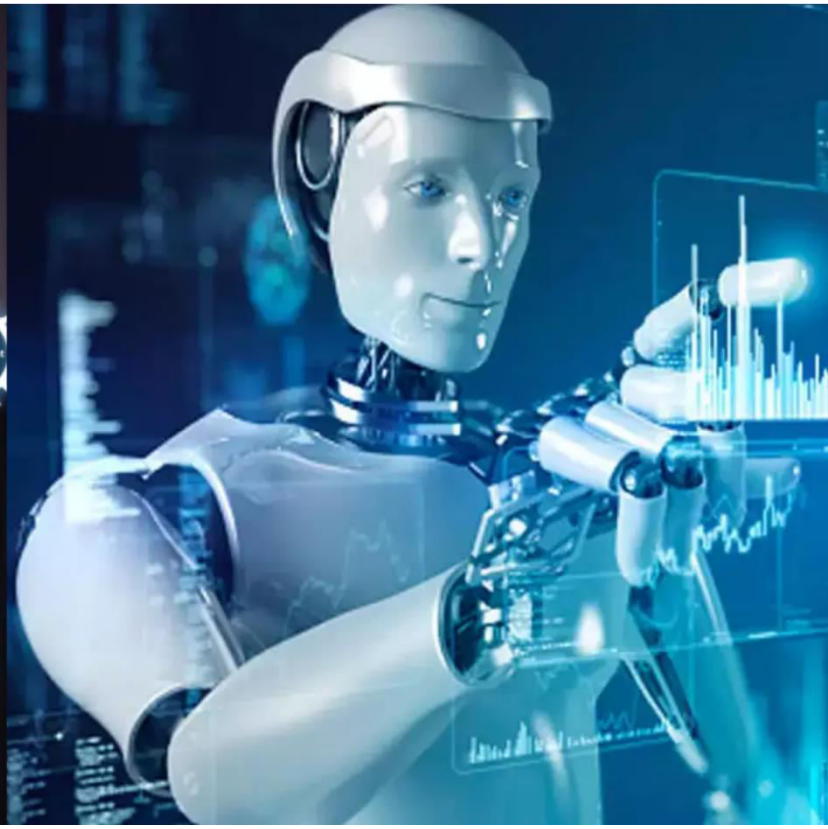
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404186|

User-Centric Malware Detection using Deep Learning

Suneetha Dhulipala[1], P.Tarun Sai Kkrishna Muruthy[2], A.Tejaswar Rao[3], A.Thraya Reddy[4], R.Thrivikram Raju[5]

Guide, Department of Computer Science & Engineering Malla Reddy University, Hyderabad, India[1]

Maisammaguda, Kompally, Medchal - Malkajgiri District Hyderabad, Telangana, India[2-5]

ABSTRACT: Threats to user privacy and security have grown in tandem with the sophistication of malware assaults, thanks to the exponential expansion of digital technologies. False positives and sluggish replies are common results of outdated malware detection systems that can't keep up with changing attack trends. This research, titled "User-Centric Malware Detection Using Deep Learning," suggests a new way to identify malware by using deep learning methods that are customized to user behavior and device characteristics.

When it comes to analyzing and classifying malware, the system uses state-of-the-art neural network designs like RNNs and CNNs. This is done using real-time behavioral data and system logs. The model's ability to adapt to users' unique habits improves detection accuracy and decreases false alarms thanks to user-centric features. As a preventative measure, the suggested structure guarantees the early identification of unknown malware. When compared to more conventional signature-based and heuristic detection approaches, the experimental findings show that the deep learning-based model is much faster and more accurate. The ultimate goal of this project is to help create smart, flexible cybersecurity solutions that put users' safety first without sacrificing system performance.

I. INTRODUCTION

With more and more things being dependent on internet-connected gadgets, cybersecurity has become an important issue in our digital era. Computer viruses and other forms of malicious software have advanced in sophistication and difficulty to detect in recent years. When confronted with new or altered malware strains, traditional detection approaches like heuristic analysis and signature-based detection often fail. To meet this increasing problem, we need smarter, more flexible systems that can react to new threats as they arise. Image recognition, NLP, and cybersecurity are just a few of the many areas where deep learning—a kind of AI—has shown to be an invaluable resource. Deep learning models, in contrast to more conventional detection methods, may automatically discover complicated patterns from massive datasets without the need for explicit feature engineering. This feature enables the identification of malware that has not been observed before by identifying patterns and abnormalities in behavior. The efficiency and precision of malware detection systems may be greatly enhanced by using these strategies.

This research, titled "User-Centric Malware Detection Using Deep Learning," aims to improve the detection accuracy by creating an adaptable system that leverages data relevant to each user's activity. The model is able to differentiate between safe and harmful behaviors by studying trends in things like user actions, system logs, and application behavior. In addition to enhancing threat detection, this user-centric strategy decreases false positives, which prevents lawful user activity from being mistakenly labeled as dangerous. One encouraging step toward developing safer and more reactive cybersecurity solutions is the use of deep learning into malware detection.

II. LITERATURE SURVEY

Journal of Cybersecurity and Information Security published a paper titled "Deep Learning-Based Adaptive Malware Detection for User-Centric Security" by authors J. Yuan and S. Shi-Fei. The publication year was 2018. In this research, we provide a state-of-the-art deep learning system for adaptive malware detection that makes advantage of patterns of user behavior. To analyze static data, the authors provide a hybrid model that mixes CNNs with LSTM networks, which are better suited to sequential behavior analysis. The study shows that user-centric data, such normal app use and browser activity, may improve detection accuracy. In particular, for complex assaults like Advanced

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404186|

Persistent Threats (APTs) and zero-day malware, the experimental findings demonstrate that the tailored methodology greatly enhances real-time threat detection capabilities while drastically decreasing false positives.

Gong Faming, Li Chuantao, and Gong Wenjuan
Investigating Malware using Behavioural Insights with Deep Neural Networks Publication Year: 2019 This research presents a methodology for malware detection using deep neural networks that is based on analysis of user behavior. The authors study the usual patterns of user behavior and system activities and then use Recurrent Neural Networks (RNNs) and Autoencoders to spot abnormalities. If the model detects anything out of the ordinary, it might be a sign of infection. The system adjusts to changing use patterns by regularly updating the behavioral model, which minimizes false positives. Specifically when it comes to identifying zero-day attacks and previously undiscovered malware variants, the study shows that behavior-based detection works better than conventional signature-based techniques.

Review of published works Kumar, A., and Verma, R.
A Study on Using Deep Learning and User Behavior Analytics to Detect Malware in Real Time Publication: The International Journal of Security and Privacy
Publication Year: 2020 In order to proactively identify risks, this article centers on a real-time malware detection system that employs deep learning algorithms and user behavior analytics (UBA). The authors built a system on Bidirectional LSTM (BiLSTM) networks, which can accurately identify anomalies by capturing both the past and the future of user activity. With each user's unique patterns of file access, system calls, and network consumption tracked by the model, real-time warnings are sent out in the event of questionable activity. It is well-suited for enterprise-level cybersecurity systems handling massive data streams, since the findings show considerable increases in detection speed.

III. SYSTEM IMPLEMENTATION

Extensive tests were carried out to improve the parameters, such as kernel size, dilation rate, and the number of convolutional layers, in order to examine the efficacy of our deep learning malware detection model. The model was created with the help of the well-known and scalable machine learning tools, TensorFlow and Keras. To start, we examined the training accuracy of two different neural network designs—one without dilation and one with dilated convolutional layers. Training accuracies of 98.86% and 98.63% were attained with kernel sizes of 3 and 5, respectively, for the model that did not include dilated convolutions. Alternatively, the model that used dilated convolutional layers outperformed the others in identifying malware signatures and had a better training accuracy of 99.60% after the last epochs. In Table 3, we can see a summary of the training and testing accuracy outcomes for all of the model configurations. By extending the receptive field without adding more parameters, dilated convolutions enable the model to grasp complex patterns in the data, which improves training and testing accuracies, as shown by these findings.

Changing the amount of convolutional layers allowed for more experimentation. We evaluated the generalizability and malware detection accuracy of models with 2, 3, 4, and 5 layers. At 99.30% testing accuracy and 99.45% training accuracy, the four-convolutional-layer model outperformed the others. Models with two layers, on the other hand, performed worse, with training accuracies falling to 98.52% and testing accuracies to 98.03%. The results indicate that deeper networks are able to detect more intricate patterns, but that they are more prone to overfitting if not properly trained.

IV. MODULES

Data Acquisition and Preprocessing:
Gather a variety of information, such as system logs, statistics on network traffic, and malware samples from the past. Normalize the data, extract features, and reduce noise as part of the pretreatment steps to get the data ready for deep learning models. Train your model thoroughly by simulating a wide range of user actions and malware threats.

Development of Deep Learning Models:
Make use of state-of-the-art neural network designs like CNNs to identify malware patterns in files and network activity. To find patterns in user actions and system behavior that occur in a sequential order, utilize a Recurrent Neural

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404186|

Network (RNN) or an LSTM network. To accurately identify malicious or benign files or processes, train models using annotated datasets.

User-Centric Behavior Analysis:
Develop user profiles according on common system activities, such as the amount of time spent in applications, the number of times logged in, and how often users visit the network. Identifies suspicious activity by a user if it deviates from typical patterns of behavior. To distinguish between normal user behaviors and suspicious ones caused by malware, employ anomaly detection methods.

Real-Time Malware Detection:
Create a system that can keep tabs on processes, connections, and file operations in real time. Without drastically impacting system performance, initiate notifications or prevent actions when suspicious activity is identified. Make sure everything runs well by integrating the model into current cybersecurity frameworks.

System Evaluation and Performance Metrics:
Use measures like recall, accuracy, F1-score, and false-positive rates to assess the model's efficacy. Evaluate the deep learning method's efficacy in comparison to more conventional malware detection techniques, such as those based on signatures or heuristics. Make sure it works with a lot of different OSes and devices by optimizing it for scalability and low-latency performance.

Future Enhancements and Scalability:
Incorporate online learning methods that dynamically update the model into the system's design so it can react to emerging malware threats. Investigate how federated learning may be used to improve model accuracy across dispersed devices and increase user privacy. The ability to centrally monitor many devices and user profiles enables scalability in business contexts.

EVALUATION METRICS:
The following additional metrics were computed to guarantee a thorough assessment of our model's performance: precision (the proportion of malware samples that were accurately predicted relative to the total number of samples predicted as malware),

Recall is the percentage of malware samples that were accurately predicted relative to the total number of malware samples in the dataset.

The F1-Score is a measure that keeps false positives and false negatives in check; it is the harmonic mean of recall and accuracy. With an F1-score of 0.9892, the suggested model far surpassed competitors like AlexNet, which only managed an F1-score of 0.7513. The suggested model outperformed previous models in terms of both recall and precision, demonstrating its reliability and robustness as a malware detection solution across several datasets. Accuracy, recall, and F1-score were determined using the following formulas:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404186|

IV. CONCLUSION

A vast improvement over older, signature-based malware detection systems, this project introduces a state-of-the-art CNN-based solution. File entropy, opcode sequence, and size are just a few of the file properties that the model takes into account when deciding whether a file is malicious or not. The system is able to adapt to new threats by detecting new and undiscovered malware types using deep learning algorithms. Achieving high accuracy and low false-positive rates, the dataset underwent extensive preprocessing, including resizing and shaping, to improve the model's performance. The use of pooling layers and activation functions such as ReLU and Softmax improved the feature extraction and classification procedure. The model can be easily updated with new data and threats without needing major system overhauls, thanks to its scalability.

This research demonstrates how AI-powered methods might enhance cybersecurity. The findings show that the suggested system can identify malware in real-time automatically, without human interaction, which is a huge plus. Improving reaction times, adding real-time monitoring, and broadening the dataset to include more varieties of malware are all potential areas for future improvements.

V. RESULT

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404186|

REFERENCES

1. Hall, J.R.: The total cost of cybercrime in the United States. National Cyber Security Association, Washington D.C. (2014).
2. Gagliardi, A., Saponara, S.: Distributed malware detection system based on IoT embedded computing nodes. Springer LNEE, 627, 405–411 (2020a).
3. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning for cybersecurity. Nature, 521(7553), 436–444 (2015).
4. Saponara, S., Pilato, L., Fanucci, L.: Early malware detection system using behavioral analysis for improved cybersecurity in networks. SPIE Real Time Cybersecurity Processes, 9139, 913903 (2014).
5. Celik, T., Özkaramanlı, H., Demirel, H.: Malware detection without traditional signatures: A machine learning-based approach. In: 2007 15th European Signal Processing Conference, IEEE, pp. 1794–1798 (2007).
6. Rafiee, A., Dianat, R., Jamshidi, M., Tavakoli, R., Abbaspour, S.: Malware detection using wavelet analysis and entropy characteristics. In: IEEE 3rd International Conference on Computer Research and Development, vol. 3, pp. 262–265 (2011).
7. Vijayalakshmi, S.R., Muruganand, S.: Real-time malware detection in network traffic using anomaly detection techniques. In: IEEE 2nd International Conference on Communication and Electronics System (ICCES), pp. 167–171 (2017).
8. Gagliardi, A., Saponara, S.: AdViMaD: Advanced video malware detection system for real-time cybersecurity monitoring. Energies, 13(8), 2098 (2020b).
9. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Malware classification using deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105. MIT Press, Cambridge (2012).

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details