

An Efficient Data Segmentation and Replication Technique for Cloud using Fruit Fly Optimized Centrality Measures

S.Periyanchi, K.Chitra

Abstract— Cloud computing is a creating worldview to give dependable and versatile framework permitting the clients (data proprietors) to store their data and the data customers (clients) can get to the data from cloud servers. This paradigm decreases stockpiling and upkeep cost of the data proprietor. However, cloud data storage still gives rise to security related problems. In case of shared data, the data face both cloud-specific and insider threats. In this work, we propose FOA (fruit fly optimization algorithm) optimized centrality measure fragmentation and replication of information in the cloud for optimum performance and security that consider both security and performance issues. FOA is a technique for deducing global optimization based on the foraging character of the fruit fly. The sensory perception of the fruit fly is superior than that of other species, particularly the sense of smell and vision. In our methodology, we divide a data files and replicate the fragmented data over the cloud nodes using FOA centrality measures. Every one of the cloud node just store a single information data fragment that ensures even if there arise an occurrence of a successful attack, no important information is shown to the attacker. We also compare the performance of the our methodology with other standard replication schemes. Observed results shows higher level of security and performance improvements.

I. INTRODUCTION

Cloud Computing denotes accumulating and accessing information and programs over the internet instead of your computer's hard drive. The cloud is just a image for the Internet. Cloud drops technology is about securing data over the cloud because when the users outsource their data to unauthorized executive control, it gives rise to security related concerns. The information gives and take may manifest because of third party by different users and nodes inside the cloud

. On contacting the data by third party users and processes must be not permitted, otherwise, one weak individual will put the entire cloud at threat. Also, the security practice must be unease in the improvement of data healing time.

Some of the threats are 1) neighboring person which makes a way to an enemy to bypass the user lines. 2) The offsite information storage cloud utility needs users to move information in clouds idea and shared location that may origin a choice of security concerns. 3) Grouping and

elasticity of the cloud, agree to the physical properties to be shared among lots of users. These shared properties may be reallocates to other users for the result in data compromise at some instant time.4) A multi-tenant virtual atmosphere may be the outcome in VM to get away the limitations of virtual machine monitor(VMM) which can hold up to other VMs may contact to third party data. 5) In cross tenant virtualized set-up, due to inappropriate media purification, the consumer records can also gets leaked.

Data security tactics:

In the Cloud drops technology, a file is partitioned into trashes, and duplicates the garbage data over the cloud nodes. Every one of the nodes supplies only a single bit of a particular data file that guarantees that if there is any successful attack, where the attacker cannot get a meaningful data.

The nodes are accumulating the trashes are divided by a certain space by using graph T-coloring to make illegal an enemy of deducing the spots of the trashes.

In this piece of writing, we suggest a partitioning and duplicating of Data which is about securing data over the cloud while ensuring protection and security is optimized by an fruit fly optimization algorithm. Therefore, these services will result for the rise of security concerns. We present FOA optimized centrality measure based fragmentation and reproduction of data in the cloud for optimum enactment and security that divide user files into fragments and replicates them at planned locations inside the cloud

The remainder of an article is ordered as follows. Section 2 provides an overview of the literature methodology. Section 3 presents our network model and fruit fly algorithm for cloud drops technology. Section 4 discusses the proposed model with an algorithm and Section 5 shows results and discussion. Section 6 concludes this thesis.

II LITERATURE REVIEW

G.n.Gan et al [2] suggests an optimized algorithm based on genetic annealing algorithm for scheduling the task. In this way the QOS request of various kind of assignment is thinking about in cloud, this causes effectively conveyance of assets in the cloud condition. L. Guo et al[3] talked about a Multi-target Task Assignment issue by utilizing Particle Swarm Optimization (PSO) calculation in Cloud Computing. In this writing,

Revised Version Manuscript Received on April 12, 2019.

S.Periyanchi, Research scholar, Bharathiyar university,Coimbatore T.N ,India.(Email : periyanchi@yahoo.com)

Dr.K.Chitra, Asst professor ,Dept of CSE, Govt arts college,Melur, T.N ,India.



the both expense and time are enhanced where it expanded the productivity and lessens the expense of the system.

Suraj et al[14] focus on calculation and communication cost. Monir and Mohammed et al [15] suggests Divisible Load theory (DLT) to partition the entire load into fractions and scheduling them to reduces the entire total load cost among available processors.

DLT normally used in the area of aerospace ,image processing, non linear processing as equipment to access computation load. It analyzes the optimal fraction of work load that is to be allocated to each processor.

Durga devi et al[5] propose a hybrid optimization algorithm by combining Metaheuristic Swarm Optimization (MSO) and Cuckoo Search (CS)] .This optimization used for data scheduling in cloud environment.over all objective for applying optimization is ti reduce span time,response time and increase a resource utilization. T. Loukopoulos et al. [16] present a Genetic Algorithm (GA) based data replica strategy to tackle continuously changing read/write demands . Mazhar Ali et al [1] introduced a DROPS methodology ,entire data was divided and stored in a multiple node of cloud and jointly considers security and recovery time

III PREFACE

Before go into the clarification of propose system, we present the related thoughts in the accompanying for the solace of the perusers.

3.1 Data partition

The data security of a large - scale cloud depends upon the security of the framework all in all and the data security of isolated cloud nodes. A compelling assault into a solitary node extreme for unfortunate casualty node, yet additionally danger for different nodes. The data on injured individual hub completely known to assailant due to putting away of whole document in a single node.

To reduce a quantity compromise data ,a data files divided into fragment and stored on a separate nodes [17]. A successful attack o single node only give a access to portion of data .

Likewise, if an aggressor is finding out about the areas of the information fragments, the probability of finding entire information fragments on most of the nodes is uncommonly low

3.2 CM-Centrality Measures

The centrality measures have been used for investigating network and recognizing focal nodes of these networks[22].Such networks are meant by directed or undirected and weighted or unweighted graphs. A directed graph, likewise named digraph G, which is our thought in this work ,it is spoken to by (V, E) pair, where V is a lot of hubs and E is the arrangement of directed associations among hubs. The greatest utilized centrality measures to recognize a focal hub in a digraph, which we will audit in the accompanying subcategories, are: degree, betweenness and closeness.

3.2.1DCM-Degree Centrality Measure

The DC characterized as local centrality measure since it is determined by just its directed associations. The DC of a node is determined by including of its approaching (id-indegree) and active (od-outdegree) association loads:

$$CD(v)=\sum(id(v)+od(v))$$

where the id(v) is the summation of association loads entering hub v, and the od(v) is the summation of association loads leaving hub v

3.2.2 BCM-Betweenness Centrality Measure

The BC characterized as worldwide centrality measure since it is determined dependent on the briefest ways between hub combines in the graph. The BC of a node is determined by including the extent of most limited ways between nodes combines that experience that nodes .for the directed graph G=(V, E), the BC of a node v is characterized as:

$$CB=\sum_{s \neq v \neq t \in V} \sigma_{st}(v) / \sigma_{st}$$

Where σ_{st} speaks to number of most limited ways from node s to node t and $\sigma_{st}(v)$ is the quantity of briefest ways from s to t that goes through node v.

3.2.3 CCM-Closeness Centrality Measure

Like betweenness, the CCM is a global centrality measure determined dependent on the most limited ways idea. It discovers how much a node is near every single other node in the graph. For a directed graph G=(V, E), the closeness centrality of a node v is characterized as:

$$Cc(v)=\frac{1}{\sum_{t \in V} d_G(v,t)}$$

Where $t \neq v$, and $d_G(v,t)$ is the most brief way between nodes v and t. As to cloud network , the CCM is a proportion of how rapidly a node speaks with different nodes in the cloud network .

3.2.4 SCM-Strangeness Centrality Measure

The strangeness of a node n is the greatest separation to any node from a node n [17]. A node is increasingly focal in the network, on the off chance that it is less offbeat. Officially, the capriciousness can be given as:

$$S(v_a) = \max_b d(v_a, v_b);$$

where $d(v_a, v_b)$ speaks to the space between node v_a and node v_b . It additionally noticed that in our systems the centrality measures presented above appear to be important and significant than utilizing straightforward jump tally sort of measurements.

IV. PROPOSED FOA TECHNICHS

This section tells how to deal with the cloud with some optimization problems. Here the fruit fly In order to understand the FOA the basics and an algorithm has been proposed in below.

Basic idea of FOA

It is an effective strategy for deriving centrality measure enhancement dependent on the chasing conduct of the fly. On contrasting, the tactile perception of the fruit

fly is superior to that of different species, particularly the feeling of smell and representation. The olfactory organ of a fruit fly can assemble various scents from the climatic air, but then a nourishment source 40km away. Accordingly, the fruit fly flies to the staple, utilizes its intense mental picture or vision to find the nourishment and where its colleagues assemble, and afterward it makes a trip toward that path to achieve the goal.

Fruit fly individualities for searching food are described, as below:

4.1 General FOA algorithm

Step1: An initial random location of a fruit fly is given as Init X; Init Y

Step2: Random path and distance of searching for food using the sense of smell of a fruit fly individual.

$$X_f = X + \text{Random Value}$$

$$Y_f = Y + \text{Random Value}$$

Step3: As the site of foodstuff cannot be known, the space (Spc) to the origin is estimated before the decision value of smell concentration (Smf) is calculated; this result is the reciprocal of space.

$$Spc_f = \sqrt{x^2 + y^2}$$

$$Sm_f = \frac{1}{spc_f}$$

Step4: The smell concentration value or concentration decision function (Smf) substituted in (FF)fitness function

$$Smell_f = \text{Func}(Sm_f)$$

Step5: For finding the maximum value, determine the fruit fly with the high smell concentration

$$[bestSmell\ bestguide] = \max(Smell)$$

Step6:

To retain best concentration value and x ,y ranges ,fly goes towards the area by vision

$$Smell_{best} = bestSmell$$

$$X = X(bestguide)$$

$$Y = Y(bestguide)$$

This steps must be processed in an iterative optimization manner, which repeatedly executes the steps 2-5, and judge whether the smell concentration is better than the previous iterative smell concentration, if yes, execute Step 6.

From the above algorithm, it describes the basic characteristics of fruit fly on searching food in a short time with the considerations of smell and vision. This characters of fruit fly helps to optimize the best food in a shortest distance with the comparison of several foodstuffs. Likewise in our proposed methodology, by using this sensible optimization of fruit fly we can easily find out the following parameters with an effective manner in a cloud drops technology.

- FOA for solving strangeness
- FOA for solving closeness
- Fruit Fly Optimization for finding solution for global centrality analysis

4.1.1 FOA for solving strangeness Centrality Measures

In this chapter , Fly optimization is used to find out the strangeness value of centrality measures respectively:

$$Y = 3 - X^2$$

and in the previous, the strangeness value is set to 3 according to fruit fly optimization.the initial value for random postion is [0,10].for iteration the fly path and space zone is set to [-1,1].

Therefore the algorithm for finding the strangeness centrality measure of an iterative search solution will be developed with the following steps which is similar to the general algorithm.

Randomized initial fruit fly swarm location

$$X=10*\text{rand}();$$

$$Y=10*\text{rand}();$$

For an individual fruit fly the random path and space for searching is

$$X_f(i)=X +2*\text{rand}()-1;$$

$$Y_f(i)=Y +2*\text{rand}()-1;$$

In the case of food location not known, strangeness of the origin calculated first and then smell concentration decision value calculated as reciprocal of strangeness

$$Spc_f(i)=(X_f(i)^2+Y_f(i)^2)^{0.5};$$

$$Sm_f(i)=1/ spc_f(i);$$

To find smell concentration of individual location of fly ,smell concentration decision value into concentration decision value function is

$$Smell_f(i) =3- Sm_f(i)^2;$$

To search a fly with maximum smell concentration (finding the strangeness of centrality measures) among the all nodes.

$$[bestSmell\ bestguide]=\text{strangeness}(Smell);$$

To Keep the best smell concentration value and x, y coordinate, and at this moment, the fruit fly swarm will use vision to fly towards that location so that the strangeness centrality measures between the nodes can be find out.

$$X=X(bestguide);$$

$$Y=Y(bestguide);$$

$$Smell_{best}=bestSmell;$$

This steps must be processed in an iterative optimization manner, which repeatedly executes the steps 2-5, and judge whether the smell concentration (i.e) the strangeness centrality measures is superior than the previous iterative centrality measure, if yes, then execute Step 6.

4.1.2 FOA for solving closeness centrality measures

In this chapter, optimization is used to find out the closeness value of centrality measures with functions of respectively:

$$Y = -5 - X^2$$

and in the past, the closeness value is set to -5 as per FOA methodology . initial random zone for fruit fly is [0,10],for iterative fruit fly random fly path space zone is [-1,1]

An algorithm for closeness centrality measures is given below

Step 1: X=10*rand();

$$Y=10*\text{rand}();$$

Step 2: X_f(i)=X +2*rand()-1;

$$Y_f(i)=Y +2*\text{rand}()-1;$$



An Efficient Data Segmentation And Replication Technique For Cloud Using Fruit Fly Optimized Centrality Measures

Step 3: $spc_f(i) = (X_f(i)^2 + Y_f(i)^2)^{0.5}$;
 $Sm_f(i) = 1 / spc_f(i)$;
 Step 4: $Smell_f(i) = -5 + Sm_f(i)^2$; // measuring closeness centrality
 Step 5: [bestSmell bestguide]=closeness(Smell);
 Step 6: $X = X(\text{bestguide})$;
 $Y = Y(\text{bestguide})$;
 $Smell_{best} = \text{bestSmell}$;

This entire steps will be processed in an iterative optimization manner, which repeatedly executes the steps 2-5,

and judge whether the smell concentration (i.e) the strangeness centrality measures is superior than the previous iterative centrality measure, if yes, then execute Step 6.

4.1.3 Fruit fly applied for solving global CA

In this chapter use a FOA to search global maxima value, the function is:

$$y = \sin(x)/x$$

From the above expression, we can evaluate the global centrality measures by using the fruit fly optimization algorithm which has been given below

Step 1: $X = 10 * \text{rand}()$;
 $Y = 10 * \text{rand}()$;
 Step 2: $X_f(i) = X + 2 * \text{rand}() - 1$;
 $Y_f(i) = Y + 2 * \text{rand}() - 1$;
 Step 3: $spc_f(i) = (X_f(i)^2 + Y_f(i)^2)^{0.5}$;
 $Sm_f(i) = 1 / spc_f(i)$;
 Step 4: $Smell_f(i) = \sin(Sm_f(i)) + Sm_f(i)$; // for finding the best global centrality
 Step 5: [bestSmell bestguide]=global(Smell);
 Step 6: $X = X(\text{bestguide})$;
 $Y = Y(\text{bestguide})$;
 $Smell_{best} = \text{bestSmell}$;

Therefore, it can be accomplished that the proposed FOA is a more powerful method to solve the cloud security problem by comparing with all the previous methods. The proposed methodology solves the third party problem of cloud computing data during transmission to create a safe cloud computing network for the all users. The experiment verifies that the cloud computing security by using FOA is designed in this paper is secure and effective.

V. PERFORMANCE EVALUATION & RESULTS

Our proposed work as pursues, when client sends the information record to cloud. The cloud supervisor framework plays out: (a) division of data, (b) FOA upgraded centrality based nodes choice and stores one part over every one of the selected node, and (c) nodes determination for fragments replication. The cloud administrator keeps up record of the piece position and is thought to be a safe entity. Once document is portioned into fragments cloud trough discovers cloud nodes for section situation. The choice completely based giving equivalent significance to both security and recovery time. In request to offer a superior access time pick a most focal node for information arrangement. For that

reason, our technique utilizes the idea of centrality to lessen access time utilizing FOA. The centralities decide how focal a node depends on various measures as talked about in segment 3 & 4. We select node with three CM, specifically: (a) DCM, (b) BCM, and (c) CCM.



Figure 1 data access delay comparison (sec)

We execute evaluation of the system model developed in Section iv. Our FOA method compared with results of well known replication methodology, namely: (i) Greedy algorithm, and (j) Genetic Replication Algorithm (GRA). The main performance indicators communication delay of node. In figure 1, Data access delays measured as an average time elapsed from the moment of sending data file request and having the requested data arrived. As expected, access delay becomes smaller for replicas located by FOA optimized centrality node selection compare with other techniques

VI CONCLUSIONS

We proposed the FOA based data replication methodology, considering both security level retrieval time. The client data was divided and the fragments are spread over multiple nodes. The divided and replicated data are stored with node distance of FOA optimized centrality measures. Our method ensured that no significant information accessible by attacker when there is a attack. It also ensures no node have a more than one copy of data fragment of same. The performance of the FOA replication methodology was compared with classical replication techniques. The results of the simulations shows that the simultaneous focus on the security and performance, resulted in increased security level and performance improvement

REFERENCES

1. Mazhar Ali et al, "Drops: Division And Replication Of Data In Cloud For Optimal Performance And Security," IEEE Transactions On Cloud Computing.
2. G. n. Gan et al "Genetic simulated annealing algorithm for task scheduling based on cloud computing environment," 2012 National Conference on Integrated Systems
3. L. Guo et al, "Multi-Objective Task Assignment in Cloud Computing by Particle Swarm Optimization," International Conference on Communications, Networking and Mobile Computing,
4. P. Durgadevi et al, "Task Scheduling using Amalgamation of Metaheuristics Swarm Optimization Algorithm and Cuckoo Search in Cloud Computing Environment", Journal for Research, vol.1,no. 9, 2015.
5. K. Hashizume et al, "An analysis of security issues for cloud computing," Journal of cluster computing, Vol. 6, No. 1, 2015, pp. 1-13.
6. M. Hogan et al, "NIST cloud computing standards roadmap," NIST Special Publication, July 2011.



8. W. A. Jansen, "Cloud hooks: Security and privacy issues in cloud computing," In 44th Hawaii IEEE International Conference on System Sciences (HICSS), 2011, pp. 1-10.
9. Y. Tang, P. et al, "Secure overlay cloud storage with access control and assured deletion," Journal of cluster and Secure Computing, Vol. 9, No.5, Nov. 2013, pp. 903-917.
- 10.M. Tu et al, "On the optimal placement of secure data objects over Internet," cloud sensor journal, pp. 14-14, 2003.
- 11.W. T. Pan, "A new fruit fly optimization algorithm: taking the financial distress model as an example," Knowledge-Based Systems, vol. 26, pp. 69-74, 2013.
- 12.S. M. Lin, "Analysis of service satisfaction in web auction logistics service using a combination of fruit fly optimization algorithm and general regression neural network", Neural Computing and Applications, vol 22 (3), pp. 783-791, 2013.
- 13.Suraj Pandey1 et al., "A Particle Swarm Optimization based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments,".
- 14.M. Abdullah et al, "Cost-based Multi-QoS Job Scheduling Using Divisible Load Theory in Cloud Computing", Procedia Computer Science, vol. 18, pp. 928-935, 2013.
- 15.T. Loukopoulos et al, "Static and adaptive distributed at a replication using genetic algorithms," Journal of data security, Vol. 64, No. 11, 2004, pp.1270- 285.