Soft Computing (2021) 25:3195–3223
https://doi.org/10.1007/s00500-020-05373-x

METHODOLOGIES AND APPLICATION

Improving security using SVM-based anomaly detection: issues and challenges

Mehdi Hosseinzadeh[1,2] · Amir Masoud Rahmani[3] · Bay Vo[4] · Moazam Bidaki[5] · Mohammad Masdari[6] · Mehran Zangakani[7]

Published online: 17 October 2020

Abstract

Security is one of the main requirements of the current computer systems, and recently it gains much importance as the number and severity of malicious attacks increase dramatically. Anomaly detection is one of the main branches of the intrusion detection systems which enables to recognize the newer variants of the security attacks. This paper focuses on the anomaly detection schemes (ADS), which have applied support vector machine (SVM) for detecting intrusions and security attacks. For this purpose, it first presents the required concepts about the SVM classifier and intrusion detection systems. It then classifies the ADS approaches and discusses the various machine learning and artificial intelligence techniques that have been applied in combination with the SVM classifier to detect anomalies. Besides, it specifies the primary capabilities, possible limitations, or advantages of the ADS approaches. Furthermore, a comparison of the studied ADS schemes is provided to illuminate their various technical details.

Keywords SVM · Multiclass SVM · Anomaly intrusion detection · Feature selection · Security · PCA

Communicated by V. Loia.

✉ Bay Vo
vd.bay@hutech.edu.vn

Mehdi Hosseinzadeh
hosseinzadeh.m@iums.ac.ir

Amir Masoud Rahmani
rahmani@srbiau.ac.ir

Moazam Bidaki
mBidaki@Iau-neyshabur.ac.ir

Mohammad Masdari
M.Masdari@Iaurmia.ac.ir

Mehran Zangakani
Mehranzangakany@Gmail.com

[1] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

[2] Health Management and Economics Research Center, Iran University of Medical Sciences, Tehran, Iran

[3] Department of Computer Science, Khazar University, Baku, Azerbaijan

[4] Faculty of Information Technology, Ho Chi Minh City University of Technology (HUTECH), Ho Chi Minh City, Vietnam

[5] Computer Engineering Department, Urmia Branch, Islamic Azad University, Urmia, Iran

[6] Department of Computer Engineering, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran

[7] Afagh Higher Education Institute, Urmia, Iran

1 Introduction

Almost all TCP/IP layers are vulnerable to some kinds of malicious behaviors and security attacks, which may be conducted by internal or external attackers (Yan et al. 2015; Singh et al. 2016). However, network hacking and attacking methods are evolving every day to keep security pressure on the computing technologies and networks such as the Internet of Things (IoT) (Qi et al. 2017; Alaba et al. 2017), wireless body area networks (WBANs) (Yessad et al. 2018; Masdari et al. 2017), eHealthcare systems (Yaseen et al. 2018; Masdari and Ahmadzadeh 2017), and cloud computing (Ghomi et al. 2017; Masdari and Zangakani 2019).

Intrusion detection and handling are critical components of the security infrastructure of organizations, aimed to inspect events happening in the system/network for any violation of the security policies (Tang et al. 2018; Yuan et al. 2010). Generally, intrusion detection solutions, according to their detection method, can be categorized into anomaly detection (Peng et al. 2018), signature/misuse detection (Sallay et al. 2013; Subbulakshmi et al. 2011; Teng et al. 2018; Peddabachigari et al. 2007; Qazanfari et al. 2012; Tang et al. 2010; Senthilnayaki et al. 2015; Heba et al. 2010; Nskh et al. 2016), and hybrid IDS schemes (Mulay et al. 2010). The signature-based schemes (Bostani and Sheikhan 2017; Khreich et al. 2017; Jiang and Yasakethu 2013; Laamari and Kamel 2014; Bamakan et al. 2016; Ning and Jianhua 2012; Mehmod and Rais 2016; Zhang et al. 2012; Kabir et al. 2018; Wang et al. 2017; Ganapathy et al. 2012; Hu et al. 2014) try to find intrusions by only relying on the signatures of the previously known threats. However, they cannot deal with the emerging attacks and encrypted traffics and require a timely updating of their signature database, which is an arduous task. On the other hand, anomaly intrusion detection methods create profiles of normal activities and consider any deviation from them as intrusions (Liao et al. 2013). They can detect new attacks (Wressnegger et al. 2013), but finding all normal behaviors is difficult. The third category of the intrusion detection schemes applies both of the before mentioned, signature detection and anomaly detection methods to benefit from both of them. The hybrid schemes' performance depends on their architecture and the integration method of the applied intrusion detection methods. This article is mainly focused on the anomaly detection schemes and the hybrid intrusion detection systems.

Generally, anomaly detection is the process of recognizing any possible abnormal data items that are not according to the normal patterns of other data in a system (Chandola et al. 2009). In general, data anomalies can be classified into the following categories (Ahmed et al. 2016):

- Point anomalies: A single data instance is considered to be anomalous regarding the rest of the data.
- Contextual/Conditional anomalies: Some data may be anomalous in some fields, while they maybe not in the others.
- Collective anomalies: A single data instance may not be anomalous by itself, but its occurrence with other data instances may be considered anomalous.

Anomaly detection-based intrusion detection is an active research area in the intrusion detection context, and many techniques ranging from machine learning (Zaman and Lung 2018; Wani et al. 2019; Anton et al. 2018) and artificial intelligence to pattern recognition are benefited in the literature to deal with it (Saied et al. 2016; Gautam and Om 2016; Feng et al. 2019; Muna et al. 2018; Sindhu et al. 2012; Mazini et al. 2018; Hodge and Austin 2004). SVM is an interesting machine learning method that can be used for binary classification of the multi-dimensional labeled data. For this purpose, it uses some training data points known as support vectors for finding one or more optimal hyperplanes (Patel et al. 2013). Furthermore, for the classification of the nonlinear data, SVM applies various kernel functions (Abraham et al. 2007) that map data to a higher dimension for a better classification. Several types of the SVM classifiers are introduced in the literature, such as one-class SVM, radial SVM, least-square SVM, multiclass SVM, and so on (Bamakan et al. 2016; Al-Qatf et al. 2018; Liu et al. 2006). The SVM classifier is extensively utilized to deal with the intrusion detection problem (Amraee et al. 2018; Yang et al. 2019; Khamis et al. 2020; Rasheed and Tang 2019; Cid-Fuentes et al. 2018). However, despite several survey articles, such as Zarpelão et al. (2017), Modi et al. (2013) and Elshoush and Osman (2011), which are presented in the literature to study the intrusion detection schemes, there is not an in-depth survey paper to investigate SVM-based ADS schemes and focus on their achievements and techniques.

To this end, this paper presents an extensive survey and taxonomy of the different SVM-based ADS schemes. It first provides the basic concepts regarding SVM and intrusion detection. It then classifies the SVM-based ADS schemes regarding the type of the SVM classifier applied in them. Afterward, it summarizes the main contributions and capabilities of the SVM-based ADS schemes and illuminates their advantages and any possible shortcomings.

Besides, each category of the studied ADS schemes compares their utilized datasets, evaluation metrics, kernel functions, and feature extraction methods. Finally, the leading techniques and approaches employed in the SVM-based ADS strategies are discussed to highlight future

research areas. According to our investigations, this is the first review paper that aims to investigate the SVM-based ADS schemes. Our article's main contributions can be listed as follows:

- Introducing key concepts and essential knowledge about SVM classifier, intrusion detection, and anomaly detection.
- Classification of the studied approaches, regarding the type of SVM classifier and other machine learning techniques.
- Providing a systematic review of the ADS schemes, illuminating their significant contributions, and limitations.
- Comparison of the studied ADS approaches to highlight their applied workload datasets, simulators software, and evaluated metrics.
- Highlighting the open challenges in the anomaly detection context.

The remainder of this survey article will be as follows: Sect. 2 presents the research methodology applied in this paper, Sect. 3 describes the SVM classifier's properties, and Sect. 4 discusses the anomaly-based intrusion detection. Section 5 puts forward a classification and survey of the ADS frameworks, while Sect. 6 presents a comparison of the various features of the investigated schemes. At last, Sects. 7 concludes the article and discusses future research in the anomaly detection area. Table 1 presents the acronyms applied in the remaining of our review article.

Table 1 Abbreviations and acronyms

Abbreviation	Description
ACO	Ant colony optimization
ADS	Anomaly detection system
ANN	Artificial neural network
BSM	Basic security module
DR	Detection rate
FNR	False-negative rate
FPR	False-positive rate
GA	Genetic algorithm
LDA	Linear discriminant analysis
OAA	One against all
OAO	One against one
OC-SVM	One-class SVM
PSO	Particle swarm optimization
RBF	Radial basis function
SVM	Support vector machine
PCA	Principal component analysis
SOM	Self-organizing map

2 Research methodology

This section supplies the research methodology applied in this systematic literature review (Tian et al. 2018). It also discusses how the applied reference papers employed in this article, are searched, selected, and finally refined. This survey only uses the articles which have been introduced in the main libraries of Table 2. At first, for searching the SVM review papers in the ADS context, these search terms are used in the Google scholar:

- allintitle: survey anomaly intrusion detection SVM
- allintitle: review anomaly intrusion detection SVM
- allintitle: survey anomaly intrusion detection support vector machine
- allintitle: review anomaly intrusion detection support vector machine

However, we found no review paper satisfying these conditions. For finding the survey articles in the intrusion detection context, the following search strings are applied in the Google scholar:

- allintitle: survey intrusion detection
- allintitle: review intrusion detection
- allintitle: overview intrusion detection

By using these terms, some survey articles are found and referenced in the introduction section. Also, for finding original research articles in the SVM-based intrusion detection context, the following strings are searched:

- intrusion detection SVM
- intrusion detection Support vector machine

The results found by these searches are screened, and documents such as thesis, patents, and invalid papers that were not from Table 2's publishers, are removed. The remaining articles are classified as anomaly detection schemes and misuse detection schemes. Figure 1 exhibits the percentage of the SVM-based ADS schemes, which have been published since 2010. As shown in this figure, fewer SVM-based misuse intrusion detection schemes have been proposed to handle intrusions, and most of these schemes are devoted to the anomaly detection context. Using this strategy, we found several papers for conducting the study, which will be reviewed in the next section. Also, the initial search is refined by a quality assessment checklist, which contains the following questions:

- Is the research methodology put forward in this article?
- Does the research methodology fulfill the requirements of the problem under study?
- Is this study analyzed as it should be?

Table 2 The applied publishers

Publication	URL address
IEEE Xplore	http://www.ieee.org/web/publications/xplore/
ScienceDirect—Elsevier	http://www.elsevier.com
SpringerLink	http://www.springerlink.com
Sage	http://journals.sagepub.com
Google Scholar	http://Scholar.google.com
ACM	http://www.acm.org
Wiley	http://onlinelibrary.wiley.com
Inderscience	http://www.inderscience.com
Emerald	http://www.emeraldinsight.com
Hindawi	https://www.hindawi.com
OXFORD academic	https://academic.oup.com

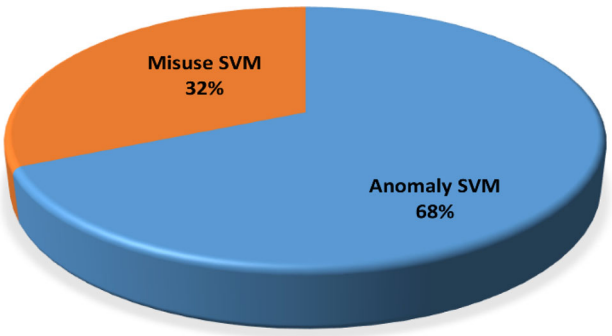

Fig. 1 Percentage of the SVM-based misuse detection and anomaly detection schemes

The number of SVM-based ADS schemes presented in the literature from 2010 up to 2019 is shown in Fig. 2. As depicted in this figure, the number of these schemes is increasing, and as a result, this context can be considered as an active research area in the security literature. Table 3 exhibits the main research questions of this paper, which will be answered in the remainder of this paper.

3 SVM

SVM is one of the supervised learning models used for regression analysis and classification problems. It uses the nearest training data points of each data to build one or more hyperplanes to classify high-dimensional data. It maximizes the margin between them as much as possible (De la Hoz et al. 2015). The one-class SVM tries to compute a hyperplane that a pre-specified fraction of the training data samples fall beyond that, while the hyperplane has the most margin to the origin (Fig. 3).

When data are linearly separable, the SVM's hyperplanes can be computed efficiently. Otherwise, as shown in Fig. 4, data should be mapped using a kernel function into space with a higher dimension in which a hyperplane can be easily detected to separate them (Sindhu et al. 2012). To mitigate the training time and enhance the SVM classifier performance, some of the SVM-based ADS schemes have applied the existing kernel functions such as radius base function (RBF), and some others have developed new ones (Yi et al. 2011). Also, some of the schemes have employed multiple kernel functions, while some others have only used one kernel function (Song et al. 2011).

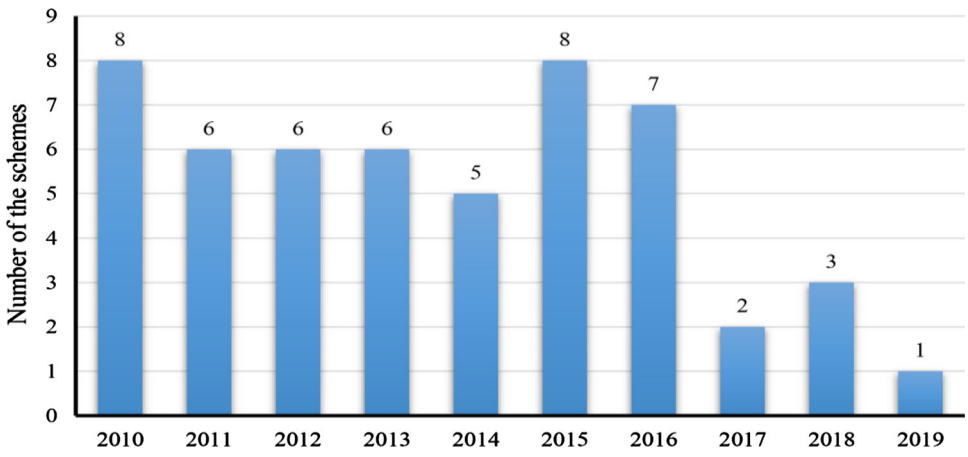

Fig. 2 Publication year of the SVM-based ADS schemes

Table 3 Research questions

Index	Question	Reason
1	Which kind of SVM classifier is applied by each IDS scheme? How is it adapted to detect different types of attacks?	ADS approaches have applied various types of SVM to classify the traffic data and events happening in the system. Since SVM is a binary classifier, it is interesting to know that how it is applied to solve the multiclass anomaly ADS problems
2	Which kinds of kernel functions are applied in the investigated ADS schemes?	ADS schemes apply different kernel functions for the classification of nonlinear data, and this question illuminates further technical details about each system
3	Which classifiers are utilized in combination with SVM in the investigated ADS schemes?	For improving the classification performance, some schemes have applied other classifiers in combination with SVM. By having information about these classifiers, future research area can be recognized. Also, the achieved results can be used to enhance the existing schemes
4	Which techniques are used in each ADS scheme for the selection of features?	Considering a high number of features in the ADS datasets, achieving an optimal feature set is very important
5	Which datasets are applied in each ADS scheme?	Identifies the key datasets which are used in the ADS context. This is helpful in designing new ADS solutions, especially when the authors want to evaluate their scheme on multiple datasets
6	Which metrics are used to analyze each ADS scheme?	It indicates the metrics mainly used in the anomaly detection context to verify the improvements made by each scheme
7	Is k-fold cross-validation conducted to evaluate the performance of the studied ADS schemes?	Cross-validation is very important for proper evaluation of the SVM-based ADS schemes. Since by randomly selecting some part of a dataset for training and using the other part for testing, we may incorrectly achieve a high level of performance

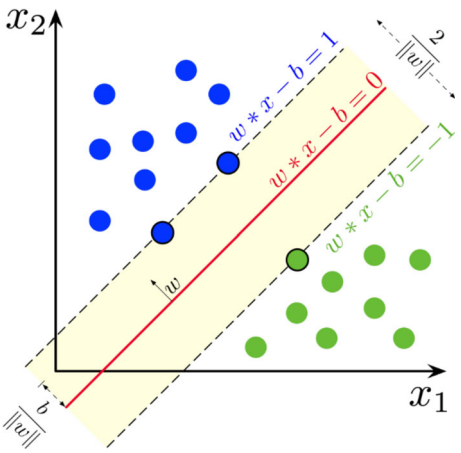

Fig. 3 SVM classifier

Generally, two main types of SVMs are multiclass SVMs and binary SVMs, in which the latter can be further classified into linear and nonlinear classes. Also, multiclass SVMs can be classified into the one against all SVM (OAA-SVM), one against one SVM (OAO-SVM), and DAGSVM. In the first case or the OAA-SVM, only one binary SVM will be used for each class of data, to separate class members from other classes, and a data item can be placed in a class if its SVM accept it and remaining SVMs do not accept it. Although this method can be suitable for

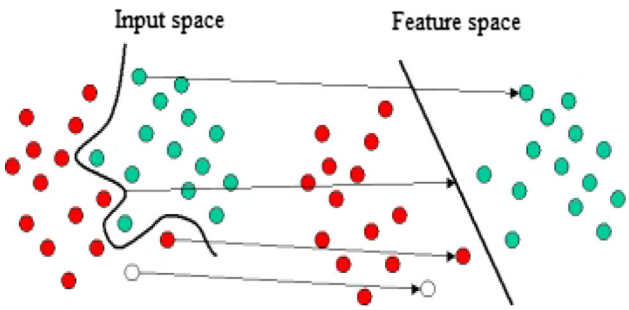

Fig. 4 Mapping to a higher dimension

tightly clustered data, it may leave some parts of the feature space unclassified, especially in cases in which a data item is accepted by more than one SVM or when all SVMs have rejected it. The OAA-SVM applies the majority voting method for performing classification. On the other hand, OAO-SVM uses one SVM classifier for each pair of classes and applies fewer SVMs, in which each SVM will be trained based on data of its two categories. But, this method can suffer from the limitations of the binary SVMs. At last, in the DAGSVM method, nodes are binary SVMs, and each classifier builds $k*(k-1)/2$ hyperplanes and applies a graph traversal method for testing.

4 Anomaly-based intrusion detection

This section provides essential background concepts regarding the anomaly detection and handling approaches.

4.1 Challenges

Anomaly detection is not an easy task, and there are several challenging issues to accurately recognizing anomalies, which can be listed as follows:

- Each data point in a high-dimensional dataset may appear to be anomalous, known as data-snooping bias.
- Datasets required for training and validation of ADS are not publicly available, or they do not exist at all.
- Datasets may contain noisy data, which tends to be an anomaly.

4.2 ADS location

Based on the location of anomaly detectors, ADS schemes can be categorized as host-based and network-based approaches. Typically, the host-based schemes act independently and gather system events to detect any possible anomalies. Besides, they do not require any specialized hardware (Zhou et al. 2010). On the other hand, network ADS approaches continuously sniff the inbound and outbound traffic packets and analyze them for detecting any possible anomalies. Also, they often need special-purpose hosts or hardware (Masdari and Jalali 2016). Besides, network ADS schemes can be used in a hierarchical topology or can be integrated with the host-based ADS schemes to increase their effectiveness.

Furthermore, regarding the data that the host-based ADS schemes require to operate, they can be categorized into white-box schemes, black-box schemes, and gray-box schemes. White-box systems perform a statistic evaluation of the source codes, but black-box schemes try to find the control-flow attacks using the system call sequences. Also, gray-box approaches extract runtime information from the monitored process when a system call is invoked.

4.3 ADS vulnerabilities

ADS approaches are not immune to the security attacks themselves, and various attacks can fool them. For instance, in the mimicry attacks, the attacker monitors the system calls' order, to create a series of system calls that can appear normal to fool the ADS. Also, to prevent the detection of attacks, the attackers may attempt to find the required settings such as specific thresholds that can set off the alarms.

4.4 ADS steps

Figure 5 depicts the general architecture of a network ADS which performs the following steps in the anomaly detection process:

- Capturing transferred data packets from the monitored network or using existing datasets for the intrusion context.
- Data preprocessing.
- Extracting or finding a set of features describing the events happening in the monitored network or host.
- Training the SVM classifier using normal activities and profiles.
- Classifying the network traffic and system events by using the learned models.
- Evaluating the classification results using appropriate metrics.
- Producing the required response according to the SVM results.

4.5 Datasets

Online ADS schemes capture the required data directly from the monitoring environment (a network or a host), while offline ADS schemes use the logged events, in the form of a dataset. For instance, DARPA is one of the datasets that have been used for many years in the ADS literature (Cheng et al. 2012) and consists of two primary data: TCPDUMP and BSM. TCPDUMP is gathered from a simulated LAN and contains data packets. Besides, BSM (basic security module) has the logs of system calls' executions (Gautam and Om 2016). The training data of the first and third weeks of DARPA, which have not any security attacks, can be used to support the training of the ADS schemes.

On the other hand, the second week of the DARPA training data has a subset of attacks to be used in misuse detection. Also, regarding the shortcomings of the DARPA dataset, other datasets such as KDD Cup and NSL-KDD are created from it by removing the redundant records. The KDD Cup dataset contains 41 features, in which nine features are nominal variables, and 32 features are continuous variables. It has a label that specifies that the record is normal or attack type. KDD Cup consists of five million training records and two million testing records. The DARPA-based datasets contain the following types of attacks:

- DoS Attack: The attacker overloads the victim with many requests, exhausting its resources. As shown in Fig. 6, this attack can be conducted against various network layers, different protocols, and applications.

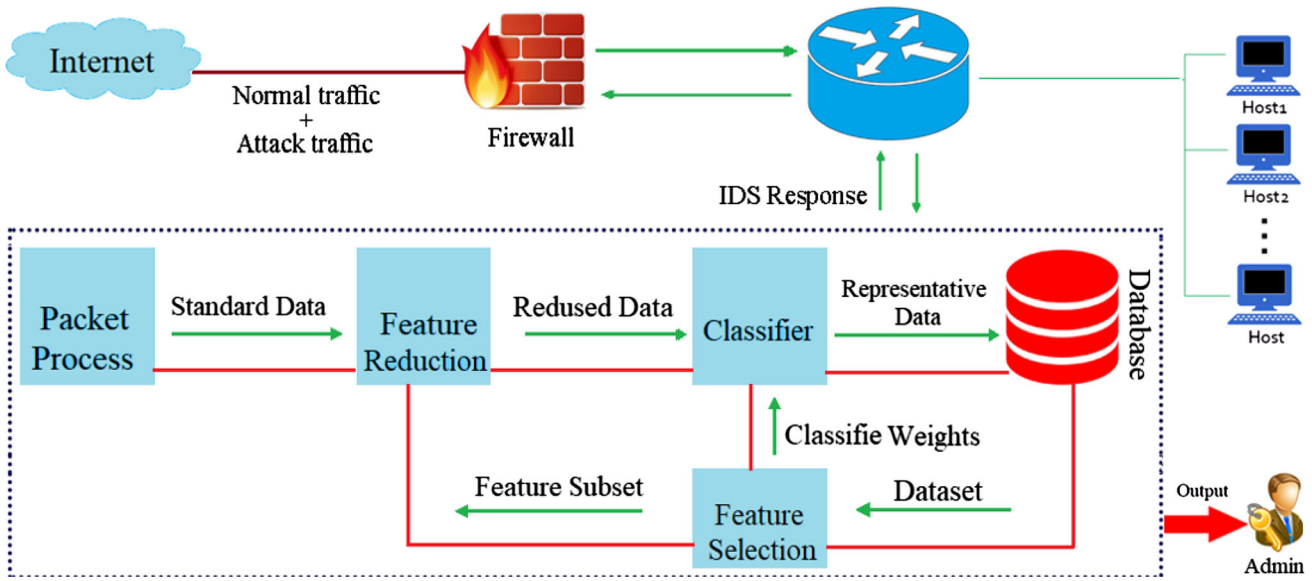

Fig. 5 Architecture of a network-based ADS security system

- U2R Attack: Unauthorized access into a system as an admin user.
- R2L Attack: The intruder uses system vulnerability to act as a regular user.
- PROBE Attack: The intruder scans the network computers to find their vulnerability.

4.6 Dimension reduction

For dealing with high-dimensional intrusion datasets, feature selection/extraction (Nguyen et al. 2010; Gong et al. 2011) should be conducted to remove irrelevant, redundant, and noisy features (Guo et al. 2010). Feature selection techniques can lessen the dataset's features, keeping a subset of principal, and related attributes aimed to enhance the ADS performance (Kabir et al. 2018). In general, these methods are categorized as filter-based (Ambusaidi et al. 2016), wrapper-based (Sindhu et al. 2012), and hybrid approaches. The filter-based techniques such as information gain, principal component analysis (PCA) (Yi et al. 2011), entropy, and so on focus on the intrinsic properties of the features. On the other hand, wrapper-based approaches apply a classifier to evaluate a subset of features. Even though this method is computationally expensive than the filtering method, it is more dependable (Saied et al. 2016).

4.7 ADS training

An SVM classifier should be trained using the training part of a dataset, and then it should be evaluated using the testing part of that dataset. To further validate the SVM, *N*-fold cross-validation can be used in which data samples should be split into N equal segments. Then, performance evaluations should be conducted in the N phases, in each of which, $(N-1)$ folds of data will be applied for training purposes, and one remaining fold will be employed for the testing.

Regarding the learning method applied in the intrusion detection process, the ADS schemes can also be classified as unsupervised, semi-supervised, and supervised learning methods (Khan et al. 2012). The unsupervised anomaly detection techniques detect anomalies using unlabeled data. Although they often have a high detection rate, they may suffer from a high false-positive rate (FPR) (Sallay et al. 2013). The supervised learning techniques require a fully labeled dataset for training classifiers, which may not be available in some context. Also, semi-supervised ADS can employ partially labeled datasets. Generally, supervised ADS approaches for intrusion detection try to find the relationships among the features and their classes (Wang et al. 2008). But, there is an initial training period for ADS, and it must be frequently retrained, to apply new data items in the training data for recognizing new normal behaviors. Most ADS methods use non-incremental learning algorithms. But in this case, with the accumulation of new data samples, their training time increases, and they will have difficulties in tuning themselves with changing situations. On the other hand, incremental learning can rapidly learn from new samples and meet the requirements of real-time ADS.

4.8 ADS evaluation metrics

The outcome of an ADS scheme in the testing phase can be as follows:

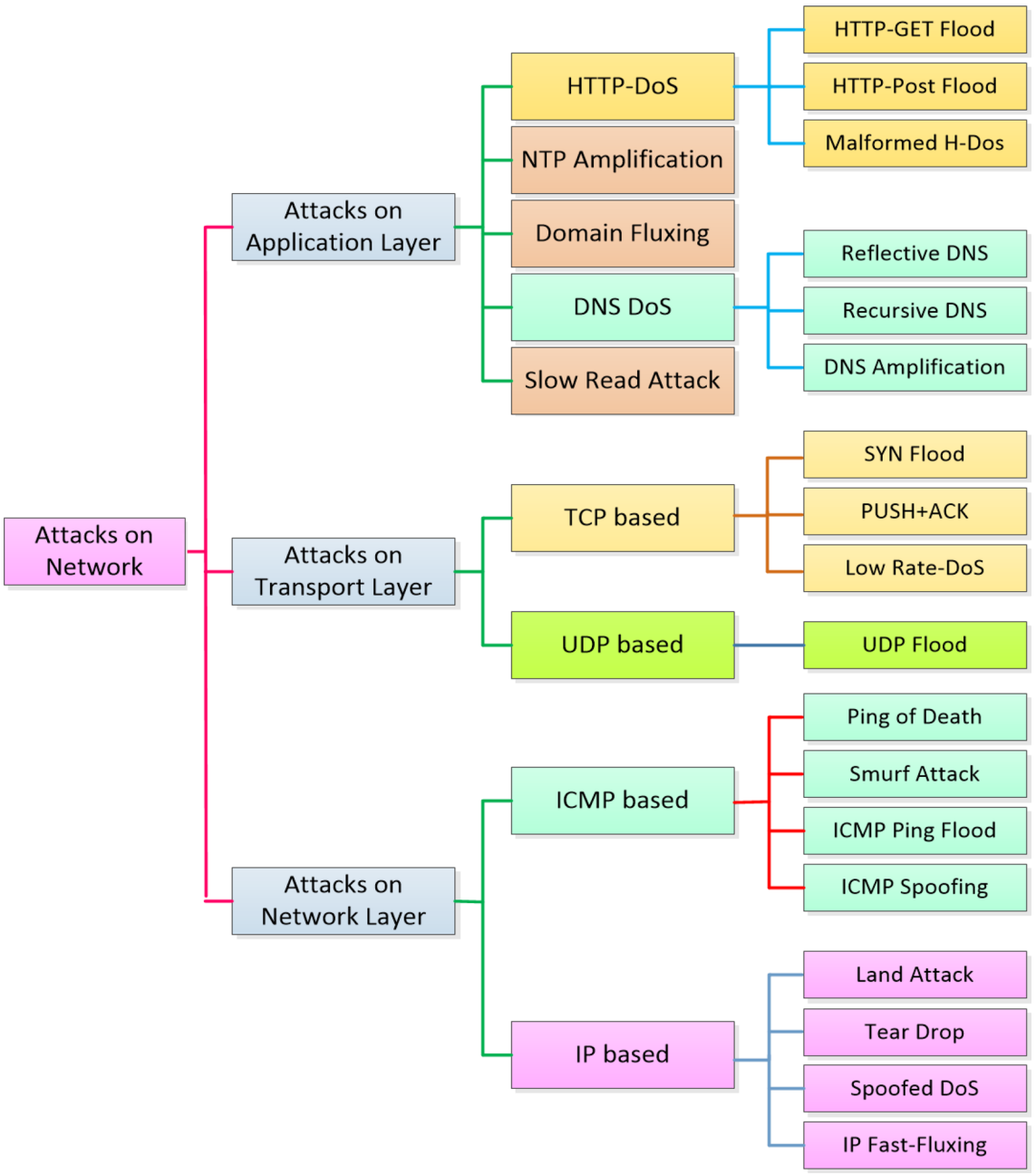

Fig. 6 Security attacks on the TCP/IP (Masdari and Jalali 2016)

- True Negatives (TN): Normal traffic is correctly recognized as normal traffic.
- True Positives (TP): Attack traffic is correctly detected as attack traffic.
- False Positives (FP): Normal traffic is misclassified as attack traffic.
- False Negatives (FN): Attack traffic is falsely misclassified as normal traffic.

Figure 7 indicates the metrics which have been applied in the performance evaluation of the ADS schemes. To have an efficient ADS scheme, both FP and FN should be reduced, while TP and TN should be maximized. However, the ADS approaches often try to balance the ability to detect new attacks and generating a low rate of FP. The proposed SVM-based ADS schemes can have one of the outputs specified in Fig. 8. In the first case, which happens in the simplest form of the ADS, traffic is only classified into the normal and abnormal categories, without specifying the type of attacks. In the second case, not only can the ADS detect unusual traffic, but it also recognizes the kind of attacks conducted against the network. In the third case,

01	$Specificity = \frac{TN}{TN + FP} (\times 100\%)$
02	$Recall - \frac{TP}{TP + FN} (\times 100\%)$
03	$Percision = \frac{TP}{TP + FP} (\times 100\%)$
04	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} (\times 100\%)$
05	$True\ Negative\ rate = \frac{TN}{TN + FP}$
06	$True\ Positive\ rate = \frac{TP}{TP + FN}$
07	$False\ Positive\ rate = \frac{FP}{TN + FP}$
08	$F - Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall}$

Fig. 7 ADS evaluation metrics

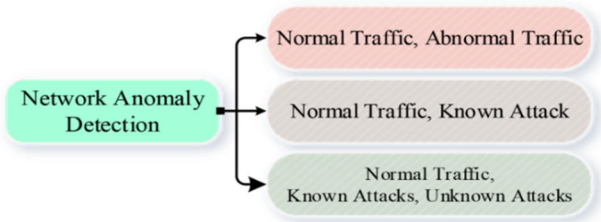

Fig. 8 ADS outputs

the ADS determines the known class of attacks and also finds new attacks.

4.9 ADS response

Also, regarding the final ADS response, they can be classified as passive or active approaches. In the passive mode, ADS logs the events and notifies the security or network administrator. In contrast, in the active mode, in addition to sending the required alerts, it can conduct necessary actions to block the intruders.

5 Study of the proposed SVM-based ADS schemes

This section shall briefly survey the state-of-the-art SVM-based ADS schemes introduced in the literature since 2010. Figure 9 exhibits the classification of these ADS approaches regarding the type of utilized SVM classifier and other techniques. As shown in this figure, different types of SVM are applied by outlined systems to benefit from their capabilities, for example, to conduct multiclass classification. Moreover, in some approaches, to boost the detection rate and accuracy of the ADS, SVM parameters and its applied kernel functions are trained through using meta-heuristic algorithms. Also, to increase the detection rate, some of the ADS approaches have exploited the SVM with other classifiers such as decision trees, ANN, and naïve byes. Furthermore, some of the outlined schemes have applied feature extraction methods or utilized meta-heuristic algorithms for feature selection. At last, some other ADS approaches have tried to handle the imbalanced datasets, which makes the SVM training process imperfect.

5.1 One-class SVM-based ADS schemes

This subsection focuses on the ADS schemes, such as Dong and Peng (2018) and Ioannou and Vassiliou (2019), which have employed the OC-SVM classifier. The authors in Renjit and Shunmuganathan (2011) have provided effective ADS and used a local agent in the hosts to gather data from its system and employed an SVM classifier to find anomalies in this data. When the agent detects that its system is under attack, it disconnects the system from the network. The local agent in each host applies a mobile agent to collect data from the neighbor nodes before it allows the system to send data to its neighbors. As an advantage, the local agent can remove the local system from the network when the system is under security attacks. Figure 10 shows the anomaly detection process applied in this scheme.

In Agarwal and Mittal (2012), the authors try to detect network anomalies by presenting a hybrid ADS approach by using SVM and the entropy. DARPA dataset is used to evaluate this method. In this scheme, the first normalized entropy of features is computed then they are used to train the SVM model for classifying the network traffic. The authors indicated that this scheme could handle attack traffic with high detection accuracy and fewer false alarms.

Xie and Zhang (2012) introduced an SVM-based anomaly ADS scheme, which mitigates the number of required training samples as well as training time. Before the classifier training, data should be normalized, and all types should be converted into binary data. Afterward, it extracts the relevant intrusion features and converts the raw data into vectors, which will be stored in the vector library. The authors have evaluated this scheme on the KDD Cup dataset and showed that it could achieve a higher detection rate than the basic SVM.

In Zhang et al. (2015), the authors have provided an anomaly handling scheme using OC-SVM that applies the

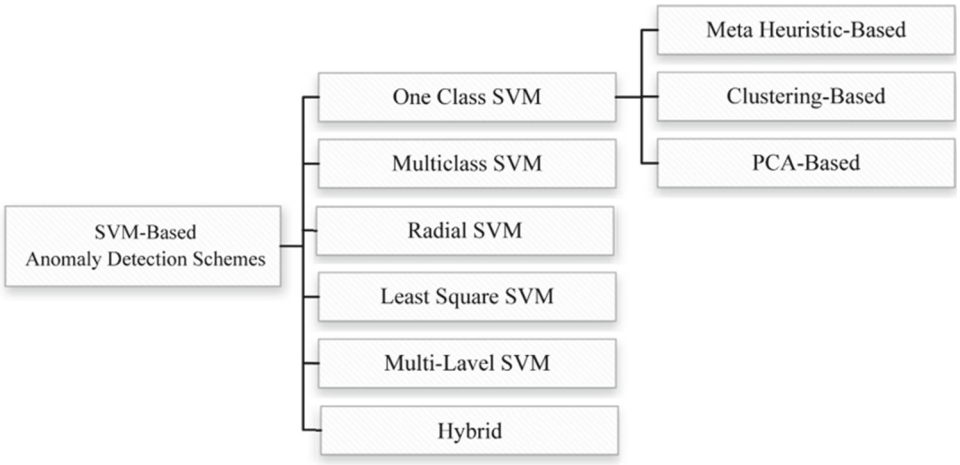

Fig. 9 Taxonomy of the SVM-based ADS schemes

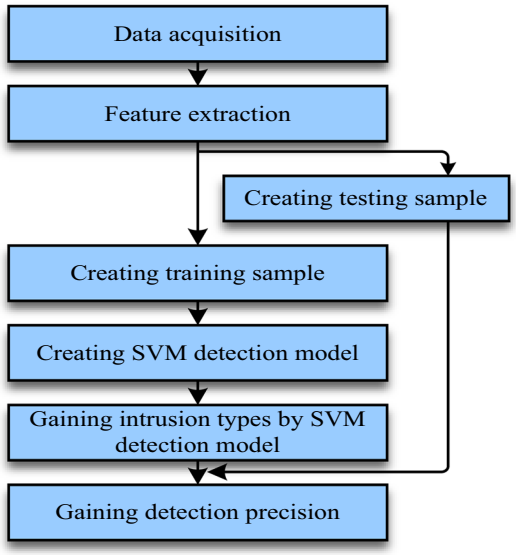

Fig. 10 Intrusion detection process by SVM in Renjit and Shunmuganathan (2011)

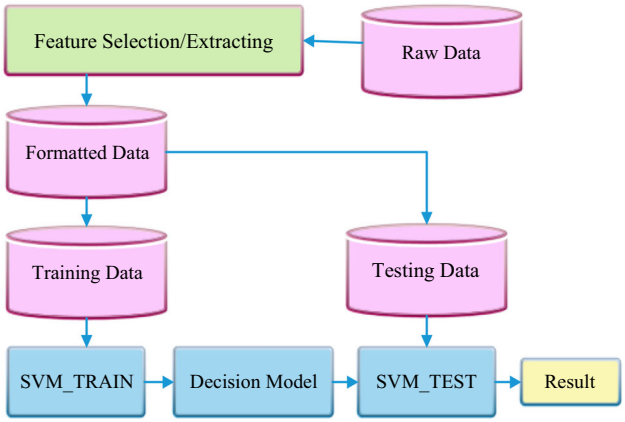

Fig. 11 OC-SVM-based ADS model in Zhang et al. (2015)

data of the network's normal connections for training purposes. This ADS model is compared with the C-SVM and probabilistic neural network classifiers and applied the radial basis function (RBF) kernel in them. This anomaly detection model is evaluated on the KDD Cup dataset, and the authors have selected a random sample (3%) of normal data records in the raw training data. To test the ADS model's capability to recognize various types of attacks, they randomly chose different kinds of records in the testing data. Nonetheless, some attacks such as U2R and R2L were selected because of their low number of records in the KDD Cup. Besides, they achieved higher detection rates and improved precision, recall, and F-value metrics. Figure 11 shows the architecture of this ADS model.

In Ramamoorthi et al. (2011), the authors proposed an anomaly detection mechanism to detect DDoS attacks using ESVM or enhanced SVM with string kernels. Besides, they use the normal behaviors of users for training and apply normal and attacks traffic data for testing the ESVM. As an advantage, this scheme can detect network layer and application layer-based DoS attacks with high accuracy.

In Khreich et al. (2017), the authors employed OC-SVM to introduce anomaly detection-based ADS. They train OC-SVM on the fixed-size feature vectors achieved from the system call traces. They also apply a window of N system calls and for each window associate a feature vector. Besides, they compute the frequency of the n grams, which are a sequence of n items from a system call trace. They structured variable n grams into the fixed-size vectors and assigned weights to them, considering their frequency. Afterward, the fixed-size feature vectors are used to train the OC-SVM with different kernels.

In Amraee et al. (2018), the authors tried to deal with the anomaly detection in the distributed networks by introducing a fully distributed SVM-based approach. They solved the SVM optimization using an efficient gradient-based algorithm to obtain an estimate for the hyper-planes parameters. They achieved high performance with low

complexity and communication loads in the anomaly detection process. They applied synthetic and real datasets in their simulations and indicated that their distributed approach increases the nodes' performance the same as a central node with access to all required data.

The OC-SVM classifier is sensitive to the noisy data in the training step, which causes a shift in the boundary of OC-SVM toward the noisy data. For dealing with this issue, in Tian et al. (2018), an ADS solution denoted as Ramp-OCSVM is introduced, which, instead of the Hinge loss function, applies the ramp loss. In the performance evaluations, the first experiment was conducted on a two-dimensional synthetic dataset, the second one is carried out on a few UCI datasets, and at last, UNSW-NB15 and NSL-KDD are used. The results showed that the Ramp-OCSVM outperforms other models based on detecting anomalies.

In Zhang et al. (2019), the authors proposed an SVM-based ADS which extracts and optimizes the training features. It applies the Kullback–Leibler divergence and cross-correlation calculated by the data and control plane traffics for SVM training. This training approach can increase the detection accuracy of their scheme. They also validated its performance according to realistic traffics. They exhibited that their ADS solution can recognize short-duration attacks in the network.

Table 4 exhibits the comparison of the OC-SVM-based ADS schemes. It compares the evaluated metrics, applied ADS datasets, and kernels utilized in the SVM classifier.

The ADS approach provided in Miao et al. (2018), formulated a distributed online OC-SVM classifier and used a decentralized cost function. To prevent sending data, they used an approximate function, instead of the kernel and minimized the cost functions using stochastic gradient descent. They carried out experiments on the real and synthetic datasets and indicated that their methods could have low misdetection and high TPR. Besides, they claimed that their approach achieves good accuracy with short running time and incurs lower overheads in terms of memory and CPU.

In Tang et al. (2019), the authors proposed a DDoS detection approach based on V-SVM. This scheme tries to normalize the feature data, and by using PCA, it reduces the data dimension. It applies a kernel and uses the parameter V control support vector, and the number of error vectors establishes a V-SVM-based DDoS attack classification model to detect attacks. The simulations' results exhibited that this ADS method can improve the accuracy, reduces the FNR while ensuring the timeliness and stability of the achieved results.

5.1.1 PCA-based ADS schemes

PCA or principal component analysis is known as an effective filter-based feature selection method that converts a set of observations of the correlated variables to values of linearly uncorrelated variables, denoted as principal components (De la Hoz et al. 2015). This section focuses on the PCA-based anomaly ADS schemes (Reddy et al. 2016; Liu et al. 2010). For instance, in Heba et al. (2010), the authors presented an ADS approach that used PCA on the NSL-KDD dataset to choose the optimal feature subset and reduced the data dimensions from 41 to 23. The authors validated the effectiveness of their ADS through simulations performed K-fold cross-validation on the NSL-KDD dataset. As an advantage, this scheme can improve the testing and training processes.

Also, the ADS scheme proposed in Thaseen and Kumar (2014) integrates PCA and the SVM, and it optimizes the parameters of the RBF kernel function. In this scheme, SVM builds the classification model according to the training data achieved from the PCA. It lessens the required time for testing and training and enhances the accuracy of ADS. For evaluation purposes, cross-validation is carried out on the KDD Cup dataset. Concerning the malicious minority behaviors like U2R and R2L, creating the equal size of data samples for training and testing, can improve classification accuracy. Furthermore, since the classifier input needs a reduced feature set, low resources will be used, and minimum training and testing overhead will be incurred. However, further performance analysis is necessary to evaluate the capabilities and merits of this solution.

The ADS approach in Chu et al. (2019) presents an ADS for early detection of the advanced persistent threat (APT) attacks, which are organized by a group of hackers. This scheme applies PCA for feature sampling, to enhance the detection accuracy. This ADS solution is evaluated on the NSL-KDD dataset. The results indicated that the SVM classifier has the highest detection rate, and this scheme can be used to reduce the APT attacks' impact.

Table 5 exhibits the comparison of the PCA and rough set ADS schemes. It mainly compares the utilized evaluation metrics, datasets, kernels, and feature extraction methods in the outlined ADS schemes.

5.1.2 Meta-heuristic-based ADS schemes

Different meta-heuristic algorithms are benefited in combination with the SVM classifier to improve ADS capabilities and effectiveness (Enache and Sgarciu 2014). Often, meta-heuristic algorithms are incorporated for the training of the SVM classifiers, tuning their optimal

Table 4 Comparison of OC-SVM-based ADS schemes

Schemes	Evaluated metrics	ADS datasets	Kernel functions	Feature extraction	Advantages/limitations
Renjit and Shunmuganathan (2011)	Detection Rate, FPR			PCA, LDA	Distributed ADS scheme, Considering the practical scenarios for ADS, Comparison with two other ADS solutions
Xie and Zhang (2012)	Detection rate, Fall-out rate, Miss rate	KDD Cup	Gaussian		Selecting learning vectors in the training step, reducing the training samples, reducing the training time
Zhang et al. (2015)	Detection Rate, Recall	KDD Cup	Gaussian	Mutual Information	
Ramamoorthi et al. (2011)	Classification rate		Linear, Polynomial, RBF, String kernel	Synthetic traffic generation	Using various kernels for classifications, Comparison with multiple kernels, Only detects normal traffics and attack, traffics, not type of the attacks
Khreich et al. (2017)	AUC, FPR, TPR	ADFA-LD	Gaussian, linear		A complete comparison with the neural, networks, STIDE, and HMM
Amraee et al. (2018)	Precision accuracy curve, AUC score, ROC curve	Iris Thyroid Occupancy SMTP Mammography			Providing distributed anomaly detection algorithm, Applies gradient-based training, Comparison with other SVM classifiers
Tian et al. (2018)	Recall, Precision, F1 score, Recall, FPR	NSL-KDD UNSW-NB15	Gaussian		Improves SVM to deal with the noisy data, Extensive comparisons with other types of the SVM classifier, No comparison with different classifiers
Zhang et al. (2019)	TPR, FPR, Precision, F-score, Overall success rate	Real traffic captured from an edge router in King Saud University	RBF	Entropy	
Miao et al. (2018)	Memory consumption CPU time AUC TPR FPR F1	UCI Datasets	Gaussian		Using Distributed Online SVM for ADS, Extensive comparisons and evaluations using multiple datasets, Does not support multiclass classification

Table 5 Comparison of the PCA-based ADS schemes

Schemes	Evaluated metrics	Kernel functions	Datasets	Feature extraction	Advantages/limitations
Reddy et al. (2016)	Accuracy Training Time Testing Time	RBF	DARPA KDD Cup	Rough Set	Simple ADS approach, Needs further evaluations
Liu et al. (2010)	Accuracy FPR	Sigmoid	KDD Cup	Rough Set	Needs further evaluations, Employs binary PSO which is an old optimization algorithm
Heba et al. (2010)	Accuracy	Gaussian nonlinear polynomial	KDD Cup	PCA	Reducing testing time, Reducing training time
Thaseen and Kumar (2014)	Accuracy	RBF	KDD Cup	Chi squared	Reducing testing time, Reducing training time, Increasing accuracy, It can deal with unbalanced datasets

parameters, and for choosing more crucial features from the intrusion detection dataset Enache and Sgârciu (2015a).

An ADS scheme denoted as PSO-OCSVM is proposed, in Shang et al. (2015), which applies the PSO algorithm for the training of the SVM classifier. They established an improved feature vector extraction method and data preprocessing for industrial Modbus TCP protocol. Furthermore, they mitigated the training delay in their ADS model and increased its accuracy. This mechanism trains the anomaly detection model with only the one class of the samples for the anomaly detection. Nonetheless, they have not compared their approach with other ADS schemes, but it outperforms the Grid-OCSVM.

In Enache and Sgârciu (2015b), the authors applied the binary BA to conduct features selection using the wrapper method and to train the SVM parameters. The authors have employed ten-fold cross-validations to evaluate this scheme, and experiments are performed on the NSL-KDD, which indicate some improvements in terms of FPR and detection rate.

In Aslahi-Shahri et al. (2016), wrapper-based ADS is proposed, which employs the genetic algorithm (GA) and SVM classifier to reduce the dataset's features. The feature distribution is done to put four features in the highest priority, another four in the next level of preference, and the other 2 in the last level of priority. The block diagram of this ADS approach is given in Fig. 12. The authors showed that their hybrid scheme could achieve better results regarding the TPR and FPR metrics.

In Enache and Patriciu (2014), the authors employed information gain for feature selection in an anomaly detection model. Also, the authors have employed RBF kernel function with the SVM, and their parameters are adjusted by using the PSO and ABC algorithms. They used the NSL-KDD dataset with the WEKA software. They indicated that the ABC algorithm achieves a better performance than PSO, in terms of metrics such as detection rate, FPR, and accuracy.

In Feng et al. (2014), the authors presented CSOACN, which applied the SVM and clustering using self-organized ACN. They combined the SVM with CSOACNs to use their strengths. Although they have evaluated this scheme using the KDD Cup, further analysis is needed via other benchmark datasets. The block diagram of active learning in this approach is depicted in Fig. 13.

The ADS method, introduced in Al Shorman et al. (2019), uses OC-SVM for providing an unsupervised evolutionary method for botnet detection in IoT. This scheme uses the gray wolf optimization algorithm to detect IoT botnet attacks to optimize the SVM's parameters. For performance evaluations, a dataset with real data is used, and the authors exhibited that their ADS can outperform other approaches by using metrics like TPR, FPR, Geometric-mean, and detection time.

In Bostani and Sheikhan (2017) put forward MI-BGSA, a hybrid wrapper-based ADS approach that uses binary GSA and applies a mutual information approach for filter-based feature selection. It extracts proper features using a multi-objective function to increase the detection rate and reduce FPR. However, this scheme is only evaluated on the NSL-KDD and is not compared with the other ADS approaches.

Table 6 exhibits the comparison of the meta-heuristic SVM-based ADS schemes and compares their applied evaluation metrics, datasets, kernel functions, and feature extraction methods.

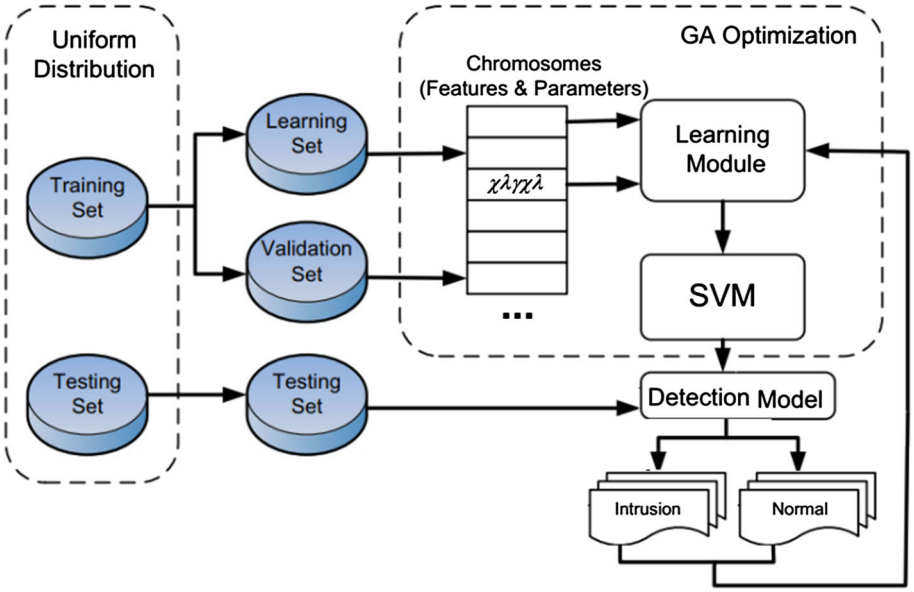

Fig. 12 ADS architecture in Aslahi-Shahri et al. (2016)

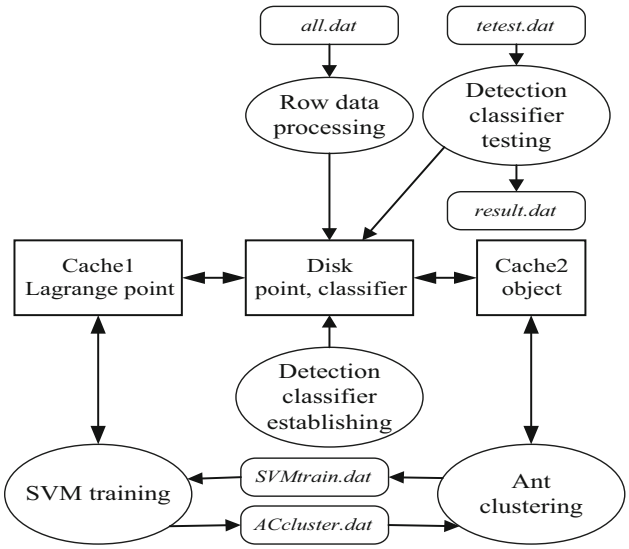

Fig. 13 Active learning of SVM in Feng et al. (2014)

5.1.3 Clustering-based ADS schemes

To handle large datasets and enhance ADS performance, some ADS schemes have exploited clustering algorithms in combination with the SVM. This method reduces the training time of the classification process by providing a reduced and smaller dataset to the classifier. Clustering methods are employed by different ADS schemes such as Emadi and Mazinani (2018) to get higher performance, dealing with the unlabeled data, and handling large datasets. For instance, in Li and Zhao (2011), the authors provided a fuzzy SVM-based ADS model, which has utilized rough set theory for dimension reduction of the KDD Cup dataset. Also, the rough set upper and lower approximation is used to find the relation of features. They reduced the dataset attributes from 41 to 20, and this can improve training and classification performance. This fuzzy scheme benefits from a fuzzy membership function based on the affinity among sample points to decrease the outlier's effects. The authors evaluated their solution in terms of FPR, detection rate, and detection efficiency and achieved better results than an SVM-based ADS solution, which has not applied feature selection. Figure 14 indicates the architecture of this scheme.

Some of the ADS schemes perform clustering of network traffic connections into normal and abnormal classes based on some features. Often, the distance measure to find the similarity of features is the Euclidean distance function. In Sani and Ghasemi (2015), the authors have provided a learning method of an appropriate distance function according to a set of supervision information. This scheme solves a semidefinite optimization problem, to reduce the distance between the similar, and increases it among the different features. They evaluated the performance of this scheme in terms of FPR on the Kyoto2006 + dataset.

The ADS approach proposed in Ashok et al. (2011) applied the SVM classifier along with the K-means method to handle data anomalies. They computed the information measure (IM) by calculating the relation between each attack class and features. Then, they have used the K-means algorithm to remove the redundant features by categorizing five attack clusters. Afterward, they divide each cluster into several triangles based on several features obtained by IM. They have reduced the dimensionality of feature set by calculating the triangular area by simplifying attack class with that cluster. They carried out experiments

Table 6 Comparison of the meta-heuristic ADS schemes

Schemes	Evaluated metrics	Kernel functions	Datasets	Feature extraction	Advantages/limitations
Enache and Sgarciu (2014)	Accuracy, Detection rate Testing Time, FPR	RBF	NSL-KDD, KDD Cup		Bat algorithm is improved with Lévy flights Only is compared with other meta-heuristic algorithms and it should be compared with different classifiers
Enache and Sgârciu (2015)	Accuracy Detection rate FPR	RBF, Polynomial	NSL-KDD, KDD Cup	Information Gain, Entropy	Bat algorithm is improved with Lévy flights Should be compared with other classifiers
(Tian and Gu 2010)	Training Time	Polynomial, Sigmoid	Five datasets from UCI	Mutual Information	Compared with K-NN, K-Means, PCA, and SOM, It does not improve the PSO algorithm which has been applied in the ADS
Lin et al. (2012)	Accuracy		KDD Cup	Information Gain	Further evaluation is needed
Shang et al. (2015)	Detection rate	Sigmoid Gaussian	UCI	Information Gain	Not enough evaluations
Enache and Sgârciu (2015)	Detection rate FPR	RBF	NSL-KDD KDD Cup		Improving binary bat algorithm, No comparison with other classifiers
Aslahi-Shahri et al. (2016)	Recall, Accuracy FPR, Precision ROC, F-measure (Liu et al. 2008)		KDD Cup	PCA	It should be compared with other ADS schemes.
Enache and Patriciu (2014)	Accuracy Training Time Testing Time	RBF	NSL-KDD	Information Gain	Further experiments and evaluations are needed to verify the achieved results No improvement to the applied meta-heuristic algorithms
Feng et al. (2014)	Detection rate Training Time FPR FNR	RBF	KDD Cup		Using active learning SVM for achieving high performance Using CSOACN or clustering based on self-organizing ant colony network It is only compared with other types of SVM, not with other classifiers
Bostani and Sheikhan (2017)	Accuracy Detection rate Execution time FPR	RBF	NSL-KDD	Mutual Information	Conducting comprehensive experiments with other optimization algorithms and feature selection methods, The applied meta-heuristic algorithm is not improved
Al Shorman et al. (2019)	TPR FPR Geometric-mean	RBF	NN-BaIoT		Anomaly detection in the IoT environment which have been focused by fewer schemes, A complete comparison of the proposed scheme

on the KDD Cup dataset and evaluated the FPR and the best detection rate.

In Chitrakar and Chuanhe (2012), the authors benefited from K-means and K-medoADS clustering algorithms with the Naïve Bayes classifier to provide an ADS approach. The need for large samples in this scheme is reduced by using the SVM while keeping the clustering results of the K-medoADS. As an advantage, the authors have conducted evaluations using the Kyoto2006 + dataset, which is newer than the conventional DARPA-based datasets. They evaluated the performance of their scheme using metrics such as FPR, accuracy, and detection rate of their solution. Table 7 exhibits the comparison of the clustering assisted SVM-based ADS schemes and focuses on their applied evaluated metrics, ADS datasets, kernel functions, and feature extraction methods.

To deal with the security issues in the application layer of the industrial control systems, in Shang et al. (2018), the authors analyzed the rule of Modbus/TCP protocol. Besides, they presented an ADS approach by using the

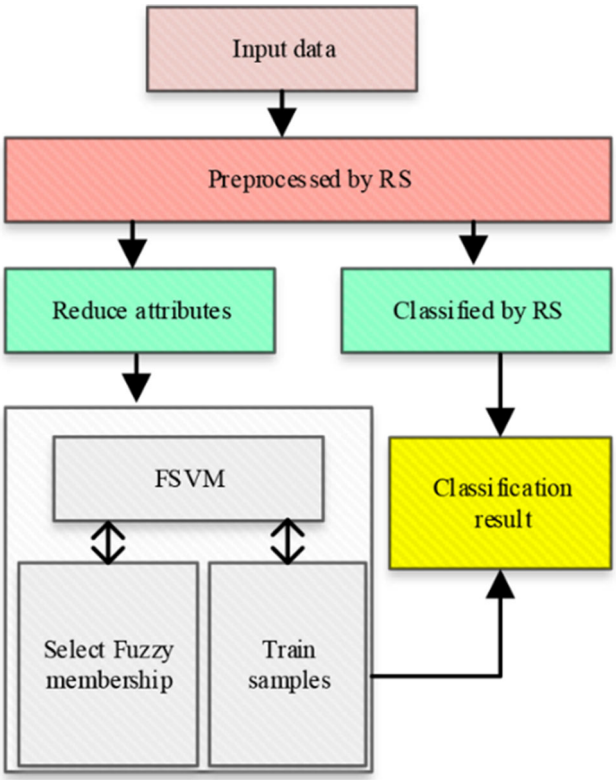

Fig. 14 The block diagram of the RS-FSVM in Li and Zhao (2011)

SVM classifier and fuzzy C-means (FCM) clustering method to compute the distance between the cluster center and industrial network communication. By conducting several experiments, they indicated their method could mitigate the training time and enhance the accuracy without needing labeled data.

5.2 Multiclass SVM-based ADS schemes

Also, multiclass SVMs can be classified into the one against all SVM (OAA-SVM), one against one SVM (OAO-SVM), and DAGSVM. In the first case or the OAA-SVM, only one binary SVM will be used for each class of data, to separate class members from other classes, and a data item can be placed in a class if its SVM accept it and remaining SVMs do not accept it. Although this method can be suitable for tightly clustered data, it may leave some parts of the feature space unclassified, especially in cases in which a data item be accepted by more than one SVM or when all SVMs reject. The OAA-SVM applies the majority voting method for performing classification. On the other hand, OAO-SVM uses one SVM for each pair of classes and employs fewer SVMs, and each SVM will be trained based on data of its two categories. But, this method can suffer from the limitations of the binary SVMs. At last, in the DAGSVM method, the classifier builds $k*(k-1)/2$ hyper-planes and applies a graph traversal method for testing, and its nodes are binary SVMs.

Regarding several classes of attacks that should be detected by an ADS solution, various multiclass SVM-based ADS schemes such as Mewada et al. (2010) are designed and introduced. For example, the multiclass ADS approach in Mulay et al. (2010), utilizes classifiers such as hierarchical multiclass SVM and decision tree to provide ADS, which can mitigate the testing and training time. In this scheme, the decision tree structure determines the division of the feature space and has a direct impact on classifier performance. They studied the hierarchical multiclass SVM and tree-structured multiclass SVM, which consider the distribution of classes along with the distances of class centers. These types of SVMs help to build a pattern recognition model for ADS and classify security attacks and increase the efficiency of ADS by decreasing the training and testing time of data. The authors did not provide any experimental results for their scheme.

The authors in Aburomman and Reaz (2017), applied the two-class SVM for multiclass classification purpose. They proposed WOAR-SVM, a new ADS method, to improve the multiclass classification. They have evaluated several methods for creating a multiclass SVM classifier from some binary SVMs. They applied a weighted OAA-SVM and employed the differential evolution (DE) optimization algorithm to tune the SVM and reduce the prediction errors of each binary classifier. Also, it enables the integration of binary hypotheses into a multiclass hypothesis, where each binary classifier may feature a unique set of parameters. Besides, to evaluate the accuracy of this solution, the authors performed experiments using the NSL-KDD dataset.

Table 8 provides a comparison of the multiclass SVM-based ADS schemes. It mainly compares the evaluated metrics, utilized datasets, kernel functions, and feature extraction methods applied in the outlined ADS schemes.

5.3 Least square SVM-based schemes

LS-SVM or Least squares SVMs are supervised kernel-based learning methods, which apply equality constraints instead of inequality restrictions to solve linear equations in the classification problems and have lower computation overhead. Also, in Ambusaidi et al. (2016), the authors introduced FMIFS, a supervised feature selection algorithm that applies the mutual information to choose optimal features and handle the linearly and nonlinearly dependent features. They used FMIFS for mitigating the redundancy of features and combined it with the LS-SVM. As an advantage, this ADS scheme is evaluated on multiple datasets such as KDD Cup, Kyoto 2006 + , and NSL-KDD

Table 7 Comparison of the clustering assisted SVM-based ADS schemes

Schemes	Evaluated metrics	Kernel functions	Datasets	Feature extraction	Advantages/limitations
Li and Zhao (2011)	Detection Rate accuracy FPR FNR		KDD Cup	Rough Set	Using fuzzy SVM for classification, Only is compared with standard SVM, further comparison with other schemes and classifiers are needed
Sani and Ghasemi (2015)	Detection Rate		Kyoto 2006+ KDD Cup	PCA	This scheme is evaluated on two datasets, but it is compared with only one ADS scheme
Ashok et al. (2011)	Detection Rate FPR Recall		KDD Cup		It is compared with two other schemes; however, further experiments are needed
Chitrakar and Chuanhe (2012)	Accuracy Detection Rate FPR ROC		Kyoto 2006+	Mutual Information	It is compared with the Naïve Bayes classifier with the K-means and K-Medoids clustering methods
Shang et al. (2018)	Accuracy Training time Testing time				Focuses on the industrial networks which are less investigated by other schemes, Further comparisons with different classifiers and ADS schemes are needed

Table 8 Comparison of the multiclass SVM-based ADS schemes

Schemes	Evaluated metrics	Kernel functions	Datasets	Feature extraction	Advantages/limitations
Mewada et al. (2010)	Detection rate	Cauchy, ANOVA, RBF	KDD Cup		Primitive ADS solution, Its idea should be further expanded and should also be evaluated
Mulay et al. (2010)					Only proposing the hierarchical multiclass SVM and tree-structured multiclass SVM algorithms for ADS problem, No experiments and evaluations
Aburomman and Reaz (2017)	Accuracy, FPR, FNR, PPV, F1	Gaussian	NSL-KDD		A comprehensive multiclass ADS approach, Using the DE algorithm for finding an optimal model selection vector, Complete evaluation of the proposed ADS approach

datasets in terms of FPR, accuracy, detection rate, and F-measure.

5.4 Radial SVM-based schemes

Many ADS approach has been proposed in the literature to deal with the anomaly detection problem in the literature using radial SVM, which will be studied in this section.

The ADS approach introduced in Kuang et al. (2012) by Kuang et al. uses a localized kernel PCA for data preprocessing. It decreases the time-shift sensitivity problem in the SVM. It benefits from an improved version of the SVM classifier to find the main features. In this scheme, the GA algorithm is used for training the kernel parameters in the SVM. They exhibited that the KPCA outperforms the PCA by exploring higher-order information of the inputs, and it

better extracts the principal components. Nevertheless, this scheme is not compared with other classifiers to indicate its performance.

In Enache et al. (2015) carried out a wrapper-based selection of features that combined Naïve Bayes and SVM classifiers. They employed a binary version of the BA, which is enhanced by using Lévy flights to improve the exploration capability. By cross-validation and experiments conducted on the NSL-KDD, the authors indicated that this algorithm could improve the detection rate, execution time, and FPR of this ADS scheme.

Table 9 exhibits the comparison of the radial SVM-based ADS schemes. It mainly compares the evaluated metrics, applied ADS datasets, and utilized kernel functions in the employed SVM classified.

5.5 Hybrid ADS schemes

Different ADS schemes such as Erfani et al. (2016), Liu et al. (2011), Dixit et al. (2018), Anton et al. (2019) and Hasan et al. (2019) have benefited from the SVM classifier in combination with other classifiers and techniques to handle anomalies. This subsection provides a brief study on such ADS schemes.

In Lin et al. (2012), the authors present a wrapper-based ADS which tries to select the essential features of the KDD Cup dataset using SVM classifier and decision tree (DT), which can achieve the decision rules. They also applied simulated annealing (SA) for feature selection training of SVM and DT classifiers for enhancing the detection accuracy. The parameter for the decision tree and SVM is adjusted through SA. The authors have employed ten-fold cross-validation to evaluate the accuracy of classification for the KDD Cup. The flowchart of this ADS approach is shown in Fig. 15.

In Zhang and Shen (2005), Zhang and Shen try to handle the online intrusion detection as a text processing problem and detect the system calls in the privileged processes. They employed a modified TF-IDF (term frequency-inverse document frequency), which is a text processing method, regarding the correlation between the processes, the time information, and the consequences of the attacks. Also, inspired by online SVM, they modified the robust SVM and one-class SVM. Besides, the authors have conducted evaluations on the BSM audit data of the DARPA dataset. As an advantage, it can reduce the support vectors leading to mitigation in the required training time and running time while keeping the detection accuracy. Figure 16 depicts the block diagram of the ADS proposed in this scheme.

In Wang et al. (2016), Wang et al. put forward an effective method called exemplar extraction to extract subsets from the original dataset before building the detection models. The affinity propagation and K-means algorithms are used to find the exemplars. They presented exemplar extraction, to extract a smaller set of representative exemplars which summarize a large training dataset. Consequently, the training will be shorter, and the test will

Table 9 Comparison of the least square SVM and radial SVM-based ADS schemes

Schemes	Evaluated metrics	Kernel functions	ADS datasets	Feature extraction	Advantages/limitations
Ambusaidi et al. (2016)	Accuracy Detection Rate FPR, Training Time, Testing Time		KDD Cup, NSL-KDD, Kyoto 2006+	Mutual Information	Comprehensive comparisons with several datasets
Kuang et al. (2012)	Detection Rate	RBF	KDD Cup	PCA, Entropy	Tuning Kernel parameters with GA, Needs further comparisons against other ADS schemes
Enache et al. (2015)	Accuracy Detection Rate Training Time Testing Time	RBF			Improving Bat algorithm with Lévy Flights, Conducting comparisons only with the Naïve Bayes classifier

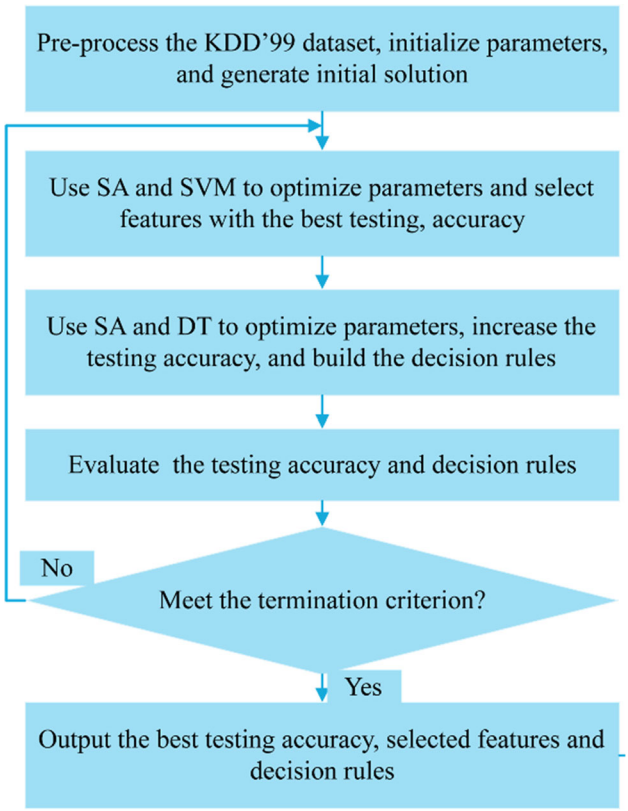

Fig. 15 Flowchart of the proposed solution in Lin et al. (2012)

be performed on a compressed detection model. Also, they used PCA to reduce the dimension of data and employed IG for attribute selection. Also, K-NN and OC-SVM classifiers are utilized for anomaly detection. As an advantage, the authors have applied multiple datasets to validate their work in the MATLAB software. In this process, two HTTP datasets are extracted from a real network environment and the KDD Cup dataset. Figure 17 depicts the steps carried out in this ADS scheme.

In Erfani et al. (2016), Erfani et al. provided a hybrid ADS model using an unsupervised deep belief network trained to extract proper features for the training of the OC-SVM. This model's performance can be enhanced with a deep auto-encoder while reducing its training and testing time. As an advantage, the authors have provided an extensive evaluation of their scheme using various types of SVM, with linear and RBF kernels. It applies a high-dimensional unlabeled dataset. The DBN is trained as a dimensionality reduction algorithm, and the derived features are taken as input to train the OC-SVM. This scheme reduces the complexity and scalability issues of the SVM, in training with large-scale datasets by mitigating the impact of irrelevant features. DBNs are appropriate feature detectors for anomaly detection, taking only unlabeled data to capture higher-order correlations among features, generating an accurate model, and imposing minimal computationally and memory complexity. As an advantage, by using a deep architecture, this scheme can deliver better generalization, since basic kernels such as linear kernels can be used instead of more expensive kernels such as RBF without affecting the accuracy. However, this study is just conducted on the WSN datasets, and further tests are required on other datasets.

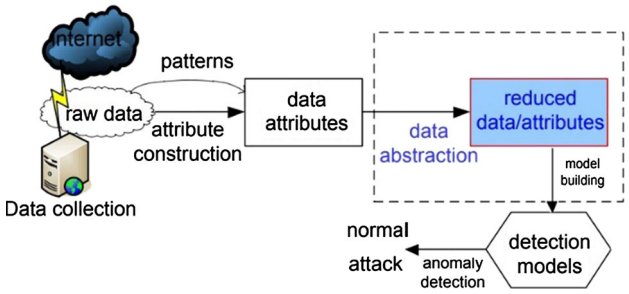

Fig. 17 ADS steps in Wang et al. (2016)

The approach introduced in Enache and Sgârciu (2015c), proposed a feature selection method using a wrapper-based technique that applies an enhanced binary BA with SVM and C4.5 decision tree classifiers. They tried to increase the FPR and detection rate of their anomaly detection approach and used the NSL-KDD dataset in ten-fold cross-validations. Figure 18 exhibits the block diagram of this ADS model.

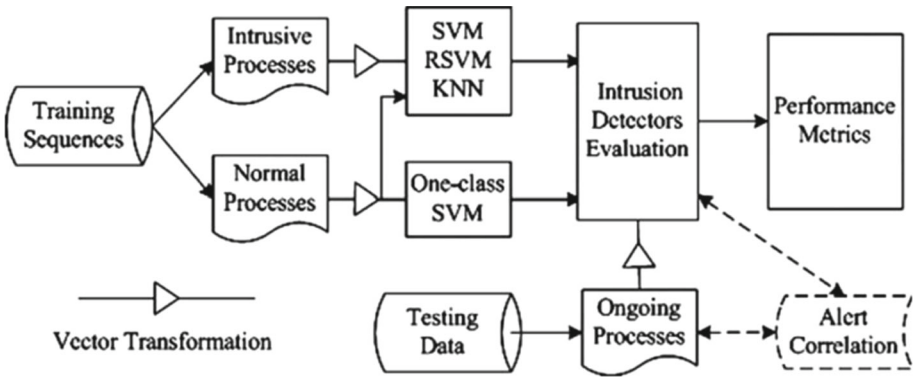

Fig. 16 ADS block diagram in Zhang and Shen (2005)

The ADS solution proposed in Kim et al. (2014), have applied both anomaly detection and misuse detection methods to deal with intrusions. In this scheme, the misuse detection part uses the known attack data, and the anomaly detection part utilizes the normal traffic data. For this purpose, this scheme decomposes the normal training data into disjoint subsets misuse detection model, and also an ADS model is created for each of them. It applies the SVM and C4.5 decision tree classifiers, where the decision tree is used to implement misuse detection and decomposes the training data of normal profiles into smaller categories. Afterward, the OC-class SVM is employed to have an anomaly detection model for each decomposed subset, and it can use information about the known attack in building normal behavior profiles. This scheme improves the performance of the detection rate and speed of unknown attacks. They have applied Gaussian kernel in the SVM and carried out their simulations on the NSL-KDD dataset. However, the normal data can be decomposed evenly into subsets without degrading the misuse detection performance.

For detecting intrusions, Chen et al. (2005) utilized the ANN and SVM classifiers with two encoding methods. They claimed that SVM with term frequency-inverse document frequency could achieve the best performance. The authors also applied BSM data from the DARPA dataset in their experiments.

In Wang et al. (2010), the authors presented a hybrid model that integrates anomaly detection with the provisioning of high-level data about the anomalies. This ADS approach uses C-SVM and OC-SVM to find and filter unknown anomalous connections. Besides, SOM or self-organizing map is used in this scheme. SOM is an unsupervised ANN model to map high-dimensional data in two-dimensional lattices and discover hidden patterns in the training dataset. Also, since attacks that have common properties belong to the same cluster, this scheme conducts clustering and categorizing similar malicious connections in four attack classes. Their model used the KDD Cup and showed a high detection rate with a low FPR. The architecture of this ADS is shown in Fig. 19.

The scheme presented in Anil and Remya (2013), exploited SOFM, a self-organized feature map to improve the extraction of features from the KDD Cup. Also, a tournament-based GA is used to find features, and SOFM is employed to achieve similar groups from the dataset to reduce its size and mitigate the training time of the SVM. As an advantage, this ADS benefits from low computational time and a high anomaly detection rate.

In Ergen and Kozat (2019), the authors focus on unsupervised anomaly detection and used long short-term memory (LSTM) ANN. Furthermore, they used a decision function using the support vector data description (SVDD) and OC-SVMs. They also trained and optimized the LSTM and OC-SVM using quadratic programing and gradient-based methods. For this purpose, they modify the objective criteria of the OC-SVM and SVDD algorithms in proving the convergence.

They indicated that variable-length sequences could be processed while maintaining good performance. Their conducted experiments showed their method could outperform other conventional approaches.

In Serkani et al. (2019), a hybrid ADS scheme is provided using the LS-SVM and C5.0 decision tree classifiers, in which the latter is used for feature selection. It is performed done by pruning the decision tree and removing the features with the least predictor significance. Afterward, the LS-SVM is used, and the final features are the ones with the highest surface area under the ROC curve. They used the UNSW-NB15 and KDD Cup 99 datasets and indicated that this approach improves accuracy and reduces the FPR compared to the other ADS works.

In Injadat et al. (2018), the authors provided an effective ADS solution and used the Bayesian optimization for tuning the classifiers such as SVM, K-NN, and Random Forest. Then, they evaluated the performance of these

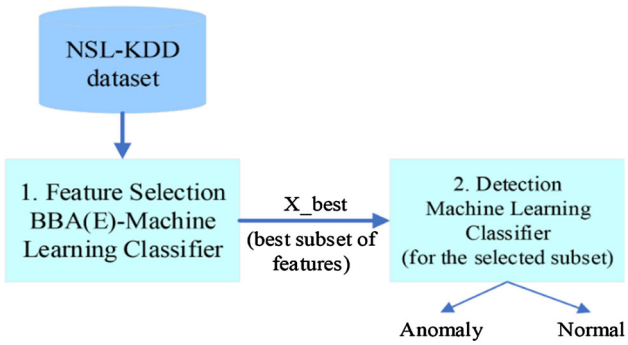

Fig. 18 The ADS model in Enache and Sgârciu (2015)

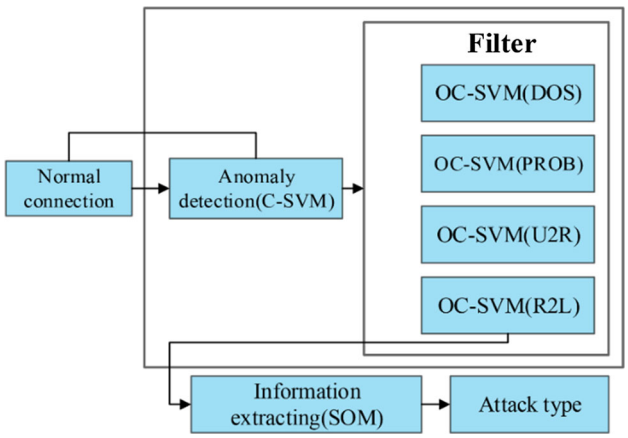

Fig. 19 ADS architecture in Wang et al. (2010)

classifiers in the MATLAB environment. For performance evaluations, the ISCX 2012 dataset is applied, and metrics such as precision, accuracy, recall, and FPR are evaluated.

Table 10 compares the hybrid SVM-based ADS approaches, regarding their evaluated metrics, ADS datasets, and kernel functions utilized in the outlined schemes.

6 Discussion

This section is aimed to provide a comparison of different properties of the investigated ADS schemes. On the other hand, it gives some information about the following issues:

- Various types of SVMs are utilized to recognize anomalies.
- Kernel functions which are employed in the studied ADS schemes.
- Other classifiers applied in combination with the SVM to deal with anomalies.
- The feature extraction methods applied in the outlined ADS schemes.
- The datasets which are used to analyze the efficiency of the investigated ADS schemes.
- Evaluation metrics utilized to verify the effectiveness of the proposed ADS solutions.

We further compared various types of SVM classifiers employed in the investigated ADS solutions in Fig. 20. As shown in this figure, the most common type of SVM applied in the ADS is OC-SVM, multiclass SVM, and hybrid schemes, which use other classifiers in combination with SVM.

Figure 21 depicts the percentage of the proposed ADS schemes using various meta-heuristic algorithms in combination with the SVM classifier to improve the performance of ADS schemes. As outlined before, these algorithms can be efficiently used for feature selection of benchmark datasets, training of the SVM, kernel functions, and other employed classifiers. However, only a few meta-heuristic algorithms are employed in the outlined anomaly-based ADS approaches. Thus, in the future, different newly proposed meta-heuristic algorithms should be applied in this context to enhance the ADS capabilities further.

Also, all of the employed meta-heuristic algorithms are single objective algorithms, and multi-objective optimization algorithms have not been adequately studied. Consequently, in future studies, multi-objective optimization algorithms can be benefited to solve ADS problems with several conflicting objectives.

Employing a proper dataset is very critical, and since ADS solutions are context-aware, designing special-purpose datasets in each different context should be considered in the future. Figure 22 exhibits the properties of an ideal dataset, which should be considered in selecting an appropriate dataset for anomaly detection. Figure 23 exhibits the datasets applied in the studied schemes, and it indicates the number of solutions that have used each dataset. As shown in this figure, the primary datasets used in this content are DARPA, KDD Cup, NSL-KDD, and Kyoto 2006 + . However, these datasets are quite old, and in the subsequent studies, newer datasets should be adapted for the evaluation of ADS schemes. Also, the DARPA-based datasets often suffer from a lack of attack diversity and only contain a limited type of attack. Furthermore, only a handful of schemes have evaluated their proposed model using multiple datasets. In the future, this issue can be further challenged to create more general and useful ADS models.

One of the crucial features in the investigated ADS schemes is the evaluation metrics that have been utilized to indicate the effectiveness of the proposed ADS solutions. Figure 24 exhibits the evaluation metrics used in the SVM-based ADS schemes. As depicted in this figure, metrics such as detection rate, accuracy, FPR, train duration, and testing time are mostly used to evaluate the performance of the ADS schemes.

Also, Fig. 25 indicates various kernel functions applied in the ADS schemes for classification of nonlinear data by converting low dimensional data to high dimensional one, to detect numerous attacks and intrusions accurately. As shown in this figure, RBF, Gaussian, and polynomial are the three most widely benefited kernel functions by the studied ADS schemes. To achieve better results, tuning the parameters of these kernel functions can be further examined.

Figure 26 exhibits the other classifiers that have been combined with the SVM classifier to improve the classification performance. As shown in this figure, classifiers such as DT and SOM are mostly applied in conjunction with different types of SVM classifiers. Furthermore, Fig. 27 depicts the number of schemes which have successfully benefited from each kind of feature extraction methods. As shown in this figure, PCA, mutual information, and entropy methods are utilized for feature extraction in the investigated SVM-based ADS schemes.

7 Conclusions and future research directions

Anomaly detection is a classification problem aimed to find the nonconforming patterns in data. In the security and intrusion detection contexts, anomaly detection approaches can deal with new cyber threats launched against the hosts and computer networks. Several interesting pieces of research have been conducted in the anomaly intrusion

Table 10 Comparison of the hybrid ADS schemes

Schemes	Evaluated metrics	Kernel functions	Datasets	Feature extraction	Advantages/limitations
Lin et al. (2012)	Accuracy		KDD Cup	Information Gain	Further evaluations are needed with other datasets Efficient combination of SVM and DT classifiers
Zhang and Shen (2005)	Accuracy Training Time FPR support vectors	RBF	DARPA's BSM data set		Evaluating the online training for various kinds of SVMs without accuracy deterioration Reducing training time
Wang et al. (2016)	Detection Rate FPR		KDD Cup Two HTTP datasets	PCA Information Gain	Extracting subsets from the original massive data, Extensive evaluations on the SVM and KNN
Erfani et al. (2016)	AUC Training Time Testing Time	Linear kernels		Deep belief networks	Good integration of SVM and Deep Learning Extensive comparison in various datasets
Kim et al. (2014)	Accuracy Training Time Testing Time ROC	Gaussian	NSL-KDD	Entropy	Successful combination of SVM and decision tree classifiers Compared with SVM and decision tree, but not with other ADS schemes
Chen et al. (2005)	Detection Rate FPR ROC	Gaussian Polynomial	DARPA's BSM data set		Using SVM and ANN for anomaly detection Extensive comparison of the proposed models using t-tests and ADS metrics
Wang et al. (2010)	Accuracy FPR FNR Detection Rate	RBF	KDD Cup		Combining SVM and SOM Needs further evaluation and comparison with other ADS schemes
Anil and Remya (2013)	Detection Rate	Polynomial, inner product, and sigmoid	KDD Cup	Information Gain Entropy	Combination of SVM, SOM, and GA Only is compared with SVM, further comparisons are needed
Ergen and Kozat (2019)	ROC		Hong Kong Exchange rate, Http, Alcoa stock price		Unsupervised anomaly detection, Extensive comparisons are conducted, Required mathematical proofs are provided
Serkani et al. (2019)	Accuracy, TPR, FPR, AUC		UNSWNB15, KDD Cup	Information Gain	Combination of the LS-SVM and decision tree, Extensive performance Evaluations
Injadat et al. (2018)	Precision, Accuracy, Recall, FPR	Gaussian	ISCX		They evaluated the performance of each classifier using the Bayesian optimization with their different parameters. No combination of the classifiers.

detection field, using techniques provided by artificial intelligence, data mining, etc. Support vector machine (SVM) is a binary classifier successfully applied in the regression analysis and classification fields. Also, SVM is used by many research works in the literature to deal with the anomaly detection problem in the security context.

This paper addresses the SVM-based anomaly detection systems (ADS) by giving a thorough survey of them. For

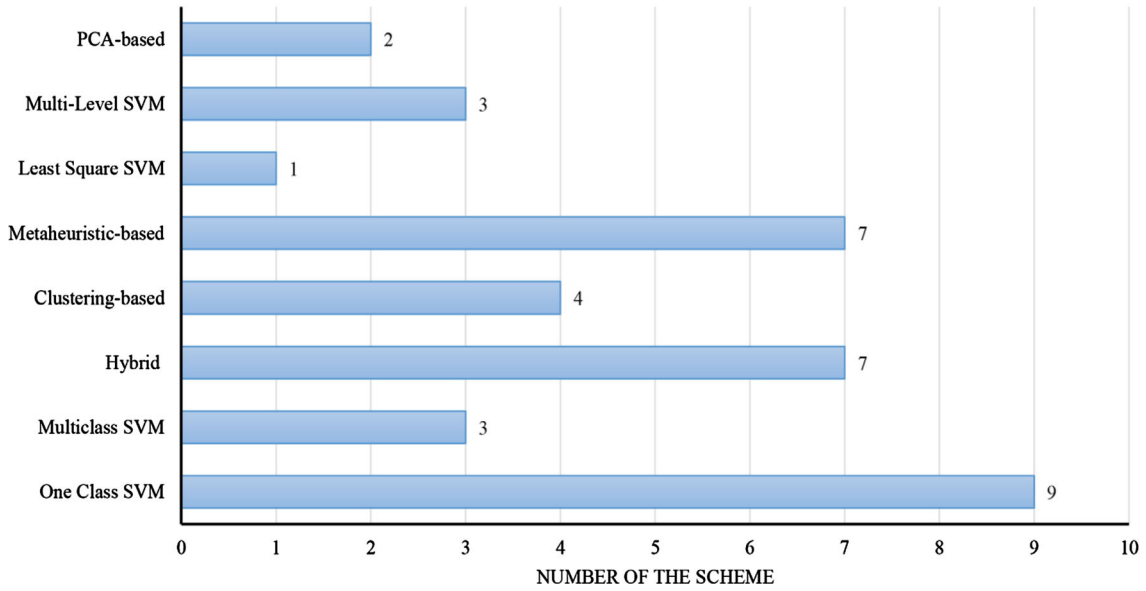

Fig. 20 Number of the ADS schemes designed using each kind of the SVM

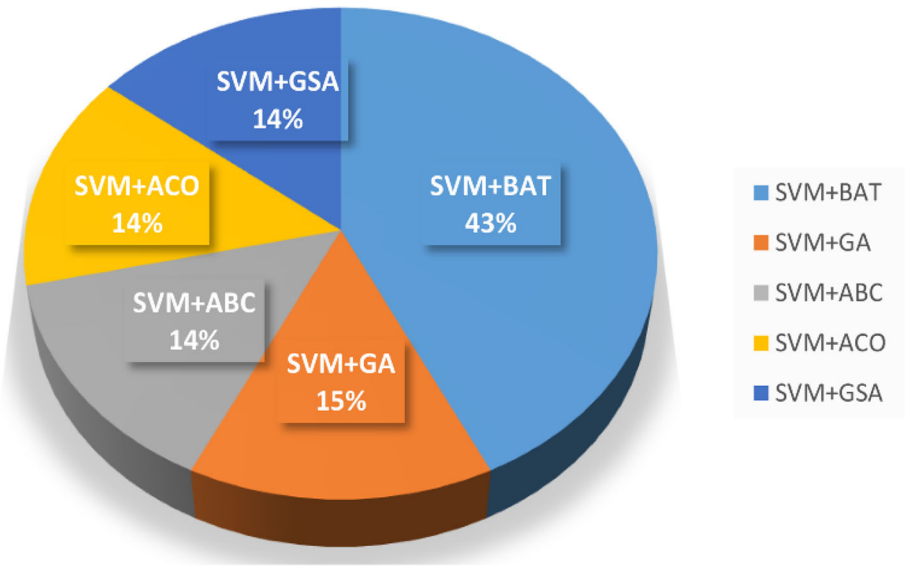

Fig. 21 Meta-heuristic algorithms applied in the SVM-based ADS schemes

this purpose, it puts forward the background concepts and knowledge about the SVM classifier and anomaly detection systems (ADS). After that, a taxonomy of the SVM-based ADS schemes is provided according to the type of the SVM classifier and feature selection techniques. Then, it summarizes their main contributions and innovations. Also, any possible limitations and advantages of the investigated ADS schemes are discussed. However, in spite of the several anomaly detection approaches designed and provided in the literature, there are some open research problems and challenging issues which lie ahead of the researchers:

- For adapting to the changing legal behaviors, improving detection accuracy, and reducing the FP, the future ADS solutions must provide support for incremental learning and online training.
- With the emerging technologies like IoT, the need for low overhead, fast, and high-performance ADS solutions become apparent to run on resource-limited IoT devices.
- Security-related anomalies must be found very quickly before they attack computer networks and hosts. Thus, required training time and execution complexity should be considered in designing an ADS.
- Most studied SVM-based ADS schemes are designed for network anomaly detection, and only a few numbers

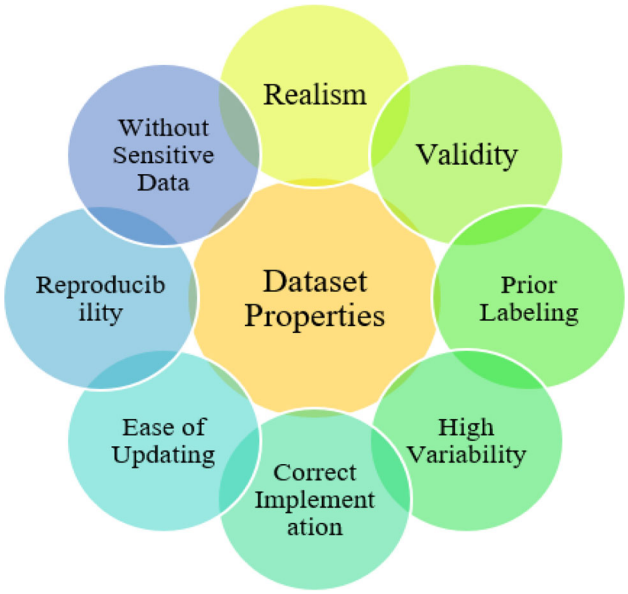

Fig. 22 Properties of an ideal dataset

of works have been designed for detecting anomalies in the hosts.

- The evaluated ADS schemes can handle either host-based anomalies or network anomalies. Accordingly, designing an ADS scheme to be effective in both cases can be challenged in the future.
- Since normal behaviors are different in each context, there are not general ADS to be used in all contexts.

Moreover, high-dimensional anomaly datasets can cause serious challenges; some of them can be listed as follows:

- Curse of dimensionality: The feature subspace can grow exponentially as the input dimension increases.
- Irrelevant features: Such features in the applied input data may cause noise and can prevent the classifiers from truly recognizing anomalies. This problem makes choosing a subset of features and finding relevant attributes a challenging issue.
- As outlined in the previous section, feature selection based on the multi-objective optimization algorithms should be investigated in the future for the SVM-based ADS schemes.
- The benchmark datasets such as the KDD-based datasets suffer from the enormous volume, high dimension, and skewed distribution of data. Also, they may fail in simulating real-world computer network traffic. Thus, an ADS which have been tested using these datasets may have unacceptable performance in the real environments; for handling this problem, the similarity of the applied datasets with the actual network traces should be evaluated.

Regarding other types of intrusion detection methods, the following remarks can be made:

- Regarding the results of Fig. 1, it can be seen that fewer researches have been conducted in the SVM-based misuse detection context, which should be further investigated in the future.
- From the studied ADS schemes, only a few of them support both anomaly detection and misuse detection. Thus, further investigations about the hybrid misuse and anomaly detection systems should be made in the future, for recognizing both known and unknown attacks and intrusions.
- Because of the inherent uncertainty of the intrusion detection process, fuzzy logic can be employed to enhance the ADS schemes' performance. Nonetheless, only a small number of fuzzy ADS schemes based on the SVM classifier are presented in the literature, and in

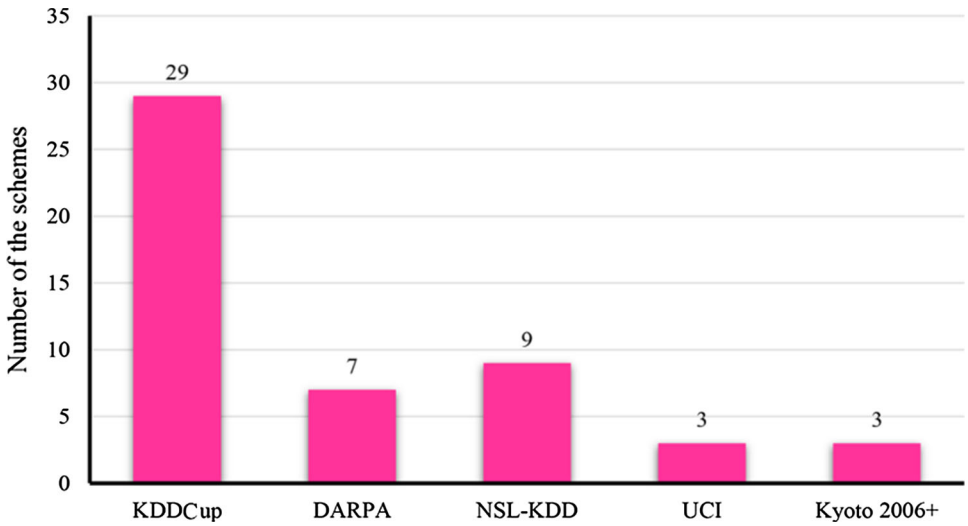

Fig. 23 Datasets used in the SVM-based ADS approaches

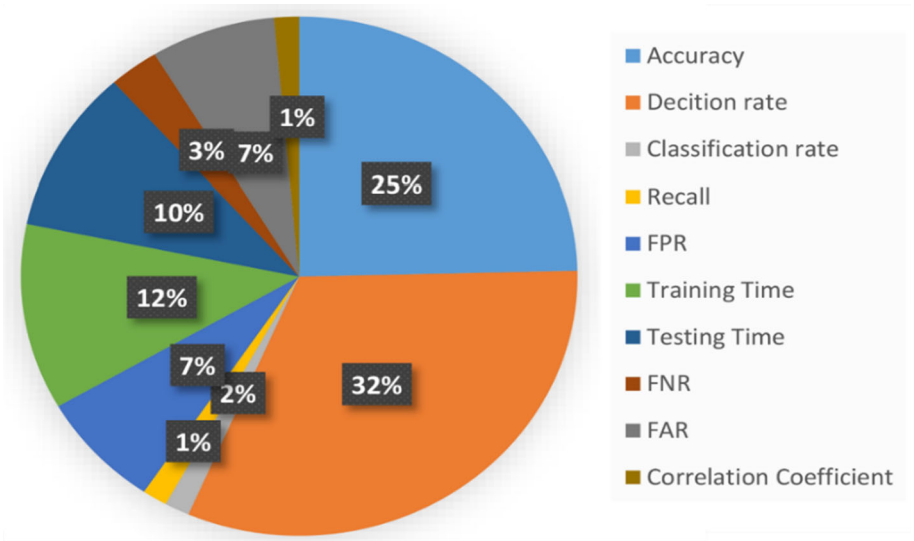

Fig. 24 Evaluation metrics employed in the SVM-based ADS schemes

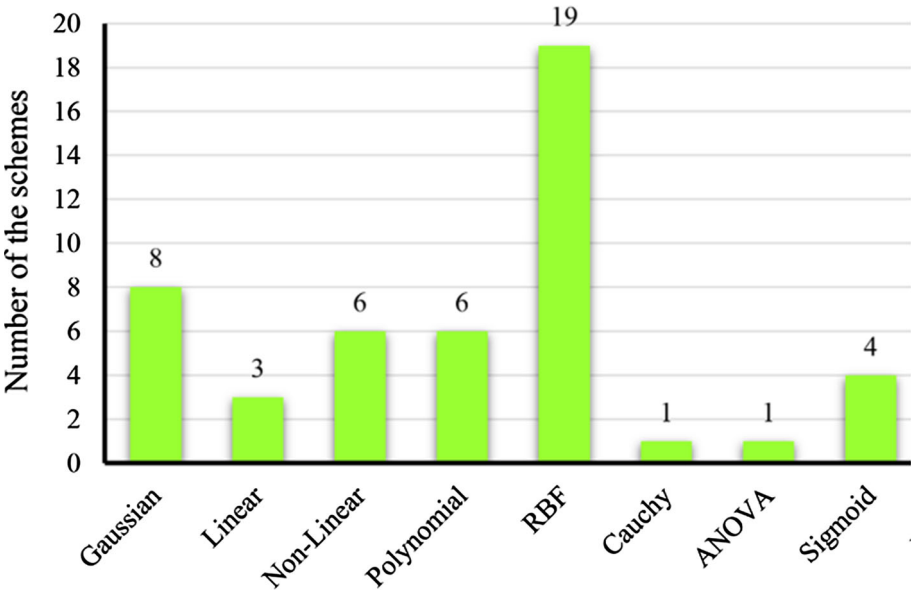

Fig. 25 Kernels employed in the SVM-based ADS schemes

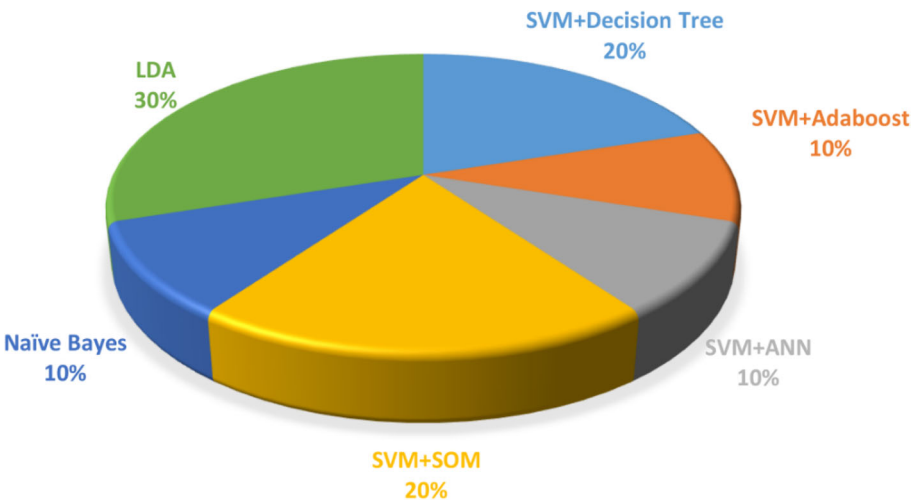

Fig. 26 Classifiers applied in combination with the SVM-based ADS

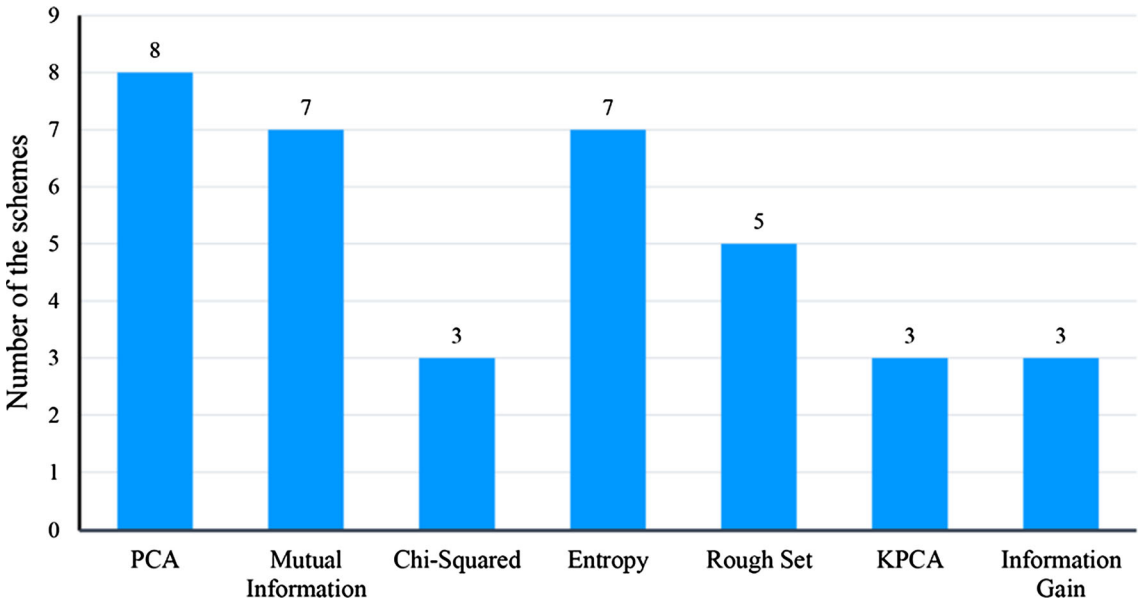

Fig. 27 Feature extraction methods utilized in conjunction with the SVM-based ADS

the subsequent investigations, it should be focused further.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Abraham A, Jain R, Thomas J, Han SY (2007) D-SCIDS: Distributed soft computing intrusion detection system. J Netw Comput Appl 30:81–98

Aburomman AA, Reaz MBI (2017) A novel weighted support vector machines multiclass classifier based on differential evolution for intrusion detection systems. Inf Sci 414:225–246

Agarwal B, Mittal N (2012) Hybrid approach for detection of anomaly network traffic using data mining techniques. Procedia Technol 6:996–1003

Ahmed M, Mahmood AN, Hu J (2016) A survey of network anomaly detection techniques. J Netw Comput Appl 60:19–31

Al Shorman A, Faris H, Aljarah I (2019) Unsupervised intelligent system based on one class support vector machine and Grey Wolf optimization for IoT botnet detection. J Ambient Intell Hum Comput 1–17

Alaba FA, Othman M, Hashem IAT, Alotaibi F (2017) Internet of things security: a survey. J Netw Comput Appl 88:10–28

Al-Qatf M, Lasheng Y, Al-Habib M, Al-Sabahi KJIA (2018) Deep learning approach combining sparse autoencoder with SVM for network intrusion detection. IEEE Access 6:52843–52856

Ambusaidi MA, He X, Nanda P, Tan Z (2016) Building an intrusion detection system using a filter-based feature selection algorithm. IEEE Trans Comput 65:2986–2998

Amraee S, Vafaei A, Jamshidi K, Adibi P (2018) Abnormal event detection in crowded scenes using one-class SVM. SIViP 12:1115–1123

Anil S, Remya R (2013) A hybrid method based on genetic algorithm, self-organised feature map, and support vector machine for better network anomaly detection. In: 2013 Fourth international conference on computing, communications and networking technologies (ICCCNT). pp 1–5

Anton SD, Kanoor S, Fraunholz D, Schotten HD (2018) Evaluation of machine learning-based anomaly detection algorithms on an industrial Modbus/TCP data set. In: Proceedings of the 13th international conference on availability, reliability and security. pp 1–9

Anton SDD, Sinha S, Schotten HD (2019) Anomaly-based intrusion detection in industrial data with SVM and random forests. In: 2019 International conference on software, telecommunications and computer networks (SoftCOM). pp 1–6

Ashok R, Lakshmi AJ, Rani GDV, Kumar MN (2011) Optimized feature selection with k-means clustered triangle SVM for Intrusion Detection. In: 2011 Third international conference on advanced computing (ICoAC). pp 23–27

Aslahi-Shahri B, Rahmani R, Chizari M, Maralani A, Eslami M, Golkar M, Ebrahimi A (2016) A hybrid method consisting of GA and SVM for intrusion detection system. Neural Comput Appl 27:1669–1676

Bamakan SMH, Wang H, Yingjie T, Shi Y (2016) An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization. Neurocomputing 199:90–102

Bostani H, Sheikhan M (2017) Hybrid of binary gravitational search algorithm and mutual information for feature selection in intrusion detection systems. Soft Comput 21:2307–2324

Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: a survey. ACM Comput Surv CSUR 41:1–58

Chen W-H, Hsu S-H, Shen H-P (2005) Application of SVM and ANN for intrusion detection. Comput Oper Res 32:2617–2634

Cheng C, Tay WP, Huang G-B (2012) Extreme learning machines for intrusion detection. In: The 2012 international joint conference on neural networks (IJCNN). pp 1–8

Chitrakar R, Chuanhe H (2012) Anomaly detection using Support Vector Machine classification with k-Medoids clustering. In: 2012 Third Asian Himalayas international conference on internet (AH-ICI). pp 1–5

Chu W-L, Lin C-J, Chang K-N (2019) Detection and classification of advanced persistent threats and attacks using the support vector machine. Appl Sci 9:4579

Cid-Fuentes JA, Szabo C, Falkner K (2018) Adaptive performance anomaly detection in distributed systems using online SVMs. IEEE Trans Dependable Secure Comput

De la Hoz E, De La Hoz E, Ortiz A, Ortega J, Prieto B (2015) PCA filtering and probabilistic SOM for network intrusion detection. Neurocomputing 164:71–81

Dixit M, Moholkar A, Limaye S, Limaye D (2018) Naive Bayes and SVM based NIDS. In: 2018 3rd International conference on inventive computation technologies (ICICT). pp 527–532

Dong H, Peng D (2018) Research on abnormal detection of ModbusTCP/IP protocol based on one-class SVM. In: 2018 33rd Youth academic annual conference of chinese association of automation (YAC). pp 398–403

Elshoush HT, Osman IM (2011) Alert correlation in collaborative intelligent intrusion detection systems—a survey. Appl Soft Comput 11:4349–4365

Emadi HS, Mazinani SM (2018) A novel anomaly detection algorithm using DBSCAN and SVM in wireless sensor networks. Wirel Pers Commun 98:2025–2035

Enache A-C, Patriciu VV (2014) Intrusions detection based on support vector machine optimized with swarm intelligence. In: 2014 IEEE 9th international symposium on applied computational intelligence and informatics (SACI). pp 153–158

Enache A-C, Sgarciu V (2014) Enhanced intrusion detection system based on bat algorithm-support vector machine. In: 2014 11th International conference on security and cryptography (SECRYPT). pp 1–6

Enache A-C, Sgârciu V (2015a) Anomaly intrusions detection based on support vector machines with an improved bat algorithm. In: 2015 20th international conference on control systems and computer science (CSCS). pp 317–321

Enache A-C, Sgârciu V (2015b) An improved bat algorithm driven by support vector machines for intrusion detection. In: International joint conference. pp 41–51

Enache A-C, Sgârciu V (2015c) A feature selection approach implemented with the Binary Bat Algorithm applied for intrusion detection. In: 2015 38th International conference on telecommunications and signal processing (TSP). pp 11–15

Enache A-C, Sgarciu V, Petrescu-Niţă A (2015) Intelligent feature selection method rooted in Binary Bat Algorithm for intrusion detection. In: 2015 IEEE 10th Jubilee international symposium on applied computational intelligence and informatics (SACI). pp 517–521

Erfani SM, Rajasegarar S, Karunasekera S, Leckie C (2016) High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. Pattern Recogn 58:121–134

Ergen T, Kozat SS (2019) Unsupervised anomaly detection with LSTM neural networks. IEEE Trans Neural Netw Learn Syst

Feng W, Zhang Q, Hu G, Huang JX (2014) Mining network data for intrusion detection through combining SVMs with ant colony networks. Future Gener Comput Syst 37:127–140

Feng F, Liu X, Yong B, Zhou R, Zhou Q (2019) Anomaly detection in ad-hoc networks based on deep learning model: a plug and play device. Ad Hoc Netw 84:82–89

Ganapathy S, Yogesh P, Kannan A (2012) Intelligent agent-based intrusion detection system using enhanced multiclass SVM. Comput Intell Neurosci 2012:9

Gautam SK, Om H (2016) Computational neural network regression model for Host based Intrusion Detection System. Perspect Sci 8:93–95

Ghomi EJ, Rahmani AM, Qader NN (2017) Load-balancing algorithms in cloud computing: a survey. J Netw Comput Appl 88:50–71

Gong S, Gong X, Bi X (2011) Feature selection method for network intrusion based on GQPSO attribute reduction. In: 2011 International conference on multimedia technology (ICMT). pp 6365–6368

Guo Y, Wang B, Zhao X, Xie X, Lin L, Zhou Q (2010) Feature selection based on Rough set and modified genetic algorithm for intrusion detection. In: 2010 5th international conference on computer science and education (ICCSE). pp 1441–1446

Hasan M, Islam MM, Zarif MII, Hashem M (2019) Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet of Things 7:100059

Heba FE, Darwish A, Hassanien AE, Abraham A (2010) Principle components analysis and support vector machine based intrusion detection system. In: 2010 10th International conference on intelligent systems design and applications (ISDA). pp 363–367

Hodge V, Austin J (2004) A survey of outlier detection methodologies. Artif Intell Rev 22:85–126

Hu W, Gao J, Wang Y, Wu O, Maybank S (2014) Online adaboost-based parameterized methods for dynamic distributed network intrusion detection. IEEE Trans Cybern 44:66–82

Injadat M, Salo F, Nassif AB, Essex A, Shami A (2018) Bayesian optimization with machine learning algorithms towards anomaly detection. In: 2018 IEEE global communications conference (GLOBECOM). pp 1–6

Ioannou C, Vassiliou V (2019) Classifying security attacks in IoT networks using supervised learning. In: 2019 15th International conference on distributed computing in sensor systems (DCOSS). pp 652–658

Jiang J, Yasakethu L (2013) Anomaly detection via one class svm for protection of scada systems. In: 2013 International conference on cyber-enabled distributed computing and knowledge discovery (CyberC). pp 82–88

Kabir E, Hu J, Wang H, Zhuo G (2018) A novel statistical technique for intrusion detection systems. Future Gener Comput Syst 79:303–318

Khamis SA, Foozy CFM, Ab Aziz MF, Rahim N (2020) Header based email spam detection framework using support vector machine (SVM) technique. In: International conference on soft computing and data mining. pp 57–65

Khan SA, Daachi B, Djouani K (2012) Application of fuzzy inference systems to detection of faults in wireless sensor networks. Neurocomputing 94:111–120

Khreich W, Khosravifar B, Hamou-Lhadj A, Talhi C (2017) An anomaly detection system based on variable N-gram features and one-class SVM. Inf Softw Technol 91:186–197

Kim G, Lee S, Kim S (2014) A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Syst Appl 41:1690–1700

Kuang F, Xu W, Zhang S, Wang Y, Liu K (2012) A novel approach of KPCA and SVM for intrusion detection. J Comput Inf Syst 8:3237–3244

Laamari MA, Kamel N (2014) A hybrid bat based feature selection approach for intrusion detection. In: Bio-inspired computing-theories and applications. Springer, pp 230–238. https://doi.org/10.1007/978-3-662-45049-9_38

Li L, Zhao K-n (2011) A new intrusion detection system based on rough set theory and fuzzy support vector machine. In: 2011 3rd International workshop on intelligent systems and applications (ISA). pp 1–5

Liao H-J, Lin C-HR, Lin Y-C, Tung K-Y (2013) Intrusion detection system: a comprehensive review. J Netw Comput Appl 36:16–24
Lin S-W, Ying K-C, Lee C-Y, Lee Z-J (2012) An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection. Appl Soft Comput 12:3285–3290
Liu Y, An A, Huang X (2006) Boosting prediction accuracy on imbalanced datasets with SVM ensembles. In: Pacific-Asia conference on knowledge discovery and data mining. pp 107–118
Liu Y, Huang X, An A, Yu X (2008) Modeling and predicting the helpfulness of online reviews. In: 2008 Eighth IEEE international conference on data mining. pp 443–452
Liu H, Jian Y, Liu S (2010) A new intelligent intrusion detection method based on attribute reduction and parameters optimization of SVM. In: 2010 Second international workshop on education technology and computer science (ETCS). pp 202–205
Liu W, Ren P, Liu K, Duan H-x (2011) Intrusion detection using SVM. In: 2011 7th International conference on wireless communications, networking and mobile computing (WiCOM). pp 1–4
Masdari M, Ahmadzadeh S (2017) A survey and taxonomy of the authentication schemes in Telecare Medicine Information Systems. J Netw Comput Appl 87:1–19
Masdari M, Jalali M (2016) A survey and taxonomy of DoS attacks in cloud computing. Secur Commun Netw 9:3724–3751
Masdari M, Zangakani M (2019) Green cloud computing using proactive virtual machine placement: challenges and issues. J Grid Comput 1–33
Masdari M, Ahmadzadeh S, Bidaki M (2017) Key management in wireless body area network: challenges and issues. J Netw Comput Appl 91:36–51
Mazini M, Shirazi B, Mahdavi I (2018) Anomaly network-based intrusion detection system using a reliable hybrid artificial bee colony and AdaBoost algorithms. J King Saud Univ Comput Inf Sci
Mehmod T, Rais HBM (2016) Ant colony optimization and feature selection for intrusion detection. In: Advances in machine learning and signal processing. Springer, pp 305–312. https://doi.org/10.1007/978-3-319-32213-1_27
Mewada A, Gedam P, Khan S, Reddy MU (2010) Network intrusion detection using multiclass support vector machine. Spec Issue IJCCT 1:172–175
Miao X, Liu Y, Zhao H, Li C (2018) Distributed online one-class support vector machine for anomaly detection over networks. IEEE Trans Cybern 49:1475–1488
Modi C, Patel D, Borisaniya B, Patel H, Patel A, Rajarajan M (2013) A survey of intrusion detection techniques in cloud. J Netw Comput Appl 36:42–57
Mulay SA, Devale P, Garje G (2010) Decision tree based support vector machine for intrusion detection. In: 2010 International conference on networking and information technology (ICNIT). pp 59–63
Muna A-H, Moustafa N, Sitnikova E (2018) Identification of malicious activities in industrial internet of things based on deep learning models. J Inf Secur Appl 41:1–11
Nguyen HT, Petrović S, Franke K (2010) A comparison of feature-selection methods for intrusion detection. In: International conference on mathematical methods, models, and architectures for computer network security. pp 242–255
Ning L, Jianhua Z (2012) Intrusion detection research based on improved PSO and SVM
Nskh P, Varma MN, Naik RR (2016) Principle component analysis based intrusion detection system using support vector machine. In: IEEE international conference on recent trends in electronics, information & communication technology (RTEICT). pp 1344–1350
Patel A, Taghavi M, Bakhtiyari K, JúNior JC (2013) An intrusion detection and prevention system in cloud computing: a systematic review. J Netw Comput Appl 36:25–41
Peddabachigari S, Abraham A, Grosan C, Thomas J (2007) Modeling intrusion detection system using hybrid intelligent systems. J Netw Comput Appl 30:114–132
Peng H, Sun Z, Zhao X, Tan S, Sun Z (2018) A detection method for anomaly flow in software defined network. IEEE Access 6:27809–27817
Qazanfari K, Mirpouryan MS, Gharaee H (2012) A novel hybrid anomaly based intrusion detection method. In: 2012 Sixth international symposium on telecommunications (IST). pp 942–947
Qi J, Yang P, Min G, Amft O, Dong F, Xu L (2017) Advanced internet of things for personalised healthcare systems: a survey. Pervasive Mob Comput 41:132–149
Ramamoorthi A, Subbulakshmi T, Shalinie SM (2011) Real time detection and classification of DDoS attacks using enhanced SVM with string kernels. In: 2011 International conference on recent trends in information technology (ICRTIT). pp 91–96
Rasheed W, Tang TB (2019) Anomaly detection of moderate traumatic brain injury using auto-regularized multi-instance one-class SVM. IEEE Trans Neural Syst Rehabil Eng
Reddy RR, Ramadevi Y, Sunitha KN (2016) Effective discriminant function for intrusion detection using SVM. In: 2016 International conference on advances in computing, communications and informatics (ICACCI). pp 1148–1153
Renjit JA, Shunmuganathan K (2011) Multi-agent-based anomaly intrusion detection. Inf Secur J A Glob Perspect 20:185–193
Saied A, Overill RE, Radzik T (2016) Detection of known and unknown DDoS attacks using Artificial Neural Networks. Neurocomputing 172:385–393
Sallay H, Ammar A, Saad MB, Bourouis S (2013) A real time adaptive intrusion detection alert classifier for high speed networks. In: 2013 12th IEEE international symposium on network computing and applications (NCA). pp 73–80
Sani RA, Ghasemi A (2015) Learning a new distance metric to improve an SVM-clustering based intrusion detection system. In: 2015 International symposium on artificial intelligence and signal processing (AISP). pp 284–289
Senthilnayaki B, Venkatalakshmi K, Kannan A (2015) Intrusion detection using optimal genetic feature selection and SVM based classifier. In: 2015 3rd international conference on signal processing, communication and networking (ICSCN). pp 1–4
Serkani E, Gharaee-Garakani H, Mohammadzadeh N (2019) Anomaly detection using SVM as classifier and decision tree for optimizing feature vectors. ISeCure-The ISC Int J Inf Secur 11:159–171
Shang W, Li L, Wan M, Zeng P (2015) Industrial communication intrusion detection algorithm based on improved one-class SVM. In: 2015 World congress on industrial control systems security (WCICSS). pp 21–25
Shang W, Cui J, Song C, Zhao J, Zeng P (2018) Research on industrial control anomaly detection based on FCM and SVM. In: 2018 17th IEEE international conference on trust, security and privacy in computing and communications/12th IEEE international conference on big data science and engineering (TrustCom/BigDataSE). pp 218–222
Sindhu SSS, Geetha S, Kannan A (2012) Decision tree based light weight intrusion detection using a wrapper approach. Expert Syst Appl 39:129–141
Singh K, Singh P, Kumar K (2016) A systematic review of IP traceback schemes for denial of service attacks. Comput Secur 56:111–139
Song G, Guo J, Nie Y (2011) An intrusion detection method based on multiple kernel support vector machine. In: 2011 International

conference on network computing and information security (NCIS). pp 119–123
Subbulakshmi T, BalaKrishnan K, Shalinie SM, AnandKumar D, GanapathiSubramanian V, Kannathal K (2011). Detection of DDoS attacks using Enhanced Support Vector Machines with real time generated dataset. In: 2011 Third international conference on advanced computing (ICoAC). pp 17–22
Tang P, Jiang R-a, Zhao M (2010) Feature selection and design of intrusion detection system based on k-means and triangle area support vector machine. In: Second international conference on future networks, 2010. ICFN'10. pp 144–148
Tang X, Tan SX-D, Chen H-B (2018) SVM based intrusion detection using nonlinear scaling scheme. In: 2018 14th IEEE international conference on solid-state and integrated circuit technology (ICSICT). pp 1–4
Tang X, Cao R, Cheng J, Fan D, Tu W (2019) DDoS attack detection method based on V-support vector machine. In: International symposium on cyberspace safety and security. pp 42–56
Teng S, Wu N, Zhu H, Teng L, Zhang W (2018) SVM-DT-based adaptive and collaborative intrusion detection. IEEE/CAA J Autom Sin 5:108–118
Thaseen IS, Kumar CA (2014) Intrusion detection model using fusion of PCA and optimized SVM. In: 2014 International conference on contemporary computing and informatics (IC3I). pp 879–884
Tian J, Gu H (2010) Anomaly detection combining one-class SVMs and particle swarm optimization algorithms. Nonlinear Dyn 61:303–310
Tian Y, Mirzabagheri M, Bamakan SMH, Wang H, Qu Q (2018) Ramp loss one-class support vector machine; a robust and effective approach to anomaly detection problems. Neurocomputing 310:223–235
Wang X-Y, Zhang H-M, Gao H-H (2008) Quantum particle swarm optimization based network intrusion feature selection and detection. IFAC Proc Vol 41:12312–12317
Wang F, Qian Y, Dai Y, Wang Z (2010) A model based on hybrid support vector machine and self-organizing map for anomaly detection. In: 2010 International conference on communications and mobile computing (CMC). pp 97–101
Wang W, Liu J, Pitsilis G, Zhang X (2016) Abstracting massive data for lightweight intrusion detection in computer networks. Inf Sci
Wang H, Gu J, Wang S (2017) An effective intrusion detection framework based on SVM with feature augmentation. Knowl Based Syst 136:130–139
Wani AR, Rana Q, Saxena U, Pandey N (2019) Analysis and detection of DDoS attacks on cloud computing environment using machine learning techniques. In: 2019 Amity international conference on artificial intelligence (AICAI). pp 870–875
Wressnegger C, Schwenk G, Arp D, Rieck K (2013) A close look on n-grams in intrusion detection: anomaly detection vs. classification. In: Proceedings of the 2013 ACM workshop on artificial intelligence and security. pp 67–76
Xie Y, Zhang Y (2012) An intelligent anomaly analysis for intrusion detection based on SVM. In: 2012 International conference on computer science and information processing (CSIP). pp 739–742
Yan Q, Yu FR, Gong Q, Li J (2015) Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: a survey, some research issues, and challenges. IEEE Commun Surv Tutor 18:602–622
Yang M, Rajasegarar S, Erfani SM, Leckie C (2019) Deep learning and one-class SVM based anomalous crowd detection. In: 2019 International joint conference on neural networks (IJCNN). pp 1–8
Yaseen M, Saleem K, Orgun MA, Derhab A, Abbas H, Al-Muhtadi J, Iqbal W, Rashid I (2018) Secure sensors data acquisition and communication protection in eHealthcare: review on the state of the art. Telemat Inform 35:702–726
Yessad N, Omar M, Tari A, Bouabdallah A (2018) QoS-based routing in wireless body area networks: a survey and taxonomy. Computing 100:245–275
Yi Y, Wu J, Xu W (2011) Incremental SVM based on reserved set for network intrusion detection. Expert Syst Appl 38:7698–7707
Yuan J, Li H, Ding S, Cao L (2010) Intrusion detection model based on improved support vector machine. In: 2010 Third international symposium on intelligent information technology and security informatics (IITSI). pp 465–469
Zaman M, Lung C-H (2018) Evaluation of machine learning techniques for network intrusion detection. In: NOMS 2018-2018 IEEE/IFIP network operations and management symposium. pp 1–5
Zarpelão BB, Miani RS, Kawakani CT, de Alvarenga SC (2017) A survey of intrusion detection in Internet of Things. J Netw Comput Appl 84:25–37
Zhang Z, Shen H (2005) Application of online-training SVMs for real-time intrusion detection with different considerations. Comput Commun 28:1428–1442
Zhang X, Jia L, Shi H, Tang Z, Wang X (2012) The application of machine learning methods to intrusion detection. In: 2012 Spring congress on engineering and technology (S-CET). pp 1–4
Zhang M, Xu B, Gong J (2015) An anomaly detection model based on one-class svm to detect network intrusions. In: 2015 11th International conference on mobile ad-hoc and sensor networks (MSN). pp 102–107
Zhang Y, Yang Q, Lambotharan S, Kyriakopoulos K, Ghafir I, AsSadhan B (2019) Anomaly-based network intrusion detection using SVM. In: 2019 11th International conference on wireless communications and signal processing (WCSP). pp 1–6
Zhou CV, Leckie C, Karunasekera S (2010) A survey of coordinated attacks and collaborative intrusion detection. Comput Secur 29:124–140