ISSN 2278-3091
Volume 9, No.3, May - June 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse210932020.pdf
https://doi.org/10.30534/ijatcse/2020/210932020

Shortest Path of a Graph using Centrality Measures

G.Vadivu[1], C.Fancy[2], S.Sivasankari[3], K.Sornalakshmi[4]

[1] Professor, Department of Information Technology, SRM Institute of Science and Technology, India. vadivug@srmist.edu.in
[2]Assistant Professor, Department of Information Technology, SRM Institute of Science and Technology, India.fancyc@srmist.edu.in
[3]Assistant Professor, Department of Information Technology, SRM Institute of Science and Technology, India.sivasans2@srmist.edu.in
[4]Assistant Professor, Department of Information Technology, SRM Institute of Science and Technology, India.sornalak@srmist.edu.in

ABSTRACT

When it comes to finding the shortest path in a graph, most people think of Dijkstra's algorithm. While Dijkstra's algorithm is indeed very useful, there are some other parameters that can be used to find the shortest path while communicating in a weighted network. In a graph network, there are different types of centrality measures used to find the importance of a node. In that Degree Centrality, Closeness Centrality and Betweenness Centrality are useful for identifying the amount of importanceof a node in a graph. Centrality measures are used to find nodes that act as a bridge from one part of a network to another part of a network. In this paper, results shows that there is better possibility to find shortest path using Degree Centrality or Closeness Centrality or Betweenness centrality compared with Dijkstra's algorithm.

Key words: Centrality, Graph, Path, Social Network Analysis

1. INTRODUCTION

To find the shortest path between two nodes in the graph is a problem of finding a way/path between these two vertices in such a way, its sum of distances of its sum of edge distance is minimized. The shortest path problem can be categorized as undirected, directed, signed and multi-graph. Undirected graph considered as bi-directional edges is considered in this paper for identifying shortest path between two nodes.

Two vertices are said to be adjacent when both vertices are connected with a common edge. A path in an undirected graph is defined as a sequence of vertices connecting the vertices from a source to destination vertex.

The extended work of shortest path finding problem is termed as single-pair shortest path problem. In single source shortest path problem, it finds shortest path from a source node to all the other nodes in the directed graph. Another extension is single destination shortest path which finds shortest paths from all the nodes in a graph to single destination node. Another extended work is "all pairs shortest path problem", which finds shortest route between each and every pair of vertices in the graph. The problem of finding the shortest route between two vertices on a social network can be labelled as a unique case of shortest path problems in the graphs.

The Social network is a logical unit that have been useful to study the relationships between an individual, a group and an organization or even with the entire society as nodes. An accepted truth of social network is an approach to understand the social interaction is that social components has to be primarily extracted and has to be investigated by the properties of relationships between the nodes, instead of investigating the properties of the nodes themselves. As different types of relations, the network analytics and configurations were useful to a wider range of research domain.

Social network analysis is the method of analyzing the social structure with the use of networks in the form of graph. It features the networked structure with respect to nodes and the connections termed as edges or links which connects them.

2.RELATED WORKS

The influential nodes in social networks could be perfectly identified with the help of various centrality measures. Researchers have done modifications in the existing centrality calculations and tried the newer algorithms to improvise the efficiency[1]. The author has focused on modular centrality. It aims at calculating the local as well as global influence of the nodes. They aimed at selecting a standard centrality measure. Evacuate all the between network joins from the secluded system and Compute the local network. Process the Local part of the Modular centrality utilizing the standard centrality.

Some authors have done research with the implementation of knowledge domain visualization[2]. They involved in refining the lucidity of networks. In this a triangular inequality function is utilized. The next issue they have focused on is the integration of diverse networks. The final issue they have focused on is the prominent nodes in the complex networks. Upon experimentation, it is understood that the simple recognizable proof of such defining moments is a significant and fundamental advance toward compelling recognition of paradigmatic changes in an information area. The dynamic KDViz technique rearranges the assignments of following critical changes of an information area's cocitation organize after some time.

Habitually, recommender frameworks for the web manage applications that have two measurements, clients and things [3]. The authors have proposed a multidimensional methodology, called DaVI (Dimensions as Virtual Items), that comprises of embedding logical and foundation data as new client thing sets.To assess its viability, we utilized the DaVI approach with two distinctive top-N recommender calculations, Item-based Collaborative Filtering and Association Rules based, and ran a broad arrangement of tests in three diverse true informational collections. The experimental outcomes unequivocally demonstrate that our methodology empowers the utilization of existing two-dimensional suggestion calculations in multidimensional information, misusing the valuable data of these information to improve the prescient capacity of top-N recommender frameworks.

The author focused on smart grid and its applications in paper [4]. Smart grids are modernized power frameworks with data innovation support. Brilliant Grids are the most encouraging advancement in the vitality and utilities market. Smart frameworks are being introduced in numerous nations and it is required to have multi-overlay benefits in productive vitality the executives. The Smart Grids get ongoing meter information with high speed and volume. In such situation, close to ongoing proficient investigation of spilling savvy meter information and fast dynamic is huge. Right now, overview the current procedures and means for constant vitality information the board in keen matrices. Understudies and educators utilize progressively point by point data about their qualities, shortcomings, and individual scholarly execution to comprehend why understudies face a learning hole and attempt to comprehend the examples to beat the hazard in future [5]. So, the information is gathered from understudies with more prominent security by making them answer addresses that in a roundabout way causes us to decide the character, conduct and execution of the understudy, and encourages the school to give incredible guarantee of value instruction for all.

The paper [6] presents a comparison at different centrality measures for chart based key expression extraction. Through investigations completed on three standards informational collections of various dialects and spaces, we show that basic degree centrality accomplishes results tantamount to the generally utilized Text Rank calculation, and that closeness centrality acquires the best outcomes on short records. Utilizing three standard datasets of various dialects and spaces, we demonstrated that degree centrality, regardless of being thoughtfully the least difficult measure, accomplishes results practically identical to the generally utilized Text Rank calculation. Besides, results show that closeness fundamentally outflanks the other centrality gauges on short records.The paper [7] presents a conventional meaning of progression and sums up a current thought of dependability for hub centrality gauges in weighted diagrams. It is indicated that the every now and again utilized proportions of degree, closeness what's more, eigenvector centrality are steady and consistent while betweenness centrality is not one or the other. Numerical tests in manufactured and genuine world systems show that both security and coherence are alluring by and by since they suggest various degrees of strength within the sight of loud information. Specifically, a steady option of betweenness centrality is appeared to show versatility against commotion while saving its thought of centrality.

Stability and coherence, as formal portrayals of the power of centrality measures, were presented. The most every now and again utilized centrality measures were demonstrated to be steady and ceaseless except for betweenness centrality. We represented the heartiness ramifications of solidness and progression in uproarious arbitrary and true systems. At long last, we indicated that the elective stable form of betweenness centrality conveys a comparable centrality thought to the first one.

The paper [8] deals with proposing a recommendation system. Citations are significant in scholarly dispersal. To assist analysts with checking the fulfilment of references while creating a paper, we present a reference suggestion framework called RefSeer. Specialists can utilize it to discover related attempts to refer while writing papers. It can likewise be utilized by analysts to check the culmination of a paper's references. RefSeer presents both points based worldwide proposal and furthermore reference setting based neighbourhood suggestion. By assessing the nature of suggestion, we show that such proposal framework can prescribe references with great exactness and review. We likewise show that our proposal framework is productive and adaptable.

Social Network Analysis (SNA)[9] is the system of performing investigation on information found in social natural circles to comprehend the connections and structures of fundamental system. With the disturbing ascent of web-based life, a wide range of ideas have been achieved through the conduct of individuals inside these systems. Of incredible disarray was the manner by which to distinguish and grant credits to individuals who extraordinarily impact the informal community. This paper looks to draw out the different systems in a social setting. It is known there can be various systems that can be drawn from similar informational collection, the key target accordingly is to help in the assessment of scholarly employees in tertiary foundations and demonstrate that the individuals who occupy the vacant space for example auxiliary gap in the system perform better. For some time it has been realized that scholastic systems investigation has been to a great extent actualized dependent on reference charts not different properties like territories of premium, interdisciplinary work. The significant disadvantage with this sort of investigation is that, the productivity and homophily that exist between the employees outside of their references couldn't be revealed. While a few individuals have low number of references it anyway doesn't suggest they are least dedicated. A few individuals perform better in gathering introductions while slacking in references, in this manner by entreating the reference diagram it involves that these individuals couldn't be completely spoken to by their exertion and commitment to the general public. With this proposed model; the hidden ties between employees depends on their regions of intrigue which has been an untamed space. This will be investigated against Butts' auxiliary gap hypothesis and check whether it applies to scholastic establishments. Key inquiries to be addressed are the way these zones of premium have impacted individual associates inside their system and how it has earned them, if any, advantageous posts in foundations that have different individuals with various premiums in one division. In an all-encompassing investigation the administration quality of individuals in auxiliary opening can likewise be resolved.

The objective of the paper [10] is to identify the shortest path between 2 points in a surface on the space. It was implemented using numerical method based random changes.

The paper [11] represents the SCAMS result in clustering multipath. It is used as an alternate way to cluster multipath to improve accuracy in performance.

3.PROPOSED WORK

In the proposed work, undirected graph with weight for the edges was created using networkx. Python Data frames are used to read the graph data in the form of csv. The following code segment used to read the csv file, and to extract different columns in the heading of SOURCE, TARGET and WEIGHT.

```
df=pd.read_csv("~/grpah_input.csv")
source=df['SOURCE']
dest=df['TARGET']
weight=df['WEIGHT']
nodes=sorted(set(df['SOURCE']))
G=nx.Graph()
```

The below mentioned function is used to create the nodes read from csv data input file.

```
defcreate_node(G):
fori in nodes:
G.add_node(i)
return G
```

The following command is used to create the graph with the nodes alone, without any edges.

```
G=create_node(G)
```

The below mentioned commands are used to add the edges between the nodes, and returns the graph with nodes and edges.

```
Defcreate_edge(source,dest):
fori in range(0,len(source)):
G.add_edge(source[i],dest[i],weight=weight[i])
return G
```

The below line is used to call the function to create the edges and receives the Graph in G.

```
G=create_edge(source,dest)
```

Python and networkx commands are used to create a graph:

```
posi = nx.spring_layout(G)
nx.draw_networkx_nodes(G, posi, node_size=700)
elarger = [(u, v) for (u, v, d) in G.edges(data=True) if d['weight'] >= 3]
esmaller = [(u, v) for (u, v, d) in G.edges(data=True) if d['weight'] < 3]
nx.draw_networkx_edges(G, posi, edgelist=elarger, width=4)
nx.draw_networkx_edges(G, posi, edgelist=esmaller, width=2, alpha=0.5, edge_color='b', style='dashed')
nx.draw_networkx_labels(G, posi, font_size=12, font_family='sans-serif')
```

In this system, spring layout was used to visualize the graph (figure 1) in better manner. Edges are displayed based on their weight if it is >= 3, then displayed with dark lines otherwise dashed lines. Node labels are displayed with mentioned font size and font name. T

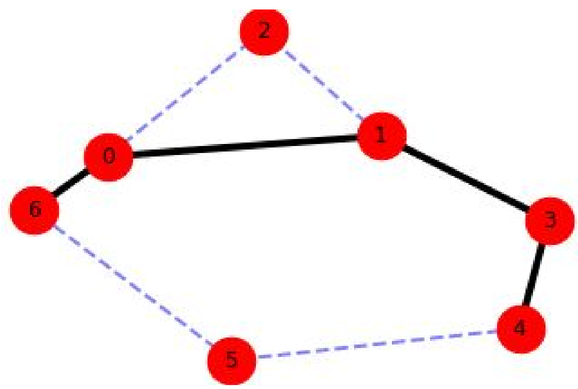

Figure 1: Graph with thick or dashed line based on weight

```
s1=[p for p in nx.all_shortest_paths(G,0,6)]
print("All Shortest Paths without considering weights ",s1)
```

All possible shortest paths were calculated based on nx.all_shortest_paths between the nodes 0 and 6. The output is given below:

All Shortest Paths without considering weights [[0, 6]], Figure 1 shows this.

The same source and destination nodes were give for Dijkst ra algorithm, got the shortest path and length as shown be low which is more appropriate to choose the shortest path.

```
print("Dijkstra Shortest (considering the weights) Path ",s3)
s4=nx.dijkstra_path_length(G,0,6)
print("Dijkstra Shortest (considering the weights) Path Length is ",s4)
```

Dijkstra Shortest (considering the weights) Path [0, 1, 3, 4, 5, 6]
Dijkstra Shortest (considering the weights) Path Length is 1 3

Social network analysis is the process of analyzing social structures in the form of network graph. In that, centrality measures are more useful to find the most influencing central node in the network. There are many centrality measures that are used in social network analysis. Among that, we have used Betweenness centrality, Closeness Centrality and Degree centrality to find the shortest path between two specified nodes along with weight.

The Betweenness Centrality algorithm measures the shortest path between every pair of vertices in a network graph, using BFS(Breadth First Search) algorithms. Then each node receives a point, based on number of shortest paths that passes through the vertex. The vertices that frequently fall on those shortest paths will be having a high betweenness centrality score. Vertices with high betweenness gives the opinion that end to be the brokers in social networks. These vertices combines the different opinion, transfer the ideas between network groups, and often will get power from their ability to make an introductions and pull opinions.

bc=nx.betweenness_centrality(G)

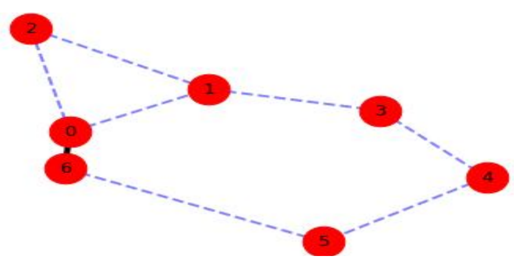

Figure 2: Graph with thick or dashed line based on betweenness centrality

s1=[p for p in nx.all_shortest_paths(G,0,6)]
print("All Shortest Paths without considering weights ",s1)
s3=nx.dijkstra_path(G,0,6)
print("Shortest Path based on Betweenness Centrality ",s3)
s4=nx.dijkstra_path_length(G,0,6)
print("Betweenness Centrality based Shortest Path Length is ",s4)

All Shortest Paths without considering weights [[0, 6]]
Shortest Path based on Betweenness Centrality [0, 2, 1, 3, 4, 5, 6]
Figure 2 shows the betweenness Centrality based Shortest Path Length is 4.5331766

Closeness centrality is the way to identify the nodes which are capable of spreading data more efficiently via network graph. Further the closeness centrality of the node calculates the average distance to every other nodes. Nodes having highest closeness score means that the shortest distance to all other nodes in a graph .

cc=nx.closeness_centrality(G)

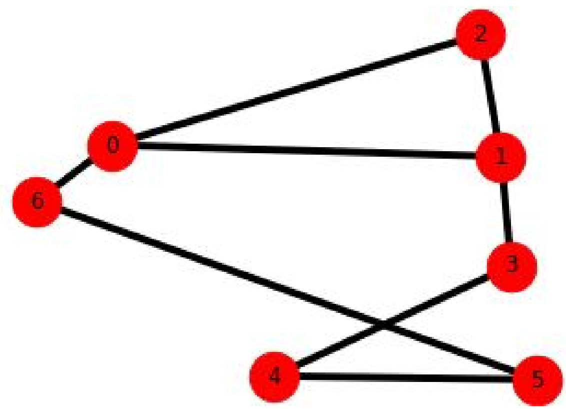

Figure 3: Graph with thick or dashed line based on closeness centrality

All Shortest Paths without considering weights [[0, 6]]

Shortest Path based on Closeness Centrality [0, 2, 1, 3, 4, 5, 6]

Figure 3 shows the Closeness Centrality based Shortest Path Length is 14.063581500000002

Degree Centrality

Degree centrality is defined as number of links connected with node and number of edges a node has. If the network graph is directed, then 2 separate measures of *degree centrality* are defined, named as in degree and out degree.In degree is a count of the number of edgesdirected towards the node andout degree is the measure of number of edges that the node directed to other nodes. The degree centrality calculates the degree as the sum of in degree and out degree.

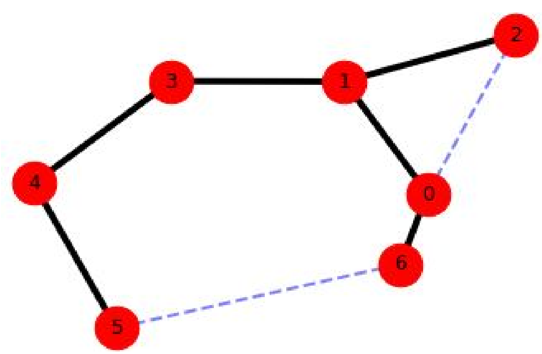

Figure 4: Graph with thick or dashed line based on degree centrality

All Shortest Paths without considering weights [[0, 6]]

Shortest Path based on Degree Centrality [0, 2, 1, 3, 4, 5, 6]

Figure 4 shows the Degree Centrality based Shortest Path Length is 9.8327

Table 1: Edge Weights

SOURCE	TARGET	WEIGHT
0	1	3
1	2	2
2	0	1
3	1	4
4	3	3
5	4	2
6	5	1
0	6	6

The sample data(See Table 1) that contains the edge weights. The comparative results of algorithm for finding shortest path is provided(See Table 2). The result shows that the shortest path between 0 and 6 based on Dijkstra algorithm is 6. The betweenness centrality value along with distance measure shows less path length value.

Table 2: Comparison of Algorithms to find Shortest Path length

Algorithm	Shortest Path	Path Length
Dijkstra Algorithm	[0,6]	6
Betweenness Centrality	[0, 2, 1, 3, 4, 5, 6]	4.5
Closeness Centrality	[0, 2, 1, 3, 4, 5, 6]	14.06
Degree Centrality	[0, 2, 1, 3, 4, 5, 6]	9.83

Sample data(See Table 3) used for different paths with same path length. The result given in "TABLE IV." shows that the shortest path between 0 and 6 based on Dijkstra algorithm is 13. The betweenness centrality, degree centrality values along with distance measure shows less path length value.

Table 3: Sample Data – Edge Weights

SOURCE	TARGET	WEIGHT
0	1	3
1	2	2
2	0	1
	1	4
4	3	3
5	4	2
6	5	1
0	6	13

Table 4:.Comparison of Algorithms based on same path

Algorithm	Shortest Path	Path Length
Dijkstra Algorithm	[0, 2, 1, 3, 4, 5, 6]	13
Betweenness Centrality	[0, 2, 1, 3, 4, 5, 6]	4.5
Closeness Centrality	[0, 2, 1, 3, 4, 5, 6]	14.06
Degree Centrality	[0, 2, 1, 3, 4, 5, 6]	9.83

The algorithms are compared based on similar path (See Table 4).

4.CONCLUSION

All the pairs of vertices in a connected network graph, there exists at least one shortest path exists between the vertices such a way either the number of edges that passes through or the sum of their weights in the edges are minimized. Further betweenness centrality, closeness centrality and degree centrality was calculated for each vertex. The shortest path was identified and analyzed based on these centrality measures as well as Dijkstra shortest path. The results shows that while finding the shortest path between two nodes, the values of centrality measures can also be considered for faster transmission of data through the nodes having higher degree centrality, higher closeness centrality and higher betweenness centrality nodes.

REFERENCES

1. Ghalmane, Z., El Hassouni, M., Cherifi, C.**Centrality in modular networks**, EPJ Data Sci. 8, 15 2019.
2. Chen.**Searching for intellectual turning points: Progressive knowledgedomain visualization,** Proc. Nat. Acad. Sci. USA, vol-101, pp. 5303–5310, 2004. https://doi.org/10.1073/pnas.0307513100
3. M.A.Domingues, A.M.Jorge, and C.Soares. **Dimensions as virtualitems: Improving the predictive ability of top-Nrecommendersystems,** Inf. Process. Manage. Vol- 49, No. 3, pp. 698–720, 2013. https://doi.org/10.1016/j.ipm.2012.07.009
4. K.Sornalakshmi, G.Vadivu.**A Survey on Realtime Analytics Framework for Smart Grid Energy Management,** International Journal of Innovative Research in Science Engineering and Technology, March 2015.
5. G Vadivu, V. Dilli Ganesh**. Ego Network Analysis for Predicting Students Performance**, International Journal of Trend in Scientific Research and Development, 2015.
6. Florian Boudin.**A Comparison of Centrality Measures for Graph-Based Key phrase Extraction**, International Joint Conference on Natural Language Processing, pages 834–838,Nagoya, Japan, 14-18 October 2013.
7. S. Segarra and A. Ribeiro.**Stability and continuity of centrality measures in weighted graphs**, 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, QLD, 2015, pp. 3387-3391. https://doi.org/10.1109/ICASSP.2015.7178599
8. W. Huang, Z. Wu, P. Mitra, and C. L. Giles, RefSeer.**A citation recommendation system**, in Proc. IEEE/ACM Joint Conf. Digit. Libraries (JCDL), Sep. 2014, pp. 371–374.
9. Nyasha F. Thusabantu, Dr. G Vadivu, Mr. L. Thusabantu. **Implementing the Structural Hole**

Theory to Analyze Academic Networks in Tertiary Institutions, International Journal of Trend in Scientific Research and Development, 2015.

10. Ahmad Tayyar.**A shortest path method on a surface in space,** International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.2, 2019, pp. 208-210. https://doi.org/10.30534/ijatcse/2019/16822019
11. Jojo F. Blanza, Lawrence Materum. **Wireless Propagation Multipath Clustering: On Simultaneously Solving the Membership and the Number of Clusters**, International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.5, 2019, pp. 1914 – 1919. https://doi.org/10.30534/ijatcse/2019/15852019