**ANALISIS INTERAKSI TWEET VAKSINASI PADA JEJARING SOSIAL TWITTER DENGAN AKUN PRFMNEWS**

****

**Rafif Ridho - 05111840000058**

**Muhammad Ilham Bayhaqi - 0511184000069**

**DEPARTEMEN INFORMATIKA**

**FAKULTAS TEKNOLOGI ELEKTRO DAN INFORMATIKA CERDAS**

**INSTITUT TEKNOLOGI SEPULUH NOPEMBER**

**2021**

Metodologi Penelitian

Penelitian ini merupakan penelitian Social Network Analysis dengan mengambil data tweet dengan kata kunci “vaksin”, “vaksinasi”, “astrazeneca” ataupun “sinovac” pada twitter dimulai dari akun sumber yaitu @prfmnews yang merupakan akun twitter salah satu radio yang memberikan informasi seputar Bandung dengan jumlah pengikut mencapai 261 ribu pengguna dan telah bergabung pada twitter mulai Juni 2010. Dari setiap data tweet yang terkumpul akan dicari hubungan interaksi yang ada berupa like dan retweet. Untuk memperkaya dataset yang dimiliki, maka dari hubungan interaksi yang didapat akan dilakukan teknik snowball sampling dimana akan dipilih akun-akun yang berpotensi yaitu akun yang pernah melakukan interaksi baik like maupun retweet lebih dari sekali terhadap akun lain yang ada akan dijadikan sebagai akun sumber berikutnya. Pada snowball sampling ini, kami membatasi untuk kedalaman dari sampling hingga kedalaman 3 tingkat dikarenakan pada kedalaman ini, data yang diperoleh dirasa sudah mencukupi untuk dianalisis.

Dataset yang terkumpul akan dilakukan visualisasi dan diekstraksi knowledgenya dengan menerapkan analisis untuk sentralitas, aktor populer dan klasifikasi aktor-aktor yang terlibat terhadap aktor yang berperan penting dalam penyebaran informasi. Pada analisis sentralitas ini diterapkan empat jenis sentralitas sebagai berikut.

* Degree Centrality

Degree centrality digunakan untuk mencari aktor yang menjadi “celebrities”. Degree centrality ini menunjukkan jumlah koneksi yang dimiliki oleh node.

* Eigenvector Centrality

Eigenvector centrality merupakan versi “rekursif” dari Degree centrality. Eigenvector centrality ini memberikan bobot yang lebih tinggi ke node yang memiliki keterhubungan yang lebih tinggi.

* Closeness Centrality

Closeness centrality menunjukkan node-node yang dapat dikatakan sebagai “Gossip Mongers” pada jaringan tersebut. Closeness centrality menunjukkan kedekatan suatu individu dengan individu yang lainnya.

* Betweenness Centrality

Betweenness centrality digunakan untuk mencari boundary spanners atau node yang menjadi penghubung dari 2 atau lebih komunitas yang ada.

Sedangkan untuk penentuan aktor populer digunakan perhitungan menggunakan pagerank. Pagerank merupakan perhitungan sentralitas yang didasarkan pada in-degree dari tiap node sehingga dapat menghasilkan aktor-aktor yang merupakan aktor kunci pada jaringan sosial tersebut.

Kemudian untuk analisis selanjutnya adalah analisis pada komunitas yang ada. Analisis pada komunitas ini ditujukan agar dapat mengetahui bagaimana kedekatan dari aktor populer terhadap aktor yang lainnya. Aktor kunci yang diperoleh pada analisis sentralitas akan dijadikan sebagai acuan untuk menjadi atribut kelas yang digunakan pada pengelompokan aktor-aktor yang ada. Pengelompokan aktor ini akan menggunakan teknik semi-supervised learning dengan menggunakan gaussian field dan harmonic function. Pemilihan teknik klasifikasi ini didasarkan pada kemampuannya yang telah digunakan untuk melakukan pengelompokan berbagai macam kasus seperti kasus random walks, electric network dan spectral graph theory. Selain itu, dalam penelitian ini juga dilakukan analisis terhadap subgraph yang ada yaitu dengan mencari triads yang ada. Analisis dengan pencarian triads ini memungkinkan untuk melihat bagaimana tingkat kekuatan hubungan dari aktor-aktor yang ada dan juga bagaimana tingkah laku aktor-aktor tersebut di dalam jejaring sosial twitter.

Analisis selanjutnya merupakan analisis yang dilakukan terhadap jejaring sosial dengan data yang sama dengan data yang terkumpul sebelumnya namun dengan melakukan pembentukan graf dengan 2 mode pada node-nya yaitu node yang merepresentasikan akun user dan node yang merepresentasikan tweet user. Dari graf ini, maka kita dapat mengetahui user manakah yang memiliki banyak tweet yang banyak direspon oleh user lain. Adapun untuk gambaran diagram aliran yang digunakan dalam penelitian ini terlihat pada diagram berikut.



.

Gambar 1. Diagram alir penelitian yang digunakan

Hasil dan Pembahasan

Penelitian ini terdiri dari 2 kali rentang waktu pengambilan data pada twitter dengan rincian sebagai berikut.

* Pengambilan data pertama (batch-1) diambil untuk rentang tanggal 1 Agustus 2020 hingga 31 Januari 2021.
* Pengambilan data kedua (batch-2) diambil untuk rentang tanggal 15 Februari 2021 hingga 15 Mei 2021.

Pada setiap rentang pengambilan, pengambilan data dilakukan dengan menggunakan snowball sampling membentuk dataset yang digunakan untuk visualisasi. Pada pengambilan data ini digunakan library twint atau twitter intelligence tool yang mampu untuk mengambil data tweet dari akun sumber tertentu dengan menerapkan beberapa filter yang dibutuhkan yaitu rentang tanggal pengambilan data dan kata kunci yang digunakan. Dari penggunaan twint ini menghasilkan data-data tweet yang terkait dan telah disimpan pada file dengan format json sehingga langkah selanjutnya adalah mencari interaksi yang terhadap tweet-tweet tersebut. Penggunaan twint memiliki batasan yaitu tidak dapat digunakan untuk mencari interaksi like dan retweet yang ada sehingga peneliti menggunakan web scraping dengan menggunakan library selenium untuk mencari interaksi like dan retweet-nya. Untuk visualisasinya graf-nya sendiri, digunakan library dari networkx dimana akan dibentuk graf berarah dengan node yang merepresentasikan akun twitter dan edge yang merepresentasikan interaksi yang ada baik merupakan interaksi like dan retweet oleh satu akun twitter ke akun twitter lainnya.

Pada penelitian ini untuk tujuan mempermudah pengamatan, maka setiap node yang terbentuk diberi sebuah ukuran dimana ukuran ini sebanding dengan ukuran degree masuknya dan juga setiap node memiliki warna untuk membedakan ukurannya dari warna gelap untuk node dengan ukuran yang kecil hingga warna cerah untuk node dengan ukuran besar. Sedangkan untuk edge juga memiliki bobot yang sebanding dengan banyaknya interaksi berupa jumlah like dan retweet antara suatu akun dengan akun lainnya. Untuk penempatan penggambaran nodenya, kami menggunakan algoritma *ForceAtlas2* dimana algoritma ini merupakan algoritma tata letak ruang yang mengelompokkan node-node berdasarkan komunitasnya sehingga graf lebih mudah untuk diamati. Selain itu, dikarenakan jumlah data yang diambil cukup banyak, maka perlu dilakukan pembatasan pada visualisasinya. Pembatasan yang diberlakukan adalah dengan membatasi untuk menampilkan hanya akun-akun yang memiliki bobot edge lebih dari ambang batas yang ditentukan. Ambang batas yang diambil bergantung pada banyak data yang ada dimana semakin banyak data yang didapat akan digunakan batas yang lebih besar begitu pula sebaliknya. Dari hasil percobaan yang dilakukan, kami memilih untuk menerapkan bobot minimal 5 pada percobaan batch-1 dan bobot minimal 2 untuk percobaan pada batch-2.

1. Analisis Sentralitas pada Batch-1

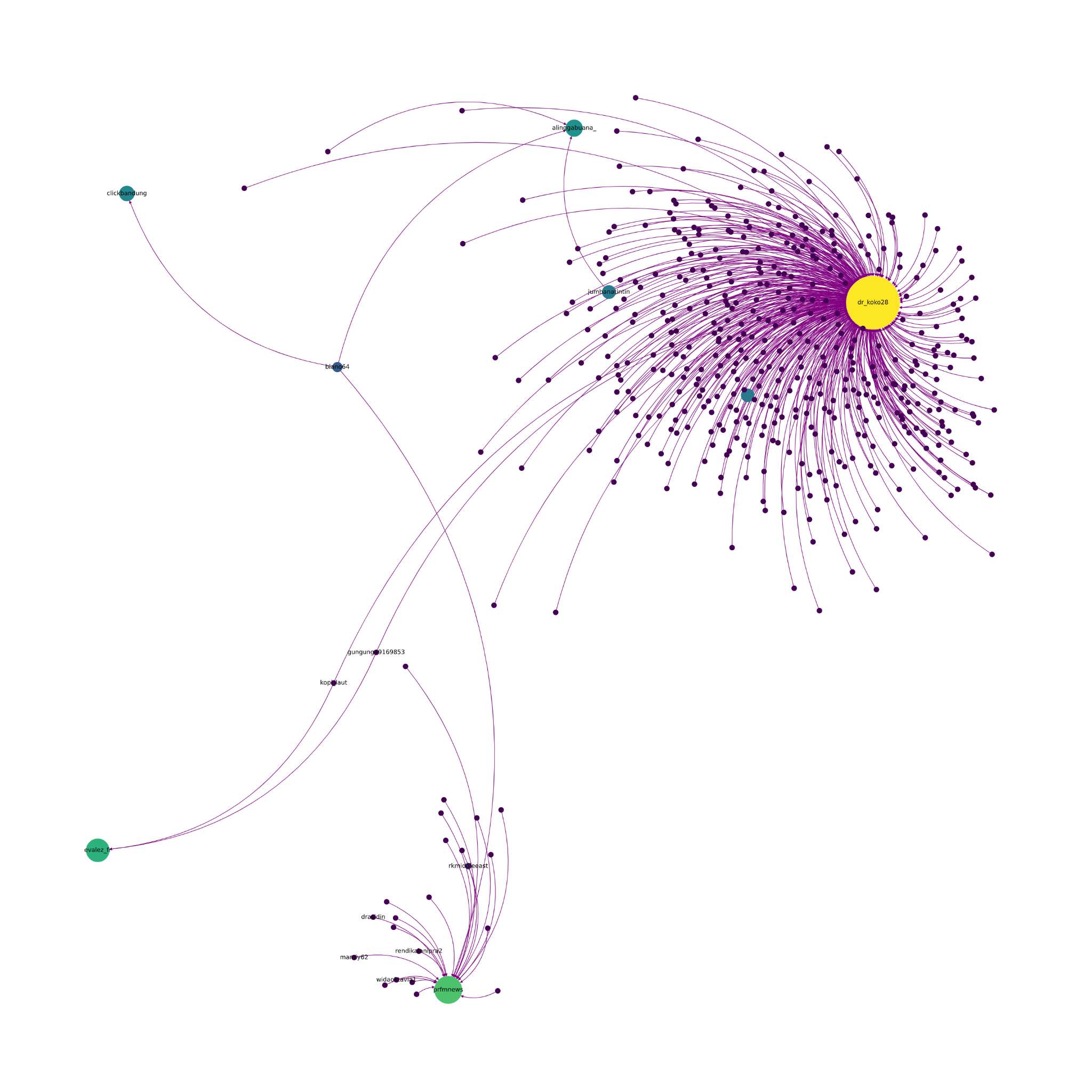
Pada pengambilan data batch-1 dengan menggunakan twint dan selenium menghasilkan data total sesuai dengan tabel berikut.

Tabel 1. Jumlah data yang didapat berdasarkan kedalam sampling pada batch-1.

|  |  |  |  |
| --- | --- | --- | --- |
| Kedalaman | Tweet | Nodes | Edges |
| 1 | 258 | 501 | 500 |
| 2 | 465 | 680 | 705 |
| 3 | 806 | 6833 | 7080 |

Pada akun sumber @prfmnews diperoleh 258 tweet dengan 501 nodes dan 500 edges dimana semua interaksi merupakan interaksi like dan retweet yang dilakukan hanya antara pengguna dengan akun prfmnews. Kemudian pada kedalaman kedua kita lakukan pengambilan kembali sehingga total tweet menjadi 465 tweet dengan 680 nodes dan 705 edges dan kedalaman ketiga, berubah menjadi total 806 tweet dengan 6833 nodes dan 7080 edges.

Dari keseluruhan data yang terkumpul, dilakukan pembersihan dengan menerapkan pembatasan bobot edge dengan bobot minimal 5 sehingga jaringan menjadi hanya terdiri dari 484 nodes dengan 485 edges. Untuk visualisasi jaringan yang terbentuk setelah pembersihan dapat dilihat pada gambar berikut.



Gambar 2. Jaringan sosial yang terbentuk pada batch-1 setelah dilakukan pembatasan bobot

Pada visualisasi tersebut, nodes memiliki ukuran dan warna yang merepresentasikan banyaknya interaksi like dan retweet yang terjadi pada tweet akun node. Semakin besar ukuran node dan semakin cerah warna pada nodes menunjukkan node tersebut memiliki interaksi yang lebih banyak dari akun lainnya. Kemudian dari hasil visualisasi tersebut, kami melakukan beberapa perhitungan untuk sentralitas yang jaringan. Untuk tabel perhitungan nilai sentralitas dapat dilihat pada tabel 2.

Tabel 2. Hasil perhitungan sentralitas batch-1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Username** | **Degree** | **Eigenvector** | **Closeness** | **Betweenness** |
| dr\_koko28 | 455 | 0.000312 | 0.942029 | 0.000000 |
| prfmnews | 23 | 0.000016 | 0.047619 | 0.000000 |
| alinggabuana\_ | 5 | 0.999999 | 0.006211 | 0.000000 |
| evalez\_fr | 2 | 0.000001 | 0.004141 | 0.000000 |
| clickbandung | 1 | 0.000001 | 0.002070 | 0.000000 |
| biano64 | 3 | 0.000000 | 0.000000 | 0.000000 |

Tabel 2 merupakan tabel hasil perhitungan sentralitas pada batch-1. Dalam tabel ini memuat beberapa jenis sentralitas seperti *degree centrality, eigenvector centrality, closeness centrality*, dan *betweenness centrality*. Pada tabel tersebut dilakukan pengurutan secara *descending* terhadap *closeness centrality,* dilanjutkan oleh *eigenvector centrality* kemudian *degree centrality* dan terakhir adalah *betweenness centrality*.

Pada pengukuran pertama yaitu *degree centrality*, didapatkan bahwa akun dengan nama @dr\_koko28 memiliki *degree centrality* yang paling besar yaitu dengan 455 *degree* diikuti oleh akun @prfmnews yang merupakan akun sumber asal dengan perolehan 23 *degree*, akun @alinggabuana\_ dengan 5 *degree*, kemudian @biano dengan 3 *degree* dan akun evalez\_fr dengan 2 *degree*. Dari hasil perhitungan ini menunjukkan bahwa akun @dr\_koko28 merupakan aktor dengan koneksi terbanyak dalam jaringan ini. Ini juga menunjukkan bahwa akun tersebut merupakan akun yang memiliki pengaruh yang paling besar terhadap akun lainnya.

Pengukuran kedua didasarkan pada *eigenvector centrality*. Bila dilihat dari pengukuran ini, nilai terbesar diperoleh oleh akun @alinggabuana\_ dengan eigenvector centralitynya mencapai nilai 0.999999 kemudian disusul oleh akun @dr\_koko28 dengan nilai 0.000312 dan @prfmnews dengan nilai 0.000016. Untuk akun lainnya hanya memiliki nilai perhitungan *eigenvector* yang sangat kecil atau kurang dari nilai 0.00002. Perhitungan *eigenvector* ini merupakan menunjukkan keterhubungan oleh suatu akun dengan akun lainnya. Dari sini terlihat bahwa akun @alinggabuana\_ memiliki keterhubungan yang sangat besar dengan akun lainnya.

Pengukuran ketiga merupakan pengukuran dari *closeness centrality*. Pada *closeness centrality* ini terlihat bahwa nilai kedekatan yang tertinggi dimiliki oleh akun dengan nama @dr\_koko28 dengan nilai 0.942029 kemudian dilanjutkan oleh akun @prfmnews dengan nilai 0.047619, akun @alinggabuana\_ dengan nilai 0.006211, akun @evalez\_fr dengan nilai 0.004141 , dan akun @clickbandung dengan nilai 0.00207. Nilai sentralitas ini menunjukkan seberapa dekat suatu aktor dengan aktor lainnya dan terlihat bahwa akun @dr\_koko28 merupakan akun yang paling dekat dengan akun-akun lainnya.

Pengukuran sentralitas terakhir didasarkan oleh perhitungan *betweenness centrality*. Pada pengukuran ini terlihat bahwa seluruh akun memiliki nilai *betweenness centrality* sebesar 0. Pengukuran ini akan menunjukkan seberapa tingkat suatu akun akan menyebarkan informasi ke akun lainnya. Dengan begitu, dari hasil ini menunjukkan tidak ada akun yang berperan secara langsung sebagai penyebar informasi dari suatu akun terhadap akun lainnya.

Tabel 3. Hasil perhitungan ranking popularitas aktor batch-1

|  |  |
| --- | --- |
| **Username** | **Pagerank** |
| dr\_koko28 | 0.425878 |
| prfmnews | 0.021853 |
| alinggabuana\_ | 0.019263 |
| evalez\_fr | 0.0023584 |
| clickbandung | 0.001608 |
| biano64 | 0.001104 |

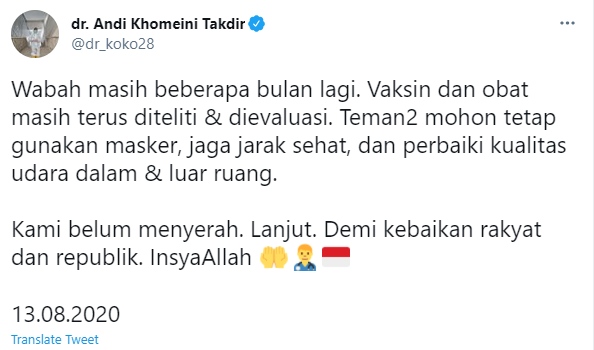
Tabel 3 merupakan tabel hasil perhitungan popularitas aktor pada batch-1. Untuk mengukur popularitas, dilakukan perhitungan menggunakan algoritma *pagerank* sehingga terlihat tingkat popularitas tiap akun kemudian dilakukan pengurutan. Dari tabel tersebut terlihat bahwa terdapat akun dengan popularitas yang sangat jauh tingginya dibandingkan dengan akun lainnya yaitu akun @dr\_koko28 dengan nilai *pagerank* sebesar 0.425878. Pada posisi setelahnya diperoleh akun @prfmnews dengan nilai 0.021853, akun @alinggabuana dengan nilai 0.019263 dan akun @evalez\_fr dengan popularitas 0.0023584. Dari hasil peringkat ini, maka bisa diambil akun-akun yang menjadi aktor kunci untuk menentukan komunitas yang ada yaitu @dr\_koko28, @prfmnews, @alinggabuana, dan @evalez\_fr.

Dari perhitungan sentralitas, terlihat bahwa akun @dr\_koko28 merupakan akun yang sangat mencolok dalam jaringan sosial ini. Untuk setiap nilai yang diperoleh oleh akun @dr\_koko28 ini terdapat perbedaan yang sangat besar bahkan bila dibandingkan dengan akun sumber utama yang diamati yaitu @prfmnews. Hal ini membuat kami merasa bahwa akun ini perlu dianalisis lebih jauh lagi.



Gambar 3. Profil akun twitter @dr\_koko28

Dari penelusuran lebih lanjut, kami menemukan bahwa akun @dr\_koko28 merupakan akun dengan nama dr. Andi Khomeini Takdir merupakan seorang dokter spesialis penyakit dalam dan merupakan pendiri dari JDN (Junior Doctors Network) Indonesia yang memiliki jumlah followers mencapai 171 ribu. Selain memiliki jumlah follower yang banyak, akun @dr\_koko28 ini cukup sering untuk membagikan tweet seputar informasi terbaru mengenai covid-19 dan juga membagikan tips seputar cara untuk menjaga pola hidup sehat terlebih lagi pada masa pandemi.



Gambar 4. Salah satu tweet oleh @dr\_koko28 terkait update informasi vaksin

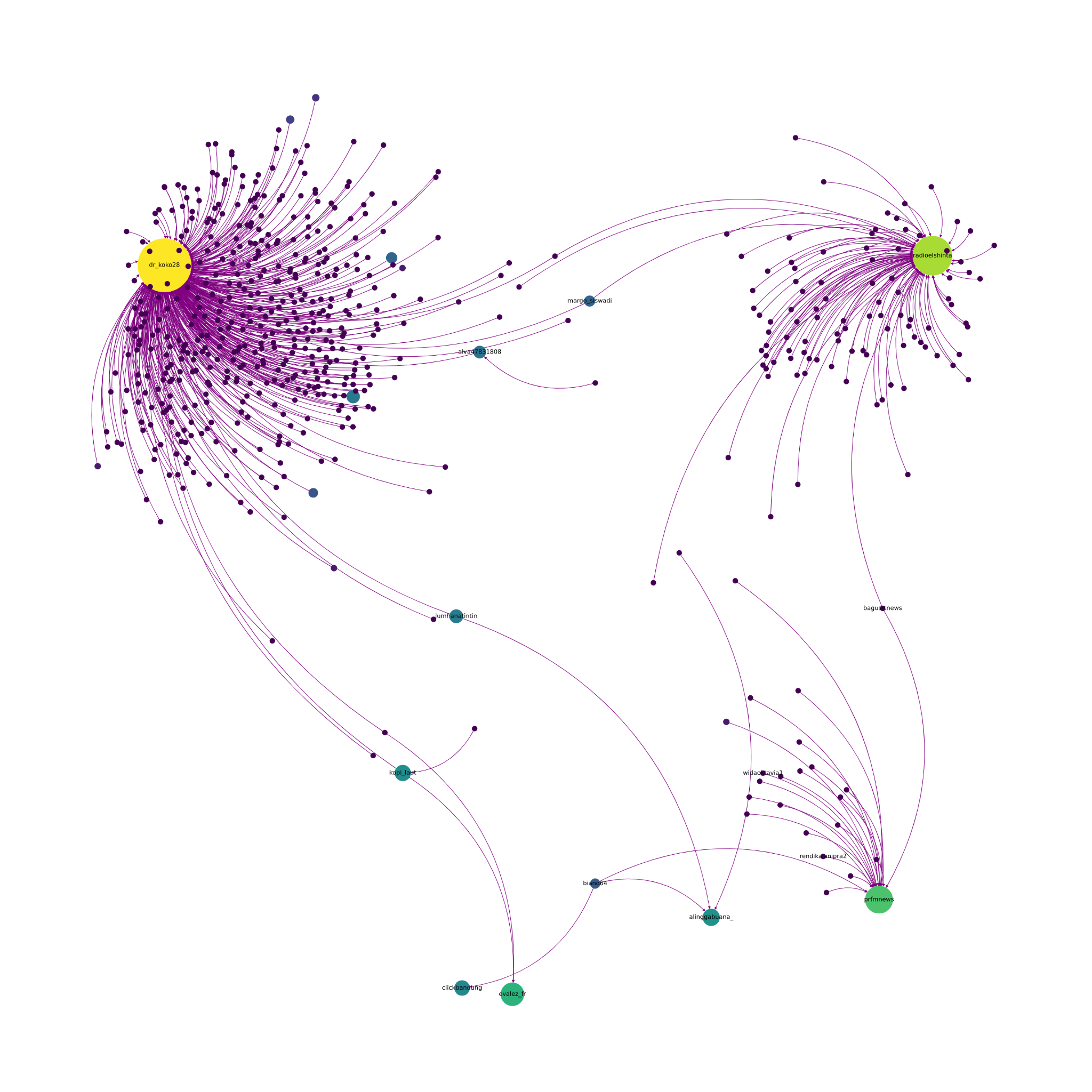
Pada gambar 4, merupakan salah satu contoh tweet yang dilakukan oleh @dr\_koko28. Tweet ini dipublikasikan pada 13 Agustus 2020 pada 6:45 AM waktu SE Asia Standard Time atau UTC +7 melalui Twitter android. Tweet ini sendiri telah di likes oleh 510 orang dan dilakukan retweet 84 kali per 30 Mei 2021. Dalam tweet ini, didapatkan informasi bahwa vaksin dan obat terkait covid masih akan diteliti. Selain membagikan update informasi, tweet ini juga mengajak masyarakat untuk selalu menerapkan protokol kesehatan.

Dari hasil pengamatan ini, kami mencoba untuk menggali lebih dalam pada node @dr\_koko dengan menerapkan teknik snowball sampling dengan akun ini sebagai akun awal. Seperti pengambilan data sebelumnya, interaksi yang digunakan masih sama yaitu menggunakan like dan retweet sebagai dasar acuan. Dari pengambilan sampel ini diperoleh data sesuai pada tabel 4.

Tabel 4. Jumlah data yang didapat berdasarkan kedalaman sampling dengan akun awal @dr\_koko28.

|  |  |  |  |
| --- | --- | --- | --- |
| Kedalaman | Tweet | Nodes | Edges |
| 1 | 229 | 5853 | 5852 |
| 2 | 572 | 5977 | 6015 |
| 3 | 1455 | 7776 | 7930 |

Pada tabel 4, dilakukan pengambilan data menggunakan snowball sampling dengan akun awal @dr\_koko28 dan mendapati total tweet diperoleh sebesar 1455 tweet dengan 7776 nodes dan 7930 edges. Dari hasil ini, setelah dilakukan penggabungan dengan dataset dari pengambilan data batch-1, dihasilkan jumlah total node dan edge pada jaringan sosial yang terbentuk sebesar 8697 nodes dan 9158 nodes. Langkah selanjutnya yang dilakukan adalah membentuk visualisasi jaringan seperti pada gambar 5.



Gambar 5. Jaringan sosial setelah dilakukan snowball sampling lanjutan pada akun @dr\_koko28.

Pada gambar 5, telah terlihat bagaimana visualisasi yang diperoleh dengan melakukan penggabungan snowball sampling kedua pada batch-1 dengan snowball sampling awal pada batch-1. Nodes pada jaringan sosial tersebut memiliki ukuran dan warna berdasarkan *InDegree* yang dimiliki dimana akan berukuran besar dan cerah ketika nilai *InDegree*-nya besar dan akan berukuran kecil dan gelap ketika nilai *InDegree*-nya kecil. Pada visualisasi ini juga telah dilakukan pembatasan terhadap edge yang terbentuk dengan menerapkan pembobotan minimal sebesar 5 interaksi untuk mempermudah pengamatan. Dari pembatasan bobot ini, jumlah node dan edge yang ditampilkan masing-masing sebanyak 586 nodes dengan 589 edges. Langkah selanjutnya adalah melakukan perhitungan sentralitas.

Tabel 5. Hasil perhitungan sentralitas pada batch-1 setelah dilakukan sampling kembali pada akun @dr\_koko28

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **username** | **degree** | **eigenvector** | **closeness** | **betweenness** |
| dr\_koko28 | 455 | 0.000327 | 0.777793 | 0.000000 |
| radioelshinta | 102 | 0.000070 | 0.174359 | 0.000000 |
| prfmnews | 23 | 0.000016 | 0.039316 | 0.000000 |
| alinggabuana\_ | 5 | 0.999999 | 0.005128 | 0.000000 |
| evalez\_fr | 2 | 0.000009 | 0.003846 | 0.000000 |
| kopi\_laut | 3 | 0.000001 | 0.001709 | 0.000006 |
| alva47831808 | 2 | 0.000001 | 0.001709 | 0.000003 |
| clickbandung | 1 | 0.000001 | 0.001709 | 0.000000 |
| biano64 | 3 | 0.000000 | 0.000000 | 0.000000 |
| bagusrtnews | 2 | 0.000000 | 0.000000 | 0.000000 |

Tabel 5 merupakan tabel hasil perhitungan sentralitas pada batch-1 setelah data ditambahkan dengan data dari akun awal @dr\_koko28. Sama seperti pada tabel 2, tabel ini memuat beberapa jenis sentralitas seperti *degree centrality, eigenvector centrality, closeness centrality*, dan *betweenness centrality*. Pada tabel tersebut dilakukan pengurutan secara *descending* terhadap *closeness centrality,* dilanjutkan oleh *eigenvector centrality* kemudian *degree centrality* dan terakhir adalah *betweenness centrality*.

Dari pengukuran *degree centrality*, ada beberapa node baru yang terbentuk yaitu node akun @radioelshinta dengan degree sebesar 102, @kopi\_laut dengan nilai degree 3, @alva47831808 dengan nilai degree 2, dan bagusrtnews dengan nilai degree 2. Pada pengukuran ini, tetap akun @dr\_koko menjadi akun yang memiliki pengaruh paling besar, namun terdapat tambahan pula oleh akun @radioelshinta yang memiliki pengaruh yang cukup besar.

Pengukuran kedua didasarkan pada *eigenvector centrality*. Bila dilihat dari pengukuran ini, juga tidak terdapat perubahan nilai dari akun akun yang sebelumnya telah tercantum pada tabel 2, hanya saja terdapat nilai baru yaitu untuk akun @radioelshinta dengan nilai 0.00009 yang juga masih terbilang kecil bila dibandingkan oleh akun @alinggabuana\_.

Pengukuran ketiga merupakan pengukuran dari *closeness centrality*. Pada pengukuran ini, perubahan jumlah node yang terlibat menyebabkan perubahan nilai dari *closeness centrality.* Pada akun @dr\_koko28 *closeness centrality*-nya berubah dari nilai 0.942029 menjadi 0.777793, akun @radioelshinta menjadi memiliki nilai 0.174359, akun @prfmnews berkurang dari 0.047619 menjadi 0.039316, akun @alinggabuana\_ berubah dari 0.006211 menjadi 0.005128, akun @evalez\_fr dari 0.004141 menjadi 0.003846 , dan akun @clickbandung dengan nilai 0.00207. Dari sini terlihat bahwa selain akun @dr\_koko, akun @radio\_elshinta memiliki kedekatan juga dengan node lainnya.

Pengukuran sentralitas terakhir didasarkan oleh perhitungan *betweenness centrality*. Pada pengukuran ini sebelumnya, semua nilai pada node bernilai 0 atau tidak ada influencer sedangkan pada pengukuran kedua ini, terdapat 2 akun yang berperan sebagai influencer meskipun nilainya sangat kecil yaitu akun @kopi\_laut dengan nilai 0.000006 dan akun @alva47831808 dengan nilai 0.000003.

Tabel 6. Hasil perhitungan ranking popularitas aktor pada batch-1 setelah dilakukan sampling kembali pada akun @dr\_koko28

|  |  |
| --- | --- |
| **Username** | **Pagerank** |
| dr\_koko28 | 0.352196 |
| radioelshinta | 0.078864 |
| prfmnews | 0.017765 |
| alinggabuana\_ | 0.015932 |
| evalez\_fr | 0.002683 |
| alva47831808 | 0.001689 |
| kopi\_laut | 0.001689 |
| clickbandung | 0.001330 |

Tabel 6 menjelaskan data popularitas yang didapatkan setelah melakukan penambahan data sampling. Terdapat sebuah akun yang masuk dalam daftar popularitas yaitu akun @radioelshinta dengan nilai *pagerank* sebesar 0.078864.

1. Analisis Komunitas pada Batch-1

Setelah melakukan analisis pada sentralitas pada jaringan sosial batch-1, perlu dilakukan analisis terhadap hubungan dari tiap aktor yang berpengaruh terhadap aktor lainnya atau melakukan analisis pada komunitas. Analisis pada komunitas pada penelitian ini dilakukan dengan 2 cara. Cara pertama adalah dengan melakukan perhitungan *harmonic function* terhadap akun-akun populer sehingga dapat terlihat kedekatan tiap akun yang ada terhadap tiap akun popular yang ada. Cara kedua adalah dengan melakukan analisis pada triad yang ada pada jaringan sosial yang terbentuk.

Pada cara pertama ini, merupakan analisis dengan memperhatikan nilai yang telah didapatkan oleh perhitungan sentralitas. Pada perhitungan sebelumnya, kami telah menemukan aktor-aktor yang menjadi aktor kunci untuk menentukan komunitas yang ada. Pada awalnya terdapat aktor-aktor populer seperti yang terdapat pada tabel 3. Dengan menerapkan klasifikasi menggunakan harmonic function, maka dapat diperoleh komunitas yang ada seperti pada tabel 7 berikut.

Tabel 7. Pengelompokan nodes berdasarkan komunitas pada batch-1

|  |  |
| --- | --- |
| **Aktor kunci** | **Nodes pada komunitas** |
| dr\_koko28 | 453 |
| prfmnews | 23 |
| alinggabuana\_ | 5 |
| evalez\_fr | 3 |

Dari perolehan komunitas yang ada, bila digambarkan dengan menggunakan jaringan sosial untuk tiap komunitas yang terbentuk berdasarkan warna akan ditunjukkan pada gambar 6 berikut.



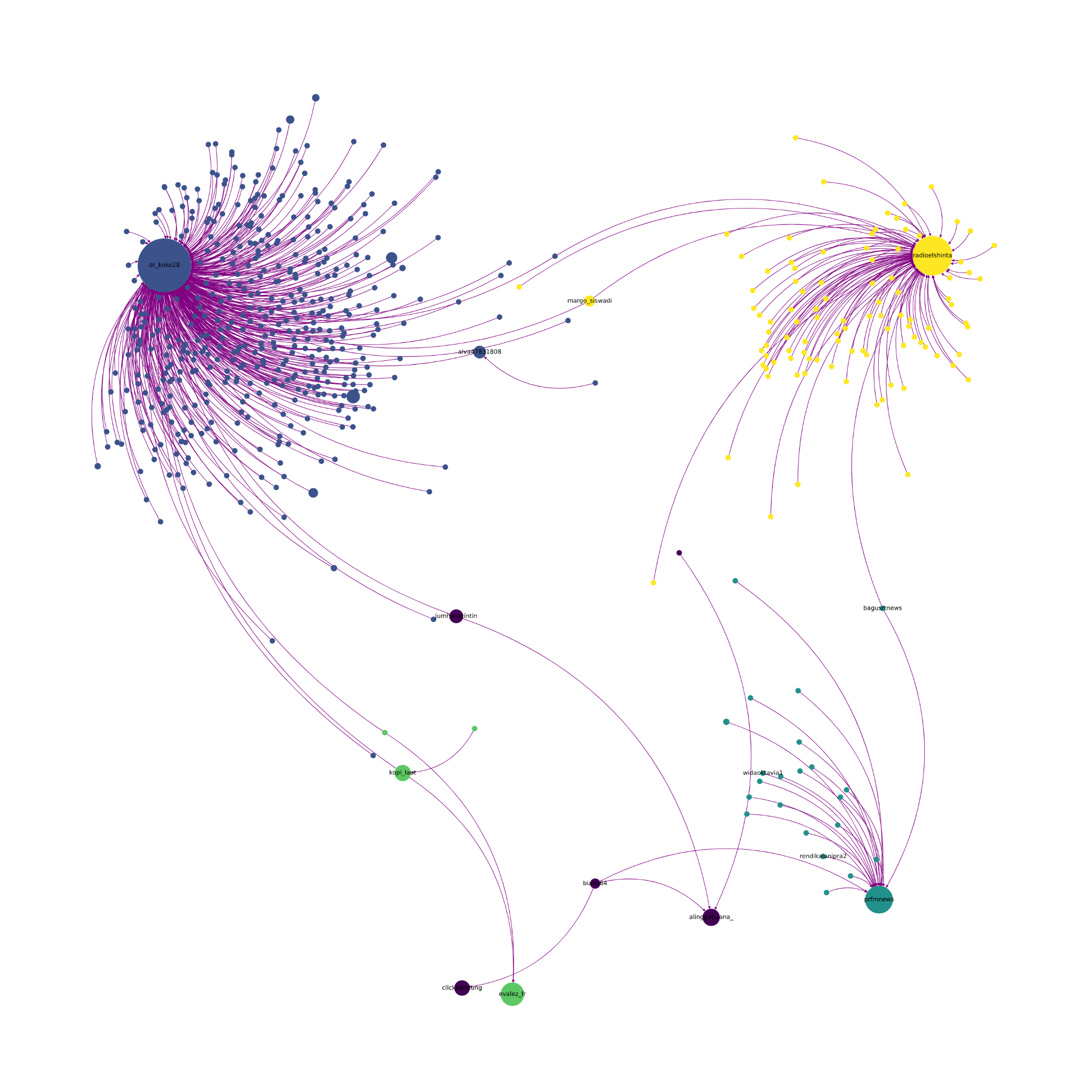
Gambar 6. Pengelompokan node pada jaringan sosial batch-1

Selanjutnya kami melakukan perhitungan yang sama dengan perhitungan pada tabel 7 dengan data yang telah ditambahkan oleh sampling terhadap akun @dr\_koko28. Dengan menggunakan hasil pada hasil perhitungan popularitas menggunakan *pagerank* yang terdapat pada tabel 6, maka kami menambahkan akun-akun yang merupakan akun populer sebagai akun yang dapat menjadi pusat komunitas pada jaringan sosial yang terbentuk. Untuk komunitas yang terbentuk saat ini berubah dari data pada tabel 7 menjadi jumlah yang sesuai pada tabel 8.

Tabel 8. Pengelompokan nodes berdasarkan komunitas setelah akun @dr\_koko28 dilakukan sampling tambahan pada batch-1

|  |  |
| --- | --- |
| **Aktor kunci** | **Nodes pada komunitas** |
| dr\_koko28 | 453 |
| radioelshinta | 101 |
| prfmnews | 23 |
| alinggabuana\_ | 5 |
| evalez\_fr | 4 |

Pada tabel 8, terbentuk sebuah komunitas baru dengan aktor kunci yaitu akun @radioelshinta dengan jumlah komunitas sebesar 101 node. Selain itu juga terjadi sedikit perubahan pada jumlah komunitas yang terbentuk sebelumnya. Perubahan yang terjadi adalah pada komunitas dengan aktor kunci @evalez\_fr yang bertambah dari yang sebelumnya berjumlah 3 kini menjadi 4.



Gambar 7. Pengelompokan aktor setelah akun @dr\_koko28 dilakukan snowball sampling lanjutan.

Dari perolehan komunitas yang ada, bila digambarkan dengan menggunakan jaringan sosial untuk tiap komunitas yang terbentuk berdasarkan warna akan ditunjukkan pada gambar 7. Pada gambar ini, setiap node dalam komunitas yang sama akan memiliki warna yang sama dengan total terdapat 5 jenis komunitas yang terbentuk.

Tabel 9. Hasil analisis Triad pada pengambilan batch-1

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 003 | 012 | 021C | 021D | 021U | 030C | 030T | 102 |
| 33132153 | 125982 | 3 | 9 | 108693 | 0 | 0 | 0 |
| 111D | 111U | 120C | 120D | 120U | 201 | 210 | 300 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Pada deteksi komunitas kedua adalah dengan menganalisis berdasarkan hubungan triad yang terdapat pada jaringan. Dari tabel 9, telah diperoleh hasil analisis triad pada jaringan batch-1. Terlihat bahwa triad dengan code 003 sangat besar yaitu sebanyak 33.132.153 pasangan triad yang memiliki kode 003, 125.982 pasangan triad yang memiliki kode 012, dan 108.693 kode 021U, 9 pasangan triad dengan kode 021D dan 3 pasangan triad dengan kode 021C. Pada hasil sensus ini, terdapat 33.366.837 triad yang termasuk jenis *vacuously transitive*, 3 triad yang termasuk dalam kategori *intransitive* dan 0 triad yang termasuk dalam kategori *transitive*. Untuk jenis *vacuously transitive*, selain triad dengan kode 021U dan 021D tidaklah menarik dimana hanya merupakan interaksi dyad dan interaksi yang tidak berhubungan. Disini terlihat dari 021U menunjukkan sangat banyak aktor egosentris yang sering menerima perhatian banyak aktor tetapi tidak pernah membalas dan pada 021D menunjukkan hanya sedikit aktor yang berinteraksi dengan aktor lain tetapi tidak mendapatkan balasan. Untuk jenis *intransitive* menunjukkan bahwa hanya sedikit aktor yang membangun interaksi dengan melalui perantara. Dan dari jaringan ini terlihat untuk jenis *transitive* tidak ada aktor yang memiliki hubungan kuat dalam interaksinya.

1. Analisis Sentralitas pada Batch-2

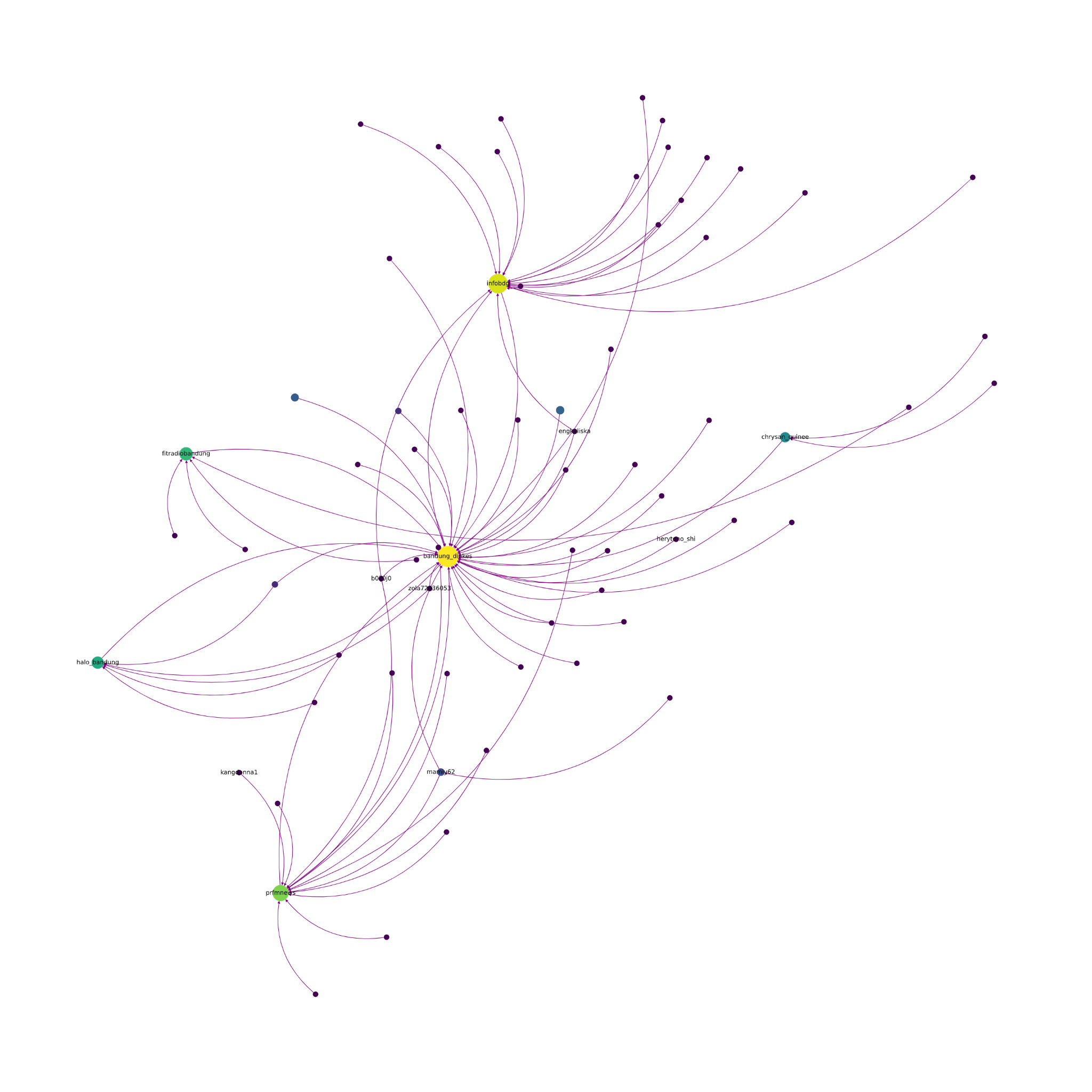
Pada pengambilan data batch-2 dengan menggunakan twint dan selenium menghasilkan data total sesuai dengan tabel berikut.

Tabel 10.Jumlah data yang didapat berdasarkan kedalam sampling pada batch-2

|  |  |  |  |
| --- | --- | --- | --- |
| Kedalaman | Tweet | Nodes | Edges |
| 1 | 156 | 68 | 67 |
| 2 | 239 | 198 | 204 |
| 3 | 580 | 380 | 405 |

Pada akun sumber @prfmnews diperoleh 156 tweet dengan 58 nodes dan 67 edges dimana semua interaksi merupakan interaksi like dan retweet yang dilakukan hanya antara pengguna dengan akun prfmnews. Kemudian pada kedalaman kedua kita lakukan pengambilan kembali sehingga total tweet menjadi 239 tweet dengan 198 nodes dan 204 edges dan kedalaman ketiga, berubah menjadi total 580 tweet dengan 380 nodes dan 405 edges.

Dari keseluruhan data yang terkumpul, dilakukan pembersihan dengan menerapkan pembatasan bobot edge dengan bobot minimal 5 sehingga jaringan menjadi hanya terdiri dari 68 nodes dengan 76 edges. Untuk visualisasi jaringan yang terbentuk setelah pembersihan dapat dilihat pada gambar 8.



Gambar 8. Jaringan sosial yang terbentuk pada batch-2

Setelah dilakukan visualisasi, tahap selanjutnya adalah analisis sentralitas dari jaringan sosial pada batch-2. Analisis sentralitas ini menggunakan 4 jenis sentralitas yaitu *degree centrality, eigenvector centrality, closeness centrality*, dan *betweenness centrality.* Untuk hasil perhitungannya telah dijabarkan pada tabel 11.

Tabel 11. Hasil perhitungan sentralitas pada batch-2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Username** | **Degree** | **Eigenvector** | **Closeness** | **Betweenness** |
| bandung\_dinkes | 36 | 0.707109 | 0.663366 | 0.037087 |
| infobdg | 19 | 0.408252 | 0.496296 | 0.010176 |
| prfmnews | 14 | 0.408253 | 0.465278 | 0.006784 |
| halo\_bandung | 6 | 0.408237 | 0.416149 | 0.001357 |
| fitradiobandung | 5 | 0.000005 | 0.059701 | 0.003618 |
| chrysan\_quinee | 3 | 0.000002 | 0.029851 | 0.001809 |
| mamiy62 | 3 | 0.000001 | 0.014925 | 0.000905 |
| b0ij0j0 | 3 | 0.000000 | 0.000000 | 0.000000 |
| zola72036053 | 2 | 0.000000 | 0.000000 | 0.000000 |
| engkaliska | 2 | 0.000000 | 0.000000 | 0.000000 |

Dari tabel 9 terdapat 4 jenis sentralitas yang dibahan. Jenis yang pertama adalah *degree centrality*. *Degree* tertinggi didapatkan oleh akun @bandung\_dinkes dengan nilai sebesar 36 kemudian @infobgd dengan nilai sebesar 19, @prfmnews dengan nilai sebesar 14 dan nilai *degree* lainnya memiliki kurang dari 10. Dari data ini terlihat bahwa akun @bandung\_dinkes merupakan akun dengan pengaruh terbesar pada batch-2.

Jenis sentralitas kedua yang dihitung adalah *eigenvector centrality* dimana merupakan versi rekursif dari *degree centrality*. Dari hasil ini didapatkan bahwa keterhubungan tertinggi dimiliki oleh akun @bandung\_dinkes dengan nilai 0.707109, kemudian akun dengan keterhubungan yang cukup besar adalah akun @infobdg, @prfmnews, dan @halo\_bandung yang berkisar pada angka 0.408 dengan selisih yang cukup kecil.

Kemudian untuk perhitungan *closeness centrality* menunjukkan seberapa kedekatan akun dengan akun lainnya. Akun @bandung\_dinkes memiliki nilai.kedekatan yang paling tinggi yaitu mencapai 0.663366. Akun selanjutnya yang memiliki kedekatan yang cukup tinggi adalah akun @infobandung dengan nilai 0.496296, akun @prfmnews dengan nilai 0.465278 dan akun @halo\_bandung dengan nilai 0.416149. Untuk akun-akun lainnya memiliki kedekatan yang cukup rendah yaitu dengan nilai dibawah 0.06.

*Betweenness centrality* merupakan sentralitas yang mengukur node yang berpotensi sebagai akun influencer atau akun yang menyebarluaskan informasi terhadap komunitas lainnya. Dalam perhitungan *betweenness centrality* pada batch-2 ini, yang berperan sebagai penyebar informasi terbesar adalah akun @bandung\_dinkes dengan nilai 0.037087 selanjutnya oleh akun @infobdg dengan nilai 0.010176 dan akun @prfmnews dengan nilai sebesar 0.006784 sedangkan untuk akun lainnya memiliki nilai *betweenness centrality* yang kurang dari 0.002.

Tabel 12. Hasil perhitungan ranking popularitas aktor pada batch-2

|  |  |
| --- | --- |
| **Username** | **Pagerank** |
| bandung\_dinkes | 0.422519 |
| infobdg | 0.180727 |
| prfmnews | 0.173933 |
| halo\_bandung | 0.068520 |
| fitradiobandung | 0.009706 |
| chrysan\_quinee | 0.005956 |
| mamiy62 | 0.004081 |
| herytono\_shi | 0.002206 |
| b0ij0j0 | 0.002206 |
| kangdonna1 | 0.002206 |

Tabel 12 merupakan tabel hasil perhitungan popularitas aktor pada batch-2 yang telah diurutkan secara *descending*. Pada tabel ini terlihat ada 4 akun yang memiliki popularitas yang lebih tinggi dari nilai 0.01 yaitu pada akun @bandung\_dinkes dengan nilai popularitas 0.422519, akun @infobdg dengan nilai popularitas 0.180727, akun @prfm\_news dengan nilai popularitas 0.173933 dan akun @halo\_bandung dengan nilai popularitas 0.68520.

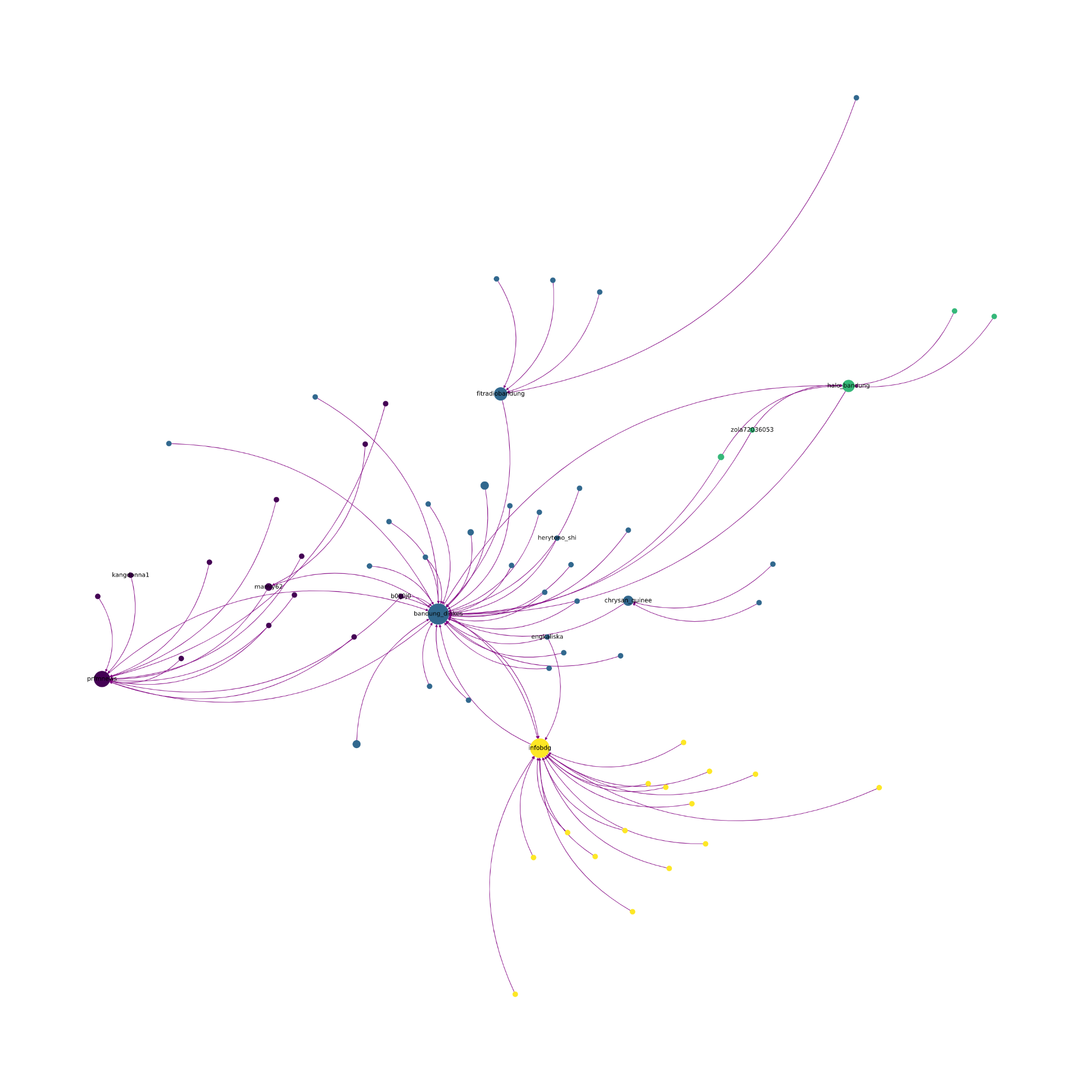
1. Analisis Komunitas pada Batch-2

Seperti halnya yang telah dilakukan pada batch-1, pada batch ini juga dilakukan analisis untuk deteksi komunitas. Analisis pertama dilakukan dengan menggunakan *harmonic function* terhadap akun akun populer. Dari akun-akun yang memiliki cukup popularitas dijadikan sebagai akun kunci untuk mencari komunitas-komunitas yang terbentuk sehingga akan terbentuk 4 komunitas dengan akun kunci yaitu @bandung\_dinkes, @infobdg, @prfmnews, dan @halo\_bandung.

Tabel 13. Pengelompokan nodes berdasarkan komunitas pada batch-2

|  |  |
| --- | --- |
| **Aktor kunci** | **Nodes pada komunitas** |
| bandung\_dinkes | 33 |
| infobdg | 16 |
| prfmnews | 14 |
| halo\_bandung | 5 |

Tabel 13 menunjukkan jumlah node yang terhubung dalam komunitas sesuai dengan aktor kunci pada batch-2. Untuk hasil visualisasi dari komunitas yang terbentuk dapat dilihat pada gambar 9 dengan node yang memiliki warna yang sama menandakan bahwa keduanya dikelompokkan pada komunitas yang sama.



Gambar 9. Pengelompokan node pada jaringan sosial batch-2

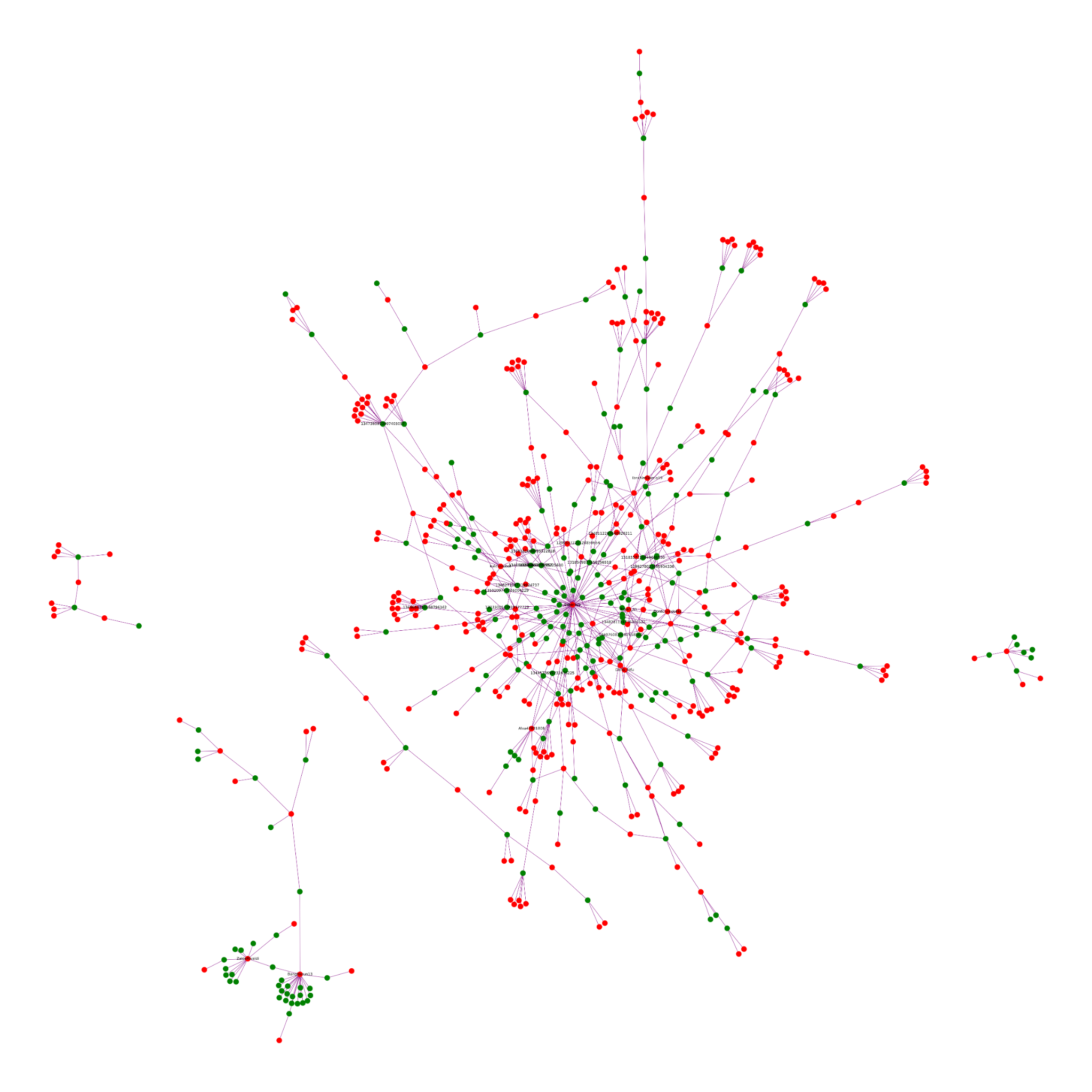
Tabel 14. Hasil analisis Triads pada pengambilan batch-2

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 003 | 012 | 021C | 021D | 021U | 030C | 030T | 102 |
| 46083 | 3179 | 8 | 1 | 650 | 0 | 0 | 75 |
| 111D | 111U | 120C | 120D | 120U | 201 | 210 | 300 |
| 111 | 0 | 0 | 6 | 0 | 3 | 0 | 0 |

Kemudian, pada cara selanjutnya adalah untuk menganalisis berdasarkan hubungan triads yang terdapat pada jaringan. Dari tabel 14, telah diperoleh hasil analisis triad pada jaringan batch-2. Terlihat bahwa triad dengan code 003 sangat besar yaitu sebanyak 46083 pasangan triad yang memiliki kode 003, 3179 pasangan triad yang memiliki kode 012, dan 650 kode 021U, 111 pasangan triad memiliki kode 111D, 75 pasangan triad memiliki kode 102, 8 pasangan triad dengan kode 021D. Pada hasil sensus ini, terdapat 49.988 triad yang termasuk jenis *vacuously transitive*, 122 triad yang termasuk dalam kategori *intransitive* dan 6 triad yang termasuk dalam kategori *transitive.* Untuk jenis *vacuously transitive*, selain triad dengan kode 021U dan 021D tidaklah menarik dimana hanya merupakan interaksi dyad dan interaksi yang tidak berhubungan. Disini terlihat dari 021U menunjukkan cukup banyak aktor egosentris yang sering menerima perhatian banyak aktor tetapi tidak pernah membalas. Pada 021D menunjukkan hanya sedikit aktor yang berinteraksi dengan aktor lain tetapi tidak mendapatkan balasan. Untuk jenis *intransitive* menunjukkan bahwa cukup sedikit aktor yang membangun interaksi dengan melalui perantara. Dan dari jaringan ini terlihat untuk jenis *transitive* terdapat 6 triad yang memiliki hubungan kuat dalam interaksinya.

1. Analisis pada Jaringan 2 Mode

Dalam pembentukan 2 mode ini, kami menggunakan 2 jenis node yaitu node akun user dan node tweet. Dari data kedua jenis tersebut, kami menggunakan *undirected graph* dalam melakukan visualisasinya. Dalam representasi ini, kami mendefinisikan edge dari jaringan ini dalam dua kasus. Kasus pertama yaitu ketika seorang user membuat sebuah tweet yang berkaitan dengan kata kunci yang ada, maka akan dibentuk sebuah relasi antara node akun user dan tweet yang dibuat. Kasus kedua adalah bila suatu user menyukai sebuah tweet ataupun melakukan retweet terhadap suatu tweet, maka juga akan dibentuk relasi edge antar keduanya. Kemudian kami melakukan pemotongan untuk edge yang memiliki bobot kurang dari 2 dan menyebabkan terbentuknya banyak subgraf. Untuk mengatasi ini, kami melakukan implementasi dengan menggunakan metode metode island dengan batasan minimal 10 node yaitu metode yang melakukan pembuangan terhadap subgraf yang memiliki anggota node yang sedikit. Untuk visualisasinya penggunaan 2 mode ini pada batch-1 dapat dilihat pada gambar 10 dibawah. Pada gambar 10 ini, node berwarna merah merepresentasikan node akun user dan node dengan warna hijau merepresentasikan tweet yang ada.



Gambar 10. Jaringan sosial dengan menggunakan jaringan 2 mode pada batch-1

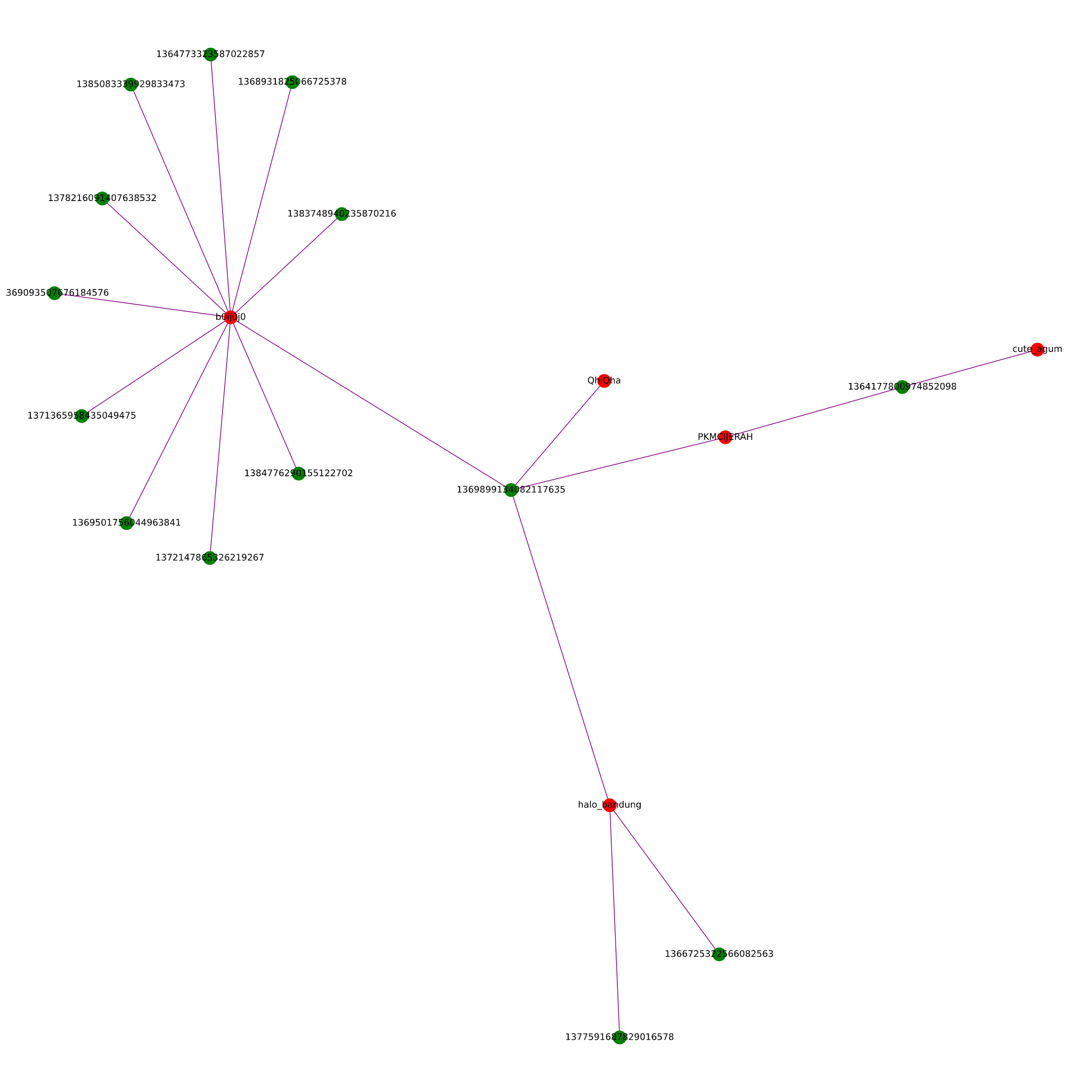
Dari graf yang terbentuk pada gambar 10, terlihat sebuah akun yang paling tampak yaitu akun @bekeraza. Akun ini memiliki degree centrality sebesar 62 yang merupakan degree centrality terbesar bila dibandingkan dengan node lainnya dan dengan nilai eigenvector centrality sebesar 0.682816, closeness centrality sebesar 0.549568 dan betweenness centrality-nya sebesar 0.235757.



Gambar 11. Salah satu aktivitas yang dilakukan oleh akun @bekeraza

Kemudian kami melakukan penelusuran pada akun @bekeraza ini. Akun ini merupakan akun asal kota Makassar. Salah satu aktivitas yang terlihat dari akun ini ditunjukkan pada gambar 11. Pada gambar tersebut terlihat ternyata akun ini merupakan akun yang sering sekali melakukan retweet terhadap tweet yang membahas masalah seputar kesehatan. Jadi, tingginya degree centrality yang ditunjukkan pada node ini bukanlah dikarenakan akun ini banyak membuat tweet yang di retweet atau like orang lain, namun lebih pada aktif untuk melakukan interaksi pada tweet yang dibuat orang lain.

Selanjutnya, kami juga melakukan visualisasi pada jaringan 2 mode yang dengan model pembentukan yang serupa dengan pembentukan jaringan 2 mode pada batch-1 yaitu dengan menjadikan akun user dan tweet sebagai jenis node yang ada. Pada batch-2 ini juga dilakukan hal yang sama seperti melakukan pemotongan edge dan menerapkan metode island dengan batas yang sama pula. Hasilnya, terbentuklah graf yang bisa dilihat pada gambar 12.



Gambar 12. Jaringan sosial menggunakan 2 mode network pada batch-2

Dari graf 2 jenis yang terbentuk pada gambar 12, kami melakukan hal serupa dengan sebelumnya yaitu melakukan penelusuran lebih pada akun yang terlihat mencolok yaitu akun @b0ij0j0. Pada akun ini dalam graf tersebut memiliki nilai degree centrality sebesar 11, eigenvector centrality sebesar 0.687907 closeness centrality sebesar 0.816993 dan betweenness centrality sebesar 0.600000. Akun ini merupakan akun yang berlokasi di kota Bandung.



Gambar 13. Salah satu aktivitas yang dilakukan oleh akun @b0ij0j0

Pada gambar 13 tersebut, menunjukkan salah satu aktivitas yang dilakukan oleh akun @b0ij0j0 pada twitter. Dari salah satu aktivitas ini, terlihat bahwa akun ini juga bukan merupakan akun yang rajin untuk membuat tweet sendiri. Namun akun ini memang merupakan akun yang sering untuk berinteraksi dengan tweet yang dibuat oleh akun lainnya. Namun bila dibandingkan dengan akun @bekeraza yang terdapat pada batch-1, akun ini memang bukanlah akun yang hanya berinteraksi dengan tweet yang membahas seputar kesehatan. Hal ini bisa terlihat pada gambar 13 tersebut bahwa akun ini berinteraksi dengan akun radio dengan topik yang general.

1. Analisis perbedaan pada batch-1 dan batch-2

Pada hasil analisis dari pengambilan data pada batch-1, diperoleh akun yang sangat berpengaruh pada jaringan sosial tersebut adalah akun @dr\_koko28. Setelah dilakukan sampling tambahan dengan akun pusat @dr\_koko28 didapatkan juga sebuah aktor yang cukup berpengaruh juga yaitu @radioelshinta. Kemudian pada analisis dengan pengambilan data pada batch-2 ditemukan akun berpengaruh yang berbeda dibandingkan dengan analisis pada batch-1 yaitu akun @dinkes\_bandung dan akun @infobdg. Bila diamati dari asal tiap-tiap akun yang berpengaruh tersebut, terdapat perbedaan yang cukup jelas yaitu pada akun batch-1, akun @dr\_koko28 dan akun @radioelshinta merupakan akun yang berkaitan dengan informasi yang ada di dalam skala nasional atau berpusat di Jakarta. Sedangkan akun @dinkes\_bandung dan akun @infobdg masing-masing merupakan akun yang beroperasional khusus untuk wilayah Bandung dan sekitarnya. Bila dari asal operasional tersebut dikaitkan dengan waktu pengambilan data dimana data pada batch-1 merupakan data sebelum vaksinasi dilakukan secara massal dan data pada batch-2 merupakan data yang diambil setelah adanya vaksinasi massal, maka dapat terlihat bahwa sebelum adanya vaksinasi, skala penyebaran informasi mengenai vaksinasi sendiri kebanyakan disebarkan dalam skala nasional, sedangkan ketika vaksinasi mulai dilakukan, akun-akun informasi yang berskala lokal pada wilayah Bandung mulai terlihat lebih aktif dalam memberikan informasi seputar vaksinasi.

Pada analisis triad didapatkan bahwa pada batch-1 dan batch-2 memiliki beberapa kesamaan pada hubungan interaksinya. Kesamaan ini terlihat dari jumlah perbandingan dari masing-masing kategori triad yang didapatkan. Pada kategori *vacuously transitive*  pada 021U menunjukkan bahwa banyak aktor egosentris yang sering menerima perhatian banyak aktor tetapi tidak pernah membalas tetapi terdapat perbedaan dimana pada batch-2 jumlahnya relatif jauh lebih rendah dibandingkan pada batch-1. Untuk 021D, keduanya menunjukkan hasil yang sama yaitu hanya sedikit aktor yang berinteraksi dengan aktor lain tetapi tidak mendapatkan balasan. Pada *intransitive* untuk batch-1 dan batch-2 memiliki jumlah yang sedikit namun untuk perbandingannya, pada batch-2 memiliki nilai yang lebih besar yaitu sebesar 122 triad dibandingkan pada batch-1 yang hanya berjumlah 3 triad. Untuk kategori *transitive*, terlihat perbedaan dimana terdapat 6 triad yang memiliki hubungan kuat pada batch-2 dan tidak ada triad yang memiliki hubungan kuat pada batch-1.

Kesimpulan

Tujuan utama dari penelitian ini adalah untuk melakukan ekstraksi knowledge yang bisa didapatkan dari analisis jejaring sosial dengan kata kunci berupa “vaksin”, “vaksinasi”, “astrazeneca” ataupun “sinovac” dengan akun awal yaitu akun @prfmnews yang merupakan akun twitter salah satu radio di Bandung. Dengan melakukan berbagai macam analisis pada sentralitas dan ranking popularitas maka dapat terlihat komunitas-komunitas yang terbentuk. Dari hasil data pada pengambilan batch-1 dapat disimpulkan bahwa 2 komunitas paling besar merupakan komunitas dengan pusat akun @dr\_koko28 dan @radioelshinta yang notabene adalah akun beroperasi di Jakarta atau akun skala nasional justru lebih berpengaruh dibandingkan akun lokal Bandung sendiri. Selain itu dari hasil pengambilan batch-2, tweet yang berasal komunitas dengan akun-akun pusat berasal dari lokal Bandung seperti akun @dinkes\_bandung dan akun @infobdg justru baru mulai mendapatkan perhatian lebih bagi netizen khususnya warga Bandung sendiri. Pada perhitungan sensus triad, terlihat bahwa untuk tingkat interaksi yang terjadi pada jejaring sosial pada kedua batch ini memanglah lemah namun pada batch-2 terjadi beberapa peningkatan dalam tingkat kekuatan hubungan yang terjadi. Untuk hasil penelusuran pada jaringan 2 mode, kami menemukan bahwa akun akun yang memiliki tingkat degree yang tinggi ternyata bukanlah akun yang sering melakukan tweet yang tweetnya sering di like ataupun di retweet, namun lebih kepada akun yang sering melakukan retweet dan like pada sering melakukan interaksi dengan berbagai tweet yang dipublikasikan oleh media informasi.