OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

A THOROUGH STUDY ON SEARCH OVER ENCRYPTED DATA AND DATA SECURITY IN THE CLOUD

Shivpriya Mahamuni[1], Prof. H. P. Channe[2]
PG Student, Dept. of Computer Engineering, Pune Institute Computer Technology, Pune[1]
Assistant Professor, Dept. of Computer Engineering, Pune Institute Computer Technology, Pune[2]
mahamunishivpriya@gmail.com[1], hpchanne@pict.edu[2]

Abstract: There has been immense research performed in the area of cloud and cloud computing. The cloud platform has been known the world over for its ease of use and simplicity. The cloud can be utilized for the purpose of storage of any kind and also is capable of proving a very powerful computational power at a relatively low price. Therefore, there has been increased interest in this platform which has been attributed to its growth in recent years. The cloud technology helps reduce the dependency of the user on a specific fixed set of hardware and storage and move it over to the cloud, which can provide ubiquitous storage capability without the need of maintaining and scaling up the hardware. This has to lead to a lot of individuals and organization including the healthcare facilities, to migrate to the cloud platform. This introduces a new concern of security, therefore, all the data uploaded on the cloud is encrypted, and this encrypted data is impossible to search and reduces the Quality of Experience. Therefore, this paper utilizes the reverse circle encryption technique to achieve efficient search over the encrypted data.

Keywords– *Data Security, Search over encrypted data, Cloud Computing*

I INTRODUCTION

The cloud computing platform has been one of the fastest growing technologies in the world. It has been on the rise since the past couple of decades. There has been an increased amount of interest due to the fact that it offers a greater degree of convenience. This is also part of the reason why a lot of organizations and individuals are interested and are readily adopting this new standard. The cloud allows the owner to restructure their processes so that the control and maintenance of their data handed over to the cloud administrators.

There are a lot of activities that can be done on the cloud, which is basically just a server with almost unlimited potential, and as a user it does a lot more than just store the data, the user can access it from anywhere in the world, with any device that can connect to the internet, which is a lot more convenient than the local storage. There is a lot of data handling and processing that goes on in the background to provide such a service to the user. The cloud is a very beneficial and innovative service.

Apart from the individual users, there are also organizations that benefit from this platform and are doing so in increasing numbers. There are very enticing opportunities for the business sector as there are various innovative ways the cloud has transformed the landscape, such as SaaS or software as service which allows accessing any application that is being stored onto the cloud which saves a lot of time and resources. As the application does not need to be installed on the client machine, it is not limited by its spec but this also ensures that the client side would not be put under any type of excessive computational load.

Then there is PaaS which stands for Platform as a service, this type of cloud services are useful for companies as it is used to allow the employees to develop and implement various applications on the cloud without any slowdowns or crashes of any kind on their personal machine that would cost them a lot of money and resources.

And ultimately, IaaS which is Infrastructure as Service, this is where the companies which are big with numerous branches and employees require a platform for their implementation and various other services and their integration. This is a very useful feature that has been deployed by major companies to provide effective and good connection and is therefore used primarily for streaming purposes.

The cloud is one of the most essential aspects of the internet age is very useful in its applications. It makes it highly convenient for the user to be able to access their storage from anywhere in the world and not worry about being limited to the local storage in your device. It is also very resilient to crashes and corruption which is highly likely to happen on local storage, this would mean losing a lot of data, which is very unlikely to happen on the data stored on the cloud platform.

There is one downside to this practice that is security. When any data is uploaded on the cloud, the user hands over the management and the responsibility to safeguard the data onto the cloud provider. Therefore, it is advisable to only upload the private and sensitive data on to a trusted cloud platform. It is also preferable to encrypt the data to safeguard it against unauthorized access and provide some sort of security in the event of a breach. Encryption has been always the go-to preference for maintaining the security of the data, due to the fact that even if an unauthorized person gains access to the data, it would be scrambled and of no use to the attacker.

The encryption is one of the most efficient systems to safeguard the data but this is also not free from drawbacks, as encryption makes the data unreadable even for the system. This makes it highly difficult to index and maintain the data for search purposes. This defeats the purpose and renders a few of the key points enabling convenience and ease of use that is associated with the cloud to be of no use. As the unique encryption cannot be cracked without the specific key with the owner, it becomes impossible to be able to search the said data on the cloud.

There are a few algorithms that are used to execute a search on the encrypted data, such as keeping the keywords as plaintext, and using them to achieve indexing and ultimately to perform a search. This is a very unique approach and is very ingenious to keep the data safe and the keywords to be able to perform the search. But this too is not a completely safe approach as the plaintext keywords can be identified by an intruder.

II LITERATURE SURVEY

K. Strang et al. [1] explains that there has been an exponential growth in the area of Big Data as an upcoming and popular technology. Due to the immense growth in this industry, there has been a growing concern about the security aspect of this technology. But even considering the security risks that this platform pertains, there has been significantly less amount of research done on the security aspect of this technology. Therefore, the authors concluded in the survey that there has to be a lot more research done for security in the Big Data Environment as there has been growing concern in the community.

P. Sreekumari et al. [2] elaborates on a system designed by the authors to provide an efficient and privacy-preserving technique for the search and retrieval of data from an encrypted cloud. The authors develop a fuzzy system for the searching process to retrieve sensitive information without any lapse in security. One of the main drawbacks of this system is the lack of important parameters that define the state of the data in the cloud, such as verifiability, security, and efficiency.

S. Lighari et al. [3] proposes a novel approach for the analysis and management of security in a Big Data environment. As big corporations generate a lot of data in the form of log files, Pcap files, DNS logs, etc, this is usually stored in the data warehouse. As this large amount of data is generated every day, it accumulates in the storage and turns massive in size. This large amount of data cannot be processed, which will be a large security risk for the corporation, therefore, the authors design an approach based on apache spark for the analysis and management of this data.

Revathy P et al. [4] state that due to the explosion of the cloud platform, there has been an exponential urge for big corporations to store massive amounts of their data on the cloud. As there has been an increase in the amount of data being produced every day, Big Data has been growing traction as the data can be analyzed for unlocking great insights from that data. Hadoop is the application of choice for processing Big Data in Realtime scenarios. The authors have also stated that the big corporation has been utilizing the Hadoop Framework extensively.

S. Wang et al. [5] states that there has been immense growth in the sector of big data and Data mining for a couple of years. The authors propose a technique for the utilization of big location data as means for providing detection of unlawful activities and personalized advertisements, due to the fact that most of the people hang out in groups at a particular location together. The authors, therefore, developed a technique for the analysis of big location data and preserving the privacy of the user at the same time.

P. Pandiaraja et al. [6] elaborates that the authors have encountered some inconsistencies while performing searches on the cloud as most of the data that is being stored on the cloud is sensitive data. This data needs to be safeguarded and in case of a public cloud, the users utilize encryption to keep the data out of the hands of malicious users. Therefore, the authors have presented a technique that allows for searching over the encrypted data without any loss of security. The only drawback is that the system uses the Apriori algorithm which has some obvious bottlenecks with performance.

D. Jutla et al. [7] expresses concerns over the state of the cloud environment as it has been growing exponentially, which is the cause of a greater concern towards the privacy of the data being stored on the cloud. The researchers develop a technique that maintains the privacy of the cloud and its data as well as can provide master management to the cloud to help maintain suitable data in a Big Data Environment. There are seven components to this system that can collectively preserve the privacy of the system. One drawback of this technique is the increased space and time complexity.

A. Mishra et al. [8] explains the various techniques used to define the large amount to data which is being called as the Big Data. Big Data has been highly vital from the security point of view as a lot of data can easily allow for a greater insight against cyber-attacks. The data after an attack can disclose valuable information pertaining to the attackers and their motives. Due to the fact that most of the attackers utilize a particular modus operandi, it would be very easy to determine the attacker. The technique has not been tested extensively in a real-time scenario.

Y. Reddy et al. [9] expresses that there has been an exponential increase in the volume of data that is being generated every day. This data is very difficult to produce as the traditional techniques cannot be applicable to this ginormous size of data. The volume of the data is large but due to the velocity being high the data becomes very difficult to manage efficiently. Therefore, the authors have discussed the challenges faced with securing the sensitive data of the users and developing a secure model for the cloud environment. The one drawback is that there has not been a designate4d honeypot that can trap attackers and malicious users.

M. Chibba et al. [10] states that there has been an exponential increase in the amount of data generated and processed every day. This is a matter of high concern as most of the data has the opportunity to leak, which can result in the breach of privacy of a certain user. This is what can create a sense of unjust among the masses and a lot of tension for the individual. The authors have designed a technique for the increase in the privacy of the data which can be achieved by Collaboration, consultation, and co-operation.

III CONCLUSION

Cloud security is one of the most pressing topics in today's world of technological advancements. There has to be a way designed to protect cloud platform against the threats, internal as well as external. To maintain the integrity and security of the data on the cloud it is encrypted prior to uploading it on to the cloud. This technique of securing the data is one of the best and the most efficient but it also doesn’t allow traditional systems to be able to search over the encrypted data on the cloud.

Therefore, this paper proposes a technique for performing and enhancing search over the encrypted data. The proposed system utilizes a trap door to keep the intruder at bay. This proposed technique is faster and a lot more efficient than conventional systems. The trap door designed to be efficient and detects any intruders and safeguards your data accurately.

REFERENCES

[1] K. Strang and Z. Sun, “Meta-Analysis of Big Data Security and Privacy”, IEEE International Conference on Big Data, 2016.

[2] P. Sreekumari, “Privacy-Preserving Keyword Search Schemes over Encrypted Cloud Data: An Extensive Analysis”, 4th IEEE International Conference on Big Data Security on Cloud, 2018.

[3] S. Lighari and D. Hussain, “Hybrid model of rule-based and clustering analysis for big data security”, First International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT), 2017.

[4] Revathy P and R. Mukesh, “Analysis of Big Data Security Practices”, 3rd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 2017.

[5] S. Wang and R. Sinnott, “Privacy-protected Place of Activity Mining on Big Location Data”, IEEE International Conference on Big Data (BIGDATA), 2017.

[6] P. Pandiaraja and P. Kumar, “Efficient Multi-keyword Search Over Encrypted Data in Untrusted Cloud

Environment" Second International Conference on Recent Trends and Challenges in Computational Models, 2017.

[7] Dawn N. Jutla and Peter Bodorik, "PAUSE: A Privacy Architecture for Heterogeneous Big Data Environments", IEEE International Conference on Big Data (Big Data), 2015.

[8] A. Mishra and Y. Singh, "Big Data Analytics for Security and Privacy Challenges", International Conference on Computing, Communication and Automation, 2016.

[9] Y. Reddy, "Big Data Security in Cloud Environment", 4th IEEE International Conference on Big Data Security on Cloud, 2018.

[10] M. Chibba and A. Cavoukian, "Privacy, Consumer Trust and Big Data: Privacy by design and the 3 C'S", ITU Kaleidoscope: Trust in t