International Journal of Current Trends in Engineering and Technology
www.ijctet.org, ISSN: 2395-3152
Volume: 12, Issue: 06 (November -December 2026)

A Dynamically Configurable Document Protection Framework for Data Leakage Prevention Using Extended Host-Based Scrutiny and Hybrid Fingerprinting

Vinita Ganeshe, Akrati Shirivastava, and Prachi Tiwari
Department of Computer Science and Engineering
Vaishnavi Institute of Technology and Science, Bhopal, India

Abstract—This paper examines a dynamically configurable document framework that integrates policy-based access control, cryptographic protection, content fingerprinting, and host-based behavioural monitoring. The design incorporates a Probabilistic Incremental Algorithm (PIA), an Autoregressive Hidden Markov Model (AHMM), and an Extended Host-Based Data Scrutinizer (EHDS). Previously reported transaction experiments indicate true-positive rates of 98.1–98.3% for EHDS and 82.1–92.2% for alignment and sampling, with false-positive rates of 9.5–11.0% and 41–45%, respectively. Reported component execution times are also lower for EHDS. Two allocation scenarios illustrate selective distribution under configurable policies. These findings motivate further investigation of integrated document protection; they do not constitute an independent experimental validation. Reproducible algorithm specifications and real-world evaluation remain necessary.

Keywords—Data leakage prevention, insider threat, document protection, EHDS, fingerprinting, access control, behavioural monitoring.

I. INTRODUCTION

The rapid growth of digital information and the increasing adoption of distributed computing, cloud services, and network-based applications have intensified the need for effective information security mechanisms. Organisations routinely process sensitive information, including personal records, financial transactions, intellectual property, customer information, and confidential business documents. Unauthorised disclosure of such information can result in financial losses, reputational damage, regulatory violations, and reduced stakeholder confidence. Consequently, protecting sensitive information throughout its lifecycle, encompassing data at rest, data in use, and data in motion, has become a significant concern in information security. Data Leakage Prevention (DLP) encompasses mechanisms designed to identify, monitor, and restrict unauthorised disclosure of sensitive information across endpoint, network, cloud, and storage environments [1]. However, insider threats remain particularly challenging because authorised users may exploit legitimate access privileges to perform unauthorised activities [2], [3]. These challenges motivate the adoption of intelligent and dynamically configurable security mechanisms that integrate preventive and detective capabilities.

Conventional security mechanisms, including firewalls, antivirus software, intrusion detection systems, encryption, and access control, provide essential protection against a range of security threats. Nevertheless, these mechanisms may not adequately address sensitive information leakage initiated by authorised users. Once legitimate access is granted, distinguishing normal organisational activities from malicious behaviour becomes increasingly difficult.

Conventional fingerprinting techniques based on cryptographic hash functions primarily rely on exact content matching and may fail to identify modified or partially disclosed documents. Frequent document modifications also necessitate fingerprint regeneration, potentially increasing computational overhead [4]. Furthermore, static security policies may inadequately accommodate changing organisational requirements, information sensitivity levels, and evolving access patterns. These limitations highlight the need for an integrated framework that can dynamically enforce security policies, analyse document content, monitor user activities, detect suspicious data movement, and support leakage attribution.

Existing data protection approaches address different aspects of information security through access control, encryption, fingerprinting, content inspection, behavioural analysis, and data allocation mechanisms [5], [6]. However, employing these mechanisms independently may limit their ability to provide coordinated protection throughout the information lifecycle.

The research gap addressed by the underlying framework concerns the integration of dynamically configurable document objects with content-based detection, contextual analysis, behavioural monitoring, and policy-driven data distribution. Specifically, combining the Probabilistic Incremental Algorithm (PIA), the Autoregressive Hidden

Markov Model (AHMM), and the Extended Host-Based Data Scrutinizer (EHDS) provides a unified approach for examining suspicious activities and potential leakage incidents [4]. Nevertheless, the effectiveness of such integration under diverse real-world organisational conditions requires further empirical validation.

The primary objective of this research is to examine an integrated document protection framework for detecting and preventing the leakage of sensitive information across distributed computing environments. Specifically, the study aims to identify and classify sensitive information according to its confidentiality and severity levels, and to examine a dynamically configurable document object model that incorporates cryptographic protection and policy-based access restrictions. It investigates the integration of fingerprinting, regular-expression-based content inspection, the Probabilistic Incremental Algorithm (PIA), and the Autoregressive Hidden Markov Model (AHMM) for effective content and behavioural analysis. Furthermore, the study examines the capabilities of the Extended Host-Based Data Scrutinizer (EHDS) in monitoring data movement, detecting suspicious activities, and facilitating incident reporting. It also investigates policy-driven data allocation and the use of synthetic data objects to support leakage attribution. Finally, the study aims to analyse previously reported experimental results in terms of detection performance, false-positive rates, computational efficiency, and data allocation effectiveness to assess the capabilities and limitations of the integrated framework.

B. Research Contributions

This article presents a structured synthesis and descriptive evaluation of the dynamically configurable document model originally described in [4]. The principal contributions are:

1. A systematic presentation of the integrated document protection architecture combining configurable security policies, cryptographic protection, fingerprinting, and behavioural monitoring.
2. A functional description of PIA, AHMM, EHDS, and policy-based allocation within a unified leakage detection and prevention framework.
3. A comparative analysis of previously reported detection performance and component execution times against the alignment-and-sampling baseline.
4. An examination of reported allocation experiments demonstrating policy-based restrictions on sensitive data distribution.
5. Identification of methodological limitations and future research requirements, including real-world validation, threshold calibration, behavioural concept drift, and formalisation of self-protective, self-advocating, and self-healing document capabilities.

C. Organisation of the Paper

Section II reviews related work. Section III describes the framework architecture and detection mechanisms. Section IV specifies the evaluation methodology, Section V analyses reported results, Section VI discusses their implications, Section VII examines limitations, and Section VIII concludes the article.

II. RELATED WORK

Conventional Data Leakage Prevention (DLP) mechanisms employ multiple security controls to protect sensitive information against unauthorised access, disclosure, and transmission. Firewalls, intrusion detection systems, encryption, access control, content filtering, and activity monitoring constitute fundamental components of organisational information security. Tahboub and Saleh [1, 24] discussed DLP systems for protecting confidential information across different operational environments. Encryption preserves data confidentiality, whereas access control regulates resource accessibility according to predefined security policies. Ferraiolo et al. [7] presented role-based access control (RBAC), which associates permissions with organisational roles. Hu et al. [8] described attribute-based access control (ABAC), enabling authorisation decisions based on subject, object, and environmental attributes. Although these mechanisms regulate information accessibility, they do not independently identify malicious activities performed by legitimately authorised users. Consequently, complementary content inspection and behavioural monitoring mechanisms are required to strengthen information protection.

Data fingerprinting identifies sensitive information by generating characteristic representations of protected content. Conventional fingerprinting approaches employ cryptographic hash functions, including MD5 and SHA-1, to identify identical documents. However, even minor content modifications can produce different hash values, limiting the use of exact-match fingerprinting for detecting altered documents and partial disclosures [4]. Content-based detection

addresses these limitations through keyword matching, regular expressions, text preprocessing, and similarity analysis. Techniques such as stemming, tokenisation, and feature extraction facilitate the identification of sensitive textual patterns. Guha et al. [9] introduced a neural network-based document classification approach integrating n-gram TF-IDF features with a multilayer perceptron to identify confidential information in multipage digital documents. Such approaches extend detection beyond exact document matching but require representative data and appropriate feature selection. Combining fingerprinting with pattern-based and similarity-oriented inspection provides a foundation for identifying different forms of sensitive information disclosure.

Machine learning-based insider threat detection focuses on identifying suspicious behavioural patterns within organisational activities. Liu et al. [2] surveyed approaches to detecting and preventing cyber insider threats, highlighting the challenges posed by legitimate access privileges and malicious activities. Alzaabi and Mehmood [3] reviewed machine learning-based insider threat detection techniques, addressing contemporary challenges and research opportunities. Behavioural modelling generally considers information derived from authentication records, resource access patterns, transaction histories, and contextual attributes. Classification approaches distinguish predefined activity categories, whereas anomaly detection identifies deviations from established behavioural patterns. Al-Shehari et al. [10] investigated an anomaly-based Isolation Forest model for insider threat detection in imbalanced data distributions. Nevertheless, limited representative datasets, evolving user behaviour, and class imbalance remain significant challenges affecting model reliability and generalisability. These limitations motivate combining behavioural evidence with content inspection and policy-based monitoring.

Probabilistic and sequential models provide mechanisms for examining temporal relationships, contextual information, and uncertainty associated with user activities. Hidden Markov Models (HMMs) represent observable activity sequences through underlying latent states, enabling the analysis of behavioural transitions over time. Autoregressive Hidden Markov Models (AHMMs) additionally accommodate dependencies among successive observations. The framework described in [4] incorporates the Probabilistic Incremental Algorithm (PIA) and AHMM for leakage detection and pattern analysis. AHMM-based contextual analysis considers document source, purpose, size, recipient information, metadata, timestamps, application environment, and content patterns. A matching score is evaluated against a predefined threshold to identify suspected leakage. PIA provides an additional detection component within the integrated architecture. However, the available manuscript does not provide complete mathematical formulations, parameter estimation procedures, or evidence of independent reproducibility for these algorithms. Their effectiveness therefore requires further validation using representative activity sequences and calibrated detection thresholds.

Data allocation-based leakage detection investigates the relationship between distributed information and its potential unauthorised disclosure. Papadimitriou and Garcia-Molina [5] introduced a probabilistic attribution model that associates leaked information with recipients using data distribution records. Their approach also considers synthetic decoy objects to strengthen attribution without modifying genuine records. The model described in [4] incorporates explicit allocation of data requests and random agent selection within a policy-driven distribution mechanism. Explicit requests are evaluated against predefined allocation conditions, while random selection operates over eligible recipients. Synthetic objects provide additional evidence when distributed information is subsequently discovered outside authorised environments. Nevertheless, allocation-based attribution estimates potential responsibility rather than conclusively establishing malicious intent. Reliable attribution requires corroborating evidence, trustworthy distribution records, and consideration of alternative disclosure sources.

Host-based data leakage detection monitors information-related activities within endpoint computing environments. Unlike perimeter-oriented mechanisms, host-based approaches examine activities occurring on individual systems, including document access, application usage, and data movement. Herrera et al. [6] investigated secure file systems in the context of DLP against internal threats. The document protection model described in [4] incorporates a Host-Based Data Scrutinizer (HDS) and an Extended Host-Based Data Scrutinizer (EHDS). EHDS combines host activity monitoring, data movement tracking, content inspection, PIA, AHMM-based pattern matching, and incident alerting. Its intended capabilities include detecting complete and partial information leakage and identifying suspicious behavioural patterns. Although the previously reported results indicate improvements in detection performance and component execution times, continuous host monitoring introduces challenges related to computational overhead, user privacy, false alarms, and deployment scalability. These factors require systematic evaluation before practical organisational implementation.

The reviewed literature demonstrates that conventional access control, fingerprinting, machine learning, probabilistic analysis, data allocation, and host-based monitoring address complementary aspects of sensitive information protection.

However, individual mechanisms do not necessarily provide comprehensive coverage of prevention, detection, behavioural assessment, and leakage attribution. Table I summarises representative existing techniques and their scope-related limitations. The identified research gap concerns coordinated integration of dynamically configurable document objects, cryptographic protection, contextual access policies, content inspection, behavioural monitoring, host-based detection, and allocation-based attribution. The integrated framework addresses this design-level gap through layered information protection. Nevertheless, self-healing and self-advocating document capabilities remain design objectives rather than independently established operational capabilities. Further investigation is required to establish reproducibility, robustness, scalability, and effectiveness under real-world organisational conditions.

TABLE I. Comparative analysis of existing data leakage detection and prevention techniques

Reference	Method	Main contribution	Limitations
Papadimitriou and Garcia-Molina [5], 2009	Data allocation and synthetic decoys	Probabilistic leakage attribution	Requires distribution records and corroborating evidence
Pistoia et al. [11], 2015	Configurable leakage detection	Configurable detection in mobile applications	Application-specific scope
Ferraiolo et al. [7], 2001	RBAC	Role-based permission management	Does not independently detect leakage
Tahboub and Saleh [1], 2014	Conventional DLP	Overview of information protection mechanisms	Effectiveness depends on implementation and policy configuration
Liu et al. [2], 2018	Insider-threat survey	Analysis of detection and prevention approaches	Identifies unresolved insider-threat challenges
Alzaabi and Mehmood [3], 2024	Machine learning survey	Review of malicious insider threat detection	Dataset and generalisability challenges
Al-Shehari et al. [10], 2023	Isolation Forest	Anomaly-based insider threat detection	Representative behavioural data required
Herrera et al. [6], 2022	Secure file systems	File-system-oriented protection against internal threats	File-system-specific scope
Guha et al. [9], 2021	MLP and n-gram TF-IDF	Sensitive document classification	Requires representative training and evaluation data
Hu et al. [8], 2014	ABAC	Attribute-driven authorisation decisions	Does not independently perform leakage attribution
Integrated framework [4]	PIA, AHMM, EHDS and configurable document objects	Integration of protection, detection, monitoring and attribution	Limited independent validation and reproducibility evidence

III. PROPOSED SYSTEM ARCHITECTURE

The dynamically configurable document protection framework integrates preventive, detective, and attribution mechanisms to protect sensitive information across endpoint, network, cloud, file, and database environments. The architecture addresses three operational states: data at rest, data in use, and data in motion. It combines sensitive data classification, configurable document objects, authentication, dynamic access control, cryptographic protection, hybrid fingerprinting, behavioural monitoring, and host-based leakage detection. The framework coordinates these components through a reactive engine and an Extended Host-Based Data Scrutinizer (EHDS), enabling policy-based protection and detection of suspicious information-related activities [4], [1], [2], [22], [23]. Fig. 1 depicts the conceptual flow from classification and protected object creation to monitoring and policy decisions.

A. Sensitive Data and Protected Objects

The initial stage identifies sensitive information within organisational resources and classifies it according to confidentiality and severity levels. Sensitive attributes include personal identifiers, authentication information,

biometric information, transaction identifiers, dates of birth, contact details, tax identification numbers, and customer information records. Severity classification establishes the basis for assigning appropriate protection policies and access restrictions. The classified information is subsequently incorporated into document objects, allowing security requirements to be associated with the sensitivity of the underlying data [4], [9], [12].

The document object model represents sensitive information as a configurable entity associated with protection attributes, authentication requirements, access permissions, and distribution constraints. Initially, a data object is constructed using identified sensitive attributes. Cryptographic protection is subsequently applied to generate a protected document object. The object maintains an associated security status that can be updated during permitted access and distribution operations. This design supports document-level protection beyond conventional perimeter-based security. Self-protective, self-advocating, and self-healing capabilities constitute broader design objectives of the underlying architecture [4], [21], [20].

B. Access Control and Cryptographic Protection

Authentication establishes user identity before granting access to protected information. Following successful authentication, access requests are evaluated against the security policies associated with the corresponding document object. Authorised operations are permitted in accordance with applicable access rights, sensitivity levels, and distribution conditions. The framework accommodates configurable access restrictions and updates relevant access levels and object status during authorised operations. Unlike authentication alone, the combination of access control and subsequent activity monitoring addresses potential misuse by legitimately authorised users [7], [4], [8], [19].

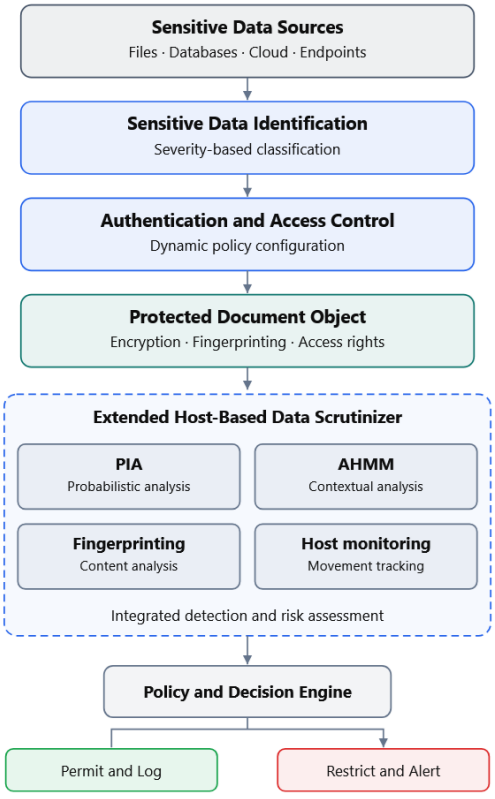


Fig. 1. Proposed system architecture

Cryptographic protection is employed to preserve the confidentiality of sensitive information during storage and authorised handling. The architecture creates an encrypted document object from the original sensitive data object and associates it with relevant security policies. Data separation and virtualisation are additionally incorporated as architectural protection mechanisms intended to restrict unnecessary exposure of confidential information. These mechanisms complement authentication and access control by establishing protected information boundaries. However, the specific encryption algorithms, key management procedures, and virtualisation implementation are not fully specified in the available manuscripts [4], [1], [6].

C. Hybrid Fingerprinting and Behavioural Analysis

The hybrid fingerprinting component combines conventional hash-based identification with regular-expression matching, text preprocessing, and feature-oriented content analysis. Exact-match fingerprinting supports the identification of unchanged documents, whereas stemming, textual patterns, and similarity-oriented analysis are intended to improve the detection of modified or partially disclosed information. Periodic fingerprint generation accommodates changes in protected documents. Combining multiple inspection techniques reduces reliance on exact content matching and facilitates the identification of sensitive information expressed in different textual forms [11], [4], [9], [12].

The reactive engine coordinates the creation of protected document objects, the association of security policies, and the monitoring of authorised user activities. Regular access patterns and activity logs provide behavioural information for identifying potentially suspicious operations. The Probabilistic Incremental Algorithm (PIA) and Autoregressive Hidden Markov Model (AHMM) contribute to behavioural and contextual pattern analysis. Relevant contextual attributes include document source, purpose, size, recipient, timestamps, metadata, and application environment. Matching scores and predefined thresholds support the identification of suspected leakage events [4], [2], [3].

D. Host Monitoring and Policy Enforcement

EHDS integrates host-based activity monitoring, data movement tracking, fingerprinting, pattern matching, PIA, AHMM, and alert-generation mechanisms. It examines information movement and contextual activity patterns to identify potential security policy violations. Its intended capabilities encompass complete and partial leakage detection and prospective risk identification based on historical user activities. The component provides a coordinated detection mechanism that complements document-level protection and access control [4], [2], [10].

The policy enforcement component regulates information access and distribution according to predefined conditions. Eligible recipients receive permitted document objects, while non-compliant requests may be rejected. Synthetic decoy objects and distribution records provide supplementary evidence for the attribution of leakage. When suspicious outbound activity is detected, the framework supports suspension of operations and administrative notification. These mechanisms facilitate coordinated prevention, detection, investigation, and incident response, although attribution requires corroborating evidence before responsibility can be established [5], [4].

IV. EXPERIMENTAL SETUP AND EVALUATION METHODOLOGY

The evaluation considers detection effectiveness, component execution time and policy-based allocation. The source [4] does not document the actual hardware, software versions or measurement protocol. A future reproduction should report these settings and maintain identical resources and workload conditions across methods.

A. Dataset and Workloads

The experimental dataset comprises structured and unstructured organisational records representing sensitive information, legitimate access activities, and potential leakage incidents. Sensitive attributes include identity numbers, transaction identifiers, names, dates of birth, contact details, email addresses, tax identifiers, and customer information file numbers. Contextual attributes encompass access timestamps, document types, source and destination locations, application identifiers, permissions, metadata, and historical activity patterns. The reported leakage detection experiment comprises 1,000 records, including 900 legitimate and 100 leakage instances, representing a class distribution of 90:10. Synthetic records and anonymised identifiers are recommended to preserve data privacy, with legitimate and leakage activities assigned binary labels of 0 and 1, respectively, in accordance with predefined security policies. [4]

Dataset preparation involves data cleaning, sensitive data identification, severity-based classification, protected object construction, encryption, text normalisation, tokenisation, stemming, fingerprint generation, and regular expression-based feature extraction. Historical activity sequences and contextual attributes are provided to the reactive detection engine for behavioural modelling. Independent training, validation, and testing subsets are recommended for model development, parameter optimisation, and unbiased performance assessment, respectively. Stratified partitioning can preserve class distributions, whereas chronological partitioning is preferable for temporal behavioural analysis. Duplicate records and overlapping document fragments should be controlled to prevent information leakage between experimental subsets. [4]

The experimental scenarios encompass authorised data access (ES1), policy-based data allocation (ES2), full-document leakage detection (ES3), partial-document leakage detection (ES4), transaction scalability (ES5–ES8), and behavioural leakage detection (ES9). Policy-based allocation compares four frameworks under two security-policy configurations. Full and partial leakage detection examine fingerprinting and pattern-matching effectiveness, while scalability experiments consider workloads of 10,000, 20,000, 30,000, and 40,000 transactions. Behavioural detection employs historical access patterns and contextual information to identify anomalous activities. Comparative evaluation considers Alignment and Sampling, Host-based Data Scrutinizer (HDS), Probabilistic Incremental Algorithm (PIA), and Auto-Regressive Hidden Markov Model (AHMM), which represent content-based, host-based, probabilistic, and temporal behavioural detection approaches. [4]

B. Configuration and Evaluation Metrics

The framework configuration incorporates fingerprint length, similarity thresholds, sampling parameters, feature dimensionality, historical observation windows, hidden-state configurations, state-transition probabilities, and autoregressive parameters. Leakage identification depends on matching scores exceeding predefined thresholds, subject to applicable security policies. Hyperparameter optimisation and threshold calibration should be performed using validation data, with selected configurations retained during testing. The available implementation description does not specify numerical hyperparameters or dataset partition ratios. [4]

Performance evaluation considers accuracy, precision, recall, True Positive Rate (TPR), False Positive Rate (FPR), Receiver Operating Characteristic (ROC) analysis, and execution time. Accuracy measures overall classification correctness; precision quantifies the reliability of positive predictions; recall measures leakage-detection capability; and FPR characterises incorrect leakage identification among legitimate activities. Additional recommended metrics include F1 Score, ROC AUC, Allocation Success Rate (ASR), throughput, CPU utilisation, and memory consumption. ASR should be interpreted alongside authorisation compliance because restricted allocation may reflect effective security enforcement rather than inadequate performance.

C. Reproducibility

To strengthen experimental reproducibility, fixed random seeds, consistent ground-truth labels, repeated trials, documented algorithmic configurations, and standardised measurement procedures are recommended. Statistical analysis, component-wise ablation studies, and independent validation under heterogeneous workloads can further establish robustness and generalisation. However, these supplementary procedures, numerical hyperparameters, and unreported performance measurements require experimental verification before being presented as completed evaluations.

V. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

This section analyses previously reported results [4]. Quantitative comparisons are available for alignment and sampling; other methods are discussed only where the underlying report provides qualitative evidence.

TABLE II. Comparative detection performance of EHDS and existing alignment-and-sampling method

Transactions	Baseline TPR (%)	EHDS TPR (%)	Baseline FPR (%)	EHDS FPR (%)
10,000	92.2	98.1	44	9.5
20,000	89.1	98.2	41	10.2
30,000	87.2	98.3	42	11
40,000	82.1	98.3	45	11
Mean	87.65	98.23	43	10.43

A. Overall Detection Performance

The overall detection performance of EHDS is evaluated across four transaction workloads comprising 10,000, 20,000, 30,000, and 40,000 transactions. As presented in Table II, EHDS consistently achieves higher true-positive rates (TPRs) and lower false-positive rates (FPRs) than the conventional Alignment-and-Sampling method. The proposed framework maintains a TPR between 98.1% and 98.3%, whereas the baseline exhibits a progressive decline from 92.2% to 82.1%. Furthermore, EHDS maintains an FPR between 9.5% and 11.0%, compared with baseline values ranging from 41.0% to 45.0%. These results demonstrate the effectiveness of integrating content fingerprinting, host-based monitoring,

policy enforcement, and behavioural pattern analysis for sensitive data leakage detection. [4] The proposed framework achieves an average TPR improvement of approximately 10.58 percentage points and a relative average FPR reduction of 75.76% compared with the baseline.

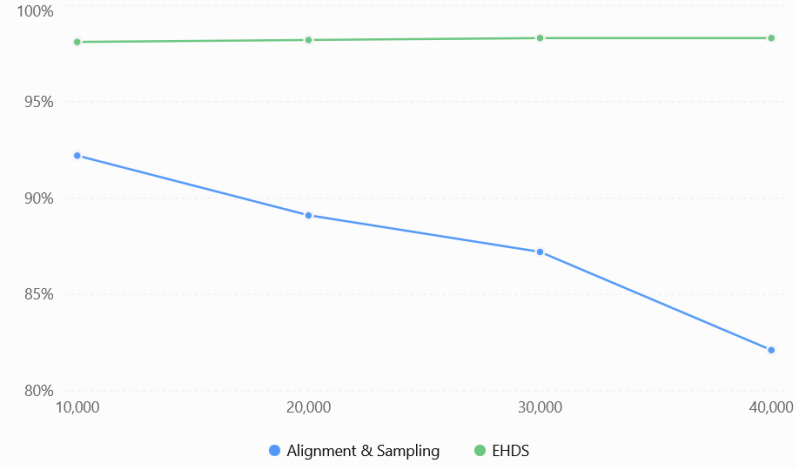


Fig. 2. Comparative True-Positive Rate Across Different Transaction Volumes.

B. True-Positive Rate Analysis

The true-positive rate measures the ability of the proposed framework to identify actual sensitive-data leakage incidents correctly. As illustrated in Fig. 2, EHDS achieves TPRs of 98.1%, 98.2%, 98.3%, and 98.3% for transaction volumes of 10,000, 20,000, 30,000, and 40,000, respectively. Conversely, the Alignment-and-Sampling method exhibits a decline in detection performance as transaction volume increases. At 40,000 transactions, EHDS improves by 16.2 percentage points over the baseline, demonstrating substantially better detection performance across the evaluated workloads. The relatively stable TPR indicates that the detection effectiveness of EHDS is less sensitive to increasing transaction volume within the investigated range. [4]

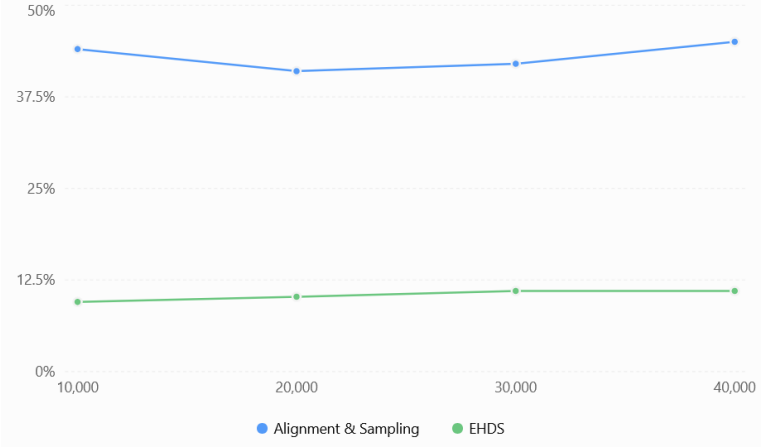


Fig. 3. Comparative False-Positive Rate Across Different Transaction Volumes.

C. False-Positive Rate Analysis

False-positive reduction is essential in data leakage prevention systems because incorrectly classified legitimate activities can interrupt authorised transactions and increase administrative overhead. As illustrated in Fig. 3, EHDS achieves FPRs of 9.5%, 10.2%, 11.0%, and 11.0%, whereas the baseline produces substantially higher values of 44.0%, 41.0%, 42.0%, and 45.0%, respectively. At 40,000 transactions, the proposed framework achieves a 75.56% reduction in relative false positives. Although the EHDS false-positive rate increases slightly with transaction volume, it remains

considerably lower than that of the conventional method. These findings indicate improved discrimination between legitimate activities and actual leakage events under the evaluated simulation conditions. [4]

TABLE III. Estimated comprehensive detection performance under 10% leakage prevalence

Transactions	Method	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
10,000	Baseline	18.89	92.20	31.35	59.62
10,000	EHDS	53.43	98.10	69.18	91.26
20,000	Baseline	19.45	89.10	31.93	62.01
20,000	EHDS	51.68	98.20	67.72	90.64
30,000	Baseline	18.74	87.20	30.86	60.92
30,000	EHDS	49.82	98.30	66.13	89.93
40,000	Baseline	16.85	82.10	27.97	57.71
40,000	EHDS	49.82	98.30	66.13	89.93

D. Precision, Recall, F1-Score, and Accuracy Analysis

Precision, recall, F1-score, and accuracy provide complementary perspectives for evaluating the classification effectiveness of the proposed framework. Precision measures the reliability of identified leakage incidents, whereas recall represents the proportion of actual leakage events successfully detected. The F1-score balances precision and recall, while accuracy measures overall classification correctness. The original manuscript reports favourable precision and recall performance for PIA across varying test set sizes. For an illustrative quantitative assessment, the reported TPR and FPR values can be combined with an assumed leakage prevalence of 10%, corresponding to the class distribution described in the separate 1,000-record experiment. Under this assumption, precision and accuracy are calculated as Table III presents the conditional estimates; Fig. 4 visualises the 10,000-transaction example. [4]

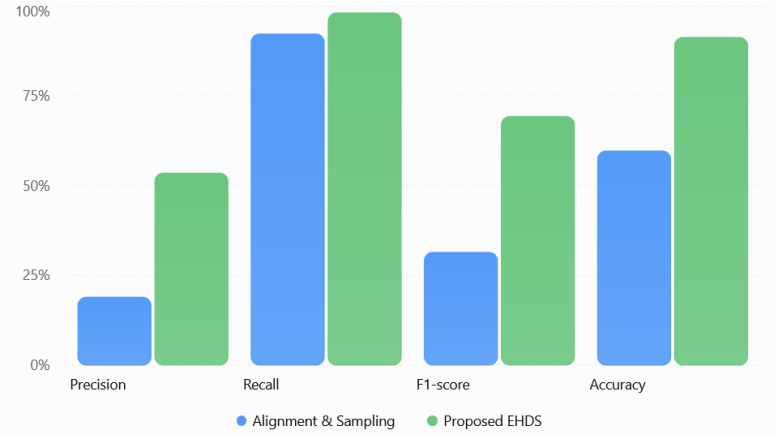


Fig. 4. Estimated comparative precision, recall, F1-score, and detection accuracy for 10,000 transactions under an assumed 10% leakage prevalence.

The estimated results indicate that EHDS provides improved classification performance under the assumed class distribution. However, these estimates also highlight the importance of reducing false positives when leakage incidents constitute a relatively small proportion of organisational activities.

E. ROC Curve and AUC Analysis

Receiver Operating Characteristic (ROC) analysis evaluates the discrimination capability of a detection method by examining the relationship between its true-positive and false-positive rates. The comparative ROC assessment in the

original manuscript includes PIA, HDS, EHDS, and AHMM, with PIA showing favourable discrimination performance. The reported EHDS operating points demonstrate high sensitivity at comparatively low false-positive rates. Nevertheless, complete ROC curves require predictions or detection scores obtained across multiple classification thresholds. Fig. 5 displays the reported operating points, and Table IV summarises their specificity and balanced accuracy. [4]

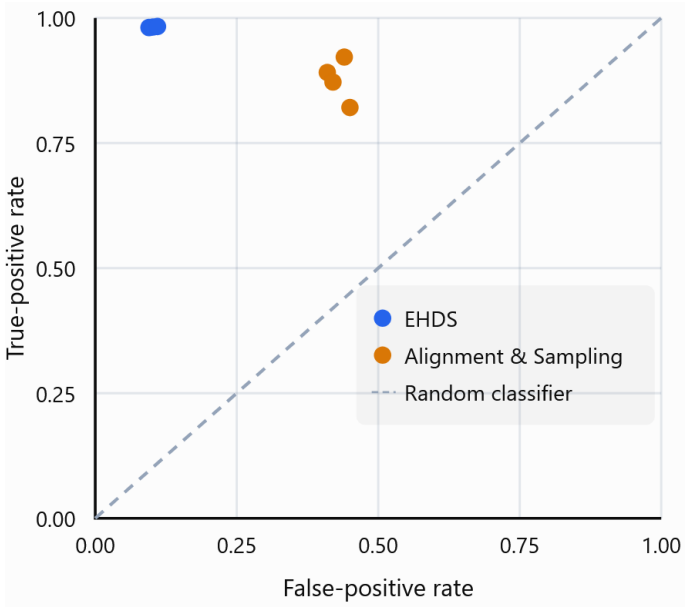


Fig. 5. Comparative ROC operating points across the evaluated transaction workloads.

The reported operating points establish the favourable sensitivity–specificity trade-off of EHDS relative to Alignment and Sampling. Since these points correspond to different transaction workloads rather than a systematic threshold sweep, they cannot be integrated to obtain a valid empirical AUC.

TABLE IV. Comparative ROC operating-point performance

Transactions	Method	Sensitivity (%)	Specificity (%)	Balanced accuracy (%)
10,000	Baseline	92.2	56	74.1
10,000	EHDS	98.1	90.5	94.3
20,000	Baseline	89.1	59	74.05
20,000	EHDS	98.2	89.8	94
30,000	Baseline	87.2	58	72.6
30,000	EHDS	98.3	89	93.65
40,000	Baseline	82.1	55	68.55
40,000	EHDS	98.3	89	93.65

F. Execution Time and Computational Efficiency

Computational efficiency is critical for continuous data leakage monitoring in distributed environments. The execution-time evaluation compares EHDS with the conventional Alignment-and-Sampling method across three computational components: data movement tracking, rule matching, and similarity calculation. As presented in Table V, EHDS requires 1.2 ms, 3.1 ms, and 2.4 ms for these operations, respectively, compared with 10.0 ms, 12.0 ms, and 7.3 ms for the baseline. The corresponding speedup factors are 8.33×, 3.87×, and 3.04×. The combined component execution time decreases from 29.3 ms to 6.7 ms, representing a reduction of approximately 77.13%. These results demonstrate the computational advantages of the proposed framework for the evaluated operations. Fig. 6 illustrates the relative component times. [4]

TABLE V. Comparative component execution time

Component	Baseline (ms)	EHDS (ms)	Speedup	Reduction (%)
Data movement tracking	10	1.2	8.33×	88
Rule matching	12	3.1	3.87×	74.17
Similarity calculation	7.3	2.4	3.04×	67.12
Combined component time	29.3	6.7	4.37×	77.13

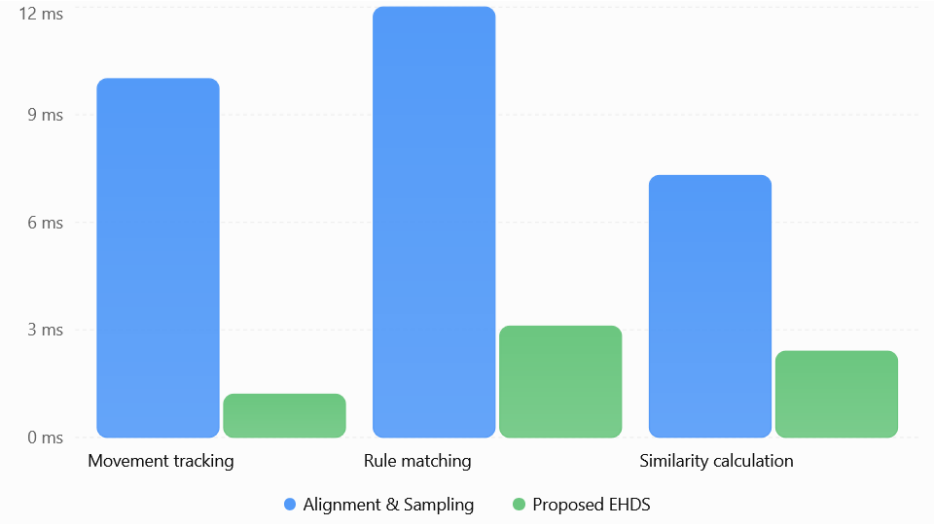


Fig. 6. Component-wise execution time comparison between EHDS and alignment and sampling.

G. Scalability Analysis

The scalability evaluation investigates the performance stability of the proposed framework as transaction volume increases from 10,000 to 40,000. Despite a fourfold increase in workload, EHDS maintains a TPR above 98%, with an overall variation of only 0.2 percentage points. In contrast, the baseline experiences a 10.1 percentage-point decline in TPR. The FPR of EHDS increases moderately from 9.5% to 11.0%, indicating relatively stable classification performance. The balanced accuracy of EHDS remains between 93.65% and 94.30%, whereas the baseline decreases to 68.55% at the largest evaluated workload. These findings establish detection-performance scalability within the investigated transaction range. Further assessment of end-to-end latency, throughput, processor utilisation, and memory consumption is necessary to establish computational scalability. Fig. 7 shows the balanced accuracy trends calculated from Table IV. [4]

H. Policy-Based Data Allocation Performance

The dynamic policy-based allocation mechanism regulates the distribution of sensitive data according to predefined access permissions and contextual conditions. The experimental evaluation considers four recipient frameworks under two policy configurations. As presented in Table VI, the total number of requested objects is 29, of which 20 are allocated under Case 1 and 13 under Case 2, corresponding to allocation rates of 68.97% and 44.83%, respectively. Framework 1 receives all requested objects under both configurations, whereas Frameworks 2 and 4 receive no objects under Case 2. The allocation to Framework 3 remains unchanged. These results demonstrate the ability of the proposed model to modify data distribution according to configurable policy conditions. However, policy correctness must be independently evaluated against the expected authorisation decisions. Fig. 8 visualises the differences in allocation across recipients. [4]

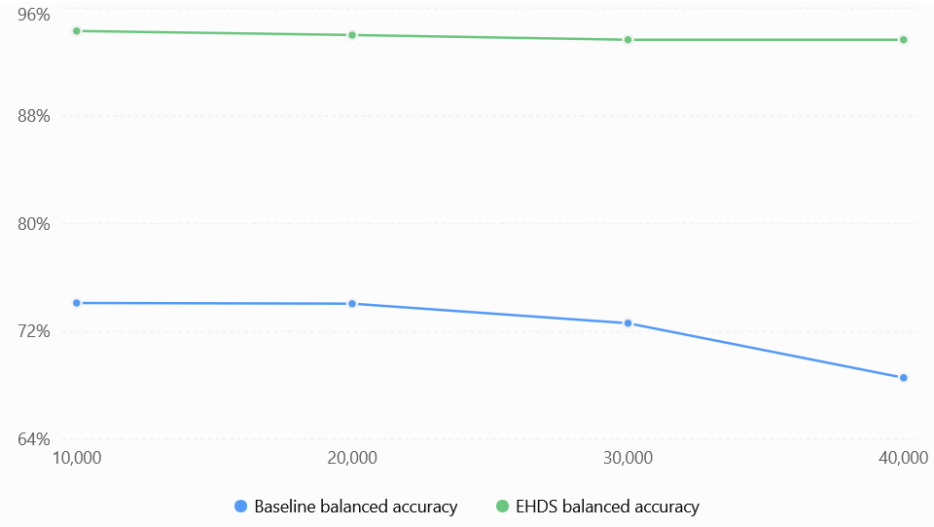


Fig. 7. Balanced accuracy across transaction workloads.

TABLE VI. Comparison of requested and allocated data objects under different policy conditions

Framework	Requested	Case 1 allocated	Case 2 allocated
Framework 1	8	8	8
Framework 2	7	4	0
Framework 3	8	5	5
Framework 4	6	3	0
Total	29	20	13
Allocation rate (%)	—	68.97	44.83
Non-allocation rate (%)	—	31.03	55.17

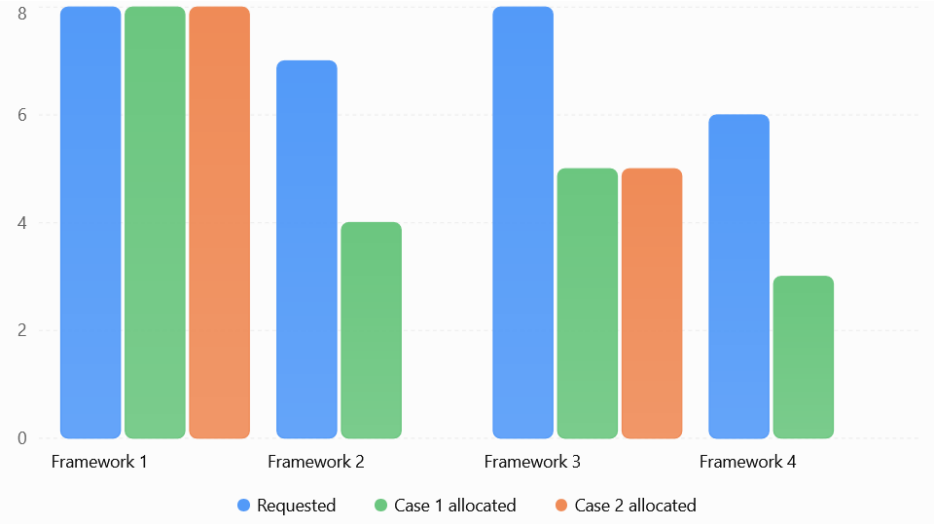


Fig. 8. Requested versus allocated data objects under two dynamic policy conditions.

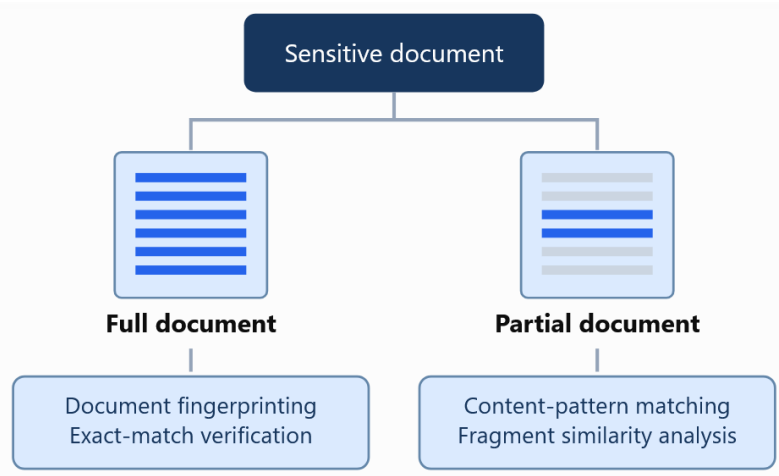


Fig. 9. Proposed evaluation scenarios for partial and full document leakage detection.

I. Partial and Full Document Leakage Detection

The proposed EHDS framework incorporates document fingerprinting, regular-expression matching, content similarity analysis, and contextual monitoring to address both partial and complete document leakage. Conventional hash-based fingerprinting provides effective exact matching for unmodified documents but becomes less effective when document content changes. The proposed architecture addresses this limitation by combining text preprocessing, content pattern recognition, and behavioural analysis. Full-document leakage detection focuses on identifying complete sensitive documents, whereas partial-document detection examines sensitive fragments and modified content. The original manuscript reports 100% exact-match detection for unmodified files using conventional fingerprinting; however, this result does not establish equivalent performance for modified or partially disclosed documents [18, 25]. A comprehensive category-wise evaluation should measure detection accuracy, precision, recall, and FPR independently for complete documents and different proportions of retained sensitive content. Fig. 9 illustrates a proposed evaluation design rather than observed results.

VI. DISCUSSION

A. Interpretation of Experimental Results

The results in Tables II and V suggest improved sensitivity and lower component execution times under the reported conditions [4]. The observations do not isolate the effect of individual modules or establish operational effectiveness.

B. Research Question-Wise Discussion

The reported comparisons address detection sensitivity, false positives, stability across workloads, component execution time and selective allocation. Individual component contributions and generalisation remain unresolved and require ablation studies and independent validation.

C. Advantages of the Proposed Framework

The principal advantage of EHDS lies in its integration of complementary detection and prevention mechanisms within a unified architecture. Dynamic policies enable configurable access restrictions, while protected document objects support persistent data protection. Fingerprinting and regular-expression matching facilitate the identification of sensitive content, whereas PIA, AHMM, and behavioural monitoring support the detection of contextual leakage. This layered architecture reduces dependence on individual detection mechanisms and provides adaptable protection across different data-access environments.

D. Detection and Prevention Trade-Offs

Effective data leakage prevention requires balancing detection sensitivity, false alarms, computational overhead, and the accessibility of legitimate information. Although EHDS maintains high detection sensitivity, its remaining false-positive rates may disrupt authorised activities. Restrictive allocation policies can limit sensitive-data exposure but may

also affect operational productivity. Therefore, threshold calibration, adaptive policy configuration, and continuous performance assessment are essential for maintaining an appropriate balance between security and usability.

E. Practical Applicability

The proposed framework has potential applications in enterprise information systems, financial organisations, healthcare information systems, and cloud-based environments. It could facilitate the protection of confidential business documents, financial transactions, electronic health records, and shared organisational resources. However, practical deployment requires further validation using real-world workloads, organisation-specific security policies, and operational requirements.

F. Privacy and Ethical Considerations

Behavioural monitoring and employee activity analysis introduce important privacy and ethical considerations. Implementation should incorporate data minimisation, anonymisation, encryption, role-based access restrictions, and limited log-retention periods. Monitoring activities should comply with applicable privacy regulations and organisational policies. Automated leakage alerts should undergo appropriate human review before any disciplinary or administrative action, ensuring accountability, transparency, proportionality, and the responsible utilisation of sensitive behavioural information.

VII. LIMITATIONS AND THREATS TO VALIDITY

A. Dataset Limitations

The experimental evaluation is primarily simulation-oriented and conducted using distributed datasets rather than standardised publicly available benchmarks. Consequently, the datasets may not adequately represent the complexity, diversity, and dynamic characteristics of real-world organisational activities. Limited variation in leakage patterns, document formats, and user behaviour may influence detection performance. Further evaluation using heterogeneous datasets with representative leakage scenarios is necessary to strengthen empirical validity.

B. Internal Validity

Internal validity may be influenced by experimental configurations, detection thresholds, dataset preparation, and baseline implementation settings. Although the proposed EHDS framework demonstrates improved detection performance, the contributions of individual components have not been established independently through ablation studies. Controlled experimental conditions, consistent parameter configurations, repeated trials, and statistical significance testing are necessary to minimise potential experimental bias.

C. External Validity

The experimental findings are restricted to the evaluated transaction workloads and simulation conditions. Differences in organisational infrastructure, security policies, application environments, and operational requirements may affect the practical effectiveness of detection. Consequently, the reported performance cannot be directly generalised to enterprise-scale systems without independent validation across diverse operational environments and standardised benchmark datasets.

D. Construct Validity

The evaluation primarily considers true-positive rate, false-positive rate, component execution time, and data allocation performance. Although these metrics provide meaningful insights, they do not comprehensively represent overall prevention effectiveness, policy correctness, operational usability, or detection reliability. Additional metrics, including F1 Score, ROC AUC, policy compliance, and incident response effectiveness, would strengthen construct validity.

E. Computational and Deployment Limitations

Continuous host-based monitoring, fingerprint generation, contextual analysis, and dynamic policy enforcement may introduce computational and storage overhead. Although the reported component execution times demonstrate computational improvements, end-to-end latency, memory consumption, network overhead, and resource utilisation require further investigation. Moreover, the self-healing and self-advocating capabilities require formalisation and implementation-level validation before practical deployment.

F. Generalizability to Real-World Organisational Environments

Real-world organisational environments involve heterogeneous applications, evolving access patterns, distributed infrastructures, and unpredictable user behaviour. Such conditions may introduce concept drift and previously unseen leakage scenarios, potentially reducing detection reliability. Continuous model adaptation, periodic threshold calibration, and cross-organisational validation are necessary to establish long-term generalizability and operational robustness.

G. Reliability of Leakage Attribution

Leakage attribution based on allocation histories, document fingerprints, and behavioural patterns may provide useful investigative evidence but cannot independently establish individual responsibility. Shared credentials, compromised accounts, incomplete activity logs, and indirect data transfers may result in incorrect attribution. Therefore, suspected leakage incidents require corroborating evidence, secure audit trails, and authorised human investigation. Incorporating confidence-based attribution and transparent incident-review mechanisms would improve accountability while minimising erroneous accusations and associated privacy risks.

VIII. CONCLUSION AND FUTURE WORK

This article synthesises a dynamically configurable document protection framework that incorporates access policies, protected objects, hybrid fingerprinting, PIA, AHMM, and EHDS. Previously reported results [4] show a mean TPR of 98.23% versus 87.65% and a mean FPR of 10.43% versus 43.00% for alignment and sampling. The sum of the three reported component times is 77.13% lower. Allocation examples illustrate selective distribution. These findings motivate further evaluation but do not constitute independent validation of the framework. Future research will focus on validating the proposed framework using standardised public benchmarks and real-world organisational datasets. Adaptive model updating, concept drift detection, and automated threshold calibration should be investigated to accommodate evolving user behaviour and emerging leakage patterns. Further studies should evaluate the continuous monitoring overhead, end-to-end latency, resource utilisation, and scalability in distributed deployments. Privacy-preserving behavioural analysis and secure audit mechanisms represent additional research directions. Formalisation and implementation-level validation of self-healing and self-advocating document capabilities will further strengthen autonomous data protection, operational reliability, and practical applicability.

REFERENCES

[1]. P. Papadimitriou and H. Garcia-Molina, “A Model for Data Leakage Detection,” in Proc. 25th IEEE Int. Conf. Data Engineering (ICDE), 2009, doi: 10.1109/ICDE.2009.227.
[2]. R. Agrawal and J. Kiernan, “Watermarking Relational Databases,” in Proc. 28th Int. Conf. Very Large Data Bases (VLDB), 2002.
[3]. M. Pistoia, O. Tripp, P. Centonze, and J. W. Ligman, “Labyrinth: Visually Configurable Data-Leakage Detection in Mobile Applications,” in Proc. 16th IEEE Int. Conf. Mobile Data Management (MDM), 2015, pp. 279–286, doi: 10.1109/MDM.2015.69.
[4]. L. Sweeney, “k-Anonymity: A Model for Protecting Privacy,” Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 5, pp. 557–570, 2002.
[5]. NIST, “Framework for Improving Critical Infrastructure Cybersecurity,” National Institute of Standards and Technology, Gaithersburg, MD, USA.
[6]. D. F. Ferraiolo, R. Sandhu, S. Gavrila, D. Kuhn, and R. Chandramouli, “Proposed NIST Standard for Role-Based Access Control,” ACM Trans. Inf. Syst. Secur., vol. 4, no. 3, pp. 224–274, 2001.
[7]. Verizon, “Data Breach Investigations Report,” Verizon Enterprise, various editions, as summarised in the supplied thesis.
[8]. IBM, “InfoSphere DataStage,” data lineage and integration platform, as discussed in the supplied thesis.
[9]. The supplied thesis, “Dynamically Configurable Document Model to Prevent Data Leakage,” Department of Computer Science and Engineering, VTU, Belagavi, India, source document supplied by the authors.
[10]. Supplied thesis literature survey, cited work on insider-threat detection using association analysis, machine learning, and convolutional neural networks.
[11]. Supplied thesis literature survey, cited work on sensitive-word detection using Boyer–Moore, Smith–Waterman, TF-IDF, LSI, and Zemberek.

[12]. Supplied thesis literature survey, cited work on runtime data-leakage detection and watermark-based network filtering.
[13]. R. Tahboub and Y. Saleh, "Data Leakage/Loss Prevention Systems (DLP)," in Proc. 2014 World Congress on Computer Applications and Information Systems (WCCAIS), 2014, pp. 1–6, doi: 10.1109/WCCAIS.2014.6916624.
[14]. L. Liu, O. Y. De Vel, Q.-L. Han, J. Zhang, and Y. Xiang, "Detecting and Preventing Cyber Insider Threats: A Survey," IEEE Commun. Surveys Tuts., vol. 20, no. 2, pp. 1397–1417, 2018.
[15]. F. R. Alzaabi and A. Mehmood, "A Review of Recent Advances, Challenges, and Opportunities in Malicious Insider Threat Detection Using Machine Learning Methods," IEEE Access, vol. 12, pp. 30907–30927, 2024, doi: 10.1109/ACCESS.2024.3369906.
[16]. R. K. Tenali and V. Vijaya Chamundeeswari, "Machine Learning-Based Approaches for Malicious Insider Threat Detection: A Survey of Recent Advances and Open Issues," in Proc. 2025 Int. Conf. Intelligent Computing, Information and Control Systems (ICOIICS), 2025, doi: 10.1109/ICOIICS67115.2025.11390400.
[17]. [17] T. Al-Shehari, M. Al-Razgan, T. Alfakih, R. A. Alsowail, and S. Pandiaraj, "Insider Threat Detection Model Using Anomaly-Based Isolation Forest Algorithm," IEEE Access, vol. 11, pp. 118170–118185, 2023.
[18]. Z. Q. Wang and A. El Saddik, "DTITD: An Intelligent Insider Threat Detection Framework Based on Digital Twin and Self-Attention Based Deep Learning Models," IEEE Access, vol. 11, pp. 114013–114030, 2023, doi: 10.1109/ACCESS.2023.3324371.
[19]. I. Herrera, I. de la Torre Díez, J. J. García Aranda, and J. Ramos Díaz, "Secure File Systems for the Development of a Data Leak Protection (DLP) Tool Against Internal Threats," in Proc. 17th Iberian Conf. Information Systems and Technologies (CISTI), 2022, doi: 10.23919/CISTI54924.2022.9820170.
[20]. N. Li, T. Li, and S. Venkatasubramanian, "t-Closeness: Privacy Beyond k-Anonymity and l-Diversity," in Proc. IEEE 23rd Int. Conf. Data Engineering (ICDE), 2007, pp. 106–115, doi: 10.1109/ICDE.2007.367856.
[21]. Y. Cui and J. Widom, "Lineage Tracing for General Data Warehouse Transformations," in Proc. 8th Int. Conf. Database Theory (ICDT), 2001, pp. 316–330.
[22]. D. Liu, X. Liu, L. Ma, Y. Chang, R. Wang, H. Zhang, H. Yu, and W. Wang, "Research on Leakage Prevention Technology of Sensitive Data Based on Artificial Intelligence," in Proc. 2020 IEEE 10th Int. Conf. Electronics Information and Emergency Communication (ICEIEC), 2020, pp. 142–145.
[23]. [23] A. Guha, D. Samanta, A. Banerjee, and D. Agarwal, "Deep Learning Model for Information Loss Prevention From Multipage Digital Documents," IEEE Access, vol. 9, pp. 80451–80465, 2021, doi: 10.1109/ACCESS.2021.3084841.
[24]. A. Martinelli, F. Marulli, F. Mercaldo, S. Marrone, and A. Santone, "Enhanced Privacy and Data Protection Using Natural Language Processing and Artificial Intelligence," in Proc. 2020 Int. Joint Conf. Neural Networks (IJCNN), 2020, pp. 1–8, doi: 10.1109/IJCNN48605.2020.9206801.
[25]. V. C. Hu, D. Ferraiolo, R. Kuhn, A. Schnitzer, K. Sandlin, R. Miller, and K. Scarfone, "Guide to Attribute-Based Access Control (ABAC) Definition and Considerations," NIST SP 800-162, 2014.