Enhancing Intrusion Detection Systems by Integrating Support Vector Machine and Random Forest Classifiers Using Dempster-Shafer Theory

Bilal Waheed[1], Maulana Bintang Irfansyah[2], Idris Winarno[3], Akhmad Alimudin[4]

[1,2,3]*Informatics and Computer Engineering Department, Politeknik Elektronika Negeri Surabaya, Surabaya, Indonesia*

[1]*B.S.c Electrical engineering Department, Federal Urdu University of Arts, Sciences & Technology, Pakistan*

[1]bilal677@pasca.student.pens.ac.id(*)

[2]maulanabin@pasca.student.pens.ac.id, [3,4][idris, alimudin]@pens.ac.id

Received: 2025-10-01; Accepted: 2025-11-20; Published: 2026-01-09

Abstract— The role of Intrusion Detection System (IDS) in defending the networks against the increasing stages of cyberattacks is critical. However, single machine learning models tend to be ineffective due to data imbalances, variability in attack patterns, and ambiguity in network traffic. To address these issues, this paper introduces a hybrid IDS architecture based on Support Vector Machine (SVM) and Random Forest (RF) classifiers, with evidence fusion using the Dempster-Shafer Theory (DST). The strategy leverages the high-quality boundary detection of SVM and the strength of an ensemble of RF. At the same time, DST offers a principled approach to dealing with uncertainty and integrating conflicting evidence. As a benchmark, the CSE-CIC-IDS2018 dataset, a set of labelled benign and multiple-attack traffic flows, was used. Stratified train-test partitioning with a fixed random seed, feature standardization, and class balancing using the Synthetic Minority Oversampling Technique (SMOTE) were considered data pre-processing steps for RF. An SVM trained and optimized with RAPIDS cuML on GPUs was trained and optimized via a grid search over selected hyperparameters. To unify the strengths of models, we further integrated classifier outputs using Dempster-Shafer Theory (DST), which transforms probabilistic outputs into belief assignments and yields an ultimate decision based on the belief assignment with the highest value. The models exhibit high predictive ability, as demonstrated by the experimental results. The DST-based fusion outperformed the two individual classifiers, achieving 97.84% accuracy, 97.41% precision, 94.96% recall, and 96.17% F1-score. In this paper, we show that combining classifiers using DST results in a substantial, computable gain over single-model methods. It is the novelty of using DST-based IDS fusion as well to enhance robustness and balanced detection. These results confirm the value of DST-based fusion in improving IDS performance. These results confirm that combining SVM and RF with DST yields a more robust, reliable IDS. In addition to improving the capability for precise threat discovery, this approach also has certain implications for uncertainly evolving network circumstances, highlighting its suitability for actual-world cybersecurity applications.

Keywords— Intrusion Detection Systems; Support Vector Machine; Random Forest; Dempster-Shafer Theory; Cybersecurity.

I. INTRODUCTION

Due to our reliance on a range of technologies in today's world, cyberattacks are becoming increasingly common and advanced, making it crucial to identify threats promptly and effectively. Rules-based or signature-based intrusion detection systems are repeatedly rigid and create numerous false positives [1]. One of the key elements of the security systems is the Intrusion Detection System (IDS), which monitors network traffic to detect possible attacks or abnormal behaviour, thereby preventing them before they occur [2]. Advanced IDS models that address uncertainty and strengths multiple classifiers are essential for improving detection [3]. Since James P. Anderson developed his model for classifying illegal access, IDS technology has experienced incredible improvements since the 1980s. The initial IDS systems subjected known attack patterns to tests based on signatures and predefined rules. They could only detect known threats, and hence they did not help detect new or modified threats. Due to this difficulty, systems that detect unusual or abnormal activities were introduced in the 1990s [4]. As cybersecurity threats have become more sophisticated, researchers have focused on machine-learning-based IDS to improve automation, flexibility, and attack-detection accuracy [5].

IDS is a valued security precaution [6]. In recent times, hackers have employed sophisticated methods to compromise computer systems and steal confidential information. By effectively using an intrusion detection system, companies can improve security, reduce the likelihood of successful attacks, and limit the damage caused by malicious activity [7].

Compared with the traditional approach, machine learning (ML) can process information in real time and detect emerging trends in web traffic. This characteristic implies that attackers are constantly developing new ways of working. Intrusion detection systems, assisted by ML, enhance threat recognition and continue to improve over time, which is why they are helpful in current security strategies [8]. When detecting threats, IDS systems might be misled by various sensors. Sometimes, analysts can be unclear about whether the signalling is an anomaly, but it is minimized. Since it is unreliable, incident response would be slower and less effective [9]. This kind of uncertainty increases the risks of false positives and missed threats. The Dempster-Shafer Theory is a system of evidence integration that enhances the accuracy and confidence of

DOI : https://doi.org/10.25139/inform.v11i1.11075

intrusion detection decisions by combining evidence from multiple sources [5]. SVM and RF with Dempster-Shafer Theory enhance detection accuracy and reduce decision uncertainty in Intrusion Detection System environments [10].

Hybrid intrusion detectors that combine classifiers such as SVMs and RFs have been reported to improve detector performance by leveraging complementary models [11]. Notably, hybrid arrangements of SVM and RF are likely to yield better detection accuracy than using individual models. However, many of these ensemble systems rely on pure voting/averaging fusion, which does not address conflicts or uncertainties in classifier responses [12]. On the other hand, the present study integrates SVM and RF predictions using DST to provide a more stringent approach to evidence use in the presence of uncertainty and disagreement. The approach thus combines the abilities of the two classifiers, thus authorizing decision-making in intrusion detection. DST provides a formal representation of evidence fusion that can process partially received data or inconsistent data, in contrast to the traditional approach, which relies on probability-based estimates to narrow estimates. DST not only enables attributing mass to particular hypotheses (and assigning a belief to the frame of discernment, which involves uncertainty), but also applies Dempster's rule to integrate multiple sources of evidence and resolve conflicts [13]. It has an advantage in intrusion detection systems because classifier outputs may be inconsistent or uncertain, leading to false detections. In this way, it increases the acceptance and comprehensibility of the result.

The use of SVM and RF in modern Intrusion Detection Systems is driven by their ability to detect complex attack patterns and reduce false positives. Support Vector Machines (SVMs) are effective at defining data classes, whereas Random Forests (RFs) enhance accuracy by integrating decision trees. Their scalability and dynamism are appropriate in scalable dynamic cybersecurity scenarios [14]. SVMs work very well with high-dimensional data [15]. A more imbalanced dataset and a multiclass (which is typical of Intrusion Detection Systems) ensemble algorithm like the Random Forest is more applicable [16]. RFs are flexible in handling multifaceted data sources, and SVMs have a significant impact on IDS accuracy on standard datasets [17]. The combination of these models will increase the resilience of Intrusion Detection Systems, resulting in improved detection across various network traffic scenarios [16], [18]. Most existing Intrusion Detection systems can use only a single classification model or simplistic ensemble methods, which are unlikely to be effective at managing the complexity of today's network traffic. These systems often exhibit ambiguity and conflicting outcomes, especially when faced with different attack patterns or imbalanced information [19].

Random Forest and SVM find considerable differences in network traffic, and the former identifies more deviations. They enable quick threat identification and enhance the security of industries in sensitive locations [20]. To provide more effective protection against attacks, A. H. Farooqi and his colleagues proposed using a voting classifier comprising DT, RF, and XGB. Findings from numerous studies indicate that integrating classifiers enhances the accuracy of security detection [21]. Recent research has shown that DST enhances intrusion detection systems by managing logical uncertainties [22]. Given the outcomes of various classifiers, such as SVM and RF, DST might help make reasonable decisions when different information is available. It improved the IDS's accuracy and performance across different conditions. This was achieved using DST, which made it very easy to manage the system's complex security issues[5]. The latest development has combined DST and deep learning technologies to develop an evidentiality classifier. This method, even in the absence of intrusion detection systems, showed that DST could help to improve classification. The classification results were enhanced by addressing uncertainty and collecting data from multiple sources [23]. It revealed DST's capacity to enhance decision-making efficiency in multiple ways and yielded effective results across a wide range of machine learning tasks [24].

This study presents a hybrid system based on the Dempster-Shafer Theory to combine SVM and RF, unlike traditional IDS methods that use only one classifier, thereby enabling the system to handle uncertainty and unreliable evidence [5][25]. A more evidence-based and flexible solution is the Dempster Rule of Combination, which combines multiple sources to provide more reliable intrusion detection [26]. The Dempster-Shafer theory combines SVM and RF to enhance IDS performance. This method addresses uncertainty by integrating classifiers to increase detection rates and minimize false alarms. Through this method, decisions become more reliable, especially in complex, ever-changing situations [16]. To achieve higher IDS detection accuracy, a dual-model intrusion detection system is developed that combines SVMs with a random forest. By applying the Dempster-Shafer Theory, the system gathers and utilizes classifier outputs to address data uncertainty. Results on the CSE-CIC-IDS2018 dataset prove that the model efficiently detects attacks. Although some hybrid models have been considered, including stacking classifiers with feature selection and multi-stage pipelines [27], only a few studies have examined hybrid IDS frameworks that integrate evidence-based fusion models such as the Dempster-Shafer Theory [28]. This gap is essential to fill, as network threats are constantly evolving; therefore, detection systems should balance accuracy with performance across minority attack classes. IDS is critical, but modern network traffic volumes and complex attacks pose significant challenges for accurate, real-time detection. Single classifiers, such as SVM or RF, are highly prone to model uncertainty and often underperform on imbalanced datasets. In this paper, an improved intrusion detection system incorporating Support Vector Machine and Random Forest classifiers is proposed within the context of the DST. The suggested hybrid approach integrates the advantages of all constituent models into a formal fusion.

The research study aims to make and evaluate a hybrid IDS using the SVM and RF classifiers, with the DST as a decision-fusion model, to increase detection power. Three contributions in this work. To address the drawbacks of single models, it first

DOI : https://doi.org/10.25139/inform.v11i1.11075

frames a hybrid IDS model that combines SVM and RF via DST. Second, it will be able to determine the performance of the proposed approach relative to individual classifiers on the CSE-CIC-IDS2018 dataset. Third, it confirms the relevance of DST-based decision fusion as a proper technique for handling unequal and overlapping network traffic and for optimizing IDS resiliency in a work setting.

II. Research Methodology

As shown in Fig. 1, the proposed IDS includes data collection and pre-processing, feature selection, SVM and Random Forest classification, and decision fusion. The workflow is helpful because it provides effective precision in detection. It has never been done and thus represents a novelty of this work, which integrates varied classifiers via DST fusion and goes beyond simple ensemble techniques such as majority voting or weighted averaging.

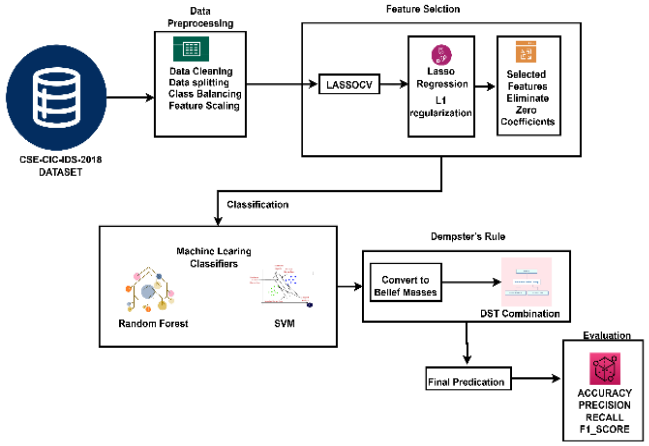


Fig.1. Research Methodology

A. Dataset Description

It was performed on the CSE-CIC-IDS2018 dataset, which consisted of 14 attack-situation files, including DDoS-HOIC, LOIC-UDP, SSH brute-force, SQL injection, and infiltration. The merged data set contained over 9.6 million network flow records and 78 numerical variables, making it highly representative of traffic variety. The intrusion detection data is comprehensive and marked, and it is trained at the Canadian Institute of Cybersecurity [27]. The data consists of real-life traffic and has approximately 16 million entries (including both different types of attacks and 15 categories: DoS, brute-force, botnet, and infiltration). Pre-processing of the selected features was performed to remove noise, perform dimension reduction, and balance class distributions for training and evaluation. Experimental analysis was carried out using the CSE-CIC-IDS2018 dataset, a source of numerous labelled network traffic datasets.

B. Data Pre-processing

In the data pre-processing, correlation-based feature selection, label encoding, normalization, and handling of missing values were performed. Outliers were also identified and removed to enhance the model's resilience. Statistical techniques and domain expertise guided the selection of features, retaining the most discriminative ones. Data cleaning involved removing redundant header rows, eliminating non-informative identifiers such as IP addresses and timestamps, and replacing invalid values (NaN and Infinity) with zeros. Ensured consistent numerical input for the learning models.

Lasso feature selection was used to select the 46 most critical flow features in the CSE-CIC-IDS2018 dataset. The Synthetic Minority Over-sampling Technique (SMOTE) was proposed to address the observed class imbalance in the dataset, where the initial distribution was 71.4% to 28.6% in favour of benign and attack cases, respectively. To combine the outputs of a support vector machine (SVM) and a random forest (RF) classifier to reduce the amount of uncertainty in the model of either individual classifier, a Dempster-Shafer theory (DST)-based fusion mechanism was developed. The methodology is based on the notebook's findings. The linkage is that DST combines the predictions of two robust but complementary models, SVM and RF, to overcome the deficiency of data with unequal representation, a problem initially addressed by SMOTE. An imbalance exists in the dataset.

To address the significant class imbalance, the SMOTE was applied to the training data. The minority attack samples were oversampled at a ratio of 1.0 until the minority sets of benign and attack samples had nearly equal sizes of 5.5 million each, resulting in a balanced training set by definition. All numbers were standardized using StandardScaler to ensure they had similar distributions. This step reduced the 11 million balanced training samples and 1.9 million test samples (78 features each) to a standard scale, thereby improving consistency and merging for the SVM and the random forest classifier.

The experimental setup is presented in Table 2. It includes dataset partitioning, pre-processing techniques, software packages, hardware design, hyperparameter search spaces for Random Forest and SVM, the fusion strategy, and the decision rule based on DST.

TABEL I
EXPERIMENTAL CONFIGURATION FOR REPLICATION

Item	Specification
Dataset	CSE-CIC-IDS2018 (traffic flows: normal + multiple attack classes)
Data Partition	80% training / 20% testing; stratified sampling; random seed = 42
Pre-processing	Standardization with StandardScaler (scikit-learn v1.3.0); imbalance handled by SMOTE (imblearn v0.11.0, k=5)
Libraries	scikit-learn v1.3.0; cuML v23.6.0; imbalanced-learn v0.11.0
System Setup	Intel Core i7 processor; NVIDIA A100 GPU (40 GB); 16 GB RAM
Random Forest	Grid n_estimators ∈ {100, 200, 300}; max_depth ∈ {10, 20, None}; min_samples_split ∈ {2, 5, 10}
SVM	C ∈ {0.1, 1, 10}; kernel ∈ {linear, rbf}; gamma ∈ {scale, 0.01, 0.001}
Fusion Strategy	Probabilistic outputs combined using Dempster–Shafer Theory
Decision Criterion	Predicted label = class with maximum belief score after fusion

DOI : https://doi.org/10.25139/inform.v11i1.11075

C. Data Challenges

The dataset had a severe class imbalance: the most common class was benign traffic (class 0), while all attack categories were class 1. This imbalance is a significant problem for machine learning models, as classifiers often learn to associate with the most common class, which in this case was benign traffic. In this case, we have adopted a binary classification scheme in which benign traffic is the most common category (class 0), and all attack categories are class 1. This configuration approximates a real-world IDS use-case in which the main objective is to differentiate between normal and malicious activity, and stratified sampling was used to ensure that the distribution of labels (71.4 % benign, 28.6 % attacks) was the same between training and testing sets (7.7 million samples and 1.9 million samples, respectively).

D. Feature Selection

To maximize the performance of the Intrusion Detection System on the CSE-CIC-IDS2018 dataset, we selected features using the Least Absolute Shrinkage and Selection Operator (LASSO) regression. In regression models, LASSO uses L1 regularisation to produce sparse solutions. By applying Lasso-based feature selection before classification, the proposed framework reduces dimensionality and improves efficiency, a feature often overlooked in IDS hybrid models [29]. Adding a penalty equal to the coefficients makes LASSO want the coefficients of the least critical features to become zero. As a result, LASSO reduces the feature set, making the IDS easier to use and faster. The goal was to identify key network traffic features that significantly aid in distinguishing between good and bad behaviour.

After pre-processing, 78 numeric features were retained for model training. To further enhance efficiency and reduce redundancy, Lasso feature selection was employed to extract the most discriminative features before classification. To reduce redundancy and improve computational effectiveness, Lasso regression cross-validation (LassoCV) was applied to the balanced, consistent training data. The algorithm incurs a cost when it discards less informative features by setting them to zero, retaining only the most predictive features. Of the 78 original attributes, 46 discriminatory features were employed, comprising statistics on packet length, flow inter-arrival time, flag counts, and header lengths. The features are significant traffic features that distinguish between benign and malevolent network behaviours. Lasso regression reduced the 78 features to 46, identifying the most redundant and weak predictors; the remaining features are the most discriminative and unique to benign and malicious traffic. The selected aspects were packet length statistics, flow inter-arrival times, the number of headers and flags, and throughput-related metrics, which are well-known indicators of abnormal behaviour, such as DoS/DDoS, brute-force, and infiltration.

The attributes selected in this research can be divided into several categories. The first one is packet-based statistics, which include the minimum, maximum, mean, and standard deviation of forward and backward packet lengths. The second one is flow characteristics, which involve the flow period, the rate of bits flowing through the flow, and the gaps between packet arrivals on either side of the flow. The other significant one is the flag counts, protocol-level indicators for PSH, URG, FIN, SYN, RST, ACK, CWE, and ECE. The rest of the data in the dataset are header-type data and segment-type data, such as forward header length, backward header length, and segment size. Finally, the feature also accounts for active and inactive flows, with the active mean, active maximum, and idle minimum. The combination of these 46 properties reflects the statistical, temporal, and protocol-specific characteristics of network traffic, and, depending on this combination, network intrusion detection is likely to succeed.

The 46 features identified were consistently applied to the training (11 million samples) and testing (1.9 million samples) sets, and remained the same across all evaluation stages. Random Forest and Support Vector Machine are two generated classifiers that were evaluated individually, and the evidence fusion was performed using the Dempster-Shafer Theory (DST). Scikit-learn was used to implement RF, and GPU-accelerated SVM from cuML (RAPIDS) was used to manage large-scale training effectively. To improve classification performance, a grid search with cross-validation was used to optimize hyperparameters for both models.

E. Assessment Criteria

Accuracy, precision, recall, and F1-score, as defined in Equations (1) to (4) [30], are used to measure the performance of the proposed IDS. These measures estimate the accuracy, reliability, sensitivity, and the overall trade-off of the classifiers in intrusion detection.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

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

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Where, the TP is correctly identified instances of attacks (true positives), the TN is correctly labelled normal cases (true negatives), the FP is normal traffic which is identified as an attack (false positives, and the FN is attack traffic that is not recognized as such (false negatives)

The four metrics calculated using the values of the confusion matrix are: TP (true positives - correctly identified attacks), TN (true negatives - correctly identified benign flows), FP (false positives - benign flows mistaken by the classifier), and FN (false negatives - attacks not determined by the classifier). Based on these counts, the following performance measures are [30]: Accuracy (Equation 1) is the ratio of the total correct forecasts. Measures of accuracy reflect overall correctness but may be misleading when class distributions are uneven. A model that predicts the majority class may achieve high accuracy while showing little sensitivity to the minority class of interest. Precision (Equation 2) measures the reliability

DOI : https://doi.org/10.25139/inform.v11i1.11075

of optimistic predictions. High precision implies that when the model makes a positive call (e.g., an attack), it tends to be right. The proper focus in the case of costly false alarms is precision, such as blocking legitimate traffic or causing an expensive incident response. Recall (Equation 3) is also known as sensitivity. Recall measures how well the model identifies true positives: a high recall indicates that the model can identify a small number of attacks. Given the high cost of missing a positive instance (a missed intrusion), it is crucial to prioritize recall. F1-score (Equation 4) is the harmonic mean of precision and recall. The harmonic mean is a summary statistic that is more sensitive to extreme imbalance between recall and precision than the arithmetic mean, so the F1 score is a single-number statistic that biases towards the model with a good trade-off between false alarms and correctly identifying positives. It is beneficial in cases of unequal data sets, where reporting accuracy would be inferior [31].

F. Confusion Matrix

A confusion matrix is a tool for visualizing a classification model's performance. It calculates TP, FP, TN, and FN, providing information on the model's accuracy. The confusion matrix can be used to obtain key metrics of accuracy, precision, recall, and F1-score.

G. Model Integration Using Dempster-Shafer Theory

The final SVM and RF predictions were combined using DST to increase decision reliability. Using the theory, each class was assigned a belief mass according to the classifier's confidence. The outputs were then fused using Dempster's rule of combination to arrive at a final decision [32]. This integration improves detection accuracy, particularly for ambiguous or borderline cases. The Dempster-Shafer Theory was employed to derive a reliable and consistent classification by combining the probability results of RF and GPU-accelerated SVM classifiers. It is possible to apply DST, a well-established theory, to combine multiple sources of evidence and resolve the conflicts between them. The RF and SVM probabilistic outputs, and the BBAs, were used to determine whether a specific network instance was an intrusion (with a class label of 1).

The Dempster–Shafer fusion framework integrates evidence from two classifiers (SVM and Random Forest) to achieve a more reliable decision. Each model's output probability is first transformed into basic belief assignments, where the belief for the "Benign" and "Attack" classes is given by and, respectively, as shown in Equation (5). The conflict between the two evidence sources is quantified using the conflict coefficient, defined in Equation (6). To ensure normalization, the denominator prevents division by zero and represents the non-conflicting evidence mass. The combined belief masses for each hypothesis are then computed using Dempster's rule: for the "Benign" class, Equation (7), and for the "Attack" class, Equation (8). The mass of uncertainty follows in Equation (9). Finally, a decision rule selects the combination chosen by the most classifiers; otherwise, the results of the two classifiers are combined to yield a single, more valid response. Given two evidence sources A and B, with BPAs

$$\hat{y} = \begin{cases} 0, if\ m_{AB}(\{0\}) \geq m_{AB}(\{1\}) \\ 1, otherwise \end{cases} \quad (5)$$

Conflict coefficient: K

$$K = m_A(\{0\})m_B(\{1\}) + m_A(\{1\})m_B(\{0\}) \quad (6)$$

$$m_{AB}(\{0\}) = \frac{m_A(\{0\})m_B(\{0\}) + m_A(\{0\})m_B(\Theta) + m_A(\Theta)m_B(\{0\})}{1-K} \quad (7)$$

Normalization factor: $1-K$ in order to prevent complete conflict, the denominator cannot be zero. Combined belief masses (Dempster's rule):

$$m_{AB}(\{1\}) = \frac{m_A(\{1\})m_B(\{1\}) + m_A(\{1\})m_B(\Theta) + m_A(\Theta)m_B(\{1\})}{1-K} \quad (8)$$

$$m_{AB}(\Theta) = \frac{m_A(\Theta)m_B(\Theta)}{1-K} \quad (9)$$

$$m(\{0\}) = 1 - p_1, \qquad m(\{1\}) = p_1, \qquad m(\Theta) = 0 \quad (10)$$

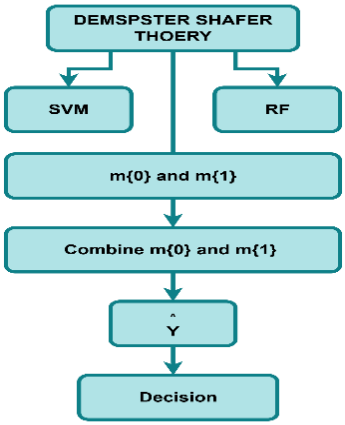


Fig 2. DST Fusion Process for SVM and RF Classifiers

Fig.2 illustrates how DST gradually combines the strengths of the two classifiers to provide a common-sense approach to handling their uncertainty and divergence, thereby achieving more reliable intrusion detection. The process also allows this system to combine evidence from the two classifiers, thereby reducing uncertainty and increasing the reliability of the decisive decision. The strategy combines the unique strengths of RF and SVM. It works especially well when the classifiers' insights differ slightly, making the final results more reliable in the face of network events. This approach strengthens agreement among the classifiers while explicitly managing uncertainty when conflicts arise.

SVM and RF are complementary base classifiers, and their uncertainties can be handled in the DST fusion process. SVM provides prediction probabilities, i.e., chunks of predictions, which serve as the primary input to DST. RF is also able to provide prediction probabilities (P(Attack) on the DST input. DST is the final, formal layer in the decision-making process. It converts the outputs of every classifier into a Belief Mass Function (Basic Belief Assignment, BBA). Using the Dempster-Shafer Rule of combination, the Combinations of the

DOI : https://doi.org/10.25139/inform.v11i1.11075

masses of belief of the SVM and the RF are combined to produce the resultant mass.

III. Result and Discussion

The CSE-CIC-IDS2018 dataset, which multitasks across multiple attack types, was also subjected to the hybrid approach, and the accuracy, precision, recall, and F1-score were evaluated. When comparing machine learning models for object classification, it is also necessary to choose the most accurate model and the most suitable approach to the problem. Integrated models of RF, SVM, and DST are widely popular and proficient at providing considerable benefits. SVM is a robust classification model that performs well when the classes in the dataset are well separated. Random Forest is an ensemble learning method that is highly flexible and robust, and in most cases can produce high-quality results even when the data is very complex and high-dimensional. The decision model, which uses DST-integrated methods, handles uncertainty and conflicting information in the decision process by combining basic machine learning methods with a probabilistic model. Tables II show the classification results on the CSE-CIC-IDS2018 dataset using RF, SVM, and the suggested DST fusion model.

Tabel II
Performance Comparison Table

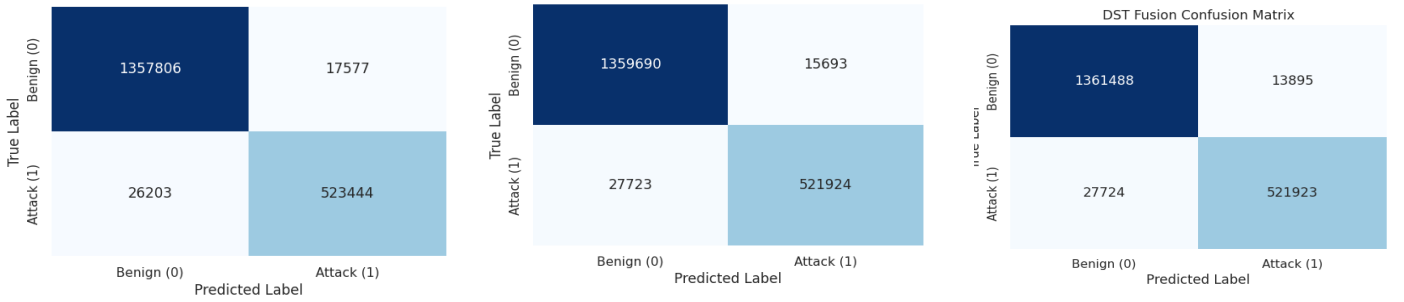
Model	Accuracy	Precision	Recall	F1-Score
SVM	0.977447	0.9709810	0.949562	0.960069
RF	0.977257	0.967511	0.952328	0.959859
DST Fusion	0.978380	0.974068	0.949560	0.961658

The performance of the proposed IDS was assessed using standard evaluation metrics (accuracy, precision, recall, F1-score, and confusion matrix analysis). Table 2 presents the comparative results for the SVM, Random Forest, and the DST-based fusion model. The SVM classifier achieved an accuracy of 97.74%, precision of 97.10%, recall of 94.96%, and F1-score of 96.01%. The Random Forest model also yielded similar results: 97.73% accuracy, 96.75% precision, 95.23% recall, and 95.99% F1-score. The DST fusion method achieved the highest overall performance, with an accuracy of 97.84%, precision of 97.41%, recall of 94.96%, and F1-score of 96.17%, surpassing the single classifiers in all metrics.

Since the size of the test set (1,925,030 samples) is so enormous, an absolute gain of DST Fusion over Random Forest by 0.001123% is an Accuracy weight, which at this sample size will change the number of correctly classified tests by some 2,160 instances. The statistical significance of this difference is most probably significant.

The most significant operational gain is recorded in the Precision measure, where DST Fusion achieved a comparative 0.7% relative gain compared to the RF. In a real-world security operations centre. Precision is essential, as it has a direct, proportional impact on the False Positive Rate. False positives result in wasted time and resources for security analysts searching for harmless traffic. The system uses DST Fusion to provide a more reliable alarm signal and reduce alert fatigue.

Fig.3 shows the confusion matrices for the SVM, Random Forest, and Dempster-Shafer fusion approach, evaluated on the CSE-CIC-IDS2018 Dataset. In the SVM classifier, the majority of benign cases were correctly classified, while a few were incorrectly classified as attacks. Likewise, most attack samples were correctly detected, but some were misclassified as benign. The same trend can be observed in the Random Forest model, though with slightly better detection of attack traffic than the SVM. The fusion model has established the best overall performance, with the fewest false positives and the highest accuracy in both benign and attack classification.



(a) Random Forest Model (b) Support Vector Machine Model (c) DST-Based Fusion (SVM + RF) Model

Fig.3. The Confusion Matrices Model Performance

Comparing the performance of the SVM, RF, and DST-integrated models with the optimum performance of more traditional algorithms, such as logistic regression and k-nearest neighbours (KNN), and more advanced algorithms, such as deep learning, should be done. These traditional and modern models offer insight into the benefits and drawbacks of the models under consideration. Ensuring that SVMs perform better than ensemble methods such as RF requires clean, well-separated data, and even then, performance is not always as good. On the other hand, RF is more resilient and generalizes to a broader range of classification problems, as it can accommodate greater data complexity and is more apt to be chosen when numerous classification tasks are required.

In situations where data is missing or conflicting, the DST-integrated model may be better than the others, as it incorporates uncertainty and is thus more integrated. It is also possible to compare these models with existing technologies, e.g., DNNs, and next-generation algorithms, e.g., gradient boosting models (GBMs), to determine whether older models

DOI : https://doi.org/10.25139/inform.v11i1.11075

can compete or whether newer models can achieve a significant improvement in prediction accuracy.

The above findings indicate that both the SVMs and the Random Forests are suitable detectors. They have their disadvantages, but they complement each other: SVMs tend to sacrifice recall for precision, whereas random forests tend to sacrifice precision for recall. Simultaneously, the number of false positives slightly rises with the help of the Random Forest. When the two models are merged using DST-based fusion, their strengths are combined and enhanced, resulting in improved overall performance, particularly in accuracy and F1-score.

The best results were obtained with the fusion model developed using the DST model, which achieved 97.84% accuracy, 97.41% precision, and 96.17% F1-score. It also minimized false positives to 13,895 without negatively impacting recall at 94.96. The results suggest that the DST can be used to integrate the strengths of both SVMs and Random Forests to achieve more objective and reliable classification. The suggested DST technique is better than current IDS combination techniques, e.g., majority voting and weighted averaging, which are commonly associated with a precision/recall trade-off. DST is a dynamically combined system that uses evidence from multiple classifiers, unlike fixed-weight systems, and considers uncertainty and conflict to retain the accuracy of SVM and the recall of Random Forest, while maximizing F1-score. The evidence-based fusion framework, in turn, is a more adaptable and resilient IDS solution.

The gain from integrating Dempster-Shafer Theory can be clarified by its ability to accommodate uncertainty and combine evidence from diverse classifiers. Whereas SVM focuses on identifying the best possible decision boundaries, Random Forest focuses on capturing highly complex feature interactions via ensemble learning. In contrast, DST integrates the probabilistic outputs of both systems by redistributing belief mass based on confidence levels and the degree of conflict. It reduces uncertainty in ambiguous cases, especially when one classifier is lowly confident, making the system more reliable in its decision-making. The runoff of compensatory learning between the margin-based reasoning provided by SVM and the feature-level diversity provided by RF enables DST to produce even more stable predictions, particularly for mixed or borderline traffic behaviour.

Nevertheless, several limitations remain. The implementation at hand is optimized for batch fusion and could perform poorly in high-speed or large-scale network deployments. Moreover, because the models are trained on familiar attack models, unidentified or novel threats would not be adaptable to them. The next step in merging DST with deep learning models, such as CNNs or transformer-based models, should be pursued in future studies to promote generalization and representation learning in dynamic settings. The online DST-based fusion, which updates belief assignments after each incoming data packet, is another promising approach for enhancing the robustness and flexibility of current intrusion detection systems.

In summary, evaluating and comparing SVM, RF, and DST-integrated models across a wide range of performance metrics yields essential insights into the strengths and weaknesses of each method. Confusion matrix analysis provides a comprehensive understanding of classifier performance across various decision thresholds. The DST-integrated model is proficient at managing uncertainty, whereas SVM and RF are proficient at handling diverse data challenges. Applying feature scaling improved SVM performance by increasing margin separation. It reduced the presence of high-magnitude features, thereby improving the balance in the detection of various attack types. The use of Lasso-based feature selection remarkably reduced model complexity while preserving analytical strength. This not only provides better training competence on the large-scale IDS dataset but also enhances interpretability by identifying the key traffic features most suggestive of attacks. By reducing the feature space from 78 to 46 dimensions, the framework reached a more efficient representation of traffic features without compromising predictive performance. This dimensionality reduction enhanced the scalability and speed of training and also improved the model interpretability.

IV. CONCLUSION

An intensive exploration of the total Precision and F1 Score of the Dempster-Shafer Theory (DST) fusion model, which is a combination of SVM and Random Forest prediction based on the joint set of data. This performance gain can largely be attributed to complementary learning, in which the DST framework leverages agreement between the two dissimilar base classifiers. The fusion strategy reduces False Positives (FP) by giving more weight to conflicting or marginally supported predictions (quantified by the conflict term K), resulting in substantially greater precision and more dependable positive (Attack) decisions. But because the currently employed system fixes uncertainty to zero to create a mass of beliefs, it has a significant scaling issue when considering millions of SMOTE-balanced flows. After the features are selected using Lasso, the problem scale becomes huge when there are millions of SMOTE-balanced flows. A better IDS system combining SVM and RF classifiers with the DST has been proposed in the paper. Combining the results of both classifiers, the fusion technique produces more definite and accurate intrusion detection. Randomly, feature selection is performed using Lasso regression, enabling the model to be much more accurate and practical with fewer features. The proposed solution is typically better or as good as the best single models, as demonstrated by experimental validation on the CSE-CIC-IDS2018 dataset. It is also extremely useful in addressing uncertainty and other attacks. DST integration also improves the system's ability to handle complex, ambiguous threats, and it is an effective approach to developing intelligent next-generation IDS solutions. This model may be used in the future to apply to real-time systems and to combine with deep learning methods to improve performance further. The results of the experiment make it clear that such a fusion is more accurate, achieves higher F1 scores, and produces fewer false alarms, especially for minority attack classes, thereby demonstrating the system's

DOI : https://doi.org/10.25139/inform.v11i1.11075

novelty and applicability. Additional extensions will be implemented dynamically in the future to introduce non-zero mass uncertainty, adding novelty and facilitating more studies combining DST with larger-scale deep learning baseline models.

ACKNOWLEDGEMENT

Bilal Waheed, a student at Politeknik Elektronika Negeri Surabaya, expresses his sincere thanks to the Kemitraan Negara Berkembang Scholarship program as it has helped him with funding throughout the period of conducting this study. He also wishes to express genuine gratitude to the faculty members and colleagues at PENS who provided excellent guidance, technical contributions and constructive feedback which played a significant role in ensuring that this study is successful.

REFERENCES

[1] H. Asad, S. Adhikari, and I. Gashi, "A perspective–retrospective analysis of diversity in signature-based open-source network intrusion detection systems," *Int J Inf Secur*, vol. 23, no. 2, pp. 1331–1346, 2024.
[2] J. Azimjonov and T. Kim, "A comprehensive empirical analysis of data sets, regression-based feature selectors, and linear SVM classifiers for intrusion detection systems," *IEEE Internet Things J*, vol. 11, no. 21, pp. 34676–34693, 2024.
[3] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges," *Cybersecurity*, vol. 2, no. 1, pp. 1–22, 2019.
[4] W. Qiu, Y. Ma, X. Chen, H. Yu, and L. Chen, "Hybrid intrusion detection system based on Dempster-Shafer evidence theory," *Comput Secur*, vol. 117, p. 102709, 2022.
[5] L. Diana, P. Dini, and D. Paolini, "Overview on intrusion detection systems for computers networking security," *Computers*, vol. 14, no. 3, p. 87, 2025.
[6] A. A. Hagar, D. G. Chaudhary, A. Al-Bakhrani, and B. W. Gawali, "Big data analytic using machine learning algorithms for intrusion detection system: A survey," in *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)*, 2020, pp. 6063–6084.
[7] N. Mohamed, "Artificial intelligence and machine learning in cybersecurity: a deep dive into state-of-the-art techniques and future paradigms," *Knowl Inf Syst*, pp. 1–87, 2025.
[8] R. Zuech, T. M. Khoshgoftaar, and R. Wald, "Intrusion detection and big heterogeneous data: a survey," *J Big Data*, vol. 2, no. 1, p. 3, 2015.
[9] H. Bakır and Ö. Ceviz, "Empirical enhancement of intrusion detection systems: a comprehensive approach with genetic algorithm-based hyperparameter tuning and hybrid feature selection," *Arab J Sci Eng*, vol. 49, no. 9, pp. 13025–13043, 2024.
[10] A. Ali, S. Naeem, S. Anam, and M. M. Ahmed, "Machine learning for intrusion detection in cyber security: Applications, challenges, and recommendations," *UMT Artif. Intell. Rev*, vol. 2, no. 2, pp. 41–64, 2022.
[11] H. A. Al Essa and W. S. Bhaya, "Ensemble learning classifiers hybrid feature selection for enhancing performance of intrusion detection system," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 1, pp. 665–676, 2024.
[12] V. Sharma and D. J. Shah, "Ensemble Learning Classifiers and Hybrid Feature Selection for Enhancing Intrusion Detection System Performance," 2025.
[13] C. F. Cheang, Y. Wang, Z. Cai, and G. Xu, "Multi-VMs Intrusion Detection for Cloud Security Using Dempster-shafer Theory.," *Computers, Materials & Continua*, vol. 57, no. 2, 2018.
[14] M. A. M. Hasan, M. Nasser, B. Pal, and S. Ahmad, "Support vector machine and random forest modeling for intrusion detection system (IDS)," *Journal of Intelligent Learning Systems and Applications*, vol. 6, no. 01, pp. 45–52, 2014.
[15] E. C. Uwazie, A. A. Obiniyi, M. Olalere, and P. N. Achi, "Comparison of Random Forest, K-Nearest Neighbor, and Support Vector Machine Classifiers for Intrusion Detection System," in *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG)*, IEEE, 2024, pp. 1–6.
[16] A. Halimaa and K. Sundarakantham, "Machine learning based intrusion detection system," in *2019 3rd International conference on trends in electronics and informatics (ICOEI)*, IEEE, 2019, pp. 916–920.
[17] Y. Zhou, G. Cheng, S. Jiang, and M. Dai, "Building an efficient intrusion detection system based on feature selection and ensemble classifier," *Computer networks*, vol. 174, p. 107247, 2020.
[18] B. A. Tama, S. Y. Lee, and S. Lee, "A Systematic Mapping Study and Empirical Comparison of Data-Driven Intrusion Detection Techniques in Industrial Control Networks.," *Archives of Computational Methods in Engineering*, vol. 29, no. 7, 2022.
[19] A. H. Farooqi, S. Akhtar, H. Rahman, T. Sadiq, and W. Abbass, "Enhancing network intrusion detection using an ensemble voting classifier for internet of things," *Sensors*, vol. 24, no. 1, p. 127, 2023.
[20] Y. Bi, D. Bell, H. Wang, G. Guo, and J. Guan, "Combining multiple classifiers using dempster's rule for text categorization," *Applied Artificial Intelligence*, vol. 21, no. 3, pp. 211–239, 2007.
[21] A. A. Megantara and T. Ahmad, "A hybrid machine learning method for increasing the performance of network intrusion detection systems," *J Big Data*, vol. 8, no. 1, p. 142, 2021.
[22] R. Kavya and J. Christopher, "Interpretable systems based on evidential prospect theory for decision-making," *Applied Intelligence*, vol. 53, no. 2, pp. 1640–1665, 2023.
[23] Ü. Çavuşoğlu, "A new hybrid approach for intrusion detection using machine learning methods," *Applied Intelligence*, vol. 49, no. 7, pp. 2735–2761, 2019.
[24] R. Zhao, Y. Mu, L. Zou, and X. Wen, "A hybrid intrusion detection system based on feature selection and weighted stacking classifier," *IEEE Access*, vol. 10, pp. 71414–71426, 2022.
[25] E. C. Uwazie, A. A. Obiniyi, M. Olalere, and P. N. Achi, "Comparison of Random Forest, K-Nearest Neighbor, and Support Vector Machine Classifiers for Intrusion Detection System," in *International Conference on Science, Engineering and Business for Driving Sustainable Development Goals, SEB4SDG 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/SEB4SDG60871.2024.10629939.
[26] Z. Chen, M. Simsek, B. Kantarci, M. Bagheri, and P. Djukic, "Machine learning-enabled hybrid intrusion detection system with host data transformation and an advanced two-stage classifier," *Computer Networks*, vol. 250, p. 110576, 2024.
[27] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "CSE-CIC-IDS2018 Dataset," *Canadian Institute for Cybersecurity, University of New Brunswick*, 2018.
[28] Y. Tang, D. Wu, and Z. Liu, "A new approach for generation of generalized basic probability assignment in the evidence theory," *Pattern Analysis and Applications*, vol. 24, no. 3, pp. 1007–1023, 2021.
[29] I. H. Putro and T. Ahmad, "Feature Selection Using Pearson Correlation with Lasso Regression for Intrusion Detection System," in *2024 12th International Symposium on Digital Forensics and Security (ISDFS)*, IEEE, 2024, pp. 1–6.
[30] M. Conciatori, A. Valletta, and A. Segalini, "Improving the quality evaluation process of machine learning algorithms applied to landslide time series analysis," *Comput Geosci*, vol. 184, p. 105531, 2024.
[31] H. Kaur and D. K. Sandhu, "Evaluating the Effectiveness of the Proposed System Using F1 Score, Recall, Accuracy, Precision and Loss Metrics Compared to Prior Techniques," *Int. J. Commun. Networks Inf. Secur*, vol. 15, no. 4, pp. 368–383, 2023.
[32] T. M. Chen and V. Venkataramanan, "Dempster-Shafer theory for intrusion detection in ad hoc networks," Nov. 2005. doi: 10.1109/MIC.2005.123

DOI : https://doi.org/10.25139/inform.v11i1.11075