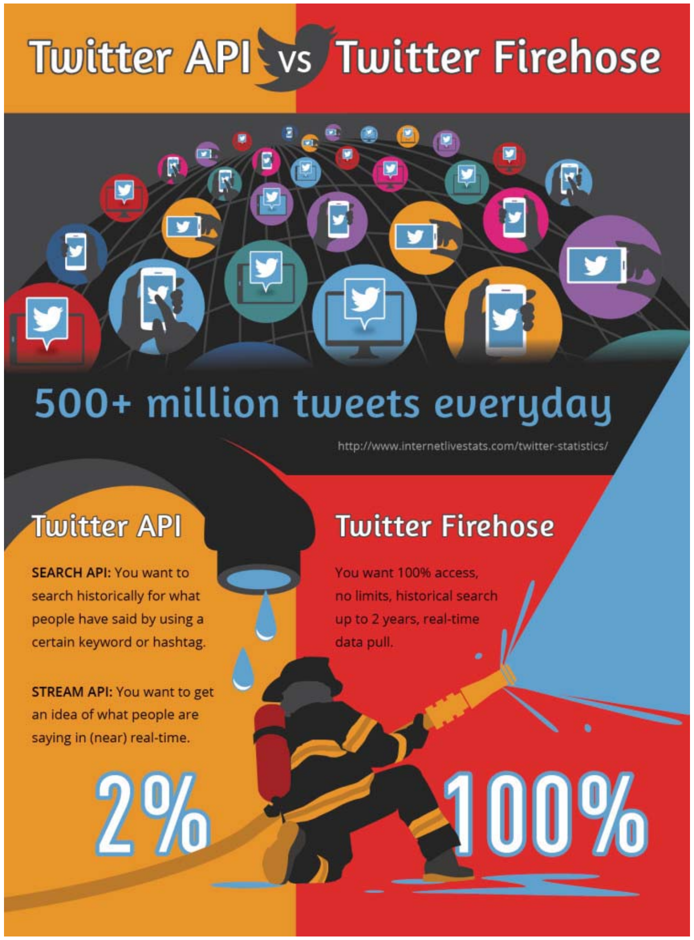
- Mirza Basim Baig and Leman Akoglu: Correlation of Node Importance Measures: An Empirical Study through Graph Robustness. In *WWW* 2015 (Web Science Track).
  - http://www3.cs.stonybrook.edu/~leman/pubs/15-www-node-centrality-correlations.pdf

End of the Road

- Measuring tie strength in implicit social networks
  - An axiomatic approach
- A guide to selecting a network similarity method
  - An empirical guide

What's Next: Incomplete Networks

- Networked representations of real-world phenomena are often partially observed
- Where does the partial observability come from?
  - Data comes in increments
  - Data is expensive or difficult to collect
  - Even when your data is complete, you may not have the computational resources to examine all of the data



What's Next? Incomplete Networks

- **Research question:** Given my networked representation, how do I increase the observability of the phenomenon under study?
- Two approaches: model aware vs. model agnostic
- Preliminary work is promising
  - NIPS Workshop on Networks 2015
    - http://arxiv.org/pdf/1511.06463v1.pdf
  - SIAM SDM 2016 Tutorial
    - http://eliassi.org/sdm16tut.html
  - IEEE/ACM ASONAM 2016
    - Camera-ready forthcoming at http://eliassi.org/pubs.html
  - Manuscript on using multi-armed bandits for this problem currently under review

What's Next? Culture Analytics

- UCLA IPAM Long Program on Culture Analytics
  - http://www.ipam.ucla.edu/programs/long-programs/culture-analytics/

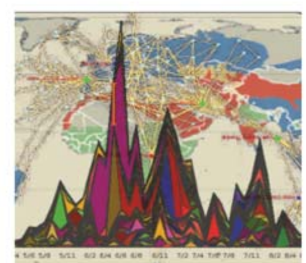
Culture Analytics

MARCH 7 - JUNE 10, 2016

OVERVIEW | PARTICIPANT LIST | SEMINAR SERIES | ACTIVITIES

Overview

The explosion in the widespread use of the Internet and social media and the ubiquity of low cost computing have increased the possibilities for understanding cultural behaviors and expressions, while at the same time have facilitated opportunities for making cultural artifacts both accessible and comprehensible. The rapidly proliferating digital footprints that people leave as they crisscross these virtual spaces offer a treasure trove of cultural information, where culture is considered to be expressive of the norms, beliefs and values of a group. This program encourages the exploration of the unsolved mathematical opportunities that are emerging in this cultural information space. Many successful approaches to the analysis of cultural content and activities have been developed, yet there is still a great deal of work to be done. In this program, we aim to promote a vigorous collaboration across disciplines and devise new approaches and novel mathematics to address these problems of culture analytics, by bringing together leading scholars in the social sciences and humanities with those in applied mathematics, engineering, and computer science.



ORGANIZING COMMITTEE

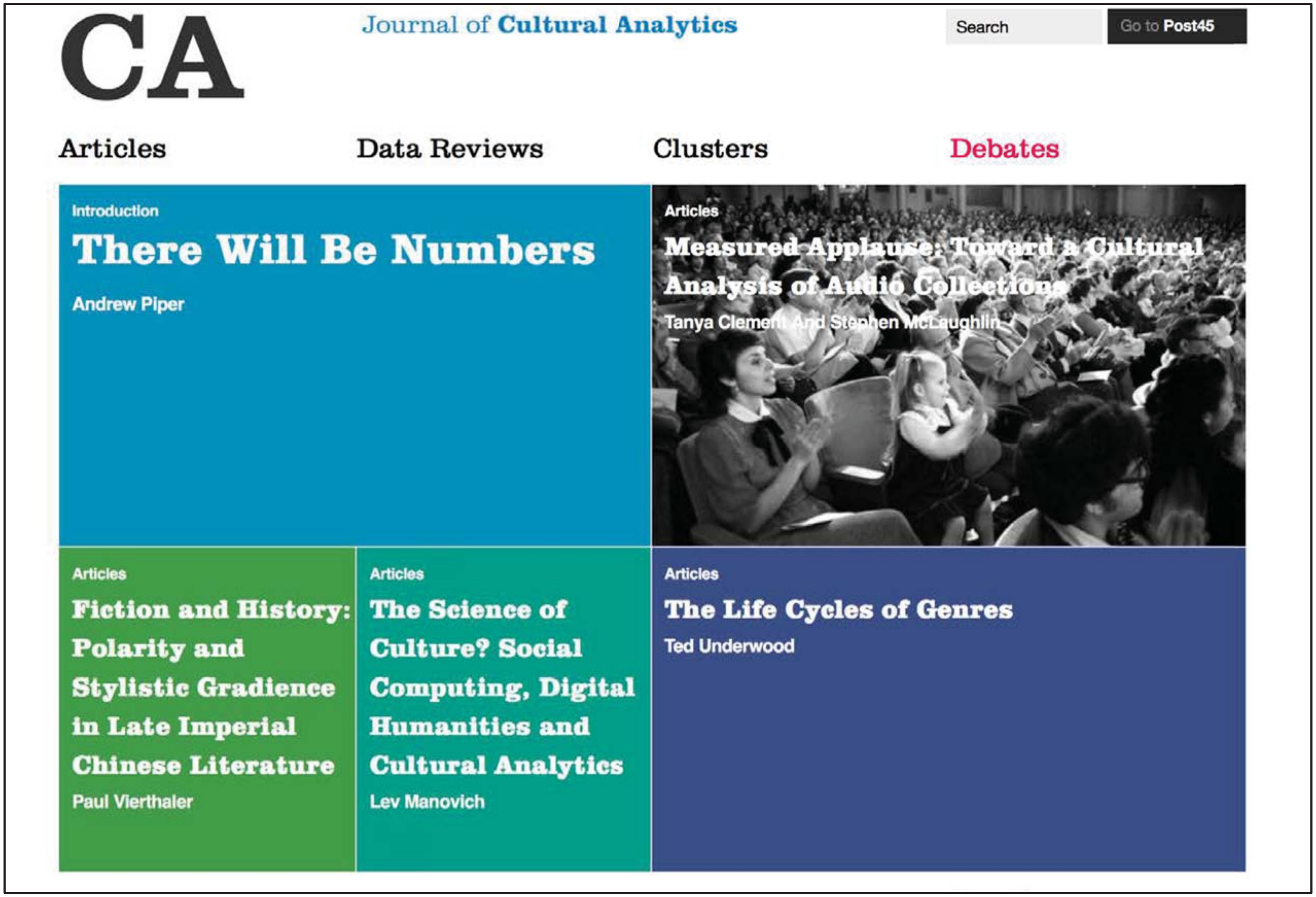
Tina Eliassi-Rad (Rutgers University, Computer Science)
Mauro Maggioni (Duke University, Mathematics and Computer Science)
Lev Manovich (The Graduate Center, CUNY, Computer Science)
Vwani Roychowdhury (University of California, Los Angeles (UCLA), Electrical Engineering)
Timothy Tangherlini (University of California, Los Angeles (UCLA), Germanic Languages and Literatures, Scandinavian Section)

What's Next? Some Problems in Culture Analytics (from IPAM Culture Analytics Organizing Committee)

- Are there axioms of culture? If so, can one develop a formal treatment of these?
- Identifying, defining and measuring cultural complexities
- Detect culturally meaningful phase transitions
- Develop a calculus of culture (what does this mean?)
- Measure the impact of culture on social conflict, inequality, public health, and the environment
- Measure the cultural impact of globalization
- In this digital age, are cultures accelerating? If so, how, by how much? Can it be precisely measured?
- Many more...

What's Next? Journal of Cultural Analytics

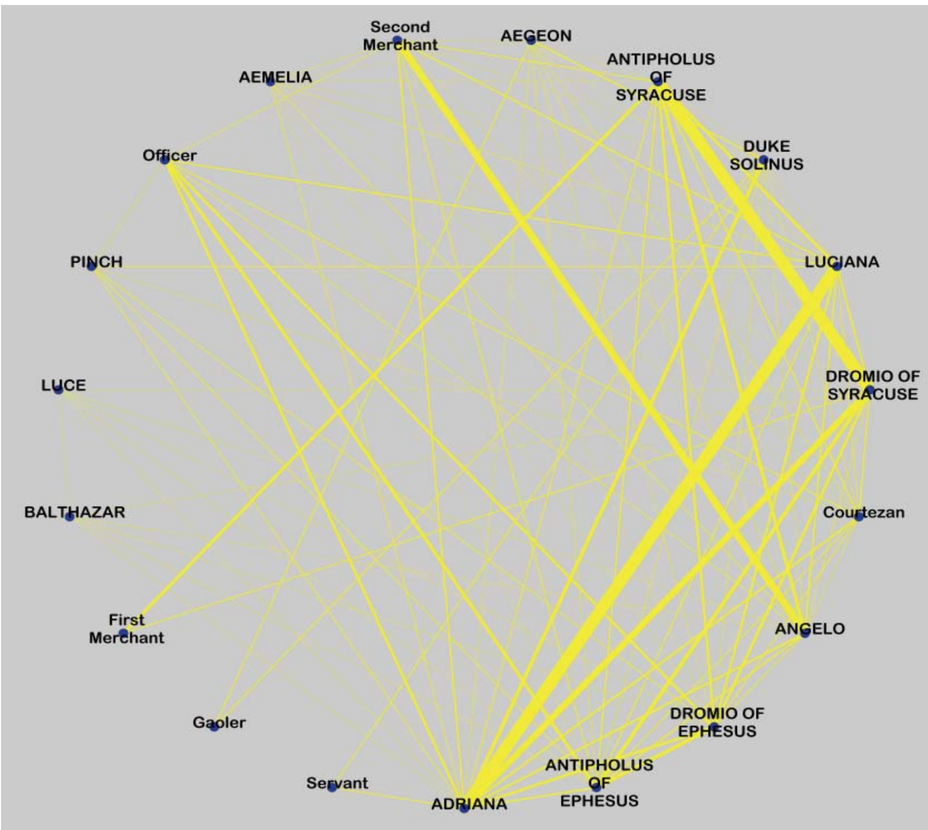
- http://culturalanalytics.org/



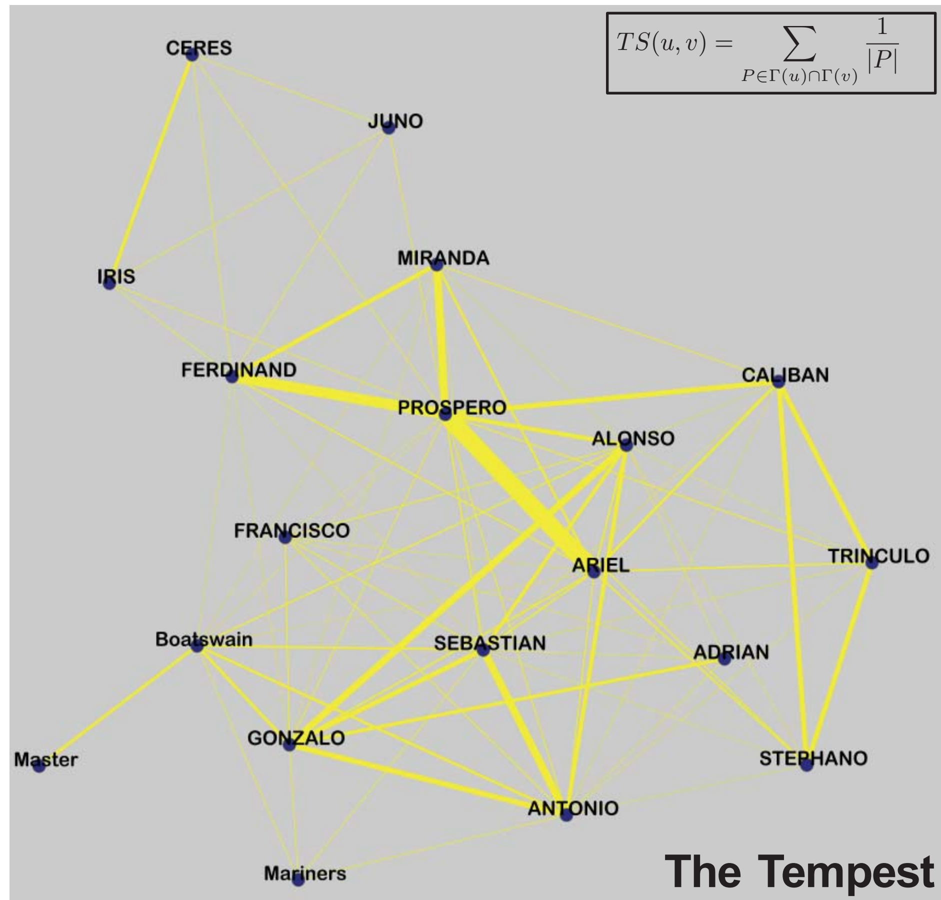
Thank You

Papers @ http://eliassi.org/pubs.html

Contact: tina@eliassi.org



A Comedy of Errors



The Tempest

Supported by NSF, DTRA, DARPA, IARPA, DOE/LLNL & WaPo Labs.