INVESTIGATION INTO MACHINE LEARNING TECHNIQUES FOR NOVELTY DETECTION IN WIRELESS SENSOR DATA

Arul Jothi S¹, Jayasree B S², Harini S², Nivedha K², Selva Keerthana B G² and Gokul R² ¹Assistant Professor, Department of Computer Science and Engineering, PSG College Of Technology, Coimbatore (Affliated To Anna University, Chennai, India) saj.cse@psgtech.ac.in ²Student, Department of Computer Science and Engineering, PSG College Of Technology, Coimbatore (Affliated To Anna University, Chennai, India)

19z322@psgtech.ac.in

ABSTRACT

Since many years ago, anomaly detection has been utilized to locate and separate abnormal components from data. Anomalies have been found using a variety of ways. Machine Learning (ML), which is one of the increasingly important approaches, is crucial in this domain. In order to determine which model works best or how to increase accuracy by making changes to the cur-rent model so that it can adapt to different datasets, this research compares the performance of various machine learning and deep learning models for outlier detection on the IBRL (Intel Berkeley Research Lab) dataset and find which model suits the best or how to improve the accuracy by making changes in the existing model so that the model could adapt to various datasets also.

Keywords

anomaly detection, wireless sensor network, IBRL

1.INTRODUCTION

Outlier detection is a key consideration in the development and deployment of machine learning and deep learning algorithms. Models are often developed and leveraged to perform outlier detection for different organizations that rely on large datasets to function. When creating and deploying machine learning algorithms, outlier detection is a crucial factor to take into account. For many organizations whose operations depend on massive datasets, models are frequently created and used to do outlier detection. Because outliers have the potential to badly skew the overall result of an analysis and because their behavior may be precisely what is desired, it is crucial to recognize and deal with them while evaluating data. In order to understand trends and relationships between data points, machine learning algorithms learn from the data. The process of finding outliers, unexpected events, observations, or peculiar patterns that deviate from expected behavior is known as anomaly detection.

An important tool for preserving data quality is outlier detection since it allows for the removal and analysis of abnormal data and errors. It has various uses in business and aids in seeing suspicious network traffic patterns that could indicate a hack, spotting a cancerous tumor in an MRI scan, detecting fraud in credit card transactions, and defect identification in operating systems.

2. RELATED WORK

In sensor networks, anomaly detection is a well-researched topic. Numerous studies have developed numerous methods to cope with anomaly detection. These studies have condensed the elementary issue of anomaly detection into a presentation of all scenarios and methods for dealing with it.

One-class support vector machine is proposed in [\[2\]](https://www.researchgate.net/publication/321664272_Data_driven_hyperparameter_optimization_of_one-class_support_vector_machines_for_anomaly_detection_in_wireless_sensor_networks) for anomaly detection in wireless sensor networks". This study explores the use of data-driven hyperparameter optimization with OCSVM to find outliers in WSNs. Rather than having measures like area under the receiver's operating characteristic or geometric mean accuracy (G-mean), the information of each sample's of the farthest and the closest neighbors is employed to generate the objective cost (AUROC). The suggested method produced a low percentage of false alarms and a high degree of detection accuracy.

Author in [\[3\]](https://www.researchgate.net/publication/342732975_Deep_Learning_for_Anomaly_Detection_A_Review) presents a thorough taxonomy of the research on deep anomaly detection, spanning improvements in three supreme categories and eleven intricate categories of the approaches. They went over core assumptions, criterium of choice, benefits, and drawbacks before talking about how they approach the aforementioned problems. They also went over a number of potential prospects for the future and fresh approaches to dealing with the problems.

In paper [\[6\]](https://www.researchgate.net/publication/326134783_A_distributed_anomaly_detection_model_for_wireless_sensor_networks_based_on_the_one-class_principal_component_classifier) one-class principal component classifier was proposed. For WSNs, a distributed anomaly detection approach was put out in this work based on clusters. To improve the efficiency and effectiveness of detection, the model makes use of the spatial correlation of sensing data in a small area (i.e., cluster). The suggested model seeks to effectively utilize the limited resources of sensors in order to overcome the shortcomings of existing detection algorithms. By reacting to dynamic changes in the sensed environment, it also seeks to increase effectiveness.

Author in [\[7\]](https://www.researchgate.net/publication/326846706_Unsupervised_anomaly_detection_for_unlabelled_Wireless_Sensor_Networks_Data) proposed a method using OCSVM classifier. In this study, the efficacy performance in terms of detection accuracy, false positive rate, detection rate, and false negative rate is evaluated using an unsupervised classification model OCSVM to detect anomalies. One PCA version known as CCIPCA(Candid-Covariance Free Incremental Principal Component Analysis) dimension reduction is utilized to lessen the computational complexity of the CESVM (Centered Hyper-ellipsoidal Support Vector Machine) classifier which in turn lessens the resource limitation experienced in sensor nodes. According to the results, the performance of the CESVM and CCIPCA anomaly detection schemes is comparable in the majority of the WSNs dataset.

Incremental Principal Component Analysis and Support Vector Machine (OCSVM) was published in [\[8\]](file:///C:/Users/HP/Downloads/Lightweight%20Anomaly%20Detection%20Scheme%20Using%20Incremental%20Principal%20Component%20Analysis%20and%20Support%20Vector%20Machine), This paper concentrates on designing a lightweight anomaly detection scheme to yield data collection which is reliable while ingesting less energy using one-class learning schemes and dimension reduction concepts. This paper's goal is to design and develop a lightweight anomaly detection by efficiently detecting anomalous data while consuming energy efficaciously. One-class support vector machine (OCSVM) is used as an anomaly detection algorithm owing to its highlights in classifying unlabeled data while the multivariate data can be detected by the hyper-ellipsoid variance.

In paper [\[9\]](https://eudl.eu/doi/10.4108/eai.13-7-2018.161368), ML Based Hybrid Model was proposed for Detection of Faults In Wireless Sensor Data. In this research, a hybrid approach for knowledge discovery based on machine learning is proposed. For the study, the Intel Berkeley Research Lab (IBRL) dataset is taken into account. Three models were used to identify anomalies in the dataset using the spatialtemporal connection, including:

1) Minimum Covariant Determinant (MCD)Isolation Forests (IF)

2) Isolation Forests (IF)

3) Histogram Based Outlier Score (HBOS).

Additionally, the electrical configuration of the Wireless Sensor Networks component parts has been used to identify errors in the dataset's outliers. The outcomes demonstrate that the suggested hybrid model with IF performed better than average with a respectable precision value. Among the sensors which were deployed in the IBRL dataset, the experiment was also able to identify the problematic sensors.

Author in [\[11\]](https://www.researchgate.net/publication/348849888_Anomaly_Detection_in_Wireless_Sensor_Network_using_Inverse_Weight_Clustering_and_C50_Decision_Tree) proposed a model based on Inverse Weight Clustering and C5. 0 Decision Tree. In this paper, Inverse Weighted Clustering (IWC) combined with C5. 0 decision tree algorithms have been used as a model where IWC is used to cluster sample data into groups, label them and then used a C5. 0 classifier to train and test the model to distinguish between the anomalous and normal activities in a wireless sensor network. The experiment was carried out on three different datasets (University of North Carolina Greensboro (UNCG), Intel Berkeley Research Lab (IBRL) and Bharatpur Airport dataset. The results show that IWC+C5. 0 is the efficient technique for detecting anomalies on IBRL, with a higher accuracy rate.

Author in [\[15\]](https://www.researchgate.net/publication/329534739_Outlier_Detection_in_Wireless_Sensor_Networks_Using_Model_Selection-Based_Support_Vector_Data_Descriptions) proposed a method using Model Selection-Based Support Vector Data Descriptions. First, the space and time complexity of detecting outliers is reduced by using the Toeplitz matrix random feature mapping. In order to make the algorithm stable when the feature dimensions are low, a unique model selection technique is developed. Using this strategy, it is possible to choose a decision model that is approximately optimal and to prevent both overfitting and underfitting issues. This yields less time complexity and improved accuracy for outlier detection in WSNs when compared to earlier methods.

Author in [\[16\]](https://journals.sagepub.com/doi/full/10.1177/1550147716686161) proposed a method using Support Vector Data Description. This research proposes a better method for the detection of anomalies in energy-constrained wireless sensor networks based on support vector data description, which minimizes computing complexity. The primary goal was to increase computational complexity starting with the training and decision-making phases. In order to increase training speed, reduction of training data and the SMO algorithm that utilizes second-order approximation were combined. In the testing phase, a quick decision strategy for an unobserved sample was put up to quicken testing.

From the literature survey of all these papers, The detailed investigation suitable for outlier detection is focused in the next section. Since the focus of Deep Learning techniques on anomaly detection is less on wireless sensor data, we are investigating DL techniques. Apart from ML techniques, the author in that paper has discussed the DL techniques used for anomaly detection but limited work has been done on wireless sensor data. We are going to investigate DL techniques that are suitable for wireless sensor data.

3.INVESTIGATIONS ON CONVENTIONAL METHODS

Conventional methods in this paper are considered as statistical methods and machine learning methods. Statistical methods are very complex for high dimensional data and unsupervised data.

Our next focus is on Machine Learning models. Many machine learning models have been studied for detecting anomalies in wireless sensor data. Some methods include SVDD (Support Vector Data Description) and clustering techniques, IPCA and SVM etc. In SVDD, the difficulty is that the calculations are challenging in the training phase $[16]$. One of the limitations of the model that involves IPCA and SVM lightweight anomaly detection scheme [\[8\]](file:///C:/Users/HP/Downloads/Lightweight%20Anomaly%20Detection%20Scheme%20Using%20Incremental%20Principal%20Component%20Analysis%20and%20Support%20Vector%20Machine) is the capacity of the model to universally ensure the anomalous character of typical occurrences during the anomaly identification process. Also machine learning techniques are not suitable for unsupervised data and to detect anomalies in multivariate time-series data.

In real time data, the data has to be trained already in machine learning techniques whereas in deep learning techniques, it will learn the features and train itself. Deep learning techniques have efficient computational cost and processing cost when compared to machine learning techniques.

Since many years, Deep learning methods have been used for many purposes but this paper focuses on how deep learning models are more efficient in detecting outliers compared to machine learning models. Deep learning methods enhance a generic feature learning objective function to understand how data samples are represented. Despite not being designed with anomaly detection in mind, trained representations can nonetheless help by being forced to acknowledge certain significant underlying patterns in data.

3.1. Autoencoders

Autoencoders are unsupervised machine learning algorithms that use artificial neural networks. The main idea behind autoencoders is to discover a low dimensional feature space that is capable of reconstructing the given data instances. Autoencoders are mainly composed of 3 components. The encoder is the one that compresses the given data instances and produces the low dimensional feature space. The bottleneck is the one that retains the compressed input data instances. The decoder attempts to recreate the actual data from the extracted low-dimensional feature space. It is essential to choose the autoencoder's parameters so as to reduce the reconstruction error. However the anomaly data will have a poor reconstruction error as compared to the regular data instances. Hence the anomaly score can be directly measured as the reconstruction score.

The typical neural network function can be used to represent the encoding network after passing through an activation function.

$$
k = \sigma(Wi + b) \tag{1}
$$

The decoding network has a similar formulation using a different bias and weights.

$$
i' = \sigma'(W'k + b')\tag{2}
$$

The loss function or reconstruction error is

$$
L(i, i') = ||i - i'||^2 = ||i - \sigma'(W'(\sigma(Wx + b)) + b')||^2 \tag{3}
$$

Various specialized autoencoders have been developed. Sparse autoencoders instead of reconstructing the input data instances from all the hidden nodes, the k-highest hidden nodes are chosen for reconstruction. This model has been tried on image datasets for the purpose of classifying images in [\[1\].](https://arxiv.org/abs/1312.5663?context=cs) The Denoising Autoencoder reconstructs the input data from input data instances by introducing corruptions into the input data. Classification based on this model is performed in [\[10\]](https://www.jmlr.org/papers/volume11/vincent10a/vincent10a.pdf) for image and audio datasets such as tzanetakis and MNIST. The paper [\[12\]](https://ieeexplore.ieee.org/document/8422402) uses autoencoders which are executed on each sensor. The autoencoder's input and the output is provided as training data to the cloud and performs anomaly detection locally. The cloud learns the efficient parameters (weights and bias) from the training data fed by the autoencoder.

Infrequent regularities and the presence of outliers in the training set may cause the learned feature representations to be skewed, which is one of the drawbacks of autoencoders. Since autoencoders is usually applied in feature extraction and classification, the parameters and type of autoencoder must be chosen appropriately for carrying out anomaly detection.

3.2. Generative Adversarial Network (GAN)

GANs are an efficient group of neural networks that are used for unsupervised learning. A generator and a discriminator are both present in GANs. Generator attempts to deceive the Discriminator by creating false samples of data. The Discriminator attempts to distinguish between the genuine and fraudulent data. Both Generator and Discriminator compete with each other throughout the training phase. The procedures are repeated multiple times, and each time, the Generator and Discriminator become better at their work[\[4\]](https://arxiv.org/abs/1611.02163) .

The generative model is trained to attempt to optimize the probability that the Discriminator will make a mistake while still capturing the distribution of the data. The Discriminator is created using a model that determines the probability that the sample it got originated from the training instances rather than the generator. The Discriminator is seeking to reduce its reward V (D, G) in the minimax game that the GANs are designed as, while the Generator is aiming to maximize its loss by minimizing the Discriminator's reward. The formula below can be used to explain it mathematically:

$$
V (D, G) = E_{i \sim \text{Pdata}(i)} [logD(i) + E_{j \sim P(j)} [log(1 - D(G(j)))] \tag{4}
$$

Where G represents generator, D denotes discriminator, Pdata(i) is the distribution of real data distribution, $P(i)$ is the distribution of generator, i is the sample from Pdata(i), j represents sample from $P(i)$, $D(i)$ is the Discriminator network, $G(i)$ is the Generator network.

When it comes to training GAN models for the execution, it is crucial, and there may be certain frequent difficulties. Failure of the setup and mode collapse are the two primary difficulties encountered when training a GAN model. It draws attention to some of the prevalent issues, such as those involving perspective, counting, and global organization. In paper [\[13\]](https://arxiv.org/abs/1606.03498), A variety of our innovative architectural components and training methods are employed in conjunction with the generative adversarial networks (GANs) framework. The superiority of the created images was proved by a test. This approach generates MNIST samples that are identical to real data. Author in paper[\[4\]](https://arxiv.org/abs/1611.02163) addresses the issue of mode collapse that is frequently encountered, stabilizes the training of GANs featuring complex repeating generators, and improves the variety and coverage of the generator's distribution of data.

4. EXPERIMENTAL ANALYSIS

4.1. Datasets

4.1.1. Intel Berkeley Research lab (IBRL)

The dataset carries information concerning data gathered from fifty-four sensors utilized in the Intel Berkeley research laboratory between February twenty eighth and April 5th, 2004. each thirty-one seconds, Mica2Dot sensors equipped with clapboards gathered timestamped topological data alongside measurements of the humidity, temperature, light, and voltage. information was gathered exploitation the TinyOS-based TinyDB in-network query processing technology.

Figure 1. Arrangement of Sensors in IBRL

date	time			epoch moteid temperature humidity- light- voltage-			
yyyy-mm-	hh:mm:ss:	$-int$	-int	-real	real	real	real
dd	XXXX						

Figure 2. Schema for IBRL Dataset

4.1.2. Secure Water Treatment (SWaT)

The EPIC testbed was run for 11 days, seven of which were spent in normal operation and four in attack simulations, to gather the data. gathered data from 51 sensors and actuators, including network traffic. Information with labels for typical and aberrant behaviour. The EPIC testbed was run for 11 days, seven of which were spent in normal operation and four in attack simulations, to gather the data. gathered data from 51 sensors and actuators, including network traffic. Information with labels for typical and aberrant behavior.

4.1.3. Water Distribution (WADI)

A test site for the design of safe cyber-physical systems for water distribution. The data was gathered over the course of 16 days, split into two periods of time: two days of attack scenarios and 14 days of regular operation collected information from all 123 sensors and actuators.

4.1.4. Networked Aquatic Microbial Observing System (NAMOS)

9 buoys with temperature and chlorophyll awareness sensors (fluorimeters) have been positioned in Lake Fulmor, James Reserve, in August 2006, and have been there for extra than 24 hours. Chlorophyll sensors on buoy wide variety 103 are used to degree it for 104 samples.

4.1.5. Sensorscope Lausanne Urban Canopy Experiment (LUCE)

Between July 2006 and May 2007, a project this project is collected. The instrument was engineered on a WSN of one hundred ten sensing element nodes that were put in on the EPFL field to measure necessary environmental parameters, reminiscent of the close temperature, surface temperature, and relative humidity.

The various machine learning models reminiscent of DWT+OCSVM, DWT+SOM, PCCAD, CESVM-DR were used to observe outlier in wireless sensing element networks such as NAMOS, LUCE and IBRL. The experimental outcome of that is shown within the below table.

Figure 3. Accuracy of ML models applied on sensor datasets

From the Figure 3 we could see that PCCAD has high accuracy compared to other models when applied to IBRL and LUCE dataset. The other machine learning models can be tuned with correct parameters so that it produces better results for IBRL and LUCE datasets. On the other hand, CESVM-DR gives high accuracy for NAMOS dataset. The accuracy of DWT+OCSVM and PCCAD is very low compared to other models for NAMOS dataset. The memory utilization of DWT+OCSVM and DWT+SOM is $O(n)$ and $O(mn+nd)$ severally where n is that the range of data observations, m is Low-rank estimation of the kernel Gram matrix and d is reduced dimension of the information vector. PCCAD's memory utilization is O(nd). The procedure complexness of PCCAD is O(Nd) where N is the calculation of CCIPCA. The CESVM-DR incorporates a mathematical complexity of $O(P+m2d+dn2)$ where P is the linear improvement downside calculation.

Datasets	Models	FPR
IBRL	DWT+OCSVM	1.9
	DWT+SOM	1.09
	PCCAD	0.3
	CESVM-DR	1.6
LUCE	DWT+OCSVM	1.9
	DWT+SOM	1.09
	PCCAD	0.09
	CESVM-DR	2
NAMOS	DWT+OCSVM	12.8
	DWT+SOM	0.5
	PCCAD	11.5
	CESVM-DR	0

Table 2. False Positive Rates of ML models applied on sensor datasets.

Figure 4. False Positive Rates of ML models applied on sensor datasets

From the Figure 4., DWT+OCSVM and PCCAD has higher False Positive Rate (FPR) for NAMOS dataset whereas it produces very less FPR when DWT+SOM and CESVM-DR is used. The PCCAD has very less FPR for both IBRL and LUCE datasets when compared to other models.

Since machine learning models need loads of training time and involve computational overhead, the deep learning models comparable to EGAN, VAE and UAE are wont to find anomalies in wireless detector networks such as SWaT and WADI. The precision, recall rate and F1 score are visualized in the Table. 3.

Datasets	Methods	Precision	Recall	F1
SWaT	EGAN	40.57	67.73	51
	VAE	94	78.5	85.5
	UAE	96.5	77.8	86.1
WADI	EGAN	11.33	37.84	$\sqrt{7}$
	VAE	85.3	62.1	71.8
	UAE	91. 6	64	75.

Table 3. Recall rate, F1 score, and precision of DL models applied on sensor datasets.

Figure 5. Recall rate, F1 score, and precision of DL models applied on sensor datasets

From Figure 5. it is clear that EGAN shows less precision, recall rate and F1-score for both SWaT and WADI datasets. The deep learning models UAE and VAE shows almost similar precision, recall rate and F1-score for SWaT dataset. Incase of WADI dataset UAE gives better precision than VAE.

5. PROPOSED MODEL

Figure 6. Proposed Model

In our proposed model, we have a certain number of nodes and a cluster head among the nodes. Applying deep learning on sensor nodes where training and testing is done on the sensor node, the complexity increases. So the data is trained on the cloud. The data is trained in such a way that it would be able to identify normal and abnormal data. During training phase, If a particular sensor is known to give wrong readings, The outliers are removed and then the data is forwarded to the cluster head with the hyperparameters. Then, the hyperparameters are updated which include weight, bias, learning rate, optimizer etc. with the help of tensorboard. This process is called hyperparameter tuning. While the parameters are passed, the packet size or load should be ensured that it is low to avoid overhead. The outlier detection process is performed on the cluster head. The aggregated data is preprocessed and the model is tested on the data. Variational Autoencoder model is applied on the data. The reconstruction error is calculated from the input and the output data. The filter/threshold/outlier score is fixed from the reconstruction error and it helps to classify whether a particular data is normal or an outlier. The normal data is forwarded and the outlier is reverted.

6. CONCLUSION AND FUTURE WORK

It is essential but difficult to develop an efficient outlier detection strategy for WSNs because of the constrained resources. The objective is to increase sensor connectivity by reducing energy consumption and memory use while maintaining accurate data at the base station.

We presented an anomaly detection approach using various models on various datasets.

In the future, If the threshold that we calculate from the reconstruction error is not that efficient, then we can check whether the threshold can be fixed by taking the mean or standard deviation or any distribution other than normal distribution. The existing models identified threshold based on ROC curve and Precision, Recall curve.

Another objective is instead of removing the outlier, we try to develop a method whether we could replace the outlier with some measures. We focus also on the detection ability by making some changes in our proposed approach to make the model suitable for various datasets.

REFERENCES

- [1] Alireza Makhzani & Brendan Frey(2014) "K-sparse autoencoders", arXiv:1606.03498v1,1312.5663v2
- [2] Daniel Fährmann, Naser Damer, Florian Kirchbuchner and Arjan Kuijper (2022) "Lightweight Long Short-Term Memory Variational Auto-Encoder for Multivariate Time Series Anomaly Detection in Industrial Control Systems", DOI: 10.3390/s22082886
- [3] Guansong Pang, Chunhua Shen, Longbing Cao, Anton van den Hengel (2020) "Deep Learning for Anomaly Detection: A Review",ACM Computing Surveys, DOI: 10.1145/3439950
- [4] Luke Metz, Ben Poole, David Pfau, and Jascha Sohl-Dickstein (2017) "Unrolled generative adversarial networks", arXiv:1611.02163v4
- [5] Manuel Lopez-Martin, Belen Carro, Antonio Sanchez-Esguevillas and Jaime Lloret (2017), "Conditional Variational Autoencoder for Prediction and Feature Recovery Applied to Intrusion Detection in IoT"
- [6] Murad A. Rassam, Mohd Aizaini Maarof and Anazida Zainal (2018) "A distributed anomaly detection model for wireless sensor networks based on the one-class princpal component classifier", [International Journal of Sensor Networks](https://www.researchgate.net/journal/International-Journal-of-Sensor-Networks-1748-1287) 27(3):200, DOI[:10.1504/IJSNET.2018.093126](http://dx.doi.org/10.1504/IJSNET.2018.093126)
- [7] Nurfazrina Mohd Zamry, Anazida Zainal, Murad A. Rassam (2018) "Unsupervised Anomaly Detection for Unlabelled Wireless Sensor Networks Data", International Journal of Advances in Soft Computing and its Applications, Volume: 10, No.2
- [8] Nurfazrina M. Zamry, Anazida Zainal, Murad A. Rassam, Eman H. Alkhammash, Fuad A. Ghaleb, and Faisal Saeed (2021) "Lightweight Anomaly Detection Scheme Using Incremental Principal Component Analysis and Support Vector Machine", DOI: 10.3390/s21238017.
- [9] P. Raghu Vamsi and Anjali Chahuan,"Machine Learning Based Hybrid Model for Fault Detection in Wireless Sensors Data", 05 November 2019, DOI: 10.4108/eai.13-7- 2018.161368
- [10] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre-Antoine Manzagol. "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion", Journal of Machine Learning Research 11, 3371–3408
- [11] Pramod Kumar Chaudhary, Arun Kumar Timalsina (2021) "Anomaly Detection in Wireless Sensor Network using Inverse Weight Clustering and C5.0 Decision Tree", Volume 7
- [12] Tie Luo and Sai G.Nagarajan (2018) "Distributed Anomaly Detection using Autoencoder Neural Networks in WSN for IoT", [IEEE International Conference on](https://ieeexplore.ieee.org/xpl/conhome/8411665/proceeding) [Communications \(ICC\),](https://ieeexplore.ieee.org/xpl/conhome/8411665/proceeding) DOI: [10.1109/ICC.2018.8422402](https://doi.org/10.1109/ICC.2018.8422402)
- [13] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen, Xi Chen (2016) "Improved Techniques for Training GANs", arXiv:1606.03498v1
- [14] Van Vuong Trinh; Kim Phuc Tran; Truong, Thu Huong, (2017) "Data driven hyperparameter optimization of one-class support vector machines for anomaly detection in wireless sensor networks", [International Conference on Advanced](https://ieeexplore.ieee.org/xpl/conhome/8119419/proceeding) [Technologies for Communications \(ATC\),](https://ieeexplore.ieee.org/xpl/conhome/8119419/proceeding) DOI: [10.1109/ATC.2017.8167642](https://doi.org/10.1109/ATC.2017.8167642)
- [15] Zhan Huan, Chang Wei and Guang-Hui Li (2018) "Outlier Detection in Wireless Sensor Networks Using Model Selection-Based Support Vector Data Descriptions",DOI: 10.3390/s18124328
- [16] Zhen Feng, Jingqi Fu, Dajun Du, Fuqiang Li, Sizhou Sun (2017) "A new approach of anomaly detection in wireless sensor networks using support vector data description", Volume: 13, Issue: 1, DOI: 10.1177/2F1550147716686161

Authors

Arul Jothi S

Jayasree B S

Harini S

Arul Jothi S completed her Master Degree at Kumaraguru College of Technology in 2010. She was working as Assistant Professor since 2010. Currently she is working as Assistant Professor at PSG College of Technology, Coimbatore, India. In this career she has published 7 international Journals which are Scopus indexed and 4 conference papers. She is pursuing her PhD under the research title "Improving data accuracy in Wireless Sensor Networks". She is an active professional member in ACM, CSI and IJFERP.

Jayasree B S is pursuing her final year Bachelor of engineering in the department of Computer science and engineering at PSG College of technology, Coimbatore. She is working on a project "Anomaly detection in wireless sensor networks with variational autoencoders" as a part of her academics.

Harini S is pursuing her final year Bachelor of engineering in the department of Computer science and engineering at PSG College of technology, Coimbatore. She is working on a

project "Anomaly detection in wireless sensor networks with variational autoencoders" as a part of her academics.

Nivedha K is pursuing her final year Bachelor of engineering in the department of Computer science and engineering at PSG College of technology, Coimbatore. She is working on a project "Anomaly detection in wireless sensor networks with variational autoencoders" as a part of her academics.

Nivedha K

Selva Keerthana B G

Gokul R

Selva Keerthana B G is pursuing her final year Bachelor of engineering in the department of Computer science and engineering at PSG College of technology, Coimbatore. She is working on a project "Anomaly detection in wireless sensor networks with variational autoencoders" as a part of her academics.

Gokul R is pursuing his final year Bachelor of engineering in the department of Computer science and engineering at PSG College of technology. He is working on a project "Anomaly detection in wireless sensor networks with variational autoencoders" as a part of his academics.