**Proceedings of**
**INTERNATIONAL CONFERENCE ON COMPUTING, COMMUNICATION AND ENERGY SYSTEMS (ICCCES-16)**
**In Association with IET, UK & Sponsored by TEQIP-II**
29th -30th, Jan. 2016

Paper ID: CSEIT13

# LOG-BACKUP UTILITY: A SECURE LOG MINING SYSTEM

Samratsinh Dhumal
Komal Rajput
Akshay Patil
Pavan Phatak
Vaibhav Kadam
Bachelor of Technology, Department of IT, K.E Society'sRajarambapu Institute of Technology, Sakharale, Maharashtra
Sadanand Howal
Assistant Professor, Department of IT, K.E Society'sRajarambapu Institute of Technology, Sakharale, Maharashtra

**Abstract—The purpose of the paper is to take back-up of all the activities conducted on a war ship. These activities will be stored in form of logs in our system. Winmon is software showing graphical interface of war ship. Log Backup Utility will interact with Winmon and store the generated logs. Different activities carried out on war ship will be captured through sensors while in LAN; logs will be collected by our system and stored at single server. The collected logs will be categorized by using K-Means algorithm. Software will auto enable USB ports for authorized user. The system will be such that all the logs will be stored in a single clustered database. All the event logs will be reviewed and this information will be used for further warfare techniques. This will help for next successful voyage of the warship.**

**Keywords— back-up, logs, Winmon, LAN, sensors, K-Means, USB ports, data mining, log mining**.

## I. INTRODUCTION

Log Backup Utility is a system which takes back-up of all logs created by the Clients i.e. computers connected in the LAN and/or number of sensors in different compartments of a warship. Log Backup Utility system provides Authentication as it has only one type of user i.e. Administrator. Winmon [1] is a graphical interface which will capture these activities and display images of the compartments. The system which is proposed and the graphical interface Winmon will together retrieve the logs [5] on a single server. The collected logs will be stored in a single database. After storing the logs, K-Means [3] Clustering Algorithm will be applied to the stored logs and will get sorted. After Clustering, logs will be stored in a proper sorted database. Also the proposed system is a secure and well protected system. Initially all USB ports are disabled, but when Administrator logs-in in the Log Backup Utility System all the USB ports will be enabled and Administrator can take back-up directly on the external storage. All the USB ports will be disabled again as soon as the Administrator logs-out.

Tasks to be fulfilled:-

1. To retrieve different event logs.
2. To provide secure authentication and authorization to the administrator.
3. To store and sort those logs using clustering techniques.
4. To provide secure and easy access to the required logs.
5. To auto enable USB ports only to authorized users.

## II. LITERATURE SURVEY

The major Literature Survey was done by Larsen & Toubro [1], the company providing us internship. They provided us details for understanding the project. They first gave a short introduction of the project. Then they provided a requirement manual which gave an idea for developing the system. Then we were given a brief explanation of the project by the company's assistant manager.The warship contains many sensors [7] per compartment. There are many compartments on a single deck and the ship contains such 4-5 decks.Major problem with project is to collect this huge number of event log and further to sort out them in a specific category. The book- DATA MINING – Concepts and Techniques by Jiawei Han, Micheline Kamber, and JianPei [4] has mentioned various data mining method. We have chosen clustering technique to sort log events. The collection, sorting and storing of these logs will be done with the help of some instances from DATA MINING. The grouping of these logs will be done by using Data Clustering algorithms [2]. Also it has an efficient and well equipped algorithm to search the logs, thus it provides a well-equipped window for searching specific logs. In our project, with the help of this algorithm the logs will be categorized and stored accordingly.

By reading and analyzing IEEE conference paper by Tapas Kanugo, David M. Mount,Nathan S. Netanyahu "An Efficient k-Means Clustering Algorithm: Analysis and Implementation" [3] we got an idea of the Log Mining[6] thus it helped us for the implementation of our paper. This literature survey gives us more new and important concepts that are very helpful for our project. This literature survey gives the direction to our work.

## III. PROPOSED SYSTEM

### A. Algorithm

Algorithm for K-Means:-
Input: z: the number of clusters,
M: a data set containing n objects.
Output: A set of z clusters.

Method:
(1) Arbitrarily choose z objects from M as the initial cluster centers;
(2) Repeat
(3) (Re)assign each object to the cluster towhich the object is the most similar,based on the mean value of the objects in the cluster;
(4) Update the cluster means, i.e., calculate the mean value of the objects for each cluster;
(5) Until no change.

Algorithmic steps for K-Means:-

Input: P= {$p_1$, $p_2$ … $p_n$} (collection of logs to be categorized)

Q (numbers of categories)

maxiters ( iteration limit )

Output: L= {$l_1$, $l_2$… ly} (collection of centroid of clusters)

M={ $m_{(p)}$|P=1,2,3,4,…..,n } (collection of clusters labels P)

foreach $l_i$ЄL do

$l_i$←pj Є P(e.g. random selections )

end

foreach $p_i$ЄP do

$m_{(l_i)}$ ←argmin distance ($l_i$,$l_j$) Є p{1…...Q}

end

Changed← false;

iter← 0;

Repeat

foreach $l_i$ЄL do

UpdateCluster($l_i$);

end

foreach $l_i$ЄL do

minDist←argminDistance($l_i$,$l_j$) j Є {1.…Q};

ifminDist ≠ m($p_j$) then

m($l_i$) ←minDist;

Changed← true;

end

end

iter++;

Until Changed = true and iter ≤ MaxIters;

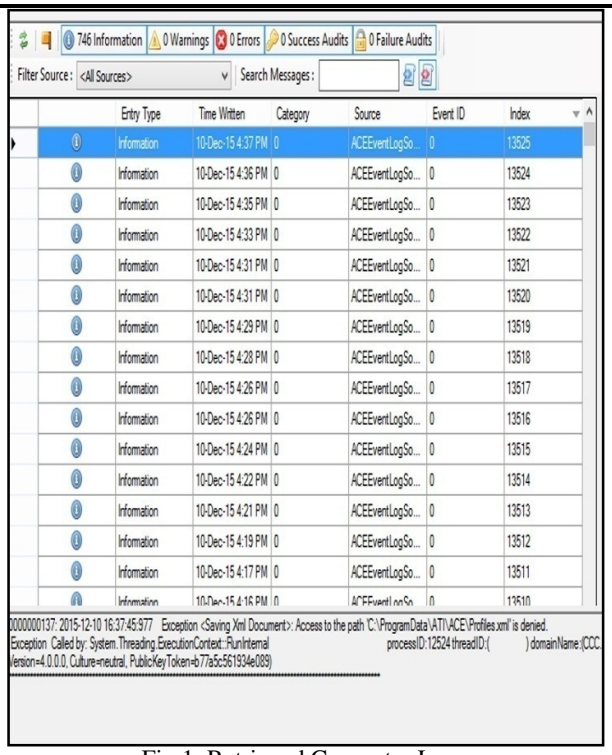


Fig 1: Retrieved Computer Logs

Figure 1 represents, the event logs retrieved from the different computers connected in LAN. The different logs retrieved are,

1. Application Logs
2. System Logs
3. Security Logs
4. Error Logs
5. Audit Logs
6. Warning Logs

B. Block Diagram

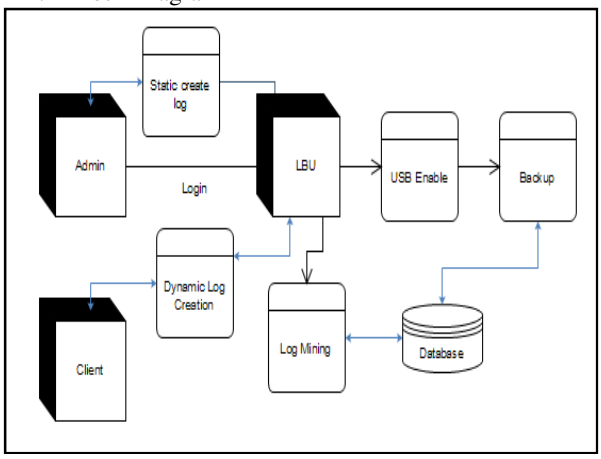


Fig 2: Modular Diagram

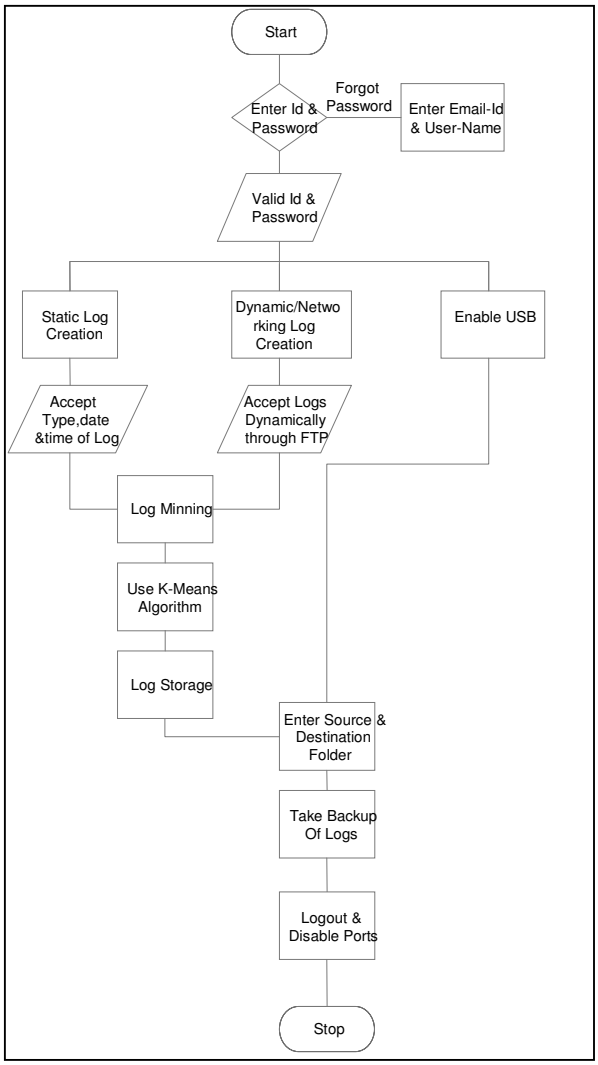


Fig 3: System Flow

In System Flow figure the algorithm i.e. step by step procedure from start to stop is represented. It is drawn with the help of the tool: Edraw Max [8]. System Flow figure is used for designing, analyzing, and handle a process in different modules of the system. It starts from start and ask user for username and password for secure authentication.

C. Tables

The table is a database table which stores the logs in it. The logs are stored according to following attributes.

Table 1: Database Table of Log Storage

| Sr.No | Attributes | Data type | Length |
|---|---|---|---|
| 1. | User | varchar | 25 |
| 2. | Level | integer | 25 |
| 3. | Source | varchar | 20 |
| 4. | event_id | integer | 20 |
| 5. | log_name | varchar | 20 |
| 6. | Opcode | varchar | 20 |
| 7. | Logged | varchar | 25 |
| 8. | task_category | varchar | 20 |
| 9. | Keyword | varchar | 15 |
| 10. | Computer | varchar | 25 |

D. Implementation Details

Log Backup Utility is a system which is useful to take back-up of all logs created by the Clients i.e. computers connected in the LAN and different sensors in the different compartments of the warship.

i. Secure Authentication:

Log Backup Utility system has Authentication and has only one type of user i.e. Administrator. Initially all USB ports are disabled, but when Administrator logs-in in the Log Backup Utility System all USB ports will be enabled and Administrator can take back-up directly on the external storage. All the USB ports will be disabled again as soon as the Administrator logs out.

ii. Static Log Creation:

As Log Backup Utility System is being developed in .NET framework, it gives better Graphical User Interface (GUI). Log Backup Utility System can allows Administrator to create logs manually by providing the simple form which contains information about the activity. These logs are termed as static logs and the process is called static log creation. These static logs will be collected on the server.

iii. Dynamic Log Backup and Clustering using K-means:

Any ongoing activity is sensed and recorded as log and stored in the database. Log Backup Utility internally uses K-Means Clustering Algorithm to sort the log. Dynamic Logs will be collected through the parallel processing and collected on the server. After collecting the logs K-Means Clustering Algorithm is applied to both the collected logs and gets separated out and sorted. After Clustering logs will be stored in the database.

## IV. TEST BED

We have tested this system using simulation of the system on the following requirements:-

- Windows Desktop PC
- Core i3 Processor

Table 2: Example of Retrieved Logs in Simulation of system

| Sr. No. | Type Of Log | Number Of Logs On a Single Machine | Total Number of Machines | Total Collected Logs |
|---|---|---|---|---|
| 1 | Application Logs | 17625 | 20 | 459608 |
| 2 | Error Logs | 125 | 20 | 1211 |
| 3 | System Logs | 2127 | 20 | 42545 |
| 4 | Audit Logs | 1894 | 20 | 37880 |
| 5 | Warning Logs | 8538 | 20 | 200670 |
| 6 | Security Logs | 7604 | 20 | 143054 |

## V. ANALYSIS

Previously, there were no such systems to record any type of logs for any warship; also if any data is needed from the previous records, we cannot retrieve that data. This system which will be introduced helps to get the logs of occurring activities on a specific warship. The warship contains many sensors per compartment. There are many compartments on a single deck and the ship contains such 4-5 decks. The sensors are used to sense all the activities occurring in a particular compartment. Winmon is a graphical interface which will capture these activities and will also display images of the compartments. Activity logs occurred in those compartments will be retrieved by using this software. These logs will be then collected on a server system and then will be sorted and stored effectively. Security is also a major factor provided by the system. This system can be accessed only by the officials. At first all the naval system's USB slots are disabled. After successful login of an official the USB slots are enabled and the data can be retrieved and after the logout the USB slots are disabled again.

After implementation of this project firstly all the logs will be stored in a single clustered database. All the event logs will be reviewed and this information will be used for further warfare techniques. This will also help for next successful voyage of the warship.

## VI. CONCLUSION &FUTURE WORK

This project helps the navy to use previous information to improve the tactics and warfare techniques. Also, as the logs are stored in a secure database, they cannot be accessed by unofficial people. Thus the security factor is maintained. Thus, it provides secure, reliable and helpful software for storing, retaining and properly using the logs of a warship. So the systems efficiently retrieve store and logs. Also it sorts all these logs using data mining methods.

## REFERENCES

[1] Larsen and Toubro "www.larsentoubro.com" Accessed 23 July, 2015.

[2] Website of Wiki Books "www.en.wikibooks.org" Accessed 17 August, 2015.

[3] IEEE conference paper by Tapas Kanugo, David M. Mount,and Nathan S. Netanyahu "An Efficient k-Means Clustering Algorithm: Analysis and Implementation".

[4] A book by Jiawei Han, MichelineKamber, JianPei: DATA MINING – Concepts and Techniques 3rd Edition.

[5] IEEE conference paper by Levy, E. Department of Comput. Sci., Texas Univ., Austin, TX, USA "Log-Driven backup: A recover scheme for large memory database system".

[6] IEEE conference paper by Jianli Duan Sch. Of Sci., Quingdao Technol. Univ., Quingdao, China "Research on web log mining analysis".

[7] IEEE conference paper by Anis KOUBAA, Mário ALVESand Eduardo TOVAR "Wireless Sensor Networks: A Technical Overview".

[8] Tool: - Edraw Max.