

Optimizing firewall timing for brute force mitigation with random forests

Ahmad Turmudi Zy¹, Isariato¹, Anggi Muhammad Rifa'i¹, Abdul Ghofir²,
Muhammad Najamuddin Dwi Miharja¹, Ananto Tri Sasongko¹

¹Department of Informatics Engineering, Faculty of Engineering, Universitas Pelita Bangsa, Bekasi, Indonesia

²Department of Informatics Engineering, Faculty of Engineering, President University, Bekasi, Indonesia

Article Info

Article history:

Received Aug 23, 2024

Revised Feb 25, 2025

Accepted Mar 15, 2025

Keywords:

Brute force attack mitigation

Cybersecurity frameworks

Firewall optimization

Pattern recognition

Random forest

ABSTRACT

Mitigating brute force attacks remains a critical challenge in cybersecurity, requiring intelligent and adaptive solutions. This research introduces an approach to optimizing firewall deployment timing for enhanced brute force mitigation using pattern recognition techniques with the random forest algorithm. Leveraging the UNSW-NB15 dataset, comprehensive preprocessing and exploratory data analysis (EDA) were performed to ensure the dataset's suitability for machine learning applications. The study utilized a structured workflow, splitting the dataset into training and testing subsets to rigorously evaluate the model's performance. The proposed random forest model achieved a high accuracy of 98.87%, supported by precision, recall, and F1-scores that confirm its effectiveness in distinguishing normal and attack traffic. The confusion matrix further validated the model's robustness, highlighting its potential in improving the efficiency of firewall deployment. These findings demonstrate the critical role of advanced machine learning techniques in enhancing cybersecurity defenses, particularly in mitigating brute force attacks through optimized, data-driven strategies.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Ahmad Turmudi Zy

Department of Informatics Engineering, Faculty of Engineering, Universitas Pelita Bangsa

Bekasi, Indonesia

Email: turmudi@pelitabangsa.ac.id

1. INTRODUCTION

In the modern digital era, cybersecurity is a pressing concern for organizations due to the growing frequency and complexity of cyberattacks [1]. Organizations must remain ahead of developing risks because cybercriminals are always coming up with new ways to get around security safeguards. Brute force assaults have grown particularly prevalent among the several types of cyberattacks [2]. In order to obtain unauthorized access to systems, attackers in these attacks methodically try a large number of password or encryption key combinations. Brute force attacks are a significant danger due to their simplicity and the potential for serious outcomes [3].

Brute force attacks can have serious consequences. Sensitive information may be exposed as a result of data breaches brought on by attackers once they get access [4]. This can seriously harm the targeted organization's credibility and reputation in addition to causing financial loss through fraud or theft [5]. Furthermore, hacked systems frequently have their integrity compromised, which increases their susceptibility to other vulnerabilities and possible disruptions [6]. Simple password restrictions and account lockout policies are two examples of typical strategies for thwarting brute force attacks; however, they

frequently fall short. These approaches can't keep up with the continuously changing nature of cyber threats and are usually reactive rather than proactive [7].

Given these challenges, there is an urgent need for more advanced and dynamic solutions to enhance the effectiveness of brute force mitigation [8]. Organizations must adopt strategies that are not only capable of detecting and responding to brute force attacks in real time but also capable of adapting to new and emerging attack patterns [9]. This calls for the integration of cutting-edge technologies and innovative approaches to cybersecurity, ensuring that defenses remain robust and effective in the face of an ever-changing threat landscape [10].

To tackle these cybersecurity challenges, advanced machine learning techniques have proven to be a promising solution. Machine learning enables the analysis of large datasets, allowing for the identification of patterns that could indicate malicious activity [11]. Among the various machine learning algorithms, the random forest algorithm has gained significant attention in the realm of cybersecurity. This algorithm is renowned for its robustness and accuracy, making it particularly well-suited for detecting and mitigating cyber threats [12]. Its ability to handle large datasets with numerous features allows it to analyze complex data structures and uncover subtle patterns that may signify a brute force attack [13].

The random forest algorithm operates by generating multiple decision trees during the training process and then combining their outputs to form a final prediction [14]. Each decision tree is constructed from a distinct subset of the training data, with the final classification determined by a majority vote among the trees. This ensemble technique is highly effective in reducing overfitting, a common issue where a model performs well on training data but struggles with new, unseen data [15]. By averaging the predictions from several trees, random forests improve the model's ability to generalize, making it more reliable in real-world situations where new and unexpected patterns may emerge [16].

In addition to the inherent strengths of the random forest algorithm, the integration of pattern recognition techniques further enhances its effectiveness in cybersecurity applications [17]. Pattern recognition involves identifying regularities and anomalies in data, which is crucial for detecting complex attack patterns that traditional methods might overlook. For instance, brute force attacks often exhibit specific behavioral patterns, such as repeated login attempts within a short period. By leveraging pattern recognition, machine learning models can be trained to recognize these patterns and differentiate between normal user behavior and potential attacks [18]. This combination of random forests and pattern recognition provides a powerful toolset for proactively identifying and mitigating brute force attacks, thereby bolstering the overall security posture of an organization [19].

The integration of random forests and pattern recognition techniques into cybersecurity strategies presents a formidable approach to enhancing brute force attack mitigation [20]. Random forests, with their ensemble learning capabilities, can effectively manage and analyze vast amounts of network traffic data, identifying patterns that may indicate an ongoing brute force attack [21]. By combining these capabilities with pattern recognition techniques, it becomes possible to detect even the most subtle and sophisticated attack patterns. This synergy allows for more accurate and timely identification of brute force attacks, which is critical for deploying countermeasures promptly and effectively [22].

By leveraging these advanced technologies, cybersecurity strategies can move beyond traditional reactive approaches. Instead of responding to attacks after they have occurred, organizations can implement proactive measures that anticipate and prevent potential threats [23]. The ability to analyze and identify attack patterns in real-time significantly enhances the speed and accuracy of the response, reducing the window of opportunity for attackers [24]. This research specifically focuses on optimizing the timing of firewall deployments using random forests and pattern recognition. By determining the most effective times to activate firewalls, it is possible to preemptively block malicious traffic, thereby minimizing the risk of successful brute force attacks [25].

The UNSW-NB15 dataset from [26], which contains comprehensive network traffic data, serves as the foundation for training and evaluating the random forest model. This dataset is particularly valuable because it includes a wide variety of normal and malicious traffic patterns, providing a robust training ground for the model. Training the random forest model on this dataset enables it to effectively differentiate between benign and malicious traffic with a high degree of accuracy. The successful application of this technology can significantly improve defense mechanisms against brute force attacks, providing a proactive rather than reactive approach to cybersecurity [27].

The purpose of this research is to integrate findings into practical cybersecurity frameworks, enhancing the efficiency and effectiveness of network defense mechanisms using the random forest algorithm. Brute force attacks, as persistent and evolving threats, demand proactive solutions beyond traditional reactive methods [28]. This study demonstrates how advanced machine learning techniques, particularly random forest, can address critical cybersecurity challenges by bridging the gap between theoretical models and practical applications [29].

This research contributes by introducing a novel methodology for optimizing firewall deployment timing through pattern recognition and random forests, enabling proactive detection and mitigation of brute force attacks. It also validates the integration of machine learning with dynamic firewall strategies to improve real-time response capabilities against cyber threats. Additionally, it emphasizes the significance of leveraging high-quality datasets, such as UNSW-NB15, to accurately identify attack patterns while minimizing false positives and negatives.

Driven by the increasing sophistication of brute force attacks, this study offers a scalable and adaptable solution that integrates seamlessly into existing security frameworks. Optimizing firewall deployment timing ensures efficient resource utilization, avoids disruptions to legitimate traffic, and preemptively blocks malicious activities. This approach strengthens immediate defenses and provides a foundation for future advancements in intelligent and automated cybersecurity solutions. By addressing these objectives, this research contributes to the development of adaptive and resilient cybersecurity technologies, equipping organizations with effective tools to counter emerging threats and protect critical network infrastructures.

2. RELATED WORKS

The evolution of research in optimizing firewall deployment timing for enhanced brute force mitigation has gained significant traction in recent years. Traditional approaches to mitigating brute force attacks primarily relied on static rules and signature-based detection methods, which often fall short in adapting to the dynamic nature of cyber threats. Recent studies have highlighted the limitations of these conventional methods and the need for more adaptive and intelligent solutions. For instance, dynamic rule adjustment and real-time traffic analysis have been proposed as more effective strategies for detecting and responding to brute force attacks [30], [31]. These advancements underscore the necessity of optimizing firewall deployment timing to enhance the responsiveness and efficiency of mitigation efforts.

The integration of machine learning into cybersecurity frameworks has revolutionized the approach to threat detection and mitigation [18]. Machine learning algorithms, particularly supervised learning models, have been extensively used to analyze network traffic and identify malicious patterns. Studies have demonstrated that machine learning can significantly improve the accuracy and speed of detecting various types of cyberattacks, including brute force attacks [32]. The ability of machine learning models to learn from historical data and recognize complex attack patterns makes them invaluable in the realm of cybersecurity. Research efforts have focused on developing models that can adapt to evolving threats, thereby enhancing the overall robustness of network defense mechanisms.

Among various machine learning algorithms, the random forest algorithm stands out for its robustness and efficiency in handling large datasets with numerous features. As an ensemble learning method, random forest builds multiple decision trees during the training phase and aggregates their outputs to enhance classification accuracy and mitigate overfitting [33]. This method has shown particular effectiveness in cybersecurity, where generalizing from diverse, high-dimensional data is essential. Studies on random forest have highlighted its superior performance in detecting and mitigating brute force attacks, making it a favored option in this field [12], [34].

Several studies have explored similar themes in optimizing firewall deployment and enhancing brute force attack mitigation using advanced techniques. For instance, research on adaptive firewall policies that leverage real-time data analysis and machine learning for dynamic rule adjustments has shown promising results [35]. Additionally, studies employing various ensemble learning methods, including random forest, have highlighted their effectiveness in improving detection accuracy and response times [36]. These works collectively emphasize the importance of continuous innovation and the integration of intelligent algorithms in developing more resilient and proactive cybersecurity strategies. The findings from these studies provide a solid foundation for further research and development in optimizing firewall deployment timing for enhanced brute force mitigation.

Based on the reviewed literature, it is clear that optimizing firewall deployment timing plays a crucial role in improving brute force mitigation. Traditional methods relying on static rules and signature-based detection is insufficient in addressing the dynamic nature of cyber threats. Machine learning, particularly the random forest algorithm, has demonstrated its effectiveness in analyzing high-dimensional data and identifying complex attack patterns, offering significant improvements in detection accuracy and adaptability. This study aims to optimize firewall deployment timing for enhanced brute force mitigation using pattern recognition with the random forest algorithm. By leveraging its ability to process diverse data and generalize effectively, this research seeks to provide a more accurate, efficient, and adaptive solution to counter evolving cyber threats.

3. **METHOD**

The methodology for optimizing firewall deployment timing to improve brute force attack mitigation involves five crucial stages, as depicted in Figure 1. First, relevant data from the UNSW-NB15 dataset is systematically collected to ensure a comprehensive representation of both normal and attack traffic. This data undergoes preprocessing to ensure quality and consistency, addressing missing values and encoding categorical features.

Next, appropriate models, particularly the random forest algorithm, are selected for training and testing to distinguish between normal and attack traffic. An ablation study follows to assess the impact of different components on the model's performance, especially the contribution of pattern recognition techniques to detection accuracy. Finally, the results are thoroughly analyzed to evaluate the proposed method's efficiency and effectiveness in optimizing firewall deployment timing to mitigate brute force attacks. Figure 1 visually represents this methodology, illustrating the progression through these five stages.

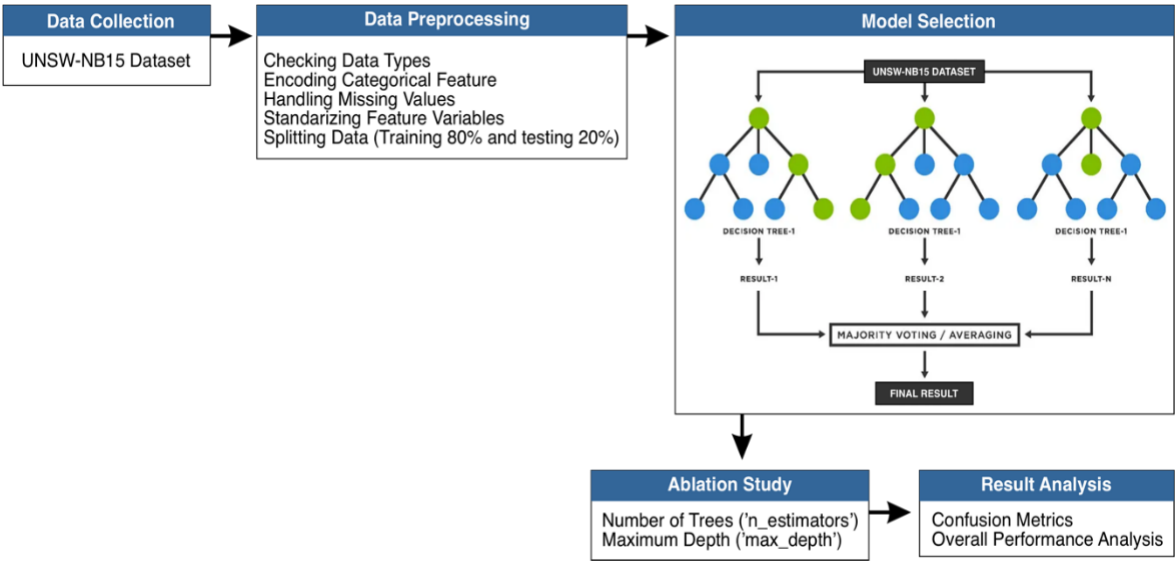


Figure 1. Methodology enhanced brute force mitigation purpose

3.1. **Data collection**

The initial phase involves systematically collecting relevant data from the UNSW-NB15 dataset, a comprehensive resource designed for cybersecurity research, particularly in intrusion detection and prevention systems [26]. The dataset encompasses a diverse set of features that characterize network traffic, including connection duration, the protocol employed, the state of the connection, the number of packets transmitted and received, and the total bytes exchanged show in Table 1. Additionally, it contains computed features such as the source-to-destination time-to-live (TTL) value, the bit rate between source and destination, inter-packet arrival times, and jitter.

The dataset consists of both numeric and categorical features. Numeric features include connection duration, packet counts, byte counts, and data transfer rates, while categorical features encompass protocol type, service type, and connection state. The target variable, label, indicates whether the traffic is normal (0) or an attack (1), with an additional attack_cat column specifying the attack category.

To prepare the dataset for machine learning, categorical features are converted into numeric format using techniques such as label encoding. Any missing values are addressed by either filling them in or removing the affected rows. The dataset is then divided into features (X) and the target variable (y), where X includes all columns except label and attack_cat, and y represents the label. The features are standardized to bring them to a comparable scale, with a mean of zero and a standard deviation of one. This preprocessing step ensures the data is well-prepared for training machine learning models, facilitating precise and efficient detection and mitigation of brute force attacks.

3.2. **Data preprocessing**

Data preprocessing is a critical step in the machine learning pipeline to ensure the quality, consistency, and readiness of the dataset for training the model. It plays a vital role in improving the

performance and accuracy of machine learning algorithms by eliminating noise and handling missing or inconsistent values. This step involves several sub-processes, checking data types, encoding categorical features, handling missing values, standardizing feature variables, and splitting data, which are essential to prepare the dataset effectively before model training.

- Checking data types: the initial phase of preprocessing involves inspecting the data types of each column in the dataset. This step is crucial to identify which features are categorical and which are numerical. Categorical features need to be encoded into a numeric format, while numerical features need to be standardized to ensure uniform scaling.
- Encoding categorical features: in the dataset, categorical features such as 'proto' (protocol type), 'service' (network service on the destination), and 'state' (state and condition of the protocol) were identified. These features cannot be directly used by most machine learning algorithms that require numeric input. Therefore, they were encoded using the LabelEncoder. Label encoding converts categorical values into a numeric format where each unique category is assigned a distinct integer value. This transformation enables the random forest model to process and learn from these features effectively.
- Handling missing values: the dataset was then checked for missing values, which can negatively impact the performance of the machine learning model if not handled properly. Missing values can arise due to various reasons, such as incomplete data collection or data corruption. In this research, missing values were addressed by either dropping rows with missing data or filling them using appropriate methods like forward fill or mean imputation. In this instance, rows with missing values were dropped to maintain the dataset's integrity and ensure that the model is trained on complete data.
- Standardizing feature variables: the last phase of data preprocessing involves standardizing the feature variables, a critical step to ensure that each feature equally influences the model's learning process. Standardization adjusts the data so that it has a mean of zero and a standard deviation of one. This process is particularly vital for algorithms that rely on distance measures, like random forests, as it prevents features with larger scales from overwhelming the learning process. The StandardScaler from scikit-learn was utilized to standardize the numerical features, thereby improving the model's performance and accelerating convergence.
- Splitting data: post-preprocessing, the dataset is split into features (X) and target variable (y), with 80% of the data allocated for training and 20% for testing. The features are all columns except 'label' and any non-numeric columns like 'attack_cat', while 'label' represents the target variable indicating whether the traffic is normal or an attack. This splitting prepares the data for the subsequent training and testing phases.

Table 1. Head sample dataset

Feature	sample_1	sample_2	sample_3	sample_4	sample_5
id	0	1	2	3	4
dur	1	2	3	4	5
proto	tcp	tcp	tcp	tcp	tcp
Service	-	-	-	ftp	-
state	FIN	FIN	FIN	FIN	FIN
spkts	6	14	8	12	10
dpkts	4	38	16	12	6
sbytes	258	734	364	628	534
dbytes	172	42014	13186	770	268
rate	74.08749	78.473372	14.170161	13.677108	33.373826
Sport_ltm	1	1	1	1	1
dst_src_ltm	1	2	3	3	40
ftp_login	0	0	0	1	0
ftp_cmd	0	0	0	1	0
http_mthd	0	0	0	0	0
src_ltm	1	1	2	2	2
srv_dst	1	6	6	1	39
ips_ports	0	0	0	0	0
attack_cat	normal	normal	normal	normal	normal
label	0	0	0	0	0

3.3. Model selection

During the model selection phase, the random forest algorithm was selected for its robustness, accuracy, and capability to manage large datasets with numerous features. First introduced by Breiman in 2001, random forests are an ensemble learning technique that builds multiple decision trees and aggregates their outputs to enhance classification accuracy and minimize overfitting. This algorithm is particularly well-suited for differentiating between normal and attack traffic in network data.

Once the data is split, the random forest model is initialized and trained using the training set. Key parameters, such as the number of trees in the forest ($n_estimators$), are set to 100, meaning the model will construct 100 decision trees. The algorithm works by bootstrap sampling, where random samples of the training data are drawn for each tree. Each tree is built using a subset of features at each node, which increases diversity and reduces overfitting. Node splitting is based on criteria like Gini impurity, calculated as in (1) [15].

$$Gini(D) = 1 - \sum_{i=1}^n P_i^2 \quad (1)$$

Where P_i is the probability of class i in the dataset D (1).

After all trees are constructed, they make independent predictions. The final prediction is determined through majority voting: each tree casts a vote for a class, and the class that receives the most votes is selected. This ensemble approach enhances the model's accuracy and stability, reducing variance and mitigating overfitting. The training phase involves feeding the training data into the model to learn patterns distinguishing normal from attack traffic, ensuring robust and reliable classification performance.

3.4. Ablation study

An ablation study was carried out to evaluate how various components influenced the model's performance. This process involved systematically modifying or removing specific components of the model to assess their influence on overall accuracy and robustness. Key aspects such as the number of trees ($n_estimators$), the maximum tree depth (max_depth), and preprocessing techniques like feature scaling and encoding were thoroughly examined. The findings from this study were instrumental in fine-tuning the model, enhancing its capability to detect brute force attacks effectively. This approach ensured that the model was optimized for peak performance in practical applications.

3.5. Result analysis

The results were analyzed comprehensively to evaluate the efficiency and effectiveness of the proposed approach in optimizing firewall deployment timing for enhanced brute force mitigation. Key performance metrics such as accuracy, precision, recall, and F1-score were calculated using the following formulas [11]. The TP represents true positives, TN true negatives, FP false positives, and FN false negatives.

- Accuracy measures the overall correctness of the model by evaluating the proportion of total correct predictions out of all predictions made:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

- Precision assesses the model's ability to identify only the relevant instances of attacks by calculating the ratio of correctly predicted attack instances to the total predicted attack instances:

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

- Recall assesses the model's ability to identify all actual attack instances by determining the ratio of correctly predicted attack instances to the total number of actual attack instances.

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

- F1-score offers a balanced measure between precision and recall, making it particularly valuable in cases of imbalanced class distribution. It is calculated as the harmonic mean of precision and recall.

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (5)$$

4. RESULTS AND DISCUSSION

The results of this study demonstrate the effectiveness of using the random forest algorithm for optimizing firewall deployment timing to mitigate brute force attacks. The model achieved high accuracy in distinguishing between normal and attack traffic, as evidenced by key performance metrics. The calculated accuracy was 98.87%, indicating a high overall correctness in the model's predictions. Precision was found to be 98.99%, reflecting the model's ability to correctly identify attack instances with

minimal false positives. The recall rate stood at 99.77%, showcasing the model's capability to capture almost all actual attack instances, and the F1-score was 99.38%, indicating a robust balance between precision and recall.

The confusion matrix further elucidated the model's performance. It showed in Table 2 that out of 35,069 total instances, 10,827 normal instances were correctly classified, while 342 were misclassified as attacks. Conversely, 23,845 attack instances were correctly identified, with only 55 being wrongly classified as normal traffic. This high true positive rate, coupled with minimal false negatives and false positives, underscores the model's reliability in real-world scenarios.

Table 2. Confusion matrix

		Predicted		Total
		0	1	
Actual	0	10827	342	11169
	1	55	23845	23900
Total		10882	24187	35069

An ablation study further dissected the contributions of different components within the model. By systematically removing or modifying parts of the model, it was observed that certain features significantly enhanced the model's performance, thereby fine-tuning its effectiveness in detecting brute force attacks. The insights gained from the ablation study helped in refining the model, ensuring that the most critical components were optimized for better performance.

The confusion matrix provided additional insights into the model's ability to classify normal and attack traffic accurately. It revealed a high true positive rate with minimal false negatives and false positives, underscoring the model's reliability in real-world scenarios. This comprehensive analysis confirms the proposed approach's efficacy in enhancing cybersecurity measures through optimized firewall deployment timing, ultimately contributing to more robust network defense mechanisms against evolving brute force attacks. This research highlights the potential of machine learning, particularly the random forest algorithm, in addressing critical cybersecurity challenges. Continuous innovation and rigorous evaluation, as demonstrated in this study, are essential for developing effective solutions to counteract the ever-evolving landscape of cyber threats.

5. CONCLUSION

This study introduces a robust approach to optimizing firewall deployment timing for enhanced brute force attack mitigation by leveraging pattern recognition with the random forest algorithm. The experimental results demonstrate the model's exceptional performance, achieving an accuracy of 98.87%, precision of 98.99%, recall of 99.77%, and an F1-score of 99.38%. These metrics highlight the model's ability to accurately identify attack traffic while maintaining a low rate of false positives and effectively detecting nearly all actual attack instances. The ablation study further provided critical insights into the contributions of specific model components, enabling fine-tuning to enhance the algorithm's overall effectiveness. The findings confirm the efficacy of integrating machine learning techniques into cybersecurity frameworks to strengthen network defense mechanisms. By optimizing firewall deployment timing, this research contributes to developing proactive and adaptive strategies for mitigating brute force attacks. Future work could extend this methodology to address other cyber threats, refine the model to handle emerging attack patterns, and explore real-time deployment scenarios. These advancements will play a pivotal role in fostering more resilient and secure network infrastructures to meet the demands of an ever-evolving cybersecurity landscape.

FUNDING INFORMATION

This work was supported by a research Grant from Department of Research and Community Service, Universitas Pelita Bangsa, for supporting and funding this research. The author declares no conflict of interest. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ahmad Turmudi Zy	✓	✓	✓	✓	✓	✓		✓	✓	✓		✓		✓
Isariato		✓				✓		✓		✓	✓			
Anggi Muhammad Rifa'i	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓		
Abdul Ghofir					✓		✓		✓		✓			
Muhammad		✓	✓	✓		✓		✓		✓	✓	✓		
Najamuddin Dwi														
Miharja														
Ananto Tri Sasongko					✓		✓			✓		✓		✓

C : Conceptualization	I : Investigation	Vi : Visualization
M : Methodology	R : Resources	Su : Supervision
So : Software	D : Data Curation	P : Project administration
Va : Validation	O : Writing - Original Draft	Fu : Funding acquisition
Fo : Formal analysis	E : Writing - Review & Editing	

CONFLICT OF INTEREST STATEMENT

The authors declare that have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY

The data that support the findings of this study are openly available in UNSW-NB15 at <http://doi.org/10.1109/MilCIS.2015.7348942>, reference number 26.

REFERENCES

[1] A. Djenna, S. Harous, and D. E. Saidouni, "Internet of things meet internet of threats: new concern cyber security issues of critical cyber infrastructure," *Applied Sciences*, vol. 11, no. 10, May 2021, doi: 10.3390/app11104580.

[2] J. Luxemburk, K. Hynek, and T. Cejka, "Detection of HTTPS brute-force attacks with packet-level feature set," in *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2021, pp. 0114–0122, doi: 10.1109/CCWC51732.2021.9375998.

[3] M. Z. Hussain, Z. M. Hanapi, A. Abdullah, M. Hussin, and M. I. H. Ninggal, "An efficient secure and energy resilient trust-based system for detection and mitigation of sybil attack detection (SAN)," *PeerJ Computer Science*, vol. 10, Aug. 2024, doi: 10.7717/peerj-cs.2231.

[4] A. F. Ootom, W. Eleisah, and E. E. Abdallah, "Deep learning for accurate detection of brute force attacks on IoT networks," *Procedia Computer Science*, vol. 220, pp. 291–298, 2023, doi: 10.1016/j.procs.2023.03.038.

[5] S. S. Nalegaev and N. V. Petrov, "Simple criteria to determine the set of key parameters of the DRPE method by a brute-force attack," *Physics Procedia*, vol. 73, pp. 281–286, 2015, doi: 10.1016/j.phpro.2015.09.137.

[6] A. S. Edu, M. Agoyi, and D. Agozie, "Digital security vulnerabilities and threats implications for financial institutions deploying digital technology platforms and application: FMEA and FTOPSIS analysis," *PeerJ Computer Science*, vol. 7, Aug. 2021, doi: 10.7717/peerj-cs.658.

[7] A. Joshi, M. Wazid, and R. H. Goudar, "An efficient cryptographic scheme for text message protection against brute force and cryptanalytic attacks," *Procedia Computer Science*, vol. 48, pp. 360–366, 2015, doi: 10.1016/j.procs.2015.04.194.

[8] J.-S. Cho, Y.-S. Jeong, and S. O. Park, "Consideration on the brute-force attack cost and retrieval cost: A hash-based radio-frequency identification (RFID) tag mutual authentication protocol," *Computers & Mathematics with Applications*, vol. 69, no. 1, pp. 58–65, Jan. 2015, doi: 10.1016/j.camwa.2012.02.025.

[9] S. Jacob, Y. Qiao, Y. Ye, and B. Lee, "Anomalous distributed traffic: detecting cyber security attacks amongst microservices using graph convolutional networks," *Computers & Security*, vol. 118, Jul. 2022, doi: 10.1016/j.cose.2022.102728.

[10] Taskeen and S. Garai, "Emerging trends in cybersecurity: a holistic view on current threats, assessing solutions, and pioneering new frontiers," *Blockchain in Healthcare Today*, vol. 7, no. 1, Apr. 2024, doi: 10.30953/bhty.v7.302.

[11] A. M. Rifai, S. Raharjo, E. Utami, and D. Ariatmanto, "Analysis for diagnosis of pneumonia symptoms using chest X-ray based on MobileNetV2 models with image enhancement using white balance and contrast limited adaptive histogram equalization (CLAHE)," *Biomedical Signal Processing and Control*, vol. 90, Apr. 2024, doi: 10.1016/j.bspc.2023.105857.

[12] N. Mishra and S. Pandya, "Internet of Things applications, security challenges, attacks, intrusion detection, and future visions: a systematic review," *IEEE Access*, vol. 9, pp. 59353–59377, 2021, doi: 10.1109/ACCESS.2021.3073408.

[13] V. Giraddi, S. Giraddi, N. D G, A. Bidaragaddi, and S. G. Kanakareddi, "Machine learning approach to intrusion detection: performance evaluation," *Procedia Computer Science*, vol. 235, pp. 1851–1859, 2024, doi: 10.1016/j.procs.2024.04.176.

[14] Q. Du and J. Zhai, "Application of artificial intelligence sensors based on random forest algorithm in financial recognition models," *Measurement: Sensors*, vol. 33, Jun. 2024, doi: 10.1016/j.measen.2024.101245.

[15] M. Chen and Z. Liu, "Predicting performance of students by optimizing tree components of random forest using genetic algorithm," *Heliyon*, vol. 10, no. 12, Jun. 2024, doi: 10.1016/j.heliyon.2024.e32570.




[16] J. Zhang, "Impact of an improved random forest-based financial management model on the effectiveness of corporate sustainability decisions," *Systems and Soft Computing*, vol. 6, Dec. 2024, doi: 10.1016/j.sasc.2024.200102.

[17] A. K. Dey, G. P. Gupta, and S. P. Sahu, "An efficient cyber assault detection system using feature optimization for IoT-based cyberspace," *Procedia Computer Science*, vol. 235, pp. 757–766, 2024, doi: 10.1016/j.procs.2024.04.072.




- [18] A. M. Rifa'i, E. Utami, and D. Ariatmanto, "Analysis for diagnosis of pneumonia symptoms using chest x-ray based on resnet-50 models with different epoch," in *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Dec. 2022, pp. 471–476, doi: 10.1109/ICITISEE57756.2022.10057805.
- [19] A. Raza, K. Munir, M. S. Almutairi, and R. Sehar, "Novel class probability features for optimizing network attack detection with machine learning," *IEEE Access*, vol. 11, pp. 98685–98694, 2023, doi: 10.1109/ACCESS.2023.3313596.
- [20] S. Facchinetti, S. A. Osmetti, and C. Tarantola, "Network models for cyber attacks evaluation," *Socio-Economic Planning Sciences*, vol. 87, Jun. 2023, doi: 10.1016/j.seps.2023.101584.
- [21] E. Irshad and A. B. Siddiqui, "Context-aware cyber-threat attribution based on hybrid features," *ICT Express*, vol. 10, no. 3, pp. 553–569, Jun. 2024, doi: 10.1016/j.icte.2024.04.005.
- [22] W. F. Urmi *et al.*, "A stacked ensemble approach to detect cyber attacks based on feature selection techniques," *International Journal of Cognitive Computing in Engineering*, vol. 5, pp. 316–331, 2024, doi: 10.1016/j.ijcce.2024.07.005.
- [23] M. F. Safitra, M. Lubis, and H. Fakhurroja, "Counterattacking cyber threats: a framework for the future of cybersecurity," *Sustainability*, vol. 15, no. 18, Sep. 2023, doi: 10.3390/su151813369.
- [24] A. N. Kia, F. Murphy, B. Sheehan, and D. Shannon, "A cyber risk prediction model using common vulnerabilities and exposures," *Expert Systems with Applications*, vol. 237, Mar. 2024, doi: 10.1016/j.eswa.2023.121599.
- [25] C. Mironeanu, A. Archip, C.-M. Amarandei, and M. Craus, "Experimental cyber attack detection framework," *Electronics*, vol. 10, no. 14, Jul. 2021, doi: 10.3390/electronics10141682.
- [26] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *2015 Military Communications and Information Systems Conference (MilCIS)*, Nov. 2015, pp. 1–6, doi: 10.1109/MilCIS.2015.7348942.
- [27] M. Nadeem, A. Arshad, S. Riaz, S. S. Band, and A. Mosavi, "Intercept the cloud network from brute force and DDoS attacks via intrusion detection and prevention system," *IEEE Access*, vol. 9, pp. 152300–152309, 2021, doi: 10.1109/ACCESS.2021.3126535.
- [28] G. Uçtu, M. Alkan, İ. A. Doğru, and M. Dörterler, "A suggested testbed to evaluate multicast network and threat prevention performance of next generation firewalls," *Future Generation Computer Systems*, vol. 124, pp. 56–67, Nov. 2021, doi: 10.1016/j.future.2021.05.013.
- [29] E. Pellegrino *et al.*, "Machine learning random forest for predicting oncosomatic variant NGS analysis," *Scientific Reports*, vol. 11, no. 1, Nov. 2021, doi: 10.1038/s41598-021-01253-y.
- [30] M. Rodríguez, Á. Alesanco, L. Mehavilla, and J. García, "Evaluation of machine learning techniques for traffic flow-based intrusion detection," *Sensors*, vol. 22, no. 23, Nov. 2022, doi: 10.3390/s22239326.
- [31] S. Zhang, X. Xie, and Y. Xu, "A brute-force black-box method to attack machine learning-based systems in cybersecurity," *IEEE Access*, vol. 8, pp. 128250–128263, 2020, doi: 10.1109/ACCESS.2020.3008433.
- [32] I. F. Kilincer, F. Ertam, and A. Sengur, "Machine learning methods for cyber security intrusion detection: datasets and comparative study," *Computer Networks*, vol. 188, Apr. 2021, doi: 10.1016/j.comnet.2021.107840.
- [33] Y. Zhang, J. Liu, and W. Shen, "A review of ensemble learning algorithms used in remote sensing applications," *Applied Sciences*, vol. 12, no. 17, Aug. 2022, doi: 10.3390/app12178654.
- [34] M. Alduailij, Q. W. Khan, M. Tahir, M. Sardaraz, M. Alduailij, and F. Malik, "Machine-learning-based DDoS attack detection using mutual information and random forest feature importance method," *Symmetry*, vol. 14, no. 6, May 2022, doi: 10.3390/sym14061095.
- [35] A. Hamarsheh, "An adaptive security framework for internet of things networks leveraging SDN and machine learning," *Applied Sciences*, vol. 14, no. 11, May 2024, doi: 10.3390/app14114530.
- [36] B. A. Tama and S. Lim, "Ensemble learning for intrusion detection systems: a systematic mapping study and cross-benchmark evaluation," *Computer Science Review*, vol. 39, Feb. 2021, doi: 10.1016/j.cosrev.2020.100357.

BIOGRAPHIES OF AUTHORS






Ahmad Turmudi Zy    is a lecturer with expertise in security systems, sentiment analysis, computer networking, and applied systems. He holds a Master's degree from President University and is currently affiliated with Universitas Pelita Bangsa. His research focuses on enhancing the security and efficiency of networked systems and the application of sentiment analysis in various domains. He can be contacted at email: turmudi@pelitabangsa.ac.id.






Isarianto    is a lecturer in Universitas Pelita Bangsa for Information Technology from 2022. Master Degree from President University in Master Science of Information Technology (M.Sc.) in 2019. His research interests are in information retrieval, artificial intelligent, information system, network security, and computer network. He can be contacted at email: isarianto@pelitabangsa.ac.id.






Anggi Muhammad Rifa'i    is a dedicated lecturer in the Informatics Engineering program at Universitas Pelita Bangsa, with expertise in artificial intelligence, earned a Master's degree from AMIKOM University Yogyakarta. His research areas include data science, machine learning, deep learning, image processing, and computer vision, particularly in the field of bioinformatics. He has published several papers in international journals and conferences and is currently a researcher at Universitas Pelita Bangsa. He can be contacted at email: anggimuhammad@pelitabangsa.ac.id.






Abdul Ghofir    is a master and bachelor degree in Information Technology from President University, Indonesia, in 2015 and 2010. Currently, he is teaching at President University for computer network, network security, and socket programming. The research was about cache distribution in squid proxy server, network defense system monitoring, digital image secured, and smart electricity meter, since 2015. He can be contacted at email: geoff@president.ac.id.



Muhammad Najamuddin Dwi Miharja    is an informatics enthusiast with a focus on medical informatics. He began his academic journey at Universitas Islam Indonesia in 2014, specializing in medical informatics. Currently, he is pursuing his undergraduate degree in Informatics at Institut Teknologi Adhi Tama Surabaya (ITATS). He is passionate about integrating technology into healthcare and is always eager to explore new advancements in the field. He can be contacted at email: najamuddind.dwi@gmail.com.



Ananto Tri Sasongko    holds a Doctor of Computer Science from Universitas Indonesia, Indonesia, in 2022. He received his M.Sc. (Information Management) from the George Washington University, Washington, DC, USA, in 1993. He also got a B.Eng. (Electrical Engineering) in 1990 from Universitas Gadjah Mada, Indonesia. He has currently an Assistant Professor at Informatics Engineering at Universitas Pelita Bangsa, Indonesia, since 2015. His research includes wireless ad hoc networks, artificial intelligence, machine learning, data science, and information systems. He can be contacted at email: ananto@pelitabangsa.ac.id.