JOURNAL OF BIG DATA SCIENCE

P-ISSN 3025-6615

Volume 1 Number 1
September 2023

Scientometrics Study and Authorship Network Analysis in Universitas Bina Darma Lecturer Publications

Ferdi Aditya[1]*, Deni Erlansyah[2], Dendi Triadi[3], Rezki Syaputra[4]

[1,2,3,4]Data Science Interdisciplinary Research Center, Universitas Bina Darma
*f.aditya@binadarma.ac.id

Ahmad Yani Street No.12, Palembang, South Sumatra 30624, Indonesia

Received September 25th, 2023; **Revised** September 29th, 2023; **Accepted** September 30th, 2023

Abstract

Scientometrics is the study of measurement and analysis of science, innovation and technology through scientific publications. One form of measurement that can be taken is the network of authors measurement. This study uses author network analysis as a measurement tool performed in scientific studies. The purpose of this study was to observe the Authorsip network formed among professors at Universitas Bina Darma, in order to determine which professors and departments are the most productive in producing yearbook articles or magazine. The method used in this study is the centrality of graphic degrees. Software used to view Gephi 0.9.2. The data used in this study are published data for the year 2015-2020. Based on the results of this study, it can be concluded that the agent with the highest central value is the EU with a value of 28, where the EU is the agent. with the largest number of publications. Meanwhile, the actor who has an influence or relationship and frequently collaborates on publications with the highest score on Betweenness Centrality is AM with a score of 61500.94.

Keywords: *Centrality Detection, Social Network Analysis, SNA, Data Visualization, Gephi.*

1. INTRODUCTION

One indicator of the progress of science and technology in a country is the number of published and utilized research results. Data on international scientific publications in Indonesia indexed by Scopus as of 2017 are at 12,098 publications, Indonesia is currently ranked third at the Asean level. Publication is one of the tasks that must be carried out by a lecturer, Scientometrics is the study of measurement and analysis of science, technology and innovation [1], [2]. Authors in a publication can be modeled as a network (graph) in which the main object is the author/author represented as a set of nodes, with relationships relationship between one and two authors when writing together. is a representation of the relationship (edge) [3], [4].

Social network analysis (SNA) looks at social relationships related to network theory, including nodes representing individual actors in the network and relationships representing relationships between networks. Individual [5], [6]. SNA is a method used to analyze the structure of social networks with various elements in an interconnected social environment. The network analysis approach has been widely used in various domains, such as: analyzing and detecting research network communities on Research Gate social media [7]. In addition, social network analysis has also been widely used in various social studies such as friendship networks [8], [9], hoax detection [10], and YouTube video searches with a network analysis approach with various approaches such as degree centrality [11], betweenness centrality, closeness centrality and community detection [12]–[14].

Centrality in SNA is a measure to see the position of an actor/group in a sociogram. Actor Degree Centrality is the number of direct relations owned by an actor. Betweeness Centrality is one way to measure centrality in a social network [15]. Interaction between 2 or more actors sometimes depends on other actors in the network. Actors who act as intermediaries between 2 or more actors are often

considered to have a bigger role in the flow of information because they control the interactions between these actors. Betweenness of an actor is the number of presences of an actor in the geodesic (shortest path) of each other pair of actors compared to the number of geodesic pairs of actors in the network. Individuals with the highest intermediate value are considered to have the most control over the flow of information in the network [15]. Another measure of centrality is closeness. Closeness measures the closeness between actors/nodes. The original idea of this measure is that an agent is called the center of the network if it can interact more easily and quickly with other agents. [15]. In terms of information flow, a hub close to other actors is more efficient because they can access information faster.

The centrality method in social network analysis can be used to determine the structure of the authentication network. Authors in a publication can be referred to as actors (nodes), while faculties can be referred to as networks (graphs). Actors here can be seen in relation to other actors in a network, how central the prominent actors (nodes) are in the network so that we can find out which lecturers and faculties are the most productive in producing publications. Based on this phenomenon, this article will present an authoritative network analysis that is applied to determine scientometrics and lecturer performance in the publication of scientific papers.

2. RESEARCH METHODOLOGY

The research was carried out by visualizing the author's network structure pattern from the dataset obtained through the crawler results and the data was visualized into a sociogram using Gephi software version 0.9.2. The dataset used is the authorship network data of Universitas Bina Darma Lecturers.

In this research, the authors used dataset groups from 2015 to 2020 with 1,250 authors and 631 articles. Before this study entered into further discussion, the authors carried out the evaluation stages of the research results. This evaluation is useful as a benchmark for planned activities to determine the state of an object. From this evaluation.

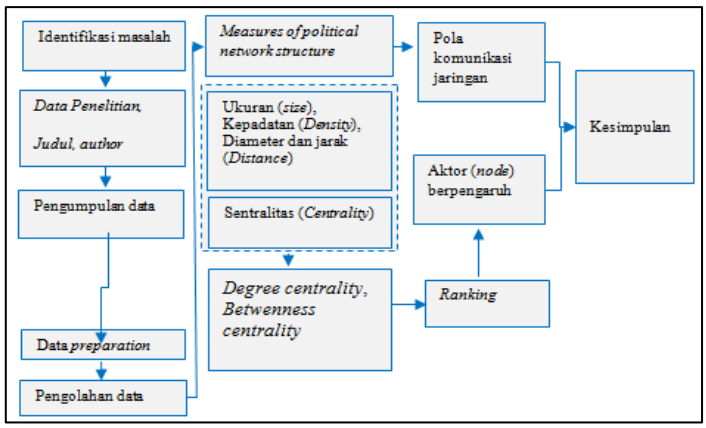

Figure 1 : Research Framework

From the research framework above, it can be explained as follows:

1) Problem identification is a process in which the author observes phenomena from social networks according to the background of the research
2) Lecturers at Binadarma University are research objects which are the material.
3) Data collection is done by manually inputting data.
4) Processing data using Centrality. Next, the visualization results will calculate the values of network properties, including the number of nodes, edges, average, average path length, and number of communities.
5) After the next data processing, the analysis will be conducted to calculate the size, mutuality, distance and finally the centrality. To identify nodes or actors that influence the number of network interactions, it is necessary to measure centrality. The calculated center values are degree centrality, neighborhood centrality, and intermediate centrality.

6) From the results of the centrality measurement, a ranking is made to rank the centrality score of the agent. From the outcome evaluation process of the applied research framework as described above, the final step is to draw conclusions.

Based on Publication data from the Universitas Bina Darma LPPM taken from 2015 to 2020, there were 632 articles from Universitas Bina Darma lecturers listed in the Universitas Bina Darma Lecturer Research Performance Report. From these articles, 1291 authors were registered from various faculties, see Figure 2.

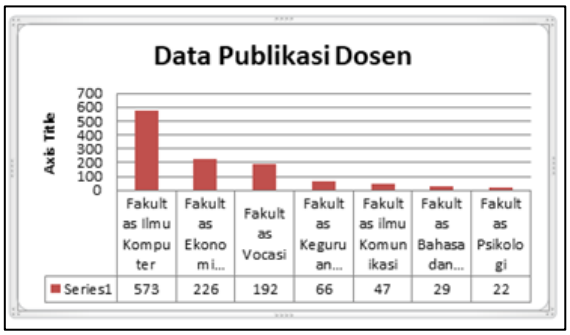

Figure 2 : Publication Data of Universitas Bina Darma

The network that was formed from Universitas Bina Darma Lecturers based on writings in Proceedings and Journals basically shows that there are several writers who have a big role in the network that occurs. See Figure 3, where the figure shows the form of the Universitas Bina Darma lecturer authorship network based on the number of publications and the relationship between lecturers through joint publications. A writer who has a big role if he has a good measure of centrality. The existence of writers with good centrality will usually form sub-groups in the existing network.

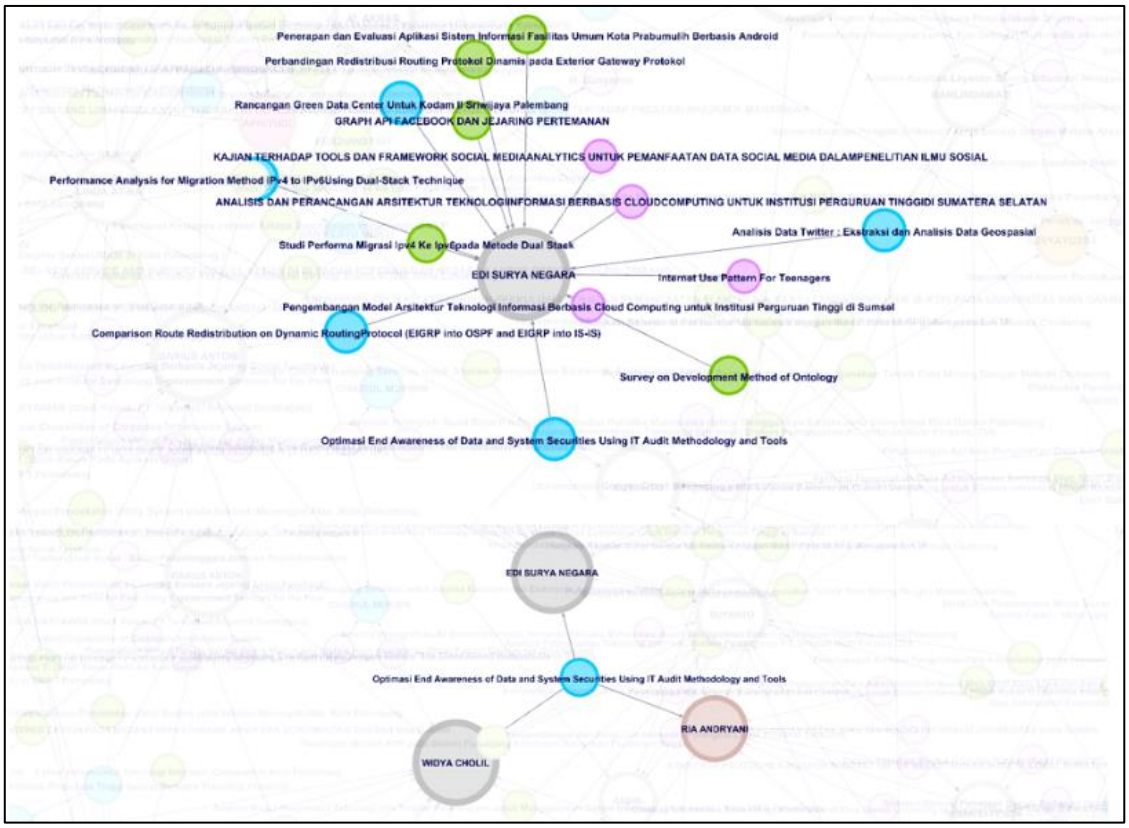

Figure 3 : Social Network Between Universitas Bina Darma Lectures in Publication

3. RESULTS AND DISCUSSION

This research has succeeded in measuring the productivity of Universitas Bina Darma in terms of publications using the scientometrics method while to determine the authorship network structure the centrality algorithm is used, where the results of this study it can be concluded that the actor who has the highest degree centrality value is UE with a value of 28, where UE is the actor with the highest number of publications. Meanwhile, the actor who has influence or relationship and often collaborates in publications with the highest score on Betweenness Centrality is AM with a score of 61500.94.

3.1 Lecturer Productivity in Producing Publication

3.1.1 Degree Centrality

This analysis aims to identify the most central or influential actors in a network. The measures used in this analysis are degree centrality, both degrees of entry and exit, closeness centrality and betweenness centrality. The following table of degree centrality can be seen below.

Table 1 : Degree Centrality

No	Author Name	Faculty	DC	BC
1	U.E	Computer Science	28	47851.60
2	A.M	Computer Science	27	61500.94
3	M.I	Computer Science	24	46789.29
4	L.A.A	Computer Science	23	50313.80
5	HAR	Communication	22	34502.47
6	Y.K	Computer Science	22	52545.97
7	A H.M	Computer Science	20	50748.43
8	E.S.N	Computer Science	18	22649.88
9	CH. D	Technique	18	27120.89
10	D.M	Economy and Business	18	21084.37

From Table 1 above it can be seen that among the authors who have a fairly high DC are in a group that has a fairly large BC value (Ari Muzakir), but if measured from the largest DC (Usman Ependi), the BC value (47851) is not balanced , with this means the author with the highest number of articles is not necessarily connected to many nodes in other large grubs. See Figure 4.

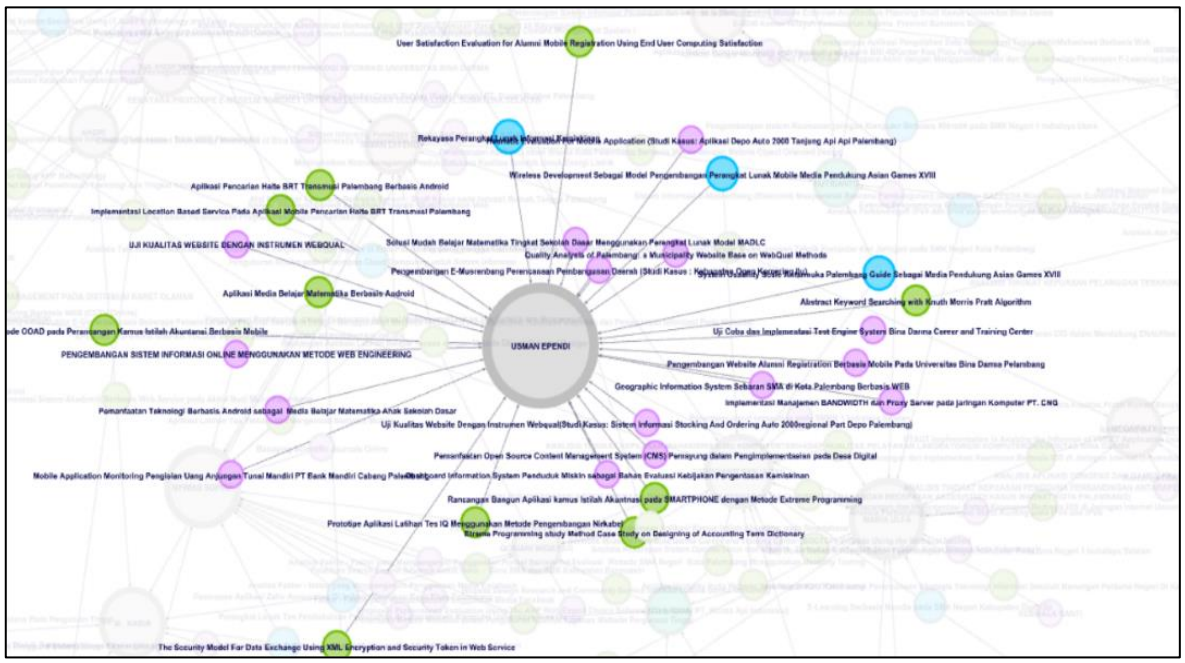

Figure 4 : Social Network Between Universitas Bina Darma Lectures with a Degree Centrality Value

3.1.2 Betweenness Centrality

Betweenness centrality is a measurement at the points where there are points that are present when we walk across from one point to another. The point that has a high betweenness centrality is considered as the most 'strong' influential point in the network. If these points with high betweenness centrality are removed from the network, it will disrupt all communication between other points because they are on the path with the maximum number of passes. The following table Betweeness Centrality can be seen in Table 2 below.

Table 2 : Betweeness Centrality

No	Author Name	Faculty	DC	BC
1	A.M	Computer Science	27	61500.94
2	Y.N.K	Computer Science	22	52545.97
3	A.H.M	Computer Science	20	50748.43
4	L.A.A	Computer Science	23	50313.80
5	U.S	Computer Science	28	47851.60
6	M.I.H	Computer Science	24	46789.29
7	L.Y.S	Economy and Business	15	44005.98
8	S.A	Economy and Business	13	39074.34
9	HAR	Communication	22	34502.47
10	KUR	Communication	11	30778.13

From Table 2 above it can be seen that among the writers who have BC with the highest score Ari Muzakir (61500), based on BC which means Ari Muzakir is connected to many nodes in a large group, and also as an intermediary actor between other writers and if seen in his DC score namely 27 hereby provides evidence that apart from being connected to a large group and as an intermediary node, Ari Muzakir is also one of the authors with the largest number of connected studies and as an intermediary node between other nodes.

3.2 Determine Faculty Productivity in Producing Publications

3.2.1 Degree Centrality

This analysis aims to identify the most central or influential actors in the network. The measures used in this analysis were centrality, both entry and exit, centrality of proximity, and centrality of in-between. The following 3-degree center chart can be seen below.

Table 3 : Degree Centrality of Faculty

No	Faculty	DC	BC
1	Computer Science	70	2415.0
2	Economy and Business	37	666.0
3	FKIP	21	430.0
4	Vocation	21	210.0
5	Technique	20	190.0
6	Language and Literature	11	265.0
7	Communication	10	45.0
8	Psychologist	6	15.0

And based on the Degree Centrality Table, the results are the same. The Faculty of Computer Science is the faculty with the highest number of DC researches of 70 with BC, which is balanced with the previous table of 2415.

3.2.2 Betweenness Centrality

Betweenness centrality is a measurement at the points where there are points that are present when we walk across from one point to another. The point that has a high betweenness centrality is considered as the most 'strong' influential point in the network. If these points with high betweenness centrality are removed from the network, it will disrupt all communication between other points because they are on the path with the maximum number of passes. The following Betweeness Centrality can be seen in Table 4 below:

Table 4 : Betweenness Centrality of Faculty

No	Faculty	DC	BC
1	Computer Science	70	2415.0
2	Economy and Business	37	666.0
3	FKIP	21	430.0
4	Language and Literature	11	265.0
5	Vocation	21	210.0
6	Technique	20	190.0
7	Communication	10	45.0

Based on betweeneess centrality from Table 4. Above it can be seen that the BC score is 2415 with a total of 70 DCs, with this the Faculty of Computers is a faculty that is connected with the most nodes and other faculties, along with a high number of DCs of 70

4. CONCLUSION

Based on the results of this study, the structure of the authorship network for Universitas Bina Darma lecturers can be concluded that Universitas Bina Darma lecturers still do very little collaboration in research, both Journals or Proceedings and multidisciplinary scientific fields and across universities. Actors (lecturers) who have the highest degree, namely those with the highest DC values, this shows that UE (DC = 28*), AR (DC = 27*) and MIH (DC = 24*).

Is the actor who publishes the most Journals and Proceedings. Actors (lecturers) who have relationships or connect and often cooperate in publishing have the highest BC values, namely those with the greatest BC values, this shows that AM (BC = 61500), YKN (BC = 52545), and AHM (BC = 50748) is the actor who does the most cooperation or collaboration in publishing Journals and Proceedings. We can conclude that the faculties that produce the most publications are the computer science faculties, the economics and business faculties and the vocational faculties..

REFERENCES

[1] A. Darko, A. P. C. Chan, X. Huo, and D.-G. O. Manu, “A Scientometric Analysis and Visualization of Global Green Building Research,” *Build. Environ.*, vol. 149, pp. 501–511, 2019.

[2] B. Zhong, H. Wu, H. Li, S. Sepasgozar, H. Luo, and L. He, “A Scientometric Analysis and Critical Review of Construction Related Ontology Research,” *Autom. Constr.*, vol. 101, pp. 17–31, 2019.

[3] A. Higaki, T. Uetani, S. Ikeda, and O. Yamaguchi, “Co-authorship Network Analysis in Cardiovascular Research Utilizing Machine Learning (2009–2019),” *Int. J. Med. Inform.*, vol. 143, p. 104274, 2020.

[4] V. M. Patel *et al.*, “Collaborative Patterns, Authorship Practices and Scientific Success in Biomedical Research: a Network Analysis,” *J. R. Soc. Med.*, vol. 112, no. 6, pp. 245–257, 2019.

[5] D. N. Sari, D. Syamsuar, and E. S. Negara, “Structure Community Analysis on Social Network,” in *In The 6th International Conference on Information Technology and Business Application (ICIBA2017)1*, Pusat Penerbitan dan Percetakan Universitas Bina Darma Press (PPP-UBD Press) Palembang, 2017, pp. 1–7.

[6] J. Kim and M. Hastak, “Social Network Analysis: Characteristics of Online Social Networks After a Disaster,” *Int. J. Inf. Manage.*, vol. 38, no. 1, pp. 86–89, 2018.

[7] E. S. Negara, D. Kerami, I. M. Wiryani, and T. B. M. Kusuma, “Researchgate Data Analysis to Measure the Strength of Indonesian Research,” *Far East J. Electron. Commun.*, vol. 17, no. 5, pp. 1177–1183, 2017.

[8] R. Andryani, E. S. Negara, and D. Triadi, “Social Media Analytics: Data Utilization of Social Media for Research,” *J. Inf. Syst. Informatics*, vol. 1, no. 2, pp. 193–205, 2019.

[9] W. Anjar *et al.*, “Data Mining: Algoritma dan Implementasi,” *Yayasan Kita Menulis*, 2020.

[10] D. F. Brianna, E. S. Negara, and Y. N. Kunang, “Network Centralization Analysis Approach in the Spread of Hoax News on Social Media,” in *In 2019 International Conference on Electrical*

Engineering and Computer Science (ICECOS), IEEE, 2019, pp. 303–308.
[11] R. Amanda and E. S. Negara, "Analysis and Implementation Machine Learning for YouTube Data Classification by Comparing the Performance of Classification Algorithms," *J. Online Inform.*, vol. 5, no. 1, pp. 61–72, 2020.
[12] E. S. Negara, D. Triadi, and R. Andryani, "Topic Modelling Twitter Data with Latent Dirichlet Allocation Method," in *International Conference on Electrical Engineering and Computer Science (ICECOS)*, BATAM, 2019, pp. 386–390.
[13] T. Sutabri, A. Suryatno, and E. S. Negara, "Improving Naïve Bayes in Sentiment Analysis for Hotel Industry in Indonesia," in *In 2018 Third International Conference on Informatics and Computing (ICIC)*, IEEE, 2018, pp. 1–6.
[14] E. S. Negara and R. Andryani, "A Review on Overlapping and Non-Overlapping Community Detection Algorithms for Social Network Analytics," *Far East J. Electron. Commun.*, vol. 18, no. 1, pp. 1–27, 2018, doi: 10.17654/ec018010001.
[15] J. Zhang and Y. Luo, "Degree Centrality, Betweenness Centrality, and Closeness Centrality in Social Network," in *In Proceedings of the 2017 2nd International Conference on Modelling, Simulation and Applied Mathematics (MSAM 2017)*, 2017, pp. 300–303.