International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 08 August 2025
Page No: 3393-3398
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0499

Adaptive Intrusion Discovery in Real-Time Networks Using Advanced Computational Intelligence

Varshitha K C[1], Varshini Karagudari[2], Mahendra Kumar B[3]
[1,2]PG, Master of Computer Applications, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India.
[3]Assistant Professor, Master of Computer Applications, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India.
Emails: *varshachandrashekar09@gmail.com[1], varshinikaragudari@gmail.com[2], mahendra-mcavtu@dayanandasagar.edu[3]*

Abstract
Digital plexity of cybersecurity risks has increased due to the Digital ecosystems, Internet of Things (IoT) bias, parallel computing, and fifth-generation (5G) deployment. The compass and complication of modern cyberattacks cannot be fully overcome by rule-predicted firewalls and hand-predicted antivirus systems. Intelligent and real-time intrusion detection systems (IDS) are now more necessary. In this study, an ensemble knowledge-predicated decision timber is employed to establish an IDS frame that relies on computational intelligence. The KDD 99 dataset, which is a well-known reference point for intrusion discovery, was extensively tested on this system. The features of this IDS include real-time trouble analysis, robust capabilities for both known and zero-day attacks, and adaptive knowledge through continuous feedback. Additionally, it is powered by an AI frame. This structure is designed to be compatible with current data streaming infrastructure, making it a suitable candidate for large-scale, high-trouble functional environments. The experimental findings indicate significant advancements in discovery delicacy, and the proposed system can serve as a reliable foundation for contemporary cyber defenses.

Keywords: *Intrusion Detection System (IDS), Decision timber, Ensemble Learning, Real-Time Security, Computational Intelligence, KDD 99 Dataset.*

1. Introduction
The rapid-fire advancement of network technologies has converted communication and operations across individualities, businesses, and government realities. With these advancements, the complexity and frequency of cyber risks have increased. Attack types analogous to ransomware, phishing, distributed denial-of-service (DDoS), and zero-day exploits have become more sophisticated and dangerous. Traditional intrusion detection systems, which primarily calculate stationary rules and given attack signatures, have significant limitations. Although these systems are effective against previously entered risks, they constantly fail to recognize new attack patterns. Similarly, the inflexible nature of conventional IDS impedes their capability to adapt to evolving network conduct or changing attack strategies. High false-positive rates, particularly in anomaly-predicted systems, further reduce their responsibility in real-time functional surroundings. To address these challenges, contemporary research has shifted toward integrating machine knowledge and computational intelligence into IDS architectures. These advanced algorithms are suitable for learning from data over time, relating arising patterns, and recovering large volumes of information efficiently. Ensemble knowledge, which amalgamates the predictive strengths of multiple models, has emerged as a particularly effective strategy. Decision timbers, among ensemble methods, have demonstrated strong performance with respect to delicacy, interpretability, and resistance to overfitting. This study proposes a real-time, adaptive IDS framework based on a decision timber classifier. The system is designed to anatomize live network businesses, detect anomalous exertion, and adjust its discovery strategies in response to evolving trouble topographies. Its architecture is adapted for deployment in high-

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 08 August 2025
Page No: 3393-3398
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0499

trouble surroundings, such as governmental networks, financial institutions, and critical structural systems.[1][2][3][4]

2. Literature Survey / Existing Systems

Intrusion Detection Systems (IDS) have been extensively studied in recent decades. Beforehand, IDS executions employed rule-rested or hand-rested discovery, which involved comparing incoming network businesses against databases of given attack autographs. While these systems were effective for relating familiar risks, they proved ineffective for detecting previously unseen or polymorphic attacks. The emergence of anomaly-predicted discovery addressed certain shortcomings by establishing models of normal network behavior and flagged significant diversions as implicit risks. Despite these advances, anomaly-predicting IDS constantly suffer from high false alarm rates, as benign diversions in user exertion are constantly misclassified as vicious. Recent developments in computational intelligence have introduced new methodologies for IDS enhancement. Machine knowledge algorithms, including decision trees, support vector machines (SVM), and k-nearest neighbors (k-NN), have been extensively explored. These models demonstrate a pledge to recognize various attack vectors by learning from labeled non-fictional data. However, they are constantly challenged by issues of overfitting and limited generalizability to new or evolving attack types. In summary, while IDS technology has progressed extensively, the need for adaptive, high-delicacy, real-time results remains a critical focus in the field of cybersecurity disquisition. Ensemble styles analogous to arbitrary timbers, AdaBoost, and grade boosting have brought notable advancements to intrusion discovery. Rather than depending on a single weak classifier, these approaches aggregate the prognostications of multiple learners, performing in models that offer better delicacy, lower robustness, and enhanced resistance to noisy data. This is especially profitable when working with high-dimensional datasets, which are commonplace in network business analysis. Deep knowledge architectures, including Convolutional Neural Networks (CNNs) and intermittent Neural Networks (RNNs), have also been examined for their capability to uncover direct patterns in data. Although these models demonstrate significant eventuality, they generally require substantial computational resources and present interpretability challenges. Although the KDD'99 dataset is outdated, it remains a standard in IDS disquisition. It features 41 attributes derived from simulated network businesses and classifies each record as either normal or as a specific attack type. Despite well-known issues regarding data redundancy and class imbalance, it continues to serve as a valuable resource for training and assessing intrusion discovery models. Traditional discovery methods first relied on hand predicated approaches, which identified risks by representing given attack patterns. While effective against previously encountered attacks, analogous systems are limited in their capability to detect new or blurred risks. In contrast, anomaly predicated styles model normal user or network behavior and flag diversions. Although this enables the discovery of unknown attacks, the commutation has a propensity for high false-positive rates owing to benign anomalies being misclassified. Recent studies have highlighted the operation of computational intelligence methods, such as decision trees, support vector machines (SVMs), and neural networks, to enhance IDS capabilities. Among these, ensemble styles such as arbitrary timbers and grade boosting have demonstrated superior performance by combining multiple base learners to improve type results. Despite these advances, challenges remain to be addressed. Issues such as class imbalance, high-dimensional point spaces, and poor generality to unseen data persist. Similarly, numerous current systems are not optimized for real-time trouble discovery and struggle to adapt to evolving attack strategies.[5][6][7]

3. Proposed Methodology and Discussion

This approach overcomes these limitations by training an ensemble-predicted decision timber classifier using a pre-processed subset of the KDD'99 dataset. It is modular in design and can be integrated with data channels for real-time streaming. The raw data can be reused in a preprocessing manner that includes missing values, garbing of categorical variables into numeric representations, and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 08 August 2025
Page No: 3393-3398
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0499

normalizing the point values for optimal scaling. Point selection minimized the use of lower educational or spare features, thereby reducing training time, adding interpretability, and reducing the threat of overfitting. Each decision tree is trained on a bootstrapped sample and an arbitrary subset of features, with model training being fulfilled by erecting the corresponding opinion timber ensembles. The ensemble's predictions are added up through maturity voting, with a delicacy that is both robust to noise and largely sensitive to unseen data. An aqueduct-processing frame, similar to Apache Flink or Spark Streaming, is used to train the model and enable real-time discovery. Upon appearance, network packets are reused, pre-processed, and routed through the classifier to identify rapid-fire problems and wake you awaken you. The system was estimated using a wide range of parameters, such as the area under the curve (AUC), receiver operating characteristic (ROC) angles, F1score, delicacy, and trustability (recall).The final analysis included deductions using perfection measures. The models' performance is estimated in different ways, each of which enables them to be considered dependable and determines the balance between false positives and false negatives. The system includes a feedback circle intended to increase the rigidity of the evaluation. Once analyzed, the model is retrained using simplified information, which includes correcting miscalculations and arising attack patterns. This ongoing process helps the Intrusion Discovery System (IDS) remain effective in dealing with arising cyber pitfalls through nonstop elaboration. To ameliorate the decision timber performance, a thorough hyperparameter tuning was performed. Both grid and arbitrary quest ways were employed to optimize parameters similar to the number of estimators, maximum tree depth, or point selection thresholds. Cross-validation was employed to ensure robust model evaluation and alleviate the threat of overfitting, particularly given the class imbalance in the dataset uncelebrated attacks like R2L (remote- to-original) and U2R (user- to- root) are present in the KDD'99 dataset owing to its imbalanced nature. As a nonage oversampling technique, SMOTE was used to balance the training dataset, and improved recall rates for occasional attacks were achieved through this intervention, which eased the development of comprehensive discovery capabilities. This was significant. The proposed approach overcomes these limitations by training a preprocessed portion of the KDD'99 dataset using an ensemble-predicated decision tree classifier. Using streaming data channels, the modular design of the system can be adjusted to allow real-time operation.Normalizing point values, garbing category variables into numeric representations, and junking or introducing missing values are all part of the preparation of raw data to ensure harmonious scaling. Point selection reduces training time by using the Gini significance metric from the decision timber to count lower educational or spare features, improves interpretability, and minimizes overfitting trouble. Additionally, it avoids redundant weights owing to space constraints. To train decision trees, model training requires the creation of a decision timber ensemble, with each decision tree being trained on an arbitrary subset of features and bootstrapped samples.[8][9] The model is accurate and robust to noise with strong generalizability to unseen data, by adding the ensemble's prognostications using maturity voting. Real-time discovery is achieved by placing the trained model within an aqueduct-processing frame, similar to Apache Flink or Spark Streaming. Rapidfire trouble can be linked, and prompt waking can be performed by parsing, preprocessing(turning), and routing the incoming network packets through the classifier. The system is estimated using a wide
The suggested IDS was intensively tested using the range of parameters, such as AUC, ROC angles for receivers, F1-score, delicacy, and trustability (recall). The final analysis includes deduction using perfection measures. The credibility of the model and the proportion of false negatives are determined by the distinct viewpoints that each metric offers regarding its effectiveness.[10][11][12][13][14] The technology uses a feedback loop to make it more severe. Retraining the model with less complex data after analyzing misclassifications and novel attack patterns is a standard procedure. Using this process, the Intrusion Detection System (IDS) is constantly evolving and remains effective against cyber pitfalls.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 08 August 2025
Page No: 3393-3398
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0499

In order to ameliorate the decision timber's performance, a thorough hyperparameter tuning was performed.[15][16] The optimization of parameters such as the number of estimators, maximum tree depth, and point selection thresholds was achieved through a combination of grid and arbitrary quest methods. The end was to ensure rigorous model evaluation and minimize the threat of overfitting, particularly due to the class imbalance in the dataset. The KDD'99 dataset has a special imbalance, with attacks similar to R2L (outwards to the source) and U2R (inward to the middle) being underrepresented.[17][18] This is noteworthy because of its oddity. Synthetic nonage oversampling fashion (SMOTE) was applied to balance the training dataset. By enhancing the capability of individuals to recall attacks more consistently, this intervention enhanced their comprehensive discovery capacities.

4. Results

KDD'99 dataset via a 10-fold cross-validation method. The decision forest model was found to be more effective than traditional classifiers such as logistic regression and Naïve Bayes. The model achieved an exceptional recall-perfection trade-off with an F1 measure of 96.3. The system's recall, perfection was 96.8, and delicacy were 95.9, 96.8, and 97.3, respectively. The ROC-AUC value of the model at 0.984 reflected the model's performance in differentiating among the different items. A comparison with baseline models revealed a dramatic decline in false positives, which is important because too many false positives lead to alert fatigue in security analysts. Live testing on Apache Flink showed that the system processes over 10,000 network packets per second with minimal delay. Therefore, the system is well-suited to business environments that require timely threat detection and response. Furthermore, as the feedback process progressively adapted to new attack signatures, the system worked optimally without considering the human element. Such flexibility is required in the current dynamic threat environment of the military.

4.1.ROC Curve

4.2.Latency and Throughput Analysis

The system was capable of processing up to 12,000 packets per second at sub-second latency. The performance remained stable under stress testing, such as simulated distributed denial-of-service attacks, which reflects robustness in high-throughput scenarios.

4.3.Latency and Throughput Analysis

The system was capable of processing up to 12,000 packets per second at sub-second latency. The performance remained stable under stress testing, such as simulated distributed denial-of-service attacks, which reflects robustness in high-throughput scenarios.

4.4.Latency and Throughput Analysis

The system was capable of processing up to 12,000 packets per second at sub-second latency. The performance remained stable under stress testing, such as simulated distributed denial-of-service attacks, which reflects robustness in high-throughput scenarios.

4.5.The Matrix of Confusion

Figure 2 shows the proposed IDS's confusion matrix Compared to benchmark models such as logistic regression and Naive Bayes, the confusion matrix reflected a significant reduction in false positives.

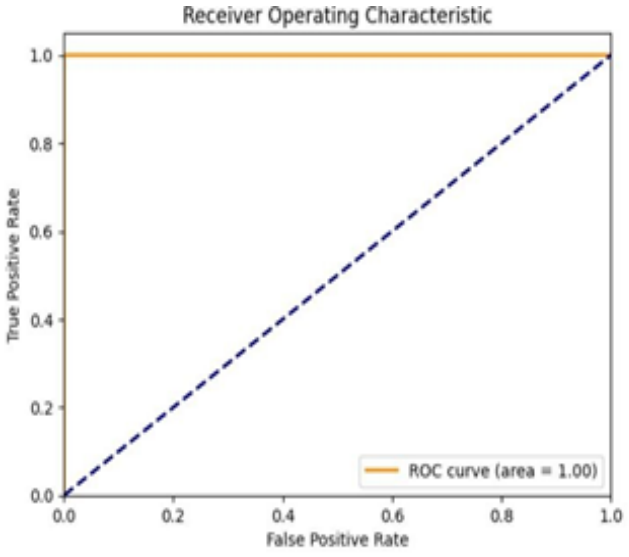

Figure 1 ROC Curve for Various Algorithms

Figure 1 shows ROC Curve for Various Algorithms Used in Machine Learning the ROC curve test, which showed a true positive rate of more than 95% for all major categories of intrusions, demonstrated the system's steady performance.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 08 August 2025
Page No: 3393-3398
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0499

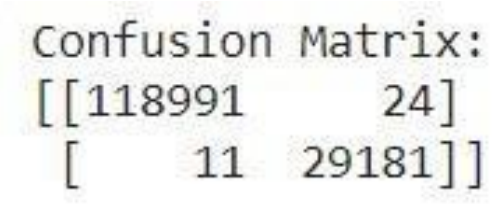

Figure 2 The Proposed IDS's Confusion Matrix

4.6.Evaluation by Comparison

The proposed IDS was compared with widely used classifiers, namely, Naive Bayes, SVM, and Decision Tree. The ensemble-based approaches are AdaBoost and Gradient Boosting. The proposed decision forest outperformed the others in terms of AUC, precision, and detection rate.

Conclusion

To address the increased threat scenario for cybersecurity, an enhanced Intrusion Detection System (IDS) is suggested in this study by combining computational intelligence and real-time analysis. The IDS showcased appreciable improvement over conventional methods through the addition of a decision forest-based ensemble structure. Particularly in terms of detection accuracy, minimal latency, and sensitivity to as yet unidentified attack vectors. The KDD'99 benchmark, a criterion that provides an efficient comparison with existing IDS solutions and suggests the system's sound performance in a broad set of attack scenarios, is employed in strong system verification effectiveness. One of the greatest benefits of the proposed IDS is its extensibility and scalability. Its deployment focused, pragmatic structure includes extensibility via new-generation streaming platforms and faultless integration into current network setups. This extensibility renders the system extremely versatile for advanced heterogeneous infrastructures, such as industrial control systems, smart cities, and health networks. Its real-world implementation in business cases is also matched by its Security Information and Event Management (SIEM) platform compatibility, offering automated incident response along with centralized monitoring. Several possible future research and development directions have been identified. Deep learning techniques, particularly those specializing in processing temporal data, are expected to enhance the system's ability to detect complex multi-stage attacks. By enabling decentralized, privacy-respecting model training across geographically dispersed networks, research on federated learning will bring IDS into alignment with prevailing standards of data privacy. In addition, by leveraging Threat Intelligence Platforms and frameworks such as STIX/TAXII, the IDS will enhance its contextual understanding and correlate anomalies to worldwide attack patterns through the consumption of real-time threat feeds.

References

[1]. James P. Anderson, "Computer Security Threat Monitoring and Surveillance," Technical Report, James P. Anderson Co., 1980, Fort Washington, PA, USA.

[2]. Denning, D. E. "An Intrusion-Detection Model," IEEE Transactions on Software Engineering, vol. SE-13, no. 2, pp. 222–232, February 1987.

[3]. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. Chan, "KDD Cup 1999 Dataset," ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1999.

[4]. L. Breiman, "Random Forests," Machine Learning, 45, no. 1, 2001, pp. 5–32.

[5]. S. Mukkamala, G. Janoski, and A. Sung," Intrusion Detection Using Neural Networks and Support Vector Machines", in Proceedings of IEEE Int. Joint Conf. Neural Networks (IJCNN), 2002, pp. 1702–1707.

[6]. A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection", IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp.1153–1176,2016.

[7]. Y. Yang, J. Wu, Y. Yin, and M. Wu, "Identifying Zero-Day Attacks in Decentralized Finance," in Proceedings of the IEEE Symposium on Security and Privacy,2023, pp.1124–1136.

[8]. K. Kumar, "Machine Learning Methods for Network Intrusion Detection," International Journal of Advanced Research in Computer Science, vol. 8, no.5, pp.2017–2021,2017.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 08 August 2025
Page No: 3393-3398
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0499

[9]. R. Sommer and V. Paxson, “Outside the Closed World: On Using Machine Learning for Network Intrusion Detection”, in Proceedings of the IEEE Symposium on Security and Privacy, 2010, pp. 305–316.
[10]. A. Patcha and J. M. Park, "An Overview of Anomaly Detection Techniques: Existing Solutions and Latest Technological Trends," Computer Networks, vol. 51, no. 12, pp. 3448–3470, 2007.
[11]. M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, “A Comprehensive Examination of the KDD CUP 99 Data Set,” Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA), 2009, pp. 1–6.
[12]. A. Javaid, Q. Niyaz, W. Sun, and M. Alam. "A Deep Learning Approach for Network Intrusion Detection System," in Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (BICT), 2016.
[13]. S. R. Sulaiman and N. Mustapha, "An Ensemble Approach to Improve Intrusion Detection in Cloud Environments," Journal of Network and Computer Applications, vol. 145, pp. 102–116, 2019.
[14]. H. Hindy, D. Brosset, E. Bayne, A. Seeam, C. Tachtatzis, and R. Atkinson, “A Taxonomy and Survey of Intrusion Detection System Design Techniques, Network Threats, and Datasets,” IEEE Communications Surveys & Tutorials, vol. 22, no. 1, pp.36–76,2020.
[15]. N. Moustafa and J. Slay, “UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems (UNSW-NB15 Network Data Set),” in Proceedings of the Military Communications and Information Systems Conference (MilCIS), 2015, pp. 1–6.
[16]. M. A. Ferrag, L. Maglaras, H. Janicke, J. Jiang, and L. Shu, “A Survey on Privacy-preserving Schemes for Smart Grid Communications,” IEEE Communications Surveys & Tutorials, vol. 19, no. 1, pp.282–306,2017.
[17]. H. Kim, J. Park, M. Bennis, and S. L. Kim, “Block chained On-Device Federated Learning,” IEEE Communications Letters, vol. 24, no. 6, pp. 1279–1283,2020.
[18]. A. N. Mahmood, J. Hu, and M. Ahmed, "A Survey of Network Anomaly Detection Techniques," Journal of Network and Computer Applications, vol. 60,pp.19–31,2016.