Data Centric Access Security- an Improvised Technique to Achieve Security in Cloud Environment

Smita Sharma, R.P.Singh

Abstract - Cloud computing innovation has now turned into a good substitute to traditional computing technological developments This advancement of technology offers a new concept of a charge-per-use access resource model based primarily on virtualization technology. Because of the numerous advantages they offer, cloud computing concepts are achieving rapid adoption. This would comprise price-efficiency, time involved, and effective use of assets in computation. Given these advantages, there are many barriers to the widespread acceptance of this emerging technology, particularly data privacy and security issues. In extension to the conventional security hazards faced by internet-connected computer systems, cloud systems have relevant privacy and security problems due to the virtualization and multi-tenancy environment of the cloud. A further research target is focused on principles of Trust Computing (TC). Although these techniques offer users with mechanisms for evaluating and assessing security, they do not yield enough controlling functionality for users. In addition, Data Centric Security (DCS) is an evolving strategy intended to protect the data itself against migration towards the cloud.

Keywords : Cloud Computing , Data Security , Data Centric Security, Encryption ,Trusted Computing , CRT .

I. INTRODUCTION

Security is one of cloud computing's main user concerns towards the adoption of cloud computing. Cloud Service Provider (CSP). While this is typically handled on the basis of contractual or Service Level Agreements (SLA), the CSP may be able to obtain the information as well as give it to third parties. However , one must believe the CSP to enforce the access control rules for certain users that are decided by the information holder. In inter-cloud environments , where information can migrate from one CSP to another, the situation turned into more complicated. Users might lose their control over the data. This condition tends to improve strategies to data protection and implementing a data approach where data is self-protected wherever it resides. Encryption has been the most widely adopted cloud data security technique. Encrypting data prevents undesired accesses. It can be implemented efficiently through the new emerging data centric approach. In conventional data security strategies, the database that stores the data provides required security. The mechanisms used to improve security of the data and to handle the protected data are managed by the database administrators. This type of technique can be defined as a system-centered strategy that is not adequate in the insecure trustworthy cloud environment to protect client data. It is assumed that a data-centric strategy is more adequate and suitable for cloud based services. The data centric security illustrates that the security enforcement focus is all over the data.

For cloud computing, the data-centric protection approach is to protect data from within as per the their quality and identification to ensure maximum data security at all points of the data life cycle, irrespective of the context in which the data is saved[4].

II. CLOUD COMPUTING: ISSUES AND CHALLENGES

In cloud computing services, customers are focused on moving their sensitive data and applications from their own private computing environments to a cloud environment which is shared by different users and which is commonly accessible via a public network. Cloud computing security concerns are primarily associated to the key elements of software over which cloud computing relies. These elements includes the following :

A.Cloud Web services and Applications

These are by far the most commonly used applications for using various cloud services.

B. Virtualization

Behind the presence of cloud computing virtualization is the core state of the art technology. Both PaaS and SaaS are based on IaaS-level infra-structure virtualization.

C.**Cryptography**

At present, these methods are the most common approaches for achieving a reasonable degree of cloud computing security standards.

Any defined vulnerability for the above three core computer technology components can therefore be regarded as cloud computing systems limitations. There are several significant security problems and vulnerabilities that emerge due to the essence of cloud computing which are as follows :

A .Unauthorized access to the management interface

Management interfaces are normally open to authentic users and potentially unauthorized attackers via public networks in cloud computing, whereas traditional data centers are generally only accessed explicitly or via private networks through licensed administrators. Data access, although, is primarily controlled through a web application or software technology, so the cloud data system tends to be prone to these technologies weaknesses.

B. Data recover issue[5]

Because of the existence of hardware-level virtualization and synchronization of cloud services, space and storage units that former customers leased can be reassigned to new clients. Such new clients may be able to restore data from these storage and memory areas that may hold confidential data from former customers.

Revised Manuscript Received on November 05, 2019.

Smita Sharma*, Department of Computer Science & Eng, SSSUTMS, Sehore India,, sharmasmita34@gmail.com

R.P.Singh , Vice Chancellor SSSUTMS, Sehore, India

C. Virtual machine (VM) prototype image vulnerability[7, 6]

Cloning a prototype image of a fully customizable VM typically creates a new VM because it reduces effort and time. Some clients will therefore offer similar configurations for VMs. By becoming a cloud user with administrative privileges, an intruder may access information about cloud service template images. After the intruder has obtain access rights to the images of the template, he / she may check for deficiencies in the images that other clients may also utilize.

D. Data penetrate possibility [7] Another weakness related to VM prototype representations is that cloud providers can utilize prototypes that other users have built for new clients. Such models may include hidden wormhole generated by an intruder intending to be a customer and allowing the intruder to use virtual machines of other clients.

E. Injection weekness [7]

Because almost all the cloud services utilize services of web application, harmful codes can be entered into a cloud platform by using the loopholes in such services of web application to obtain access to those web servers through the use of these programs.

F. Challenges in security measurement and monitoring [7]

Consumers need to be able to detect and track their cloud services and assets ' security condition. Nonetheless, offering cloud customers with such functionality remains a challenge as the conventional standard resources available are still not appropriate for the cloud computing framework[7]. Because the cloud framework has complex and efficient centralized infrastructure that can include multiple cloud service providers, the cloud infrastructure requires new decentralized monitoring technologies to meet this complexity**.**

H. Digital key management and random numbers difficulty[10] - Different types of keys and random numbers are important for cryptographic functions in a cloud plate form. Managing and storing various keys in a cloud computing framework are critical aspects as there is no complete physical separation between different clients ' storage facilities. Effectiveness in random number generation relies more on the hardware clock that is utilize by the generator of random numbers. In a cloud computing platform, where a variety of cloud users used the identical generation tools concurrently in distant sessions, an absence of such consistency can be encountered.

I. Issue of Cloud interoperability[8] This issue is how different cloud providers enable data owners to easily transfer their data back to their available local assets from one cloud provider to another as necessary. A database owner can stay on a certain cloud provider without compatibility between cloud providers and can not probably shift to other cloud providers.

G. Monitoring patterns of activity [9] A cloud user's activity trends can be studied either on the identical cloud by other clients or by the cloud service provider. This analysis may be a phase ahead concerning a safety threat but it may also be utilized to uncover business operations that could not be identified in normal circumstances.

III. DATA INTEGRITY AND PRIVACY PROTECTION FROM CLOUD SERVICE PROVIDERS

The following two approaches are widely used by cloud providers to protect customer data :-

A. Trusted Computing (TC)

The IT group, specifically the Trusted Computing Group (TCG), is trying to build a range of techniques to ensure that computer systems follow the optimal mode of operation[12]. The modern approach allow consumers to determine cloud providers ' reliability through the implementation of a combination of hardware and software techniques. Remote server authentication is a techniques which helps customers to certify hosts. The TCG is attempting to build their result that depends on the establishment of a standardized hardware unit.

B. Data Centric Security (DCS)

Data security approaches for the cloud computing model are categorized according to two principles in the DCS approach: the first category is dependent on the level at which the protection is given, and the second category depends on who is liable for ensuring the security.

In Fig 3.1, the degrees at which protection functions can be given throughout reference to information are outlined in data-centric security.In Fig 3.2 different thresholds of system-centric security are highlighted. Generally , approaches focused on providing security beyond the data layer are categorized as system-centric security.

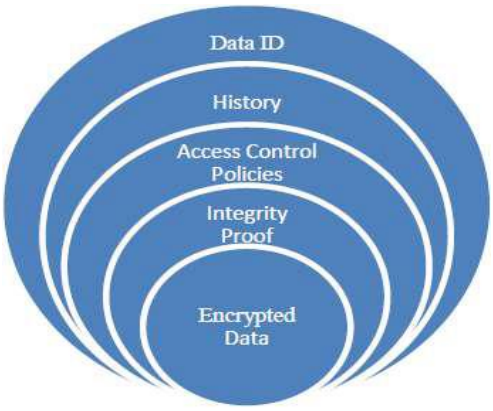

Fig 3.1: Data Centric Security

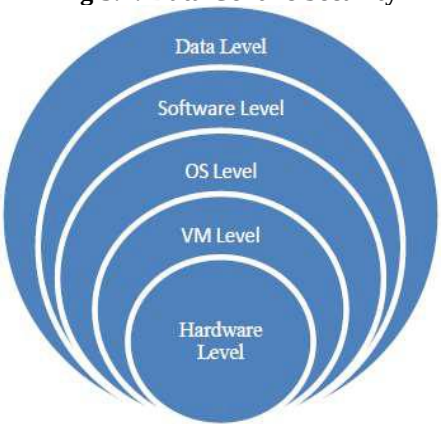

Fig 3.2: System Centric Security

Cloud computing has three levels of responsibilities for privacy and security that are as follows:-

A.Service provider level

In this point, the cloud provider provides security and sponsorship.

B. Trusted Computing level

Security is supported and funded by a third party at this stage.

C. Data-centric security level

Security is supported and funded by the information holder at this stage.

These three security classifications should be merged and adhered to the database framework for a robust security result Nevertheless, it is important to reduce the dependency of one level of security on that of another. Nevertheless, the information-centered security level must only be enforced by the data holder from the obligation sense, as the software in the cloud is held by the data owner irrespective of which cloud framework provides customers with services.

IV. CHARACTERISTICS OF DCS

All of the DCS strategy's features are relevant to security and privacy concerns of cloud computing data protection for users. Data security has been one of cloud computing's major challenges. To enhance protection, privacy, honesty, and information quality should be observed.

A data holder is usually the best person to decide his or her own data's security standards. Primarily, information are categorized depending on the data holder Depending on this, the access control procedures and security resources required to manage the data are established in compliance with those data and security standards.

A. Data set

Data is created in various forms in a computing environment, which can be separated into physical and software layers. The data set may be a frame, a folder or a directory including files[14] in terms of space and propagation. The data set is the compilation of data in reference to security that can be secured and handled safely as irrespective of the data. Data can typically be divided into three sections of cloud computing: structured, unstructured and semi-structured data. Structured data are represented in a rigid structure. Unstructured records does not have a pre determined template of information and/or do not exist as a database array. Where as Semi-structured data is a type of structured data with a hierarchical structure of material information provided in fields and records and differentiated by symbols usually referred to as tags. Semi-structured data usually include XML, HTML or email. The data set can thus be utilize to reference to any of the data sources listed above. It is important to determine for which form of data set the solution will be used for, in a realistic development of the DCS approach.

B. Lifecycle of data

Every data set has several system phases in the cloud computing framework. The lifecycle relates to the overall phase of the processing of any set of data usually passing through the following phases:

- **Creation of Data**

This will be the first phase of the information lifecycle where in the computing framework the data can be interpreted in multiple forms of digital set of data. The data can be developed on the holder side of cloud computing before it is forward to the cloud or processed in the cloud. From this first phase, problems of security and privacy should be addressed when developing a DCS strategy.

- **Transmission of Data**

The data can be transferred from one location to another within or between the server or data holder domain at this phase. Data privacy and confidentiality should be maintained for all data transfer conditions for avoiding unauthorized clients from accessing or exploiting data. With the DCS method, even though the data owner implemented security methods to the information prior to sending the data in the cloud, the data is protected in any direct communication between the data holder and the service provider, or between virtual machines and different cloud computing platforms.

- **Storage/backup of Data**

The data is saved in the cloud framework in a hardware storage devices; typically users of cloud share the hardware resources and effectively separate their data. At this phase, maintaining data privacy and authenticity includes archiving the information without preventing the cloud service provider from performing the necessary recovery and archive techniques.

- **Process /Use of Data**

If computational activities interact with data contents and include the CPU, the data is said to be in use and stored. During this phase, the key challenge will be how to preserve information privacy with less effect on its quality of use.

- **Sharing of Data**

At this phase, by allowing other cloud clients to utilize this data, the data holder expands usage of data. Such cloud uses are licensed clients who can make use the data in compliance with the data holder's authentication control policy.

- **Destoy/Deletion of Data**

The data will be removed permanently from the database, either because the data holder does not want to save the data in the cloud framework, or the data holder will manage data with some other cloud service provider. When these information already include confidential data, the data holder must be sure that the confidential data cannot be released, even though the lost data are recovered.

C. Conventional Advantages of DCS

It is assumed that this technique will yield additional benefits. Confidentiality of information is maintained. The data is identity-protected and no licensed agency can broken the security in the database or anywhere else. Confidentiality of information is secured even though there are breaches of privacy in the cloud. It can be treated more secure to outsource encrypted information to the cloud then to keep this internally unencrypted. As a result, the data is secure even if all the database service is hacked. The risk management challenge is decreased[2]. Every data set's security criteria can be efficiently defined irrespective of other data sets or existing security efficiencies[2]. Auditability and transparency are given by software codes linked to the data. The control logs are added to the information and the logs are secured from changes.

The metadata for standard monitoring and tracking features added to the software can provide visibility of information usage and position in the cloud[1]. Data access control protection is maintained as information access is implemented under the protected protocols concealed within the information. Information privacy is preserved throughout their lifecycle. Authorized users must independently check the validity and reliability of each set of data. Data security and confidentiality is based on the stability of the encryption approaches utilized more than cloud service providers value and do not need a TTP. Cloud Providers and users are limited by the regulatory authority concerns. Because information is protected and is not usually publicly disclosed, it might not be susceptible to rules imposed on data on the internet in different parts of the world. The risk of data disclosure via virtual machine (VM) will not reveal confidential data to malicious parties since the data in the network is coded and it can be decrypted outside of the cloud framework.

V. CHALLENGES RELATED TO THE IMPLEMENTATION OF DCS APPROACH

The implementation of the DCS approach to the cloud computing system has several issues to address. This is expected as the DCS technique manages to provide a sustainable solution to improve information security and privacy in the cloud, and all related information security concerns must be addressed by the software specification. Some of these challenges can be encountered in the method's design stage and others issues can be only found in the execution stage. The encrypted information must always stay encrypted in the cloud environment in the DCS methodology. Information security and availability are compromised if the information continues to be in encrypted form in the cloud to meet this criteria. Techniques need to be configured to handle the associated user access policies and the keys that are use to encrypt data that will be accessible, particularly when the encrypted data is exchanged between trusted clients.

VI. IMPLEMENTATION OF DCS TECHNIQUE IN CLOUD COMPUTING FRAMEWORK

The key algorithm used in the suggested solution is Chinese Remainder Theorem (CRT).

A.Chinese Remainder Theorem and its Functions

The CRT was created to study mathematical issues in historical documents published by a Chinese mathematician[11]. It is recognized to be one of the earliest mathematics proofs used during the initial century to construct calendars. Mathematicians all over the world had improved the CRT until it achieved its present formula[3].

Chinese Remainder Theorem: For every given integer, a1, a2,..., ak, the subsequent simultaneous coherence scheme has a special X solution, in which all0 < n= n1n2... nk, such that even the non-negative integer numbers n1, n2,...., nk are fairly prime.

$$\begin{aligned} X &\equiv a_1 \bmod n_1 \\ X &\equiv a_2 \bmod n_2 \\ &\vdots \\ X &\equiv a_k \bmod n_k \end{aligned}$$

……(4.1)

If X is specified in the above-mentioned convergence, the formula shall measure the a_i:

$$a_i = X \bmod n_i$$

…….(4.2)

as i=1,2…..k .

The identical solution X can be determined by the corresponding formulas for the reciprocal convergence:

$$X = \sum_{i=1}^{k} a_i M_i M_i^{-1} \pmod{M}.$$

---------(4.3)

where

$M = n_1 n_2 \dots n_k$,

$M_i = M/n_i$,

M_i^{-1} is the multiplicative inverse of $M_i \bmod n_i$, i. e. $M_i M_i^{-1} \equiv 1 \pmod{n_i}$

Since M_i is comparatively prime to n_i, in mod a_i, there can be a distinct multiplicative inverse The multiplicative reverse equation in modular is therefore an integral part of deciding the CRT approach. The Extended Euclidean Algorithm (EEA) could be used effectively to measure the multiplicative inverse, and is an important part of the Garner's algorithm that is commonly need to find the answer for CRT.

B. Granting and Revoking Strategy

A new conceptual model is applied to the CRT to allow a new client u_{k+1} access to a directory r, which can be seen in Equation (4.4). The holder must recheck the shared value X'_r in this kind of scenario.

$$\begin{aligned} X'_r &\equiv \left(E_{K_{pub\,1}}(C_r \parallel K_s)\right) \bmod n_1 \\ X'_r &\equiv \left(E_{K_{pub\,2}}(C_r \parallel K_s)\right) \bmod n_2 \\ &\vdots \\ X'_r &\equiv \left(E_{K_{pub\,k}}(C_r \parallel K_s)\right) \bmod n_k \\ X'_r &\equiv \left(E_{K_{pub\,k+1}}(C_r \parallel K_s)\right) \bmod n_{k+1} \end{aligned}$$

……(4.4)

The X'_r answer from the above concurrent convergence can be determined from the X_r, which has already been added to the file, with much less dynamic calculation if the current X'_r convergence model has the similar modular of the earlier X_r module plus the advance modular.

VII. ANALYSIS OF RESULT

This segment addresses the workload, processing and calculation implications and deployment problems. The overhead calculation is measured as the calculation time required. The conclusions were based on the belief that all of the values of the parameters are earlier obtained. Each segment emphasizes on the data holder and trusted consumer side of the information.

A.At the Side of Data Owner

Next, the data archive is encrypted by an AES and the corresponding encrypted data file becomes the DCS file's final part. Table I displays overhead related to storage and processing times for encrypting various data file types using the 256-bit key size AES algorithm. The workload space for all data files is about 300 bytes. Through the rise in file size, the computing latency is raised to exceed as much as a minute for a 571 MB size data file.Therefore, choosing a robust encryption technique and a lengthy key to achieve better security is significant.

Table - I. Time and Space Overheads for the AES Encryption Technique

Name of File	File Size in bytes	File Size after performing encryption in bytes	Increase in file size in bytes	AES encryption in seconds
Any.docx	14,074	14,386	312	<1
m.jpg	65,812	66,130	318	<1
Wildlife.wmv	26,246,026	26,246,338	312	1.6
Case.avi	137,237,016	137,237,330	314	14.8

Relative to word files, we can note according to the above table that media files acquire more space and time overheads. AES encryption calculation of the X_r value, calculation of the HMAC for keywords, calculation of the index of the encoded file, and encryption of the index are the main actions that the information holder executes to build a DCS folder. Table II shows the complete processing times of various file types provided by these processes. Calculation of the X_r rate of five clients and addition of five coded keywords are performed.

Table II. Complete Time Overhead Enforced for Generating DCS file

Input file size in bytes	AES in sec	Calculation of Xr for 5 users in seconds	Calculationof file digest and Signature in sec	HMAC SHA 256 for 5 Keywords in sec	Total time in seconds
14,074	<1	0.029	0.003	0.199	~ 1
65,812	<1	0.029	0.005	0.199	~ 1
42,750,493	2.3	0.029	0.720	0.199	3.248
137,237,016	14.8	0.029	5.050	0.199	20.078

The average processing time for data of less than 50 MBs was somewhat 4 seconds and for a database of 130-MB volume was approximately 20 seconds. The average processing time for a 571 MB document was nearly about 1.7 minutes. From the details shown in Table II, it is noticeable that there is little effect on calculating the specifications needed for identity management, sharing of keys, and search capabilities.

B. At the Side of Trusted User

A trusted user can execute three functions; compute $C_r||K_s$ value from X_r value, check an encrypted data file's validity and reliability, and decrypt the encrypted document. The processing times for various sizes of DCS data files are shown in Table III. The highest processing time is required for AES decryption technique, preceded by confirmation of the validity and reliability of the encrypted file. There is a limited execution time when calculating the $C_r||K_s$ value. It is possible to recreate the actual encrypted file from only a DCS file on the server side and then return the actual encrypted data file rather than the DCS file as well as its signed list. Therefore, the client side prevents the cost of copying the encoded data from the DCS data file into the current reconstructed folder.

Table III. Processing Time of the Activities at the Client Side for various DCS files

Input DCS file name	File Size in bytes	Generating actual encrypted file in ms	Calculating $C_r||K_s$ from X_r in ms	AES decryption in sec
Any.docx.dcs	15,337	214	29	<1
m.jpg.dcs	67,082	350	29	<1
Wildlife.wmv.dcs	26,247,290	1448	29	3.5
Case.avi.dcs	137,238,282	11520	29	32

VIII. CONCLUSIONS AND FUTURE WORK

This implemented paper discusses the principle of Data Centric Security (DCS) as a significant strategy to improving the data privacy and security of files stored by users in cloud computing environments, primarily in the public cloud plate-form where information can often be transferred from one cloud server to another cloud sever and accessible by different client. A potential change can be introduced on the proposed approach to address further technical issues that the DCS strategy experiences hence it can be extended to different types of cloud service. Among the most difficult aspects would be how to maintain the key features of the initial DCS solution in this implemented paper once it is extended and applied to include several cloud computing systems and several specific cloud computing plate-forms.

REFERENCES

1. L. Chen and D. Hoang, "Active data-centric framework for data protection in cloud environment," in ACIS 2012: Proceedings of the 23rd Australasian Conference on Information Systems 2012, Geelong, Australia, 2012, pp. 1-11.
2. S. Ransom and C. Werner, "Towards Data-Centric Security in Ubiquitous Computing Environments," in Database and Expert Systems Application, 2009. DEXA '09. 20th International Workshop on, 2009, pp. 26-30.
3. O. Knill, "A multivariable Chinese remainder theorem," arXiv preprint arXiv:1206.5114, 2012.
4. S. Ransom and C. Werner, "Towards Data-Centric Security in Ubiquitous Computing Environments," in Database and Expert Systems Application, 2009. DEXA '09. 20th International Workshop on, 2009, pp. 26-30.

5. T. J. Lehman and S. Vajpayee, "We've Looked at Clouds from Both Sides Now," in SRII Global Conference (SRII), 2011 Annual, 2011, pp. 342-348.
6. Y. Chen, V. Paxson, and R. H. Katz, "What's new about cloud computing security?," University of California, Berkeley Report No. UCB/EECS-2010-5 January, vol. 20, pp. 2010-5, 2010.
7. B. Grobauer, T. Walloschek, and E. Stocker, "Understanding Cloud Computing Vulnerabilities," Security & Privacy, IEEE, vol. 9, pp. 50-57, 2011.
8. T. Dillon, W. Chen, and E. Chang, "Cloud Computing: Issues and Challenges," in Advanced Information Networking and Applications (AINA),
9. 2010 24th IEEE International Conference on, 2010, pp. 27-33.
10. [9] Y. Chen, V. Paxson, and R. H. Katz, "What's new about cloud computing security?," University of California, Berkeley Report No. UCB/EECS-2010-5 January, vol. 20, pp. 2010-5, 2010.
11. K. Popovic and Z. Hocenski, "Cloud computing security issues and challenges," in MIPRO, 2010 Proceedings of the 33rd International Convention, 2010, pp. 344-349.
12. S. Iftene, "General secret sharing based on the Chinese remainder theorem," ed: Citeseer, 2006.
13. S. A. Almulla and Y. Chan Yeob, "Cloud computing security management," in Engineering Systems Management and Its Applications (ICESMA), 2010 Second International Conference on, 2010, pp. 1-7.
14. F. Rocha, S. Abreu, and M. Correia, "The Final Frontier: Confidentiality and Privacy in the Cloud," Computer, vol. 44, pp. 44-50, 2011.
15. R. K. L. Ko, M. Kirchberg, and B. S. Lee, "From system centric to datacentric logging - Accountability, trust & security in cloud computing," in Defense Science Research Conference and Expo (DSR), 2011, 2011, pp. 1-4.
16. A.-R. Sadeghi, "Trusted Computing — Special Aspects and Challenges SOFSEM 2008: Theory and Practice of Computer Science." vol. 4910, V. Geffert, J. Karhumäki, A. Bertoni, B. Preneel, P. Návrat, and M. Bieliková, Eds., ed: Springer Berlin / Heidelberg, 2008, pp. 98-117.

AUTHORS PROFILE

Smita Sharma has completed her B.Tech and M.Tech in Computer Science & Engineering from Rajiv Gandhi Technical University , Bhopal (M.P.) . She is currently a PH.D scholar in the Department of Computer Science and engineering in SSSUTMS . Her interests of research are cloud security , artificial intelligence and image processing. She has published more than 10 research papers in reputed international journals and conferences related to these research areas. She has about six years of teaching experience. She is a member of ACM . She is currently working as an Assistant Professor in Computer Science & Engineering department in University of Information Technology, Rajiv Gandhi Technical University , Bhopal (M.P.).

Dr. R.P. Singh is former Director and Prof. Electronics and Communication at Maulana Azad National Institute of Technology, (MANIT) Bhopal. Dr. Singh Graduated and Post Graduated in Electronic Engineering from Institute of Technology (now IIT), B.H.U. Varanasi in 1971 and 1973, respectively. He did his Ph.D. from Barakatullah University Bhopal in 1991. He has 39 years of teaching, research, and administrative experience in Maulana Azad College of Technology (MACT)/MANIT out of which 22 years as Professor. He was Head of the Department at of Electronics, and Computer Science and Engineering Department at MANIT Bhopal .He has worked as Professor In-charge Academic and Chairman Admission Committee Dean (Academic) & Dean (R/D) at MACT /MANIT, Bhopal. He has published 125 papers in National / International reputed and indexed Journals including SCI. He was Chairman of Computer Society of India. Bhopal.