Modularity Vitality for Bipartite Networks and Projections

Thomas Magelinski & Kathleen M. Carley
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
{tmagelin, kathleen.carley}@cs.cmu.edu.edu

Abstract

Modularity Vitality has recently been identified as an efficient and powerful way of identifying "community bridges" and "community hubs" in unipartite networks [6]. In this work, we expand this line of analysis to bipartite networks by deriving efficient calculations for Bipartite Modularity Vitality, Modularity Vitality on projected networks over the projected nodes, and Modularity Vitality on projected networks over the non-projected nodes. These measures of contribution to community structure aid in the identification of central nodes in bipartite networks, and in the interpretation of bipartite communities.

Community hubs contribute positively to community structure through intra-community links, while community bridges contribute negatively through inter-community links. Modularity Vitality measures node contributions to community structure by taking the difference of a network partition's modularity and that of the network partition if a node were to be removed. An efficient procedure for calculating modularity vitality has been derived. One application of this measure showed that community hubs in hashtag networks are more helpful in interpreting the hashtag community than other common measures, such as degree [5].

Enhanced methods of interpreting communities is of particular interest in the field of Social Cybersecurity [4, 10]. In this field, social media communities are analyzed to understand public discourse and how it may be manipulated. It is very common to study communities of users, hashtags, and URLs [7, 10, 11, 2]. Modularity vitality has been used to find users central to coordination networks on Twitter [8]. However, these networks are often bipartite, i.e. User-to-Hashtag networks, and thus require specialized analysis. At the same time, multiple definitions of Modularity have been defined for bipartite networks. Here, we study Barber's definition, and derive an efficient calculation for Bipartite Modularity Vitality [3]. We show that Bipartite Modularity Vitality can be used to identify community bridges and hubs in both node-sets simultaneously through demonstration on a User-to-URL Twitter Network. Community hubs within the URL nodeset are representative URLs, or URLs which can be used to best interpret the community they belong to. Community hubs within the user nodeset can be used similarly.

While pure bipartite analysis is a natural choice, projection is still extremely common in practice [10, 9]. Arthur thus defined a Projected Modularity, which accounts for the artifacts present in the unipartite graph from projection [1]. From this we derive two Projected Modularity Vitalities. First, the vitality over the resulting nodeset. That is, a measure of community contribution for each node present in the unipartite network. This measure can be used as Modularity Vitality has been previously, though it is now appropriate for projected networks. Second, the vitality over the projected nodeset. This vitality measures community contribution for each node in the nodeset that was projected over.

To demonstrate these vitalities, we project the User-to-URL network over the URLs, giving a User-to-User network, with weights indicating the number of times two users shared the same URL. Our first Projected Modularity Vitality identifies users which are hubs within their community, and are thus a reasonable starting point to interpret their group. The second Projected Modularity Vitality identifies URLs which are contributing the most to the community structure of our Users. Analysis of these

values together allow for similar insights to the bipartite analysis, while utilizing unipartite methods of community discovery.

Acknowledgments

This work was supported in part by the Office of Naval Research (ONR) Award N000141512797 Minerva award for Dynamic Statistical Network Informatics, and the Center for Computational Analysis of Social and Organization Systems (CASOS). Thomas Magelinski was also supported by the IDeaS Center as a Knight Fellow. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the ONR.

References

[1] R. Arthur. Modularity and projection of bipartite networks. *Physica A: Statistical Mechanics and its Applications*, page 124341, 2020.

[2] M. Babcock, D. M. Beskow, and K. M. Carley. Beaten up on twitter? exploring fake news and satirical responses during the black panther movie event. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 97–103. Springer, 2018.

[3] M. J. Barber. Modularity and community detection in bipartite networks. *Physical Review E*, 76(6):066102, 2007.

[4] K. M. Carley, G. Cervone, N. Agarwal, and H. Liu. Social cyber-security. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 389–394. Springer, 2018.

[5] T. Magelinski, M. Bartulovic, and K. M. Carley. Canadian federal election and hashtags that do not belong. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 161–170. Springer, 2020.

[6] T. Magelinski, M. Bartulovic, and K. M. Carley. Measuring node contribution to community structure with modularity vitality. *IEEE Transactions on Network Science and Engineering*, 8(1):707–723, 2021.

[7] T. Magelinski, D. M. Beskow, and K. M. Carley. Graph-hist: Graph classification from latent feature histograms with application to bot detection. In *AAAI*, pages 5134–5141, 2020.

[8] T. Magelinski, L. H. X. Ng, and K. M. Carley. A synchronized action framework for responsible detection of coordination on social media. *arXiv preprint arXiv:2105.07454*, 2021.

[9] T. Magelinski, Z. K. Stine, T. Marcoux, N. Agarwal, and K. M. Carley. Artifacts of crisis: Textual analysis of euromaidan. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 329–339. Springer, 2020.

[10] J. Uyheng, T. Magelinski, R. Villa-Cox, C. Sowa, and K. M. Carley. Interoperable pipelines for social cyber-security: assessing twitter information operations during nato trident juncture 2018. *Computational and Mathematical Organization Theory*, pages 1–19, 2019.

[11] L. Yang, T. Sun, M. Zhang, and Q. Mei. We know what@ you# tag: does the dual role affect hashtag adoption? In *Proceedings of the 21st international conference on World Wide Web*, pages 261–270, 2012.