
Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org

Volume 26, Issues 4, pp. 1754–1762, DOI 10.12694/scpe.v26i4.4441

LARGE-SCALE INTELLIGENT NETWORK ATTACK DETECTION BASED ON HIERARCHICAL SYMBOLIC DYNAMIC FILTERING

WEI LI*, BO FENG †, AND LINA WANG‡

Abstract. Smart grid technology enhances grid security, reliability, and efficiency. In order to ensure efficient and reliable power distribution, it must address new vulnerabilities brought about by digital communication technology. In this paper, a new Energy Efficient Anomaly Detection (EEAD) technique is proposed, which uses HSDF pre-processing and HMM learning. A number of subsystems are initially created within the system. Hierarchical symbolic dynamic filtering (HSDF) converts time series data into symbol sequences and then learns the causal relationship between the nominal characteristics of subsystems. Then the converted sequences will be fed to the Hidden Markov model (HMM) which detects the anomaly by calculating the occurrence probability of the current observation based on the trained network. Simulation results on an IEEE 118 bus system to verify the performance of the suggested method under various operating conditions such as False Positive Rate, Detection rate, Accuracy, and True Positive Rate.

Key words: Hierarchical symbolic dynamic filtering, Hidden Markov model, Energy Efficient Anomaly Detection, Smart grid, Network

1. Introduction. The smart grid is still in its early stages of development, but as a cyber-physical network and essential service, it is vulnerable to threats that are not anticipated and that arise when attackers insert inaccurate, fraudulent, or malware information. The smart grid technology became added to enhance the prevailing energy gadget via modernization. Various strength control and operation strategies are utilized in the clever grid era to reap the very best feasible benefits [1]. These management and operations technologies include smart meter implementations and consumer applications, smart inverters, production meters, generators that generate renewable energy, and resources. Various energy savings in the grid centers will be installed.

Smart grid safety, security, and tracking will be improved by advanced monitoring and SCADA (supervisory control and data acquisition) systems. However, energy systems are susceptible to cyberattacks that may compromise the security of the grid [2]. Therefore, cyberattacks can cause computer viruses and anomalies that compromise the security and resilience of smart grid systems. Cyberattacks can damage equipment by overloading it, or generating erroneous requests and generating large amounts of energy. Additionally, malicious attacks can also cause false-negative results that affect false overload conditions in the power grid [3]. Therefore, real-time detection of cyberattacks is essential to ensure the reliable performance of critical infrastructures, including the smart grid.

Targeted cyberattack detection and resilience to attacks require constant monitoring of the online system. The literature suggests solutions to protect electrical systems and work. However, they are theoretically expensive, technically difficult, and not suitable for large and complicated circuits [4]. These challenges present opportunities for information analytics methods like machine learning that use AI to tackle complex structured datasets to detect and prevent cyberattacks. Attack detection measures should be investigated as such malicious activity adversely affects the safe and reliable operation of SGs [5].

In order to resolve these drawbacks, this research proposes a novel Energy Efficient Anomaly Detection (EEAD) technique, which promotes safety by detecting cyberattacks on the smart grid. This proposed EEAD technique system consists of the following major contributions:

- In this paper, a new Energy Efficient Anomaly Detection (EEAD) technique is proposed, which uses HSDF pre-processing and HMM learning.

*Shijiazhuang Institute of Railway Technology, Shijiazhuang, Hebei, 050041, China (WeiLi728@163.com)
†Shijiazhuang Institute of Railway Technology, Shijiazhuang, Hebei, 050041, China (Corresponding author, BoFeng27@ 126.com)
‡School of Future Information Technology, Shijiazhuang University, Shijiazhuang, Hebei, 050035, China (LinaWang81@163.com)

- A number of subsystems are initially created within the system. Hierarchical symbolic dynamic filtering (HSDF) converts time series data into symbol sequences and then learns the causal relationship between the nominal characteristics of subsystems.
- Then the converted sequences will be fed to the Hidden Markov model (HMM) which detects the anomaly by calculating the occurrence probability of the current observation based on the trained network.

The remaining section of the research is structured as defined. The literature review is represented in section 2. The proposed method is represented in section 3. Section 4 represents the experimental results and Section 5 summarizes the conclusion.

2. Literature survey. Energy efficient anomaly detection techniques has increased the risks imposed by serious cyber security attacks such as timed attacks, coordinated. Many studies have been conducted to solve this problem. Among those, some of the techniques have been reviewed in this section.

In 2019 Karimipour, H., et al [6] Presented an unsupervised online anomaly identification algorithm that develops effective computational methods for identifying relationships of cause and effect between subsystems by using SDF, time series data partitioning, and feature extraction techniques. Results confirm system performance with 99% accuracy, 98% TPR, and less than 2% FPR.

In 2019 Sakhnini, J., et al [7] presented a feature selection (FS) method that analyses three different supervised learning methods. Each technique can be used in three different ways. These techniques have been based on IEEE 14-bus, 57-bus, and 118-bus systems to evaluate their versatility. A simulation study reveals that combining supervised learning and FS heuristic methods improves the performance of classification algorithms for detecting FDI attacks.

In 2019 Geris, S. et al [8] presented an Anomaly detection method based on feature clustering combined with a linear correlation coefficient algorithm (FGLCC). The suggested method uses decision trees as classifiers. To verify the performance, we applied the proposed method to an IEEE 39 bus system. The outcome confirms the higher accuracies (96%) and detection rate (97%) with a minimum false positive rate (1.65%) comparable to existing techniques.

In 2020 Acosta, M.R.C., et al [9] presented a DR-based ML scheme to detect SCA threads in SG networks. To overcome the computational complexity caused by the multi-dimensional space of large-scale energy systems, we apply the KPCA method to convert the data into low-dimensional space. The numerical outcome demonstrates that the suggested scheme outperforms modern approaches and improves accuracies in detecting stealthy cyberattacks in smart grid measurements.

In 2020 Al-Abassi, A., et al [10] presented a deep learning-based technique called Ensemble Stacked AutoEncoder (ESAE) that aims to address the problem of information imbalance. This technique develops a deep represents learning model to generate accurate balanced represents, which leads to high performance on unbalanced information. Using IEEE 14-bus, 30-bus, and 57-bus system test cases, the suggested technique is evaluated for all degrees of data imbalance.

In 2020 Gunduz, M.Z. et al [11] presented IoT-based smart grid threats, and possible solutions are analyzed. It provides a detailed overview of the smart grid cyber security state, focusing on types of cyberattacks. Special emphasis is placed on discussing and researching network vulnerabilities, attack mitigations, and security requirements. The aim is to gain a deeper understanding of cybersecurity vulnerabilities and solutions and to provide guidance for future research directions of cybersecurity in smart grid applications.

In 2020 Dou, C., et al [12] presented a mechanism that combines variable-mode decomposition (VMD) and machine learning. For the cause of figuring out the traits of FDIA, VMD is used to decompose the gadget country time collection into a hard and fast of additives with specific frequencies. The simulation outcomes reveal the effectiveness and robustness of our method.

In 2021 Monday, H.N., et al [13] presents a method for detecting distributed denial of service (DDoS) attacks on smart grid infrastructure. For the cause of figuring out the traits of FDIA, VMD is used to decompose the gadget country time collection into a hard and fast of additives with specific frequencies. The simulation outcomes reveal the effectiveness and robustness of our method. The experimental consequences show that the proposed technique detects DDoS assaults with a higher detection rate and a totally low false alarm rate.

In 2021 Khazaei, J. et al [14] Presented A two-level mixed linear programming (BMILP) model has been

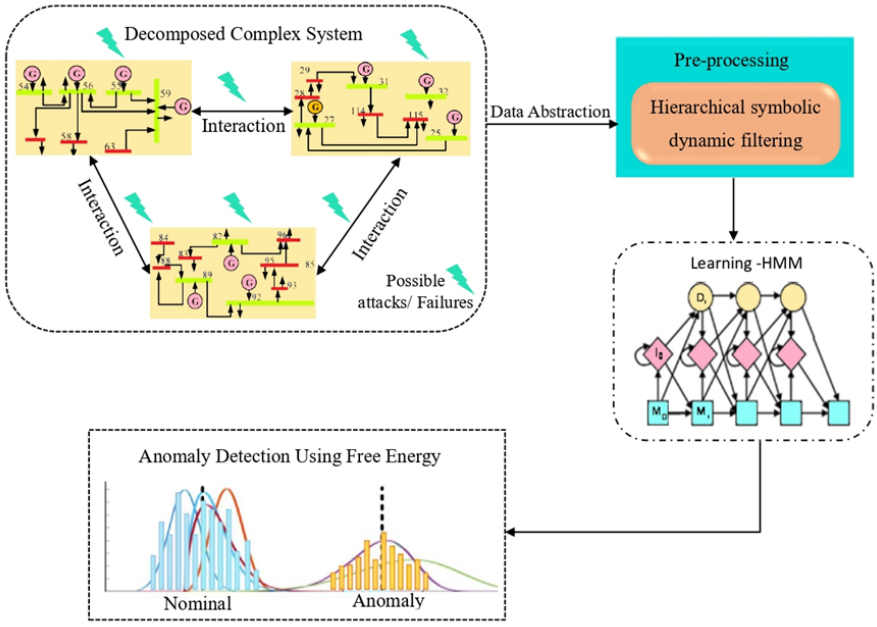

Fig. 3.1: Overall block diagram of the proposed EEAD Method

designed for accurately simulating false data information (FDI) for the purpose of traversing various transmission lines and causing power outages in large-scale networks. Compared with the present study, the suggested model assumes that the attackers have limited access to the measurement buses. It models attacks on targeted transmission lines that go undetected using the existing DC state estimation method.

In 2023 Bahadoripour, S., et al [15] presented a multimodal network attack detection model proposed to analyze the network and sensor methods of the ICS environment and construct an abstract generic representation based on these methods. The results of using the Safe Water Treatment (SWaT) technique demonstrate that the suggested model can existing unimodal models by achieving an accuracy of 0.99, recovering 0.98 and f-measure of 0.98, showing the effectiveness of using both methods in a combined model to detect network attacks.

From the above reviews, is found that these methods possess some drawbacks such as the computational complexity caused by the multi-dimensional space of higher-scale energy systems. In order to overcome these drawbacks a novel Energy Efficient Anomaly Detection technique is proposed in this section.

3. Proposed method. This paper proposes an Energy Efficient Anomaly Detection (EEAD) technique, which uses HSDF pre-processing and HMM learning. A number of subsystems are initially created within the system. Hierarchical symbolic dynamic filtering (HSDF) converts time series data into symbol sequences and then learns the causal relationship between the nominal characteristics of subsystems

Then the converted sequences will be fed to the Hidden Markov model (HMM) which detects the anomaly by calculating the occurrence probability of the current observation based on the trained network. The overall block diagram for the suggested technique has been given in Figure 3.1.

3.1. Preprocessing. Pre-processing is a method used to enhance specific aspects and remove unwanted distortions from the input image. Here, the Hierarchical Symbolic Dynamic Filtering (HSDF) has been proposed in Energy Efficient Anomaly Detection (EEAD) technique.

3.1.1. Hierarchical Symbolic Dynamic Filtering (HSDF). Inference is an approach to estimating the probability of assigning a slow-time epoch τ_n to a class $D^m \in D(here, D = \{D^m \forall m = 1, \cdots, R\}$ is the set

of present classes) or a newly created class D^{P+1}. Let the symbol sequence for the current slow-time epoch be $\widetilde{U_{\tau_n}}$. Afterward, the probability of class D^m for the present epoch τ_n is provided by $\mu(\widetilde{U_{\tau_n}})$ as demonstrated in the previous article. $Pr(D^m, U^m|\widetilde{U_{\tau_n}})$ which is similar to $Pr(D^m, U^m|\widetilde{U_{\tau_n}})$ in this instance because all current classes are finished and identified the symbols sequence , can be utilized to represent the following probability for class determination. Using this configuration, acquire the information that follows:

$$Pr(D^m|\widetilde{U_{\tau_n}}) \propto \mu(U_{\tau_n}|U^m)\forall m = 1, \cdots, R \tag{3.1}$$

It uses the Chinese Restaurant Process (CRP) to introduces the occurrence of a new type D^{R+1} with the CRP hyperparameter γ_n as follow (noted that the epoch-specific hyperparameter test τ_n).

$$\mu_{\gamma_n}(D^{R+1}|\widetilde{U_{\tau_n}}) = \gamma_n \sum_{m=1}^{R} \mu(\widetilde{U_{\tau_n}}|U^m) \Rightarrow \sum_{m=1}^{R} \mu_{\gamma_n}(D^m|\widetilde{U_{\tau_n}}) = (1-\gamma_n) \sum_{m=1}^{R} \mu(\widetilde{U_{\tau_n}}|U^m) \tag{3.2}$$

Here, we introduce the concept of sticky into the proposed algorithm based on the fact that practical systems usually cannot oscillate their operating point or internal parameter conditions every slow period. In the current context, this means that if a slow-time epoch τ_{n-1} belongs to a class, $D^R \in D$, there is a high probability that the stream data is new at epoch τ_n also belongs to D^R. This concept is incorporated into the formula by giving a positive trend for the final seen classes D^R as follow:

$$\mu_{\gamma_n}(D^R|\widetilde{U_{\tau_n}}) = \max\{\frac{r}{1-r} \sum_{m=1}^{R} +1\mu_{\gamma_n}(D^m|\widetilde{U_{\tau_n}}), \mu_{\gamma_n}(D^r|\widetilde{U_{\tau_n}})\} \tag{3.3}$$

Here, $0 < r < 1$ is the coefficient of adhesion. Note that the rationale for this fit is to ensure some minimal probability for the final seen classes D^R and this context, the suggested wording confirms,

$$\frac{\mu_{\gamma_n}(D^r|\widetilde{U_{\tau_n}})}{\sum_{m=1}^{R+1} \mu_{\gamma_n}(D^m|\widetilde{U_{\tau_n}})} \geqslant r \tag{3.4}$$

This can be verified by considering the extreme cases, here $\mu_{\gamma_n}(D^r|\widetilde{U_{\tau_n}})$ before applying the stickiness factors. According to numeric simulation outcomes in the remaining section, the "stickiness" adjustment substantially decreases "hunting behaviour" in the process of class detection and development. Quantitatively, "predatory behaviour" in detection and layering processes has been substantially reduced by the "tracking" adjustment. The numeric simulation outcomes in the remaining section will illustrate.

Lastly, the $\mu_{\gamma_n}(D^m|\widetilde{U_{\tau_n}})$ factors are normalized to obtained the posterior probabilities $Pr(D^m|\widetilde{U_{\tau_n}})$ for each class as occurs:

$$Pr(D^m|\widetilde{U_{\tau_n}}) = \frac{\mu_{\gamma_n}(D^m|\widetilde{U_{\tau_n}})}{\sum_{\mu_{\gamma_n}}(D^m|\widetilde{U_{\tau_n}})} \tag{3.5}$$

A random sample is generated from this distribution for identity determination and class generation at test time τ_n.

3.2. Learning HMM (Hidden Markov Model). Here, the pre-processing of hierarchical index dynamic filtering checks whether the anomaly detection is an attack or not. Hidden Markov models (HMMs) are doubling stochastic procedures characterized by undetected (hidden) state processes that can be discovered through an additional set of stochastic events generated by a set of observations. A set of hidden states $s = \{s_1, \cdots, s_N\}$ derived from observations of the network, here D is the no.of states in the modelling image, permits HMM to characterize system dynamics. Strictly speaking, an HMM can be officially explained by the unknown parameters $\theta = \{\pi, A, B\}$, Where $\pi \in RN$ is the initial probability vector that determines the initial probabilities of the system in various states; $A \in R^{N \times N}$ is the transition probability matrix associated with the change of state of the latent variables; B is the issue opportunity matrix representing the probabilities of predicting a particular price in array S.

The hidden state sequences that expose a potential state s_i of $x^{(n)}$ across duration are particularly represented as $X^{(n)} = [X_1^{(n)}, \cdots, X_l^{(n)}]$, where $1 \leqslant i \leqslant M$. The probabilities of transition from a state s_n to the state s_n for $i, j \in \{1, \cdots, M\}$, is provided by the formula $a_{m,n} = R(X_i = s_m | X_{i-1} = s_n)$, which is utilized to determine the entries of the transition probabilities of matrix A. Lastly, the density function of the probability distribution of the time-sample Z_i at time t, while Z_i is in the state s_m, determines the entries of the emission probabilities of matrix B, which is expressed by $b_{m,i} = R(Z_i | X_i = s_m)$. A mixture of Gaussian distributions is assumed as the emission probability distribution B in the proposed analysis, with M multivariable standard densities. Often, changes in vegetation life cycles influence the set of states s. The anomalous discrimination within the AD-primarily based totally HMM that arbitrates among the 2 hypotheses is described as follows:

H0: Anomalies are absent
H1: Anomalies are present,

where under hypothesis H1, a given graph $Z^{(n)}$ is considered the capacity of a given package is then followed as

$$\begin{aligned} &R(Z^{(n)}|\theta) = \textstyle\sum_{allX^{(n)}} R(Z^{(n)}|X^{(n)}) R(X^{(n)}, \theta) = \sum_{X_1^{(n)}, \cdots, X_l^{(n)}} \pi_{X_1}^{(n)} b_{X_1}^{(n)}, \\ &aX_1{}^{(n)}, x_2^{(n)} \cdots a_{X_{I-1}}^{(n)}, X_I^{(n)}, I \end{aligned} \tag{3.6}$$

The HMM method parameter vector was determined by increasing the logging capability, in order to accurately represent and learning the temporal framework of the fundamental information for AD.

$$\hat{\theta} = arg_\theta \max \log \sum_{n=1}^{G} R(Z^{(n)}|\theta) \tag{3.7}$$

where $\hat{\theta}$ is an improved parameter vector to describe X.

4. Result and discussion. This segment presents the experimental analysis of the suggested approach to Energy Efficient Anomaly Detection (EEAD) techniques. Here, it describes the Performance metrics, Case study, and Testing System.

4.1. Performance Metrics. In this section, the performance of the suggested technique under various operating conditions such as Detection Rate, True Positive Rate, Accuracy, and False Positive Rate.

4.1.1. Accuracy. The accuracy of all correctly predicted categories to the dataset's actual classifications represents the prediction algorithm's accuracy. Equation 4.1 determines the model's accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

4.1.2. True Positive Rate. The true positive result is one in which the model correctly predicts the positive outcome. Equation 4.2 determines the model's true positive rate.

$$\text{True Positive Rate(TPR)} = \frac{TP}{TP + FN} \tag{4.2}$$

4.1.3. False Positive Rate. In the positive class, the FPR measures the percentage of incorrect predictions. Equation 4.3 determines the model's False positive rate.

$$\text{False Positive Rate} = \frac{FP}{FP + TN} \tag{4.3}$$

4.1.4. Detective Rate. It is the proportion of true positive to all non-self-results discovered by the detecting array, here TP and FN are the totals for true positive and false negative samples, respectively. Equation 4.4 determines the model's Detection rate.

$$\text{Detection rate} = TPTP + FN \tag{4.4}$$

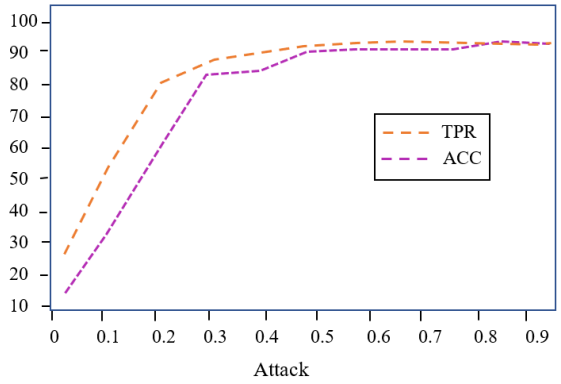

Fig. 4.1: Performance of Accuracy and True Positive Rate of Attack

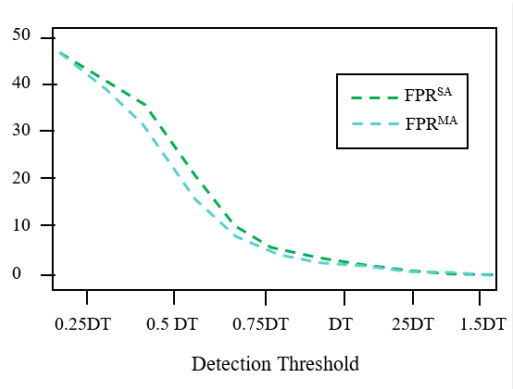

Fig. 4.2: Performance of False Positive rate of single and multi-attack

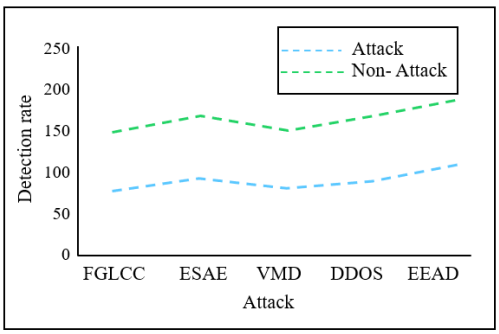

Fig. 4.3: Performance of Attack in Detection Rate

4.2. Case Studies. In this section, case research below distinct working situations is simulated to confirm the effectiveness of the suggested method. Case 1 is modelled as a physical network system with multiple agents based on the IEEE 118 bus model, each containing a generator. Figure 4.4 shows that stored energy represents energy that can be injected into the system from various smart grids. The attack strategy involves overloading

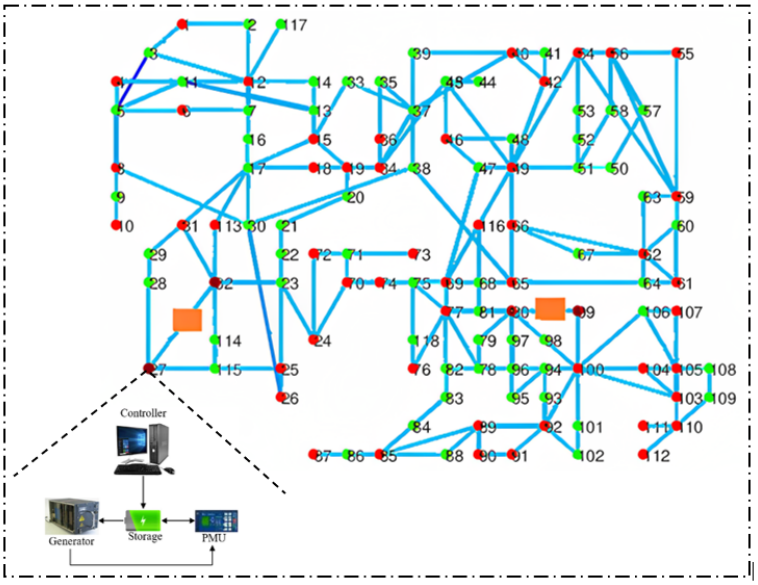

Fig. 4.4: IEEE 118 Bus System

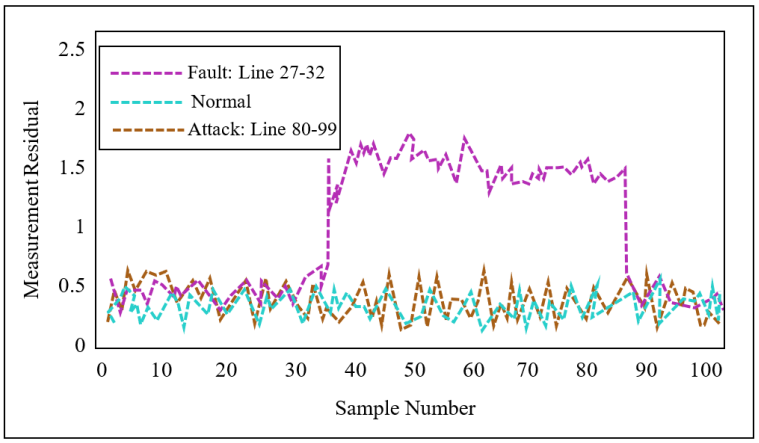

Fig. 4.5: Measurement residual and cyber-attack

lines 27-32 and 80-99. Figure 5 shows the attack area. Residues normalize under normal operating conditions due to errors and network attacks presented. It can be seen that all the measurement residues resulting from the cyberattacks have approximately the same amplitude as those measured under normal operating conditions, which implies that conventional residue testing can't detect stealthy cyber-attacks.

4.3. Testing System. A description of the case studies is provided by Matpower. The case studies are all assumed to be fully observable.

To ensure the accuracies of ancient information, the measurement model has been secured. Providing meter protection for large smart grids is very costly due to their thousands of meters. It recognizes crucial meters and safeguards them based on the most effective PMU placement to decrease expenses. It also imagines that a typical day won't bring any changes to the network topology.

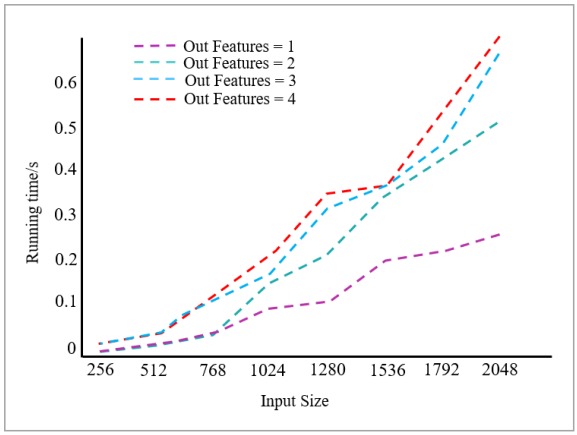

Fig. 4.6: Comparisons of the Running time

4.4. Computational complexity comparison. The computational complexity of the self-recognition module was tested using various parameters. Specifically, we compared complexity based on runtime and different input sizes by simply changing the k parameter. As shown in Figure 4.6, the execution time keeps changing, but the linear complexity can usually be maintained for different input sizes. This is mainly because in the branch I of the self-perception model, linear projection can be effectively used to obtain a k-dimensional sequence that can be computed in relation to latent nodes. So, the complexity can be greatly reduced to O(n).

5. Conclusion. In this section, a novel Energy Efficient Anomaly Detection (EEAD) technique is proposed, which uses HSDF pre-processing and HMM learning. A number of subsystems are initially created within the system. Hierarchical symbolic dynamic filtering (HSDF) converts time series data into symbol sequences and then learns the causal relationship between the nominal characteristics of subsystems. Then the converted sequences will be fed to the Hidden Markov model (HMM) which detects the anomaly by calculating the occurrence probability of the current observation based on the trained network. Simulation results on an IEEE 118 bus system to verify the performance of the suggested technique under various operating conditions such as False Positive Rate, Detection rate, Accuracy, and True Positive Rate.

Acknowledgments. This work was supported by China University Industry University Research Fund (2021BCB02005) Ministry of Education Vocational Education Reform and innovation funding (HBKC217034).

REFERENCES

[1] X. Wang, X. Luo, Y. Zhang and X. Guan, Detection and isolation of false data injection attacks in smart grids via nonlinear interval observer. IEEE Internet of Things Journal, 6(4), pp.6498-6512, 2019.
[2] J. Ding, A. Qammar, Z. Zhang, A. Karim, and H. Ning, Cyber Threats to Smart Grids: Review, Taxonomy, Potential Solutions, and Future Directions. Energies, 15(18), p.6799, 2022.
[3] H. Karimipour, A. Dehghantanha, R.M. Parizi, K.K.R. Choo, and H. Leung, A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids. IEEE Access, 7, pp.80778-80788, 2019.
[4] L. Haghnegahdar, and Y. Wang, A whale optimization algorithm-trained artificial neural network for smart grid cyber intrusion detection. Neural computing and applications, 32, pp.9427-9441, 2020.
[5] M.I. Oozeer, and S. Haykin, Cognitive dynamic system for control and cyber-attack detection in smart grid. IEEE Access, 7, pp.78320-78335, 2019.
[6] H. Karimipour, S. Geris, A. Dehghantanha, and H. Leung, Intelligent anomaly detection for large-scale smart grids. In 2019 IEEE Canadian conference of electrical and computer engineering (CCECE) (pp. 1-4). IEEE, 2019, May.
[7] J. Sakhnini, H. Karimipour, and A. Dehghantanha, Smart grid cyber-attacks detection using supervised learning and heuristic feature selection. In 2019 IEEE 7th international conference on smart energy grid engineering (SEGE) (pp. 108-112). IEEE. 2019, August.
[8] S. Geris, and H. Karimipour, Joint state estimation and cyber-attack detection based on feature grouping. In 2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE) (pp. 26-30). IEEE, 2019, August.

[9] M.R.C. Acosta, S. Ahmed, C.E. Garcia, and I. Koo, extremely randomized trees-based scheme for stealthy cyber-attack detection in smart grid networks. IEEE access, 8, pp.19921-19933, 2020.
[10] A. Al-Abassi, J. Sakhnini, and H. Karimipour, Unsupervised Stacked Autoencoders for Anomaly Detection on Smart Cyber-physical Grids. In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 3123-3129). IEEE, 2020, October.
[11] M.Z. Gunduz, and R. Das, Cyber-security on smart grid: Threats and potential solutions. Computer networks, 169, p.107094, 2020.
[12] C. Dou, D. Wu, D. Yue B. Jin, and S. Xu, A hybrid method for false data injection attack detection in smart grid based on variational mode decomposition and OS-ELM. CSEE Journal of Power and Energy Systems, 8(6), pp.1697-1707, 2020.
[13] H.N. Monday, J.P. Li, G.U. Nneji, A.Z. Yutra, B.D. Lemessa, S. Nahar, E.C. James, and A.U. Haq, The Capability of Wavelet Convolutional Neural Network for Detecting Cyber Attack of Distributed Denial of Service in Smart Grid. In 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP) (pp. 413-418). IEEE. 2021, December.
[14] J. Khazaei, and M.H. Amini, Protection of large-scale smart grids against false data injection cyberattacks leading to blackouts. International Journal of Critical Infrastructure Protection, 35, p.100457. 2021.
[15] S. Bahadoripour, E. MacDonald, and H. Karimipour, A Deep Multi-Modal Cyber-Attack Detection in Industrial Control Systems. arXiv preprint arXiv:2304.01440, 2023.

Edited by: Bradha Madhavan
Special issue on: High-performance Computing Algorithms for Material Sciences
Received: Jul 29, 2024
Accepted: Nov 6, 2024