Enhancing Network Security with RNA-Based Intrusion Detection and Cuckoo Search Algorithm

Omar Fitian Rashid[1][a], Hind Moutaz Al-Dabbas[2][b] and Humam Al-Shahwani[3][c]
[1]Department of Geology, College of Science, University of Baghdad, Baghdad, Iraq
[2]Department of Computer Science, College of Education for Pure Science/Ibn Al-Haitham, University of Baghdad, Baghdad, Iraq
[3]Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq

Keywords: Intrusion Detection System, RNA Encoding, Cuckoo Search Algorithm, Pattern Matching, CICIDS2017 Dataset, Network Security, Anomaly Detection.

Abstract: Intrusion detection systems are an important component of defensive measures protecting computer systems and networks from abuse. The CICIDS2017 dataset is a rich dataset that contains the recent attack patterns on the network, therefore it suitable for IDS models. This paper focuses on the development of a new anomaly-based IDS utilizing RNA encoding coupled with the Cuckoo Search Algorithm for pattern-matching optimization on the CICIDS2017 network traffic dataset. The proposed method involves four steps: picking up the records from the dataset at random, encoding the 83 attributes into RNA sequences, identifying the keys for optimal behaviour through the Cuckoo Search Algorithm, and employing a matching model to categorize records as benign or malicious. The integration of the Cuckoo Search Algorithm helps in adjusting the behaviour key selection process using global exploration (Levy flights) and local optimization that enhances accuracy and the detection rate. Performance analysis of the proposed method was done based on detection rate (DR) for each thread, DR for all threads, false alarm rate (FAR), accuracy, RNA encoding time, and keys matching time. This novel approach has the potential to improve the response to real-time network security threats and the ability to counter new forms of cyber threats.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

Cyberattacks represent an ever-growing threat that has become a real priority for most organizations. Where attackers used various scenarios to trick defence systems into entering the system and private information or causing harm. Detecting attacks has become necessary to save information privacy within computer systems. Where intrusion can have severe consequences, such as financial losses, theft of sensitive data, or even critical infrastructure disruptions (Bilot et al., 2023). An Intrusion Detection System (IDS) is an effective method to secure systems from cyber-attacks (Singh et al., 2022). Emerging technologies such as Big Data, Internet of Things, Cloud Computing, Wireless Sensor Networks, etc. (Pirozmand et al., 2020; Shivhare et al. 2020). Where IDS can be placed between the network switch and firewall, employing port mirroring technology to monitor incoming and outgoing network traffic enables the detection of intrusions. Figure 1 shows a running of IDS connectivity (Azam et al., 2023).

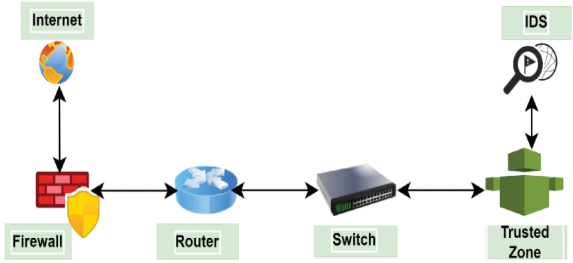

Figure 1: Implementation of IDS.

[a] https://orcid.org/0000-0002-8186-0795
[b] https://orcid.org/0000-0002-3912-2051
[c] https://orcid.org/0000-0003-1248-1991

Rashid, O. F., Moutaz Al-Dabbas, H. and Al-Shahwani, H.
Enhancing Network Security with RNA-Based Intrusion Detection and Cuckoo Search Algorithm.
DOI: 10.5220/0014372600004848
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 2nd International Conference on Advances in Electrical, Electronics, Energy, and Computer Sciences (ICEEECS 2025)*, pages 106-111
ISBN: 978-989-758-783-2
Proceedings Copyright © 2026 by SCITEPRESS – Science and Technology Publications, Lda.

To improve the effectiveness of the IDS and increase the speed of pattern matching, the RNA encoding with Cuckoo Search Algorithm (CSA) based IDS approach is presented in this paper. RNA encoding is applied to the network traffic data to map it into a biological format in order to effectively process the 83 attributes of the dataset. The method of finding out the most typical patterns in this encoded data is done by CSA, an optimization algorithm which mimics the actions of parasitic cuckoo birds. CSA effectively searches for the behaviour keys within the search space and uses Levy flights for the global search. In contrast, random walks are used for local optimization, guaranteeing that only the best patterns are used for classification. The inclusion of CSA into the IDS framework brings in the following benefits. First, CSA performs dynamic pattern matching by going through a population of behaviour keys and thus can respond to changes in the dataset and the threat landscape. This leads to increased DR and decreased FAR as compared to the matching algorithms used in the past.

Second, CSA has a faster computational time than encoding and key matching to make the system suitable for real-time intrusion detection in large-scale networks (Yang and Deb, 2010). The following are the advantages that can be obtained from the use of CSA in the IDS (Gupta and Kalra, 2020; Li et al., 2024):

- Avoiding Local Minima: The other benefit of CSA over the other conventional matching algorithms is that the CSA is not easily stuck in the local optimum. Levy flights are used to ensure that the algorithm can search throughout the whole search space, hence increasing the probability of identifying the best patterns for intrusion detection all over the world. This move also helps in minimizing false positive results and negative results.
- Better Classification: CSA can potentially improve the performance of discriminating malicious traffic from benign traffic by identifying better pattern-matching patterns. This leads to a higher DR and lower FAR due to the fact that the optimization step is in a position to fine tune the matching keys for data.
- Flexibility: CSA, on the other hand, is capable of real-time matching, unlike other methods that are fixed in as much as the traffic flow is concerned. This means that it can be used to always update the behaviour keys that are used in your IDS as new types of attacks surface hence making the IDS more robust.
- Reduced Computation Time: CSA has been established to be efficient in the time complexity and hence can be used in real time intrusion detection. CSA is advantageous because unlike the situation where one has to think of all the patterns and, after that, eliminate the unsuitable ones, CSA is able to identify the suitable patterns right from the onset.

This paper is structured as follows: Section 2 presents the literature review of anomaly-based intrusion detection and optimization algorithms. In section 3, the authors provide the proposed method with a focus on RNA encoding as well as the CSA for extracting key behaviours. In section 4, the experimental setup and the performance evaluation of the proposed system are illustrated using the CICIDS2017 dataset. Finally, Section 5 provides the conclusion of the paper and points to further research.

2 LITERATURE REVIEW

Several new works deal with IDS based on different techniques. A new IDS is implemented by (Jayalatchumy et al. 2024) to enrich the intrusion detection performance, where the proposed method starts by reducing data methodologies, then applies the Crow search method to select the most significant features, and finally, the ensemble classifier method to classify the standard and invader labels. Multiple feature selection methods are applied for anomaly IDS (Eljialy et al., 2024) based on a software-defined networking dataset, where the proposed method uses different feature extraction methods to select the high-scoring features. A novel feature selection method is suggested (Tripathi et al., 2024) based on machine learning algorithms to detect intrusion threats more accurately, and this system is done with target features with a big effect on the target variable, where the achieved results showed an almost 51% reduction in irrelevant features. A genetic algorithm with a clustering method is applied for IDS (Fouad and Hameed, 2022) where the proposed method is used to recognize the incoming network traffic and classify it as either normal or attack. This system has two genetic algorithm models; the first one is used to handle digital features, and the second handles all features. Abdulboriy and Sun Shin, (2024) proposed an improvement method for IDS based on incremental majority voting, and this enhancement is done based on leveraging the collective decision-making power of multiple machine learning models such as the KNN classifier, SoftMax Regressor and Adaptive Random Forest classifier. An anomaly

detection method is suggested (Li et al., 2022); this method is different from the existing and traditional generative adversarial networks method, where the proposed method includes a generator and a discriminator and is used to map the different traffic feature data from separate datasets. Nallakaruppan et al., 2024 executed a comparative test for different classification techniques for IDS based on machine learning. Then, they proposed a new security framework to overcome the intrusions on the host side. A new feature selection system is built for IDS (Rashid et al., 2029) by using DNA encoding and Short Tandem Repeats (STR) positions, where the performance of this method shows that the proposed method classification results are faster than the previous IDS methods.

A new network IDS is proposed (Saikam and Ch, 2024) based on the combination of deep networks and hybrid sampling, where this system starts by reducing the noise samples and then uses Deep Convolutional Generative Adversarial Networks to increase sample size; also, a deep network model is established based on DenseNet169. A novel IDS using an artificial neural network and genetic algorithm is suggested by (Cengiz et al., 2024), where the optimal weights are generated by applying a genetic algorithm and then used for the artificial neural network. A new IDS is built based on a few-shot class-incremental learning technique, and this system can learn new attacks by using a few samples; also, the proposed method contains a feature extraction model and classifier learning (Du et al., 2024). A new IDS idea is proposed by (Subhi et al., 2024) based on Deoxyribonucleic Acid (DNA) sequences and Shortest Tandem Repeat (STR), where DNA encoding is suggested to encode network traffic into DNA sequences, then used Teiresias algorithm to extract the STR values, and finally applied Horspool algorithm to classify testing dataset based on keys and their positions. Han et al. (2023) suggested a novel IDS by using binary particle swarm-wrapped feature selection, where the proposed method can enhance achieved accuracy results by increasing the relationship between training and feature selection method. A hierarchical IDS model is proposed by (Xu et al., 2023) based on applying many deep learning methods with attention mechanism, and this led to achieving different advantages such as reducing the noise based on the Stacked Convolutional Denoising Autoencoders model, extracting spatial features by using the CNN method, and the classification results are achieved by applying SoftMax classifier. Osa et al. (2024) designed the IDS method based on a Deep Neural Network, where data imbalance is managed by using SMOTE and Random Sampling, and the evaluation is done by using a single Jupyter notebook in the Google Collaboratory environment.

3 MATERIALS AND METHODS

This paper proposes a new anomaly IDS model using RNA encoding with CSA in order to enhance the pattern matching of the new proposed model based on the CICIDS2017 dataset; the reason for choosing this dataset over other datasets, this set includes the latest network attacks. The proposed system consists of four steps, which are listed in Figure 2.

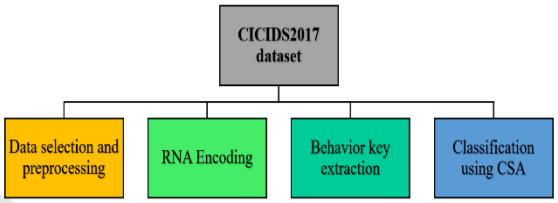

Figure 2: The proposed anomaly intrusion detection method steps

The first step of the proposed method is data selection and pre-processing, where a random sample of the CICIDS2017 dataset is chosen to include all labels that represent different types of network traffic, which provides for normal, Benign, Bot, DDoS, DoS GoldenEye, DoS Hulk, DoS Slowhttptest, DoS slowloris, FTP-Patator, Heartbleed, Infiltration, PortScan, SSH-Patator, Web Attack (Brute Force), Web Attack (SQL Injection), and Web Attack (XSS). The selected data is then divided into the data used in training the model and the data used for testing.

The second step of the proposed method is RNA encoding, where all record attribute values convert into RNA characters; this is done by understanding the dataset features and the possible values for each attribute. Where this data set has 83 features and all possible attribute values are numerical values, which means the possible values range from 0 to 9 and floating points. In order to apply RNA encoding, each value is converted to two RNA characters, where these two characters can handle all possible values. An example of encoding 0 can be converted to “CT” and so on for other possible values.

The third step of the suggested method is behaviour key extraction. This is done by using the Cuckoo Search Algorithm (CSA) to automatically extract the behaviour keys. This step involves finding the best patterns in the RNA sequences that would assist in distinguishing the traffic that is normal from

that which is an attack. CSA is iteratively applied to the current procedure to improve the selection of the patterns for classification and to maintain only the best ones.

The final step is classification using CSA. The Cuckoo Search Algorithm is used subsequently for the behaviour keys extraction in order to perform pattern matching. CSA generates several patterns that are employed in the identification of whether or not the testing records constitute a threat. This means that if any of the extracted keys are present in the RNA sequence of a record, it is considered a benign record. If it does not, then it is categorised as a threat. CSA is adapted to our IDS framework to enhance the process of pattern matching. CSA is a nature-inspired optimization algorithm which is based on the brood parasitism of some species of cuckoos. In this behaviour, cuckoos lay their eggs in the nests of other birds of different species with the aim of replacing the genuine eggs of the host birds. Any strange egg present in the nest is removed, or the host bird stops attending the nest. This process is similar to how CSA investigates and optimizes potential solutions, which are improved until the optimal solution is discovered. Cuckoo Search Algorithm used for pattern matching based on the following stages:

- Initializing P|otential Patterns: each pattern is a potential RNA sequence or behaviour key that may be used for the differentiation of benign and malicious traffic. The algorithm starts with a population of candidate solutions, or 'patterns,' of which each is a key extracted from the RNA-encoded data set. These initial patterns are generated randomly and are spread over a very large area of possibility of the RNA sequences.
- Fitness Evaluation and Optimization: the following step is to measure the fitness of each pattern that has been developed. In our IDS, the fitness of a pattern is defined by the ability of a pattern to classify traffic into either benign traffic or malicious traffic categories. The fitness function includes the detection rate (DR), false alarm rate (FAR), and the total accuracy of classification. The patterns that exhibit higher accuracy and DR are preferred, while the patterns that lead to higher FAR are penalized.
- Levy Flights and Exploration: CSA has the capability of performing Levy flights, and this has been considered one of the strengths of this algorithm as it helps to perform the global search. Levy flights consist of making a random walk or large jumps across the solution space and thus avoid being trapped in local optima. They are crucial in this global exploration to help identify such patterns that are otherwise not easily discernible in the regular search algorithms.
- Replacement and Elitism: in each iteration, the current patterns which are not very effective are substituted by new patterns that are based on the Levy flight process. It uses both random search and elitism in order to enable the algorithm to search for the solutions while at the same time exploiting the promising solutions. This aids in ensuring that only the best of patterns (nests) are retained and improved on all the time. During the training of the gradient boosting, the algorithm finds the best set of patterns which yields the best performance.

Finally, when the set of behaviour keys is optimized for the final time using the Cuckoo Search Algorithm, the keys are used to classify the network traffic. These optimized patterns are employed to search the RNA sequences available in the testing dataset. If any of the extracted and optimized keys is present in a record, then such a record is considered to be benign. If there is no match with the key, then the record is labelled as a threat. This process is highly efficient especially due to CSA's effectiveness in the matching process.

4 RESULTS AND DISCUSSIONS

As criteria used for assessing the performance of the proposed RNA encoding approach integrated with the Cuckoo Search Algorithm, the following parameters were used: the detection rate (DR) for each attack separately, the DR for the total attack records, the false alarm rate (FAR), the accuracy of the proposed method, the time required to encode the records into RNA characters, and the needed time to classify the records either to benign or threat. The obtained DR results for each attack separately are shown in Table 1 and Figure 3.

Table 1: The detection rate results for each attack type.

Attack name	Detection Rate
Bot	86%
DDoS	93%
DoS GoldenEye	81%
DoS Hulk	92.44%
DoS Slowhttptest	96%
DoS slowloris	96.3%
FTP-Patator	95.5%
Heartbleed	87.5%
Infiltration	100%
PortScan	89.07%
SSH-Patator	99%

Brute Force	92%
Sql Injection	100%
XSS	93%

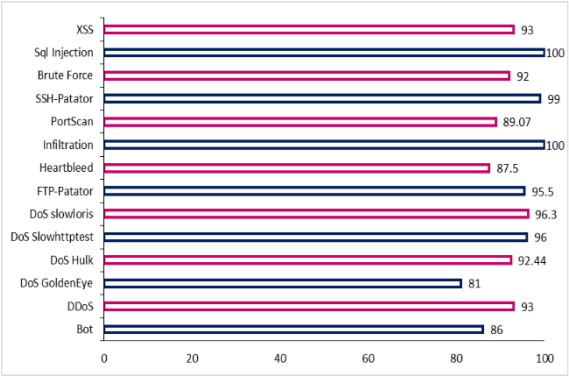

Figure 3: The detection rate results for each attack type.

As shown in Table 1 and Figure 3, the achieved detection rate for all attacks separately is good. Where the highest DR result is achieved for SQL injection attack and is equal to 100%, and the lowest DR results are equal to 81% for DoS GoldenEye attack.

The obtained DR, FAR, and accuracy for the suggested anomaly IDS method are listed in Table 2 and Figure 4.

Table 2: Obtained DR, FAR, and accuracy results.

	Results
Detection Rate	92.84%
False alarm rate	19.6%
Accuracy	92.22%

As mentioned in Table 2 and Figure 4, the obtained DR, FAR, and accuracy by applying the proposed method are equal to 92.84%, 19.6%, and 92.22%, respectively.

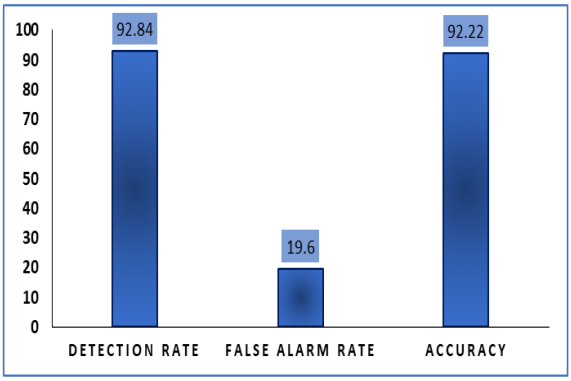

Figure 4: The detection rate results for each attack type.

Finally, the performance of the proposed method is calculated based on the time needed for RNA encoding and classification in terms of seconds for all random testing records and for one record only, and these times are listed in Table 3 and Figure 5.

Table 3: The required times for both RNA encoding and classification

	Results
RNA encoding time for all 10000 records	74.5 sec
Classification time for all 10000 records	30 sec
RNA encoding time for one record	0.00745 sec
Classification time for one record	0.003 sec

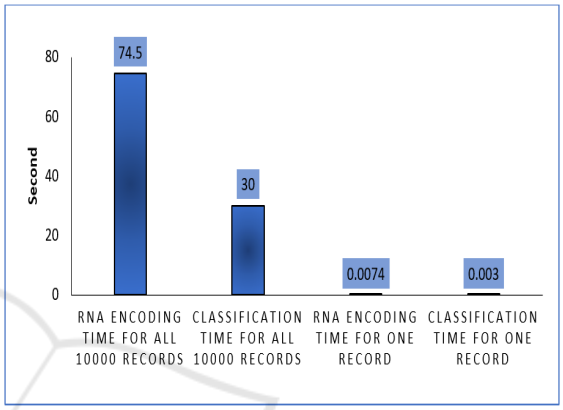

Figure 5: The detection rate results for each attack type.

Table 3 and Figure 5 show the time needed to encode one record and all testing records (10000 records) to RNA characters are equal to 0.0074 and 74.5 seconds, respectively, while the time required for applying the CSA algorithm for classification based on extracted keys for one record and all testing records are equal to 0.003 and 30 seconds respectively.

5 CONCLUSIONS

In this paper, an IDS that employs RNA encoding and CSA for pattern-matching optimization was demonstrated as an anomaly-based IDS. Through the incorporation of CSA into the pattern-matching phase, the detection rate was boosted, the false alarms were minimized and the classification ability of the system was also enhanced. CSA enabled the flexibility in the ‘‘behavior keys’’ that only the optimal patterns would be used in classification. The results clearly show that the proposed method achieves high accuracy and real-time detection rate for intrusion detection. Future work could be performed by integrating CSA with other optimization algorithms or by trying different

encoding schemes. Moreover, the system could be expanded to cover more intricate network topology which will enhance the overall security of the system in cyber security.

ACKNOWLEDGEMENTS

If any, should be placed before the references section without numbering.

REFERENCES

Abdulboriy, A. J., Shin, S. (2024). An Incremental Majority Voting Approach for Intrusion Detection System Based on Machine Learning. *in IEEE Access*, vol. 12, pp. 18972-18986.

Azam, Z., Islam, M. M., Huda, M. N. (2023). Comparative Analysis of Intrusion Detection Systems and Machine Learning-Based Model Analysis Through Decision Tree. *in IEEE Access*, vol. 11, pp. 80348-80391.

Bilot, T., Madhoun, N. E., Agha, K. A., Zouaoui, A. (2023). Graph Neural Networks for Intrusion Detection: A Survey. *in IEEE Access*, vol. 11, pp. 49114-49139.

Cengiz, K., Lipsa, S., Dash, R. K., Ivković, N., Konecki, M. (2024). A Novel Intrusion Detection System Based on Artificial Neural Network and Genetic Algorithm With a New Dimensionality Reduction Technique for UAV Communication. *in IEEE Access*, vol. 12, pp. 4925-4937.

Du, L., Gu, Z., Wang, Y., Wang, L., Jia, Y. (2024). A Few-Shot Class-Incremental Learning Method for Network Intrusion Detection. *in IEEE Transactions on Network and Service Management*, vol. 21, no. 2, pp. 2389-2401.

Eljialy, A. E. M., Uddin, M. Y., Ahmad, S. (2024). Novel Framework for an Intrusion Detection System Using Multiple Feature Selection Methods Based on Deep Learning. *Tsinghua Science and Technology*, Vol. 29 Issue. 4. Page: 948 - 958.

Fouad, N., Hameed, S. M. (2022). Genetic Algorithm based Clustering for Intrusion Detection *Iraqi. Journal of Science*, 58(2B): 929-938.

Gupta, A., Kalra, M. (2020). Intrusion Detection and Prevention system using Cuckoo search algorithm with ANN in Cloud Computing. *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Waknaghat, India, pp. 66-72.

Han, Y., Wang, Y., Cao, Y., Geng, Z., Zhu, Q. (2023). A Novel Wrapped Feature Selection Framework for Developing Power System Intrusion Detection Based on Machine Learning Methods. *in IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 11, pp. 7066-7076.

Jayalatchumy, D., Ramalingam, R., Balakrishnan, A., Safran, M., Alfarhood, S. (2024). Improved Crow Search-Based Feature Selection and Ensemble Learning for IoT Intrusion Detection. *in IEEE Access*, vol. 12, pp. 33218-33235.

Li, S., Cao, Y., Hadi, H. J., Hao, F., Hussain, F. B., Chen, L. (2024). ECF-IDS: An Enhanced Cuckoo Filter-Based Intrusion Detection System for In-Vehicle Network. *in IEEE Transactions on Network and Service Management*, vol. 21, no. 4, pp. 3846-3860.

Li, Z., Wang, P., Wang, Z. (2024). FlowGANAnomaly: Flow-Based Anomaly Network Intrusion Detection with Adversarial Learning. in Chinese. *Journal of Electronics*, vol. 33, no. 1, pp. 58-71.

Nallakaruppan, M. K., Somayaji, S. R. K., Fuladi, S., Benedetto, F., Ulaganathan, S. K., Yenduri, G. (2024). Enhancing Security of Host-Based Intrusion Detection Systems for the Internet of Things. *in IEEE Access*, vol. 12, pp. 31788-31797.

Osa, E., Orukpe, P. E., Iruansi, U. (2024). Design and implementation of a deep neural network approach for intrusion detection systems. *e-Prime Advances in Electrical Engineering, Electronics and Energy*, Volume 7.

Pirozmand, P., Ghafary, M. A., Siadat, S., Ren, J. (2020). Intrusion detection into cloud-fog-based iot networks using game theory. *Wireless Communications and Mobile Computing*, 1-9.

Rashid, O. F., Othman, Z. A., Zainudin, S. (2019). Features Selection for Intrusion Detection System Based on DNA Encoding. *In: Piuri V., Balas V., Borah S., Syed Ahmad S. (eds) Intelligent and Interactive Computing. Lecture Notes in Networks and Systems*, vol 67. Springer, Singapore.

Saikam, J., Ch, K. (2024). EESNN: Hybrid Deep Learning Empowered Spatial–Temporal Features for Network Intrusion Detection System. *in IEEE Access*, vol. 12, pp. 15930-15945.

Shivhare, A., Singh, V. K., Kumar, M. (2020). Anticomplementary triangles for efficient coverage in sensor network-based IoT. *IEEE Systems Journal*, 14(4), 4854-4863.

Singh, V. K., Singh, C., Raza, H. (2022). Event classification and intensity discrimination for forest fire inference with IoT. *IEEE Sensors Journal*, 22(9), 8869-8880.

Subhi, M. A., Rashid, O. F., Abdulsahib, S. A., Hussein, M. K., Mohammed, S. M. (2024). Anomaly Intrusion Detection Method based on RNA Encoding and ResNet50 Model. *Mesopotamian Journal of CyberSecurity*, 4 (2), pp. 120 - 128.

Tripathi, G., Singh, V. K., Sharma, V., Vinodbhai, M. V. (2024). Weighted Feature Selection for Machine Learning Based Accurate *Intrusion Detection in Communication Networks. in IEEE Access*, vol. 12, pp. 20973-20982.

Xu, H., Sun, L., Fan, G., Li, W., Kuang, G. (023). A Hierarchical Intrusion Detection Model Combining Multiple Deep Learning Models With Attention Mechanism. *in IEEE Access*, vol. 11, pp. 66212-66226.

Yang, X., Deb, S. (2010). Engineering Optimisation by Cuckoo Search. *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 1, no. 4.