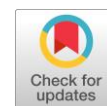


An extended approach of weight collective influence graph for detection influence actor



Galih Hendro Martono ^{a,1}, Azhari Azhari ^{b,2,*}, Khabib Mustofa ^{b,3}

^a Department of Informatics, Universitas Bumigora, Indonesia

^b Department of Computer Science and Electronics, Universitas Gadjah Mada, Indonesia

¹ galih.hendro@mail.universitasbumigora.ac.id; ² arisn@ugm.ac.id; ³ khabib@ugm.ac.id

* corresponding author

ARTICLE INFO

Article history

Received February 10, 2022

Revised March 15, 2022

Accepted March 31, 2022

Available online March 31, 2022

Keywords

Node central

Centrality measure

Collective influence

Key actor

Weight collective influence graph

ABSTRACT

Over the last decade, numerous methods have been developed to detect the influential actors of hate speech in social networks, one of which is the Collective Influence (CI) method. However, this method is associated with unweighted datasets, which makes it inappropriate for social media, significantly using weight datasets. This study proposes a new CI method called the Weighted Collective Influence Graph (WCIG), which uses the weights and neighbor values to detect the influence of hate speech. A total of 49, 992 Indonesian tweets were and extracted from Indonesian Twitter accounts, from January 01 to January 22, 2021. The data collected are also used to compare the results of the proposed WCIG method to determine the influential actors in the dissemination of information. The experiment was carried out two times using parameters $\delta=2$ and parameter $\delta=4$. The results showed that the usernames *bernacleboy* and *zack_rockstar* are influential actors in the dataset. Furthermore, the time needed to process WCIG calculations on HPC is 34-75 hours because the larger the parameter used, the greater the processing time.



This is an open access article under the [CC-BY-SA](#) license.



1. Introduction

Hate speech is one of the significant topics of discussion related to social media analysis. It is mainly associated with the freedom of users to share content and opinions on existing social media platforms. Freedom of opinion in social media has also led to an increase in the number of hate speech through social media. This increase is one of the challenges faced by the government in uncovering influential actors of hate speech on social media and issuing a legal law in the form of the Information and Transaction Law of Electronics (UU ITE). One of the methods used to uncover influential actors is the centrality measure, which means the vital node in a network. It represents the influence of a person whose presence on the network tends to dominate others. There are numerous benefits associated with detecting actors on the network, such as explaining the network's dynamics. Some of the traditional methods widely used to detect the influence of a node in a network include Betweenness Centrality (BC), Degree Centrality (DC), and Closeness Centrality (CC) [1]. These metrics are evaluated according to the number of node connections, relationship with neighbours, and the path that crosses the node [2]. DC only looks at the target node's information with low accuracy and time complexity [3]. DC and CC have a better result than BC, which has high time complexity [4]. Other methods include Cross-Face Centrality [5]. Rahim et al. [6] collaborated centrality and similarity measurements used for friendly recommendation. Detection of the central or influential actors in a network is mainly out on social media, such as Twitter [7]–[12]. Besides, influential actors implement other fields, such as, cybercrime [13]–[15] and e-commerce [16].

Mittal and Bhatia [17] conducted a research to develop traditional methods using Closeness Centrality in multiple social networks. Laishui et al. [18] used eigenvector centrality based on node similarity to rank nodes in multilayer and temporal networks. Nakajima et al. [8] and Takagi et al. [12] used BC to evaluate nodes using the statistic approach. Sarbani et al. [19] implemented degree, betweenness, eigenvector, subgraph, and edge clustering coefficient centralities using the PPI network dataset. The results showed that the Edge clustering coefficient is suitable for the dataset. Conversely, subgraph and eigenvector centralities showed poor performance. Similar research conducted by [20] proposed Neighborhood Closeness Centrality (NCC) in the PPI network. The computational results on the *S. cerevisiae* and *E. coli* dataset showed that NCC has the best accuracy and better performance than the random walk model. Several shortcomings are associated with the traditional method, such as the closeness centrality. Complexity and processing time was generated due to the large size of data and the Dynamic Nature of social media, which has less accuracy in DC [21], with numerous stability problems. The process of overcoming the noise contained in the network was carried out in [22], [23]. In addition, the centrality method based on the heuristic is unable to optimize the global function of the influential node, thereby guaranteeing its performance [24]. Another method was also developed by Morone et al. [25], known as Collective Influence (CI). This method measures the centrality of a node by analyzing the associated number of ties. The CI uses DC to calculate the number of neighboring node bonds. This implies the greater the number of connected nodes, the more significant its dissemination opportunity.

Based on these problems, Morone et al [25] proposed a new Collective Influence (CI) method, which introduces an influential node based on the number of connected neighbouring nodes. After introducing the CI method, many studies were carried out to improve its performance [24], [26]–[29]. Research conducted by Morone et al. [24] added a *heapification function* as a subtree at the node i to sort the largest to smallest CI values in the existing subtree nodes. Similarly, Kobayashi et al [26] stated that each connected node is considered a separate community used to calculate the centrality. Furthermore, Teng et al [27] proposed developing the CI method by adding a Linear Threshold Model (LTM). Another researcher by Kong et al. [28] introduced the concept of the probably-established subcritical path (PSP) to determine the distribution of information based on the path traversed by nodes. The results showed that PSP-based CI performed better than CI-TM. Other studies Wu et al [29] also proposed Enchanted Collective Influence (ECI), a CI method developed by adding features to overcome the local network topology's loop density and degree diversity. However, the concept of the CI method can only be used for unweighting graphs. Based on this, the purpose of this research is to introduce a new method of detecting influential actors in a network, specifically in hate speech known as WCIG. The contributions are as follows; (1) The WCIG method handles weighted graphs by adding a weight parameter based on the user's number of followers; (2) Implementation of WCIG with different parameters δ , is used to determine the coverage of neighbouring nodes.

2. Method

This research proposes the WCIG method, which was later used by Morone et al [25] to develop CI. DC is one of the traditional methods used to determine the influential actors in a network. This method is based on calculating the number of nodes connected to the degree. The formula used to calculate the degree of centrality is stated as follows [30].

$$C_D(i) = \frac{\sum_{j=1}^n a(j,i)}{n-1} \quad (1)$$

Where n is the number of nodes in the graph and $a(j, i) = 1$ assuming a relationship between nodes j and i . Conversely, $a(j, i) = 0$ supposing then it simply implies no relationship between node j and node i . The value $C_D(i) = 0$ indicates that the node i is an isolated node because it is not connected to any node. The degree centrality method is the basic concept used to calculate the centre of a node based on its relationship with others. This method was further developed by Opsahl [31] using a weighted graph. It was concluded that each node offers a different contribution, therefore, it is necessary to

consider their weights. This method combines the number of relations or connectedness between nodes (degree) with its strengths. Due to this, Opsahl et al [31] used the tuning parameter (α) to determine the number of relations compared to the weights. The Opsahl method is based on the Eq. (2) [31].

$$C_D^{w\alpha}(i) = k_i \left(\frac{s_i}{k_i}\right)^\alpha = k_i^{(1-\alpha)} s_i^\alpha \tag{2}$$

Here $C_D^{w\alpha}$ is the centrality of a node i with weight w , k_i which are several degree in node i , s_i is weight in node i and α is a tuning parameter. This is similar to the degree centrality method, proposed by Opsahl [31] which is based on the connectivity between two nodes (local). According to Opsahl, it is necessary to calculate the connection or influence of each node connected to the network (global). This method was finally developed by Morone [25], and it is called Collective Influence (CI). The CI calculation is based on the number of ties formed between a node and its neighbor [25]. Moreover, the greater the number of nodes connected to it, the higher the value of its centrality. In addition, the collective influence method considers the neighboring node network. It spreads faster when connected to the centre and has a more excellent value. Furthermore, this method is also based on the number of nodes connected to it and the neighboring value (∂). The formula for Collective Influence is stated in Eq. (3) [24].

$$CI_\ell(i) = (k_i - 1) \sum_{j \in \partial(i, \ell)} (k_j - 1) \tag{3}$$

Where k denotes the number of ties (degree), i, j are the nodes used in the graph, and $j \in \partial(i, l)$ is the range of neighboring nodes (j) of the source (i) at the radius l . The value CI_ℓ is calculated for each node based on its connectivity. Next, they are sorted based on the CI_ℓ value, with the largest identified as an influential actor in the network. This process is repeated until none of the nodes is connected or an isolated graph is plotted. In Fig. 1 it is known that the WCIG method is not only used to determine the centrality value between two nodes rather it also calculate the centrality of the connected nodes. For example, in Fig. 1, parameter $\partial = 2$ (marked as a line with red color), the WCIG value of node A is based on the connection between nodes B, C, D, E, W, O, H, I, K, and M. The greater the neighboring value (∂), the wider the influential node. However, this tends to reduce computational performance [32]. The parameter ∂ is used to determine the extent of the relationship between them. An illustration is shown in Fig. 1.

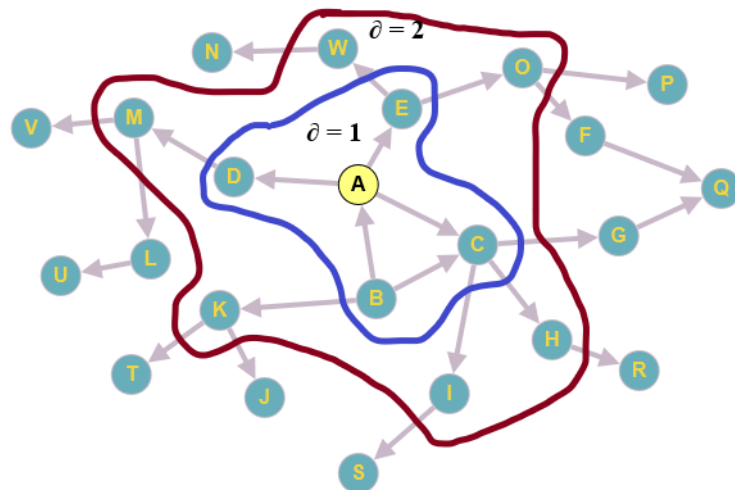


Fig. 1. Illustration of the neighbour node with parameter ∂ in a graph

This study proposes a new method called Weighted Collective Influence Graph (WCIG). It was developed because the CI approach is only based on the presence or absence of relations, thereby leading to information loss. The network's topology is difficult to explain [31]. In addition, according to

Rachman [33], interactions between users on social media have varying intensities. Therefore, the higher the interactive session power, the greater the flow of information. Furthermore, the use of weights to determine actor centrality provides better accuracy when compared to unweighted graphs [34]. This research is the development of a CI method using DC, which calculates the value of the centrality of a node based on its relationship. The development of the WCIG method is carried out by adding weights to the calculation based on Opsahl, which combines the number of bonds and its weights in the graph. The tuning parameter (α) determines the number of relations compared to the weights, and its formula is shown in Eq. (4). The traditional CI method was developed from the WCIG by adding degree and weight. The measure of the WCIG is shown in the following formula.

$$WCIG(i) = (k_i - 1)^{(1-\alpha)} s_i^\alpha \sum_{j \in \partial Ball(i, \ell)} (k_j - 1)^{(1-\alpha)} s_j^\alpha \quad (4)$$

Where k represents the number of connected nodes (degree), S_i represents nodes' weight, S_j represents nodes' weight j and α is a tuning parameter. In this study, the two attributes used in the WCIG calculation, are the number of connected nodes (degree) and the total weight of each. The number of ties is obtained based on the retweet interaction that exists between each user. In addition, when user A retweets user B, the users are connected. The total weight is obtained from followers' interactions. In disseminating information on social media Twitter, retweets and followers' interactions play a significant role. Meanwhile, with the retweet interaction, users indirectly disseminate information to those connected to them. This is also realized through followers' interaction because everyone that follows a user tends to be able to see the posted information. The proposed WCIG method is used to determine the value of every node per iteration. In addition, it further deletes those with the most significant WCIG value. This research aims to determine the impact of information dissemination by removing the most influential node followed by the subsequent ones. Iterations are repeated until an isolated node is formed. The pseudocode of the WCIG method is shown in Algorithm 1, Fig. 2.

```

Algorithm CentralityNodeWCIG ( $G, m$ )
Input  $G = (V, E, W)$ ;  $m$  = number of node centrality
Output list rank of node of influence actor
 $\forall i, \quad WCIG(i) = (k_i - 1)^{(1-\alpha)} s_i^\alpha \sum_{j \in \partial Ball(i, \ell)} (k_j - 1)^{(1-\alpha)} s_j^\alpha$ 
 $j \leftarrow 0$ ;
while  $\forall i, \quad WCIG(i) = (k_i - 1)^{(1-\alpha)} s_i^\alpha \sum_{j \in \partial Ball(i, \ell)} (k_j - 1)^{(1-\alpha)} s_j^\alpha$ 
 $j + +$ ;
 $WCIGmax_j \leftarrow \max(WCIG(i))$ 
 $G \leftarrow \text{RemoveEdges}(WCIGmax_j)$ 
 $\forall i, \quad WCIG(i) = (k_i - 1)^{(1-\alpha)} s_i^\alpha \sum_{j \in \partial Ball(i, \ell)} (k_j - 1)^{(1-\alpha)} s_j^\alpha$ 
endwhile
Return listrankofnode ( $\forall i, WCIGmax_j$ )

```

Fig. 2. The pseudocode of the WCIG method

3. Results and Discussion

3.1. Dataset

Data were collected through Twitter Streaming API and the dataset used in this research is mainly in the Indonesian Language. Crawling was carried out by retrieving tweets containing hate speeches based on keywords obtained from <https://hatebase.org/> such as babi, banci, bule, cabo, celeng, cina,

kafir, pelacur, lonte, munafik, perempuan jalang, perempuan nakal, singkek, sundal, bencong, bagong, waria, dan binal. As many as 18 keywords were crawled from January 01 to 22, 2021. The summary of the number of tweets from crawling and description of the total dataset is shown in Table 1, and Table 2 respectively.

Table 1. Summary of data from crawling

Keyword	Period of data collecting			Total
	Jan 4 th	Jan 13 th	Jan 22 th	
Babi	0	0	6,380	6,380
Banci	38	0	3,074	3,112
Bule	0	0	6,380	6,380
Cabo	49	0	102	151
Celeng	48	60	764	872
Cina	0	0	6,380	6,380
Kafir	0	0	6,380	6,380
Pelacur	8	47	1,728	1,783
Lonte	0	0	3,667	3,667
Munafik	0	0	21,370	21,370
Perempuan Jalang	15	9	8	32
Perempuan Nakal	68	24	45	137
Singkek	23	0	38	61
Sundal	60	65	997	1,122
Bencong	53	0	2,409	2,462
Bagong	32	8	2,692	2,732
Waria	0	0	557	557
Binal	0	0	1,412	1,412

Table 2. Description of the dataset

Information	Values
Number of data	49,992
Number of nodes	45,493
Number of edges	50,848
Average degree	1.11
Network diameter	16

3.2. Experiment Scenario

The pre-processing technique, consisting of data cleaning, tokenization, filtering, stemming, and stop-word was carried out before implementing the WCIG method. First, the data cleaning process includes deleting unused information, missing values and ensuring that those used adheres to the graphic format. Afterwards, the process of tokenization aims to separate constituent words. Filtering is carried out by removing the word "RT" and symbols in the tweet data. The stemming process is used to change implicating words into essential words. Meanwhile, the stop-word removal is the process of removing terms that are considered less critical such as conjunctions "dan," "jika," "atau," "tetapi," etc. The pre-processing results were further used to generate data graph. This is carried out by connecting user based on retweet and mention (marked with symbol "@" or "RT"). When the tweet contains both symbols then the user is connected to those that retweeted. Afterwards, the detection influence actor is implemented with WCIG, and the result is visualized through a graphical representation of the relations between nodes. The test was carried out two times with parameters $\partial = 2$ and $\partial = 4$. The adoption of various parameters is to determine changes in the WCIG value of each user in respect to its increase in the same dataset. The larger the parameter ∂ used, the greater the coverage of calculated neighboring nodes. The dataset used has a network diameter of 16, this simply implies that 1 node is connected to 16 neighboring ones (16 levels). The results of the implementation are further analyzed to determine the most influential users; besides, this is carried out using a High-Performance Computer (HPC). The detail of the experiment is shown in Fig. 3.

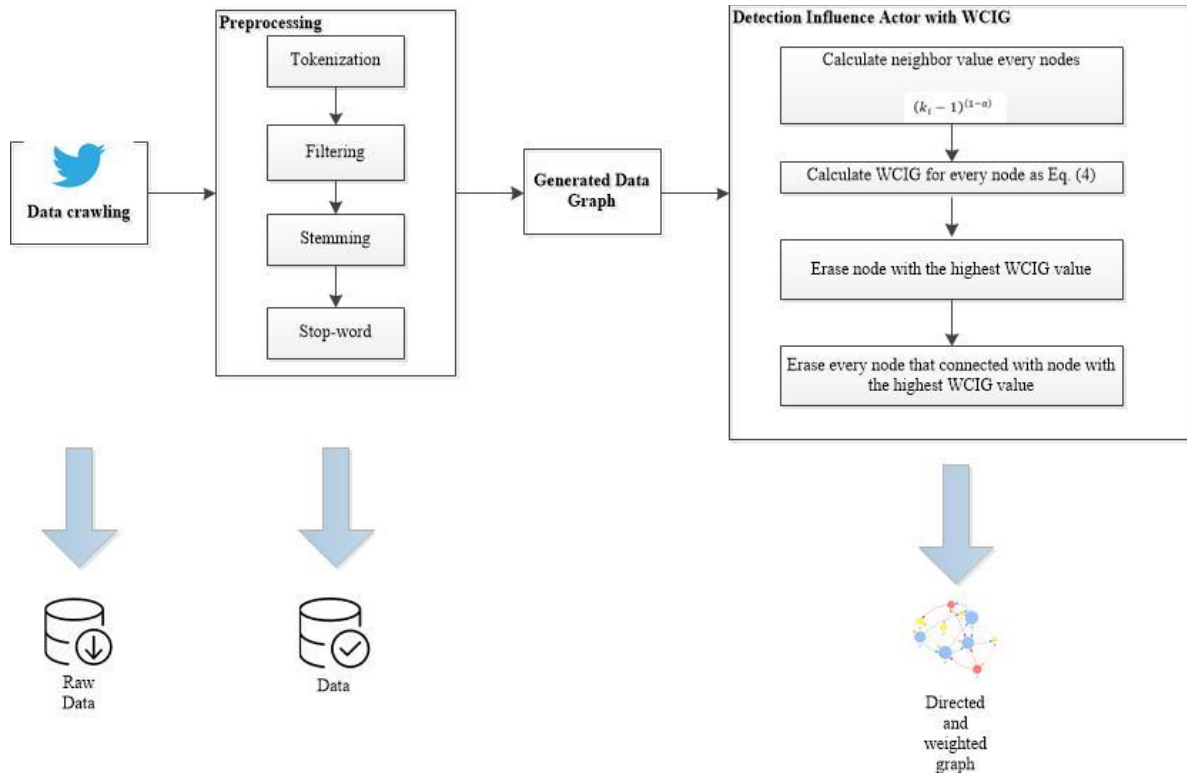


Fig. 3. Detail experiment

3.3. Results

The results are divided into two sections. The first implements WCIG and CI methods using parameters $\partial = 2$ and $\partial = 4$. Furthermore, Kendall's Tau test was carried out to determine the suitability between the results of both methods.

3.3.1. Experiment Result of WCIG and CI

This section presents experimental results using the WCIG method and compares it with CI. The WCIG calculation is based on the number of ties that connects one user to another (retweet interactions), the total weight (interactions with followers) and the number of connected users. The process allows the expansion of information dissemination from neighboring nodes. This concept determines the impact of an influential user, not only based on its connection with others, rather, it also includes all those connected to them. This is in contrast to CI, which calculates centrality based on connected nodes and relationship with neighboring nodes. This research is implemented using Python language with four experiments, namely WCI with parameter $\partial = 2$, and 4, as well as CI with parameter $\partial = 2$, and 4. Implementation with 2 different ∂ parameters is used to determine the effect and differences in the nodes' coverage. The larger the parameter used, the better the results because more nodes are counted. The experiment results are 10 users who influence the spread of hate speech on Twitter. The comparison of the results of the WCIG experiment with the CI is shown in Table 3.

Table 3 shows the results of each experiment yield by 10 users in implementing the two methods with different parameters. Several users appear in each experiment, namely bernacleboy, zack_rockstar, dr_koko28, hafizzismailz, anwarrahmad, Republikaonline, fahmirusliMFT, and AzzamIzzulhaq. Meanwhile, kakti_64 appeared in the WCI experiment with parameters 2 and 4, and CI with parameter $\partial = 2$. Rudyroutepecker users appeared in WCI and CI experiments with $\partial = 4$, respectively. Norazambudin and Fundulus users only appeared on WCI experimental results with parameter $\partial = 2$ and 4.

Table 3. Comparison between WCIG and CI

Ranking	Parameter $\partial = 2$		Parameter $\partial = 4$	
	WCI	CI	WCI	CI
1	bernacleboy	Zack_rockstar	bernacleboy	Zack_rockstar
2	Zack_rockstar	bernacleboy	Zack_rockstar	bernacleboy
3	dr_koko_28	Rudyroutepecker	dr_Koko_28	dr_koko_28
4	Hafizismailz	Hafizismailz	Hafizismailz	Hafizismailz
5	Anwarrahmad	dr_koko_28	Republikaonline	Republikaonline
6	Republikaonline	Kakti64	Rudyroutepecker	Rudyroutepecker
7	fahmirusliMFR	Republikaonline	fahmirusliMFR	Kakti64
8	Kakti64	Anwarrahmad	Anwarrahmad	Anwarrahmad
9	Norazambudin	fahmirusliMFR	Fundoelous	AzzamIzzulhaq
10	AzzamIzzulhaq	AzzamIzzulhaq	AzzamIzzulhaq	fahmirusliMFR

An example of information dissemination is evident in username bernacleboy, with 2,271 types of retweet interactions and 1,732 followers. The spread of information started when bernacleboy posted a tweet on Twitter, which was seen by all 1,732 followers. Furthermore, other users retweeted the tweet, while 2,271 others engaged in interactions. This caused led to the occurrence of information dissemination at level 1 of the interaction session. Everyone following the 2,271 users that retweeted the tweet were able to see the outcome thereby spreading information. This interaction is referred to as level 2 information dissemination. As long as other users retweet the information, more people read it and indirectly spread the information. The interaction causes the user with the username bernacleboy to be the most influential person and the first to disseminate the information. Interestingly, this occurred in the form of the username _bernacleboy’s followers that also shared the information both internally and externally, directly or indirectly. Visualization of the graphical representation is shown in Fig. 4 where the yellow-colored node represents the username bernalceboy and the black colored ones depicts other connected users.

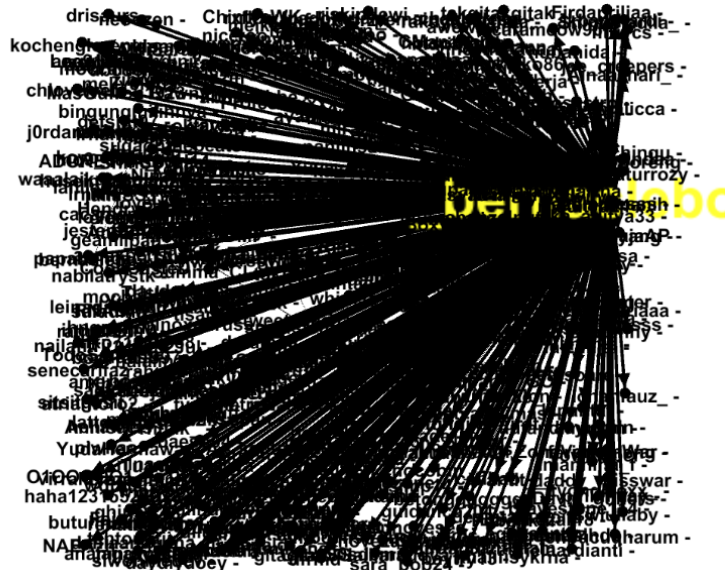


Fig. 4. Visualization of a graph from user bernacleboy

As illustrated in Fig. 4, in the first iteration process, it is evident that bernacleboy is the most influential user, therefore it is important to delete the 2,271 connected users. This process affects the connected nodes based on their degree. Next, the WCIG value will be recalculated based on the updated value of each node. This iteration shows that the username zack_rockstar has the highest WCIG value. Therefore, connected users are deleted, and the iteration is repeated until an isolated node is formed. The tests showed that the greater the parameter value ∂ , the more time needed to process the WCIG method. When parameter $\partial = 2$, the processing time is 120,557.94 seconds. The number of iterations needed to increase the value ∂ to 4 was 148,883.42 at an average time of 35.11 seconds. The condition

also applies to the number of iterations needed to process the WCIG method from 3,433 iterations for parameter $\partial = 2$ to 3,485 for parameters $\partial = 4$. The processing time for the CI method is similar to the WCIG. CI with parameter $\partial = 2$ takes 115,907.43 second at an average iteration time of 34.63 second and 3,347 iterations. CI with parameter $\partial = 4$ takes 140,390.31 seconds with average iteration time of 41.51 seconds and 3,382 iterations. The time complexity required to complete this algorithm is $O(N^3)$. The time complexity quite different with previous methods such as CI which is $O(N \log N)$ [26]. A summary of the tests is shown in Table 4.

Table 4. Summary of implementation WCIG and CI

Computer Specifications	Method Parameter	WCI		CI	
		$\partial = 2$	$\partial = 4$	$\partial = 2$	$\partial = 4$
CPU Intel®	Number of iterations	3,433	3,485	3,347	3,382
Core (TM) i9-7900X CPU @ 3.3GHz, RAM 16Gb, GeForce RTX 1080Ti	Time for all iterations (second)	120,557,94	148,883.42	115,907.43	140,390.31
	Average time needed for every iteration (second)	35.11	42.72	34.63	41.51

3.3.2. Kendall's Tau Correlation Coefficient

Kendall's Tau coefficient is a non-parametric statistic used to measure the degree of correspondence between two rankings [3]. Its correlation coefficient test is used to compare the suitability or relationship between the WCI and the CI methods. Assuming variable X is the WCI method, and Y is CI methods are $X = \{X_1, X_2, X_3, \dots, X_n\}$ and $Y = \{Y_1, Y_2, Y_3, \dots, Y_n\}$. The formula used to calculate Kendall Tau Correlation is as follows [35].

$$\tau = \frac{n_c - n_d}{\frac{n \times (n-1)}{2}} \quad (5)$$

Where n_c , n_d are the concordant and discordant pairs, respectively and n is the total number of pairs. The hypothesis obtained is as follows; (1) H_0 : There is no match between the results of the WCI and CI methods; (2) H_1 : There is a match between the results of the WCI and CI methods. The implementation is carried out with the SPSS 26 application on WCI and CI outputs using the parameters $\partial = 2$ and 4. The Kendall's Tau correlation coefficient test results for the WCI and CI methods using the parameter $\partial = 2$ and 4 are shown in Table 5 and Table 6.

Table 5. Kendall's tau coefficient correlation WCI and CI parameter $\partial = 2$

			WCI	CI
Kendall's tau_b	WCI	Correlation Coefficient	1.000	0.491*
		Sig. (2-tailed)	-	0.036*
		N	11	11
	CI	Correlation Coefficient	0.491*	1.000
		Sig. (2-tailed)	0.036*	-
		N	11	11

*Correlation is significant at the 0.05 level (2-tailed)

Table 6. Kendall's tau coefficient correlation WCI and CI with parameter $\partial = 4$

			WCI	CI
Kendall's tau_b	WCI	Correlation Coefficient	1.000	-0.822**
		Sig. (2-tailed)	-	0.001
		N	10	10
	CI	Correlation Coefficient	-0.822**	1.000
		Sig. (2-tailed)	0.01	-
		N	10	10

**Correlation is significant at the 0.01 level (2-tailed)

The first test of Kendall's Tau correlation shown in Table 5 obtained a correlation coefficient value of 0.491 between the WCI and CI variables. It can be concluded that the relationship between the WCI and CI variables is strong. In addition, the significance value or sig. (2-tailed) between both variables is $0.036 < 0.05$, it can be concluded that H_0 is rejected and H_1 is accepted. The result means that there is a match between both variables. Similar with the second test in Table 6, the correlation coefficient value between the WCI and CI variables is -0.822. It can be concluded there is a solid negative relationship between both variables. The significance value or sig. (2-tailed) between WCI and CI variables is $0.001 < 0.01$, and it can be concluded that H_0 is rejected and H_1 is accepted.

4. Conclusion

This research shows the implementation of the WCIG method, which is used on weighted and directed datasets instead of the CI method. Furthermore, this weight represents the number of interactions between one user and another. In addition, the neighbor value is also one of the WCIG methods parameter that significantly contribute to detecting influential actors, especially in hate speech. The result showed that there are correlation between WCIG and CI using Kendall's Tau coefficient. Furthermore, the time complexity between WCIG and CI is different, which is $O(N^3)$ and $O(N \log N)$, respectively. The time complexity showed that time to process CI is slightly faster compare than WCIG. Although in the experiment the different not too much. There are several limitations associated with the use of WCIG methods. One of them is related to the weight used, which is obtained from the number of followers. Therefore, future research needs to be carried out using the WCIG method with other Twitter interactions such as tweets, retweets, follows, and mentions. Other approaches, such as fuzzy can be used to determine the interaction between users. Furthermore, it is hoped that the WCIG method is continuously used to detect influential actors and the associated communities in a dataset.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding statement. None of the authors have received any funding or grants from any institution or funding body for the research.

Conflict of interest. The authors declare no conflict of interest.

Additional information. No additional information is available for this paper.

References

- [1] K. Zhang, H. Li, L. Qin, and M. Wu, "Closeness centrality on BBS reply network," in *Proceedings - 2011 International Conference of Information Technology, Computer Engineering and Management Sciences, ICM 2011*, 2011, vol. 2, pp. 80–82, doi: [10.1109/ICM.2011.88](https://doi.org/10.1109/ICM.2011.88).
- [2] B. C. B. De Figueiredo, F. G. Nakamura, and E. F. Nakamura, "A group-based centrality for undirected multiplex networks: a case study of the Brazilian Car Wash Operation," *IEEE Access*, vol. 9, 2021, doi: [10.1109/ACCESS.2021.3086027](https://doi.org/10.1109/ACCESS.2021.3086027).
- [3] Y. Yang, X. Wang, Y. Chen, and M. Hu, "Identifying Key Nodes in Complex Networks Based on Global Structure," *IEEE Access*, vol. 8, pp. 32904–32913, 2020, doi: [10.1109/ACCESS.2020.2973241](https://doi.org/10.1109/ACCESS.2020.2973241).
- [4] K. Das, S. Samanta, K. De, and M. Pal, "Complete neighbourhood centrality and its application," in *International Conference on Computational Intelligence and Networks, CINE 2020*, 2020, pp. 0–5, doi: [10.1109/CINE48825.2020.234386](https://doi.org/10.1109/CINE48825.2020.234386).
- [5] M. H. Ibrahim, R. Missaoui, and J. Vaillancourt, "Cross-Face Centrality: A New Measure for Identifying Key Nodes in Networks Based on Formal Concept Analysis," *IEEE Access*, vol. 8, pp. 206901–206913, 2020, doi: [10.1109/ACCESS.2020.3038306](https://doi.org/10.1109/ACCESS.2020.3038306).
- [6] M. Rahim, A. Ejaz, Q. Rajput, and K. A., "A Comparative Study of Similarity and Centrality Measures for Friends Recommendation," in *International Conference on Information Science and Communication Technology (ICISCT)*, 2019, doi: [10.1109/CISCT.2019.8777419](https://doi.org/10.1109/CISCT.2019.8777419).

- [7] A. J. Oroh, Y. Bandung, and L. M. Zagi, "Detection of the Key Actor of Issues Spreading Based on Social Network Analysis in Twitter Social Media," in *IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, 2021, pp. 206–212, doi: [10.1109/APWiMob51111.2021.9435268](https://doi.org/10.1109/APWiMob51111.2021.9435268).
- [8] K. Nakajima, K. Iwasaki, T. Matsumura, and K. Shudo, "Estimating Top-k betweenness centrality nodes in online social networks," in *Proceedings - 16th IEEE International Symposium on Parallel and Distributed Processing with Applications, 17th IEEE International Conference on Ubiquitous Computing and Communications, 8th IEEE International Conference on Big Data and Cloud Computing, 11t*, 2019, pp. 1128–1135, doi: [10.1109/BDCLOUD.2018.00167](https://doi.org/10.1109/BDCLOUD.2018.00167).
- [9] Q. Liao and S. Sinthupinyo, "Exclusiveness Centrality: A Novel Centrality Measure to Find Behavioral Influencers on Twitter," in *Proceedings of the 10th International Conference on Electronics, Computers and Artificial Intelligence, ECAI 2018*, 2019, doi: [10.1109/ECAI.2018.8679095](https://doi.org/10.1109/ECAI.2018.8679095).
- [10] A. K. Yadav, R. Johari, and R. Dahiya, "Identification of Centrality Measures in Social Network using Network Science," in *International Conference on Computing, Communication, and Intelligent Systems, ICCIS 2019*, 2019, pp. 229–234, doi: [10.1109/ICCCIS48478.2019.8974553](https://doi.org/10.1109/ICCCIS48478.2019.8974553).
- [11] M. Jaouadi and L. Ben Romdhane, "Influence maximization problem in social networks: An overview," in *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*, 2019, doi: [10.1109/AICCSA47632.2019.9035366](https://doi.org/10.1109/AICCSA47632.2019.9035366).
- [12] K. Takagi and K. Tago, "Comparison of measurements by the betweenness centrality and subjective experiment on the word priority of Tweets," in *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/TOP/SCI)*, 2018, pp. 1936–1941, doi: [10.1109/SmartWorld.2018.00324](https://doi.org/10.1109/SmartWorld.2018.00324).
- [13] M. C. K. Michel and M. C. King, "Cyber Influence of Human Behavior : Personal and National Security, Privacy, and Fraud Awareness to Prevent Harm," in *IEEE International Symposium on Technology and Society (ISTAS)*, 2019, pp. 1–7, doi: [10.1109/ISTAS48451.2019.8938009](https://doi.org/10.1109/ISTAS48451.2019.8938009).
- [14] S. Zhang, W. Si, T. Qiu, and Q. Cao, "Toward More Effective Centrality-Based Attacks on Network Topologies," in *IEEE International Conference on Communications*, 2020, vol. 2020-June, pp. 0–5, doi: [10.1109/ICC40277.2020.9148785](https://doi.org/10.1109/ICC40277.2020.9148785).
- [15] C. Mitchell, R. Agrawal, and J. Parker, "The Effectiveness of Edge Centrality Measures for Anomaly Detection," in *IEEE International Conference on Big Data, Big Data 2019*, 2019, pp. 5022–5027, doi: [10.1109/BigData47090.2019.9006468](https://doi.org/10.1109/BigData47090.2019.9006468).
- [16] H. M. Kavitha, S. Duncan, P. Ravikumar, and E. Varun, "Online Shopping Customer Behaviour Analysis using centrality measures," in *1st IEEE International Conference on Advances in Information Technology, ICAIT 2019 - Proceedings*, 2019, pp. 223–227, doi: [10.1109/ICAIT47043.2019.8987252](https://doi.org/10.1109/ICAIT47043.2019.8987252).
- [17] R. Mittal and M. P. S. Bhatia, "Cross-Layer Closeness Centrality in Multiplex Social Networks," in *2018 9th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2018*, 2018, doi: [10.1109/ICCCNT.2018.8494042](https://doi.org/10.1109/ICCCNT.2018.8494042).
- [18] L. Lv, K. Zhang, T. Zhang, X. Li, J. Zhang, and W. Xue, "Eigenvector centrality measure based on node similarity for multilayer and temporal networks," *IEEE Access*, vol. 7, pp. 115725–115733, 2019, doi: [10.1109/ACCESS.2019.2936217](https://doi.org/10.1109/ACCESS.2019.2936217).
- [19] S. Roy, P. Dey, and D. Kundu, "Social Network Analysis of Cricket Community Using a Composite Distributed Framework : From Implementation Viewpoint," *IEEE Trans. Comput. Soc. Syst.*, vol. 5, no. 1, pp. 64–81, 2018, doi: [10.1109/TCSS.2017.2762430](https://doi.org/10.1109/TCSS.2017.2762430).
- [20] G. Li, M. Li, J. Wang, Y. Li, and Y. Pan, "United Neighborhood Closeness Centrality and Orthology for Predicting Essential Proteins," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 17, no. 4, pp. 1451–1458, 2020, doi: [10.1109/TCBB.2018.2889978](https://doi.org/10.1109/TCBB.2018.2889978).
- [21] X. Zhao, S. Guo, and Y. Wang, "The Node Influence Analysis in Social Networks Based on Structural Holes and Degree Centrality," in *2017 IEEE International Conference on Computational Science and Engineering*

- (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), 2017, no. 1, pp. 708–711, doi: [10.1109/CSE-EUC.2017.137](https://doi.org/10.1109/CSE-EUC.2017.137).
- [22] S. Segarra and A. Ribeiro, “A Stable Betweenness Centrality Measure In Networks,” in *2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, 2014, no. 3, pp. 3859–3863, doi: [10.1109/ICASSP.2014.6854324](https://doi.org/10.1109/ICASSP.2014.6854324).
- [23] S. Segarra and A. Ribeiro, “Stability and continuity of centrality measures in weighted graphs,” in *IEEE Transactions on Signal Processing*, 2016, vol. 64, no. 3, pp. 543–555, doi: [10.1109/TSP.2015.2486740](https://doi.org/10.1109/TSP.2015.2486740).
- [24] F. Morone, B. Min, L. Bo, R. Mari, and H. A. Makse, “Collective Influence Algorithm to find influencers via optimal percolation in massively large social media,” *Nature*, vol. 94, no. 1, pp. 1–22, 2016, doi: [10.1103/PhysRevE.94.012305](https://doi.org/10.1103/PhysRevE.94.012305).
- [25] F. Morone, K. Roth, B. Min, H. E. Stanley, and H. A. Makse, “Model of brain activation predicts the neural collective influence map of the brain,” in *Proceedings of the National Academy of Sciences*, 2016, vol. 114, no. 15, pp. 3849–3854, doi: [10.1073/pnas.1620808114](https://doi.org/10.1073/pnas.1620808114).
- [26] T. Kobayashi and N. Masuda, “Fragmenting networks by targeting collective influencers at a mesoscopic level,” Nature Publishing Group, 2016. doi: [10.1038/srep37778](https://doi.org/10.1038/srep37778)
- [27] X. Teng, S. Pei, F. Morone, and H. A. Makse, “Collective Influence of Multiple Spreaders Evaluated by Tracing Real Information Flow in Large-Scale Social Networks,” in *Nature*, 2016, vol. 6, no. October, doi: [10.1038/srep36043](https://doi.org/10.1038/srep36043).
- [28] D. Kong, C. Fu, J. Yang, D. Xu, and L. Han, “The Impact of the Collective Influence of Search Engines on Social Networks,” in *IEEE Access*, 2017, vol. 5, pp. 24920–24930, doi: [10.1109/ACCESS.2017.2767075](https://doi.org/10.1109/ACCESS.2017.2767075).
- [29] T. Wu, L. Chen, L. Zhong, and X. Xian, “Enhanced collective influence: A paradigm to optimize network disruption,” in *Physica A: Statistical Mechanics and its Applications*, 2017, vol. 472, no. 2006, pp. 43–52, doi: [10.1016/j.physa.2016.12.036](https://doi.org/10.1016/j.physa.2016.12.036).
- [30] H. Li, “Centrality Analysis of Online Social Network Big Data,” in *2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA)*, 2018, pp. 38–42, doi: [10.1109/ICBDA.2018.8367648](https://doi.org/10.1109/ICBDA.2018.8367648).
- [31] T. Opsahl, F. Agneessens, and J. Skvoretz, “Node centrality in weighted networks: Generalizing degree and shortest paths,” in *Social Networks*, 2010, vol. 32, no. 3, pp. 245–251, doi: [10.1016/j.socnet.2010.03.006](https://doi.org/10.1016/j.socnet.2010.03.006).
- [32] F. Morone, B. Min, R. Mari, and H. A. Makse, “Influence maximization in complex networks through optimal percolation,” in *Nature*, 2016, doi: [10.1038/nature14604](https://doi.org/10.1038/nature14604).
- [33] Z. A. Rachman, W. Maharani, and Adiwijaya, “The analysis and implementation of degree centrality in weighted graph in Social Network Analysis,” in *2013 International Conference of Information and Communication Technology, ICoICT 2013*, 2013, pp. 72–76, doi: [10.1109/ICoICT.2013.6574552](https://doi.org/10.1109/ICoICT.2013.6574552).
- [34] H. Kretschmer and T. Kretschmer, “A New Centrality Measure for Social Network Analysis Applicable to Bibliometric and Webometric Data,” in *COLLNET Journal of Scientometrics and Information Management*, 2013, vol. 1, no. 1, pp. 1–7, doi: [10.1080/09720502.2007.10700947](https://doi.org/10.1080/09720502.2007.10700947).
- [35] H. Rastogi and M. Jain, “A correlative study of centrality measures across real-world networks,” in *Proceedings of the 4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020*, 2020, pp. 987–994, doi: [10.1109/I-SMAC49090.2020.9243484](https://doi.org/10.1109/I-SMAC49090.2020.9243484).