



# Defining Reputation in Signed Networks

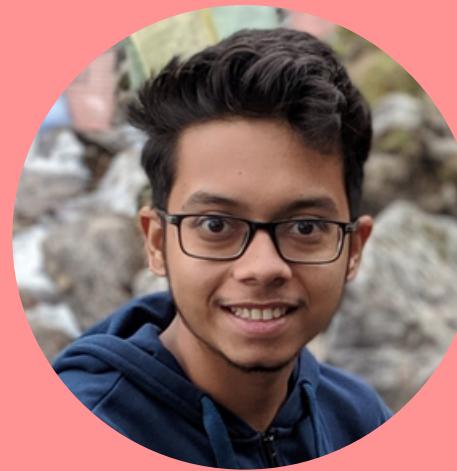
— 01

Final Presentation





# Group Members



**Pushkar Mujumdar**

18110132



**Harsh Patel**

18110062



**Shivam Sahni**

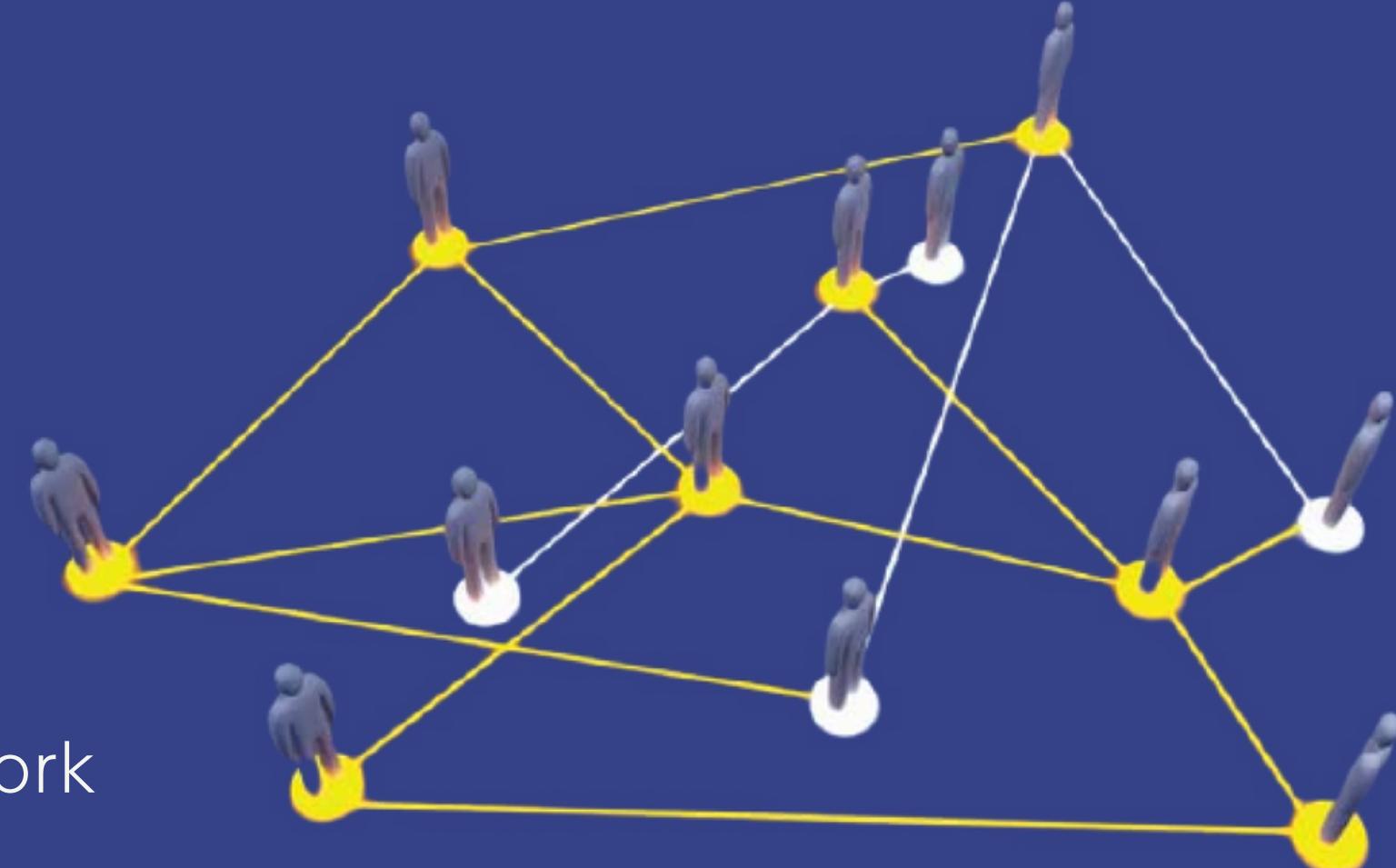
18110159

— 02

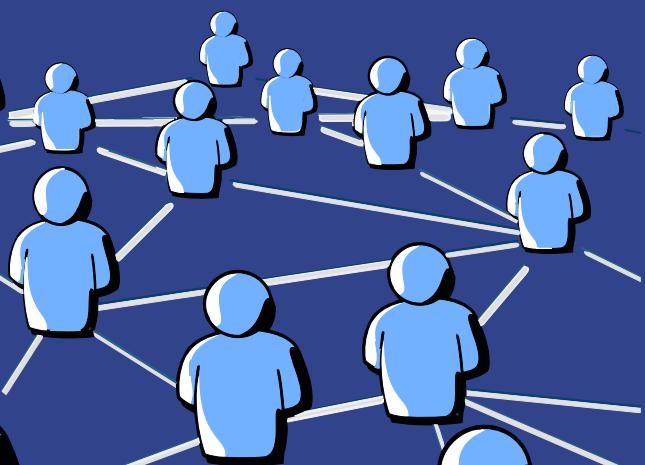


# Signed Networks ?

- Powerful data representation in Social Network Analysis
- Captures relations between entities: trust/distrust , like/dislike...
- Social balance Theory!



— 03



*"An enemy of my friend is my enemy"*



# Dataset



- > Bitcoin OTC Trust Weighted Network.
- > Who-trusts-whom network of Bitcoin users.
  - + Positive edges represent trust
  - Negative edges represent distrust.
- > We reduce the range of edge weights [-1,1]

Dataset statistics	
Nodes	5,881
Edges	35,592
Range of edge weight	-10 to +10
Percentage of positive edges	89%



# What we mean by Reputation?

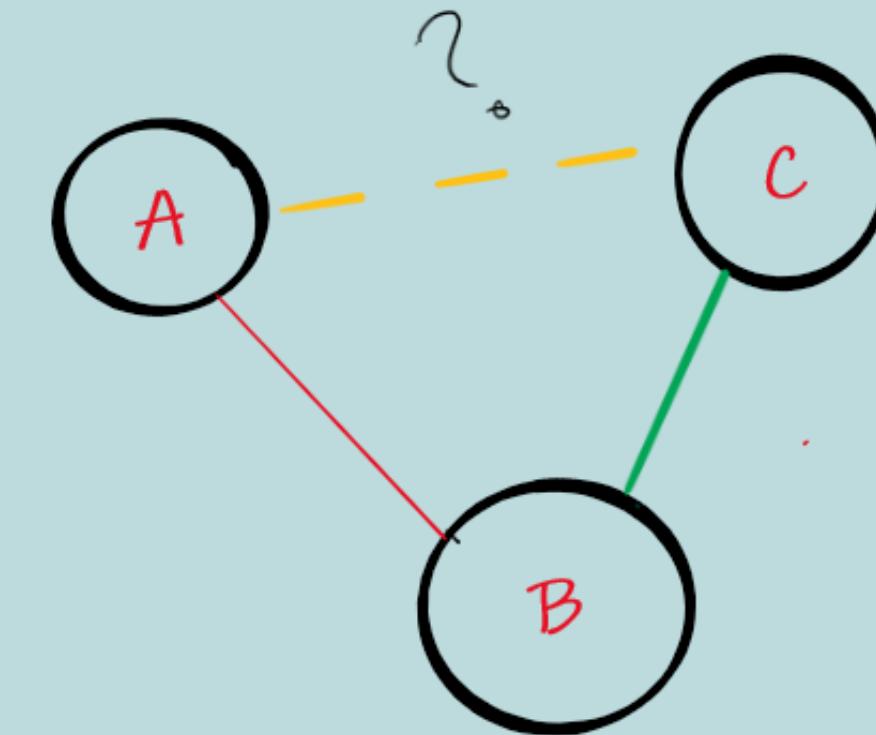
How Non-Fraudulent a User is?

How risky is it to transact with a User?

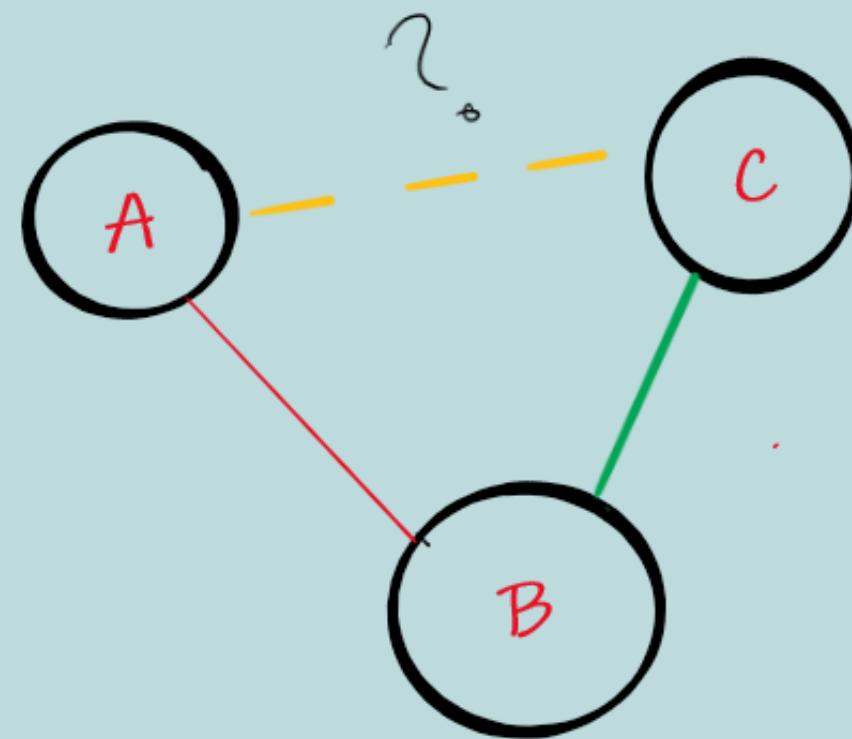
How Trustworthy a User is?

One very important question to answer ---->

Is an Uncomplete graph enough to estimate an answer to  
these questions?



# Social Balance Theory



Here, green edge represents trust and red edge distrust



# Link Prediction



Fairness



Goodness



— 05

$$g(v) = \frac{1}{|in(v)|} \sum_{u \in in(v)} f(u) \times W(u, v) \quad (1)$$

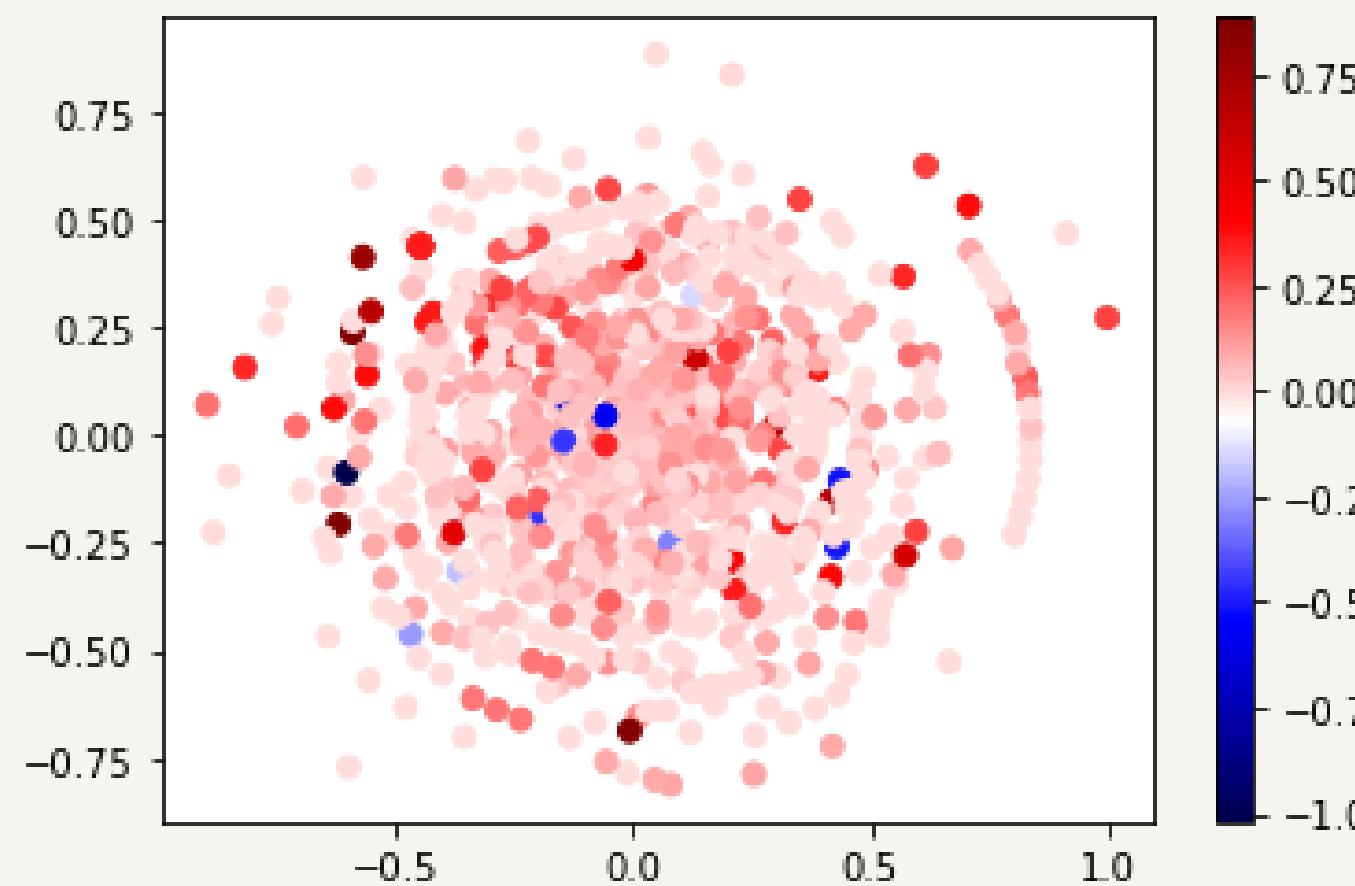
$$f(u) = 1 - \frac{1}{|out(u)|} \sum_{v \in out(u)} \frac{|W(u, v) - g(v)|}{R} \quad (2)$$

$$E(u, v) = f(u) * g(v)$$



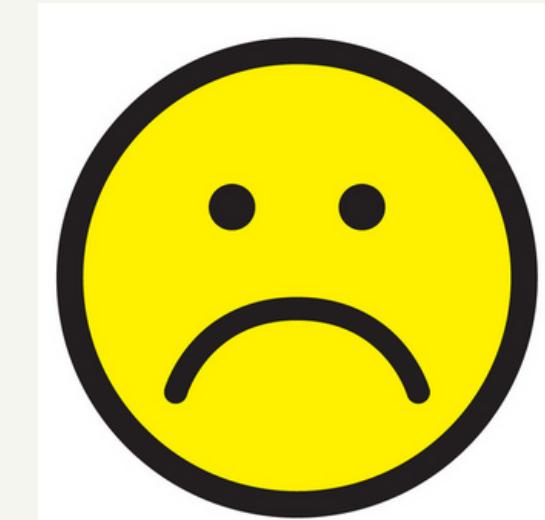
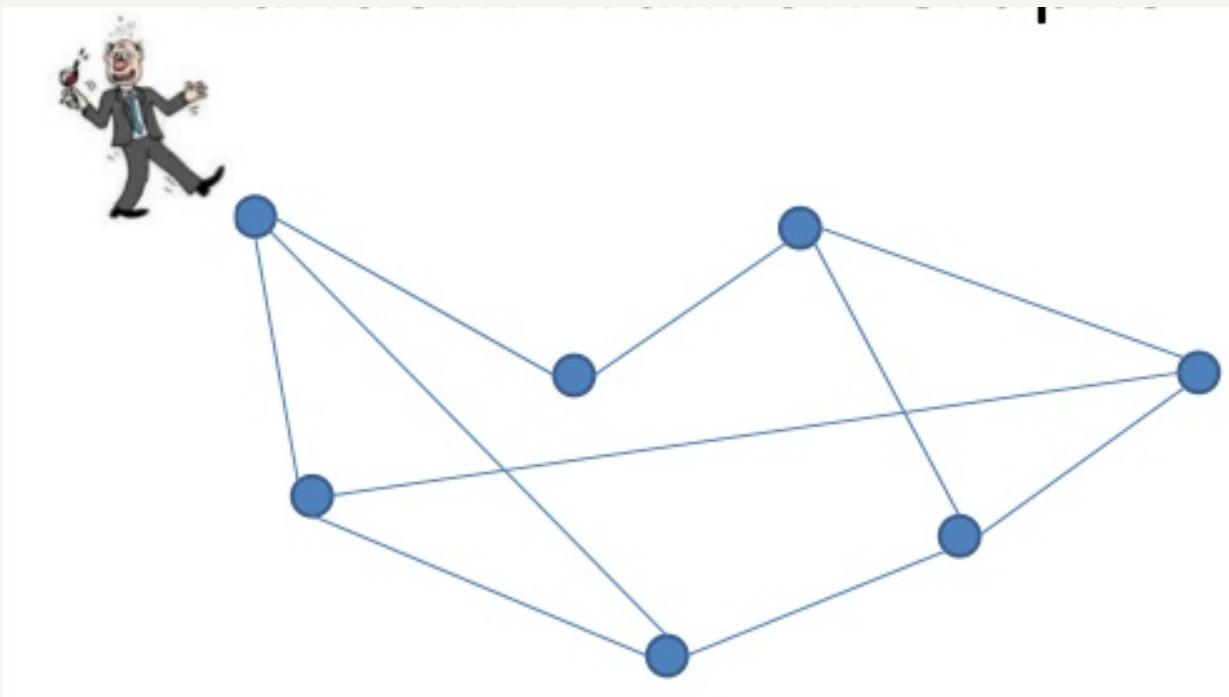
# Reputation using Goodness

We define Reputation as the 'Goodness' of the node in the Complete Graph  
(After Link Prediction)



Reputation of nodes (10% of the dataset)

# SignRank



The probability of reaching a node  $i$  with positive emotion at time  $t + 1$  is given by:

$$\pi_i^{+(t+1)} = \alpha \left( \sum_{j \in IN_{(i)}^+} \pi_j^{+(t)} p_{ji} + (1 - \lambda) \sum_{j \in IN_{(i)}^-} \pi_j^{-(t)} p_{ji} + \frac{\lambda}{2N} \sum_{j=1}^N \pi_j^{-(t)} \right) + \frac{1-\alpha}{2N}$$

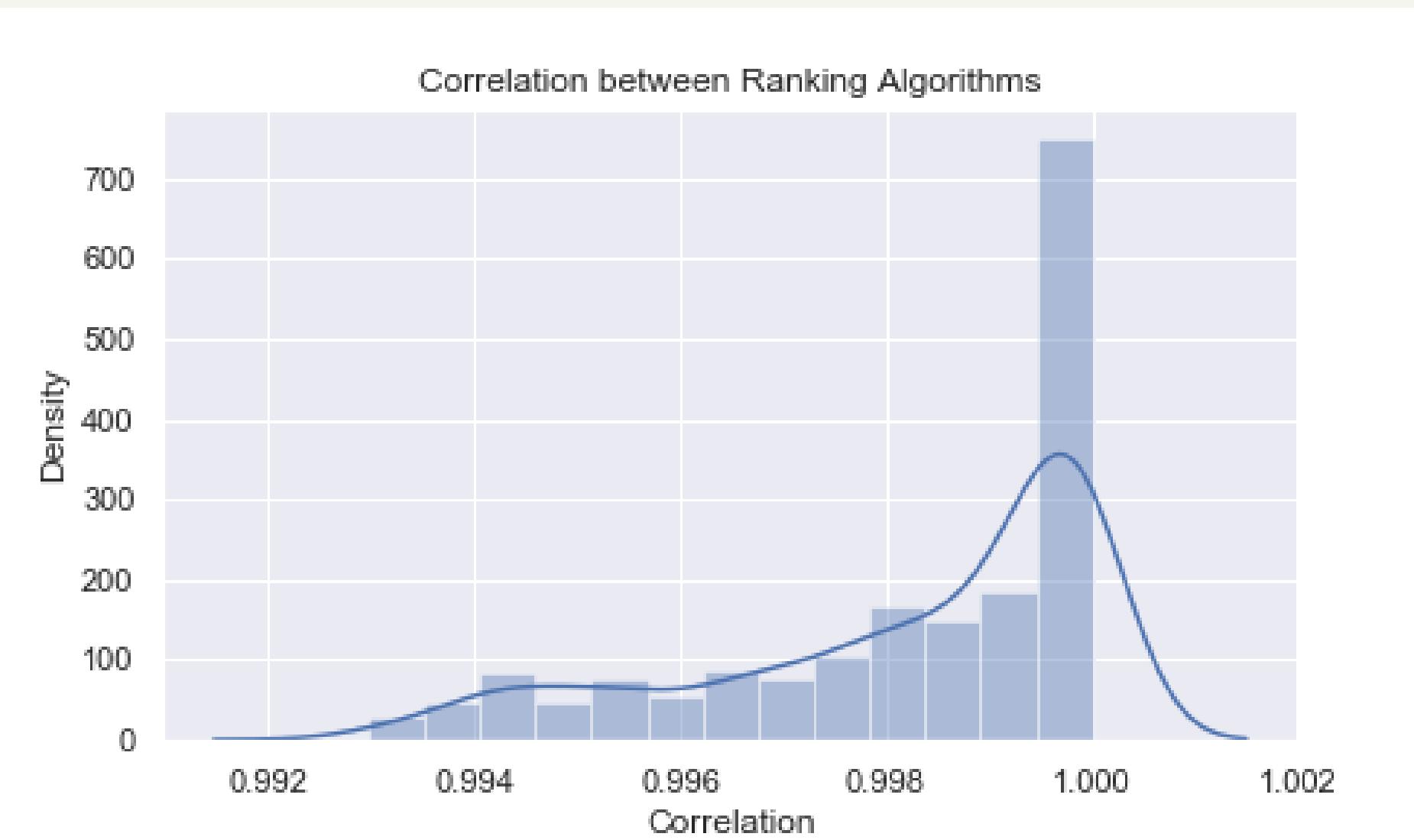
Similarly, the probability of reaching a node  $i$  with negative emotion at  $t + 1$  is given by:

$$\pi_i^{-(t+1)} = \alpha \left( \sum_{j \in IN_{(i)}^-} \pi_j^{+(t)} p_{ji} + (1 - \lambda) \sum_{j \in IN_{(i)}^+} \pi_j^{-(t)} p_{ji} + \frac{\lambda}{2N} \sum_{j=1}^N \pi_j^{-(t)} \right) + \frac{1-\alpha}{2N}$$

# Reputation using SignRank

$$R(i) = \frac{\pi_i^+}{\pi_i^-}$$

# Comparing performance



$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Spearman's rank correlation coefficient

# Next Research Question

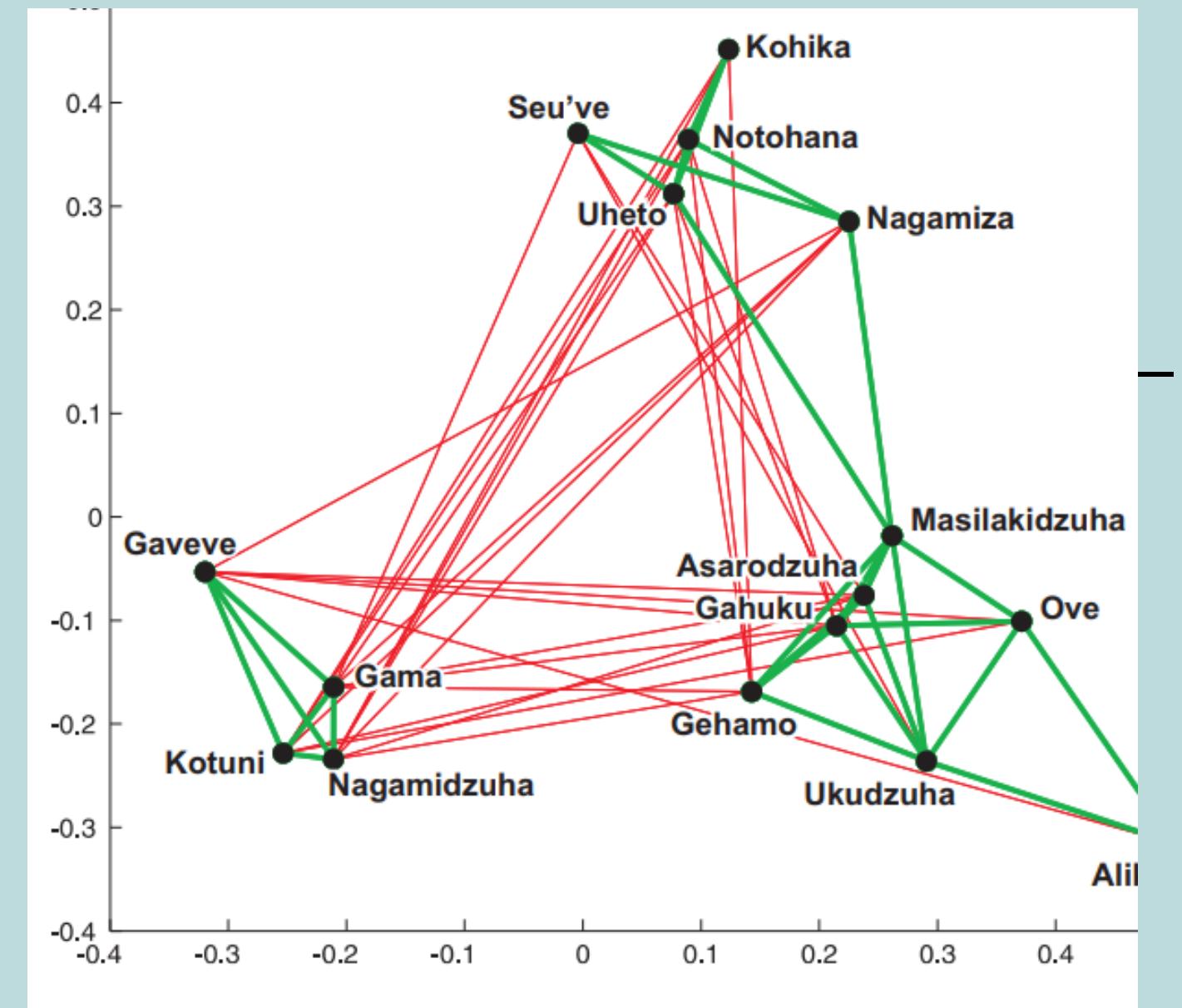
How do we define a community?

How Trustworthy a Community is?

Do we have any Rival communities?

Is this what we call a good **CLUSTERING**?

Communities in Bitcoin Network!



Kunegis et al. (2010)



# Designing Metric

- To estimate trust-worthiness of a community
- Makes use of notion of Agreements

— 06



edges inside community



edges outside community

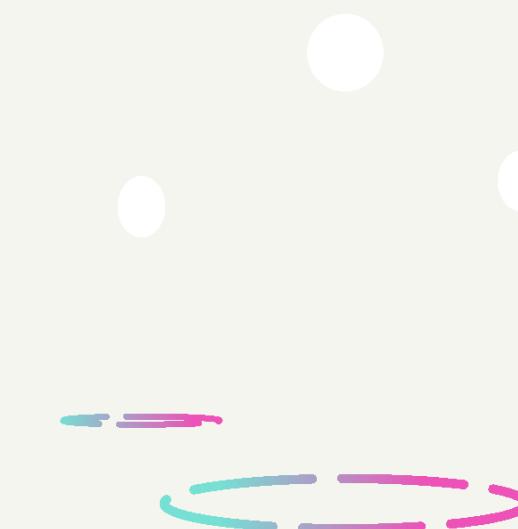
$$Trust_{C_i} = \frac{\sum_{edges} w_{(inside+ve)} + \sum_{edges} |w|_{(outside-ve)}}{N_{edges} C_i}$$





# Clustering

- > Spectral Clustering
  - > The usual Laplacian doesn't work here.



Why?

- > Signed Laplacian Matrix

$$\bar{L} = \bar{D} - A,$$

here,

$$\bar{D}_{ii} = \sum_{j \sim i} |A_{ij}|.$$

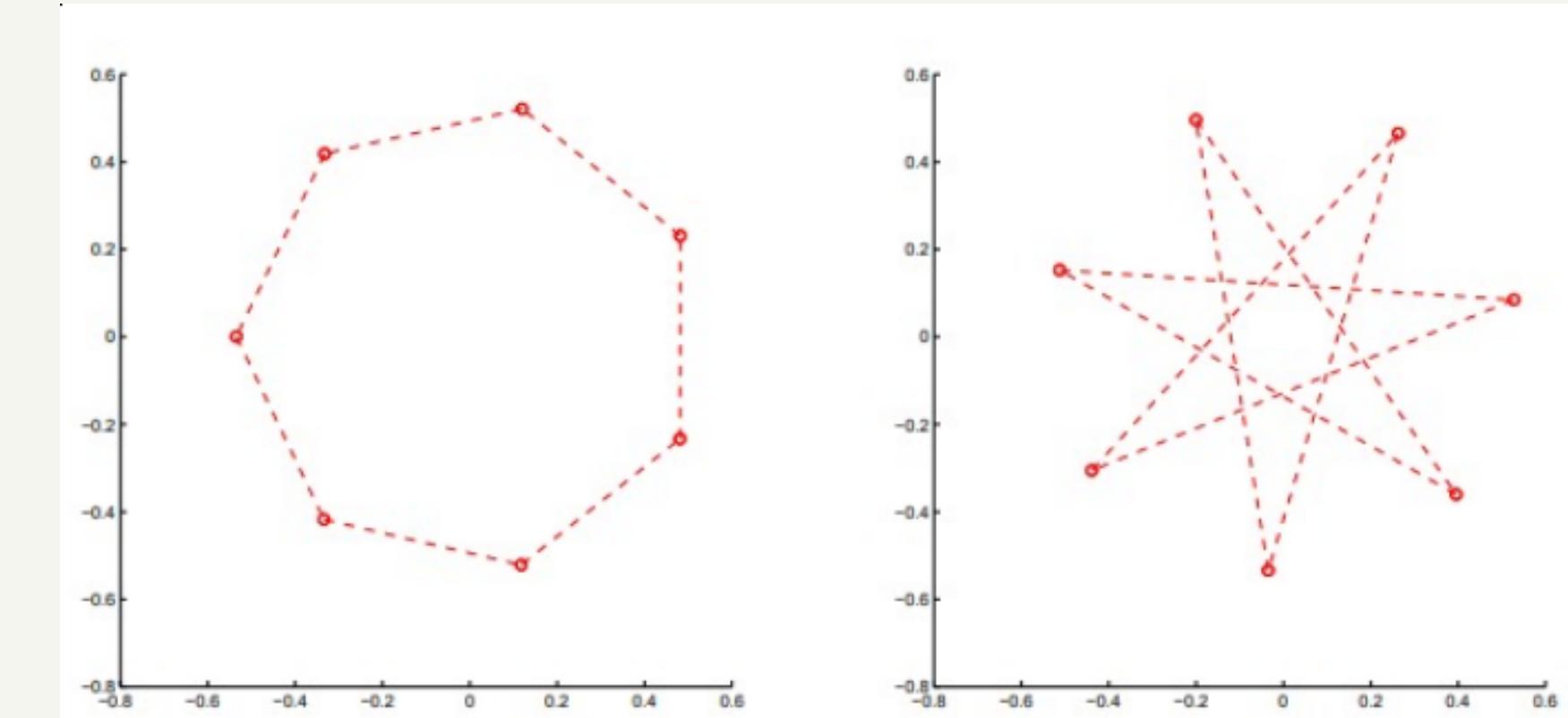


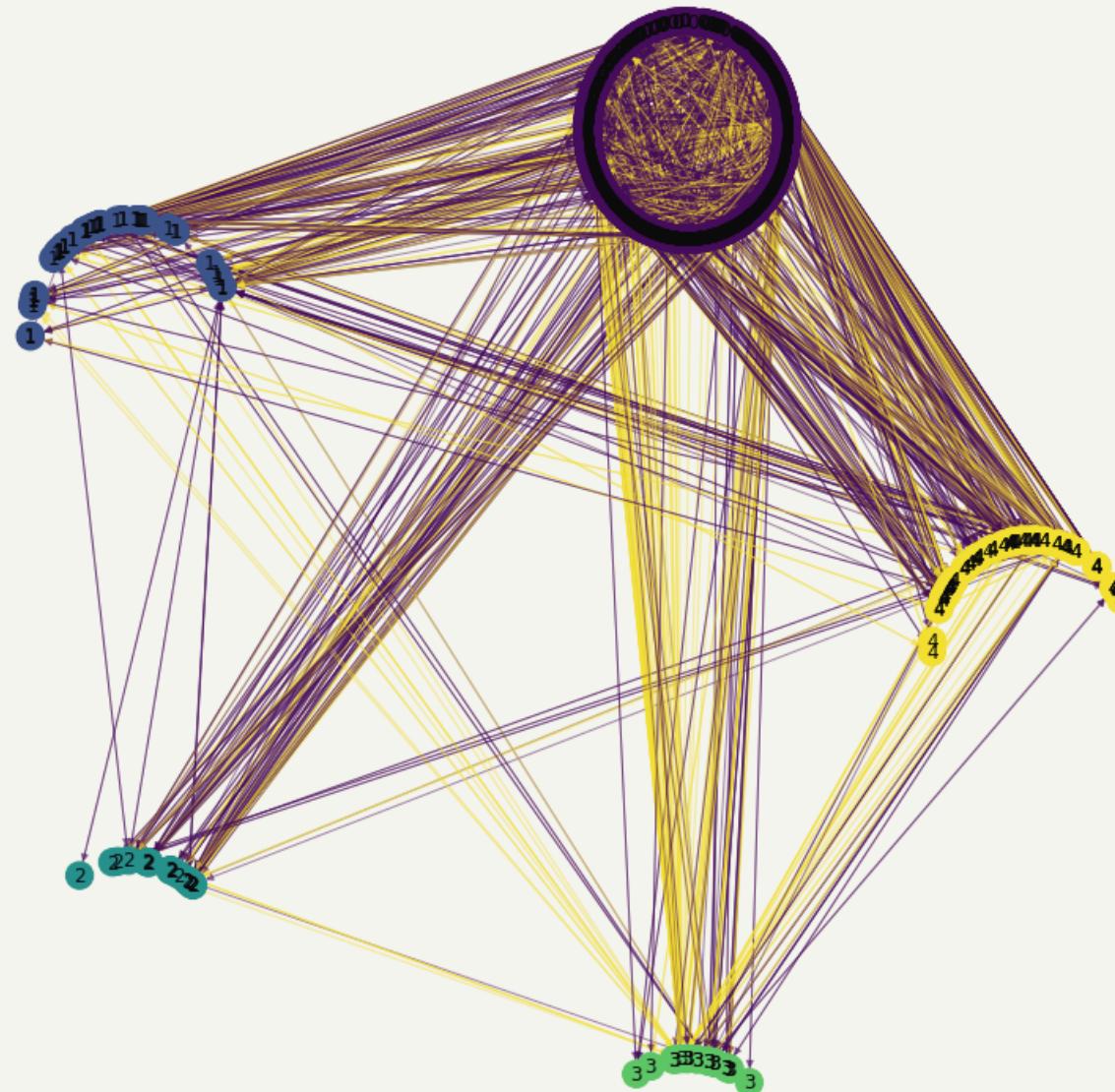
Figure: Example supporting signed spectral clustering. Here, the edges are negative weight. On the left is the embeddings generated from regular laplacian while on the right are the embeddings generated from signed laplacian.

— 06 —





# Spectral Clustering Results!



Cluster No.	No. of Nodes	Inside +ve edges	Outside -ve edges	Trust Ci
1	660	798	342	0.070695
2	69	23	216	0.022101
3	50	3	132	0.017225
4	20	5	33	0.013065
5	18	0	57	0.000846
<b>Total Nodes:</b>	<b>817</b>	<b>Total Trust metric of the Network:</b>		<b>0.123934</b>

- Yellow edges: +ve trust
- Red edges: -ve trust
- Number of clusters optimized using elbow method.
- More +ve trust edges trust inside clusters.
- More -ve trust edges between clusters.

Figure: Signed Spectral Clustering over the original graph





# Correlation Clustering



- What makes a good clustering?

— 06

## Minimizing Disagreements

$\text{sum}_W \{ \text{outside} <+> + |\text{inside} <-> | \}$

---

## Maximizing Disagreements

$\text{sum}_W \{ \text{inside} <+> + |\text{outside} <-> | \}$

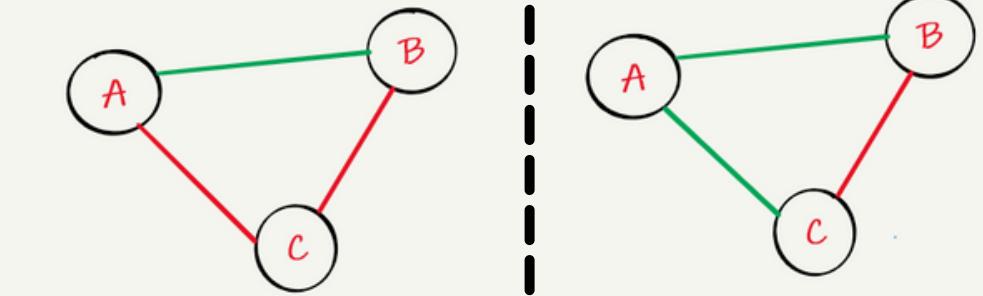




# Constant Factor Approximation for minimizing disagreements



- Disagreements □ Erroneous Triangles



Social Balance Theory is back in action!

— 06

- Allow only  $\delta$ -good nodes in the cluster!

$$|N^+(v) \cap \mathcal{C}| \geq (1 - \delta)|\mathcal{C}|$$

$$|N^+(v) \cap (V \setminus \mathcal{C})| \leq \delta|\mathcal{C}|$$

$\delta$ -Clean Clusters make less mistakes!

Fix  $\delta \leq 0.25$  □ Bound: Atmost 8\*OPT





# Correlation Clustering Results!

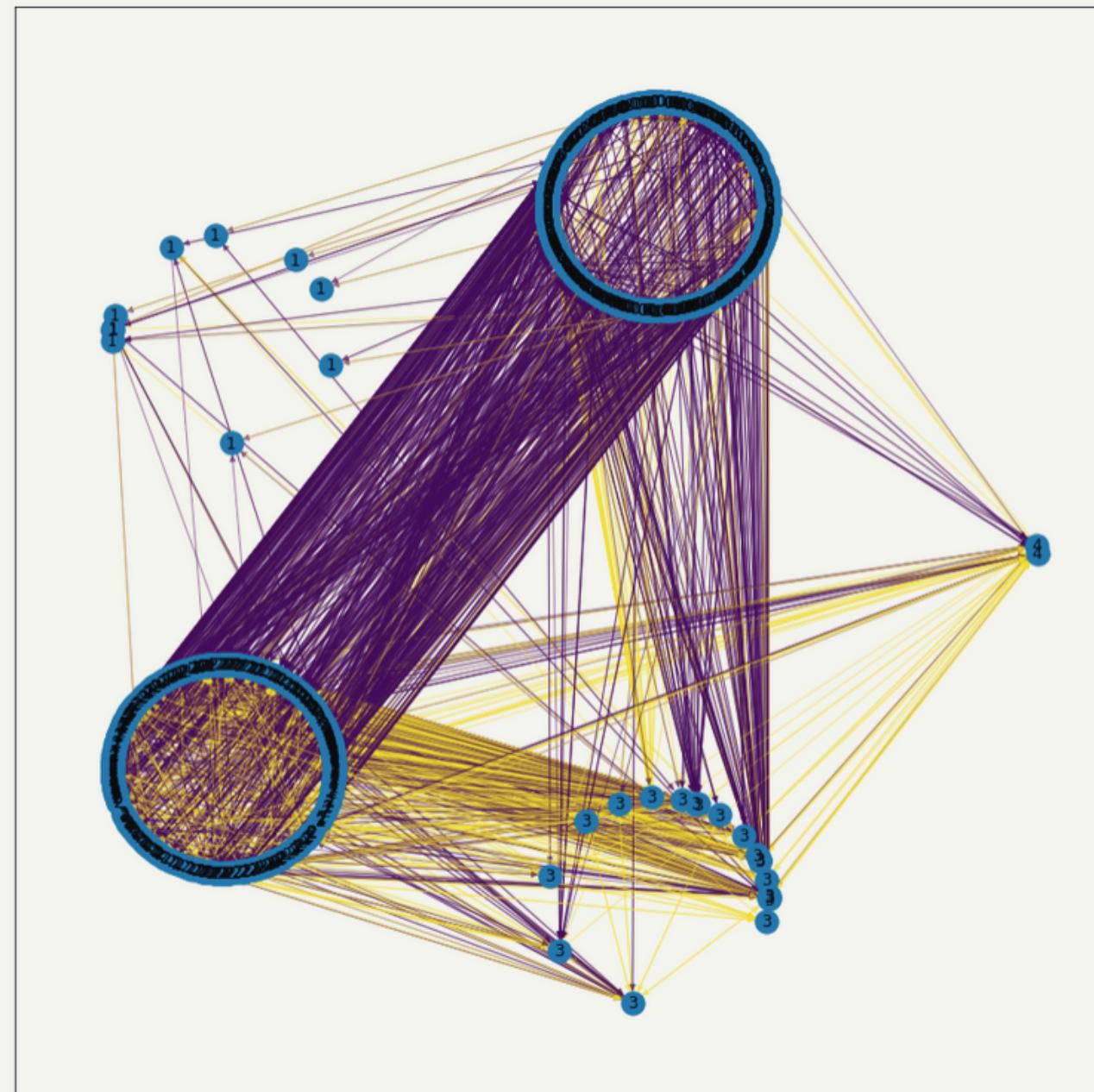


Figure: Correlation Clustering over the original graph

Cluster No.	No. of Nodes	Inside +ve edges	Outside -ve edges	Trust Ci
1	380	598	582	0.095065
2	19	47	260	0.055009
3	2	2	20	0.036433
4	406	0	520	0.024462
5	10	0	15	0.000722
<b>Total Nodes:</b>	<b>817</b>	<b>Total Trust metric of the Network:</b>		<b>0.211691</b>

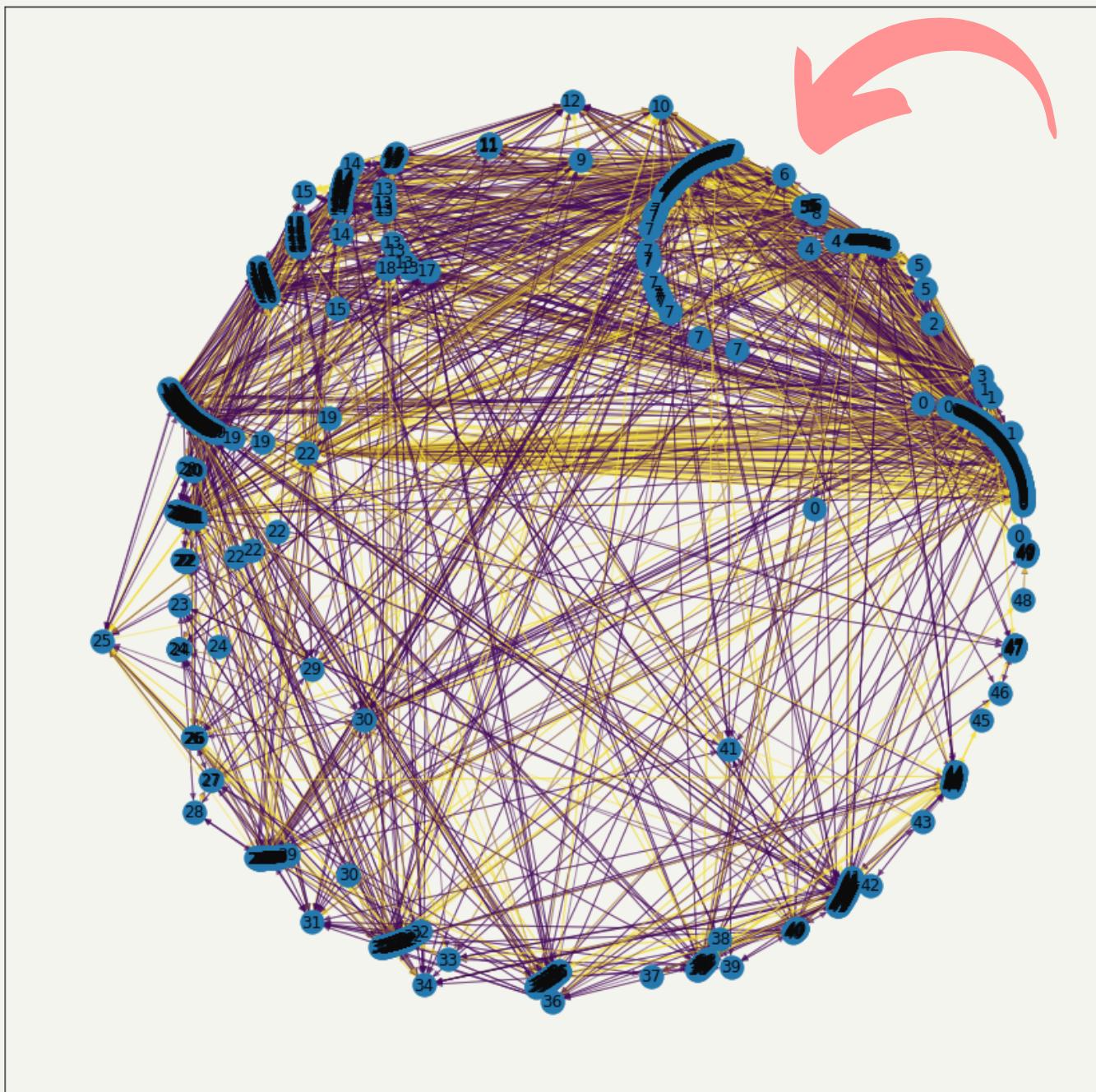
- Yellow edges: +ve trust
- Purple edges: -ve trust
- Cleanliness Parameter  $\delta = 0.2$
- More -ve trust edges between clusters.





# Correlation Clustering Results!

## With Link Prediction



Cluster No.	No. of Nodes	Inside +ve edges	Outside -ve edges	Trust Ci
1	125	10245	39788	0.020745
2	148	14136	45778	0.01974
3	52	1914	18040	0.011124
4	76	2691	24335	0.010605
5	36	594	12270	0.009615
.	.	.	.	.
.	.	.	.	.
50	1	0	253	0.005589
<b>Total Nodes:</b>		<b>817</b>	<b>Total Trust metric of the Network:</b>	

- Yellow edges: +ve trust
- Purple edges: -ve trust
- Cleanliness Parameter  $\delta = 0.05$
- More -ve trust edges between clusters.



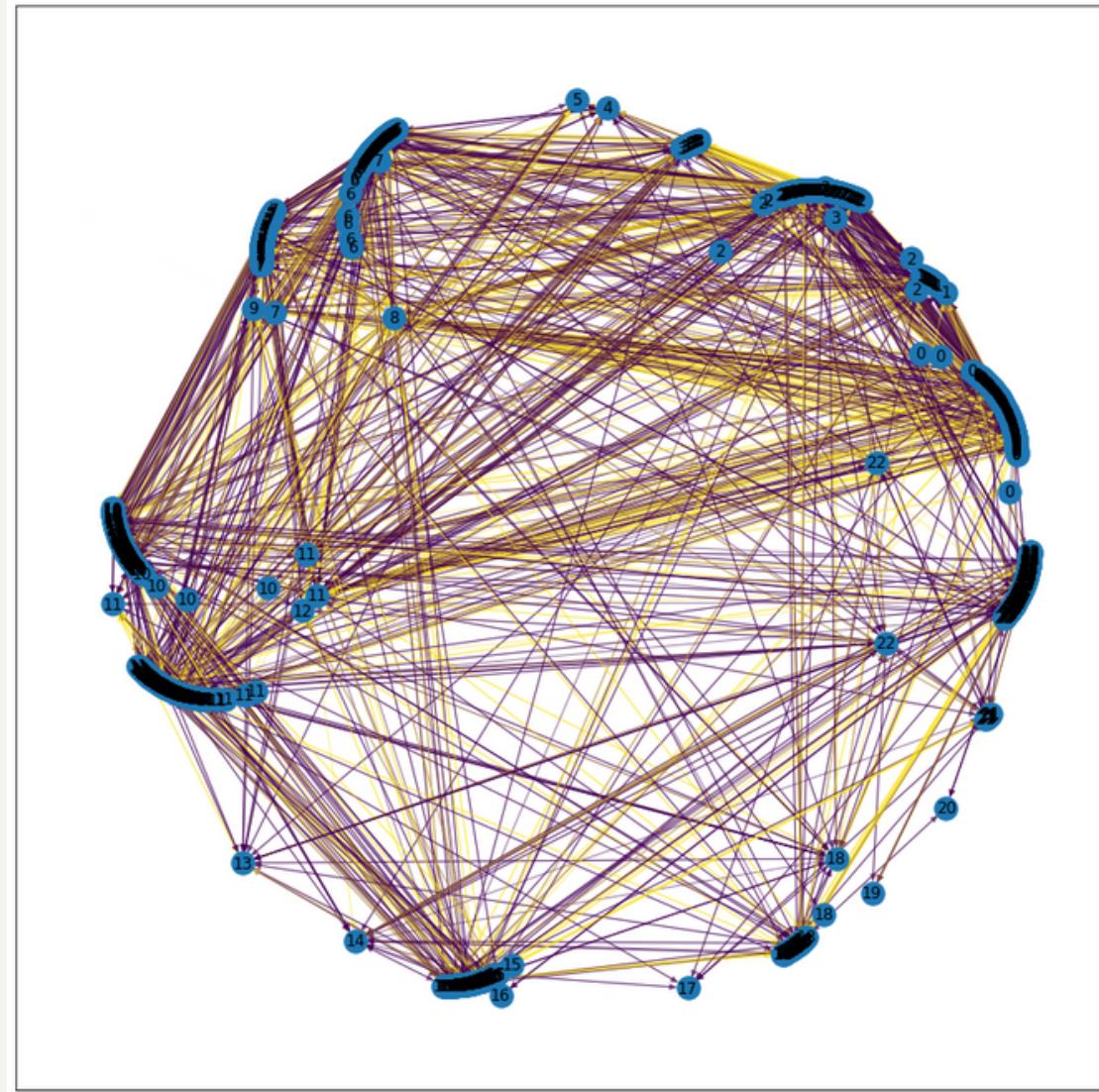
Figure: Correlation Clustering over the complete graph



# Is this a Resilinet Network?



Let's remove the  
**Most trustworthy**  
**community!**



Re-predict the link weights!

Cluster No.	No. of Nodes	Inside +ve edges	Outside -ve edges	Trust Ci
1	96	6672	29403	0.01932
2	90	5118	26974	0.014159
3	80	1971	22066	0.012576
4	56	1373	16783	0.011443
5	86	3207	24379	0.011293
.	.	.	.	.
.	.	.	.	.
23	1	0	352	0.004836
<b>Total Nodes:</b>		<b>692</b>	<b>Total Trust metric of the Network:</b>	
<b>0.1918919</b>				

- Trust levels differ!
- Overall Trust Reduces...?

Figure: Correlation Clustering over the Disrupted graph



# References

- [1] Bansal, N., Blum, A. \& Chawla, S. Correlation Clustering. *Machine Learning* 56, 89–113 (2004).  
<https://doi.org/10.1023/B:MACH.0000033116.57574.95>
- [2] Santo Fortunato,Community detection in graphs,*Physics Reports*,Volume 486, Issues 3–5,2010,Pages 75-174,ISSN 0370-1573,<https://doi.org/10.1016/j.physrep.2009.11.002>.
- [3] G. Palla, I. Derényi, I. Farkas, T. Vicsek Uncovering the overlapping community structure of complex networks in nature and society *Nature*, 435 (2005), pp. 814-818
- [4] Paul Thagard, “What Is Trust?” *Psychology Today*, Sussex Publishers, [www.psychologytoday.com/us/blog/hot-thought/201810/what-is-trust](http://www.psychologytoday.com/us/blog/hot-thought/201810/what-is-trust).
- [5] S. Kumar, F. Spezzano, V. S. Subrahmanian and C. Faloutsos, "Edge Weight Prediction in Weighted Signed Networks," 2016 IEEE 16th International Conference on Data Mining (ICDM), 2016, pp. 221-230, doi: 10.1109/ICDM.2016.0033.
- [6] Cong Wan, Yanhui Fang, Cong Wang, Yanxia Lv, Zegie Tian, Yun Wang, and Huaming Wu. 2019. SignRank: A Novel Random Walking Based Ranking Algorithm in Signed Networks. *Wirel. Commun. Mob. Comput.* 2019 (2019). DOI:<https://doi.org/10.1155/2019/4813717>
- [7] Esmailian, P., Jalili, M. Community Detection in Signed Networks: the Role of Negative ties in Different Scales. *Sci Rep* 5, 14339 (2015).  
<https://doi.org/10.1038/srep14339>
- [8] Kunegis, Jérôme \& Schmidt, Stephan \& Lommatzsch, Andreas \& Lerner, Jürgen \& De Luca, Ernesto \& Albayrak, Sahin. (2010). Spectral Analysis of Signed Graphs for Clustering, Prediction and Visualization. *Proc SDM*. 559-. 10.1137/1.9781611972801.49.
- [9] S. Kumar, B. Hooi, D. Makhija, M. Kumar, V.S. Subrahmanian, C. Faloutsos. REV2: Fraudulent User Prediction in Rating Platforms. 11th ACM International Conference on Web Searchand Data Mining (WSDM), 2018.



Thankyou

— 08

