Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 46 No. 08
August 2025

Deep Network Based Classifiers using Regularization Techniques for detecting cyber security threats in IOTs devices

Amani Ali Ahmed Ali[1,2], Abdulmalek Abduljabbar saeed Alqobaty[1], Talal Ahmed Ali[1]

Taiz University, Taiz, Yemen[1]

Al-Hikma University, Taiz, Yemen[2]

Abstract

Introduction: AI (Artificial intelligence) is now woven into nearly every facet of our lives, sparking innovation, and trans forming industries. Over the last few years, the ML (machine learning) and IoT (Internet of Things) have surged in prominence, becoming essential across numerous fields, ranging from manufacturing and transportation to agriculture and healthcare. Their applications span from fore casting sales to bolstering security measures and devices of monitoring. However, this growing focus on IoT technique has brought about important security concerns that must be addressed to ensure the reliability of these systems.

Objectives: Our study seeks to address a significant gap in the area by presenting a comprehensive strategy. By thoroughly exploring neural networks (NN) along with validating our findings against the MALNET-IMAGE, Virus-MNIST, Malimg and collection datasets, author aspires to push the boundaries of IoT threat detection. This unified framework, shaped by insights from the existing research we reviewed, is developed to provide exceptional resilience against the evolving landscape of threats.

Methods: A new model of deep network based on different classifiers and stacked through RBM strategy (Restricted Boltzmann Machine) for cybersecurity threat detection is suggested and investigated in this paper. Additionally, classifiers layer has been inserted which can classify the images. Authors then ensure the structurer of deep neural network (DNN) against over-fitting due to mighty dropconnect and dropout performance. Training with both techniques of regularization, with randomly chosen weights / activations subsets have been dropped.

Results: The evaluation over the datasets of MALNET-IMAGE, Virus-MNIST, Malimg and collection datasets to manage threat level displays a classification error rate development once utilizing deep network trained with dropconnect or dropout.

Conclusions: The merits of this strategy over traditional strategies are as follows: (1) experimental outcomes show that the proposed model exceeds traditional classification and other approaches of DNN; (2) the networks of the recommended architectures have deep construction and therefore the competence of extraction of feature is powerful than traditional classifiers.

Keywords: Intrusion Detection Systems, Deep learning, Dropconnect, Dropout, Feature extraction, Over-fitting.

1. Introduction

It is recently proved that the frameworks of detection /classification reliant over learning of DNN along with their derivatives such as DBN (Deep Belief Network), SAE (Stacking Auto-Encoder), CNN and RBM are appropriate accomplishment with precision in a few subjects specifically in threat detection [1]. In this part, author delves into the pivotal research surrounding IoT security, particularly high lighting the significance of NN. [2] offer methods of DL, relevant datasets, and comparison analysis with the target of identify breaches of cyber security. More precisely, authors analyze the current status of DL-based detection systems of intrusion.

authors investigated hybrid DL modules aimed at improving the precision of IoT detection of attack. Their findings support the principle of combining different ML mod ules to identify complex threat patterns effectively [3]. Similarly, authors conducted a detailed systematic review of ML methods in the domain of security of IoT. Their extensive study highlighted the diverse threats that IoT systems encounter, emphasizing the necessity for a strong defense strategy [4].

Researchers investigated the utilization of Recurrent Neural Networks (RNN) to identify threats of malware within IoT environments, further demonstrating the implications of DL strategies within this area [5].

Additionally, researchers took on a groundbreaking project to analyze interactions among devices of smart home, which resulted in a significant real-world dataset of IoT traffic [6]. Their work highlights the inherent vulnerabilities present in devices of IoT and emphasizes the urgent requirement for efficient detection strategies.

Researcher developed a DL ensemble particularly designed for detecting abnormal network and cyber-attacks. Their study emphasizes the possibility of ensemble techniques to boost the protection of IoT [7]. In contrast, researcher tackled the complex of identifying unknown attacks of security while assessing classifiers of ANN. Their results highlight the merits and disadvantages of each technique, encouraging a balanced application of these classifiers [8].

Researcher conducted an insightful study over IoT protection, utilizing the dataset of N-BaIoT to demonstrate the abilities of Spiking NN in classifying IoT malware [9]. Additionally, researcher introduced the innovative GANIBOT mod ule, a framework of semi-supervised Generative Adversarial Network (GAN), which represents a new advance within IoT anomaly detection botnets [10].

researcher introduced ADEPT, a system designed to determine and detect related stages of attacks, highlighting the complex nature of breaches of IoT protection [11]. In a complementary effort, researchers utilized the IoT23 dataset to investigate the diverse threats present within IoT infrastructures [12]. Their concentrate on the condition for standardized threats lays a vital groundwork for detection studies of future intrusion.

During the last several years, the progression of security within environments of IoT has seen remarkable developments. An increasing number of researchers are turning to ensemble learning as an efficient strategy to bolster the outcomes of detection systems of intrusion (IDs) [13–16]. Traditional security measures, like firewalls, have proven to be insufficient. The fusion of ML into IDs has led to the emergence of ensemble approaches which have significantly improved recall metrics along with precision [17]. There's been a noticeable uptick in research within intrusion detection research, especially with the embedding of ML strategies. For instance, authors conducted an assessment of detection of DDoS attack via leveraging machines learning over the CICIDS2017 dataset [18]. Mean while, researcher others explored DL approaches for detection of intrusion within IoT, while others performed a benchmarking of detection systems of intrusion utilizing techniques of Decision Tree (DT) [19,20]. Authors presented a productive module of data aimed at addressing the imbalance in IDS [21]. Additionally, researcher provided a thorough investigation of the dataset of CICIDS2017 [22]. Researcher investigated both methods of unsupervised and supervised learning using PyCaret with the dataset of CICIDS 2017 [23].

Within the cloud security domain, existing challenges have been highlighted, prompting the presentation of an ensemble IDS techniques. This strategy, which incorporates both classification and selection of feature, has established new standards in accuracy while simultaneously reducing rates of false alarm [24]. When considering Industrial IoT, a promising prototype of deep federated learning emerged. This ensemble-based classifier not only made the training approach more efficient but also outperformed traditional central strategies in relation to speed [16]. An approach of meta heuristic ensemble was formed for the primary concern of tackle the specific vulnerabilities of IoT networks, enhancing the classifier of RF achievement beyond its peers [14]. The merging of stacked ensemble learning and selection of feature has opened up new avenues, notably identifying a strong synergy between the random forest techniques (RF) and boosting machine of light gradient (LGBM) [15].

2. Objectives

The existing research over IoT protection is abundant, yet there's a significant gap among the findings that creatively employ NN in a holistic manner. A significant prior work tends to concentrate on either DL or ML, rarely blending the two in a meaningful way. This gap presents a unique opportunity for our study to deliver a significant contribution. Our study seeks to handle a significant research gap through presenting a comprehensive strategy. By thoroughly exploring NN and validating our findings against the MALNET-IMAGE, Virus-MNIST, Malimg and collection datasets, author aspires to push the boundaries of threat monitoring within IoT frameworks. This unified framework, shaped by observations from the existing studies we reviewed, is produced to provide exceptional resilience against the progressing threat scenario. Within IoT protection, an abundance of insights, methods, and conclusions have emerged that enhance our understanding. Several

of the leading impactful studies within the range have been gathered in this document, utilizing their insights to guide our study. As IoT environments advance, so will the associated threats; our aim is to stay ahead of these challenges, guaranteeing that IoT architectures remain resilient, secure, and reliable.

In the continuously developing field of network protection of IoT, researcher have seen a wave of significant developments that leverage the functionalities of different networks with CNNs. Researchers took advantage of CNN's skill in extracting spatial features alongside deep learning talent for detecting temporal patterns to generate a hybrid IDS (Intrusion Detection Systems). Their strategy, enhanced with techniques of dropout along with normalization layers of batch, demonstrated impressive precision and rates of detection across four different datasets, particularly excelling at reducing rates of false alarms.

In this manuscript, authors investigated Dropconnect and Dropout for proposed architecture of deep network used to the detection issue of cybersecurity threat. The two systems are just workable for fully-connected layers. Authors evaluated the studies of experimental on famous known datasets. Figure 1 shows our proposed architecture involving the steps of pre-processing, and afterward an automatic extractor of features utilizing techniques of Dropconnect or Dropout.

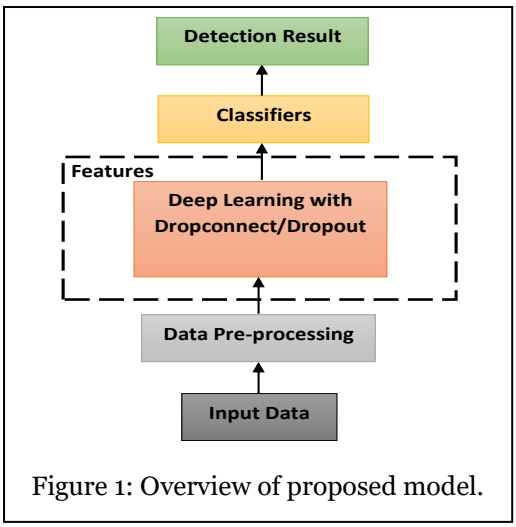

Figure 1: Overview of proposed model.

3. Methods

The suggested structure presents the composition of deep network model and RBM, and further presents the deep network classification structure built upon classifiers fusion.

The raw materials for both influence of the ultimate model and training are features. Theoretically, the more availability of hidden layers, the more features the DNN can extract, and therefore the more difficult the function learned. In like manner, the model of DNN can be displayed in detail. Nonetheless, the DNN has been gradually exchanged by a module of shallow learning, like boosting and SVM, because of issues that happen when the *W* weights are initialized with a random number during a network of multi-layer. If the weights are assigned to be excessively large, at that point the procedure of training will end in the local optimum. At the point when the weights are assigned excessively small, dispersion of gradient will happen, also the weights alter gradually owing to the small gradient, getting the optimal solution is challenging too. To tackle these issues, an initialization of layer-by-layer of the DNNs can get initial weights which are near into the optimal solution [25]. Initialization of layer-by-layer is gained by unsupervised learning which may be performed automatically. This work suggests a new technique of classifying threats. The basic deep network model comprises of pre-processing, pre-training and fine-tuning. The pre-training introduction is the difference among NN and deep network.

Proposed deep network is a probabilistic generative design comprised of single layer of observable neurons and numerous layers of stochastic hidden variables. Proposed deep network is opposite the model of traditional discriminative which has been stacked through RBM. This structure can possibly be developed using effective method by layer-by-layer of greedily learning [26], like a RBM, to configure the deep network. The supervised learning of greedy layer-wise had higher detection precision rate than does a NN. The consequence of the prior layer passed into the subsequent layer. To learn in the networks, two stages were essential which are: an unsupervised learning of feature whilst the second was a supervised classifiers learning. All RBM layers are unsupervised trained. Into various spaces of features, the input must be mapped. The information must be preserved as much as potential. The classifiers layer has been inserted as the highest layer within the recommended architecture as a supervised classifier. The suggested deep network learning had been retrieved from the brain structure of human. All deep network layers act as a model of LR (logistic regression).

The suggested structure input data involve the three-dimensional (3D) vector got during stage of pre-processing. Pre-processing stage involves binarization, resizing and median filtering. The layer of RBM had been trained one by one within stage of pre-training. The duplicate of hidden variables within the prior layer is the succeeding visible variable. In manner of layer-wise, the parameters have been transferred. Also from the prior layer, the features have been learned. The classifiers within the highest layer have undergone training through fine-tuning, where the function of cost has been revised by back propagation for optimizing the weights [27]. The configuration of proposed DNN is exposed within Figure 2.

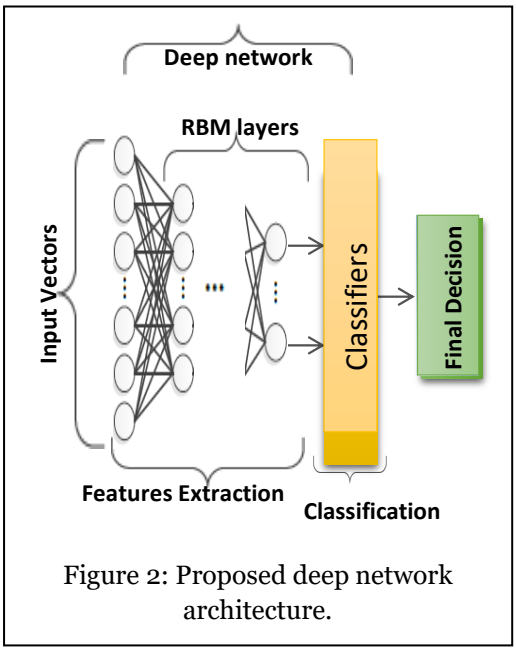

Figure 2: Proposed deep network architecture.

RBM

RBM [28] are NNs of generative stochastic among an assortment of hidden neurons $H= \{h_1, \ldots, h_{n-1}, h_n\}$, and an assortment of visible neurons $V = \{v_1, \ldots, v_{m-1}, v_m\}$. The units of hidden and visible have been connected in conjunction with a matrix of *W* weight. The framework of model has been restricted by not permitting connections of intra-layer among the units. An energy function has been presented to figure out the RBM state that has been upgraded from the Hopfield network energy function within a structure of nonlinear dynamic. So, the system objective function is transformed into an extreme value issue and the sturcture of RBM can be simply analyzed [29]. Formally, the function of energy and the RBM probabilistic semantics have been calculated below:

$$E(V,H) = -\sum_{j=1}\sum_{i=1} V_j H_i W_{ji} - \sum_{j=1} V_j x_j - \sum_{i=1} H_i y_i \quad (1)$$

$$P(V) = \frac{1}{Z}\sum_{H} e^{-E(V,H)} \quad (2)$$

Here the connection among hidden neurons H_i and *Vi* visible neurons indicated by W_{ji} , their biased indicated by y_i and x_i, respectively, and the splitting function indicated by Z.

Regularization Techniques

Over-fitting is capability occur owing to the big parameter count within the recommended architecture of networks along with consequently dropconnect or dropout (refer to figure 3) has been practiced so that authors can prevent the proposed system from issue of over-fitting and boost its

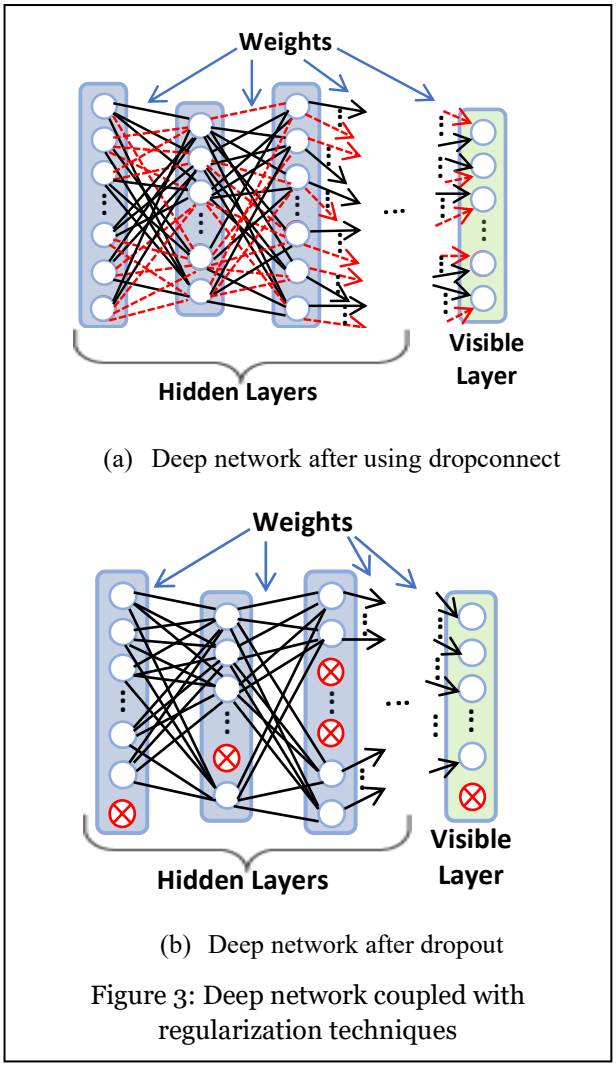

Figure 3: Deep network coupled with regularization techniques

effectiveness. Dropout resembles dropconnect, but used to the *W* weights.

The approach of dropout, fundamentally suggested in [30] as a regularization type for fully-connected layers, has been triumphal used with a few kinds of DNN [31]

and it demonstrates its achievement during multiple missions of detection [32]. In circumstances of training with Dropout, all elements of a layer's result are left with *p* probability, otherwise has been adjusted to zero along with *1-p* probability.

Dropconnect [33] that is a dropout generalization sets rather an arbitrarily chosen weights subset inside the network into zero, such that dropout mainly sets or drops to zero complete columns or rows of the matrix of *W* weight. Oppositely, dropconnect is more well drops and grained matrix elements of individual weight.

Experiments, Outcomes and Discussion

The recommended detection-oriented framework for cyber intrusions in IOTs devices based on fusion classifier has been offered in this part. The approach of dropconnect and dropout was applied through training over the deep network which has been conducted on MALNET-IMAGE, Virus-MNIST, Malimg and collection datasets. Therefore, authors are proficient at assessment the regularization techniques effectiveness separately for DNN contrasted with the prior studies.

Experiments on Datasets

The operation on the MALNET-IMAGE, Virus-MNIST, Malimg and collection datasets were investigated to cover every shape of threats in IOT, so that the measurement of the operational performance of unsupervised feature learning method utilizing deep network coupled with dropconnect/dropout. Due to Intrusion position inside IOT devices, all threats can have numerous forms. Prior inputting into the proposed models, the pixel values into the span of [0, 1] has been scaled. In Table 1, the datasets details have been presented.

The MALNET-IMAGE comprises of 1,262,024 samples. Each sample within the 363 images of test per class and 1,451 images of training per class has been normalized by 28 × 28 size of pixels. In addition, study in [34] can be referred for more findings concerning the classes for every MALNET-IMAGE shapes. The Virus-MNIST threats classification mission comprises of white and black images, shaping 10 classes. Each sample within the 1,038 images of test and 4,150 images of training has been normalized by 28 × 28 size of pixels. In addition, study in [35] can be referred for more findings concerning the classes for every Virus-MNIST shapes. The Malimg comprises of 9,458 samples. Each sample within the 75 images of test and 302 images of training has been normalized by 28 × 28 size of pixels. In addition, study in [36] can be referred for more insights regarding the classes for every Malimg shapes and structure. The collection dataset comprises of 9978 samples. Each sample within the 80 images of test per class and 319 images of training per class has been normalized by 28 × 28 size of pixels. In addition, study can be referred for more information pertaining to the classes for every collection dataset shapes.

Table 1: Datasets details

Datasets	Dataset distribution	IDR (Intrusion Detection Rate)/Forms	Total of shapes (classes)	Total per class
MALNET-IMAGE	Training	1,009,619	696	1,451
	Testing	252,405		363
Virus-MNIST	Training	41,504	10	4,150
	Testing	10,376		1,038
Malimg	Training	7,566	25	302
	Testing	1,892		75
collection dataset	Training	7,982	25	319
	Testing	1,996		80

Experiments Setting

Our best robust features extractor network was chosen for the main concern of boost our outcome of classification. This extractor of feature is formed by four layers RBM along with 1000 hidden units within all layers. The adopted deep network architecture applied in experiment over Malimg and collection datasets were 784-1000- 1000- 1000-1000-25, such that, it forms a network coupled with pixels size of 28 × 28 as input images giving 784 input dimensionality with four hidden layers. The end layer which is the top layer within the suggested models is involved by 25 units (classes) giving the final finding of the structure. The found error rate on Malimg and collection dataset were 1.82% and 1.95% respectively. The same with MALNET-IMAGE dataset was 784-1000- 1000- 1000-1000-696. The end layer which is the top layer within the suggested models is involved by 696 units (classes) giving the final outcome of the systems. The found error rate was 2.34%. The same with Virus-MNIST dataset was 784-1000- 1000- 1000-1000-10. The end layer which is the top layer within the suggested models is involved by 10 units (classes) giving the final outcome of the systems. The found error rate was 1.54%. Two regularization strategies have been executed for structure of deep network named dropconnect and dropout with the target of optimize this design. In the experiment, dropconnect/dropout has been applied

separately within every proposed network layers. In both strategies authors tentatively removed 45% and 25% from hidden layers and apparent layer respectively. These weights/units have been arbitrarily selected only in the stage of training. The comparison between dropconnect and dropout with No-Drop has been accomplished in the part that follows.

4. Results and Discussion

The rate of error of classification gets via the DNN trained with no units of dropping (No-Drop) over the Malimg dataset rising to 1.82% using the group of tests is efficient whenever contrasted with prior results presented within the literature. Appling approach of dropout, authors get 0.9%, relative development in error rate of sample which is dropped by 0.92%. While used dropconnect to the suggested deep network model, in hidden and apparent layers, decreases the frequency of error to 0.43% accomplishing a gain of 0.47% on the approach of dropout. The rate of error of classification gets with the DNN trained along with no units of dropping (No-Drop) over the collection dataset rising to 1.95% using the group of tests is efficient whenever contrasted with prior results presented within the literature. Appling approach of dropout, authors get 1.03%, relative development in error rate of sample which is dropped by 0.92%. While used dropconnect to the suggested deep network model, in hidden and apparent layers, decreases the frequency of error to 0.56% accomplishing a gain of 0.47% on the approach of dropout. The rate of error of classification gets with the DNN trained along with no units of dropping (No-Drop) over the MALNET-IMAGE dataset rising to 2.34% using the group of tests is efficient whenever contrasted with prior results presented within the literature. Appling strategy of dropout, authors get 1.42%, relative development in error rate of threat which is dropped by 0.92%. While used dropconnect to the suggested deep network model, in hidden and apparent layers, decreases the incidence of error to 0.95% accomplishing a gain of 0.47% on the approach of dropout. The rate of error of classification gets with the DNN trained along with no units of dropping (No-Drop) over the Virus-MNIST dataset rising to 1.54% using the group of tests is efficient whenever contrasted with prior results presented within the literature. Appling strategy of dropout, authors get 0.89%, relative development in error rate of threat which is dropped by 0.46%. While used dropconnect to the suggested deep network model, in hidden and apparent layers, decreases the incidence of error to 0.95% accomplishing a gain of 0.11% on the approach of dropout.

It is noticed that dropconnect mainly shows a better job than dropout during this detection mission of intrusion within IOT environment. Therefore, dropconnect blocked the overfitting of fully-connected layer more proficiently than the others during this experiment. The two strategies develop on No-Drop during these threats datasets of MALNET-IMAGE, Virus-MNIST, Malimg and collection datasets were displayed within Table 2. Compared with accuracies rate of sample detection acquired from state-of-the-art, those rates were statistically noticeably paramount. A comparative examination of the proficiency of proposed architecture with different techniques utilizing cybersecurity threats datasets MALNET-IMAGE, Virus-MNIST, Malimg and collection were also discussed. According to Table 2, this has been demonstrated that the recommended structure with dropconnect and dropout outperforms the other common models of various techniques with both the omission also the event of dropout once experimented over MALNET-IMAGE, Virus-MNIST, Malimg and collection datasets.

Table 2. Performance contrast with the used datasets

Authors	Methods	Detection Rate (%)
[37]	DA	88
[37]	RBM	53
[37]	DBM	74
[37]	DBN	63
	Proposed model	99

5. Conclusion

In this manuscript, a novel deep network structure built on classifiers has been recommended for classifying and detection intrusion in IOTs devices. Two well-known techniques of applying regularization to deep network have been exploited and studied to block overfitting. However, this issue as yet exists as a challenge to master once it transacts with working in areas offering a little data quantity or training extremely enormous neural networks. After unsupervised and supervised pre-training, the model of deep network could learn successfully the features. So, authors inquired into a learning of feature concentrated on method by the DNN module for detection problem of cybersecurity threats in IOTs

devices where the approach of dropconnect and dropout have been utilized throughout the training stage using the goal that authors could block the suggested model from over-fitting. A layer of classifiers is added on top for classifying the threat in images. That was demonstrated that for the threat level one on MALNET-IMAGE, Virus-MNIST, Malimg and collection datasets, the findings have been favorable with an error rate of classification utilizing dropconnect and dropout. Ultimately, authors concluded that applying techniques of regularization during a deep network had the option to better significantly the finding of threats than No-Drop. Regarding to the suggested structure outcomes, authors suggested that the framework of deep network be designed with four hidden layers, with all having 1000 hidden units. The same model might be employed also on other security. The same model might be employed too on other application like face recognition, recognition of speech and others which its requirement is incrementing with the system software availability within applications.

Future work, authors will develop the model of deep network regarding its time consumption and accuracy. Authors will keep developing the recommended structure with the recent accomplishments in the part of dimensionality reduction method.

Refrences

[1] Hinton, G. E. (2012). A practical guide to training restricted Boltzmann machines. In Neural networks: Tricks of the trade 7700 (pp. 599-619). Springer, Berlin, Heidelberg.

[2] Barik, K., Misra, S., Konar, K., Fernandez-Sanz, L., & Koyuncu, M. (2022). Cybersecurity deep: approaches, attacks dataset, and compara tive study. Applied Artificial Intelligence, 36(1), 2055399.

[3] Sahu, A. K., Sharma, S., Tanveer, M., & Raja, R. (2021). Internet of Things attack detection using hybrid Deep Learning Model. Computer Communications, 176, 146-154.

[4] Ahmad, R., & Alsmadi, I. (2021). Machine learning approaches to IoT security: A systematic literature review. Internet of Things, 14, 100365.

[5] Woźniak, M., Siłka, J., Wieczorek, M., & Alrashoud, M. (2020). Recurrent neural network model for IoT and networking malware threat detection. IEEE Transactions on Industrial Informatics, 17(8), 5583 5594.

[6] Anagnostopoulos, M., Spathoulas, G., Viaño, B., & Augusto Gonzalez, J. (2020). Tracing your smart-home devices conversations: Areal world IoT traffic data-set. Sensors, 20(22), 6600.

[7] Dutta, V., Choraś, M., Pawlicki, M., & Kozik, R. (2020). A deep learning ensemble for network anomaly and cyber-attack detection. Sensors, 20(16), 4583.

[8] Al-Zewairi, M., Almajali, S., & Ayyash, M. (2020). Unknown security attack detection using shallow and deep ANN classifiers. Electronics, 9(12), 2006.

[9] Umair, M., Tan, W. H., & Foo, Y. L. (2023, July). Efficient Malware Classification with Spiking Neural Networks: A Case Study on N-BaIoT Dataset. In 2023Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN) (pp. 231-236). IEEE.

[10] Saurabh, K., Singh, A., Singh, U., Vyas, O. P., & Khondoker, R. (2022, August). Ganibot: A network flow based semi supervised generative adversarial networks model for iot botnets detection. In 2022 IEEE International Conference on Omni-layer Intelligent Systems (COINS) (pp. 1-5). IEEE.

[11] Sudheera, K. L. K., Divakaran, D. M., Singh, R. P., & Gurusamy, M. (2021). ADEPT: Detection and identification of correlated attack stages in IoT networks. IEEE Internet of Things Journal, 8(8), 6591-6607.

[12] Booij, T. M., Chiscop, I., Meeuwissen, E., Moustafa, N., & Den Hartog, F. T. (2021). ToN_IoT: The role of heterogeneity and the need for standardization of features and attack types in IoT network intrusion data sets. IEEE Internet of Things Journal, 9(1), 485-496.

[13] Turukmane, A. V. (2023). Forecasting the IoT-based cyber threats using the hybrid forage dependent ensemble classifier. Concurrency and Computation: Practice and Experience, 35(2), e7460.

[14] Dey, A. K., Gupta, G. P., & Sahu, S. P. (2023). A metaheuristic-based ensemble feature selection framework for cyber threat detection in IoT enabled networks. Decision Analytics Journal, 7, 100206.

[15] Mishra, A. K., & Paliwal, S. (2023). Mitigating cyber threats through integration of feature selection and stacking ensemble learning: the LGBM and random forest intrusion detection perspective. Cluster Computing, 26(4), 2339-2350.

[16] Jahromi, A. N., Karimipour, H., & Dehghantanha, A. (2023). An ensemble deep federated learning cyber-threat hunting model for Industrial Internet of Things. Computer Communications, 198, 108-116.

[17] Mahfouz,A., Abuhussein, A., Venugopal, D., & Shiva, S. (2020). Ensemble classifiers for network intrusion detection using a novel network attack dataset. Future Internet, 12(11), 180.

[18] Ahmed, A. S., Kurnaz, S., & Khaleel, A. M. (2023). Evaluation DDoS Attack Detection Through the Application of Machine Learning Techniques on the CICIDS2017 Dataset in the Field of Information Security. Mathematical Modelling of Engineering Problems, 10(4).

[19] Jose, J., & Jose, D. V. (2023). Deep learning algorithms for intrusion detection systems in internet of things using CIC-IDS 2017 dataset. International Journal of Electrical and Computer Engineering (IJECE), 13(1), 1134-1141.

[20] Azalmad, M., El Ayachi, R., & Biniz, M. (2023, July). Unveiling the Performance Insights: Benchmarking Anomaly-Based Intrusion Detection Systems Using Decision Tree Family Algorithms on the CICIDS2017 Dataset. In International Conference on Business Intelligence (pp. 202-219). Cham: Springer Nature Switzerland.

[21] Barkah, A. S., Selamat, S. R., Abidin, Z. Z., & Wahyudi, R. (2023). Data Generative Model to Detect the Anomalies for IDS Imbalance CICIDS2017 Dataset. TEM Journal, 12(1).

[22] Oyelakin, A., Ameen, A. O., Ogundele, T. S., Salau-Ibrahim, T., Abdulrauf, U. T., Olufadi, H. I., ... & Muhammad-Thani, S. (2023). Overview and exploratory analyses of CICIDS 2017 intrusion detection dataset. Journal of Systems Engineering and Information Technology (JOSEIT), 2(2), 45-52.

[23] Krsteski, S., Tashkovska, M., Sazdov, B., Radojichikj, L., Cholakoska, A., & Efnusheva, D. (2023, April). Intrusion detection with supervised and unsupervised learning using pycaret over cicids 2017 dataset. In Computer Science On-line Conference (pp. 125-132). Cham: Springer International Publishing.

[24] Krishnaveni, S., Sivamohan, S., Sridhar, S. S., & Prabakaran, S. (2021). Efficient feature selection and classification through ensemble method for network intrusion detection on cloud computing. Cluster Computing, 24(3), 1761-1779.

[25] Bengio, Y., & Delalleau, O. (2011). On the expressive power of deep architectures. International Conference on Algorithmic Learning Theory (pp. 18-36). Springer, Berlin, Heidel berg.

[26] Bengio, Y., Lamblin, P., Popovici, D., & Larochelle, H. (2007). Greedy layer-wise training of deep networks. In Advances in neural information processing systems (pp. 153 160).

[27] Hecht-Nielsen, R. (2002). Theory of the backpropagation neural network. International Joint Conference on Neural Networks. vol. I, (pp. 593-605). IEEE.

[28] Mohamed, A. R., Sainath, T. N., Dahl, G., Ramabhad ran, B., Hinton, G. E., & Picheny, M. A. (2011). Deep belief networks using discriminative features for phone recognition. International conference on acoustics, speech and signal pro cessing (ICASSP) (pp. 5060-5063). IEEE.

[29] Hinton, G. E. (2002). Training products of experts by minimizing contrastive divergence. Neural computation, 14(8), 1771-1800.

[30] Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., &Salakhutdinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580.

[31] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to pre vent neural networks from overfitting. The journal of machine learning research, 15(1), 1929-1958.

[32] Zhang, S., Bao, Y., Zhou, P., Jiang, H., & Dai, L. (2014). Improving deep neural networks for LVCSR using dropout and shrinking structure. International Conference on Acous tics, Speech and Signal Processing (ICASSP) (pp. 6849-6853). IEEE.

[33] Wan, L., Zeiler, M., Zhang, S., Le Cun, Y., & Fergus, R. (2013). Regularization of neural networks using dropconnect. International conference on machine learning 28 (pp. 1058 1066).

[34] Freitas, S., Duggal, R., & Chau, D. H. (2022, October). MalNet: A large-scale image database of malicious software. In Proceedings of the 31st ACM International Conference on Information Knowledge Management (pp. 3948-3952).

[35] Noever, D., & Noever, S. E. M. (2021). Virus-MNIST: A benchmark malware dataset. arXiv preprint arXiv:2103.00602.

[36] Nataraj, L., Yegneswaran, V., Porras, P., & Zhang, J. (2011, Octo ber). A comparative assessment of malware classification using binary texture analysis and dynamic analysis. In Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence (pp. 21-30).

[37] Ferrag, M.A.,Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. Journal of Information Security and Applications, 50, 102419.