



The Slashdot Zoo: Mining a Social Network with Negative Edges



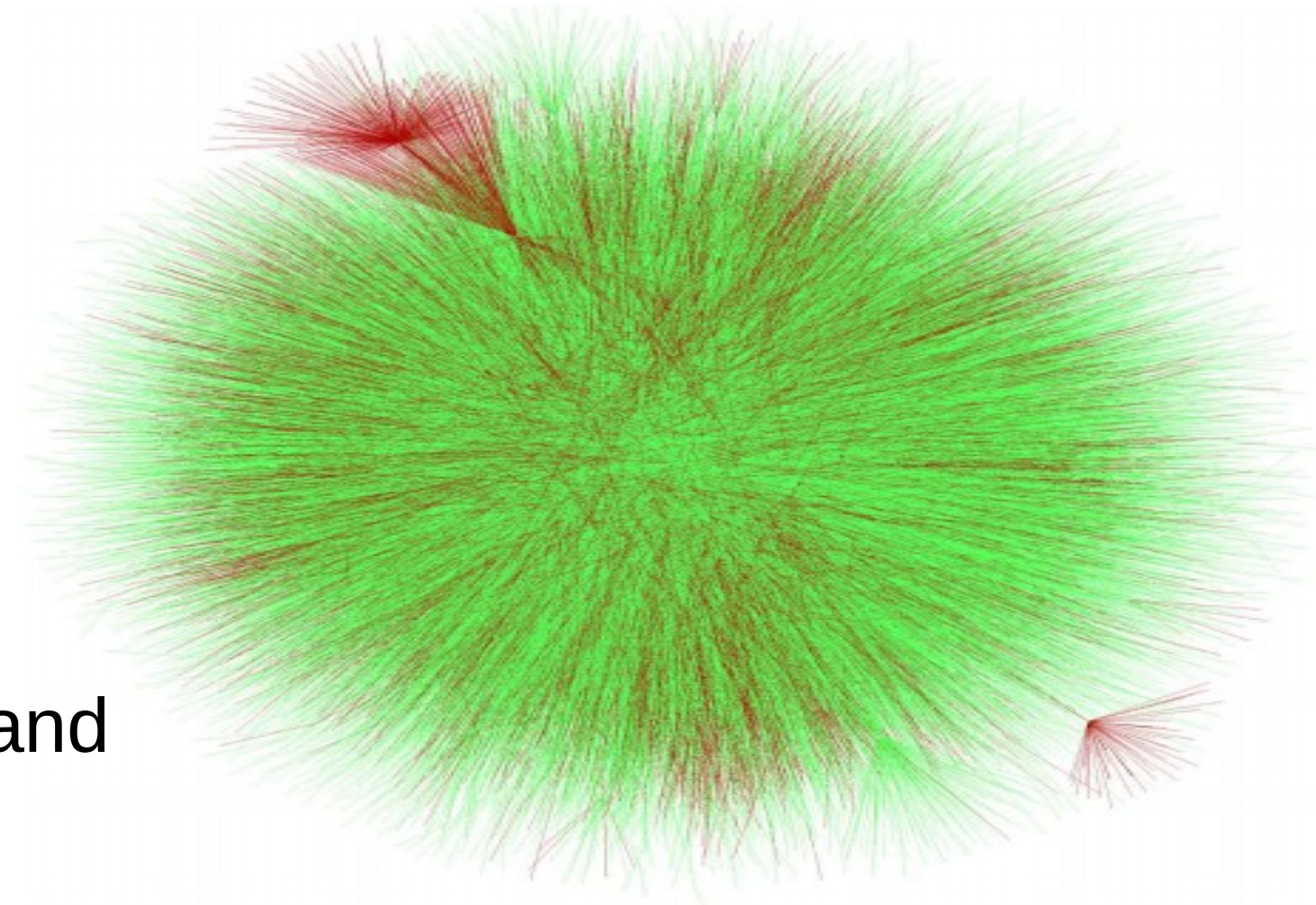
Tong Niu
tong.niu.cn@outlook.com

Outline

- **Introduction**
 - **Slashdot zoo & experiments**
 - **On the global level**
 - **On the node level**
 - **On the link level**
- **Conclusion**
- **Discussion**

Introduction

- Directed graphs with Weighted edges
 - Positive
 - negative
- The Slashdot Zoo
 - Users can tag other users as friends and foes , providing positive and negative endorsements

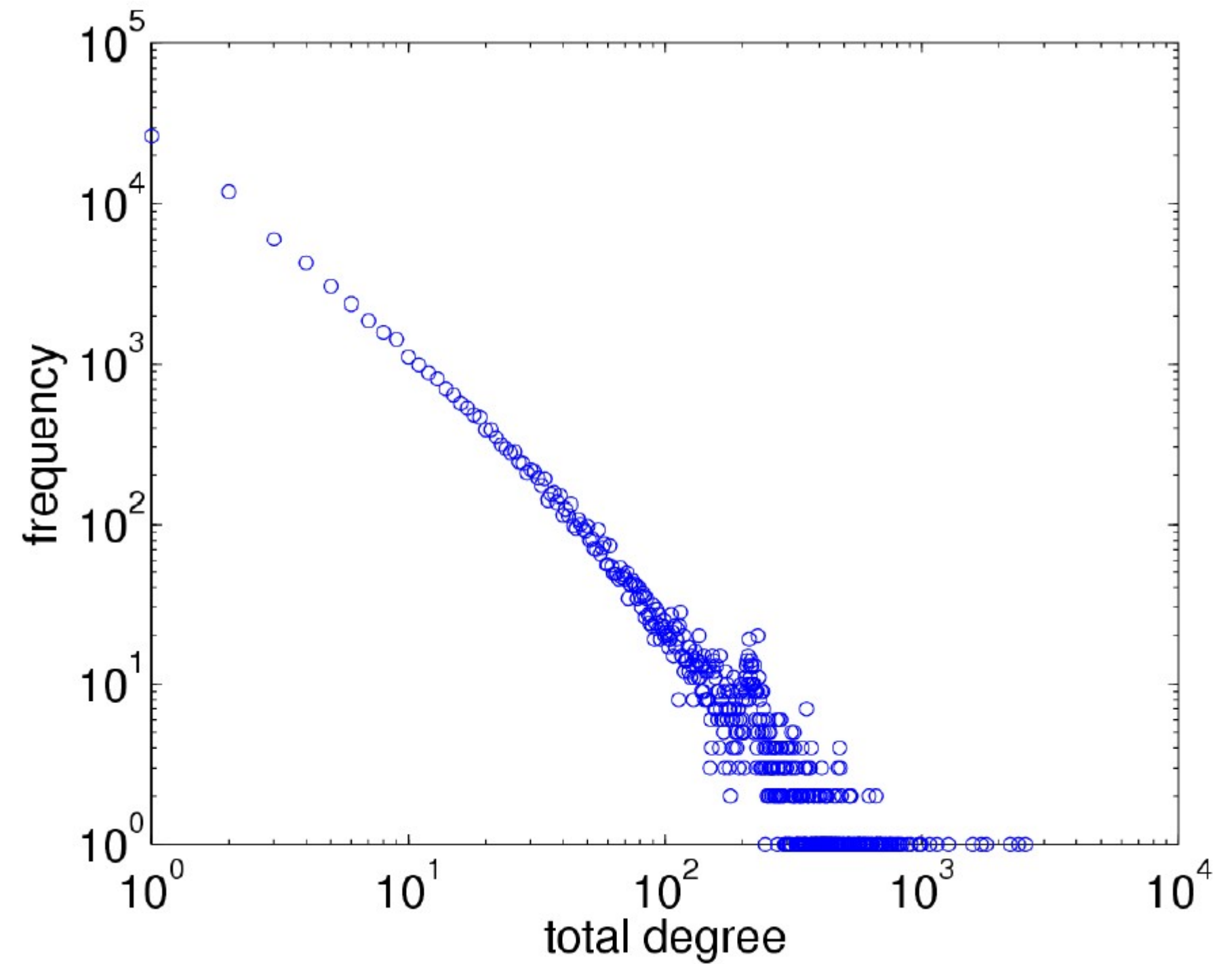


red: "foe"

green: "friend"

Introduction

- The Slashdot zoo
 - Follows a power law



Introduction

- Signed Graphs
 - Each edge has a positive or negative sign
 - Balanced , if the product of edge signs around every circle is positive

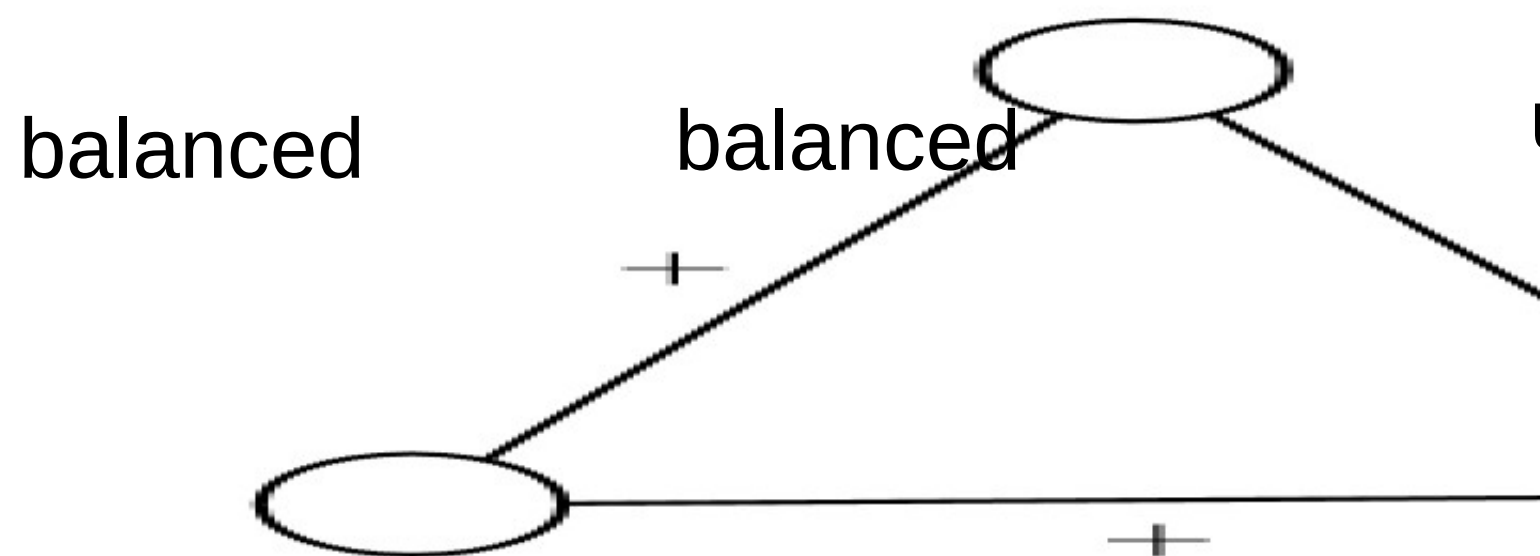


figure 1

figure 2

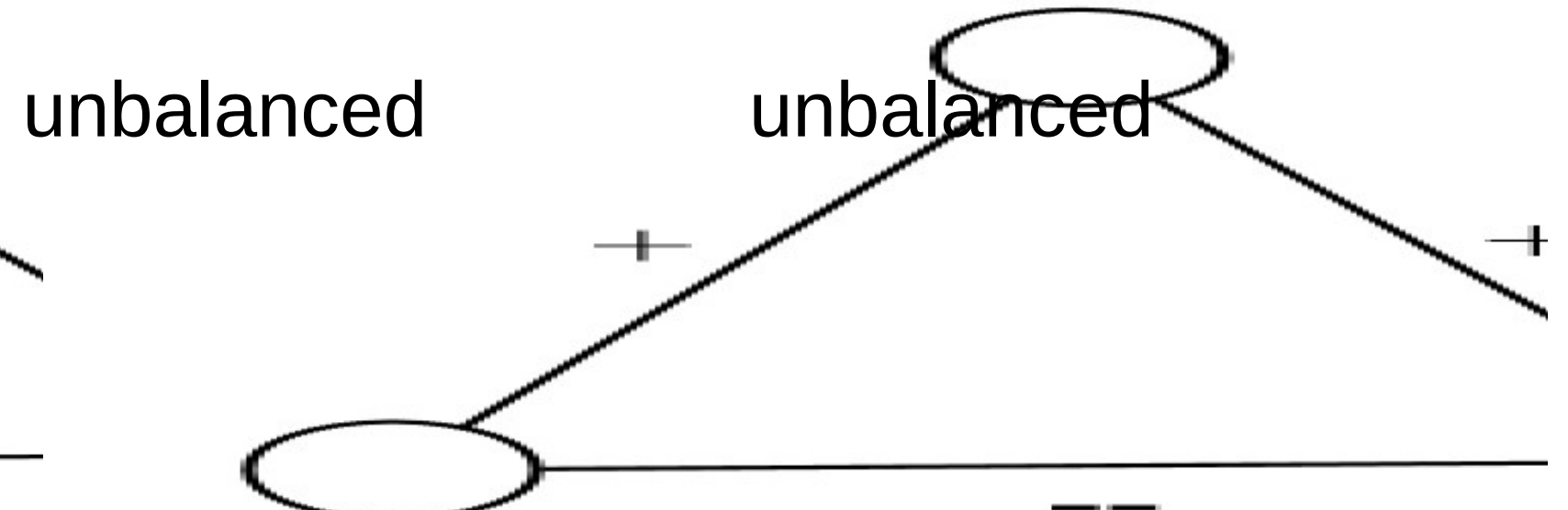


figure 3

figure 4

Introduction

- Transitivity
 - “The friend of my friend is my friend”
- Multiplicative Transitivity
 - for signed graphs
 - “the enemy of my enemy is my friend”—triangle is balanced

Global level

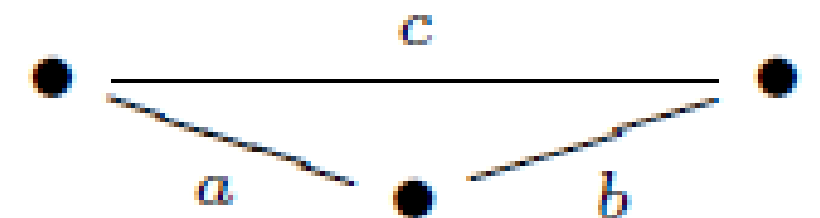
- Clustering coefficient
 - A measure of transitivity
 - Define the proportion of all incident edge pairs that are completed by a third edge to form a triangle
 - Taking values between 0 and +1

$$C(G) = \frac{\|\bar{A} \circ \bar{A}^2\|_+}{\|\bar{A}^2\|_+} \quad (1)$$

Matrix that contains its entry (i, j) as a sum of paths of length 2

A : unsigned adjacency matrix

$\|A\|_+$: sum of all matrix elements

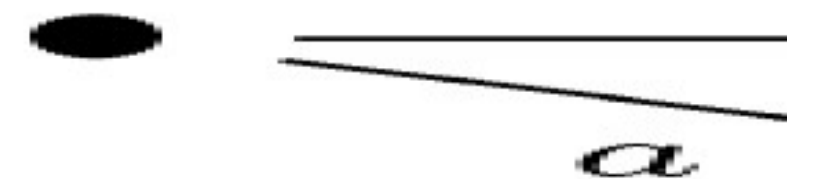


Global level

- Signed clustering coefficient
 - A measure of multiplicative transitivity
 - Taking values between -1 and +1

$$C_s(G) = \frac{\|A \circ A^2\|_+}{\|\bar{A}^2\|_+} \quad (2)$$

Matrix that contains at its entry (i,j) a sum of paths of length 2
 A : signed adjacency matrix
 $\|A\|_+$: sum of all matrix elements



Global level

- Relative signed clustering coefficient
 - Taking values between -1 and +1
 - +1 : all triangles are oriented coherently
 - -1 : the sign does not hold multiplication rule

$$S(G) = \frac{C_s(G)}{C(G)} = \frac{\|A \circ A^2\|_+}{\|\bar{A} \circ \bar{A}^2\|_+} \quad (3)$$

- VS random graph

Clustering coefficient	3.19% (0.0095%)
Directed clustering coefficient	5.62% (0.0191%)
Signed clustering coefficient	2.44%
Directed signed clustering coefficient	4.44%

Node level

- Popularity and centrality
 - Identify “trolls”
 - Centrality: denote what extent the node is central to the graph
 - Central nodes are usually well connected to other nodes, and at the short distance to most other nodes

Node level

- Fans minus freaks(FMF)
 - A user is a fan of his friends and the freak of his foes
- PageRank(PR)
 - Defined on directed graphs with nonnegative edge weights
 - If there are a lot of edges through a node, the PR value of this node will be high.(importance)
- Signed spectral Ranking(SR)
 - Extend to the case of negative edge weights

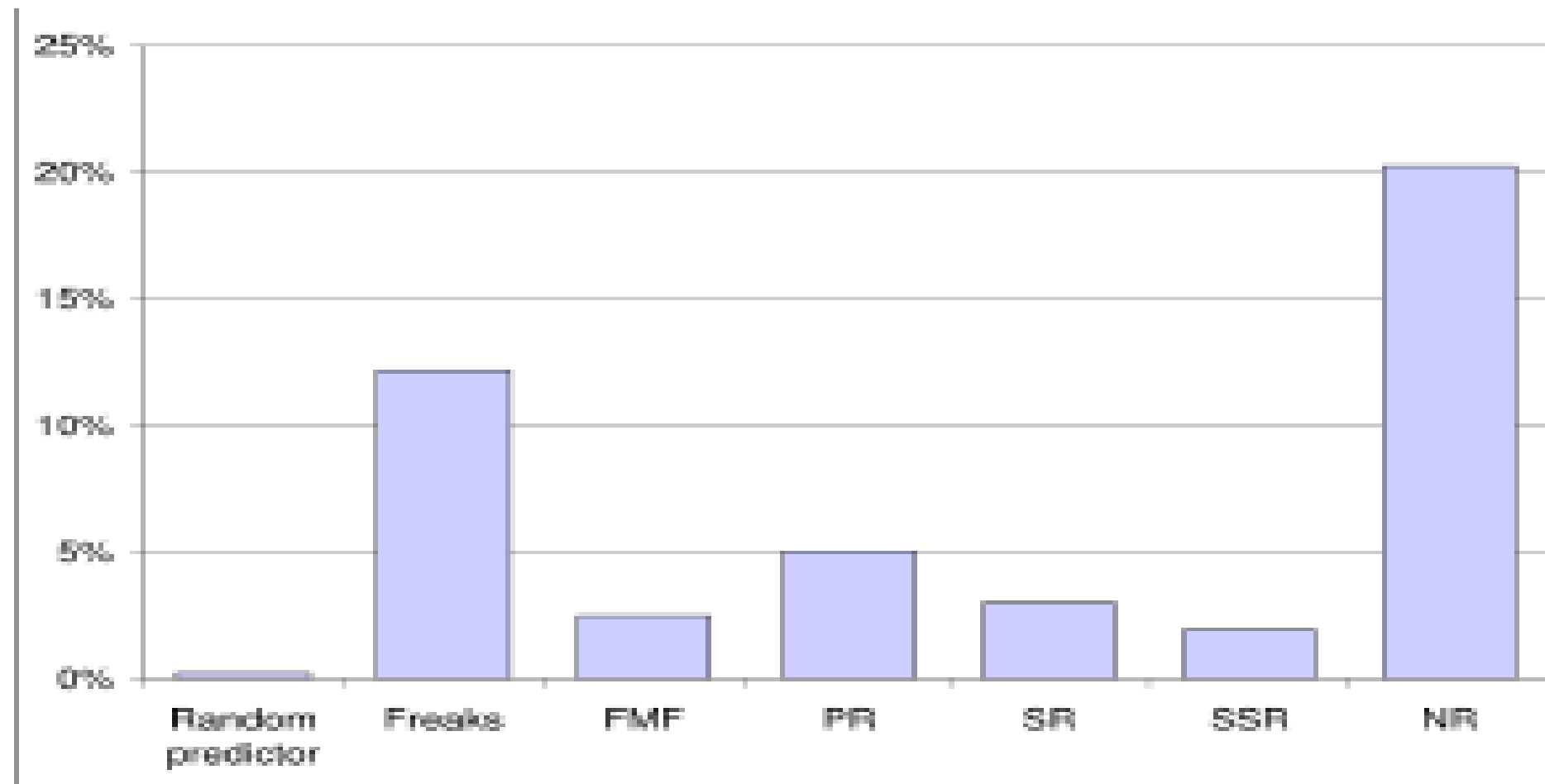
Node level

- Signed symmetric spectral ranking(SSR)
 - Outlinks by users are counted towards the popularity of a user
 - It models the idea “popular users will refrain from having many foes, and negative edges are more prevalent between unpopular users”
- Negative Rank(NR)
 - Identify unpopular users
 - Not useful for finding central or prominent users

$$NR = SR - \beta \cdot PR$$

- **The best performance is attained at $\beta = 1$**
The best performance is attained at

Node level

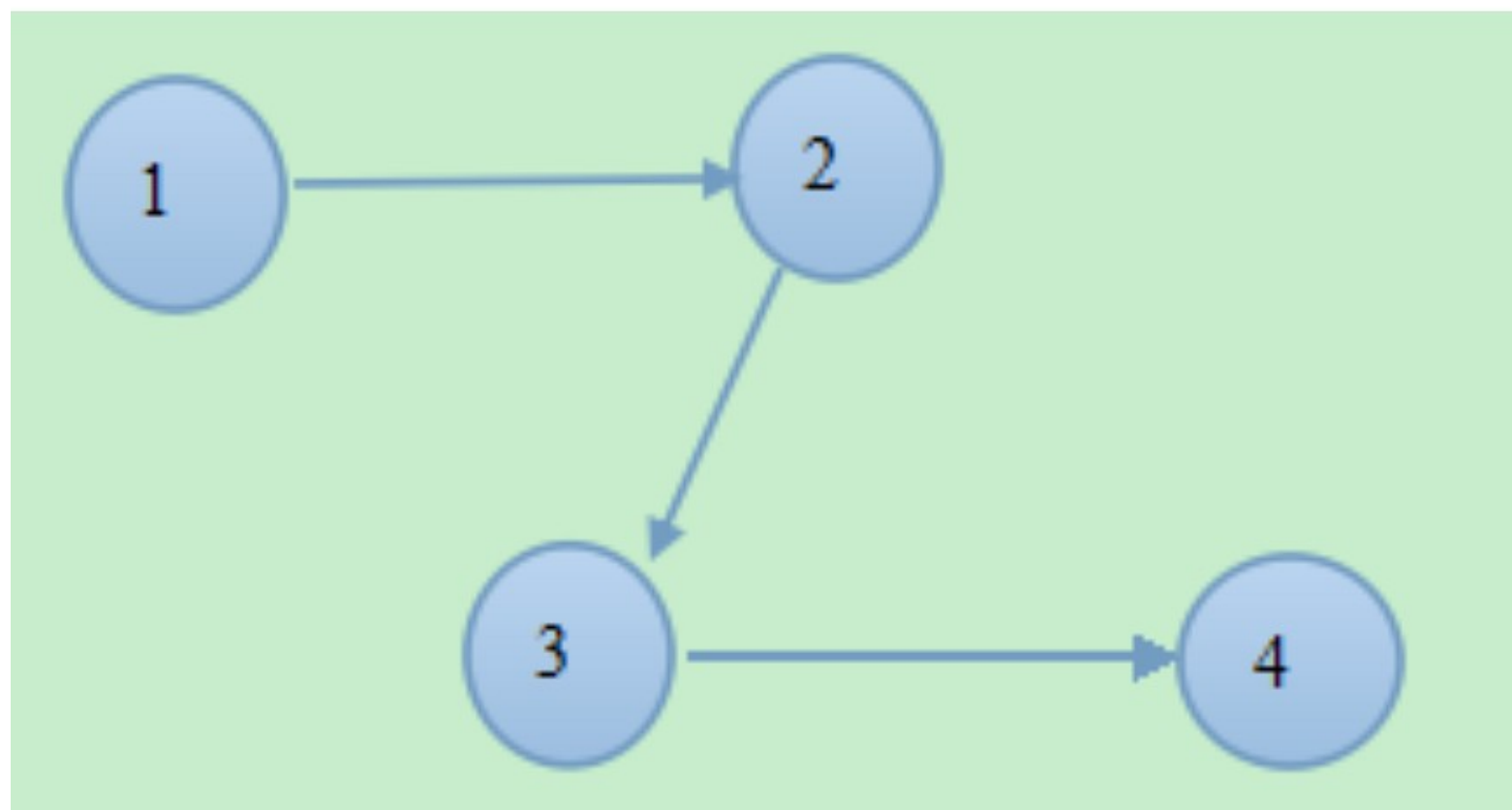


The mean average precision of measures of identifying trolls

- NR performs best
- SR performs badly
- SR is not a good indicator of popularity, but has to be combined with PR to give NR.

Link level

- ~~old link prediction algorithms with nonnegative edge are only able to predict the presence of an edge~~
- ~~Need to define new algorithm with signed edge and are able to predict missing edges' sign~~
 - ~~Implemented by similarity function~~
 - ~~A , adjacency matrix for proximity measure~~



0	0	1	0
0	0	0	1
0	0	0	0
0	0	0	0

Link level

- In real world is a huge calculation – dimensionality reduction
- Reduced by performing a sparse singular value decomposition

$$A \approx A_k = U_k D_k V_k^T$$

- k is the reduced dimension

Conclusion

Analyze social network on graphs with negative edge weights—
Slashdot zoo on three levels and shows that the network exhibits
multiplicative transitivity and importance of signed techniques.

Discussion

