Analysis of Social Network for Telecommunication Companies

Aktham Sawan
Department of computer science
Al-Quds University
Palestine
Aktham.sawan@students.alquds.edu

Rashid Jayousi
Department of computer science
Al-Quds University
Palestine
rjayousi@staff.alquds.edu

ABSTRACT

Social Network Analysis (SNA) is created to analyze social network data and to find influencers users in these networks, so main companies in data mining filed (such as IBM, SAS, R and python) create their own SNA algorithms to find these important users. The aim of this research is to customize SNA algorithm for telecom companies because the current algorithms are not just design for telecom networks but for any type of networks. When the current algorithms such as IBM SNA is used to find influencers customers from the Palestinian mobile service provider company, Jawwal, many high value customers do not include in the result, and these results just contain 55% from input customers.

In the new proposed algorithm, relation strength and extenders have been used to enhance final results.

To test the new algorithm, 300M records, which belong to around 4 million customers over three months period have been collected from as case study; moreover, the same data set have been used to check IBM SNA models results. In this research, many experiments have been tested based on (call duration, call count and ratio between call duration and call count); in addition, two groups size have been used (15 and 20) and Oracle SQL-PL/SQL have been used to develop the new algorithm.

Parameters that have approved by Jawwal as the following.

1. Ratio for call count multiply by 0.3.
2. Call duration multiply by 0.7.
3. Maximum group size of customers are 20 customers.

Results for the new algorithm have increased coverage of NW to be 75.9% instead of around 55% for IBM algorithm; moreover, all high value customers have included in the results for he ne alg i hm. New novelty ideas have created in this research such as,

1. Extenders, this type of customers are used for customer who is influencer in one group and follower in the other group.
2. Relation strength, this algorithm are used to create groups and assign followers to their most related influencer.
3. Super Group, it is a high level of groups that connect related groups in one group and find super influencer inside it.

KEYWORDS

Social network analysis, algorithms. Telecommunication Companies

1. INTRODUCTION

Social networks become progressively the most important channel to connect people, because social media companies offer their services that make the whole world such as a small piece. Recently big companies work to analyze social network that includes nodes (individuals or customers) and who these nodes are exchange data between each other such as (voice, data, SMS, video values, ideas, visions, financial exchange, kinship, friendship, dislike, trade or conflict, ..Etc.) [1-5].

The general definition of SNA (social network analysis) is how to create groups for nodes based on their links and communications; in addition, how to draw relationships between these nodes [3], [6].

Several domains for social network analysis (SNA) such as (computer science, sociology, mathematics and physics) lead to different methodologies and tools, and many programs create to study and manipulate such networks [4].

SNA is using in marketing fields such as Retention, attrition, prevention, Churn, Segmentation, Acquisition new customers, Fraud detection, up sell and cross selling [3].

Social communities are recognizing depends on two aspects. Firstly depends on the behavioral relationships between nodes (customers).

Secondly depends on nodes that are segments and measures to find influencer and disseminator nodes. By using SNA customers can be targeted depending on their behavior and changing their status within communities, (e.g., an influencer for a group is Churn then followers for this node should be targeted to save them from churn) [7], [8].

Below points, explain some of SNA advantages:

1. Identifying which individuals are playing significant roles (knowledge brokers, leaders, information managers, etc.).
2. Distinguishing information, bottlenecks, breakdowns and structural holes as well as isolated units, individuals, and teams.
3. Creating opportunities to fast-track flows of knowledge through organizational and functional boundaries.
4. Strengthening the effectiveness and efficiency in the existing channels.
5. Raising alertness of important networks and groups to find ways to improve performance.
6. Enhancing strategies. [7], [8].

The present thesis aims to create a new SNA algorithm that customized for telecommunication companies. The main idea for this algorithm is using relation strength as a main factor for creating groups to enhance the results of coverage and quality. The new algorithm increases network coverage comparing with current algorithm; moreover, results for the new algorithm include most of high valued customers comparing with minimal number of them in results than coming from current algorithms.

2. RELATED WORK

In general, the current SNA algorithms are not suitable for all field, so many researchers worked to enhance these algorithms. The Authors in [9] proposed a collective model using user based and content-based approach beside with a method for centrality measurement to enhance centrality in SNA. Authors in [10] designed a hybrid model based on ANN and Fuzzy techniques for opinion recommendation system that focused in communities and users respectively to enhance and classify social data. Authors in [11] improved SNA results to discover and rank terrorists and their target lists by using Analytical Hierarchy Process (AHP).

Authors in [12] suggested a new method called social network based on engineering education (SNEE), to analyze the students relations and evaluate students learning, also to organize method of education in social NW. In [13] authors worked to create a new algorithm. The main aim of this algorithm was to find better central nodes (influence nodes) that calculated by using centrality measurement in SNA algorithm; in addition, they did that by using fuzzy interface system. In [14] author developed a new method by using a multidimensional scale algorithm to differentiate between various factors and methods in social NW.

Researchers in [15] focused on improving performance of algorithm by assigning communities to SN and using hybrid approach to do that. Researchers in [16] worked to use SNA instead of using current tools to analyze transportation, because the old methods consumed time and so expensive; moreover, it needs rigorous data to have reliable results; however, SNA can deal with the connectivity and complexity issue of NW more effectively.

Authors in [17] proposed the efficient method OWL DL to classify different users. This method utilized semantic in web technology for describing target rules, then they used FOFA and SNA to define users' domain. Authors in [18] focused on the huge amount of data in social networks that could cause too many elements for the visualization of data; in addition, it was a challenge when discovering the information. Authors represent Chord diagrams to raise the level of abstraction. In paper [19] author proposed a new methodology to involve human interactions and analyzed it by designing online SNA software that analyzes and elicits the behavior of requirements to create stronger online SNA sites.

Can be found from theses researchers that authors worked to enhance either SNA algorithm or its performance, because the current SNA algorithms need to be more customized and specified fields such as telecommunications companies that this research aim to customize SNA for it.

3. BACKGROUND

IBM, SAS, R and Python companies created SNA models to analyze networks even social NW or telecom companies NW. These models focused on groups and do not focused on individual users.

It created groups consisting of individuals who are related with each other based on their calls or SMS transactions in telecom.

3.1 SNA form IBM

Results of SNA that built from IBM Company have groups; each group contains influence and dissemination users for incoming and outgoing calls and some of followers.

Marketing campaigns target influencers and disseminators users. [20].

Entities that should be used as input for the SNA model are.

- Users: - members in networks.
- Value: - weight of relationship between users like the following **Fig 1** (value of relation between users, in telecom for example call duration plus number of SMS) [5].

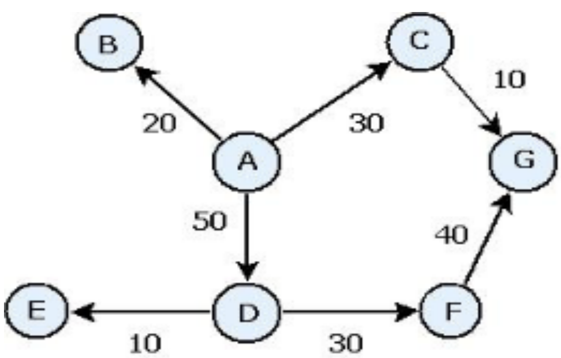

Figure 1: Sample of related nodes [5].

Can be seen in **Fig 1** that A calls B, C and D in different weights, based on these weights many measures can be determined by using the SNA model [7].

1. When SNA model creates a group based on inputs each group has two main characteristics (density and degree)
2. Density: - equal number of relations between users in a group, divide in all possible relations.

   For example:-as appears in Fig 2. The density in group A equals 7/ (7*6) =0.17 and the density in group B equals 42/ (7*6) =1.0. Relation between individuals in group B stronger than in group A.
3. Degree: - each node in a group has in degree and out degree, **Table 1** represents in degree and out degree for each node.

Can be found that node A has a greater number of out degree, which means it the central node; in addition, can be found that node G has a greater number of in degree nodes, which means itsthe prestige node [20].

When SNA model creates groups, it used the below steps.

1. Determine similarity for nodes to include it in one group

   1.1 As can be seen in **Fig 2**. Can be found that node A has relations with (B, C, D, E, F, G), and node B has relations with (A, C, D, E, F, G) and node C has relations with (A, D, H, I, J).

Table 1: Sample of in and out degree for one group

Node	Degree	In-degree	Out-degree
A	3	0	3
B	1	1	0
C	2	1	1
D	3	1	2
E	1	1	0
F	2	1	1
G	2	2	0

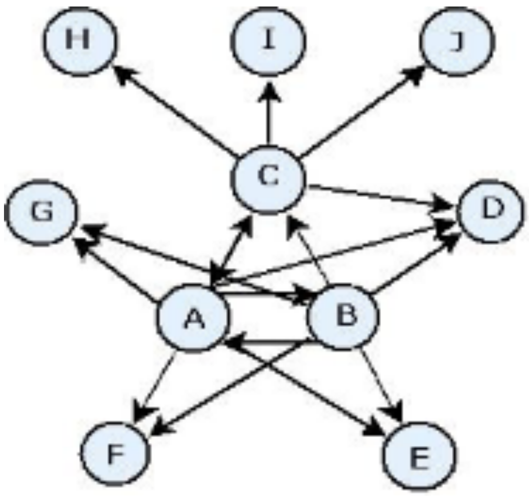

Figure 2: Sample of relations between nodes [5].

1.2 From that can be found the node A and B calling the same 5 nodes(C, D, E, F, G); in addition can be found that node A and C call the same one node (D). Therefore, node A and B are more similar than node A and C and may have same group [5].

2. Partitioning node into groups to partition subscribers to groups, three parameters use for that.
   2.1 Firstly, minimum group size, to determine minimum number of subscriber for each group.
   2.2 Secondly, maximum group size, to determine maximum number of subscriber for each group.
   2.3 Thirdly, coverage threshold, to determine relationship weights between subscribers in the same group based on similarity. For example if coverage is set (40%) that means stronger (40%) of relation between subscribers will be in the same group. Sometimes some groups contains more than maximum group size, in such cases SNA divides Groups into sub groups that suitable with the maximum group size. [7].
3. Describer groups and members in these groups.
   3.1 Each group has density, in degree and out degree. If group in Figure Fig.1 was taken as example. Can be found that density equal 14/42=33%, in degree equal 7/42 =17%, out degree equal 7/42=17%. In users side authority and dissemination users will be determined. High in-degree in the group will be authority user in our example is (G) .High out-degree will be dissemination in our example is (A) [20].

3.2 SNA from R

SNA from R Company are used to categorize main actors within groups. To identify actors, several centrality metrics calculate in NW as the following:-

1. Degree: - calculating number of connections for each node (in-degree, out-degree).
2. Betweenness: - calculating total of shortest paths between nodes.
3. Closeness: - finding distance between one node and other nodes Eigenvector: - finding centrality with weighted degree (count more incoming links from highly central nodes). [21], [22].
4. Density: - calculating number of real connections in NW divide to number of possible connections in NW.
5. Reciprocity: - finding amount of nodes that are symmetric.
6. Mutuality: - calculating number of complete transactions.
7. Transitivity: - summation number of transitive trios.
8. Graph mean: - finding average of graph based on graph density.
9. Centralization: - calculating this measure based on degree, closeness and Betweenness. [22]

R Company has diversity of tools to analysis SNA based on graphs such as bipartite networks and random graphs; moreover, it is easy to calculate a degree and weight networks by using built-in visualization tools; In addition, tools from R Company that used in built-in graphics are extracted immediately a statistical analysis to achieve network and econometrics SNA in the same roof. [21]

3.2. SNA from SAS

SAS Company has a wide diversity of capable tools such as SAS/Internet that build reports for integrating and utilizing groups. In SNA, it has SNA influential tools that identify link between members in graphical and mathematical analysis. It's focused on links between members persons of interest". This approach of analysis has vast practical prominence in fields like fraud and epidemiology analysis [23].

Main measures in SNA from SAS.

1. In-Degree: - calculating number of incoming links for each subscriber.
2. Out-Degree: - calculating number of outgoing links for each subscriber.
3. Centrality: - summation of in and out degree.
4. Density: - calculating density of each group based on number of links in group divide to maximum probable links that can be exist in each group.
5. Prestige: - calculating prestige based on divide number of culms to number of rows in NW. [24].

3.4. SNA from python

Python Company create an SNA algorithm to find influencers nodes that connected with followers nodes via links. Many measures and packages are used to analyze the NW and find relations between nodes. Measures that use in SNA algorithm. [25] Degree (in, out): calculating incoming/outgoing edges for each nodes.

1. Centrality: - summation of in and out edged.
2. Density: - calculating density based on number of links in group divide to maximum probable links that can be exist in it.
3. Closeness: - finding distance between nodes and other nodes.
4. Betweenness: - calculating total of shortest paths between nodes.
5. Eigenvector: - calculating this measure based on centrality with weighted degree (count more incoming links from highly central nodes). [26].

4. RESEARCH METHODOLOGY

The current research involved customizing SNA for telecom companies. Input data for this research was collected from two companies in telecom field. Main ideas in SNA from SPSS IBM have been used in this new model.

4.1. Data collection

Around 300 million records were collected and analyzed from Jawwal Network data. All incoming, outgoing calls and SMS for whole Jawwal customers in the last three months were included in input data for this model. Input data contained sender and receiver number with values between them, but these numbers were encrypted for security purposes. Values between customers were calculated based on calls duration and count of calls and SMS, every two SMSs counted as one minute of call. Systems data were removed such as monitoring and broadcasting messages.

4.2. Method

The main goal of SNA models is to create groups for customers and find important customers (influencers and disseminator) as can be seen in **Fig 3**. The below steps were used to find them.

As can be seen in **Fig 3** the subscriber in red is influencer subscriber and the others subscriber in blue are flowers, arrows explain who are they communicate in one group.

1. Calculating the degree of centrality for each node based on the degree distribution for them. Values between customers contain calls duration and the number of calls and SMS (two SMSs counted as one call).
   Example if customer number 599??33?3 C1 has 5 outgoing calls(relations) with other customers.
   To calculate degree of centrality in this group, an equation between calls duration and count of calls were used as following. C1 is a customer and has the following outgoing values to other customers
   C2 (5 duration, 2 count of calls),C3(10 duration, 2 count of calls) ,C4(5 duration, 2 count of calls) ,C5(8 duration, 4 count of calls) ,C6(7 duration, count of calls).
   Degree of centrality equals summation of calls duration multiply by 0.7 and summation of calls counts multiply by 0.3. that equals =(35*0.7) +(12*0.3) =24.5+4 =28.5.

Figure 3: Influencers and disseminator in the groups

By using mathematics notations.
Two arrays were used as following
Duration Array values are DARY{5,10,5,8,7}
Count Array values are CARY{2,2,2,4,2}
to calculate degree of centrality for C1

Deg_of_Cent= $\sum_{i=1}^{n-1} DARY[i] * 0.7 + \sum_{i=1}^{n-1} CARY[i] * 0.3$

2. Ranking all nodes in networks, adjacency matrix for directed graph was used for whole four million customers based on degree of centrality.
3. Calculating relation strength between every two nodes. Example if total outgoing calls for A equals 100 minutes and A calls 5 numbers as following (B 15 minutes, C 10minutes, D 15 minutes, F 20 minutes and E 40 minutes).Relation strength between A and B =15/100 =15%. Maximum relation strength for A is 40% and it's between A and E

By using mathematics notations.
Array values are ARY{15,10,15,20,40}

To calculate relation strength

Rel_STR= $\sum_{i=1}^{n-1} ARY[i]$

Algorithm 1: to calculate relation strength

```
Res←0
For i←0 to n-1
Res←ARY[i]/X
Return Res
```

4. Finding the most important customer for each customer.
5. Creating communities for most important customer, who has maximum number of flowers in networks.
6. Calculating closeness centrality to determine disseminator customer, who has the highest number of outgoing relations with the other customers in the same group, and influencer customer, who has the highest number of incoming relations with the other customers in the same group.
7. Ranking customers for each group based on incoming and outgoing relations.
8. Calculating many measures for each group (such as size, density, max rank for dissemination, min rank for dissemination , group max rank for influencer ,group min rank for influencer ,group in degree and group out degree).

A new type of customers were created in this research, called Extender customer. This customer is disseminator or influencer in a group and follower in another one as can be seen in **Fig 4.**

A new type of groups was created in this research based on Extender customer, called Super group. To calculate super groups the following steps were used.

1. Calculating the number of extenders in each normal group.
2. Calculating the degree of centrality of normal groups Example if group number1 G1 has five customers, to calculate degree of centrality in this group, an equation between calls duration and count of calls were used as following. G1 is a group and has the following customers with their data.

C1 (10 duration, 5 count of calls), C2 (20 duration, 10

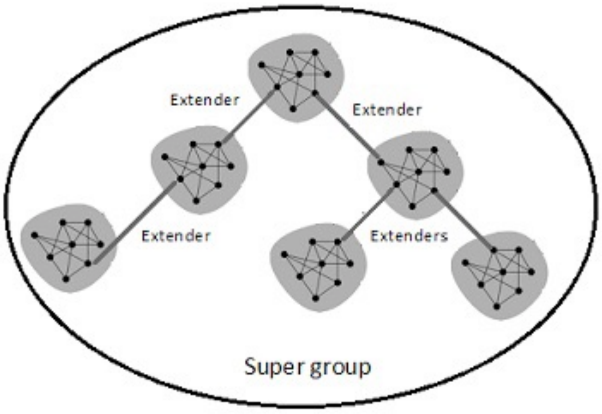

Figure 4: Super Group

count of calls), C3 (15 duration, 10 count of calls), C4 (25 duration, 5 count of calls), C5 (35 duration, 15 count of calls).
Degree of centrality equal summation of calls duration of all customers multiply by 0.7 and summation of calls counts for all customers in a group multiply by 0.3. that equals (105*0.7) +(45*0.3) =73.5+13.5=87
By using mathematics notations.
Two arrays were used as following
Duration Array values are

DARY{10,20,15,25,35}

Count Array values are

CARY{5,10,10,5,15}

to calculate degree of centrality for G1

Deg_of_Cent= $\sum_{i=1}^{n-1} DARY[i] * 0.7 + \sum_{i=1}^{n-1} CARY[i] * 0.3$.

3. Ranking normal groups that were calculated based on number of extender and degree of centrality of each group.
4. Generating Super groups by using the highest rank of normal group (that have maximum number of extenders) and hierarchical network design that was mixed with extenders as can be seen in **Fig 4.**
5. Calculating closeness centrality for each super group to determine disseminator customers, who have the highest number of outgoing relations with the other customer in a super group, and influencer customer, who have the highest number of incoming relations with the other customers in a super group.
6. Ranking customers based on incoming and outgoing relations to find influencers and disseminator.
7. Calculating many measures about each super group(such as size, density, max rank for dissemination, min rank for dissemination , group max rank for influencer ,group min rank for influencer ,group in degree and group out degree).

4.3. VALIDATION

Different parameters have been used to validate results in this research, and these parameters are calls duration, calls count and mix between calls duration and calls count; moreover, groups size have been used in two sizes (15 and 20) because the average of community size in Jawwal is 18.2 customers.

Real campaigns have been created to validate quality of results for SNA model in this research comparing with results of SNA model from IBM SPSS.

5. EXPERIMENT AND RESULTS

In this research, hundreds of experiments were used to find optimal parameters that should be included in final version.
All current data mining models such as association, clustering and classification were tested to find if it suitable to use it in any stage of this research, but results were approved that these model not suitable for this type of algorithms. From many parameters, mix between calls duration and calls count with groups size equal twenty have been given the best results in this research.

SNA algorithm in this research VS SNA SPSS from IBM based on centrality ratio between calls duration and calls count with group size equal 20.

In this experiment both proposed algorithms and SNA SPSS from IBM model were compared by using different parameters as the following.

1. Degree of centrality was calculated based on groups sizes that should be minimum 2 and maximum 20 and data that should be based on mix between calls count and calls duration.
2. Closeness centrality was calculated based on summation of volume of

incoming and outgoing calls that happened between nodes. The main points that can be seen in Table 2 , when proposed algorithm and SNA SPSS from IBM are compared.

Table 2: Results for group size between 2 and 20 and centrality based on ratio

Measure	IBM SPSS	Proposed model
Nodes in groups	2,198,340	3,039,095
Links in groups	4,713,664	6,715,705
Number of groups	186,652	276,430
Average group size	11.77	11.95
Groups density	0.21	0.22
Average of In and Out degree	1.99	2.03

3. Results of proposed model cover around 75.9 % from subscribers in input network, but results of SNA SPSS from IBM just cover 54.9% from subscribers in same input network. Between volume and Count.
4. Links between nodes in proposed model are 6.7 M, but in SPSS SNA are 4.7 M, that mean number of links increase by 42.
5. Number of groups in proposed model are 276K, but in SNA SPSS are 186K, that mean number of groups increase by 48

The main points that can be seen in Table 3, for super groups when group size between 2 and 20 and centrality based on ratio between volume and calls count.

Table 3: Super group's results for group size between 2 and 20 and centrality based on ratio between volume and calls count

Measure	Proposed model
Nodes in groups	2,793,238
Links in groups	7,067,607
Number of Extenders	266,105
Average super group size	140.35
Average influencers in super group	13.3
Average of In and Out degree	2.52

1. Extenders customers, new type of customers (influencer in a group and follower in other one), are 266K.
2. Average of super group size is 140 customers.
3. Each super group contains of around 13.3 groups and influencers.
4. Links between nodes are 7 M link.

Main findings from experiments in this research are the following.

1. The highest coverage of network was retrieved when calls count was used to calculate degree of centrality.
2. The lowest coverage of network was retrieved when calls duration was used to calculate degree of centrality.
3. When calls duration was used to calculate degree of centrality, groups were more representative for telecommunication companies' real groups.
4. Balanced ratio between call duration and number of call was created (0.7 for volume, 0.3 for number of transactions).
5. The middle coverage of network was retrieved when ratio between call duration and number of call was used to calculate degree of centrality; in addition, it gave suitable groups structured for customers in telecom companies.

6. CONCLUSION

SNA algorithms that created from commercial companies such as IBM focused to create a general algorithm to be used for many fields such as (social media, telecom and transportation), but when these algorithm were used for telecom, final results were covered around 55% from input network. These algorithms created for general use and did not specify for telecom companies.

These findings were the main motivation to create a SNA algorithm that specified for telecommunication companies uses; in addition, it created by using business ideas that coming from this domain. When the new algorithm have been created many points have taken in consideration such as increase number of users in final results and cover all high valued customers inside it.

Different parameters have been used to validate results in this research, and these parameters are calls duration, calls count and mix between calls duration and calls count; moreover, groups size have been used in two sizes (15 and 20) because the average of community size in Jawwal is 18.2 customers that have used as case study, and the new algorithm have created by used Oracle Sql-PL/SQL.

Result from experiments in this research are the following.
The highest coverage of network was retrieved when
calls count was used to calculate degree of centrality. The lowest coverage of network was retrieved when calls duration was used to calculate degree of centrality. When calls duration was used to calculate degree of centrality, groups were more representative for telecommunication company's real groups. Balanced ratio between call duration and number of call was created (0.7 for volume, 0.3 for number of transactions).The middle coverage of network was retrieved when ratio between call duration and number of call was used to calculate degree of centrality; in addition, it gave suitable groups structured for customers in telecom companies, also final output have increased to be 75.9% instead of 55% when SNA SPSS from IMB has been used; moreover, all high valued customers have included inside final output.

These results have approved and the objectives have achieved; moreover, new novelty ideas have created in this research such as, firstly Extenders, new type of customers (influencer in a group and follower in other one). Secondly, Relation Strength that use to create groups and assign followers to their most related influencer. Thirdly Super Group that use as a new layer to connect related groups in one super group.

The new algorithm has applied in Mobily Company in Saudi Arabia, and the same positive results have been found such as Jawwal. When new parameters were added, the results be more efficient and network coverage was increased.

When new parameters have been added the results have been more efficient and number of used in final results have increased, so future work can be focused to add and test another parameters to enhance network coverage and groups distributions.

REFERENCES

[1]Rob Cross, Rew Parker and Stephen P. Borgatti.. 2010. A bird's eye view: Using social network analysis to improve knowledge creation and sharing

[2] Jennifer Roberts. 2015. A brief introduction to social network analysis. http://web.mit.edu/vdb/www/6.977/l-jenn.pdf.

[3]Dan McKenzie, 2010, Social network analysis, https://www.sas.com.

[4] David Combe , Christine Largeron, El˝od Egyed-Zsigmond2 and Mathias Géry. 2010 .A comparative study of social network analysis tools," Laboratoire Hubert Curien, vol. 551.

[5]IBM, 2012. Ibm spss modeler social network analysis 15 user guide.. ftp://public.dhe.ibm.com.

[6] J.C.Thomas, W.A Kellogg and T. Erickson. 2001. The knowledge management puzzle: Human and social factors in knowledge management. IBM Systems Journal, vol. 40.

[7]Olivier Serrat. 2009. Social network analysis. http://betterevaluation.org/sites/default/files/social-network-analysis.pdf:

[8] Scott Ponce.. 2011. The overseas development institute's research and policy in development. http://www.fao.org.

[9]Warih Maharani and Alfian Akbar Gozali..2015. Collaborative social network analysis and content-based approach to improve the marketing strategy of smes in Indonesia. International Conference on Computer Science and Computational Intelligence, vol. 59, pp. 373-381.

[10] Nilesh KUMAR Dubey, 2016. Fuzzy and ann based mining approach testing for social network analysis. Second International Conference on Information and Communication Technology for Competitive Strategies, vol. 6, no. 3, pp. 48–52.

[11]Pankaj Choudhary, 2016. Ranking terrorist nodes of 9/11 network using analytical hierarchy process with social network analysis. .International Symposium of the Analytic Hierarchy.

[12]Goran Putnik, Cátia Filipa Veiga Alves, M. L. Rocha Varela and Vaibhav Shah, . 2016. Analyzing the correlation between social network analysis measures and performance of students in social network based engineering education, International Journal of Technology and Design Education, vol. 26, no. 3, pp. 413-437.

[13] Fereshte-Azadi Parand, Mohsen Gorzin and Hossein Rahimi. 2016. Combining fuzzy logic and eigenvector centrality measure in social network analysis. International Journal of Technology and Design Education, vol. 459, no. 3, pp.24-31.

[14] John Stevens. 2010. Comparing social networks: Comparing multiple social networks using multi-dimensional scaling.

[15] Mohammed Alandoli, Mahmoud Al-Ayyoub, Mohammad Al-Smadi, and Yaser Jararweh, 2016. Using dynamic parallelism to speed up clustering based community detection in social networks. Future Internet of Things and Cloud Workshops (FiCloudW), IEEE International CONFERENCE. pp. 240- 245.

[16]Islam H. El-adaway, Ibrahim S. Abotaleb and Eric Vechan. 2016. Social network analysis approach for improved transportation planning. Journal of Infrastructure Systems.

[17] Jeungmin Lee, Hansaem Park, Kyunglag Kwon And In-Jeong Chung. 2016.. An approach for user interests extraction sing decision tree and social network analysis.. Advanced Multimedia and Ubiquitous Engineering, pp. 551–560.

[18] Amin Jalali.2016. Supporting social network analysis using chord diagram in process mining. Perspectives in Business Informatics Research, pp. 16- 32.

[19] Faiza Ghafoor and Muaz A. Niaz. 2016. Using social network analysis of human aspects for online social network software.

[20]IBM.2015. Ibm spss modeler social network analysis 18.0 user guide.

[21]Drew Conway. 2009. Social network analysis in r.new York university -department of politics. http://files.meetup.com/sna.pdf.

[22]George Zhang, 2014. Social network analysis with r sna package,. http://files.meetup.com/1406240/sna.pdf.

[23]Shane Hornibrook and Charlotte. 2007. Degrees. of separation: Social network analysis using the sas system,.

[24]Efraim Darom. 2012. Social network analyses with sas/iml. http://www.sascommunity.org/seugi/SEUGI1997/DAROMDATAMIN.PDF.

[25]Marcelo Ponce. 2015. A basic introduction to complex networks, using PYTHON.

[26] Scott A.Hale,. 2016. Network analysis, oxford internet institute.