IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 11, November 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011017 |

A Novel Approach for Network Traffic and Attacks Analysis Using Big Data in Cloud Environment

Ravindra Changala[1], M Sunny Kumar[2], Fardeen Abdul Aziz[3], MD Ibrahim Zia Ul Huda[4], B Hari[5]

[1]Assistant Professor, Department of IT, Guru Nanak Institutions Technical Campus, Hyderabad, India

[2,3,4,5] Department of IT, Guru Nanak Institutions Technical Campus, Hyderabad, India

ABSTRACT: Recently mining of data from online life is pullingin more thought due to the shoot in the advancement ofLarge Information. In security, Large Information deals with a collection of monstroushigh level information for researching, imagining and to draw the piecesof information for the assumption and expectation of advancedattacks. Big Data Analytics (BDA) is the term made byspecialists to depict the specialty of managing, dealing with andassembling a lot of information for future assessment. Information is beingmade at a disturbing rate. The fast improvement of the Web,Internet of Things (IoT) and other imaginative advances are the standardresponsible social occasions behind this continued with progression. Theinformation established is a connection of the earth, it is conveyed out of, alongthese lines can utilize the information moved away from designs to comprehendthe inward activities of that framework. This has turned into a criticalcomponent in network safety where the goal is to getassets. Also, the creating assessment of data hasmade enormous data a high worth goal. At this moment,examine progressing investigation works in digital protectionsimilar to gigantic data and component how Large datais gotten and the way in which gigantic data can in like manner be used as agadget for network protection. All the while, a Major Information basedconcentrated log examination system is completed torecognize the framework traffic occurred with aggressors throughDDOS, SQL Infusion and Brute force attack. The log record isnormally communicated to the united cloud server and largedata is begun in the examination cycle.

KEYWORDS: DDOS, SQL, big data, traffic analysis, network attacks, cyber security, IoT.

I.INTRODUCTION

In the current period, information development has accomplished quickprogression in organizations and adventures which made the termLarge Information particularly standard. The augmentation indata improvement is outstandingly fast as thedata is made from a variety of sources, formodel, online life, pictures in electronic position,high level accounts, business record, etc. The leading group ofthis enormous proportion of data known as Large Information is atroublesome endeavor. This data can get pay tothe undertakings as authentic assessment of this Enormous Informationprompts suitable cognizance of the client requirementsto take decision on the key reason. On the other hand, hacking of theenormous data prompts veritable risk as there is alikelihood of the expansion of noxious programming in theworking structures and the applications. In this manner to ensureabout the Huge data about the computerized perils updatedmethod is proposed and executed at the present time.

Top security associations joined to bestow information to oneanother attempting to gather understanding from the shareddata. Their goal was to give reliable securitycontraptions to their clients, and to achieve that, theyexpected to accept in whatever amount as could be anticipated frompropelling perils that were developed consistently. Theyunderstood the power of joint exertion for more unmistakable benefit. This was expected because with the risingof polymorphic malware and other propelling risks, theyrequired an extraordinary arrangement information on these risks in order to totallyappreciate what they were overseeing and how to killagainst it. The ordinary techniques of requestingmalware were turning out to be purposeless. SecIntel Tradedata allowed them the opportunity to get critical pieces ofinformation from voluminous data. Human assessmentfurthermore, traditional strategies.

Past region showed how security can be polishedwith Large information. This part familiarizes how with guarantee about thegiant information against various ambushes.Right when the informationgets massive, depicting secure with it turns as extremely irritating.In [58], producers considered the security issues related withenormous information and scattered figuring. They saw the waythat most affiliations reallocate data set as huge informationinto the cloud.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011017 |

Appropriated handling anyway,everything has different hazards related with it. The objective in wasto track down security weaknesses in the cloud to edifymerchants about late weaknesses. They saw that gathering,respectability and receptiveness in a specific solicitation as the most enormoussecurity gives a cloud supplier face. Portrayal rightas of now infers the security of information against unapprovedimpedance or use. Reliability would be the assumption forunapproved and inappropriate information change. Receptiveness would be apart of identical to information recuperation from stuff, programming andconstruction, messes up, and additionally from information will questions.Notwithstanding, portrayal is the most colossal perspective withregard to huge information assurance. Two or three information secretstrategies exist with the most striking ones being gotten tocontrol and encryption.

II.RELATED WORK

The object of our project is we can protect a data from the attackers. We can detect a attackers easily. We can a added a sql attackers we can detect. We can added a servers to protect an owner and big data.

The information created is an impression of nature, it is delivered out of, along these lines can utilize the information escape frameworks to make sense of the internal activities of that framework. This has become a significant component in digital security where the objective is to ensure resources.

The framework of the model for the computer networkattack analysis is to analyze the network attacks in thevirus propagation model by establishing the relationship function of the framework daemon, so as to determinethe structure of the framework analysis unit, and to clarifythe mutual instructions and file relations of each analysisunit.Computer network attack under the virus propagationmodel adopts client/server mode to initiate computernetwork attack. The principle of network client/servicemode is that a host provides service (server side) and anotherhost receives service (client). A server as a hostusually calls a default port and listens. If a client has aconnection request on the port of the server, the correspondingprogram on the server will run automaticallyto respond to the client's request.

Alongside the information produced by IoT gadgets, the development of Bring Your Own Device (BYOD) has made associations powerless to different assault vectors. Every one of these gadgets produce information. In this way associations are beginning to hold onto BDA as a device in their cyber security approach. Examining the information that goes through the system is basic to secur the association. Be that as it may, a few organizations despite everything have reservations on utilizing huge information investigation as it will in general be a costly endeavor. BDA likewise will in general be an unpredictable field and requires mastery. Besides, workers are not happy with
individual data accumulated as this may include following client action. There are open difficulties of the most proficient method to separate the IoT framework information, individual information and touchy information and the security of every one of them utilizing Big information examination.
Bertino et al [4] introduced the security and protectionissues for large information concerning the secrecy, protection,and reliability. In information privacy, the difficultiesdistinguished were consolidating a huge number of accesscontrol strategies and authorizing control arrangements in largeinformation sources. Cyber security assignments, for example,client confirmation, get to control, and client checking is notedto be key in recognizing the dangers and halting them. Thecreator noticed that both security and protection can beaccomplished by utilizing trend setting innovations, forexample, cryptography. Mishra et al [5] analyzed security andprotection challenges related to Big information investigationfor ensuring database stockpiling and exchange log documents,and secure calculations in disseminating structures.
The Y. Gahiet al [6] featured the advantages of large information,examination and assessed security and protection challenges inBig information conditions utilizing different BDA devices, forexample, Hadoop, MapReduce, and HDFS. Security andprotection challenges related to Big information conditionswere likewise recorded as arbitrary circulation, security of largeinformation calculations, and access control. K. Abouelmehdi etal[7] analyzed large information developing issues of securityand protection comparable to the utilization of hugeinformation scientific apparatuses, for example, Hadoop. B.
Matturdi et al [8] displayed an audit of large informationsecurity and protection challenges while putting away, lookingand dissecting. In B. Nelson et al [9], the creators led an efficient writing audit covering security and protection for largeinformation by ordering approaches regarding secrecy,information respectability, security, information investigation,representation, information organization, and streamspreparing.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011017 |

III.LITERATURE SURVEY

Karthiban, M. K., & Raj, J. S , "Big data analytics fordeveloping secure internet of everything", Storage and processing of information is the major application of big data analytics. Internet of Everything (IoE) is the smart connection between people, data, things and processes. This paper studies the available frameworks used for developing secure Internet of Everything with big data analytics. Big data is a collection of data generated from the sensors embedded in the surrounding physical objects. This information can be used for analysis of the surroundings and development based on the inference. Internet of Everything uses this data for automation of the electronic equipment in the surrounding environment. However, with the increasing level of automation, the vulnerability to attack also increases. This paper presents a detailed analysis of big data analytics that is used for developing a secure internet of everything.

D. Rawat and K. Z. Ghafoor, "Smart Cities Cybersecurity and Privacy", Smart Cities Cybersecurity and Privacy examines the latest research developments and their outcomes for safe, secure, and trusting smart cities residents. Smart cities improve the quality of life of citizens in their energy and water usage, healthcare, environmental impact, transportation needs, and many other critical city services. Recent advances in hardware and software, have fueled the rapid growth and deployment of ubiquitous connectivity between a city's physical and cyber components. This connectivity however also opens up many security vulnerabilities that must be mitigated. Smart Cities Cybersecurity and Privacy helps researchers, engineers, and city planners develop adaptive, robust, scalable, and reliable security and privacy smart city applications that can mitigate the negative implications associated with cyber-attacks and potential privacy invasion. It provides insights into networking and security architectures, designs, and models for the secure operation of smart city applications.

N. Miloslavskaya and A. Tolstoy , "Application of big data, fast data, anddata lake concepts to information security issues",tToday we witness the appearance of some additional to Big Data concepts: data lakes and fast data. Are they simply the new marketing labels for the old Big Data IT or really new ones? Thus the key goal of the paper is to identify the relationship between these three concepts, giving special attention to their application to information security (IS) issues. The reason lies in the fact that volumes of IS-related information is one thing, but the real problem for securing enterprises' IT infrastructure assets is the speed with which things related to IS happen.

A. D. Mishra and Y. B. Singh, "Big data analytics for security and privacychallenges", the digitalization of our day-to-day activities has resulted in a huge volume of data. This data, called Big Data, is used by many organizations to extract valuable information either to take marketing decisions, track specific behaviours or detect threat attacks. The processing of such data is made possible by using multiple techniques, called Big Data Analytics, which allow getting enormous benefits by dealing with any massive volume of unstructured, structured and semi-structured content that is fast changing and impossible to process using conventional database techniques. However, while Big Data represents an immense opportunity for many industries and decisions makers, it also represents a big risk for many users.

This risk arises from the fact that these analytics tools consist of storing, managing and efficiently analyzing varied data gathered from all possible and available sources. The consequence is that people become widely vulnerable to exposure because of combining and exploring specific behavioral data. That is, it is possible to collect more data than it should have which leads to many security and privacy violations. Therefore, research community has to consider these issues by proposing strong protection techniques that enable getting benefits from big data without risking privacy. In this paper, we highlight the benefits of Big Data Analytics and then we review challenges of security and privacy in big data environments. Furthermore, we present some available protection techniques and propose some possible tracks that enable security and privacy in a malicious big data context.

B. Nelson and T. Olovsson, "Security and privacy for big data: A systematicliterature review", big data is currently a hot research topic, with four million hits on Google scholar in October 2016. One reason for the popularity of big data research is the knowledge that can be extracted from analysing these large data sets. However, data can contain sensitive information, and data must therefore be sufficiently protected as it is stored and processed. Furthermore, it might also be required to provide meaningful, proven, privacy guarantees if the data can be linked to individuals. To the best of our knowledge, there exists no systematic overview of the overlap between big data and the area of security and privacy. Consequently, this review aims to explore security and privacy research within big data, by outlining and providing structure to what research currently exists. Moreover, we investigate which papers connect security and privacy with big data, and which categories these papers cover. Ultimately, is security and privacy research for big data different from the rest of the research within the security and privacy domain? To answer these questions, we perform a systematic literature review (SLR), where we collect recent papers from top conferences, and categorize them in order

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011017 |

to provide an overview of the security and privacy topics present within the context of big data. Within each category we also present a qualitative analysis of papers representative for that specific area. Furthermore, we explore and visualize the relationship between the categories. Thus, the objective of this review is to provide a snapshot of the current state of security and privacy research for big data, and to discover where further research is required.

M. Paryasto, A. Alamsyah, B. Rahardjo, et al, "Big-data securitymanagement issues", big data phenomenon arises from the increasing number of data collected from various sources, including the internet. Big data is not only about the size or volume. Big data posses specific characteristics (volume, variety, velocity, and value -4V) that make it difficult to manage from security point of view. The evolution of data to become big data rises another important issues about data security and its management. NIST defines guide for conducting risk assessments on data, including risk management process and risk assessment. This paper looks at NIST risk management guidance and determines whether the approach of this standard is applicable to big data by generally define the threat source, threat events, vulnerabilities, likelihood of occurence and impact. The result of this study will be a general framework defining security management on Big Data.

IV. PROPSOED SYSTEM

In proposed system we can added a brute force attack and sql injection attack for showing a attacks. We can detect a attacks from a log records from the big data. From the server the owner can upload from the server .

The Internet of Everything connected device count may reach a hundred billion by the year 2020. Voluminous data will be generated by IoE. Several issues and challenges may be associated with such data with respect to information and communication technology (ICT). Big data analytics is essential for processing this generated data and to put them to use . Big Data is sometimes referred as facts that is collected regarding reality. The generation of these facts is done by sensors implanted in physical articles that surround us. Larger the information gathered, more it can be used for enhancing technology. Since monitoring and generating information is a continuous process, it is essential to keep these devices active and connected to the internet always to ensure uninterrupted updating of data to the server. This makes us vulnerable to attacks that can damage the safety and security of the information. Laney et al. related big data to three words namely Velocity, Volume and Variety (3 Vs). Variability, Veracity and Value of the data are also prominent properties of big data. It represents the rate of generation of data, amount of data generated and that data is available in different forms respectively. Most of this big data is stored on cloud platform. Even though this platform is supposed to be trustworthy and secure, several companies use this data to study the browsing and purchase habits of the user. This can be a major privacy concern for many users. The data is exposed to data leakage and security attack vulnerabilities. These data breaches affect the reputation of high-profile companies if not addressed properly.

Past territory displayed how security can be practiced with Big data makers considered the security issues related. They perceived the way that most redistribute database into the cloud. Appropriated processing at any rate, despite everything has various related with it. The target in was to find security in the cloud in order to venders. They saw that grouping in a particular request as the most huge security gives a cloud provider face right presently implies the protection of data against unapproved impedance or use. Trustworthiness would be the expectation of unapproved data change. Regardless, characterization is the most huge point of view with respect to enormous data protection. A couple of data mystery techniques exist with the most striking ones being gotten to control networks.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011017 |

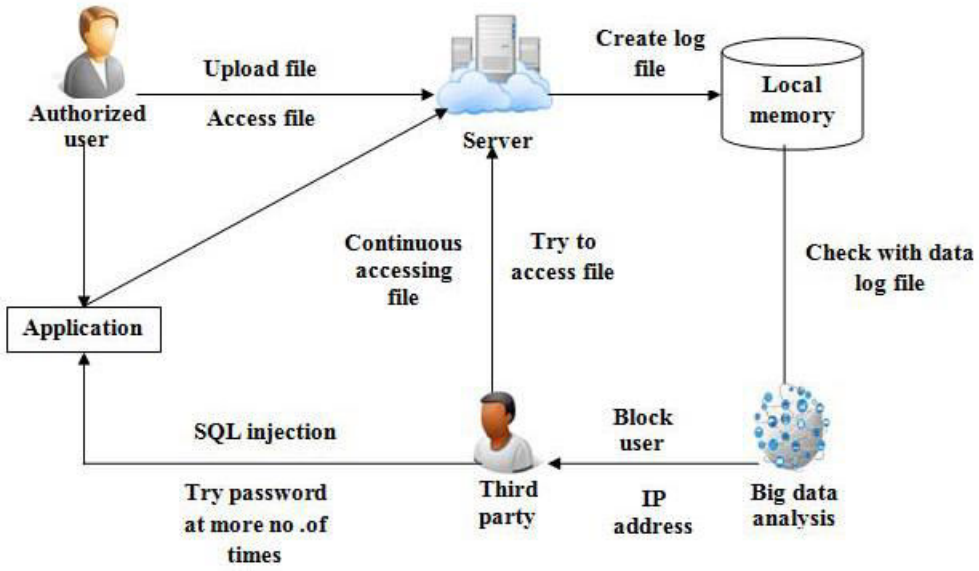

Fig1 . Architecture of the approach

In this project cloud user has a login in the database.Authorized user can upload a files in the server.Authorized user can also have a access file. Authorized user can send request to the application.Application have a continuous accessing file.Third party can attack a SQL injection.In third party can attack a more no of times in the application. Third party can also try a access a file. Big data analysis has a have a block users and ip address in the third party user. Server can have a log files in a local memory. Big data Analysis have check with data log files in the local memory.

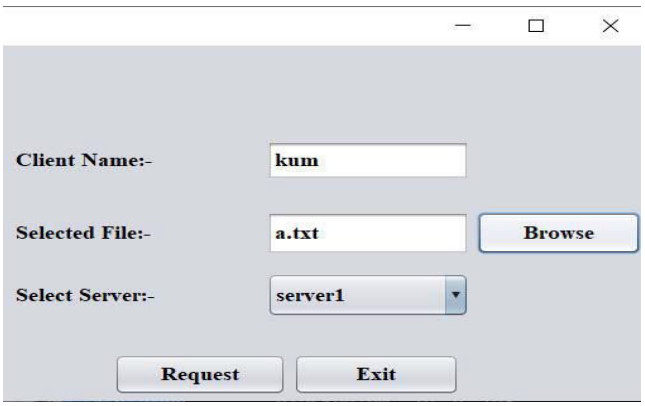

Fig. 2. Cloud prototype to detect DDOS attack

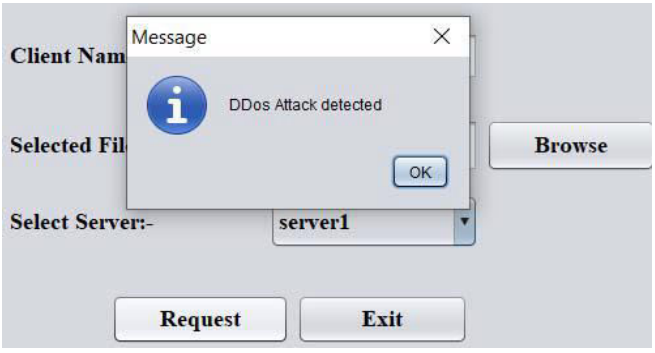

Fig. 3. Detection of DDOS attack

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011017 |

The above pictures shows how cloud environment is useddetect ddos ,brute force,sql injection attacks.The cloudprototype is used to execute the attacks.The servers are createdand client or user registration is done.Using the cloudprototype created the servers are attacked and the attacks areidentified successfully.

Attacks Categories	Original Instances	Percentage (Original)	Distinct Instances	Percentage (Distinct)
DoS	229,855	91.78%	23,570	80.23%
U2R	70	0.03%	70	0.24%
R2L	16,345	6.53%	3,056	10.40%
Probe	4,166	1.66%	2,682	9.13%
Total Attacks	250,436	100%	29,378	100%

Table 1. Original instances and distinct instances of four main categories of attacks

The conventional method can make a reasonable analysis of common network attacks, but the reliability of theanalysis is low under the virus propagation model. This paper proposes a new research method of computernetwork attack analysis based on the virus propagation model. Based on the relationship between the frameworkand daemon, the framework of the model for computer network attack analysis is set up, the attack analysis technologyof computer network is determined, and the construction of the model for computer network attack analysisis completed. Computer attack objects and computer attack process are analysed, and computer networkattack analysis is carried out. Using the coverage test and the uncertainty test, the parameters of the reliability calculationvariables are measured and the reliability calculation formula is replaced. It is concluded that the designed method ofcomputer network attack analysis is 47.15% more reliable than the conventional analysis method, and is suitable for thenetwork attack analysis under the virus propagation model.

V. CONCLUSION & FURTURE ENHANCEMENTS

The proposed strategy is executed to tackle thesecurity and protection issues of huge information. Right now,ongoing examination works in digital security, according tolarge information and feature how huge information is ensuredand how Big information can likewise be utilized as a devicefor digital security. Hence the task infers that through thisframework the assaults and log record are distinguishedindependently. Utilizing huge information, the assault andsquare the IP address is recognized and have actualized a BigData based brought together log investigation framework todistinguish the system traffic happened by aggressors throughDDOS, SQL Injection and Bruce Force assault. The logdocument is consequently transmitted to the incorporated cloudserver and bi g information is started for investigation process.

As a future work, we can add a encryption techniques for owner1 to owner 2 while sharing a data. We can give high security for the attackers cannot attacks a information. We can give a many security to protect a networks.

REFERENCES

[1] D. Laney, “3d data management: Controlling data volume, velocity andvariety,” META Group Research Note, vol. 6, no. 70, 2001.
[2] N. Miloslavskaya and A. Tolstoy, “Application of big data, fast data, anddata lake concepts to information security issues,” in Future Internet of Thingsand Cloud Workshops (FiCloudW), IEEE International Conference on, pp.148–153, 2016.
[3] Ravindra Changala, “A Survey1 on Clustering Techniques to Improve Energy Efficient Routing in Wireless Sensor Networks” in International Journal of Applied Engineering Research, 10(58), pp.-1-5,2015.
[4] E. Bertino, “Big data-security and privacy,” in Big Data (BigDataCongress), 2015 IEEE International Congress on, pp. 757–761, 2015.
[5] A. D. Mishra and Y. B. Singh, “Big data analytics for security and privacychallenges,” in Computing, Communication and Automation (ICCCA), 2016International Conference on, pp. 50–53, 2016.
[6] Y. Gahi, M. Guennoun, and H. T. Mouftah, “Big data analytics: Securityand privacy challenges,” in Computers and Communication (ISCC), 2016 IEEESymposium on, pp. 952–957, 2016.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011017 |

[7] K. Abouelmehdi, A. Beni-Hssane, H. Khaloufi, and M. Saadi, "Big dataemerging issues: Hadoop security and privacy," in Multimedia Computing andSystems (ICMCS), 2016 5th International Conference on, pp. 731–736, 2016.
[8] Ravindra Changala, "Secured Activity Based Authentication System", in Journal of innovations in computer science and engineering (JICSE), Volume 6, Issue 1, Pages 1-4, September 2016.ISSN: 2455-3506.
[9] B. Nelson and T. Olovsson, "Security and privacy for big data: A systematicliterature review," in Big Data (Big Data), 2016 IEEE International Conferenceon, pp. 3693–3702, 2016.
[10] N. Miloslavskaya, A. Tolstoy, and S. Zapechnikov, "Taxonomy forunsecure big data processing in security operations centers," in Future Internetof Things and Cloud Workshops (FiCloudW), IEEE International Conferenceon, pp. 154–159, 2016.
[11] Ravindra Changala, "Retrieval of Valid Information from Clustered and Distributed Databases" in Journal of innovations in computer science and engineering (JICSE), Volume 6, Issue 1,Pages 21-25, September 2016.ISSN: 2455-3506.
[12] M. Paryasto, A. Alamsyah, B. Rahardjo, et al., "Big-data securitymanagement issues," in Information and Communication Technology(ICoICT), 2014 2nd International Conference on, pp. 59–63, 2014.
[13] Ravindra Changala, "Intrusion Detection System Using Genetic Algorithm" published in International Journal of Emerging Trends in Engineering and Development [IJETED], Impact Factor 2.87, ISSN NO:2249-6149, Issue 2,Vol. 4 May 2012.
[14] Karthiban, M. K., & Raj, J. S. (2019). BIG DATA ANALYTICS FORDEVELOPING SECURE INTERNET OF EVERYTHING. Journal ofISMAC, 1(02), 129-136.
[15] R. Clarke, "Quality assurance for security applications of big data," inIntelligence and Security Informatics Conference (EISIC), 2016 European, pp.1–8, 2016.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details