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.): 1990

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 advancement of digital technology has brought remarkable improvements to society and industry, enabling global connectivity, innovation, and efficiency. However, this rapid evolution has also introduced a new dimension of risk in the form of cybersecurity threats. As organizations and individuals increasingly rely on digital platforms, the frequency and sophistication of cyberattacks have surged dramatically (Alenezi et al., 2020). One of the most critical threats among them is malware. Malware is any software intentionally designed to disrupt, damage, or gain unauthorized access to computer systems. These threats not only compromise personal data and business operations but can also cripple critical infrastructure. Cyber-physical systems (CPS), which blend software with physical processes in areas like healthcare, manufacturing, and energy, are particularly vulnerable to these threats (Chowdhury et al., 2023). A single malware breach in such systems can have severe consequences ranging from financial loss to public safety risks.

Recent reports indicate a significant escalation in the volume of attacks, with malware incidents rising by 358% and ransomware alone increasing by 435% since 2020, highlighting the urgent need for new detection strategies (Cybersecurity Ventures, 2023). This literature review explores the nature of malware, traditional and machine learning-based detection methods, and their integration into a real-world web-based application. The objective is to critically assess existing methodologies, identify gaps, and recommend a hybrid model that leverages the strengths of various detection techniques for robust and scalable malware mitigation (Gandhi et al., 2023; Sharma et al., 2023).

1.2 Literature Review and Analysis

Malware exists in multiple forms, each with unique characteristics and propagation techniques. Alenezi et al. (2020) classify malware into several categories, including viruses that attach to legitimate software and replicate when executed, worms that self-replicate across networks without user intervention, and trojans that masquerade as safe programs to deceive users. Among the most damaging forms of malware is ransomware, which encrypts user data and demands payment for decryption. The infamous Colonial Pipeline attack is a stark example, where attackers halted fuel supply across the southeastern United States, causing widespread disruption and financial loss (Beerman et al., 2021).

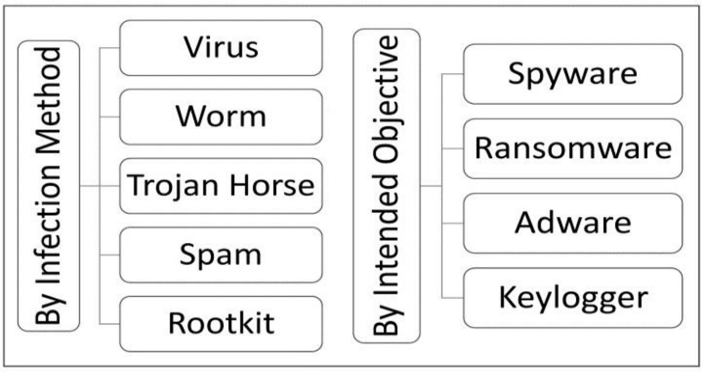

Figure 1: Classification of Malware (Alenezi et al., 2020)

Traditional malware detection techniques have served as the first line of defence for many years. **Signature-based** detection identifies malware by comparing files against a database of known patterns. While this method is fast and effective for previously identified threats, it falls short against zero-day attacks and evolving threats (Nataraj et al., 2023). **Sandboxing** involves executing suspicious files in a controlled environment to observe behaviour. Although thorough, this method demands significant computational resources. **Behavioural analysis**, which monitors system activity for abnormal patterns, provides an additional layer of protection but often suffers from false positives (Sharma et al., 2023).

To address the limitations of traditional approaches, researchers are increasingly turning to **Artificial Intelligence (AI)** and **Machine Learning (ML)**. **Supervised learning** algorithms like Decision Trees, Random Forests, and Support Vector Machines (SVMs) show high accuracy when trained on large, well-labelled datasets (Chowdhury et al., 2023). **Unsupervised** learning techniques, such as clustering and autoencoders, are better suited for detecting anomalies without prior knowledge, although they may produce more false alarms (Lee et al., 2023). A promising advancement is the emergence of **hybrid** models, which combine supervised and unsupervised learning. These models offer a balanced approach: detecting known malware accurately while also identifying novel or evolving threats through anomaly detection mechanisms.

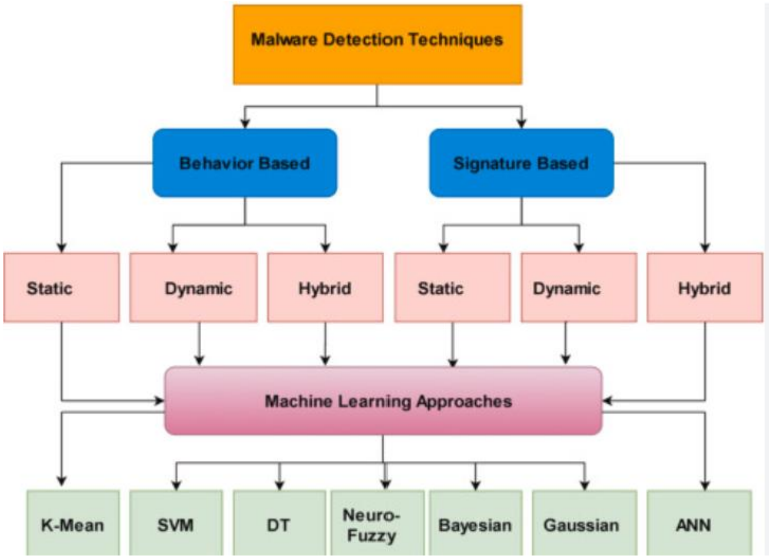

Figure 2: Different Techniques for Malware Detection (Lee et al., 2023)

Integrating these detection systems into a **web-based platform** introduces additional design considerations. Roberts et al. (2023) advocate using **Flask, a lightweight Python framework,** for its simplicity and compatibility with popular ML libraries like Scikit-learn and TensorFlow. In contrast, Mitchell et al. (2023) propose using the **MERN stack (MongoDB, Express.js, React.js, Node.js)** for better performance and real-time interaction, though they caution about the complexity of integrating Python-based ML tools into JavaScript environments. Nataraj et al. (2023) further suggest incorporating existing antivirus engines as APIs or backend components to enhance detection speed.

Despite these advancements, questions remain about real-time efficiency, scalability, and model adaptability in live environments with multiple concurrent users. There is a growing need to balance model complexity and computing resource consumption, especially in systems serving a large user base.

1.3 Research Methodology

Following **key design principles outlined in books** by Creswell (2014), Kothari (2004), and Bishop (2006), this project adopts a structured methodology to explore, analyse, and propose a malware detection system suitable for development as a real-world web application. The approach begins with a comprehensive review of peer-reviewed literature, evaluating both conventional and AI-based detection methods. Key references from Alenezi et al. (2020), Gandhi et al. (2023), Sharma et al. (2023), (Lee et al., 2023) and Chowdhury et al. (2023) serve as the theoretical foundation.

The research proceeds by comparing traditional detection methods against AI-based approaches. Signature-based methods are found to be fast but ineffective against novel threats. Sandboxing provides deep insight but is slow and computationally intensive. Behavioural monitoring offers real-time protection but may result in false positives. AI-based techniques, particularly supervised learning, show higher accuracy when large sets of labelled data is available. Unsupervised learning, while more flexible, can misinterpret benign anomalies as malicious behaviour.

Gandhi et al. (2023) and Chowdhury et al. (2023) recommend hybrid AI models that merge supervised classification with unsupervised anomaly detection. These models can reduce false positives while improving adaptability to unknown malware. Also, they stress that **pairing strong feature extraction techniques with powerful classification algorithms can help to solve typical issues faced by traditional detection methods**. Effective implementation of such models can translate raw software characteristics into meaningful input for ML algorithms.

The experimental phase will begin with the collection of diverse malware datasets from open-source repositories. These samples will undergo preprocessing to remove noise and ensure consistent formatting. Nataraj et al. (2023) emphasize the need for both static analysis, which examines the software without executing it, and dynamic analysis, which observes software behaviour during execution. Sharma et al. (2023) recommend validating models in simulated environments to test detection reliability in scenarios that mimic real-world attacks.

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, challenges include dataset diversity, resource overhead, and generalization across multiple environments. To address this, the proposed system will focus on modularity and efficient deployment using Flask for backend integration.

2.Ethics Practices

2.1 Case Study Scenarios

This section explores five real-world ICT scenarios, analysing the ethical implications, involved stakeholders, and alignment with the ACS Code of Ethics.

The 2021 **Colonial Pipeline Ransomware Attack** involved the DarkSide hacking group exploiting an inactive VPN account. The attack encrypted vital systems and halted operations, leading to fuel shortages and economic disruption. Although the company paid a $4.4 million ransom, the decryption tool proved ineffective (Beerman et al., 2021; Hall, 2021).

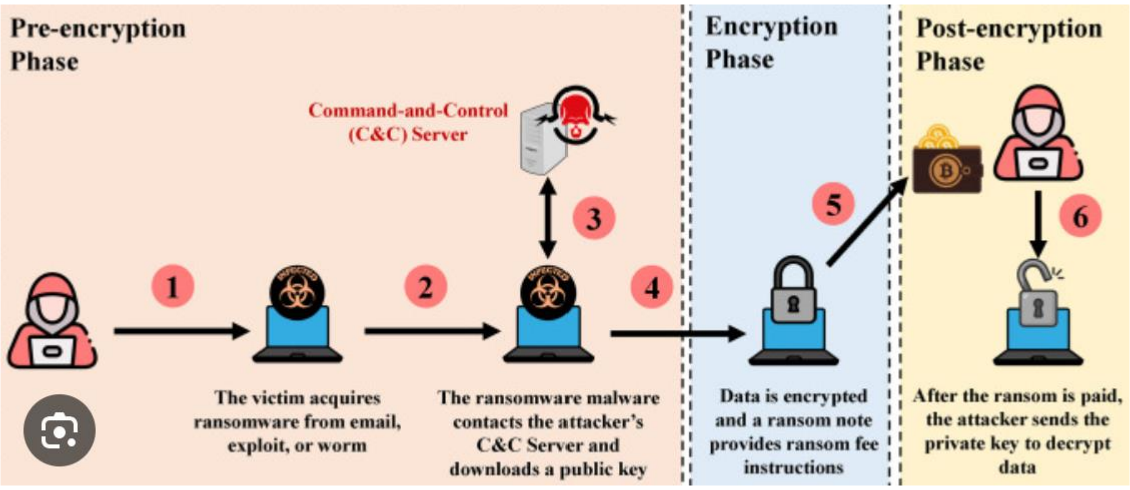
Pre-encryption Phase
Encryption Phase
Post-encryption Phase
Command-and-Control (C&C) Server
1
2
3
4
5
6
The victim acquires ransomware from email, exploit, or worm
The ransomware malware contacts the attacker's C&C Server and downloads a public key
Data is encrypted and a ransom note provides ransom 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 25) highlights cultural insensitivity, where a young female professional was assigned to share an office with a male colleague despite her religious beliefs. This created an uncomfortable work environment and diminished her sense of respect and inclusion (ACS Code of Professional Conduct Case Studies, 2014). **Peter's Religious Practice (ACS Case 31)** concerns a lack of accommodation for prayer time, showing how the absence of a designated space can negatively impact an individual's emotional and spiritual well-being (ACS Code of Professional Conduct Case Studies, 2014). **Anna's Accessible Interface (ACS Case 32)** stands as a positive example of inclusive design. She developed a culturally appropriate user interface for Aboriginal communities, ensuring usability and respect for cultural norms (ACS Code of Professional Conduct Case Studies, 2014). The **2017 Equifax Data Breach** involved hackers exploiting an unpatched Apache Struts vulnerability, allowing them to access the personal data of over 148 million individuals over a 78-day period. The breach highlighted serious failures in vulnerability management and communication transparency (Kabanov & Madnick, 2020).

2.2 Ethical Dilemma

Each case reveals significant ethical dilemmas. The Colonial Pipeline incident underscores the conflict between public safety and negotiating with criminals. The ethical debate around whether paying ransoms legitimizes cybercrime or serves as a necessary act to restore critical services (Beerman et al., 2021; Hall, 2021). Ilnaz and Peter's experiences question whether organizational policies truly accommodate cultural and religious diversity. Anna's case emphasizes the ethical obligation to ensure digital systems are inclusive and equitable, particularly for marginalized communities. The Equifax breach presents a dilemma involving corporate responsibility to protect sensitive user data and act transparently when breaches occur (Kabanov & Madnick, 2020).

2.3 ICT Involvement

ICT professionals played a vital role in preventing and addressing these dilemmas. In the Colonial Pipeline case, cybersecurity teams, alongside federal agencies like the FBI, worked to secure networks and trace the attackers (Beerman et al., 2021). For Ilnaz and Peter, ICT project managers, HR professionals, and facility teams needed to collaborate on inclusive workplace policies. Anna's project illustrates the contributions of UI/UX designers and developers in creating culturally aware interfaces. Following the Equifax breach, IT teams implemented automated SSL certificate renewals and improved vulnerability scanning to ensure more comprehensive coverage. Compliance officers revised internal protocols, reflecting a broader institutional response to regain public trust (Kabanov & Madnick, 2020).

2.4 Application of the ACS Code of Ethics

Multiple ACS ethical principles were violated in the aforementioned cases. Under **Public Interest (1.2.1),** Colonial Pipeline and Equifax failed to safeguard critical infrastructure and consumer data (Beerman et al., 2021; Hall, 2021; Kabanov & Madnick, 2020). For **Quality of Life (1.2.2),** the lack of cultural sensitivity in Ilnaz's and Peter's cases disrupted personal dignity and workplace harmony (ACS Code of Professional Conduct Case Studies, 2014). **Honesty (1.2.3)** was compromised in both the Colonial Pipeline and Equifax breaches due to delayed public disclosures and poor communication. In terms of **Competence (1.2.4),** Colonial's reliance on an inactive VPN and Equifax's poor vulnerability scanning revealed major technical deficiencies. **Professionalism (1.2.6)** was absent in the lack of foresight, inclusivity, and transparency shown by the organizations involved. These breaches reinforce the importance of embedding ethics into all levels of ICT planning and execution. These mistakes are important lessons to avoid similar issues in future.

2.5 Adopting and maintaining Equity and Accessibility

Promoting inclusive systems is both a social and professional responsibility. Ilnaz's experience (Case 25) reveals the need for respectful workplace arrangements. Peter's challenge (Case 31) shows how overlooking small accommodations can impact well-being. Anna's solution (Case 32) is a model of best practice, demonstrating how technology can serve underrepresented communities. Katherina's volunteer work (Case 28) in providing ICT access to individuals with disabilities exemplifies how digital tools can uplift lives. In Case 24, inconsistent disability data compromised policy-making, illustrating the importance of accurate, inclusive data design. These cases affirm the need to prioritize accessibility and fairness from the start of any ICT initiative (ACS Code of Professional Conduct Case Studies, 2014).

2.6 Conclusion

This research review and ethical analysis underscore the growing importance of intelligent, ethical, and inclusive approaches to cybersecurity. Hybrid machine learning models show great promise in improving malware detection by blending the strengths of supervised and unsupervised learning. Flask emerges as a lightweight yet effective backend framework for integrating AI into web-based systems. However, implementation success hinges on careful dataset preparation, algorithm selection, and attention to performance trade-offs.

At the same time, the ethical review reinforces that technology development must be aligned with values of respect, equity, and public responsibility. From the Colonial Pipeline and Equifax breaches to workplace inclusion scenarios, these case studies offer essential lessons in ICT professionalism. The ACS Code of Ethics provides a critical framework for navigating these issues, ensuring systems are not only technically robust but also socially responsible. Future phases of this project will expand on the development of hybrid models, specific training algorithms, and establish a fully developed prototype system capable of addressing real-world cybersecurity challenges.

3.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. Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
4. Chowdhury, D., Stevens, L., & Grant, P. (2023). Cyber-physical systems anomaly detection using machine learning. *IEEE Transactions on Cybersecurity, 38*(4), 523–541.
5. Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
6. Cybersecurity Ventures. (2023). *2023 official cybercrime report*. https://cybersecurityventures.com
7. Gandhi, V., Kumar, S., & Kumar, S. (2023). Detection and classification of malware using machine learning techniques. *Journal of Information Security and Applications, 71*, 103223. https://doi.org/10.1016/j.jisa.2023.103223
8. 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.
9. 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
10. Kothari, C. R. (2004). *Research methodology: Methods and techniques* (2nd ed.). New Age International.
11. Lee, C., Wang, H., & Kim, S. (2023). Anomaly-based malware detection using autoencoders and decision trees. In *Proceedings of the International Conference on Cyber Threats* (Vol. 28, No. 1, pp. 215–232).
12. Mitchell, S., Brown, L., & Carter, P. (2023). Evaluating MERN stack for AI integration. *Web Systems and Security Journal, 17*(4), 89–106.
13. Nataraj, L., Yegneswaran, V., & Porras, P. (2023). Dynamic pattern recognition using signature analysis. *Journal of Computer Security Research, 31*(1), 45–62.
14. Roberts, T., Lee, J., & Adams, R. (2023). Efficiency of Flask in AI model deployment. *International Journal of Web Applications, 34*(2), 101–119.
15. Sharma, P., Kaur, J., & Singh, H. (2023). Malware detection using behaviour analysis. *Journal of Cybersecurity Techniques, 29*(3), 181–197.
16. Australian Computer Society. (2014). *ACS code of professional conduct: Case studies* (Version 2.1). https://www.acs.org.au