

SKRIPSI

**DETEKSI KOMUNITAS AKTIF MENGGUNAKAN
ALGORITMA CLUSTERING FACETNET**



Vinieta Abhinandaniya Anggayashti

NPM: 2014730066

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
2018**

UNDERGRADUATE THESIS

**ACTIVE COMMUNITY DETECTION USING FACETNET
CLUSTERING ALGORITHM**



Vinieta Abhinandaniya Anggayashti

NPM: 2014730066

**DEPARTMENT OF INFORMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
2018**

LEMBAR PENGESAHAN



DETEKSI KOMUNITAS AKTIF MENGGUNAKAN ALGORITMA CLUSTERING FACETNET

Vinieta Abhinandaniya Anggayashti

NPM: 2014730066

Bandung, 31 Mei 2018

Menyetujui,

Pembimbing

A handwritten signature in black ink, appearing to be 'Veronica Sri Moertini', written over a diagonal line.

Dr. Veronica Sri Moertini

Ketua Tim Penguji

A handwritten signature in black ink, appearing to be 'Thomas Anung Basuki', written over a diagonal line.

Dott. Thomas Anung Basuki

Anggota Tim Penguji

A handwritten signature in blue ink, appearing to be 'Elisati Hulu', written over a vertical line.

Elisati Hulu, M.T.

Mengetahui,

Ketua Program Studi

A handwritten signature in black ink, appearing to be 'Mariskha Tri Adithia', written over a diagonal line.

Mariskha Tri Adithia, P.D.Eng



PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

DETEKSI KOMUNITAS AKTIF MENGGUNAKAN ALGORITMA CLUSTERING FACETNET

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
Tanggal 31 Mei 2018



Vinieta Abhinandaniya Anggayashti
NPM: 2014730066

ABSTRAK

Teknologi kini sudah semakin berkembang, salah satunya di bidang komunikasi. Saat ini, hampir semua orang menggunakan telepon atau media sosial dalam berkomunikasi. Hal ini menghasilkan data telepon dan media sosial yang melimpah. Data telepon atau media sosial dihasilkan dari interaksi antar penggunanya. Data ini dapat dianalisis untuk tujuan tertentu, salah satunya untuk menemukan komunitas aktif. Komunitas aktif adalah kumpulan individu yang aktif dalam berinteraksi.

Data telepon dan data media sosial memiliki ukuran yang sangat besar dan memiliki suatu pola. Data telepon dan data media sosial dapat digambarkan dengan graf. Graf adalah sejenis diagram yang dibentuk dari beberapa simpul dan sisi. Simpul pada graf dapat menggambarkan satu individu. Sisi pada graf dapat menggambarkan interaksi antar satu individu atau lebih. Banyaknya interaksi antar individu juga dapat digambarkan menggunakan graf. Setiap sisi memiliki bobot yang menggambarkan banyaknya interaksi antar individu.

Dalam menganalisis suatu komunitas dari data graf, dapat digunakan salah satu algoritma yang berbasis teknik *clustering*, yaitu algoritma FacetNet. Algoritma FacetNet dapat menganalisis data graf dan menemukan komunitas aktif dari data graf tersebut.

Pada skripsi ini, telah dibangun perangkat lunak untuk menganalisis data graf dan mendeteksi komunitas aktif dari data graf. Eksperimen yang dilakukan pada skripsi ini menggunakan beberapa data graf yang didapatkan dari website <http://konect.uni-koblenz.de/>. Eksperimen dilakukan untuk menguji kebenaran hasil analisis algoritma FacetNet dan menguji kinerja perangkat lunak yang telah dibangun. Hasil eksperimen yang telah dilakukan adalah algoritma FacetNet dapat mendeteksi komunitas aktif dengan benar dan lamanya eksekusi perangkat lunak dipengaruhi oleh banyaknya simpul pada data graf.

Kata-kata kunci: Graf, Komunitas, Clustering, Algoritma FacetNet

ABSTRACT

As years goes by, technologies grow on a lot of sector, one of the sectors is from communication tech. At this point, almost every human being use their handphone and their social media to do a communication. This create a lot of data on the interaction that are made from the user (which right now, a ton of it). This data can be analysed for lot of purpose, one of the purposes is to find a community which active to interact between each other (between individual).

From what we know the data that came from in is gadget (such as handphone, etc) and from application that use for social media, have resulted a big data, that have their own unique pattern. Both of those data are represented as graph. Graph is a diagram consisting of node and edge. Nodes on a graph represent one individual each and on the other side edge on graph can show interaction from a lot of individual. Interaction that came from a lot of individual can be explained using graph too. Every edge have their own weight to shown how many interaction between each individual.

Every community is analyse by the data that came from graph, can use as on of the algorithm that base from clustering technique, called FacetNet algorithm. This algorithm can be use for analyse graph data and find the active community from the data.

On this thesis, a software is made for analyse graph data and use to detect active community from the data. An experiment use on this thesis use a serveral data from graph, the data came from website <http://konect.uni-koblenz.de/>. To find the accurate data from FaceNet and to test the performance of the software we already build, this experiment made for test the capability of the software. The result from the test is that the algorithm from FacetNet can detect the activity community from a lot of interaction with the accuration and the performance of the software is depending by how many node on the graph data.

Keywords: Graph, Communities, Clustering, FacetNet Algorithm

*Dipersembahkan untuk Ayah di surga, Mama, dan kedua adikku
tercinta*

KATA PENGANTAR

Puji syukur penulis sampaikan kepada Allah SWT karena atas berkat dan rahmatNya penulis bisa menyelesaikan penyusunan skripsi dengan judul "Deteksi Komunitas Aktif Menggunakan Algoritma Clustering FacetNet". Penulis menyadari bahwa di dalam skripsi ini masih terdapat banyak kekurangan. Penulis juga menyadari bahwa penyusunan skripsi ini tidak terlepas dari bantuan berbagai pihak. Pada kesempatan ini, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua yang sudah memberikan semangat, motivasi, dan dukungan dalam berbagai hal dari hari pertama sampai hari terakhir penulis berada di kampus UNPAR serta untuk kedua adik penulis (Kinta dan Agha) yang selalu memberikan semangat.
2. Ibu Veronica Sri Moertini selaku dosen pembimbing yang telah memberikan waktu, nasihat, bimbingan, kritik, masukan, dan tambahan wawasan selama proses pembuatan skripsi ini sehingga skripsi ini dapat selesai tepat waktu.
3. Ibu Imelda yang telah membantu mengajarkan dan memberikan wawasan tentang algoritma FacetNet dan telah meluangkan waktunya untuk penulis.
4. Muhammad Adrian Putra Zubir yang selalu menemani, memberikan semangat, memberikan candaan, dan membantu penulis dalam hal apapun serta selalu mendengar berbagai keluhan kesah dari penulis.
5. Nancy Valentina, Nadhila Herdanti Dwi Larasati, dan Farhan Fariz Fazrian yang selalu memberi semangat, menemani serta membantu penulis saat melakukan penyusunan skripsi ini serta selalu bisa memberikan masukan jika penulis kebingungan dalam hal apapun.
6. Reynaldo Immanuel dan Muhammad Irfan yang membantu dalam penyusunan skripsi terutama saat penulis kebingungan karena menemukan error pada program serta semua anggota Tim Belajar yang selalu menyemangati penulis.
7. Tim Misi Budaya Yunani Listra UNPAR dan Keluarga UKM Listra UNPAR yang selalu memberi semangat dan selalu mengatakan bahwa penulis bisa menyelesaikan skripsi initepat waktu.
8. Rekan-rekan mahasiswa dari Jurusan Teknik Informatika angkatan 2014 dan senior yang selalu mau memberikan jawaban atas pertanyaan penulis terkait skripsi ini.
9. Semua pihak yang tidak mungkin disebutkan satu-persatu yang sudah memberikan bantuan dan dukungan dalam pengerjaan skripsi ini.

Akhir kata, penulis memohon maaf jika terdapat kesalahan dan kekurangan dalam skripsi ini. Semoga skripsi ini dapat bermanfaat bagi pihak yang membutuhkan.

Bandung, Mei 2018

Penulis

DAFTAR ISI

KATA PENGANTAR	xv
DAFTAR ISI	xvii
DAFTAR GAMBAR	xix
DAFTAR TABEL	xxi
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Batasan Masalah	2
1.5 Metodologi Penelitian	3
1.6 Sistematika Pembahasan	3
2 LANDASAN TEORI	5
2.1 Graf	5
2.1.1 Definisi Graf	5
2.1.2 Jenis-Jenis Graf	7
2.1.3 Representasi Graf	9
2.2 <i>Clustering</i>	13
2.2.1 Definisi <i>Clustering</i>	13
2.2.2 Metode-Metode <i>Clustering</i>	13
2.3 Algoritma FacetNet	14
2.3.1 Definisi Algoritma FacetNet	14
2.3.2 Formulasi Algoritma FacetNet	14
2.3.3 Algoritma Iterasi	16
2.3.4 <i>Community Membership</i>	17
2.3.5 <i>Soft Modularity</i>	17
3 KOMPUTASI PADA ALGORITMA FACETNET	19
3.1 Contoh Komputasi pada Algoritma FacetNet	19
3.2 Penghitungan <i>Soft Modularity</i>	45
4 ANALISIS DAN PERANCANGAN PERANGKAT LUNAK	49
4.1 Tahapan Algoritma FacetNet	49
4.2 <i>Use Case</i>	50
4.3 Struktur Data Graf	52
4.4 Perancangan Diagram Kelas	53
4.5 Rancangan Algoritma untuk Metoda Kelas	58
4.6 Perancangan Antarmuka	66
5 IMPLEMENTASI DAN PENGUJIAN PERANGKAT LUNAK	69

5.1 Implementasi Perangkat Lunak	69
5.2 Implementasi Antarmuka	76
5.3 Pengujian Perangkat Lunak dengan Data Kecil	79
5.4 Eksperimen Perangkat Lunak	85
6 KESIMPULAN DAN SARAN	87
6.1 Kesimpulan	87
6.2 Saran	87
DAFTAR REFERENSI	89
A KODE PROGRAM	91

DAFTAR GAMBAR

2.1	Graf G_1 dan Graf G_2	5
2.2	Graf G_3 : Graf Berarah	8
2.3	Graf G_4 : Graf Berbobot	8
2.4	Graf G_5 : Graf Berbobot Tak-Berarah	10
2.5	Adjacency List dari Graf Tak Berarah G_1	12
2.6	Adjacency List dari Graf Berarah G_3	12
2.7	Adjacency List dari Graf G_4	12
3.1	Graf	19
4.1	<i>Use Case</i>	51
4.2	Contoh Fail Masukan	52
4.3	Diagram Kelas Sederhana	53
4.4	Diagram Kelas: Kelas FacetNet	54
4.5	Diagram Kelas: Kelas GUI	57
4.6	Diagram Kelas: Kelas DrawGraph	58
4.7	Rancangan Antarmuka Halaman Awal	66
4.8	Rancangan Antarmuka Mencari Fail	66
4.9	Rancangan Antarmuka Hasil Penghitungan Soft Modularity	67
4.10	Rancangan Antarmuka Hasil Proses Algoritma FacetNet	67
4.11	Rancangan Antarmuka Melihat Visualisasi Graf	68
5.1	Tampilan Halaman Awal	76
5.2	Tampilan Halaman Pemilihan Fail	77
5.3	Tampilan Halaman Setelah Proses Selesai	77
5.4	Tampilan Halaman Setelah Menutup Kotak Pesan	78
5.5	Tampilan Halaman Visualisasi Graf Masukan	78
5.6	Graf dengan 5 Simpul dan 5 Sisi	79
5.7	Hasil Uji Jumlah Komunitas 2	79
5.8	Graf dengan 10 Simpul dan 11 Sisi	80
5.9	Hasil Uji Jumlah Komunitas 2	81
5.10	Graf dengan 15 Simpul dan 17 Sisi	82
5.11	Hasil Uji Jumlah Komunitas 2	83

DAFTAR TABEL

2.1	Tabel Representasi Graf G_1	10
2.2	Tabel Representasi Graf G_5	10
2.3	Tabel Representasi Graf G_3	11
2.4	Tabel Representasi Graf G_4	11
2.5	Tabel Representasi Graf G_1	11
2.6	Tabel Representasi Graf G_3	11
3.1	Tabel Pembagian Komunitas	45
4.1	Use Case: Lihat Graf	51
4.2	Use Case: Deteksi Graf	52
5.1	Tabel Hasil Uji Graf 5 Simpul	80
5.2	Tabel Pembagian Komunitas Graf 15 Simpul	80
5.3	Tabel Hasil Uji Graf 10 Simpul	81
5.4	Tabel Pembagian Komunitas Graf 15 Simpul	82
5.5	Tabel Hasil Uji Graf 15 Simpul	84
5.6	Tabel Hasil Eksperimen	85

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Saat ini teknologi sudah sangat berkembang dalam berbagai bidang. Salah satu perkembangan teknologi adalah pada bidang komunikasi. Kini banyak sekali media komunikasi yang ditawarkan. Telepon adalah salah satu hasil perkembangan teknologi di bidang komunikasi. Dengan munculnya telepon, orang-orang kini memilih menggunakan telepon dibandingkan mengirim surat untuk berkomunikasi karena kemudahan dan kecepatan yang ditawarkan oleh telepon. Banyaknya penggunaan telepon menghasilkan data telepon yang melimpah. Data telepon ini dapat dianalisis dan dimanfaatkan dalam berbagai bidang pekerjaan.

Selain telepon, media sosial adalah hasil perkembangan teknologi dalam bidang komunikasi yang saat ini lebih dipilih oleh orang-orang sebagai media komunikasi mereka. Hampir setiap orang saat ini sudah menggunakan *smartphone* yang dapat mengakses internet di manapun dan kapanpun. Penggunaan media sosial sebagai media komunikasi dipilih oleh orang-orang karena internet menawarkan kemudahan, kecepatan, dan biaya yang lebih murah. Media sosial juga menghasilkan banyak sekali data dalam berbagai bentuk dan dalam berbagai ukuran. Data yang didapat dari penggunaan media sosial ini juga sangat berguna dan dapat dimanfaatkan dalam berbagai bidang pekerjaan.

Data telepon dan data media sosial adalah contoh data yang berukuran sangat besar dan memiliki pola. Kedua data ini tersedia sangat banyak dan dapat dianalisis untuk beberapa tujuan. Salah satu tujuan yang bisa didapat dari menganalisis kedua data ini adalah untuk mendeteksi atau mencari komunitas aktif. Definisi komunitas dalam Kamus Besar Bahasa Indonesia (KBBI) adalah kelompok organisme (orang dan sebagainya) yang hidup dan saling berinteraksi di dalam daerah tertentu. Definisi aktif dalam Kamus Besar Bahasa Indonesia (KBBI) adalah giat (bekerja, berusaha). Dari kedua definisi tersebut dapat disimpulkan bahwa komunitas aktif adalah kelompok organisme atau individu yang hidup dan saling berinteraksi dengan giat atau rajin di dalam daerah tertentu. Jika ditemukan komunitas aktif yang relevan, maka hal ini dapat dimanfaatkan sebagai media pemasaran perusahaan sesuai dengan bidangnya. Perusahaan dapat menentukan target pemasaran mereka dengan mengetahui informasi tentang komunitas aktif sesuai bidang perusahaan tersebut.

Graf dibentuk oleh simpul dan sisi yang menghubungkan simpul-simpul. Setiap simpul dapat dihubungkan dengan satu atau lebih sisi. Setiap sisi dapat memiliki atau tidak memiliki bobot. Graf berbobot adalah graf yang memiliki bobot pada setiap sisinya. Pada data telepon atau data media sosial dapat dilakukan pra-olah data untuk menyimpan data tersebut dalam format graf berbobot.

Data yang dianalisis untuk mendeteksi suatu komunitas dapat berupa data graf. Hal ini karena data graf memiliki pola yang dapat menggambarkan suatu komunitas. Satu simpul pada graf dapat menggambarkan satu individu atau objek. Satu sisi pada graf yang menghubungkan satu simpul dengan simpul lain dapat menggambarkan suatu interaksi dari dua simpul tersebut. Bobot pada sisi tersebut dapat menggambarkan banyaknya interaksi yang dilakukan oleh kedua simpul. Untuk menganalisis data graf dapat digunakan satu algoritma berbasis *clustering*, yaitu algoritma FacetNet.

Algoritma FacetNet (*A Framework for Analyzing Communities and Their Evolutions in Dynamic Networks*) adalah algoritma yang dapat digunakan untuk mendeteksi komunitas aktif pada data graf serta perubahannya seiring berjalannya waktu. Algoritma FacetNet menggunakan nilai peluang sebagai penentu komunitas untuk setiap individu. Algoritma ini menyediakan skema iterasi untuk menganalisis data graf sehingga dapat menentukan komunitas yang baik. Algoritma FacetNet juga dapat menganalisis perubahan komunitas seiring berjalannya waktu.

Fokus pada penelitian ini adalah menganalisis data graf untuk menemukan komunitas aktif. Sebuah perangkat lunak akan dibuat untuk mendeteksi komunitas aktif tersebut. Perangkat lunak akan menerima data graf dan menganalisis data graf tersebut menggunakan algoritma FacetNet. Perangkat lunak akan menghasilkan data berupa tabel yang berisi informasi tentang komunitas yang ditemukan.

1.2 Rumusan Masalah

Berdasarkan latar belakang pada subbab 1.1, maka terdapat beberapa rumusan masalah yang dapat dirumuskan sebagai berikut:

1. Bagaimana cara kerja algoritma FacetNet untuk mendeteksi komunitas aktif?
2. Bagaimana cara membangun perangkat lunak untuk mendeteksi komunitas aktif dari data graf dengan memanfaatkan algoritma FacetNet?
3. Bagaimana cara melakukan eksperimen pada perangkat lunak untuk menguji kinerja perangkat lunak pendeteksi komunitas aktif?

1.3 Tujuan

Berdasarkan rumusan masalah pada subbab 1.2, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Mempelajari algoritma FacetNet untuk mendeteksi komunitas aktif.
2. Membangun perangkat lunak untuk mendeteksi komunitas aktif dari data graf dengan memanfaatkan algoritma FacetNet.
3. Melakukan eksperimen untuk menguji kinerja perangkat lunak pendeteksi komunitas aktif.

1.4 Batasan Masalah

Batasan-batasan masalah yang diambil dalam melakukan penelitian adalah sebagai berikut:

1. Data graf yang digunakan adalah data graf berbobot tak berarah.
2. Data graf yang digunakan untuk eksperimen merupakan data yang diunduh dari website <http://konect.uni-koblenz.de/>.
3. Data graf hanya dianalisis untuk menemukan komunitas aktif pada satu waktu. Perubahan komunitas pada data graf tidak dianalisis pada penelitian ini.

1.5 Metodologi Penelitian

Metodologi yang dilakukan dalam melakukan penelitian ini adalah sebagai berikut:

1. Melakukan studi literatur

Studi literatur dilakukan untuk mempelajari beberapa teori yang dibutuhkan dalam melakukan penelitian ini. Teori-teori yang dibutuhkan adalah teori mengenai graf, teknik clustering, dan algoritma FacetNet.

2. Membangun perangkat lunak

Proses pembangunan perangkat lunak dimulai dengan membaca data graf dan menentukan jumlah komunitas yang diinginkan. Kemudian data graf tersebut akan dideteksi komunitasnya menggunakan algoritma FacetNet sesuai jumlah komunitas yang diinginkan.

3. Melakukan pengujian terhadap perangkat lunak

Pengujian perangkat lunak dilakukan untuk menguji kebenaran hasil analisis dan kinerja dari perangkat lunak.

4. Menulis dokumen

Penulisan dokumen akan dibagi menjadi 6 bab utama yang akan dijelaskan pada subbab Sistematika Pembahasan.

1.6 Sistematika Pembahasan

Pada laporan penelitian ini akan terdapat enam bab yang ditulis secara sistematis. Enam bab tersebut dapat dijelaskan secara singkat sebagai berikut:

1. Bab 1 merupakan bab pendahuluan yang berisi latar belakang, rumusan masalah, tujuan, batasan masalah, metodologi penelitian, dan sistematika pembahasan.
2. Bab 2 berisi dasar teori hasil dari studi literatur mengenai graf, teknik dan algoritma *clustering*, serta algoritma FacetNet.
3. Bab 3 berisi komputasi algoritma FacetNet.
4. Bab 4 adalah bab analisis dan perancangan perangkat lunak. Pada bab ini terdapat *use case*, penjelasan struktur data graf, perancangan diagram kelas, *pseudo code*, dan perancangan *user interface*.
5. Bab 5 merupakan bab implementasi dan pengujian perangkat lunak. Pada bab ini akan dicantumkan potongan kode program, penjelasan cara melakukan pengujian perangkat lunak, dan kesimpulan hasil pengujian.
6. Bab 6 berisi kesimpulan dari hasil penelitian ini dan saran yang mungkin akan berguna untuk penelitian-penelitian selanjutnya.