Proceedings on Big Data Analytics
& Innovation (Peer-Reviewed),
Volume 1, 2016, pp.177-190

Improved Influence Factor Scheme for Detecting Influential Nodes in Mobile Phone Network

ELIZABETH N. ONWUKA, BALA A. SALIHU & SHERIFF MURTALA
Federal University of Technology, Minna, Nigeria

ABSTRACT The number of mobile phone users is increasing tremendously. Network of users are formed using the call (or social) interactions between these mobile phone users. Such networks could be represented using social network graphs where the nodes represent persons and the edges are the communications between them. In such networks, communities of nodes with certain commonalities could be identified using community detection techniques. It should be noted that in every community there are usually nodes that have high influence, referred to as influential nodes. Knowledge of such nodes helps to understand the communities better and to relate with the community members. For example, removal of influential nodes from a criminal community will collapse the community and probably also the network they belong to. Also, influential nodes could be used to feed information to an entire network. Therefore, it is important to accurately identify nodes that are prominent in a network. For these reasons, work on techniques for identifying influential nodes in communities is currently receiving attention in the research arena. One such technique in literature is the influence factor scheme, which indicates how important an individual node can be in a network. The scheme integrates betweenness centrality, closeness centrality and eigenvector centrality. However, the use of eigenvector centrality in the scheme strongly affects the measure of influence across the network by limiting the detection of influential nodes to the neighbouring nodes around the most influential nodes within the largest component (community) of the network. It neglects the fact that there could be an influential node in other smaller components in the network. This can be misleading, especially in a massive social network like the mobile phone network that contains several neighbourhoods with hundreds or thousands of nodes and edges. This is because it is not necessarily true that every node that is connected to the most important nodes is truly important. This limitation makes it difficult to detect the real influential nodes in large social networks. Principal component centrality is a variant of eigenvector centrality that considers every component (community) in a network when searching for influential nodes across a network graph. In this research, we present an improved influence factor scheme that incorporates closeness centrality, betweenness centrality and principal component centrality to identify nodes that are truly influential in a mobile phone network. The improved scheme has better accuracy, precision and specificity. Furthermore, in terms of accessibility, the improved scheme outperforms the existing scheme because information through the detected influential nodes reached all members of the communities in the network.

Keywords: social network, mobile phone network, influential nodes detection, centrality measures

1. Introduction

Since the invention of mobile communication and other services attached to it, many people find it better and cheaper to communicate using the medium than wired communication thereby attracting more subscribers to use mobile communication network. A survey carried out by international telecommunication union (ITU) shows that the population of mobile phone subscribers increased from 738 million in the year 2000 to 7 billion in 2015 and within this same time the proportion of population covered by a 2G mobile cellular network rose from 58% to 95% with more remote areas captured (Reserved, 2016). In developing countries, at least one member of every household communicates using a mobile phone. Each subscriber enjoys making calls and receiving calls from other users, and also enjoys the same for short message and Internet services. Telecommunication networks have really made the world a global village in the sense that peoples' social reach has expanded even across borders. The log of activities of each user is stored on the user's phone and also recorded with the Mobile Network Operators (MNOs). The information collected by the MNOs is referred to as Call Detail Record (CDR).

CDR contains metadata that describe a specific instance of a telecommunication transaction (calls, messages and Internet services) but does not include the content of that transaction, for example, CDR for a particular call contains both the caller and receiver's number, the time stamp (date and time), the duration of that call and other relevant information. CDR may capture thousands or millions of users within a specific time and place and it can be used to create a network of mobile phone subscribers. CDR is a huge repository of human behavioral data and it belongs to the group of data being currently described as Big Data. Inter-relationship network between humans at various spheres, generally called social networks, can be reconstructed with CDR. A mobile phone network is a social structure that represents the interconnection of mobile phone subscribers based on call detail record (CDR). The idea of forming a social interaction between mobile phone users support researchers in different area of studies like personal mobility prediction, fraud detection in telecommunication (Pinheiro, 2012), urban planning and development, geographical partitioning (Blondel *et al*., 2015) and intelligence gathering for national security (Farley, 2003). Considering the benefits of knowing the most influential nodes in a group or clusters, it is important to develop a technique for identifying the most influential nodes in any given group. This is because identification of such nodes gives a good insight into that group. The major problem in this area is how to accurately determine the genuine influential nodes (individuals) in a social network.

Related work

2.1 Background

There is a rapidly growing literature on influential nodes discovery in social networks, which indicates that a lot of study had been carried out in this field (Borgatti,

2006) (Probst, 2013) (Zhang *et al*., 2013) (Ilyas and Radha, 2010) (Ilyas and Radha, 2011) (Sathik and Rasheed, 2009) (Ahsan, *et al*., 2015) (Singh *et al*., 2013). However, due to the challenges of getting mobile phone data, little studies have been carried out on discovering communities and important mobile subscribers in mobile phone network. A mobile phone network is treated like any other social network that has a tree network structure. Social network is usually modelled as a graph, $G = (V, E)$ where V is a set containing all nodes (actors) in the network and E is also a set containing all edges (links) between two elements (pairs) of set V. If the direction of the edges is considered the graph is said to be directed and undirected otherwise. Also, when the weight of the edges is considered, the graph is said to be weighted and binary (unweighted) otherwise.

Exploring social network data requires basic concepts of graph representation, analysis and visualization (Abraham, 2012). These concepts include centrality measures, shortest path problems, clustering techniques and network density. This is necessary when interpreting result in order to have a good understanding of the social interactions between nodes in a network. Due to the rich resources in social network analysis, it serves as a tool for analyzing and visualizing big data (Lieberman, 2014). Some major areas of study in social network analysis are community structure, detection of cliques and discovery of key nodes and neighbours. Recently, more attention has been given to detection of influential nodes in social network. This is added to the fact that researchers and investigators have taken full advantage of social network analysis to unravel the operation of terrorists and criminals (Farley, 2003). This crime investigation application becomes more necessary now that communication networks has changed the way people live and transact business. It is intuitively believed that criminals rely on this network for planning criminal activities of all sorts. In this study, we focused on identifying important and interesting nodes in a mobile phone network.

2.2 Influential Nodes

Influential nodes are set of nodes whose roles are very important in the spread of influence across the network. These nodes have the tendency to influence other nodes either constructively or destructively. Influential nodes and "key nodes" seem to be the same. Recently, (Probst, 2013) (Singh *et al*., 2013) presented an overview of existing techniques of finding important and influential nodes in social networks. In this subsection, we discuss some of the previous studies that had been done in this area of research. For clarity, we classify the methods of influential nodes detection into two categories: centrality measures and non-centrality approach.

Centrality measures

In graph theory and network analysis, the most important tool is centrality measure. Centrality measures are considered as structural measures of influence that indicate a node's position in a social network. Degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality are the four widely used centrality measures in determining the relative importance of a node within a network. Although these measures have limitations, they have been proven to be the basis of other methods for identifying key nodes within a social network (Landherr *et al.*, 2010).

i) Degree Centrality: Degree centrality is defined as the number of edges incident upon a node. In other words, this measure indicates how many nodes can be directly reached by a particular node. The degree centrality of a user, *v* is given by

$$DC(v) = \deg(v) \tag{1}$$

$$\deg(v, G) = |\{u \in V : (u, v) \in E\}| \tag{2}$$

Nodes with high degree centrality scores might be considered important. But one major flaw of this centrality measure is that it relies on direct connections between nodes. Using this individual centrality alone to determine the key nodes will result in the selection of nodes that have high number of direct connections.

ii) Closeness Centrality

Bavelas defined closeness centrality of a user as the reciprocal of the sum of its distances from all other nodes (Bavelas, 1950). This measure is effective in describing the hierarchy among members within a group and can also be used to indicate how fast a node can reach every other node in the network. The weakness of closeness centrality is that it is unsuitable for disconnected graphs.

iii) Betweenness Centrality: This expresses the number of times a user acts as a bridge along the shortest path between two other nodes. Nodes with high betweenness are responsible for controlling the spread of information across the graph. However, they might not be responsible for causing maximum disconnection (fragment) within the network (Borgatti, 2006).

iv) Eigenvector Centrality: Eigenvector centrality (also called eigencentrality) is a measure of how well a particular node is connected to other influential nodes. This is one of the oldest centrality measures developed to assist social analyst to recognize the behavior of people (Seeley, 1949). To determine eigenvector centrality, it is imperative to first find the adjacency matrix, A of the graph, G. with $A = a(v,u)$; and $a(v,u) = 1$ if there exist a link between nodes "v" and "u" and $a(v,u) = 0$ if otherwise for a binary network. The eigenvector centrality of a node, *v* is expressed mathematically as

$$x_v = \frac{1}{\lambda} \sum_{u \in M(v)} x_u \tag{3}$$

where M(v) is the set of neighbours of node, *v*. In matrix representation, eigenvector centrality is given as

$$x_v = \frac{1}{\lambda} A x_u \quad (4)$$

where λ is the eigenvalue (constant) and x_v is the corresponding eigenvector.

v) Other Centrality-Based Approaches: The number of centrality measures extend beyond the four metrics discussed earlier. It is quite interesting that most of the new measures were related in one way or the other to the four most popular centrality measures with a little modification. Ilyas and Radha introduced a new centrality called principal component centrality (PCC), a variant of eigenvector centrality (Ilyas and Radha, 2010). PCC is based on principal component analysis (PCA) and karhunen loeve transform (KLT) which handles graph adjacency matrix as a covariance matrix. Contrary to Eigenvector centrality, PCC provides more features for centrality computation. Moreover, an investigation was carried out to detect influential nodes in two separate datasets using eigenvector centrality and principal component centrality (Ilyas and Radha, 2011). Eigenvector centrality usually considers the most influential user within the largest community in a network and consequently ranks the neighbours of the influential node and ignores other nodes in the remaining small communities that have low eigenvector scores. In the case of PCC, it considers both the nodes in the largest community and other nodes with zero eigenvalues in the remaining small communities.

Despite the introduction of these new centrality measures. The fact still remains that an individual centrality measure might not be the most appropriate for a given network application. A centrality measure is applied depending on specific purpose and the position of a user in a network. For instance, nodes that are most spreaders of virus act as regulators in the network. Another different purpose is identifying nodes that can maximally disrupt the social network. This has opened up more fascinating research fields on group and improved centrality measures that can be universal in identifying the most influential nodes (Everett and Borgatti, 1999). Some studies also considered combining two or more centralities measures in getting a general set of influential nodes. Sathik and Rasheed proposed an algorithm to identify sets of key players based on centrality measures (Sathik and Rasheed, 2009). The authors addressed the key player problems (Borgatti, 2006), using closeness centrality, degree centrality and betweenness centrality.

Lately, in order to adequately discover real influential nodes. Ahsan *et al*. described a scheme that combines closeness centrality, betweenness centrality and eigenvector centrality to determine the influence factor of actors in an online social network obtained from Facebook (Ahsan, *et al*., 2015). The study shows that these three centrality measures are important in measuring the influence of each user and as well as the influence of the entire social network.

Non Centrality Approach

In this subsection, we would be looking at previous studies that employed other techniques different from centrality approach in detecting influential nodes.

$$IF(v) = \frac{2CC_{norm}(v)BC_{norm}(v)PCC_{norm}(v)}{CC_{norm}(v) + BC_{norm}(v)} \quad ; \quad 0 < IF(v) < 1 \tag{9}$$

where $IF(v)$ is the influence factor of node v, $CC_{norm}(v)$ is the normalized closeness centrality of node v, $BC_{norm}(v)$ is the normalized betweenness centrality of node v and $PCC_{norm}(v)$ is the normalized principal component centrality of node v. The normalization for each centrality measure is done using the expression

$$Z_{norm}(v) = \frac{Z(v) - Z_{\min}}{Z_{\max} - Z_{\min}} \tag{10}$$

A node will have an influence factor value between 0 and 1. Where 0 describes node as insignificant and 1 defines node as highly influential.

Results

The data processing was done using the Data laboratory tab in Gephi (Bastian *et al.*, 2009), while other analysis was carried out using Python 2.7 on a Dell Latitude computer system with Intel(R) Core(TM) i5 CPU M 540@ 2.53GHz processor and 4GB RAM. NodeXL, a free open-source template for Microsoft® Excel® was used in evaluating the performance of the improved scheme.

4.1 Processed Data

Nodobo dataset contains unsuccessful calls which make up about 30% of the total number of calls. These calls are either calls missed by the call receiver or outgoing calls that failed to connect due to low airtime or weak service signal of network operators. The details are presented in Table 1.

Table 1. Details of the call records

Call Status	Outgoing calls	Incoming calls	Missed calls	Total
Successful calls	5,976	2,998	Nil	8,974
Unsuccessful calls	2,068	169	1,824	4,061

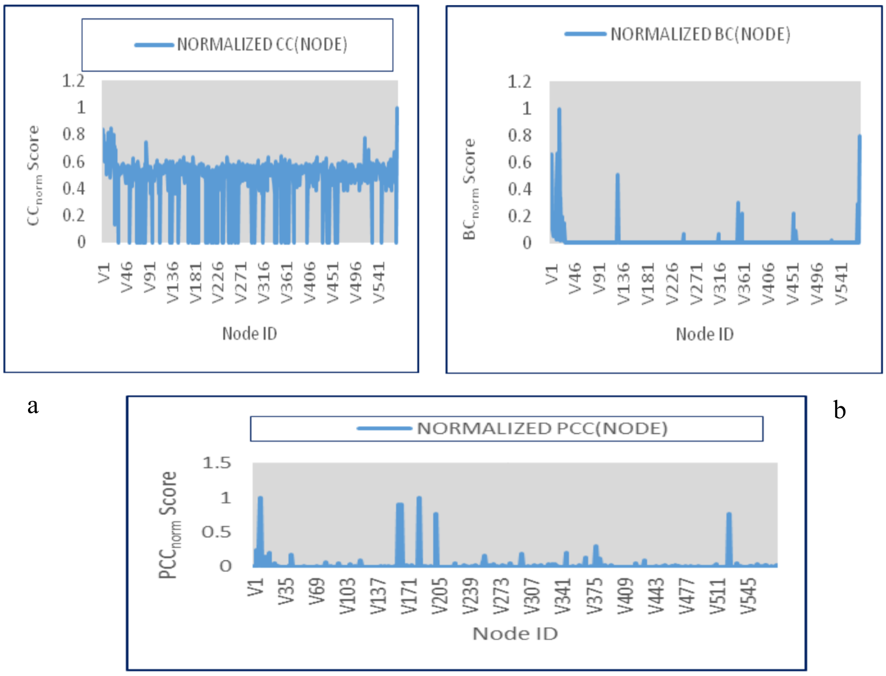
Thus, 691 distinct links with 577 distinct nodes were discovered. As mentioned earlier, privacy of mobile phone users is critical in this analysis. In order to achieve this, each phone number is represented with a new identity number. The new identity number starts with letter "V" and a number ranging from 1 to 577 is attached. To keep track of the 27 seed nodes they are represented using V1, V2, V3……, V27. Next, the communication links (edges) and their corresponding duration (weight) were labelled.

4.2 Extracted Features

The centrality measures for each individual node were determined. Their minimum, average and maximum scores are listed in Table 2. The individual score is normalized and the distributions of the normalized scores among the nodes are illustrated in Figure 1. The top ten node with high closeness centrality, betweenness centrality and principal component centrality scores are presented in Table 3.

Table 2. Minimum, average and maximum scores of the centrality measures

CENTRALITY MEASURES	MINIMUM SCORE	AVERAGE SCORE	MAXIMUM SCORE
CC	0.0464450	0.208766	0.3818100
BC	0	0.00740253	0.529191
PCC	0.00000829	0.018398	1.0025730



a b c

Fig. 1: Distribution of normalized centrality scores across the network (a) normalized closeness centrality scores (b) normalized betweenness centrality scores (c) normalized principal component centrality scores

Table 3: Top ten nodes with high closeness centrality scores, high betweenness centrality scores and high principal component centrality scores

Rank	1	2	3	4	5	6	7	8	9	10
CC	V577	V18	V1	V12	V3	V22	V51 4	V13	V14	V86
BC	V18	V577	V13	V3	V12 6	V12	V20	V2	V35 0	V57 3
PCC	V8	V183	V163	V16 0	V52 5	V20 2	V7	V37 8	V4	V17

4.3 Evaluation of the Improved Influence Factor Scheme

The improved scheme was evaluated by using it to detect the influential node in the constructed graph. Based on the individual influence factor, thirty-nine nodes were detected while the remaining five hundred and thirty-eight nodes have zero influence. The influential nodes detected are shown in Fig. 2. The minimum and maximum influence factor for the mobile phone network is 0.057797486 and 1.85216E-06 respectively.

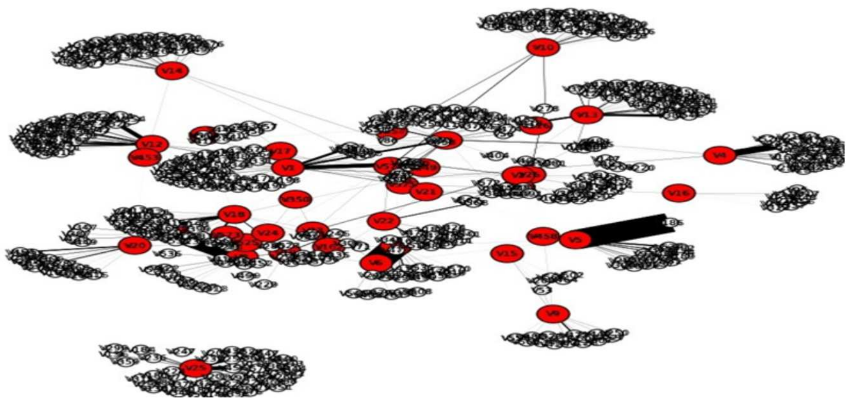


Fig. 2: Location of the identified influential nodes (in red) based on improved Influence factor scheme.

To gain insight into the performance of the improved scheme, a scatter plot was used to show the influence factor of each node for both schemes. This is depicted in Figure 3. According to the existing scheme, sixty nodes were detected as being influential nodes. The two set of influential nodes identified by both scheme is shown in Figure 4. Table 4 summarises the observation from the investigation of the two schemes and Table 5 presents the statistical measures used in comparing the two schemes.

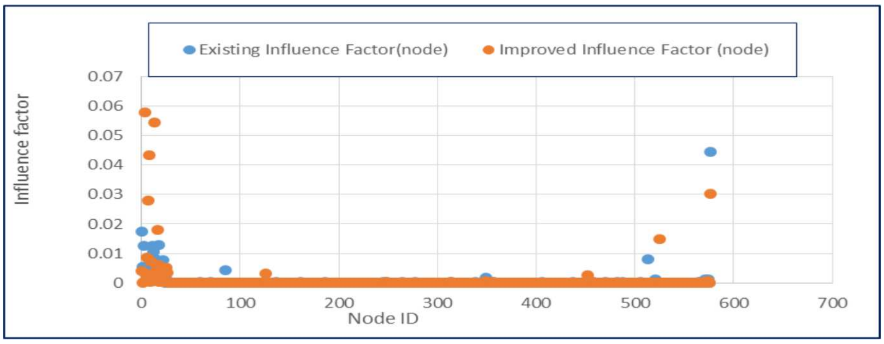


Fig. 3: Comparison between the existing and improved scheme.
Fig 4. Location of the two set of influential nodes detected using both scheme.

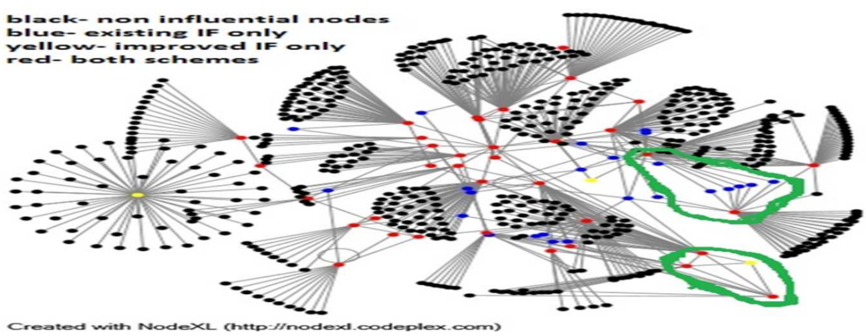
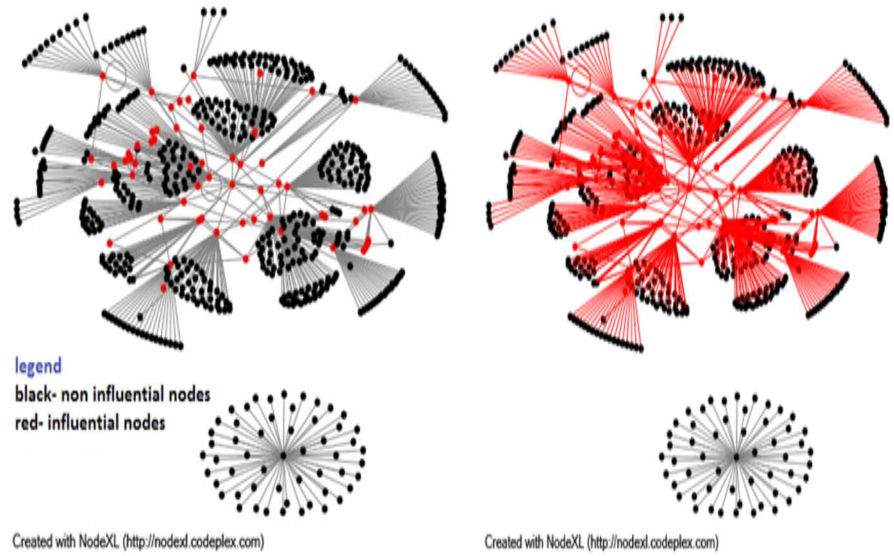


Table 4. Investigation of both schemes

Scheme	True Positive	True Negative	False Positive	False Negative
Existing Scheme	44	2	16	1
Improved Scheme	37	16	2	8

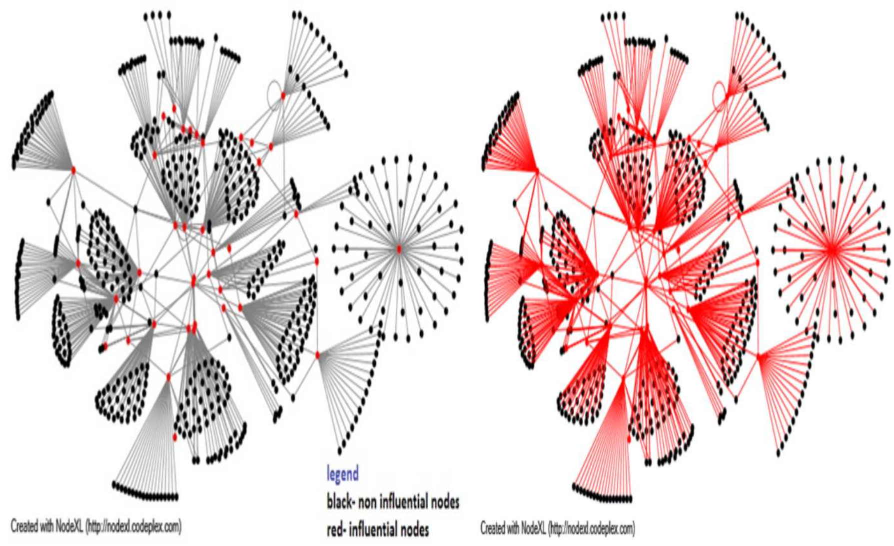
Table 5. Statistical measure of the two schemes

	EXISTING IF SCHEME	**IMPROVED IF SCHEME**
ACCURACY	73.01	84.12
PRECISION	73.33	94.87
SENSITIVITY (Prob. of Detection)	97.78	82.22
F1 MEASURE	83.81	88.10
SPECIFICITY(TNR)	11.11	88.89



a b

Fig. 5 (a): Location of the 60 influential nodes identified by the existing influence factor scheme and (b) nodes accessible by the influential nodes.



a b

Fig.6 (a): Location of the 39 influential nodes detected by the improved influence factor scheme and **(b)** nodes accessible by the influential nodes.

5.0 Discussion

This section discusses the result of the performance of the improved influence factor scheme on a mobile phone network created from the call record dataset. We discovered that identifying high number of influential nodes does not matter especially when these nodes are detected because they are linked to the top influential nodes. The normalized betweenness centrality score distribution discloses that a few selection of nodes are responsible for transferring information from one node to another. This is in contrast with closeness centrality, a node can be surrounded by other nodes but does not necessarily mean that it allows the flow of information. Such node is said to have a betweenness centrality score of 0 and this explains why large number of nodes in the mobile phone networks have 0 betweenness value. The principal component centrality with tuning parameter 'p' equal to 300 (approximately 52% of the 577 largest eigenvalues), nodes with 0 eigenvalues now have a significant principal component centrality value.

The improved scheme concentrated more on nodes that are highly important and reduced the detection of nodes that are not necessary important (with low IF score) by ignoring them. Also, the improved scheme gives a high influence factor than the existing scheme, though it has a less count of detected influential nodes when compared to the existing scheme. It is important to note that most of the nodes that are not discovered by the improved scheme have low IF scores in the existing scheme. These nodes are seen to be influential only because they are connected to top influential nodes and not that they are necessarily important, this can be confirmed from their positions in the network graph. Ignoring these nodes will not affect the network in any way.

Thorough investigation of the two set of influential nodes detected by both schemes reveals that some influential nodes detected only by the existing scheme as shown in Figure 4 (in blue) are connected to the same pair of influential nodes and they are not connected to any other nodes, this is attributable to the fact that the existing scheme incorporated eigenvector which has the weakness of considering only the influential nodes and neighborhoods. The improved scheme outperforms the existing scheme in terms of accuracy, precision, F1 measure and specificity. However, it underperforms in sensitivity (probability of detection).

To further compare the improved scheme, the percentage of nodes that are reachable through the influential nodes identified using each scheme is determined. The nodes detected according to the existing influence factor were able to reach 90.64% of the nodes while the improved scheme was able to reach 100% of the nodes. The existing scheme ignored the smaller component of the graph; therefore no influential node was detected in it which means that there is no way of accessing that component. But the improved scheme detected an influential node in the smaller component through which the nodes in that component can be accessible. Hence in a large graph with many components, the improved scheme will detect all influential nodes in every component.

6.0 Conclusion

The influential nodes detection in a mobile phone network is a difficult job as huge amount of mobiles subscribers (nodes) are connected to the mobile phone network every seconds. In this paper, a method to identify the most influential nodes based on influence factor measure is developed. The basic components of the proposed approach are closeness centrality, betweenness centrality and principal component centrality. The proposed scheme integrates these three centrality measures to improve the detection of influential nodes across the mobile phone network. The results obtained from the experimental analysis and comparison with the existing influence factor scheme showed that the improved scheme is more accurate and precise in identifying influential nodes that can maximally spread influence across the entire mobile phone network, however the probability of detection is slightly lower. The specific recommendations for further studies based on limitations in this research is the collection of dataset which captures more components, more nodes (users) and more connection links (edges). More so, finding what can be done to correct the sensitivity of the improved scheme.

Correspondence
Sheriff Murtala
Department of Telecommunications Engineering
Federal University of Technology, Minna, Nigeria

References

Abraham, A. (ed.) (2012) Computational social networks: Mining and visualization. London: Springer London.

Agarwal, N., Liu, H., Tang, L. and Yu, P.S. (2008) 'Identifying the influential bloggers in a community' In Proceedings of the 2008 international conference on web search and data mining (pp. 207-218). ACM.

Ahsan, M., Singh, T. and Kumari, M. (2015) 'Influential node detection in social network during community detection', In Cognitive Computing and Information Processing (CCIP), 2015 International Conference on (pp. 1-6). IEEE.

Bastian, M., Heymann, S. and Jacomy, M. (2009) Gephi: an open source software for exploring and manipulating networks. ICWSM, 8, pp.361-362.

Bavelas, A. (1950) 'Communication patterns in Task-Oriented groups', The Journal of the Acoustical Society of America, 22(6), pp. 725–730. doi: 10.1121/1.1906679.

Blondel, V.D., Decuyper, A. and Krings, G. (2015) 'A survey of results on mobile phone datasets analysis', EPJ Data Science, 4(1). doi: 10.1140/epjds/s13688-015-0046-0.

Borgatti, S.P. (2006) 'Identifying sets of key players in a social network', Computational and Mathematical Organization Theory, 12(1), pp. 21–34. doi: 10.1007/s10588-006-7084-x.

Brandes, U. (2001) 'A faster algorithm for betweenness centrality*', The Journal of Mathematical Sociology, 25(2), pp. 163–177. doi: 10.1080/0022250x.2001.9990249.

Canali, C. and Lancellotti, R. (2012) 'A quantitative methodology based on component analysis to identify key users in social networks' , International Journal of Social Network Mining, 1(1), pp.27-50.

Catanese, S., Ferrara, E. and Fiumara, G. (2012) 'Forensic analysis of phone call networks', Social Network Analysis and Mining, 3(1), pp. 15–33. doi: 10.1007/s13278-012-0060-1.

Eirinaki, M., Monga, S.P.S. and Sundaram, S. (2012) 'Identification of influential social networkers', International Journal of Web Based Communities, 8(2), pp.136-158.

Erlandsson, F., Bródka, P., Borg, A. and Johnson, H. (2016) 'Finding Influential Users in Social Media Using Association Rule Learning', Entropy, 18(5), p.164.

Everett, M.G. and Borgatti, S.P. (1999) 'The centrality of groups and classes', The Journal of Mathematical Sociology, 23(3), pp. 181–201. doi: 10.1080/0022250x.1999.9990219.

Farley, J.D. (2003) 'Breaking Al Qaeda cells: A mathematical analysis of counterterrorism operations (A guide for risk assessment and decision making)', Studies in Conflict & Terrorism, 26(6), pp. 399–411. doi: 10.1080/10576100390242857.

Ferrara, E., De Meo, P., Catanese, S. and Fiumara, G. (2014) 'Detecting criminal organizations in mobile phone networks', Expert Systems with Applications, 41(13), pp. 5733–5750. doi: 10.1016/j.eswa.2014.03.024.

Freeman, L.C. (1977) A set of measures of centrality based on betweenness. Sociometry, pp.35-41.

Goldenberg, J., Han, S., Lehmann, D.R. and Hong, J.W. (2009) 'The role of hubs in the adoption process'. Journal of Marketing, 73(2), pp.1-13.

Han, B., Li, J. and Srinivasan, A. (2014) 'Your friends have more friends than you do: Identifying influential mobile users through random-walk sampling', IEEE/ACM Transactions on Networking, 22(5), pp.1389-1400.

Heidemann, J., Klier, M. and Probst, F. (2010) 'Identifying key users in online social networks: A pagerank based approach.

Hinz, O., Skiera, B., Barrot, C. and Becker, J.U. (2011) 'Seeding strategies for viral marketing: An empirical comparison', Journal of Marketing, 75(6), pp. 55–71. doi: 10.1509/jm.10.0088.

Ilyas, M.U. and Radha, H. (2010) 'A KLT-inspired node centrality for identifying influential neighborhoods in graphs'. In Information Sciences and Systems (CISS), 2010 44th Annual Conference on (pp. 1-7). IEEE.

Ilyas, M.U. and Radha, H. (2011) 'Identifying influential nodes in online social networks using principal component centrality'. In 2011 IEEE International Conference on Communications (ICC) (pp. 1-5). IEEE.

Landherr, A., Friedl, B. and Heidemann, J. (2010) 'A critical review of centrality measures in social networks', Business & Information Systems Engineering, 2(6), pp. 371–385. doi: 10.1007/s12599-010-0127-3.

Lieberman, M. (2014) Visualizing big data: Social network analysis. In Digital research conference.

McDiarmid, A., Bell, S., Irvine, J. and Banford, J. (2013) 'Nodobo: Detailed mobile phone usage dataset'. Unpublished paper, accessed at http://nodobo.com/papers/iet-el. pdf on, pp.9-21.

Narayanam, R. and Narahari, Y. (2011) 'A shapley value-based approach to discover influential nodes in social networks', IEEE Transactions on Automation Science and Engineering, 8(1), pp.130-147.

Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Menezes, M.A. de, Kaski, K., Barabási, A.-L. and Kertész, J. (2007) 'Analysis of a large-scale weighted network of one-to-one human communication', New Journal of Physics, 9(6), pp. 179–179. doi: 10.1088/1367-2630/9/6/179.

Ortiz-Arroyo, D. and Hussain, D.A. (2008) 'An information theory approach to identify sets of key players'. In Intelligence and Security Informatics (pp. 15-26). Springer Berlin Heidelberg.

Perkins III, F.C., Convergys CMG Utah Inc., 2002. System and method for processing call detail records. U.S. Patent 6,396,913.

Pinheiro, C.A.R. (2012) Community detection to identify fraud events in telecommunications networks. SAS SUGI Proceedings: Customer Intelligence.

Probst, F. (2013) Customer Relationship Management in a Digitally Connected World (Doctoral dissertation).

Reserved, I.A.R. (2016) ITU: Committed to connecting the world. Available at: http://www.itu.int (Accessed: 16 October 2016).

Sathik, M.M. and Rasheed, A.A. (2009) A centrality approach to identify sets of key players in an online weblog. International Journal of Recent Trends in Engineering, 2.

Schult, D.A. and Swart, P. (2008) Exploring network structure, dynamics, and function using NetworkX. In Proceedings of the 7th Python in Science Conferences (SciPy 2008) (Vol. 2008, pp. 11-16).

Seeley, J.R. (1949) 'The net of reciprocal influence; a problem in treating sociometric data', Canadian Journal of Psychology Revue Canadienne de Psychologie, 3(4), pp. 234–240. doi: 10.1037/h0084096.

Shetty, J. and Adibi, J. (2005). 'Discovering important nodes through graph entropy the case of enron email database' ,In Proceedings of the 3rd international workshop on Link discovery (pp. 74-81). ACM.