# Network Analysis:

## The Hidden Structures behind the Webs We Weave

## 17-213 / 17-668

# Graph Signatures and Dynamics of Social Ties

Tuesday, September 12, 2023

Patrick Park & Bogdan Vasilescu

Carnegie Mellon University
School of Computer Science

S3D
Software and Societal Systems Department

# 2-min Quiz, on Canvas

# Quick Recap – Last Thursday's Lecture

Structural Balance: triads of friends and enemies

But, most real world social networks are not perfectly balanced

Many different triadic relationships exist

Triadic closure – two nodes that are connected to the same set of other nodes have a higher probability of forming an edge

Q: Why do social networks exhibit triadic closure?

Local clustering coefficient (probability that two neighbors of a node are connected) measures the extent of triadic closure in a network

# Quick Recap – Last Thursday's Lecture

Edge vs. Social Tie ⟵ more today

Content of the tie can partly shape the structure of the network

Information diffusion: valued information diffuses through strong ties

Q: Will word about the exquisite cake from La Gourmandine spread like wildfire at the party?

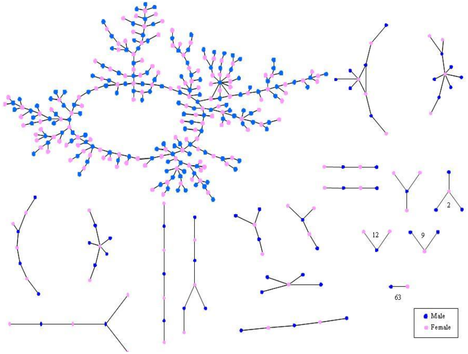
A: Not necessarily

# Case Study: Graph Signature of Social Ties

# Graph Signature of Social Ties

As the high school romantic relationship network example demonstrates, sometimes certain relationship types in specific social contexts (e.g., school) leave a visible structural marker

in high school context

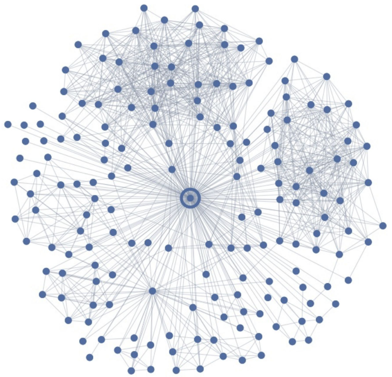


# Graph Signature of Social Ties

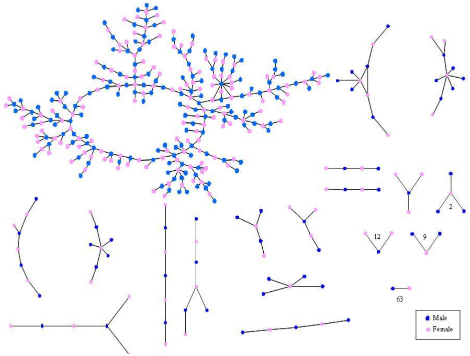
As the high school romantic relationship network example demonstrates, sometimes certain relationship types in specific social contexts (e.g., school) leave a visible structural marker

The same type of relationship can leave different structural markers in different social contexts

in Facebook



in high school context

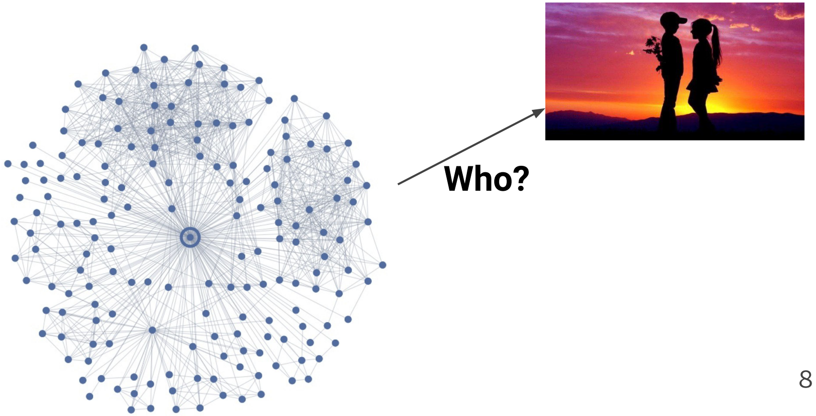


# Graph Signature of Social Ties

An Illustrative Problem:

Predict the significant other (romantic partner / spouse) of a Facebook user solely from the user's friendship graph

**Q**: Can you think of a graph characteristic that can hint at romantic partners or spouses?



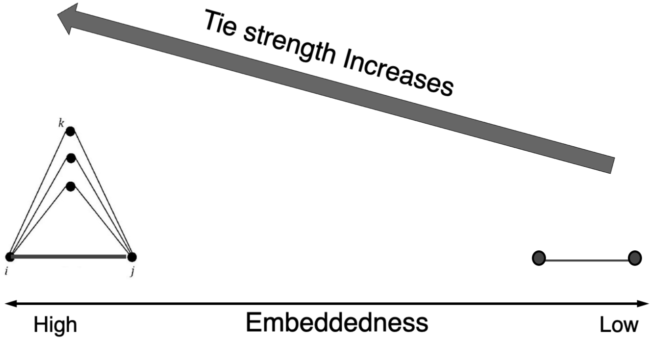
# Graph Signature of Social Ties

An Illustrative Problem:

Predict the significant other (romantic partner / spouse) of a Facebook user solely from the user's friendship graph

A network analyst who learned about strong ties and triadic closure may reason:

- A social tie that is highly embedded tends to be strong



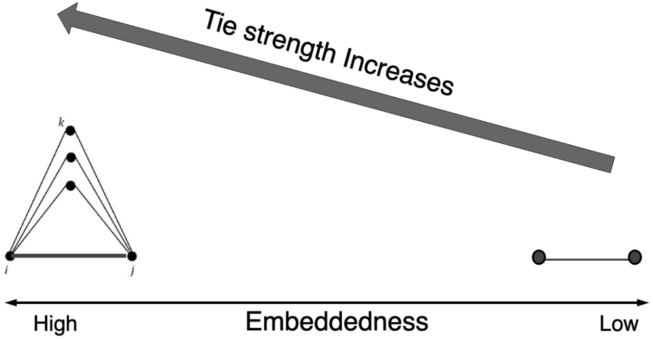
# Graph Signature of Social Ties

The Problem:

Predict the significant other (romantic partner / spouse) of a Facebook user solely from the user's friendship graph

A network analyst who learned about strong ties and triadic closure may reason:

- A social tie that is highly embedded tends to be strong
- A partner is one of the strongest ties who knows many friends of the partner



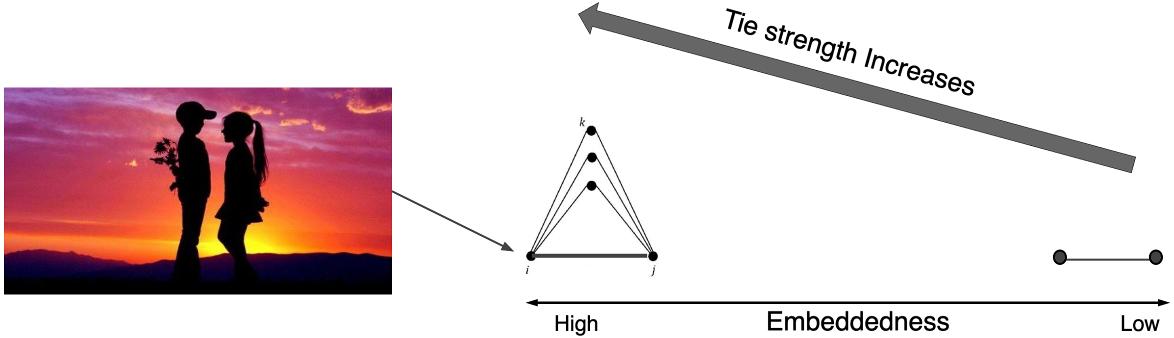
# Graph Signature of Social Ties

The Problem:

Predict the significant other (romantic partner / spouse) of a Facebook user solely from the user’s friendship graph

A network analyst who learned about strong ties and triadic closure may reason:

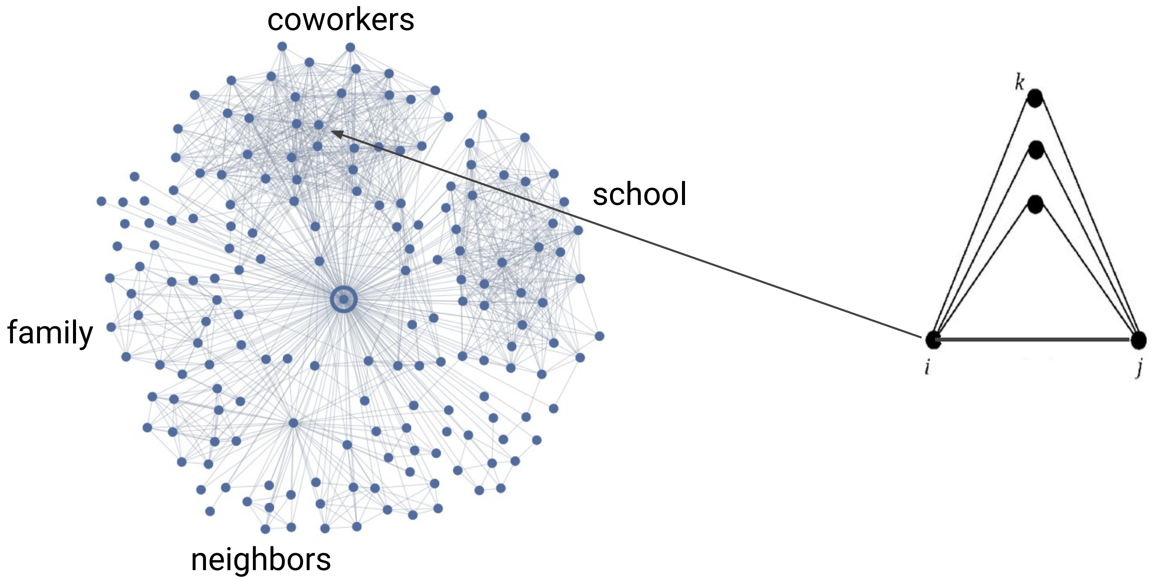
- A social tie that is highly embedded tends to be strong
- A partner is one of the strongest ties who knows many friends of the partner
- Therefore, the node with **highest embeddedness** is likely to be the partner



# Graph Signature of a Significant Other

In practice, the friend with highest embeddedness is someone who is highly connected in the largest cluster

- Example: coworker, college friend, often not the significant other



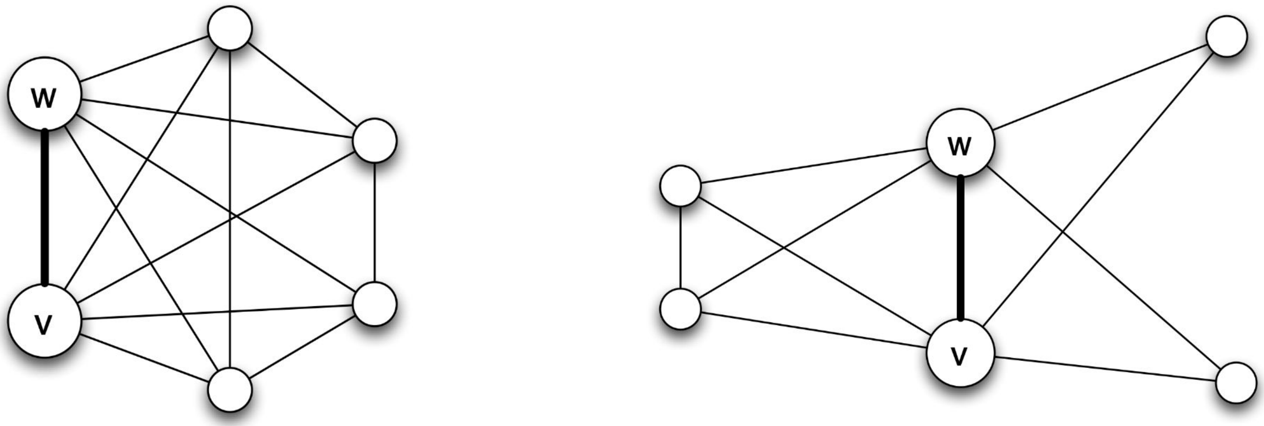
# Graph Signature of a Significant Other

Backstrom and Kleinberg draw insight from the psychology literature on the characteristics of intimate ties

- a sense of intimacy, voluntary investment in the companionship
- an interest in **being together** as much as possible through interactions **in multiple social contexts** over a long period
- a sense of **mutuality** and support for partner's needs

They focus on the fact that many couples are together in multiple social contexts

# Graph Signature of a Significant Other

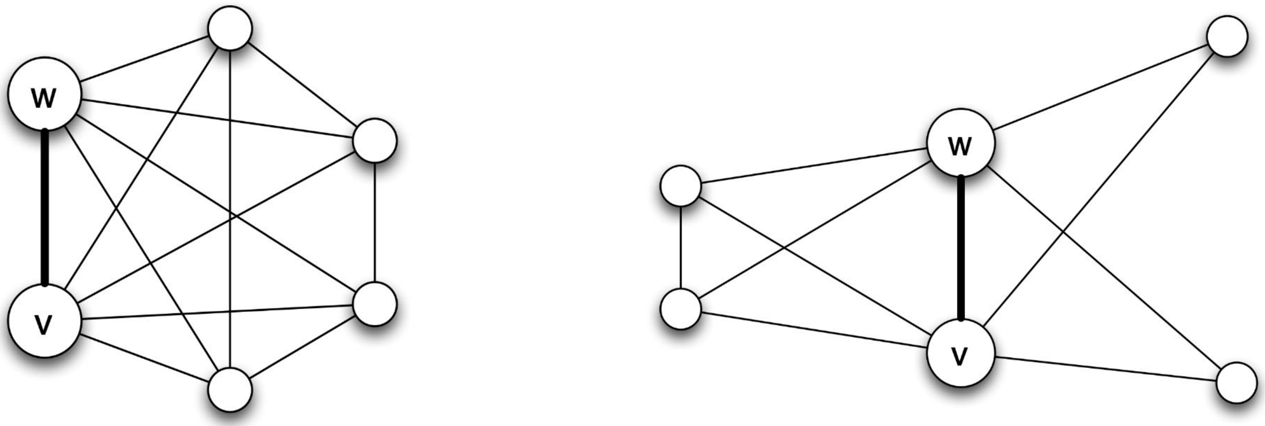



Instead of just counting mutual friends, look at their structure.

- How well connected are the common endpoints of edge $e$?
- If not well connected, suggests something about $v$-$w$ relationship.
- $v$-$w$ cannot be easily "explained" by any one social focus.

Source: [Jon Kleinberg's slide presentation]

# Graph Signature of a Significant Other

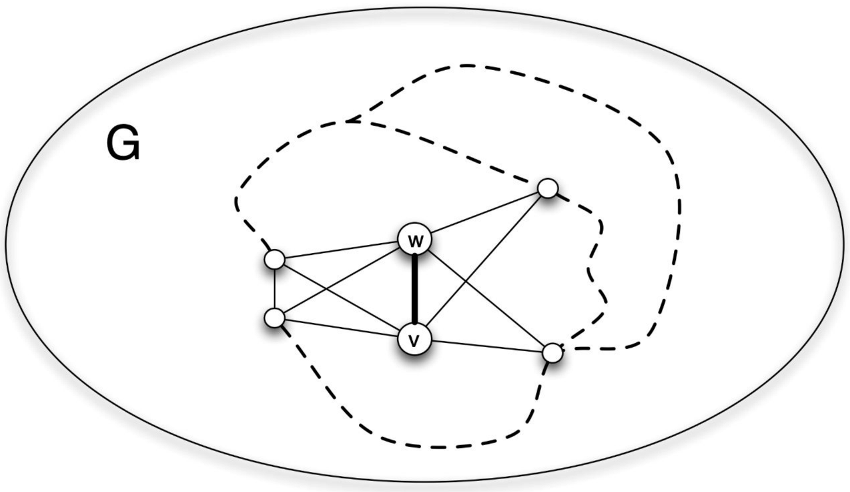



*w-v* tie on the left is highly embedded, but in a single, social context

*w-v* tie on the right participates in three different social contexts

Together, they constitute a local bridge connecting these different contexts

Intuitively, the tie on the right is more likely to be partners




$C_{vw}$ = common neighbors of $v$ and $w$.

Sum of distances between pairs in $C_{vw}$, after deleting $v$ and $w$:

$$\sum_{s,t \in C_{vw}} d_{G - \{v,w\}}(s, t).$$

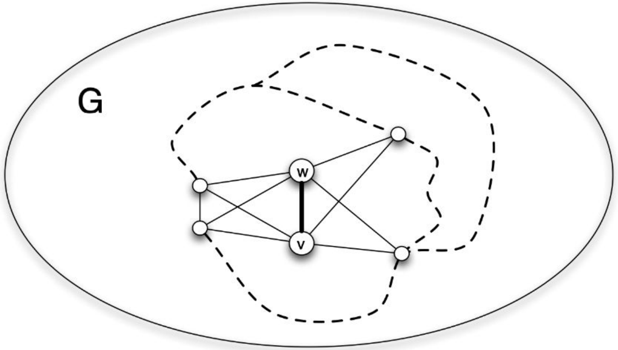
The <u>dispersion</u> of edge $(v, w)$ with respect to distance function $d$.

- Should use 0-1-valued metric; normalize by $|C_{vw}|$.



Source: [Jon Kleinberg's slide presentation]

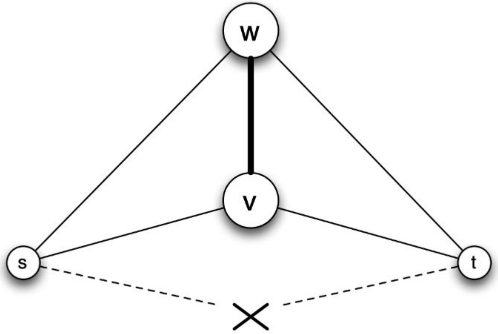
Can use many possible functions $d$.
$disp(v, w) = \sum_{s,t \in C_{vw}} d_{G - \{v,w\}}(s, t)$.




- $d(s, t) = \begin{cases} 0 \text{ if } (s, t) \text{ is an edge} \\ 1 \text{ otherwise} \end{cases}$
- $d(s, t) = \begin{cases} 0 \text{ if shortest } s\text{-}t \text{ path avoiding } v, w \text{ has} \leq k \text{ edges} \\ 1 \text{ otherwise} \end{cases}$

Can also normalize the dispersion: $\dfrac{disp(v, w)}{|C_{vw}|^{\alpha}}$.

- Analogue of clustering coefficient [Watts-Strogatz 98] is $k = 1$ and $\alpha = 2$.
- Searching over choices of $k, \alpha$ shows $k = 2$ and $\alpha = 1$ nearly optimal.




Source: [Jon Kleinberg's slide]

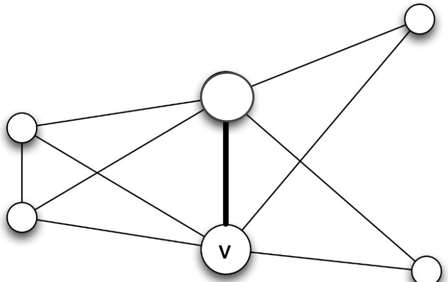
# Evaluating the Methods

For evaluation, use 1.3 million Facebook users who:

- Declare a relationship partner in their profile (symmetric).
- Have between 50 and 2000 friends.
- Are at least 20 years old.

For each user $v$, rank all friends $w$ by competing metrics:

- Embeddedness of $v$-$w$ edge.
- Dispersion of $v$-$w$ edge.
- Number of photos in which $v$ and $w$ are both tagged.
- Number of times $v$ viewed $w$'s profile in last 90 days.




For what fraction of all users $v$ is the top-ranked $w$ the relationship partner?

Source: Jon Kleinberg's slide

A random guess for a user with 100 friends

= 1% accuracy

Highest dispersion

= 50.6% accuracy

| type | embed | dispersion | photo | profile view |
|---|---|---|---|---|
| all | 0.247 | 0.506 | 0.415 | 0.301 |
| married | 0.321 | 0.607 | 0.449 | 0.210 |
| married (female) | 0.296 | 0.551 | 0.391 | 0.202 |
| married (male) | 0.347 | 0.667 | 0.511 | 0.220 |
| relationship | 0.132 | 0.344 | 0.347 | 0.441 |
| relationship (female) | 0.139 | 0.316 | 0.290 | 0.467 |
| relationship (male) | 0.125 | 0.369 | 0.399 | 0.418 |

Notes:

- Embeddedness vs. dispersion
- Structural vs. activity-based
- Married vs. in a relationship
- Female vs. male
- Combining all via machine learning: 0.716 married, 0.682 relationship
- Approx 34-38% of dispersion's incorrect guesses are family members.

Source: Jon Kleinberg's slide

Prediction performance much higher for married couples, compared to unmarried relationships

**Why?**

| type | embed | dispersion | photo | profile view |
|---|---|---|---|---|
| all | 0.247 | 0.506 | 0.415 | 0.301 |
| married | 0.321 | 0.607 | 0.449 | 0.210 |
| married (female) | 0.296 | 0.551 | 0.391 | 0.202 |
| married (male) | 0.347 | 0.667 | 0.511 | 0.220 |
| relationship | 0.132 | 0.344 | 0.347 | 0.441 |
| relationship (female) | 0.139 | 0.316 | 0.290 | 0.467 |
| relationship (male) | 0.125 | 0.369 | 0.399 | 0.418 |

Notes:

Embeddedness vs. dispersion

Structural vs. activity-based

Married vs. in a relationship

Female vs. male

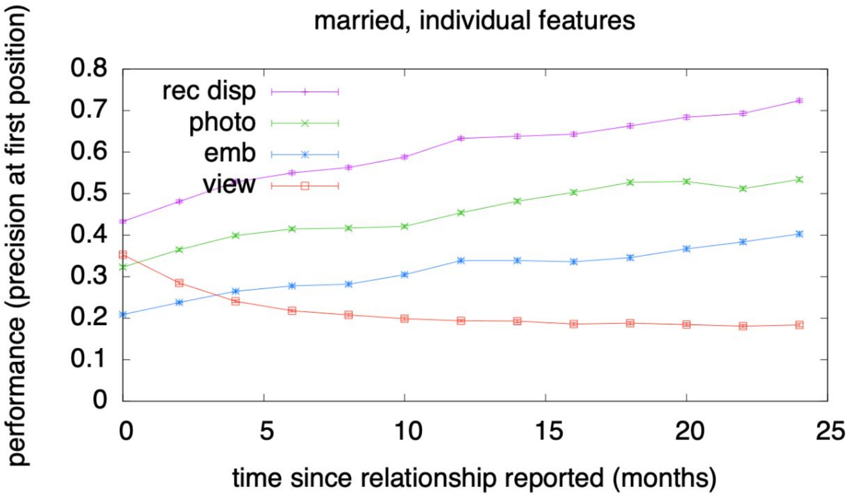
Combining all via machine learning: 0.716 married, 0.682 relationship

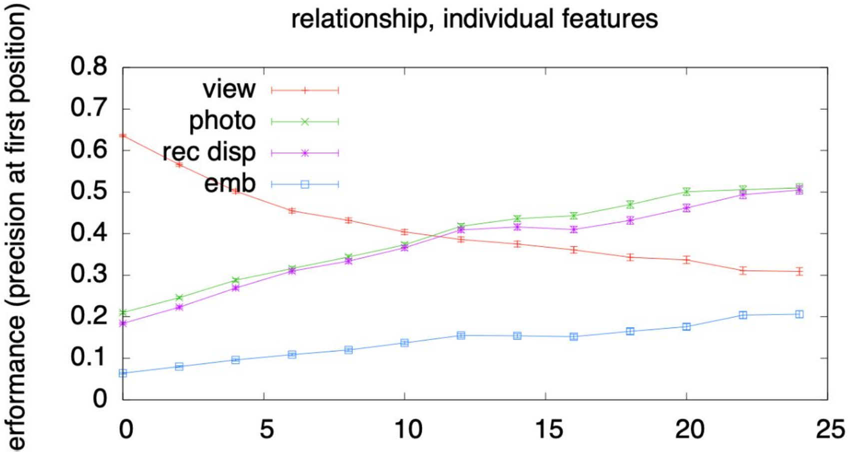
Approx 34-38% of dispersion's incorrect guesses are family members.

Source: Jon Kleinberg's slide

Because it takes time for a couple to share multiple social contexts

Recall, intimate ties have an interest in being together as much as possible through interactions in multiple social contexts **over a long period**




Source: [Jon Kleinberg's slide](#)

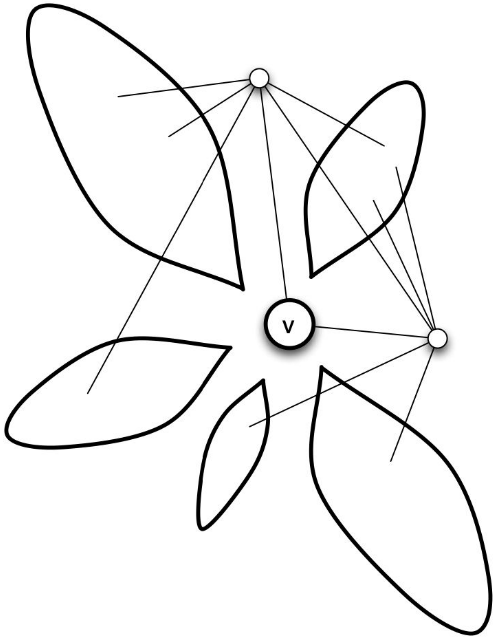
# Graph Signature of a Significant Other

So, a significant other is a person who navigates the social world with you as a single/common unit, a companion

Lesson 1: Seek insights from the social and try to map them on to quantitative features in the graph

Being together in multiple contexts→ network dispersion

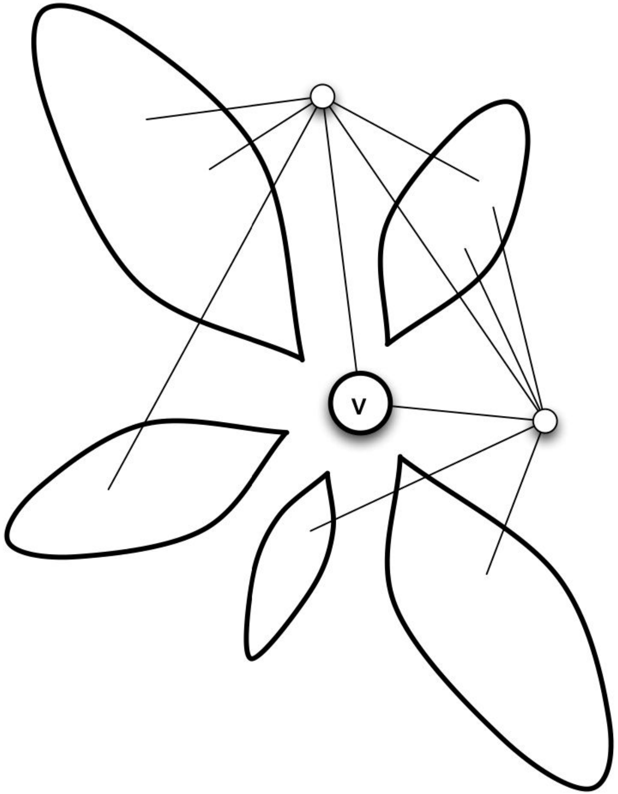
Lesson 2: Analyze those graph features and circle back to evaluate how well they capture the relationships within a social context



# Graph Signature of a Significant Other

Q: Suppose *i* and *j* are partners in real life

If *j* gets the highest dispersion score from *i*'s network, but *i* does not get the highest dispersion score in *j*'s network, what do you think this mismatch suggests of their romantic relationship?

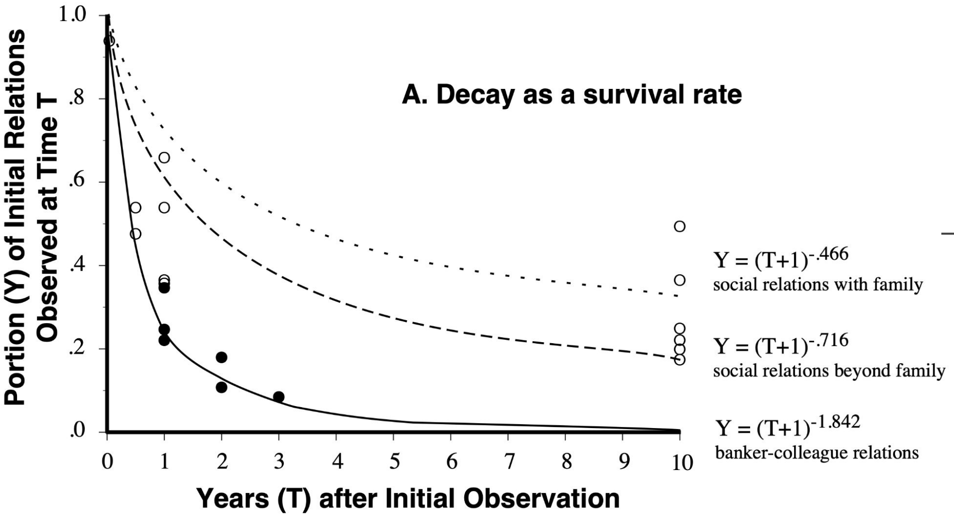


# The Dynamics of Social Ties

# Persistence and Decay of Social Ties

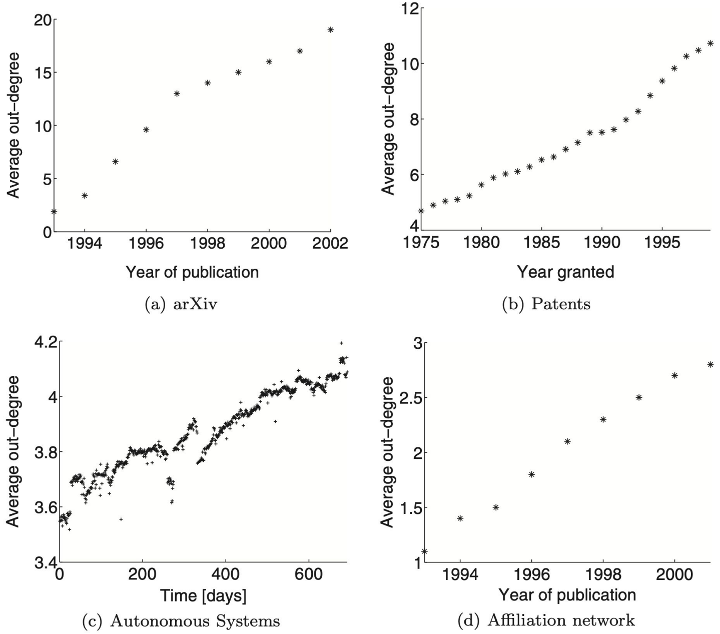
People form relationships and those relationships can persist or subside over time

The evolution of a social network is closely related to such ebbs and flows of social ties




Burt 2000

Network densification Leskovec et al. 2007



# Interdependence and Persistent Social Ties

Then what factors influence how long a tie persists (commitment to a relationship)?

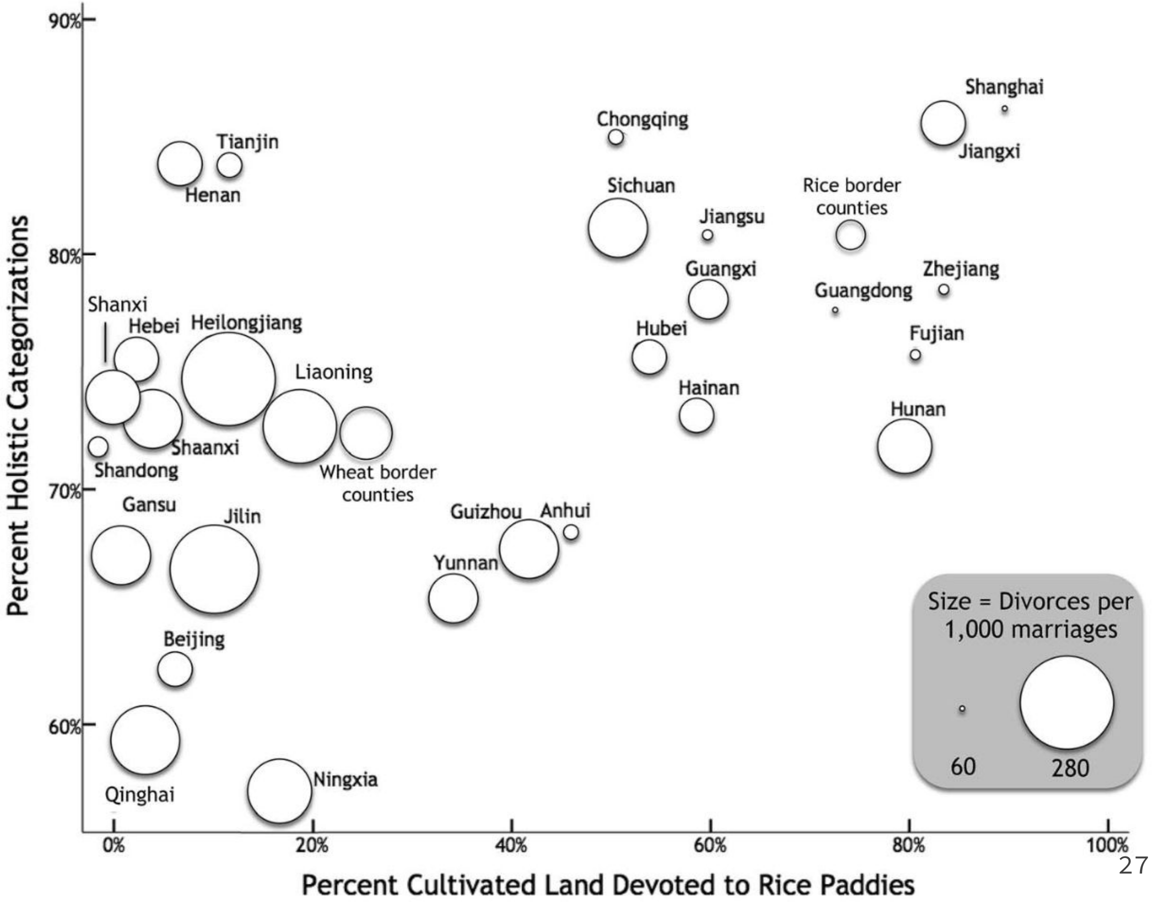
- Historically, more interdependent modes of production seems to have influenced people's thinking styles and social organization, including how people relate with one another

Rice farming requires highly interdependent, coordinated labor, compared to wheat farming

# Interdependence and Persistent Social Ties

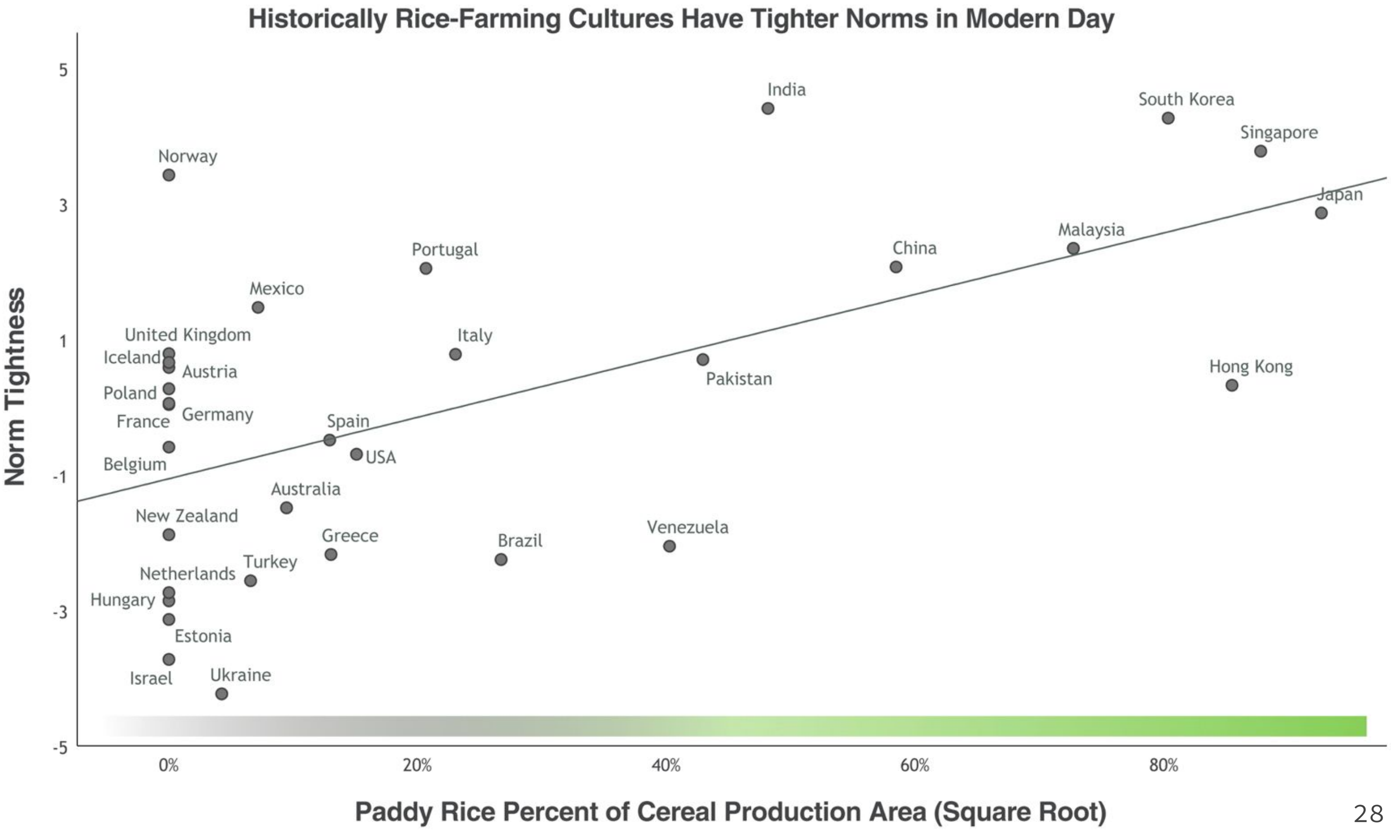
Even within a same country, the intensity of interdependent labor shows a correlation with holistic thinking styles

[Talhelm et al. 2014]



# Interdependence and Persistent Social Ties

Even across countries, rice farming cultures have “tighter” norms – stronger group pressure on individual conformity




Talhelm et al. 2020

# Uncertain Environments and Social Ties

These historical differences may have contributed to systematic differences in generalized trust and commitment to relationships

Survey of Japanese and American respondents

Q: “Do you think you can put your trust in most people, or do you think it’s always best to be on your guard?”

A: “People can be trusted”  **47% American vs.  26% Japanese**

# Uncertain Environments and Social Ties

Japanese society enforces stricter norms within groups, which provide security to their members

→ Strong trust for in-group members (norm violation is met with harsh sanctions)
→ Much weaker trust to outsiders/strangers (relatively weaker norms to ensure security)

In the extreme, if everyone distrusts outsiders, individually optimal choice is to rather stay in the community and increase commitment to existing ties

→ Strong ingroup trust: low transaction cost
→ Static relationships: high opportunity cost

# Uncertain Environments and Social Ties

Individualist cultures (e.g., U.S.) where the environment forced self-sufficiency and lower interdependent modes of subsistence (think the wild west):

→ Necessary to learn to trust strangers

→ High transaction cost (due to thin trust)

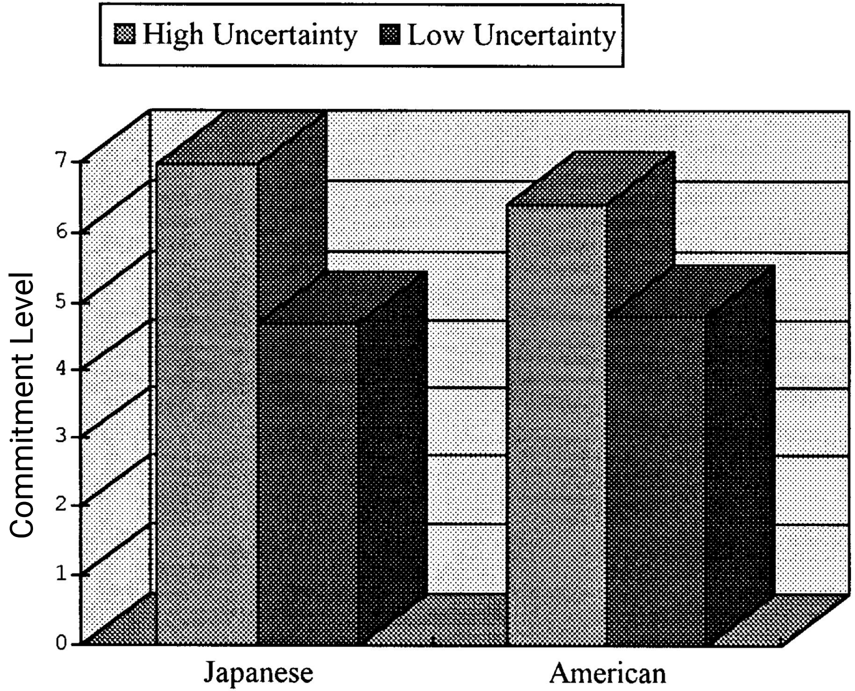
→ Low opportunity cost (possibility of more beneficial interactions)

# Uncertain Environments and Social Ties

In experimental settings where everyone transact with strangers (i.e., **no in-group security**), Japanese and the U.S. participants showed similar levels of commitment to their partners

**Both groups form long-term, committed relationships when uncertainty is high**.

(uncertainty = experimentally manipulated risk of being taken advantage of)




Yamagishi et al. 1998

# Uncertain Environments and Social Ties

It is not so much a matter of culture:

It is more a matter of **structure**

- In a society where **in-group cohesion** is strong, general trust becomes less critical

It is also more a matter of circumstances

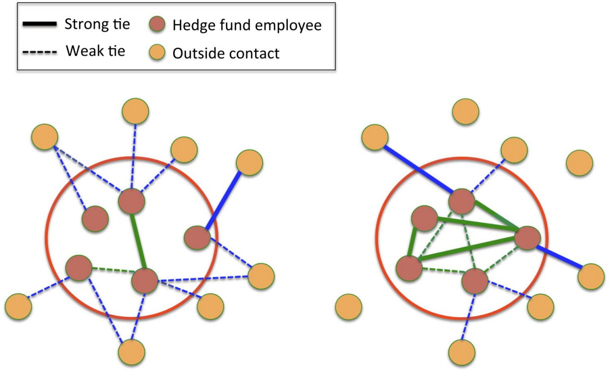
- Does the environment force interdependent modes of subsistence?
- Is there high uncertainty in the environment?

# Uncertain Environments and Social Ties

In real-world settings, people tend to shrink their communication ties to fewer, strong ties (“turtling up”)

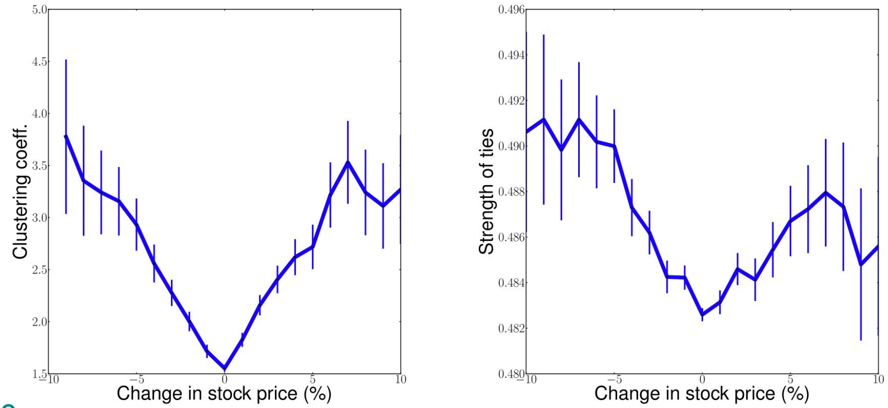
A shock leads people to revert to their trusted ingroup (higher clustering and higher average tie strength)

This tendency grows more salient with the magnitude of the shock




(a) Open network

(b) Turtled-up network




(a) Average clustering ($\nu(C_{s,d})$)

(b) Strength of ties ($S_{s,d,.1}$)

[Romero et al. 2019](#)

# Summary

An interpersonal tie influences and is influenced by the broader network structure

- Social support differs by type of relationship
- Topic-alter dependency can affect information diffusion
- Social tie can create a graph signature
- Dynamics of social ties hold implications for network structure

# Where We Are in the Course

Basic building blocks of networks: nodes, links, dyads, triads

Basic tools for analyzing networks: graph theory, BFS, random graph model

Universal properties (natural sciences) vs context and nuance (social sciences)

Fundamental properties of networks. Many types of networks display:

Q: What have we seen so far?

# Where We Are in the Course

Basic building blocks of networks: nodes, links, dyads, triads

Basic tools for analyzing networks: graph theory, BFS, random graph model

Universal properties (natural sciences) vs context and nuance (social sciences)

Fundamental properties of networks. Many types of social networks display:

- Short paths connecting nodes
  - Random Wikipedia articles https://www.thewikigame.com
  - Co-authorship distance https://www.csauthors.net/distance
- Triangles formed by common neighbors
- Similarity between neighbors ← more next time