RESEARCH ARTICLE OPEN ACCESS

Assessment System in Newspaper Based On Text Mining

Ms.J.Shantha Lakshmi Revathy[1], N.Sai Priya[2]

[1]Asst. Professor, *Department Of Computer Science, Velammal College Of Engineering and Technology, Madurai, Tamilnadu, India*

[2]PG Student, *Department Of Computer Science, Velammal College Of Engineering and Technology, Madurai, Tamilnadu, India*

ABSTRACT:
The newspaper is a publication which includes news report, articles and advertisements. Online newspaper consist of day-to-day events and the process of handling the sensible data is a major issue. The system is designed as an application which consist of top ranking newspaper with their wikileaks information and the main goal of this phase is to protect the sensible information. The future work of this application is hosted in the cloud and mined text can easily be retrieved. The application consist of the editorial view, where the new editor should be registered and they can upload or delete the sensible news only after they receive an OTP which is a six digit number is generated by SMTP protocol in their mail. This OTP can thus be used to ensure authentication and the TLS V1.0 used in the SMTP protocol provides a greater security and efficiency in generating the OTP. The AES cryptographic algorithm is used to provide security to the sensible data from the hackers. The Admin and the user log will automatically generate into jar format, hence the user entry time, exit time and OTP will be updated in the database. A plug in is created which provides an assessment on the number of views made for a single news by the double click operation. This assessment is represented as bar charts and hence the sensible data status can be regularly verified.

Index Terms: One Time Password (OTP), Simple Mail Transfer Protocol(SMTP), Advanced Encryption Standard (AES) ,Transport Layer Security(TLS)

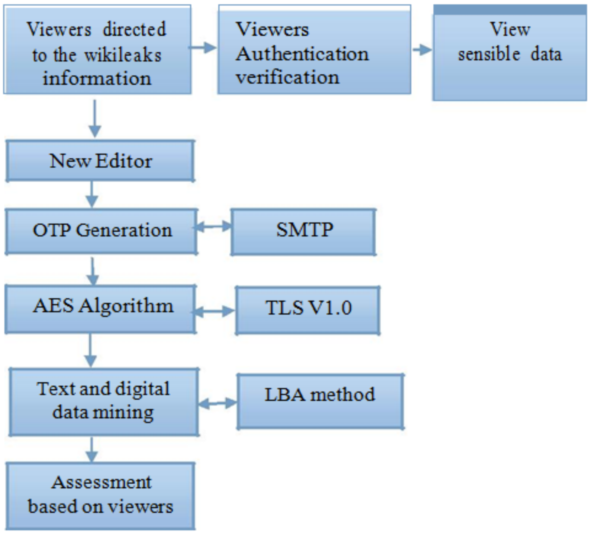
I. INTRODUCTION

An online newspaper is the online version of a newspaper, either as a stand-alone publication or as the online version of a printed periodical. Going online created more opportunities for newspapers, such as competing with broadcast journalism in presenting breaking news in a more timely manner. The credibility and strong brand recognition of well-established newspapers, and the close relationships they have with advertisers, are also seen by many in the newspaper industry as strengthening their chances of survival. The movement away from the printing process can also help decrease costs. Wiki leaks is an organization that facilitates the anonymous leaking of secret information through its website. Wiki Leaks is most commonly known for all of the negative effects that it has had. However, there are many benefits to having sensitive information about corruption or potentially embarrassing information exposed to the public. The main objective of this project is to provide authentication to sensible data. The editor side enhances authentication and it provides the system from the hacker. The editor side provides authentication by disabling the right click where there may be a chance for the hacker to change over the existing news. Thus the changes in data is only possible for the authenticated user who have already registered.

II. NEED OF SYSTEM

The Wikileaks information should ensure that only users with limited IP address can access the data. This system also enhances the status of each user and provides sensible data based on their IP address. The latest version of TLS 1.0 is used which is more secure than SSL version 3.0. This version of TLS enhances higher security and reliability when compare with the 3.0 version of SSL. This system proposes an enhanced security using the AES algorithm over the sensible data. Viewers after directed to the wiki leaks information are about to view the sensible data and only the authenticated users who have received the OTP can upload the sensible information. The protocol for mail submission is the same, but uses port 587. SMTP connections secured by SSL, known as SMTPS, default to port 465 (nonstandard, but sometimes used for legacy reasons).Although electronic mail servers and other mail transfer agents use SMTP to send and receive mail messages, user-level client mail applications typically use SMTP only for sending messages to a mail server for relaying.

III. SYSTEM ARCHITECTURE


Viewers directed to the wikileaks information
Viewers Authentication verification
View sensible data
New Editor
OTP Generation
SMTP
AES Algorithm
TLS V1.0
Text and digital data mining
LBA method
Assessment based on viewers

IV. MODULE DESCRIPTION

Modules

1. Log Generators
2. Logging Client or Logging Relay
3. Logging Cloud
4. Log Monitor

Log Generators

These are the computing devices that generate log data. Each organization have to adopt the cloud-based log management service that has a number of log generators. Each of these generators is up to with logging capability. The log files generated by these hosts are not stored locally rather they are stored temporarily till the time they are pushed to the logging client.

Logging Client or Logging Relay

The logging client is a collector that receives groups of log records generated by one or more log generators, and prepares the log data so that it can be pushed to the cloud for long term storage. The log data is transferred from the generators to the client in batches, either on a schedule, or as and when needed depending on the amount of log data waiting to be transferred. The logging client incorporates security protection on batches of accumulated log data and pushes each batch to the logging cloud. When the logging client pushes log data to the cloud it acts as a logging relay. The terms logging client and logging relay interchangeably. The logging client or relay can be implemented as a group of collaborating hosts. For simplicity however,there is a single logging client.

Logging Cloud

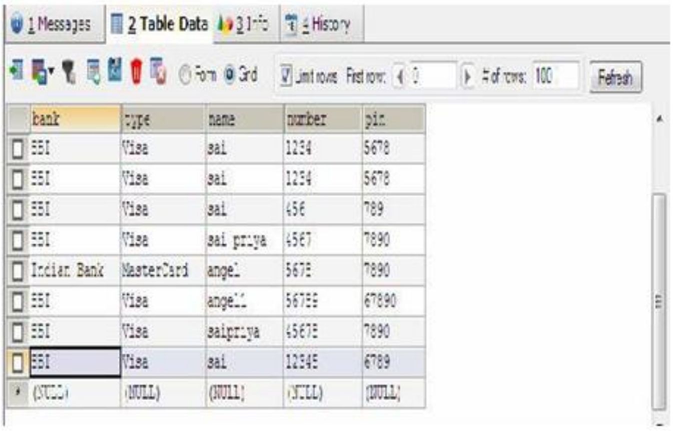
The logging cloud provides long term storage and maintenance service to log data received from different logging clients belonging to different organizations. The logging cloud is maintained by a cloud service provider. Only those organizations that have subscribed to the logging cloud’s services can upload data to the cloud. The cloud, on request from an organization can also delete log data and perform log rotation. Before the logging cloud will delete or rotate log data it needs a proof from the requester that the latter is authorized to make such a request. The logging client generates such a proof. However, the proof can be given by the logging client to any entity that it wants to authorize.

Log Monitor

These are hosts that are used to monitor and review log data. They can generate queries to retrieve log data from the cloud. Based on the log data retrieved, these monitors will perform further analysis as needed. They can also ask the log cloud to delete log data permanently, or rotate logs.

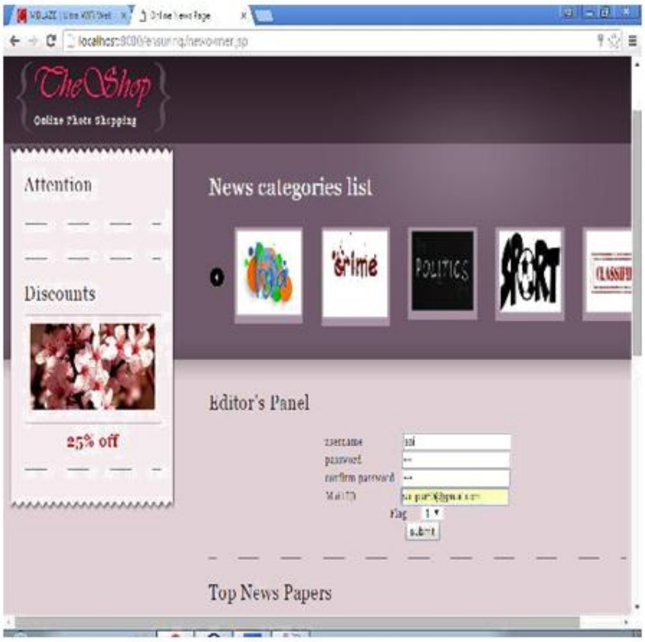
V. RESULTS AND DISCUSSION

The wikileaks information that is to be published in the newspaper is highly sensible and hence they should be less in size. Thus the wikileaks information is mined based on the text mining concept and only the relevant and the important information alone is displayed in the future, the plan to refine the log client implementation so that it is tightly integrated with the OS to replace current log process. In addition, to address privacy concerns current implementation allows access to log records that are indirectly identified by upload-tag values. The plan to investigate practical homomorphic encryption schemes that will allow encryption of log records in such a way that the logging cloud can execute some queries on the encrypted logs without breaching confidentiality or privacy. Thus based on this method the number of viewers who have read the wikileaks information can de assessed and provides a greater way of security and reliability.A number of OTP systems also aim to ensure that a session cannot easily be intercepted or impersonated without knowledge of unpredictable data created during the previous session, thus reducing the attack surface further.The viewers those who view the sensible data is assessed based on the doble click operation. A plugin is created in the jscript which is used to count on the number of clicks made on each sensible information. A bar chart is generated which is used to provide the number of views and since the right click operation is not allowed it is difficult for the hacker to easily retrieve the sensible data.



bank	type	name	number	pin
SBI	Visa	sai	1234	5678
SBI	Visa	sai	1234	5678
SBI	Visa	sai	456	789
SBI	Visa	sai priya	4567	7890
Indian Bank	MasterCard	angel	5678	7890
SBI	Visa	angel1	56789	67890
SBI	Visa	saipriya	45678	7890
SBI	Visa	sai	12345	6789
(NULL)	(NULL)	(NULL)	(NULL)	(NULL)

Fig 1: Database maintains the uploaded information of the editor



Editor's Panel

username sai
password
confirm password
Mail ID [illegible]
Flag 1
submit

Fig 2: Editor Registration

The editor who wants to upload the wiki leaks information should be registered with a valid mail ID

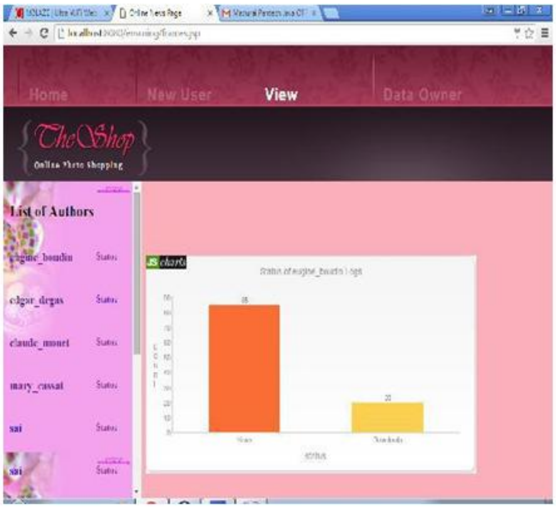


Fig 3: Bar chart representing the number of views made Hence an assessment is made by a bar chart generating the number of views made on each and every data

VI. CONCLUSION

Thus the sensible data is being secured from the unauthenticated user using AES algorithm. Thus AES algorithm can be implemented which is used for log monitor and log generator. The proposal of a comprehensive scheme that addresses security and integrity issues not just during the log generation phase, but also during other stages in the log management process, including log collection, transmission, storage and retrieval. The hacking can thus easily be determined based on the viewer entry and exit time.This system also provides only authenticated editors to enter into the system and to upload the wikileaks information. The status of sensible data can also be identified based on how many users have read the data and this status is provided as a bar chart.Thus sensible data can be uploaded only by the authenticated users and hence misusing the sensible data is highly impossible. The OTP which is generated by the TLS V1.0 by the SMTP protocol is used which provides greater security and reliability. Thus the process of editing the wikileaks information can be performed only by the authenticated users an d the sensible data cannot be hacked by the viewers.

REFERENCES

[1]. Soon Dong Park, Joong Chae Na, Young-Hwan Kim, dong Kyue Kim, "Efficient OTP(One Time Password) Generation using AES-based MAC," Journal of Korea Multimedia Society, vol. 11, No. 6, pp. 845-851, June. 2008.

[2]. IETF RFC 4226, HOTP: An HMAC-Based One-Time Password Algorithm, Dec. 2005.

[3]. Federal Information Processing Standards (FIPS), "Announcing the Advanced Encryption Standard (AES)," National Institute of Standards and Technology (NIST), November 2001.

[4]. J Bonneau, "Cache-Collision Timing Attacks Against AES", CHES 2006, 6th International Workshop, Yokohama, Japan, Oct. 2006.

[5]. V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-Based Encryption for Fine-Grained Access Control of Encrypted Data", Proc of 13th ACM Conference on Computer and Communications Security, pp. 89-98, ACM, 2006.

[6]. S. McGovern, "Information Security Requirements for a Coalition Wide Area Network", Thesis in Naval Postgraduate School, Monterey, California (June 2001), http://cisr.nps.edu/downloads/theses/01thesis-mcgovern.pdf, NPS/CISR, 2001, (retrieved 23.4.2013).

[7]. PCI Security Standards Council. (2006, Sep.) Payment Card Industry (PCI) Data Security Standard—Security Audit Procedures Version 1.1 [Online]. Available: https://www.pcisecuritystandards.org/pdfs/pci−audit−procedur es−v1-1.pdf

[8]. Hua Li and Jianzhou Li, "A High Performance Sub-Pipelined Architecture for AES", IEEE International Conference on Computer Design: VLSI in Computers and Processors 2005 (ICCD 2005), pp. 491-496, Oct. 2005.

[9]. K. Kent and M. Souppaya. (1992). Guide to Computer Security Log Management, NIST Special Publication 800-92 [Online]. Available: http://csrc.nist.gov/publications/nistpubs/800-92/SP800-92.pdf

[10]. Richa Garg and Renu Vig, "An Efficient Montgomery Multiplication Algorithm and RSA Cryptographic Processor", IEEE International Conference on Computational Intelligence and Multimedia Applications, vol. 2, pp. 188 - 195, Dec. 2007.

[11]. T. P. Pedersen "Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing", Advances in Cryptology -CRYPTO, vol. 576, pp.129 -140 1991

[12]. M.K. Lee, J.K. Min, S.H. Kang, S.H. Chung, H. Kim and D.K. Kim, "Efficient Implementation of Pseudorandom Function for Electronic Seal Protection Protocols," LNCS, Vol. 4298, pp. 173-186 Springer, 2007.

[13]. IETF RFC 4226, HOTP: An HMAC-Based One-Time Password Algorithm, Dec. 2005.

[14]. Bruce Schneier, "SHA-1 broken," Feb. 2005, Available : http://schneier.com/blog/archives/2005/02/shal-broken.html

[15]. J. Golic, and M. Salmasizadeh and E. Dawson, Fast Correlation Attacks on the Summation Generator," Journal of Cryptology, Vol. 13, No. 2, pp.245-262, 2000.