JOURNAL OF SCIENCE
SAKARYA UNIVERSITY

Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University | http://www.saujs.sakarya.edu.tr/en/

Title: A DLP Module Design based on Plug-in for MS Word

Authors: Hussein AL-SANABANI, Murat İSKEFİYELİ
Recieved: 2019-12-08 23:15:20

Accepted: 2020-06-02 22:56:06

Article Type: Research Article
Volume: 24
Issue: 4
Month: August
Year: 2020
Pages: 770-781

How to cite
Hussein AL-SANABANI, Murat İSKEFİYELİ; (2020), A DLP Module Design based on Plug-in for MS Word. Sakarya University Journal of Science, 24(4), 770-781, DOI: https://doi.org/10.16984/saufenbilder.655984
Access link
http://www.saujs.sakarya.edu.tr/en/pub/issue/55932/655984

Sakarya University Journal of Science 24(4), 770-781, 2020

A DLP Module Design Based on Plug-in for MS Word

Hussein AL-SANABANI*[1], Murat İSKEFİYELİ[2]

Abstract

Inadvertent Data leakage by insiders is considered a serious problem for many organizations. Organizations are increasingly implementing Data Leakage/Loss Prevention solutions also know as (DLP), to protect the confidentiality of their data. Currently, DLP solutions have difficulties to identify confidential data as well as lack the ability to allow users to distinguish confidential from non-confidential data. Moreover, they are limited to work outside organizations. In order to solve this problem, it is important to introduce a DLP-Plugins model where the data owners can identify the privacy of the files during their entire lifecycle (creating, editing, etc.) by classifying them. This model uses security measures such as data encryption and access control to prevent accidental leakage of the classified files by the insiders. The proposed model guarantees that the right user will have access to the correct files according to their security access privilege inside or outside the organization. By always keeping classified files encrypted this will protect them all the time and everywhere. The DLP-Plugins model guarantees the usability for the users, all that will be required is to simply open and close the file as they do normally. As an example of the DLP-Plugins model, we have built a DLP-Plugin for Microsoft Word.

Keywords: Data loss prevention, data leakage prevention, encryption, access control, plugin

1. INTRODUCTION

The rate at which data in digital form enters and leaves organizations today is very high. On a daily basis, a typical enterprise can send and receive millions of email messages and downloads, via various channels an enterprise saves and transfers hundreds or even thousands of files [1].

Customers, business partners, regulators and shareholders expect enterprises to protect their sensitive data that they hold [1]. Leaked data can cause serious damage to an enterprise including but not limited to loss of customer loyalty and

* Corresponding Author: hussein.sanabani@ogr.sakarya.edu.tr
[1] Sakarya University, Department of Computer Engineering, Sakarya, Turkey.
ORCID: https://orcid.org/0000-0001-6580-4470
[2] Sakarya University, Department of Computer Engineering, Sakarya, Turkey.
ORCID: https://orcid.org/0000-0002-8210-5070. E-mail:miskef@sakarya.edu.tr

employee confidence which can lead to lawsuits, loss of competitive advantage, political crises, and company closure among others [2], [3]. Because of this, enterprise data is one of the most important assets an enterprise has; protection of this data must therefore be given the first priority [3].

In information security, Data leakage (or data loss) is referred to as unwanted exposure of information [4]. It is one of the most serious security issues that intentionally or unintentionally expose private or sensitive data to an unauthorized entity [2], [5].

Symantec reported that in 2014 there were 1523 data breaches in total compared to 1211 in 2015 and 1209 in 2015. However, the total number of identities exposed in these breaches was 1.2 billion in 2014, 564 million in 2015 and 1.1 billion in 2016 [6]. According to Report by the Global State of Information Security Survey 2015 shows that in 2014, security breaches reached 42.8 million rising by 48% from 2013 [7]. The most-cited culprits of incidents were the employees. More so in 2016 and 2017, about 30% of breaches were reported to have come from current employees while in 2016 and 2017 about 28% and 26% of breaches respectively were reported to have come from former employees [8]. As opposed to outside crime, 32% of the respondents mentioned insider crime as the most damaging and costly breaches in the 2014 US State of Cybercrime Survey [9].

This clearly indicates the extent of the data leakage problem in all kinds of organizations and thus has to be solved. The first step towards solving is by organizations understanding what confidential data is held in the organization, how this data is managed, and how to protect it from unauthorized access [1]. As a result, various data loss prevention (DLP) solutions have been developed to cope with this problem.

According to [5] a DLP solution is defined as : "a system that is designed to detect and prevent the unauthorized access, use, or transmission of confidential information". DLP detect and prevent unauthorized access to confidential data by using deep analysis for both content and surrounding context around confidential data [10]. [11] also defines DLP solutions as "Products that, based on central policies, identify, monitor, and protect data at rest, in motion, and in use, through deep content analysis".

DLP solutions help to identify, monitor, protect and reduce the risks of confidential data leakage. DLP solutions are used not only to detect and deter an unauthorized user from getting access to confidential data but also to protect confidential data from being inadvertently shared [3]. From the outside lots of different technologies such as Intrusion Detection and Prevention systems, anti-malware and firewalls exist to protect data, DLP systems are designed to do the job from within [3].

According to [2], [12] DLP solutions can be classified as active DLP solutions or inactive DLP solutions. Active DLP solution tracks the confidential data while a user is typing instead of parsing a file after it has been created and saved. While inactive (content-aware and/or context-aware) DLP solution perform both content inspection and/or contextual analysis to identify sensitive data while at-rest, in-use, or in-motion to prevent it from leakage, so it needs to parse all file types.

In [5] DLP solutions can handle data Leakage using two main modes:

Detective modes: The DLP solutions try to detect leakage occurrences and take the suitable corrective action to handle any leakage occurrence that happened by using Context-based inspection, Content-based inspection and Content tagging to identify the sensitive data.

Preventive modes: The DLP solutions prevent leakage before they occur by using several preventive approaches like Access control, Disabling functions, Encryption and Awareness.

The problem: Most of the current DLP solutions have difficult to identify the sensitive data using Content-based (regular expressions, statistical algorithms, pattern matching, keyword comparison, document fingerprint, etc.) or Context-based (file type/size, sender,

header/metadata information, source, destination, etc.) inspection because it is very difficult to identify the sensitive data automatically and also all these techniques need the capability to parse various file types. In addition, they are usually not able to identify the confidential data if they are compressed, encrypted, or obfuscated [3], [13]. Moreover, Most of the DLP solutions don't have the ability to allow insiders to identify the sensitive data as they are the creators of it and the most familiar to what it is containing. They also restrict users to work inside the organization's borders.

The solution: To cope with this problem, this paper introduces DLP-Plugins model that will be added to the legacy applications like Microsoft Office, pdf readers, text editor, media players, and mail applications. This model will let the data owners to identify the sensitive data by classifying it when it was created or edited according to their security level (their Privilege) that they have. The classified data will be protected by using two preventive approaches (access control and encryption). As a result, DLP-Plugins model will guarantee that the sensitive data will be standing protected all the time and everywhere at rest, in motion, and in use.

The contribution: According to the best of our knowledge DLP-Plugins model is the first DLP solution that use Plugins model to protect data against inadvertent data leakage. It doesn't need effort to parse all the files types, but on the other hand, it requires to build Plugin for the wanted application. This model also allows the data owners to identify the confidentiality of data as they are the most knowledgeable on what the data contains. It doesn't matter whether the file is encrypted, compressed, or obfuscated because the confidential data will be protected all the time by encryption. However, when the classified (protected) data need to be modified (edited), a plaintext data will be only available for the authorized users inside the authorized computers inside or outside the organization. This will provide more flexibility to expand organization workplace while guaranteeing the security of the data and without putting a burden on the user. To prove this, we have built a DLP-Plugin for Microsoft Word as an instant of this DLP-Plugins model.

The rest of this paper is structured as follows. In Section 2 we review the related work. Section 3 presents the motivation for our approach (DLP-Plugins model), then Section 4 describes the proposed model. At last, Section 5 gives conclusions and proposes future directions.

2. RELATED WORK

Papers Of late surveys [5] have shown much interest in research concerning data loss prevention. Data leakage was addressed by current DLP solutions according to what, where and how to protect. What to protect focuses on data-at-rest (DAR), data-in-motion (DIM) and data-in-use (DIU). Where to protect concerns Endpoint and Network. While how to protect describes the leakage handling approaches [5].

A technique for data leak protection (DLP) based on monitoring confidential information as it travels inside a file system on a computer system is described in [14]. The basis of this method is the idea that every travel from confidential to non-confidential item increases the security level of the destination item to that of the source item. Therefore the system can identify the confidential information by spreading labels over all confidential items to avoid hidden passages for information leakage. [4], [5], [15], [16], [17] mention about content tagging. This technique is used to tag the file holding confidential data to identify it, and the enterprise policies will be enforced based on the assigned tag. Once the tag is assigned, the tag stays with the content as it is moved or copied or included in or attached to other files or file types even with the most extreme modification of content, like changing format, encrypting and compressing. Tags can be allocated to files in two methods: manually by the author of the confidential data or automatically by the DLP solution. This technique can identify the file but not the contained confidential data [4].

The content tagging technique is the nearest to our methodology, because it also considers the classification (tagging) of confidential data. What

distinguishes our work is that in our work the data owners identify the sensitive data by classifying it at the time of creation or during modification. And this classified data will be protected by encrypting it from the outset and always remain encrypted.

An active DLP model proposed by [2] can track the sensitive data while the user is typing as opposed to parsing a file after it has been created and saved. [18], [19] discusses Data Leakage Detection using Image and Audio Files. The Goal of this system is to find which data of the distributor (owner of data) has been leaked and if leaked, detects the agent (trusted party) who leaked that data. Basing on data-driven usage control concept, Data loss prevention solution for Microsoft Windows operating systems to allow fine-grained policy-based protection is presented in [13].

In [20] Microsoft has built a DLP solution for Office 365 this DLP solution use context-based inspection, content-based inspection and content tagging (label) to identify the sensitive data and take an action according to the predefined policy. However, our DLP-Plugins model can work together with this solution in perfect harmony.

In general, our work can be distinguished from previous work in that our DLP solution is not an independent solution. It is a plugins model directed to protect certain applications files. This plugin will prevent the potential inadvertent leakage incidents before they occur by taking proper preventive measures such as data encryption and access control in order to put the most effort into preventing potential inadvertent disclosure in the first place. Moreover, DLP-Plugins model relies on the role of the human factor for distinguishing the sensitive data. Finally, this model only focused on unintentional (accidental) data leakage.

3. MOTIVATION

The CIA triad of information security means confidentiality, integrity and availability [21]. We limit our discussion on confidentiality. By definition, Confidentiality of information is typically assurance that sensitive information is accessed only by authorized users [22]. This task can be achieved by various mechanisms such as device control, encryption and access control [5]. To secure voluminous amounts of personal data from malicious insider and outsider attacks these simplest measures have been used [5]. However, according to [4] the easiest way to deter data leakage is by using DLP solution that relies on security policy and access rights (access control) because they have been in use long enough and follows well established foundations. All of this motivates us to focus on using encryption and access control in DLP-Plugins model to protect the confidentiality of the data.

In [23], [24] Lior Arbel, director of strategic data security solutions at Websens said that “data categorization is one of the key ways that DLP solutions use to determine which data needs heightened levels of security and what does not. Furthermore, in order to protect data they would need to classify the data first and then run discovery on that data”. Based on this our DLP-Plugins model relies on classifying the data from the outset when it was created, or at any point of its lifecycle.

Christian Toon, head of information risk at Iron Mountain emphasizes that to achieve an effective DLP implementation, the human element should not be ignored: “Data loss prevention technologies are only as good as the employees using them” [24]. For that, we assigned the classification process of the sensitive data to the authorized users (data owners). However, an organization can attempt to force its employees to comply with its regulations by using control mechanisms, surveillance, and monitoring. But this approach has proven to be ineffective in several cases such as the incident of Edward Snowden [25]. Organizations can also rely on the acceptance and the cooperation of their employees because there is no never-known-to-fail method to prevent data leakage [25]. This means that intentional data leakage is harder to prevent, for that reason our method is focused only on unintentional (accidental) data leakage.

4. THE PROPOSED MODEL

In this section, we describe the classification method used by the DLP-Plugins model, the DLP-Plugins design components, how the model works and its implementation in Microsoft Word and finally we show the performance of the model.

4.1. Classification method

In any organization, corporate institution or Government department, data and information are classified according to some criteria ranging from highly to less confidential data. This criterion will define who is permitted to use and distribute the data [15]. Accordingly [26]–[28] the most common example that is used by governments and organizations is :Top Secret, Secret, Confidential and Restricted. However most governments and organizations have their own rules to state the security levels, determine the level for the data and who has permission to handle this level. Consequently, it is well known that the mostly used organizational structure type by both companies and governments is hierarchy (pyramidal). Based on this, we use hierarchy classification strategy as shown in Figure 1.

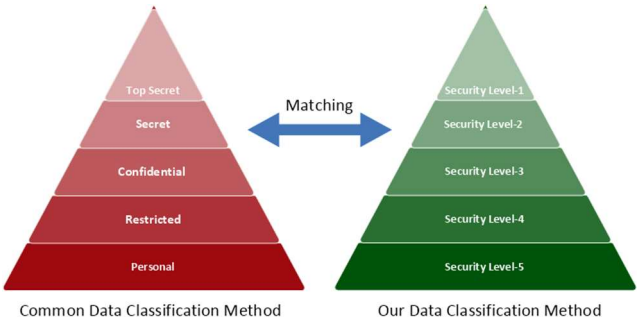

Figure 1. Hierarchy Classification Strategy

We classified the Data into five levels of security sequenced from Top Secret to Personal. The user with security Level-1 (Top Secret) has all privileges and can access to all other security levels. The user that has Level-2 can have access at this level and all levels that are bigger than Level-2, i.e. Level-3, Level-4 and Level-5. Also, the user that has Level-3 can have access at this level and all levels bigger than three, this means that he can have access to the Level-3, Level-4, and Level-5. The same for the Level-4 and Level-5 security levels. However, the security Level-5 (Personal) is different for each user i.e. each user can see only his/her own data that classified as Level-5 but he/she cannot see the security Level-5 for the other users because it contains their personal information. This proposed classification model can be adapted or changed to fit organization's and government's requirements.

4.2. DLP-Plugins components

The DLP-Plugins model that we build has three main Components:

4.2.1. DLP-Administrator Panel

DLP-Administrator Panel is a website that provides a central control mechanism to manage users. For example (add new user or delete existing user, change the privilege of existing user in a specific computer, etc.) and for adding or deleting a new DLP-Plugin.

4.2.2. DLP-Web Service

DLP-Web Service is a web service used to check both user and computer identity to know the user security level that he/she has, then send appropriate encryption and decryption keys through the encrypted channel to DLP-Plugin according to the security level that the user has.

4.2.3. DLP-Plugin

DLP-Plugin is the plugin that will be added to the legacy application to protect their data by decrypting the files when opening them and encrypting them when closing them by the authorized users. This DLP-Plugin will communicate with DLP-Web Service to check the user identity and to get his/her security level (appropriate encryption and decryption keys). Also, it will let the users classify their data or change its security level according to the security level that they have.

The Figure 2 shows that the first and second components will be fixed to all the DLP-Plugins that will be added to the legacy applications but the third components will be varying according to

the diversity of the applications that will be added to them. As an example of these DLP-Plugins, we have developed one DLP-Plugin for Microsoft word.

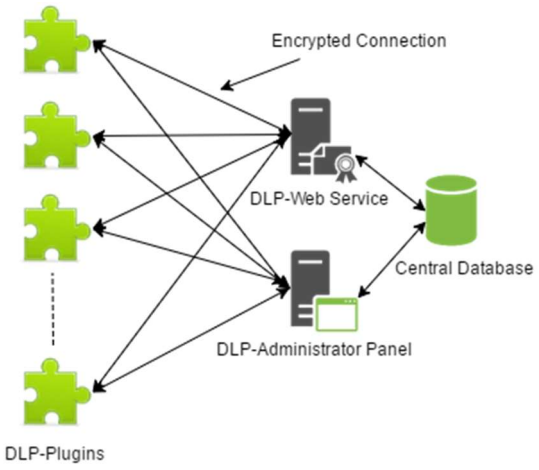

Figure 2. DLP-Plugins components

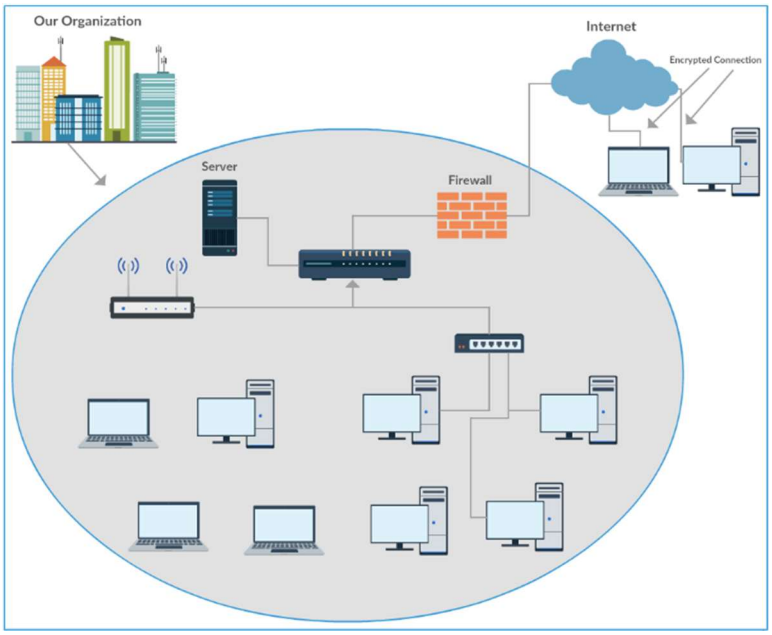

Figure 3. An example of a network topology of an organization

To achieve this we build DLP-Plugin for Microsoft Word application to protect Word documents from accidental leakage by internal workers or to have access by unauthorized users. This Plugin is an add-in feature for Microsoft Word that allows the DLP functions. It works with all Windows desktop versions of Microsoft Word (2010, 2013,2016, 365) except the versions that are prior to Word 2010. To understand how the model works we divided it into three processes:

4.3. How the DLP-Plugins model work (ex: Microsoft Word)?

Assume we have organization that has a network topology like this in the Figure 3:

Suppose we want to protect the organization Word documents from unintentional leakage to the outside of the organization. And also we want to protect these documents from having illegal access by unauthorized insiders but at the same time, we want to provide legal access while guaranteeing the usability for the insiders from inside or outside the organization.

4.3.1. Word application startup process

When a word application is started the DLP-Plugin inside the word application will start to recognize the computer and user identity. After getting the identity, the DLP-Plugin will connect to the DLP-Web Service to get his security level

(encryption and decryption keys) according to his identity. Now that the computer has its privileges (security level) it means it has appropriate encryption and decryption keys.

Figure 4 below depicts how the computer request his security level by sending his identity to the DLP-Web Service and getting his security level (encryption and decryption keys) through the encrypted channel.

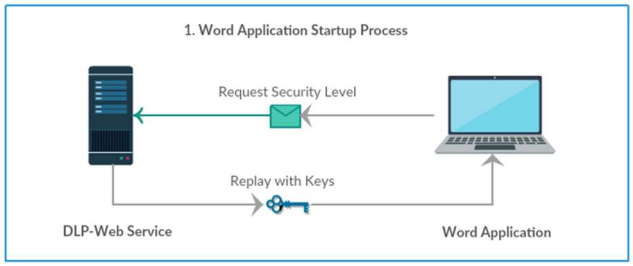

Figure 4. Word application startup process

4.3.2. Word document opening process

Every classified document is by default encrypted. From the first process described above, the user acquires his lawful decryption and encryption keys. When the user starts to open this classified document the DLP-Plugin will check the security level of the document and if this security level is legal to that user it means the DLP-Plugin has appropriate key to decrypt this document. So the DLP-Plugin will decrypt and open the document for that user. If the document security level is not legal to that user the document will remain encrypted because DLP-Plugin doesn't have appropriate key to decrypt this document. This will happen quite easily without any additional burden to the user. The user just double clicks on the document that he/she wants to open and the decryption process will be done in the background without feeling any different. This process is described in the Figure 5 below:

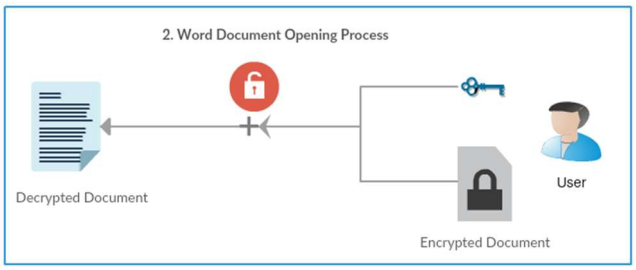

Figure 5. Word document opening process

4.3.3. Word document closing process

While the document is open, the user can change the security level of the document to any level that he has. When closing the document, the DLP-Plugin will encrypt the document according to its security level and its appropriate key. This is described in the Figure 6 below:

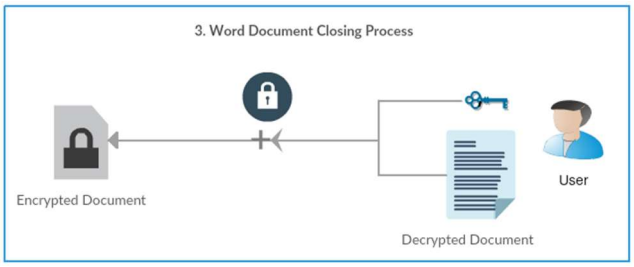

Figure 6. Word document closing process

These three processes provide a compatible protection for the sensitive data from inadvertently leaking to the outside of the organization and also preventing access by unauthorized users.

The pseudocode for these processes is shown in Algorithm 1.

Algorithm 1 Word application processes

```
Input: (computerID, userID)
   // Requst SecurityLevel From DLP-Web Service
   userSecurtiyLevel := RequstSecurityLevel(computerID,userID)
   InitializeUserInterfaceWithUserSecurtiyLevel(userSecurtiyLevel)
   While ture
      ListenToUserOpen_CloseDocumentEvent(event)
         If event.Type = OPEN_Document
            docSecurtiyLevel := GetDocumentSecurtiyLevel(event.Doc)
            if  CheckIslegalForDoc(userSecurtiyLevel,  docSecurtiyLevel)
            = true
               doc := DecryptDoc (event.Doc, userSecurtiyLevel)
               ShowDocment(doc)
            Otherwise
               ShowDocment(event.Doc) // Will show encrypted data
            end if
         end if
         if event.Type = Close_Document
```

```
        docSecurtiyLevel := GetDocumentSecurtiyLevel(event.Doc)
        if docSecurtiyLevel ≠ null
            encryptedDoc:= encryptDoc(event.Doc, userSecurtiyLevel)
            SaveAndCloseDocument(encryptedDoc)
        Otherwise
            CloseDocument(event.Doc)
        end if
    end if
end while
```

Figure 7 shows how the document for both the authorized user and the unauthorized user will appear. The document in the figure was classified as Level-1. The user on the left side of the figure is authorized because he has this level so the document appears to him normally. But the user on the right side of the figure doesn't have Level-1 so he will see just encrypted data.

4.4. DLP-Plugins model Performance (ex: Microsoft Word)

The execution results are taken on a machine having Intel Core i3 (2.67 GHz) processor with 4 GB RAM and Windows 10 64-bit operating System and Microsoft Word 2013 32-bit. The C# 4.5 .NET Framework platform is used for implementation. The .NET Framework build-in Cryptography Dynamic-link library (DLL) for (AES, RC2, TDES) and Bouncy Castle Cryptography DLL for (RC4, Blowfish, Twofish) are used for encryption algorithm implementation. We do the three tests for each process described before in section 4.3. and all tests were run 4 times and their average was calculated, see tables for each test.

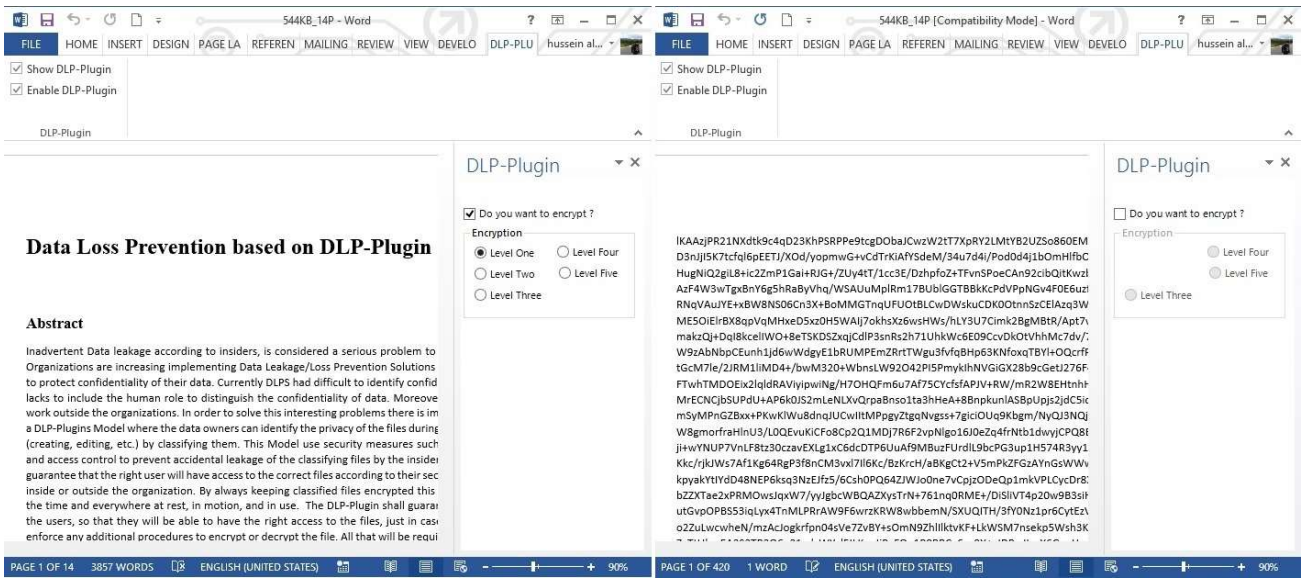

Figure 7. The appearances of the document for both authorized and unauthorized users

4.4.1. Frist test for startup the Word application

The time to get the computer identity and connect to the DLP-Web Service varies from 0.01 second to 2.0 seconds depend on the internet connection.

4.4.2. Second test for opening document

Table 1 shows the overall time of opening and decrypting the document for different file sizes and number of pages using different encryption algorithms. The overall time is a summation of three times. The first one is the time of reading the encrypted data from the document. The second one is the time of decrypting the ciphertext to get the plaintext (XML string). The third one is the time of parsing the XML string to make the document and open that document. The key size of AES, RC4, RC2, Blowfish and Twofish is 128 bits and the key size of TDES is 112 bits.

Table 1. The overall time of opening and decrypting the document for different file sizes and number of pages using different encryption algorithms

File Size	Number of Pages	AES-128	RC4-128	RC2-128	TDES-112	Blowfish-128	Twofish-128
3.45MB	9	2.695263	2.802374	2.981562	3.652868	4.309394	3.924538
544KB	14	1.204576	1.297097	1.469328	1.606416	1.703509	1.559582
847KB	50	2.050756	2.08891	2.244023	2.535857	2.759254	2.794405
1.01MB	70	2.436801	2.433337	2.696202	2.995279	3.413842	3.408304
1.60MB	100	3.377409	3.309007	3.481061	3.926474	4.511645	4.614573
1.70MB	150	4.194024	4.068522	4.567168	5.437092	5.668988	5.676459
1.77MB	200	4.501686	5.031383	5.319095	6.192539	6.639504	7.018375
2MB	300	6.088659	6.856815	7.222979	8.49772	9.526075	9.267652
Average	111.625	3.318647	3.485931	3.747677	4.355531	4.816526	4.782986

Figure 8 summarizes the results presented in Table 3 for opening document test.

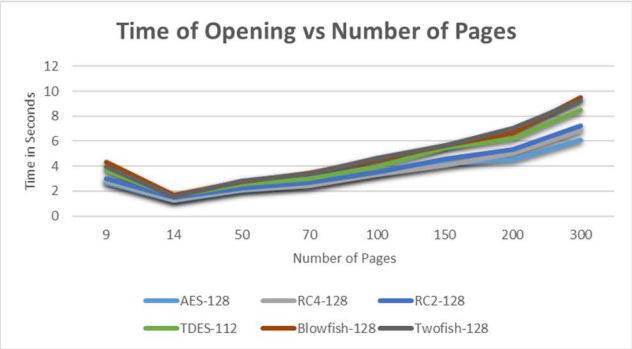

Figure 8. Time of opening vs number of pages

Observation: The results reveal that the time of opening the encrypted document depends on file size and number of pages. Increasing the number of pages or file size will increase the opening time of the document. It can be clearly seen that AES performs better followed by RC4 with key size 128 bits among the encryption algorithms tested.

4.4.3. Third test for closing document

Table 2 shows the overall time of closing and encrypting the document for different file sizes and number of pages using different encryption algorithms. The overall time is a summation of three times. The first one is the time of reading the XML string that represents the document. The second one is the time of encrypting the plaintext (XML string) to get ciphertext. The third one is the time of saving and closing the document. The key size of AES, RC4, RC2, Blowfish and Twofish is 128 bits and the key size of TDES is 112 bits.

Table 2. The overall time of closing and encrypting the document for different file sizes and number of pages using different encryption algorithms.

File Size	Number of Pages	AES-128	RC4-128	RC2-128	TDES-112	Blowfish-128	Twofish-128
3.45MB	9	2.690997	2.70322	2.564246	3.130119	3.694704	3.538402
544KB	14	0.608243	0.665121	0.771921	0.785535	0.995787	0.970647
847KB	50	1.319755	1.471099	1.49678	1.571272	1.724479	1.939608
1.01MB	70	1.520164	1.730294	1.644743	1.920754	2.057411	2.136468
1.60MB	100	2.263135	2.42154	2.467695	2.697342	3.111858	2.990853
1.70MB	150	2.922747	2.90495	3.402629	3.528995	4.196691	4.167968
1.77MB	200	2.996991	3.407393	3.750937	3.88995	4.706226	4.964365
2MB	300	3.739411	4.003328	4.3916	4.844758	5.351529	5.30996
Average	111.625	2.25768	2.413368	2.561319	2.796091	3.229835	3.252284

Figure 9 summarizes the results presented in Table 2 for closing document.

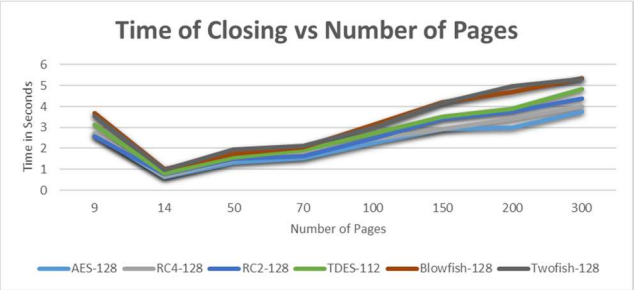

Figure 9. Time of closing vs number of pages

Observation: The result shows that the time of closing and encrypting the document depends on file size and number of pages. Increasing the number of pages or file size will increase the closing time of the document. As with opening the document, AES performs better followed by RC4 with key size 128 bits among the encryption algorithms tested.

4.5. Discussion

For We can see that the time for closing is smaller than the time for opening this shows that the major overhead is a result of parsing the XML to make the word document. In general, we can ignore the time required to close the file because the closing processing is run in the background while the Word application can do other processes. The second and the third tests revealed that the time for opening and closing document is proportional to the number of pages and file size. As the number of pages or the file size increases, the time for opening and closing the document also increases proportionally to a number of pages or file size and vice versa. Further the results show that the AES and RC4 encryption algorithms are the fastest and suitable among the other algorithms.

5. CONCLUSIONS AND FUTURE WORK

This paper introduces a DLP-Plugins model that use two preventive approaches: Access control and Encryption to prevent the probable unintentional data leakage incidents before they occur. This DLP-Plugins model let the authorized insiders (data owners) identify the sensitive data by classifying them during its entire lifecycle rather than scanning both content and context of the data as most of the commercial DLP solutions do because it easily avoids the imposed high overhead of scanning. DLP-Plugins model provides more flexibility to work outside of the organization and guarantees the usability for the users that they have the right to access the classified documents easily, just open and close the document as they do normally.

We also realize that in our approach it is too easy for authorized users to intentionally leak data. But as it's known that intentional data leakage is impossible to prevent it and that is the problem that all the current DLP solutions faced, so the organizations should rely on the acceptance and the cooperation of their employees. Consequently, our DLP-Plugins model trusts on the approval and the cooperation of the employees and focused only on the unintentional leaking of data. However DLP-Plugins model can work together with current DLP solutions in perfect harmony.

The implementation of DLP-Plugin for Microsoft Word and the performance results show that the proposed DLP-Plugins model is feasible, easy to use and practical using current technologies. The result revealed that the AES and RC4 Encryption algorithms perform better and thus most suitable among the other algorithms.

As a future work, we suggest developing DLP-Plugins for protecting all types of office documents (Excel, PowerPoint, Access, etc.), for E-mail Program to protect all E-mail message, for PDF reader to Protect PDF file and also for images and videos player to protect all images and videos files. Also, we suggest integrating those DLP-Plugins with the current DLP solutions in the market.

Research and Publication Ethics

This paper has been prepared within the scope of international research and publication ethics.

Ethics Committee Approval

This paper does not require any ethics committee permission or special permission.

Conflict of Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this paper.

REFERENCES

[1] S. Liu and R. Kuhn, "Data loss prevention," IT Prof., vol. 12, no. 2, pp. 10–13, Mar. 2010.

[2] J.-S. Wu, Y.-J. Lee, S.-K. Chong, C.-T. Lin, and J.-L. Hsu, "Key Stroke Profiling for Data Loss Prevention," in 2013 Conference on Technologies and Applications of Artificial Intelligence, pp. 7–12, 2013.

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

[4] S. Alneyadi, E. Sithirasenan, and V. Muthukkumarasamy, "A survey on data leakage prevention systems," J. Netw. Comput. Appl., vol. 62, pp. 137–152, Feb. 2016.

[5] A. Shabtai, Y. Elovici, and L. Rokach, A Survey of Data Leakage Detection and Prevention Solutions. Boston, MA: Springer US, 2012.

[6] Symantec, "Internet Security Threat Report," vol. 22, no. April, 2017.

[7] PwC, "Managing cyber risks in an interconnected world: Key fndings from The Global State of Information Security Survey 2015," 2014.

[8] PwC, "The Global State of Information Security Survey 2018," pwc.com, 2018. [Online]. Available: https://www.pwc.com/us/en/services/consulting/cybersecurity/library/information-security-survey.html. [Accessed: 09-May-2020].

[9] N. Mickelberg, Kevin; Schive, Laurie; Pollard, "US cybercrime: Rising risks, reduced readiness Key fndings from the 2014 US State of Cybercrime Survey," 2014.

[10] S. Alneyadi, E. Sithirasenan, and V. Muthukkumarasamy, "Detecting Data Semantic: A Data Leakage Prevention Approach," in 2015 IEEE Trustcom/BigDataSE/ISPA, vol. 1, pp. 910–917, 2015.

[11] R. Mogull and M. Rothman, "Understanding and Selecting a Data Loss Prevention Solution," 2017.

[12] B. Reed and N. Wynne, "Magic Quadrant for Content-Aware Data Loss Prevention," 2016.

[13] T. Wuchner and A. Pretschner, "Data Loss Prevention Based on Data-Driven Usage Control," in 2012 IEEE 23rd International Symposium on Software Reliability Engineering, pp. 151–160, 2012.

[14] M. Petkovic, M. Popovic, I. Basicevic, and D. Saric, "A Host Based Method for Data Leak Protection by Tracking Sensitive Data Flow," in 2012 IEEE 19th International Conference and Workshops on Engineering of Computer-Based Systems, pp. 267–274, 2012.

[15] M. H. Matthee, "Tagging Data to Prevent Data Leakage (Forming Content Repositories)," SANS Inst., pp. 1–26, 2016.

[16] I. McAfee, "McAfee Host Data Loss Prevention 2.2.1 Product Guide." McAfee, Inc, pp. 1–80, 2008.

[17] I. McAfee, “McAfee Data Loss Prevention 11.0.300 Product Guide.” McAfee, Inc, pp. 1–211, 2020.

[18] S. S. Dandavate, P.P.; Dhotre, “Data Leakage Detection using Image and Audio Files,” Int. J. Comput. Appl., vol. 115, no. 8, pp. 1–4, 2015.

[19] P. S. V. Kale, Sandip A.; Kulkarni, “Data Leakage Detection,” Int. J. Adv. Res. Comput. Commun. Eng., vol. 1, no. 9, pp. 668–678, 2012.

[20] Microsoft, “Overview of data loss prevention,” microsoft.com, 2019. [Online]. Available: https://docs.microsoft.com/en-us/microsoft-365/compliance/data-loss-prevention-policies?view=o365-worldwide. [Accessed: 09-May-2020].

[21] J. Andress, “What is Information Security?,” in The Basics of Information Security, Elsevier, pp. 1–22, 2014.

[22] B. Guttman and E. Roback, An Introduction to Computer Security : The NIST Handbook, vol. SP800, no. 12. 1995.

[23] L. Arbel, “Data loss prevention: the business case,” Comput. Fraud Secur., vol. 2015, no. 5, pp. 13–16, May 2015.

[24] T. Caldwell, “Data loss prevention – not yet a cure,” Comput. Fraud Secur., vol. 2011, no. 9, pp. 5–9, Sep. 2011.

[25] B. Hauer, “Data and Information Leakage Prevention Within the Scope of Information Security,” IEEE Access, vol. 3, pp. 2554–2565, 2015.

[26] M. Diri, Mustafa; Gülçiçek, “Türkiye’de Kamu Hizmetinin Görülmesinde Kullanılmakta Olan Gizlilik Derecesi Tanımları: Uygulamadaki Sorunlar ve Çözüm Önerileri,” Maliye Derg., vol. 162y, pp. 497–537, 2012.

[27] Office Cabinet UK, Government Security Classifications April 2014, pp. 1–35, 2013.

[28] Office Cabinet UK, International Classified Exchanges, no. 1–23. 2015.