Intrusion Detection System in IoT Based on GA-ELM Hybrid Method

Elijah M. Maseno [1],*, Zenghui Wang [2], and Fangzhou Liu [3]

[1] Department of Computer Science, University of South Africa, Florida, South Africa
[2] Department of Electrical Engineering, University of South Africa, Florida, South Africa
[3] Research Center of Intelligent Control and Systems, Harbin Institute of Technology, Harbin, China
*Correspondence: 13090879@mylife.unisa.ac.za (E.M.M)

***Abstract*—In recent years, we have witnessed rapid growth in the application of IoT globally. IoT has found its applications in governmental and non-governmental institutions. The integration of a large number of electronic devices exposes IoT technologies to various forms of cyber-attacks. Cybercriminals have shifted their focus to the IoT as it provides a broad network intrusion surface area. To better protect IoT devices, we need intelligent intrusion detection systems. This work proposes a hybrid detection system based on Genetic Algorithm (GA) and Extreme Learning Method (ELM). The main limitation of ELM is that the initial parameters (weights and biases) are chosen randomly affecting the algorithm's performance. To overcome this challenge, GA is used for the selection of the input weights. In addition, the choice of activation function is key for the optimal performance of a model. In this work, we have used different activation functions to demonstrate the importance of activation functions in the construction of GA-ELM. The proposed model was evaluated using the TON_IoT network data set. This data set is an up-to-date heterogeneous data set that captures the sophisticated cyber threats in the IoT environment. The results show that the GA-ELM model has a high accuracy compared to single ELM. In addition, Relu outperformed other activation functions, and this can be attributed to the fact that it is known to have fast learning capabilities and solves the challenge of vanishing gradient witnessed in the sigmoid activation function.**

***Keywords*—intrusion detection system, extreme learning machine, genetic algorithm, TON_IoT data sets, hybrid**

I. Introduction

IoT has been adopted in different fields, which include but are not limited to smart homes, health care, farming, power grids, and smart wearables. These devices use different protocols to connect to the internet, which increases the complexity of the internet. The integration of a large number of electronic devices exposes IoT technologies to various forms of cyber-attacks [1, 2]. Cybercriminals have shifted their focus to IoT as it provides a broad network intrusion surface area. These Cyber-attacks are targeted toward breaching sensitive information, disrupting services, and breaching financial sector records. Cybercriminals have invested heavily to improve the efficiency and effectiveness of their tools. The focus of these tools is on how to evade the existing cyber defense mechanisms undetected. To achieve this, cybercriminals have sorted to invest in machine learning algorithms to develop their arsenals. Cybercriminals are leveraging ML algorithms to develop complex and intelligent cyber-attacks. These tools have the capability of learning their environment and evolving when need be. One of the standard techniques in cyber defense is Intrusion Detection Systems (IDS). IDS is used as a tool for the cyber defense to monitor and secure networks. Most of the existing IDS have proven to be ineffective in protecting IoT devices [3]. To improve the performance of the existing IDS, researchers have proposed the integration of two or more machine learning algorithms [4]. The proposed hybrid intrusion detection systems have proven to be superior to stand alone traditional intrusion detection systems. This work proposes the optimization of ELM using GA. Research shows that a major limitation of ELM is that it randomly selects the input weights and the biases. One of the commonly proposed techniques for overcoming this limitation is the use of metaheuristic algorithms [5]. Huang and Jiang *et al*. [6], Alexandre and Cuadra *et al*. [7], Matias and Araújo *et al*. [8] proposed the application of a Genetically Optimized Extreme Learning Machine (GA-ELM) in different fields. In these studies, GA was used for the optimization of input parameters or optimization of ELM structure. It was observed that the researchers proposed the use of the sigmoid activation function. The Sigmoid activation function is known to be slow in learning and in addition it suffers from vanishing gradient problem. Ali and Zolkipli *et al*. [9], in most of the existing works, the choice of activation function is always not justified but randomly selected. The study further proposes the investigation of GA-ELM using the different activation functions. GA and ELM have been widely used in the field of intrusion detection separately. Researchers have observed that most of the available data sets used for the validation of intrusion detection systems have limitations and cannot be reliably used to evaluate modern intrusion detection systems [10, 11]. Moustafa [10] observed that these data sets do not capture the complex cyber threats of

Manuscript received September 25, 2022; revised November 12, 2022; accepted January 4, 2022; published July 5, 2023.

doi: 10.12720/jait.14.4.625-629