IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

DEVELOPMENT OF A SOFTWARE FOR LOGINDEXING WITH ALERT CUSTOMIZATION (LIAC)

Sabarish Krishnan
B. TECH ISDF Student,
Department of Computer Science and Engineering,
DR.MGR Educational and Research Institute, Chennai, Tamil Nadu, India

Abstract: As the Cyber space is becoming more and more dangerous, security tools are coming up with new and effective methods to protect the confidential data. Those tools will provide a basic security measure which will not even be relevant to the user's situation most of the time. If you want to have a better control over the security you should ideally reach out for SIEM tools l like Splunk if you are a giant organization who is ready to spend a lot of money. But what about the normal people who is concerned about their data security? This is where LIAC comes into picture. The main objective of the tool is to parse the logs that are generated on their system – pass it to a set of customizable rules – and to trigger the alert. These rues can be changed according to the e user' s particular needs. The entire project is expected to be completed by the end of January 2023. The Programming languages used include ~ Front end: HTML/CSS, Java Script, API: python with flask, Back end: Apache server, phpMyAdmin for managing MySQL database

Keywords: Log parsing, monitoring, Action prevention, custom use- case creation

I. INTRODUCTION

The number of devices connected to the internet and computer networks around your workspace are constantly growing. As the number of devices connected to a network increase so does the amount of traffic and complexity of the network. Most devices in a network run some sort of application or service which produce log messages. These log messages can contain vital information about the system or software execution and information about security threats to a device or the network. A log file is a collection of messages created when an event occurs during a program's execution. These events can be anything from a failed authentication of a user to an error condition within the program. The log message is created to describe the event that occurred and then stored within the programs log file. These log files can then be parsed for information which can provide more information about how the program is executing or if there have been any security incidents.

Examples include:

- Application logs can identify what transactions have been performed, at what time, and for whom. Those logs may also describe the hardware and operating system resources that were used to execute that transaction.
- Network devices, such as firewalls, intrusion detection/prevention systems, routers, and switches are generally capable of logging information. These logs have value of their own to network administrators, but they also may be used to enhance the information in application and other logs.

Because of this, a good centralized log management system is becoming a crucial part of any network today. The centralized log managements system is used to collect and store the logs from different devices and allows easier access to the log data. It can be used in small office/home office (SOHO) with just a couple of devices. With proper management, these logs can be of great benefit in a variety of scenarios, to enhance security, system performance and resource management, and regulatory compliance. In particular, a log management infrastructure can capture information and aid analysis about the following:

- Access - Who is using services.
- Change Monitoring - How and when services are modified.
- Cost Allocation - When applicable, who should pay how much for services.
- Malfunction - When services fail.
- Resource Utilization - How much capacity is used by services.
- Security Events - What activity occurred during an incident, and when. □ User Activity - What people are doing with services.

1.1 BENIFITES

The following are example scenarios illustrating how the information in logs can be critical to resolving a security or operational issue:

- Internet access is very slow for 10–20-minute periods at random times throughout the day. Router logs identify a high rate of transmission errors on the campus's Internet connection at those times. The network administrator calls the campus's ISP to repair the connection.
- Internet access is very slow, but everything seems to be working correctly. Firewall and router logs determine that a particular PC in the dorms is under attack from over 500 addresses located around the world, most likely a botnet. This allows the network administrator to have that PC's Internet traffic blocked, relieving the campus's networkwide congestion.
- A student has been dropped from all of her classes and accuses the Registrar and threatens to sue. An application log shows that her user ID had been used to drop the classes two hours after she had dropped her boyfriend, and that the session originated from her boyfriend's dorm room. She remembers that she had given her password to her ex.
- A server containing sensitive information (Social Security numbers) was compromised by hackers who installed an FTP service and pornographic images. To determine who, how and extent of the compromise the logs from the server operating system, the FTP software, firewall and IDS system are all analysed as part of the investigation.

1.2 MOTIVATION

There are various applications (known as log file analyzers or log files visualization tools) that can digest a log file of specific vendor or structure and produce easily human readable summary reports. Such tools are undoubtedly useful, but their usage is limited only to log files of certain structure. Although such products have configuration options, they can answer only built-in questions and create built-in reports. And these applications can be only used my big organizations to monitor and secure their infrastructure. But what about general people who is concerned about their system security or security of their private area network?

The LIAC tool will let you take the complete control and customize your environment according to your personal needs. For instance, if you are a person who always enters the password right during system login – you can set an alert which notify even a single login failure happens and also ley you to enforce an action like immediately shutting down the system

II. LITERATURE SURVEY

2.1 ISSUES ON WINDOWS EVENT LOG

Due to the rapidly increasing connectivity and dependency over the internet by individuals and corporations to carry out their businesses, security breaches are increasing day by day. Security and privacy are becoming a greater concern for the modern world. In this context, log data are very useful as it is used to track the history of an intruder in day-to-day work and providing evidence for further investigation. [6]

2.2 WINDOWS POWESHELL IN ACTION

Windows PowerShell is the next-generation scripting environment created by Microsoft. It's designed to provide a unified solution for Windows scripting and automation, able to access the wide range of technologies such as .NET, COM, and WMI through a single tool. Since its release in 2006, PowerShell has become the central component of any Windows management solution. PowerShell can be used for text processing, general scripting, build management, creating test frameworks, and so on. [9]

2.3 CENTRALIZED LOG MANAGEMENT FOR COMPLEX COMPUTER NETWORK

In modern computer networks log messages produced on different devices throughout the network is collected and analyzed. The data from these log messages gives the network administrators an overview of the network's operation, allows them to detect problems with the network and block security breaches. In this thesis several different centralized log management systems are analyzed and evaluated to see if they match the requirements for security, performance and cost which was established. These requirements are designed to meet the stakeholder's requirements of log management and allow for scaling along with the growth of their network. [1]

2.4 DESIGN OF SECURE LOG MANAGEMENT OVER CLOUD

A Log is consisting of much helpful data regarding activities or events of systems and networks and these data having number of attributes and own syntax. These logs are made-up of events which has been done by users on systems or in networks. This information is very expensive for organizations. These logs are used for finding problems, to optimize performance, to record all events, and to investigate malicious activity in systems or networks. So, protection from attackers is required. Hence organization should maintain integrity, confidentiality, security of logs. we propose more effective secure cloud-based log management to decrease cost and provide security of logs from attackers. [4]

2.5 UTILIZING A SQL DATA STORE FOR SCALABLE LOG ANALYSIS

A potential problem for persisting large volume of data logs with a conventional relational database is that loading massive logs produced at high rates is not fast enough due to the strong consistency model and high cost of indexing. As a possible alternative, a modern SQL data store, which sacrifices transactional consistency to achieve higher performance and scalability, can be utilized. [7]

2.6 EFFICIENT WAY OF WEB DEVELOPMENT USING PYTHON AND FLASK

Web is the most frequently and rapidly used networking aid which satisfies the requirements of all type of users and which provides a solution for any type of problem. For designing and developing such well-defined and well structured, we have to choose a proper technology. Therefore, a dynamic web application or portal can be developed by using flask and python. Various applications such as hand gestures can be easily developed by python which helps various people. Good development of a web page or an application can be easily attracting users which leads to success of the project. The technological needs of a web development project can be achieved by using "python" & "flask". [8]

2.7 ALERT MANAGEMENT AND CORRELATION

Alert management includes functions to cluster, merge and correlate alerts. The clustering and merging functions recognize alerts that correspond to the same occurrence of an attack and create a new alert that merges data contained in these various alerts. The correlation

function can relate different alerts to build a big picture of the attack. The correlated alerts can also be used for cooperative intrusion detection and tracing an attack to its source. [5]

III. PROBLEM STATEMENT

The overall goal of this research is to invent and design a good model of generic processing of log files with alert customization in it. The following list sums up areas involved by this research. There isn't any tool in the market trend which does this, in generic SEIM tool one has to hire a cyber security professional to write the use cases based on their environment. Organization would cost certain crores per year to monitor the client's environment, but in here you are the client and you are allowed to write – experiment – create your own use cases. This will cost you less comparatively but a downside is that the person who is likely to use the LIAC should have certain level of knowledge in the field of security

IV. EXISTING SYSTEM

The current tools are considered as a useful tool by security researchers for threat management of a company. They are mostly favoured by public companies and large organizations that place a huge importance on the compliance and regulations. Many similar log management and alert management tools are there as an opensource projects with appetizing frontend but none of them will provide much control and customization over the logs and the usecase creation.

V. PROPOSED SYSTEM

LIAC is a local web hosting software which primarily contains three main pages and a documentation page act as a guide which tells the user on how to use the software. They are alert page, usecase-management page and the create-usecase page. Each page is responsible for processing and displaying specific tasks. The database will be stored with different type of logs in a processable format. The create-usecase page will let the user to create a custom usecase based on the threats that user's environment. This page is highly customizable because the user is holding the raw logs and can manipulate to its extends. The usecase-management page will display all the created / built-in usecases along with the fields. Alerts will be triggered and displayed in the alert page if all the fields in the usecase matched the logs on the database. It also sends an email to the client, i.e., the user whenever an alert is triggered.

VI. FUTURE SCOPE

Only the windows security logs can be manipulated. The future scope is to let the user manipulate with other windows logs. The input fields in the "custom use-case" page are limited and can't be left empty, which restricts the user with limited amount of input. The future scope is to let the user select a variety of input and select the needed one using a drag and drop mechanism.

VII. MODULES INVOLVED

7.1 DIFFERENT LOG SOURCES

There are many sources of log information, they are Application log, System log, Network log and Security log

7.1.1 APPLICATION LOG

They have the potential to identify what transactions have been performed, at what time, by whom, and on what object. Those logs may also describe the client and server hardware and operating system resources that were used to execute that transaction. Applications should log their activity in a manner that correlates well with the business processes the applications support, particularly any operations that modify permissions or access rights. These logs should include:

- The business operation that was requested
- Whether the request was accepted or denied
- The time and date the operation was
- Who initiated the operation
- System and network resources used
- Any information needed for business process controls
- Client hardware and software characteristics

7.1.2 SYSTEM LOG

Many components of the IT infrastructure generate logs., such as web, database, authentication, print, etc. provide detailed information about their activity and are an integral part of system administration. When related to application logs, they provide an additional layer of detail that is not observable from the application itself. Service logs can also aid in intrusion analysis, when an intrusion bypasses the application itself. Examples of these components include: Operating Systems, Web servers, Database servers, Print servers, File servers, Authentication servers, DHCP servers, DNS servers, electronic mail server logs

7.1.3 NETWORK LOG

Network devices, such as firewalls, intrusion detection/prevention systems, routers, and switches are generally capable of logging information. As before, these logs add purpose of their own to network administrators, but they also can enhance the information in operating system, service, and application logs. Examples of these components include: Routers, Switches, Wireless access points, Network-based firewall, Host-based firewalls, Intrusion detection and prevention systems, Telephone Switches

These logs typically describe flows of information through the network, but not the individual packets contained in that flow. A flow is the traffic that corresponds to a logical connection between two processes in the network. Examples of flows include a connection to a web server, a remote login session, or a Domain Name System lookup. Information logged for a flow should include: Network (IP) addresses or telephone numbers of the end points, Service identifiers (port numbers) for each of the end points, Whether the flow was accepted or denied, Date, time, and duration of the flow, Number of packets and bytes used by the flow.

7.1.4 SECURITY LOG

The Security Log, in Microsoft Windows, is a log that contains records of login/logout activity or other security-related events specified by the system's audit policy. Auditing allows administrators to configure Windows to record operating system activity in the Security Log. Local Security Authority Subsystem Service writes events to the log. The Security Log is one of the primary tools used by Administrators to detect and investigate attempted and successful unauthorized activity and to troubleshoot problems

Following are the main elements of an event log:

- **Log name:** Name of the event log to which events from different logging components will be written. Events are commonly logged for system, security, and applications.
- **Event date/time:** Includes the date and time when the event occurred.
- **Task category:** Identifies the type of recorded event log. Application developers can also define task categories to serve as extra information about the event.
- **Event ID:** This Windows identification number helps network administrators uniquely identify a specific logged event.
- **Source:** Name of the program or software causing the event log.
- **Level:** Event level represents the severity of the recorded event log. These include information, error, verbose, warning, and critical.
- **User:** Name of the user who logged onto the Windows computer when the event occurred.
- **Computer:** Name of the computer logging the event.

7.2 DATA MANAGEMENT PROCESS

Data management is the practice of collecting, organizing, and accessing data to support productivity, efficiency, and decision-making. The data management process includes a wide range of tasks and procedures, such as: Collecting, processing, validating, and storing data - Integrating different types of data from disparate sources, including structured and unstructured data - Ensuring high data availability and disaster recovery - Governing how data is used and accessed by people and apps.

7.2.1 DATA TARGET – EVENT VIEWER

The initial step is to target which piece of data that we are focusing on and where we can get that data from. As for the first step the targeted data is windows security logs that are logged in our local system by LSASS (Local Security Authority Subsystem Service). Since this is just an experimental phase, just a single domain of log is taken into test. Now that we targeted our data let's see where we can find this data.

Event Viewer is a component of Microsoft's Windows NT operating system that lets administrators and users view the event logs on a local or remote machine. Applications and operating-system components can use this centralized log service to report events that have taken place, such as a failure to start a component or to complete an action. The Event Viewer uses event IDs to define the uniquely identifiable events that a Windows computer can encounter. The event viewer can be easily found by searching it on the start menu. After it's been boot up, navigate to windows logs -> security on the top left corner to see all the security related logs that have been logged by your operating system.

7.2.2 DATA COLLECTION – POWERSHELL SCRIPT

PowerShell is an object-oriented automation engine and scripting language with an interactive command-line shell that Microsoft developed to help IT professionals configure systems and automate administrative tasks. So, there is a manual way to fetch the data that we needed but there lies a problem. When we want to work with live logs that's generating on your machine you have to automate the process which results in automatically fetching our required data from event viewer. This is where we are going to use PowerShell to make our job easy. This PS script will fetch our data from event viewer and store it in the desirable location that we need to proceed further.

7.2.3 DATA STORAGE – MYPHPADMIN DATABASE

A data store is a repository for persistently storing and managing collections of data which include not just repositories like databases, but also simpler store types such as simple files, emails. Here we are using MYSQL to store our data. Initially we use MYSQL workbench, a visual database design tool. But further proceeding in the project we find it more convenient on working with phpMyAdmin is the web application written primarily in PHP. It's used for managing MySQL database.

7.2.4 DATA ANALYSIS – API

Application Programming Interface are mechanisms that enable two software components to communicate with each other using a set of definitions and protocols. API architecture is usually explained in terms of client and server. The application sending the request is called the client, and the application sending the response is called the server. REST stands for Representational State Transfer. REST defines a set of functions like GET, PUT, DELETE, etc. that clients can use to access server data. Clients and servers exchange data using HTTP.

The main feature of REST API is statelessness. Statelessness means that servers do not save client data between requests. Client requests to the server are similar to URLs you type in your browser to visit a website. The response from the server is plain data, without the typical graphical rendering of a web page. Flask is a widely used micro web framework for creating APIs in Python. It is a simple yet powerful web framework that is designed to get started quickly and easily, with the ability to scale up to complex applications. So, flask is used as an API in this project to GET/POST data and render HTML template.

7.3 APPLICATION PROGRAM INTERFACE

Before anything we use Flask to connect with our database to ensure that the API can properly fetch the needed data and parse it for further process. Our heart and the uniqueness of the project lies in letting the user creating their custom use- cases depending upon their needs. For that we created a rough algorithm which lets the user input certain fields and prints the output as alert if it

can find it in the database. And it resulted as expected, now the hard part is to attach the frontend and make the idea work like an uninterrupted cycle.

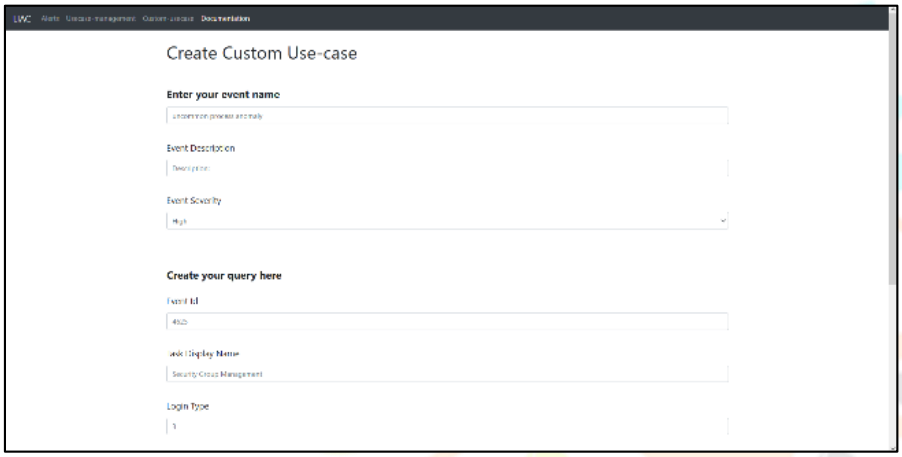
7.4 FRONTEND PROCESS

LIAC is designed to be a web interface security software, so the first and foremost thing is to get an idea of how many web pages does it going to contain and how these pages should look like. So LIAC mainly consist of 5 different pages, they are: - Home, Alert management, Use-case management, Custom use-case & Documentation.

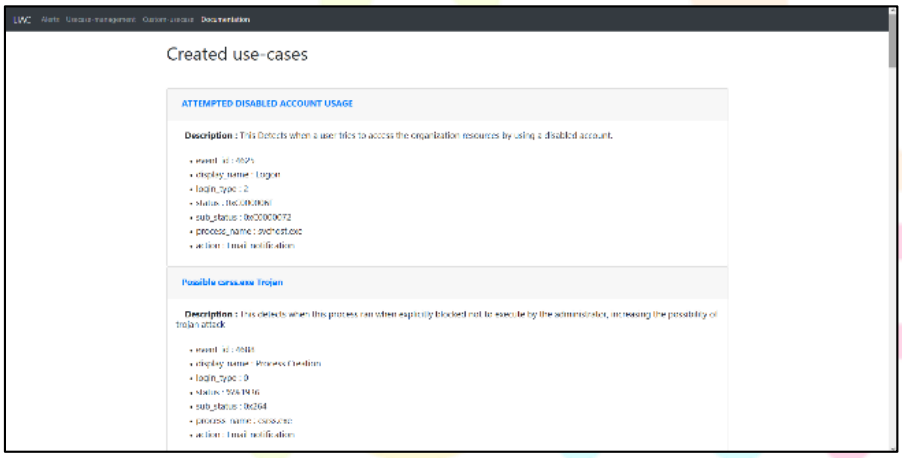
7.4.1 DEPLOY THE SOFTWARE ON LOCAL WEBHOST

HTTP is the protocol for websites. The internet uses it to interact and communicate with computers and servers. Let me give you an example of how you use it every day. When you type the name of a website in the address bar of your browser and you hit enter. What happens is that an HTTP request has been sent to a server. We will write code that will take care of the server-side processing. Our code will receive requests. It will figure out what those requests are dealing with and what they are asking. It will also figure out what response to send to the user. Localhost is the default name of the computer you are working on. The term is a pseudo name for 127.0.0.1, the IP address of the local computer. This IP address allows the machine to connect to and communicate with itself. Therefore, localhost (127.0.0.1) is used to establish an IP connection to the same device used by the end-user. Flask allows us to locally deploy our code on our local host.

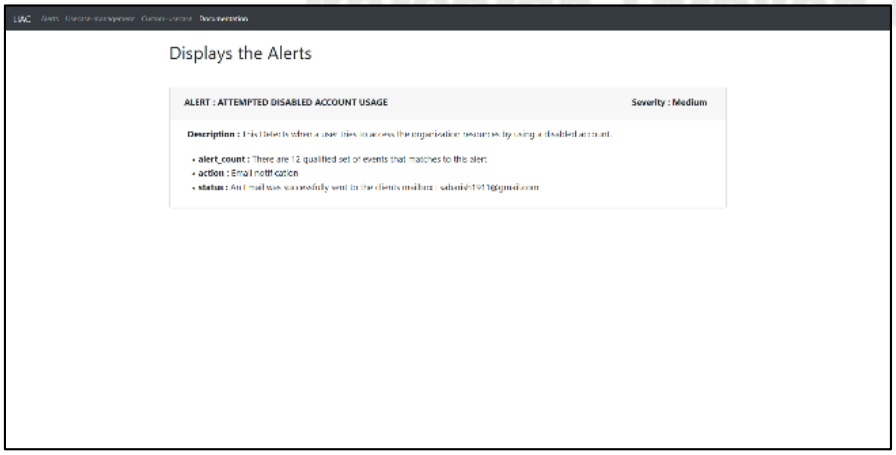
VII. RESULT & PERFORMANCE EVALUATION



(Fig 7. 1 LIAC – custom use- case page)



(Fig 7. 2 LIAC use- case management page)



(Fig 7. 3 LIAC alert page)

As you can see all the intended pages LIAC software were created and functioning as expected. So, the flow goes like ~

step 1: In the "custom use- case" page (fig 7.1) you will fill the field details that you find appropriate to create a custom use - case.
Step 2: All the use- case created by the user is displayed on the "use - case management" page (fig 7.2).
Step 3: The alert will trigger along with the action to be executed for a specific alert in the "alert" page (fig 7.3) only if any use- case activity matches activity happened in your windows machine.

VIII. CONCLUSION

Log Indexing with Alert Customization is a local web hosting software in which the software will look for possible alerts in various windows logs based on the user defined custom usecases. It primarily contains three main pages and a documentation page act as a guide which tells the user on how to use the software. They are alert page, usecase-management page and the create-usecase page. Each page is responsible for processing and displaying specific tasks. The database will be stored with different type of logs in a processable format. The create-usecase page will let the user to create a custom usecase based on the threats that user's environment. This page is highly customizable because the user is holding the raw logs and can manipulate to it extends. The usecase-management page will display all the created / built-in usecases along with the fields. Alerts will be triggered and displayed in the alert page if all the fields in the usecase matched the logs on the database. It also sends an email to the client, i.e., the user whenever an alert is triggered.

ACKNOWLEDGEMENT

We are really very thankful and we will also like to display our sincere recognition for our Head of the Department Dr S Geetha and our project guide Srilakshmi for the guidance and support in making this Research possible. I also like to thank my friend/colleague at Ernst & Young LLP, Azeeb C Rusthique who initially supported and believed my idea. Their Valuable support & guidance from initial to final level helps us to achieve to complete this research paper. Our sincere thanks to all the faculties who helped us to complete this and gave their valuable advice and made it easy to complete this.

REFERENCE

[1] Marcus Hanikat - Centralized log management for complex computer networks
[2] Log Management for the University of California: Issues and Recommendations
[3] Jan Valdman - Log File Analysis
[4] Harshal N. Kolhe & Imran R. Shaikh - Design of Secure Log Management Over Cloud
[5] Ali A. Ghorbani, Wei Lu and Mahbod Tavallaee - Alert Management and Correlation
[6] Prasanta Kumar Sahoo - Research Issues on Windows Event Log
[7] Khalid Mahmood - Utilizing a SQL Data Store for Scalable Log Analysis
[8] Vangala Rama Vyshnavi, Amit Malik - Efficient Way of Web Development Using Python and Flask.
[9] Bruce Payette Richard Siddaway - WINDOWS POWESHELL IN ACTION