2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Analyzing the Progression of Alzheimer's Disease in Human Brain Networks

Anjan Chowdhury
CSCR, ISI-Kolkata
anjan_r@isical.ac.in

Swarup Chattopadhyay
SCSE, XIM University
swarupc@xim.edu.in

Kuntal Ghosh
MIU, ISI-Kolkata
kuntal@isical.ac.in

***Abstract*—In this study, we present a new hybrid model that integrates the anatomical and topological characteristics of a brain network. The aim is to effectively capture the structural and/or topological alterations that take place in networks as individuals transition from a healthy control state to the stage of Alzheimer's disease. The utilisation of a brain atlas allows for the assessment of the Euclidean distance between two specific regions of interest (ROIs) inside the brain. This distance is considered a metric of anatomical distance, providing a quantitative representation of an anatomical characteristic. Conversely, the measurement of topological similarity, which assesses a characteristic of topology, is determined by calculating the cosine distance between nodes following the embedding of the whole real brain network into a vector space of dimensionality d. The empirical findings obtained using real-brain network data indicate that the hybridization approach well captures the observed topological variations during the transition from a healthy cognitive state (HC) to Alzheimer's disease (AD).**

***Index Terms*—Alzheimer's disease, Brain Network, Structural property, Node embedding**

I. Introduction

Alzheimer's disease. Alzheimer's disease (AD) is the most common neurodegenerative brain disease and one of the severe forms of dementia occurring among the elderly [1]. Patients with AD show symptoms of impaired memory, speech, analytical thinking, and other critical cognitive functions that affect the performance of daily activities.

Importance of the study of the progression of AD. Thus, it is crucial to identify the changes in structural patterns in the underlying brain networks corresponding to AD patients in order to provide preventive solutions. Therefore, it becomes challenging to discover the underlying patterns of connectivity that cause the slow but progressive changes in functional connectivity in the AD brain networks [2]. **rs-fMRI data to capture the progression of AD.** Recent studies use the resting-state functional magnetic resonance imaging (rs-fMRI) technique to investigate the course of Alzheimer's disease [3]. Investigations are conducted to capture the alteration in functional changes in these various functional parts of the brain [4] during the progression of AD.

Brain network modeling to capture the progression. Presently, network modeling is sharply gaining popularity in the area of network neurosciences for modeling or simulating real-brain networks [5]. One of the important aspects of network modeling is to generate a synthetic target brain network from a real source brain network. Using this generation process, one can easily find out the changes in connection patterns across the nodes in the brain network. The more similar the real and synthetic target brain networks are, the more accurate they are at capturing the underlying connectivity patterns. Current approaches to network modelling techniques [5] include both anatomical and topological relationships to synthesize real-brain networks.

Node embedding as topological features. Sometimes, topological features computed using local properties alone may fail to capture some essential information inside the whole network, which in turn results in a less accurate model. In order to circumvent this, we have employed popular node embedding techniques for computing topological features to more accurately capture the global structural information of a network in this article. Accordingly, we have employed two types of methods, viz., **(i)** methods based on **random walks** and **(ii)** methods based on **graph neural networks (GNN)** for computing **embedded** topological features of a network. In light of the foregoing, we propose a novel model that combines both the anatomical information and **embedded** topological information of the brain to generate the connection probability for simulating the real-brain network and capture the topological changes in the progression from healthy controls (HC) to AD brain networks.

Summary of the contributions. Thus, in summary, in this paper we propose a novel method with the combination of both anatomical and topological similarity to generate a synthetic AD brain network from a real HC brain network in order to capture the changes in the connection pattern while progressing from HC to AD. The complete flowchart of the process of synthesizing an AD brain network from a real HC brain network is represented in Fig. 1.

II. Related Works

Graph Theoretical Analysis in brain imaging. Graph theory offers many network modeling techniques to simulate

ASONAM '23, November 6-9, 2023, Kusadasi, Turkey

ACM ISBN 979-8-4007-0409-3/23/11…$15.00
http://dx.doi.org/10.1145/3625007.3627496

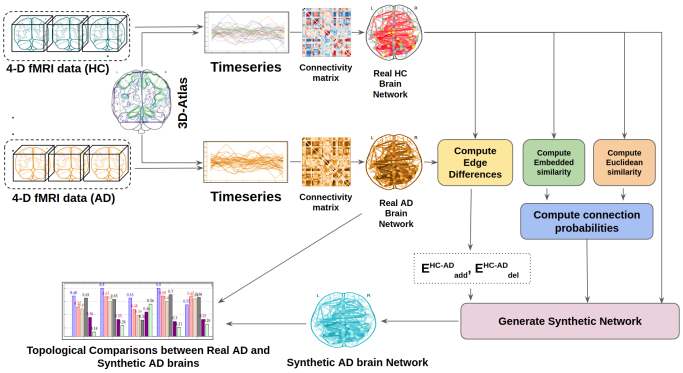


Fig. 1. Full workflow: simulating **synthetic** Alzheimer's Disease (AD) brain from **real** Healthy Control (HC) brain.

the evolutionary processes of real-world complex networks [6], [7]. Network modeling can infer the reasons governing interconnections and explain the mechanisms underlying the creation of a network [8]. It also helps us to identify many important topological properties that can be used for analyzing complex human brain networks [9].

Machine learning techniques in computational neurology. Recently, machine learning approaches are increasingly being used in differentiating between healthy and diseased situations by analyzing physiological patterns (bio-markers) [10], [11]. Recent studies by [12] have used DTI tractography to construct connection matrices and graph matrices from DTI data of Alzheimer's patients. Recent survey paper on this topic is [13].

III. Data Collection and Preprocessing

Dataset gathering. We have collected the rs-fMRI datasets consisting of 120 participants from the Alzheimer's Disease Neuroimaging Initiative (ADNI, http://adni.loni.ucla.edu) repository. Out of these 66 belong to AD and the rest belong to HC.

Preprocessing the Data. We have used Data Processing Assistant for Resting-State fMRI (DPARSF, http://www.rfmri.org/DPARSF) software to preprocess the raw fMRI data. Several steps to preprocess the raw fMRI data include the removal of the first ten volumes [14], distortion correction [15], slice timing corrections [16], motion correction [17] and spatial normalizations [18].

Generation of Brain Networks. We masked the rs-fMRI data with the AAL atlas [19] to extract BOLD signals from 116 regions. Then, for each subject, we computed a 116×116 correlation matrix containing pairwise correlation values. Then generated an average correlation matrix by taking the average of individuals in each group. Then we apply a threshold Th to the correlation matrices to generate binary adjacency matrices for each group. Finally, these binary adjacency matrices represent the graph or network corresponding to HC (G^{HC}_{avg}) and AD (G^{AD}_{avg}) brains.

IV. Proposed Framework

We now describe how we generate the synthetic AD brain network using the real HC brain network that can be found in Algorithm 1. The detailed steps are:

Step 1: Anatomical similarity computation. We have used the standard Euclidean distance as the anatomical distance to compute the anatomical similarity (Euclidean similarity (EDS)) between regions in a brain. M_{EDS} represents the matrix of Euclidean similarity for all pairs of nodes.

Step 2: Topological similarity computation. How two nodes share structural similarity in a network can be determined by topological similarity computation. We used several node embedding techniques to compute topological similarity matrix (M_{TPS}) for all pairs of nodes.

Step 3: Connection probability computation. The connection probability between any node pair (u, v) in a network can be defined as:

$$P_{con}(u,v) = TPS(u,v)^{k1} * EDS(u,v)^{k2} \quad (1)$$

Where $k1$ and $k2$ are two preferential parameters corresponding to topological similarity and anatomical similarity, respectively, $TPS(u,v)$ and $EDS(u,v)$ denote the topological similarity and anatomical (euclidean) similarity between nodes u and v respectively. Finally, P_{con_all} is generated for all (u, v).

Step 4: Modification of HC network. In this step, we modify G^{HC}_{avg} using the generated connection probabilities and the edge differences calculated above. The proposed simulation process will end or stop if N^{AD}_{add} connections are successfully established and N^{AD}_{del} connections are successfully deleted.

Step 5: Returning the generated synthetic network. Finally, the modified network G^{HC}_{avg} is returned, which is actually the synthetic AD network (G^{AD}_{syn}).

A. Proposed Model for Generating Connection Probabilities

Motivation of the proposed construction. Prior research [20] employed only anatomical distance to calculate connection probabilities, followed by a penalized exponential decay model to synthesize real-brain networks. They observed that just penalizing connection probability based on anatomical distance would be insufficient to reproduce the topological properties of real-brain functional networks. Hence there must be some relationship, or trade-off, between distance penalization and one or more other elements, allowing realistically real-brain network structure to form. To preserve the balance or trade-off, we investigated embedded topological similarity and paired it with anatomical similarity. The preferential parameters ($k1$, $k2$) aid in the optimization of the connection probability, as described in Eqn. (1). A power-law based anatomical similarity preference, as well as a power-law function of a topological similarity term, were incorporated in the best-fitting of these connection probability models. The empirical findings imply that using a two-parameter connection probability model with trade-offs improves the accuracy of the model.

Construction of the proposed variants. We employ here three different **random walk-based** node embedding techniques like DW [21], N2V [22], and LINE [23] as well as four different well-known **GNN-based** node embedding techniques such as GCN [24], GraphSAGE [25], ChebyNet [26], and GAT [27] to embed each node in a brain network into vector

Algorithm 1: Synthesizing AD networks from real HC network

Input : G^{HC}_{avg}: Real HC brain network, N^{AD}_{add}, N^{AD}_{del}: number of edges to be added and deleted respectively to generate synthetic AD brain network from real HC brain network, a set of coordinates C representing the location of the brain regions, parameters $k1$ and $k2$.
Output: G^{AD}_{syn}: Synthetic AD brain network.

/* Producing anatomical similarity matrix */
$M_{EDS} \leftarrow$ genAllPairAS(C); Step 1;
/* Producing topological similarity matrix */
$M_{TPS} \leftarrow$ genAllPairTS(G^{HC}_{avg}); Step 2;
/* Producing connection probability matrix */
$P_{con\_all} \leftarrow$ genAllPairCP($M_{TPS}, M_{EDS}, k1, k2$); Step 3;
/* Producing G^{AD}_{syn} from G^{HC}_{avg} by iteratively add/delete edges to/from G^{HC}_{avg} using P_{con_all} until reaches to N^{AD}_{add} or N^{AD}_{del} */
$c_1 \leftarrow 0, c_2 \leftarrow 0$; Step 4;
while *NOT* ($c_1 \geq N^{AD}_{add}$ *AND* $c_2 \geq N^{AD}_{del}$) **do**
  $r \leftarrow$ genRandNum$(0, 1)$;
  if $r \geq 0.5$ **then**
    /* Find a node pair with no edge and maximum connection probability */
    $(u, v) \leftarrow$ getMaxConnNoEdgePair($G^{HC}_{avg}, P_{con\_all}$) ;
    /* add an edge between u and v */
    $G^{HC}_{avg}(V^{HC}, E^{HC}) \leftarrow G^{HC}_{avg}(V^{HC}, E^{HC} \cup (u, v))$;
    $c_1 \leftarrow c_1 + 1$;
  else
    /* Select an edge with minimum connection probability */
    $(u, v) \leftarrow$ getMinConnEdge($G^{HC}_{avg}, P_{con\_all}$) ;
    /* Delete the edge between (u, v) */
    $G^{HC}_{avg}(V^{HC}, E^{HC}) \leftarrow G^{HC}_{avg}(V^{HC}, E^{HC} - \{(u, v)\})$;
    $c_2 \leftarrow c_2 + 1$;
/* Return modified G^{HC}_{avg} (synthetic brain network) */
return G^{HC}_{avg}; Step 5;

space. Thus, we have ended up with a vector representation $\vec{u}_{EMB}$ of each node u of real brain networks G^{HC}_{avg} for HC. We have used cosine similarity as the topological similarity measure because of its robustness and prediction accuracy in many applications [28] which can be defined by: $EBS(u, v) = (\vec{u}_{EMB}.\vec{v}_{EMB})/(\|\vec{u}_{EMB}\| * \|\vec{v}_{EMB}\|)$, Where, "." denotes the dot product of two vectors. Finally, the proposed connection probability between two nodes u and v in a real brain network is expressed by the following equation:

$$P^{EMB}_{con}(u, v) = \left(\frac{\vec{u}_{EMB}.\vec{v}_{EMB}}{\|\vec{u}_{EMB}\| * \|\vec{v}_{EMB}\|}\right)^{k1} * EDS(u, v)^{k2} \quad (2)$$

The proposed method or the proposed connection probability, as defined above, varies by incorporating different embedding techniques, as previously discussed. We abbreviated the respective proposed variants as PV_i, $i \in \{1, 2, ..., 7\}$ corresponding to the different embedding methods, viz., DW, N2V, LINE, GCN, SAGE, ChebyNet, and GAT, respectively.

B. Performance evaluation: Modified Similarity Index (MSI)

We used Similarity Index (SI) [5] for measuring the performance evaluation of our study. We modified the original similarity index by adding an extra 1 to the denominator to avoid generation of zero in the denominator part, and called it MSI defined as: MSI = $1/(1 + (E_{LE} + E_{AC} + E_T + E_M + E_{GE} + E_R))$. We include six important topological properties for a reasonable and convincing evaluation. E_{LE}, E_{AC}, E_T, E_M, E_{GE}, E_R refer to the relative error between real and synthetic brain networks corresponding to the local efficiency, avg. clustering coefficient, transitivity, modularity, global efficiency and rich club coefficient respectively. If the value of MSI is large, then it is highly likely that the real-world network G^{AD}_{avg} is similar to G^{AD}_{syn} network.

Models	E_{LE}	E_{AC}	E_{GE}	E_M	E_T	E_R	MSI
CN	0.09	0.15	0.01	0.15	0.16	0.04	0.63
JI	0.10	0.15	0.05	0.17	0.0	0.06	0.65
PA	0.07	0.05	0.33	0.07	0.41	0.06	0.50
RA	0.01	0.15	0.03	0.13	0.06	0.04	0.70
AA	0.09	0.15	0.01	0.14	0.03	0.04	0.68
PV_1	0.01	0.02	0.02	0.03	0.02	0.03	0.88
PV_2	0.01	0.01	0.01	0.02	0.02	0.02	0.92
PV_3	0.25	0.20	0.22	0.19	0.08	0.03	0.51
PV_4	0.05	0.05	0.04	0.01	0.01	0.06	0.82
PV_5	0.02	0.03	0.02	0.01	0.01	0.04	0.88
PV_6	0.01	0.01	0.01	0.05	0.01	0.11	0.83
PV_7	0.01	0.0	0.0	0.02	0.02	0.05	0.90
Random	0.31	0.25	0.16	0.19	0.12	0.06	0.47

TABLE I

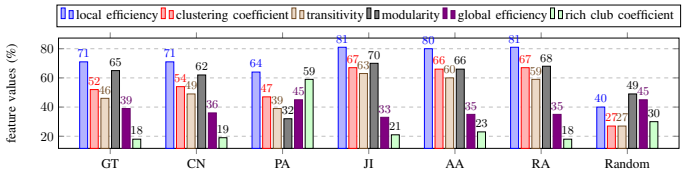
COMPARISON OF MSI-VALUES FOR ALL THE MODELS (BASELINES AND OUR PROPOSED MODEL). LARGER MSI VALUE INDICATE MORE SIMILARITY BETWEEN THE SYNTHETIC AD AND REAL AD BRAIN NETWORKS. RED, GREEN AND BLUE CELLS INDICATE THE WORST, BEST AND SECOND BEST PERFORMER RESPECTIVELY.

V. BASELINES AND HYPER-PARAMETERS

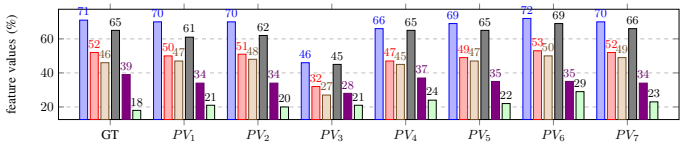
Baselines. We have used five baseline methods that use the same anatomical similarity (EDS) but different topological similarities [29], viz.Common Neighbour (CN), Jaccard Index (JI), Preferential Attachment (PA), Adamic–Adar (AA) while calculating the connection probabilities. In addition to these baselines, we have also included a random model where edges are successively added or deleted *randomly* unlike the others.
Hyper-parameters. For all the models, We set the values of $k1$ and $k2$ to 20 and 8, respectively. For the random walk-based models, we keep the default values for their respective hyperparameters. For all the GNN, we set the number of layers to 2, learning rate to 0.001 and epochs to 200.

VI. RESULTS AND FINDINGS

Comparison of G^{AD}_{avg} and G^{AD}_{syn} using network properties. Fig. 2a shows the bar-plot of six important network properties for different competitive models. Subsequently the bar-plot of the generated topological features using the proposed variants is depicted in Fig. 2b. The leftmost group of bars in Fig. 2a and 2b represent the ground truth. It is clear from Fig. 2b that in most of the cases, the proposed variants (viz. PV_2 (N2V) and PV_7 (GAT)) produce similar topological features as observed in the ground truth. Furthermore, $Random$ is the worst performer.
Comparison using MSI. Table I shows the values of relative errors of the respective topological properties between the G^{AD}_{syn} and G^{AD}_{avg} networks as well as the MSI values. It is easy to see that our proposed variants have the highest MSI scores and the random model has the lowest MSI scores.
Comparison using degree distribution. The experimental result given in Fig. 3 shows that the degree distribution of the synthetic networks generated through the proposed variants (Fig. 3b) closely resembles the degree distribution of the target networks compared to the other baseline models (Fig. 3a).

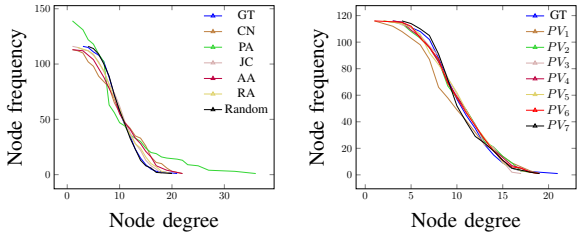


(a) real AD brain network (leftmost) vs synthetic AD brain networks produced by the baselines.



(b) real AD brain network (leftmost) vs synthetic AD brain networks produced by the proposed variants.

Fig. 2. Comparison of topological properties: real AD brain network (leftmost) vs synthetic AD brain networks produced by the various models.



(a) Degree distribution of real and synthetic brain networks for AD group: GT vs baselines

(b) Degree distribution of real and synthetic brain networks for AD group: GT vs proposed variants

Fig. 3. Comparing the complementary cumulative degree distribution of networks generated through various generative models.

VII. Conclusion and Future Work

We have proposed a generative model (and its variants) that combines structural and functional similarity to capture the topological differences of the brain network in the progression from HC to AD. In the future, we aim to investigate the progression from HC to Mild Cognitive Impairment (MCI), which is a stage in between HC and AD and also the changes in the connectivity patterns in various sub-regions of the brain. **Reproducibility.** The code can be found at https://github.com/anjangit000/ADProgression.

References

[1] A. Association *et al.*, "2018 alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 14, no. 3, pp. 367–429, 2018.

[2] B. M. Tijms, A. M. Wink, W. de Haan, W. M. van der Flier, C. J. Stam, P. Scheltens, and F. Barkhof, "Alzheimer's disease: connecting findings from graph theoretical studies of brain networks," *Neurobiology of aging*, vol. 34, no. 8, pp. 2023–2036, 2013.

[3] N. T. Doan, A. Engvig, K. Persson, D. Alnæs, T. Kaufmann, J. Rokicki, A. Córdova-Palomera, T. Moberget, A. Brækhus, M. L. Barca *et al.*, "Dissociable diffusion mri patterns of white matter microstructure and connectivity in alzheimer's disease spectrum," *Scientific reports*, vol. 7, no. 1, pp. 1–12, 2017.

[4] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature reviews neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[5] R. F. Betzel, A. Avena-Koenigsberger, J. Goñi, Y. He, M. A. De Reus, A. Griffa, P. E. Vértes, B. Mišic, J.-P. Thiran, P. Hagmann *et al.*, "Generative models of the human connectome," *Neuroimage*, 2016.

[6] C. V. Cannistraci, G. Alanis-Lobato, and T. Ravasi, "From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks," *Scientific reports*, 2013.

[7] M. Jalili, Y. Orouskhani, M. Asgari, N. Alipourfard, and M. Perc, "Link prediction in multiplex online social networks," *Royal Society open science*, vol. 4, no. 2, p. 160863, 2017.

[8] M. E. Newman, A.-L. E. Barabási, and D. J. Watts, "The structure and dynamics of networks." 2006.

[9] D. S. Bassett and O. Sporns, "Network neuroscience," *Nature neuroscience*, vol. 20, no. 3, pp. 353–364, 2017.

[10] K. Hahn, N. Myers, S. Prigarin, K. Rodenacker, A. Kurz, H. Förstl, C. Zimmer, A. M. Wohlschläger, and C. Sorg, "Selectively and progressively disrupted structural connectivity of functional brain networks in alzheimer's disease—revealed by a novel framework to analyze edge distributions of networks detecting disruptions with strong statistical evidence," *Neuroimage*, vol. 81, pp. 96–109, 2013.

[11] M. Mapstone, A. K. Cheema, M. S. Fiandaca, X. Zhong, T. R. Mhyre, L. H. MacArthur, W. J. Hall, S. G. Fisher, D. R. Peterson, J. M. Haley *et al.*, "Plasma phospholipids identify antecedent memory impairment in older adults," *Nature medicine*, vol. 20, no. 4, pp. 415–418, 2014.

[12] C.-Y. Lo, P.-N. Wang, K.-H. Chou, J. Wang, Y. He, and C.-P. Lin, "Diffusion tensor tractography reveals abnormal topological organization in structural cortical networks in alzheimer's disease," *Journal of Neuroscience*, vol. 30, no. 50, pp. 16 876–16 885, 2010.

[13] R. Sharma, T. Goel, M. Tanveer, C. Lin, and R. Murugan, "Deep learning based diagnosis and prognosis of alzheimer's disease: A comprehensive review," *IEEE Transactions on CDS*, 2023.

[14] P. Park, Byeon, "Funp (fusion of neuroimaging preprocessing) pipelines: a fully automated preprocessing software for functional magnetic resonance imaging," *Frontiers in neuroinformatics*, 2019.

[15] C. Hutton, A. Bork, O. Josephs, R. Deichmann, J. Ashburner, and R. Turner, "Image distortion correction in fmri: a quantitative evaluation," *Neuroimage*, vol. 16, no. 1, pp. 217–240, 2002.

[16] R. Sladky, K. J. Friston, J. Tröstl, R. Cunnington, E. Moser, and C. Windischberger, "Slice-timing effects and their correction in functional mri," *Neuroimage*, vol. 58, no. 2, pp. 588–594, 2011.

[17] J. D. Power, B. L. Schlaggar, and S. E. Petersen, "Recent progress and outstanding issues in motion correction in resting state fmri," *Neuroimage*, vol. 105, pp. 536–551, 2015.

[18] J. Ashburner and K. J. Friston, "Nonlinear spatial normalization using basis functions," *Human brain mapping*, 1999.

[19] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain," *Neuroimage*, 2002.

[20] C. C. Kaiser, Marcus Hilgetag, "Modelling the development of cortical systems networks," *Neurocomputing*, vol. 58, pp. 297–302, 2004.

[21] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *ACM SIGKDD*, 2014, pp. 701–710.

[22] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *ACM SIGKDD*, 2016, pp. 855–864.

[23] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei, "Line: Large-scale information network embedding," in *WWW*, 2015.

[24] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *ICLR*, 2017.

[25] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 1025–1035.

[26] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," *Advances in neural information processing systems*, vol. 29, 2016.

[27] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio *et al.*, "Graph attention networks," *stat*, 2017.

[28] M. Gulistan, M. Mohammad, F. Karaaslan, S. Kadry, S. Khan, and H. A. Wahab, "Neutrosophic cubic heronian mean operators with applications in multiple attribute group decision-making using cosine similarity functions," *International Journal of Distributed Sensor Networks*, 2019.

[29] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," *Physica A: statistical mechanics and its applications*, vol. 390, no. 6, pp. 1150–1170, 2011.