Optimal Machine Learning Based Intrusion Detection System in Wireless Sensor Networks for Surveillance Applications

Sibi Amaran[1,*], Ramalingam Madhan Mohan[1] and Rethnaraj Jebakumar[2]

[1]*Department of Computer Science and Engineering, Annamalai University, Chidambaram, India*
[2]*Department of CSE, SRM Institute of Science and Technology, Kattankulathur, Chennai, India*
E-mail: sibi.amaran@gmail.com; madhanmohan_mithu@yahoo.com; jebakumr@srmist.edu.in
[*]*Corresponding Author*

Received 24 August 2021; Accepted 29 September 2021;
Publication 15 November 2022

Abstract

Security is considered as a major design issue in wireless sensor network (WSN) and can be solved by the use of intrusion detection systems (IDS). In this view, this paper devises a new k-means clustering with optimal support vector (KM-OSVM) based IDS for WSN. The KM-OSVM model incorporates preprocessing, clustering, classification, and parameter tuning. Primarily, data preprocessing and K-means clustering technique are applied to group the data instances into a set of clusters. Besides, SVM based classification technique is employed to allot class labels, and the parameters in SVM are optimally adjusted by the use of crow search optimization (CSO) algorithm, shows the novelty of the work. The experimental outcome of the KM-OSVM model is examined using UNSW-NB15 and CICIDS2017

Journal of Mobile Multimedia, Vol. 19_2, 437–450.
doi: 10.13052/jmm1550-4646.1924

datasets. The obtained outcomes demonstrated that the KM-OSVM model ensured better performance with the maximum accuracy of 95.12% and 98.98% respectively. Therefore, the KM-OSVM model can be employed as an effective tool to achieve security in the resource constrained WSN.

Keywords: Intrusions, security, WSN, machine learning, K-means clustering, parameter tuning.

1 Introduction

Generally, wireless sensor networks (WSNs) are fault-tolerant, distributed, infrastructure-less, scalable, and dynamic by their nature [1, 2]. The WSNs are exposed to many kinds of security threats which is reduce an entire efficiency of this network. In the key management, authentication protocols and secured routing protocol give secured broadcast as lack of reliability transfer of messages. Specifically, these methods are protecting the network from outside attacks then depict failure besides inside attack. It is a purpose for providing secret, authentication, and integrity data [3–6]. From the outer attack, if the intruder attempts to obtain allow the data, this technique is hidden confidential data. From the inside attack, a sensor node which is several sensor networks begins carrying out maliciously with no attempts that exist able to allowing data present from received messages. [7] describe many feasible attacks on WSNs as DoS, routing attack, Sybil attack, and so on.

An IDS is a security system utilized for detecting abnormal performance from a network. It can be supposed that IDSs are "not fit" to secure WSNs. It is true for any extended as IDS methods are generally computation costly. However, when it is regarded as WSN that efforts to track the progress of enemies, these networks are given the most helpful data to make their approach for beating the enemy from that region. Also, it will be a quick modification from the technology and maintaining from mind the future viewpoints; the abilities of sensor nodes are increasing quickly. In the sensor nodes are further memory and survival time specifically utilized to transmit multimedia data and underwater applications. Because of new development from the sensor technologies, this network is made visible and will be utilized by day-to-day life. Therefore, it is a necessity of secure WSN which makes sure secured communication and certain delivery of packets from network. IDS-based models are most effective for detecting abnormal performance of sensor nodes if it is reason DoS attacks, perform as Sybil nodes, and otherwise act some other malicious action.

In recent years, several models are presented for designing IDS for WSN. [8] established an evolutionary model for extracting ID rules. For extracting diverse rules and controlling the count of rulesets, rules are verified and removed based on the distance among rules from the similar kind of ruleset and rules from several kinds of rulesets. [9] presented an advanced hybrid IDS (AHIDS) which automatic detect WSNs attacks. The AHIDS employs the cluster structure and improved LEACH protocol for reducing an energy utilization of sensor nodes. An AHIDS utilizes anomaly detections and misuse detections dependent upon fuzzy rule set and MLPNN. Because of the benefits of negative selection algorithm (NSA) from the classifier domain, [10] introduced a WSN-NSA ID method dependent on enhanced V-detector technique for WSN. The V-detector technique is altered by changing detector generation rule and optimize detector, and PCA is utilized for reducing detection feature. [11] established an ID method dependent upon fuzzy association rule that utilize fuzzy association rule for constructing classifications, and utilizes a few equivalent metrics for evaluating the consistency of some new instances with several rule sets. The class equivalent to an optimal matching ruleset is stated as label of instance.

[12] concentration on detected a different kind of anomaly from WSN that appears concurrently from the group of neighbouring nodes and takes to an essential time period. By presented distributed segments based recursive kernel density estimate, the global probabilities density functions are track and their variance among all 2-time periods is always measured for decision making. [13] effort on a new model to handle information from segments-based approach. A group of neighboring data segment as arbitrary variables, it can be defined individuals performing abnormally with exploit its spatial predictability and, simulated with spatial analyses, especially examines for implementing a prediction difference detector from the WSN. [14] presented a metric utilizing an FL model dependent upon Sugeno fuzzy inferences method to evaluate the number of realism of the state-of-art IDS datasets. According to the presented metric outcomes, it can be intended and generate a synthetically realistic next generation IDS dataset, and an initial analyses is shown for assisting from designs of future IDSs.

This paper devises a new k-means clustering with optimal support vector (KM-OSVM) based IDS for WSN. The KM-OSVM model incorporates preprocessing, clustering, classification, and parameter tuning. Primarily, data preprocessing and K-means clustering technique are applied to group the data instances into a set of clusters. Besides, SVM based classification technique is employed to allot class labels, and the parameters in SVM

are optimally adjusted by the use of the crow search optimization (CSO) algorithm. The experimental outcome of the KM-OSVM method is examined using UNSW-NB15 and CICIDS2017 dataset.

The rest of the paper is organized as follows. Section 2 discusses the proposed model and Section 3 provides the experimental validation of the proposed model. Lastly, Section 4 concludes the paper.

2 The Proposed KM-OSVM Model

The workflow involved in the presented KM-OSVM model is illustrated in Figure 1. The figure demonstrates that the input networking data is initially pre-processed. Then, K-means clustering technique is utilized to cluster the networking data. Afterward, the SVM model is employed for the classification of clustered data. Lastly, the CSO algorithm is employed to determine the optimum parameter values of SVM.

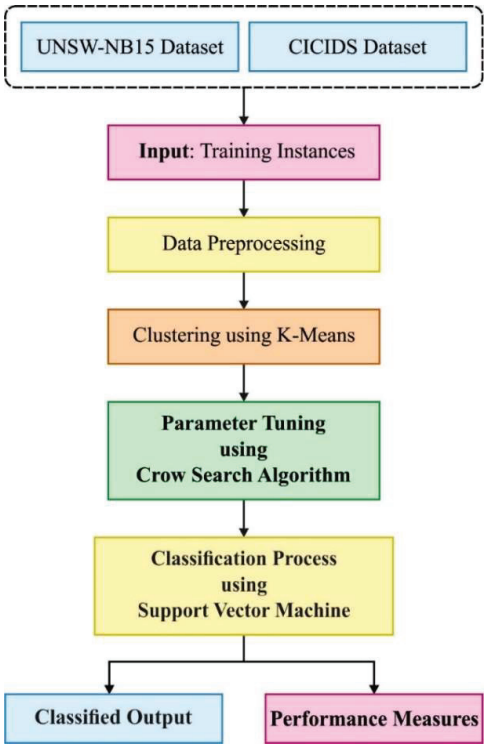


Figure 1 Working process of KM-OSVM model.

2.1 Data Pre-processing

The steps for data pre-processing are as: Data transfer, recognition method requires every input record from the layout of the vectors of real-numbers. So, the symbolic features from data set must be converted to numeric values. In the numerical normalization, a value of typical word is normalization from 0 and 1 interval, and the value X'_{ij} is normalization as:

$$X''_{ij} = \frac{X'_{ij} - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Where, $X_{min} = \min\{X'_{ij}\}$ and $X_{max} = \{X'_{ij}\}$.

2.2 K-means Clustering Technique

In K-means is most simply unsupervised learning techniques to solve the clustering problem dependent upon an easy iterative model to find a local better solution [15]. In separated clustering, an essential purpose of k-means technique for determining the more appropriate center of all clusters. It is assumed that the amount of clusters k is identified previously for resolving the clustering problem. Important steps contained from k-means clustering technique are:

1. Arbitrarily initialized k cluster center dependent upon the data ranges of N data objects.
2. Allocate all objects to a collection that is the neighboring center.
3. Upgrade the positions of all centers by computing the mean value of an objects allocated.
4. Repeating steps (2) and (3) still the maximal count of iterations are attained otherwise still the cluster centers no extensive transfer.

In K-means is one of the clustering technique as is most flexible, easy, direct, simply applied with quick implementation, measurable and effectual from a huge information collection. Though k-means has been a widely helpful clustering technique, it undergoes many disadvantages. The count of clusters k should be identified in progress. Therefore, inappropriately the technique is prone for obtaining stuck from local minimal (correspondingly from local maximum and saddle points) solutions. The performance of k-means technique is severely based on selective or arbitrary first centroids. Data clustering is not appropriate for clustering with various methods and densities.

2.3 Data Classification Using SVM Model

SVM is a margin based classifier model in that a better hyperplane is separating many classes basically feasible succeeding the rule of structural risk minimize. These allow SVM to take a great generalization ability and flexibility for over-fitting problems [16]. Also, the SVM manages non-linear classifier problem by electing kernel function for mapping the original feature space to e few higher dimensionality feature space in which samples are linearly separable. Also, SVM is to act new detection. The hyperplane of the SVM is given in Figure 2.

Assume it has a trained dataset $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, yn)\}$, where $x_i \in \mathbb{R}^p$, $i = 1, 2, \ldots, n$ is a feature vectors and $y_i \in \{-1, +1\}$, $i = 1, 2, \ldots, n$ is the equivalent binary responses. The SVM classifier method is certainly a quadratic optimization problem is:

$$\min_{\alpha} \frac{1}{2} \sum_{i}^{N} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^{n} \alpha_i \; s.t. \sum_{i=1}^{n} \alpha_i y_i = 0, \; 0 \leq \alpha_i \leq C, \tag{2}$$

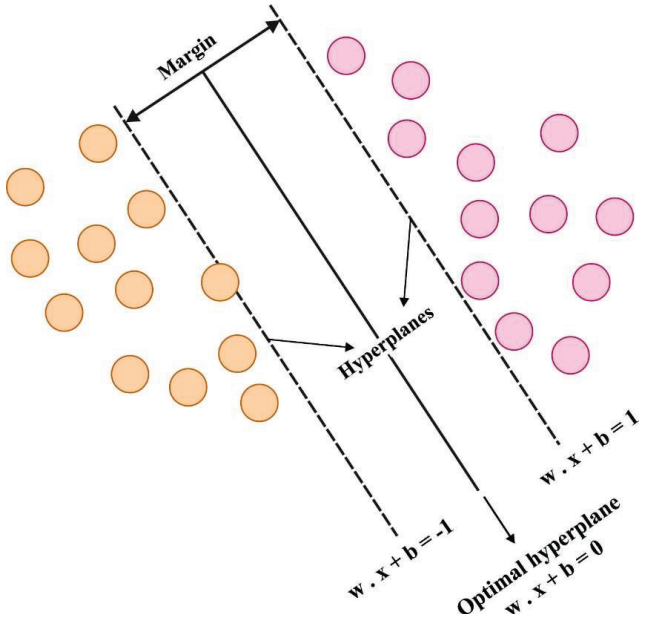


Figure 2 Hyperplanes of SVM.

where $\alpha_i (1 \leq i \leq N)$ refers the Lagrange multiplier equivalent to an instance x_i, $K(\cdot, \cdot)$ refers the kernel functions, and C refers the penalty parameters define the trade-off among the maximized of classifier margin and the minimized of misclassified error. With resolving the optimized problem from Equation (2), the decision functions are attained as:

$$D(x) = sign \left[\sum_{i=1}^{N} \alpha_i^* y_i K(x_i, x) + b^* \right]. \tag{3}$$

Lastly, to an unclassified test sample, it is classified as normal to D(x) $= +1$ or intrusion to D(x) $= -1$.

2.4 Parameter Tuning Using CSO Algorithm

In order to adjust the parameters involved in the SVM model, CSO algorithm is applied to it. The crows are regarded as the smartest creature between birds. It contains a huge brain related to its body size. In line with, a brain to body connection, their brain is slightly lesser than humans. A verifies to the smartness of crows are many several. It is illustrated self-aware from the mirror testing and takes the ability to generate tools. A crow was memorizing faces and alert one another as existence of threats [17]. Also, it can employ tools; share data from difficult methods, and remembers their confidential feed places. It's monitoring several birds, tracking where the birds retained their feed confidentially and stealing if the bird left their place. If the crow prepared robbery, it is the role it safe, i.e., altering hidden spots for abstaining in being a victim later. In fact, it employs their skill of thief for speculating the activity of thief and is decided the more secured manner for defending their feed in theft. The principles of CSA are:

- The crow exists as the group.
- The crows retain from mind the place of their confidential places.
- The crows following one another to do stealing.
- The crows secured their hideouts in being taken by chance.

Most probably it is a N-dimension environment including many crows. An entire crows are C and the position of crow u at time (iterations) iter from the search space (SS) is defined by a vector,

$$V^{u,itera}(p = 1, 2, \ldots, C;\ \text{itera} = 1, 2, \ldots, \text{itera}_{\max}) \tag{4}$$

Where $V^{u,itera} = [V_1^{u,itera}, V_2^{u,itera}, \ldots, V_c^{u,itera}]$ and $\text{iter}_{\max}$ implies the iteration with maximal count. All the crows are a memory that the position

of its confidential place is stored. In the iteration, a position of confidential position of the crow u is signified as $s^{u,itera}$. It is an optimal position that crow u attained previously. During the memory of all crows, the position of its optimal knowledge is saved. The crows begin search to for optimal sources from the surroundings. Assume that at iteration, crow v requires to carry their confidential location, $s^{v,itera}$. In this iteration, crow u definite for tracing crow v to the confidential place of crow v. During this stage, 2 actions might occur.

Event 1: The crow v is no purpose that crow u is tracked. So, crow u will attain for the confidential place of crow v. During this stage, a novel position of the crows u is established as,

$$V^{u,itera+1} = V^{u,itera} + k_j \times fll^{u,itera} \times (s^{v,itera} - V^{u,itera}) \tag{5}$$

Where, k_{j} refers the arbitrary numbers through uniform distribution amongst [0, 1] and $fll^{u,itera}$ denotes the flying length of a crow u at iteration. Minimal value of fll outcomes from the local search and maximal values affect global search.

Event 2: A crow v distinguishes as crow u is tracked. Accordingly, for defending their confidential location in stealing, crow v is deceived crow u through traveling to another place of SS. Then, event 1 and 2 are explained as,

$$V^{u,itera+1} = \begin{cases} V^{u,itera} + k_j \times fll^{u,itera} \times (s^{v,itera} - V^{u,itera}) k_j \\ \qquad \geq AWP^{v,itera} \\ \textit{a random location otherwise} \end{cases} \tag{6}$$

Where $AWP^{v,itera}$ represents the probabilities of attentiveness of crow v at iterations.

3 Experimental Validation

The experimental evaluation of the KM-OSVM model is performed against UNSW-NB15 dataset [18], which has a set of 56000 samples under the Normal class and 119,341 instances under the Attack class. Then, the CICIDS2017 dataset [19] has a set of 2,273,097 instances under benign class and 557,646 instances under Attack class. Besides, we have used 10-fold cross validation technique to split the dataset into training and testing parts. A comparative IDS results analysis of the KM-OSVM model with existing methods takes place in Table 1 and Figure 3. The set of measures used to

Table 1 Performances of existing with proposed method on UNSW-NB15 dataset

Method	Accuracy (%)	DR (%)	FAR (%)
KM-OSVM	95.12	99.71	3.22
OSVM	94.10	98.67	4.12
GAA-ADS	92.80	91.30	5.10
Fuzzy SOM	90.61	71.81	9.42
TSDL	89.71	92.46	12.76
MSCNN	91.40	92.30	15.50
KDE-CL	91.01	99.69	27.48
NB_SVM	93.75	94.73	7.33
NB_SVM2	93.35	95.27	8.78

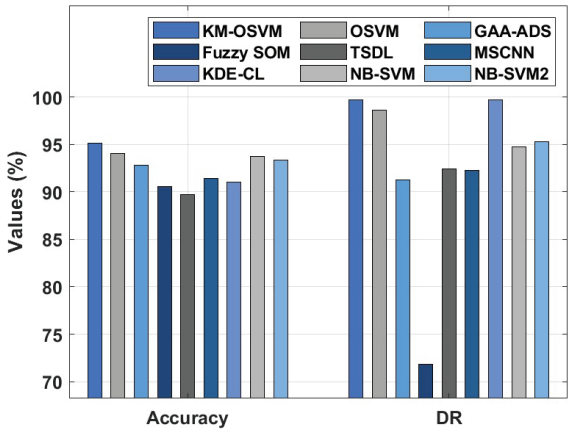


Figure 3 Result analysis of KM-OSVM model on UNSW-NB15 dataset.

examine the outcomes are accuracy, detection rate (DR), and false acceptance rate (FAR). From the resultant values, it is observable that the TSDL model has accomplished ineffective performance with the accuracy of 89.71%, DR of 92.46%, and FAR of 12.76%. Next, the fuzzy SOM model has obtained certainly raised outcomes with the accuracy of 90.61%, DR of 71.81%, and FAR of 9.42%.

Eventually, the KDE-CL model has depicted slightly enhanced performance with an accuracy of 91.01%, DR of 99.69%, and FAR of 27.48%. Meanwhile, the MSCNN model has demonstrated moderate results with an accuracy of 91.4%, DR of 92.3%, and FAR of 15.2%. Furthermore, the GAA-ADS model has showcased somewhat manageable results with an accuracy of 92.8%, DR of 91.3%, and FAR of 5.1%. In the meantime, the NB_SVM2, NB_SVM, and OSVM models have exhibited closer performance with the

Table 2 Performances of existing with proposed method on CICIDS2017 dataset

Method	Accuracy (%)	DR (%)	FAR (%)
KM-OSVM	98.98	99.68	1.23
OSVM	98.93	99.61	1.36
DBN	98.24	99.00	2.10
AE-QDA	94.20	96.40	6.30
Jrip	94.47	93.40	4.47
DNN	92.92	92.38	3.24
BRS	97.96	96.38	1.42
NB_SVM	98.92	99.46	3.00
NB_SVM2	98.84	99.51	3.55

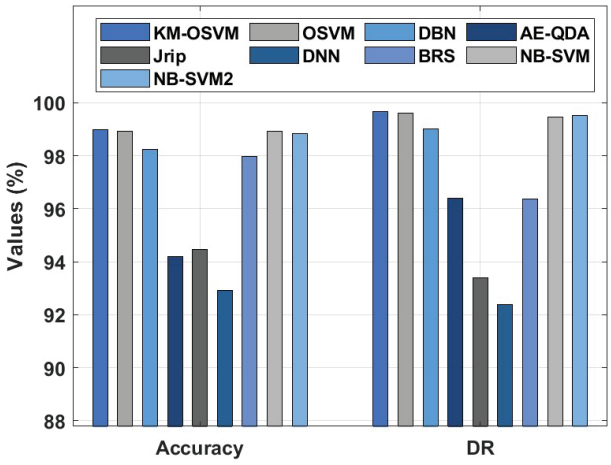


Figure 4 Result analysis of KM-OSVM model on CICIDS2017 dataset.

accuracy of 93.35%, 93.75%, and 94.1% respectively. But the KM-OSVM model has ensured supreme results with an accuracy of 95.12%, DR of 99.71%, and FAR of 3.22%.

A comparative IDS results analysis of the KM-OSVM model with existing methods takes place in Table 2 and Figure 4. The set of measures used to examine the outcomes are accuracy, DR, and FAR. From the resultant values, it is observable that the DNN model has accomplished ineffective performance with the accuracy of 92.92%, DR of 92.38%, and FAR of 3.24%. Next, the AE-QDA model has obtained certainly raised outcomes with the accuracy of 94.20%, DR of 96.40%, and FAR of 6.30%. Eventually, the

Jrip model has depicted slightly enhanced performance with an accuracy of 94.47%, DR of 93.40%, and FAR of 4.47%. Meanwhile, the BRS model has demonstrated moderate results with an accuracy of 97.96%, DR of 96.38%, and FAR of 1.42%. Furthermore, the DBN model has showcased somewhat manageable results with an accuracy of 98.24%, DR of 99%, and FAR of 2.1%. In the meantime, the NB_SVM2, NB_SVM, and OSVM models have exhibited closer performance with the accuracy of 98.84%, 98.92%, and 98.93% respectively. But the KM-OSVM model has ensured supreme results with an accuracy of 98.98%, DR of 99.68%, and FAR of 1.23%.

4 Conclusion

This paper has devised a new KM-OSVM based IDS for WSN. At the beginning stage, the input networking data is initially pre-processed. Then, K-means clustering technique is utilized to cluster the networking data. Afterwards, the SVM method is utilized for the classification of clustered data. Lastly, the CSO algorithm is employed to determine the optimum parameter values of SVM. The usage of CSO algorithm increases the detection performance of the SVM model. The experimental outcome of the KM-OSVM model is examined using UNSW-NB15 and CICIDS2017 datasets. The obtained outcomes demonstrated that the KM-OSVM model ensured better performance with the maximum accuracy of 95.12% and 98.98% respectively. In future, the detection rate and computational complexity of the KM-OSVM technique can be improved by the use of feature selection approaches.

References

[1] Arjunan, S. and Sujatha, P., 2018. Lifetime maximization of wireless sensor network using fuzzy based unequal clustering and ACO based routing hybrid protocol. Applied Intelligence, 48(8), pp. 2229–2246.

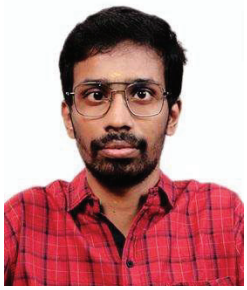
[2] Uthayakumar, J., Vengattaraman, T. and Amudhavel, J., 2017. A simple data compression algorithm for anomaly detection in Wireless Sensor Networks. International Journal of Pure and Applied Mathematics, 117(19), pp. 403–410.

[3] Arjunan, S. and Pothula, S., 2019. A survey on unequal clustering protocols in Wireless Sensor Networks. Journal of King Saud University-Computer and Information Sciences, 31(3), pp. 304–317.

[4] Uthayakumar, J., T. Vengattaraman, and J. Amudhavel. "Data compression algorithm to maximize network lifetime in wireless sensor networks." JARDCS (2017): 2156–2167, (Scopus Indexed).
[5] Arjunan, S., Pothula, S. and Ponnurangam, D., 2018. F5N-based unequal clustering protocol (F5NUCP) for wireless sensor networks. International Journal of Communication Systems, 31(17), p. e3811.
[6] Kadiravan, G., Sariga, A. and Sujatha, P., 2019, March. A Novel Energy Efficient Clustering Technique for Mobile Wireless Sensor Networks. In 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN) (pp. 1–6). IEEE.
[7] Roosta, T., Shieh, S.P., Sastry, S.: Taxonomy of Security Attacks in Sensor Networks and Countermeasures. In: Proc. of 1st IEEE Int. Conf. on System Integration and Reliability Improvements (2006).
[8] Lu, N., Sun, Y., Liu, H., Li, S. Intrusion Detection System Based on Evolving Rules for Wireless Sensor Networks. J. Sens. 2018, 2018, 1–8.
[9] Singh, R., Singh, J., Singh, R. Fuzzy Based Advanced Hybrid Intrusion Detection System to Detect Malicious Nodes in Wireless Sensor Networks. Wirel. Commun. Mob. Comput. 2017, 2017, 1–14.
[10] Sun, Z., Xu, Y., Liang, G., Zhou, Z. An Intrusion Detection Model for Wireless Sensor Networks with an Improved V-Detector Algorithm. IEEE Sens. J. 2018, 18, 1971–1984.
[11] Tajbakhsh, A., Rahmati, M., Mirzaei, A. Intrusion detection using fuzzy association rules. Appl. Soft. Comput. 2009, 9, 462–469.
[12] Xie, M., Hu, J., Guo, S., Zomaya, A.Y. Distributed Segment-Based Anomaly Detection with Kullback–Leibler Divergence in Wireless Sensor Networks. IEEE Trans. Inf. Forensic Secur. 2017, 12, 101–110.
[13] Xie, M., Hu, J., Guo, S. Segment-based anomaly detection with approximated sample covariance matrix in wireless sensor networks. IEEE Trans. Parallel Distrib. Syst. 2015, 26, 574–583.
[14] Haider, W., Hu, J., Slay, J., Turnbull, B.P., Xie, Y. Generating realistic intrusion detection system dataset based on fuzzy qualitative modeling. J. Netw. Comput. Appl. 2017, 87, 185–192.
[15] Chen, J., Qi, X., Chen, L., Chen, F. and Cheng, G., 2020. Quantum-inspired ant lion optimized hybrid k-means for cluster analysis and intrusion detection. *Knowledge-Based Systems*, *203*, p. 106167.
[16] Gu, J. and Lu, S., 2021. An effective intrusion detection approach using SVM with naïve Bayes feature embedding. *Computers & Security*, *103*, p. 102158.

[17] Manimurugan, S., Majdi, A.Q., Mohmmed, M., Narmatha, C. and Varatharajan, R., 2020. Intrusion detection in networks using crow search optimization algorithm with adaptive neuro-fuzzy inference system. *Microprocessors and Microsystems*, *79*, p. 103261.
[18] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[19] https://www.unb.ca/cic/datasets/ids-2017.html

Biographies



Sibi Amaran, Research Scholar in the Department of COMPUTER SCIENCE & ENGG. in Annamalai University for 4 Years and research area is Wireless Sensor Networks.

Ramalingam Madhan Mohan, working as Associate Professor in the Department of COMPUTER SCIENCE & ENGG. in Annamalai University. My Teaching and research Experience is 22 years and research area is Wireless Sensor Networks.

Rethnaraj Jebakumar, working as Associate Professor in the Department of Computing Technologies in SRM Institute of Science and Technology. My Teaching and Research Experience is 18 years and research area is Wireless Sensor Networks, Mobile Ad hoc Networks, Cloud Computing, Big Data, Data Mining and IOT.