

Penggunaan Metode berbasis Graph untuk Mining Frequent Sequential Access Pattern Pada Studi Kasus : Website iGracias Universitas Telkom

Rahmi Rohdiniyah ^{#1}, Ibnu Asror M.T. ^{*2}, Gede Agung Ary W M.T. ^{#3}

Fakultas Informatik Universitas Telkom
Jl. Telekomunikasi No.1, Bandung Indonesia.

¹ rahmirohiniyah@gmail.com

² iasror@telkomuniversity.ac.id

² degunk@telkomuniversity.ac.id

Abstract

The use of website in education, especially in university, is mainly to save various pieces of information provided by the university. For that reason, it is important to perform a quality test on the website, so, it would be able to give the best service to the users. One of the techniques to test the quality of a website is by using web usage mining. Web usage mining is one of the branches of web mining that is used to find useful information or knowledge from user routine on a website. This paper performs a quality test on Telkom University's Igracias website using a graph-based method for mining frequent sequential access patterns of website. This method is applied because there is only a node for each page or item, which means less memory usage. Besides, graph mining can be used to find access pattern using user's access behavior, even for the complex one. Those are the strengths of using graph. From this implementation method be produced frequent sequential access user group

Keywords: web mining, web usage mining, metode berbasis graph, frequent sequential access patterns.

Abstrak

Penggunaan *website* pada bidang pendidikan, khususnya sebuah universitas, bertujuan untuk menyimpan berbagai informasi yang ada pada lingkungan universitas tersebut. Untuk itu, perlu dilakukan perbaikan struktur untuk memelihara kualitas dari web. Salah satu teknik yang dapat digunakan adalah dengan menggunakan *web usage mining*. *Web usage mining* merupakan salah satu cabang dari *web mining* yang digunakan untuk menemukan informasi atau pengetahuan yang bermanfaat dari pola navigasi *user* pada sebuah *website*. Pada publikasi ini menggunakan metode berbasis graph untuk *frequent sequential access patterns* dan menggunakan Igracias Universitas Telkom sebagai studi kasusnya. Dimana pengguna terbesar igracias Universitas Telkom terdiri dari mahasiswa, dosen dan pegawai. Metode ini memiliki kelebihan untuk menemukan *behavior* pola pengaksesan *user*. Dari implementasi metoda ini didapat pola akses group user secara berurutan.

Kata Kunci: *web mining, web usage mining, metode berbasis graph, frequent sequential access patterns.*

I. INTRODUCTION

INTERNET sebagai sumber informasi pada saat ini terus mengalami perkembangan yang pesat, baik dalam hal jumlah pengguna internet maupun sumber informasi yang tersedia. Penggunaan internet dimanfaatkan oleh para pengembang, untuk menyampaikan informasi melalui website. Salah satu bidang yang turut serta melakukan pemanfaatan internet sebagai sarana pembantu untuk mendapatkan informasi yaitu pada bidang pendidikan. Penggunaan website pada universitas merupakan salah satu contoh penerapan internet oleh bidang pendidikan. Website universitas merupakan sarana digital yang dapat diakses seluruh pengguna internet maupun khusus untuk entitas akademik dan non-akademik yang ada pada universitas tersebut. Untuk itu mengetahui pola pengaksesan menjadi hal penting bagi pengembang. Agar dapat mengembangkan menu yang sudah ada berdasarkan pengaksesan yang dilakukan oleh user. Teknik web mining dapat digunakan untuk melihat frekuensi perilaku *browsing* pengguna yang tercatat dalam *web server logs*, *browser logs*, *proxy server logs*[1].

Web mining merupakan salah satu teknik yang digunakan untuk melakukan penggalian data tentang web. *Web mining* dibagi menjadi tiga klasifikasi, yaitu *web content mining*, *web usage mining* dan *web structure mining*. *Web usage mining* dapat digunakan untuk melihat pola perilaku user pada suatu *website*. Karena *web usage mining* bertujuan untuk menemukan informasi atau pengetahuan yang bermanfaat dari pola navigasi user pada web. Pada publikasi ini digunakan metode berbasis *graph* untuk *mining frequent sequential access patterns* suatu *website* dan dengan studi kasus *website* Igracias Universitas Telkom.

Penggunaan metode *graph* untuk *mining frequent sequential access patterns* dilakukan pada tahap *pattern discovery*. Metode ini digunakan karena kelebihan dari menggunakan *graph* adalah hanya ada satu *node* untuk sebuah *page* atau *item* yang mana dengan begitu menggunakan sedikit memori[1]. Selain itu, *graph mining* dapat digunakan untuk menemukan *behavior* pola pengaksesan dengan menggunakan tingkah laku pengaksesan pengguna yang kompleks sekalipun [1]. Metode ini menggunakan *minimum support* untuk melakukan *pruning* sebagai batas frekuensi minimum yang harus ditentukan terlebih dahulu oleh penggunaannya. *Minimum support* yang digunakan sebaiknya tidak terlalu besar dan tidak terlalu kecil [2]. Sehingga menghasilkan pola pengaksesan yang memang sering dilakukan oleh pengguna.

II. LITERATURE REVIEW

A. Web Mining

Web mining merupakan teknik pada *data mining* dimana datanya tidak struktur atau semi-struktur yang bertujuan untuk mengetahui dan mengekstraksi informasi atau *knowledge* yang berpotensi dari *web*[3]. Penggunaan *web mining* berguna untuk mencari informasi atau pengetahuan bermanfaat dari struktur *web hyperlinks*, halaman *web*, dan data penggunaan *web*. *Web mining* memiliki tiga klasifikasi, yaitu *web content mining*, *web structure mining* dan *web usage mining*.

B. Web Usage Mining

Web usage mining digunakan untuk menganalisis *web log files* untuk menemukan pola pengguna dalam mengakses sebuah halaman *web*. *Web usage mining* merupakan fokus utama dalam *web mining* yang mempelajari interaksi dari pengguna *web*[5]. *Web usage mining* digunakan untuk menemukan pola yang bermakna dari data yang dihasilkan oleh transaksi *client-server* pada satu atau lebih *web server*.

Web usage mining dikategorikan ke dalam tiga fase[6]:

1. *Preprocessing* : Berdasarkan klien , server dan server proxy , *preprocessing* adalah pendekatan pertama untuk mengambil data mentah dari sebuah *web* dan memproses data. Secara otomatis mengubah data mentah untuk digunakan pada proses selanjutnya.
2. *Pattern Discovery* : Berdasarkan dari data *preprocessing*, data mentah yang ada digunakan untuk menemukan pengetahuan dan menerapkan teknik yang akan digunakan pada *machine learning* yang menggunakan prosedur dari data mining.

Pattern Analysis : *Pattern Analysis* adalah proses setelah penemuan pola. *Pattern Analysis* memeriksa apakah pola yang ada pada *web* benar dan memandu proses ekstraksi informasi atau pengetahuan dari *web*.

C. Preprocessing

Beberapa *database* mempunyai data yang tidak konsisten dan terdapat noise. *Preprocessing* dilakukan agar data tersebut ditransformasi. *Log server logs* yang akan digunakan pada publikasi ini akan terlebih dahulu melalui proses *preprocessing*. Hasil *preprocessing* akan membuat data menjadi lebih terintegrasi dan konsisten. Tahapan dari *preprocessing* tersebut adalah untuk :

- 1) Membersihkan data, berupa filterisasi pada data yang dihasilkan *web page* berdasarkan *automatic request* (misalnya file gambar) yang secara spesifik tidak di *request* oleh *user*.
- 2) Identifikasi *user* berbeda. Pada identifikasi *user* akan digunakan informasi ID *login* sebagai idenitas unik dari tiap *user*.

Identifikasi session, dilakukan untuk mengetahui sesi tunggal yang dilakukan oleh *user*. Umumnya sesi akan dibagi berdasarkan waktu.

D. Event Log

Server log merupakan informasi yang berisikan daftar kegiatan yang dibuat secara otomatis dan dikelola oleh server[7]. *Web server log* berisikan : *remote host field (IP client)*, *authuser field*, waktu dan tanggal akses, *method request*, halaman yang diakses, protocol yang digunakan, jumlah *byte* yang diakses, halaman *referrer*, dan *agent* yang digunakan. Sebuah *access log* memiliki format yang berbeda-beda, salah satu contoh dari *access log* adalah seperti gambar 1[8]:

```
129.119.4.17 - - [16/Aug/1999:00:00:11 -0500] "GET /img/XredSeal.gif HTTP/1.1" 301 328
"http://www.seas.smu.edu/" "Mozilla/4.0 (compatible; MSIE 4.01; Windows NT) "
129.119.4.17 - - [16/Aug/1999:00:00:11 -0500] "GET /img/ecom.gif HTTP/1.1" 304 -
"http://www.seas.smu.edu/" "Mozilla/4.0 (compatible; MSIE 4.01; Windows NT)".
```

Gambar. 1. Contoh Server Access Log

Event log merupakan catatan history yang memuat rangkaian aktivitas yang dilakukan pengguna dalam sebuah sistem informasi. *Event log* pada sistem informasi memuat catatan proses yang dilakukan oleh tiap pengguna[9]. Setiap *event log* juga memiliki berbagai atribut seperti *case_id*, *activity*, *time*, *cost* dan *resource*. Namun tidak semua atribut tersebut selalu ada pada suatu *event log*. Misalnya adalah tidak adanya atribut *resource* maupun *cost*, tergantung dari data yang diperoleh dari suatu sistem informasi[10].

1	CASE_ID,ACTIVITY,EKSEKUTOR,WAKTU,GROUP_EKSEKUTOR,USERGROUPNAME
2	293,Presensi >> Berita Acara Perkuliahan,53013,9/9/2014 3:35:59 PM,3,DOSEN
3	337,Presensi >> Kehadiran,54871,9/9/2014 3:36:32 PM,12,MAHASISWA
4	,Presensi >> Berita Acara Perkuliahan,82601,9/9/2014 3:37:58 PM,3,DOSEN
5	,Data Mahasiswa >> Bantuan Jadwal >> Jadwal Dosen,48406,9/9/2014 3:38:09 PM,38,ADMIN DATA MAHASISWA FAKULTAS
6	353,dashboard,434,9/9/2014 3:38:13 PM,20,PEGAWAI
7	351,Presensi >> Kehadiran,54160,9/9/2014 3:38:24 PM,12,MAHASISWA
8	,Registrasi >> Pengaturan >> Jadwal Registrasi,440,9/9/2014 3:38:30 PM,33,ADMIN REGISTRASI FAKULTAS
9	354,dashboard,48349,9/9/2014 3:38:43 PM,20,PEGAWAI
10	319,Presensi >> Cetak Daftar Hadir,82574,9/9/2014 3:38:46 PM,3,DOSEN

Gambar. 2 Event Log iGracias Universitas Telkom

Contoh gambar 2.4 merupakan salah satu contoh *event log* yang tidak memiliki atribut *cost* ataupun *resource*. Pada *event log* tersebut, memuat catatan history aktivitas yang dilakukan pengguna iGracias. Atributnya pun terdiri dari *case_id*, *activity*, *eksekutor*, *waktu*, *group_eksekutor*, dan *usergroupname*. Banyak juga catatan history seperti *event log* dicatat dalam bentuk *server log*. Namun informasi yang terkandung dalam *server log* tidak jauh berbeda dengan format pada *event log*. Pada *server log* mau pun *event log* memiliki *authuser field*, *waktu* dan *page* yang dituju *user*.

E. Graph Mining Frequent Sequential Access Pattern

Metode *graph* adalah salah satu metode yang akan digunakan pada publikasi ini untuk menentukan seberapa besar frekuensi pengaksesan suatu *web page*. Dalam sebuah *graph* terdiri dari dua komponen yaitu, *node* dan *edge*. *Node* disini mendeskripsikan sebuah *web page* sedangkan *edge* merupakan tautan antar halaman web yang diakses oleh *user*[1]. Semakin sering suatu web diakses maka *edge* yang mengarah pada *node* tersebut semakin banyak dengan kata lain bobot akses terhadap web page tersebut semakin besar.

Graph mining frequent sequential access patterns merupakan suatu proses *mining* untuk mencari tahu frekuensi sebuah *pattern* pada saat mengakses sebuah website. Metode ini menggambarkan sebuah *subsequence* atau kumpulan item pada suatu dataset yang terjadi secara berulang dengan frekuensi tertentu. Sebagai contoh, *user A* mengakses *web page X* lalu *web page Y* dan yang terakhir *web page Z*. Jika frekuensi ini sering terjadi dan tersimpan pada *web server log* sebuah website maka contoh tersebut dapat kita katakan sebagai *frequent sequential pattern*. Pada publikasi ini akan dilakukan proses *web usage mining* berdasarkan *event log* untuk mencari *frequent sequential pattern* dengan metode *graph*[1].

1) Construct Web Usage Graph

Sebuah *graph* akan dibangun berdasarkan daftar dari sebuah akses halaman *web* dengan sesi yang berbeda. *Graph* tersebut terdiri dari *vertices (nodes)* dan *edges (links)* dimana *nodes* merepresentasikan sebuah *web pages* dan *edges* merepresentasikan akses sekuensial antar *page*. Jumlah *nodes* pada *graph* sama dengan jumlah *web pages* yang berbeda-beda yang diakses selama semua sesi berlangsung. Setiap *node* pada *graph* mengandung bobot sebagai nilai *node count* yang menyimpan jumlah kejadian tertentu suatu *pages* dengan sesi dan *node id* berbeda. Setiap *edge* pada *graph* mengandung bobot sebagai nilai *link count* yang merepresentasikan frekuensi dari *edge*, *edge id* dan daftar *session id*. Daftar sesi bersangkutan dengan *path* atau *edge*[1].

2) Pruning Web Usage Graph

Pada tahap ini akan dilakukan penghapusan semua *node* yang memiliki *node count* kurang dari *minimum support threshold* dan kemudian menghubungkan kembali *edges*-nya. Pada saat pemangkasan *graph*, jika ada *edge* yang terdapat diantara dua *frequent node* (*node* yang memiliki *node count* lebih dari *minimum support threshold*), maka *edge* tersebut akan tetap utuh. Namun jika ada *edge* yang terdapat diantara dua *infrequent node* (*node* yang memiliki *node count* kurang dari *minimum support threshold*), maka *edge* tersebut dihilangkan. Semua *infrequent node* dihapus dan menghubungkan kembali *frequent node* dengan *edge* yang berasal dari *sequence infrequent node* awal yang tersambung. Kemudian diperiksa kembali, jika masih ada *edge* yang memiliki *link count* kurang dari *minimum support threshold* pada pemangkasan *web usage graph*

maka *edge* tersebut akan dihilangkan. Tahap ini menghasilkan pemangkasan *graph* yang hanya memiliki *frequent nodes* dan *edges*[1].

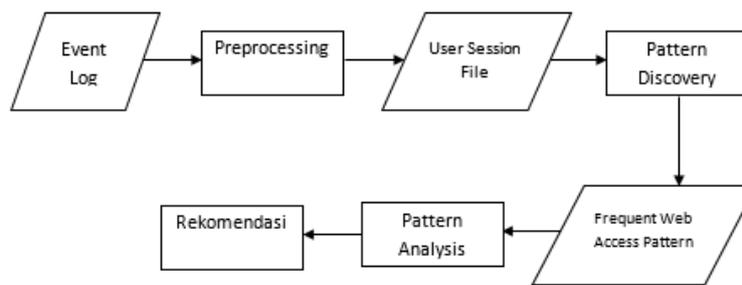
3) Mining Frequent Sequential Access Pattern

Proses *mining* ini dimulai dengan pencarian urutan terbesar dari *nodes*, sebagai contoh *path* terpanjang dengan *length* maksimum. Dimana *length* adalah jumlah *nodes* yang terlibat dalam urutan. Dengan melakukan *traversing node* dari *graph* hasil *pruning*, kita akan mendapatkan frekuensi urutan setiap *node* dimana hal itu merepresentasikan suatu pola akses web. Frekuensi urutan tersebut akan menjadi *minimum link count* dari seluruh *edges* yang terlibat. Dengan cara tersebut, lintasi semua *path* yang ada pada *graph* hasil *pruning* dan daftarkan semua pola frekuensi bersama dengan frekuensinya lalu susun berdasarkan *length*-nya. Frekuensi urutan dengan *length* 1 merupakan semua *nodes* pada *graph* hasil *pruning* dan *node count* akan mewakili frekuensinya[1].

III. RESEARCH METHOD

A. Gambaran Umum Sistem

Pada publikasi ini perencanaan sistem yang akan dibuat seperti tercantum pada gambar 3 berikut:



Gambar. 3. Rancangan Sistem

Penjelasan mengenai gambaran umum yang terdapat pada gambar 3, aktifitas yang dilakukan adalah sebagai berikut:

1. Dataset berupa event log yang didapatkan dari Igracias Universitas Telkom dari September 2014 sampai Juni 2015.
2. Dilakukan tahap *preprocessing* pada dataset yang ada sehingga dapat digunakan sesuai dengan kebutuhan system.
3. Dilakukan tahap *pattern discovery* untuk mendapatkan *frequent web access pattern* dengan menggunakan metode *graph mining frequent sequential access patterns*.
4. Dilakukan tahap *pattern analysis* untuk memeriksa pola dengan melakukan kuisisioner terhadap pengguna. Tahap ini juga dilakukan pengujian terlebih dahulu untuk mendapatkan *minimum support* yang baik agar menghasilkan pola yang sesuai dengan hasil evaluasi melalui kuisisioner.

Rekomendasi merupakan tahap pemberian rekomendasi kepada Sistem Informasi dari hasil keluaran system yang sudah dilakukan evaluasi kuisisioner.

IV. RESULTS AND DISCUSSION

Berikut adalah hasil pengujian yang dilakukan sesuai dengan skenario yang telah ditentukan dan disertai dengan analisis hasil pengujian.

1) Pengujian dengan menggunakan beberapa *minimum support*

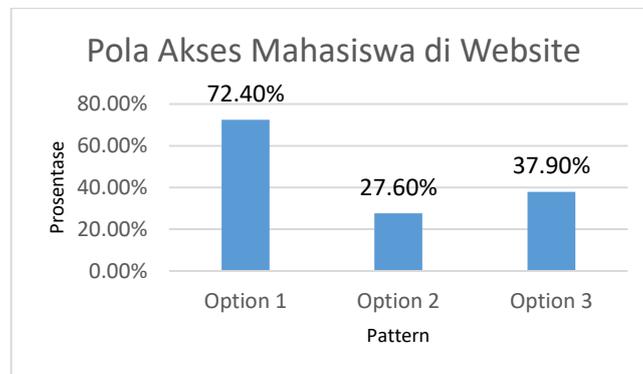
Pengujian terhadap beberapa *minimum support* bertujuan untuk mendapatkan *minimum support* yang cukup. *Minimum support* digunakan pada tahap *pruning* yang bertujuan untuk menghasilkan node yang kurang sering diakses agar terlihat pola pengaksesan yang sering dilakukan. Dari pola tersebut, akan dapat dilihat *behavior* dari pengguna.

Pada pengujian ini akan digunakan tiga *minimum support*, yaitu 40%, 50% dan 60% untuk data mahasiswa dan dosen. Sedangkan data pegawai akan digunakan *minimum support* 30%, 40% dan 50%. Data pegawai menggunakan *minimum support* yang berbeda dikarenakan jumlah data pegawai jauh lebih sedikit daripada data mahasiswa dan dosen. *Minimum support* yang digunakan sebaiknya tidak terlalu tinggi dan juga tidak terlalu rendah. Untuk itu, *minimum support* yang dipakai berada diantara tengah tengah.

2) Pengujian dengan membandingkan hasil *web access pattern* dengan hasil kuesioner

Selain melakukan uji coba pada sistem, dilakukan juga evaluasi dengan kuesioner kepada *user*. Dimana akan diketahui kondisi sesungguhnya pada hasil kuesioner tersebut. Dan dari hasil kuesioner akan dibandingkan dengan hasil dari sistem. Sehingga akan terlihat perbandingan antara sistem dengan hasil kuesioner nya.

Kuesioner ini dilakukan dengan teknik *sampling area* yang dikutip dari Uma Sekaran et al. bahwa teknik ini menggunakan sampel antara 5 sampai 20 responden. Igracias merupakan sumber informasi yang digunakan luas diseluruh kawasan Universitas Telkom. Untuk ini pada publikasi ini digunakan teknik *sampling*. Teknik *sampling* digunakan pada tiap-tiap group user, yakni pada group mahasiswa, dosen dan karyawan.

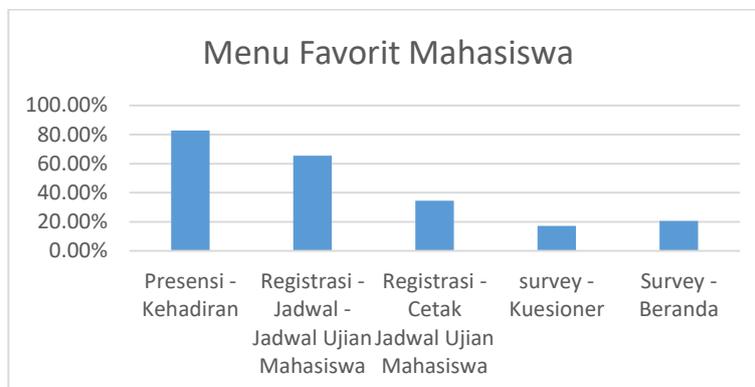


Gambar. 4. Hasil Kuesioner Mahasiswa

Berikut merupakan keterangan dari gambar 4 :

- Option 1 : Presensi - Kehadiran >> Registrasi - Jadwal - Jadwal Ujian Mahasiswa >> Survey - Kuesioner >> Survey - Beranda >> Registrasi - Cetak Jadwal Ujian Mahasiswa
- Option 2 : Presensi - Kehadiran >> Survey - Kuesioner >> Survey - Beranda >> Registrasi - Cetak Jadwal Ujian Mahasiswa

- Option 3 : Registrasi - Jadwal - Jadwal Ujian Mahasiswa >> Registrasi - Cetak Jadwal Ujian Mahasiswa



Gambar. 5. Hasil Kuesioner Menu Favorit Mahasiswa

Pada proses pengambilan data melalui kuesioner pada mahasiswa, terdapat 29 responden. Pilihan pada kuesioner tersebut diambil dari frekuensi terbanyak pada data hasil web access patterns.

Begitu juga dilakukan pada group dosen dan karyawan hasil akhir akan ditampilkan pada tahap selanjutnya.

3) Analisis

Tabel berikut merupakan hasil behavior pola navigasi setiap pengguna, yang akan diambil hanya lima data yang sering dilakukan oleh user.

Tabel 1 Hasil Web Access Pattern Data Mahasiswa

No.	Pattern
1	login >> dashboard >> Presensi - Kehadiran >> Registrasi - Jadwal - Jadwal Ujian Mahasiswa >> Survey - Kuesioner >> Survey - Beranda >> Registrasi - Cetak Jadwal Ujian Mahasiswa
2	login >> dashboard >> Presensi - Kehadiran >> Registrasi - Jadwal - Jadwal Ujian Mahasiswa >> Survey - Kuesioner >> Survey - Beranda
3	login >> dashboard >> Presensi - Kehadiran >> Registrasi - Jadwal - Jadwal Ujian Mahasiswa >> Registrasi - Cetak Jadwal Ujian Mahasiswa
4	login >> dashboard >> Registrasi - Jadwal - Jadwal Ujian Mahasiswa >> Registrasi - Cetak Jadwal Ujian Mahasiswa
5	login >> dashboard >> Presensi - Kehadiran

Tabel 2 Hasil Web Access Pattern Data Dosen

No.	Pattern
1	login >> dashboard >> Presensi - Jadwal Dosen >> Presensi - Cetak Daftar Hadir >> Presensi - Cetak Presensi >> Presensi - Laporan Perkuliahan >> Presensi - Berita Acara Perkuliahan >> Presensi - Submit BAP >> Presensi - Berita Acara Perkuliahan
2	login >> dashboard >> Presensi - Cetak Presensi >> Presensi - Laporan Perkuliahan >> Presensi - Berita Acara Perkuliahan >> Presensi - Submit BAP >> Presensi - Berita Acara Perkuliahan
3	login >> dashboard >> Presensi - Cetak Presensi >> Presensi - Laporan Perkuliahan >> Presensi - Berita Acara Perkuliahan >> Presensi - Submit BAP
4	login >> dashboard >> Presensi - Cetak Presensi >> Presensi - Laporan Perkuliahan >>

	Presensi - Berita Acara Perkuliahan
5	login >> dashboard >> Presensi - Cetak Presensi >> Presensi - Laporan Perkuliahan

Tabel 3 Hasil Web Access Pattern Data Pegawai

No.	Pattern
1	dashboard >> Manajemen SDM - Pegawai - Data Pegawai >> Manajemen SDM - Pegawai - Lihat Kehadiran >> Manajemen SDM - Penggajian - Detail Penggajian >> Manajemen SDM - Riwayat Pelatihan Pegawai >> Manajemen SDM - Penggajian - Tunjangan Pegawai >> Manajemen SDM - Cuti - Pengajuan Cuti
2	dashboard >> Manajemen SDM - Pegawai - Data Pegawai >> Manajemen SDM - Pegawai - Lihat Kehadiran >> Manajemen SDM - Cuti - Pengajuan Cuti >> Manajemen SDM - Riwayat Pelatihan Pegawai
3	dashboard >> Manajemen SDM - Penggajian - Tunjangan Pegawai >> Manajemen SDM - Riwayat Pelatihan Pegawai >> Manajemen SDM - Cuti - Pengajuan Cuti
4	dashboard >> Manajemen SDM - Penggajian - Tunjangan Pegawai >> Manajemen SDM - Penggajian - Detail Penggajian
5	Manajemen SDM - Pegawai - Data Pegawai >> Manajemen SDM - Pegawai - Lihat Kehadiran

Tanda “>>” mendandakan proses yang dilakukan selanjutnya. Urutan penyebutan dari satu sampai lima menandakan pola tersebut paling sering dilakukan oleh group yang bersangkutan.

Pada table 1,2, dan 3 terlihat pola-pola yang pengaksesan menu yang berbeda untuk setiap group user (mahasiswa, dosen, dan pegawai). Pola mahasiswa lebih banyak di penjadwalan, registrasi, dan jadwal ujian. Sedangkan pola pada dosen lebih banyak di presensi, submit bap, dan laporan perkuliahan. Sedangkan pada data pegawai lebih banyak di lihat kehadiran, detail penggajian, pengajuan cuti.

V. Conclusion

Hasil percobaan *minimum support* yang cukup untuk mendapatkan behavior pola pengaksesan setiap data berbeda. Setelah dilakukan uji coba dengan menggunakan *minimum support* yang kecil akan menghasilkan *sequence patterns* yang panjang. Sedangkan menggunakan *minimum support* yang besar hasil *sequence patterns* akan lebih pendek. Untuk itu, dipilih *minimum support* yang sekiranya dapat mewakili pola pengaksesan yang sering dilakukan oleh *user*. Hasil percobaan pada data dosen, mahasiswa dan pegawai memiliki *minimum support* yang sama. Hal itu dikarenakan perbedaan dimana jumlah data pegawai lebih sedikit dibandingkan dengan jumlah data dosen dan mahasiswa.

Metoda berbasis *graph* berhasil mendapatkan pola-pola pengaksesan group user, hal ini dibuktikan dari perbedaan yang mencolok untuk masing-masing group user. Pola dosen sangat berbeda dengan pola mahasiswa, dan pola dosen juga memiliki pola berbeda dengan pola pegawai. Dan diperkuat hasil sistem dengan kuesioner yang diberikan ke perwakilan group didapatkan pola yang sama.

REFERENCES

- [1] D. K. Singh, “Graph based Approach for Mining Frequent Sequential Access Patterns of Web pages,” vol. 40, no. 10, pp. 33–37, 2012.
- [2] K. K. Soni, J. Agrawal, S. Sharma, and S. Agrawal, “Association Rule Mining Based on Density and Regional Minimum Support,” *2015 Fifth Int. Conf. Commun. Syst. Netw. Technol.*, pp. 968–975, 2015.
- [3] S. Vijayarani and M. E. Suganya, “Research issues in web mining,” *Int. J. Comput. Technol.*, vol. 2, no. 3, pp. 55–64, 2015.
- [4] J. Punin, M. Krishnamoorthy, and M. Zaki, “Web usage mining—Languages and algorithms,” *Explor. Data Anal. Empir. Res.*, pp. 266–281, 2003.

- [5] M. Heydari, R. Alsaqour, K. Imran, and K. Vaziry, "A Weighted Graph Web Usage Mining Method to Evaluate Usage of Websites Department of Computer Science , Faculty of Information Science and Technology , University," vol. 5, no. 9, pp. 1606–1616, 2011.
- [6] M. Valera and U. Chauhan, "An efficient web recommender system based on approach of mining frequent sequential pattern from customized web log preprocessing," *2013 4th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2013*, 2013.
- [7] P. Sharma, S. Yadav, and B. Bohra, "A review study of server log formats for efficient web mining," *Proc. 2015 Int. Conf. Green Comput. Internet Things, ICGCIoT 2015*, pp. 1373–1377, 2016.
- [8] J. Tian, S. Rudraraju, Z. Li, S. Member, and I. C. Society, "on Workload and Failure Data Extracted from Server Logs," *October*, vol. 30, no. 11, pp. 754–769, 2004.
- [9] A. A. Hermawan, "ANALISIS KONTEKS PROSES BISNIS BERDASARKAN ' EVENT LOG ' BUSINESS PROCESS CONTEXT ANALYSIS BASED ON ' EVENT LOG ,' " vol. 4, no. 3, pp. 133–150, 2014.
- [10] T. Pustaka, M. Pendugaan, K. Pada, and P. Penggunaan, "2. Tinjauan Pustaka 2.1," no. Aminudin 2008, pp. 1–14, 2009.

