Malaysian Politicians' Connection Pattern on Twitter using SNA: A Case of Najib Razak

NUR AMALINA DIYANA SUHAIMI
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Malaysia
nuramalinadiyana@graduate.utm.my

SYAZA IFFAH MOHAMMAD SALLEH
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Malaysia
syazaiffah@graduate.utm.my

SITI AMANINA FARHANAH ABDUL HAKIM
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Malaysia
sitiamaninafarhanah@graduate.utm.my

PRITHEEGA MAGALINGAM
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Malaysia
mpritheega.kl@utm.my

NURAZEAN BINTI MAAROP
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Malaysia
nurazean.kl@utm.my

MOHANA SHANMUGAM
College of Computing and Informatics
Universiti Tenaga Nasional
Malaysia
mohana@uniten.edu.my

2021 International Congress of Advanced Technology and Engineering (ICOTEN) | 978-1-6654-1224-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICOTEN52080.2021.9493501

***Abstract*— Najib Razak is one of the most prominent politicians in Malaysia whose popularity has risen worldwide over the years due to his political sharp-witted strategy and various political scandals. He is also identified as one of the most followed Malaysian politicians on social media, especially Twitter. Hence, this study aims to apply Social Network Analysis (SNA) to further examine the interactions between Twitter users and the relationship formed with Najib Razak. A complete network of Najib Razak's Twitter account is used to study the connection pattern, influence, and groups developed between account users in the network. Netlytic is used to extract the data on Twitter, and based on the extracted dataset, it is discovered that 1004 nodes that represent Twitter users, follows and mentions the @najibrazak Twitter account. The dataset was further analyzed using R to explore the interaction and the connection patterns were visualized using Gephi. Based on the findings, the connectivity, centrality and clustering of the top 10 most influential Twitter users that contribute to the discussion and mention of Najib Razak on Twitter were determined. The previous work using Najib Razak's twitter account focused on finding the relations between public and politicians by analyzing the issues discussed through language processing at topical and lexical level. Unlike the previous achievement, the results from this proposed SNA technique can be further analyzed to gather greater insights on the hidden relationship built between politicians to strengthen their position and distinguish their possible future followers for further investigations.**

Keywords*— *Najib Razak, Twitter, Centrality, Connectivity, Clustering, Social Network Analysis, Politician, Malaysia

I. INTRODUCTION

Since its establishment in 2006, Twitter has become one of the most commonly used social media, linking users and enabling online interactions [1]. Najib Razak is a Malaysian politician who served from April 2009 to May 2018 as Malaysia's 6th Prime Minister. Mohammad Najib bin Tun Haji Abdul Razak, or also known as his pseudonym Najib Razak or tweet handle @najibrazak has one of the most Twitter followers with a total of 4,278,817 followers as of September 2020, according to Malaysia's Twitter usage statistics by Socialbakers. Many online media, however, reported that he had the largest number of fake followers on Twitter [2]. The explanation behind the formation of @najibrazak's Twitter account in 2008 is due to the political party, Barisan Nasional coalition's failure to gain a two-thirds of majority for the very first time, according to Liow [3]. When they started using the social media site as a way to connect and engage with their supporters and attempting to reach a broader Malaysian audience, an assertive move by the coalition party can be observed [3].

FRGS Grant, vote number R.K130000.7856.5F016

Via Twitter, Najib Razak is able to post personal opinions or messages ("tweets") on his page at any time to his followers and supporters, allowing for a connection that does not require a face-to-face interaction. He also able to communicate faster with the masses by retweeting, mentioning, or simply tweeting on Twitter, like other users. Retweeting is an act of sharing someone else's tweets. Meanwhile, mentioning is tagging another Twitter user in a tweet, and the 'like' button is used to show appreciation of relatableness of a particular tweet. Besides personal opinions and messages, Najib Razak often tweets his view on political news or current trending topics in the country.

Following that, in this study specifically, the aim is to examine the interactions and connections that appear in the network of Najib Razak's Twitter account by applying social network analysis (SNA) techniques. The connectedness of nodes and importance of actors in the targeted network will help us understand the relationship and factors influencing the connectivity between Twitter user accounts with Najib Razak. Other than that, the clustering measure can be used to comprehend the groups or subgroups developed between account users in the network that are highly connected to Najib Razak. Meanwhile, the betweenness centrality of each actors in the targeted network is used to determine how the messages of

Najib Razak can be further elaborated, discoursed and passed among Twitter users. Furthermore, the insight and the analysis result obtained assisted in identifying Twitter users who are the conversational hubs.

The past work using Najib Razak's Twitter account or the tweet handle @najibrazak analyses tend to focus more on political issues communicated through the post and the voice of power was determined using the language processing method. In this research, more network measures are explored to find its suitability to investigate the connection trend between users during the peak of any political issues. The related work in depth, motivation, SNA techniques proposed for the analysis and the results are discussed further in the following sections.

II. RELATED WORK

According to Himelboim [4], Twitter is among the most popular sites in social media, which appears to have more than 313 million monthly active users as of 30th of June 2016. Twitter enables users to build, tag, and post contents online. The Twitter API also allows posts, supporters, followers, and user profiles to be monitored and collected as data, allowing users to retrieve association data as well as individual profile information among users. In enforcing his or her will on other social actors, Castells [5] described power as the institutional ability of a social actor. However, power relationships are destructive because of the complex and inconsistent existence of society. Castells also stated that by representing competing values and desires and involving a multitude of social actors in producing and debating social activities, the Internet is said to have intensified the counter-power phase.

In specific, Twitter enables instant, rapid, and ubiquitous distribution of information. Even so, Twitter has shown the potential for social hierarchies to crumble and have a huge effect on the relationship between elected leaders and the people through the offering of a personal forum where political candidates communicate with voters in a more direct manner. More significantly, social media helps politicians to establish a deeper relationships with electors by reaching out to a specific demographic of voters. It also helps candidates to focus on how they want to be perceived by the average voter [6].

A previous recent study that was done on Twitter analysis by Kasmani [7] in 2019, studied the impact of Najib Razak's Twitter handle @najibrazak during the 2018 election. From the study, the political discourse of Najib Razak's tweets from 1st December 2016 to 28th February 2017 is found to be mainly related to the government's political structure, ideology, institutions, political actors, and political events from a topical perspective. It is determined that through the topical level and lexical level analysis conducted, Najib Razak had posts regularly on Twitter about the authority and government agenda, whereby the public shall follow and comply with the government's way of ruling the country. Other than that, Najib's tweets consisted of the sound of sympathy when he posted in the page, a few pictures of impacted individuals who were in need that were taken during his political visits. The study also found that Najib Razak often interacts with elite Twitter users [7] including politicians. However, other than the specificity on the phrases of the sample tweets that are analyzed by the author and the list of people mentioned by Najib Razak's tweets, the details on the connectivity and interactions between Najib Razak and Twitter users were not revealed. It is found that there was no unique SNA method proposed in this area of research that focuses specifically on Twitter of Najib Razak or tweet handle @najibrazak analysis.

Thus, the SNA elements and structures are adapted in this paper to explore the relationship that connect the followers to the Twitter account of @najibrazak and identify the measures that can reveal the density, interconnectedness, centrality and clustering of Twitter users in the Najib Razak network. Below are some of the related work that further explains the SNA measures which will be used in this paper.

A. Density & Interconnectedness

Density measures the interconnectedness of individuals in a network [4]. It can range from low density, a community of people is lightly connected to high density, where users are tightly interconnected [8-10]. The intensity of information that flows within it is influenced by the degree to which a network is firmly related. Carley [11] found that contact between people contributes to mutual knowledge, and knowledge sharing interjects to even more interaction, a finding that has important consequences for a group's cohesion and its interactions beyond its borders with individuals. Meanwhile, Granovetter [12] found that strong and repetitive links usually bind closely interrelated people.

Burt [13] observed that knowledge transmission is best for networks in which individuals are very strongly connected. Likewise, Coleman [14, 15] showed that a significant consequence of strongly established close relationships is an improvement in trust between people, which can lead to a higher transfer of knowledge. Other than that, Zubcsek et al. [16] found that the strength of contact within knowledge communities among individuals is larger than in other areas of the network. In a study of the role of social network structures in the dissemination of data on Digg and Twitter, Lerman and Ghosh [17] found that the rate at which data is spread across a network depends on its size. Thus, in this study, the analysis on the density were performed to discover the "potential connection" that exists between two different nodes in the Najib's name-network. It will also help us to understand how "close" the network, as closer the network entities, it can be concluded that the Twitter users in the network are associated and know one another.

B. Degree Centrality

A node value that affects a network is called centrality [18, 19]. In concept, centrality is a value reflecting how many node-to-other node connections exist. There are several approaches to describe centrality for each of the node results, such as the centrality of degree, centrality of betweenness, centrality of proximity, and centrality of eigenvectors.

Degree centrality is characterized as the calculation of the number of node-owned connections or the number of node-linked relationships. PageRank is an algorithm that can be used to provide the centrality score of nodes based on the connection that occurs between nodes. Based on the score, the ranking of the other 'pages' or nodes can be identified further [20]. These numbers can act as weight that reflects closeness or strength of contact between actors in each of the link formed [21]. Evaluating the relationship formed between actors is applicable by implementing link weight. In the [21], weighted links or relationship weight representing closeness and the degree of nodes that values the interaction frequency between nodes makes the outcome of the social network research more reliable and meaningful compared to unweighted ties.

Other than that, the centrality of the eigenvector would provide the most powerful node in a network. Among several network nodes, the most prominent node is the node with the highest eigenvector value. The essential factor in determining the most significant node in a network with centrality calculation using proprietor centrality is the importance of the proprietor amongst the adjacent nodes. Consequently, even though there is a node with low centrality of weight or degree, it may still be the most powerful node with the centrality of the eigenvector. By considering the centrality importance of its neighbouring vector, the accuracy of the most powerful user decision in a network will ideally improve [22, 23]. The PageRank analysis in this paper will determine the top important Twitter users that are connected to the prime node, @najibrazak.

In this research, the analysis also includes identifying users with high betweenness centrality as it gives information on how the important nodes sustain the communication flow and influence the connectivity between Twitter users within the Najib Razak Twitter name-network when they are removed. Node with the highest betweenness centrality may act as a bridge between renown politicians and other Twitter users within the network as without them, the whole network may break into subgroups and interrupt any interaction between the groups. The application of betweenness centrality measure may reveal numerous understandings on the network especially when focusing on the middle-man impact on a communication link [24].

C. Clustering

Studies related to social media networks have also highlighted the significance of network clusters. To create clusters of politically based users, O'Connor [25] identified retweeting trends. Next, Himelboim [4] found that users had similar content and site URLs within a cluster that represented a common political mindset. Other than that, the blogosphere [26], Facebook [27], and preferences for books ordered on Amazon.com have also found a common trend of homophilous polarisation [28]. Besides, Choi and Park [29] found parallels in conversation topics within clusters.

Gomez-Rodriguez, Leskovec, and Krause [30] who studied data distribution between websites, blogs, and social media for the ongoing news events highlighted the role of network clusters in information flow. They discovered that in a couple of moments or days, clusters of news media sites and blogs frequently appear, generating an intensified number of information pathways in the network by linking individuals and the content. However, Twitter clusters are formed by individuals and they decide whom to respond to, mention, or retweet. Social media networks differ terms of the density of connections within them. Whereas, the concentrated groups of people who often share an opinion or ideology are found to be the key social characteristics which regulate the rate and type of information transmitted through social media.

According to Newman & Givran [31], network structure modularity is a measurement of the consistency of clustering, a network separated into a collection of interconnected substrings. Modularity captures the degree to which clusters are separated from each other (a set of values from 0 to 1), differentiated between fragmented and integrated networks. It also helps to differentiate between two very different kinds of extremely dense networks. Concerning their clustering, high-density networks may be distinct from one another. Networks may all have high overall network density scores of one or two (or more) compact clusters. Yet networks that lack substantial interconnection with two or more dense clusters are somewhat different from those that have a single large dense cluster. These segregated networks are extremely modular.

In comparison, a single dense group, a unified community, is a high-density network with low modularity. Networks where participants, regardless of cluster affiliation, when strongly interconnected are therefore classified as less modular. Between these network shapes, density alone does not distinguish. When paired with modularity, density can be differentiated as integrated and fragmented network patterns. As discussed earlier, these two separate network architectures are significantly dissimilar in terms of information flow. A few lightly interconnected and highly interconnected clusters characterize and reveal highly modular networks.

A clique on the other hand is defined as a small subset of a network whereby k-number of nodes that are close and intensely tied to each other are extracted from the network for analysis. In this Twitter network, it is common to form cliques based on different attributes such as age, gender, similar likes, hobbies, favourite books or food, etc. [32] . The closeness or connectedness between the nodes based on these similarities and identified number of nodes in a clique can be used to explore the network characteristics. Therefore, modularity analysis in this study will aid to separate the network in the form of clusters or groups according to their level of density (low-high). Besides, by performing the clique analysis, it enables us to identify the most central gang or circle of nodes that are closely tied to one another in the Najib Razak name-network.

III. MOTIVATION

This research emphasises only on Twitter users, that have mentioned Najib Razak or @najibrazak on their Twitter page or communication text. Around 1,000 rows of Twitter data were extracted on November 28th, 2020 via Netlytic and that made up to a name-network consisting of 1,004 nodes and 2,626 edges. The focus is given on determining who-follows-who and mentions @najibrazak besides identifying how the most influential nodes can affect the relationship between Twitter users and Najib Razak in the name-network.

Next is to recognize the path connections and the distance between Twitter user accounts and Najib Razak. SNA measures are used for this purpose and to further investigate the motive that influence the strength of @najibrazak node. It has also been set to identify the type of nodes or users that make up clusters, highly connected to Najib Razak and involving them on the latest political topics. Clusters in the networks with high modularity are stated to have dense connections within modules between the nodes but in separate components. In some scenarios, the structure of the network may cause sparse connections between nodes. In a nutshell, the Najib Razak name-network can be analyzed further by finding the network pattern using SNA techniques and fulfilling these motives. SNA also used in this research to construct, visualize, and analyze Twitter communication networks. The analytical questions to achieve the motives above are listed below:

- What are the relationships between Twitter user accounts with Najib Razak in the name-network?
- What influences the connectivity between Twitter user accounts and Najib Razak in the name-network?
- Which clusters and subgroups of Twitter users are highly connected to Najib Razak in the name-network?

To explore further, the data used and the method to extract the data are explained in the following section.

IV. DATASET USED

The name-network or 'who-mention-whom' dataset linked to Najib Razak was extracted for this study based on Twitter data referenced by Netlytic [33]. The Najib Razak Twitter Name-Network is a communication network that focuses on mining personal names in texts related to Najib Razak. In this research, Netlytic is used to extract the name-network of Najib Razak. It is an analyzer of cloud-based text and social networks that can summarize and discover contact networks from articles on social media that are publicly accessible. Public APIs are used by Netlytic to allow users to capture posts from social media sites or other public sources (such as Twitter, YouTube, RSS Feed, or text/CSV). SNA researchers may use Netlytic to discover common topics and to identify and explore emerging discussion topics.

After completing the initial needful information on the Netlytic page, the dataset then was extracted from Twitter and made available. The whole process required less time and steps for researchers to follow. The extracted dataset consists of Twitter user information such as the number of followers of @najibrazak mentions and the contents. The understanding of Najib Razak by Twitter users can be further analyzed from the dataset this research provides crucial insights into the overall network of relationships, including interconnectedness, centrality, clusters, and subgroups.

V. METHODOLOGY

A Gephi software is used to perform the SNA on the Najib Razak Twitter name-network dataset. Gephi is a leading software for visualization and exploration of all forms of graphs and networks. It is a guide for data analysts and scientists interested in investigating and interpreting graphs and networks as it helps to expose patterns and trends intuitively, highlight outliers, and tells stories with their data [34].

After obtaining the extracted dataset, the first step involved is opening the dataset in Gephi software to conduct initial exploratory analysis. The data table was checked for errors, besides multiple edges and loops were removed from the network before visualizing it in the form of a graph. It is also determined that no data cleaning process is required as the dataset is complete without any errors. Hence, it is ready to analyze the relationship and to gain an insight into different fields, including commonalities or variations that occur in the Najib Razak name-network.

The first step is to upload the dataset into Gephi and to visualize the network. Since it is a large network, the degree centrality of each node and the degree distribution of the network are calculated. The output is expected to classify famous, strong, prominent, and most influential individuals in the Najib Razak Twitter name-network dataset. Next, to understand the factors that influence the ties between Twitter users within the Najib Razak Twitter name-network dataset, further analyses using the network density and the betweenness centrality measure were conducted. Finally, the community detection analysis is done to extract information of clusters and subgroups that are highly connected to Najib Razak in the name-network. These SNA methods have not been applied to Najib Razak's name-network and the results obtained were compiled and discussed in the following section.

VI. RESULTS AND DISCUSSION

The novel findings related to the research questions are divided into three separate sections and discussed in detail.

A. Objective 1: Relationships between Twitter user accounts and Najib Razak in the Twitter name-network

The Twitter dataset that has been extracted via Netlytic is analyzed using Gephi software to understand the relationship between Twitter users and Najib Razak. Fig. 1 shows a large network formed and filtered by degree distribution to attain top 15 Twitter users with high degree. The size of the nodes in the graph represents the degree distribution where the highest the degree of the node, the larger the size. Thus, Najib Razak is found to be the node with highest degree in the network. Meanwhile, Fig. 2 depicts the line graph that denotes degree distribution from the lowest to the highest count of nodes based on the value of the degree. It is shown that, 328 nodes, that is the highest number of nodes represents those Twitter users with only 1 connection in this network.The degree centrality and distribution of the dataset are analyzed to reveal and assess the important nodes through the number of direct connections that one node has to other nodes. The premise of degree centrality is the primary indicator of value or power within the network that depends on the number of connections. The objective of this analysis is to measure and determine the top 10 most powerful and influential users or individuals in the Najib Razak's Twitter name-network (i.e., the "celebrities"). The degree centrality of the Najib Razak Twitter name-network reveals a long-tailed distribution that explains the network is skewed to the right whereby higher the value of connections between nodes, lesser the number of counts. Thus, it can be seen that there are only a few Twitter users who are prominent, have a high degree and identified as influential nodes in this network.

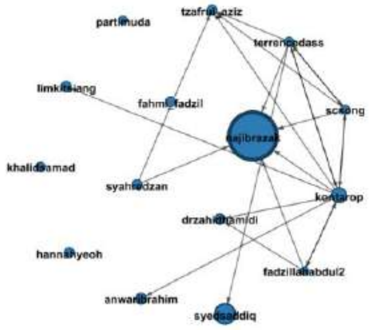


Fig. 1. Najib Razak Twitter name-network based on degree distribution

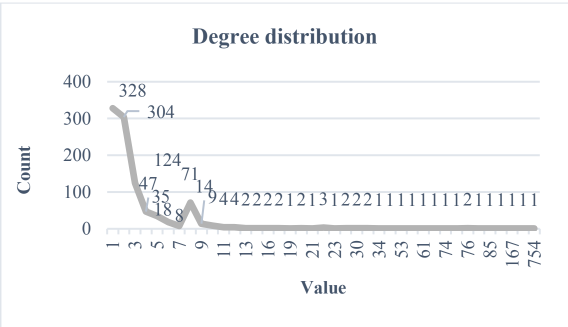


Fig. 2. Degree Distribution Analysis of Najib Razak's Twitter name-network

Based on the analysis, the top most prominent Twitter users is listed in Table I. The highest degree node is, @najibrazak himself with 754 links or connections, followed by @syedsaddiq with 263 connections, @anwaribrahim with 95 connections, and @drzahidhamidi with 85 connections. It can be noted that the top 10 most influential or powerful Twitter users in the network based on the degree distribution are all Malaysian politicians from both government and opposition parties.

TABLE I. TOP 10 TWITTER USERS' RANK BASED ON DEGREE CENTRALITY VALUE

No	ID	Label	Degree
1	n912	@najibrazak	754
2	n973	@syedsaddiq	263
3	n784	@anwaribrahim	95
4	n823	@drzahidhamidi	85
5	n870	@limkitsiang	79
6	n702	@syahredzan	76
7	n828	@fahmi_fadzil	76
8	n987	@tzafrul_aziz	75
9	n842	@hannahyeoh	74
10	n865	@khalidsamad	69

The same pattern is also observed in the network based on the in-degree measurement of the Najib Razak Twitter name-network as in the top 10 most influential or powerful users, all are noted to be Malaysian politicians. These users with high in-degree measure can be viewed as conversational hubs that occur in some scenarios such as the names that others have mentioned, replied to, or retweeted their posts concerning Najib Razak. Meanwhile, the number of out-degree, defined as connection flowing from a selected node to a range of other network members, is another significant measure for networks. A high level of out-degree relative to in-degree will often indicate that a specific node is not considered a direct source of information. Nevertheless, these nodes can provide pathways to several essential outlets, supporting others as an information aggregator or hub.

The top 10 most influential or powerful users in the dataset based on the out-degree measurement are also analyzed. These users can be interpreted as the Twitter users that have mentioned or initiated interaction with Najib Razak. In other words, high out-degree means that the user tweets many times about Najib Razak, trying to draw the interest of other users by referencing or reacting to them regarding the Najib Razak-related topics. Next, using R, an analysis of the Top 10 Twitter users by the ranking result of PageRank, Hubs, and Authorities were summarized. It is seen that Syed Saddiq, @syedsaddiq, holds the first place and becomes the most significant node with the highest PageRank score. He also has the largest hub and authority value. He is called the 'hub' or person that has connected to the largest number of nodes in the network. Other Twitter users ranked 2nd until the 10th are as listed in the Table II. Compared to the ranking based on degree centrality, the page rank, hubs and authorities scores show Najib Razak holds only the 7th place with 759 PageRank score, 36 hub, and 758 authorities. Different SNA measures may show unique results and rankings that lead to new findings.

It is noted that at the time of the research, Syed Saddiq has been closely linked to Najib Razak due to a recent meeting between both politicians. The meeting was said to take place to discuss Budget 2021, prior to the voting of the budget in the parliament. [35-37]. This would also explain why Tengku Zafrul Aziz, @tzafrul_aziz, the Minister of Finance ranked 2nd on influence and importance, followed by other local politicians, as Twitter users are interested to know what would be the outcome from voting of the Budget. Najib Razak has also been noted to have been vocal in his statements regarding the budget and have been noted to question whether the budget would be able to support Malaysia's economy.

TABLE II. TOP 10 TWITTER USERS' RANK BASED ON PAGERANK, HUBS, AND AUTHORITIES

No	Label	PageRank	Hubs	Authorities
1	@syedsaddiq	837	75	834
2	@tzafrul_aziz	834	66	763
3	@limkitsiang	763	41	762
4	@drzahidhamidi	762	40	761
5	@fahmi_fadzil	761	39	760
6	@anwaribrahim	760	37	759
7	@najibrazak	759	36	758
8	@hannahyeoh	758	31	757
9	@khalidsamad	757	8	756

10	@syahredzan	472	2	472

Continuing the exploration of the user ranking, deeper analysis and explanation have been done to reveal the connectivity between them in the following section.

B. Objective 2: Connectivity between Twitter users within Najib Razak name-network

Before analyzing what influences the connectivity between Twitter users within the Najib Razak name-network, the density of the network is considered as it will help in understanding how connected the components of the network are relative to how they could be linked in reality. Other than that, the average distance between the pairs of nodes can also be identified further.

Network density defines the component of a network's possible links that are real connections. A "potential connection" is a connection between two nodes that may theoretically occur, whether or not it does. On the opposite, an "actual connection" exists. In Gephi, density measures how close the network is to complete. Network density is represented by the number between 0 and 1, where a complete network has all possible edges and density equals to 1. The network density of the Najib Razak Twitter name-network is 0.003, a low value, suggesting that the Twitter users that converse and discuss “Najib Razak” on Twitter are mostly individuals that do not know each other. This is an interesting finding where prominent people tend to influence the connectivity of unknown individuals with each other in the Najib Razak Twitter name-network.

Besides degree centrality, the betweenness centrality measures the extent to which a node plays the associating role or acts as a bridge between two or more communities that otherwise would not be able to communicate with each other. The more Twitter users depend on a middle node to make relations or connections with other individuals, the greater the centrality of that node becomes. The betweenness centrality is based on the premise that if a user presides over a communication bottleneck, the person may gain power. Table III shows the top 10 Twitter users based on the betweenness centrality ranking.

TABLE III. TOP 10 TWITTER USERS’ RANK BASED ON BETWEENNESS CENTRALITY VALUE

No	ID	Label	Betweenness Centrality
1	n10	@kontarop	5538.67
2	n13	@terrencedass	4373.25
3	n136	@alphaque	2818.00
4	n37	@thanussha5	2348.50
5	n31	@dean24924388	2075.00
6	n30	@ayahanda_md	1486.17
7	n2	@rameshraoaks	607.33
8	n40	@le_coco19	415.00
9	n34	@shreddedvincent	371.67
10	n45	@pokdi10	216.00

These Twitter users fall on the shortest path between other pairs of users in the network. Users with a high betweenness centrality follow the shortest paths to reach other users within the network and because of their position within the network, they have considerable control over information diffusion. These users are important in disseminating information over a network. According to [38], users with high betweenness are often nominated as leaders because without them, the whole network may be separated into componenents on its own without any associations.

Fig. 3 represents the top 2 Twitter users that emerged based on the highest betweenness centrality value . Users with high betweenness centrality often connect accounts found in different groups (i.e., network clusters), that supports the information flow across groups. Hence, in this research, information about these top 2 Twitter users will be further studied and examined to recognize who they are and what can be learned to understand how they influence the connectivity of Twitter users within the Najib Razak Twitter name-network.

Fig. 3. Networks showing the top two Twitter users with the highest betweenness centrality. Green nodes represent the two Twitter users: @kontarOP and @terrencedass

The user with the highest betweenness centrality score, @kontarOP has been noted to have a total of 2,744 followers and is following 2,912 other Twitter users. The user joined Twitter in February 2012. @kontarOP’s Twitter account content or posts are also studied to examine the interaction related to Najib Razak’s topics. This would support the understanding of the key role and influence of that user where his central node position between users in a group plays a great role in reaching out to other users that would not otherwise interact directly regarding Najib Razak.

When @kontarOP’s posts on Twitter that are related to the topic of Najib Razak were studied, it can be observed that @kontarOP regularly interacts through the regular posts’ interval, and managed to act as the bridge between other users to communicate with each other regarding Najib Razak. Meanwhile, @terrencedass, the user with the second highest betweenness centrality score, have a total of 3,360 followers and is following 1,044 other Twitter users where his Twitter account usage started in March 2012. In contrast to @kontarop’s regular posts or tweets on Twitter regarding Najib Razak, @terrencedass has been identified to communicate and interact with other Twitter users regularly while mentioning Najib Razak and prompting discussions on the related topic with other Twitter users. His regular tweet has made him act as a bridge between other Twitter users by flowing all kinds of attractive new information.

Another SNA metric to analyze the factor that influences the connectivity between Twitter users within Najib Razak Twitter name-network is eigenvector centrality. Messages can spread broadly if they are retweeted, or passed along by a few influential users. As such, being followed by one popular Twitter user bestows more influence than being followed by many brand-new Twitter users with few followers. Eigenvector centrality accounts for this by going beyond the number of followers a user has. It measures the collective influence of each follower. Being recognized by someone seen as powerful contributes heavily to one's perceived influence. Eigenvector centrality elevates those users followed by a smaller group, but more influential, number of followers.

In other words, eigenvector centrality is a more sophisticated view of centrality, i.e., a person with few connections could have a very high eigenvector centrality if those few connections were themselves very well connected, or in short, a connection to a popular individual is more important than a connection to a loner. Table IV lists the top 10 Twitter users based on the eigenvector centrality ranking. The networks of the top 2 Twitter users with high eigenvector centrality value is analyzed. Twitter user, @najibrazak is omitted from the discussion as he has posted some conversation or protective arguments on himself. Thus, the information on @syedsaddiq will be further studied and examined to understand how this user is influencing the connectivity between other Twitter users within the Najib Razak Twitter name-network.

TABLE IV. TOP 10 TWITTER USERS' RANK BASED ON EIGENVECTOR CENTRALITY RANKING

No	ID	Label	Eigenvector centrality
1	n912	@najibrazak	1
2	n973	@syedsaddiq	0.345698
3	n784	@anwaribrahim	0.125062
4	n987	@tzafrul_aziz	0.115087
5	n823	@drzahidhamidi	0.113649
6	n870	@limkitsiang	0.104074
7	n828	@fahmi_fadzil	0.095658
8	n702	@syahredzan	0.093155
9	n842	@hannahyeoh	0.093124
10	n865	@khalidsamad	0.086526

Syed Saddiq bin Syed Abdul Rahman or his Twitter handle, @syedsaddiq, has a total of 1 Million followers and is following 1,147 other Twitter users based on Fig. 4. He is noted to have joined Twitter in January 2012. As a Malaysian politician, he served under the former Prime Minister Tun Mahathir bin Mohamad as Minister of Youth and Sports in the Pakatan Harapan (PH) administration from July 2018 until the fall of the PH administration in February 2020. Naturally, the eigenvector centrality would be high since he is a public figure that would attract numerous popular Twitter accounts, which will precede to more influences when discussing about Najib Razak.

Syed Saddiq has also been embroiled in a scandal involving Najib Razak after posting a picture of him with the ex-Prime Minister in November 2020. The post as per Fig. 5 has garnered nationwide attention and uproar; this may have also increased and influenced the connectivity between other Twitter users with @syedsaddiq as he was passing more information about his acquaintance with Najib Razak.

Fig. 4. @syedsaddiq Twitter profile

Fig. 5. @syedsaddiq's post mentioning Najib Razak on Twitter

Hence, it can be concluded that the factors that would influence the connectivity between other Twitter users and Najib Razak are the frequency of their interactions with Najib Razak or regarding the topic of Najib Razak, as noted through the number of tweets sent out by the top two users with the highest betweenness centrality, @kontarop and @terrencedass in relation to Najib Razak. The regular interactions would spark conversations or the interests of other Twitter users, hence increasing the engagement with Najib Razak on Twitter. Meanwhile, Syed Saddiq has increased his own followers by sharing his latest agenda with the prominent person, Najib Razak and at at the same time retweeting about this by the other followers made the others important as well. This shows wide spread connectivity of the Najib Razak's name-network.

C. Objective 3: Clusters and subgroups of Twitter users which are highly connected to Najib Razak

Modularity is designed to quantify the power of a network's division in the form of clusters or groups where networks with high modularity have dense connections within modules between the nodes, but the separate modules has sparse connections [39]. In this analysis, to study the clustering of Twitter users related to Najib Razak Twitter name-network, the modularity class has been analyzed and it can be observed that 14 clusters are predominantly connected to Najib Razak as per the modularity class result in Fig.6. The average clustering coefficient of the data set is 0.125.

In measuring modularity, the Louvain method [40] is used. One thing that is highlighted is that high modularity does not indicate a "good" partition. In terms of computing time, the Louvain Method outperforms other methods such as Clauset, Newman, and Moore, Pons and Latapy, and Wakita and Tsurumi [41]. It enables us to evaluate networks of unparalleled scale at a more vast and large scope, with more attributes or factor that can be taken into consideration. Hence, by focusing on ad-hoc networks with a defined group structure, the Louvain approach has shown to be very effective. This is because of its hierarchical structure, which is indicative of the renormalization method, making it possible to consider various resolutions in clusters. Table V shows the modularity class filtered and presents the top 5 densed cluster. The filtering process involved in Gephi is partitioning the nodes in the network by modularity value and then dragging the filters option by ego network, setting where the node selected is Najib Razak.

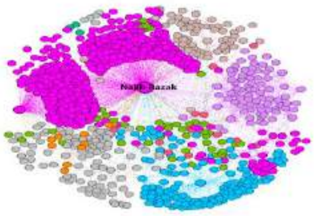
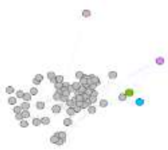


Fig. 6. Clustering Network of Najib Razak Twitter name-network

TABLE V. MODULARITY CLASS OF NAJIB RAZAK TWITTER NAME-NETWORK

No	Modularity Class	Cluster	Modularity %
1	12		49
2	4		16.33
3	10		7.37
4	11		6.77
5	13		6.77

The top five densed cluster is further analysed one by one to identify the position of important users. For this analysis, the SNA metrics such as the eccentricity, clustering, and eigenvector centrality are used. Ecentricity is the maximum of all the distances of a particular node from every other node in the graph. When the maximum distance is smaller, the node is considered to be closer to every other node and it is more central. The top 5 most influential Twitter users in their respective modularity class based on the SNA metrics mentioned above and its results are shown in Table VI – Table X.

TABLE VI. TOP FIVE INFLUENTIAL TWITTER USERS FOR MODULARITY CLASS 12 – CLUSTER 1

ID	**Label**	**Eccentricity**	**Modularity class**	**Clustering**	**Eigenvector centrality**
n912	@najibrazak	1	12	0.000548	1
n973	@syedsaddiq	1	12	0.001088	0.345698
n33	@shamsul__kamal	2	12	0.160714	0.007217
n395	@esshimself	1	12	0.166667	0.006439
n178	@lohemithren	1	12	0.25	0.001422

TABLE VII. TOP FIVE INFLUENTIAL TWITTER USERS FOR MODULARITY CLASS 4 – CLUSTER 2

ID	**Label**	**Eccentricity**	**Modularity class**	**Clustering**	**Eigencentrality**
n30	@ayahanda_md	5	4	0.111828	0.020256
n13	@terrencedass	5	4	0.038312	0.014406
n10	@kontarop	4	4	0.008483	0.008171
n12	@scsong	5	4	0.035989	0.004861
n9	@fadzillahabdul2	4	4	0.022449	0.002077

TABLE VIII. TOP FIVE INFLUENTIAL TWITTER USERS FOR MODULARITY CLASS 10 – CLUSTER 3

ID	**Label**	**Eccentricity**	**Modularity class**	**Clustering**	**Eigencentrality**
n784	@anwaribrahim	1	10	0.002128	0.125062
n823	@drzahidhamidi	1	10	0.002941	0.113649
n870	@limkitsiang	1	10	0.001947	0.104074
n828	@fahmi_fadzil	1	10	0.000175	0.095658
n702	@syahredzan	1	10	0.013684	0.093155

TABLE IX. TOP FIVE INFLUENTIAL TWITTER USERS FOR MODULARITY CLASS 11 – CLUSTER 4

ID	**Label**	**Eccentricity**	**Modularity class**	**Clustering**	**Eigencentrality**
n929	@partimuda	1	11	0.006531	0.071519
n990	@unimalaya	1	11	0	0.025363
n37	@thanussha5	7	11	0.142857	0.01853
n136	@alphaque	6	11	0.185897	0.009442

n45	@pokdi 10	8	11	0.03508 8	0.001252

TABLE X. TOP FIVE INFLUENTIAL TWITTER USERS FOR MODULARITY CLASS 13 – CLUSTER 5

ID	Label	Eccentri city	Modula rity class	Cluster ing	Eigencentr ality
n987	@tzafrul_ aziz	0	13	0.0057 66	0.115087
n868	@kwspm alaysia	0	13	0.0064 94	0.030058
n795	@barisan asional	0	13	0.0071 43	0.027982
n681	@simtzet zin	1	13	0.1078 43	0.020726
n85	@ieka93	1	13	0.0277 78	0

From the results, it is found that many of the users, noted (e.g., Cluster 1 – @syedsaddiq; Cluster 2 - @ayahanda_md, @terrencedass, @kontarOP; Cluster 3 - @anwaribrahim, @drzahidhamidi, @limkitsiang, @fahmi_fadzil; Cluster 5 - @tzafrul_aziz) and also various political parties (e.g., Cluster 4 - @partimuda, Cluster 5 - @barisanasional) communicating topically regarding Najib Razak via Twitter. These are the nodes with highest eigenvector centrality measure. This indicates that these nodes are connected to a node with high degree centrality and at the same time they themselves are positioned as a central node connecting to many links. These are the Twitter users whom are highly connected to Najib Razak. This can be used as an evidence to scrutinize further their acquaintance with Najib Razak and his activities.

Upon further investigations related to finding the sub-groups of Twitter users, Table XI shows the top 5 largest cliques with the same size obtained from the network. It is determined that Syed Saddiq and Najib Razak appear the most in these large cliques as the clique starts from them. There are also other users such as @alwan_kaker, @Rameshraoaks, @Mahazalimreturn, and @ayahanda_md who repeatedly appear in the list of the cliques. According to [42], nodes that belong to all the cliques are said to be the most central nodes that connect with other users in the network. Other than that, the size of the cliques with five members also shows an excellent cohesion in linkages between nodes.

TABLE XI. TOP FIVE LARGEST CLIQUES

No	Cliques
1	@syedsaddiq - @terrencedass - @alphaque- @sharaadkuttan - @iamjoelee
2	@najibrazak - @alwan_kaker - @rameshraoaks- @mahazalimreturn - @ayahanda_md
3	@najibrazak - @alwan_kaker - @rameshraoaks - @mahazalimreturn - @razmanrahim
4	@najibrazak - @ayahanda_md - @hamdanrukun - @kontarop - @dean24924388
5	@najibrazak - @ayahanda_md - @rameshraoaks @mahazalimreturn - @_c0mm4nd3r

VII. CONCLUSION

This section concludes all the findings obtained in this research and discusses the limitations and recommendation for future study. From the results that have been presented, the research questions and related objectives have been satisfied. It can be proven that based on the degree, in-degree and out-degree results that have been presented, the Twitter users who are trying to direct their ties to prominent nodes and initiating connections by mentioning Najib Razak's name on Twitter and discussing him can be considered as a measure of significance. These users with high in-degree measurement can be viewed as conversational hubs, as others have mentioned, replied to, or retweeted their posts concerning Najib Razak.

Meanwhile, to understand the factors that influence the connectivity between Twitter users within Najib Razak name-network, it has been identified that the connectivity of Twitter users to the topic on Najib Razak is determined using betweenness centrality value. Users with a high betweenness centrality provide the shortest paths between other users within the network and because of their position within the network, they have considerable control over information diffusion. From this analysis, it also can be concluded that the number of tweets and topical mentions are the factors that increase the betweenness centrality. Hence, the increasing connectivity to Najib Razak in the name-network. Another metrics that was used to analyze the influence of the connectivity is eigenvector centrality. Messages can spread broadly if retweeted, or passed along, by a few influential users. Hence, the factors of the connectivity can be significantly identified by using these metrics and to discover the relationship of the top 10 most influential users.

Lastly, identifying the clusters and subgroups of Twitter users is achieved by applying the modularity class to the network, and this analysis managed to identify 14 different clusters that have an association with Najib Razak on Twitter. Other than that, from the 14 clusters, the top 5 most significant cliques were obtained which allows us to detect the most central nodes that connected well in the network with other Twitter users. One of the limitations that has been identified in this paper is that due to the unavailability of few attributes such as ID, following, followers, and several mentions in the dataset. The analysis and possibilities to explore other suitable metrics that can be instilled in this study are limited. In the future, it is recommended to include exploration on another network in the social media platform of Najib Razak such as Facebook and Instagram accounts, and comparison of the findings among the 3 platforms are observed and studied. Besides, more attributes such as user tweets, timestamp, and the number of likes should be extracted to explore the network and to strengthened further the analysis results. Following that, various strategies performed by the politicians or specific individuals with special interest to embark on the social media platform relating to political matters can be further analyzed.

ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Higher Education, Government of Malaysia and Research Management Centre, Universiti Teknologi Malaysia for supporting this work through the FRGS Grant, vote number R.K130000.7856.5F016 and Registration Proposal No: FRGS/1/2018/ICT04/UTM/02/8 and UTM.

References

[1] G. Enli, "Twitter as arena for the authentic outsider: Exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election," *European journal of communication,* vol. 32, no. 1, pp. 50-61, 2017.

[2] Socialbakers. "Most popular Twitter accounts in Malaysia." https://www.socialbakers.com/statistics/twitter/profiles/malaysia (accessed January, 2021).

[3] J. C. Liow, "Malaysia's March 2008 general election: understanding the new media factor," *The Pacific Review,* vol. 25, no. 3, pp. 293-315, 2012/07/01 2012, doi: 10.1080/09512748.2012.685096.

[4] I. Himelboim, M. A. Smith, L. Rainie, B. Shneiderman, and C. Espina, "Classifying Twitter topic-networks using social network analysis," *Social media+ society,* vol. 3, no. 1, p. 2056305117691545, 2017.

[5] M. Castells, "Communication, power and counter-power in the network society," *International journal of communication,* vol. 1, no. 1, p. 29, 2007.

[6] R. D. Waters and J. M. Williams, "Squawking, tweeting, cooing, and hooting: Analyzing the communication patterns of government agencies on Twitter," *Journal of public affairs,* vol. 11, no. 4, pp. 353-363, 2011.

[7] M. F. Kasmani, "A political discourse analysis of the twitter posts of@ najibrazak prior to 2018 general elections," *SEARCH (Malaysia),* vol. 11, no. 2, pp. 129-143, 2019.

[8] S. Matei, "Analyzing Social Media Networks with NodeXL: Insights from a Connected World by Derek Hansen, Ben Shneiderman, and Marc A. Smith: Burlington, MA: Morgan Kauffman, 2011. 284 pages. ISBN: 978-0-12-382229-1," ed: Taylor & Francis, 2011.

[9] J.-S. Lin and I. Himelboim, "Political brand communities as social network clusters: winning and trailing candidates in the GOP 2016 primary elections," *Journal of Political Marketing,* vol. 18, no. 1-2, pp. 119-147, 2019.

[10] O. Zorzi, "Granovetter (1983): The strength of weak ties: A network theory revisited," in *Schlüsselwerke der Netzwerkforschung*: Springer, 2019, pp. 243-246.

[11] K. Carley, "A theory of group stability," *American sociological review,* pp. 331-354, 1991.

[12] M. S. Granovetter, "The strength of weak ties," *American journal of sociology,* vol. 78, no. 6, pp. 1360-1380, 1973.

[13] R. S. Burt, *Brokerage and closure: An introduction to social capital*. Oxford university press, 2005.

[14] G. Delanty, "The foundations of social theory," *Social Theory,* vol. 19, 2009.

[15] J. S. Coleman, *Foundations of social theory*. Harvard university press, 1994.

[16] P. P. Zubcsek, I. Chowdhury, and Z. Katona, "Information communities: the network structure of communication," *Social Networks,* vol. 38, pp. 50-62, 2014.

[17] K. Lerman and R. Ghosh, "Information contagion: An empirical study of the spread of news on digg and twitter social networks," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2010, vol. 4, no. 1.

[18] S. Wasserman and K. Faust, "Social network analysis: Methods and applications," 1994.

[19] T. W. Valente, *Social networks and health: Models, methods, and applications*. Oxford University Press New York, 2010.

[20] T. D. Huynh, "Extension of PageRank and application to social networks. (Extension de PageRank et application aux réseaux sociaux)," 2015. [Online]. Available: https://tel.archives-ouvertes.fr/tel-01187929

[21] H. Kretschmer and T. Kretschmer, "A new centrality measure for social network analysis applicable to bibliometric and webometric data," *Collnet Journal of Scientometrics and Information Management,* vol. 1, no. 1, pp. 1-7, 2007.

[22] B. Ruhnau, "Eigenvector-centrality—a node-centrality?," *Social networks,* vol. 22, no. 4, pp. 357-365, 2000.

[23] P. Bonacich and P. Lloyd, "Eigenvector-like measures of centrality for asymmetric relations," *Social networks,* vol. 23, no. 3, pp. 191-201, 2001.

[24] P. Magalingam, S. Davis, and A. Rao, "Using shortest path to discover criminal community," *Digital Investigation,* vol. 15, pp. 1-17, 2015/12/01/ 2015, doi: https://doi.org/10.1016/j.diin.2015.08.002.

[25] B. T. O'Connor, M. Krieger, and D. Ahn, "TweetMotif: Exploratory Search and Topic Summarization for Twitter," in *ICWSM*, 2010.

[26] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 U.S. election: divided they blog," presented at the Proceedings of the 3rd international workshop on Link discovery, Chicago, Illinois, 2005. [Online]. Available: https://doi.org/10.1145/1134271.1134277.

[27] S. P. Robertson, R. K. Vatrapu, and R. Medina, "The social life of social networks: Facebook linkage patterns in the 2008 US presidential election," in *Proceedings of the 10th annual international conference on digital government research: Social networks: Making connections between citizens, data and government*, 2009: Citeseer, pp. 6-15.

[28] V. Krebs, "The social life of books: Visualizing communities of interest via purchase patterns on the WWW. Orgnet. com," ed, 2004.

[29] S. Choi and H. W. Park, "An exploratory approach to a Twitter-based community centered on a political goal in South Korea: Who organized it, what they shared, and how they acted," *New Media & Society,* vol. 16, no. 1, pp. 129-148, 2014.

[30] M. Gomez-Rodriguez, J. Leskovec, and A. Krause, "Inferring networks of diffusion and influence," *ACM Transactions on Knowledge Discovery from Data (TKDD),* vol. 5, no. 4, pp. 1-37, 2012.

[31] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical review E,* vol. 69, no. 2, p. 026113, 2004.

[32] M. Jamali and H. Abolhassani, "Different aspects of social network analysis," in *2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings)(WI'06)*, 2006: IEEE, pp. 66-72.

[33] "Netlytic - Social Media Text and Social Networks Analyzer." https://netlytic.org/ (accessed 25 November 2020).

[34] "Gephi - The Open Graph Viz Platform." https://gephi.org/ (accessed 27 November, 2020).

[35] T. Velusamy. "“He's a convicted criminal!” Parti MUDA Had A Meeting With Najib & Malaysians Are NOT Happy'." World of Buzz. https://worldofbuzz.com/hes-a-convicted-criminal-parti-muda-had-a-meeting-with-najib-malaysians-are-not-happy/. (accessed 10th December, 2020).

[36] A. Mohanakrishnan. "Najib meeting on windfall tax, says Muda co-founder." The Vibes Dotcom. https://www.thevibes.com/articles/news/7329/najib-meeting-on-windfall-tax-says-muda-co-founder (accessed 10th December, 2020).

[37] FMTReporters. "'Najib questions whether 2021 budget will prevent economic downturn'." https://www.freemalaysiatoday.com/category/nation/2020/11/22/najib-questions-if-2021-budget-will-prevent-economic-downturn/ (accessed 10th December, 2020).

[38] L. C. Freeman, D. Roeder, and R. R. Mulholland, "Centrality in social networks: II. Experimental results," *Social networks,* vol. 2, no. 2, pp. 119-141, 1979.

[39] A. Fornito, A. Zalesky, and E. Bullmore, *Fundamentals of brain network analysis*. Academic Press, 2016.

[40] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti, "Generalized louvain method for community detection in large networks," in *2011 11th international conference on intelligent systems design and applications*, 2011: IEEE, pp. 88-93.

[41] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment,* vol. 2008, no. 10, p. P10008, 2008.

[42] C. Reffay and T. Chanier, "How Social Network Analysis can help to Measure Cohesion in Collaborative Distance-Learning," in *CSCL*, 2003.