

Article

# Smart Traffic Shaping Based on Distributed Reinforcement Learning for Multimedia Streaming over 5G-VANET Communication Technology

Adel A. Ahmed , Sharaf J. Malebary , Waleed Ali  and Omar M. Barukab 

Information Technology Department, Faculty of Computing and Information Technology-Rabigh, King Abdulaziz University, Jeddah 25729, Saudi Arabia

\* Correspondence: aaaabdullah1@kau.edu.sa; Tel.: +96-65-6388-4738

**Abstract:** Vehicles serve as mobile nodes in a high-mobility MANET technique known as the vehicular ad hoc network (VANET), which is used in urban and rural areas as well as on highways. The VANET, based on 5G (5G-VANET), provides advanced facilities to the driving of vehicles such as reliable communication, less end-to-end latency, a higher data rate transmission, reasonable cost, and assured quality of experience (QoE) for delivered services. However, the crucial challenge with these recent technologies is to design a real-time multimedia traffic shaping that maintains smooth connectivity under the unpredictable change of channel capacity and data rate due to handover for rapid vehicle mobility among roadside units. This research proposes a smart real-time multimedia traffic shaping to control the amount and the rate of the traffic sent to the 5G-VANET based on distributed reinforcement learning (RMDRL). The proposed mechanism selects the accurate decisions of coding parameters such as quantization parameters, group of pictures, and frame rate that are used to manipulate the required traffic shaping of the multimedia stream on the 5G-VANET. Furthermore, the impact of the aforementioned three coding parameters has been comprehensively studied using five video clips to achieve the optimal traffic rate value for real-time multimedia streaming on 5G communication. The proposed algorithm outperforms the baseline traffic shaping in terms of peak-signal-to-noise-ratio (PSNR) and end-to-end frame delay. This research will open new comfortable facilities for vehicle manufacturing to enhance the data communication system on the 5G-VANET.

**Keywords:** 5G-VANET; RMDRL; QP; GOP; FR**MSC:** 68T01; 68M10

**Citation:** Ahmed, A.A.; Malebary, S.J.; Ali, W.; Barukab, O.M. Smart Traffic Shaping Based on Distributed Reinforcement Learning for Multimedia Streaming over 5G-VANET Communication Technology. *Mathematics* **2023**, *11*, 700. <https://doi.org/10.3390/math11030700>

Academic Editor: Wujie Zhou

Received: 30 December 2022

Revised: 24 January 2023

Accepted: 28 January 2023

Published: 30 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Recently, the exponential growth of real-time applications and the massive development of vehicles have proceeded to develop an Intelligent Transport System (ITS) which has a capability to utilize machine learning algorithms to predict the suitable traffic rate for multimedia streaming over a vehicular ad-hoc network (VANET). In particular, the multimedia stream cannot be transmitted through 5G-VANET without using comparison techniques which utilize advanced video coding (e.g., H.264) to accommodate the channel capacity and bandwidth requirements [1,2]. However, in order to produce encoded video with a suitable traffic rate, the video coding mechanism needs to search for the optimal value of coding parameters such as the quantization parameters (QP), the group of pictures (GOP), and the frame rate (FR). On the other hand, a good quality of experience (QoE) for the end user can be achieved using three factors which are the initial delay, the quality switch, and the stall frame [3]. As a result, the real-time multimedia stream may not necessarily be effective for the end user's perceptual quality, since the video coding mechanism is dependent mainly on searching for the optimal coding parameters [4–6]. Therefore, machine learning (ML) techniques can be used to predict the optimal coding parameters

and produce valuable insights and guide the automated decision-making processes which might attract the consideration of the manufacturing and intellectual institutions. Consequently, new technology in the communication field is expected to utilize machine learning to develop smart vehicles that can act as intelligent vehicles which have the capability to collaborate with the nearest Roadside Unit (RSU) and the central gateway to provide very useful data communication for vehicles. Unfortunately, traditional network technologies such as 4G are not qualified to manage the constraints of Information Technology Solutions (ITS) in an efficient, accessible, smooth, and low cost manner. For instance, real-time video watching needs very fast data communication which can be provided by 5G and 4G cannot meet the requirements of latency of ITS applications and those of vast bandwidth and short delay (1ms over-the-air) which resolve the shortcomings of the 4G technology [7–9].

### 1.1. VANET Based on 5G

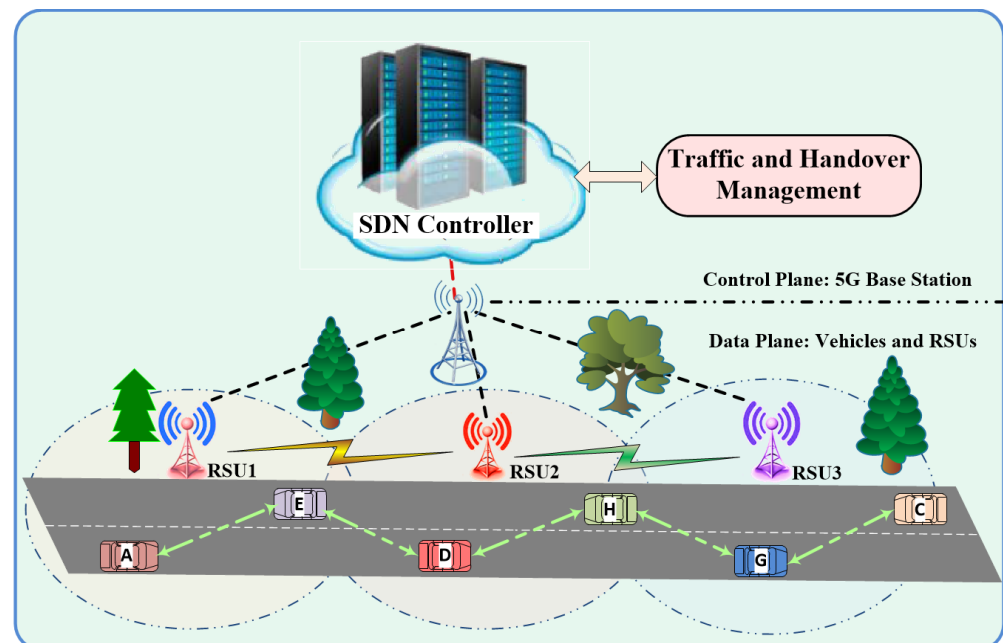
The architecture of a 5G mobile system relies on virtualized and softwarized networks such as Software-Defined Networking (SDN), Network Function Virtualization (NFV), and Multi-access Edge Computing (MEC). In order to support the growing data traffic, connected devices, heterogeneous network management, and vehicle mobility, 5G-VANET based on SDN should be deployed. To address the shortcomings of conventional networks, such as the need to reconfigure and troubleshoot connections for each vehicle in a VANET, and also the need to make efficient use of network resources and reduce path-recovery delay as a result of distributed mechanisms, the SDN paradigm offers a logical, centralized, and programmable approach of designing networks [10,11]. For 5G-VANET supervision, SDN can adopt an intelligent and harmonious strategy, which makes it easier to have logically centralized management over heterogeneous networks. As shown in Figure 1, the only goal of the 5G-VANET based on SDN is to separate the control plane, also known as the network operating system, which is hosted in the 5G base station from the data plane (forwarding plane), which is hosted in the RSU and network vehicles. The data plane actually transports packets from one location to another while the control plane decides how they should move through the network. OpenFlow, a programmable network protocol that enables the controller to create the forwarding tables and track the packet statistics of routers and switches, is the most well-known control protocol for SDN [7,12–14]. As a result, the 5G base station may connect with RSU and vehicles via the OpenFlow protocol, and it also carries out the commands given by the SDN controller. The dynamic routing policy can also be used to manage traffic. When a packet enters the network and reaches a 5G base station, the base station's proprietary firmware contains instructions that instruct it where to forward the packet.

The 5G-VANET has to face different issues, including the geographical obstacles of urban zones, the vehicle mobility patterns, the frequent topology changes, fast handover problem, and the real-time application requirements of the delivery delay and data loss. Furthermore, the safety applications in the 5G-VANET rely on real-time transmission and it uses a V2V communication system to avoid traffic accidents, which are mainly caused by human errors. The data loss ratio in the real-time routing plays a crucial role in the 5G-VANET. Moreover, it is indispensable that the handover should be executed impeccably so that the customer is not cognizant of any difference. Consequently, the handover problem should be efficiently studied while an optimal routing solution for 5G-VANET is being developed.

### 1.2. Traffic Shaping over 5G-VANET

The traffic shaping efficiency of the packet through the 5G-VANET network should be considered in each vehicle because the routing decision will not only affect the transmission of the current packet, but also subsequent future packets as the vehicles and RSU along the route will need to choose the optimal path to forward packets toward the destination. The entire traffic shaping decision in the VANET can be seen as a collection of many small decisions that in the long term affect the usability and availability efficiency of the entire

network. Moreover, it is difficult to use a simple rule to determine the packet forwarding policy because of frequent changes of channel capacity and bandwidth in VANETs [15–17]. This decision-making process matches well with a reinforcement learning (RL) model. Fortunately, these issues can be solved by using reinforcement learning, whose algorithms look for a policy that links system states to the agent actions that should perform if those states materialize. The ideal input/output pairs are never offered in reinforcement learning, and system evaluation frequently happens simultaneously with learning. The issue that arises when an agent has to learn behavior through trial-and-error interactions with a dynamic environment is known as reinforcement learning. Hence, applying suitable RL into the traffic shaping algorithm will enable the vehicles to learn from their experience and make a better decision in the long term. This utilizes intelligence into the entire routing process and the operation of the VANET network. As can be shown in Figure 1, a real-time video streaming (forwarding) on VANETs means that the messages of multimedia application are transmitted based on the limitation of end-to-end deadlines, which is useful in many applications, including vehicle traffic monitoring, traffic accidents and collisions avoiding, and advanced crash warning.



**Figure 1.** The 5G-VANET Architecture.

High data loss and erratic connection availability have an impact on real-time multimedia traffic shaping. Due to these issues, we must use a machine learning technique to predict the channel capacity and ensure strong resistance to congestion and channel error. Therefore, rigorous resource management system design, topology considerations, and real-time multimedia coding metric selection are required for the 5G-VANET real-time traffic shaping [18–23].

### 1.3. Problem Statement and Research Motivation

By averaging the available data rate, traditional traffic shaping algorithms over the 5G-VANET periodically convert burst traffic into fixed-rate traffic or convert burst traffic into controlled maximum rate (e.g., Token Bucket). However, due to radio signal interference, noise, radio channel contention, and the wireless transceiver's modulation capabilities, such traffic shaping algorithms may face a number of difficulties, including rapid topology changes, a high node mobility, and dynamic changes in the coverage area. Furthermore, vehicles in the intersection can be connected to a very weak RSU which will make unstable data communication. The local optimal problem will definitely affect a routing protocol that

performs calculations per-hop. This paper's major goal is to create a real-time multimedia traffic shaping system based on distributed reinforcement learning (RMDRL) that takes into account the aforementioned difficulties.

#### 1.4. Research Contribution

The following contributions are reported in this research work:

- This paper proposes a smart real-time multimedia traffic shaping mechanism based on distributed reinforcement learning that will be applicable for the 5G-VANET. The idea of using distributed reinforcement learning in real-time traffic shaping is a novel idea since the task of traffic shaping will be distributed among RSU, vehicles, and mobile devices.
- It comprehensively studies the impact of adapting the three parameters of video coding (QP, GOP, FR) on achieving the optimal traffic rate value for real-time multimedia streaming on the 5G-VANET which will increase the fidelity criteria and reduce the video bitrate with a high quality.
- The proposed mechanism provides a high QoE in terms of the PSNR, and it also ensures short frame latency.

The rest of this paper is arranged as follows: Section 2 presents the related works on traffic shaping. The system design of the smart traffic shaping is explained in Section 3. Section 4 describes the simulation experiments and performance evaluation. Finally, Section 5 presents the conclusion and future work.

## 2. Related Works on Traffic Shaping

Most of the research studies in the literature focus on the traffic shaping solutions based on network adaptation and channel estimation. In contrast, this research solves the traffic shaping based on the optimal adjusting of video coding parameters using the machine learning algorithm. Ahmed [24] proposed an adaptive traffic shaping of multimedia streaming for real-time routing over the next generation of Wireless Multimedia Sensor Networks. Wu et al. [25] developed a routing protocol that utilizes a reinforcement learning algorithm (QLAODV) to manage the network state information and to enhance correctness information. Furthermore, Shin et al. [26] proposed a reinforcement learning based network selection algorithm which learns the channel circumstance to reduce the latency of IoT devices for massive, connected IoT networks. Akbari and Tabatabaei [27] proposed an energy-efficient mechanism based on fuzzy logic and reinforcement learning, in which they utilized the remaining energies of the nodes on the routes, the available bandwidth, and the distance to the sink. The most interesting research was proposed by Rossi et al. [28]. In order to jointly control the data compression operations (distortion costs) at the sources and the routing (transport costs) of the compressed information in a distributed IoT, the authors of [28] proposed a distributed algorithm based on the alternating direction method of multipliers (ADMM). To estimate the essential data for routing protocols, Lai et al. [29] introduced the machine learning-assisted route selection (MARS) system. This system can predict vehicle movement, and then it can select routing paths with better transmission capacities for packet transmission. In addition, Immich et al. [30] proposed an intelligent quality-driven and network-aware method (called AntArmour) that dynamically allotted a specific level of redundancy using an ant colony optimization scheme. As a result, AntArmour is able to protect high-definition video streams that are being transmitted live in real-time. Ben Ameer et al. [31] proposed a combination of traffic shaping techniques and TCP congestion management versions to mitigate the negative effects of congestion events limiting the downsides of concurrent HAS streams in the home gateway. For multimedia traffic via SDN, Al Jameel et al. [32] suggested a reinforcement learning-based routing scheme. To improve the end-user QoE, the RL agent learns to prioritize the bandwidth and select a path with a low packet loss ratio, end-to-end delay, and jitter. Additionally, Marwah et al. [33] developed a junction-aware vehicle selection method for multi-path video streaming in a vehicular network that makes use of the SDN paradigm's capabilities

to track and collect network statistics in order to determine the best way. A deep learning (DL) technique based on the unidirectional long short-term memory (LSTM) was also proposed by Abdellah et al. [34] to anticipate traffic in V2X networks, allowing time series prediction models to predict future values as a function of past values. This makes predictions more accurate, which improves decision-making. Vergados et al. [35] performed an evaluation of the MPEG Dynamic Adaptive Streaming over HTTP (DASH) algorithm over a vehicular scenario, with respect to the effect of vehicle speed, on different QoE metrics. Moreover, Esmaily and Kravlevska [36] proposed the principal design criteria for creating and deploying experimental environments for network slicing in 5G. Furthermore, the authors of [36] described the most common small-scale state-of-the-art testbeds for network slicing with their characteristics.

The limitations of the previous literature studies in [24–36] can be specified in two facts: First, the traffic shaping that has been proposed in the related works optimized the number of packets and the frame queue at the sender of the network, which increase the loss ratio and hence decrease the fidelity of the reconstructed video at the receiver. Second, the parameters’ configuration optimization of real-time video coding was not considered in the related works, though it is more important than the channel network optimization. Therefore, the proposed mechanism utilizes machine learning, especially DRL, to optimize the real-time video coding mechanism.

### 3. System Design of Smart Traffic Shaping

The VANET is a particular case of a wireless multi-hop network, which needs serious consideration for the design of efficient real-time traffic shaping protocols for multimedia streaming. This is mainly due to the various challenges in VANETs such as fast topology changes, high node mobility, and the dynamic change of wireless coverage area due to interferences of radio signals, noise, radio channel contentions, and the modulation capabilities of the wireless transceiver. Therefore, traditional real-time multimedia traffic shaping protocols must be modified to outfit VANET’s characteristics using a novel idea of the design of real-time traffic shaping. Hence, the proposed mechanism is designed to resolve the VANET challenges using several functions including an optimal multimedia encoding model, discovering the impact of encoding parameters on traffic shaping, and to design traffic shaping based on a DRL model as illustrated in Figure 2.

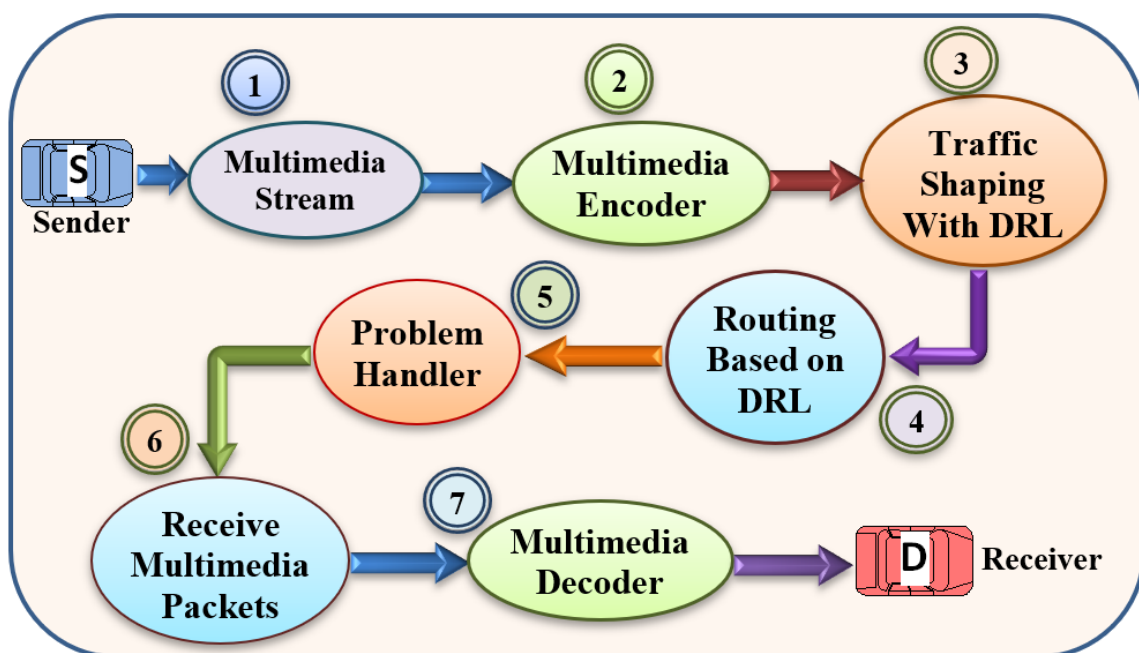


Figure 2. RMDRL System Model.

### 3.1. Multimedia Encoding Model

The first stage of the proposed mechanism is to initiate the video coding model that might use an advanced video coding (H.264) or a high-efficiency video coding (H.265). In this research, the coding H.264 has been used because it is the most preferred coding used by 83% of the industry developers of multimedia devices as it has been mentioned in Bitmovin's "Video Developer Report 2021/2022" [37]. This research offers the ideal resolution parameters and encoding configuration for the real-time multimedia coding that result in a minimal processing complexity with an allowance for energy consumption based on various experiments using the open-source Xvid software [38]. The five fundamental operational elements that make up the encoding complexity are entropy coding, in-loop filtering, transform/quantization, and inversion, motion estimation, and compensation, as well as intra prediction. The discovery of the ideal video coding parameters is shown in Table 1. The input video resolution is set to the Common Intermediate Format (CIF), which is encoded at 30 frames per second with an image size of  $352 \times 288$  and a  $16 \times 16$  macroblock. The CIF input resolution was chosen for this study because of its small size and good frame fidelity, which is advantageous for high mobility vehicle limits. Additionally, by adjusting the search range (SR) parameter, the number of reference frames (NRF), the motion estimation (ME) algorithm, the quantization parameter (QP), the group-of-pictures (GOP) size, the entropy coding based on context-based adaptive binary arithmetic coding (CABAC), the adaptive deblocking filter (Deblock), rate-distortion optimization (RDO), and Chroma ME, the proposed system facilitates the real-time multimedia streams to the available bandwidth. For H264/AVC, the ideal complexity coding values are recommended in [39]. According to the findings in [39], the GOP should be greater than eight, QP should be between twenty-six and thirty-four, and SR should be between four and eight. These values were shown to be the best for achieving the optimal encoding complexity in a multi-hop wireless medium. Furthermore, it was discovered in [39] that setting SR and NRF both to one result in the best encoding complexity. Figure 3 shows how the proposed RMDRL can manage the target traffic rate based on GOP, QP, and frame rate (FR).

**Table 1.** H.264 Encoder Configuration Options.

Complexity Parameter	H.264 Configuration Option	Proposed Optimized Value
Motion Estimation	-me (Dia, Hex, Umh, Esa)	Hex or Dia
Quantization parameter (QP)	-qp (0–51)	DRL uses four values 10, 20, 30, and 40
GOP	-keyint (1–250)	DRL uses specific 14 values 10, 20, 40, 80, 100, 120, 140, 160, 180, 200, 220, 240, and 250.
Coding-based Adaptive Binary Arithmetic Coding (CABAC)	-no-cabac	Disable
Deblock Filter	-no-deblock; -nf	Disable/turn off
Rate Distortion Optimization (RDO)	-subme (1–9)	Disable RDO < 6
Chroma Motion Estimation	-no-chroma-me	Disable
Frame Rate (FR)	-fps (15–30)	DRL uses four values 15, 20, 25, and 30
Search Range	-merange	1
Number of Reference Frames	-ref	1

### 3.2. Traffic Shaping Model Based on DRL

The output of the encoding model is the variable bit rate (VBR) of data traffic which causes packet losses due to unpredicted wireless channel capacity between the source and the destination. In order to reduce burstiness in the traffic stream that is supplied to the network, we should build the adaptive traffic shaper so as to reshape the VBR traffic. The adaptive traffic shaper's major goals are to reduce rather than to eliminate the burstiness of the multimedia data stream and establish a balance between channel capacity and buffering delay. Dynamic traffic shaping based on DRL has been proposed to forecast the best traffic rate for the multimedia stream on the VANET in order to accomplish this goal. Figure 4

depicts the proposed dynamic traffic shaping based on DRL. Since the priority of frames relies on the importance of each type of frames, the I-packets should have the highest priority to be processed, and the target bitrate (BRT) should be estimated based on the packet’s priority. The average data rate,  $r_0$ , for an  $MmNn$  GOP sequence can be written as [4]:

$$r_0 = \frac{R_I + (n - \frac{n}{m}) \times R_B + (\frac{n}{m} - 1) \times R_P}{n} \tag{1}$$

where  $R_I$ ,  $R_P$ , and  $R_B$  are the I-frame, P-frame, and B-frame data rates, respectively;  $n$  is the number of frames in the GOP and  $m$  is the I-P or P-P frame interval.

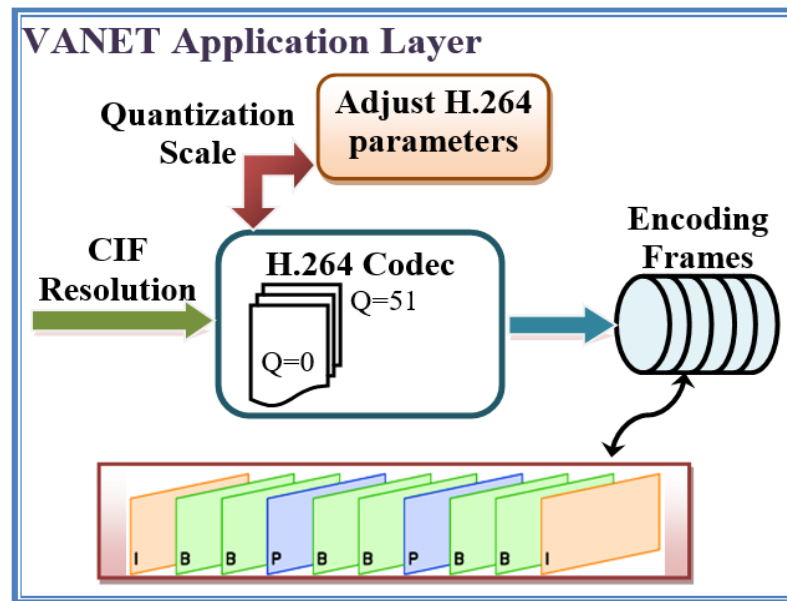


Figure 3. Block Diagram of Proposed Encoding Model.

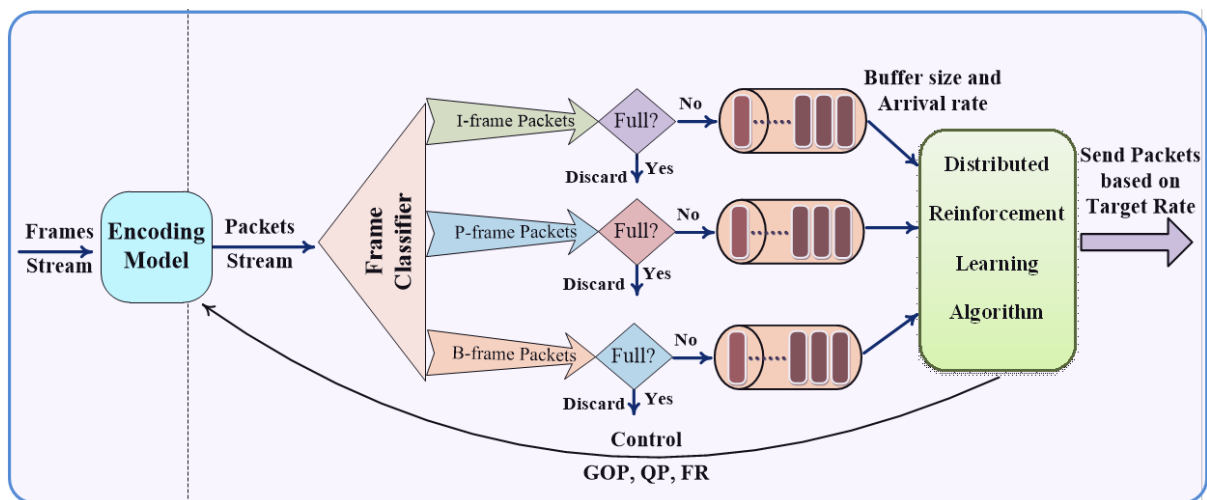


Figure 4. Block Diagram of the Proposed Traffic Shaping.

Formally, the reinforcement learning model consists of: (a) a discrete set of environmental states,  $S$ ; (b) a discrete set of agent actions,  $A$ ; and (c) a set of scalar reinforcement rewards,  $R$ . As can be described in Algorithm 1, we model the traffic shaping problem in the VANETs as follows: The entire VANET including the vehicles, the links capacity between vehicles, and the packets constitute the environment. Each link between two nodes is considered as an agent. Each link capacity of a node in the network is considered a state of the agent,  $s$ . The set of link capacity performance (bad, poor, good, very good, excellent)

in the network is the state space,  $S$ . A link capacity in a certain node can be controlled by setting the multimedia coding parameters (GOP,  $F_R$ , QP) to obtain a suitable bitrate. Hence, the possible set of actions ( $A$ ) that are allowed at each node is controlling the three parameters (GOP, FR, QP), which are investigated carefully with an example in Section 4.1 of this paper. The state transitions are equivalent to selecting the three parameters so as to generate the target bitrate that should be suitable to transmit the real-time multimedia with the available bitrate. We distribute the reinforcement learning task to each node since it is not possible to have a global view of the VANET connection capacity transitions. Therefore, in Q-learning, the Q-value  $Q(s, a)$  ( $s \in S, a \in A$ ) is used to estimate the future rewards if the agent performs a specific action  $a$  while it is in a specific state.

In the proposed model, the rewards can be determined using the available and required link bitrate (link capacity). If the possible action makes the predicted link capacity higher than the required link capacity, this is considered as a good reward (1), otherwise it is considered as a bad reward (0). By taking a series of actions in response to a changing environment, RL aims to maximize an agent's reward. This is exactly called the  $\epsilon$ -greedy action selection policy, which balances between the exploration and exploitation in the proposed mechanism. In order to build a table of Q-values for each environmental state and each possible action, the agent will explore various states through transition from state to state until it reaches the goal which is called an episode. Each time the agent arrives at the goal state, the program goes to the next episode. The proposed traffic shaping model based on DRL performs the following steps:

**Step 1:** Initialize the matrix Q.

The Q-matrix must be constructed and initialized to 0. The columns represent the actions while the rows represent the states. As it will be explained in Section 4.1, this research suggests 224 possible actions for adjusting the three parameters (GOP, FR, QP) on five video clips. Furthermore, it suggests five states which are bad, poor, good, very good, and excellent.

**Step 2:** Choose and perform action.

For each episode, a random initial state and a possible action are selected by the agent so as to go to the next state. After that, the maximum Q value for the next state based on all possible actions is estimated as:

$$Q_N(s, a) \leftarrow (1 - \alpha)Q_C(s, a) + \alpha \left\{ R(s, a) + \gamma \text{Max}_{x \in S} \tilde{Q}_x(s, a) \right\} \quad (2)$$

where  $\alpha$  denotes the learning rate, and  $\gamma$  denotes the discount factor. Both are crucial factors affecting the Q-learning algorithm. In Equation (2), the symbol  $Q_N$  denotes the new value of Q,  $Q_C$  is the current Q value, and  $\text{Max}_{x \in S} \tilde{Q}_x$  is the maximum predicted reward by the given new state and possible actions. The main objective of Equation (2) is to calculate the Q-matrix which represents the brain of the agent. The learning rate  $\alpha$  limits how quickly learning can occur. When  $\alpha$  has a value of 0, the algorithm learns nothing as it keeps using the old value and a small  $\alpha$  value results in a slow learning process. A high  $\alpha$  value makes the algorithm emphasize the new information. The discount factor controls the value of future rewards. The low values of  $\gamma$  lead the algorithm to consider the immediate rewards value, while higher values of  $\gamma$  cause the learning algorithm to count future rewards more strongly. In this research, a lot of experiments and analysis are performed to determine the most suitable learning rate  $\alpha$  which is found to be equal to 0.7. The discount factor  $\gamma$  is calculated based on signal-to-interference-plus-noise ratio (SINR) as follows:

$$\gamma = \lambda \times \left( \frac{\text{SINR}}{\text{SINR}_{\text{MAX}}} \right), \text{ where } 0 < \lambda < 1 \quad (3)$$

where  $\lambda$  is a predefined value, and  $\text{SINR}_{\text{MAX}}$  is  $-40$  dB [23]. The main contribution of this approach is that the discount factor becomes dynamic because of the changes in the SINR parameter. In contrast to the traditional RL approach that has a static discount factor, the present approach uses a dynamic discount factor that aims to provide a more accurate



estimation on the Q-values. The bitrate and hence the channel capacity are affected by SINR which can be predicted using the shadowing model at any vehicle node (e.g., Vehicle A) as follows:

$$SINR(A) = \frac{P_r(d)}{I + N}, \text{ where } P_r(d) = P_r(d_0) + 10\beta \log\left(\frac{d}{d_0}\right) + X_\sigma \quad (4)$$

where  $P_r(d)$  denotes the received signal at the vehicle A in dB,  $I$  denotes the total multipath interference, and  $N$  represents the total noise that affect the received signal at the vehicle A. In addition,  $d$  denotes the distance between RSU and the vehicle A;  $d_0$  represents the reference distance;  $\beta$  is the path loss exponent; and  $X_\sigma$  represents the zero-mean Gaussian distributed random variable in dB. In order to shape the traffic into an optimal bitrate, the three parameters (GOP, FR, QP) have been controlled. The state transitions are equivalent to selecting those three parameters in order to estimate a target bitrate (BRT) which is proportionally decreased with the increasing of QP, GOP, and the decreasing of FR. The following equation shows the correlations among the actions taken to achieve the combination between the three parameters as:

$$C\left(s, a_{(GOP,FR,QP)}\right) = W_1 \times \left(\frac{GOP}{GOP_{MAX}}\right) + W_2 \times \left(\frac{FR}{FR_{MAX}}\right) + W_3 \times \left(\frac{QP}{QP_{MAX}}\right) \quad (5)$$

where  $W_1 + W_2 + W_3 = 1$

The values  $W_1, W_2,$  and  $W_3 \in [0, 1]$  are tuning the weights assigned to each parameter which provide the optimal selection for the total calculation of combining the three parameters. It is interesting to note that the  $W_1, W_2,$  and  $W_3$  parameters have been determined based on comprehensive experiments as can be seen in Section 4. The optimal values of  $W_1, W_2,$  and  $W_3$  are 0.2, 0.2, and 0.6, respectively. The function  $C\left(s, a_{(GOP,FR,QP)}\right)$  is used in the estimation of the reward function.

**Step 3: Reward estimation.**

The reward function  $R(s, a_{(GOP,FR,QP)})$  is calculated based on the average bitrate ratio as follows:

$$\text{Bitrate Ratio}(BRR) = \frac{\text{Predicted Bitrate}}{\text{Required Bitrate}} = \frac{BR_P}{BR_R} = \frac{\beta \times \log_2(1 + SINR)}{A \times \left(e^{-B \times C(s, a_{(GOP,FR,QP)})}\right)} \quad (6)$$

$$R(s, a_{(GOP,FR,QP)}) = \begin{cases} 1 & \text{if } BRR \geq 0 \\ 0 & \text{if } BRR < 0 \end{cases}$$

where  $BR_P, BR_R,$  and  $\beta$  represent the predicted bitrate, the required bitrate, and the channel bandwidth in bps, respectively. The consistent factors  $A$  and  $B$  are estimated based on the motion complexity of the video clips as will be explained in Section 4. If the bitrate ratio ( $BRR$ ) is greater than 0, the reward is stored in the  $R$  queue as 1, otherwise it is stored as 0. The agent performs Equation (6) at each state with the appropriate action (e.g., change the three parameters) and it will stop if it achieves the target bitrate.

**Step 4: Q-Learning evaluation.**

Since the  $\epsilon$ -greedy policy is used as the action selection policy, the greediest action has been chosen at a particular state. The function  $Q(s,a)$  has been maximized and the Q-Table has been updated until the learning is stopped. When assessing the reliability of the RMDRL method, convergence is a crucial factor. In fact, reinforcement learning does not grantee convergence. However, Watkins and Dayan [40] demonstrate that Q-Learning converges to the best action-values with probability 1, provided that all actions are regularly sampled in all states and that action-values are represented discretely. Fortunately, the RMDRL method meets every requirement for convergence. According to the suggested algorithm, each connection employs the multimedia coding parameters (GOP, FR, QP) to generate a suitable bitrate, and a link's capacity is equivalent to a state. The action-values (Q-values) are obviously separately represented in the RMDRL model. We can therefore demonstrate that our suggested approach converges to the best action-values.

### 3.3. Real-Time Routing Protocol

The previous traffic shaping model determined the most suitable bitrate to stream a real-time multimedia based on DRL. After that, the proposed RMDRL can use any real-time routing protocol to forward the multimedia packets through several hops towards the destination. For instance, real-time routing employed by one of the present authors, namely RTVP [24] has been used to forward the multimedia packets through several hops towards the destination. The RTVP routing is built with features such as the corona mechanism, neighborhood management, power management, routing management, packet classifier, dynamic multipath forwarding, path interference handler, and next hop cost. The entire network is separated into coronas that are centered on the sink in a corona structure. The appropriate power and saving conditions for the transceiver for each sensor node are specified by the power management. The neighborhood management selects a set of forwarding applicant nodes and keeps track of all next hop candidate nodes in a neighborhood table. The routing management generates a forwarding decision, determines the best next hop forwarding, and activates a routing problem handler.

---

#### Algorithm 1: RMDRL pseudocode

---

**Input:** Video clip, Available Bitrate (BR), Learning rate ( $\alpha$ ), Learning episodes number (n), QP, GOP, FR.  
**Output:** High fidelity of the reconstructed video at the receiver.  
**Start Algorithm (RMDRL)**

**Phase 1, Multimedia Stream**

- 1 | **While (new multimedia session start) do**
- 2 |     *Get the vehicle Addresses (); //the source and destination IP address*
- 3 |     *Create the real-time multimedia steam (); //The video stream should be converted to MP4*

**Phase 2, Multimedia Encoder**

- 4 |     *Get video clip();*
- 5 |     *Adjust H.264 Encoder Parameters as shown in Table 1 ();*
- 6 |     *Initialize (GOP, FR, QP);*
- 7 |     *Initialize Real-time Routing Protocol ();*
- 8 |     *Create video data traffic files using x.264 and mp4trace(); //Trace file contains three types of packet I, P, and B.*

**Phase 3, Traffic Shaping with DRL**

- 9 |     *Initialize Q-Table:  $Q(s,a) = 0$  for  $\forall s \in S, \forall a \in A$  (GOP, FR, QP);*
- 10 |     *for episode  $\leftarrow 1$  to n do;*
- 11 |     *Chose action at a given state s and calculate C(GOP, FR, QP) according to Equation (5);*
- 12 |     
$$Q_N(s,a) \leftarrow (1 - \alpha)Q_C(s,a) + \alpha \left\{ R(s,a) + \gamma \text{Max}_{x \in S} \tilde{Q}_x(s,a) \right\}$$
- 13 |     *Move to next State S;*
- 14 |     *end;*
- 15 |     *Adjust the value of (GOP, FR, QP) based on the output of DRL; // as shown in Figure 4*

**Phase 4, Real-time Routing Protocol**

- 16 |     *Forward I frame packets ();*
- 17 |     *Forward P frame packets ();*
- 18 |     *Forward B frame packets ();*

**Phase 5, Problem Handler**

- 19 |     *Solve Routing Problem ();*
- 20 |     **End; //While loop**

**Phase 6, Receive Multimedia Packets**

- 21 |     *Receive all packets at the destination ();*

**Phase 7, Mutimeia Decoder and Performance Evaluation**

- 22 |     *Reconstruct the video clip ();*
- 23 |     *Calculate PSNR and Frame Delay (); //QoE evaluation*
- 24 | **End; //Algorithm**

---

#### 4. Simulation Experiments and Performance Evaluation

The experimental work in this research is divided into two parts: (a) the study of the impact of the main three parameters (GOP, QP, FR) on the traffic shaping rate, and (b) the evaluation of the performance of RMDRL using the NS-2 simulator framework.

##### 4.1. Impact of Video Coding on Traffic Shaping

The impact of video coding has been tested based on the three parameters which have been mentioned in this paper (i.e., GOP, QP, FR). First, the impact of each single parameter on traffic shaping has been evaluated. After that, the combination of the two parameters GOP and FR has been tested. Finally, the three parameters have been considered to study the impact of video coding on traffic shaping. The experiment is implemented using an opensource of video coding, namely Xvid [38]. Five of the uncompressed YUV videos have been downloaded from the video coders' website [41] as described in Table 2. In order to generate the video coding in mp4 format, x264 program, which is part of the Xvid software has been utilized with an optimal configuration that is presented in Table 1. Additionally, these video clips' motion content can be divided into three groups based on how complicated the scenes are: high motion/scene complexity, medium motion/scene complexity, and low motion/scene complexity. Using the five video clips in Table 2, we found that the GOP parameter varies within 14 discrete values which are 10, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 220, 240, and 250. Furthermore, the QP parameter is found to vary within four discrete values which are 10, 20, 30, and 40. The FR parameter is observed to vary within four discrete values which are 15, 20, 25, and 30. Therefore, 224 experiments have been performed with all the possible combinations of values of the three parameters so as to obtain the target bitrate that is suitable for real-time multimedia streaming on the 5G-VANET. The results show that the three parameters should be considered to adjust the traffic shaping.

**Table 2.** Experiment configuration.

Video Clip in CIF Resolution (352 × 288)	Number of Frames	Motion Content
Akiyo	300	Low
Foreman	300	High
Highway	1200	High
Mobile	300	High
Bus	300	Medium

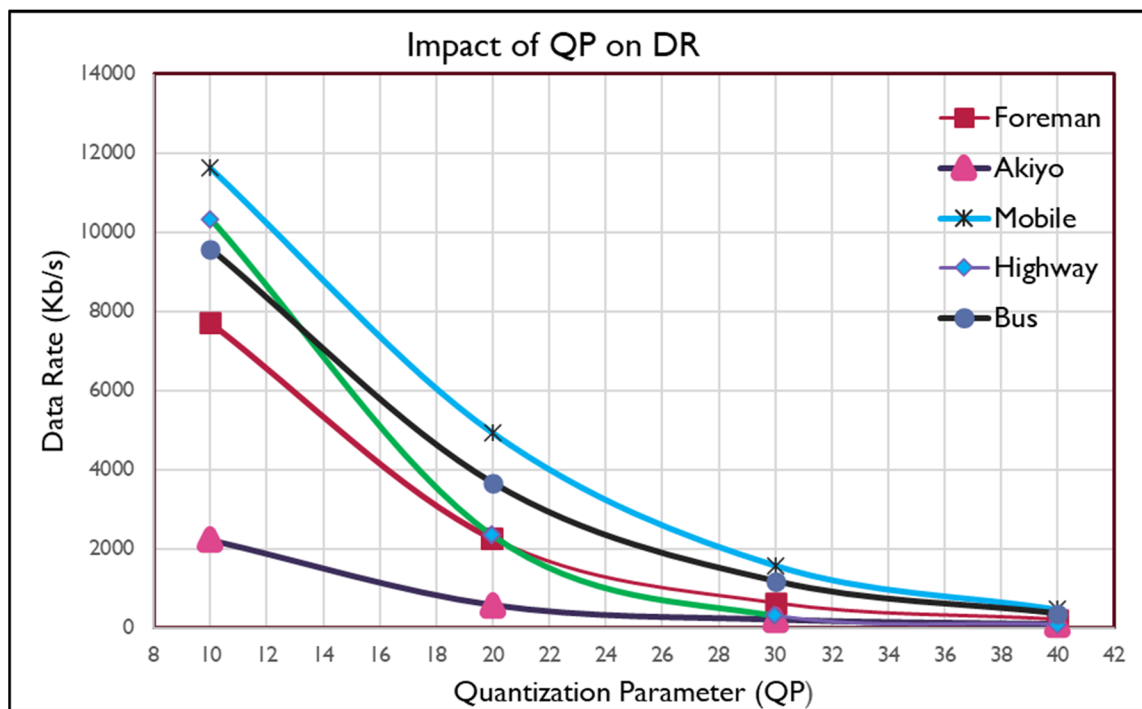
- Results and Discussion

Figure 5 consists of five parts of experiments that explain the impact of the three important video coding parameters on traffic shaping. Figure 5a shows the impact of the QP parameter on the data rate that is essential for real-time multimedia streaming. The data rate is decreased exponentially with the increase in the QP parameter as shown in Figure 5a. In addition, Figure 5a shows that the decreasing rate of multimedia stream by increasing the QP parameter is between 25 times and 110 times. This is mainly because when the QP is increasing, most of the luminance component of the pixels is zero which decreases the frame size. Moreover, Figure 5b shows the impact of the GOP parameter on the data rate, which is decreased linearly with the increasing of the GOP. Figure 5b also shows that the decreasing rate of multimedia stream by increasing the GOP parameter is between 1.1 and 1.6 times. This is mainly due to the fact that the GOP influences the number of I-frames only. Figure 5c shows the impact of the FR parameter on the data rate, which is decreased linearly with the decrease in the FR parameter. Figure 5c also shows that the FR parameter does not have a high impact on the data rate of the multimedia stream parameter which is decreasing by 1.1–1.6 times. This is mainly due to the fact that the FR influences the number of frames sent in a specific period. In addition, Figure 5d shows the impact of the combination of the two parameters GOP and FR on the data rate of real-time multimedia

traffic shaping. As can be shown in the figure, the data rate of traffic shaping is decreased by 2.1–3.2 times when the GOP and FR parameters are combined. The linear decreasing of the data rate due to using the combination of GOP and FR experiences better enhancement compared with using each parameter individually. This is primarily due to the decreasing of both the I-frame and the frame rate at the same time. However, the decreasing rate of the combination of GOP and FR is much less than the impact of QP alone. Finally, Figure 5e shows the impact of the combination of the three parameters which are GOP, FR, and QP on the data rate of real-time multimedia traffic shaping. As can be shown in Figure 5e, the data rate of traffic shaping is decreased by 86–400 times when the three parameters are combined. Furthermore, the figure shows that the decreasing rate is exponential and goes 10 times less than when using the QP alone. This is primarily due to the fact that most of the luminance component of the pixels is zero and the I-frame that increases the frame rate is much lesser. The important achievement in the combination of the three parameters can be derived and represented with the following equation:

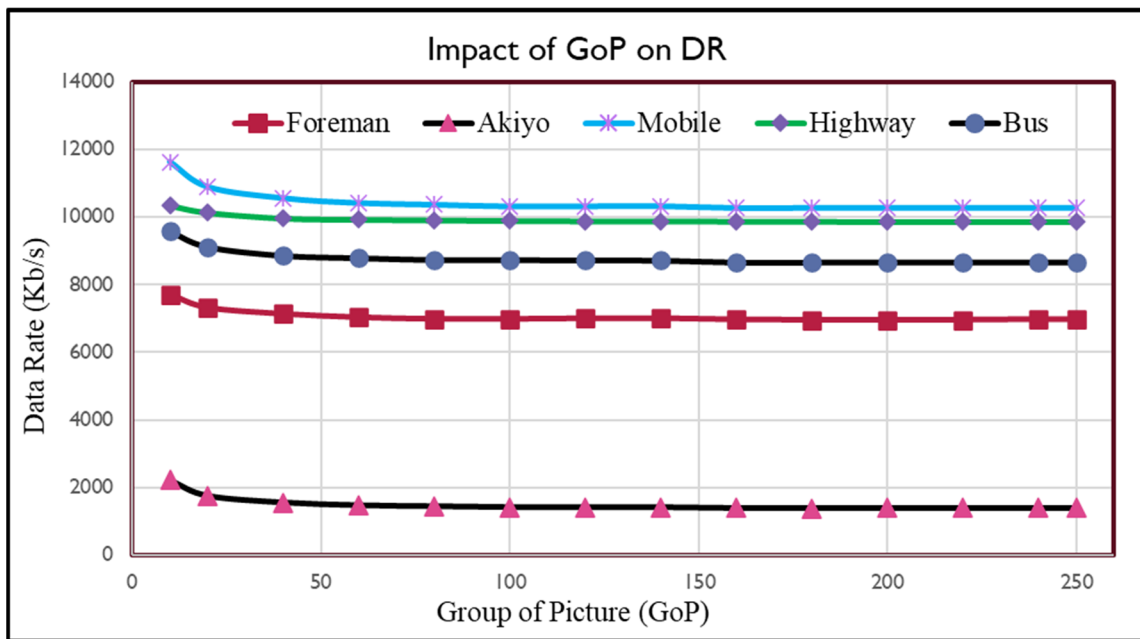
$$Data\ Rate = A \times \left( e^{-(B \times C(GOP,FR,QP))} \right) \tag{7}$$

The Equation (7) has been derived based on the average of the data rate for the five video clips. As a result, this equation has been used in Equation (6) to measure the reward value of DRL.

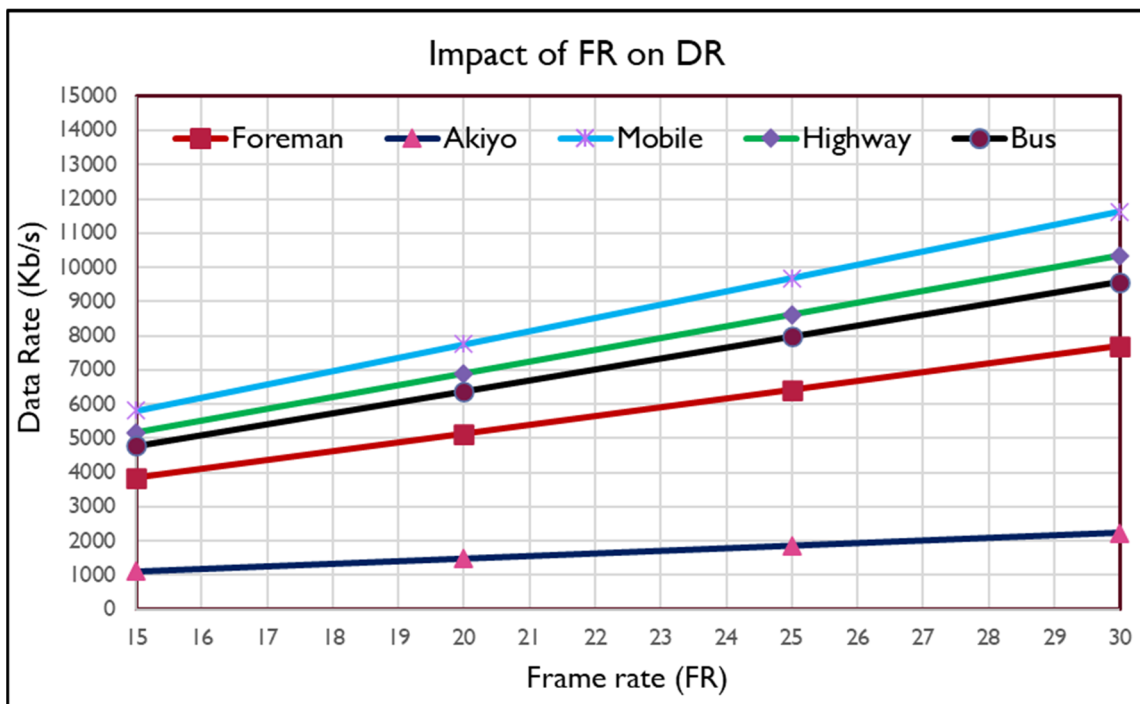


(a)

Figure 5. Cont.

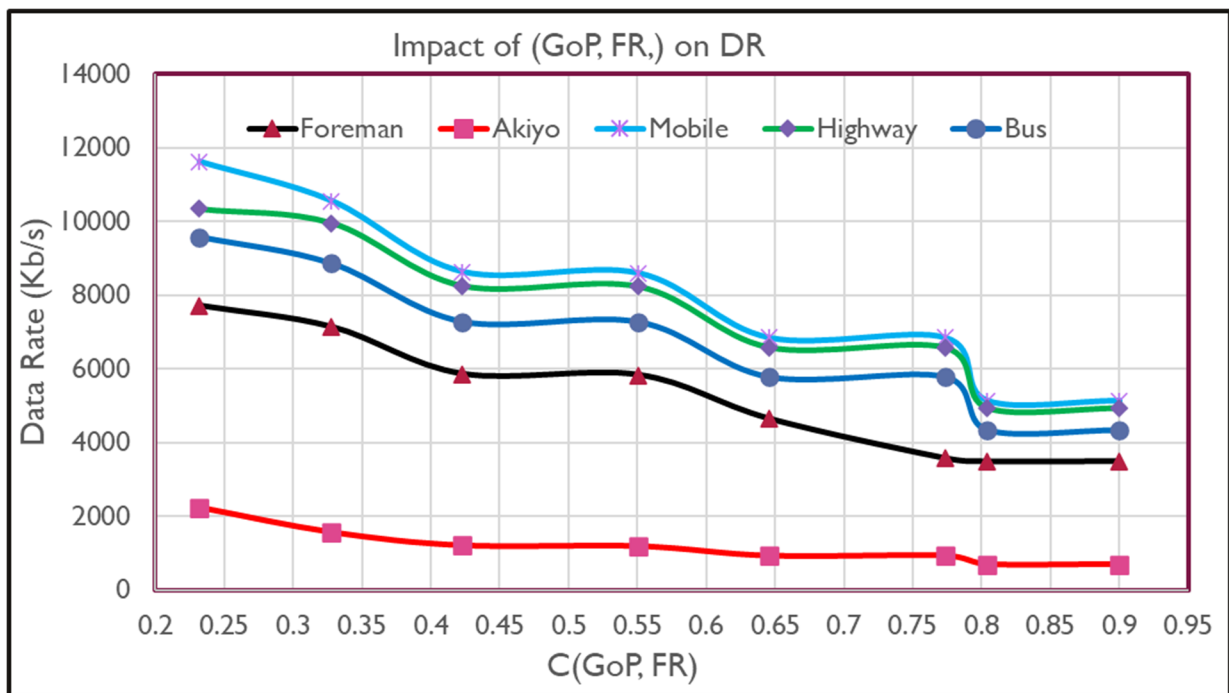


(b)

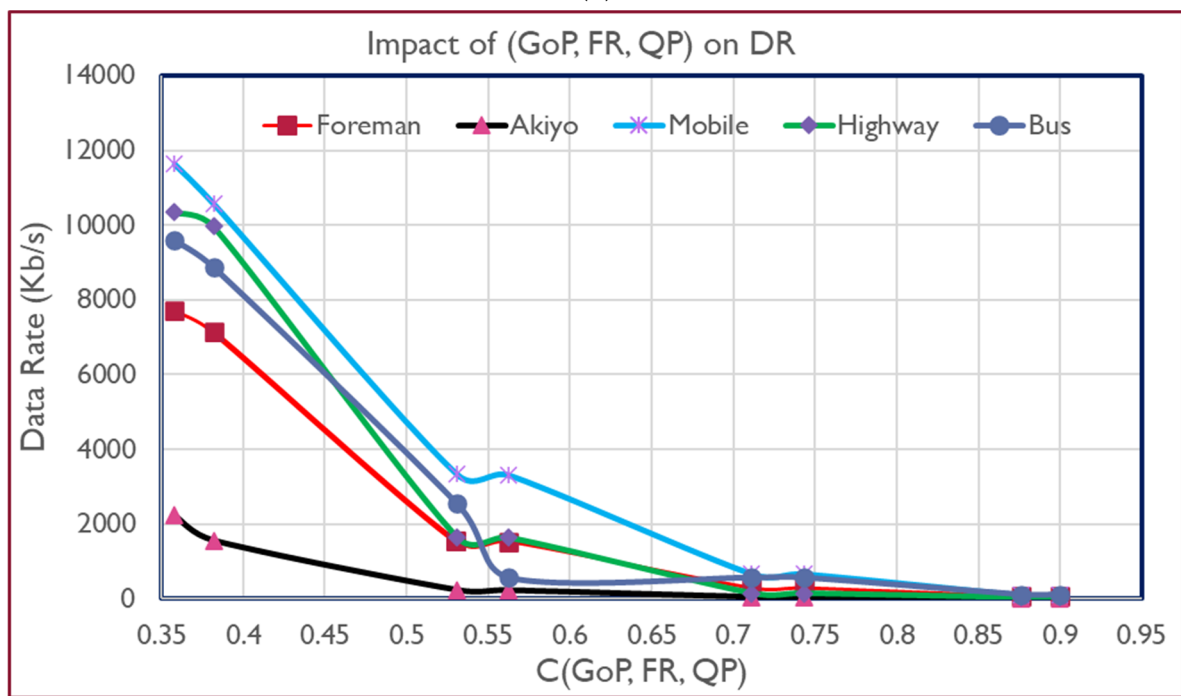


(c)

Figure 5. Cont.

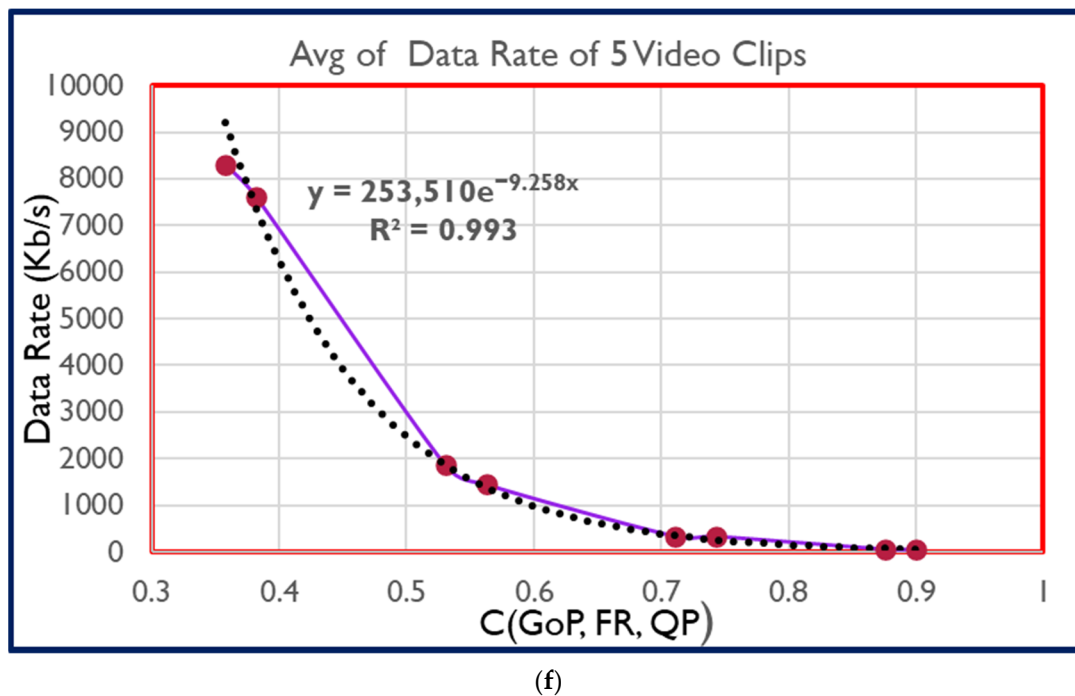


(d)



(e)

Figure 5. Cont.



**Figure 5.** Impact of three parameters on the data rate of traffic shaping: (a) QP impact; (b) GOP impact; (c) FR impact; (d) Impact of the combination of the two parameters GOP and FR; (e) Impact of the combination of the three parameters GOP, FR, and QP; (f) The average of the data rate and the equation derivation for the impact of C(GOP, FR, and QP).

#### 4.2. Performance Evaluation of RMDRL

The first stage of using DRL is creating the Q-matrix based on Equation (2) at each agent. The Q-matrix consists of five states and 224 actions which relies on the configuration of the three parameters. The R-table is calculated based on Equation (6). In this experiment, the five video clips that have been mentioned in Table 2 are used to evaluate the QoE performance in terms of the PSNR and the frame delay. First, the pre-processing for the video clip is implemented using the x.264 and mp4trace programs which generate video data traffic for NS-2 network simulation. In this experiment, the network parameters have been configured exactly as in the previous work of the ARTVP [24]. The network topology consists of 121 nodes, 13–19 hops between the source and the destination, and the number of data packets and control packets depends on the number of frames that are generated from the video coding program. Second, the received frames are used to construct the video clip at the destination node. Finally, the output video file post-processing uses the etmp4 and ffmpeg programs to rebuild the output multimedia based on the input and output tracing files that are generated during network simulation. The QoE performance in terms of the PSNR and the frame delay are calculated for the RMDRL model and compared with corresponding quantities in the ARTVP model. The PSNR is defined as the subjective evaluation ratio between the maximum possible intensity value of an image and the power of the corrupting noise for the received image. If the PSNR is less than 27, this means the presence of visible noise or smoothing of many edges. By contrast, the case when PSNR is greater than 35, corresponds to very low noise, wherein the received image is considered as excellent. Additionally, the data generated from the sender and the receiver trace files is used to calculate the frame delay, which is calculated as the interval between the time the first packet of a frame is delivered and the time the last packet of the same frame is received.

- Results and Discussion

Table 3 shows the sample output of the five video frames containing artefacts for ARTVP and RMDRL. As can be seen in Table 3, RMDRL produces better imperceptible video compared with ARTVP. In addition, Figure 6 presents the QoE performance of

RMDRL and the baseline traffic shaping (e.g., ARTVP) in terms of the PSNR and the frame delay. As shown in Figure 6a, RMDRL experiences on average 10.23% higher PSNR compared with that experienced by ARTVP. Furthermore, Figure 6b shows that RMDRL experiences on average 25.5% less frame delay compared with that of ARTVP. The impressive QoE performance of the RMDRL algorithm that provides perfect reconstructed video fidelity at the receiver is mainly achieved due to the following reasons: Firstly, RMDRL adjusts the coding parameters (GOP, FR, QP) to fit the unpredictable change of the channel capacity which can manipulate the frame size and hence the total number of packets that should be sent through the channel. By contrast, ARTVP focuses on the channel and buffering optimization which is limited by the availability of bandwidth and it is affected by the SINR. Secondly, the frame delay is optimized in RMDRL due to less retransmission of the frame packets, less packet size, and less bitrate requirement. Finally, the smart distributed reinforcement learning causes RMDRL to be fast in selecting the appropriate coding optimization parameters to fit the unpredictable change in the channel capacity and hence the congestion of packets at the sending buffer is prevented.

**Table 3.** Sample output of video frames containing artefacts for ARTVP and RMDRL.



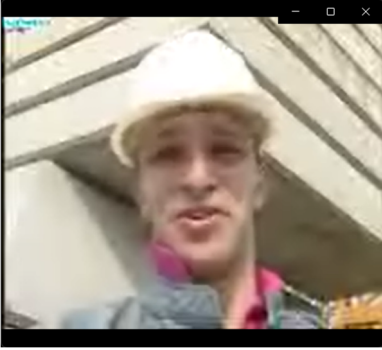







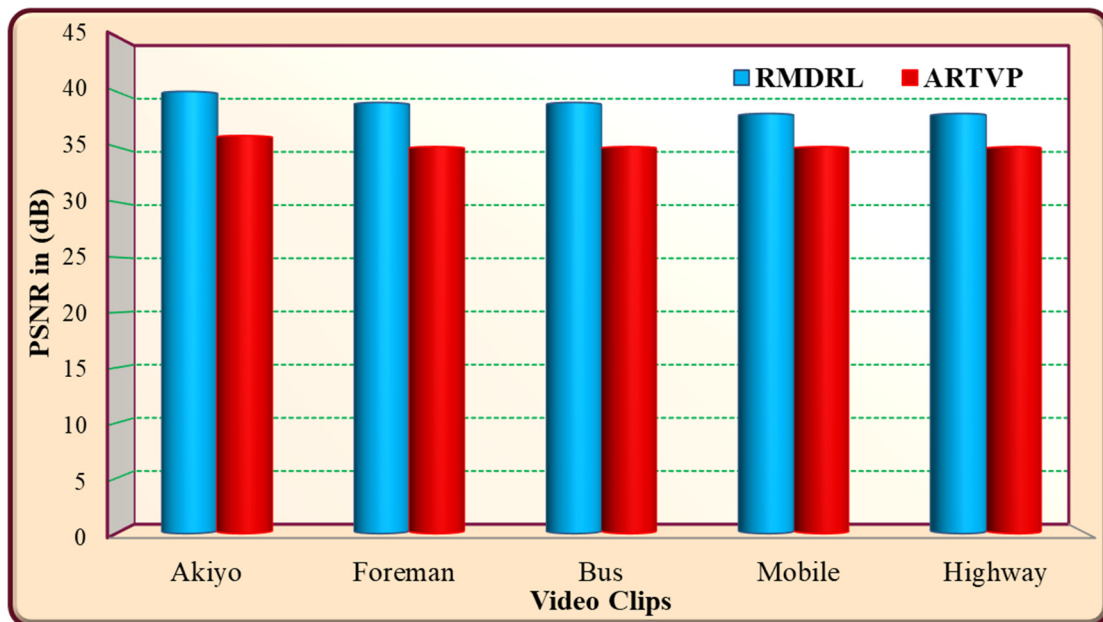
Video Clip in CIF Resolution (352 × 288)	ARTVP	RMDRL
Akiyo		
Foreman		
Highway		



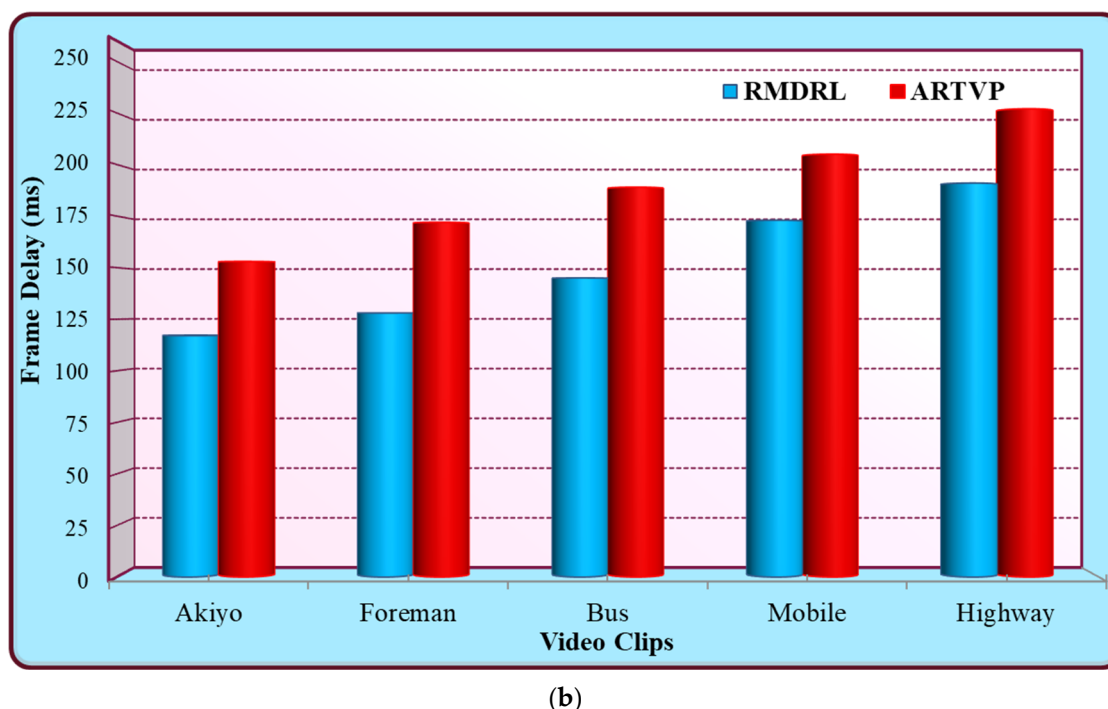
Table 3. Cont.

Video Clip in CIF Resolution (352 × 288)	ARTVP	RMDRL
Mobile		
Bus		



(a)

Figure 6. Cont.



**Figure 6.** QoE performance for RMDRL and ARTVP: (a) PSNR; (b) Frame delay.

## 5. Conclusions and Future Work

This research proposed a smart real-time multimedia traffic shaping mechanism based on distributed reinforcement learning (RMDRL) which makes the accurate decisions of selecting the GOP, FR, and QP parameters that are used to manipulate the coding of real-time multimedia stream on the 5G-VANET. In addition, the impact of adapting the three parameters of video coding QP, GOP, and FR has been investigated to achieve the optimal traffic rate value for real-time multimedia streaming on the 5G-VANET, which will increase the throughput and reduce the video bitrate with a high quality. The findings in the experimental results show the efficiency and effectiveness of RMDRL performance in terms of the perfect video fidelity of the reconstructed video at the receiver using the PSNR, which is greater than 35 dB and a frame delay that is less by 25% compared with the baseline traffic shaping protocol ARTVP. The future work of this research is developing a smart routing protocol based on distributed reinforcement learning over the 5G-VANET.

**Author Contributions:** Conceptualization, A.A.A. and W.A.; methodology, A.A.A.; software, A.A.A.; validation, S.J.M., O.M.B., and W.A.; formal analysis, S.J.M.; investigation, W.A.; resources, A.A.A.; data curation, O.M.B.; writing—original draft preparation, A.A.A.; writing—review and editing, W.A.; visualization, A.A.A.; supervision, Adel A.; project administration, A.A.A.; funding acquisition, A.A.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** King Abdulaziz University-Institutional Funding Program for Research and Development-Ministry of Education: IFPIP: 213-830-1443.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This research work was funded by Institutional Fund Projects under grant no. (IFPIP: 213-830-1443). The authors gratefully acknowledge technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

**Conflicts of Interest:** The authors declare that no conflict of interest exists.

## References

1. Bentaleb, A.; Taani, B.; Begen, A.C.; Timmerer, C.; Zimmermann, R. A survey on bitrate adaptation schemes for streaming media over HTTP. *IEEE Commun. Surv. Tutor.* **2018**, *21*, 562–585. [\[CrossRef\]](#)
2. Ahmed, A.A.; Alzahrani, A.A. A comprehensive survey on handover management for vehicular ad hoc network based on 5G mobile networks technology. *Trans. Emerg. Telecommun. Technol.* **2019**, *30*, e3546. [\[CrossRef\]](#)
3. Taha, M.; Ali, A.; Lloret, J.; Gondim, P.R.; Canovas, A. An automated model for the assessment of QoE of adaptive video streaming over wireless networks. *Multimed. Tools Appl.* **2021**, *80*, 26833–26854. [\[CrossRef\]](#)
4. Alam, M.F.; Atiquzzaman, M.; Karim, M.A. Traffic shaping for MPEG video transmission over the next generation internet. *Comput. Commun.* **2000**, *23*, 1336–1348. [\[CrossRef\]](#)
5. Trestian, R.; Comsa, I.S.; Tuysuz, M.F. Seamless multimedia delivery within a heterogeneous wireless networks environment: Are we there yet? *IEEE Commun. Surv. Tutor.* **2018**, *20*, 945–977. [\[CrossRef\]](#)
6. Vega, M.T.; Perra, C.; Liotta, A. Resilience of video streaming services to network impairments. *IEEE Trans. Broadcast.* **2018**, *64*, 220–234. [\[CrossRef\]](#)
7. Barakabitze, A.A.; Walshe, R. SDN and NFV for QoE-driven multimedia services delivery: The road towards 6G and beyond networks. *Comput. Netw.* **2022**, *214*, 109133. [\[CrossRef\]](#)
8. Uzakgider, T.; Cetinkaya, C.; Sayit, M. Learning-based approach for layered adaptive video streaming over SDN. *Comput. Netw.* **2015**, *92*, 357–368. [\[CrossRef\]](#)
9. Hossain, M.B.; Wei, J. Reinforcement Learning-Driven QoS-Aware Intelligent Routing for Software-Defined Networks. In Proceedings of the 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Ottawa, ON, Canada, 11–14 November 2019; pp. 1–5.
10. Rekkas, V.P.; Sotiroudis, S.; Sarigiannidis, P.; Wan, S.; Karagiannidis, G.K.; Goudos, S.K. Machine Learning in Beyond 5G/6G Networks—State-of-the-Art and Future Trends. *Electronics* **2021**, *10*, 2786. [\[CrossRef\]](#)
11. Nassef, O.; Sun, W.; Purmehdi, H.; Tatipamula, M.; Mahmoodi, T. A survey: Distributed Machine Learning for 5G and beyond. *Comput. Netw.* **2022**, *207*, 108820. [\[CrossRef\]](#)
12. Kaur, J.; Khan, M.A.; Iftikhar, M.; Imran, M.; Haq, Q.E.U. Machine Learning Techniques for 5G and Beyond. *IEEE Access* **2021**, *9*, 23472–23488. [\[CrossRef\]](#)
13. Karunathilake, T.; Förster, A. A Survey on Mobile Road Side Units in VANETs. *Vehicles* **2022**, *4*, 482–500. [\[CrossRef\]](#)
14. Hsieh, Y.L.; Wang, K. Dynamic overlay multicast for live multimedia streaming in urban VANETs. *Comput. Netw.* **2012**, *56*, 3609–3628. [\[CrossRef\]](#)
15. Nakano, T.; Nakagawa, R.; Yamai, N. Mitigating Congestion with Decentralized Traffic Shaping for Adaptive Video Streaming over ICN. In Proceedings of the 17th Asian Internet Engineering Conference, Hiroshima, Japan, 19–21 December 2022; pp. 36–43.
16. Kua, J.; Armitage, G.; Branch, P. A survey of rate adaptation techniques for dynamic adaptive streaming over HTTP. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 1842–1866. [\[CrossRef\]](#)
17. Anand, D.; Togou, M.A.; Muntean, G.M. A Machine Learning Solution for Automatic Network Selection to Enhance Quality of Service for Video Delivery. In Proceedings of the 2021 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Chengdu, China, 4–6 August 2021; pp. 1–5.
18. Lekharu, A.; Moulai, K.; Sur, A.; Sarkar, A. Deep Learning Based Prediction Model for Adaptive Video Streaming. In Proceedings of the 2020 International Conference on COMMunication Systems & NETWORKS (COMSNETS), Bangalore, India, 7–11 January 2020; pp. 152–159.
19. Godfrey, D.; Kim, B.S.; Miao, H.; Shah, B.; Hayat, B.; Khan, I.; Sung, T.E.; Kim, K.I. Q-learning based routing protocol for congestion avoidance. *Comput. Mater. Contin.* **2021**, *68*, 3671–3692. [\[CrossRef\]](#)
20. Gao, Z. 5G Traffic Prediction Based on Deep Learning. *Comput. Intell. Neurosci.* **2022**, *2022*, 3174530. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Wang, Z.; Hu, J.; Min, G.; Zhao, Z.; Chang, Z.; Wang, Z. Spatial-Temporal Cellular Traffic Prediction for 5G and Beyond: A Graph Neural Networks-Based Approach. *IEEE Trans. Ind. Inform.* **2022**, *2022*, 2190–2202.
22. Zhou, S.; Wei, C.; Song, C.; Pan, X.; Chang, W.; Yang, L. Short-Term Traffic Flow Prediction of the Smart City Using 5G Internet of Vehicles Based on Edge Computing. *IEEE Trans. Intell. Transp. Syst.* **2022**, *2022*, 1–10. [\[CrossRef\]](#)
23. Ahmed, A.A. An effective handover management based on SINR and software-defined network over urban vehicular ad hoc networks. *Trans. Emerg. Telecommun. Technol.* **2019**, *30*, e3787. [\[CrossRef\]](#)
24. Ahmed, A.A. A real-time routing protocol with adaptive traffic shaping for multimedia streaming over next-generation of Wireless Multimedia Sensor Networks. *Pervasive Mob. Comput.* **2017**, *40*, 495–511. [\[CrossRef\]](#)
25. Wu, C.; Kumekawa, K.; Kato, T. Distributed reinforcement learning approach for vehicular ad hoc networks. *IEICE Trans. Commun.* **2010**, *93*, 1431–1442. [\[CrossRef\]](#)
26. Shin, K.S.; Hwang, G.H.; Jo, O. Distributed reinforcement learning scheme for environmentally adaptive IoT network selection. *Electron. Lett.* **2020**, *56*, 462–464. [\[CrossRef\]](#)
27. Akbari, Y.; Tabatabaei, S. A new method to find a high reliable route in IoT by using reinforcement learning and fuzzy logic. *Wirel. Pers. Commun.* **2020**, *112*, 967–983. [\[CrossRef\]](#)
28. Rossi, M.; Centenaro, M.; Ba, A.; Eleuch, S.; Erseghe, T.; Zorzi, M. Distributed learning algorithms for optimal data routing in IoT networks. *IEEE Trans. Signal Inf. Process. Over Netw.* **2020**, *6*, 179–195. [\[CrossRef\]](#)

29. Lai, W.K.; Lin, M.T.; Yang, Y.H. A machine learning system for routing decision-making in urban vehicular ad hoc networks. *Int. J. Distrib. Sens. Netw.* **2015**, *11*, 374391. [[CrossRef](#)]
30. Immich, R.; Cerqueira, E.; Curado, M. Efficient high-resolution video delivery over VANETs. *Wirel. Netw.* **2019**, *25*, 2587–2602. [[CrossRef](#)]
31. Ben Ameer, C.; Mory, E.; Cousin, B. Combining traffic-shaping methods with congestion control variants for HTTP adaptive streaming. *Multimed. Syst.* **2018**, *24*, 1–18. [[CrossRef](#)]
32. Al Jameel, M.; Kanakis, T.; Turner, S.; Al-Sherbaz, A.; Bhaya, W.S. A Reinforcement Learning-Based Routing for Real-Time Multimedia Traffic Transmission over Software-Defined Networking. *Electronics* **2022**, *11*, 2441. [[CrossRef](#)]
33. Marwah, G.P.K.; Jain, A.; Malik, P.K.; Singh, M.; Tanwar, S.; Safirescu, C.O.; Mihaltan, T.C.; Sharma, R.; Alkhayat, A. An Improved Machine Learning Model with Hybrid Technique in VANET for Robust Communication. *Mathematics* **2022**, *10*, 4030. [[CrossRef](#)]
34. Abdellah, A.R.; Muthanna, A.; Essai, M.H.; Koucheryavy, A. Deep Learning for Predicting Traffic in V2X Networks. *Appl. Sci.* **2022**, *12*, 10030. [[CrossRef](#)]
35. Vergados, D.J.; Kravetska, K.; Michalas, A.; Vergados, D.D. Evaluation of HTTP/DASH Adaptation Algorithms on Vehicular Networks. In Proceedings of the 2018 Global Information Infrastructure and Networking Symposium (GIIS), Thessaloniki, Greece, 23–25 October 2018; pp. 1–5.
36. Esmaily, A.; Kravetska, K. Small-scale 5G testbeds for network slicing deployment: A systematic review. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 6655216. [[CrossRef](#)]
37. Bitmovin's Industry Report, Video Developer Report 2021/22. Available online: <https://go.bitmovin.com/video-developer-report-2021> (accessed on 20 January 2023).
38. Open Source Video Codec. Available online: [www.xvid.com/download/](http://www.xvid.com/download/) (accessed on 13 September 2022).
39. Ahmed, A.A. An optimal complexity H.264/AVC encoding for video streaming over next generation of wireless multimedia sensor networks. *Signal Image Video Process.* **2016**, *10*, 1143–1150. [[CrossRef](#)]
40. Watkins, C.J.C.H.; Dayan, P. Q-Learning. *Mach. Learn.* **1992**, *8*, 279–292. [[CrossRef](#)]
41. YUV Sequences. Available online: <http://trace.eas.asu.edu/yuv/> (accessed on 14 September 2022).

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.