Reputation-based security model for detecting biased attacks in big data

Vinod Desai[1], **Dinesha Hagare Annappaiah**[2]
[1]Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning),
BLDEA's V P Dr PG Halakatti College of Engineering and Technology, Vijayapura, India
[2]Department of Computer Science and Engineering, Nagarjuna College of Engineering and Technology, Bangalore, India

Article Info

Article history:

Received Jun 28, 2022
Revised Oct 14, 2022
Accepted Oct 24, 2022

Keywords:

Cyber attack
IoT
Malicious
RBAD
Security

ABSTRACT

As internet of things (IoT) devices are increasing since the emergence of these devices in 2010, the data stored by these devices should have a proper security measure so that it can be stored without getting in hands of an attacker. The data stored has to be analyzed whether the data is safe or malicious, as the malicious data can corrupt the whole information. The security model in big data has many challenges such as vulnerability to fake data generation, troubles with cryptographic protection, and absent security audits. As cyber-attacks are increasing the main objective of each organization is to secure the data efficiently. This paper presents a model of reputation security for the detection of biased attacks on big data. The proposed model provides various evaluation models to identify biased attack in malicious IoT devices and provide a secure communication metric for big data. The results show better rates in terms of attack detection rate, attack detection failure rata, system throughput and number of dead nodes when the attack rate is increased when compared with the existing reputation-based security (ERS) model. Moreover, this model reputation-based biased attack detection (RBAD) increases the security of the IoT devices in the big data and reduces the biased attack coming from various malicious nodes.

Corresponding Author:

Vinod Desai
Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning),
BLDEA's V P Dr PG Halakatti College of Engineering and Technology
RPX8+9XQ, Ashram Rd, Adarsh Nagar, Vijayapura, Karnataka 586103, India
Email: desaivinod2021@gmail.com

1. INTRODUCTION

Internet of things (IoT) [1] is widely used in our day-to-day life and is used in different areas having different applications. IoT applications can range from small sensors to large business applications. Using the sensors [2] and different IoT devices. Moreover, the security in big data is very challenging as it is concerned with attacks [3] that can originate either from online or offline spheres, hence, we can collect data and store the data using different protocols such as hypertext transfer protocol (HTTP) [4], message queue telemetry transport protocol (MQTT) [5], and constrained application protocol (CoAP) [6]. The data is stored so it can be utilized for the improvement of the device. Once the data has been stored it has to analyzed or has to be processed so it can be classified using an attribute. For the success of the IoT applications security, privacy, and trust play an important role. Big data [7], [8] plays a big role in IoT applications as it analyzes data systematically by extracting the data from the datasets. The attacks include theft of the data that is stored online, ransomware, or distributed denial of service (DDoS) attacks [9] that can crash the whole server. This issue can even be bad for the companies where the stored data is confidential or even sensitive, such as the details of the customer, numbers of the credit card, or even the details of the contact of the customers. These attacks lead to the organization's big loss and can cause financial problems. There are several ways through which the security measures in big data can be used to protect the data. One of the most common security measures is to encrypt the data while transmitting the data from the sender to the receiver on online platforms. Encrypted data is useless to hackers if don’t have a proper encryption key to decode the data. Therefore, encryption is one of the methods using which the data can be protected from attackers and the data is completely protected in this process. Another method of security in big data is to create trust between the sender and the receiver. Many methods such as data mining [10], supervised [11] and unsupervised [12] machine learning algorithms have been used for the trust in data quality. Trust plays an important role in the security [12]–[14]. Two ways using which the trust method can be applied to the big data. The first method is to first trust the receiver whether he/she can handle the data without getting leaked, the second method is to make a trust between the sender and the receiver so that when the sender sends the data it should maintain a quality. After seeing all the above problems, we proposed a model which provides an evaluation model to identify biased attack in malicious IoT devices and provide a secure communication metric for big data. The organization of the paper has been formatted according to the following. In section 2 few existing systems along with the benefits and drawbacks have been given. In section 3 the proposed methodology has been explained. In this Section, the detail of all the evaluation methods along with the final reputation metric and communication metric has been explained. In section 4, the results obtained by the proposed model are explained. In Section 5, the conclusion and future work of the whole research work has been given.

2. LITERATURE SURVEY

Najib *et al.* [15], a survey on trust design models for IoT machines has been done. This survey gives an idea about the existing systems and methods that are currently being used for the trust models. It gives a classification of the trust models using five operations that includes the trust algorithm, trust metric, trust propagation, trust source, and trust architecture. Xiao *et al.* [16], a trust model using the blockchain has been used for mobile edge computing (MEC) which prevents different kinds of attacks. This model calculates the execution of the edge devices and sends that information to the nearby edge devices. It uses a method where it selects a miner of blockchain, which applies different protocols to get a recording of the block. They have proposed a reinforcement learning algorithm that improves the model efficiency. The security of the model is also calculated and the efficiency of the edge utility is provided. Wang *et al.* [17], they have addressed the biased load attack in smart grids using a feature selection model which selects the interval state of the node and eliminates all the unnecessary nodes which provide less threshold. Also, a matrix has been used to detect the attack in the sensors. Debe *et al.* [18], this model was built for the Ethereum blockchain and other technologies to keep the trust between the public fog nodes and IoT devices. The model is evaluated using performance, security, and cost. This model helps to keep the trust as long as possible as compared to the existing trust models. Alshammaria and Alsubhi [13], they have proposed a trust model to identify biased attack using malicious nodes. In this model they have provided security to the services using trust algorithms that can identify the nodes. The results show that this model provides good accuracy while detecting the malicious biased attacks. Ghafoorian *et al.* [19], it describes a model which is based on role-based access control (RBAC), which can prevent any security threat and has less execution time compared to the RBAC. This paper explains the security methods required for the trust-based system. This model is evaluated using the Advogato dataset. Liu *et al.* [20], a method of blockchain-authorized group-validation method is proposed for the vehicles with distributed identification based on the secret sharing and dynamic proxy mechanisms. The validation values are used for the combined authentication. The node with a higher-reputation which does the edge computing is stored in the blockchain and uploads the final values of the authentication to the server. Zhang *et al.* [21], a scheme has been proposed based on the probabilistic skyline computation technique. This scheme first assigns a trust score to each individual based on the performance without showing the information to the other node, and then it selects the subset of the reliable individuals for a particular work. This scheme helps to preserve the individual’s identity and it also gives a higher efficiency during the extensive simulations. Yuan and Li [22], a mechanism for the IoT edge devices which are based on multi-source feedback information fusion has been proposed. Two mechanisms have been used in this paper; the multi-score feedback mechanism and the lightweight trust evaluating mechanism. An algorithm named feedback information fusion is used which overcomes the traditional trust schemes. This mechanism provides higher reliability and efficiency. Nguyen *et al.* [23], it proposes a trust-worthy access control using the blockchain. This blockchain protects the clouds from illegal offloading. To protect this offloading from various attacks a deep reinforcement learning (DRL) algorithm has been used using the Q-network. Xu *et al.* [24], another offloading method using blockchain-based secure computation has been proposed. It uses the DRL algorithm and for the management of the trust, short-term trust variability and long-term reputation are considered. For a more complete reputation, a threevalued subjective logic (3VSL) is used. To store the data, validate, and update different kinds of blockchains have been used. To implement intelligent and secure computation in vehicular cloud networks, the DRL algorithm has been used. Mohammadi *et al.* [25], they have reviewed various exiting trust based IoT recommended techniques. In Figure 1, an example of how the trust has been evaluated according to the [25] has been given.

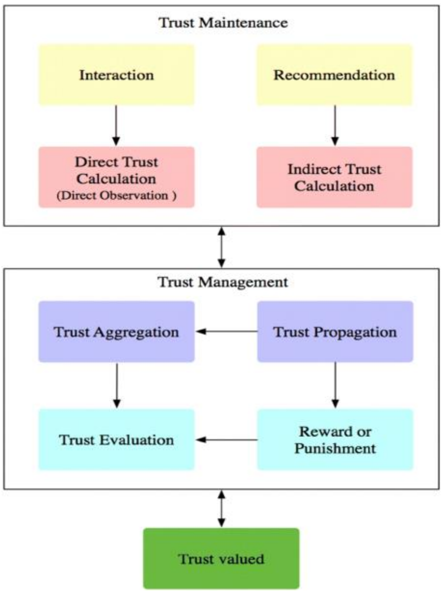

Figure 1. Example of a suggestion based on trust IoT modeling [25]

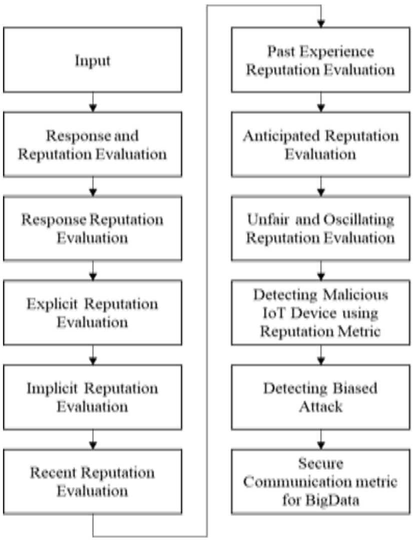
Input
Response and Reputation Evaluation
Response Reputation Evaluation
Explicit Reputation Evaluation
Implicit Reputation Evaluation
Recent Reputation Evaluation
Past Experience Reputation Evaluation
Anticipated Reputation Evaluation
Unfair and Oscillating Reputation Evaluation
Detecting Malicious IoT Device using Reputation Metric
Detecting Biased Attack
Secure Communication metric for BigData

Figure 2. Proposed architecture

3. METHOD

In the Figure 2, the proposed architecture has been given. In this architecture there are various evaluation methods which have been used to detect the biased attack. In this model, first the input is passed where the data contains malicious nodes which are constantly oscillating from one node to another node. The input goes through various evaluation models which will check which node has more reputation and after each evaluation it goes to the next evaluation process. In the final step we provide a secure communication metric for big data which will provide a reputed node as a final output without any malicious nodes. Further, each evaluation method has been given below in the sub sections.

3.1. Response and reputation evaluation

In this part, we evaluate the model using the response and reputation in the cluster-based wireless sensor networks (WSN). In this method, we first compute the reputation level of the WSNs using the cluster head. We validate that the sensor keeps the information of the trust level of the entire session when a connection between a software device to another software device is made. To minimize the cost of the storage between the IoT devices and to allocate the weights to a specific session, we use the exponential mean update process to store the result of the trust level. To describe the overall process of the reputation level, let $\mathrm{Sec}_o^u(x,y)$ where x and y are the different IoT devices having o communication in the u^{th} session period. The overall process can be described using the following (1):

$$\mathrm{Sec}_o^u(x,y) = \gamma * \mathrm{Sec}_{rec}(x,y) + (1-\gamma) * \mathrm{Sec}_{o-1}^u(x,y) \tag{1}$$

Where the Sec_{rec} depicts the reputation level parameter of the most current operation. The response-based security model, in which the sensor gives the response about the quality of experience with the other IoT devices is given using the following (2):

$$\mathrm{Sec}_{rec} = \begin{cases} 0, \text{if the interaction is completely untrustable} \\ 1, \text{if the interaction is completely trustable,} \end{cases} \tag{2}$$

in (1), the γ is improved using the variance

$$\gamma = W + d * \frac{\mu_o^u(x,y)}{1+\beta_o^u(x,y)} \tag{3}$$

the accumulated result is calculated using the following (4):

$$\beta_o^u(x,y) = d * \mu_o^u(x,y) + (1-d) * \beta_{o-1}^u(x,y) \tag{4}$$

in (4), d is a constant which shows how the sensor behaves if there is any failure in $\mu_o^u(x,y)$. The $\mu_o^u(x,y)$ is calculated using the following (5):

$$\mu_o^u(x,y) = |\mathrm{Sec}_{o-1}^u(x,y) - \mathrm{Sec}_{rec}| \tag{5}$$

Therefore, the model provides less reputation to the current variance than the previously collected variance if d is decreased. Moreover, it can be seen that as the $\mu_o^u(x,y)$ increases the γ also increases. From this, we can say that the significance of the higher threshold is given to the current response. The W in (3) shows the reputation weight which is the threshold of the value that helps to prevent γ to become a fixed parameter. The variation in the reputation response in the different IoT devices x and y is found by communicating the sensor device of x and y given using the (6):

$$\mathbb{V}_o^u(x,y) = \sqrt{\frac{\sum_{p\in\mathcal{H}}\left(\mathrm{Sec}_o^u(x,p) - \mathrm{Sec}_o^u(y,p)\right)^2}{|\mathcal{H}(x,y)|}} \tag{6}$$

in (6) the p depicts the common IoT devices, $\mathcal{H}(x,y)$ shows the interaction between the common sensor device with the sensor device x and y. To establish a connection between the sensor device x and $\left(\mathbb{R}_o^u(x,y)\right)$, it first compares $\mathbb{V}_o^u(x,y)$ with the connection $\mathcal{I}$. The connection is then updated using the following (7):

$$\mathbb{R}_o^u(x,y) = \begin{cases} \mathbb{R}_{o-1}^u(x,y) + \frac{1-\mathbb{R}_{o-1}^u(x,y)}{X}, \text{if } \mathbb{V}_o^u(x,y) < \mathcal{H}, \\ \mathbb{R}_{o-1}^u(x,y) - \frac{1-\mathbb{R}_{o-1}^u(x,y)}{Y}, \text{else,} \end{cases} \tag{7}$$

in (7), the X is used to show the award parameter, and Y is used to show the penalty parameter, where both of these parameters can be changed in the dynamic environment based on the security requirement.

3.2. Response reputation evaluation

In this part, the evaluation of the response from the reputation is done using the response data. In the current models of trust-based security for the WSNs, the trust data that is given using a good sensor device is assumed to be true and if the response is given from a malicious sensor device, then it is given as false. Since there may be a chance where the good sensor device can send the wrong trust data and the malicious sensor device may send correct data to hide the malicious behavior of the sensor device. Therefore, it is very important to create a method where the response reputation can be evaluated. To compute the trust level, the response given using the sensor device with good performance or the data having the quality that can be trustworthy can be trusted and safe is used. Hence, more weight is given to a sensor device that has high reliability and less weight is given to the sensor device which has low reliability. Assume $\mathrm{\mathbb{F}_o^u(x, y)}$ explains the response trust of the sensor device y from the sensor device x, then the value can be evaluated using the following (8):

$$\mathrm{\mathbb{F}_o^u(x, y)} = \begin{cases} 1 - \frac{\log\left(\mathrm{Sec_o^u(x,y)}\right)}{\log\theta}, \text{if } \mathrm{\mathbb{R}_o^u(x, y)} > \theta, \\ \qquad 0, \text{else} \end{cases} \tag{8}$$

in (8), the $\log\theta$ shows that the parameter is least tolerable.

3.3. Explicit reputation evaluation

In this part, explicit reputation evaluation is done using the intra-cluster communication and inter-cluster communication used in a different kind of WSNs. Let $\mathrm{\mathbb{L}_o^u(x, y)}$ show the explicit reputation data that the sensor device x has on sensor device y with the interaction o in the $\mathrm{u^{th}}$ session. Then the explicit reputation is calculated using the following (9):

$$\mathrm{\mathbb{L}_o^u(x, y)} = \mathrm{Sec_o^u(x, y)} \tag{9}$$

by using (9), if sensor device y gives a better performance, then the sensor device x can b considered as a reputation parameter. This helps the sensor device y to have trust in the sensor device x.

3.4. Implicit reputation evaluation

In this part, the implicit reputation evaluation is done using the intra-cluster communication and inter-cluster communication used in a different kind of WSNs. To achieve secure transmission of data from the sensors, the sensor device demands the other sensor to send the response data of the sensor device which is working with the specific sensor device. The sensor device then collects the response from the different sensor devices for calculating the implicit reputation using the following equation. In (10), $Z = \mathbb{S}(y)$ shows the sensor device set which has already been communicated to the sensor device y.

$$\mathbb{G}_o^u(x, y) = \begin{cases} \frac{\sum_{p\in Z-\{x\}} \mathbb{F}_o^u(x,p) * \mathbb{L}_o^u(x,y)}{\sum_{p\in Z-\{x\}} \mathbb{F}_o^u(x,p)}, if\ |Z - \{x\}| > 0, \\ \qquad 0, \qquad\qquad\qquad if\ |Z - \{x\}| = 0. \end{cases} \tag{10}$$

3.5. Recent reputation evaluation

In this part, the recent reputation evaluation is done using the intra-cluster communication and inter-class cluster communication used in a different kind of WSNs. The reputation parameter can be calculated using the explicit and implicit reputation metrics. The explicit reputation is given a higher reputation than the sensor device which does the computing has more interaction with the target sensor device. Assume $\mathbb{C}_o^u(x, y)$ is the reputation parameter, such that the sensor device x has trust in the sensor device y, then:

$$\mathbb{C}_o^u(x, y) = \delta * \mathbb{L}_o^u(x, y) + (1 - \delta) * \mathbb{G}_o^u(x, y) \tag{11}$$

in (11), the δ gives the weight of the reputation parameter which is calculated using the (12):

$$\delta = \frac{\mathbb{T}^u(x,y)}{\mathbb{T}^u(x,y) + \vec{\mathbb{T}}^u(x,y)} \tag{12}$$

in (12), the $\mathbb{T}^u(x, y)$ is used to show the number of times the sensor device x has communicated with the sensor device y in the u^{th} session. The $\vec{\mathbb{T}}^u(x, y)$ is used to show the mean size of the communication that has been calculated when the sensor x and sensor y have a reputable connection. The $\vec{\mathbb{T}}^u(x, y)$ is calculated using the (13):

$$\vec{\mathbb{T}}^u(x, y) = \frac{\sum_{p\in Z-\{x\}} \mathbb{F}_o^u(x,p) * \mathbb{T}^u(p,y)}{|Z - \{x\}|} \tag{13}$$

3.6. Past experience reputation evaluation

In this part, the past experience reputation evaluation is done using the intra-cluster communication and inter-class cluster communication used in a different kind of WSNs. As time flows, the current reputation parameter will change to the old reputation parameter. Like the reputation parameter is calculated, the experience parameter is calculated using the exponentially averaging update method. Let $\mathbb{L}_o^u(x,y)$ gives the past reputation parameter that the sensor device x has on the sensor device y:

$$\mathbb{L}_o^u(x,y) = \frac{\varphi * \mathbb{L}_{o-1}^u(x,y) + \mathbb{C}_{o-1}^u(x,y)}{2}, \tag{14}$$

in (14), the $\varphi(0 \le \varphi \le 1)$ is the award parameter and $\mathbb{L}_0^0(x,y) = 0$. From the past experience reputation parameter, the present malicious IoT devices communicating with the sensor device cannot give the result by looking at the past performance. For the current sensor device which is ideal, it has to communicate in a proper manner having a significant number of connections so that the reputation parameter can replace the past reputation parameter.

3.7. Anticipated reputation evaluation

In this part, the anticipated reputation evaluation is done using the intra-cluster communication and inter-class cluster communication used in a different kind of WSNs. To evaluate the anticipated reputation parameter, both the current reputation parameter and the past experience reputation parameter are used. Let $F_o^u(x,y)$ defines the anticipated reputation parameter that the sensor device y has on the sensor device x. We can calculate the anticipated reputation using the (15).

$$F_o^u(x,y) = \begin{cases} 0, if\ neither\ \mathbb{L}\ or\ \mathbb{C}\ is\ available \\ \alpha\mathbb{C}_o^u(x,y) + (1-\alpha)\mathbb{L}_o^u(x,y)\ if\ either \mathbb{L}\ or\ \mathbb{C}\ is\ available \end{cases} \tag{15}$$

Here, α is set to zero initially. Though, it can be changed according to the dynamic environment using the ω which is the deviation factor as shown as show in (16).

$$\alpha = \begin{cases} \alpha + 0.1, if\ \mathbb{C}_o^u(x,y) - \mathbb{L}_o^u(x,y) > \omega, \\ \alpha - 0.1, if\ \mathbb{C}_o^u(x,y) - \mathbb{L}_o^u(x,y) < -\omega, \\ \alpha, if\ -\omega < \mathbb{C}_o^u(x,y) - \mathbb{L}_o^u(x,y) < \omega. \end{cases} \tag{16}$$

Using the ω we can make it work in the dynamic environment, which will tell us how fast can the sensor device can come out from the past reputation parameter. The value of ω must be set to a very small value as the malicious sensor device can use this parameter to come out from their earlier malicious performance.

3.8. Unfair and oscillating reputation evaluation

In this part, unfair and oscillating reputation evaluation is done using the intra-cluster communication and inter-class cluster communication used in a different kind of WSNs. The biased attack can come from a malicious IoT device. The malicious IoT devices may purposefully vary between the states so that they can affect the reputation parameter which affects the overall performance of the network. We try to accumulate the unfair and oscillating reputation parameter to calculate the oscillation using the (17):

$$\mathbb{D}_o^u(x,y) = \begin{cases} \mathbb{D}_{o-1}^u(x,y) + \frac{\mathbb{C}_o^u(x,y) - \mathbb{L}_o^u(x,y)}{\rho}, \mathrm{if} \mathbb{C}_o^u(x,y) - \mathbb{L}_o^u(x,y) > \tau \\ \mathbb{D}_{o-1}^u(x,y) + \mathbb{L}_o^u(x,y) - \mathbb{C}_o^u(x,y), \mathrm{if} \mathbb{C}_o^u(x,y) - \mathbb{L}_o^u(x,y) > -\tau \\ \mathbb{D}_{o-1}^u(x,y), \mathrm{otherwise}, \end{cases} \tag{17}$$

in (17), τ defines the acceptance parameter of the reputation error in evaluating the reputation parameter. The $\rho(\rho > 1)$ defines the penalty parameter for the acceptance in the reputation parameter. Hence the unfair and oscillating reputation of the sensor device can be calculated using (17) and the equation can be established as follows:

$$\overline{\overline{\mathbb{D}}}_o^u(x,y) = \begin{cases} 0, \mathrm{if}\ \mathbb{D}_o^u(x,y) > \mathbb{D} \\ \cos\left(\frac{\pi}{2} * \frac{\mathbb{D}_o^u(x,y)}{\max \mathbb{D}_o^u(x,y)}\right), \mathrm{otherwise}, \end{cases} \tag{18}$$

3.9. Reputation metric for detecting malicious IoT device

In this part, the secure metric reputation evaluation is done using the intra-cluster communication and inter-class cluster communication used in a different kind of WSNs. The security metric for the reputation parameter $\mathcal{F}_o^u(x, y)$ is calculated using the anticipated reputation parameter and the oscillating reputation parameter and is calculated using (19).

$$\mathcal{F}_o^u(x, y) = F_o^u(x, y) * \overline{\overline{\mathbb{D}}}_o^u(x, y) \tag{19}$$

From (19), it can be said that the sensor device having the highest anticipated reputation parameter with less oscillation reputation parameter results in having less overall reputation results. The sensor device which purposefully changes its state between the oscillating reputation parameter will have less reputation parameter because of the less oscillating reputation parameter. If a sensor device wants to gain a higher accumulated reputation parameter, then it should not show any oscillation of the reputation parameter. Hence (19) is used to select the sensor device having a higher reputation parameter which is composed of all the security parameter that is required to achieve the effective security method for a different kind of WSN application.

3.10. Secure communication metric for big data collection environment

In this part, the secure communication metric for the big data collection environment is evaluated. From (19), we detect the good sensor device and malicious sensor device. When we start sending the packets to a sensor device that is currently ideal, it consumes energy above the sensor device in the cluster network having a higher reputation parameter. Therefore, to balance the load between the cluster head, first the evaluation of the traffic i.e., consumption of energy above the sensor device is calculated using the (20).

$$\mathcal{T}^u(x, y) = \mathbb{T}^u(x, y) + \textstyle\sum_{p \in Z-\{x\}} \mathbb{F}_o^u(x, p) * \mathbb{T}^u(p, y) \tag{20}$$

After the evaluation of the traffic, the selection of the cluster head to transmit the packet is evaluated using the (21):

$$\min \textstyle\sum_{p \in Z-\{x\}} \mathcal{T}^u(x, p) \tag{21}$$

suppose, if any new sensor does not have a reputation value, then the probability of the sensor device is calculated using the (22).

$$\mathcal{P}^u(x, y) = \begin{cases} \frac{\mathcal{F}_o^u(x,y)}{\sum_{p \in V} \mathcal{F}_o^u(x,p)}, \text{if } \sum_{p \in V} \mathcal{F}_o^u(x, p) \neq 0, \\ \text{arbitrarily choose any sensor device, else.} \end{cases} \tag{22}$$

To select the sensor device from V, this model uses (22) having a high reputation parameter which has a high probability to get selected. When the reputation parameter is set to zero or is zero then the sensor device is selected randomly. The proposed method has a performance higher when compared to the existing reputation-based models and is shown in the results and discussions section.

4. RESULTS AND DISCUSSIONS

In this section, the results have been evaluated using a simulator. In the simulator, various parameters are considered for simulation. Each network parameter has its value which has been set accordingly. The Network simulation area considered for the detection of the attack was 100×100m. The number of edge servers considered was 4 servers. The number of IoT devices was considered from the range of 500 to 1000 devices. As the IoT devices kept increasing the value of the number of malicious IoT devices was set to 10, 20, 30, and 40 percent. The MAC protocol used in the simulation process was IEEE 802.11b. The ratio propagation of each device was set to 6 meters and the sensing range of each device was set to 3 meters. Each sensor device's energy was set in the range from 0.05 to 0.2 Joules. The dissipation of the radio energy was set to 50 ni/bit. The length of the control packet and data packet was set to 248 bits and 2,000 bits respectively. The transmission speed of the IoT device was 100 bits and the bandwidth was set to the value of 1,0000 bits. The sensing event of each device was 0.1 seconds. The Idle energy consumption of each IoT device was set to the value of 50 ni/bit and the amplification energy was set to 100 pJ/bit/m2. The model has been compared with the existing reputation-based security (ERS) models [16], [22].

4.1. Attack detection rate vs Attack rate in sensor devices

In this section, the experimental results for the detection rate of a malicious attack on the IoT device have been given. In Figure 3. it can be seen that the reputation-based biased attack detection (RBAD) detects more malicious attacks in the IoT devices when compared with the existing ERS model. In the Figure. 3 the RBAD model performs 16%, 29%, 19.5% and 19.48% better than the existing ERS model for 10%, 20%, 30% and 40% attack rate respectively.

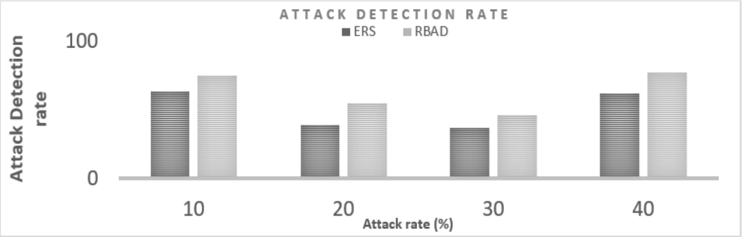

Figure 3. Attack detection rate vs attack rate

4.2. Attack detection failure rate vs Attack rate in sensor devices

In this section, the experimental results for the failure of the attack detection rate in the IoT device have been given. In Figure 4. The RBAD shows less failure to detect malicious attacks in the IoT device when compared with the existing ERS model. In Figure 4. the RBAD model performs 32.4%, 26.2%, 14.28 and 39.4% better than the existing ERS model for 10%, 20%, 30%, and 40% attack rates respectively.

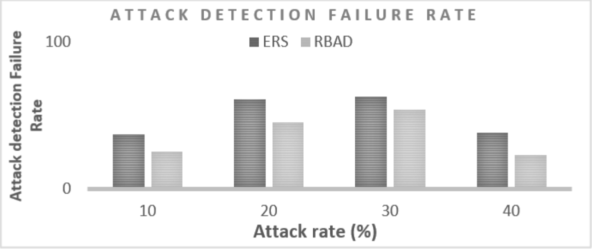

Figure 4. Attack detection failure rate vs attack rate

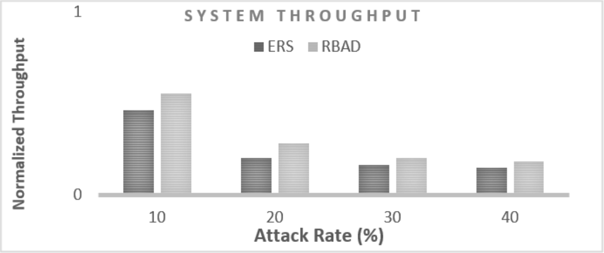

Figure 5. Throughput vs attack rate

4.3. Performance of the system throughput of the model

In this section, the experimental results of the system throughput in the IoT devices have been given. Figure 5. shows the performance of the throughput of the RBAD model and the ERS model. The RBAD model shows a higher throughput when compared with the existing ERS model. The ERS model showed a throughput of 0.46, 0.20, 0.16, 0.14 whereas the RBAD model showed a throughput of 0.55, 0.28, 0.20, 0.18 for 10%, 20%, 30% and 40% attack rate respectively. From the results, it can be clearly stated that our model performs better than the existing model.

4.4. Dead nodes vs attack rate

In this section, the experimental results for the dead nodes vs attack rate in the IoT device have been done. In Figure 6 The number of dead nodes with an attack rate of 10% in the IoT device is shown. Similarly Figures 7-9. show the Number of dead nodes with an attack rate of 20%, 30%, and 40% respectively. In all the figures the number of dead nodes concerning the simulation time has been plotted and it can be seen that our model reduces the dead nodes concerning the time which provides more security to the model.

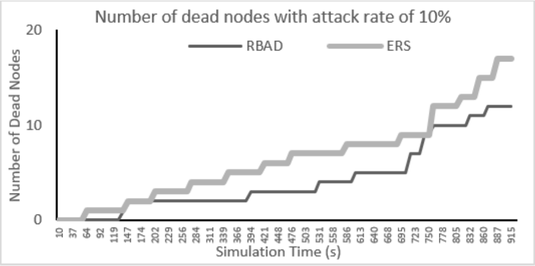

Figure 6. Number of dead nodes with an attack rate of 10%

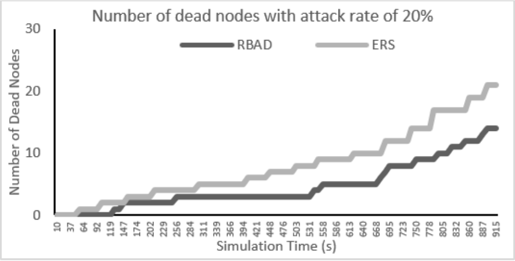

Figure 7. Number of dead nodes with an attack rate of 20%

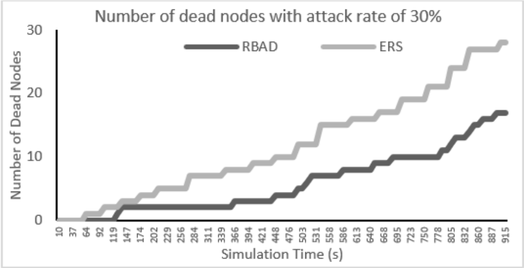

Figure 8. Number of dead nodes with an attack rate of 30%

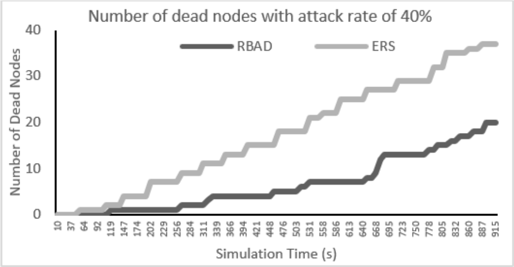

Figure 9. Number of dead nodes with an attack rate of 40%

5. CONCLUSION

In this paper, we have presented a model of RBAD to detect the biased attack in big data. In this method, the reputation-based method has been proposed. Experiments on the model using the data of the IoT devices model shows better performance when compared with the existing ERS model. Various results have been discussed about the attack rate, attack failure rate, system throughput of the model, and dead nodes vs attack rate. The results have attained an outcome that the RBAD model has performed well when compared with the existing ERS model. Hence, the RBAD model provides more security to the IoT devices in the big data.

REFERENCES

[1] N. Shaikh, "Internet of things (IoT), applications, open issues, and challenges: a comprehensive survey.," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 6, pp. 4540–4545, Jun. 2022, doi: 10.22214/ijraset.2022.44990.

[2] Z. El-Rewini, K. Sadatsharan, N. Sugunaraj, D. F. Selvaraj, S. J. Plathottam, and P. Ranganathan, "Cybersecurity attacks in vehicular sensors," *IEEE Sens. J.*, vol. 20, no. 22, pp. 13752–13767, Nov. 2020, doi: 10.1109/JSEN.2020.3004275.

[3] M. Aljabri et al., "Intelligent techniques for detecting network attacks: review and research directions," *Sensors*, vol. 21, no. 21, p. 7070, Oct. 2021, doi: 10.3390/s21217070.

[4] R. Estepa, J. E. Diaz-Verdejo, A. Estepa, and G. Madinabeitia, "How much training data is enough? a case study for HTTP anomaly-based intrusion detection," *IEEE Access*, vol. 8, pp. 44410–44425, 2020, doi: 10.1109/ACCESS.2020.2977591.

[5] O. Sadio, I. Ngom, and C. Lishou, "Lightweight security scheme for MQTT/MQTT-SN protocol," in *2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*, Oct. 2019, pp. 119–123. doi: 10.1109/IOTSMS48152.2019.8939177.

[6] L. Canuto, L. Santos, L. Vieira, R. Goncalves, and C. Rabadao, "CoAP flow signatures for the internet of things," in *2019 14th Iberian Conference on Information Systems and Technologies (CISTI)*, Jun. 2019, pp. 1–6. doi: 10.23919/CISTI.2019.8760759.

[7] S. Sen and C. Jayawardena, "Cybersecurity and network performance modeling in cyber-physical communication for big data and industrial IoT technologies," in *2019 IEEE Bombay Section Signature Conference (IBSSC)*, Jul. 2019, pp. 1–8. doi: 10.1109/IBSSC47189.2019.8973080.

[8] F. Jemili and H. Bouras, "Intrusion detection based on big data fuzzy analytics," in *Open Data, IntechOpen*, 2022. doi: 10.5772/intechopen.99636.

[9] R. Arthi and S. Krishnaveni, "Design and development of IOT testbed with DDoS attack for cyber security research," in *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*, May 2021, pp. 586–590. doi: 10.1109/ICSPC51351.2021.9451786.

[10] L. Mohan, S. Jain, P. Suyal, and A. Kumar, "Data mining classification techniques for intrusion detection system," in *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*, Sep. 2020, pp. 351–355. doi: 10.1109/CICN49253.2020.9242642.

[11] A. Lakshmanarao, A. Srisaila, and T. S. Ravi Kiran, "Machine learning and deep learning framework with feature selection for intrusion detection," in *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, Mar. 2022, pp. 1–5. doi: 10.1109/IC3IOT53935.2022.9767727.

[12] U. Ghugar and J. Pradhan, "NL-IDS: trust based intrusion detection system for network layer in wireless sensor networks," in *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Dec. 2018, pp. 512–516. doi: 10.1109/PDGC.2018.8745870.

[13] S. T. Alshammari and K. Alsubhi, "Building a reputation attack detector for effective trust evaluation in a cloud services environment," *Appl. Sci.*, vol. 11, no. 18, p. 8496, Sep. 2021, doi: 10.3390/app11188496.

[14] V. Desai and H. A. Dinesh, "Efficient reputation-based cyber attack detection mechanism for big data environment," *Indian J. Sci. Technol.*, vol. 15, no. 13, pp. 592–602, Apr. 2022, doi: 10.17485/IJST/v15i13.2102.

[15] W. Najib, S. Sulistyo, and Widyawan, "Survey on trust calculation methods in internet of things," *Procedia Comput. Sci.*, vol. 161, pp. 1300–1307, 2019, doi: 10.1016/j.procs.2019.11.245.

[16] L. Xiao et al., "A reinforcement learning and blockchain-based trust mechanism for edge networks," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5460–5470, Sep. 2020, doi: 10.1109/TCOMM.2020.2995371.

[17] X. Wang, X. Luo, M. Zhang, Z. Jiang, and X. Guan, "Detection and localization of biased load attacks in smart grids via interval observer," *Inf. Sci. (Ny).*, vol. 552, pp. 291–309, Apr. 2021, doi: 10.1016/j.ins.2020.12.027.

[18] M. Debe, K. Salah, M. H. U. Rehman, and D. Svetinovic, "IoT public fog nodes reputation system: a decentralized solution using ethereum blockchain," *IEEE Access*, vol. 7, pp. 178082–178093, 2019, doi: 10.1109/ACCESS.2019.2958355.

[19] M. Ghafoorian, D. Abbasinezhad-Mood, and H. Shakeri, "A thorough trust and reputation based RBAC model for secure data storage in the cloud," *IEEE Trans. Parallel Distrib. Syst.*, vol. 30, no. 4, pp. 778–788, Apr. 2019, doi: 10.1109/TPDS.2018.2870652.

[20] H. Liu, P. Zhang, G. Pu, T. Yang, S. Maharjan, and Y. Zhang, "Blockchain empowered cooperative authentication with data traceability in vehicular edge computing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4221–4232, Apr. 2020, doi: 10.1109/TVT.2020.2969722.

[21] X. Zhang, R. Lu, J. Shao, H. Zhu, and A. A. Ghorbani, "Secure and efficient probabilistic skyline computation for worker selection in MCS," *IEEE Internet Things J.*, vol. 7, no. 12, pp. 11524–11535, Dec. 2020, doi: 10.1109/JIOT.2020.3019326.

[22] J. Yuan and X. Li, "A reliable and lightweight trust computing mechanism for iot edge devices based on multi-source feedback information fusion," *IEEE Access*, vol. 6, pp. 23626–23638, 2018, doi: 10.1109/ACCESS.2018.2831898.

[23] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "Secure computation offloading in blockchain based IoT networks with deep reinforcement learning," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 4, pp. 3192–3208, Oct. 2021, doi: 10.1109/TNSE.2021.3106956.

[24] S. Xu, C. Guo, R. Q. Hu, and Y. Qian, "Blockchain-inspired secure computation offloading in a vehicular cloud network," *IEEE Internet Things J.*, vol. 9, no. 16, pp. 14723–14740, Aug. 2022, doi: 10.1109/JIOT.2021.3054866.

[25] V. Mohammadi, A. M. Rahmani, A. M. Darwesh, and A. Sahafi, "Trust-based recommendation systems in internet of things: a systematic literature review," *Human-centric Comput. Inf. Sci.*, vol. 9, no. 1, p. 21, Dec. 2019, doi: 10.1186/s13673-019-0183-8.

BIOGRAPHIES OF AUTHORS

Vinod Desai is presently working as Assistant Professor in Dept. Of Computer Science and Engineering (Artificial Intelligence and Machine Learning) in B.L.D.E. A's V.P. Dr.P.G.Halakatti College of Engineering and Technology Vijayapura. He has 11 years of teaching experience and 1 year of Industry experience as quality analyst. He is pursuing his Ph.D in Cyber Security and Machine Learning under Visvesvaraya Technological University Belagavi in Computer Science and Engineering. He has completed his MTech and BE from VTU Belagavi. He published 3 patents. He published various national and international journals. He presented papers in various conferences like IEEE. He can be contacted at email: desaivinod2021@gmail.com.

Dr. Dinesha Hagare Annappaiah is a fledgling quixotic entrepreneur indulged passionately in serving and uplifting society through advance research, innovation, and knowledge imparting in the arena of Engineering and Technology. He is the founder of an IT industry named Cybersena (R&D) India Private Limited, India in the domain of Cybersecurity and forensic investigations. He is presently a Post-Doctoral fellow in the Cyber forensic domain and also working as HOD, CSE, NCET, Bangalore. He has completed his research in Ph.D. (Cloud Computing Security), Post-Graduation in MTech in Software Engineering and Under Graduation in B.E., and Diploma in Computer Science and Engineering. He can be contacted at email: sridini@gmail.com.