M.Sharmiladevi *et al*, International Journal of Computer Science and Mobile Applications,
Vol.8 Issue. 3, March- 2020, pg. 23-31 **ISSN: 2321-8363**
Impact Factor: 5.515

Data Loss Prevention in Detecting and Preventing Data Breaches: An Overview

M.Sharmiladevi[1]; Rafion Houdhoyfi[2]; Vignesh Ramamoorthy H[3]
[1, 2]Under Graduates, [3]Assistant Professor
Department of Information Technology and Cognitive Systems
Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India
[1] sharmiladevim18mit017@skasc.ac.in
[2] rafionhoudhoyfi18mit022@skasc.ac.in
[3] hvigneshram@gmail.com

Abstract

In modern virtual economy, statistics enters and leaves our on-line world at record rates. A typical corporation sends and receives thousands and thousands of e-mail messages and downloads, saves, and transfers hundreds of files via diverse channels on an everyday basis. Enterprises also hold touchy facts that customers, business partners, regulators, and shareholders count on them to protect. Unfortunately, groups constantly fall sufferer to big information loss, and high-profile information leakages related to sensitive private and corporate records continue to seem Data loss may want to substantially damage a company's competitiveness and reputation and can also invite lawsuits or regulatory effects for lax security. Therefore, groups need to take measures to recognize the touchy information they keep, how it's controlled, and how to prevent it from being leaked or compromised.

Keywords: Data, Statistics, confidential, leakages, digital

1. Introduction

Data in each enterprise is one of the most vital assets, Therefore the safety of this statistics ought to take the first priority. Although the businesses have safety measurements and technical parries inclusive of firewalls, nevertheless the statistics leakage occurs. The records leakage happens when sensitive statistics is found out to unauthorized parties whether it is intentionally or not. The information leaked may reason severe threats to a business enterprise. The loss of private or sensitive facts can severely effect a enterprise's reputation, clients and employee confidence, competitive benefit and in a few cases lead to the closure of the company, or political crises together with Wiki leaks Data leakage problem need to be solved the usage of the Data Leakage/Loss Prevention System. DLP solutions assist identifying, monitoring, protective and decreasing the risks of sensitive-facts leakage. It is used to detect and prevent unauthorized consumer from getting sensitive information, and even to guard personal data that can be by chance shared. In this paper we can first talk about the present safety techniques utilized in Data safety and within the 2nd section we are able to speak about records leakage prevention systems, finally. We will compare between them.

2. Data Loss Prevention and its Importance

Data Loss Prevention (DLP) is that the practice of detecting and preventing data breaches, ex filtration, or unwanted destruction of sensitive data. Organizations use DLP to guard and secure their data and suits regulations. Data loss refers to an occasion during which important data is lost to the enterprise, like during a ransom ware attack. Data loss prevention focuses on preventing illicit transfer of knowledge outside organizational boundaries.

Organizations generally use DLP to:

- Protect Personally Identifiable Information (PII) and comply with applicable regulations.
- Protect Intellectual Property essential for the organization.
- Achieve data visibility in huge organizations.
- Secure cell body of workers and implement security in Bring Your Own Device (BYOD) environments.
- Secure data on far off cloud systems.

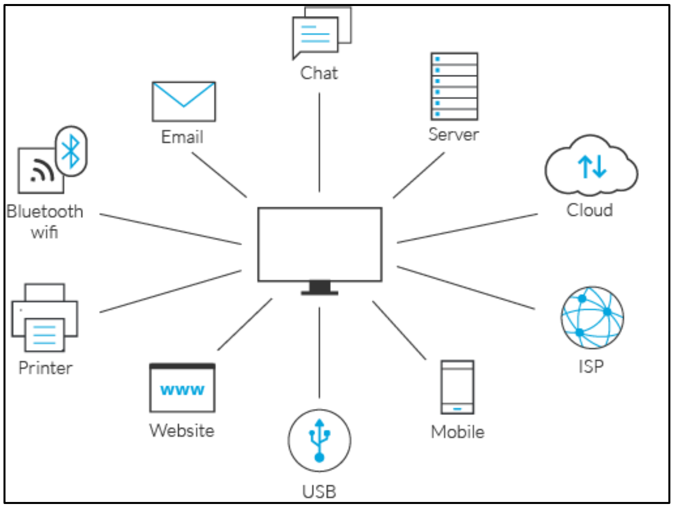

Figure 1: Overview of Data Loss Prevention

Importance of Data Loss Prevention

According to a Gartner CISO survey, records loss prevention (DLP) may be a pinnacle priority for CISOs. Data loss prevention (DLP) is typically defined as any solution or procedure that identifies confidential data, tracks that statistics because it moves through and out of the corporate and prevents unauthorized disclosure of records by means of growing and enforcing disclosure policies. Since confidential statistics can reside on quite few computing devices and move through many network access points (wire line, wireless, VPNs, etc.), there are many solutions that are tackling the effort of knowledge loss, facts healing and facts leaks.

(i) Causes of Data Leak

- **Insider threats** — a malicious insider, or an attacker who has compromised a privileged user account, abuses their permissions and tries to move data outdoor the organization.

- **Extrusion by means of attackers** — many cyber-attacks have sensitive facts as their target. Attackers penetrate the safety perimeter using techniques like phishing, malware or code injection, and advantage get entry to to sensitive facts.
- **Unintentional or negligent statistics exposure** — many information leaks occur because of personnel who lose sensitive facts in public, offer open Internet get admission to information, or fail to restriction get admission to in keeping with organizational policies.

(ii) Components of Data Loss Prevention

Manage – Define enterprise facts usage policies, report information loss incidents, and set up incident response capability to enable corrective actions to remediate violations. Data loss prevention isn't always just an era issue; it is also a policy and policy control issue. Enterprise records usage guidelines ought to deal with issues consisting of how get entry to information is determined; how facts get right of entry to be authenticated; and how regulations are enforced. Management functionalities must also include facts loss reporting functionality and incident remediation workflow control.

Discover – Define the sensitivity of enterprise statistics, create an inventory of sensitive statistics, find sensitive facts wherever it is stored, and manage data cleanup. This includes discovering and inventorying sensitive statistics at relaxation in file servers, databases, record and facts management, email repositories, and net content and applications; and scanning for sensitive information stored at the endpoint together with laptops, desktops, and workstations at remote workplaces that allows you to stock, secure, or relocate that facts.

Monitor – Monitor the use of sensitive statistics, recognize sensitive information usage pattern, and gain corporation visibility. This could encompass monitoring records in motion by way of inspecting community communications including email, Instant Messaging (IM), web, FTP, P2P and others for confidential records in violation of records safety policy; and monitoring facts at the end factors consisting of downloading to nearby drives, coping to USB or other removable media devices, burning to CD/DVDs, and printing or faxing electronically.

Protect - Enforce protection regulations to proactively steady facts and save you sensitive records from leaving an enterprise. Automatic safety of sensitive records across endpoint, network and storage systems. This includes protecting records at relaxation with automated encryption, quarantine, and remove; Restrict printing, saving, copying, accessing, movement of and downloading of sensitive records to detachable media or other drives; and stopping information in motion from being dispatched in violation of statistics security coverage or encrypting information for steady exchange.

3. Types of DLP

Network: Network (statistics in motion) technology is typically set up at network egress points close to the perimeter. It analyzes community traffic to detect sensitive information that is being dispatched in violation of data protection policies. Multiple security control factors may document pastime to be analyzed via a central management server.

- **End point:**

  Endpoint (statistics in use) structures run on internal end-consumer workstations or servers. Like network-primarily based structures, endpoint-based technology can address internal in addition to outside communications. it can therefore be used to manipulate records drift between agencies or forms of users (e.g. 'Chinese walls'). They can also control e mail and Instant Messaging communications earlier than they attain the corporate archive, such that a blocked communication (i.e., one that was never sent, and therefore now not concern to retention rules) will not be identified in a subsequent legal discovery situation. Endpoint systems have the advantage that they can display and control access to physical devices (inclusive of mobile gadgets with data garage capabilities) and in some cases can access data earlier than it is encrypted. Some endpoint-primarily based structures provide software controls to block attempted transmissions of confidential data and provide immediately user feedback. They must be set up on each notebook within the network, can't be used on mobile devices (e.g., mobile telephones and PDAs) or where they can't be practically installed (as an example on a computing device in an Internet cafe).

- **Data identification:**

  DLP includes techniques for figuring out confidential or sensitive information. Sometimes pressured with discovery, statistics identity is a manner via which corporations use a DLP era to determine what to appearance for.

  Data is classified as either based or unstructured. Structured facts is living in fixed fields inside a file which includes a spreadsheet, whilst unstructured information refers to free-form textual content or media in text documents, PDF documents and video. An anticipated 80% of all records are unstructured and 20% dependent.

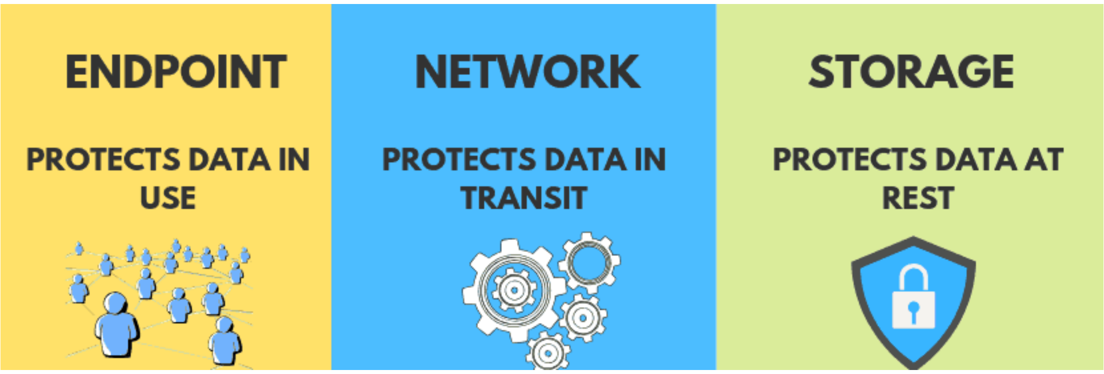

Figure 2: Types of Data Loss Solutions

- **Data leak detection:**

  Sometimes a facts distributor offers sensitive records to at least one or more 0.33 parties. Sometime later, a number of the data is found in an unauthorized place (e.g., on the net or on a user's laptop). The distributor should then check out the source of the leak.

- **Data at rest:**

  "Data at rest" specially refers to vintage archived statistics. This facts is of fantastic issue to agencies and government institutions absolutely because the longer information is left unused in storage, the more likely it might be retrieved by means of unauthorized individuals. Protecting such information involves methods such as get admission to control, facts encryption and facts retention policies.

- **Data in use:**

  "Data at rest" specially refers to vintage archived statistics. This facts is of fantastic issue to agencies and government institutions absolutely because the longer information is left unused in storage, the more likely it might be retrieved by means of unauthorized individuals. Protecting such information involves methods such as get admission to control, facts encryption and facts retention policies.

- **Data in motion:**

  "Data in motion" is information this is traversing via a community to an endpoint. Networks can be inner or external. DLP systems that protect records in-motion display sensitive records visiting throughout a network through various conversation channels.

4. Strategies and Tools for Data Loss Prevention

- ***Standard Security Measures*****:** an agency ought to have the fundamental statistics safety infrastructure to assist its operations, such as firewalls, intrusion detection, prevention systems, anti-malware and anti-virus protection, and vulnerability management or risk management systems.
- ***Mature Security Measures*****:** Some businesses may decide they need introduced tracking and threat protection that various superior protection analytics tools offer. This may include security system learning, honeypots, network visitors analyzers, records integrity controls, user identification exams or activity-based totally verification, and more to locate irregular statistics access.

(i) DLP Specific Tools

This equipment will particularly block attempts to copy or transmit sensitive data to an unauthorized location, intentionally or unintentionally. DLP particular gear can help in many regions including:

- Simplified management of DLP policy
- Educate and alert customers without concerning IT/safety personnel
- Monitor sensitive emails earlier than leaving the organization
- Identify PII, HIPAA, SOX, PCI DSS, GDPR or different compliance related statistics

- Use record matching to save you sensitive documents or records from leaving the organization
- Protect information in motion thru SMTP, HTTP, and FTP facts

Strong protection coverage will aid in making DLP work properly. Regular audits should take place, and protection incident and remediation should be well documented and addressed. To get began with DLP, an business enterprise need to perform a class of its dependent and unstructured records sets so that its policies can country what information is classified as sensitive.

Data loss is an unsettling problem across the world. Organizations of all kinds and sizes are at risk for facts loss. When you're making DLP an active a part of your safety strategy, you can gain aggressive advantage. Your maximum sensitive facts will be blanketed by means of this gear which will, in turn, protect your emblem and shareholder value. Your DLP tools and techniques can even help save you the permanent emblem reputation damage you might incur from a information loss incident or records breach. If you have questions on DLP strategies or tools, our safety professionals are here to offer steerage and considerate discourse on how to stand up DLP.

5. Techniques for Preventing Data Loss

- *Compliance*
  If your business is regulated, compliance is a critical piece of a DLP plan. Starting at mandatory regulation helps to ensure custom data policies don't contradict compliance. For example, Healthcare companies must comply with HIPAA compliance. To process credit cards, you need a PCI-DSS compliance plan.

- *Organize Data by Risk and Vulnerability*
  To start, identify, and organize data by type. The next step is to analyze each data type. Look at the level of risk each category presents, then their vulnerability to establish a list of targets. Start with the riskiest and most vulnerable data types. Then, implement a set of policies and technology to protect them.

- *Define User Roles*
  Users should have clearly-defined roles to make sure that users can only access data that's necessary for their job. Examples might include a "Sales Agent" job that allows access to payment card data. Another example might be "Senior Technical Support," which has access to bug reports and engineering fixes. Data can be restricted from specific user roles as well. A Senior Technical Support agent doesn't usually need access to payment card data.

- *Involve Key Stakeholders*
  No one knows the business like the people doing the work. Take time to involve leaders from different parts of the company. They often see vulnerabilities that upper management can overlook and also anticipate problems or conflicts with new policies. Employees are more likely to buy into a new plan if they are involved in developing it.

- *Create policies, implement technology*
  It should go without saying that DLP plans should be enacted. If the DLP plan requires a review of all firewall settings, make sure employees are tasked with that job. Create a timeline for buying, testing, and installing new DLP software.

- *Automate*
  Humans make mistakes, and our ability to see problems can be limited. Automating data loss tools and policies helps reduce the risk of human error. Some tasks cannot be automated. Wherever possible, try to create systems that protect data automatically. For example, spam-filtering software can prevent phishing

attacks. Users are protected because they never have a chance to see (or be manipulated by) a phishing email.

- *Educate*
  Education takes leader involvement to the next step. Leaders who help develop DLP plans already understand the policy. The next step is to train individual team members. A good data loss prevention strategy is to teach what, then why. Explain the new systems and software and how to use it. Then have a conversation about why this change is essential. Help them understand how protecting data is a win for the company and its customers.

- *Document*
  Write down a master plan of the DLP solution. Proper documentation is essential for a couple of reasons. First, it helps keep the project on task. Everyone can refer back to the agreements and plan that was made. Second, it helps keep a record of what's been implemented. The documentation is especially helpful if there are instructions. At a minimum, leave a brief note with the "how" and "why" for each part of the DLP plan.

- *Measure*
  Once the DLP plan is in place, check back regularly, and review the progress. Most data loss prevention tools have reporting metrics. Check to see how many intrusion attempts have been blocked. Review server logs to verify that data is being used appropriately.

- *Delete Unnecessary Data*
  It can be tempting to hold on to data forever. We never know when we might need it! Old, unused data can be a liability. If server logs are no longer required after seven years, delete them. If you can't bear to delete them, archive them in long-term, secure storage. That data might seem valuable. But the value can easily be outstripped by its vulnerability. When weighed against the cost of a data breach, old data might not seem quite so important.

- *DLP Statistics*

  Cybercrime is big business. Here are a few sobering facts about Data Loss and the consequences.:

  - Data loss is estimated to have generated over a half billion dollars profit for cybercriminals in 2018.
  - Security breaches increased by 11% in 2018.
  - 43% of cyber-attacks target small businesses.
  - Ransomware statistics show that is costs businesses more than $75 billion per year.
  - 83% of IT security professionals have experienced phishing attacks. Source: (Wombat Security)
  - It takes an average of 50 days between discovery and reporting of a breach.
  - DLP is a culture, not a silver bullet
  - Building a data loss prevention strategy is a critical component for today's digital businesses.

The landscape of digital crime is continually evolving. A quick, one-and-done solution may be effective today, but obsolete tomorrow. Instead, develop policies as living documents. If a vulnerability is discovered, a flexible and growth-oriented security policy can better adapt to new threats. Eventually, data protection will grow to become a core feature of company culture.

6. Future of DLP

In present day times, it would not be easy in an effort to find any enterprise, which doesn't acknowledge the significance of virtual security. With each of enterprise segments; from logistics & distribution to healthcare, from telecom to trading, from engineering to advertising and designing; all rely on numerous digital gear for gathering records and data in order to perform their organizational work approaches and business-oriented tasks. Nowadays, huge quantity of statistics is produced, and this has by no means been the case ever with information being produced in such big quantities and at such excellent speed. Such massive quantity of records and facts together with omnipresent internet connectivity is something that still lures the hackers, so records leak dangers increase manifolds. This has result in new breaches being made public every other day and thus, organizational security has climbed the enterprise priority ladders, becoming one in every of the pinnacle issues today. The conventional approach closer to records security covered the strategies mainly aimed towards a company's integrated IT infrastructure, which includes antivirus software, get proper of access to control, DLP and firewalls, which no longer ensures total organization protection against statistics breaches.

7. Conclusion

A DLP technique needs a whole lot of idea and planning. A technical solution will allow you to have far more manipulated of your information. However points to recall is DLP will never be perfect. End of the day if a user wants a piece of records he or she will be able to see in front of their display screen they can just copy the whole thing onto a bit of paper or take an picture of the display with a camera, so there's no perfect remedy.

Now we are not pronouncing do not hassle will a DLP solution because it does give you a big amount of control and visibility to your records and provides other blessings together with consumer education and awareness, potential to see where sensitive facts resides in the network, permits you to meet rules and more. However be conscious that even though the solution performs a BIG part in DLP, still it is a subset of the procedure and an investment that wishes to be maintained.

References

[1] Leman-Meza, B., Burns, P., Eavenson, M., Palaniswami, D., & Sheth, A. (2005). An ontological approach to the document access problem of insider threat, proceedings of the IEEE international conference on intelligence and security informatics (ISI) 2005 (pp. 486–491). Georgia: Atlanta.

[2] Anderson, R.H. (1999). Research and development initiatives focused on preventing, detecting, and responding to insider misuse of critical defense information systems (RAND CF-151-OSD). Technical report, RAND Corporation.

[3] Anderson, R.H., Bozek, T., Longstaff, T., Meitzler, W., Skroch, M., & VanWyk, K. (2000). Research on mitigating the insider threat to information systems #2 (RAND CF-163-DARPA). Technical report, RAND Corporation.

[4] Brackney, R.C., & Anderson, R.H. (2004). Understanding the insider threat (RAND CF-196-ARDA). Technical report, RAND Corporation.

[5] Cappelli, D., Moore, A., & Trzeciak, R. (2012). The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes (Theft, Sabotage, Fraud). Addison-Wesley Professional.

[6] Chivers, H., Clark, J. A., Nobles, P., Shaikh, S. A., & Chen, H. (2013). Knowing who to watch: Identifying attackers whose actions are hidden within false alarms and background noise. Information Systems Frontiers, 15(1). doi:10.1007/s10796-010-9268-7.

[7] Farahmand, F., & Spafford, E. H. (2013). Understanding insiders: An analysis of risk-taking behavior. Informations System Frontiers, 15(1). doi:10.1007/s10796-010-9265-x.

[8] Hanley, M., & Montelibano, J., (2011). Insider Threat Control: Using Centralized Logging to Detect Data Exfiltration Near Insider Termination, Technical Note, CERT.

[9] Hunker, J., & Probst, C. (2011). Insiders and insider threats: An overview of definitions and mitigation techniques. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 2(1), 4–27.

[10] Kandias, M., Mylonas, A., Virvilis, N., Theoharidou, M., & Gritzalis, D. (2010). An insider threat prediction model, proceedings of the trust, privacy, and security in digital business 7th international conference (TrustBus2010) (LNCS 6264) (pp. 26–37). Spain: Bilbao.

[11] Kim, S., Cho, N. W., Lee, Y. J., Kang, S., Kim, T., Hwang, H., et al. (2013). Application of density-based outlier detection to database activity monitoring. Information Systems Frontiers, 15(1). doi:10.1007/s10796-010-9266-9.

[12] Matthew, S., Petropoulos, M., Mgo, H., & Upadhyaya, S. (2010). A data-centric approach to insider attack detection in database systems, in Proceedings of Recent Advances in Intrusion Detection: 13th International Symposium, RAID 2010, Ottowa, Ontario, Canada, (LNCS 6307), 382–401.

[13] Maybury, M., Chase, P., Cheikes, B., Brackney, D., Matzner, S., Hetherington, T., et al. (2005). Analysis and detection of malicious insiders. In Proceedings of the 2005 Intl. Conference on Intelligence Analysis.

[14] McCormick, M. (2008). Data theft: a prototypical insider threat. In S. Stolfo, S. Bellovin, S. Hershkop, A. Keromytis, S. Sinclair, & S. Smith (Eds.), Insider attack and cyber security: beyond the hacker (pp. 52–67). New York: Springer.

[15] Moore, A., Cappelli, D., Caron, T., Shaw, E., & Trzeciak, R. (2009). Insider Theft of Intellectual Property for Business Advantage: A Preliminary Model, First International Workshop on Managing Insider Security Threats (MIST 2009).

[16] Moore, A., Hanley, M., & Munide, D. (2012). A Pattern for Increased Monitoring for Intellectual Property Theft by Departing Insiders, Technical Report, CERT.

[17] Panigrahi, S., Sural, S., & Majumdar, A. K. (2013). Two-stage database intrusion detection by combining multiple evidence and belief update. Information Systems Frontiers, 15(1). doi:10.1007/s10796-010-9252-2.

[18] Phyo, A. H., & Furnell, S. M. (2004). Detection-oriented classification of insider IT misuse, proceedings of the 3rd security conference. Nevada: Las Vegas.

[19] Predd, J., Pfleeger, S. L., Hunker, J., & Bulford, C. (2008). Insiders behaving badly. IEEE Security and Privacy, 6(4), 66–70.

[20] Schultz, E. E. (2002). A framework for understanding and predicting insider attacks. Computers & Security, 21(6), 526–531.

[21] Shaw, E., & Stock, H. (2011). Behavioral Risk Indicators of Malicious Insider Theft of Intellectual Property: Misreading the Writing on the Wall, White Paper, Symantec.

[22] Shaw, E., Ruby, K. G., & Post, J. M. (1998). The insider threat to information systems: The psychology of the dangerous insider. Security Awareness Bulletin, 2, 1998.

[23] Verizon. (2012). 2012 Data Breach Investigations Report. www.verizonbusiness.com/about/events/2012dbir/. Accessed 6 February 2013.

[24] Wood, B.J. (2000). An Insider Threat Model for Adversary Simulation, SRI International, Cyber Defense Research Center, System Design Laboratory, Albuquerque, New Mexico.

[25] Wood, S., & Wiskoff, M.F. (2002). Americans who spied against their country since WorldWar II. Technical Report PERS-TR-92-005, Defense Personnel Security Research and Education Center (PERSEREC).

[26] https://www.researchgate.net/publication/266617827_Data_LeakageLoss_Prevention_Systems_DLP

[27] https://ieeexplore.ieee.org/abstract/document/6916624

[28] https://www.veracode.com/security/guide-data-loss-prevention

[29] https://www.data-recovery-solutions.com/blog/common-causes-of-data-loss/

[30] https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=904672

[31] https://cipher.com/blog/why-data-loss-prevention-dlp-matters-to-your-security-strategy/

[32] http://www.internet-computer-security.com/Security%20Guides/Other/Data-Loss-Prevention.html