
Social Network Analysis for User Interaction Analysis on Social Media Regarding E-Commerce Business

Nugroho Agung Prabowo ^{1,*}, Bambang Pujiarto ², Firmantya Safri Wijaya ³, Lutfiana Gita ⁴, Denny Alfandy ⁵

Department of Information Technology, Faculty of Engineering, Universitas Muhammadiyah Magelang, Magelang, Indonesia

¹ naprabowo@ummgl.ac.id *; ² bpujiarto@ummgl.ac.id; ³ Alifwijaya440@gmail.com; ⁴ gitaasmara250@gmail.com;

⁵ denialfan163@gmail.com

* Corresponding author

(Received July 3, 2021 Revised August 11, 2021 Accepted August 29, 2021, Available online September 1, 2021)

Abstract

Promotion is necessary for an e-commerce business to introduce its products. Social media is one of the media that can be employed. Social media platforms give a wealth of information, one of which is User Generated Content (UGC). UGC is a user's social media history that is viewable by other users. Social media analysis is required to determine the pattern of interaction between a business and its customers based on user-generated content (UGC) that is extensively shared on social media. This data can be utilized to assist businesses with product promotion on social media. Social Network Analysis is the technique used to analyze the interaction patterns of UGC in social media (SNA). Social network modeling can assist e-commerce enterprises in comprehending the patterns of interaction that exist on social media. The findings of this study indicate that the superior social network is the Lazada interaction social network. Additionally, the research identifies the key companies in each e-commerce sector.

Keywords: User Generated Content; Social Network Analysis; Key Cast; Network Properties; E-Commerce

1. Introduction

High internet penetration in Indonesia has led to the development of online e-commerce businesses. Online e-commerce business is a means of commercial transactions between and between organizations and individuals digitally with the use of the internet, web and online business applications. The e-commerce business has changed many things in the business world, where the focus is not only on goods but on services, information and intelligence focus on customers.

Social media is important in e-commerce business in cyberspace because social media provides insight in e-commerce business marketing. Social media has become a tool for creating, communicating and delivering the value that organizations provide to consumers. Consumers globally use social media to search for product information and share product-related comments on social media.

All information provided by users on social media is called User Generated Content (UGC). UGC is data left by users that can be accessed by the public online. The social media that holds a lot of UGC is Twitter. Twitter is a social media that has many benefits so that it has very strong information power, including; key communication for politics, dissemination of information related to natural disasters and for business for marketing, and customer service.

Social media analysis is needed to see the pattern of interaction between the company and its customers from UGC which is widely spread on social media. The data on social media is unstructured and has a large volume, so sophisticated tools and techniques are needed to extract information from social media. The method that can be used is Social Network Analysis (SNA) which is part of the Social Computing technique to extract information on unstructured data and has a large volume. big. SNA is a study that studies human relationships by utilizing graph theory.

SNA helps to understand social relationships that represent users with nodes (nodes) and relationships between users are represented by lines (edges) on the Online Social Network (OSN). SNA can be used to study network patterns of organizations, ideas, and people who are connected in various ways in an environment. So SNA can be applied by

businesses to gain insight and knowledge from the market and community through the media. Businesses can take advantage of the social networking properties of SNA to make comparisons with other businesses regarding relationships on social networks.

SNA is an analytical approach that can identify social structures and explain the position of key stakeholders or what can be called key actors. Key players can be the most important actors in disseminating information, especially for businesses to promote their products on social media because they have the greatest effect in disseminating information on social networks. So that this can be used as insight for companies in helping product marketing, especially for e-commerce businesses that use social media to promote products. So the purpose of this study is to determine the properties of social networks that are formed in each e-commerce and to find out the key players in social networks for information dissemination.

2. Literature Review

Kaplan and Haenlein [25] define social media as a collection of Internet-based apps built on Web 2.0's conceptual and technological basis that facilitate the creation and exchange of user-generated content. Social media, alternatively referred to as consumer-generated media, is a term that refers to user-generated material. Social media can take on a variety of forms, including virtual communities, weblogs, microblogs, wikis, image or video sharing, social networking sites, social bookmarking, and other forms of social media. Recent years have seen a rise in social media studies. While some researchers examine the impact of social media on public relations, others examine social media in the context of a travel-related setting and demonstrate that social media reviews are very credible. While user-generated material on social media platforms might be fabricated by someone with a vested interest, many people believe the content is trustworthy since it is based on real, independent experiences. As a result, social media has functioned as a new type of word of mouth for products/services or providers and has been shown to be critical for customer decision-making in the e-commerce context.

In November 2005, Yahoo! used the phrase "social commerce" to refer to a collection of collaborative online purchasing features such as shared wish lists, user ratings, and other user-generated content sharing of online product information and suggestions. According to Stephen and Toubia [47], social commerce is a type of internet-based social media that enables individuals to actively participate in marketing and selling items and services through online marketplaces and communities. Marsden [33] defines social commerce as a subset of electronic commerce that entails the use of social media, or online platforms that facilitate social interaction and user contributions, to facilitate online purchasing and selling of items and services. We can see from this description that social commerce has two critical components – social interaction and user participation – that originate from the characteristics of social media.

The term "social interaction" refers to a bond [53] or a relationship of social contact. It originates in social capital theory and manifests itself as a structural dimension of social capital. The term "social interaction" refers to the bond formed between two actors through reciprocal activity. It demonstrates the relationships between individuals or groups, which is aided by social media. Within the framework of social media, all of the activities in which actors engage are based on their social media interactions. They have an effect on knowledge transmission because social media engagement relationships operate as a conduit for and facilitate information flow. This can also be said of social trading. Thus, social media engagement ties can be critical in motivating users to interact more and, as a result, make transactions more social.

The term "user contribution" refers to user engagement on social media platforms. Commitment is defined as a psychological concept that encapsulates the desire and drive to continue participating. Given the enormous number of individuals who continue to utilize social media, we may anticipate that the greater the level of user commitment, the greater the level of user contribution. Commitment has been conceptualized as a relational dimension of social capital in social capital theory. It has been demonstrated to have an effect on attitudes and future purchasing intentions. Commitment to a website has been identified as a critical component for influencing online purchase attitudes, and commitment to a virtual community has also been identified as critical for brand loyalty and customer behavior.

Numerous research on user adoption and acceptance have been conducted, with theories and models such as the theory of reasoned action, the theory of planned behavior, the technology acceptance model, and the integrated theory of technology acceptance and usage serving as the foundation. While the majority of research examines usage behavior in isolation, some studies have examined the relationship between system use in different contexts, implying that user behavior in one context can influence user behavior in another related context through the transfer of attitudes toward the system. Numerous studies have also discovered that early use of information technology systems may have an effect on others' future post-use aspirations. With the increasing integration of systems and platforms, it is vital to pay attention to the interaction between user behavior across contexts.

Stewart [48] makes use of entities to explain how confidence is transferred across distinct targets. Song et al. [46] used entities to shed light on how usage behaviors are transferred between technologies and settings. The term "entity" refers to the degree to which a collection of distinct entities is deemed to form a coherent whole. This is a

critical factor for comparing groups, since the perception of entities has a significant impact on the creation of one's impression of information from individuals or groups, which in turn impacts the information processing method. Because groups have traditionally been thought of as collections of entities, social media users on social networks can also be thought of as groupings. Lickel et al. [30] assert that social interactions within a group can result in perceptions of the entity. That is, social networking groups with a high level of social contact will also have powerful entities that influence the formulation and processing of group members' impressions and information, which has an effect on their perceptions and attitudes.

Big data is data that is larger beyond the capabilities of ordinary database systems to process. Big data is extremely vast, rapidly changing, and does not fit into traditional database structures. By implementing advanced analytical technologies and processes, companies can leverage big data to produce novel insights, products, and services that boost the organization's economic value. Big Data is not a discrete technology, approach, or effort in and of itself. Big Data is a broad term that encompasses a broad range of topics in business and technology.

Big Data refers to technologies and projects that utilize data that is too diverse, constantly changing, or massive to be efficiently handled by conventional technology, expertise, or infrastructure. In other words, Big Data is data that has an extraordinary volume, velocity, or variety that cannot be managed using standard procedures. Big Data refers to data creation, storage, information mining, and analysis activities that are characterized by their size, velocity, and variety. According to (Eaton, Dirk, Tom, George, and Paul), "Big Data" refers to data that cannot be processed or evaluated using conventional technologies. According to (Dumbill, 2012), Big Data is data that exceeds the capabilities of standard database systems to process. The data is either too vast and/or too quickly, or it does not fit inside the structure of the present database system. To extract value from data, it must choose a different method of processing it.

User Generated Content (UGC) is data or content that is generally available by other users and is developed by individuals who are not professionals in this field (Moens et al., 2014). Online UGC data is critical as a source of information for enterprises seeking to discover value, necessitating the use of cutting-edge procedures and technology to facilitate data retrieval, storage, administration, and analysis. Social computing is a new and rapidly evolving computerized paradigm at the intersection of computer science and social science that entails a multidisciplinary approach to analyzing and modeling social behavior across multiple media and platforms in order to create intelligent and interactive methods and applications.

Social network analysis (SNA) is a method for modeling users as points (nodes) and their interactions as lines (edges). This approach is necessary because it opens up new avenues for understanding individuals or groups through their patterns of social interaction. SNA can be used to investigate the network patterns of organizations, ideas, and individuals who are connected in a variety of ways in an environment. SNA has various network qualities that enable it to map relationships, which is extremely beneficial for enhancing management knowledge production in businesses. SNA considers the following network properties: nodes, edges, average degree, dimension, and average path length.

Table. 1. Definition of Network Properties

Network Properties	Explanation
Nodes	Nodes represent the positions held by actors in the network
Edges	Edges reflect relationships between actors or entities that occur in the network
Average Degree	Average degree can be determined by the number of relationships on one node divided by the number of relationships that occur in one social network
Diameter	Diameter is the furthest distance between two adjacent nodes
Average Path Length	Average path length is the average geodetic distance, or the average path traversed by each node to other nodes

Centrality measurement is used to determine the actor who plays the most important role in a social network, this shows the degree of a person's center. There are four measurements of centrality, namely: degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality.

Table. 2. Definition of Centrality

Centrality	Explanation
Degree Centrality	Degree centrality is determined by the number of edges associated with the nodes
Betweenness Centrality	Betweenness centrality identifies nodes that will be intermediaries for information
Closeness Centrality	Closeness centrality is the average distance from a given node to all other nodes in the social network
Eigenvector Centrality	Eigenvector centrality shows the most important nodes in the network based on the connections that nodes have and the nodes associated with these nodes

Based on several theories and studies, the framework of thought in this research will be described as follows:

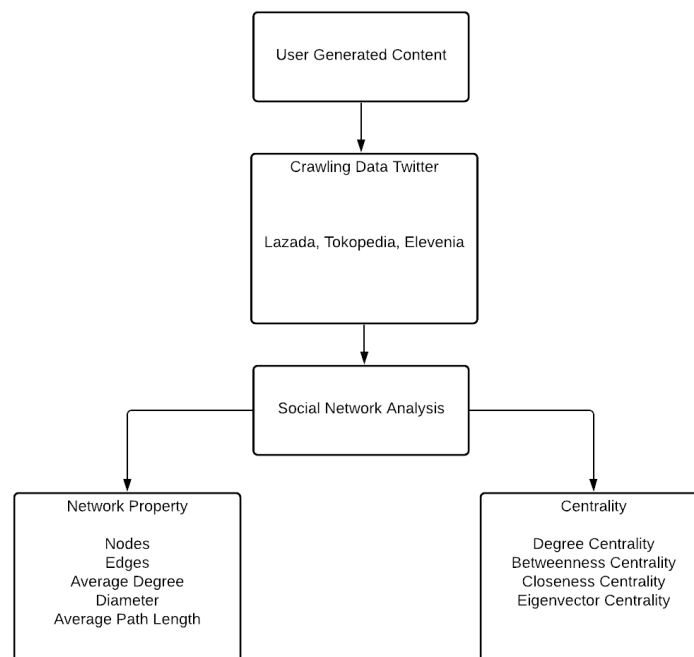


Fig. 1. Thinking Framework

Sources of data obtained in the study included in the category of User Generated Content (UGC). UGC is data that is generally visible to other users, where the content contains a number of creativity and is made by people who are not professionals in this regard. UGC data can be obtained using data crawling. This research takes data on Twitter social media because it is open-source so that data can be taken openly. The data taken are users (actors) who interact on Twitter social media related to e-commerce Lazada, Tokopedia and Elevenia.

Social Network Analysis (SNA) is one of the Social Computing methods to extract information on unstructured and large volume data. Social network is a modeling of users which is symbolized by points (nodes) and interactions between these users are represented by lines (edges). SNA has several network properties to map relationships which are very helpful for improving the creation of management knowledge in organizations. Network properties can be used as a comparison between social networks in each user interaction regarding the e-commerce business. Centrality measurement is used to determine the actor who plays the most important role in a social network, this shows the degree of a person's center.

3. Research Methods

The method in this study is a qualitative research method. Qualitative research is research that understands social phenomena or phenomena by providing detailed exposure to them. Based on the research objectives, this research is

included in descriptive research. Descriptive research is designed to collect data to describe the characteristics of a person, activity or situation. The research is cross sectional, namely data collection is carried out in one period, then the data is processed, analyzed and then drawn conclusions.

Sources of data obtained by researchers are secondary data sources. The data taken is a tweet on Twitter social media in which it is about business e-commerce Lazada, Tokopedia and Elevenia taken from 15 July 2020 to 30 July 2020. The data analysis technique used is Social Network Analysis (SNA). The stages in the research are as follows:

1) Data Collection

Online conversation data retrieval is done by crawling Twitter social media data. This stage is carried out by crawling using an R language-based application, namely R Studio which is connected to the Twitter API. The retrieved tweets contain the keywords “lazada”, “tokopedia” and “elevenia”.

2) Preprocessing Data

The dirty tweets, mentions, replies and retweets that have been collected are then preprocessed to remove irrelevant tweets to make the analysis process easier. Next, determine the actors who interact on each tweet to serve as nodes in the social network.

3) Network Modeling

Data that has gone through preprocessing is then reprocessed using the Gephi application to visualize the network model using an undirected graph, which is a network that does not take into account the direction of the relationship.

4) Identify Network Properties

Each network model that has been processed with the Gephi application has several properties that will be calculated values. The network properties to be calculated are: nodes, edges, average degree, diameter, and average path length.

5) Centrality Analysis

To determine the key actors in the social network, a centrality analysis was carried out. Centrality analysis is carried out using the Gephi application to see the value of each centrality for each actor in the social network. The calculation of centrality is done, namely: degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality.

4. Result and Discussion

4.1. E-commerce Social Network Property Comparison Results

The analysis that can be done regarding the social network that is formed is the calculation of the social network properties. Calculation of user interaction social network properties regarding Lazada, Tokopedia and Elevenia e-commerce on Twitter social media for a period of one month using the Gephi application. The results of the calculation of social networks that can be used as comparisons are as follows:

Table 3. Comparison of Social Network Properties for each e-commerce

Network Properties	Lazada	Tokopedia	Elevenia
<i>Nodes</i>	47.398	9.897	3.755
<i>Edges</i>	50.385	10.301	5.358
<i>Average Degree</i>	21,3	2,08	2,04
<i>Diameter</i>	24	14	10
<i>Average Path Length</i>	7,25	4,08	3,88

The table 3 above shows a comparison of social network properties formed on user interactions on Twitter social media regarding the Lazada, Tokopedia and Elevenia e-commerce businesses in the research time span. The first compared property is node. The higher the node, the more actors involved in the social network. This in business can indicate that many actors are aware of the existence of the e-commerce business. The highest node network property value is Lazada, which is 47,398 nodes, which shows that there are 47,398 actors talking on Twitter social media, followed by Tokopedia with 13,129 nodes and Elevenia with 3,755 nodes. The second property comparison is edge. Edge shows the interactions that occur between actors. The higher the edge value will indicate that many conversations are about e-commerce business on Twitter social media. Lazada has the highest edge value of 50,385 edges. Meanwhile, Tokopedia and Elevenia have 10,590 edges and 5,358 edges.

The third comparison of network properties is the average degree. Average degree shows the average number of relationships that the actor has in the social network. The higher the average degree value, the better because each actor in the network has many relationships so that it will expand the spread of information. The e-commerce business that has the largest average degree value is Lazada with a value of 2.13, followed by Tokopedia and Elevenia with a value of 2.08 and 2.04. The next comparative analysis is diameter. Diameter is the maximum distance between nodes. The smaller the diameter, the faster the information will circulate in the social network. The smallest diameter is in user conversations about Elevenia's e-commerce business with a value of 10. Meanwhile, Tokopedia and Lazada have a value of 14 and 24. Another measure that can measure the speed of information circulating between actors in a social network is the average path length. Average path length is the average distance between nodes. The smallest average path length value is in user interaction regarding Elevenia's e-commerce business with a value of 3.88, followed by Tokopedia and Lazada with a value of 4.08 and 7.25. The following are the results of the social network visualization regarding the e-commerce business of Lazada, Tokopedia and Elevenia:

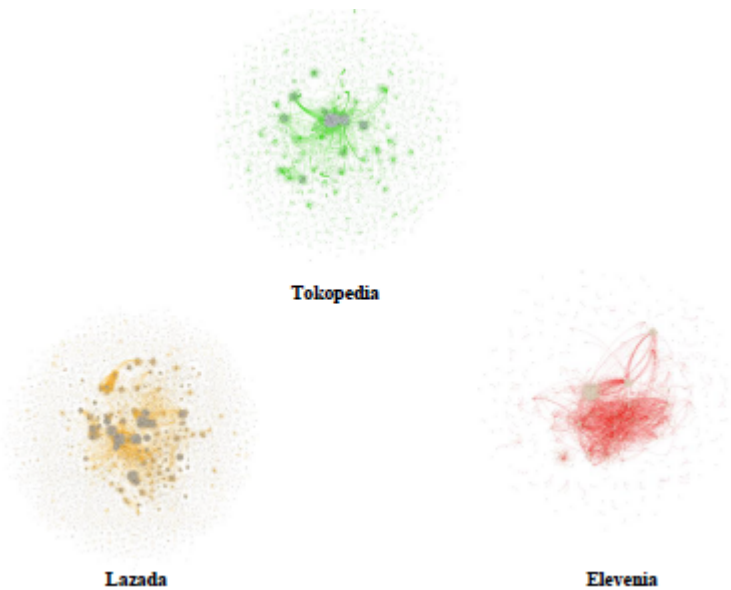


Fig. 2. Visualization of e-commerce Business Social Network.

The three images above are the result of visualization of the social network of e-commerce business Lazada, Tokopedia and Elevenia within a period of one month which shows the relationship between nodes and edges that are incorporated in the social network of each e-commerce. Visualization is made using the Gephi application using the Open Ord layout. The type of graph used is an undirected graph, which is a graph that does not consider the direction of the relationship between nodes.

4.2. Determination of Key Players in e-commerce

In determining the key players in each e-commerce on social media, Twitter can use centrality measurements. Centrality measurement is used to determine the actor who plays the most important role in a social network, this shows the degree of a person's center. In determining the key players, four values of centrality are used, namely: degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality.

Based on the results of the four centralities, it was found that the key player for Lazada's e-commerce social network is user "C". The value of degree centrality owned by "C" is 5,988, this shows that "C" account has a total relationship of 5,988 with other accounts. The higher the degree of centrality value, it can be said that the account has many relationships so that it can affect other accounts. User "C" has the highest betweenness value of 0.26. The higher the betweenness value indicates that the account acts as an intermediary between other actors. User "C" is also the closest to other actors so that if the account disseminates information it will be faster because it has the highest closeness value of 0.22. The user who has the highest eigenvector value is user "C" with a value of 1.0. The user who has the highest eigenvector value indicates that the user has relationships with many important actors in the social network.

The key player for the Tokopedia e-commerce social network is the "A" user because it has the highest score on the four centralities. "A" has a degree value of 410, a betweenness value of 0.05, a closeness value of 0.30, and an eigenvector value of 0.06. Meanwhile, the key player for the Elevenia e-commerce social network is user "B" because it has the highest score on the four centralities. User "B" has a degree value of 111, a betweenness value of 0.03, a closeness value of 0.31, and an eigenvector value of 0.04.

5. Conclusion

Based on the calculation results of social network properties, namely node, edge, average degree, diameter, and average path length, the social network property values for each e-commerce are obtained. After comparison, the e-commerce that has the highest value on social network properties is Lazada with a total of three out of five social network properties. Lazada has a higher social network property value regarding nodes which indicates that many users talk about Lazada on social media, the edge which proves a lot of interaction between users about Lazada on social media, and the average degree which indicates that each actor in the network has many relationships so that will expand the dissemination of information about e-commerce Lazada.

Through the calculation of centrality, namely degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality, the key players of each e-commerce are obtained. These key players can help spread information about each e-commerce on Twitter social media faster and more widely. Based on the results of the four centralities, it was found that the key player for Lazada's e-commerce social network is user "C", for e-commerce Tokopedia is the "B" user, and for Elevenia e-commerce user "A".

6. Suggestion

For e-commerce businesses that have not excelled in node and edge social network properties, it can be improved by being more active in posting interesting tweets and adding accounts on social media for certain categories. To increase the value of the average degree of social network property, e-commerce businesses can hold competitions or campaigns related to certain topics that invite friends or other people to join. E-commerce businesses can also take advantage of key players in disseminating information so that it spreads faster and wider on Twitter social media. E-commerce businesses can give prizes in the form of promos or products to key players so that they can work well together in disseminating information on Twitter social media. For further research, you can conduct social network research on e-commerce businesses that have not been researched by the author and can use sentiment analysis to analyze opinions that occur on social media regarding e-commerce business.

References

- [1] Alalwan, Ali Abdallah., Rana, Nripendra P., Dwivedi, Yogesh K., and Algharabat, Raed. (2017). Social Media in Marketing: A Review and Analysis of The Existing Literature. *Journal of Telematics and Informatics*.
- [2] Alamsyah, Andry. 2013. The Role of Social Network Analysis for Knowledge Management. *Indonesian Journal of Management*, 12(4), 309-314.
- [3] Alhadj, Reda. and Rokne, Jon. 2014. *Encyclopedia of Social Network Analysis and Mining*. New York: Springer Science+Business Media Publisher.
- [4] Gunther, Wendy Arianne., Mehrizi, Mohammad H. Rezazade., Huysman, Marleen., and Feldberg, Frans. 2017. Big Data Debating: A Literature Review on Realizing Value from Big Data. *Journal of Strategic Information Systems*.
- [5] Hanneman, Robert A. and Riddle, Mark. 2005. *Introduction to Social Network Methods*. Riverside, United States: University of California.
- [6] Jackson, Matthew O. 2008. *Social and Economic Networks*. Princeton, NJ 08540, United States: Princeton University Press
- [7] Laudon, Kenneth C. and Traver, Carol Guercio. 2014. *E-commerce. Business. Technology. Society*. (10th ed.). One Lake Street, Upper Saddle River, New Jersey: Pearson Education, Inc.
- [8] Mbaru, Emmanuel K. and Barnes, Michele L. 2017. Key players in conservation diffusion: Using Social Network Analysis to Identify Critical Injection Points. *Journal of Biological Conservation*, 222-232.
- [9] Mincer, Marcin. and Niewiadomska-Szynkiewicz, Ewa. 2012. Application of Social Network Analysis to the Investigation of Interpersonal Connections. *Journal of Telecommunications and Information Technology*, 2, 83-91.
- [10] Moens, Marie-Francine., Li, Juanzi., and Chua, Tat-Seng. 2014. *Mining User Generated Content*. 6000 Broken Sound Parkway NW, Suite 300, Boca Raton: CRC Press.
- [11] O'Reilly. 2012. *Big Data Now: 2012 Edition*. 1005 Gravenstein Highway North, Sebastopol: O'Reilly Media, Inc.
- [12] O'Reilly. 2012. *The Twitter Book*. 1005 Gravenstein Highway North, Sebastopol: O'Reilly Media, Inc.

- [13]Oktora, Rio. and Alamsyah, Andry. 2014. Interaction Patterns and Actors Most Playing a Role in the 2013 JGTC Event Through Twitter Social Media (Study Using Social Network Analysis Method). *Indonesian Journal of Management*, 14(3), 201-210.
- [14]Olmedilla, M., Martínez-Torres, M.R., and Toral, S.L. 2016. Harvesting Big Data in Social Science: A Methodological Approach for Collecting Online User-Generated Content. *Journal of Computer Standards and Interfaces*, 46, 79–87.
- [15]Rahayu, Rita., and Day, John. 2015. Determinant Factors of E-commerce Adoption by SMEs in Developing Country: Evidence from Indonesia. *Journal Procedia - Social and Behavioral Sciences*, 195, 142 – 150.
- [16]Rios, Sebastián A., Aguilera, Felipe., Nuñez-Gonzalez, J. David., and Graña, Manuel. 2017. Semantically Enhanced Network Analysis for Influencer Identification in Online Social Networks. *Journal of Neurocomputing*, 1-11.
- [17]Sapountzi, Androniki. and Psannis, Kostas E. 2016. Social Networking Data Analysis Tools & Challenges. *Journal of Future Generation Computer Systems*.
- [18]Now, Uma. and Bougie, Roger. 2013. *Research Methods for Business: A Skill-Building Approach* (6th ed.). The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom: John Wiley & Sons Ltd.
- [19]Sujarweni, V. Wiratna. 2015. *Business & Economics research methodology* (1st ed.). Bantul, Yogyakarta: PT Pustaka Baru
- [20]Tavakolifard, Mozghan. and Almeroth, Kevin C. 2012. Social Computing: An Intersection of Recommender Systems, Trust/Reputation Systems, and Social Networks. *IEEE Networks*, 53-58.
- [21]Yadav, Mayank. and Rahman, Zillur. 2017. Measuring Consumer Perception of Social Media Marketing Activities in E-Commerce Industry: Scale Development & Validation. *Journal of Telematics and Informatics*.