

Analisis Jejaring Sosial Pengguna Twitter Paling Influential Yang Telah Wafat Menggunakan Centrality Method

Abstrak—Beriringan dengan perkembangan pesat *social network*, jumlah dan keragaman *platform-platform* di internet yang menawarkan fitur untuk berhubungan dengan orang lain, juga turut berkembang. Salah satu dari *social network* tersebut adalah Twitter. Twitter memungkinkan terbentuknya pemetaan antara individu-individu dan relasi yang terjadi di antara individu-individu tersebut. Bagaimana hubungan tersebut dapat terjadi, konsekuensi dari tiap hubungan, dan analisis hubungan antar individu, dapat dipelajari menggunakan teknik *Social Network Analysis* atau Analisis Jejaring Sosial. Tak lepas dari sisi kemanusiaan, pengguna-pengguna Twitter yang merupakan suatu individu biasanya meninggalkan akun mereka dalam status aktif, ketika mereka meninggal dunia. Kasus influentialnya seorang pengguna twitter ketika wafat, dapat dipelajari dalam SNA. Analisis Jejaring Sosial adalah salah satu cabang informatika sosial yang secara spesifik melakukan analisis pola hubungan atau fenomena hubungan pada suatu *social network*. Salah satu dari contoh implementasinya adalah *centrality measurement* yang digunakan untuk menentukan pengguna yang paling berpengaruh dalam persebaran informasi. *Centrality measurement* merupakan nilai sentralitas pada relasi antara pengguna yang dipetakan kedalam bentuk Graf. Salah satu metode untuk menghitung *centrality measurement* adalah dengan mencari *node* yang paling sering dilewati oleh *shortest path*, kemudian menghitung *node* tersebut sebagai *node* yang memiliki *centrality* paling tinggi.

Kata kunci : *Social Network Analysis, Centrality Method, Node, Graf*

Abstract— Along with the rapid development of social networks, the number and variety of platforms on the internet that offer features for connecting with other people has also been growing. One of those social networks is Twitter. Twitter allows the formation of mapping between individuals and the relationships that occur between individuals. How these relationships can occur, the consequences of each relationships, and the analysis of relationships between individuals, can be studied using *Social Network Analysis* techniques. Related to humanity aspects, Twitter users who are individuals usually leave their accounts in an active status when they pass away. The influential case of a Twitter users when they pass away, can be

studied in SNA. *Social Network Analysis* is a branch of social informatics that specifically analyzes relationship patterns or relationship phenomena on a social network. One example of its implementation is *centrality measurement* which is used to determine the most influential users in the distribution of information. *Centrality measurement* is the value of centrality in the relationship between users, that is mapped into a graph. One method of calculating *centrality measurement* is by finding the node that is most often passed by the shortest path, and then counting that particular node, as the node with the highest centrality.

Keywords : *Social Network Analysis, Centrality Method, Node, Graph*

I. PENDAHULUAN

A. Latar Belakang

Social Network merupakan suatu jaringan yang membantu kumpulan dari individu, untuk berkomunikasi dalam bentuk relasi antar individu tersebut, dan juga dalam bentuk informasi relasional [6]. *Social network* merupakan peta dari interaksi dan keterhubungan antar individu, berisi relasi antar individu tersebut yang berbentuk graf [1]. Dengan menggunakan *Social Network Analysis* (Analisis Jejaring Sosial), kita mampu untuk mempelajari bagaimana suatu hubungan dapat terjadi, konsekuensi dari tiap hubungan yang terjadi, dan menganalisis hubungan antar individu. Salah satu bentuk implementasi analisis jejaring sosial adalah, *centrality measurement* atau perhitungan sentralitas [6]. *Centrality measurement* adalah perhitungan nilai sentralitas, yang dilakukan untuk *men-sorting* individu dengan level pengaruhnya masing-masing, dalam suatu penyebaran informasi. Metode yang dapat digunakan dalam perhitungan centrality terkomputasi antara lain adalah, *degree centrality* (DC), *betweenness centrality* (BC), *closeness centrality* (CC), *k-shell centrality* (KS), dan *local centrality* (LC) [6]. Penelitian ini akan menggunakan metode *centrality method* untuk mencari pengguna twitter yang paling *influential* pada periode tertentu, dengan status sudah wafat.

B. Tujuan

1. Merepresentasikan relasi antara individu atau pengguna yang terdapat pada jejaring sosial Twitter ke dalam sebuah bentuk graf yang menggambarkan hubungan antar individu atau pengguna dalam relasi tersebut.

2. Menganalisis dan mengimplementasikan metode *Centrality* untuk menentukan user yang paling berpengaruh, melalui banyaknya pengguna lain yang me-mention dengan ucapan berbela sungkawa.

C. Dataset

Dataset yang digunakan adalah data yang diambil dari twitter dengan *crawling*. Data yang diambil adalah tweet beserta propertinya, seperti : seberapa banyak di-retweet atau di-like, siapa user lain yang di-reply atau me-reply, dan siapa user lain yang di-mention pada tweet tersebut. Dari data yang dikumpulkan akan dilakukan analisis *centrality method*.

II. DASAR TEORI

A. Social Network

Social Network atau jejaring sosial adalah peta dari suatu struktur sosial yang terdiri atas individu (atau organisasi) yang dilambangkan sebagai “*node*”, yang terhubung pada satu atau lebih relasi[7]. Sebuah jejaring sosial terjadi ketika semua *node* memiliki hubungan dengan *node* lainnya. Fitur yang terdapat dalam jejaring sosial yaitu adanya informasi relasional. Oleh karena itu, jejaring sosial lebih fokus pada informasi relasional daripada atribut *node* [9]. Hubungan antar node digambarkan dengan sebuah graf, dan setiap *node* dapat dihubungkan satu sama lain dengan sebuah *edge*. Grafik yang merepresentasikan data sosial, memungkinkan untuk dilakukannya analisis informasi yang terdapat pada hubungan sosial [2].

B. Twitter

Twitter adalah *platform microblogging* dan layanan *social networking* yang memungkinkan pengguna untuk melakukan “*follow*” antar satu sama lain. Pengguna juga dapat mengirim dan membaca “*tweet*”, yaitu pesan teks terdiri atas maksimal 140 karakter. Layanan twitter juga digunakan oleh berbagai organisasi untuk memasarkan produk.

1. Follow

Follow adalah suatu fungsi yang melakukan penandaan bahwa suatu user yang melakukan “*follow*” akan menerima *tweet* dari user lain yang di “*follow*”.

2. Following

Following adalah suatu derajat atau angka yang merepresentasikan jumlah *user* lain yang di-*follow* oleh suatu *user*.

3. Followed

Followed adalah suatu derajat atau angka yang merepresentasikan jumlah *user* lain yang mem-*follow* oleh suatu *user*.

4. Unfollow

Follow adalah suatu fungsi yang melakukan penandaan bahwa suatu *user* menghentikan penerimaan *tweet* dari *user* lain yang sedang di “*follow*”.

5. Tweet

Tweet adalah suatu fungsi yang melakukan pengunggahan suatu post secara publik (yang juga disebut *Tweet*), yang bisa berisi kombinasi dari media gambar, video, atau karakter dengan batas jumlah saat ini 280 buah.

6. Reply

Reply adalah suatu fungsi yang melakukan pengunggahan suatu post untuk membalas suatu *Tweet*, yang bisa berisi kombinasi dari media gambar, video, atau karakter dengan batas jumlah saat ini 280 buah.

User lain yang mendapatkan *Reply*, akan mendapatkan pemberitahuan, dengan properti dari reply tersebut bertanda “*Replying to*”.

7. Mention

Mention adalah suatu *Tweet* yang mengandung suatu tag berupa username dari user lain.

User lain yang mendapatkan mention, akan mendapatkan pemberitahuan, dengan properti dari tweet tersebut bertanda “*Mentioned you*”.

8. Retweet (RT)

Retweet adalah suatu aksi yang melakukan penyebarluasan suatu tweet dari user lain maupun tweet diri sendiri yang sudah lampau.

Tweet yang di-retweet akan memiliki suatu perubahan pada propertinya, berupa berubahnya angka di sebelah lambang retweet, berdasarkan jumlah berapa kali tweet tersebut di-retweet.

9. Like

Like adalah suatu aksi yang melakukan penambahan properti dari suatu tweet, yang berupa berubahnya angka di sebelah lambang like, berdasarkan jumlah berapa kali tweet tersebut di-like.

10. Quote Retweet

Quote Retweet adalah suatu aksi yang melakukan retweet, namun diberi tambahan kutipan berupa string yang ditulis oleh user yang me-retweet, dengan batas 280 character.

C. Social Network Analysis

Social Network Analysis (SNA) adalah sebuah cabang studi informatika sosial yang mempelajari tentang keterhubungan antar individu dengan mengimplementasikan

teori graf [8]. Pada *social network*, individu atau organisasi direpresentasikan sebagai *node*. Sedangkan relasi yang terjadi antar individu direpresentasikan dengan *edge*. Sebuah jaringan sosial juga melakukan pemetaan banyak orang dimana terdapat relasi antar individunya [7].

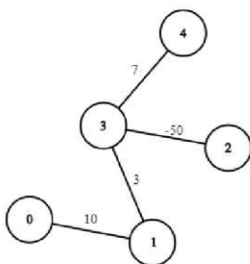
D. Graf

Graf adalah cara untuk merepresentasikan relasi antar pengguna pada struktur jejaring sosial. Graf dapat dinotasikan sebagai himpunan (V, E) , dimana V merupakan himpunan dari simpul-simpul (*vertex* atau *node*) dan E merupakan dari himpunan sisi-sisi (*edges*) yang menghubungkan sepasang simpul [5]. Graf dapat diklasifikasikan terhadap dua aspek, yaitu:

1. Berdasarkan keberadaan bobotnya

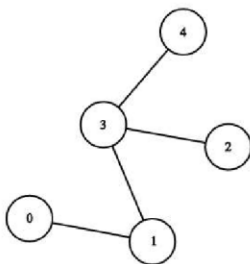
Berdasarkan keberadaan bobotnya, graf dapat diklasifikasikan menjadi dua, yaitu graf berbobot dan graf tidak berbobot. Bobot dapat merepresentasikan jumlah interaksi, kekuatan hubungan, jarak antara *node*, dan lainnya. Sedangkan graf tidak berbobot hanya merepresentasikan suatu hubungan antar *node*-nya saja.

Gambar 1



Graf berbobot

Gambar 2

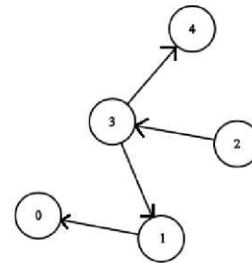


Graf tidak berbobot

2. Berdasarkan keberadaan arahnya

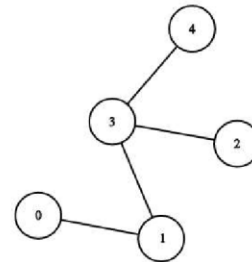
Berdasarkan keberadaan arahnya, graf dapat diklasifikasikan menjadi dua, yaitu graf berarah dan graf tidak berarah. arah dapat merepresentasikan arah interaksi, dan relasi antara dua *node*. Sedangkan graf tidak berarah hanya merepresentasikan suatu hubungan antar *node*-nya saja.

Gambar 3



Graf berarah

Gambar 4



Graf tidak berarah

E. Representasi graf dalam Twitter sebagai social networking platform

Graf pada umumnya terdiri atas *node* dan *edge*. *Node* pada Twitter, dapat merepresentasikan pengguna, dan *edge* dapat merepresentasikan hubungan antar pengguna.

Gambar 5



Representasi graf pada relasi following-followed twitter

'shulhannur' merupakan follower pengguna 'joseprb' dan 'muharik'. Dapat juga dikatakan bahwa pengguna 'joseprb' dan 'muharik' mem-follow pengguna 'shulhannur'. Informasi yang berupa tweet dari pengguna 'shulhannur' dapat muncul di homepage atau timeline pengguna 'muharik' saja, karena pengguna 'joseprb' bukan follower dari pengguna 'shulhannur'.

F. Adjacency Matrix

Adjacency Matrix atau Matriks ketetanggaan adalah jenis matriks yang digunakan dalam analisis jejaring sosial. Matriks ketetanggaan menggambarkan tabel yang baris dan kolomnya menunjukkan relasi antar *node*. Pengukuran dilakukan berdasarkan jumlah *node* (n).

Dalam Matriks ketetanggaan, 1 merepresentasikan hubungan, artinya jika *node* 1 dan *node* 2 bernilai 1 maka *node* 1 memiliki hubungan pertemanan dengan *node* 2.

Sebuah Matriks Ketetanggaan $M(G) = [a_{ij}]$ dari graf $G = (V, E)$ dimana $V = \{V_1, \dots, V_n\}$ didefinisikan dengan [5] :

$$a_{ij} = \begin{cases} 1, & \text{jika } (V_i, V_j) \in E(G) \\ 0, & \text{jika } (V_i, V_j) \notin E(G) \end{cases} \quad (1)$$

G. Centrality Measurement

Dalam graph theory dan SNA, terdapat berbagai centrality measurement terhadap node dalam grafik yang menentukan tingkat kepentingan relatif dari node dalam suatu graf (contohnya dalam kasus ini, seberapa influentialnya seseorang dalam jaringan sosial) [6].

Representasi dari centrality measurement pada penelitian ini dapat diambil dari beberapa twitter user yang sudah meninggal, dan berita tentang kematian user tersebut di-tweet oleh user lainnya yang menggunakan string berisi kalimat bela sungkawa, contohnya "RIP". Kemudian dari kumpulan tweet tersebut diukur seberapa central user yang menge-tweet berdasarkan beberapa faktor seperti, berapa kali tweet tersebut di-retweet, berapa kali tweet tersebut di-like, berapa kali tweet tersebut di-reply dan lain lain.

Penelitian ini akan menggunakan dua method centrality measuement yaitu *closeness centrality*, dan *eigenvector centrality*

H. Closeness Centrality

Closeness Centrality adalah salah satu metode *centrality measurement* dalam suatu jejaring sosial, yang berfokus pada seberapa dekat suatu node dengan node lainnya. *Closeness Centrality* akan menghitung *centrality measurement* sebuah node berdasar jumlah jarak terpendek antara node tersebut dengan node lainnya [4].

Sebuah *Closeness Centrality* C_c dari node i , yang dipengaruhi oleh jarak antara node i dan j , dan besarnya *distance* antara kedua node tersebut, didefinisikan dengan :

$$C_c(i) = \frac{1}{\sum_{j=1}^N d(i, j)} \quad (2)$$

I. Eigenvector Centrality

Eigenvector Centrality adalah salah satu metode *centrality measurement* dalam suatu jejaring sosial, yang berfokus pada memberikan nilai relatif kepada semua node dalam jaringan tersebut, dengan prinsip bahwa node yang memiliki nilai relatif tertinggi, adalah node yang tetangga-tetangganya mempunyai nilai Eigenvector Centrality yang terbesar [4].

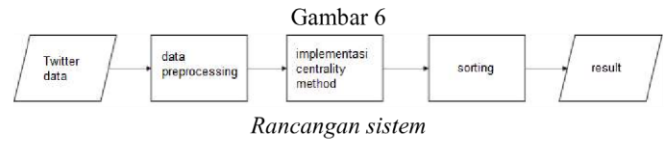
Sebuah *Eigenvector Centrality* C_e , dari node ke- i v_i , yang dipengaruhi oleh sebuah konstan λ , matriks ketetanggaan A , dan simpul lainnya yang dilambangkan j , didefinisikan dengan :

$$C_e(v_i) = \frac{1}{\lambda} \sum_{j=1}^n A_{ji} C_e(v_j) \quad (3)$$

III. PERANCANGAN DAN IMPLEMENTASI ANALISIS

A. Sistem

Rancangan sistem untuk menentukan urutan pengguna yang paling influential dalam penyebaran informasi di Twitter dengan menggunakan Centrality Method, secara umum dapat digambarkan sebagai berikut:



Data yang akan diinputkan berupa dataset yang diperoleh dari Twitter menggunakan API. Data set yang diambil terdiri Tweets, dan properti propertinya, seperti : seberapa banyak di-retweet atau di-like, siapa user lain yang di-reply atau me-reply, dan siapa user lain yang di-mention pada tweet tersebut.

Setelah crawling dataset, data di-preprocessing, kemudian dilakukan proses merge duplicate untuk menghilangkan atribut yang identik. Kemudian dataset direpresentasikan dalam matriks ketetanggaan $n \times n$, dengan n merupakan jumlah node yang terambil.

Pada penelitian ini akan dilakukan perhitungan centrality untuk menentukan pengguna yang paling berpengaruh yang sudah wafat. Hasil perhitungan tersebut akan menampilkan daftar sorting pengguna berdasarkan banyaknya mention berisi ucapan bela sungkawa.

B. Sorting

Proses sorting dilakukan berdasarkan pengguna yang memiliki nilai centrality paling tinggi, node yang memiliki centrality paling tinggi merupakan node yang paling influential dalam persebaran informasi.

IV. HASIL

A. Twitter Data Crawling

Proses crawling dilakukan menggunakan kode python yang mengimplementasikan atau mengimport library twint, yang kemudian diexport menjadi file csv yang berisi 5001 baris dataset, masing masing berupa tweet dan propertinya

Gambar 7

```

[ ] # !pip3 install twint
!pip3 install --user --upgrade git+https://github.com/twintproject/twint.git@origin/master#egg=twint

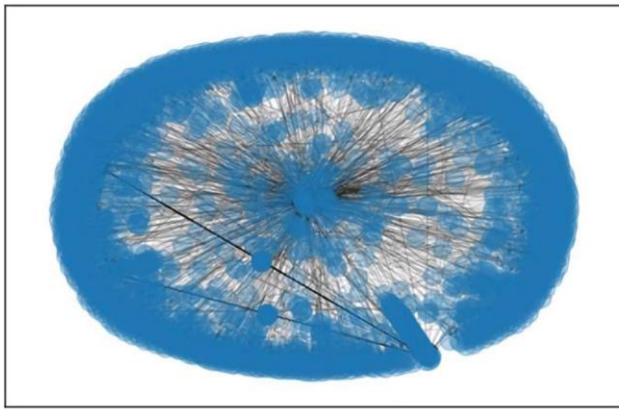
[ ] import twint
import pandas as pd
import numpy as np
import nest_asyncio
import networkx as nx
import matplotlib as plt
nest_asyncio.apply()
  
```

Twitter Data Crawling

A. Analisis

Dari file cxv tersebut diperoleh masing masing user yang menge-tweet sebagai node, beserta properti tweetnya sebagai arah dan bobot dari vertex yang menyambung antara kedua node. Kemudian hasil dari centrality measurement divisualisasikan dalam bentuk graf berikut

Gambar 8



Graf hasil

Kemudian dilakukan proses terakhir yaitu sorting berdasarkan nilai eigenvector dan closeness centrality-nya

Gambar 9

```
10 most influential users in Prince Philip's death are:
Eigenvector Centrality:
[('RoyalFamily', 0.9444678811836785),
 ('ClarenceHouse', 0.19472491045352974),
 ('KensingtonRoyal', 0.10725489720167683),
 ('piersmorgan', 0.05414810344162327),
 ('SkyNews', 0.05102417439691422),
 ('BBCNews', 0.047900245352205176),
 ('BBCBreaking', 0.03852845821807806),
 ('superscuba83', 0.029156671083950966),
 ('MailOnline', 0.02915667108395096),
 ('TowerOfLondon', 0.02603274203924193)]
```

10 user dengan nilai eigenvector centrality tertinggi dalam memberitakan kasus kematian Prince Philip

Dengan sorting closeness centrality berikut, dapat dilihat bahwa terdapat user yang menduduki posisi yang sama, dan terdapat juga user yang bertukar posisi,

Lebih spesifik lagi user yang bertukar posisi adalah superscuba83 dengan Mailonline. superscuba83 memiliki nilai *eigenvector centrality* yang lebih tinggi dibandingkan Mailonline. Sedangkan Mailonline memiliki nilai *closeness centrality* yang lebih tinggi dibandingkan superscuba83.

Gambar 10

```
Closeness Centrality:
[('RoyalFamily', 0.11996943018335313),
 ('ClarenceHouse', 0.024711040255081706),
 ('KensingtonRoyal', 0.013164038034504355),
 ('piersmorgan', 0.006775607811877242),
 ('SkyNews', 0.006001922490797834),
 ('BBCNews', 0.005978477481068154),
 ('BBCBreaking', 0.004782781984854523),
 ('MailOnline', 0.0035870864886408927),
 ('superscuba83', 0.0031885213232363493),
 ('TowerOfLondon', 0.0031885213232363493)]
```

10 user dengan nilai closeness centrality tertinggi dalam memberitakan kasus kematian Prince Philip

DAFTAR PUSTAKA

- [1] Boyd, Danah M. and Ellison Nicole B.. "Social Network Sites: Definition, History, and Scholarship". Computer-Mediated Communication. 2008.
- [2] C. N. Utami, "Analisis dan Implementasi Community Detection Menggunakan Algoritma Girvan and Newman dalam Social Network," IT Telkom, Bandung, 2014.
- [3] Geisberger Robert, et al. "Better Approximation of Betweenness Centrality". Institute for Theoretical Computer Science. Algorithmics II Universität Karlsruhe (TH). Germany
- [4] Prabowo and Dwi "Analisis perbandingan Node Centrality pada jaringan manusia riil terhadap jaringan teoritis (Random Network dan Scale-Free Network)". 2016.
- [5] Ruohonen Keijo. Graph Theory. 2013.
- [6] Scott John. "Social Network Analysis Theory and Application". 2011.
- [7] Susanto Budi, et al. "Penerapan Social Network Analysis dalam Penentuan Centrality". Fakultas Teknologi Informasi. Universitas Kristen Duta Wacana. Yogyakarta
- [8] Tsevetovat Maksim and Kouznetsov Alexander."Social Network Analysis for Startups". O'Reilly. 2011.
- [9] Wasserman Stanley and Faust Katherine. "Social Network Analysis : Methods and Applications". United State. America. Cambridge University Press. 1994.