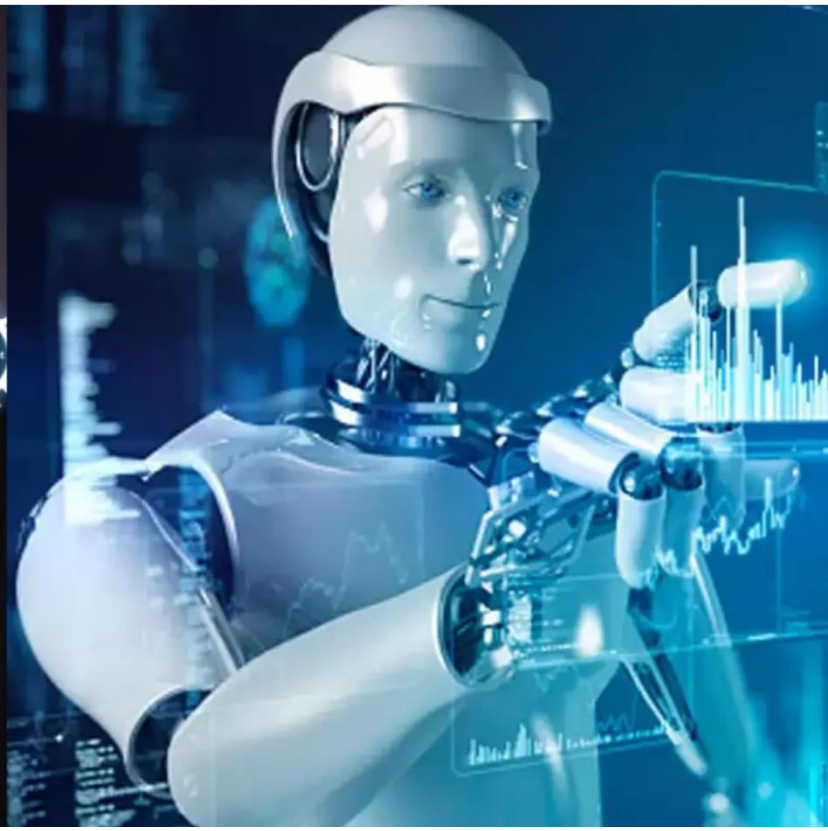
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

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: 9.935

Volume 15, Issue 5, May 2026

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

Impact Factor 9.935

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

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505142|

Hybrid Multi-Factor Cyber Attack Detection using Forensic Data Analytics

Himanshu Shukla[1], Shikha Goswami[2], Rajeev Singh[3], Govind Verma[4]

M.Tech. Student, Department of Information Technology, College of Technology, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, India[1]

Assistant Professor, Department of Information Technology, College of Technology, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, India[2,4]

Professor, Department of Computer Engineering, College of Technology, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, India[3]

ABSTRACT: Cyber security risks in modern organizations have increased significantly due to the widespread use of digital platforms, cloud computing, and interconnected systems. Traditional intrusion detection systems (IDS) based on rule-based and signature-based approaches are often ineffective against dynamic and unknown cyber-attacks. Although machine learning techniques improve detection capability, challenges such as class imbalance and high false positive rates remain. This study proposes a Hybrid Multi-Factor Cyber Attack Detection System using forensic data analytics. The system analyzes cyber forensic activity data, including login attempts, suspicious behaviour, and abnormal user activities. SMOTE (Synthetic Minority Oversampling Technique) is applied to address dataset imbalance before model training. Decision Tree and Random Forest classifiers are integrated with rule-based analysis to develop the hybrid detection model. Experimental results demonstrate that the proposed hybrid approach achieves improved suspicious activity detection with a recall, outperforming standalone machine learning models. The study highlights the effectiveness of combining forensic analytics, multi-factor analysis, and hybrid intelligence to strengthen cyber-attack detection in modern cyber security environments.

KEYWORDS: Cyber Security, Machine Learning, Hybrid Detection, Multi-Factor Authentication, Cyber Forensics, Random Forest, Decision Tree

I. INTRODUCTION

The rapid growth of cloud computing, digital communication, and internet-connected devices has transformed the modern technological environment. However, this digital transformation has also increased vulnerabilities to cyber-attacks, unauthorized access, brute-force attacks, phishing, insider threats, and data breaches. Organizations increasingly depend on intrusion detection systems (IDS) to monitor suspicious activities and protect sensitive information. Traditional intrusion detection systems primarily rely on signature-based and rule-based detection methods. Although these approaches are effective against known threats, they fail to detect zero-day attacks and evolving attack patterns. According to Ahmed et al. [1], traditional anomaly detection techniques often suffer from limited adaptability and high false positive rates. Machine learning (ML) has emerged as a powerful solution for intrusion detection because of its ability to learn patterns from large datasets and identify abnormal behavior automatically. Xin et al. [2] stated that machine learning-based IDS improves detection capability by analyzing complex traffic behavior and activity patterns. However, machine learning models face several practical limitations such as class imbalance, overfitting, lack of interpretability, and reduced detection capability for minority attack classes. Deep learning approaches have further improved intrusion detection performance by automating feature extraction and identifying sequential attack patterns. Yin et al. [3] proposed recurrent neural network-based IDS models capable of detecting sophisticated attacks. Despite improved accuracy, deep learning models require high computational resources and large-scale datasets, making real-time implementation difficult. One of the most critical problems in cyber security datasets is class imbalance, where normal user activities significantly outnumber malicious activities. This imbalance causes machine learning models to favor majority classes while ignoring attack instances. He et al. [4] introduced adaptive synthetic sampling methods to improve minority class learning. However, standalone balancing techniques are often insufficient in highly dynamic

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

Impact Factor 9.935

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

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505142|

cyber environments. To address these limitations, hybrid intrusion detection systems combining machine learning, rule-based analysis, and forensic analytics have gained attention. Hybrid systems improve attack detection capability while leveraging domain knowledge and behavioral analysis. This research proposes a Hybrid Multi-Factor Intrusion Detection System integrating rule-based suspicious activity analysis with machine learning classification models to improve cyber-attack detection.

II. LITERATURE WORK

Several researchers have proposed machine learning and hybrid approaches for intrusion detection systems. Ahmed et al. [1] discussed anomaly-based intrusion detection techniques and identified challenges such as high false positive rates and limited adaptability. Their study highlighted the need for intelligent detection mechanisms. Xin et al. [2] analyzed the role of machine learning in cyber security and concluded that ML-based IDS improves detection capability compared to traditional approaches. However, the study also reported limitations associated with imbalanced datasets. Yin et al. [3] proposed a deep learning intrusion detection framework using recurrent neural networks (RNN). Their approach achieved improved classification accuracy but required significant computational resources. He et al. [4] introduced the Adaptive Synthetic Sampling (ADASYN) technique to handle class imbalance problems in machine learning datasets. The approach improved minority class learning by generating synthetic samples. Sommer and Paxson [5] argued that machine learning models may not perform effectively in operational environments due to dynamic attack behavior and lack of contextual knowledge.

Thaseen and Ganesh [6] developed a hybrid intrusion detection framework combining feature selection and classification techniques. Their work demonstrated improved detection accuracy and reduced false positive rates. Aljawarneh et al. [7] proposed feature selection techniques for network intrusion detection systems and reported enhanced model efficiency and reduced computational overhead. Tama et al. [8] developed a two-stage hybrid intrusion detection model using ensemble learning techniques. Their study achieved better classification performance than standalone models. Ferrag et al. [9] reviewed deep learning applications in cyber security and emphasized challenges related to resource utilization, scalability, and real-time implementation. Sharma and Kumar [10] proposed a hybrid authentication attack detection model integrating machine learning with rule-based logic. Their research highlighted the effectiveness of multi-factor analysis in detecting suspicious login behavior. The literature indicates that hybrid approaches combining machine learning and rule-based analysis provide improved intrusion detection performance, especially in authentication-based attack scenarios.

Despite significant advancements in intrusion detection systems, existing machine learning models still struggle to effectively detect minority attack instances due to dataset imbalance and evolving attack behavior. Traditional systems often generate high false positives and fail to identify authentication-based attacks efficiently. This research is a hybrid multi-factor cyber-attack detection system using forensic data analytics. The study aims to analyze authentication-based cyber activities, evaluate machine learning models, and improve suspicious activity detection performance by combining rule-based and machine learning approaches.

III. METHODOLOGY

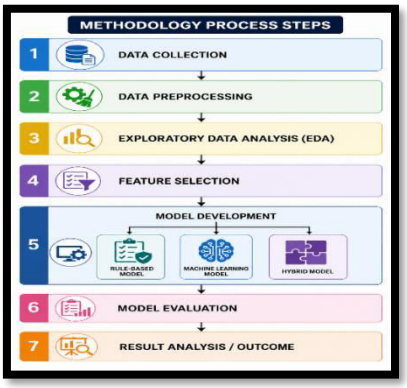


Figure 1 Hybrid Multi Factor Detection Methodology

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

Impact Factor 9.935

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

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505142|

3.1 Dataset Collection
The dataset used in this research consists of cyber forensic activity records containing authentication-related events such as login attempts, IP address behavior, failed login counts, session activity, and suspicious user patterns.

3.2 Data Preprocessing
The preprocessing stage includes:
a. Handling missing values
b. Removing duplicate records
c. Encoding categorical variables
d. Feature normalization
e. Data balancing for minority attack classes
SMOTE (Synthetic Minority Oversampling Technique) was applied to balance minority attack instances and improve suspicious activity detection capability.

3.3 Feature Selection
Features such as Login Attempts, Hour, Activity Type, and authentication behavior were selected for machine learning classification and suspicious activity analysis.

3.4 Machine Learning Models
1) The dataset was divided into training and testing sets using the train_test_split() method. Decision Tree and Random Forest classifiers were trained using balanced data generated through SMOTE. Performance evaluation was conducted using Accuracy, Precision, Recall, and F1-score metrics.
2) Decision Tree
Decision Tree classification was used because of its simplicity and interpretability in identifying suspicious activities.
3) Random Forest
Random Forest was implemented to improve classification stability and reduce overfitting by combining multiple decision trees.

3.5 Rule-Based Detection
A rule-based suspicious activity detection mechanism was developed using:
- Multiple failed login attempts
- Login attempts from unknown IP Addresses
- Abnormal access timing
- Repeated authentication failures
- Suspicious session behavior

3.6 Hybrid Model
The hybrid model combines machine learning classification, rule-based suspicious activity detection, and multi-factor authentication behavior analysis. The final prediction was generated using a logical OR operation between machine learning predictions and rule-based suspicious activity detection. If either approach classified an activity as suspicious, the hybrid system labeled the activity as an attack. This integration improved attack detection capability and enhanced recall for minority attack instances.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The performance of the hybrid model system was evaluated using accuracy, precision, and recall.
4.1 Decision Tree Performance
The Decision Tree model achieved high classification accuracy for normal activities but showed lower recall for attack instances due to class imbalance.
4.2 Random Forest Performance
Random Forest improved overall classification stability and reduced overfitting. However, minority attack detection remained challenging.
4.3 Hybrid Model Performance
The hybrid model demonstrated improved attack detection capability by combining rule-based analysis with machine learning classification.

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

Impact Factor 9.935

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

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505142|

4.4 Performance Metrics

The performance of the Hybrid Multi-Factor Cyber Attack Detection system was evaluated using standard classification metrics such as Accuracy, Precision and Recall. These metrics help in measuring the effectiveness of machine learning models in identifying both normal and malicious activities within cyber forensic datasets.

Accuracy

Accuracy represents the overall correctness of the model by measuring the ratio of correctly classified instances to the total number of instances.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

A high accuracy value indicates that the model performs well in overall classification. However, in imbalanced cyber security datasets, accuracy alone may not provide reliable performance evaluation because models may favour majority classes.

Precision

Precision measures the proportion of correctly predicted attack instances among all predicted attack instances.

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

Higher precision indicates that the model produces fewer false alarms and improves reliability in cyber-attack prediction.

Recall

Recall measures the capability of the model to correctly detect actual attack instances.

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

Recall is one of the most important metrics in intrusion detection systems because missing an attack can lead to serious security threats.

The hybrid model achieved better recall compared to standalone machine learning models. These performance metrics collectively provide a comprehensive analysis of the hybrid intrusion detection system and demonstrate its effectiveness in detecting cyber-attacks within imbalanced forensic datasets.

The hybrid multi-factor model combining machine learning and rule-based suspicious activity analysis achieved the best performance with 61% accuracy, 54% precision, and 94% recall. The results indicate that the hybrid approach significantly improves attack detection capability compared to standalone machine learning models. Although the hybrid model generated slightly higher false positives, the improved recall is highly beneficial in cybersecurity applications where detecting attacks is more important than maximizing overall accuracy. The experimental analysis demonstrates that combining forensic analytics, machine learning, and behavioural rule-based detection provides an effective solution for authentication-based cyber-attack detection in imbalanced datasets.

Table 1 : Performance Table of Model used

Model	Accuracy	Precision	Recall
Multi-Factor Attack Detection	51%	19%	57%
Decision Tree	82%	78%	85%
Random Forest	79%	75%	82%
Hybrid Model	61%	54%	94%

V. CONCLUSION

This research proposed a Hybrid Multi-Factor Cyber Attack Detection system using forensic data analytics, machine learning, and rule-based suspicious activity analysis. The study evaluated Rule-based, Decision Tree, Random Forest, and hybrid detection approaches on cyber forensic datasets. Experimental results demonstrated that standalone machine learning models achieved high overall accuracy but failed to effectively detect minority attack instances. The hybrid model significantly improved recall and suspicious activity detection capability by combining forensic behavioral analysis with machine learning classification. The research concludes that hybrid intrusion detection systems provide a

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

Impact Factor 9.935

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

Volume 15, Issue 5, May 2026

|DOI: 10.15680/IJIRSET.2026.1505142|

more practical and effective approach for cyber security environments where attack detection is more important than overall accuracy.

VI. FUTURE WORK

Future research can focus on improving the hybrid intrusion detection system through real-time deployment in practical cyber security environments. Deep learning techniques can be integrated to enhance attack detection accuracy and automate feature extraction. Explainable AI methods may also be used to improve the interpretability of intrusion detection decisions. The system can be extended for detecting advanced persistent threats (APT) and complex cyber-attacks. Additionally, the hybrid model can be implemented in cloud computing and IoT environments to provide scalable and secure cyber defense solutions. Further optimization can also help reduce false positives and improve system efficiency.

REFERENCES

1. M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," Journal of Network and Computer Applications, vol. 60, pp. 19–31, 2016. doi: https://doi.org/10.1016/j.jnca.2015.11.016.
2. I. Ahmad, A. W. Malik, and R. A. Khan, "Intrusion detection in IoT networks: Hybrid machine learning and rule-based approach," Sensors, vol. 23, no. 6, p. 2890, 2023. doi: https://doi.org/10.3390/s23062890.
3. I. Alharkan and M. Alharbi, "Hybrid IDS for cybersecurity using ensemble learning and anomaly detection," Journal of Big Data, vol. 10, no. 1, p. 45, 2023. doi: https://doi.org/10.1186/s40537-023-00712-9.
4. H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," in IEEE International Joint Conference on Neural Networks, 2008, pp. 1322–1328. doi: https://doi.org/10.1109/IJCNN.2008.4633969.
5. R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in IEEE Symposium on Security and Privacy, 2010, pp. 305–316. doi: https://doi.org/10.1109/SP.2010.25.
6. S. Thaseen and V. Ganesh, "Hybrid intrusion detection system using machine learning and rule-based approach," Procedia Computer Science, vol. 115, pp. 404–411, 2017. doi: https://doi.org/10.1016/j.procs.2017.09.092.
7. C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," IEEE Access, vol. 5, pp. 21954–21961, 2017. doi: https://doi.org/10.1109/ACCESS.2017.2762418.
8. S. Mittal and R. Sharma, "Hybrid machine learning-based intrusion detection system for imbalanced datasets," Computers & Security, vol. 92, p. 101744, 2020. doi: https://doi.org/10.1016/j.cose.2020.101744.
9. P. Sharma and S. Kumar, "Rule-based and machine learning hybrid approach for detection of authentication attacks," Journal of Cyber Security and Mobility, vol. 10, no. 2, pp. 123–145, 2021. doi: https://doi.org/10.13052/jcsm.2021.10.2.
10. S. Tiwari and R. Singh, "Improving detection of imbalanced cyber forensic datasets using hybrid models," Journal of Cloud Computing, vol. 13, no. 1, p. 120, 2024. doi: https://doi.org/10.1186/s13677-024-00685-x.
11. X. Zhang and Y. Chen, "Hybrid machine learning-based anomaly detection for cyber security applications," PLOS ONE, vol. 19, no. 3, p. e0308206, 2024. doi: https://doi.org/10.1371/journal.pone.0308206.
12. V. Kumar and S. Gupta, "Enhancing intrusion detection using hybrid models with feature selection and ensemble learning," Scientific Reports, 2025. doi: https://doi.org/10.1038/s41598-025-87028-1.

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