

Analisis Jejaring Sosial Terhadap Keberadaan Komunitas Berdasarkan Ucapan Bela Sungkawa Menggunakan In-Degree 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 *Community Detection* yang digunakan untuk mendefinisikan keberadaan suatu komunitas spesifik dalam persebaran informasi. *Community Detection* pada dasarnya merupakan pengelompokkan dan pengklasifikasian individu-individu yang terdapat pada jejaring sosial. Pada penelitian ini, akan diimplementasikan *Community Detection* menggunakan metode *In-Degree Centrality*, untuk mendeteksi keberadaan komunitas berdasarkan afiliasinya pada tweet ucapan bela sungkawa di Twitter menggunakan Python.

Kata kunci : *Social Network Analysis, Community Detection, In-Degree Centrality*

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 Community Detection, which is used to define the existence of a specific community, in the distribution of information. Community Detection is basically classifying a group of individuals in a specific social network. In this research, Community Detection will

be implemented using a method called In-Degree Centrality, to detect the existence of communities, based on their affiliation using their mourning tweet, on Twitter using Python.

Keywords : *Social Network Analysis, Community Detection, In-Degree Centrality*

I. PENDAHULUAN

A. Latar Belakang

Entertainment atau hiburan, adalah salah satu industri terbesar di dunia. *Entertainment* dapat diklasifikasikan menjadi berbagai cabang, contohnya olahraga, perfilman, sinema elektronik, musik dan beberapa klasifikasi lainnya, tergantung faktor pengklasifikasiannya. *Entertainment Industry* hampir selalu dijalankan oleh individu-individu yang berkompeten, dan individu-individu terkait tersebut sebagian besar merupakan kontributor aktif di jejaring sosial. Tak lepas dari sisi kemanusiaan, individu-individu tersebut biasanya meninggalkan efek yang besar, ketika mereka meninggal dunia. Contohnya jika seseorang atlet Bola Basket ternama pada industri hiburan olahraga meninggal karena kecelakaan mobil, akan bermunculan individu-individu maupun organisasi- organisasi di jejaring sosial yang mereka pernah aktif, untuk menunjukkan perasaan berbela sungkawa. Sehingga dapat diimplikasikan bahwa integritas pada dunia industri hiburan, sangat erat kaitannya dengan jejaring sosial.

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, *Community Detection* atau *Community Search* [10]. *Community Detection* adalah perhitungan nilai keberadaan dan pola suatu kumpulan individu, yang dilakukan untuk *men-sorting* keterikatan antar individu, untuk diklasifikasikan bahwa kumpulan individu tersebut merupakan suatu komunitas atau bukan. Metode yang dapat digunakan dalam menemukan komunitas terkomputasi antara lain adalah, *In-Degree Centrality*, yang akan digunakan dalam penelitian ini, untuk mencari komunitas yang memiliki

pola kesamaan pada penyampaian ucapan bela-sungkawa di Twitter, dengan menganalisis String yang mereka Tweet.

B. Tujuan

1. Merepresentasikan relasi similaritas, dan modularitas 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 *InDegree Centrality* untuk menentukan komunitas apa saja yang berada, pada suatu kumpulan dataset, melalui banyaknya pengguna lain yang men-Tweet dengan ucapan bela sungkawa yang ditujukan pada orang yang sama.

C. Dataset

Dataset yang digunakan adalah data yang diambil dari twitter dengan *crawling* menggunakan library *twint* pada Python. Data yang diambil adalah 5001 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 *In-Degree Centrality*.

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 penyebaran 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].

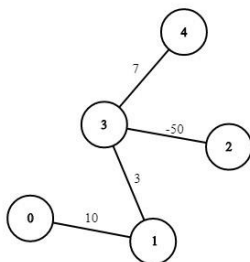
D. Graf

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

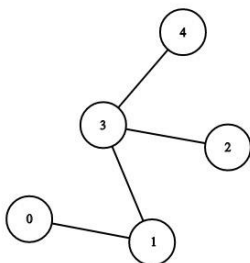
Berdasarkan keberadaan bobotnya, graf dapat diklasifikasikan menjadi dua, yaitu graf berbobot dan graf tidak berbobot. Bobot dapat merepresentasikan jumlah interaksi, kekuatan hubungan, jarak antara *node*, dan lainnya. Sedangkan graf tidak berbobot hanya merepresentasikan suatu hubungan antar *node*-nya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

Gambar 1



Graf berbobot

Gambar 2



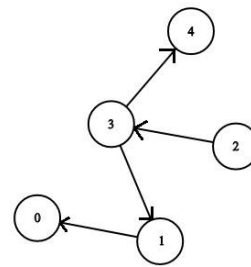
Graf tidak berbobot

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

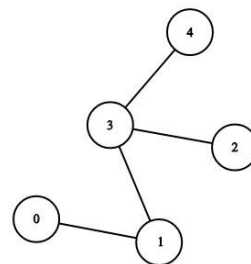
relasi antara dua *node*. Sedangkan graf tidak berarah hanya merepresentasikan suatu hubungan antar *node*-nya saja.

Gambar 3



Graf berarah

Gambar 4

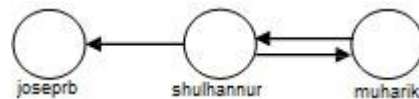


Graf tidak berarah

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] :

1, jika $(V_i, V_j) \in E(G)$

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

G. Community Detection

Dalam graph theory dan SNA, terdapat berbagai Community Detection terhadap node dalam grafik yang menentukan tingkat kepentingan relatif dari graf yang terbentuk (contohnya dalam kasus ini, adanya komunitas berisi individu-individu yang menge-tweet ucapan bela sungkawa, yang ditujukan kepada orang yang sama) [6].

Masalah yang penting dalam menemukan komunitas pengguna atau community detection, adalah melakukan analisis berdasarkan isi pesan dan hubungan dengan pengguna lain. Suatu komunitas adalah pola dengan bobot yang besar pada hubungan internal dan rendah pada hubungan eksternal. Hubungan ini dapat diketahui oleh kesamaan konten pada tweet antar pengguna, hubungan following-followed di antara mereka dan juga kesamaan lainnya pada data pribadi mereka seperti lokasi, jenis kelamin, usia, dll. struktur kedekatan ini kemudian dapat digunakan untuk berbagai keperluan seperti menargetkan skema konten hiburan, menghitung kepopuleran suatu konten hiburan, mendeteksi trend hiburan di masyarakat, dan lain sebagainya. [11]

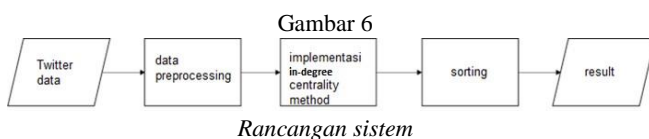
Representasi dari Community Detection pada penelitian ini dapat diambil dari beberapa Tweet yang menggunakan string berisi kalimat bela sungkawa, contohnya “RIP”, “rip”, “rest in piece”, “Rest in Piece”, “Rest In Piece”, “REST IN PIECE”, “Turut berduka”, dan lain lain. Kemudian dari kumpulan tweet tersebut diukur seberapa similar tweet-tweet tersebut, jika similar, maka akan dianggap sebagai suatu komunitas.

Penelitian ini akan menggunakan satu *Community Detection* method yaitu *In-Degree Centrality*.

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 mereply, 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 nxn, dengan n merupakan jumlah node yang diambil.

Pada penelitian ini akan dilakukan perhitungan Community Detection untuk menentukan seberapa similar Tweet yang menggunakan string berisi kalimat bela sungkawa. Hasil perhitungan tersebut akan menampilkan daftar komunitas berdasarkan seberapa similar Tweet yang menggunakan string berisi kalimat bela sungkawa.

B. Sorting

Proses sorting dilakukan berdasarkan komunitas yang memiliki anggota paling banyak, dan komunitas yang memiliki anggota paling banyak, merupakan komunitas yang paling similar 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 yng berisi 5001 baris dataset, masing masing berupa tweet dan propertinya

Gambar 7

```

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

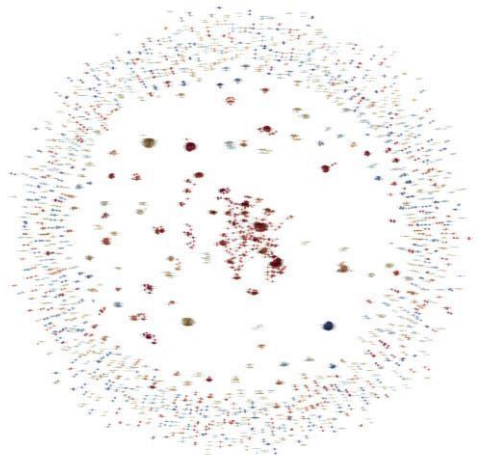
[ ] import twint
import pandas as pd
import numpy as np
import nest_asyncio
import networkx as nx
import matplotlib.pyplot as plt
nest_asyncio.apply()
  
```

Twitter Data Crawling

A. Analisis

Dari file csv tersebut diperoleh masing masing user yang menge-tweet sebagai node, beserta isi string tweetnya sebagai arah dan bobot dari vertex yang menyambung antara kedua node. Kemudian hasil dari community detection divisualisasikan dalam bentuk graf berikut

Gambar 8



Graf hasil

Kemudian dilakukan proses terakhir yaitu sorting berdasarkan jumlah individu di community tersebut

Gambar 9

harry_m51	1053
Ciaran_THFC	1053
jeongkiits	1052
ffairypixies	1052
jeannene54	1051
Lover1Trump	1051
bandanaghosh18	1050
MusicSurama	1050
dk_1_	1049
panya_bookue	1049
endrwks	1048
ranbooisfunny	1048
manckmade	1047
MindBodySoleUK	1047
ninerronnie	1046
OldSchool80s	1046
ollikainenotso	1045
sulkahattu	1045
kirjavainen	1045
sabinatraitior	1044
sbasquad	1044
anthony68915732	1043
RealJaguarKenya	1043

23 community dengan banyaknya individu tertinggi dalam memberitakan ucapan Bela Sengkawa

DAFTAR PUSTAKA

- [1] Boyd, Danah M. and Ellison Nicole B.. "Social Network Sites: Definition, History, and Scholarship". ComputerMediated 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.
- [10] Wanyun Cui, Yanghua Xiao, Haixun Wang, Yiqi Lu, and Wei Wang. 2013. Online search of overlapping communities. In Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data (SIGMOD '13). ACM, New York, NY, USA, 277-288. DOI=<https://dx.doi.org/10.1145/2463676.2463722>
- [11] Utami, Nisa Citra. Analisis dan Implementasi Community detection Menggunakan Algoritma Girvan and Newman dalam Social Network. IT Telkom. 2014