IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 17 August 2023, accepted 4 September 2023, date of publication 11 September 2023,
date of current version 14 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3313630

RESEARCH ARTICLE

Abnormal Transactions Detection in the Ethereum Network Using Semi-Supervised Generative Adversarial Networks

YOUSEF K. SANJALAWE[1], AND SALAM R. AL-E'MARI[2]
[1]Cybersecurity Department, School of Information Technology, American University of Madaba, Amman 11821, Jordan
[2]Information Security Department, Faculty of Information Technology, University of Petra, Amman 11196, Jordan

Corresponding author: Salam R. Al-E'mari (salam.ammari@uop.edu.jo)

ABSTRACT Numerous abnormal transactions have been exposed as a result of targeted attacks on Ethereum, such as the Ethereum Decentralized Autonomous Organization attack. Exploiting vulnerabilities in smart contracts, malicious users can pursue their own illicit objectives through abnormal transactions. Consequently, identifying these malevolent users, implicated in fraudulent activities and their attribution, becomes exceedingly complex. Cryptocurrency transactions used for malicious purposes, employing pseudo-anonymous accounts to send and receive ransom payments and accumulating funds under various identities, further highlight the need to control and detect these abnormal transactions for maintaining a high level of security within the Ethereum network. Although existing Intrusion Detection Systems (IDSs) help mitigate abnormal transaction occurrences, their performance necessitates improvement. To address this issue, this study presents a novel approach, named Abnormal Transactions Detection Using a Semi-Supervised Generative Adversarial Network (ATD-SGAN), which efficiently detects abnormal attacks within the Ethereum network. ATD-SGAN leverages a semi-supervised generative adversarial network for this purpose. The results demonstrate that ATD-SGAN significantly enhances the performance of state-of-the-art IDSs. It achieves an increase in detection accuracy from 3.78% to 11.05% and reduces the false alarm rate from 42.29% to 0.15%. Moreover, ATD-SGAN notably improves the F1-measure, ranging from 10.39% to 3.79%, compared to the current IDSs.

INDEX TERMS Abnormal transactions, ethereum, feature selection, intrusion detection system, network security.

I. INTRODUCTION

Illicit activities, including money laundering, phishing, and fraud, have cast a shadow over the advancements made in cryptocurrencies and the accompanying advantages they offer, as highlighted in a study on detecting such activities on the blockchain network [1]. Due to the substantial volume of sensitive data they handle, these technologies are vulnerable to a range of malicious actions, attacks, and security threats that pose risks to the availability and integrity of information and services. Unlawful behaviors have had a substantial impact on financial systems like Ethereum, introducing unprecedented challenges. The pseudonymous nature of blockchain networks allows criminals to conceal their true identities, making it an attractive feature for carrying out abnormal or malicious activities.

Furthermore, the increasing reliance on the Ethereum network for various aspects of our lives, including cryptocurrencies and decentralized apps (DApps), has resulted in a surge in Ethereum transactions. However, this has also captured the attention of attackers who exploit vulnerabilities in Ethereum contracts, transactions, and the Ethereum Virtual Machine (EVM) to devise a range of attack methods aimed at stealing Ethereum or disrupting the Ethereum market [2]. It is important to note that Ethereum, which is built on blockchain technology, comprises two types of transactions. The first type involves external transactions used for cryptocurrency exchanges, while the second type entails internal transactions

The associate editor coordinating the review of this manuscript and approving it for publication was Kashif Saleem.

executed by smart contracts or DApps. Internal transactions require gas to execute, and this gas is acquired through external transactions using the Ether cryptocurrency [2], [3].

Although blockchain networks are secure, they are exposed to security vulnerabilities. Consequently, intruders have emerged in Ethereum networks and made thefts of millions of Ethers. For instance, a Decentralized Autonomous Organization (DAO) attack occurred in 2016 and over $50M was stolen [4], [5]. In addition, $13M of Ether was stolen by a parity multi-sig wallet attack in July 2017 and a new version from this attack stole $155M of Ether in November 2017 [6]. Further, in 2018 integer flow attack stole $2.3 M of Ether [7]. While $48.7 M of Ether was stolen by an unknown address account in South Korea through a cryptocurrency exchange [8], and $48.7 M of Ether was stolen by a 51% attack in 2020 [9]. Besides that, several attacks attempted to steal cryptocurrencies from the Ethereum network or other malicious actions. All the above-mentioned attacks generate a huge number of abnormal transactions, therefore; detection of these abnormal transactions led to detecting the attacks that target the Ethereum network. On the other hand, the conventional IDS are unable to detect abnormal transactions because the Ethereum network has a new complex environment and infrastructure. Therefore, it is essential to propose an IDS approach mainly to detect abnormal transactions in the Ethereum network, the security, detection, and protection of the various communication infrastructures using Intrusion Detection Systems (IDSs) are of critical importance.

To this end, IDS is a security tool that monitors network traffic for signs of cyber-attacks or malicious activity. It can be used to detect and prevent attacks on Ethereum networks, as well as to identify and alert on suspicious activity. There are several types of IDS that can be used for Ethereum, including network-based, host-based, and wireless IDS. The importance of an IDS in Ethereum lies in its ability to provide an additional layer of security to protect against attacks and malicious activity. By continuously monitoring network traffic and alerting on suspicious activity, an IDS can help to identify and prevent potential threats before they can cause harm [10]. Generative Adversarial Networks (GANs) comprise a pair of neural networks, namely the generator and the discriminator, which collaborate in the identification and classification of network data. The generator generates synthetic data to train the discriminator, whose role is to discern between real and synthetic data. As the generator and discriminator undergo training, the generator becomes proficient in generating lifelike synthetic data, while the discriminator becomes adept at distinguishing between real and synthetic data. GANs can enhance the precision of an IDS by augmenting the volume of available training data and aiding the IDS in generalization, enabling the detection of novel attack types [11], [12].

However, there is still a need to improve the performance of the existing IDS in terms of detection accuracy and reduce the false-positive rate as these approaches do not evaluate on a real dataset and are usually evaluated using only self-prepared datasets. The characteristics of self-prepared datasets in terms of attack coverage, accuracy, and validity are not revealed. Therefore, it is imperative to develop resilient IDSs that can improve detection accuracy, decrease the false alarm rate, and enhance the detection rate when identifying anomalous transactions within the Ethereum network. This paper aims to accomplish the following objectives:

- To adopt a multi-digraph theory to extract a set of features for the Ethereum transactions.
- Proposed multi-objective function to reduce the dimensionality of the dataset and improve detection performance.
- Proposed ensemble feature selection mechanism to select the most significant features that contribute to detecting abnormal transactions in the Ethereum network efficiently.
- To adapt automatic data augmentation mechanism to avoid overfitting and achieve impressive detection performance from few labeled transactions used in training
- To evaluate ATD-SGAN approach

A. PAPER ORGANIZATION

The structure of this paper is as follows: Section II provides an overview of the related works on intrusion detection in the Ethereum network; Section III introduces the proposed IDS approach; Section IV presents the experimental results; Section V discusses the outcomes; and finally, Section VI concludes the paper.

II. RELATED WORKS

Several research studies have been undertaken to identify abnormal transactions using blockchain networks. Furthermore, the present work utilizes a learning model based on the anomaly detection approach. In contrast, machine learning and deep learning techniques can aid IDSs in automatically detecting both new and existing attacks without the need for human intervention by optimizing feature selection. In recent times, numerous machine learning and deep learning algorithms, such as support vector machines and artificial neural networks, have been incorporated into IDSs to bolster system security. Moreover, a Convolutional Neural Network (CNN) incorporating a self-attention mechanism has been employed to construct the ABCNN (Attention-based Convolutional Neural Network) model for the purpose of identifying vulnerabilities in smart contracts. The ABCNN model utilized a self-curated dataset by manually collecting 8632 verified smart contracts from Etherscan. This dataset was prepared to facilitate the training and evaluation process. The ABCNN model demonstrated superior performance with a reduced missing rate and faster execution time. Additionally, it successfully detected three types of attacks, namely: (i) Reentrancy, (ii) Arithmetic issues, and (iii) Time manipulation [13].

The ESCORT model employed deep learning networks and Transfer Learning (TL) to effectively identify both known

and unknown vulnerabilities, addressing the scalability and generalization limitations present in previous research efforts. ESCORT utilizes a multi-output neural network architecture comprising two main components: (i) A shared feature extractor that learns the semantics of the input smart contract, and (ii) Multiple branch structures where each branch focuses on learning a specific vulnerability type using the extracted features from the feature extractor. The research paper provides a thorough assessment of ESCORT's performance on different smart contracts, successfully detecting six vulnerability types as well as identifying new vulnerability types through the application of TL. ESCORT achieved a better detection accuracy rate in the empirical results [14]. In addition, a new model to detect abnormal transactions in the Bitcoin network is proposed by utilizing the K-Nearest Neighbours (KNN) algorithm and by testing the model on the Elliptic dataset which has 203,769 nodes and 234,355 edges. In addition, the Elliptic dataset classifies the data into three categories illicit, licit, or unknown. On the other hand, the Elliptic dataset has 166 features for each node where 94 features represent local information about the transaction and 72 features are called aggregated features. However, the proposed approach has higher accuracy, but the rate of detection and precision is not efficient [15].

Besides, a supervised machine learning-based anomaly detection method was used, in [16], to identify malicious nodes by analyzing the transaction behavior of accounts. Supervised machine learning models were applied to two different types of accounts: Externally Owned Accounts (EOA) and smart contract accounts. These models achieved a detection accuracy of 96.54% with a false-positive ratio of 0.92% for EOA accounts and 96.82% with a false-positive ratio of 0.78% for smart contract accounts. During the period from 20 January 2020 to 24 February 2020, the method identified 85 new malicious EOA and 1 malicious smart contract address. When tested on these addresses, the model's accuracy was 96.21% with a false-positive ratio of 3%. Moreover, the authors proposed a framework to detect abnormal entities in the Ethereum network through several machine learning methods: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Stacking, and AdaBoost in [1]. First, the dataset was gathered from Etherscan.io, and then all instances of this dataset were labeled by Ethereum community members (i.e. experts). After that, the re-sampling technique was used to handle the nature of the imbalanced dataset. As a result, the proposed framework achieves high performance in the classification of the Ethereum entities for the Stacking, and AdaBoost learning methods. Furthermore, a fraud detection model was proposed to identify illicit accounts on the Ethereum blockchain in [17]. The model utilized three machine learning algorithms: decision tree (j48), Random Forest (RF), and KNN. A dataset comprising 42 features was obtained from Kaggle.com and subsequently, the correlation coefficient was employed to select the most impact features. A new dataset was then constructed, containing only 6 selected features. The experimental results demonstrated significant enhancements in time measurements across all three algorithms, while the Random Forest algorithm exhibited improved performance in terms of the F-measure.

Within the framework of the escalating adoption of Ethereum and the subsequent proliferation of smart contract-driven decentralized applications (DApps), the frequency of malicious attacks targeting this ecosystem has surged. Notably, frontrunning attacks exploit transaction latency within the pending pool by manipulating gas prices, thereby posing a serious threat to DApp security. Thus, the authors in [18] proposed a model-based defense mechanism based on Multi-Layer Perceptron (MLP). The proposed model aims to discern whether a transaction exhibits indicators of a frontrunning attack. By involved the extraction of transaction-specific features, which are then transformed into feature vectors for real-time analysis, and extensive experiments on a comprehensive transaction dataset. In addition, the study in [19] introduced a model that combines Generative Adversarial Networks (GAN) and Deep Recurrent Neural Networks (RNN) for cyber threat identification within the Ethereum blockchain. The proposed model follows a two-phase approach, the first phase of the model utilized GAN to produce fake transactions by leveraging genuine Ethereum transactions as a foundation. Subsequently, the second phase employed a bi-directional Long Short-Term Memory (LSTM) mechanism to detect adversarial transactions during a cyber threat hunting process. As a result, the model achieved in the first phase an accuracy of 82.51% in generating transactions closely resembling authentic Ethereum transactions. In the second phase, the model demonstrated high performance with a 99.98% accuracy rate in identifying adversarial transactions. Furthermore, in [20], a method is introduced to detect fraudulent activities within the Ethereum blockchain through the analysis of transaction records. This approach entails the use of web crawlers to gather labeled fraudulent addresses. Subsequently, a transaction network is constructed using the available public transaction ledger. For the purpose of identifying fraudulent transactions, a specialized algorithm based on network embedding is employed. This algorithm is tailored for networks structured by transaction amounts. It extracts features from the nodes in the network. Notably, the study [20] adopts a Graph Convolutional Network (GCN) model for the classification of addresses into legitimate or fraudulent categories. The experimental results showcase an impressive accuracy of 95%, underscoring the system's effectiveness in pinpointing fraudulent transactions within the Ethereum blockchain. Despite the challenge posed by an unlabeled dataset for evaluating the approach's performance, the trimmed k-means algorithm successfully identified known instances of anomalies.

On the other hand, the arena of mitigating Distributed Denial-of-Service (DDoS) attacks has become a focal point for extensive research endeavors. Concurrently, emerging

technologies, with blockchain at the forefront, present promising avenues for groundbreaking solutions. In [21], the authors introduced Cochain-SC, an inventive approach anchored in blockchain technology. Cochain-SC pioneers a two-tiered mitigation framework that encompasses both intra-domain and inter-domain scenarios of DDoS attacks. Harnessing the capabilities of software-defined networks (SDN) in conjunction with the secure decentralization facilitated by blockchain, Cochain-SC devises a pioneering strategy that amalgamates these technologies to achieve robust, collaborative, and effective mitigation outcomes. This entails the fusion of SDN-based intra-domain mechanisms responsible for classifying and mitigating flows with blockchain-enabled inter-domain cooperation facilitated by smart contracts. In addition, Co-IoT is a novel blockchain-based framework designed for collaborative DDoS mitigation. By leveraging the capabilities of SDN and blockchain technology, particularly Ethereum's smart contracts, Co-IoT aspires to foster collaborative efforts among SDN-based domains. The framework's decentralized approach facilitates the exchange of attack-related information, aiming to enhance flexibility, efficiency, security, and cost-effectiveness in combating large-scale DDoS attacks [22]. Furthermore, The field of safeguarding blockchain nodes faces a significant breakthrough with the introduction of BrainChain, an innovative and scalable solution designed to counteract the most extensive DDoS attack witnessed, specifically the Domain Name System (DNS) amplification attack. BrainChain is meticulously crafted within SDN to protect and enhance the resilience of blockchain nodes. This scheme is composed of four pivotal components namely: (i) The Flow Statistics Collection scheme (FS), (ii) The Entropy-Based scheme (ES), (iii) The Bayes Network-Based Filtering scheme (BF), and (iv) The DNS Mitigation scheme (DM). where the empirical assessment affirms the formidable capabilities of BrainChain in promptly and accurately identifying and countering DNS amplification attacks [23].

Despite the numerous research studies conducted to detect abnormal transactions in two prominent blockchain networks, Bitcoin and Ethereum, these studies continue to face certain common challenges. One such challenge is the absence of a definitive ground truth to evaluate the effectiveness of any proposed model. Additionally, there are multiple cybersecurity concerns across different layers that further complicate the detection process.

III. PROPOSED IDS APPROACH

This paper proposes an abnormal transactions-based detection approach, called ATD-SGAN, in the Ethereum network using SGAN. The proposed approach enhances the detection performance while detecting abnormal transactions in the Ethereum network. The aim of this approach is to detect abnormal transactions using a semi-supervised learning method and deep learning. The proposed approach consists of five main stages, namely: (i) Ethereum data gathering: which aims to propose gathering transactions and labeling the transactions, (ii) data pre-processing: which aims to increase the quality of the dataset by eliminating noisy data, (iii) feature extraction: aims to extract feature based multi-digraph theory, (iv) ensemble feature selection: aims to select a mutual feature from two bio-inspired algorithms and (v) abnormal transactions detection: to detect abnormal transaction utilized SGAN in Ethereum network as shown in Figure 1.

A. BLTE DATASET

Benchmark Labelled Transactions Ethereum (BLTE) is a benchmark dataset that gathers based on a real Ethereum network called Ethereum Classic (ETC) network [24], and it is a real chain, public, open-source, and distributed platform. ETC has many tables, but ATD-SGAN chooses the Transactions table with seventeen features as shown in Fig 5. These transactions are performed by EOA which deals with external transactions and records them on blockchain to exchange cryptocurrency transactions [25]. According to Figure 1 the first stage (Ethereum Dataset Generation), and second stage (pre-processing) have been implemented and discussed in former research minutely [24].

However, the seventeen features of the transaction table suffer from two main problems. The first problem is that these features, in their current form, do not contribute to the detection of abnormal transactions and need to be further analyzed to derive new features that contribute to the detection of abnormal transactions, while the second problem is that the transaction table is unable to be automatically labeled (i.e in case of the transaction does not exist in Etherscamdb). The first problem is tackled in the feature extraction stage (refer to Section III-C) while the second problem is tackled in the abnormal transactions' detection stage (refer to SectionIII-E). ATD-SGAN obtains abnormal transactions from Etherscamdb,[1] which is open-source and available on GitHub.[2]

B. DATA PREPROCESSING STAGE

Data pre-processing is a significant stage that can enhance the performance of intrusion detection systems [26]. The proposed model begins cleansing data by eliminating the irrelevant features including any feature that has a null value or the same value for all Ethereum transactions in BLTE. Consequently, thirteen out of seventeen features are confirmed in BLTE. Table 1 presents the description for each feature.

Therefore, the BLTE dataset holds great importance in evaluating detection systems that heavily depend on labeled data, specifically transactions. Consequently, within the BLTE dataset, every transaction has been categorized as either a normal or abnormal transaction. Each transaction involves two addresses: the sender and the receiver. Abnormal transactions are characterized by originating from intrusion

[1] https://etherscamdb.info
[2] https://github.com/MrLuit/EtherScamDB

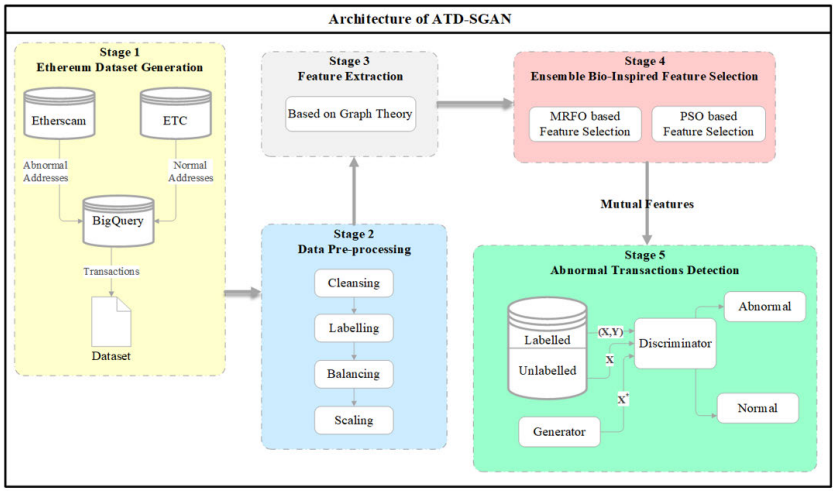

FIGURE 1. The main stages of ATD-SGAN.

TABLE 1. Transactions table.

Field	Description
Hash	transaction hash value
Nonce	TThe count of transactions executed by the sender's account
Transaction index	The position of the transaction within a block
From address	Source account
To address	Target account
Value	The transferred value in Wei, which represents the smallest unit of Ether.
Gas	Amount of gas by source
Gas price	The gas price in Wei, as specified by the source
Receipt cumulative gas used	The gas consumption of this transaction when executed within a block
Receipt gas used	The cumulative gas usage of this particular transaction
Block timestamp	The timestamp associated with the block utilized by this transaction
Block number	Block number of the transaction
Block hash	Hash of the block used by a transaction
From scam	A value of 1 denotes that the sender address is identified as a scam, while a value of 0 indicates a normal address
To scam	A value of 1 signifies that the receiver address is classified as a scam, whereas a value of 0 represents a normal address
Scam	A value of 1 signifies that the transaction is considered abnormal, while a value of 0 indicates a normal transaction

TABLE 2. Statistics of the BLTE dataset.

Trx Type	No. Trx	Ratio
Abnormal	14145	46%
Normal	16605	54%
Total	30750	100%

source addresses, being targeted addresses, or involving both. Let training data be (T) = (I, Trx, S, R), where I, Trx, S, and R are a transaction Id, transaction, sender, and receiver, respectively. Further, Trx $\in$ {0,1} is a binary classification, 0 indicates normal Trx and 1 indicates abnormal Trx. Eventually, BLTE has sixteen features, thirteen after the cleansing step, and inserting three features from the labeling step that is from scam to scam and scam. Then, balancing data is a significant concern when preparing a dataset to increase the classification accuracy of the model. The reduction method is one approach to processing imbalanced data [27].

The ATD-SGAN approach leverages the instance selection technique to reduce the count of a specific class in the training data since the generated dataset in BLTE contains a lower count of abnormal transactions as compared to normal transactions; therefore, the instance selection is used to reduce the number of normal transactions as it does not affect the model performance [28]. BLTE reduced the size of normal transactions for compatible abnormal transactions in the dataset by Local Density-based Instance Selection (LDIS) and Table 2. summarizes information on the total number of Ethereum transactions in the BLTE.

Ultimately, the scaling step refers to converting values of attributes in a dataset in a specific range. The two main methods of scaling are standardization and normalization. Standardization transforms attribute values based on Gaussian distribution, while normalization transforms attribute values to a common scale with a specific range. Whereas machine learning algorithm always benefits from the normalization method to convert the values in a dataset without distorting variation in its range [29].

There are various normalization methods such as Min-Max, Z-score, and so on. ATD-SGAN applies the common one in a Min-Max normalization due to its enhanced speed learning model, which scales data between 0 and 1 according to Equation (1) where a symbol X is a numerical value, X_{max}, X_{min} is the maximum and minimum values of the attribute, respectively. While $X_{norm} \in [0,1]$ is a new value for the

attribute [30]:

$$X_{norm} = \frac{X + X_{min}}{X_{max} - X_{min}} \quad (1)$$

However, each dataset consists of features (attributes), such as, in BLTE dataset, All degree, In degree, Out degree, Unique in degree, Unique out degree, Avg amount incoming, Avg amount outgoing, Total amount incoming, Total amount outgoing, Max amount incoming, and Total Amount (described in Table 3). In fact, feature engineering is the process of selecting and creating important features from raw data that can be used to train a machine learning model. In the context of an intrusion detection system, feature engineering is the process of identifying and creating features that can accurately classify malicious activity. This is important for ensuring that the IDS can effectively detect and respond to threats while minimizing false positives, or instances of normal activity being incorrectly classified as malicious. By carefully selecting and creating relevant features, it is possible to build an IDS that is effective at detecting and responding to real threats while minimizing disruptions to normal network activity [31], [32].

In machine learning, statistical analysis, and deep learning, in particular, feature selection is the process of selecting a subset of relevant features for use in model construction. The goal of feature selection is to select the most useful features in predicting the response while excluding irrelevant or redundant features. The chosen features should be able to effectively predict the response variable and should not be highly correlated with each other. Therefore, the following subsections discuss the feature extraction and selection used in this paper in detail.

C. FEATURE EXTRACTION STAGE

Feature extraction is a process to extract subset features from input data that improves the accuracy of learned models [33]. According to existing studies, several features have been extracted using weighted multi-digraph from Bitcoin and Ethereum networks as mentioned in Table 1. In the ATD-SGAN approach, 22 features have been extracted from sixteen features from BLTE based on multi-graph theory. Whereas the Ethereum network is a graph $G = (N, E)$, where (N) presents as a node of the Ethereum address, and (E) is a transaction edge that links between two Ethereum addresses if the edge has weight or value, then it is a weighted graph; otherwise, it unweighted graph. However, there are several types of graphs, but the Ethereum network is a weighted multi-digraph in which each node has multiple weighted edges from the source node to the target node, this paper used the Neo4j graph database to present and analyze the relationship (transactions) between Ethereum addresses (nodes) for BLTE dataset. Figure 2 demonstrates an example for one Ethereum node, where node 2 is a sender and other nodes are receivers. Furthermore, each edge has four essential values tuples (s, u, v, t) where s is the source node (sender), u is the target node (receiver), v is (value) of the transaction and t is (timestamp) of the transaction occurred [34], [35]. Indeed, multi-digraph allows multiple transactions between nodes which is necessary to record the information of the Ethereum node.

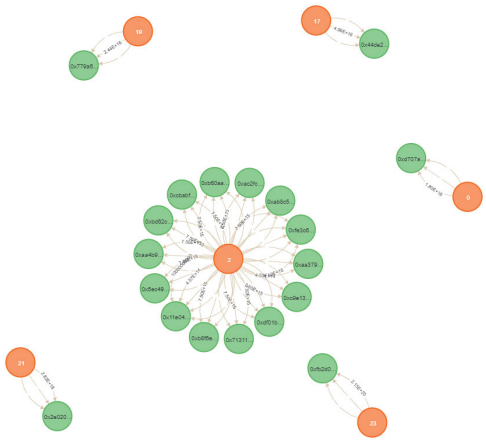

FIGURE 2. Sample of multi-digraph of Ethereum nodes.

Table 3 presents the results of feature extraction. However, the 22 features do not contribute to the detection of abnormal transactions. Thus, an ensemble Bio-Inspired feature selection mechanism has been proposed to select the most significant features that contribute to the detection accuracy of abnormal transactions. On the contrary, the existing approaches that ignore or select the features are based on simple heuristics.

D. ENSEMBLE BIO-INSPIRED FEATURE SELECTION STAGE

To enhance the performance of ATD-SGAN, it is crucial to eliminate irrelevant features after the process of feature extraction. This helps to reduce redundant data, improve the accuracy of the prediction model, and decrease the training phase duration [36]. Feature selection is employed by ATD-SGAN, utilizing two bio-inspired algorithms: (i) Manta Ray Foraging Optimization (MRFO), and (ii) Particular Swarm Optimization (PSO). Subsequently, ATD-SGAN identifies the common features from the outcomes of these two algorithms using a novel multi-objective function. The subsequent subsections provide a comprehensive explanation of the feature selection mechanism employed by ATD-SGAN.

1) FEATURE SELECTION BASED MRFO

MRFO is a bio-inspired meta-heuristic algorithm for feature subset selection, which is used to enhance the feature selection stage due to its effectiveness in feature selection, as well as it needs a smaller number of iterations and configuration settings to converge (i.e., reaching the optimum), which are major concerns in any feature selection problem [37], [38], [39], [40]. Thus, the MRFO is utilized to reduce the dimensionality of the dataset consequently decreasing the complexity of detection computation and enhancing the overall detection performance by

TABLE 3. Features code.

Id	Feature name	Description
0	All degree	The number of all transactions that were sent and received from other nodes.
1	In degree	The number of received transactions.
2	Out degree	The number of sent transactions.
3	Unique in degree	The number of received transactions from unique addresses.
4	Unique out degree	The number of sent transactions to unique addresses.
5	Avg amount incoming	The average amount received from other addresses.
6	Avg amount outgoing	The average amount that was sent to other nodes.
7	Total amount incoming	The average amount received from other addresses.
8	Total amount outgoing	The average amount was sent to other addresses.
9	Max amount incoming	The maximum amount of transaction has been received.
10	Total Amount	The total amount of transactions that have been sent and received.
11	Min amount incoming	The minimum amount of transaction has been received.
12	Min amount outgoing	The minimum amount of transaction has been sent.
13	Avg time incoming	The average time of received transactions
14	Avg time outgoing	The average time of sent transactions
15	Total amount outgoing	The total amount of time of sent transactions.
16	Active Duration	The time difference between the first and last transactions related to a given address.
17	Clustering coefficient	The measure of connectivity amongst neighbors of a given address.
18	Mean time interval	The mean interval time two transactions related to a given address
19	Max time interval	The maximum interval time for two transactions related to a given address.
20	Min time interval	The minimum between two transactions related to a given address.
21	Avg gas price	The average gas price of a given address.

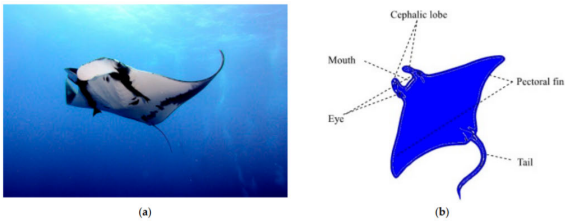

FIGURE 3. a) Manta Ray's body, (b) Manta Rays' physical construction [39].

avoiding irrelevant or duplicated features (if any). The MRFO algorithm was inspired by the feeding strategies of Manta rays, which are the largest deep-sea creatures, Figure 3 demonstrates the body of the Manta ray. The properties of Manta rays are body flat, swimming smoothly, largemouth, and plankton is the main food [39].

The MRFO algorithm draws inspiration from three foraging strategies observed in Manta rays: (i) Chain, (ii) Cyclone, and (iii) Somersault [41]. In the Chain feed strategy, Manta rays observe the position of plankton and swim toward it. The concentration of plankton in a particular position plays a crucial role, as a higher concentration signifies a better solution. The equations describing this strategy are presented in Equation (2) [41].

$$x_i^d(t+1) = \begin{cases} x_i^d(t) + r.(x_{best}^d(t) - x_i^d(t)) \\ +\alpha.(x_{best}^d(t) - x_i^d(t)), & \text{if 1} \\ x_i^d(t) + r.(x_{(i-1)}^d(t) - x_i^d(t)) \\ +\alpha.(x_{best}^d(t) - x_i^d(t)), & \text{i=2,...,N} \end{cases} \tag{2}$$

where $x_i^{d(t)}$ refers to the position of t^{th} individual at time t in d^{th} dimension. while r refers to a random vector within the range of [0, 1], and α is a weight coefficient as shown in Equation (3), $x_{best}^{d(t)}$ is the plankton with high condensation [41].

$$\alpha = 2.r.\sqrt{|\log(r)|} \tag{3}$$

Manta rays exhibit fascinating behavior when they detect a patch of plankton with a high concentration in deep water. They form a long foraging chain and swim in a spiral pattern toward the food, known as the cyclone feed strategy. The equations describing this strategy can be observed in Equation (4) and Equation (5) [40].

$$x_i^d(t+1) = \begin{cases} x_{best}^d(t) + r.(x_{best}^d(t) - x_i^d(t)) \\ +\beta.(x_best^d(t) - x_i^d(t)), & \text{if 1} \\ x_{best}^d(t) + r.(x_{i-1}^d(t) - x_i^d(t)) \\ +\beta.(x_best^d(t) - x_i^d(t)), & \text{i=2,...,N} \end{cases} \tag{4}$$

$$\beta = 2.e^{r_1 * \frac{T-t+1}{T}} \tag{5}$$

In the MRFO algorithm, the weight coefficient β is employed, while T represents the maximum number of iterations. The variable r_1 denotes a random number within the range $\in$ [0, 1]. Each individual within the algorithm performs a random search, utilizing the best-found plankton position as a reference point. Moreover, to enhance the exploration capability of the MRFO algorithm, Equation (6) and Equation (7) are employed. These equations compel the manta rays to search for new positions by assigning a random position as their reference point, thus enabling a more extensive global

search [40].

$$x_{rand}^{d} = Lb^{d} + r * (Ub^{d} - Lb^{d}) \tag{6}$$

$$x_i^d(t+1) = \begin{cases} x_{rand}^d(t) + r.(x_{rand}^d(t) - x_i^d(t)) \\ +\beta.(x_r and^d(t) - x_i^d(t)), & \text{if 1} \\ x_{rand}^d(t) + r.(x_(i-1)^d(t) - x_i^d(t)) \\ +\beta.(x_{rand}^d(t) - x_i^d(t)), & \text{i=2,...,N} \end{cases} \tag{7}$$

The variable $x_{\text{rand}}{}^{\text{d}}$ represents a random position within the search space, where Lb^{d} and Ub^{d} denote the lower and upper limits, respectively, of the d^{th} dimension. In the MRFO algorithm, the chain foraging strategy is employed if the random value exceeds 0.5. Conversely, if the random value is less than or equal to 0.5, the MRFO algorithm utilizes the cyclone foraging strategy as defined in Equation (3). The position is then updated to find the best solution according to Equation (7), while the somersault strategy, outlined in Equation (8), is utilized [42].

$$X_i^d(t+1)b = x_i^d(t)b + bs.(r_2.x_{best}^d - r_3.x_i^d(t)) \\ , i = 1, 2 \ldots, N \tag{8}$$

In the MRFO algorithm, the variable S represents the Somersault factor, while r_2 and r_3 denote random numbers within the range of [0, 1] [41]. The value of S is fixed at 2. The somersault feed strategy is characterized by random, frequent, localized, and cyclical movements, enabling manta rays to maximize their intake of plankton. This strategy involves utilizing the best-known plankton position as a pivot, and each swimmer swims back and forth around the pivot while somersaulting to reach new positions.

The effectiveness of the MRFO algorithm has been demonstrated in solving real-world engineering problems. It has been evaluated and compared with eight benchmark algorithms, showcasing superior performance in solving engineering problems. The MRFO algorithm has also been successfully applied in feature selection for S-shaped and V-shaped transfer functions. An evaluation was conducted on 18 UCI datasets, and the MRFO algorithm outperformed existing methods in terms of accuracy and the number of selected features. The results demonstrate the effectiveness of the MRFO algorithm compared to state-of-the-art methods in terms of accuracy and selected features [40].

2) FEATURE SELECTION BASED PSO

The PSO algorithm, originally proposed by James Kennedy and Russell Eberhart in 1995, draws inspiration from the collective behavior observed in bird and fish swarms [43], [44]. It aims to optimize problems by iteratively refining candidate solutions [45], [46]. PSO operates based on the concept of a global best solution, which is continually updated during each iteration to converge toward the optimal solution. To achieve optimal feature selection in BLET, PSO employs a fitness function that leads to improved feature selection, precision, and true negative rate. The algorithm initiates particles with random positions, evaluating their fitness at each position. These particles then update their positions and velocities based on historical data, aiming to converge toward the optimal position. This process is demonstrated by Equation (9) and Equation (10), which illustrate the updating of particle positions and velocities, respectively [47].

$$x_{i,j} = x_{i,j} + V_{i,j} \tag{9}$$

$$V_{i,j} = u * V_{i,j} + c_1 * rand_1 * (LB_i - x_(i,j)) \\ + c_2 * rand_2 * (GB_i - x_{i,j}) \tag{10}$$

The inertia weight value often fluctuates during iterations within the range of [0, 1]. LB_i represents the current best local solution at iteration number I, while GB_i represents the current best global solution at iteration number I. The variables rand_1 and rand_2 are random numbers within the range of [0, 1], while c_1 and c_2 typically denote two constants [47].

TABLE 4. Values of the objectives' weights.

Wight	Substitution	Value
W_1	$W_1 = \frac{2(3+1-1)}{3(3+1)}$	0.5
W_3	$W_2 = \frac{2(3+1-2)}{3(3+1)}$	0.3333
W_2	$W_3 = \frac{2(3+1-3)}{3(3+1)}$	0.1667

3) PROPOSED MULTI-OBJECTIVE FUNCTION

Proposed a new multi-objective function based on the scalarization method that combines the multi-objective into the single solution utilized weights and it was incorporated into the fitness function [48]. The bio-inspired algorithms RMFO and PSO seek to combine the three objectives namely: (i) high accuracy, (ii) smaller false-positive rate, and (iii) a smaller number of subsets features as shown in Equation (11).

$$Fitness = W_1 * Accuracy - W_2 * \\ FPR - W_3 * Numfeatures \tag{11}$$

However, the Rank-Sum (RS) weights method is utilized in this paper to calculate weights because it is commonly used. Equation (12) can be used to compute RS weights [49].

$$W_i = \frac{2(n+1-i)}{n(n+1)} \tag{12}$$

where W_i indicates the variable weight value, n indicates the number of the total weights, and i is the weight number based on its order in Equation (11) and Table 4 illustrates the value for each Wight.

The proposed approach combines the MRFO and PSO algorithms to effectively select relevant features from a given dataset. Initially, the dataset is divided into separate training and testing sets, and a subset of features is generated. Subsequently, the bio-inspired algorithm generates candidate feature subsets using a multi-objective function. MRFO is employed to maximize classification performance while minimizing the number of selected features, guided by the proposed multi-objective approach. These candidate feature

subsets are then evaluated using a KNN classifier trained on the transformed training and testing sets. The search for additional feature subsets continues until a stopping criterion, based on the desired number of selected features, is met. Finally, the approach identifies the mutual features shared by both the MRFO and PSO algorithms to be utilized in the subsequent abnormality detection stage.

E. ABNORMAL TRANSACTIONS DETECTION

Detecting abnormal transactions plays a crucial role in the proposed approach, which aims to customize a prediction model using the selected features from earlier stages. To identify abnormal transactions within the Ethereum network, the proposed approach adopts the SGAN algorithm. The dataset of abnormal transactions is obtained from donors through Etherscamdb. It is worth noting that abnormal transactions constitute a small proportion within the Ethereum network, indicating a high likelihood of encountering unlabeled abnormal transactions. Furthermore, in the scenario where a new abnormal behavior emerges, it won't be automatically included in the database for recognizing intrusion addresses. Semi-supervised learning, by internalizing hidden patterns within the data, aims to generalize from a limited set of labeled data points to accurately classify new, unseen examples. The scarcity of labeled datasets poses a significant challenge in both machine-learning research and real-world applications. Despite the abundance of unlabeled data available (such as images, videos, and text on the internet), assigning class labels to them is often cost-prohibitive, impractical, and time-consuming. To tackle the aforementioned challenges, the SGAN algorithm can provide a solution by automatically assigning accurate labels to transactions. This is made possible due to the enhanced discriminator present in the SGAN, which offers improvements over the vanilla GAN approach.

In SGAN, the discriminator takes a random noise vector z and produces a fake example. On the other side, the discriminator received three types of inputs namely: (i) real labeled transactions (normal and abnormal transactions resulting from BLTE), (ii) real unlabelled transactions (from BLTE without label), (iii) fake transactions generated from the generator. Then, the discriminator classifies the unlabelled and fake transactions which aims to distinguish fake transactions from the real ones, and for real transactions (unlabelled) identifies the correct class (normal or abnormal). However, turning the discriminator from being a binary classifier to a multi-classifier might look like a trivial change (vanilla GAN) but it implies more significance than it receives at first glance. As a result, the training of SGAN was conducted in this form to ensure the classification accuracy is close to a supervised classifier while using labeled and unlabelled transactions. On the other hand, the generator aims to serve as a source of additional information (the fake transactions it produced), which may help the generator learn the relevant pattern in transactions, improving the classification accuracy.

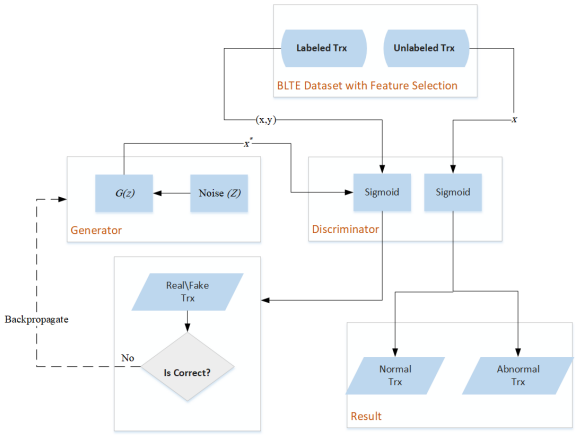

FIGURE 4. Flowchart of ATD-SGAN approach.

Although the BLTE training dataset has 24600 labeled data transactions which are 80% of the BLTE dataset, only a small fraction of these transactions are used in training and pretends that all the remaining transactions are unlabelled. After training, testing set 6150 which is 20% of the BLTE dataset related to Table 2 was used to assess the effectiveness of the classification model in generalizing the previously unseen transactions (unlabelled). By using SGAN, the discriminator becomes a well-trained and robust classifier that may achieve impressive classification performance from few labeled transactions as possible, thereby dimensioning the dependency of the classification process on a huge volume of labeled transactions. Moreover, Figure 4 presents the flowchart of the ATD-SGAN approach.

IV. EXPERIMENT AND RESULTS

This section describes the aspects related to the design methodology and the implementation of the proposed ATD-SGAN approach which is thoroughly explained in Section III. ATD-SGAN approach aims to improve the abnormality detection in the Ethereum network, in terms of detection accuracy, recall, false alarm rate, and F1-measure. Referring to the background of the aforementioned bio-inspired feature selection in Section III-C, the MRFO and PSO algorithm design and evaluation based on the proposed multi-objective function are presented in this section.

A. EVALUATION METRICS

In order to assess the effectiveness of an IDS, various evaluation metrics can be utilized to gauge its performance. The performance metrics are calculated using a confusion matrix derived from the output of the two-class classifier. The confusion matrix provides detailed information about the classification outcomes. Each column in the matrix represents a predicted class instance, while each row represents an actual class instance in the real-world scenario. The equations utilized to assess the performance of feature selection are illustrated in Table 5. The symbols used in the equations

are as follows: TP represents the count of true positives, FN represents the count of false negatives, TN represents the count of true negatives, and FP represents the count of false positives [50], [51], [52], [53], [54], [55].

B. EXPERIMENTAL SETUP

1) IMPLEMENTATION ENVIRONMENT

The ATD-SGAN is implemented using Python programming language, which is characterized by its easiness and implementation robustness, as it is rich in libraries that allow developers to implement machine learning and others easily from out of the box, friendly syntax, and many researchers and developers support python and view it as a standard programming language [56]. In detail, Python version 3.8, and Spyder Editor version 5.2 for facilitating interactive code writing, execution, and result visualization. Moreover, Table 6 presents the main libraries used to implement the ATD-SGAN.

The efficient execution of the proposed approach relies on hardware components that offer ample computational power and memory. This includes a capable multi-core processor, Intel Core i7, with a clock speed of at least 2.0 GHz. Dedicated Graphics Processing Unit (GPU) support is also essential, with GPUs like the NVIDIA GeForce GTX 1080 being preferable. Moreover, 32 GB of RAM has been utilized to handle the computational requirements effectively. Additionally, involved a Solid State Drive (SSD) to enhance data loading speed and storage efficiency, with a capacity of 1 TB.

2) HYPERPARAMETERS OF ATD-SGAN

Fine-tuning of hyperparameters is a crucial step in achieving success in ML and DL models [57]. In order to thoroughly assess the performance of ATD-SGAN, it is necessary to fine-tune multiple hyperparameters. The evaluation experiments are conducted in phases to evaluate the performance of the ATD-SGAN approach, utilizing various hyperparameters as follows:

- **Loss Function:** The binary cross-entropy loss function is used, serving as the objective function in the neural network. This loss function is well-suited for binary classification problems, which is the case for abnormal transaction detection.
- **Activation Function:** Sigmoid activation functions are applied at each node after the linear combination of inputs. The sigmoid function is commonly used in binary classification tasks, as it maps the output to a probability-like range between 0 and 1.
- **Optimizer:** The Adam optimizer is chosen for updating the model's parameters during training. Adam is known for its adaptive learning rate and momentum properties, making it efficient for a wide range of optimization tasks.
- **Learning Rate:** The learning rate, set at 10^{-4}, defines the step size taken during parameter updates. This value is carefully selected to balance convergence speed and stability.
- **Batch Size:** A batch size of 32 is used, indicating the number of events processed in a single update. Larger batch sizes can lead to better hardware utilization and smoother gradient updates.
- **Epochs:** The number of training epochs spans a range from 500 to 8000. This broad range allows for observing how the model's performance evolves over extended training periods.

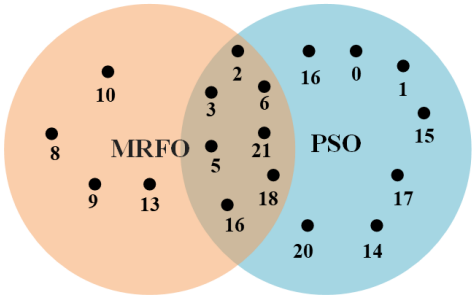

FIGURE 5. Ensemble feature selection results.

Hyperparameter fine-tuning facilitates the acquisition of profound insights into the intricate interactions that exist between hyperparameters and their discernible impact on performance outcomes. This discernment holds pivotal importance in the interpretation of the obtained results of the model. Besides, the comprehensive evaluation of hyperparameters serves as a testament to the depth of assessment undertaken in unraveling the full potential of ATD-SGAN. Moreover, it serves as an unequivocal demonstration of the model's adaptability to varying configurations, effectively enhancing the sphere of intrusion detection.

Furthermore, ATD-SGAN applies MRFO and PSO to both training and test data in order to select a subset of relevant features. The bio-inspired algorithms generate candidate feature subsets, starting with a random subset of features created by a new multi-objective function, as proposed in Equation (11). Furthermore, Table 7 shows the parameters of MRFO and PSO used in the experiments of ATD-SGAN.

C. RESULTS

1) ENSEMBLE FEATURE RESULTS

Let D is a BLTE dataset with 22 features $D = \{F_0, F_1, F_2, \ldots, F_{22}\}$, R is the feature subset from D by RMFO algorithm $R \subseteq D$, and P is the feature subset from D by PSO algorithm $P \subseteq D$. Then, the intersection between two sets R and P presents the mutual feature selection, where $\forall R, P : R \cap P \equiv \{F|F \in R \Lambda F \in P\}$. Figure5 depicts the results of ensemble feature selection based on mutual features.

Consequently, $R = \{2, 3, 5, 6, 8, 9, 10, 13, 16, 18, 21\}$, $P = \{0, 1, 2, 3, 5, 6, 14, 15, 16, 17, 18, 20, 21\}$, and $S = R \cap P = 2, 3, 5, 6, 16, 18, 21$. In summary, a total of 7 features out of 22 are selected as a result of the mutual feature step.

TABLE 5. Evaluation metrics.

Metric	Description	Equation
Accuracy	It measures the percentage of correct predictions in the IDS's model.	$AC = \frac{TP + TN}{TP + TN + FP + FN}$ (13)
False Alarm	It quantifies the proportion of normal points that are incorrectly identified as attacks. A high FPR indicates a lower effectiveness of the IDS	$FPR = \frac{FP}{TN + FP}$ (14)
Recall	It represents the ratio of detected intrusions to the total number of attacks on the system.	$R = \frac{TP}{TP + FN}$ (15)
Precision	It quantifies the proportion of accurately predicted attacks to the total number of attacks within the system.	$P = \frac{TP}{TP + FP}$ (16)
F1-measure	It calculates the average of precision and recall.	$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$ (17)

TABLE 6. Python l Used to implement the ATD-SGAN.

Library	Version	Description
TensorFlow	1.14.0	It is an open-source platform that is used to build neural networks.
NumPy	1.15.4	Scientific computing library for the multidimensional array object.
Pandas	0.25.1	It is manipulated and analyzed data to convert a CSV file to the data frame type.
Sklearn	0.21.3	It is a machine-learning library.
Matplotlib	3.1.1	It is an assistance to create visualizations.
Requests	2.25.1	Simple API to request data by HTTP protocol.
Google Cloud		1.22.0 Google API to connect Python with google BigQuery.
JSON	2.0.9	To read JSON file in python.
Networkx	2.5	It packages to create and manipulate graph data structure.
OS	3.7.4	It is a module that utilizes the operating system to interact with the files system.

TABLE 7. Values of MRFO and PSO control parameters.

Parameter	Value
Number of Iteration	5
Fitness function	Equation (11)
Population Size	22
Random values, r1, r2	[0,1]
Inertia Weight	$0.4 \sim 0.9$
Cognitive factor, c1	1.4
Social factor, c2	1.4

According to Table 4.5, the mutual feature selection is Out degree, Unique in degree, Avg amount incoming, Avg amount outcoming, Active Duration, Mean time interval, Avg gas price. Table 8 presents the sample of results.

2) ATD-SGAN PERFORMANCE ACROSS DIFFERENT EPOCHS

No doubt that detection accuracy is a vital metric for any IDS since it indicates the robustness of the IDS against intrusions, attacks, or abnormal behaviors. However, to demonstrate the robustness and reliability (in terms of accuracy) of the ATD-SGAN, it was run with different training epochs (500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, and 8000, respectively). This extensive experimentation aims to capture the model's performance across different training durations. The outcomes reveal an intriguing trend in terms of detection accuracy, false alarm rate, and F1 measure. As the number of training epochs increases, there is a progressive improvement in all performance metrics. This observation highlights the model's capacity to continuously learn and adapt to the dataset, resulting in heightened accuracy, reduced false alarms, and enhanced F1-measure as the training progresses as shown in Table 9.

3) ENHANCING DETECTION PERFORMANCE

The research contributions, particularly the ensemble feature selection process and the incorporation of the SGAN model, are central to the heightened detection performance of the ATD-SGAN approach. The mutual feature selection step successfully narrows down the feature set from the original 22 to a compact set of 7 essential features. This parsimonious selection not only improves the model's efficiency but also signifies the effectiveness of the multi-objective function utilized for feature selection. Moreover, the SGAN's role in automatic data augmentation equips the model with augmented and diverse data, essential for the training process. The synergy of these contributions culminates in a highly accurate and robust intrusion detection system. To ensure a fair and meaningful comparison, all the aforementioned IDSs were assessed using the BLTE dataset. The results were obtained for each IDS, and the evaluation metrics were computed accordingly. Table 10, presents the results obtained from state-of-the-art IDSs and ATD-SGAN using the BLTE dataset based on seven subset feature selections S (refer to Section IV-C) and compared to the performance using all

TABLE 8. Ensemble bio-inspired feature selection results.

Algorithm	No. Run	Accuracy	False Alarm	No. of Features	Features codes
MRFO	1	95.74%	3.33%	11	{0,1, 6, 9, 10, 11, 12, 13, 15, 19, 20}
	2	94.68%	1.66%	13	{0, 1, 5, 6, 9,11, 12, 14, 15, 16, 17, 20, 21}
	3	93.62%	8.33%	12	{2, 3, 5, 6, 9, 10, 12, 15, 16, 18, 19,21}
	4	92.55%	8.33%	13	{1,2, 6, 5, 7, 9, 10, 11, 13, 15, 17,18, 21}
	5	96.81%	5.00%	11	{2, 3,5, 6, 8, 9, 10, 13, 16, 18, 21}
PSO	1	97.87%	5.00%	13	{0, 1, 2, 3, 5, 6,14, 15, 16, 17, 18, 20, 21}
	2	96.81%	1.66%	15	{2, 3, 5, 6, 8, 10, 11, 12, 13, 14, 15, 17, 18, 20, 21}
	3	96.81%	3.33%	16	{1, 2, 3, 5, 6, 7, 9, 10, 11, 12, 14, 16, 17, 18, 19, 21}
	4	95.74%	8.33%	13	{1, 3, 5, 9, 7, 8,10, 11, 12, 13, 17, 18, 21}
	5	96.81%	8.33%	15	{0, 1, 3, 4, 8, 9, 10, 12, 13, 14, 15, 16, 19, 20, 21}

TABLE 9. Values of performance Metrics of ATD-SGAN with different epochs.

Epochs	Accuracy	False Alarm	F1-measure
500	89.87%	15.86%	89.58%
1000	92.61%	12.18%	92.60%
2000	93.71%	10.32%	93.76%
3000	94.64%	8.80%	94.70%
4000	95.95%	6.76%	96.04%
5000	96.55%	5.70%	96.66%
6000	96.87%	5.17%	96.99%
7000	97.49%	4.18%	97.67%
8000	97.82%	3.50%	97.97%

features of the original dataset. Furthermore, the state-of-the-art approaches selected are based on the related works (refer to Section II).

As shown in Table 10, the results ensure the superiority of ATD-SGAN over the other state-of-art IDSs in terms of the average detection accuracy, false alarm rate, and F1-measure, as it obtained the highest average detection accuracy (i.e., 95.06%) and the highest average f1-measure (95.11%), and lowest false alarm rate (i.e., 8.05%). Overall, the comparison result revealed that ATD-SGAN detection accuracy on the previously seen transactions in the testing dataset is far superior two comparable with other models trained on the same number of labeled transactions.

V. DISCUSSION

In the above-mentioned sections, the ATD-SGAN has been compared with LR, RF, KNN, SVM, MLP, LSMT, and CNN in terms of average detection accuracy, false alarm, and F1 measure. The obtained comparison results ensure that the ATD-SGAN outperformed the other state-of-the-art IDSs in all evaluation metrics. However, this section provides a discussion of enhancement resulting from the ATD-SGAN on the other state-of-the-art IDSs.

Figure 6 (a) depicts the enhancement percentage of the ATD-SGAN on the other state-of-the-art approaches in terms of the average accuracy in detecting abnormal transactions existing in the BLTE dataset across all runs' experiments. However, the enhancement percentages in terms of average detection accuracy look slight if they are taken alone without bearing in mind other metrics used in the evaluation. In fact, if the enhancement percentage resulting from all of the evaluation metrics is considered together, of course, the enhancement will be clearly significant. The false alarm rate is another important evaluation metric that is usually calculated to indicate the degree of effectiveness of any IDS. It denotes the ratio in classifying normal transactions wrongly as abnormal transactions; this means the IDS with the lowest value of false alarm is the best IDS. However, using the BLTE dataset, the ATD-SGAN declines the false alarm rate to LR, SVM, KNN, RF, MLP, LSTM, and CNN, respectively. Figure 6 (b) presents the enhancement percentages of the ATD-SGAN with other state-of-the-art IDSs in terms of false alarms. Besides that, the F1-measure is commonly used to assess the success of a binary classifier, especially when the count of one class is less than another, herein since the BLTE dataset contains binary classes (i.e., two-class instances: (i) normal, and (ii) abnormal transaction), the precision is an important metric to be used in evaluating the ATD-SGAN. However, Figure 6 (c) shows the enhancement percentages of ATD-SGAN with other state-of-the-art IDSs in terms of F1-measure. It can be seen in Figure 6 that the ATD-SGAN also the ATD-SGAN enhanced the F1-measure of the compared other IDSs approaches.

Substantially, concluded from the above findings, the ATD-SGAN is indeed an applicable IDS to address research gaps. In detail, the use of multi-digraph theory to extract the most important set of features from the generated BLTE dataset (refer to Section III-B) has increased the overall performance by decreasing the selected number of features used to train and test the classifier, then in detecting abnormal transactions, respectively. Besides, it was discovered that the proposed multi-objective function (refer to Section III-C), which is implicitly achieved in the research objective number two in this paper, has a direct positive effect on the feature selection algorithm (i.e., MRFO), and consequently on detection process as well. In other words, the use of multi-objectives as a fitness function also ensures the proper efficient selection of a set of features. It also assesses the feature subset if it meets the objectives (i.e., the highest accuracy and recall and the lowest number of features) or not, effectively.

Although deep learning is carried out more efficiently than machine learning, especially when learning a huge volume of data, it still suffers from challenges, which might result in data loss or overfitting problems. The ATD-SGAN proves that it

TABLE 10. Comparison results of ATD-SGAN with state-of-the-art approaches.

Ref	Approach	No of Features	Average Accuracy	Average False Alarm	Average F1-measure
[1]	LR	7	87.20%	13.95%	86.99%
		22	68.40%	38.78%	73.75%
[1]	SVM	7	88.80%	12.98%	88.52%
		22	70.00%	36.84%	75.41%
[13]	CNN	7	91.60%	8.06%	91.63%
		22	76.37%	33.13%	71.61%
[1], [17]	RF	7	85.60%	11.30%	86.15%
		22	72.00%	38.20%	78.13%
[15], [16], [17]	KNN	7	87.20%	12.20%	87.30%
		22	70.77%	36.00%	75.95%
[18]	MLP	7	89.20%	11.11%	89.16%
		22	72.00%	33.33%	77.42%
[19]	LSTM	7	90.00%	9.68%	90.04%
		22	76.00%	33.33%	81.25%
Proposed model	ATD-SGAN	7	95.06%	8.05%	95.11%
		22	79.33%	13.92%	77.83%

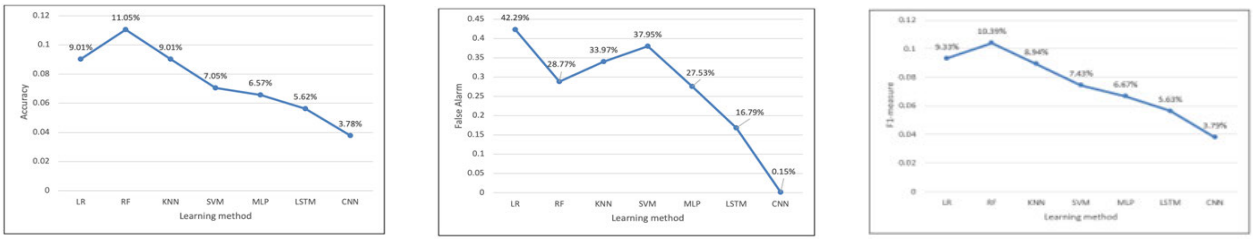

FIGURE 6. Enhancement percentages of ATD-SGAN with other state-of-the-art IDSs.

overcomes these issues by using the Semi-supervised GAN model, which is an unsupervised learning method of DL that automatically generates new augmented data similar to the existing one. Also, the ATD-SGAN is not used SGAN only for generating new data instances, but it is also used to classify (detect) unlabelled data (i.e., testing data). However, selecting and extracting the right features can significantly improve the performance of a deep-learning classifier. Some of the ways in which feature selection and extraction can affect a deep-learning classifier include:

1) Reducing the dimensionality of the data: By selecting the most relevant features and extracting them, you can help reduce the complexity of the data, which can make the training process more efficient and reduce the risk of overfitting.
2) Improving generalization: Removing irrelevant or redundant features can help the classifier learn more generalizable patterns in the data, improving its performance on unseen data.
3) Enhancing interpretability: Extracting meaningful features from the data can help you better understand and interpret the model's decisions, which can be particularly useful in applications where interpretability is important.
4) Decreasing computational complexity: Removing unnecessary features can reduce the number of parameters in the model, decreasing the computational complexity of training and inference.

While feature selection and extraction can be beneficial for deep learning classifiers, it's important to find the right balance. Removing too many features can lead to a loss of important information that the classifier needs to make accurate predictions. On the other hand, ensemble feature selection involves using the predictions of multiple models to identify the most relevant features for a deep learning classifier. This method can have several beneficial impacts on deep learning classifiers, including:

1) Improved accuracy: Combining the predictions of multiple models can help identify a more robust set of relevant features, which can improve the accuracy of the classifier.
2) Reduced risk of overfitting: By aggregating the predictions of multiple models, ensemble feature selection can help prevent the classifier from overfitting to any one particular model, resulting in a more generalizable model.
3) Enhanced interpretability: Ensemble feature selection can help identify a smaller and more interpretable set of features, making it easier to understand and interpret the classifier's decisions.
4) Increased efficiency: By selecting a smaller and more relevant set of features, ensemble feature selection can make the training process more efficient and reduce the computational complexity of the classifier.

Our proposed solution revolves around the utilization of Semi-Supervised Generative Adversarial Networks for the

detection of anomalous transactions within the Ethereum network. We believe that the strengths of our approach lie in several key areas:

- Real Dataset Utilization: Unlike many existing approaches that rely on self-prepared datasets, we employ real-time Ethereum transactions to evaluate the effectiveness of our IDS. This helps in establishing the real-world applicability of our method and allows for a more accurate assessment of its performance.
- Enhanced Detection Accuracy: Our approach seeks to improve the detection accuracy of anomalous transactions through the utilization of state-of-the-art generative adversarial networks. By incorporating both labeled and unlabeled data, our IDS aims to achieve a more refined classification, thus reducing false negatives and positives.
- Transparent Evaluation: In our paper, we emphasize transparency in evaluation by thoroughly discussing the strengths and limitations of our method. We present a comprehensive analysis of our results, including the areas where our approach excels and where further refinement is needed.
- Practical Significance: Our research strives to contribute to the development of resilient IDSs that can make tangible improvements in the detection of anomalous transactions in the Ethereum network. By addressing the research problem's core aspects, we aim to bridge the gap between existing methodologies and the practical requirements of a real-world blockchain environment

In conclusion, the ATD-SGAN approach proves to be highly effective in securing not only the Ethereum network but also other types of blockchain networks. By successfully detecting abnormal transaction attacks, this IDS ensures the network's resilience. When implemented on real Ethereum transactions, the ATD-SGAN efficiently classifies them as normal or abnormal, enabling miners to identify and distinguish fake transactions. As a result, the network becomes more resistant to attacks and abnormal transactions. The ATD-SGAN meets the criteria for delivering strong security measures and effective decision-making capabilities. Additionally, this IDS surpasses the state-of-the-art IDSs in terms of accuracy, recall, false alarm rate, precision, and F1-measure, showcasing its exceptional performance across a range of evaluation metrics.

VI. CONCLUSION AND FUTURE WORKS

Throughout this study, we have introduced a new approach, called ATD-SGAN, that employs Semi-Supervised Generative Adversarial Networks to detect anomalous transactions. This approach capitalizes on the integration of real-time Ethereum transaction data, thereby bridging the gap between existing methodologies and the practical requirements of real-world blockchain environments. Our method's strengths include its utilization of real datasets, which contrasts with the reliance on self-prepared datasets that often lack transparency in terms of attack coverage and accuracy.

The implications of our research are manifold. Firstly, our approach significantly improves detection accuracy by leveraging the power of generative adversarial networks and semi-supervised learning. Secondly, the utilization of real-time Ethereum transactions establishes the relevance of our findings in a rapidly evolving and dynamic blockchain environment. Moreover, our transparent evaluation approach, addressing both strengths and limitations, contributes to the scholarly discourse by fostering transparency and encouraging further advancements. In terms of insights, our study underscores the value of embracing real datasets for evaluating blockchain-based security solutions. The complexities of real-world transactions and the presence of varying attack scenarios challenge us to create more resilient IDSs that can withstand evolving threats.

Additionally, the insights drawn from our results shed light on the intricacies of anomaly detection within blockchain networks, prompting future researchers to delve deeper into refining IDSs and their applications. The ATD-SGAN was compared with LR, RF, KNN, SVM, MLP, LSTM, CNN, and ATD-SGAN using the BLTE dataset, and it outperformed all of them, as it achieved 95.06%, 8.05%, and 95.11% of average accuracy, average false alarm, and average F1-measure, respectively. Particularly the ATD-SGAN can be applied to secure the Ethereum network, and other types of Blockchain networks in general, without being vulnerable to abnormal transaction attacks. When this IDS is implemented on real Ethereum transactions, these transactions are efficiently classified into normal or abnormal ones; thus, the miner can distinguish whether the transaction is fake or not, and consequently, it will have the ability to figure out the abnormal account. Therefore, a miner can secure its network from attacks and abnormal transactions. The ATD-SGAN then satisfies the requirements of achieving high security and efficient self-decision. Despite the successful implementation of the proposed ATD-SGAN to detect abnormal transactions in the Ethereum network, there is still a margin for improvement. The following is a brief list of recommendations that can be improved or provide a basis for future research:

- ATD-SGAN has been designed for binary classification of Ethereum transactions (normal or abnormal). However, the ATD-SGAN can be extended to multi-class anomaly detection problems in the Ethereum network.
- Applying mutual features based on proposed multi-objective function in other network datasets to enhance IDS performance wherein feature selection plays a significant role in detection performance.
- ATD-SGAN approach can be extended to detect other intrusion attacks such as phishing, malware, spam, and botnets.
- Design a real-time approach to detecting abnormal transactions in Blockchain networks.

- Hybridizing the ATD-SGAN with signature-based IDS to enhance the overall detection performance.

ACKNOWLEDGMENT

The authors express their gratitude to the University of Petra and the American University of Madaba, Jordan, for administrative and technical support.

REFERENCES

[1] F. Poursafaei, G. B. Hamad, and Z. Zilic, "Detecting malicious Ethereum entities via application of machine learning classification," in *Proc. 2nd Conf. Blockchain Res. Appl. Innov. Netw. Services (BRAINS)*, Sep. 2020, pp. 120–127.

[2] H. Zhu, W. Niu, X. Liao, X. Zhang, X. Wang, B. Li, and Z. He, "Attacker traceability on Ethereum through graph analysis," *Secur. Commun. Netw.*, vol. 2022, Jan. 2022, Art. no. 3448950.

[3] Q.-B. Nguyen, A.-Q. Nguyen, V.-H. Nguyen, T. Nguyen-Le, and K. Nguyen-An, "Detect abnormal behaviours in Ethereum smart contracts using attack vectors," in *Proc. 6th Int. Conf. Future Data Secur. Eng.* Nha Trang, Vietnam: Springer, Nov. 2019, pp. 485–505.

[4] R. Brandon. (2016). *How an Experimental Cryptocurrency Lost (and Found) $53 Million*. [Online]. Available: https://www.theverge.com/2016/6/17/11965192/ethereum-theft-dao-cryptocurrency-million-stolen-bitcoin

[5] T. Chen, Z. Li, Y. Zhang, X. Luo, A. Chen, K. Yang, B. Hu, T. Zhu, S. Deng, T. Hu, J. Chen, and X. Zhang, "DataEther: Data exploration framework for Ethereum," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jul. 2019, pp. 1369–1380.

[6] J. Frank, C. Aschermann, and T. Holz, "ETHBMC: A bounded model checker for smart contracts," in *Proc. 29th USENIX Secur. Symp.*, 2020, pp. 2757–2774.

[7] L. Brent, A. Jurisevic, M. Kong, E. Liu, F. Gauthier, V. Gramoli, R. Holz, and B. Scholz, "Vandal: A scalable security analysis framework for smart contracts," 2018, *arXiv:1809.03981*.

[8] D. Canellis. (2019). *Hackers Steal $48.7 m in Ethereum From South Korean Cryptocurrency Exchange Upbit*. [Online]. Available: https://thenextweb.com/hardfork/2019/11/27/ethereum-upbit-cryptocurrency-exchange-hackers-stolen-million-hot-wallet

[9] TSMIT. (2020). *Hackers May Have Just Stolen $1 Million From the Ethereum Classic Blockchain in a '51%' Attack*. MIT Technology Review. [Online]. Available: https://www.technologyreview.com

[10] A. H. H. Kabla, M. Anbar, S. Manickam, T. A. Al-Amiedy, P. B. Cruspe, A. K. Al-Ani, and S. Karuppayah, "Applicability of intrusion detection system on Ethereum attacks: A comprehensive review," *IEEE Access*, vol. 10, pp. 71632–71655, 2022.

[11] G. Andresini, A. Appice, L. De Rose, and D. Malerba, "GAN augmentation to deal with imbalance in imaging-based intrusion detection," *Future Gener. Comput. Syst.*, vol. 123, pp. 108–127, Oct. 2021.

[12] J. Lee and K. Park, "GAN-based imbalanced data intrusion detection system," *Pers. Ubiquitous Comput.*, vol. 25, no. 1, pp. 121–128, Feb. 2021.

[13] Y. Sun and L. Gu, "Attention-based machine learning model for smart contract vulnerability detection," *J. Phys., Conf. Ser.*, vol. 1820, no. 1, Mar. 2021, Art. no. 012004.

[14] O. Lutz, H. Chen, H. Fereidooni, C. Sendner, A. Dmitrienko, A. R. Sadeghi, and F. Koushanfar, "ESCORT: Ethereum smart COntRacTs vulnerability detection using deep neural network and transfer learning," 2021, *arXiv:2103.12607*.

[15] A. Elbaghdadi, S. Mezroui, and A. El Oualkadi, "K-nearest neighbors algorithm (KNN): An approach to detect illicit transaction in the Bitcoin network," in *Integration Challenges for Analytics, Business Intelligence, and Data Mining*. IGI Global, 2021, pp. 161–178.

[16] N. Kumar, A. Singh, A. Handa, and S. K. Shukla, "Detecting malicious accounts on the Ethereum blockchain with supervised learning," in *Proc. 4th Int. Symp., Cyber Secur. Cryptogr. Mach. Learn. (CSCML)*. Sheva, Israel: Springer, Jul. 2020, pp. 94–109.

[17] R. F. Ibrahim, A. M. Elian, and M. Ababneh, "Illicit account detection in the Ethereum blockchain using machine learning," in *Proc. Int. Conf. Inf. Technol. (ICIT)*, Jul. 2021, pp. 488–493.

[18] M. Varun, B. Palanisamy, and S. Sural, "Mitigating frontrunning attacks in Ethereum," in *Proc. 4th ACM Int. Symp. Blockchain Secure Crit. Infrastruct.*, May 2022, pp. 115–124.

[19] E. Rabieinejad, A. Yazdinejad, R. M. Parizi, and A. Dehghantanha, "Generative adversarial networks for cyber threat hunting in Ethereum blockchain," *Distrib. Ledger Technol., Res. Pract.*, vol. 2, no. 2, pp. 1–19, Jun. 2023.

[20] R. Tan, Q. Tan, P. Zhang, and Z. Li, "Graph neural network for Ethereum fraud detection," in *Proc. IEEE Int. Conf. Big Knowl. (ICBK)*, Dec. 2021, pp. 78–85.

[21] Z. A. El Houda, A. S. Hafid, and L. Khoukhi, "Cochain-SC: An intra- and inter-domain DDoS mitigation scheme based on blockchain using SDN and smart contract," *IEEE Access*, vol. 7, pp. 98893–98907, 2019.

[22] Z. A. El Houda, A. Hafid, and L. Khoukhi, "Co-IoT: A collaborative DDoS mitigation scheme in IoT environment based on blockchain using SDN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[23] Z. A. E. Houda, A. Hafid, and L. Khoukhi, "BrainChain—A machine learning approach for protecting blockchain applications using SDN," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[24] S. Al-E'mari, M. Anbar, Y. Sanjalawe, and S. Manickam, "A labeled transactions-based dataset on the Ethereum network," in *Proc. Int. Conf. Adv. Cyber Secur.* Singapore: Springer, Feb. 2020, pp. 61–79.

[25] F. Scicchitano, A. Liguori, M. Guarascio, E. Ritacco, and G. Manco, "Blockchain attack discovery via anomaly detection," in *Proc. Consiglio Nazionale Delle Ricerche, Istituto di Calcolo e Reti ad Alte Prestazioni*, 2019, pp. 1–12.

[26] B. Riyaz and S. Ganapathy, "A deep learning approach for effective intrusion detection in wireless networks using CNN," *Soft Comput.*, vol. 24, no. 22, pp. 17265–17278, Nov. 2020.

[27] K. Yoon and S. Kwek, "A data reduction approach for resolving the imbalanced data issue in functional genomics," *Neural Comput. Appl.*, vol. 16, no. 3, pp. 295–306, May 2007.

[28] M. Blachnik and M. Kordos, "Comparison of instance selection and construction methods with various classifiers," *Appl. Sci.*, vol. 10, no. 11, p. 3933, Jun. 2020.

[29] S. Rao, P. Poojary, J. Somaiya, and P. Mahajan, "A comparative study between various preprocessing techniques for machine learning," *Int. J. Eng. Appl. Sci. Technol.*, vol. 5, no. 3, pp. 431–438, 2020.

[30] J.-M. Jo, "Effectiveness of normalization pre-processing of big data to the machine learning performance," *J. Korea Inst. Electron. Commun. Sci.*, vol. 14, no. 3, pp. 547–552, Jan. 2019.

[31] J. B. Awotunde and S. Misra, "Feature extraction and artificial intelligence-based intrusion detection model for a secure Internet of Things networks," in *Illumination of Artificial Intelligence in Cybersecurity and Forensics*. Cham, Switzerland: Springer, 2022, pp. 21–44.

[32] S. Ullah, J. Ahmad, M. A. Khan, E. H. Alkhammash, M. Hadjouni, Y. Y. Ghadi, F. Saeed, and N. Pitropakis, "A new intrusion detection system for the Internet of Things via deep convolutional neural network and feature engineering," *Sensors*, vol. 22, no. 10, p. 3607, May 2022.

[33] S. S. Funai and D. Giataganas, "Thermodynamics and feature extraction by machine learning," *Phys. Rev. Res.*, vol. 2, no. 3, Sep. 2020, Art. no. 033415.

[34] D. Guo, J. Dong, and K. Wang, "Graph structure and statistical properties of Ethereum transaction relationships," *Inf. Sci.*, vol. 492, pp. 58–71, Aug. 2019.

[35] D. Lin, J. Wu, Q. Yuan, and Z. Zheng, "T-EDGE: Temporal weighted multidigraph embedding for Ethereum transaction network analysis," *Frontiers Phys.*, vol. 8, p. 204, Jun. 2020.

[36] J. Brownlee, "How to choose a feature selection method for machine learning," *Mach. Learn. Mastery*, vol. 10, Nov. 2019.

[37] S. Chattopadhyay, A. Dey, and H. Basak, "Optimizing speech emotion recognition using Manta-Ray based feature selection," 2020, *arXiv:2009.08909*.

[38] Y. Duan, C. Liu, S. Li, X. Guo, and C. Yang, "Manta ray foraging and Gaussian mutation-based elephant herding optimization for global optimization," *Eng. Comput.*, vol. 39, no. 2, pp. 1085–1125, Apr. 2023.

[39] M. G. Hemeida, S. Alkhalaf, A.-A.-A. Mohamed, A. A. Ibrahim, and T. Senjyu, "Distributed generators optimization based on multi-objective functions using Manta Rays Foraging Optimization Algorithm (MRFO)," *Energies*, vol. 13, no. 15, p. 3847, Jul. 2020.

[40] K. K. Ghosh, R. Guha, S. K. Bera, N. Kumar, and R. Sarkar, "S-shaped versus V-shaped transfer functions for binary Manta Ray Foraging Optimization in feature selection problem," *Neural Comput. Appl.*, vol. 33, pp. 11027–11041, Jan. 2021.

[41] W. Zhao, Z. Zhang, and L. Wang, "Manta ray foraging optimization: An effective bio-inspired optimizer for engineering applications," *Eng. Appl. Artif. Intell.*, vol. 87, Jan. 2020, Art. no. 103300.
[42] B. Tran, B. Xue, and M. Zhang, "A new representation in PSO for discretization-based feature selection," *IEEE Trans. Cybern.*, vol. 48, no. 6, pp. 1733–1746, Jun. 2018.
[43] R. A. Ibrahim, A. A. Ewees, D. Oliva, M. Abd Elaziz, and S. Lu, "Improved salp swarm algorithm based on particle swarm optimization for feature selection," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 8, pp. 3155–3169, Aug. 2019.
[44] H. B. Nguyen, B. Xue, I. Liu, and M. Zhang, "Filter based backward elimination in wrapper based PSO for feature selection in classification," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2014, pp. 3111–3118.
[45] O. Almomani, "A feature selection model for network intrusion detection system based on PSO, GWO, FFA and GA algorithms," *Symmetry*, vol. 12, no. 6, p. 1046, Jun. 2020.
[46] D. A. Putri, D. A. Kristiyanti, E. Indrayuni, A. Nurhadi, and D. R. Hadinata, "Comparison of naive Bayes algorithm and support vector machine using PSO feature selection for sentiment analysis on E-Wallet review," *J. Phys., Conf. Ser.*, vol. 1641, no. 1, Nov. 2020, Art. no. 012085.
[47] E.-S. M. El-Kenawy and M. Eid, "Hybrid gray wolf and particle swarm optimization for feature selection," *Int. J. Innov. Comput. Inf. Control*, vol. 16, no. 3, pp. 831–844, 2020.
[48] Y.-M. Xia, X.-M. Yang, and K.-Q. Zhao, "A combined scalarization method for multi-objective optimization problems," *J. Ind. Manag. Optim.*, vol. 17, no. 5, pp. 2669–2683, 2021.
[49] N. Gunantara, "A review of multi-objective optimization: Methods and its applications," *Cogent Eng.*, vol. 5, no. 1, Jan. 2018, Art. no. 1502242.
[50] S. Al-E'mari, M. Anbar, Y. Sanjalawe, S. Manickam, and I. Hasbullah, "Intrusion detection systems using blockchain technology: A review, issues and challenges," *Comput. Syst. Sci. Eng.*, vol. 40, no. 1, pp. 87–112, 2022.
[51] Y. Sanajalwe, M. Anbar, and S. Al-E'mari, "COVID-19 automatic detection using deep learning," *Comput. Syst. Sci. Eng.*, vol. 39, no. 1, pp. 15–35, 2021.
[52] N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad, "Survey on SDN based network intrusion detection system using machine learning approaches," *Peer-Peer Netw. Appl.*, vol. 12, no. 2, pp. 493–501, Mar. 2019.
[53] S. Tug, W. Meng, and Y. Wang, "CBSigIDS: Towards collaborative blockchained signature-based intrusion detection," in *Proc. IEEE Int. Conf. Internet Things (iThings) IEEE Green Comput. Commun. (GreenCom) IEEE Cyber, Phys. Social Comput. (CPSCom) IEEE Smart Data (SmartData)*, Jul. 2018, pp. 1228–1235.
[54] Y. Sanjalawe and T. Althobaiti, "DDoS attack detection in cloud computing based on ensemble feature selection and deep learning," *Comput., Mater. Continua*, vol. 75, no. 2, pp. 3571–3588, 2023.
[55] T. Althobaiti, Y. Sanjalawe, and N. Ramzan, "Securing cloud computing from flash crowd attack using ensemble intrusion detection system," *Comput. Syst. Sci. Eng.*, vol. 47, no. 1, pp. 453–469, 2023.
[56] S. Raschka, J. Patterson, and C. Nolet, "Machine learning in Python: Main developments and technology trends in data science, machine learning, and artificial intelligence," *Information*, vol. 11, no. 4, p. 193, Apr. 2020.
[57] P. T. Sivaprasad, F. Mai, T. Vogels, M. Jaggi, and F. Fleuret, "Optimizer benchmarking needs to account for hyperparameter tuning," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 9036–9045.

YOUSEF K. SANJALAWE received the Ph.D. degree in cloud computing and cybersecurity from Universiti Sains Malaysia (USM), Penang, Malaysia, in 2020. He is currently an Assistant Professor with the Department of Cybersecurity, School of Information Technology, American University of Madaba (AUM). He was a field supervisor of Ph.D. students in different fields, including cybersecurity, cloud computing, the IoT, fog computing, optimization, and AI. His main research interests include AI, cybersecurity, blockchain, optimization, cloud computing, and the IoT.

SALAM R. AL-E'MARI received the bachelor's and master's degrees in computer science from Yarmouk University, Jordan, and the Ph.D. degree in cybersecurity from Universiti Sains Malaysia (USM), Penang, Malaysia, in 2022. She is currently an Assistant Professor with the Department of Information Security, University of Petra (UoP). She has made significant contributions to various domains, including blockchain, deep learning, network security, and other computer science disciplines.