Analisis Jaringan Pengguna *Platform* Musik dalam Membentuk Tren Musik

¹Rahman Taufik, ²Muhaqiqin, ³Ridho Sholehurrohman, dan ⁴Igit Sabda Ilman

^{1,2,3,4}Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung Jalan Prof. Dr. Ir. Soemantri Brojonegoro No.1 Gedong Meneng, Bandar Lampung, Provinsi Lampung, Indonesia e-mail: ¹rahman.taufik@fmipa.unila.ac.id, ²muhaqiqin@fmipa.unila.ac.id, ³ridho.sholehurrohman@fmipa.unila.ac.id, ⁴igit.sabda@fmipa.unila.ac.id,

Abstract — The social media with music content is currently growing very quickly, many interactions that occur are based on the same interest in content. The relationship occurs between users of the music platform based on the influence of the same interest in music that can be seen in the form of a network. This study uses the idea-based percolation method (PIB) to identify social groups based on the listener's musical genre, the Louvain method to identify group modularity, and the clique-percolation method to identify overlapping communities. The generated network includes a bipartite network that is formed when listeners are connected to a music group and a unipartite network that is formed when listeners connect with listeners who have the same musical taste. The results show that there is a network of groups based on the relationship of preferred music tastes, there are 13 communities that are formed based on the value of modularity and there are overlapping communities with 10 nodes forming a k-clique. Furthermore, the results represent that in a music network, the influence of music taste is always initiated by one user, then spreads to other users.

Keywords: Community Detection; Music Platform; Network; Percolation Idea-Based.

1. PENDAHULUAN

Perkembangan media sosial saat ini berkembang dengan sangat cepat, banyak interaksi yang terjadi didasari oleh minat yang sama dalam hal konten. Diantara media sosial yang populer dan terus berkembang salah satunya adalah media sosial dengan konten musik. Platform musik seperti Last FM, Soundcloud, JOOX, Spotify dan lain-lain, berlomba-lomba memberikan layanan yang terbaik dan mendorong interaksi sosial dengan memberikan kebebasan terhadap pengguna untuk berbagi lagu yang mereka sukai kepada publik. Dengan kebiasaan berbagi lagu, baik satu lagu maupun kumpulan lagu yang biasa disebut *playlist*, bukan tidak mungkin bisa mempengaruhi minat orang lain terhadap musik, dan secara tidak langsung akan menciptakan tren musik dalam kurun waktu tertentu.

Hubungan yang terjadi antara pengguna platform musik berdasarkan pengaruh minat musik yang sama, dapat dilihat dalam bentuk jaringan. Suatu jaringan terdiri dari beberapa *node* dan koneksi yang disebut link atau *edge*, koneksi ini menghubungkan satu *node* dengan *node* lainnya. Dengan visualisasi jaringan, gambaran interaksi antara *node* dengan *node* akan terlihat, selain itu penyebaran informasi, kecepatan informasi dan informasi apa saja yang berpengaruh akan terlihat dan dapat dianalisis. Pemanfaatan analisis jaringan ini digunakan untuk mengetahui informasi apa saja yang dibutuhkan sehingga terciptanya tren musik yang berkembang di masyarakat.

Untuk mendapatkan data-data pengguna dengan minat musik yang akan divisualisasikan dalam bentuk jaringan, dalam studi ini digunakan situs Last.fm. Dibandingkan dengan platform musik lainnya, Last.fm ini adalah platform musik yang sudah lama terbentuk (dari tahun 2002), sehingga data dari tahun ke tahun akan mudah ditemukan dan data yang didapatkan dalam jumlah yang besar. Kelebihan lainnya dari fitur last.fm adalah detail selera musik dari setiap profil pengguna dan fitur like yang diberikan oleh pengguna lain jika mereka menyukai playlist musik pengguna tersebut. Berdasarkan ketertarikan suatu playlist yang paling banyak disukai oleh pengguna lain, hal ini menjadikannya sebagai rekomendasi tren musik dari tahun ke tahun.

Tujuan penelitian ini adalah untuk menganalisis hubungan yang terjadi antara pengguna platform musik berdasarkan ketertarikan mereka terhadap playlist seseorang sehingga membentuk tren musik yang berkembang di masyarakat. Percobaan yang dilakukan dalam studi ini meliputi analisis struktur jaringan, perhitungan *centrality*, deteksi komunitas dan penyebaran informasi.

2. METODOLOGI PENELITIAN

Analisis jaringan untuk situs web musik dilakukan oleh Lambiotte dan Ausloos [2] yang didedikasikan untuk pendengar yang memiliki kebiasaan berbagi musik. Penelitian ini menggunakan data dari www.audioscrobbler.com dan www.musicmobs.com, serta metode percolation idea-based (PIB) untuk mengidentifikasi kelompok sosial berdasarkan genre musik pendengar. Jaringan yang terbentuk adalah jaringan bipartite dan unipartite. Jaringan bipartite terbentuk ketika pendengar terhubung dengan grup musik, sedangkan jaringan unipartite terbentuk ketika pendengar terhubung dengan pendengar yang mempunyai selera musik yang sama. Dengan menggunakan teknik PIB untuk jaringan pendengar, selera musik individu dan jaringan grup musik, hasil menunjukan grafik eksponensial yang berarti bahwa jaringan selera musik terbentuk dari pilihan individu dan tren kolektif.

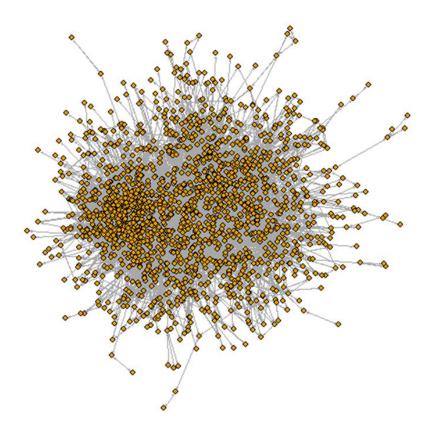
Penelitian jaringan situs web musik berdasarkan pendengar lebih lanjut dikembangkan dengan meningkatkan metodologi pemetaan kelompok-kelompok musik ke dalam genre menggunakan catatan pendengar [2]. Penelitian ini mencoba untuk menemukan jawaban atas masalah keberadaan tren musik yang memungkinkan untuk diklasifikasikan dan pemetaan jaringan genre dan sub-genre musik. Data tambahan diambil dari www.last.fm untuk mendepatkan deskripsi properti musik seperti genre / sub-genre dalam bentuk tag dan label musik. Dari database diperoleh 2394 genre musik yang berbeda, meskipun begitu hanya 142 genre musik yang diambil dikarenakan penyederhanaan dan genre musik yang tidak relevan. Menggunakan metodelogi yang sama dengan penelitian sebelumnya [3] Jaringan terbentuk oleh pemilihan tag genre oleh sesama pendengar yang menghubungkan antar genre dan memberikan definisi yang obyektif terhadap gagasan sub-genre.

2.1. Dataset

Dataset yang digunakan dalam studi ini diperoleh dari International Workshop on Information Heterogeneity and Fusion in Recommender Systems (Hetrec) tahun 2011, data ini dapat diunduh di http://ir.ii.uam.es/hetrec2011. Data tersebut merepresentasikan informasi-informasi tentang pengguna, hubungan antara pengguna, serta tagging data musik seperti judul lagu dan nama artist. Informasi yang digunakan dari dataset ini berupa id pengguna, id artist, nama artist, id genre, nama genre, timestamp, tanggal, bulan, tahun, serta relasi antar pengguna. Informasi-informasi tersebut diambil dan direpresentasikan sebagai node dan edge. Profil pengguna dengan playlist musik yang dimilikinya direpresentasikan sebagai node, dengan jumlah sebanyak 186479 node. Ketertarikan atau hubungan pengguna terhadap playlist pengguna lainnya direpresentasikan sebagai edge, dengan jumlah sebanyak 25434 edge. Dataset yang diambil merepresentasikan waktu selama tahun 2005 sampai tahun 2011. Berdasarkan hubungan satu arah antara satu pengguna yang mengikuti pengguna lainnya, jaringan ini direpresentasikan dalam bentuk graf berarah (directed) dan tidak berbobot (unweighted).

2.2. Visualisasi Jaringan

Untuk membentuk graf berarah berdasarkan *node* dan *edge* dari dataset yang diambil, digunakan aplikasi R untuk mendapatkan hasil visualisasi graf dan informasi-informasi yang nantinya dianalisis. Berikut hasil visualisasi jaringan pada gambar dibawah ini.



Gambar 1. Visualisasi graf

2.2. Struktur Jaringan

Struktur jaringan merepresentasikan bentuk jaringan berdasarkan distribusi data yang terjadi. Untuk mendapatkan struktur jaringan, dilakukan pengukuran jaringan meliputi *average path length*, *clustering coefficient*, dan *degree distribution* [4]. Berikut adalah hasil statistik pengukuran struktur jaringan.

a) Average path length

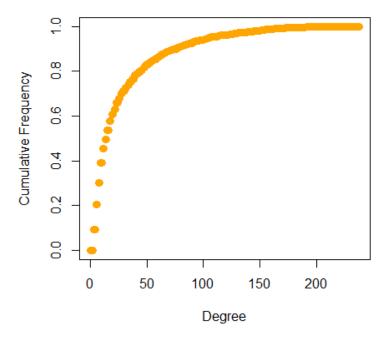
Dalam jaringan, terdapat jarak antara dua *node* yang didefinisikan sebagai distance (d), jarak maksimal diantara semua pasangan *node* yang didefinisikan sebagai diameter (D), dan jarak rata-rata atas semua pasangan *node* yang didefinisikan sebagai *average path length* (L). Dalam kasus tertentu *average path length* menentukan "ukuran" efektif pada suatu jaringan, sebagai contoh jaringan *small-world* yang memiliki ukuran *average path length* yang kecil. Untuk studi kasus ini, diameter jaringan yang didapatkan adalah 9, dengan dengan *average path length* sebesar 3,518628

b) Clustering coefficient

Clustering coefficient (C) didefinisikan sebagai fraksi rata-rata pasangan tetangga nodes yang juga merupakan tetangga satu sama lain. Contoh, node i yang terhubung dengan node lain disebut tetangga, dan setiap tetangga dari node i dapat saling terhubung, maka tetangga yang saling terhubung tersebut membentuk kelompok atau clustering. Meskipun begitu sebagian besar jaringan kompleks nyata tidak sepenuhnya acak, karena kebanyakan jaringan di dunia nyata yang berskala besar memiliki kecenderungan ke arah pengelompokan dan secara global mereka terhubung jauh dengan kelompok lainnya. Adapun hasil clustering coefficient untuk studi kasus ini adalah 0,2015625.

c) Degree distribution

Karakteristik penting dari suatu *node* adalah *degree*, *degree* (ki) pada *node* biasanya didefinisikan jumlah total yang terhubung. Semakin besar *degree* "semakin penting" *node* tersebut di suatu jaringan. Rata-rata *degree* atas *node* penting disebut juga rata-rata / average *degree* pada suatu jaringan (k). Penyebaran *node degree* dicirikan dengan distribution function P(k) yang berarti probabilitas suatu tersebut *node* akan dipilih. Jika semua *node* memiliki *degree* sama maka *degree distribution* tersebut adalah single sharp spike (delta distribution). Sedangkan untuk jaringan nyata berskala besar distribusi *degree*-nya dapat membentuk distribusi poisson yaitu distribusinya yang jatuh secara eksponensial atau membentuk distribusi power law yaitu distribusinya yang terus meningkat. Adapun hasil *degree distribution* untuk studi kasus ini sebagai berikut



Gambar 2. Distribusi degree

Berdasarkan hasil *average path length* yang memiliki nilai 3,518628 dan *clustering coefficient* yang memiliki nilai 0,2015625, jaringan ini dapat dinyatakan kedalam model jaringan small-world. Sedangkan dilihat dari *degree* distribusi yang membentuk distribusi power law, jaringan ini termasuk kedalam scale free model. Untuk menjelaskan distribusi derajat power law, Barabasi-Albert [4] berpendapat bahwa model jaringan terbentuk secara dinamis dengan penambahan *node* baru ke dalam jaringan secara kontinyu, hal ini menyiratkan terdapat suatu *node* dengan jumlah *degree* terbanyak. Seperti dalam studi kasus ini pengguna dengan jumlah rujukan terbanyak, secara alami akan bertambah diikuti oleh pengguna baru lainnya, fenomena ini biasa disebut "rich get richer".

3. HASIL DAN PEMBAHASAN

3.1. Pengukuran Centrality

Untuk mengetahui keberadaan *node* "penting" yang memainkan peran kunci dalam membawa *nodes-nodes* lain dalam membentuk suatu jaringan, digunakan pengukuran *centrality* yang terdiri dari *degree centrality*, *eigencentrality*, *closeness centrality* dan *betweenness centrality* [5]. Berikut adalah hasil pengukuran *centrality*:

a) Degree centrality

Degree centrality merujuk pada node dengan degree terbesar yang ada didalam jaringan. Degree ini berdasarkan jumlah tetangga yang terdekat dengan node tersebut. Degree centrality mengukur pengaruh suatu node terhadap node lainnya dalam suatu jaringan. Ukuran degree centrality dari keseluruhan jaringan adalah 0,0571403.

```
> head(sort_degree)
1543 1281 831 179 1503 1023
119 110 106 97 95 91
```

Dalam studi kasus ini didapatkan urutan *node* dengan *degree centrality* teratas, dimana *node* yang paling tinggi ditujukan oleh *node* dengan id 1543. *Node* dengan id 1543 merupakan *node* yang berperan penting dan memiliki pengaruh dalam jaringan musik ini.

b) Betweenness centrality

Betweenness centrality merujuk pada node yang bertindak sebagai jembatan sepanjang jalur terpendek diantara dua node lainnya, node yang memiliki probabilitas tinggi pada jalur terpendek dipilih secara acak antara dua simpul yang memiliki hubungan yang tinggi. Jika suatu node memiliki betweenness centrality yang tinggi maka hal ini menunjukan bahwa node tersebut sering dilewati atau terletak pada shortest path antara node yang lainnya.

Ukuran betwenness *centrality* dari keseluruhan jaringan ini adalah 0,06878413, dimana diperoleh *node* dengan betwenness *centrality* terbesar adalah *node* dengan id 1543.

c) Eigencentrality

Eigencentrality merujuk pada ukuran pengaruh suatu node tergantung pada jumlah tetangga dan kualitas koneksinya didalam jaringan. Terdapat persamaan antara eigencentrality dengan degree centrality yaitu mengukur pengaruh suatu node berdasarkan jumlah edges yang dimiliki oleh suatu node, tetapi perbedaannya kualitas node dipengaruhi oleh nilai eigen yang dimiliki oleh tetangga lain dari node tersebut.

```
> head(sort_eigen)
     1300     1023     179     1247     129     1356
1.0000000 0.9351501 0.9277137 0.8683462 0.8320409 0.8270248
```

Ukuran *eigencentrality* jaringan ini adalah 0,9150933, dimana *node* dengan nilai eigen terbesar adalah *node* dengan id 1300. Berbeda dengan *node* id 1543 (*node degree* terbesar), *node* id 1300 dipengaruhi oleh *node* lainnya yang memiliki nilai eigen tinggi.

d) Closeness centrality

Closeness centrality merujuk pada seberapa dekat atau jauh node sentral ke semua node lainnya, semakin sentral sebuah node, semakin rendah jarak total ke semua node lainnya.

Ukuran *closeness centrality* dari keseluruhan jaringan ini adalah 0,245886, dimana diperoleh *node* dengan *closeness centrality* terbesar adalah *node* dengan id 831, sedangkan untuk *node* dengan id 1543 berada diurutan kedua, hal ini menunjukan bahwa *node* dengan id 1543 masih memiliki peran penting terhadap *node* lainnya yang ada didalam jaringan musik ini.

3.1.1 Analisis Centrality

Berdasarkan perhitungan degree centrality, eigencentrality, closeness centrality dan betweenness centrality, didapatkan node penting yang memiliki pengaruh penting terhadap node lainnya dalam suatu jaringan. Node yang paling penting dalam jaringan studi kasus ini ditunjukan oleh node dengan id 1543, node ini menjadi penting dan berpengaruh karena memiliki nilai degree centrality, closeness centrality dan betweenness centrality yang paling besar, bahkan untuk degree centrality dan closeness centrality node ini merupakan node dengan urutan pertama. Meskipun node dengan id 1300 merupakan node dengan nilai eigen terbesar, tetapi dilihat dari jumlah degree yaitu 89, maka node dengan id 1300 bukan termasuk node yang paling penting.

Mengikuti studi kasus musik yang merepresentasikan *node* sebagai pengguna, pengguna dengan id 1543 merupakan pengguna yang paling banyak dirujuk oleh pengguna lain dalam hal selera musik. Informasi playlist musik yang dimiliki oleh pengguna id 1543 didominasi oleh artist atau musisi dengan genre Alternative Pop dan Progressive Rock, dari hal tersebut menunjukan bahwa tren musik yang terjadi selama tahun 2005 sampai 2011 adalah musik dengan genre Alternative Pop dan Progressive Rock.

3.2. Deteksi Komunitas

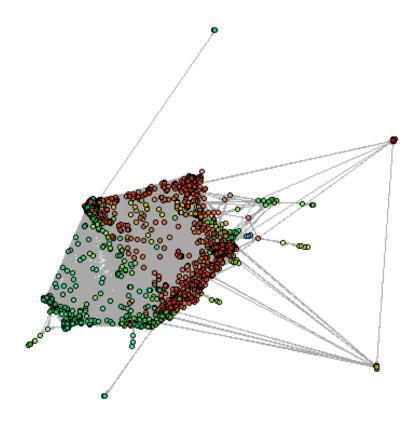
Konsep dasar deteksi komunitas adalah setiap kumpulan data yang 'serupa' dapat membentuk komunitas. Dalam studi ini, digunakan metode deteksi komunitas berdasarkan Louvain dan Clique Percolation untuk menentukan adanya overlapping communities.

a) Metode Louvain

Metode ini berdasarkan pada gagasan pencarian nilai modularitas yang maksimal untuk mengukur kualitas partisi tertentu dari suatu graf. Pencarian nilai modularitas berdasarkan kepadatan *edge* yang diharapkan dapat membentuk subgraf dalam hal ini partisi pada deteksi komunitas. Fokus metode ini digunakan untuk mendeteksi komunitas graf berarah, tidak berarah atau berbobot.

```
IGRAPH clustering multi level, groups: 13, mod: 0.45
 groups:
  $'1'
   [1] "111" "206"
                                     "435"
                                            "603"
                                                            "855"
  [11] "1862"
  $`2`
    [1] "5"
                "6"
                               "10"
                                      111411
                                              0150
   [11] "34"
                "37"
                       "41"
                               "50"
                                      "51"
                                                             "61"
                "71"
                       "74"
                               "77"
   [31] "97"
                "101"
                       "107"
                                              "113"
                               "108"
                                      "109"
                                                     "114"
                                                             "115"
                                                                            "121"
               "123" "125"
                               "127" "139"
   ... omitted several groups/vertices
```

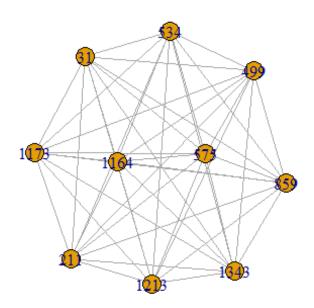
Nilai modularitas yang didapatkan dari graf ini adalah 0,45 dengan jumlah komunitas yang didapatkan adalah 13 komunitas. Komunitas dapat direpresentasikan dalam bentuk graf dengan pewarnaan untuk masing-masing komunitas.



Gambar 3. Visualisasi graf berdasarkan deteksi komunitas

b) Metode Clique Percolation

Metode ini berdasarkan pada gagasan k-clique sebagai subgraf komplit yang terdiri dari k *nodes* (jumlah *node*) dengan *edge* yang saling terhubung antar *node*. Jumlah k-clique ditentukan berdasarkan nilai clique maksimal yang didapatkan dari graf tersebut. Komunitas k-clique didefinisikan sebagai persatuan k-clique yang dapat terbentuk melalui pengulangan proses k-clique yang berdekatan. Fokus metode ini digunakan untuk menganalisis struktur komunitas yang overlapping baik untuk graf berarah, tidak berarah atau berbobot.



Gambar 4. Visualisasi graf berdasarkan overlapping community

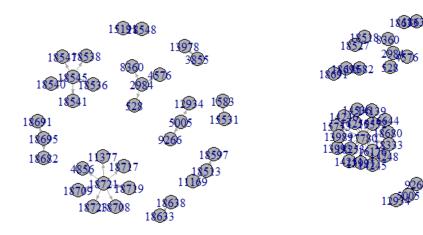
node-node yang melakukan overlapping dengan node lain antara lain adalah node 31, 534, 499, 859, 575, 1343, 1213, 1164, 211, dan 1173. Jumlah overlapping yang didapatkan adalah 10 nodes, masing-masing node tersebut memiliki relasi ke semua nodes lainnya sehingga membentuk subgraf komplit.

3.3. Penyebaran Informasi

Penyebaran informasi dalam graf yang terbentuk merepresentasikan bagaimana proses penyebaran informasi tersebut dilakukan dan informasi apa saja yang disebarkan. Dalam studi kasus ini model cascade dipilih sebagai information propagation dikarenakan jaringan music online terbentuk berdasarkan pengguna yang dapat mengontrol ketertarikan (hubungan) mereka terhadap beberapa playlist music pengguna lainnya.

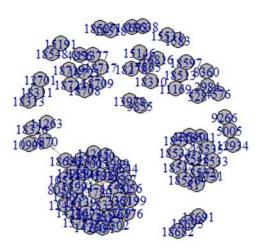
```
> is.cascade(cascades)
[1] TRUE
> summary(cascades)
# cascades: 1170
# nodes: 12523
# nodes in cascades: 12523
# possible edges: 178605
Summary statistics for cascade length and number of ties:
         length ties
1.00000 0.000000
1st Qu. 1.00000 0.000000
Median
        3.00000 1.000000
Mean
       10.70342 9.061538
3rd Qu. 8.00000
                   6.000000
       472.00000 470.000000
```

Berdasarkan keterangan diatas, sebanyak 12523 *nodes* dalam jaringan dapat memperoleh informasi dan sebanyak 178605 *edges* memungkinkan terjadinya hubungan penyebaran informasi berdasarkan cascade.



Gambar 5. Visualisasi cascade A.

Gambar 6. Visualisasi cascade B



Gambar 7. Visualisasi cascade C

Graf diatas (gambar 5, 6 dan 7) merupakan contoh penyebaran informasi cascade yang terjadi berdasarkan jumlah *nodes*. Untuk gambar (5) merepresentasikan cascade diantara 25 *nodes*, gambar (6) merepresentasikan cascade diantara 50 *nodes*, dan gambar (7) merepresentasikan cascade diantara 75 *nodes*. Ketiga graf diatas menunjukan bahwa penyebaran informasi untuk jaringan ini selalu dimulai dari 1 *node*, walaupun jumlah *node* yang mendapatkan informasinya berbeda-beda untuk setiap *node*.

Representasi dari penyebaran informasi cascade untuk studi kasus ini menunjukan bahwa penyebaran playlist musik biasanya dimulai dari 1 orang, lalu menyebar ke orang lain dalam jumlah yang tidak tentu. Selain itu terdapat beberapa pengguna yang banyak dirujuk oleh banyak pengguna dalam hal selera musik (contoh ini dapat dilihat pada gambar 7 dimana terdapat 1 *node* yang banyak dirujuk oleh *node* lain).

4. KESIMPULAN

Dalam penelitian ini, jaringan yang terbentuk didasari oleh hubungan antara pengguna platform musik berdasarkan ketertarikan mereka terhadap playlist musik seseorang. Struktur jaringan ketertarikan musik ini membentuk scale-free network dimana terdapat fenomena rich-get-richer yaitu hubungan pengguna yang semakin lama semakin banyak dirujuk. Dalam kasus ini, pengguna yang banyak dirujuk oleh pengguna lain dalam hal playlist musik adalah pengguna dengan id 1543 dilihat dari perhitungan *centrality* jaringan. Playlist musik yang dimiliki oleh pengguna dengan id 1543 ini didominasi oleh musik dengan genre Alternative Pop dan Progressive Rock. Berdasarkan banyaknya jumlah pengguna lain yang mengikuti selera musik pengguna dengan id 1543, genre musik Alternative Pop dan Progressive Rock merupakan tren musik yang berlansung saat itu (kurun waktu tahun 2005 sampai 2011).

Meskipun terdapat tren musik yang banyak dijadikan rujukan oleh banyak pengguna. Sebagian besar pengguna memiliki ketertarikan terhadap banyak artis dengan selera musik yang beragam, hal tersebut menunjukan adanya kelompok-kelompok yang terbentuk berdasarkan hubungan selera musik yang disukai. Untuk mendapatkan jumlah komunitas dan anggotanya, digunakan metode deteksi komunitas untuk jaringan ini yaitu: (i) terdapat 13 komunitas yang terbentuk berdasarkan nilai modularitas menggunakan metode Louvain dan (ii) terdapat overlapping communities dengan 10 *node* membentuk k-clique berdasarkan metode clique-percolation.

Untuk melihat penyebaran informasi dalam hal ketertarikan playlist musik didalam jaringan ini, model cascade digunakan berdasarkan pengguna yang dapat mengontrol ketertarikan (hubungan) mereka. Dapat dilihat ratarata penyebaran informasi selalu dimulai oleh satu *node*, dengan jumlah *node* yang dipengaruhinya berbedabeda. Hal ini merepresentasikan bahwa dalam jaringan musik, pengaruh selera musik selalu diawali oleh satu pengguna, kemudian menyebar ke pengguna lainnya.

DAFTAR PUSTAKA

- [1] **S.** Fortunato, "Community detection in graphs," *Physics Reports*, 486(3–5), 75–174, 2010 https://doi.org/10.1016/j.physrep.2009.11.002
- [2] R. Lambiotte, & M. Ausloos, "Uncovering collective listening habits and music genres in bipartite networks," *Physical Review E*, 72(6), 2005. https://doi.org/10.1103/PhysRevE.72.066107.
- [3] R. Lambiotte, & M. Ausloos, "On the genre-fication of music: a percolation approach," *The European Physical Journal B*, 50(1–2), 183–188, 2006, https://doi.org/10.1140/epjb/e2006-00115-0.
- [4] X. F. Wang, & G. Chen, "Complex networks: Small-world, scale-free and beyond," *IEEE Circuits and Systems Magazine*, 3(1), 6–20. 2003, https://doi.org/10.1109/MCAS.2003.1228503.
- [5] J. Zheng & Y. Luo, "Degree Centrality, Betweenness Centrality, and Closeness Centrality in Social Network," Conference: 2017 2nd International Conference on Modelling, Simulation and Applied Mathematics (MSAM2017), 2017