ISSN:0975-9646

Ant Colony Optimization Technique for Secure Various Data Retrieval in Cloud Computing

K.Sriprasadh [1]
Research Scholar
Bharath University

M.Prakash Kumar [2]
PG Scholar
Department of Computer Science and Engineering
Jai Mathajee Engineering College

Abstract **- Data retrieval is the largest task in any large database, in the world's largest data bases like cloud data retrieval is the one of the major issue. Retrieving the data and processing the query over cloud server is very difficult. Many searching technique are used for retrieving the data from cloud servers. It can be retrieved through an optimization technique. There are many data retrieval techniques like Boolean Symmetric Searchable Encryption, Secure Ranked Keyword Search over Encrypted Data and Privacy-Preserving Multi-keyword Ranked Search over Encrypted Cloud Data, Over Encrypted Data in Cloud Computing, etc. In this Ant Hill optimization technique is prominent techniques which give hands to solve this problem in efficient way comparing other optimization technique listed above . Ant hill method has not been used in reliability design, yet it is a method that is expressly designed for combinatorial problems. An ant optimization technique for data retrieval is devised and tested on a well known suite of problems from the literature. It is shown that the ant colony method performs with little variability over problem instance. It is competitive with the best known heuristics for data retrieval.**

Key words: Data retrieval, Ant hill Optimization ,cloud agent

1. INTRODUCTION

Cloud computing based on the services, the services are known to be Information as a Service (IAAS) , Platform as a Service (PAAS), Software as a Service (SAAS) in this Information as a service plays a vital role where the large information are stored for as the database where it can be retrieved by the user, for the future use.

While retrieving the information there is chance of the data loss or wrong data can be retrieved to avoid this algorithm can be applied to solve this problem.

II.IAAS

Cloud inforamtion as a service (IaaS) is a highly automated offering where compute resources, complemented by storage and networking capabilities, are offered to the customer on-demand. Here the customers can store the data and can retrieve their whenever they require.So it's not surprising that subcategories of cloud based storage services fall under IaaS monitor.

A. Storage as a Service

Storage is one of those necessities that only grow over time. It can be constant struggles to maintain enough storage capacity and manage it effectively. These solutions have interactive self service portals that allow administrators to provision storage, transfer data to different tiers of storage ,dispatch specific data sets to different media (such as disk or tape),and add or remove storage as needed. Storage – as – a –service providers also have the latest storage technologies and virtually limitless capacity.

Tiers generally include fast storage for system disk and bulk storage for file serving. And as with other types of IaaS, storage enterprises pay only for what they use one of the concerns that organization have about moving storage to the cloud is security. After all, storage systems contain sensitive information about the organizations and its users or customers. Cloud –based storage has the security controls to ensure that all data is stored securely in data center facilities, with extremely high availability.

B. Disaster Recovery and Backup as Service

The idea behind moving disaster recovery to the cloud is to ensure that organizations have uninterrupted access to data and applications ,regardless of emergencies such as a power solutions always include redundancy and automatic failover to ensure ongoing access reducing downtime to nearly zero

Many solutions also employ continuous data protection (CDP) which allows to multiple versions of all data sets to be recovered. This gives users the ability to restore the data to any point in time. Data and applications are stored in secure offsite facilities

There are two basic options when it comes to disaster recovery as a service backup and restore to the cloud. With the first option organizations retain application and data on their own premise .but back up and data to the cloud to restore it to hardware on their own premise when disaster occurs.

With the second option data is restored to virtual machine in the cloud For mission –critical applications and resources that must be recovered quickly and complete the best.

III. INFRASTRUCTURE-AS-A-SERVICE (IAAS) SECURITY ISSUES

IaaS provides a pool of resources such as servers, storage, networks, and other computing resources in the form of virtualized systems, which are accessed through the Internet. Users are entitled to run any software with full control and management on the resources allocated to them. With IaaS, cloud users have better control over the security compared to the other models as long there is no security hole in the virtual machine monitor. They control the software running in their virtual machines, and they are responsible to configure security policies correctly. However, the underlying compute, network, and storage infrastructure is controlled by cloud providers. IaaS providers must undertake a substantial effort to secure their

systems in order to minimize these threats that result from creation, communication, monitoring, modification, and mobility.[1] Here are some of the security issues associated to IaaS.

A. Virtualization

Virtualization allows users to create copy, share, migrate, and roll back virtual machines, which may allow them to run a variety of applications. However, it also introduces new opportunities for attackers because of the extra layer that must be secured. Virtual machine security becomes as important as physical machine security, and any flaw in either one may affect the other. Virtualized environments are vulnerable to all types of attacks for normal infrastructures; however, security is a greater challenge as virtualization adds more points of entry and more interconnection complexity. Unlike physical servers, VMs have two boundaries: physical and virtual.

B .Virtual machine monitor

The Virtual Machine Monitor (VMM) or hypervisor is responsible for virtual machines isolation; therefore, if the VMM is compromised, its virtual machines may potentially be compromised as well. The VMM is low-level software that controls and monitors its virtual machines, so as any traditional software it entails security flaws. This useful feature can also raise security problems. An attacker can compromise the migration module in the VMM and transfer a victim virtual machine to a malicious server. Also, it is clear that VM migration exposes the content of the VM to the network, which can compromise its data integrity and confidentiality.[1]

C. Shared Resource

VMs located on the same server can share CPU, memory, I/O, and others. Sharing resources between VMs may decrease the security of each VM. For example, a malicious VM can infer some information about other VMs through shared memory or other shared resources without need of compromising the hypervisor. Using covert channels, two VMs can communicate bypassing all the rules defined by the security module of the VMM. Thus, a malicious Virtual Machine can monitor shared resources without being noticed by its VMM, so the attacker can infer some information about other virtual machines.[2]

D. Public VM Image Repository

In IaaS environments, a VM image is a prepackaged software template containing the configurations files that are used to create VMs. Thus, these images are fundamental for the overall security of the cloud. One can either create her own VM image from scratch, or one can use any image stored in the provider's repository. For example, Amazon offers a public image repository where legitimate users can download or upload a VM image. Malicious users can store images containing malicious code into public repositories compromising other users or even the cloud system. For example, an attacker with a valid account can create an image containing malicious code such as a Trojan horse. If another customer uses this image, the virtual machine that this customer creates will be infected with the hidden malware. Moreover, unintentionally data leakage can be introduced by VM replication. Some confidential information such as passwords or cryptographic keys can be recorded while an image is being created. If the image is not "cleaned", this sensitive information can be exposed to other users. VM images are dormant artifacts that are hard to patch while they are offline.[2]

E. Virtual machine rollback

Furthermore, virtual machines are able to be rolled back to their previous states if an error happens. But rolling back virtual machines can re-expose them to security vulnerabilities that were patched or re-enable previously disabled accounts or passwords. In order to provide rollbacks, we need to make a "copy" (snapshot) of the virtual machine, which can result in the propagation of configuration errors and other vulnerabilities.[3]

F. Virtual machine life cycle

Additionally, it is important to understand the lifecycle of the VMs and their changes in states as they move through the environment. VMs can be on, off, or suspended which makes it harder to detect malware. Also, even when virtual machines are offline, they can be vulnerable; that is, a virtual machine can be instantiated using an image that may contain malicious code. These malicious images can be the starting point of the proliferation of malware by injecting malicious code within other virtual machines in the creation process.

G. Virtual Networks

Network components are shared by different tenants due to resource pooling. As mentioned before, sharing resources allows attackers to launch cross-tenant attacks. Virtual Networks increase the VMs interconnectivity, an important security challenge in Cloud Computing. The most secure way is to hook each VM with its host by using dedicated physical channels. However, most hypervisors use virtual networks to link VMs to communicate more directly and efficiently. For instance, most virtualization platforms such as provide two ways to configure virtual networks: bridged and routed, but these techniques increase the possibility to perform some attacks such as sniffing and spoofing virtual network. Data related Vulnerabilities

Data can be co-located with the data of unknown owners (competitors, or intruders) with a weak separation. Data may be located in different jurisdictions which have different laws. Incomplete data deletion – data cannot be completely removed. Data backup can be done by untrusted third –party providers. Information about the location of data usually is unavailable or not disclosed to users. Data is often stored, processed and transferred in clear plain text.[3]

III. THE CONVENTIONAL ANT COLONY OPTIMIZATION ALGORITHM FOR CLOUD DATA RETRIEVAL

The Ant Colony Optimization (ACO) algorithm is a meta-heuristic that has a combination of distributed computation, *autocatalysis* (positive feedback), and constructive greediness to find an optimal solution for combinatorial optimization problems. This algorithm tries to mimic the ant's behavior in the real world. Since its introduction, the ACO algorithm has received much attention and has been incorporated in many optimization problems, namely the

network routing, traveling salesman, quadratic assignment, and resource allocation problems.[5]

The ACO algorithm has been inspired by the experiments run by Goss et al. using a colony of real ants. They observed that real ants were able to select the shortest path between their nest and food resource, in the existence of alternate paths between the two. The search is made possible by an indirect communication known as *stigmergy* amongst the ants. While traveling their way, ants deposit a chemical substance, called *pheromone*, on the ground. When they arrive at a decision point, they make a probabilistic choice, biased by the intensity of pheromone they smell. This behavior has an autocatalytic effect because of the very fact that an ant choosing a path will increase the probability that the corresponding path will be chosen again by other ants in the future.[5] When they return back, the probability of choosing the same path is higher (due to the increase of pheromone). New pheromone will be released on the chosen path, which makes it more attractive for future ants. Shortly, all ants will select the shortest path.

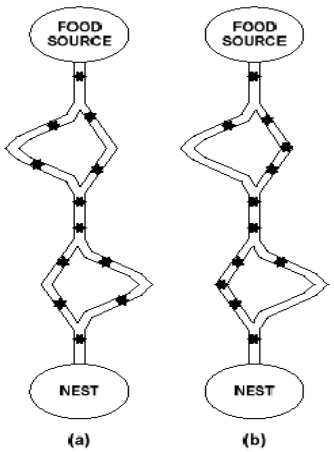

Figure1: **Double bridge experiment. (a) Ants start exploring the double bridge. (b) Eventually most of the ants choose the shortest path in principle capable of building a solution (i.e., of finding a path between nest and food resource), it is only the colony of ants that presents the "shortest path finding" behavior. In a sense, this behavior is an emergent property of the ant colony.**

Figure 1 shows the behavior of ants in a double bridge experiment. In this case, because of the same pheromone laying mechanism, the shortest branch is most often selected. The first ants to arrive at the food source are those that took the two shortest branches. When these ants start their return trip, more pheromone is present on the short branch than the one on the long branch. This will stimulate successive ants to choose the short branch. Although a single ant is

This behavior was formulated as Ant System (AS) by Dorigo et al. Based on the AS algorithm, the Ant Colony Optimization (ACO) algorithm was proposed [41]. In ACO algorithm, the optimization problem is formulated as a graph $G = (C; L)$, where C is the set of components of the problem, and L is the set of possible connections or transitions among the elements of C. The solution is expressed in terms of feasible paths on the graph G, with respect to a set of given constraints. The population of agents (ants) collectively solves the problem under consideration using the graph representation. Though each ant is capable of finding a (probably poor) solution, good quality solutions can emerge as a result of collective interaction amongst ants. Pheromone trails encode a long-term memory about the whole ant search process. Its value depends on the problem representation and the optimization objective.

```
while (termination criterion not satisfied)
        ant generation and activity();
                pheromone
                evaporation();
        daemon actions(); "optional"
        end while
end Algorithm
```

Figure 2: Ant Colony Algorithm.

Informally, the behavior of ants in ACO algorithm can be summarized as follows. A colony of ants concurrently and asynchronously moves through adjacent states of the problem by moving through neighbor nodes of G. They move by applying a stochastic local decision policy which makes use of the information contained in the local node and ant's routing table. By moving, ants incrementally build solutions to the optimization problem. When the solution is being built, every ant evaluates the solution and puts the information about its goodness on the pheromone trails of the connection used. This pheromone information will direct the search of future ants, until a feasible solution is found.[6]

A. The ants in ACO algorithm have the following properties

In order to get more insight on the algorithm, the same ant colony optimization can be implemented in cloud data retrieval, in various steps. Here a cloud agent can be act as ant, which is meant for data retrieval and search results can be stored in the storage area and after completion of retrieval the stored data can be displayed as output from the cloud.

1. Each cloud agent searches for the data which taken for the retrieval
2. Cloud agent has its own memory which enable to store the searched data in it for latter, based on the memory capacity different types of data can be searched out and stored for the future retrieval purpose [7].
3. An cloud agent C can be assigned a start state S^c and retrieving storage buffer can be set as E^c
4. Cloud agent start from a initial state and move over all to the feasible data locations, building the solutions in an incremental way. The procedure stops when at least one queried data has found.
5. The Cloud agent locates a data in a node f ca n move to node g chosen in a feasible neighborhood N^c through probabilistic decision rules. This can be formulated as follows : An cloud agent c in state sr=<Sr-1;f> can move to any node in its feasible neighborhood N^c ;defined as N^{ki} $j \mid (j \in Ni) \Lambda (< sr, j > \in S)\} sr \in S$, with S is a set of all states.

6. A probabilistic rule is function of the following
   a) The data stored in a local node, data structures $A_f = [A_{fg}]$ called ant routing table obtains from. pheromone trails and heuristic values
   b) the ant's own memory from previous iteration, and the problem constraints.
7. When moving from node f to neighbor node g, the agent update the pheromone trails $t_{(fg)}$ on the edge (f,g).
8. Once the data is retrieved from the cloud ,the agent can retrace the same path backward , update pheromone trails and close the operation

By this kind operation a cloud agent retrieves all types of data from cloud storage.[5]

IV. SECURITY IN DATA RETRIEVAL:

Data retrieval can be done through this proposed system securely now the security part can be explained ,here the cloud agent is representing the user who searches for his own data [10] , if it is any agent to attempt to retrieve any other data which is irreverent to his search area the agent with restricted. As ant searches for its food when the food is hidden in any other unwanted matters, similarly the cloud agent will not consider any other area which is irrelevant to search data

V. CONCLUSION AND FUTURE WORK

In this paper the secure data retrieval from cloud database without the loss of data is been elaborated, in future the data retrieval will be biggest task in environment of cloud extended Big Data systems ,this technique also can be implemented in Big Data concepts with some technical changes.

REFERENCES

[1] Cloud Computing Security Issues in Infrastructure as a Service Pankaj Arora* RubalChaudhry Wadhawan Er. Satinder Pal Ahuja M.Tech CSE, IGCE. Asstt.prof (CSE), IGCE Associate Professor & HOD (CSE),IGCE Punjab Technical univ. Punjab technical univ. Punjab technical Univ.

[2] An analysis of security issues for cloud computing Keiko Hashizume1*, David Rosado2, Eduardo Fernández-Medina2 and Eduardo B Fernandez1 Journal of Internet Services and Applications, Springer 27 Feb 2013

[3] As Survey on Data Retrieval Techniques in Cloud Computing JCIT4141PPL, S.Balasubramaniam, Dr. K. kavitha

[4] Infrastructure as a Service Security :Challenges and Solutions Wesam Dawoud #1, Ibrahim Takouna 2, Christoph Meinel #3# Hasso Plattner Institute Potsdam, Germany 1 esam.dawoud@hpi.uni-potsdam.de 3 meinel@hpi.uni-potsdam.de Ministry of Education & Higher Education Palestine 2 itakouna@gmail.com

[5] A Characteristics Study of Ant Colony Optimization Algorithms for Routing Problems Bhanu Pratap Singh1, Sohan Garg2 1Research Scholar Mewar University-Ghaziabad 2 Professor-IIMT Management College-Meerut

[6] Ant Colony Optimization a book by Marco Dorigo and Thomas Stutzle

[7] International Journal of Web &Semantic Technology (IJWesT) Vol.3,No.2,April 2012 Ant Colony Optimization :A solution of Load balancing in Cloud

[8] A. Colorni, M. Dorigo et V. Maniezzo, Distributed Optimization by Ant Colonies, actes de la première conférence européenne sur la vie artificielle, Paris, France, Elsevier Publishing, 134-142, 1991.

[9] D. Martens, M. De Backer, R. Haesen, J. Vanthienen, M. Snoeck, B. Baesens, "Classification with Ant Colony Optimization", IEEE Transactions on Evolutionary Computation, volume 11, number 5, pages 651—665, 2007.

[10] Anthill: A Framework for the Development of Agent-Based Peer-to-Peer Systems Ozalp ¨ Babaoglu Hein Meling Alberto Montresor