ZKG INTERNATIONAL ISSN: 2366-1313

CYBER SECURITY IN POWER SYSTEMS USING META HEURISTIC AND MACHINE LEARNING ALGORITHMS

[1]Bhimavarapu Revathi, [2]G. V. Jyothirmai, [3]G. Indu, [4]S. Rishitha, [5]G. Sindhuja

[1]Assistant Professor, Department of CSE(DS), Malla Reddy Engineering College for Women (Autonomous Institution – UGC, Govt. of India), Hyderabad, INDIA.

[2345]UG,Department of CSE(DS), Malla Reddy Engineering College for Women (Autonomous Institution – UGC, Govt. of India), Hyderabad , INDIA.

ABSTRACT Supervisory Control and Data Acquisition system linked to Intelligent Electronic Devices over a communication network keeps an eye on smart grids' performance and safety. The lack of algorithms protecting the power system communication protocols makes them vulnerable to cyberattacks, which can result in a hacker introducing false data into the operational network. This can result in delayed attack detection, which might harm the infrastructure, cause financial loss, or even result in fatalities. Similarly, attackers may be able to feed the system with fake information to hoax the operator and the algorithm into making bad decisions at crucial moments. This paper attempts to identify and classify such cyber-attacks by using numerous deep learning algorithms and optimizing the data features with a metaheuristic algorithm. We proposed a Restricted Boltzmann Machine-based nature-inspired artificial root foraging optimization algorithm. Using a publicly available dataset produced in Mississippi State University's Oak Ridge National Laboratory, simulations are run on the Jupiter Notebook. Traditional supervised machine learning algorithms like Artificial Neural Networks, Convolutional Neural Networks, and Support Vector Machines are measured with the proposed algorithm to demonstrate the effectiveness of the algorithms. Simulations show that the proposed algorithm produced superior results, with an accuracy of 97.8% for binary classification, 95.6% for three-class classification, and 94.3% for multi-class classification. Thereby outperforming its counterpart algorithms in terms of accuracy, precision, recall, and f1 score.

INDEX TERMS Artificial neural network, artificial root foraging, cyber security, deep learning

1 INTRODUCTION

The extraordinarily intricate architectural design of the electrical power systems must be handled cautiously and with the best control strategy feasible to ensure both the protection of human life and the system's safety [1]. The system becomes more complex as the control process must run more quickly [2]. Automated devices are introduced to modern power systems to make operating them easier. The number of pieces of protective equipment that are part of the system is directly impacted by operational demand and consumer count [3]. Recent years have seen the development of automated systems for connected power module protection, automation, and control [4]. Protective device performances have somewhat improved as a result of developments in algorithms and power systems architecture [5], [6].

However, the likelihood of security problems increases as the number of connections to the power system modules intensifies. Hence the quality of control is expected to be in the higher range for modern power systems. The contemporary power systems are implemented with various International Electrotechnical Commission (IEC) standards [7], [8] and are generally operated with six significant components, as depicted in Figure 1. Generators, transformers, and safety equipment are all part of the power system's electrical components. These primary hardware ranges and ratings change depending on the loads connected to the network. The protection mechanisms built into the electrical system also differ depending on the linked equipment's location and nature [9]. The control components include the synchronization model and operational modules for transmitting the required signal to the digital modules used for the operation. The power system's information and communication devices, which transmit control signals between linked systems and components across wired or wireless networks, are represented by digital modules [10]. The convergence network regulates the power flow in the connected system by analyzing the load requirement and the power system state. The importance of the convergence networks increases when the power system is linked to Distributed Energy Resources (DERs) [11], [12]. The regulatory components ensure that the integration of power is constantly smooth and efficient

2 RELATED SERVICES

The smart grid protection strategy uses

local measures or external devices to build a smart grid protection system that is both effective and efficient. However, one of the key issues is the ability to connect physical and digital components to suit the configuration of the system. Measurement of data source authentication system was developed to analyze the data flow of a power system by extracting the features through an ensemble empirical mode decomposition model with the Fast Fourier Transform (FFT) technique. The experiment was conducted with a back-propagation neural network for data classification. An accuracy of 80.9% is achieved, and comparatively, it is better than the traditional long short-term memory (LSTM) model's accuracy of 77.8% [17]. To train the neural network algorithms, a sizable dataset is required. The performance of a neural network algorithm's prediction process is influenced by the amount of training data present in the network. The authors of [18] generated a power system dataset based on IEC 61850 Generic Object-Oriented Substation Event (GOOSE) communication for developing a reliable cybersecurity system. The compone

nts of the power system are divided into numerous categories to monitor the load demand in different areas. Due to environmental conditions, the associated field will see variations in demand in particular. The system is more vulnerable to cyber threats since the scattered devices are connected through different channels [19]. The testbedbased power system quality analysis is one of the familiar methods widely used for observing the response of the power system in different scenarios. The test bed generates different kinds of cyber security issues to analyze and formulate a defending algorithm. An OMNeT++-based simulation technique was structured [20] to analyze the nature of cyberattacks in a bidirectional communication network. The model was integrated with Power Systems Computer Aided Design (PSCAD) for the power simulation. The physical power systems are open to dynamic data injection attacks. An example is the ease with which the energy consumption values on smart meters could be altered. So, an interval state estimation method was developed to analyze the possible variations in the readings with respect to time. A kernel quantile regression is also incorporated in the work to estimate the uncertainties in renewable and electric load forecasting applications [11]. The cyberattack on the Internet of Things (IoT)-based smart grids may affect the costly and important systems that are connected to the power system. The hospital equipment and electric train are some of the costlier and most needed systems that always depend upon the quality of the power supply. Therefore, a blockchainbased technique was equipped with Hilbert-Huang transform to estimate power quality through the data collected from voltage and current sensors. The experimental work founds satisfied with the performance of the proposed model on false data injection attacks.

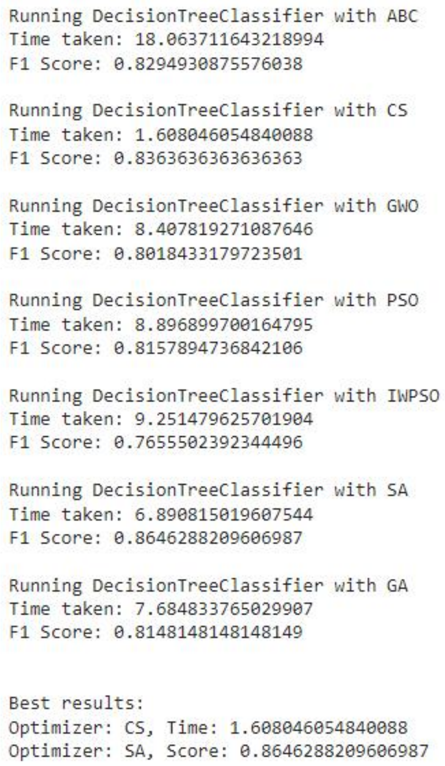
3 METHODOLOGY

The optimization process may be presented as the process of determining the best method to use existing resources while not breaking any restrictions that may exist. This strategy consists of multiple steps: mathematically defining a system model that reproduces its behavior, determining its variables and constraints, establishing the objective function, and, finally, seeking the states that produce the most desirable results by maximizing or minimizing the objective function. The optimization search strategy can be performed using any of its appropriate categories, such as quantum-based techniques, meta-heuristicbased approaches, and multi-objective-based techniques [12]. However, the main purpose of solving complicated optimization issues is to find a solution, regardless of how good it is. When at least a solution is found, numerous methods can be used to enhance it. This is the fundamental principle behind developing metaheuristic optimization algorithms. Meta means upper level or beyond, while heuristic means to know, find, or direct an investigation, which is where the word heuristics originates. On the other hand, heuristics represents a collection of rules applied while addressing a problem based on experience [13]. Metaheuristics are approximate methods that combine basic heuristic principles to produce a more efficient exploration and exploitation of research space [14], where the search space is the space that includes all the possible solutions that are bounded by the physical system limitations. The dimensions of the search space depend on the number of optimization variables that represent the set of the required parameter. Voß et al. [15] define a metaheuristic as a repeated process that leads and modifies tasks while employing subordinate heuristics to facilitate obtaining optimal or near-optimal outcomes. The MA can function with single or many solutions using a minimization or maximizing approach at each iteration. Metaheuristic algorithms have been created to deal with the increasing complexities of the problem, particularly with the inclusion of uncertainties into the system, which may surpass the constraints of traditional algorithms.

The power networks are large-scale, dynamic systems with various users, cables, transformers, and generation units. A power system's primary goal is to safely, consistently, and cost-effectively supply users with enough high-quality energy. Several system factors are control variables, such as the generator's active power, the compensator's reactive power, and the common bus voltage. They may be regulated independently and directly influence power flows and the system's stability. Other variables, such as the voltages of the load bus, the reactive power, and the power flows in the branches, are included as dependent variables [16,17]. Changing the control variables combinations can lead to the power-balance equation, but only certain combinations permit achieving the predefined objectives. Determining the best combinations which provide the desired state can be achieved by solving the optimal power flow problem (OPF) issue [18]. By making optimal adjustments to the control variables, the considered objective function may include reducing fuel consumption, power loss, and attenuating voltage deviation (VD), respecting the system restrictions. Both deterministic (traditional) and metaheuristic optimization algorithms can be used to overcome this issue [19]. Gradient, Newton's approach, linear programming (LP), and quadratic programming (QP) are examples of classical optimization methods applied to OPF issues. Unfortunately, due to the high nonlinearity and nonconvexity issues, these approaches cannot give a global solution and only obtain local solutions [20]

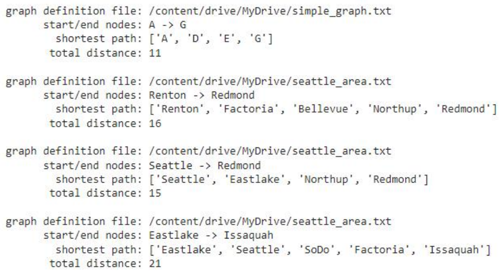
4 RESULTS

The experiment was performed in a Jupyter notebook on a 16GB RAM Intel 7 processor system. The proposed RF-RBM technique was tested against conventional CNN, ANN, and SVM algorithms because those were found to be successful models in several intrusion detection studies [37], [49]. In this, the SVM is a machine learning-based technique, whereas CNN and ANN are deep learning-based techniques. We utilize the hyperparameters given in Table 3 for the simulations, and we classify the network intrusion through different algorithms. One of the most used neural network algorithms, CNN, can provide a higher accuracy rate when the training data samples are plentiful. However, because CNN learns characteristics from a large dataset, preprocessing of the training data is minimal.



```
Running DecisionTreeClassifier with ABC
Time taken: 18.063711643218994
F1 Score: 0.8294930875576038

Running DecisionTreeClassifier with CS
Time taken: 1.608046054840088
F1 Score: 0.8363636363636363

Running DecisionTreeClassifier with GWO
Time taken: 8.407819271087646
F1 Score: 0.8018433179723501

Running DecisionTreeClassifier with PSO
Time taken: 8.896899700164795
F1 Score: 0.8157894736842106

Running DecisionTreeClassifier with IWPSO
Time taken: 9.251479625701904
F1 Score: 0.7655502392344496

Running DecisionTreeClassifier with SA
Time taken: 6.890815019607544
F1 Score: 0.8646288209606987

Running DecisionTreeClassifier with GA
Time taken: 7.684833765029907
F1 Score: 0.8148148148148149


Best results:
Optimizer: CS, Time: 1.608046054840088
Optimizer: SA, Score: 0.8646288209606987
```

Fig 1: Acuracy Values



```
graph definition file: /content/drive/MyDrive/simple_graph.txt
      start/end nodes: A -> G
        shortest path: ['A', 'D', 'E', 'G']
       total distance: 11

graph definition file: /content/drive/MyDrive/seattle_area.txt
      start/end nodes: Renton -> Redmond
        shortest path: ['Renton', 'Factoria', 'Bellevue', 'Northup', 'Redmond']
       total distance: 16

graph definition file: /content/drive/MyDrive/seattle_area.txt
      start/end nodes: Seattle -> Redmond
        shortest path: ['Seattle', 'Eastlake', 'Northup', 'Redmond']
       total distance: 15

graph definition file: /content/drive/MyDrive/seattle_area.txt
      start/end nodes: Eastlake -> Issaquah
        shortest path: ['Eastlake', 'Seattle', 'SoDo', 'Factoria', 'Issaquah']
       total distance: 21
```

Fig 2: Shortest path for route

Three layers make up a conventional CNN: a convolution layer, a pooling layer, and a fully connected layer. The convolution layer is set up to separate the kernel's learnable parameters from the input data. The kernel clarifies to the layer the kind of information that is available [17] and [19]. Data is forwarded by the kernel to different neurons in the pooling layer, which lowers the spatial complexity of the retrieved information in the convolution layer. All of the CNN's neurons are interconnected in the fully connected layer with their biases toward comprehending the data that has been gathered

5 CONCLUSION

In this study, we present a nature-inspired restricted Boltzmann machine algorithm to detect and classify the types of attacks in the smart grids' SCADA systems. The fundamental notion is that the artificial root foraging optimization method is designed on the biological root growth optimization algorithm. To demonstrate the optimization capability, the dataset features were fine-tuned using the artificial root foraging algorithm before the neural network algorithm. The proposed RF-RBM algorithm is compared to three cutting-edge neural network algorithms in the experimental study, which was conducted in three categories: binary classification, three-class classification, and multi-class classification. The outcomes of the experiments demonstrate that the proposed algorithm RF-RBM is best suited for cyberattack detection and classification in SCADA systems for smart grids. This is shown by the excellent accuracy, sufficient precision, respectable recall, and a high f1 score demonstrated by the proposed algorithm.

6 REFERENCES

1. De Leon-Aldaco, S.E.; Calleja, H.; Aguayo Alquicira, J. Metaheuristic Optimization Methods Applied to Power Converters:
A Review. IEEE Trans. Power Electron. 2015, 30, 6791–6803. [CrossRef]

2. Dantzig, G.B. Linear Programming. Oper. Res. 2002, 50, 42–47. [CrossRef]

3. Bertsekas, D.P. Nonlinear Programming. J. Oper. Res. Soc. 1997, 48, 334. [CrossRef]

4. Bellman, R. Dynamic Programming. Science 1966, 153, 34–37. [CrossRef]

5. Fletcher, R. Practical Methods of Optimization; John Wiley & Sons, Ltd.: Chichester, UK, 2000; ISBN 9781118723203.

6. Box, F. A Heuristic Technique for Assigning Frequencies to Mobile Radio Nets. IEEE Trans. Veh. Technol. 1978, 27, 57–64.[CrossRef]

7. Devarapalli, R.; Bhattacharyya, B.; Sinha, N.K. An Intelligent EGWO-SCA-CS Algorithm for PSS Parameter Tuning under System Uncertainties. Int. J. Intell. Syst. 2020, 35, 1520–1569. [CrossRef]

8. Wolpert, D.H.; Macready, W.G. No Free Lunch Theorems for Optimization. IEEE Trans. Evol. Comput. 1995, 1, 67–82. [CrossRef]

9. Tsai, C.-W.; Chiang, M.-C.; Ksentini, A.; Chen, M. Metaheuristic Algorithms for Healthcare: Open Issues and Challenges. Comput.Electr. Eng. 2016, 53, 421–434. [CrossRef]

10. Shah, P.; Sekhar, R.; Kulkarni, A.J.; Siarry, P. Metaheuristic Algorithms in Industry 4.0; CRC Press: Boca Raton, FL, USA, 2021; ISBN 9781003143505.

11. Dai, Y.; Zhou, X.; Chu, X.; Li, C.; Su, Z.; Zhu, Z.; Cui, P.; Qi, J.; Wang, Y. Effect of Entrainer Thermodynamic Properties on the Separation of Ternary Mixtures Containing Two Minimum Boiling Azeotropes by Extractive Distillation. Ind. Eng. Chem. Res. 2022, 61, 15273–15288. [CrossRef]

12. Wang, Y.; Bu, G.; Wang, Y.; Zhao, T.; Zhang, Z.; Zhu, Z. Application of a Simulated Annealing Algorithm to Design and Optimize a Pressure-Swing Distillation Process. Comput. Chem. Eng. 2016, 95, 97–107. [CrossRef]

13. Byles, D.; Mohagheghi, S. Sustainable Power Grid Expansion: Life Cycle Assessment, Modeling Approaches, Challenges,

andOpportunities. Sustainability 2023, 15, 8788. [CrossRef]

14. Massoud Amin, S.; Wollenberg, B.F. Toward a Smart Grid: Power Delivery for the 21st Century. IEEE Power Energy Mag. 2005, 3,34–41. [CrossRef]

15. Agustriyanto, R.; Zhang, J. Obtaining the Worst Case RGA and RDGA for Uncertain Systems via Optimization. Proc. Am. Control Conf. 2007, 5360–5365. [CrossRef]

16. Han, Q.; Wen, M. An Uncertain Optimization Model for Repairable Inventory System. In Proceedings of the 2014 Prognostics and System Health Management Conference (PHM-2014 Hunan), Zhangjiajie, China, 24–27 August 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 378–382.

17. Fu, X.; Guo, Q.; Sun, H. Statistical Machine Learning Model for Stochastic Optimal Planning of Distribution Networks Considering a Dynamic Correlation and Dimension Reduction. IEEE Trans. Smart Grid 2020, 11, 2904–2917. [CrossRef]

18. Fu, X. Statistical Machine Learning Model for Capacitor Planning Considering Uncertainties in Photovoltaic Power. Prot. Control Mod. Power Syst. 2022, 7, 5. [CrossRef]

19. Lee k, Y.; EI-Sharkawi, M. Modern Heuristic Optimization Techniques; Lee, K.Y., El-Sharkawi, M.A., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2008; ISBN 9780470225868.

20. Rezk, H.; Olabi, A.G.; Wilberforce, T.; Sayed, E.T. A Comprehensive Review and Application of Metaheuristics in Solving the Optimal Parameter Identification Problems. Sustainability 2023, 15, 5732. [CrossRef]

21. Rezk, H.; Olabi, A.G.; Sayed, E.T.; Wilberforce, T. Role of Metaheuristics in Optimizing Micrograms Operating and Management Issues: A Comprehensive Review. Sustainability 2023, 15, 4982. [CrossRef]