iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83352

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Research and Development of Adaptive Mathematical Modeling for Malware Prediction and Analysis

Assylbek Dias Orynbekuly[1], Galymzhanov Alisher Maratovich[2], Beketova Gulzhanat Sakitzhanovna[3]
[1, 2]*Master's Student, Almaty University of Power Engineering and Telecommunications, Almaty, Kazakhstan*
[3]*PhD, Associate Professor, Almaty University of Power Engineering and Telecommunications, Almaty, Kazakhstan*

Abstract: ***The purpose of this study is to develop and numerically analyze an improved mathematical model for the spread of malware in network structures based on a modified SIR immune response model. The research methodology is based on constructing a model that takes into account the processes of infection, recovery, and loss of immunity of network nodes, as well as applying the fourth-order Runge–Kutta numerical method to calculate the dynamics of malware propagation. During the simulation, key parameters of cyber threat propagation were determined, including infection, recovery, and immunity loss rates. The obtained results showed that the maximum infection level reaches 34.7% of the total number of network nodes, while the malware propagation peak occurs after 32.5 conditional time units. The research results confirm that response speed, timely software updates, and continuous monitoring of network activity have a significant impact on reducing the scale of infection. It is concluded that the proposed model can be used as a tool for the quantitative assessment of cyber threats and for justifying measures aimed at increasing the resilience of network systems to malicious attacks. The practical significance of the study lies in the possibility of applying the obtained results in developing strategies for the optimal allocation of information security resources and preventing the widespread propagation of malware.***
Keywords: Cybersecurity, mathematical modeling, malware, SIR model, cyber threat.

I. INTRODUCTION

In the context of the rapid growth and development of cyber threats, traditional protection methods, which are often reactive in nature, are becoming less effective. There is an urgent need for modern tools capable of modeling and forecasting the spread of malware, which is crucial for proactive defense and the optimization of cybersecurity resources.

One of the key problems is the insufficient understanding of how factors such as the infection rate, the effectiveness of implemented countermeasures, and the probability of reinfection affect the dynamics of threat propagation in a network environment. This limits the ability to make well-grounded decisions when developing security strategies.

Mathematical modeling of malware propagation in networks has a long history. The fundamental basis was established in [1], where epidemiological models were first applied to describe the spread of computer viruses. Subsequently, various models were developed and adapted to specific conditions and types of networks, including models that take into account network saturation [2] and node scanning [3].

SIR models and their modifications, such as SEIRS [4] and SIRS models with time delay [5], are widely used and provide methods for analyzing malware propagation in wireless and other network environments. Modern research is shifting its focus toward complex network structures [6–7] and the human factor [8], which makes the models more realistic and comprehensive in describing everyday network interactions. Threat modeling related to risk assessment for cyber-physical systems is also an important research area [9]. Another significant aspect of mathematical modeling in cybersecurity is the integration of vulnerabilities and threats into a unified system. An example of such an approach is the ICAR model proposed by Hemberg et al. [10], which uses category theory to establish mathematical relationships.

Traditional protection mechanisms are unable to keep pace with the rapidly changing tactics of cyberattacks, in which attackers use advanced technologies such as machine learning and deep learning to evade detection. In response, researchers are exploring the use of machine learning and game theory to develop more effective cybersecurity solutions [11–12].

Commercial antivirus products remain one of the main means of protecting computer systems. Many researchers [13–18] propose using deep learning for malware classification as a key component of next-generation protection systems.

Many authors have focused on adversarial learning-based attacks, but only a few have proposed defense mechanisms [19] that involve learning undesirable patterns. In 2015, Papernot et al. [20] introduced a defense against attacks based on defensive distillation, which is grounded in adversarial learning.

The purpose of this study is to examine modern approaches to modeling malware propagation in order to identify their strengths and weaknesses, to develop an improved model that takes into account the unique characteristics of modern cyber threats and network infrastructure, and, based on the obtained results, to propose specific recommendations for improving existing strategies and developing new approaches to protection against cyber threats.

II. METHODS

Let us consider a mathematical model of malware propagation in a network using a modified SIR model, which consists of three groups: Susceptible, Infected, and Recovered.

$$\begin{cases} \frac{dS}{dt} = -\beta SI + \gamma R \\ \frac{dI}{dt} = \beta SI - \delta I \\ \frac{dR}{dt} = \delta I - \gamma R \end{cases}, \quad (1)$$

where **S** is the number of susceptible nodes, **I** is the number of infected nodes, **R** is the number of protected nodes, **β** is the infection rate, **δ** is the recovery rate, and **γ** is the rate of immunity loss. Susceptible nodes **S** can become infected with a probability proportional to the number of infected nodes **I**. Infected nodes **I** can recover at a certain rate. Recovered nodes **R** may become susceptible again over time.

The application of the fourth-order Runge–Kutta method to the SIR model (1) begins with defining a function for each equation of the system:

$$\begin{cases} f_1(t,S,I,R) = -\beta SI + \gamma R \\ f_2(t,S,I,R) = \beta SI - \delta I \\ f_3(t,S,I,R) = \delta I - \gamma R \end{cases}. \quad (2)$$

The required parameters **S**, **I**, and **R** are expressed as follows:

$$S_{n+1} = S_n + \frac{h}{6} \cdot (k_{1,1} + 2k_{2,1} + 2k_{3,1} + k_{4,1}),$$

$$I_{n+1} = I_n + \frac{h}{6} \cdot (k_{1,2} + 2k_{2,2} + 2k_{3,2} + k_{4,2}),$$

$$R_{n+1} = R_n + \frac{h}{6} \cdot (k_{1,3} + 2k_{2,3} + 2k_{3,3} + k_{4,3}),$$

$$t_{n+1} = t_n + h,$$

Where $h = \frac{b-a}{N}$ is the step size. The coefficients are determined as follows:

$$k_{1,i} = f_i(x_n, S_n, I_n, R_n),$$

$$k_{2,i} = f_i\left(t_n + \frac{h}{2}, S_n + \frac{h}{2}k_{1,1}, I_n + \frac{h}{2}k_{1,2}, R_n + \frac{h}{2}k_{1,3}\right),$$

$$k_{3,i} = f_i\left(t_n + \frac{h}{2}, S_n + \frac{h}{2}k_{2,1}, I_n + \frac{h}{2}k_{2,2}, R_n + \frac{h}{2}k_{2,3}\right),$$

$$k_{4,i} = f_i\left(t_n + h, S_n + hk_{3,1}, I_n + hk_{3,2}, R_n + hk_{3,3}\right),$$

Here $i = 1,2,3$ for S, I, R respectively.

For the specific system (2), the coefficients ... will be calculated as follows:

$$k_{1,1} = -\beta S_n I_n + \gamma R_n,$$

$$k_{1,2} = \beta S_n I_n - \delta I_n,$$

$$k_{1,3} = \delta I_n - \gamma R_n,$$

$$k_{2,1} = -\beta\left(S_n + \frac{h}{2}k_{1,1}\right)\left(I_n + \frac{h}{2}k_{1,2}\right) + \gamma\left(R_n + \frac{h}{2}k_{1,3}\right),$$

$$k_{2,2} = \beta\left(S_n + \frac{h}{2}k_{1,1}\right)\left(I_n + \frac{h}{2}k_{1,2}\right) - \delta\left(I_n + \frac{h}{2}k_{1,2}\right),$$

$$k_{2,3} = \delta\left(I_n + \frac{h}{2}k_{1,2}\right) - \gamma\left(R_n + \frac{h}{2}k_{1,3}\right),$$

$$k_{3,1} = -\beta\left(S_n + \frac{h}{2}k_{2,1}\right)\left(I_n + \frac{h}{2}k_{2,2}\right) + \gamma\left(R_n + \frac{h}{2}k_{2,3}\right),$$

$$k_{3,2} = \beta\left(S_n + \frac{h}{2}k_{2,1}\right)\left(I_n + \frac{h}{2}k_{2,2}\right) - \delta\left(I_n + \frac{h}{2}k_{2,2}\right),$$

$$k_{3,3} = \delta\left(I_n + \frac{h}{2}k_{2,2}\right) - \gamma\left(R_n + \frac{h}{2}k_{2,3}\right),$$

$$k_{4,1} = -\beta(S_n + hk_{3,1})(I_n + hk_{3,2}) + \gamma(R_n + hk_{3,3}),$$

$$k_{4,2} = \beta(S_n + hk_{3,1})(I_n + hk_{3,2}) - \delta(I_n + hk_{3,2}),$$

$$k_{4,3} = \delta(I_n + hk_{3,2}) - \gamma(R_n + hk_{3,3}).$$

This method makes it possible to numerically solve the system of differential equations of the SIR model with high accuracy, taking into account the nonlinearity of interactions between different groups — susceptible, infected, and recovered nodes — during the spread of malware across the network.

For the numerical implementation of the SIR model, the coefficient values of the system of equations (1) and the initial conditions of the target variables presented in Table 1 will be used.

The values $\beta = 0.3$, $\delta = 0.1$, and $\gamma = 0.05$ are based on the analysis of modern malware and its ability to spread rapidly. These values take into account the average propagation rate of different types of malware, ranging from relatively slow worms to fast-spreading botnets, as well as both automated protection tools and manual intervention by administrators, security update cycles, and the emergence of new malware versions.

Table I. Modeling Parameters

Definition	**Value**
Infection rate, β	0.3
Recovery rate, δ	0.1
Immunity loss rate, γ	0.05
Initial number of nodes in the network, N	1000
Initial number of susceptible nodes, S	995
Initial number of infected nodes, I	5
Number of protected nodes, R	0

These parameters were selected based on an analysis of data on the real-world spread of malware and consultations with cybersecurity experts. They provide a realistic representation of malware propagation dynamics in modern network environments, taking into account both the technical aspects of virus spread and the organizational factors that influence response and recovery rates.

III.RESULTS AND DISCUSSION

Figure 1 shows the spread of malware in the network. A rapid increase in the number of infected nodes is observed, followed by a peak, after which the number of infected nodes begins to decrease.

The maximum number of infected nodes reaches approximately 347, which accounts for about 34.7% of the entire network. This occurs approximately 32.5 time units after the beginning of the spread. By the end of the simulated period, which is 100 time units, about 442 nodes, or 44.2% of the network, are in the recovered state.

To minimize the initial spread of malware, early detection and rapid response systems should be implemented. Regular updates and patches can increase the recovery rate δ and reduce network vulnerability. After the main wave of infection has declined, monitoring should continue, as some infected nodes may remain in the network. Increasing user awareness can reduce the infection rate β and improve the overall resilience of the network. Considering the possibility of immunity loss γ, it is important to continuously adapt protective measures to new threats.

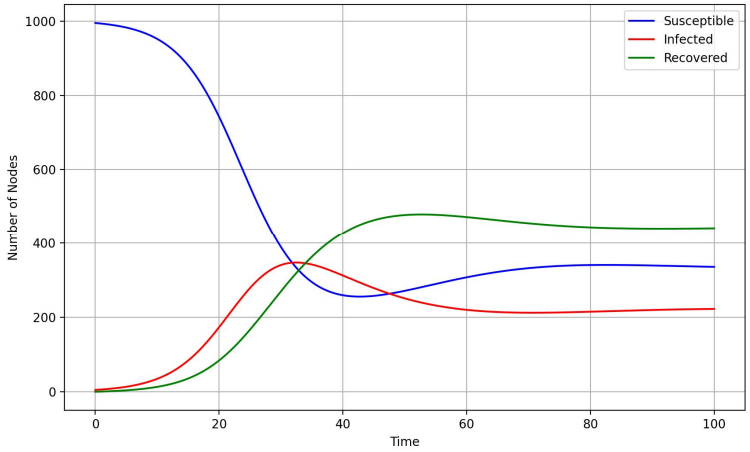


Fig. 1 Results of the SIR Model

Figure 2 shows how rapidly malware spreads across the network over time, reaching the infection peak and then beginning to decline. The infection peak indicates the moment when the virus spreads most intensively. After this point, containment and elimination measures begin to have a significant impact.

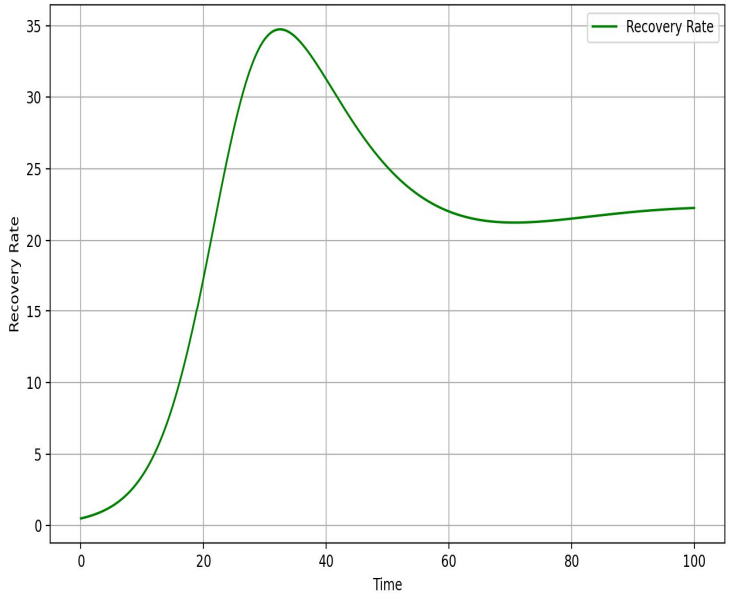


Fig 2. Dependence of the Infection Rate on Time

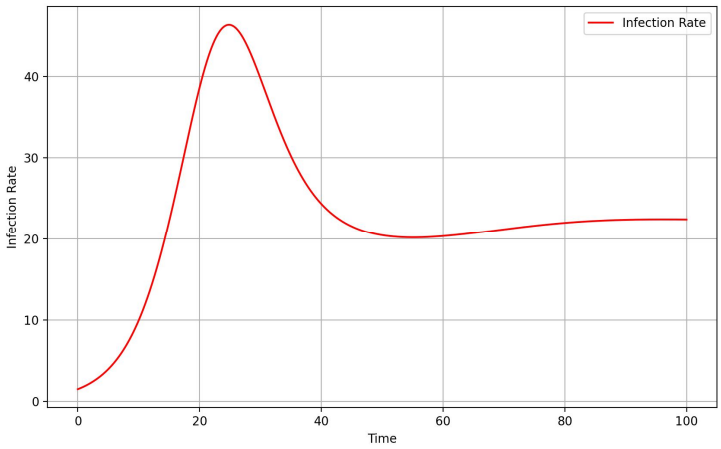


Figure 3. Dependence of the Recovery Rate on Time

In Figure 4, the distribution of node states at the end of the simulation provides a clear representation of the final state of the network after the infection wave has passed. A high percentage of recovered nodes indicates that the threat was successfully contained and eliminated. The low number of infected nodes at the end of the period shows that the infection was almost completely suppressed.

The model demonstrates typical infection behavior, where the initial phase is characterized by a rapid increase in infections, followed by a recovery phase.

The effectiveness of recovery and threat containment measures is confirmed by the high percentage of recovered nodes.

It is important to continue monitoring and maintaining security measures in order to prevent repeated attacks or new threats.

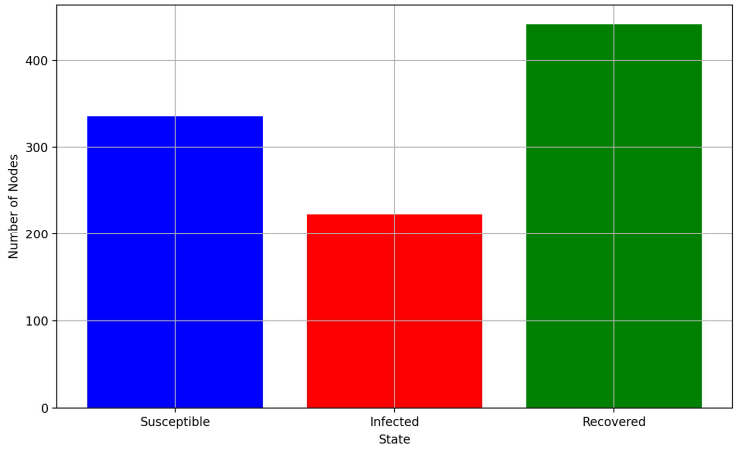


Figure 4. Distribution of Node States at the End of the Simulation

Our study demonstrates that effective management of the parameters of the SIR model can significantly influence the dynamics of malware propagation in networks. Without timely intervention and updates, the infection peak may be substantial, which emphasizes the need for rapid response and the development of strategies to improve security. The basic reproduction number R_0 shows the potential for rapid infection spread, requiring measures aimed at reducing the infection rate and increasing the recovery rate. An increase in the recovery rate δ proved effective in reducing the infection peak and protecting the network from prolonged malicious attacks. The immunity loss parameter γ indicates the need for continuous adaptation of protective measures and updates to increase resilience against new threats.

These results highlight the need for preventive measures and continuous monitoring in real-world scenarios, which can help protect the network from cyber threats. They emphasize the importance of dynamic adaptation and parameter management within the model to improve cybersecurity and network resilience.

IV. CONCLUSIONS

The article presents a comprehensive analysis of the application of a modified SIR model for modeling and understanding malware propagation in network infrastructures. The results of the study emphasize the importance of a clear understanding and management of indicators such as infection rate, recovery rate, and immunity loss rate, which are key components in developing effective strategies to combat cyber threats.

The modeling demonstrates the critical role of rapid response measures, such as updating network security protocols and training users, in reducing infection peaks and protecting the network. The obtained data indicate that even after the main wave of infection has declined, the network may remain vulnerable, which highlights the need for continuous monitoring and adaptation of security measures.

The research findings can be practically applied to strengthen cybersecurity in various sectors, including corporate networks and government information systems. The use of the identified strategies helps increase system resilience against new types of threats.

Thus, this study makes a significant contribution to understanding the dynamics of malware propagation and proposes relevant solutions for improving the cybersecurity of modern networks.

REFERENCES

[1] J. O. Kephart and S. R. White, «Directed-graph epidemiological models of computer viruses», in Proc. IEEE Comput. Soc. Symp. Res. Security Privacy, 1991, pp. 343-359, https://doi.org/10.1109/RISP.1991.130801.

[2] C. C. Zou, W. Gong, and D. Towsley, «Code red worm propagation modeling and analysis», in Proc. 9th ACM Conf. Comput. Commun. Security, 2002, pp. 138-147, https://doi.org/10.1145/586110.586130.

[3] Y. Wang, C. Wang, and C. C. Zou, «Modeling the propagation and defense of internet e-mail worms», IEEE Trans. Dependable Secure Comput., vol. 4, no. 2, pp. 105-118, Apr.-Jun. 2007, https://doi.org/10.1109/TDSC.2007.1001.

[4] B. K. Mishra and N. Jha, «SEIQRS model for the transmission of malicious objects in computer network», Appl. Math. Model., vol. 34, no. 3, pp. 710-715, Mar. 2010, https://doi.org/10.1016/j.apm.2009.06.011.

[5] L. X. Yang and X. Yang, «A new epidemic model of computer viruses», Commun. Nonlinear Sci. Numer. Simul., vol. 17, no. 11, pp. 5324-5331, Nov. 2012, https://doi.org/10.1016/j.cnsns.2012.05.030.

[6] R. Pastor-Satorras and A. Vespignani, «Epidemic spreading in scale-free networks», Phys. Rev. Lett., vol. 86, no. 14, p. 3200, Apr. 2001, https://doi.org/10.1103/PhysRevLett.86.3200.

[7] C. H. Nwokoye and V. Madhusudanan, "Epidemic models of malicious-code propagation and control in wireless sensor networks: An in-depth review," Wireless Personal Communications, vol. 125, pp. 1827–1856, 2022, doi: 10.1007/s11277-022-09636-8.

[8] N. P. Dong, H. V. Long, and N. T. K. Son, "The dynamical behaviors of fractional-order SE_1E_2IQR epidemic model for malware propagation on wireless sensor network," Communications in Nonlinear Science and Numerical Simulation, vol. 111, 2022, Art. no. 106428, doi: 10.1016/j.cnsns.2022.106428.

[9] S. M. Al-Tuwairqi and W. S. Bahashwan, "The impact of quarantine strategies on malware dynamics in a network with heterogeneous immunity," Mathematical Modelling and Analysis, vol. 27, no. 2, pp. 282–302, 2022, doi: 10.3846/mma.2022.14391.

[10] A. Valence, «ICAR, a categorical framework to connect vulnerability, threat and asset managements», Cryptography and Security, Jun. 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2306.12240.

[11] W. Zhang, Z. Wang, Z. Zhang, and J. Zou, "Delay effect on a malware propagation model incorporating user awareness," in Proc. International Conference on Cyber-Physical Social Intelligence (ICCSI), 2022, pp. 555–560, doi: 10.1109/ICCSI55536.2022.9970556.

[12] A. Wolsey, "The State-of-the-Art in AI-Based Malware Detection Techniques: A Review, «Cryptography and Security», Oct. 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2210.11239.

[13] A. Chernikova, N. Gozzi, N. Perra, S. Boboila, T. Eliassi-Rad, and A. Oprea, "Modeling self-propagating malware with epidemiological models," Applied Network Science, vol. 8, 2023, Art. no. 52, doi: 10.1007/s41109-023-00578-z.

[14] R. Pascanu, J. W. Stokes, H. Sanossian, M. Marinescu, and A. Thomas, «Malware classification with recurrent networks», in Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP), 2015, pp. 1916–1920, https://doi.org/10.1109/ICASSP.2015.7178304.

[15] W. Huang and J. W. Stokes, «Mtnet: A multi-task neural network for dynamic malware classification», in Proc. Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA), 2016, pp. 399–418, https://doi.org/10.1007/978-3-319-40667-1_20.

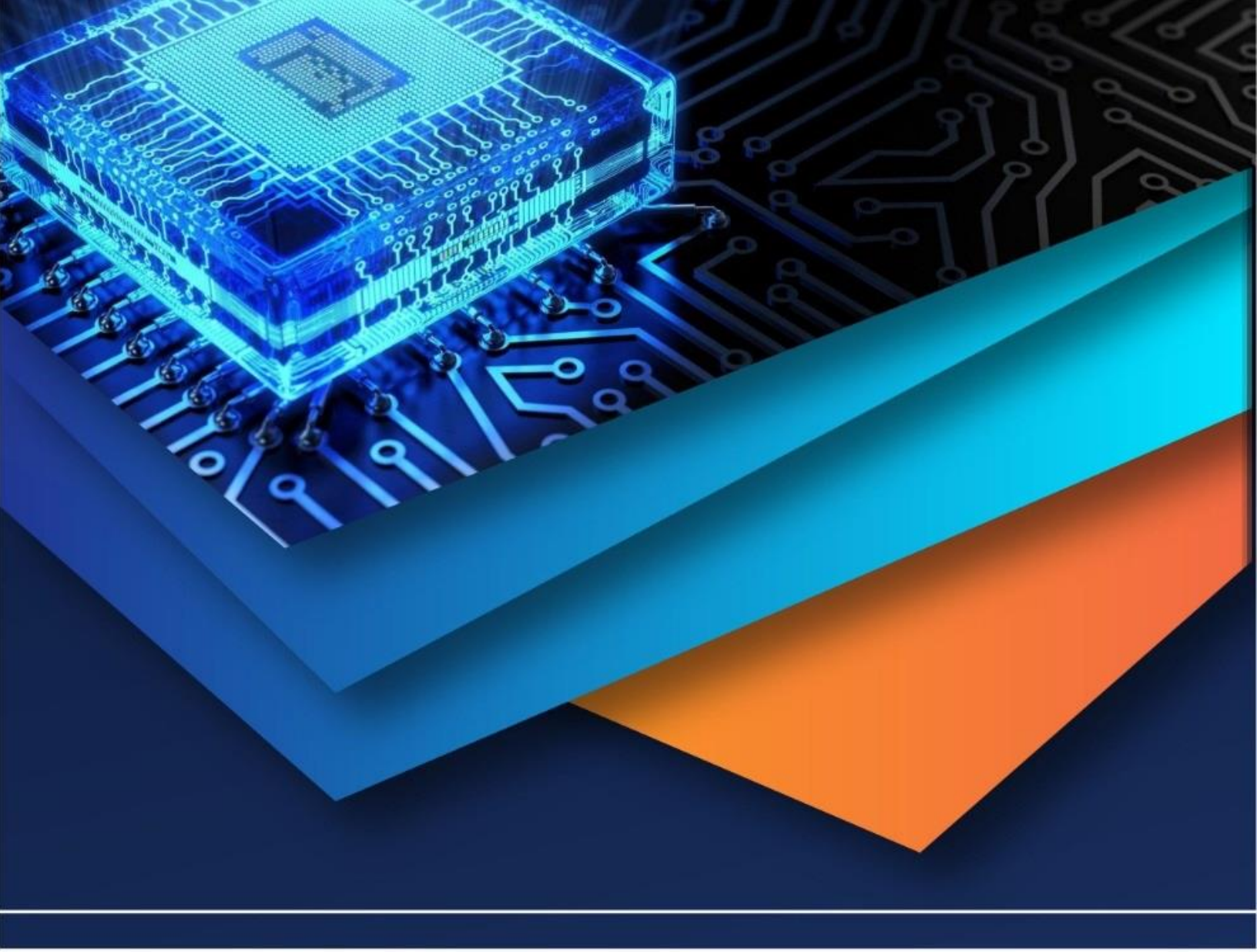
[16] B. Athiwaratkun and J. W. Stokes, «Malware classification with LSTM and GRU language models and a character-level CNN», in Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP), 2017, pp. 2482–2486, https://doi.org/10.1109/ICASSP.2017.7952603.

[17] L. Tian, F. Shang, and C. Gan, "Optimal control analysis of malware propagation in cloud environments," Mathematical Biosciences and Engineering, vol. 20, no. 8, pp. 14502–14517, 2023, doi: 10.3934/mbe.2023649.

[18] I. J. Goodfellow, J. Shlens, and C. Szegedy, «Explaining and harnessing adversarial examples», in Proc. Int. Conf. Learn. Representations (ICLR), 2015.

[19] M. T. Jafar, L.-X. Yang, G. Li, Q. Zhu, C. Gan, and X. Yang, "Malware containment with immediate response in IoT networks: An optimal control approach," Computer Communications, vol. 228, 2024, Art. no. 107951, doi: 10.1016/j.comcom.2024.107951.

[20] O. A. M. Omar, H. M. Ahmed, T. A. Nofal, A. Darwish, and A. M. Sayed Ahmed, "Analysis and optimal control of propagation model for malware in multi-cloud environments with impact of Brownian motion process," Mathematical and Computational Applications, vol. 30, no. 1, 2025, Art. no. 8, doi: 10.3390/mca30010008.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY