ADV MATH SCI JOURNAL
Advances in Mathematics: Scientific Journal **9** (2020), no.6, 3459–3475
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.6.26 Spec. Issue on RDESTM-2020

FRAMEWORK TO ENHANCE THE QOS AND SECURITY IN CLOUD ENVIRONMENT

DEEPAK AHLAWAT, AMANDEEP KAUR, AND DEEPALI GUPTA[1]

ABSTRACT. The paper discusses the genetic algorithm-based clustering and compare it with existing clustering techniques in terms of accuracy. Ratio of Distribution of clusters, Memory Consumed and processing times of GA-Based clustering is also compared with existing clustering algorithms. A Framework to optimize the storage capabilities in cloud environment is also designed. By applying the framework, QoS parameters are studied in two scenarios viz., without security and with security, for the various clustering techniques.

1. INTRODUCTION

Clustering is an important data mining tool for examining data. Data clustering is a technique in various fields of computer science and related fields. While data mining may be considered as the main source of clustering, it is widely used in other fields like data-mining, energy research, machine learning, computer networking, pattern recognition, and hence many research works has been done from the very beginning, researchers have been dealing with their complexity and computational costs and as a result, they are dealing with clustering algorithms for increased scalability and speed [1,2].

[1]*corresponding author*

2010 *Mathematics Subject Classification.* 62H30, 68T10, 68M25.

Key words and phrases. Improved Rank Similarity, Hybrid Similarity, GA, PDR.

1.1. **Clustering by means of similarity measures.** The similarity between two objects suggests that how similar is the data in the two objects. Similarity between two objects can be found by many measures such as, by finding the difference between the contents of the data objects. In context of text mining or finding the similarity between the text files, some of the similarity measures taken are like Cosine similarity, Gaussian similarity, Jaccard similarity, etc. These types of similarity measures are used because they convert the text data into vectors and then assigning them appropriate weights. After finding weights and finding distance between documents, the conclusions can be made easily that if the distance between the data objects is small then there is greater similarity otherwise there is less similarity between the data objects. Similarity based study can be very useful in some cases, for example, in plagiarism checking software, the plagiarism checking software find the similarity of words in the document and the data stored in its repository [3].

1.1.1. *Cosine Similarity.* Generally, the similarity is a concept of matching; i.e., how close or alike are the objects that are being compared. Similarity can be calculated by using the concept of vectors [4]. The measure between two vectors is determined by cosine similarity technique. In other words, documents having vector space. Basically, this technique is used to determine the cosine angle between two vectors. The angle measured between used to evaluate the orientation measurement. It is noted that Cosine similarity technique is used to calculate the degree not an angle between the documents. The similarity in Cosine can be constituted as below:

$$\cos\theta = \frac{(\vec{A}.\vec{B})}{\left\|\vec{A}\right\| \left\|\vec{B}\right\|}.$$

The documents in the vector space model are presented using the different terms named as term frequency (TF), inverse document frequency (IDF) or weighing scheme (TF-IDF). In the case of information retrieval, the cosine similarity of the two documents ranges from 0 to 1 because the term frequency (tf-idf weight) cannot be negative. The angle between the two frequency vectors cannot be greater than 90° [5].

1.1.2. *Link-Based Similarity.* Similarity measures may be categorized into similarity based either on content or on link-based similarity. Content-based measures quantify object similarities by comparing the contents of the objects concerned, such as texts and multimedia. On the other hand, link-based measures interpret the relationships between objects as links, and use the link information to quantify similarities. The more neighbors two objects have in common, the greater would be the similarity between the two [6].

1.1.3. *Jaccard Similarity*. Jaccard Similarity being a coefficient is also termed as the Jaccard Index which is utilized for computing the diversity and similarity of sample sets. It computes the similarity among the sample sets and is termed as the intersection set categorized by union size of the sample sets. Jaccard distance computes the dissimilarity among the sample sets and is corresponding to the Jaccard coefficient [7]. It finds the most common between two document sets

$$\text{Jaccard Index} = \frac{\text{Intersection of docuements}}{\text{Union of documents}}.$$

1.1.4. *Improved Rank Similarity*. Improved Rank similarity is the sum of cosine similarity between the Doc and Link Similarity (Rank Similarity) between the Doc [8]. totalImproved = totalcountcos[i]+rankcos[i]; Thus,

$$\text{Improved Rank Similarity} = \text{Cosine Similarity} + \text{Rank Similarity}.$$

1.1.5. *Hybrid Similarity.* The hybrid similarity is the combination of Cosine similarity and Jaccard similarity. The aim of designing hybrid similarity is to enhance the clustering mechanism. The elements are grouped in such a manner that documents above threshold similarity value are kept in one cluster and rests of the documents are kept in another cluster. The hybrid similarity is the combination of Cosine similarity and Jaccard similarity [9].

$$\text{Hybrid Similarity} = \text{Cosine Similarity} + \text{Jaccard Similarity}.$$

1.2. **Genetic Algorithm.** Today a vast amount of data is there and this data is present in different formats. To make useful use of the data, we divide this data into small groups and each group is considered as population. Genetic algorithm provides us an optimal solution, by iteratively applying the genetic

operators the optimal solution is found. A set of strings called chromosomes are combined. One fitness function is to be created and crossover of parent values are considered. If the fitness function is satisfied then mutation occurs otherwise the mutation doesn't take place. This process iteratively runs and leads to an optimal solution. After recombination (crossover and mutation) of candidate solution selection, a new population of candidate solutions can be created [10].

2. Related Work

Sitikhu et al. (2019) presented a comparison between different methods of semantic similarity. The study is conducted on two short text news articles. The techniques used are Cosine similarity with term frequency-inverse document frequency vectors, cosine similarity with word-to-vector vectors, soft cosine similarity with word-to-vector vectors. Among these three vectors, cosine with term frequency-inverse document frequency had the highest accuracy among the stated three techniques. The results were cross-validated and compared with the text in news articles. The accuracy word-to-vector can be increased by using the Document-to-Vector model instead of the word-to-vector model [11].

Haghighi et al. (2019) studied a dynamic resource management approach to clustering which is energy efficient also. The approach followed by the researcher is based on meta-heuristic and used in cloud environment in Iaas Platform. It has certain challenges like reducing the power consumption of Data-Centers and still following the SLA constraints. Researcher studied the Quality of service under SLA contraints in the variable nature and complex nature of cloud [12].

Basu et al. (2019) proposed a solution in the cloud environment by using the genetic algorithm. The researcher aims to reduce the overall power consumption of nodes in cloud environment, so that nodes are up to the mark and no node overloaded or lightly loaded. The objective function is defined by the author and genetic algorithm is applied on it in cloud environment [13].

Neysiani et al. (2019) described a solution to produce suggestions based on previous transactions. The solution is based on genetic algorithm and uses association rules and collaborative filtering. The collaborative filtering and genetic

algorithm with association rules provides the suggestions based on previous transactions. The experimental result shows that the algorithm proposed the researcher outperforms the MOPSO algorithm [14].

Li et al. (2019) presented cloud framework with the feature of security in media called secure media. The focus the researcher is on protecting multimedia data and services. The framework has two levels of data storage protection and one security access control protocol which is adapted to video services, i.e., it contains three levels in total. As revealed in the results, Sec-ABAC protocol is quite secure and more efficient when compared to the traditional ABAC protocol. However, some disadvantages were observed in this work such as, the deployment of the framework on Amazon Web Services is not included and also there is no analyse done on the performance in the results [15].

Zhu et al. (2018) in the work proposed a new Algorithm based on Cosine Similarity and called it CSA (Cosine Similarity Algorithm). The algorithm monitors the droplet generation frequency in a precise and quick manner. In the method when a video clip is captured of microscopic droplet generation with a high-speed camera, then the measurements can be by measuring the cosine similarities between the frames in the video clip. The work states that by evaluating cosine resemblance among the frames in the video clip, a microscopic droplet generation video clip taken by a high-speed camera can be calculated solidly. The experimental results clearly show that CSA system has proven to be very effective for the unfailing, practical, and keeping track online of complex droplet generation processes. This method has tremendous potential to be very significant in advancing the droplet microfluidic research and development [16].

P-Manickam et al. (2011) has done a comparative study of networking protocols; DSDV, DSR and AODV taking various parameters into account; throughput, end-to-end delay, packet delivery ratio (PDR). The network chosen by the researcher is 500X500 and the simulation is done on NS-2 simulator [17].

3. Proposed Work And Results

The proposed work first compares the precision, recall and hence accuracy of various clustering algorithms with the introduction of GA in Hybrid Clustering,

TABLE 1. Accuracy Comparison

Total Passed Queries	Accuracy K-Means (%)	Accuracy Cosine Similarity (%)	Accuracy Improved Rank Similarity (%)	Accuracy Hybrid Similarity (%)	Accuracy Using GA (%)
50	51.08	62.42	66.44	67.96	82.30
100	52.78	64.02	68.88	69.12	83.72
500	54.28	66.95	70.46	70.07	85.73
1000	55.58	68.87	72.18	70.85	87.11
1500	56.12	71.27	74.76	71.94	88.00

termed as GA-Based Clustering. After this the simulation is done in CloudSim and a framework is designed to employ this GA-Based clustering in Cloud. At last the results are compared for Throughput, Delay, PDR and Energy Consumption.

3.1. **GA-Based Clustering.** The focus is on creating premium clusters when there are no reference clusters for data. The proposed solution uses similarity indexes for the relation between the data element. It has considered hybrid similarity and on this Hybrid Similarity the Genetic Algorithm is applied. Genetic fitness function is created, if the child value is greater then the average then it is shifted to next cluster or mutated to next cluster otherwise remains in the same cluster.

3.1.1. *Accuracy*. For clustering, we have to find the best match between the class labels and the cluster labels,

$$\text{Precision} = \frac{\text{True Positive}}{(\text{True Positive}+\text{False Positive})} \tag{3.1}$$

$$\text{Recall} = \frac{\text{True Positive}}{(\text{True Positive}+\text{False Negative})} \tag{3.2}$$

$$\text{F-Measure} = \frac{(2 * \text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{3.3}$$

$$\text{Accuracy (Percent)} = (\text{F-Measure}) * 100. \tag{3.4}$$

It can be clearly seen from the results that Accuracy with GA is the Highest. So,

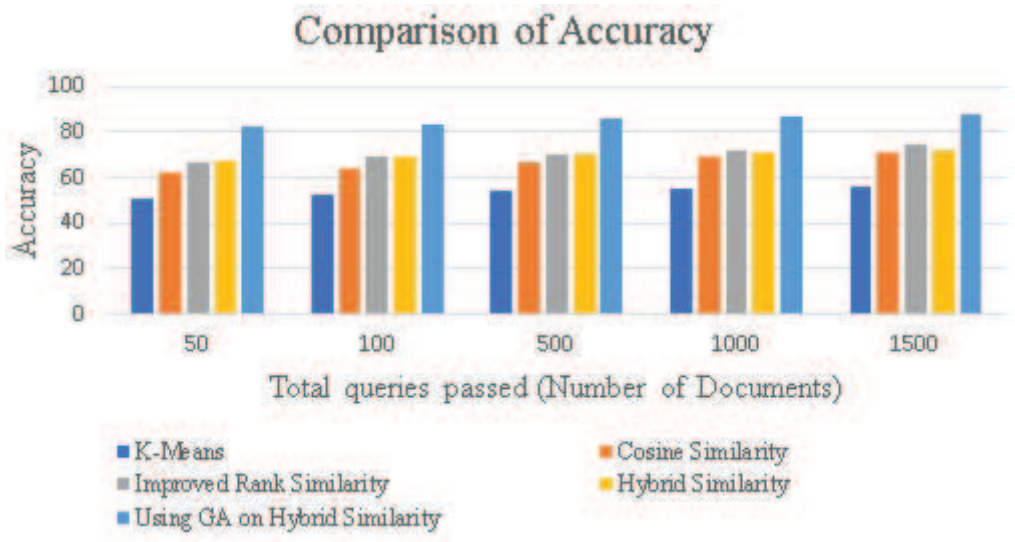

FIGURE 1. Comparison of Accuracy

genetic algorithm can be used to enhance the accuracy of the clustering algorithms like K-Means and Cosine similarity clustering. Results show accuracy of GA-Based clustering has an improvement of 27.97 Figure 1 shows the graphical representation of Accuracy comparison of various techniques. The high accuracy indicates that the proposed system has more accurate clusters then the previous techniques which will enhance the search and other process more accurate with the proposed technique.

The other findings with the GA-Based clustering are

1.Ratio of Distribution It is the total number of documents in the cluster with respect to the total number of documents collected at the time of initialization of the clustering.

$$\text{Ratio of Distribution} = \frac{\text{Total number of documents in one cluster}}{\text{Total number of document}}.$$

Table 2 demonstrate that the documents are distributed uniformly with a little margin after the application of Genetic Algorithm. It is apparent from the execution part that Genetic algorithm has been applied to cluster 1. Table 2 shows the ratio of distribution for cluster 1, 2 alone and then cluster 1 and 2 with GA. It can be concluded that there is an optimization of 50% in clusters after applying the genetic algorithm.

TABLE 2. RoD Comparison

Total Number of Documents	RoD Cluster 1	RoD Cluster 2	RoD Cluster 1 with GA	RoD Cluster 2 with GA
50	0.75	0.25	0.45	0.55
100	0.72	0.28	0.47	0.53
500	0.69	0.31	0.48	0.52
1000	0.67	0.33	0.51	0.49
1500	0.64	0.36	0.50	0.50

TABLE 3. Comparison of Memory Consumed

Number of Documents	Memory Consumed without GA (per document) KB	Memory Consumed with GA (per document) KB
50	47093	19985
100	9518	1657
500	332	228
1000	147	97
1500	110	90

2. Memory Consumed:
Below Table 3 shows the comparison of memory consumed in the clustering per document. It can be seen that the average memory consumed per document without GA is 11.17 MB and with GA is 4.30 MB. As the memory consumed in GA-Based Clustering is much less than the clustering without GA, it can be concluded that GA-Based clustering is more economical than without GA clustering.

3. Time of Formation and Optimization Time: It is the total time consumed in clustering the data files. The evaluation is done in milliseconds. The graphs are created by taking the time in seconds.
Table 6 shows the time of cluster formation and cluster optimization. It is obvious that the cluster formation time will be much more as compared to the cluster optimization time. As the numbers of documents are reduced when it comes to optimizing the set, it takes less time.
The maximum processing time is 1800 seconds for 1500 documents, while the maximum optimization time is 178 seconds.

3.2 Framework to optimize the storage capability of Cloud

Figure 2, summarizes the Framework developed for the thesis work. The first

TABLE 4. Time of Formation Vs Optimization Time

Number of Documents	Time of Formation (seconds)	Optimization Time (seconds)
50	6.0	1.0
100	11.0	2.0
500	62.0	18.0
1000	420.0	88.0
1500	1800.0	178.0

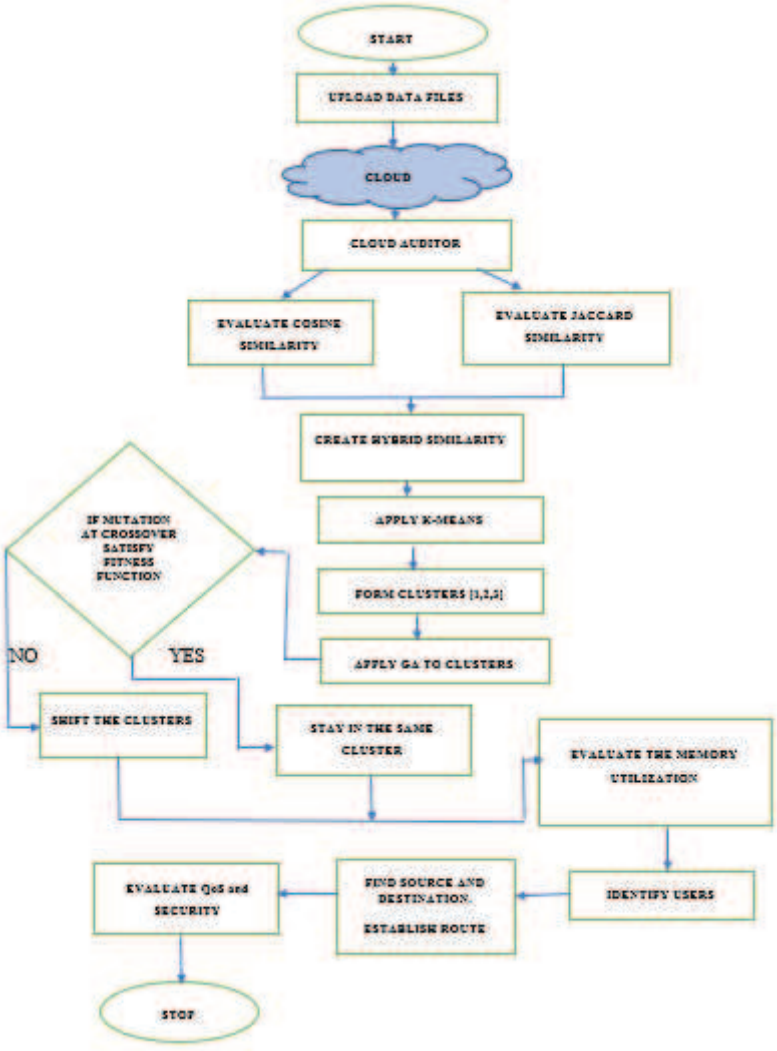

FIGURE 2. Framework to Optimize the Storage Capabilities in Cloud

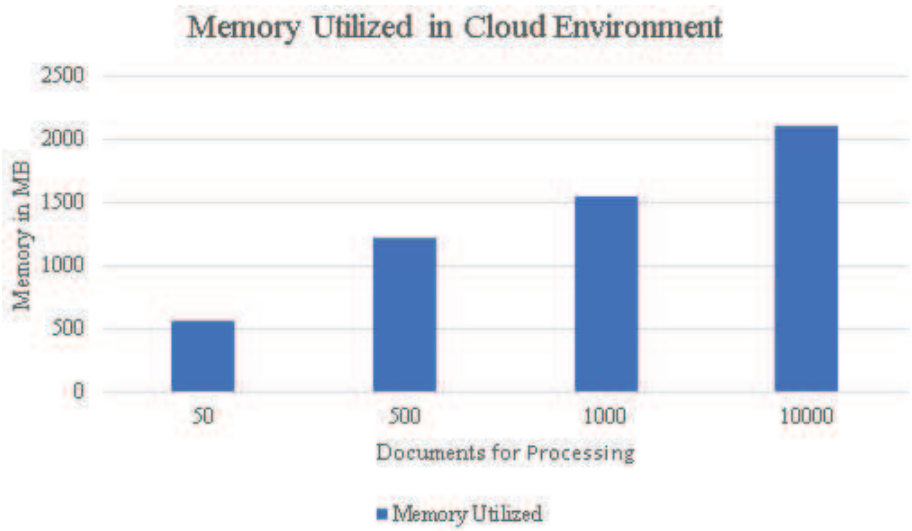

FIGURE 3. Memory Utilized in Cloud Environment

step is to upload the data files to the cloud. After that the Cloud Auditor audits the files. The files are the compared and processed to calculate the Cosine Similarity and Jaccard Similarity. Another type of similarity is then calculated i.e., Hybrid Similarity, which is the sum of Cosine and Jaccard Similarity. On the Hybrid Similarity, K-Means Algorithm is applied to create 3 clusters viz., [1], [2], [3]. Then Genetic Algorithm is applied on these clusters and see for the fitness Function. Two case arises, First, If Mutation at crossover satisfies the fitness function, then stay in the same cluster, Second, If Mutation at crossover doesn't satisfies the fitness function, then shift the cluster. From the three created clusters in the cloud, evaluate the Memory and Bandwidth utilized. Also, after that create a network architecture by identifying Users and then find source and destination. Evaluate the QoS and Security parameters like Throughput, Delay, PDR and Energy Consumption to see the results. A better throughput, less delay, more PDR and less Energy consumption is desired to optimize the storage capabilities in Cloud. Figure 3 shows the Memory Utilized during processing of different number of documents in Cloud scenario. The X-axis in the figure is for total documents for processing and Y-axis shows the memory in MB. 10000 documents have been taken for the processing. It is being concluded that the memory usage enhances with an enhancement in the document to execute.

10,000 documents are considered for the execution and around 2100 Mb 2.1GB memory is used on the hard disk.

3.2.1 Storage Optimization in Cloud Environment

As we can observe that enterprises are shifting towards the cloud, the architecting, administration and management of data in cloud has emerged as a critical subject for consideration. Today, there are too many options to choose from so, it becomes very crucial and significant for enterprises to understand every aspect of structuring and storing data. The main reason of enterprising shifting to Cloud is that the cloud storage services are inexpensive and based on pay-as-you-go, but it should be assured that the application performance is not compromised. Cloud storage can be used with different sized and different types of data. It may be used for sharing and also for recovery of data in case of some disaster. According to the amount of data the cloud storage can be arranged in different proposed tiers. According to Microsoft Azure and Amazon's AWS [18], the Storage in Cloud environment can be viewed from two points-Throughput and Scalability. To optimize the storage, the resources must be scalable and should have more throughput. Throughput is termed as no. of units of information a system can process at a particular time interval, or in other words, number of packets transferred in unit time. When many computers run concurrently then the throughput of the system act as a parameter to measure the compare the effectiveness of the system. A greater value of throughput increases the efficiency of the system and is desirable. The results of throughput are shown in the next section with the integration of security in the code. As shown in Figure 3, it can be seen as there is more demand for memory, it can be allocated from the available memory to the memory in use. Memory utilized increases with the number of documents and the free memory is allocated. As the demand increases it can be scaled according to the needs. 3.3 QoS and Security If the resources are optimally used then it will optimize the storage capabilities. The proposed framework is compared for QoS and Security by taking throughput, end-to-end delay, PDR and energy consumption. The proposed framework first considers the transfer of packets without security and after that the transmission by introducing security into it.

Data Communication Based Results are based on throughput, Delay, Packet Delivery Ratio and Energy Consumption for evaluating the QoS and Security Parameters.

The proposed work model evaluates the users based on the Quality of Service (QoS) and Security. There are certain assumptions which are undertaken in this evaluation as follows

(a) Users in the cloud network share the resources between them.
(b) There is a source user who supplies the data to a demanding user also termed as the destination user.
(c) The user can't transfer the data directly to the demanding user and hence intermediate users are also required based on the location.
(d) The nearby location users will be selected to transfer the data from one end to another.

The algorithm used for making the study are AODV (Ad-Hoc On-Demand Distance Vector Routing) and Firefly algorithm (Fitness function). The network is designed in such a manner that a track record is managed at every set of transfer. The security pattern of the proposed framework keeps a record of every transfer so that the nodes can be verified each and every time the data is transferred through the path. The security is measured in terms of health of the deployed nodes and communicating nodes. Each communicating nodes is evaluated for health check up with as the checkpoints are applied for the simulation check-up in order to see that everything is going find in the network or not. The route discovery process aims to transfer the data from one end to another with the nearest node in place so that the nodes or the communicating nodes which are far from the transmission line are not allowed to access the current system or architecture. This architecture saves a lot time cost and reduces the security breach issue in the network.

The network keeps a record of the network node and the communication behaviour to establish route more precisely and quickly when a data transfer is attempted next time. If the network does not face any intrusion in the network, it surely will transfer a greater number of packets. It not only saves time which is required for the route search but also tracks the consumption of the power in the network. If an unexpected power consumption is found to be observed in the network, an analysis of the stored network is evaluated using machine

learning architecture. The training part of the machine learning algorithm takes the consumed power of the network per path delivery along with the path labels as the name associated target value of the consumption and hence a supervised mechanism is applied. Against the simulation result, if the identified path value does not match with the training value, the identified path is termed to be suspected and hence analysis of each involved cloud member is analysed. In the simulation, network is iterated five times.

3.3.1 Throughput Throughput is termed as no. of units of information a system can process at a particular time interval, or in other words, number of packets transferred in unit time. When many computers run concurrently then the throughput of the system act as a parameter to measure the compare the effectiveness of the system. A greater value of throughput increases the efficiency of the system and is desirable.

$$\text{Throughput} = \frac{\text{Total Number of Packets}}{\text{Total Time Taken}}.$$

The throughput of the proposed is evaluated with the integrated security aspects and parameters. The throughput of the proposed system is found to be comparatively high when the system is analysed with no security integration.

3.3.2 Packet Delivery Ratio (PDR) Another parameter is PDR (Packet Delivery Ratio), the quantitative relation between the numbers of transferred packets from a traffic source and numbers of received packets to a traffic sink is known as the Packet Delivery Ratio (PDR). A high packet delivery ratio means that a greater number of packets reaching the destination. It is desirable that PDR should be high.

$$\text{PDR} = \frac{\text{Total Packets at destination end}}{\text{Total Packets at Source End}}.$$

3.3.3 Delay Delay refers to the time taken for a packet to be transmitted across a network from source to destination.

3.3.4 Energy Consumption Energy consumption is the amount of energy or power used. Energy consumption is studied in two scenarios viz., without security and with security. When there is no security then the packet travels to destination via intermediate nodes and there is no machine intelligence so it may be possible that the packet again comes to the same failed node and it has

TABLE 5. Comparison of Data Communication Parameters

	Cosine with Security	Proposed Technique without Security	Proposed Technique with Security
Throughput	30.2	24.2	44.3
PDR	71%	60%	85%
Delay	141ms	121ms	104ms
Energy Consumption	148mj	126mj	106mj

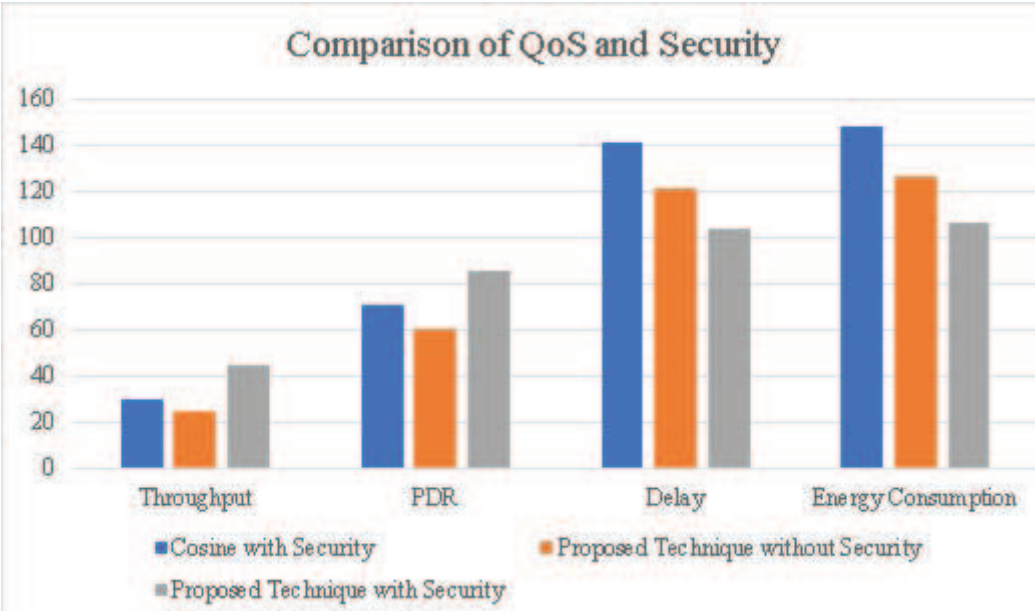

FIGURE 4. Comparison of QoS and Security

to cover a greater number of nodes to reach destination. In security scenario, a smaller number of nodes required to reach the destination and the energy consumption is less.

3.3.5 Comparison of QoS and Security Various clustering techniques are studied and compared here. Cosine Similarity with security, proposed technique without security and proposed technique with security are compared. Table 5

The throughput for the proposed technique with security is 44.3 and without security is 24.2 and that of cosine similarity with security is 30.2. Throughput increases by 83.05 % by implying security in proposed framework. The PDR for

the proposed technique with security is 85% and without security is 60% and that of cosine similarity with security is 71%. PDR increases by 41.67% by implying security in proposed framework. The Delay for the proposed technique with security is 104ms and without security is 121ms and that of cosine similarity with security is 141ms. Delay decreases by 14.05% by implying security in proposed framework. The energy consumption for the proposed technique with security is 106mj and without security is 126mj and that of cosine similarity with security is 148mj. Energy Consumption decreases by 15.87% by implying security in proposed framework.

4. Conclusion

Genetic Algorithm-Based clustering can be used to enhance the accuracy of the clustering algorithms like K-Means and Cosine similarity clustering. Results show accuracy of GA-Based clustering has an improvement of 27.97% over Cosine similarity-based clustering. The Accuracy of GA-Based clustering is greater than Improved Rank and Hybrid Similarity also. The RoD of GA-Based is also better than other techniques. Memory Consumption and Optimization time is also less in GA-Based Clustering. The proposed framework shows that the memory utilized in cloud is scalable and the utilization increases with the increase in the number of documents for processing. For 10000 documents approx. 2.1 GB of memory is utilized. Throughput increases by 83.05% by implying security in proposed framework and 46.68% more than Cosine with security. PDR increases by 41.67% by implying security in proposed framework and 19.71% more than Cosine with security. Delay decreases by 14.05% by implying security in proposed framework and 26.24% less than Cosine with security. Energy Consumption decreases by 15.87% by implying security in proposed framework and 28.37% less than Cosine with security.

References

[1] A. A. Alazeez, S. Jassim, H. Einckm: *An Enhanced Prototype-based Method for Clustering Evolving Data Streams in Big Data*, ICPRAM 2017 173–183.

[2] I. M. Masciari, G. M. Mazzeo, C. Zaniolo: *Clustering Goes Big: CLUBS-P, an Algorithm for Unsupervised Clustering Around Centroids Tailored for Big Data Applications*,

In 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP) 2018 Mar 21, 558–561.

[3] S. GARCIA, J. LUENGO, F. HERRERA: *Data pre-processing in data mining*, New York: Springer, 2015.

[4] K. SRINIVAS, B. K. RANI, A. GOVRDHAN: *Applications of data mining techniques in healthcare and prediction of heart attacks*, International Journal on Computer Science and Engineering, (IJCSE), **2**(2) (2010), 25–35.

[5] J. JAYABHARATHY, S. KANMANI, N. SIVARANJANI: *Correlation based multi-document summarization for scientific articles and news group*, Proceedings of the International Conference on Advances in Computing, Communications and Informatics 2012 Aug 3, 1093–1099.

[6] P. CALADO, M. CRISTO, M. A. GONCALVES, E. S. DE MOURA, B. RIBEIRO-NETO, N. ZIVIANI: *Link-Based Similarity Measures for the Classification of Web Documents*, Journal of the American Society for Information Science and Technology, **57**(2) (2006), 208-–221.

[7] S. NIWATTANAKUL, J. SINGTHONGCHAI, E. NAENUDORN, S. WANAPU: *Using of Jaccard Coefficient for Keywords Similarity*, Proceedings of the International Multi Conference of Engineers and Computer Scientists 2013, **1** (2013), 13-15.

[8] D. AHLAWAT, A. KAUR, D. GUPTA: Enhancement of the Accuracy and QoS in Clustering of Data, ICRITO- IEEE sponsored International Conference at Amity University, Noida, 4-5 June 2020.

[9] D. AHLAWAT, D. GUPTA: *An Enhanced Mechanism of Big Data Clustering in Cloud*, 4th International Conference on Cyber Security and privacy in communication networks (ICCS) 2018, Special Session 'Analysis of Big Data and Mining Data'. (Elsevier), Jaipur, 2018.

[10] J. MCCALL: *Genetic algorithms for modelling and optimisation*, Journal of Computational and Applied Mathematics, **184**(1) (2005), 205–222.

[11] P. SITIKHU, K. PAHI, P. THAPA, S. SHAKYA: *A Comparison of Semantic Similarity Methods for Maximum Human Interpretability*, IEEE International Conference on Artificial Intelligence for Transforming Business and Societyy, 2019.

[12] M. A. HAGHIGHI, M. MAEEN, M. HAGHPARAST: *An Energy-Efficient Dynamic Resource Management Approach Based on Clustering and Meta-Heuristic Algorithms in Cloud Computing IaaS Platforms*, Wireless Personal Communications, **104**(4) (2019), 1367–91.

[13] S. BASU, G. KANNAYARAM, S. RAMASUBBAREDDY, C. VENKATASUBBAIAH: *Improved Genetic Algorithm for Monitoring of Virtual Machines in Cloud Environment*, Smart Intelligent Computing and Applications, **15** (2019), 319–326.

[14] B. S. NEYSIANI, N. SOLTANI, R. MOFIDI, M. H. NADIMI-SHAHRAKI: *Improving Performance of Association Rule-Based Collaborative Filtering Recommendation Systems using Genetic Algorithm*, International Journal of Information Technology and Computer Science, **11**(2) (2019), 48–55.

[15] H. LI, C. YANG, J. LIU: *A novel security media cloud framework*, Computers and Electrical Engineering, **74** (2019), 605–615.

[16] X. ZHU, S. SU, M. FU, J. LIU, L. ZHU, W. YANG, G. JING, Y. GUO: *A Cosine Similarity Algorithm Method for Fast and Accurate Monitoring of Dynamic Droplet Generation Processes*, www.nature.com/scientificreports, (2018) 1–14.

[17] P. MANICKAM, T. GURU BASKAR, M. GIRIJA, D. MANIMEGALAI: *Performance Comparisons of Routing Protocols in Mobile Ad Hoc Networks*, International Journal of Wireless and Mobile Networks (IJWMN), **3**(1) (2011), 98–106.

[18] https://blogs.dxc.technology/2018/02/15/five-keys-to-cloud-storage–optimization/

MAHARISHI MARKANDESHWAR UNIVERSITY
SADOPUR, AMBALA, HARYANA, INDIA
Email address: deepakahlawat1983@gmail.com

MAHARISHI MARKANDESHWAR UNIVERSITY
SADOPUR, AMBALA, HARYANA, INDIA
Email address: amandeepkaur@mmumullana.org

CHITKARA UNIVERSITY INSTITUTE OF ENGINEERING AND TECHNOLOGY
CHITKARA UNIVERSITY, PUNJAB, INDIA
Email address: deepali.gupta@chitkara.edu.in