Trust as a Metric for Resiliency in Signed Social Networks

Harsh Patel
harsh.patel@iitgn.ac.in
Indian Institute of Technology
Gandhinagar
India

Shivam Sahni
shivam.sahni@iitgn.ac.in
Indian Institute of Technology
Gandhinagar
India

Pushkar Mujumdar
pushkar.mujumdar@iitgn.ac.in
Indian Institute of Technology
Gandhinagar
India

ABSTRACT
Recent technological advancements have resulted in a surge in online trading, raising severe concerns about theft and fraud, especially on platforms like Bitcoin OTC (over-the-counter), where users’ identities remain anonymous. To mitigate the risk, it has become essential to capture the reputation of users based on their trade histories. The who-trusts-whom signed network of people has the capability to reflect the nature of such positive and negative relations between the users. It can be used to analyze linkage patterns, strength, and resiliency of such platforms. Due to the dynamic nature of trust between individuals, these trust networks are often vulnerable to link or node failures, making it critical to understand the stability of such systems. In this paper, we consider the problem of quantifying the resiliency of signed networks with the help of trustworthy community structures. We propose a metric for computing the Trustworthiness of a community structure. Using the trustworthiness scores of all communities structures, we generate a pipeline for assessing the resiliency of a signed network. We also show how these generated resiliency scores are concordant with the true nature of the network.

CCS CONCEPTS
• Information systems → Clustering; Reputation systems; Data analytics; Social tagging systems.

KEYWORDS
signed networks, resiliency, community detection, link prediction

ACM Reference Format:
Harsh Patel, Shivam Sahni, and Pushkar Mujumdar. 2018. Trust as a Metric for Resiliency in Signed Social Networks. In Woodstock ’18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY, ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3122445.3122456

1 INTRODUCTION
Resilience is the ability of a complex system to withstand the disturbances caused due to any internal failure or natural changes and recover quickly to maintain its functionality. Understanding how complex networks demonstrate resilience has become quite critical because of the increasing chances of crises in various systems like transportation, communication, and finance [6].

Signed networks are a powerful way of data representation that effortlessly reflect the mixture of positive and negative relations among entities. These signed networks are even used to capture weighted social relationships between people, like the strength of trust/mistrust between two people on a particular scale. In this paper, we tackle the problem of quantifying the trustworthiness of individual communities and the resiliency of the overall network by utilizing the signed trust relationships in the network.

The interplay between the positive and negative edges of a signed network leads to the formation of conflicting communities. Thus, it has become essential to exploit the signed information to identify more stable communities to understand the strength of these networks. There have been previous works like [12], and [11], focusing on measuring the stability of a community in signed social networks.

Transaction networks like the Bitcoin OTC are often incomplete graphs as there may be tacit relations of trust/mistrust that do not appear in the graph. The problem of predicting weights of such links is quite significant and has been well explored by the community using approaches like random-walk based [14], and generative adversarial network (GAN) framework based [3]. In our work, we use the Fairness-Goodness approach developed in [9], which uses the Social Balance Theory [7] to predict the unknown links in the network. According to the Social Balance Theory, “A friend of an enemy is also an enemy”. The work by [5] shows that currently available large online social networks are concordant with the social balance theory. This edge weight prediction is often helpful in improving the traditional tasks in real-world signed networks like network analysis [11] and community detection [13].

We use the predicted edge weights in order to determine a trustworthiness score of a community. Community detection plays a significant role in understanding the underlying patterns in complex networks. However, it has been significantly under-explored for signed networks as compared to unsigned ones [4]. Few previous works include modified Spectral clustering [10], Correlation clustering [2] for community detection in signed networks.

Using such graph clustering approaches over the Bitcoin OTC Trust Network, we identify the trustworthiness of the communities in this transaction network. This trustworthiness metric is further used to quantify resiliency. We show that these resiliency scores are a good estimate for understanding the stability of the network. With this work, we present a significant base for further study of resiliency quantification problem in signed social networks.
2 DATASET
We use the Bitcoin OTC Trust Weighted Network [8] for our experiments. The directed edge e(A, B) represents the reputation of node B w.r.t. the node A. The positive edges represent trust between users, and the negative edges represent distrust between users.

| Statistic             | Value       |
|-----------------------|-------------|
| Nodes                 | 5881        |
| Edges                 | 35592       |
| Range of Edge weight  | -10 to +10  |
| Percentage of Positive Edges | 89%     |

Table 1: Dataset Statistics

For ease in experimentation, we reduce the range of edge weights to −1 to +1 for better interpretation of results.

3 BACKGROUND

3.1 Link Prediction
Suppose we are given a graph with three nodes A, B and C. Given that there is a positive relation between A and B, and a negative relation between B and C, and the relation between A and C is not specified. So without link prediction it would be taken as a neutral relation, which might be erroneous. Here, we cannot ignore the fact that according to the Social Balance Theory, the relation between A and C would likely be negative. Thus, link prediction helps in augmenting the signed network and enables it to capture the social interactions robustly.

A simple algorithm suggested in [9], defines two statistics, Fairness and Goodness. The Fairness of a node quantifies how fair a node is, at judging the Goodness of other nodes. The Goodness defines how good or reputed a node actually is. The mutually recursive definitions of Fairness and Goodness of a node are as follows:

\[
g(v) = \frac{1}{|in(v)|} \sum_{u \in in(v)} f(u) \ast W(u, v)
\]

\[
f(u) = 1 - \frac{1}{|out(u)|} \sum_{v \in out(u)} \frac{|W(u, v) \ast g(v)|}{R}
\]

where \(in(v)\) is the in-degree of node \(v\) and \(out(u)\) is the out-degree of node \(u\), \(W(u, v)\) is the edge weight between node \(u\) and \(v\), and \(R\) is the range normalizing constant.

Now, the weight of the edge from a node \(u\) to node \(v\) is predicted as:

\[
E(u, v) = f(u) \ast g(v)
\]

3.2 Spectral Clustering
Spectral clustering allows us to perform dimensionality reduction on the graph to obtain node embeddings in fewer dimensions. We can subsequently employ clustering techniques on these embeddings to cluster the graph into desired number of clusters. The motivating idea behind spectral clustering is Graph Drawing.

3.2.1 Graph Drawing. The objective of graph drawing is to generate embeddings for nodes such that a node is close to its neighbours [10].

Let \(X_i \in \mathbb{R}^2\) be the coordinates of node \(i\) in the drawing. For positive weighted edges, we get the following vertex equation:

\[
X_i = (\sum_{j=1}^{i} A_{ij})^{-1} \sum_{j=1}^{i} A_{ij}X_j
\]

On simplifying, we obtain,

\[
DX = AX
\]

\[
LX = 0
\]

where \(D\) is the diagonal matrix, \(A\) is the adjacency matrix and \(L = D - A\). We call the L matrix as the unsigned Laplacian Matrix.

To extend this equation to Signed Networks, we need to look at some subtleties. First, we want a node to be close to its positive weight neighbors and far from its negative weight neighbors. Second, we need to change the normalising part in the vertex equation.

For signed network the equation now becomes:

\[
X_i = (\sum_{j=1}^{i} |A_{ij}|)^{-1} \sum_{j=1}^{i} A_{ij}X_j
\]

on simplifying,

\[
\tilde{D}X = AX
\]

\[
\tilde{L}X = 0
\]

where \(\tilde{D}\) is defined as \(\tilde{D}_{ii} = \sum_j |A_{ij}|\) and we obtain the signed Laplacian Matrix, \(\tilde{L} = \tilde{D} - A\).

Figure 1: The figure depicts the embeddings for a Graph containing only negative edges by using unsigned Laplacian embeddings and signed Laplacian embeddings respectively from left to right. We see that the nodes are more separated in the right plot. Thus, the signed Laplacian embeddings perform better here [10].

3.3 Correlation Clustering
Correlation clustering is an algorithm majorly used in a scenario where the relations between the entities of a network are known. The algorithm tries to put similar objects into the same clusters and dissimilar into different clusters.

We use a constant factor approximation algorithm, described in [2], which minimizes the number of Disagreements, i.e., the absolute sum of negative edges inside and the sum of positive edges...
outside each cluster. This iterative algorithm tries to estimate the δ-clean clusters, where δ-clean implies that for each node \( v \in C \), the cluster has at least \((1 - \delta) \cdot |C|\) positive-weighted neighbors inside \( C \) and at most \( \delta \cdot |C|\) positive-weighted neighbors outside \( C \).

One significant advantage of using correlation clustering is that we do not need to fix the optimal number of clusters for our network. The algorithm iteratively determines the best clustering while minimizing the disagreements as defined before. However, the cleanliness parameter \( \delta \) plays a vital role in the clustering performance.

### 4 OUR APPROACH

In this section, we first propose a resiliency metric that helps in capturing the trustworthiness of any individual community in the network. We also devise a few experiments to support the idea of how this metric corresponds to the resilient nature of this network.

\[
\text{Trustworthiness}_{C_i} = \frac{\sum_{\text{edges}} w_{\text{inside+ve edges}} + \sum_{\text{edges}} |w|_{\text{outside-ve edges}}}{\max(N_{\text{nodes inside}} + N_{\text{nodes total}} - N_{\text{nodes outside}})}
\]

This metric is defined over the basic notion of what is popularly known in these settings as Agreements, which means more no. of positive edges inside and negative edges outside a cluster.

We also add a normalizing term similar to the normalizing term seen in the formula for the expansion of a set of nodes in a network [1]. The intuition behind it is that we want our communities to be large and have higher agreements.

#### 4.1 Resiliency Computation Pipeline

For subsequent experiments in this section, 25% of the network is considered for ease of experimentation. We first perform link prediction using the Fairness Goodness algorithm as described in 3.1 on the Bitcoin Trust network. This allows us to capture all the trust relations, that are not explicitly present in the network. We provide the performance evaluation of this algorithm in section 5.1. We then use the spectral 3.2 and correlation clustering 3.3 algorithms to determine all the conflicting communities in the trust network. Further using the trustworthiness metric, the trust level of each community is calculated. In section 5.2, we also perform a validation experiment to verify the effectuality of signed network clustering.

The results for spectral and correlation clustering are shown in Tables 2 and 3 respectively. The total trustworthiness of the network is defined as the sum of trustworthiness values of all the communities determined using the above approach. Using the clustering approaches, we were able to detect the best community in the network on the basis of its trustworthiness.

Now, to quantify and analyse the resiliency of this trust network, we perform the following disruption event: We firstly remove the most trustworthy community from the network and then re-evaluate the total trustworthiness of the network by using the similar heuristics as before. All the edge weights are re-calculated using the link prediction algorithm and the trust levels of the community structures are re-determined by applying both the clustering over the disrupted network. The spectral and correlation clustering results over the disrupted network are shown in Table 4 and 5 respectively.

We here hypothesize that the change in the total trustworthiness of the network due to the above disruption determines how resilient our trust network is. We give a detailed discussion of these results in section 6.

#### 4.2 Comparison with baseline

To check for the efficacy of the resiliency scores, we utilise the temporal nature of the Bitcoin OTC network to generate an year-wise distribution of resiliency scores for the network. As a baseline, we also compute resiliency scores using the following method: the disruption event remains the same as in section 4.1, but, to compute the resiliency scores, we compare the mean absolute error (MAE) of the link predicted values for the entire network before and after the disruption.

We claim that if we see a similar trend in both the techniques for computing resiliency scores over the years, then our hypothesis: the change in total trustworthiness of the network is a good estimate to determine the resiliency of the trust network, is significant. The year-wise distribution of the MAE scores and the change in total trustworthiness of the network \( R^+ \) due to the disruption event are shown in figure 2. We provide a detailed discussion of these results in section 6.
5 PERFORMANCE EVALUATION

5.1 Link Prediction

To evaluate the performance of link prediction algorithm, we use the temporal nature of the Bitcoin OTC Trust Network. This allows us to partition the entire network according to the presence of its entities in a particular time-frame. We evaluate the absolute error between the predicted and the expected edge weights for different split proportions as shown in Table 6. The average MAE score 0.16 is quite low, validating that the link prediction is able to effectively capture the unknown edge weights.

5.2 Signed Clustering

To verify that the clustering techniques on signed networks indeed get us a better metric score, we formed a sub-graph from the original graph by only considering the positive edge weights. Thereafter, we applied the spectral clustering on the sub-graph and compared the results with the signed clustering techniques. The results are summarised in Table 7. The metric scores for each type of clustering allow us to inference that the signed clustering techniques benefit us in getting a better metric score. We claim that higher metric score is representative of a better clustering because of the presence of more agreements in a correctly clustered network.

6 INFERENCES

This section includes the discussion on the results from the experiments performed in section 4.1 and 4.2. The tables 2 and 3 show the trustworthiness scores of the clusters obtained from spectral and correlation clustering, respectively. We observe that the total trust metric computed by correlation clustering is higher than that of spectral clustering. Through this, we claim that correlation clustering performs better than spectral clustering, as it is able to cluster the graph into better communities, i.e., communities having higher trustworthiness scores. The tables 4 and 5 show the trustworthiness scores of the clusters obtained on the disrupted network described in section 4.1 using spectral and correlation clustering, respectively. We observe a radical change in the total trustworthiness of the network which reflects that after the disruption event, the strength of the trust network reduces significantly. Thus, we can interpret that the change in the trustworthiness metric is a good estimate to determine the resiliency of this network. Higher the change in the total trustworthiness, lesser the strength and resiliency of the network.

\[ R = \text{Trustworthiness}_{\text{original}} - \text{Trustworthiness}_{\text{disrupted}} \]

To benchmark the performance of the trustworthiness as a metric for resiliency, in 4.2, we compare our results with a baseline metric: MAE of the link predicted edge weights of the trust network before and after the disruption event. Figure 2 shows the trend of \( R \) and MAE values over the years. We observe a similar trend in both the metrics, which validates that our resiliency metric is concordant with the expected behavior of the network on foreign disruption events.

7 CONCLUSION AND FUTURE WORK

Disruptions are quite common in real-world networks and could lead to a severe effect on both the life and strength of the network. Thus, resiliency estimation is an important task while analyzing any realistic network. With this work, we present a clustering-based approach to quantify the resiliency of signed social networks by exposing them to a custom disruption event. We investigate the performance of our resiliency metric using various experiments on a real-world signed trust network. Our approach can be extended to assess and quantify the resiliency of various real-world signed networks. The current approach lacks in benchmarking the performance of the resiliency metric on possible realistic disruption events. We would like to devise constrained disruptions based on the level of impact on the network for a more generic assessment. We hope that our work serves as a motivation for others to contribute in the field of resiliency estimation for signed social networks.
REFERENCES

[1] Sanjeev Arora, Satish Rao, and Umesh Vazirani. 2009. Expander flows, geometric embeddings and graph partitioning. *J. ACM* 56, 2 (2009), 5.

[2] Nikhil Bansal, Avrim Blum, and Shuchi Chawla. 2004. Correlation Clustering. *Machine Learning* 56, 1 (01 Jul 2004), 89–113. https://doi.org/10.1023/B:MACH.000033116.57574.95

[3] Roshni Chakraborty, Ritwika Das, and Nilotpal Chakraborty. 2020. Link Prediction in Signed Networks. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media (Virtual Event, USA)* (HT ’20). Association for Computing Machinery, New York, NY, USA, 235–236. https://doi.org/10.1145/3372923.3404805

[4] Pouya Esmailian and Mahdi Jalili. 2015. Community Detection in Signed Networks: the Role of Negative ties in Different Scales. *Scientific Reports* 5, 1 (23 Sep 2015), 14339. https://doi.org/10.1038/srep14339

[5] Giuseppe Facchetti, Giovanni Iacono, and Claudio Altafini. 2011. Computing global structural balance in large-scale signed social networks. *Proceedings of the National Academy of Sciences* 108, 52 (2011), 20953–20958. https://doi.org/10.1073/pnas.1109521108 arXiv:https://www.pnas.org/content/108/52/20953.full.pdf

[6] Ilaria Giannoccaro, Vito Albino, and Anand Nair. 2018. Advances on the Resilience of Complex Networks. *Complexity* 2018 (13 Aug 2018), 8756418. https://doi.org/10.1155/2018/8756418

[7] Fritz Heider. 1946. Attitudes and Cognitive Organization. *The Journal of Psychology* 21, 1 (1946), 107–112. https://doi.org/10.1080/00223980.1946.9917275 arXiv:https://doi.org/10.1080/00223980.1946.9917275 PMID: 21010780.

[8] Srijan Kumar, Bryan Hooi, Disha Makhija, Mohit Kumar, Christos Faloutsos, and VS Subrahmanian. 2018. Rev2: Fraudulent user prediction in rating platforms. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. ACM, 333–341.

[9] Srijan Kumar, Francesca Spezzano, V. S. Subrahmanian, and Christos Faloutsos. 2016. Edge Weight Prediction in Weighted Signed Networks. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*. 221–230. https://doi.org/10.1109/ICDM.2016.0033

[10] Jérôme Kunegis, Stephan Schmidt, Andreas Lommatzsch, Jürgen Lerner, Ernesto W De Luca, and Sahin Albayrak [n.d.]. Spectral Analysis of Signed Graphs for Clustering, Prediction and Visualization. 559–570. https://doi.org/10.1137/1.9781611972801.49 arXiv:https://epubs.siam.org/doi/pdf/10.1137/1.9781611972801.49

[11] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. *Signed Networks in Social Media*. Association for Computing Machinery, New York, NY, USA, 1361–1370. https://doi.org/10.1145/1753326.1753352

[12] R. Sun, C. Chen, X. Wang, Y. Zhang, and X. Wang. 5555. Stable Community Detection in Signed Social Networks. *IEEE Transactions on Knowledge and Data Engineering* 01 (dec 5555), 1–1. https://doi.org/10.1109/TKDE.2020.3047224

[13] R. Sun, C. Chen, X. Wang, Y. Zhang, and X. Wang. 5555. Stable Community Detection in Signed Social Networks. *IEEE Transactions on Knowledge and Data Engineering* 01 (dec 5555), 1–1. https://doi.org/10.1109/TKDE.2020.3047224

[14] V. A. Traag and Jeroen Bruggeman. 2009. Community detection in networks with positive and negative links. *Phys. Rev. E* 80 (Sep 2009), 036115. Issue 3. https://doi.org/10.1103/PhysRevE.80.036115

[15] Cong Wan, Yanhui Fang, Cong Wang, Yanxia Lv, Zehao Tian, and Yun Wang. 2019. SignRank: A Novel Random Walking Based Ranking Algorithm in Signed Networks. *Wireless Communications and Mobile Computing* 2019 (09 Apr 2019), 4813717. https://doi.org/10.1155/2019/4813717