Enhanced Intrusion Detection in Software-Defined Networking using Advanced Feature Selection: The EMRMR Approach

Raed Basfar
Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia
raedbasfar@gmail.com (corresponding author)

Mohamed Y. Dahab
Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia
mdahab@kau.edu.sa

Abdullah Marish Ali
Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia
ammali@kau.edu.sa

Fathy Eassa
Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia
feassa@kau.edu.sa

Kholoud Bajunaied
Department of Finance, University of Business and Technology, Jeddah, Saudi Arabia
k.bajunaid@ubt.edu.sa

Received: 13 October 2024 | Revised: 3 November 2024 | Accepted: 5 November 2024

ABSTRACT

Most traditional IP networks face serious security and management challenges due to their rapid increase in complexity. SDN resolves these issues by the separation of control and data planes, hence enabling programmability for centralized management with flexibility. On the other hand, its centralized architecture makes SDN very prone to DDoS attacks, hence necessitating the use of advanced and efficient IDSs. This study focuses on improving IDS performance in SDN environments through the integration of deep learning techniques and novel feature selection methods. This study presents an Enhanced Maximum Relevance Minimum Redundancy (EMRMR) approach that incorporates a Mutual Information Feature Selection (MIFS) strategy and a new Contextual Redundancy Coefficient Upweighting (CRCU) strategy to optimize feature selection for early attack detection. Experiments on the inSDN dataset showed that EMRMR achieved better precision, recall, F1-score, and accuracy compared to the state-of-the-art approaches, especially when fewer features are selected. These results highlight the efficiency of the proposed EMRMR approach in the selection of relevant features with minimal computational overhead, which enhances the real-time capability for IDS in SDN environments.

Keywords-software-defined networking; distributed denial of service; deep learning; enhanced maximum relevance minimum redundancy; mutual information feature selection; contextual redundancy coefficient upweighting

I. INTRODUCTION

Security is one of the main obstacles to the adoption and deployment of a Software Defined Network (SDN) across various networks, despite its advantages [1]. The network's central controller is susceptible to a single point of failure because it serves as its beating heart [2]. If the attacker can take advantage of the controller system, he can either control or obstruct the entire network as he desires. One of the most significant risks to SDN networks is DDoS attacks [3], as they can target any SDN layer, including the data, control, and application layers. In addition, DDoS attacks may target the communication links that connect the data link to the control layer. Several mitigation approaches have recommended a backup controller to reduce the damage caused by DDoS attacks. However, as a secondary controller is also vulnerable to DoS/DDoS attacks, this is not a workable solution [4].

Intrusion Detection Systems (IDSs) are standard security tools to keep an eye on and identify hostile activity within an organization's network, raising alerts in case they detect attacks or if the observed traffic from the incoming or leaving network matches suspicious activity [5, 6]. Given that security concerns rank among the most critical problems of SDNs, much research has focused on developing IDSs as an essential solution [7]. Statistical, Machine Learning (ML), and Deep Learning (DL) techniques are frequently used for anomaly-based detection solutions [8, 9]. SDN's centralized control plane design offers fresh ways to thwart DDoS attacks. This serves as a motivation for this study to employ DL approaches to mitigate the issue of DDoS attacks in SDNs.

Feature selection is an essential preprocessing step that is necessary for the effectiveness of anomaly detection models [10]. By removing irrelevant and redundant features, these methods can preserve the most representative attributes of the initial dataset [11]. Optimized subset characteristics shorten the classifier's execution time while simultaneously increasing accuracy and detection rate. Therefore, fewer features can help to develop a lightweight model with low computing overhead and prediction latency and detect attacks in real-time environments. Furthermore, since feature selection techniques help to avoid the curse of dimensionality, the model is less likely to experience overfitting [12]. To achieve high model performance utilizing ML/DL tasks, several studies have focused on feature selection strategies to eliminate noisy and meaningless features [13]. Three general ways can be used to select features: filter, wrapper, and embedded methods [14-16].

Current DDoS attack prevention measures are ineffective in SDNs, although several feature selection techniques employ ML to detect DDoS attacks [17-19]. The lack of an SDN network intrusion dataset is one of the significant drawbacks of previous studies [20]. Some datasets were produced using a standard network and not an SDN design [21, 22]. However, this adaption might not be sufficiently suitable for actual SDN detection [23]. The security risks associated with SDNs are different from those that typically affect legacy networks in terms of their nature. For example, the SDN controller receives a request for a policy when any unmatched flow is triggered at the open flow switches. The intruder can launch a new type of DDoS attack by sending massive amounts of mismatched flows that overload the controller's resources. As both malicious and legitimate traffic is sent to the SDN controller for decision-making, the attack traffic also imitates the same typical behavior. As a result, the DDoS class on the SDN network does not always share the essential characteristics of DDoS attacks on traditional networks. Furthermore, omitting the most critical parameters while employing inappropriate feature selection techniques can waste a large amount of data.

Given the effectiveness of DL in several domains, combining SDN and DL can improve IDS performance and network security [24, 25]. However, the requirement for an IDS to be lightweight and have high detection rates is increasing with network speeds [26]. Feature selection is an important step in achieving optimal intrusion detection performance. Effective feature subsets can shorten training and testing times, allowing for lightweight IDS that ensure high detection rates and are appropriate for online and real-time attack detection [27]. Mutual Information (MI) is a popular feature selection technique that has been used to classify features and assess the most relevant to DDoS attacks [28].

Integrating DL with SDN opens an attractive perspective in enhancing IDS by improving detection accuracy and adaptability to evolving network threats [29]. Given the ever-increasing network speeds, a lightweight and high-performance IDS is urgently needed, particularly for real-time detection. Feature selection plays a very important role, as it reduces processing load while retaining all critical information, enabling possible fast and effective detection [30, 31]. All features are ranked according to their importance using techniques such as MI, which optimizes IDS for quick and efficient response, making it appropriate for real-time security applications [32].

Mutual Information Feature Selection (MIFS) is widely used to improve IDS performance in various environments, including SDN. It considers nonlinear relationships between features and class labels, making it suitable for handling complex and nonlinear data patterns. It can also handle noisy and incomplete data, reducing their dimensionality and improving IDS accuracy. However, the relevance-redundancy trade-off is a common issue, as including redundant features can negatively affect IDS performance. The current calculation of the redundancy coefficient is not suitable for the detection of attacks whose behavior is constantly changing, as data lacking sufficient attack patterns can make it difficult to perceive common characteristics. This study proposes an enhanced feature selection technique called Contextual Redundancy Coefficient Gradual Upweighting (CRGU) MIFS, which evaluates candidate features individually instead of comparing them with common characteristics of already selected features. This study proposes an improved redundancy-relevancy tradeoff technique for the MIFS goal function, integrates it into the training phase of the IDS model for the ISDN, and conducts an experimental evaluation to measure the accuracy of the improved model and compare it with existing solutions. The objectives of this study can be described as follows.

- Propose an improved redundancy-relevancy tradeoff technique for the MIFS goal function.
- Integrate the improved MIFS into the training phase of the IDS model for the ISDN.
- Conduct an experimental evaluation to measure the accuracy of the improved model and compare it with existing solutions.

II. RELATED WORKS

Conventional IP networks, still in widespread use today, have become more complicated and challenging to administer. Network complexity increases when IT operators must access network devices such as switches and routers independently using vendor-specific commands to implement any high-level network policies, such as Quality of Service (QoS) or routing policies. IP-based network devices also include vertical integration. Embedded within the same network device are the control plane, which makes decisions, and the data plane, which determines how to route network traffic based on directives from the control plane. Connecting the control and data planes can limit the network's ability to adapt to the dynamic nature of the network. Furthermore, in traditional networks, the rapid expansion of networking can lead to a considerable decrease in network innovation and an increase in maintenance expenses [23]. Furthermore, there is an increase in the number of middle-box devices, such as firewalls, load balancers, detection and defense systems, etc., as all devices are dispersed throughout the network [23]. In [33], it was stated that 57 network companies have reported a significant growth in middle-box devices, and the number now matches that of other required network equipment, such as routers.

The developing network design, commonly referred to as SDN, promises faster failover and central network control, thereby addressing many of the constraints of traditional IP networks. By separating the control layer from the underlying infrastructure components, SDN aims to eliminate vertical integration. With the help of a centralized controller, decoupling the two layers improves network flexibility and makes network management easier. Regardless of the underlying network technology, the new paradigm enables operators to manage the entire network using software APIs connected to the SDN controller through the northbound interface. The ability of the SDN system to provide global visibility motivates numerous companies, such as Microsoft, Huawei, and Google, to use the new paradigm in their network data centers [33].

IDSs are essential for protecting SDN systems from cyber threats. Feature selection is a critical component of an IDS design aimed at identifying the most relevant features for effective intrusion detection. Various feature selection algorithms have been proposed to find the optimal set of features for IDSs in various environments, including SDN [34]. These algorithms use statistical measures, such as correlation and information gain, or population-based heuristic search approaches, such as particle swarm optimization, ant colony optimization, simulated annealing, and genetic algorithms. Some algorithms use unsupervised feature subset selection methods, fuzzy rough set theory, and mutual information-based feature selection algorithms such as MIFS. Other methods include MIFS-U, mRMR, and multi-objective evolutionary wrappers [35-38]. Attribute evaluation techniques, such as ReliefF, Chi-squared, Correlation Feature Selection (CFS), and Principal Component Analysis (PCA), are used to facilitate the feature selection process. These techniques employ search techniques such as BestFirst, ExhaustiveSearch, GreedyStepwise, RandomSearch, and Ranker for feature ranking.

MI-based feature selection methods have garnered significant attention in IDS research due to their ability to capture dependencies between variables. Numerous studies have investigated the application of MI in feature selection for IDS in SDN systems. In [39], the Normalized Mutual Information Feature Selection (NMIFS) method was introduced, which is a filter-based approach that utilizes MI. Similarly, in [11], the Joint Mutual Information Maximization (JMIM) and Normalized Joint Mutual Information Maximization (NJMIM) methods were proposed to address the overestimation of feature significance in MI-based feature selection. In [5], the performance of an IDS model using an MI-based feature selection algorithm, called MMIFS, was examined. In [40], a feature selection method was developed that combined MI and Pearson's correlation coefficient to design an effective IDS. In [10, 11] MI-based feature selection algorithms were also proposed for IDS. Furthermore, in [12, 41], the importance of MI in feature selection for IDS was emphasized. These studies underscore the importance of MI in selecting relevant features for intrusion detection. Moreover, in [13], a cross-correlation-based feature selection method was compared with MI-based selection, highlighting the relevance of feature selection techniques in IDS. Additionally, in [42], the importance of feature selection in ML-DL-based IDS was emphasized to enhance effectiveness and scalability.

The advantages of the proposed EMRMR approach over existing schemes for intrusion detection in SDN environments are highlighted by comparing the key metrics and techniques used in some related feature selection methods. Table I shows a comparison of existing schemes with the proposed one.

TABLE I. COMPARISON WITH EXISTING SCHEMES

Feature selection technique	Dataset	Performance metrics	Advantages	Limitations
MIFS + CRCU	inSDN	Precision, Recall, F1-score, Accuracy	Efficient feature selection, low computational cost, high detection accuracy	Limited testing on larger datasets
MRMR	Custom	Precision, F1-score	Simple implementation, moderate relevance	High redundancy in feature selection
ReliefF	inSDN	Accuracy, Recall	Effective for high-dimensional data	Ineffective in early attack detection
Random Forest (RF) feature selection	UNSW-NB15	Precision, Recall, Accuracy	Reduces redundancy in high-dimensional features	Computationally intensive
Chi-Square	CICIDS2017	F1-score	Strong statistical relevance	Less adaptive to SDN-specific data

III. METHODOLOGY

MIFS is a well-known feature selection method that can efficiently choose pertinent features regardless of the data distribution, making it appropriate for early detection situations in which the data lack sufficient attack patterns. Figure 1 illustrates the research process for feature selection.

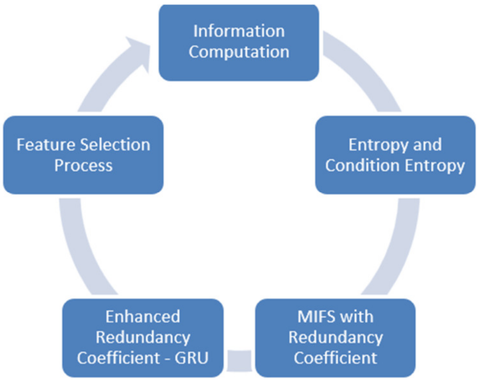
Information Computation
Entropy and Condition Entropy
MIFS with Redundancy Coefficient
Enhanced Redundancy Coefficient - GRU
Feature Selection Process

Fig. 1. Research process for feature selection.

MI is a measure of the amount of information that two discrete variables exchange with one another. Equation (1) provides the MI computation.

$$I(X;Y) = H(X) - H(X|Y) =$$

$$\sum_{y \epsilon Y} \sum_{x \epsilon X} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (1)$$

where $p(x)$ and $p(y)$ are the marginal distributions of x and y, $p(x,y)$ is the joint distribution of x and y, $H(X)$ is the entropy of X, and $H(X|Y)$ is the conditional entropy of X given Y. The following equation is used to compute the entropy $H(X)$.

$$H(X) = -\sum_{x_i \in X} p(x_i) \log(p(x_i)) \quad (2)$$

The following equation can be used to compute the conditional entropy $H(X|Y)$.

$$(X|Y) =$$

$$-\sum_{y_j \in Y} p(y_j) \sum_{x_i \in X} p(x_i|y_j) \log\left(p(x_i|y_j)\right) \quad (3)$$

Equation (4) represents the general formula for the linear combinations of Shannon information terms [30].

$$J(X_k) =$$

$$I(X_k;Y) - \beta \sum_{X_j \epsilon S} I(X_j;X_k) + \gamma \sum_{X_j \epsilon S} I(X_j;X_k|Y) \quad (4)$$

Terms (5) and (6), stand for the relevancy and redundancy terms, respectively, in this equation. Both terms are weighed by parameters β and γ, which have values between 0 and 1. The redundancy term is represented by the sum of marginal redundancy, which is expressed in (7), and conditional redundancy, which is expressed in (8).

$$I(X_k;Y) \quad (5)$$

$$\beta \sum_{X_j \epsilon S} I(X_j;X_k) + \gamma \sum_{X_j \epsilon S} I(X_j;X_k|Y) \quad (6)$$

$$\beta \sum_{X_j \epsilon S} I(X_j;X_k) \quad (7)$$

$$\gamma \sum_{X_j \epsilon S} I(X_j;X_k|Y) \quad (8)$$

The conditional mutual information between the candidate and other features in the selected set S given the class label Y is expressed by (7), whereas the mutual information between the candidate feature and the class label Y is expressed by (5).

$$\beta = \gamma = \frac{1}{|S|} \quad (9)$$

$$\beta = \frac{|S|}{|F|} \quad (10)$$

$$\gamma = \frac{|F-S1}{|F|} \quad (11)$$

A. *Improved (Situational) Upweighting of Redundancy Coefficient*

The β and γ parameters are used in the Contextual Redundancy Coefficient Upweighting (CRGU) technique. The relative contextual relevancy in relation to the label is represented by the coefficient's size, which also indicates the degree of trust in the contextual redundancy term. A CRCU is proposed that calculates the redundancy using (10) and (11), as opposed to the current MIFS that calculates it using (9). Every time a new feature is introduced to the chosen set, the CRCU progressively increases the weights rather than updating the value linearly. Because it also considers the class label, such conditional relevancy is less affected by contextual redundancy.

$$CRGU_{MIFS(x_k)} =$$

$$MI(x_k,y) - \frac{|S|}{|F|}\left(\sum_{s_j \epsilon S} I(x_k,x_j)\right) + \frac{|F-S|}{|F|}\left(\sum_{X_j \epsilon S} I(X_j;X_k|Y)\right)$$

where $|F|$ and $|S|$ denote feature counts in the original set and the selected set, respectively.

B. *Mutual Information Feature Selection based on CRGU*

The integration of the CRGU into the MIFS is seen in (12). It guarantees that, given the features that have already been chosen, the feature with the highest MI with the class label at the end of each iteration is added to the set S. The MI value for every feature in the original set F is first calculated using this technique. Following selection, the feature with the highest MI value is kept in S. Equation (12) is used to add the following features to the chosen set. Subsequently, the $J(X)$ value of each feature in S is used to rank them from highest to lowest. The features that rank higher than n are kept, while the remaining features are eliminated. The number of required features determines the value of n.

IV. RESULTS AND DISCUSSION

The study used a dataset specifically designed for SDN systems, with a focus on assessing the performance of the MIFS approach combined with the CRCU technique. The objective of the experiment is to evaluate the detection accuracy and the real-time practicality of the IDS. A thorough examination was carried out to assess how feature selection affects the performance of several deep learning models in effectively identifying DDoS attacks. The results emphasize enhancements in detection rates, showcasing the effectiveness of the proposed technique in improving SDN security.

The inSDN dataset [20] was used in the experimental evaluation. It is specifically tailored to address the distinct issues of identifying intrusions and malevolent actions within SDN settings. Unlike conventional network datasets, inSDN represents unique operating characteristics and attack vectors exclusive to SDNs. This makes it an ideal basis for creating and testing IDSs customized for this technology. The dataset consists of several data sources, such as network traffic statistics, controller logs, and flow rules. The collected key aspects encompass packet size, inter-arrival time, protocol type, source and destination IP addresses, port numbers, flow durations, frequency of control messages, types of control commands sent, response times, and special anomaly indications such as spikes in mismatched flow requests. The dataset is carefully annotated to distinguish between regular network traffic, different forms of DDoS attacks, and other irregularities, ensuring a thorough coverage of possible security risks. Data gathering entails the utilization of a regulated SDN testbed, employing network simulation tools such as Mininet, as well as real-world SDN controllers, such as OpenDaylight or ONOS. This process involves modeling a wide range of attack scenarios, including volumetric DDoS attacks, protocol exploits, and application layer attacks. Preprocessing involved normalizing the data, selecting features based on MIFS, and dividing the data into training and testing sets. This ensures the dataset's resilience and suitability for training ML models, evaluating feature selection methods, and creating real-time detection systems. The inSDN dataset is an essential resource for researchers and practitioners, enabling advances in network security by offering a thorough and practical basis for the development and testing of IDSs in SDN environments.

Table I presents the performance metrics of the proposed EMRMR technique across different feature counts. The table illustrates the precision, recall, F1-score, and accuracy for feature counts of 5, 10, 15, 20, 25, and 30. The EMRMR method achieved outstanding results with a feature count of 5, boasting precision, recall, F1-score, and accuracy all at 0.99. This indicates an exceptional ability to identify the most relevant features for classification with minimal redundancy. As the feature count increases, the metrics remain consistently high, with slight variations. For instance, with 20 features, the method maintains high performance with a precision of 0.89, recall of 0.88, F1-score of 0.89, and accuracy of 0.90. These results underscore the robustness and effectiveness of the EMRMR method in feature selection, ensuring high classification accuracy while efficiently managing feature space. This stability across different feature counts highlights EMRMR's adaptability and reliability in various scenarios, making it a valuable tool for enhancing the performance of ML models.

TABLE II. EVALUATION RESULTS OBTAINED BY THE PROPOSED EMRMR

Feature count	Precision	Recall	F1-score	Accuracy
5	0.99	0.99	0.99	0.99
10	0.88	0.88	0.88	0.89
15	0.87	0.87	0.87	0.89
20	0.89	0.88	0.89	0.9
25	0.86	0.85	0.86	0.87
30	0.89	0.88	0.89	0.9

Figure 2 provides a comparison of the precision of various feature selection methods across different feature counts, including MRMR [43], EMRMR (proposed), [44], [45], and [19]. The EMRMR method demonstrates better performance, particularly with fewer features. At a feature count of 5, EMRMR achieves 0.99 precision, significantly higher than MRMR (0.87), [44] (0.84), [45] (0.83), and [19] (0.81). When the feature count increases to 10, EMRMR maintains a high precision of 0.88, which is slightly lower than MRMR's 0.89 but still competitive with [44] (0.85), [45] (0.84), and [19] (0.83). For feature counts of 15 and 20, EMRMR shows a consistent precision of 0.87 and 0.89, respectively, aligning closely with MRMR (0.87 and 0.88) and surpassing [44] (0.86) and [19] (0.84-0.85). With 25 features, EMRMR's precision is 0.86, comparable to MRMR (0.89) and slightly lower than [44] (0.84), [45] (0.84), and [19] (0.84). At 30 features, EMRMR maintains a high precision of 0.89, significantly outperforming MRMR, which drops to 0.80 and remains competitive with [44] (0.86), [45] (0.85), and [19] (0.83).

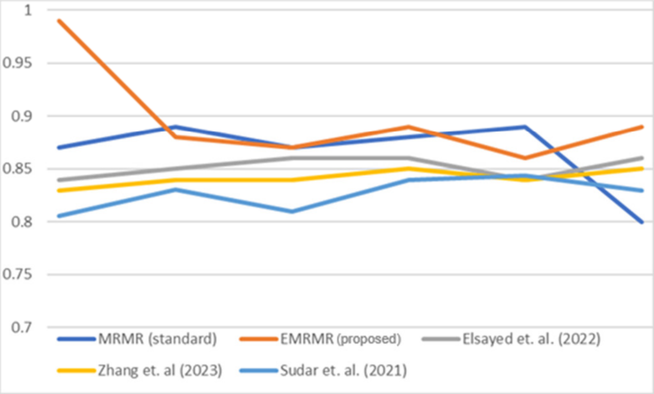

Fig. 2. Precision achieved by various methods across feature counts.

This comparison underscores EMRMR's consistent high precision across varying feature counts, particularly excelling at lower feature counts, making it a robust and reliable method for feature selection in comparison to MRMR and other contemporary methods. Even as the feature count increases, EMRMR maintains high precision. The good performance of EMRMR can be attributed to its enhanced mechanism, as described previously. EMRMR improves on the standard MRMR approach by incorporating additional steps that better handle feature relevance and redundancy. Specifically, EMRMR uses a more sophisticated evaluation of feature interactions, ensuring that selected features provide maximum discriminatory power while minimizing redundancy. This results in a more efficient and effective feature selection process, particularly evident at lower feature counts, where the impact of redundant features can be more pronounced. Furthermore, the ability of EMRMR to maintain high precision with fewer features underscores its robustness in selecting the most relevant features without compromising accuracy. This is crucial for applications with limited computational resources or where interpretability is important. By effectively balancing relevance and redundancy, EMRMR provides a significant advantage in feature selection, leading to improved model performance.

Figure 3 presents a comparison of recall values for various feature selection methods across different feature counts, including MRMR [43], EMRMR (proposed), [44], [45], and [19]. The results highlight the good performance of the

EMRMR method, especially at lower feature counts. For a feature count of 5, EMRMR achieves a near-perfect recall of 0.99, significantly outperforming MRMR (0.68), [44] (0.846), [45] (0.857), and [19] (0.806). As the feature count increases to 10, EMRMR maintains a high recall of 0.88, matching the performance of MRMR and exceeding the recall of [44] (0.853), [45] (0.861), and [19] (0.830). At feature counts of 15 and 20, EMRMR continues to show strong performance with recall values of 0.87 and 0.88, respectively, surpassing MRMR (0.81 and 0.79) and maintaining competitive performance with [44] (0.840 and 0.830), [45] (0.820 and 0.840), and [19] (0.810 and 0.840). With 25 features, EMRMR and MRMR both achieve a recall of 0.85, comparable to [44] (0.830), [45] (0.830), and [19] (0.844). At 30 features, EMRMR maintains a high recall of 0.88, significantly higher than MRMR (0.5), and competitive with [44] (0.860), [44] (0.850), and [19] (0.829).

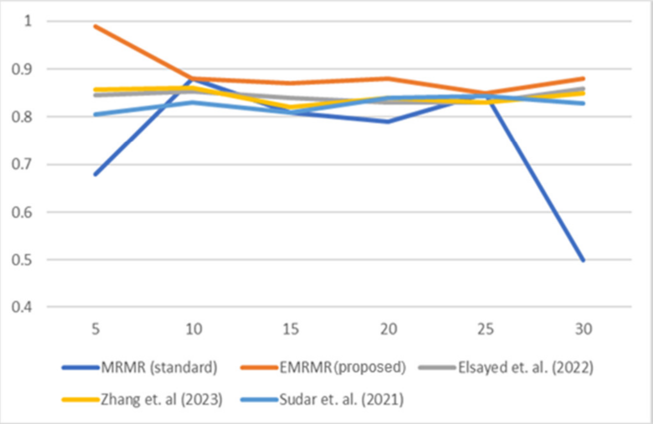

Fig. 3. Recall achieved by various methods across feature counts.

This comparison underscores EMRMR's capability to maintain high recall values across varying feature counts, particularly excelling at lower counts. The enhanced mechanism of EMRMR, which effectively balances feature relevance and redundancy, enables it to achieve such high recall, highlighting its robustness and reliability in accurately identifying relevant features. This makes EMRMR a highly effective method for feature selection, ensuring comprehensive detection of relevant instances in various applications. Such an improvement can be attributed to its enhanced mechanism, as described previously. EMRMR refines the standard MRMR approach by incorporating additional steps that more effectively handle feature relevance and redundancy. This advanced evaluation process ensures that selected features maximize discriminatory power while minimizing redundancy, leading to more accurate and comprehensive detection of relevant instances. The result is a more effective feature selection process, particularly evident at lower feature counts where the impact of irrelevant or redundant features can be more pronounced.

Figure 4 presents a comparison of F1 scores for various feature selection methods across different feature counts, including MRMR [43], EMRMR (proposed), [44], [45], and [19]. The results underscore the exceptional performance of the EMRMR method, particularly at lower feature counts. At a feature count of 5, EMRMR achieves an outstanding F1 score of 0.99, significantly higher than MRMR (0.69), [44] (0.87), [45] (0.90), and [19] (0.88). With 10 features, EMRMR maintains a high F1 score of 0.88, which is slightly higher than MRMR (0.86) and comparable to [44] (0.84), [45] (0.83), and [19] (0.86). For feature counts of 15 and 20, EMRMR continues to demonstrate strong performance with F1 scores of 0.87 and 0.89, respectively, surpassing MRMR (0.81 and 0.80) and remaining competitive with other methods such as [44] (0.85 and 0.86), [45] (0.83 and 0.87), and [19] (0.85 and 0.86). With 25 features, EMRMR achieves an F1 score of 0.86, similar to MRMR (0.84) and in line with [44] (0.84), [45] (0.84), and [19] (0.83). At 30 features, EMRMR maintains a high F1 score of 0.89, significantly outperforming MRMR (0.40) and demonstrating superior performance compared to [44] (0.75), [45] (0.77), and [19] (0.76).

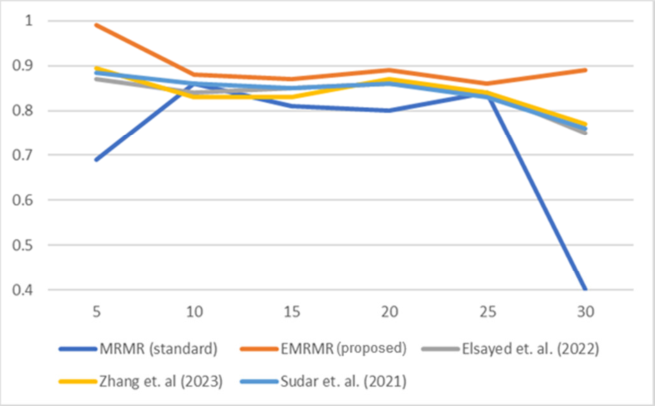

Fig. 4. F1 score achieved by various methods across feature counts.

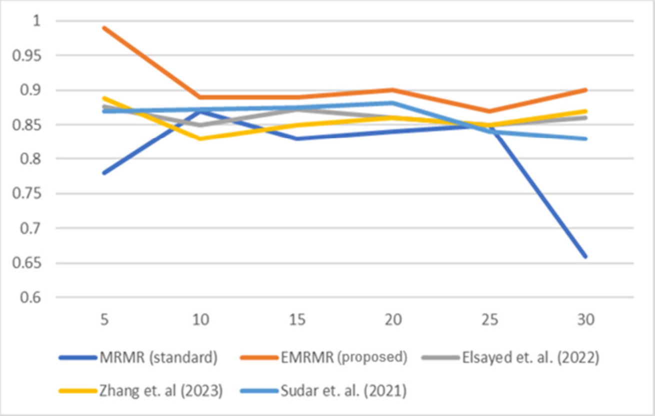

Fig. 5. Accuracy achieved by various methods across feature counts.

Figure 5 compares the accuracy of various feature selection methods across different feature counts, including MRMR [43], EMRMR (proposed), [44], [45], [19]. The results clearly highlight the superior performance of the EMRMR method, particularly at lower feature counts. At a feature count of 5, EMRMR achieves an exceptional accuracy of 0.99, significantly higher than MRMR (0.78), [44] (0.88), [45] (0.89), and [19] (0.87). With 10 features, EMRMR maintains a high accuracy of 0.89, slightly higher than MRMR (0.87) and comparable to [44] (0.85), [45] (0.83), and [19] (0.87). For feature counts of 15 and 20, EMRMR continues to demonstrate strong performance with accuracies of 0.89 and 0.90 respectively, surpassing MRMR (0.83 and 0.84) and remaining competitive with [44] (0.87 and 0.86), [41] (0.85 and 0.86), and [21] (0.88). With 25 features, EMRMR achieves an accuracy of 0.87, similar to MRMR (0.85) and in line with [44] (0.85), [45] (0.85), and [19] (0.84). At 30 features, EMRMR maintains a high accuracy of 0.90, significantly outperforming MRMR (0.66) and demonstrating superior performance compared to [44] (0.86), [45] (0.87), and [19] (0.83).

This comparison underscores EMRMR's consistent and good accuracy across various feature counts. The significant improvement observed with EMRMR, particularly at lower feature counts, can be attributed to its enhanced mechanism, which effectively balances feature relevance and redundancy. By selecting features that provide maximum discriminatory power while minimizing redundancy, EMRMR enhances the overall accuracy of the model. Its robustness and effectiveness in feature selection make EMRMR a highly reliable method to improve model performance in diverse applications. EMRMR enhances the standard MRMR approach by improving the way feature relevance and redundancy are estimated. This refined evaluation process ensures that the selected features offer maximum discriminatory power while minimizing redundancy, leading to more accurate and comprehensive feature selection. This results in improved model accuracy, showcasing the robustness and effectiveness of EMRMR in various applications. The consistently high performance of EMRMR across different feature counts underscores its reliability and superiority as a feature selection method, making it an excellent choice for enhancing ML model accuracy.

V. CONCLUSION

The results of this study highlight the substantial improvements achieved by the EMRMR technique in the context of intrusion detection for SDN. By integrating advanced feature selection strategies, specifically MIFS and CRCU, EMRMR effectively balances feature relevance and redundancy. This results in superior performance metrics compared to traditional methods such as MRMR, as well as other contemporary techniques. The EMRMR mechanism allows it to maintain high precision, recall, F1-score, and accuracy, particularly at lower feature counts, demonstrating its robustness in selecting the most relevant features without compromising accuracy. Its ability to reduce computational burden and enhance detection rates makes it suitable for real-time and lightweight IDSs. The significant improvements observed with EMRMR can be attributed to its effective evaluation of feature interactions, ensuring maximum discriminatory power while minimizing redundancy. This study confirms that employing contextual feature estimation in MIFS can improve IDS accuracy, which in turn enhances network security. Future work may explore the integration of EMRMR with other ML models and its application to various cybersecurity challenges, further validating its effectiveness and broadening its applicability in the field. Future works should involve extending the EMRMR approach by considering adaptive weighting mechanisms that can help further improve real-time detection in SDN environments. Testing the framework with more extensive and various datasets will be pursued to enhance generalizability. Additionally, integration of the EMRMR approach with other machine learning models will be studied to further improve IDS performance.

REFERENCES

[1] Y. Maleh, Y. Qasmaoui, K. El Gholami, Y. Sadqi, and S. Mounir, "A comprehensive survey on SDN security: threats, mitigations, and future directions," *Journal of Reliable Intelligent Environments*, vol. 9, no. 2, pp. 201–239, Jun. 2023, https://doi.org/10.1007/s40860-022-00171-8.

[2] S. Mehraban and R. K. Yadav, "Traffic engineering and quality of service in hybrid software defined networks," *China Communications*, vol. 21, no. 2, pp. 96–121, Oct. 2024, https://doi.org/10.23919/JCC.fa.2022-0860.202402.

[3] A. A. Bahashwan, M. Anbar, S. Manickam, T. A. Al-Amiedy, M. A. Aladaileh, and I. H. Hasbullah, "A Systematic Literature Review on Machine Learning and Deep Learning Approaches for Detecting DDoS Attacks in Software-Defined Networking," *Sensors*, vol. 23, no. 9, Jan. 2023, Art. no. 4441, https://doi.org/10.3390/s23094441.

[4] A. A. Najar and S. Manohar Naik, "Cyber-Secure SDN: A CNN-Based Approach for Efficient Detection and Mitigation of DDoS attacks," *Computers & Security*, vol. 139, Apr. 2024, Art. no. 103716, https://doi.org/10.1016/j.cose.2024.103716.

[5] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm," *IEEE Transactions on Computers*, vol. 65, no. 10, pp. 2986–2998, Jul. 2016, https://doi.org/10.1109/TC.2016.2519914.

[6] R. Chaganti, W. Suliman, V. Ravi, and A. Dua, "Deep Learning Approach for SDN-Enabled Intrusion Detection System in IoT Networks," *Information*, vol. 14, no. 1, Jan. 2023, Art. no. 41, https://doi.org/10.3390/info14010041.

[7] J. C. Correa Chica, J. C. Imbachi, and J. F. Botero Vega, "Security in SDN: A comprehensive survey," *Journal of Network and Computer Applications*, 2020, https://doi.org/10.1016/j.jnca.2020.102595.

[8] A. D. R. L. Ribeiro, R. Y. C. Santos, and A. C. A. Nascimento, "Anomaly Detection Technique for Intrusion Detection in SDN Environment using Continuous Data Stream Machine Learning Algorithms," in *2021 IEEE International Systems Conference (SysCon)*, Vancouver, Canada, Apr. 2021, pp. 1–7, https://doi.org/10.1109/SysCon48628.2021.9447092.

[9] S. Zavrak and M. Iskefiyeli, "Flow-based intrusion detection on software-defined networks: a multivariate time series anomaly detection approach," *Neural Computing and Applications*, vol. 35, no. 16, pp. 12175–12193, Jun. 2023, https://doi.org/10.1007/s00521-023-08376-5.

[10] F. Amiri, M. Rezaei Yousefi, C. Lucas, A. Shakery, and N. Yazdani, "Mutual information-based feature selection for intrusion detection systems," *Journal of Network and Computer Applications*, vol. 34, no. 4, pp. 1184–1199, Jul. 2011, https://doi.org/10.1016/j.jnca.2011.01.002.

[11] M. Bennasar, Y. Hicks, and R. Setchi, "Feature selection using Joint Mutual Information Maximisation," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8520–8532, Dec. 2015, https://doi.org/10.1016/j.eswa.2015.07.007.

[12] Z. Chkirbene, A. Erbad, R. Hamila, A. Mohamed, M. Guizani, and M. Hamdi, "TIDCS: A Dynamic Intrusion Detection and Classification System Based Feature Selection," *IEEE Access*, vol. 8, pp. 95864–95877, 2020, https://doi.org/10.1109/ACCESS.2020.2994931.

[13] G. Farahani, "Feature Selection Based on Cross-Correlation for the Intrusion Detection System," *Security and Communication Networks*, vol. 2020, no. 1, 2020, Art. no. 8875404, https://doi.org/10.1155/2020/8875404.

[14] D. Kshirsagar and S. Kumar, "Towards an intrusion detection system for detecting web attacks based on an ensemble of filter feature selection techniques," *Cyber-Physical Systems*, vol. 9, no. 3, pp. 244–259, Jul. 2023, https://doi.org/10.1080/23335777.2021.2023651.

[15] J. Maldonado, M. C. Riff, and B. Neveu, "A review of recent approaches on wrapper feature selection for intrusion detection," *Expert Systems with Applications*, vol. 198, Jul. 2022, Art. no. 116822, https://doi.org/10.1016/j.eswa.2022.116822.

[16] M. A. Siddiqi and W. Pak, "Optimizing Filter-Based Feature Selection Method Flow for Intrusion Detection System," *Electronics*, vol. 9, no. 12, Dec. 2020, Art. no. 2114, https://doi.org/10.3390/electronics9122114.

[17] Z. Ling and Z. J. Hao, "An Intrusion Detection System Based on Normalized Mutual Information Antibodies Feature Selection and Adaptive Quantum Artificial Immune System," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 18, no. 1, pp. 1–25, Jan. 2022, https://doi.org/10.4018/IJSWIS.308469.

[18] N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad, "Survey on SDN based network intrusion detection system using machine learning approaches," *Peer-to-Peer Networking and Applications*, vol. 12, no. 2, pp. 493–501, Mar. 2019, https://doi.org/10.1007/s12083-017-0630-0.

[19] K. Muthamil Sudar and P. Deepalakshmi, "An intelligent flow-based and signature-based IDS for SDNs using ensemble feature selection and a multi-layer machine learning-based classifier," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 3, pp. 4237–4256, Jan. 2021, https://doi.org/10.3233/JIFS-200850.

[20] M. S. Elsayed, N. A. Le-Khac, and A. D. Jurcut, "InSDN: A Novel SDN Intrusion Dataset," *IEEE Access*, vol. 8, pp. 165263–165284, 2020, https://doi.org/10.1109/ACCESS.2020.3022633.

[21] T. Linhares, A. Patel, A. L. Barros, and M. Fernandez, "SDNTruth: Innovative DDoS Detection Scheme for Software-Defined Networks (SDN)," *Journal of Network and Systems Management*, vol. 31, no. 3, Jun. 2023, Art. no. 55, https://doi.org/10.1007/s10922-023-09741-4.

[22] J. Buzzio-García *et al.*, "Exploring Traffic Patterns Through Network Programmability: Introducing SDNFLow, a Comprehensive OpenFlow-Based Statistics Dataset for Attack Detection," *IEEE Access*, vol. 12, pp. 42163–42180, 2024, https://doi.org/10.1109/ACCESS.2024.3378271.

[23] G. A. N. Segura, A. Chorti, and C. B. Margi, "Centralized and Distributed Intrusion Detection for Resource-Constrained Wireless SDN Networks," *IEEE Internet of Things Journal*, vol. 9, no. 10, pp. 7746–7758, Feb. 2022, https://doi.org/10.1109/JIOT.2021.3114270.

[24] A. M. El-Shamy, N. A. El-Fishawy, G. Attiya, and M. A. A. Mohamed, "Anomaly Detection and Bottleneck Identification of The Distributed Application in Cloud Data Center using Software–Defined Networking," *Egyptian Informatics Journal*, vol. 22, no. 4, pp. 417–432, Dec. 2021, https://doi.org/10.1016/j.eij.2021.01.001.

[25] M. H. H. Khairi, S. H. S. Ariffin, N. M. A. Latiff, A. S. Abdullah, and M. K. Hassan, "A Review of Anomaly Detection Techniques and Distributed Denial of Service (DDoS) on Software Defined Network (SDN)," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2724–2730, Apr. 2018, https://doi.org/10.48084/etasr.1840.

[26] T. Jafarian, M. Masdari, A. Ghaffari, and K. Majidzadeh, "SADM-SDNC: security anomaly detection and mitigation in software-defined networking using C-support vector classification," *Computing*, vol. 103, no. 4, pp. 641–673, Apr. 2021, https://doi.org/10.1007/s00607-020-00866-x.

[27] J. Li, M. S. Othman, H. Chen, and L. M. Yusuf, "Optimizing IoT intrusion detection system: feature selection versus feature extraction in machine learning," *Journal of Big Data*, vol. 11, no. 1, Feb. 2024, Art. no. 36, https://doi.org/10.1186/s40537-024-00892-y.

[28] W. F. Urmi *et al.*, "A stacked ensemble approach to detect cyber attacks based on feature selection techniques," *International Journal of Cognitive Computing in Engineering*, vol. 5, pp. 316–331, Jan. 2024, https://doi.org/10.1016/j.ijcce.2024.07.005.

[29] S. Chakraborty, A. K. Turuk, and B. Sahoo, "Federated Learning enabled software-defined optical network with intelligent control plane architecture," *Computers and Electrical Engineering*, vol. 118, Aug. 2024, Art. no. 109329, https://doi.org/10.1016/j.compeleceng.2024.109329.

[30] S. Yu *et al.*, "FreeEM: Uncovering Parallel Memory EMR Covert Communication in Volatile Environments," in *Proceedings of the 22nd Annual International Conference on Mobile Systems, Applications and Services*, Tokyo, Japan, Jun. 2024, pp. 372–384, https://doi.org/10.1145/3643832.3661870.

[31] P. Rajesh Kanna and P. Santhi, "Exploring the landscape of network security: a comparative analysis of attack detection strategies," *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. 8, pp. 3211–3228, Aug. 2024, https://doi.org/10.1007/s12652-024-04794-y.

[32] M. A. Khadse and D. M. Dakhane, "A Review on Network Covert Channel Construction and Attack Detection," *Concurrency and Computation: Practice and Experience*, Art. no. e8316, https://doi.org/10.1002/cpe.8316.

[33] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-Defined Networking: A Comprehensive Survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, Jan. 2015, https://doi.org/10.1109/JPROC.2014.2371999.

[34] J. Li *et al.*, "Feature Selection: A Data Perspective," *ACM Computing Surveys*, vol. 50, no. 6, pp. 1–45, Nov. 2018, https://doi.org/10.1145/3136625.

[35] Q. Al-Tashi, S. J. Abdulkadir, H. M. Rais, S. Mirjalili, and H. Alhussian, "Approaches to Multi-Objective Feature Selection: A Systematic Literature Review," *IEEE Access*, vol. 8, pp. 125076–125096, 2020, https://doi.org/10.1109/ACCESS.2020.3007291.

[36] M. Labani, P. Moradi, and M. Jalili, "A multi-objective genetic algorithm for text feature selection using the relative discriminative criterion," *Expert Systems with Applications*, vol. 149, Jul. 2020, Art. no. 113276, https://doi.org/10.1016/j.eswa.2020.113276.

[37] R. Gandhi, U. Ghose, and H. K. Thakur, "Revisiting Feature Ranking Methods using Information-Centric and Evolutionary Approaches: Survey," *International Journal of Sensors Wireless Communications and Control*, vol. 12, no. 1, pp. 5–18, Jan. 2022, https://doi.org/10.2174/2210327911666210204142857.

[38] N. Hoque, H. A. Ahmed, D. K. Bhattacharyya, and J. K. Kalita, "A Fuzzy Mutual Information-based Feature Selection Method for Classification," *Fuzzy Information and Engineering*, vol. 8, no. 3, pp. 355–384, Sep. 2016, https://doi.org/10.1016/j.fiae.2016.09.004.

[39] P. A. Estevez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized Mutual Information Feature Selection," *IEEE Transactions on Neural Networks*, vol. 20, no. 2, pp. 189–201, Oct. 2009, https://doi.org/10.1109/TNN.2008.2005601.

[40] C. Suman, S. Tripathy, and S. Saha, "Building an Effective Intrusion Detection System using Unsupervised Feature Selection in Multi-objective Optimization Framework." arXiv, May 16, 2019, https://doi.org/10.48550/arXiv.1905.06562.

[41] B. A. Manjunatha, P. Gogoi, and M. T. Akkalappa, "Data Mining based Framework for Effective Intrusion Detection using Hybrid Feature Selection Approach," *International Journal of Computer Network and Information Security*, vol. 11, no. 8, pp. 1–12, Aug. 2019, https://doi.org/10.5815/ijcnis.2019.08.01.

[42] A. V. Turukmane and R. Devendiran, "M-MultiSVM: An efficient feature selection assisted network intrusion detection system using machine learning," *Computers & Security*, vol. 137, Feb. 2024, Art. no. 103587, https://doi.org/10.1016/j.cose.2023.103587.

[43] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, Dec. 2005, https://doi.org/10.1109/TPAMI.2005.159.

[44] M. S. E. Sayed, N. A. Le-Khac, M. A. Azer, and A. D. Jurcut, "A Flow-Based Anomaly Detection Approach With Feature Selection Method Against DDoS Attacks in SDNs," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 4, pp. 1862–1880, Sep. 2022, https://doi.org/10.1109/TCCN.2022.3186331.

[45] L. Zhang, K. Liu, X. Xie, W. Bai, B. Wu, and P. Dong, "A data-driven network intrusion detection system using feature selection and deep learning," *Journal of Information Security and Applications*, vol. 78, Nov. 2023, Art. no. 103606, https://doi.org/10.1016/j.jisa.2023.103606.