A novel Technique: Data Leakage Hindering in Cloud computing using Swarm Intelligence

C. Suresh Kumar
Associate Professor
Dept of Computer Science
Madurai Kamaraj University College
Madurai, TamilNadu, India

K. Iyakutti
Professor
Dept of Physics & Nanotechnology
SRM University,
Chennai, Tamilnadu, India

Abstract

Cloud computing is based on network based computing with the intention of takes place over the Internet. The cloud computing is a culmination of numerous attempts at large scale computing with seamless access to virtually limitless resources. It has many of the assistances such as cost effective, increased storage, shared resources, easy backup and restoration. Although there would be some challenges at some point in of the data diffusion in the cloud computing. The main objective of this context is to spotlight the issue of data security and privacy, probable to be data leakages in the cloud services. This paper initiates a novel scheme, Impeding Data Leakage technique in the cloud computing. This novel technique involved in hampering the leakage of data due to intrusions in the transmission pathway and not to detect or prevent the intrusion, it would lead to halt the entire data leakage with that meticulous path. The objective of this paper is to integrate ant colony optimization technique (ACO) with artificial bee colony technique (ABC), for the intention of hampering the communication path which is interrupted/ the data is leaked. Thus, the ABC crafts the probable and conscientious path to the server and the ACO responses for the requested path and it would back forward the information if the data leakage arises. The leakage of data would be originated with the security restraints by the ACO. Therefore, the ABC and ACO are integrated to make a halt to the data leakage in the cloud service.

Keywords: Cloud Computing, Data leakage, Data leakage hindering, Data leakage Prevention

1. Introduction

Cloud computing is often misunderstood as a complex arbitrary technology but it is greatly simplified if the generic misused cloud term is removed and replaced with service. However, service computing doesn’t have that catchy brand name that is easy to market with fluffy white clouds, so the term never caught on. But that is exactly what Cloud Computing is defined as a set of computer resources providing a service. These services are neatly bundled by companies or cloud service providers and offered for lease or metered by use to the public as a tenancy business model. The cloud computing architecture is designed for elasticity so resources are conserved and used on an as needed basis. This flexibility benefits both the provider and end user allowing both the scale the services as required. Often the service provider can oversubscribe yet provide more than adequate and fair services to the user. To the end user, the Cloud simply appears like an application or device to them performing the needed task to their satisfaction. Yet, below the cloud cover, there are three distinct facets: SaaS, PaaS and IaaS. Thereby eliminating the need for cloud computing users will make the computing resources available on demand. The elimination of an up-front commitment by Cloud users, thereby allowing companies to start small and increase hardware resources only when there is an increase in their needs. The ability to pay for use of computing resources on a short-term basis as needed and release them as needed, thereby rewarding conservation by letting machines and storage go when they are no longer useful.

1.1. Objectives of Cloud Computing

- Protect data from unauthorized access, disclosure, modification, and monitoring.
- Protect information resources from supply chain threats.
Cloud computing has now emerged to become one of the best methods for companies wanting to revamp and enhance their IT infrastructures. However, there are certain issues and problems associated with cloud computing. Needless to say, it is very advantageous for everyone to adapt to new technology, but it is also wise to recognize some of the risks associated with this technology, so as to avoid the possibility of future issues. Here, we bring you information on the risks associated with cloud computing, along with suggestions on how to deal with the same. Generally speaking, most cloud computing service providers are already familiar with the issues involved and can deal with them right at the beginning. This makes the process more of less safe for you. But it also implies that you make wise decisions while choosing your service provider. The users need to clarify all your doubts and issues with your provider before choosing them. Security is one of the most major issues of cloud computing. Being entirely based on the Internet makes it vulnerable to hack attacks. But logically speaking, the entire modern IT world today is invariably connected to the Internet. Hence, the level of vulnerability here is much the same as everywhere else. Of course, the fact that cloud computing is a distributed network also makes it easier for companies to quickly recover from such attacks. Data loss is a very serious problem in Cloud computing. If the vendor closes due to financial or legal problems there will be a loss of data for the users. The users won’t be able to access those data’s because data is no more available for the users as the vendor shut down.

Data leakage is declared as “the unauthorized transmission of information from an organization to an external destination or recipient”. It should be noted that data leakage is not a problem only associated with the cloud. Hackers, spyware and inadvertent data breaches can also lead to data leakage from a company server. Data loss is self explanatory and it can happen if someone losses a briefcase. But, on a business scale, data loss can be disastrous to the business and its clients. Some of the common causes of data loss are Human error, File corruption, Hardware Damage and Site-Damage. Most important of them all is that, there should be a good degree of encryption provided by the vendor to the user that only the user should be able to access the data and not the malicious User. Data leakage prevention like most security efforts is a complex challenge. It is meant to prevent the deliberate and inadvertent release of sensitive information. Too many companies are trying to cure the symptoms rather than prevent them in the first place. Part of the protection equation is being overlooked. Data Leakage Prevention is one part of a comprehensive enterprise data security program that includes comprehensive security best practices for the protection of mission-critical enterprise data repositories. The security must be able to both foil attackers who are financially motivated and won’t be deterred by minimalist security and prevent the accidental release of data. Data security will go nowhere without robust, proactive database security.

This context focuses on the data leakages on the cloud services and not to prevent, but it would hamper the leakage of information. This paper embraces the integrated swarm intelligence techniques to halt the data leakages. Swarm Intelligence indicates a recent computational and behavioral metaphor for solving distributed problems that originally took its inspiration from the biological examples provided by social insects (ants, termites, bees, wasps) and by swarming, flocking, herding behaviors in vertebrates. Any attempt to design algorithms or distributed problem-solving devices inspired by the collective behavior of social insects and other animal societies. The swarm intelligence comes from termite nest building, bee’s societies, bird flocking, fish schooling and ant colony. These are all intriguing, extremely fascinating behaviors that in spite of the specific diversity seem to be related to few in variant properties: Control is fully distributed among a number of individuals, Communications among the individuals happen in a localized way; System-level behaviors appear to transcend the behavioral repertoire of the single individual; The overall response of the system is quite robust and adaptive with respect to changes in the environment. Swarm intelligence designs is allocating computing resources to a number of relatively simple units, no centralized control and interact in a relatively simple and localized way. The swarm lives distributed in some space Communication is a key aspect to get nonlinear behavior, and communication happens locally.

Thus, this perspective technique acquires the behavior of the swarm intelligence to hinder the leakages of the data during transmission in cloud computing. The major objective of this paper is to entirely stop leaking of data by using integrated swarm intelligence (SI) techniques, which includes two SI algorithms Ant Colony Optimization (ACO) and Artificial Bee Colony (ABC). These two algorithms are integrated to impede the leakage and
the entire communication path is rerouted and the transmitted information is back-forwarded to the cloud client. The novel technique is described in the following section.

2. Novel Technique for Hindering the Data Leakage

The real problem is that the risk of data loss does not increase with cloud computing, it increases when employees leverage their own devices, such as smart phones and tablets that happen to leverage cloud for back-end services. This is going to be a constant fit for most in enterprise IT over the next several years. Users won’t give up their personal devices at work, and they don’t want to carry around two phones and two tablets. Thus, those in enterprise IT need to get better at managing these devices, and thus keep data secure. While cloud computing is often the scapegoat of those looking for “data leaks,” data has really been leaking for some time. Cloud or no cloud. USB drives have been around for years, and much of the data that walked out of companies was on those kinds of drives, as well as stolen laptops. So, while it’s cool to kick the cloud in the face, IT has larger problems to address. The risk of data loss is an exception to the availability discussion on the preceding slide. Users may be able to tolerate an occasional service interruption, but non-recoverable data losses can kill a business. Most cloud computing services use distributed and replicated global file systems which are designed to insure that hardware failures will not result in any permanent data loss, but I believe there is still value in doing a traditional off site backup of one’s data, whether that data is in use by traditional servers or cloud computing servers. When looking for solutions, make sure you find ones that backs up data from the cloud.

Cloud computing definitely makes sense if your own security is weak, missing features, or below average. Ultimately, if the cloud provider’s security people are better than cloud users, the web-services interfaces don’t introduce too much new vulnerability, and the cloud provider aims at least as high as you do, at security goals, then cloud computing has better security. However, some problems cannot be tackled with traditional hardware and software. There are some drawbacks in traditional techniques in computing tasks have to be, well-defined, fairly predictable and computable in reasonable time. Thus, the alternate computing would be introduced such as DNA based computing, Bio-computing, neural networks and Quantum computing. These kind of computing are involved in the field of process optimization, Robotics and telecommunications. Obviously, these computing would have some complexity like:

- Compromises between different goals
- Some natural mechanisms are not well understood
- Well-defined problems cannot be solved in better way
- Computing Sometimes fails

Thus the implementation of swarm intelligence in the cloud computing is undertaken in this paper. The swarm intelligence signifies the biological computing by the behavior of social insects. This paper introduces a novel technique, Impeding Data Leakage Technique (IDL) which is worn to hamper the data leakages in the cloud services using the swarm intelligence, not to prevent it but it would halt the outflow of information to make a secure service. IDL technique includes an integrated swarm intelligence technique known to be Ant Colony Optimization (ACO) and Artificial Bee Colony (ABC). The ACO and ABC techniques are integrated to construct an impede scheme of data leakage in cloud computing. The entire server behaviors would be knobbed by the ACO and the client behaviors would be handling with the ABC.

3. Impeding Data Leakage Technique (IDL)

The technique Impeding Data Leakage (IDL) is extended to hinder the outflow of the information during communication within the cloud services. The data might be leaked due to the intrusion occurs within the environment. A Secure cloud is always a reliable source of information thus protecting the cloud is a very important task for security professionals who are in charge of the cloud. Some of the ways by which a cloud can be protected are Protection of data, making sure data is available for the users, delivering high performance for the users, using Intrusion Detection System on Cloud to monitor any malicious activities, to make sure the application used by the user is safe to use, the cloud server...
must provide a support system for the user, user should be able to recover any loss of data in the cloud. The intruder would assault the information during the communication path from the server to the cloud users. The loss of control would leads to the data outflow in the communication channels.

The cloud service is more secure than the other computing services, but in this case it may cause some data loss during transmission. Thus, this context constructs a technique to completely hinder the data loss, known to be IDL. The term “Impede” declares that completely halts the data loss and prevents the communication channels. The major objective of this paper is to hinder the leakage by redirecting the paths based on the swarm intelligence techniques (ABC & ACO). The client requests are exceeds to the cloud server with the possible and responsible routes within the cloud environment. In this state, the requests are analyzed and the communication route is constructed by using the technique of swarm intelligence known to be Artificial Bee Colony (ABC). The transmission route is generated by the ABC only if the possible routes to the cloud server are responsible in affirm and also the responsible routes are evaluated that there might be any occurrence of data leakage, if outflow occurs then next responsible route is chosen for the request transmission. After transmitting requests, the cloud server responses to the client requests using Ant Colony Optimization (ACO). The server responses are transferred through optimized path to the client, as same as ABC the optimized path restrains any leakage of information and if leakage occurs the ACO chooses the next optimized path to transfer the information. ACO and ABC swarm intelligence techniques is based on the idea of stigmergy. stigmergy is a mechanism of communication by modifying the environment. Ant Colony Optimization and Artificial Bee Colony use artificial stigmergy. This kind technique used for solving problems which can be expressed as finding good paths through graphs.

ACO algorithm acquires the idea of the behavior of each ant in nature. Each ant tries to find a route between its nest and a food source. The natural behavior of each ant is:

- Wander randomly at first, laying down a pheromone trail
- If food is found, return to the nest laying down a pheromone trail
- If pheromone is found, with some increased probability follow the pheromone trail
- Once back at the nest, go out again in search of food

However, pheromones evaporate over time, such that unless they are reinforced by more ants, the pheromones will disappear. Ants are practically blind but they still manage to find their way to and from food. These observations inspired a new type of algorithm called ant algorithms. Ant systems are a population based approach. In this respect it is similar to genetic algorithms. There is a population of ants, with each ant finding a solution and then communicating with the other ants. Thus this ant algorithm is implemented in this paper to finding a solution for the data leakage within the cloud services. In this context, the ants are assumed to be transferring contrivance or routing device, which would, wanders for the shortest and lossless path within the services and then transfer the information. Ant systems are a population based approach. In this respect it is similar to genetic algorithms. There is a population of ants, with each ant finding a solution and then communicating with the other ants. This technique is interested in exploring the search space, rather than simply plotting a route. The communication system needs to allow the ants to explore paths and follow the best paths with some probability in proportion to the intensity of the pheromone trail. The system does not want them simply to follow the route with the highest amount of pheromone on it, else our search will quickly settle on a probably sub-optimal solution. The probability of an ant following a certain route is a function, not only of the pheromone intensity but also a function of what the ant can see (visibility). The pheromone trail must not build unbounded. Therefore, they need “evaporation”.

Figure 3: Impeding Data Leakage System
The figure illustrated that the ACO involved in the routing of cloud based service. The shortest and loss free path is evaluated using these ants. An ant would wander to the source from the destination to search for the undeviating route with leakage free for the transmission. Then the possible route is determined and the information is transferred through this path by the server.

Likewise, the ABC algorithm integrated with this system to acquire the requesting routes to the server. ABC is developed based on inspecting the behaviors of real bees on finding nectar and sharing the information of food sources to the bees in the hive. Agents in ABC: The Employed Bee, the Onlooker Bee and the Scout.

- **The Employed Bee:** It stays on a food source and provides the neighborhood of the source in its memory.
- **The Onlooker Bee:** It gets the information of food sources from the employed bees in the hive and select one of the food source to gathers the nectar.
- **The Scout:** It is responsible for finding new food, the new nectar, and sources.

A colony of bees can extend itself over long distance in multiple directions. Scout bees search for food randomly from one patch to another. The exchange of information among the bees is the most important occurrence in the formation of the collective knowledge. Communication among the bees related to the quality of the food sources occurs in the dancing area. Thus the bees’ communication in the dancing area which includes the information about:

- Distance from the source to destination
- Quality of the information
- Energy consumption

In this paper, the ABC algorithm is involved in the client area of the cloud service to hinder the data leakage within the transmission path. Thus the artificial bees randomly search for destination routes from source which doesn’t hold any leakage of data. Then it would compute the quality of the information and the energy required to reach the destination through the possible path. In this scheme,

- The **employee bee** acquires the probable responding route to transfer the information to the server.
- The **scout bee** involved in acquiring the next possible responding route for the communication.
- The **onlooker bee** contains the situation of seeking for the response from the server.

Finally, the bees found the loss free multiple routes for the communication. The information gathered from the source is transmitted through these routes. Thus the data leakage is hampered using this
technique. The possible and responding route is chosen and then the search for loss less route to transfer the data. By using this technique, the data leakage is hindered entirely throughout the communication path.

4. Conclusion

This paper concluded that the data leakage would be hindered during the communication within the cloud service. The cloud service has many of the assistances for the web users to store their information within the cloud server. But in some cases there would occur leakage of data due to the unauthorized user or intrusions within the communication path. Thus a novel technique would involve in this context to hinder the leakage entirely within the particular chosen route known to be IDL (Impeding Data Leakage) Technique. This technique acquires swarm intelligence based algorithm ABC and ACO to hamper the data loss. These two algorithms are integrated within the communication path of the cloud service to halt the data loss. Thus the performance provides a better and quality outcome based on this technique.

5. References


[5] Bijayalaxmi Purohit, Pawan Prakash Singh, "Data leakage analysis on cloud computing ", IJERA


