

ANALISIS DAN IMPLEMENTASI *PROBABILISTIC PARTNERSHIP INDEX (PPI)* PADA *LAPLACIAN CENTRALITY* DALAM *SOCIAL NETWORK ANALYSIS*

ANALYSIS AND IMPLEMENTATION OF *PROBABILISTIC PARTNERSHIP INDEX (PPI)* ON *LAPLACIAN CENTRALITY* IN *SOCIAL NETWORK ANALYSIS*

Muliadi Angga Wicaksono¹, Kemas Rahmat S. W., S.T., M.Eng.², Alfian Akbar G., S.T., M.T.³

Fakultas Informatika, Universitas Telkom, Bandung

¹anggawicaksono@live.com, ²bagindokemas@telkomuniversity.ac.id, ³alfian@tass.telkomuniversity.ac.id

Abstrak

Social network digunakan sebagai media penghubung antar individu untuk dapat saling berkomunikasi tak terbatas oleh jarak dan waktu. Salah satu *social network* yang populer saat ini adalah *Twitter*. Adanya hubungan relasi atau interaksi yang terjadi antar individu dalam *social network* dapat dipetakan dalam sebuah graf. Graf yang akan digunakan adalah graf tak berarah dan berbobot. Relasi atau hubungan yang terjadi antar individu yang terbentuk dalam graf tersebut dapat dihitung dengan menerapkan *centrality measurement*. *Centrality measurement* digunakan untuk mengukur tingkat pengaruh individu dalam suatu jaringan sosial. Setiap hubungan atau relasi yang terbentuk perlu dilakukan pembobotan agar hasil *centrality* yang diperoleh dapat menggambarkan kondisi sebenarnya. Metode pembobotan yang digunakan adalah *Probabilistic Partnership Index (PPI)*. Sedangkan untuk mengukur *centrality* tersebut menggunakan metode *Laplacian centrality*. Dengan menerapkan *PPI* sebagai metode pembobotan pada pengukuran *Laplacian centrality* diharapkan akan diperoleh nilai *centrality* yang menggambarkan kondisi sebenarnya sehingga dapat diketahui tingkat individu yang berpengaruh pada suatu jaringan sosial.

Kata Kunci : *Social Network, centrality measurement, laplacian centrality, probabilistic partnership index (PPI)*.

Abstract

Social network is used as a media connecting between the individual to be able to communicate unlimited by time and distance. One of the most popular social networks is *Twitter*. The existence of relationships or interactions that occur between individuals in a social network can be mapped in a graph. Graph that will be used is an weighted-undirected graph. The relationship that occurs between individuals that are formed in the graph can be computed by applying the centrality measurement. Centrality of measurement used to measure level of influence of individuals in a social network. Every relationship that formed should be weighted so that the results obtained centrality can describe the actual conditions. Weighting method used is a *Probabilistic Partnership Index (PPI)*. As for the centrality measure using *Laplacian centrality* method. By applying *PPI* as weighting method of measurement is expected to be obtained *laplacian centrality* centrality value that describes the actual conditions in order to obtain the individual level of influence in social networks.

Key words : *Social network, centrality measurement, laplacian centrality, probabilistic partnership index (PPI)*.

1. Pendahuluan

Saat ini kemajuan teknologi bertumbuh kembang dengan cepat, termasuk *social network* melalui *social media* seperti *facebook, twitter, path, instagram* dan lain-lain. Perkembangan *social network* ini tidak lepas dari kebutuhan pengguna yang memerlukan komunikasi tak terbatas oleh jarak dan waktu. *Social network* adalah suatu struktur sosial yang dibentuk oleh individu, sehingga antar individu akan dapat saling berhubungan [9]. Hubungan yang terjadi dapat diukur dan dianalisis dengan memanfaatkan teori graf yang didefinisikan sebagai *social network analysis* [7]. *Social network analysis* dapat digunakan untuk mendapatkan informasi yang implisit, seperti hubungan interaksi antar *user* [7]. *Twitter* sebagai bentuk *social network* dapat digunakan *user* sebagai media untuk dapat berhubungan, bertukar dan mengalirkan informasi, sehingga dapat diimplementasikan dalam *social network analysis* [14][17].

Dalam *social network analysis* terdapat *centrality measurement* yang digunakan untuk mengukur dan menentukan pentingnya hubungan antar node dalam suatu jaringan sosial [3][8]. Dengan menggunakan *centrality measurement*, dapat diketahui *user* yang paling berpengaruh dalam jaringan yang direpresentasikan dengan nilai *centrality* yang tertinggi [14]. Sehingga, *user* tersebut dapat dimanfaatkan untuk mengalirkan informasi dalam jaringan sosial. Salah satu contoh metode *centrality measurement* adalah *laplacian centrality*. Pada penelitian sebelumnya telah disebutkan bahwa metode *laplacian centrality* dapat digunakan untuk mengukur *centrality* suatu aktor yang direpresentasikan pada graf berbobot [10][11]. Akan tetapi, pada proses pembobotan masing-masing aktor masih menganut metode biasa, yaitu dengan menerapkan nilai bobot yang sama pada relasi antar *node*-nya sehingga nilai total relasinya dianggap kurang relevan. Oleh karena itu dibutuhkan metode pembobotan yang

digunakan untuk mengukur kekuatan antar *node* agar diperoleh hasil pengukuran *centrality* yang menggambarkan kondisi sebenarnya. Dalam analisis *bibliometric* atau *webometric* dikenal beberapa metode yang disebut *collaborative linkage indexes* yang dapat digunakan untuk mengukur kekuatan antar *node* pada graf, yaitu: *Jaccard Index*, *Salton-Ochiai Indexes*, *Probabilistic Afinity Index (PAI)* dan *Probabilistic Partnership Index (PPI)* [16].

Dalam tugas akhir kali ini akan menerapkan metode *Probabilistic Partnership Index (PPI)*. PPI dapat digunakan dalam memeriksa jaringan individu dalam pertukaran kompleks [16]. Dalam penelitian sebelumnya metode *Probabilistic Partnership Index* ini dapat diterapkan dalam pengukuran *degree centrality* [14]. Oleh karena itu, dengan menerapkan *Probabilistic Partnership Index (PPI)* bersama pada *laplacian centrality* diharapkan dapat diperoleh pengukuran *centrality* yang sesuai kondisi sebenarnya agar dapat diketahui tingkat *user* yang paling berpengaruh pada jaringan sosial.

Perumusan masalah yang akan dibahas pada penelitian ini adalah

- Bagaimana melakukan pembobotan dengan *Probabilistic Partnership Index* pada perhitungan *laplacian centrality*?
- Bagaimana pengaruh perubahan bobot terhadap pengukuran *laplacian centrality*?
- Bagaimana menentukan peringkat *user* yang paling berpengaruh di *Twitter*?

Dari perumusan diatas dibuat tujuan dari tugas akhir ini :

- Menganalisis pengaruh hasil pembobotan dengan metode *probabilistic partnership index* jika dibandingkan dengan pembobotan biasa pada pengukuran *laplacian centrality*.
- Menganalisa pengaruh perubahan bobot interaksi *following*, *mention* dan *reply* terhadap metode *probabilistic partnership Index* pada pengukuran *laplacian centrality*.
- Mengetahui 10 tingkat *user* yang paling berpengaruh pada *social network*.

Metode penelitian yang digunakan adalah dengan cara studi literatur, yaitu mempelajari materi yang berhubungan dengan konsep *social network analysis*, *laplacian centrality* serta *probabilistic partnership index*. Selanjutnya dilakukan analisis dan perancangan sistem. Dibuat juga dataset untuk pengujian sistem. Kemudian mengimplementasikan sistem yang telah dirancang dan melakukan pengujian terhadap sistem. Lalu dilakukan pengujian untuk tiap skenario pengujian sehingga mendapatkan hasil sistem lalu ditarik kesimpulan.

2. Dasar Teori

2.1 Social Network

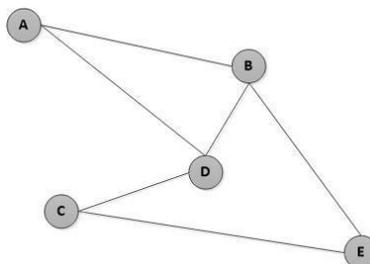
Menurut Freeman [7], jaringan sosial atau *social network* adalah suatu struktur sosial yang yang terdiri dari individu (atau organisasi) yang disebut "*node*" yang terhubung oleh satu atau lebih tipe tertentu saling ketergantungan seperti persahabatan, kekerabatan, kepentingan bersama, hubungan kepercayaan, dll.

Jejaring sosial memiliki cakupan yang luas pada pemanfaatannya yang dapat diimplementasikan diberbagai bidang, seperti bidang sains, ekonomi, sosiologi dan teknologi. Seiring dengan perkembangan dunia teknologi saat ini, jejaring sosial banyak dimanfaatkan sebagai media penyebaran informasi dan komunikasi yang efektif dan efisien.

Adanya relasi antar individu dalam jejaring sosial dapat digunakan sebagai representasi hubungan yang terbentuk antar individu. Hubungan yang terjadi dapat diukur dan dianalisis dengan memanfaatkan teori graf yang didefinisikan sebagai *social network analysis* [7]. Seperti yang telah dikemukakan oleh Otte dan Rousseau, hubungan antar individu menunjukkan bahwa *social network analysis* dapat digunakan dalam pengambilan informasi, termasuk di dalamnya hubungan antar individu dalam bentuk "*node*" [8]. *Node* adalah aktor individual dalam jaringan, *edge* adalah hubungan yang terjadi antara masing-masing *node*.

2.1.1 Representasi Social Network

Social network akan lebih mudah dipahami ketika direpresentasikan ke dalam bentuk graf dan matriks. Berikut contoh representasi *social network* dalam bentuk graf dan matriks.



Gambar 2.1: Representasi graf dari *social network*

Gambar 2.1 menunjukkan representasi pertemanan dalam bentuk graf. *Node* direpresntasikan dengan *vertex*, sedangkan relasi yang terbentuk antar *vertex* direpresentasikan dengan *edge*. *Social network* juga dapat direpresentasikan dalam bentuk matriks, contoh representasi matriks dari *social network* ditunjukkan pada tabel 2.1 sebagai berikut:

Tabel 2.1: Representasi matriks dari *social network*

vertex	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	1
C	0	1	0	1	1
D	1	1	1	0	1
E	0	1	1	1	0

Tabel 2.1 merepresentasikan hubungan pertemanan dalam bentuk matriks dari *social network*. Setiap hubungan antar *node* akan memiliki satu nilai. Nilai 1 menunjukkan adanya hubungan antara *node*, sebaliknya nilai 0 menunjukkan tidak adanya hubungan yang terjadi antar *node*. Sedangkan bagian diagonal matriks tersebut bernilai 0 karena *node* tidak mungkin berhubungan dengan dirinya sendiri.

2.2 Social Network Analysis

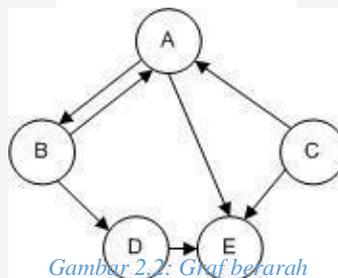
Menurut Scott [13], SNA adalah sekumpulan metode yang digunakan untuk menginvestigasi aspek relasi pada struktur sosial. Sedangkan menurut Krebs [15] SNA didefinisikan sebagai proses pemetaan dan pengukuran relasi antara orang yang satu dengan orang lain. Dari pendapat tersebut dapat ditarik kesimpulan secara garis besar SNA merupakan metode yang dapat digunakan untuk menganalisa interaksi sosial dalam suatu kelompok dengan cara melihat perilaku individu-individu tersebut dalam berhubungan satu sama lain.

2.3 Graf

Secara informal, graf adalah himpunan benda-benda yang disebut simpul (*vertex* atau *node*) yang terhubung oleh sisi (*edge*). Dari berbagai jenis graf yang ada, pada *social network analysis* digunakan suatu jenis graf yang terdiri dari *node* atau *vertex* untuk merepresentasikan suatu aktor. *Vertex* atau *node* tersebut saling terhubung satu sama lain [12]. Hubungan yang terjadi antar *vertex* direpresentasikan sebagai *edge* terdapat dua cara sebagai berikut:

2.3.1 Graf berarah

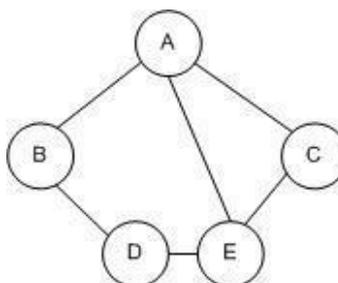
Graf berarah atau *directed graph* terdiri dari pasangan yang terurut yang menggambarkan hubungan yang detail dan jelas. *Directed graph* ditandai dengan anak panah pada ujung *edge*-nya.



Gambar 2.2: Graf berarah

2.3.2 Graf tidak berarah

Graf tidak berarah atau *undirected-graph* adalah graf sederhana yang menghubungkan antar dua *node*, dimana *edge* yang menghubungkan antara *node* tidak terdapat anak panah.



Gambar 2.3: Graf tidak berarah

Directed graph dan *undirected graph* dapat memiliki bobot pada masing masing *edge*-nya. Graf ini disebut dengan graf berbobot (*weighted graph*). Sedangkan graf yang tidak memiliki bobot disebut graf tak berbobot (*unweighted graph*). Pada tugas akhir kali ini, graf yang akan digunakan dalam analisis jejaring sosial adalah graf tak-berarah dan berbobot (*undirect-weighted graph*) karena *laplacian centrality* adalah metode pengukuran *centrality* untuk graf berbobot dalam bentuk graf tidak berarah [11].

2.4 Centrality Measurements

Hal utama dalam melakukan pembobotan dalam SNA adalah dengan mengukur *centrality* (*centrality measurement*). *Centrality measurement* mengukur kemampuan setiap *vertex* dalam mengalirkan informasi dan menentukan keterputusan suatu graf. Terdapat berbagai metode *centrality measurement* seperti yang telah diungkapkan oleh Freeman, yaitu *Degree Centrality*, *Betweenes Centrality*, *Closenees Centrality* [7].

2.4.1 Degree Centrality

Degree centrality adalah *centrality measurement* yang mengukur jumlah koneksi yang dimiliki oleh sebuah *node*. Dapat juga didefinisikan sebagai jumlah derajat relasi yang terhubung pada suatu *node* [4].

2.4.2 Closeness Centrality

Closeness centrality adalah dengan mengukur jarak rata-rata *node* dengan semua *node* yang lain dalam jaringan. *Closeness centrality* digambarkan dengan seberapa dekat suatu aktor dengan aktor lainnya dalam sebuah jaringan [13].

2.4.3 Betwenees Centrality

Betweenes centrality merupakan pengukuran yang menekankan pada kemampuan *node* dalam posisinya sebagai penghubung *node* satu dengan yang lainnya melalui *shortest path* yang melewati *node* tersebut [13].

2.4.4 Laplacian Centrality

Laplacian centrality merupakan salah satu *centrality measurement* yang memandang *centrality* dari *vertex* berkaitan dengan kemampuan jaringan untuk menanggapi penghapusan *vertex* dari jaringan [10][11].

a. Laplacian Matriks

Laplacian matriks pada graf berbobot (G) dirumuskan sebagai berikut:

$$L(G) = D(G) - A(G)$$

dimana

$$D(G) = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{bmatrix}$$

$$A(G) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

dengan $d_i = \sum_{j=1}^n a_{ij}$ dan d_i dikatakan *sum-weight* dari *vertex* v_i dimana a_{ij} adalah tetangga dari v_i

b. Laplacian Energy

Laplacian energy pada jaringan dirumuskan sebagai berikut:

$$LE(G) = \sum_{v=1}^n d_v + 2 \sum_{x-y \in E} w_{xy}$$

c. 2-Walks

Terdapat tiga tipe *2-walks*, yaitu:

1. *Closed 2-walks* yang mengandung *vertex* v

Gambar 2.4: Laplacian centrality

$$N_{\text{non-closed}}(v) = \sum_{u \in V} N_{\text{non-closed}}(v, u)$$

2. *Non-closed 2-walks* yang mengandung *vertex v* sebagai salah satu *end-vertex*

$$N_{\langle v, v \rangle} = \sum_{u \in V} (\sum_{w \in V} A_{uw})$$

3. *Non-closed 2-walks* yang mengandung *vertex v* sebagai titik tengah

$$N_{\langle v, v \rangle} = \sum_{u \in V} A_{uv} A_{vu}$$

d. Delta Laplacian Energy

Delta laplacian centrality adalah penurunan relatif *Laplacian energy* dalam jaringan yang disebabkan oleh penonaktifan *vertex* dari jaringan yang menggambarkan besarnya pengaruh *vertex* terhadap *laplacian energy* dari jaringan yang dirumuskan sebagai berikut:

$$(\Delta H)_v = \Delta H(v) - \Delta H(v) = 4N_{\langle v, v \rangle} + 2N_{\langle v, v \rangle} + 2N_{\langle v, v \rangle}$$

Dengan H merupakan jaringan yang diperoleh dari penghapusan *vertex v* dari jaringan G.

e. Laplacian centrality pada vertex

Laplacian Centrality pada *Vertex* dirumuskan sebagai berikut:

$$L_c(v) = \frac{(\Delta H)_v}{\Delta H(v)}$$

2.3 Collaborative Linkage Indexes

Collaborative linkage indexes, dikenal juga dengan *similarity coeficient*. *Collabrative* adalah metode menghasilkan suaty *paper* ilmiah dari sekelompok peneliti. Indikator *co-authorship* adalah salah satu indikator yang digunakan untuk mengalisis jumlah *paper* yang dihasilkan oleh gabungan peneliti. *Collaborative linkage index* perlu diukur karena seiring dengan banyaknya *paper* yang dihasilkan oleh para peneliti. Adapaun metode *collaborative linkage indexes* yang ada saat ini, yaitu [16]:

2.3.1 Jaccard Indexes dan Salton-Ochiai Indexes

Jaccard Indexes dan *Salton-Ochiai Indexes* digunakan untuk mengukur hubungan yang terjadi dalam waktu yang hampir bersamaan. Kedua *index* dapat dihitung baik berdasarkan dengnan artikel yang ditulis atau hubungan yang dihasilkan pada *co-authorship* [16].

Jaccard Index adalah jumlah artikel *co-athored* antara dua sektor i dan j dibagi oleh gabungan artikel dari kedua sektor.

2.3.2 Probabilistic Affinity Index (PAI)

Probabilistic Affinity Index (PAI) merupakan rasio dari jumlah link yang diamati dengan jumlah link yang diharapkan (*expected value*) yang menunjukkan kecenderungan relative pada *co-aothorsip* [16].

2.3.4 Probabilistic Partnership Index (PPI)

Probabilistic Partnership Index (PPI) digunakan dalam memeriksa jaringan individu dalam pertukaran kompleks. Setiap hubungan yang terjadi di dalam jaringan dapat diproyeksikan untuk standar model masing-masing. Selain itu, PPI juga dapat dinyatakan sebagai perbedaan standar jumlah *link* yang dapat diharapkan [16]. PPI dirumuskan sebagai berikut:

$$PPI_{ij} = \frac{R_{ij} - \hat{E}[R_{ij}]}{\sigma}$$

Dimana R_{ij} adalah relasi yang terbentuk antara *node i* dan *node j*, sedangkan $\hat{E}[R_{ij}]$ adalah nilai *expected value*, dan σ adalah standar deviasi dari jumlah *link* antara *node i* dan *j* yang dicari menggunakan metode *monte carlo*.

Dari proses perhitungan akan didapat nilai PPI. Nilai PPI = 0 mengindikasikan bahwa jumlah *link* yang diamati sebanding dengan nilai *expected value*. Sedangkan jika nilai PPI > 0 mengindikasikan bahwa jumlah *link* yang diamati lebih besar dibandingkan dengan nilai *expected value*. Nilai ini akan dinormalisasi $\frac{PPI_{ij}}{2 + |PPI_{ij}|} = 2 / (1 + \exp(-0,183 PPI_{ij})) - 1$. Sehingga nilai PPI yang digunakan adalah nilai yang sudah dilakukan proses normalisasi.

2.4 Social Media

Social media adalah sebuah media yang digunakan oleh individu untuk saling berkomunikasi, berbagi informasi, bersosialisasi satu sama lain secara *online* tanpa terbatas oleh ruang dan waktu [1][2].

2.4.1 Social Media Twitter

Twitter adalah salah satu *social media* yang berbentuk *microblogging* yang dimiliki dan dioperasikan oleh *Twitter Inc.* *Twitter* memungkinkan penggunaanya untuk berkomunikasi dan berbagi informasi dengan mengirim dan membaca pesan teks hingga 140 karakter [6]. Di dalam *twitter* dikenal berbagai jenis fitur yang dapat digunakan oleh penggunaanya untuk dapat saling berkomunikasi dan berbagi informasi, yaitu:

a. *Following*

Following adalah fitur yang dapat digunakan oleh pengguna untuk dapat mengikuti pengguna lain. Apabila seorang pengguna mengikuti pengguna lain, maka pengguna tersebut akan memperoleh kicauan yang dibuat oleh pengguna lain tersebut. Selain itu dengan fitur ini, kedua pengguna yang saling mengikuti satu sama lain, dapat saling melakukan *Direct Message*.

b. *Mention*

Mention merupakan salah satu fitur yang ada pada *twitter* yang digunakan ketika seorang pengguna ingin berinteraksi dengan pengguna lain. Pengguna yang akan menggunakan fitur *mention* ini ditandai dengan '@' diikuti *username* pengguna lain tersebut pada saat membuat kicauan.

c. *Retweet*

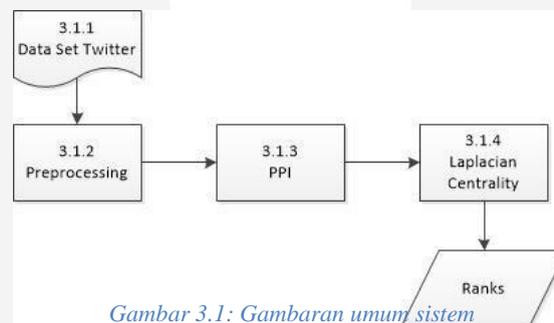
Retweet adalah fitur yang ada pada *twitter* yang digunakan oleh pengguna untuk kembali mem-*posting tweet* yang dilakukan oleh pengguna lain sehingga dapat dibaca atau dibagikan kepada pengguna lain yang mengikuti pengguna tersebut [5]. Dengan adanya fitur ini, pengguna dapat menyebarkan informasi dari pengguna lain secara mudah tanpa merubah isi kicauan aslinya.

d. *Reply*

Reply adalah fitur yang ada pada *Twitter* yang digunakan oleh *user* untuk membalas kicauan yang dikirim oleh *user* lain. Fitur *reply* ini hanya akan dapat dilihat oleh pengguna penerima *reply* dan pengguna lain yang mengikuti keduanya. Kicauan dapat dikatakan *Reply* apabila dalam kicauan yang dibuat diawali dengan *@username*.

3. Perancangan Sistem

Secara umum sistem yang akan dibangun memiliki dua proses utama yaitu, melakukan pengukuran nilai *centrality*, serta melakukan pembobotan. Dalam membangun sistem ini, untuk mengukur nilai *centrality* dengan mengimplementasikan metode *laplacian centrality*, sedangkan untuk melakukan fungsi pembobotan dengan mengimplementasikan metode pembobotan *probabilistic partnership index (PPI)*. Masukkan sistem ini berupa sebuah matriks hasil *preprocessing* dari *data set Twitter* yang diperoleh dengan melakukan *retrieve* data menggunakan *NodeXL* yang merupakan *template* yang terdapat pada aplikasi perangkat lunak *Microsoft Office Excel*. Dari sistem ini, akan dihasilkan nilai *centrality node* suatu jaringan yang dari hasil tersebut dapat diketahui hasil peringkat berdasarkan nilai *laplacian centrality*. Hasil nilai *centrality* dan peringkat ini kemudian akan dianalisis untuk dapat ditarik kesimpulan. Gambaran umum sistem ditunjukkan pada gambar 3.1 di bawah ini.



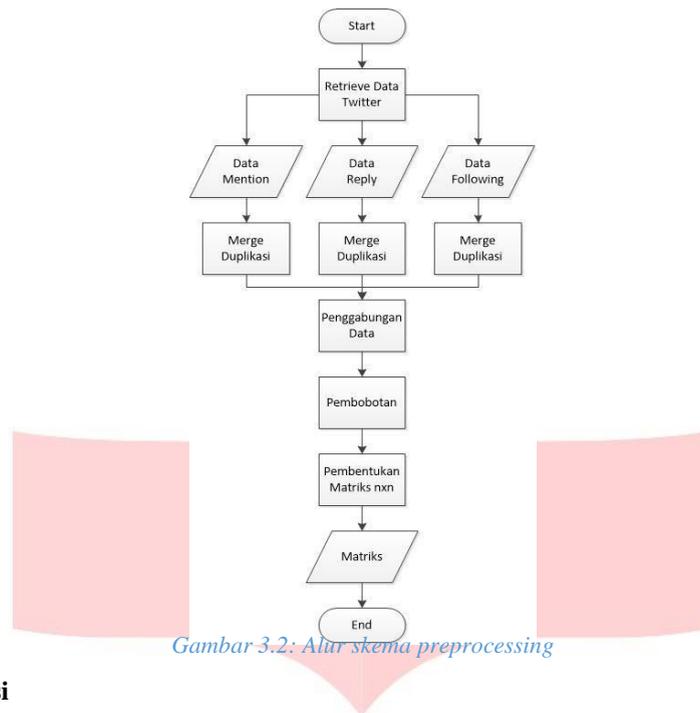
Gambar 3.1: Gambaran umum sistem

3.1 Data set

Data set yang digunakan merupakan *data set twitter* yang berasal dari *Twitter @anggawicaksono* yang diperoleh dengan cara melakukan *retrieve* menggunakan *NodeXL*. Pemilihan akun *@anggawicaksono* yang dipilih dengan alasan memudahkan dalam melakukan pengambilan *data set*, karena dalam pengambilan *data set* dengan menggunakan *NodeXL* harus diketahui *username* dan *password* *Twitter* untuk proses autentifikasi pada *Twitter API*. *Data set* yang diambil adalah *data set* berupa relasi *mention*, *reply*, dan *following*. *Data set* yang diperoleh dari hasil *retrieve* ini kemudian akan dilakukan *preprocessing* agar sesuai dengan kebutuhan masukan sistem.

3.2 Preprocessing

Data set yang diperoleh dari hasil *retrieve* akan dilakukan *preprocessing* dengan gambaran diagram alir seperti pada gambar 3.2 sebagai berikut:



Gambar 3.2: Alur skema preprocessing

3.2.1 Merge Duplikasi

Proses ini dilakukan untuk menghilangkan relasi-relasi yang duplikat pada *data set* yang diperoleh dari hasil *retrieve NodeXL*. *Merge duplikasi* dilakukan pada semua *data set*, yaitu *data mention*, *data reply*, dan *data following*.

3.2.2 Penggabungan Data

Dari ketiga *data set* yang telah dilakukan *merge duplikasi*, kemudian digabungkan menjadi satu *data set* yang berisi relasi *mention*, *reply*, dan *following*.

3.2.3 Pembobotan

Pada awalnya data hasil *retrieve twitter* yang belum memiliki bobot, namun karena pada metode *laplacian centrality* mengharuskan data yang digunakan harus memiliki bobot, maka akan dilakukan proses pembobotan dengan menambahkan kolom bobot relasi dengan menginisialisasikan besar bobot masing-masing relasi *mention*, *reply*, dan *following*. Setiap interaksi yang terjadi, maka bobot akan bertambah sebesar *x*, *y* ataupun *z*. penambahan bobot didasarkan pada penelitian sebelumnya [17] yang memberikan besaran bobot untuk relasi *following* adalah +1, *reply* +1, dan *mention* +1. Total bobot antara *vertex 1* dan *vertex 2* merupakan hasil penjumlahan bobot setiap relasi *mention*, *reply*, dan *following* yang didapat dari hasil *retrieve twitter*.

Berikut contoh skenario jumlah bobot yang dihasilkan dari interaksi yang terjadi antara *vertex A* dan *vertex B*. *User A follow user B*, *User B follow user A*, *user A* melakukan *mention* kepada *user B* sebanyak 1 kali, *user B* melakukan *reply* kepada *user A* sebanyak 1 kali. Maka

$$\begin{aligned}
 C_{AB} &= \text{[Mention]} + \text{[Reply]} + \text{[Following]} + \text{[Following]} \\
 &= 1 + 1 + 1 + 1 \\
 &= 7
 \end{aligned}$$

3.2.4 Pembentukan Matriks

Tahap selanjutnya yaitu melakukan pembentukan matriks *nxn* agar dapat digunakan sebagai masukan sistem.

3.3 Pembobotan Probabilistic Partnership Index

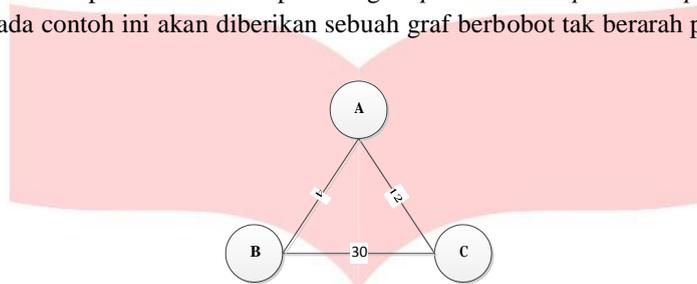
Melakukan perhitungan bobot *probabilistic partnership index (PPI)* dengan terlebih dahulu menghitung nilai *expected value* dan *standar deviasi*. Perhitungan *expected value* dan *standar deviasi* dilakukan berdasarkan perhitungan distribusi normal sehingga diperoleh nilai untuk perumusan PPI seperti pada persamaan (13). *Expected value* dan *standar deviasi* didapatkan dengan menggunakan metode *monte carlo*. *Monte carlo* diterapkan dengan melibatkan angka acak untuk memodelkan sistem. Nilai data yang akan diambil, dibangkitkan dengan menggunakan pembangkit angka acak dengan menggunakan fungsi *rand* yang terdapat pada *Matlab*. Pembangkit angka acak akan membangkitkan peubah acak yang berguna untuk menentukan nilai *expected value* dan *standar deviasi* pada suatu relasi. Nilai *expected value* yang diperoleh untuk setiap *node* kemudian akan dihitung rata-ratanya kemudian diterapkan dalam perhitungan nilai PPI suatu *node*.

3.4 Pengukuran Laplacian Centrality

Melakukan pengukuran *laplacian centrality* pada setiap *vertex* yang berada pada satu jaringan yang sama. Hal yang dilakukan adalah membuat *laplacian matriks* dengan menggunakan persamaan (1). Kemudian melakukan perhitungan *laplacian energy* jaringan tersebut dengan menggunakan persamaan (4). Setelah itu melakukan perhitungan *2-walks* yang terdiri dari tiga tahap yaitu *2-walks type 1* untuk menghitung nilai relasi antar *node* dengan menggunakan persamaan (5), *2-walks type 2* untuk menghitung ketetangaan dengan menggunakan persamaan (6), dan *2-walks type 3* untuk menghitung keterpusatan *node* dengan menggunakan persamaan (7). Kemudian dari hasil ketiga tahap *2-walks* tersebut akan digunakan untuk menghitung *delta laplacian energy* dengan menggunakan persamaan (8). Terakhir untuk mendapatkan nilai *laplacian centrality node* dengan menggunakan persamaan (9).

3.5 Contoh Implementasi

Untuk memahami proses-proses yang akan diimplementasikan menurut penjelasan di atas, akan diberikan sebuah contoh permasalahan dan implementasi dari perhitungan *probabilistic partnership index* dan pengukuran *laplacian centrality*. Pada contoh ini akan diberikan sebuah graf berbobot tak berarah pada gambar 3.3 sebagai berikut:



Gambar 3.3 Graf berbobot tak berarah

Dari graf di atas akan direpresentasikan ke dalam matriks *adjacency nxn* pada tabel 3.1 sebagai berikut:

Tabel 3.1 Representasi matriks berbobot tak berarah

	A	B	C
A	0	4	12
B	4	0	30
C	12	30	0

Dari tabel 3.1 di atas, dapat ditunjukkan nilai bobot relasi antara *vertex A* dan *vertex B* adalah bernilai 4. Kemudian akan diubah bobotnya dengan menerapkan metode pembobotan *probabilistic partnership index*. Sebagaimana seperti pada tugas akhir sebelumnya [14] untuk menghitung nilai *probabilistic partnership index*, dengan mencari *expected value* dan *standar deviasi* yang didapatkan dengan metode *monte carlo* [16]. Sebagai contoh akan dihitung nilai PPI untuk relasi A-B, sesuai persamaan (14) sebagai berikut:

Diketahui

$$\begin{aligned}
 N_{ij} &= 4 \\
 N_{ij}^{\#} &= \frac{1}{N} \sum_{k=1}^N \sum_{l=1}^N A_{kl} A_{li} A_{kj} \\
 &= \frac{92}{6} = 15,33 \\
 N_{ij}^N &= 100, \text{ banyaknya nilai pembangkit acak} \\
 N_{ij}^{\#} &= 14,4600 \\
 \hat{E}[A_{ij}] &= 0,2766 \\
 N_{ij}^{\#} &= \sqrt{\frac{N_{ij}^{\#} - N_{ij}^2}{N}} \cdot N_{ij} \\
 &= \sqrt{(14,4600 - 15,33)^2 \cdot \frac{100}{14,4600}} \\
 &= 1.2628 \\
 PP_{ij} &= \frac{A_{ij} - \hat{E}[A_{ij}]}{\hat{\sigma}} \\
 &= \frac{4 - 0,2766}{1,2628} \\
 &= 2,9492
 \end{aligned}$$

$$N_{PP} = 2 / (1 + \exp(-0,185 + 0,6667)) - 1 = 0,2634$$

Dari perhitungan tersebut dapat diperoleh nilai PPI untuk relasi *vertex A-B* adalah 0,2634 yang merepresentasikan hubungan ilmiah yang terbentuk antara *vertex A* dan *vertex B*. cara perhitungan yang sama diterapkan untuk semua node untuk memperoleh nilai PPI pada masing-masing relasi. Sehingga, nilai PPI yang didapatkan adalah dapat dilihat pada tabel 3.2 sebagai berikut:

Tabel 3.2 Tabel hasil perhitungan PPI

	A	B	C
A	0	0,2634	0,5092
B	0,2634	0	0,7727
C	0,5092	0,7727	0

Dari, matriks hasil pembobotan PPI di atas, akan dilakukan penghitungan nilai *laplacian centrality* untuk *vertex A* sesuai persamaan (9) berikut ini:

a. Membuat *laplacian matriks*, berdasarkan persamaan (1).

$$L = D - A = \begin{bmatrix} 0,7726 & 0 & 0 & 0 & 0,2634 & 0,5092 \\ 0 & 1,0361 & 0 & 0 & 0,2634 & 0,7277 \\ 0 & 0 & 1,2818 & 0,5092 & 0,7277 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 0,7726 & -0,2634 & -0,5092 \\ -0,2634 & 1,0361 & -0,7727 \\ -0,5092 & -0,7727 & 1,2819 \end{bmatrix}$$

$$= \begin{bmatrix} 0,7726 & -0,2634 & -0,5092 \\ -0,2634 & 1,0361 & -0,7727 \\ -0,5092 & -0,7727 & 1,2819 \end{bmatrix}$$

b. Perhitungan *laplacian energy* pada jaringan, berdasarkan persamaan (4).

$$L(E) = \sum_{i=1}^n d_i^2 + 2 \sum_{i < j} w_{ij}^2$$

$$L(E) = 0,7726^2 + 1,0361^2 + 1,2818^2 + 2(0,2634^2 + 0,7277^2) + 2(0,5092^2) = 0,5969 + 1,0735 + 1,6433 + 2(0,0694 + 0,2593) + 2(0,5970) = 3,3135 + 2(0,3287) + 2(0,5970) = 4,6667$$

c. Perhitungan *2-Walks*

Type 1, berdasarkan persamaan (5).

$$N_{2,1}(G) = \sum_{i,j \in V} d_i d_j = 0,7726^2 + 1,0361^2 = 0,0694 + 0,2593 = 0,3287$$

Type 2, berdasarkan persamaan (6).

$$N_{2,2}(G) = \sum_{i,j \in V} (d_i + d_j) w_{ij} = (0,7726 + 1,0361) \times 0,2634 + (1,0361 + 1,2818) \times 0,7277 = 0,2634 \times 0,7727 + 0,5092 \times 0,7277 = 0,3711$$

Type 3, berdasarkan persamaan (7).

$$N_{2,3}(G) = \sum_{i,j \in V} d_i d_j w_{ij} = 0,2634 \times 0,5092 = 0,1341$$

d. Perhitungan *delta laplacian energy*, berdasarkan persamaan (8).

$$\begin{aligned} (\Delta L)_d &= L_2(G) - L_1(G) \\ &= 4N_{2,1}(G) + 2N_{2,2}(G) + 2N_{2,3}(G) \\ &= 4(0,3287) + 2(0,3711) + 2(0,1341) \\ &= 2,7769 \end{aligned}$$

c. Perhitungan *laplacian centrality*, berdasarkan persamaan (9).

$$L_c(G) = \frac{(\Delta L)_d}{L(E)}$$

$$= \frac{2,7769}{5,1651} = \text{0,5376}$$

Dari hasil perhitungan di atas, maka nilai *centrality vertex A* adalah 0,5376. Perhitungan yang sama juga diterapkan untuk menghitung nilai *centrality node* lainnya. Sehingga didapatkan nilai *centrality vertex A*, *vertex B* dan *vertex C* dalam jaringan sosial tersebut adalah seperti yang ditunjukkan pada tabel 3.3 sebagai berikut:

Tabel 3.3 Tabel perolehan nilai laplacian centrality

Vertex	Nilai Laplacian Centrality
A	0,5376
B	0,7992
C	0,9436

4. Pengujian Sistem

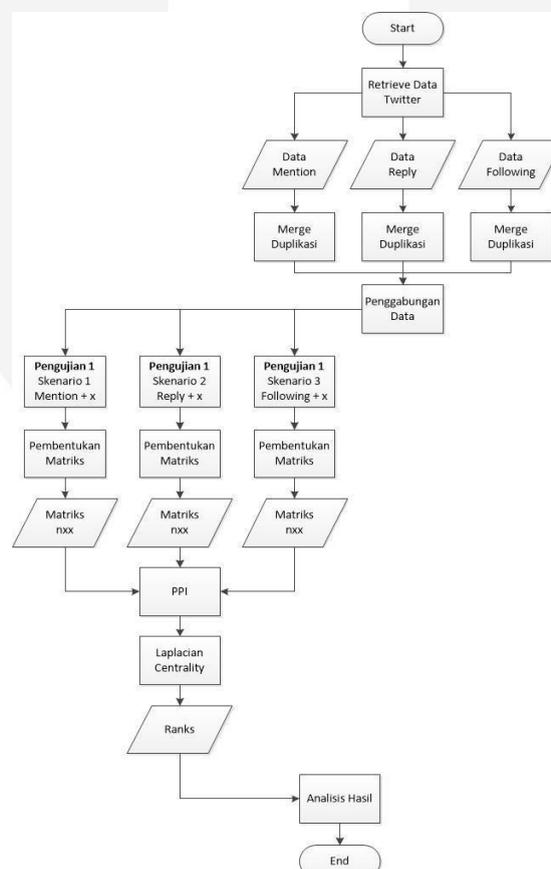
Pengujian pada Tugas Akhir ini dilakukan untuk mengetahui hasil dari pengukuran *laplacian centrality* pada graf berbobot tak berarah dan menganalisa pengaruh penggunaan metode pembobotan *probabilistic partnership index* serta perubahan nilai bobot pada masing-masing relasi *mention*, *reply* dan *following*.

4.1 Data Pengujian

Data yang digunakan dalam pengujian Tugas Akhir ini berupa data *Twitter* pada akun @anggawicaksono yang diperoleh dengan menggunakan *NodeXL*. Pengambilan data dilakukan sebanyak 15 kali selama 15 hari berturut-turut, hal ini disebabkan karena *server Twitter* membatasi waktu pengambilan data yang dikenal dengan *Twitter rate limit*. Pengambilan data relasi *mention* dan *reply* dilakukan masing-masing 15 kali, sedangkan pengambilan data relasi *following* dilakukan hanya sekali saja karena *user* hanya sekali saja pula melakukan interaksi *following*. Dari semua data tersebut kemudian digabung berdasarkan setiap relasi *mention*, *reply* dan *following*. Kemudian data tersebut akan dilakukan *preprocessing* seperti penjelasan pada bab sebelumnya.

4.2 Skenario Pengujian

Untuk mencapai tujuan yang telah dijelaskan di atas, akan dilakukan skenario pengujian pada *data set* hasil *retrieve* menggunakan *NodeXL*. Skenario pengujian akan digambarkan dalam bentuk diagram alir pada gambar 4.1 sebagai berikut:

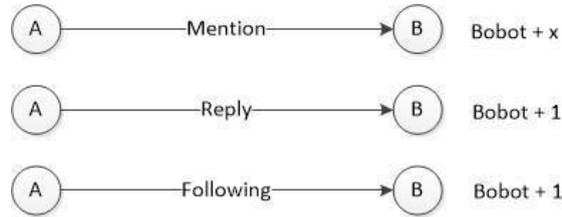


Gambar 4-1: Alur Skema pengujian

4.3 Hasil Pengujian Sistem

4.3.1 Skenario 1

Pengujian pada skenario perubahan bobot *mention* dilakukan sebanyak 11 kali dengan bobot *reply* dan *following* bernilai 1, dan bobot *mention* bernilai 1 berurut hingga 11.



Gambar 4-2: Skenario perubahan nilai bobot *mention*

Dari hasil pengujian, didapatkan hasil 10 peringkat *vertex* dengan nilai *centrality* tertinggi, adalah sebagai berikut:

Rank	Uji1	Uji2	Uji3	Uji4	Uji5	Uji6	Uji7	Uji8	Uji9	Uji10	Uji11
1	7	7	7	7	7	7	7	7	7	7	7
2	665	316	67	316	316	316	316	316	316	316	316
3	316	67	665	665	276	67	665	665	67	67	67
4	393	276	316	276	67	665	276	67	276	276	276
5	67	665	276	67	665	276	67	276	665	665	665
6	129	393	129	393	393	393	393	393	393	393	3
7	302	129	393	129	129	129	129	3	129	129	393
8	312	302	302	302	302	3	3	129	3	3	129
9	53	3	3	531	3	302	302	302	302	302	302
10	531	102	531	3	9	102	531	102	9	102	9

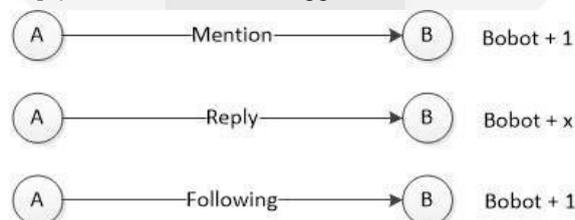
Keterangan: 3. alfagil; 7. anggawicaksono; 9. billyfuad; 53. AgainstAIDS; 67. infobdg; 102. annisarah; 129. fikryajihady; 276. arbyyasin; 302. fauqalanhar; 312. dariusputra; 316. Infoitt; 393. herryadhin; 394. ArdhiListiyanto; 531. eldofadliadi; 665. FarhanRyandi.

Dilihat pada Tabel 4.4, posisi ranking dari setiap percobaan pengujian yang dilakukan, terdapat adanya peningkatan dan penurunan posisi ranking. Posisi ranking ke-1 tidak mengalami perubahan pada setiap percobaan disebabkan karena selisih nilai *laplacian centrality* yang dihasilkan memiliki selisih yang signifikan dibandingkan *vertex* yang lainnya. Sedangkan pada posisi ranking yang lainnya, posisi ranking mengalami perubahan. Meskipun terdapat beberapa posisi ranking yang cenderung ditempati oleh *vertex* yang sama seperti pada posisi ranking ke 2 dan 6. Selain itu, perubahan tersebut disebabkan oleh pembangkit nilai acak pada metode *monte carlo* yang digunakan untuk menghasilkan nilai *expected value* dan *standar deviasi* pada perhitungan pembobotan PPI yang menyebabkan berubah-ubahnya nilai *laplacian centrality*, sehingga memungkinkan adanya perubahan pada posisi ranking *vertex* pada jaringan tersebut.

Dengan demikian, dapat diambil kesimpulan bahwa perubahan nilai bobot *mention* pada pengukuran nilai *laplacian centrality* dengan melibatkan pembobotan PPI dapat mempengaruhi posisi ranking pada jaringan tersebut.

4.3.2 Skenario 2

Pengujian pada skenario perubahan bobot *reply* dilakukan sebanyak 11 kali dengan bobot *mention* dan *following* bernilai 1, dan bobot *reply* bernilai 1 berurut hingga 11.



Gambar 4-3: Skenario perubahan nilai bobot *reply*

Dari hasil pengujian, didapatkan hasil 10 peringkat *vertex* dengan nilai *centrality* tertinggi, adalah sebagai berikut:

Rank	Uji1	Uji2	Uji3	Uji4	Uji5	Uji6	Uji7	Uji8	Uji9	Uji10	Uji11
1	7	7	7	7	7	7	7	7	7	7	7
2	665	316	665	316	276	316	316	316	316	316	316
3	316	393	316	665	316	67	67	276	665	665	276
4	393	67	276	67	665	393	276	665	67	276	393
5	67	665	67	276	67	276	665	393	393	67	665
6	129	129	129	393	129	665	393	129	276	393	67
7	302	302	393	102	393	129	129	67	129	129	129
8	312	276	9	129	302	102	302	302	102	3	302
9	53	394	102	3	3	302	3	102	302	102	102
10	531	531	302	9	394	531	9	9	9	302	3

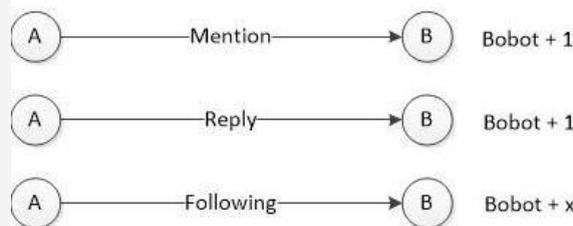
Keterangan: 3. alfagil; 7. anggawicaksono; 9. billyfuad; 53. AgainstAIDS; 67. infobdg; 102. annisarah; 129. fikryajihady; 276. arbyyasin; 302. fauqalanhar; 312. dariusputra; 316. Infoitt; 393. herryadhin; 394. ArdhiListiyanto; 531. eldofadliadi; 665. FarhanRyandi.

Dilihat pada Tabel 4.7, posisi ranking dari setiap percobaan pengujian yang dilakukan, terdapat adanya peningkatan dan penurunan posisi ranking. Posisi ranking ke-1 tidak mengalami perubahan pada setiap percobaan disebabkan karena selisih nilai *laplacian centrality* yang dihasilkan memiliki selisih yang signifikan dibandingkan *vertex* yang lainnya. Sedangkan pada posisi ranking yang lainnya posisi ranking mengalami perubahan. Meskipun terdapat beberapa posisi ranking yang cenderung ditempati oleh *vertex* yang sama seperti pada posisi ranking ke 2. Selain itu, perubahan tersebut disebabkan oleh pembangkit nilai acak pada metode *monte carlo* yang digunakan untuk menghasilkan nilai *expected value* dan *standar deviasi* pada perhitungan pembobotan PPI yang menyebabkan berubah-ubahnya nilai *laplacian centrality*, sehingga memungkinkan adanya perubahan pada posisi ranking *vertex* pada jaringan tersebut.

Dengan demikian, dapat diambil kesimpulan bahwa perubahan nilai bobot *reply* pada pengukuran nilai *laplacian centrality* dengan melibatkan pembobotan PPI dapat mempengaruhi posisi ranking pada jaringan tersebut.

4.3.3 Skenario 3

Pengujian pada skenario perubahan bobot *following* dilakukan sebanyak 11 kali dengan bobot *mention* dan *reply* bernilai 1, dan bobot *following* bernilai 1 berurut hingga 11.



Gambar 4-4: Skenario perubahan nilai bobot *following*

Dari hasil pengujian, didapatkan hasil 10 peringkat *vertex* dengan nilai *centrality* tertinggi, adalah sebagai berikut:

Rank	Uji1	Uji2	Uji3	Uji4	Uji5	Uji6	Uji7	Uji8	Uji9	Uji10	Uji11
1	7	7	7	7	7	7	7	7	7	7	7
2	665	316	316	316	316	316	316	316	316	316	316
3	316	276	276	393	276	665	665	276	665	665	665
4	393	665	665	665	665	276	276	129	276	276	276
5	67	393	393	276	393	393	393	393	67	393	393
6	129	67	67	67	67	67	129	665	393	67	67
7	302	129	129	129	129	129	67	67	129	129	129
8	312	302	302	302	302	302	302	302	302	302	302
9	53	3	3	3	3	3	531	3	3	3	3
10	531	9	102	102	9	531	3	9	531	394	531

Keterangan: 3. alfagil; 7. anggawicaksono; 9. billyfuad; 53. AgainstAIDS; 67. infobdg; 102. annisarah; 129. fikryajihady; 276. arbyyasin; 302. fauqalanhar; 312. dariusputra; 316. Infoitt; 393. herryadhin; 394. ArdhiListiyanto; 531. eldofadliadi; 665. FarhanRyandi.

Dilihat pada Tabel 4.10, posisi ranking dari setiap percobaan pengujian yang dilakukan, terdapat adanya peningkatan dan penurunan posisi ranking. Posisi ranking ke-1 tidak mengalami perubahan pada setiap percobaan disebabkan karena selisih nilai *laplacian centrality* yang dihasilkan memiliki selisih yang signifikan dibandingkan *vertex* yang lainnya. Sedangkan pada posisi ranking yang lainnya posisi ranking mengalami perubahan. Selain itu, perubahan tersebut disebabkan oleh pembangkit nilai acak pada metode *monte carlo* yang digunakan untuk menghasilkan nilai *expected value* dan *standar deviasi* pada perhitungan pembobotan PPI. yang menyebabkan berubah-ubahnya nilai *laplacian centrality*, sehingga memungkinkan adanya perubahan pada posisi ranking *vertex* pada jaringan tersebut.

Dengan demikian, dapat diambil kesimpulan bahwa perubahan nilai bobot *following* pada pengukuran nilai *laplacian centrality* dengan melibatkan pembobotan PPI dapat mempengaruhi posisi ranking pada jaringan tersebut.

4. Kesimpulan dan Saran

Kesimpulan yang dapat diambil dari tugas akhir ini antara lain:

1. Pada pengujian *laplacian centrality* dengan melibatkan pembobotan *probabilistic partnership index*, nilai *laplacian energy* suatu *vertex* berbanding lurus dengan perolehan nilai *laplacian centrality*. Namun, dalam beberapa kasus pengujian, perolehan nilai *centrality* mengalami kenaikan atau penurunan. Hal ini disebabkan karena pengaruh pembangkit nilai acak pada metode *monte carlo* yang digunakan untuk menghasilkan nilai *expected value* dan *standar deviasi* pada perhitungan metode pembobotan *Probabilistic Partnership Index*.
2. Pada pengujian nilai *laplacian centrality* penambahan bobot pada interaksi *mention*, *reply*, dan *following* berbanding lurus dengan perolehan nilai *centrality*. Pada *vertex* dan atau *vertex* tetangganya yang tidak melakukan salah satu, dua, atau ketiga interaksi tersebut akan memperoleh nilai *laplacian centrality* yang kecil di saat *vertex* lain yang tidak terhubung langsung dengan *vertex* tersebut melakukan salah satu, dua, atau ketiga interaksi tersebut.
3. *Vertex* yang paling berpengaruh pada jaringan dari hasil pengukuran *Laplacian Centrality* secara berturut-turut adalah *anggawicaksono*, *FarhanRyandi*, *infoitt*, *herryadhin*, *fikryaljihady*, *fauqalanhar*, *dariusputra*, *AgainstAIDS*, *eldodfadliadi*.

Saran yang dapat diberikan dan khususnya berkaitan dengan kasus pada tugas akhir ini antara lain

1. Untuk penelitian lebih lanjut, dapat menggunakan *data set* dari social network yang lain seperti *facebook*.
2. Menggunakan metode pembobotan lain selain *probabilistic partnership index*. Seperti *Jaccard Index*, *Salton-Ochiai*.
3. Untuk penelitian lebih lanjut, dapat menggunakan *centrality measurement* yang lain.

Daftar Pustaka:

- [1] A. Sitaram. (2010). Predicting the Future With Social Media. Social computing lab HB Labs, Palo Alto, California, USA.
- [2] Anonymous, "Jejaring Sosial" [Online]. Available: http://id.wikipedia.org/wiki/Jejaring_sosial
- [3] Crnovrsanin, Tark; Carlos D., Correa., & Kwan-Liu Ma (2009). Social Network Discovery based on Sensitivity Analysis. Advanced Analysis and Mining, pp. 107-112.
- [4] Eunice E. Santos, Chair et al. (2007). "Effective and Efficient Methodologies for Social Network Analysis". Virginia, USA
- [5] Hong, L., Convertino, G., & Chi, E. H. (2011). Language Matters in Twitter: A Large Scale Study. Fifth International AAAI Conference on Weblogs and Social Media
- [6] Java, A., Finin, T., Song, X., & Tseng, B. (2007). Why We Twitter: Understanding Microblogging Usage and Communities. ACM
- [7] L. Freeman, "The Study of Social Network," 2002. [Online]. Available: http://www.insana.org/INSNA/na_inf.htm [Accessed 08 April 2014]
- [8] Otte, E and Rousseau, R (2002). "Social network analysis: a powerful strategy, also for the information sciences," Information Science, vol 28, pp. 443-445
- [9] Pan, L (2007). "Effective and Efficiency Methodologies for SNA". USA: Virginia Polytechnic Institute and State University [1]
- [10] Qi, X., Duval, R. D., Christensen, K., Fuller, E., Spahiu, A., Wu, Q., et al (2013). Terrorist Networks, Network Energy and Node Removal: A New Measure of Centrality Based on Laplacian Energy. Social Networking, 19-31
- [11] Qi, Xi. Fuller, E. Wu, Qin. Wu, Yezhou. Zhang, Cun. (2012). "Laplacian Centrality: A New centrality measure for weighted networks," Information Science, vol 194, pp. 240-253
- [12] Ruohen, K. (2012). Graph Theory
- [13] Scoot John. (1992). Social network Analysis Theory and Application. Newberry Park CA :Sage
- [14] Sharafina N, (2013). "Implementasi Probabilistic Partnership Index (PPI) pada Analisis Jejaring Sosial dengan Pendekatan Kretschmer" Fakultas Informatika, Institut Teknologi Telkom, Bandung
- [15] V. Krebs. (2010). "How do to Social Network Analysis,". [Online]. Available: <http://www.orgnet.com/sna.html> . [Accessed 10 April 2014].
- [16] Yamashita, Yasuhiro and Yoshiko Okubo, (2006). "Patterns of Scientific Colaboration between Japan and France: Inter-sectoral Analysis using Probabilistic Paetnership Index (PPI)," Akademia Kiado, Budapest and Springer. Dordrecht, Scientometrics, vol 68, no.2, pp. 2

- [17] Sulaskin, A, (2012). “Analisis Degree Centrality dalam Social Network Analysis menggunakan Probabilistic Affinity Index (PAI) pada Graf Berarah-berbobot”, Fakultas Informatika, Intitut Teknologi Telkom, Bandung
- [18] Kretschmer, H and Kretschmer, T, (2010), “A New Centrality Measure for Social Network Analysis Applicable to Bibilometric and Webometric Data”, Departement of Library and Information Science, 26-D-10117, Humboldt-University, Berlin.

