Advanced Production and Industrial Engineering
R.M. Singari and P.K. Kankar (Eds.)

doi:10.3233/ATDE220742

Deduction of Edge Signs in Bitcoin Alpha Social Network Modelled as a Signed Graph

Jannat Kaur BABBAR[1], Keshav JINDAL[2], Parigya JAIN[3] and Payal[4]

[1, 2, 3, 4] *Department of Applied Mathematics, Delhi Technological University*

Abstract. Signed graphs have a wide array of applications in the social networking domain as the industry and platforms highly rely on forming trust links between users so as to smooth out the process of interacting virtually. Signed graphs facilitate this process by providing a mode of representing such networks as graphs that have edges with a positive/negative sign that help in defining the nature of relationship between the nodes(that can represent the users of the platform or any respective representatives of the platform) of the graph.
In this paper we have dealt with the platform of bitcoin alpha that is now coming into notice due to cryptocurrency's rising popularity. Trading online can be risky and thus the entire platform is functional on the principle of trust/distrust between such anonymous users. We have attempted to formulate the social network of bitcoin alpha platform into a signed graph and predict the links to establish trust/distrust between any two users in the entire graph using concepts of balanced and unbalanced graph theories, and fairness and goodness measures of vertices. Fairness of a user denotes how reliable the rating given by that particular user to others is, whereas goodness measures how likeable or trustworthy a particular user of the website is. Using these metrics, we have attempted to solve this problem.

Keywords. Signed Network, Balance theory, Positive Links, Trust, Distrust

1. Introduction

With the growth of Web 2.0 in 2000's, a new type of media emerged on the internet in the form of websites called social media. These were the websites where the user generated content in the form of text posts, images, videos, comments etc. With time, these social media websites have shown immense growth. For example, the daily active users of facebook in 2011 were around 417 million and today in 2021, that number has reached 1.9 billion. Viewing the users as nodes and relationships between them as edges, these social media networks can be represented as graphs and the concepts of graph theory can help us understand them in a better way. A large amount of existing research [1] on social media networks has focused mainly on positive relationships. But the edge of the graph can have both positive or negative signs. Consider two users u and v in a social network and let say there is a link between them, then either that link has a sign positive or a sign negative. We call such networks as signed social networks because they contain positive and negative links. A positive link indicates trust, like, or consent to the setting, and a negative link indicates that the setting is suspicious, disliked, or disapproved. This represents the real world scenario in an even better way. In recent times, research on signed networks has attracted a lot of attention. The increasing availability of large scale social media data on the Internet makes it useful for various social media analytics tasks such as community search and link prediction, and also various traditional data mining tasks [2, 3, 4, 5] like recommendations and feature selections. The rest of the paper is ordered as follows. In section 2, signed graph has been explained. Then, in section 3, differences between balanced and unbalanced graphs are explained. Then, in section 4, review of some related work. In section 5, formulated the problem for this paper. In section 6, describe the dataset that is used to solve the formulated problem. In section 7, analyse the problem statement and discuss the approach and solution to solve the problem. A conclusion of the findings is included in section 8.

2. SIGNED GRAPH

In the field of graph theory in mathematics, a signed graph is a graph where each edge has a positive or a negative sign. Fig - 1 shows an example of a signed graph. The edges marked with green colour are of positive sign and the edges marked with red colour are of negative sign. The name signed graph was first given by Frank Harary in a mathematical paper of his in 1953 [1]. A signed graph G, or briefly an s-graph, consists of a set E of n points $P_1, P_2, \ldots, P_n$ together with two disjoint subsets L^+, L^- of the set of all unordered pairs of distinct points. The elements of the sets L^+, L^- are called positive lines and negative lines respectively. Whenever there is a group of objects or a group of people and we want to define some kind of interpersonal relationships among them, a signed graph comes handy. The easiest approach to study such a group of people is to make a graph that has the objects or the people as the nodes and the edges between 2 nodes define the relationship between them. The presence of an edge between two users x and y means that x and y are related in some fashion. Now further, we can define the sign of the edge. The edge can either have + sign or a - sign. A positive edge sign indicates trust, like, or consent to the setting, and a negative edge sign indicates that the setting is suspicious, disliked, or disapproved. This represents the real world scenario in an even better way.

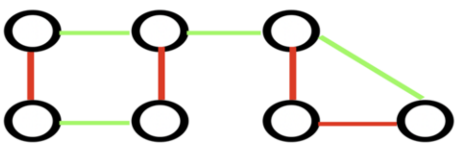

Fig - 1 An Example of a Signed Graph

3. BALANCED AND UNBALANCED SIGNED GRAPH

A signed graph can be marked as a balanced graph if all the possible cycles that are present in the signed network are ultimately positive. The inference of positive here is that if we try to segregate the nodes of the graph into separate groups and then the nodes in the same group form a positive edge with each other and the nodes in a different group form a negative edge with each other. Unbalanced graph is the one that has a minimum of one cycle existing in the network that denotes a negative sign.

4. RELATED WORK

In [6], a real life graph network of a social media website was considered and modelled as a signed graph network. The links here were studied for mainly two theories that are status theory and structural balance theory. In [7, 8, 9, 10] also, social media networks were considered as signed graphs and links were predicted. In [11] the authors considered a two-layer social network and then developed an optimal algorithm to predict the signs of next links in any of the layers. The node based features used were reputation, that would mainly denote the popularity of a node in the social network, and optimism, where a more optimistic node is the one that is likely to point positively to other nodes. Support vector machines (SVMs) , Naive Bayes and K-nearest neighbour (KNN) were used as classifiers. Then, many machine learning based methods were analysed and compared with the baseline edge sign prediction methods used in single-layer networks. It was found that the node based and meta path based feature set, i.e. the feature set that included both these features performed best with an accuracy of 84%, whereas baseline methods had an accuracy of around 75%. In [12], a comparison was made with the theories of social psychology, and how the results obtained were inclined with the ones suggested by balance theory and status theory. A transfer learning-based algorithm was used in [13], [14]. Four explicit topological features were computed. These were based on degree of node, count of triads, how far nodes are spread etc. Apart from the explicit topological features, some latent or hidden topological features were also added to the feature vector. These hidden features were calculated using the NMTF(Non-Negative Matrix Tri Factorization) technique proposed by the authors. Signed social networks are those that have social relationships that can take one among two forms: positive or negative. Many methods for tasks including community discovery, multitask feature learning, node categorization, link prediction, and spectral graph analysis have been developed to mine signed networks. Researchers have recently begun to learn low-dimensional vector representations for a social network, thanks to the invention of network representation learning. SNE adopts the log-bilinear model and includes two signed type vectors to capture the positive or negative relationship of all the edges that come along the path. SiNE computes a new objective function guided by social theories to understand signed network embedding; it offers to add a virtual node to strengthen the training process. SIDE provides a linearly scalable method that supports balance theory along with random walks to obtain the low dimensional vector for the directed signed network. SIGNet merges balance theory with a specialised random and new sampling techniques in directed signed networks. The crux of these techniques are to define an objective function that includes sociological theory and then using some machine learning techniques to improve look-up embedding. Apart from these, Graph Neural Networks(GNNs) have been quite useful in many graph analysis and graph classification problems. They also come very handy in the semi-supervised node classification task. The GNN concept was initially suggested in [15-16], that expanded neural networks to handle data represented in graph domains. Recently, techniques like Convolution, Attention and other mechanisms have been tested by researchers.

5. FORMULATION OF PROBLEM

Bitcoin-alpha is a platform where people can trade cryptocurrency with other users of the platform. Since the users are anonymous, it can become tricky to trust other users for trade, thus to solve this issue, there is a system to rate a user’s trust that can be viewed as a trust score and then every user has an associated trust score that can be used as a reference for any user before indulging in trade with them. So now, to view this problem as a signed graph problem, let us assume that the users of the bitcoin alpha platform represent the nodes, this network of nodes will form a connected graph where every user forms a link with every other user. Now the problem to be solved here will be to predict the nature of link between any two users or we can say to predict the sign of the edge between two nodes. A positive link will indicate a trustworthy relation between two users and a negative link will indicate potential fraud in trade between the users. The graph formed in this problem will be a directed signed graph i.e. let us say a user u gives a positive rating to another user v but v gives a negative rating to user u, then the directed edge from u to v is positive but v to u is negative. Now we consider all the possible scenarios that can take place while drawing a parallel of users of bitcoin alpha represented as a signed graph network and justify the link prediction for these scenarios. Here considering a,b,c as three users of bitcoin alpha, we fix the signs between the user a , b and between the user a , c. We consider all the possible edge signs that can happen between these users. Now, we compute the edge sign from b to c and from c to b. So, for example, in Fig - 2, b trusts a and a does not trust c. So, now, to compute the sign from b to c, b will make use of the node a and hence, will not trust c. So, the edge sign from b to c is +. Now, for the edge sign c to b, we make use of balance theory in signed graphs, and as there is a cycle formed, so the product of edge signs should be positive. And thus, the edge sign c to b is -. Similarly, the edge signs of all other scenarios can be computed.

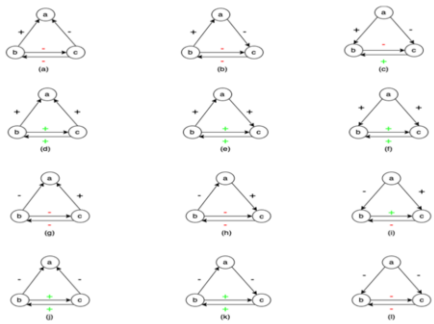

Fig - 2 All possible scenarios with three nodes in a signed graph

6. DATASET DESCRIPTION

In our paper, we have used the dataset collected from the Bitcoin Alpha website. It's a cryptocurrency exchange website where users can come and trade several cryptocurrencies like bitcoin, ethereum, litecoin etc. But now as the users are anonymous on the internet, these exchanges allow their users to give ratings to other users depending on how they feel about their trustworthiness. On Bitcoin Alpha, users can give each other a rating between -10 to 10. Let say, there are two users, u and v, then if u gives a rating of -10 to v, then it means that u thinks v as a fraudster and if u gives a rating of 10 to v, then it means that u trusts v as he trusts himself.

7. ANALYSIS AND DISCUSSION

We have represented the bitcoin alpha dataset as a signed network, wherein every node represents a user of the website, who is able to give trust ratings to other users, as well as receive ratings from them. We take the edge sign to be positive if the rating to a particular user lies between +1 to +10, and negative if it lies between -10 to 0. The problem statement is to predict the sign of edges between two users in the bitcoin alpha dataset. In a signed network, where users have the privilege of giving ratings to other users, it is also important to know the reliability of each rating, apart from knowing the average rating of every user. In order to use these two measurements so as to predict edge signs accurately, we will be using the concepts of Fairness and goodness for every node of the graph. Fairness in social networks that give users the option of giving ratings to others, there is a high chance of people creating multiple accounts in order to increase their own rating and decrease the rating of other users who are rated well. In order to differentiate between such users, we can use the concept of fairness to measure how trustworthy a particular person using the website is. Rating given by a person who has high fairness should be given more importance. Thus this should be used as a criteria while predicting the edge sign. Goodness is a concept that mainly defines how well a person is rated by other users. If a user has high goodness, it would mean that more fair people find them to be trustworthy, while less goodness implies that they are not trusted much among the fair people. In our work, we have calculated the goodness and fairness of each node, in order to predict the sign of the unknown edges. While calculating the goodness of an edge, the average indegree of each vertex is calculated, where the incoming edges are weighted by the amount of fairness of the node from where the edge is originating. In this way, the rating of a person is given more importance if it comes from a fair user. If we take two vertices x and y, with an edge originating at x, and going towards y, it would basically mean that a user x is rating a user y. Hence, we can predict the sign of the edge between them as the sign of

$$F(x) * G(y) \qquad (1)$$

Where F(x) denotes the fairness of vertex x and G(y) denotes the goodness of vertex y. Fig. 3 shows, Graph plotted with Percentage Accuracy vs Percentage of Edges Removed from Training Data.

8. CONCLUSION

From the Bitcoin Alpha dataset, we chose our training set to include 90%, 80%, 70%, 60% and 50% of the known edges. Goodness and fairness features were calculated for these edges, and that formed our training data. The rest of the edges were part of our test set, and using the values for fairness and goodness calculated, we predicted the signs of the test set edges. The average accuracy came out to be 86.6%. We used fairness and goodness metrics to predict the edge sign of unknown edges in a signed graph, with enhanced accuracy. Hence, the outcome is that these two metrics can be used for the edge sign prediction problem and they give fairly good accuracy. Social network sites where users can give each other ratings on parameters such as trust/distrust, friend/foe can be modelled as a signed graph. Using edge sign prediction techniques, we can deduce how this graph will evolve in future. A use case of this edge sign prediction problem was found in social networks like Bitcoin Alpha, where all users are anonymous and it's important for users to know the current rating of another user before trading. Whether or not another user can be trusted for trading can be found out using their current rating. Modelling Bitcoin Alpha as a signed network, we have performed the task of predicting the edge signs using balance theory and the fairness and goodness measures of nodes.

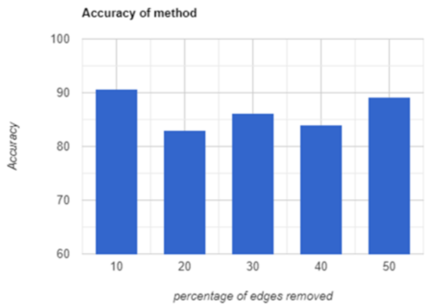

Fig - 3 Graph plotted with Percentage Accuracy vs Percentage of Edges Removed from Training Data

REFERENCES

[1] Harary, Frank, and Jerald A. Kabell. "A simple algorithm to detect balance in signed graphs." Mathematical Social Sciences 1.1 (1980): 131-136.

[2] M. Burke and R. Kraut. Mopping up: Modelling wikipedia promotion decisions. In CSCW, pages 27–36, 2008.

[3] A. Argyriou, T. Evgeniou, and M. Pontil. Multi-task feature learning. In NIPS, pages 41–48, 2006.

[4] A. K. Menon and C. Elkan. Link prediction via matrix factorization. In Machine Learning and Knowledge Discovery in Databases, pages 437–452. Springer, 2011.

[5] X. Zhu, Z. Ghahramani, J. Lafferty, et al. Semi-supervised learning using gaussian fields and harmonic functions. In ICML, volume 3, pages 912–919, 2003.

[6] Y. Freund and R. E. Schapire. A decision-theoretic generalisation of on-line learning and an application to boosting. In EuroCOLT, pages 23–37, 1995.

[7] Jure Leskovec, Daniel Huttenlocher, Jon Kleinberg. Signed Networks in Social Media. arXiv:1003.2424, 2010.

[8] Jiliang Tang, Shiyu Chang, Charu Aggarwal and Huan Liu. Negative Link Prediction in Social Media. arXiv:1412.2723, 2014

[9] S. Papadopoulos, Y. Kompatsiaris, A. Vakali, and P. Spyridonos. Community detection in social media. Data Mining and Knowledge Discovery, 24(3):515–554, 2012.

[10] D. Wang, D. Pedreschi, C. Song, F. Giannotti, and A. Barab ́asi. Human mobility, social ties, and link prediction. In KDD, pages 1100–1108. ACM, 2011.

[11] S.-H. Yang, A. J. Smola, B. Long, H. Zha, and Y. Chang. Friend or frenemy?: predicting signed ties in social networks. In SIGIR, pages 555–564. ACM, 2012.

[12] Mahdi Jalili, Yasin Orouskhani, Milad Asgari, Nazanin Alipourfard and Matjaz Perc. Link prediction in multiplex online social networks. In The Royal Society, 2017.

[13] Jure Leskovec, Daniel Huttenlocher, Jon Kleinberg. Predicting positive and negative links in online social networks. In WWW'10, pages 641-650, 2010

[14] W. Dai, Q. Yang, G.-R. Xue, and Y. Yu. Boosting for transfer learning. In ICML, pages 193–200, 2007.

[15] J. Ye, H. Cheng, Z. Zhu, and M. Chen. Predicting positive and negative links in signed social networks by transfer learning. In WWW, pages 1477–1488, 2013.

[16] Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G.: The graph neural network model. IEEE Transactions on Neural Networks 20(1), 61–80 (2009). https://doi.org/10.1109/TNN.2008.2005605