Unsupervised Graph Network Discrimination

In recent years, network analysis techniques have evolved(Chen et al., 2019; Williamson & Tec, 2019) in proportion to the rapid growth of real world networks. Much research has been done on networks such as social networks (Bródka et al., 2011; Culotta & Cutler, 2016), communication networks (Mccallum, 2007), citation networks etc. As time progresses, these networks become bigger and more complex, consequently holding vast amounts of interesting information which can be used for various purposes. Network analysis techniques include, but are not limited to, using random graph models to capture or derive the properties of real world networks (Williamson & Tec, 2019), subgraph isomorphism (Cordella et al., 2004), graph simulation etc. A significant portion of the research conducted on networks has been on centralities (Newman, 2010) to deduce, “which” are the more important nodes in the graph i.e.. central nodes or nodes that propagate most information, and how traffic flows through them. In turn, calculated centralities of various network graphs have been used in many other research fields as well. For example, analyzing social networks, such as tweet classification (Hussain & Islam, 2016), detecting political discussion practices (Miller et al., 2015) and many other such applications (Cohn et al., 2019; Rossman et al., 2010; Yang & Liu, 2008); has involved the generation of multiple centralities. Centralities have also been used in analyzing road traffic networks, such as finding road network patterns (Zhang et al., 2011), tourism management (Lee et al., 2013) etc.. Another important application of centrality measurements is in analyzing biological networks, as can be seen in (Bell et al., 1999; Joyce et al., 2010; Narayanan, 2005; Park & Kim, 2009).

Centralities such as Degree, Closeness, Betweenness, Crossclique, Pagerank, etc. have all been proposed and developed (Crucitti et al., 2006) over the years to answer “Which node is the most influential?” in various different applications (Landherr et al., 2010). “Significance” or “importance” of nodes varies from one context to another. For example, in some cases, it is essential to identify which node(s) propagate more information locally or globally in a graph(Newman, 2010), whereas in other cases, detecting the central node(s) might be of more value (Chiu et al., 2016). Centralities can, therefore, be seen as features/characteristics of a graph; thus, it is possible to discriminate between networks based on their centrality values (Wang & Krim, 2012) //*citation added manually.*

In this paper, our main concern is “What information can we extract from a given, unknown graph of a network?” Calculating various centralities from a graph is possible, but this raises the question of what can be concluded from this information? Is it possible to discriminate between types of networks based on only the centrality values of their nodes? Or, Can unsupervised learning techniques applied to unknown graphs produce meaningful context about the network in question? Being able to find meaningful insights about an unidentified network has many practical uses, such as deducing, whether it is a criminal network or not, or if the network was part of a successful political campaign, or whether the network is an ego network of a socially influential person.

Bell, D. C., Atkinson, J. S., & Carlson, J. W. (1999). Centrality measures for disease transmission networks. *Social Networks*, *21*(1), 1–21. https://doi.org/10.1016/S0378-8733(98)00010-0

Bródka, P., Skibicki, K., Kazienko, P., & Musiał, K. (2011). A degree centrality in multi-layered social network. *Proceedings of the 2011 International Conference on Computational Aspects of Social Networks, CASoN’11*, 237–242. https://doi.org/10.1109/CASON.2011.6085951

Chen, H., Yin, H., Chen, T., Viet, Q., Nguyen, H., & Xue, W. P. (2019). Exploiting Centrality Information with Graph Convolutions for Network Representation Learning. *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, 590–601. https://doi.org/10.1109/ICDE.2019.00059

Chiu, C. C., Balkunid, P., & Weinberg, F. (2016). When managers become leaders : The role of manager network centralities , social power , and followers ’ perception of leadership. *The Leadership Quarterly*. https://doi.org/10.1016/j.leaqua.2016.05.004

Cohn, A. M., Amato, M. S., Zhao, K., Wang, X., Cha, S., Pearson, J. L., Papandonatos, G. D., & Graham, A. L. (2019). Discussions of Alcohol Use in an Online Social Network for Smoking Cessation: Analysis of Topics, Sentiment, and Social Network Centrality. *Alcoholism: Clinical and Experimental Research*, *43*(1), 108–114. https://doi.org/10.1111/acer.13906

Cordella, L. P., Foggia, P., Sansone, C., & Vento, M. (2004). A (sub)graph isomorphism algorithm for matching large graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *26*(10), 1367–1372. https://doi.org/10.1109/TPAMI.2004.75

Crucitti, P., Latora, V., & Porta, S. (2006). Centrality in networks of urban streets. *Chaos*, *16*(1). https://doi.org/10.1063/1.2150162

Culotta, A., & Cutler, J. (2016). *Predicting Twitter User Demographics using Distant Supervision from Website Traffic Data*. *55*, 389–408.

Hussain, J., & Islam, M. A. (2016). Evaluation of graph centrality measures for tweet classification. *2016 International Conference on Computing, Electronic and Electrical Engineering, ICE Cube 2016 - Proceedings*, 126–137. https://doi.org/10.1109/ICECUBE.2016.7495209

Joyce, K. E., Laurienti, P. J., Burdette, J. H., & Hayasaka, S. (2010). A new measure of centrality for brain networks. *PLoS ONE*, *5*(8). https://doi.org/10.1371/journal.pone.0012200

Landherr, A., Friedl, B., & Heidemann, J. (2010). A Critical Review of Centrality Measures in Social Networks. *Business & Information Systems Engineering*, *2*(6), 371–385. https://doi.org/10.1007/s12599-010-0127-3

Lee, S. H., Choi, J. Y., Yoo, S. H., & Oh, Y. G. (2013). Evaluating spatial centrality for integrated tourism management in rural areas using GIS and network analysis. *Tourism Management*, *34*, 14–24. https://doi.org/10.1016/j.tourman.2012.03.005

Mccallum, A. (2007). *Topic and Role Discovery in Social Networks with Experiments on Enron and Academic Email*. *30*, 249–272.

Miller, P. R., Bobkowski, P. S., Maliniak, D., & Rapoport, R. B. (2015). *Talking Politics on Facebook : Network Centrality and Political Discussion Practices in Social Media*. https://doi.org/10.1177/1065912915580135

Narayanan, S. (2005). *The Betweenness Centrality Of Biological Networks A Study of Betweenness Centrality*.

Newman, M. (2010). Networks: An Introduction. In *Networks: An Introduction*. https://doi.org/10.1093/acprof:oso/9780199206650.001.0001

Park, K., & Kim, D. (2009). *Localized network centrality and essentiality in the yeast – protein interaction network*. 5143–5154. https://doi.org/10.1002/pmic.200900357

Rossman, G., Esparza, N., & Bonacich, P. (2010). I’d like to thank the academy, team spillovers, and network centrality. *American Sociological Review*, *75*(1), 31–51. https://doi.org/10.1177/0003122409359164

Williamson, S. A., & Tec, M. (2019). Random clique covers for graphs with local density and global sparsity. *35th Conference on Uncertainty in Artificial Intelligence, UAI 2019*.

Yang, B., & Liu, J. (2008). Discovering global network communities based on local centralities. *ACM Transactions on the Web*, *2*(1). https://doi.org/10.1145/1326561.1326570

Zhang, Y., Wang, X., Zeng, P., & Chen, X. (2011). Centrality characteristics of road network patterns of traffic analysis zones. *Transportation Research Record*, *2256*, 16–24. https://doi.org/10.3141/2256-03