International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)
ISSN: 2305-0543 (Online), ISSN: 2411-6173 (Print)
Vol. 10(38), Oct. 2020, PP. 4762-4767

# Improving micro-macro businesses on social networks using business intelligence solutions

1st* Gholamreza Shah Mohammadi
Islamic azad university (E-Campus)
Tehran, Iran
shah_mohammadi@yahoo.com

2nd Sadaf Rakhshade Tali
Islamic azad university (E-Campus)
Tehran, Iran
s.rakhshandehtali@gmail.com

***Abstract*—Throughout history, people in every society have always been thinking of buying and selling goods and creating markets. By the advent of the Internet, the concept of e-commerce has made it possible to design e-business positions around the world. With the boom of social communities, e-business owners have targeted these communities and have always sought to grow their customers among social network users and used these networks to advertise their products and services Increasing data has led to the development of a set of techniques and tools called business intelligence to convert bulk data from different sources into meaningful information to facilitate decision-making in e-commerce. This research introduces a criterion for identifying central nodes of the network to improve micro / macro businesses in social networks The benchmark suggested as a solution to business intelligence, identifies the most important nodes in the social network using the criteria of PageRank, Betweenness centrality, Degree centrality, Closeness centrality, and it recommends to apply these nodes for business intelligence strategies. The intimated solution of this research is evaluated in a data mining process on data located on the Facebook social network. The results show that the proposed hybrid criterion is capable of identifying central nodes in social networks and can be used in line with business intelligence goals and strategies**

***Keywords— Social Network, Business Intelligence, Hybrid Criteria, Page Rank, Betweenness centrality, Degree centrality, Closeness centrality***

## I. Introduction

Many enterprises and local authorities are still using traditional methods for acquiring knowledge to make strategic decisions, by collecting data from questionnaires[1].Firms need to grow optimally with the available technology and capabilities to continue their activities in the competitive arena. Nowadays, IT-based analytic and business applications have joined the modern mainstream of business intelligence, and traditional vendors have turned to new, innovative approaches of BI [2]. Business Intelligence (BI) is a set of techniques and tools that help to transform big data from different sources into meaningful information to support decision making and improve organizational performance. Today, business intelligence is considered as a major driving force for organizational performance, and most organizations use business intelligence to analyze data. Technically, business processes include Extract, Transform & Load (ETL), Data Warehouse, Online Analytical Processing (OLAP), Data Mining, Decision Model and Visualization.

Information management and decision making are two key necessities to make changes in legal environments. Business intelligence reduces information dispersion, user interaction, easy access to information, timely dissemination of information, and decisions about organizational change in business. Business intelligence affects both necessity and management of information in decision making. Information management also provides structural information for decision making [3]. Domo, a provider of real-time business data access services, interprets the amount of data generated in a minute from 2016 on social media entitled "Data Never Sleeps" [4]. While data flowing rapidly in the digital environment, social networks enable people to access information quickly, yet it is difficult to distinguish between the different information that is constantly being collected [5].

## II. Background

### A. Business Intelligence

Business Intelligence (BI) is an umbrella term for strategies, technologies and information systems used by companies to extract big and diverse data based on value chains and knowledge to support a wide range of operational, tactical and strategic businesses of decision-making [6].

Business intelligence is recognized as an effective factor in the performance of organizations, extracting information and knowledge hidden in data sources. It proposes plans that lead to the achievement of organizational goals. Investigating research highlights the important role adaptation of knowledge management based on business intelligence solutions in managing the challenges of organizations [7].

### B. Business Intelligence and Data Mining

Data mining means the extraction of cache information or specific patterns and relations in large volumes of data in one or more big database. Data mining or knowledge discovery in database is the process of extracting the information needed and used by the user from a large set of data. This huge data set has its own challenges from storage to processing and memory recovery [8]. Data mining involves the use of sophisticated analytical tools to discover valid and unknown patterns and dependencies in massive datasets. As a result, data mining technology is more than just data collection and management and involves analysis and prediction [9].



**Article History:**
Received Date: May. 14, 2020, Accepted Date: Jul. 11, 2020, Available Online: Sep. 01, 2020

Data mining in e-commerce is a vital way to change the position of the company and determine the investment with the information the business needs. Many companies have recently adapted to e-business and started processing their data and storing data in databases, of course, the only way to get the greatest use of this data for making decisions is business intelligence [10]. In recent years, business needs in the academic field have attracted much attention to the analysis of social networks. Today, this powerful tool is being used not only by IT professionals but also researchers in other disciplines such as education, biology, communication science, economics and so on benefit from it as a key technique [11]. Social Media sites like Facebook, Twitter, LinkedIn, and Google + contain an outsized volume of unprocessed information. By analyzing this knowledge, new information will be gained. The traditional data processing techniques won't be applicable for processing the increasing number of dynamic and unstructured data[12].

### C. *Methods for identifying central nodes in the social network*

The centrality of the node in the social network indicates the importance and influence of that node in the network. It can be said that high-centrality nodes are the most influential nodes in the network, which usually have a greater impact on information propagation. For evaluating the centrality of nodes in social networks, ranking criteria, closeness, betweenness and Eigenvector are the most commonly used [13, 14, 15, and 16].

#### *1) Degree -centrality measure*

The first and simplest concept of centrality is the degree, which is defined as the number of event links per node. In a directional grid, two distinct criteria of centrality are usually defined, they are the input and output degrees. Degree centrality of V vertex to the V graph G = (V, E) with | V | vertex and | E | edges are defined as 1relation.

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

#### *2) Betweenness centrality measure*

Betweenness centrality criterion specifies the number of times a node acts as a bridge along the shortest path between two other nodes and it has been introduced as a criterion for quantifying one person's control over the relationship between other characters of betweenness centrality (BC) measure in a network. A node that is connected to more nodes will exist in many of the shortest possible paths between pairs of other nodes.

$$C_B(V_i) = \sum_{i \neq j \neq r \epsilon V}^{N} \frac{k_{jr}(V_i)}{k_{jr}} \tag{2}$$

#### *3) Closeness Centrality measure*

Closeness Centrality is a measure that determines how long the information will spread from node to node. According to this definition, one node in the network has the highest closeness criterion if its total distance with all other nodes is the smallest value [17]. The value of Closeness Centrality is obtained from the relation (3).

$$C_u(u) = \frac{1}{\sum_{v \neq u} d(u,v)} \tag{3}$$

#### *4) Page Rank*

One of the spectral indices that are nowadays discussed for use in Google's ranking algorithm. Page Rank works by counting the number and quality of links on a page to get an approximate estimation on the importance of a website. The basic assumption is that more important websites are likely to receive more links from other websites. This property is calculated from the relation (4).

$$p = \alpha pA + (1 - \alpha)v \tag{4}$$

So that A is the li-normalized proximity matrix of the graph, $\alpha \in [0. 1.]$ It is a conversion factor and v is a priority vector. It is claimed that the page rank score is a probability of distributing web pages, which has 11 norms, but if A has empty rows, it is not necessarily true. Scientific papers suggest several ways to solve the problem by randomly converting matrix A into 16. One common solution is substituting empty rows with its own priority vector V, but other solutions have been proposed [18].

## III. Related Work

In 2019, Robles et al [19] presented an article titled 30 Years of Smart Models in Management and Business: A Bibliometric Review. This study provides a conceptual analysis of existing definitions of intelligence in the literature by defining key bibliometric performance indicators, identifying authors and research areas and evaluating development contexts using SciMAT as bibliographic analysis software. The study evaluated 6,392 original research articles published between 1998 and 2017 using this software. Research topics that have contributed the most to these articles are Business Intelligence, Business Information System, Collective Intelligence, Competitive Intelligence, Data Warehouse, Innovation, Knowledge Discovery, Machine Learning, Market Intelligence, National information, Neural network, organizational performance management, query, semantic web, social network, supply chain management, information technology, text mining and web search.

Elaine et al [20] presented an article in 2019 on the Acceptance, Use and Success of a Business Intelligence System (A Principal Review). This study provides a comprehensive overview of what has been considered in terms of acceptance, use and success in the business intelligence system using a principled review. The number of sources studied is shown in Figure (1).

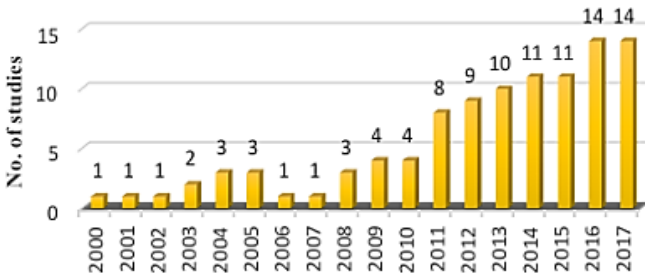


Fig. 1. Publications from 2000 to 2017 in the field of business intelligence [20].

This study shows that little empirical research has focused on user issues and reports that the user is one of the key challenges in the business intelligence system.



IEEE Office - SEET Complex, Federal University of Technology, Owerri, Nigeria, https://aeuso.org
EISSN: 2305-0543, PISSN: 2411-6173

In 2018, Muntean presented an article on business intelligence issues for sustainable projects. This article introduces sustainability as an integral part of a company's business and emphasizes the need for integration at all levels: business model, performance management system, business intelligence project and data model. In this paper, the business intelligence framework that facilitates complex analytical processing is nowadays known as the Business Analytics (BA) framework. The results also show that in big data environments, business intelligence is taking on new forms, and business intelligence tools, which have now become big data analytics, can understand this huge amount of data. Whereas in 2017, 53% of companies acquired big data analytics.

In 2017, Santasso presented an article on a data warehouse with Big Data technology for higher education. This article provides a comparison of traditional data warehouse and modern data warehouse. The goal in traditional data warehousing is to aggregate data collected for a particular business area and support the decision-making process while structured, semi-structured and unstructured data processing from diverse sources of information with the volume of information is more than the ability of traditional tools to store, manage and analyze them.

## IV. Suggested Method

This research seeks a solution from business intelligence to help to improve micro-macro businesses on social networks. It uses data mining and text mining techniques in this regard. The proposed method is a business intelligence approach in which a data mining process is also followed. Therefore, this method is first described following business intelligence, then the data mining process is provided in detail.

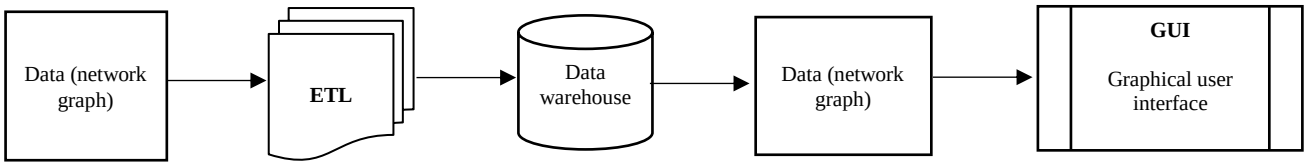


Fig. 2. Conceptual model of research methodology based on business intelligence processes

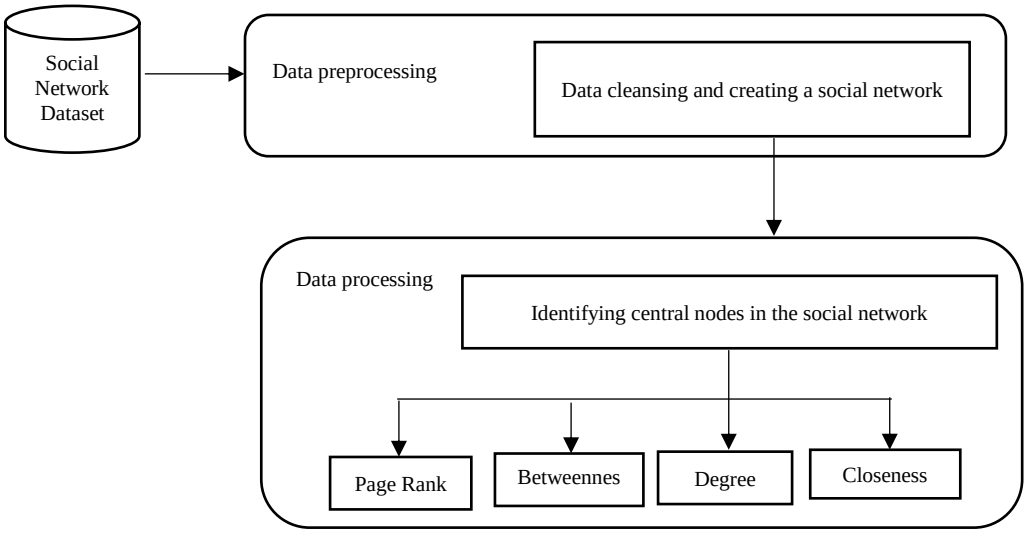


Fig. 3. Conceptual model of the proposed research approach based on data mining processes

The proposed method is suggested following the data mining process over several steps. These steps are presented in Figure 3 in the form of a conceptual model.

Following the conceptual model form of the proposed approach presented in Fig 3, a hybrid measure is applied to identify the central nodes. This criterion is calculated by summing the values of degree, betweenness, rank and closeness to each node using the relation (5).

$$I_n = w_p \times P_n + w_{bc} \times BC_n + w_d \times D_n + w_c \times C_n \quad (5)$$

In relation (5) the values are:

- I: Proposed hybrid criterion for determining the degree of importance
- n: Node ID
- $w_p$ : Weight intended for Page Rank measure
- $P_n$ : PageRank value of node n
- $w_{bc}$ : Weight intended for betweenness measure
- $BC_n$ : Betweenness value of node n
- $w_d$ : weight intended for degree measure
- $D_n$ : Rating value of the nth node
- $w_c$ : Weight intended for the closeness measure
- $C_n$ : closeness value of the nth node



IEEE Office - SEET Complex, Federal University of Technology, Owerri, Nigeria, https://aeuso.org
EISSN: 2305-0543, PISSN: 2411-6173

Weight values are calculated based on the relationship (6). According to this equation, the sum of the weight values is equal to one. Weight values intended are weighted for page rank criteria, betweenness, degree, and closeness respectively. In other words, the highest weight is assigned to the PageRank and the lowest weight to the closeness measure.

$$W = w_p + w_{bc} + w_d + w_c = 1 \quad (6)$$

$$w_p \geq w_{bc} \geq w_d \geq w_c$$

### A. *Implementation and results*

MATLAB R2018b software was used to simulate and evaluate the results of this research. All steps were performed with coding. The data employed in the simulation model is social network dataset. The only standard dataset available from social networks is the Facebook Social Network dataset, which is used as a standard dataset applicable to social network evaluations and can be downloaded from the Stanford University site [21]. This dataset contains multiple subgraph data from the main Facebook graph with 3980 nodes. These subgraphs have 107, 348, 414, 698, 1684 nodes respectively. To avoid the complexity of the data volume and to better represent the outcomes in the simulation model of this research, a graph with 348 nodes is selected. Figure 4, this graph shows the Facebook social network.

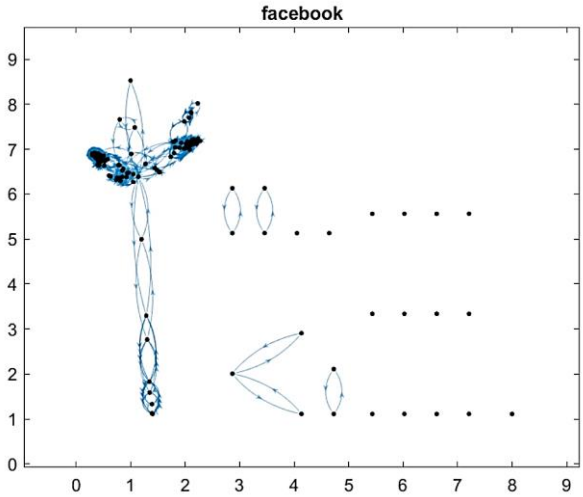


Fig. 4. Graph created with Facebook social network dataset

Considering the proposed criterion of this study, which combines several criteria for identifying central nodes in the social network graph, the central graph nodes are identified by the proposed method and compared with each of these criteria. Results are presented in the form of diagrams and figures. Figure (5) shows the results of identifying central nodes using the proposed hybrid criterion in the Facebook graph. The values of this criterion are calculated according to the method presented in Section 4 (Suggested Method) and after sorting thirty nodes with the most centrality value are selected and displayed.

The results, thirty nodes with the highest page rank criterion, degree, betweenness and closeness shown in Figure 6-7.

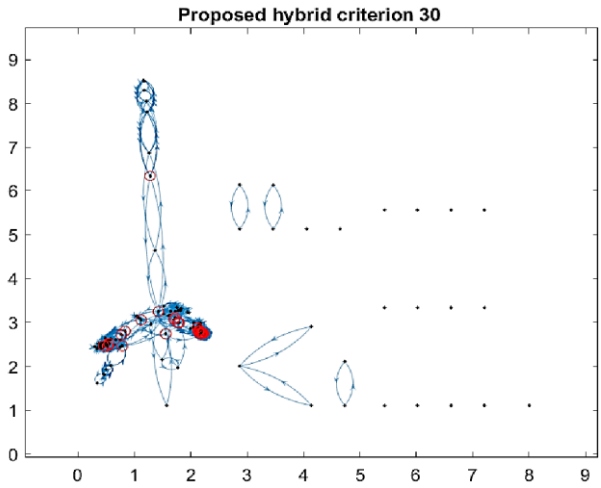


Fig. 5. Identification of central nodes using the measure of importance in the Facebook graph

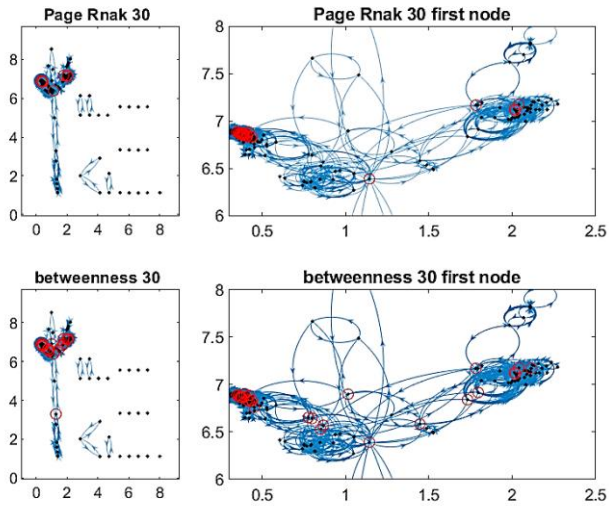


Fig. 6. Degree of identification and representation of thirty nodes with higher page rank and betweenness

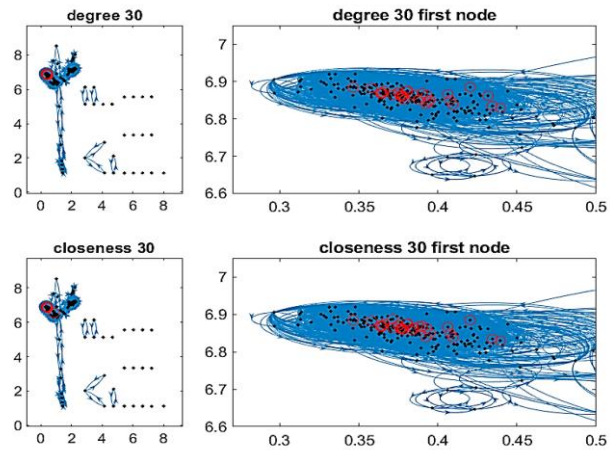


Fig. 7. Identifying and displaying thirty nodes with higher degree and closeness.

### B. *Evaluation of the results*

The results, thirty nodes with the highest criteria for the proposed hybrid centrality, PageRank, degree, betweenness



IEEE Office - SEET Complex, Federal University of Technology, Owerri, Nigeria, https://aeuso.org
EISSN: 2305-0543, PISSN: 2411-6173

and closeness are shown in Figures 5 to 7. As can be seen in the figure, the criterion of importance can identify the central nodes in all parts of the graph. These nodes can be used to achieve business goals. For example, to advertise a new product and express the characteristics of a product, these nodes can provide the fastest information. Also, to evaluate the results, the values of thirty nodes were compared with the highest criterion of page rank centrality, degree, betweenness and closeness and proposed hybrid measure (Fig. (8)).

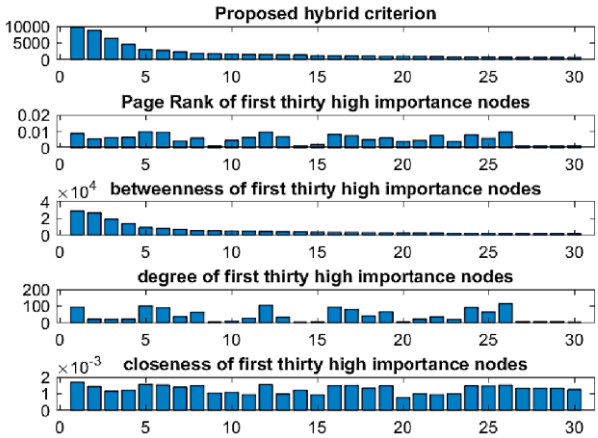


Fig. 7. compares 4 centrality measure with each other

As can be seen in the results, all four criteria of page rank, betweenness, degree and closeness can identify central nodes in the social graph. According to the definitions provided by these criteria, the method of identifying central nodes in each criterion and accordingly in the results is different. Indeed, how to select central nodes by any criteria as follows:

- Degree: Based on the number of links created on a node (input and output link).
- Betweenness: Based on the number of times a node acts as a bridge along the shortest path between two nodes.
- Closeness: How long it will take to spread the information from node to node.
- Page Rank: Works based on the number and quality of links on a page.

Therefore, the suggested criterion of this research, considering the capability of all four measures, identifies the central nodes of the social graph and provides more comprehensive results. These central nodes in social networks can be used as candidate nodes for business intelligence purposes. In other words, these criteria can be effective in e-commerce strategies. Consequently, the recommended solution in applying business intelligence for the central nodes identified by the proposed hybrid measure is:

- Since nodes identified by the measure of degree, betweenness, and page rank are based on the number and quality of an input or output links of a node, In a statement of the advancement and introduction of products includes business calling plans, identified nodes of greater importance can be used (with more impact on rank, betweenness and page rank metrics).
- In applications where the duration of information dissemination from node to other nodes is considered (with greater closeness effect).

## V. CONCLUSION

This research seeks to improve active businesses on social networks through solutions offered in business intelligence. The main purpose of this research is to present a model based on business intelligence solutions adopting social network data analysis and in this regard, data mining and text mining techniques on social network graph are used. This paper employs a hybrid measure including centrality, page rank, betweenness, degree and closeness criteria to identify central nodes in social networks. Applying this criterion, we can identify central nodes in the social network and use them in business intelligence strategies such as advertising, product introduction, and business recall.

## REFERENCES

[1] Vassakis K., Petrakis E., Kopanakis I., Makridis J., Mastorakis G. (2019) Location-Based Social Network Data for Tourism Destinations. In: Sigala M., Rahimi R., Thelwall M. (eds) Big Data and Innovation in Tourism, Travel, and Hospitality. Springer, Singapore

[2] Richardson, Joao Tapadinhas,( 2017 ), Magic Quadrant for Business Intelligence and Analytics Platforms, Published: 16 February, Gartner Reprint,ID: G00301340

[3] Shah J. Miah, Shahadat KhanShaheb Ali,(2017), Aalysis of interaction between Businesis in Telligence and SMES: Learn from each other, JISTEM - Journal of Information Systems and Technology Management, Vol. 14, No. 2, May/Aug., 2017 pp. 151-168

[4] Lars Degerstedt, Södertörn University & Pelle Snickars, Umeå University , (2017), More Media, More People—On Social & Multimodal Media Intelligence, Citation for the original published paper (version of record): Degerstedt, L., Pelle, S, HUMAN IT 13(3): 54–84

[5] Harun Bayer , Mustafa Aksogan , Enes Celik, Adil Kondiloglu, (2017), Big data mining and business intelligence trends, Journal of Asian Business Strategy Volume 7, Issue 1(2017): 23-33

[6] Darren Edge, Jonathan Larson, Christopher White, (2018), Bringing AI to BI: Enabling Visual Analytics of Unstructured Data in a Modern Business Intelligence Platform, ACM 978-1-4503-5621-3/18/04…$15.00, https://doi.org/10.1145/3170427.3174367

[7] Tanaporn Panrungsri, Esther Sangiamkul, (2017), Business Intelligence Model for Disaster Management: A Case Study in Phuket, Thailand, WiPe Paper – Operational applications and perspectives, Proceedings of the 14th ISCRAM Conference – Albi, France, pages:727-738

[8] Richa Gupta ,"Journey from Data Mining to Web Mining to Big Data", International Journal of Computer Trends and Technology (IJCTT) ,page 18-20, volume 10 number 1 , Apr 2014

[9] Castellón González, Pamela, and Juan D. Velásquez. "Characterization and detection of taxpayers with false invoices using data mining techniques."Expert Systems with Applications 40.5 (2013): 1427-1436

[10] Ismail.M, Mohammed Mansur Ibrahim, Zayyan Mahmoud Sanusi, Muesser Nat, (2015), Data Mining in Electronic Commerce:Benefits and Challenges, Int. J. Communications, Network and System Sciences, Volume 8, pages: 501-509.

[11] Instagram (the Team), Kevin Systrom, Available at : https://www.instagram.com/about/us/

[12] Karrupusamy, P., Chen, J., & Shi, Y. (Eds.). (2020). Sustainable Communication Networks and Application. Lecture Notes on Data Engineering and Communications Technologies. doi:10.1007/978-3-030-34515-0

[13] Anne-Marie Kermarrec, Erwan le Merrer, Bruno Sericola, Gilles Tredan, "Second order centrality: Distributed assessment of nodes criticity in complex networks", Computer Communications, Vol. 34, pp-619-628, (2011).

[14] Theoretic Impact of Nodes during Percolation in Networks", Vol. 8, (2013) .

[15] Mohammad Reza Faghani, "A Study of XSS Worm propagation and Detection Mechanisms in online Social Networks", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, Vol. 8, pp-1815-1826, (2013) .

[16] Zhonghua Liu, Cheng Jiang, Juyun Wang, Hua Yu, "The node importance in actual complex networks based on a multi-attribute ranking method", Knowledge-Based Systems, Vol. 84, pp-56-66, (2015)

[17] [51] Linton C. Freeman, Centrality in Social Networks Conceptual Clarification, Social Networks, 1 (1978/79) 215-239 Available at: www.leonidzhukov.net/hse/2018/sna/.../freeman79-centrality.pdf

[18] Lv, L., Zhang, K., Zhang, T., Bardou, D., Zhang, J., & Cai, Y. (2019). PageRank centrality for temporal networks. Physics Letters A. doi:10.1016/j.physleta.2019.01.041

[19] López-Robles, J. R., Otegi-Olaso, J. R., Porto Gómez, I., & Cobo, M. J. (2019). 30 years of intelligence models in management and business: A bibliometric review. International Journal of Information Management, 48, 22–38. doi:10.1016/j.ijinfomgt.2019.01.0132019

[20] Noor UL-Ain, Vaia Giovanni, William DeLone,( 2019) , Business intelligence system adoption, utilization and success -A systematic literature review, Proceedings of the 52nd Hawaii International Conference on System Sciences, ISBN: 978-0-9981331-2-6

[21] https://snap.stanford.edu/data/egonets-Facebook.html