INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 2)
Available online at: www.ijariit.com

Framework for enhancing privacy of files in web using random key generation technique

Anjana K.
anjanakarunakaran8@gmail.com
Rajalakshmi Engineering College, Chennai, Tamil Nadu

Kumar P.
hod.cse@rajalakshmi.edu.in
Rajalakshmi Engineering College, Chennai, Tamil Nadu

ABSTRACT

In general online websites contain millions of articles related to different domains. The major flaw found in these websites is the absence of security for files uploaded without any protection. It is made in such a way as it is simply available for any users to change the content of these files. Thus, to overcome the above drawback we propose the system "Framework for Enhancing Privacy of the Files in Web Using Random Key Generation Technique" which makes all the files protected using key generation and encryption techniques so that no other invalid user is allowed to access the file. Even the authorised user must request to the admin to get the key for the particular file from the database to access it. Thus the system provides secured data access with original information uploaded by the data owner, it also avoids the duplication of the same content from different users of the web.

Keywords*—*AES, Cloud, Key Generation

1. INTRODUCTION

Nowadays the use of information available in online web such as Wikipedia, Google is trending to mine the information from the pages in the web for the user's knowledge. The major drawback in the present system is the absence of high security to files stored in the database of the websites. In order to resolve we proposed a model "Framework for Enhancing Privacy of the Files in Web Using Random Key Generation Technique" was the system protects all the files uploaded by generating individual keys for each file in a database of the website. If any user needs to edit or download or upload the content need to register and send a request to the admin to get the keys for the file and then they can access the file content with the permission. We also have a facility where the admin can create a set of offensive or unwanted words list which should not be added in the document at the time of uploading the file so that it reduces the illegal or unwanted information in the content.

Cloud computing provides the power to deal with big data on the various domain where the data owners are ready to outsource their data on the web with privacy policies for the content. The most common privacy available for large datasets is by using encryption and decryption techniques for data before outsourcing.

2. LITERATURE SURVEY

In general, the accessibility to the files or content is done through the query access were the ranking is done through the query process. The general issue is the query access for encrypted data which is done using index-based tree traversal with multi-keyword. [1]

The relevancy of data retrieved for the searched keyword issues in major drawback which directly attains in the cloud which results in traffic in retrieving unwanted files which are not related to the issues searched. Thus these issues are solved by the searchable secure encryption (SSE) [2] and open secure encryption (OSE) [3] which enhances the search through the encrypted data with the keyword. But the only option for the search is Boolean search which consumes much time to get only the related files from the list.

The cloud consists of many primitive and spatial database which are encrypted for privacy. The k-NN is a scheme which is used were the domain of classification, ordering and clustering are involved in the search. The multiple keys for access and privacy of the content are given to both user and data owner which causes and impact of access to content without any rules causes uneven authorisation to files [4][5].

The different keys for a single file cause mapping issues which can cause confusion in the retrieval of the files from the cloud for the users. The mapping function is solved through the index matching and tree-based search without giving any access to the key of the file thus by avoiding the unwanted key access to the users of the file and by having a reasonable search time for the file. The workflow for searching the file using the keyword is one of the issues faced during access of the file. Another main issue is maintaining the key credentials of each file which enhances the security of file [6] [7].

The data protection is one of the most vital areas of cloud were it acquires a set of techniques followed to protect data. The Data protection is the data protection one of the most vital areas of cloud were it acquires a set of techniques followed to protect data which is commonly called as secure socket scheme (SSS) where it includes all sort of protection schemes. The data must be protected for an authorised user to access only going through secured access [8] [10].

The extraction of the file is also based on extraction algorithm which is been used as an iterative and non-iterative learning format for searching the files from the database. The system also varies according to the location of the information using a supervised learning algorithm [11]. The system also carries the ranked search through the keywords were the files are arranged according to the keyword which is frequently searched using TF-IDF algorithm. The system also uses encryption for fire protection were the access of the file is restricted for the users of web [12] [13].

The data available is also handled using iterative graph were the details are summarised through a graph with the details of the keyword and content of each file. The system also uses all sort of iterative indexing using the ranking algorithm [14]. The effectiveness of the system is also done through the most cited files used by the users of the site. The flow of searching and privacy of the files are maintained through the encryption algorithm available were the key generation is also done through the number generator algorithm for having an isolated content without the unauthorised user to access the file content from cloud without the permission of the data owner who posted the file to the cloud for outsourcing of the content [15].

Thus from all the above survey issues, we have proposed work to enhance the security of content in the web without any changes or offensive words added to the content on the web.

3. PROPOSED WORK

Many user-centered platforms are available for exploring knowledge all around the world. The main issue is the security for the contents in the web is not secured completely which resulted in unauthorized access and changes of the original content without the knowledge of the data owner who uploaded the content. The main motive of the proposed system is to provide security to the content on the web.

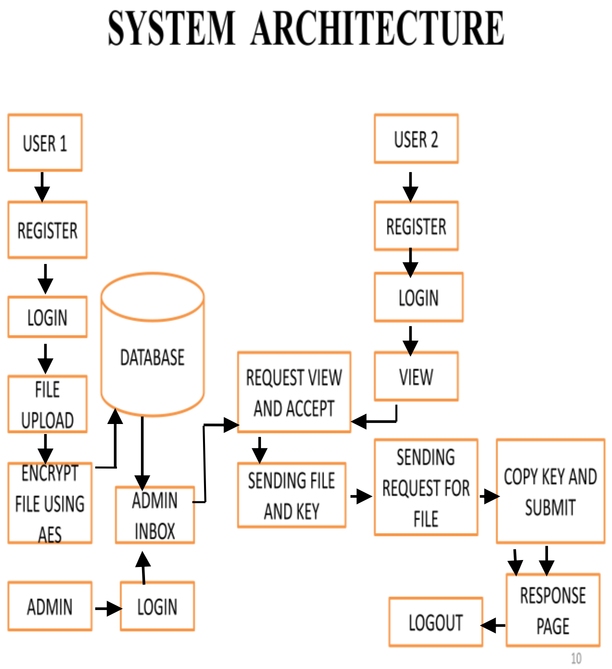

Fig. 1: System architecture

Figure 1 shows the entire system flow architecture of the proposed work. The proposed work consists of two actors as one is the user and another is the admin. The user1 will register and login into the system once they login they upload the file they need with a random key generation technique were a key length of five characters will be obtained randomly for the identity of the file. This file and the corresponding key will store in the database. The next step is another user who needs the file will log in from their account and gives a request to admin for getting the key from the admin. Now the admin will log in from their account and will see all the request from a different user for different files and once the admin accepts the corresponding key will be sent to the user who has requested for the file. Now the user can download the file for their need. Thus our proposed system enhances the security of the contents in the web using the key generated for each file which made it authorized for the users to access.

The proposed system consists of different actions which are designed within five modules for execution of the entire system. The five modules are connected with one another for a flow in the system. The five modules are as follows,

- User Interface Design
- Login and File Upload
- Blocked Words
- Request/Response File
- View/Read File

3.1 Module description

3.1.1 User interface design: The user interface design is the first module was the user1 has to register with all the details of the user. In this module, the registration process is done by creating a user id and password which will be used by the user for the entire process.

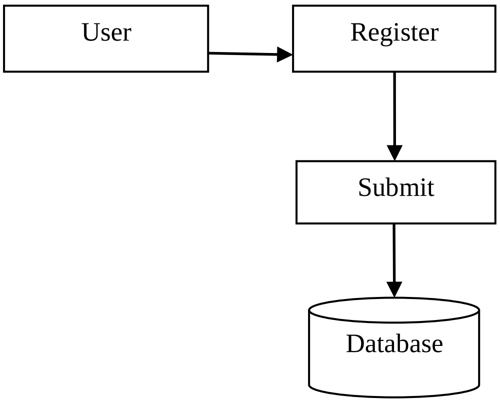

Fig. 2: User interface design

3.1.2 Login and file upload: The next module is the login and file upload module was the next step after registration is to login with the user id and password which enters the next page for file upload and search access. In this module, the Random Key Generation algorithm is used to create a random key for each file which is stored in the database. The file content is been encrypted using Advance Encryption Standard algorithm were the whole content is been encrypted and stored for more security.

In this module, the two major techniques are used for providing more security to the content on the web. The other accessibility in the system is a search option which is made common for both files and technology in the web.

In this system there are two technologies used were one is done for the fire protection and accessibility and another one is done for the content security in the file. The key is generated using the Random Key Generation Technique were the random keys are generated for the files which are uploaded in the web. The key generation is done randomly for each file where the keys will not be repeated for any of the files. The next technology used is the Advanced Encryption Standard algorithm were the system encrypts the entire content is been encrypted and made inaccessible for the user.

This module has many other facilities as a search box for various options which can be used for searching files and technology related files from the web without searching in general. This makes the system to ensure more security as well as good performance to the system. The system also has search index were a various group of the category of files are searched using this option without any common search for the files.

The entire system works on the flow of uploading and accessing the files on the web. The technology used in this module helps the system to work as a flow where the key generation and encryption process is done without any flaws. These two major technology increases the security and the accessibility to the content on the web.

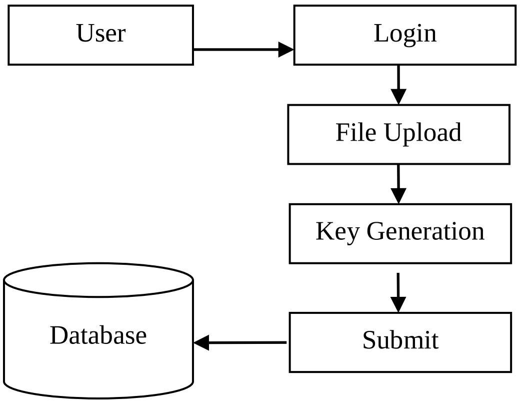

Fig. 3: Login and file upload

3.1.3 Blocked Words: In this module, the entire process is done by the admin were it can be accessed only by the admin of the website for safe and proper access. In this module the admin is allowed to add unwanted words or phrases to the website which makes an action of checking the content of file before uploading in web and matches each of the word in the file content along with the words in the blocked list if in case any of the words in the block list matches with the fie content then an error message will be shown to the user who uploads and blocks the user from posting the file until the particular word is taken out of the content.

This improves the quality of the content on the web without using any unwanted or offensive words in the content of the web. The system improves the welfare of the users of the web with quality content without any unwanted content added to the file.

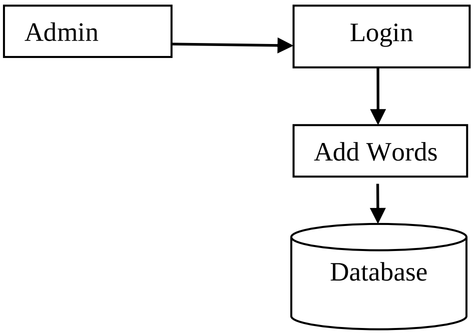

Fig. 4: Blocked words

3.1.4 Request/Response for file: In this module, the entire process is between the users and the admin of the webpage. In this module, the entire system works for the security of the file on websites. Initially, the user who wants a file from the website will log in and search for the file from the available files, once they get the file the have an option of giving request to the admin for getting the key for the particular file and they submit their request from their login. The next step is the admin part where the admin gets a notification from the user who has requested for the file. If once the admin accepts the request the key for that file requested will be sent to the user login. Now the user will get the key from database in their profile and they can now

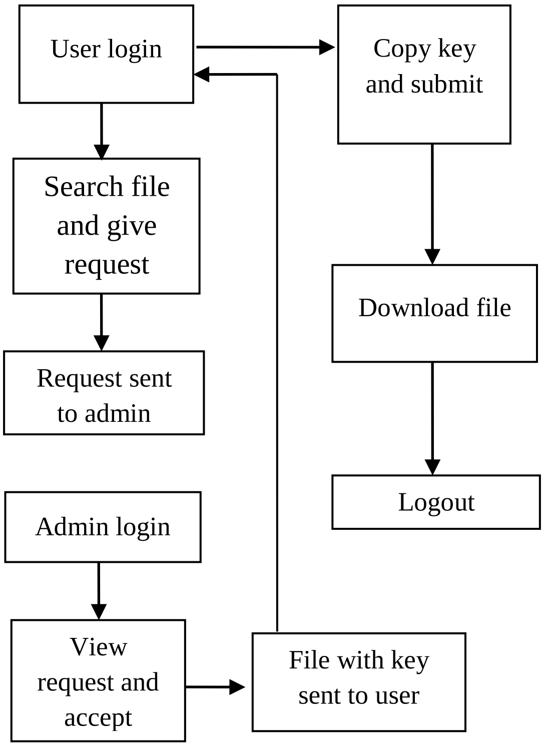

Fig. 5: Request / response for file

Access the file using the key by downloading the file for any changes to be made or for any other access. Thus this process will secure the content with the help of key and gives access only to the authorized user of the website

3.1.5 View/Read file: In this module, the final action is done where the user is able to access the file they requested for using the key for the file from the admin. In this module, the user who requested for the file will now be able to access the file using the key where they will copy the key and paste in the search box once they submit the key the file will be available for the user. The user must download the file for their access and use it. This the final step of action followed by the user. The system has well-secured access to all the files on the web with a key following. The system also has an option of downloading the content without accessing the file in the data owner account. In our system, the user cannot modify any of the content in the file without downloading the file from the web. Thus editing or adding any unwanted content from others profile is not possible.

These are the modules of the entire system which are followed in a sequence for the entire system to work. The proposed work generally gives high security to the content of the web than the existing system. All these modules are interlinked with each other were one module is the continuation of the other in execution. Thus the proposed work made secured access to the files in websites with the key generation technique and encryption standard. The system works on high security with the content access with all the users of the websites.

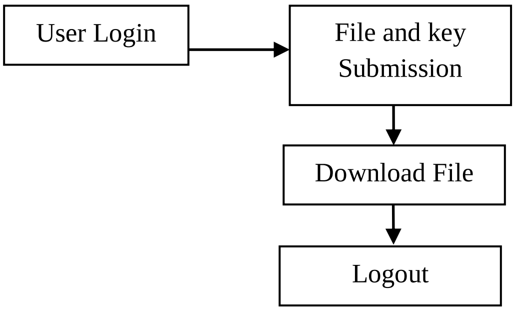

Fig. 6: View/read file

4. EXPERIMENTAL RESULTS

The experimental results for the proposed model are based on the software and hardware resources which are used for the system execution. The hardware and the software resources used in this system is latest and has many techniques used for execution.

4.1 Hardware resources

The hardware resources used in our proposed system helps in functioning of the entire process which contains WINDOWS7 operating system with a dual-core processor. It needs a minimum of 4GB RAM and 250GB Hard Disk for the working of the system. Since the proposed work is done on the web there are no such requirements for the system.

4.2 Software resources

The software resources used in our proposed system helps in functioning of the entire process which runs in the background has JAVA in the front end which comprise of the J2EE servlet for the function. The backend of the system works with MY SQL which gives the query based database for the system. The IDE used for the system is Eclipse which serves as the platform for the functioning.

These are the software and hardware resources used in the proposed model. The system has a setup where the working is done using the internet and chrome access. The resources are used with interlinked access and the system works in a better flow and gives security to the system.

4.3 System input

The input for the proposed system is the files or content uploaded to the web. The next input is the action of giving request to the admin for requesting the key for the particular file. The request is been sent by the user for the file to the admin of the webpage. These data will be stored in the database of the webpage.

4.4 System output

The system output is shown in step by step flow process were each action of the system is been shown as a step of the system. The steps are executed as a sequence for a more visual explanation.

5. CONCLUSION

The web services are highly demanded in the present system where the people are highly depending on the information which are on the websites. The web services are more sensitive for public usage were the quality of the information is more important for the users. One of the major challenges in web service and cloud storage is the security concern were in the existing system the security for the content in the web is less were easily the users are allowed to access the file and modify the content they need without any term of permission from the data owner. This major flaw is been addressed and solved by our proposed system were the files stored on the web will be protected using a key which will identify each file. The key is been generated using random key generation technique were the keys for each file is been randomly generated and it is very difficult for the user to judge the key. The system also has an encryption standard was the content of each file is been encrypted and stored in the database for more security. Thus the security of the content in the web is been enhanced through the proposed work and its performance in security is highly efficient than the existing system. In future, the security can be developed more using various other security algorithms and techniques in the present system.

6. REFERENCES

[1] Xiao Feng Ding, Peng Liu and Hai Jin (2017) “Privacy-Preserving Multi-keyword Top-*k* Similarity Search Over Encrypted Data” IEEE Transactions on Dependable and Secure Computing.

[2] Cong Wang, Ning Cao, Jin Li, Kui Ren, and WenjingLou (2017) “Secure Ranked Keyword Search over Encrypted Cloud Data” IEEE Cloud Computing Service.

[3] Ning Cao, Cong Wang, Ming Li, Kui Ren, and Wenjing Lou (2017) “Privacy-Preserving Multi-keyword Ranked Search over Encrypted Cloud Data” IEEE Cloud Computing Security.

[4] Ke Cheng, Liangmin Wang, Yulong Shen, Hua Wang, Yongzhi Wang, Xiaohong Jiang, and Hong Zhong (2017)“Secure *k*-NN Query on Encrypted Cloud Data with Multiple Keys” IEEE Transaction on Cloud Computing.

[5] Jingyun Wang, Brendan Flanagan, Hiroaki Ogata(2017) “Semi-automatic construction of ontology based on data mining technique” 6th IIAI International Congress on Advanced Applied Informatics.

[6] Tomoaki Urata, Akira Maeda (2017) “An Entity Disambiguation Approach Based on Wikipedia for Entity Linking in Microblogs” 6th IIAI International Congress on Advanced Applied Informatics.

[7] Tao Jiang, Hongzhi Yu, Xiangzhen He, Xianghe Meng (2017) “Mining Tibetan-Chinese Bilingual Entities from Wikipedia” IEEE conference.

[8] K.D.C.G. Kapugama, S.A.S. Lorensuhewa, M.A.L. Kalyan(2016) “ Enhancing Wikipedia Search Results Using Text Mining” International Conference on Advances in ICT for Emerging Regions.

[9] Masato Tokuhisa, Yuuki Ishihara, Shuuhei Kimura (2016) “Recommending Paragraphs of Wikipedia Pages as a Travel Guide” IEEE 9th International Workshop on Computational Intelligence and Applications.

[10] Kire Jakimoski (2016) “Security Techniques for Data Protection in Cloud Computing” International Journal of Grid and Distributed Computing.

[11] Wengen Li, Jiabao Zhao (2016) “Text Rank algorithm by exploiting Wikipedia for short text keywords extraction” 3rd International Conference on Information Science and Control Engineering.

[12] Muhidin Mohamed, MouradOussalah (2016) “An Iterative Graph-based Generic Single and Multi-Document Summarization Approach using Semantic Role Labeling and Wikipedia Concepts” IEEE Second

International Conference on Big Data Computing Service and Applications.

[13] KeitaTsuji (2016) “Books Cited in Wikipedia: Possibility to Use their Nippon Decimal Classification Categories for Book Recommendation”5th IIAI International Congress on Advanced Applied Informatics.

[14] Sruthi V, Surekha Mariam Varghese (2015) “Secure Multi-Keyword Top-K Retrieval Over Encrypted Cloud Data Using Holomorphic Encryption” *IOSR* Journal of Computer Engineering.