Identification of Alzheimer's Disease Hub Genes Based on Improved HITS Algorithm

Rasha mahdi abd ul kader
Department of Computer Science, College of Basic Education, University of Diyala, Diyala, Iraq
rashamahdi@uodiyala.edu.iq
Orcid.org/0009-0006-1333-8934

Atyaf ismaeel Altameemi
Department of Computer Science, College of Education for Pure Sciences, University of Diyala, Diyala, Iraq
dr.atyaf.ismaeel@uodiyala.edu.iq
Orcid.org/0000-0003-0619-1592

Rana jassim mohammed
Department of Computer Science, Presidency of Diyala University, Diyala, Iraq
ad.ranamohammed@uodiyala.edu.iq
Orcid.org/0000-0002-2080-9670

Qusay Kanaan Kadhim
Department of Computer Science, College of Science, University of Diyala, Diyala, Iraq
dr.qusay.kanaan@uodiyala.edu.iq
Orcid.org/0000-0003-2814-2409

DOI: http://dx.doi.org/10.31642/JoKMC/2018/110105

Received Aug. 30, 2023. Accepted for publication Nov. 2, 2023

Abstract—Alzheimer's disease is a severe, neurodegenerative condition that gradually breaks memories, thinking abilities, and the ability to carry out even the most basic tasks. The hub genes of AD were examined in this study. They understand how interactions between proteins and non-protein substances are crucial to understanding how proteins work. Network investigations of protein-protein interactions, in particular, help understand biological issues. This article offers a novel approach to identifying essential proteins using weighted PPI networks and Hyperlink-Induced Topic Search (HITS) algorithm. We discovered the top 10 hub genes linked to AD using a protein network analysis: AKT1, TGFB1, GRB2, NFKB1, PIK3CA, PIK3R1, TNF, IFNG, VEGFA, and TP53. It was discovered by gene enrichment that most gene activities might be categorized as vital to the plasma membrane, including engagement in signaling cascades, G-protein composite reliability activation, and cell contact. The prioritized genes were determined by the convergent functional genomics ranking AKT1, TGFB1, GRB2, NFKB1, PIK3CA, PIK3R1, TNF, IFNG, VEGFA, and TP53. To better understand AD pathophysiology and find new biomarkers or medication targets for AD treatment, these molecular pathways hub genes will be helpful.

Keywords—HITS algorithm, PPI networks, Hub genes, Alzheimer's disease, Functional enrichment

I. INTRODUCTION

The majority of instances of dementia related to aging—50%–60%—are caused by Alzheimer's Disease (AD), which is characterized by a progressive memory loss coupled with additional cognitive deficiencies in judgment, abstraction, language, attention, and visuoconstructive ability [1]. Around the world, AD affects 36 million people, and by 2050, it's predicted that 110 million people will be affected [2,3] Therefore, AD is incurable, and current knowledge of its pathogenic processes is incomplete. Later, the same team used several bioinformatics techniques to evaluate independent data and found that most of those 18 proteins were associated with the concentrations of A or tau proteins in CSF [4,5] Since the publication of these two papers, several profiling strategies have suggested protein panels with significant diagnostic potential, but repeatability has remained the primary concern [6,7].Hu and colleagues have addressed the issue of reproducibility by identifying a number of utilizing two sizable, well-characterized clinical samples, we examined the causes of activation linked to the development of AD [8,9]. O'Bryant and colleagues [10,11]good diagnostic accuracy across cohorts was also noted. Cross-validation across different cohorts has also been employed by plasma proteomics researchers to solve the over-fitting issue in high dimensional investigations. It has been proposed it due to AD being a disease associated with protected system malfunction and mitochondrial dysfunction [12,13], concentrating on genes implicated in pertinent pathways could aid in the finding of biomarkers[14 ,15]. Only a small number of earlier investigations, though, have modeled using biological data. Various efforts have been made to far to predict hub proteins using Hub protein discovery methods based on biological experiments and networks. Despite the use of conventional experimental techniques like gene knockout

stages [16,17] , RNA interference and conditional knockouts [18,19], they take more time and money, yet they can predict hub proteins accurately. Technologies with high throughput, like the yeast two-hybrid approach, have been developed [20], mass spectrometry analysis [21,22]. Based on existing PPI data, some academics have proposed various computational ways to overcome these experimental limits. According to certain research, the centrality-lethality rule—which states that essential proteins in PPI networks tend to be strongly connected proteins [23] . The PPI networks can entirely breakdown in the absence of highly linked protein nodes, which would be fatal to the organism itself. There are several different network centrality measurements, like Degree Centrality (DC) [24], Neighborhood Centrality (NC), Information Centrality (IC), Eigenvector Centrality (EC), Betweenness' Centrality (BC), Closeness Centrality (CC), Sub-graph Centrality (SC), and Local Average Connectivity (LAC)[25] . To determine essential proteins, certain centrality methods are applied because of these research' findings, but they have some limitations because there are many false positive and false negative outcomes with PPI data. As a result, numerous techniques for detecting hub proteins have been suggested.

It was first introduced by Kleinberg in 1998, and is known as "hypertext induced topic search" (HITS) in the field of web structure mining[26]. Prior to joining them in the link structure, Kleinberg separated authority pages and hub pages from network pages. The former offers the best information on search-related issues; the more network pages cite it, the greater its authority value. The latter offers significant linkages and the more authoritative pages it cites, the greater its hub value. Web searches frequently use the HITS algorithm, which effectively addresses various real-world including web community[26].An innovative computational technique is presented in this study, to identify essential proteins using weighted PPI networks and the HITS algorithm. First, we create a directed network from the original undirected PPI network. After that, we weighted PPI networks using biological data and network topological properties and assess three factors: protein functions, false positives and false negatives, and protein locations. The biological information employed in this technique includes data on gene expression, and Gene Ontology (GO) annotation. As an illustration understanding PPI networks' topological characteristics [27]. Then, using the authority and hub principles produced using HITS method, the rank the proteins. This research used protein-protein interactions to discover the hub genes responsible for Alzheimer's disease and HITS algorithm. Finally, we obtained the top 10 genes related to AD.

II. MATERIALS AND METHODS

A proposal is used the dataset of PPI network of human is integrated from HPRD (the Human Protein Reference Database), have 37437 genes. The datasets GWAS (Genome-wide association study) downloaded the GWAS association results for AD [26]. The GWAS have 668 genes involved in AD.Methods of detailed description of all the algorithms, formulas and methods that helped to present proposal used the HITS algorithm and extra-biological information to identify and evaluate hub genes, as shown in Figure (1). It consists of three stages:

- Constructing Weighted Protein-Protein Interaction Network
- Hypertext Induced Topic Search **(**HITS) Algorithm
- Evaluating Identifying Hub Gene based on HITS Algorithm

A. Constructing Weighted Protein-Protein Interaction Network

The most common way to depict a protein-protein interaction network is as an undirected graph G = (V, E), where V is a set of vertices that stands for proteins and E is a collection of all interactions between proteins[26]. We start from the assumption that protein connections are interactive and transform the undirected PPI network G = (V, E) into a bidirectional network G'= (V, E') that is similar to it in order to defy accepted wisdom. Mentioning certain networks in biology, like kinase networks, are exempt from the mathematical process that converts an undirected graph into a directed graph. The high number of false positives in high-throughput PPI networks, and false negatives will affect the prediction accuracy. In order to resolve this issue, we weigh edges separately using network topological traits and biological information. Nodes with high-quality biological data are anticipated to be directed to using nodes that indicate higher topological data that vice versa, based on the HITS algorithm. The building of the weighted PPI network is illustrated with an example in Figure (2).

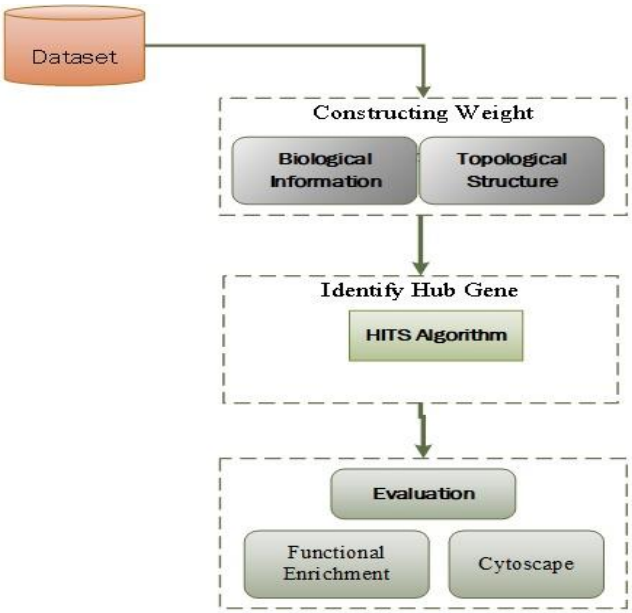


Figure 1. Block diagram of the proposal for the prediction of Alzheimer's diseases

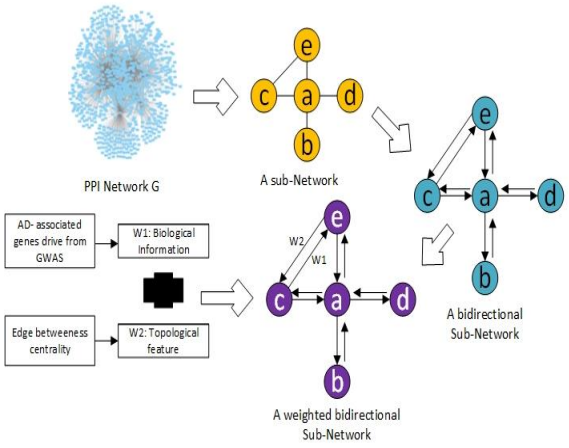


Figure 2. An example of a weighted PPI network

Network topology weighted edge: Edge Betweenness Centrality (EBC) is a method of measuring how much effect a node has on the information flow in a graph. It is frequently used to identify nodes that connect two distinct regions of a graph. The number of times a node is observed to be located along the shortest path connecting two other nodes, in addition to betweenness, is a measure of centrality [28]. It is equal to the total number of shortest routes between every vertex and every other vertex that go via that node. If we assume that the shortest paths are used for item transfer, a node with high betweenness centrality has a significant impact on the transmission of items through the network. The quickest path is taken by the brain to digest information in order to conserve time and energy, according to studies. In order to reflect the shortest paths in our model, we chose proximity and betweenness centrality as measurements [28]. Betweenness Centrality CB for the graph G = (V, E) (v) is given in Equation (1). EBC (^{V}) can be defined ,as follows [28]:

$$EBC(^{V}) = \sum_{s \neq v \neq t \neq} \frac{\sigma st(v)}{\sigma st} \qquad (1)$$

Where σst is the total number of shortest paths from node s to node t and σst (^{V}) is the number of those paths that pass through ^{V},as shown in Figure (3).

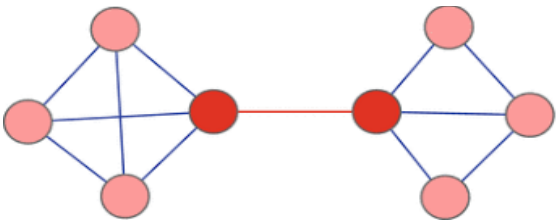
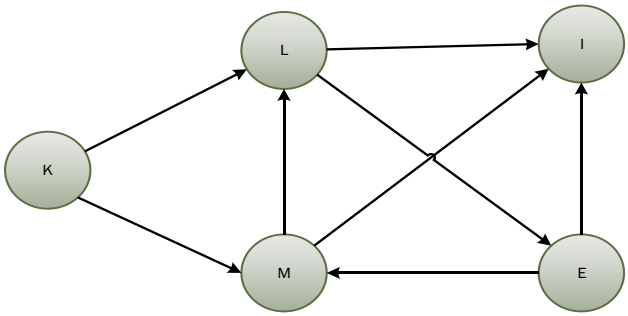


Figure 3. Betweenness centrality [28]

Biological information weighted edge: Several genes linked to AD were found via Genome-Wide Association Studies (GWASs). The biological pathway is a method using which a gene's information is utilized in a creation from a useful gene produce.

B. HITS algorithm

The Hyperlink- Induced Topic Search algorithm, an iterative algorithm, was first suggested to evaluate a significance of online pages [26]. The HITS algorithm rates the web page via analysing all of its in-links and out-links, which is reliant on the search query. Each page receives two attributes in the HITS algorithm: the hub and the authority. Many high-quality hub pages will link to the authority, which is a high-quality authority page. The values of each page's authority that the page hub links to together make up its value [29] . Hub is a top-notch hub page that links to numerous top-notch authority pages. The total hub values that point to the page make up the page authority value. Illustrates how to determine the hub and authority's value, as shown in Figure (4).



a(L)=h(K) +h(M) a(M)=h(K)+h(E)

h(L)=a(I)+a(E) h(M)=a(L)+a(I)

Figure 4. An example of how to calculate hub and authority values simply [26]

The Hub (h) score and Authority (a) score for a node is calculated the technique below[26]:

1. Begin with a hub score and authority score of 1 for each node
2. Implement the Authority Updating Rule

$$at(v) = \sum(u, v) \in E\ ht\text{-}1(u) \qquad (2)$$

3. Run the Hub Update Rule

$$ht(v) = \sum(u, v) \in E\ at\text{-}1\ (u) \qquad (3)$$

4. To normalize the numbers, divide the values by the square roots of the sum of the squares of all Hub scores and all Authority scores
5. Continue as necessary from steps two and three

C. Identifying Hub Gene based on HITS Algorithm

The suggested HITS approach is applied, and weighted PPI networks are constructed. The ultimate score is calculated by combining the authority value and the hub value via Eq. (4). The notice that in every change in weight some common genes appear and new genes also appear, this lead to conclusion that it is better to make the value α half to each of the h=0.5 and a=0.5 to make the network balanced, as presented in the Table

(1), to fully assess each gene's significance, where [0, 1] utilized to modify a ratio (α) between these two marks. Only when the value of is equal to 0 does the sorting score depend on the topological data. A sorting score is calculated using a biological data if the value of is 1. The definition of gene scores (v) states that they anticipate that various parameters will have an impact on its performance α. To reduce the selection pressure of the parameter and make it easier to apply gene scores to various organisms to discover hub genes. The balance of the effect is set to 0.5. The proposed approach described in pseudo-code as two phases in Algorithm (1). The first stage uses edge betweenness centrality to weigh PPI networks with biological documents and topological features. The HITS method is used in the second stage to find the hub gene, which is based on Eq. (4).

$$\text{Gene Score } (v) = \alpha \times a\,(v) + (1 - \alpha) \times h\,(v) \qquad (4)$$

Where $\alpha \in [0, 1]$ is employed for modify a ratio of these two scores, where ν represents a node. Algorithm to weighted PPI networks [30].

Algorithm (1): HITS algorithm

Input: DG: Directed Weighted Graph, G: Network Graph, V: Nodes, E: Edge, h: hub, a: authority, t= no of iteration
Output: Identify hub genes related to Alzheimer's disease
Step1
1: Transform G to Bidirectional Graph G` (V, E`)
2: for each interrelating protein pair (a, b) in PPI do
3: Compute EBC /* Find shortest path between of two nodes*/
4: Calculate the ratio of related gene to disease (DG) // 0 if no gene related to disease, 1 if two genes related to disease, 0.5 if one gene related to disease //
5: end for
6: for each interacting protein pair (a,b) in G` do
7: edge (a,b) = EBC (a,b)
8: edge (b,a) =DG (b,a)
9: end for
Step2
10: for t in [1, max_iter] do
11: for each v in V do
12: Run the Authority Update Rule by using Eq.(2)
13: Run the hub Update Rule by using Eq.(3)
14: $at = \frac{a_t}{\max(a_t)}$
15: $ht = \frac{h_t}{\max(h_t)}$
16: t = t + 1
17: until $|at - a\ t\text{-}1| + |ht - ht{-}1| < \gamma$ // γ: Denote collect the highest-ranked pages // from a text-based search
18: return (at, ht)
19: end for
20: end for
21: Calculate the score of each gene via Eq.(4)
22: Return hub genes

III. RESULTS AND DISCUSSION

The HITS algorithm's effectiveness needed to be demonstrated, the performance of the algorithm was evaluated by changing the values (α) .

That suggests that the HITS algorithm was successful in locating hub genes, it determined the top ten hub genes linked to AD: AKT1, TGFB1, GRB2, NFKB1, PIK3CA, PIK3R1, TNF, IFNG, VEGFA and TP53.

TABLE I. SHOWS THE WEIGHT FOR H AND A

h Ratio	a Ratio	Top 10 Genes
0.5	0.5	AKT1,TGFB1,GRB2,NFKB1,PIK3CA,PIK3R1,TNF,IFNG,VEGFA,TP53
0.6	0.4	AKT1,TGFB1,NFKB1,GRB2,IFNG,TNF,PIK3CA,PIK3R1,TP53,VEGFA
0.4	0.6	PIK3R1,TGFB1,AKT1,PIK3CA,NFKB1,TNF,VEGFA,GRB2,IFNG,TP53
0.7	0.3	AKT1,TGFB1,IFNG,NFKB1,GRB2,TP53,TNF,VEGFA,PIK3CA,IL1B
0.3	0.7	PIK3R1,PIK3CA,TGFB1,TNF,AKT1,VEGFA,GRB2,NFKB1,IGF1R,IL6
0.8	0.2	AKT1,TGFB1,NFKB1,IFNG,TP53,GRB2,TNF,VEGFA,TRAF6,FYN
0.2	0.8	PIK3R1,PIK3CA,TGFB1,TNF,VEGFA,GRB2,NFKB1,AKT1,CAV1,IGF1R
0.0	1.0	PIK3R1,PIK3CA,TNF,CAV1,IRS1,PDGFRB,CXCL8,ESR1,IGF1R,IL6
1.0	0.0	AKT1,TGFB1,IFNG,NFKB1,TP53,GRB2,TRAF6, VEGFA,TNF,FYN
0.9	0.1	AKT1,TGFB1,NFKB1,IFNG,TP53,GRB2,VEGFA,TNF, TRAF6 ,FYN
0.1	0.9	PIK3R1,PIK3CA,TNF,TGFB1,CAV1,VEGFA,CXCL8,ESR1,IGF1R,IL6

A. *Functional Enrichment Analysis*

The FunRich analysis is a predominant bioinformatics tool for annotations of genes and their products, in term Biological Pathways (BP).In general, the tool includes 200 biological pathways, and when the 10hub genes are entered into the tool to be analyzed in the biological pathways the results appeared 33 biological pathways containing 100% of the ten hub genes, as shown in Figure (5).

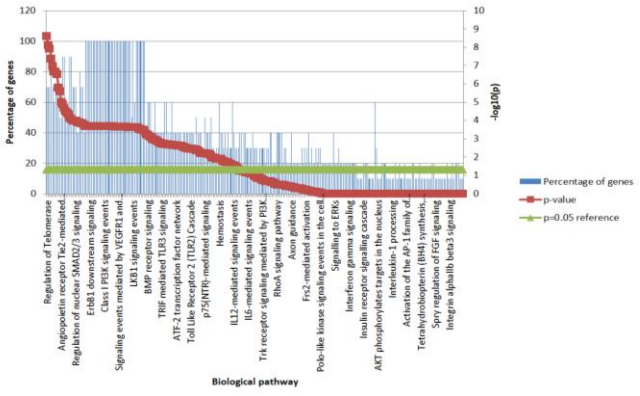


Figure 5. Hub genes in the biological pathways of Alzheimer's disease

B. Prioritizing Hub Genes

Using the CFG technique, which took into account several types of evidence connected to AD, the probable driver genes were ranked from AD-specific modules [31] .

A CFG scores ranged starting zero to five, with five denoting the utmost importance. Five evidences for AD were found: 1) Genetic correlation A gene received one point if it possessed at most one locus that was expressively related with Alzheimer's according to the International Genomics of Alzheimer's Project's statistical data in overview [IGAP]; otherwise, it received a zero value. 2) Genetics controls gene expression a gene received one point if it was linked to Expression Quantitative Trait Loci (eQTLs) in IGAP data that indicated an AD-risk; otherwise, it received zero points. 3) Protein interaction one point was if a gene made a physical connection to any of the Alzheimer disease genes known. (APP, PSEN1, PSEN2, APOE, or MAPT); otherwise, zero points were given. 4) Interaction of proteins if any of the AD core genes were physiologically interacting with a gene (APP, PSEN1, PSEN2, APOE, or MAPT). 5) Mouse brain with early signs of AD change. When a gene's expression in the hippocampus of 2-month-old AD mice differed from that of one point was awarded for age-matched wild-type mice; otherwise, 0 points were given, as shown in Table (2).

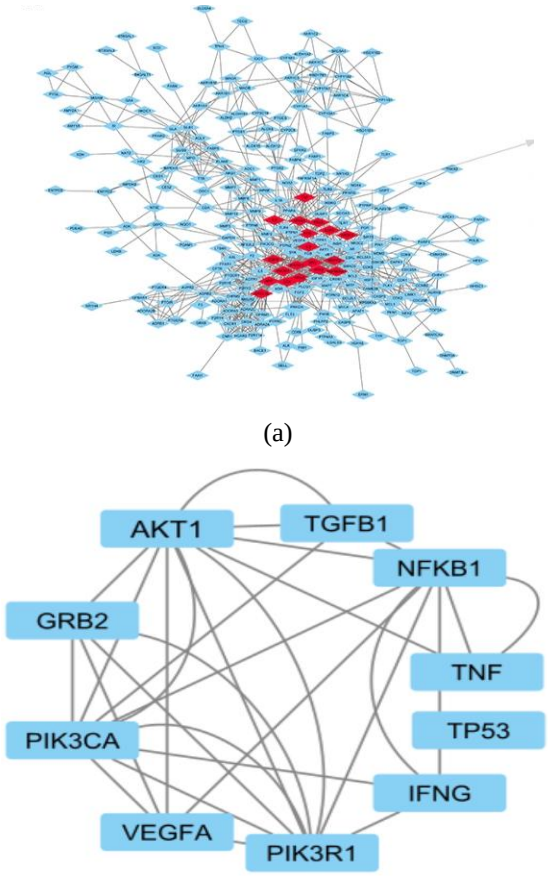
TABLE II. THE CFG APPROACH ASSIGNS THE 10 IDENTIFIED CANDIDATE GENES A PRIORITY

Gene	eQTL	GWAS	PPI	Early_DEG	Pathology cor (Aβ)	Pathology cor (Tau)	CFG
AKT1	0	0	PSEN2,MAPT, APOE	NA	-0.223,ns	-0.266,ns	1
TGFB1	1	0	APP,MAPT, APOE	NA	0.871,***	0.681,**	3
GRB2	5	0	APP,PSEN1, PSEN2,MAPT, APOE	NA	0.110,ns	0.138,ns	2
NFKB1	1	0	APP,PSEN2, MAPT	NA	0.731,***	0.247,ns	3
PIK3CA	2	0	APP,PSEN1, PSEN2,MAPT, APOE	NA	-0.170,ns	-0.485,ns	2
PIK3R1	2	0	PSEN1,PSEN2, MAPT,APOE	NA	0.005,ns	-0.141,ns	2
TNF	NA	0	-	NA	NA	NA	0
IFNG	NA	0	MAPT	NA	NA	NA	1
VEGFA	2	0	APP,PSEN2, APOE	yes	-0.215,ns	-0.758,**	4
TP53	1	0	APP,PSEN1, PSEN2	NA	NA	NA	2

*C. **Protein–protein Interaction Network (PPI)***

The PPI network was visualized using the Cytoscape software. The PPI network contained 5664 nodes and 37219 edges, as shown in Figure (6). In all, the 10 hub genes with strong connections were AKT1, TGFB1, GRB2, NFKB1, PIK3CA, PIK3R1, TNF, IFNG, VEGFA, and TP53.

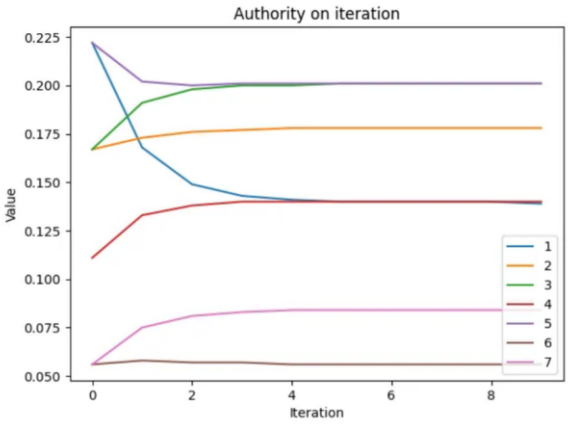
Now we all knew that after enough iteration, hub and authority would always converge to a specific value, as shown in Figure (7).



(a)

(b)

Figure 6. (a) An Overview of the PPI. The network has 3397 nodes and 37219 edges. The hub genes, which are displayed as large blue nodes, are interacted with other genes linked to Alzheimer's disease by the red nodes. (b) top 10 hub genes.



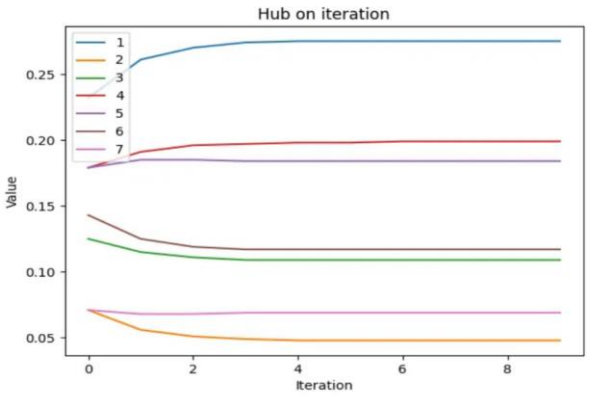


Figure 7. Performance Measurement for convergence

In order to identify the relationship between total edges and computation time, we perform 100 iterations with a varied number of total edges. You can see that the relationship between the inference of edges and computing time is almost linear, which is really good, as shown in Figure (8).

The edges' connections to one another, which are why it isn't entirely linear, will also, have a small impact on how long it takes to compute. For example, 500 iterations are still not enough. The computation time will also be slightly affected by the connections between the edges, which are why it is not totally linear.

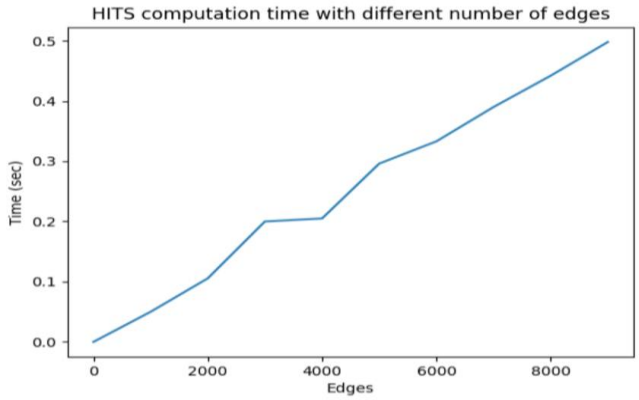


Figure 8. Measure the HITS computation

IV. CONCLUSION

Understanding the basic processes that underlie cellular life is significant for identifying key genes. This article introduced a novel computer method to determine important proteins via a weighted PPI network and the HITS algorithms.

A weighted PPI network is used to discover essential proteins using biomedical information and network topology. DEGs with up-regulation showed a strong link with osteoclast differentiation, lysosome, neutrophil degranulation, and immune response.

In contrast, DEGs showed a major connection to the ribosome, translation, peptide production, and oxidative phosphorylation through down-regulation. Consequently, we discovered AKT1, TGFB1, GRB2, NFKB1, PIK3CA, PIK3R1, TNF, IFNG, VEGFA, and TP53as essential genes in AD determined by PPI network analysis. This HITS method is used in this study for extensive bioinformatics investigation to deepen our understanding of the molecular basis of AD. It may provide a hint for developing a medication for AD patients.

REFERENCES

[1] H. Gui, Q. Gong, J. Jiang, M. Liu, and H. Li, "Identification of the Hub Genes in Alzheimer's Disease," Comput. Math. Methods Med., vol. 2021, doi: 10.1155/2021/6329041.

[2] Z. Han, H. Huang, Y. Gao, and Q. Huang, "Functional annotation of Alzheimer's disease associated loci revealed by GWASs," PLoS One, vol. 12, no. 6, 2017, doi: 10.1371/journal.pone.0179677.

[3] S. Taha Ahmed and S. Malallah Kadhem, "Using Machine Learning via Deep Learning Algorithms to Diagnose the Lung Disease Based on Chest Imaging: A Survey," Int. J. Interact. Mob. Technol., vol. 15, no. 16, p. 95, 2021, doi: 10.3991/ijim.v15i16.24191.

[4] J. J. Wee and S. Kumar, "Prediction of hub genes of alzheimer's disease using a protein interaction network and functional enrichment analysis," Genomics and Informatics, vol. 18, no. 4, pp. 1–8, 2020, doi: 10.5808/GI.2020.18.4.E39.

[5] S. T. Ahmed and S. M. Kadhem, "Early Alzheimer ' s Disease Detection Using Different Techniques Based on Microarray Data : A Review," iJOE – Vol. 18, No. 04, 2022, no. Mci.

[6] H. Gao, Y. Tao, Q. He, F. Song, and D. Saffen, "Functional enrichment analysis of three Alzheimer's disease genome-wide association studies identities DAB1 as a novel candidate liability/protective gene," Biochem. Biophys. Res. Commun., vol. 463, no. 4, pp. 490–495, 2015, doi: 10.1016/j.bbrc.2015.05.044.

[7] M. S. Croock, S. D. Khudhur, and A. K. Taqi, "Edge Detection and Features Extraction for Dental X-Ray," Eng. Tech.Journal, vol. 34, no. 13, pp. 2420–2432.

[8] N. Mahendran, P. M. Durai Raj Vincent, K. Srinivasan, and C. Y. Chang, "Machine Learning Based Computational Gene Selection Models: A Survey, Performance Evaluation, Open Issues, and Future Research Directions," Front. Genet., vol. 11, no. December, pp. 1–25, 2020, doi: 10.3389/fgene.2020.603808.

[9] I. S. Abed, "Lung Cancer Detection from X-ray images by combined Backpropagation Neural Network and PCA," Eng. Technol. J., vol. 37, no. 5A, pp. 166–171, 2021, doi: 10.30684/etj.37.5a.3.

[10] J. Long, G. Pan, E. Ifeachor, R. Belshaw, and X. Li, "Discovery of Novel Biomarkers for Alzheimer's Disease from Blood," Dis. Markers, vol. 2016, 2016, doi: 10.1155/2016/4250480.

[11] H. A. R. Akkar and S. A. Salman, "Detection of Biomedical Images by Using Bio-inspired Artificial Intelligent," Eng. Technol. J., vol. 38, no. 2A, pp. 255–264, 2020, doi: 10.30684/etj.v38i2a.319.

[12] S. Perera, K. Hewage, C. Gunarathne, R. Navarathna, D. Herath, and R. G. Ragel, "Detection of Novel Biomarker Genes of Alzheimer's Disease Using Gene Expression Data," MERCon 2020 - 6th Int. Multidiscip. Moratuwa Eng. Res. Conf. Proc., pp. 1–6, 2020, doi: 10.1109/MERCon50084.2020.9185336.

[13] Matheel Emaduldeen Abdulmunim Zainab Fawzy Abed, "Classification Mammogram Images Using ID3 decision tree Algorithm Based on Contourlet Transform," Eng. &Tech.Journal, vol. 33, no. 3, p. 2015, 2015, [Online]. Available: https://www.iasj.net/iasj/article/105290.

[14] M. Hu and J. Wang, "Identification of Hub Genes and Immune Cell Infiltration Characteristics in Alzheimer's Disease," J. Healthc. Eng., vol. 2021, 2021, doi: 10.1155/2021/7036194.

[15] S. T. Ahmed, Q. K. Kadhim, H. S. M. Alsultani, and W. S. Abd almahdy, "Applying the MCMSI for Online Educational Systems Using the Two-Factor Authentication," Int. J. Interact. Mob. Technol., vol. 15, no. 13, p. 162, 2021, doi: 10.3991/ijim.v15i13.23227.

[16] X. Pang et al., "The Bioinformatic Analysis of the Dysregulated Genes and MicroRNAs in Entorhinal Cortex, Hippocampus, and Blood for Alzheimer's Disease," Biomed Res. Int., vol. 2017, 2017, doi: 10.1155/2017/9084507.

[17] M. K. Kadhim, A. K. A. Hassan, and M. K. Kadhim, "Proposed intelligent system using hybrid material techniques: A review," AIP Conf. Proc., vol. 2235, no. May, 2020, doi: 10.1063/5.0007485.

[18] L. Wu, W. Wang, S. Tian, H. Zheng, P. Liu, and W. Wu, "Identification of Hub Genes in Patients with Alzheimer Disease and Obstructive Sleep Apnea Syndrome Using Integrated Bioinformatics Analysis," Int. J. Gen. Med., vol. 14, no. November, pp. 9491–9502, 2021, doi: 10.2147/IJGM.S341078.

[19] H. A. Abdulmohsin, H. Bahjat, A. Wahab, A. Mohssen, and J. Abdul, "A New Hybrid Feature Selection Method Using T-test and Fitness Function," Comput. Mater. Contin., 2021, doi: 10.32604/cmc.2021.014840.

[20] T. Yan, F. Ding, and Y. Zhao, "Integrated identification of key genes and pathways in alzheimer's disease via comprehensive bioinformatical analyses," Hereditas, vol. 156, no. 1, pp. 1–13, 2019, doi: 10.1186/s41065-019-0101-0.

[21] T. W. Rowe, I. K. Katzourou, J. O. Stevenson-Hoare, M. R. Bracher-Smith, D. K. Ivanov, and V. Escott-Price, "Machine learning for the life-time risk prediction of Alzheimer's disease: a systematic review," Brain Commun., vol. 3, no. 4, 2021, doi: 10.1093/braincomms/fcab246.

[22] A. K. Hassan and S. N. Mohammed, "A novel facial emotion recognition scheme based on graph mining," Def. Technol., vol. 16, no. 5, pp. 1062–1072, 2020, doi: 10.1016/j.dt.2019.12.006.

[23] A. El-Gawady, M. A. Makhlouf, B. S. Tawfik, and H. Nassar, "Machine Learning Framework for the Prediction of Alzheimer's Disease Using Gene Expression Data Based on Efficient Gene Selection," Symmetry (Basel)., vol. 14, no. 3, 2022, doi: 10.3390/sym14030491.

[24] A. Sharma and P. Dey, "A machine learning approach to unmask novel gene signatures and prediction of Alzheimer's disease within different brain regions," Genomics, vol. 113, no. 4, pp. 1778–1789, 2021, doi: 10.1016/j.ygeno.2021.04.028.

[25] D. A. Salem, R. A. A. A. A. Seoud, and H. A. Ali, "A New Gene Selection Technique Based on Hybrid Methods for Cancer Classification Using Microarrays," Int. J. Biosci. Biochem. Bioinforma., vol. 1, no. 4, pp. 261–266, 2011, doi: 10.7763/ijbbb.2011.v1.49.

[26] S. W. 1 and F. W. Xiujuan Lei, "Identification of essential proteins based on improved HITS algorithm," Genes (Basel)., vol. 10, no. 2, 2019, doi: 10.3390/genes10020177.

[27] S. T. Ahmed and S. M. Kadhem, "Early Alzheimer's disease detection using different techniques based on microarray data: A review," Int. J. Online Biomed. Eng., vol. 18, no. 04, pp. 106–126, Mar. 2022, doi: 10.3991/ijoe.v18i04.27133.

[28] T. Inoha, K. Sadakane, Y. Uno, and Y. Yonebayashi, "Efficient Computation of Betweenness Centrality by Graph Decompositions and Their Applications to Real-World Networks," IEICE Trans. Inf. Syst., vol. E105D, no. 3, pp. 451–458, 2022, doi: 10.1587/transinf.2021FCP0003.

[29] S. T. Ahmed and S. M. Kadhem, "Optimizing Alzheimer ' s disease prediction using the nomadic people algorithm," Int. J. Electr. Comput. Eng., vol. 13, no. 2, pp. 2052–2067, 2023, doi: 10.11591/ijece.v13i2.pp2052-2067.

[30] D. Otasek, J. H. Morris, J. Bouças, A. R. Pico, and B. Demchak, "Cytoscape Automation: Empowering workflow-based network analysis," Genome Biol., vol. 20, no. 1, pp. 1–15, 2019, doi: 10.1186/s13059-019-1758-4.

[31] M. Ayalew et al., "Convergent functional genomics of schizophrenia: From comprehensive understanding to genetic risk prediction," Mol. Psychiatry, vol. 17, no. 9, pp. 887–905, 2012, doi: 10.1038/mp.2012.37.