International Journal of Advance Research in Science and Engineering
Vol. No.10, Issue No. 03, March 2021
www.ijarse.com
IJARSE
ISSN 2319 - 8354

CAUSES OF DATA LEAKAGES AND ITS PREVENTIVE MEASURES

Mohammed Inuwa Ali[1], Adzapeni Eric Agbeko Kofi[2]

Department of Computer Science , Lovely Professional University (India)

Abstract

Today, data plays a very important aspect of our life. In order to give information on the behaviour of any material, be it human or things around us, samples of data need to be collected, the more we think of information, the more we think of data, the term Data Leakage Prevention (*DLP) refers to the protection against data leakage. Data leakage/loss refers to an event in which important data is breached or leakage to the business, such as in a varus/malware attack. Data leakage prevention aims to prevent the unlawful transfer of data across organizational boundaries.*

Keywords*: Cybercriminals, data leakage, insider threats, malware*

Introduction

Data leakage is the unauthorized transmission of data from an organization to an external destination or recipient. When we think of data leakage, we think of data held on stolen or misplaced laptops or data that is leaked via email. However, the vast majority of data loss does not occur on electronic media; it occurs through printers, cameras, copiers, removable USB drives and even the dumpster dive for discarded documents.

Data loss refers to intentional or accidental data loss. It is not possible to access the data because the data is lost. In general, data loss can occur in any device that stores data. Data loss occurs due to hardware failures such as memory or disk failure, power outages or due to an environmental disaster such as a flood. In addition, malware and viruses can cause data loss on purpose.Data loss prevention aims to focus on these situations and take steps to avoid data loss. Making backup files and using anti-malware software are two options to prevent data loss.

Data leaks can be unintentional or intentional,for example, consider the case of an employee who takes an office report home and forgets it inside a bus or train, and someone else takes the file, is a data leakage scenario that occurs unintentionally oran employee who sends an important

organizational file to someone who has no legal right to obtain that information is an intentional data leak.

There is never a day without a confidential data breach that makesaheadline. Data leaks is a huge data security issue, that damage to any organization, regardless of size, can be severe. And can tarnished reputation, make massive financial penalties or cripple a lawsuits, it's a threat that any organization will want to protect itself from.

Types of Data Leaks

There are two types of data leaks and it is important to understand that the problem may be initiated through an external or internal source. All areas must besafeguarded to ensure that the most common data leakage threats are avoided.

1. The accidental breach

Unauthorized data leaks do not necessarily mean intentional or malicious. The good news is that the majority of data leaks are accidental. For example, an employee may unintentionally choose the wrong recipient when sending an email containing confidential data. Unfortunately, unintentional data leaks can still result in the same penalties and reputational damage that they do not mitigate legal liabilities.

When we think of data leaks, we think of data held on stolen or misplaced laptops or data that is leaked via email. However, the vast majority of data loss does not occur on an electronic medium, it occurs through printers, cameras, copiers, removable USB drives and even the dumpster dive for discarded documents. While an employee may have signed an employment contract that effectively signifies trust between employer and employee, there is nothing to stop them from data leaking and confidential information outside if they are unhappy or have promised a large payment by cybercriminals.

2. Electronic communications with malicious intent

Many organizations give employees access to the Internet, email and instant messaging as part of their role. The problem is that all of these media are capable of transferring files or accessing external sources on the Internet. Malware is often used to target these media and with a high success rate. For example, a cybercriminal could very easily send a legitimate corporate email account and request that sensitive information be sent to it. The user would unwillingly send the information, which could contain financial data or sensitive pricing information.Phishing attacks are another method of cyberattack with a high success rate of data leaks. Simply clicking on a link and visiting a web page that contains malicious code could allow an attacker to access a computer or network to retrieve the information they need.

Damage caused by a data leakage

- Government-imposed fines for non-compliance with privacy and security regulations
- Expensive repairs to computer systems
- Purchase of new security and cyber-security software
- The cost of hiring outside of public relations, legal and forensic advisors
- Loss of consumer confidence and damage to the organization's brand

A data leakage response team may include:

- IT security personnel
- Lawyers to determine legal responsibilities and liability
- Customer service associates will answer questions from affected customers
- Human resources personnel if the breach involves employee data
- A data protection officer
- Forensic consultants to trace the attack or uncover hidden malware

The inclusion of C-level senior managers in the data breach response team ensures that the data response plan receives the support and resources they need. The involvement of senior managers demonstrates a commitment to data readiness and encourages the participation of mid- and lower-level employees. However, to obtain senior management support, management managers must explain the consequences of a security breach based on its potential impact on the organization and how a data breach response plan can help the organization avoid liability costs, government fines and loss of revenue.

Causes of data leakage

Three common causes of data leaks are;

1. Insider threats - a malicious insider or attacker who has compromised a privileged user account abuses its permissions and attempts to move data outside the organization.
2. Attacker Extrusion - many cyber-attacks have sensitive data as their target. Attackers penetrate the security perimeter using techniques such as phishing, malware or code injection, and gain access to sensitive data.
3. Unintentional or negligent exposure to data - many data leaks occur as a result of employees losing sensitive data in public, providing open Internet access to data, or not restricting access by organizational policy.

Why data leakage prevention is considered an absolute necessity for companies of all sizes.

1. **Increasing external threats and attacks**

Organizations take data loss very seriously. However, as data thieves become more sophisticated every day, and many are finding new ways to access networks more frequently, organizations are facing increasing pressure to actively continue to look for new threats.

2. **Internal Threats**

Disgruntled employees are a prime example of insider threats - individuals who deliberately intend to cause harm to a company from within. They may do it themselves or try to get help from an outsider to carry out the attack.Since they already have access to data and may also have sensitive information about different staff members within the company, the attack can be more dangerous than an attempted breach from outside the organization. This is especially true if the disgruntled employee happens to be a high-ranking executive, as they typically have access to twice as much sensitive information compared to other employees.

3. **Accidental sharing of information**

The person in question may not intend to harm the company or jeopardize the company's data. They could simply be a victim of social engineering, a favourite method used by data thieves.The attacker typically studies the target (the organization) and chooses a victim (the employee) as a means. The usual tactics are to study the victim in depth and involve them in their plans, and they are completely unaware of it.They invariably try to have the victim accidentally reveal sensitive information without them realizing it.

4. **Poorly maintained BYOD**

Bring your own device (BYOD) strategies have helped many industries operate more efficiently. However, there are still industries that have not adopted Byod at all or have a poorly deployed and maintained BYOD solution.

Unfortunately, BYOD makes it easier for employees to inadvertently share sensitive information via their personal cell phones and tablets. They may not be aware of the level of data security that is either sitting idle inside the device or during data transmission.

5. **Cloud-based storage and services**

While we talk about the challenges of BYOD, another point to note is that employees may use their personal storage devices and personal cloud-based storage services such as Google Drive or Dropbox to store and share confidential company-owned information that is otherwise not supposed to leave the company's network and infrastructure.It is possible, especially in non-technology companies, that these individuals may not be aware of the proper protocols. It is the company's responsibility to

implement iron-clad network security measures to ensure that employees have the proper permissions and authorizations to access data, and ensure that it is only shared within the company's networks. Data loss through BYOD can be a common occurrence when security protocols have not been defined.

6. **Bad reputation within the industry**

If you don't have proper security measures in place and have the misfortune of repeated attacks as a result, your company could quickly earn a bad reputation.Data loss prevention is a topic that new business owners should pay special attention to when starting a business. Every new piece of data created, stored, used and shared from day one is sensitive information. Laying a solid foundation at the beginning will result in a little less worry down the road.

Today there are a myriad of techniques on the market that offer different types of content inspection. One thing to consider is that while many DLP providers have developed their own content engines, some use third-party technology that is not designed for DLP. For example, rather than creating patterns that match credit card numbers, a DLP provider may allow a search engine provider's technology to match credit card numbers. When evaluating DLP solutions, pay particular attention to the types of models detected by each solution versus a true body of sensitive data to confirm the accuracy of its content engine.

A complete DLP solution is a dedicated suite of products that identifies, monitors and protects critical data in your organization while it is used (at the end point), in motion (on the network) and at rest (in storage) using technologies such as classification, fingerprinting and accurate data matching. DLP solutions will also include centralized management, policy creation and law enforcement capabilities.

Some data loss prevention solutions may require considerable effort to be deployed, which can also result in significant ongoing maintenance costs. To make sure you choose the right solution for your organization's data protection needs, here are some key things to consider when evaluating options.

Preventing data loss solves three main goals that are common pain points for many organizations: privacy and compliance, intellectual property protection, and data visibility.

1. **Privacy/Compliance:** Does your organization select and store identifiable personal information (IPI), protected health information (ISP) or payment card information (CPI)? If this is the case, you are more than likely subject to compliance regulations, such as HIPAA (for PHI) and GDPR (for the personal data of EU residents), which require you to protect your customers' sensitive data. DLP can identify, classify and mark sensitive data and monitor activities and events surrounding this data. In addition, reporting capabilities provide the details needed for compliance audits.

2. **Intellectual Property Protection**: Will your organization have important intellectual property and trade or state secrets that could endanger your organization's financial health and branding in the event of loss or theft? DLP solutions such as Digital Guardian that use context-based classification can classify intellectual property in structured and unstructured forms. With the policies and controls in place, you can protect yourself from the unwanted exfiltration of this data.

3. **Data visibility:** Is your organization looking for additional visibility into data movement? A comprehensive enterprise DLP solution can help you see and track your endpoint, network and cloud data. This will give you visibility into how individual users in your organization interact with data.

A DLP solution can also monitor the parameters or channels through which this data flows. DLP solutions analyze data repositories, such as file sharing and servers, and then analyze and catalogue content. Some DLP products, or modules in DLP software suites, provide automated reports for incident response. They can also block the evacuation of sensitive data from the organization, or encrypt it before it is sent, according to the rules established by the organization.

However, technology is only one component of the DLP. Effective data security requires best practices of the DLP that include detailed policies and procedures for processing and storing sensitive data and dealing with security breaches. The effectiveness of the DLP also depends on IT staff's knowledge of data security requirements and end-user awareness of data security practices.

While these are the three main use cases, DLP can address a variety of other pain points, including internal threats, Office 365 data security, user and feature behaviour analysis, and advanced threats.

Understanding the types of data and the level of protection each type needs is a crucial first step and the need to understand where this data resides, how and where it is used, and by whom.

A complete DLP solution should have features that will automatically discover and classify data on your network, from device to cloud. It should also help you get visibility into where your data is going and who is using it, which will help you better understand what protection you need and where, as well as identify any gaps that may exist in your current processes.

With a multitude of ways in which data can leave your organization, whether through malicious actions or inadvertently, it's important to understand which protocols a DLP solution can analyze and act against. Can it control USB ports so that sensitive data cannot be downloaded to a USB stick or other external device? Does it cover apps that users can or can't use on specific devices? An effective DLP solution should be able to apply policies for information leaving the network across a wide range of protocols - email, webmail, instant messaging, wikis, blogs, FTP, cloud services, and more.

The understanding of organization's needs in key areas and what potential DLP solutions can provide will help will help in identifying the solution that best suits the environment and available resources.

Below are the recommended guidelines for developing an effective DLP program:

1. **Implement a single, centralized DLP program.** Many organizations implement inconsistent and ad hoc DLP practices and technologies that are implemented by various departments and business units. This inconsistency results in a lack of visibility on data assets and low data security. In addition, employees tend to ignore the DLP programs of departments that the rest of the organization does not support.

2. **Assess internal resources.** To create and execute a DLP plan, organizations need staff with expertise in DLP, including DLP risk analysis, data breach response and reporting, data protection laws, training and awareness of DLP. Some government regulations require organizations to employ in-house staff or use external consultants with knowledge of data protection. For example, the GDPR includes provisions that affect organisations that sell goods or services to consumers in the European Union (EU) or monitor their behaviour. The GDPR requires the appointment of a Data Protection Officer (DPO) or staff who can assume the responsibilities of the DPO, including conducting compliance audits, monitoring DLP performance, training employees on compliance requirements, and the liaison role between the organization and compliance authorities.

3. **Make an inventory and evaluation.** An assessment of the types of data and their value to the organization is an important first step in the implementation of a DLP program. The aim is to identify the relevant data, where it is stored and whether it is sensitive data (intellectual property, confidential information or data covered by regulation). Some DLP products, such as McAfee DLP, can quickly identify information assets by analysing file metadata and cataloguing the result, or if necessary, by opening files to analyze their content. The next step is to assess the risk associated with each type of data in the event of a leakage of the data. The data output points and the likely cost to the organization in the event of data loss must also be taken into account. Loss of information about employee benefits programs does not present the same level of risk as the loss of 1,000 patient medical records or 100,000 bank account numbers and passwords.

4. **Implement in stages**. The DLP is a long-term process that is best implemented in stages. The most effective approach is to prioritize data types and communication channels. Similarly, consider implementing DLP software components or modules as needed, based on the organization's priorities, rather than at once. Risk analysis and data inventory help establish these priorities.

5. **Create a classification system.** Before DLP policies can be created and executed, the company needs a data classification framework or taxonomy for structured and unstructured data. Data security categories may include confidential, internal, public, personally identifiable information (PII), financial data, regulated data, intellectual property and others. DLP products can analyze data using a pre-configured taxonomy, which the company can then customize, to help identify key data

categories. While DLP software automates and speeds up classification, humans select and customize categories.

6. Make employees aware. Employee awareness and acceptance of safety policies and procedures is essential to the DLP. Education and training efforts, such as courses, online training, periodic emails and posters, can improve employees' understanding of the importance of data security and strengthen their ability to follow recommended DLP best practices. Sanctions for data breaches can also improve compliance, especially if they are clearly defined. The SANS Institute offers a variety of data security training and awareness resources.

When you're looking for DLP providers, you should set your evaluation criteria

Some of these questions are:

1. What types of deployment architectures are available?
2. Do they support Windows, Linux and OS X with feature parity?
3. What deployment options do they offer? Do they provide managed services?
4. Do you have to defend yourself against mainly internal or external threats? Or both?

5. Do you need to conduct content or context-based inspections and classifications? Will your users be able to self-upgrade documents? Do you need a mix of multiple methods?

6. Are you most concerned with protecting structured or unstructured data?
7. Do you plan to see and apply the data movement based on policies, events or users?

10. How fast do you need to deploy your DLP program?

11. Will you need additional staff to manage your DLP program?

These clearly define the roles and responsibilities of those involved in your organization's DLP program. The establishment of duties and duties based on functions will provide checks and balances.

Policies

A number of data protection laws are already in place, not to mention all of the pending legal requirements and potential laws that are being drafted around the world. A typical DLP policy contains three elements:

- Location: Where this policy will be applied
- Condition: Essentially, the parameters that the strategy seeks to prevent data loss
- Action: If a situation meets the conditions set, a measure is taken to prevent losses

A DLP policy is configured to detect information protected by GDPR. The location is where the personal information is stored.

Conditions may include:

- Data is not used in the way it has been agreed by the user

- Older data that needs to be deleted to stay in compliance
- Personal data stored in another unsecured location

The shares correspond to the condition. For example, data may be deleted if it is deemed to be contrary to GDPR regulations, or personal data may be blocked when it is found in an unverified environment.

Best Practices for Establishing a Data Leakage Prevention Policy

A data leakage prevention policy can help organizations prevent unauthorized access to data and protect themselves from potential damage. While no protection is bulletproof, there are best practices that can help establish a successful data protection policy:

Identify the data that the policy is primarily intended to protect. Most often, data are categorized according to vulnerability and risk factors. Taking the time to understand and categorize the data can lead to greater organizational knowledge.

Establish criteria for evaluating data loss prevention providers. But creating an evaluation framework with the right questions can help lead to an informed purchasing decision.

Clearly define the role of those who will be involved in preventing data loss. It is not just a question of who will monitor the use of the data and make the rules. Segregation of responsibilities helps prevent abuse.

Keep it simple at first. Choose a specific type of data or risk to process. The goal is to secure the most critical data and get a measurable victory early, and then build on that.

Get help from the organization's management. Each head of department or unit has a role to play in developing a data loss prevention policy that aligns with the corporate culture. It is a strategy that affects all departments and functions.

Educate everyone in the organization about how and why the data loss prevention policy is in place. Many executives see employees as the weak link in data loss prevention, but do not consider safety education a priority.Carefully document data loss prevention processes. A written policy should focus on data protection.

Set and share steps for success. Data loss prevention measures will determine the return on investment of policies and solutions. They can also help determine effectiveness.

Watch for the use of the data before blocking it. First, set up data loss prevention tools to report sensitive data loss. Make sure that all rules that block data transfer will not disrupt the workflow.

Conclusion

DLP is a program, not a product. Installing a DLP tool is only the first step in preventing data loss. While you can get quick ways of prevention and understanding that DLP is a program that you need to work at all time and it will help you achieve lasting success.

To be successful,work with business unit leaders to define the DLP strategies that will govern your organization's data. This will ensure that the various business units are aware of the policies in place and how they might be affected. Keep in mind that there is no single good way to develop DLP policies.

Set success indicators and share relationships with business leaders. Determine the key performance indicators (Performance INDICATORs) that you need to measure and monitor closely to determine the success of your DLP program and areas for improvement. Share these settings with your organization's leaders to show the positive impact of DLP and its business value.

Carefully document your processes. This will help you to consistently implement policies, will give you a registration document for when exams are needed, and will also be useful when ingesting new team members or employees.

Reference

https://www.verizon.com/about/investors/annual-report

https://www.sciencedirect.com/science/article/abs/pii/S1084804516000102

https://www.hindawi.com/journals/wcmc/2018/5823439/

https://link.springer.com/chapter/10.1007/978-1-4614-2053-8_3