Mesopotamian journal of Cybersecurity
Vol.2023, **pp.** 48–56
DOI: https://doi.org/10.58496/MJCS/2023/009; ISSN: 2958-6542
https://mesopotamian.press/journals/index.php/CyberSecurity

Research Article

An efficient cyber security system based on flow-based anomaly detection using Artificial neural network

J.Jasmine Hephzipah[1,], Ranadheer Reddy Vallem[2,*],M.Sahaya Sheela[3,] ,G.Dhanalakshmi[4,]

[1]*Department of Electronics & Communication Engineering, R.M.K.Engineering College, RSM Nagar, Kavaraipettai-601206,TamilNadu,India,*
[2]*Research Scholar, Chaitanya Deemed To Be University, Kishanpura, Hanamkonda, Warangal -506001, Telangana, India,:*
[3]*Associate Professor,* [2]*Assistant Professor, Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu-600062, India,*
[4]*Professor,Department of Department of Electronics and Communication Engineering, Siddhartha Institute of Technology and Sciences, Hyderabad-500088, Telangana,*

ARTICLE INFO

Article History
Received 5 Jun 2023
Accepted 1 March 2023
Published 5 March 2023

Keywords
Crime detection
internet security
classification
Defect transmission rate
cyber security

ABSTRACT

Cyber security is developing factor for protecting internet resources by handing various monitoring feature based support to improve the security. Increasing internet cries in the defined facts for need of advance met in cyber security. Most internet attacker's theft the information through malicious activities, false data injection, hacking and make soon creating procedures. In most cases cyber sercuity failed to detect the malicious activities because the monitoring feature analyses improper to predict the result in previous machine learning algorithms. TO resolve this problem to propose an advance cyber security based on flow-based anomaly detection using Min max game theory optimized artificial neural network (MMGT-ANN). The reprocessing was carried out with KDD crime dataset. Then Data driven network model is applied to monitor the feature margins and defect scaling rate. Based on the feature scaling rate Transmission Flow defect rate is estimated and applied with Min max Game theory to select the feature limits. Then features are trained with optimized ANN to detect the crime rate. By the attention of the proposed system achieves higher performance in precision rate to attain higher detection accuracy with lower time complexity compared to the other system.

1. INTRODUCTION

The growth of information and networking technology has supported the organizations in many ways. Such development has been adapted by different organizations and sectors

Network security is used to secure the computer networks from unauthorized access, exploitation, modification and manages the network administrator or communication medium. It depends on layers of security and it contains several components including network monitoring, data analysis through deep learning models. It aims a variety of threats and breaks them from arriving or increasing on your network security performance

Internet security is used to protect your data and computing power from damage or theft. Data security means that data stored on a computer cannot be read by unauthorized individuals. For the past 15 years, people have been ignorant about computer and cyber security. A security problem is the execution of malicious code on a computer due to system or application vulnerabilities or threats that lead to improper user actions and system failures. The computer is extended and connected to the Internet, browsing the Internet, e-commerce, social networks and e-mail. This mainframe faces various types of problems like viruses, spyware, malware, phishing, spam, scams etc. A virus program damages a computer and replicates itself, usually resulting in data loss [1]. Spyware is a type of malicious code that monitors your Internet activity without your knowledge. Phishing, scams, malware and spam scams all fall into one category and force users to click on items such as links, links and images. The most common and persistent threats to security stop in your inbox. They come from various sources and appear canonical. Most cyber attacks are quick, easy and cheap, making them difficult to detect and track. Cyber security is used to improve integrity, confidentiality, reliability and authentication to avoid financial loss, personal loss, privacy loss, data loss and computer outages.

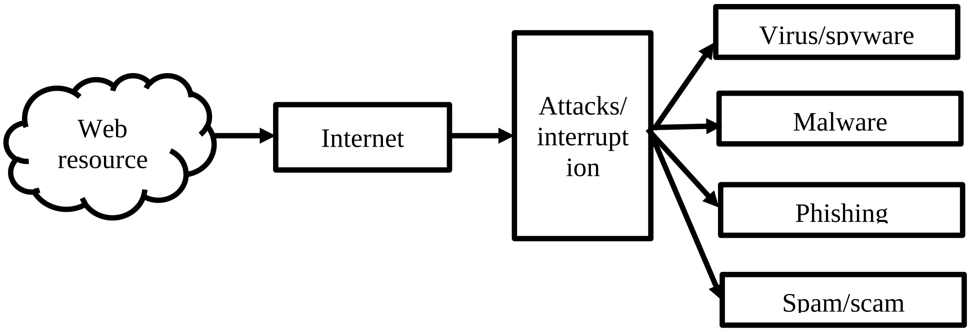

Figure 1. internet point of attacks

The internet technology is an ever growing sector of this decade and the users perform their day to day activities through them. The computers are affected by means of virus, spyware, malware, phishing, spam, scam etc. as shown in figure 1. There are virus programmers designed to damage the computers which can replicate themselves and can cause loss or corruption of data. There are malicious users or hackers who steal the user sensitive information to perform phishing attack [2-4].

So the internet security is one of the most valuable factors which are based on the principles of confidentiality and integrity. These factors become quite challenging in packet sniffers, malicious routers, covert channels, eavesdroppers and client side script attackers on internet.

Machine learning and deep learning models are used to analyze data transmission behavior to detect intrusion. It is used in many research fields, and there are things like machine learning approaches. It avoCrime attack dimensional curses because it reduces their time in learning, reduces data collection costs, reduces collection and improves classification performance. Facilities selection is one of the critical building blocks of data processing. It is. It reduces the evolutionary by eliminating inappropriate and unwanted features. Must be capable of determining optimal feature set validity [5]**.**

2. LITERATURE SURVEY

[6] The authors proposed a new approach to quantify the relevance and importance of texts describing cyber security. Predefined 'important' texts are analyzed according to their textual similarity with different repositories, and maximum similarity is calculated considering the textual importance. [7] The authors present coalitional Insurance as a talented another or complement to the outdated cover plans offered by It resources are mostly affected by cybercrimes. Below the future cyber cover perfect, several broadcast operators form an cover union, and the union premiums take into account system susceptibilities and damage distribution. [8] The authors propose a measurement data source recognition algorithm based on feature extraction techniques like KNN, ANN, including ensemble empirical mode decomposition and real-time scaled data classification with machine learning. However, power system cyber security is also addressed by data spoofing attacks. [9] The authors conducted a methodical appraisal of CC values and their adoption. Barriers to CC adoption are explored based on an examination of present cybersecurity assessment trends using Random forest Algorithm (RFA). Additionally, share experience and lessons learned with the CC from the recent development of the Australian Cyber Criteria Assessment Program to develop profiles when defining security requirements. [10] The authors present small business cybersecurity and emphasize aligning this study with the popular NIST Cybersecurity Framework (CSF). Based on the literature summarizes the main challenges SMEs face in implementing good cybersecurity and summarizes key recommendations. [11] The authors proposed several measurements to monitor abnormal load deviations and suspicious branch flow changes. A formal two-step approach is proposed to detect False Data Injection (FDI) cyber-attacks. The first phase determines whether the system is attacked, and the second phase identifies the target branch. [12] The authors introduce Cyber Threat Intelligence (CTI), which provides intelligence analysis to identify known and unknown threats using various characteristics such as threat actor capabilities, motivations, and indicators of compromise (IoC). Analyze and predict threats to improve Internet supply chain security. Combine CTI and ML techniques to analyze and predict threats based on CTI attributes. [13] The authors present recent analytical research on social networks and Internet traffic, using common concepts to classify Internet hosts or applications and users or tweets, using similarity, communication, and shared security goals. The ability to do so depends on the broader use of diverse networks or social trends rather than being identified in isolation.

[14] The authors provide a complete appraisal of modelling, imitation, examination approaches, categories of cyberattacks, and state-of-the-art CPPS cybersecurity countermeasures. Consideration of other cyberattack discovery and justification mechanism arrangements for entire power systems is given. [15] The authors proposed a systematic approach for security validation of OT-focused CPSs using UPPAAL model validation to improve network security. As a result of valid security attacks, these are considered insecure environments. That's it's essential to identify the weaknesses in your CPS.

3. PROPOSED SOLUTION

Classification is a standard data mining technique based on machine learning that usually uses the information to classify a data set between different classes. Many methods serve as a classification of information mining. Therefore, this work presents the idea of implementing an intelligent anti-jamming system based on deep learning. This method does not require prior knowledge of interference patterns and network models. We use MMGT-ANN based on learning to improve behavioural processes.

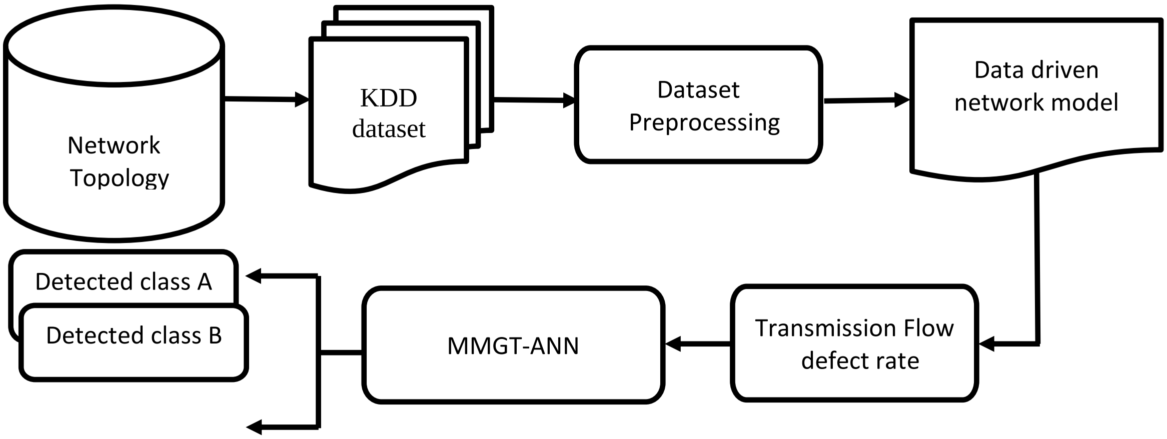

Fig.2: proposed architecture diagram - MMGT-ANN

Use collective intelligence with deep reinforcement to improve sensor performance and reduce detection latency. Figure 2 shows the proposed architecture diagram - MMGT-ANN. It also includes usage and occupancy of channels and conversion rates of channels. The proposed framework was evaluated for parameters such as key generation, delay, power, packet loss, transactions and timing related parameters.

In wireless communication systems, security quality has a very important impact on communication performance. A communication system's ability to achieve channel efficiency is highly dependent on collecting accurate channel position information. In fact, it is very difficult to properly assess security. Least squares and least squares error are two common channel estimates. In addition to the frequently used traditional model-based security method, a new DL-based scheme has recently been used, which provides an optimized ANN based detection and improve the cyber security. In essence, traditional communication systems are usually based on long-established mathematical algorithms. Therefore, the implementation of deep learning methodology can inspire new data-driven perspectives for the wireless universe.

3.1 Data driven network model

Initially the pre-processing ins carried out to normalize the dataset,. This preserves the feature monitoring approach from monitored attain noiseless data. The data drives model creates for data collection to observe he feature limits to create graph point in Graph G and vertivs V in edge point E which if collective feature set

$E = \{e_{ij} / i, j \in N\}$, that represent N number of nodes

$$e_{ij} = \begin{cases} 1, & if \;\; d(i, j) \le R_n \\ 0 & otherwise \end{cases} \quad ,$$

$d(i, j)$ represent the node and its feature limits variation $i, j \in N$.

Based on the data driven approach, the feature rate is collected and preprocessed to reduce the feature dimension.

3.2 Transmission Flow defect rate

In this stage , feature are analysed with observed margins to identify the defect flow during the packet transmission. This Computes the feature limits in defect scale rate W_t is estimated by frequency set Fs (F (f(x))

$$f(x) \rightarrow W_t\left(Val(fs_1) \leftrightarrow Val(fs_2)\right)$$

$$f\ (x) = \frac{\text{vs}}{|A_{\text{vs}}|}$$

The frequency limits are points the variation in mean factor f(x) at
$f\ (x) = \frac{\text{vs}}{|A_{\text{vs}}|}$ -(f (x)-f (x-1)).

Frequency set Fs (A_{vs}) $= \frac{\sum_{n=1}^{A_{\text{vs}}}\{vs_n:(fs_{iC1} \rightarrow vs_n) \neq 0\}}{\sum_{n=1}^{A_{\text{vs}}} vs_n} - \ x^{n-1} f(x)$

TO attain the difference level by inverse definition F (x^{I}) to get the actual feature weight.

$FSA^{|}$ (vs)=1-$\frac{\sum\{FSA(\{Val_i \exists Val_j\}):(Val_j \complement vs_i)\}}{|vs_i|}$

The feature thresholds f (x) by feature scaling will be multiple to source the alteration by means of opposite function to get the actual feature limits. It chooses the marginal weighting function by means of the cumulative difference of the 'R' feature weights

3.3 Min max-game theory

The maximum and minimum scale detect range are variated based on this game theory approach/Once the feature value is activated, a mini-max game is initiated between the generator and the discriminator.

• The strategy of the game is explained by the equation,

$$\min_G \max_D \mathbb{E}_{x \sim p_{data}}[\log(D(x))] - \mathbb{E}_{z \sim p_{datz}}[\log(1 - D(G(z)))]$$

Based on the z- input error, p_{data}-data distribution, D-discriminator function, G-Generator function, after the mini-max game is over, the perceived output is the difference in behavior between the normal nodes and the attacking nodes. Classifies attacking hosts and excludes them from data transmission.

3.4 Artificial neural network

The ANN was constructed and features are feed into input layer crime margins to create sequence feature limits. The crime limits be linear value into $CRM^{|}{}_{T1}, CRM^{|}{}_{T2} \ldots .n.$ The hidden layer is a layer between the input layer and output layer. The inoput feature lmits are feed forward into subjective levels as $h_1, h_2 \ldots n$. Feature variation margins. The 16 cross layers is used to cross evaluate the Logical neurons to carry the crime feature limits.

$$C(\theta) = \sum_i [y_T]_i \ \log(\, a^L\,(x_T)_i) + (1 - [y_T]_i) \log(\, 1 - (a^L\,(x_T)_i))$$

- θ is the set of the neural network parameters,
- x_T is the training data vector,
- y_T is the label vector (Attacker/Normal),
- a^L (x_T) is the output of the neural network at the last layer L.
- In the hidden layer, sigmoid and hyperbolic tangent (Tanh) functions are discarded for implementation.

  $[\sigma(z)]_k = \frac{1}{1+e^{-z_k}}$, $\frac{e^{z_k} - e^{-z_k}}{e^{z_k} + e^{-z_k}}$,

  z-input, k-entry count
- The output layer is implemented using a smooth maximum activation problem with gradient descent given in the following equation.

$$[\sigma(z)]_k = \frac{e^z k}{\sum_j e^z j}$$

The output layers produce class margins to detect the crime rates under different class by reference. This classifies the defect dependencies on class margins rate, the actual value of comparison be trained with ideal crime rate to produce efficient result.

4. RESULT AND DISCUSSION

The proposed implementation is tested on python language with an anaconda environment using publicly available cloud KDDcup99 dataset. Crime Attack can be effectively detected by comparative parameters such as classification accuracy, sensitivity, specificity, false ratio and time complexity with the help of a confusion matrix.

TABLE I: ENVIRONMENT AND VALUES PROCESSED

Parameters	**Values**
Simulation Tool	Anaconda, Jupiter notebook
Simulation language	Python
Name of the dataset	KDDcup99 dataset
No of users/ records	500/ 2500
Number of classes	High / medium / low

Table 1 shows the Environment and values processed in simulation environment. The comparison algorithms are, Random Forest Algorithm (RFA), Support Vector Machine (SVM), K-Nearest Neighbour (KNN) are compared with proposed MMGT-ANN.

TABLE II: PERFORMANCE ON INTRUSION DETECTION ACCURACY VS. NO OF SERVICES

Intrusion Detection Accuracy in % vs No of Services			
Comparison methods/ services	**10 Services**	**20 Services**	**30 Services**
RFA	70.9	73.6	78.3
KNN	76.2	77.1	81.8
SVM	78.7	82.4	87.5
MMGT-ANN	91.9	96.6	98.3

Table 2 describes the Crime Attack accuracy performance vs no of services with different techniques like RFA, KNN, SVM and the proposed MMGT-ANN produce high performance than any other methods.

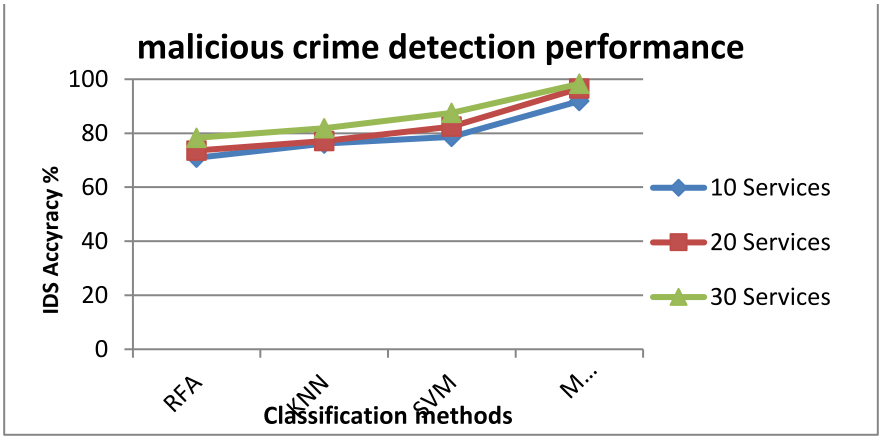

Fig. 3: Impact of malicious Crime Attack accuracy performance

Figure 3 denotes impact of intrusion detection classification accuracy performance with various services like 10, 20 and 30. The proposed technique MMGT-ANN method attained 98.3% for 30 services also the previous RFA attained 78.3%, KNN 81.8%, and SVM attained 87.8%. Nonetheless, the proposed method produces better performance than other techniques.

TABLE III: IMPACT OF SENSITIVITY PERFORMANCE

Comparison methods/ services	**10 Services**	**20 Services**	**30 Services**
RFA (%)	68.6	71.8	77.1
KNN (%)	73.2	75.6	80.7
SVM (%)	76.7	81.5	86.2

MMGT-ANN (%)	89.3	94.5	97.1

Table 3 describes the impact of sensitivity performance the proposed compared with previous techniques.

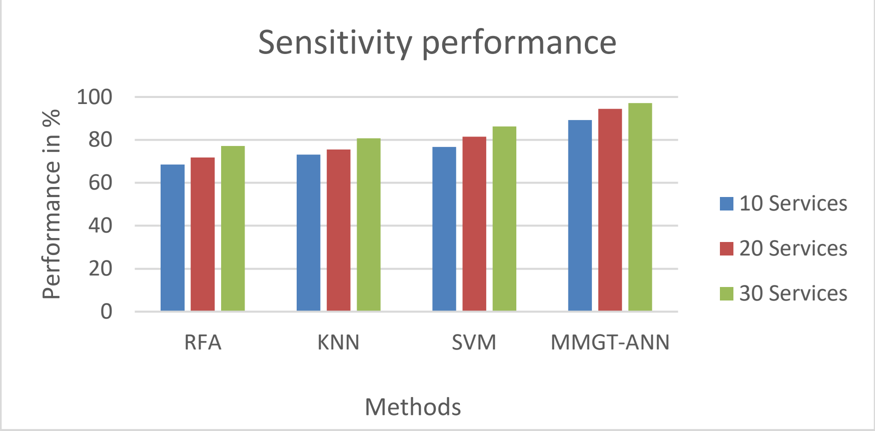

Fig.4: Analysis of sensitivity performance

Figure 4 shows the sensitivity performance for CRIME ATTACK detection using MMGT-ANN algorithm. The proposed algorithm provide result is 96% of sensitivity performance for 30 services; similarly the exiting algorithm provide results are RFA is 76% of Sensitivity performance, KNN is 81% of Sensitivity performance and SVM is 85% of Sensitivity performance for 30 services.

TABLE IV: IMPACT OF SPECIFICITY PERFORMANCE

Comparison methods/ services	10 Services	20 Services	30 Services
RFA (%)	71.2	72.6	78.4
KNN (%)	78.7	76.8	83.6
SVM (%)	79.4	82.8	87.4
MMGT-ANN (%)	90.8	95.2	96.8

Table 4 describes the analysis of specificity performance measures in different number of services such as 10, 20, and 30 services. The proposed technique provide better result than previous approaches.

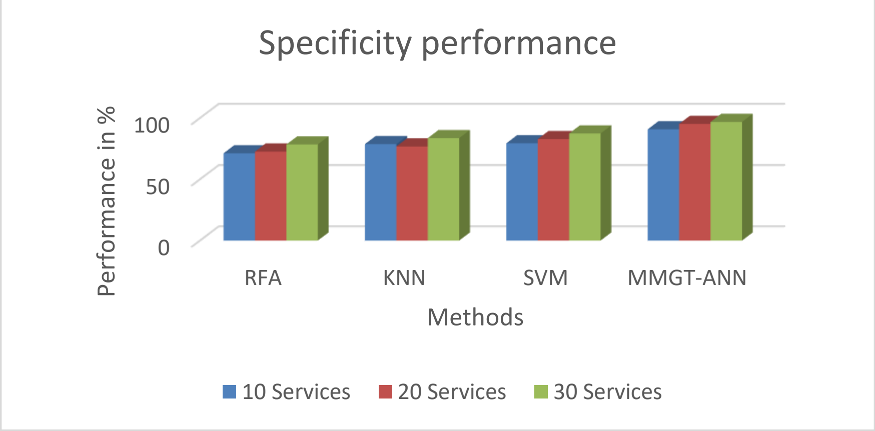

Fig.5: Analysis of Specificity performance

Figure 5 illustrate the analysis of specificity performance the proposed and previous approaches comparison result presented. The proposed MMGT-ANN algorithm has 97% of Specificity performance for 30 services; similarly the existing algorithm results are RFA is 77% of specificity performance, KNN is 82% of specificity performance and SVM is 86% of specificity performance for 30 services.

TABLE V: ANALYSIS ON FALSE CLASSIFICATION RATIO

False Classification Ratio in % vs No of Services			
Comparison methods/ services	**10 Services**	**20 Services**	**30 Services**
RFA	28.4	25.1	20.7
KNN	23.5	21.4	16.9
SVM	20.7	16.9	12.6
MMGT-ANN	6.7	2.4	1.3

Analysis of false classification ratio the proposed comparison with previous methods performance is listed in table 5.

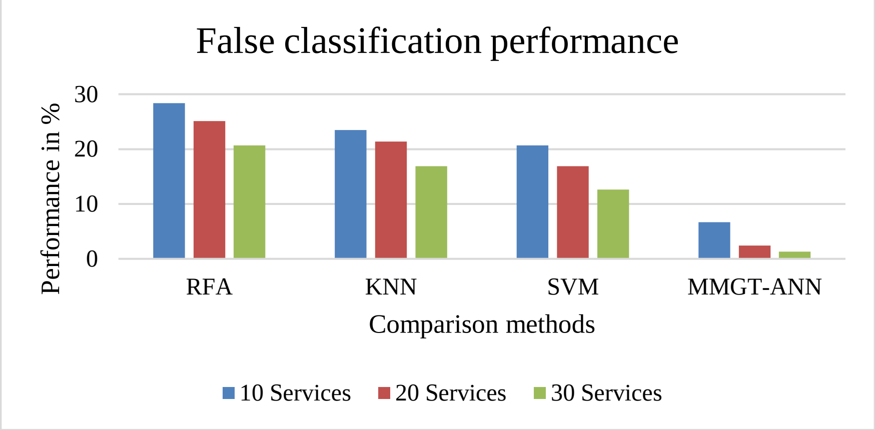

Fig.6: Impact of false classification ratio

Figure 6 illustrates impact of false classification ratio performance for Crime Attack with various services like 10, 20 and 30 services. In this graph, X-axis is a comparison methods moreover Y-axis performance gradually decrease with each method. The proposed MMGT-ANN method achieves 1.3% false classification performance for 30 services besides RFA achieves 20.7% of false classification performance, KNN method achieves 16.9%, and SVM method achieves 12.6%.

TABLE VI: IMPACT OF TIME COMPLEXITY PERFORMANCE

Time complexity in seconds vs No of Services			
Comparison methods/ services	**10 Services**	**20 Services**	**30 Services**
RFA	62.4	66.1	67.2
KNN	48.6	52.8	58.6
SVM	31.6	35.5	41.2
MMGT-ANN	17.5	21.6	28.3

Table 6 describes the impact of time complexity performance vs no of services. The proposed MMGT-ANN has 28.3 sec for Crime Attack classification besides RFA has 67.2 sec, KNN has 58.6 sec and SVM has 41.2 sec for Crime Attack classification.

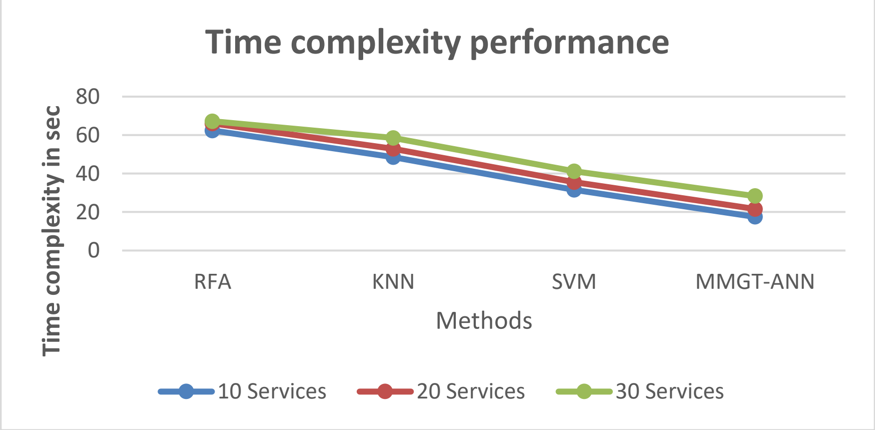

Fig.7: Result of time complexity performance

Figure 7 denotes result of time complexity performance the proposed MMGT-ANN technique compared with other methods like RFA, KNN and SVM. In figure X-axis presents comparison methods besides Y-axis presents time complexity performance in seconds with each methods. However the proposed method produced less time complexity result than previous techniques.

TABLE VII. PERFORMANCE ON VARIOUS MEASURES

Comparison methods/ services	Detection Rate %	False Ratio %	Time Complexity in sec
RFA	78.6	18.6	62.5
KNN	81.3	16.2	48.4
SVM	87.2	10.6	31.5
MMGT-ANN	98.2	1.3	17.9

Table 7 denotes the proposed MMGT-ANN performance of various measures based on detection rate, false ratio and time complexity. The proposed MMGT-ANN techniques gives better performance than addition methods.

5. CONCLUSION

An efficient cyber security system based on flow-based anomaly detection using Artificial neural network produce high performance compared to the other system. Tis archives high detection accuracy based on crime feature evaluation and reduce the time and false detection rate. The min max gaming approach reduces the dimensionality of non-deterministic characteristic ranges to choose from crime rate scaling factor. Importance of min max game theory analysis helps in crime attack detection rate. This proposed MMGT-ANN makes it an overwhelming choice in crime attack with high intrusion detection accuracy of 94.2%, false rate of 0.4%, and occasional degree time complexity of up to 12 seconds. Intrusion detection proves detection as well as different methods.

Funding

Non.

Conflicts Of Interest

The authors declare no conflicts of interest.

Acknowledgment

Authors would like to thank the anonymous reviewers for their efforts.

REFERENCES

[1]. O. Mendsaikhan, H. Hasegawa, Y. Yamaguchi and H. Shimada, "Quantifying the Significance and Relevance of Cyber-Security Text Through Textual Similarity and Cyber-Security Knowledge Graph," in IEEE Access, vol. 8, pp. 177041-177052, 2020, doi: 10.1109/ACCESS.2020.3027321.

[2]. P. Lau, L. Wang, Z. Liu, W. Wei and C. -W. Ten, "A Coalitional Cyber-Insurance Design Considering Power System Reliability and Cyber Vulnerability," in IEEE Transactions on Power Systems, vol. 36, no. 6, pp. 5512-5524, Nov. 2021, doi: 10.1109/TPWRS.2021.3078730.

[3]. S. Liu et al., "Model-Free Data Authentication for Cyber Security in Power Systems," in IEEE Transactions on Smart Grid, vol. 11, no. 5, pp. 4565-4568, Sept. 2020, doi: 10.1109/TSG.2020.2986704.

[4]. N. Sun et al., "Defining Security Requirements with the Common Criteria: Applications, Adoptions, and Challenges," in IEEE Access, vol. 10, pp. 44756-44777, 2022, doi: 10.1109/ACCESS.2022.3168716.

[5]. Z. Zhao, Y. Huang, Z. Zhen and Y. Li, "Data-Driven False Data-Injection Attack Design and Detection in Cyber-Physical Systems," in IEEE Transactions on Cybernetics, vol. 51, no. 12, pp. 6179-6187, Dec. 2021, doi: 10.1109/TCYB.2020.2969320.

[6]. Z. Zhang, H. A. Hamadi, E. Damiani, C. Y. Yeun and F. Taher, "Explainable Artificial Intelligence Applications in Cyber Security: State-of-the-Art in Research," in IEEE Access, vol. 10, pp. 93104-93139, 2022, doi: 10.1109/ACCESS.2022.3204051.

[7]. M. N. Y. Marican, S. A. Razak, A. Selamat and S. H. Othman, "Cyber Security Maturity Assessment Framework for Technology Startups: A Systematic Literature Review," in IEEE Access, vol. 11, pp. 5442-5452, 2023, doi: 10.1109/ACCESS.2022.3229766.

[8]. A. Chidukwani, S. Zander and P. Koutsakis, "A Survey on the Cyber Security of Small-to-Medium Businesses: Challenges, Research Focus and Recommendations," in IEEE Access, vol. 10, pp. 85701-85719, 2022, doi: 10.1109/ACCESS.2022.3197899.

[9]. X. Li and K. W. Hedman, "Enhancing Power System Cyber-Security With Systematic Two-Stage Detection Strategy," in IEEE Transactions on Power Systems, vol. 35, no. 2, pp. 1549-1561, March 2020, doi: 10.1109/TPWRS.2019.2942333.

[10]. A. Yeboah-Ofori et al., "Cyber Threat Predictive Analytics for Improving Cyber Supply Chain Security," in IEEE Access, vol. 9, pp. 94318-94337, 2021, doi: 10.1109/ACCESS.2021.3087109.

[11]. J. Ye et al., "Cyber–Physical Security of Powertrain Systems in Modern Electric Vehicles: Vulnerabilities, Challenges, and Future Visions," in IEEE Journal of Emerging and Selected Topics in Power Electronics, vol. 9, no. 4, pp. 4639-4657, Aug. 2021, doi: 10.1109/JESTPE.2020.3045667.

[12]. H. Tao et al., "TrustData: Trustworthy and Secured Data Collection for Event Detection in Industrial Cyber-Physical System," in IEEE Transactions on Industrial Informatics, vol. 16, no. 5, pp. 3311-3321, May 2020, doi: 10.1109/TII.2019.2950192.

[13]. R. Coulter, Q. -L. Han, L. Pan, J. Zhang and Y. Xiang, "Data-Driven Cyber Security in Perspective—Intelligent Traffic Analysis," in IEEE Transactions on Cybernetics, vol. 50, no. 7, pp. 3081-3093, July 2020, doi: 10.1109/TCYB.2019.2940940.

[14]. R. V. Yohanandhan, R. M. Elavarasan, P. Manoharan and L. Mihet-Popa, "Cyber-Physical Power System (CPPS): A Review on Modeling, Simulation, and Analysis With Cyber Security Applications," in IEEE Access, vol. 8, pp. 151019-151064, 2020, doi: 10.1109/ACCESS.2020.3016826.

[15]. C. -C. Chan, C. -Z. Yang and C. -F. Fan, "Security Verification for Cyber-Physical Systems Using Model Checking," in IEEE Access, vol. 9, pp. 75169-75186, 2021, doi: 10.1109/ACCESS.2021.3081587.