SAM Centrality: a Hop-Based Centrality Measure for Ranking Users in Social Network

Abdul Samad[1,*], Mamoona Qadir[2], Ishrat Nawaz[3], Muhammad Arshad Islam[4], Muhammad Aleem[4]

[1]Capital University of Science and Technology, Islamabad Pakistan
[2]Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan Pakistan
[3]The Islamia University of Bahawalpur, Bahawalpur Pakistan
[4]FAST-National University of Computer and Emerging Sciences, Islamabad Pakistan

Abstract

Majority of researcher are attracted by the social network analysis due to the rush of people towards social network. Along with many problems, social network analysis is facing an interesting problem that is ranking of users in social network which is gaining more attention due to the increasing number of social users. Measuring centrality of nodes in a social graph, have been important issue in social network analysis. Lot of centrality methods have been proposed in this regard. In this paper, hop based centrality measures called SAM is purposed. To investigate the measure, we applied on various dataset. In comparisons, on all these social graphs, we obtain better results than other centrality measures (i.e., Degree, PageRank, Betweeness and Closeness) using SIR model.

1. Introduction

Social network is defined as a string of people, groups and their confidential connections [42]. It may be online through interactions on networking sites (such as Facebook, Twitter, LinkedIn etc.) or offline through face-to-face contacts [40] on public places (such as Universities, Schools, Colleges, Conferences etc.). Moreover, family network (formed by member of family) [35], farmer network (formed by farmers from the village) [49], business network (formed by business menŠs in the business) [19], employee network (formed by employees in the organizations) [48] and player network (formed by players on the ground) [32] are other examples of offline social networks. Nowadays, majority of people are using social apps on smart phones to connect with the people of similar interest. Lately, popular online social networks like Facebook, Twitter and LinkedIn have gained immense attractiveness in our life. Colleagues, individual, groups, etc. all are connected with each others on social networks. So, social network is house of

spreading news [21], marketing product [44], targeting groups [4] etc. To carry out these types of tasks in social network, identifying the central or influential or significant node is challenging task. Centrality specify the most influential or central or significant node in the social network. The centrality of nodes, or detection of influential nodes which are most central than others, has been a basic issue in social network analysis. Suppose there are 75 participants in a meeting. Chair person of the meeting is assuming to be a central. Vice chancellor of the university among the professors is assuming as a influential. Monitor in the class room is assuming as an influential student. Main branch of the bank among other banks in a country is assuming as central branch. For sure, the detection of some nodes with high centralities, which are considered significant, is most valuable in numerous fields, such as detection of vital proteins and candidates for drug target [13], catastrophic outages prevention in the internet or power grids [34] [33] [2], detection of individuals or groups for advertisement of e-commerce items [26] [31], eruption of epidemics controlling [38] [12] and so on.

*Corresponding author. Email: writetosamadalvi@gmail.com

As time is most important [23], with the passage of time lot of centrality measures have been invented, some of them are Closeness [5], Degree [16], Betweenness [15] [8], Coreness [24], PageRank [9], H-Index [30], DPRank [28], LeadRank [29] etc. Lu et al. [6] have done a great survey for the identification of influential vertices in the complex networks. The significance of the node is basically impressed by the topological structure of the social network it affiliated to. Indeed, the majority of centrality measures use the structural information to identify the vital nodes in networks. In general, it is duty of centrality measure to calculate and assign ranking score to every node in the network, where ranking scores are expected to give ranking of nodes according to their significance. According to Lu et al. [6], on the base of structural information, centrality measures are categorized into path-based and neighborhood-based centralities. Neighborhood-based centrality measures indicate that whether two vertices are connected indirectly or directly by any path. In other words, centrality measures that belong to this category are used to find the capability of node to get in touch with all other nodes. Likewise, path-based centrality measures tries to find the shortest path from the starting node to ending node.

In reality, occasionally huge amount of data arise in the network. Therefore, due to high time complexity, it is difficult to measure centrality. Centrality matrices also perform an important role in detection of most powerful and vital elements in huge amount of data. Indeed, every centrality measure identifies influential nodes from different perspective. Time complexity is also another important component of the social network analysis. Based on the different concepts of influence of the nodes or edges, variety of centrality measure were proposed and exerted in suitable area. In this paper, we have investigated the influence of the node via hop distance from other nodes. We argued that as hop distance of node u increased from v the influence of node u decreased. The key contribution of our paper is the proposal of new centrality measure SAM based on the hop distance of nodes. SAM centrality measure is proposed to rank the users in the social network.

Reminder of this paper is structured as follows: State-of-the-art centrality measures from the literature are discussed in section 2; In section 3, definition and methodology of SAM centrality is presented; Results of applying centrality measures on real datasets are shown in section 4; Finally, we conclude the contribution of this paper along with future direction in section 5.

2. Literature Review

A well-known task in finding influential nodes in social network is to measure the centrality of every node. According to some centrality measures, the nodes with
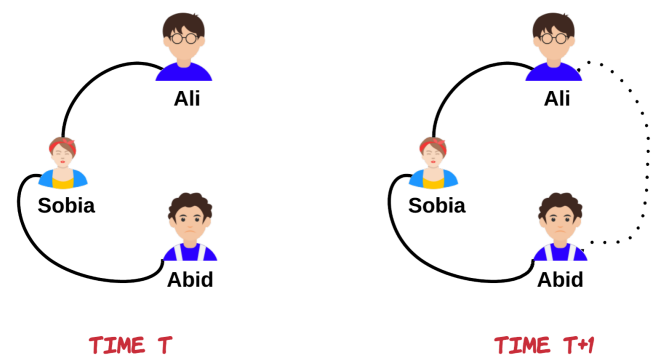

Figure 1. Example of Link Prediction

maximum centrality have more capability to influence the other nodes in the network. Likewise, minimum centrality shows that nodes are not much powerful to influence other nodes the network. For the accurate prediction of vital nodes in the social network, it is more desirable to use centrality measures to discover the most influential nodes. In the last few decades, majority of centrality measures have been proposed as shown below.

In 1950, during the study of communication network, the initial work on centrality measure and its application was defined by bavelas [6]. In 1953, Shimbel proposed stress centrality to determine the quantity of communication which works on the shortest paths [45]. Katz [20] presented Katz centrality to determine the power of a node in the network. For determining the influence of the node, Katz centrality considers all the possible paths in the network. Moreover, Katz stated that the path with shortest length plays an important role in finding influential nodes. In 1965, beauchamp [7] revised the meanings of bavelas's [6] centrality measure and find out the limitations of centrality measure. Besides, proposed an improved centrality measure to extend its utility. After a while, in 1996, Sabidussi [39] professed about Beauchamp [7] improved centrality measure and gave a new definition of centrality. Moreover, new centrality measure is tested for fulfilling the constraints of that definition. Nieminen [37] altered some postulates of sabidussi [39] centrality index and proposed another centrality measure based on degrees of nodes for undirected network graph. In 1978, Freeman [17] revived the research in the field of find influential nodes by proposal of three centrality measures. The first measure is the total effect which indicates the total effect of the node on other nodes in the network. The second measure is the immediate effect which indicates the quickness with which the nodeŠs total effect perceived. The third measure is mediative effect which presented the degree of centralization for the whole network. Stephenson et al. [46] invented

the information centrality which indicates that the transmitted information between two nodes in the connected network can be helpful to find out vital nodes in the graph. In 1991, Freeman et al. [18] again presented a novel idea to find the influential nodes by using the network flow. This measure was same as Freeman [17] but little bit different from the original one. In 1994, Borgatti [47] modified the centrality measure of Freeman [17] and proposed betweenness centrality for the undirected network graph to further directed network graph. In 1999, Borgatti and Evertt [14], extended the state-of-the-art three centrality measures to apply for the individual as well as the groups. Recently, Samad et al. [41], evaluated state of the art centrality measures in order to recommend paper citation.

3. METHODOLOGY: SAM CENTRALITY

First we have proposed a centrality measure namely SAM. In addition, we give a brief picture of real world datasets with ground truth. As the PageRank, Betweeness, Degree and Closeness are targeted by the majority of authors and most commonly used, we treat them as a benchmark centrality measures to compare with our proposed centrality. Also, we have tested our centrality measure on other real world datasets without ground truth.

3.1. Degree Centrality

In a social graph G(V,E), where V represents the set of nodes and E represents the set of ties between nodes, the degree of node xi, is defined as the count of directly connected adjacent nodes of $k_i = \sum_j a_{i,j}$. In mathematics, $a_{i,j}$ represents i^{th} row and j^{th} column of social graph. Where, $a_{i,j} = 1$ means there is a link between nodes and $a_{i,j} = 0$ means nodes are not connected to each other. Furthermore, degree centrality is widely used to identify influential nodes in different networks graphs. If node has more connected neighbors, the more chances to become an influential node. For the purpose of finding influential nodes in different network graphs, normalized degree centrality is calculated by Equation 1.

$$ND_u = \frac{k_u}{n-1} \tag{1}$$

Where, k_u is the number of neighbors of node i, $n = |V|$ is total nodes in the graph and $n-1$ is the possible maximum possible degree of any node.

3.2. Betweeness Centrality

Betweenness centrality was firstly introduced by Bavelas [27] in 1948. Important of node is decided on the stay of node on shortest path surrounded by pair of

nodes in the network. Usually, between v_s (source node) and v_t (target node), many shortest paths exists. The influence of v_u is computed by counting all paths on which v_u exists. Hence, betweenness of v_u is computed as follows in Equation 2.

$$B_u = \sum_{y \neq z \epsilon N} \frac{\sigma_{st}(u)}{\sigma_{st}} \tag{2}$$

Where $\sigma_{st}(u)$ representing the total number of shortest paths between node s and node t. While, σ_{st} representing the number of those shortest paths which are passing through u. Moreover, betweenness centrality is known as global centrality measure. Furthermore, the normalized betweenness centrality is defined as follows in Equation 3.

$$NB_u = \frac{2}{(n-1)(n-2)} \sum_{y \neq z \epsilon N} \frac{\sigma_{st}(u)}{\sigma_{st}} \tag{3}$$

Where n is the size of network.

3.3. PageRank Centrality

PageRank is used by search engine of Google to rank the web pages. It identifies the rank of websites through random walk in network that is based on the connection of different web pages. It identifies the significance of web page by considering the quality and quantity of linked pages. At first, every node assigned one unit as a PR value. After that, each node equally shares PR value with outgoing neighbors. In mathematics, PR value of each node at t step is calculated as follows in Equation 4.

$$P_u = \sum_1^n a_{i,j} \frac{P_j(t-1)}{k_j^{out}} \tag{4}$$

Where n representing the total nodes in the network, k_j^{out} representing the out degree edges of node k_j. The process will stop when PR values of every node will reach the steady state.

3.4. Closeness Centrality

In case of identify the influence of node u, Closeness centrality summarize the distance of node u to all other vertices in the network and defined as follows in Equation 5.

$$C_u = \frac{1}{\sum_{v \epsilon N} d_{u,v}} \tag{5}$$

Where $N = |V|$ is set of nodes in the network. On the other hand, d_u, v is distance between u and v. For comparisons with other centralities closeness centrality is defined as follows in Equation 6.

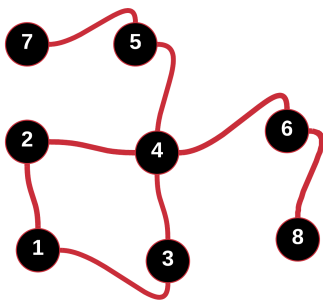$$NC_u = \frac{n-1}{\sum_{v \epsilon N} d_{u,v}} \tag{6}$$

Figure 2. Example of Social Network

Where $n-1$ is representing all other nodes in the network except u for which closeness centrality is computing. Closeness centrality is considered better index than degree as it consider both direct and indirect links. The limitation of this measure is that it is not good for disconnected networks, because finite distance cannot exists between two nodes.

3.5. SAM Centrality

Hop distance Analysis. A lot of researchers invested their research effort to understand online social networks (such as Facebook, Twitter etc) [43]. Social networks have many properties like small-world, degree distribution, scale-free [36], cluttering [27]. Use of the separation metrics(i.e., diameter, radius, eccentricity, average path length etc), which are based on shortest paths of all pairs [11], is mandatory to quantify these properties. These separation metrics are widely used in social networks to study and analyze the overall graph structure. Some commonly used metrics are as follows:

1. Graph Radius: Graph radius is the minimum eccentricity of any node in the graph. Considering Figure 2, Graph radius is 2 as node 4 has minimum eccentricity 2 in the graph.

2. Graph Diameter: Graph diameter is the maximum eccentricity of any node in the graph. Considering Figure 2, Graph diameter is 4 as nodes 1,7 and 8 have maximum eccentricity 4.

3. Eccentricity: Consider shortest paths of node 1 to all other nodes as shown in Figure 2, then the longest shortest path is 4 from node 7 and 8 that is the eccentricity of node 1.

4. Average Path Length: Average path length is defined as the average of all paths.

5. Shortest Path: Considering Figure 2, their are two paths between node 2 and 6, at first path node 4 exist, at second path node 1, 3 and 4 exist. So, the shortest path between nodes 2 and 6 is 2.

6. Hop Distance: Hop distance [25] is the number of intermediate nodes between source and destination nodes. Considering Figure 2, there is 2 hop distance between nodes 1 and 4.

Furthermore, these metrics are performed in the analysis Facebook social graph by nadeem at al. [1]. They have analyzed almost 957K unique Facebook users. According to the results, they have found the following stats:

- 22 is the Highest degree of any node.
- Out of all, 26.96% nodes have only 1 degree.
- By confirming power law, there are large number of nodes having smaller degree (0-4) and only few nodes have large degrees (10-22).
- Diameter of network is 34.
- Average path length is almost 14.
- Average degree of nodes is 3.051. Which means, most of the nodes are indirectly connected to each others.
- Average eccentricity is 26.

Method. In this paper, we have proposed a SAM centrality based on hop distance in the network. This finds the influence of user by considering the hop distance and number of connected nodes on hop distance. For example, consider a graph in Figure 3, where node 1 is directly connected with nodes 2 and 3 and both nodes are at hop $h = 1$ from node 1. Likewise, nodes 4, 5, 6 and 7 are indirectly connected with node 1 by hop distance $h = 2$. Here, nodes 4, 5, 6 and 7 will be given less weight then nodes 2 and 3, which are directly connected. So, this process will end at hop h. In this way, centrality of every node will be computed as follows in Equation 7.

$$SAM_u = \sum_{i=1}^{h} \frac{\sigma_i(x)}{N-1} \times i \qquad (7)$$

Where h is maximum number hop distance of node u whose centrality is calculated. While, $\sigma_i(x)$ is number of neighbors that are connected with node u at i^{th} hop distance. Besides, N is total number of nodes in graph. Moreover, the proposed centrality is normalized i.e., value 1 shows the that the node is influencing the 100% network and 0 shows that no ability to influence the network. Put simply, we assume that in start network all N-1 nodes are connected directly with center node, therefore, the centered node would get value 1. A node with value 0 means node is disconnected from rest of the network or having no adjacent neighbor node.
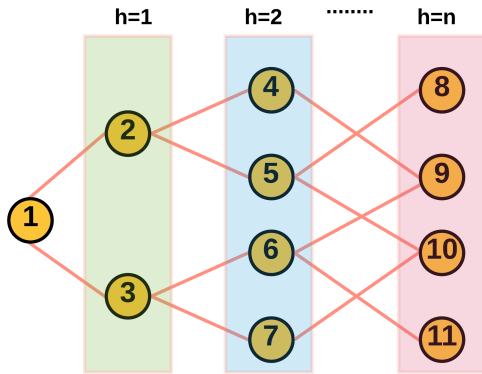
Figure 3. Example of Hope Distance Between Nodes

4. Experiments and Results

In experimental section, to demonstrate the utility of SAM centrality, we apply SAM centrality to several datasets and compare the results with the state-of-the-art centrality measures, such as PageRank, Degree, Closeness and Betweeness. The parameter a for the PageRank is set to 0.85 as recommended in the literature by Page et al. [10]. As we consider only un-weighted networks, if certain networks have weights on the edges, the weights of the edges will be erased. Besides, multiple edges between two nodes will be simplified into single edge.

Table 1. Notations Used in Results

Notation	Description
SAMrank	Ranking given by SAM centrality.
Drank	Ranking given by Degree centrality.
Brank	Ranking given by Betweeness centrality.
Prank	Ranking given by PageRank centrality.
Crank	Ranking given by Closeness centrality.
DScore	Centrality Score of Degree centrality.
BScore	Centrality Score of Betweeness centrality.
PScore	Centrality Score of PageRank centrality.
CScore	Centrality Score of Closeness centrality.
SAMScore	Centrality Score of SAM centrality.

Here, we tested SAM centrality on four real networks (such as Zachary's Karate Club, Dolphin, Jakarta and Rhodes) with ground truth (i.e., which node is most important in the network). We have seen that SAM centrality finds correct results on these benchmark datasets.

4.1. Dataset 1: Zachary's Karate Club

Zachary's Karate Club was the initial social network. From 1970 to 1972, Wayne W Zachary studied the Karate Club social network for three years. The network

graph consists of 34 nodes and 78 edges. The nodes, representing the members of the Karate Club, while edges representing the relationship between members who interacted to each other outside the club. During his study, a clash between administrator of the club "John A" (as ♯34) and instructor of the club "Mr. Hi" (as ♯1) was started, as a result the club had divided into two clubs. Few members formed another club along with Mr. Hi (instructor), while the remaining members of the club along with John A (Administrator) found new instructor for the club.

For the Karate Club social network, the normalized centrality score of all centrality measures is listed in Table 2. The top individuals from the network are presented in bold. It can easily be seen in Table 2, Administrator (♯34) and Instructor (♯1) getting the high scores from all centrality measures. Moreover, Figures 4 and 5 are showing the results of ranking and centrality score given by each centrality measure. Here you can see that, most of the nodes have got low ranking given by other centrality measures.

Table 2. Centrality Score of Top 10 Nodes, From Zachary's Karate Club, Sorted by SAM Centrality

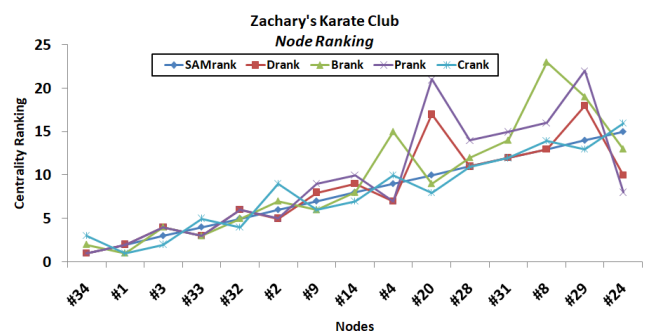| Node	D_u	B_u	P_u	C_u	SAM_u
♯34	0.515	0.304	0.101	0.550	0.697
♯1	0.485	0.438	0.097	0.569	0.692
♯3	0.303	0.144	0.057	0.559	0.626
♯33	0.364	0.145	0.072	0.516	0.626
♯32	0.182	0.138	0.037	0.541	0.576
♯2	0.273	0.054	0.053	0.485	0.571
♯9	0.152	0.056	0.030	0.516	0.551
♯14	0.152	0.046	0.030	0.516	0.551
♯4	0.182	0.012	0.036	0.465	0.525
♯20	0.091	0.032	0.020	0.500	0.520

Figure 4. Ranking of top 15 nodes from Zachary's Karate Club, sorted by SAM centrality

4.2. Dataset 2: Dolphin

Dolphin is an undirected social network of frequent interaction between 62 dolphins in New Zealand. Here
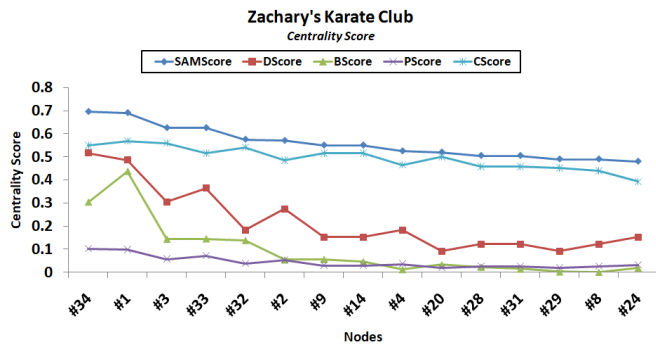
Figure 5. Centrality score of top 15 nodes from Zachary's Karate Club, sorted by SAM centrality

Figure 6. Ranking of top 15 nodes from Dolphin dataset, sorted by SAM centrality

Figure 7. Centrality score of top 15 nodes from Dolphin dataset, sorted by SAM centrality

link indicate the frequent interaction between nodes. The data is collected between 1994 and 2001. The network consists of 62 nodes and 159 edges.

Table 3 represents the centrality score of all nodes from all centrality measures. The dolphin ♯36 getting high scores from all centrality measures. Moreover, Figures 6 and 7 are showing the results of ranking and centrality score of all centrality measures. In case of ranking, you can see, there are more ups and downs. This shows that, all centrality measures are not agrees to each other as they given high and low ranks.

Table 3. Centrality Score of Top 10 Nodes, From Dolphin Dataset, Sorted by SAM Centrality

Node	D_u	B_u	P_u	C_u	SAM_u
♯36	0.115	0.248	0.021	0.418	0.477
♯40	0.131	0.143	0.022	0.404	0.477
♯37	0.180	0.139	0.030	0.399	0.496
♯20	0.148	0.103	0.025	0.391	0.477
♯14	0.197	0.062	0.032	0.377	0.492
♯1	0.131	0.213	0.025	0.372	0.445
♯33	0.164	0.057	0.028	0.365	0.467
♯7	0.082	0.118	0.016	0.365	0.419
♯28	0.082	0.067	0.015	0.365	0.421
♯8	0.098	0.022	0.017	0.363	0.433

4.3. Dataset 3: Jakarta

In 2009, a terrorist attack took place in Jakarta, where two hotels were hit by a group of 28 terrorist. In this attack, nine people were killed by these terrorists. Jakarta social network representing the relations between terrorist who were directly or indirectly associated with Jakarta terrorist attack.

The normalized centrality score, for the Jakarta social network, is listed in Table 4. The top individual ♯177 is getting high centrality score from all centrality measures. Furthermore, Figures 8 and 9 are showing the results of ranking of nodes as well as centrality score.
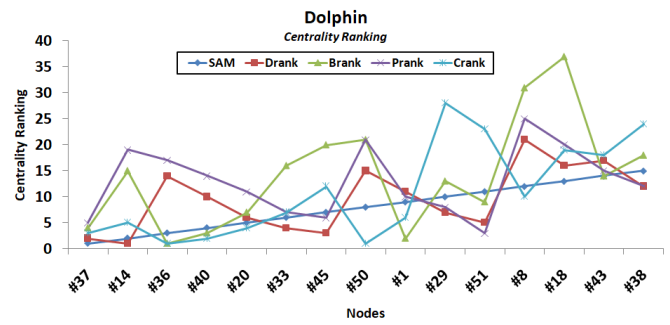
In case of ranking, all centrality measures are little bit agree on top few nodes. Besides, there is huge difference in centrality score.

Table 4. Centrality Score of Top 10 Nodes, From Jakarta Dataset, Sorted by SAM Centrality

labelTab:3					
Node	D_u	B_u	P_u	C_u	SAM_u
177	0.145	0.072	0.053	0.031	0.258
1561	0.091	0.043	0.031	0.031	0.236
2371	0.091	0.026	0.030	0.031	0.233
2370	0.073	0.018	0.024	0.031	0.221
1595	0.073	0.023	0.026	0.031	0.214
2374	0.073	0.012	0.025	0.031	0.211
2391	0.055	0.020	0.024	0.031	0.183
2373	0.036	0.001	0.014	0.030	0.170
598	0.036	0.001	0.014	0.030	0.168
2392	0.018	0.000	0.009	0.030	0.160

4.4. Dataset 4: Rhodes

Rhodes is a social network of Greek terrorist group known as N17 (November 17). That was collected from the reporting (Abram and smith 2004; Irwin et. al 2002). In the social graph, relations indicated that reporting has confirmed the link between two
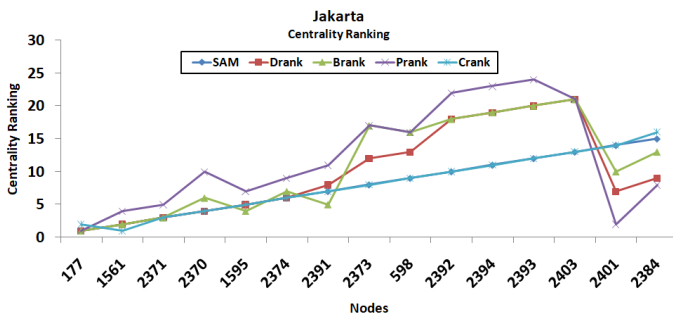
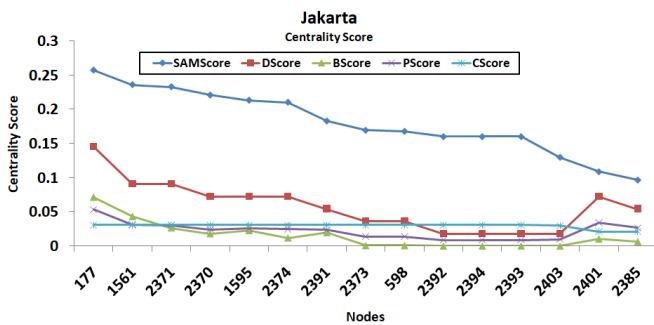**Figure 8.** Ranking of top 15 nodes from Jakarta dataset, sorted by SAM centrality

Table 5. Centrality Score of Top 10 Nodes, From Rhodes Dataset, Sorted by SAM Centrality

			labelTab:1		
Node	D_u	B_u	P_u	C_u	SAM_u
Pavlos	0.32	0.11	0.04	0.48	0.60
Christodoulos	0.25	0.08	0.03	0.48	0.57
Dimitris	0.25	0.08	0.03	0.48	0.57
Alexandros	0.23	0.10	0.03	0.48	0.56
Savas	0.25	0.18	0.04	0.45	0.55
Nikitas	0.14	0.02	0.02	0.43	0.49
Fotis	0.14	0.02	0.02	0.43	0.49
Yiannis	0.14	0.02	0.02	0.43	0.49
Sardanopoulos	0.14	0.01	0.02	0.41	0.47
Vassilis	0.09	0.01	0.01	0.41	0.45

Figure 9. Centrality score of top 15 nodes from Jakarta dataset, sorted by SAM centrality
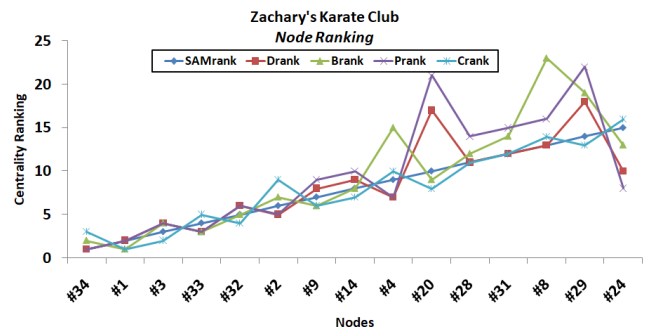

Figure 10. Ranking of top 15 nodes from Rhodes dataset, sorted by SAM centrality
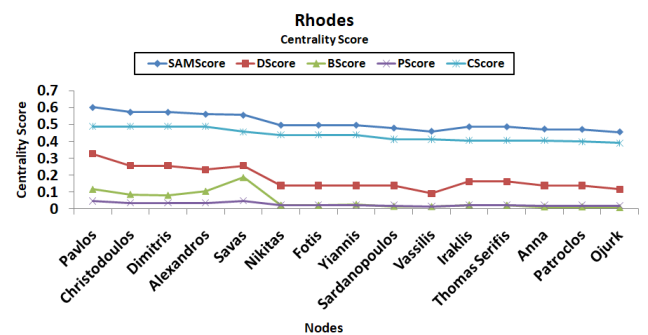
Figure 11. Centrality score of top 15 nodes from Rhodes dataset, sorted by SAM centrality

terrorists at some point few years ago. N17 have attacked many kinds of targets from 1975 via bombing and assassination. The network of N17 shows that it is small group of 63 links and 22 nodes.

The normalized centrality score of Rhodes network is listed in Table 5. Where, "Pavlos" and "Christodoulos" getting the high scores from all centrality measures. Moreover, Figures 10 and 11 are representing the results of ranking and centrality score given by each centrality measure. Here you can see that, most of the nodes have got low ranking given by other centrality measures. In case of centrality score, closeness centrality is close to our proposed centrality SAM. Besides, not much difference between state-of-the-art centralities for the top few nodes.

4.5. More Datasets

To show the utility of SAM centrality, further we test SAM centrality on five more datasets (i.e., Astro, CondMat, GrQc, HepPh and HepTh), whose stats are shown in Table 6. As previously used datasets (i.e., Zachary's Karate Club, Dolphin, Jakarta and Rhodes) have ground truth, datasets used here in this section have no ground truth. Therefore, we used SIR

(i.e., susceptible-infected-recovered) model [3] for the ranking criteria. Here, every node is called susceptible except the first which is called infected node (i.e., the initial node). At every step, the infected nodes will infect their neighbors which are still in susceptible states with β (i.e., the probability value). After infecting the neighbors, infected node will enter into recovered state with γ (i.e., a probability value). Here we set the probability value $\gamma = 1$. When there would be no infected node, the spreading process will be end. For comparing SAM centrality with other centrality measures, we analyze KendallŠs tau correlation [22]. Here, high t means that the performance of centrality measures is better. We test centrality measures with two normally used infected probabilities (i.e., $\beta = \beta c$ and $\beta = 1.5\beta c$), where βc is the threshold. The results of correlation are shown in Table 7.

Table 6. Stats of Five More Datasets

Dataset	Nodes	Edges	Triangles
AstroPh	18772	198110	1351441
CondMat	23133	93497	173361
GrQc	5242	14496	48260
HepPh	12008	118521	3358499
HepTh	9877	25998	28339

Table 7. KENDALL'S Correlation with Probability β

β	Network	D_u	B_u	P_u	C_u	SAM_u
βc	Astro	0.62	0.71	0.59	0.78	0.87
1.5β		0.68	0.73	0.61	0.80	0.89
βc	CondMat	0.69	0.66	0.52	0.75	0.85
1.5β		0.71	0.67	0.54	0.78	0.86
βc	GrQc	0.70	0.44	0.24	0.66	0.83
1.5β		0.71	0.44	0.83	0.67	0.81
βc	HepPh	0.70	0.41	0.41	0.68	0.79
1.5β		0.70	0.42	0.40	0.69	0.80
βc	HepTh	0.66	0.50	0.46	0.61	0.81
1.5β		0.68	0.51	0.45	0.60	0.82

5. Conclusion

In this paper, we have proposed a hop-based centrality measure SAM. We have experimented on four datasets and compare our centrality with other state-of-the-art four centrality measures. Our purposed centrality SAM punishes the hop distance heavily in order to rank the nodes in the network. The results show that SAM is better to find t he i mportance o f n ode i n t he network. in future, we will try to find out the inverse distance of nodes in order to rank the users of the network.

References

[1] AKHTAR, N., JAVED, H. and SENGAR, G. (2013) Analysis of facebook social network. In *2013 5th International Conference and Computational Intelligence and Communication Networks* (IEEE): 451–454.

[2] ALBERT, R., ALBERT, I. and NAKARADO, G.L. (2004) Structural vulnerability of the north american power grid. *Physical review E* **69**(2): 025103.

[3] ANDERSON, R.M., ANDERSON, B. and MAY, R.M. (1992) *Infectious diseases of humans: dynamics and control* (Oxford university press).

[4] BAGLIONE, S.L. and TUCCI, L.A. (2019) Perceptions of social media's relevance and targeted advertisements. *Journal of Promotion Management* **25**(2): 143–160.

[5] BAVELAS, A. (1950) Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America* **22**(6): 725–730.

[6] BAVELAS, A. (1950) Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America* **22**(6): 725–730.

[7] BEAUCHAMP, M.A. (1965) An improved index of centrality. *Behavioral science* **10**(2): 161–163.

[8] BRANDES, U. (2001) A faster algorithm for betweenness centrality. *Journal of mathematical sociology* **25**(2): 163–177.

[9] BRIN, S. and PAGE, L. (1998) The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems* **30**(1-7): 107–117.

[10] BRIN, S. and PAGE, L. (1998) The anatomy of a large-scale hypertextual web search engine* 1. *Computer networks and ISDN systems* **30**(1-7): 107–117.

[11] CHIEN, T.Q. and VAN, H.H. (2017) Applying algorithm finding shortest path in the multiple-weighted graphs to find maximal flow in extended linear multicomodity multicost network. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems* **4**(11).

[12] COHEN, R., HAVLIN, S. and BEN-AVRAHAM, D. (2003) Efficient immunization strategies for computer networks and populations. *Physical review letters* **91**(24): 247901.

[13] CSERMELY, P., KORCSMÁROS, T., KISS, H.J., LONDON, G. and NUSSINOV, R. (2013) Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review. *Pharmacology & therapeutics* **138**(3): 333–408.

[14] EVERETT, M.G. and BORGATTI, S.P. (1999) The centrality of groups and classes. *The Journal of mathematical sociology* **23**(3): 181–201.

[15] FREEMAN, L.C. (1977) A set of measures of centrality based on betweenness. *Sociometry* : 35–41.

[16] FREEMAN, L.C. (1978) Centrality in social networks conceptual clarification. *Social networks* **1**(3): 215–239.

[17] FREEMAN, L.C. (1978) Centrality in social networks conceptual clarification. *Social networks* **1**(3): 215–239.

[18] FREEMAN, L.C., BORGATTI, S.P. and WHITE, D.R. (1991) Centrality in valued graphs: A measure of betweenness based on network flow. *Social networks* **13**(2): 141–154.

[19] GORAYA, M.A.S., JING, Z., SHAREEF, M.A., IMRAN, M., MALIK, A. and AKRAM, M.S. (2019) An investigation of the drivers of social commerce and e-word-of-mouth

intentions: Elucidating the role of social commerce in e-business. *Electronic Markets* : 1–15.

[20] KATZ, L. (1953) A new status index derived from sociometric analysis. *Psychometrika* **18**(1): 39–43.

[21] KELLY, M.P., MARTIN, N., DILLENBURGER, K., KELLY, A.N. and MILLER, M.M. (2019) Spreading the news: History, successes, challenges and the ethics of effective dissemination. *Behavior Analysis in Practice* **12**(2): 440–451.

[22] KENDALL, M.G. (1938) A new measure of rank correlation. *Biometrika* **30**(1/2): 81–93.

[23] KHAN, S.U.R., ISLAM, M.A., ALEEM, M., IQBAL, M.A. and AHMED, U. (2018) Section-based focus time estimation of news articles. *IEEE Access* **6**: 75452–75460.

[24] KITSAK, M., GALLOS, L.K., HAVLIN, S., LILJEROS, F., MUCHNIK, L., STANLEY, H.E. and MAKSE, H.A. (2010) Identification of influential spreaders in complex networks. *Nature physics* **6**(11): 888.

[25] LE, D., NGUYEN, B. and TRAN, H. (2019) A simulation analysis of the connectivity of multi-hop path between two arbitrary nodes in cognitive radio ad hoc networks. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems* **6**(21).

[26] LESKOVEC, J., ADAMIC, L.A. and HUBERMAN, B.A. (2007) The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)* **1**(1): 5.

[27] LI, L., ALDERSON, D., DOYLE, J.C. and WILLINGER, W. (2005) Towards a theory of scale-free graphs: Definition, properties, and implications. *Internet Mathematics* **2**(4): 431–523.

[28] LIU, M., XIONG, Z., MA, Y., ZHANG, P., WU, J. and QI, X. (2018) Dprank centrality: Finding important vertices based on random walks with a new defined transition matrix. *Future Generation Computer Systems* **83**: 376–389.

[29] LÜ, L., ZHANG, Y.C., YEUNG, C.H. and ZHOU, T. (2011) Leaders in social networks, the delicious case. *PloS one* **6**(6): e21202.

[30] LÜ, L., ZHOU, T., ZHANG, Q.M. and STANLEY, H.E. (2016) The h-index of a network node and its relation to degree and coreness. *Nature communications* **7**: 10168.

[31] LU, Y. (2012) Recommender systems. In *Physics Reports* (Citeseer).

[32] MCLEAN, S., SALMON, P.M., GORMAN, A.D., DODD, K. and SOLOMON, C. (2019) Integrating communication and passing networks in football using social network analysis. *Science and Medicine in Football* **3**(1): 29–35.

[33] MOTTER, A.E. (2004) Cascade control and defense in complex networks. *Physical Review Letters* **93**(9): 098701.

[34] MOTTER, A.E. and LAI, Y.C. (2002) Cascade-based attacks on complex networks. *Physical Review E* **66**(6): 065102.

[35] NGUYEN, A.W., CHATTERS, L.M. and TAYLOR, R.J. (2016) African american extended family and church-based social network typologies. *Family relations* **65**(5): 701–715.

[36] NGUYEN, M.T. (2019) An energy-efficient framework for multimedia data routing in internet of things (iots). *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems* **6**(19).

[37] NIEMINEN, J. (1974) On the centrality in a graph. *Scandinavian journal of psychology* **15**(1): 332–336.

[38] PASTOR-SATORRAS, R. and VESPIGNANI, A. (2002) Immunization of complex networks. *Physical review E* **65**(3): 036104.

[39] SABIDUSSI, G. (1966) The centrality index of a graph. *Psychometrika* **31**(4): 581–603.

[40] SAMAD, A. (2019) *Evaluation of Textual and Topological Similarity Measures for Citation Recommendation*. Ph.D. thesis, CAPITAL UNIVERSITY.

[41] SAMAD, A., ISLAM, M.A., IQBAL, M.A. and ALEEM, M. (2019) Centrality-based paper citation recommender system. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems* **6**(19).

[42] SAMAD, A., QADIR, M. and NAWAZ, I. (2019) Sam: a similarity measure for link prediction in social network. In *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)* (IEEE): 1–9.

[43] SATYANARAYANA, S.K., SOOD, K., TAO, Y. and YU, S. (2014) Security and privacy in online social networks: A survey. *EAI Endorsed Trans. Indust. Netw. & Intellig. Syst.* **1**(1): e3.

[44] SHAREEF, M.A., MUKERJI, B., DWIVEDI, Y.K., RANA, N.P. and ISLAM, R. (2019) Social media marketing: Comparative effect of advertisement sources. *Journal of Retailing and Consumer Services* **46**: 58–69.

[45] SHIMBEL, A. (1953) Structural parameters of communication networks. *The bulletin of mathematical biophysics* **15**(4): 501–507.

[46] STEPHENSON, K. and ZELEN, M. (1989) Rethinking centrality: Methods and examples. *Social networks* **11**(1): 1–37.

[47] WHITE, D.R. and BORGATTI, S.P. (1994) Betweenness centrality measures for directed graphs. *Social networks* **16**(4): 335–346.

[48] ZHANG, L., GUO, X., LEI, Z. and LIM, M.K. (2019) Social network analysis of sustainable human resource management from the employee trainingâĂŹs perspective. *Sustainability* **11**(2): 380.

[49] ZIEGENHORN, R. (2018) *Networking the farm: The social structure of cooperation and competition in Iowa agriculture* (Routledge).