

HIERARCHICAL MULTI-VIEWPOINT SELF ORGANIZING MAP PADA PENGELOMPOKAN PENGGUNA UNTUK MENGETAHUI PROFIL UNDUH DI LINGKUNGAN KAMPUS

Tesa Eranti Putri, Chastine Fatichah, Diana Purwitasari
Institut Teknologi Sepuluh Nopember Surabaya
tesaep@gmail.com

Abstrak. Fasilitas internet kampus terkadang disalahgunakan untuk mengunduh data yang tidak terkait akademis, sehingga mengganggu pengguna yang memerlukan internet untuk kepentingan akademis. Guna memonitor pengunduhan di kampus, diperlukan profil unduh. Profil dapat dibentuk dari pengelompokan pengguna. Penelitian ini mengajukan pemodelan untuk pengelompokan pengguna berupa *Self Organizing Map hierarchical multi-viewpoint*. Pengelompokan dilakukan berdasarkan jumlah transaksinya, dilihat dari banyak viewpoint. Setiap viewpoint membentuk peta tersendiri, disusun berjenjang (*hierarchical*), kemudian dilatih menggunakan gabungan fitur viewpoint sekarang dengan viewpoint di bawahnya. Pengujian dilakukan dengan analisis manual hasil pengelompokan. Dari pengujian, diperoleh fitur viewpoint yang memberikan gambaran profil unduh yang paling jelas dan lengkap adalah domain email.

Kata kunci. *profil unduh, pengelompokan pengguna, web usage mining, Self Organizing Map Hierarchical Multi-viewpoint*

Fasilitas internet gratis di kampus terkadang disalahgunakan untuk melakukan pengunduhan berkas data berukuran besar yang tidak berhubungan dengan akademis. Pemakaian bandwidth internet kampus untuk mengunduh ini berpotensi menurunkan kualitas kecepatan akses bagi pengguna lain yang memerlukan layanan internet untuk kepentingan akademis. Oleh karena itu, diperlukan informasi berupa profil unduh pengguna internet di lingkungan kampus.

Profil unduh pengguna diartikan sebagai suatu deskripsi yang menyatakan perilaku unduh pengguna tertentu. Sebagai contoh, terdapat sekelompok pengguna internet di kampus yang memiliki profil selalu melakukan pengunduhan antara jam 13.00-15.00 di sekitar area rektorat kampus dan sekelompok pengguna tersebut sebagian besar berasal dari jurusan tepat di samping rektorat. Profil unduh pengguna dapat diketahui dari pengelompokan pengguna berdasarkan preferensinya dalam mengunduh. Pengelompokan dapat dilakukan dengan teknik *web usage mining* [1]. Dokumen *web access log* yang dihasilkan oleh *proxy server* kampus dapat dimanfaatkan sebagai sumber untuk proses *web usage mining* ini.

Dalam sepuluh tahun terakhir, terdapat beberapa penelitian terkait dengan eksplorasi dokumen *web access log* sebagai sumber *web*

usage mining. Eksplorasi tersebut antara lain untuk pengenalan pola navigasi pengguna dan tren jam sibuk di kampus [2] dan peningkatan layanan jaringan internet dengan melakukan analisis tingkat *error* pada dokumen *web access log* [1]. Contoh lainnya adalah untuk memberi prediksi rekomendasi situs yang serupa berdasarkan pola navigasi pengguna [3][4]. Sebagian besar dari penelitian tersebut memanfaatkan algoritma SOM untuk melakukan *web usage mining* pada dokumen *web access log*. SOM memiliki kelebihan dari segi visual berupa pemetaan data berdimensi tinggi ke peta berdimensi rendah, dengan tetap mempertahankan topologi dan relasi yang tersembunyi antar data. Namun, SOM juga memiliki beberapa kelemahan. Berdasarkan dari penelitian sebelumnya [3][5], kelemahan dari algoritma SOM ini antara lain:

- 1) Kecepatan konvergensi peta SOM bergantung pada ukuran vektor *input / masukan* (vektor fitur), ukuran peta, dan inialisasi parameter pelatihan SOM.
- 2) Fitur yang mewakili suatu obyek data dimampatkan nilainya, sehingga kurang bisa mendeskripsikan karakteristik data secara mendetil
- 3) Tidak adanya pembatas yang jelas antar cluster dengan topik serupa yang bertetangga, sehingga peta SOM terlihat

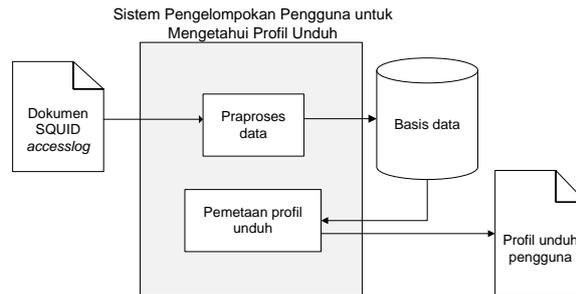
kurang tematis dan tidak bisa dilakukan analisis kuantitatif distribusi data.

Sebagai solusi untuk mengatasi kelemahan SOM di atas, dikembangkan *Multiple SOM* yang menambahkan jumlah peta yang digunakan, dari satu menjadi banyak (*multiple*) [5]. Beberapa penelitian yang menggunakan konsep *multiple SOM* ini antara lain algoritma *Multi-SOM* [6] dan *Hierarchical-SOM* (HSOM) [7]. Algoritma *Multi-SOM* membagi fitur dataset ke dalam banyak sudut pandang (*multi-viewpoint*) berbeda dan masing-masing *viewpoint* direpresentasikan dalam sebuah peta SOM, sehingga satu dataset yang sama dapat dianalisis dari berbagai *viewpoint*. Sedangkan algoritma HSOM menyusun dua buah peta SOM dalam sebuah susunan berjenjang (*hierarchical*), kemudian menggabungkan indeks *node* pemenang dari pelatihan SOM lapisan pertama dengan fitur tambahan yang diperoleh setelah pelatihan SOM lapisan pertama selesai sebagai vektor *input* untuk pelatihan SOM lapisan kedua.

Penelitian ini mengajukan kombinasi algoritma *Multi-SOM* dan HSOM membentuk SOM *hierarchical multi-viewpoint* dalam rangka melakukan pengelompokan pengguna untuk mengetahui profil unduh. Dengan SOM *hierarchical multi-viewpoint* ini, fitur yang mendeskripsikan preferensi unduh pengguna akan dipecah ke dalam *viewpoint-viewpoint* dan setiap *viewpoint* membentuk peta SOM tersendiri yang akan disusun secara *hierarchical*. Kemudian setiap peta SOM *viewpoint* tersebut dilatih dan digunakan untuk mengelompokkan pengguna berdasarkan *viewpoint* pada peta tersebut. Pelatihan akan menggunakan vektor fitur yang merupakan gabungan dari vektor fitur pada *viewpoint* tersebut dengan fitur berupa indeks *node* pemenang dari *viewpoint* di bawahnya sebagai vektor *input*. Diharapkan dengan pemodelan ini, profil unduh pengguna dapat diperoleh sekaligus mengatasi kelemahan-kelemahan dari SOM.

I. METODOLOGI

Secara keseluruhan, sistem pengelompokan pengguna untuk mengetahui profil unduh ini terdiri atas dua modul utama, yaitu: penyaringan data dan pe. Diagram yang menggambarkan alur kerja sistem ini dapat dilihat pada **Gambar 1**.



Gambar 1 Alur kerja sistem pengelompokan pengguna untuk mengetahui pola unduh

Praproses Data

Dokumen *SQUID access log* mentah berukuran sangat besar (± 3 GB per dokumen) dan tidak semuanya berisi transaksi unduhan. Modul praproses data bertujuan untuk melakukan penyaringan pada dokumen mentah tersebut, sehingga mengurangi kemungkinan terjadinya bias dalam pemetaan SOM akibat adanya data tidak valid (*noise*) sekaligus mereduksi waktu proses dan jumlah data yang harus diproses sistem karena hanya mengambil data transaksi unduhan saja. Sebelum disaring, setiap dokumen akan dipecah ke dalam dokumen yang lebih kecil berukuran 100 MB dengan menggunakan *tool file splitter* di luar sistem.

Dalam penelitian ini, terdapat dua macam sumber ketidakvalidan suatu baris data dalam dokumen *SQUID access log*, yaitu [3] eksternal dan internal. Sumber ketidakvalidan eksternal berasal dari faktor di luar dokumen *SQUID access log*, seperti baris data yang terpotong akibat proses pemecahan dokumen dengan *tool file splitter*. Sedangkan sumber ketidakvalidan internal berasal dari atribut yang berada di dalam dokumen *SQUID access log* sendiri, yaitu ketika nilai dari satu atau lebih atribut tidak memenuhi kriteria data yang telah ditentukan sebelumnya. Sebagai contoh: suatu baris transaksi yang tercatat dikatakan tidak valid jika transaksi tersebut bukanlah transaksi unduhan.

Langkah selanjutnya dalam praproses data adalah ekstraksi atribut, yaitu mengambil dan memproses data transaksi yang masuk ke dalam basis data untuk memperoleh bentuk data bersih yang dapat digunakan untuk tahap selanjutnya. Pada tahapan ini, pengguna akan dikelompokkan ke dalam segmen-segmen berdasarkan jurusan asal pengguna dan waktu transaksi unduhannya. Contoh segmentasi pengguna ini dapat dilihat pada **Tabel 1**.

Tabel 1 Segmentasi pengguna berdasarkan waktu dan jurusan

ID segmentasi	Hari	Jam	ID Jurusan
1	Senin	00.00-00.59	1
2	Senin	01.00-02.00	1
...
4536	Minggu	23.00-23.59	27

Setiap transaksi unduhan akan ditandai dengan nomor segmentasi yang sesuai. Kemudian setiap transaksi juga akan diberi label sesuai dengan fitur-fitur yang akan digunakan sebagai *viewpoint* pada penelitian ini. Fitur-fitur yang dipilih untuk dijadikan *viewpoint* yaitu: tipe data yang diunduh, lokasi pengunduhan, dan domain email pengunduh. Contoh data transaksi yang sudah ditandai dengan jenis fitur pengunduhan yang bersesuaian dapat dilihat pada **Tabel 2**.

Dari tabel transaksi yang sudah ditandai tersebut, akan dikonstruksi vektor fitur sejumlah *viewpoint* untuk digunakan dalam proses pemetaan profil unduh. Berdasarkan **Tabel 2**, dapat dikonstruksi vektor fitur sebagai berikut:

1) Vektor fitur tipe data

Vektor fitur ini berupa jumlah transaksi pengunduhan berdasarkan tipe datanya. Vektor fitur yang dihasilkan dapat dilihat pada persamaan (1) sebagai berikut.

$$x_{TD}8 = \{ 1, 2 \} \tag{1}$$

dimana:

$x_{TD}8$ = vektor fitur untuk *viewpoint* tipe data untuk segmen pengguna ke-8. Angka pertama merupakan jumlah transaksi untuk tipe data

video (x_{TD1}) dan angka kedua merupakan jumlah transaksi untuk tipe data *compressed / archive* (x_{TD2}).

2) Vektor fitur lokasi

Vektor fitur ini berupa jumlah transaksi pengunduhan berdasarkan lokasi pengunduhannya. Vektor fitur yang dihasilkan dapat dilihat pada persamaan (2) sebagai berikut.

$$x_{L}8 = \{ 1, 1, 1 \} \tag{2}$$

dimana:

$x_{L}8$ = vektor fitur untuk *viewpoint* lokasi unduh untuk segmen pengguna ke-8. Angka ke-1 hingga ke-3 berturut-turut adalah jumlah transaksi untuk lokasi perpustakaan dan rektorat (x_{L1}), untuk lokasi FTIF (x_{L2}), dan untuk lokasi FMIPA (x_{L3}).

3) Vektor fitur domain email

Vektor fitur ini berupa jumlah transaksi pengunduhan berdasarkan fakultas pemilik domain email dari pengguna yang mengunduh. Vektor fitur yang dihasilkan dapat dilihat pada persamaan (3) sebagai berikut.

$$x_{DE}8 = \{ 3, 0, 0, 0, 0 \} \tag{3}$$

dimana:

$x_{DE}8$ = vektor fitur untuk *viewpoint* fakultas domain *email* untuk segmen pengguna ke-8. Angka ke-1 hingga ke-5 berturut-turut adalah jumlah transaksi untuk domain *email* pengunduh yang berasal dari FMIPA, FTI, FTSP, FTK, dan FTIF.

Tabel 2 Contoh transaksi unduhan yang sudah ditandai

ID Transaksi	ID Segmentasi	ID Domain Email	ID Fakultas	ID Tipe Data	ID IP	ID Lokasi
1	8	51	1	2	37	3
2	8	51	1	1	23	2
3	8	51	1	2	17	1

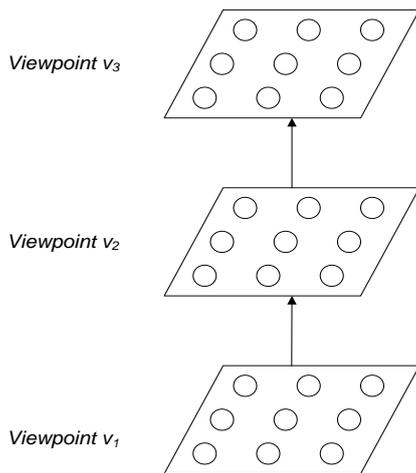
Pemetaan Profil Unduh

SOM pada umumnya hanya terdiri atas dua lapisan saja, yaitu lapisan *input* dan lapisan *output* berupa peta SOM berukuran satu atau dua dimensi. Namun, dengan arsitektur *hierarchical multi-viewpoint*, peta SOM *output* yang digunakan tidak hanya satu lapis saja dan

lapisan-lapisan peta *output* ini disusun secara *hierarchical*.

Jumlah lapisan peta SOM yang digunakan pada penelitian ini adalah tiga lapis, disesuaikan dengan jumlah fitur yang akan dijadikan *viewpoint*. Peta SOM pada tahapan ini disusun mulai dari *viewpoint* v_1 sebagai lapisan paling bawah dan *viewpoint* v_3 sebagai lapisan teratas.

Agar hasil pemetaan lebih berfokus pada pengaruh vektor fitur daripada oleh faktor lain, maka ukuran dimensi peta pada tiap lapisan *viewpoint* disamakan, yaitu 5 x 5. Parameter-parameter lain seperti vektor bobot (w_j), parameter *learning rate* (α) dan jumlah iterasi (T) juga akan disamakan untuk setiap peta *viewpoint*. Konstruksi susunan peta *viewpoint* dapat dilihat pada **Gambar 2**.



Gambar 2 Konstruksi SOM *hierarchical multi-viewpoint*

Setelah pelatihan SOM pada satu *viewpoint* selesai dilakukan, maka peta yang telah terbentuk ini akan dibagi-bagi ke dalam area-area yang mana setiap area ini mewakili satu tipe pola unduh tertentu. Pengelompokan area ini dilakukan terlebih dahulu sebelum naik ke pelatihan SOM lapisan selanjutnya. Tujuan dari pembagian area terlebih dahulu ini adalah agar karakteristik pola unduh global yang sudah terbentuk pada peta sebelumnya dapat dibawa naik ke lapisan berikutnya.

Modul pembagian area ini pada dasarnya membentuk *cluster-clusteer* yang di dalamnya berisi *node-node* yang memiliki karakteristik yang serupa. Jika umumnya karakteristik tersebut dilihat dari kedekatan jarak di antara *node* yang bertetangga [8], maka pada modul ini, keserupaan karakteristik tersebut diukur dari kesamaan posisi indeks bobot yang maksimal.

Vektor bobot w_j pada setiap *node* yang telah melalui proses pelatihan SOM akan mengalami perubahan nilai sedemikian hingga data-data yang mirip akan tertarik masuk ke dalam *node* tersebut. Hal ini berarti vektor bobot w_j menyimpan karakteristik pengelompokan data yang berada di dalamnya.

Modul pembagian area memanfaatkan karakteristik pengelompokan yang terbentuk pada vektor bobot w_j untuk membagi peta ke dalam area-area. Suatu *node* akan terkelompok bersama *node* lain dalam sebuah area jika *node-node* tersebut memiliki indeks posisi bobot dominan yang sama. Hal ini berarti jika *node* A memiliki bobot dominan pada posisi pertama dan *node* B juga memiliki bobot dominan pada posisi pertama yang sama, maka *node* A dan B akan terkelompok menjadi satu area.

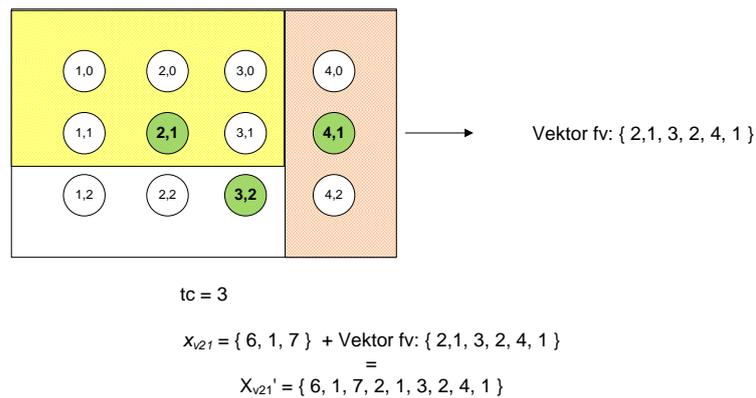
Langkah selanjutnya setelah semua area terbentuk adalah melakukan integrasi vektor fitur untuk digunakan sebagai *input* lapisan *viewpoint* di atasnya. Langkah ini dilakukan jika *viewpoint* sekarang belum mencapai *viewpoint* teratas (v_3). Algoritma integrasi vektor fitur yang diuraikan pada [7] akan dimodifikasi dengan mempertimbangkan pembagian area yang terbentuk pada peta.

Pada modul ini, setelah peta terbagi ke dalam area, dilakukan perhitungan jumlah *node* yang terkelompok di masing-masing area. Kemudian, ditetapkan sebuah nilai *threshold* (tc). Nilai tc ini menentukan jumlah minimal *node* yang terkelompok ke dalam sebuah area. Jika jumlah *node* yang terkelompok dalam area kurang dari nilai tc , maka area tersebut akan diabaikan. Dari area yang tersisa, selanjutnya dihitung jumlah data yang terkelompok untuk setiap *node* (k_j) per area. Kemudian, pada setiap area, dicari *node* yang memiliki nilai k_j paling besar untuk diambil nilai koordinat posisi *node* tersebut untuk diintegrasikan dengan vektor fitur selanjutnya. Ilustrasi proses integrasi vektor fitur ini dapat dilihat pada **Gambar 3**.

Pengujian

Pengujian yang dilakukan bertujuan untuk mengetahui apakah fungsionalitas sistem pengelompokan untuk mengetahui profil unduh sudah berjalan dengan baik. Selain itu, pengujian ini juga mencari fitur *viewpoint* yang mampu memberikan hasil pengelompokan yang menunjukkan gambaran profil unduh di kampus dengan lengkap.

Uji coba dilakukan pada 4536 segmen pengguna berdasarkan jurusan dan waktu unduh yang mencakup 16.452 data transaksi unduh dari kampus ITS antara tanggal 1-9 Juli 2014. Untuk setiap skenario dan *viewpoint*, parameter pelatihan, ukuran peta *output*, dan matriks vektor bobot yang digunakan akan disamakan.



Gambar 3 Ilustrasi Contoh Integrasi Tambahan Vektor Fitur

Pengujian dijalankan pada parameter *learning rate* $\alpha = 0,5$ dan jumlah iterasi $T = 40$. Setiap peta pada setiap lapisan *viewpoint* diinisialisasi dengan ukuran 5×5 . Ukuran dan nilai vektor bobot diinisialisasi pada pelatihan *viewpoint* lapisan I saja. Vektor bobot ini kemudian digunakan untuk setiap lapisan *viewpoint* yang dilatih. Susunan *viewpoint* ditetapkan dari bawah ke atas dengan urutan: lokasi – domain email – tipe data. Ukuran vektor fitur untuk setiap *viewpoint* berbeda, disesuaikan dengan jenis *viewpoint* yang digunakan, sesuai dengan yang diuraikan pada subbab Praproses Data.

II. HASIL DAN PEMBAHASAN

Peta hasil pengelompokan *viewpoint* I lokasi, *viewpoint* II domain email, dan *viewpoint* III tipe data berturut-turut dari atas ke bawah dapat dilihat pada Gambar 4, Gambar 5, dan **Gambar 6**.

Viewpoint I terbagi ke dalam 2 area, yaitu Q2 dan Q3. Area Q2 memiliki 4420 segmen pengguna yang terkelompok di dalamnya sementara Q3 hanya 116. Hal ini menunjukkan transaksi unduh didominasi oleh pengunduhan yang terjadi di daerah kampus FTIF (fitur ke-2 pada vektor fitur lokasi). Berdasarkan hasil penelusuran dan analisis kontennya, ternyata diperoleh bahwa area Q2 merupakan area berisi profil unduh pengguna dari FMIPA ITS yang memiliki jumlah transaksi sedang (± 30 transaksi). Sedangkan, area Q3 berisi profil unduh pengguna FTIF yang memiliki jumlah transaksi tinggi (*heavy-hitter*, ± 50 transaksi) dan bercampur dengan profil unduh pengguna yang jarang mengunduh. Hal ini dikarenakan

vektor fitur berisi jumlah transaksi dari pengguna FTIF lebih banyak dan dominan, sehingga pengguna yang jarang mengunduh yang jumlah transaksinya lebih sedikit banyak yang terkumpul di sini.

Viewpoint II domain email terbagi ke dalam 4 area yang berbeda, yaitu: Q1 (84 anggota), Q2 (66 anggota), Q5 (116 anggota), dan Q7 (4270 anggota). Q7 adalah area yang muncul dari fitur tambahan yang diperoleh dari *viewpoint* I. Berdasarkan hasil pengujian pada *viewpoint* I dan setelah menelusuri konten Q7, Q7 yang memiliki *node* dengan jumlah anggota 3596 menunjukkan profil unduh pengguna yang jarang mengunduh. Area Q1 berisi profil pengunduh segmen pengguna yang berasal dari FMIPA dengan jumlah transaksi pengunduhan sedang. dan area Q5 berisi profil pengunduh *heavy hitter* dari FTIF yang kini berdiri sendiri karena profil unduh pengguna yang jarang melakukan pengunduhan masuk ke fitur tambahan. Sedangkan area Q2 merupakan profil unduh pengguna dari FTI yang jumlah transaksinya sedang.

Viewpoint III tipe data terbagi ke dalam 2 area, Q2 (310 anggota) dan Q3 (4226 anggota). Q3 adalah fitur yang muncul dari sebagai tambahan dari *viewpoint* II. Karena pada data transaksi unduhan didominasi oleh tipe data archive yang merupakan fitur x_{TD2} , maka data pada area Q2 yang terbentuk dari fitur x_{TD2} tersebut secara tidak langsung menunjukkan profil unduh pengguna yang pernah mengunduh dan area Q3 yang memiliki *node* yang berjumlah banyak merupakan profil unduh dari pengguna yang jarang mengunduh.

Q3 (2)	Q3 (0)	Q3 (0)	Q3 (1)	Q3 (2)
Q3 (0)	Heavy hitter (FTIF) dan yang jarang mengunduh	Q3 (1)	Q3 (3)	Q2 (0)
Q3 (14)	Q3 (2)	Q3 (6)	Sedang (FMIPA)	Q2 (14)
Q2 (69)	Q3 (16)	Q3 (0)	Q2 (27)	Q2 (15)
Q2 (4126)	Q2 (76)	Q3 (32)	Q2 (62)	Q2 (31)

Gambar 4 Hasil Uji Coba Viewpoint I Lokasi

Q2 (2)	Sedang (FTI)	Q2 (7)	Q2 (0)	Sedang (FMIPA)
Q2 (1)	Q2 (2)	Q2 (0)	Q1 (18)	Q1 (5)
Q2 (3)	Q2 (3)	Q2 (3)	Q1 (16)	Q1 (44)
Q2 (0)	Q2 (9)	Q2 (32)	Q7 (5)	Jarang mengunduh
Q5 (69)	Heavy hitter (FTIF)	Q7 (84)	Q7 (400)	Q7 (185)
	Q5 (17)		Q7 (3596)	Q7 (3596)

Gambar 5 Hasil Uji Coba Viewpoint II Domain Email

Q2 (2)	Q2 (1)	Q2 (5)	Q2 (4)	Q2 (16)
Q2 (3)	Q2 (7)	Q2 (2)	Q2 (9)	Q2 (15)
Q2 (10)	Q2 (5)	Q2 (5)	Q2 (31)	Q2 (41)
Q2 (9)	Q2 (13)	Q2 (3)	Q2 (91)	Q3 (197)
Q2 (17)	Q2 (13)	Q2 (10)	Q3 (116)	Jarang mengunduh
				Q3 (315)

Gambar 6 Hasil Uji Coba Viewpoint III Tipe Data

III. SIMPULAN

Penelitian ini mengajukan pemodelan berupa *Self Organizing Map (SOM) Hierarchical Multi-viewpoint* untuk melakukan pengelompokan pengguna guna mengetahui profil unduh di lingkungan kampus. Pemodelan ini diterapkan pada dataset yang terdiri atas 4536 segmen pengguna berdasarkan jurusan dan waktu unduh yang mencakup 16.452 data transaksi unduh dari kampus ITS antara tanggal 1-9 Juli 2014. Hasil uji coba menunjukkan pemodelan dengan SOM *hierarchical multi-viewpoint* ini mampu memberikan gambaran profil unduh dari pengguna-pengguna yang berada di lingkungan kampus ITS dengan jelas, Selain itu, diperoleh kesimpulan bahwa fitur yang digunakan sebagai *viewpoint* yang memberikan gambaran profil unduh paling lengkap adalah fitur domain email.

Ke depannya, data transaksi yang digunakan dalam penelitian dapat ditambah rentang waktu pengamatannya, sehingga hasil pengelompokan pengguna dapat menggambarkan profil unduh pengguna di lingkungan kampus dengan lebih lengkap.

IV. DAFTAR PUSTAKA

[1] Tyagi, N.K., Solanki, A.K., dan Wadhwa, M. 2010. *Analysis of Server Log by Web Usage Mining for Website Improvement*. IJCSI International Journal of Computer Science Issues Vol 7:4, pp. 17-21.

[2] Wang, X., Abraham, A., dan Smith, K. A. 2005. *Intelligent Web Traffic Mining and Analysis*. Journal of Network and Computer Applications 28, pp. 147-165.

[3] Putri, Tesa E. 2011. *Implementasi Analisa Weblog untuk Rekomendasi Link Favorit dengan Kombinasi Algoritma SOM dan Fuzzy*. (Skripsi S1, Institut Teknologi Sepuluh Nopember).

[4] Mishra, A.K., Mishra, M.K., Chaturvedi, V., Gupta, S.K. dan Singh, J. 2013. *Web Usage Mining Using Self Organizing Map*. IJARCSSE International Journal of Advanced Research in Computer Science and Software Engineering Vol 3:6, pp. 532-539.

[5] Khanchouch, I., Boujenfa, K., dan Limam, M. 2013. *An Improved Multi-SOM Algorithm*. IJNSA International Journal of Network Security & Its Applications Vol 5:4, pp. 181-186.

[6] Polanco, X., dan Hoffmann, M. 2000. *Visual research topic discovery with MULTI-SOM model of analysis*. Francia (2000).

[7] Smith, Toby. 2009. *Adapting to Increasing Data Availability using Multi-Layered Self Organizing Maps*. Proceedings of 2009 International Conference on Adaptive and Intelligent Systems, pp. 108-113.