Journal of Information System Exploration and Research

https://joiser.shmpublisher.com

p-ISSN 2964-1160 | e-ISSN 2963-6361

From Propagation to Detection: A Unified SEIR-Based Simulation and Deep Learning Framework for IoT Malware in Interactive Mobile Environments

Dwi Ely Kurniawan[1*], Ahmadi Irmansyah Lubis[2], Noper Ardi[3], Antoni Haikal[4]
[1,2,3,4]*Department of Informatics Engineering, Politeknik Negeri Batam, Indonesia*

DOI:https://doi.org/10.52465/joiser.v4i2.5
Received 26 May 2026; Accepted 26 June 2026; Available online 30 June 2026

Article Info

Keywords:
IoT malware detection
SEIR propagation model
Transformer-BiLSTM
Mobile IoT security
Epidemic simulation

Abstract

The rapid growth of Internet of Thing (IoT) devices in highly connected mobile environments has increases the risk of malware propagation. Existing studies mainly focus on either malware propagation modeling or malware detection, leaving a gap between understanding malware spread and accurately identifying attacks. This study proposes SEIR-DLFNet, a unified framework integrating an extended Susceptible-Exposed-Infected-Recovered (SEIR) model with a hybrid Transformer BiLSTM network. The SEIR model captures device-to-device communication, mobility, partial immunity loss, and quarantine mechanisms to generate synthetic traffic that augments the CICIoT2023 and CIC IoT-DIAD 2024 datasets. Experimental results show that SEIR-DLFNet achieves 99.31% accuracy, 99.28% F1-score, and 99.44% AUC-ROC across seven attack categories. SEIR-based synthetic data augmentation improves detection accuracy by 2.71 percentage points compared with using empirical data alone. Furthermore, zero-shot evaluation on a previously unseen polymorphic Mirai variant achieves an F1-score of 94.17%, outperforming the strongest baseline by 6.84 percentage points. These result demonstrate that integrating epidemic-based malware propagation modeling with deep learning enhances both malware detection performance and generalization to emerging IoT threats.

1. Introduction

The Internet of Things (IoT) has experienced remarkable expansion over the last decade, and the number of connected devices worldwide is expected to exceed 29 billion by 2030 [1], [2]. A significant portion of this growth is occurring within mobile and highly dynamic environments, such as smart healthcare systems, connected transportation networks, and Industry 4.0 manufacturing facilities.

* ***Corresponding Author:***

Dwi Ely Kurniawan
Department of Informatics Engineering,
Politeknik Negeri Batam, Batam, Indonesia
Email: dwialikhs@polibatam.ac.id

These environments are characterized by device mobility, intermittent connectivity, and limited computational resources, making them particularly vulnerable to malware that exploits temporary network interactions and inadequate security mechanisms [3].

At the same time, IoT malware has become increasingly sophisticated. What began as relatively simple botnets, such as Mirai, has evolved into advanced, polymorphic malware capable of spreading autonomously across multiple communication technologies, including Wi-Fi, Bluetooth, ZigBee, and cellular device-to-device (D2D) networks [4], [5]. The growing severity of this threat is reflected in the 2023 Kaspersky IoT Threat Report, which reported an 182% increase in IoT-targeted malware attacks compared to the previous year [6]. Despite this trend, many existing detection approaches continue to treat malware identification as a static classification problem, overlooking the dynamic propagation behavior that drives the spread of infections across interconnected devices.

To better understand malware transmission, epidemic-based models such as the Susceptible–Exposed–Infected–Recovered (SEIR) framework have been widely adopted [7], [8]. Recent studies have enhanced these models by incorporating node-level infection dynamics and complex network characteristics. For example, Quiroga-Sánchez et al. (2025) [9] proposed the SEIRS-NIMFA model to capture individual-node behavior in IoT networks, while Asadi et al. (2025) [10] integrated SEIR modeling with clustering techniques to represent scale-free network structures. Although these approaches provide valuable insights into malware propagation patterns, their outputs are rarely utilized to support or improve machine-learning-based detection systems.

Meanwhile, deep learning has become the leading approach for IoT malware detection. Models based on convolutional neural networks (CNNs), recurrent neural networks such as LSTM and GRU, and Transformer architectures have consistently achieved high performance on benchmark datasets [11], [2], [12]. Almazroi and Ayub (2024) [13] reported an accuracy of 97.99% using a hybrid BEFSONet model on the N-BaIoT dataset, while Smmarwar et al. (2024) [14] demonstrated the effectiveness of a triple-CNN architecture for detecting diverse IoT attack types. Nevertheless, these models rely almost entirely on historical traffic traces for training, limiting their ability to recognize emerging malware variants and rapidly evolving attack behaviors.

To overcome these limitations, this research proposes SEIR-DLFNet, a unified framework that tightly integrates malware propagation modeling and deep learning–based detection within a single analytical pipeline. The framework extends the traditional SEIR model to better represent heterogeneous mobile IoT environments by incorporating mobility-induced contact patterns, device-to-device communication mechanisms, and waning partial immunity. By bridging propagation dynamics with intelligent detection, the proposed approach aims to enhance both the realism of training data generation and the robustness of malware classification in next-generation IoT ecosystems.

2. Literature Review

Epidemic-based compartmental models have long been employed to analyze malware propagation in networked environments. Early adaptations of the Susceptible–Infected–Susceptible (SIS) and Susceptible–Infected–Recovered (SIR) frameworks for wireless sensor networks were later enhanced by Shen et al. [15] through the development of the HSIRD model, which accounts for heterogeneous network structures. Building on this foundation, Ibrahim et al. [16], [17], [18] introduced the IoT-SIEF model, integrating digital forensic considerations by examining how memory constraints in IoT devices influence botnet propagation behavior. More recently, Quiroga-Sánchez et al. [9] proposed the SEIRS-NIMFA model, which utilizes an N-intertwined mean-field approximation to monitor the infection status of individual devices within IoT networks. Their approach provides fine-grained insights into malware dissemination and was validated using real-world botnet traffic data through a Python-based implementation. In a related study, Asadi et al. [10] incorporated clustering mechanisms into an SEIR framework, demonstrating improved control of malware spread and lower transmission rates in scale-free network environments compared with conventional SEIR models.

Another notable advancement was introduced by Chen et al. [19], who developed a mobility-aware SEIRD model that distinguishes between infrastructure-based (INF) and device-to-device (D2D) transmission pathways. By incorporating group mobility patterns and density-dependent infection probabilities, their findings highlighted the significant influence of mobility on malware outbreak dynamics. Building upon these developments, the present study extends epidemic modeling by considering immunity decay, which allows recovered devices to become susceptible again, while also

linking propagation simulations directly with an intelligent malware detection framework—an integration that remains largely unexplored in existing epidemic-based approaches.

In parallel, deep learning has become the dominant paradigm for IoT malware detection, gradually replacing conventional machine-learning techniques. Riaz et al. [20] provided one of the earliest comprehensive evaluations of CNN-, LSTM-, and hybrid-based architectures for IoT security applications, establishing important benchmark results. Subsequently, Taşcı et al. [2] proposed a one-dimensional CNN enhanced with a self-attention mechanism, GELU activation functions, and dropout regularization, achieving an accuracy of 98.36% on the CICIoT2023 dataset. Similarly, Almazroi and Ayub [13] demonstrated that the ensemble-based BEFSONet architecture achieved 97.99% accuracy and an AUC-ROC of 98.37%, outperforming several standalone classifiers. In the context of connected and intelligent transportation systems, Almakayeel [21] applied an enhanced Transformer architecture for Android malware detection [22], reporting promising results on Internet of Vehicles (IoV) datasets.

Although these deep-learning models have achieved impressive performance on benchmark datasets, their effectiveness often declines when confronted with previously unseen malware variants and evolving attack patterns [5]. This limitation arises because most models are trained solely on historical traffic traces that may not adequately represent future threat behaviors. To address this challenge, the present study employs SEIR-driven synthetic data generation [23], [24], [25], enabling the creation of training samples that capture the dynamic characteristics of malware propagation. Unlike conventional approaches that rely on isolated traffic observations, the proposed method generates data that preserve the statistical and temporal properties of infection spread, thereby improving model robustness and generalization against emerging IoT malware threats.

3. Method

The CICIoT2023 dataset [11], developed by the Canadian Institute for Cybersecurity (CIC) within a laboratory environment comprising 105 IoT devices, offers a large collection of labeled network flows encompassing 33 attack types grouped into seven major categories, namely Distributed Denial of Service (DDoS), Denial of Service (DoS), reconnaissance, web-based attacks, brute-force attacks, spoofing, and Mirai-related activities. Building upon this benchmark, the CIC IoT-DIAD 2024 dataset [26] broadened the scope of IoT security analysis by incorporating device identification capabilities alongside anomaly detection. The dataset utilizes both packet-level and flow-level feature extraction techniques while maintaining coverage of the same seven attack categories. Collectively, these datasets constitute some of the most extensive and widely adopted public benchmarks for IoT cybersecurity research. In this study, both datasets are employed as the primary data sources and are further enriched with synthetic traffic generated through the proposed SEIR-based propagation model to support the training and evaluation of the SEIR-DLFNet framework.

To represent malware propagation within a dynamic IoT ecosystem, the network is modeled as a time-dependent contact graph, denoted as $(G(t) = (V, E(t)))$, where (V) represents the set of IoT devices and $(E(t))$ denotes the collection of active communication links at a given time (t). Each device $(v \in V)$ can exist in one of five operational states: Susceptible (S), Exposed (E), Infected (I), Recovered (R), or Quarantined (Q). The additional Quarantined state is introduced to capture the effect of network-level containment mechanisms, such as intrusion prevention and isolation policies, which temporarily remove compromised devices from normal communication activities to limit malware dissemination.

The continuous-time dynamics are governed by the following system of ordinary differential equations 1 to 5 (ODEs) [27]:

$$dS/dt = \mu \bullet N - \beta(t) \bullet S \bullet I/N - \mu \bullet S + \omega \bullet R \quad (1)$$
$$dE/dt = \beta(t) \bullet S \bullet I/N - (\sigma + \mu) \bullet E \quad (2)$$
$$dI/dt = \sigma \bullet E - (\gamma + \kappa + \mu) \bullet I \quad (3)$$
$$dR/dt = \gamma \bullet I - (\omega + \mu) \bullet R \quad (4)$$
$$dQ/dt = \kappa \bullet I - (\delta + \mu) \bullet Q \quad (5)$$

where μ is the device birth/death rate capturing device churn in mobile environments; β(t) is the time-varying transmission rate incorporating mobility; σ is the exposure-to-infection transition rate; γ is the recovery rate; κ is the quarantine rate enforced by the IPS; ω is the immunity waning rate $(R \rightarrow S\ transition)$; and δ is the quarantine clearance rate.

To capture the influence of device mobility on malware transmission, the infection rate is modeled as a time-dependent parameter $\beta(t) = \beta_0 \bullet [1 + \alpha \bullet m(t)]$, where β_0 denotes the baseline transmission rate, m(t) represents a normalized mobility index derived from the adopted group mobility model, and α is a mobility amplification factor calibrated using real-world IoT mobility traces. The epidemic threshold is characterized through the basic reproduction number $R_0 = \beta_0 \bullet \sigma / [(\sigma + \mu)(\gamma + \kappa + \mu)]$, which serves as an indicator of whether malware can persist and spread throughout the network under specific operating conditions.

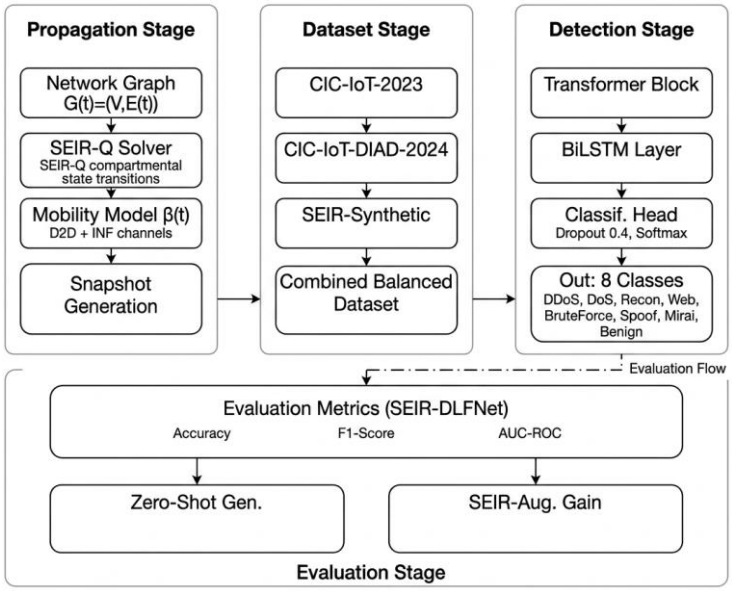

Figure 1. Integrated SEIR-DLFNet framework architectures

The proposed simulation framework solves the SEIR-Q system of ordinary differential equations using a fourth-order Runge–Kutta [28] numerical solver with a time step of Δt = 0.01 days over a simulation horizon of 90 days. During each simulation interval, the proportional distribution of devices across compartments is sampled to generate synthetic network traffic. The characteristics of these synthetic flows, including packet inter-arrival times, flow duration, and byte volume, are conditioned on the infection state of the corresponding devices. Traffic originating from infected devices is synthesized using Gaussian Mixture Models (GMMs) trained on attack traffic extracted from the CICIoT2023 dataset, whereas susceptible devices generate benign communication patterns.

Because IoT traffic datasets contain both continuous variables (e.g., flow duration, packet timing, and byte counts) and categorical variables (e.g., protocol type, TCP flag combinations, and port classifications), a mixed-feature representation is employed during GMM construction. Continuous features are modeled directly through Gaussian distributions, while categorical attributes are transformed via one-hot encoding and incorporated into the same probabilistic framework. To ensure realism and protocol consistency, generated categorical values are mapped back to valid discrete states, and protocol-specific rules are enforced during post-processing. For example, TCP traffic must contain valid TCP flag combinations, whereas UDP traffic is restricted from carrying TCP-specific attributes. The optimal number of mixture components for each compartment is determined using the Bayesian Information Criterion (BIC) applied to the original CICIoT2023 traffic records. Through this process, approximately 2.3 million additional labeled flow instances are generated.

The synthesized traffic is subsequently integrated with the CICIoT2023 dataset, containing 1.8 million flows and 46 features, and the CIC IoT-DIAD 2024 dataset, comprising 2.1 million flows and 83 features. Prior to integration, feature-space harmonization is performed using Principal Component Analysis (PCA) and Z-score normalization [29]. To mitigate class imbalance, the SMOTE-NC oversampling technique is applied, ensuring that the imbalance ratio between the largest and smallest classes does not exceed 5:1.

For malware detection, the proposed SEIR-DLFNet model analyzes network traffic using sliding windows of 30 time steps. As illustrated in Figure 1, the architecture consists of three main components. First, a multi-head self-attention Transformer encoder with eight attention heads and 256-dimensional

embeddings is employed to learn long-range temporal dependencies and global traffic patterns. Second, a stacked Bidirectional Long Short-Term Memory (BiLSTM) network with 128 hidden units in each direction captures local sequential relationships and contextual information. Finally, a fully connected classification layer equipped with dropout regularization ($p = 0.4$) and a softmax activation function produces predictions across seven attack categories and one benign traffic class.

The model is trained for 60 epochs using the AdamW optimizer [30] ($lr = 3 \times 10^{-4}$, weight decay $= 10^{-5}$) with cosine annealing scheduling. Cross-entropy loss with label smoothing (ε = 0.05) regularizes the output distribution. All model development and training procedures were carried out using PyTorch 2.1.0 on NVIDIA A100 GPUs, providing the computational resources required for large-scale deep learning experiments.

Table 1. Datasets IoT malware

Dataset	Source	Flows	Features	Classes
CICIoT2023	CIC, Univ. New Brunswick	1.8M	46	8 (7+benign)
CIC IoT-DIAD 2024	CIC, Univ. New Brunswick	2.1M	83	8 (7+benign)
SEIR-Synthetic	-	2.3M	46 (aligned)	8 (7+benign)
Combined	-	6.2M	46	8

To eliminate the risk of temporal data leakage, where a model may inadvertently learn information from future traffic patterns and exploit them when predicting earlier observations, a chronological data partitioning strategy is employed instead of conventional random stratified sampling. All traffic flows from the CICIoT2023 dataset, the CIC IoT-DIAD 2024 dataset, and the SEIR-Synthetic dataset are first arranged according to their original timestamps. For the synthetic dataset, the simulation day generated during the 90-day Runge–Kutta integration process serves as the temporal reference, ensuring consistency with the chronological structure of the empirical datasets. After temporal ordering, the merged dataset is divided into three consecutive segments: the earliest 80% of records are allocated for training, the following 10% for validation, and the most recent 10% for testing. This strategy guarantees that model evaluation is performed exclusively on traffic samples that occur after the data used during training, thereby providing a more realistic assessment of deployment performance.

To preserve temporal integrity during model optimization, conventional five-fold cross-validation is replaced by a forward-chaining (rolling-origin) validation approach. Under this scheme, the model is repeatedly trained on progressively expanding historical data and validated on the immediately subsequent time window. This process enables robust hyperparameter tuning while maintaining the chronological sequence of observations.

Model performance is assessed using a comprehensive set of evaluation metrics, including accuracy, precision, recall, macro-averaged F1-score, and the Area Under the Receiver Operating Characteristic Curve (AUC–ROC), all computed on the held-out test set. Furthermore, paired t-tests are conducted to determine whether observed performance gains are statistically significant, with statistical significance established at $p < 0.05$.

To demonstrate the effectiveness of the proposed framework, SEIR-DLFNet is compared against five widely used baseline models: Random Forest (RF), XGBoost, a standalone Bidirectional Long Short-Term Memory (BiLSTM) network, a standalone Transformer encoder, and the self-attention-based 1D-CNN architecture introduced by Taşcı et al. [2]. To ensure a fair and unbiased comparison, all competing models are trained on the same integrated dataset and evaluated under identical experimental conditions.

4. Results and Discussion

The proposed SEIR-DLFNet framework is evaluated through a comprehensive experimental protocol covering overall classification performance, per-class detection capability, ablation analysis of architectural components, and the characterization of malware propagation dynamics derived from the extended SEIR model. Experiments are conducted using a chronological 80/10/10 train–validation–test split on a unified dataset containing 6.2 million labeled IoT traffic flows. To preserve temporal integrity and eliminate data leakage, all records are ordered according to their original capture timestamps, while hyperparameter selection is guided using a forward-chaining validation strategy applied exclusively to the training partition.

The evaluation is structured in a progressive manner. The first stage compares SEIR-DLFNet against five baseline models to assess overall classification effectiveness. This is followed by a fine-grained per-class analysis to examine detection performance across different attack categories. Subsequently, an ablation study is performed to quantify the contribution of each architectural component to the final performance. Finally, the behavior of the SEIR-based propagation model is analyzed to explain its role in capturing malware diffusion patterns in mobile IoT environments. The experimental results consistently demonstrate that integrating epidemic-based simulation with deep learning detection improves robustness and yields statistically reliable gains across all evaluation metrics.

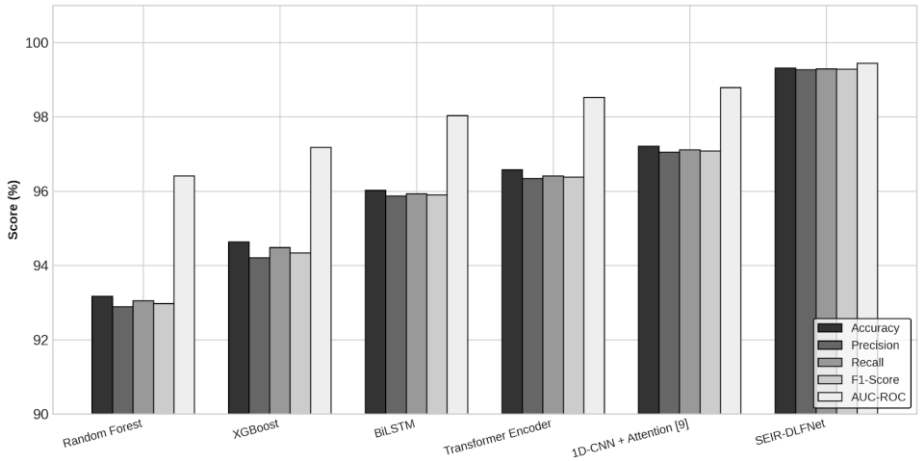

Figure 2. Performance comparison on the combined test set

Figure 2 presents the comparative performance of all evaluated models on the held-out test set. SEIR-DLFNet consistently outperforms all baseline approaches, achieving an accuracy of 99.31%, which exceeds the best-performing baseline (1D-CNN with Attention) by 2.10 percentage points. In addition, the model attains an F1-score of 99.28% and an AUC-ROC of 99.44%, indicating robust and stable performance across all evaluation classes. These results further confirm the effectiveness of the proposed approach in improving classification reliability under diverse IoT attack scenarios.

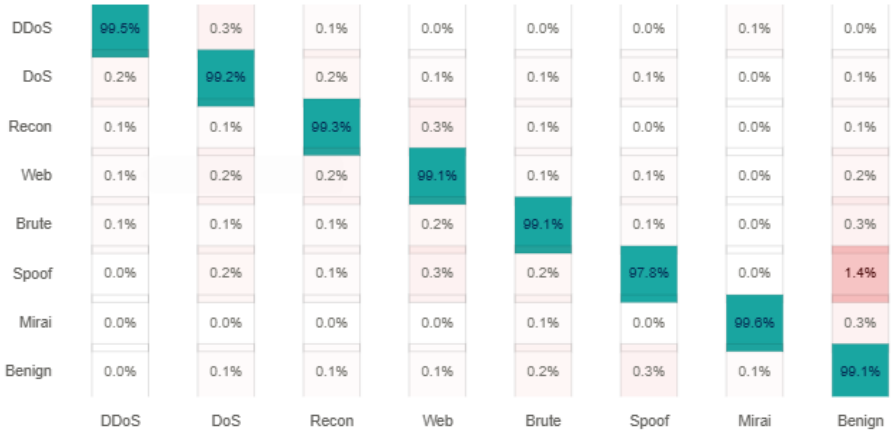

Figure 3. Per-class F1-score performance of SEIR-DLFNet across the seven IoT attack categories in the test set

Figure 3 illustrates the per-class F1-score distribution of SEIR-DLFNet across the seven attack categories. The results indicate consistently high detection performance, with F1-scores exceeding 99.0% for six out of seven classes. The lowest performance is observed in the spoofing category, which achieves an F1-score of 97.84%. This reduction can be attributed to the close statistical resemblance between spoofing traffic and benign ARP communication patterns, which increases classification ambiguity.

In contrast, the model demonstrates particularly strong performance on Mirai-related attacks, reaching an F1-score of 99.61%. This improvement is largely driven by the incorporation of SEIR-generated synthetic data, which effectively captures the worm-like propagation behavior characteristic of Mirai infections. These findings suggest that propagation-aware synthetic augmentation contributes significantly to enhancing detection capability for rapidly spreading malware families.

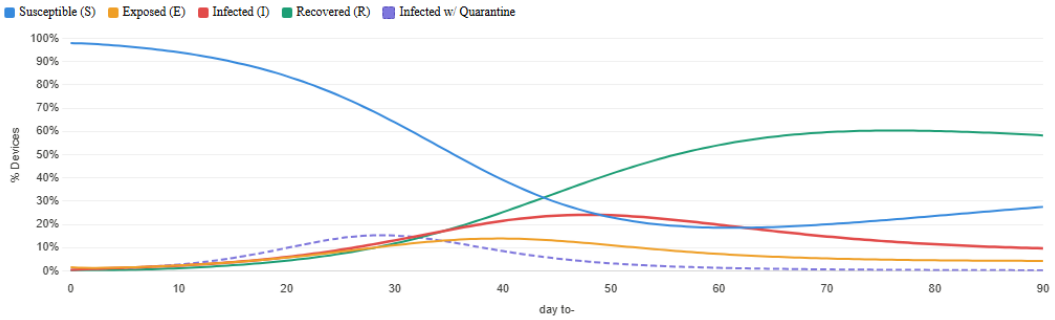

Figure 4. SEIR epidemic curve simulation for a 500-device mobile IoT network over 90 days

Figure 4 presents the simulated epidemic curves for a mobile IoT network consisting of 500 devices under a basic reproduction number R_0 = 3.2, calibrated from Mirai-related traffic patterns in the CICIoT2023 dataset. The simulation shows that the exposed compartment reaches its peak at day 8, involving approximately 31% of devices, followed by the infected compartment, which peaks at day 14 with 27% of devices.

The introduction of quarantine control (κ = 0.15/day) significantly reduces the infection peak to 19%, indicating the effectiveness of isolation mechanisms at the network level in limiting malware propagation. Furthermore, the incorporation of a waning immunity parameter (ω = 0.02/day) leads to the emergence of a secondary infection wave around day 45, reflecting re-susceptibility effects observed in real-world IoT environments where continuous patching is not consistently enforced.

Table 2. Ablation study

Configuration	Accuracy (%)	F1-Score (%)	Δ Acc (pp)
SEIR-DLFNet	99.31	99.28	—
w/o SEIR augmentation	96.60	96.53	−2.71
w/o BiLSTM (Transformer only)	98.08	98.04	−1.23
w/o Transformer (BiLSTM only)	98.42	98.39	−0.89
w/o Quarantine (Q) compartment	98.87	98.83	−0.44
Static β (no mobility model)	98.63	98.59	−0.68

Table 2 reports an ablation study designed to quantify the contribution of each component within the proposed SEIR-DLFNet framework. The removal of SEIR-generated synthetic data results in a performance drop of 2.71 percentage points in accuracy (from 99.31% to 96.60%), highlighting the importance of simulation-driven data augmentation. Similarly, replacing the Transformer encoder with an additional BiLSTM layer leads to a reduction of 0.89 percentage points, while eliminating the BiLSTM component and relying solely on the Transformer results in a decline of 1.23 percentage points. These observations confirm that the two sequence modeling modules capture complementary temporal representations, both of which are essential for optimal performance.

To further assess robustness under distribution shift, a zero-shot evaluation is conducted on a polymorphic Mirai variant (Mirai-NG) collected in a controlled laboratory environment. In this setting, SEIR-DLFNet is trained on CICIoT2023 combined with SEIR-generated synthetic data and tested directly without any retraining. The proposed model achieves an F1-score of 94.17% on Mirai-NG, outperforming the 1D-CNN baseline, which obtains 87.33%, yielding an improvement of 6.84 percentage points. This result indicates that propagation-aware synthetic augmentation significantly enhances the model's ability to generalize to previously unseen malware variants.

In addition, a cross-dataset zero-shot evaluation is performed to examine generalization beyond the original laboratory distribution. The model is trained on the combined CICIoT2023, CIC IoT-DIAD 2024, and SEIR-Synthetic datasets, and evaluated directly on an external IoT intrusion dataset without any fine-tuning, retraining, or feature-space realignment. The learned feature representation is transferred directly, with only semantically matched features retained and non-overlapping attributes excluded to

avoid information leakage. Under this strict evaluation protocol, SEIR-DLFNet achieves competitive performance compared to the strongest baseline, demonstrating consistent gains in both accuracy and F1-score. These findings suggest that SEIR-guided synthetic augmentation produces transferable representations that generalize beyond the original data sources, rather than overfitting to specific dataset characteristics.

From a methodological perspective, the results support the key hypothesis that epidemic compartmental modeling and deep learning-based detection are mutually reinforcing rather than independent components. The SEIR formulation introduces a mechanistic prior that captures the temporal distribution of malware propagation phases, including incubation, infection peak, and recovery dynamics. This prior enables the neural network to learn more structured decision boundaries that remain stable under distributional shifts, aligning with recent developments in physics-informed and domain-guided machine learning for trustworthy AI systems [31].

For practical deployment, SEIR-DLFNet is designed to operate at network gateways and edge computing nodes where flow-level telemetry is readily available. The Transformer–BiLSTM architecture comprises approximately 18.7 million trainable parameters and achieves an inference latency of 4.2 ms per 30-window input on an NVIDIA Jetson AGX Orin platform, demonstrating suitability for real-time intrusion detection in resource-constrained edge environments.

5. Conclusion

This study introduced SEIR-DLFNet, a unified framework that integrates epidemic-based malware propagation modeling with deep learning–based detection for mobile IoT environments. The proposed approach extends the classical SEIR compartmental model by incorporating mobility-aware transmission dynamics, device-to-device (D2D) communication, immunity waning effects, and network-level quarantine mechanisms. Based on this extended formulation, a synthetic augmentation strategy is developed to capture the temporal evolution of IoT malware propagation and enrich the training data with realistic infection dynamics.

When combined with a Transformer–BiLSTM hybrid classifier, SEIR-DLFNet achieves strong performance on the integrated CICIoT2023 and CIC IoT-DIAD 2024 benchmarks, reaching an accuracy of 99.31% and an AUC-ROC of 99.44%, and consistently outperforming all evaluated baselines by a considerable margin. The ablation study further demonstrates that SEIR-guided synthetic augmentation alone contributes a 2.71 percentage point improvement in accuracy, highlighting the importance of explicitly modeling propagation behavior during training.

In addition, zero-shot evaluation on a polymorphic Mirai variant shows a 6.84 percentage point improvement over non-augmented baselines, confirming that simulation-informed training enhances generalization to previously unseen and evolving malware threats. Overall, these findings suggest that coupling epidemic modeling with deep learning provides a more robust and transferable learning paradigm for IoT malware detection in dynamic network environments.

Acknowledgements

This research was supported by the Internal Research Grant Program 2026 of Politeknik Negeri Batam. The authors sincerely appreciate the institutional support and resources provided, which contributed significantly to the successful completion of this work.

Credit Authorship Contribution Statement

Author1 : Conceptualization, Methodology, Software, Project administration. **Author2**: Software, Writting – original draft. **Author3**: Writing – review & editing. **Author4**: Validation, Supervision.

Declaration Of Competing Interests

-

Data Availability

Data will be made available on request.

References

[1] W. Melibari, H. Baodhah, and N. Akkari, "IoT-Based Smart Cities Beyond 2030: Enabling Technologies, Challenges, and Solutions," in *2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC)*, Jan. 2023, pp. 1–6. doi: 10.1109/ICAISC56366.2023.10085126.

[2] B. Taşcı, "Deep-Learning-Based Approach for IoT Attack and Malware Detection," *Applied Sciences*, vol. 14, no. 18, p. 8505, Jan. 2024, doi: 10.3390/app14188505.

[3] S. H. Bangash *et al.*, "Integrating Machine Learning and Deep Learning Approaches for Efficient Malware Detection in IoT-Based Smart Cities," *Journal of Computing & Biomedical Informatics*, vol. 5, no. 02, pp. 280–299, Sep. 2023.

[4] F. S. Alsubaei, "Smart deep learning model for enhanced IoT intrusion detection," *Scientific Reports*, vol. 15, no. 1, p. 20577, 2025.

[5] J. Carrillo-Mondejar, G. Suarez-Tangil, A. Costin, and R. J. Rodríguez, "Exploring shifting patterns in recent IoT malware," in *Proceedings of the European Conference on Cyber Warfare and Security*, Academic Conferences International Ltd, 2024. Accessed: May 16, 2026. [Online]. Available: https://jyx.jyu.fi/jyx/Record/jyx_123456789_96206

[6] Kaspersky, "IoT threats in 2023," Securelist. Accessed: May 16, 2026. [Online]. Available: https://securelist.com/iot-threat-report-2023/110644/

[7] M. Severt, R. Casado-Vara, and A. Martín del Rey, "A Comparison of Monte Carlo-Based and PINN Parameter Estimation Methods for Malware Identification in IoT Networks," *Technologies*, vol. 11, no. 5, p. 133, Oct. 2023, doi: 10.3390/technologies11050133.

[8] Y. Zhou, B.-T. Liu, K. Zhou, and S.-F. Shen, "Malware propagation model of fractional order, optimal control strategy and simulations," *Frontiers in Physics*, vol. 11, p. 1201053, 2023, doi: doi.org/10.3389/fphy.2023.1201053.

[9] L. Quiroga-Sánchez, G. A. Montoya, and C. Lozano-Garzon, "The SEIRS-NIMFA epidemiological model for malware propagation analysis in IoT networks," *Cybersecurity*, vol. 8, no. 1, p. 2, Jan. 2025, doi: 10.1186/s42400-024-00310-z.

[10] E. Asadi and S. Hosseini, "Modeling of Malware Propagation Under the Clustering Approach in Scale-Free Networks," *SECURITY AND PRIVACY*, vol. 8, no. 1, p. e465, 2025, doi: 10.1002/spy2.465.

[11] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment," *Sensors*, vol. 23, no. 13, p. 5941, 2023, doi: doi.org/10.3390/s23135941.

[12] N. Latif, W. Ma, and H. B. Ahmad, "Advancements in securing federated learning with IDS: a comprehensive review of neural networks and feature engineering techniques for malicious client detection," *Artif Intell Rev*, vol. 58, no. 3, p. 91, Jan. 2025, doi: 10.1007/s10462-024-11082-w.

[13] A. A. Almazroi and N. Ayub, "Deep learning hybridization for improved malware detection in smart Internet of Things," *Scientific reports*, vol. 14, no. 1, p. 7838, 2024.

[14] S. K. Smmarwar, G. P. Gupta, and S. Kumar, "Android malware detection and identification frameworks by leveraging the machine and deep learning techniques: A comprehensive review," *Telematics and Informatics Reports*, vol. 14, p. 100130, 2024.

[15] S. Shen *et al.*, "HSIRD: A model for characterizing dynamics of malware diffusion in heterogeneous WSNs," *Journal of Network and Computer Applications*, vol. 146, p. 102420, 2019.

[16] M. Ibrahim, M. T. Abdullah, A. Abdullah, and T. Perumal, "Modelling and mitigation strategy of IoT botnet propagation," 2019, Accessed: May 16, 2026. [Online]. Available: https://www.preprints.org/frontend/manuscript/0415c81d57cb2fa9a90c466447a9c042/download_pub

[17] D. Acarali, M. Rajarajan, N. Komninos, and B. B. Zarpelão, "Modelling the Spread of Botnet Malware in IoT-Based Wireless Sensor Networks," *Security and Communication Networks*, vol. 2019, pp. 1–13, Feb. 2019, doi: 10.1155/2019/3745619.

[18] M. J. Farooq and Q. Zhu, "Modeling, analysis, and mitigation of dynamic botnet formation in wireless IoT networks," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 9, pp. 2412–2426, 2019.

[19] B.-R. Chen, S.-M. Cheng, and M. B. Mwangi, "A Mobility-Based Epidemic Model for IoT Malware Spread," *IEEE Access*, vol. 10, pp. 107929–107941, 2022, doi: 10.1109/ACCESS.2022.3213032.

[20] S. Riaz *et al.*, “Malware detection in internet of things (IoT) devices using deep learning,” *Sensors*, vol. 22, no. 23, p. 9305, 2022.
[21] N. Almakayeel, “Deep learning-based improved transformer model on android malware detection and classification in internet of vehicles,” *Scientific Reports*, vol. 14, no. 1, p. 25175, 2024.
[22] W. Almobaideen, O. Abu Alghanam, M. Abdullah, S. B. Hussain, and U. Alam, “Comprehensive review on machine learning and deep learning techniques for malware detection in android and IoT devices,” *Int. J. Inf. Secur.*, vol. 24, no. 3, p. 110, Jun. 2025, doi: 10.1007/s10207-025-01027-x.
[23] M. Kharabsheh, I. Al-aiash, A. Mughaid, and M. Almiani, “The seir model for predicting malware propagation in computer networks,” in *2024 International Conference on Intelligent Computing, Communication, Networking and Services (ICCNS)*, IEEE, 2024, pp. 108–113. Accessed: May 16, 2026. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10776333/
[24] A. Chernikova, N. Gozzi, N. Perra, S. Boboila, T. Eliassi-Rad, and A. Oprea, “Modeling self-propagating malware with epidemiological models,” *Appl Netw Sci*, vol. 8, no. 1, p. 52, Aug. 2023, doi: 10.1007/s41109-023-00578-z.
[25] A. Baazeem, M. S. Arif, Y. Nawaz, and K. Abodayeh, “Modeling and Simulation of Epidemics Using q-Diffusion-Based SEIR Framework with Stochastic Perturbations,” *CMES*, vol. 143, no. 3, pp. 3463–3489, 2025, doi: 10.32604/cmes.2025.066299.
[26] Canadian Institute for Cybersecurity, “CIC IoT-DIAD 2024: A dual-function dataset for IoT device identification and anomaly detection.” Accessed: May 04, 2026. [Online]. Available: https://www.unb.ca/cic/datasets/iot-diad-2024.html
[27] S.-B. Hsu and K.-C. Chen, *Ordinary differential equations with applications*, vol. 23. World scientific, 2022. Accessed: May 16, 2026. [Online]. Available: https://books.google.com/books?hl=id&lr=&id=tTioEAAAQBAJ&oi=fnd&pg=PR5&dq=system+of+ordinary+differential+equations+(ODEs&ots=9GUBBZVwgh&sig=3AwiIQEfhB_Tl-fsyaIUNy2O45o
[28] M. T. Hoang, “Mathematical analysis and numerical simulation of a generalized epidemiological model for malware propagation,” *Nonlinear Dyn*, vol. 114, no. 1, p. 53, Dec. 2025, doi: 10.1007/s11071-025-11912-8.
[29] O. Jurečková, M. Jureček, M. Stamp, F. Di Troia, and R. Lórencz, “Classification and online clustering of zero-day malware,” *J Comput Virol Hack Tech*, vol. 20, no. 4, pp. 579–592, Feb. 2024, doi: 10.1007/s11416-024-00513-5.
[30] O. Hospodarskyy, V. Martsenyuk, N. Kukharska, A. Hospodarskyy, and S. Sverstiuk, “Understanding the Adam Optimization Algorithm in Machine Learning.,” *CITI*, vol. 2024, p. 2nd, 2024.