IMPLEMENTASI OLAP UNTUK ANALISIS LOG PADA PROSES BUSINESS INTELLIGENCE SEBAGAI DATA PENDUKUNG DECISION SUPPORT SYSTEM

(Studi kasus Institut Teknologi Sumatera)

(Tesis)

Oleh
HERU RUWANDAR

PROGRAM PASCASARJANA MAGISTER TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022

IMPLEMENTASI OLAP UNTUK ANALISIS LOG PADA PROSES BUSINESS INTELLIGENCE SEBAGAI DATA PENDUKUNG DECISION SUPPORT SYSTEM

(Studi kasus Institut Teknologi Sumatera)

Oleh
HERU RUWANDAR

Tesis

Sebagai Salah Satu Syarat untuk Mencapai Gelar
MAGISTER TEKNIK ELEKTRO

Pada

Program Pascasarjana Magister Teknik Elektro
Fakultas Teknik Universitas Lampung

PROGRAM PASCASARJANA MAGISTER TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022

ABSTRACT

IMPLEMENTATION OF OLAP FOR LOG ANALYSIS IN BUSINESS INTELLEGENCE PROCESS AS DATA SUPPORTING DECISION SUPPORT SYSTEM

By

HERU RUWANDAR

Nowaday, internet is the most popular media to obtain information and can be used anywhere with a variety of used devices. Internet packet traffic can be controlled and logged by the firewall. This log of internet traffic is called the internet access log and generally stored in a log file. The log data can be used for internet usage analysis needs. Log data records internet traffic at any time and is stored in a log file, so that the data content becomes very large and the size becomes large. The analysis requires a tool that can accommodate big data processing, namely OLAP (On-Line Analytical Processing). This study aims to obtain new information related to internet usage using the OLAP method to analyze access logs on the ITERA internet network. Log data is processed using Python, starting from the ETL process (selecting, cleaning and shorting) to visualization. This study has been completed in obtaining new information related to internet use, including: 1) the highest traffic time is between 7 AM to 5 PM West Indonesian Time where that time is working hours in ITERA, 2) social media are the most accessed applications by internet users namely Facebook, TikTok, WhatsApp, Instagram, YouTube, and Telegram, and 3) the most access location in ITERA is GKU by which up to 33%. The result can be used as supporting data for making decision or policies related to the internet and supporting internet infrastructure at ITERA.

Keywords: Log file, OLTP, OLAP, firewall, python

ABSTRAK

IMPLEMENTASI OLAP UNTUK ANALISIS LOG PADA PROSES BUSINESS INTELLIGENCE SEBAGAI DATA PENDUKUNG DECISION SUPPORT SYSTEM

Oleh

HERU RUWANDAR

Internet merupakan media untuk mencari informasi yang paling popular saat ini dan dapat dipakai dimanapun dengan bermacam-macam perangkat yang digunakan. Lalu lintas paket internet dapat dikontrol dan dicatat oleh *firewall*. Pencatatan lalu lintas internet inilah yang disebut *log* akses internet dan umumnya disimpan dalam *log file*. Data *log* tersebut dapat digunakan untuk kebutuhan analisa penggunaan internet. Data *log* mencatat lalu lintas internet setiap saat dan disimpan ke dalam *log file*, sehingga isi data menjadi sangat banyak dan ukuran menjadi besar. Untuk analisa memerlukan alat yang dapat mengakomodir pemrosesan data besar yaitu OLAP (*On-Line Analytical Processing*). Penelitian ini mempunyai tujuan untuk mendapatkan informasi baru terkait penggunaan internet menggunakan metode OLAP untuk menganalisa *log* akses pada jaringan internet ITERA. Data *log* diproses menggunakan Python, mulai dari proses ETL (*selecting, cleaning and shorting*) sampai dengan visualisasi. Penelitian ini berhasil mendapatkan informasi baru terkait penggunaan internet antara lain: 1) waktu trafik tertinggi adalah di antara pukul 07.00 WIB sampai dengan pukul 17.00 WIB di mana waktu tersebut adalah jam kerja di lingkungan ITERA, 2) aplikasi sosial media menjadi aplikasi yang paling banyak di akses oleh pengguna internet yaitu Facebook, TikTok, WhatsApp, Instagram, YouTube, dan Telegram, dan 3) lokasi akses terbanyak di ITERA adalah GKU yang penggunaanya sampai 33%. Informasi ini dapat dijadikan data pendukung untuk pengambilan keputusan atau kebijakan terkait internet dan infrastruktur pendukung internet di ITERA.

Kata kunci: *Log file*, OLTP, OLAP, *firewall*, Python

Judul Tesis	:	**IMPLEMENTASI OLAP UNTUK ANALISIS LOG PADA PROSES BUSINESS INTELLIGENCE SEBAGAI DATA PENDUKUNG DECISION SUPPORT SYSTEM**
Nama Mahasiswa	:	Heru Ruwandar
Nomor Pokok Mahasiswa	:	1825031007
Progam Studi	:	Magister Teknik Elektro
Fakultas	:	Teknik

MENYETUJUI

1. Komisi Pembimbing

Pembimbing I	Pembimbing II
Dr. Ing. Ardian Ulvan, S.T., M.Sc.	**Dr. Eng. Mardiana, S.T., M.T.**
NIP. 19731128 199903 1 005	NIP. 19720316 199903 2 002

2. Ketua Progam Studi Magister Teknik Elektro

Misfa Susanto, S.T., M.T., Ph.D.
NIP. 19710525 199903 1 001

MENGESAHKAN

1. Komisi Penguji

Ketua Komisi Penguji (Pembimbing I)	: **Dr. Ing. Ardian Ulvan, S.T., M.Sc.** .....
Sekretaris Komisi Penguji (Pembimbing II)	: **Dr. Eng. Mardiana, S.T., M.T.** .....
Anggota Komisi Penguji (Penguji I)	: **Misfa Susanto, S.T., M.T., Ph.D.** .....
Anggota Komisi Penguji (Penguji II)	: **Dr. Ing. Melvi, S.T., M.T.** .....

2. Dekan Fakultas Teknik

Dr. Eng. Helmy Fitriawan,S.T., M.Sc.
NIP 19750928 200112 1 002

3. Direktur Progam Pascasarjana

Prof. Dr. Ir. Ahmad Saudi Samosir, S.T., M.T.
NIP 19710415 199803 1 005

Tanggal Lulus Ujian Tesis: **2 Agustus 2022**

SURAT PERNYATAAN

Dengan ini Saya menyatakan bahwa sesungguhnya tesis yang saya buat sebagai syarat untuk mendapatkan gelar Magister Teknik pada Progam Pascasarjana Magister Teknik Elektro seluruhnya adalah benar merupakan hasil karya sendiri.

Adapun dalam bagian-bagian tertentu dalam penulisan tesis ini, saya kutip dari hasil penulisan orang lain yang sumbernya dituliskan dengan jelas sesuai dengan norma, kaidah dan etika penulisan karya ilmiah.

Tesis saya dengan judul "Implementasi OLAP Untuk Analisis Log Pada Proses *Business Intelligence* Sebagai Data Pendukung *Decision Support System*" dapat saya selesaikan berkat bimbingan dan motivasi dari pembimbing saya, yaitu:

1. Dr. Ing. Ardian Ulvan, S.T., M.Sc.
2. Dr. Eng. Mardiana, S.T., M.T.

Saya ucapkan terima kasih yang sebesar-besarnya kepada semua pihak, khususnya dosen pembimbing dan Bapak/ Ibu Dosen Progam Studi Magister Teknik Elektro Universitas Lampung yang telah banyak memberikan ilmu pengetahuan, bimbingan dan motivasi.

Apabila di kemudian hari ditemukan seluruh atau sebagian tesis yang saya buat ini bukan hasil karya saya sendiri atau adanya plagiat dalam bagian-bagian tertentu, saya bersedia menerima sanksi akademik sesuai dengan peraturan perundangan yang berlaku.

Bandar Lampung, 2 Agustus 2022

Heru Ruwandar
NPM: 1825031007

RIWAYAT HIDUP

Penulis adalah anak pertama dari tiga bersaudara yang dilahirkan di Tanjung Fajar, pada tanggal 24 Oktober 1987 dari pasangan Bapak Rusin dan Ibu Kasmirah. Pendidikan formal penulis dimulai dari Sekolah Dasar Negeri (SDN) 2 Tanjung Kemala Kec. Pubian yang diselesaikan pada tahun 1999. Kemudian pendidikan dilanjutkan di Madarasah Tsanawiyah (MTs) Nurul Ulum Payung Rejo Kec. Pubian yang diselesaikan pada tahun 2002. Selanjutnya pendidikan pada Sekolah Menengah Kejuruan Negeri (SMKN) 1 Gading Rejo Tanggamus (sekarang Pringsewu) yang diselesaikan pada tahun 2005.

Tahun 2006, penulis terdaftar sebagai mahasiswa jurusan Teknik Elektro Fakultas Teknik Universitas Lampung melalui jalur non Seleksi Penerimaan Mahasiswa Baru (non-SPMB). Penulis memilih Konsentrasi Sistem Komputer dan Informatika sebagai fokus perkuliahan dan penelitian. Selama masa kuliah penulis juga aktif dalam organisasi intra kampus yaitu Himpunan Mahasiswa Teknik Elektro (HIMATRO) sebagai Anggota Departemen Kerohanian pada tahun 2008-2009.

Penulis melaksanakan Kerja Praktek di Bulletin Board Sevice (BBS) Unilanet UPT Pusat Komputer Universitas Lampung pada tanggal 1 Juli sampai dengan 30 September 2010. Laporan kerja praktik membahas tentang penggunaan radius server pada server hotspot di UPT Puskom Unila. Terhitung sejak 1 Oktober 2010

sampai dengan 30 September 2013, penulis magang pada BBS Unilanet UPT Puskom (sekarang UPT TIK) sebagai Network dan System Administrator.

Pada tahun 2018, penulis terdaftar sebagai Mahasiswa Program Pascasarjana Teknik Elektro Universitas Lampung. Selain sebagai mahasiswa, penulis berkerja sebagai Network dan System Administrator di UPT TIK Institut Teknologi Sumatera sejak 20 Oktober 2015. Sejak tahun 2021, penulis melakukan penelitian dengan judul tesis "Implementasi Olap Untuk Analisis Log Pada Proses Business Intelligence Sebagai Data Pendukung Decision Support System" dibawah bimbingan Bapak Dr. Ing. Ardian Ulvan, S.T., M.Sc. dan Ibu Dr. Eng. Mardiana, S.T., M.T.

Bandar Lampung, 2 Agustus 2022
Penulis

HERU RUWANDAR

Sebuah karya kecil ini, kupersembahkan untuk:

Bapak dan Ibu

Rusin dan Kasmirah

Istri Tercinta

Dewi Suhika, M.Sc.

Anak-anakku tersayang:

Aydan Faiz Ruwandar

Abhivandya Ayres Ruwandar

Aiza Kamila Zhara Ruwandar

yang selalu mendoakanku, menyayangiku, semua perhatian dan pemberian yang tiada henti.

MOTTO

Ilmu itu lebih berharga daripada emas

Man jadda wa jadda
(barang siapa yang bersungguh-sungguh, pasti ia akan mendapat yang diinginkan)

SEORANG OPTIMIS MELIHAT KESEMPATAN DALAM SETIAP MALAPETAKA, SEDANG SEORANG PESIMIS MELIHAT MALAPETAKA DALAM SETIAP KESEMPATAN . (ANUM)

SAN WACANA

Assalamu'alaikum Wr. Wb.

Syukur Alhamdulillahi robbil'alamin, Penulis haturkan puji syukur atas kehadirat Allah SWT yang senantiasa melimpahkan rahmat dan hidayah, serta inayah-Nya kepada penulis sehingga penulis dapat menyelesaikan laporan Tesis dengan mempersembahkan judul tesis "Implementasi Olap Untuk Analisis Log Pada Proses Business Intelligence Sebagai Data Pendukung Decision Support System" dengan sebaik-baiknya. Shalawat beriring salam selalu tercurah kepada junjungan seluruh alam Nabi Muhammad SAW, sahabatnya, serta para pengikutnya yang selalu istiqomah diatas jalan agama islam hingga hari ajal menjemput.

Dalam penyusunan tugas akhir ini penulis banyak mendapat bimbingan, motivasi dan bantuan baik moral maupun materi oleh banyak pihak. Untuk itu dengan sepenuh ketulusan hati penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. Karomani, M. Si, Selaku Rektor Universitas Lampung.

2. Bapak Prof. Dr. Ir. Ahmad Saudi Samosir, S.T., M.T. Selaku Direktur Program Pascasarjana Fakultas Teknik Universitas Lampung.

3. Bapak Dr. Eng. Helmy Fitriawan, S.T., M.Sc. selaku Dekan Fakultas Teknik, Universitas Lampung.

4. Ibu Herlinawati, S.T., M.T. selaku Ketua Jurusan Teknik Elektro Universitas Lampung.

5. Bapak Misfa Susanto, S.T., M.T., Ph.D. selaku Ketua Program Studi Magister Teknik Elektro Universitas Lampung dan juga selaku penguji utama yang telah banyak memberikan masukan, kritik dan saran yang bermanfaat bagi penulis.

6. Bapak Dr. Ing. Ardian Ulvan, S.T., M.Sc. selaku dosen pembimbing utama tesis dan juga dosen pembimbing akademik, yang banyak memberikan waktu, ide pemikiran dan semangat serta motivasi bagi penulis.

7. Ibu Dr. Eng. Mardiana, S.T., M.T. selaku pembimbing kedua tesis, yang telah banyak memberikan waktu, pengalaman, motivasi dan pemikiran bagi penulis.

8. Ibu Dr. Ing. Melvi, S.T., M.T. selaku dosen penguji kedua yang telah banyak memberikan kritik, saran dan motivasi yang bermanfaat bagi penulis.

9. Seluruh Dosen Program Studi Magister Teknik Elektro Universitas Lampung, berkat ilmu yang telah diajarkan kepada penulis selama penulis menjalani masa studi di perkuliahan.

10. Seluruh Tenaga Kependidikan Program Studi Magister Teknik Elektro Fakultas Teknik Universitas Lampung yang telah banyak membantu penulis dalam administrasi, sehingga dapat menyelesaikan tugas akhir ini.

11. Teman-teman Program Studi Magister Teknik Elektro Unila angkatan 2018 (Pak Dwi, Elka, Nambi, dan Reza) untuk kebersamaan yang telah dijalani, berjuang sampai lulus semua. Terimakasih banyak atas dukungan dan semangat dari kalian. Semoga Allah memberikan balasan yang terbaik untuk semuanya.

12. Istri tercinta Dewi Suhika, M.Sc. yang selaku memberi semangat yang tiada henti. Dan anak-anakku tersayang (Aydan Ayres Aiza), penyemangat setiap saat.
13. Serta semua pihak yang telah membantu dalam penyusunan tesis ini. Terimakasih banyak.

Penulis menyadari bahwa Tesis ini masih jauh dari kesempurnaan, baik dari segi isi maupun cara penyajiannya. Oleh karena itu, Penulis sangat mengharapkan saran serta kritik yang bersifat membangun dari pembaca. Akhir kata sedikit harapan penulis semoga karya sederhana ini dapat berguna dan bermanfaat bagi kita semua. Aamiin Allahumma Aamiin.

Wassalamu'alaikum Wr. Wb.

Bandar Lampung, 2 Agustus 2022

Penulis,

Heru Ruwandar

DAFTAR ISI

DAFTAR GAMBAR

Gambar **Halaman**

DAFTAR TABEL

Tabel **Halaman**

1. PENDAHULUAN

1.1. Latar Belakang

Kebutuhan akan informasi yang mudah dan cepat mengharuskan seseorang untuk menggunakan internet sebagai media untuk mendapatkan informasi tersebut. Seiring berjalannya waktu, internet menjadi semakin popular digunakan di kalangan masyarakat di Indonesia bahkan dunia. Saat ini internet sudah menjadi salah satu kebutuhan penting dalam mencari informasi. Bahkan beberapa tahun belakangan ini, internet sudah menjadi sarana untuk bertransaksi jual beli menggunakan bermacam-macam aplikasi, mulai dari aplikasi *web, desktop*, sampai aplikasi pada ponsel pengguna (*mobile*).

Lalu lintas penggunaan internet diatur oleh *router* untuk mengatur jaringan, dan *Firewall* untuk mengatur keamanan lalu lintas paket internet. *Router* adalah perangkat yang akan melewatkan paket *Internet Protocol* (IP) dari suatu jaringan ke jaringan yang lain menggunakan metode pengalamatan dan protokol tertentu untuk melewatkan paket data tersebut [1]. *Firewall* didefinisikan sebagai sebuah komponen atau kumpulan komponen yang membatasi akses antara sebuah jaringan yang diproteksi dengan internet, atau antara kumpulan-kumpulan jaringan lainnya [2]. *Firewall* umumnya digunakan pada bagian terluar pada suatu jaringan. *Firewall* adalah satu titik antara dua atau lebih jaringan di mana semua lalu lintas harus

melewatinya; lalu lintas dapat dikontrol oleh *Firewall* dan dapat diautentikasi melalui perangkat, dan semua lalu lintas dicatat [3]. Pencatatan lalu lintas internet inilah yang disebut *log* akses internet.

Data *log* lalu lintas dari para pengguna umumnya disimpan dalam *log file*. Data *log file* berbeda bentuk dan format tergantung dari perangkat yang menghasilkan *log* tersebut. Untuk *Firewall*, *log* berisi data lalu lintas dan transaksi para pengguna internet. Data *log* tersebut dapat digunakan untuk kebutuhan analisa di masa mendatang. Analisa tersebut dapat menggunakan metode OLTP *(Online Transaction Processing*). OLTP merupakan suatu proses yang mendukung operasi bisnis sehari-hari.

Data *log* mencatat lalu lintas internet setiap saat dan disimpan ke dalam *log file*, sehingga isi data menjadi sangat banyak dan ukuran menjadi besar. Untuk analisa memerlukan alat yang dapat mengakomodir pemrosesan data besar. OLAP (*On-Line Analytical Processing*) dapat digunakan untuk melakukan *query* yang tidak mengubah data, pengindeksan berat untuk meningkatkan kinerja *query*, denormalisasi basis data untuk memenuhi persyaratan *query* umum dan meningkatkan waktu respons *query* [4]. OLAP sangat erat kaitannya dengan *Decision Support System* (DSS) atau Sistem Pendukung Keputusan. OLAP juga merupakan bagian dari tekno*log*i di belakang *Business Intellegence* (BI).

Penggunaan internet di lingkungan Institut Tekno*log*i Sumatera (ITERA) terus mengalami peningkatan seiring bertambahnya civitas akademika. Layanan internet di ITERA dibangun dan dikelola oleh UPT TIK (Unit Pelaksana Teknis Teknologi Informasi dan Komunikasi), sehingga pengguna dapat mengakses internet dengan mudah. Untuk dapat mengakses internet, para pengguna harus menggunakan akun

yang sudah didaftarkan. Aktifitas penggunaan internet di lingkungan ITERA semua terdokumentasi dalam data *log* yang ada di *Firewall* yang ada di UPT TIK ITERA. *Log* transaksi muncul dari aktivitas yang dicatat ketika orang berinteraksi dengan sistem dan layanan komputer [5]. Data *log* selain berisi aktifitas penggunaan internet pengguna juga berisi data identitas pengguna seperti alamat IP, *Medium Access Control (MAC) Address, Uniform Resource Locator (URL)* tujuan, dan waktu akses.

Sampai saat ini, data *log* tersebut belum dimanfaatkan lebih lanjut untuk mendapatkan informasi baru terkait penggunaan internet atau sebagai alat pendukung keputusan. Pengambilan keputusan yang terkait penggunaan internet di ITERA, saat ini masih berdasar kepada pemakaian *bandwidth* dan banyaknya pengguna yang *login* pada saat yang bersamaan (*concurrent*). Hal ini dirasa kurang cukup untuk menentukan arah pengambil keputusan guna mengambil kebijakan yang tepat.

Dalam penelitian ini, data *log* diambil dari *Firewall* yang merekam seluruh aktifitas penggunaan internet. Data *log* diambil mulai dari bulan Mei sampai dengan Juli 2021. Data *log* ini kemudian diproses menggunakan proses BI. Data *log* selanjutnya dianalisa sehingga mendapatkan informasi baru yang dapat digunakan sebagai data system pengambil keputusan/DSS ITERA.

1.2. Rumusan Masalah

Berdasarkan uraian tersebut masalah yang dijadikan penelitian dirumuskan sebagai berikut:

a. Bagaimana mengolah data transaksi secara *online* (OLTP) sehingga menghasilkan sebuah Sys*log*,
b. Bagaimana menganalisa informasi secara *online* dari data transaksi (OLAP) sehingga mendapatkan sebuah *data warehouse* (DW),
c. Bagaimana mengolah DW untuk mendapatkan berbagai informasi baru dari pengguna internet ITERA,
d. Bagaimana mengolah informasi-informasi baru dari pengguna internet menjadi data tambahan untuk pendukung system pengambil keputusan (DSS),

1.3. Tujuan Penelitian

Adapun tujuan dari penelitian yang dilakukan adalah:

a. Menganalisa data transaksi secara *online* menggunakan OLTP pada periode tertentu,
b. Menganalisa informasi secara *online* dari data transaksi (OLAP) sehingga mendapatkan sebuah DW,
c. Memperoleh data warehouse untuk mendapatkan berbagai informasi baru dari pengguna internet ITERA,
d. Mendapatkan informasi baru dari proses BI sebagai data pendukung untuk DSS,

1.4. Manfaat Penelitian

Manfaat yang ingin didapatkan dari penelitian yang dilakukan adalah:

a. Membantu memberikan tambahan informasi untuk membuat keputusan / DSS,

b. Informasi yang didapat dari penelitian ini dapat digunakan untuk menentukan kebijakan selanjutnya terkait internet dan infrastruktur pendukungnya yang sesuai dengan RENSTRA ITERA 2019-2024 bidang pendidikan menuju industri 4.0.

1.5. Batasan Masalah

Pada penelitian ini dilakukan pembatasan masalah yang dibahas sebagai berikut:

a. Proses sains data yang terlibat adalah OLTP, OLAP, DW untuk BI dan DSS,
b. Perangkat lunak yang digunakan adalah Adiscon *Log* Analyzer, Python, dan Apache Spark,
c. Data yang digunakan adalah *log* aktifitas pengguna internet ITERA mulai bulan Mei sampai dengan bulan Juli 2021,
d. Data yang digunakan adalah lalu lintas jaringan internet ITERA.

II. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Kumar S dkk [6] melakukan analisa *traffic* jaringan yang digunakan untuk menganalisa keamanan jaringan. Makalah yang ditulis mengusulkan untuk memiliki ekosistem keamanan yang sehat untuk sejumlah besar situs *web* dari serangan dan ancaman, untuk meningkatkan pengetahuan tentang pola dan tren lalu lintas dan juga untuk melakukan keputusan waktu yang otentik pada *bad traffic*. Bagian pertama, peneliti memasang *Firewall* dengan tambahan Suricata snort yang mempunyai kemampuan inspeksi paket pencegahan dan deteksi intrusi bersama dengan penyebaran sertifikat SSL yang aman. Bagian kedua dilakukan melalui penyebaran *server* syslog di jaringan. Semua algoritma analitik log ditulis dalam python bersama dengan beberapa tampilan berbasis *web*. Program ini membantu dalam menganalisis log dan mengekstraksi *fields* dan informasi penting. Log mentah diubah menjadi format yang dapat dimengerti pengguna dan ditampilkan menggunakan *dashboard* dan tampilan terminal menggunakan Kibana. Dari hasil penelitiannya didapatkan bahwa dengan adanya pembatasan sistem yang diusulkan membantu dalam mengembangkan teknik agregasi dan analisis log secara *real time* melalui tampilan *dashboard* dan terminal, dan peringatan dibuat berdasarkan berbagai kondisi berbahaya. Analisis hasil log dan pembaruan *Firewall* berdasarkan hasil log, peringatan yang dihasilkan dan penerapan sistem keamanan *Server* yang lengkap. Analisis log membantu dalam perhatian keamanan dan deteksi cepat

proses yang gagal, menentukan tren dan data yang disimpan dalam arsip data dan memfasilitasi *streaming* data dinamis. Visualisasi dalam Kibana membantu dalam memahami lanskap ancaman dan memungkinkan keputusan untuk memperkuatnya. Ini juga membantu dalam mengidentifikasi kesenjangan dan celah melalui plot [6].

Sedangkan Week dkk [3] meneliti tentang data log pada *Firewall*. Analisis *Firewall Internet Service Provider* (ISP) bisnis kecil menunjukkan bahwa jaringan diserang sekitar 276 kali per hari. *Firewall* diserang oleh 2.822 host dari 2.699 jaringan. Jaringan ini berada di 108 negara berbeda. Sekitar setengah dari serangan adalah Windows RPC dan SQL Slammer. Serangan itu tersebar merata dari waktu ke waktu. Ada lebih banyak serangan selama jam non-bisnis dan lebih dari 50% serangan terjadi pada hari Minggu, Senin, dan Selasa Waktu Standar Pasifik. Namun, karena perbedaan zona waktu, 53% dari semua serangan berasal dari China pada larut malam, Jumat hingga Sabtu. Sedikit kurang dari setengah serangan berasal dari sepuluh jaringan dan sekitar 25% berasal dari sepuluh *host*. Analisis dari penelitian ini adalah informasi berharga untuk ISP dan peneliti bisnis kecil. Informasi yang diberikan di sini menawarkan bukti kuat akan kebutuhan *Firewall* dan perlindungan jaringan serta membantu administrator jaringan fokus pada area yang merupakan ancaman tertentu. Sebagai contoh, administrator jaringan dapat memilih untuk menerapkan langkah-langkah keamanan ekstra dalam mencegah serangan Windows RPC dan SQL Slammer. Selain itu, dari analisis, pemilik dapat memilih untuk menggunakan lebih banyak sumber daya selama jam kerja non-reguler selama hari Minggu, Senin, dan Selasa [3].

Dalam penelitian lainnya, Kurniawan [7] mencari data tentang penggunaan internet di Universitas Lampung dengan menggunakan teknik *data mining* yaitu metode *clustering*. Berdasarkan hasil analisa dapat diketahui bahwa tingkah laku pengguna di Universitas Lampung berdasarkan tipe autentikasi sebagian besar bertipe tidak terautentikasi sebesar 51,2% pada Juni dan 54,61% di Juli, sedangkan sebagian besar pemakaian data bertipe terautentikasi 96,07% Juni dan 96,36% di Juli. Jumlah pengguna dan pemakaian data lebih banyak berada di dalam ruangan. Pemakaian data tidak terautentikasi sering diakses di Rektorat Lt.2 sebesar 11,42% pada Juni dan 11,91% di Juli. Sedangkan untuk pemakaian data terautentikasi sering di akses pada Juni di FKIP GD I Lt.1 sebesar 3,63% dan Juli di Puskom Tower sebesar 5,31%. Untuk pemakaian data Juni terbesar di Fakultas KIP sebesar 20,03% dan Juli di Fakultas Teknik sebesar 18,33%. Penggunaan internet berdasarkan hari menunjukan penggunaan tinggi pada Senin hingga Jumat dan mengalami penurunan penggunaan pada Sabtu, Minggu dan hari libur. Penggunaan berdasarkan jam menunjukkan penggunaan tertinggi terjadi pada pukul 09.00 WIB hingga 14.00 WIB. Penggunaan internet pada kondisi libur perkuliahan berdasarkan halaman yang diakses menunjukan kategori hiburan yang paling banyak diakses sebesar 52,52%, selanjutnya sosial/komunikasi dengan 17,86%. Penelitian ini membuat tiga kelompok dengan algoritma *k-means* dimana hasil untuk Juni pada C1 adalah nilai tinggi yang banyak terdapat di Fakultas KIP sebanyak empat akses point, pada C2 adalah nilai menengah yang banyak terdapat di Fakultas Teknik dan Fakultas MIPA masing-masing enam akses point, pada C3 adalah nilai rendah yang banyak terdapat di Rektorat sebanyak 12 akses point. Sedangkan untuk Juli pada C1 terdapat satu akses point di UPT TIK, pada C2 banyak terdapat di Fakultas

Teknik sebanyak lima akses point, pada C3 banyak terdapat di Fakultas Pertanian sebanyak 13 akses point [7].

Sementara itu, RV Imbar [8] melakukan *review* tentang *Decision Support System* (DSS) mulai dari arsitektur, perangkat kerasnya, dan *platform* sistem operasi. Hasil reviewnya antara lain DSS adalah sistem informasi yang tujuan utamanya adalah memberikan informasi kepada manajemen. Karakteristik umum dari sebagian besar atau semua DSS termasuk penggunaannya oleh manajer dan pekerja berpengetahuan lainnya, mereka menggunakan *database*, dan menggunakan model. DSS umumnya digunakan ketika komputer tidak dapat diprogram untuk membuat keputusan untuk semua kasus. DSS mendukung, tetapi tidak menggantikan, pembuat keputusan adalah manusia. Untuk arsitektur sistem informasi (SI), penulis tidak mengacu pada merek dan model komputer, *disk drive*, atau monitor yang tepat dalam sistem. Sebaliknya, penulis fokus pada tiga masalah spesifik tingkat tinggi yaitu *Interoperability, Compatibility, Scalability*. Arsitektur keseluruhan dari DSS harus disusun dan dipahami sebelum keputusan pemilihan perangkat keras dan perangkat lunak dibuat. Sifat arsitektur ini bergantung pada DSS. Untuk menyusun arsitektur DSS harus mempertimbangkan spektrum DSS yang akan digunakan organisasi: a. Keputusan strategis, taktis (pengendalian manajemen), dan operasional, b. Keputusan tidak terstruktur, semi-terstruktur, dan terstruktur, c. Semua tingkat manajemen dan pekerja berpengetahuan dalam organisasi, d. Semua fungsional utama, produk atau lini bisnis utama, dan divisi geografis organisasi [8].

A Setiawan [9] melakukan penelitian tentang DSS untuk menentukan jenis supplier dengan metode *Analytical Hierarchy Process* (AHP). Pada penelitianya, penulis

membuat sebuah sistem informasi yang dapat menentukan penilaian AHP. Pemilihan supplier untuk melakukan pembelian barang, departemen pembelian mempunyai 4 kriteria yang harus dipenuhi, diantaranya : Kualitas barang (0.4), Harga barang (0.3), Ketepatan waktu pengiriman dan (0.2), dan Status supplier (0.1). Sistem pengambilan keputusan yang dibuat akan didasarkan pada keputusan yang dihasilkan oleh proses AHP. Dari penelitiannya didapatkan, dengan memasukkan jenis barang yang akan dibeli, kemudian *user* memilih data pemasok dan memasukkan nilai kriteria dari setiap pemasok. Sistem akan menghitung secara otomatis menggunakan AHP dari nilai kriteria yang dimasukkan tadi. Hasilnya akan terlihat pada laporan pemasok mana yang akan dipilih sebagai pemasok sesuai hasil perhitungan [9].

Selanjutnya Syarli dkk [10] melakukan penelitian tentang BIS (*Business Intelligence System*) menggunakan *OnLine Analitycal Processing* (OLAP) dengan melakukan inventarisasi data transaksi yang terjadi antara gudang farmasi dan 11 puskesmas yang berada diwilayah Kabupaten Mamasa. Penulis mengumpulkan data penyakit dari semua puskesmas. Proses pengolahan data dilakukan dengan metode OLAP, kemudian melalui proses *validation, cleaning, transforming, agregating, and loading* proses dinajutkan ke tahap *Extraction, Transformation, Loading* (ETL). Data kemudian dibuat menggunakan dimensi. Hasil pengolahan data akan mengidentifikasi pola-pola tersembunyi dalam data dan menarik informasi baru dari data. Adapun dimensi data sebagai berikut: a. Dimensi pasien, b. Dimensi Jenis penyakit, c. Dimensi jenis obat. Hasil dari penelitian ini adalah data-data penyakit dan penderita tahun 2016 dan 2017. Data-data ini kemudian

diproses menggunakan ETL (OLAP), hasilnya divisualisasikan agar informasinya dapat dengan mudah dipahami. Visualisasi informasi melalui grafik sifatnya interaktif dan menarik sehingga akan memudahkan dalam membaca informasi sementara hasil *print out* (*hardcopy*) akan menjadi dokumentasi bagi manajemen. Informasi yang dihasilkan dari sistem tidak menjadi satu-satunya alat penentu kebijakan tetapi dapat digunakan sebagai bahan pertimbangan untuk pengambil keputusan [10].

Pemrosesan log telah dilakukan oleh He [15] dengan membuat sebuat algoritma yang digunakan memisahkan data log dan juga bisa digunakan data sintetis atau *real-world datasets*. Zulfadhilah [17] melakukan penelitian data log menggunakan data mining dengan algoritma K-Means. Sedangkan Shilin [22] membuat pengurai log menggunakan pendekatan *novel clustering-based* untuk segera dan tepat mengidentifikasi dampak masalah sistem dengan memanfaatkan urutan log (urutan peristiwa log) dan sistem *Key Performance Indicators* (KPI).

Dari penelitian sebelumnya maka peneliti membuat suatu penelitian yang berkaitan dengan hal tersebut dengan judul "Implementasi OLAP Untuk Analisis Log Pada Proses *Business Intelligence* Sebagai Data Pendukung *Decision Support System*". Pada penelitian ini, pemrosesan data log dengan OLAP diharapkan dapat dilakukan secara lebih cepat untuk mengetahui karakteristik pengguna internet di lingkungan Itera. Keluaran data dan informasi baru yang didapat, diharapkan mampu menjadi data pendukung untuk DSS terkait penggunaan internet.

2.2. Business Intelligence

Business inteligence (BI) bukanlah sebuah produk atau sistem, melainkan sebuah arsitektur dan koleksi operasional yang terintegrasi terhadap aplikasi pengambil keputusan dan *database* yang menyediakan pelaku bisnis kemudahan akses kepada data bisnis. BI telah menarik perhatian dari banyak organisasi mengenai kegunaan dan keuntungannya bagi organisasi tersebut. Meskipun begitu, BI tetap dihadapkan pada tantangan untuk memperoleh hasil yang maksimal dari implementasi BI tersebut. Tantangan utama dari BI berhubungan erat dengan pola bisnis yang bersifat unik bagi tiap organisasi, begitu juga dengan kebijakan dan aturan bisnis yang diberlakukan oleh perusahaan. Hal tersebut menyebabkan perusahaan tidak dapat membeli produk BI seperti barang jadi pada umumnya dan berharap dapat memenuhi setiap solusi dari kebutuhan bisnisnya sehingga BI harus dikembangkan sesuai dengan kebutuhan dan proses bisnis perusahaan. Secara umum, aplikasi BI dapat menghabiskan dana jutaan dolar dalam membangunnya. Di sisi lain, implementasi BI dapat memberikan banyak keuntungan. Tidak cuma keuntungan yang berwujud seperti peningkatan volume penjualan, tetapi juga yang tak berwujud seperti meningkatkan reputasi perusahaan. Banyak dari keuntungan tersebut, khususnya yang tak berwujud sulit untuk diukur dalam nominal uang. Oleh karena itu, untuk menjastifikasi keuntungan yang diperoleh dari implementasi BI, harus dihubungkan dengan problem bisnis dan strategi bisnis perusahaan [11].

Secara umum, BI bertujuan untuk menyajikan berbagai informasi yang disesuaikan dengan kebutuhan setiap penggunanya. Informasi tersebut dapat berasal dari mana saja, misalnya dari data histori pembelian barang oleh pelanggan, data histori reparasi, data histori komplain, dan sebagainya. Data-data tersebut kemudian diolah

dan disajikan dalam bentuk informasi yang mudah dicerna oleh penggunanya dengan satu tujuan yaitu membantu pencapaian tujuan bisnis perusahaan. BI memiliki karakteristik sebagai pendukung ketersediaan data yang relevan yang akan disajikan pada pengguna. Biasanya, BI mengintegrasikan informasi dari keseluruhan sumber informasi perusahaan sehingga pembuat keputusan dapat membuat analisis dengan berbekal pengetahuan yang lengkap dan *real time* [12].

2.3. OLTP dan OLAP

OLTP (*On-Line Transaction Processing*) menggambarkan kebutuhan sistem dalam ruang lingkup operasional dan merupakan proses yang mendukung operasi bisnis sehari-hari. Sebuah perusahaan biasanya memiliki sejumlah system-sistem OLTP yang berbeda untuk proses-proses bisnis seperti pengontrolan *inventory*, faktur pelanggan, dan *point-of-sale*. Sistem-sistem ini menghasilkan data operasional yang mendetil, bersifat *current* (terbaru), dan dapat diperbaruhi. Sistem-sistem OLTP dioptimalkan untuk sejumlah transaksi yang bersifat dapat diprediksi, berulang, dan dapat diperbaruhi. Data OLTP diatur menurut kebutuhan transaksi yang berhubungan dengan aplikasi bisnis dan mendukung keputusan sehari-hari untuk sejumlah pengguna operasional [12].

OLAP(*On-Line Analytical Processing*) adalah proses analisis sejumlah besar data *multidimension* yang bersifat dinamis. OLAP dideskripsikan sebagai sebuah teknologi yang menggunakan suatu *view* bersifat *multidimension* mengenai sejumlah data yang teragregasi dan menyediakan akses informasi yang strategis untuk kebutuhan analisis lebih lanjut. OLAP memungkinkan pengguna memperoleh pengertian dan pengetahuan yang lebih mendalam mengenai berbagai

aspek data perusahaan melalui akses data yang cepat, konsisten, dan interaktif. OLAP dapat membantu proses pengambilan keputusan untuk menanggapi kejadian yang akan datang [12]. Tabel 2.1 menampilkan perbandingan antara OLTP dan OLAP.

Tabel 2.1 Perbedaan OLAP dan OLTP [12]

OLTP	**OLAP**
Digunakan untuk mendukung kegiatan transaksi sehari-hari	Digunakan untuk mendukung kegiatan Analisis
Menggunakan *view single* dimensi	Menggunakan *view multi* dimensi
Mendukung keputusan sehari-hari	Mendukung keputusan masa depan
Tidak bergantung pada OLAP	Bergantung pada data yang tersimpan dalam sistem OLTP
Melayani pengguna operasional	Melayani pengguna managerial
Operasi *query*-nya sederhana dan berulang-ulang	Operasi *query*-nya lebih rumit, bersifat *ad hoc*, dan tidak melibatkan operasi *update* data
Menggunakan data sehari-hari	Menggunakan data yang terangkum dalam data *cube*

2.4. Data Warehouse

Sebuah *Data Warehouse* (DW) mengintegrasikan dan mengelola aliran informasi dari *database* perusahaan.

Menurut definisi, DW memiliki sifat-sifat berikut:

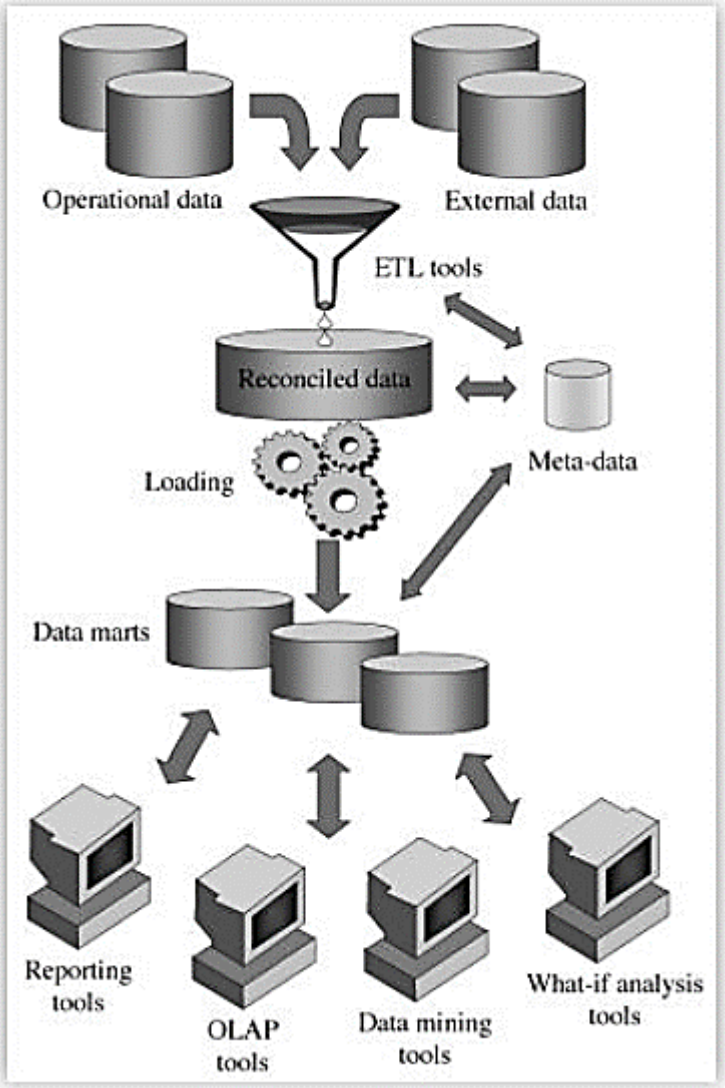
- Berorientasi pada subjek: diatur menurut visi bisnis yang berbeda;
- Terintegrasi: dari sumber data yang heterogen;
- Tidak mudah menguap: selalu dimasukkan, tidak pernah dihapus;
- Varian dalam waktu: posisi historis kegiatan dalam waktu.

Ada beberapa definisi data *warehousing* yang sedikit berbeda, yaitu:

- DW adalah kumpulan data yang berorientasi pada subjek, terintegrasi, varian waktu, dan non-volatil untuk mendukung proses pengambilan keputusan manajemen (1 *keydata*);
- DW adalah teknologi yang menggabungkan data terstruktur dari satu atau lebih sumber sehingga dapat dibandingkan dan dianalisis untuk kecerdasan bisnis yang lebih besar (Informatica);
- Basis data besar (biasanya ditempatkan di sekelompok *server*, atau komputer mini atau *mainframe*) yang berfungsi sebagai gudang terpusat dari semua data yang dihasilkan oleh semua departemen dan unit organisasi besar. Perangkat lunak penambangan data tingkat lanjut diperlukan untuk mengekstrak informasi yang berarti dari gudang data (Kamus Bisnis) [12].

DW bertindak sebagai pusat penyimpanan informasi yang berasal dari satu atau lebih sumber data. Data mengalir dari sistem transaksional dan *database* relasional lainnya ke DW dan umumnya terdiri dari data terstruktur, semi-terstruktur, dan tidak terstruktur. Data ini dimuat, diproses, dan dikonsumsi secara teratur. Pengguna seperti ilmuwan data, analis bisnis, dan pembuat keputusan menggunakan alat BI, klien SQL, dan spreadsheet untuk mengakses data yang diproses di DW [13].

Gambar 2.1 memberikan visi global lingkungan DW. Beberapa elemen penting dari lingkungan DW sudah dapat dilihat, seperti proses ETL, metadata, *data mart*, dan alat OLAP.



Gambar 2.1. Gambaran lingkungan DW secara umum [13]

2.5. Decision Support System

Decision Support System (DSS) atau Sistem Pendukung Keputusan / SPK, secara umum didefinisikan sebagai sebuah sistem yang mampu memberikan kemampuan baik kemampuan pemecahan masalah maupun kemampuan pemgkomunikasian untuk masalah semi-terstruktur. Secara khusus, SPK didefinisikan sebagai sebuah sistem yang mendukung kerja seorang manajer maupun sekelompok manajer dalam

memecahkan masalah semi-terstruktur dengan cara memberikan informasi ataupun usulan menuju pada keputusan tertentu [14].

Konsep DSS diperkenalkan kira-kira pada kurun waktu 1970-an. Pada kurun waktu tersebut DSS masih dalam proses *Research dan Development*. Sedangkan aplikasinya secara meluas dimulai pada kira-kira akhir tahun 1980-an dan awal tahun 1990-an. Dan pada masa yang akan datang DSS masih akan berkembang terus dan memerlukan berbagai perbaikan dan penyempurnaan yang disesuaikan dengan keperluan dan perkembangan teknologi informasi. Di antara perkembangan DSS yang akan terjadi dimasa yang akan datang meputi aspek-aspek: *integrated architecture*, *connectivity*, *document data* dan *intelligence*.

2.5.1 Definisi DSS

Definisi DSS sampai saat ini masih tergantung kepada dari sudut mana DSS tersebut dipandang. Namun pada umumnya DSS bisa didefinisikan dengan melibatkan aspek-aspek sebagai berikut:

- Sistem yang berbasasis komputer
- Membantu memecahkan masalah seorang manager
- Masalah semi terstruktur
- Interaktif diantara sistem dan manager
- Menggunakan analisis data

Kedua aspek yang terakhir adalah berasaskan aplikasi teknologi yang kemudian disebut dengan DDM (dialog, data dan modelling).

2.5.2 Tujuan DSS

Menurut Peter G. W. Keen dan Michael S. Scott Morton mengemukakan bahawa perinsip dasar konsep DSS adalah struktur masalah, dukungan keputusan dan efektivitas keputusan [13]dalam[14].

Dari ketiga konsep tersebut maka disusunlah tujuan DSS, sebagai berikut:

- DSS dapat membantu manger dalam membuat keputusan untuk memecahkan masalah semi struktural,
- DSS dapat mendukung terhadap penilaian manager
- DSS dapat meningkatkan efektivitas dan efisiensi seorang manager dalam mengambil suatu keputusan.

Agar berhasil mencapai tujuannya maka sistem tersebut harus: sederhana, *robust*, mudah untuk dikontrol, mudah beradaptasi, lengkap pada hal-hal penting, mudah berkomunikasi dengannya [14].

2.6 Metode yang digunakan

Pada penelitian ini data log diambil dari *Firewall Router* Fortigate yang berada di ITERA selama 3 bulan yang dimulai dari Mei - Juli 2021. Data log ini merupakan data seluruh kegiatan lalu lintas internet di ITERA. Data log diambil menggunakan RSyslog dan ditampilkan dengan Adiscon Log Analyzer dalam betuk baris log yang masih sulit untuk dipahami. Kemudian dilakukan pemilihan data yang akan dipakai menggunakan proses ETL. Pemilihan data log dengan mengambil atribut alamat IP, waktu akses, negara tujuan, negara asal, dan aplikasi yang dipakai. Data hasil pemilihan ini selanjutnya di simpan dalam *database*. Proses ETL juga digunakan untuk mengelompokkan data yang sudah diproses sebelumnya.

2.7 Evaluasi

Untuk mengevaluasi apakah algoritma pemilihan data log dalam penelitian ini, yaitu dengan cara melihat visualisasi yang ditampilkan oleh plot Python. Hasil yang baik adalah dapat ditampilkannya visualisasi kebiasaan dan pola perilaku penggunaan internet di ITERA. Negara mana saja yang paling sering dikunjungi, kapan waktu yang paling banyak dipakai untuk memakai internet, dan dimana para pengguna memakai internet di lingkungan kampus ITERA.

III. METODE PENELITIAN

3.1. Waktu Dan Tempat Penelitian

Tempat yang digunakan dalam penelitian ini adalah ITERA (Institut Teknologi Sumatera), dimulai pada bulan Mei 2021 sampai dengan Juni 2021, organisasi tesis dimulai dari penyusunan proposal sampai dengan pelaporan hasil penelitian.

3.2. Perangkat dan peralatan penelitian

Dalam penelitian ini untuk mendapatkan pengguna log dan memprosesnya menjadi informasi, maka alat dan bahan yang digunakan dapat dilihat pada table 3.1:

Table 3.1 Perangkat dan peralatan penelitian

No.	Perangkat dan peralatan	Keterangan
1	Fortigate 601E	*Firewall* yang diambil data log
2	RSyslog	*Software* pengambil dan pengirim data log
3	Adiscon Log Analyzer	Menampilkan data mentah dan memasukkan ke *database*
4	Python *programming language*	*Software* untuk Proses ETL dan visualisasi
5	Apache Spark	*Software* ETL

3.3. Spesifikasi Sistem

Sistem yang dikembangkan dalam penelitian ini terdiri dari beberapa bagian, yaitu spesifikasi OLTP (*On-Line Transaction Processing*), *Data Warehouse* (DW), dan OLAP (*On-Line Analytical Processing*) yang merupakan bagian dari *Business Intelligence* (BI):

a. OLTP
    i. Dapat mencatat data transaksi internet pengguna
    ii. Menggunakan Fortigate 601E sebagai *Firewall* yang diambil data log aktifitas penggunanya untuk keseluruhan lalu lintas internet
    iii. Dapat mengkategorikan data transaksi sesuai dengan jenis *url* yang diakses dan waktu akses.
    iv. Dapat mengkonversi data transaksi kedalam *database*

b. DW
    i. Merupakan *database* yang berisi kumpulan data transaksi yang sudah diproses OLTP
    ii. Berisi data yang sudah di ekstraksi, dirubah kedalam bentuk *database* yang siap digunakan (ETL/*extracting, transforming, loading*)

c. OLAP
    i. Berisi kumpulan data transaksi yang sudah diproses OLTP

ii. Merupakan kumpulan semua data-data transaksi dari seluruh pengguna internet ITERA yang berada dalam *database*

iii. Dapat melakukan *query* ke semua data *warehouse*

3.4. Prosedur Penelitian

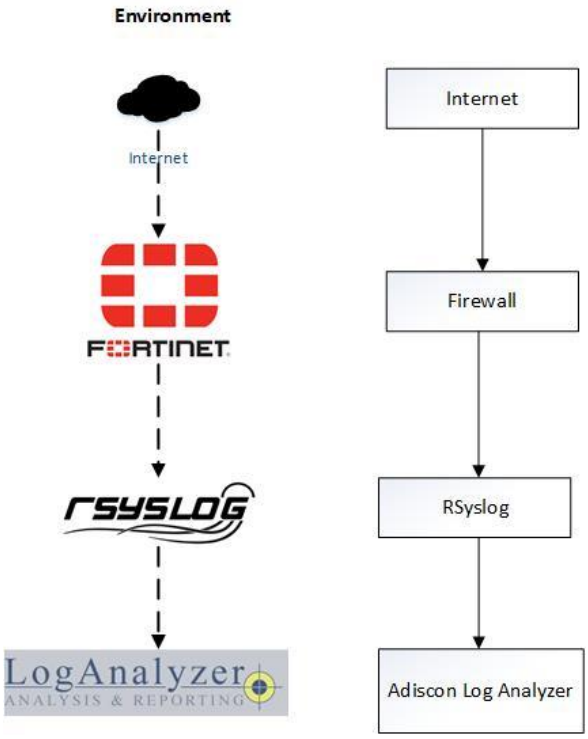
Prosedur penelitian yang digunakan termasuk dalam penelitian *Research and development* (R&D). Menurut Sugiyono [14], penelitian dan pengembangan adalah metode penelitian yang digunakan untuk menghasilkan dan mengembangkan sebuah produk serta menguji validitas produk yang dihasilkan.

Penelitian ini menghasilkan data pengguna internet yang berupa banyaknya negara tujuan akses, negara asal akses, waktu kapan yang paling banyak memakai internet, dan aplikasi apa saja yang banyak dipakai. Adapun prosedur penelitian yang dilakukan yaitu mengumpulkan data log pengguna internet dari *Firewall* menggunakan Adiscon Log Analyzer, menyeleksi data log mana data yang dibutuhkan dan tidak, mengkategorikan data yang didapat, mengkorelasikan antar kategori, dan terakhir adalah memvisualisasikan hasil yang didapat. Semua algoritma analitik log ditulis dalam python. Program ini membantu dalam menganalisis log dan mengekstraksi bidang dan informasi penting. Log mentah diubah menjadi format yang dapat dimengerti pengguna dan ditampilkan menggunakan dashboard dan tampilan plot Python.

3.5. Rancangan Model Sistem

Model sistem yang digunakan dalam penelitian ini dibagi menjadi beberapa bagian sebagai berikut:

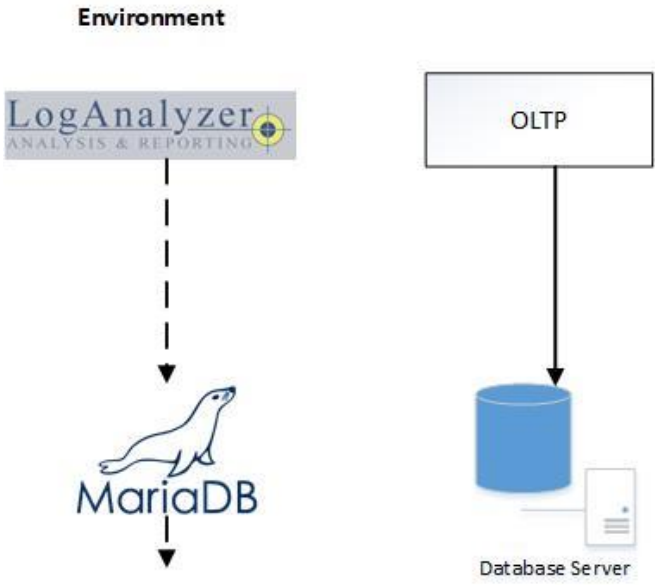
3.5.1 OLTP



Gambar. 3.1 Blok diagram OLTP

Pada tahap OLTP seperti terlihat pada gambar 3.1, data transaksi internet dihasilkan oleh *Firewall* kemudian diambil oleh RSyslog. Selanjutnya data dikirimkan ke syslog *server* dalam bentuk teks dan diterima oleh Adiscon Log Analyzer sebagai perangkat lunak OLTP sehingga menghasilkan sebuah file log. File log tersebut di-extract dan disimpan ke *database* yang dapat dianalisis secara realtime ataupun dengan waktu tertentu menggunakan *software* Adiscon Log Analyzer.

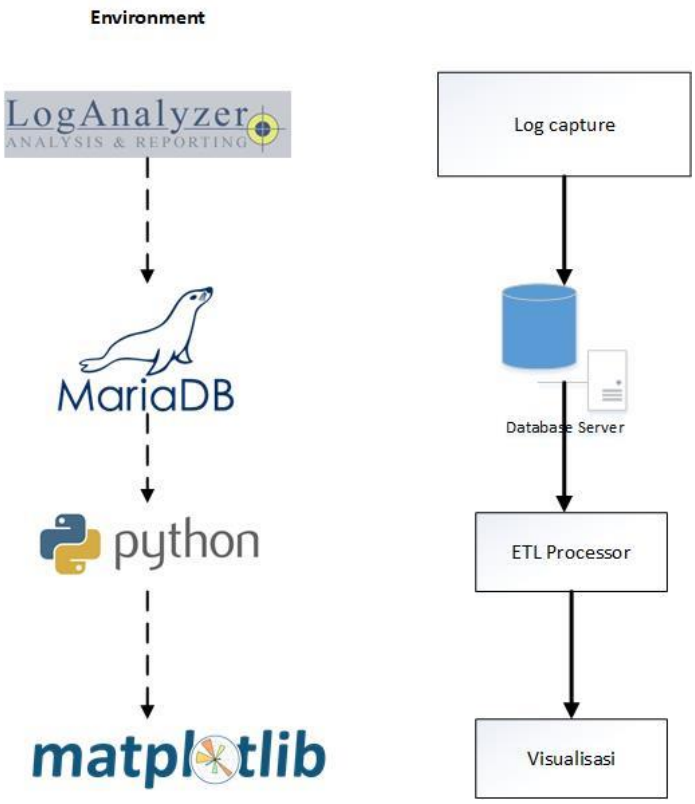
3.5.2 *Data Warehouse*



Gambar. 3.2 Blok diagram *Data Warehouse*

Gambar 3.2 memperlihatkan data log yang sudah ada dalam *database* OLTP, kemudian dilakukan proses ETL. Hasil proses ETL selanjutnya disimpan dalam *database* yang berada pada *log store*. *Database* ini berisi semua data transaksi dari jaringan internet ITERA yang selanjutnya akan digunakan untuk *data warehouse*.

3.5.3 OLAP

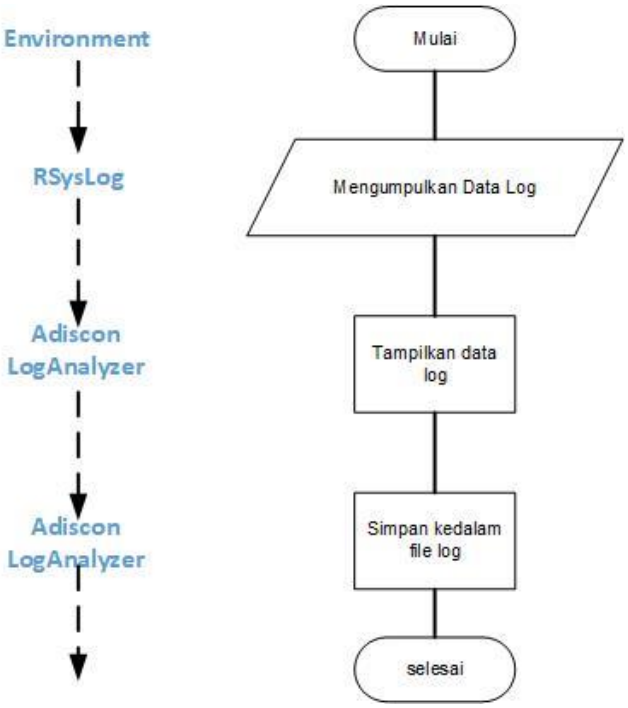


Gambar. 3.3 Blok diagram OLAP

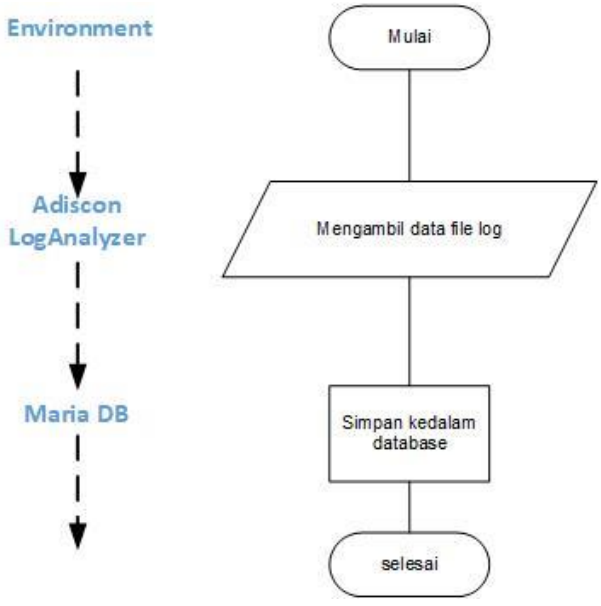
Gambar 3.3 menunjukkan data log yang sudah diekstrak dan disimpan di *database*, selanjutnya diproses menggunakan algoritma pembersihan log menggunakan pemrograman Python. Atribut yang digunakan adalah alamat IP, waktu akses, dan aplikasi yang diakses dan digunakan. Menampilkan hasil olahan data menggunakan visualisai grafik yang lebih mudah dipahami.

3.6. Diagram Alir penelitian

Adapun diagram alir penelitian ini juga terbagi dalam beberapa tahap, sebagaimana ditunjukkan pada Gambar 3.4 – Gambar 3.7:

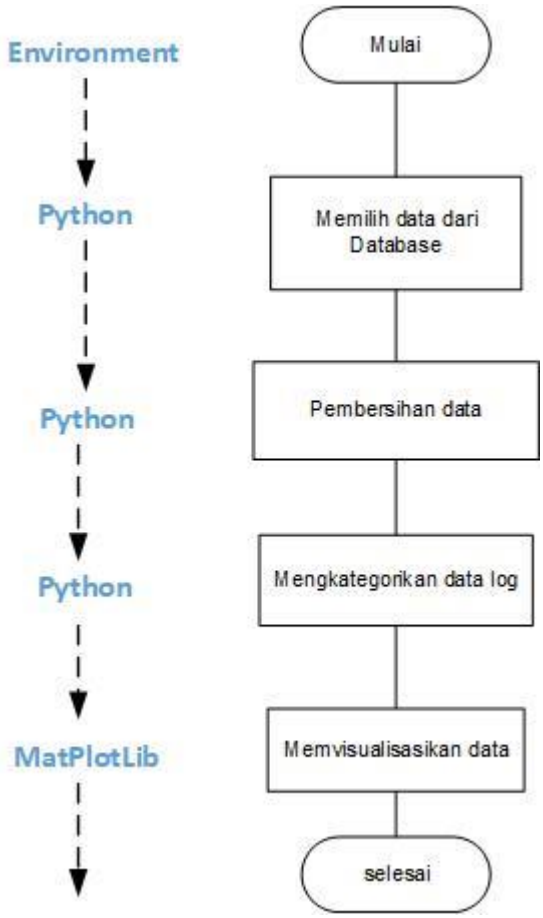


Gambar 3.4 Diagram alir penelitian OLTP



Gambar 3.5 Diagram alir penelitian DW

Penelitian dimulai dengan pengumpulan data dari *Firewall* di tunjukkan pada gambar 3.4. Data log diambil menggunakan perangkat lunak RSyslog, kemudian dikirim ke perangkat lunak Adiscon Log Analyzer menggunakan protokol TCP port 514. Data log tersebut dapat ditampilkan dan otomatis disimpan kedalam file log. File log tersebut kemudian dimasukkan kedalam *database* yang dijadikan sebagai *data warehouse* seperti terlihat pada gambar 3.5.

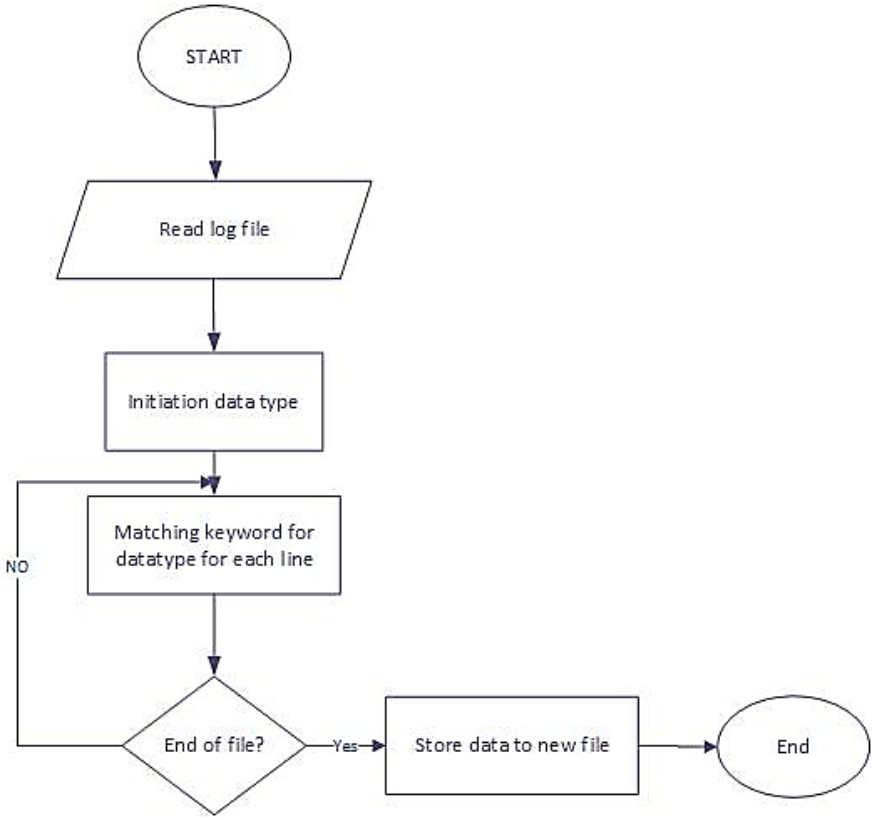


Gambar 3.6 Diagram alir penelitian OLAP

Gambar 3.6 menunjukkan alur penelitian OLAP. Data yang ada pada *database* kemudian dilakukan proses ETL yaitu dengan mengkategorikan sesuai dengan atribut yang akan di ambil menggunakan Python. Data yang sudah terkategorikan

tersebut divisualisasikan menggunakan Matplotlib agar data dapat dianalisa dengan lebih mudah.

Algoritma pembersihan data log diperlihatkan pada gambar 3.7. Langkah pertama adalah membaca file log, kemudian inisiasi *keyword type* data yang akan diambil. Selanjutnya mencocokkan *keyword* yang akan diambil pada tiap baris data log. Jika menemukan *keyword* yang sama, kemudian atribut data tersebut disimpan pada file yang baru untuk dilakukan analisa selanjutnya.



Gambar 3.7 Algoritma untuk pembersihan log dengan pemrograman Python

3.7. Tahapan Penelitian

Secara umum penelitian ini dilakukan dengan dua tahapan sebagai berikut:

1. Pengambilan data log (OLTP dan DW)
    a. Pengambilan log transaksi dari *Firewall* menggunakan RSyslog. Data yang diambil adalah data penggunaan internet dijaringan internet ITERA
    b. Memasukkan data log ke dalam *database* menggunakan Adiscon Log Analyzer

2. OLAP: Proses ETL (*Extract, Transform, Load*) menggunakan Python
    a. *Selecting*, yaitu memilih data log dari DW mengunakan skrip python. Data yang dipilih adalah semua aktifitas pengguna internet di ITERA.
    b. *Cleaning* dan *sorting*, yaitu proses pembersihan data log yang sudah dipilih untuk mengambil atribut-atribut yang dibutuhkan menggunakan algoritma pada gambar 3.7. Atribut yang di ambil adalah: waktu akses, IP sumber, IP tujuan, negara sumber, negara tujuan, dan aplikasi yang digunakan.
    c. Klasifikasi / Kategorisasi, adalah mengkategorikan atribut yang sesuai dengan kategori yang akan dibuatkan grafik/ visualisasinya.
    d. Visualisasi, adalah membuat grafik/ visualisasi atribut yang sudah dikategorisasi sebelumnya menggunakan Matplotlib.

V. SIMPULAN DAN SARAN

5.1 SIMPULAN

Penelitian berhasil mendapatkan data dan informasi mengenai kapan waktu tersibuk akses internet, IP sumber, IP tujuan, aplikasi yang diakses, dan lokasi akses. Data ini dapat digunakan untuk pendukung kebijakan mengenai koneksi dan akses internet bagi sivitas akademika ITERA. Berdasarkan hasil analisa dan pembahasan dapat diambil simpulan:

1. Penelitian ini berhasil menganalisa secara *online* menggunakan *Log* Analyzer sebagai OLTP dan juga menyimpan data *log* di *database* sebagai *Data Warehouse*.
2. Trafik tertinggi penggunaan akses internet adalah diantara pukul 07.00 wib sampai dengan pukul 17.00 wib, dimana waktu tersebut adalah jam kerja di lingkungan ITERA.
3. Untuk tujuan akses, para pengguna banyak yang mengakses aplikasi sosial media, *update OS Windows*, aplikasi *streaming*, dan *online meeting*. Aplikasi-aplikasi tesebut memakai IP dari negara Amerika Serikat/AS, sehingga AS merupakan negara yang paling banyak dikunjungi oleh pengguna internet di ITERA.

4. Aplikasi sosial media menjadi aplikasi yang paling banyak di akses oleh pengguna internet. Hal ini dapat dilihat dari hasil penelitian yang menunjukkan bahwa Facebook merupakan aplikasi yang paling banyak diakses selama 24 jam, disusul oleh TikTok, WhatsApp, Instagram, YouTube, dan Telegram.
5. Lokasi yang paling banyak menggunakan internet di lingkungan ITERA adalah gedung GKU, dimana gedung ini menjadi tempat perkuliahan hampir semua program studi dan juga sebagai tempat UPT TIK berada yang merupakan unit yang mengatur koneksi internet di ITERA dan aula tempat terlaksananya kegiatan-kegiatan di ITERA.
6. Informasi baru yang didapatkan dari penelitian adalah terkait waktu trafik tertinggi, aplikasi yang di akses, dan dimana lokasi akses terbanyak. Informasi ini dapat dijadikan data pendukung untuk pengambilan keputusan atau kebijakan terkait internet dan infrastruktur pendukung internet di ITERA.

5.2 SARAN

1. Pada penelitian ini, data hasil masih berupa hasil proses data lalulintas internet dan belum dilakukan analisa apakah data tersebut benar dilakukan oleh pengguna atau *virus, bot,* atau lainnya. Untuk penelitian selanjutnya, diharapkan dapat melakukan analisa yang lebih dalam untuk memastikan data *log* yang benar-benar dari pengguna internet.

2. Untuk analisa yang lebih mendalam, dapat digunakan *machine learning* atau *deep learning*.

DAFTAR PUSTAKA

[1] A. D. Prasetyo. "LKP : Penerapan Jaringan Wireless Berbasis Router Mikrotik dengan Menggunakan Metode VTP (Vlan Trunk Protocol) di PT. Pertamina Pabrik Aspal Gresik". Skripsi, S1 Teknik Komputer, Universitas Dinamika, Surabaya, 2017.

[2] Nanan A. *Optimalisasi Firewall Pada Jaringan Komputer Berskala Luas*. JSI (Jurnal sistem Informasi). Vol 1 no 1. Universitas Suryadarma. 2014.

[3] J. Week et al. *A Firewall Data Log Analysis of Unauthorized and Suspicious Traffic*. Journal of Information System Security. Vol. 7 Issue 3, p2-15. 14p. September 2011.

[4] M Geaur Rahman. *OLAP-based Decision Support System for Business Data Analysis*. International Journal of Computer Applications (0975 – 8887), Volume 181, No. 18. September 2018.

[5] S. Dumais, R. Jeffries, D. M. Russel, D. Tang and J. Teevan, "Understanding User Behavior Through Log Data and Analysis" in *Ways of Knowing in HCI*, New York, Springer New York, 2014, pp. 349-372.

[6] S. Kumar et al. *Analysis of Network Traffic and Security through Log Aggregation*. International Journal of Computer Science and Information Security (IJCSIS), Vol. 16, No. 6, June 2018.

[7] A Kurniawan. *Analisa Perilaku User Dengan Menggunakan Big Data Jaringan Universitas Lampung*. Skripsi, S1 Teknik Elektro, Universitas Lampung, Bandar Lampung, 2019.

[8] R.V. Imbar. *Decision Support System Architecture, Hardware, and Operating System Platforms*. Jurnal Sistem Informasi Vol. 2 No. 1, page: 41-50, Maret 2007.

[9] A. Setiawan. *Implementasi Aplikasi Decision Support System Dengan Metode Analytical Hierarcy Process (Ahp) Untuk Penentuan Jenis Supplier*. Gaung Informatika, Vol. 2 No. 2, hal: 93-104. Juli 2009.

[10] Syarli dkk. *Perancangan Business Intelligence System Pada Gudang Farmasi Dinas Kesehatan Kabupaten Mamasa*. Jurnal Keteknikan dan Sains (JUTEKS) – LPPM UNHAS Vol. 1, No.1, Juni 2018.

[11] S Darudiarto dkk. *Business Intelligence: Konsep Dan Metode*. Jurnal Communication and Information Technology (CommIT), Vol. 4 No. 1, hlm. 63 – 67. Universitas Bina Nusantara. Mei 2010.

[12] Imelda. *BUSINESS INTELLIGENCE*. Majalah Ilmiah UNIKOM, Vol.11 No. 1, hlm 111-122. Universitas Komputer Indonesia. 2013.

[13] F. Almeida, *Concepts and Fundaments of Data Warehousing and OLAP*. Porto, Portugal, ISSUU Publishing, 2017.

[14] G. R. Rachman. (2005). Teaching Materials & Files. Available: url: http://galihrakacita.staff.gunadarma.ac.id/Downloads/folder/0.2.

[15] Pinjia H, dkk. "*Towards Automated Log Parsing for Large-Scale Log Data Analysis*". IEEE Transactions on Dependable and Secure Computing, Volume 15, Issue: 6, page 931-944. 2018. Available Online: https://ieeexplore.ieee.org/document/8067504

[16] Sugiyono, *Metode Penelitian dan Pengembangan*. Cetakan ke 4, Penerbit ALFABETA, Bandung. 2017.

[17] M. Zulfadhilah, I. Riadi, Y. Prayudi. *Log Classification using K-Means Clustering for Identify Internet User Behaviors*. International Journal of Computer Applications, pp 0975 – 8887, Volume 154 – No.3, November 2016.

[18] H. Rana, M. Patel. *A Study of Web Log Analysis Using Clustering Techniques*. International Journal of Innovative Research in Computer and Communication Engineering. Vol. 1, Issue 4, June 2013.

[19] M. W. Talakua, Z. A. Leleury, A. W. Talluta. *Analisis Cluster Dengan Menggunakan Metode K-Means Untuk Pengelompokkan Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014*. Jurnal Ilmu Matematika dan Terapan. Volume 11 Nomor 2. Hal. 119 – 128. Desember 2017.

[20] A. Iswardani, I. Riadi, *Denial Of Service Log Analysis Using Density K-Mans Method*, Journal of Theoretical and Applied Information Technology, Vol. 83, no. 2, pp. 299–302, 2016

[21] D. R. Ningrat, D.A.I. Maruddani, T.Wuryandari. *Analisis Cluster Dengan Algoritma K-Means Dan Fuzzy C-Means Clustering Untuk Pengelompokan Data Obligasi Korporasi*. Jurnal Gaussian, Volume 5, Nomor 4. Halaman 641-650. 2016. Available Online: Http://Ejournal-S1.Undip.Ac.Id/Index.Php/Gaussian

[22] S. He, et al. *Identifying Impactful Service System Problems via Log Analysis*. ESEC/FSE '18. Lake Buena Vista, FL, USA. November 4–9, 2018.