International Journal of Data Informatics and Intelligent Computing (IJDIIC)
Vol. 3, No. 3, September 2024, pp. 1~7
DOI: 10.59461/ijdiic.v3i3.121

Advancing Cyber Resilience for Autonomous Systems with Novel AI-based Intrusion Prevention Model

Sujatha Krishna[1], Paryati[2]
[1]Information Technology Department, College of Computing and Information Sciences, University of Technology and Applied Sciences-Shinas, Al-Aqr, Shinas 324, Oman
[2]University Development "Veteran" Yogyakarta, UPN "Veteran" Yogyakarta, Yogyakarta, Indonesia

Article Info

Article history:

Received April 29, 2024
Revised June 18, 2024
Accepted July 06, 2024

Keywords:

Autonomous Systems
Intrusion Prevention
Cyber Threats
Gannet optimized Mutated k-Nearest Neighbour (GO-MKNN)
Intrusion Prevention Systems (IPS).

ABSTRACT

Autonomous systems rely on intricate algorithms and are networked, which makes them more susceptible to cyber-attacks. The contexts of traditional intrusion prevention systems (IPS) are frequently difficult with the ever-changing nature of cyber threats. To overcome these limitations, we propose the Gannet optimized-mutated k-nearest Neighbour (GO-MKNN) approach as a novel customized intrusion prevention paradigm for self-governing systems. To improve detection accuracy and adaptability, the GO-MKNN algorithm combines the optimization powers of Gannet algorithms with the durability of the MKNN approach. Initially, this study obtained a dataset from CICIDS2017 CAN-intrusion, which was utilized for automobile attacks, and suggested designing additional IDS for the CAN system to train our suggested model. Following dataset collection, data cleaning and normalization were performed. Python was utilized to simulate our proposed method. The suggested method's effectiveness was evaluated in terms of Precision (%), Accuracy (%), F1-score (%) and recall (%). The experimental findings of the research may contribute to the development of a strong framework for intrusion detection that would guarantee the dependability and safety of autonomous vehicles.

Corresponding Author: Paryati (e-mail: upnyaya@gmail.com)

1. INTRODUCTION

Modern autonomous systems, including self-driving automobiles, unmanned drones, completely automated manufacturing operations and automated urban infrastructures, discover cyber resilience as an important component. The essence of cyber protection autonomous devices ensures the systems are resilient to attacks that verify dependability, protection and efficiency. Self-driving vehicles are difficult because they rely on sensors driven by data algorithms and network-based connections [1]. The networks play a role in industries like national defence, transport, and medical care, and the nature of threats ranges from information theft and security breaches to physical damage and operational interruptions [2]. Robust regulations and processes promote cyber-security resilience, which is necessary for the administration of self-sufficient systems. The synchronized immediate reaction to cyber catastrophes was censured by efficient management that helped to minimize possible harm and expedite the recapture of operations. Furthermore, the notion of security by design becomes essential for autonomous devices to engage in more interactions with networks that possess differing degrees of security [3]. The rapid rate of security threats is evolving, and the growing sophistication of attack methods is essential. The development of artificial intelligence models capable of anticipating and repelling cyber-attacks, block-chain-based technologies for safe and open operations and quantum computing transform encryption techniques [4]. The study's aim is to propose a Gannet-optimized mutated k-Nearest Neighbour (GO-MKNN) algorithm for strong security plans that shield autonomous technologies against cyber-attacks and weaknesses, guaranteeing their secure and dependable operation in dynamic environments. These sections make up the article's structure: section 2, literature review: section 3, methodological framework: section 4, outcomes: section 5, conclusion.

2. LITERATURE REVIEW

Detecting potential cyber-attacks against "Autonomous Mobility Systems (AMS)" on the vehicle and network phases, digital elements and realistic AMS operation scenarios [5]. Transportation was significantly altered by the growing cyber connection of vehicles among infrastructure and motor vehicles. The use of robotics, "Unmanned Aerial Vehicles (UAV)", and self-driving automobiles was creating new applications, indicating the turning point of autonomous vehicles. For precise and efficient resilience and robustness, application-aware robustness measurements, comprehensive fault assessment platforms and compact fault mitigation techniques [6]. Study [7] proposed an "ARC (Autonomous Response Controller)' for the security system. Where ARC employed a statistical "Hierarchical Risk Correlation Tree (HRCT)" to quantify the economic threat posed by "cyber-physical system (CPS)" assets by cyber-attacks to anticipate the avenues of attackers to accomplish specific targets, system resilience can be enhanced by “HMADS (Hierarchical Multi-Agent Dynamic System)”cyber-security surveillance and response, the multi-agent hierarchy utilized for a robust system of control model [8]. They constitute a hierarchy-based structure with a three-layer multi-agent network that has the ability to recognize and respond to cyber assaults.

Study [9] examined how competitive choices in strategy are strengthened in environments with automated systems, conventional wisdom, consumer expectations, and collaborative leadership. A study [10] examined an optimal routing protocol-based safe and sustainable health information transmission system for "IoT (Internet of Things)". Mobile computers were spreading throughout the private and public sectors, becoming more significant in the field of healthcare, not exclusively for sensory stimulation that possesses interpersonal interaction, documentation and presentation [11]. Medical care possessed a high degree of safety and conservation of energy with the "ECC-EERP (Elliptic Curve Cryptography-Based Energy-Efficient Routing Protocol)" to reduce the entire quantity of energy used by Web Service Network (WSN) [12]; it connected private keys and public keys to protect and decode internet traffic.

3. METHOD

3.1. System model

The framework paradigm of our suggested system effectively detects harmful CAN data from the Automatic Vehicles system. The AI-based suggested architecture consists of Automatic Vehicles and several sensors that are connected to one another, such as$\{ ast_1, t_2, \ldots., t_m\} \in T$. Using the CAN system, $t_1$and $t_2$interact to exchange actual-time data from the system of Automatic Vehicles.

$$t_1 \frac{via\ CAN\ traffic}{Monitored\ data} t_2 \tag{1}$$

An attacker can compromise the effectiveness of the Automatic Vehicles systems by taking advantage of the CAN data and exploiting the (t_l) susceptible instrument, $t_l \in T$.

$$t_l \frac{via\ manipulated\ CAN\ traffic}{forged\ data} t_2 \tag{2}$$

AI can effectively divide CAN data into harmful (τ) and not harmful (τ') types and optimizes the protection of the Automatic Vehicles system's actual time data being monitored, and that was needed to address the protection concerns.

$$\mathbb{C} \rightarrow \tau\ ,\ \tau' \tag{3}$$

$$P = max \sum_{j=1}^{m} Secure\ (D) \tag{4}$$

The locations of the AI classification algorithm, harmful and not harmful CAN data of Automatic Vehicles are represented by (τ) and (τ'). They develop several AI classifiers using CICIDS2017 information to accomplish the target function. The normalization issues, missing values and class imbalance issues from the datasets are noted and corrected during the pre-processing stage of the data. The Automatic Vehicles CAN data (τ) and (τ') are classified by AI classifiers in an effort to accurately identify the data. The suggested AI intelligence paradigm was evaluated using a range of performance measures such as f1-score, accuracy, recall, and precision. Figure 1 depicts the proposed methodology. The following provides an extensive description of every layer.

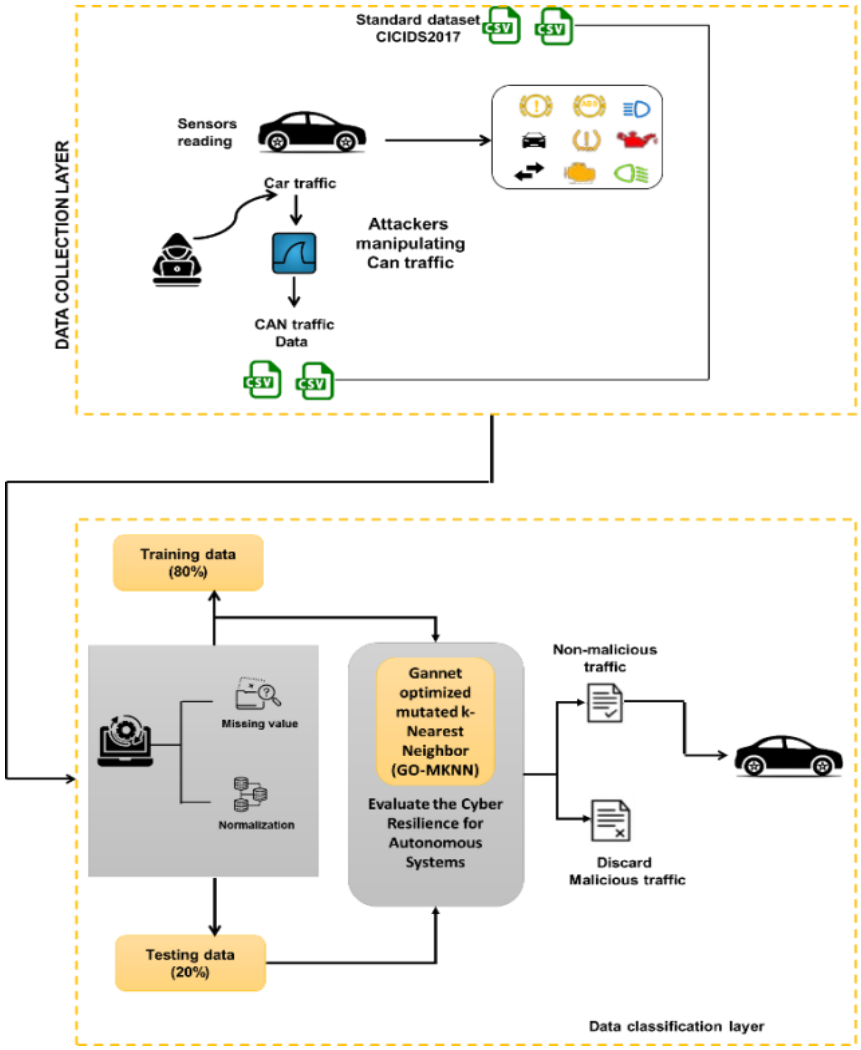


Figure 1. Proposed methodology

3.2. Dataset description

An external and inter-autonomous vehicle dataset was utilized in these suggested IDS to assess this work. The CICIDS2017 CAN-intrusion dataset, which was utilized for automobile attacks and suggested designing an additional IDS for the CAN system, was gathered. The dataset, which consists of 56661 rows and 78 columns, constitutes details on forwarding packets, flow packets, total bytes of flow and backward package dimension and other dataset properties that are useful for identifying certain security threats and generating alarms using IDS[13]. The dataset is split into training and testing datasets of 80% and 20%.

3.3. Data pre-processing Layer

Data normalization was carried out by scaling characteristics in a common range. Scaling the values of attributes in a dataset was accomplished using the data preparation technique known as normalization. The traditional standard population represents the dispersion mean of 0 and a standard error of 1. The initial stage was splitting the dataset into testing and training sets. Equation (5) specifies the uniformity:

$$Y = \frac{X_j - \mu}{\sigma}(\mathrm{j}) \tag{5}$$

Here $X_j$represents a single data point, μ represents a collection of data mean and σ denotes the dataset deviation from the mean.

3.4. Data Classification Layer

To accurately categorize (τ) and (τ') of autonomous vehicles CAN data, the training dataset was run by the Gannet optimized-mutated k-nearest Neighbor (GO-MKNN) algorithm.

3.4.1. Gannet optimization (GO)

Gannet Optimization (GO) incorporates ideas from the behaviour of gannets, seabirds recognized for their adept hunting. They are utilized in conjunction with cyber resilience for autonomous systems that strengthen autonomous systems' defences against cyber-attacks by utilizing tactics derived from nature to ensure safe, effective functioning under dynamic conditions. The GO uses a mathematical representation of the distinctive feeding actions of gannets to locate the best area in the field of search. GO was divided into two stages: the discovery stage and the extraction stage. The population-centered adaptive optimization methods are similar to GO. In the exploratory stage, gannets use the U-formed dip in Equation (6) and the V-curved dip in Equation (7) to find the prey. GOA establishes an information matrix NW that was modified during the exploration stage, as indicated in Equation (8). NW_jUsed to store the modifications in the location of the Gannet, substituting $W_j$using NW_jprovided a particular person NW_jsurpasses the current state of W_j.

$$b = 2 * cos(2 * \pi * q2) * (1 - s/S) \tag{6}$$

$$a = \begin{cases} 2 * \left(-\frac{1}{\pi} * w + 1\right) * (2 * \pi * q3) * \left(1 - \frac{s}{S}\right), w\epsilon(0, \pi) \\ 2 * \left(\frac{1}{\pi} * w + 1\right) * (2 * \pi * q3) * \left(1 - \frac{s}{S}\right), w\epsilon(2\pi, \pi) \end{cases} \tag{7}$$

$$NW_j^{s+1} = \begin{cases} w_j^s + v1 + (2 * q4 - 1) * b * \left(w_j^s - w_q^s\right), r \geq 0.5 \\ w_j^s + u1 + (2 * q5 - 1) * a * \left(w_j^s - \frac{1}{M}\sum_{i=1}^{M} w_j^s\right), r \geq 0.5 \end{cases} \tag{8}$$

Here, s represents the total number of repetitions, and $q2, q3, and$ q4 represent arbitrary numbers that range from 0 to 1. Where S represents the maximum amount of repetitions;$v1$ and $u1$ indicate the arbitrary number between $-b$ to $b$and$-a$ to a respectively; M represents the total population size; w_q^s signifies a randomly picked person; and w_j^s is the j th individuals in the population at the moment. The Gannet constitutes additional excursions in the stage of extraction to capture the species that possess desperate attempts. The Gannet's recapture potential was strong, expressed in Equation (9), which represents quickly seizing the prey. However, the energy level drops, no longer seizing the fleeing species that resort to a periodic motion to locate the next meal. Equation (10) represents the location modification ratio during the extraction stage.

$$Cap = \frac{1}{\frac{n * u^2}{0.2 + (2 - 0.2) * (1 + \frac{s}{S})}} \tag{9}$$

$$NW_j^{s+1} = \begin{cases} s * Cap * \left|w_j^s - w_{best}^s\right| * \left|w_j^s - {w_{best}}^s\right| + w_j^s, Cap \geq d \\ w_{best}^s - \left(w_j^s - w_{best}^s\right) * O * s, Cap < d \end{cases} \tag{10}$$

$$O = Levy(dim) = 0.01 * \frac{\mu * \sigma}{|v|^{\frac{1}{\beta}}} \tag{11}$$

Here, μ and σ represent the arbitrary integers ranging from 0 to 1, β indicates an integer value that represents the Gannet's kilograms and ${w_{best}}^s$indicates the strongest performer in a particular community.

3.4.2. Mutated K-Nearest Neighbour (MKNN)

For increased flexibility in autonomous systems, the Mutated K-Nearest Neighbour (MKNN) blends mutation techniques with the conventional KNN algorithm. The systems are strengthened against cyber-attacks by cyber resilience that guarantees system stability and uninterrupted operation. The Euclidean method was frequently applied; however, the best scenarios for the sample vector's measurement criteria are consistent. The amount of distance measured by Euclid was correlated with the value of the test component when all of the components and variables were distinct kinds. The weight $t_l(L = 1,2, \ldots., C)$of every directional sub-scalar was hard to ensure in real-world applications because the dimension's value has a narrow variation spectrum. The weight coefficients of coordinates exhibiting significant changes are less than those exhibiting minor changes, and sampling variability across several attributes was assessed within the same region. Equation (12) represents the normalized Euclidean distance:

$$c(w_j, w_i) = \sqrt{\sum_{l=1}^{C} \frac{(w_{jl}, w_{il})^2}{t_l}} \tag{12}$$

The frequently employed deep learning algorithm is refined by numerous academics, primarily in the area of l parameter modification. Nevertheless, the investigation has been carried out on the choice of proximity operations, duration length. The sample-based distributions, information features and statistical needs describe conventional problems of optimization. They create sensible and practical MKNN classification algorithms utilizing the predictive power of meta-heuristic strategy.

3.5. Gannet optimized mutated k-Nearest Neighbour (GO-MKNN)

A powerful hybrid strategy that has the potential to completely transform the security and effectiveness of autonomous systems was developed by Gannet, an optimized and mutated k-Nearest Neighbour algorithm that combined with cyber resilience for autonomous systems. The mutated k-Nearest Neighbour algorithm is recognized for its flexibility and durability in task classification combined with Gannet optimization that mimics seabird foraging behaviour to improve exploration and exploitation that achieves previously unattainable levels of reliability and diversity in making decisions. The integration of cyber resilience tactics enhances the system's capacity to identify, counter, and recover from hostile intrusions, guaranteeing uninterrupted and secure functioning in ever-changing surroundings. When combined, hybrid solutions constitute the new era of safe and trustworthy autonomy by improving the efficiency and dependability of autonomous systems, which also possess resistance to cyber-attacks.

4. EXPERIMENTAL RESULTS

We assess the recommended approach and calculate the effectiveness of the procedure using measures like F1-score (%), accuracy (%), Precision (%), and Recall (%). We also present an efficacy comparison between our proposed strategy and other current approaches. Decision trees, random forests, extra trees, and XGboost are some of the current techniques [14].

Accuracy provides a strong evaluation of the system's effectiveness by evaluating the model's preciseness assessment by computing the ratio of successfully anticipated attack indicator occurrences to the total occurrences. The precision of a model indicated how accurate its anticipated results were. The ratio of accurately predicted positive results to the total number of anticipated advantages is assessed. The comparison of accuracy and precision between the suggested and conventional approaches is displayed in Figures 2(a) and (b). As opposed to other approaches, the recommended GO-MKNN method achieves an accuracy of 99% and a precision of 98.76%. Our suggested approach effectively ensured the cars' safety, dependability, and robust security.

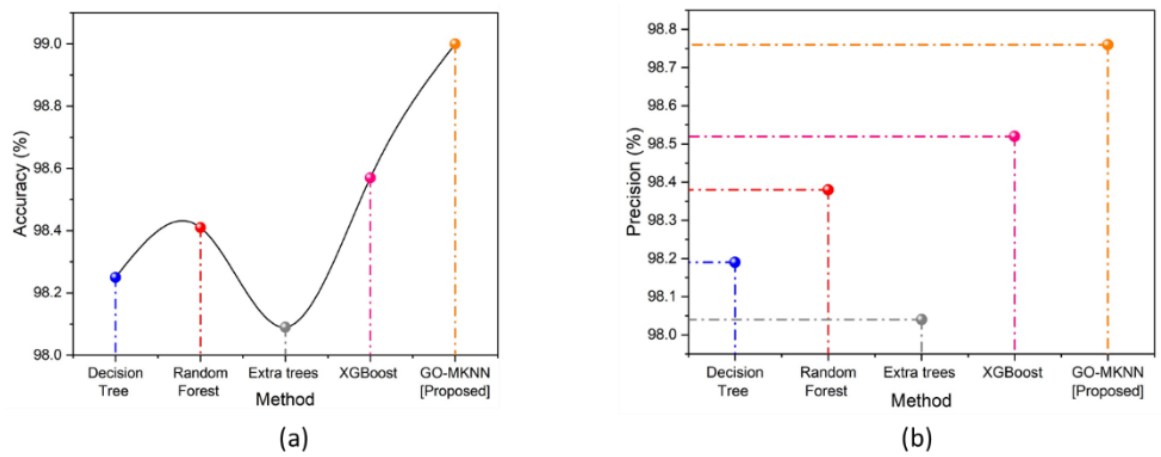
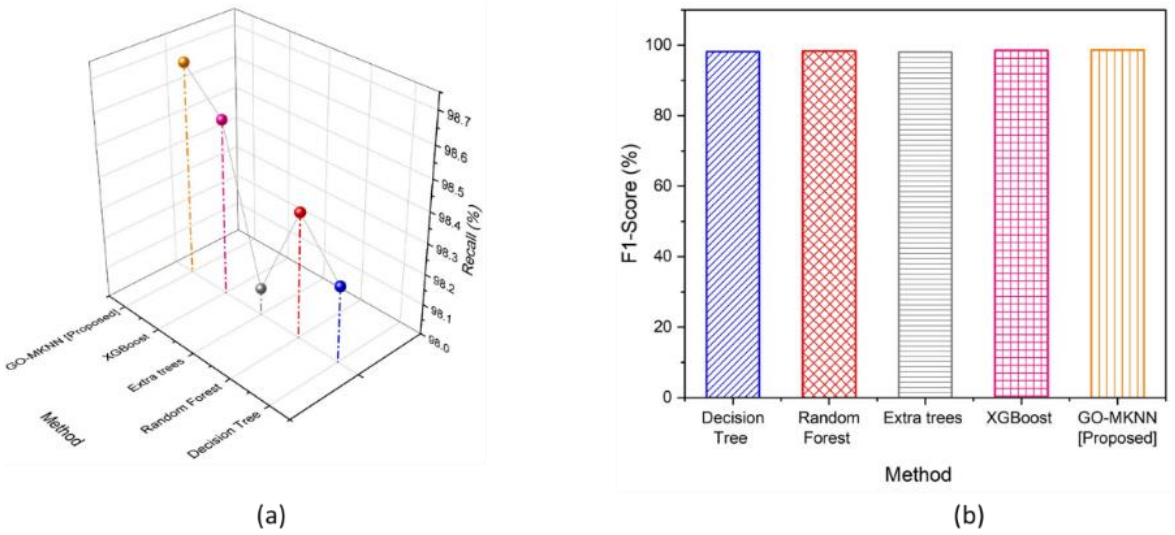


Figure 2. Result of (a) Accuracy and (b) Precision

The capacity of a system to precisely recognize every instance of a specific class, in this case, assaults or intrusions, is measured by recall. The F1-score is the average of recall and accuracy. Figures 3(a) and (b) present a comparative evaluation of recall and f1-score between the approach and traditional methods. In contrast to various methods, the suggested GO-MKNN method attains a recall of 98.69% and an F1-Score of 98.65%. Our proposed method provided optimization of strong security, reliability, and vehicle safety. Table 1 depicts outcome values.



(a) (b)

Figure 3. Result of (a) Recall (b) F1-Score

Table 1. Outcome Values

Methods	Recall (%)	Accuracy (%)	F1-Score (%)	Precision (%)
DT	98.25	98.25	98.18	98.19
RF	98.41	98.41	98.37	98.38
XGBoost	98.57	98.57	98.52	98.52
Extra trees	98.09	98.09	98.03	98.04
GO-MKNN [Proposed]	98.69	99	98.65	98.76

5. CONCLUSION

In this study, we introduced a novel approach called Gannet-optimized mutated k-nearest Neighbor (GO-MKNN) that demonstrates the customized intrusion prevention paradigm for autonomous systems. The experimental results showed Accuracy (99%), Precision (98.76%), F1 Score (98.65%) and Recall (98.69%). When the results of the suggested procedure were compared to other previously employed algorithms, it became evident from the assessments that the recommended approach was superior in terms of robust security and cyberattack prevention. The systems are complex and can be challenging to attain complete durability due to multiple risks. In future research, the robustness of autonomous vehicles will be significantly enhanced by utilizing AI and AL (Automated learning) to foresee attacks. The creation systems anticipate risks in addition to responding to them and constitute an essential domain of development.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest in this work.

REFERENCES

[1] A. Kott and P. Theron, "Doers, Not Watchers: Intelligent Autonomous Agents Are a Path to Cyber Resilience," *IEEE Secur. Priv.*, vol. 18, no. 3, pp. 62–66, May 2020, doi: 10.1109/MSEC.2020.2983714.

[2] O. Onishchenko, K. Shumilova, S. Volyanskyy, Y. Volyanskaya, and Y. Volianskyi, "Ensuring Cyber Resilience of Ship Information Systems," *TransNav, Int. J. Mar. Navig. Saf. Sea Transp.*, vol. 16, no. 1, pp. 43–50, 2022, doi: 10.12716/1001.16.01.04.

[3] Z. Lian, P. Shi, and C.-C. Lim, "Adaptive Resilient Control for Cyber-Physical Systems Under Cyberattack and Input Saturation," *IEEE Trans. Ind. Informatics*, vol. 19, no. 5, pp. 6513–6524, May 2023, doi: 10.1109/TII.2022.3198699.

[4] S. Colabianchi, F. Costantino, G. Di Gravio, F. Nonino, and R. Patriarca, "Discussing resilience in the context of cyber physical systems," *Comput. Ind. Eng.*, vol. 160, p. 107534, Oct. 2021, doi: 10.1016/j.cie.2021.107534.

[5] B. Zou, P. Choobchian, and J. Rozenberg, "Cyber resilience of autonomous mobility systems: cyber-attacks and resilience-enhancing strategies," *J. Transp. Secur.*, vol. 14, no. 3–4, pp. 137–155, Dec. 2021, doi: 10.1007/s12198-021-00230-w.

[6] Z. Wan, K. Swaminathan, P.-Y. Chen, N. Chandramoorthy, and A. Raychowdhury, "Analyzing and Improving Resilience and Robustness of Autonomous Systems," in *Proceedings of the 41st IEEE/ACM International Conference on Computer-Aided Design*, New York, NY, USA: ACM, Oct. 2022, pp. 1–9. doi: 10.1145/3508352.3561111.

[7] H. A. Kholidy, "Autonomous mitigation of cyber risks in the Cyber–Physical Systems," *Futur. Gener. Comput. Syst.*, vol. 115, pp. 171–187, Feb. 2021, doi: 10.1016/j.future.2020.09.002.

[8] C. Rieger, C. Kolias, J. Ulrich, and T. R. McJunkin, "A Cyber Resilient Design for Control Systems," in *2020 Resilience Week (RWS)*, IEEE, Oct. 2020, pp. 18–25. doi: 10.1109/RWS50334.2020.9241300.

[9] N. K. Rajagopal *et al.*, "Future of Business Culture: An Artificial Intelligence-Driven Digital Framework for Organization Decision-Making Process," *Complexity*, vol. 2022, pp. 1–14, Jul. 2022, doi: 10.1155/2022/7796507.

[10] E. Refaee *et al.*, "Secure and Scalable Healthcare Data Transmission in IoT Based on Optimized Routing Protocols for Mobile Computing Applications," *Wirel. Commun. Mob. Comput.*, vol. 2022, pp. 1–12, Jun. 2022, doi: 10.1155/2022/5665408.

[11] R. K. Kaushal *et al.*, "Using Mobile Computing to Provide a Smart and Secure Internet of Things (IoT) Framework for Medical Applications," *Wirel. Commun. Mob. Comput.*, vol. 2022, pp. 1–13, Oct. 2022, doi: 10.1155/2022/8741357.

[12] R. Natarajan, G. H. Lokesh, F. Flammini, A. Premkumar, V. K. Venkatesan, and S. K. Gupta, "A Novel Framework on Security and Energy Enhancement Based on Internet of Medical Things for Healthcare 5.0," *Infrastructures*, vol. 8, no. 2, p. 22, Feb. 2023, doi: 10.3390/infrastructures8020022.

[13] J. Thaker, N. K. Jadav, S. Tanwar, P. Bhattacharya, and H. Shahinzadeh, "Ensemble Learning-based Intrusion Detection System for Autonomous Vehicle," in *2022 Sixth International Conference on Smart Cities, Internet of Things and Applications (SCIoT)*, IEEE, Sep. 2022, pp. 1–6. doi: 10.1109/SCIoT56583.2022.9953697.

[14] S. Mewada, D. Sreenivasa Chakravarthi, S. J. Sultanuddin, and S. Kant Gupta, "Design and Implementation of a Smart Healthcare System Using Blockchain Technology with A Dragonfly Optimization-based Blowfish Encryption Algorithm," in *The Data-Driven Blockchain Ecosystem*, Boca Raton: CRC Press, 2022, pp. 137–153. doi: 10.1201/9781003269281-10.

BIOGRAPHIES OF AUTHORS

Sujatha Krishna received the B.E. and M. Tech degrees in Computer Science and Engineering from Visvesvaraya Technological University, Karnataka, India. She received a PhD degree in Computer Science and Engineering from REVA University, Karnataka, India. She is currently working as a Lecturer at the University of Technology and Applied Sciences-Shinas, Sultanate of Oman. Her research interests include big data, data mining, machine learning and privacy-preserving algorithms. She can be contacted at email: sujatha.krishna@utas.edu.om.

Paryati is a Lecturer and Assistant Professor at the National Development University "Veteran" Yogyakarta, Indonesia. She has completed a bachelor's degree in Informatics Management and Computer Engineering Study Program. She completed her master's in computer science at Gajah Mada University as the fastest-graduating student and cum laude. She is currently completing her doctoral preparation. She has been a certified Microsoft Innovative educator trainer since 2011. Her research interests are genetic algorithms, artificial intelligence, expert systems, data mining, machine learning, deep learning systems, fuzzy logic, data analysis and smart city information systems. She has published 97 papers at international and national conferences as well as in the indexed journals Scopus, Springer, IEEE, IGI Global, Taylor & Francis, and others. She has 4 book chapters at Taylor & Francis, 3 books at Betham publisher, 7 books at LAMBERT publisher. She is a peer reviewer in various international journals. She is a member of technical functional bodies in APTIKOM, HMI, IASR and others. She has published 13 COPYRIGHT Books and 4 PATENTS. She is a reviewer for the journals ABAARJ, AJASR, AJCR, AJOMCOR, AJPAM, AJPAS, ARJOM, CJAST, GPH, IJB, JAMCS, MULTIDICIPLINARY, JPRI. She has been a speaker for 5 webinars, 5 international conferences, 5 international workshops, and soft skills workshops at home and abroad. She has received various research and community service grants at national and international levels. She can be contacted at email: upnyaya@gmail.com.