# Novel Graph-Based Machine Learning Technique to Secure Smart Vehicles in Intelligent Transportation Systems

Brij Bhooshan Gupta, *Senior Member, IEEE*, Akshat Gaurav, *Graduate Student Member, IEEE*, Enrique Caño Marín, and Wadee Alhalabi

***Abstract*— Intelligent Transport Systems (ITS) is a developing technology that will significantly alter the driving experience. In such systems, smart vehicles and Road-Side Units (RSUs) communicate through the VANET. Safety apps use these data to identify and prevent hazardous situations in real-time. Detection of malicious nodes and attack traffic in Intelligent Transportation Systems (ITS) is a current research subject. Recently, researchers are proposing graph-based machine learning techniques to identify malicious users in the ITS environment, through which it is easy to analyze the network traffic and detect the malicious devices. Therefore, graph-based machine learning techniques could be a technique that efficiently detect malicious nodes in the ITS environment. In this context, this article aims to provide a technique for resolving authentication and security issues in ITS using lightweight cryptography and graph-based machine learning. Our solution uses the concepts of identity based authentication technique and graph-based machine learning in order to provide authentication and security to the smart vehicle in ITS. By authenticating smart vehicles in ITS and identifying various cyber threats, our proposed method substantially contributes to the development of intelligent transportation communication environment.**

***Index Terms*— VANET, V2V, V2G, ITS, IBE, graph-based machine learning, support vector machine.**

## I. Introduction

NUMEROUS traffic accidents have resulted in deaths and severe injuries, and patients are often unable to get medical facilities timely. Therefore, Researchers created the idea of an Intelligent Transport Systems (ITS) network to improve road safety by standardizing communication protocols for smart and electrified vehicles [1]. Through the Intelligent Transport Systems (ITS) [2], smart automobiles and electric vehicles connect, sharing data that might be utilised to save human lives in the event of a pandemic [3], and additionally resulting in new opportunities towards a more sustainable mobility leveraging the use of technology [4], as, for example, the emission reductions derived from the alleviation of the traffic volume [5] and its impact on citizens' health". The United States was the first regulatory body to see the promise of intelligent transportation systems, awarding DSRC 75 MHz of 5.9 GHz spectrum. IEEE develops the WAVE standards for VANET based on the DSRC standards. Mobile vehicles equipped with OBUs and fixed RSUs comprise the VANET. There are now two forms of communication supported by smart vehicles, namely V2V (many smart cars) and V2G (smart vehicles communicating with the RSU). as shown in Figure 1. However, as technology improves and the automobile industry grows, many new forms of communication arise [6].

Until now, the communication network met the criteria of the smart vehicle's environment [7], [8]. However, when the number of facilities and vehicles increases, the communication network is unable to offer electric vehicle devices with ultra-low latency services, security, and high dependability. Due to the fact that electric vehicles rely on connectionless transmission, they are susceptible to a variety of cyber-attacks. One of these kinds of attack is the DDoS attack. DDoS attacks disrupt V2V and V2G connections, preventing the smart vehicle from processing information received from its neighbours or transmitting critical information to its neighbours. As a result of the DDoS attack, the passengers' safety is jeopardized. To counter this issue, researchers suggest that the electric vehicle environment has to be powered by machine learning security mechanisms. AI and machine learning-enabled electric vehicles provide a number of benefits, including increased device reliability and security, as these technologies are increasingly being adopted across sectors. When there are a high number of relevant characteristics, the machine learning algorithm performs well. However, as the number of irrelevant characteristics in the dataset grows, the accuracy of the machine learning algorithm decreases. As a result, researchers are constantly looking for the best technique to identify important characteristics for machine learning algorithms. Recently, researchers have developed a graph-based machine learning [9]–[11] for picking

Manuscript received 15 November 2021; revised 28 March 2022 and 1 May 2022; accepted 4 May 2022. Date of publication 30 May 2022; date of current version 2 August 2023. This work was supported by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant KEP-PhD-5-611-38. The Associate Editor for this article was W. Wei. *(Corresponding author: Brij Bhooshan Gupta.)*

Brij Bhooshan Gupta is with the International Center for AI and Cyber Security Research and Innovations and the Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan, also with Lebanese American University, Beirut 1102, Lebanon, and also with the Department of Computer Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: gupta.brij@gmail.com).

Akshat Gaurav is with the Ronin Institute, Montclair, NJ 07043 USA (e-mail: akshatgaurav470@gmail.com).

Enrique Caño Marín is with the Computer Science Department, University of Alcalá, 28805 Madrid, Spain (e-mail: enrique.cano@outlook.com).

Wadee Alhalabi is with the Department of Computer Science and the Immersive Virtual Reality Research Group, King Abdulaziz University, Jeddah 22254, Saudi Arabia, and also with the Department of Computer Science, Dar Al-Hekma University, Jeddah 22246, Saudi Arabia (e-mail: wsalhalabi@kau.edu.sa).

Digital Object Identifier 10.1109/TITS.2022.3174333

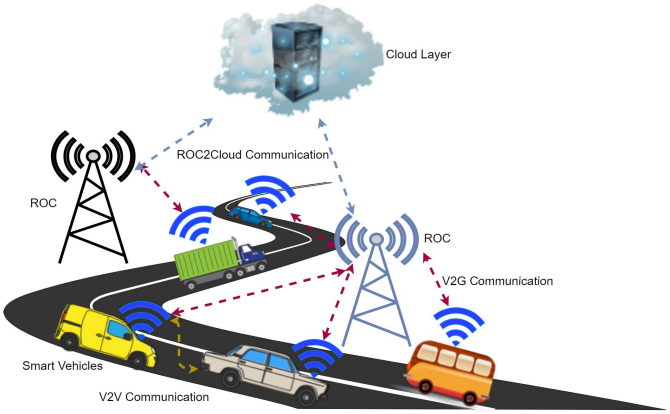



Fig. 1. ITS architecture.

the best feature of a dataset [12]. The use of graph-based characteristics for detecting malicious nodes has been documented in several research [9], [10], [13], [14]. If there is a good mix of critical and not-so-important elements in the data, machine learning algorithms can learn more successfully, as opposed to the opposite. Malicious node detection methods rely heavily on feature selection, according to several research in the literature. There have been a few research focusing on flow-based characteristics. Key to this sector is finding the optimal graph-based features to reveal hidden network structures that could disclose the communication patterns of malicious hosts.

In this context, using graph-based machine learning and identity-based encryption (IBE), this study discusses the development of a secure data exchange mechanism in conjunction with a cyber-attack detection approach. Furthermore, the proposed technique guarantees that no personal or private information will be revealed by using an identity-based encryption signature mechanism for electric car access control. The evaluation of network irregularities and the exclusion of dangerous packets is carried out using graph-based machine learning.

Section II disscusess the prior work in the area of attack detection in ITS. That's followed by a section discussing the methodology. Section IV gives the details about the simulation and results. Section V, summarizes the paper.

## II. Related Work

The Intelligent Transport Systems (ITS) [15], [16] enables a intelligent transportation system, which increases road safety [17], [18]. However, since smart vehicles have limited computing capacity, they share personal and private information, making them vulnerable to many kinds of cyberattacks such as data leakage, DoS, and DDoS attacks [19]–[24]. The researchers suggested several detection techniques for detecting various kinds of cyberattacks [25] in the Intelligent Transport Systems (ITS) environment [26]–[28]. However, the majority of cyber attack detection methods are ineffective in the electric car situation due to the vehicle's low resources. However, with today's communication networks, it's basic to use sophisticated detection methods such as machine learning approaches for identifying various kinds of assaults. Machine learning methods extract characteristics from raw data, which improves the detection process's reaction time [29], [30].

Conventional attack detection methods cannot be used in VANETs since the topology is not continuous. Therefore, a graph-based attack detection method was presented by researchers. However, there is yet no study on graph-based DDoS detection in VANET scenarios. Nevertheless, several other disciplines have presented attack detection methods based on graphs. For example, autors in [31] proposed a distributed graph-based technique to monitor IoT devices. The proposed approach classified communication traffic as normal or abnormal. Each node does its own anomaly detection. In another work, a network graph intrusion detection technique was suggested by the authors [32] in order to find malicious nodes in that network.

Author [33] proposed Graph-based Outlier Detection in Internet of Things for DoS attack detection. when a DoS assault is detected in real time, the author's GODIT technique uses a revolutionary graph-based outlier detection in IoT approach to quickly analyse the real-time graph data and identify it.Graph-stream anomaly detection techniques such as GODIT can outperform typical machine learning algorithms, according to data collected from an IoT equipped smart home. In other work, for the identification of botnets, the authors [12] suggest a graph-based ML model that first examines the relevance of graph properties before building a generalised model based on the chosen key features. A variety of feature sets are examined using five filter-based feature assessment measures drawn from diverse theories, such as coherence, correlation, and knowledge. The proposed graph-based botnet detection was evaluated with different supervised ML algorithms on two heterogeneous botnet datasets. Using features decreases training time and model complexity while increasing the detection rate of bots.

Other then graph based ML techniques there are several access control (AC) and attack detection systems that may be used for attack detection, but none of them consider the unique benefits of driving an electric car. FGAC technique has not been extensively studied in the context of electric vehicles. Data leaking is also a significant problem in the context of electric vehicles. As long as access control measures are in place, only authorised individuals will have access to the data.These approaches provide the data generator rule-based control over the data. Recent years have seen the development of a variety of access control systems, including IBAC, ABAC, and CBAC [34]–[38]. The authors in [39] proposed a graph-based technique to analyze the communication between different hosts in the network. The proposed approach is moderately efficient but not memory efficient because as the number of users increases, the size of the graph also increases.

The authors in [12] proposed a graph-based botnet detection technique. The proposed technique can detect a variety of botnets and is capable of detecting zero-day cyber attacks. The author analyzes the efficiency of the graph-based feature section technique using different machine learning algorithms. While all of the AC and attack detection systems described above are effective, they do not take into account the benefits

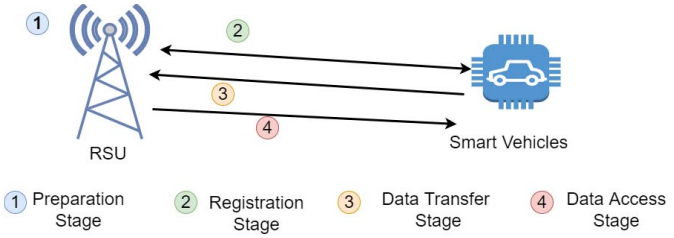



Fig. 2. State diagram for authentication stage.

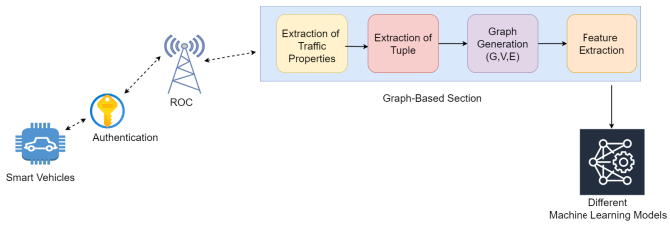



Fig. 3. Graph-based machine learning model.

of the electric car environment. The topic of FGAC in the context of electric vehicles has received little attention.

## III. Methodology

This section gives the details of the system model and system components of our proposed approach. Our proposed approach consists of; first stage uses the concepts of identity-based encryption system to authenticate the smart vehicles; Second stage applies the techniques of graph based machine learning to detect the malicious user/vehicle.

### A. Stage 1: Authentication Technique

This section explains the authentication method in depth. In Figure 2, we depict the suggested stage's state diagram. RSU performs all sophisticated encryption and key generation computations in our suggested solution. The details of the step performed by RSU in this stage are as following:

- **Preparation Stage**: The RSU determines a security parameter for the production of the master secret and public parameters in this first step of our suggested technique.
- **Registration Stage**: As part of our planned strategy, this is the second phase. It requests the public parameters when it is within the RSU's range, and the RSU creates a secret key based on that vehicle's unique ID.
- **Data Transfer Stage**: To get started with producing data, the smart car uses the RST to do so.
- **Data retrieve Stage**: This is the last part of the communication. At this point, the RSU verifies the smart vehicle's identity before granting access.

*1) IBE Technique:* We used Biliner map technique in the authentication of new vehicles in the ITS environment. The various stages of the authentication technique are as following:

- **Key Generation**: RSU generates the public parameters as well as the master key at this step. In order to keep our method modest, our proposed approach uses less number of public parameters than in [40].

$$(Parameters, Maseter_{key}) \xleftarrow{\$} \mathcal{G}(1)^S \quad (1)$$

$$Parameters = (x, x^y), \quad y \in \mathbb{Z}_p^* \quad (2)$$

$$Master_{key} = y \quad (3)$$

where $x$ cyclic group generator.

- **Encryption**: The public parameters and the unique ID are used by the smart vehicles to register with the RSU. For registration, smart vehicles encrypt their information and send it to the RSU.

$$R_{Inforamtion} = (x^{rID}, Y^r, \hat{e}(x,x)^r.I_{text}); \quad r \in \mathbb{Z}_p^* \quad (4)$$

where, $\hat{e}$ is biliner function, $I_{text}$ is personal data if vehicle, and $R_{information}$ is final registration information sent by the smart vehicle to the RSU.

- *Decryption*: At RSU, the details of the smart vehicles are stored in decryption form and the details are decrypted by the following equation:

$$I_{Information} \leftarrow (x^{rID}, (Y^r)^t, x^{\frac{1}{TD+yt}}); \quad y, t \in \mathbb{Z}_p^* \quad (5)$$

**Algorithm 1** IBE Technique

**Input**: Unidentified Vehicle
**Output**: Verification of Vehicle
**Begin**:
**Key Generation stage**:
Security parameter selection
Generation of Public parameters and Master Key
**Encryption Stage**:
$T_{t1} \xleftarrow{\$} x^{rID}$
$T_{t2} \xleftarrow{\$} Y^r, r \in \mathbb{G}$
$T_{t3} \xleftarrow{\$} \hat{e}(x,x)^r.$
$I_{text}\ R_{Inforamtion} = (x^{rID}, Y^r, \hat{e}(x,x)^r.I_{text})$
**Decryption Stage**:
Private key generation
$I_{Information} \leftarrow (x^{rID}, (Y^r)^t, x^{\frac{1}{TD+yt}})$
**End**

### B. Graph-Based Machine Learning Technique

This is the malicious node detection stage of our proposed approach. In this stage, firstly, the network traffic is converted into graph and then features of the graph are used to detect whether the malicious vehicle. The steps of this stage are as following:

*1) Network Flow Collection:* In our proposed approach, all the smart devices are connected through other devices by RSU. As RSU has high computation power and memory, the malicious node detection algorithm is implemented on the RSU. As represented in the Figure 3, the first step is to collect and process the network traffic for one time window. We prepare the graph structure from the network traffic. We calculate two tuples from the network traffic;

$$f_i = \{S_{ip}, S_N\}; i \in \{1, 2, ..n\} \quad (6)$$

$$f_i \in \mathcal{T} \quad (7)$$

where $S_{ip}$ is the source IP address and $S_N$ is packets sent. $\mathcal{T}$ is tuple set, $n$ is total number of smart vehicles. As in

our proposed approach, the communication occurs in between RSU and smart vehicles, there is no need to calacute the destination IP address and packets send. This reduces the memory requirements and increases the speed of operation.

*2) Graph Generation:* This is the second stage of our proposed approach in which the graph is prepared by using the tuple values calculated in the Equation 6 and Equation 7. The final graph is represented by $\mathcal{G}(\mathcal{V}, \mathcal{E})$. where

$$\mathcal{G} = Graph \tag{8}$$

$$\mathcal{V} = Vertices \tag{9}$$

$$\mathcal{E} = Edge \tag{10}$$

*3) Feature Extraction:* This is the third stage graph-based machine learning technique. At this stage RSU extracts important features from the network graph. This way, we extract the following standard features from the network graph :

- **Vertex Degree**: The total number of edges that link to a particular vertex is known as the vertex degree. In-degree is the total number of incident edges to a vertex. As in our proposed approach, all the smart vehicles are communicated to the RSU, we only calculate in degree of RSU. At the time of attack, the malicious user tries to flood the RSU with the fake requests; hence, at the time of attack, the in-request value increases. The value of in-degree of RSU is calculated by the following equation:

$$RSU_i = \sum \mathcal{E}(i), \ \forall i \in Vertices$$

$$\mathcal{E}(i) = \begin{cases} 1, & if \ i \in Edge \\ 0, & otherwise \end{cases}$$

- **Out degree Weight**: It is the measure of the total packets sent by any smart vehicle to the RSU. It is also an important factor for the identification of malicious nodes because the malicious nodes try to send a large number of packets at the time of attack. The value of out-degree weight is calculated by the following equation:

$$RSU_{weight} = \sum \mathcal{E}(i), \quad \forall i \in Vertices$$

- **Out degree centrality**: Measures of centrality, such as the Out Degree centrality, are simple to understand and may be used to any data collection. Centrality is measured by the number of smart cars it is able to interact with. Degree centrality is the ratio of the number of outbound connections to the number of nodes in a network.
- **Betweenness Centrality**: It is the number of shortest routes via a vertex that determines its betweenness centrality. The number of times a vertex serves as a bridge between two other vertices is known as betweenness centrality. The value of betweenness centrality (BC) is calculated by the following equation [41]:

$$BC = \sum_{x \neq y \neq v \in Vertices} \frac{\epsilon_{xy}(v)}{\epsilon_{xy}}$$

where $\epsilon_{xy}(v)$ is the shortest path from 'x' to 'y' passing through 'v' and $\epsilon_{xy}$ is the shortest path from 'x' to 'y'.

- **Local Clustering Coefficient**: A node's local clustering coefficient reveals the density of its immediate surroundings. If you want to know how near a node's neighbours are to each other, you may use the local clustering coefficient measure. Its value is calculated by the following equation [42]:

$$C = \frac{CP_a}{N_a(N_a - 1)}$$

where $N_a$ is the value of number of neighbours of a node and $CP_a$ is the number of connected pairs.

- **Closeness Centrality**: Another fascinating and potent centrality measure is "closeness centrality". The shortest route between a vertex v and every other vertex v's in a graph is used to determine the vertex v's closeness centrality. It is calculated by the following equation [43]:

$$CC_v = \frac{1}{\sum p_{uv}}$$

where $p_{uv}$ is the shortest path from 'u' to 'v'.

- **Eigen Centrality**: When looking at the relevance of nodes, Eigen centrality [44] is a good indicator of how important a particular node is in relation to its neighbours. EC goes beyond degree centrality in that nodes are affected not just by the number but also the significance of their neighbours. It is calculated by the following equation [13]:

$$EC_x = \frac{1}{\alpha} \sum \sigma_{x,y} EC_y$$

$$\sigma_{x,y} = \begin{cases} 1, & x \ connected \ to \ y \\ 0, & x \ not \ connected \ to \ y \end{cases}$$

- **Katz Centrality**:By considering both immediate and non-immediate surrounding nodes, Katz centrality estimates the relative significance of each node in a network. It is calculated by following equation [45]nodes:

$$KC_v = \sum \alpha \mathcal{A}$$

$$\mathcal{A} = \begin{cases} 1, & edge \ between \ v \ u \\ 0, & otherwise \end{cases}$$

- **Local Clustering Coefficient**: A node's proximity to other nodes is measured. The LCC metric measures the proximity of a node's neighbours to one another [42].

*4) Machine Learning Models:* In order to detection of malicious users in the graph, we used six different machine learning techniques. The details of the machine learning techniques used for testing the graphs are as follows:

- *SVM*: SVM [46] created by Vapnik *et al.* [47], [48] is a machine learning algorithm used in regression [49], [50] and pattern recognition [50], [51]. kernel mapping [52] is used to transform the input data into a linearly-separable feature space in the kernel mapping technique of an SVM In an SVM, the decision function is proportional to how many SVs and their weights there are, as well as how many kernels, such as Gaussian and polynomial, have been selected in advance [52], [53].

– *Logistic Regression*: Logistic regression models the probability of a discrete outcome given an input variable. Logistic regression is the most used model for a binary outcome, and so on. When there are more than two possible outcomes, multinomial logistic regression may be employed. If you're seeking to figure out which group a new sample belongs in, you'll want to utilise logistic regression as an analytical approach. [54].
– *DTC*: There are several ways to solve classification and regression issues using Decision Trees, although they are most often employed for Classification difficulties. Internal nodes reflect the dataset's attributes and branches represent decision rules; each leaf node represents an outcome. This is a tree-based classifier. The Decision Node and the Leaf Node are the two nodes of a Decision Tree. Leaf nodes, on the other hand, are the result of such choices and do not have any more branches to follow them.
– *Random Forest*: The supervised learning approach, Random Forest, is a well-known machine learning algorithm. Machine Learning (ML) applications include both Classification and Regression tasks. An ensemble of classifiers is used to combine and enhance the model's performance, and this approach is based on the notion of ensemble learning. The algorithm was created by Breiman and Cutler [55]. Decision trees are used in RF to forecast. Decision trees are formed during the training phase and then used for class prediction, taking into account the voted classes of each individual tree. This is the class with the highest margin function (MF) [56].
– *Gradient Boosting*: Gradient boosting is one of the best machine learning techniques. It is common to think of machine learning algorithms as having two main types of errors: bias and variance. When used as one of the strategies for boosting, gradient boosting decreases bias error in the model [57]. For example, the gradient boosting approach may be used to generate an additive model with the lowest possible loss function. In this approach, the gradient boosting technique keeps increasing the number of decision trees that reduce the loss function throughout each step. To improve performance, a shrinkage parameter known as the learning rate may be used to lower the new decision tree's contribution at each iterative step. A higher number of minor steps is more accurate than a smaller number of significant steps in gradient boosting, according to the shrinking method's underlying premise [58].
– *Multinomial Naive Bayes*: One of the most prominent Bayesian learning methods in Natural Language Processing is the Multinomial Naive Bayes algorithm (NLP). The Bayes theorem is used to predict the tag of a piece of text, such as an email or a newspaper article. For a given sample, it produces the tag with the highest possibility of being present. Several algorithms make up the Naive Bayes classifier, but they all have one thing in common: they classify features independently of one another. The inclusion or deletion of one characteristic is unaffected by the presence or absence of another.

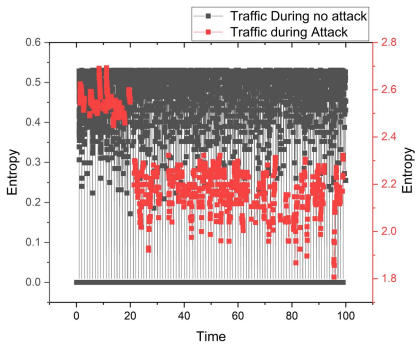


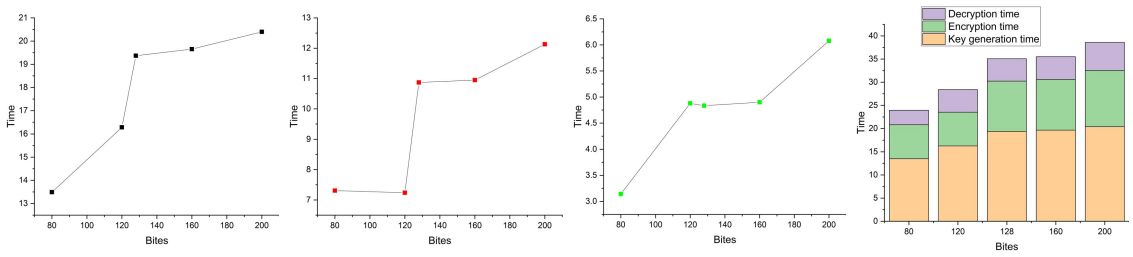

Fig. 4. Entropy variation at attck and no attack time.

## IV. Results and Analysis

The graph-based machine learning technique is examined in this section. We generate network traffic data with the help of tools like OMNET++, Vines, and SUMO. OMNET++ is used to run the simulation. In this case, the attack packets flood the victim's network with erroneous traffic, causing significant damage. As a result, the attacker node sends out packets every one second, but the legitimate nodes send out packets every five seconds, as shown. Simulating the whole process takes 100 seconds, during which time all log data is collected. As a result, the variation of entropy is represented in Figure 4. Instead of employing particular protocols, we utilise a general routing protocol for our simulations. Since the dataset includes a high number of null values, they are eliminated during the preparation process. We used graph-tool [59] to convert the network data into graph. With everything divided up into training and testing sets, the suggested approach can be thoroughly tested on both datasets.

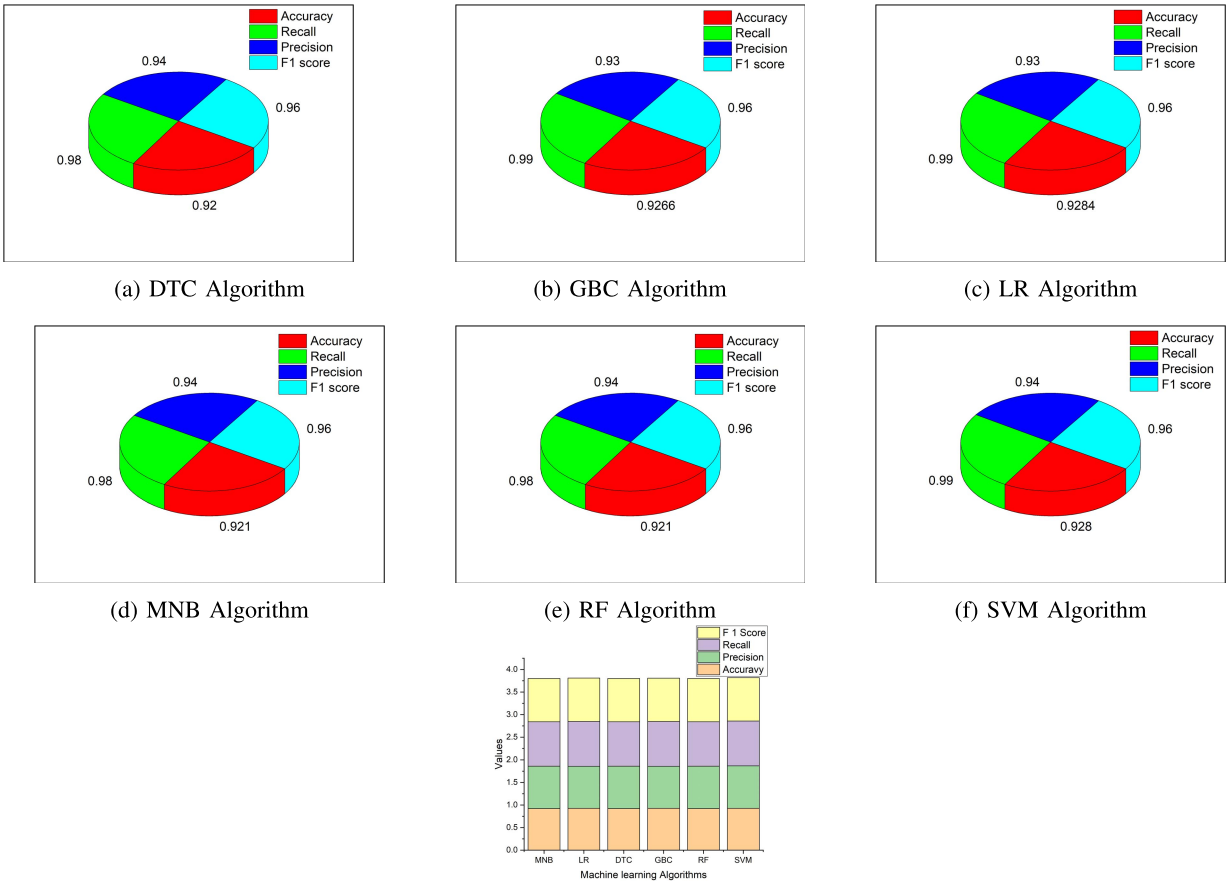
Our suggested IBE approach is tested in five distinct situations, and we compute the key setup time, encryption time, decryption time, and total time for each test scenario to establish its performance. For several test instances, the results are shown in Figure 5. Messages of varying lengths are generated for each test scenario. For the first, second, third, fourth, and fifth tests, we utilise 80-byte, 120-byte, 128-byte, 160-byte, and 200-byte messages.

– **Key Set up stage**: This is the starting point of the proposed authentication approach. RSU selects the public parameters and the master secret. The value of key generation time for different test scenarios are represetned in Figure 5a. From Figure 5a, it is clear that the key setup time is directly proportional to the bits of the key. Because of this, we must choose a key length that minimises the amount of time lost during the generating process.
– **Encryption Time**: In the second stage of our proposed approach, smart vehicle sends their information to the RSU with the help of public parameters and unique IDs. Figure 5b depicts a variety of test cases that employ different lengths of encryption keys. Similar to the key setup time, the encryption time is also depending upon the keyword bits.
– **Verification Time**: Finally, we've reached the end of our recommended strategy. Cloud service provider's public




(a) Key Selection (b) Encryption (c) Decryption (d) Total Time

Fig. 5. Performance evaluation of proposed approach.




(a) DTC Algorithm (b) GBC Algorithm (c) LR Algorithm

(d) MNB Algorithm (e) RF Algorithm (f) SVM Algorithm

(g) Comparison of different Algorithms

Fig. 6. State-of-the-art comparison.

key was utilised by the receiver of the data in order to verify its digital signatures. Equation 5 depicts the decryption process whereas Figure 5c depicts the time required to complete it. As shown in Figure 5d, key length may effect total time; consequently, in order to reduce total time, we need to choose an optimal key length for the current case.

### *A. State-of-Art Comparison*

This part compares our suggested approach to various frameworks and methods that are already available. We compare the state-of-the-art (as represented in Figure 6) using the following parameters:

– **Precision**:

$$P = \frac{\delta_{tp}}{\delta_{tp} + \delta_{fp}} \tag{11}$$

– **Recall**:

$$R = \frac{\delta_{tp}}{\delta_{tp} + \delta_{fn}} \tag{12}$$

– **F1-score**:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{13}$$

– **Accuracy**:

$$Accuracy = \frac{\delta_{tp} + \delta_{tn}}{\delta_{tp}} \tag{14}$$

Figure 6 compares the different machine learning techniques with respect to precision, recall, F1 score, and accuracy. From Figure 6b and Figure 6c, it is clear that the GBC and LR techniques have relatively less precision than other machine learning algorithms in the detection of malicious users. However, their precision is higher than other machine learning algorithms. Also, LR algorithm has highest accuracy, compared to all the other machine learning algorithms.

## V. Conclusion

Intelligent Transport Systems (ITS) place a premium on passenger safety. The early work in this field was mainly focused on developing new methods and equipping vehicles with new devices that might improve their performance. However, when the number of cars increases, the limited resources available to Intelligent Transport Systems (ITS) are unable to meet the security needs. The majority of research on the evolution of electric car technology has glossed over this point. Few researchers have addressed this problem and developed AI, machine learning, and deep learning-based security frameworks. These methods, however, are inefficient owing to the resource-constrained nature of electric cars. In this regard, we propose a system for providing AC in an Intelligent Transport System environment while also safeguarding it against various cyber threats. Our system is built on methods from IBE signature scheme, entropy measurement, and graph-based machine learning technique. Access control is managed through IBE, and cyberattacks are detected via entropy measurement and graph-based machine learning techniques that analyze the characteristics of data flow. We got acceptable results showing that our suggested framework is capable of communicating effectively with Intelligent Transport Systems (ITS). Our work is distinguished by a development framework that makes effective use of the smart vehicle system's resources. Our suggested framework is a critical first step toward the development of methods for securing Intelligent Transport Systems (ITS) using graph-based machine learning techniques. Our future research should focus on enhancing the suggested framework via the incorporation of new machine learning technologies.

## Acknowledgment

The authors acknowledge with thanks DSR technical and financial support.

## References

[1] M. Benadda and G. Belalem, "Improving road safety for driver malaise and sleepiness behind the wheel using vehicular cloud computing and body area networks," *Int. J. Softw. Sci. Comput. Intell.*, vol. 12, no. 4, pp. 19–41, 2020.

[2] L. Li, X. Li, Z. Li, D. D. Zeng, and W. T. Scherer, "A bibliographic analysis of the IEEE transactions on intelligent transportation systems literature," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 2, pp. 251–255, Jun. 2010.

[3] A. H. Sodhro *et al.*, "Towards 5G-enabled self adaptive green and reliable communication in intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5223–5231, Aug. 2021.

[4] C. T. C. Trapp, D. K. Kanbach, and S. Kraus, "Sector coupling and business models towards sustainability: The case of the hydrogen vehicle industry," *Sustain. Technol. Entrepreneurship*, vol. 1, no. 2, May 2022, Art. no. 100014.

[5] C. Medina-Molina and M. D. L. S. Rey-Tienda, "The transition towards the implementation of sustainable mobility. Looking for generalization of sustainable mobility in different territories by the application of QCA," *Sustain. Technol. Entrepreneurship*, vol. 1, no. 2, May 2022, Art. no. 100015.

[6] M. Noor-A-Rahim *et al.*, "6G for vehicle-to-everything (V2X) communications: Enabling technologies, challenges, and opportunities," 2020, *arXiv:2012.07753*.

[7] M. S. Omar *et al.*, "Multiobjective optimization in 5G hybrid networks," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1588–1597, Jun. 2018.

[8] H. Fatemidokht, M. K. Rafsanjani, B. B. Gupta, and C.-H. Hsu, "Efficient and secure routing protocol based on artificial intelligence algorithms with UAV-assisted for vehicular ad hoc networks in intelligent transportation systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4757–4769, Jul. 2021.

[9] A. A. Daya, M. A. Salahuddin, N. Limam, and R. Boutaba, "BotChase: Graph-based bot detection using machine learning," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 1, pp. 15–29, Mar. 2020.

[10] W. Wang, Y. Shang, Y. He, Y. Li, and J. Liu, "BotMark: Automated botnet detection with hybrid analysis of flow-based and graph-based traffic behaviors," *Inf. Sci.*, vol. 511, no. 2, pp. 284–296, Feb. 2020.

[11] B. M. Rahal, A. Santos, and M. Nogueira, "A distributed architecture for DDoS prediction and bot detection," *IEEE Access*, vol. 8, pp. 159756–159772, 2020.

[12] A. Alharbi and K. Alsubhi, "Botnet detection approach using graph-based machine learning," *IEEE Access*, vol. 9, pp. 99166–99180, 2021.

[13] S. Chowdhury *et al.*, "Botnet detection using graph-based feature clustering," *J. Big Data*, vol. 4, no. 1, pp. 1–23, Dec. 2017.

[14] B. Venkatesh, S. H. Choudhury, S. Nagaraja, and N. Balakrishnan, "BotSpot: Fast graph based identification of structured P2P bots," *J. Comput. Virol. Hacking Techn.*, vol. 11, no. 4, pp. 247–261, Nov. 2015.

[15] J. Shu, L. Zhou, W. Zhang, X. Du, and M. Guizani, "Collaborative intrusion detection for VANETs: A deep learning-based distributed SDN approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4519–4530, Jul. 2021.

[16] X. Xu, W. Wang, Y. Liu, X. Zhao, Z. Xu, and H. Zhou, "A bibliographic analysis and collaboration patterns of IEEE transactions on intelligent transportation systems between 2000 and 2015," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 8, pp. 2238–2247, Aug. 2016.

[17] S. S. Manvi and S. Tangade, "A survey on authentication schemes in VANETs for secured communication," *Veh. Commun.*, vol. 9, pp. 19–30, Jul. 2017.

[18] G. Guo and S. Wen, "Communication scheduling and control of a platoon of vehicles in VANETs," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 6, pp. 1551–1563, Dec. 2015.

[19] F. Sakiz and S. Sen, "A survey of attacks and detection mechanisms on intelligent transportation systems: VANETs and IoV," *Ad Hoc Netw.*, vol. 61, pp. 33–50, Jun. 2017.

[20] L. Xiao, X. Wan, C. Dai, X. Du, X. Chen, and M. Guizani, "Security in mobile edge caching with reinforcement learning," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 116–122, Jun. 2018.

[21] Z. Lu, G. Qu, and Z. Liu, "A survey on recent advances in vehicular network security, trust, and privacy," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 760–776, Feb. 2018.

[22] I. Cvitic, D. Perakovic, B. B. Gupta, and K.-K.-R. Choo, "Boosting-based DDoS detection in Internet of Things systems," *IEEE Internet Things J.*, vol. 9, no. 3, pp. 2109–2123, Feb. 2022.

[23] P. Arthurs, L. Gillam, P. Krause, N. Wang, K. Halder, and A. Mouzakitis, "A taxonomy and survey of edge cloud computing for intelligent transportation systems and connected vehicles," *IEEE Trans. Intell. Transp. Syst.*, early access, Jun. 7, 2021, doi: 10.1109/TITS.2021.3084396.

[24] A. H. Sodhro, J. J. P. C. Rodrigues, S. Pirbhulal, N. Zahid, A. R. L. de Macedo, and V. H. C. de Albuquerque, "Link optimization in software defined IoV driven autonomous transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3511–3520, Jun. 2021.

[25] M.-P. Pelletier, M. Trépanier, and C. Morency, "Smart card data use in public transit: A literature review," *Transp. Res. C, Emerg. Technol.*, vol. 19, no. 4, pp. 557–568, 2011.

[26] J. Liang, Q. Lin, J. Chen, and Y. Zhu, "A filter model based on hidden generalized mixture transition distribution model for intrusion detection system in vehicle ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 7, pp. 2707–2722, Jul. 2020.

[27] Y. Zeng, M. Qiu, Z. Ming, and M. Liu, "Senior2local: A machine learning based intrusion detection method for VANETs," in *Proc. Int. Conf. Smart Comput. Commun.* Tokyo, Japan: Springer, 2018, pp. 417–426.

[28] A. H. Sodhro, G. H. Sodhro, M. Guizani, S. Pirbhulal, and A. Boukerche, "AI-enabled reliable channel modeling architecture for fog computing vehicular networks," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 14–21, Apr. 2020.

[29] M. Qiu *et al.*, "Data allocation for hybrid memory with genetic algorithm," *IEEE Trans. Emerg. Topics Comput.*, vol. 3, no. 4, pp. 544–555, Dec. 2015.

[30] R. Rauter, D. Globocnik, E. Perl-Vorbach, and R. J. Baumgartner, "Open innovation and its effects on economic and sustainability innovation performance," *J. Innov. Knowl.*, vol. 4, no. 4, pp. 226–233, Oct. 2019.

[31] M.-O. Pahl, F.-X. Aubet, and S. Liebald, "Graph-based IoT microservice security," in *Proc. NOMS - IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2018, pp. 1–3.

[32] Q. Ding, N. Katenka, P. Barford, E. Kolaczyk, and M. Crovella, "Intrusion as (anti) social communication: Characterization and detection," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2012, pp. 886–894.

[33] R. Paudel, T. Muncy, and W. Eberle, "Detecting DoS attack in smart home IoT devices using a graph-based approach," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2019, pp. 5249–5258.

[34] Y. Liu *et al.*, "Capability-based IoT access control using blockchain," *Digit. Commun. Netw.*, vol. 7, no. 4, pp. 463–469, Nov. 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2352864820302844

[35] X. Yang and W. Ding, "Researches on data encryption scheme based on CP-ASBE of cloud storage," *Int. J. High Perform. Comput. Netw.*, vol. 14, no. 2, pp. 219–228, 2019.

[36] S. Kaushik and C. Gandhi, "Capability based outsourced data access control with assured file deletion and efficient revocation with trust factor in cloud computing," *Int. J. Cloud Appl. Comput.*, vol. 10, no. 1, pp. 64–84, Jan. 2020.

[37] A. Tewari and B. B. Gupta, "Cryptanalysis of a novel ultra-lightweight mutual authentication protocol for IoT devices using RFID tags," *J. Supercomput.*, vol. 73, no. 3, pp. 1085–1102, Mar. 2017.

[38] F. Mirsadeghi, M. K. Rafsanjani, and B. B. Gupta, "A trust infrastructure based authentication method for clustered vehicular ad hoc networks," *Peer–Peer Netw. Appl.*, vol. 14, pp. 2537–2553, Oct. 2020.

[39] S. Lagraa, J. Francois, A. Lahmadi, M. Miner, C. Hammerschmidt, and R. State, "BotGM: Unsupervised graph mining to detect botnets in traffic flows," in *Proc. 1st Cyber Secur. Netw. Conf. (CSNet)*, Oct. 2017, pp. 1–8.

[40] H. Yan, Y. Wang, C. Jia, J. Li, Y. Xiang, and W. Pedrycz, "IoT-FBAC: Function-based access control scheme using identity-based encryption in IoT," *Future Gener. Comput. Syst.*, vol. 95, pp. 344–353, Jun. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167739X1830997X

[41] U. Brandes, "A faster algorithm for betweenness centrality," *J. Math. Sociol.*, vol. 25, no. 2, pp. 163–177, 2001.

[42] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, nos. 66–84, pp. 440–442, 1998.

[43] T. Opsahl, F. Agneessens, and J. Skvoretzc, "Node centrality in weighted networks: Generalizing degree and shortest paths," *Social Netw.*, vol. 32, no. 3, pp. 245–251, 2010.

[44] A. N. Langville and C. D. Meyer, "A survey of eigenvector methods for web information retrieval," *SIAM Rev.*, vol. 47, no. 1, pp. 135–161, 2005.

[45] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.

[46] L. Zhang, W. Zhou, and L. Jiao, "Wavelet support vector machine," *IEEE Trans. Syst., Man, Cybern., B (Cybern.)*, vol. 34, no. 1, pp. 34–39, Feb. 2004.

[47] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[48] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Washington, DC, USA: Springer, 1999.

[49] A. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statist. Comput.*, vol. 14, no. 3, pp. 199–222, 2014.

[50] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neural Comput.*, vol. 12, no. 5, pp. 1207–1245, May 2000.

[51] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[52] B. Schölkopf *et al.*, "Input space versus feature space in kernel-based methods," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 1000–1017, Sep. 1999.

[53] A. J. Smola, B. Schölkopf, and K.-R. Müller, "The connection between regularization operators and support vector kernels," *Neural Netw.*, vol. 11, no. 4, pp. 637–649, Jun. 1998.

[54] T. Edgar and D. Manz, *Research Methods for Cyber Security*. Washington, DC, USA: Syngress, 2017.

[55] L. Breiman and A. Cutler, "Random forests-classification description," Dept. Statist., Berkeley, CA, USA, Tech. Rep. 2, 2007.

[56] Y. Liu, Y. Wang, and J. Zhang, "New machine learning algorithm: Random forest," in *Proc. Int. Conf. Inf. Comput. Appl.* Chengde, China: Springer, 2012, pp. 246–252.

[57] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Frontiers Neurorobot.*, vol. 7, p. 21, Dec. 2013.

[58] S. Touzani, J. Granderson, and S. Fernandes, "Gradient boosting machine for modeling the energy consumption of commercial buildings," *Energy Buildings*, vol. 158, pp. 1533–1543, Jan. 2018.

[59] T. de Paula Peixoto. *Graph-Tool*. Accessed: Feb. 12, 2022. [Online]. Available: https://graph-tool.skewed.de/

**Brij Bhooshan Gupta** (Senior Member, IEEE) received the Ph.D. degree from the Indian Institute of Technology (IIT) Roorkee, India. At present, he is working as the Director with the International Center for AI and Cyber Security Research and Innovations and a Full Professor with the Department of Computer Science and Information Engineering (CSIE), Asia University, Taiwan. He is also serving as a Distinguished Research Scientist with LoginRadius Inc., USA, which is one of leading cybersecurity companies in the world, especially in the field of customer identity and access management (CIAM). He is also a Visiting/Adjunct Professor with several universities worldwide. In more than 16 years of his professional experience, he has published over 400 papers in journals/conferences including 30 books and eight patents with over 14000 citations. His research interests include information security, cyber physical systems, cloud computing, blockchain technologies, intrusion detection, AI, social media and networking. He has received numerous national and international awards including the Canadian Commonwealth Scholarship in 2009 and the Faculty Research Fellowship Award in 2017 from MeitY, Government of India, IEEE GCCE Outstanding and WIE Paper Awards, and the Best Faculty Award in 2018 and 2019, NIT, Kurukshetra. He is also selected in the 2021 and 2020 Stanford University's ranking of the world's top 2% scientists. He is also selected as the 2021 Distinguished Lecturer in IEEE CTSoc. He is also serving as the Member-in-Large and the Board of Governors for the IEEE Consumer Technology Society (2022-204). He is also leading *IJSWIS*, *IJSSCI*, and *IJCAC*, IGI Global, as the Editor-in-Chief. Moreover, he is also serving as a Lead Editor for a Book Series with CRC, World Scientific, and IET Press. He also served as a TPC Member and the Organized/Special Session Chair for ICCE-2021 and GCCE 2014–2021, the TPC Chair for the 2018 INFOCOM: CCSNA Workshop, and the Publicity Co-Chair for 2020 ICCCN.

**Akshat Gaurav** (Graduate Student Member, IEEE) received the master's degree in computer engineering (cyber security) from the National Institute of Technology, Kurukshetra, Haryana, India. He is currently a Cyber Security Researcher at Ronin Institute. He is also working in DDoS attack detection, intrusion detection, the IoT security, cloud/fog computing, and cryptography. He has published research papers in renowned journals and transections like IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (6.492 impact factor), *Neural Computing and Applications* (5.606 impact factor), *Journal of King Saud University-Computer and Information Sciences* (13.473 impact factor), and *Technological Forecasting and Social Change* (8.593 impact factor). He received the Best Paper Award at the International Conference on Smart Systems and Advanced Computing (SysCom 2021). In addition to his research, he has actively participated in other activities, such as volunteering at the International Conference on Cyber Security, Privacy, and Networking (ICSPN 2021), and a TCP Member at the IEEE Fourth International Workshop on Data-Driven Intelligence for Networks and Systems (DDINS 2022). He is also an active reviewer in leading journals like IEEE TRANSACTIONS OF INTELLIGENT TRANSPORT SYSTEMS (6.492 impact factor), IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS (10.215 impact factor), IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS (4.102 impact factor), *International Journal of Computer Communications* (Elsevier) (3.167 impact factor), *Journal of Signal Processing Systems* (Springer) (1.348 impact factor), *International Journal of Big Data Research* (Elsevier) (3.578 impact factor), *Journal of Innovation & Knowledge* (Elsevier) (9.269 impact factor), and *Journal of Sustainable Technology and Entrepreneurship* (Elsevier).

**Enrique Caño Marín** received the B.Sc. degree in industrial engineering and the M.Sc. degree in robotics and automation. He is currently pursuing the Ph.D. degree in information and knowledge engineering with the Computer Science Department, University of Alcalá. He is a Global IT Business Analyst at Roche. His research interests include social network analysis (SNA), artificial intelligence (AI), machine learning, and applied automation.

**Wadee Alhalabi** received the master's and Ph.D. degrees in electrical and computer engineering from the University of Miami in 2004 and 2008, respectively. He is currently a Professor of computer science at King Abdulaziz University (KAU). Prior to joining KAU, he worked in industry including Saudi Arabia Monetary Agency and Jeddah Desalinations and Power Generation plant as an Electronics Technician and a Security Systems Engineer. He has published more than 60 articles in virtual reality, image processing, rehabilitation engineering, and machine learning. His research interests include virtual reality and machine learning. Currently, he is a Former Member of the University Scientific Council and a member of the Promotion Committee. Currently, he is the Editor-in-Chief of the *Journal of King Abdulaziz University-Computing and Information Technology Sciences*.