Cluster Computing
https://doi.org/10.1007/s10586-022-03668-2

Survey of Techniques on Data Leakage Protection and Methods to address the Insider threat

Isabel Herrera Montano[1] · José Javier García Aranda[2] · Juan Ramos Diaz[2] · Sergio Molina Cardín[2] · Isabel de la Torre Díez[1] · Joel J. P. C. Rodrigues[3,4]

Received: 7 February 2022 / Revised: 25 April 2022 / Accepted: 21 June 2022

Abstract
Data leakage is a problem that companies and organizations face every day around the world. Mainly the data leak caused by the internal threat posed by authorized personnel to manipulate confidential information. The main objective of this work is to survey the literature to detect the existing techniques to protect against data leakage and to identify the methods used to address the insider threat. For this, a literature review of scientific databases was carried out in the period from 2011 to 2022, which resulted in 42 relevant papers. It was obtained that from 2017 to date, 60% of the studies found are concentrated and that 90% come from conferences and publications in journals. Significant advances were detected in protection systems against data leakage with the incorporation of new techniques and technologies, such as machine learning, blockchain, and digital rights management policies. In 40% of the relevant studies, significant interest was shown in avoiding internal threats. The most used techniques in the analyzed DLP tools were encryption and machine learning.

Keywords Data leak Protection · Data leak Prevention · DLP · Internal threat · Classified Information Security · DRM

1 Introduction

In terms of information security, insider threat refers to the risk posed by an organization's employees, partners, or customers to the organization's information [1]. Data leakage is the disclosure of information to unauthorized entities or individuals [2], commonly caused by an intentionally or unintentional threat to the insider [3], [4], [5]. Data leakage protection (DLP) systems or DLPS are designed primarily to monitor data flow in an organization and apply predefined measures on terminal devices or networks within the organization [2]. The measures range from logging activities, sending alerts to end users and administrators, to quarantining data or blocking it altogether. DLP tools can monitor data at rest and in motion to detect sensitive information [3], [6].

In both corporate and hospital environments, the security of classified information is vital, the cost to companies of the lack of DLP technologies is estimated at over $200 per employee per year, and the human factor accounts for 35% of the causes of security breaches, including malicious and unintentional activities of both employees and third parties [7]. Not all sectors are equally affected by the costs of data leakage, the most sensitive being the healthcare and

✉ Isabel Herrera Montano
isabel.herrera.montano@uva.es

José Javier García Aranda
jose_javier.garcia_aranda@nokia.com

Juan Ramos Diaz
juan.ramos_diaz.ext@nokia.com

Sergio Molina Cardín
sergio.molina_cardin.ext@nokia.com

Isabel de la Torre Díez
isator@tel.uva.es

Joel J. P. C. Rodrigues
joeljr@ieee.org

[1] Department of Signal Theory and Communications and Telematics Engineering, University of Valladolid, Paseo de Belén, 15, 47011 Valladolid, Spain

[2] Department of Innovation, Nokia, Maria Tubau Street, 9, 28050 Madrid, Spain

[3] College of Computer Science and Technology, China University of Petroleum (East China), 266555 Qingdao, China

[4] Instituto de Telecomunicações, 6201-001 Covilhã, Portugal

Published online: 14 July 2022

banking sectors due to the large volume of personal data they both handle [8]. The Spanish report [9] shows several aspects that give rise to data leakage in the healthcare sector, with malicious insider threats and unintentional employee actions being evident. Motivated by all of the above, the main objective of this work is to survey the literature to detect existing techniques to protect data leakage, and to identify the methods used to address the insider threat.

Studies similar to this focus on reviewing the functions of DRM products popular in 2011 and available on the market, quantitatively evaluating the impact of the use of these products [10]; analyzing the existing digital forensics and incident management literature with the aim of contributing to the knowledge gaps in incident management in the cloud environment [11]; outline lines of research based on a systematic review focused on blockchain technology applied to eHealth [12]; examining the state of the art in security, privacy, and big data protection research [13]; in [14] a survey about sensitive data leakage prevention and anti-theft technologies for protecting the information security of e-government users; and in [15] study monitoring strategies for confidential documents based on virtual file system (VFS), in [16] a systematic review of the literature focused on management functions in information security is carried out. The recent studio [17] presents a review focused on the mobile agent model for data leakage prevention. The review only considered papers published in the journal "Communications and Network" and conference papers published between 2009 and 2019. Mobile agent-based distributed intrusion prevention and detection systems were analyzed in terms of their design, capabilities, and shortcomings. Other studies focus on reviewing blockchain strategies for secure and shareable computing, examining the state of blockchain security in the literature, from the point of view of information system security issues, classified into three levels: process level, data level, and infrastructure level [18], survey the literature to analyze how blockchain systems can overcome potential cybersecurity barriers to achieve intelligence in Industry 4.0 [19].

Research on data protection has increased with the introduction of telecommuting due to the pandemic and the need to move data to external devices and networks. Similar work has been found to exist in reviews related to data protection, but it is worth noting that there is no recent study focused on grouping the work developed in the last ten years on DLP tools, where special attention is given to the techniques used in DLP tools and methods to combat the insider threat. The main contributions of this article are the following: (1) it highlights the most used techniques in DLP tools, (2) it summarizes the methods found in the literature to face the insider threat, with the aim of promoting the transformation of protection against data leaks in this sense, to make it more secure, and (3) exposes the limitations, advances, and applications of DLPS, in order to encourage the development of new tools.

This paper addresses the following research questions:

RQ1. What techniques or technologies are used as DLP tools?
It is solved in Sections 2 and 5, giving a presentation of the main tools found in Section 2 and an analysis of their frequency of use in relevant studies in Section 5.
RQ2. How is the insider threat addressed in the DLP tools found in the literature?
The answer to this question is presented in section 3, which summarizes how insider threat is addressed in the literature analyzed.
RQ3. What are the highlights the most used techniques in DLP tools, limitations, advances, and applications of DLPS in different fields, in order to encourage the development of new tools, and 2) it exposes the methods found in the literature to face the insider threat, with the aim of promoting the transformation of data leakage protection in this sense, to make it more secure advances and applications of DLP systems?
This question is answered in Section 5.3, where the main advances and applications of DLP systems in the period studied are presented.

The rest of the document is organized as follows: Sect. 2 describes the main techniques and technologies used in DLP. Section 3 presents the methods to address insider threats found in the literature and Sect. 4 describes the methodology followed for the literature review. Section 5 discusses the results obtained and the main limitations, advances, and applications of DLPS. Finally, this article is concluded, and future work is presented.

2 Techniques and Technologies in DLP

Several studies propose novel DLPS integrated by different techniques and technologies to try to ensure optimal protection of confidential information, this section gives an overview of the most used techniques and technologies in the papers relevant to this study.

2.1 Overview of techniques most commonly used in DLPS

2.1.1 Inteligents documents

This technique consists of encapsulating within the document both the data it contains and the security mechanisms

to control the use of such data [20], [21]. The security mechanisms can be content deletion, content editing, content reading, or an authorized user to perform each operation. This technique makes it possible to record where, when, and by whom the content of the document is accessed [22]. It is a technique generally used in DRM systems and very useful in combination with DLPS.

2.1.2 Encryption

The most widely used technique in DLPS is cryptography, this is because it is the main basis of security and is based on the conversion of data from a readable format to an encrypted format. Any encryption algorithm is equivalent to a mutating substitution algorithm, the substitution unit being the concept of "block", and the substitution table being something nonfixed (and therefore mutating). The robustness of the algorithm is given by the mutability, which prevents statistical attacks [3].

2.1.3 Hash

A widely used DLPS approach is exact file hash matching. This method is based on the verification of outbound traffic by comparing the hash values of the intercepted traffic and existing sensitive data [2]. If a match is detected between the values, a leak is detected by the system. This approach presents the problem that any modification of the original document may result in a completely different hash value, which would not allow the system to detect the confidential document [20].

2.1.4 Virtual file system (VFS)

A VSF is an abstraction layer on top of a real file system (RFS), that is, an intermediate layer between system calls and the RFS driver [15]. They also provide the ability to perform operations before and after reading, writing, etc. In exchange for this intermediate "translation" between the applications and the actual file system, some of the original RFS performance is lost.

2.1.5 Challenges or context-based keys

Challenges replace a stored key with a calculated key, eliminating the security problem in key storage and distribution [3], [21], [22], in turn, allowing the user to be identified through biometric data, the location of the computer by nearby Wi-Fi signals or GPS, among other benefits that this technique allows.

2.1.6 Minifilters

Minifilters are low-level applications that run in Windows kernel mode and perform value-added functions (backup, encryption, monitoring, etc.) on filesystem operations (read, write, metadata modification, etc.) [23], [24], [25].

2.1.7 Biometric information

This technique is widely used in DLPS to identify the user accessing the information and thus try to ensure that it is a legitimate user with permissions to access the information [26], [27], [28].

2.1.8 Hypervisor

Hypervisor-based memory introspection, the approach looks for the presence of sensitive raw data in memory on both client and server machines, transcending the dependency on pre-existing security perimeters. This solution presents a high computational cost as a hypervisor-based tool consists of deploying one or more virtual machines to monitor system calls, which consumes too much hardware resources, such as memory and processing [29].

2.2 DRM for document protection

Digital Rights Management (DRM) systems, this term refers to a set of policies, techniques and tools that guide the proper use of digital content. A DRM system is based on ensuring that only intended recipients can view sensitive files regardless of their location. Thus, ensuring data protection beyond the boundaries controlled by DLP systems, so that an organization is always in control of its information [30], [31].

The integration of DLPS and DRM policies ensures that vulnerabilities are minimized and that an organization can immediately deny access to any file, regardless of its location [6]. In [31], [32] and [33], the enterprise digital rights management (eDRM) system is presented, which provides persistent protection for documents using cryptographic methods and also includes features for document protection that are easy to use for the enterprise. In the study [34] the authors reveal the importance of DRM solutions to prevent unauthorized users, inside or outside the boundaries of the organization, from reading an accidentally sent document. As well as, their limitations towards certain types of documents, in addition to preventing the file from further propagation on the external network once filtered, nor an expert hacker from attempting to decrypt the file's content. In [35] DRM systems are compared with the proposed DLPS (UC4Win). In [36] the authors reveal some of the problems

faced by DRM systems as a document security solution, expose that they are difficult or inapplicable to the organization's IT infrastructure and that they rely on certain plugins and these plugins may be used.

2.3 DLPS in the literature

Table 1 summarizes the contributions of the works found in the literature focused on the development and implementation of DLPS, as well as the techniques and technologies employed.

3 Methods to address the Insider threat

The main concern of recent times, in information security, is the internal threat posed by employees, partners, and collaborators of the organizations originating confidential information. One of the main measures adopted in the literature is the control of information use, which goes beyond access control [35] allowing to restrict operations that allow data leakage of confidential information and to regulate its use.

The authors of [54] highlight the importance of strengthening the security of the confidential document management system in the face of the threat of company employees to confidential information; to address this situation, they propose a security model for confidential documents with a distribution control strategy. The first is based on storing the content encrypted with a symmetric encryption algorithm, ensuring that only the authorized user is able to decrypt the content; access control information is stored that allows to know the degree of authority of the user to use such confidential information and records each operation that the user performs; in addition, a hash function is used to ensure the integrity of the content. To control the distribution of confidential information, a client-server strategy is used in which a client will not be able to distribute confidential documentation without permission from the server, in which the control policies defined by the administrator are used and a monitor is installed on the client's computer that allows the server to control the operations performed by the user and prohibit unauthorized operations.

In the study [35] a DLPS based on usage control and dynamic data flow monitoring (UC4Win) is presented. This system can to monitor process calls to the Windows API in order to prevent or modify data flows that pose a threat of confidential information leakage.

In [55] a scheme based on mandatory kernel-level encryption on write operation and decryption on open operation is proposed through middleware to ensure that data remain encrypted in memory. In addition, usage control policies are established, such as read-only, save, export, write, backup, and impression rights. For access control, a method of mutual authentication and key agreement between client and server is proposed, using the SM2 algorithm for its management.

In [26] an approach is presented to control the use of confidential documentation, through the capture of biometric signals from users who interact with the object (document), correlating this information with the content accessed by users, without storing biometric information, but the correlation between the two. In this way, when a loss of information occurs, the organization will be able to know which user accessed the information, minimizing the risk of an attack on the biometric data.

The authors of [23] propose a DLPS based on widows file system mini-filters to control the use of classified documentation by controlling OS I/O operations. The proposed system will block I/O operations from any external storage device. In addition, a strategy is adopted to restrict the movement of classified information by adding the process that performs the read request on the path where the classified information is stored to a blacklist and blocking subsequent write attempts from that process.

The authors of [56] propose a Document Semantic Signature (DSS) approach to address the insider threat. To obtain the DSS, the content of a document is extracted and summarized, updating the DSS dynamically whenever the information is modified. The DLPS monitors the newly generated information by tracking its transfer or exfiltration by comparing the DSS of such information and the DSS of sensitive information. The study takes into account the possibility that an employee with access to confidential information can change the content using synonyms to evade the DLPS, which is based on keyword-based leak detection, and the proposed system addresses this problem. The system was tested with a public dataset achieving encouraging results.

In the study conducted by the authors of [57], a prototype of an anti-leakage system based on the enterprise cloud is presented. The system uses keyword-based content monitoring and filtering techniques. Once the keywords, which represent confidential information in a document to be sent, are detected, the user and the network administrator are alerted of the possible data leakage, and a trace is left in a log where the incidence is written.

In [27] it is proposed to use eye tracking technology for information protection. This technology allows obtaining user behavior information such as gaze location, gaze tracking, and points of interest. This technology in information security can be used to identify the user interacting with confidential information through biometric eye data, obtain metadata of the user performing operations of creating, sending, modifying, and receiving confidential information for use in cases of conflicts detected by the DLP system, in

Table 1 Summary of relevant papers

Title	Contribution	Techniques and Technologies
TaintEraser: Protecting Sensitive Data Leaks Using Application-Level Taint Tracking [37]	TaintEraser is presented, which allows you to track user data that you consider sensitive in the different applications that interact with them.	TaintEraser
CLOUD SHREDDER: Removing the Threat to On-Road Data Disclosure on Laptops in the Cloud Computing Era [38]	This study presents a new approach to eliminate the threat of ubiquitous Internet access and cloud computing "Cloud Shredder".	Cloud Shredder
Hypervisor-based Background Encryption [39]	This study proposes a hypervisor-based approach that enables instant disk encryption without interfering with user activities.	BitVisor (Hypervisor + Encryption)
Hypervisor-based protection of sensitive files in a compromised system [40]	Special purpose hypervisor intended to protect sensitive files on a compromised operating system.	Filesafe (Hypervisor)
A Survey on Data Loss Prevention Techniques [6]	A form of DLP is presented by storing them securely with: "On-the-fly encryption security in storage" (O-E-Sis).	O-E-sis (Encryption)
Architectural Design and Realization for Management of End Point DLP [41]	Software solution developed for DLP terminal protection. It is an architecture designed with kernel hook in the Windows operating system and coded in the C language.	System-Call-Table (SSDT Hooking)
Designing and developing a free Data Loss Prevention system [42]	The authors of this study propose the use of the MyDLP and OpenDLP tools as a free DLP solution.	MyDLP + OpenDLP
MLDED: Multi-Layer Data Exfiltration Detection System [43]	The proposal of this study is MLDED, which is a multi-level data exfiltration detection system. This system allows detecting data exfiltration outside the organization through Hashing, Keyword Extraction, and Tagging techniques.	MLDED (Hashing + Keyword Extraction + Tagging)
The Design and Implementation of User Autonomous Encryption Cloud Storage System Based on Dokan [44]	This study proposes Dokan for the design and implementation of an encrypted cloud storage system.	Dokan (VFS) + Encryption + OpenStack
Linebased end-to-display encryption for secure documents [45]	A line-based encryption method for the design of an end-to-end display cipher is presented. With this technique, data loss can be avoided by using pixel-domain encryption and additional hardware to decrypt the graphics stream, decrypting data between the endpoint and the display, ensuring that data at rest are always encrypted.	On-screen encryption
Biometric/Cryptographic Keys Binding Based on Function Minimization [46]	This study proposes a cryptographic system in which the key depends on biometric patterns, so it is necessary to have a valid biometric template in the system to generate the key.	Biometric cryptosystem (Encryption)
Enterprise Digital Rights Management for Document Protection [31]	This paper studies an enterprise Digital Rights Management System (eDRM) based on cryptography that protects documents and provides useful functions for information security.	eDRM (Encryption)
Hypervisor-Based Sensitive Data Leakage Detector [47]	"HyperSweep" is the DLPS proposed in this study, based on virtual machine memory introspection technology, aiming to check the memory contents of a guest system for sensitive information.	HyperSweep (Virtual Machine + Hypervisor KVM)
Secure data storage and intrusion detection in the cloud using MANN and dual encryption through various attacks [48]	An intrusion detection system using Machine Learning (ML) is proposed to detect if a document contains intrusive data, being this the case, it is stored in a secure site.	ML + Encryption + Steganography techniques
Using malware for the greater good: Mitigating data leakage [34]	DocGuard is the method proposed in this study to protect against accidental data leakage. It consists of making antimalware and antivirus software active on storage systems, detect the leaked file and block access to it.	DocGuard
Off-line enterprise rights management leveraging biometric key binding and secure hardware [28]	This study presents a modification of the biometric key binding scheme proposed in [46] which is mainly used to protect document encryption keys that are stored personal devices.	Biometric cryptosystem
A combined data storage with encryption and keyword-based data retrieval using the scds-tm model in cloud [49]	"SCDS-TM" is the cloud storage system proposed in this study. This system attempts to ensure data confidentiality, integrity, and functionality through elliptic curve cryptography and proof of correctness of the storage.	SCDS-TM (Encryption)
Transparent Encryption with Windows Minifilter Driver [24]	Implementation of a DLP solution that protects any data that is about to leave an endpoint using a Windows Minifilter driver framework. The application provides a plain text view of files even if they are stored in encrypted form on disk.	Minifilter + Encryption

Table 1 (continued)

Title	Contribution	Techniques and Technologies
Data loss prevention (DLP) using MRSH-v2 algorithm [20]	Analysis of the different types of methods for implementing DLP tools, determining that exact file matching is the best approach, using for demonstration the implementation of the MRSH-v2 algorithm to show the capabilities of this method.	MRSH-v2
E-REA Symmetric Key Cryptographic Technique [50]	The authors of this study propose a cryptographic algorithm to protect data from malicious attacks and provide security during data transfer. It is a symmetric key encryption algorithm that represents the improvement of the reverse encryption algorithm.	Encryption
A Forecasting-Based DLP Approach for Data Security [51]	A DLP solution is presented that allows classification of users accessing confidential information according to the number of accesses, based on past data, to predict the future through statistical models.	Statistical analysis
Design and Development of a Dynamic and Efficient PII Data Loss Prevention System [52]	Introduces the DLP tool, PII Guardian, designed to detect and prevent known attacks. PII Guardian provides preventive actions by classifying detected data breaches according to their impact.	Rule-based approach
Cloud security framework and key management services collectively for implementing DLP and IRM [53]	This study proposes an open source framework in the cloud area for building a data loss prevention tool.	Cloud security framework + Key management + Encryption

addition it can serve to improve the security and integrity of documents based on the information of which parts of the document are of greatest interest to the user.

The authors of the study [25] propose as a solution to the internal threat a free DLPS that is based on detecting confidential information at the exit of the USB ports by means of automatic learning and blocks the copy operation, for this purpose it integrates modules in the kernel space (minifilters). The system is developed for the Windows OS as it is the most widely distributed in business environments.

The study [1] focuses on the insider threat that can be intentionally caused by an employee, for this, they propose "Efficient DLP-Visor" which is a context-based DLPS. The system is a thin hypervisor that intercepts call in kernel space. The proposed DLPS makes it possible to detect data leaks even though the employee in question is the system administrator himself. Basically, the System works as follows: The administrator sets a File System path where sensitive information is stored, the DLPS logs any process that opens or reads a document from that path as critical, and any file written by that process is logged as sensitive, as well as any process that receives information from a critical process. DLPS tracks critical processes by capturing kernel mode calls and blocks the relevant operations of those processes.

In [3] and [21] a DLPS for the protection of confidential information is proposed. The proposed system allows access control, through the development of the encryption key, through the combination of a set of parameters; these parameters can be biometric identification of the user accessing the information, geographic location, electronic fingerprint of the device, date, and time, among others. Although this proposal does not specifically present usage control, it is robust due to the ability to require several parameters to generate the decryption key, thus ensuring that the content remains encrypted as long as the established criteria are not met, since the key is never stored.

It has also been seen in the results obtained that DRM tools are based on the control of copies of protected information and therefore gain value for the control of use and protection of information from the threat of collaborators and partners. The proposal of [33] and [32] allows the implementation of an information system independent of the servers containing the control policies that were necessary to access with conventional systems. It controls the use and access to the document through a license (document xml apart from the confidential information) containing the security rules and the configuration of the various security modules necessary for the management of the document. The rules are encrypted by means of public and private keys stored and known by the user.

The authors of [58] analyze three models of traditional document security management, exposing the limitations of each of them, and to try to overcome them they propose a system based on storage in the private enterprise cloud, with a system of authorization and encryption of documents in a virtual machine that encrypts all the document that is written in it, as well as light clients with a common terminal in the virtual machine that will guarantee that all written documents are encrypted and to decrypt them will have to be done through the same encryption system that will guarantee that the user leaves a trace of the operation carried out. External users will need an electronic certificate to decrypt the document.

In [36] the main problems of different solutions for information protection within an organization are identified, among which DLP and DRM solutions are described. A solution based on active documents and DRM is proposed

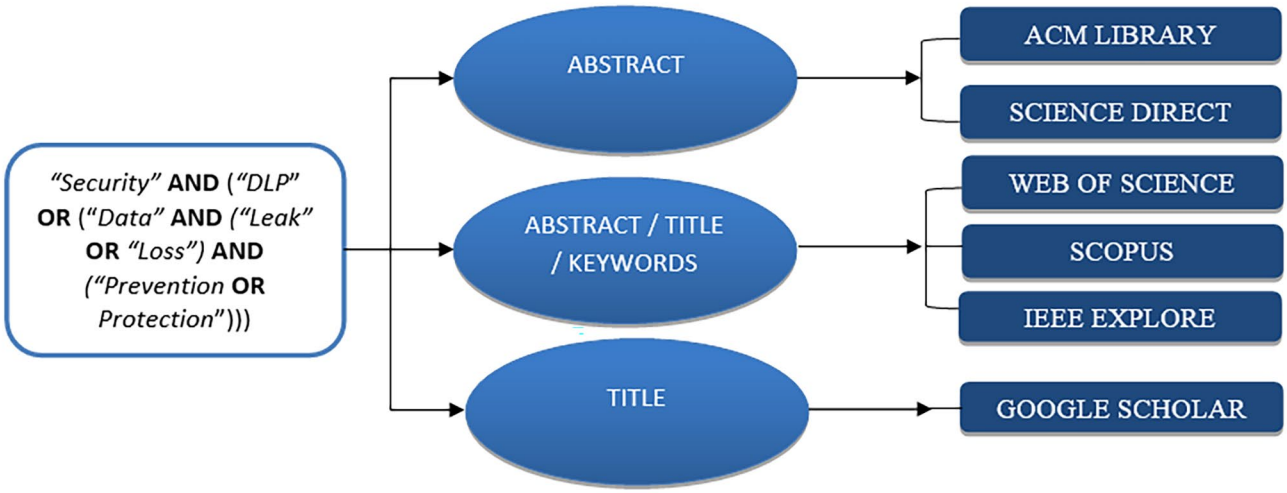

Fig. 1 Search criteria in different databases

that allows the control of document usage, mainly copy, paste, cut, delete, and print operations, inside and outside the organization of origin. The transfer channels considered in this work were removable storage, e-mails, and shared folders. This work does not implement the system, but proposes an idea on how to solve the problem of data leakage with active documents.

Given the persistent concern in organizations and enterprises regarding the internal threat to data leakage protection, it has attracted the interest of the research community in an attempt to circumvent it. The recent study conducted in [59], presents a system CITD for the detection of insider threats based on the behavior of workers according to their role and machine learning. The system was tested in three real organizations to reduce false positives that allow improvements in the tool.

4 Methodology

This paper utilizes the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method [60] in a literature review to analyze existing techniques and technologies for DLP focused on electronic document and classified information leakage. Three stages of PRISMA application are shown in this study: literature search; selection of relevant articles; and data extraction.

4.1 Literature search

For this research, the search was focused on articles related to the techniques and technologies used for DLP published in impact journals, conference articles, and book section, mainly in scientific databases such as Google Scholar, Science Direct, IEEE Xplore, Web of Science, Scopus and ACM Digital Library, from 2011 to April 2022, these databases cover relevant scientific information in multiple engineering fields, allowing access to articles published in scientific and academic journals, repositories, archives and other collections of scientific texts.

The following keywords were used for the literature search: "Security" AND ("DLP" OR ("Data AND ("Leak" OR "Loss") AND ("Prevention OR Protection"). These terms are searched in Abstract/Title/Keywords from 2011 to 2022. Figure 1 shows the search strategy used in this research, the search criteria used are provided by the search engine of each of the scientific databases.

4.2 Study selection and relevant papers

Once the terms have been entered in the search engines of the databases, the articles to be analyzed are selected by reading the titles of the results obtained (in this case 158). Repeated entries in more than one database were eliminated (56 articles). Selection criteria were applied in the analysis of the abstracts of 102 articles to classify those that were completely analyzed, the selection criteria were as follows: (1) Studies of novel proposals of techniques and technologies for DLP. (2) Studies of analysis of techniques and technologies for DLP; 65 articles were obtained for complete analysis, then those studies aimed at systems for malware and rootkit protection, image cryptography and steganography were eliminated, as well as reviews of techniques and technologies since they are related works to this, but not relevant to the analysis. A total of 42 articles remained for analysis.

The procedure described is shown in in Fig. 2 the PRISMA diagram, where the paper selection process can

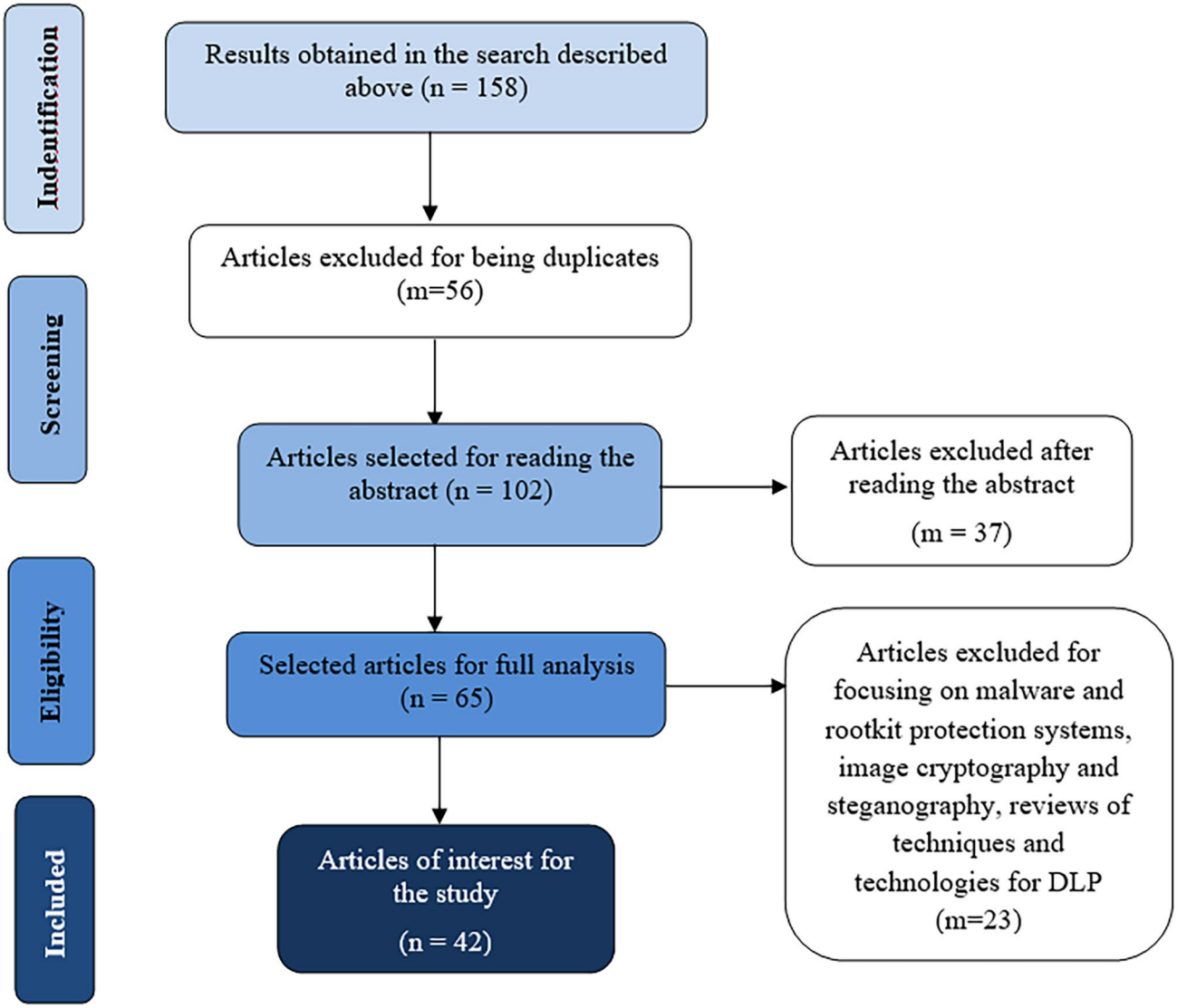

Fig. 2 PRISMA Methodology

be seen and how, out of a total of 158 papers found, a total of 42 papers papers were relevant for analysis in this paper.

5 Discussion of results

This section discusses the results, after applying the above methodology, classifying the relevant studies according to year and type of publication, analyzing the number of relevant publications for each year reviewed and their origin, to determine where the greatest dissemination of the topic in question is to be found. We analyze the use of the main techniques in DLPS, discuss their limitations, advances and applications according to the reviewed literature.

5.1 Classification according to year and type of publication

Figure 3 shows the frequency of publications by year of the relevant studies found during the period 2011–2022. It is observed that the year 2019 reaches the highest number of publications in this period. In general, the number of papers published per year ranges from 1 to 8 with a statistical mode of 3 and a mean of 4 approximately, which means that in the years 2011 and 2019 the mean was exceeded. We can appreciate that approximately 60% of the relevant articles for this study were found in the period between 2017 and 2022, which shows a significant interest in recent years in the security of sensitive digital information. Figure 4 shows the number of published papers according to their origin, it is observed that 60% of the relevant papers come from

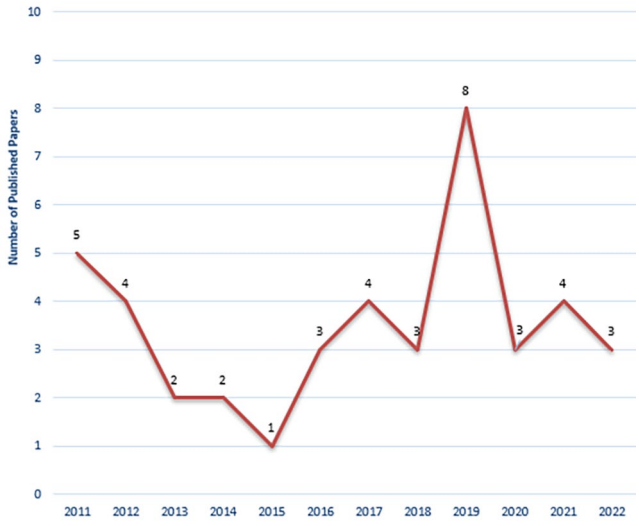

Fig. 3 Frequency of papers published per year

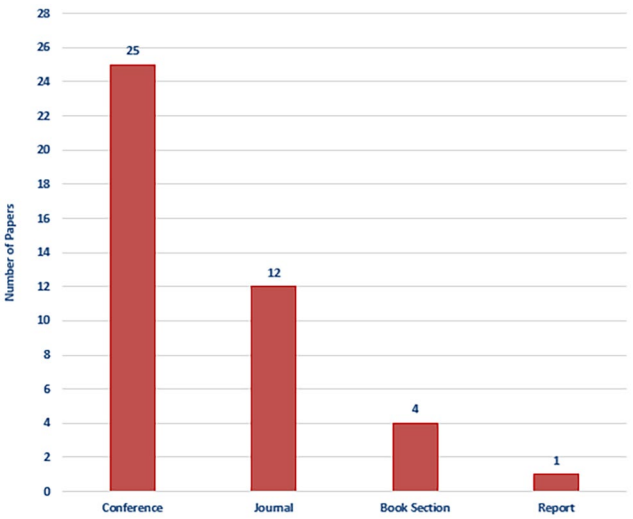

Fig. 4 Paper / Article source

congresses and approximately 30% of them from journals, demonstrating the deep interest in the academic field for protection against data leakage.

5.2 Analysis of the use of the main techniques and technologies in DLPS

Figure 5 shows the most frequently used techniques in the literature. It can be seen that among the most used is cryptography with 40% of use and ML is present in 12% of the articles studied, being evident the progress of DLPS in the use of this technique for the classification of sensitive documentation. Others, such as hypervisor, biometric information capture, and intelligent documents, are present in 10% of the 42 relevant papers to this study. In the literature it has been seen that these techniques and technologies are widely used in combination with each other, For example, in systems where mini-filters and VFS or middleware are used, documents are often encrypted for storage in memory. Also, when active documents are used, hash algorithms are incorporated to guarantee the integrity of the information, as well as ML to classify the information according to the degree of confidentiality to apply security and access policies accordingly. In DLPS, these and other techniques used as a complement can undoubtedly guarantee maximum security to confidential information.

5.3 Limitations, advances, and applications

Limitations that have emerged over the years are the almost complete dependence on the quality of the security policies used and the precise definition of the data to be protected, as well as the necessary over-approaches in the dynamic monitoring of the data flow [35]. In [36] four challenges facing document security are identified, one of them being human negligence, DLPS are not able to overcome this challenge since as a means of security they rely on user, password and security policies to ensure the security of information, without taking into account that the user himself may be the one who provides the data leakage, they themselves are the tools to perform the security policy of any organization so a user and password is not enough. The tracking of unmarked documents [37] or not classified as confidential also represented a major limitation in the DLPS at the time.

Some of these problems have already been solved with the incorporation of new techniques and technologies to DLPS, such as ML for document classification, the recent study [61] proposes a multilayer framework for insider threat detection based on a hybrid method composed of two predictive models with an accuracy level higher than 97%, another application of ML in data protection are network intrusion detection systems, which can be seen in studies [62], [63], [64]. DRM systems for tracking sensitive information outside the organization, biometric information for user identification, and context-based keys to determine the date, place and time of information access. An important advance is the incorporation of blockchain to protect the DLPS logs where the information of detected anomalies is stored, storing these DLPS logs in the Hyperledger Fabric ledger in real time, thus preventing the manipulation of these logs by authorized users to try to eliminate evidence of data leakage [65].

In terms of DLPS applications, the studies reviewed focus on the security of sensitive information at the enterprise level and as such, most of the trends and developments lean in this area. However, the authors of [21] propose a DLP solution using context-based encryption to prevent information leakage in drones. In the poster [66] the authors propose

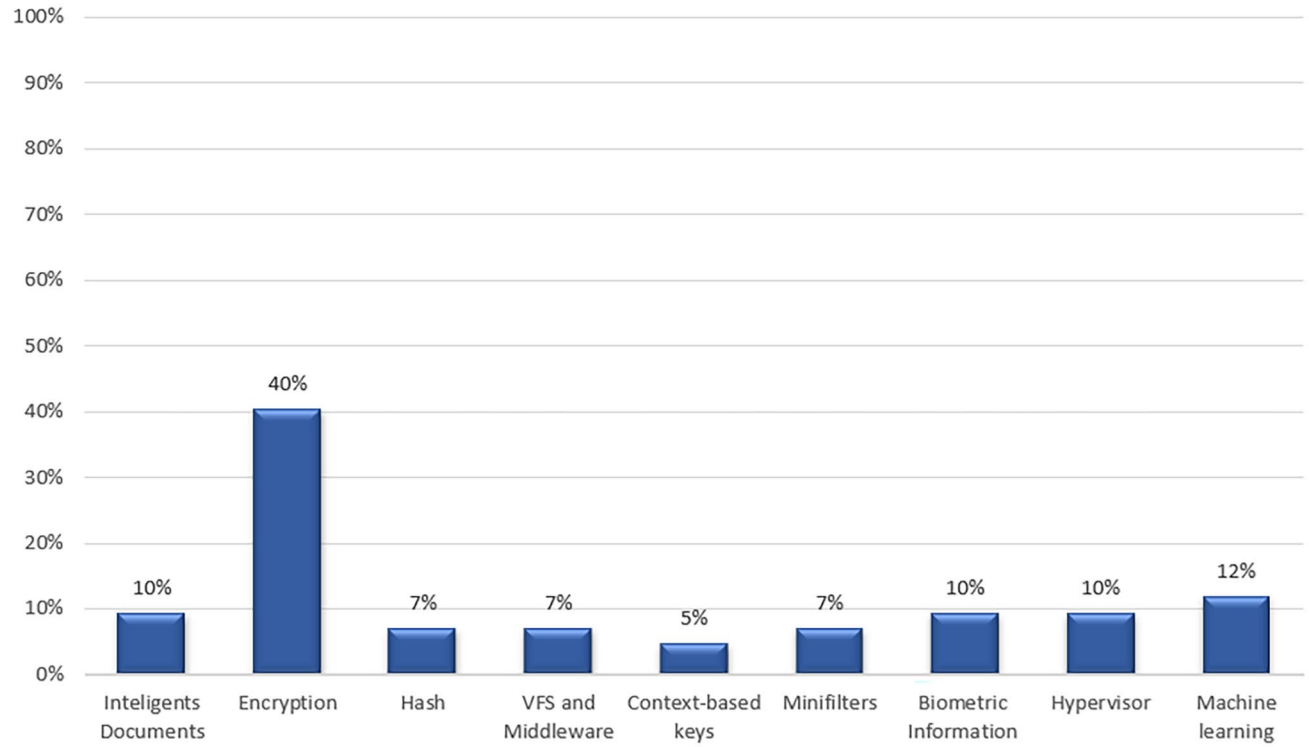

Fig. 5 Percentage of use of the most frequent techniques and technologies

a data leak detection tool for a health information system based on memory introspection. A recent study proposes a blockchain-based architecture that allows the secure transfer of electronic health records between different health care systems, verifying the integrity and consistency of requests and responses to electronic health records [67].

6 Conclusions

This research focuses on a literature survey where a total of 42 relevant studies were obtained. The survey allowed answering three research questions that met the objective proposed in this study. A deep interest in evading insider threat was detected in more than 40% of the analyzed studies. In addition, it is given that the DLPS with the highest incidence in this regard have access control and control of the use of confidential information by controlling the operations that allow data leakage (copy, opening, writing and reading), as well as policies of privacy. DRM for the case of partners and collaborators. These tools mainly use biometric information capture techniques, interception of calls in kernel space using hypervisor, VFS, middleware, and mini-filters. As well as security policies encapsulated in documents. In the analysis of the techniques and technologies that are the most used, We found the encryption technique with 40% use in the studies analyzed.

Significant progress is seen in DLP tools with the incorporation of techniques such as ML for the classification of sensitive information and detection of anomalous activity, in addition to blockchain for the protection of DLPS records. No article was found in the literature that provides the open access code of DLPS for reuse and improvement by other researchers. Few studies focused on data security in the healthcare sector and only one applying DLP on the Internet of Things (IoT) was found in the search results.That is why we propose as future lines of work to carry out studies on the security and protection of the electronic health record, as well as the development and implementation of a DLPS focused on the insider threat, based on the experience of the works found that meet the requirements of being lightweight, unobtrusive, where access to information does not depend on user data and saved passwords, with free access to the source code so that other researchers can adapt it to their needs and provide validations and improvements. To this end, we propose to carry out a study of the techniques and technologies that allow the development of virtual file systems, for the implementation of a secure file system as a DLP tool. As well as, the study of lightweight encryption and decryption algorithms suitable for the needs of a virtual file system. Another line of research that DLPS intends to adopt is its application to IoT, since this technology is advancing every day and most of them are high collectors of personal data.

Acknowledgements This research has been carried out in a collaborative stay between the Telemedicine and e-Health group of the University of Valladolid and the Instituto da Telecomunicações da Delegação da Covilhã, Portugal. We thank Nokia Spain for the close collaboration to achieve successful results.

Author contributions All authors contributed to the conception and design of the study. The first draft of the manuscript was written by Isabel Herrera Montano and all authors commented on earlier versions of the manuscript. All authors read and approved the final manuscript.

Funding This research has been partially supported by the "Centro para el Desarrollo Tecnológico Industrial (CDTI)" of the Spanish Ministry of Science and Innovation in the framework of the project "Technologies for the security of digital relationships in a hyperconnected world (Secureworld)" number 18.IP.MJ. Authors: Mrs. Isabel Herrera Montano, Mr. José Javier García Aranda, Mr. Juan Ramos Diaz, Mr. Sergio Molina Cardín, Mrs. Isabel de la Torre Díez. It is also partially funded by the FCT/MCTES through national funds and, where appropriate, EU co-financed funds under project UIDB/50008/2020; and by the Brazilian National Council for Scientific and Technological Development - CNPq, through grant no. 313036/2020-9. Author: Prof. Joel J. P. C. Rodrigues. Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data Availability Not required in the review article.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

References

1. Kiperberg, M., Amit, G., Yeshooroon, A., Zaidenberg, N.J.: Efficient DLP-visor: An efficient hypervisor-based DLP. In: 2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing (CCGrid). pp. 344–355. IEEE (2021)
2. Alneyadi, S., Sithirasenan, E., Muthukkumarasamy, V.: A survey on data leakage prevention systems. J. Netw. Comput. Appl. **62**, 137–152 (2016). https://doi.org/10.1016/j.jnca.2016.01.008
3. Holgado, P., García, A., García, J.J., Roncero, J., Villagrá, V.A., Jalain, H.: Context-based Encryption Applied to Data Leakage Prevention Solutions. In: Proceedings of the 14th International Joint Conference on e-Business and Telecommunications. pp. 566–571. SCITEPRESS - Science and Technology Publications (2017)
4. Morrow, B.: BYOD security challenges: Control and protect your most sensitive data. Network Security. 5–8 (2012). (2012). https://doi.org/10.1016/S1353-4858(12)70111-3
5. Barlette, Y., Jaouen, A., Baillette, P.: Bring Your Own Device (BYOD) as reversed IT adoption: Insights into managers' coping strategies. Int. J. Inform. Manage. **56**, 102212 (2021). https://doi.org/10.1016/j.ijinfomgt.2020.102212
6. Raj, S.R., Cherian, A., Abraham, A.: A Survey on Data Loss prevention Techniques. Int. J. Sci. Res. **2**, 2319–7064 (2013)
7. Meizlik, D.: The ROI of Data Loss Prevention (DLP). (2008)
8. Brook, C.: DATAINSIDER Digital Guardian's Blog, https://digitalguardian.com/blog/whats-cost-data-breach-2019
9. CCN-CERT: Ciberamenazas y Tendencias Edición 2017 CCN-CERT IA-16/17. In: Centro Criptológico Nacional de España. p. 86: (2017)
10. Zeng, W., Van Moorsel, A.: Quantitative Evaluation of Enterprise DRM Technology. Electronic Notes in Theoretical Computer Science. 275, 159–174 (2011). https://doi.org/10.1016/j.entcs.2011.09.011
11. Ab Rahman, N.H., Choo, K.-K.R.: A survey of information security incident handling in the cloud. Computers & Security. **49**, 45–69 (2015). https://doi.org/10.1016/j.cose.2014.11.006
12. Alonso, S.G., Arambarri, J., López-Coronado, M., de la Torre Díez, I.: Proposing New Blockchain Challenges in eHealth. J. Med. Syst. **43**, 64 (2019). https://doi.org/10.1007/s10916-019-1195-7
13. Georgiadis, G., Poels, G.: Enterprise architecture management as a solution for addressing general data protection regulation requirements in a big data context: a systematic mapping study. Springer, Berlin Heidelberg (2021)
14. Palazov, A.: Some Technologies for Information Security Protection in Weak-Controlled Computer Systems and Their Applicability for eGovernment Services Users. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 117–122 (2011)
15. Hu, C., Chen, F., Zheng, H.: Researches on the Security Protection and Inspection Method for Confidential Documents Based on Linux Operating System. In: Proceedings of the 3rd International Conference on Machine Learning and Soft Computing - ICMLSC 2019. pp. 249–252. ACM Press, New York, New York, USA (2019)
16. Soomro, Z.A., Shah, M.H., Ahmed, J.: Information security management needs more holistic approach: A literature review. Int. J. Inform. Manage. **36**, 215–225 (2016). https://doi.org/10.1016/j.ijinfomgt.2015.11.009
17. Kayode, A.B., Dayo, A.O., Uthman, A.A.: A Review on Distribution Model for Mobile Agent-Based Information Leakage Prevention. Commun. Netw. **13**, 68–78 (2021). https://doi.org/10.4236/cn.2021.132006
18. Leng, J., Zhou, M., Zhao, J.L., Huang, Y., Bian, Y.: Blockchain Security: A Survey of Techniques and Research Directions. IEEE Trans. Serv. Comput. 1–1 (2021). https://doi.org/10.1109/TSC.2020.3038641
19. Leng, J., Ye, S., Zhou, M., Zhao, J.L., Liu, Q., Guo, W., Cao, W., Fu, L.: Blockchain-Secured Smart Manufacturing in Industry 4.0: A Survey. IEEE Trans. Syst. Man Cybernetics: Syst. **51**, 237–252 (2021). https://doi.org/10.1109/TSMC.2020.3040789
20. Husham Ali, B., Jalal, A.A., Al-Obaydy Al-Obaydy, W.N.I.: Data loss prevention (DLP) by using MRSH-v2 algorithm. Int. J. Electr. Comput. Eng. (IJECE). **10**, 3615 (2020). https://doi.org/10.11591/ijece.v10i4.pp3615-3622
21. Garcia, A., Holgado, P., Garcia, J.J., Roncero, J., Villagrá, V., Jalain, H.: Sistema de cifrado basado en contexto aplicado a prevención de fuga de datos. In: Proceedings XIII Jornadas de Ingenieria Telematica - JITEL2017. pp. 93–100. Universitat Politècnica València, Valencia (2017)

22. Garcia Aranda, J.J.A.: EP 2 709 333 A1 EUROPEAN PATENT APPLICATION, (2014)
23. Buda, A., Colesa, A.: File System Minifilter Based Data Leakage Prevention System. In: 2018 17th RoEduNet Conference: Networking in Education and Research (RoEduNet). pp. 1–6. IEEE (2018)
24. Porízek*, D.: Transparent Encryption with Windows Minifilter Driver. (2019)
25. Thombre, S.: Freeware Solution for Preventing Data Leakage by Insider for Windows Framework. International Conference on Computational Performance Evaluation, ComPE 2020. 44–47 (2020). (2020). https://doi.org/10.1109/ComPE49325.2020.9200160
26. Alruban, A., Clarke, N., Li, F., Furnell, S.: Biometrically Linking Document Leakage to the Individuals Responsible. In: Furnell S., Mouratidis H., Pernul G. (eds) Trust, Privacy and Security in Digital Business. pp. 135–149 (2018)
27. Shokishalov, Z., Wang, H.: Applying Eye Tracking in Information Security. Procedia Comput. Sci. **150**, 347–351 (2019). https://doi.org/10.1016/j.procs.2019.02.062
28. Catuogno, L., Galdi, C., Riccio, D.: Off-line enterprise rights management leveraging biometric key binding and secure hardware. J. Ambient Intell. Humaniz. Comput. **10**, 2883–2894 (2019). https://doi.org/10.1007/s12652-018-1023-9
29. Vojnak, D.T., Eordevic, B.S., Timcenko, V.V., Strbac, S.M.: Performance Comparison of the type-2 hypervisor VirtualBox and VMWare Workstation. In: 2019 27th Telecommunications Forum (TELFOR). pp. 1–4. IEEE (2019)
30. Subramanya, S.R., Yi, B.K.: Digital rights management. IEEE Potentials. **25**, 31–34 (2006). https://doi.org/10.1109/MP.2006.1649008
31. Reddy, R.S.C., Gopu, S.R.: Enterprise Digital Rights Management for Document Protection. In: 2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA). pp. 321–326. IEEE (2017)
32. Munier, M., Lalanne, V., Ricarde, M.: Self-protecting documents for cloud storage security. In: Proc. of the 11th IEEE Int. Conference on Trust, Security and Privacy in Computing and Communications, TrustCom-2012. pp. 1231–1238. IEEE (2012)
33. Munier, M.: A Secure Autonomous Document Architecture for Enterprise Digital Right Management. In: 2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems. pp. 16–23. IEEE (2011)
34. Guri, M., Puzis, R., Choo, K.-K.R., Rubinshtein, S., Kedma, G., Elovici, Y.: Using malware for the greater good: Mitigating data leakage. J. Netw. Comput. Appl. **145**, 102405 (2019). https://doi.org/10.1016/j.jnca.2019.07.006
35. Wuchner, T., Pretschner, A.: Data Loss Prevention Based on Data-Driven Usage Control. In: 2012 IEEE 23rd International Symposium on Software Reliability Engineering. pp. 151–160. IEEE (2012)
36. Aaber, Z.S., Crowder, R.M., Fadhel, N.F., Wills, G.B.: Preventing document leakage through active document. In: 2014 World Congress on Internet Security, WorldCIS 2014. pp.53–58.Infonomics Society(2014)
37. Zhu, D.Y., Berkeley, U.C., Song, D., Wetherall, D.: TaintEraser: Protecting Sensitive Data Leaks Using Application-Level Taint Tracking. In: ACM SIGOPS Operating Systems Review. pp. 142–154 (2011)
38. Zhang, N., Jing, J., Liu, P.: CLOUD SHREDDER: Removing the Laptop On-road Data Disclosure Threat in the Cloud Computing Era. In: 2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications. pp. 1592–1599. IEEE (2011)
39. Omote, Y., Chubachi, Y., Shinagawa, T.: Hypervisor-based Background Encryption. In: 27th Annual ACM Symposium on Applied Computing. pp. 1829–1836 (2012)
40. Wang, J., Yu, M., Li, B., Qi, Z., Guan, H.: Hypervisor-based protection of sensitive files in a compromised system. In: 27th Annual ACM Symposium on Applied Computing. pp. 1765–1770 (2012)
41. Topaloglu, M., Ucar, E., Umut, I.: AWERProcedia Information Technology & Computer Science Architectural Design and Realization for Management of end Point. 03, 167–172 (2013)
42. Koutsourelis, D., Katsikas, S.K.: Designing and developing a free Data Loss Prevention system. In: Proceedings of the 18th Panhellenic Conference on Informatics - PCI '14. pp. 1–5. ACM Press, New York, New York, USA (2014)
43. Allawi, M.A.A., Hadi, A., Awajan, A.: MLDED: Multi-layer Data Exfiltration Detection System. In: 2015 Fourth International Conference on Cyber Security, Cyber Warfare, and Digital Forensic (CyberSec). pp. 107–112. IEEE (2015)
44. Yin, J., Yang, J., Chen, Y.: The Design and Implementation of User Autonomous Encryption Cloud Storage System Based on Dokan. In: Proceedings of the 2016 International Conference on Computer Science and Electronic Technology. pp. 917–928. Atlantis Press, Paris, France (2016)
45. Burg, S., Channakeshava, P., Bringmann, O.: Linebased end-to-display encryption for secure documents. In: 2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA). pp. 1–6. IEEE (2016)
46. Riccio, D., Galdi, C., Manzo, R.: Biometric/Cryptographic Keys Binding Based on Function Minimization. In: 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). pp. 144–150. IEEE (2016)
47. Chang, S.-H., Mallissery, S., Hsieh, C.-H., Wu, Y.-S.: Hypervisor-Based Sensitive Data Leakage Detector. In: 2018 IEEE International Conference on Software Quality, Reliability and Security (QRS). pp. 155–162. IEEE (2018)
48. Anitha Ruth, J., Sirmathi, H., Meenakshi, A.: Secure data storage and intrusion detection in the cloud using MANN and dual encryption through various attacks. IET Inform. Secur. **13**, 321–329 (2019). https://doi.org/10.1049/iet-ifs.2018.5295
49. Divya, S.V., Shaji, R.S., Venkadesh, P., A COMBINED DATA STORAGE WITH ENCRYPTION, AND KEYWORD BASED DATA RETRIEVAL USING SCDS-TM MODEL IN CLOUD: Malaysian J. Comput. Sci. **32**, 163–185 (2019). https://doi.org/10.22452/mjcs.vol32no3.1
50. Dhanuja, B., Prabadevi, B., Bhavani Shankari, K., Sathiya, G.: E-REA Symmetric Key Cryptographic Technique. In: 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE). pp. 1–8. IEEE (2020)
51. Gupta, K., Kush, A.: A Forecasting-Based DLP Approach for Data Security. Presented at the (2021)
52. Fugkeaw, S., Worapaluk, K., Tuekla, A., Namkeatsakul, S.: Design and Development of a Dynamic and Efficient PII Data Loss Prevention System. In: Communications and Network. pp. 23–33 (2021)
53. Ahmad, S., Mehfuz, S., Beg, J.: Cloud security framework and key management services collectively for implementing DLP and IRM. Materials Today: Proceedings. (2022). https://doi.org/10.1016/j.matpr.2022.03.420
54. Zheng, S., Liu, J.: A global strategy for controlling document distribution in confidential document management system. In: 2011 IEEE 3rd International Conference on Communication Software and Networks. pp. 410–415. IEEE (2011)
55. Ma, Z.: CPSec DLP: Kernel-Level Content Protection Security System of Data Leakage Prevention. Chin. J. Electron. **26**, 827–836 (2017). https://doi.org/10.1049/cje.2017.05.002
56. Alhindi, H., Traore, I., Woungang, I.: Data Loss Prevention Using Document Semantic Signature. In: Lecture Notes on Data

Engineering and Communications Technologies. pp. 75–99 (2019)
57. Chen, Z., Wang, J., Yang, Y., Yang, G., Wen, L., Chen, L.: Research on Key Technology of Enterprise Private Cloud Anti-Leakage. In: 2019 International Conference on Intelligent Computing, Automation and Systems (ICICAS). pp. 829–834. IEEE (2019)
58. Liu, N.: Cloud Technology in the Security Management of Enterprise Document. In: 2011 Second International Conference on Innovations in Bio-inspired Computing and Applications. pp. 267–269. IEEE (2011)
59. Erola, A., Agrafiotis, I., Goldsmith, M., Creese, S.: Insider-threat detection: Lessons from deploying the CITD tool in three multi-national organisations. J. Inform. Secur. Appl. **67**, 103167 (2022). https://doi.org/10.1016/j.jisa.2022.103167
60. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G.: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. **339**, b2535–b2535 (2009). https://doi.org/10.1136/bmj.b2535
61. Al-Mhiqani, M.N., Ahmad, R., Abidin, Z.Z., Abdulkareem, K.H., Mohammed, M.A., Gupta, D., Shankar, K.: A new intelligent multilayer framework for insider threat detection. Computers & Electrical Engineering. **97**, 107597 (2022). https://doi.org/10.1016/j.compeleceng.2021.107597
62. Awan, M.J., Masood, O.A., Mohammed, M.A., Yasin, A., Zain, A.M., Damaševičius, R., Abdulkareem, K.H.: Image-Based Malware Classification Using VGG19 Network and Spatial Convolutional Attention. Electronics. **10**, 2444 (2021). https://doi.org/10.3390/electronics10192444
63. Khalaf, B.A., Mostafa, S.A., Mustapha, A., Mohammed, M.A., Mahmoud, M.A., Al-Rimy, B.A.S., Abd Razak, S., Elhoseny, M., Marks, A.: An Adaptive Protection of Flooding Attacks Model for Complex Network Environments. Security and Communication Networks. 1–17 (2021). (2021). https://doi.org/10.1155/2021/5542919
64. Azizan, A.H., Mostafa, S.A., Mustapha, A., Foozy, C.F.M., Wahab, M.H.A., Mohammed, M.A., Khalaf, B.A.: A Machine Learning Approach for Improving the Performance of Network Intrusion Detection Systems. Annals of Emerging Technologies in Computing. **5**, 201–208 (2021). https://doi.org/10.33166/AETiC.2021.05.025
65. Lee, G., Son, M., Choi, N., Hong, S., Kim, H.: Blockchain based Removable Storage Device Log Management System. In: 2020 22nd International Conference on Advanced Communication Technology (ICACT). pp. 276–279. IEEE (2020)
66. Mallissery, S., Wu, M.-C., Bau, C.-A., Huang, G.-Z., Yang, C.-Y., Lin, W.-C., Wu, Y.-S.: POSTER: Data Leakage Detection for Health Information System based on Memory Introspection. In: Proceedings of the 15th ACM Asia Conference on Computer and Communications Security. pp. 898–900. ACM, New York, NY, USA (2020)
67. Ajayi, O., Abouali, M., Saadawi, T.: Blockchain architecture for secured inter-healthcare electronic health records exchange. Adv. Intell. Syst. Comput. 1263 AISC. 161–172 (2021). https://doi.org/10.1007/978-3-030-57796-4_16

Isabel Herrera Montano Master in Information and Telecommunication Technologies Research at the University of Valladolid (UVa) in 2020. Currently, member of the Telemedicine and eHealth research group at UVa (http://sigte.tel.uva.es) and PhD student at UVa. She participates in R&D projects related to cybersecurity, information security, machine learning and data analytics. She is also author of articles in SCI journals.

Jose Javier Garcia Aranda is Telecommunication Engineer since 1996 and Ph.D in Telecommunications from The Universidad Politecnica de Madrid in 2015. Currently he is Innovation projects leader at NOKIA Corp. His professional career includes Philips telecommunications, Telefonica Research & Development, Alcatel-Lucent and NOKIA. His current research interests include cyber security, parallel computing, fast image and video coding and computational complexity.

Juan Ramos Diaz Computer Science Engineer by the Universidad Complutense de Madrid. Currently he is working in the Innovation Department in Nokia Spain where he is developing multiple R+D projects related with distributed and parallel computing, cybersecurity and web development.

Sergio Molina Cardín Master's degree in Telecommunications Engineering by the Universidad Politécnica de Madrid (UPM) in 2020. Currently working in the Innovation Department in Nokia TECSS (Spain) developing multiple R&D projects related with several topics such as distributed and parallel computing, cybersecurity or image and video processing

Isabel de la Torre Díez Currently, she is a Professor in the Department of Signal Theory and Communications at the University of Valladolid, Spain. She is leader of GTe Research Group (http://sigte.tel.uva.es). She is author or coauthor of more than 210 papers in SCI journals, peer-reviewed conferences proceedings, books and international book chapters. She has coauthored 21 registered innovative software. She has been involved in more than 100 Program committees of international conferences until 2021. She has participated/coordinated in 45 funded European, national and regional research projects.

Joel J. P. C. Rodrigues [Fellow, IEEE & AAIA] is with the College of Computer Science and Technology, China University of Petroleum, Qingdao, China; Senac Faculty of Ceará, Brazil; and senior researcher at the Instituto de Telecomunicações, Portugal. Prof. Rodrigues is an Highly Cited Researcher (Clarivate), N. 1 of the top scientists in computer science in Brazil (Research.com), the leader of the Next Generation Networks and Applications (NetGNA) research group (CNPq), Member Representative of the IEEE Communications Society on the IEEE Biometrics Council, and the President of the scientific council at ParkUrbis – Covilhã Science and Technology Park. He was Director for Conference Development - IEEE ComSoc Board of Governors, an IEEE Distinguished Lecturer, Technical Activities Committee Chair of the IEEE ComSoc Latin America Region Board, a Past-Chair of the IEEE ComSoc Technical Committee (TC) on eHealth and the TC on Communications Software, a Steering Committee member of the IEEE Life Sciences Technical Community and Publications co-Chair. He is the editor-in-chief of the International Journal of E-Health and Medical Communications and editorial board member of several high-reputed journals (mainly, from IEEE). He has been general chair and TPC Chair of many international conferences, including IEEE ICC, IEEE GLOBECOM, IEEE HEALTHCOM, and IEEE LatinCom. He has authored or coauthored about 1000 papers in refereed international journals and conferences, 3 books, 2 patents, and 1 ITU-T Recommendation. He had been awarded several Outstanding Leadership and Outstanding Service Awards by IEEE Communications Society and several best papers awards. Prof. Rodrigues is a member of the Internet Society, a senior member ACM, and Fellow of AAIA and IEEE.