Div-HeteRec over Learn-HeteRec is more pronounce in the Hyderabad Meetup dataset. This is due to the fact that, Bangalore has more of information technology (IT) culture and majority of the groups are pertaining to IT only. Hence, there is a lack of diversity amongst groups in Bangalore. As a result of this diversity in recommended groups does not make much sense for this data.

Next we analyze the results for *Tag* to *Group* recommendation. Both RWR and Collaborative Filtering using NMF fail drastically in this task. Also, the Uni-HeteRec model which was performing satisfactorily in previous task fails. The same (uniform) parameter values which were doing reasonably well in previous task do not work in this setting. This emphasizes the fact that, for different tasks, different relations carry different priorities. However, unlike the previous case where Div-HeteRec was performing best here we find that Learn-HeteRec outperforms all other approaches. This is in-line with the intuition that the tags of groups are closely related to each other and suggesting a diverse set of tags should hurt the performance.

Summarizing our observations, we find that

- Making diverse recommendations in cases where they are intuitive also improves the performance of a recommender system.
- learning the importance of relations always helps; different recommendation tasks will require different weightages to be assigned to various relations.

7 CONCLUSION

In this paper, we introduced the need for diversity in recommendations in a heterogeneous information network (HIN). We proposed a sublinear vertex reinforced random walk based mechanism for integrating diversity. This emulates rich getting richer mechanism by increasing transitions to nodes which have been visited more frequently in the past. Further, in a HIN, different relation types have significantly different priorities for a given recommendation task. In a multivariate random walk framework we proposed a cross entropy cost based learning framework for systematically learning these priorities. Finally, we demonstrated the effectiveness of our approach using real-world Meetup dataset. So, naturally for recommendations requiring diversity the proposed Div-HeteRec is performing the best and for situations where diversity is not essential, the proposed Learn-HeteRec is ideally suited. It is interesting to explore, in the future, a hybrid model that aptly combines both diverse and non-diverse requirements. We make the code and datasets used in our experiments publicly available.

REFERENCES

[1] Jaime Carbonell and Jade Goldstein. 1998. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval.* ACM, 335–336.

[2] Wai-Ki Ching, Ximin Huang, Michael K Ng, and Tak Kuen Siu. 2013. Markov Chains: Models, Algorithms and Applications. (2013).

[3] Marina Drosou and Evaggelia Pitoura. 2010. Search result diversification. *ACM SIGMOD Record* 39, 1 (2010), 41–47.

[4] Avinava Dubey, Soumen Chakrabarti, and Chiranjib Bhattacharyya. 2011. Diversity in ranking via resistive graph centers. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM.

[5] Rachid Guerraoui, Anne-Marie Kermarrec, Tao Lin, and Rhicheek Patra. 2017. Heterogeneous recommendations: what you might like to read after watching interstellar. *Proceedings of the VLDB Endowment* 10, 10 (2017), 1070–1081.

[6] Taher H Haveliwala. 2002. Topic-sensitive pagerank. In *Proceedings of the 11th international conference on World Wide Web.* ACM, 517–526.

[7] Daniel S Himmelstein and Sergio E Baranzini. 2015. Heterogeneous network edge prediction: a data integration approach to prioritize disease-associated genes. *PLoS computational biology* 11, 7 (2015), e1004259.

[8] Jon M Kleinberg. 1999. Hubs, authorities, and communities. *ACM computing surveys (CSUR)* 31, 4es (1999), 5.

[9] Rong-Hua Li and Jeffery Xu Yu. 2013. Scalable diversified ranking on large graphs. *IEEE Transactions on Knowledge and Data Engineering* 25, 9 (2013), 2133–2146.

[10] Qiaozhu Mei, Jian Guo, and Dragomir Radev. 2010. Divrank: the interplay of prestige and diversity in information networks. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining.* Acm, 1009–1018.

[11] Sharad Nandanwar and M Narasimha Murty. 2016. Structural neighborhood based classification of nodes in a network. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* ACM, 1085–1094.

[12] Sharad Nandanwar and M Narasimha Murty. 2018. Overlap-Robust Decision Boundary Learning for Within-Network Classification. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence.* AAAI.

[13] Zaiqing Nie, Yuanzhi Zhang, Ji-Rong Wen, and Wei-Ying Ma. 2005. Object-level ranking: bringing order to web objects. In *Proceedings of the 14th international conference on World Wide Web.* ACM, 567–574.

[14] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. *The PageRank citation ranking: Bringing order to the web.* Technical Report. Stanford InfoLab.

[15] Robin Pemantle. 1992. Vertex-reinforced random walk. *Probability Theory and Related Fields* 92, 1 (1992), 117–136.

[16] Tuan-Anh Nguyen Pham, Xutao Li, Gao Cong, and Zhenjie Zhang. 2015. A general graph-based model for recommendation in event-based social networks. In *Data Engineering (ICDE), 2015 IEEE 31st International Conference on.* IEEE, 567–578.

[17] Filip Radlinski, Paul N Bennett, Ben Carterette, and Thorsten Joachims. 2009. Redundancy, diversity and interdependent document relevance. In *ACM SIGIR Forum*, Vol. 43. ACM, 46–52.

[18] Chuan Shi, Xiangnan Kong, Yue Huang, S Yu Philip, and Bin Wu. 2014. Hetesim: A general framework for relevance measure in heterogeneous networks. *IEEE Transactions on Knowledge and Data Engineering* 26, 10 (2014), 2479–2492.

[19] Yizhou Sun, Rick Barber, Manish Gupta, Charu C Aggarwal, and Jiawei Han. 2011. Co-author relationship prediction in heterogeneous bibliographic networks. In *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on.* IEEE, 121–128.

[20] Yizhou Sun and Jiawei Han. 2012. Mining heterogeneous information networks: principles and methodologies. *Synthesis Lectures on Data Mining and Knowledge Discovery* 3, 2 (2012), 1–159.

[21] Yizhou Sun, Jiawei Han, Xifeng Yan, Philip S Yu, and Tianyi Wu. 2011. Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. *Proceedings of the VLDB Endowment* 4, 11 (2011), 992–1003.

[22] Hanghang Tong, Jingrui He, Zhen Wen, Ravi Konuru, and Ching-Yung Lin. 2011. Diversified Ranking on Large Graphs: An Optimization Viewpoint. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '11).* ACM, New York, NY, USA, 1028–1036. DOI: http://dx.doi.org/10.1145/2020408.2020573

[23] Stanislav Volkov. 2001. Vertex-reinforced random walk on arbitrary graphs. *Annals of probability* (2001), 66–91.

[24] Stanislav Volkov. 2006. Phase transition in vertex-reinforced random walks on with non-linear reinforcement. *Journal of Theoretical Probability* 19, 3 (2006), 691–700.

[25] Xiao Yu, Xiang Ren, Yizhou Sun, Quanquan Gu, Bradley Sturt, Urvashi Khandelwal, Brandon Norick, and Jiawei Han. 2014. Personalized entity recommendation: A heterogeneous information network approach. In *Proceedings of the 7th ACM international conference on Web search and data mining.* ACM, 283–292.

[26] Cheng Xiang Zhai, William W Cohen, and John Lafferty. 2003. Beyond independent relevance: methods and evaluation metrics for subtopic retrieval. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval.* ACM, 10–17.

[27] Mi Zhang and Neil Hurley. 2008. Avoiding monotony: improving the diversity of recommendation lists. In *Proceedings of the 2008 ACM conference on Recommender systems.* ACM, 123–130.

[28] Shenghuo Zhu, Kai Yu, Yun Chi, and Yihong Gong. 2007. Combining content and link for classification using matrix factorization. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval.* ACM, 487–494.

[29] Xiaojin Zhu, Andrew B Goldberg, Jurgen Van Gael, and David Andrzejewski. 2007. Improving Diversity in Ranking using Absorbing Random Walks.. In *HLT-NAACL.* 97–104.