Computer Systems
Science and Engineering

Tech Science Press

Doi:10.32604/csse.2025.062413

ARTICLE

Evaluation and Benchmarking of Cybersecurity DDoS Attacks Detection Models through the Integration of FWZIC and MABAC Methods

Alaa Mahmood and İsa Avcı*

Department of Computer Engineering, Karabuk University, Karabuk, 78000, Türkiye

*Corresponding Author: İsa Avcı. Email: isaavci@karabuk.edu.tr

Received: 18 December 2024; Accepted: 01 April 2025; Published: 25 April 2025

ABSTRACT: A Distributed Denial-of-Service (DDoS) attack poses a significant challenge in the digital age, disrupting online services with operational and financial consequences. Detecting such attacks requires innovative and effective solutions. The primary challenge lies in selecting the best among several DDoS detection models. This study presents a framework that combines several DDoS detection models and Multiple-Criteria Decision-Making (MCDM) techniques to compare and select the most effective models. The framework integrates a decision matrix from training several models on the CiC-DDOS2019 dataset with Fuzzy Weighted Zero Inconsistency Criterion (FWZIC) and Multi-Attribute Boundary Approximation Area Comparison (MABAC) methodologies. FWZIC assigns weights to evaluate criteria, while MABAC compares detection models based on the assessed criteria. The results indicate that the FWZIC approach assigns weights to criteria reliably, with time complexity receiving the highest weight (0.2585) and F1 score receiving the lowest weight (0.14644). Among the models evaluated using the MABAC approach, the Support Vector Machine (SVM) ranked first with a score of 0.0444, making it the most suitable for this work. In contrast, Naive Bayes (NB) ranked lowest with a score of 0.0018. Objective validation and sensitivity analysis proved the reliability of the framework. This study provides a practical approach and insights for cybersecurity practitioners and researchers to evaluate DDoS detection models.

KEYWORDS: Cybersecurity attack; DDoS attacks; DDoS detection; MABAC; FWZIC

1 Introduction

Distributed denial of service (DDoS) attacks have become a pervasive and disruptive threat in the digital landscape. The DDoS attacks are designed to overwhelm and incapacitate targeted systems, rendering them inaccessible to legitimate users. By flooding a network, website, or online service with overwhelming traffic or malicious requests, DDoS attacks disrupt normal operations, causing significant downtime, financial losses, and damage to an organization's reputation [1,2]. DDoS attacks come in multiple forms [3–5]: protocol, volumetric, and application layer attacks.

The DDoS attack model is continually evolving to keep pace with technological advancements. Attackers constantly devise new methods to circumvent service providers' defenses, driven by the evolution of distributed Denial of Service (DoS) techniques [6,7]. As the sophistication and scale of DDoS attacks continue to evolve, organizations must implement robust and proactive defense mechanisms [8]. These mechanisms include implementing traffic monitoring and anomaly detection systems, utilizing mitigation techniques such as rate limiting and traffic filtering, and leveraging the services of specialized DDoS mitigation providers [9]. Moreover, the effective identification and mitigation of DDoS attacks heavily rely

on cooperation and the exchange of information among various entities, including organizations, Internet Service Providers (ISPs), and security communities [10–12].

Many efforts were made to evaluate the DDoS detection model. A study by [13] used True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN), and F-measures to evaluate the DDoS detection model [14]. The main factors the DDoS model should provide are accuracy, False Acceptance Rate (FAR), sensitivity, and specificity. Measurement and accuracy are the main requirements for DDoS detection models [15]. The study in [16] used accuracy as a key indicator for comparison among the DDoS detection models. While the study by [17] relied on the misclassification rate to determine the best DDoS detection models [18], the detection rate and FP rate were used to assess and benchmark these models. It is challenging to compare DDoS detection models across multiple evaluation criteria simultaneously, and the benchmarking process is hindered by the difficulty of comparing different criteria that involve trade-offs and disputes [18].

Two primary factors must be considered when assessing the effectiveness of DDoS detection models: reliability and computational complexity [19]. Nevertheless, the existing method of comparing the proposed model with previous models in the reviewed studies fails to consider all evaluation and benchmarking criteria. Instead, it focuses solely on one review aspect, overlooking the others. This approach lacks flexibility in addressing the conflict or tradeoff between the requirements [20,21]. The contradictory nature of the situation and the tradeoff involved are the primary challenges encountered while evaluating and measuring DDoS detection methods [22–24].

The second problem that impacted the evaluation and benchmarking process was the significance of each criterion. When evaluating DDoS detection models, multiple criteria are considered, and the relevance of each criterion varies depending on the specific objectives of the model. However, the significance of one evaluation criterion may be heightened while reducing the importance of another criterion, depending on the model's aims [24,25]. Hence, a necessary compromise and clash will arise between assessment and benchmarking standards due to the varying significance of each criterion in distinct models [26,27]. A challenge occurs during the benchmarking process of DDoS detection models when many criteria and sub-criteria are considered simultaneously [28,29]. This challenge is attributed to the tradeoff between the requirements, each of which holds varying degrees of importance [30–33]. Previously, Multi-Criteria Decision-Making (MCDM) may have been considered appropriate for these purposes [34,35]. Various techniques can be used to manage MCDM and solve practical problems. These methods help structure issues for Decision Makers (DMs) and analyze, rank, sort, and score many alternatives [36–39].

This study employed the Fuzzy Weighted Zero Inconsistency Criterion (FWZIC) to calculate the weights of evaluation criteria, as it yields more consistent results than the Analytical Hierarchy Process (AHP) and other MCDM weighting methods. Additionally, benchmarking and ranking alternatives were performed using the Multi-Attributive Border Approximation Area Comparison (MABAC) method, one of the most widely applied methods for solving MCDM problems [39]. Additionally, MABAC ranks accurately and quickly, enabling the selection of optimal options with precision. Much of the research reviewed demonstrated that FWZIC yields satisfactory utility values when combined with MABAC because both methods can effectively handle uncertainties related to the problem statement. Neither FWZIC nor MABAC requires much user experience, even for those unfamiliar with MCDM literature. It is recommended that MABAC be employed in conjunction with various scenarios, such as individual and group situations [35]. Two primary decision-making situations are emphasized: the first scenario involves a single decision-maker. Group Decision-Making (GDM) involves multiple decision-makers [40]. GDM refers to a situation in which individuals collaboratively select choices from a set of options. No member of the group is affected by the decision. Individual and group dynamics, including social factors, influence the outcome. The

methodologies employed in collective decision-making systematically collect and incorporate elements from experts, encompassing their expertise across multiple areas. Regarding group cases, each expert is presented with judgment criteria that require subjective assessment.

Furthermore, as shown before, the expert assigns a specific weight to each criterion. Ultimately, the assessment and comparison of DDoS detection models indicate a requirement to incorporate FWZIC and MABAC methods. The proposal entails assigning weights to criteria (reliability, time complexity rate).

The motivation is that attacks require detection efficiency. Many classifiers work to detect attacks. However, selecting the most suitable classifier based on several criteria is challenging. Choosing the best is considered one of the primary motivations for this study; this selection will undoubtedly contribute to improving the speed and accuracy of detection.

The primary aim of this study is to propose a framework for enhancing, evaluating, and benchmarking DDoS detection classifiers. This study is organized into four sections. Section 1 establishes the theoretical framework, including the proposed solution. Section 2 offers evaluation and benchmarking procedures. Section 3 analyzes the findings, validates them, and discusses the proposed methodology. Section 4 provides a brief overview of the research.

2 Materials and Methods

Fig. 1 illustrates the components of the proposed framework.

The proposed framework Fig. 1 is divided into three parts: the decision matrix, the FWZIC method, and the MABAC method. They will be explained in detail below.

2.1 Evaluation and Benchmarking of DDOS Attack Detection Classifiers through the Integration of FWZIC and MABAC Methods

The evaluation and benchmarking framework that was developed relies on MCDM techniques. This study's strategy is formulated by combining FWZIC to assign weights and MABAC for ranking, enabling the identification of optimal alternatives within the suggested decision matrix. The literature analysis on MCDM techniques highlights FWZIC and MABAC as suitable methods for benchmarking and ranking DDoS detection models. The mathematical model of MABAC is proposed for addressing specific issues, such as simultaneously handling multiple evaluation criteria within the suggested decision matrix, even in cases of conflict among the requirements. Furthermore, FWZIC is utilized to assign weights to criteria, addressing the understanding of the significance of these criteria within the proposed decision matrix. Therefore, integrating FWZIC and MABAC methodologies is suitable for assessing and comparing DDoS detection models and their hierarchical ranking.

2.2 Suggested Decision Matrix

Table 1 shows the suggested decision matrix. The rows represent the classifier's metrics (criteria), while the columns represent the classifiers (alternatives). The values of this matrix are obtained after executing the first part of Fig. 1.

To get the decision matrix. The CiC-DDoS2019 dataset has been used. Preprocessing this dataset may reveal irregularities, missing values, outliers, and similar issues that necessitate correction before analysis or modeling endeavors (see Fig. 2). After processing the dataset, it was divided into two parts: 80% for training and 20% for testing.

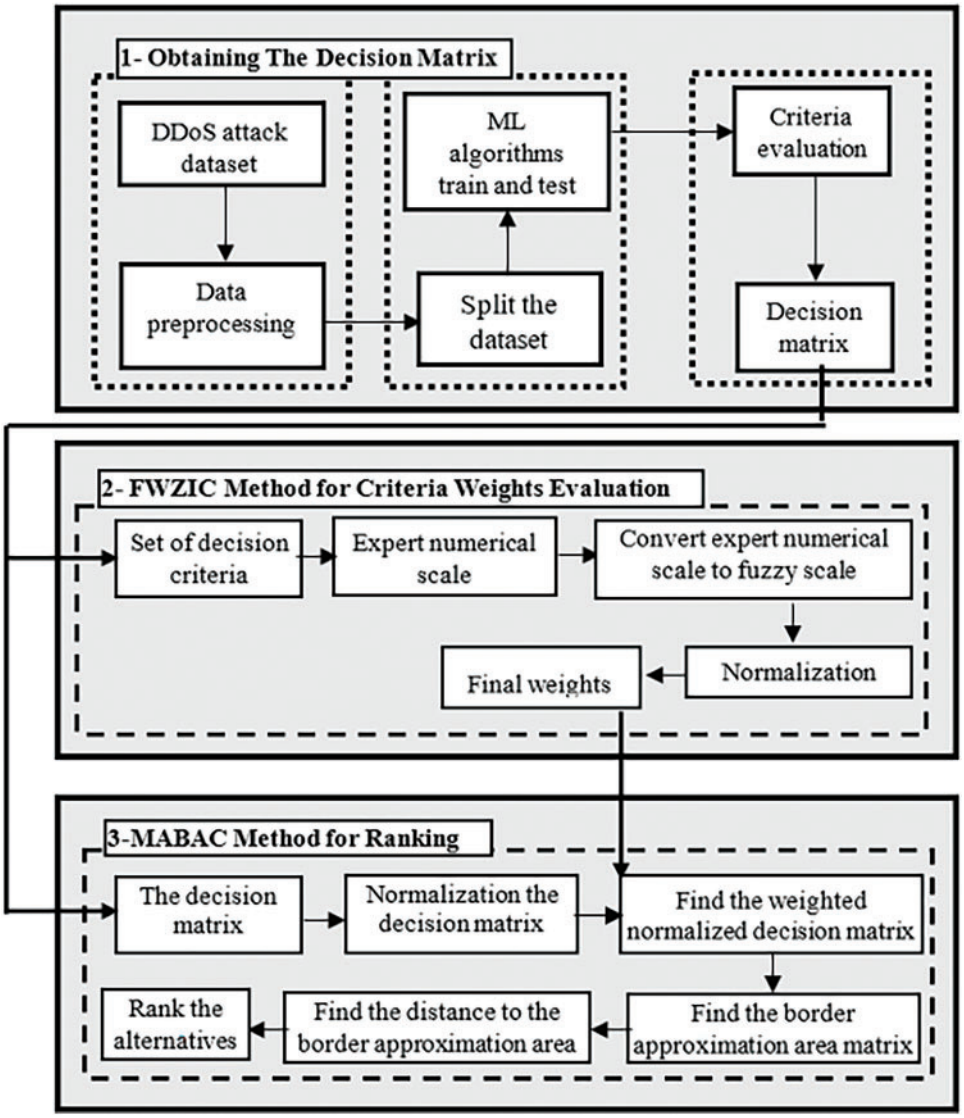

Figure 1: The proposed framework

Table 1: The decision matrix

Alternatives	C1	C2	Cn
A1	(A1, C1)	(A1, C2)	(A1, Cn)
A2	(A2, C1)	(A2, C2)	(A2, Cn)
An	(An, C1)	(An, C2)	(An, Cn)

Different five machine learning algorithms, DPRCT, SVM, LR, NB, KNN, and Stacked Classifiers (SCs), were used to train and test the CiC-DDoS2019 dataset; SC is an ensemble technique that aims at improving the accuracy of results in models by combining multiple classifiers instead of using a single classifier [35]. After that, a set of metrics (ACC, REC, F1, and T) was applied to the classifiers to assess their quality and

performance criteria. The decision matrix can be constructed from the measurements obtained. All work was done using Python.

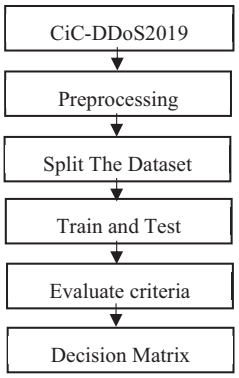

Figure 2: Steps to obtain the suggested decision matrix

2.3 Applying the FWZIC Method for Assigning Criteria Weights

The second part of Fig. 1 represents the implementation of the FWZIC method, which can be summarized in five steps as follows:

1. Collection of evaluation criteria (from the decision matrix)
   - Examination of the pre-agreed set of evaluation criteria.
   - Classify all the criteria, sub-criteria, and relevant indicators based on their behavioral patterns and the assessment employed.
2. The technique of Structured Expert Judgment (SEJ)
   - Identify Experts: The term 'expert' cannot be defined by any quantitative measure of resident knowledge. In this context, the term 'expert in a given subject' refers to an individual who possesses specialized knowledge in a specific field and is recognized by others as an authority.
   - Select Experts: Following the identification of the collection of experts, the selection process for the experts to be utilized in the study is undertaken. Broadly speaking, selecting at least four specialists for a specific subject is necessary.
   - Develop the Evaluation Form: Creating an assessment form is essential, as it is a tool for gathering expert consensus.
   - Experts are identified and selected in the relevant subject domains (e.g., DDoS detection models) to evaluate and determine the importance of the criteria. A SEJ panel has been created, and a form has been designed to gather the collective agreement of all SEJ panelists for each criterion.

Transforming a language scale into a corresponding numerical scale is necessary to facilitate further research. Therefore, during this stage, the experts reported that the level of importance or significance for each criterion on the Likert linguistic scale is transformed into a corresponding numerical scale, as depicted in Table 2. The Likert scale is predicated on the notion that varying levels of relevance are associated with the evaluation criteria. The spectrum of importance levels spans from the lowest to the highest.

3. Constructing the evaluation decision matrix

During this process, the evaluation decision matrix is fabricated. Table 3 presents the fundamental components of this matrix, including the criteria and alternatives.

Table 2: The criteria's importance

Numerical scoring scale	Linguistic scoring scale
1	Not important
2	Slightly important
3	Moderately important
4	Important
5	Very important

Table 3: The evaluation decision matrix

Experts	Criteria		
	C1	C2	Cn
E1	Imp*(E1/C1)	Imp(E1/C2)	Imp(E1/Cn)
E2	Imp(E2/C1)	Imp(E2/C2)	Imp(E2/Cn)
E3	Imp(E3/C1)	Imp(E3/C2)	Imp(E3/Cn)
Em	Imp(Em/C1)	Imp(Em/C2)	Imp(Em/Cn)

Note: * The variable "Imp" denotes the level of importance.

4. Utilization of a decentralized membership function

The fuzzy membership function is applied to the data of the previous step, followed by a defuzzification procedure. This transformation aims to enhance precision and facilitate further data analysis. The fuzzy technique offers the benefit of handling unclear situations by employing ambiguous numbers instead of crisp ones to ascertain the relative value of the criteria. Due to their conceptual and computational simplicity, the primary type of fuzzy number used in fuzzy MCDM is the Triangular Fuzzy Number (TFN). The TFNs are represented as A = (a, b, c), as shown in Fig. 3.

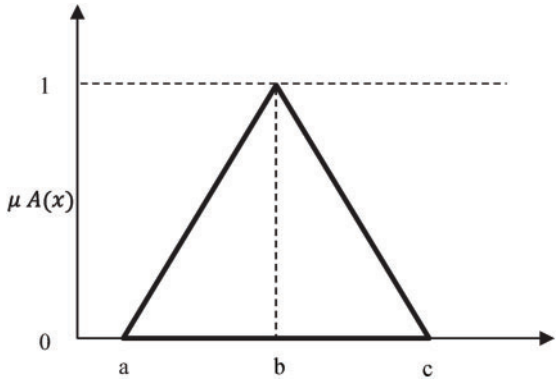

Figure 3: The triangular fuzzy numbers (TFNs)

The TFN membership function (x) is defined as:

$$\mu A\,(x) = \begin{cases} 0 & \text{if } x < a \\ \dfrac{x-a}{b-a} & \text{if } a \le x \le b \\ \dfrac{c-x}{c-b} & \text{if } b \le x \le x \quad \text{where } a \le b \le c \\ 0 & \text{if } x > c \end{cases} \tag{1}$$

Based on the findings shown in Table 4, it is recommended that all linguistic variables be converted into triangular fuzzy numbers. This translation assumes that the fuzzy number is equivalent to the variable linked to each criterion for expert K.

Table 4: Linguistic terms and their equivalent (TFNs)

Linguistic terms	TFNs
Not important	(0.00, 0.10, 0.30)
Slight important	(0.10, 0.30, 0.50)
Moderately important	(0.30, 0.50, 0.75)
Important	(0.50, 0.75, 0.90)
Very important	(0.75, 0.90, 1.00)

5. Make the weight coefficients associated with the evaluation criteria ultimate values

The weight coefficients for the evaluation criteria $(w1, w2, \ldots, wn)$ are concluded using the fuzzification data acquired in the preceding step. The ratio of fuzzification data can be calculated using Eq. (2), which is commonly employed in conjunction with TFNs, as illustrated in Table 5.

$$\frac{\text{Imp}\,(Em/Cn)}{\sum_{j=1}^{k} \text{Imp}\,(Em/Cj)} \tag{2}$$

k is the number of criteria, m = 1, 2, 3 (the current expert), n = 1, 2, 3 (the current criteria).

Table 5: Fuzzy expert decision matrix

Experts	Criteria		
	C1	***C2***	***Cn***
E1	$\frac{\text{Imp}(E1/C1)}{\sum_{j=1}^{n}\text{Imp}(E1/Cj)}$	$\frac{\text{Imp}(E1/C2)}{\sum_{j=1}^{n}\text{Imp}(E1/Cj)}$	$\frac{\text{Imp}(E1/Cn)}{\sum_{j=1}^{n}\text{Imp}(E1/Cj)}$
E2	$\frac{\text{Imp}(E2/C1)}{\sum_{j=1}^{n}\text{Imp}(E2/Cj)}$	$\frac{\text{Imp}(E2/C2)}{\sum_{j=1}^{n}\text{Imp}(E2/Cj)}$	$\frac{\text{Imp}(E2/Cn)}{\sum_{j=1}^{n}\text{Imp}(E2/Cj)}$
Em	$\frac{\text{Imp}(Em/C1)}{\sum_{j=1}^{n}\text{Imp}(Em/Cj)}$	$\frac{\text{Imp}(Em/C2)}{\sum_{j=1}^{n}\text{Imp}(Em/Cj)}$	$\frac{\text{Imp}(Em/Cn)}{\sum_{j=1}^{n}\text{Imp}(Em/Cj)}$

The mean values must be found to get the values of the evaluation criteria's weight coefficients $(w1, w2, \ldots, wn)$. Each column in the fuzzy expert decision matrix in Table 5 has its elements added together,

and the sum is divided by the number of experts. For instance, $w1$ will represent the final fuzzy weight of the criteria $C1$.

2.4 MABAC for Benchmarking and Ranking DDoS Detection Models

Compared to the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) and VlseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR) methods, MABAC is a recent method introduced MABAC is grounded in the measurement of the distance of alternatives from the Border Approximation Area (BAA). It follows a systematic process with a straightforward computational procedure, making it competent in addressing real-world decision-making problems. The steps of the MABAC method are as follows:

1. Get the decision matrix obtained from Section 2.1.
2. Normalizing the decision matrix.
   Eq. (3). For beneficial criteria (ACC, PREC, REC, and F1):

$$Normalized\ value = \frac{Xij - \min(Xj)}{\max(Xj) - \min(Xj)} \tag{3}$$

   Eq. (4). For non-beneficial criteria (T):

$$Normalized\ value = \frac{\max(Xj) - Xij}{\max(Xj) - \min(Xj)} \tag{4}$$

   where (Xij) represents the current element of the decision matrix, while min and max represent the highest and lowest values for the same criterion.
3. Determine the weighted normalized decision matrix as in Eq. (5).
   - Assign weights Wj to each criterion.
   - Compute the weighted normalized values:

$$vij = Wj * xij \tag{5}$$

The weight values obtained from the FWZIC method will be multiplied by the normalized decision matrix. Table 6 shows the output of the decision matrix.

Table 6: Weighted normalized decision matrix [41,42]

Alternatives	Criteria				
	ACC	**PREC**	**REC**	**F1**	**T**
DT	ACC (1) $*$ $w1$	PREC (1) $*$ $w2$	REC (1) $*$ $w3$	F1 (1) $*$ $w4$	T (1) $*$ $w5$
SVM	ACC (2) $*$ $w1$	PREC (2) $*$ $w2$	REC (2) $*$ $w3$	F1 (2) $*$ $w4$	T (2) $*$ $w5$
LR	—-	—-	—-	—-	—-
NB	—-	—-	—-	—-	—-
KNN	—-	—-	—-	—-	—-
SC	ACC (n) $*$ $w1$	PREC (n) $*$ $w2$	REC (n) $*$ $w3$	F1 (n) $*$ $w4$	T (n) $*$ $w5$

Note: $(w1, w2, w3, w4,$ and $w5)$ are the weights obtained from the FWZIC method for each criterion (ACC, PRC, REC, F1, and T) in sequence.

4. Calculate the border approximation area (BAA).

For simplicity, let us assume the BAA is the mean value of the weighted normalized decision matrix for each criterion by applying Eq. (6). Table 7 shows the BAA matrix.

$$Gj = \frac{\sum Cj}{N} \tag{6}$$

Gj represents the current element of a BAA array.
$\sum Cj$ The sum of elements of column j for the weighted normalized decision matrix.
N The number of alternatives.

5. Compute the distance of each alternative from the BAA as in Eq. (7).

Calculate the distance Qij:

$$Qij = vij - Gj \tag{7}$$

vij represents the current element of the weighted normalized decision matrix.

6. Rank the alternatives.

- Compute the overall performance score for each alternative.

$$Qi = \sum_{j=1}^{n} Qij \tag{8}$$

Qi represents the sum of the elements of each row in the matrix.

Rank the alternatives in descending order of Qi, where a higher score indicates a better alternative.

Table 8 presents the algorithm of the proposed work, which is divided into three sections. The first section explains how to obtain the decision matrix, the second section shows the FWZIC method, and the third section offers the MABAC method.

Table 7: BAA matrix [42,43]

$G1 = \frac{\sum C1}{N}$	$G2 = \frac{\sum C2}{N}$	$G3 = \frac{\sum C3}{N}$	$G4 = \frac{\sum C4}{N}$	$G5 = \frac{\sum C5}{N}$

Table 8: Proposed work algorithm

Step 1: Decision matrix
1. **Input:** DDoS attack dataset (CiC-DDoS2019).
2. **Preprocess** the dataset to clean and structure the data.
3. **Split** the dataset into training and testing sets.
4. **Train and test** machine learning (ML) algorithms.
5. **Evaluate** criteria based on the ML model results.
6. **Generate** the decision matrix.
Step 2: FWZIC
7. **Define** a set of decision criteria.
8. **Assign** an expert numerical scale to the criteria.
9. **Convert** the expert numerical scale into a fuzzy scale.
10. **Normalize** the fuzzy scale values.
11. **Compute** the final weights of the criteria.

(Continued)

Table 8 (continued)

Step 3: MABAC

12. **Input:** Decision matrix and final weights.
13. **Normalize** the decision matrix.
14. **Find** the weighted normalized decision matrix.
15. **Compute** the border approximation area matrix.
16. **Calculate** the distance to the border approximation area.
17. **Rank** the alternatives based on computed distances.

Output:

18. **Obtain** ranked alternatives for decision-making.

3 Results

The results of the suggested framework are as follows: Section 3.1 presents the evaluation findings of the proposed decision matrix. Section 3.2 presents the outcomes of the FWZIC, which are used to calculate the weight of the evaluation criteria. Section 3.3 presents the findings of MABAC for benchmarking and rating the outcomes of DDoS detection models.

3.1 Results of the Classifiers and the Decision Matrix

The current section presents the results of six classifiers evaluated according to the identified criteria in Table 9. The numerical values in the matrix cells represent the performance values of each detection classifier, depending on the corresponding criteria.

Table 9: Decision matrix results

Alternatives	Criteria				
	ACC	PREC	REC	F1	T
DT	93.11	93.12	93.11	93.11	15.7
SVM	99.32	99.33	99.32	99.32	9.1
LR	99.13	99.14	99.13	99.13	12.5
NB	76.22	75.25	76.22	80.02	4.1
KNN	97.82	97.81	97.82	97.84	3.9
SC	99.57	99.46	99.45	99.45	143.5

The SC with the classification group (DT, SVM, LR, and KNN) outperformed all other classifiers, achieving an accuracy of 99.57%, the highest value obtained in this study. The LR appears consistently in high-performing stacks, indicating its strength across different combinations and its effectiveness in complementing other classifiers. This analysis confirms that utilizing diverse classifiers in stacking can significantly enhance accuracy, underscoring the importance of carefully selecting model combinations to achieve optimal results. Conversely, the NB single classifier was more conservative, reaching the lowest accuracy of 76.22%. The algorithms are ordered from best to worst performance, according to accuracy: SC > SVM > LR > KNN > DT > NB.

Regarding execution time, the KNN algorithm was the fastest (3.9 s), whereas the SC took the longest. It takes the longest (143.5 s) because it deals with several classifiers. This time is relatively long compared with the rest of the classifier implementation times. Also, this conflicts with one of the interests of this work, in which the time factor is essential for eliminating the attack as quickly as possible.

The superiority of the single linear SVM classifier over the other single classifiers highlights the dataset's inherently linear nature. This conclusion arises from the linear SVM's ability to effectively establish clear boundaries among different categories via linear techniques. Therefore, classifiers based on the linear data separation technique are the best choice for this problem. In contrast, the naive Bayes (NB) single classifier, which relies on a probabilistic approach, demonstrated the lowest performance with the given problem because it is incompatible with the dataset structure. This contradiction suggests the limitations of probability-based classifiers when applied to challenges of this nature.

All the other mixed classifiers accuracy values except the value of the combination "DT + SVM + LR + KNN" are less than the accuracy value of the SVM single classifier (99.32%), so the work does not rely on them; also, their implementation time is larger than the SVM implementation time due to the integration of more than one classifier in Table 10.

Table 10: Stacked classifiers ACC

Stacked classifiers	Accuracy
DT + SVM + LR + KNN	99.57%
DT + SVM + LR	99.22%
KNN + NB	97.52%
LR + NB	98.78%
LR + KNN	98.73%
SVM + NB	99.17%
SVM + KNN	99.31%
DT + NB	92.86%
DT + SVM	99.17%

3.2 FWZIC results

The weights of the evaluation criteria utilize the FWZIC approach. The first process begins with FWZIC by identifying the evaluation criteria (ACC, PREC, REC, F1, and T). The relevance level of each evaluation criterion is determined by gathering the viewpoints of three professional specialists using a specially designed evaluation form. The recommendations provided by experts are subsequently converted into a standardized scoring scale, as seen in Table 4, Section 2.2. Afterward, the expert decision matrix (EDM) is built, as shown in Table 11. The numbers in Table 11 indicate the significance levels for each evaluation attribute, as determined by expert judgment. As in step four of the FWZIC steps, the EDM is transformed into a Fuzzy Matrix, as illustrated in Table 12. The procedure involves converting crisp values into fuzzy numbers with equal values. The final weight is ultimately determined by applying it throughout the defuzzification process. Table 13 displays the ultimate weights assigned to the five evaluation criteria of the DDoS detection models.

Table 11: The EDM

Expert	Accuracy	Precision	Recall	F1 score	Time (s)
E1	5	4	4	2	2
E2	5	4	4	3	3
E3	4	4	4	3	3

Table 12: Fuzzy-EDM

Expert	Accuracy	Precision	Recall	F1 score	Time (s)
E1	(0.75, 0.9, 1)	(0.5, 0.75, 0.9)	(0.5, 0.75, 0.9)	(0.1, 0.3, 0.5)	(0.1, 0.3, 0.5)
E2	(0.75, 0.9, 1)	(0.5, 0.75, 0.9)	(0.5, 0.75, 0.9)	(0.3, 0.5, 0.75)	(0.3, 0.5, 0.75)
E3	(0.5, 0.75, 0.9)	(0.5, 0.5, 0.9)	(0.5, 0.75, 0.9)	(0.3, 0.5, 0.75)	(0.3, 0.5, 0.75)

Table 13: Final weights

Evaluation criteria	Weight
T	0.2585
ACC	0.22833
PREC	0.22027
REC	0.14647
F1	0.14644

The final weights indicate that all experts have close opinions in Table 9. The highest weight values correspond to time and accuracy, two crucial elements in detecting DDoS attacks. Coordination between these two criteria is essential. High accuracy is required in detecting the attacks, but the time the classifier takes during the detection is also crucial. This dispute is referred to as a "criteria conflict", which involves calibrating all criteria to achieve optimal results.

3.3 MABAC Results

The MABAC method calculates the ranking of the DDoS detection models based on determined criteria. The MABAC results are in Table 14. They demonstrate that the SVM algorithms achieved the best ranking among other algorithms. This algorithm is the most suitable for our work (DDoS attack detection) because it has high accuracy and a short execution time compared to other algorithms.

So after training and testing with the CiCDDoS2019 dataset, passing them to many classifiers, taking the opinions of the experts, and putting it all in MCDM, it turns out that the SVM classifier is the most ideal for working in the environment of DDoS attacks, no wonder, it has many characteristics and good compatibility between all criteria so, the sound tuning of this classifier has made it stand out from the rest. This result, in turn, does not diminish the status of the other classifiers; if the goal is high detection accuracy, then the SC classifier can be chosen, of course, provided the time factor is not a significant concern in this case.

Table 14: Detection models ranking results

Detection models	Score	Rank
SVM	0.04440	6
KNN	0.01518	5
LR	0.01366	4
DT	0.01356	3
SC	0.01182	2
NB	0.00184	1

3.4 Validation

Below is a description of the two methods used to verify the results.

3.4.1 Objective Validation

Objective validation is a step that validates the work we have done. The classifiers were divided into two groups: Group 1, comprising SC, NB, and DT, and Group 2, consisting of LR, SVM, and KNN. The average and variance for each group were calculated. In general, the objective validation results, as presented in Table 15, indicate that the average and variance values provide evidence supporting the validity and systematic ranking of the groups, as determined by the MABAC results of the DDoS detection models. It shows that SVM remained within the group with lower average and variance values in Table 15.

Table 15: Objective validation results

Alternatives	Accuracy	Precision	Recall	F1 score	Time		
SC	0.995700	0.994600	0.994500	0.994500	0.996500		
NB	0.765492	0.243414	0.233585	0.195375	0.028571		
DT	0.935121	0.063744	0.063751	0.063751	0.109408		
	0.9002042	0.102386	0.099112	0.0863753	0.3793263	**Average**	0.3134808
	0.1210906	0.126224	0.1207406	0.099633	0.5390366	**Variance**	0.201345
LR	0.995581	0.003217	0.003218	0.003218	0.087108		
SVM	0.997489	0.001307	0.001307	0.001307	0.063415		
KNN	0.982424	0.01659	0.01639	0.016189	0.027178		
	0.9918315	0.007038	0.0069717	0.0069047	0.0592337	**Average**	0.2143959
	0.0082025	0.0083272	0.0082123	0.008097	0.030183	**Variance**	0.0126044

3.4.2 Sensitivity Analysis

Sensitivity Analysis examination in MCDM helps to understand the stability and robustness of the proposed solution and comprehend which parameter changes affect preferences for alternatives. It shows how various conditions influence the stability of ranking in Fig. 4.

It shows the alternatives and their rankings for the proposed solution and the other three scenarios in Fig. 4. As seen in all scenarios, the experts agreed that the proposed solution yielded the best result for the SVM classifier, which ranked 6th. Their opinions again agreed with our proposed solution regarding the NB classifier, which took the lowest rank (1). As for the remaining classifiers (DT, LR, KNN, and SC), the

experts' opinions varied. For instance, the rank of the DT was not fixed in all scenarios with the proposed solution; it took ranks 2, 3, and 4. The same goes for the remaining classifiers (LR, KNN, and SC). It changes in all scenarios, indicating their susceptibility to alterations in terms of criterion or weight.

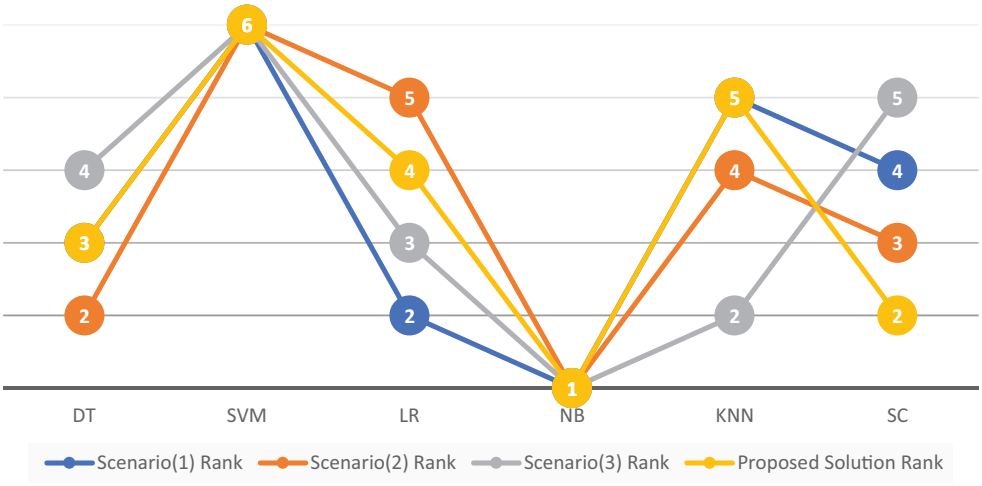

Figure 4: Sensitivity analysis

4 Conclusion

The infected devices used in DDoS attacks act as soldiers, executing commands from an attacker simultaneously. This attack exhausts the server's resources, preventing legitimate users from accessing it or making it invisible. A considerable amount of research has been conducted in this field; however, selecting the appropriate classifier for the current work is challenging, particularly when multiple criteria are involved. It cannot depend on a classifier that achieves high accuracy while ignoring other criteria.

Therefore, the contribution of this research was to enhance the accuracy of the detection by adding the stacked classifier, benchmarking, and evaluating (choosing the best among a group of classifiers) by using two methods of the MCDM, which are FWZIC and MABAC, under the presence of several criteria (accuracy, F1 score, Precision, Recall, and Time). A questionnaire was directed at three experts with a long history in cybersecurity to set the criteria weights. Two methods were also used to verify the results: objective validation and sensitivity analysis.

The FWZIC and MABAC methods proved that the SVM algorithm was superior to the rest and more suitable for the current work. Objective validation divided the classifiers into two groups. The average and variance were calculated for each group, where the SVM classifier was in the group with the lowest variance and average, so it is superior to the rest. Sensitivity Analysis also yielded the highest rank for the SVM classifier in all ranking scenarios, indicating that it is superior to the rest. This progress represents the benchmarking aspect of this work, which aims to select the most suitable classifier for the current study. Another notable achievement of this work was the Stacked classifier, which demonstrated that combining these algorithms (DT, SVM, LR, and KNN) yields the best results, with a high accuracy of 99.57%, representing a significant enhancement over this work.

FWZIC and MABAC have demonstrated their ability to select the most suitable classifiers based on several criteria, making them the most appropriate options in this field. Stacked classifiers achieve the highest accuracy due to their ability to integrate multiple models from each classifier. The SVM classifier is more

efficient and superior to the rest due to its short execution time and high accuracy. Therefore, this classifier is handy for applications on high-priority websites, such as renowned e-commerce sites or those related to national security. For other websites, where time is not a critical factor, the stacked classifier can be used with high accuracy, albeit at the expense of time.

This study's limitations are the data set and classifiers used to make it more comprehensive. Another set of datasets and classifiers can be added to make the framework more comprehensive.

Acknowledgement: This study was produced from the doctoral thesis "Cyber Security Defense Mechanism against Distributed Denial of Service Attacks" at Karabük University Graduate Education Institute.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Alaa Mahmood, İsa Avcı; data collection: Alaa Mahmood; analysis and interpretation of results: Alaa Mahmood, İsa Avcı; draft manuscript preparation: Alaa Mahmood, İsa Avcı. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The dataset is available at https://www.unb.ca/cic/datasets/ddos-2019.html (accessed on 1 January 2025). The University of New Brunswick in Canada is affiliated with the Canadian Institute of Cybersecurity.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

Abbreviations

DDoS	Distributed denial of service
MCDM	Multi-criteria decision making
FWZIC	Fuzzy weighted zero inconsistency criterion
MABAC	Multi-attributive border approximation area comparison
EDM	Expert decision matrix
BAA	Border approximation area
ML	Machine learning
DT, SVM, LR,	Decision tree, support vector machine, logistic regression,
KNN, NB, SC	k-nearest neighbor, naive bayes, stacked classifier
ACC, PREC, REC, F1, T	Accuracy, precision, recall, F1 score, time
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
VIKOR	VlseKriterijumska Optimizacija I Kompromisno Resenje

References

1. Wani S, Imthiyas M, Almohamedh H, Alhamed KM, Almotairi S, Gulzar Y. Distributed denial of service (DDoS) mitigation using blockchain—a comprehensive insight. Symmetry. 2021;13(2):227. doi:10.3390/sym13020227.
2. Roopak M, Tian GY, Chambers J. Multi-objective-based feature selection for DDoS attack detection in IoT networks. IET Netw. 2020;9(3):120–7. doi:10.1049/iet-net.2018.5206.
3. Koca M, Ali Aydin M, Sertbaş A, Zaim AH. A new distributed anomaly detection approach for log IDS management based on deep learning. Turk J Elec Eng Comp Sci. 2021;29(5):2486–501. doi:10.3906/elk-2102-89.
4. Mishra A, Gupta BB, Joshi RC. A comparative study of distributed denial of service attacks, intrusion tolerance and mitigation techniques. In: 2011 European Intelligence and Security Informatics Conference; 2011 Sep 12–14; Athens, Greece. p. 286–9. doi:10.1109/EISIC.2011.15.

5. Abhishta A, van Heeswijk W, Junger M, Nieuwenhuis LJ, Joosten R. Why would we get attacked? An analysis of attacker's aims behind DDoS attacks. J Wirel Mob Netw Ubiquitous Comput Dependable Appl. 2020;11(2):3–22. doi:10.22667/JOWUA.2020.06.30.003.
6. Liu C, Huang J. DDoS defense systems in large enterprises: a comprehensive review of adoption, challenges, and strategies. J Artif Intell Mach Learn Manag. 2018;2(1):1–21.
7. Singh R, Tanwar S, Sharma TP. Utilization of blockchain for mitigating the distributed denial of service attacks. Secur Priv. 2020;3(3):e96. doi:10.1002/spy2.96.
8. Guntamukalla DVR. Mitigation against distributed-denial of service attacks using distribution and self-learning aegis system [dissertation]. Kingsville, TX, USA: Texas A & M University-Kingsville; 2017.
9. Ahmed Issa AS, Albayrak Z. DDoS attack intrusion detection system based on hybridization of CNN and LSTM. Acta Polytech Hung. 2023;20(2):105–23. doi:10.12700/APH.20.2.2023.2.6.
10. Ali J, Roh BH, Lee B, Oh J, Adil M. A machine learning framework for prevention of software-defined networking controller from DDoS attacks and dimensionality reduction of big data. In: 2020 International Conference on Information and Communication Technology Convergence (ICTC); 2020 Oct 21–23; Jeju, Republic of Korea. p. 515–9. doi:10.1109/ictc49870.2020.9289504.
11. Uddin R, Kumar SAP, Chamola V. Denial of service attacks in edge computing layers: taxonomy, vulnerabilities, threats and solutions. Ad Hoc Netw. 2024;152(11):103322. doi:10.1016/j.adhoc.2023.103322.
12. Avcı İ., Koca M. Predicting DDoS attacks using machine learning algorithms in building management systems. Electronics. 2023;12(19):4142. doi:10.3390/electronics12194142.
13. Suresh M, Anitha R. Evaluating machine learning algorithms for detecting DDoS attacks. In: Advances in Network Security and Applications: 4th International Conference, CNSA 2011; 2011 Jul 15–17; Chennai, India. Berlin/Heidelberg, Germany: Springer; 2011. p. 441–52.
14. Tuan TA, Long HV, Son LH, Kumar R, Priyadarshini I, Son NTK. Performance evaluation of Botnet DDoS attack detection using machine learning. Evol Intell. 2020;13(2):283–94. doi:10.1007/s12065-019-00310-w.
15. Jyoti N, Behal S. A meta-evaluation of machine learning techniques for detection of DDoS attacks. In: 2021 8th Internatioanl Conference Computing for Sustainable Global Development (INDIACom); 2021 Mar 17–19; New Delhi, India. p. 522–6.
16. Jadhav V, Devale P, Jadhav R, Molawade M, Mohite S, Bidwe RV. Bug predictive models based on data analytics and soft computing techniques: a survey. In: 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom); 2023 Mar 15–17; New Delhi, India. p. 822–31.
17. Ojugo A, Eboka AO. An empirical evaluation on comparative machine learning techniques for detection of the distributed denial of service (DDoS) attacks. J Appl Sci Eng Technol Educ. 2020;2(1):18–27. doi:10.35877/454ri.asci2192.
18. Gao Y, Feng Y, Kawamoto J, Sakurai K. A machine learning based approach for detecting DRDoS attacks and its performance evaluation. In: 2016 11th Asia Joint Conference on Information Security (AsiaJCIS); 2016 Aug 4–5; Fukuoka, Japan. p. 80–6. doi:10.1109/AsiaJCIS.2016.24.
19. Butt HA, Al Harthy KS, Ali Shah M, Hussain M, Amin R, Rehman MU. Enhanced DDoS detection using advanced machine learning and ensemble techniques in software defined networking. Comput Mater Contin. 2024;81(2):3003–31. doi:10.32604/cmc.2024.057185.
20. Özçelik İ, Brooks R. Distributed denial of service attacks: real-world detection and mitigation. London, UK: Chapman and Hall/CRC; 2020. doi:10.1201/9781315213125.
21. Fachkha C, Bou-Harb E, Debbabi M. Inferring distributed reflection denial of service attacks from darknet. Comput Commun. 2015;62(39):59–71. doi:10.1016/j.comcom.2015.01.016.
22. Kumar I. Emerging threats in cybersecurity: a review article. Int J Appl Nat Sci. 2023;1(1):1–8.
23. Liu Z, Jin H, Hu YC, Bailey M. Practical proactive DDoS-attack mitigation via endpoint-driven in-network traffic control. IEEE/ACM Trans Netw. 2018;26(4):1948–61. doi:10.1109/TNET.2018.2854795.
24. Al-Omari M, Abu Al-Haija Q. Performance analysis of machine learning-based intrusion detection with hybrid feature selection. Comput Syst Sci Eng. 2024;48(6):1537–55. doi:10.32604/csse.2024.056257.

25. Brooks RR, Yu L, Ozcelik I, Oakley J, Tusing N. Distributed denial of service (DDoS): a history. IEEE Ann Hist Comput. 2022;44(2):44–54. doi:10.1109/MAHC.2021.3072582.
26. Karami M, Park Y, McCoy D. Stress testing the booters: understanding and undermining the business of DDoS services. In: Proceedings of the 25th International Conference World Wide Web; 2016 Apr 11–15; Montreal, QC, Canada. p. 1033–43.
27. Yusof AR, Udzir NI, Selamat A. Systematic literature review and taxonomy for DDoS attack detection and prediction. Int J Digit Enterp Technol. 2019;1(3):292. doi:10.1504/IJDET.2019.097849.
28. Gavrić Ž, Simić D. Overview of DOS attacks on wireless sensor networks and experimental results for simulation of interference attacks. Ing Inv. 2018;38(1):130–8. doi:10.15446/ing.investig.v38n1.65453.
29. Kalutharage CS, Liu X, Chrysoulas C, Pitropakis N, Papadopoulos P. Explainable AI-based DDOS attack identification method for IoT networks. Computers. 2023;12(2):32. doi:10.3390/computers12020032.
30. Mirkovic J, Reiher P. A taxonomy of DDoS attack and DDoS defense mechanisms. SIGCOMM Comput Commun Rev. 2004;34(2):39–53. doi:10.1145/997150.997156.
31. Distributed Denial of Service (DDoS) Attacks/tools. Archived on wayback machine, 2018 May 24. [cited 2025 Jan 1]. Available from: https://web.archive.org/.
32. The Spread of the Code-Red Worm (CRv2). Archived on wayback machine, 2017 Nov 20. [cited 2025 Jan 1]. Available from: https://web.archive.org/.
33. Annamalai M, Dhanushiya S, Mythreyi O, Mishra JS. Enhancing multicriteria decision-making through cryptographic security systems. In: Multi-criteria decision-making and optimum design with machine learning. Boca Raton, FL, USA: CRC Press; 2024. p. 224–32.
34. Tešić D, Marinković D. Application of fermatean fuzzy weight operators and MCDM model DIBR-DIBR II-NWBM-BM for efficiency-based selection of a complex combat system. J Decis Anal Int Comp. 2023;3(1):243–56. doi:10.31181/10002122023t.
35. Almahdi EM, Zaidan AA, Zaidan BB, Alsalem MA, Albahri OS, Albahri AS. Mobile-based patient monitoring systems: a priori tisation framework using multi-criteria decision-making techniques. J Med Syst. 2019;43(7):219. doi:10.1007/s10916-019-1339-9.
36. Koca M, Avci I. A novel hybrid model detection of security vulnerabilities in industrial control systems and IoT using GCN + LSTM. IEEE Access. 2024;12(1):143343–51. doi:10.1109/ACCESS.2024.3466391.
37. Bhol SG. Applications of multi criteria decision making methods in cyber security. In: Choudhury A, Kaushik K, Kumar V, Singh BK, editors. Cyber-physical systems security. Studies in big data. Vol. 154. Singapore: Springer; 2025. p. 233–58. doi:10.1007/978-981-97-5734-3_11.
38. Khatari M, Zaidan AA, Zaidan BB, Albahri OS, Alsalem MA. Multi-criteria evaluation and benchmarking for active queue management methods: open issues, challenges and recommended pathway solutions. Int J Info Tech Dec Mak. 2019;18(4):1187–242. doi:10.1142/S0219622019300039.
39. Sahoo SK, Goswami SS. A comprehensive review of multiple criteria decision-making (MCDM) methods: advancements, applications, and future directions. Decis Mak Adv. 2023;1(1):25–48. doi:10.31181/dma1120237.
40. Avcı İ, Koca M. A novel security risk analysis using the AHP method in smart railway systems. Appl Sci. 2024;14(10):4243. doi:10.3390/app14104243.
41. Sałabun W. How the normalization of the decision matrix influences the results in the VIKOR method? Procedia Comput Sci. 2020;176:2222–31. doi:10.1016/j.procs.2020.09.259.
42. Khalid KM. Selection of the best village crop potential using the multi-attribute border approximation area comparison (MABAC) method. J Artif Intell Eng Appl (JAIEA). 2023;3(1):394–407. doi:10.59934/jaiea.v3i1.341.
43. Rahim N, Abdullah L, Yusoff B. A border approximation area approach considering bipolar neutrosophic linguistic variable for sustainable energy selection. Sustainability. 2020;12(10):3971. doi:10.3390/su12103971.