Data Leakage Prevention: E-Mail Protection via Gateway

Kamaljeet Kaur, Ishu Gupta and Ashutosh Kumar Singh

1Govt. Sr. Sec. School, Ambala, Haryana, India
2, 3Department of Computer Applications, National Institute of Technology, Kurukshetra, Haryana, India

er.kamal1986@gmail.com, ishugupta23@gmail.com

Abstract. Protection of digital assets and intellectual property is becoming a challenge for most of the companies. Due to increasing availability of database services on internet, data may be insecure after passing through precarious networks. To protect intellectual property (IP) is a major concern for today's organizations, because a leakage that compromises IP means, sensitive information of a company is in the hands of biggest competitors. Electronic information processing and communication is replacing paper in many applications increasingly. Instead of paper, an email is being used for communication at workplace and from social media logins to bank accounts. Nowadays an email is becoming a mainstream business tool. An email can be misused to leave company’s sensitive data open to compromise. So, it may be of little surprise that attacks on emails are common. So, here we need an email protection system (EPS) that will protect information to leave organization via mail. In this paper, we developed an algorithm that will offer email protection via gateway during data transfer. This algorithm matches the patterns with the keywords stored in the database and then takes the actions accordingly to protect the data. This paper describes why email protection is important? How companies can protect their confidential information from being leaked by insiders.

1. Introduction
In information security, data leakage is unwanted disclosure of confidential information. Companies placed important data like Social Security no’s, credit card information, data related to their financial property etc. in their databases; therefore protection of data must have the first priority [1] [2]. Data security is comprised through unintentional and unwise behaviors of employees and IT professionals. The survey shows that employees around the world are responsible to put corporate & personal data at risk [3]. When an authorized user of company discloses confidential information to external entities called as a data leakage.

Data leakage has great impact on the company’s reputation [4]. According to survey report done by IBM, 46% of the companies suffered from reputational damage due to data leakage, including confidential employee information, customer information etc. Data leakage incidents usually occur when employees performed everyday tasks, such as sending e-mail that inadvertently contains sensitive information [5]. The main reason for using email is possibly the convenience and the speed with which it can be transmitted, irrespective of geographical distances. Sensitive information, such as account statements, credit card scores, and information about sanctions, is being exchanged through
emails [6]. According to industry analysts, email volume in companies is growing by more than 30% annually, and the average user receives 7MB of data per day via email. As a result of this growth, handling an email is becoming a critical business.

As shown in Figure 1, three main principles of Information Security are involved in maintaining the Confidentiality, Integrity, and Availability of information resources. A fault in any one of these three key areas will weaken the security posture of an email system and open the door to misuse. Therefore, the contents of emails are more important and valuable than ever, and their security has raised many concerns [7].

![Figure 1: Three main principles of security.](image)

2. Threats to an E-Mail

Employees put company’s sensitive data and assets at risk despite policies that define correct procedures. The following examples show how employees intentionally and unintentionally leak sensitive data.

- **Usage of Unauthorized applications**: In companies, usage of personal emails can place sensitive data and personal information at risk. According to a survey report, 63 percent of employees admit that they use work computer for personal usage. These applications do not follow corporate security standards. As a result, data leakage by an employee is high.

- **Misuse of Corporate Computers**: Employees intentionally use company’s computers in many ways that undermine IT security policies which includes sharing of work devices and sensitive information with non-employees. These behaviors can result in leaking out the IP of company, which poses serious threats to company’s security and profitability.

- **Misuse of Passwords and Login/Logout Procedures**: When an employee leaves a system logged in and with a password attached to it that invites an attacker to steal the sensitive data at their leisure. If employee used that computer for personal use which means information is now willingly available to the attacker.

3. Why E-Mail Protection is Important?

Email is used by millions of people to communicate around the globe and is a critical application for many businesses. Email messages passes through intermediate computers before reaching their final destination and it is relatively easy for attackers to intercept and read messages. The backups of these can remain up to several months on their server, even if we delete them from our mailbox. Employee of an organization accesses important files and sends it by their personal mails or by any other web service to the outside of the network. A total of 95% of such losses are unintentional. In fact, in 2015, over 200 billion emails were sent & received each day worldwide [8] [9].
- **E-Mail Protection via Gateway**
  Data leakage prevention solution can be configured on three levels  
  i) *Individual user level*  
  ii) *Domain level*  
  iii) *At gateway level.*

- **Individual user level:** Some companies facilitates to their employees that they can complete their work from home, in this situation their personal computers should be connected to their office PC via RSA token.

- **Domain level:** A company can prevent data leakage on domain level by enforcing the rule that an employee must be restricted to the company’s domain. But still there are departments in companies that have rights to send information outside of the network like the sales department of the company.

- **Gateway level:** In a work place, gateway computer route all the traffic from one network to another. When an employee of the company sends an email to the outside of the network, firstly mail passes through gateway computer then it leaves the network [10] [11].

4. **Proposed System**

Usage of email is increasing day by day and it is vulnerable in many ways. Security of an email is the main concern for companies & it includes confidentiality that ensures information will not expose to unauthorized entities.

To maintain confidentiality and integrity of an email, we proposed an Email Protection System. Our proposed EPS consider two parameters i.e. data states and deployment. In general data consist three states i.e. data at rest, data in use and data in motion. In this paper our focus is on the state when data moves from one network to another. So we deployed our Email Protection System on gateway that will protect an email consisting confidential information.

As shown in Figure 2, in companies there are sensitive departments that consists sensitive information and they have right to send any information outside of the network for example: Sales department of the company. To keep information safe, administrator enforces an Email Protection policy by selecting important keywords of the documents into a table stored in data base. Our EPS will automatically check all outgoing mails from the sales department and take action based on what the administrator has specified in the policy.
5. Proposed Algorithm
EPS checks the actual mail (including attachment) that is ready to leave the network, whether it is sending in our domain or outside of the network. After finishing, it checks the department, i.e. which department is sending an Email like Sales department or any other. Sales department creates a spreadsheet with name, credit card information, and social security no’s and prepares it monthly. Department attach this spreadsheet into an email and sends it to a business partner.

EPS opens the email consisting spreadsheet inside it which is created by the sales department. Protection System read it through, checks all of the cells of spreadsheet and apply the policy that administrator put in place of matching keywords. EPS not only scans the subject and copy but also scan the content inside it. EPS will match all the keywords of Spreadsheet with the table having important keywords of sensitive documents that are stored in company’s database by using PATTRENMATCHING ALGORITHM. After finding matching words, administrator can take appropriate action.

Figure 2. Email protection system deployed on gateway level
Algorithm: Pattern-Matching Algorithm

**Input:** T1, T2 (Pattern), Dept, and Mail  
//T2: Master Pattern Table  
//Dept: sensitive department like sales department of company  
**Output:** J, the Pattern match join table of T1 and T2

1. begin  
2. J = 0;  
3. If match(Mail.To, Company Domain) AND not match(Mail.dept, Dept) then  
4. Return J;  
5. end if  
6. If match((Mail.To, Company Domain) AND match(Mail.dept, Dept)) OR not match(Mail.To, Company Domain) then  
7. for all t1 ∈ Mail do  
8. for all t2 ∈ T2 do  
9. If match(t1, t2) then  
10. J = J Join (t1 ↔ t2);  
11. end if  
12. end for  
13. If J ≠ 0 then  
14. Apply Action ( );  
15. end if  
16. Return J;  
17. end for  
18. end if  
19. end

6. Necessary Actions for an E-Mail Protection

After finding anything malicious, administrator can take necessary actions which are as follows:

- **Block an email:** Once scanning is complete, extracted keywords of an outgoing mail are matched with the keywords stored in the database, the administrator can block an Email if found anything malicious.

- **Encrypt an Email:** To protect sensitive information, the administrator can encrypt an email by using some encryption algorithms.

- **Quarantined an Email:** On finding suspicious message or attachment, administrator has right to isolate it apart from the mailbox so that appropriate measures can be taken to read or retrieve it.

7. Conclusion

Data leakage is the root cause that damages the reputation of the company. Employees of the company leak the information and put assets at risk. This paper described how emails are moving through intermediate computers before reaching their final destination, where confidential data is located and then apply the right set of controls to that data. This paper explained threats to an email and why there is a necessity to protect emails? The big advantage to deploy Email Protection System on Gateway is to protect company's sensitive information from the hands of competitors.
Our EPS maintains the three principles Confidentiality, Integrity and Availability to protect emails from malicious intent which violates the company’s security. This Proposed System has the ability to block or quarantine inbound or outbound emails. We can’t imagine a company that does not use an email for communication and equally it is difficult to think about a company that would not get any benefit from an email security gateway. Future efforts can be made in implementing the procedure that can tackle the real world scenario.

8. References

[1] A. Shabtai, Y. Elovici and L. Rokach, "A Survey of Data Leakage Detection and Prevention Solutions," Springer Briefs in Computer Science, Springer, 2012.

[2] J. Kumar and A. K. Singh, "Dynamic resource scaling in cloud using neural network and black hole algorithm," 2016 Fifth International Conference on Eco-friendly Computing and Communication Systems (ICECCS), Bhopal, 2016, pp. 63-67.

[3] H. Taneja, Kapil and A. K. Singh, "Preserving Privacy of Patients based on Re-identification Risk," Fourth International Conference on Eco-friendly Computing and Communication Systems (ICECCS), 2015, pp. 448-454.

[4] I. Gupta, and A. K. Singh, "Privacy and Security Architecture for Cloud Data," technical report, NIT Kurukshetra, India, 2016.

[5] S. Chhabra and A. K. Singh, "Dynamic data leakage detection model based approach for MapReduce computational security in cloud," 2016 Fifth International Conference on Eco-friendly Computing and Communication Systems (ICECCS), Bhopal, 2016, pp. 13-19.

[6] I. Gupta, and A. K. Singh, "Data Leakage Detection in Cloud Environment," technical report, NIT Kurukshetra, India, 2017.

[7] S. Gill, G. Rupnar, V. Ramteke, D. Patil and V. M. Wadhai, "Email Security Protocol," International Journal of Computer Trends and Technology (IJCTT), Mar-Apr 2011.

[8] P. Zilberman, S. Dolev, G. Katz, Y. Elovici and A. Shabtai, "Analyzing group communication for preventing data leakage via email," Proceedings of 2011 IEEE International Conference on Intelligence and Security Informatics, Beijing, 2011, pp. 37-41.

[9] R. Tahboub and Y. Saleh, "Data Leakage/Loss Prevention Systems (DLP)," 2014 World Congress on Computer Applications and Information Systems (WCCAIS), Hammamet, 2014, pp. 1-6.

[10] S. Peneti and B. P. Rani, "Data leakage prevention system with time stamp," 2016 International Conference on Information Communication and Embedded Systems (ICICES), Chennai, 2016, pp. 1-4.

[11] M. A. Fayssel and S. S. Haque, "Towards Cyber Defense: Research in Intrusion Detection and Intrusion Prevention Systems," International Journal of Computer Science and Network Security (IJCSNS), Vol.10, No.7, July 2010, pp. 316-325.