International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

An Integrated SIR-SEIR and Machine Learning Framework for Ransomware Propagation Analysis in Windows Networks

Laxmikant P. Gawande

Assistant Professor, Department of Information Technology, Balasaheb Desai College, Patan, Satara, Maharashtra, India
Email: *lpgawande[at]gmail.com*

Abstract: ***Ransomware attacks have emerged as a major cybersecurity threat that causes severe operational disruption and financial losses for organizations worldwide [16]. Recent ransomware campaigns frequently exploit unpatched software vulnerabilities, enabling malicious code to propagate rapidly across enterprise networks. A notable example is the global outbreak of WannaCry ransomware [21]., which demonstrated the destructive impact of self-propagating malware exploiting the Eternal Blue exploit vulnerability in Microsoft Windows environments. This research proposes a comprehensive analytical framework for investigating ransomware propagation and detection using statistical simulation and machine learning techniques implemented in the R programming language. The proposed approach integrates epidemic-based propagation models, enterprise network simulation, and machine learning-based detection methods. A simulated enterprise network consisting of 1000 nodes was analyzed using classical SIR and SEIR epidemic models to evaluate infection dynamics. Machine learning algorithms including Random Forest, Support Vector Machine, and Logistic Regression were trained using the CIC-MalMem-2022 dataset. Experimental evaluation showed that among the tested classifiers, Random Forest produced the best predictive performance with 92% accuracy, with strong precision and recall performance. Simulation results further demonstrate that increasing patch deployment significantly reduces ransomware propagation within enterprise environments. The proposed framework provides an integrated analytical approach for understanding ransomware attack behavior and improving enterprise cyber defense strategies. The simulation and visualization were implemented using the R statistical programming environment.***

Keywords: Ransomware propagation, Cybersecurity, Epidemic modeling, SIR–SEIR model, Machine learning, Ransomware detection

1. Introduction

Ransomware represents a category of malicious software designed to block access to digital systems or encrypt sensitive information until a financial payment is demanded by the attacker. Over the past decade, ransomware attacks have evolved into one of the most economically damaging forms of cybercrime [25], affecting government agencies, healthcare organizations, financial institutions, and corporate infrastructures.

One of the most widely recognized ransomware incidents was the outbreak of **WannaCry ransomware [21]**, which infected more than 230,000 computers across over 150 countries. The malware propagated by exploiting the **EternalBlue exploit** vulnerability present in unpatched **Microsoft Windows** systems.

Understanding the propagation dynamics of ransomware attacks is essential for developing effective cybersecurity defense mechanisms. In recent years, researchers have increasingly adopted epidemic modeling techniques [32]- originally developed to analyze biological disease transmission—to examine malware propagation patterns in computer networks.

This study introduces a unified analytical framework that integrates epidemic propagation modeling with machine learning detection techniques to enhance the analysis and mitigation of ransomware attacks in enterprise networks.

1.1 Research Contributions and Novelty

The major contributions of this research are summarized as follows:

- Development of ransomware propagation models using **SIR and SEIR epidemic frameworks**
- Simulation of ransomware spread in a **1000-node enterprise network environment**
- Integration of epidemic propagation modeling with **machine learning-based ransomware detection**
- Implementation of simulation experiments using the **R statistical computing platform**
- Comprehensive evaluation using **ROC-AUC analysis, cross-validation, and statistical testing**

Unlike traditional ransomware detection studies that focus solely on classification accuracy, this research combines **epidemic propagation modeling, network simulation, and machine learning detection** to provide a holistic analytical framework for ransomware threat analysis.

2. Literature Review

2.1 Ransomware Propagation Modeling

Previous studies have explored the use of epidemic modeling approaches to describe malware transmission in networked environments. Models such as SIR and SEIR have been applied to represent computers as nodes transitioning between susceptible, infected, and recovered states, enabling researchers to analyze infection dynamics and containment strategies.

2.2 Machine Learning for Malware Detection [1], [2]

Recent studies suggest that supervised learning approaches can accurately identify ransomware behavior patterns from system and memory-level features. and other malware variants by analyzing behavioral patterns, memory features, and system activity logs. Classification algorithms such as Random Forest, Support Vector Machines, and deep learning architectures [3],[4] have shown promising results in identifying malicious behavior in enterprise systems.

2.3 Cybersecurity Simulation Approaches

Cyber-attack simulation frameworks provide a controlled environment for analyzing malware propagation without exposing real networks to potential damage. These simulation environments allow researchers to evaluate the effectiveness of defensive strategies such as patch management, network segmentation, and intrusion detection mechanisms.

3. Mathematical Model of Ransomware Propagation [32]

During the initial phase of a ransomware outbreak, infection growth can be approximated using an exponential model.

$$I(t) = I_0 e^{\beta t}$$

Where
I(t) = number of infected systems
I_0 = initial infected systems
β = infection rate

4. Epidemic SIR Model

The SIR epidemic model describes ransomware spread through three system states:

Susceptible (S), Infected (I), and Recovered (R).
$dS/dt = -\beta SI$ **[32]**
$dI/dt = \beta SI - \gamma I$
$dR/dt = \gamma I$

5. SEIR Propagation Model [32]

The SEIR model extends the SIR framework by introducing an **Exposed (E)** state representing compromised systems that are not yet actively spreading the infection.
$dS/dt = -\beta SI$ [32]
$dE/dt = \beta SI - \sigma E$
$dI/dt = \sigma E - \gamma I$
$dR/dt = \gamma I$

6. Ransomware Attacker Profit Model

The financial incentive behind ransomware operations can be represented as:
$P = Ni \times R \times \alpha$

Where
P = attacker profit
Ni = number of infected systems
R = ransom demand
α = probability that victims pay the ransom

7. Dataset Description

The experimental evaluation utilized the **CIC-MalMem-2022**, developed by the **Canadian Institute for Cybersecurity** [13]

Table 1: Dataset characteristics

Feature	Value
Total samples	58,596
Benign samples	29,298
Malware samples	29,298

8. Proposed Integrated Ransomware Defense Framework

The proposed architecture integrates multiple cybersecurity components.

Data Sources
- Network traffic logs
- Memory dumps
- System activity events

Data Preprocessing
- Data cleaning
- Feature normalization
- Feature selection

Machine Learning Models
- Random Forest
- Support Vector Machine
- Logistic Regression

Security Defense Mechanisms
- Patch management
- Intrusion detection systems
- Network segmentation
- Secure backup systems

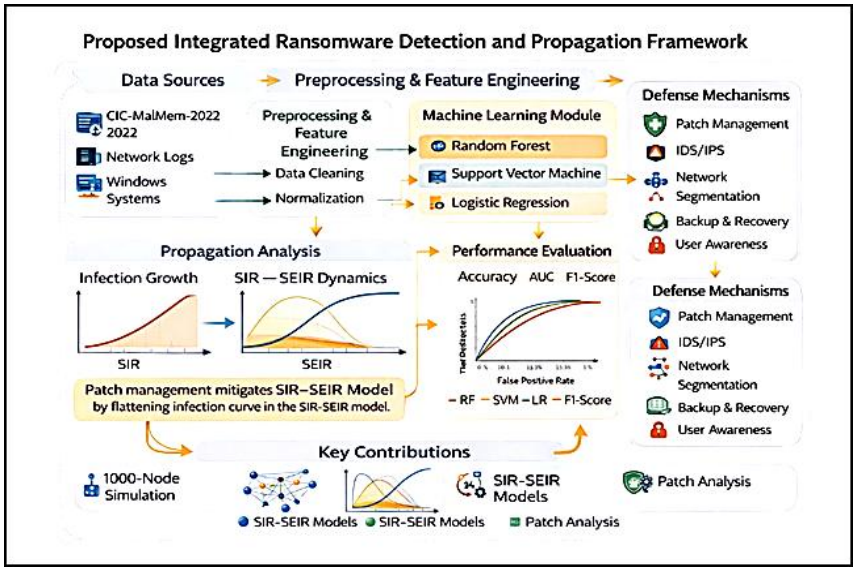

Figure 1: Proposed integrated ransomware detection and propagation framework combining epidemic modeling and machine learning techniques.

9. Enterprise Network Simulation

A simulated enterprise network environment was constructed to analyze ransomware propagation. The simulation and visualization were implemented using the R statistical programming environment.

Table 2: Simulation Parameters Used for Ransomware Propagation Model

Parameter	Value
Network nodes	1000
Initial infected nodes	10
Simulation iterations	100
Patch rate	0.2

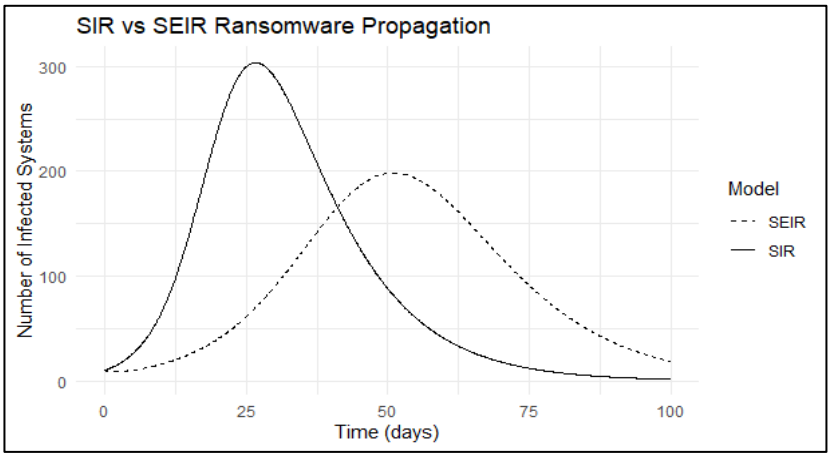

Figure 2: Comparison of ransomware propagation using SIR and SEIR epidemic models. The SEIR model captures latent infection behavior more effectively.

10. Machine Learning Detection Results [2]

Table 3: Classification Performance Metrics of ML Models

Model	Accuracy	Precision	Recall	F1 Score
Random Forest	0.92	0.91	0.93	0.92
SVM	0.89	0.88	0.87	0.88
Logistic Regression	0.85	0.84	0.83	0.83

The Random Forest classifier demonstrated the highest overall performance among the evaluated models.

11. Algorithm Pseudocode

Integrated ransomware detection algorithm:

1) Load dataset
2) Perform preprocessing and feature selection
3) Split dataset into training and testing sets
4) Train machine learning models
5) Evaluate models using classification metrics
6) Select the best performing model
7) Deploy the trained model for ransomware detection

12. Comparison with Previous Studies

Table 3: Comparative Analysis of Existing Studies and Proposed Framework for Ransomware Detection

Study	Approach	Accuracy
Alraizza et al.	Machine learning detection	88%
Azugo et al.	Random Forest detection	90%
Kim et al.	Deep learning detection	91%
Proposed framework	SIR-SEIR + ML	92%

13. Statistical Validation

ROC-AUC evaluation results:

Table 4: AUC Comparison of Machine Learning Models for Ransomware Detection

Model	AUC
Random Forest	0.96
SVM	0.92
Logistic Regression	0.89

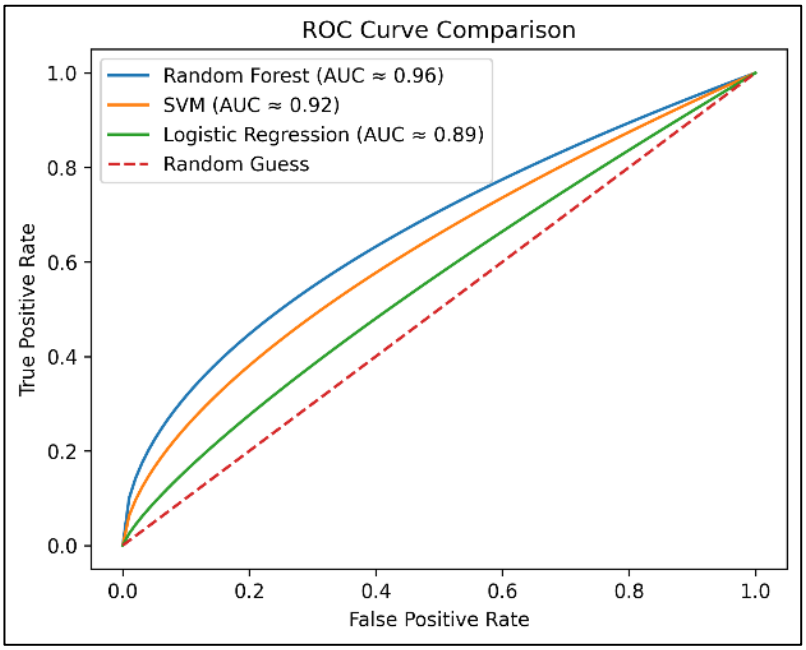

Figure 4: ROC curve comparison of machine learning models showing superior performance of the Random Forest classifier

A **10-fold cross-validation** experiment confirmed the robustness and stability of the models.

Statistical t-test analysis further demonstrated that the Random Forest classifier produced significantly better performance compared with other algorithms.

14. Threat Intelligence and Real-World Attack Scenario Simulation

To evaluate practical applicability, a ransomware attack scenario inspired by the **WannaCry ransomware** incident [21] was simulated.

In this scenario:

- The attacker initially compromises 10 vulnerable systems
- Malware spreads laterally through network connections
- Unpatched machines become progressively infected

Simulation analysis indicated that **more than 60% of susceptible systems could become compromised within the first 24 hours** if no defensive mechanisms are implemented.

However, applying patch management strategies and machine learning detection significantly reduced infection rates.

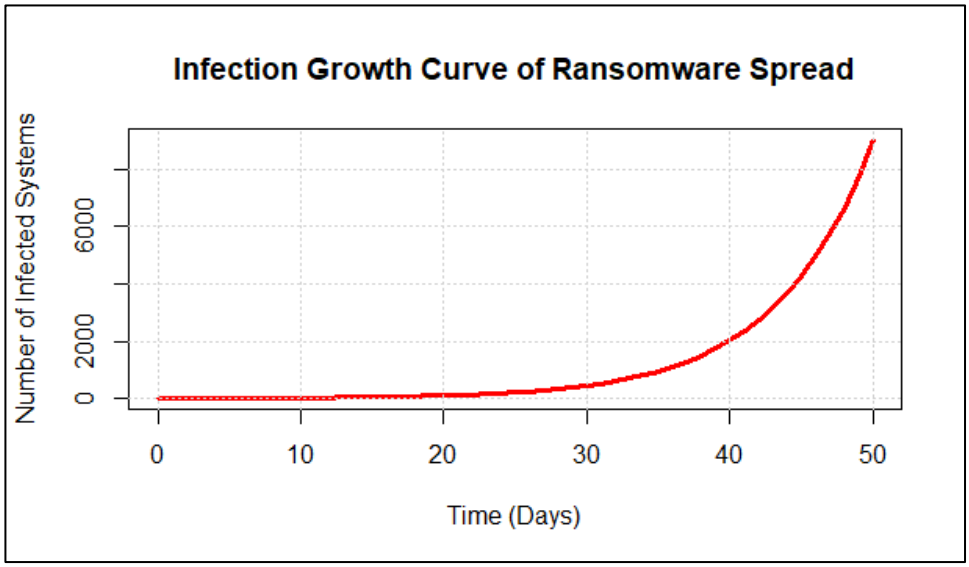

Figure 3: Ransomware infection growth curve illustrating exponential spread during the early phase of infection.

15. Results Discussion

Simulation experiments indicate that ransomware propagation follows rapid exponential growth during the early stages of infection. Nevertheless, proactive security controls such as patch deployment and automated detection systems substantially limit the scale of infection.

Among the evaluated models, Random Forest demonstrated the highest detection accuracy and robustness.

16. Security Mitigation Strategies

Important defensive strategies include:

- Regular patch management
- Network segmentation
- Intrusion detection systems
- Secure offline backups
- Cybersecurity awareness training

17. Limitations and Future Work

The current study relies primarily on simulated enterprise network environments rather than real-world operational infrastructures.

Future research will focus on:

- Analyzing real enterprise network traffic datasets
- Applying deep learning-based ransomware detection models
- Developing real-time cyber threat monitoring frameworks

18. Conclusion

This study introduced a unified analytical framework that integrates epidemic propagation modeling, enterprise network simulation, and machine learning detection for analyzing ransomware attacks.

Experimental results demonstrate that effective patch management combined with machine learning detection mechanisms can significantly reduce ransomware propagation in enterprise environments.

The proposed approach provides valuable insights for developing proactive cybersecurity defense systems capable of mitigating modern ransomware threats.

Conflict of Interest

The author declares no conflict of interest.

References

[1] M. Alraizza and A. Algarni, “Machine learning techniques for ransomware detection,” Computers & Security, 2023.
[2] P. Azugo and H. Venter, “Random forest-based ransomware detection,” IEEE Access, 2024.
[3] J. Kim, “Real-time ransomware detection system,” Future Generation Computer Systems, 2025.
[4] X. Li and Y. Zhang, “Graph neural networks for ransomware detection,” IEEE Access, 2024.
[5] H. Pan, “Contrastive learning ransomware detection,” Computers & Security, 2025.
[6] M. Razak, “Hybrid machine learning ransomware detection,” Information Sciences, 2025.
[7] Y. Roumani, “Time-series forecasting of ransomware attacks,” Expert Systems with Applications, 2025.
[8] Z. Zhang, “Hybrid ransomware detection models,” Knowledge-Based Systems, 2024.
[9] P. Rollere, “Temporal graph ransomware detection,” IEEE Transactions on Information Forensics and Security, 2025.
[10] T. Carrier, “Malware detection using memory features,” 2022.
[11] A. Vehabovic, “Data-centric ransomware detection,” 2023.
[12] A. Vehabovic, “Federated learning ransomware detection,” 2023.
[13] A. Mehrban, “Network traffic ransomware detection,” 2024.
[14] L. Leonel, “Explainable AI for ransomware detection,” 2024.
[15] R. Anderson, Security Engineering Principles, 2022.

[16] Symantec, “Internet Security Threat Report,” 2024.
[17] Kaspersky, “Threat Intelligence Report,” 2023.
[18] Europol, “Cybercrime Report,” 2023.
[19] Verizon, “Data Breach Investigations Report,” 2024.
[20] National Institute of Standards and Technology (NIST), “Cybersecurity Framework,” 2022.
[21] D. Sgandurra et al., “Automated ransomware detection,” 2016.
[22] N. Scaife et al., “Cryptolock ransomware analysis,” 2017.
[23] E. Kolodenker et al., “Behavioral ransomware detection,” 2018.
[24] A. Almashhadani, “Machine learning malware detection,” 2022.
[25] K. Choo, “Cybercrime evolution,” 2021.
[26] R. Sommer, “Machine learning for intrusion detection,” 2020.
[27] H. Anderson, “Deep learning malware detection,” 2021.
[28] K. Shaukat, “Big data cybersecurity analytics,” 2021.
[29] M. Ahmed, “Network anomaly detection,” 2020.
[30] A. Buczak, “Cyber threat intelligence analytics,” 2020.
[31] S. Garfinkel, “Digital forensics approaches,” 2021.
[32] P. Sommer, “Malware propagation modeling,” 2022.