International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 3 | Mar -2017 www.irjet.net p-ISSN: 2395-0072

STRATEGIES FOR DATA LEAKAGE PREVENTION

Anirudh Garg, Ashish Kumar, Pradeep Kumar H. S.

Department of Information Science and Engineering, National Institute of Engineering, Mysore, India
Assistant Professor, Department of Information Science and Engineering, National Institute of Engineering, Mysore, India

---

Abstract - *In today's world knowledge is power so data leakage has become a major concern for any organization or an individual who doesn't want their information or data to be leaked to outsiders who they did not want to divulge the information, but they cannot be sure that the employees working under them will have the best interest of the organization and won't put their personal agendas in front so to we want to prevent any confidential information to not leave the premises of the organization to ensure that we protect the data with the help of a firewall which will scan all the outgoing messages sent by the employee if the transfer is not authorized by the administrator the employee will be reported and there will be repercussions for his/her actions this is a very important procedure in present day organizations because by the time we detect that data has been leaked it may be too late to take any action so we need to be one step ahead of the problem. It involves a firewall ,keyword-recognition and identifying the guilty agent.*

1.INTRODUCTION

In this paper, we develop a model for protecting the information that we want to protect and we are going to see how the data can be leaked and what measures our data leakage prevention implementations reduce these threats, we are also going to make a software which will help us prevent data leakage. As we have seen data leakage is a major concern because in present information is power, so we have designed a software in which the admin or the administrator (distributor) can send the message like an e-mail, message via some other service to an agent or multiple agents now if that agent wants to send that data to some unauthorized person or we can say if he wants to leak the data to someone outside the company which may or may not cause harm to the company itself, as he is not authorized to send this data to anyone, the firewall that we have installed in the pc of the employees will scan the whole document and will determine whether the file or information that the employee is trying to share is in his jurisdiction or if he is trying to send a data with confidential information, this will be determined by scanning the document and determining how many keywords have matched with the document and if the firewall gets a red-flag(if it gets a common keyword) the employee's ID will be sent to the administrator and then it's up to the administrator to decide if he wants to take action and if he does how severe the consequences.

2. LITERATURE SURVEY

Now a day's most of the businesses be it minute or multinational all are making use of the network in some form or the other this can be seen by the booming use of the cloud to store the data of the companies in the cloud storage which are more flexible i.e. they can be accessed remotely and can add storage value to it, then to be stored in physical drives, cloud is being used by most of the notable companies , marketplace and academic world. number of notable commercial and individual cloud computing services, e.g., form Amazon, Google, Microsoft, Yahoo, and Sales force. Also, top database vendors, like Oracle, as they are adding cloud support to their databases.

The providers are enjoying the opportunity to be able to build a market to sell cloud space or storage to the people who are willing to buy for a relatively low-cost as compared to physical storages, these low-cost and pay for use cloud can only work if the providers are able to provide the required amount of protection to the data. quick provisioning, quick flexibility, everywhere network contact, hypervisor defense against network vulnerability, economical failure recovery and data storage solution, on-request security checks, synchronized detection of system altering and rapid re-construction of services.

The cloud provides this compensation, until some of the risks are better understood. The basic concept of the cloud , based on the services they offer, form application service provisioning, grid and service computing, to Software as a Service. Despite of the specific architecture, the dominant concept of this computing model is that customers' data, which can be of individuals, organizations or enterprises, is processed remotely in unknown machines about which the user not aware. The ease and efficiency of this approach, however, comes with privacy and security risks. Confidentiality of data is the main hurdle in implementation of cloud services[1].

3.EXISTING SYSTEMS

Data Leakage Prevention (DLP) is a very integral part of our society because we need to not only punish who are trying to access restricted information but we have to take every precaution that it does not happen in the first place, so that why we have some of the data leakage prevention strategies which are in place in our present society.

1. IDS/IPS

IDS - IDS is an acronym for intrusion detection system, it is a device or software application that monitors networks and system activities for malicious activities such as Any detected activity or violation is typically reported either to an administrator or collected centrally using a security information and event management (SIEM) system. A SIEM system combines outputs from multiple sources, and uses alarm filtering techniques to distinguish malicious activity from false alarms[2].

IPS - IPS is an acronym for intrusion prevention system, it monitors networks, system activities for malicious activities, examines network traffic flows to detect and prevent vulnerability exploits. Vulnerability exploits usually come in the form of malicious inputs to a target application or service that attackers use to interrupt and gain control of an application or machine[3].

IPS systems are expansion of intrusion detection systems because they both monitor network traffic, system activities for malicious activity, IPS are able to prevent or block intrusions that are detected. This can perform actions such as indicating an alarm, leaving the malicious packets.

2. Anti-Malware

Malware is short for malicious software, they are any software which hinder with the functioning of a device or can be used to enter into a device without consent or can be used to gather information which the user is not ready to share(simply put access restricted information). Type of malwares is virus, worms, Trojans, spyware, backdoors, etc. Malware have two categories:

- Malware which modifies the recourses such as memory, BIOS Code, PCI devices expansions EEPROMS [4].
- Malware that does not modify any of those resources, but only the resources which are dynamic by nature, like e.g. data sections, such as by modifying some function pointers in some kernel data structures, so that the attackers code gets executed instead of the original system or application.

Antivirus or Anti-virus software also sometimes called anti-malware software are used by computer and such devices to prevent them from getting infested with malwares.

They work in the real-time environment i.e. they will keep on scanning the systems and if and when anti-malware scans a malware it scans the severity of the threat and take action accordingly, user can also configure the anti-malware so that user know which action will be taken either it will be cleaned, quarantined or deleted.

4.PROPOSED SYSTEM

In the proposed system we use a firewall for the data leakage prevention, in this approach the administrator or the distributor sends a message to all the agents that are authorized to access the information and no one else, we have to be sure that the agents to whom we have provided the information should not or must not release this information to someone outside the authorized circle, as we cannot be sure what an an-authorized person may do with this information.

So we use a firewall approach to ensure that such a thing does not happen in the first place because in case of data leakage detection it can only be done when data has already been leaked , and by the time we find out who has leaked the information and to where it can already be too late to do anything.

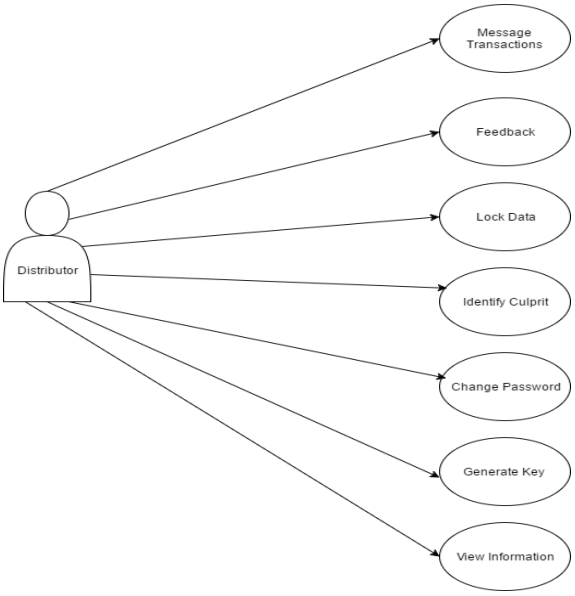

Fig -1: Use Case Diagram for Distributor

So we use data leakage prevention policy in which the administrator sends the message to the authorized parties and that document is scanned by the firewall because firewall is a software which keeps track of all the transactions and keep any unauthorized person to access our information but in our case the firewall has been modified such that it keeps the unauthorized user from sending or divulging sensitive information to someone outside the authorization policy[5].

The firewall stores the message sent by the administrator or the distributor to the agent and when that agent if he tries to send that information outside the work area via e-mal, or any other messaging service the firewall will scan the message and compare it with the message the distributor sent if the message is found consistent, an alert will be sent

to the administrator about the attempt in leaking of data (attempt because the firewall will prevent the agent to deliver the message) and the id of the agent who is responsible for the attempt[6].

After the alert has been sent the administrator can see which of his agent or agents have tried to leak the information and can take actions according to his will and what the company bylaws allow, because if they go unpunished they/he/her are most likely to try and leak data again and this time they might be successful.

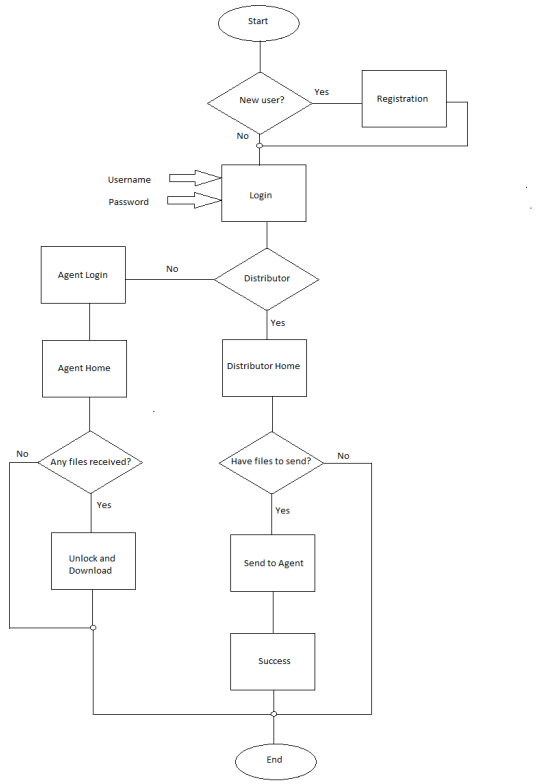

Fig -2:Flowchart for Agent

5.METHODOLOGY

In this paper we provide a description of all the steps that the data leakage prevention strategy will follow as to not all data to be accessed by unauthorized user outside the circle to which the administrator has determined :

1. The request is sent to the distributor or the distributor may initiate the process.
2. The request may be implicate or explicate.
3. If implicate a message is generated.
4. If explicate it is checked whether he is a legitimate user to which the information can be divulged(it is chosen by the distributor whether the agent is legitimate or not).
5. The message is sent after the verification.
6. The messaged is accessed by the recipient.
7. The recipient tries to send this message to anyone outside the organization or the authorized personal.
8. The message is scanned to see if it contains any confidential information.
9. If not the message is sent.
10. If yes, then the message is not sent and an alert is sent to the distributor.
11. The recipient is then identified and is reveled to the distributor.

These are the steps that are to be followed by the data leakage prevention program as to prevent leakage and identify the culprit so as to prevent such action from happening in the future.

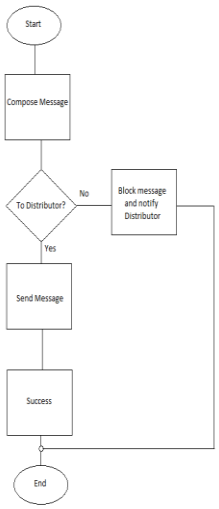

Fig -3:Flowchart for Distributor

In the prevention system the distributor or the administrator has full access the prevention system denies access to both agent and the target of the sent file from the agent outside the organization.

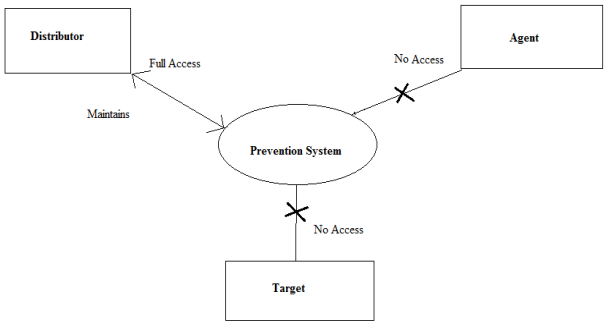

Fig -4: Control Flow Diagram for Prevention System

6.CONCLUSION

This paper help us to comprehend how much power is there in information and there are people in the same organization or company who may try to steal this information or pass it on to someone who may or may not use it to harm the organization, so we should take as many precautions as we can so that we does not end up being a victim of data leakage, and the first precaution that we need to take is to prevent the data getting leaked in the first place, and because sometimes data will get leaked so in that case we have data leakage detection but we don't know how much damage would have been done to the clients and the company so it is best to be prepared. As firewall is to filter who are allowed to access the computer and what we are trying to send outside the network because of this, Data leakage prevention with the help of a firewall will help us as it will not allow anyone who does not have authorization able to send a message outside the company or organization , and if he tries he will be caught and punished.

REFERENCES

[1] https://en.wikipedia.org/wiki/Intrusion_detection_system

[2] http://www.iosrjournals.org/iosr-jce/papers/vol1-issue3/S0132836.pdf

[3] https://www.paloaltonetworks.com/documentation/glossary/what-is-an-intrusion-prevention-system-ips

[4] https://vxheaven.org/lib/pdf/Introducing%20Stealth %20Malware%20Taxonomy.pdf

[5] https://digitalguardian.com/blog/cisos-guide-data-loss-prevention-dlp-strategy-tips-quick-wins-and-myths-avoid

[6] https://www.digitalocean.com/community/tutorials/what-is-a-firewall-and-how-does-it-work