Journal of Scientific & Industrial Research
Vol. 82, May 2023, pp. 522-528
DOI: 10.56042/jsir.v82i05.1084

Enhancing Network Forensic and Deep Learning Mechanism for Internet of Things Networks

J Avanija[1], K E Naresh Kumar[2], Ch Usha Kumari[3], G Naga Jyothi[4], K Srujan Raju[5] & K Reddy Madhavi[1]*

[1]School of Computing, Mohan Babu University, Tirupati, 517 102, Andhra Pradesh, India

[2]Department of CSE, RGMCET, Nandyal 518 501, Andhra Pradesh, India

[3]Dept of ECE, GRIET, Hyderabad, 500 090, Andhra Pradesh, India

[4]Department of ECE, Madanapalli Institute of Technology and Science, Madanapalli 517 325, Andhra Pradesh, India

[5]Department of CSE, CMR Technical Campus, 501 401, Hyderabad, Telangana, India

Received 03 June 2022; revised 12 September 2022; accepted 06 October 2022

The integration of intelligence into everyday products has been possible due to the ongoing shrinking of hardware and a rise in power efficiency. The Internet of Things (IoT) area arose from the tendency to add computational capabilities to so-called non-intelligent daily items. IoT systems are attractive targets for cyber-attacks because they have many applications. Adversaries use a variety of Advanced Persistent Threat (APT) strategies and trace the source of cyber-attack events to safeguard IoT networks. The Particle Deep Framework (PDF), which is proposed in this study, is a novel Network Forensics (NF) that encompasses the digital investigative phases for spotting & tracing attack activity in IoT networks. The suggested framework contains three novel functionalities for dealing with encrypted networks, such as collecting network data flows & confirming their integrity, using a PSO algorithm, "Bot-IoT "& "UNSW NB15" datasets. The suggested PDF is related to several deep-learning methods. Experimental outcomes show that the proposed framework is very good at discovering & tracing cyber-attack occurrences when compared to existing approaches. The proposed design is implemented using neural network technology. The proposed design has 10% accuracy when compared with the existing structure. This paper is expected to offer a quick reference for researchers interested in understanding the use of network forensics and IOT.

Keywords: Attack tracing, Botnets, IOT, Network forensics, Particle swarm optimization

Introduction

The Internet of Things exploded in popularity in recent years. According to Gartner, the number of connected IoT devices worldwide is estimated to reach 20.4 billion in 2020, up 145% from 2017. This new diversified area will continue growing as businesses benefit from IoT services. The IoT market has seen rapid expansion in recent years, with estimates indicating that this trend will continue.[1] The smart home sector had 664 million in 2017, which is one of the most popular IoT applications. Innovative home applications include smart lighting, refrigerators, stoves, thermostats, & locks, to name a few. The 'smart city' is another IoT application that has been envisaged for various European countries on a greater scale. Cost and precision are essential drivers of the current trends in industrial, agricultural, & health applications. It's difficult to defend against such cyber-attacks since there are no widely acknowledged design standards for IoT devices.[2] Several protocols, like Lo Ra & Zigbee, could interact in an IoT deployment, increasing complexity & heterogeneity.[3–5] Several forensic methods have been offered to solve these problems, but they do not consider all aspects of the inquiry.[6–8] Most of them employ a public ledger system in which diagnostic & communication data is shared among various organizations, including police & insurance companies. One defining feature of IoT devices is always on.

As a result, gathering network flow from an IoT network produces a significant amount of data. Automated processes are frequently used to undertake analysis of the obtained data to eliminate human mistakes, with deep learning being one such prominent automated way. An investigator can quickly find patterns in network data that denote the occurrence of an attack using deep learning models.[9,10] Furthermore, a machine learning model

*Author for Correspondence
E-mail: kreddymadhavi@gmail.com

must first be trained before it can be employed. DDO attacks between 2018 and 2023 are shown in Fig. 1. Data must be used, and settings for the model's hyperparameters must be chosen while training a model. Both data and hyperparameters are significant during the training phase. However, the hyperparameter is reduced since it specifies the model's abstract structure & training conditions. Nevertheless, the scientific community has approved no single-method strategy as the optimal method. The NF procedures-based optimization is required to promptly analyze security occurrences in IoT networks.

The important contributions of the paper are listed below. The first was PDF, a revolutionary NF framework. Then, based on PSO, which is employed to estimate the hyperparameters of "Deep Neural Network" (DNN). Finally, the performance of created DNN is evaluated & compared to that of other classification models. "Bot-IoT" & the "UNSW-NB15" were employed for the evaluations & comparisons.[11,12] Several research scholars have done surveys on IOT and NF, which are listed in Table 1.

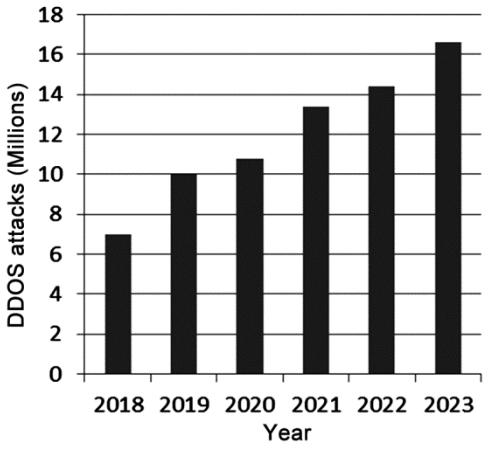

Fig. 1 — DDOS attacks from 2018-2023

The shadow-based IOT technique proposed is useful for limited applications.[13] And also author explained how digital forensics is useful for IOT designs and the drawbacks of using shadow technique-based IOT. When we observe Fig. 1, there is a timeline between the IOT and people's usage. There was a vast increase in IOT usage; meanwhile, security is a drawback. We cannot hide the document quickly, and the networks should have high security to hide the data to overcome all these drawbacks. Some more authors are also explained in detail in Table 1. In the present era, network forensics for IOTs using deep learning methods plays a paramount role in the development of security and mobile networking applications. The following three sorts of interactions can be found in an IoT-based system.

T2U: Things to Users

a) A user accesses an IoT device using a cloud service to remotely control a device (target) (intermediate) and gateway (intermediate), or the other way around;
b) Local IoT device through a gateway or the other way around.

T2C: Things to Cloud: An Internet of Things device posts via gateways

Relevant Study

Digital Forensics (DF) conducting in-depth investigations of attacks & collecting traces left by intruders after suspicious events are detected. Forensics is carried out to investigate attacks & collect traces left by the attackers. These investigations help prove the authenticity of the attack and prevent future threats.[13] The 5 main forensics steps are:

(a) Identification phase: This gives information that either crime has been present or not. This method identifies and detects by IDS.

Table 1 — Literature survey

Paper	Methods	Importance
Smart DF for IOT[13]	Shadow IOT technique	An in-depth analysis of numerous tools and strategies for a quick digital inquiry framework is provided by this study.
Roadmap of DF[14]	Forensic methods	The MLP model gives PDF a 97.9% accuracy rate.
A distributed logging method for IoT forensics[15]	PSO technique	It is used for digital forensic finding
Findings of IoT Devices for Forensic Investigation [16]	Neural network	—
A distributed logging scheme for IoT[17]	Malware detection methods	IoT devices lack adequate security, making them increasingly susceptible to malware.
A novel framework based on deep learning for Internet of Things networks: A particle deep framework[18]	DNA-based IOT technique	—

(b) Evidence Collection phase: In this, forensic experts find evidence from cloud service models such as SaaS, IaaS, and PaaS.
(c) Analysis phase: This forensic expert observes the report, then correlates and comes to a conclusion.
(d) Preservation phase: It protects data integrity and necessitates a high volume of data will be preserved for further verification with high security.
(e) Reporting phase: Forensic people investigate based on the information given in this stage.

The Application domain is used to categorize DF. A DF in the condition of network management is termed NF. A DF in the condition of cloud computing is termed Cloud Forensic (CF). A DF in the condition of the web is termed web forensic, a DF in the context of mobile is termed mobile forensic, and finally, for internet for things is termed as IOT forensic. Several researchers have created forensic frameworks for the IOTs, and different types of forensics are shown in Fig. 2.[13–15] Cebe *et al.* Created a Block 4 Forensic acquisition methodology for collecting vehicle data. IoT Dots is a new acquisition model created by Babun *et al*.[16,17]

Particle Swarm Optimization (PSO)

It is a swarm-based optimization technique first established by Eberhart.[18] It is called metaheuristic because it doesn't make any assumptions about the underlying problem and is used to find "good enough" solutions in an acceptable amount of time.[18] PSO is frequently used to calculate the value of a variable.

PARTICLE DEEP FRAMEWORK

The stages of the novel NF framework, dubbed PDF, based on PSO with deep learning in drawing attack origins & detecting them in IoT networks are shown in Fig. 3a. The algorithm stages are:

a. Network Capturing: The devices were set up in promiscuous mode, allowing them to perceive traffic on the local network. Network packet captures are performed using network capturing programmers like Wireshark, tcpdump, and Ettercap.
b. Data Collection: BoT-IoT and UNSW-NB15 are examples of data sets of this stage where data is captured in a manner that can be further processed and scrutinized.
c. PSO adapting hyperparameters: Compared to other evolutionary algorithms, the PSO algorithm was chosen to adapt hyperparameters of the deep model because the situation can readily discourse local-optimum problematic and swiftly converges to acquire the best fitness values.
d. MLP for attack identification: The MLP deep learning algorithm is trained and tested using the hyper-parameters determined by the PSO algorithm and data collected in Stage 2. The MLP was used by 7 stages, with the No. of neurons: 20, 40, 60, 80, 40, 10, and 1, as best outputs in view

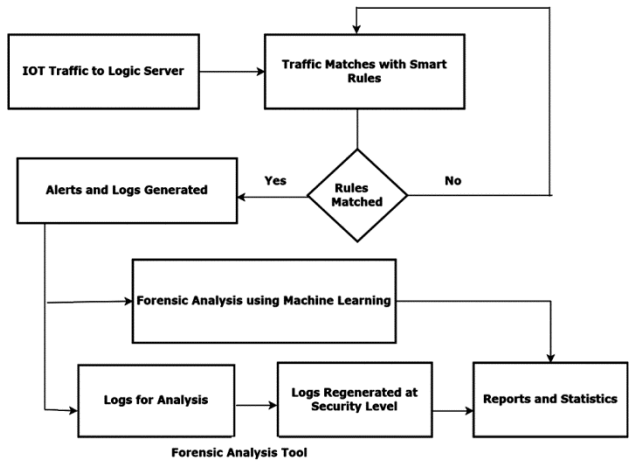

Fig. 2 — Types of Forensics in IOT

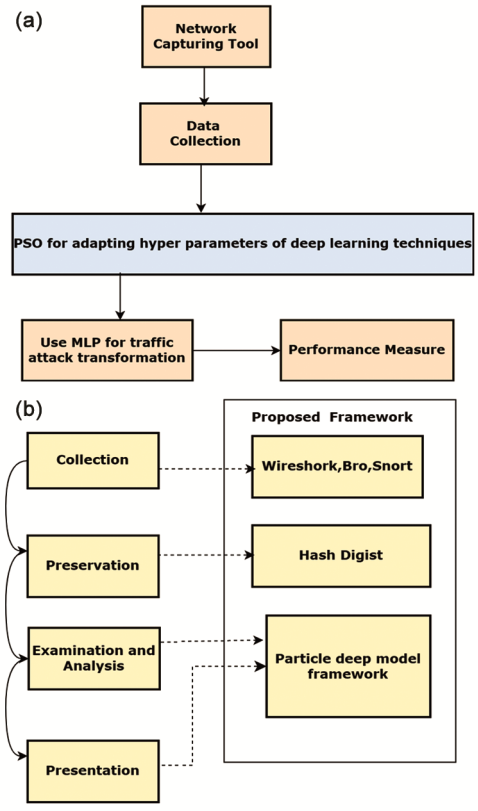

Fig. 3 — Proposed (a) NF using PSO and (b) NF stages

of the accuracy of detecting the false alarm rate produced.

e. Performance: Finally, by executing the trained deep MLP model with Testing & Validating data and acquiring performance measures.

PSO Algorithm for DL Parameter Estimations

It is a metaheuristic evolutionary method that generates a predetermined number of Particles (P) begun randomly, set to traverse the search space of a variable. It was inspired by observations of animal swarming in its natural environment (v) During a particle's propagation across the search space, the output of an objective function is assessed at each new point $-V_{t+1}$, which may vary depending on the issue being optimized. As indicated in Eqs 1–4, the particle is defined by four values: velocity- vt, current position- xt, local best- xlbest, and swarm best position -xgbest.

$$P = p_1, p_2, \ldots\ldots p_n, n \epsilon N \quad \ldots (1)$$

$$\forall p_n \in P, Pn = \{x_t, v_t, xlbest, xgbest\} \quad \ldots (2)$$

$$V_{t+1} = V_t + \theta_1 * rand * (xlbest - x_t) + \theta_2 * rand * (xgbest - x_t), rand\varepsilon[0,1] \quad \ldots (3)$$

$$X_{t+1} = v_t + x_t \quad \ldots (4)$$

In the above equation, the term v_{t+1}, refers to particle's new velocity is determined by prior velocity vt, (random) fraction of learning rates 1, 2, and the (distance) between its present position and the swarm's best places. Based on its prior position x_t and current velocity, Eq. 4 calculates a particle's updated position x_{t+1}. For hyperparameter modification, there are various reasons to employ the PSO algorithm rather than another evolutionary metaheuristic technique. The PSO is explained in detail in Algorothm.1.

Proposed PDF for NF

The innovative PDF is a significant addition to the NF discipline, overlapping stages of NF, collection, investigation, preservation, presentation, and analysis, as shown in Fig. 3(b).

Algorithm: 1 PSO maximum Algorithm

Step 1: P← No. of particles
Step 2: $\forall p \in P$, $p.Xlbest = p.x_0$, $p.Xgbest = \infty$
Step 3: epoch←load_epoch
Step 4: e←0
Step 5: While e < epochs
do
for each p∈P
do
Vt + 1 = Vt + θ1* rand * (Xlbest − xt) + θ2 * rand *(Xgbest − xt), randε[0,1]
Xt+1 = vt +xt
if X_{t+1} > Xlbest then
$X_{lbest} = X_{t+1}$
End
Step 6: If X_{t+1} > Xlgest then
Xgbest = X_{t+1}
End
End
End
Return
P.globle.best()

Suggested framework has the advantage of the various layers of a deep NN, which improve the model's performance while keeping the execution time under control. In Algorithm 2, we illustrate PDF iteration. The pre-selected layers & no. of neurons are first loaded into the Neural Network (NN). At first, the batch size, No. of epochs, and learning rate hyperparameters [b,e,lr] all randomly started. Then, with a predetermined No. of particles & iterations a particle swarm is produced.

Algorithm 2: Particle deep model for hyper-parameter estimation of deep learning

Data: nn ← load NN design();
[b,e,lr] ← initialize random hyper-parameters();
Hyper-parameters ← [b,e,lr];
PS ← construct ps(n particles, swarm epochs);
i ← 0;
for each h1 ∈ hyper parameters
do
while PS.swarm epochs ≠ 0
do
h1 ← P S. maximize (nn.AUC, h1) using algorithm1;
end
nn.save opt hyperparam((h1));
end
nntrain NN training set();

The optimized hyperparameters' value is then obtained using process 1, & the NN's AUC value is maximized (h1P S. maximize (nn. AUC, h1)). The approach is repeated on each hyperparameter to be optimized, with the final NN being trained using the determined values. We evaluated the optimized deep MLP model using the "Bot-IoT" dataset, which is a recent dataset that includes both IoT & non-IoT traces & assaults.[19] Data was divided into 2 sets: training and testing, each with 80% and 20% of the total.

Table 2 — Evaluation measures

	Optimized NN with compressed input	Unoptimized NN	Optimized NNwith 13-features input
Neurons per layer	1,10,80,60,40,10,201	11,04,08,06,04,02,013	11,04,08,06,04,02,013
F-measure	0.9730	0.9990	0.9990
Recall	0.9470	0.9990	0.9990
FPR	0.8100	0.8840	00
FNR	0.0520	$9.260*10^{-5}$	$9.540*10^{-5}$
Batch size	30640	3500	732.0
Learning rate	0.00150	0.20	0.00150
Precision	0.9990	0.9990	1
Accuracy	0.9470	0.9990	0.9990

With min-max in the range [0,1], the training & testing sets were standardized. Because there is no conventional technique for picking optimal hyperparameters, we manually selected the values & trained the deep MLP model first, then used the PDF as described in the Algorithm.

Results and Discussion

The "Bot-IoT" combines "IoT & non-IoT" traffic to mimic smart home positioning. The former is produced using Node-Red and complete collection for training & testing of suggested PDF. The current Bot-IoT dataset was chosen.[19–22] totaling 72,00,00,000 records & 16.7 GB in CSV format. We used the Bot-IoT dataset's "10-best feature" version. We chose 6 measures for evaluation as shown in Table. 2. They are recall, accuracy, FPR, F-measure, precision & FNR.[22]

Precision: It is the fraction calculated by the formulae

Precision = TP/(FP +TP)

Accuracy: It is calculated by

Accuracy = (TN +TP)/(TN +TP + FN + FP).

The FPR & FNR are fractions of records mistakenly categorized as "positive" (FP/(TN + FP)) or "negative"

(F N/(TP + FN)).

Recall: Fraction of records that is calculated by

Recall = T P/(T P + F N).

The evaluation methods of different NN with different compressor outputs are illustrated in Table 2. The F-measure value is very low for the optimized NN with compressed input and is high for the 13 features input. Also, the accuracy rate is very high in the optimized NN of 13 features input. Also, in this, precision is good. The FPR & FNR of these were also calculated. The unoptimized NN and NN with compressed input that has been optimized are shown in Fig. 4 (a & b) . The proposed result has a high accuracy rate. A 13-input optimized NN is shown in Fig. 5 while Fig. 6 compares the proposed deep learning for PDFs with the existing design. The

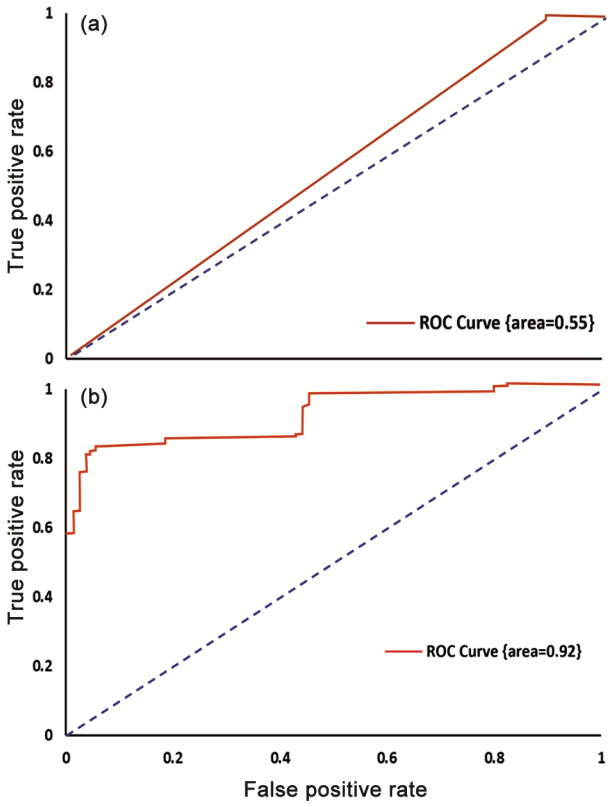

Fig. 4 — (a) Unoptimized NN (b) NN with compressed input that has been optimized

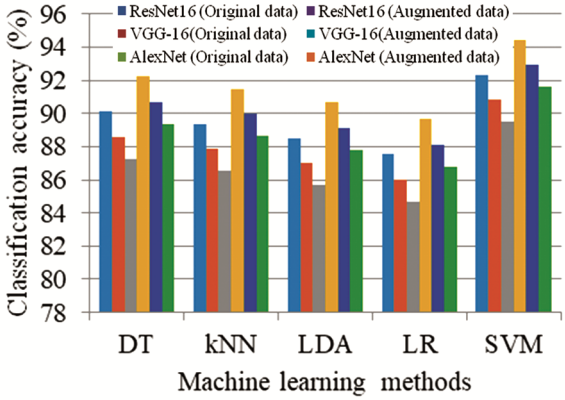

Fig. 5 — A 13 input optimized NN

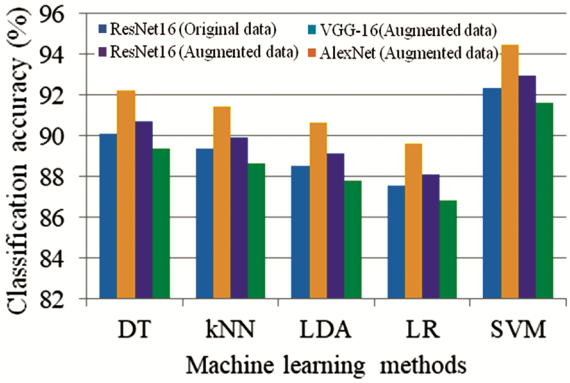

Fig. 6 — Proposed deep learning for PDFs

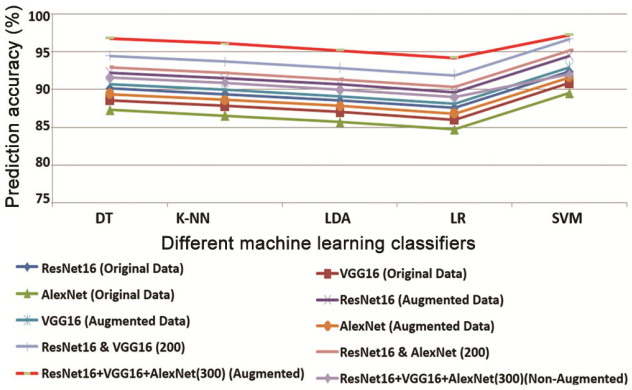

Fig. 7 — Accuracy graph for NN

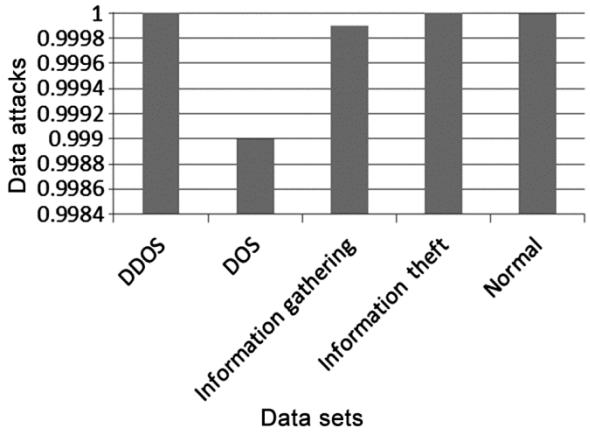

Fig. 8 — Bot-IoT dataset of data attack

accuracy graph comparison of different NN is demonstrated in Fig. 7. The Bot-IoT dataset of data attack is shown in Fig. 8, and Fig. 9 shows the accuracy graph of other RBF-NN cases.

The proposed NN has a high accuracy rate when compared with other designs. The proposed deep learning for PDF has a 96% accuracy rate when compared with other DT, KNN, LDA, SVM, and LR

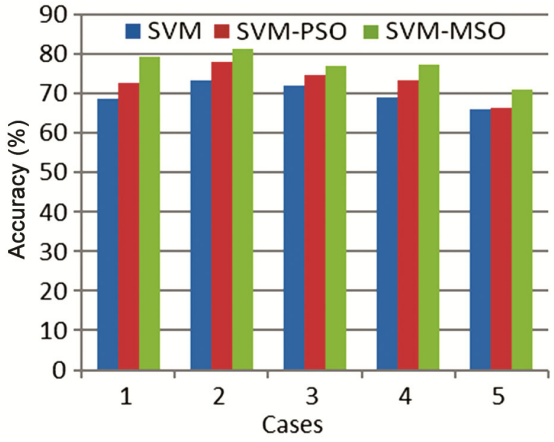

Fig. 9 — Accuracy graph of different RBF-NN cases

methods, shown in Fig. 6. In this, different data sets of DDOS, DOS, and normal data sets are considered and are passed to the proposed PDF NN design. For this, we have considered five different cases using RBF-NN. Case-1 has 15 data sets, and case-2 has considered 30 data sets and so on.

Conclusions

As a result of the fast use of IoT systems by humans & industry, attacks on IoT networks have proliferated. The PDF is a novel NF framework for the detection & analysis of cyber-attacks in IoT networks, which is presented in this paper. The components of the PDF were first discussed, as well as their relevance to the forensic stages. The PDF's fundamental technique uses Deep Learning as the underlying model & PSO to modify its hyper-parameters, with the Bot-IoT as validation. The PDF can classify documents at a rate of 14,762 per second & has 99.9% attack detection accuracy. It also has a minimal no. of false negatives & positives. We plan to expand the PSO's capabilities in the future by modifying it to analyze numerous hyper-parameters. The proposed design has 10% more accuracy when compared with the existing design.

Conflict of Interest

The authors declare no conflict of interest.

References

1 Koroniotis N & Moustafa N, Enhancing network forensics with particle swarm and deep learning: The particle deep framework, *arXiv preprint arXiv:2005.00722* (2020).

2 Koroniotis N, Moustafa N & Sitnikova E, Forensics and deep learning mechanisms for botnets in the internet of things: A survey of challenges and solutions, *IEEE Access*, **7** (2019) 61764–61785.

3 Ronen E, Shamir A, Weingarten A O & O'Flynn C, IoT goes nuclear: Creating a ZigBee chain reaction, *2017 IEEE Symposium on Security and Privacy* (IEEE), 2017.

4 Meffert C, Clark D, Baggili I & Breitinger F, Forensic state acquisition from internet of things (FSAIoT) A general framework and practical approach for IoT forensics through IoT device state acquisition, *Proc 12th Int Conf Avail Reliability, and Security* (2017), https://doi.org/10.1145/3098954.3104053.

5 Raghunath K K M, Koti M S, Sivakami R, Kumar V V, NagaJyothi G & Muthukumaran V, Utilization of IoT-assisted computational strategies in wireless sensor networks for smart infrastructure management, *Int J Syst Assur Eng Manag* (2022) 1–7, https://doi.org/10.1007/s13198-021-01585-y.

6 Hassan M A, Samara G & Fadda M A, IoT Forensic Frameworks (DFIF, IoTDOTS, FSAIoT): A Comprehensive Study, arXiv preprint arXiv:2203.15705 (2022).

7 Hossain M, Karim Y & Hasan R, FIF-IoT: A forensic investigation framework for IoT using a public digital ledger, *2018 IEEE Int Cong Internet Things* (IEEE) 2018.

8 Hossain M, Hasan R & Zawoad S, Probe-IoT: A public digital ledger based forensic investigation framework for IoT, *IEEE INFOCOM 2018 - IEEE Conf Comput Commun Workshops* (Honolulu, HI, USA) 2018, 1–2, doi: 10.1109/INFCOMW.2018.8406875.

9 Shone N, Ngoc T N, Phai V D & Shi Q, A deep learning approach to network intrusion detection, *IEEE Trans Emerg Topics Comput*, **2(1)** (2018) 41–50.

10 Prabakaran S & Mitra S, Survey of analysis of crime detection techniques using data mining and machine learning, *J Phys Conf Ser* (IOP Publishing) **1000(1)** (2018).

11 Koroniotis N, Moustafa N, Sitnikova E & Turnbull B, Towards the development of realistic botnet dataset in the internet of things for network forensic analytics: Bot-iot dataset, *Future Gener Comput Syst*, **100** (2019) 779–796.

12 Moustafa N & Slay J, UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set), *2015 Military Commun Inf Syst Conf* (IEEE) 2015.

13 Cebe M, Erdin E, Akkaya K, Aksu H & Uluagac S, Block4forensic: An integrated lightweight blockchain framework for forensics applications of connected vehicles, *IEEE Commun Magaz*, **56(10)** (2018) 50–57.

14 Babun L, Sikder A K,Acar A & Uluagac A S, Iotdots: A digital forensics framework for smart environments, arXiv *preprint arXiv:1809.00745* (2018).

15 Yuan X, Li C & Li X, Deep Defense: identifying DDoS attack via deep learning, *2017 IEEE Int Conf Smart Comput* (IEEE) 2017.

16 Brun O, Yin Y, Gelenbe E, Kadioglu Y M, Augusto-Gonzalez J & Ramos M, Deep learning with dense random neural networks for detecting attacks against IoT-connected home environments, in *Security in Computer and Information Sciences*, Euro-CYBERSEC 2018, (Springer International Publishing London, UK) 2018, 79–89. https://doi.org/10.1007/978-3-319-95189-8_8

17 Kennedy J & Eberhart R, Particle swarm optimization, *Proc ICNN'95- IEEE Int Conf Neural Netw* (IEEE) 1995, 1942–1948, DOI: 10.1109/ICNN.1995.488968

18 Wang D, Tan D & Liu L, Particle swarm optimization algorithm: an overview, *Soft Comput*, **22(2)** (2018) 387–408.

19 Parsopoulos K E, Particle Swarm Methods, in *Handbook of Heuris*, edited by R Martí, P Panos & M Resende (Springer Cham) 2015, DOI 10.1007/978-3-319-07153-4_22-1.

20 Elbagoury M B, Maskeliunas R & Salem A B M M, A Hybrid Liar/Radar-based deep learning and vehicle recognition engine for autonomous vehicle pre-crash control, *East-Eur J Enterp Technol*, **5(9)** (2018) 6–17, 10.15587/1729-4061.2018.141298.

21 Zhao T, Ekim Y, Joel P & Giorgio R, Automated vehicle safety guarantee, verification and certification: A survey, arXiv preprint arXiv:2202.02818 (2022).

22 Dequaire, Julie, Deep tracking in the wild: End-to-end tracking using recurrent neural networks, *Int J Robot Res*, **374(5)** (2018) 492–512.