ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

Insider Data Theft Mitigation Techniques

[I]Shivani Singla [II]Tejbir Rana
[I]M.Tech,CSE, [II]Assistant Professor(CSE)
[I,II]Shivalik Institute of Engineering & Technology

Abstract
Cloud computing is gradually changing the way we use computers. Unlike previous methods of storing data on hard drives on computers, more and more data is being stored on virtual drives in cloud. Be it in the form of social networking, cloud storage or any other online services we are becoming more and more dependent on cloud services. The cloud however has some issues security being the most crucial one. Data theft from cloud may become hazardous to personal or financial life of a cloud user. Amongst the security issues, insider data theft is one of the most crucial as an insider knows more about the cloud and it loopholes than anyone from outside and once data is stolen, it is almost impossible to track the culprit. In this review paper, we discuss about the risk of insider data theft and countermeasures to check data theft.

Index Terms
Fog computing, insider data theft, data leakage, anomalous behavior pattern detection, encryption, user behavior profiling, decoy

I. Introduction

Cloud computing can be defined as a set of IT services provided over a network to a customer on a leased basis with the facility to scale up or down service requirements of the customer. A third party provider usually delivers Cloud Service, and it is usually also owns the infrastructure. Advantage of Cloud computing are: pay-what-you-use feature, fast and easy deployment, budget-friendliness, scalability, resilience, flexibility and efficiency. Regardless of the prospective gain achieved from cloud computing, the organizations are dawdling to recognize it due to security issues and challenges associated with it. For years the Internet has been represented on network diagrams by a cloud symbol until 2008 when a variety of new services started to emerge that allowed access to the computing resources over the Internet: Cloud computing. Cloud computing encompasses activities ranging from the use of social networking sites and other forms of interpersonal computing to access online software applications, data storage and processing power. Since its inception, cloud computing has added capabilities dynamically without any investment in new infrastructure, providing training to new personnel, or licensing any new software. But as more and more information on individuals and companies are placed in the cloud, concerns are beginning to grow about just how safe an environment it is. Figure 1 shows security as the number one on the list of challenges/issues in the IT cloud services as per the survey conducted by IDC Enterprise Panel.

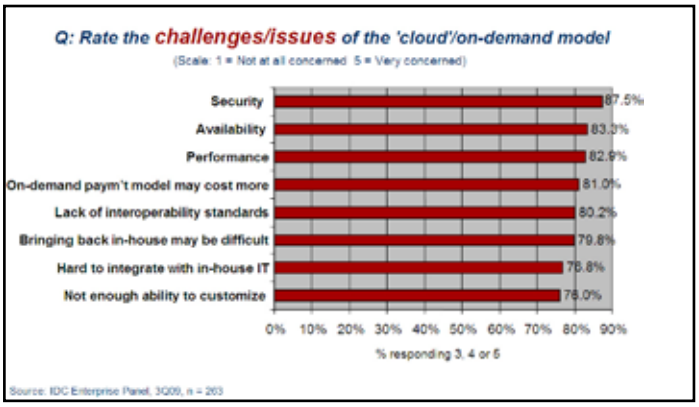

Fig. 1: Result of IDC survey ranking security challenges, 2009 [1]

A thorough understanding of security threats: their types and perspective, is required to counter the attacks and to win the trust of the customers to use this promising environment. In this paper section II discusses threat due to malicious insider in detail. Section III primarily discusses about security measures for mitigating insider data theft attacks.

II. Data Breach By Malicious Insider

Breach of security occurs from outside of the organizations as well as from within. According to CyberSecurity Watch Survey conducted in 2011 on 607 professionals, businesses, consultants and government executives insiders are responsible for 21% of the total cyber-attacks. 33% of the respondents contemplated that the attacks by the insider were more costly and damaging to organizations [2]. The most common inside attacks are unauthorized access to and use of corporate information (63%), unintended disclosure of private or sensitive data (57%), virus, worms, or other malicious codes (37%), and theft of intellectual property (32%). The cloud computing vulnerabilities to malevolent insider are: imprecise roles and responsibilities, poor enforcement of role definitions, non applicability of need-to-know principle, AAA vulnerabilities, system or OS vulnerabilities, insufficient physical security procedures, impracticality of processing data in encrypted form, application vulnerabilities or poor patch management [3]. Malicious disruption of an organization's sensitive information resources could lay the entire victim organization's operation on the line. There are three types of cloud-related insider threats: the rogue administrator, insiders who exploit cloud vulnerabilities, and the insiders who use the cloud to conduct disreputable activity [4]. Rogue administrator has dispensation to steal unprotected files, brute-force attack over passwords, and download customers' data from the victim organization. Insiders who exploit cloud vulnerabilities try to gain unauthorized access to confidential data in an organization; they could make a fortune by selling the sensitive information, or use the information for their future businesses. Insiders who use the cloud to conduct nefarious activity carry out attacks against its own employer's IT infrastructure. Since the insiders are familiar with the IT operations of their own companies, the attacks are generally difficult to be traced using forensic analysis.

III. Data Protection Countermeasures

Data breaches caused by insiders could be either unintentional or deliberate. It is advisable to apply proper security tools to deal with insider threats as it is complicated to make out the insiders' behavior. The tools include: data loss prevention systems, anomalous behavior pattern detection tools, format preserving and encryption tools, user behavior profiling, decoy technology, and authentication and authorization technologies such as multifactor and 4-eyes solutions [5-7]. These tools provide functions such as real-time detection on monitoring traffic, audit trails recording for future forensics, and trapping malicious activity into decoy documents.

A. Data Leakage Prevention

Preventing accidental or malicious loss of data by insiders or outsiders is the main purpose of DLP solutions. With appropriate implementation of the mechanism organizations can control the access of their sensitive data.

A comprehensive DLP solution that protects data in motion, data at rest and data in user require complex and significant amount of preparation activities. Among these activities, data classification, risk assessment and policy development are the most critical ones and involve both the commitment from senior management and IT security personnel [8].

A comprehensive DLP solution is usually a combination of Network DLP, Endpoint DLP, Embedded DLP components and employee training program. The following addressed several techniques / processes to mitigate the data leakage threats: Secure Content Management, Embedded DLP in Applications, Thin Client Restriction on Removable Media, Application Proxy Firewalls, Secured Data Transmission via Internet Training and Awareness

B. Anomalous Behavior Pattern Detection Tools

Anomaly detection is the identification of items, events or observations which do not conform to an expected pattern or other items in a data set [9]. Typically the anomalous items will translate to some kind of problem such as bank fraud, a structural defect, medical problems or finding errors in text. Anomalies are also referred to as outliers, novelties, noise, deviations and exceptions. Three broad categories of anomaly detection techniques exist. Unsupervised anomaly detection techniques detect anomalies in an unlabeled test data set under the assumption that the majority of the instances in the data set are normal by looking for instances that seem to fit least to the remainder of the data set. Supervised anomaly detection techniques require a data set that has been labeled as "normal" and "abnormal" and involves training a classifier (the key difference to many other statistical classification problems is the inherent unbalanced nature of outlier detection). Semi-supervised anomaly detection techniques works by constructing a model which represents normal behavior using a normal training data set which is given, and then the likelihood of a test instance is tested which is generated by the learnt model.

Several anomaly detection techniques have been proposed in literature. Some of the popular techniques are: distance based techniques (k-nearest neighbor, local outlier factor), one class support vector machines, replicator neural networks, cluster analysis based outlier detection, pointing at records that deviate from learned association rules.

C. Encryption

Computer encryption is based on cryptography, which has been used as long as humans have wanted to keep information secret. Algorithms or ciphers provide a way in which to craft a message and give a certain range of possible combinations. A key, on the other hand, helps a person or computer figure out the one possibility on a given occasion [10].Computer encryption systems generally belong in one of two categories:Symmetric-key encryption: If any computer wants to send a message over the network it will first encrypt the message using the security key (a separate code which is available with each individual computer). There is a requirement in Symmetric-key encryption that you know which computer you are going to communicate with so that a key can be installed on each one. In order to decode the information each computer must know the secret code. This code is used for finding the key which would then be used to decode the information.

Public-key encryption: In symmetric key encryption is that two users attempting to communicate with each other need a secure way to do so; otherwise, an attacker can easily pluck the necessary data from the stream. Also known as asymmetric-key encryption, public-key encryption uses two different keys at once - a combination of a private key and a public key. The private key is not communicated to any other computer and is known to your computer only, whereas the public key is given to those computers which want to securely communicate with our computer. The public key, provided by the originating computer, and its own priva te key are used for decoding an encrypted message sent by another computer.

D. User Behavior Profiling

It is expected that access to a user's information in the Cloud will exhibit a normal means of access. A model is prepared as to how, when and how much a user accesses their information in the Cloud. This 'normal means of access' can be continuously checked to determine whether abnormal access to a user's information is occurring. Search for specific files is likely to be targeted and limited the reason being that valid users of a computer system are familiar with the files on that system and where they are located. Search by a masquerade is liable to be extensive and untargeted because of his unfamiliarity with the structure and contents of the file system. Based on this key assumption, user search behavior is profiled and user models are developed trained with a one class modeling technique, namely one-class support vector machines. In a one-class modeling technique a classifier can be built without having to share data from different users. The data and privacy of the user is thus preserved. Abnormal search behaviors that exhibit deviation from the user baseline are monitored. A potential subterfuge attack is signaled by such detection [11].

E. Fog Computing Decoy Technology

Decoy documents, honeyfiles, honeypots, 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. The adversary will be befuddled and confused by serving such decoys into falsely believing they have ex-filtrated useful information. On integrating this technology with user behaviour profiling technology a user's information in the Cloud can be secured. Decoy information may be returned by the Cloud and delivered in such a way as to appear completely legitimate and normal on noticing abnormal access to a cloud service. On the contrary when decoy information is being

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

returned by the Cloud to the genuine user, being the owner of the information, he would readily identify the decoy information and 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. The Cloud security system would deliver unbounded amounts of bogus information to the adversary in case the access is correctly identified as an unauthorized access, thus securing the user's true data from unauthorized exposé [12]. Two purposes are then served by the decoy: (1) validating data access authorization on detection of abnormal information access, and (2) perplexing the attacker with spurious information.

Traps are placed within the file system. The advantages of placing decoys in a file system are threefold: (1) the detection of masquerade activity (2) the confusion of the attacker and the additional costs incurred to distinguish real from bogus information, and (3) the deterrence effec which, although hard to measure, plays a significant role in preventing masquerade activity by risk-averse attackers.

IV. Conclusion and Future Work

Cloud computing is in continual development in order to make different levels of on-demand services available to customers. While people enjoy benefits cloud computing brings, security in clouds is a key challenge. Much vulnerability in clouds still exists and hackers continue to exploit these security holes. Flaws that may endanger the security of cloud users must be identified if we are to provide better quality of service. In this paper, we examined the security vulnerabilities in clouds from the perspective of insider theft, included related real world exploits, and introduced countermeasures to those security breaches. In the future, we will continue to contribute to the efforts in studying cloud security risks and the countermeasures to cloud security breaches.

References

[1]. F. Gens. , New IDC IT Cloud Services Survey: Top Benefits and Challenges, IDC eXchange, 2010.

[2]. 2011 CyberSecurity Watch Survey, CERT Coordination Center at Carnegie Mellon University, 2011.

[3]. D. Catteddu, G. Hogben, Cloud Computing Benefits, Risks and Recommendations for Information Security, The European Network and Information Security Agency (ENISA), 2009.

[4]. Insider Threats Related to Cloud Computing, CERT, July 2012.

[5]. Steve Katz, Tackling the Insider Threat, February 2009

[6]. White Paper, Cloud Security Risks and Solutions, BalaBit IT Security, July 2010.

[7]. S. J. Stolfo, M. B. Salem, A. D. Keromytis, Fog computing: Mitigating Insider Data Theft Attacks in the Cloud, IEEE Symposium on Security and Privacy Workshops, 2012 : 125-128.

[8]. ITO, Data Leakage Prevention Issue 5.

[9]. Chandola, V., Banerjee, A., Kumar V. , Anamoly Detection: A Survey, 2009

[10]. Tyson, Jeff, How Encryption Works? , 2014.

[11]. M. Ben-Salem, S. J. Stolfo, Modeling user search-behavior for masquerade detection, Proceedings of the 14th International Symposium on Recent Advances in Intrusion Detection. Heidelberg: Springer, September 2011: 1–20.