An ensemble model to detect packet length covert channels

Muawia A. Elsadig, Ahmed Gafar
Scientific Research Department, Deanship of Scientific Research, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Article Info

Article history:

Received Sep 24, 2022
Revised Jan 21, 2023
Accepted Feb 4, 2023

Keywords:

Classification algorithms
Covert channel detection
Deep learning
Machine learning
Packet length covert channel

ABSTRACT

Covert channel techniques have enriched the way to commit dangerous and unwatched attacks. They exploit ways that are not intended to convey information; therefore, traditional security measures cannot detect them. One class of covert channels that difficult to detect, mitigate, or eliminate is packet length covert channels. This class of covert channels takes advantage of packet length variations to convey covert information. Numerous research articles reflect the useful use of machine learning (ML) classification approaches to discover covert channels. Therefore, this study presented an efficient ensemble classification model to detect such types of attacks. The ensemble model consists of five machine learning algorithms representing the base classifiers. The base classifiers include naive Bayes (NB), decision tree (DT), support vector machine (SVM), k-nearest neighbor (KNN), and random forest (RF). Whereas, the logistic regression (LR) classifier was employed to aggregate the outputs of the base classifiers and thus to generate the ensemble classifier output. The results showed a good performance of our proposed ensemble classifier. It beats all single classification algorithms, with a 99.3% accuracy rate and negligible classification errors.

Corresponding Author:

Muawia A. Elsadig
Scientific Research Department, Deanship of Scientific Research, Imam Abdulrahman Bin Faisal University
Dammam, Saudi Arabia
Email: muawiasadig@yahoo.com

1. INTRODUCTION

A covert channel is a channel that is exploited to transfer secret data covertly in a way violates system security policy [1], [2] a sender uses covert channels to communicate with a receiver in order to send him\her a covert message that cannot be seen by others [3]. In 1973, the concept of covert channel was presented by Lampson [4] to be executed in stand-alone machine environment. Where, in 1987, this concept was expanded to include computer networks [5]. Recent research indicated that the existence of network covert channels, which offer significant assistance for undertaking covered communications including transmitting secret data and/or stealing confidential information, is one of the major risks to internet security [6]. Covert channels in networks, are techniques that aim to hide information in normal network traffic to stay undiscovered [7], and have recently attracted more attention [8]. Actually, the environment for building various types of covert channels on computer networks has gotten rich [9], [10]. In addition, even the new technologies in this area do not pay much attention in the design phases to the weaknesses that can be used to establish different kinds of covert channel threats [11]. Some factors that play an important role in the development of covert channel attacks on networks were presented in [12]. These involve i) the rapid advancement of network techniques that help in introducing different scenarios for committing serious attacks that based on covert channels, ii) switch techniques that offer the capability for secret messages to change their locations, which complicates the way to detect such attacks, and iii) internal control protocol techniques that utilize micro protocol approaches to secure dynamic route and reliable communication to covert traffic.

Covert channels that use data packet lengths as a carrier have become popular approaches to hiding secret messages. In which the difference of data packet lengths is exploited to send a secret data. In the literature, there are many techniques that exploit network packet lengths which lead to form different types of covert channel attacks. Voice over internet protocol (VoIP) traffic is an enrich environment for building packet length covert channels, as it includes continued transmission of huge amount of data that attract creating such covert channels. Especially mobile VoIP, which becomes popular for transmitting large amounts of data, and therefore mobile VoIP traffic becomes a potential target to establish covert channels with high throughput.

Unfortunately, packet length covert channels can be applied, even if encryption methods are implemented and can exploit several network protocols. Resolving such covert channels can be attained by identifying and developing some indicators to help in building adequate detection approaches. Simply it is possible to eliminate this kind of threats, packet length covert channels, by having all packets taking the same length but this approach diminishes the network capacity [13] and therefore is not considered as an adequate solution to such type of covert channels. The use of machine learning methods to combat covert channel attacks has become a hot topic for network security [14]. Eventually, the ongoing development of different techniques to construct covert channel attacks is clearly noticeable. Moreover, covert communication can be exploited to carry out different malicious activities that pose massive risks to our data and confidentiality. Therefore, more research efforts are required to develop adequate solutions capable to counter these threats.

In this paper stacking ensemble model has been introduced to predict the presence of any attack caused by covert channels that exploiting packet lengths. This model is a machine learning design that uses stacking method to aggregate the results of the base classifiers to generate better results. The ensemble model has five classifiers as the base classifiers, whereas the logistic regression (LR) classifier is the meta-classifier that is used to combined the base classifiers' output to form the ensemble model output. This section provides brief information to covert channels, illustrating their basic concept with focusing on packet-length-based covert channels. The next section highlighted the main common types of covert channels. Subsequently, relevant work is given in section 3 which provides general overview in covert channels that use packet lengths and sheds lights on some detection methods that were presented to counter them. Section 4 gives a description to our proposed classifier. It illustrates the methods to construct and evaluate the proposed approach. Section 5 provides a discussion of the results obtained, whereas the study is concluded in section 6.

2. COVERT CHANNEL TYPES

Covert channels are mainly divided into two types: storage covert channels (SCC) and timing covert channels (TCT) [15]. In storage channels, storage location is used to hold a covert message [16] such as protocol fields, whereas in timing channels, a covert message is modulated using the timing characteristics of the network traffic. In other words, covert storage channels use an object to send a covert message while covert timing channels use timing aspects to convey covert messages. The construction of covert channels by combining both storage and timing techniques leads to a third type that takes the advantages of both categories, timing and storage, and therefore it has the capability to pose real challenge. It is known as hybrid channels.

3. RELATED WORK

To avoid being detected, the attackers are motivated to construct hidden channels to convey secret data. These channels violate the security policy of the system and cannot be detected using traditional intrusion detection methods. The construction of covert channel over network has become modern way to leak information and poses real challenges. For covert channels that exploit packet lengths, the detection of such channels may be more difficult, as some of them generate covert traffic that is highly imitated normal traffic, and therefore the detection methods that watch the variations between normal and covert traffic may fail to catch them.

3.1. Overview

The concept of covert channels that exploit the variations of network packet lengths to pass secret messages was initially proposed by Padlipsky *et al.* [17] and Girling [5]. In these covert channels which presented in [5], [17], the transmission parties (sender and receiver) should share some rules before starting their transmission session. However, these approaches are vulnerable to detection. They cause notable changes that can be caught by detection methods and therefore it is easy to identify such channels [18].

In the covert channel proposed in [19], a matrix of unique packet lengths should be shared between the two parts that intend to initiate covert communication before starting their communication session. However, their proposed covert channel can be detected [20]. Ji *et al.* [20] proposed a packet length covert channel which has the ability to deliver normal network traffic; therefore, detection methods fail to discover it. However, Nair *et al.* [21] proposed a detection method that is capable to detect their channel. A packet length-based covert channel proposed by Ji *et al.* [22] that was expected to cause no abnormal behaviours and therefore the authors indicated that their approach has great resistance to detection methods.

Hussain and Hussain [23] proposed a covert channel with high capacity of covert message. It exploits packet lengths and data payload. However, the construction of this type of covert channel is complicated and may be vulnerable to detection due to the use of the data payload [24].

Abdullaziz *et al.* [25] proposed a covert channel based on network packet length. One packet can carry 1 bit of a covert message. This covert channel has high resistance to be detected and does not require a shared key as many packet-length covert channels require; however, it has low bandwidth [26] and data shifting may occur as a result of user datagram protocol (UDP) packet drop [27].

Sabeti and Shoaei [26] introduced two approaches that are expected to have greater resistance against detection methods. In these approaches, two lengths are used to modulate one bit and it may necessary to change their order to fit the secret message. However, long covert messages require huge number of network packets as one bit of a covert message requires pair of network packets. Our ensemble approach that has been proposed by this study is expected to detect these channels because swapping of the packets causes abnormality that can be discovered by our proposed approach.

Liang *et al.* [28] constructed a packet length covert channel using mobile VoIP by mapping the length of the packet with covert data. As the authors stated, the throughput, undetectability, and reliability of the scheme were evaluated. However, in the event that a malicious party is able to decrypt the packet content, the original ordering of the packets can be obtained and therefore this breaks the covert communication.

3.2. Detection using machine learning

Packet length covert channels can be eliminated by getting all packet lengths equal in size. However, this method is computationally expensive and thus will affect network performance. The process of reconstructing network packets by examining the size of each packet and then applying either padding or splitting technique definitely causes more overheads. Therefore, developing detection methods that capable to spot such attack is appropriate alternative.

The use of machine learning and deep learning technologies has been demonstrated in a number of domains, including information security. Numerous machine learning techniques have been used to identify harmful behaviour and exhibit high performance. Tang *et al.* [14] indicated that the application of machine learning techniques to thwart covert communication has become a hot topic in network security. Readers who are interested can view some recent studies that use machine learning or deep learning techniques to thwart covert channel attacks in [29]–[34]. As an example, Dua *et al.* [6] proposed a two-layered system to locate and detect storage covert channels that use the IPv6 protocol. deep neural networks (DNN) and the one-vs-rest technique with support vector machine (SVM) are used in their suggested system. The authors use their own dataset, which was produced using the pcapStego tool and the CAIDA dataset, to evaluate their system. With a prediction time of 0.0719 seconds and high accuracy that reaches 99.7, their suggested system is suitable for use in real-time applications as the authors stated.

Each learning algorithm has its advantages and drawbacks, not a perfect one at all; therefore, ensemble approaches may improve performance by aggregating different classification models so to boost up their advantages and to reduce their limitations. Stacking is an ensemble technique where a meta-classifier is used to combining the outputs of different classifiers to enhance the prediction process. This method produces powerful models that capable to improve classification accuracy. Yang *et al.* [35] proposed a detection approach based on stacking technique, which is an aggregation method that combines weak learners in order to produce a more powerful classifier. The base classifiers for their model are k-nearest neighbors (KNN), SVM, and random forest (RF), while a neural network (NN) model is used to aggregate them. The suggested approach is aimed to detect covert channel attacks that exploiting DNS. In comparison with many existing approaches, this approach was performing better. In the same context, Yang *et al.* [36] proposed a detection method based on stacking technique to detect domain name system (DNS) covert channel. Their results showed the effectiveness of their model to identify the targeted covert channels. The authors of this model stated that their model is capable for identifying even unknown covert channel traffic. Furthermore, Yang *et al.* [37] proposed stacking model to detect a DNS covert channel. Their model showed excellent performance compared to existing approaches, as the authors indicated.

Cassavia *et al.* [38] proposed a deep ensemble-based model to detect covert channels that exploits time to live (TTL) filed of IPv4 protocol. The model's ability to detect the presence of these covert channels was proved by the authors, who also emphasized that the model was built to be lightweight and only require a small number of training cases. Whereas, Li *et al.* [39] stated that stacking classifications models may be used as a more affordable alternative to deep learning classification approaches and they presented a stacking-based ensemble classification approach. In which, the employed primary classifiers (base classifiers) are KNN, LR, SVM, and naive Bayes (NB) whereas RF was used as a secondary classifier (meta-classifier). Their proposed model showed better performance compared to the other models in both middle and small size of dataset with acceptable time-consuming.

A review article reviewed some techniques of network steganalysis and steganography with focusing on their shortcomings and novelties. The article reported that more research work is required on network steganalysis [40]. This is to fill the gap as the network steganography techniques have witnessed rapid increases as a result of the advanced development of network technologies.

4. METHOD

This research presents an ensemble classification model for covert channel attacks that are exploited network packet lengths to initiate illegal ways to exchange hidden data. In this type of covert channels, the covert sender exploits the packet length values to convey a covert message he/she wants to send. Odd represent 1 and even represent 0 or the opposite scenario; therefore, it may require to change the packet lengths to fit the covert message. The covert receiver watches the lengths of the received packets to decipher the covert message.

4.1. Dataset

A dataset has been created according to the above description of our target covert channel. Wireshark 3.4.2 was employed to capture network traffic. Then, the recorded traffic was divided into two parts. The first part was kept as is to represent normal traffic. The Scapy 2.4.4 and Python 2.7.18 were used to modify the rest of the captured traffic to construct covert traffic. Accordingly, a dataset of 300 instances was developed. It contains 150 cases of normal traffic and 150 cases of covert traffic.

4.2. Data preprocessing and feature extraction

In this study, a bag of words method is used for feature extraction purposes which is a flexible way for feature extraction. This method is applied after the preprocessing phase. The bag-of-words method is a simplified representation of text using criteria such as occurrence of words (word frequency). It has been used in many fields such as natural language processing, and computer vision [41]. Many recent works employed this method to enhance classification accuracy such as [42]–[46]. The preprocessing phase includes applying a filter to consider only the most important tokens that have significant influence in prediction accuracy. Orange 3.28 was used to implement both phases, data preprocessing and feature extraction. Orange is a data mining package to analyse data [47].

4.3. The proposed classifier

The proposed ensample model uses a stacking strategy to aggregate the output of a group of classifies to enhance the prediction accuracy. It is commonly known that ensemble classifiers achieve a higher accuracy rate compared to single classification models. The base classifiers of our ensemble model were selected carefully based on their reputation in reflecting good performance. They include KNN, decision tree (DT), SVM, NB and RF, while LR classifier is used as meta-model to aggregate the outputs of the base classifiers.

4.4. Experiments and evaluation

The base classifiers of our proposed ensemble model were fully trained and tested using the created dataset that consists of 300 samples of normal (overt) traffic and malicious (covert) traffic. To build the ensemble model, LR classifier is used to combine the base classifiers whereas a cross validation method is applied for the model evaluation. The cross-validation method is a sampling method that uses different parts of the given dataset to train and test a machine learning model on different iterations. It is a resampling method used to evaluate machine learning models [48]. Interested readers for more information on cross-validation methods are referred to [49]. In our case, the 10-fold cross-validation method was used. 90% of the dataset used for training, and the rest for testing. In the 10-fold cross-validation method, the dataset is divided into ten parts (folds). Then the model is trained and tested ten times. In each time, different fold is used for testing and the rest of the dataset for training. Each time the testing and training data are changed, which leads to a more comprehensive validation. The performance of our proposed ensemble model compared to individual classifies was examined by computing the confusion matrix that gives the number of cases that are correctly classified compared to misclassified cases. Furthermore, some other evaluation measures were calculated: the classification accuracy, sensitivity, precision, the area under the curve (AUC) and specificity were calculated.

5. RESULTS AND DISCUSSION

As it shown by Table 1, the reported results indicated that our proposed ensemble classifier achieved high accuracy reached 99.3% to detect packet length covert channel attacks. It outperformed all ensemble base classifiers (DT, KNN, SVM, RF, and NB). In addition, the ensemble classier reported high performance when considering other performance evaluation metrics that include AUC, recall, specificity and precision. Figure 1 visualizes the accuracy of the proposed stacking ensemble model compared to other classifiers, which clearly depicts that the new model has outperformed the others. The classification errors, that were evaluated using confusion matrix, are displayed in Table 2, which illustrates good performance of the proposed classifier by casing negligible classification errors that tend to zero. This confirms our assumption that ensemble models have the capabilities to improve perdition accuracy. More evaluation was done by computing the receiver operating characteristic (ROC) curves of all classifiers. Figures 2 to 7 show these curves for KNN, RF, NB, SVM, DT and the proposed classifier, respectively. ROC curves have proven that the proposed classifier outperformed all other classifiers by showing better performance. The performed experiments and the evaluation measures employed have proven that our proposed classifier has enhanced the detection accuracy of packet length covert channel. It achieved the high accuracy with the least classification errors compared to other classifiers. It is noteworthy to mention that RF classifier achieved good classification accuracy that reaches 98.7% followed by NB which achieved 97.7% but still our proposed classifier outperforms them by reaching 99.3% accuracy rate.

Table 1. Classification performance

Model	AUC	F1 specificity	Sensitivity (recall)	Precision	Accuracy
KNN	96.7%	86.5%	100%	67.1%	84.3%
DT	93.6%	93.5%	96%	91.1%	93.3%
NB	100%	97.7%	99.3%	97.1%	97.7%
RF	99.9%	98.7%	99.3%	98%	98.7%
SVM	99.7%	95.2%	100%	90.9%	95%
The ensemble classifier	100%	99.3%	99.3%	99.3%	99.3%

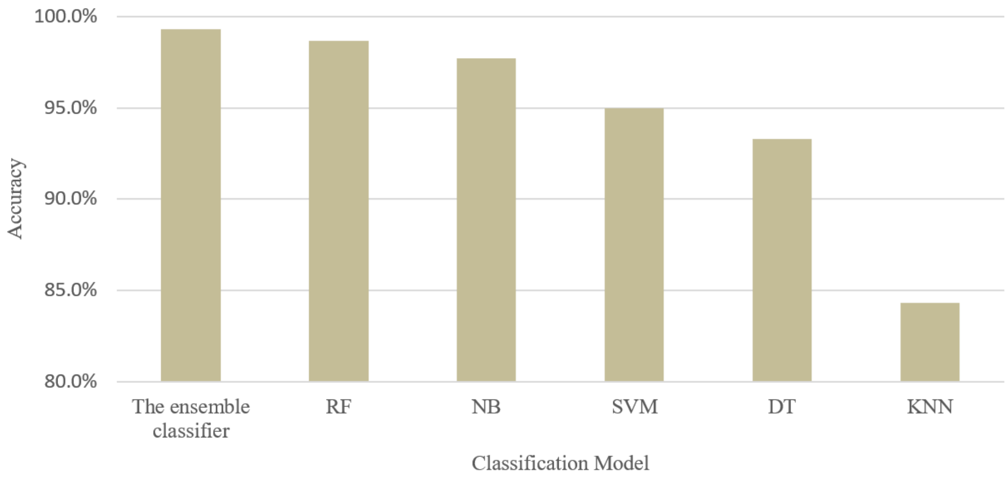

Figure 1. Accuracy of the classifiers

Table 2. Classification errors

Model	False positive (FP)	False negative (FN)
NB	0.04	0.01
RF	0.01	0.02
SVM	0.10	0.00
KNN	0.13	0.00
DT	0.09	0.04
The ensemble classifier	0.01	0.01

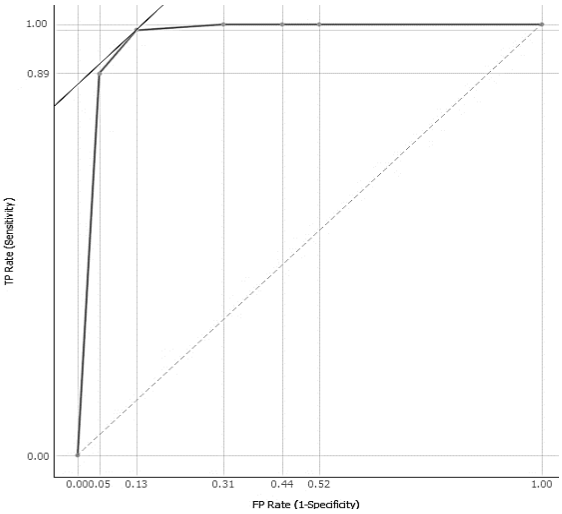

Figure 2. KNN receiver operating characteristic curve

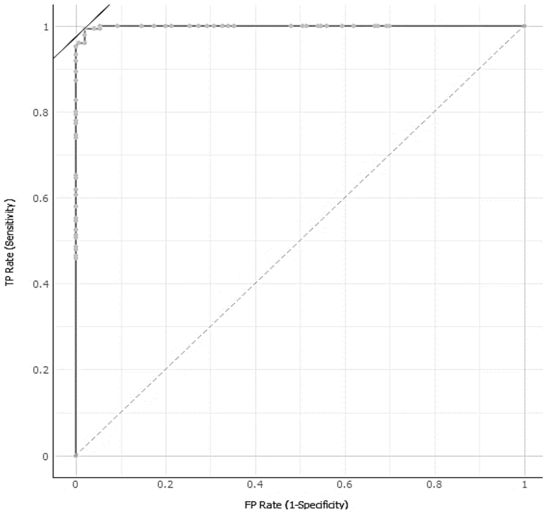

Figure 3. RF receiver operating characteristic curve

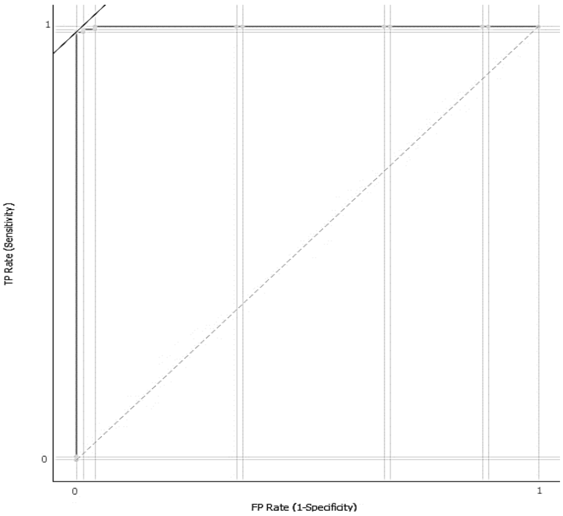

Figure 4. NB receiver operating characteristic curve

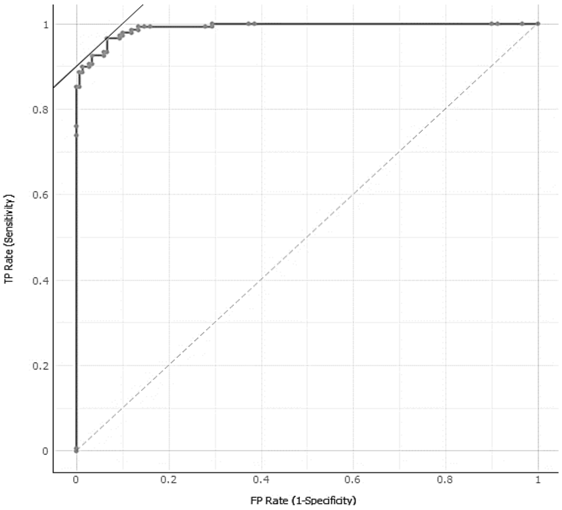

Figure 5. SVM receiver operating characteristic curve

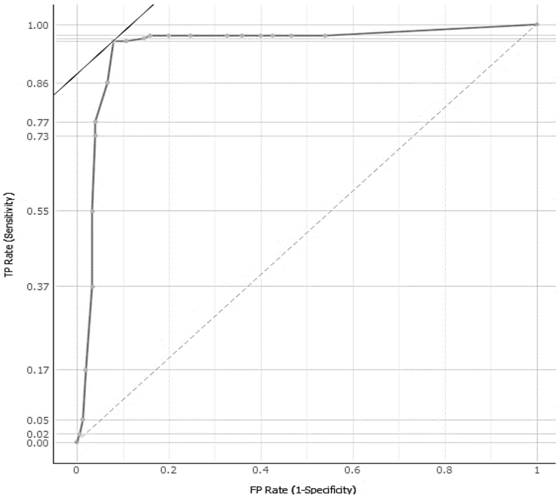

Figure 6. DT receiver operating characteristic curve

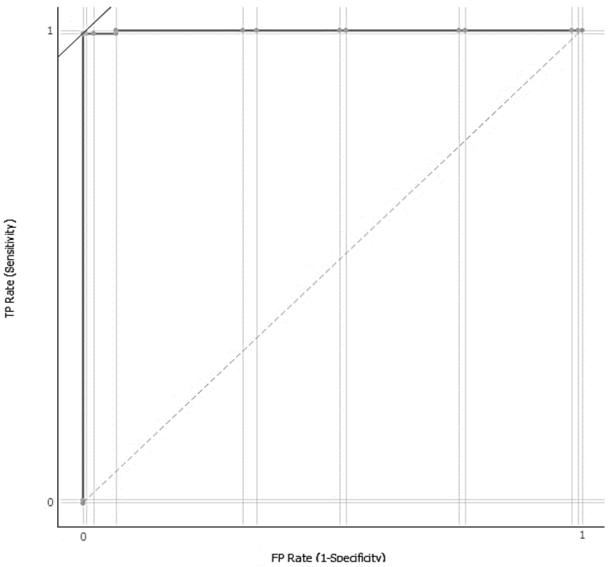

Figure 7. The ensemble classifier receiver operating characteristic curve

6. CONCLUSION

Machine learning technology has proven its usefulness in various fields, especially in information security. Many machine learning approaches have been applied to detect malicious activities and reflected high performance. In this study, an ensemble classification model was presented to discover the existence of covert channels that exploit data packet lengths to pass covert messages. This attack exploits the variations of packet lengths to leak secret information. It is an undetectable and dangerous attack which is capable of generating malicious traffic (covert traffic) that is typically look like normal traffic (overt traffic), and thus detection approaches would fail to spot the variations. The developed ensemble detection scheme achieved outstanding performance compared to the single classification models. Its detection accuracy reached 99.3% with neglected classification errors. Up to our knowledge, the proposed ensemble model has achieved a classification accuracy rate that outperformed all existing detection methods presented to discover such attack.

ACKNOWLEDGEMENTS

The author would like to thank the Deanship of Scientific Research for supporting this research work, project No. 2019-342.

REFERENCES

[1] C. Alcaraz, G. Bernieri, F. Pascucci, J. Lopez, and R. Setola, "Covert channels-based stealth attacks in industry 4.0," *IEEE Systems Journal*, vol. 13, no. 4, pp. 3980–3988, Dec. 2019, doi: 10.1109/JSYST.2019.2912308.

[2] F. Iglesias, F. Meghdouri, R. Annessi, and T. Zseby, "CCgen: injecting covert channels into network traffic," *Security and Communication Networks*, pp. 1–11, May 2022, doi: 10.1155/2022/2254959.

[3] C. Zhang, L. Zhu, C. Xu, Z. Zhang, and R. Lu, "EBDL: Effective blockchain-based covert storage channel with dynamic labels," *Journal of Network and Computer Applications*, vol. 210, Jan. 2023, doi: 10.1016/j.jnca.2022.103541.

[4] B. W. Lampson, "A note on the confinement problem," *Communications of the ACM*, vol. 16, no. 10, pp. 613–615, Oct. 1973, doi: 10.1145/362375.362389.

[5] C. G. Girling, "Covert channels in LAN's," *IEEE Transactions on Software Engineering*, no. 2, pp. 292–296, Feb. 1987, doi: 10.1109/TSE.1987.233153.

[6] A. Dua, V. Jindal, and P. Bedi, "Detecting and locating storage-based covert channels in internet protocol version 6," *IEEE Access*, vol. 10, pp. 110661–110675, 2022, doi: 10.1109/ACCESS.2022.3215132.

[7] J. Hielscher, K. Lamshöft, C. Krätzer, and J. Dittmann, "A systematic analysis of covert channels in the network time protocol," in *The 16th International Conference on Availability, Reliability and Security*, Aug. 2021, pp. 1–11, doi: 10.1145/3465481.3470075.

[8] K. Lamshöft, T. Neubert, J. Hielscher, C. Vielhauer, and J. Dittmann, "Knock, knock, log: Threat analysis, detection and mitigation of covert channels in syslog using port scans as cover," *Forensic Science International: Digital Investigation*, vol. 40, Apr. 2022, doi: 10.1016/j.fsidi.2022.301335.

[9] M. A. Elsadig and Y. A. Fadlalla, "Network protocol covert channels: countermeasures techniques," in *2017 9th IEEE-GCC Conference and Exhibition (GCCCE)*, May 2017, pp. 1–9, doi: 10.1109/IEEEGCC.2017.8447997.

[10] M. A. Elsadig and Y. A. Fadlalla, "Survey on covert storage channel in computer network protocols: detection and mitigation

techniques," *International Journal of Advances in Computer Networks and Its Security*, vol. 6, no. 3, pp. 11–17, 2016.
[11] M. A. Elsadig and A. Gafar, "Covert channel detection: machine learning approaches," *IEEE Access*, vol. 10, pp. 38391–38405, 2022, doi: 10.1109/ACCESS.2022.3164392.
[12] A. Epishkina and K. Kogos, "A traffic padding to limit packet size covert channels," in *2015 3rd International Conference on Future Internet of Things and Cloud*, Aug. 2015, pp. 519–525, doi: 10.1109/FiCloud.2015.20.
[13] A. Epishkina, K. Kogos, and D. Frolova, "A technique to limit hybrid covert channel capacity via random increasing of packets' lengths," *Procedia Computer Science*, vol. 190, pp. 231–240, 2021, doi: 10.1016/j.procs.2021.06.029.
[14] Z. Tang, J. Wang, H. Li, J. Zhang, and J. Wang, "Cognitive covert traffic synthesis method based on generative adversarial network," *Wireless Communications and Mobile Computing*, pp. 1–14, Jun. 2021, doi: 10.1155/2021/9982351.
[15] L. Zhang, T. Huang, W. Rasheed, X. Hu, and C. Zhao, "An enlarging-the-capacity packet sorting covert channel," *IEEE Access*, vol. 7, pp. 145634–145640, 2019, doi: 10.1109/ACCESS.2019.2945320.
[16] M. A. Elsadig and Y. A. Fadlalla, "A balanced approach to eliminate packet length-based covert channels," in *2017 4th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, Nov. 2017, pp. 1–7, doi: 10.1109/ICETAS.2017.8277839.
[17] M. A. Padlipsky, D. W. Snow, and P. A. Karger, "Limitations of end-to-end encryption in secure computer networks," 1978.
[18] M. A. Elsadig and Y. A. Fadlalla, "Packet length covert channels crashed," *Journal of Computer Science and Computational Mathematics*, pp. 59–66, Dec. 2018, doi: 10.20967/jcscm.2018.04.001.
[19] Z. Peng, "Coverting channel based on packet length," *Computer Engineering*, 2008.
[20] L. Ji, W. Jiang, B. Dai, and X. Niu, "A novel covert channel based on length of messages," in *2009 International Symposium on Information Engineering and Electronic Commerce*, 2009, pp. 551–554, doi: 10.1109/IEEC.2009.122.
[21] A. S. Nair, A. Sur, and S. Nandi, "Detection of packet length based network steganography," in *2010 International Conference on Multimedia Information Networking and Security*, 2010, pp. 574–578, doi: 10.1109/MINES.2010.126.
[22] L. Ji, H. Liang, Y. Song, and X. Niu, "A normal-traffic network covert channel," in *2009 International Conference on Computational Intelligence and Security*, 2009, pp. 499–503, doi: 10.1109/CIS.2009.156.
[23] M. Hussain and M. Hussain, "A high bandwidth covert channel in network protocol," in *2011 International Conference on Information and Communication Technologies*, Jul. 2011, pp. 1–6, doi: 10.1109/ICICT.2011.5983562.
[24] M. A. Elsadig and Y. A. Fadlalla, "Packet length covert channel: a detection scheme," in *2018 1st International Conference on Computer Applications and Information Security (ICCAIS)*, Apr. 2018, pp. 1–7, doi: 10.1109/CAIS.2018.8442026.
[25] O. I. Abdullaziz, V. T. Goh, H.-C. Ling, and K. Wong, "Network packet payload parity based steganography," in *2013 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (CSUDET)*, May 2013, pp. 56–59, doi: 10.1109/CSUDET.2013.6670985.
[26] V. Sabeti and M. Shoaei, "New high secure network steganography method based on packet length," *The ISC International Journal of Information Security*, vol. 12, no. 1, 2020.
[27] J. O. Seo, S. Manoharan, and A. Mahanti, "A discussion and review of network steganography," in *2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, Aug. 2016, pp. 384–391, doi: 10.1109/DASC-PICom-DataCom-CyberSciTec.2016.80.
[28] C. Liang, Y. Tan, X. Zhang, X. Wang, J. Zheng, and Q. Zhang, "Building packet length covert channel over mobile VoIP traffics," *Journal of Network and Computer Applications*, vol. 118, pp. 144–153, Sep. 2018, doi: 10.1016/j.jnca.2018.06.012.
[29] S. Al-Eidi, O. Darwish, Y. Chen, and M. Elkhodr, "Covert timing channels detection based on image processing using deep learning," in *Advanced Information Networking and Applications*, Springer International Publishing, 2022, pp. 546–555.
[30] S V Hayrapetyan and K. S. Zaytsev, "Network covert channels detection method for packet data transmission networks security increase," *International Journal of Open Information Technologies*, vol. 10, no. 8, pp. 30–38, 2022.
[31] S. Al-Eidi, O. Darwish, Y. Chen, and G. Husari, "SnapCatch: automatic detection of covert timing channels using image processing and machine learning," *IEEE Access*, vol. 9, pp. 177–191, 2021, doi: 10.1109/ACCESS.2020.3046234.
[32] M. Guarascio, M. Zuppelli, N. Cassavia, G. Manco, and L. Caviglione, "Detection of network covert channels in IoT ecosystems using machine learning," 2022.
[33] F. Massimi and F. Benedetto, "Deep learning-based detection methods for covert communications in e- health transmissions," in *2022 45th International Conference on Telecommunications and Signal Processing (TSP)*, Jul. 2022, pp. 11–16, doi: 10.1109/TSP55681.2022.9851366.
[34] A. Dua, V. Jindal, and P. Bedi, "DICCh-D: detecting IPv6-based covert channels using DNN," in *Communications in Computer and Information Science*, Springer Nature Switzerland, 2022, pp. 42–53.
[35] P. Yang, X. Wan, G. Shi, H. Qu, J. Li, and L. Y. Yang, "Identification of DNS covert channel based on stacking method," *International Journal of Computer and Communication Engineering*, vol. 10, no. 2, pp. 37–51, 2021, doi: 10.17706/IJCCE.2021.10.2.37-51.
[36] P. Yang, Y. Li, and Y. Zang, "Detecting DNS covert channels using stacking model," *China Communications*, vol. 17, no. 10, pp. 183–194, Oct. 2020, doi: 10.23919/JCC.2020.10.013.
[37] P. Yang, X. Wan, G. Shi, H. Qu, J. Li, and L. Yang, "Naruto," in *Proceedings of the 2020 The 2nd World Symposium on Software Engineering*, Sep. 2020, pp. 109–115, doi: 10.1145/3425329.3425336.
[38] N. Cassavia, L. Caviglione, M. Guarascio, A. Liguori, and M. Zuppelli, "Ensembling sparse autoencoders for network covert channel detection in IoT ecosystems," in *Lecture Notes in Computer Science*, Springer International Publishing, 2022, pp. 209–218.
[39] H. Li, Y. Jin, J. Zhong, and R. Zhao, "A fruit tree disease diagnosis model based on stacking ensemble learning," *Complexity*, pp. 1–12, Sep. 2021, doi: 10.1155/2021/6868592.
[40] J. O. Seo, S. Manoharan, and A. Mahanti, "Network steganography and steganalysis-a concise review," in *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, 2016, pp. 368–371, doi: 10.1109/ICATCCT.2016.7912025.
[41] C. A. Hirway, E. Fallon, K. Flanagan, and P. Connolly, "Determining receipt validity from e-mail subject line using feature extraction and binary classifiers," *International journal of simulation: systems, science and technology*, May 2022, doi: 10.5013/IJSSST.a.23.02.03.
[42] L. S. Shankar, A. Sravani, T. S. Kumar, S. Rajender, and C. Z. Basha, "Convolution neural network (CNN) based computerized classification of adulterated fruits with SIFT and bag of words (BOW)," in *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Jan. 2022, pp. 1068–1073, doi: 10.1109/ICSSIT53264.2022.9716553.
[43] Y. Barve, J. R. Saini, K. Pal, and K. Kotecha, "A novel evolving sentimental bag-of-words approach for feature extraction to detect misinformation," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 4, 2022, doi:

10.14569/IJACSA.2022.0130431.
[44] J. Lee *et al*., "Unsupervised machine learning for identifying important visual features through bag-of-words using histopathology data from chronic kidney disease," *Scientific Reports*, vol. 12, no. 1, Mar. 2022, doi: 10.1038/s41598-022-08974-8.
[45] N. Bayat, E. Rastegari, and Q. Li, "Human gait recognition using bag of words feature representation method," *arXiv preprint arXiv:2203.13317*, Mar. 2022.
[46] A. Kadriu, L. Abazi, and H. Abazi, "Albanian text classification: bag of words model and word analogies," *Business Systems Research Journal*, vol. 10, no. 1, pp. 74–87, Apr. 2019, doi: 10.2478/bsrj-2019-0006.
[47] J. Demsar *et al.*, "Orange: data mining toolbox in Python," *Journal of Machine Learning Research*, vol. 14, no. 1, pp. 2349–2353, 2013.
[48] R. Wazirali, "An improved intrusion detection system based on KNN hyperparameter tuning and cross-validation," *Arabian Journal for Science and Engineering*, vol. 45, no. 12, pp. 10859–10873, Dec. 2020, doi: 10.1007/s13369-020-04907-7.
[49] M. W. Browne, "Cross-validation methods," *Journal of Mathematical Psychology*, vol. 44, no. 1, pp. 108–132, Mar. 2000, doi: 10.1006/jmps.1999.1279.

BIOGRAPHIES OF AUTHORS

Muawia A. Elsadig received the bachelor degree in computer engineering, the M.Sc. degree in computer networks, and the Ph.D. degree in computer science (information security). Currently, he is an assistant professor of cybersecurity and researcher, Deanship of Scientific Research, Imam Abdulrahman Bin Faisal University (IAU), Dammam KSA. In addition, he has engaged in teaching different computer courses in the college of Computer Science and Information Technology (CSIT) and the college of applied studies. He worked for different accredited international universities and has a rich record of publications at recognized international journals and conferences. He received many awards for his research achievements. Dr. Elsadig has many years of teaching experience and considerable industry contributions. He contributed as a reviewer for many international reputable journals. His research interests include the area of information security, network security, cybersecurity, wireless sensor networks, bioinformatics, and information extraction; ranging from theory to design to implementation. He can be contacted at email: muawiasadig@yahoo.com.

Ahmed Gafar received the B.Sc. and M.Sc. degrees from Omdurman Islamic University, in 2009 and 2011, respectively. He is currently working as a Lecturer with the Deanship of Scientific Research, Imam Abdulrahman Bin Faisal University, Saudi Arabia. His research interests include statistical prediction models, machine learning, and its associated applications. He can be contacted at email: agmohamed@iau.edu.sa.