International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-5, January 2020

# Data Obfuscation Technique for Security in Cloud Computing

**Diwakar Ramanuj Tripathi**

***Abstract:*** ***Cloud computing is at present a day's changes into the most crushed in marvels to use for enormous scale alliance or for person who need unmistakable structure organizations with least expense. Usually person's data is dealt with on open Cloud which is accessible to everybody for get to. These focal raises some issue in reverse to adaptable organizations gave by cloud suppliers, similar to Confidentiality, Integrity, Availability, Authorization and some more. Starting late, Lots of decisions are open to guarantee the data and most perfect ways is to use encryption. Encryption can't give enough confirmation while contemplating heaps of customers' fragile data. It in like manner eats up greater chance to perform encryption and translating process for every single inquiry. Moreover it is definitely not an OK practice to contemplate customer driven considering the way that once customer data is moved on Cloud premises, customer doesn't have direct control over this data. To empty load of Cloud server likewise to give acceptable security to customer's data, we propose a methodology by joining the two procedures viz. Jumbling and Encryption in this assessment paper. Customer data may be scrambled if protection is needed for its records or reports and the SaaS Cloud organization is reviewed using obscurity techniques. Using this two-way approach, we may assume that the proposed agreement provides sufficient protections for darkened access and safety of even open data on cloud servers. Our point is additionally to give appropriate assessment on data disordering for security in cloud computing.***

***Key words:*** ***cloud computing, Confidentiality, privacy, data security.***

## I. INTRODUCTION

Cloud Computing has beginning late rose as new point of view for empowering and ignoring on organizations the Internet. Cloud Computing is the utilization of computing assets, for example, equipment and programming that are disregarded on as organization the web. The US National foundation of standard and improvement has given the full scale criticalness of cloud computing that is "Cloud computing is a model for connecting with certain, advantageous, on-request plan access to a typical pool of configurable computing assets (e.g., systems, servers, collecting, applications, and organizations) that can be quickly provisioned and discharged with unimportant association exertion or organization supplier facilitated exertion. This cloud model is made out of five crucial properties, three assistance models, and four sending models".

Manuscript published on January 30, 2020.
* Correspondence Author
**Dr. Diwakar Ramanuj Tripathi*,** Information Technology Consultant, Nagpur.



Cloud computing is starting late tolerating inconceivable game plan of thought among customers. Another strategy for describing the cloud computing is to take a gander at its five crucial properties as notice in table I, on demand self-administration, far reaching framework get to, resource pooling, quick flexibility and assessed administrations. In On-request self-organization, client gets Organizations gave by cloud agreeing to his essential with no human correspondence. Sweeping framework gets to makes cloud administrations open over the web so customer can pick up induction to any cloud administration using framework through any customer. Trademark resource pooling requires resources available acr oss the cloud and it goes everywhere specific customers go. No need to learn where to take care of the advantages.
Fast versatility indicates that Cloud management capabilities as indicated by customer request can be given easily and de ftly and are available to the customer at any level, assessed a dministration screen, control and reporting, providing conve nience to both the supplier and the user of the aid used.
Cloud computing gives different organizations; as shown in table I, these organizations sent three models, programming as an organization, orchestration as an organization, structure as an organization. In SaaS, software companies' customers can use application that starts running on a cloud system right now. Those applications from any area can be available. Saas Occasion is a CRM program from salesforce.com. In PaaS, The cloud supplier gives the stage as organization to the client where he can deal with its application and use it without regulating cloud structure. Model is Google Apps, IaaS kind of administration, cloud provider give establishment where the customer can manage its establishment near to application for its motivation. The best occurrence of Iaas is amazon web organizations.
Data jumbling methods incorporate three essential properties: reversibility, specification, and move. Figure 1 offers a significant level outline of data lack of definition systems and the different properties each help. We before long depict these properties, underlining reversibility, which is vital to data security. Data muddling methods incorporate three guideline properties: reversibility, detail, and move. Figure 1 offers a raised level delineation of data tangling strategies and the different properties each help. We before long depict these properties, concentrating on reversibility, which is basic to data security.

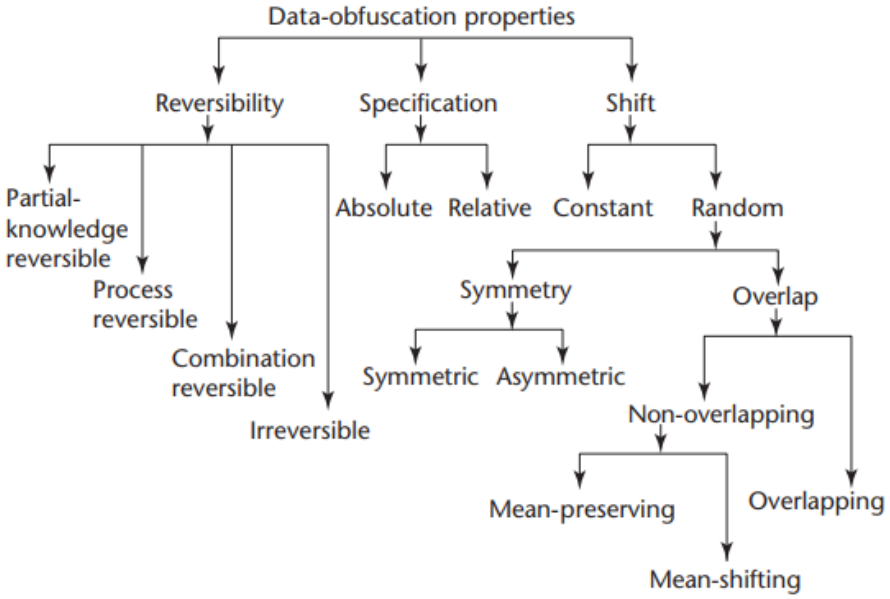


**Figure 1: Data obfuscation properties**

The three primary properties—reversibility, particular, and move—are each additionally characterized by different sub properties.

- ✓ **Challenges to Security and Privacy**
  - ➢ **Privacy Issues**
    - A. Misuse of Cloud computing
    - B. Malicious Insiders
  - ➢ **Security Issues**
    - A. Multi Tenancy
    - B. Access
    - C. Availability

### 1.1 Existing data security approaches

The database society has since quite a while ago perceived that aggressors can utilize particular equations known as trackers to bargain really a modest question set. For the most part, trackers permit assailants register database bits of knowledge without requiring some improvement commonality with the database substance, to the extent that the structure's requests use an optional specific procedure to pick get subsets.

Data randomization purposely baffles trackers from reproducing a database through dreary requests. Data mining examiners use randomization techniques to make a precise gathered data model without the data record's careful data. Data randomization for the most part works on a subset of data source tables, fields, and chronicles to keep the database's authentic attributes. The end customer can change data randomization to get the fundamental characteristics, subsequent to disturbing them with to a great extent a self-confident adaptable or maybe data discretization. While researchers have used data randomization basically to databases and data mining, the thought is particularly similar to subjective data obscurity techniques; data randomization strategies are stretched out effectively to make helpful data sets which is generally solely spread to end clients, whose data confusion underpins.

***Data anonymization*** attempts to arrange data into fixed or potentially adaptable between times. Each datum segment will be superseded by the class between time of its, alongside reasonable interim decisions can ensure that the factual data is kept up. Latanya Sweeney just as her associates made a security insurance method which guarantees that each data product will interface with at the very least k different sections, regardless of whether the documents are legitimately associated with outside data. This strategy requires a speculation just as concealment technique to obtaining the important secrecy level: speculation replaces a value with a significantly less particular worth, while concealment doesn't dispatch a value on the whole.

Data swapping keenly swaps sections inside one zone in records set along these lines the interesting report passages are unmatched; however the data are kept up all through the one of kind fields. Clients can lead swapping these sorts of the swapped qualities are close to each other, in this way approximating the data in the non obfuscated data records. This strategy thusly presents a kind of data jumbling. The 3 methodologies referenced as of now have a spot with the grouping of puzzle protecting data mining techniques. Data jumbling is a speculation of present strategies and they have utilizes past data mining. The "Data muddling models" sidebar (p. forty one) prospect records model projects of data muddling. Data jumbling offers a group of systems which contains the 3 strategies expressed above and a few others, and also offers a standard for classifying the distinctive muddling techniques relying upon the characteristics and measurements we today portray.

### 1.2 Need of Data Obfuscation

By and by days encryption is called as the way to making the data development guaranteed, it's the control of data using a figuring in such a manner, that when found by an unapproved solitary it won't look good yet at whatever point decoded sufficiently subject to the kind of cryptography estimation utilized, will twist up absolutely strong for any person. Regardless of the way that it's greatly helpful it before long can't resolve the security issue in the organization provider"s side showed by the cloud computing as the data required in the provider"s side ought to be in the decoded sort of its thusly it could be handled.

Cloud computing chips away at dispersed computing in which the authority center ideas with the aggregate of the materials for instance hardware, amassing and programming in the kind of Software as a Service (SaaS), Platform as a Service(PaaS) or maybe Infrastructure as a Service(IaaS) that the customer can pick by spending for it notwithstanding remotely get to these materials working with any device including PC telephone, tablet or mobile phone which will require not have a top end assurance. It's different central focuses, for instance, the organizations are incredibly modest, may be picked by the solicitations of the individual and may be stopped accurately proportional to per criticalness moreover, in this way attracts a ton of customers and has now become widely during that time yet obviously inferable from the inadequacies of its the cloud continues being not without a doubt the principal choice of a couple. At whatever point an individual is using cloud computing game plans he's critical to move the data being set up on the program provider's side, process the data immediately the paper is returned. This particular expert association fundamentally offers a virtual room on the customer together with others customers moreover that steamed the clients the best. Encryption is regularly used to keep up a key good ways from data from falling into an unseemly hands while being moved anyway when it lands at the objective of its must be changed into its fundamental structure basically being kept in similar additional room that is emphatically given by an outcast and besides may be used by contention as well. This' where confusion of data can assume a critical job. Muddling is basically veiling of data in these sorts of way that the data gets unusable for an aggressor or perhaps an unapproved faculty anyway does free the attributes of its which might be used to process the data in this structure without affecting the impacts if the data is deobfuscated into the exemplary type of it. Obscurity is a sub scope of encryption and could be called semi encryption. Since obscurity empowers the data to keep the characteristics of its, it could be incredibly helpful device in cloud computing security. Attributable to this specific point different jumbling techniques have been investigated and furthermore the absolute best strategy in connection to cloud computing security was chosen and significantly furthers a reason for more examination was proposed in "Muddling as a degree for Cloud Computing.

The issue over mystery of delicate and individual information has brought about the usage of various strategies for stowing away, scrambling and muddling touchy data of databases. The need for mystery has brought about the improvement of various (data muddling) DO procedures that offer security upkeep in the cost of information misfortune. Most of the systems take into account certain areas and function admirably for a little arrangement of uses. In the absence of a standard for ordering DO techniques, examination and execution assessment of the different systems isn't straight advance. The space of consideration in this specific examination is data mining. A ton of data mining applications include learning by means of group examination.

### 1.3 Obfuscation Techniques

Jumbling ordinarily alludes to the various strategies used to shroud delicate data. "Scrubbing" may be used conversely with confusion. The strategies referenced here have 2 essential objectives: Secure sensitive data from disclosure and allow available test data which have the form of comparative data as secret.

- **Masking**

Software covering covers vulnerable characters or perhaps fields with a meaningless character such as "X." Masking guarantees that the software image is displayed alongside images on the papers. When printing a receipt, every day instances of concealment include using Xs on all but the last four digits of a charge card number.

- **Substitution**

Substitution recharges data zone with taking after substance which is not connected to the underlying data. A true instance of substitution is to displace the actual first and last names with names heedlessly chosen from a substantial once-over of legitimate first and last names developed specifically for substitution use. Substitution ensures the fundamental data shape while hiding the particular fragile data**.**

**Substitution and rearranging records**

Rearranging are in a general sense equivalent to beside that revising uses the explanation data itself rather than an outside overview. Revising moves data between lines subsequently the data shape is spared anyway the hidden nuances of the tricky data are concealed.

- **Number and date fluctuation**

Variance modifies number or likely dates data by modifying the zone with similar data which is a self-assertive portion of the secret data. The percent shift is selected to hold the crisp inside generous stretches out of the plastic new data because of the field close to the use of the field. Distinction keeps the data shape while disguising the basic tricky data.

- **Gibberish age**

Garbage advancement is required once the touchy data you need to stow away has related data like correspondence which can perceive the underlying data. An ordinary outline is bank records. You can jumble the record information of individuals in the database tables in spite of the fact that documents are associated with pictures or perhaps different copied (.pdf records) of the month to month clarifications coordinated to those buyers. Those set away announcements contain all the information you need to stow away. To prevent this touchy data from being revealed, nonsense age replaces the characterized data with unpredictable "trash" data reports of equivalent size.

- **Encryption**

Encryption will keep the underlying data set up and open to anyone with the decoding key component. This' not all that engaging since the data will probably be made data pointless for improvement and testing purposes.

- **Data Generation**

Data age makes fake or nonexistent data without any planning or perhaps unique other imperativeness sources which is helpful for testing purposes.

You can locate a few more jumbling techniques not recorded here just as the strategy for every one of these recorded could fluctuate from truly simple to scientifically.

## II. LITERATURE REVIEW

**Khaled M Khan (2019)** this paper proposes a data jumbling approach in re-appropriating structure expansion to cloud computing. It is basically established on separating the lines and sections of systems to adjust their certifiable estimation joined with including sporadic uproar and improving to ensure arrangement and assurance. In our philosophy, confused systems are sent to servers with no open key encryption. While it figures on frameworks, the server can't expel or get certified characteristics either from cluttered systems or from enrolled duplication results, however customers can remove genuine handled characteristics using an amazingly irrelevant computing effort from results made by the server.

**Muhammad Hataba and Ahmed El-Mahdy(2018)[3]** This paper displays an investigation of programming affirmation subject to the possibility of security by absence of clearness, code jumbling is at present a fervently discussed issue in the field of cutting edge right organization, guaranteeing against making sense of and modifying. Obscurity ends up being helpful in conditions where depending upon cryptographic procedures isn't adequate, this is typical in remote execution conditions where the item is executed on a surprising revealed undermining condition, for instance, the new computing stages: cloud-computing perspective and mobile phones. Confusion is outstanding among malware and disease planners yet also game designers and industrials who need to guarantee their authorized development. They use it to mask the movement of their code while executing in an uncontrolled area. In this paper, we talk about comparative thoughts yet for the differing inspiration driving cloud security. We examine the front line in strategies and figuring's for programming jumbling. We furthermore address how to assess the nature of these strategies by methods for a strong course of action of estimations.

**Jayeshkumar Madhubhai Patel and Krunal Suthar(2018)[4]** Cloud figuring is correct presently every day's gotten generally destroyed in wonders to use for a massive scale connection or for person who need assorted structure organizations with least expense. Person data is typically handled on an open cloud that is available to everyone to get there. This key raises some concerns that Cloud providers have provided in reverse to adaptable organizations, similar to confidentiality, fairness, availability, authorization and some more. Piles of alternatives now accessible for a day and the absolute best course are to use encryption to ensure the data. Encryption can not provide adequate security when worrying about the delicate data of the client, as it also expends more remarkable opportunities to process encryption and unscrambling. In this paper, we propose a system for combining methods, i.e., to clear the magnitude of the Cloud server, also to keep adequate protection to the data of the client in Cloud state. Lack of definition and encryption. Customer data may be encoded out of chance requiring protection for documenting or monitoring, and Cloud's DaaS organization needs to be checked using perplexity frameworks. Using this two-way approach, we may conclude that the proposed agreement provides ample protection to ensure that even the data accessible on cloud servers get to dark and guarantee security. We would also like to provide a valid reliability testing system, a better access management mechanism that reduces the size of the Client as a company supplier as well.

**Dr L. Lawrence Arockiam and S. Monikandan(2017)[5]** In the open cloud environment, data security in the cloud is the most important check. Clients are moved to cloud to collect correctly, safely and scalable. Cloud service providers (CSP) and other cloud companies are exposing data due to security concerns. This paper proposes an assurance protocol as a Security Service Algorithm (SSA), called MONcrypt to protect the data from unapproved implementation in cloud collection. This proposed security procedure depends on the strategy of tangling the data. Safety as a Service (SEaaS) leverages the MONcrypt SSA. Clients may take advantage of this SEaaS security organization to test their data at any level. Redirections for evaluating the protection of proposed MONcrypt SSA were powered in cloud condition (Amazon EC2). A safety evaluation instrument is used to measure the safety of the planned and current indeterminate consistency techniques. MONcrypt distinguishes and confuses current methodologies such as Base32, Base64, and Hexadecimal Encoding. The suggested solution provides better implementation and unfathomable protection when there are different and current techniques of disordering. MONcrypt diminishes the size of the data being transferred to cloud storage instead of the present system.

## III. PROPOSED OBFUSCATION TECHNIQUE:

ARO OBFUS CT The proposed confounding approach for testing the numerical data in the cloud array is used. Right when the client needs to clutter the fragile numerical data, this proposed technique is fair and mandatory by then. This method is a symmetrical form in cryptography. There are two keys used to encrypt and disentangle this proposed say. Alternatively, both keys are properties of whole numbers. With these two keys, the jumbling of numerical data is conceivable by the proposed ARO Obfus CT for ensuring the data out in the open cloud. The proposed ARO Obfus Cryptographic Techniques (CT) uses five common, coherent activities on numerical data, for example, mul), (pow), (pivot), (mod), (ascii). In cloud side, the two enigma keys are generated and sent to the clients. Such keys are retained as a service in the company supplier called Key Management. (KMaaaz). The entire work and outcome are separate and the current methods are separate. For confusing method the size of the given plain substance is settled. The simple material is duplicated with the K1 model calculation and preserved in the vitrine. For the extended value the square value is decided. The model made with K2 is expanded by one and placed into the qualities of the square. Every time, these attributes are relinquished to clear for K2. Binding 256 finds the mod-value in the going with arrange. For every mod worth the ascii character is built.

Such ASCII respects are the same content of the plaint content in the figure. The pseudo code is provided underneath for the proposed ARO Obfus CT.

**Pseudocode for ARO_Obfus CT for Numerical Data:**

ARO_Obfus(PT)

1. start
2. PT← plaintext
3. N← size of(PT)
4. Get a key $K_1$ from cloud for ARO_Obfus CT //Multiple the $K_1$ into PT(i)
5. MT(i)<-- PT(i)*$K_1$i=0,1,2… <N //find square SQ value for MT(i)
6. SQ(i) ← pow(MT(i),2) i=0,1,2… <N //Rotate the SQ at K number of times
7. Get a key $K_2$from cloud for ARO_Obfus CT //Rotate the RTN at $K_2$ number of times
8. RTN(i) ← rotate(SQ(i),K2+j) j=1,2… <=N //Find the module MOD for RTN by 256
9. MOD(i) ← RTN(i)%256 //Convert the MOD into ASCII code to produce Ciphertext CT
10. CT(i) ← ascii(MOD(i))
11. CT← cipher Text
12. End

## 4. RESULT AND DISCUSSION

The analysis method for the proposed strategy of obfuscation is done below the example data and the programmed test keys created.

**Step 1:** Consider the following plaintext which is the age of employees

**PT<-- 35 56 47 56 51 48**

**Step 2:** Find the total size of values in the PT and put as N.

**N<--6**

**Step 3:** The generated K1 value is multiplied by plain text (PT) and put as MT. The K1 value is 12 here. Check K1= 12, multiply the K1 in PT.

**Step 4:** Calculation of the square value for MT values: Find the Square SQ(i) for MT(i)

**Step 5:** The main K2 is generated, and the sample K2 is 4 here. For number of K2 times the square value is rotated from right to left. And K2 is one route incremented. Rotate the SQ(i) by K2 time numbers from right to left (back to front) Test K2= 4, for consecutive values in SQ(i), K2+i, i=1,2,3,... N.

**Step 6:** Rotated RTN(i) is ,

**Step 7:** Find RTN(i) Modulus by 256. The values for the Mod are calculated by dividing the values rotated by 256. And for each mod-value the ascii character is made. Those asci characters are the original numeric plaintext ciphertext. MOD(i) = 256 per cent RTN(i)

**Step: 3**

| PT(i) | MT(i)=PT(i)*K1 |
|---|---|
| 35 | 420 |
| 56 | 672 |
| 47 | 564 |
| 56 | 672 |
| 51 | 612 |
| 48 | 576 |

**Step: 4**

| MT(i) | SQ(i)= Pow(MT(i),2) |
|---|---|
| 420 | 176400 |
| 672 | 451584 |
| 564 | 318096 |
| 672 | 451584 |
| 612 | 374544 |
| 576 | 331776 |

**Step: 5**

| SQ(i) | K2=4 |
|---|---|
| 176400 | K2=4 |
| 451584 | K2=5 |
| 318096 | K2=6 |
| 451584 | K2=7 |
| 374544 | K2=8 |
| 331776 | K2=9 |

**Step: 6**

| SQ(i) | RTN(i) |
|---|---|
| 176400 | 640017 |
| 451584 | 515844 |
| 318096 | 318096 |
| 451584 | 445158 |
| 374544 | 443745 |
| 331776 | 776331 |

**Step: 7**

| RTN(i) | MOD(i) |
|---|---|
| 640017 | 17 |
| 515844 | 4 |
| 318096 | 144 |
| 445158 | 230 |
| 443745 | 97 |
| 776331 | 139 |

**Step 8:** Convert MOD(i) into ASCII Code to produce the ciphertext CT

**CT =1$1ga,**

The proposed method of obfuscation performs properly and generates the ciphertext with the blend of a wide range of character codes from ASCII. The following findings are derived from the above tests and the inputs of test data.

**Plaintext to Ciphertext:**

The Plain text is: 35 56 47 56 51 48

The CipherText : 1$1ga,

In the plain text, the numerical data ' 56 ' appears multiple times and the situation of those data is 2 and 4. The ciphertext character in this plaintext's proportionate situation is ' $g. ' This results in similar data having different ascii characters in the plaintext.

**Ciphertext to Plaint text:**

The CipherTextis: 1$1ga,

The Plain text is: 36 56 47 56 51 48

In the situation of 1 and 3 the ascii character in the plaintext ' 1 ' appears multiple times. The plaintext is 36 and 47 which are close to those positions. A similar character in the cipher text is known to lack similar data or opportunity in the plaintext. And maybe this is different.

**The Data size reduced:**

The data size for the same plaintext listed above is 17 bytes. (The plaintext would be: 35 56 47 56 51 48). Whatever it may be, the cipher text data size (the cipher text is: 1$1ga) is 6 bytes for the same plaintext. It decreases by 33 per cent (1/3).

## IV. CONCLUSION

In this article, another ambiguity method, ARO Obfus CT, is introduced and finished to affirm data confidentiality out in the open cloud gathering. This proposed method passed on the least size of the data when dealing with the supplier's jumbled data.

This paper provides data on different threats in cloud computing situation related to security and assurance in cloud condition of the touchy data of the client. Apparently the request is withheld from both the disclosures and therefore the protection is upgraded. Experts have suggested different methods for addressing problems using different methodologies that restrict the problem of data security and cloud safety to some degree. We explored the main focuses and constraints of existing systems to fully explain the security and protection issues. These are the open questions to track.

## REFERENCES


1. Mell, P., Grance, T.: The NIST Definition of Cloud Computing. National Institute of Standards and Technology, Information Technology Laboratory (2011), http://csrc.nist.gov/groups/SNS/cloud-computing/
2. Khan, Khaled. (2019). Data Obfuscation for Privacy and Confidentiality in Cloud Computing. 10.1109/QRS-C.2015.41.
3. Hataba, Muhammad & El-Mahdy, Ahmed. (2018). Cloud Protection by Obfuscation: Techniques and Metrics. Proceedings - 2012 7th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, 3PGCIC 2012. 369-372. 10.1109/3PGCIC.2012.18.
4. Suthar, Krunal & Patel, Jayeshkumar. (2018). ObfuCloud: An Enhanced Framework for Securing DaaS Services Using Data Obfuscation Mechanism in Cloud Environment. 333-343. 10.1007/978-981-10-5523-2_31.
5. Monikandan, S. & Lawrence, Dr. L. Arockiam. (2017). Confidentiality Technique to Enhance Security of Data in Public Cloud Storage using Data Obfuscation. Indian Journal of Science and Technology. 8. 10.17485/ijst/2015/v8i24/80032.
6. Subashini, S., Kavitha, V.: A survey on security issues in service delivery models of cloud computing. Journal of Network and Computer Applications 34(1), 1–11 (2011)
7. Top Threats to the Cloud Computing V1.0, Cloud Security Alliance, http://www.cloudsecurityalliance.org/topthreats/2010
8. Babu, J., Kishore, K., Kumar, K.E.: Migration from Single to Multi-Cloud Computing. International Journal of Engg. Research and Tech. 2(4) (April 2013)
9. Chandran, S., Angepat, M.: Cloud Computing: Analyzing the Risk involved in Cloud Computing Environment (2011)
10. Munir, K., Palaniappan, S.: Security threats/attacks present in cloud environment. IJCSNS 12(12) (2012).
11. Munir, K., Palaniappan, S.: Secure Cloud Architecture. ACIJ 4(1) (2013).


## AUTHORS PROFILE

**Dr. Diwakar Ramanuj Tripathi,** Received the graduation (B.Sc.) degree in Computer Science, Master degree (MCA) in computer Application and Doctor of Philosophy (**Ph.D.**) in Computer Science. He is a Microsoft Certified I.T. professional (**MCITP**), Microsoft Certified Technology Specialist (**MCTS**) and Microsoft Certified Trainer (**MCT**) with 10 + years' of experience in a field of Computer Science. He has presented research paper in various international and national level conferences and also published research paper in various reputed journal. He is currently working as an Information Technology Consultant in Nagpur (Maharashtra).