COS70008 – Technology Innovation Project and Research

Developing a Web Based System for predicting and analysing Malicious Attacks using a Hybrid Machine Learning Model

Assignment – 1

Research Paper Review and Ethics Practices

Student Name: Arun Ragavendhar Arunachalam Palaniyappan

Student ID: 104837257

Date: 20 / 03 / 2025

Contents

Word Count (excluding Table of Contents, References, etc.): 2604

Abbreviations

CPS: Cyber Physical System
AI: Artificial Intelligence
ML: Machine Learning
SVM: Support Vector Machines
DHS: Department of Homeland Security
ICT: Information and Communication Technology
ACS: Australian Computer Society
CISA: Cybersecurity and Infrastructure Security Agency

List of Figures

1.Research Paper Review

1.1 Introduction

The rapid growth of digital technology has transformed society and industry, provided numerous benefits, but, has also introduced significant cybersecurity threats (Alenezi et al., 2020). Increasing dependence on digital platforms has led to more frequent and sophisticated cyberattacks, especially involving malware. Malware is malicious software designed to harm, disrupt, or gain unauthorized access to computer systems, threatening personal data, business operations, and critical infrastructure (Gandhi et al., 2021).

Cyber-physical systems (CPS), which merge digital and physical parts, are critical in industries such as healthcare, manufacturing, transportation, and energy (Chowdhury et al., 2023). A successful attack on these systems can have serious consequences. Malware attacks worldwide have spiked by around 358%, and ransomware alone has climbed by 435% since 2020, highlighting the urgent need for new detection strategies (Cybersecurity Ventures, 2023).

This preliminary literature review examines different malware types, detection methods (traditional and machine learning), and approaches to integrate these methods into a web application (Gandhi et al., 2023; Sharma et al., 2023). The goal is to identify effective techniques, uncover knowledge gaps, blend established methods with innovative improvements and establish a foundation for selecting a suitable cybersecurity solution to build upon (Nataraj et al., 2023).

1.2 Literature Review and Analysis

Malware comes in many types, each spreading and causing harm in its own way. According to Alenezi et al. (2020), the most common examples include viruses, which attach themselves to good programs and copy themselves when those programs run; worms, which make copies of themselves automatically without needing user help; and trojans, which pretend to be safe software so that users install them by mistake. Ransomware is another major type—it locks or encrypts files and demands money to unlock them, with a big example being the Colonial Pipeline attack explained by Beerman et al. (2021).

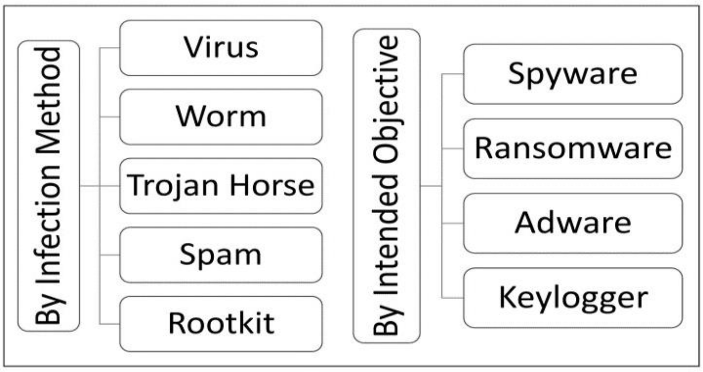

Figure 1: Classification of Malware (Alenezi et al., 2020)

Over the years, different **traditional detection** methods have been used to spot these threats. **Signature-based detection** is quick and accurate for threats that are already known but fails if the malware is totally new (often called a zero-day attack). Nataraj et al. (2023) explain that this approach involves looking for distinct patterns—or "signatures"—in harmful files. Another approach is **sandboxing,** where suspicious files run inside a safe environment so experts can watch for harmful behaviour (Nataraj et al., 2023). While sandboxing is very thorough, it can be time-intensive and use a lot of computing power. Meanwhile, **behavioural analysis** looks at live software activity to catch odd actions that might hint at malware. Although this method finds newly emerging threats, Sharma et al. (2023) note that it can accidentally flag safe programs as harmful.

As a result, usage of **Artificial Intelligence (AI)** and **Machine Learning (ML)** are becoming more common for smarter detection. **Supervised learning** (using Decision Trees, Random Forests, and Support Vector Machines, etc.) is very accurate if it has large, well-labelled datasets to learn from (Chowdhury et al., 2023). In contrast, **unsupervised methods** (such as autoencoders or clustering) excel at spotting strange patterns that might signal a hidden, unknown threat but may have more false positives as well (Lee et al., 2023).

An increasingly popular solution is to use **hybrid ML** models, which combine supervised and unsupervised techniques. These models give high accuracy for known malware and can also detect new or hidden types with reasonable success.

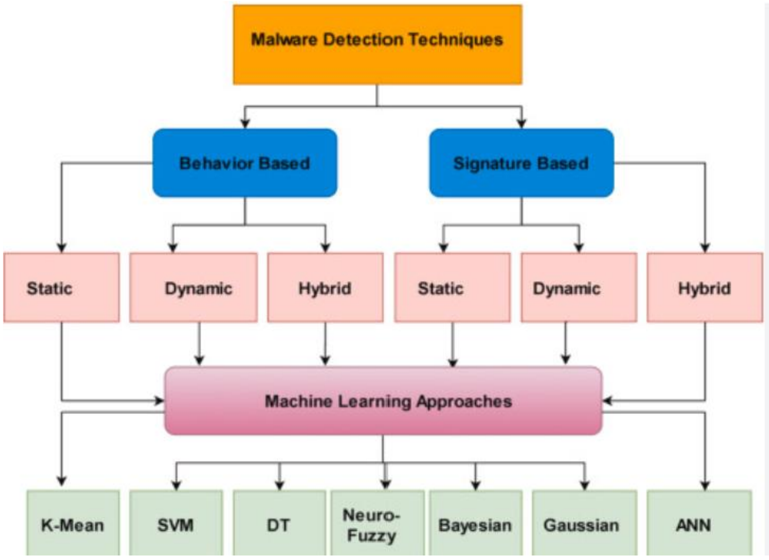

Figure 2: Different Techniques for Malware Detection (Lee et al., 2023)

Turning these detection ideas into a **web-based application** adds another layer of complexity. Roberts et al. (2023) suggest using **Flask (a lightweight Python framework)** because it is simple to set up and works easily with popular ML libraries. On the other hand, Mitchell et al. (2023) highlight the **MERN stack (MongoDB, Express.js, React.js, Node.js)** for its speed and ability to handle real-time data, but note that mixing Python-based AI tools with a JavaScript-driven setup can be tricky.

Finally, integrating classic antivirus or signature-based scanners directly into an app can be fast as it can be integrated directly as a part of the backend or as standalone API (Nataraj et al., 2023). Despite growing evidence that hybrid ML solutions work well, there are still open questions about how these methods perform in real-time and on a large scale, especially when many users are involved.

1.3 Research Methodology

In line with **research books focusing on key design principles** (Creswell, 2014; Kothari, 2004) and standard machine learning guidelines (Bishop, 2006), this study follows a clear, step-by-step plan to investigate malware, review detection tools, and propose an effective method for a real-world web application. Initially, works by Alenezi et al. (2020), Gandhi et al. (2023), Sharma et al. (2023), Nataraj et al. (2023), and Chowdhury et al. (2023) were closely examined to understand how different approaches manage both well-known and emerging malware.

This method review synthesizes and narrows down insights from the literature review by contrasting several non-AI techniques with AI methods for malware detection and examines integration into web applications via Flask-built internal APIs.

A review of traditional (non AI) approaches reveals both strengths and trade-offs: signature-based detection quickly pinpoints recognized threats but lacks the flexibility to handle unfamiliar malware; sandboxing offers rich insights by testing suspicious files in a secure environment, though it demands considerable computing power; and behavioural analysis can spot emerging risks by tracking unusual actions in real time, but often raises false positives by flagging harmless software as malicious.

In contrast, AI-driven methods adapt better to changing threat patterns, particularly when supported by ample labelled data in a supervised learning scenario. Unsupervised methods detect odd behaviours without labelled data but tend to raise more false alarms. Gandhi et al. (2023) and Chowdhury et al. (2023) recommend combining these supervised and unsupervised methods into hybrid ML models to improve accuracy, adapt to unseen malware, and reduce needless alerts. Also, they stress that **pairing strong feature extraction techniques with powerful classification algorithms can help to solve typical issues faced by traditional detection methods** such as lower and inconsistent detection rates, false alarms, and failure to recognise evolved or brand-new malware.

Putting this plan into action starts with gathering a wide set of malware samples from open repositories, then carefully cleaning and preparing those files. Nataraj et al. (2023) explain that both static (analysing file properties without execution) and dynamic analysis (observing software behaviour in controlled environments) approaches are useful for feature extraction, ensuring a complete view of how the malware operates. After that, Sharma et al. (2023) propose testing these methods in a simulation of real-world malware attack scenarios to confirm they are reliable and effective.

From these steps, the **inference is that hybrid AI models deliver strong results by blending supervised classification (for known threats) and unsupervised anomaly detection (for unknown threats).** However, one main obstacle might be the diversity of malware datasets, because using too few or very similar samples can limit the model's overall accuracy and flexibility. Still, the expected finding is that the hybrid ML approach can outperform others in spotting both familiar and never-before-seen attacks, while also limiting false alarms.

2.Ethics Practices

2.1 Case Study Scenarios

This section examines five real-life ICT incidents, describing each event, the ethical issues involved, the ICT activities or professions implicated, and the relevant ACS Code of Ethics requirements.

Colonial Pipeline Ransomware Attack: In 2021, the DarkSide hacking group exploited an unused VPN account at Colonial Pipeline. This security gap allowed attackers to encrypt critical files and shut down fuel supply, resulting in widespread fuel shortages, long gas station lines, and economic disruption. The company paid a $4.4 million ransom even though the decryption tool was ineffective (Beerman et al., 2021; Hall, 2021).

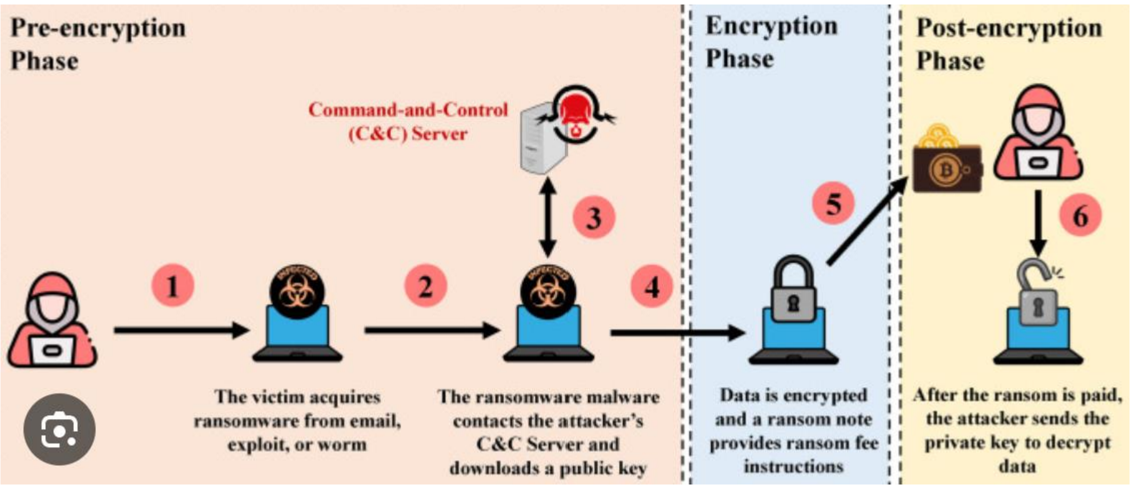
Pre-encryption Phase
Encryption Phase
Post-encryption Phase
Command-and-Control (C&C) Server
1
2
3
4
5
6
The victim acquires ransomware from email, exploit, or worm
The ransomware malware contacts the attacker's C&C Server and downloads a public key
Data is encrypted and a ransom note provides ransom fee instructions
After the ransom is paid, the attacker sends the private key to decrypt data

Figure 3: The effect of the Darkside Ransomware on the Colonial Pipelines CPS (Beerman et al., 2021)

Ilnaz's Workplace Challenge (ACS Case No. 25): Ilnaz, a young professional, was forced to share an office with a male colleague despite her cultural and religious beliefs, leaving her feeling uncomfortable and undervalued (ACS Code of Professional Conduct Case Studies, 2014).

Peter's Religious Practice (ACS Case No. 31): Peter had difficulty finding a quiet space for daily prayers at work due to the absence of a designated area, which affected his well-being (ACS Code of Professional Conduct Case Studies, 2014).

Anna's Accessible Interface (ACS Case No. 32): Anna developed a user interface tailored for remote Aboriginal communities by carefully selecting culturally appropriate images and design elements, ensuring the system was both accessible and user-friendly (ACS Code of Professional Conduct Case Studies, 2014).

Equifax Data Breach (2017): In 2017, cybercriminals exploited an unpatched Apache Struts vulnerability in Equifax's Automated Consumer Interview System. The attackers remained undetected for 78 days and accessed the personal data of approximately 148 million U.S. consumers, including names, Social Security numbers, and birth dates (Kabanov & Madnick, 2020).

2.2 Ethical Dilemma

Each case study brings forth specific ethical dilemmas that require careful consideration:

For the Colonial Pipeline attack, the dilemma involves whether organizations should pay ransoms to restore services quickly or refuse such payments despite potential prolonged disruption, alongside the challenge of communicating openly during crises (Beerman et al., 2021; Hall, 2021). Ilnaz's Workplace Challenge raises the question of whether workplace policies are adaptable enough to honour cultural and religious differences, protecting individual dignity. Peter's Religious Practice spotlights the need for workplaces to accommodate personal religious practices without compromising operational needs. Anna's Accessible Interface underscores the ethical responsibility to design technology that is inclusive and accessible for disadvantaged or vulnerable groups. The Equifax Data Breach presses the need for ethical responsibility to protect sensitive consumer data and ensure prompt disclosure of breaches, thereby maintaining public trust (Kabanov & Madnick, 2020).

2.3 ICT Involvement

ICT professionals and organizations play a key role in resolving these ethical dilemmas:

In the Colonial Pipeline attack, agencies such as the FBI and cybersecurity experts collaborated to trace the breach, secure the network and mitigate its impact, highlighting the importance of coordinated government and private sector responses (Beerman et al., 2021).

In the cases of Ilnaz and Peter, HR professionals, facilities managers, and ICT project teams must collaborate to create work environments that respect cultural and religious needs. Anna's case involves UI/UX designers, software developers, and usability experts who ensure that technology is accessible, inclusive and culturally sensitive.

Once the breach was discovered, Equifax patched the unpatched Apache Struts vulnerability and upgraded their vulnerability scanning processes to cover all directories, ensuring that no critical subdirectories were overlooked. IT managers moved from a manual to an automated SSL certificate renewal process to prevent future lapses, thus restoring the effectiveness of the IDS/IPS in monitoring encrypted traffic. Compliance officers strengthened oversight and enforced stricter checks on both vulnerability management and certificate renewals. They also carried out a restructuring of internal security protocols and leadership changes, aiming at safeguarding sensitive data and rebuilding public trust (Kabanov & Madnick, 2020).

2.4 Application of the ACS Code of Ethics

In each of the case studies, one or more of the Australian Computer Society's (ACS) ethical codes were not followed.

1.2.1 Public Interest

Colonial Pipeline: The breach showed a failure to secure critical public infrastructure, leading to fuel shortages and economic disruption in the south east parts of the US (Beerman et al., 2021; Hall,2021).
Equifax: Negligence in protecting sensitive consumer data compromised nearly 148 million records and severely eroded public trust (Kabanov & Madnick, 2020).

1.2.2 Quality of Life

Ilnaz's Workplace Challenge: Forcing Ilnaz to share an office despite her cultural and religious needs disrupted the quality of her personal life (ACS Code of Professional Conduct Case Studies, 2014).
Peter's Religious Practice: The absence of an appropriate prayer space at work negatively impacted Peter's well-being, showing a disregard for individual cultural requirements (ACS Code of Professional Conduct Case Studies, 2014).

1.2.3 Honesty

Colonial Pipeline: Delayed disclosure of the breach and ineffective communication undermined stakeholder trust and the public perception of the company (Beerman et al., 2021; Hall, 2021).
Equifax: The failure to promptly notify consumers about the breach further compromised transparency and honesty (Kabanov & Madnick, 2020).

1.2.4 Competence

Colonial Pipeline: Overlooking an unused VPN account revealed gaps in technical vigilance, threat and risk assessment and periodic server maintenance schedules (Beerman et al., 2021; Hall, 2021).
Equifax: The vulnerability scanning process failed to cover critical subdirectories, and manual SSL certificate management led to expired certificates, both evidencing significant lapses in professional competence (Kabanov & Madnick, 2020).

1.2.6 Professionalism

Colonial Pipeline: The decision to pay a ransom despite an ineffective decryption tool, along with poor crisis handling, reflects a lack of professional judgment (Beerman et al., 2021; Hall, 2021).
Ilnaz's Workplace Challenge & Peter's Religious Practice: Inadequate accommodations for cultural and religious needs in the workplace indicate a breach of professional responsibility to create respectful, supportive environments (ACS Code of Professional Conduct Case Studies, 2014).
Equifax: The delayed and insufficient response to the breach demonstrates a failure to uphold the highest standards of professionalism in ICT leadership (Kabanov & Madnick, 2020).

Using technology in a responsible way mandates following the ACS Code of Ethics. These mistakes are important lessons to avoid similar issues in future.

2.5 Adopting and maintaining Equity and Accessibility

Inclusive design is essential for fairness and accessibility in systems. ACS Case No. 25 shows Ilnaz facing cultural and religious challenges when required to share an office with a male colleague, highlighting the need to respect diverse norms. Case No. 31 features Peter's struggle to observe his prayers at work, emphasizing religious accommodation. In Case No. 32, Anna designs a UI for remote Aboriginal communities using culturally sensitive visuals, ensuring usability for disadvantaged groups. Case No. 28 shows Katherina's voluntary ICT support for people with disabilities, reflecting its role in improving quality of life. Case No. 24 reveals how inconsistent disability coding in legacy systems hinders policy decisions, stressing the need for accurate, inclusive data. Collectively, these cases underscore the importance of designing systems that promote equity and accessibility (ACS Code of Professional Conduct Case Studies, 2014).

2.6 Conclusion

This research underscores the urgent need for robust cybersecurity measures, spotlighting the broad range of malware types and exploring both time-tested and next-generation detection strategies. A review of five case studies—including those on the Colonial Pipeline ransomware attack, Ilnaz's workplace challenge, Peter's religious practice, Anna's accessible interface, and Katherina's support for disability groups—provided valuable insights into the ethical challenges and learnings from diverse real-world scenarios. After weighing traditional, supervised, unsupervised, and hybrid machine learning methods, it becomes clear that hybrid solutions deliver the strongest combination of accuracy, adaptability, and real-time responsiveness. Events like the Colonial Pipeline ransomware attack and the Equifax data breach demonstrate the high stakes of inadequate defense mechanisms and reinforce the demand for dependable, ethically sound detection tools.

Despite these advantages, several challenges remain. Achieving instant detection without overloading system resources is a considerable hurdle. Additionally, limited dataset diversity can hinder a model's ability to detect novel threats, and scaling small-scale proofs of concept into fully operational developments is often more complex than anticipated. To address these gaps, this study advocates hybrid machine learning as a key strategy, blending supervised and unsupervised approaches within a web-based framework. Technologies such as Flask enable quick deployment and seamless AI integration, tackling many of the pitfalls found in single-method or traditional solutions. At the same time, adhering to the ACS Code of Ethics ensures transparent, responsible data handling, plus alignment with professional standards.

Building on these insights, upcoming research will delve more deeply into project specifics like malware dataset collection, cleaning, and preparation, as well as the specific algorithms and AI model combinations best suited for hybrid approaches. Detailed steps for integrating these solutions into a web application will also be outlined, giving practical guidance on how to move from prototype to full-scale development.

4.References

1. Alenezi, M. N., Alabdulrazzaq, H., Alshaher, A. A., & Alkharang, M. M. (2020). Evolution of malware threats and techniques: A review. *International Journal of Communication Networks and Information Security, 12*(3), 326–334.
2. Beerman, J., Berent, D., Falter, Z., & Bhunia, S. (2021). A review of the Colonial Pipeline ransomware attack. *Cybersecurity Journal, 12*(3), 245–261.
3. Chowdhury, D., Stevens, L., & Grant, P. (2023). Cyber-physical systems anomaly detection using machine learning. *IEEE Transactions on Cybersecurity, 38*(4), 523–541.
4. Cybersecurity Ventures. (2023). *2023 official cybercrime report.* Retrieved from https://cybersecurityventures.com
5. Gandhi, V., Kumar, S., & Kumar, S. (2023). Detection and classification of malware using machine learning techniques.

6. Hall, T. (2021). Examining the Colonial Pipeline ransomware incident and its impact on national security. *International Journal of Cyber Threat Intelligence, 9*(2), 87–105.
7. Lee, C., Wang, H., & Kim, S. (2023). Anomaly-based malware detection using autoencoders and decision trees. *International Conference on Cyber Threats, 28*(1), 215–232.
8. Mitchell, S., Brown, L., & Carter, P. (2023). Evaluating MERN stack for AI integration. *Web Systems and Security Journal, 17*(4), 89–106.
9. Nataraj, L., Yegneswaran, V., & Porras, P. (2023). Dynamic pattern recognition using signature analysis.
10. Reeder, J. R. (2021). Cybersecurity's Pearl Harbor moment: Lessons learned from the Colonial Pipeline ransomware attack. *U.S. Cyber Defense Review, 7*(4), 112–130.
11. Roberts, T., Lee, J., & Adams, R. (2023). Efficiency of Flask in AI model deployment. *International Journal of Web Applications, 34*(2), 101–119.
12. Sharma, P., Kaur, J., & Singh, H. (2023). Malware detection using behaviour analysis.
13. Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
14. Kothari, C. (2004). *Research methodology: Methods and techniques* (2nd ed.). New Age International.
15. Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
16. Australian Computer Society. (2014). *ACS Code of Professional Conduct case studies (Version 2.1)*. Australian Computer Society.
17. Kabanov, I., & Madnick, S. (2020). *A systematic study of the control failures in the Equifax cybersecurity incident* (Working Paper CISL# 2020-19). MIT Sloan School of Management. https://ssrn.com/abstract=3957272