Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 6296841, 9 pages
https://doi.org/10.1155/2022/6296841

Research Article

Minimum Latency-Secure Key Transmission for Cloud-Based Internet of Vehicles Using Reinforcement Learning

V. Akilandeswari,[1] **Ankit Kumar**,[2] **S. Thilagamani**,[3] **V. Subedha**,[4] **V. Kalpana**,[5] **Kiranjeet Kaur**,[6] **and Evans Asenso**[7]

[1]*Department of Information Technology, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India*
[2]*Department of Computer Engineering and Applications, GLA University, Mathura, Uttar Pradesh, India*
[3]*Department of Computer Science and Engineering, M. Kumarasamy College of Engineering, Thalavapalayam, Karur, Tamilnadu, India*
[4]*Department of CSE, Panimalar Institute of Technology, Pidarithangal, Tamil Nadu, India*
[5]*Department of Computer Science and Engineering, K. Ramakrishnan College of Technology, Samayapuram, Trichy, Tamilnadu, India*
[6]*Department of CSE, University Centre for Research & Development, Chandigarh University, Mohali, Punjab 140413, India*
[7]*Department of Agricultural Engineering, School of Engineering Sciences, University of Ghana, Accra, Ghana*

Correspondence should be addressed to Evans Asenso; easenso@ug.edu.gh

Received 13 July 2022; Revised 22 August 2022; Accepted 1 September 2022; Published 26 September 2022

Academic Editor: Vijay Kumar

The Internet of Vehicles (IoV) communication key management level controls the confidentiality and security of its data, which may withstand user identity-based attacks such as electronic spoofing. The IoV group's key is updated with a defined frequency under the current key management method, which lengthens the time between crucial changes and encryption. The cluster key distribution management is used as the study object in this paper, which is based on the communication security on the Internet of Vehicles cluster. When vehicles enter and exit the cluster, the Internet of Vehicles must update the group key in real-time to ensure its forward and backward security. A low-latency IoV group key distribution management technology based on reinforcement learning is proposed to optimize the group owner vehicle according to factors such as changes in the number of surrounding vehicles and essential update records and the update frequency and the key length of its group key. The technology does not require the group leader vehicle to predict the nearby traffic flow model. The access-driven cache attack model reduces the delay of encryption and decryption and is verified in the simulation of the IoV based on advanced encryption standards. The simulation results show that, compared with the benchmark group key management scheme, this technology reduces the transmission delay of key updates, the calculation delay of encryption and decryption of the IoV, and improves the group key confidentiality.

1. Introduction

The shared group key in the vehicle network (VANET, vehicular ad hoc networks) cluster performs symmetric encryption on the communication of group members [1], which is the key to ensuring the communication security and user privacy in the process of cooperative driving of cluster vehicles, congestion avoidance, and entertainment services [2, 3]. Due to the high node mobility, VANET is a specific type of wireless multicast network that must accommodate quick topology changes [4]. Intervehicle connectivity is developing into a potential area for investigation, standardisation, and growth as more cars are outfitted with wireless communication and computing technology. A subclass of mobile ad hoc networks (MANETs) and vehicular ad hoc networks (VANET) are networks that are created by moving vehicles [5]. It can be put into systems that are connected to health, where it can save a lot of lives

every day, and nonsecurity applications for commercial purposes. However, due to the dynamic and open nature of the Internet of Vehicles (IoV) cluster, the cluster is vulnerable to various internal and external network attacks. IoV systems are subject to a variety of assaults, including those on authenticating and identity, availability, secrecy, routing, and data validity, which leads to a number of difficult security and privacy requirements. In recent years, various security researchers have worked hard to secure the privacy and security of the Internet of Vehicles. Typical assaults on accessibility include denial-of-service and channel interference. This kind of attack primarily makes use of the bandwidth and transmission power restrictions to bring down the IoV system [6, 7].

Attackers use timing attacks, cache attacks [8], and other means to steal the group key, conduct eavesdropping, electronic spoofing, and other attacks, resulting in user privacy: spills and traffic accidents. Therefore, the Internet of Vehicles needs to update the group key in real-time to ensure the forward and backward security of the group key when vehicles join and leave the cluster [9].

While the IoV cluster uses the group key to perform node authentication and communication encryption for vehicles, which has attracted extensive attention, how to effectively manage the IoV key is an urgent problem that needs to be solved. Reference [10] proposes an anonymous key management scheme for the Internet of Vehicles based on critical public infrastructure (PKI, public key infrastructure), which authenticates vehicle nodes and protects data integrity through anonymous digital certificates. However, keys and anonymous certificates make the certification centre pay a high management cost to ensure system security and impose a burden on vehicle node storage and communication overhead. Reference [11] proposes using cluster-based peer-to-peer batch key agreement, but vehicle nodes can only belong to one cluster. When a new node joins the cluster, the cluster leader needs to renegotiate keys with cluster members, which is challenging to apply to high dynamic Sexual Internet of Vehicles. Researchers have developed a STRIDE model for data security that groups security threats into sexual categories including spoof, tampering, repudiation, and information disclosure. The IoV system is susceptible to the threats outlined above and can be targeted in a variety of ways, including network jamming, eavesdropping, and infiltration. The IoV system could be harmed by these attacks, which could weaken its resistance and reliability or, in the worst case, bring it to a standstill and result in accidents. The GKA scheme proposed by Reference [12] uses pseudonyms for cluster authentication and key update. It uses aliases to construct digital signatures while simultaneously authenticating cluster vehicles to prevent the leakage of vehicle trajectory privacy. Reference [9] proposes to use a software-defined network to manage the Internet of Vehicles. Different clusters independently build a decentralized network for communication and use a significant network to authenticate and collect the keys of the groups. Reference [13] proposed to perform distributed batch anonymous authentication on vehicles by computing the hash message authentication code instead of the authentication revocation list for the clustered decentralized network. This scheme protects user privacy and reduces the authentication delay; Feriani and Hossain [14]. It is proposed that based on the trust mechanism, after the vehicle with a high reputation is certified by the trust centre, anonymous key negotiation can be carried out to ensure the confidentiality and integrity of the message while protecting privacy.

As the leading research direction of machine learning, reinforcement learning is widely used in vehicle networking communication security [15, 16]. Security is a key challenge because the Internet of Things has such a significant impact on its consumers' lives. By enabling hackers to gain direct control of automobiles, a failure in the IoV system could endanger safety and could result in road accidents. In an IoT setup, like a smart grid, all sensors, gadgets, and machines can effectively and safely control energy use while humans act as observers. While inside the car, automotive monitors like brake sensors, limited fuel, and tyres pressure sensors, among others, are fitted, and the exterior of the car has security mechanisms like CCTV and parking sensors. Reference [17, 18] proposed using reinforcement learning to assist mobile devices in optimizing relay scheme, transmit power, and communication frequency to resist hostile interference. This scheme can reduce energy consumption and ensure communication reliability. In the case of the unknown channel model and spoofing attack model, Reference [19, 20] uses reinforcement learning to optimize the authentication threshold of wireless channel physical layer characteristics such as received signal strength and achieve high-precision authentication in the dynamic game of malicious senders. Reference [21] used reinforcement learning-based on the long-term memory effect for joint user access control and battery energy prediction, improving network throughput while saving energy. Reference [22] proposes using reinforcement learning to perform joint buffer allocation and interference alignment according to the current buffer usage and channel state to solve multi-user interference in wireless communication and improve the transmission rate.

This paper is based on the communication security on the Internet of Vehicles cluster and takes the cluster key distribution management as the research object. In this network, vehicles with common interests in the same area (such as destinations, hobbies, and friend groups) build clusters, cooperate to form driving patterns based on vehicle queues, share road conditions and driving information, improve with a small and constant distance between vehicles road capacity and energy efficiency, and provide multimedia services. Since vehicles frequently join and leave the cluster, to ensure the safety of cars in the group and prevent electronic spoofing and other attacks, some information in the cluster, such as road conditions and multimedia needs to be updated in real-time. Compared with traditional PKI, this paper proposes a certificateless signature scheme, which abandons digital certificates and signs the group key through a random prime number secretly stored between the trust centre and the vehicle to ensure data integrity. The unit (RSU, roadside unit) broadcasts the group key update

information, realizes the synchronous update of the group key, reduces the communication overhead of the vehicle node updating the group key, and reduces the transmission delay of the group key update. Currently, services for the Internet of Vehicles are mostly provided through cloud computing technology. A computer model known as "cloud computing" is one in which customers have 24/7 access to the Internet and computing capabilities are offered on a pay-as-you-go basis. Because of the IoV's limitations in terms of bandwidth, transmission power, and mobility, IoV routing algorithms are typically somewhat complicated. This complexity causes holes and weaknesses in the IoV routing method as a result. When the group leader's update strategy has no previous effect, the update strategy of the group key is only related to the current state of the cluster, so the optimization process of the critical update strategy is constructed as a Markov decision process. This paper proposes a low-latency group key distribution management technology based on reinforcement learning to obtain the optimal required update frequency and critical length in a dynamic cluster environment. Under the premise, the technology uses reinforcement learning to optimize the update frequency and crucial length of the group key according to the number of cluster vehicles, the change in the number of cluster vehicles, the binding update decision in the previous period, and the security level of the current cluster communication. In the learning of swarm communication, the optimal key update strategy is obtained, the computational delay of encryption and decryption of cluster communication and the probability of the group key being stolen are reduced, and the quality of cluster communication service is improved.

2. System Model

2.1. Key Management Model. Key production, distribution, usage storing, rotation, backup or recovery, and revocation destruction are all steps in the lifecycle of activities that make up key management model. When used in combination with 5G and beyond, AI technologies like machine learning (ML) and reinforcement learning (RL) offer special alternatives to dependability issues in low-latency vehicle communication networks. Numerous IoV applications demand adaptable and clever resources systems. Applications for information services, obstacle detection, and accident prevention, for instance, the need for information exchange through regular access to services with a high data transmission rate [23, 24]. The low-latency key management technology of the Internet of Vehicles is mainly composed of an On-Board unit (OBU), a roadside unit, and a Trusted Authority (TA) assembled in the vehicle. The OBU communicates with the RSU and surrounding vehicles using the dedicated short-range communication (DSRC) protocol [25]. RSU is really impenetrable and secure. RSU is unable to sign delivered messages simultaneously in the names of OBU and LCA. It is assumed in this analysis that RSUs have no computational, energy, or capacity constraints. RSUs can therefore serve as middlemen for the transmission of communications from OBUs. Because OBUs and RSUs can communicate wirelessly, each vehicle must have a tamper-proof device installed. Security, encryption, and decryption operations can be carried out by OBU. Moving vehicles' On-Board units (OBUs) communicate information such as location, the current situation, traffic, speed, direction, and accident occurrences, continually and occasionally. The aforementioned data will be processed, gathered, and provided to the cloud platform and drivers using RSUs. As a communication relay and broadcast node, RSU connects with TA through a wired network and communicates with OBU through a wireless network. When all vehicles enter the TA registration for the first time, the TA distributes the vehicle identification ID, generates a random prime number n for developing the group key and encrypted group key, generates the public key VPK and the private key VSK, and registers {ID, VPK, n}, and store {ID, VSK, n} in the OBU of the vehicle. In addition, TA registers RSU's identity RID and its public key RPK and stores its private key RSK in RSU. Table 1 shows the important parameter symbols.

As shown in Figure 1, the Internet of Vehicles organizes M vehicle nodes $\{\mathrm{Veh}_i\}_{1\le i\le M}$ to build a cluster in a particular area based on a dynamic clustering algorithm [26] and provides them with services such as cooperative driving and congestion avoidance. The moving vehicles joining and leaving the cluster have certain randomness, so the group key needs to be updated in real-time. It is assumed that the number of vehicle nodes entering and leaving the set in unit time obeys a Poisson distribution with mean λ, and the cluster vehicle identity is $\{\mathrm{ID}_i\}_{1\le i\le M}$. To ensure the security of communication and privacy within the cluster, a symmetric encryption mechanism is used for intracluster communication. In addition, a certificateless signature key management scheme is used to update the group key and cluster head (cluster head) Veh_F for the cluster. The tremendous growth in the smart automotive sectors has recently led to a huge rise in demand in Internet of Vehicles (IoV) technology. Vehicles can interact with their surroundings and public networks thanks to Internet of Vehicles technologies. Additionally, it enables moving objects to communicate and gather data on other moving objects and highways. The group key needs to be updated in real-time since the moving vehicles joining and departing the cluster have some degree of randomness. The cluster's group key and cluster head (cluster head) are updated using a key management mechanism. The IoV group's key is updated with a fixed frequency under the current key management method, which lengthens the time between critical updates and encryption.

The RSU is responsible for the secondary encryption of the request and forwards it to the TA, and the TA feeds back the new key according to the request. Specifically, the cluster leader Veh_F observes the number of vehicles in the cluster, the number of vehicles joining and leaving the group, the average driving speed of the collection, the critical update decision at the last D moments and the security level at the current moment, and generates a key update request. The binding update request mainly contains the cluster vehicle ID of the updated key and the length of the new key x_2, namely, $x_2 \| \{\mathrm{ID}_i\}_{1\le i\le M}$. $\mathrm{Veh_F}$ encrypts the critical update

Table 1: Important parameter symbols.

Parameter	Meaning
M^k	The number of vehicles in the cluster at time k
v^k	The average speed of cluster vehicles
g^k	Cluster critical update decisions for the first D moments
ρ^k	Cluster communication security level
x_1^k	Update key decision
x_2^k	Update key length
λ	Poisson distribution mean
VPK/VSK	Vehicle's public/private key
RPK/RSK	RSU public/private key
n	Random prime is chosen by TA
$E(\cdot)$	Encryption
$h(\cdot)$	Hash function
T/T^*	Send/receive timestamp

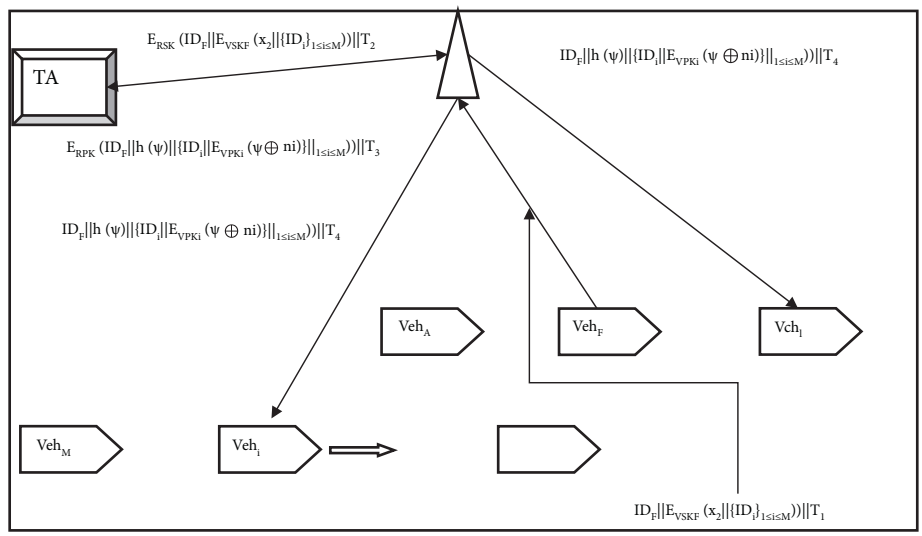

Figure 1: Key management model.

request with the private key VSK_F, adds a timestamp T_1 and sends it to the RSU. After the RSU receives the necessary update request at T_1^*, it calculates $T_1^* - T_1$, which is used to verify the validity of the binding update request and prevent replay attacks. If $T_1^* - T_1 < \Delta T_1$, the RSU uses the private key RSK to encrypt the key update request again and sends the key update request to the TA through the wired network.

TA receives the key update request forwarded by the RSU and first calculates if $T_2^* - T_2 < \Delta T_2$ to verify the validity of the key update request. The TA decrypts the key update request using the public keys RPK and VPK_F and obtains the new key length x_2 and identity $\{ID_i\}_{1\le i\le M}$, while authenticating Veh_F and RSU. According to $\{ID_i\}_{1\le i\le M}$ corresponding to $\{n_i\}_{1\le i\le M}$, TA uses the HCPA-GKA [27] algorithm based on the Chinese remainder theorem to generate a group key ψ of length x_2 and calculate the hash The value $h(\psi)$ generates a digital signature. TA uses $\{n_i\}_{1\le i\le M}$ and ψ to perform XOR operation to encrypt the group key, and uses $\{VPK_i\}_{1\le i\le M}$ to encrypt the XOR result to ensure the security of key transmission, that is, to generate $h(\psi)\left\|\left\{E_{VPK_i}(\psi\oplus n_i)\right\}_{1\le i\le M}\right.$. TA encrypts the $ID_F\|h(\psi)\|\left\{E_{VPK_i}(\psi\oplus n_i)\right\}_{1\le i\le M}$ with RPK and sends it to the RSU, which decrypts it and broadcasts it to the swarm vehicle $\{Veh_i\}_{1\le i\le M}$.

The group member Veh_i receives the critical update message by RSU and uses the private key VSK_i to decrypt the $E_{VPK_i}(\psi\oplus n_i)$ part corresponding to ID_i to obtain $\psi\oplus n_i$. Then, the group members use the random prime number n_i stored in the OBU to decrypt $\psi^* = \psi\oplus n_i\oplus n_i$ to obtain the group key ψ^* and calculate its hash value $h(\psi^*)$ to verify the authenticity of the new key ψ^*.

If $h(\psi^*) \neq h(\psi)$, the received vital update information is a forged message; otherwise, update ψ^* as a new key for intracluster communication.

2.2. Attack Model. Attack graph creation computation time begins when network complexity increases proportionally in a linear path to the number of subnets. The length of time it takes to generate an attack graph varies with network size. This paper mainly considers the internal attack on the Internet of Vehicles. The attacker can be a legitimate vehicle

node in the network, or a controlled malicious vehicle node, which obtains the group key through side-channel attacks such as timing attacks and cache attacks [28], and then implements an electronic spoofing attack [29]. The act of spoofing involves hiding a message or identification so that it looks to be coming from a reliable, approved source. Sensitive personal or business data may be stolen, credentials may be gathered for use in fraud or future attacks, malware may be transmitted via harmful links or attachments, trust relationships may be used to gain unwanted network access, and access limits may be disregarded. They may even conduct a man-in-the-middle (MITM) assault or a denial-of-service (DoS) attack. The effect of IoV interactions among RSUs, cars, and other TPMs is implied by the routing algorithm and its quality, and IoV routing methods are usually rather complex due to the IoV's limits of bandwidth, transmission power, and mobility. As a result, this complication creates gaps and vulnerabilities in the IoV routing mechanism. When an attacker pretends to be a legitimate equipment or user to steal information, transmit malware, or get around security systems, this is called spoofing. Man-in-the-middle (MITM), in this instance, the attacker places himself in the middle of two clients that are speaking and spoofs both of their addresses. By doing this, each victim transmits a data packet to the hacker rather than directly to its intended recipient. A Denial-of-Service (DoS) attack aims to bring down a computer system or network so that its targeted recipient is unable to access it. DoS attacks achieve this by providing the victim with an excessive amount of traffic or content that causes a crash. More specifically, attackers use spyware to observe the delay side-channel information during the encryption process, monitor cache hits and misses, and steal encryption keys. The attacker pretends to be an authenticated vehicle or RSU and sends malicious or forged information, such as traffic density and collision avoidance information, to surrounding vehicles through legal identities. Even cars pretending to be legal entities send this information to surrounding vehicles.

Assuming that the probability of a successful internal attack by each vehicle node is y, the chance of being attacked by a cluster with M vehicles is $1-(1-y)^M$. Among them, the probability of a successful cluster attack is positively correlated with the number of cluster vehicles and critical update frequency and negatively correlated with the key length.

3. Key Management Technology Based on Reinforcement Learning

Different models, classifications, and training approaches for machine learning are frequently utilized for prediction issues and intelligent management. Reinforcement learning (RL) will give behavior guidance in IoV applications to increase resilience and scalability. In IoV networks, it can provide routes from source to destination or route optimization. Wireless local area networks (WLANs) will gain new capabilities to meet user needs from flexible network structures, such as the software-defined networking (SDN) model while attaining higher levels of effectiveness and mobility in those complicated circumstances. By identifying workable configurations through learning, machine learning (ML) approaches used in conjunction with SDN may enhance network resource consumption and management. Throughput maximisation and latency reduction may be guaranteed with the use of machine learning (ML) and the Software-Defined Network (SDN) in the Internet of Things (IoV). By offering more dependable and enhanced routing services, ML and SDN will improve IoV network performance. A distribution cache replacement approach based on the popularity of the material will be possible with a reinforcement learning scheme like the Q-learning technique. It also has the ability to predict the unknowable popularity of cached contents. With the joining and leaving of vehicles, the merging and splitting of the cluster, and the change of communication security level, the group leader needs to update the group key of the intracluster communication in real-time [30, 31]. The dynamic optimization process of the required length and update frequency of the group key, the security level of the group communication, and the calculation delay of communication encryption and decryption can be regarded as a Markov process. This paper proposes a low-latency IoV group key distribution management technology based on reinforcement learning. The group leader optimizes the length and update frequency of the group key and obtains the optimal key update strategy through continuous trial and error and knowledge.

In this technique, the cluster performs intrusion detection based on physical layer characteristics, such as channel status and received signal strength [32, 33], and Veh_F evaluates the communication security level of the cluster-based on the feedback detection results ρ. When the cluster detects that the vehicle is under attack, such as spoofing, $\rho=0$; otherwise, $\rho=1$. The cluster leader Veh_F comprehensively considers information such as the number of vehicles in the cluster M, the cluster average speed v, the critical update decision in the first D moments and the security level ρ of the cluster communication at the last moment, and selects the critical update according to the long-term learning benefit of the cluster [34, 35]. The policy $x \in A$ includes whether to send a key update request $x_1 \in \{0, 1\}$ and the updated vital length $x_2 \in [1, L_A]$ bit, that is, $x=[x_1, x_2]$. Among them, L_A is the upper limit value of the critical length, and A is the essential management action space. When $x_1=0$, Veh_F does not send a key update request, and the cluster continues to use the previous key for communication encryption; when $x_1=1$, Veh_F requests the TA to update the group key of length x_2.

At time k, Veh_F observes the number of vehicles M^{k-1} and M^k at time $k-1$ and time k in the cluster, the average speed of the cluster vehicles , the critical update decision $g^k=[x_1^i]_{k-D\le i\le k-1}$ and the security level ρ^{k-1} of the cluster communication at time $k-1$, and construct the current cluster state s^k.

$$s^k=\left[M^{k-1}, M^k, v^k, g^k, \rho^{k-1}\right]. \tag{1}$$

Let $Q(s\ x)$ denote the long-term payoff of Veh_F taking action x in state s; $\mathrm{Veh_F}$ selects a more efficient group key by using an ε-greedy strategy based on the current cluster state

s^k through the corresponding Q value. New strategy $x^k = [x_1^k, x_2^k]$. Specifically, Veh_F selects the key update strategy with the largest long-term benefit $Q(s^k, x)$ in the current state s^k with the probability of $1 - \varepsilon$; arbitrarily selects a key update strategy from A with the probability of ε, namely,

$$p\left(x^k = x^*\right) = \begin{cases} 1 - \epsilon, & x^* = \underset{x \in A}{\operatorname{argmax}} Q\left(S^k, x\right), \\ \dfrac{\epsilon}{|A|}, & \forall x^* \in A. \end{cases} \tag{2}$$

At time k, if Veh_F sends a key update request forwarded by RSU to TA, the request contains cluster vehicle identity $\{ID_i\}_{1 \le i \le M}$ and key length x_2^k. TA receives the critical update request and generates a group key ψ is returned to the RSU; after the RSU receives the encrypted new key, it broadcasts it to the cluster vehicles to update the entire cluster key.

When the cluster performs a critical update, Veh_F obtains the current security level ρ^k of intra

cluster communication through the feedback of cluster members and evaluates the calculation of cluster encryption and decryption of the group key of length x_2^k. Veh_F evaluates the benefit u^k of this critical update by the number of cluster vehicles M^k, the cluster communication security level ρ^k, and the essential length x_2^k.

$$u^k = \rho^k M^k - c_1 x_2^k M^k - c_2 x_1^k M^k. \tag{3}$$

Among them, c_1 and c_2 represent the weights of the computation delay and critical update delay of cluster communication encryption and decryption, respectively.

Veh_F observes the state s^{k+1} at the time of cluster $k + 1$ and updates the Q value corresponding to the state-action pair (s^k, x^k) based on the Bellman equation.

$$Q\left(s^k, x^k\right) \leftarrow (1 - \alpha) Q\left(s^k, x^k\right) + \beta\left(u^k + \max_{x \in A} Q\left(s^{k+1}, x^k\right)\right). \tag{4}$$

The specific process of the IoV key management technology based on reinforcement learning is shown in Algorithm 1.

4. Simulation Experiments

3D technology is now a prominent technology that is desired and used in a variety of sectors. The current standard traffic simulation software has a scene interface with an excellent 3D effect, which allows users to generate simulation scenarios more logically and stereoscopically. Trans-Modeler, a simulation programme that enables 3D modelling, offers an amazing 3D impact when 3D models must be addressed during modelling. This paper uses Python to build a dynamic cluster simulation scenario to verify the advantages of the low-latency IoV group key distribution management technology based on reinforcement learning in the computing delay and communication security level of encryption and decryption. Reference [29] for the setting of stimulation parameters are as follows: the group leader and vehicles within a range of 300 m form a cluster, the number of vehicle nodes in the initial cluster is 50, and the average speed of vehicles in the cluster $v \in [0, 120]$ km/h, and the same as that of the cluster vehicles. The number is inversely proportional [30]. Cluster vehicles use the Advanced Encryption Standard (AES) symmetric encryption algorithm to encrypt intracluster communication, with optional key length $x_2 \in \{128,192,256\}$ bits. The packet size for intracluster vehicle communication is 1 kB. The learning rate of Q-learning is $\alpha = 0.2$, and the discount factor $\beta = 0.8$.

Figure 2 and Table 2 provide the average performance of reinforcement learning-based key management (RLKA) over 3000-time slots, examining the average number of vehicles joining and leaving the cluster per time slot $\lambda \in \{1, 2, \ldots, 5\}$. Figure 3 and Table 3 depict the effect of encryption and decryption calculation delay and cluster communication security level. It is clear from Figures 2 and 3 that when λ increases from 1 to 5, the calculation delay of vehicle encryption and decryption of 1 kB data packet in the cluster increases from 432 μs increased to 497 μs, the latency increased by 15.0%, and the cluster communication security level decreased from 0.963 to 0.794, and the performance decreased by 17.5%. When $\lambda = 3$, compared with the GKA scheme, the proposed technique's encryption and decryption calculation delay are reduced by 18.1%, and the security level is improved by 24.1%. This is because, compared with the fixed vital length and update frequency of the GKA scheme, the RLKA technology proposed in this paper can be based on the number of vehicles in the current cluster, the number of vehicles joining and leaving the group, and the safety feedback of cluster vehicle intrusion detection [31]. Prescan simulation software uses a basic vehicle dynamics model, which is unable to drive the intelligent vehicle precisely in the vertical direction or to effectively reflect the vehicle's dynamic properties in that direction. CarSim simulation tool is advised when the necessary study incorporates vehicle dynamics. In many ways, the domestic traffic flow operation law differs from that of other nations, particularly when the signal light is yellow because of differences in driver behavior and vehicle operation. When a driver sees a yellow light in the real world, he is more likely to speed through the junction than to stop and slow down. The software programmes, like Vissim and Prescan, do take this into account but do not set a variable driving reaction model, and the simulation result is quite different from the real scenario [32]. Therefore, reasonably update of the key and control of the length of the key, reduce the probability of successful attack by malicious nodes, ensure the security of IoV cluster communication, and obtain minor encryption and decryption calculation delay overhead. Since the IoV cluster requires the cluster key to conduct node identification and connection cryptography for automobiles, which has garnered considerable attention, it is urgently necessary to find a solution to the issue of how to maintain the IoV key. The routing algorithm and its quality indicate the impact of IoV interactions between RSUs, automobiles, and other TPMs, and IoV routing techniques are typically somewhat complicated because to the IoV's bandwidth limitations. The

(1) Initialize α, β, $Q=0$, s^1, $k=1$, and K.
(2) Observe the number of vehicles M^{k-1} and M^k at time $k-1$ and time k, the vehicle speed v^k, the critical update situation g^k in the first D time and the communication security level ρ^{k-1} at time $k-1$.
(3) Construct the state $s^k = [M^{k-1}, M^k, v^k, g^k, \rho^{k-1}]$ of the cluster at the current k time.
(4) Select the key update strategy $x^k = [x_1^k, x_2^k]$ according to formula (2).
(5) When $x_1^k = 1$, do the following:
    (a) Request TA to generate a group key ψ of length x_2^k;
    (b) RSU broadcasts the critical update message to cluster vehicle $\{\text{Veh}_i\}_{1 \le i \le M}$.
(6) Evaluate the current communication security level ρ_k.
(7) Obtain benefit u_k according to equation (3).
(8) Update $Q(s^k, x^k)$ according to equation (4).
(9) If $k=K$, end the study; otherwise, $k=k+1$, go to step 2.

Algorithm 1: Low-latency IoV key distribution management algorithm based on reinforcement learning.

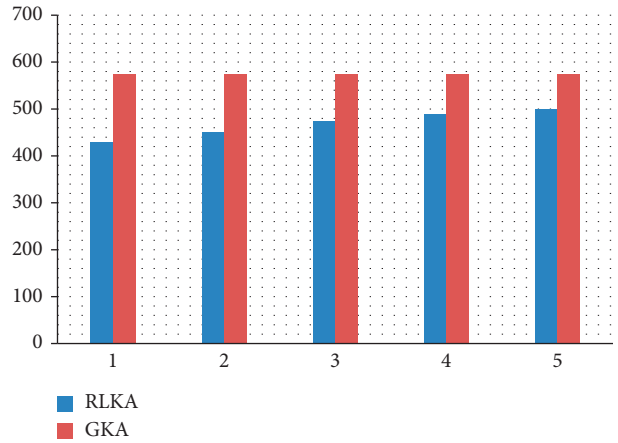

Figure 2: Calculation delay of vehicle encryption and decryption in the cluster w.r.t. time slot.

Table 2: Calculation delay of vehicle encryption and decryption.

Algorithm	Time slot 1	Time slot 2	Time slot 3	Time slot 4	Time slot 5
RLKA	430	450	475	490	500
GKA	575	575	575	575	575

foundational framework for creating a smart IoV environment is provided by the 5G network. In order to reach high performance, it pushes the vehicle network capabilities. The Internet of Everything (IoE) is capable of a new evolution thanks to 5G technology [36].

To verify the advantage of RLKA technology using certificates signature in terms of communication overhead, this paper uses MATLAB to build a simulation of the transmission delay of essential updates. Reference [13] for simulation parameter setting, the wireless transmission rate between vehicle node and RSU is 600 Mbit/s, and the data size of vehicle ID, timestamp T, and request key length x_2 is 4 bytes. As shown in Figure 4 and Table 4, when the number of cluster vehicles increases from 20 to 100, due to the increase in the data volume of critical update information, the transmission delay of cluster key update information rises

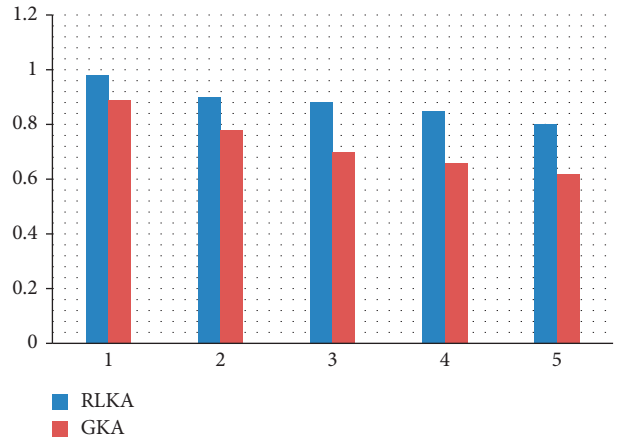

Figure 3: Cluster communication security level w.r.t. time slot.

Table 3: Cluster communication security level.

Algorithm	Time slot 1	Time slot 2	Time slot 3	Time slot 4	Time slot 5
RLKA	0.98	0.9	0.88	0.85	0.8
GKA	0.89	0.78	0.7	0.66	0.62

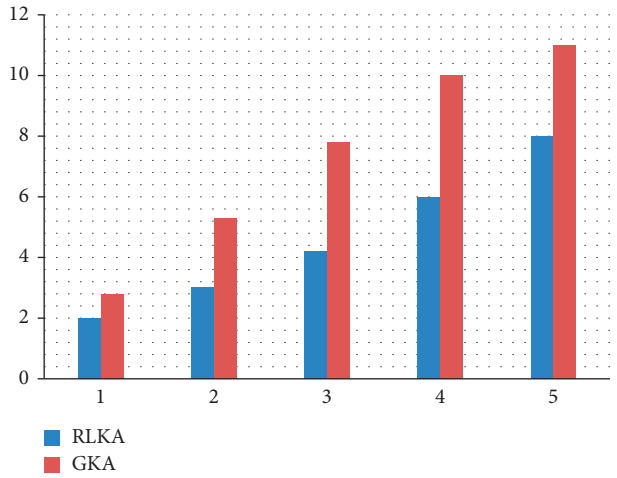

Figure 4: Key update transmission delay.

Table 4: Transmission delay with increasing the number of vehicles in a cluster denoted by N.

Algorithm	$N = 20$	$N = 40$	$N = 60$	$N = 80$	$N = 100$
RLKA	2	3	4.2	6	8
GKA	2.8	5.3	7.8	10	11

linearly with the number of cluster vehicles. When the number of cluster vehicles is 60, compared with the GKA [16] scheme, the critical update transmission delay of the proposed technique is reduced by 31.0%. Compared with the PKI-based key management scheme of the Internet of Vehicles, this paper uses the certificateless signature key management technology, abandons the digital certificate, avoids the communication delay and communication load caused by the transmission of the digital certificate, and realizes the group key through RSU broadcast.

5. Conclusion

This paper proposes a low-latency IoV group key distribution management technology based on reinforcement learning. The group leader vehicle observes cluster characteristics such as the number of surrounding vehicles and vehicle speed. It combines the critical update decision at the last moment with the current IoV communication security. Construct the cluster state according to the situation, use reinforcement learning to optimize the update frequency and critical length of the group key, realize the resistance to attacks such as crucial theft and spoofing through low-latency cluster key update, improve the security level of cluster communication, and reduce the number of intracluster attacks—communication delay. Researchers were having trouble figuring out how to apply deep learning methodologies to IoV applications due to issues with big data simulations, control computation, and resource management. These issues call for significant improvements in order to create a coherent system that satisfies both service excellence and customer quality requirements. From this research, the simulation results demonstrate that, when compared to the GKA scheme, the proposed RLKA technology can guarantee communication security while also lowering the transmission delay of group key updates, lowering the computational uncertainty of communication encryption and decryption within the cluster, enhancing the confidentiality of the group key, and enhancing the quality of cluster communication services.

Data Availability

The data shall be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Z. Wu and D. Yan, "Deep reinforcement learning-based computation offloading for 5G vehicle-aware multi-access edge computing network," *China Communications*, vol. 18, no. 11, pp. 26–41, 2021.

[2] D. Kwon and J. Kim, "Optimal trajectory learning for UAV-BS video provisioning system: a deep reinforcement learning approach," in *Proceedings of the 2019 International Conference on Information Networking (ICOIN)*, Berlin Heidelberg, June 2019.

[3] H. Yang, X. Xie, and M. Kadoch, "Intelligent resource management based on reinforcement learning for ultra-reliable and low-latency IoV communication networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4157–4169, 2019.

[4] H. Badis and A. Rachedi, "Modeling tools to evaluate the performance of wireless multi-hop networks," in *Modeling and Simulation of Computer Networks and Systems*, pp. 653–682, Morgan Kaufmann, Burlington, Massachusetts, USA, 2015.

[5] M. Rath, B. Pati, and B. K. Pattanayak, "An overview on social networking: design, issues, emerging trends, and security," *Social Network Analytics: Computational Research Methods and Techniques*, vol. 21, 2018.

[6] K. Mershad and H. Artail, "A framework for secure and efficientdata acquisition in vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 2, pp. 536–551, 2013.

[7] N. Shah and S. Valiveti, "Intrusion detection systems for theavailability attacks in ad-hoc networks," *Int J Electron ComputSci Eng (IJECSE, ISSN*, vol. 1, no. 3, pp. 1850–1857, 2012.

[8] S. Deshmukh, K. Thirupathi Rao, and M. Shabaz, "Collaborative learning based straggler prevention in large-scale distributed computing framework," *Security and Communication Networks*, vol. 2021, pp. 1–9, Article ID 8340925, 2021.

[9] X. Hu, S. Xu, L. Wang et al., "A joint power and bandwidth allocation method based on deep reinforcement learning for V2V communications in 5G," *China Communications*, vol. 18, no. 7, pp. 25–35, 2021.

[10] C. Pan, Z. Wang, Z. Zhou, and X. Ren, "Deep reinforcement learning-based URLLC-aware task offloading in collaborative vehicular networks," *China Communications*, vol. 18, no. 7, pp. 134–146, 2021.

[11] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: a tutorial," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3039–3071, 2019.

[12] C. Sharma, A. Bagga, B. K. Singh, and M. Shabaz, "A novel optimized graph-based transform watermarking technique to address security issues in real-time application," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–27, Article ID 5580098, 2021.

[13] H. Ding and K.-C. Leung, "Resource allocation for low-latency NOMA-V2X networks using reinforcement learning," in *Proceedings of the IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Vancouver, BC, Canada, May 2021.

[14] A. Feriani and E. Hossain, "Single and multi-agent deep reinforcement learning for AI-enabled wireless networks: a tutorial," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1226–1252, 2021.

[15] T. Li, W. Liu, Z. Zeng, and N. N. Xiong, "DRLR: a deep reinforcement learning based recruitment scheme for massive data collections in 6G-based IoT networks," *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 14595–14609, 2022.

[16] C. Sharma, B. Amandeep, R. Sobti, T. K. Lohani, and M. Shabaz, "A secured frame selection based video watermarking technique to address quality loss of data: combining graph based transform, singular valued decomposition, and hyperchaotic encryption," *Security and Communication Networks*, vol. 2021, pp. 1–19, Article ID 5536170, 2021.

[17] X. Wang, Y. Wang, Q. Cui, K.-C. Chen, and W. Ni, "Machine learning enables radio resource allocation in the downlink of ultra-low latency vehicular networks," *IEEE Access*, vol. 10, pp. 44710–44723, 2022.

[18] J. Tian, Q. Liu, H. Zhang, and D. Wu, "Multiagent deep-reinforcement-learning-based resource allocation for heterogeneous QoS guarantees for vehicular networks," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 1683–1695, 2022.

[19] F. Naeem, S. Seifollahi, Z. Zhou, and M. Tariq, "A generative adversarial network enabled deep distributional reinforcement learning for transmission scheduling in Internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4550–4559, 2021.

[20] J. Chen, L. Chen, and M. Shabaz, "Image fusion algorithm at pixel level based on edge detection," *Journal of Healthcare Engineering*, vol. 2021, pp. 1–10, Article ID 5760660, 2021.

[21] G. Sun, G. O. Boateng, D. Ayepah-Mensah, G. Liu, and J. Wei, "Autonomous resource slicing for virtualized vehicular networks with D2D communications based on deep reinforcement learning," *IEEE Systems Journal*, vol. 14, no. 4, pp. 4694–4705, 2020.

[22] M. F. Khan and K.-L. A. Yau, "Route selection in 5G-based flying ad-hoc networks using reinforcement learning," in *Proceedings of the 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, Penang, Malaysia, August 2020.

[23] J. Li, J. Tang, J. Li, and F. Zou, "Deep reinforcement learning for intelligentcomputing and content edge service in ICN-based IoV," in *Proceedings of the 2021 IEEEInternational Conference on Communications Workshops (ICC WorkShops)*, pp. 1–7, Montreal, Quebec, June 2021.

[24] W. Ejaz, M. Naeem, S. K. Sharma et al., "IoV-based deployment and scheduling of charging infras-tructure in intelligent transportation systems," *IEEE Sensors Journal*, vol. 21, no. 14, pp. 15504–15514, 2021.

[25] W. J. Yun and J. Kim, "Attention-based reinforcement learning for real-time UAV semantic communication," in *Proceedings of the 2021 17th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 1–6, Piscataway, New Jersey USA, Semtember 2021.

[26] S. Sharma and B. Singh, "Cooperative reinforcement learning based adaptive resource allocation in V2V communication," in *Proceedings of the 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 489–494, Noida, India, March 2019.

[27] Y. Du, P. Gao, J. Yang, F. Shi, and M. Shabaz, "Experimental analysis of mechanical properties and durability of cement-based composite with carbon nanotube," *Advances in Materials Science and Engineering*, vol. 2021, pp. 1–12, Article ID 8777613, 2021.

[28] S. Y. Lien, S. C. Hung, D. J. Deng, C. L. Lai, and H. L. Tsai, "Low latency radio access in 3GPP local area data networks for V2X: stochastic optimization and learning," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4867–4879, 2019.

[29] X. Du, H. Van Nguyen, C. Jiang, Y. Li, F. R. Yu, and Z. Han, "Virtual relay selection in lte-V: a deep reinforcement learning approach to heterogeneous data," *IEEE Access*, vol. 8, pp. 102477–102492, 2020.

[30] Z. Ding and J. Xiang, "Overview of intelligent vehicle infrastructure cooperative simulation technology for IoV and automatic driving," *World Electric Vehicle Journal*, vol. 12, no. 4, p. 222, 2021.

[31] C. Pan, Z. Wang, H. Liao et al., "Asynchronous federated deep reinforcement learning-based URLLC-aware computation offloading in space-assisted vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2022.

[32] Y. Lu, S. Maharjan, and Y. Zhang, "Adaptive edge association for wireless digital twin networks in 6G," *IEEE Internet of Things Journal*, vol. 8, no. 22, pp. 16219–16230, 2021.

[33] A. Nassar and Y. Yilmaz, "Deep reinforcement learning for adaptive network slicing in 5G for intelligent vehicular systems and smart cities," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 222–235, 2022.

[34] F. Yang, S. Wang, J. Li et al., "An overview of Internet of vehicles," *China Commun*, vol. 11, no. 10, pp. 1–15, 2014.

[35] X. Li, Q. Li, D. Kong, X. Zhang, and X. Wang, "Learning Based Trajectory Design for Low-Latency Communication in UAV-Enabled Smart Grid Networks," in *Proceedings of the 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*, pp. 1–5, Victoria, BC, Canada, November 2020.

[36] C. Choe, J. Ahn, J. Choi, D. Park, M. Kim, and S. Ahn, "A robust channel access using cooperative reinforcement learning for congested vehicular networks," *IEEE Access*, vol. 8, pp. 135540–135557, 2020.