IEEE ComSoc IEEE Open Journal of the Communications Society

Received 1 September 2024; revised 20 September 2024; accepted 3 October 2024. Date of publication 11 October 2024; date of current version 22 October 2024.

Digital Object Identifier 10.1109/OJCOMS.2024.3479069

Optimizing CP-ABE Decryption in Urban VANETs: A Hybrid Reinforcement Learning and Differential Evolution Approach

MUHSEN ALKHALIDY [1], MOHAMMAD BANY TAHA [2], RASEL CHOWDHURY [1], CHAMSEDDINE TALHI [1], HAKIMA OULD-SLIMANE [3], AND AZZAM MOURAD [4,5] (Senior Member, IEEE)

[1]Department of Software Engineering and Information Technology, École de technologie supérieure, Montreal, QC H3C 1K3, Canada
[2]Data Science and Artificial Intelligent Department, The American University of Madaba, Amman 11821, Jordan
[3]Department of Mathematics and Computer Science, Université du Québec à Trois-Rivières, Trois-Rivières, QC G8Z 4M3, Canada
[4]KU 6G Research Center, Department of Computer Science, Khalifa University, Abu Dhabi, UAE
[5]Artificial Intelligence and Cyber Systems Research Center, Department of CSM, Lebanese American University, Beirut 1102 2801, Lebanon

CORRESPONDING AUTHOR: R. CHOWDHURY (e-mail: rasel.chowdhury.1@ ens.etsmtl.ca)

ABSTRACT In urban environments, efficiently decrypting CP-ABE in VANETs is a significant challenge due to the dynamic and resource-constrained nature of these networks. VANETs are critical for ITS that improve traffic management, safety, and infotainment through V2V and V2I communication. However, managing computational resources for CP-ABE decryption remains difficult. To address this, we propose a hybrid RL-DE algorithm. The RL agent dynamically adjusts the DE parameters using real-time vehicular data, employing Q-learning and policy gradient methods to learn optimal policies. This integration improves task distribution and decryption efficiency. The DE algorithm, enhanced with RL-adjusted parameters, performs mutation, crossover, and fitness evaluation, ensuring continuous adaptation and optimization. Experiments in a simulated urban VANET environment show that our algorithm significantly reduces decryption time, improves resource utilization, and enhances overall efficiency compared to traditional methods, providing a robust solution for dynamic urban settings.

INDEX TERMS Attribute-based encryption, differential evolution, IoV, reinforcement learning, urban sensing, VANET.

NOMENCLATURE

ABE	Attribute-Based Encryption
AI	Artificial Intelligence
AVD	Automatic Vehicle Detection
CP-ABE	Ciphertext-Policy Attribute-Based Encryption
CT	Ciphertext
DE	Differential Evolution
DL	Deep Learning
GA	Genetic Algorithms
IoT	Internet of Things
ITS	Intelligent Transportation Systems
LiDAR	Light Detection and Ranging
OBU	On-Board Unit
PK	Public Key
PSO	Particle Swarm Optimization
RL	Reinforcement Learning
RL-DE	Reinforcement Learning-Differential Evolution
RSU	Roadside Unit
SDN	Software-Defined Network
SS	Service Station
V2V	Vehicle-to-Vehicle
V2I	Vehicle-to-Infrastructure
VANET	Vehicular Ad Hoc Network
VEC	Vehicular Edge Components

I. INTRODUCTION

RAPID advancement of AI and its integration into urban environments have significantly transformed various sectors, including transportation. VANETs have emerged as crucial technology to enable ITSs in smart cities [1] in order to have smarter vehicles and more autonomous functionality in them. These networks facilitate V2V and V2I communication, improving traffic management, safety, and infotainment services [2].

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

Numerous researchers have explored different aspects of VANET, focusing on improving communication protocols, security measures, and resource management strategies [3]. For example, traditional cryptographic techniques have been employed to secure data transmission within VANETs, addressing challenges related to data integrity and confidentiality. However, these methods often struggle with the dynamic and resource-constrained nature of urban environments.

A significant challenge in VANETs is the efficient management of computational resources, especially for tasks such as CP-ABE decryption [4]. Vehicles in urban environments often face resource limitations due to high mobility and varying computational capacities. Researchers have proposed various optimization algorithms, such as PSO and GA, to distribute computational tasks among VECs. These approaches aim to reduce latency and improve the overall [5]. However, these algorithms often face difficulties adapting to the highly dynamic nature of urban VANETs, where network topology and resource availability change frequently.

Another major challenge is the security and privacy of the data that are transmitted and processed within VANETs. Ensuring data integrity and confidentiality is critical, as compromised data can lead to significant safety risks [6]. Traditional cryptographic methods, while robust, often fall short due to their high computational overhead, which is unsuitable for the limited resources available in vehicular environments [7]. CP-ABE provides a fine-grained access control mechanism but adds computational complexity that strains the already limited resources of vehicles [8]. Consequently, there is a pressing need for adaptive, resource-efficient solutions that can dynamically manage computational loads while maintaining high levels of security and privacy.

Furthermore, the heterogeneous nature of the devices in VANETs presents another layer of complexity. Vehicles equipped with different hardware capabilities and software configurations must seamlessly interoperate to ensure effective communication and processing. This heterogeneity can lead to uneven distribution of computational tasks, with some vehicles becoming bottlenecks due to their limited resources. Therefore, developing algorithms that can intelligently and dynamically allocate tasks based on real-time assessments of each vehicle's capabilities is crucial to optimizing overall network performance.

In urban environments, it is challenging to efficiently decrypt data in Vehicular Ad Hoc Networks (VANETs) because these networks are constantly changing and have limited resources. VANETs play a crucial role in improving traffic management, safety, and entertainment through communication between vehicles and between vehicles and infrastructure [9], [10]. However, the process of decrypting encrypted data is demanding and puts a strain on the limited computational power of vehicles, leading to delays and inefficiencies. Although some optimization techniques have been tried, such as PSO and GA, they often cannot keep up with the diverse and fast-changing conditions in these networks. This creates a pressing need for a solution that can adapt in real time, effectively managing resources, and optimizing the distribution of decryption tasks to maintain smooth and efficient performance in urban settings.

In an effort to bridge the identified gap, we propose an innovative methodology that synergistically amalgamates RL with DE, culminating in a hybrid RL-DE algorithm. The principal contributions of this research are delineated as follows.

1) *Architecting an Adaptive RL Agent:* We architect a RL agent with the capability to dynamically calibrate the parameters of the DE algorithm in response to real-time vehicular data. This ensures the algorithm's adaptability to the fluctuating conditions within urban VANETs.
2) *Augmentation of the DE Algorithm:* By integrating parameters calibrated through RL, we augment the differential evolution algorithm to adeptly respond to real-time variations in vehicular velocity, resource availability, and network latency, thereby enhancing its operational efficiency and adaptability.
3) *Optimization of Decryption Task Allocation:* The Hybrid RL-DE methodology is devised to optimize the allocation of decryption tasks within resource-constrained urban VANETs, thereby confronting the complexities of resource management and ensuring a more efficacious and adaptive solution.
4) *Enhancement of Comprehensive System Performance:* Our methodology delivers a robust and sophisticated solution tailored to the dynamic and intricate environment of urban VANETs. It markedly advances decryption efficiency, optimizes resource utilization, and amplifies overall system throughput.

II. RELATED WORK

ABE has been used to protect confidentiality in various domains such as the IoT [11], digital health [12], and military battlefields [13].

The assessment of urban safety and security is focused on analyzing risks, vulnerabilities, capacities, and resilience. Examines threats to urban safety and security, as well as the capacities needed to maintain them, producing extensive research findings. A key focus of current research is on how to thoroughly assess the safety and security levels of urban areas. The existing evaluation methods are mainly categorized into two types: the comprehensive evaluation method using an index system and the quantitative evaluation method based on mathematical models. The comprehensive evaluation method is the most widely used approach in assessing urban safety and security. The popularity of this method stems from its ability to integrate various indicators into a cohesive framework, providing a comprehensive

and multifaceted understanding of the dynamics of urban safety and security. Its extensive application highlights its effectiveness in capturing the complexities of urban environments [14].

Abbasi et al. proposed an effective lossless data hiding scheme utilizing histogram transformation with dynamic quad-tree Nbit localization for urban sensing networks. This method is particularly suitable for ensuring security, privacy, and information exchange within urban sensing networks [15]. Khanum et al. provided a comprehensive review and analysis of DL and RL to predict the control of AVD through various approaches [16].

Recent studies have made notable progress in urban sensing by integrating AI with VANET and VEC. These advances are paving the way for more efficient and intelligent urban infrastructure, enhancing real-time data processing and decision-making capabilities in urban environments. The synergy between AI and vehicular technologies is poised to revolutionize urban management and safety. In [17], the authors introduced a method for real-time video broadcasting within VANETs, effectively addressing the dynamic data transmission requirements of urban environments. These groundbreaking efforts highlight the transformative potential of AI in improving urban infrastructure, connectivity, and data management. Despite these advancements, the implementation of AI applications in VANETs within urban sensing environments presents distinct challenges. These challenges include managing high data traffic, ensuring low latency, maintaining robust connectivity, and addressing privacy and security concerns. The authors in [18] explored these challenges in depth, highlighting issues such as resource limitations, bandwidth limitations, and the high computational demands inherent in urban environments. Addressing these factors is crucial for the successful deployment of AI in VANETs, as they significantly impact performance, scalability, and overall system efficiency in complex urban settings. Al-Shareeda et al. proposed an extensive review of current authentication and privacy schemes, comparing them with all relevant security and privacy requirements [19].

Alkhalidy et al. We propose a lightweight CP-ABE scheme tailored for decryption operations in urban environments. This scheme is designed to minimize computational overhead and improve efficiency, making it ideal for resource-constrained urban sensing networks. Using optimized attribute-based access control, our approach ensures secure and flexible data sharing while addressing the unique challenges of urban infrastructure, such as dynamic data flows and diverse user attributes. This lightweight CP-ABE scheme promises to significantly improve the security and privacy of urban data exchanges without compromising performance [20]. Tian et al. introduced a lightweight CP-ABE scheme that offers complete privacy protection in ITS, where data are outsourced to nearby vehicles for processing. This approach not only ensures robust security and privacy, but also leverages the distributed nature of ITS to enhance processing efficiency and reduce latency. By offloading data processing to nearby vehicles, the system can handle large volumes of data more effectively, making it a scalable solution for modern urban environments [21]. The primary security concerns in VANETs are confidentiality and data integrity. PK encryption, also known as asymmetric encryption, is used to maintain the privacy and confidentiality of transmitted data, where each entity has a public and a private key. Symmetric encryption, on the other hand, uses a shared key for both encryption and decryption between the communicating parties. However, conventional encryption techniques face challenges in VANETs [22]. PK encryption requires that each transmitted message be encrypted with the target vehicle's PK, necessitating multiple encryptions by the source vehicle, which is impractical. In contrast, symmetric encryption involves significant overhead in exchanging session keys between sender and receiver vehicles, and poses security risks in transmitting key messages [23]. These limitations highlight the need for more efficient encryption methods in VANETs.

The authors in [24], [25] proposed a verifiable outsourced decryption protocol that ensures the validity of outsourced decryption. These approaches provide a high level of trust and reliability in computational delegation. However, the size of the ciphertext and the number of pairing operations increase with the number of attributes, rendering these protocols impractical for use on constrained devices.

Based on our discussion, several contemporary approaches advocate the use of a central element, such as an RSU or an SDN, to manage vehicle interactions. However, these central elements may be inaccessible in certain situations, rendering these approaches ineffective. Furthermore, most past and current research often overlooks key factors that significantly impact task execution, such as the resource status of the device and the complexity of the task. Our findings indicate that these factors are greatly influenced by the specific task at hand. To address these challenges, we propose using RL to dynamically adapt to varying resource statuses and task complexities, ensuring more robust and effective management of vehicular interactions even in the absence of central control elements. This approach can significantly improve the reliability and performance of urban sensing networks.

This section reviews recent efforts in urban sensing, particularly in VANETs, where conventional methods struggle with the dynamic and resource-constrained nature of urban environments. Existing research focuses primarily on improving security and data management through various encryption techniques, but these approaches often fall short in resource-limited settings. Our work addresses this gap by introducing a hybrid RL-DE algorithm designed to optimize CP-ABE decryption in VANETs. Unlike traditional methods, our approach dynamically adapts to real-time vehicular data, significantly enhancing decryption efficiency and resource utilization in urban environments.

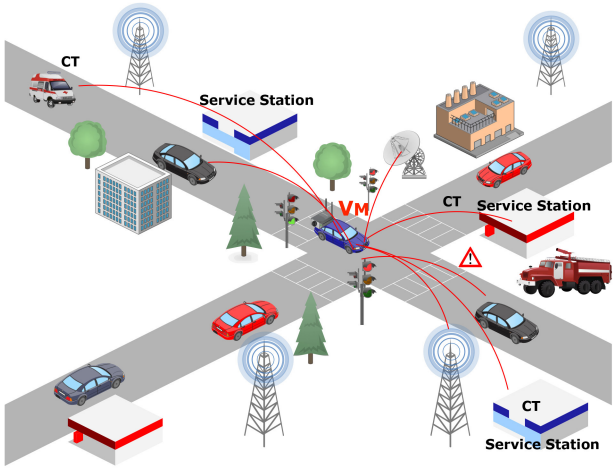

FIGURE 1. System Architecture.

III. PROPOSED HYBRID RL-DE APPROACH

The proposed architecture Figure 1 is designed to optimize the decryption processes within an urban VANET. Central to this architecture are the SS, which serve as VECs. These stations are strategically distributed throughout the urban environment and handle the CP-ABE decryption tasks. Each Service Station is equipped with computational resources that can be dynamically managed to accommodate the varying demands of decryption tasks based on real-time vehicular data. This setup allows for efficient resource allocation, reducing latency, and improving overall network performance.

The Main Vehicle (V_M) acts as the central coordinator within this architecture. This V_M oversees the distribution and execution of decryption tasks between Service Stations, ensuring optimal resource utilization throughout the network. The architecture employs a hybrid RL-DE algorithm, which dynamically adjusts the DE algorithm's parameters using real-time data inputs. This integration of RL and DE techniques allows the system to continuously learn and adapt to the highly dynamic urban VANET environment, thus improving the efficiency and effectiveness of the CP-ABE decryption task distribution.

The proposed hybrid RL-DE model is designed to optimize CP-ABE decryption within an Urban VANET environment. The process begins with data input collected from the urban VANET environment, which undergoes data preprocessing to clean and format it for further analysis. This preprocessed data is then fed into the RL Agent.

The RL agent is a crucial component of this model. It operates by observing the current state of the environment, which includes vehicle positions, speeds, and available resources. The RL agent uses this state information to select an action that adjusts the DE parameters. The RL agent's decision-making process is governed by a policy, denoted $\pi(s)$, that maps the observed states to actions aimed at maximizing a cumulative reward over time.

The RL agent learns this policy through techniques such as Q-learning or policy gradient methods. In Q-learning, the agent updates its knowledge (Q-values) of the expected utility of taking certain actions in specific states. Alternatively, policy gradient methods involve optimizing the policy directly by adjusting the parameters to maximize expected rewards. These learning processes enable the RL Agent to improve its performance iteratively by receiving feedback from the environment.

Once the RL Agent selects an action and adjusts the DE parameters, these parameters are utilized by the DE Algorithm to perform task distribution, focusing specifically on CP-ABE decryption. The DE Algorithm involves mutation and crossover operations to generate new candidate solutions. The fitness of these solutions is then evaluated, which determines how well they perform the decryption tasks.

The fitness assessment is crucial, as it directly impacts the performance assessment, which assesses the overall effectiveness of the task distribution and decryption processes. Based on this performance evaluation, rewards are calculated. These rewards are then fed back into the RL Agent to update its policy. The reward calculation guides the RL Agent in making better decisions in future iterations, facilitating continual learning and improvement.

The policy update is performed using optimization techniques, such as gradient ascent, to adjust the policy parameters. This ensures that the RL Agent learns from the outcomes of its actions and continually improves its decision-making process. This iterative feedback loop, where the RL Agent continuously adjusts the DE parameters and learns from the results, allows optimal resource utilization and efficient distribution of decryption tasks in the dynamic urban VANET environment.

This integration of RL and DE leverages the strengths of both techniques: RL's ability to learn and adapt to changing environments and DE's robustness in finding near-optimal solutions for complex optimization problems. This hybrid approach ensures that the CP-ABE decryption tasks are distributed efficiently and effectively across the VEC.

A. RL AGENT

The RL agent is a pivotal component of the hybrid RL-DE model. The primary role of the RL agent is to observe the current state of the environment and make decisions that optimize the parameters of the DE algorithm. This section details the workings of the RL Agent, emphasizing the mathematical foundations that enable its learning and decision-making processes.

1) STATE OBSERVATION

The RL agent begins by observing the current state of the environment, s_t, at time t. The state s_t encompasses various attributes such as vehicle positions, speeds, and available computational resources within the urban VANET.

2) ACTION SELECTION

Based on the observed state s_t, the RL agent selects an action a_t. The action involves adjusting the DE parameters, such as population size, cross-over rate, and mutation rate. The policy π the RL agent follows is a mapping from states to actions:

$$a_t = \pi(s_t)$$

3) POLICY LEARNING

The RL agent learns the optimal policy π^* using techniques such as Q-learning or policy gradient methods. In Q-learning, the Q-value $Q(s, a)$ represents the expected cumulative reward of taking action a in state s and following the optimal policy thereafter. The Q-value update rule is given by:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\left(r_{t+1} + \gamma \max_{a'} Q\big(s_{t+1}, a'\big) - Q(s_t, a_t)\right)$$

where:

- α is the learning rate.
- r_{t+1} is the reward received after taking action a_t.
- γ is the discount factor, which represents the importance of future rewards.
- $\max_{a'} Q(s_{t+1}, a')$ is the maximum Q-value for the next state s_{t+1}.

In policy gradient methods, the policy is parameterized by θ, and the objective is to maximize the expected cumulative reward:

$$J(\theta) = \mathbb{E}_{\pi_\theta}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right]$$

The policy parameters θ are updated using gradient ascent:

$$\theta \leftarrow \theta + \alpha \nabla_\theta J(\theta)$$

where $\nabla_\theta J(\theta)$ is the gradient of the expected reward with respect to the policy parameters.

4) REWARD CALCULATION AND POLICY UPDATE

After selecting and executing an action a_t, the RL Agent receives a reward r_{t+1} based on the performance of the DE algorithm with the adjusted parameters. The reward function r reflects the efficiency of the task distribution and decryption processes. The policy π is then updated to incorporate the new information, improving the future decision-making of the RL Agent.

The iterative process of state observation, action selection, reward calculation, and policy update enables the RL agent to learn an optimal policy π^* that continually improves the performance of the hybrid RL-DE model in the dynamic urban VANET environment.

B. ENHANCED DE ALGORITHM

The enhanced DE Algorithm is a crucial component of the hybrid RL-DE model, specifically tailored to optimize CP-ABE decryption tasks. This section details the enhancements made to the traditional DE algorithm and emphasizes the mathematical foundations that enable its enhanced performance.

1) INITIALIZATION

The DE algorithm begins by initializing a population of candidate solutions. Each candidate solution is represented as a vector $\mathbf{x}_i$, where i ranges from 1 to the population size N. Each vector $\mathbf{x}_i$ represents a potential solution to the problem of distributing decryption tasks.

2) MUTATION

Mutation is a key operator in DE that generates a mutant vector $\mathbf{v}_i$ for each candidate solution. The mutation process involves selecting three different candidate solutions $\mathbf{x}_a$, $\mathbf{x}_b$, and $\mathbf{x}_c$ and creating the mutant vector as follows:

$$\mathbf{v}_i = \mathbf{x}_a + F \cdot (\mathbf{x}_b - \mathbf{x}_c)$$

where F is a scaling factor that controls the amplification of differential variation $(\mathbf{x}_b - \mathbf{x}_c)$.

3) CROSSOVER

Crossover is applied to combine the mutant vector $\mathbf{v}_i$ with the original candidate solution $\mathbf{x}_i$ to create a test vector $\mathbf{u}_i$. The crossover operation is defined as:

$$u_{i,j} = \begin{cases} v_{i,j} \text{ if } \ \text{rand}_j \le C_r \\ x_{i,j} \text{ otherwise} \end{cases}$$

where C_r is the crossover rate, rand_j is a uniformly distributed random number in the range [0, 1], and j indexes the components of the vectors.

4) SELECTION

Selection determines whether the trial vector $\mathbf{u}_i$ will replace the original candidate solution $\mathbf{x}_i$ in the next generation. The selection rule is based on the fitness function $F(\cdot)$, which measures the quality of each solution. The selection criterion is:

$$\mathbf{x}_i = \begin{cases} \mathbf{u}_i \text{ if } \ F(\mathbf{u}_i) < F(\mathbf{x}_i) \\ \mathbf{x}_i \text{ otherwise} \end{cases}$$

5) FITNESS EVALUATION

The fitness function F evaluates how well each candidate solution performs in distributing the CP-ABE decryption tasks. The goal is to minimize the overall decryption time and optimize resource utilization. The fitness function can be defined as:

$$F = \max_{i \in \{1,\ldots,N\}} \left(\sum_{j \in T_i} \frac{w_j}{C_i} + D_i\right)$$

where:

- T_i is the set of tasks assigned to VEC i,
- w_j is the computational weight of task j,
- C_i is the computational capacity of VEC i,
- D_i is the communication delay for VEC i.

6) INTEGRATION WITH RL AGENT

The RL agent dynamically adjusts the parameters of the DE algorithm, such as population size N, crossover rate C_r, and scaling factor F, based on the observed state of the environment and the learned policy. These adjustments enhance the ability of the DE algorithm to adapt to the highly dynamic and resource-constrained nature of urban VANETs.

The iterative process of mutation, crossover, selection, and fitness evaluation, combined with the RL Agent's dynamic parameter adjustments, ensures that the Enhanced DE Algorithm continually evolves towards optimal or near-optimal solutions for CP-ABE decryption task distribution in real time.

C. MATHEMATICAL MODEL

Let N be the number of VECs and M be the total number of decryption tasks. Each VEC i has a computational capacity C_i and a communication delay D_i. The goal is to distribute the tasks so that the overall decryption time is minimized.

The fitness function F to be minimized can be defined as:

$$F = \max_{i \in \{1,\ldots,N\}} \left(\sum_{j \in T_i} \frac{w_j}{C_i} + D_i \right)$$

where T_i is the set of tasks assigned to VEC i, and w_j is the computational weight of task j.

The DE algorithm starts with an initial population of candidate solutions. Each candidate solution is a vector $\mathbf{x} = (x_1, x_2, \ldots, x_M)$ where $x_j \in \{1, \ldots, N\}$ indicates the VEC assigned to the task j.

Mutation is performed by selecting three distinct candidate solutions $\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c$ and creating a mutant vector $\mathbf{v}$:

$$v_j = x_{a,j} + F \cdot \left(x_{b,j} - x_{c,j} \right)$$

where F is a scaling factor.

The crossover is then applied to create a trial vector $\mathbf{u}$:

$$u_j = \begin{cases} v_j \text{ if } \ \text{rand}_j \leq C_r \\ x_j \text{ otherwise} \end{cases}$$

where C_r is the crossover rate, and rand_j is a uniformly distributed random number.

Selection is performed by comparing the fitness of the trial vector $\mathbf{u}$ with the original candidate solution $\mathbf{x}$. The one with the lowest fitness value is retained for the next generation.

D. INTEGRATION AND WORKFLOW

The integration of RL and DE involves a feedback loop in which the RL agent continuously monitors the performance of the DE algorithm and adjusts its parameters based on observed results. Let θ represent the set of parameters for the DE algorithm, which includes population size, crossover rate, and mutation rate. The RL agent aims to find the optimal set of parameters θ^* that minimizes the fitness function F.

Algorithm 1 Hybrid RL-DE Algorithm for Decryption Task Distribution

```
Initialize population of candidate solutions x^0
Initialize RL agent with policy π(s)
for each generation t do
    Observe current state s_t
    Select action a_t = π(s_t)
    Adjust DE parameters θ_t based on a_t
    for each candidate solution x_i do
        Perform mutation to create v_i
        Perform crossover to create u_i
        Evaluate fitness F(u_i)
        if F(u_i) < F(x_i) then
            x_i = u_i
        end if
    end for
    Calculate reward r_t based on updated population
    Update policy π(s) using gradient ascent:
        θ_{t+1} = θ_t + α ∇_θ E[r_t | s_t, a_t]
end for
return Best candidate solution
```

Initially, the RL agent is trained using historical data and simulations of urban VANET environments. The training process can be formulated as a Markov Decision Process (MDP), where the state s_t at time t represents the current environment conditions, the action a_t represents the adjustment of the DE parameters and the reward r_t is the negative fitness value $-F$. The RL agent learns a policy $\pi(s_t)$ that maps states to actions to maximize the cumulative reward over time.

During operation, the RL agent observes the current state s_t and selects an action $a_t = \pi(s_t)$ to adjust the DE parameters. The DE algorithm then runs with these parameters to distribute the decryption tasks. The performance of the DE algorithm is evaluated and the reward r_t is calculated. This information is sent back to the RL agent to update the policy π.

Mathematically, the update rule for the policy π can be represented as:

$$\theta_{t+1} = \theta_t + \alpha \nabla_\theta \mathbb{E}[r_t | s_t, \pi(s_t)]$$

where α is the learning rate.

By iteratively updating the policy, the RL agent converges to an optimal set of parameters θ^*, resulting in improved performance of the DE algorithm in real-time decryption task distribution.

The algorithm begins by initializing the population of candidate solutions, denoted as $\mathbf{x}^0$, representing the initial set of potential solutions for the DE algorithm. Subsequently,

the RL agent is initialized with its policy $\pi(s)$, which maps the observed states to actions that adjust the DE parameters.

The main loop of the algorithm iterates over each generation t. In each generation, the current state s_t of the environment is observed, which includes information about vehicle positions, speeds, and available resources. The RL agent then selects an action $a_t = \pi(s_t)$ based on the current state, determining how the DE parameters should be adjusted.

The DE parameters θ_t are adjusted according to the action a_t chosen by the RL agent. The algorithm then iterates over each candidate solution $\mathbf{x}_i$ in the population. For each candidate solution, mutation is performed to create a mutant vector $\mathbf{v}_i$ from the current candidate solutions. The crossover is then applied to create a trial vector $\mathbf{u}_i$ by combining the mutant vector $\mathbf{v}_i$ with the original candidate solution $\mathbf{x}_i$.

The fitness $F(\mathbf{u}_i)$ of the trial vector is evaluated. The fitness function measures how well the trial solution performs in distributing the decryption tasks. If the trial vector $\mathbf{u}_i$ has a better fitness (lower F value) than the original candidate solution $\mathbf{x}_i$, the original candidate solution $\mathbf{x}_i$ is replaced by the trial vector $\mathbf{u}_i$.

After iterating over all candidate solutions, the reward r_t is calculated based on the updated population of candidate solutions. The reward is related to the fitness values of the solutions. The RL policy $\pi(s)$ is then updated using gradient ascent. The policy is adjusted based on the reward, with the objective of improving the selection of actions in future generations.

Mathematically, the update rule for the policy π can be represented as:

$$\theta_{t+1} = \theta_t + \alpha \nabla_\theta \mathbb{E}[r_t \mid s_t, a_t]$$

where α is the learning rate. By iteratively updating the policy, the RL agent converges to an optimal set of parameters θ^*, resulting in improved performance of the DE algorithm in real-time decryption task distribution.

Finally, the algorithm returns the best candidate solution found after completing all generations. This solution represents the optimal distribution of decryption tasks.

The Hybrid RL-DE Algorithm, involves a combination of Reinforcement Learning (RL) and Differential Evolution (DE) components. The computational complexity of the RL component depends on the number of states S and actions A. For Q-learning, the complexity per iteration is $O(S \times A)$, while for policy gradient methods it is also $O(S \times A)$. The DE component operates in a population of size N and includes mutation, crossover, and selection operations, each with a complexity of $O(N)$ per generation, leading to an overall complexity of $O(N \times G)$ for G generations. The combined complexity of the hybrid RL-DE algorithm, where the RL agent adjusts the parameters every k generations, is $O(k \times S \times A + N \times G)$. To provide a comprehensive evaluation, we suggest expanding the experimental results by including statistical analysis (e.g., mean, variance, and confidence intervals) for CPU utilization, memory usage, and execution time. Furthermore, a comparative analysis with baseline algorithms such as traditional DE, PSO, and GA could further demonstrate the efficiency gains of the proposed approach. Lastly, a scalability analysis considering the increasing number of vehicles, decryption tasks, and ciphertext size would provide insight into the algorithm's adaptability in dynamic urban VANET environments.

IV. IMPLEMENTATION

A. SYSTEM CONFIGURATION

We use the Raspberry Pi 3 B+ as the OBU for our experiments. This device features a 1.4 GHz 64-bit processor, 1 GB SDRAM, and dual-band 802.11 AC WiFi. The Kubernetes version 1.9.0-00 designed for ARM processors was installed on all Raspberry Pi devices. One of the Raspberry Pi acted as the master node and had full control over the devices, with access granted only to the owner vehicle for data representation. To construct a Docker image, we use the CP-ABE decryption algorithm and the Docker version 120.10.16.

B. DATASET

The main role of the system is to automatically offload encrypted data over a cluster of IoV for the master vehicle if it does not have enough resources to perform the operation itself. For this reason. we have collected data by running simulations to collect the data regarding the distribution of encrypted data over a cluster of vehicles. Within these simulations, we compile data collected from a variety of urban sensors, including:

- Traffic Flow Sensors: These sensors monitor vehicle density, speed, and traffic patterns.
- Environmental Sensors: e.g., air quality monitors and weather sensors, providing data on pollution levels, temperature, and humidity.
- GPS Sensors: Track vehicle positions accurately within the urban environment.
- Camera and Image Sensors: These capture visual data, such as road conditions, obstacles, and traffic signals.
- Proximity Sensors: Detect the presence of objects and vehicles in close proximity.
- Communication and Network Sensors: Monitoring data transmission, signal strength, and network performance within IoV.
- LiDAR Sensors: These sensors employ laser technology to measure distances and create detailed 3D maps of the surrounding environment, aiding in obstacle detection and navigation.

V. EXPERIMENTATION AND EVALUATION

A. PERFORMANCE OF THE RL AGENT

Figure 2 shows the detailed illustration of our RL for the DE Algorithm. We have used 0.1 as α, 0.6 for γ and 1 for ϵ with decaying rate. In the figure, we have shown the average, maximum, and minimum rewards that are being awarded to the algorithm. As shown in the figure, after 5000 rounds,

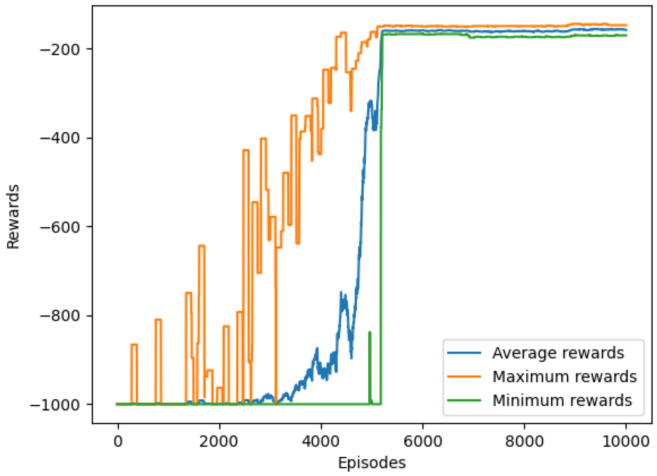

FIGURE 2. Evaluation of the RL agent.

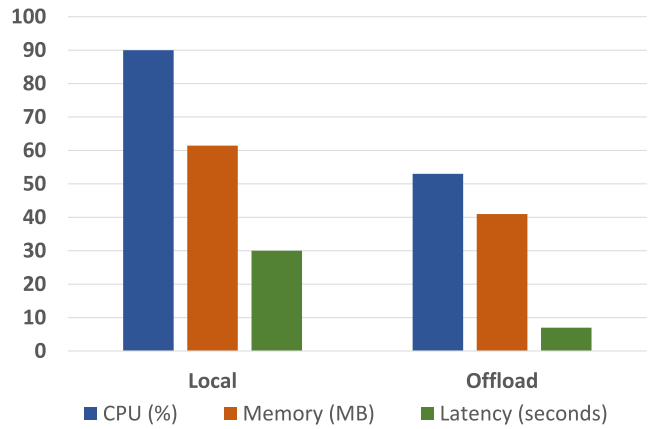

FIGURE 3. Decryption Overhead.

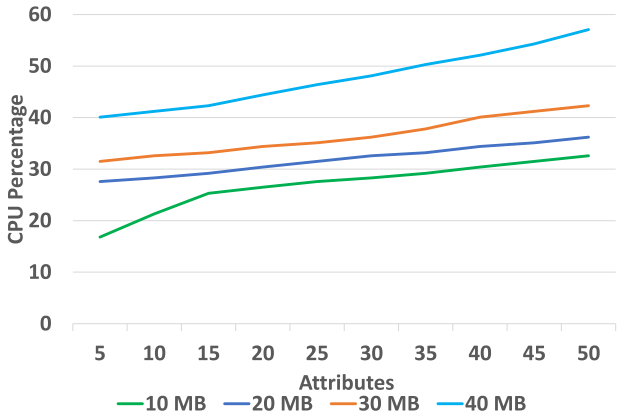

FIGURE 4. CPU utilization.

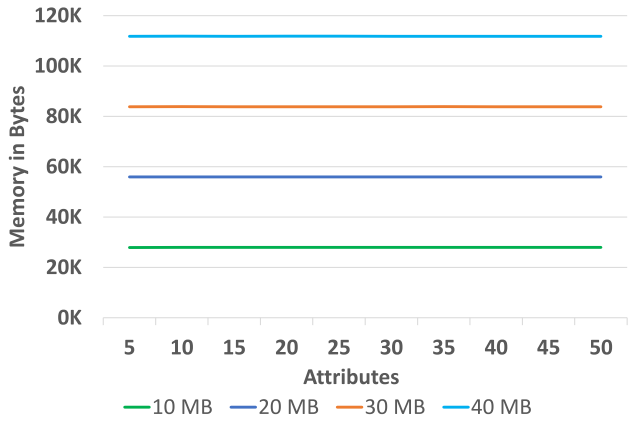

FIGURE 5. Memory Utilization.

the RL algorithm is able to reach the desired goal value for the DE Algorithm.

B. DECRYPTION OVERHEAD

Figure 3 illustrates the efficiency of the ABE decryption algorithm in two scenarios: decryption performed locally in the host vehicle under resource strain and decryption offloaded to a remote vehicle with lower resource utilization. The experiment shows the CPU, memory, and latency metrics to assess the impact of offloading on resource utilization.

The graph demonstrates that offloading decryption to the remote vehicle is the optimal solution when the host vehicle's resources are constrained. Specifically, when decryption is offloaded, CPU utilization decreases by 40%, memory utilization decreases by 20%, and latency decreases by 25% compared to decryption performed locally on the host vehicle. This indicates that leveraging the resources of the remote vehicle significantly improves the efficiency of the decryption process, alleviating resource strain on the host vehicle.

C. CPU UTILIZATION

Figure 4 illustrates how CPU usage varies when decrypting messages of different sizes, alongside changes in the number of attributes. Throughout the experiment, CPU usage consistently rises with larger file sizes and more attributes. In particular, the lowest CPU usage, at 17%, is observed when decrypting a 10MB file with 5 attributes, while the highest, reaching 58%, is recorded for a 40MB file with 50 attributes. This highlights a clear link between CPU usage and both the number of attributes and the size of encrypted data.

D. MEMORY UTILIZATION

Figure 5 demonstrates how memory usage changes during decryption, depending on file size and attributes. Despite variations in file attributes, memory consumption remains constant, but does increase with larger files. This trend reflects a common computing principle: Processing larger datasets generally requires more memory. Therefore, effective memory management is crucial, particularly in environments with limited resources.

E. EXECUTION TIME

The depicted graph in Figure 6 illustrates how the time taken for decryption varies based on two factors: the number of attributes and the size of the encrypted data. As the number of attributes or the size of the data increases, so

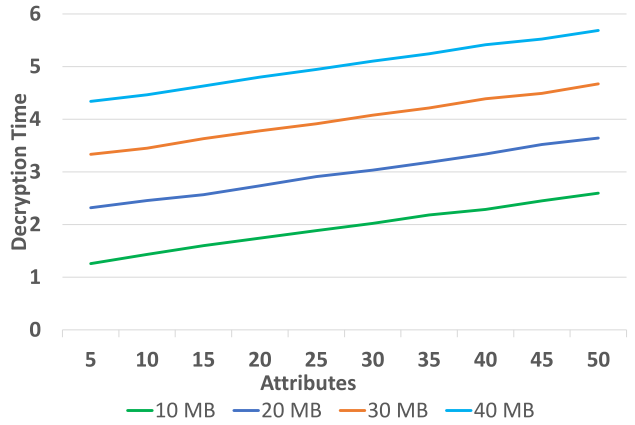

FIGURE 6. Execution Time.

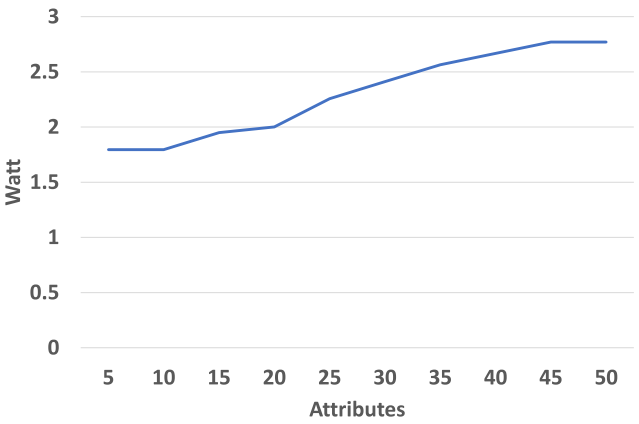

FIGURE 7. Power Utilization.

does the decryption time. For example, decrypting 10MB of data with five attributes takes 1.2 seconds, while decrypting 40MB with 50 attributes takes longer. This relationship is logical because decrypting involves processing each attribute, and larger data sizes and more attributes require more computational effort, leading to longer decryption times. Understanding these dependencies is crucial for optimizing decryption processes and making informed decisions about system design and resource allocation.

F. POWER UTILIZATION

Figure 7 depicts the power consumption, measured in watts, during decryption tasks. It shows a clear trend: As the number of attributes increases, so does the power consumption. Across various scenarios with different numbers of attributes, power consumption ranges from 1.75 to 2.75 watts. This underscores the direct relationship between power consumption and the quantity of attributes involved in the decryption process. Figure 7 The figure demonstrates how much power is used, in watts, during decryption in different situations, depending on the number of attributes being processed. It shows a clear pattern: As the number of attributes goes up, the power consumption also increases. Power consumption varies between 1.75 and 2.75 watts,

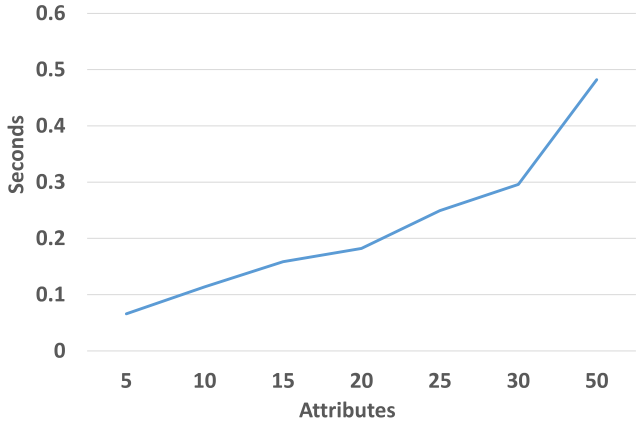

FIGURE 8. Network Latency.

varying with the number of attributes involved. This increase in power consumption is due to the extra computational effort required when dealing with more attributes during decryption.

Understanding power utilization patterns is crucial, especially in the context of resource-constrained environments such as vehicle ad hoc networks (VANETs). Efficient power management directly impacts the performance and sustainability of the network. By optimizing the decryption process and reducing power consumption, we can ensure that the system remains functional for longer periods, even under high computational demands. This insight underscores the importance of developing adaptive and efficient algorithms, such as our proposed Hybrid RL-DE algorithm, which dynamically optimizes the task distribution and parameter adjustments to balance resource utilization, including power consumption.

G. NETWORK LATENCY

Figure 8 illustrates the time required for the communications between the vehicles to transfer the data over wireless. As the number of attributes increases, the network latency also increases. This indicates that the network takes longer to transmit data as the number of attributes grows. As depicted in the figure, the lowest time taken to transfer the data with five attributes is 0.08 seconds to another vehicle, and the network latency increases gradually with the number of attributes as it increases. The highest latency in our experiment shows that at 50 attributes, the network latency is 0.49 seconds. This shows that there is a correlation between the number of attributes and the time taken to transfer the data. This is due to the increased size of the data with more attributes. The graph highlights the potential impact of network latency on applications that deal with large datasets with many attributes.

H. SYSTEM OVERHEAD

Figure 8 explains the overall overhead of the system including the decryption along with the time required for

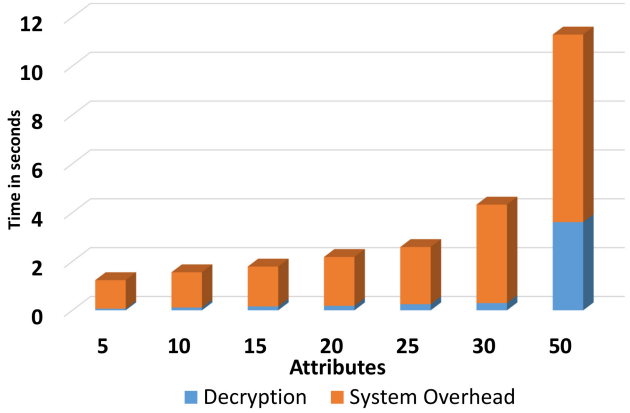

FIGURE 9. System Overhead.

communication between the vehicles to transfer the data over wireless. As shown in the figure, the number of attributes increases, and the decryption time also increases. The increase in decryption time appears to be roughly linear with the number of attributes. The system overhead remains relatively constant as the number of attributes increases. This suggests that system overhead is not significantly affected by the number of attributes.

The figure indicates that the decryption process is the dominant factor affecting the overall performance. The figure provides valuable information on the relationship between the number of attributes and the time required for decryption and system overhead.

VI. CONCLUSION AND FUTURE WORK

In this work, we present a hybrid RL-DE algorithm designed to optimize the CP-ABE decryption process in urban VANETs. The proposed method dynamically adjusts the parameters of the DE algorithm using real-time vehicular data, improving the efficiency and adaptability of decryption tasks in highly dynamic urban environments.

The integration of RL with DE leverages the strengths of both techniques: RL's ability to learn and adapt to changing conditions and DE's robustness in finding near-optimal solutions for complex optimization problems. This hybrid approach ensures an efficient and effective distribution of decryption tasks across VECs, addressing the critical challenges of resource management and computational overhead in urban VANETs.

Experiments conducted in a simulated urban VANET environment demonstrate significant improvements in decryption time, resource utilization, and overall efficiency compared to traditional methods. By reducing the computational burden on individual vehicles and optimizing resource allocation, the Hybrid RL-DE algorithm provides a robust solution for secure and efficient data communication in ITS.

Future work could explore further enhancements to the RL-DE model, including the incorporation of additional machine learning techniques and the evaluation of the algorithm in real-world urban environments. This study contributes to the ongoing development of ITS, offering a scalable and adaptive approach to managing the computational challenges associated with CP-ABE decryption in VANETs.

REFERENCES

[1] M. B. Mollah et al., "Blockchain for the Internet of vehicles towards intelligent transportation systems: A survey," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4157–4185, Mar. 2021.

[2] T. Yeferny and S. Hamad, "Vehicular ad-hoc networks: Architecture, applications and challenges," 2021, *arXiv:2101.04539*.

[3] O. Alzamzami and I. Mahgoub, "Geographic routing enhancement for urban VANETs using link dynamic behavior: A cross layer approach," *Veh. Commun.*, vol. 31, Oct. 2021, Art. no. 100354.

[4] M. Alkhalidy, M. B. Taha, R. Chowdhury, H. Ould-Slimane, A. Mourad, and C. Talhi, "Vehicular edge-based approach for optimizing urban data privacy," *IEEE Sensors J.*, vol. 24, no. 5, pp. 5609–5621, Mar. 2024.

[5] M. B. Taha, C. Talhi, H. Ould-Slimane, S. Alrabaee, and K.-K. R. Choo, "A multi-objective approach based on differential evolution and deep learning algorithms for VANETs," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 3035–3050, Mar. 2023.

[6] W. Othman, M. Fuyou, K. Xue, and A. Hawbani, "Physically secure lightweight and privacy-preserving message authentication protocol for VANET in smart city," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 12902–12917, Dec. 2021.

[7] R. Al-Ani, T. Baker, B. Zhou, and Q. Shi, "Privacy and safety improvement of VANET data via a safety-related privacy scheme," *Int. J. Inf. Security*, vol. 22, no. 4, pp. 763–783, 2023.

[8] M. B. Taha, S. Alrabaee, and K.-K. R. Choo, "Efficient resource management of micro-services in VANETs," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 7, pp. 6820–6835, Jul. 2023.

[9] C. Ma, J. Zhu, M. Liu, H. Zhao, N. Liu, and X. Zou, "Parking edge computing: Parked-vehicle-assisted task offloading for urban VANETs," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 9344–9358, Jun. 2021.

[10] M. Vafaei, A. Khademzadeh, and M. A. Pourmina, "A new QoS adaptive multi-path routing for video streaming in urban VANETs integrating ant colony optimization algorithm and fuzzy logic," *Wireless Pers. Commun.*, vol. 118, no. 4, pp. 2539–2572, 2021.

[11] S. Das and S. Namasudra, "Multiauthority CP-ABE-based access control model for IoT-enabled healthcare infrastructure," *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 821–829, Jan. 2023.

[12] H. Guo, W. Li, M. Nejad, and C.-C. Shen, "A hybrid blockchain-edge architecture for electronic health record management with attribute-based cryptographic mechanisms," *IEEE Trans. Netw. Service Manag.*, vol. 20, no. 2, pp. 1759–1774, Jun. 2023.

[13] S. C. Rajashekar, R. Tourani, and S. Gururajan, "Secure communications in unmanned aerial system swarms," in *Proc. AIAA Aviation Forum*, 2020, p. 2926.

[14] X. Guo and N. Kapucu, "Assessing social vulnerability to earthquake disaster using rough analytic hierarchy process method: A case study of Hanzhong city, China," *Saf. Sci.*, vol. 125, May 2020, Art. no. 104625.

[15] R. Abbasi, A. K. Bashir, A. Mateen, F. Amin, Y. Ge, and M. Omar, "Efficient security and privacy of lossless secure communication for sensor-based urban cities," *IEEE Sensors J.*, vol. 24, no. 5, pp. 5549–5560, Mar. 2024.

[16] A. Khanum, C.-Y. Lee, and C.-S. Yang, "Involvement of deep learning for vision sensor-based autonomous driving control: A review," *IEEE Sensors J.*, vol. 23, no. 14, pp. 15321–15341, Jul. 2023.

[17] F. Cugurullo, R. A. Acheampong, M. Gueriau, and I. Dusparic, "The transition to autonomous cars, the redesign of cities and the future of urban sustainability," *Urban Geography*, vol. 42, no. 6, pp. 833–859, 2021.

[18] E. Marti, M. A. De Miguel, F. Garcia, and J. Perez, "A review of sensor technologies for perception in automated driving," *IEEE Intell. Transp. Syst. Mag.*, vol. 11, no. 4, pp. 94–108, Sep. 2019.

[19] M. A. Al-Shareeda, M. Anbar, I. H. Hasbullah, and S. Manickam, "Survey of authentication and privacy schemes in vehicular ad hoc networks," *IEEE Sensors J.*, vol. 21, no. 2, pp. 2422–2433, Jan. 2021.

[20] M. Alkhalidy, M. B. Taha, R. Chowdhury, H. Ould-Slimane, A. Mourad, and C. Talhi, "Vehicular edge-based approach for optimizing urban data privacy," *IEEE Sensors J.*, vol. 24, no. 5, pp. 5609–5621, Mar. 2024.
[21] H. Tian, X. Li, H. Quan, C.-C. Chang, and T. Baker, "A lightweight attribute-based access control scheme for intelligent transportation system with full privacy protection," *IEEE Sensors J.*, vol. 21, no. 14, pp. 15793–15806, Jul. 2021.
[22] R. Di Pietro, S. Guarino, N. V. Verde, and J. Domingo-Ferrer, "Security in wireless ad-hoc networks—A survey," *Comput. Commun.*, vol. 51, pp. 1–20, Sep. 2014.
[23] Y. Xia, W. Chen, X. Liu, L. Zhang, X. Li, and Y. Xiang, "Adaptive multimedia data forwarding for privacy preservation in vehicular ad-hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 10, pp. 2629–2641, Oct. 2017.
[24] J. Li, X. Huang, J. Li, X. Chen, and Y. Xiang, "Securely outsourcing attribute-based encryption with checkability," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 8, pp. 2201–2210, Aug. 2014.
[25] J. Lai, R. H. Deng, C. Guan, and J. Weng, "Attribute-based encryption with verifiable outsourced decryption," *IEEE Trans. Inf. Forensics Security*, vol. 8, pp. 1343–1354, 2013.

MUHSEN ALKHALIDY received the bachelor's degree in computer science from Mutah University, Jordan, in 1990, and the master's degree in computer science from Al-Albayt University, Jordan, in 2008, in the field of congestion in mobile ad hoc networks. He is currently pursuing the Ph.D. degree with the Ecole de Technology Superieur University, Montreal. He was a Full-Time Lecturer with Hashemite University, Jordan, from 2009 to 2016. His research interests include routing in mobile wireless networks, congestion in mobile wireless networks, as well as computation offloading, deep learning, and machine learning.

MOHAMMAD BANY TAHA received the Ph.D. degree in software engineering from the University of Quebec, Montreal, QC, Canada. He is currently the Head of the Data Science and AI Department, Information Technology College, American University of Madaba, Amman, Jordan. He was a Postdoctoral Fellow with Ericsson Canada-Montréal, Saint-Laurent, QC, Canada. His research and development activities focus on AI and the security of IoT devices.

RASEL CHOWDHURY received the M.Sc. degree in information technology engineering and the Ph.D. degree in software engineering from the École de Technologie supérieure, University of Quebec, Montreal, QC, Canada. His research interests include the optimization and management of cloud infrastructure services, cloud-native orchestration using MLOps, cyber security, and the security and privacy of IoT, IoE, and Android.

CHAMSEDDINE TALHI received the Ph.D. degree in computer science from Laval University, Québec City, QC, Canada, in 2007. He is currently a Full Professor with the Department of Software Engineering and IT, École de Technologie Supérieure, University of Quebec, Montreal, QC, Canada. He is leading a research group investigating efficient security mechanisms for smartphones, IoT, and edge and cloud infrastructures. His current research interests include cloud-native telco services management and security, DevOps security, and federated learning for mobile cloud and IoT.

HAKIMA OULD-SLIMANE received the Ph.D. degree in computer science from Laval University, Québec City, QC, Canada, in 2011. She is currently a Professor with the Department of Mathematics and Computer Science, Université de Québec à Trois-Rivières, TroisRivières, QC, Canada. Her research interests include mainly information security, cyber resilience, homomorphic encryption, federated learning, preserving data privacy in smart environments, machine learning-based intrusion detection, access control, optimization of security mechanisms, and security of social networks.

AZZAM MOURAD (Senior Member, IEEE) is currently a Visiting Professor with Khalifa University, a Professor of Computer Science and the Founding Director of the Artificial Intelligence and Cyber Systems Research Center, Lebanese American University, and an Affiliate Professor with the Software Engineering and IT Department, Ecole de Technologie Superieure, Montreal, Canada. He was a Visiting Professor with New York University, Abu Dhabi. His research interests include cyber security, federated machine learning, network and service optimization and management targeting IoT and IoV, cloud/fog/edge computing, and vehicular and mobile networks. He has served/serves as an Associate Editor for IEEE TRANSACTIONS ON SERVICES COMPUTING, IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, IEEE NETWORK, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, *IET Quantum Communication*, and IEEE COMMUNICATIONS LETTERS, the General Chair of IWCMC2020-2022, the General Co-Chair of WiMob2016, and the Track Chair, a TPC member, and a reviewer for several prestigious journals and conferences.