Hindawi
Mobile Information Systems
Volume 2023, Article ID 9767513, 1 page
https://doi.org/10.1155/2023/9767513

Retraction

Retracted: Vulnerability Assessment of Asphalt Plant through Machine Learning Techniques

Mobile Information Systems

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] A. Haider, S. Khan, A. Mohamed, S. Khan, and R. Khan, "Vulnerability Assessment of Asphalt Plant through Machine Learning Techniques," *Mobile Information Systems*, vol. 2022, Article ID 9496123, 9 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9496123, 9 pages
https://doi.org/10.1155/2022/9496123

Research Article

Vulnerability Assessment of Asphalt Plant through Machine Learning Techniques

Abid Haider,[1] **Sarmadullah Khan**,[2] **Abdullah Mohamed**,[3] **Shahbaz Khan**,[4] **and Razaullah Khan**[4]

[1]*Department of Electrical Engineering, CECOS University, 25100, Pakistan*
[2]*School of Computer Science and Informatics, De Montfort University, LE1 9BH, UK*
[3]*Research Centre, Future University in Egypt, New Cairo 11835, Egypt*
[4]*Institute of Manufacturing, Engineering Management, University of Engineering and Applied Sciences, 19060, Pakistan*

Correspondence should be addressed to Abid Haider; abid@cecos.edu.pk

Received 10 February 2022; Revised 22 April 2022; Accepted 27 April 2022; Published 14 May 2022

Academic Editor: Hafiz Tayyab Rauf

Many businesses throughout the globe have recently realized the value of Supervisory Control and Data Acquisition (SCADA) systems. Many critical infrastructures, such as electricity grids, asphalt plants, and wastewater disposals, are controlled by these systems. With the introduction of Fourth Industrial Revolution, 4IR or Industry 4.0, today's SCADA systems cannot be separated from the outside world, making them more susceptible to hostile assaults. Conventional security systems including different antivirus software and firewalls are unable to safeguard SCADA systems as they are of distinct requirements. For this, different machine learning algorithms, i.e., SVM, KNN, and random forest, are tested to cover the anomaly detection along with security protection for SCADA systems. The dataset used in this research study was made locally in an asphalt plant by using different sensor data grouped in two classes: one is natural signal values, and the other one is attack class in which different sensor values are found out of range while in operation. Amongst the above-mentioned algorithms, KNN outperformed with an accuracy rate of 89% for anomaly detection and any kind of external attack can be detected and notified to the control room for on-time actions.

1. Introduction

SCADA (Supervisory Control and Data Acquisition) systems manage and monitor industrial and essential infrastructure activities like electricity, gas, water, trash, railroads, and transportation. Those very systems for controlling critical national infrastructures were once considered secure because they had proprietary restrictions and limited connectivity. SCADA's approach is no longer immune to cyber threats because of increased connectivity to the Internet and business networks. In fact, the threat to critical infrastructure is significantly bigger than conventional computer vulnerabilities in terms of impact and scope of assault. A cyber-attack on a sewage system in Queensland resulted in the release of 800,000 gallons of raw sewage into nearby parks and rivers, killing marine life and producing odour and discoloration of the water [1, 2]. The SQL Slammer server malware recently targeted the nuclear power plant in the USA, causing a nearly five-hour outage of the nuclear plant's safety monitoring system [3]. Because all of these commodities are necessary for the smooth operation of day to day lives, their protection and security are vital as well as a national priority.

It is essential to assess the security risk of SCADA systems and build proper security solutions to defend them from assaults to fully understand how to safeguard them [4–6]. The lack of adequate modelling tools to assess the privacy of the SCADA system is a major issue in the study and innovation of intrusion detection systems for the SCADA system. A SCADA testbed allows us to build a rudimentary model of a SCADA system while also testing real-world assaults and experimenting with different security solutions.

Because of the scope and cost of building independent SCADA systems, including the possible danger and disruption of operations provided by the essential infrastructure, conducting security tests on an actual SCADA system is impracticable.

Current ICT (Information and Communication Technology) systems for utilities are based on the assumption that there is a connection between a company's network and the SCADA system's network. These networks should be designed in such a way that they can provide operating and commercial telecommunications services while also meeting a set of technical standards and features. Remote control, tele protection, operational telephone, and operating video are examples of operational services. All of them are linked to one other, either directly or indirectly in electric power systems with the technological technique for power generation. Because of their importance, these services have rigorous standards for reliability, availability, and latency.

A lot of critical infrastructures (CIs) and businesses rely on Supervisory Control and Information Acquisition systems, which are widely utilized in critical infrastructure (CIs). Oil pipelines, treatment facilities, and chemical factories are just a few examples. To keep the control network, separate from other parts of the system, including the Internet, SCADA systems have traditionally used a security concept known as the "air gap." In the actual world, true solitude is difficult to achieve. In the first place, complete isolation may result in the use of antiquated software.

The vendor's security upgrades cannot readily be applied to the program if it does not have Internet access. Second, implementing real isolation is difficult since CI is very often geographically dispersed. Furthermore, SCADA devices have made use of proprietary software, hardware, and communication protocols that have given a misleading impression of security through concealment. SCADA systems may now be integrated with the Internet and business networks because of the widespread usage of standardized communication protocols. As a result of their extensive deployment regions, dispersed operating mode, and increasing interconnectivity, SCADA systems are now vulnerable to a wide range of risks. Because of the TCP/IP stack's extensive usage, SCADA systems have adopted it. Protocols such as the Modicum Communication Bus (Modbus), Distributed Network Protocol (DNP3) [4, and IEC 60870-5-104], and TCP/IP are widely used. These protocols have been in development for over two decades and are well-known for their susceptibility to low-tech network assaults. There is no IT system that is not protected by intrusion detection systems (IDSs), using conventional IDS to identify intrusions calls and for a database of attack signatures each signature representing a distinct attack and its features. Its major drawback is the need for human analysis of vulnerabilities and threats to come up with unique signatures using this approach. A strong option for developing outlier detection algorithms about typical behaviour and adjusting on their own to deviations is machine learning (ML) technology, which can do even without being preprogramed again or given an explicit pattern to work from.

2. Literature Review

Internet connectivity makes SCADA systems more susceptible to global cyber security attacks since it allows for remote access and scalability. As a result, the number of security vulnerabilities, problems, and case studies published in the literature was kept tracked.

The DDoS assault is one of the ongoing threats to the Internet. As a result, it is a continuous source of worry for information technology and computer security professionals. DoS attacks come in various forms like SYN flood, ICMP flood, and UDP flood [7], to name a few. In reality, several different studies are being conducted to identify distributed denial of service (DDoS) assaults, including (1) appliances from commercial hardware; (2) methods of machine learning (for both low and high-rate DDoS assaults, the PRCD method uses a partial rank correlation-based detection (PRCD) technique [8]); (3) to differentiate between malicious TCP flows and begin DDoS attacks: the mean inter-space delay variation measurement; and (4) deep learning approaches.

Cherdantseva et al. [9] used a well-established institutional research approach to examine the most recent advances in cyber security risk assessment for SCADA systems. Their research included a wide variety of SCADA security and risk studies, including the application of approximately 24 different risk assessment techniques to the SCADA systems. They proposed the following intuitive classification for the techniques they examined. First-rate: Methods tailored to the activity and in-depth explanations of the guidelines. Formula-based techniques and model-based methods are included in the second class, and class 3 is quantitative and qualitative research methods. When it comes to identifying system flaws and assessing security against possible attacks, Turk et al. [10] provided an in-depth look at the methods available. Simulation frameworks, testbeds, simulating SCADA assaults, mathematical modelling, probabilistic modelling, and also risk modelling and assessment are examples of such methods discussed in the research. System developers and service providers may use it to test their systems before putting them into operation, and end-users can use it to comprehend security provisions and adhere to all legal obligations. Forensic science and ethical hacking were also discussed in detail by the writers. These methods include scanning and penetration testing, machine learning, honeypots, intrusion prevention systems, network intrusion detection systems, and network intrusion detection systems (IDS).

Cyber-attacks and the resulting damages were discussed in detail in [11, 12]. Countermeasures should be implemented by water and sanitation facilities to avoid or minimize the harm caused by assaults on their control systems. According to the findings of the research, the following are the major problems facing the hygiene and sewerage sector. The degree to which their business systems, as well as control systems, are interdependent industrial control equipment comes in a broad variety of configurations. They also discussed potential countermeasures, such as choosing safety standards, evaluating gaps, and analysing vulnerabilities/

risks, that may be utilized to overcome these difficulties. Last but not least, they stressed that institutions should make the most of their limited funding to create and execute security-enhancing initiatives over time. Policies, procedures, training, and increasing awareness may all be used to implement cyber-security more affordably. High- and low-rate assaults may both be detected using the two-layer filtering technique described in [13]. The researchers utilized the ns-2 simulation program to evaluate the performance of the suggested approaches. High-rate DDoS assaults are detected using the first layer's average filter using metric1. Low-rate DDoS assaults are detected using a second layer called discrete Fourier transform using metric2. Both high- and low-rate DDoS assaults may be detected using the suggested techniques since they are simple to implement. However, detection accuracy is poor when high- and low-rate assaults occur at the same time. DDoS detection using deep learning in a network context is suggested in [14] and shown to be much superior to conventional machine learning methods. Layers in the model include an input, a forward-recursive layer, a reverse-recursive layer, a fully connected-hidden layer, and a fifth-output layer. There were three types of neural networks used: recurrent, long-term, and short-term (CNN). To obtain input data for the training model, all attack packets are mixed together with a random number of legal packets.

To use a supervised learning approach, Support Vector Machines, the authors in [15] presented an automated defensive strategy for identifying DDoS assaults (SVM). Sixty percent of the sample was randomly selected, 809 of which were considered normal and 809 were considered aberrant. The categorization accuracy improved significantly (by around 10%) as a consequence of the research. In [16], authors investigated hydroelectric power plants' SCADA systems for vulnerabilities and found ways to secure wireless information systems' architecture.

The research used the optimized network engineering tool to conduct a simulation-based examination of SCADA system vulnerabilities, including DDoS assaults. They experimented with two different possibilities: (1) a model that does not assault the network's infrastructure and (2) DDoS assaults that are simulated in the model. There are two main goals of the research in [17]: defect detection and data flow timing coordination for smart power grids. According to their research, DoS (denial of service) and MITM (man in the middle) assaults are the most frequent types of cyber-attacks (MIM). A NED file and programming logic were utilized with the OMNeT++ emulator. The proposed framework's strength is evident in its capacity to include a wide range of attack scenarios while also being able to offer a highly accurate behavioural analysis during simulated cyber-attacks.

3. Objectives

The main objectives of this research are as follows:

(1) To propose robust machine learning models that can differentiate between the normal signals and attack signals

(2) To increase the accuracy and performance with low computational power requirements

4. Methodology

Real-world Asphalt plant pipeline datasets are utilized to evaluate the advantages of ML-based methods for anomaly identification in SCADA systems. A dataset is explored first, and then anomaly detections are targeted as mentioned as follows.

4.1. The Asphalt Plant Pipeline Dataset. The suggested system is applied on an asphalt plant, with the batch mixed type of asphalt plant being the primary emphasis. This plant was chosen for the intended study because it has a variety of workstations for research and is trimmed to minimize disturbances. The manufacturing process is extensively investigated, and all disturbances are classified according to each workstation. Different sensors were selected for this research which are elevator speed control, Dreyer and Heater thermocouple, weight sensor, Bitumen weight sensor, Baslet Elevator, hot bin mixer, and dust collector. The unheated materials kept in the cold containers are initially fed onto the conveyor or bucket elevator by releasing cold-feed valves in batch mix asphalt plants. It transports the materials to the dryer, where they are heated and then dried. The exhaust stack and dust collectors are in charge of removing undesirable dust from the dryer exhaust. The heated and processed aggregates are delivered by hot elevator to the screening section, in which different sized particles are processed and kept in hot bins (temporary storage). When necessary, the hot bins open in regulated volumes into the weighing box. The aggregates are subsequently deposited in the Pugmill or mixing tank with the appropriate filler ratio. The fully prepared asphalt is transported out from the mixing chamber for distribution. The dataset has been made in a way that first all normal and natural values of these sensors were collected up to 10,000 values and then external inference signals were added to each sensor values and recorded values up to 10,000.

4.2. Preprocessing. Preprocessing in machine learning is the process of encoding data so that it can be read by a computer in a numerical state. Using data preprocessing methods, end-products are created standardized/normalized from raw data that includes no null values and many more. Any algorithm development or computer vision job necessitates the use of data preparation for machine learning and deep learning. Some of the preprocessing that has been on the dataset is given as follows.

4.2.1. Replacing Null Values. In the dataset, some of the values of features were missing so there was a need for some preprocessing on them because machine learning models do not work with null values. To overcome this problem, null values are replaced by the average value of that row and then replaced.

4.2.2. Normalization. Normalization means to scale down the values of features. A preprocessing method called "feature scaling" is utilized to normalize the data collection. If certain datasets are completely overrun by others, then the

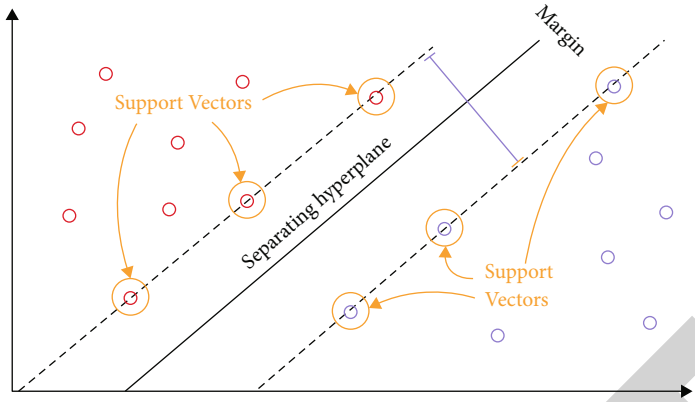

Figure 1: SVM principle working.

machine learning models will ignore the previously ignored data. In this instance, the datasets reflect different characteristics.

4.3. Machine Learning Models. Machine learning is a data analysis technique that automates the construction of analytical models. It is a subfield of artificial intelligence that depends on the idea that algorithms can learn from input, understand trends, and make decisions with or little user intervention. Machine learning already is not at all like machine learning in the past, thanks to advancements in technology. It was prompted by pattern detection and the idea that computers would learn without being trained to do tasks. In this research, some of the well-known machine learning models are used. Let us have a brief discussion about them.

4.3.1. Support Vector Machine. SVM is a type of classifier derived from statistical learning theory by Vapnik and Chervonenks introduced by Boser, Vapnik in 1992 to solve binary classification problems. And then they are extended to nonlinear regression problems. From the real value hypothesis, binary classification is done. SVM tunes the solution based on optimization theory. The simplest model of SVM is to find the maximal margin hyperplane in a chosen kernel-induced feature space. SVM is also a population algorithm for classification. SVM is based on the concept of decision planes, which defines the decision boundaries as discussed earlier. Figure 1 shows how the SVM algorithm works.

SVM uses kernel function, which finds the linear hyperplane that separates classes with the maximum margin. Figure 2 illustrates how data points (that is, support vectors) belong to two different classes (red versus blue) separately using the full margin judgment boundary.

4.3.2. K-Nearest Neighbor Algorithm. K-nearest neighbor algorithm is the kind of arrangement equation that is used for task identification. As shown in Figure 3, all cases are stored in K-nearest neighbors which are then grouped based on the closeness estimates that are available in them. Many choices in favor of neighbor groups are prioritized in the KNN characterization method for new occurrences. The most well-known class amongst its K-nearest neighbors receives the post. When there are an infinite number of measures available, the best characterization is achieved when the estimation of K is set large enough that the limit increments are perfected. KNN stands for the number of the nearest neighbors. The number of nearby neighbors is the most important factor. If the number of classes is 2, K is almost always an odd number. The measurement is regarded as the nearest neighbor calculation when $K = 1$. This is the simplest scenario. Assume the P1 is the stage that the symbol would anticipate. To start, locate the nearest highlight P1 and then the mark of the closest direct route that leads to P1.

Assume that P1 is the stage at which the mark must plan. To begin, locate the k closest highlight P1 and then characterize focuses based on the dominant portion vote of its k neighbors. Each item casts a vote in favor of their party, with the class with the most votes being considered the expectation. You find the difference between focuses using distance estimates such as Euclidean distance, hamming distance, Manhattan distance, and Minkowski distance to find the closest comparable focuses. The following are some of the most important advancements made by KNN:

(i) Should save information on your hard disc

(ii) Adjust the value of K to the desired number of neighbors

(iii) Measure the distance between the query example and the present example based on the outcomes at each data point in the data collection

(iv) Apply the distance and index of the illustration to an organized list

(v) Arrange the sorted list of distances and indices from the smallest to the largest using the distances (in an ascending order)

(vi) Choose the first K entries from the sorted list

(vii) Assemble the labels for the K entries you have selected

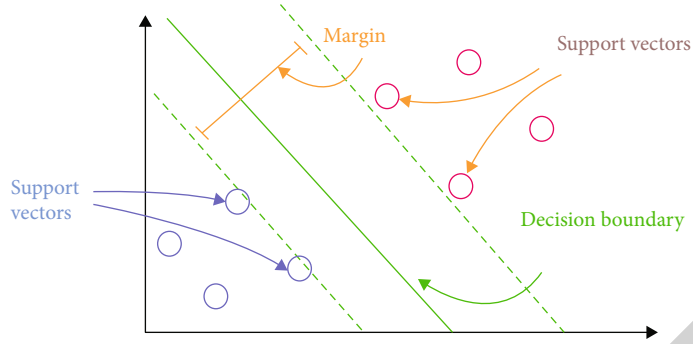

Figure 2: Classification by SVM.

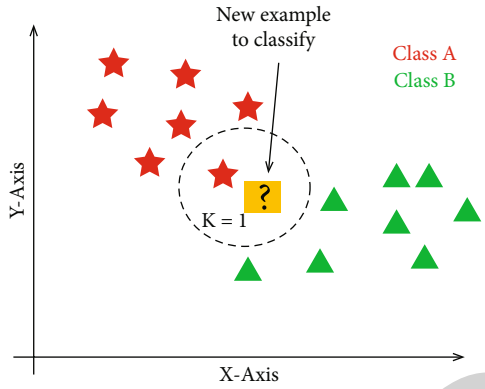

Figure 3: KNN classification process.

(viii) The mean of the K labels can be returned if there is a decline

(ix) Return the mode of the K labels if the K labels have been classified

4.3.3. Random Forest. Regression and classification issues may be solved using a random forest, which is a machine learning method. Many classifiers are used to give answers to difficult problems using ensemble learning. In a random forest method, each choice is represented by a single tree. The random forest method generates a "forest," which is then trained by bagging or bootstrapping. Bagging is a machine learning ensemble meta-algorithm that enhances accuracy. Based on the decision trees' predictions, the (random forest) method decides on a final result. Using the average of different trees output, it makes predictions. As the trees grow, so does the accuracy of the results.

A random forest algorithm's building blocks include decision trees. Making decisions with the help of decision trees is easier since it is arranged in a tree-like fashion. Two types of nodes make up a decision tree, and they are the decisions and the leaf nodes. A decision tree method separates training data into branches that are then divided into still further branches. This pattern repeats itself till a node of a leaf is obtained at the end of the tree. There is no way to separate the leaf node any further. The whole process can be seen in Figure 4.

4.4. Training the Dataset. The dataset contains 4966 rows and 129 columns in which one column has class labels which are Normal and Attack. Normal belongs to when every sensor value is ok and there is no inference of another extra signal value. On the other side, Attack labels belong to when there is an abnormality in the sensor's values and other signals. During the training, 80% of the data is used for train data and 20% of data is used for testing. The whole architecture of the proposed models is given below in Figure 5.

As can be seen in the above figure, the first step is to preprocess the data; preprocessing on the data has to be done and explained above. After preprocessing, the next step is feature selections; there are above 120 features. In this research, different combinations of features were tried and got the results for each combination. After those different classifiers are used and fit on the preprocessed data, in the end, the results are obtained which are Normal signals and Attack signals.

5. Results

SVM, KNN, and decision trees models all are utilized throughout the project's training and testing phases. The datasets utilized for these models were gathered locally, as described in the dataset portion above. Let us go over the findings of each model one at a time and discuss them thoroughly. However, there are a few important evaluation parameters to address first, and they are listed below.

5.1. Confusion Matrix. Machine learning categorization performance may be measured using a confusion matrix. Using this table, you can see in Figure 6 how well a classification model performs on a test dataset for which the actual values have been determined.

5.1.1. Precision. True positives, as well as false positives together, make up precision, which measures how accurate a test is. Precision examines the sample to determine how many false positives were included. As long as there are not any false positives (FPs), the model is considered to be 100 percent accurate. Precision will appear worse when more FPs are thrown into the mix. Both positive and negative values from the confusion matrix are required to

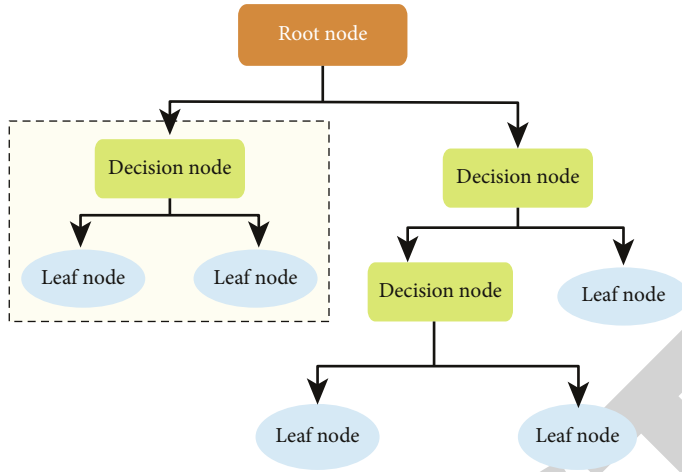

Figure 4: Random forest process.

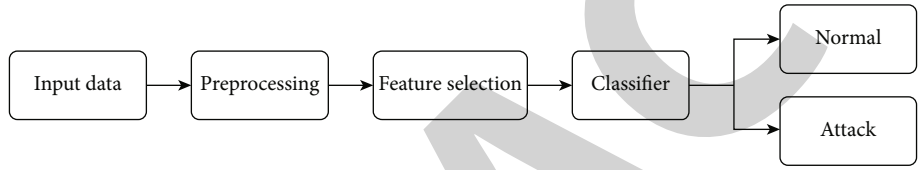

Figure 5: Proposed Methodology.

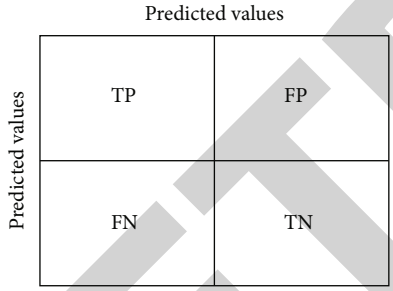

Figure 6: Confusion matrix.

determine the accuracy of a model. The formula for precision is given as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (1)$$

5.1.2. Recall. Recall, on the other hand, takes a different path. As opposed to counting how many times the model got it wrong, recall counts how many times the model got it right. The formula for the recall is given as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (2)$$

5.1.3. F1 Score. When it comes to a dataset, the *F*-score (sometimes referred to as the *F*1-score) indicates how well a model performed on it. It is used to assess binary classification methods, which sort instances into "positive" and "negative" categories. Model precision and recall are combined within *F*1-score. The formula to calculate the *F*1 score is given below.

$$F1\ \text{score} = 2 * \text{Precision} * \frac{\text{Recall}}{\text{Precision}} + \text{Recall}. \quad (3)$$

5.1.4. Evaluation of SVM. As discussed in the models in this research, the SVM model is used for the classification of Normal and Attack signals. The confusion matrix for SVM for the local dataset is given in Figure 7.

As can be seen in Figure 7, the diagonal values are not good enough for classification because the dataset is unbalanced. The only solution for this problem is to balance the dataset and then apply the classification model to it. After balancing the dataset, the results of the confusion matrix for SVM are as follows:

As shown in Figure 8, zero means Natural/Normal signal and one means Attack signal. From this confusion matrix, there are important evaluation parameters derived which are given in Table 1.

5.1.5. Evaluation of Random Forest. As in this research, three different models are used for the identification of signals which are Attack and Natural. So, the confusion matrix for a random forest is also calculated to check the model performance. Firstly, the dataset was unbalanced so the results were not good enough that can be seen in Figure 9.

As can be seen in the above figure, the results are worst as all data are misclassified by random forest, this is because

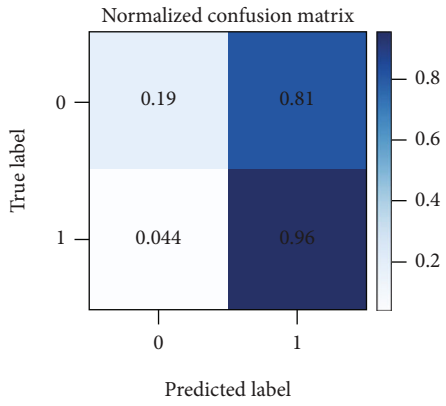

Figure 7: Confusion matrix for SVM for unbalanced dataset.

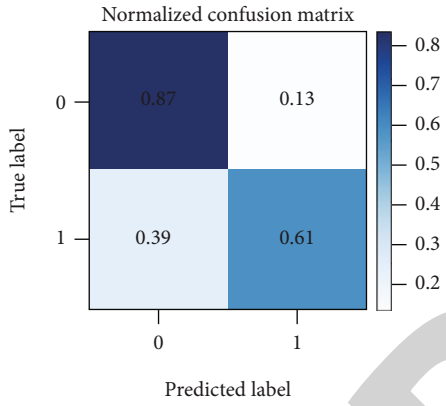

Figure 8: Confusion matrix for SVM after balancing dataset.

Table 1: Classification report for SVM.

Classes	Precision (%)	Recall (%)	*F*1 score (%)	Accuracy (%)
Natural	69	87	77	74
Attack	82	61	70	

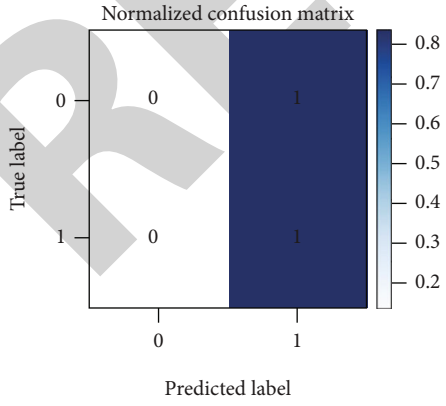

Figure 9: Confusion matrix for random forest with imbalanced dataset.

of an imbalanced dataset. After balancing the dataset, the results become good enough for generalization and classification of signals that can be seen in Figure 10.

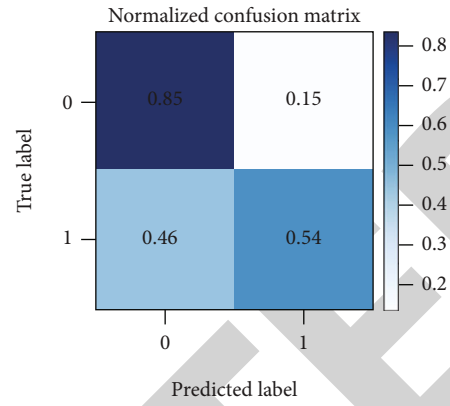

Figure 10: Confusion matrix for random forest with balance dataset.

Table 2: Evaluation parameters for random forest.

Classes	Precision (%)	Recall (%)	*F*1 score (%)	Accuracy (%)
Natural	65	85	74	70
Attack	79	54	64	

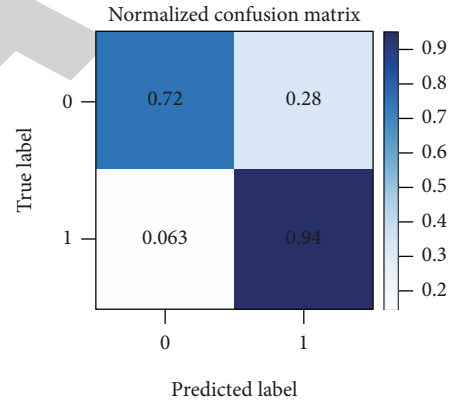

Figure 11: Confusion matrix for KNN when data is unbalanced.

Table 3: Evaluation parameters for KNN.

Classes	Precision (%)	Recall (%)	*F*1 score (%)	Accuracy (%)
Natural	75	72	74	89
Attack	93	94	93	

The rest of the evaluation parameters which are derived from the confusion matrix for the random forest is mentioned below in Table 2, including precision, recall, *F*1 score, and an accuracy score.

5.1.6. Evaluation of KNN. The third and last model that is used for the classification of Natural signal and Attack signal is KNN which is also a good classifier in machine learning, and in this research, KNN showed a good accuracy as compared to the other models that can be seen in Figure 11 of confusion matrix.

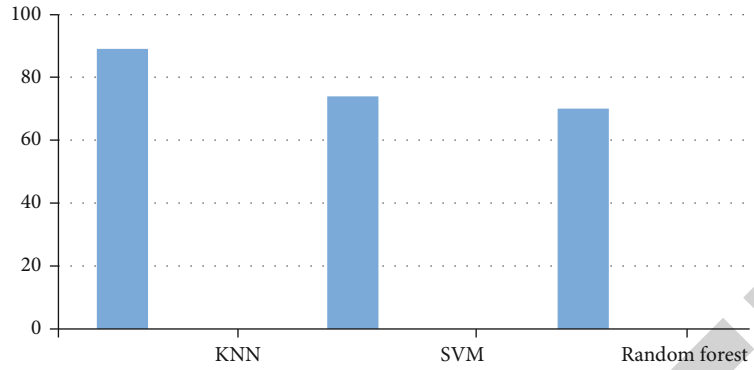

Figure 12: Comparison amongst KNN, SVM, and random forest.

Also, the classification report is generated for KNN which is shown in Table 3.

5.1.7. Comparison of Models. As can be seen in Results, the best model for generalization on unseen data is KNN which has 89% accuracy as compared to other models which are SVM and random forest which have an accuracy of 74% and 70%, respectively. Figure 12 shows a comparison between different models.

6. Conclusion

In this research study, there are three machine learning methods: (i) SVM, (ii) KNN, and (iii) random forest classifiers, to identify anomaly detection and cyber-attack risks based on the training done on local asphalt dataset to strengthen the security framework of the SCADA system. The KNN classifier has the greatest results, with 89% accuracy, while SVM and random forest get 74% and 70% accuracy, respectively. In the future, these algorithms can be used to evaluate a variety of datasets, including one that is relevant to SCADA systems. Other machine learning models can also be examined using different setup settings and datasets.

Data Availability

The datasets used and analyzed during the current study is of local asphalt plant and are not available for public use as the company can not release its network data because of the obligations of confidentiality laws and user privacy restrictions.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Slay and M. Miller, "Lessons Learned from the Maroochy Water Breach," in *Intl. Federation for Information Processing Publications*, no. 253pp. 73–82, Springer, Boston, MA, 2008.

[2] M. Abrams and J. Weiss, *Malicious Control System Cyber Security Attack Case Study - Maroochy Water Services, Australia*, Technical Report, Mitre, 2008.

[3] X. Dong, H. Lin, R. Tan, R. K. Iyer, and Z. Kalbarczyk, "Software defined networking for smart grid resilience: opportunities and challenges," in *Proceedings of the 1st ACM Workshop on Cyber Physical System Security*, pp. 61–68, New York, NY, USA, 2015.

[4] A. Bessani, P. Sousa, M. Correia, N. F. Neves, and P. Veríssimo, "The crutial way of critical infrastructure protection," *IEEE Security & Privacy*, vol. 6, no. 6, pp. 44–51, 2008.

[5] M. Brundle and M. Naedele, "Security for process control systems: an overview," *IEEE Security & Privacy*, vol. 6, no. 6, pp. 24–29, 2008.

[6] D. Dzung, M. Naedele, T. P. von Hoff, and M. Crevatin, "Security for industrial communication systems," *Proceedings of the IEEE*, vol. 93, no. 6, pp. 1152–1177, 2005.

[7] R. Ozdemir and H. Canbolat, "A SCADA System in a Construction Chemicals Manufacturing Plant," *International Journal of Scientific anf Technological Research*, vol. 2, no. 1, pp. 16–23, 2016.

[8] S. Toklu and M. Şimşek, "Two-Layer Approach for Mixed High-Rate and Low-Rate Distributed Denial of Service (DDoS) at-Tack Detection and Filtering," *Arabian Journal for Science and Engineering*, vol. 43, no. 12, pp. 7923–7931, 2018.

[9] Y. Cherdantseva, P. Burnap, A. Blyth et al., "A review of cyber security risk assessment methods for SCADA systems," *Computers & Security*, vol. 56, pp. 1–27, 2016.

[10] R. J. Turk, *Cyber incidents involving control systems*, Citeseer, 2005.

[11] S. Panguluri, W. Phillips, and J. Cusimano, "Protecting water and wastewater infrastructure from cyber attacks," *Frontiers in Earth Science*, vol. 5, no. 4, pp. 406–413, 2011.

[12] S. Zahra, W. Gong, H. A. Khattak, M. A. Shah, and H. Song, "Cross-Domain Security and Interoperability in Internet of Things," *IEEE Internet of Things Journal*, 2021.

[13] S. Toklu and M. Şimşek, "Two-Layer Approach for Mixed High-Rate and Low-Rate Distributed Denial of Service (DDoS) Attack Detection and Filtering," *Arabian Journal for Science and Engineering*, vol. 43, no. 12, pp. 7923–7931, 2018.

[14] C. Li, Y. Wu, X. Yuan et al., "Detection and defense of DDoS attack-based on deep learning in OpenFlow-based SDN," *International Journal of Communication Systems*, vol. 31, no. 5, pp. 1–15, 2018.

[15] L. M. S. Hoyos, E. G. A. Isaza, J. I. Vélez, and O. L. Castillo, "Distributed denial of service (DDoS) attacks detection using machine learning prototype," in *Distributed Computing and Artificial Intelligence, 13th International Conference*, vol. 474, pp. 33–41, Springer, Cham.

[16] J. D. Markovic-Petrovic and M. D. Stojanovic, "Analysis of SCADA system vulnerabilities to DDoS attacks," in *2013 11th international conference on telecommunications in modern satellite, cable and broadcasting services (telsiks)*, pp. 591–594, Nis, Serbia, 2013.

[17] D. Bhor, K. Angappan, and K. M. Sivalingam, "Network and power-grid co-simulation framework for Smart Grid wide-area monitoring networks," *Journal of Network and Computer Applications*, vol. 59, pp. 274–284, 2016.