

Penerapan Algoritma Katz Pada Struktur Komunitas Jejaring Sosial *Twitter*

Application of the Katz Algorithm to the Structure of the Twitter Social Networking Community

Diah Novita Sari^{1*}, Kurniati²

^{1,2}Manajemen Informatika, Politeknik Negeri Sriwijaya

*diah.novita.sari@polsri.ac.id

Abstrak

Social Network Analysis (SNA) merupakan sebuah studi yang mempelajari tentang hubungan manusia dengan memanfaatkan teori graf. SNA dapat menemukan beberapa komunitas dan besar pengaruh sentralitas pada komunitas dalam jejaring sosial. Namun seberapa besar ukuran dari jaringan komunitas, kepadatan jaringan, dan resiprositas dari jaringan komunitas serta berapa banyak jumlah akun sebagai *important user* pada suatu komunitas dalam jaringan sosial sampai saat ini belum pernah dilakukan. Penelitian ini bertujuan untuk mengidentifikasi ukuran dari jaringan komunitas, kepadatan jaringan, dan resiprositas dari jaringan komunitas dan mengidentifikasi jumlah akun sebagai *important user* pada suatu komunitas dalam jaringan sosial. Sampel yang digunakan adalah pengguna *twitter* yang melakukan *posting* atau melakukan *mention* dengan kata kunci “Jokowi”. Penarikan data sampel menggunakan teknik *systematic random sampling*. Analisis yang dilakukan menggunakan metode algoritma Katz. Berdasarkan hasil penelitian, bahasa pemrograman *python* dan algoritma katz berhasil menemukan ukuran dari jaringan komunitas dengan kata kunci “jokowi” sejumlah 4052 *nodes*. Kepadatan jaringan rendah yaitu 0,0002 jika dibulatkan menjadi 0 (nol). Resiprositas pada penelitian ini yaitu berlangsung dua arah. Diameter jaringan pada penelitian ini adalah 6 dan jarak / *Average Path length* yaitu 2.23 dan terdapat 12 akun sebagai *important users*.

Kata kunci: Social Network Analysis, Algoritma Katz, Twitter.

Abstract

Social Network Analysis (SNA) is a study that studies human relations by utilizing graph theory. SNA can find several communities and the degree of centrality influences the social networks communities. But how big is the size of the community network, network density, and reciprocity of the community network and how many number of accounts as *important users* in a community in a social network has never been done. This study aims to identify the size of the community network, network density, and reciprocity of the community network and identify the number of accounts as *important users* in a community in a social network. The sample used is *Twitter* users who post or mention the keyword “Jokowi”. Withdrawal of sample data using *systematic random sampling* technique. The analysis was performed using the Katz algorithm method. Based on results of the researches, the *Python* programming language and the Katz algorithm has succeeded in finding the size of the community network. The size of the community network with the keyword “jokowi” is 4052 *nodes*. Low network density, which is 0.0002 when rounded up to 0 (zero). Reciprocity in this research is going in two directions. The network diameter in this study is 6 and the average path length is 2.23 and there are 12 accounts as *important users*.

Keywords: Social Network Analysis, Katz Algorithm, Twitter.

PENDAHULUAN

Social Network Analysis (SNA) merupakan sebuah studi yang mempelajari tentang hubungan manusia dengan memanfaatkan teori graf. Dengan pemanfaatan teori graf ini membuat SNA mampu memeriksa struktur dari hubungan sosial dalam suatu kelompok untuk mengungkap hubungan informal antar individu.

Twitter merupakan satu inovasi fenomenal dalam teknologi komunikasi internet abad ke-21. Twitter dikenal sebagai penyedia media sosial yang tidak hanya dikategorikan sebagai media sosial tetapi juga termasuk microblogging. Di Twitter, pengguna diberi kebebasan menggunakan fitur yang ada di dalamnya dengan syarat mematuhi beberapa kebijakan yang disediakan oleh penyedia layanan [1]. Twitter dikenal sebagai layanan microblogging yang memudahkan penggunaannya hanya sekedar membaca, menulis, ataupun posting pesan di forum publik [2]. Secara historis kehadiran dan kemunculan media sosial Twitter yang menawarkan ruang tertentu atau maksimal 140 karakter [3].

Dalam jaringan pertemanan *Twitter* terdapat berbagai macam komunitas virtual. Media sosial, khususnya Twitter, menawarkan kesempatan untuk berinteraksi lebih dekat dengan penggunaannya dan dapat menjadi sarana untuk menciptakan jaringan komunitas [4]. Komunikasi berjejaring ini pada akhirnya membentuk berbagai komunitas virtual [5]. Komunitas virtual merupakan kumpulan pengguna yang memiliki kesamaan dan terbentuk melalui ruang siber serta relasi yang terjadi diantara mereka termediasi secara elektronik [6]. Sehingga data jaringan pertemanan di *twitter* yang didapat bisa membentuk suatu graph jaringan pertemanan yang akan dianalisis struktur komunitasnya. Sentralitas dalam komunitas virtual bisa dimanfaatkan untuk kepentingan marketing, informasi kriminal, dan lain sebagainya.

Social Network Analysis (SNA) untuk menemukan beberapa komunitas dan besar pengaruh sentralitas pada komunitas dalam jejaring sosial. SNA adalah salah satu metode analisis dalam sebuah penelitian yang memiliki konsentrasi pada relationship research dan

sering digunakan dalam mengukur sebuah hubungan dan menggambarkan beberapa informasi secara individu [7]. Selain itu, SNA merupakan metode yang digunakan untuk memetakan dan menggambarkan struktur jaringan di dalam percakapan media sosial [8]. SNA juga dapat digunakan dalam operasi media sosial, terutama ketika mempertimbangkan pemodelan interaksi yang digunakan untuk menganalisis jaringan hubungan interaksi media sosial [9]. SNA memiliki beberapa properti jaringan untuk memetakan hubungan yang sangat membantu untuk meningkatkan penciptaan *management knowledge* [10].

Namun seberapa besar ukuran dari jaringan komunitas, kepadatan jaringan, dan resiprositas dari jaringan komunitas serta berapa banyak jumlah akun sebagai important user pada suatu komunitas dalam jaringan sosial sampai saat ini belum pernah dilakukan. Dengan menggunakan algoritma Katz penulis akan dapat menganalisa seperti apa struktur komunitas jejaring sosial khususnya *twitter*.

METODE PENELITIAN

Tahap perencanaan merupakan tahap awal pada penelitian ini. Pada tahap ini beberapa persiapan mulai dilakukan seperti *clustering data*, menentukan jenis komunitas apa saja yang akan diteliti dan metode apa yang akan digunakan dalam penelitian ini. Sampel yang akan digunakan pada penelitian ini adalah pengguna *twitter* yang melakukan *posting* atau melakukan *mention* dengan kata kunci "Jokowi". Batasan maksimal *tweet* yang ditarik yaitu 6000 *tweet* dan 100 batasan *tweet* per kata.

Teknik pengumpulan data yang digunakan adalah data primer dan data sekunder. Data primer pada penelitian ini berupa dokumen elektronik dari jaringan pertemanan pada media sosial *twitter* dengan kata kunci "Jokowi". Penarikan data sampel menggunakan teknik *systematic random sampling*. Data ditarik menggunakan *software* aplikasi Python dan ditarik hanya satu kali yaitu ketika mengumpulkan data dengan kata kunci "Jokowi". Sedangkan data sekunder pada penelitian ini yaitu berupa beberapa jurnal dan buku yang berkaitan dengan penelitian.

Selanjutnya data yang telah terkumpul dilakukan analisis dengan teknik membaca, mendeskripsikan, menganalisis, meringkas pada sumber utama dan pendukung.

Guna mendapatkan jawaban dari permasalahan tersebut, untuk menganalisis data dalam penelitian ini penulis menggunakan metode sebagai berikut:

- a. Deskripsi, merupakan pemaparan data yang diperoleh, dalam praktiknya tidak terbatas pada pengumpulan data saja, tetapi juga meliputi penjelasan dan analisis terhadap data tersebut.
- b. Analisis, merupakan metode untuk menganalisis struktur komunitas dari data jaringan pertemanan *twitter* yang diperoleh. Analisis yang dilakukan dalam penelitian ini menggunakan metode algoritma Katz. Sentralitas Katz dapat dinyatakan sebagai persamaan di bawah ini:

$$K_i = |(\sum_{k=0}^{\infty} \alpha^k A^k)|_i \dots\dots\dots (1)$$

Dimana K_i adalah sentralitas Katz dari node i , A^k mewakili kekuatan kedekatan maktrik ke- k , α adalah faktor atenuasi dan I adalah kolom vektor dimana semua nilai adalah 1.

Parameter yang digunakan untuk menganalisis struktur komunitas pada jaringan pertemanan *twitter* adalah sebagai berikut:

1. Ukuran (Size), ukuran berkaitan dengan jumlah anggota dari jaringan. Ukuran dari suatu jaringan menentukan karakteristik suatu jaringan. Jaringan dengan ukuran kecil, antar-aktor (*node*) lebih kohesif dibandingkan dengan jaringan ukuran besar. Struktur relasi di antara aktor juga berbeda antara jaringan dengan ukuran kecil dan besar
2. Kepadatan (Densitas/*Density*), merupakan perbandingan jumlah *link (ties)* yang ada dalam jaringan dengan jumlah *link* yang mungkin muncul. Kepadatan memperlihatkan intensitas antar anggota jaringan dalam berkomunikasi. Jaringan dengan kepadatan tinggi adalah jaringan dimana anggotanya saling berinteraksi satu sama

lain. Sebaliknya, jaringan dengan kepadatan rendah ditandai oleh minimnya interaksi antar-anggota jaringan. Interaksi hanya didominasi oleh aktor tertentu saja, tidak merata kesemua anggota (aktor). Rumus menentukan kepadatan (densitas) sebagai berikut:

$$D = \frac{l}{N(N-1)} \dots\dots\dots (2)$$

Dimana D adalah kepadatan (densitas), dan l adalah jumlah *link (ties)* aktual dalam jaringan, N adalah ukuran jaringan (jumlah aktor dalam jaringan). Angka dalam kepadatan (densitas) ini yaitu 0 hingga 1, dimana makin besar nilai menunjukkan makin tingginya kepadatan (densitas) dari suatu jaringan.

3. Resiprositas (Mutualitas/ *Reciprocity*), merupakan rasio dari *link (ties)* dua arah dengan total jumlah *link* dalam jaringan. Ukuran ini menggambarkan apakah relasi anggota jaringan (aktor/node) berlangsung dua arah ataukah searah. Rumus untuk menentukan resiprositas sebagai berikut:

$$R = \frac{(A_{ij}=1) \text{ dan } (A_{ji}=1)}{(A_{ij}=1) \text{ atau } (A_{ji}=1)} \dots\dots\dots (3)$$

Dimana R adalah resiprositas, A_{ij} adalah *link* dari aktor satu ke aktor lainnya. Angka resiprositas dari 0 hingga 1, dimana 1 memperlihatkan resiprositas sempurna, dimana masing-masing aktor saling memilih atau saling berinteraksi.

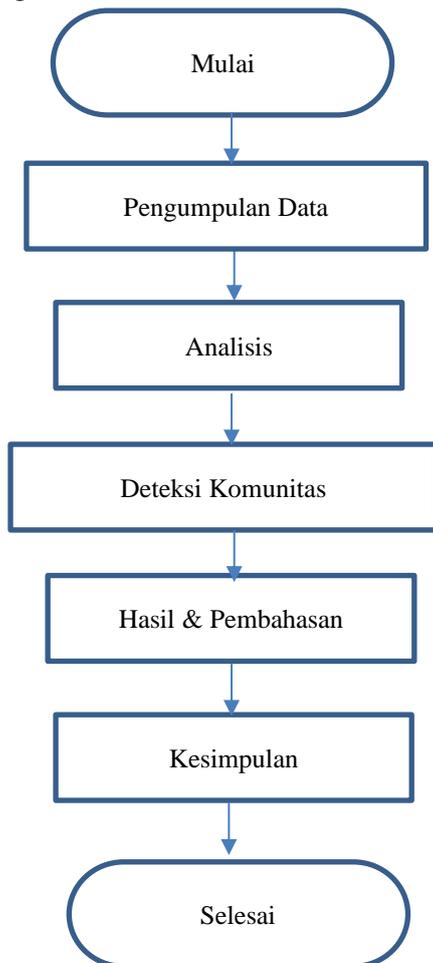
4. Sentralisasi (*Centralization*), merupakan ukuran mengenai struktur jaringan yang berhubungan dengan pemusatan suatu jaringan. Sentralisasi merujuk kepada seberapa memusat suatu jaringan pada beberapa aktor (*node*). Apakah relasi didalam suatu jaringan menyebar ke banyak orang ataukah memusat ke beberapa orang saja. Rumus untuk menentukan sentralisasi suatu jaringan adalah sebagai berikut:

$$CD = \frac{\sum(Max(C_{Di}) - C_{Di})}{n^2 - 3n + 2} \dots\dots\dots (4)$$

Dimana CD adalah sentralisasi, $\Sigma(Max(C_{Di}))$ adalah skor sentralisasi tingkatan (*degree*) maksimal dari aktor, C_{Di} adalah skor sentralitas tingkatan (*degree*) dari masing-masing aktor, dan n adalah ukuran jaringan.

5. Diameter dan Jarak (*Distance*) adalah jarak terjauh diantara dua aktor dalam suatu jaringan. Jaringan dengan diameter besar ditandai oleh persebaran aktor yang mengumpul. Sedangkan yang dimaksud dengan jarak adalah hasil rata-rata langkah (*path*) yang dibutuhkan oleh semua aktor untuk bisa saling berinteraksi.

Adapun tahapan diagram alir pada penelitian ini terdiri dari beberapa tahapan hingga mendapatkan hasil analisis yang akurat. Diagram alir pada penelitian ini dapat dilihat pada gambar 1.



Gambar 1 Flow Chart Penelitian

HASIL DAN PEMBAHASAN

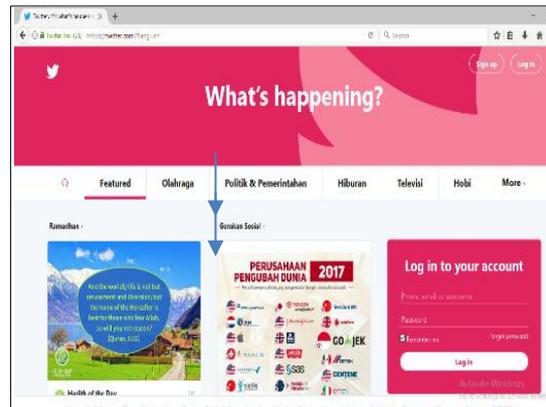
3.1 Crawling Data

Crawling data merupakan proses pengumpulan data yang akan digunakan untuk menjelajahi *web* dan mengunduh halaman *web* secara otomatis untuk mengumpulkan informasi [11].

Algoritma untuk *crawling data* yaitu sebagai berikut:

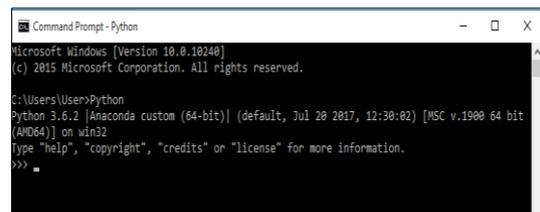
1. Import modul python yang akan digunakan;
2. Input *consumer key* dan *consumer secret*;
3. Input *Query*;
4. proses autentikasi *Twitter API*;
5. proses penarikan data twitter;
6. proses penyimpanan data yang telah didapatkan.

Pada proses ini penulis terlebih dahulu melakukan pendaftaran aplikasi di *Twitter* pada akun *twitter* melalui <https://twitter.com/>.



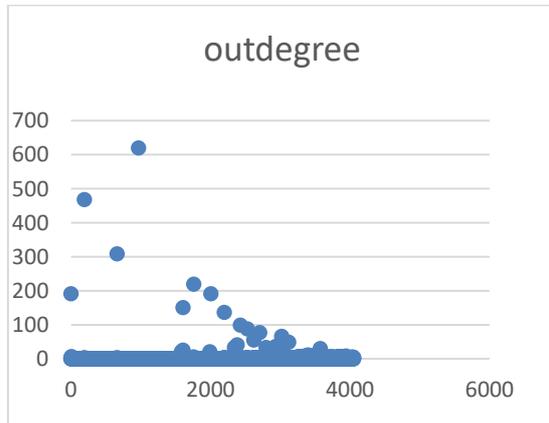
Gambar 2 Halaman untuk Login ke Twitter

Setelah melakukan registrasi maka langkah selanjutnya penulis melakukan proses *Crawling Data* yang dijalankan *melalui command prompt* dan dapat dilihat pada gambar 3.



Gambar 3 Running Aplikasi Python

Pada Gambar 8 di atas menunjukkan persebaran *indegree* keseluruhan jaringan. *Indegree* merupakan *link* yang mengarah ke aktor dan posisi aktor sebagai objek penerima. Nilai tertinggi *indegree* yaitu 503 dan nilai terendahnya adalah 0.



Gambar 9 Outdegree Distribution

Outdegree merupakan jumlah *link* yang keluar dari aktor aktor dan posisi aktor sebagai subjek. Nilai tertinggi *outdegree* yaitu 619 dan nilai terendahnya adalah 0.

2. Sentralitas Kedekatan (*Closeness Centrality*)

Perhitungan sentralitas kedekatan dapat dilihat dibawah ini:

$$C_c = \frac{4052-1}{0.00} = 0 \dots\dots\dots (6)$$

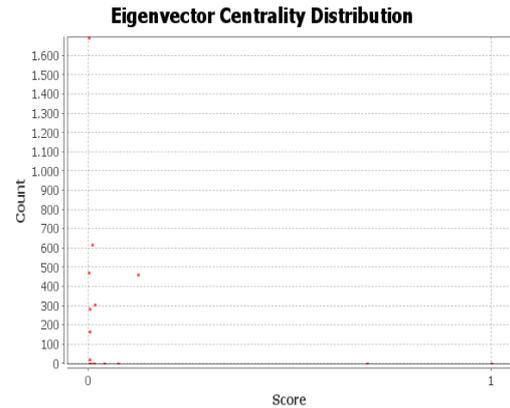
Berdasarkan hasil perhitungan diatas, angka kedekatan ialah 0 (nol). Hal ini menunjukkan bahwa jarak rata-rata aktor dengan seluruh aktor lain dalam jaringan sangat dekat (baik).

3. Sentralitas Keperantaraan (*Betweenness Centrality*)

Nilai sentralitas keperantaraan normal adalah 0-1 dimana mendekati 1 yang paling bagus.

4. *Eigenvector*

Menggambarkan seberapa penting orang yang mempunyai jaringan dengan aktor dengan cara melihat seberapa banyak jaringan yang dipunyai oleh orang/organisasi. institusi yang mempunyai relasi dengan aktor.



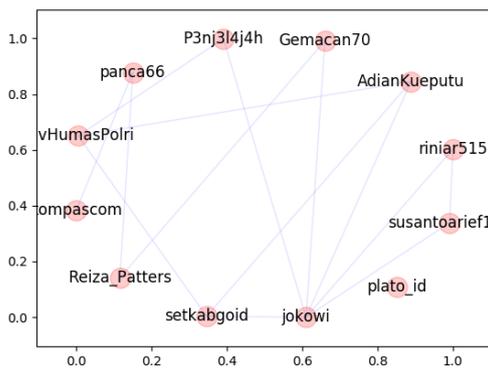
Gambar 10 Eigenvector Centrality Distribution

Sentralitas *eigenvector* ini mempunyai nilai normal 0 hingga 1, yang berarti bahwa aktor tertinggi memiliki *eigenvector* tertinggi pula.

Tabel 1 Important User

No	User	Eigenvector
1	jokowi	0.666
2	setkabgoid	0.090
3	DivHumasPolri	0.082
4	Reiza_Patters	0.079
5	kompascom	0.055
6	panca66	0.049
7	riniar515	0.044
8	susantoarief141	0.044
9	AdianKueputu	0.042
10	Gemacan70	0.042
11	plato_id	0.039
12	P3nj3l4j4h	0.039

Pada Tabel 1 terdapat 12 akun sebagai *important users*. Akun tersebut adalah akun “Jokowi” dengan *eigenvector* 0.666, akun “setkabgoid” dengan *eigenvector* 0.090, akun “DivHumasPolri” dengan *eigenvector* 0.082, akun “Reiza_Patters” dengan *eigenvector* 0.079, akun “kompascom” dengan *eigenvector* 0.055, akun “panca66” dengan *degreeeigenvector* 0.049, akun “riniar515” dengan *eigenvector* 0.044, akun “susantoarief141” dengan *eigenvector* 0.044, akun “AdianKueputu” dengan *eigenvector* 0.042, akun “Gemacan70” dengan *eigenvector* 0.042, akun “plato_id” dengan *eigenvector* 0.039, dan akun “P3nj3l4j4h” dengan *eigenvector* 0.039. Berikut merupakan visualisasi *user* dengan *eigenvector* tertinggi.



Gambar 11 Graph Important User

3.4 Analisis Level Kelompok (Group):
Identifikasi Sub-Group Dalam Jaringan

Analisis jaringan dalam level kelompok umumnya dipakai untuk mengidentifikasi keanggotaan aktor (node) dalam kelompok-kelompok kecil (grup) dalam jaringan.

1. Komponen

Komponen adalah pengelompokan yang memasukkan semua aktor (node) yang mempunyai link (edges) dengan jaringan. Komponen bisa dibagi kedalam dua jenis. Pertama komponen kuat (strong component). Komponen kuat ditandai oleh adanya hubungan langsung antara aktor. Kedua komponen lemah (weak component) yang ditandai dengan hubungan tidak langsung diantara aktor. Dari hasil yang didapatkan terdapat 2 komponen lemah (weak component) dan 4051 komponen kuat (strong component).

3.5 Analisis Level Sistem : Struktur dan Karakter

1. Ukuran (Size)

Jaringan dengan ukuran kecil, antar-aktor (node) lebih kohesif dibandingkan dengan jaringan ukuran besar. Struktur relasi di antara aktor juga berbeda antara jaringan dengan ukuran kecil dan besar (Carolan, 2013:101).Ukuran dari jaringan komunitas dengan kata kunci “jokowi” adalah sejumlah 4052 nodes.

2. Kepadatan (Densitas/Density)

Angka dalam kepadatan (densitas) ini yaitu 0 hingga 1, dimana makin besar nilai menunjukkan makin tingginya kepadatan (densitas) dari suatu jaringan.

$$D = \frac{4552}{4052(4052-1)} = \frac{4552}{16414652} = 0,0002..... (7)$$

Kepadatan jaringan pada penelitian ini adalah 0,0002 jika dibulatkan menjadi 0 (nol). Hal ini berarti kepadatan jaringan rendah.

3. Resiprositas (Mutualitas/Reciprocity)

Angka resiprositas dari 0 hingga 1, dimana 1 memperlihatkan resiprositas sempurna, dimana masing-masing aktor saling memilih atau saling berinteraksi.

$$D = \frac{4051}{4052} = 0,99..... (8)$$

Resiprositas pada penelitian ini yaitu 0.99 jika dibulatkan menjadi 1. Hal ini menunjukkan bahwa masing-masing aktor pada jaringan saling berinteraksi.

4. Sentralisasi (Centralization)

Sentralisasi merupakan ukuran mengenai berikut perhitungan sentralisasi pada penelitian ini:

$$CD = \frac{871}{20720704-13658} = 0.004 \%..... (9)$$

Berdasarkan hasil perhitungan diatas didapatkan sentralisasi sebesar 0.004%.

5. Diameter dan Jarak (Distance)

Yang dimaksud dengan diameter adalah jarak terjauh diantara dua aktor dalam suatu jaringan). Jaringan dengan diameter besar ditandai oleh persebaran aktor yang mengumpul. Sedangkan yang dimaksud dengan jarak adalah hasil rata-rata langkah (path) yang dibutuhkan oleh semua aktor untuk bisa saling berinteraksi. Diameter jaringan pada penelitian ini adalah 6 dan jarak / Average Path length yaitu 2.224815380301, jika dibulatkan menjadi 2.23.

KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini adalah berdasarkan penelitian yang telah dilakukan, bahasa pemrograman phyton dan algoritma katz berhasil menemukan ukuran dari jaringan komunitas. Ukuran dari jaringan komunitas dengan kata kunci “jokowi” adalah sejumlah 4052 nodes. Kepadatan jaringan rendah yaitu 0,0002 jika dibulatkan menjadi 0 (nol). Resiprositas pada penelitian ini yaitu

berlangsung dua arah. Diameter jaringan pada penelitian ini adalah 6 dan jarak / *Average Path length* yaitu 2.23.

Terdapat 12 akun sebagai *important users*. Akun tersebut adalah akun "Jokowi" dengan *degree* 0.666, akun "setkabgoid" dengan *degree* 0.090, akun "DivHumasPolri" dengan *degree* 0.082, akun "Reiza_Patters" dengan *degree* 0.079, akun "kompascom" dengan *degree* 0.055, akun "panca66" dengan *degree* 0.049, akun "riniar515" dengan *degree* 0.044, akun "susantoarief141" dengan *degree* 0.044, akun "AdianKueputu" dengan *degree* 0.042, akun "Gemacan70" dengan *degree* 0.042, akun "plato_id" dengan *degree* 0.039, dan akun "P3nj314j4h" dengan *degree* 0.039.

Sedangkan saran dari penelitian ini adalah dalam penelitian ini hanya menganalisis struktur komunitas yang ada pada suatu jaringan pertemanan di jejaring social *Twitter*. Maka dari itu diperlukan penelitian lebih lanjut tentang analisis struktur komunitas dengan menggunakan sampel pertemanan di jejaring sosial lainnya seperti facebook.

Analisis dalam penelitian ini hanya menggunakan satu algoritma saja yaitu algoritma Katz. Maka dari itu diperlukan beberapa algoritma secara bersamaan untuk mengetahui perbandingan hasil yang didapat.

REFERENSI

- [1] S. Asmara and dkk, "Twitter dan Public Sphere: Studi Fenomenologi Tentang Twitter Sebagai Media Alternatif Komunikasi Politik," *Jurnal Ilmiah Ilmu Komunikasi*, vol. 2, no. 2, pp. 75-84, 2020.
- [2] R. A. Setiawan and dkk, "Analisis Komunikasi Sosial Media Twitter sebagai Saluran Layanan Pelanggan Provider Internet dan Seluler di Indonesia," *Journal of Information Systems Engineering and Business Intelligence*, vol. 3, no. 1, pp. 16-25, 2017.
- [3] A. Husnusyifa, "Twitter merupakan jenis media sosial microblogging yang memfasilitasi pengguna untuk menulis pendapatnya . Secara historis , kehadiran dan munculnya media sosial Twitter yang menyediakan ruang tertentu atau media sosial lainnya , di Twitter pengguna lain," *IDEA: Jurnal Humaniora*, vol. 2, no. 2, pp. 120-133, 2019.
- [4] Z. F. Nurhadi, "Model Komunikasi Sosial Remaja Melalui Media Twitter," *Jurnal ASPIKOM*, vol. 3, no. 3, pp. 539-549, 2017.
- [5] A. Y. Widyaningrum, "Kajian Tentang Komunitas Virtual: Kesempatan dan Tantangan Kajian di Bidang Ilmu Komunikasi," *Jurnal Komunikatif*, vol. 10, no. 2, pp. 141-152, 2021.
- [6] D. W. Prayugo, "PENGARUH KOMUNITAS VIRTUAL TERHADAP MINAT BELI ONLINE PADA GRUP FACEBOOK BUBUHAN SAMARINDA," *Ejournal Ilmu Komunikasi*, vol. 6, no. 1, p. 143 – 157, 2018.
- [7] A. Aktayeva and dkk, "Cognitive Computing Cybersecurity: Social Network Analysis," in *3rd International Scientific Conference on Convergent Cognitive Information Technologies*, Switzerland, 2020.
- [8] Eriyanto, *Analisis Jaringan Media Sosial*, Jakarta: KENCANA, 2021.
- [9] N. A. Azmi and dkk, "Social Media Network Analysis (SNA): Identifikasi Komunikasi dan Penyebaran Informasi Melalui Media Sosial Twitter," *Jurnal Media Informatika Budidarma*, vol. 5, no. 4, pp. 1422-1430, 2021.
- [10] D. Kurniawan and dkk, "Pemanfaatan Social Network Analysis (Sna) Untuk Menganalisis Kolaborasi Karyawan Pada Pt. Arum Mandiri Group," *Jurnal Transformatika*, vol. 17, no. 2, pp. 149 - 159, 2020.
- [11] R. Hanifah and dkk, "Implementasi Web Crawling Untuk Mengumpulkan Web Crawling Implementation for Collecting," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 5, no. 5, pp. 531-536, 2018.