

A Systematic Survey of Ransomware Detection Through Pipelines: From Runtime Telemetry to Explainable and Response-Aware Defense

Mohammed Abd Alkareem Abd ¹, Wasan Saad Ahmed ²

¹ Department of Computer Science, College of Science, University of Diyala, 32001, Iraq.
scicompms242503@uodiyala.edu.iq , wasan@uodiyala.edu.iq

Abstract

Ransomware has evolved from simple file encryption into an advanced cyberattack that includes stealthy intrusion, behavioral manipulation, lateral movement, data exfiltration, and encryption-based impact. Much research has focused on ransomware detection. However, previous methods are fragmented and focus on detecting a single aspect, such as static analyzers, API-call modelling, deep learning, pre-encryption detection, explainable artificial intelligence (XAI), or response inference. However, many solutions achieve high detection accuracy in controlled environments. However, they are not adaptable in real time, lack context awareness, are not explainable, or have no integrated response functions. Recent research suggests that most available detection methods rely on a single telemetry source, such as API sequences, DLL usage, mutex behavior, or sandbox logs, and that ransomware detection is more of a classification problem than a continuous behavioral reasoning process. The limited interpretability of AI models has led to the use of XAI techniques to enhance transparency and trust while maintaining high detection performance. At the same time, Work has been conducted on early detection, cyber threat intelligence (CTI), ontology-based reasoning, and adversary behavior modelling. This survey presents a systematic review and classification of ransomware detection research, focusing on four key areas: telemetry collection, feature engineering for ransomware behaviors, machine learning and deep learning, and intelligent response mechanisms, each of which is broken down into subcategories. This survey describes and combines results from the following research approaches: dynamic API Modelling, CTI-enhanced analysis, early detection strategies, and MITRE ATT&CK-based behavioural modelling, to present a full picture of intelligent, explainable, and response-aware ransomware defence systems and highlight issues and future research directions.

Keywords: Ransomware Detection, Behavioral Analysis, Wazuh, Explainable AI, Zero-Day Detection.

Article history: Received: 15 May 2026, Accepted: 11 Aug 2026, Published: 15 Sep 2026.

This article is open-access under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Ransomware has changed significantly in recent years. Its functionality no longer includes encrypting files and demanding ransom. Ransomware now follows a series of well-orchestrated attack stages, designed to remain as stealthy as possible before escalating privileges, moving laterally, exfiltrating data, and encrypting it to cause maximum disruption to the victim organization. Each phase aims to go undetected as long as possible while exerting maximum pressure on the victim. The transformation

has rendered real-time defence and early detection more difficult. This complexity challenges traditional detection methods, especially signature- or static-code-based methods. Ransomware can easily evade static analysis through techniques such as obfuscation, packing, polymorphism, and encryption. As a result, studies have increasingly focused on the runtime rather than code structure. With the many ways of detecting malicious intent during runtime rather than at compile time, such as API calls, DLL use, mutex creation, registry interactions, and

memory interactions, dynamic analysis has become commonplace [1].

This runtime-monitoring focus led to the acceptance of API call sequences as features for ransomware detection. Over time, learning models can learn from API traces to detect malicious patterns. Many of these approaches, however, treat APIs as mere names and fail to account for the parameters sent to these calls or the contextual knowledge that makes them meaningful. Recent research indicates that adding granular-level information and Cyber Threat Intelligence indicators is crucial for improving theft detection. At the same time, the field has seen the emergence of highly effective classification models thanks to machine learning and deep learning. This has led to high detection accuracy, with convolutional networks, recurrent architectures, graph-based models, and even vision transformer-based designs used to achieve the highest possible accuracy on experimental datasets [3]. Nonetheless, these models often become black boxes. They can identify malicious behavior, but they cannot explain why, making them less useful in operational environments where analysts need to understand the "why" behind alerts. In response, some recent studies have begun incorporating Explainable Artificial Intelligence (XAI) methods into detection systems. Tools like SHAP and LIME enable analysts to see the behavioral characteristics that influenced a detection decision, making automated systems more transparent and increasing analysts' trust in the system [4].

However, another critical area of research is ransomware detection prior to encryption. Ransomware performs an API frenzy during its preparation stage, which can serve as an early indicator of malicious activity. Because of limited data and a short time window, some studies resort to generative models, such as WGAN-GP, to augment API data and make the model more robust. These works also go beyond detection, mapping observed behaviors to defenses using structured frameworks such as MITRE ATT&CK and D3FEND [4], to name a few. Generally, ransomware detection deep learning tools are short on rapid development and

face the emergence of zero-day variants, as noted in the reviews.

Overall, deep learning tools for ransomware detection are plagued by zero-day variants and the proliferation of Ransomware-as-a-Service (RaaS) platforms, as observed in the reviewed articles. With RaaS, attackers without technical expertise can run advanced campaigns, fueling the rapid advancement of ransomware and undermining conventional security measures [5]. Additionally, threat-centric analyses of ransomware using the MITRE ATT&CK framework reveal that ransomware is not a single malicious action, but a series of tactics and techniques. Real-world campaigns like REvil show that attackers follow multiple attack phases—from initial access to impact—and that it is necessary to examine sequences of events rather than individual events and to correlate behavior [6].

While significant strides have been made in these research areas, the literature remains disjointed. Most studies focus on one of the following areas: telemetry collection, learning models, early detection, explainability, defensive reasoning, and adversary analysis. Few attempts link these parts into a comprehensive, real-time detection and response process. This fragmentation motivates this survey. It does not rehash past ransomware detection work using algorithms or datasets. Instead, it offers a new way to view ransomware detection research through a holistic detection-reasoning perspective, starting with telemetry collection and progressing through behavioral correlation, intelligent detection, explainability, and response guidance. Reading the literature through this lens helps us see how the contributions interrelate and identify key gaps [7].

2. Systematic Review Methodology

2.1 Review Methodology

To ensure transparency, reproducibility, and unbiased selection of relevant studies, this survey followed systematic literature review (SLR) principles. The review process aimed to discover, assess, and summarize state-of-the-art research on ransomware detection pipelines focusing on behavioral telemetry, feature representation, machine learning, explainable artificial Intelligence (XAI), and automated responses. The overall review process

involved five stages: search strategy, study selection, eligibility assessment, quality evaluation, and data extraction. The literature survey followed the PRISMA process, as shown in Fig. 1 [8].

2.2 Search Strategy

We conducted an extensive literature search across major scientific databases, including IEEE Xplore, Scopus, Web of Science, ScienceDirect, SpringerLink, and Google Scholar. These databases were chosen because they include most peer-reviewed publications on cybersecurity, machine learning, and ransomware detection. We searched only publications from 2019 to 2025 to ensure the latest advancements in ransomware detection technologies were included. The search used keyword combinations with different Boolean operators. Search terms were mainly the following [9]:

1. "Ransomware Detection"
2. "Behavioral Ransomware Detection"
3. "API-based Ransomware Detection"
4. "Machine Learning AND Ransomware"
5. "Deep Learning AND Ransomware"
6. "Explainable AI AND Ransomware"
7. "Zero-Day Ransomware"
8. "Behavioral Monitoring"
9. "MITRE ATT&CK"
10. "Cyber Threat Intelligence"

2.3 Inclusion Criteria

The inclusion criteria ensured that only relevant, high-quality studies were included in this review. Publications had to be peer-reviewed journal articles or conference papers, in English, and published in the last 7 years (2019-2025). In addition, the chosen research must cover the detection, behavioural analysis, or any cybersecurity techniques related to machine learning, deep learning, explainable artificial Intelligence (XAI), behavioural monitoring, telemetry analysis, or automated response mechanisms for ransomware. We excluded studies with inadequate methodological and experimental details that would not allow meaningful comparisons. [10].

2.4 Exclusion Criteria

Several exclusion criteria ensured the review's consistency and relevance. Publications that did not

specifically discuss ransomware but discussed generic malware detection were excluded. We eliminated duplicate data from multiple databases, and excluded editorials, posters, theses, workshop summaries, and short papers because they usually did not provide extensive technical data. Moreover, research that did not clearly explain its approach or was not related to runtime behavioral analysis was not considered for further analysis. [11].

2.5 Quality Assessment

After the initial screening, each eligible study underwent quality assessment to determine its scientific value and relevance to the goals of this survey. Several factors were considered, including the transparency of the proposed methodology, experimental validation, descriptions of the datasets used, the performance evaluation metrics reported, and the research limitations discussed. Special emphasis was also placed on each study's applicability to the proposed ransomware detection pipeline. Only studies that met these quality criteria were kept for comparative analysis at the end of the process [12].

2.6 Data Extraction

Relevant information was systematically extracted after the final selection of eligible studies for qualitative and comparative analysis. The collected data included publication year, selected detection method, key sources of telemetry data, behavioral representation methods, machine learning models/deep learning models used, explainability approaches, response capabilities, evaluation metrics, and key limitations noted by the authors. These extracted characteristics formed the basis for the proposed taxonomy and the comparative analysis presented in the following sections [13],[14].

3. Taxonomy of Ransomware Detection Research

Although algorithms, datasets, and experimental settings vary across the reviewed studies on ransomware detection, a similar research pattern emerges. Most studies propose solutions that address individual functional stages of the ransomware detection process rather than proposing entirely different approaches. However, many papers address different aspects of the same problem, which vary

across algorithms, data sets, and experimental setups [15]. Some researchers focus solely on which program data to monitor during runtime. Others ask how to represent this data in a way machine learning can understand. Much of the literature has focused on improving classification precision by using more complex deep learning models.

However, more recent Work has questioned the value of high accuracy in the absence of accompanying explanations. Fewer studies explore the implications of detection results for defence action [16] and [17]. These directions can seem

disjointed when read individually. However, when viewed as a whole and considered in context, they form a coherent series of steps that begins with raw behavioral data and concludes with informed, defensive action. This observation suggests that ransomware detection research studies should not be classified by algorithm (e.g., ML vs DL), but rather by their position in this chain of reasoning [18]. Fig. 1 illustrates the proposed functional taxonomy of ransomware detection research, organizing existing studies according to their role within the end-to-end detection pipeline.

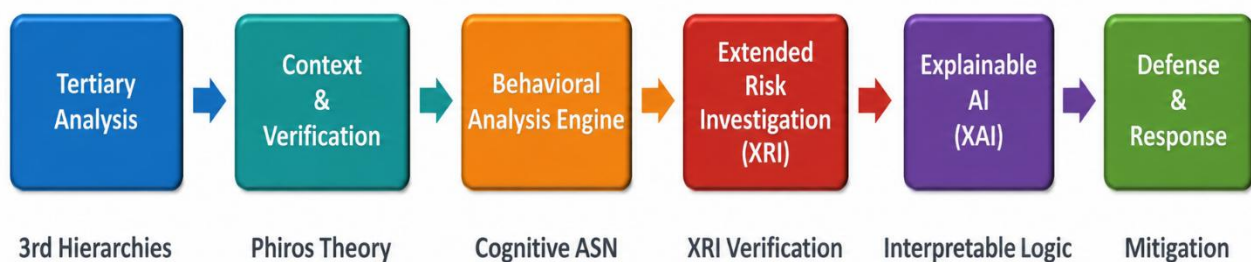


Fig. 1 A functional view of ransomware detection research across the detection pipeline.

The existing studies can be naturally subdivided into six functional layers from this perspective: 1. Collecting data during the process; 2. Representing and building features within behaviour. 3. Machine- and deep learning–based detection; 4. Pre-encryption detection, early detection, and zero-day. 5. Explainability of the detection decisions [19]. The reasoning for post-detection defense and response. This view helps to appreciate that many studies seem unrelated yet are complementary.

1. Collecting telemetry data during execution.
2. Behavioral representations, feature construction.
3. Machine and deep learning–based detection
4. Detection prior to encryption, including early and zero-day detection.
5. Explainability of detector decisions
6. Post-detection defense and Response reasoning

3.1 Telemetry Acquisition

The first set of studies aims to obtain behavioral information about ransomware at runtime, enabling accurate characterization of its behavior. Typical telemetry sources include API call sequences, loaded DLL behavior, mutex creation, registry changes, file system activity, and execution logs generated by a

sandbox. These runtime artifacts contain important information about the ransomware's behavior and are a key source for many behavioral detection frameworks [20]. Numerous studies have shown that sequences of API calls are highly effective at capturing ransomware-operating system interactions. API calls provide strong behavioral evidence, as they reflect program flow and can be used by machine learning and deep learning models to distinguish malicious activities from benign software. API names alone, however, usually do not give a complete picture of the execution context. Recent research has therefore stressed the use of API parameters, file paths, registry keys, process identifiers, and even Cyber Threat Intelligence (CTI) information to enhance semantic understanding and detection capability [21],[22].

Other research has gone beyond monitoring API calls and included behavioral data such as DLL loading, mutex creation, registry operations, and process execution traces. Multi-source telemetry helps the detection system provide a more complete picture of ransomware behavior than any single telemetry source. Likewise, telemetry data collected early, before the files are encrypted, can be sufficient to provide behavioral evidence that helps identify

ransomware before the files are irreversibly encrypted [23] [24]. However, several challenges remain. Most current methods have been tested in an offline or lab setting and do not apply to operational scenarios. Moreover, many studies still rely on a single telemetry source, which may not be robust enough to detect advanced ransomware that employs multiple evasion techniques. These observations suggest that future ransomware detection systems should combine multiple telemetry types and operate continuously in real time, improving ransomware behavior visibility, detection reliability, and deployment operations [25],[26].

3.2 Behavioral Representation and Feature Construction

After runtime telemetry is gathered, the research involves creating concrete representations from raw behavioral logs that are useful for machine learning models. N-grams, TF-IDF, embeddings, and graph-based modeling are typical feature representation methods. These techniques map low-level behavioral events to structured representations to identify ransomware characteristics. Another study reports that using several different behavioral sources - API calls, DLL activity, and Change of Mutex operations, among others - also better characterizes the behavior of ransomware than relying on a combination of one source only [27],[28].

In ransomware detection, existing behavioral representation methods have made significant strides by enabling the use of raw, unstructured telemetry data and bringing important behavioral features to the forefront. However, many strategies rely on manually designed features or unimodal behavioral representations, which can limit generalization across a wide range of ransomware families and sophisticated attack methods. In addition, manually designed features can miss complex relationships among the various types of behavioral occurrences. Hence, future Work should explore new multimodal behavioral representation approaches that can learn more complex semantic relationships and are more robust to new ransomware variants [29] and [30].

3.3 Machine and Deep Learning Detection

Today, machine learning and deep learning are the most prevalent methods for classifying ransomware

behavior because they can automatically learn deep behavioral patterns from runtime telemetry. Different approaches, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Bidirectional Long Short-Term Memory (BiLSTM) networks, Graph Neural Networks (GNNs), and Transformer-based networks, have been successfully used to distinguish ransomware from benign software. They have shown high detection accuracy across a wide range of benchmark datasets, and their ability to detect behavioural ransomware has been consistently reported.

However, they have several drawbacks. While benchmark datasets are frequently used to evaluate many machine learning and deep learning models, they are often collected in a "clean" lab environment, making their applicability in real-world settings difficult to determine. Moreover, high detection accuracy does not guarantee practical deployment, because many deep learning models are black boxes with limited explanations for their decisions. This makes model interpretability, better generalization, and real-time deployment important challenges for future ransomware-detection research.

3.4 Early and Zero-Day Detection

Thus, ransomware detection has increasingly piqued interest, particularly at the early stage, before encryption causes damage. In this context, a few studies have focused on ransomware behavior during the non-encryption phase (file enumeration, privilege escalation, process initialization, abnormal API calls, etc.) before encryption. Given the naturally limited amount of behavioral data available during this phase, many studies have leveraged data augmentation to boost model robustness and detection performance, such as using generative models. Moreover, some solutions combine behavioral analysis with cybersecurity information bases and threat intelligence systems to enable proactive defence and mitigation strategies [33].

Early detection is a major defence against the potential effects of ransomware attacks; however, it is difficult because behavioural analysis is not as advanced, with only minimal clues in the early stages of the ransomware attack lifecycle. Additionally, zero-day ransomware variants exhibit behavior that

would not be reflected in models trained on past data. These challenges indicate that adaptive learning techniques, continuous behavioral monitoring, and integration of real-time threat intelligence data can improve the ability to detect new ransomware attacks at an early stage [34].

3.5 Explainable AI

In ransomware detection systems, the focus has shifted to the need for explainable artificial Intelligence (XAI) to enhance transparency and analyst trust. Current state-of-the-art machine learning and deep learning models achieve impressive detection performance but do not provide insights into how they make decisions, hindering their use in operational cybersecurity settings. To address this challenge, explainability methods such as Shapley Additive Explanations (SHAP) and LIME (Local Interpretable Model-agnostic Explanations) have been widely used to identify the behavioural features most relevant to ransomware classification. These approaches help security analysts understand model predictions and support informed decision-making in security incident investigations [35].

While XAI significantly increases the interpretability of ransomware detection models, current studies have mostly used explainability to explain classification model results after classification is complete. Thus, explainability remains in its nascent stage and is predominantly used as a post-hoc analysis tool. Accordingly, future research should focus on closer integration of Explainable AI, a behavioral detection system, and an automatic response system to improve the operation of ransomware defense systems [36].

3.6 Post-Detection Defense and Response

Beyond ransomware detection, recent studies have also focused on how ransomware detection data can be leveraged to fuel defense strategies. These approaches include using observed behavioural indicators, correlating them with relevant cybersecurity knowledge frameworks such as MITRE ATT&CK, CAPEC, and D3FEND, and using ontology-based reasoning systems to suggest

appropriate mitigation actions. These frameworks help security analysts gain insight into adversarial tactics, techniques, and procedures (TTPs), enabling more effective incident response and threat intelligence analysis [37],[38].

While response-oriented approaches offer useful guidance on preventing ransomware attacks, most existing studies assume malicious behaviour has already been observed and do not include a detection system. This means that detection and response are typically handled separately and may not be part of a ransomware defense pipeline. Thus, future research should aim to develop an end-to-end framework that combines the four steps mentioned above, along with the other steps, to provide an intelligent system that can support real-time operational cybersecurity.

4. Related Work Organized by the Proposed Taxonomy

After proposing the taxonomy outlined in the previous section, the reviewed studies can be systematically analyzed based on their main contributions to the ransomware detection pipeline's functionality. Instead of categorizing literature solely by machine learning algorithms or experimental datasets, the proposed taxonomy classifies existing works by the functional stage each Work covers, such as telemetry acquisition, behaviour representation, intelligent detection, early and zero-day detection, explainable artificial intelligence, and post-detection defence and response [39].

This taxonomy provides a structured view of the various studies and their place in the ransomware detection process. While each study focuses on a single research goal, these studies can be considered complementary parts of an end-to-end detection pipeline. This functional framework helps identify current research trends, strengths and weaknesses, and research gaps. Therefore, the reviewed studies were categorized according to the primary contribution identified within the proposed taxonomy, as shown in Table 1, which serves as a basis for the comparative observations presented in the subsequent subsections [40].

Table 1: Taxonomy-Based Classification of Reviewed Ransomware Detection Studies

Ref.	Primary Telemetry	Behavioral Representation	Learning Model	Pre-Encryption Focus	Explainability	Defence / Response Link	Key Limitation Identified
[1] REDDS	API sequences	N-gram, TF-IDF	ML + WGAN-GP	Yes	No	Yes (Ontology)	Limited to API-only view
[2] CTIMD	API + parameters + CTI	Semantic embedding	BiLSTM	No	No	No	No Real-time deployment
[3] xRank	API + DLL + Mutex	Multi-sequence correlation	DL	No	Yes (SHAP/LIME)	No	No response integration
[4] Vitt-based detection	API behavior	Attention representation	Vision Transformer	No	No	No	Black-box decision process
[5] DL Review	—	—	—	—	—	—	Highlights the zero-day weakness of DL
[6] Zero-day detection	API behavior	Statistical features	ML	Yes	No	No	Dataset dependency
[7] Early API burst	API sequences	N-gram modeling	ML	Yes	No	No	Limited behavioural window
[8] API behavioural modelling	API calls	Sequence embedding	DL	No	No	No	Single telemetry source
[9] Explainable DL	API behavior	Feature importance	DL	No	Yes	No	No telemetry richness
[10] REvil ATT&CK	—	TTP mapping	—	—	—	Yes	No detection mechanism
[11] RaaS study	—	Threat evolution analysis	—	—	—	Indirect	No detection framework
[12] Hospital attacks	Logs + API	Incident correlation	—	No	No	Yes	Case-specific analysis
[13] API semantic study	API + parameters	Semantic modeling	DL	No	No	No	No correlation layer
[14] Swarm intelligence	API behavior	Feature optimization	ML	No	No	No	Limited explainability
[15] Generative AI	API early traces	Data augmentation	DL	Yes	No	No	No reasoning layer
[16] RDCO ontology	—	API→ATT&CK mapping	—	—	—	Yes	No detection model
[17] API injection mapping	API + DLL	ATT&CK correlation	—	No	No	Yes	No ML integration
[18] DL API sequences	API	Embedding	DL	No	No	No	Dataset-bound performance

[19] TF-IDF/N-gram study	API	Statistical encoding	ML	Yes	No	No	Weak semantic meaning
[20] Zero-day mitigation	API	Feature learning	DL	Yes	No	Yes	No explainability
[21] Intelligent detection	API + behavior	Correlated features	DL	No	No	No	No response link
[22] Explainable behavior	API	Feature attribution	DL	No	Yes	No	No early detection
[23] API+DLL study	API + DLL	Correlated sequences	DL	No	No	No	No XAI
[24] Early decision	API early traces	Statistical modeling	ML	Yes	No	No	Limited scope
[25] API call detection	API	Sequence learning	DL	No	No	No	No correlation or XAI

4.1 Studies Focusing on Telemetry Acquisition

Runtime telemetry acquisition underpins behavioral ransomware detection and, in turn, affects the quality and diversity of the behavioral data collected and, correspondingly, the results of subsequent analysis and classification steps. Therefore, many researchers have focused on identifying the most informative runtime artefacts for discriminating between a malicious category (e.g., ransomware) and a benign one. API call sequences are the most commonly used across diverse telemetry sources, as they expose the interaction between an executing program and the operating system and can provide behavioral clues for machine-learning-based detection [41].

Several studies have shown that ransomware actions can be effectively captured using only API call sequences. However, later studies have shown that the API names alone are not always enough to infer malicious intent. At runtime, for instance, the CTIMD framework can add runtime parameters to API traces (file paths, registry keys, process IDs, Cyber Threat Intelligence (CTI), etc.), providing a much greater and more nuanced semantic analysis of API calls rather than them being simple operations performed on the system [7] [9]. This contextual enrichment allows the representation of ransomware behaviour. It gives learning models a better ability to distinguish malicious behaviour from legitimate application behaviour [42].

Other studies have expanded telemetry beyond API monitoring. The xRank framework has combined API calls and DLL-loading patterns, showing that multiple telemetry sources offer complementary evidence of behaviour that can only be captured by combining these two types of telemetry, not by either alone [3]. Likewise, in early detection, students have cited a need to capture behavioral telemetry during the brief pre-encryption period, as useful signs of ransomware intent can be detected before the actual encryption process begins and thus serve as a basis for proactively mitigating ransomware attacks [43, 44].

Existing studies suggest a gradual shift from source-API communication to more fully fleshed-out multi-source behavioral observations for telemetry acquisition. However, while API-based telemetry remains a strong foundation for ransomware detection, adding behavioral sources often enhances its effectiveness by providing additional context for ransomware exploration and increasing resilience against evasion techniques used by more advanced ransomware variants. However, several current telemetry acquisition methods remain in a lab setting and often lack real-time telemetry collection frameworks suitable for operational ransomware defense [45].

Studies Focusing on Behavioral Representation and Feature Engineering

After acquiring telemetry, these studies address the issue of providing a meaningful representation of raw behaviour data for use by the machine learning model. The feature representation is an intrinsic part of the overall success of the ransomware detection system, as runtime logs are extensive, unstructured, and heterogeneous. Therefore, several behavioral encoding approaches have been proposed, including statistical representations and semantic embedding methods [46].

Also, to adapt API call sequences to conventional machine learning algorithms, several studies use statistical feature extraction methods such as N-grams or TF-IDF. As shown in the REDDS framework, with appropriate behavioral datasets and data augmentation methods, these relatively simple representations can provide highly reliable ransomware detection, particularly at the pre-encryption stage [47, 48]. The results suggest that the choice of encoding technique influences the effectiveness of behavioral representation, as does the quality and diversity of the underlying telemetry.

More recent Work has focused on a semantic-behavioral representation that reflects relationships between behavioral events in context, beyond just the frequency of behaviors. For instance, the CTIMD model uses embedding methods to map API calls into a semantic feature space, enabling it to learn semantic relationships among API calls [49] [50]. Similarly, multi-source approaches like xRank integrate API calls, DLL load events, and mutex operations into a single behavioral representation that describes ransomware behavior from multiple complementary perspectives.

Overall, the studies in this review show a shift from statistical to more semantic and multimodal behavioral features. While traditional feature engineering methods are often fast and effective for detecting certain types of ransomware, semantic representations typically provide deeper understanding. They are more effective at adapting to new ransomware variants. However, a major challenge remains in creating generalized behaviour representations that can capture heterogeneous runtime activities [51].

4.3 Studies Focusing on Machine and Deep Learning Detection Models

Finally, after creating behavioral features, models must be trained to discriminate between malicious and legitimate software behavior in the ransomware detection pipeline. The literature reviewed indicates a shift from basic machine learning methods to more sophisticated deep learning architectures that can learn complex behaviour patterns from runtime telemetry data.

Previous Work mainly used traditional classifiers with handcrafted behavioral features. More recent research has introduced deep neural architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Bidirectional Long Short-Term Memory (BiLSTM) networks, Graph Neural Networks (GNNs), and Transformer models. Dietz showed that the Vision Transformer method works well on behavioral traces as structured patterns. With attention mechanisms, it is possible to classify ransomware with very high accuracy. Similarly, some studies on deep learning summarized in Table 1 report detection accuracies above 99% when tested on benchmark behavior datasets [52].

Although these studies show promise, they share several common issues. Evaluating the robustness of most learning models against previously unseen ransomware families is hard since they are trained on and evaluated using controlled benchmark datasets. Moreover, high predictive accuracy does not imply effectiveness in real-world scenarios, especially where model decisions are incomprehensible to cybersecurity analysts. This has led to the exploration of approaches to increase model transparency and generalization, such as explainable AI and adaptive learning techniques in behavioral science.

Overall, the evaluated works show that the backbone of current ransomware detection systems is the deployment of machine learning (and, to a lesser extent, deep learning) algorithms as the primary decision-making module. However, their practical success will depend heavily on high-quality telemetry, accurate behavioral representation, and mechanisms for interpretability and Real-time deployment [53].

4.4 Research on Early and Pre-Encryption Detection

Most traditional ransomware detection models aim to detect ransomware attacks after they have started encrypting files. However, research is increasingly focused on detecting ransomware during the brief preparation phase before the attack. This research direction is inspired by the fact that ransomware typically performs various preliminary operations, such as checking the running environment, acquiring administrative privileges, initiating the process, and enumerating active files before encrypting files. These activities are strong behavioral indicators that can support proactive detection and mitigation. The body of Work defined by the acronym REDDS is among the most representative, as it intentionally captures behavioral telemetry during the pre-encryption phase and shows that it provides sufficient evidence. As mentioned above, there is not much behavioral evidence at this stage, and several studies use augmentation methods, such as WGAN-GP, to improve model robustness and handle small amounts of evidence. Other early-detection work also relies on short sequences of behavior at key stages of execution to show that spotting malicious intent is often more useful than assessing damage after encryption.

Looking back, you will see that identifying ransomware early lets you defend against it without having to react. However, malicious intent will be difficult to accurately recognize before encryption, as limited behavioral evidence will be available at that stage. Clearly, there is still a lot of room for this research to improve behavioral modelling techniques at the outset and to enhance resistance to the emergence of new ransomware versions.

4.5 Studies Focusing on Explainability and Behavioral Interpretation

As ransomware detection models become more sophisticated, explainability has become a core component of an operational cybersecurity system. The problem with deep learning models is that they deliver amazing classification results but are very difficult for analysts to understand, making them unusable for decision-making. For this, some studies

have included explainable artificial Intelligence (XAI) techniques to increase transparency and support behaviour interpretation.

The explainability techniques used in xRank, such as SHAP and LIME, reveal the behavioural characteristics that drive ransomware requests' most significant impact and help analysts understand the effects that API calls, DLL loads, and mutex behaviour, among other features, have on model predictions [55]. These techniques provide human-comprehensible explanations that not only help validate deep learning models but also enable incident investigation, rather than treating these models as black boxes. Explainability also helps researchers determine whether detection results reflect valid ransomware activity or coincidental patterns in the training data.

However, the examined sources show that explainability is mostly used as a post hoc analysis tool after performing classification. Key research focuses only on incorporating explainability into the detection process. In the future, explainability, as an integral part of a comprehensive ransomware defense, should be considered throughout ransomware behavior analysis in the pipeline rather than

4.6 Studies on the Linkage between Detection, Defense, and Response

In addition to detecting ransomware, recent studies delve deeper into how ransomware detection results can be used and how that information can best be applied to a defensive response. Beyond malware classification, these approaches aim to derive relationships between observed behavioral evidence and structured cybersecurity knowledge to support incident response, threat intelligence analysis, and automated mitigation. Ontology-based approaches, such as RDCO, anchor behavioral indicators to knowledge bases, including MITRE ATT&CK, CAPEC, CWE, and D3FEND, enabling security analysts to infer which defensive actions to take against observed behaviors [57]. Likewise, by analyzing ransomware variants such as REvil and examining their behavior in relation to adversarial tactics, techniques, and procedures (TTPs), it is

possible to develop more informed response strategies [58].

While these studies cover much ground between behavioral analysis and cyber defense, a common theme is that they generally assume an external detection mechanism has already detected a ransomware attack. As a result, detection and response remain largely separate research fields. Therefore, combining telemetry acquisition, behavioral analysis, intelligent detection, explainable artificial intelligence, and automated response in an end-to-end ransomware detection pipeline is one of the most promising directions for future research.

5. Comparative Analysis of Ransomware Detection Approaches

The taxonomy-based classification presented in Table 2 enables a comparative evaluation of the major research directions in behavioral ransomware detection. Rather than considering each approach independently, this comparison highlights how different studies address complementary stages of the ransomware detection pipeline. Examining these approaches collectively reveals their respective strengths, practical limitations, and unresolved challenges, thereby providing a clearer understanding of current research trends and opportunities for future development [59].

Table 2: Comparative Analysis of Major Ransomware Detection Approaches

Direction	Typical Input	Main Strength	Practical Limitation	What It Misses
API Telemetry Studies	API calls/parameters	Detailed behavioral visibility	Often sandbox-based	No decision reasoning
Behavioral Correlation	API + DLL + Mutex	Captures attacker intent	Complex feature design	No response link
Statistical Features (N-gram/TF-IDF)	API sequences	Effective in early traces	Weak semantic meaning	Limited generalization
Deep Learning Models	Behavioral vectors	High detection accuracy	Black-box decisions	Needs explainability
Early Detection	Pre-encryption traces	Prevents file damage	Limited data window	Needs augmentation
Explainable AI	Model outputs	Transparent reasoning	Added after detection	Not integrated in the pipeline
ATT&CK / Ontology	Observed behaviors	Guides defensive action	No detection mechanism	Needs telemetry + ML

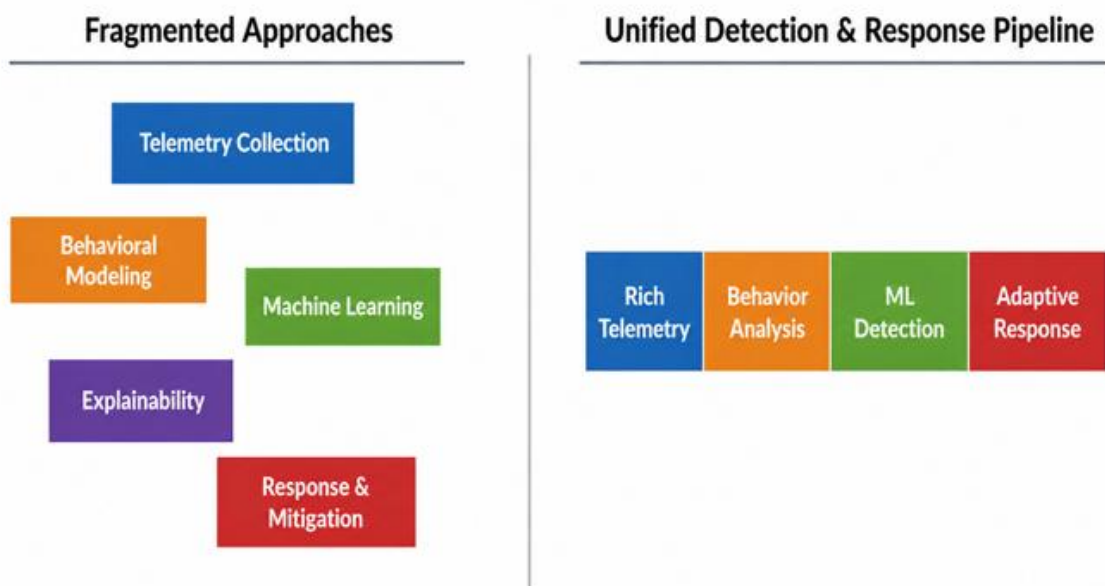


Fig. 2 Fragmentation of existing ransomware detection research and the need for an integrated detection–response pipeline.

As summarized in Table 2, no single research direction provides a complete solution for ransomware detection. While each approach addresses a specific aspect of the detection process, important capabilities such as explainability, early detection, and automated response remain fragmented across different studies. This fragmentation underscores the need for integrated ransomware detection pipelines that combine these complementary components within a unified framework. Fig. 2 illustrates how existing research directions remain functionally disconnected despite addressing complementary stages of the ransomware detection pipeline.

5.1 Telemetry Visibility vs. Decision Logic

Some studies focus primarily on the program implementation process. They carefully monitor API calls, verify parameter values, and occasionally include DLL and mutex operations to understand runtime activity better. These works can provide ample insight into ransomware behavior once it is in the system. They generally only describe behaviour. They do not explicitly state how to use this information to make a detection decision. Conversely, model-oriented studies assume that, if behavioral features are provided, learning algorithms will do the rest—thus separating observation from decision-making. Hence, effective ransomware detection requires close integration between runtime behavioral monitoring and intelligent decision-making mechanisms [60].

5.2 Counting Behavior vs. Understanding Behavior

Some works use statistical methods, such as N-grams and TF-IDF, to represent API traces. These techniques measure API co-occurrence. When applied to pre-encryption traces, these techniques perform surprisingly well! Other research trends include embedding APIs in a space that represents behavioral relationships, not just frequency [61]. These methods are more portable and better at handling new behavior as they learn relationships

between APIs. This distinction indicates a progression from counting to understanding relationships between events. These results suggest that semantic behavioral representations are more adaptable than purely statistical feature-encoding approaches, especially when facing new ransomware families.

5.3 Observing Damage vs. Preventing Damage

Dynamic analysis traditionally requires monitoring ransomware behavior for much longer periods. Early-detection studies suggest this is not necessary. They also show that, within the short period before encryption begins, sufficient behavioral evidence exists to detect the intent of a ransomware attack [62]. However, traces are short and limited at the beginning; hence, data augmentation is performed. This comparison reflects differences in the importance of observation timing and the nature of the behaviors being observed. Therefore, improving the quality of ransomware behavioral indicators for early detection is critical to accurate zero-day detection.

5.4 Accuracy vs. Interpretability

Deep learning experiments indicate that training models with appropriate behavioral features can achieve high detection accuracy. However, these models are not very informative in explaining their choices. Explainability: Explainability approaches aim to address this by showing which features led to a classification [63]. However, explainability is generally not used in detection reasoning; it is usually used afterwards. This can result in correct systems that remain ununderstood. In the future, ransomware detection systems should not only ensure predictive performance but also provide a clear understanding of their underlying reasoning to support their practical implementation.

5.5 Detection vs. Defensive Action

Most machine learning techniques stop at ransomware detection. In contrast, ontology-based and ATT&CK-compatible studies focus on providing the technical details of the behaviour to be measured. These works use the detection results but do not

explain how detection itself should be performed. These comparisons—along with the data in Table 2 and the information in the graphic in Fig. 2 demonstrate that ransomware detection research is akin to a collection of well-developed components that are rarely assembled into a system. Each piece usually responds to a need, yet the relationships between them rarely exist. This inevitably raises the outstanding research issues to be discussed in the following section.

6. Research Gaps in Existing Ransomware Detection Studies

From a comparative point of view, Part 5 highlights a disjointed solution, not a poor one— the current ransomware detection research displays this (Table 2 and Fig. 2). Although every research direction can solve a certain part of the problem, there are still

reference gaps that can be filled and cannot be used to establish a practical and unified detection system. These are named in this section. From a comparative perspective, Part 5 emphasizes a disjointed solution rather than a poor solution. As reported in the above-mentioned summarized studies, current research on ransomware detection still has some practical drawbacks that are not addressed, such as the absence of an end-to-end ransomware detection pipeline, the absence of real-time telemetry integration, weak ransomware behavioral correlation, dependency on historical ransomware detections, the absence of explainability in detection at runtime, etc. While these are separate research areas related to ransomware detection, none can create a unified, operational framework. The major gaps in the ransomware detection pipeline are summarized conceptually in Fig. 3.

Table 3: Identified Research Gaps Across Ransomware Detection Directions

Gap ID	Identified Gap	Evidence in Literature	Practical Consequence
G1	Lack of Real-time telemetry integration	Most studies rely on sandbox traces	Difficult deployment in live systems
G2	Weak correlation between behavioral sources	API is often studied separately from DLL/Mutex	Incomplete behavioral view
G3	Dependence on historical datasets	DL models trained on similar data	Weakness against zero-day variants
G4	Explainability added after detection	XAI used post-classification	Decisions not interpretable during runtime
G5	Detection not linked to response	ATT&CK/RDCO without detection model	The gap between detection and mitigation
G6	Absence of an end-to-end pipeline	Each study focuses on a single layer	No practical integrated framework

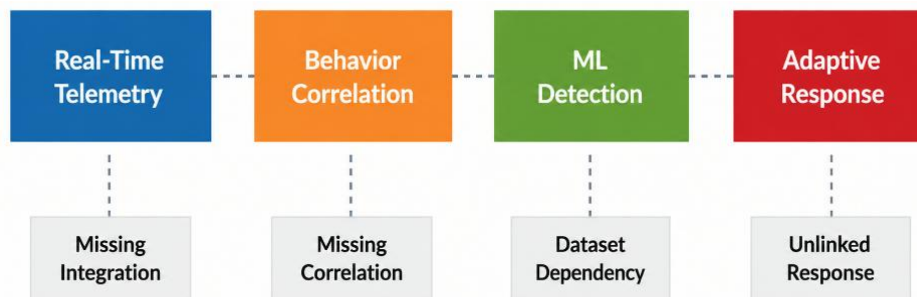


Fig. 3 Research gaps across the ransomware detection pipeline.

6.1 Gap G1 — The Forgotten Problem of Real-time Telemetry Integration

The literature surveyed indicates that many ransomware-detection papers rely on behavioral

traces collected in a managed sandbox environment. This is sensible from an experimental point of view, as sandboxes enable clean, reproducible execution traces, making data collection and evaluation easier. This practice, however, "accidentally leaves a research environment and one of operation separate from one another," she wrote. In real systems, telemetry data cannot be analyzed within a well-structured API trace. Instead, orchestration appears to create a series of events by monitoring, analyzing, and using operating system logs and endpoint sensors. This real-time data capture, filtering, and correlation is rarely discussed in detection research. Several works review monitoring APIs and how they can be used to detect ransomware behaviour [65].

However, they do not explain how to incorporate monitoring into a system that processes several thousand events per second. There is a significant difference between analyzing pre-recorded API traces and processing telemetry streams, both technically and computationally. That is because the models used for detection are often tested in an environment other than the one they will be used in. Little, however, is known about the performance of these models when embedded in real-time monitoring platforms.

6.2 Gap G2 — Behavioral Sources are Rarely Correlated in Real Operational Contexts

Another trend in the literature is the independent study of various behavioral sources. Many works report the sequence of API calls as the primary indicator of a ransomware's activity. Others extend this Work to examine DLL loading patterns and mutex activity. A smaller group attempts to use these sources to make multi-sequence representations. These techniques are clearly designed to demonstrate each source's ability to provide useful information, but are normally conducted in controlled settings, with the collection of traces where actions and sayings are aligned and available for analysis.

This is not rotaviral data; it is received across operational systems, which are treated as separate and independent data streams. In practice, it is difficult to correlate multiple heterogeneous streams in real time. This involves identifying related events, determining how to sync them, and preserving their temporal relationships. While a few papers acknowledge that

incorporating such sources can improve understanding of behaviour, detection studies rarely discuss the associated challenges. Thus, the literature clearly shows that multi-source behavioral correlation is relevant; however, it offers little insight into how to achieve it in a live monitoring environment. This gap between the theory and practice of multi-sequence analysis [66] needs to be addressed [66].

6.3 Gap G3 — Dependence on Historical Behavioral Patterns and the Zero-Day Problem

Detection approaches have consistently been shown to be highly accurate when trained and tested on available datasets using deep learning. Results indicate that learning models can detect ransomware when it is similar to previously encountered ransomware. However, a specific study on zero-day ransomware detection indicates that zero-day variants are not rare in ransomware attacks and that models trained on historical data may be unable to capture them if their semantic information is not explicitly used as features. In essence, many detection methods assume that future ransomware activity will look just like past ransomware activity. A fair assumption is that attackers will use incremental variants; however, when they attempt to adapt their approach to fly under the radar, this can be problematic. This affects not only model accuracy but also model adaptability. While many effective classifiers exist, few studies examine how they can evolve to adapt to newly observed behaviors of unseen ransomware families [67].

6.4 Gap G4 — Explainability Is Separated from the Detection Process: The ability to explain the detection process is separated from the detection itself.

In many recent studies, explainable AI is cited as a means of explaining model decisions. These works show that building analysts' trust in the system and ensuring the model learns meaningful patterns can be strengthened by knowing which APIs or behavioral aspects were used to make a decision. However, explainability is often tacked on as a post-processing step. The model is grounded in the detection process and then uses the explanation model to describe the

results. This gap is substantial and may result from this separation. If explainability is not part of the detection reasoning, it appears that there is no direct conflict between the notions of explainability and decision making at runtime. Rather, it only provides explanations for decisions made. Few studies in the literature provide a more integrated approach in which interpretability is considered a component of the detection logic. [68].

6.5 Gap G5 — Detection Results are Rarely Linked to Immediate Defensive Action

Among the common findings in the literature reviewed is "Most detection studies stop at classification. If a model identifies a ransomware attack, the system's Work is done. What to do next is left to a human analyst or another security device. Meanwhile, other research focuses on correlating ransomware activity with protection measures. Some API patterns can be correlated with particular remediation processes, such as those involving ATT&CK tactics or ontology-driven reasoning. Once ransomware activity is better understood, these studies can yield practical tips and tricks. This gap arises because these two directions rarely intersect. Detection-oriented works typically lack response logic, and vice versa: response-oriented works do not normally contain detection logic. As a result, there is no correlation between detecting undesirable behaviour and responding appropriately and defensively. This separation can significantly delay environmental response time. This is because, although a detection is not always directly linked to action, it still needs to be interpreted for mitigation to occur [69].

6.6 Gap G6 — The Absence of an End-to-End Detection and Response Pipeline

The greatest difference is evident only collectively, in the context of all preceding gaps. In the literature, telemetry collection, behavioral correlation, machine learning detection, explainability, and reasoning for response are all considered distinct research problems. Each is studied in depth and often in a very comprehensive manner. However, very little has been done to place them in a common operational framework that binds them into a continuous, logical reasoning. This makes ransomware detection

research seem like a set of old wrenches that are rarely put together. Very little Work has been done to integrate these components into a single pipeline—one for telemetry, another for modelling, another for interpretation, and another for response. It is not a defect of any particular technique. It is about consolidating existing techniques into a centralized, real-time framework that provides detection and an actionable response. [70].

7. Discussion

7.1 Evolution of Ransomware Detection Research

The systematic analysis discussed in the previous Sections has shown that ransomware detection research has evolved significantly. Initial research efforts focused on tracking applications' runtime API calls and using traditional Machine Learning algorithms to detect malicious calls and distinguish them from benign ones. As ransomware attacks have grown more sophisticated, research has increasingly focused on richer behavioral representations, deep learning architectures, explainable artificial Intelligence (XAI), and ontology-driven response mechanisms. The progression illustrates a gradual shift from detection methods to a broader behavioral analysis framework. However, the literature reviewed suggests that research along these paths has tended to develop separately, rather than as part of a holistic system to protect against ransomware attacks.

7.2 Current Challenges in Ransomware Detection

Although huge strides have been made in detecting ransomware behaviour, practical issues remain. First, most studies to date test their methods on benchmark datasets collected in laboratory settings, making it difficult to determine how they perform in operational conditions. Secondly, most detection models continue to rely on past behavioral patterns, which they struggle to withstand when faced with a new ransomware variant. Moreover, explainability methods have greatly enhanced the transparency of deep learning algorithms. However, they are usually added after the detection phase rather than being part of runtime decision-making. Lastly, limited research links intelligent detection to automated defensive responses, resulting in disjointed cybersecurity solutions [71].

7.3 Future Research Directions

Based on the results of this survey, we believe several directions may indicate future research on ransomware detection. Future frameworks should integrate multiple, heterogeneous telemetry sources in real time to provide more comprehensive behavioral analysis and enhance situational awareness. Moreover, adaptive learning algorithms that dynamically update detection models should be explored to achieve resilience against emerging zero-day ransomware attacks. Explainable AI must move from post hoc interpretation to an integral part of runtime decision-making and automated response mechanisms, tightly coupled with behaviour detection using cybersecurity knowledge frameworks such as MITRE ATT&CK and D3FEND. One of the biggest end-to-end ransomware detection pipeline advancement trends that we identified in our survey is the ability to build unified end-to-end ransomware

detection pipelines, from acquisition to representation, to intelligent detection (ID), explainability, and automated response.

8. Future Directions

The research gaps uncovered in the previous section are not to be misconstrued as an indication of a lack of ransomware detection research. Instead, they suggest that considerable progress has been made in individual research areas and that relatively insufficient effort has gone into synthesizing these efforts and advances into integrated operational concepts. Based on these research gaps, the future directions proposed in this section aim to support the development of unified, practical, and intelligent ransomware detection systems. [72]. As shown in Fig. 4, future ransomware detection systems should integrate telemetry acquisition, behavioral correlation, intelligent detection, explainability, and automated response into a unified operational pipeline.



Fig. 4 Conceptual direction for integrating telemetry, detection, explainability, and response into a unified framework.

8.1 Moving from Offline Traces to Continuous Telemetry Streams

Many ransomware detection research initiatives have used API traces collected in sandboxes. These traces are clean, clear, well-structured, and easy to analyze, making them very convenient for experimental evaluation. However, in practice, there is an obvious downside to consider when thinking about how detection systems would function in the real world. Telemetry is not sent in the correct order of API calls for analysis in operational systems. Rather, monitoring tools, operating system logs, and endpoint sensors generate a series of events. Noise, synchronicity, and relationships among these events are all unrelated [73].

The selection of events of interest, their grouping, and the rapid processing of these events for detection are issues not frequently addressed in the literature. Several studies clearly indicate which APIs and behaviors are associated with ransomware activity.

However, they do not explain how to perform this monitoring continuously in a live system that typically generates thousands of events per second.

Analyzing recorded traces is technically and computationally very different from processing live telemetry. Future studies, therefore, should not simply take humans' attitudes and behaviors for granted and should seek other ways to obtain this data. Instead, detection systems should be designed in this context, recognizing that telemetry is messy, continuous, and time-sensitive. Use ransomware detection solutions that are effective in the field, not in the lab.

8.2 Designing Systems That Correlate Behavioral Sources in Real Time

API calls, DLL activity, and mutex behaviour have been found to collectively provide a better indication of ransomware intent in multiple studies. Viewed in isolation, these sources are less informative than when analyzed together in pre-prepared datasets. In

real monitoring scenarios, however, they are not delivered in a structured pattern. API events, system interactions, and process actions are all developed individually and at different times. What is needed is a method for unwinding these streams into a sensible mix in which the same events are clearly identified as belonging to the same execution context and temporal relationships are clearly defined.

This is a practical problem that is rarely addressed in detection research, which focuses on creating a new detector. Most works assume that behavioral sources conform to the model. It is therefore evident that although the literature suggests that exploiting multi-source correlation can improve detection quality, little guidance is available on how to obtain it when events occur dynamically. Future systems need to be more than just recorders of telemetry sources. They should be able to link these events in real time and develop a logical sequence of actions over time, rather than just logs. Without this, it is difficult to restore the theoretical benefits of multi-sequence analysis in practice.

8.3 Developing models of building detection that adapt to behavioral change

Many studies on learning-based ransomware detection assume that future behaviour will be the same as that observed in the provided ransomware data sets. This works for incremental variants, where attackers tweak only a few aspects of their strategies. However, it is less useful if ransomware families drastically change to avoid detection. Several studies on zero-day variants have discussed this issue. They show that models trained on past API patterns can fail when these patterns are modified within a sequence, even when the intention remains malicious. This shows that models capture superficial patterns rather than deeper behavioral relationships. This is where the difference between pattern recognition and understanding of behaviour comes into play. Heavy reliance on frequency or sequence similarity can pose issues if attackers' execution flows vary.

However, other approaches that involve APIs as components of semantic spaces or consider relationships between different action contexts have greater potential for generalization. Future detection systems should therefore focus less on remembering "what" each did and more on capturing relationships

between actions. This requires adaptive learning approaches that enable models to adapt as new behavioral patterns are detected, rather than relying on a predefined training dataset that contains only past information.

8.4 Linking detection to defensive actions

A few works on ATT&CK techniques and ontology-driven reasoning have been found, demonstrating correlations between behaviors and mitigations. These studies offer a glimpse into how to prevent ransomware when it is known to be occurring. Less is known about how this knowledge connects directly to detection systems. Typically, detection and response are viewed and treated as separate phases, with different tools or different detection capabilities handled by human analysts. Using automated behaviour to guide response can help future detection frameworks. A system encountering such ransomware behaviour should notify the user, together with one or more suggestions, or initiate action based on the detected behaviour. This would help reduce the time from detection to mitigation and make detection a proactive measure [74].

8.5 Toward a Unified Detection-Response Pipeline

If any of these directions is pursued alone, it will form only a partial picture. Telemetry monitoring, behavioral modelling, machine learning identification, interpretability, and defensive reasoning are already good ideas in the literature. That is not a problem of lacking techniques; it is a problem of lacking integration between them. In the next couple of generations, defense against ransomware will not be a series of modules but a pipeline, with these as its stages. The structured input from behavioral correlation mechanisms would feed adaptive learning models, and telemetry collection would feed behavioral correlation mechanisms. Such models would be easy to understand, and their decisions would inform response guidance. This type of pipeline does not need to generate new ideas. Instead, it has to 'assemble' the knowledge already available in the literature into a coherent framework for use in operational contexts [75].

9. Conclusion

This survey provides a thorough overview of recent ransomware detection research using a pipeline-based approach, systematically reviewing existing studies based on their contributions to the ransomware detection system. The proposed taxonomy differs from conventional ones, which organize the literature according to algorithms or datasets, because the six stages are all interwoven: the telemetry acquisition; the behavioral representation; the detection using machine learning and/or deep learning; the early and zero-day detection; the explainable artificial Intelligence (XAI); and post-detection defence and response. This presents a clear functional structure for the individual research efforts, which collectively contribute to the growth of "smart" ransomware defense.

After systematically analyzing and comparing them, we found significant progress at each stage of the ransomware detection pipeline. However, the literature studied indicates that these research directions have proceeded largely independently of one another. Despite promising detection performance in individual studies, critical challenges remain, including Real-time integration of telemetry from various sources with cloud services, correlating the behaviour of incoming and existing threats, resistance to as-yet-undetected ransomware, explainability at runtime, and the use of intelligent telemetry with automatic responses.

Based on these observations, this survey identified six major research gaps and discussed their implications for practical research on future ransomware detection systems. Moreover, it outlines a compendium of potential future research priorities to promote the creation of end-to-end systems that integrate telemetry data collection with behavior analysis, adaptive learning, explainable AI, and automated response capabilities.

Overall, this survey shows that, for ransomware defence, refining individual detection methods is not enough: it is essential to build efficient, scalable detection–response frameworks that integrate existing technologies and fit real-world cybersecurity operations.

Acknowledgment

The author thanks the College of Engineering, University of Diyala, for their support.

Conflict of Interest

The author declares no conflict of interest.

References

- [1] Alqahtani, A., & Sheldon, F. T. (2022). A survey of crypto ransomware attack detection methodologies: An evolving outlook. *Sensors*, 22(5), 1837.
- [2] Chen, T., Zeng, H., Lv, M., & Zhu, T. (2024). CTIMD: Cyber threat intelligence enhanced malware detection using API call sequences with parameters. *Computers & Security*, 136, Article 103518.
<https://doi.org/10.1016/j.cose.2023.103518>
- [3] Ferdous, J., Islam, M. R., Mahboubi, A., & Islam, M. Z. (2024). AI-based ransomware detection: A comprehensive review. *IEEE Access*, 12, 136666–136695.
- [4] Ashwini, K., & Naga Sundara, K. B. (2024). An intelligent ransomware attack detection and classification using a dual vision transformer with Mantis Search Split Attention Network. *Computers & Electrical Engineering*, 119, Article 109509.
<https://doi.org/10.1016/j.compeleceng.2024.109509>
- [5] Kritika, E. (2025). A comprehensive literature review on ransomware detection using deep learning. *Cyber Security and Applications*, 3, 100078.
- [6] Cui, B., Hu, Y., Qu, T., He, Y., & Sun, L. (2026). A novel zero-day ransomware detection approach based on CVAE and 1D-CNN. *High-Confidence Computing*, 6(1), Article 100338.
<https://doi.org/10.1016/j.hcc.2025.100338>
- [7] Zhao, D., Wang, H., Kou, L., Li, Z., & Zhang, J. (2023). Dynamic malware detection using a parameter-augmented semantic chain. *Electronics*, 12(24), Article 4992.
<https://doi.org/10.3390/electronics12244992>
- [8] Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71.

- [9] Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339.
- [10] Lasky, J., Naughton, C., McPherson, P., Arnoth, E., & Zhang, C. (2022). Ransomware: Getting started guide and deep dive into REvil. MITRE. <https://www.mitre.org/news-insights/publication/ransomware-getting-started-guide-and-deep-dive-revil>
- [11] Kitchenham, B., & Charters, S. (2022). Guidelines for performing systematic literature reviews in software engineering.
- [12] Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2023). Systematic mapping studies in software engineering.
- [13] Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, 51(1), 7–15.
- [14] Okoli, C. (2015). A Guide to Conducting a Standalone Systematic Literature Review. *Communications of the Association for Information Systems*, 37.
- [15] Alzahrani, S., Xiao, Y., Asiri, S., Zheng, J., & Li, T. (2025). A survey of ransomware detection methods. *IEEE Access*, 13, 57943–57982.
- [16] Alvi, M. A., & Jalil, Z. (2025). XRGuard: A model-agnostic approach to ransomware detection using dynamic analysis and explainable AI. *IEEE Access*, 13, 53159–53170.
- [17] Lasky, J., Naughton, C., McPherson, P., Arnoth, E., & Zhang, C. (2022). Ransomware: Getting Started Guide and Deep Dive into REvil. MITRE.
- [18] Alzahrani, S., Xiao, Y., Asiri, S., Zheng, J., & Li, T. (2025). A survey of ransomware detection methods. *IEEE Access*, 13, 57943–57982.
- [19] Manthena, H., Shajarian, S., Kimmell, J., Abdelsalam, M., Khorsandroo, S., & Gupta, M. (2025). Explainable Artificial Intelligence (XAI) for Malware Analysis: A Survey of Techniques, Applications, and Open Challenges. *IEEE Access*, 13, 61611–61640.
- [20] Herrera-Silva, J. A., & Hernández-Álvarez, M. (2023). Dynamic feature dataset for ransomware detection using machine learning algorithms. *Sensors*, 23(3), 1053.
- [21] Chen, T., Zeng, H., Lv, M., & Zhu, T. (2024). CTIMD: Cyber threat intelligence enhanced malware detection using API call sequences with parameters: *Computers & Security*, 136, 103518.
- [22] Zhao, D., Wang, H., Kou, L., Li, Z., & Zhang, J. (2023). Dynamic malware detection using a parameter-augmented semantic chain. *Electronics*, 12(24), 4992.
- [23] Khan, Z. U. A., Mughal, M. A., Hashmi, M. U., Sarwar, R., & Ul Haq, I. (2026). Lightweight and explainable early ransomware detection using dynamic API-call features and ensemble machine learning. *Engineering Reports*, 8(4), e70734.
- [24] Cui, B., Hu, Y., Qu, T., He, Y., & Sun, L. (2026). A novel zero-day ransomware detection approach based on CVAE and 1D-CNN. *High-Confidence Computing*, 6(1), 100338.
- [25] Alhuwayshil, S., Ramachandran, S., & Kim, K. (2025). Enhancing ransomware threat detection: Risk-aware classification via Windows API call analysis and hybrid ML/DL models—*Journal of Cybersecurity and Privacy*, 5(4), 96.
- [26] Begovic, K., Al-Ali, A., & Malluhi, Q. (2023). Cryptographic ransomware encryption detection: Survey. *Computers & Security*, 132, Article 103349. <https://doi.org/10.1016/j.cose.2023.103349>
- [27] Kritika, E. (2025). A comprehensive literature review on ransomware detection using deep learning. *Cyber Security and Applications*, 3, Article 100078. <https://doi.org/10.1016/j.csa.2024.100078>
- [28] Ferdous, J., Islam, M. R., Mahboubi, A., & Islam, M. Z. (2024). AI-based ransomware detection: A comprehensive review. *IEEE Access*, 12, 136666–136695. <https://doi.org/10.1109/ACCESS.2024.3461965>
- [29] Ispahany, J., Islam, M. R., Islam, M. Z., & Khan, M. A. (2024). Ransomware detection using machine learning: A review, research limitations and future directions. *IEEE Access*, 12, 68785–68813. <https://doi.org/10.1109/ACCESS.2024.3397921>

- [30] Nguyen, P. S., Huy, T. N., Tuan, T. A., Trung, P. D., & Long, H. V. (2025). Hybrid feature extraction and integrated deep learning for cloud-based malware detection. *Computers & Security*, 150, 104233.
- [31] Onwuegbuche, F. C., Adelodun, S. O., Jurcut, A. D., & Pasquale, L. (2026). MLRan: A behavioural dataset for ransomware analysis and detection. *Journal of Network and Computer Applications*, 250, Article 104475. <https://doi.org/10.1016/j.jnca.2026.104475>
- [32] Khan, Z. U. A., Mughal, M. A., Hashmi, M. U., Sarwar, R., & Ul Haq, I. (2026). Lightweight and explainable early ransomware detection using dynamic API-call features and ensemble machine learning. *Engineering Reports*, 8(4), Article e70734. <https://doi.org/10.1002/eng2.70734>
- [33] Davidian, M., Kiperberg, M., & Vanetik, N. (2024). Early Ransomware Detection with Deep Learning Models. *Future Internet*, 16(8), 291.
- [34] Thummapudi, K., Lama, P., & Boppana, R. V. (2023). Detection of ransomware attacks using processor and disk usage data. *IEEE Access*, 11, 51395–51407. <https://doi.org/10.1109/ACCESS.2023.3279819>
- [35] Alraizza, A., & Algarni, A. (2023). Ransomware detection using machine learning: A survey. *Big Data and Cognitive Computing*, 7(3), Article 143. <https://doi.org/10.3390/bdcc7030143>
- [36] Alqahtani, A., & Sheldon, F. T. (2022). A survey of crypto ransomware attack detection methodologies: An evolving outlook. *Sensors*, 22(5), Article 1837. <https://doi.org/10.3390/s22051837>
- [37] Berrueta, E., Morato, D., Magaña, E., & Izal, M. (2022). Crypto-ransomware detection using machine learning models in file-sharing network scenarios with encrypted traffic. *Expert Systems with Applications*, 209, Article 118299. <https://doi.org/10.1016/j.eswa.2022.118299>
- [38] Tang, F., Ma, B., Li, J., Zhang, F., Su, J., & Ma, J. (2020). RansomSpector: An introspection-based approach to detect crypto ransomware. *Computers & Security*, 97, Article 101997. <https://doi.org/10.1016/j.cose.2020.101997>
- [39] Basnet, M., Poudyal, S., Ali, M. H., & Dasgupta, D. (2021). Ransomware detection using deep learning in the SCADA system of an electric vehicle charging station. In 2021, the IEEE PES Innovative Smart Grid Technologies Conference—Latin America (ISGT Latin America). IEEE. <https://doi.org/10.1109/ISGTLatinAmerica52371.2021.9543031>
- [40] Vehabovic, A., Ghani, N., Bou-Harb, E., Crichigno, J., & Yayimli, A. (2023). Ransomware Detection and Classification Strategies. *arXiv:2304.04398*.
- [41] Alzahrani, S., Xiao, Y., Asiri, S., Zheng, J., & Li, T. (2025). A survey of ransomware detection methods. *IEEE Access*, 13, 57943–57982. <https://doi.org/10.1109/ACCESS.2025.3556187>
- [42] Gulmez, S., Gorgulu Kakisim, A., & Sogukpinar, I. (2024). XRan: Explainable deep learning-based ransomware detection using dynamic analysis. *Computers & Security*, 139, Article 103703. <https://doi.org/10.1016/j.cose.2024.103703>
- [43] Kabuye, H., Issac, B., Yumlembam, R., & Neera, J. (2025). Explainable, uncertainty-aware AI-based ransomware detection. *IEEE Access*, 13, 106573–106589. <https://doi.org/10.1109/ACCESS.2025.3581424>
- [44] Davidian, M., Kiperberg, M., & Vanetik, N. (2024). Early Ransomware Detection with Deep Learning Models. *Future Internet*, 16(8), 291.
- [45] Manthena, H., Shajarian, S., Kimmell, J., Abdelsalam, M., Khorsandroo, S., & Gupta, M. (2025). Explainable artificial intelligence (XAI) for malware analysis: A survey of techniques, applications, and open challenges. *IEEE Access*, 13, 61611–61640. <https://doi.org/10.1109/ACCESS.2025.3555926>
- [46] Lin, S.-C., Yang, T.-Y., & Zhang, B.-H. (2024). Ransomware Detection by Distinguishing API Call Sequences through LSTM and BERT Models. *The Computer Journal*, 67(2), 632–641.
- [47] Davidian, M., Kiperberg, M., & Vanetik, N. (2024). Early Ransomware Detection with Deep Learning Models. *Future Internet*, 16(8), 291.

- [48] Mowri, R. A., Siddula, M., & Roy, K. (2022). A comparative performance analysis of explainable machine learning models with and without the RFECV feature selection technique towards ransomware classification. *IEEE Access*, 10, 130180–130195.
<https://doi.org/10.1109/ACCESS.2022.3228041>
- [49] Hirano, M., & Kobayashi, R. (2022). Machine learning-based ransomware detection using low-level memory access patterns obtained from a live-forensic hypervisor. In *2022 IEEE International Conference on Cyber Security and Resilience (CSR)* (pp. 323–330). IEEE.
<https://doi.org/10.1109/CSR54599.2022.9850340>
- [50] Gulmez, S., Gorgulu Kakisim, A., & Sogukpinar, I. (2024). XRan: Explainable Deep Learning-Based Ransomware Detection Using Dynamic Analysis. *Computers & Security*, 139, 103703.
- [51] Alhuwayshil, S., Ramachandran, S., & Kim, K. (2025). Enhancing ransomware threat detection: Risk-aware classification via Windows API call analysis and hybrid ML/DL models—*Journal of Cybersecurity and Privacy*, 5(4), 96.
- [52] Gulmez, S., Gorgulu Kakisim, A., & Sogukpinar, I. (2024). XRan: Explainable Deep Learning-Based Ransomware Detection Using Dynamic Analysis. *Computers & Security*, 139, 103703.
- [53] Manthena, H., Shajarian, S., Kimmell, J., Abdelsalam, M., Khorsandroo, S., & Gupta, M. (2025). Explainable Artificial Intelligence (XAI) for Malware Analysis: A Survey of Techniques, Applications, and Open Challenges.
- [54] Lin, S. C., Yang, T. Y., & Zhang, B. H. (2024). Ransomware Detection by Distinguishing API Call Sequences through LSTM and BERT Models. *The Computer Journal*, 67(2), 632–641.
- [55] Omar, C., Kannouf, N., & Mohammed, B. (2024). A systematic review and taxonomy of ransomware detection based on artificial intelligence algorithms. The 3rd International Conference on Embedded Systems and Artificial Intelligence (ESAI) will be held in 2024. IEEE.
- [56] Alhuwayshil, S., Ramachandran, S., & Kim, K. (2025). Enhancing ransomware threat detection: Risk-aware classification via Windows API call analysis and hybrid ML/DL models—*Journal of Cybersecurity and Privacy*, 5(4), Article 96.
- [57] Zhou, C., Liu, Y., Meng, W., Tao, S., Tian, W., Yao, F., & Yang, H. (2025). SRDC: Semantics-Based Ransomware Detection and Classification with LLM-Assisted Pre-training. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(27), 28566–28574.
- [58] Ferdous, J., Islam, M. R., Mahboubi, A., & Islam, M. Z. (2024). AI-based ransomware detection: A comprehensive review. *IEEE Access*, 12, 136666–136695.
<https://doi.org/10.1109/ACCESS.2024.3461965>
- [59] Zhou, C., Liu, Y., Meng, W., Tao, S., Tian, W., Yao, F., & Yang, H. (2025). SRDC: Semantics-Based Ransomware Detection and Classification with LLM-Assisted Pre-training. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(27), 28566–28574.
<https://doi.org/10.1609/aaai.v39i27.35080>
- [60] Chauhan, P., & Teotia, S. (2025). Real-Time Behavioral Analysis for Ransomware Response: A Framework Leveraging Machine Learning and Threat Intelligence Feeds. *International Journal of Performability Engineering*, 21(11), 639–650.
<https://doi.org/10.23940/ijpe.25.11.p4.639650>
- [61] Herrera-Silva, J. A., & Hernández-Álvarez, M. (2023). Dynamic feature dataset for ransomware detection using machine learning algorithms. *Sensors*, 23(3), Article 1053.
<https://doi.org/10.3390/s23031053>
- [62] Youssef, N., Elbaraway, N., & Elmaghraby, A. (2025). Transformer-based API call sequence modelling for dynamic malware detection. In *SoutheastCon 2025* (pp. 494–500). IEEE.
<https://doi.org/10.1109/SoutheastCon56624.2025.10971600>
- [63] Iwan, W. A., Suryani, V., & Yulianto, F. A. (2025). Real-time ransomware detection on the host system using machine learning and windowing thresholds. In the *2025 International Conference on Data Science and Its Applications (ICoDSA)*. IEEE.
<https://doi.org/10.1109/ICoDSA67155.2025.11157409>

- [64] Nguyen, P. S., Huy, T. N., Tuan, T. A., Trung, P. D., & Long, H. V. (2025). Hybrid feature extraction and integrated deep learning for cloud-based malware detection. *Computers & Security*, 150, Article 104233. <https://doi.org/10.1016/j.cose.2024.104233>
- [65] Zirari, K., Idrissi, H. K., El-Yahyaoui, A., Bensaid, H., & En-Nouaary, A. (2025). Deciphering Ransomware: Strategic API Usage and Behavioral Patterns for Advanced Detection Techniques. *International Journal of Interactive Mobile Technologies*, 19(10), 199–221. <https://doi.org/10.3991/ijim.v19i10.49245>
- [66] Davidian, M., Kiperberg, M., & Vanetik, N. (2024). Early Ransomware Detection with Deep Learning Models. *Future Internet*, 16(8), 291. <https://doi.org/10.3390/fi16080291>
- [67] Daengsi, T., Pornpongtechavanich, P., Boonpoor, P., Wattanachukul, K., Puangnak, K., Phanrattanachai, K., Wuttidittachotti, P., & Horkaew, P. (2026). Ransomware and artificial intelligence: A comprehensive systematic review of reviews. *Bulletin of Electrical Engineering and Informatics*, 15(2), 1848–1865. <https://doi.org/10.11591/eei.v15i2.11107>
- [68] Alzahrani, S., Xiao, Y., Asiri, S., Zheng, J., & Li, T. (2025). A Survey of Ransomware Detection Methods. *IEEE Access*, 13, 57943–57982. <https://doi.org/10.1109/ACCESS.2025.3556187>
- [69] Gurumallu, P. K., Dembala, R., Gowda, D. Y., Annaiah, A. K. M., Kumar, M. M. V., & Gohel, H. (2025). Exploring Deep Learning Approaches for Ransomware Detection: A Comprehensive Survey. *Current Computer Science*. <https://doi.org/10.2174/0126662558279673240515054741>
- [70] Gulmez, S., Gorgulu Kakisim, A., & Sogukpinar, I. (2024). XRan: Explainable Deep Learning-Based Ransomware Detection Using Dynamic Analysis. *Computers & Security*, 139, 103703. <https://doi.org/10.1016/j.cose.2024.103703>
- [71] Chen, T., Zeng, H., Lv, M., & Zhu, T. (2024). CTIMD: Cyber Threat Intelligence Enhanced Malware Detection Using API Call Sequences with Parameters. *Computers & Security*, 136, 103518. <https://doi.org/10.1016/j.cose.2023.103518>
- [72] Alvi, M. A., & Jalil, Z. (2025). XRGuard: A model-agnostic approach to ransomware detection using dynamic analysis and explainable AI. *IEEE Access*, 13, 53159–53170. <https://doi.org/10.1109/ACCESS.2025.3553562>
- [73] Manthena, H., Shajarian, S., Kimmell, J., Abdelsalam, M., Khorsandroo, S., & Gupta, M. (2025). Explainable Artificial Intelligence (XAI) for Malware Analysis: A Survey of Techniques, Applications, and Open Challenges. *IEEE Access*, 13, 61611–61640. <https://doi.org/10.1109/ACCESS.2025.3555926>
- [74] Alhuwayshil, S., Ramachandran, S., & Kim, K. (2025). Enhancing Ransomware Threat Detection: Risk-Aware Classification via Windows API Call Analysis and Hybrid ML/DL Models. *Journal of Cybersecurity and Privacy*, 5(4), Article 96. <https://doi.org/10.3390/jcp5040096>
- [75] Ramírez-Martínez, E. D., Pérez-Díaz, J. A., & Yungaicela, N. M. (2026). Generative AI for ransomware identification and mitigation: Taxonomy, challenges, and future directions. *IEEE Open Journal of the Computer Society*, 7, 1064–1082. <https://doi.org/10.1109/OJCS.2026.3690970>