# Chapter 7

# CONCLUSION AND FUTURE SCOPE

The thesis investigated extending network analytical techniques to process non-euclidean data, such as networks (graphs). This domain had received comparatively low levels of attention. There was no straightforward method to analyze network data as learning models were designed for simple euclidean data or grids. Additionally, network-structured data had issues related to the lack of independent data-points, computationally costly calculations for graph statistics, in-applicability of parallel or distributed algorithms and the curse of dimensionality.

To leverage network data's potential, there was a need for latent space representation learning techniques that could accommodate high dimensionality, and heterogeneity present in network data in the representations. These representations could then be used to improve the efficiency and performance of downstream network-based applications such as node clustering, node classification, so on.

To address such challenges, this dissertation's focus was on Latent Space Representations (L.S.R.) of networks, i.e., learning low dimension vector representations of network data. L.S.R. techniques have a data-driven mechanism for learning vector embedding of network data structures. The objectives are mentioned below:

- Survey state of the art Latent Space Representation (L.S.R.) techniques in the

literature.

- Perform extensive experiments on network data-sets from the Stanford Network Analysis Project (SNAP) for understanding the structure and behaviour of systems.
- Develop ensemble methods for L.S.R. and perform experiments on network data to compare their performance vis-a-vis existing techniques.
- Propose a convolutional implementation of G.C.N. auto-encoder architectures for L.S.R. to resolve the issues of standard graph auto-encoder architectures such as Graph Auto-Encoder (G.A.E.) and Variational Graph Auto-Encoder (V.A.E.).
- Increasing the depth of the graph convolutional network architecture for L.S.R. by use of residual blocks to address the problem of a small receptive field of G.C.N.

## 7.1 Thesis Summary with Overall Conclusion

In Chapter 2, a survey of state of the art Latent Space Representation (L.S.R.) techniques in the literature were presented. In the survey, three categories of grouping, Latent space representation techniques for networks were examined in detail. These were Probabilistic models, Statistical models, and Network representation learning models. Probabilistic models had limited flexibility and hence resulted in poor fits to data. Even though probabilistic models were "generative," they did not generate networks that shared many properties with the specific network they were fit to. Statistical models such as S.B.M. and latent variable models were regarded as the "The most promising class of statistical models for expressing networks into low-dimensional geometries". The three classes of N.R.L. techniques discussed in the literature were Adjacency preserving methods [25, 26, 27, 31, 32, 85, 86, 87, 88, 89], Multi-hop distance preserving methods and Random walk occurrence preserving methods [90, 91, 92, 93, 94, 95, 96, 97, 98, 99]. Apart from these, deep learning-based N.R.L. techniques were surveyed in the literature. A drawback of deep learning architectures was the complexity involved in building deep graph neural network models. The backbone of deep graph neural networks was

the graph convolution operation. Graph convolution relied on a fully connected layer. This layer increased the number of parameters of the model and slowed the gradient descent process. It was also observed that G.C.N. based architecture was shallow (two layers), which affected their receptive field. Thus, the survey provided an overview of research gaps in the latent space representation techniques. It also gave a technical foundation for the remaining parts of this dissertation.

In Chapter 3, Network science was applied to understand the trends in domains of scientific literature such as social networking websites (Twitter, Google+), blogs (Blog.com), photo-sharing websites (Flickr), protein-protein interactions (Protein-Net), citation networks (CORA and CiteSeer), transportation networks (High-Net), sexual contact network (Grey-Net), trade network (Trade-Net) and bill co-sponsorship network (Bill-Net). Trends identified using network analysis were the rate of information diffusion, the dominance of individual entities, the number of social contacts, and community structure. These trends would not be easily identifiable using relational databases. Although networks are present in every domain, their analysis revealed that the structural characteristics shared by them are similar, i.e., low average path, low diameter, so on. It was also noted that networks capture the behaviour of the entities present in them. Using the concepts of graph theory, it was possible to make a statistically valid analysis of these systems and provide insights into their growth. Networks across different domains saw low edge density and the presence of inequality. Networks such as Twt-Net, Gplus-Net, Flickr-Net, Wiki-Net, Blog-Net, Grey-Net, and Bill-Net are particular types of networks called "social networks." Social networks represent the sum of all professional, friendship, or family ties of the actors involved in them. Social networks were observed to have higher edge density and average degree compared to other networks. They also had high transitivity, low diameter, and negative assortativity. In the experimental setting, it was found that the two L.S.R. models for conducting the analysis (Stochastic Block model and Latent variable model) had high computational complexity, ignored attribute information, and were only suitable for data-sets with nodes in the range of $10^1 - 3$. Therefore, it was necessary to investigate models that could scale to more extensive networks ($10^3$+).

In Chapter 4, N.R.L. methods were used for network analysis as the networks in this

chapter were large, with nodes in the range of $10^3$. The frameworks in the literature were based on the optimization of a single objective function. This could lead to node embedding that is favourable for specific applications but unsuitable for others [49, 50, 51]. It was speculated that a single model was unlikely to capture the entire underlying structure of the data. Hence, to achieve optimal representations, integrating multiple models might improve the accuracy of representation learning significantly. This gap served as the principal motivation for this study. Two ensemble models were proposed viz. Weighted-Average and Mean, in which representation learning of base models (Adjacency preserving methods, Multi-hop distance preserving methods, and Random walk occurrence preserving methods) were combined by consensus. In the experimental setting, it was seen that in networks where the Gini index and transitivity were low, all three base models had near equal weight. In networks with a high Gini index, the adjacency preserving model had higher weightage in the ensemble. The networks with high transitivity gave more weightage to the Multi-hop distance preserving model, and networks with high diameter gave more weightage to Random walk occurrence preserving models. A key drawback of ensemble techniques was the computation complexity. The ensemble techniques have a computation complexity of $O(|E|d) + O(|V|d) + O(|E|d)$. This is a summation of the computation complexities of the base models. Constant time is required for performing the combination by consensus operation, i.e., weighted-average or mean. The proposed weighted-average or mean ensemble N.R.L. techniques required that all nodes in the network should be present during the training of the encoder. Hence, these approaches were inherently transductive and would not generalize to unseen nodes (non-inductive). The proposed methods discarded attribute data associated with the network nodes. Hence, “deep encoder" based techniques could be investigated for situations where all network nodes are not available at time of training, and also network nodes are associated with attribute data.

In Chapter , Graph Convolutional Network-based models were investigated. These models considered both topographical and attributed information in the vector representations. However, the drawback of the G.C.N. was the presence of a fully connected layer. The fully connected layer led to an increase in parameters as the input graph size increased. This led to slow convergence of gradient descent and increased the

time per iteration. Hence, the inquiry proposed to modify the existing Graph Auto-Encoder (G.A.E.) and Variational Graph Auto-Encoder (V.A.E.) by replacing the fully connected layers with 1-D convolutional ones. The proposed models were named as Graph Auto-Encoder with Convolutional layers (G.A.E.-F.C.) and Variational Graph Auto-Encoder with Convolutional layers (V.A.E.-F.C.). Experiments were performed to compare G.A.E. and V.A.E. with G.A.E.-F.C. and V.A.E.-F.C. The parameters for comparison were the parameters of the architectures, average time per iteration (in secs), transitivity and clustering quality measures such as Separation index, Widest within-cluster gap, Average silhouette width, Average distance between clusters and Dunn index.

In the experimental setting, the performance of G.A.E.-F.C. compared to G.A.E. and V.A.E. for the parameter transitivity showed an improvement of $3-9\%$ on Cora-Net, Cite-Net, Blog-Net, and Flickr-Net. At the same time, it was lower by $7-10\%$ on Wiki-Net and Protein-Net. Similarly, in V.A.E.-F.C., it showed an improvement of $3-7\%$ on Cora-Net, Cite-Net, Blog-Net, and Flickr-Net; however, was lower by $2-5\%$ on Protein and Wiki-Net. The performance of G.A.E.-F.C. compared to G.A.E. and V.A.E. for the parameter Dunn index showed an improvement of $3-4.6\%$ on Cora-Net and Protein-Net but was lower by $2-4\%$ on Wiki-Net, Cite-Net, Blog-Net, and Flickr-Net. Similarly, in V.A.E.-F.C., it showed an improvement of $1.8-3\%$ on Wiki-Net, Cora-Net, Cite-Net, Blog-Net, and Protein-Net but was lower by $6.6\%$ on Flickr-Net. The performance of G.A.E.-F.C. compared to G.A.E. and V.A.E. for the parameter average distance between clusters showed an improvement of $6.4-14\%$ on Wiki-Net, Cora-Net and Protein-Net but was lower by $6-23\%$ on Cite-Net, Blog-Net, and Flickr-Net. Similarly, in V.A.E.-F.C. it showed an improvement of $4.8-35\%$ on Wiki-Net, Cora-Net, Blog-Net, and Protein-Net but was lower by $15-16\%$ on Cite-Net and Flickr-Net. The performance of G.A.E.-F.C. compared to G.A.E. and V.A.E. for the parameter average silhouette width was improved by $2.3-4.6\%$ on Cora-Net, Flickr-Net, and Protein-Net but was lower by $1.2-4.9\%$ on Wiki-Net, Cite-Net, and Blog-Net. Similarly, in V.A.E.-F.C., it showed an improvement of $2.3-6\%$ on Cora-Net, Blog-Net, Flickr-Net, and Protein-Net but was lower by $1.5-2.8\%$ on Wiki-Net and Cite-Net. The performance of G.A.E.-F.C. compared to G.A.E. and V.A.E. for the parameter widest within-cluster gap showed an improvement of $6-26\%$ on Wiki-Net, Cite-Net, Blog-Net, Flickr-Net,

and Protein but was lower by 0.3% on Cora-Net. Similarly, in V.A.E.-F.C., it showed an improvement of $3-55\%$ on Wiki-Net, Blog-Net, Flickr-Net, and Protein-Net but lower by $2-5\%$ on Cora-Net and Cite-Net. The performance of G.A.E.-F.C. compared to G.A.E. and V.A.E. for the parameter separation index showed an improvement of 22% on Protein-Net however, was lower by $0.15-30\%$ on Wiki-Net, Cora-Net, Cite-Net, Blog-Net, and Flickr-Net. Similarly, in V.A.E.-F.C., it showed an improvement of $4-38\%$ on Wiki-Net, Cite-Net, Blog-Net, Flickr-Net, and Protein-Net but was lower by 32% on Cora-Net. The number of parameters in the convolutional implementations was $\frac{1}{100}-\frac{1}{35}$ that of G.A.E. and V.A.E. The average time per iteration (in secs) was $\frac{1}{156}-\frac{1}{14}$ that of G.A.E. and V.A.E. Both G.A.E.-F.C. and V.A.E.-F.C. were two layers deep and therefore had small receptive fields.

In Chapter , an attempt was made to address shallow G.C.N. architectures by the inclusion of a residual block to increase their depth. Experiments were performed to compare a standard G.C.N. with a G.C.N. having residual blocks. The parameters for comparison were the architectures' parameters, execution time (in secs), transitivity and clustering quality measures such as Separation index, Widest within-cluster gap, Average silhouette width, Average distance between clusters and Dunn index.

In the experimental setting, the performance of Residual G.C.N. compared to a standard G.C.N. for the parameter transitivity showed an improvement of $12-42\%$ on the datasets. For the parameter Dunn index, the performance of Residual G.C.N. compared to a standard G.C.N. was improved by $0.12-11\%$ on Protein-Net, Flickr-Net, and Blog-Net but was lower by $5-12\%$ on Wiki-Net, Cora-Net, and Cite-Net. For the parameter average distance between clusters, the performance of Residual G.C.N. compared to a standard G.C.N. showed an improvement of $2-5\%$ on Wiki-Net, Cora-Net, Blog-Net, and Protein-Net but was lower by $11-68\%$ on Cite-Net and Flickr-Net. For the parameter average silhouette width, the performance of Residual G.C.N. compared to a standard G.C.N. showed an improvement of $1-7\%$ on Cora-Net, Cite-Net, and Blog-Net but was lower by 4% on Wiki-Net and Flickr-Net. For the parameter widest within-cluster gap, the performance of Residual G.C.N. compared to a standard G.C.N. showed an improvement of $5-6\%$ on Wiki-Net, Cora-Net, Blog-Net, and Flickr-Net but was lower by $8-30\%$ on Cite-Net and Protein-Net. For the parameter separation

index, the performance of Residual G.C.N. compared to a standard G.C.N. showed an improvement of $7 - 8\%$ on Wiki-Net and Flickr-Net but was lower by $1 - 11\%$ on Cora-Net, Cite-Net, Blog-Net, and Protein-Net. The time for execution of Residual G.C.N. compared to a standard G.C.N. was higher by $22 - 55\%$ on the data-sets as the parameters were more by $6 - 43\%$.

## 7.2 Future Scope

This dissertation explored several vital directions for network representation learning, but many questions remain unanswered, and many future directions remain open to exploring.

### 7.2.1 Receptive field

The receptive field refers to the region in the input space that a particular node is affected by. As graphs do not have fixed structures like grids, the neighbourhood of nodes differs. The receptive field of the auto-encoder and ensemble models proposed in this investigation can be made more representative using sampling strategies.

### 7.2.2 Scalability

The proposed work does not scale to graphs with millions of nodes. The main reason is excessive parameters if multiple hidden layers are used, leading to backpropagation's high complexity. The efficiency of the model can be improved using sub-graph training techniques.

### 7.2.3 Dynamics

Models proposed in this thesis focus on static graphs, i.e., graph structures are assumed to be fixed. However, these models need to be extended to scenarios where network topology changes at any time.

# Publications

## International Journals

1. **Nerurkar**, **P**., Chandane, M. and Bhirud, S., 2019. “Understanding attribute and social circle correlation in social networks"; *Turkish Journal of Electrical Engineering & Computer Sciences*, 27(2), pp.1228-1242.

2. **Nerurkar**, **P**., Chandane, M. and Bhirud, S., 2019. "Modeling influence on a Social Network using Interaction Characteristics"; *International Journal of Computer & Mathematical Sciences*, 6(8), pp.152-160.

3. **Nerurkar**, **P**., Chandane, M. and Bhirud, S., 2019. "Measurement of netnetwork-basedd random meetings in social networks"; *Turkish Journal of Electrical Engineering & Computer Sciences*, 27(2), pp.765-779.

4. **Nerurkar**, **P**., Chandane, M. and Bhirud, S., 2019. "Exploring convolutional auto-encoders for representation learning on networks"; *Computer Science Journal*, Accepted.

5. **Nerurkar**, **P**., Chandane, M. and Bhirud, S., 2019. "Empirical Analysis of Synthetic and Real Networks"; *International Journal of Information Technology*, Accepted. Springer.

6. **Nerurkar**, **P**., Chandane, M. and Bhirud, S., 2019. "Understanding structure and behaviour of systems: A network perspective"; *International Journal of Information Technology*, Accepted. Springer.

## National Journals

1. **Nerurkar**, **P**., Chandane, M. and Bhirud, S., 2019. “Investigations on Residual Auto-encoders for unsupervised network representation learning"; *Journal of The Institution of Engineers (India): Series B*, Communicated. Springer.

## International Conferences

1. **Nerurkar**, **P**., Shirke, A., Chandane, M. and Bhirud, S., 2018. “Empirical Analysis of Data Clustering Algorithms"; *Procedia Computer Science*, 125, (pp. 770-779). Elsevier.

2. **Nerurkar**, **P**., Chandane, M. and Bhirud, S., 2019. "A Comparative Analysis of Community Detection Algorithms on Social Networks"; In *Computational Intelligence: Theories, Applications and Future Directions*-Volume I (pp. 287-298). Springer.

3. **Nerurkar**, **P**., Shirke, A., Chandane, M. and Bhirud, S., 2018. “A novel heuristic for evolutionary clustering"; *Procedia Computer Science*, 125, (pp. 780-789). Elsevier.

4. **Nerurkar**, **P**., Chandane, M., and Bhirud, S., 2019. "Representation learning for social networks using Homophily based Latent Space Model"; In *Proceedings of the International Conference on Omni-Layer Intelligent Systems* (pp. 38-43). ACM.

5. **Nerurkar**, **P**., Chandane, M. and Bhirud, S., 2019. “Community Detection Using Node Attributes: A Non-negative Matrix Factorization Approach"; In *Computational Intelligence: Theories, Applications and Future Directions-Volume I* (pp. 275-285). Springer.

# Bibliography


[1] Albert-László Barabási et al. *Network science*. 2016. xvi, 1, 3, 18

[2] Denny M. Social network analysis. *Institute for Social Science Research, University of Massachusetts, Amherst*, 2014. 1, 8, 22

[3] Denny M. Intermediate social network theory. *Institute for Social Science Research, University of Massachusetts, Amherst*, 2015. 1

[4] Matthew O Jackson. *Social and economic networks*. 2010. 1, 16, 17

[5] Alex Fout, Jonathon Byrd, Basir Shariat, and Asa Ben-Hur. Protein interface prediction using graph convolutional networks. In *Advances in Neural Information Processing Systems*, pages 6530–6539, 2017. 1

[6] Qirong Ho, Junming Yin, and Eric P Xing. Latent space inference of internet-scale networks. *The Journal of Machine Learning Research*, 17(1):2756–2796, 2016. 1

[7] Henrique F de Arruda, Filipi N Silva, Cesar H Comin, Diego R Amancio, and Luciano da F Costa. Connecting network science and information theory. *Physica A: Statistical Mechanics and its Applications*, 515:641–648, 2019. 1

[8] Alvaro Sanchez-Gonzalez, Nicolas Heess, Jost Tobias Springenberg, Josh Merel, Martin Riedmiller, Raia Hadsell, and Peter Battaglia. Graph networks as learnable physics engines for inference and control. *arXiv preprint arXiv:1806.01242*, 2018. 1

[9] Peter Battaglia, Razvan Pascanu, Matthew Lai, Danilo Jimenez Rezende, et al. Interaction networks for learning about objects, relations and physics. In *Advances in neural information processing systems*, pages 4502–4510, 2016. 1

[10] Leonardo FS Scabini, Rayner HM Condori, Wesley N Gonçalves, and Odemir M Bruno. Multilayer complex network descriptors for color-texture characterization. *Information Sciences*, 2019. 1

[11] Pranav Nerurkar, Madhav Chandane, and Sunil Bhirud. A comparative analysis of community detection algorithms on social networks. In *Computational Intelligence: Theories, Applications and Future Directions-Volume I*, pages 287–298. 2019. 1

[12] Pranav Nerurkar, Archana Shirke, Madhav Chandane, and Sunil Bhirud. A novel heuristic for evolutionary clustering. *Procedia Computer Science*, 125:780–789, 2018. 1

[13] Haochen Chen, Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. A tutorial on network embeddings. *arXiv preprint arXiv:1808.02590*, 2018. 1

[14] Rex Ying, Dylan Bourgeois, Jiaxuan You, Marinka Zitnik, and Jure Leskovec. Gnn explainer: A tool for post-hoc explanation of graph neural networks. *arXiv preprint arXiv:1903.03894*, 2019. 1

[15] István A Kovács and Katja Luck. Network-based prediction of protein interactions. *Nature communications*, 10:1240, 2019. 1

[16] Roberto Interdonato, Martin Atzmueller, Sabrina Gaito, Rushed Kanawati, Christine Largeron, and Alessandra Sala. Feature-rich networks: going beyond complex network topologies. *Applied Network Science*, 4:4, 2019. 1

[17] Igor Linkov and Benjamin D Trump. Applications of network science and systems thinking. In *The Science and Practice of Resilience*, pages 167–179. 2019. 1

[18] Vicenc Gomez, Andreas Kaltenbrunner, and Vicente Lopez. Statistical analysis of the social network and discussion threads in slashdot. In *Proceedings of the 17th international conference on World Wide Web*, page 645—654, 2008. 2, 3

[19] Santo Fortunato and Darko Hric. Community detection in networks: A user guide. *Physics Reports*, 659:1–44, 2016. xvi, 2, 3, 4, 5, 6, 18

[20] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, 1999. 2, 3

[21] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30:107–117, 1998. 2

[22] Mary McGlohon, Leman Akoglu, and Christos Faloutsos. Statistical properties of social networks. In *Social network data analytics*, pages 17–42. 2011. 3

[23] Shi Y, Gui H, Zhu Q, Kaplan L, and Han J. Aspem: Embedding learning by aspects in heterogeneous information networks. In *Proceedings of the 2018 SIAM International Conference on Data Mining*, pages 144–152, 2018. 3, 7, 28, 33

[24] Perozzi B, Kulkarni V, Chen H, and Skiena S. Don't walk, skip!: Online learning of multi-scale network embeddings. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, pages 258–265, 2017. 7, 31, 33

[25] Tang J, Qu M, Wang M, Zhang M, Yan J, and Mei Q. Line: Large-scale information network embedding. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1067–1077. International World Wide Web Conferences Steering Committee, 2015. 7, 23, 27, 33, 73, 79, 124

[26] Huang X, Li J, and Hu X. Label informed attributed network embedding. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 731–739, 2017. 7, 23, 29, 33, 73, 124

[27] Huang X, Li J, and Hu X. Accelerated attributed network embedding. In *Proceedings of the 2017 SIAM International Conference on Data Mining*, pages 633–641, 2017. 7, 23, 29, 33, 73, 124

[28] Cui P, Wang X, Pei J, and Zhu W. A survey on network embedding. *IEEE T Knowl Data En*, 2018. 8, 79

[29] Zitnik M, Agrawal M, and Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. *arXiv preprint arXiv:1802.00543*, 2018. 8, 23, 35

[30] Zitnik M and Leskovec J. Predicting multicellular function through multi-layer tissue networks. *Bioinformatics*, 33:190–198, 2017. 8, 23, 35

[31] Liao L, He X, Zhang H, and Chua TS. Attributed social network embedding. *arXiv preprint arXiv:1705.04969*, 2017. 8, 23, 73, 124

[32] Bandyopadhyay S, Kara H, Kannan A, and Murty MN. Fscnmf: Fusing structure and content via non-negative matrix factorization for embedding information networks. *arXiv preprint arXiv:1804.05313*, 2018. 8, 23, 27, 33, 73, 74, 124

[33] Hamilton W, Ying Z, and Leskovec J. Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems*, pages 1024–1034, 2017. 9, 32, 34, 43, 111

[34] Butts CT Goodreau SM Pavel NK Bender-deMoll S Morris M Handcock MS, Hunter DR. *statnet: Software Tools for the Statistical Analysis of Network Data*. The Statnet Project, 2016. 9, 21

[35] Zhang M, Cui Z, Neumann M, and Chen Y. An end-to-end deep learning architecture for graph classification. In *Proceedings of AAAI Conference on Artificial Inteligence*, 2018. 9, 11, 23, 33, 37, 109, 110

[36] Kipf TN and Welling M. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016. 9, 23, 33, 34, 88, 111, 115

[37] Chen J, Ma T, and Xiao C. Fastgcn: fast learning with graph convolutional networks via importance sampling. *arXiv preprint arXiv:1801.10247*, 2018. 9, 37, 110

[38] Wu L, Fisch A, Chopra S, Adams K, Bordes A, and Weston J. Starspace: Embed all the things! *arXiv preprint arXiv:1709.03856*, 2017. 9

[39] Donnat C, Zitnik M, Hallac D, and Leskovec J. Spectral graph wavelets for structural role similarity in networks. *arXiv preprint arXiv:1710.10321*, 2017. 9

[40] William L Hamilton, Rex Ying, and Jure Leskovec. Representation learning on graphs: Methods and applications. *arXiv preprint arXiv:1709.05584*, 2017. 9, 33, 43, 91, 113, 114

[41] Goodreau SM. Advances in exponential random graph (p*) models applied to a large social network. *Soc networks*, 29:231–248, 2007. 10, 11, 40, 43

[42] Hoff PD, Raftery AE, and Handcock MS. Latent space approaches to social network analysis. *J Am Stat Assoc*, 97:1090–1098, 2002. 10, 40, 43

[43] Yun Zhang, Kehui Chen, Allan Sampson, Kai Hwang, and Beatriz Luna. Node features adjusted stochastic block model. *Journal of Computational and Graphical Statistics*, pages 1–22, 2019. 11, 20, 40, 43

[44] Soumyasundar Pal and Mark Coates. Scalable mcmc in degree corrected stochastic block model. In *Proc. Intl. Conf. Acoust., Speech and Signal Proc*, 2019. 11, 40, 43

[45] Xiaoyan Lu and Boleslaw K Szymanski. Regularized stochastic block model for robust community detection in complex networks. *arXiv preprint arXiv:1903.11751*, 2019. 11

[46] Marina S Paez, Arash A Amini, and Lizhen Lin. Hierarchical stochastic block model for community detection in multiplex networks. *arXiv preprint arXiv:1904.05330*, 2019. 11

[47] Hannu Reittu, Ilkka Norros, Tomi Räty, Marianna Bolla, and Fülöp Bazsó. Regular decomposition of large graphs: Foundation of a sampling approach to stochastic block model fitting. *Data Science and Engineering*, pages 1–17, 2019. 11

[48] Marianna Pensky, Teng Zhang, et al. Spectral clustering in the dynamic stochastic block model. *Electronic Journal of Statistics*, 13:678–709, 2019. 11

[49] Balasubramanian M and Schwartz EL. The isomap algorithm and topological stability. *Science*, 295:7–7, 2002. 11, 27, 33, 73, 126

[50] Roweis ST and Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, 2000. 11, 27, 33, 73, 126

[51] Goyal P and Ferrara E. Graph embedding techniques, applications, and performance: A survey. *Knowl-Based Syst*, 151:78–94, 2018. 11, 73, 126

[52] Ryan A Rossi, Rong Zhou, and Nesreen K Ahmed. Deep inductive network representation learning. In *Companion of the The Web Conference 2018 on The Web Conference 2018*, pages 953–960, 2018. 11, 109

[53] Saket Gurukar and Priyesh Vijayan. Network representation learning: Consolidation and renewed bearing. *arXiv preprint arXiv:1905.00987*, 2019. 15

[54] Eitan Muller and Renana Peres. The effect of social networks structure on innovation performance: A review and directions for research. *International Journal of Research in Marketing*, 36:3–19, 2019. 15

[55] Hao Yin, Austin R Benson, and Jure Leskovec. The local closure coefficient: A new perspective on network clustering. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pages 303–311, 2019. 15

[56] Tiandong Wang and Sidney I Resnick. Consistency of hill estimators in a linear preferential attachment model. *Extremes*, 22:1–28, 2019. 15

[57] Carter T Butts. A dynamic process interpretation of the sparse ergm reference model. *The Journal of Mathematical Sociology*, 43:40–57, 2019. 16

[58] Cristian Candia, C Jara-Figueroa, Carlos Rodriguez-Sickert, Albert-László Barabási, and César A Hidalgo. The universal decay of collective memory and attention. *Nature Human Behaviour*, 3:82, 2019. 16

[59] Adam Sealfon and Jonathan Ullman. Efficiently estimating erdos-renyi graphs with node differential privacy. *arXiv preprint arXiv:1905.10477*, 2019. 16

[60] Matthew O Jackson and Brian W Rogers. Meeting strangers and friends of friends: How random are social networks? *American Economic Review*, 97:890–915, 2007. 17

[61] Kai Chen and Jiaqi Wang. Mmdetection: Open mmlab detection toolbox and benchmark. *arXiv preprint arXiv:1906.07155*, 2019. 18

[62] Ba-Dung Le, Hong Shen, Hung Nguyen, and Nickolas Falkner. Improved network community detection using meta-heuristic based label propagation. *Applied Intelligence*, 49:1451–1466, 2019. 18

[63] Zheng-Hong Deng, Hong-Hai Qiao, Ming-Yu Gao, Qun Song, and Li Gao. Complex network community detection method by improved density peaks model. *Physica A: Statistical Mechanics and its Applications*, 526:121–128, 2019. 18

[64] Andrea Lancichinetti and Santo Fortunato. Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. *Physical Review E*, 80:118–129, 2009. 18

[65] Clara Granell, Richard K Darst, Alex Arenas, Santo Fortunato, and Sergio Gómez. Benchmark model to assess community structure in evolving networks. *Physical Review E*, 92:12–19, 2015. 18

[66] Pattison P Wasserman S. Logit models and logistic regressions for social networks: I. an introduction to markov graphs and p. *Psychometrika*, 61(3):401–425, 1996. 20

[67] MEJ Park J, Newman. Solution for the properties of a clustered network. *Phys Rev E*, 72:026136, 2005. 20

[68] Prusaczyk B, Maki J, Luke DA, and Lobb R. Rural health networks: How network analysis can inform patient care and organizational collaboration in a rural breast cancer screening network. *J Rural Health*, 2018. 20

[69] Mônica da Silva Santana and Everardo Valadares de Sá Barretto Sampaio. Carbon and nitrogen stocks of soils under different land uses in pernambuco state, brazil. *Geoderma Regional*, 16:198–205, 2019. 20

[70] Diaconis P Chatterjee S. Estimating and understanding exponential random graph models. *Ann Stat*, 41:2428–2461, 2013. 21, 72

[71] Kalish Y Lusher D Robins G, Pattison P. An introduction to exponential random graph (p*) models for social networks. *Soc networks*, 29:173–191, 2007. 21

[72] Tom AB Snijders. Markov chain monte carlo estimation of exponential random graph models. *J Soc Struct*, 3:1–40, 2002. 21

[73] Snijders T Moody J Besag J Handcock MS, Robins G. Assessing degeneracy in statistical models of social networks. In *J Am Stat Assoc*, 2003. 21, 72

[74] George G Vega Yon and Kayla de la Haye. Exponential random graph models for little networks. *arXiv preprint arXiv:1904.10406*, 2019. 21

[75] Sly A Bhamidi S, Bresler G. Mixing time of exponential random graphs. In *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on*, pages 803–812, 2008. 21

[76] Butts CT Goodreau SM Morris M Handcock MS, Hunter DR. statnet: Software tools for the representation, visualization, analysis and simulation of network data. *J Stat Softw*, 24:1–11, 2008. 21

[77] Tjeerd Zandberg and Mark Huisman. Missing behavior data in longitudinal network studies: the impact of treatment methods on estimated effect parameters in stochastic actor oriented models. *Social Network Analysis and Mining*, 9:8, 2019. 21

[78] Beth Prusaczyk, Julia Maki, Douglas A Luke, and Rebecca Lobb. Rural health networks: How network analysis can inform patient care and organizational collaboration in a rural breast cancer screening network. *The Journal of Rural Health*, 35:222–228, 2019. 21

[79] Pranav Nerurkar, Madhav Chandane, and Sunil Bhirud. Measurement of network based and random meetings in social networks. *Turkish Journal Of Electrical Engineering & Computer Sciences*, 27:765–779, 2019. 21

[80] Pranav Nerurkar, Madhav Chandane, and Sunil Bhirud. Understanding attribute and social circle correlation in social networks. *Turkish Journal Of Electrical Engineering & Computer Sciences*, 27:1228–1242, 2019. 21

[81] Kenji Yamanishi, Tianyi Wu, Shinya Sugawara, and Makoto Okada. The decomposed normalized maximum likelihood code-length criterion for selecting hierarchical latent variable models. *Data Mining and Knowledge Discovery*, 33:1017–1058, 2019. 22

[82] Hoff PD. Dyadic data analysis with amen. *arXiv preprint arXiv:1506.08237*, 2015. 22

[83] David Madras, Elliot Creager, Toniann Pitassi, and Richard Zemel. Fairness through causal awareness: Learning causal latent-variable models for biased data. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 349–358, 2019. 22

[84] Geoffrey J McLachlan, Sharon X Lee, and Suren I Rathnayake. Finite mixture models. *Annual review of statistics and its application*, 6:355–378, 2019. 22

[85] Tsitsulin A, Mottin D, Karras P, and Müller E. Verse: Versatile graph embeddings from similarity measures. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*, pages 539–548, 2018. 23, 33, 36, 73, 124

[86] Ou M, Cui P, Pei J, Zhang Z, and Zhu W. Asymmetric transitivity preserving graph embedding. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1105–1114, 2016. 23, 28, 33, 73, 124

[87] Rozemberczki B, Davies R, Sarkar R, and Sutton C. Gemsec: Graph embedding with self clustering. *arXiv preprint arXiv:1802.03997*, 2018. 23, 73, 124

[88] Rozemberczki B and Sarkar R. Fast sequence based embedding with diffusion graphs. In *International Conference on Complex Networks*, 2018. 23, 73, 124

[89] Yang Z, Cohen WW, and Salakhutdinov R. Revisiting semi-supervised learning with graph embeddings. *arXiv preprint arXiv:1603.08861*, 2016. 23, 73, 124

[90] Perozzi B, Al-Rfou R, and Skiena S. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 701–710, 2014. xvi, 23, 24, 31, 33, 73, 79, 124

[91] Grover A and Leskovec J. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864, 2016. 23, 31, 33, 73, 124

[92] Sheikh N, Kefato Z, and Montresor A. gat2vec: representation learning for attributed graphs. *Computing*, 2018. 23, 73, 124

[93] Mikolov T, Sutskever I, Chen K, Corrado GS, and Dean J. Distributed representations of words and phrases and their compositionality. In *Adv Neur In*, pages 3111–3119, 2013. 23, 31, 33, 73, 124

[94] Cao S, Lu W, and Xu Q. Grarep: Learning graph representations with global structural information. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pages 891–900, 2015. 23, 27, 33, 73, 124

[95] Liu Q, Li Z, Lui J, and Cheng J. Powerwalk: Scalable personalized pagerank via random walks with vertex-centric decomposition. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 195–204, 2016. 23, 73, 124

[96] Pandhre S, Mittal H, Gupta M, and Balasubramanian Vineeth N. Stwalk: learning trajectory representations in temporal graphs. In *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, pages 210–219, 2018. 23, 73, 124

[97] Mikolov T, Chen K, Corrado G, and Dean J. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013. 23, 73, 124

[98] Lin Y, Liu Z, Sun M, Liu Y, and Zhu X. Learning entity and relation embeddings for knowledge graph completion. In *AAAI*, volume 15, pages 2181–2187, 2015. 23, 73, 124

[99] Wang Z, Ye X, Wang C, Wu Y, Wang C, and Liang K. Rsdne: Exploring relaxed similarity and dissimilarity from completely-imbalanced labels for network embedding. *Network*, 11:14. 23, 73, 124

[100] Justin Wang, Raymond KW Wong, and Thomas CM Lee. Locally linear embedding with additive noise. *Pattern Recognition Letters*, 2019. 27

[101] Zhang Jingjing. Identification of moving loads using a local linear embedding algorithm. *Journal of Vibration and Control*, pages 1077–1107, 2019. 27

[102] Liao L, He X, Zhang H, and Chua TS. Attributed social network embedding. *arXiv preprint arXiv:1705.04969*, 2017. 28

[103] Liang J, Jacobs P, Sun J, and Parthasarathy S. Semi-supervised embedding in attributed networks with outliers. In *Proceedings of the 2018 SIAM International Conference on Data Mining*, pages 153–161, 2018. xvi, 30

[104] Yang C, Liu Z, Zhao D, Sun M, and Chang EY. Network representation learning with rich text information. In *IJCAI*, pages 2111–2117, 2015. 31, 33

[105] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE Transactions on Neural Networks*, 20:61–80, 2009. 32, 33

[106] Chen CM, Tsai MF, Lin YC, and Yang YH. Query-based music recommendations via preference embedding. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 79–82, 2016. 33

[107] Dai H, Dai B, and Song L. Discriminative embeddings of latent variable models for structured data. In *International Conference on Machine Learning*, pages 2702–2711, 2016. 33

[108] Narayanan A, Chandramohan M, Chen L, Liu Y, and Saminathan S. subgraph2vec: Learning distributed representations of rooted sub-graphs from large graphs. *arXiv preprint arXiv:1606.08928*, 2016. 33

[109] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34:18–42, 2017. 33

[110] Jihun Oh, Kyunghyun Cho, and Joan Bruna. Advancing graphsage with a data-driven node sampling. *arXiv preprint arXiv:1904.12935*, 2019. 35

[111] Zhihao Jia, Sina Lin, Rex Ying, Jiaxuan You, Jure Leskovec, and Alex Aiken. Redundancy-free computation graphs for graph neural networks. *arXiv preprint arXiv:1906.03707*, 2019. 35

[112] Gerome Vivar, Hendrik Burwinkel, Anees Kazi, Andreas Zwergal, Nassir Navab, and Seyed-Ahmad Ahmadi. Multi-modal graph fusion for inductive disease classification in incomplete datasets. *arXiv preprint arXiv:1905.03053*, 2019. 35

[113] Tang J, Qu M, and Mei Q. Pte: Predictive text embedding through large-scale heterogeneous text networks. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1165–1174, 2015. 35

[114] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S Yu. A comprehensive survey on graph neural networks. *arXiv preprint arXiv:1901.00596*, 2019. xix, 36, 37, 38

[115] Jie Zhou, Ganqu Cui, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, and Maosong Sun. Graph neural networks: A review of methods and applications. *arXiv preprint arXiv:1812.08434*, 2018. xix, 36, 37, 38

[116] Yujia Li, Daniel Tarlow, Marc Brockschmidt, and Richard Zemel. Gated graph sequence neural networks. *arXiv preprint arXiv:1511.05493*, 2015. 37

[117] Yifan Hou, Hongzhi Chen, Changji Li, James Cheng, Ming-Chang Yang, and Fan Yu. A representation learning framework for property graphs. 2019. 37

[118] Guoli Yang, Cheng Zhu, and Weiming Zhang. Adaptive and probabilistic strategy evolution in dynamical networks. *Physica A: Statistical Mechanics and its Applications*, 518:99–110, 2019. 40

[119] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data, June 2014. 46, 47, 49

[120] Mira A Amati V, Lomi A. Social network modelling. *Annu Rev Stat Appl*, 5, 2018. 71

[121] Thomas G Dietterich. Ensemble methods in machine learning. In *International workshop on multiple classifier systems*, pages 1–15, 2000. 73

[122] Thomas N Kipf and Max Welling. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308*, 2016. 89

[123] Qimai Li, Zhichao Han, and Xiao-Ming Wu. Deeper insights into graph convolutional networks for semi-supervised learning. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018. xvii, 92

[124] François Chollet et al. Keras. https://keras.io, 2015. 94, 115

[125] Jiquan Ngiam, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee, and Andrew Y Ng. Multimodal deep learning. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 689–696, 2011. 111

[126] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 111