International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 1, Jan-Feb-2022, ISSN (Online): 2395-566X

IOT Network Malicious Session Detection by KNN and MothFlame Optimization Algorithm

M.Tech.Scholar Rakhi Sarathe, Asst. Prof. Sumit Sharma
Dept. of CSE
Vaishnavi Institute of Technology, Bhopal, MP

Abstract- IOT network increases the comfort of human life in various measures. This growing network attract many intruders to attack on the system hence security of IOT devices is major concern these days. This paper has proposed a IOT network intrusion detectin system that detect the session into attack and normal class. Moth flame optimization genetic algorithm was used in the work for slection of features for identifying the class representative sessions. Identification of class session was done by K-Nearest Neighbour. Expeirment was done on real dataset and result shows that proposed model Moth Flame based IOT Network Security (MFIOTNS) has improved the work efficiency optimizing various evalautin parameter values.

Index Terms- KNN, Clustering, Genetic Algorithm, Intrusion Detection.

I. INTRODUCTION

Concerns over security and privacy regarding computer networks are increasing in the world, and computer security has become a requirement as a result of the spread of information technology in daily life. The raise in the amount of Internet applications and the appearance of modern technologies such as the Internet of Things (IoT) are followed with new and recent efforts to invade computer networks and systems. The Internet of Things (IoT) is a set of interrelated devices where the devices have the ability to connect without the need for human intervention. With IoT, many things that have sensors (such as coffee makers, lights, bicycles, and many others) in areas like healthcare, farming, transportation, etc. can connect to the Internet [1]. By saving time and resources, IoT applications are changing our work and lives. It also has unlimited advantages and opens numerous opportunities for the exchange of knowledge, innovation, and growth. Every security threat within the Internet exists within the IoT as well because the Internet is the core and center of the IoT. Compared to other traditional networks, IoT nodes have low capacity and limited resources, and do not have manual controls. Also, the rapid growth and broad dailylife adoption of IoT devices makes IoT security issues very troublesome, raising the need to develop security solutions based on networks. While current systems perform well in identifying some attacks, it is still challenging to detect others.

As network attacks grow, along with a massive increase in the amount of information present in networks, faster and more effective methods of detection of attacks are required [2] and there is no doubt that there is scope for more progressive methods to improve network security. In this context, in order to provide embedded intelligence in the IoT environment, we can consider Machine Learning (ML) as one of the most effective computational models. Machine learning approaches have been used for different network security tasks such as network traffic analysis [3],[4],[5], intrusion detection[6], and botnet detection [7].

Machine Learning can be described as an intelligent device's ability to modify or automate a knowledge-based state or behavior, which is considered a critical part of an IoT solution. ML has the ability to infer helpful knowledge from data generated by devices or humans, and ML algorithms are used in tasks such as regression, and classification. Likewise, in an IoT network, ML can be used to provide security services. The use of machine learning in attack detection problems is becoming a hotly pursued subject, and ML is being used more and more in different applications in the cybersecurity field. Although many studies in the literature have used ML techniques to discover the best ways to detect attacks, only limited research exists on efficient detection methods suitable for IoT environments. Machine learning can be applied to the attack detection task via two main types of cyber-analysis: signature-based (sometimes also called misuse- based) or anomaly-based. Signaturebased techniques are designed to detect known attacks by using specific traffic characteristics (also known as "signatures") in those attacks. One of the advantages of this class of detection technique is its ability to detect all known attacks effectively without generating an overwhelming number of false alarms.

II. RELATED WORK

The authors in [13] propose an intrusion detection model based on a genetic algorithm and a deep belief network. They use the NSL-KDD dataset for detecting four types of attacks: DoS, R2L, Probe and U2R. This paper, in comparison with our work, uses an old dataset difficult to be applicable to modern IoT networks and does not

implement blockchain in their solution as an integrated mechanism for monitoring and securing IIoT networks.

In [14], an intrusion detection technique based on statistical flow features is proposed for protecting the network traffic of Internet of Things applications. The authors in this work use three machine learning techniques to detect malicious traffic events: Decision Tree, Naive Bayes and Artificial Neural Network (ANN). They use the same dataset employed by us, the UNSWNB15 dataset; however, they do not implement blockchain in their solution as an integrated mechanism for monitoring and securing IIoT networks.

A machine learning security framework for IoT systems is proposed in [15]. They built a dataset based on the NSL-KDD dataset and evaluated their proposal in a real smart building scenario. As we said in the previous related works, an old dataset may not be suitable for modern IoT networks. They use one-class SVM (Support Vector Machine) technique for detecting four types of attacks: DDoS, Probe, U2R and R2L. However, they do not use a blockchain approach for supervising IIoT networks.

The authors in [16] developed an algorithm for detecting denial-of-service (DoS) attacks using a deep-learning algorithm. They use three approaches for detecting DoS attacks: Random Forests, a Multilayer Perceptron and a Convolutional Neural Network. They use the same dataset employed by us, but they just aim to detect one attack (DoS) and do not integrate blockchain in their solution.

The authors in [20] propose a model using a machine learning algorithm to detect and mitigate botnet-based distributed denial of service (DDoS) attacks in IoT networks. The use different machine learning algorithms such as K- Nearest Neighbour (KNN), Naive Bayes model and Multi-layer Perception Artificial Neural Network (MLP ANN). They use the same dataset employed by us, but they just aim to detect one attack (DoS) and do not integrate blockchain in their solution.

In [21], the authors propose an intrusion and cyber attacks traffic identification model using Machine Learning (ML) algorithms for IoT security analysis. The authors in this work use four machine learning techniques to detect malicious traffic events: Random Forest, Random Tree, Decision Tree, Naive Bayes and BayesNet. They use the same dataset employed by us, but they do not integrate blockchain in their solution.

III.METHODOLOGY

This section provide brief of proposed Moth Flame based IOT Network Security (**MFIOTNS)**. Fig. 1 is block diagram of proposed model include dataset processing, dimension reduction, training blocks. Explanation of each block was in this section under different headings.

1. Dataset Cleaning-Input data has various features and each has its own importance, this step clean data by removing the unwanted information from the set. Such as input dataset used in this work has n fields out of those first few feature values were removed from the work, as session ID, connection type, transferring protocol, etc. information in dataset can be remove.

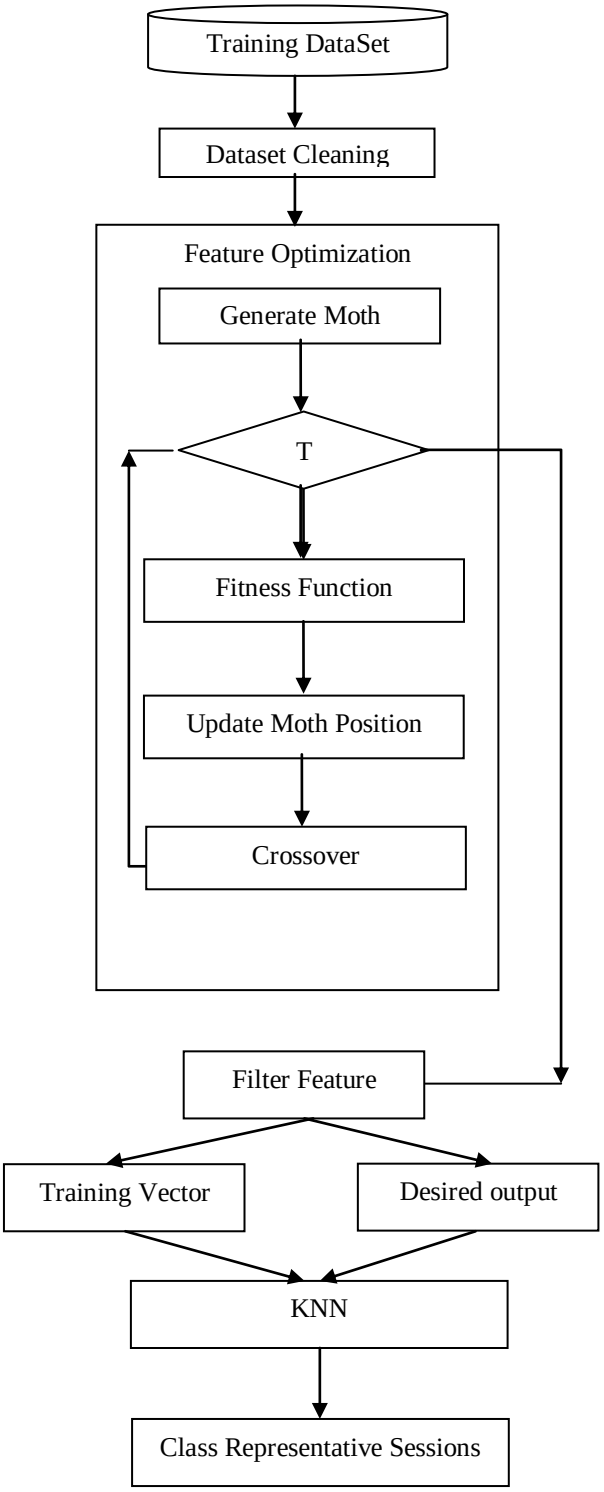

Fig. 1 Block diagram of MFOCMSD network intrusion detection.

CD←Dataset_Cleaning(RD) ----------Eq. 1

In eq. 1 RD is raw dataset and CD is clean dataset matrix. Processed dataset was arranged in matrix of row and column where each row is session and columns are feature set of a session.

2.Feature optimization
Input CD matrix was further process by Moth Flame Optimization Algorithm to reduce values of training vector and increases learning accuracy.

3.Moth Flame Optimization Algorithm: In this algorithm paper has consider each chromosome as a Moth. Objective of this algorithm was to find a Moth Flame belonging to path towards moon. Moth Flame are chromosome of this work.

4.Generate Moth Flames
Moth Flames are group of chromosome and a chromosome is possible solution of optimized feature set. So a Moth Flame is a vector of n number of elements, where n is number of column in CD. Each element is a binary value in Moth Flame vector. One shows that a feature is consider for the training and zero shows that feature is not selected for the population. So if p number of Moth Flames generate then M is Moth Flame population matrix having pxn dimension. Selection of f number of feature in vector done by random value generator function Gaussian.

M←Generate_Moth Flame(p, n, f)----------Eq. 2

5.Fitness Function
Each Moth Flame were rank as per distance. So evaluation of distance done by fitness value. Moth Flame feature vector pass training vector to the KNN (K-Nearest Neighbour) for cluster representative finding and measure the detection accuracy of the work [11]. This detection accuracy value is distance parameter in the work.

Input: M, CD
Output: F
1. Loop w=1:W // for w Moth Flames
2. Loop s=1:CD // for s training session
3. TV[s]←Training_Vector(W[w], CD[s])
4. DO[s]←Desired_Output(W[w], CD[s])
5. EndLoop
6. TNN←Train_Neural_Network(TV, DO)
7. Loop s=1:CD // for s training session
8. TV←Training_Vector(W[w], CD[s])
9. O←Predict(TV,TNN)
10. If DO[s] equals O
11. F[w]←Increment F by 1
12. EndIf
13. EndLoop
14. Endloop

In above algorithm TV is training vector, DO is desired output.

6. Update Moth Flame position
Once F value obtain by fitness function then sort f in deseding order and find best Moth Flame out of all chromosomes available in the population.

7.Crossover
Genetic algorithm success depends on change of chromosomes, hence as per changing parameter X, number of random position value of Moth Flames were modified. This operation was not done in best local Moth Flame. In this step each Moth Flame X number of positions were modified randomly from zero to one or one to zero as per best local Moth Flame feature set. These Moth Flame were further test for path distance and compared its fitness value with parent Moth Flame if child Moth Flame has better values then remove parent otherwise parent will continue. After this step if maximum ieration steps occur then jump to filter feature block otherwise evaluate fitness value of each Moth Flame Moth Flame.

8. Filter Feature
Once iteration get complete then find best Moth Flame from the last updated population. Feature having value one in chromosome consider as selected feature for training vector and other consider as unselected. Desired output matrix was also prepared in this section.

KNN Based Cluster Representativ

Feature set obtained form above algorithm were used to find the cluster representative by using KNN model. Finding of such representative help to identify the session class as by the distance vector from the representative efeature set.

IV.EXPERIMENT AND RESULTS

Experimental setup: MFOCMSD and comparing model was developed on MATLAB software. Experimental machine having 4 GB ram, i3 6th generation processor. IO dataset was taken from [15]. Comparison of **MFIOTNS** was done with cloud malicious session detection model proposed in [16].

1. Evaluation Parameter
To test our results, this work usesthe following measures Precision, Recall, and F-score. These parameters are dependent on the TP (True Positive), TN True Negative), FP (False Positive), and FN (False Negative).

V.RESULTS

Table 1. Precision value based comparison of IOT network intrusion detection models.

Dataset Size	Previsous Work	MFIOTNS
5000	0.9359	0.987
10000	0.9322	0.9814
15000	0.9312	0.9812
20000	0.9322	0.9806
25000	0.9323	0.9793

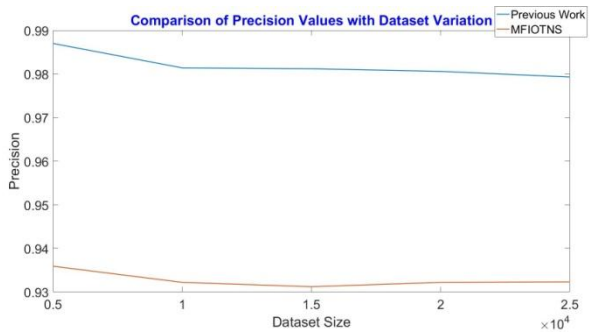

Fig. 2 Precision value based comparison.

IOT network intrusion detection models were compared on different dataset size and result shows that prospeod model has improved the precision value by 5.004% as compared to previous model proposed in [16]. It was found that that proposed modle has increases the precison value by use of moth flame feature optimization technique, as less feature has improved the clustering of KNN model.

Table 2. Recall value based comparison of IOT network intrusion detection models.

Dataset Size	Previsous Work	MFIOTNS
5000	0.8623	0.9862
10000	0.8568	0.9838
15000	0.8582	0.9825
20000	0.8586	0.9816
25000	0.8606	0.9816

Recall value parameters were compared in table 2. It was obtained that proposed model has improved the IOT intrusion detection recall parameter by % as compared to values obtained form the previous model in [16]. KNN based learning of selected feature has increases the detection recall.

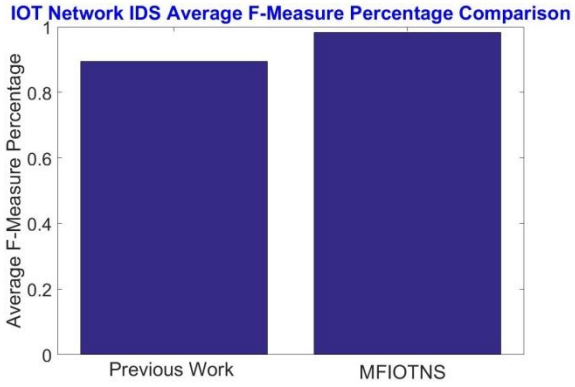

Fig. 3 F-measure value based comparison.

Table 3. F-Measure value based comparison of IOT network intrusion detection models.

Dataset Size	Previsous Work	MFIOTNS
5000	0.8976	0.9866
10000	0.8929	0.9826
15000	0.8932	0.9819
20000	0.8939	0.9811
25000	0.895	0.9805

Inverse average of precision and recall value is f-mesure parameter. Table 3 shows that use of moth flame optimization genetic algorithm for feature selection has enhanced the f-measure values of IOT network intrusion detection.

Table 4. Accuracy value based comparison of IOT network intrusion detection models.

Dataset Size	Previsous Work	MFIOTNS
5000	0.8148	0.9748
10000	0.8074	0.9673
15000	0.8079	0.966
20000	0.8090	0.9646
25000	0.8109	0.9634

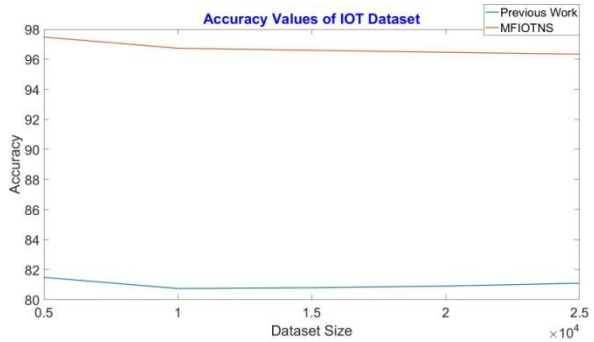

Fig. 4 Average accuracy value based comparison.

IOT network intrusion detection models were compared on different dataset size and result shows that prospeod model has improved the accuracy value by 6.25% as compared to previous model proposed in [16]. It was found that that proposed model has increases the accuracy value by use of moth flame feature optimization technique, as less feature has improved the clustering of KNN model.

VI.CONCLUSION

IOT networks are a boon for small industries, hotels, organizations, etc. But sue lack of securities measures it's a vulnerable for different type of attacks. This paper has developed a model that support such network for the intrusion detection. As input dataset has feature set that have vule set. These features are cluster into selected and rejected group by moth flame optimization algorithm. Selected features were used for the identification of feature values that act as cluster center of intrusion and non-intrusion class detection by KNN approach. Expeirment was doen on IOT dataset and result shows that proposed modle has improved the precision value of intrusion detection by % as compared to other exisiting model. In future scholars can use some other training model to increase the detection accuracy of the work.

REFERENCES

1. J. Deogirikar and A. Vidhate, "Security attacks in iot: A survey," International Conference on I-SMAC (I-SMAC), pp. 32–37, 2017.
2. T. Bodstrom and T. H ¨ am¨ al¨ ainen, "State of the art literature review ¨ on network anomaly detection with deep learning," Internet of Things, Smart Spaces, and Next Generation Networks and Systems, pp. 64–76, 2018.
3. I. Arnaldo, A. Cuesta-Infante, A. Arun, M. Lam, C. Bassias, and K. Veeramachaneni, "Learning representations for log data in cybersecurity," International Conference on Cyber Security Cryptography and Machine Learning, pp. 250–268, 2017.
4. M. Du, F. Li, G. Zheng, and V. Srikumar, "Deeplog: Anomaly detection and diagnosis from system logs through deep learning," Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pp. 1285–1298, 2017.
5. B. J. Radford, B. D. Richardson, and S. E. Davis, "Sequence aggregation rules for anomaly detection in computer network traffic," arXiv preprint arXiv:1805.03735, 2018.
6. I. Lambert and M. Glenn, "Security analytics: Using deep learning to detect cyber attacks," 2017.
7. M. Stevanovic and J. M. Pedersen, "Detecting bots using multi-level traffic analysis." IJCSA, vol. 1, no. 1, pp. 182–209, 2016.
8. Zhang, Y.; Li, P.; Wang, X. Intrusion Detection for IoT Based on Improved Genetic Algorithm and Deep Belief Network. IEEE Access 2019, 7, 31711–31722.
9. Moustafa, N.; Turnbull, B.; Choo, K.R. An Ensemble Intrusion Detection Technique based on proposed Statistical Flow Features for Protecting Network Traffic of Internet of Things. IEEE Internet Things J. 2018, 6, 4815–4830.
10. Bagaa, M.; Taleb, T.; Bernal, J.; Skarmeta, A. A machine learning Security Framework for Iot Systems. IEEE Access 2020, 8, 114066–114077.
11. Susilo, B.; Sari, R. Intrusion Detection in IoT Networks Using Deep Learning Algorithm. Information 2020, 11, 279.
12. Liu, J.; Kantarci, B.; Adams, C. Machine Learning-Driven Intrusion Detection for Contiki-NG-Based IoT Networks Exposed to NSL-KDD Dataset. In Proceedings of the ACM Workshop on Wireless Security and Machine Learning, Linz, Austria, 13 July 2020.
13. Pokhrel, S.; Abbas, R.; Aryal, B. IoT Security: Botnet detection in IoT using Machine learning. arXiv 2021, arXiv:2104.02231.
14. 21. Shafiq, M.; Tian, Z.; Sun, Y.; Du, X.; Guizani, M. Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for internet of things in smart city. Future Gener. Comput. Syst. 2020, 107, 433–442.
15. Ullah I., Mahmoud Q.H. (2020) A Scheme for Generating a Dataset for Anomalous Activity Detection in IoT Networks. In: Goutte C., Zhu X. (eds) Advances in Artificial Intelligence. Canadian AI 2020.
16. A. Fatani, M. Abd Elaziz, A. Dahou, M. A. A. Al-Qaness and S. Lu, "IoT Intrusion Detection System Using Deep Learning and Enhanced Transient Search Optimization," in IEEE Access, vol. 9.