

An Enhanced Encryption Standard Algorithm for Security of Cloud Database Files

Ibebuogu C.C Ph.D

Department of Computer Science, Imo State University, Owerri
krischynwe@yahoo.com

Ozurumba Chidimma Naomi

Methodist High School Ikenegbu Owerri, Imo State.
ozurumbanaomi@gmail.com

DOI: 10.56201/ijcsmt.v9.no1.2023.pg30.35

Abstract

The aim of this work is to design and implement an enhanced encryption standard algorithm for security of cloud database files. The objective of the research is to design a Cloud computing system that can manage big data problem, and to enhance the security of files by the use of advanced encryption standard (AES) algorithm to protect files stored in the cloud. This research was motivated due to the difficulty in effective management of multiple cloud services by various organizations, limited Storage Space for resource management, unavailability of multiple clouds providers with huge scale disseminated setups and inability to effectively solve the problems of “Big Data” and the issue of security of files within the organization. The design will adopt the top down and object-oriented analysis and design methodology (OOADM) approach, while the programming language that will be deployed for the front-end design will be JavaScript, HTML, and CSS and the back-end design will be PHP and MySQL. The result after design will secure a web based cloud computing administrative system that will provide a friendly interface for an organization to have enough space to store and manage their resources remotely without fear of an attack of the stored files.

Keywords: Encryption, Algorithm, big data, Cloud Database

Introduction

With the continuous development of society and spring up of organizations, and economic progress, when a large amount of data enters the cloud computing system, people will pay more attention to data security. In order to make the stored data in the cloud more secure, according to the characteristics of cloud computing, There is need to modified data encryption algorithm in cloud computing. In other hand, Cloud computing (Khatoun et al; 2017) is an emerging computing model applied to the Internet. It provides basic resource facilities, application systems, and software platforms as services to users. Cloud computing is also a virtualization-based architecture that virtualizes resources and builds large-scale resource pools and manages services externally.

With the development of cloud computing, amounts of user data and large-scale data are put into cloud computing systems. Because of the distributed and virtualized nature of cloud computing, users cannot intuitively determine the storage location and division of data, etc., so the security of data becomes very important. In the cloud computing, data security is generally ensured through data encryption and identity management (Hsu et al; 2013). At present, the common encryption algorithms are classified into symmetric encryption algorithm and public key encryption algorithm. Among them, DES algorithm and AES algorithm are two widely used algorithms in symmetric encryption algorithms (Abdelminaam; 2018).

Cloud Computing is a set of information communication (IT) Services that are provided to a customer over a network and these services are delivered by third party provider who owns the infrastructure. It is often provided "as a service" over the Internet, typically in the form of infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS) Sakr et al.(2014). Cloud computing is the broader concept of infrastructure convergence. This type of data center environment allows enterprises to get their applications up and running faster, with easier manageability, and less maintenance to meet business demands. For example, with cloud systems, it is possible to manage and store all smart phones or tablets applications at one location using the cloud. With these, there is no need for any memory space at the office making it possible for the security of data and applications in case device is damaged or lost Abha and Bhansali (2013).

Encryption

This is the method by which information is converted into secret code that hides information true meaning.

Big Data

Big Data refers to data sets that are too large or complex to be dealt with by traditional data processing application software.

Algorithm

An algorithm is a procedure used for solving a problem or performing a computation. Algorithm acts as an exact list of instructions that conduct specified actions step by step in either hardware or software based routine.

BIG BATA AND CLOUD COMPUTING

The term "Big Data" has recently been applied to datasets that are very huge to the extent of being difficult or awkward to work on by the use of a traditional database management system (DBMS). They can be also referred to as dataset whose size cannot be managed by just a common software tool and electronic storages, to even capture, manage, process or store files within a tolerable elapsed time is difficult.

(Kubick, 2012). According to Davenport and Dyché (2013), big data is seen as a new trend in the last decades, making various organizations like Google, eBay, Twitter, Yahoo and Facebook to become concerned about the big data and its appearance on the business scene. Big data represents higher and richer data that shows more details about behaviors, activities, and events that happened all around, so the analytics of these big data would need access to diverse types of resources within less response time. Looking at the rate at which data sizes are continually

increasing, presently extending from a few dozen terabytes (TB) to Petabytes (PB) of data in a single dataset. Accordingly, looking at the few difficulties surrounding big data, which includes: data capture, storage, sharing, visualizing, search, and analytics, currently, various initiatives are exploring large volumes of highly detailed data so as uncover evidences they could not know before now (Russom 2011).

CHARACTERISTICS OF BIG DATA

Many scholars have described three popular V's in big data i.e. Volume, Velocity, and Variety. Moreover, two more V's are added later- Veracity and Value.

- i. **Volume** – Looking at the Big Data, it could be related to be a size which is huge. Size of data shows a very vital role in decisive value out of data. Moreover, whether a specific data can truly be considered as a Big Data or not, is reliant on upon the volume of data. Hence, 'Volume' is one distinguishing factor which needs to be considered while dealing with Big Data.
- ii. **Variety** – The next aspect of Big Data is its variety. Variety refers to heterogeneous sources and the nature of data, both structured and unstructured. During earlier days, spreadsheets and databases were the only sources of data considered by most of the applications. Nowadays, data in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. are also being considered in the analysis applications. This variety of unstructured data poses certain issues for storage, mining and analyzing data.
- iii. **Velocity** – The term 'velocity' refers to the speed of generation of data. How fast the data is generated and processed to meet the demands, determines real potential in the data. Big Data Velocity deals with the speed at which data flows in from sources like business processes, application logs, networks, and social media sites, sensors, Mobile devices, etc. The flow of data is massive and continuous.
- iv. **Variability** – This refers to the inconsistency which can be shown by the data at times, thus hampering the process of being able to handle and manage the data effectively.

Moreover, big data can be described by its velocity or speed. This is basically the frequency of data generation or the frequency of data delivery. The leading edge of big data is streaming data, which is collected in real-time from the websites (Russo, 2011). Some researchers and organizations have discussed the addition of a fourth V, or veracity. Veracity focuses on the quality of the data. This characterizes big data quality as good, bad, or undefined due to data inconsistency, incompleteness, ambiguity, latency, deception, and approximations (TechAmerica, 2012).

TYPES OF BIG DATA

Big Data' could be found in three forms according to guru99.com:

1. Structured
2. Unstructured
3. Semi-structured

STRUCTURED

Whichever data that can be warehoused, retrieved and managed in the form of immovable format is characterized as a 'structured' data. Over the period of time, ability in computer science has attained countless success in emerging methods for working with such kind of data (such a way that the format is well known in advance) and also originating significance out of it. Nevertheless, today, we are predicting issues when a size of such data develops to a large magnitude, characteristic sizes are being in the range of several zettabytes. Observing these figures, one could easily comprehend why the name Big Data is given and visualize the problems involved in its storage and processing.

UNSTRUCTURED

Any data with unknown form or the structure is classified as unstructured data (guru99, 2019) secondly; to the size being massive, unstructured data poses numerous problems in terms of its processing for actualizing value out of it. A distinctive instance of unstructured data is a heterogeneous data source a mixture of simple text documents, imageries, videotapes etc. Currently establishments have wealth of data accessible with them but inappropriately, they don't know how to derive value out of it meanwhile this data is in its unprocessed form or unstructured format. A characteristic instance of Unstructured Data is the returned output by 'Google Search'

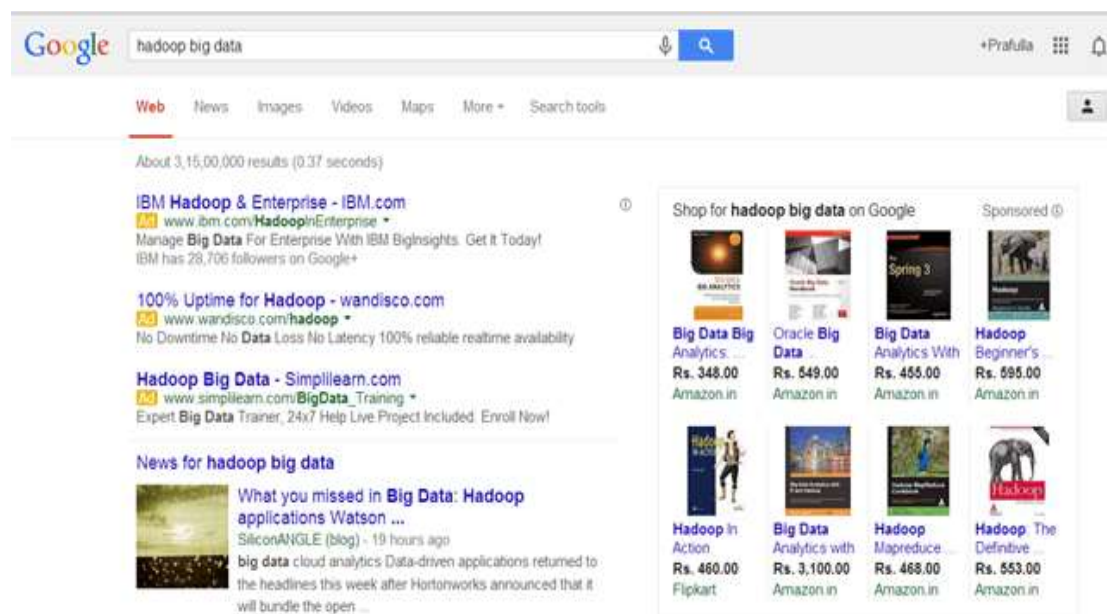


Figure 1: Example of Unstructured Data; The output returned by Google Search.

SEMI-STRUCTURED

The Semi-structured data could enclose both the forms of data. Semi-Structured data can be seen in two ways, as a structured format for instance, a table found in relational database management systems (DBMS). A characteristic instance of a semi-structured data could be illustrated in an extensible mark-up

language (XML). <rec><name>OluluGbeseemi</name><sex>Male</sex><age>35</age></rec>
<rec><name>Praise Okereke.</name><sex>Female</sex><age>41</age></rec><rec><name>
Messy Daniel</name><sex>Male</sex><age>29</age></rec>

EXAMPLES OF BIG DATA

The following are some of instances of a Big Data:

New York Stock Exchange:This makes at about one terabyte of new trade data on daily basics.

Social Media:The figure shows that 500+terabytes are being of new data get swallowed into the databases of social media site Facebook, every day.

BENEFITS OF BIG DATA PROCESSING

For various organizations to have the ability to generate Big Data which could bring in several benefits, for example, some Businesses can use outside intellect while captivating decisions: Availability of various social data from search engine and websites such as Twitter, Facebook provide quick access to easy data for organizations to fine turn their business strategies; Help in Improvement of client service: Provide traditional consumer feedback systems which are getting substituted by innovative systems designed with Big Data technologies.

In these new systems, Big Data and natural language processing technologies are being used to read and evaluate consumer responses. Initial identification of risk to the product/services, if any, and Better operational efficiency: Big Data machineries could be used for generating a presentation area or landing zone for new data before identifying what data should be moved to the data warehouse. Secondly, such integration of Big Data technologies and data warehouse helps an organization to offload rarely accessed data.

References

- Khatoon, T. Thakur, B. Singh, (2017).A provable secure and escrow-able authenticated group key agreement protocol without NAXOS trick," International Journal of Computer Applications, vol. 171, no. 3, pp. 1-8,
- Hsu, C. C. Yang, and M. S. Hwang (2013) A study of public key encryption with keyword search", International Journal of Network Security, vol. 15, no.2, pp. 71{79.
- Abdelminaam. D. S., (2018) \Improving the security of cloud computing by building new hybrid cryptography algorithms," International Journal of Electronics and Information Engineering, vol. 8, no. 1, pp.40{48,
- Davenport T. H and Dyché I. (2013) 'Big Data in big companies'.International Institute for Analytics.
- R. H. Sakr, F. Omara, O. Nomir (2014) "An Optimized Technique for Secure Data Over Cloud OS" International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) Volume 3, Issue 3, May-June 2014 ISSN 2278-6856

AbhaSachdev, MohitBhansali (2013) “Enhancing Cloud Computing Security using AES Algorithm” International Journal of Computer Applications (0975 – 8887) Volume 67– No.9,

Kubick, W.R (2012) Big Data, Information and Meaning. In: Clinical Trial Insights, retrieved from <https://www.sciencedirect.com/> pp 26 – 28.

Russom P (2011) Big Data Analytics, In: TDWI Best Practices Report, pp 1 - 40

TechAmerica (2012) Demystifying Big Data: A Practical Guide to Transforming the Business of Government. In: TechAmerica Reports, pp 1 – 40.

Guru99: Inside Facebook’s Blu - Ray Cold Storage Data Center website:

<https://datacenterfrontier.com/inside-facebook-blu-ray-cold-storage-data-center/>
Introduction to Big data: <https://www.guru99.com/what-is-big-data.html>