International Journal of Advance Research, IJOAR .org
Volume 4, Issue 4, April 2016, Online: ISSN 2320-9194

REVIEW OF MINIMIZING DATA THEFT ATTACK IN CLOUD THROUGH FOG COMPUTING

Nikhil D. Dixit[1], Prof. S.T. Khandare[2]

[1]Pursuing Master degree program in Computer Science & Engineering, in Sant Gadge Baba Amravati University, Amravati, India (M.S.) Babasaheb Naik College of Engineering, Pusad, Mob. 7350914406, Email : nikhil.dixit001@gmail.com
[1] Associate Professor, Babasaheb Naik College of Engineering, Pusad, India (M.S.) Email : khandare.shailesh@rediffmail.com

ABSTRACT

Cloud computing is a delivery platform which promises a new way of accessing and storing personal as well as business information .It provides resources to its users through the Internet. Existing data protection mechanisms such as encryption have failed in preventing data theft attacks, especially those perpetrated by an insider to the cloud provider. It also has a risk that is the involvement of a third party which makes it difficult to trust that user data is secure enough and will not be misused. To provide security, new technology called Fog computing has arrived through which user data can be secured. In this paper, we discussed this paradigm in detail and review the work that has been done using this technology.

I. INTRODUCTION

Cloud computing has become a reality which paved the way for new model of computing. The users of cloud can outsource their data and also computations. With this facility made available without capital investment, the small and medium organizations opt for out shouting their data and computations to cloud. This gives the organizations many benefits besides operational efficiency. However, it brings security risks that are to be considered carefully. Very important concern is the insider data theft. When cloud service provider is unable to prevent insider data theft attacks, it is very important security concern. When malicious insiders who can steal data illegally throw challenges, the cloud service provider may not be able to prevent them. The Cloud Security Alliance has considered this as a top threat .Many users or customers of cloud computing are aware of this kind of threat. However, they have left their concern to cloud service provider believing that the cloud service provider takes care of it. There is lack of transparency, problem in data dynamics and security related problem like authorization, authentication, and audit controls etc.

Few years back there was a Twitter incident which is a best example for data theft attack from the cloud service provider. This incident exposed the security problems in cloud computing as it could make the customers of Twitter to lose their sensitive data and documents. The documents were ex-filtrated by TechCrunch. The President of United States Barak Obama was also a victim of the data theft attack. His files also were accessed illegally .The insider attacker had stolen Twitter's admin password and gained access to Twitters corporate documents. The incident caused significant damage to Twitter and its customer across the globe. This attack was reportedly made by an outsider. However, there is possibility to have internal attacks for data theft. In their work Rocha and Coriaexplored how to steal easy passwords through malicious insider of cloud service provider (CSP). They also demonstrated how to steal private keys and the confidential data which is saved in hard disk. Once credentials are stolen, an insider can gain access to customers' data illegally. There has been much research went on cloud computing security. Especially lot of research went on cloud computing and its storage problems. Much research is on preventing unauthorized access. However, the techniques fail when it comes to an insider data theft attack. Fully holomorphic encryption solution was proposed by Van Disk and Jules as a solution to such threats to protect data.

Fog computing provides- Low latency and location awareness, it has Wide-spread geographical distribution, supports Mobility, is compromised due to the huge number of nodes. The main task of fog is to deliver data and place it closer to the user who is positioned at a location which at the edge of the network. Here the term edge refers to different nodes to which the end user is connected and it is also called edge computing. If we look according to architecture fog is situated below the cloud at the ground level. The term fog computing is given by CISCO as a new technology in which mobile devices interact with one another and support the data communication within the Internet of Things.

But here we consider Fog Computing as a paradigm through which we can provide local access to the user and with the help of decoy technology, we provide security for user data and prevent insider theft attacks.

II. SECURING CLOUDS WITH FOG

Numerous proposals for cloud-based services describe methods to store documents, files, and media in a remote service that may be accessed wherever a user may connect to the Internet.

A particularly vexing problem before such services are broadly accepted concerns guarantees for securing a user's data in a manner where that guarantees only the user and no one else can gain access to that data. The problem of providing security of confidential information remains a core.

Security problem that, to date has not provided the levels of assurance most people desire. Many proposals have been made to secure remote data in the Cloud using encryption and standard access controls. It is fair to say all of the standard approaches have been demonstrated to fail from time to time for a variety of reasons, including insider attacks, mis-configured services, faulty implementations, buggy code, and the creative construction of effective and sophisticated attacks not envisioned by the implementers of security procedures. Building a trustworthy cloud computing environment is not enough, because accidents continue to happen, and when they do, and information gets lost, there is no way to get it back. One needs to prepare for such accidents.

The basic idea is that we can limit the damage of stolen data if we decrease the value of that stolen information to the attacker. We can achieve this through a 'preventive' disinformation attack. We posit that secure Cloud services can be implemented given two additional security features:

1) User Behavior Profiling: It is expected that access to a user's information in the Cloud will exhibit a normal means of access. User profiling is a well known technique that can be applied here to model how, when, and how much a user accesses their information in the Cloud. Such 'normal user' behavior can be continuously checked to determine whether abnormal access to a user's information is occurring. This method of behavior-based security is commonly used in fraud detection applications. Such profiles would naturally include volumetric information, how many documents are typically read and how often. These simple userspecific features can serve to detect abnormal Cloud access based partially upon the scale and scope of data Transferred.

2) Decoys: Decoy information, such as decoy documents, honey files, honey pots, and various other bogus information can be generated on demand and serve as a means of detecting unauthorized access to information and

to 'poison' the thief's ex-filtrated information. Serving decoys will confound and confuse an adversary into believing they have ex-filtrated useful information, when they have not. This technology may be integrated with user behavior profiling technology to secure a user's information in the Cloud. Whenever abnormal access to a cloud service is noticed, decoy information may be returned by the Cloud and delivered in such a way as to appear completely legitimate and normal. The true user, who is the owner of the information, would readily identify when decoy information is being returned by the Cloud, and hence could alter the Cloud's responses through a variety of means, such as challenge questions, to inform the Cloud security system that it has inaccurately detected an unauthorized access. In the case where the access is correctly identified as an unauthorized access, the Cloud security system would deliver unbounded amounts of bogus information to the adversary, thus securing the user's true data from unauthorized disclosure. The decoys, then, serve two purposes: (1) Validating whether data access is authorized
When abnormal information access is detected.
(2) Confusing the attacker with bogus information.
We posit that the combination of these two security features will provide unprecedented levels of security for the Cloud. No current Cloud security mechanism is available that provides.

This level of security. We have applied these concepts to detect illegitimate data.

Access to data stored on a local file system by masqueraders, *i.e.* attackers who impersonate legitimate users after stealing their credentials. One may consider illegitimate access to Cloud data by a rogue insider as the malicious act of a masquerader. Our experimental results in a local file system setting show that combining both techniques can yield better detection results, and our results suggest that this approach may work in a Cloud environment, as the Cloud is intended to be as transparent to the user as a local file system. In the following we review briefly some of the experimental results achieved by using this approach to detect masquerade activity in a local file setting.

A. Combining User Behavior Profiling and Decoy Technology for Masquerade Detection

1) User Behavior Profiling*:* Legitimate users of a computer system are familiar with the files on that system and where they are located. Any search for specific files is likely to be targeted and limited. A masquerader, however, who gets access to the victim's system illegitimately, is unlikely to be familiar with the structure and contents of the file system. Their search is likely to be widespread and untargeted. Based on this key assumption, we profiled user search behavior and developed user models trained with a one class modeling technique, namely one-class support vector machines. The importance of using one-class modeling stems from the ability of building a classifier without having to share data from different users. The privacy of the user and their data is therefore preserved.

We monitor for abnormal search behaviors that exhibit deviations from the user baseline. According to our assumption, such deviations signal a potential masquerade attack. Our previous experiments validated our assumption and demonstrated that we could reliably detect all simulated masquerade attacks using this approach with a very low false positive rate of 1.12%.

2) Decoy Technology*:* We placed traps within the file system. The traps are decoy files downloaded from a Fog computing site, an automated service that offers several types of decoy documents such as tax return forms, medical records, credit card statements, e-bay receipts, etc. The decoy
files are downloaded by the legitimate user and placed in highly-conspicuous locations that are not likely to cause any interference with the normal user activities on the system. A masquerader, who is not familiar with the file system and its contents, is likely to access these decoy files, if he or she is in
search for sensitive information, such as the bait information.

III. ARCHITECTURE OF FOG COMPUTING

Fog computing is well suited for the geographical distribution of resources instead of having a centralized one, meaning Fog computing is the extension of Cloud computing. The difference is Fog provides proximity to its end users through dense geographical distribution and it also supports mobility. Access points or set-up boxes are used as end devices to host services at the network. In Fog computing platform multi-tier architecture is used. In first tier there is machine to machine communication and the higher tiers deals with visualization and reporting. The higher tier is represented by the cloud. The architecture is as shown in the fig shown below.

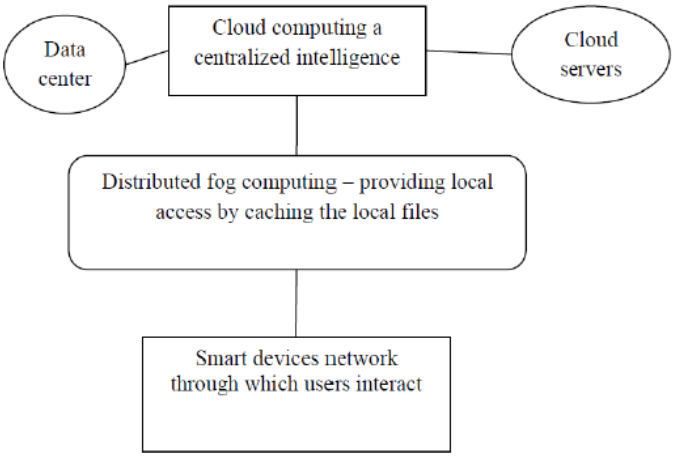

Fig. 1 : Architecture of Fog computing

IV. PROBLEMS IN FOG COMPUTING

There are many open problems that will have to be addressed to make the fog a reality. It is necessary to clearly identify these problems so future research works can focus on them. The set of open challenges for the fog to become a reality is:

1) Discovery/Sync: Applications running on devices may need either some agreed „centralized" point (e.g. establish an "upstream" backup if there are too few peers in our storage application);

2) Compute/Storage limitation**:** Current trends are improving this fact with smaller, more energy-efficient and more powerful devices (e.g. one of today's phones is more powerful than many high end desktops from 15 years ago). Still new improvements are granted for non consumer devices.

3) Management : In addition to setting up the communication routes across end nodes, IoT/ubiquitous computing nodes and applications running on top need to be properly setup and configured to operate as desired. Having potentially billions of small devices to be configured, the fog will heavily rely on decentralized (scalable) management mechanisms. That are yet to be tested at this unprecedented scale. One thing that can be predicted with certain degree of confidence is that there will be no full control of the whole fog and asymptotic declarative configuration techniques will become more common.

4) Security: The same security concerns that apply to current virtualized environments can be foreseen to affect fog devices hosting applications. The presence of secure sandboxes for the execution of droplet applications poses new interesting challenges: Trust and Privacy. Before using other devices or mini-clouds in the network to run some software, isolation and sandboxing mechanisms must be in place to ensure bidirectional trust among cooperating parties. The fog will allow applications to process user's data in third-party hardware/software. This of course introduces strong concerns about data privacy and its visibility to those third parties.

5) Standardization: Today no standardized mechanisms are available so each member of the network (terminal, edge point...) can announce its availability to host others" software components, and for others to sent it their software to be run;

6) Accountability/Monetization: Enabling users to share they spare resources to host applications is crucial to enable new business models around the concept of the fog. A proper system of incentives needs to be created. The incentives can be financial or otherwise (e.g. unlimited free data rates). On the other hand the lack of central controlling entity in the fog makes it difficult to assert if a given device is indeed hosting a component (droplet) or not;

7) Programmability: Controlling application lifecycle is already a challenge in cloud environments. The presence of small functional units (droplets) in more locations (devices) calls for the right abstractions to be in place, so that programmers do not need to deal with these difficult issues. Easy to use APIs for programmers will heavily rely on simple Management mechanisms that provide them with the right abstractions to hide the massive complexity of the fog. Some vendors like Microsoft have already taken some steps in positioning themselves in this space9.

V. CONCLUSION AND FUTURE WORK

In this paper, we present a approach to secure personal and business data in the Cloud. We have seen monitoring data access patterns by profiling user behavior to determine if and when a malicious insider illegitimately accesses someone's documents in a Cloud service. Decoy documents stored in the Cloud alongside the user's real data also serve as sensors to detect illegitimate access. Once unauthorized data access or exposure is suspected, and later verified, with challenge questions for instance, we inundate the malicious insider with bogus information in order to dilute the user's real data. Such preventive attacks that rely on disinformation technology could provide unprecedented levels of security in the Cloud and in social networks.

REFERENCES

[1] Hashizume K., Rosado D. G., Fernandez- Medina E. and Fernandez E. B. "An analysis of security issues for cloud computing". Journal of Internet Services and Applications, 2013, 4(1), pp.1-13.

[2] Luis M. Vaquero and Luis Rodero-Merino "Finding your Way in the Fog: Towards a Comprehensive Definition of Fog Computing" ACM SIGCOMM Computer Communication Review - Volume 44, Number 5, October 2014.

[3] P. Allen, "Obama's Twitter password revealed after french hacker arrested for breaking into U.S. president's account," March 2010. [Online]. Available: http://www.dailymail.co.uk/news/article-1260488/Barack- Obamas-Twitter-password-revealed-French-hacker-arrested.html

[4] D. Takahashi, "French hacker who leaked Twitter documents to TechCrunch is busted," March 2010. [Online]. Available: http://venturebeat.com/2010/03/24/french-hacker-wholeakedtwitter- documents-to-techcrunch-is-busted/

[5] Park, Y., & Stolfo, S. J. "Software decoys for insider threat", In Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security, 2012, May, pp. 93-94.

[6] The Fog Computing Paradigm: Scenarios and Security Issue Ivan Stojmenovic SIT, Deakin University, Burwood, Australia 2013.

[7] FOG Computing Mario Nemirovsky – ICREA/BSC 2014

[8] Michael Enescu - From Cloud to Fog Computing and IoT | LinuxCon + CloudOpen North America 2014 .

[9] http://www.webopedia.com/TERM/F/fog-computing.html