AN EFFICIENT DATA SECURITY MECHANISM BASED ON DATA GROUPS IN CLOUD COMPUTING ENVIRONMENT

Rajdeep Singh[1], Dr. R. K. Pateriya[2]

[1]*Research Scholar, Computer Science & Engineering, Mewar University, Chittorgarh, Rajasthan, India*

[2]*Research Supervisor, Computer Science & Engineering, Mewar University, Chittorgarh, Rajasthan, India*

Abstract: *In this paper k-means algorithm has been applied for finding the related group based on the content. In this phase the text and image data will be clustered based on the weight matrix. Then information part (IP) and content part (CP) has been calculated and evaluated. Then advanced encryption standard (AES), data encryption standard (DES) and blowfish algorithm has been applied in the combination or separately based on the threshold and data sensitivity. There are total four sensitivity and threshold levels in our approach. For the lower level sensitivity AES has been applied. For the middle level sensitivity AES + DES has been applied. For the high-level sensitivity AES + DES + Blowfish algorithm has been applied. In case of image AES + Blowfish has been applied. The results are found to be efficient in terms of generating keys automatically for individual files based on IP and CP. The time taken in the encryption and decryption process in the proposed system is found to be less due to the efficient management of data security and level wise security as per the need. The proposed framework is also efficient in the overall delay and processing due to better security management in comparison to the traditional methods. .*

Keywords: Data security, Cloud computing, data categorization, K-means.

1. INTRODUCTION

The main benefit of the cloud computing association is the resource computing and the resource controlling in case of storage [1]. It provides the data scaling more easily and it can be stored at the large-scale providing high performance the users with the computational capability. So, the above capability is strong enough for showing the strength in terms of organization, business prospective and it is also ease in the user adoptability and data configurability [2, 3].

It also provides great flexibility for the third-party data centers which shows the storage and processing phenomena in terms of storage and processing [4-6].

The service model encompasses the organization and it serves as a framework and provides the ease in resource sharing [7]. The major issues can be classified as the issues by the source means the providers who provides the clouds and the other side is the huge customers. It should be organized in such a way that the security concern should be fulfill in the two-way direction [8]. Distributed storage is comprised of many disseminated assets, yet at the same time goes about as one, either in a unified or a helpful capacity cloud design. It is exceptionally flaw tolerant through excess and conveyance of information [9]. In 2018, Torkura et al. [10] discussed about the cloud storage brokerage systems. They have suggested that the security risk assessment is important for the cloud storage brokers (CSBs). They have analyzed a threat modeling schema. It has been used for the analysis for identifying the security threats and risks in cloud brokerage systems. Their schema working mechanism is done through by generating attack trees, attack graphs, and data flow diagrams. For supporting the schema, they have presented the common configuration scoring system (CCSS). They have demonstrated the efficiency by devising CCSS base scores. In 2018, Tang et al. [11] discussed about securing cloud storage services. It has been done through the trusted Blockchain. They have suggested that ChainFS hardens the cloud-storage security against forking attacks. They have implemented ChainFS system on Ethereum and S3FS. Their results show low overhead. In 2019, Xu et al. [12] discuss about the balancing of service openness and security control. They have presented the theoretical discussion in this regard. It is based on Nash equilibrium. They

have considered two conditions based on the quantitative assessment. They have derived complete service openness investment based on their analysis. Based on this they have provided the optimal strategy for coordinating investment in both service and security. In 2019, Sharma et al. [13] discussed about the cloud computing as the depository. They have provided the store cyberspace data which may be applicable as a service to its user. They have suggested that the data security is needed as the data have been stored from different medium. They have suggested the use of multiple encryption technique for the importance of data security and privacy protection. In 2019, Almtrf et al. [14] discussed the vitality of cloud computing. They have suggested the reluctant for the cloud computing adoption. The major important aspects suggested by the authors are data security, privacy, and trust issues. They have suggested that the adoption and addressing these issues are important. They have developed a user privacy framework for emerging security model. It includes access control, encryption and protection monitor schemas. In 2019, Zhu et al. [15] discussed the use of cloud computing for the expansion of computing and storage resources with respect to IoT. They have suggested that the integrity of remote data should be verified for the data integrity and availability. They have suggested the maximum approaches uses for the data integrity are RSA based which has higher computational overhead. So, they have proposed, we propose a scheme of data integrity short signature algorithm (ZSS signature), for the verification. It supports the trusted third party (TPA) security. It reduces the computational overhead. They have shown that their approach has a higher efficiency and safety. In 2019, Singh et al. [16] discussed cloud computing in terms of e-healthcare services administration framework. They have proposed a model for the e-healthcare services administration framework. It is based on distributed computing. They have applied nominal crime and forget passwords constraints. They have suggested that it provides the high correctness rate for secure information access and recovery.

2. METHODS

Our framework has the following phases.

1. Data selection and weight matrix
2. K-Means
3. IP and CP
4. Hybrid security

2.1 Data selection and weight matrix

In this phase data selection and weight matrix assignment has been performed. First the data has been selected. Then weight based on the data size and random attribute information have been assigned.

2.2 K-Means

Then k-means clustering has been applied on the weight matrix obtained from the previous phase. In this phase the text and image data will be clustered based on the weight matrix. In our framework maximum five groups can be generated for the alikeness measure and checking.

2.3 IP and CP

IP has been calculated based on the total attributes which follow the threshold level which is generated automatically. CP shows the total value generated based on the complete attribute accumulated based on the scale. Based on the IP and CP sensitivity of the data has been generated.

2.4 Hybrid Security

Figure 1 shows the working process of the complete approach. It shows the process and encryption mechanism also. There are total four sensitivity and threshold levels in our approach. For the lower level sensitivity AES has been applied. For the middle level sensitivity AES + DES has been applied. For the high-level sensitivity AES + DES + Blowfish algorithm has been applied. In case of image AES + Blowfish has been applied.

3. RESULT & DISCUSSION

IP and CP will help in the unnecessary security complexity calculation burden as well as well as efficient in timely calculation. IP and CP have been used for data sensitivity calculation. IP shows the frequency of the repeated terms occurrences. CP shows the sensitivity of the data overall based on the attributes scale and the IP. So, based on this encryption algorithm has been applied. It will be helpful in time saving and reducing the time complexity. The hierarchy used for encryption and decryption are DES, AES + DES and AES + DES + Blowfish. In case of images only blowfish algorithm has been applied.

Figure 1 shows the IP values for 10 random samples. Figure 2 shows the CP values for 10 random samples. Figure 3 shows the IP values for 15 random samples. Figure 4 shows the CP values for 15 random samples. Figure 5 shows the IP values for 20 random samples. Figure 6 shows the CP values for 20 random samples. Figure 7 shows the IP values for 20 random samples 1. Figure 8 shows the CP values for 20 random samples 1. Figure 9 shows the flowchart of the working procedure.

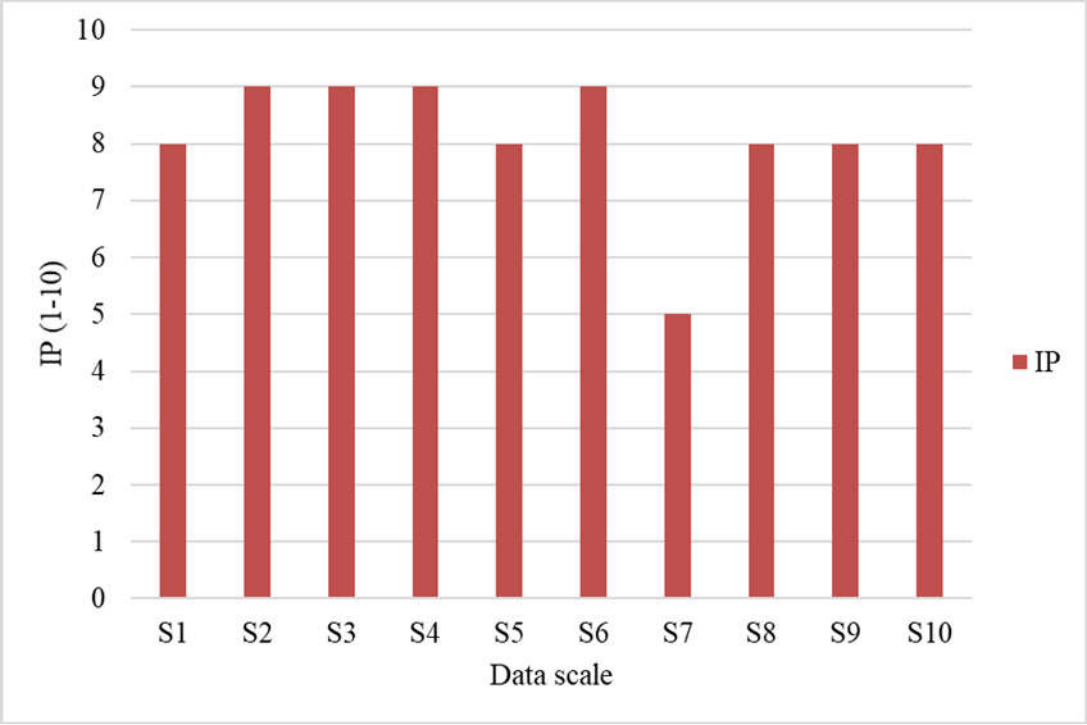

Figure 1 IP values for 10 random samples

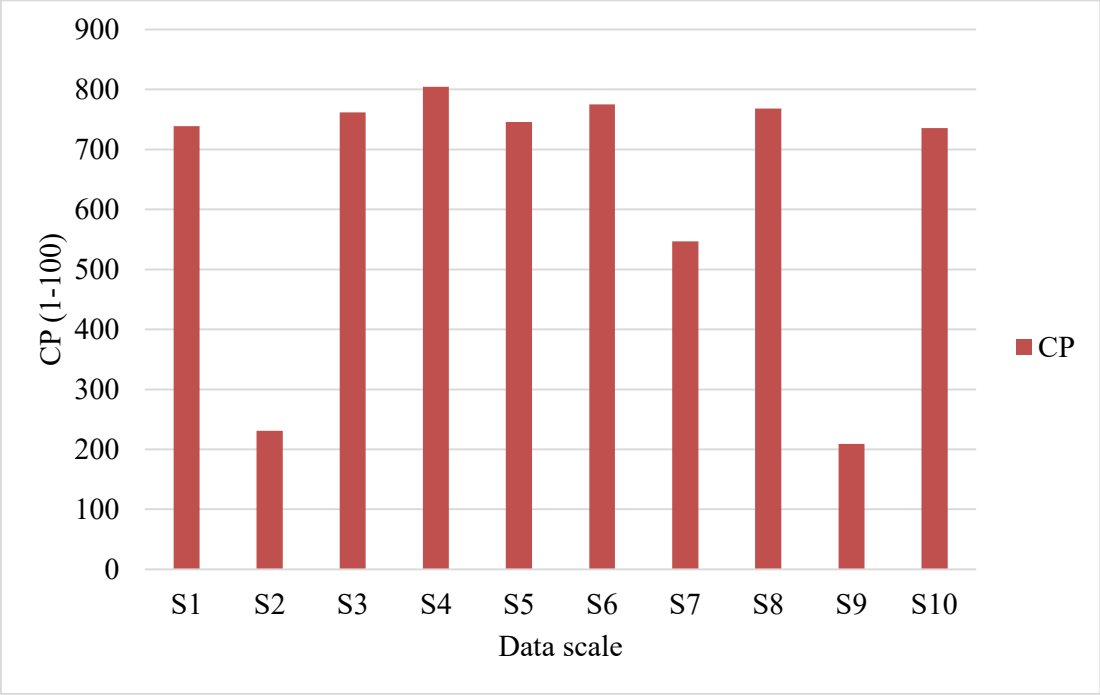

Figure 2 CP values for 10 random samples

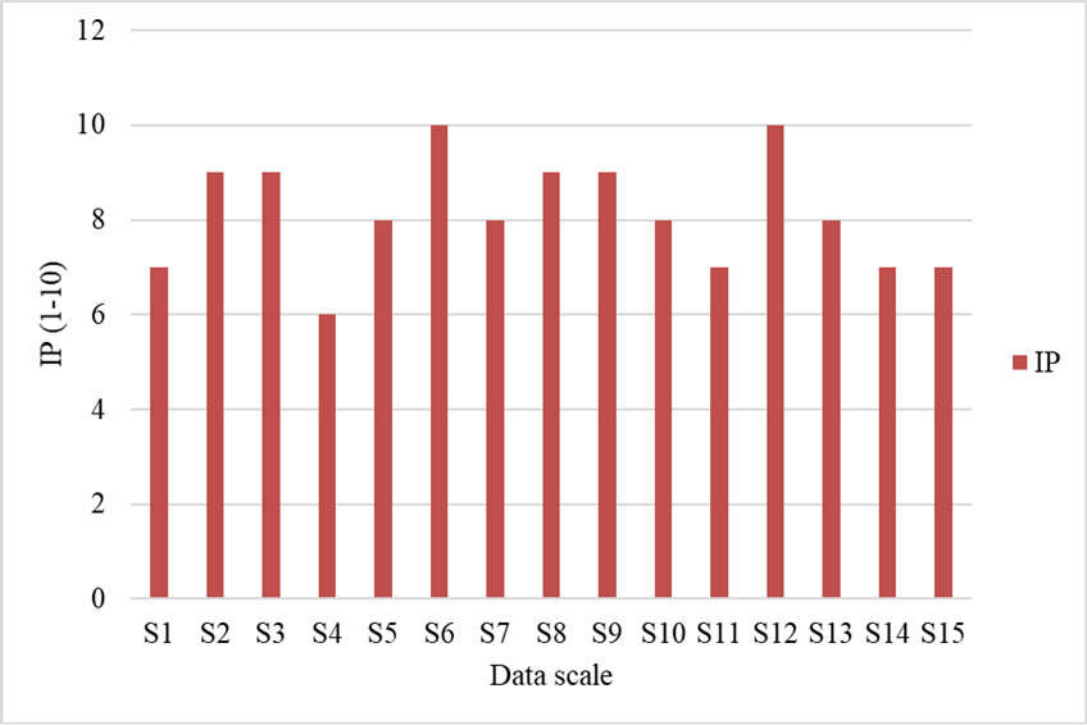

Figure 3 IP values for 15 random samples

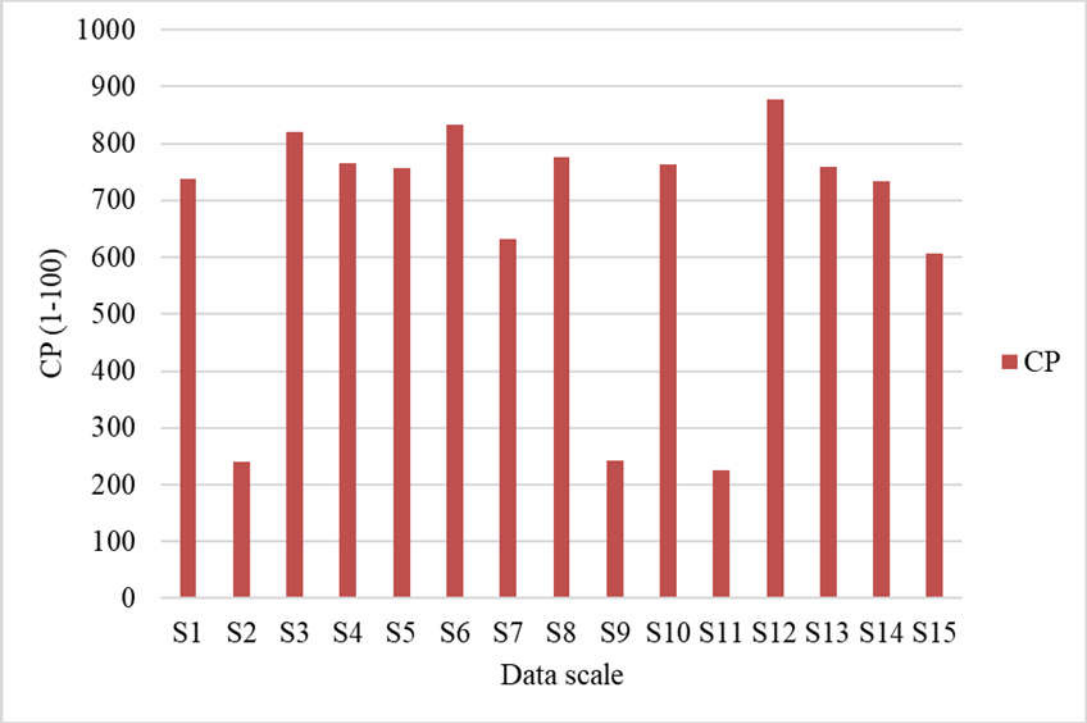

Figure 4 CP values for 15 random samples

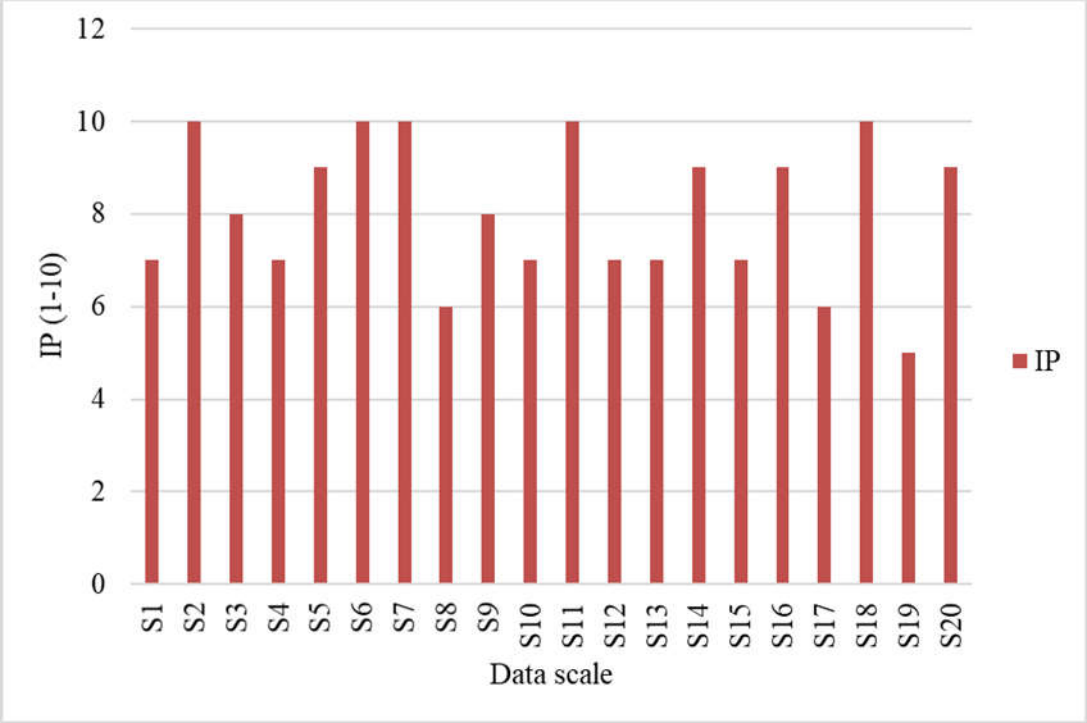

Figure 5 IP values for 20 random samples

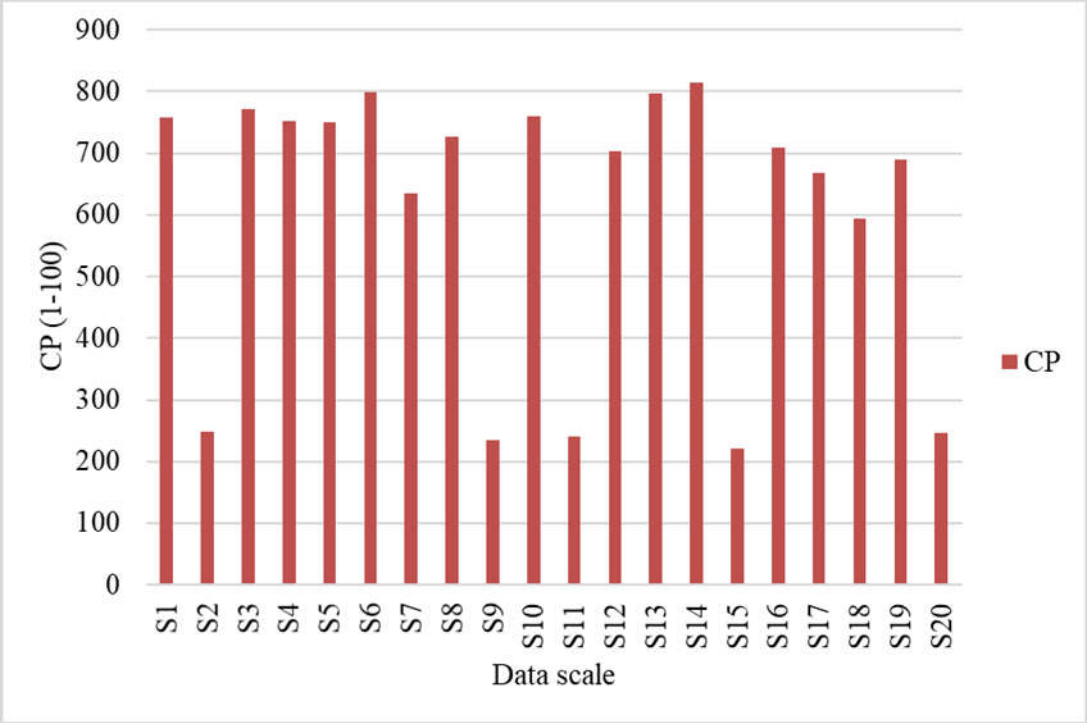

Figure 6 CP values for 20 random samples

Figure 7 IP values for 20 random samples 1

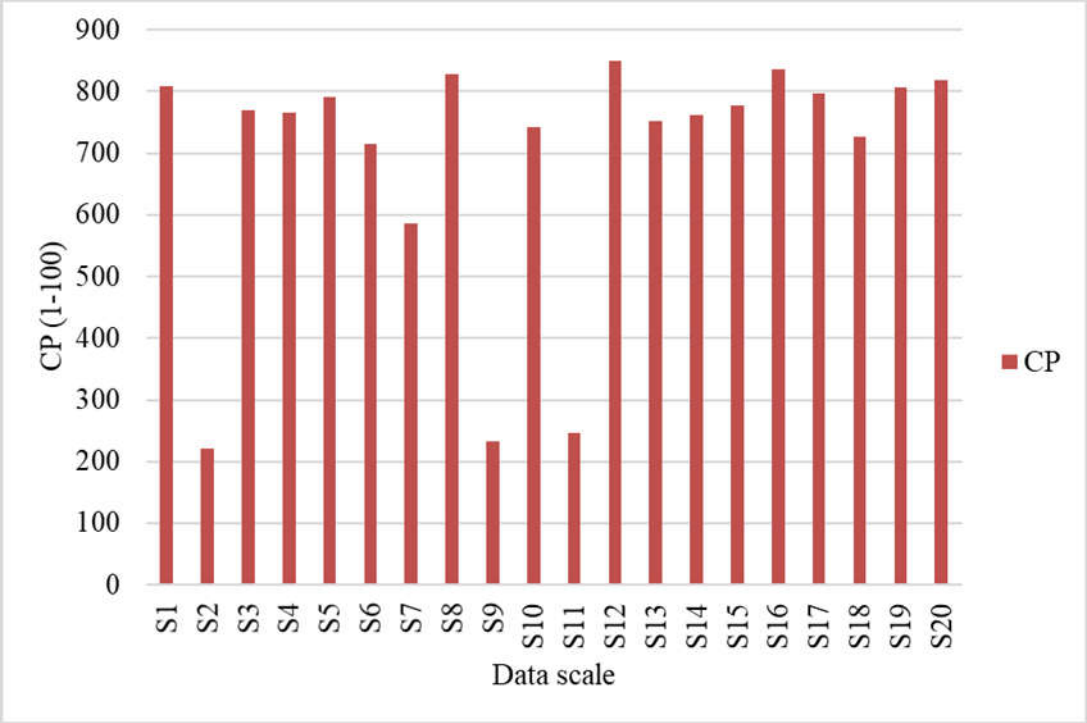

Figure 8 CP values for 20 random samples 1

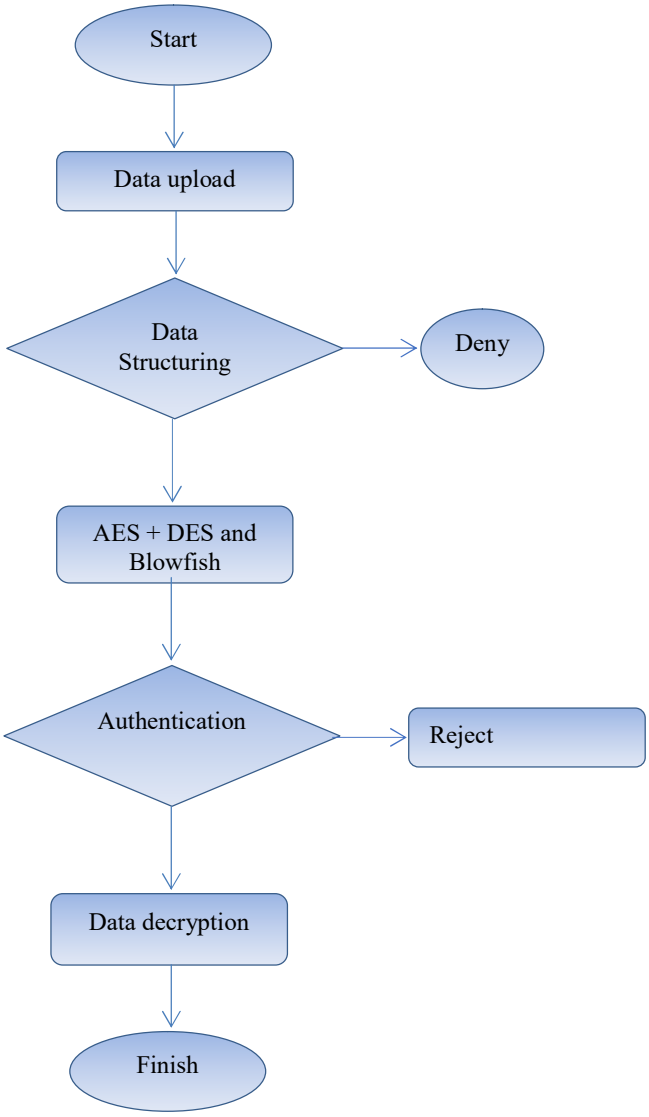

Figure 9 Flowchart of the working procedure

Figure 10 shows the attribute-based comparison. It clearly shows that the hybrid key combination of the proposed system is more efficient in producing number of keys. Figure 11 shows the key variation security comparison. It also shows a greater number of variabilities in comparison to the traditional approach.

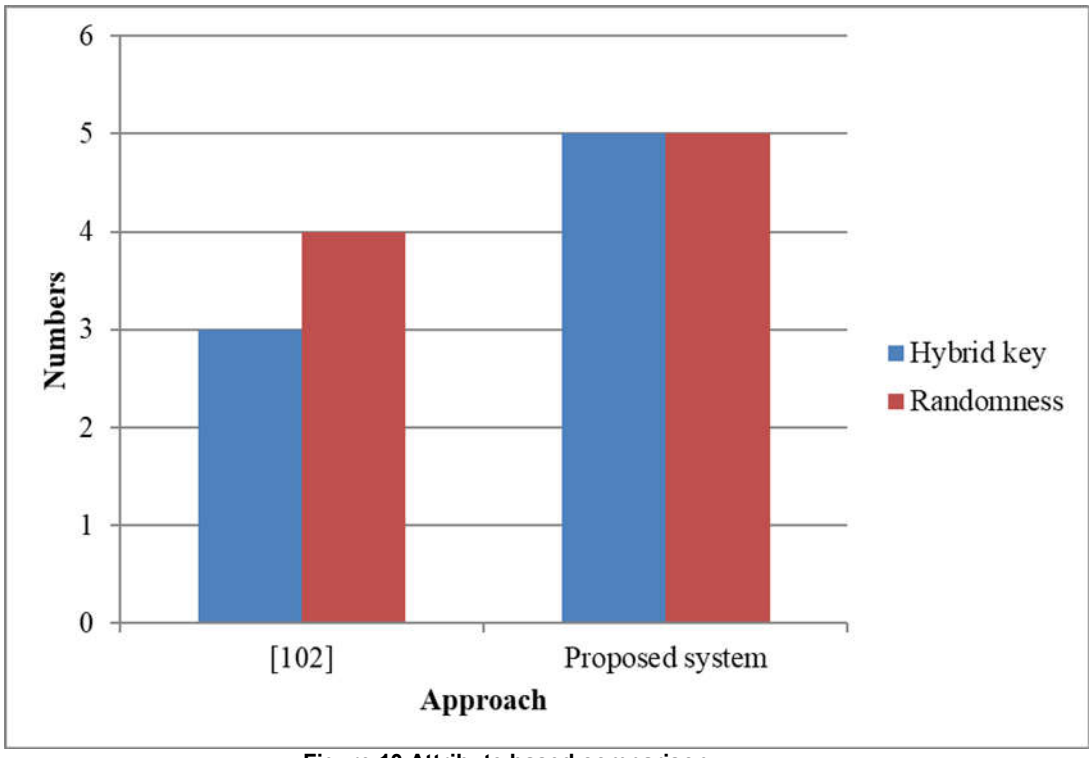

Figure 10 Attribute based comparison

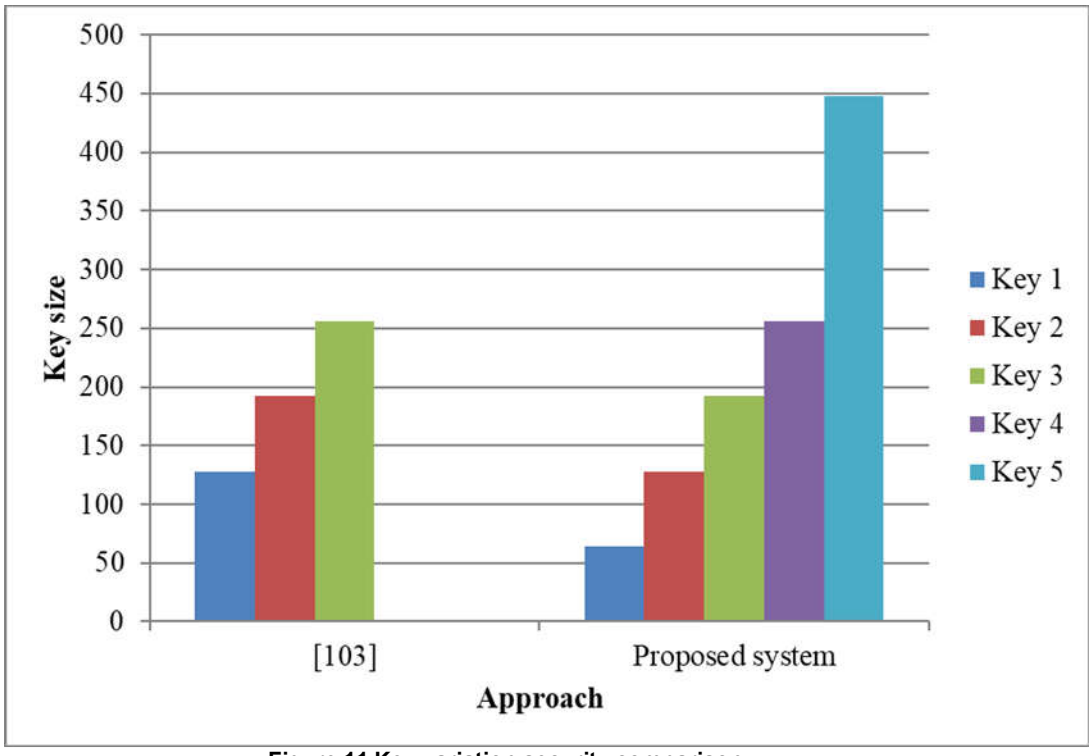

Figure 11 Key variation security comparison

4. CONCLUSION

Firstly, weight values have been assigned based on the file size and the segment content. It has been preprocessed based on random attribution scaling. Attribute weight scale is based on the scale of 1-100. The attributes scaling on the scale of 1-10. It has been prepared for partitioning based on the content alikeness and sensitivity of the data. For clustering k-means algorithm has been applied with k=5 at max as per the user partitioning requirement. Two types of data structure have been considered. Text and images have been considered. Then IP and CP have been calculated and evaluated. It has been evaluated and calculated for data sensitivity analysis. There are total four sensitivity and threshold levels in our approach. The time result and categorization show the effectiveness of our approach.

It has been evaluated and calculated for data sensitivity analysis. There are total four sensitivity and threshold levels in our approach. For the lower level sensitivity AES has been applied. For the middle level sensitivity AES + DES has been applied. For the high-level sensitivity AES + DES + Blowfish algorithm has been applied. In case of image AES + Blowfish has been applied.

The results are found to be efficient in key generations based on sensitivity key variations and attribute-based comparison.

REFERENCES

[1] *Ali M, Bilal K, Khan SU, Veeravalli B, Li K, Zomaya AY. Drops: Division and replication of data in cloud for optimal performance and security. IEEE Transactions on Cloud computing. 2015 February 5; 6(2):303-15.*

[2] *Wang S, Liang K, Liu JK, Chen J, Yu J, Xie W. Attribute-based data sharing scheme revisited in cloud computing. IEEE Transactions on Information Forensics and Security. 2016 April 6; 11(8):1661-73.*

[3] *Salunkhe SD, Patil D. Division and replication for data with public auditing scheme for cloud storage. In2016 International Conference on Computing Communication Control and automation (ICCUBEA) 2016 August 12 (pp. 1-5). IEEE.*

[4] *Londhe A, Rao PP. Platforms for big data analytics: Trend towards hybrid era. In2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS) 2017 August 1 (pp. 3235-3238). IEEE.*

[5] *Boru D, Kliazovich D, Granelli F, Bouvry P, Zomaya AY. Models for efficient data replication in cloud computing datacenters. In2015 IEEE International Conference on Communications (ICC) 2015 June 8 (pp. 6056-6061). IEEE.*

[6] *Singla S, Singh J. Cloud data security using authentication and encryption technique. Global Journal of Computer Science and Technology. 2013 August 2.*

[7] *Shaikh R, Sasikumar M. Trust model for measuring security strength of cloud computing service. Procedia Computer Science. 2015 January 1;45:380-9.*

[8] *Descher M, Masser P, Feilhauer T, Tjoa AM, Huemer D. Retaining data control to the client in infrastructure clouds. In2009 International Conference on Availability, Reliability and Security 2009 March 16 (pp. 9-16). IEEE.*

[9] *Li M, Yu S, Zheng Y, Ren K, Lou W. Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. IEEE transactions on parallel and distributed systems. 2012 March 19;24(1):131-43.*

[10] *Torkura KA, Sukmana MI, Meinig M, Cheng F, Meinel C, Graupner H. A threat modeling approach for cloud storage brokerage and file sharing systems. InNOMS 2018-2018 IEEE/IFIP Network Operations and Management Symposium 2018 April 23 (pp. 1-5). IEEE.*

[11] *Tang Y, Zou Q, Chen J, Li K, Kamhoua CA, Kwiat K, Njilla L. ChainFS: Blockchain-Secured Cloud Storage. In2018 IEEE 11th International Conference on Cloud Computing (CLOUD) 2018 July 2 (pp. 987-990). IEEE.*

[12] *Xu J, Liang C, Jain HK, Gu D. Openness and Security in Cloud Computing Services: Assessment Methods and Investment Strategies Analysis. IEEE Access. 2019 February 22; 7:29038-50.*

[13] *Sharma Y, Gupta H, Khatri SK. A Security Model for the Enhancement of Data Privacy in Cloud Computing. In2019 Amity International Conference on Artificial Intelligence (AICAI) 2019 February 4 (pp. 898-902). IEEE.*

[14] *Almtrf A, Alagrash Y, Zohdy M. Framework modeling for User privacy in cloud computing. In2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC) 2019 January 7 (pp. 0819-0826). IEEE.*

[15] *Zhu H, Yuan Y, Chen Y, Zha Y, Xi W, Jia B, Xin Y. A Secure and Efficient Data Integrity Verification Scheme for Cloud-IoT Based on Short Signature. IEEE Access. 2019 June 24; 7:90036-44.*

[16] *Singh I, Kumar D, Khatri SK. Improving The Efficiency of E-Healthcare System Based on Cloud. In2019 Amity International Conference on Artificial Intelligence (AICAI) 2019 February 4 (pp. 930-933). IEEE.*