Computers, Materials & Continua

Tech Science Press

Doi:10.32604/cmc.2025.061426

ARTICLE

TIDS: Tensor Based Intrusion Detection System (IDS) and Its Application in Large Scale DDoS Attack Detection

Hanqing Sun[1], **Xue Li**[2,*], **Qiyuan Fan**[3] **and Puming Wang**[3]

[1]School of Information Engineering, Henan University of Animal Husbandry and Economy, Zhengzhou, 450046, China
[2]School of Electronic Information Engineering, Henan Institute of Technology, Xinxiang, 453002, China
[3]School of Software, Yunnan University, Kunming, 650500, China

*Corresponding Author: Xue Li. Email: lixue@hait.edu.cn

Received: 24 November 2024; Accepted: 06 May 2025; Published: 09 June 2025

ABSTRACT: The era of big data brings new challenges for information network systems (INS), simultaneously offering unprecedented opportunities for advancing intelligent intrusion detection systems. In this work, we propose a data-driven intrusion detection system for Distributed Denial of Service (DDoS) attack detection. The system focuses on intrusion detection from a big data perceptive. As intelligent information processing methods, big data and artificial intelligence have been widely used in information systems. The INS system is an important information system in cyberspace. In advanced INS systems, the network architectures have become more complex. And the smart devices in INS systems collect a large scale of network data. How to improve the performance of a complex intrusion detection system with big data and artificial intelligence is a big challenge. To address the problem, we design a novel intrusion detection system (IDS) from a big data perspective. The IDS system uses tensors to represent large-scale and complex multi-source network data in a unified tensor. Then, a novel tensor decomposition (TD) method is developed to complete big data mining. The TD method seamlessly collaborates with the XGBoost (eXtreme Gradient Boosting) method to complete the intrusion detection. To verify the proposed IDS system, a series of experiments is conducted on two real network datasets. The results revealed that the proposed IDS system attained an impressive accuracy rate over 98%. Additionally, by altering the scale of the datasets, the proposed IDS system still maintains excellent detection performance, which demonstrates the proposed IDS system's robustness.

KEYWORDS: Intrusion detection system; big data; tensor decomposition; multi-modal feature; DDoS

1 Introduction

The rapid expansion of network infrastructure, coupled with the increasing complexity of network structures, has posed significant challenges in intrusion detection systems. On one hand, smart devices in intrusion detection systems (IDS) collect a large scale of data, the data have complex structures and are from diverse sources, which exhibit the characteristics of big data. On the other hand, the intrusion attack's technologies are becoming increasingly advanced, and the attack's methods are becoming more covert [1]. These attacks exploit network vulnerabilities to overwhelm systems, leading to congestion, paralysis, and potentially severe information leakage [2]. Therefore, it is urgent to design effective intrusion detection systems (IDS) to detect abnormal attacks in large-scale complex networks using big data and artificial intelligence technologies.

Traditional intrusion detection systems mainly rely on matrix-based techniques such as Principal Component Analysis (PCA) and Singular Value Decomposition (SVD), and traditional intrusion detection methods have shown effectiveness in simple contexts [3]. However, traditional intrusion detection methods are difficult to cope with the large-scale and multi-modal network data. Consequently, traditional intrusion detection systems often fall short of accurately identifying attacks in real-world big data scenarios. There are heterogeneous and huge volume data streams in real-world big data scenarios. This inadequacy underscores the pressing need for advanced intrusion detection systems that can not only enhance detection accuracy but also improve data quality through effective denoising techniques.

To address these challenges, this paper proposes a novel tensor-based intrusion detection system (IDS) with big data. By integrating state-of-the-art tensor decomposition techniques with advanced machine learning algorithms, the intrusion detection system aims to provide a more accurate and scalable solution for identifying Distributed Denial of Service (DDoS) attacks with big data. Tensor decomposition allows for keeping multidimensional relationships inherent in network data, facilitating a more nuanced understanding of traffic patterns [4]. This method enables the detection system to better distinguish between normal and malicious traffic, thereby enhancing the robustness and efficiency of DDoS attack detection systems. Through this approach, we seek to improve detection rates and contribute to resilient network security measures in an increasingly interconnected digital landscape. The proposed system brings several contributions to the area of network intrusion detection, including the ability to:

- Proposing to use tensors to model large-scale heterogeneous network big data in intrusion detection systems. The method integrates features from different modalities in a unified format to obtain a more comprehensive representation.
- Developing Tucker-2 Decomposition to propose HOBISVD method by employing Minimum Description Length Principle (MDLP) for feature extraction. We fuse two modalitie's features through tensor computation to obtain eigentensors. The eigentensors (factor matrices) from the tensor decomposition reveal significant interactions and anomalies within multi-modal network big data.
- Proposing a novel intrusion detection system from a perspective of big data, which utilizes tensor algebra to model and analyze the multi-model network data for capturing intricate dependencies and patterns. The system represents large-scale network data in a unified tensor and then combines the HOBISVD and XGBoost classification to effectively detect intrusion.

The organization of the remaining parts of the paper is as follows. Section 2 introduces the related work of IDS detection. Preliminary is introduced in Section 3. In Section 4, a big data-driven INS system is proposed for DDoS detection. And we illustrate a case study to verify the proposed system in Section 5. In Section 6, we summarize the paper.

2 Related Work

We analyze the existing body of work related to IDS systems, providing a comprehensive review of the current state of research in the field, including big data-driven methods and tensor decomposition methods.

2.1 Big Data Driven Method

Big data technology extracts knowledge and values from data by statistical analysis. Big data has three features: 1) objectivity, 2) accuracy, and 3) testability. Big data-based methods have become an important approach to network security. Statistical algorithms are typical big data-driven methods. The methods represent a well-established approach to anomaly detection. This type of anomaly detection identifies DDoS attacks by calculating thresholds to flag unusual behaviors. Analyzing daily traffic distributions can

pinpoint activities that significantly deviate from the norm as potential anomalies. Hamdi and Boudriga [5] introduced a novel method for the statistical analysis of DDoS attacks utilizing wavelet analysis techniques. By transforming traffic data into the frequency domain through wavelet transformation, they were able to identify specific patterns and features associated with DDoS attacks. This method effectively detects and characterizes these malicious activities by extracting frequency-domain features and incorporating statistical analysis. Tao and Yu [6] proposed a DDoS attack detection method based on information entropy within local area network environments. Under typical conditions, the IP entropy tends to be relatively high. However, during a DDoS attack, the volume of packets directed at the target IP increases significantly, leading to a decrease in IP entropy. This reduction serves as an early warning sign of potential DDoS activity. Fortunati et al. [7] proposed an enhanced anomaly detection method based on covariance analysis. This approach involves constructing a covariance matrix from network traffic data to establish a normal distribution profile. By analyzing this covariance matrix, the method can identify deviations from the expected traffic patterns. Abnormal traffic is detected by setting specific thresholds, allowing for the effective identification of anomalies indicative of potential DDoS attacks or other malicious activities.

2.2 Tensor Decomposition and Intrusion Detection System

Tensors are multidimensional arrays. Tensor models with strong expressive ability can mine abundant intrinsic information contained in massive data, and it is a promising method to solve security problems in a big data environment. In terms of the multimodal data problem in large-scale networks, researchers have presented a considerable amount of work based on tensor models for anomaly detection. In [8], a novel approach is proposed for efficient tracking of intrusion in the normal subspace arising from the decomposition of the Parallel Factor Analysis tensor. The method is based on the extraction of a normal subspace obtained by the tensor decomposition technique, considering the correlation between different metrics. Aiming to address dynamic detection issues, some researchers have proposed a series of online detection methods. In [9], the network data is first represented as a unified tensor. Then, an incremental tensor decomposition is proposed for tensor data dimensionality reduction and denoising. In the end, by combining machine learning algorithms, intrusion detection is completed. The work [10] presents an online anomaly detection system capable of handling operational network traffic of large networks.

3 Preliminary

In this section, some preliminaries will be described. The preliminary mainly includes the mathematical theories and operations used in this paper. Tensors are an important tool used in the proposed intrusion detection system, and we will focus on discussing the theory and operations related to tensors.

`Define1: Eigentensor` The eigentensor is the extension of the eigenvector, which is defined as follows. Given a tensor $A \in R^{I_1 \times I_2 \times \cdots I_M \times I_1 \times I_2 \times \cdots I_N}$, $X = [x_{11\cdots 1}, x_{11\cdots 2}, \cdots x_{I_1 I_2 \cdots I_N}] \in R^{I_1 \times I_2 \times \cdots I_N}$, if a pair ($\lambda$-X) satisfies $A *_N X = \lambda *_N X$, we call λ an eigenvalue and X an eigentensor related to λ. Here $*_N$ denotes the multimodal product.

`Define2: Matricization` Matrixization involves unfolding the tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ along the $n-mode$ and representing it as a matrix. Specially, the mapping transforms tensor elements $(i_1, i_2, \ldots, i_N)$ into matrix elements (i_n, j) as Eq. (1):

$$j = 1 + \sum_{\substack{k=1 \\ k \neq n}}^{N} (i_k - 1) J_k \quad \text{with} \quad J_k = \prod_{\substack{m=1 \\ m \neq n}}^{k-1} I_m. \tag{1}$$

Define3: Singular Value Decomposition SVD is very significant in our study. Based on the literature, the SVD formula for matrix **A** is given as Eq. (2):

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T = \sum_{k=1}^{r} \sigma_k \mathbf{u}_k \mathbf{v}_k^T = \sum_{k=1}^{r} \sigma_k \mathbf{u}_k \otimes \mathbf{v}_k. \tag{2}$$

$\otimes$ represents the tensor product, which is defined as $x \otimes y \triangleq xy^T$. The rank of **A**, denoted as $r \leq \min(m, n)$, corresponds to the dimension of space spanned by the columns or rows of **A**. Σ is a diagonal $(r \times r)$ matrix that includes the nonzero singular values of **A**. The singular values are real, non-negative, and follow the convention where $\sigma_1 > \cdots > \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_n$. The vectors u_k and v_k are the orthonormal columns of matrices $\mathbf{U}(m \times r)$ and $\mathbf{V}(n \times r)$. Specially, v_k are the eigenvectors of $\mathbf{A}^T\mathbf{A}$ and $\mathrm{u}_k = \mathrm{A}v_k/\sigma_k$. Both **U** and **V** could be expanded with additional columns to form square and orthogonal matrices of dimensions $(m \times m)$ and $(n \times n)$ matrices.

Define4: Tensor Norm The norm of the tensor $\mathcal{X} \in \mathbb{R}^{\mathrm{I}_1 \times \mathrm{I}_2 \times \cdots \times \mathrm{I}_N}$ is defined as the square root of the sum of the squares of all its elements. Mathematically, this can be expressed as:

$$\|\mathcal{X}\| = \sqrt{\sum_{\mathrm{i}_1=1}^{\mathrm{I}_1} \sum_{\mathrm{i}_2=1}^{\mathrm{I}_2} \cdots \sum_{\mathrm{i}_N=1}^{\mathrm{I}_N} \mathrm{x}_{\mathrm{i}_1\mathrm{i}_2\cdots\mathrm{i}_N}^2}. \tag{3}$$

Define5: Tensor Multiplication The n-mode product of the tensor $\mathcal{X} \in \mathbb{R}^{\mathrm{I}_1 \times \mathrm{I}_2 \times \cdots \times \mathrm{I}_N}$ and the matrix $\mathbf{U} \in \mathbb{R}^{\mathrm{J} \times \mathrm{I}_n}$ can be donated as $\mathcal{X} \times_n \mathbf{U}$, whose size is $\mathrm{I}_1 \times \cdots \times \mathrm{I}_{n-1} \times \mathrm{J} \times \mathrm{I}_{n+1} \times \cdots \times \mathrm{I}_N$. Therefore, we can obtain the calculating formula:

$$(\mathcal{X} \times_n \mathbf{U})_{\mathrm{i}_1 \cdots \mathrm{i}_{n-1}\mathrm{j}\mathrm{i}_{n+1} \cdots \mathrm{i}_N} = \sum_{\mathrm{i}_n=1}^{\mathrm{I}_n} \mathrm{x}_{\mathrm{i}_1\mathrm{i}_2\cdots\mathrm{i}_N}\, \mathrm{u}_{\mathrm{j}\mathrm{i}_n}. \tag{4}$$

In addition, the formula can also transform matrices and tensors:

$$\mathcal{Y} = \mathcal{X} \times_n \mathbf{U} \quad \Leftrightarrow \quad \mathbf{Y}_{(n)} = \mathbf{U}\mathbf{X}_{(n)} \tag{5}$$

4 The Big Data-Driven Intrusion Detection System

The era of big data brings new challenges for network security, while also providing rich resources for IDS innovation. In this work, we propose a data-driven intrusion detection system for DDoS detection. The system specializes in big data-driven intrusion detection methodologies. Fig. 1 shows the big data-driven intrusion detection system, which consists of four layers, namely i) the data layer, ii) the big data processing layer, iii) the big data rule layer, and iv) the big data application layer. Below, we will describe each layer's function, respectively.

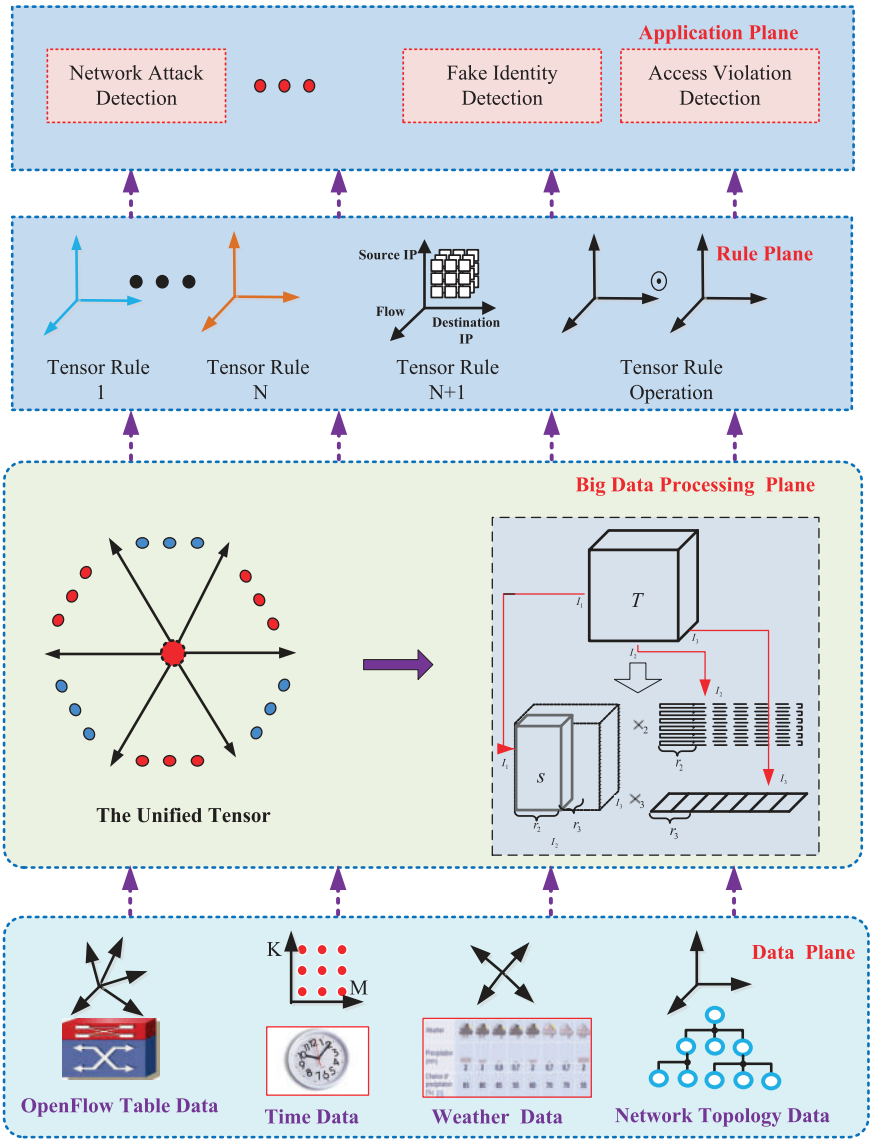

Figure 1: The proposed big data-driven Intrusion detection system. This big data-based intrusion detection system adopts a data-centric security paradigm, specifically designed for large-scale network threat analytics. The Intrusion Detection System consists of four layers, namely i) the data layer, ii) the big data processing layer, iii) the rule layer, and iv) the big data application layer

4.1 Big Data Sensing Layer

In the real world, there is a huge amount of multi-source and heterogeneous data in cyberspace, including network traffic data, network topological structure data, time-related information, and so on [11]. This is the basics of designing a big data-driven intrusion detection system. The data sensing plane is the collecting plane of the intrusion detection system. Some data stream collection tools (Sniffer, NetFlow, probe, and Flow tools) are embedded in the switches, which are a set of distributed monitors. The monitors capture the network data day and night. These data are source IP, port, destination IP, protocol type, data packet length, the number of bytes of traffic between two hosts in a fixed time, the number of data streams, flow entropy, etc. These data have various sources and formats, and the network traffic flow is related to multiple

factors. For example, it is influenced by the time. From 9:00 PM to 11:00 PM, the network traffic volume exceeds typical baseline levels, but at 2:00 AM, it is low. Also, the network traffic is influenced by weather. Hence, a holistic integration of these multi-source datasets is essential for robust network security analysis. How to fuse these data and the relationships between them is a huge challenge [12]. To cope with the problem, we use the network unified tensor (NUT) to fuse the multi-source and heterogeneous data as Fig. 2. Firstly, these data are represented as local tensors in the data collection layer. With the transmission plane, distributed local tensors are propagated to the big data processing plane, where they undergo tensor fusion operations to form a unified global tensor representation.

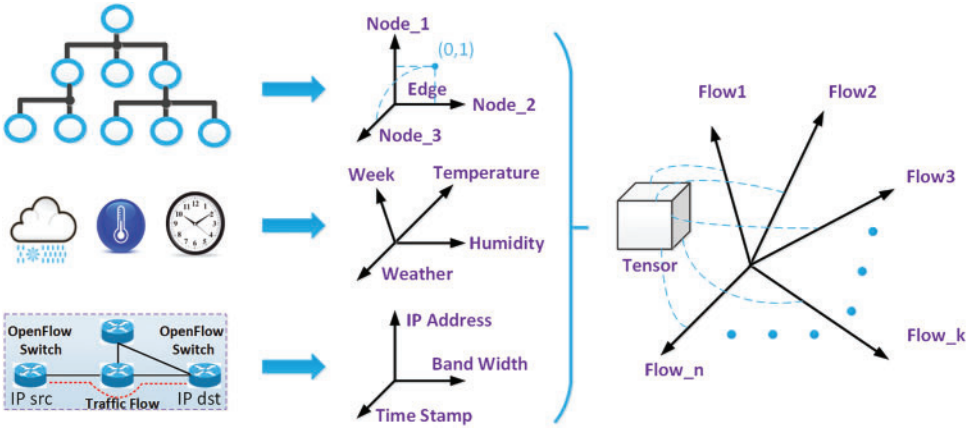

Figure 2: Initially, the raw data is transformed into distributed tensor representations within the big data infrastructure layer. Subsequently, these localized tensors are propagated to the big data processing plane through the dedicated transmission plane, where they will be fused as a global large tensor

4.2 The Proposed Method

From the viewpoint of big data, how to model these multi-source heterogeneous data is a big challenge [13]. The data models are always complex and need powerful computing methods [9]. Considering the tensor's multi-model features, in this work, we use the tensor to fuse the data in Section 4.2. Aiming at the data model, a novel tensor decomposition method is proposed. Specifically, the tensor decomposition method effectively reduces noise in network data and performs feature cutting on each mode through the Minimum Description Length Principle (MDLP) rule. Different from previous tensor decomposition algorithms such as HOSVD(Higher-Order Singular Value Decomposition), this algorithm further integrates features at the feature level through tensor mode multiplication, seamlessly achieving feature dimension reduction and denoising.

1) Definition: Inspired by the insights from [14], we developed a multi-modal feature extraction method (HOBISVD) shown in Fig. 3. HOBISVD is a variation of the Tucker-2 decomposition. It decomposes a tensor into a core tensor multiplied (or transformed) by a matrix along each mode except the first mode. The HOBISVD of a third-order tensor sets the first-factor matrices as the identity matrix. For instance, a HOBISVD can be described as Eq. (6):

$$\mathcal{X} = \mathcal{G} \times_2 B \times_3 C = [\![\mathcal{G} : I, B, C]\!], \tag{6}$$

where $\mathcal{G} \in \mathbb{R}^{I \times Q \times R}$ with $R = K$ and $\mathbf{C} = \mathbf{I}$, the $K \times K$ identity matrix. These concepts extend easily to the N-way case, we can set any subset of the factor matrices to the identity matrix.

Feature truncation is a method to optimize the obtained decomposition, thus obtaining the high-value principal components of each modality, and the result of the truncated decomposition is the approximate solution. In the work, the feature cutting on each mode is based on Minimum Description Length Principle (MDLP) rule. The truncated HOBISVD decomposition is a form of high-order PCA. Thus, in the three-way case where $\mathcal{X} \in \mathbb{R}^{I\times J\times K}$, the Eq. (7) is got,

$$\mathcal{X} \approx \mathcal{G}\times_2 B\times_3 C = \sum_{q=1}^{Q}\sum_{r=1}^{R} g_{:,qr} \circ b_q \circ c_r, \tag{7}$$

here $\mathbf{B} \in \mathbb{R}^{J\times Q}$, and $\mathbf{C} \in \mathbb{R}^{K\times R}$ are truncated factor matrices in each mode, and their column vectors are orthogonal, which are the principal components (PC) related various mode. The tensor $\mathcal{G} \in \mathbb{R}^{I\times Q\times R}$ is the core tensor.

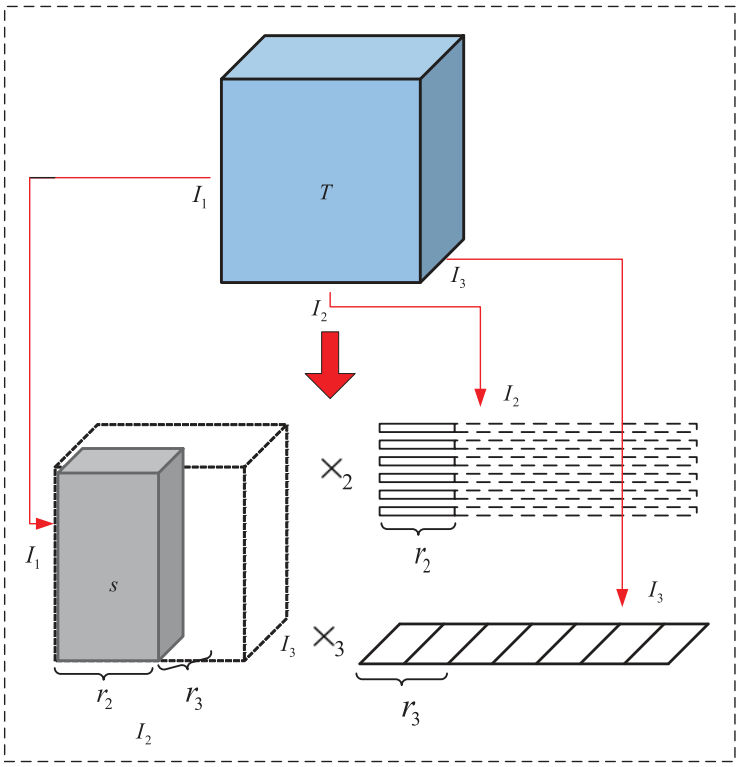

Figure 3: The HOBISVD of 3th-order tensor $\mathcal{X}$. The tensor $\mathcal{X}$ is decomposed into a core tensor S and U and V, which represent the features on the two modalities, respectively. The matrices of U and V are orthogonal, and their column vectors are the orthogonal basis of the space corresponding to each modality

Assuming that $\mathcal{X} \in \mathbb{R}^{I\times J\times K}$, the tensor decomposition can be formulated as an optimization problem as Eq. (8):

$$\underbrace{\min}_{\mathcal{G},B,C} \|\mathcal{X} - [[\mathcal{G}; B, C]]\|_{\mathrm{F}}^2, \tag{8}$$

subject to

$\mathcal{G} \in R^{I\times J\times K}$,

and $B \in R^{J\times J}$, $C \in R^{K\times K}$ are columnwise orthogonal.

Next, HOBISVD is described in detail. HOBISVD involves three key stages, i) the tensor is unfolded in two modes I2 and I3; ii) perform orthogonality-constrained matrix factorization on the unfolded matrix to obtain the eigenvector for each independent modality; iii) by using tensor N-model multiplication, feature-level fusion and dimensionality reduction are performed on two modes to obtain a serious of feature tensors.

2) Unfolding of Tensor $\mathcal{X}$: Unfolding a tensor $\mathcal{X} \in R^{I\times J\times K}$ in two modes J and K. Thus, two unfolded matrix X_2, $\mathcal{X}_3$ can be got as follows:

$$X_2 \in R^{J\times(K\cdot I)}, \tag{9}$$

$$X_3 \in R^{K\times(I\cdot J)}. \tag{10}$$

Fig. 4 illustrates the matrixization process of a 3th-order tensor $\mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_3}$ in two modes, in which $\mathbf{A}_{(1)} \in \mathbb{R}^{n_1\times(n_2\times n_3)}$, $\mathbf{X}_{(2)} \in \mathbb{R}^{n_2\times(n_3\times n_1)}$ are the matrices of the two modes.

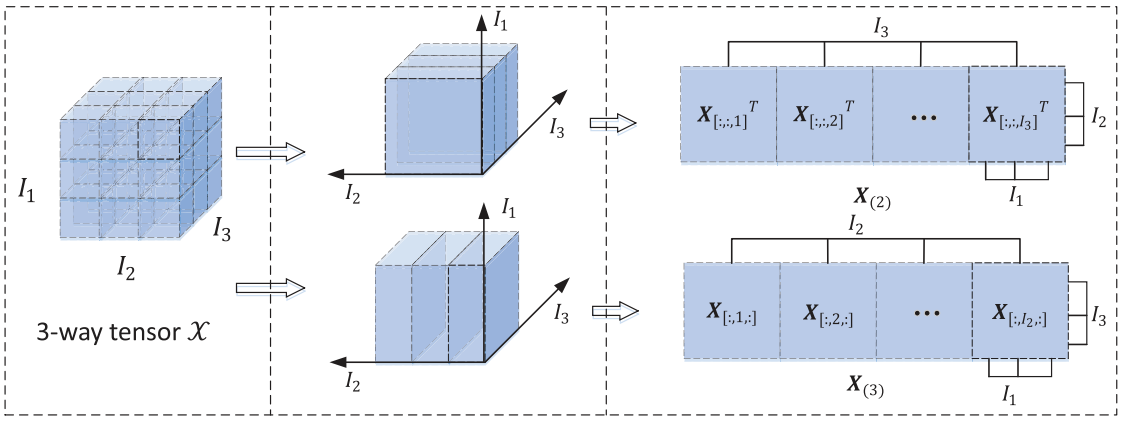

Figure 4: The unfolded processing of the tensor **X** in two modes

3) Applying Orthogonality-Constrained Matrix Factorization on the Unfolded Matrix A_2 and A_3: To achieve the optimization target illustrated in the Eq. (11) and ensure the resulting space remains orthogonal,

$$\underbrace{\min}_{\mathcal{G},B,C} \|\mathcal{X} - [[\mathcal{G}; B, C]]\|_F^2, \tag{11}$$

various methods have been devised by researchers. This article employs the singular value decomposition approach for this purpose. The singular value decomposition enables to calculation of the eigenvalues associated with each mode, which in turn allows for the determination of the optimal truncation strategy for each mode based on these eigenvalues, which will be described in the following. We apply orthogonality-constrained matrix factorization on the two unfolded tensors A_2. Three new matrix U_2, Σ_2, V_2 are got as Eq. (12):

$$\begin{aligned} X_2 &= U_2 \times \Sigma_2 \times V_2 \\ &= \sum_j^J \sigma_j \circ u_{2j} \circ v_{2j} \end{aligned}$$

$$= \sum_{j}^{r_2} \sigma_j \circ u_{2j} \circ v_{2j} + \sum_{r_2+1}^{J} \sigma_j \circ u_{2j} \circ v_{2j}. \tag{12}$$

Similarly, the same method is applied to matrix A_3, resulting in three matrices U_3, Σ_3, V_3 as Eq. (13):

$$\begin{aligned} X_3 &= U_3 \times \Sigma_3 \times V_3 \\ &= \sum_{k}^{K} \sigma_k \circ u_{3k} \circ v_{3k} \\ &= \sum_{k}^{r_3} \sigma_k \circ u_{3k} \circ v_{3k} + \sum_{r_2+1}^{J} \sigma_k \circ u_{3k} \circ v_{3k}. \end{aligned} \tag{13}$$

Fig. 5 presents an example of orthogonal matrix decomposition applied to an unfolded matrix derived from a 3rd-order tensor. In the context of tensor dimensionality reduction, several parameters need to be specified. Among these, r_2 and r_3 are particularly important. These parameters represent the number of leading singular vectors that are retained in the tensor bases U_2 and U_3. The values of r_2 and r_3 directly influence the final dimensionality of the eigentensor S, which is a crucial component in the tensor decomposition process. The selection of the eigenvector parameters r_2 and r_3 is based on retaining a specified portion of the original data from tensors X_2 and $\mathcal{X}_3$. In the following, the Minimum Description Length Principle (MDLP) based truncated method will be described [15].

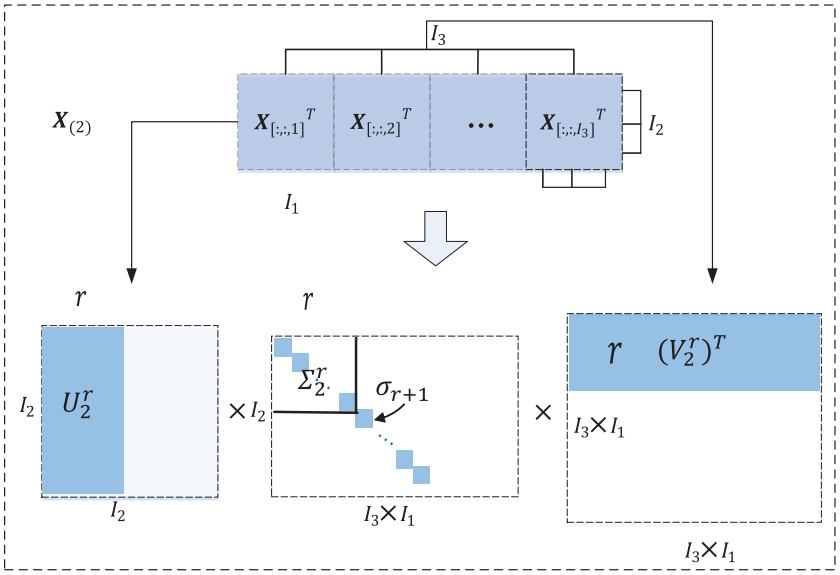

Figure 5: The process of ALS-based matrix decomposition for the unfolded matrix

4) MDLP Based Truncated Method: Enhancing data quality without significantly compromising its inherent characteristics is essential for effective noise reduction. We opt for the Minimum Description Length Principle (MDLP) as our model order selection (MOS) strategy, which aids in ascertaining the necessary rank for Singular Value Decomposition (SVD) based noise elimination. The MDLP is encapsulated

by the following principle:

$$MDL(k) = -log\left(\frac{\prod_{i=k+1}^{p}\sigma_i^{1/(p-k)}}{\frac{1}{p-k}\sum_{i=k+1}^{p}\sigma_i}\right)^{(p-k)N} + \frac{1}{2}k(2p-k)logN, \tag{14}$$

here σ_i are eigenvalues which are from SVD and $i \in \{1, 2, \cdots, p\}$. The eigenvalues are convergent and sorted in descending order, namely $\sigma_1 > \sigma_2 > \cdots > \sigma_p$. p denotes the number of eigenvalues, and N denotes the sample's number in the dataset. When *MDL(k)* gets the minimum value as Eq. (14), the best rank of the matrix A can be ascertained by Eq. (15):

$$rank(\mathbf{A}) = argmin(MDL(k)) + 1, \tag{15}$$

where $k \in \{0, 1, \cdots, p-1\}$. Indeed, as the variable k commences at 0, the aforementioned equation must incorporate an additional 1. Subsequently, only the initial $rank(\mathbf{A})$ eigenvalue is preserved, while all remaining eigenvalues are assigned a value of 0, denoted as $\sigma_1 > \sigma_2 > \cdots > \sigma_{rank(\mathbf{A})}$ and $\sigma_{rank(\mathbf{A})+1} = \cdots = \sigma_p = 0$.

We construct a canonical sinusoidal signal $y = sin(x)$ as a baseline waveform and superimpose zero-mean Gaussian white noise onto the pristine signal. We then apply singular value decomposition (SVD) for noise reduction. Fig. 6 illustrates the progression of singular values as a rank function. Numerical analysis demonstrates that when the matrix rank satisfies $rank \geq 2$, the singular values are diminutive when $rank \geq 2$. The data encompassed within the eigenvectors associated with these singular values is noise-related. Hence, we maintain the initial three singular values.

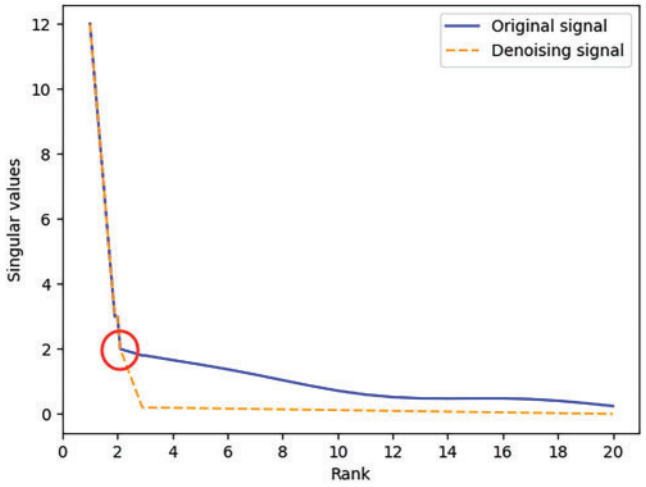

Figure 6: The progression of singular values as a function of rank

5) Tensor Multiplication Based Feature-Level Fusion and Dimensionality Reduction: According to the above step, we have obtained the two reduced U_2, and U_3 in I_2 mode and I_3 mode, which are directly isolated from each other. How to integrate the information of the two modalities is a big challenge. To address this issue, the tensor N-mode multiplication between tensors and matrices is proposed to solve the problem, as shown in Eq. (16). First, the characteristic information on modality 2 is integrated, and then the feature

information on the integrated modality is fused. Since a principal component truncation operation is used, dimensionality reduction and denoising are also achieved in the process.

$$S = X \times_2 \left(U_2^{r_2}\right)^T \times_3 \left(U_3^{r_3}\right)^T, \tag{16}$$

where X is the initial tensor, $\left(U_2^{r_2}\right)^T$ is the transpose of the r_2-dimensionally reduced U_2 tensor, $\left(U_3^{r_3}\right)^T$ is the transpose of the r_3-dimensionally reduced U_3 tensor. The Algorithm 1 shows the process of HOBISVD. In nature, HOBISVD is a mathematical method used for tensor decomposition in high-order dimensions. It can decompose a high-dimensional data tensor into the product of multiple low-rank matrices, thereby extracting important features from the data. Essentially, the method can also be applied to deep learning to reduce dimensionality and extract useful information from input data. By performing HOBISVD decomposition on input features, we can obtain feature subspaces at different levels and directions, each containing strongly correlated features to some extent in the original data.

Algorithm 1: The Proposed HOBISVD Based Dimensionality Reduction and Denoising Algorithm

Require: Tensor $\mathcal{X} \in R^{I_1 \times \cdots \times I_N}$ and feature dimension M.
Ensure: Feature matrix $S_i \in R^{M \times M}$.

for $i = 1$ to $i = N$ **do**
  Unfolding tensor A in model I_i;
  Getting X_i.
end for
for $i = 1$ to $i = N$ **do**
  If i = 1,
  $U_1 = E$;
  Else
  Computing: $X_i \approx U_i^{r_i} \times \Sigma_i^{r_i} \times V_i^{r_i}$;
  Estimating $rank(X_i)$ through MDLP based on Eq. (14);
  Getting the truncated $U_i^{r_i}$ through $rank(A_i)$.
end for
Obtain all $U_i^{r_i}$, where $i \in \{2, \cdots, N\}$.
for $i = 2$ to $i = N$ **do**
  Computing: $S_i = X \times_2 {U_i^{r_i}}^T$;
end for
Obtain $S \in R^{I_1 \times {I_2}^{r_2} \times {I_3}^{r_3} \times \cdots \times {I_N}^{r_n}}$.
Slicing S along the first mode I_1 and getting a series of features tensor $S_{[1,:,:\cdots,:]}, \cdots, S_{[I_1,:,:\cdots,:]}$.
return Feature tensors S_i.

6) XGBoost Classification Method: Aiming to deal with large-scale data, XGBoost (eXtreme Gradient Boosting) was proposed to improve the stability and accuracy of the large-scale model. It is an ensemble learning algorithm based on gradient-boosted trees (GBT). The key to XGBoost's success lies in its adaptability across diverse scenarios. XGBoost is a powerful and versatile machine-learning system known for its tree-boosting capabilities. Its influence has been acknowledged in various machine learning and data mining competitions. With tensor mode multiplication, feature-level fusion and dimensionality reduction are performed on two modes to obtain a series of feature tensors. In nature, it is a large-scale feature dataset. Aiming at the feature dataset. Aiming at the series of feature tensors, the XGBoost method is used for DDoS

detection through classification. The input is the serious of feature tensors $[S(1,,),S(2,,),\cdots,S(I,,)]$, the output is Eq. (17):

$$\hat{Y}_i = \sum_{k=1}^{K} f_k\left(S(i,,)\right), f_k \in F. \tag{17}$$

The object is Eq. (18):

$$Obj = \sum_{i=1}^{I} l\left(\hat{Y}_i, Y_i\right) + \sum_{k=1}^{K} \Omega\left(f_k\right). \tag{18}$$

The primary objective of XGBoost is to push the boundaries of machine learning limitations to offer scalable, portable, and precise solutions. When dealing with huge network data, the distributed version of XGBoost demonstrates exceptional portability, effectively addressing the challenges associated with large-scale datasets [16].

5 The Case Study

As depicted in Fig. 7, a case study is conducted to validate the proposed system. In our experiment, we used an SDN environment to simulate the experimental setup, where the OpenFlow switch collected data, which was then sent to the SDN(Software Defined Network) controller, and fused into a unified tensor model. Through HOBISVD decomposition, multimodal features were extracted, clustering was completed using XGBoost, and DDoS attack detection was achieved through clustering.

This section is structured into three distinct subsections. Section 5.1 provides an overview of the experimental dataset used. Section 5.2 focuses on detailing the pertinent evaluation metrics. Finally, Section 5.3 presents the results of the comparative experiment.

5.1 Experimental Dataset

Dataset CICDDOS 2019, curated by the Canadian Institute for Cyber Security (CIC), includes a vast array of network data with 87 features and millions of traffic instances, encompassing various types of DDoS attacks [17]. In our experimental setup, we constructed a representative subset by randomly sampling 40,000 data points from the original dataset, with a stratified distribution of 32,000 normal traffic instances and 8000 malicious attack samples, as specified in Table 1. Additionally, we eliminated several features that were deemed insignificant for the classification process from the initial set of 87. Consequently, we narrowed it down to 64 key features, as outlined in Table 2. The approximate distribution of the used data is shown in Fig. 8.

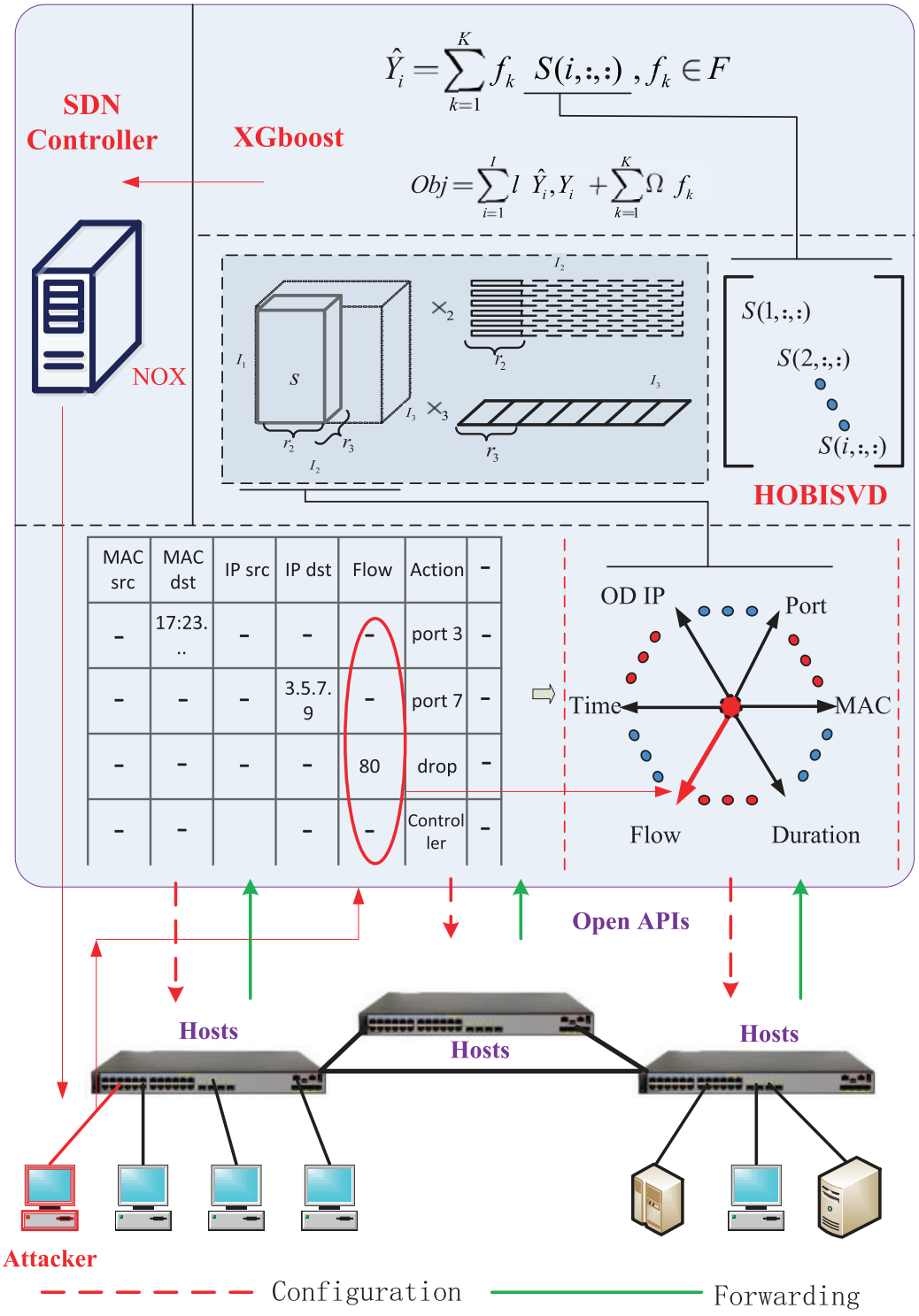

Figure 7: The case study. Firstly, the intrusion detection system collects data from the flow tables and constructs a unified tensor. Through big data processing, the tensor rule is mined and then provides intrusion detection services

Table 1: CICDDoS2019 subset construction

Order	Traffic type	Total	Order	Traffic type	Total
(1)	DNS	800	(7)	NTP	800
(2)	LDAP	800	(8)	SSDP	800
(3)	NETBIOS	800	(9)	UDP-Lag	800
(4)	SNMP	800	(10)	SYN	800
(5)	UDP	800	(11)	BENIGN	32,000
(6)	MSSQL	800			

Table 2: CICDDoS2019 network data set features used

Order	Feature	Order	Feature
(1)	Source-Port	(33)	Packet-Length-Min
(2)	Destination-Port	(34)	Packet-Length-Max
(3)	Flow-Duration	(35)	Packet-Length-Avg
(4)	Total-Fwd-Packet	(36)	Packet-Length-Std-Dev
(5)	Total-Bwd-Packet	(37)	Packet-Length-Var
(6)	Total-Length-Fwd-Packet	(38)	FIN-Flag-Count
(7)	Total-Length-Bwd-Packet	(39)	SYN-Flag-Count
(8)	Fwd-Packet-Length-Max	(40)	RST-Flag-Count
(9)	Fwd-Packet-Length-Min	(41)	PUSH-Flag-Count
(10)	Fwd-Packet-Length-Avg	(42)	ACK-Flag-Count
(11)	Fwd-Packet-Length-Std-Dev	(43)	URG-Flag-Count
(12)	Bwd-Packet-Length-Max	(44)	CWE-Flag-Count
(13)	Bwd-Packet-Length-Min	(45)	ECE-Flag-Count
(14)	Bwd-Packet-Length-Avg	(46)	Download/Upload-Ratio
(15)	Bwd-Packet-Length-Std-Dev	(47)	Avg-Packet-Size
(16)	Flow-Bytes/s	(48)	Avg-Fwd-Segment-Size
(17)	Flow-Packets/s	(49)	Avg-Bwd-Segment-Size
(18)	Flow-IAT-Avg	(50)	Subflow-Fwd-Packets
(19)	Flow-IAT-Max	(51)	Subflow-Fwd-Bytes

(Continued)

Table 2 (continued)

Order	Feature	Order	Feature
(20)	Flow-IAT-Min	(52)	Subflow-Bwd-Packets
(21)	Fwd-IAT-Total	(53)	Subflow-Bwd-Bytes
(22)	Fwd-IAT-Avg	(54)	Fwd-Win-Bytes
(23)	Fwd-IAT-Max	(55)	Bwd-Win-Bytes
(24)	Fwd-IAT-Min	(56)	Fwd-Active-Data-Packet
(25)	Bwd-IAT-Total	(57)	Fwd-Min-Segment-Size
(26)	Bwd-IAT-Avg	(58)	Avg-Time-Active-Flow
(27)	Bwd-IAT-Max	(59)	Std-Dev-Time-Active-Flow
(28)	Bwd-IAT-Min	(60)	Max-Time-Active-Flow
(29)	Fwd-Header-Length	(61)	Min-Time-Active-Flow
(30)	Bwd-Header-Length	(62)	Avg-Time-Idle-Flow
(31)	Fwd-Packet/s	(63)	Std-Dev-Time-Idle-Flow
(32)	Bwd-Packet/s	(64)	Min-Time-Idle-Flow

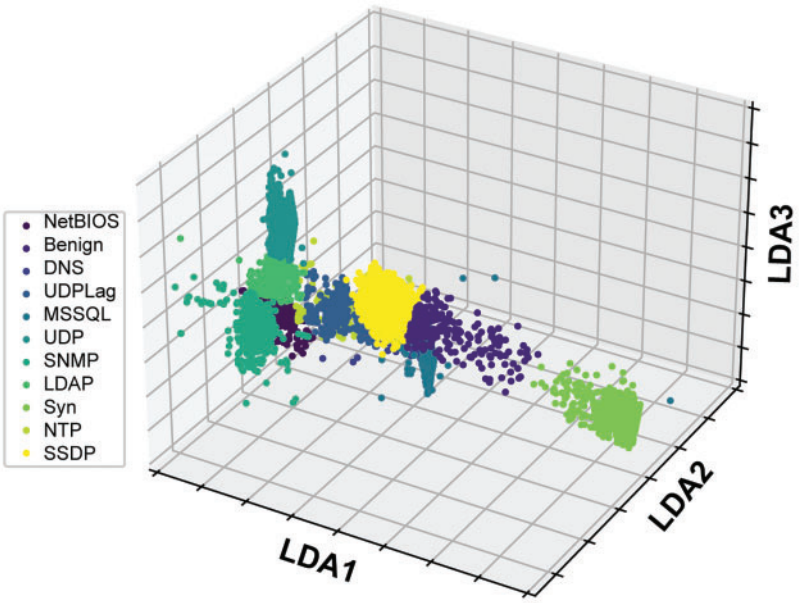

Figure 8: Data distribution of CIC-DDoS2019

5.2 Evaluation Metrics

According to the confusion matrix, we use TP, FP, TN, and FN to represent True Positive, False Positive, True Negative, and False Negative.

5.2.1 Accuracy (Acc)

Accuracy is a metric that quantifies the classifier's ability to correctly predict the entire sample, indicating the degree to which the true values align with the predicted values. A higher Accuracy value signifies better classification performance, as it suggests that a larger proportion of the predictions are accurate. Accuracy is defined in Eq. (19):

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}. \tag{19}$$

5.2.2 Precision (Pre)

Precision is a metric that assesses the classifier's ability to correctly predict the accuracy of positive samples. It measures the proportion of positive predictions that are positive, indicating how many of the predicted positive samples are true positives. A higher Precision value indicates better classification performance in terms of identifying positive instances correctly. It is defined in Eq. (20):

$$Precision = \frac{TP}{TP + FP}. \tag{20}$$

5.2.3 Recall (Rec)

Recall can measure the proportion of actual positive samples that are correctly identified by the classifier. A higher Recall value indicates better classification performance in terms of capturing positive instances within the dataset, and is is defined in Eq. (21):

$$Recall = \frac{TP}{TP + FN}. \tag{21}$$

5.2.4 F1-Score

The F1-score is a pivotal metric that encapsulates the performance of a model by harmonizing the critical aspects of Recall and Precision. It is particularly adept at providing a balanced evaluation, especially in scenarios where the dataset may be unevenly distributed. The F1-score is calculated using a formula that employs the harmonic mean, which effectively weighs both Precision and Recall, ensuring that neither metric can overshadow the other without affecting the overall score. Fi-Score is defined in Eq. (22):

$$F1\text{ - }Score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}. \tag{22}$$

5.3 Contrast Experiments

In this subsection, a series of comparative experiments will be carried out. Focusing on the CICD-DoS2019 dataset detailed in Table 1, which comprises 64 features, the data matrix is denoted as $\mathbf{X} \in \mathbb{R}^{M \times N}$, where M represents the total number of data points and $N = 64$ signifies the number of features. Within the context of the HOBISVD-based denoising algorithm, the matrix $\mathbf{X}$ is transformed into a three-way tensor $\mathcal{X} \in \mathbb{R}^{M \times N_1 \times N_2}$, with $N_1 = N_2 = 8$ and $N = N_1 \times N_2$.

5.3.1 Different Denoising Algorithm

In Table 2, various datasets of different sizes are randomly chosen to assess the impact of various denoising techniques on the classification capabilities of XGBoost. The results of these experiments are depicted in Fig. 9 and Table 3. The findings indicate that denoising enhances detection accuracy compared

to RAW (nondenoised) data, and SVD, HOSVD, and HOOI (Higher Order Orthogonal Iteration of Tensors) denoising algorithms yield lower classification efficiency and performance than HOBISVD.

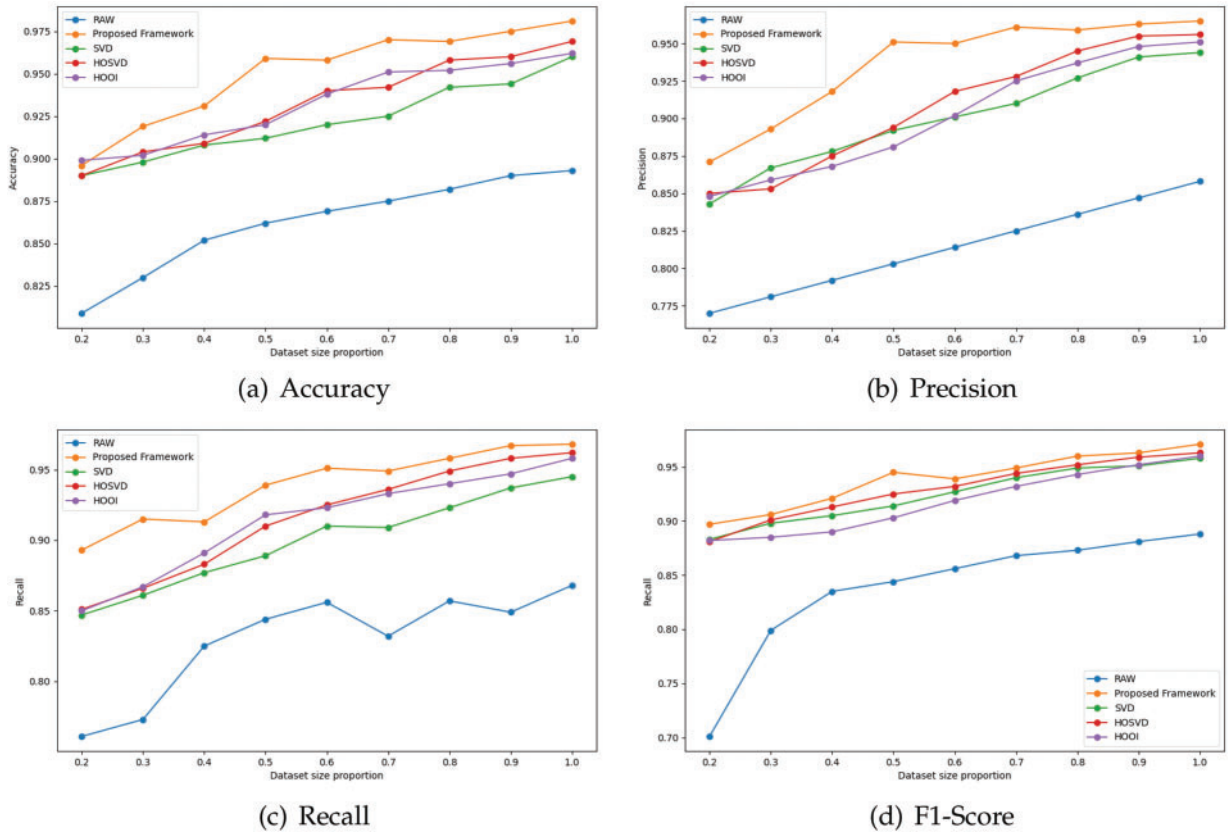

Figure 9: The effect of different denoising algorithms is compared under the truncated size $2 * 2$

Table 3: Comparison of different denoising algorithms with the truncation sizes $2 * 2$

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
HOBISVD (Proposed detection system)	**98.18**	**96.65**	**96.85**	**97.10**
RAW	89.30	85.33	86.81	88.37
SVD [18]	96.16	94.26	94.55	95.82
HOSVD [19]	97.17	95.92	96.21	96.57
HOOI [19]	96.62	95.01	95.87	96.19

Furthermore, to validate the efficacy of HOBISVD, experiments are carried out with truncation sizes ranging from $2 * 2$ to $3 * 3$. Results in Fig. 10 and Table 4 show that the truncation sizes $3 * 3$ are worse than the truncation sizes $2 * 2$ in both the performance of numerical results and the algorithm's stability. This is because, after high-order decomposition and reconstruction, the truncation sizes $2 * 2$ condense nearly all the required properties, which makes it an optimal choice without any redundancy. This also proves the correctness of the selection of $rank = 2$ in Fig. 6.

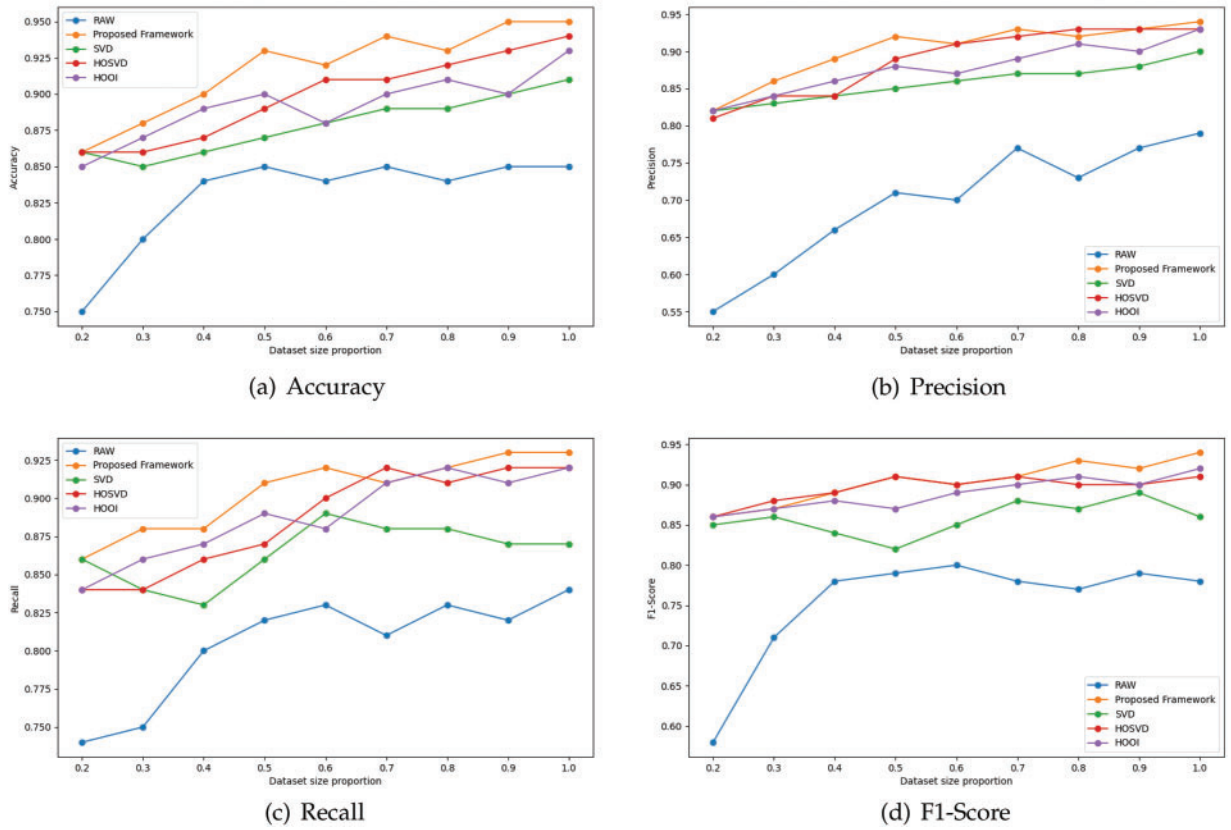

Figure 10: The effect of different denoising algorithms is compared under the truncated size $3 * 3$

Table 4: Comparison of different denoising algorithms with the truncation sizes $3 * 3$

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
HOBISVD (Proposed detection system)	**95.52**	**94.38**	**94.15**	**93.96**
RAW	85.66	88.89	80.12	83.33
SVD [18]	91.34	90.26	91.25	91.24
HOSVD [20]	93.78	**94.38**	93.27	93.82
HOOI [20]	93.81	93.33	**94.15**	93.72

5.3.2 Different Classification Algorithm

A classification algorithm is another important part of the intrusion detection system. Therefore, we set the fixed denoising algorithm as the proposed HOBISVD, and evaluate the performance of different classification algorithms by varying the size of the datasets. The compared algorithms include Linear Discriminant Analysis (LDA), Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM), and XGBoost. The results of these comparisons are illustrated in Fig. 11, leading to the following conclusions:

- As the size of the data set increases, the classification gets better because more and more features are available for classification;
- The application of HOBISVD for denoising datasets leads to superior detection outcomes across various classification algorithms;
- XGBoost demonstrates both rapid processing speed and effective detection capabilities across datasets of varying sizes;
- The proposed intrusion detection system exhibits significant robustness across datasets of different sizes, consistently delivering high classification performance.

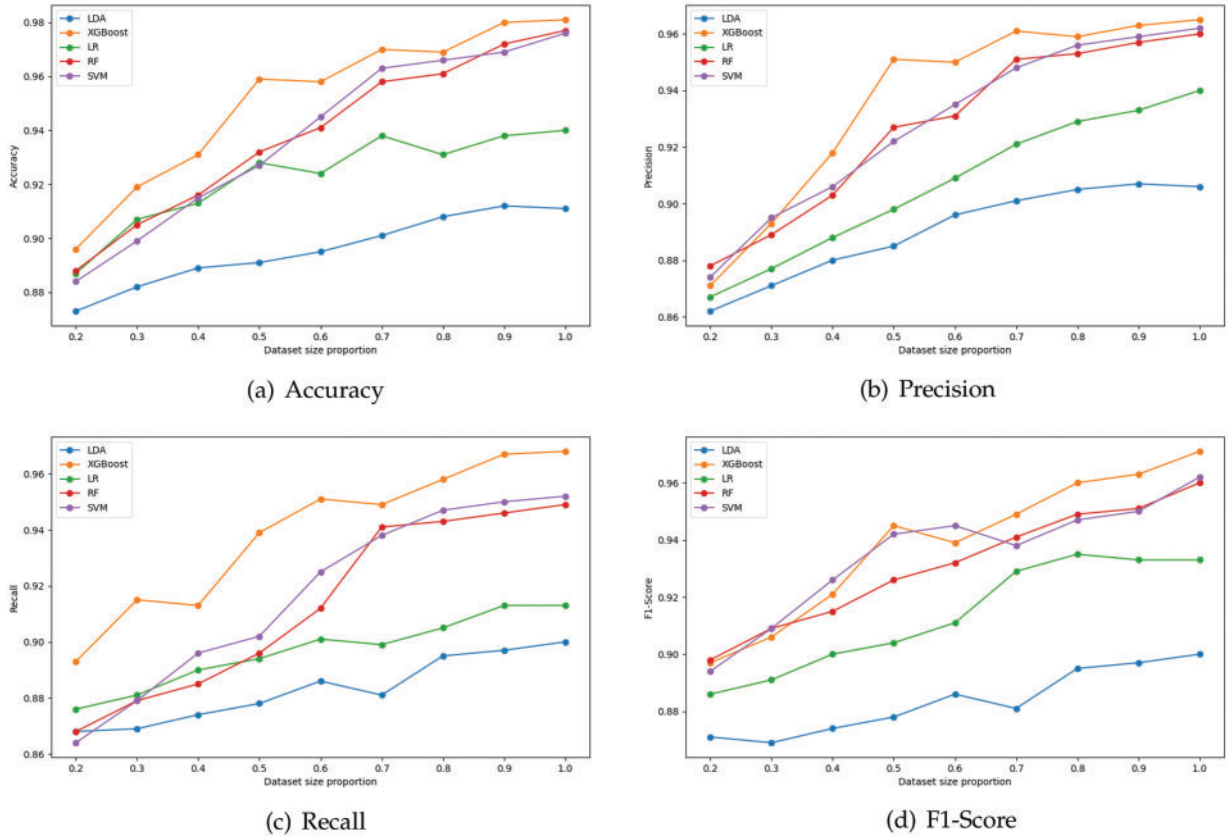

Figure 11: Classification performance of different classification algorithms

6 Summary

The exponential growth of internet users and the proliferation of smart devices have led to an exponential increase in the volume of data, forming massive data collections. The sources and formats of network data are complex and varied. This poses significant challenges in data modeling and analysis for cyber attack detection. To address the problem, this article uses a unified tensor to construct a DDoS attack detection model. Aiming at the model, a novel data analysis method is proposed for reducing the dimensionality and denoising multi-modal data through tensor decomposition. Then we seamlessly integrate the XGBoost classification algorithm to solve the DDoS attack detection problem.

Future research will focus on advancing tensor decomposition techniques, such as tensor train decomposition, to better capture the intricate relationships within network big data. Integrating tensors with advanced machine learning algorithms such as deep learning or ensemble methods could lead to more robust

and accurate network attack detection systems. Real-time tensor processing algorithms are also crucial; the co-development of (1) real-time tensor processing frameworks, (2) parallel-optimized tensor operators, and (3) streaming-enabled tensor architectures emerges as a critical triad in modern intrusion detection systems.

Acknowledgement: The authors acknowledge Jing Xu and Sizhang Li for the helpful discussions for this paper.

Funding Statement: This work was supported in part by the National Nature Science Foundation of China under Project 62166047; in part by the Yunnan International Joint Laboratory of Natural Rubber Intelligent Monitor and Digital Applications under Grant 202403AP140001; in part by the Xingdian Talent Support Program under Grant YNWR-QNBJ-2019-270.

Author Contributions: The authors confirm contribution to the paper as follows: Study conception and design: Hanqing Sun, Xue Li, Puming Wang; Data collection: Hanqing Sun, Xue Li, Qiyuan Fan; Analysis and interpretation of results: Hanqing Sun, Xue Li; Draft manuscript preparation: Hanqing Sun, Xue Li. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Not applicable.

Ethics Approval: Not applicable.

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

References

1. Srivastava A, Sinha D. FP-growth-based signature extraction and unknown variants of DoS/DDoS attack detection on real-time data stream. J Inf Secur Appl. 2025;89:103996. doi:10.1016/j.jisa.2025.103996.
2. Habib AA, Imtiaz A, Tripura D, Faruk MO, Hossain MA, Ara I, et al. Distributed denial-of-service attack detection short review: issues, challenges, and recommendations. Bulletin Electr Eng Inform. 2025;14(1):438–46. doi:10.11591/eei.v14i1.8377.
3. Li X, Cheng J, Zhang B, Tang X, Sun M. An adaptive DDoS detection and classification method in blockchain using an integrated multi-models. Comput Mater Contin. 2023;77(3):3265–88. doi:10.32604/cmc.2023.045588.
4. Alabdulatif A, Thilakarathne NN, Aashiq M. Machine learning enabled novel real-time iot targeted DoS/DDoS cyber attack detection system. Comput Mater Contin. 2024;80(3):3655–83. doi:10.32604/cmc.2024.054610.
5. Hamdi M, Boudriga N. Detecting Denial-of-Service attacks using the wavelet transform. Comput Commun. 2007;30(16):3203–13. doi:10.1016/j.comcom.2007.05.061.
6. Tao Y, Yu S. DDoS attack detection at local area networks using information theoretical metrics. In: 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications; 2013; Melbourne, Australia. p. 233–40.
7. Fortunati S, Gini F, Greco MS, Farina A, Graziano A, Giompapa S. An improvement of the state-of-the-art covariance-based methods for statistical anomaly detection algorithms. Signal Image Video Process. 2016;10:687–94. doi:10.1007/s11760-015-0796-y.
8. Streit A, Santos GH, Leão RM, e Silva EdS, Menasché D, Towsley D. Network anomaly detection based on tensor decomposition. Comput Netw. 2021;200:108503. doi:10.1016/j.comnet.2021.108503.
9. Fan Q, Li X, Wang P, Jin X, Yao S, Miao S, et al. IDAD: an improved tensor train based distributed DDoS attack detection framework and its application in complex networks. Future Gener Comput Syst. 2025;162:107471. doi:10.1016/j.future.2024.07.049.
10. Shajari M, Geng H, Hu K, Leon-Garcia A. Tensor-based online network anomaly detection and diagnosis. IEEE Access. 2022;10:85792–817. doi:10.1109/access.2022.3197651.
11. Ye Z, Luo J, Zhou W, Wang M, He Q. An ensemble framework with improved hybrid breeding optimization-based feature selection for intrusion detection. Future Gener Comput Syst. 2024;151:124–36. doi:10.1016/j.future.2023.09.035.

12. Hu X, Gao W, Cheng G, Li R, Zhou Y, Wu H. Toward early and accurate network intrusion detection using graph embedding. IEEE Trans Inform Forensics Secur. 2023;18:5817–31. doi:10.1109/tifs.2023.3318960.
13. Kumar GA, Katiyar A, Srinivasan K. Elevating IDS capabilities: the convergence of SVM, Deep learning, and RFECV in network security. In: 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE). Vellore, India: IEEE; 2024. p. 1–16.
14. Kolda TG, Bader BW. Tensor decompositions and applications. SIAM Review. 2009;51(3):455–500. doi:10.1137/07070111x.
15. Fan Q, Li X, Wang P, Jin X, Yao S, Miao S, et al. BDIP: an efficient big data-driven information processing framework and its application in DDoS attack detection. IEEE Trans Netw Serv Manag. 2025;22(1):284–98. doi:10.1109/tnsm.2024.3464729.
16. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016; San Francisco, CA, USA. p. 785–94.
17. Subrmanian M, Shanmugavadivel K, Nandhini P, Sowmya R. Evaluating the performance of LSTM and GRU in detection of distributed denial of service attacks using CICDDoS2019 dataset. In: Proceedings of 7th International Conference on Harmony Search, Soft Computing and Applications: ICHSA 2022. Seoul, Republic of Korea: Springer; 2022. p. 395–406.
18. Maranhão JPA, da Costa JPC, Javidi E, de Andrade CAB, de Sousa Jr RT. Tensor based framework for distributed denial of Service attack detection. J Netw Comput Appl. 2021;174:102894. doi:10.1016/j.jnca.2020.102894.
19. Xu J, Li X, Wang P, Jin X, Yao S. Multi-modal noise-robust DDoS attack detection architecture in large-scale networks based on tensor SVD. IEEE Trans Netw Sci Eng. 2022;10(1):152–65. doi:10.1109/tnse.2022.3205708.
20. Li S, Xu J, Liu P, Li X, Wang P, Jin X, et al. Truncated lanczos-TSVD: an effective dimensionality reduction algorithm for detecting DDoS attacks in large-scale networks. IEEE Trans Netw Sci Eng. 2024;11(5):4689–703. doi:10.1109/tnse.2024.3368048.