Network Analysis with Negative Links

Tyler Derr
Data Science and Engineering Lab
Michigan State University
derrtyle@msu.edu

ABSTRACT

As we rapidly continue into the information age, the rate at which data is produced has created an unprecedented demand for novel methods to effectively/efficiently extract insightful patterns. Then, once paired with domain knowledge, we can seek to understand the past, make predictions about the future, and ultimately take actionable steps towards improving our society. Thus, due to the fact that much of today's big data can be represented as graphs, emphasis is being taken to harness the natural structure of data through network analysis. Furthermore, many real-world networks can be better represented as signed networks, e.g., in an online social network such as Facebook, friendships can be represented as positive links while negative links can represent blocked users. Hence, due to signed networks being ubiquitous, in this work we seek to provide a fundamental background into the domain, a hierarchical categorization of existing work highlighting both seminal and state of the art, provide a curated collection of signed network datasets, and discuss important future directions.

KEYWORDS

signed networks, negative links, network analysis

ACM Reference Format:
Tyler Derr. 2020. Network Analysis with Negative Links. In *The Thirteenth ACM International Conference on Web Search and Data Mining (WSDM '20), February 3–7, 2020, Houston, TX, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3336191.3372188

1 INTRODUCTION AND MOTIVATION

Most existing network analysis research has focused on unsigned networks (or networks with only positive links). However, in many real-world social systems, relations between two nodes can be represented as signed networks with positive and negative links, where negative links can denote their foes and those they distrust, "unfriend"/"unfollow", and blocked users. In addition, systems from other domains such as chemistry, biology and ecology, physics, and political science have also been modeled as signed networks. However, the introduction of negative links in signed networks not only increases the complexity of the network representation, but also poses tremendous challenges for traditional unsigned network analysis. Hence, signed network analysis typically requires innovative and dedicated efforts to achieve state-of-the-art performance.

The contributions of this work can be summarized as follows:

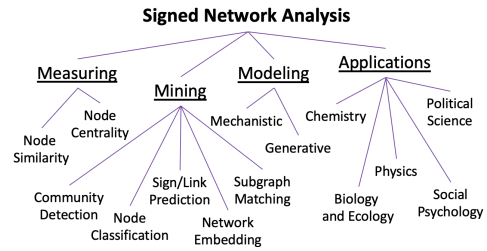

Figure 1: Overview of Signed Network Analysis Problems.

- We present an introduction into signed network analysis while concisely discussing foundational knowledge and providing a hierarchical categorization of existing problems.
- We provide a public dataset repository[1] that contains a comprehensive set of curated signed network datasets.
- Finally, we give future directions to guide those seeking to further advance the possibilities in signed network analysis.

2 BACKGROUND AND RELATED WORK

With roots in social psychology [12], signed network analysis has been gaining increased attention in recent years due to the ever-growing popularity of online social media. While in online social media some sites allow for the explicit representation of negative connections (e.g., blocking), in others we can also create implicit negative connections modeled from the users interactions.

A fundamental difference in signed networks as compared to unsigned networks is that negative and positive links have different properties. For example, positive links exhibit high local clustering, small diameters, and high transitivity, whereas the respective opposites are observed for negative links [10, 19]. In addition, the most applicable social theories to aid in our understanding and predictions have uniquely been discovered considering the interactions between positive and negative links together. For example, balance theory [3] can concisely be summarized with undirected signed triangles such as "A friend of a fiend is my friend" being balanced, while others like "An enemy of my friend is my enemy." are unbalanced; this was later further analyzed in signed bipartite networks using signed butterflies [6]. Status theory [17], is a directed signed network theory with the premise similar to that of a ladder, where positive links represent the giver expressing higher prestige over themselves and similarly negative links expressing a lower opinion of the receiver than the giver themselves. Hence, this theory can be similarly used to balance theory to add constraints such that resulting predictions should more globally adhere to the theory.

3 PROBLEMS AND METHODOLOGIES

Signed network analysis can broadly be categorized as consisting of efforts towards measuring, modeling, mining, and their applications with more detail given in Figure 1.

WSDM '20, February 3–7, 2020, Houston, TX, USA

ACM ISBN 978-1-4503-6822-3/20/02.
https://doi.org/10.1145/3336191.3372188

[1] http://www.github.com/DSE-MSU/awesome-signed-network-datasets

3.1 Signed Network Measuring

The most traditional methods in signed network measuring are centrality [13] and ranking, where methods have been developed seeking to handle the difference between "famous" and "infamous" nodes. Another development is in discovering the similarity between nodes in signed networks either coming from personalized ranking methods and dedicated relevance measurements [10].

3.2 Signed Network Modeling

Mechanistic network models are designed to construct synthetic networks that exhibit structural properties that are universally common for a given network type. In comparison, generative network models are a class of network models that take as input a given network and are able to construct synthetic networks that maintain similar properties to that of the input network. In [5] the first generative signed network model was introduced focused to better maintain balance theory (through the distribution of signed triangles) and simultaneously the link sign ratio. The previous signed network models had been mechanistic and required many parameters to be handcrafted to generate synthetic networks closely related to a desired existing network; however, having the benefit that they can generalize to construct diverse sets of signed networks.

3.3 Signed Network Mining and Applications

For signed network mining the most traditional tasks performed are community detection (i.e., clustering), link sign prediction, subgraph finding (e.g., cliques), link prediction, and node classification.

In community detection (i.e., clustering) the classical approaches were based on spectral methods [14]. More recently however, network embeddings methods have become state-of-the-art when being applied to numerous tasks (including clustering [22]) with graph neural network based approaches typically performing the best [8]. The other two primary tasks that researchers have focused their attention are node classification[1] and link sign prediction [16], which seek to predict the class of an unlabeled node or link, respectively. Efforts have also been made towards jointly predicting both link and interaction polarities [11]. In addition subgraph finding had been extended to signed networks, such as in [2].

Furthermore, signed networks have been used in various other domains outside of their traditional usage in analyzing online social networks with a lot of work (including the foundational) coming from social psychology. More specifically, in chemistry signed networks (i.e., Möbius graphs) are used with studying molecular systems [21]; in ecology for analyzing community structure [4]; in physics for modeling frustration in spin glasses [20]; and in political science for analyzing balance [6, 9] and predicting link signs [6] (e.g., predicting future congressional votes [7]).

4 FUTURE DIRECTIONS

In this section, we highlight future directions in each of the main categories. For measuring, the new direction likely to have the most impact is that of tie-strength prediction [10]. This task can either be performed with using only the network structure, or with the added assumption of having additional side information associated with the links and/or nodes. Essentially, this is the problem of taking an unweighted signed network and predicting an associated weight for each edge. For modeling, future directions can build on existing unsigned models such as Preferential Attachment and Kronecker Graphs [15], investigating models to maintain directed signed networks and status theory [17], and deep generative signed network models. Lastly, adversarial attack/defense have seen great attention recently, and so a dedicated effort (such as a rewiring attack [18]) we present as another open direction.

ACKNOWLEDGMENTS

The author would like to thank their Ph.D. advisor Dr. Jiliang Tang for the pleasure and privilege of collaborating with them. Tyler Derr and his advisor Dr. Jiliang Tang are supported by the National Science Foundation (NSF) under grant numbers IIS-1714741, IIS-1715940, IIS-1845081 and CNS-1815636, and a grant from Criteo Faculty Research Award.

REFERENCES

[1] Smriti Bhagat, Graham Cormode, and S Muthukrishnan. 2011. Node classification in social networks. In *Social network data analytics*. Springer, 115–148.
[2] Jose Cadena, Anil Kumar Vullikanti, and Charu C Aggarwal. 2016. On dense subgraphs in signed network streams. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE, 51–60.
[3] Dorwin Cartwright and Frank Harary. 1956. Structural balance: a generalization of Heider's theory. *Psychological review* 63, 5 (1956), 277.
[4] Jeffrey M Dambacher, Hiram W Li, and Philippe A Rossignol. 2002. Relevance of community structure in assessing indeterminacy of ecological predictions. *Ecology* 83, 5 (2002), 1372–1385.
[5] Tyler Derr, Charu Aggarwal, and Jiliang Tang. 2018. Signed Network Modeling Based on Structural Balance Theory. In *Proceedings of the 2018 ACM on Conference on Information and Knowledge Management*. ACM.
[6] Tyler Derr, Cassidy Johnson, Yi Chang, and Jiliang Tang. 2019. Balance in Signed Bipartite Networks. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. ACM, 1221–1230.
[7] Tyler Derr*, Hamid Karimi*, Aaron Brookhouse, and Jiliang Tang. 2019. Multi-Factor Congressional Vote Prediction. In *Advances in Social Networks Analysis and Mining (ASONAM), 2019 IEEE/ACM International Conference on*. IEEE.
[8] Tyler Derr, Yao Ma, and Jiliang Tang. 2018. Signed graph convolutional networks. In *2018 IEEE International Conference on Data Mining (ICDM)*. IEEE, 929–934.
[9] Tyler Derr and Jiliang Tang. 2018. Congressional Vote Analysis Using Signed Networks. In *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 1501–1502.
[10] Tyler Derr, Chenxing Wang, Suhang Wang, and Jiliang Tang. 2018. Relevance Measurements in Online Signed Social Networks. In *Proceedings of the 14th International Workshop on Mining and Learning with Graphs (MLG)*.
[11] Tyler Derr, Zhiwei Wang, and Jiliang Tang. 2018. Opinions power opinions: Joint link and interaction polarity predictions in signed networks. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, 363–366.
[12] Fritz Heider. 1946. Attitudes and cognitive organization. *The Journal of psychology* 21, 1 (1946), 107–112.
[13] Jérôme Kunegis, Andreas Lommatzsch, and Christian Bauckhage. 2009. The slashdot zoo: mining a social network with negative edges. In *Proceedings of the 18th international conference on World wide web*. ACM, 741–750.
[14] Jérôme Kunegis, Stephan Schmidt, Andreas Lommatzsch, Jürgen Lerner, Ernesto William De Luca, and Sahin Albayrak. 2010. Spectral analysis of signed graphs for clustering, prediction and visualization.. In *SDM*. SIAM, 559–559.
[15] Jure Leskovec, Deepayan Chakrabarti, Jon Kleinberg, Christos Faloutsos, and Zoubin Ghahramani. 2010. Kronecker graphs: An approach to modeling networks. *Journal of Machine Learning Research* 11, Feb (2010), 985–1042.
[16] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. Predicting positive and negative links in online social networks. In *Proceedings of the 19th international conference on World wide web*. ACM, 641–650.
[17] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. Signed networks in social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1361–1370.
[18] Yao Ma, Suhang Wang, Tyler Derr, Lingfei Wu, and Jiliang Tang. 2019. Attacking Graph Convolutional Networks via Rewiring. *arXiv preprint arXiv:1906.03750*.
[19] Jiliang Tang, Yi Chang, Charu Aggarwal, and Huan Liu. 2016. A survey of signed network mining in social media. *Comput. Surveys* 49, 3 (2016), 42.
[20] G Toulouse. 1987. Theory of the frustration effect in spin glasses: I. *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications* 9 (1987), 99.
[21] Nenad Trinajstic. 2018. *Chemical graph theory*. Routledge.
[22] Suhang Wang, Charu Aggarwal, Jiliang Tang, and Huan Liu. 2017. Attributed signed network embedding. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. ACM, 137–146.