

# An Efficient Fuzzy C-Means Method with Variable FV-TC for Data Sensitivity Calculation in a Cloud Computing Environment

Ashutosh Kumar Dubey

**Abstract:** In this paper an efficient fuzzy c-means (FCM) method has been used for the data sensitivity estimation. For the saturation point estimation variable fuzziness value (FV)-termination criteria (TC) have been used in the cloud computing environment. Data preprocessing has been performed along with the five attributes. Three attributes are based on the cloud user input parameters and the remaining two are the automated attributes which are calculated automatically. Then FCM has been applied. The total clusters calculated by our method are three. Then weighted product model has been applied for the cluster sensitivity calculation based on the three clusters. Our mechanism has the capability to identify the need of high, medium and low-level securities.

**Keywords:** FCM, FV, TC, Data sensitivity, Cloud computing.

## I. INTRODUCTION

The current trends emphasize the use of cloud computing in different areas and application environment. The numerous use and demand make the concern which indicates the chances of attack. The picture is also treacherous in case of multi-cloud storage architecture [1]. Some of the authors are also working in the risk categorization and efficient secure framework development [2–9].

In 2019, Halabi and Bellaiche [10] discussed the cloud computing technology advantages. Computing technology. They have suggested regarding the lack of security assurance. For this they have proposed a security risk evaluation based on quantitative mechanism. They have proposed an InterCloud setting. Their simulation results indicate the reduction in security risk. In 2019, Garg and Garg [11] discussed and provided the overview of the cloud computing technology. It has been discussed in terms of cloud computing automation. They have examined and analyses deployment and maintenance strategy to make it effective with less admin effort. In 2019, Eddermoug et al. [12] discussed about the possible security threats in terms of cloud computing. They have proposed a new model. They called it profiling and preventing security attacks. They have integrated machine learning algorithm for achieving the security. Their approach has been explored with the realistic case study. In 2019, Kunz and Mann [13] developed a program for the automatic detection of risk patterns. It has been developed for the cloud system models. They have used

Eclipse modeling framework. They have explored their approach based on the experimentation and case study for investigating the applicability and scalability. In 2019, Ghaleb et al. [14] discussed about the security monitoring system endpoints. It has been discussed in terms of cloud computing. Their suggested framework is capable of retrieving the raw data transparently. In 2019, Colombo et al. [15] discussed and introduced the framework for Data Protection as a Service (DPaaS). It has been introduced for the users belongs to the cloud computing. They have compared it by the existing Data Encryption as a Service (DEaaS). Their approach efficiency has been proved by the experiments. In 2019, Compastíe et al. [16] discussed about cloud infrastructures and the security attacks. They have focused on the critical issues of the cloud computing. They have proposed a software-defined security approach. It is based on the TOSCA language. They have shown the results based on the experiments. In 2019, Soni et al. [17] discussed about the cloud ecosystem. They have discussed and explored the cloud service providers, cloud vendors and cloud tenants with a framework. It is in terms of cloud service ecosystem. In 2019, Ramamoorthy and Baranidharan [18] discussed about the data security and storage management. They have adopted the block chain technology for cloud data protection. They have generated the secure hash blocks. They have suggested that the block chain has the capability to trace the malicious identification. In 2019, Ranjan and Agnihotri [19] discussed the attack impact in case of cloud computing. Based on this they have suggested some solution. In 2019, Sankaran et al. [20] discussed about the cloud data security. They have suggested the hybrid cryptography approach. Their proposed system includes the concept of data migration along with the distributed data hosting scheme. They have considered the cost of migration and the data access patterns. In 2019, Sharma and Mathur [21] suggested a proposition. It has been suggested based on the integration of the confidentiality along with the web-based service availability. They provided two factors-based authentications. In 2019, Nguyen and Khorev [22] discussed about the information security aspects in the cloud computing. Their model has the capability of qualitative assessment based on the information risk. The main objective of this paper is to find the data attack sensitivity based on fuzzy c-means [FCM] method with variable fuzziness value (FV)-termination criteria (TC). So that the categorize security can be applied to the data.

**Revised Manuscript Received on October 15, 2019**

Ashutosh Kumar Dubey, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India. Email: ashutosh.dubey@chitkara.edu.in

# An Efficient Fuzzy C-Means Method with Variable FV-TC for Data Sensitivity Calculation in a Cloud Computing Environment

## II. PROCEDURE FOR PAPER SUBMISSION

In this paper an efficient approach has been presented based on FCM with variable FV-TC to calculate the data sensitivity. Data preprocessing has been performed based on the text data which has been uploaded to the cloud platform. The weights have been assigned based on the five factors. The first two factors determine the data keyword occurrence weight. These keywords are the key source value from the user. Third is the sensitivity parameter which ranges between 1-5. Higher is more sensitive means need more security. It is also the input parameter of the user. The fourth parameter shows the data size scaling. The fifth parameter is the frequency aggregate of the top two keywords. Then the FCM algorithm has been applied. For the saturation point finalization variable FV-TC has been applied. The values considered here are FV (2-6) and TC (2.0E-5-6.0E-5). It is based on [23]. The flowchart for FCM with FC-TV is shown in Figure 1. Total three clusters have been considered.

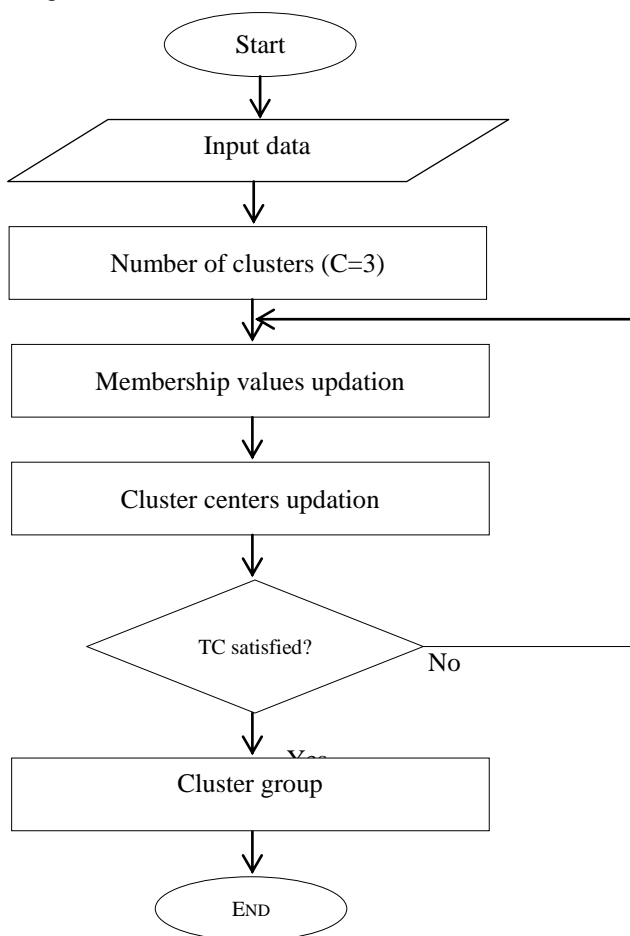


Figure 1 Flowchart for FCM with FC-TV

### Algorithm: FCM with FC-TV algorithm

- Step 1: The data weights have been considered based on the 5 attributes. Here it is denoted by D. M represents the data dimension.
- Step 2: Total number of clusters is 3. Here, it is represented by C.
- Step 3: Variable FV and TC have been considered for the saturation point calculation.
- Step 4: So, the membership matrix is calculated as follows:

$$\sum_{j=1}^C U_{ijm} = 1.0$$

Step 5: Then the degree is calculated based on the membership. It is called the degree of membership.

$$DOM_{jm} = \frac{\sum_{i=1}^D U_{ijm}^f x_{id}}{\sum_{i=1}^D M_{ijd}^f}$$

Step 6: Then the cluster distance between the center and the data point has been calculated.

$$CD_{ijm} = x_{im} - DOM_{jm}$$

Step 7: DOM has been updated iteratively.

$$U_{ijm} = \frac{1}{\sum_{c=1}^C \left( \frac{CD_{ijm}}{CD_{icm}} \right)^{\frac{2}{FV-1}}}$$

Step 8: The above process ends when the TC condition has been satisfied.

## III. RESULTS AND DISCUSSION

Total 100 text data files have been considered from web repositories. Randomly five sets have been considered. Each set consist of 20 data files. Based on the above method the results have been calculated. F1 to F5 score shows the variable combination of FV and TC. It is denoted by F1S to F5S. Means the F-score. Dt denotes the data file. Figure 2 shows the data sensitivity for the first set 1 with FV (2-6) and TC (2.0E-5-6.0E-5). Figure 3 shows the data sensitivity for the set 2 with FV (2-5) and TC (2.0E-5-5.0E-5). Figure 4 shows the data sensitivity for the set 3 with FV (2-5) and TC (2.0E-5-5.0E-5). Figure 5 shows the data sensitivity for the set 4 with FV (2-5) and TC (2.0E-5-5.0E-5). Figure 6 shows the data sensitivity for the set 5 with FV (2-5) and TC (2.0E-5-5.0E-5). The total clusters calculated by our method are three. Then weighted product model has been applied for the cluster sensitivity calculation. Based on this product model three values have been obtained, one for each cluster. It has been normalized for the comparison. It shows three levels. The highest value shows the high security is needed. The next highest shows the middle security requirement and the lower values shows the low security requirement. Figure 7 shows the cluster sensitivity for the random case. Figure 8 shows the average error in the F-score. The error percentage is also low which shows the significance of our method.

The results clearly indicate that our method has the capability in finding the risk requirement in terms of data sensitivity in the cloud computing environment. It can be helpful for the future framework creation with secure data and communication environment.

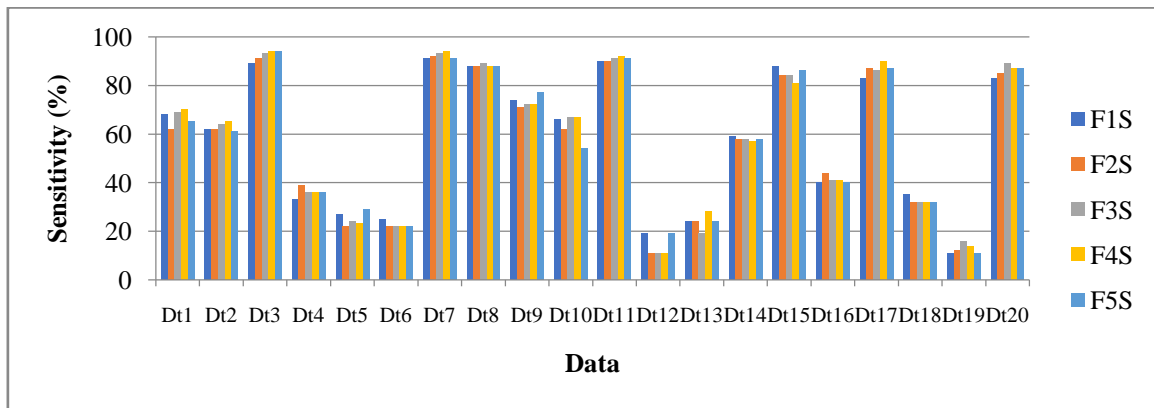


Figure 2 Data sensitivity for the set 1 with FV (2-5) and TC (2.0E-5-5.0E-5)

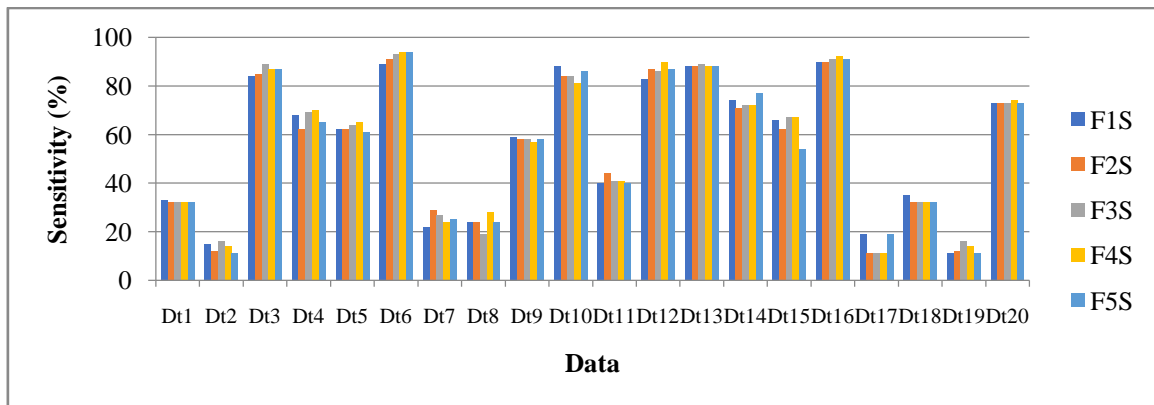


Figure 3 Data sensitivity for the set 2 with FV (2-6) and TC (2.0E-5-6.0E-5)

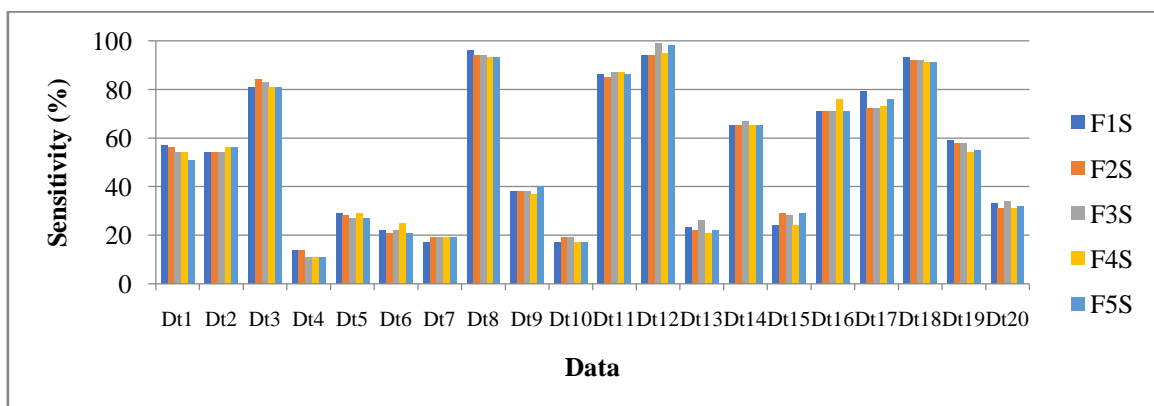


Figure 4 Data sensitivity for the set 3 with FV (2-6) and TC (2.0E-6-5.0E-5)

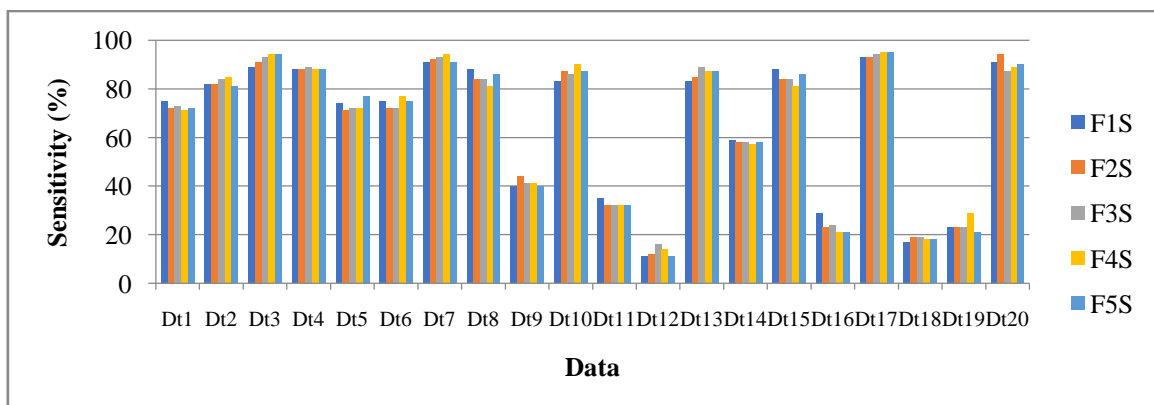


Figure 5 Data sensitivity for the set 4 with FV (2-6) and TC (2.0E-5-6.0E-5)

# An Efficient Fuzzy C-Means Method with Variable FV-TC for Data Sensitivity Calculation in a Cloud Computing Environment

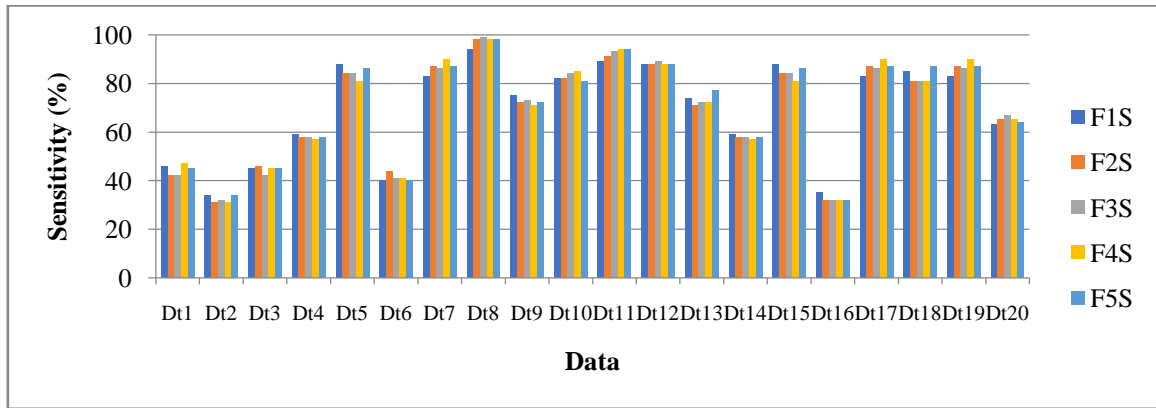


Figure 6 Data sensitivity for the set 5 with FV (2-6) and TC (2.0E-5-6.0E-5)

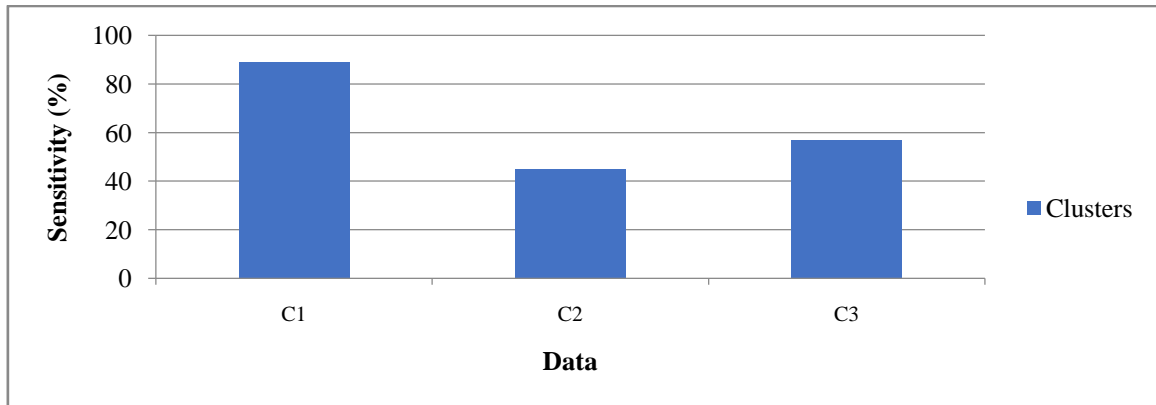


Figure 7 Cluster sensitivity of the randomly selected set

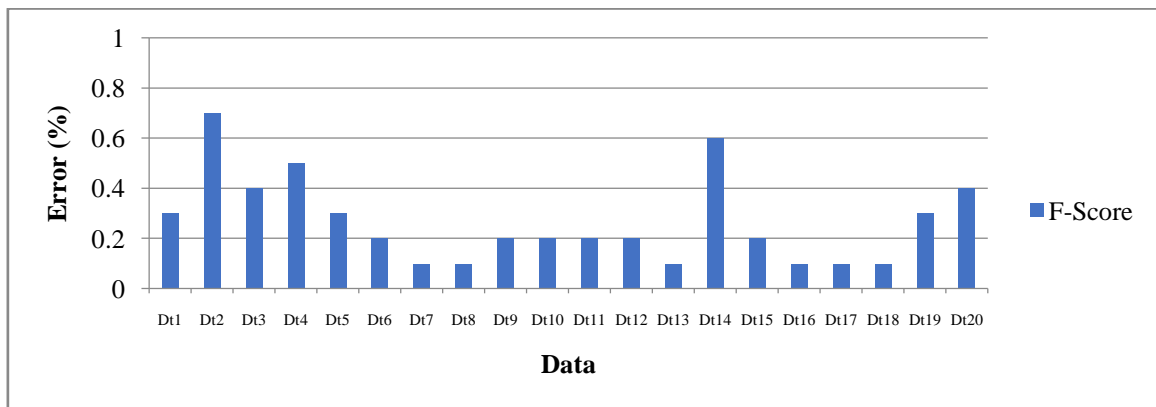


Figure 8 Average error in the F-score

## IV. CONCLUSIONS

In this paper an efficient FCM algorithm has been applied and analyzed for the data sensitivity analysis in the cloud computing environment. FCM with the variable FV-TC has the capability of clustering data efficiently. Then weighted product model has been applied for the calculation of security applicability in case of each cluster. The results obtained by our method show the performance in terms of F-score with proper data sensitivity categorization.

Some of the future suggestions are as follows:

1. For applying cryptography based on the data sensitivity achieved, strength of different cryptography algorithms can be analyzed in the future.
2. The above process may be automated based on the strength achieved.
3. Number of clusters can be increased for enhancing the

security levels.

4. In future similarity matching method may be applied along with the different clustering algorithms for better clustering.

## REFERENCES

1. Sukmana MI, Torkura KA, Graupner H, Cheng F, Meinel C. Unified cloud access control model for cloud storage broker. In international conference on information networking 2019 (pp. 60-65). IEEE.
2. Wulf F, Strahringer S, Westner M. Information security risks, benefits, and mitigation measures in cloud sourcing. In conference on business informatics 2019 (Vol. 1, pp. 258-267). IEEE.
3. Weil T. Risk Assessment methods for cloud computing platforms. In computer software and applications conference 2019 (Vol. 1, pp. 545-547). IEEE.
4. Dubey AK, Dubey AK, Namdev M, Shrivastava SS. Cloud-user security based on RSA and MD5 algorithm for resource attestation and sharing in java environment. In CSI international conference on



- software engineering 2012 (pp. 1-8). IEEE.
5. Khandelwal A, Gupta S, Mishra RS, Khandagre Y, Dubey AK. Establishing secure event detection with key pair in heterogeneous wireless sensor network. In advanced materials research 2012 (Vol. 433, pp. 3445-3450). Trans Tech Publications.
  6. Manimaran A, Durairaj M. The conjectural framework for detecting DDoS attack using enhanced entropy-based threshold technique (EEB-TT) in cloud environment. International Journal of Advanced Computer Research. 2016;6(27):230.
  7. Annane B, Ghazali O, Alti A. A new secure proxy-based distributed virtual machines management in mobile cloud computing. International Journal of Advanced Computer Research. 2019;9(43):222-31.
  8. Dalin G, Radhamani V. IRIAL-an improved approach for VM migrations in cloud computing. International Journal of Advanced Technology and Engineering Exploration. 2018;5(44):165-71.
  9. Tomar A, Dubey AK, Richhariya V. Novel sensitive information preserving mining (sipm) algorithm for association rule mining in centralized database. In international conference on emerging trends in networks and computer communications 2011 (pp. 392-397). IEEE.
  10. Halabi T, Bellaiche M. Security risk-aware resource provisioning scheme for cloud computing infrastructures. In conference on communications and network security 2019 (pp. 1-9). IEEE.
  11. Garg S, Garg S. Automated cloud infrastructure, continuous integration and continuous delivery using docker with robust container security. In conference on multimedia information processing and retrieval 2019 (pp. 467-470). IEEE.
  12. Eddermoug N, Sadik M, Sabir E, Mansour A, Azmi M. PPSA: Profiling and preventing security attacks in cloud computing. In international wireless communications & mobile computing conference 2019 (pp. 415-421). IEEE.
  13. Kunz F, Mann ZA. Finding risk patterns in cloud system models. In international conference on cloud computing 2019 (pp. 251-255). IEEE.
  14. Ghaleb A, Traore I, Ganame K. A framework architecture for agentless cloud endpoint security monitoring. In conference on communications and network security 2019 (pp. 1-9). IEEE.
  15. Colombo M, Asal R, Hieu QH, El-Moussa FA, Sajjad A, Dimitrakos T. Data protection as a service in the multi-cloud environment. In international conference on cloud computing 2019 (pp. 81-85). IEEE.
  16. Compastie M, Badonnel R, Festor O, He R. A TOSCA-oriented software-defined security approach for Unikernel-based protected clouds. In conference on network softwarization 2019 (pp. 151-159). IEEE.
  17. Soni D, Sharma V, Srivastava D. Optimization of security issues in adoption of cloud ecosystem. In international conference on internet of things: smart innovation and usages 2019 (pp. 1-5). IEEE.
  18. Ramamoorthy S, Baranidharan B. CloudBC-a secure cloud data access management system. In international conference on computing and communications technologies 2019 (pp. 217-220). IEEE.
  19. Ranjan I, Agnihotri RB. Ambiguity in cloud security with malware-injection attack. In international conference on electronics, communication and aerospace technology 2019 (pp. 306-310). IEEE.
  20. Sankaran KS, Vasudevan N, Prakash VR, Diderot PK. Access control based efficient hybrid security mechanisms for cloud storage. In international conference on communication and signal processing 2019 (pp. 0564-0567). IEEE.
  21. Sharma A, Mathur S. Concealing the user identity in cloud services. In international conference on electronics, communication and aerospace technology 2019 (pp. 372-376). IEEE.
  22. Nguyen MT, Khorev PB. Information risks in the cloud environment and cloud-based secure information system model. In international youth conference on radio electronics, electrical and power engineering 2019 (pp. 1-6). IEEE.
  23. Dubey AK, Gupta U, Jain S. Comparative study of k-means and fuzzy c-means algorithms on the breast cancer data. International Journal on Advanced Science, Engineering and Information Technology. 2018;8(1):18-29.

Professional Member of ACM. He has more than 11 years of teaching experience. He has authored a book name Database Management Concepts. He has been associated with many international and national conferences as the Technical Program Committee member. He is also associated as the Editorial Board Member/ Reviewer of many peer-reviewed journals, including Elsevier, Springer, BMI, IOS Press, Bentham Science, Thieme Publishing Group, etc. His research areas are Data Mining, Optimization, Machine Learning, Cloud Computing, Artificial Intelligence and Object-Oriented Programming.

## AUTHOR PROFILE



**Dr. Ashutosh Kumar Dubey** received his PhD degree in Computer Science and Engineering from JK Lakshmipat University, Jaipur, Rajasthan, India. He is currently in the department of Computer Science and Engineering, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India. He is also the Senior Member of IEEE and