



ARCHITECTURE FRAMEWORK AND COMPONENTS FOR BIG DATA SECURITY IN CLOUD COMPUTING.

¹Nutan Shelar, ²Sanket Bhawe, ³Priya Chandran

¹Student, ²Student, ³Asst.Professor

Bharati Vidyapeeth' Institute of Management & information Technology,
Navi Mumbai ,India.

Abstract: In this data-driven society, we are collecting a large amount of knowledge and data from people, their actions, sensors, algorithms, and also the web, and for that, the reason handling “Big Data” has become a significant challenge. Big data is that the next big thing in computing. As this data can not be processed using traditional systems, it poses numerous challenges to the research community. Privacy and security are among the important concerns with data, be it traditional data or big data. The cloud is increasingly getting used to store and process large data. Many researchers are trying to guard big data in a cloud computing environment. Traditional security mechanisms are neither efficient nor suited to the task of protecting big data within the Cloud. Protecting data within the cloud are often the same as protecting data within a conventional data center. Some components are used to secure data in cloud-like Authentication and identity, access control, encryption, secure deletion, integrity checking, and data masking are all data protection methods that have applicability in cloud computing.

Index Terms - components, big data Architecture, Big data security architecture, cloud computing, Big Data Analytics and Technologies for Big Data, MapReduce, Hadoop Map

I. INTRODUCTION

During recent years, the data production rate has been growing exponentially. Many organizations demand efficient solutions to store and analyze these big amounts of data that are preliminarily generated from various sources such as high throughput instruments, sensors, or connected devices. For this purpose, big data technologies can utilize cloud computing to supply significant benefits, like the provision of automated tools to assemble, connect, configure, and reconfigure virtualized resources on demand. Cloud services make it much easier for organizations to achieve their goals as they can deploy them easily. This shift in paradigm that accompanies the adoption of cloud computing is increasingly giving rise to security and privacy considerations referring to facets of cloud computing like multi-tenancy, trust, loss of control, and accountability. Consequently, cloud platforms that handle big data that contain sensitive information are required to deploy technical measures and organizational safeguards to avoid data protection breakdowns which may end in enormous and expensive damages. Sensitive information within the context of cloud computing encompasses data from a good range of various areas and disciplines. Data concerning health is a typical example of the type of sensitive information handled in cloud computing environments, and most people will want information associated with their health to be secure. Hence, with the proliferation of those new cloud technologies in recent times, privacy and data protection requirements are evolving to guard individuals against surveillance and database disclosure. Some samples of such protective legislation are the EU Data Protection Directive (DPD) and thus the US insurance Portability and Accountability Act (HIPAA), both of which demand privacy preservation for handling personally identifiable information.

An overview of research on the security and privacy of big sensitive data in cloud computing environments is presented in this paper. We point out new developments in the orchestration, resource control, physical hardware, and cloud service management layers of a cloud provider.

BIG DATA SECURITY AND PRIVACY ISSUES:

In Today's world, we know that all devices connected to the internet and connected, the volume of data collected, stored, and processed is increasing every day, which also brings new challenges to interns in information security. Big data analytics can be used to detect and prevent malicious intruders and advanced threats[13]. as per the researcher of procedia computer science, there are some issues in big data security in cloud computing as following 1) To guard and stop the large size of confidential business, government or regulatory data from malicious intruders and advanced threats, 2) Lack of standards and awareness about how cloud service providers securely maintaining the large disc space and erase existing big data, 3) Lack of standards about auditing and reporting of big data in the public cloud, 4) Users who do not even work for the organization (malicious intruders), but may have full control and visibility into the history of organization data (big data)[2][14]. there are many researchers are developing big data security architecture and frameworks to protect large data in the cloud. some of the architectures are like preventing and detecting intrusion [15], Securing web applications, protecting confidential data in the cloud [16], Protecting GPS data from data mining-based attacks, and securing the bank details in the cloud [17].

II. Literature Review:

In [1] the authors focuses on big data architecture which somehow results in the top five big data security risks and therefore the use of top Five best practices that ought to be considered while designing big data solutions which might thereby surmount these risks. Big data architecture, being distributive can undergo partition, replication, and distribution among thousands of knowledge and processing nodes for distributed computation thus supporting multiple features related to big data analytics like real-time, streaming, and continuous data computation together with massive parallel and powerful programming framework. These series of characteristics are put into effect via a key setup that somehow ends up in certain crucial security implications. The challenges induced by this could be handled via big data technologies and solutions that exist inside big data architecture compounds characterized for specific big data problems. Big data solutions should provide effective ways to be more proactive against fraud, management, and consolidation of information, proper security against data intrusion, malicious attacks, and lots of other fraudulent activities. particularly, this paper discusses the problems and key features that ought to be taken into consideration while undergoing the event of secured big data solutions and technologies that may handle the risks and privacy concerns (e.g. Data security, insecure computation and data storage, invasive marketing, etc.) related to big data analysis in an efficient thanks to increasing the performance impact, considering that these risks are somehow results of characteristics of huge data architecture.

The cloud is increasingly being employed to store and process big data [2]. Many researchers are trying to safeguard big data during a cloud computing environment. during this paper, they first discuss challenges and potential solutions for safeguarding big data in cloud computing. Second, they propose Meta Cloud Data Storage Architecture for safeguarding Big Data in Cloud Computing Environment. This framework ensures efficient processing of huge data in a very cloud computing environment and gains more business insights.

Big Data and cloud computing are two important issues in recent years, enables computing resources to be provided as Information Technology services with high efficiency and effectiveness [3]. Now a day's big data is one of the foremost problems that researchers try and solve it and focusing their researches thereon to urge ride the matter of how big data can be handled in recent systems and managed with the cloud of computing, and one among the foremost important issues is the way to gain perfect security for giant data in cloud computing, our paper reviews a Survey of huge data with clouds computing security and therefore the mechanisms that wont to protect and secure even have privacy for giant data with an available cloud.

Big Data is becoming a replacement technology focus both in science and in industry and motivate technology to shift to data-centric architecture and operational models. there's an important must define the fundamental information/semantic models, architecture components, and operational models that together comprise a so-called Big Data Ecosystem [4]. This paper discusses the character of huge Data that will originate from different scientific, industry, and group action domains and proposes an improved Big Data definition that features the subsequent parts: Big Data properties, data models and structures, data analytics, infrastructure, and security. The paper discusses paradigm change from the standard host or service-based to data-centric architecture and operational models in Big Data. the massive Data Architecture Framework (BDAF) is proposed to handle all aspects of the massive Data Ecosystem and includes the subsequent components: Big Data Infrastructure, Big Data Analytics, Data structures, and their models, Big Data Lifecycle Managements, Big Data Security. The paper analyses requirements and provides suggestions on how the mentioned above components can address the most Big Data challenges.

Big data, with their promise to find valuable insights for better deciding, have recently attracted significant interest from both academia and industry [5]. Voluminous data are generated from a range of users and devices and are to be stored and processed in powerful data centers. As such, there's a powerful demand for building an unimpeded network infrastructure to collect geologically distributed and rapidly generated data and move them to data centers for effective knowledge discovery. The express network should even be seamlessly extended to interconnect multiple data centers furthermore as interconnecting the server nodes within a knowledge center. during this article, they take a detailed have a look at the unique challenges in building such a network infrastructure for large data. Our study covers each and each segment during this network highway: the access networks that connect data sources, the web backbone that bridges them to remote data centers, likewise because of the dedicated network among data centers and within an information center. they also present two case studies of real-world big data applications that are empowered by networking, highlighting interesting and promising future research directions.

Network traffic could be a rich source of knowledge for security monitoring [6]. However, the increasing volume of knowledge to treat raises issues, rendering holistic analysis of network traffic difficult. during this paper, they propose an answer to address the tremendous amount of knowledge to analyze for security monitoring perspectives. they give an architecture dedicated to security monitoring of local enterprise networks. the applying domain of such a system is principally network intrusion detection and

prevention but is often used in addition for forensic analysis. This architecture integrates two systems, one dedicated to scalable distributed data storage and management and also the other dedicated to data exploitation. DNS data, NetFlow records, HTTP traffic, and honeypot data are mined and correlated in an exceedingly distributed system that leverages state-of-the-art big data solutions. Data correlation schemes are proposed and their performance is evaluated against several well-known big data frameworks including Hadoop and Spark.

In [7] the authors discussed the problem of online privacy and security may be a challenging subject because it concerns the privacy of knowledge that is increasingly more accessible via the web. In other words, people that shall access the private information of other users can do so more efficiently over the web. This study is an endeavor to deal with the privacy issue of distributed big data within the context of cloud computing. it's of great importance to safeguard and authenticate ownership data one in all the cases where data privacy is of great importance. during this paper, this privacy issue is analyzed by Petri net modeling. What today's organizations need for his or her clouds are integrated comprehensive solutions which will deliver counterintelligence. Advanced counterintelligence solutions can close security gaps by using labor-saving automation to research voluminous events occurring within the cloud, and find out system vulnerabilities through the normalization and correlation of those events. Using the proposed method, a model of security, including control of user access to databases of massive data with RMS, the multiplicity and therefore the virtual machine to stop internal threats, deleting data, insecure or incomplete data protection and control of a third-party will be provided to boost the operation in keeping with the foundations of Petri net modeling and simulation

During this data-driven society, they're collecting an oversized amount of knowledge from people, actions, sensors, algorithms, and also the web; handling "Big Data" has become a significant challenge [8]. a problem still exists regarding when data could even be called big data. How large is big data? what's the optimal solution for storing, editing, retrieving, analyzing, maintaining, and recovering big data? How can cloud computing help in handling big data-related issues? what's the role of cloud architecture in handling big data? How important is big data in business intelligence? This chapter attempts to answer these questions. First, they review a definition of giant data. Second, they describe the important challenges of storing, analyzing, maintaining, recovering, and retrieving infinite data. Third, they address the role of Cloud Computing Architecture as an answer for these important issues that cater to big data. they also discuss the definition and major features of cloud computing systems. Then they explain how cloud computing can provide an answer for large data with cloud services and open-source cloud software tools for handling big data issues. Finally, they explain the role of cloud architecture in big data, the role of major cloud service layers in big data, and also the role of cloud computing systems in handling big data in business intelligence models.

Nowadays, organizations have at their disposal an outsized volume of information with a good type of types. Technology-driven organizations want to capture processes and analyze this data at a quick velocity, to raised understand and manage their customers, their operations, and their business processes. the maximum amount as data volume and variety increases and as faster analytic results are needed, more demanding are for information architecture [9]. This data architecture should enable collecting, storing, and analyzing Big Data in Cloud Environment. Cloud Computing, ensures timeliness, ubiquity, and simple access by users. This paper proposes to develop a knowledge architecture to support Big Data in Cloud and, finally, validate the architecture with a symbol of the concept.

With the many advances in communication technologies and many other sectors, are growing up security and privacy issues [10]. In this research, a base technology called Cloud Computing (CC) is introduced to work with massive Data (BD). CC could be a technology that refers to the processing power of information within the fog, providing more "green" computational and sustainable computing. Since it's a recently investigated technology, it's many gaps in security and privacy. So, during this paper, they proposed a replacement system for Cloud Computing integrated with the Internet of Things as a base scenario for large Data. Moreover, they tried to ascertain an architecture relying on the protection of the network to boost the safety issues. an answer proposed is installing a security "wall" between the Cloud Server and also the Internet, intending to remove the privacy and security issues. As a result, they consider that CC deals more efficiently with the privacy issue of bits transferred through time. Through our proposed system, the interaction and cooperation between things and objects communicate through the wireless networks to fulfill the target set to them as a combined entity. Regarding the key goal of our research, which is that the security, a kind survey of IoT and CC presented, with attention on the safety problems with both technologies. additionally, to the current, they struggle to present the safety challenges of the mixing of IoT and Cloud Computing to produce an architecture relying on the protection of the network to enhance their security issues. Finally, they realize that through our study Cloud Computing could offer a more "green" and efficient fog environment for sustainable computing scenarios.

A. Big data:

Big Data is a very big thing in today's world refer as e flood of digital data from many digital earth sources, digitizers, scanners, numerical modeling, including sensors, mobile phones, the Internet, videos, e-mail's, and social network, etc. The data types include texts, geometries, images, videos, sounds, and combinations of every. Such data can be directly or indirectly associated with geospatial information [11]. "Big data" may be a relatively modern field of knowledge science that explores how large data sets may be de-escalated and analyzed to systematically glean insights and knowledge from them. Earlier, conventional processing solutions aren't very efficient concerning capturing, storing, and analyzing big data. Hence, companies with traditional BI solutions aren't ready to fully maximize their worth. to successfully understand what big data means, we'd like to require a glance at the 5 V's of huge data. Big data characterized by volume, velocity, variety, veracity, and value. The first V refers to the amount of information that is growing explosively and extends beyond our capability of handling large data sets; volume is that the commonest descriptor of huge Data. Velocity refers to the fast generation and transmission of information across the net as exemplified by data collection from social networks, the massive array of sensors from the micro to the macro level, and data **transmission** from sensors to supercomputers and decision-makers. Variety refers to the various data forms and during which model and structural data are archived. Veracity refers to the diversity of quality, accuracy, and trustworthiness of the info. All above four Vs are important for

the next step to the 5th V, which focuses on specific research and decision-support applications that improve our lives, work, and prosperity.

B. Cloud Computing:

Simply the Cloud computing refers to applications and services that run on a distributed network using virtualized resources and accessed by common internet protocols and network standards. it's like a resource on demand whether or not it's storage, computing, etc. Cloud follows a pay-per-usage model. you would like to pay the number of resources you utilize.

This computing service by cloud charges you based only on the number of computing resources we use.

Cloud plays a very important role within the large Data world, by providing horizontally expandable and optimized infrastructure that supports practical implementation of huge Data.

the main definition of cloud computing model, made by NIST as “Cloud computing may be a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, applications, storage, and services) which will be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and consists of 5 essential characteristics, three service models, and 4 deployment models.”[12].

Not all clouds are identical and not one style of cloud computing is true for everybody. Several different models, types, and services have evolved to assist offer the correct solution for your needs.

First of all, you would like to work out the sort of cloud computing architecture, that your cloud services are implemented on.

There are three other ways to deploy cloud services: on a public cloud, private cloud, or hybrid cloud, and Service models of cloud is classified into three types such as infrastructure as a service (IaaS), platform as a service (PaaS), serverless and software as a service (SaaS).

Cloud computing could be a big shift from the standard way businesses consider IT-related resources. Here are seven common reasons organizations are turning to cloud computing services, there are some benefits are like cost, speed, productivity, global scale, performance, security, and reliability.

C. Big Data Analytics and Technologies for Big Data

Big data analytics helps businesses to induce insights from today's huge data resources. People, organizations, and machines now produce massive amounts of knowledge. Social media, cloud applications, and machine sensor data are just a few examples. Big data will be examined to determine big data trends, opportunities, and risks, using big data analytics tools. Here are some key technologies that enable Big Data for Businesses:1) Predictive Analytics,2) NoSQL Databases,3) Knowledge Discovery Tools,4) Stream Analytic 5) In-memory Data Fabric,6) Distributed Storage,7) Data Virtualization,8) Data Integration,9) Data Preprocessing,10) Data Quality.

D. MapReduce

MapReduce is intended to be employed by programmers, instead of business users. it's a programming model, not a programming language. it's gained popularity for its easiness, efficiency, and skill to regulate “Big Data” in an exceedingly timely manner. The applications which include graph analysis, indexing and search, text analysis, machine learning, data transformation, and lots more, aren't easy to implement by making the employment of ordinary SQL which are employed by relational DBMSs. In such areas the procedural nature of MapReduce makes it easily understood by other programmers. It also has the advantage that developers do not must be troubled with implementing parallel computing – this is often handled transparently by the system. Although MapReduce is intended for programmers, nonprogrammers can exploit the worth of prebuilt MapReduce applications and performance libraries. Map Reduce programs are usually written in Java. they will even be coded in other languages like C++, Python, Ruby, R, etc. These programs may process data stored in numerous file and database systems. At Google, as an example, MapReduce was implemented on top of the Google classification system (GFS).

Data Analytics: Map Reduce Algorithm for processing log files in a distributed cloud data center

Map Reduce is a programming model or framework that processes tasks in parallel across a huge size of systems. It contains two functions as Map and Reduces. Map function splits the huge size of input data into <key, value> pairs. Intermediate <key, value> pairs will be created bases on aggregating several input key-value pairs from the Map phase. Finally, Reduce takes the intermediate key-value pairs and produces the output <key, value> pairs that can be easily understood by the end-user. In this proposed architecture, Map Reduce 132 Gunasekaran Manogaran et al. / Procedia Computer Science 87 (2016) 128 – 133 framework is used to find the number of users who were logged in to the cloud data center. Proposed MapReduce pseudo-code can efficiently process the huge size of the log file in which it contains users who were logged in with a date and the log-in time duration. As shown in Figure 3, the first process is the map phase in which each date that represents the key is assigned a value of one initially. While reducing phase, the key values are summed up to find out the number of users logged in. For example, three users were logged in on 01-02-2016, whereas two users were logged in on 02-02-2016.

E. Hadoop Map

Hadoop is an open-source software framework for storing data and running applications on clusters of commodity hardware. It provides massive storage for any kind of data, huge processing power, and the ability to handle virtually unlimited concurrent tasks or jobs. There are some important components used in Hadoop 1) Ability to store and process huge amounts of any kind of data 2) quickly 3) Computing power 4) Fault tolerance 5) Flexibility 6) Low cost 7) Scalability. HDFS (Hadoop Distributed File System): HDFS is a distributed file system that handles large data sets running on commodity hardware. It is used to scale a single Apache Hadoop cluster to hundreds and in some cases thousands of nodes. HDFS is one of the major components of Apache Hadoop, the others being MapReduce and YARN. HDFS should not be confused with or replaced by Apache HBase, which is a column-oriented non-relational database management system that sits on top of HDFS and can better support real-time data needs with its in-memory processing engine.

3). Problems of Main Topic:

Big data analytics are often wont to detect and forestall malicious intruders and advanced threats[13]. as per the researcher of procedia engineering, there are some issues in big data security within the cloud computing as following 1) to safeguard and stop huge size of confidential business, government or regulatory data from malicious intruders and advanced threats, 2) Lack of awareness and standards about how cloud service providers securely maintaining the large disc space and erase existing big data, 3) Lack of standards about auditing and reporting of massive data publicly cloud, 4) Users who don't even work for the organization (malicious intruders), but may have full control and visibility into the history of organization data (big data)[2][14]. there are such a big amount of researchers are developing big data security architecture and frameworks to shield huge data in the cloud. the protection issues in cloud computing include 1) Data security 2) Identity and access control 3) Key management 4) Virtual machine security Among these main security issues within the cloud, data security and integrity is believed to be the foremost difficult problem that could limit the employment of cloud computing. Access control and key management are all issues involved in data security. Data security within the cloud refers to data confidentiality, integrity, availability, and traceability (CIAT), and these requirements pose major problems for cloud computing. Confidentiality: it is Data confidentiality requires that information be available or disclosed only to authorized individuals, entities, or IT processes. Integrity: Data integrity ensures that the information is maintained in its original state and has not been intentionally or accidentally altered or deleted. Availability: Data availability ensures continuous access to data even within the occurrence of a natural or man-made disaster or events like fires or power outages. Traceability: Data traceability implies that the info and communications are genuine in an exceedingly transaction in which both parties involved are who they claim to be. Authentication: Authentication could be a method by which a system verifies and validates the identity of a user of the system who wishes to access it. Specifically, to realize the above requirements of CIAT, the critical security challenges of knowledge security within the cloud will be mainly outlined as follows: 1. Key management 2. Access control 3. Searchable encryption techniques 4. Remote integrity check 5. Proof of ownership.

4). Solution:

In this paper, we proposed an Architecture framework and components for large data security in cloud computing protecting Big Data in Cloud Computing Environment. Security is that the main issue in big data. handle knowledge on the cloud could be a very big challenge. We review all research papers and here we conclude that the Meta Cloud Data Storage Architecture is best thanks to protecting data on the cloud. Meta Cloud Data Storage Architecture for safeguarding big data in the cloud and Map Reduce framework is employed to seek out the number of users who were logged into the cloud data center. during this research paper, the Proposed framework protects the mapping of assorted data elements to every provider using the Meta Cloud Data Storage interface. Though this proposed approach requires high implementation effort, it provides valuable information for the cloud computing environment which will have a high impact on the following generation systems. Our future work is to increase the proposed Meta Cloud Data Storage Architecture for the real-time processing of streaming data.

Conclusion:

This paper reviewed several security and privacy issues on big data within the cloud. Also, several architectural frameworks accustomed solve security issues We also discussed several security challenges that are raised by existing in the current situation, The results that are presented within the area of cloud security and privacy are supported cloud provider activities, like providing orchestration, resource abstraction, physical resource, and cloud service management layers. Security and privacy factors that affect the activities of cloud providers about the legal processioning of consumer data were identified and a review of existing research was conducted to summarize the state-of-the-art within the field.

REFERENCES

- [1] Big Data Security – Challenges and Recommendations Renu Bhandari¹, Vaibhav Hans^{2*} and Neelu Jyothi Ahuja³
1,2*University of Petroleum and Energy Studies, India 3Centre of Information Technology, University of Petroleum and Energy Studies, Indi
- [2] Meta Cloud Data Storage Architecture for Big Data Security in Cloud Computing Gunasekaran Manogarana, Chandu Thotab
*, M. Vijay Kumar
- [3] A Survey of Big Data Cloud Computing Security Elmustafa Sayed Ali Ahmed¹ and Rashid A.Saeed² 1 Electrical and Electronic Engineering Department, Red Sea University, Portsudan, Red Sea State, Sudan 2 School of Electronics Engineering, Sudan University of Science and Technology, Khartoum, Khartoum State, Suda
- [4] Architecture Framework and Components for the Big Data Ecosystem Draft Version 0.2 Yuri Demchenko, Canh Ngo, Peter Membrey 12 September 2013
- [5] Building a Network Highway for Big Data: Architecture and Challenges Xiaomeng Yi, Fangming Liu, Jiangchuan Liu, and Hai Jin
- [6] A Big Data Architecture for Large Scale Security Monitoring Samuel Marchal* †, Xiuyan Jiang‡, Radu State* , Thomas Engel* * SnT - University of Luxembourg, Luxembourg – { Samuel.Marchal | Radu. state | Thomas.Engel }@uni.lu † TELECOM Nancy - University of Lorraine, France – samuel.marchal@univ-lorraine.fr ‡ Faculty of Sciences, Technology and Communication - University of Luxembourg, Luxembourg – kangsyoun@gmail.com

- [7] An Architecture for Security and Protection of Big Data F. Asadi Saeed Abada, H. Hamidi* b a Department of Information Technology Engineering, College of Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran b Department of Industrial Engineering, Information Technology Group, K. N.Toosi University of Technology, Tehran, Iran
- [8] Information Granularity, Big Data, and Computational Intelligence, W. Pedrycz and S.- M. Chen (eds.), Vol. 8, 2015 Chapter 13: The Role of Cloud Computing Architecture in Big Data Authors: Mehdi Bahrami and Mukesh Singhal Cloud Lab, University of California, Merced
- [9] Big Data in Cloud: A data architecture Jorge Oliveira e Sá1, César Martins2, and Paulo Simões3 1University of Minho, Portugal, jos@dsi.uminho.pt 2University of Minho, Portugal, pg21441@alunos.uminho.pt 3 ISEG – School of Economics and Management, Portugal, paulo.simões@iseg.utl.pt
- [10] Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT Christos Stergiou1, Kostas E. Psannis1,*, Brij B. Gupta2, Yutaka Ishibashi3 1Department of Applied Informatics, School of Information Sciences, University of Macedonia, Thessaloniki, Greece 2National Institute of Technology Kurukshetra, India 3Department of Scientific and Engineering Simulation Nagoya, Japan
- [11] Berkovich, S., and D. Liao. 2012. On Clusterization of Big Data Streams. In Proceedings of the 3rd International Conference on Computing for Geospatial Research and Applications, 9. ACM. Accessed November 25, 2015.
- [12] Peter Mell, and Tim Grance, “The NIST Definition of Cloud Computing,” Version 15, 10-7-09, <http://www.wheresmyserver.co.nz/storage/media/faq-files/cloud-def-v15.pdf>
- [13] Victor, N., Lopez, D., & Abawajy, J. H.. Privacy models for big data: a survey. International Journal of Big Data Intelligence, 2016 3(1), 61-75.
- [14] Thilagavathi, M., Lopez, D., & Murugan, B. S. Middleware for Preserving Privacy in Big Data. Handbook of Research on Cloud Infrastructures for Big Data Analytics, IGI Global, 2014.
- [15] Marchal S, Jiang X, State R, Engel T. A Big Data Architecture for Large Scale Security Monitoring. In 2014 IEEE International Congress on Big Data (BigData Congress), 2014 Jun 27: 56-63. IEEE.
- [16] Hongbing C, Chunming R, Kai H, Weihong W, Yanyan L. Secure big data storage and sharing scheme for cloud tenants. Communications, China. 2015 Jun;12(6):106-15.
- [17] Subashini S, Kavitha V. A metadata-based storage model for Securing data in cloud environment. In CyberC 2011 Oct 10:429-434.
- [18] Siddesh GM, Hiriyanaiyah S, Srinivasa KG. Driving Big Data with Hadoop Technologies. Handbook of Research on Cloud Infrastructures for Big Data Analytics. IGI Global, 2014 Mar 31:232.