

**Analisis dan Implementasi *Betweenness Centrality* pada *Social Network* Twitter
dengan Metode *Linear Scaling* Berbasis Pengguna**
***Analysis and Implementation Betweenness Centrality in Social Network Twitter with Linear
Scaling Method Based on User***

Diska Yunita¹, Warih Maharani², Alfian Akbar Gozali³

^{1,2,3} Fakultas Informatika, Universitas Telkom

Jalan Telekomunikasi No.1, Dayeuh Kolot, Bandung 40257

diskayunita@gmail.com¹, wmaharani@gmail.com², aaz@telkomuniversity.ac.id³

Abstrak

Social Network berisi peta individu-individu dan relasi yang terjadi di antara mereka. Analisis hubungan antar individu, bagaimana hubungan tersebut terjadi, dan konsekuensinya dapat dipelajari menggunakan teknik *Social Network Analysis*. Salah satu contoh penerapannya yaitu pada *centrality measurement* yang digunakan untuk menentukan pengguna yang berpengaruh dalam penyebaran informasi. Salah satu metode yang dapat digunakan adalah *betweenness centrality*, *node* yang paling sering dilewati *shortest path* merupakan *node* yang memiliki *centrality* paling tinggi. Pada tugas akhir ini menerapkan salah satu algoritma dari *betweenness centrality*, yaitu algoritma Geisberger. Algoritma Geisberger digunakan untuk menghitung *betweenness centrality* pada graf yang berbobot dan tidak berarah dengan menggunakan metode *Linear Scaling*. Tujuannya untuk menentukan *ranking user* yang berpengaruh dalam *social media* Twitter dan yang kedua untuk mengetahui pengaruh nilai parameter *k* dalam perhitungan nilai *centrality*. Hasil pengujian menunjukkan bahwa metode *Linear Scaling* dengan dapat digunakan untuk menentukan *ranking user* yang berpengaruh dalam penyebaran informasi di Twitter. Hasil yang kedua yaitu nilai *k* berpengaruh terhadap hasil perankingan, semakin besar nilai *k* maka hasil perankingan semakin stabil. Selain nilai *k*, faktor lain yang mempengaruhi perankingan yaitu *egde* dan penghapusan *node*.

Kata kunci: *Social Network Analysis, Betweenness Centrality, Linear Scaling*

Abstract

Social Network contains a map of individuals and relationships that occur between them. Analysis of the relationship between individuals, how these relationships occur, and the consequences can be studied using the technique of Social Network Analysis. One example application is the centrality of measurement used to determine the influential users in the dissemination of information. One method that can be used is the betweenness centrality, nodes are most often bypassed the shortest path is the node that has the highest centrality. In this final project implements one of betweenness centrality algorithm, namely Geisberger algorithm. Geisberger algorithm used to calculate the betweenness centrality in the weighted and undirected graph using Linear Scaling. The goal is to determine the ranking of influential users in social media Twitter and the second to determine the influence of the value of the parameter *k* in the calculation of the value of centrality. Test results showed that the method with the Linear Scaling can be used to determine the ranking of influential users in the deployment of information on Twitter. The second result is the value of *k* affects the on ranking results, the greater the value of *k*, the more stable the ranking results. In addition to the value of *k*, the other factors that affect the ranking of that edge and deletion of nodes.

Keywords: *Social Network Analysis, Betweenness Centrality, Linear Scaling*

1. Pendahuluan

Social Network merupakan kumpulan dari individu, relasi antar individu tersebut dan juga informasi relasional [8]. *Social network* menggambarkan peta interaksi antar individu dengan cara relasi antar individu tersebut terbentuk dalam bentuk graf [1]. Dengan menggunakan teknik *Social Network Analysis* (SNA), kita dapat mempelajari relasi yang terjadi antar individu, bagaimana relasi tersebut dapat terjadi, dan konsekuensi dari relasi tersebut. Salah satu bentuk penerapan SNA yaitu pada *centrality measurement* atau perhitungan *centrality* [8].

Centrality measurement dilakukan untuk menentukan pengguna yang memiliki pengaruh paling besar dalam penyebaran informasi. Metode yang dapat digunakan dalam perhitungan *centrality* antara lain *degree centrality*, *closeness centrality* dan *betweenness centrality* [8]. Pada tugas akhir ini akan menggunakan metode *betweenness centrality*. Berdasarkan metode *betweenness centrality* pengguna atau *node* yang paling sering dilewati *shortest path* merupakan pengguna yang paling berpengaruh dalam penyebaran informasi [3]. Namun, perhitungan *betweenness* biasanya menggunakan graf tidak berbobot. Sehingga tidak memperhitungkan besarnya bobot pada relasi antar individu.

Masalah tersebut dapat diatasi dengan menggunakan menggunakan metode *Linear Scaling*, dimana dalam perhitungan *centrality* dipengaruhi oleh jarak *node* tersebut yang merupakan bobot pada relasi antar *user*. [3] Oleh karena itu pada tugas akhir ini menggunakan metode *Linear Scaling* dalam perhitungan *betweenness centrality* untuk mengetahui *user* yang paling berpengaruh pada jejaring Twitter.

2. Landasan Teori

2.1 Social Media

Social media atau media sosial didefinisikan sebagai kelompok aplikasi berbasis internet yang dibangun menggunakan teknologi Web 2.0 yang dapat digunakan penggunaannya untuk membuat konten dan bertukar konten antar sesama pengguna [5].

Media sosial memberi kemudahan kepada pengguna untuk berkomunikasi dan berhubungan dengan satu sama lain. Sehingga popularitas media sosial terus berkembang pesat. Pada Desember 2011, tercatat ada 845 juta pengguna media sosial [4].

2.2 Social Network

Social Network atau jejaring sosial adalah suatu struktur sosial yang terdiri dari individu (atau organisasi) yang disebut “*node*”, yang terhubung oleh satu atau lebih relasi [10]. Sebuah jejaring sosial terjadi ketika semua aktor memiliki hubungan dengan aktor lainnya. Fitur yang terdapat dalam jejaring sosial yaitu adanya informasi relasional. Oleh karena itu, jejaring sosial lebih fokus pada informasi relasional daripada atribut actor [12].

Jejaring sosial menjadi sangat penting karena berisi peta dari individu-individu dan cara-cara mereka berhubungan satu sama lain. Jejaring sosial merupakan layanan berbasis web yang memungkinkan individu untuk membangun profil publik dalam sistem yang dibatasi, mengetahui daftar pengguna dan dengan siapa pengguna tersebut berhubungan, dan melihat dan melintasi daftar koneksi pengguna tersebut dan yang koneksi yang dibuat pengguna lain dalam sistem [1].

2.2.1 Social Network Twitter

Twitter merupakan microblogging dan layanan jejaring *social* yang memungkinkan pengguna untuk “*follow*” satu sama lain. Pengguna juga dapat mengirim dan menerima “*tweet*” yang merupakan pesan teks terdiri dari maksimal 140 karakter. Layanan twitter juga digunakan oleh berbagai organisasi untuk memasarkan produk.

Terdapat istilah relasi yang terdapat di Twitter yaitu, *follower* dan *following*. *Follower* merupakan pengguna yang mengirim permintaan pertemanan kepada kita. Sedangkan *following* merupakan pengguna yang kita kirim permintaan pertemanan. *Tweet* yang kita bagikan hanya dapat dilihat oleh *follower* kita saja.

2.3 Social Network Analysis

Social Network Analysis (SNA) adalah sebuah studi yang mempelajari tentang hubungan manusia dengan memanfaatkan teori graf [11]. Menurut Scott [8], SNA adalah sekumpulan metode yang digunakan untuk menginvestigasi aspek relasi pada struktur data. SNA merupakan metode yang digunakan untuk mengetahui hubungan informal antar individu dengan menganalisa struktur dari hubungan sosial dalam suatu kelompok.

Pada *social network*, individu atau orang digambarkan sebagai *node* atau titik. Sedangkan relasi yang terjadi antar individu digambarkan dengan *edge* atau *link*. Pada dasarnya sebuah jaringan sosial adalah

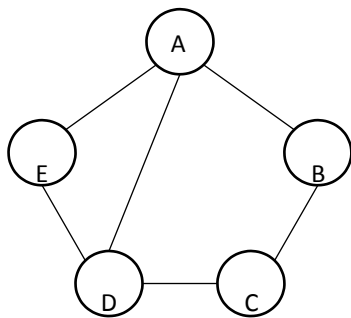
sebuah peta yang terdiri atas banyak orang dimana didalamnya terdapat relasi antar individunya [10].

2.3.1 Graf

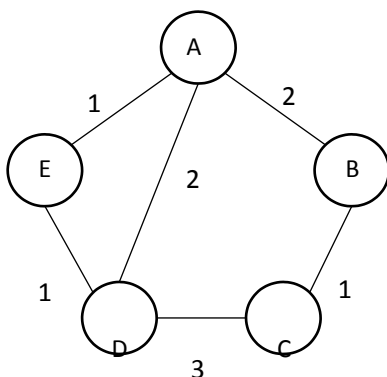
Graf adalah salah satu cara untuk menggambarkan relasi antar pengguna pada struktur jejaring sosial. Graf merupakan himpunan (V, E), V merupakan himpunan dari simpul-simpul (*vertex* atau *node*) dan E merupakan himpunan sisi (*edges*) yang menghubungkan sepasang simpul [7]. Terdapat 2 jenis representasi dari graf, yaitu:

a. Graf Berdasarkan Bobot

Berdasarkan bobot, graf dapat dikelompokkan menjadi dua macam, yaitu graf berbobot dan graf tidak berbobot. Bobot disini dapat direpresentasikan sebagai jumlah interaksi, kekuatan hubungan, jarak suatu *node*, atau yang lainnya. Sedangkan graf tidak berbobot hanya merepresentasikan suatu hubungan antar *node*-nya saja.



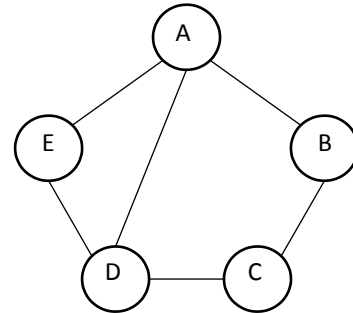
Gambar 2. 1 Graf tidak berbobot



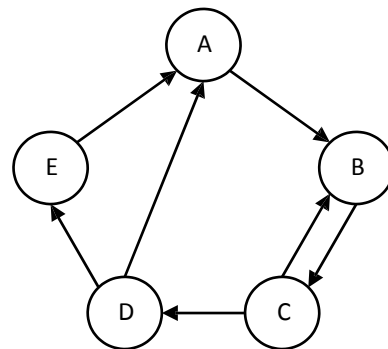
Gambar 2. 2 Graf berbobot

b. Graf Berdasarkan Arah

Graf berdasarkan arah dapat di kelompokkan menjadi 2 macam, yaitu graf berarah dan graf tidak berarah. Graf berarah tersebut merepresentasikan arah relasi yang terjadi antar *node*.



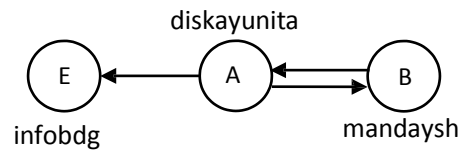
Gambar 2. 3 Graf tidak berarah



Gambar 2. 4 Graf berarah

2.3.2 Representasi Graf pada twitter

Graf pada Twitter terdiri dari *node* dan *edge*. *Node* merepresentasikan pengguna Twitter dan *edge* merepresentasikan hubungan pertemanan antar penggunanya. Graf yang digunakan pada Twitter yaitu graf tidak berarah dan berbobot (*undirected-weighted graph*). Arah pada graf tersebut digunakan untuk mengetahui *node* atau pengguna yang berlaku sebagai *follower* dan *following*. Dari graf berarah tersebut dapat diketahui pula alur informasi, karena informasi hanya diterima oleh *follower* suatu *node* atau pengguna.



Gambar 2.2 Graf pada Twitter

‘diskayunita’ merupakan *follower* pengguna ‘infobdg’ dan ‘mandaysh’. Atau dapat juga dikatakan bahwa pengguna ‘infobdg’ dan ‘mandaysh’ merupakan *following* dari pengguna ‘diskayunita’. Informasi yang berupa *tweet* dari pengguna ‘diskayunita’ hanya dapat dilihat oleh pengguna ‘mandaysh’ saja, karena

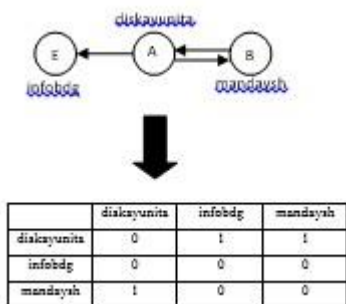
pengguna 'infobdg' bukan *follower* dari pengguna 'diskayunita'.

2.3.3 Matriks Adjacency

Jenis matriks yang biasa digunakan dalam jejaring sosial adalah matriks adjacency. Matriks tersebut menggambarkan tabel yang baris dan kolomnya menunjukkan relasi antar *node*. Pengukuran dilakukan berdasarkan jumlah *node* (n). Sebuah Matriks Adjacency A (G) = [a_{ij}] dari graf G = (V,E) dimana V = { V_1, \dots, V_n } didefinisikan dengan[7]:

$$a_{ij} = \begin{cases} 1, & \text{jika } (V_i, V_j) \in E(G) \\ 0, & \text{jika } (V_i, V_j) \notin E(G) \end{cases} \quad (2.1)$$

Dalam Matriks Adjacency 1 merepresentasikan hubungan, artinya jika *node* A dan *node* B bernilai 1 maka *node* A memiliki hubungan pertemanan dengan *node* B.



Gambar 2. 6 Pemodelan Graf dalam Matriks Adjacency

Berdasarkan Gambar 2.6 menunjukkan bahwa graf dapat dimodelkan dalam bentuk matriks *adjacency*, dimana *node* 'diskayunita' menjadi *follower* 'infobdg' dan 'mandaysh' maka pada matriks (diskayunita, infobdg) dan (diskayunita, mandaysh) memiliki nilai 1.

2.4 Centrality Measurement

Dalam teori graf dan analisa jaringan, ada berbagai ukuran sentralitas simpul dalam grafik yang menentukan tingkat kepentingan relatif dari simpul dalam sebuah graf (misalnya seberapa pentingnya seseorang dalam jaringan sosial) [8].

2.3 Betweenness Centrality

Betweenness centrality adalah perhitungan centrality dalam graf yang fokus pada hubungan suatu node dengan node lainnya. Centrality pada sebuah node bergantung pada sejauh mana ia diperlukan sebagai penghubung dalam penyebaran informasi di social network. Node yang paling sering dilewati shortest paths maka node tersebut memiliki nilai centrality yang tinggi [8].

Untuk sebuah graf G= (V, E), rumus perhitungan centrality dengan menggunakan metode Betweenness pada suatu node v [8] yaitu:

$$C_B(i) = \sum_{s \neq v} \sum_{v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (2.2)$$

Keterangan:

- $\sigma_{st}(v)$ = jumlah shortest paths dari node s ke t yang melewati node v
- σ_{st} = jumlah shortest paths dari node s ke t

2.6 Metode Linear Scaling

Geisberger membuat variasi baru dalam perhitungan *betweenness centrality* untuk graf berbobot. Metode ini merupakan pengembangan dari algoritma Brandes, inti dari pengembangan metode ini adalah mengkalikan dengan nilai $\frac{1}{\mu(s,w)}$ pada perhitungan pair-dependency [3].

$$\delta_s(v) = \sum_{w \in succ(v)} \frac{\mu(s,v)}{\mu(s,w)} \cdot \frac{\sigma_{sv}}{\sigma_{sw}} (1 + \delta(w)) \quad (2.3)$$

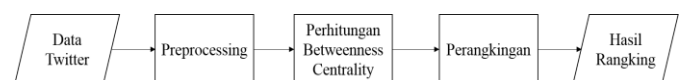
(2.3) Keterangan:

- $\mu(s, w)$ = jarak terpendek dari *node* s ke w
- $\mu(s, v)$ = jarak terpendek dari *node* s ke v

3. Perancangan dan Implementasi

3.1 Gambaran Umum Sistem

Rancangan sistem untuk menentukan rangking user yang berpengaruh dalam penyebaran informasi di Twitter dengan menggunakan metode Betweenness Centrality, secara umum dapat digambarkan sebagai berikut:



Gambar 3.1 Rancangan sistem

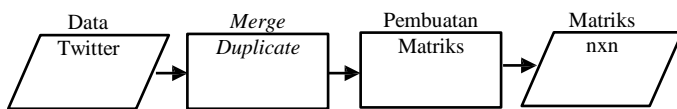
Data yang akan diinputkan kedalam sistem adalah *dataset* yang diperoleh dari *social media* Twitter yang digunakan menggunakan perangkat lunak *NodeXL*. *Data set* yang diambil terdiri dari nama-nama pengguna Twitter dan relasi yang terjadi antar pengguna tersebut yaitu *following/followed*, *mention*, *reply*, dan *retweet*.

Setelah *data set* diperoleh, selanjutnya akan dilakukan *preprocessing* terhadap datatersebut. *Data set* akan melalui proses *merge duplicate* untuk menghilangkan data yang sama. Kemudian data set direpresentasikan dalam bentuk matriks nxn, dengan n merupakan jumlah *node* yang terambil.

Pada Tugas Akhir ini akan dilakukan perhitungan *betweenness centrality* untuk menentukan pengguna yang paling berpengaruh dalam penyebaran informasi. Hasil perhitungan tersebut akan menampilkan daftar ranking pengguna berdasarkan pengaruhnya dalam penyebaran informasi di jejaring sosial Twitter.

3.2 Preprocessing

Data Twitter merupakan data set yang diperoleh dari *social network* Twitter yang diambil menggunakan perangkat lunak *NodeXL*. *Data set* terdiri atas nama-nama pengguna Twitter dan data relasi antar pengguna yang telah ditentukan, yaitu *following/followed*, *mention*, *reply*, dan *retweet*.



Gambar 3. 2 Proses preprocessing

Preprocessing dilakukan sebelum *data set* memasuki proses inti. Berdasarkan gambar 3.2 dapat dilihat alur yang terjadi saat *preprocessing*. Pada sistem ini, preprocessing yang dilakukan adalah *merge duplicate* dan merepresentasikan pengguna dan relasi yang terjadi antar pengguna dalam bentuk matriks nxn dengan n adalah jumlah *node* yang terambil.

Matriks 1	Matriks 2	Color	Width	Style	Opacity	Visibility	Label	Label Text	Label Font	Add Your Own Columns Here	Relationship	Edge Weight	
dhsaaendri01	diskayunita										Followed	4/10/2014 18:34	1
band_silverband	diskayunita										Followed	4/10/2014 18:34	1
egewirayudha	diskayunita										Followed	4/10/2014 18:34	1
Jacquelineam	diskayunita										Followed	4/10/2014 18:34	1
fathbesar	diskayunita										Followed	4/10/2014 18:34	1
asgan_forester	diskayunita										Followed	4/10/2014 18:34	1
bokadarmawan	diskayunita										Followed	4/10/2014 18:34	1
diskayunita	bokadarmawan										Followed	4/10/2014 18:34	1
diskayunita	kitsoofficial										Followed	4/10/2014 18:34	1
diskayunita	medidtarani										Followed	4/10/2014 18:34	1
nyimeng	diskayunita										Followed	4/10/2014 18:34	1
diskayunita	nyimeng										Followed	4/10/2014 18:34	1
diskayunita	gusbaipoos										Followed	4/10/2014 18:34	1
denshiday	diskayunita										Followed	4/10/2014 18:34	1
diskayunita	denshiday										Followed	4/10/2014 18:34	1
diskayunita	vian_varian										Followed	4/10/2014 18:34	1
vian_varian	diskayunita										Followed	4/10/2014 18:34	1
ilmymunull	elly_husie										Followed	4/10/2014 18:34	1
ilmymunull	diskayunita										Followed	4/10/2014 18:34	1
ilmymunull	ilmymunull										Followed	4/10/2014 18:34	1

Gambar 3.3 Hasil Tabel Relasi

3.2.1 Merge Duplicate

Merge duplicate dilakukan untuk menghilangkan data yang sama pada data set seperti pada gambar 3.3. Misalnya, pengguna A melakukan *following* terhadap pengguna B merupakan data yang sama dengan pengguna A menjadi *follower* pengguna B. hanya diperlukan 1 relasi saja yang dipilih untuk menjadi *data set*.

3.2.2 Pembuatan Matriks

Dari data set tersebut kemudian direpresentasikan dalam bentuk matriks nxn dengan n adalah jumlah *node* yang terambil untuk di analisis. Gambar 3.4 menggambarkan tabel berisi daftar pengguna yang terambil serta relasi yang terjadi. Relasi direpresentasikan dengan bilangan biner, yang artinya apabila terjadi relasi maka kolom tersebut bernilai 1, namun apabila tidak terjadi relasi kolom tersebut berisi 0.

Demetrius	passio	bel	ust	alkafz	aguparja	SRUDU0FI	andi	CCB	andi	andi	run	Heryalam	iklanonn	funkjomp	diskayyoda	ridwanakar	basuki	lit	soebandur	TRANSTV	datikom	ruswaninuz	kata2jaja	infoblog
Demetrius	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
passio	0	1	0	2	2	1	2	1	2	1	2	1	2	1	2	1	1	1	1	2	1	2	1	2
bel	0	1	0	2	2	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ust	0	1	0	2	2	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
alkafz	0	2	2	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
aguparja	0	2	2	2	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SRUDU0FI	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
andi	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
andi_CCB	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
andi_run	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Heryalam	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
iklanonn	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
funkjomp	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
diskayyoda	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
ridwanakar	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	0
basuki	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0
lit	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
soebandur	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
TRANSTV	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	2	0	0
datikom	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	2	0	1
ruswaninuz	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	1	0	1	0	0
kata2jaja	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0
infoblog	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0

Gambar 3. 4 Representasi Matriks nxn

3.3 Perhitungan Centrality

Mencari *shortest path* dan menentukan *node* yang paling sering dilewati *shortest path*. Kemudian menghitung *betweenness centrality* menggunakan metode *Linear Scaling*.

3.3.1 Perhitungan Jumlah Shortest Path

Perhitungan *shortest path* dalam teori graf menggunakan algoritma Dijkstra untuk graf berbobot

(*weighted graph*), yaitu mencari bobot paling minimum yang menghubungkan sebuah simpul yang dipilih dengan *node* lainnya[2].

3.3.2 Perhitungan *Pair-Dependency*

Pair-Dependency adalah nilai kontribusi dari semua shortest path yang dimulai dari setiap *node* dari sebuah graf secara bergiliran. Pada tiga *nodes* $s, t, v \in V$, *Pair-Dependency* dari *node* s dan t pada v sebagai *node* yang dilalui shortest path dari s ke t [3].

$$\delta_s(v) = \sum_{w \in \text{succ}(v)} \frac{\mu(s,v) \cdot \sigma_{sv}}{\mu(s,w) \cdot \sigma_{sw}} (1 + \sigma_s(w)) \quad (3.3)$$

Nilai $\delta_s(v)$ merupakan *Pair-Dependency* dari *nodes* $s, t, v \in V$ pada *node* perantara yaitu $v \in V$. Dan $\mu(s,w)$ adalah jarak terpendek dari *node* s ke w , serta $\mu(s,v)$ adalah jarak terpendek dari *node* s ke v

3.3.3 Perhitungan Nilai *Betweenness Centrality*

Nilai *betweenness centrality* dari *node* v diperoleh dari jumlah *Pair-Dependency* setiap pasangan *node* pada v [3].

$$C_B(v) = \sum_{s \neq v} \sum_{t \neq v} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (3.4)$$

Dari hasil perhitungan diperoleh nilai C_B tiap *node*, semakin tinggi nilai C_B maka *node* tersebut memiliki pengaruh yang besar dalam penyebaran informasi di media social Twitter.

3.4 Perangkingan

Proses perangkingan dilakukan berdasarkan pengguna yang memiliki nilai *centrality* paling tinggi, *node* yang memiliki *centrality* paling tinggi merupakan *node* yang paling berpengaruh dalam penyebaran informasi.

3.5 Kebutuhan Sistem

Dalam pembuatan sistem hal yang dibutuhkan yaitu,

3.5.1 Perangkat lunak

Perangkat lunak yang digunakan dalam pembuatan sistem ini adalah

1. Sistem operasi Windows 8

2. Netbeans IDE 7.3
3. NodeXL

3.5.2 Perangkat keras

Perangkat keras yang digunakan dalam pembuatan sistem ini adalah

1. Processor Intel Core i3
2. RAM 4GB
3. VGA 2GB

4. Pengujian dan Analisis

4.1 Tujuan Pengujian

Pengujian yang dilakukan dalam Tugas Akhir ini bertujuan untuk:

1. Menganalisis pengaruh perubahan parameter nilai hasil perangkingan.
2. Mengetahui ranking user yang berpengaruh dalam penyebaran informasi di Twitter.
3. Mengetahui pengaruh pemotongan relasi dan penghapusan *node* terhadap hasil perangkingan.

4.2 Data Set

Data set yang digunakan dalam Tugas Akhir ini berasal dari social media Twitter yang diambil dari pengguna Twitter @pasco_bdg. Data berupa jumlah *following* dan *followers* diambil menggunakan bantuan perangkat keras *NodeXL*. Sedangkan data berupa jumlah *mention*, *reply*, dan *retweet* diambil menggunakan bantuan situs <http://www.conweets.com/>. Dari hasil pengambilan data tersebut diperoleh jumlah *user* sebanyak 400 dan relasi sebanyak 4956.

4.3 Skenario Pengujian

4.3.1 Skenario pengujian nilai pivot (k)

Berdasarkan tujuan yang ingin dicapai, maka alur skenario pengujian yang akan dilakukan sama dengan perancangan sistem pada Gambar 3.1. Namun, pada pengujian ini dilakukan perubahan terhadap nilai k sebelum dilakukan proses perhitungan *betweenness centrality*. k menunjukkan jumlah *node* yang dijadikan sumber untuk menghitung shortest path dan *betweenness centrality*, k yang diuji yaitu 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150.

4.3.2 Skenario Pemotongan Relasi (Edge)

Pada pengujian ini, dilakukan pengelompokan node berdasarkan degree menjadi 3 kelompok dengan range 132. Kemudian dilakukan pengujian dengan 3 skenario sebagai berikut:

1. Menghilangkan relasi(edge) dari node kelompok 1
 2. Menghilangkan relasi(edge) dari node kelompok 2
 3. Menghilangkan relasi(edge) dari node kelompok 3
- Analisis Pengaruh Nilai k Dalam Perhitungan Betweenness Centrality.*

4.3.3 Skenario Penghapusan Node

Pada pengujian ini, dilakukan penghapusan terhadap node. Node yang dihapus merupakan node yang menduduki peringkat pertama pada hasil perangkingan. Pengujian ini dilakukan untuk mengetahui pengaruh suatu node pada hasil perangkingan.

4.4 Analisis Pengaruh Nilai k Dalam Perhitungan Betweenness Centrality

Pada proses pengujian ini dilakukan perubahan terhadap nilai k sebanyak 20 kali. Dilakukan analisis pengaruh nilai k terhadap hasil perangkingan.

4.4.1 Analisis Pengaruh Nilai k Terhadap Hasil Perangkingan

Pada skenario ini dilakukan pengujian sebanyak 20 kali dimana masing-masing pengujian dilakukan sebanyak 2 kali. Berikut adalah 3 tabel hasil pengujian yang dapat mewakili seluruh pengujian. Sedangkan tabel lainnya dapat dilihat pada lampiran buku Tugas Akhir ini.

Tabel 4.1 Pengaruh nilai k=1 terhadap hasil perangkingan

K=1		
Rank	Uji 1	Uji 2
1	SteakRanjang	Mahasiswa_bdg
2	KFCINDONESIA	SindikataKuliner
3	netmediatama	Niken_Fer
4	mancung80	bragaculinary
5	Resto_Guide	OMGCulinary
6	abdi_ons	lewisjaya50
7	YeahBandung	jakarta_food
8	Metro_TV	pasco_bdg
9	bragaculinary	ruswaninaxpjn
10	NyiMistik	mampirkotasolo

Berdasarkan Tabel 4.1 dapat dilihat bahwa dengan nilai k=1 perangkingan pada uji 1 dan uji 2 banyak berubah. Hal tersebut disebabkan oleh nilai k yaitu jumlah node yang dijadikan sumber untuk melakukan perhitungan betweenness centrality. Karena k yang dipilih adalah satu yang artinya pencarian shortest test path hanya dilakukan dari satu sumber saja. Pada setiap k, node sumber dipilih secara acak, sehingga setiap pengujian dilakukan perhitungan dari node sumber berbeda dengan jalur shortest path yang berbeda pula.

Skenario Penghapusan Node

Analisis Pengaruh Nilai k Dalam P

Tabel 4.2 Pengaruh nilai k=9 terhadap hasil perangkingan

K=9		
Rank	Uji 1	Uji 2
1	pasco_bdg	pasco_bdg
2	PemkotBandung	PemkotBandung
3	kata2bijak	kata2bijak
4	Mahasiswa_bdg	Mahasiswa_bdg
5	GiraffeKs	GiraffeKs
6	Jakarta_Kuliner	Jakarta_Kuliner
7	bragaculinary	detikcom
8	detikcom	Jajandibdg
9	Jajandibdg	SteakRanjang
10	winnyprtwi	winnyprtwi

Berdasarkan Tabel 4.2 dapat dilihat bahwa hasil perangkingan yang didapat dari pegujian 1 dan 2 lebih stabil. Posisi rangking 1 hingga 6 dan rangking 10 diduduki oleh *node* yang sama. Dengan nilai k yaitu 9, yang artinya dipilih 9 *node* secara acak untuk dijadikan sumber sudah cukup untuk melakukan perhitungan betweenness centrality. Pada penelitian ini, nilai k=9 merupakan hasil pendekatan yang terbaik pada studi kasus ini.

Tabel 4.3 Pengaruh nilai k=40 terhadap hasil perangkingan

K=40		
Rank	Uji 1	Uji 2
1	ridwankamil	solusijerawatan
2	pasco_bdg	pasco_bdg
3	kata2bijak	kata2bijak
4	detikcom	OfriendshipRep
5	PVJbandung	ridwankamil
6	Metro_TV	GiraffeKs
7	Kuliner_Bandung	Jakarta_Kuliner
8	OsengMercon	detikcom
9	TRANSTV_CORP	NasgorKlinik
10	InfinitoKuliner	adndamna

Berdasarkan Tabel 4.3 dapat dilihat bahwa dengan nilai k=40 maka hasil perangkingan pada pengujian 1 dan 2 lebih stabil dibanding hasil pada

table 4.1 namun tidak lebih stabil dari table 4.2. *Node* 'pasco_bdg' dan 'kata2bijak' yang tidak mengalami perubahan rangking, hal tersebut dapat terjadi karena dua hal. Pertama karena nilai *centrality* kedua *node* yang terpaut jauh dari *node* lainnya sehingga pada 2 kali pengujian *node* tersebut tetap menduduki peringkat 10 besar. kemungkinan kedua yaitu karena semua *node* memiliki relasi dengan *node* 'pasco_bdg' sehingga *node* 'pasco_bdg' merupakan penghubung yang kuat dalam penyebaran informasi dan memiliki nilai *centrality* yang besar.

Berdasarkan Tabel 4.1, 4.2, dan 4.3 dapat disimpulkan bahwa nilai k berpengaruh terhadap perangkingan, dimana semakin besar k maka semakin stabil hasil perangkingan yang didapat. Perubahan hasil perangkingan tersebut disebabkan karena semakin besar nilai k maka perhitungan *centrality* dihasilkan dari banyak *node* sumber dan semakin banyak melalui berbagai jalur shortest path yang ada. Sehingga hasil perangkingan pun semakin stabil.

4.4.1 Analisis Pengaruh Pemotongan Relasi Terhadap Hasil Perangkingan

Pada pengujian kedua ini dilakukan 3 skenario yang telah dijelaskan sebelumnya. Berikut adalah 3 tabel hasil pengujian berikut

Tabel 4.4 Hasil Pengujian Skenario 1

rank	<i>user</i>
1	pasco_bdg
2	Kuliner_Bandung
3	ridwankamil
4	infobdg
5	deliveryboy3

Dari hasil pengujian skenario pertama, dapat dilihat bahwa hasil perangkingan hanya terdapat 5 *user*. Hal tersebut terjadi karena relasi pada *node* kelompok 1 dihilangkan, sedangkan *node* pada kelompok 1 memiliki jumlah *node* yang banyak. Maka dari itu ketika dihilangkan, *node* yang memiliki nilai *centrality* hanya 5. Sedangkan *node* lain memiliki nilai *centrality* nol(0).

Tabel 4.5 Hasil Pengujian Skenario 2

rank	user
1	pasco_bdg
2	info_wanitaPlus
3	basuki_btp
4	kulinerjakarta
5	mampirkotasolo
6	infobdgkuliner
7	tamanmusikbdg
8	VirgindrinkID
9	bragaculinary
10	NyiMistik

Dari hasil pengujian skenario kedua, diperoleh 10 peringkat *node* dengan *centrality* terbesar. Akun @pasco_bdg menduduki peringkat pertama juga seperti pada skenario pertama. Hal tersebut menunjukkan bahwa nilai *centrality* @pasco_bdg terlampaui tinggi dari *node* lain sehingga ketika relasi dengan *node* kelompok 2 dihilangkan tidak mempengaruhi peringkatnya.

Tabel 4.6 Hasil Pengujian Skenario 3

rank	user
1	basuki_btp
2	Kuliner_Bandung
3	info_wanitaPlus
4	faskahtri
5	mampirkotasolo
6	detikcom
7	eva4nur
8	NomnomEatery
9	kata2bijak
10	Mochilok

Pengujian skenario penunjukan perubahan pada peringkat pertama. Hal tersebut disebabkan karena @pasco_bdg merupakan salah satu *node* kelompok 3 yang dipotong relasinya pada skenario ketiga. Namun, akun @basuki_btp, @info_wanitaplus, dan @mampirkotasolo masuk ke dalam 10 *node* dengan *centrality* terbesar. Hal tersebut menunjukkan akun tersebut memiliki relasi yang banyak pada *node* kelompok 1 dan kelompok 2.

4.4.2 Analisis Pengaruh Penghapusan Node Terhadap Hasil Perangkingan

Pada pengujian ini dilakukan 3 (tiga) skenario pengujian dimana masing masing pengujian melakukan penghapusan node yang menduduki peringkat pertama pada hasil perangkingan.

Rank	User	Centrality
1	pasco_bdg	15,426
2	Kuliner_Bandung	10,666
3	infobdgkuliner	10,666
4	GNFI	10,666
5	InfinitoKuliner	8,266
6	ridwankamil	6,32
7	arianjarw	4,666
8	WinnieH	6,400
9	WaroengNgemil	5,520
10	FitriTresnayati	4,400

Gambar 4.1 Hasil Perangkingan Data Set Lengkap

Gambar 4.1 merupakan hasil perangkingan dari pengujian sebelumnya dengan nilai $k=9$. Berdasarkan hasil tersebut, diperoleh *node* 'pasco_bdg' yang menempati peringkat pertama. Pada skenario pertama, *node* 'pasco_bdg' dihilangkan dari data set.

Rank	User	Centrality
1	ridwankamil	274.426
2	Kuliner_Bandung	46.440
3	detikcom	43.301
4	TRANSTV_CORP	32.035
5	BandungView	29.971
6	basuki_btp	25.112
7	MieTegallega	21.577
8	promo1jkt	21.027
9	Kedai_Pasco	18.576
10	infobdgkuliner	17.028

Gambar 4.1 Hasil Perangkingan Setelah Penghapusan Akun @pasco_bdg

Gambar 4.2 merupakan hasil perangkingan setelah *node* 'pasco_bdg' dihilangkan dari data set. Hasil tersebut menunjukkan bahwa hasil perangkingan berubah. Namun hanya *node* 'kuliner_bandung' saja yang tetap menduduki peringkat 2. Hal tersebut disebabkan karena *node* 'kuliner_bandung' merupakan perantara dari banyak *node* lainnya seperti *node* 'pasco_bdg'. Oleh karena itu, ketika *node* 'pasco_bdg' dihilangkan, *node* 'kuliner_bandung' tetap memiliki nilai *centrality* yang besar. Walaupun muncul *node* 'ridwankamil' yang menjadi peringkat pertama sebagai pengganti *node* 'pasco_bdg'.

Rank	User	Centrality
1	info_wanitaPlus	6.400
2	NyiMistik	1.536
3	wrgRica	1.536
4	icepasco_BDG	1.536
5	RumahSebelahID	896
6	MerryRiana	768
7	Kang_Richi	768
8	RAMENCEMEN	640
9	iklannonstop24	576
10	happy2travel2	576

Gambar 4.2 Hasil Perangkingan Setelah Penghapusan Akun @ridwankamil

Pada skenario kedua, dilakukan penghapusan *node* 'ridwankamil' yang menduduki peringkat pertama di skenario sebelumnya. Gambar 4.3 menunjukkan hasil perangkingan yang didapat. Dari hasil tersebut menunjukkan perubahan pada seluruh hasil perangkingan. Hal tersebut terjadi karena *node* pada perangkingan sebelumnya memiliki ketergantungan yang besar terhadap *node* 'ridwankamil', sehingga ketika *node* tersebut dihilangkan dari data set nilai *centrality node* lainnya turun drastis dan tidak masuk 10 peringkat *node* dengan nilai *centrality* terbesar. Dan muncul *node* 'info_wanitaPlus' sebagai peringkat pertama.

Rank	User	Centrality
1	Mahasiswa_bdg	1.809
2	mediabandoeng	1.809
3	winnyptrwi	1.809
4	FruitzeeBandung	1.206
5	infoHobbies	1.206
6	nullynuffie	1.206
7	SteakRanjang	1.206
8	Nasgorklinik	1.206
9	OfriendshipRep	1.206
10	MerryRiana	603

Gambar 4.3 Hasil Perangkingan Setelah Penghapusan Akun @info_wanitaplus

Pada skenario ketiga, perubahan data set dilakukan dengan menghapus *node* 'info_wanitaPlus' yang sebelumnya menduduki peringkat pertama. Dengan perubahan tersebut terjadi juga perubahan seluruh hasil perangkingan seperti yang terjadi pada skenario dua.

Berdasarkan Gambar 4.2, 4.3, dan 4.4 dapat disimpulkan bahwa penghapusan *node* sangat berpengaruh terhadap perangkingan, dimana setiap perubahan yang dilakukan pada dataset akan mengubah sebagian besar hasil perangkingan. Perubahan tersebut terjadi karena setiap *node* saling

memiliki ketergantungan yang tinggi dengan *node* lainnya. Sehingga dengan hilangnya satu *node* akan berpengaruh besar terhadap hasil perangkingan.

4.5 Hasil Pengujian

Dari analisis sebelumnya berdasarkan hasil perangkingan dalam mengambil nilai *k* yang digunakan, maka pada pengujian ini nilai *k* yang digunakan adalah 9, dimana *ranking* yang dihasilkan cukup stabil. Pada data Twitter ini dilakukan pengujian sebanyak 30 kali kemudian dirata-ratakan untuk mendapat 10 besar user yang memiliki nilai *centrality* tertinggi. Berikut ditampilkan nilai *centrality*:

Tabel 4.7 Hasil Rata-Rata 10 Besar pada Data Twitter

Rank	Node	Centrality Rata-Rata
1	pasco_bdg	657659,2
2	PemkotBandung	10109
3	kata2bijak	6732
4	Mahasiswa_bdg	5610
5	GiraffeKs	5049
6	Jakarta_Kuliner	4488
7	bragaculinary	4224
8	detikcom	3986,4
9	SteakRanjang	2970
10	Jajandibdg	2805

5. Kesimpulan dan Saran

5.1 Kesimpulan

Adapun beberapa kesimpulan yang dihasilkan dari pengujian dan analisis yang dilakukan, antara lain:

1. Perhitungan *centrality* dipengaruhi oleh nilai *k*. Setiap perubahan nilai *k* perangkingan *user* pun berubah, semakin besar nilai *k* atau mendekati jumlah *node* hasil perangkingan semakin stabil. Parameter nilai *k* terbaik pada Tugas Akhir ini adalah 9, dimana hasil perangkingan cukup stabil.
2. Relasi atau *Edge* yang paling mempengaruhi perangkingan yaitu relasi pada *node* yang memiliki *degree* < 132. Karena memiliki jumlah *node* yang termasuk kelompok 1 paling banyak.

3. Penghapusan *node* mempengaruhi hasil perangkangan. Setiap penghapusan *node* yang menduduki peringkat pertama akan mengubah hasil perangkangan pada perhitungan berikutnya.
4. Metode *Linear Scaling* dapat diterapkan pada data Twitter untuk menentukan *ranking user* yang berpengaruh terhadap penyebaran informasi.

5.2 Saran

Adapun saran untuk penelitian selanjutnya mengenai *Social Network Analysis* ataupun *Centrality Measurement* adalah sebagai berikut:

1. Menambahkan analisis akurasi metode *Linear Scaling*, misalkan dengan kuisioner atau *survey*.
2. Menggunakan algoritma selain Dijkstra dalam pencarian *shortest path*, misalnya Greedy.
3. Menganalisis pengaruh karakteristik data, misalnya dengan keanekaragaman jumlah data atau keanekaragaman pembobotan yang dilakukan.

6. Daftar Pustaka

- [1] Boyd, Danah M. and Ellison Nicole B.. "Social Network Sites: Definition, History, and Scholarship". *Computer-Mediated Communication*. 2008.
- [2] Brandes U.. "A Faster Algorithm for *Betweenness Centrality*". University of Konstanz. Germany. 2001.
- [3] Geisberger Robert, et al. "Better Approximation of *Betweenness Centrality*". Institute for Theoretical Computer Science. Algorithmics II Universität Karlsruhe (TH). Germany
- [4] Gundecha Pritam and Liu Huan. "Mining Social Media: A Brief Introduction". *Informis*. Arizona State University. Tempe. Arizona. 2012.
- [5] Kaplan Andreas M. and Haenlein M. "Users of the world, unite! The challenges and opportunities of Social Media". Kelley School of Business. Indiana University. France. 2010.
- [6] Mayfield Antony. "What is Social Media?". iCrossing.
- [7] Ruohonen Keijo. *Graph Theory*. 2013.
- [8] Scott John. "Social Network Analysis Theory and Application". 2011.
- [9] Smith Marc A.. "Analyzing (Social Media) Networks with *NodeXL*". University park. Pennsylvania, USA. 2009
- [10] Susanto Budi, et al. "Penerapan *Social Network Analysis* dalam Penentuan *Centrality*". Fakultas Teknologi Informasi. Universitas Kristen Duta Wacana. Yogyakarta.
- [11] Tsevetovot Maksim and Kouznetsov Alexander. "Social Network Analysis for Startups". O'Reilly. 2011.
- [12] Wasserman Stanley and Faust Katherine. "Social Network Analysis : Methods and Applications". United State. America. Cambridge University Press. 1994.