International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

Design and Analysis of Multisource Logs Forensic with Lock Technique for Cloud Security Enhancement

Kalyan Bamane, Dinesha H.A

Abstract: ***Multisource cloud log forensics (MCLF) strengthens the investigation method by means of detecting the malicious behavior of hackers thru deep cloud log evaluation. But, the accessibility attributes of cloud logs thwarts accomplishing the purpose to analyze cloud logs. Accessibility consists of the productions of cloud log get admission to, selection of suitable cloud log file, cloud log information integrity, and cloud logs trustworthiness. Hence, forensic investigators of cloud log files are dependent on cloud provider vendors (CSPs) to get entry to of diverse cloud logs. Accessing cloud logs from outside the cloud without depending at the CSPs is once more a hard, whereas the boom in cloud assaults has improved the need for MCLF to research the malicious activities of attackers. Criminals are easily hiding incriminating files within the cloud system and altering the log contents. Hence lock mechanism has been added to MCLF technique. This paper reviews the MCLF with lock technique and highlights diverse challenges and issues involved in examining cloud log data. The logging mode, the importance of MCLF, and cloud multisource-log-as-a-service are introduced. The MCLF security necessities, weakness points, and experiments are recognized to tolerate altered cloud log susceptibilities. This paper represents the design and analysis details of MCLF with Lock technique.***
Keywords: Cloud, Forensic, Logs, Multisource, Security

I. INTRODUCTION

Every event taking place in an IT environment or employer community is recorded with exceptional entries in a log record. The exercise of recording log documents is referred to as as logging [Chuvakin et al. 2013]. The log report offers proof regarding history of occasions taking place in the IT gadget and network for a distinctive duration. for instance, a community admin can find out for the network bandwidth usage for specific time period by investigating the network logs. Likewise, utility developers utilize utility logs to perceive and fasten insects internal a program code. each entry inside the log document describes the sizable facts associated with a selected event for the duration of the log file is generated. all through preliminary degree, the log record is applied for trouble shooting [Flegel 2002]. Now, the log report applied for extra functional offerings together with machine and community tracking, enhancing the overall performance of the gadget and network, recording person activity, and examining malicious behaviour [Kent and Souppaya 2014] and so fortienceh.

Manuscript published on 30 September 2019.
*****Correspondence Author(s)
Kalyan Devappa Bamane, Department of Computer Science and Engg. VTU Belgavi India . Email:kalyandbamane@gmail.com
Dr.Dinesha H.A, Department of Computer Science and Engg. VTU Belgavi India . Email:sridini@gmail.com

In modern scenario, Logs are essentially used for protection purposes due to continuous attacks at the IT gadget and network [Zuk 2011]. The logs used to file attackers' activities and community administrators investigate those assaults by means of investigating log report facts [Mao et al.2014]. In large agencies, diverse sorts of log files are created on kind of gadgets that involve the concern of powerful control of logs because of shortage of resources. To address the log management difficulty, businesses have commenced to migrate to cloud computing via using cloud logging offerings known as log-as-a-service [Saurabh and Beedgen 2014]. Log documents created on diverse organizational assets are shared to the cloud for storage and investigation using cloud garage assets and cloud log servers. Correspondingly, agencies usually run their packages in computational clouds that also want logging to examine malicious sports whilst observed. Cloud logging incorporates cloud application logs, cloud firewall logs cloud machine logs, cloud network logs, and so on. in this paper, the phrase "cloud log" is refers to logs created inside a cloud computing. these days, assaults on cloud computing are going on more frequently, which creates a anxiety among customers and corporations concerning the first-rate way to maintain their records secure from various attackers [Khan et al.2014]. Cloud log documents document diverse activities happening in the gadget and network. it's far used to examine various assaults [Vrable et al. 2012].

A nice choice is to discover the cloud log files for malicious behaviour by means of analyzing them using log enquiry methods [Lin et al. 2013; Wei et al. 2011]. The technique of investigating cloud log documents in cloud computing or through third-party examinations services is referred to as cloud log forensics (CLF) [Thorpe et al. 2012].

This paper has been organised in following manner. Section 2, represents the logs based forensic investigation details. Section 3, represents the design details of proposed MCLF with lock technique. Section 4, presents the analysis of proposed MCLF with lock techniques. Section 5, concludes the paper with future enhancements.

II. SYSTEM STUDY

This section describes the cloud forensic process, access log format, cloud log sample, and attributes of sample cloud-side logs. There are various systems that are used for attack identification and forensic IDS in cloud system.

The current cloud forensic process undertakes the following steps which can be integrated in cloud forensics considering its various service and deployment models. The figure 1 represents the current cloud forensic process which has the following steps for performing forensic reporting after successful gathering, pre-processing, investigation and analysis.

1. Detection of malicious activity
2. Logs Gathering and Integration
3. Logs pre-processing and sequencing
4. Investigation and Analysis
5. Reporting & Presentation

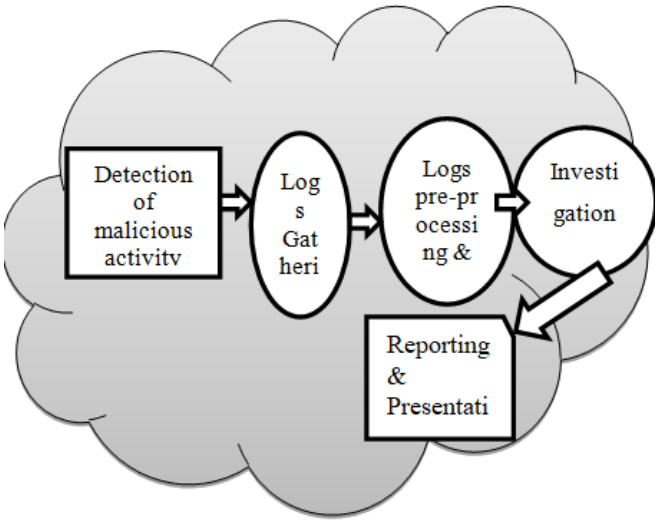

Figure 1: Present cloud forensic process

The collected logs has presented in its predefined format for further analysis. The access log format fields and meaning represented in table 1. Cloud log format is varies and its sample log presented in table II. Main attributes of cloud system logs are mentioned in table III

Table I Meaning of the Access Log Format

Fields	host	rfc931	username	date:time timezone	request	Status code	bytes
Meaning	IP address of the HTTP client which rises HTTP resource request	Identifier used to identify the client	User name used for authentication	Date and time stamp of the HTTP request	HTTP request containing (a) HTTP method = GET(b) HTTP request resource = index.html, and (c) HTTP protocol version = 1.0	Numeric code used to tell about the status of HTTP request i.e. success or failure	Numeric field used to highlight number of bytes of data transferred during the HTTP request

Table II Sample cloud log

Log attributes	Sample log data
eventID	10
userName	Bob
eventTime	2019-03-22 22:10:59
eventName	CreateUser
eventSource	iam.amazonaws.com
awsRegion	india-east-1
sourceIPAddress	192.168.14.1

Table III Attributes of sample cloud system logs

Sr	Log Attribute	Description
1	Timestamp	Timestamps in the format YYYYMMDDHHMMSS.MLS
2	User ID	ID of the user
3	URI	Uniform Resource Identifier – this is usually the tail-end of the URL
4	Source IP	Source IP address of the source that generated data request
5	Log record entry type	Type of log entry indicates the types of logs and facilitates analysis. Some of the log entry types are: E – Email A – API L – Login Q – Search query S – Download attachment

Next section describes the proposed system design.

III. SYSTEM DESIGN

This section represents the various types of logs and its sources for investigation. It also presents the proposed MCLF with Lock technique design details for multisource cloud log forensic (MCLF) with lock techniques. The MCLF collects the logs from various types of log sources such as cloud firewall, cloud application, system, security, cloud network, cloud web server, audit, and virtual machine etc. as described in table iv[2].

Table IV. Various Types of Logs sources

Types of log	Description
Cloud Firewall log	Real-time and up-to-date information of the traffic logs produced by the Firewall
Cloud Application log	Logs that are recorded by a cloud application. Cloud Application developers are responsible to specify what, when, and how to log through a cloud application implementation on a system.
System log	System logs are produced by an operating system which contains information concerning system events, drivers, operation, device change, and many more.
Security log	Logs comprise security related evidence to identify malicious behavior found in the system/network. For instance, file quarantines, malware detection, time of malicious detection, and etc.
Setup log	Setup logs capture the events occur during performing the installation of an application.
Cloud Network log	Cloud Network log is a log file that contains network related events, that is, description of the event, priority, time occurrence and much more.
Cloud web - server log	Cloud Web-server log records all happenings occur on the particular web-server such as IP address, date & time, access time, request method, and object volume etc.
Audit log	Audit log comprises user unauthorized access to the system & network for examining its accountabilities. It contains user login information, destination addresses, and timestamp etc.
Virtual machine logs	A file that comprises records of every event executed on a virtual machine (VM).

All these source details are inevitable for forensic investigation due to high intensity attacks. Many attacks are also carried to modify the log details and to hide incriminating files in cloud environment, hence the investigation becomes challenging and imperfect [21]. Many anti-forensics techniques are also impact in cloud log analysis. Hence, proposed MCLF with lock techniques is inevitable and applicable in addressing these issues. The proposed design illustrated in fig 2. Where-in which, MCLF collect logs from cloud forensic module which linked to multiple sources of management and virtualization layers. All the logs gathered and collectively stored in database for further investigation in lock mode. Any anti-forensic activities which are trying to alter the contents will be resisted by lock mechanism. Unless proper unlock key, it cannot be modified. Only authenticated/local source can add logs by proving its identity through unlock key.

Once the MCFL is ready, it can be utilised for further integration, investigation and analysis against the malicious activity detection. During investigation, log indexing, normalization, correlations, and sequencing carried out. At the end reporting will be done via sms, email, whatsapp and any other communication media tools opted by customer.

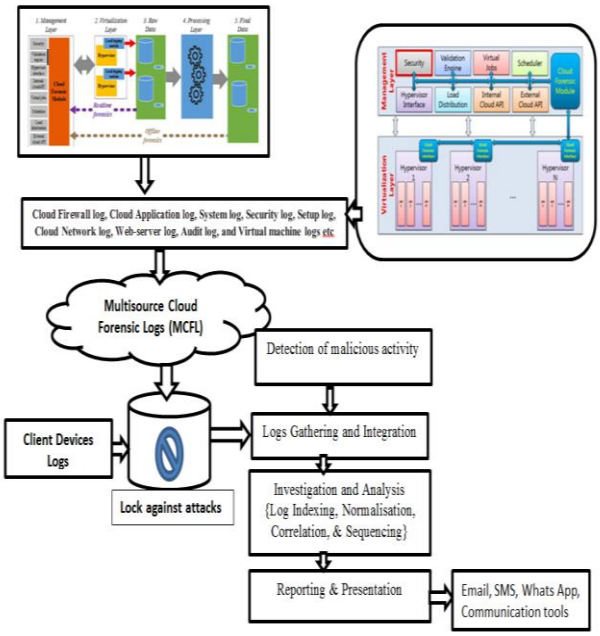

Figure 2: Proposed multisource forensic logs with lock techniques

Next section analyses the proposed techniques.

IV. SYSTEM ANALYSIS

This section describes the proposed MCLF with Lock technique analysis. The experimentation conducted by following common methodology [21]. It was also evaluated the strength by checking the probability of missing logs information.

Experimentation

The logs used for investigation are more or less labelled into csp facet and patron facet logs. The cloud logs are divided into 3 varieties of logs meta-information server, administrative, and not unusual errors. The CSP-side logs comprise server, audit and occasion logs. Administrative server logs incorporate administrative and safety logs. Likewise there are tracers, unload logs, and common mistakes log files. client aspect logs comprise all syslog's from AIX, home windows, Linux, and Solaris patron. For brevity, we've used the aforementioned logs to extract the document device state at special intervals of time. An instance of the layout of cloud-aspect logs is shown in table 3, that's founded on logs used by Salesforce [20]. There are other cloud storage services, consisting of SkyDrive, and Dropbox.15. Experimentation conducted on google pressure, created a digital machine for gaining access to this sevice. The digital device became created using VMware computing device model 10. The virtual device was created on a windows platform machine with 30GB potential. The process screen of Sysinternals software became used to gather records related to all modes of modifications that occurred at the same time as using the cloud service. Technique screen continues account of all the sports that show up with the cloud carrier proper from its set up. On this route, it received report device activity logs. Every process is identified by way of a unique manner identity (PID) by way of the method screen. Procedure display has a choice to clear out specific sorts of activities. on this, it best monitored the document device activities and registry entries

Research Methodology

The research methodology used to execute are as follows:

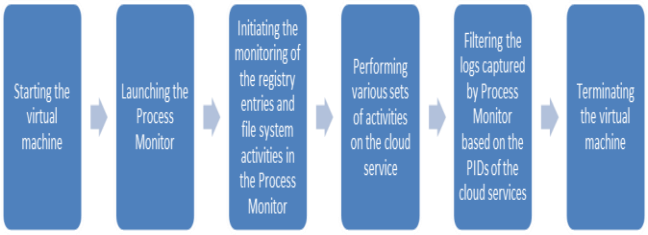

Figure 3: Research Methodology

After executing above methodology steps, FTK Imager version 3.2 from AccessData enterprise turned into beneficial for analysis of the snap shots captured using the PM. FTK Imager receives a full utilization of Google force. Alongside csp facet FTM imager information, a consumer side log which includes system logs, network logs, firewall logs, VM logs have been used for research.

The log connection engine connects inter-associated probably risk events. while these chance activities are pooled, it's miles very probably that they offer proof about a path of action that directed to digital compromise. It connects actionable occasions for each the CSP aspect and client facet. The linked events now not most effective provide perception into the possibly device misuse events but additionally pinpoint the possibly chance areas. the relationship engine also has the competence to kind and clear out log statistics. The final experimented data set logs with false positive and false negative of CSP side and customer side are represented in table V.

Table V

a.

Sl. No.	Size of log data sets (CSP side)	Percentage	
		False positive	False Negative
1	2500	13.1	10.2
2	5000	12.2	9.5
3	10000	10.2	5.9
4	25000	9.1	3.9
5	50000	7.4	2.2
6	100000	6.2	1.8

Sl. No.	Size of log data sets (CSP side)	Percentage	
		False positive	False Negative
1	2500	11.1	9.2
2	5000	10.2	8.5
3	10000	8.2	7.9
4	25000	7.1	4.9
5	50000	5.4	3.2
6	100000	4.1	2.9

b.

(a) CSP side log (b) Customer side log

Result and Discussion

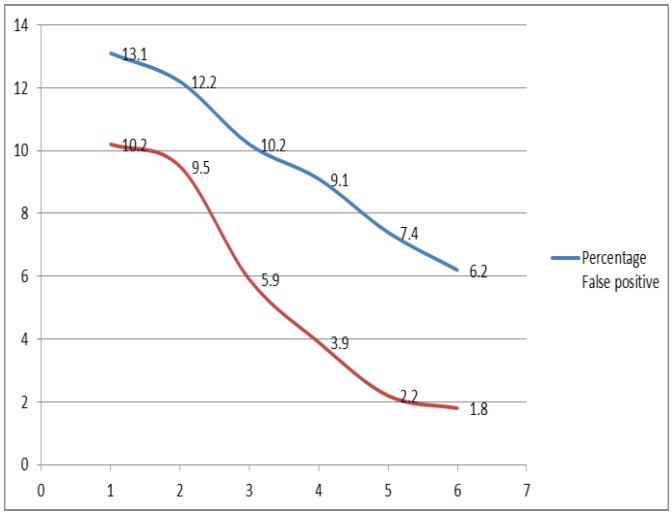

Figure 4 (a) Result CSP side: False positive (FP) and false negative (FN) rates observed for datasets of varying sizes

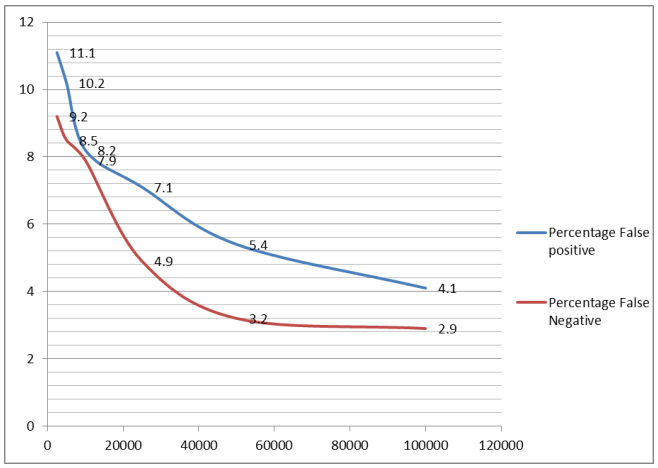

Figure 4 (b) Result Customer side: False positive (FP) and false negative (FN) rates observed for datasets of varying sizes

For this reason, it was conducted two types of evaluations to identify false negatives and false positives. False negatives and positives were calculated to be 1.8% and 6.2% respectively for the concluding investigational dataset. A graph representing the percentage of false positives and false negatives witnessed for different sizes of dataset is shown in 4a. The graph 4b represents the customer side logs concluding investigational dataset describes the percentage of false positives and false negatives witnessed for different sizes of datasets.

Evaluation of strength of proposed system

Probability of missing logs information : with reference to table 4, there are nine varieties of logs , assuming the common logs in 3 different sources, the probability of missing log can be calculated using sample space 'S'. Let S is the sample space. Two results of the event are success S and failure F. To break this one has to succeed to delete in three different sources. Thus, Sample Space S = {SSS, SSF, SFS, SFF, FSS, FFS, FFF, FSF} n (S)=8 where n (S) is the cardinality of set S. Assuming , i. Success and failure probability at different sources is independent and ii. The Probability of success at different sources is 'p' Then Success of deleting logs in all the source / missing logs probability of the event is SSS, hence defined as P (E) and is equal to P (E) = p3 Failure in breaking the multisource logs system is 1-P (E) = 1- p3.

For Instant, If probability of success at given source p=0.1 (say) then the probability of deleting the multi-source (three source) logs is 0.001. Hence the strength of forensic investigation has been achieved

V. CONCLUSION

The proposed MCLF coll the logs from various types of log sources such as cloud firewall, cloud application, system, security, cloud network, cloud web server, audit, and virtual machine etc. the proposed MCLF with lock techniques have been designed to full fill requirements of precise forensic investigation by focusing on multi-source logs. even if hackers deleted the logs details, multisource logs helps to predict and analyze the attacks intensity. MCLF is also been designed with lock mechanism so that unauthorized alteration of log file and hiding incriminating files in forensic logs can be addressed. the design and analysis of proposed MCLF with lock has been described

FUTURE ENHANCEMENT

It In future, to perform forensic investigations in an electronic system includes capturing and linking multiple streams of information. The streams comprise a network stream and a storage stream. The network stream contains a record of network activities. The storage stream contains a record of storage activities. Hence, this network and storage coupling forensic technique for cloud security enhancement and forensic techniques for mobile cloud security augment will be focused.

REFERENCES

1. Chuvakin, K. Schmidt, and Chris Phillips. 2013. Logging and Log Management: The Authoritative Guide to Understanding the Concepts Surrounding Logging and Log Management. Syngress, 460 pages.
2. SULEMAN KHAN, ABDULLAH GANI et. El , Cloud Log Forensics: Foundations, State of the Art, and Future Directions, ACM Computing Surveys, Vol. 49, No. 1, Article 7, Publication date: May 2016.
3. Flegel. 2002. Pseudonymizing unix log files. In Infrastructure Security. Springer, Berlin, 162–179.
4. Kent, S. Chevalier, T. Grance, and H. Dang. 2006. Guide to integrating forensic techniques into incident response. NIST Spec. Publ. (2006), 800–886.
5. H. Mao, C. J. Wu, E. E. Papalexakis, C. Faloutsos, K. C. Lee, and T. C. Kao. 2014. MalSpot: Multi2malicious network behavior patterns analysis. In *Advances in Knowledge Discovery and Data Mining*. Springer, Berlin, (2014), 1–14.
6. Saurabh and C. Beedgen. 2014. Master your data continous intelligence. (2014). Retrieved November 16,2015, from https://www.sumologic.com/.
7. Han, M. Shiraz, A. Gani, M. Whaiduzzaman, and S. Khan. 2014. Sierpinski triangle based data center architecture in cloud computing. *J. Supercomput.* 69, 2 (2014), 887–907.
8. Vrable, S. Savage, and G. M. Voelker. 2012. BlueSky: A cloud-backed file system for the enterprise. In *Proceedings of the 10th USENIX Conference on File and Storage Technologies*. San Jose, CA, USA, 19–19.
9. Sundareswaran, A. C. Squicciarini, and D. Lin. 2012. Ensuring distributed accountability for data sharing in the cloud. *IEEE Trans. Depend. Secure Comput.* 9, 4 (2012). 556–568.

10. Thorpe, I. Ray, T. Grandison, and A. Barbir. 2011a. The virtual machine log auditor. In *Proceeding of the IEEE 1st International Workshop on Security and Forensics in Communication Systems*. 1–7.
11. Ruan, J. Carthy, T. Kechadi, and M. Crosbie. 2011. Cloud forensics. *Advances in Digital Forensics VII*. Springer, Berlin, 35–46.
12. Birk. 2011. Technical challenges of forensic investigations in cloud computing environments. In *Workshop on Cryptography and Security in Clouds*. Zurich, Switzerland, 1–6.
13. D. Birk and C. Wegener. 2011. Technical issues of forensic investigations in cloud computing environments. In *Proceeding of the IEEE 6th International Workshop on Systematic Approaches to Digital ForensicEngineering (SADFE)*. Washington, DC, USA, 1–10.
14. Ellis. 2013. IBM Operations Analytics-Log Analysis. (2013). Retrieved November 16, 2015, from http://www-03.ibm.com/software/products/en/ibm-operations-analytics—log-analysis.
15. Burton. 2014. Real-time log management and analytics at any scale. (2014). Retrieved November 16, 2015, from https://logentries.com/.
16. Williams. 2013. Loggly, a Splunk Competitor, Raises $10.5m for Cloud-Centric Approach to Log Management.(2013). Retrieved November 16, 2015, from http://techcrunch.com/2013/09/03/loggly-a-splunkcompetitor-raises-10-5m-for-cloud-centric-approach-to-log-management/.
17. Rafael. 2013. Secure log architecture to support remote auditing. *Math. Comput. Model.* 57, 7 (2013), 1578–1591.
18. H. Beaver. 2015. Lessons on Efficient Log Analysis from Monex Insight. Case Study Report. Loggly Research. 3 pages. https://www.loggly.com/blog/lessons-efficient-log-analysis-monex-insight/.
19. R. Wyatt. 2009. Mission: Messaging: Circular Logs Vs Linear Logs. (2014). Retrieved November 16th, 2015 from http://www.ibm.com/developerworks/websphere/techjournal/0904_mismes.html.
20. Muhammad Naeem Ahmed Khan and ShahWali Ullah, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Islamabad, A log aggregation forensic analysis framework for cloud computing environments, Computer Fraud & Security, July 2017.
21. A. Pătrașcu, V.V. Patriciu, Logging for Cloud Computing Forensic Systems, INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROLISSN 1841-9836, 10(2):222-229, April, 2015.

AUTHORS PROFILE

Mr. Kalyan Bamane is serving DYPCOE Akurdi, Department of Information Technology as a Assistant Professor. He had born in 1981 at Ramanandnagar Sangli District. His schooling finished at Islamapur place in Marathi Medium (1999). He had completed B.E. in Computer Science and Engg. at RIT Sakhrale(2006). He had finished M.E. in Computer Engineering at DYPCOE, Akurdi (2011), He is pursuing his Doctorate of Philosophy in Computer Science and Engg. at Vishweshwaraya Technological University (2015-till date). He has national and international professional membership in Institution of Engineers, Indian Society for Technical Education, International Association of Computer Science and Information Technology, Institute of Research Engineers.
He had done more than 20 international and national research publications and presentations. He had participated and contributed in many national and international research and innovation platforms. He has executed more than rs.1 Lakh research grants funded by BCUD in the area of his expertise such as TTS, and email service etc. His research interest includes Cyber Security, WSN, Computer Networking, etc. He has got recognition for his vision in the field of teaching, research, innovation and Invention.

Dr. Dinesha H. A. is serving SGBIT, Belagavi as a Professor and HOD Computer Science and Engg., He has also Served i) Jain College of Engg., Belagavi as Associate Professor ii) DRDO-DIAT (DU), Ministry of Defense as an Officer In Charge Data Centre iii) PES University (PESIT), Bangalore as a Assistant Professor and Research Scholar, iv) BAE, Bangalore as a Lecturer v) VMware QA R & D as aQA Engineer. He had born in 1985 at Hagare, Belur (Tq), Hassan District. His schooling finished at Hagarerural place in Kannada Medium (2001). He had completed Diploma in Computer Science at Smt.L.V.(Govt.) Polytechnic, Hassan (2004) and B.E. in Information Science and Engg.. at Malnad College of Engineering (2007) , Hassan. He had finished M.Tech. in Software Engineering at Rashtria Vidyalaya College of Engg., Bangalore (2009), He has awarded Doctor of Philosophy in Computer Science and Engg. at Vishweshwaraya Technological University (2012-2016). He has national and international professional membership in Institution of Engineers, Indian Society for Technical Education, International Association of Computer Science and Information Technology, Institute of Research Engineers and Doctors, International Association of Engineers and Scientists and European Alliance for Innovation etc.
He had done more than 50 international and national research publications and presentations. He had participated and contributed in many national and international research and innovation platforms. He has been serving govt. bodies as a Technical Advisor, Teaching Quality Improvement Consultant, Information and Communication Tool Expert, Cyber Security Analyst, Research, Innovation and Invention Key Consultant. Major highlights are Virtualization and data center consultant DRDO, TEC member for Belagavi Smart City, Media Certification and Monitoring Committee for MLA and MP Elections, and Cyber Crime Analyst to cybercrime dept.. He has executed more than 2 crores research grants funded by DRDO in the area of his expertise such as digital forensic (KCTU), cyber security, UTM (DIAT-DRDO),server virtualizations, and email service etc. His research interest includes Cyber Security, Data Science, Decoding Ancient Technology; Computer Networking, and artificial intelligent etc. He has got recognition for his vision in the field of teaching, research, innovation and Invention.