International Journal on Engineering Artificial Intelligence Management, Decision Support, and Policies
Vol. 01, Iss. 01, S. No. 002, pp. 12-20, July 2024
ISSN(E): 3048-8788

Machine learning approach for link prediction in large stochastic online social networks (SOSNs)

Shivshankar Rajput[1], Anil Pimpalapure[2], Praveen Bhanodia[3]

(Affiliation of authors): [1,2]School of Computer Science, Eklavya University, Damoh,MP, India
[3] Computer Science Engineering, Acropolis Institute of Technology and Research, Indore ,MP, India

[1]shivshankarrajpoot@gmail.com, [2]pimpalapureanil@gmail.com, [3]praveenbhanodia@acropolis.in

How to cite this paper: Shivshankar Rajput, Anil Pimpalapure, Praveen Bhanodia, "Machine learning approach for link prediction in large stochastic online social networks (SOSNs)," *International Journal on Engineering Artificial Intelligence Management, Decision Support, and Policies*, Vol. no. 1, Iss. No 1, S No. 002, pp. 12–20, July 2024.

Received:03/07/2024
Revised: 23/07/2024
Accepted: 31/07/2024
Published : 31/07/2024

Abstract

Social networks are growing every day at a tremendous pace. A social network is an online community that allows users to interact and exchange ideas, information, activities, and interests. Millions of users are contributing to its character and behaviour, and the information being generated has a multitude of dimensional aspects that provide new opportunities and perspectives for the computation of network properties. Every individual quickly spreads messages and information throughout all linked groupings. Social networks give users the ability to make a profile, connect with others, and communicate with them through a variety of tools like chatting, updating their status, leaving comments, and sharing text, images, audio, and animated videos. With billions of users globally, social media platforms have emerged as an essential part of contemporary social interaction and communication. Along with changing how individuals share and consume information, they have also had a big impact on social, political, and cultural developments. These days, social communities, governmental agencies, and commercial enterprises all depend heavily on online social networks. Understanding networks' evolutionary nature requires the ability to predict missing links in existing networks and emerging or broken linkages in future networks. Since SNs undergo dynamic changes over time, link inference in these networks is an extremely difficult task. The dynamic character of supernovae is not well-accounted for in link prediction techniques. In this study, link prediction techniques in dynamic SNs will be thoroughly reviewed, analysed, discussed, and evaluated.

Keywords

Social Networks, Machine Learning, Social Communities

1. Introduction

Kleinberg and Liben-Nowell introduce the link prediction problem for Social Networks [SNs][1]. The challenge of link prediction is to determine, given the existing structure of a graph (sociogram or network), between which nodes and edge will develop following the disconnected nodes in future. So far, there has been some success in link prediction using traditional social network analysis criteria that are simple. This study uses a dataset gathered from messages exchanged on a dating social networking website to investigate the issue in greater detail. Social networks are crucial for grasping the nature of entities and human behavior. More than millions of peoples participating in more than one online-social networking application/sites, social network research has become significant attention [2]. A wealth of details can be found in SNs, which can be mined or examined to draw insightful conclusions.

The global connectivity enabled by online social networks has revolutionized the way we communicate, share information, and build relationships in the modern world. With the growing use and popularity of these platforms, researchers have begun employing various mathematical models to explore the dynamics and characteristics of these networks. One approach to understanding the friendships and communication patterns within these networks is the review of Stochastic Online Social Networks (SOSNs). Stochasticity, referring to randomness or uncertainty, suggests that individual behaviour and interactions are influenced by probabilistic factors [3,4]. Unlike traditional social network models, which assume consistent and predictable interactions, stochastic models account for the inherent unpredictability and variability in human behaviour within these virtual communities.

A SN is the outcome of a collection of individuals' diverse personal and social relationships. It can be used for a variety of purposes, such as political awareness campaigns, blogging, advertising, and review collection. Since SNs are always growing, they are extremely dynamic. Comprehending and optimizing the quantity of links is a primary concern in SNs, as it facilitates the effective utilization of SN services and ensures prompt information dissemination across the network. The method of determining the most likely linkages between disconnected nodes in a social network is known as link prediction[5,6,7]

With the rise of social media platforms like Facebook and Twitter, social network analysis has garnered significant interest. Social networks are dynamic structures composed of nodes (users) and edges (relationships). Link prediction, a key area in this field, aims to predict future connections based on current network characteristics [8]. Traditional methods are limited to static and homogeneous networks, necessitating advanced models for dynamic and heterogeneous networks.

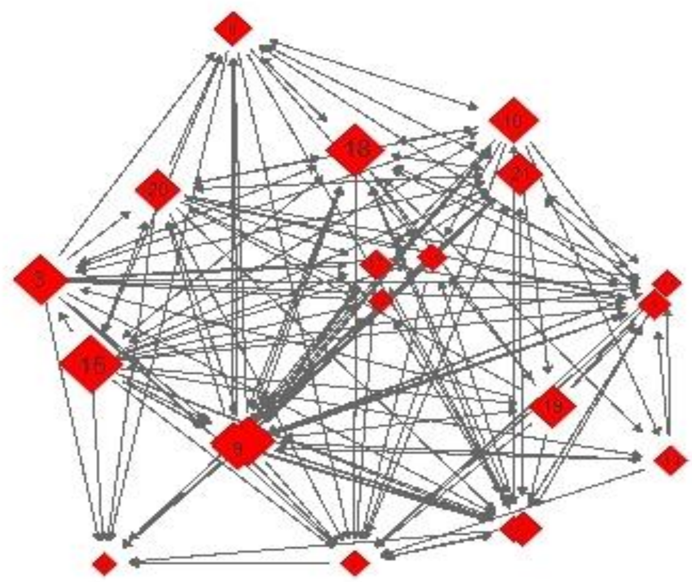

Figure 1. Social Network Representation by Directed Graph

2 Literature Review

The prediction of future links in social networks involves various methods using both local and global features. These methods typically compute a similarity measure between node pairs, which can be based on node features (like the number of adjacent nodes or common neighbors) or network topological features (like the path or distance between nodes) [9,10,11]. The performance of these methods varies because the features are independent, making it hard to define consistent similarity measures for predicting future links. To address performance gaps, probabilistic models learn from the social network, using local and global features as random variables approximated by Bayesian methods.

Alternative approaches include adjacency matrix recovery, which applies sparse matrix recovery or collaborative filtering. However, statistical models for link prediction in social networks are found to be more stable and accurate than matrix recovery approaches. Despite this, combining multiple approaches is suggested to improve link prediction [12,13].

Conventional node similarity methods and neural network techniques have been proposed for link prediction, especially for networks with multiple node types and dynamic links. For example, a multivariate NARX model was developed for predicting new and recurring links in heterogeneous networks, and a supervised link prediction model used meta-path features to support neural network models [14]. Another model used belief function theory to quantify link existence based on neighboring node information.

Network structural entities and their attributes can also derive similarity measures for potential node pairs[15,17]. A similarity measure using network motifs was developed within a supervised learning framework.

Despite extensive research, the dynamic and varied nature of online social networks (OSNs) means no single method can predict future connections accurately across all types of networks [16].

3 Machine Learning Techniques

3.1 Clusters

In order to find naturally occurring groupings within a dataset, clustering is a machine learning technique that divides data or objects into classes, groups, or clusters. This method can be used in a variety of fields, such as business research, social network analysis, and spatial data analysis. Euclidean distance is widely used by clustering algorithms to calculate the spatial separation between data objects and create clusters [18,19,20].

The complete network dataset, which consists of nodes and edges, is taken into consideration when social networks are utilised for link prediction [23]. Based on how similar certain features are among the nodes, clustering is done. These characteristics can be global (path distance between nodes) or local (node attributes). The method outlined focuses on analysing node attributes to identify similarity measures, which are subsequently utilised as clustering distance measurements [21,22].

3.2 Clustering methods

A clustering algorithm for link prediction in social networks typically involves several key steps, which include pre-processing the network data, computing similarity measures, clustering nodes based on these measures, and finally predicting potential links. Here's a general approach [65,66,67].

3.2.1. Data Pre-processing

Network Representation [23,24,25]: Represent the SNs as a graph $G2=(V,E)$, where V is the set of nodes (users) and E is the set of edges (existing links between users).

Feature Extraction [26]: Extract relevant features from the nodes and edges. Features can be local (e.g., node degree, number of common neighbors) or global (e.g., shortest path distance between nodes).

3.2.2. Similarity Measurement

- Compute Similarity [27,28]: Calculate similarity scores between all pairs of nodes using chosen similarity measures. Common measures include:
- Common Neighbors [29]: The number of shared neighbors between two nodes.
- Jaccard Coefficient: The ratio of the intersection of neighbors to the union of neighbors.
- Adamic/Adar Index [30]: A weighted measure that gives more importance to shared neighbors with fewer connections.
- Preferential Attachment [31]: The product of the degrees of two nodes.

3.2.3. Clustering

Distance Metric: Use the similarity scores as a distance metric for clustering. Nodes with higher similarity scores are considered closer[32,33].

Clustering Algorithm: Apply a clustering algorithm to group similar nodes together. Common clustering algorithms include [34,35,36,37,38]:

- K-means Clustering: Partition nodes into k clusters based on similarity.
- Hierarchical Clustering: Create a dendrogram to represent nested clusters of nodes.
- DBSCAN (Density-Based Spatial Clustering of Applications with Noise): Cluster nodes based on density, identifying core samples and expanding clusters from them [39,40].

3.2.4. Link Prediction

- Potential Link Identification: Within each cluster, identify pairs of nodes that are not currently connected by an edge but have high similarity scores. These pairs are potential candidates for future links [41].
- Prediction Model: Optionally, use a supervised learning model trained on existing links and non-links to improve prediction accuracy. Features for this model can include similarity measures and other node attributes [42,64].

4. Combined clustering approach

The proposed approach in this article combines two distinct link prediction methods to improve accuracy [43]. The methodology is twofold:

4.1 Phase One: Similarity Index Calculation

- Network nodes are processed, and neighborhood features are examined using Jaccard Coefficient (JC), Adamic/Adar Index (AA), and Preferential Attachment (PA) to generate a similarity index [44].
- Scores for non-existing links are ranked to provide the probability of future link existence [45].

4.2 Phase Two: Clustering

An unsupervised k-means clustering technique is used to classify the network into clusters [46].

The k-means clustering technique is chosen for its speed and efficiency in processing large social networks [62,63].

4.3 Methodology Steps

- Input the social network dataset in the form of an edge list G(u,v).
- Identify non-existing links between node pairs [48].
- Compute mutual neighbor nodes for every non-linked node pair.
- Compute the similarity features for each non-linked node pair.
- Sort the dataset in descending order based on similarity scores.
- Select an arbitrary centroid for the cluster node groups.
- Compute differences between centroids and similarity features, assigning nodes to the cluster where the difference is minimal.

5. Experimental results

This section deals with the assessment and comparison of the proposed solution to the link prediction problem [47,49,50]. Evaluation and comparison of PA,JC and AA shown below table 1.

Table 1. similarity measures for non-existing links and their corresponding clusters

Nodes(x,y)	JC	AA	PA	Custers
(1,4)	0.33	0.73	6	0
(1,5)	0.00	0.00	4	1
(2,4)	0.37	0.77	5	0
(2,5)	0.00	0.00	4	1
(3,5)	0.34	0.74	6	0

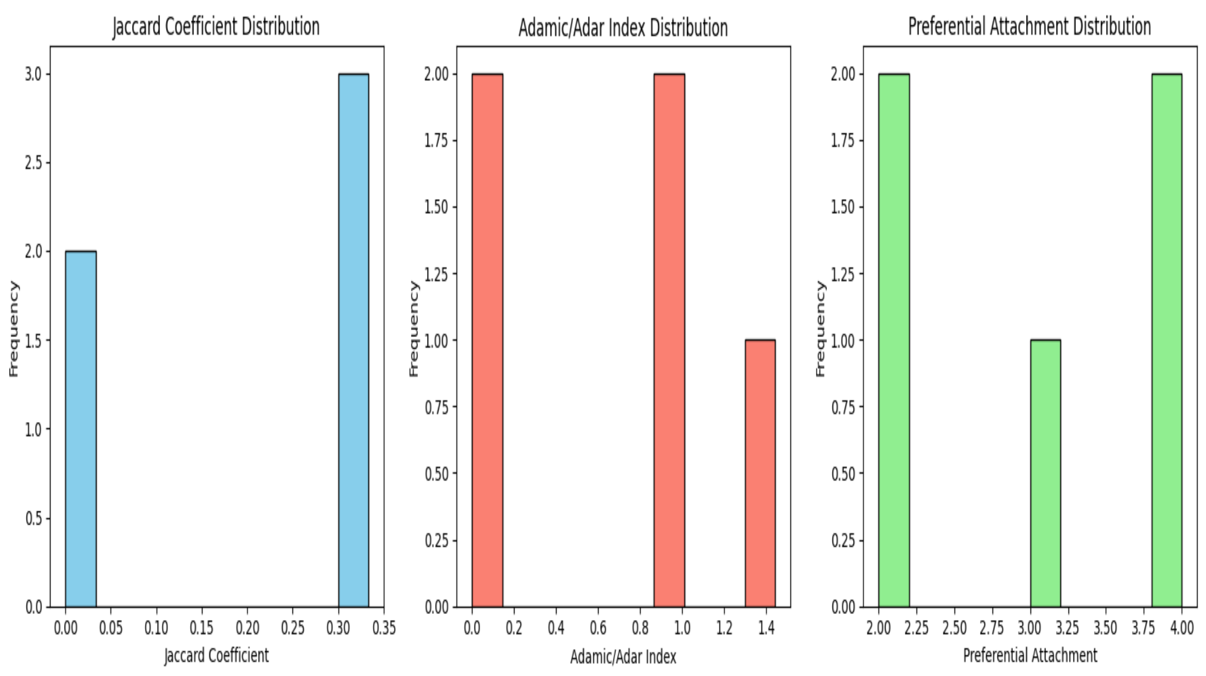

Figure 2. Representation of similarity measures using JC,AA and PA

4. Dataset used

In this paper we will use musae-wiki [68]. These datasets represent page-page networks on specific purpose. Nodes represent person and edges are mutual links between them. It is contain 19,109 nodes and 400832 likns between nodes. This is un-directed network.

7. Conclusion

This paper investigates link prediction in Online Social Networks (OSNs), highlighting the growing complexity due to the increasing size of these networks. It reviews various local feature-based similarity techniques and finds that they fall short as they focus only on local network features [51,52,53].

To address this, the article proposes a novel collaborative filtering-based approach to handle complex and large social network data. The method consists of two phases:

1. **Similarity Score Calculation[54,55]:** The first phase calculates similarity scores between non-linked node pairs using classical techniques-JC,AA and PA. This transforms the original network dataset into a feature-based network representation.

2. **Clustering and Link Prediction [56,57,58]:** The second phase uses k-means clustering on the transformed feature network to categorize nodes into clusters, predicting future links based on these clusters.

The proposed approach is more efficient than classical methods as it combines local feature-based techniques with clustering, reducing computational costs and memory usage. Future work will focus on evaluating the method's accuracy, expanding it to other social networks, and exploring additional unsupervised machine learning methods [59,60,61]. The approach also plans to incorporate considerations for outlier values that could influence results.

Funding: "This research received no external funding"

Conflicts of Interest: "This authors declare no conflict of interest."

References

[1] P. Wang, X. U. Baowen, W. U. Yurong, and X. Y. Zhou, "Link prediction in social networks: the state-of-the-art", Sci Chin, vol.58, no. (1), Jan 2015, pp. 11101-011101.

[2] Liben-Nowell, D. and Kleinberg, J. (2007), 'The link-prediction problem for social networks', Journal of the American society for information science and technology 58(7), 1019–1031.

[3] Ricci, F., Rokach, L. and Shapira, B. (2011), Introduction to recommender systems handbook, Springer.

[4] R. Albert, A.-L. Barab´asi, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (2002) 47.

[5] Z. Huang, X. Li, H. Chen, Link prediction approach to collaborative filtering,In Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries,ACM Press, New York, 2005.

[6] Hasan, Mohammad A., and Chaoji, Vineet, and Salem, Saeed and Zaki, Mohammed. (2006) Link Prediction using Supervised Learning. In Proceedings of SDM Workshop of Link Analysis, Counterterrorism and Security.

[7] Doppa, Janardhan R., and Yu, Jun, and Tadepalli, Prasad, and Getoor, Lise. (2009). Chance-Constrained Programs for Link Prediction. In Proceedings of Workshop on Analyzing Networks and Learning with Graphs at NIPS Conference.

[8] E. A. Leicht, P. Holme, M. E. J. Newman, Vertex similarity in networks, Phys. Rev. E 73 (2006) 026120.

[9] D. Sun, T. Zhou, J.-G. Liu, R.-R. Liu, C.-X. Jia, B.-H. Wang, Information filtering based on transferring similarity, Phys. Rev. E 80 (2009) 017101.

[10] D. R. White, K. P. Reitz, Graph and semigrouphomomorphisms on networks of relations, Social Networks 5 (1983) 193.

[11] P. Holme, M. Huss, Role-similarity based functional prediction in networked systems: application to the yeast proteome, J. R. Soc. Interface 2 (2005) 327.

[12] P. Jaccard, ´Etude comparative de la distribution floraledansune portion des Alpeset des Jura, Bulletin de la SocieteVaudoise des Science Naturelles 37 (1901) 547.

[13] T. Sørensen, A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons, Biol. Skr. 5 (1948) 1.

[14] T. Zhou, L. L¨u, Y.-C. Zhang, Predicting missing links via local information,Eur. Phys. J. B 71 (2009) 623.

[15] Q. Ou, Y.-D. Jin, T. Zhou, B.-H. Wang, B.-Q. Yin, Power-law strength-degree correlation from resource-allocation dynamics on weighted networks, Phys. Rev. E 75 (2007) 021102.

[16] M. Molloy, B. Reed, A critical point for random graphs with a given degree sequence, Random Structure and Algorithms 6 (1995) 161.

[17] R. Yang, C. Yang, X. Peng, and A. Rezaeipanah, "A novel similarity measure of link prediction in multi - layer social networks based on reliable paths, " Concurrency Computation:Pract. Experience, p. e6829, 2022. DOI: 10.1002/cpe.6829.

[18] S. Najari, M. Salehi, V. Ranjbar, and M. Jalili, "Link prediction in multiplex networks based on interlayer similarity," Phys. A: Stat. Mech. Appl., vol. 536, p. 120978, 2019.

[19] M. K. Manshad, M. R. Meybodi, and A. Salajegheh, "A new irregular cellular learning automata-based evolutionary computation for time series link prediction in social networks," Appl. Intell., vol. 51, no. 1, pp. 71–84, 2021.

[20] A. Zareie and R. Sakellariou, "Similarity-based link prediction in social networks using latent relationships between the users," Sci. Rep., vol. 10, no. 1, pp. 1–11, 2020.

[21] R. Xiang, J. Neville, and M. Rogati, "Modeling relationship strengthin online social networks," in International Conference on World WideWeb, 2010.

[22] Ali, Haider, et al. "Energy Efficient Heuristic Algorithm for TaskMapping on Shared-Memory Heterogeneous MPSoCs." 2018 IEEE20th International Conference on High Performance Computing andCommunications; IEEE 16th International Conference on Smart City;IEEE 4th International Conference on Data Science and Systems(HPCC/SmartCity/DSS).IEEE, 2018.

[23] Ali, Haider, et al. "Contention & Energy-Aware Real-Time TaskMapping on NoC Based Heterogeneous MPSoCs." IEEE Access 6(2018): 75110-75123.

[24] Ali, Haider, et al. "Energy Efficient Task Mapping & Scheduling onHeterogeneous NoC-MPSoCs in IoT Based Smart City." 2018 IEEE20th International Conference on High Performance Computing andCommunications; IEEE 16th International Conference on Smart City;IEEE 4th International Conference on Data Science and Systems(HPCC/SmartCity/DSS). IEEE, 2018.

[25] Sharma, Upasana, Sharma, Khatri, and S. Kumar, "Elimination basedalgorithm for link prediction on social networks," International Journalof System Assurance Engineering & Management, vol. 6, no. 1, pp. 78-82, 2015.

[26] L. A. Adamic and E. Adar, "Friends and neighbours on the Web," Social Networks, vol. 25, no. 3, pp. 211-230, 2003.

[27] S. N. Dorogovtsev and J. F. Mendes, "Evolution of networks," Advances in physics, vol. 51, pp. 1079-1187, 2002.

[28] .J. A. Hanley and B. J. McNeil, "A method of comparing the areas under receiver operating characteristic curves derived from the same cases," Radiology, vol. 148, pp. 839-843, 1983.

[29] J. L. Herlocker, J. A. Konstan, L. G. Terveen and J. T. Riedl, "Evaluating collaborative filtering recommender systems," ACM Transactions on Information Systems (TOIS), vol. 22, pp. 5-53, 2004.

[30] Kuo T T, Yan R, Huang Y Y, et al. Unsupervised link prediction using aggregative statistics on heterogeneous social networks[C]//Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2013: 775-783.

[31] M. E. Newman, "Clustering and preferential attachment in growing networks," Physical Review E, vol. 64, 025102, 2001.

[32] Liben-Nowell, D 2005, An Algorithmic Approach to Social Networks, Ph.D thesis at MIT Computer Science and Artificial Intelligence Laboratory.

[33] .Leskovec, J, Kleinberg, J & Faloutsos, C 2005, 'Graphs over Time: Densification Laws, Shrinking Diameters and Possible Explanations', Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, 21-24 August 2005.

[34] CHOUDHURY, N. & UDDIN, S. 2016. Time-aware link prediction to explore network effects on temporal knowledge evolution. Scientometrics, 108, 745-776.

[35] AHMED, N. M. & CHEN, L. 2016. An efficient algorithm for link prediction in temporal uncertain social networks. Information Sciences, 331, 120-136.

[36] PAN, L., ZHOU, T., LÜ, L. & HU, C.-K. 2016. Predicting missing links and identifying spurious links via likelihood analysis. Scientific reports, 6

[37] XIE, F., LIU, J., TANG, M., ZHOU, D., CAO, B. & SHI, M. Multi-relation Based Manifold Ranking Algorithm for API Recommendation. Advances in Services Computing: 10th Asia-Pacific Services Computing Conference, APSCC 2016, Zhangjiajie, China, November 16-18, 2016, Proceedings 10, 2016. Springer, 15-32.

[38] CHEN, H.-H., MILLER, D. J. & GILES, C. L. 2013. The predictive value of young and old links in a social network. Proceedings of the ACM SIGMOD

[39] Barabasi, A. L., Linked – The New Science of Networks, Perseus, 2002.

[40] Getoor, L., Diehl, C. P., Link Mining: A Survey. SIGKDD Explorations, Vol.7, No.2, pp.3-12, 2005.

[41] Hasan, M. A., Chaoji, V., Salem, S., Zaki, M. Link Prediction using Supervised Learning, in workshop on link discovery; issues, approaches and applications, 2005.

[42] Huang, Z., Link Prediction Based on Graph Topology: The Predictive Value of the Generalized Clustering Coefficient, in Proceedings of LinkKDD, 2006.

[43] Kautz, H., Selman, B., Shah, M. The Hidden Web. AI Magazine, Vol. 18, No. 2, 27-36, 1997.

[44] O'Madadhaim, J., Hutchins, J., Smyth, P., Prediction and ranking algorithms for event-based network data, SIGKDD Explorations, Vol.7, No.2, pp.23-30, 2005.

[45] Popescul, A., Ungar, L. H. Statistical relational learning for link prediction, in IJCAI workshop on learning statistical models from relational data, 2003.

[46] Julian, K., & Lu, W. (2016). Application of Machine Learning to Link Prediction.

[47] Caiyan, D., Chen, L., Li, B. (2107). Link prediction in complex network based on modularity. Journal of Soft Computing, 2017, 21(15): 4,197–4,214.

[48] Moradabadi, B. and Meybodi, M. R. (2108). Link prediction in stochastic social networks: Learning automata approach. Journal of Computational Science, 2018, 24: 313–328.

[49] Niladri S., Saptarshi B. , Sukumar N., Sanasam R. S. (2018). Temporal link prediction in multi-relational network. World Wide Web Journal, 2018, 21(2): 395–419.

[50] Almansoori W., Gao S., Jarada T. N., et al. Link prediction and classification in social networks and its application in healthcare and systems biology. Network Modeling Analysis Health Informatics Bioinformatics, 2012, 1: 27–36.

[51] Ibrahim, N. M. A. and Chen, L. (2015). Link prediction in dynamic social networks by integrating different types of information. Applied Intelligence, 42(4), 738–750.

[52] Bliss, C. A., Frank, M. R., Danforth, C. M. and Dodds, P. S. (2014). An evolutionary algorithm approach to link prediction in dynamic social networks. Journal of Computational Science, 5(5), 750–764.

[53] Bhat, S. Y. and Abulaish, M. (2015). HOCTracker: Tracking the evolution of hierarchical and overlapping communities in dynamic social networks. IEEE Transactions on Knowledge and Data Engineering, 27(4), 1,019–1,013.

[54] Juszczyszyn, K., Musial, K. and Budka, M. (2011, October). Link prediction based on subgraph evolution in dynamic social networks. In Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on (pp. 27–34). IEEE.

[55] enda, T., Angelova, R. and Bedathur, S. (2009, June). Towards timeaware link prediction in evolving social networks. In Proceedings of the 3rd workshop on social network mining and analysis (p. 9). ACM.

[56] Greene, D., Doyle, D. and Cunningham, P. (2010, August). Tracking the evolution of communities in dynamic social networks. In Advances in social networks analysis and mining (ASONAM), 2010 International Conference (pp. 176–183). IEEE.

[57] Gao, S., Denoyer, L. and Gallinari, P. (2011, October). Temporal link prediction by integrating content and structure information. In Proceedings of the 20th ACM International Conference on Information and Knowledge Management (pp. 1,169–1,174). ACM.

[58] Sarkar, P., Chakrabarti, D. and Jordan, M. (2012). Nonparametric link prediction in dynamic networks. arXiv preprint arXiv:1206.6394.

[59] Kunegis, J., De Luca, E.W. and Albayrak, S. 2010. The link prediction problem in bipartite networks. In International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems. 380-389. Springer, Berlin, Heidelberg. DOI= https://doi.org/10.1007/978-3-642-14049-5_39.

[60] Su Z, Zheng X, Ai J, Shang L, Shen Y (2019) Link prediction in recommender systems with confidence measures. Chaos Inter J Nonlinear Sci 29(8):083133.

[61] Daud NN, Ab Hamid SH, Saadoon M, Sahran F, Anuar NB (2020) Applications of link prediction in social networks: a review. J Netw Comput Appl 166:102716.

[62] Kumar A, Singh SS, Singh K, Biswas B (2020) Link prediction techniques, applications, and performance: a survey. Physica A 553:124289.

[63] Mishra S, Singh SS, Kumar A, Biswas B (2022) Elp: link prediction in social networks based on ego network perspective. Physica A 605:128008.

[64] Arrar, D., Kamel, N. & Lakhfif, A. A comprehensive survey of link prediction methods. J Supercomput 80, 3902–3942 (2024). https://doi.org/10.1007/s11227-023-05591-8.

[65] Y. V. Nandini, T. Jaya Lakshmi, Murali Krishna Enduri and Hemlata Sharma, Link Prediction in Complex Networks Using Average Centrality-Based Similarity Score Journal: Entropy, 2024, Volume 26, Number 6, Page 433.DOI: 10.3390/e26060433

[66] Y. V. Nandini, T. Jaya Lakshmi, Murali Krishna Endur, Link Prediction in Complex Networks: An Empirical Review Intelligent Data Engineering and Analytics, 2023, Volume 371.ISBN : 978-981-99-6705-6.

[67] Li, M., Zhou, S., Chen, G. (2023). Link Prediction of Complex Networks Based on Local Path and Closeness Centrality. In: Dinh, T.N., Li, M. (eds) Computational Data and Social Networks . CSoNet 2022. Lecture Notes in Computer Science, vol 13831. Springer, Cham. https://doi.org/10.1007/978-3-031-26303-3_5.

[68] Stanford Large Network Dataset Collection
[69] Bhanodia PK, Khamparia A, Pandey B. Supervised shift k-means based machine learning approach for link prediction using inherent structural properties of large online social network. Computational Intelligence. 2020;1–18