Strong and Weak Ties

Ana Paula
Fabrício Benevenuto
Virgílio Almeida

Como as pessoas encontram empregos?

- Professor de Sociologia em Stanford
- Abordou essa questão como parte do seu doutorado por volta de 1960
- Através de entrevistas, descobriu que a maior parte das pessoas encontra emprego por contatos pessoais
- A maioria por **conhecidos** e não por amigos – resultado inesperado!
- Por que links de amizade fracos são os mais importantes nesse caso da procura por empregos?

Mark Granovetter

Granovetter's answer

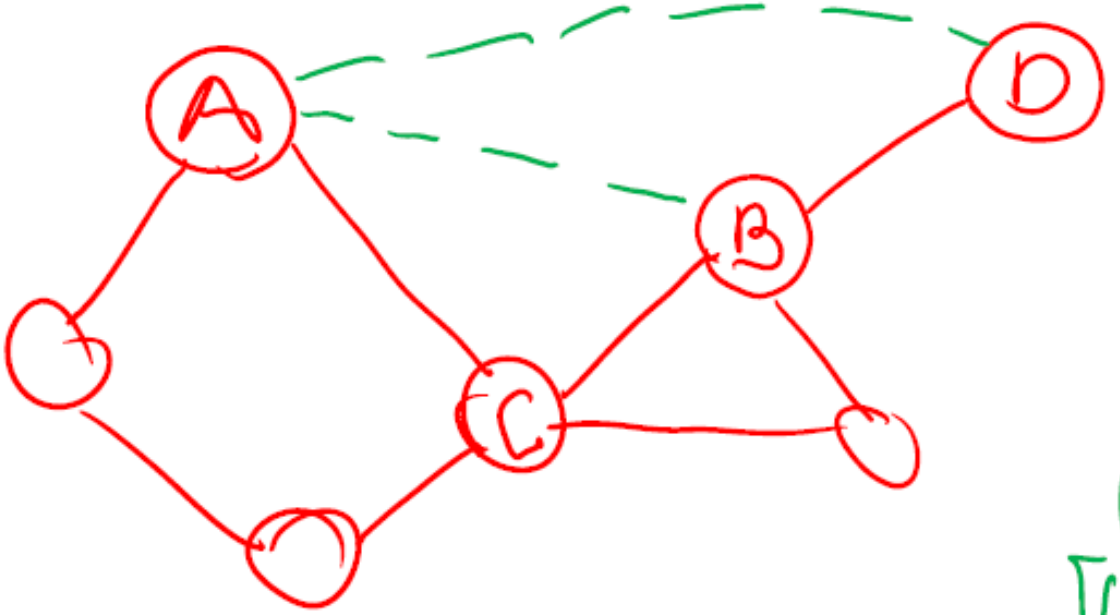
- Two perspectives on friendships
  - Structural:
    - Friendships span different portions of the network
  - Interpersonal:
    - Friendship between two people is either strong or weak

The Strength of Weak Ties

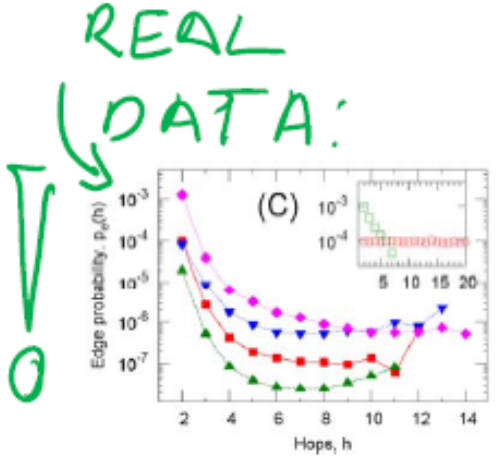
- Granovetter observed that the presence of weak ties often reduced path lengths (distance) between any two individuals—which led to quicker diffusion of information.

Triadic closure

- Which edge is more likely A-B or A-D?



- Triadic closure: If two people in a network have a friend in common there is an increased likelihood they will become friends themselves



Clustering Coefficients

- Clustering Coefficients were introduced by Watts & Strogatz in 1998, as a way to measure how close a node (or vertex) and its neighbors are from being a clique, or a complete graph within a larger graph or network.

- The clustering coefficient of a node is the number of actual connections across the neighbors of a particular node, as a percentage of possible connections. The clustering coefficient for the entire system is the average of the clustering coefficient for each node.

Coeficiente de Clusterização local

Para um nó i fração de pares de vizinhos do nó que estão eles próprios conectados.

Seja n_i o número de vizinhos do vértice *i*

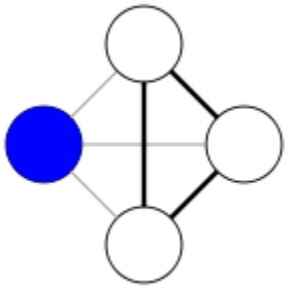
$$CC_i = \frac{\text{número de conexões entre os vizinhos de i}}{\text{número máximo de conexões possíveis entre os vizinhos de i}}$$

$$CC_{i\ direcinado} = \frac{\#\ \text{conexões direcionadas entre os vizinhos de i}}{n_i * (n_i - 1)}$$

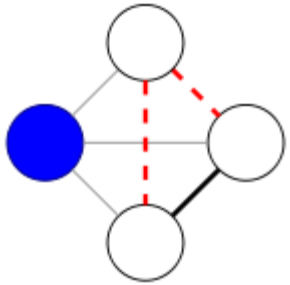
$$CC_{i\ não\text{-}direcionado} = \frac{\#\ \text{conexões não-direcionadas entre os vizinhos de i}}{n_i * (n_i - 1)/2}$$

Coeficiente de clusterização

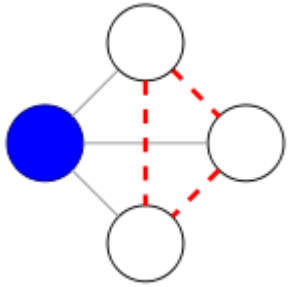
$$CC_i = \frac{\text{número de conexões entre os vizinhos de i}}{\text{número máximo de conexões possíveis entre os vizinhos de i}}$$



cc = 1



cc = 1/3



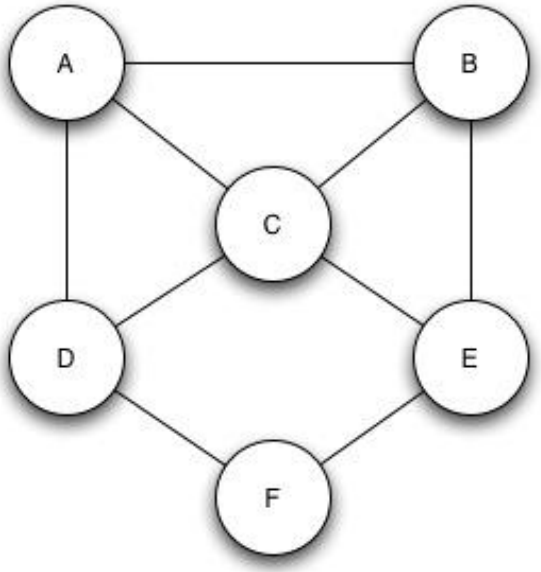
cc = 0

Os amigos dos spammers estão conectados entre si?

Coeficiente de Clusterização Global

- Média sobre todos nós *n*
  –

$$C = \frac{1}{n}\sum_{i} C_i$$



CC_A = 2/3

CC_B = 2/3

CC_C = 1/2

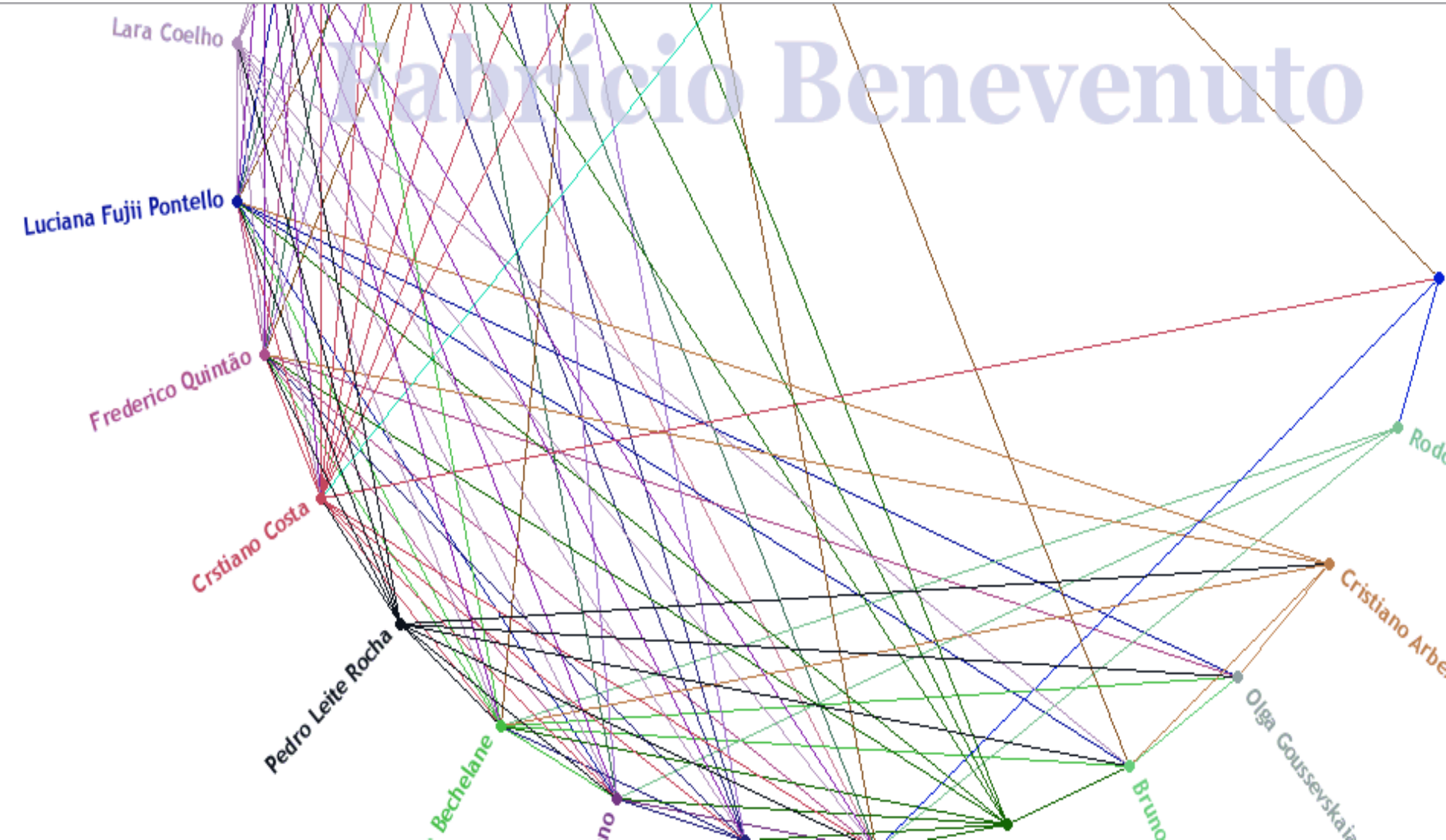
CC_D = 1/3

CC_E = 1/3

CC_F = 0

CC = 5/12

Facebook – Friend Weel



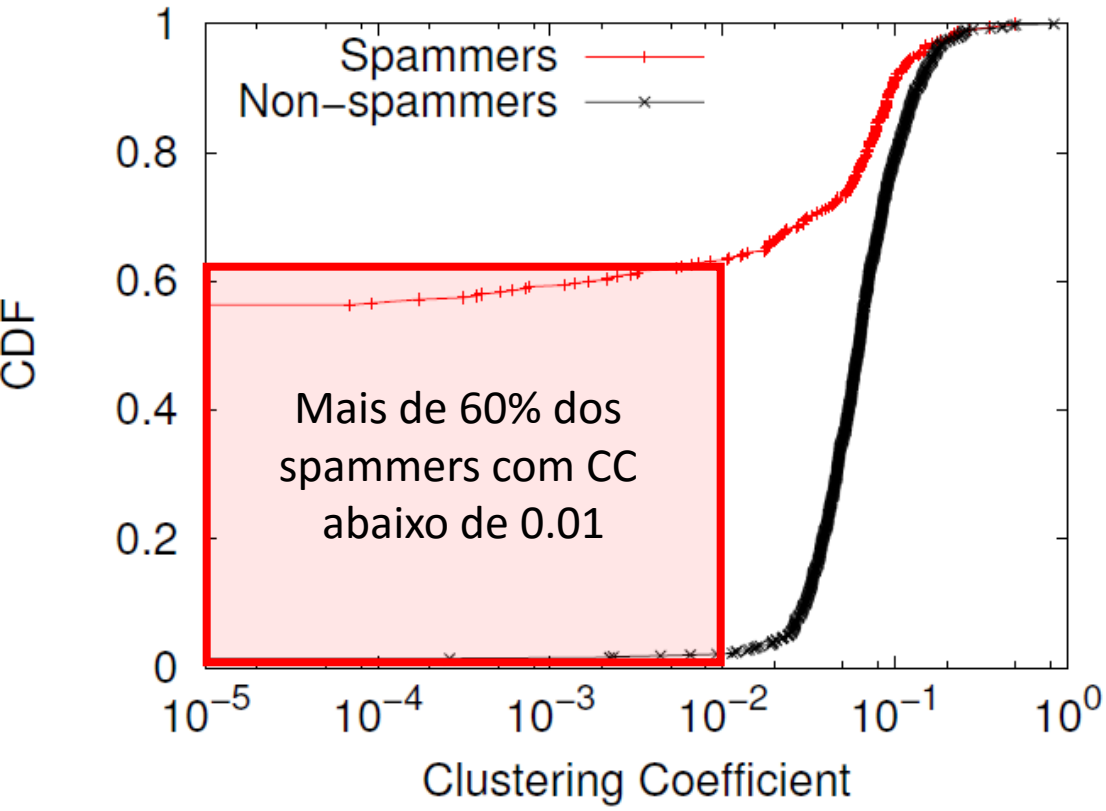
Measurement and Analysis of Online Social Networks [IMC'07]

Network	C
Web [2]	0.081
Flickr	0.313
LiveJournal	0.330
Orkut	0.171
YouTube	0.136

Redes sociais possuem coeficiente de clusterização global bem maior que redes tecnológicas

Comparação tem que ser feita com mesmo grafo com arestas embaralhadas

Útil como Feature para outros problemas



Coeficiente de clusterização
probabilidade dos vizinhos de um nodo estarem conectados

Outras 45 features capazes de diferenciar spammers

Uso de classificação (2010)

Amigos dos spammers não estão conectados entre si

Triadic closure

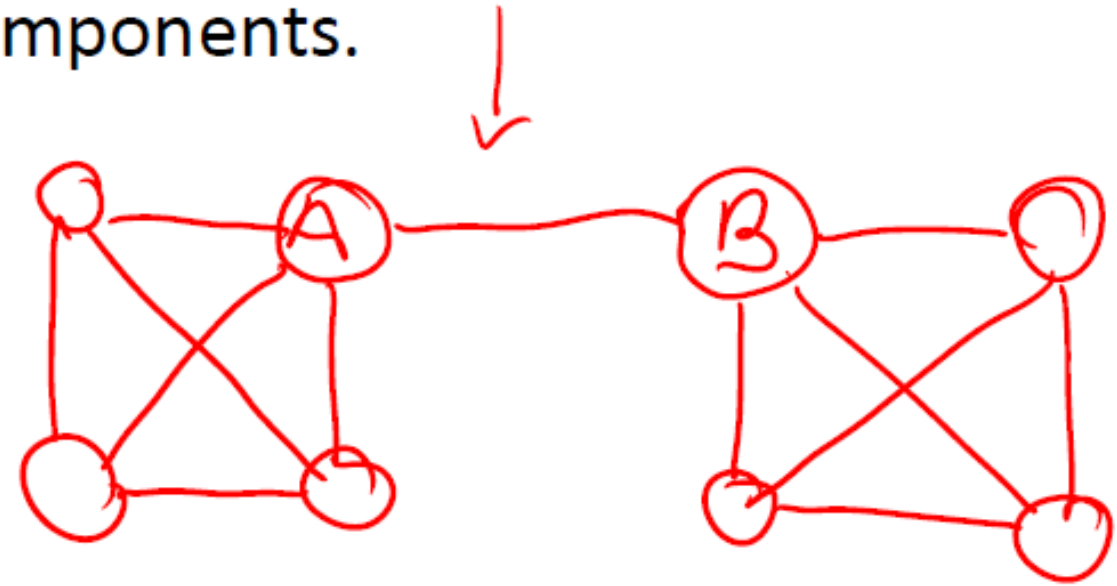
- Triadic closure == High clustering coefficient

Reasons for triadic closure:

- If B and C have a friend A in common, then:
  - B is more likely to meet C (since they both spend time with A)
  - B and C trust each other (since they have a friend in common)
  - A has incentive to bring B and C together (as it is hard for A to maintain two disjoint relationships)
- Empirical study by Bearman and Moody:
  - Teenage girls with low clustering coefficient are more likely to contemplate suicide)

Bridges and Local Bridges

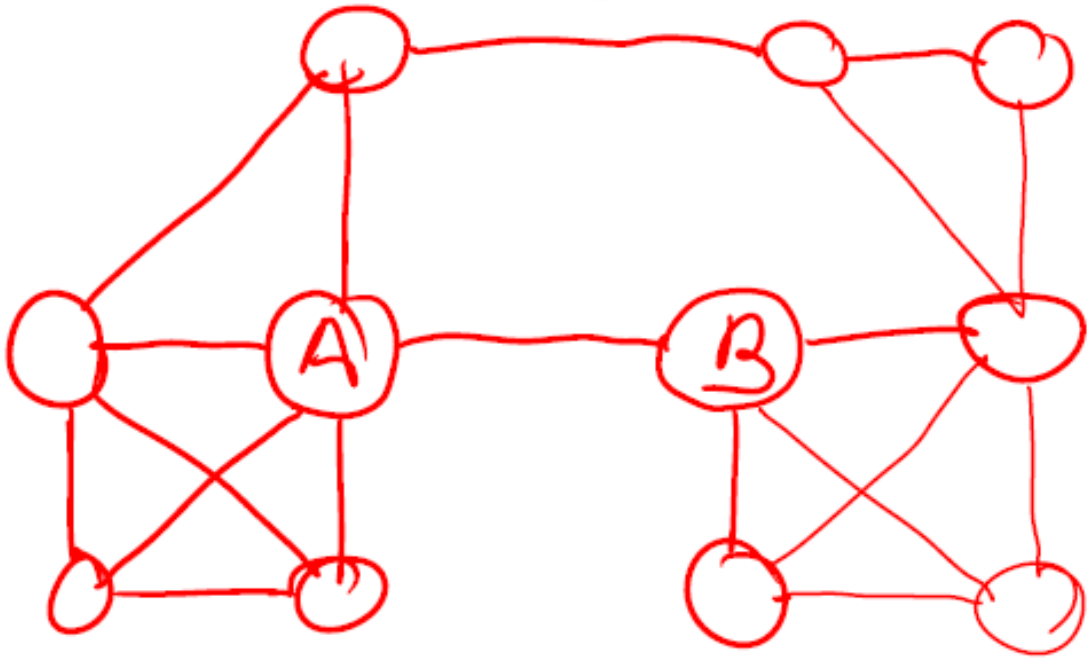
- Edge (A,B) is a bridge if deleting it would make A and B in be in two separate connected components.



Definição muito forte para dados reais

Bridges and Local Bridges

- Edge (A,B) is a local bridge A and B have no friends in common.
- Span of a local bridge is the distance of the edge endpoints if the edge is deleted.

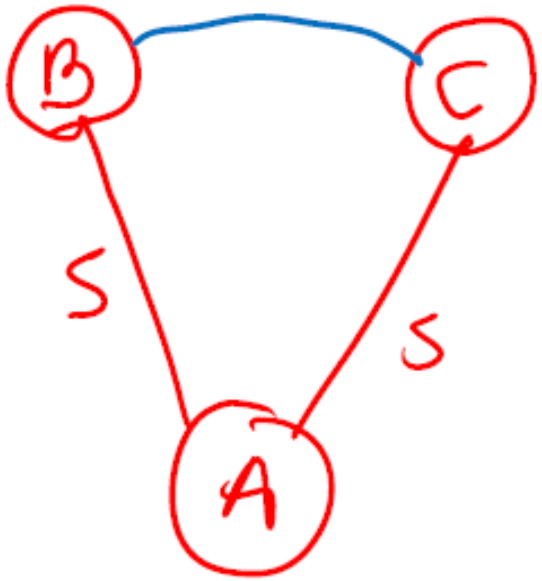


SPAN (A,B) IS 4

(local bridges with long span are like bridges)

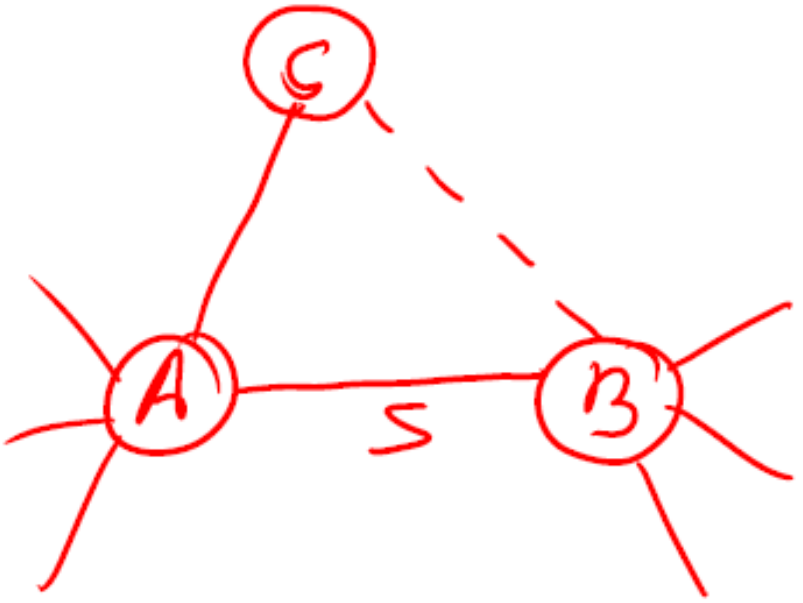
Strong Triadic Closure

- Links in networks have strength:
  - Friendship
  - Communication
- We characterize links as either Strong (friends) or Weak (acquaintances)
- **Strong Triadic Closure Property:** If A has strong links to B and C, then there must be a link (B,C) (that can be strong or weak)



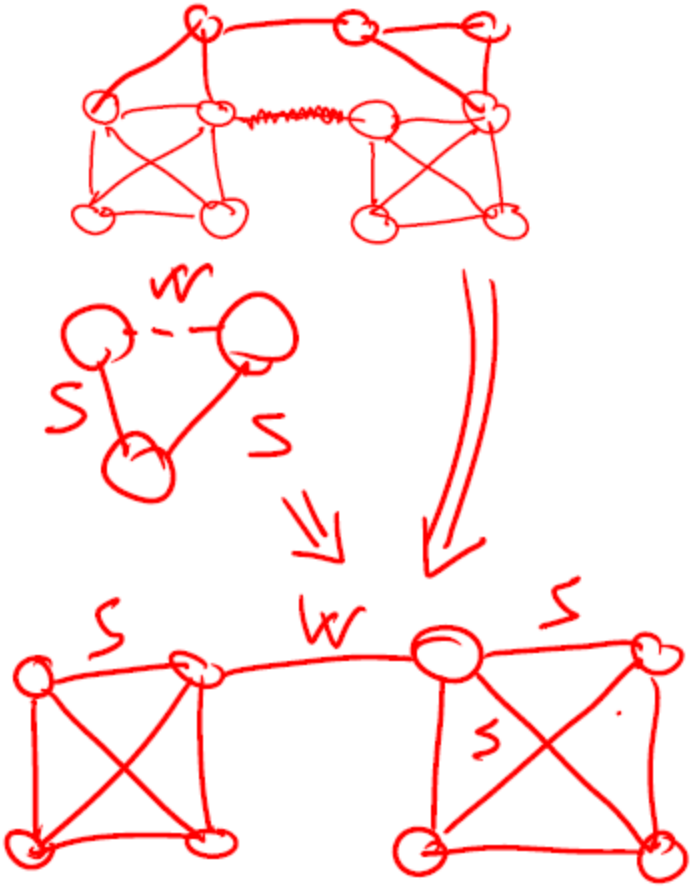
Local Bridges and Weak ties

- If node A satisfies Strong Triadic Closure and is involved in at least two strong ties, then any local bridge adjacent to A must be a weak tie.
- Proof by contradiction:
  - A satisfies Strong Triadic Closure
  - Let A-B be local bridge and a strong tie
  - Then B-C must exist because of Strong Triadic Closure
  - But then (A,B) is not a bridge



Summary of what we just did

- Defined Local Bridges:
  - Edges not in triangles
- Set two types of edges:
  - Strong and Weak Ties
- Defined Strong Triadic Closure:
  - Two strong ties imply a third edge
- ➔ Local bridges are weak ties



Tie strength and structure in real data

- For many years the Granovetter's theory was not tested
- But, today we have large who-talks-to-whom graphs:
  - Email, Messenger, Cell phones, Facebook
- Onnela et al. 2007:
  - Cell-phone network of 20% of country's population

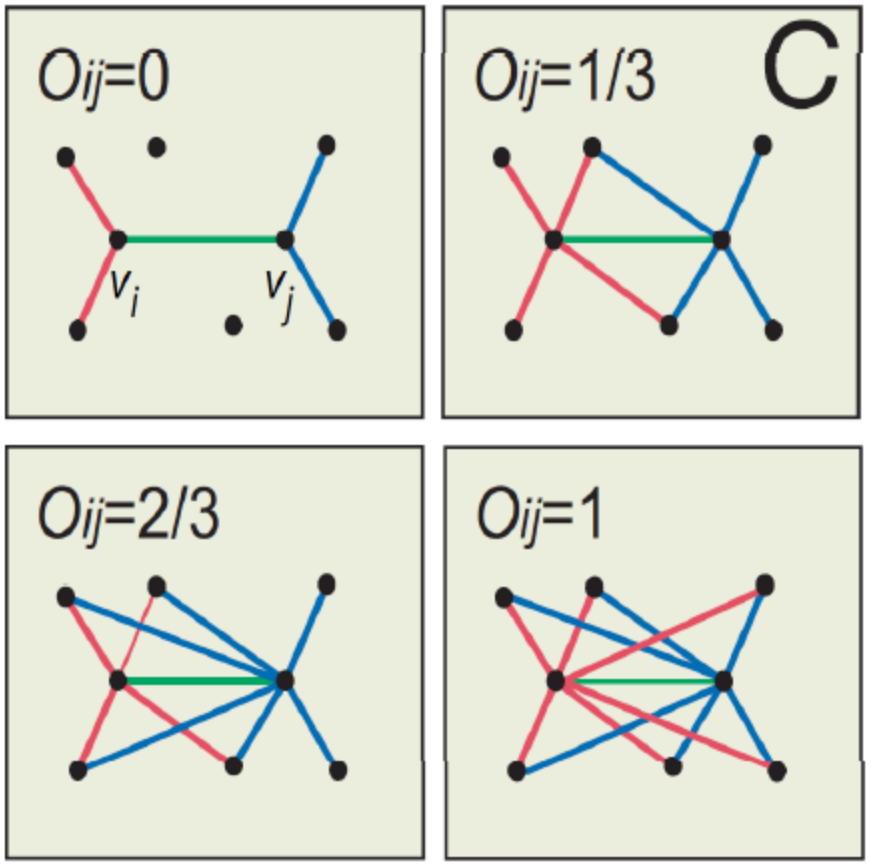
Neighborhood Overlap

- Overlap:

$$O_{ij} = \frac{n(i) \cap n(j)}{n(i) \cup n(j)}$$

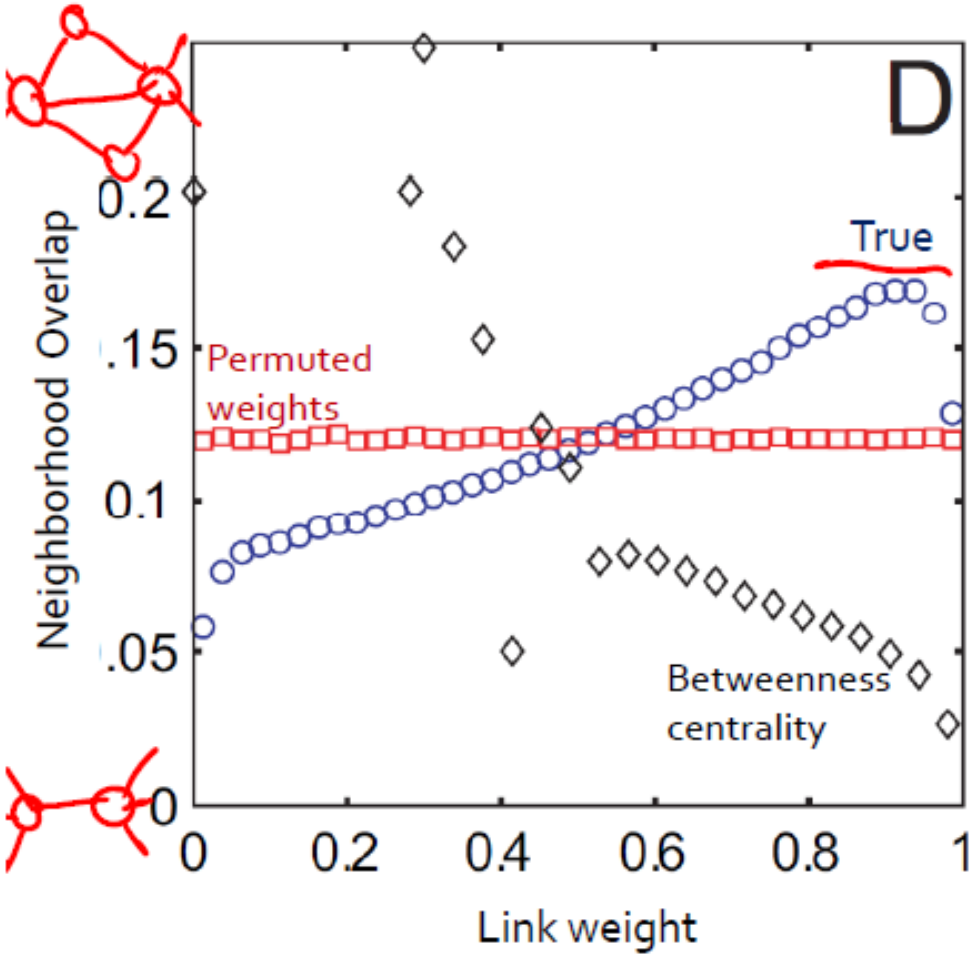
  - n(i) … set of neighbors of A

- Overlap = 0 when an edge is a local bridge



[Onnela et al. '07]

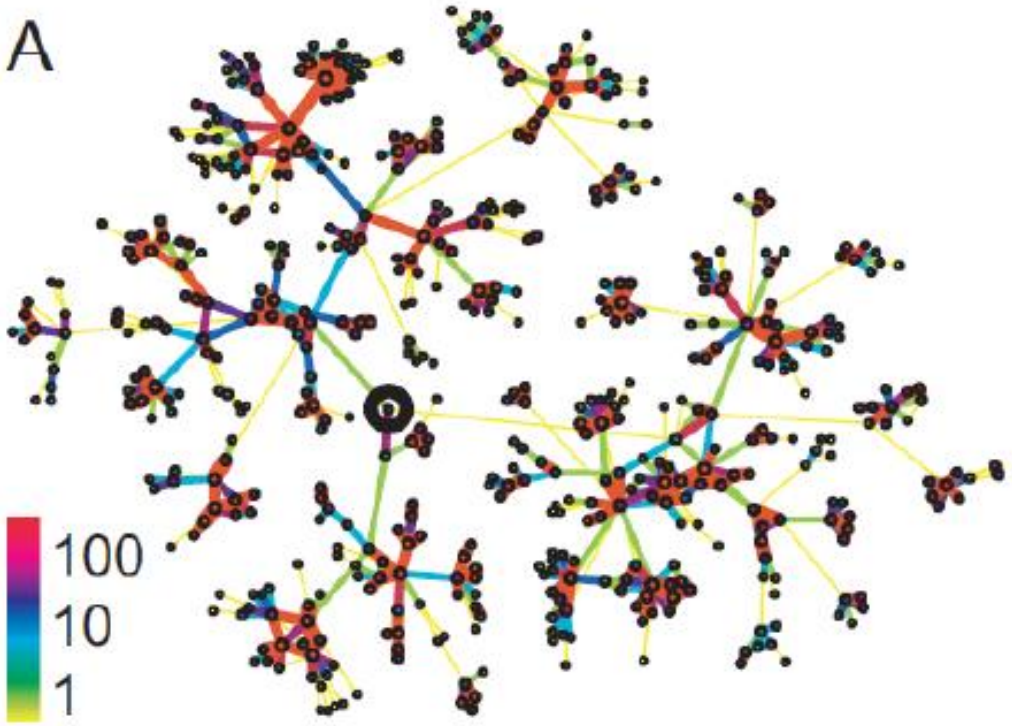
Mobile phones: Overlap vs. Weight



- Permuted weights: keep the structure but randomly reassign edge weights
- Betweenness centrality: number of shortest paths going through an edge

[Onnela et al. '07]

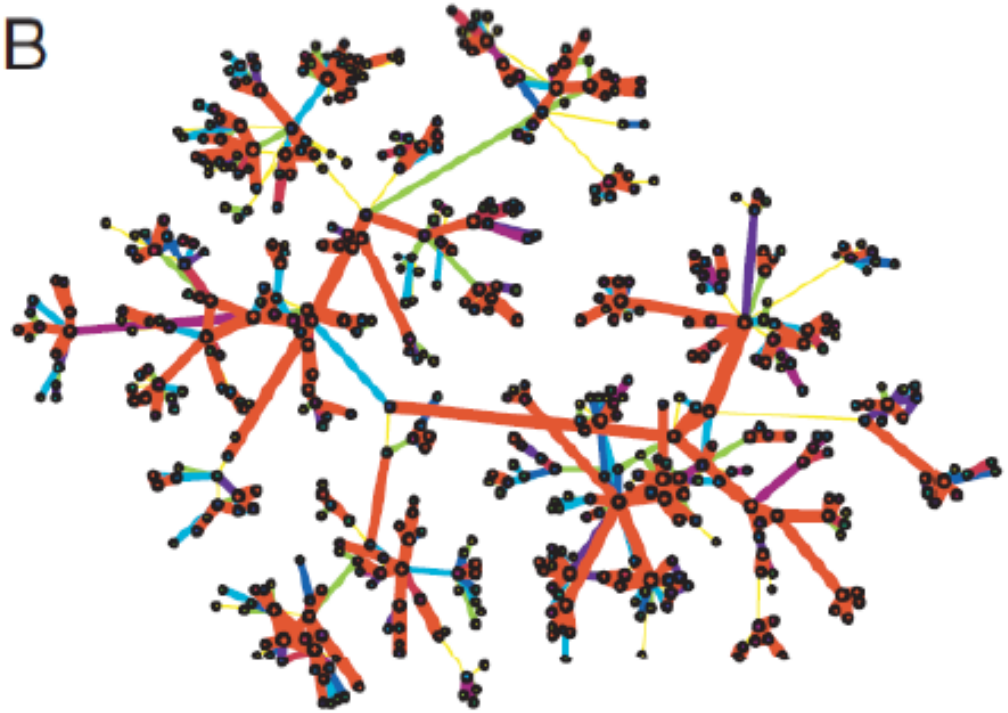
Real network tie strengths



- Real edge strengths in mobile call graph

[Onnela et al. '07]

Permuted tie strenghts

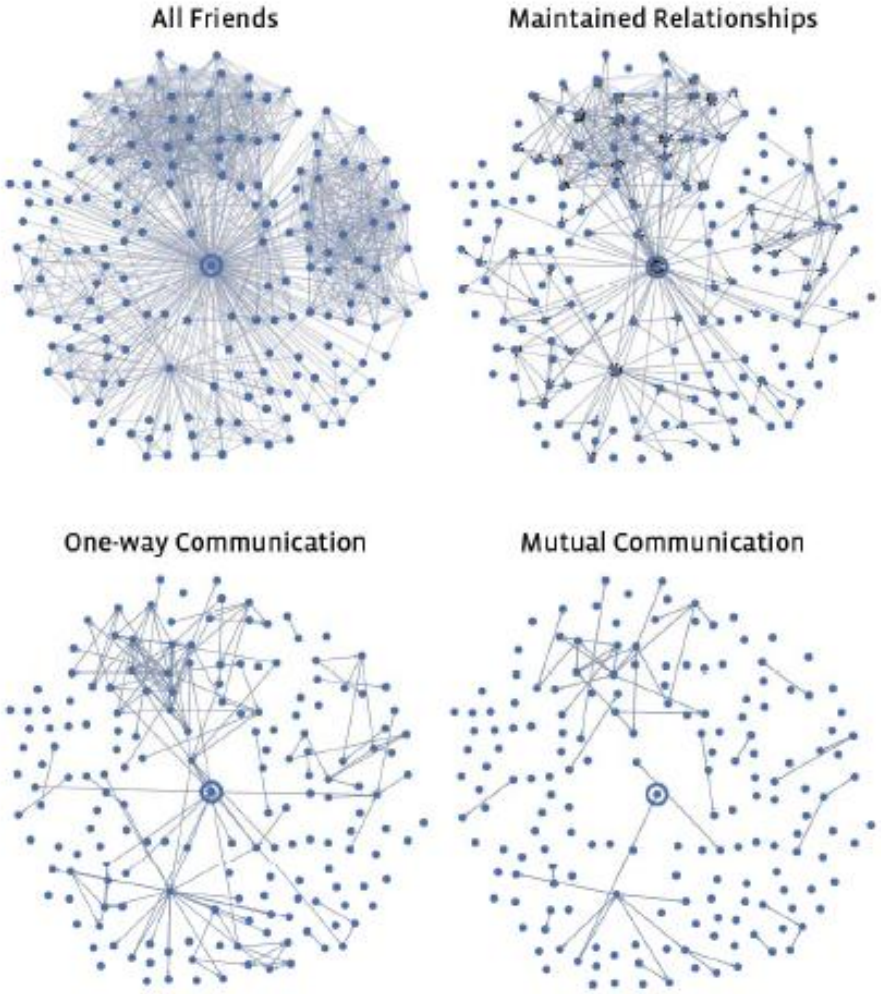


- Same network, same set of edge strengths
- But now strengths are randomly shuffled over the edges

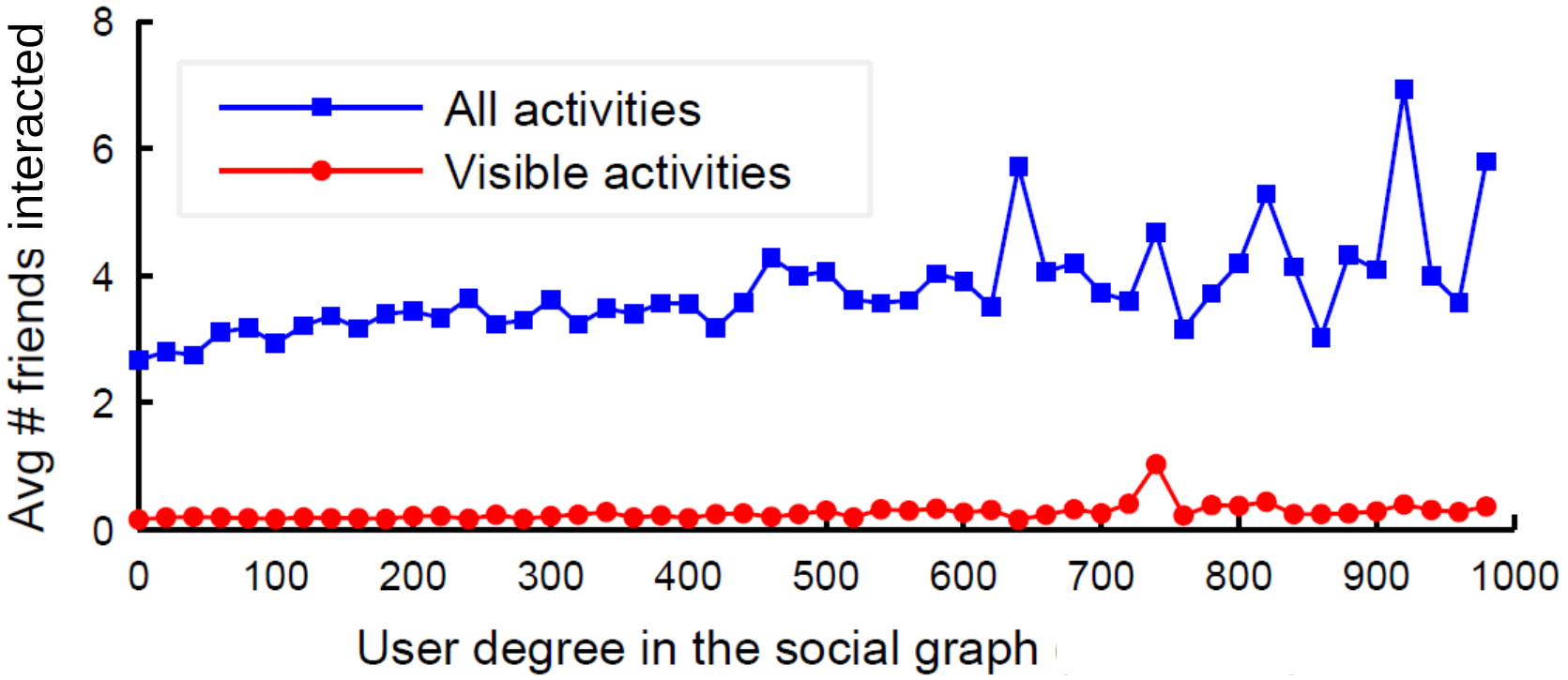
[Marlow et al. '09]

Another example: Facebook

Exemplo das interações de um usuário



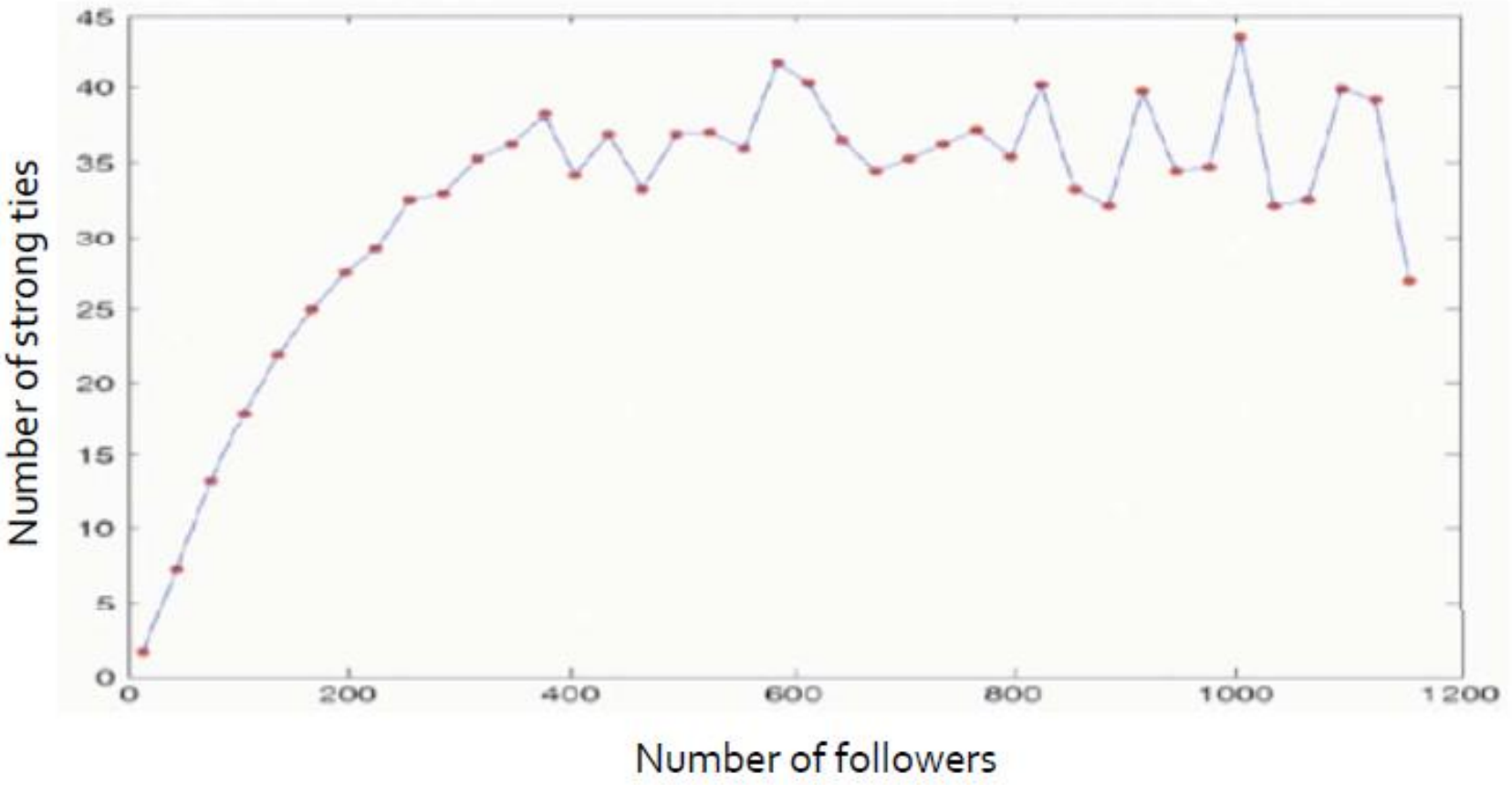
Characterizing User Behavior in Online Social Networks [IMC'09]



- Degree of interaction increases by an **order of magnitude** when incorporating silent interactions
- 85% of the active users showed **only silent interactions**!

[Huberman et al. `09]

Twitter: Strong ties vs. Followers



Strong tie no twitter = mensagem direta

Há um limite no número de strong ties, mesmno para usuários com muitos seguidores

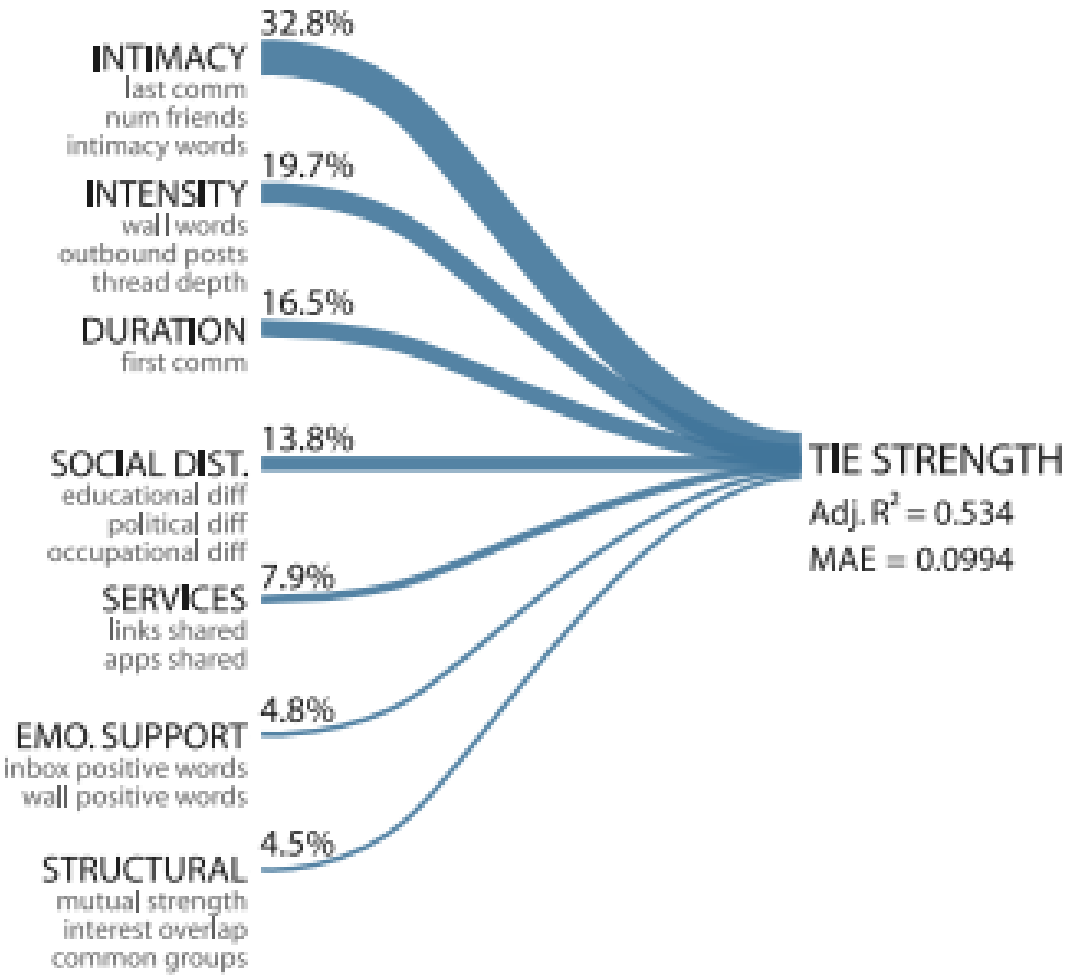
Inferring Tie Strength from Online Directed Behavior [PlosOne'13]

	Closest	*Not Closest*
Predicted Closest	582	47
Predicted Not Closest	207	742

Columns represent ground truth and rows represent predicted values.
doi:10.1371/journal.pone.0052168.t004

- Fez a pergunta: "Quem é seu melhor amigo?"
- Mediu interações públicas do Facebook
- Mostrou que mensagens privadas não acrescentam tanto

Predicting Tie Strength With Social Media [CHI'09]



- Nova métrica para medir força dos links
- combina features
- Obtém parâmetros empiricamente
- Validação por questionários

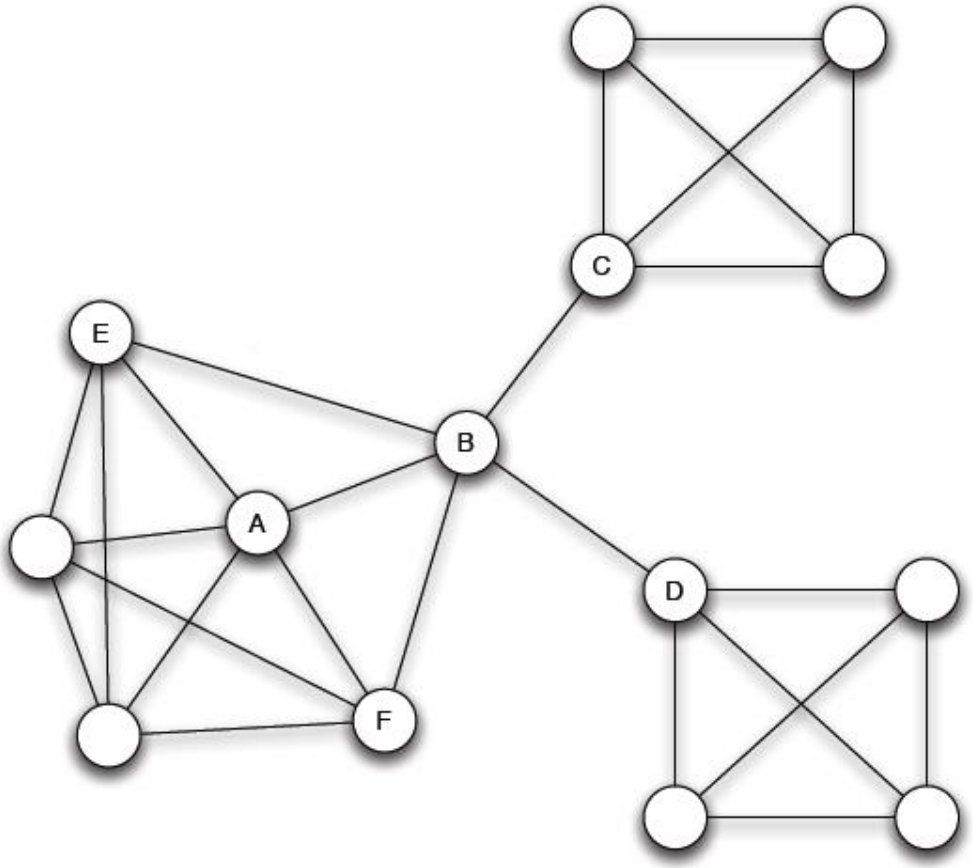
Working with Friends:Unveiling Working Affinity Features from Facebook Data [ICWSM'14]

Table 6: Ranking of most important attributes, presented by the IG (*Information Gain*) Ranking and the χ^2 (*Chi-Squared*) Ranking

Description	***IG Rank***	***IG Value***	χ^2 ***Rank***	χ^2 ***Value***
$tieStrength_3$	1	0.194	1	226.01
$tieStrength_4$	2	0.181	2	220.00
$similarity_1$	3	0.151	3	184.51
$similarity_3$	4	0.150	4	181.10
$tieStrength_2$	5	0.098	5	120.01
$popularity_1(source)$	6	0.084	8	100.93
$outgoing_1(target)$	7	0.084	6	116.81
$tieStrength_1$	8	0.083	9	98.95
$popularity_2(target)$	9	0.079	10	96.17
$Grade(target)$	10	0.075	7	104.31
$outgoing_2(target)$	11	0.073	11	91.05
$outgoing_2(source)$	12	0.069	13	82.56
$Grade(source)$	13	0.065	12	89.44
$popularity_1(target)$	14	0.048	14	62.14
$outgoing_1(source)$	15	0.040	15	46.70
$popularity_2(source)$	16	0.035	16	44.39
$similarity_2$	17	0.022	17	27.30

- Melhor métrica para interesse em trabalhar juntos [CHI'09]

Structural Holes



Nodos que criam múltiplas pontes obtêm várias vantagens

1) Early access
2) Synthesis of multiple ideas
3) Gatekeeping (regular informação que passa)
4) Tradeoff: less protected

- Dentro de uma grande empresa, quais as vantagens de se ter acesso a B?
- Maior chance de ter sucesso profissional[86,87]

Capital Social

- Estratégia para Big Brother Brasil

- Devo me fechar com um grupo ou devo buscar ser ponte entre grupos (preencher strutuctural holes)?

  - Como ponte consegue-se informação privilegiada dos grupos
  - Ponte pode decidir qual informação será propagada
  - Entretanto, há um tradeoff. Ponte seria o primeiro candidato a sair de um grupo

Capital Social

- Termo relacionado a vantagens que uma posição estrutural na rede pode te dar (ex. benefício de se tornar ponte ou de preencher structural holes)
- Mas existem muitas definições na literatura
- Pode ser entendido como ganho social devido a posição estrutural na rede
- Como isso se manifesta em redes como Twitter e Facebook?