

Article

Intelligent Stretch Optimization in Information Centric Networking-Based Tactile Internet Applications [†]

Hussain Ahmad , Muhammad Zubair Islam , Rashid Ali , Amir Haider *  and Hyungseok Kim * 

School of Intelligent Mechatronics Engineering, Sejong University, Seoul 05006, Korea; hussainahmad@sju.ac.kr (H.A.); zubair@sju.ac.kr (M.Z.I.); rashidali@sejong.ac.kr (R.A.)

* Correspondence: amirhaider@sejong.ac.kr (A.H.); hyungkim@sejong.ac.kr (H.K.)

[†] This paper is an extended version of our paper published in Proceedings of the Ninth International Conference on Green and Human Information Technology (ICGHIT 2021), Jeju, Korea, 13–15 January 2021.

Abstract: The fifth-generation (5G) mobile network services are currently being made available for different use case scenarios like enhanced mobile broadband, ultra-reliable and low latency communication, and massive machine-type communication. The ever-increasing data requests from the users have shifted the communication paradigm to be based on the type of the requested data content or the so-called information-centric networking (ICN). The ICN primarily aims to enhance the performance of the network infrastructure in terms of the stretch to opt for the best routing path. Reduction in stretch merely reduces the end-to-end (E2E) latency to ensure the requirements of the 5G-enabled tactile internet (TI) services. The foremost challenge tackled by the ICN-based system is to minimize the stretch while selecting an optimal routing path. Therefore, in this work, a reinforcement learning-based intelligent stretch optimization (ISO) strategy has been proposed to reduce stretch and obtain an optimal routing path in ICN-based systems for the realization of 5G-enabled TI services. A Q-learning algorithm is utilized to explore and exploit the different routing paths within the ICN infrastructure. The problem is designed as a Markov decision process and solved with the help of the Q-learning algorithm. The simulation results indicate that the proposed strategy finds the optimal routing path for the delay-sensitive haptic-driven services of 5G-enabled TI based upon their stretch profile over ICN, such as the augmented reality /virtual reality applications. Moreover, we compare and evaluate the simulation results of proposed ISO strategy with random routing strategy and history aware routing protocol (HARP). The proposed ISO strategy reduces 33.33% and 33.69% delay as compared to random routing and HARP, respectively. Thus, the proposed strategy suggests an optimal routing path with lesser stretch to minimize the E2E latency.

Keywords: ICN; 5G; reinforcement learning; markov decision process; AR/VR; stretch reduction; tactile internet



Citation: Ahmad, H.; Zubair Islam, M.; Ali, R.; Haider, A.; Kim, H.S. Intelligent Stretch Optimization in Information Centric Networking-Based Tactile Internet Applications. *Appl. Sci.* **2021**, *11*, 7351. <https://doi.org/10.3390/app11167351>

Academic Editor: Christos Bouras

Received: 26 June 2021

Accepted: 7 August 2021

Published: 10 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

The evolution of wireless and cellular communication has played a key role in making lives easy and smart [1]. European telecommunications standards institute and international telecommunication union have worked and improved the cellular networks to the current; fifth-generation (5G) [2,3]. The 5G provides faster communication, wider bandwidth, and higher throughput [4]. The 5G enables enhanced mobile broadband communications (eMBB), ultra-low latency and reliable communication (URLLC), and massive machine-type communication services depending on the use cases [5]. Mainly, 5G and beyond networks offer seamless communication and support various applications of tactile internet (TI) such as teleoperation, virtual reality (VR), augmented reality (AR), etc. [6]. These TI applications demand ultra-low delay of 1 ms, ultra-high reliability of 1×10^{-7} , and availability up to 99.999999% [7]. The 5G provides URLLC and eMBB services to users that improve quality of service (QoS), quality of tasks (QoT), and quality of experience (QoE) for TI applications [8]. Recently, the AR/VR applications with haptic feedback

are getting more attention in 5G-enabled TI [9]. Moreover, the TI applications and the application specific requirements and characteristics are illustrated in Table 1.

Table 1. TI requirements and characteristics for specific use cases [7].

Use Case Scenario	Traffic Type	Reliability	Latency (ms)	Average Data Rate
Immersive VR	Master → Slave Slave → Master	99.999%	<5 1–50	1–4 k pkts/s
Tele-Operation	Master → Slave Slave → Master	99.999%	1–10	1–4 k pkts/s
Automotive	Master → Slave Slave → Master	99.9%	1–10	100–2000 pkts/s 100–500 pkts/s
Internet of drones	Master → Slave Slave → Master	99.999%	50–100 case dependent	1–4 k pkts/s
Inter-personal Communication	User 1 → User 2	99.999%	50 [10] only VR users	600 Mbps
Cooperative automated driving	Master → Slave Slave → Master	99.9%	1–20 10–50	10–40 Mbps

The 5G network is an application-driven network infrastructure [11]. The research community has designed diverse network functionalities that enable 5G to meet the rapidly increasing demand for TI applications. The Multi-access edge computing (MEC) paradigm is one of the best key enablers of URLLC in the 5G ecosystem [12]. It brings the cloud-like capabilities at the edge of the network with limited abilities, i.e., caching, communication, control, and computation. The edge computing framework for 5G and beyond networks has three tiers, i.e., cloud, core, and edge devices [13]. Base station acts as an edge device for MEC-based 5G infrastructure. The massive increase in data requests results in a significant load on the network [14]. Thus, it is essential to offload some data requests and services at the edge of the network [15]. The edge device filters, parses, processes, and forwards data requests to the core network [16]. The MEC enables resource utilization, task offloading and management, content, and cache management that improves the QoE, QoS, and QoT for 5G at the edge of the network [17]. The MEC also lowers the traffic overhead at the core network [18]. Therefore, MEC is the driving factor to enable URLLC, high bandwidth, location awareness, and contextual services at the edge of the network [19]. Hence, MEC-based 5G is one of the key enablers to realize TI and allow haptic-driven AR/VR users to access the concerned application data from edge devices [20,21]. Therefore, the current MEC-based 5G system provides URLLC services for haptic-driven AR/VR applications [22,23].

Edge caching is the key enabler to cache the most requested content using information-centric networking (ICN), such as videos, close to users [24]. Recent studies also revealed that tactile users or haptic-driven TI applications such as inter-personal communication for AR/VR users are more concerned about the data presence than the geographical location of requested data [25,26]. This leads to the introduction of ICN for MEC [27,28]. The ICN enables named-based networking rather than the host addressing scheme in internet protocol (IP) [29]. The ICN uses content names for communication, which makes the communication architecture simple. The ICN is more concerned about content security rather than communication channel security [30]. Therefore, ICN contributes in achieving URLLC that improves QoS and maximizes QoE for the inter-personal communication of AR/VR users [31]. Moreover, ICN also minimizes the routing problems that lead to better network throughput in MEC-based 5G-enabled TI infrastructure [32]. Hence, it is a promising solution to enhance the MEC performance, and efficiency in terms of URLLC [33]. The ICN deals with the data cache and management in the network infrastructure. The ICN enables in-network caching in 5G core network. Therefore, ICN enables routing optimization and task scheduling in MEC-based 5G networks [34]. It lessens the forwarding of requests to the core network and tries to cache the most popular contents to the most appropriate place or node, depending on the cache strategies [35,36]. The ICN opens the 5G communication with other benefits like in-network caching, mobility, and multi-cast [37]. The ICN provides efficient caching and routing for MEC-based 5G system.

Furthermore, as discussed in Table 1 the inter-personnel communication deals with the users with different connected routers therefore, ICN enables content, cache, and routing optimization, and therefore, it increases QoS for inter-personal communication for AR/VR applications. Hence, ICN enabled MEC-based 5G system ensures high bandwidth capacity, content caching capability, and routing optimization to support low latency, and high reliability to users [38].

In ICN, the routing is performed based on three tables, i.e., content store (CS), pending interest table (PIT), and forward interest base (FIB) [39]. The caching node stores a list of caching objects in its CS. In a radio access network (RAN), the caching node has a PIT that monitors and forwards the state of interest packets for the content object [40]. Similarly, the node enlists all the possible next-hops for forwarding interest packets to the provider in FIB [41]. The routing path defines the stretch encountered while completing the request from the publisher to the consumer as illustrated in Figure 1. The path with the lesser number of hops encounters lesser stretch [42]. Routing optimization leads to the efficient path for content retrieval. Hence, optimal routing path between consumer and publisher leads to a reduction in end-to-end latency [11]. Therefore, we have proposed a strategy to select the optimal routing path i.e., intelligent stretch optimization (ISO) with lesser stretch for haptic-driven inter-personal communication for AR/VR applications.

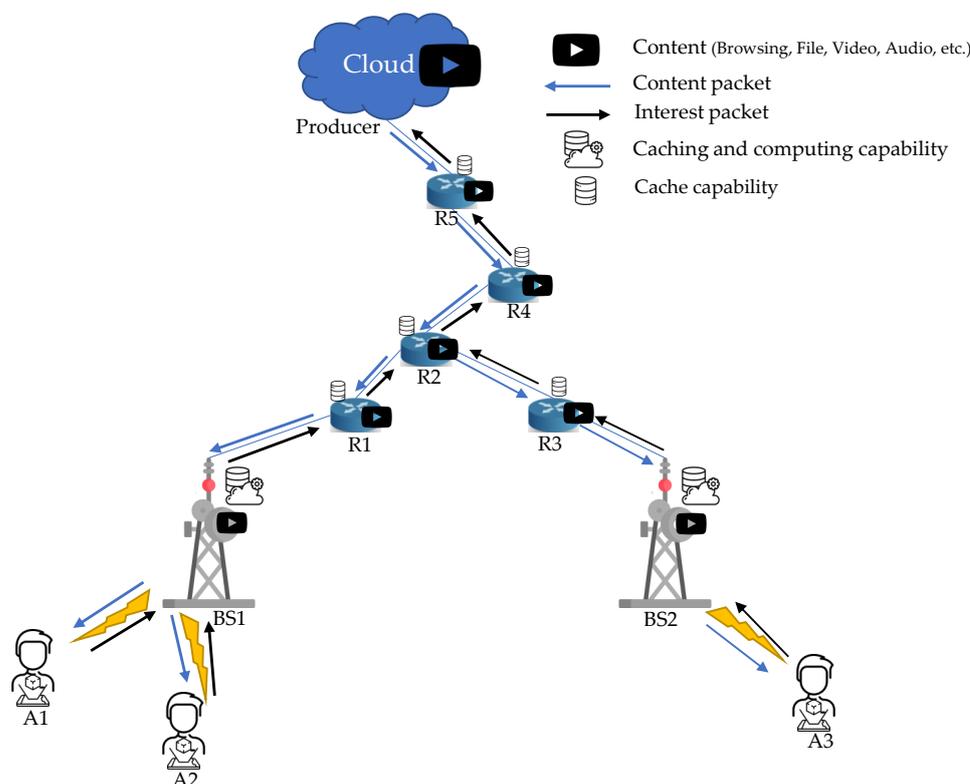


Figure 1. Overview of the basic ICN architecture for haptic-driven AR/VR applications.

The motivations of this paper are as follows:

- First, an intelligent model should be designed to evaluate stretch in ICN to obtain an optimal stretch reduction scheme for improving ICN-based 5G network performance.
- Second, reinforcement learning (RL) tends toward optimal stretch with an autonomous and intelligent decision based on previous experiences in the routing decision.

This paper proposes a RL-based ISO strategy in ICN for haptic-driven AR/VR applications in TI based on the above considerations. The contributions of this paper are as follows:

- The RL-based stretch reduction scheme has been proposed, promoting the optimal routing path selection with fewer hops. Integration of ICN, MEC, and 5G is employed to illustrate and evaluate the proposed routing scheme for stretch reduction for AR/VR users in TI.
- We propose an iterative algorithm to optimize the content retrieval path between the AR/VR user and the content provider. Furthermore, through rigorous simulation variations, the convergence and optimality of the proposed strategy is proved.
- We have pursued extensive simulation in this proposed strategy and proved that our RL-based stretch reduction strategy achieves significant routing path savings. Moreover, the proposed strategy opts for the path with the minimum stretch.

2. Related Work

Extensive work has been pursued in the field of ICN to improve the overall efficiency of network architecture. In [43–47], the researchers studied and investigated many cache strategies. Each cache strategy is discussed and evaluated in terms of a cache hit ratio (CHR), cache diversity (CD), and content redundancy (CR). The study in [44] presents a content placement strategy. The effectiveness of the proposed strategy is evaluated in terms of CHR, CD, and CR. The radio content router (RCR) caches the content according to between centrality router strategy, i.e., the router with most connections to other routers. The strategy is also evaluated in terms of CHR, CD, CR, and stretch based on the content placement. Similarly, in [45] authors have discussed and evaluated different probabilistic cache strategies in terms of the same parameters of CHR, CR, CD, and stretch. Furthermore, the authors presented a 5G-based ICN infrastructure content caching at BS, and user [37].

The authors proposed a joint optimizing the caching scheme, which discusses transcoding and routing decisions for adaptive video streaming over ICN [48]. They evaluated the work in terms of access delay and CHR. Most ICN work pursued primarily considers the model based on content providers, content caching, and content delivery. In [49] the authors have dealt with these three areas. The work mainly focuses on the content provider's selection, caching, and content delivery. The authors have proposed a cache-aware QoS routing scheme. This routing scheme discusses the proximal content available to the consumer. They have also devised the routing decision based on the appropriate FIB of the nearest proximal data providers. In [50] the Authors have investigated a new cache strategy by implementing the unsupervised learning-based proactive caching scheme for ICN networks. They introduced proactive data caching while creating clusters of the BS. Moreover, they evaluated their caching scheme based on CHR and user satisfaction ratio. In [10] authors have discussed the communication perspective regarding the VR users only keeping in view the IP protocol for communication. They worked in providing better QoS for haptic-driven VR users only. Moreover, our work improves the QoS by implementing the ISO and reduces the delay for better QoS for both AR and VR users.

In the perspective of edge network caching, ML method enables edge devices to monitor the network environment actively and intelligently by learning and predicting various network characteristics such as channel dynamics, traffic patterns, and content requests, allowing them to take proactive decisions that maximize a predefined objective such as QoS and CHR [51]. Currently approximately all research challenges are utilizing AI to obtain optimal results and performance. Similarly AI is also widely used for edge caching most of the current approaches use more than one machine learning methodology [52]. Unsupervised methods mainly clusters the users, contents, and BSs based on their location or access parameters [53,54]. Unsupervised learning approaches investigates structure or patterns in unlabeled data, such as user geolocation or content requests patterns. Given the network states, actions, and cost, RL algorithms are used to address the core problems of cache optimization [55]. The majority of published work deals with caching and main goal is to increase CHR [56]. CHR variations are intimately linked to all other caching targets [57]. When content is available in-network and not accessed via CDNs and backhaul lines, it directly affects content access latency, user QoE and data offload ratio [58].

Therefore almost no work is done towards intelligent and efficient routing scheme for ICN infrastructure. The caching strategies proposed so far are introduced for efficient content placement according to user’s request pattern. Different network topologies are used to verify the caching strategies. Each caching strategy provides different stretch results for requested content. Therefore, the caching strategies computes the stretch after the content discovery and placement. We proposed an ISO strategy that is applicable to different caching strategies.

3. Problem Statement

Currently, the cache strategies introduced so far have aimed to improve the overall efficiency of the ICN framework. This paper discusses optimal routing path selection in ICN for 5G and next-generation cellular networks to realize haptic-driven AR/VR applications. In ICN, most of the work evaluates the network infrastructure in terms of cache strategies. This paper focuses on providing better content validity while increasing the performance in terms of stretch. This paper proposes and enables intelligent routing in the whole network, mainly RCR and BS. According to our proposed ISO strategy, it is better to opt for the path with fewer hops for minimum latency and better reliability. Thus our ISO algorithm, after experiences, will learn the best optimal path with the lesser stretch between consumer and producer with the main focus for AR/VR applications in TI. After the content has been identified at a certain router in RCR, the RL agent will learn the best optimal path with less stretch. Let us suppose that the content is present in the router H_9 of RCR in Figure 2. As it can be seen from Figure 2, the content has been requested by the 2 AR and 1 VR users. Our proposed algorithm will learn the optimal path with lesser stretch for better reliability, QoT, QoE, and QoS for haptic-driven AR/VR applications H_1 to H_9 .

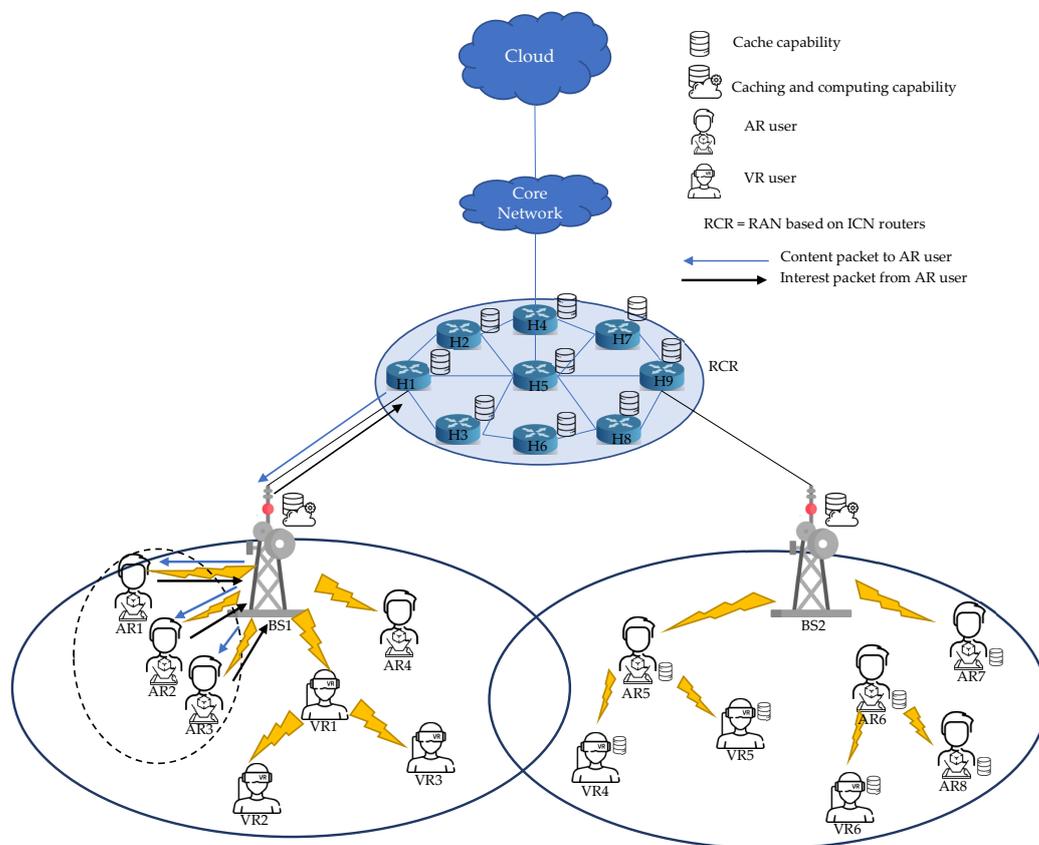


Figure 2. RL-based ISO strategy over ICN for haptic-driven AR/VR applications.

4. System Model

4.1. Methodology

This section describes the proposed ISO strategy for future ICN networks in TI. The proposed strategy achieves maximum gain when the stretch is least that increases the QoT, QoS, and QoE. Thus, it only discusses the stretch reduction via following the optimal path from consumer to producer and vice versa. Conclusively, we proposed a reinforcement learning-based model, and the primary goal is to reduce stretch for the ICN framework. Firstly, we formulate the markov decision process (MDP) of our system and then solve it with the help of Q-learning. We have described RL, Q-learning, and MDP, which we have used as a learning approach for our problem statement. Moreover, we describe our system model in detail while considering a case scenario of AR/VR application to understand our approach better.

Our ISO algorithm mainly focuses on selecting the optimal path, i.e., the path with fewer hops or ICN routers from RCR to BS. As RL learns from the experience, it will converge to a better policy after iterations. Thus, proposed ISO algorithm aims to reach the optimal routing policy after learning from different routing paths. Our scenario has the following entities:

- AR User/ VR User
- A content router in RCR having the contents that are of interest of a user
- BS

Our proposed system model has been shown in Figure 3. We supposed that router H_1 has the content according to the user connected with router H_9 . We have observed the mesh network of routers in RCR. All the routers H_1 - H_9 are connected just like a mesh network. Consumers are associated with BS1. A mesh topology connects almost every routing device. The total number of hops encountered in each step by the agent from its present or current hope H_1 to terminal hope H_9 helps to compute the stretch reward S_{ij} .

4.2. Reinforcement Learning

In RL, a node learns to take actions and then maps possible outcomes or situations for the agent's actions. Generally, the node or device does not have the knowledge about what actions to perform. The RL agent has to discover the best reward by exploring the possibilities. RL has some preliminary elements like agent and environment. Besides that, it has some policy that acts as a strategy. It characterizes the states of the respective environment to take action relatively. The other elements include reward, value function, or environment model. The reward is the parameter that helps the policy to learn it better and calculate better policy. The value function is the sum of rewards encountered in the whole episode. The environment model shows how the environment behaves and what the conditions are. The overview of our RL-based approach for ICN networks for stretch reduction is illustrated in Figure 3.

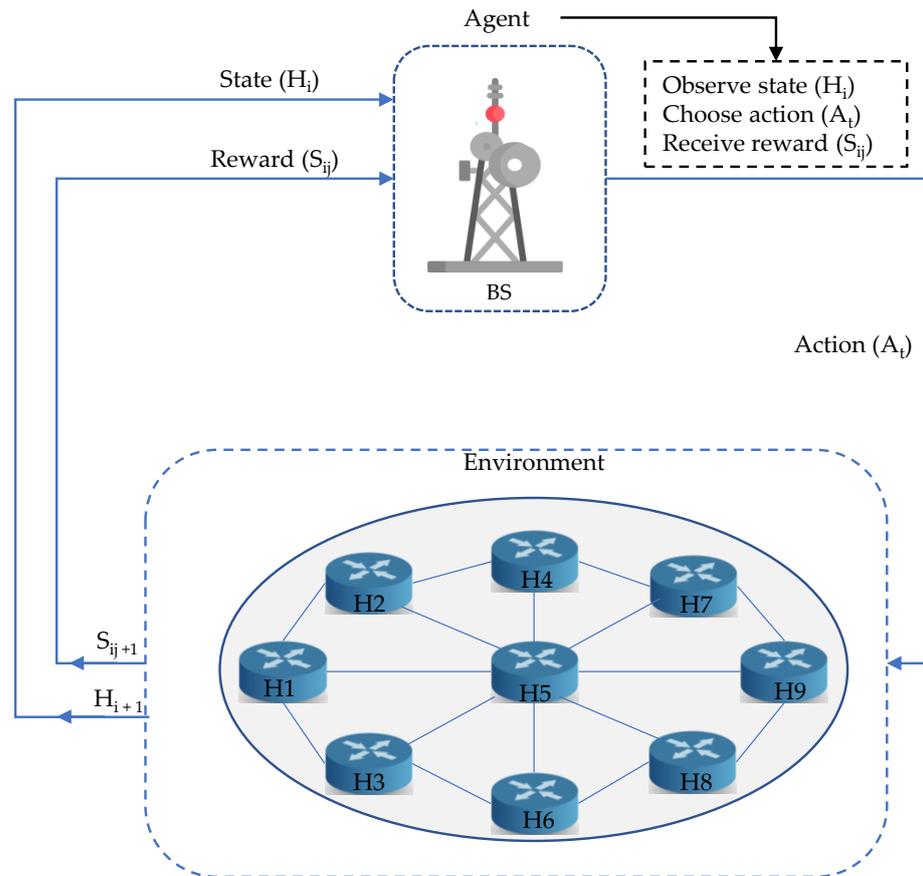


Figure 3. Overview of RL-based stretch reduction strategy.

4.2.1. Markov Decision Process

The process in which the agent observes the environment’s output in parameters like a reward, next state, and then what action would be taken next is the MDP. We have illustrated the overview of MDP in Figure 4. The Figure 4 shows that when state jumps from initial or current state H_i to next state H_{i+1} it has taken some action A_i . The action taken will lead to the reward S_{ij} . Reward is actually the routing cost in our case and state space in our scenario are routers H_1 to H_9 .

$$P[H_{i+1}|H_i] = P[H_{i+1}|H_1, \dots, H_i] \tag{1}$$

The MDP for our router H_i at time t can be described as follows:

- The router H_i senses its current state and obtains the current state.
- Based on current state of router H_i , agent selects action A_i at time t_i .
- Based on H_i the environment makes a transition to new state H_{i+1} and get a reward S_{ij} .
- That reward S_{ij} is feedback to the router and the process continues to next states with the same process.

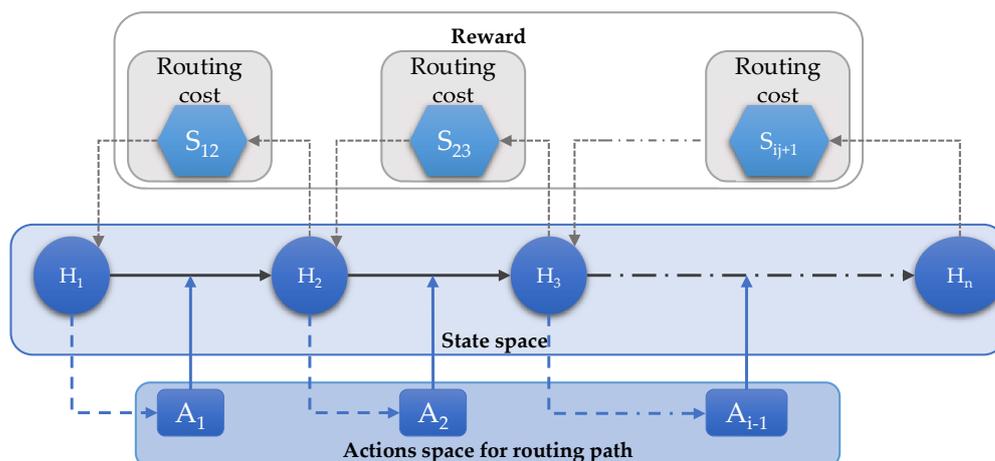


Figure 4. Framework for the stretch reduction strategy in ICN: (1) BSs estimates the state based on received observations over time t , H_i ; (2) following the policy the system select the optimal action A_i ; (3) and finally system receives reward S_{ij} correspond to the selected action.

4.2.2. Q-Learning

Q-learning is an off-policy RL algorithm to solve the MDP of the proposed strategy. We modeled a stretch reduction problem as an MDP shown in Figure 4 and proposed a Q-learning-based stretch reduction strategy.

First, we construct an MDP based on the stretch for mesh network topology. The MDP model can be defined by H, A, S . The state space H_n is the set of possible states in our network topology from router H_1 to next router. A_i is the set of actions that denotes which what are the required actions in our topology. S_{ij} is the reward function that has the feedback of reward of what action has been taken by the agent. The goal is to minimize the stretch, and reward shows the quality of step taken by the agent. To estimate the optimal value function, the optimal Q-value is selected for router H_i is defined as:

$$Q_{i+1}(s_i, a_i) = (1 - \alpha)Q_i(s_i, a_i) + \alpha \Delta Q_i(s_i, a_i) \tag{2}$$

$$\Delta Q(s_i, a_i) = r_i + \gamma(\max_a Q(s_{i+1}, a) - Q(s_i, a_i)) \tag{3}$$

We have used the Q-learning-based RL approach; therefore, it is mandatory to declare the environment with some states, actions, policy, and reward. Therefore, $H_n = [H_1, H_2, H_3, \dots, H_n]$ is state space in our infrastructure. Where $n = 9$ in our scenario. Similarly, the action space is defined as $A = [A_1, A_2, A_3, A_4, A_5]$. The action space defines the possible hop the content should follow to reach the consumer based on taking the policy. As discussed, the higher the reward minimum will be our stretch. Thus the higher reward declares the minimum stretch in the whole topology.

$$S_{ij} = 1/G_{ij} \tag{4}$$

where G_{ij} is the number of routers or hops traveled by the interest packet. Therefore, the inverse of G_{ij} help to compute stretch S_{ij} between consumer and publisher.

For routing decisions in given network topology, an agent searches for an optimal routing path with less number of routers from the BS to final router, where the content is identified with the help of PIT and FIB. The proposed ISO algorithm is discussed and illustrated in Figure 5 and Algorithm 1.

Algorithm 1: Pseudo-code for proposed ISO strategy.

Procedure: RL-BASED STRETCH REDUCTION

$H_i \leftarrow$ Set of all routers

$A_i \leftarrow$ Action space for routing

$S_{ij} \leftarrow$ Reward space

$Q(s, a) \leftarrow$ Q-table record for every episode

$Q^*(s, a) \leftarrow$ Optimal Q-value function

$\alpha \leftarrow$ Learning rate

$\gamma \leftarrow$ Discounted factor

$\epsilon \leftarrow$ Exploration/exploitation factor

Input: Initialization of $H_i, A_i, Q(s, a), \alpha, \gamma, \epsilon$

Output: $Q^*(s, a)$ for shortest path selection

Initialization of $(Q(s, a))$ with zeros

for $i = 0, \dots, T - 1$ **do**

 Observe current intersection state $s_t = H_i$;

 The agent selects action:

$a_t \leftarrow \operatorname{argmax} q(s_t, a_t)$ with probability $1 - \epsilon$

 The agent executes action A_n

 The agent calculates reward;

$r(t) = S_{ij}$ and

 Next state s_{t+1} ;

 Choose nearest s_{t+1} router based on (a_{t+1})

 Calculate reward $r(t): 1/G_{i,j}$

 Calculate Q-value and save it into Q-table

$Q(s_t, a_t) = Q(s_t, a_t) + \alpha(r_t + \gamma \times \max_a Q(s_{t+1}, a) - Q(s_t, a_t))$

 Update routing table based on Q-table

end

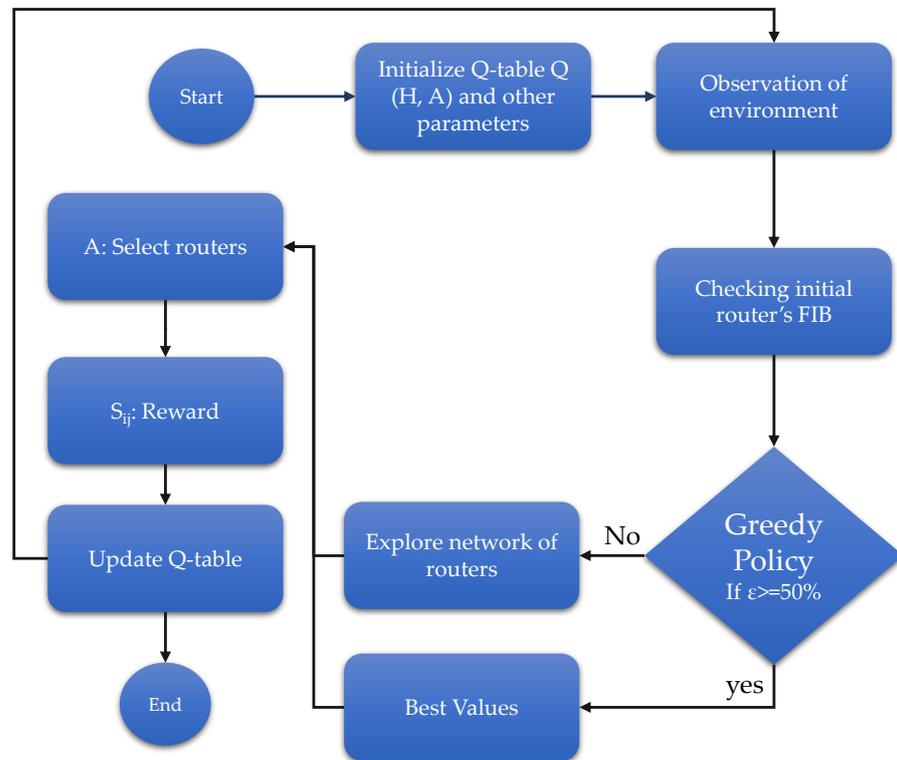


Figure 5. Q-learning algorithm for ISO strategy.

5. Performance Evaluation

We have built our simulation environment using MATLAB 2020a and consider nine content routers illustrated in Figure 2. Our proposed ISO strategy utilizes the mesh network of routers as a core network for 5G and next-generation cellular networks. We have compared our proposed strategy with random routing algorithm and history aware routing protocol (HAPR). In random algorithm the agent randomly takes the action for the routing path and there is no consistency in the routing path. Furthermore, HAPR follows the first path taken. Therefore, as the name suggests it tends to follow the same path for content retrieval and interest forwarding once the successful routing path is completed in the previous iteration. In our model AR/VR users are connected to BS1 and BS2, and the router H_9 has the desired content in its cached memory, according to consumer AR1. Our simulation results illustrate that the RL-based network will help the consumer retrieve the content from that path with less stretch that will help get a better data rate and lesser latency as compared to random and HAPR. Table 2 illustrates the default parameter values of simulation environment.

Table 2. Simulation parameters.

Parameters	Values
Simulation environment	Matlab 2020a
Network topology	Mesh network of routers
Number of AR/VR user requesting	1 at a time
RCR nodes in core network	$n = 9$
Link delay	10 ms
Routing decision	MDP (A_i)
Learning rate (α)	0.1, 0.5, 0.9
Discounted factor (γ)	0.1, 0.5, 0.9
Exploitation/exploration factor (ϵ)	0.5
Action space	[1,5]
State space	[1,9]

This section will evaluate the proposed ISO strategy in ICN systems for optimal routing path that lowers the E2E for haptic-driven AR/VR applications in TI. We computed the stretch as the performance metric for the routing path of the agent. The stretch reward is maximum for the least number of hops. Meanwhile, the proposed ISO algorithm opts for a better routing path for the least stretch. We kept different parameters of the learning rate (α) and discounted factor (γ) while keeping the exploration/exploitation factor (ϵ) constant. The learning speed of an agent is controlled and supervised by using α . Similarly, the learning rate γ controls the effect of past values on the current value or transition state. While ϵ factor controls the exploration or exploitation of the agent. On the basis of the learning rate, discounted factor, and epsilon, we have illustrated our model in terms of accumulated reward, instant episode, the average stretch of agent, and total hops covered in each episode.

We have kept the value of ϵ at 0.5 in all of our simulation results. Therefore, it means that the model explores and exploits 50%. We have computed accumulated reward as the agent covers the routing path to content router where content is placed. in Figure 6a, we have varied the learning rate from 0.1–0.9 and illustrated our agent's convergence at three different values of learning rate. The agent converges at the 121st iteration when the alpha value is 0.1, but still, the convergence is not complete, as evident from successive episodes. In comparison to the value of α at 0.5, the results are better than α value at 0.1 as it convergence at 117th episode, and hence a little better reward is attained. However, at the value of α at 0.9, convergence has been observed at the 128th episode. This shows that α value at 0.5, the proposed algorithm achieves better and optimal reward with the value of γ at 0.1.

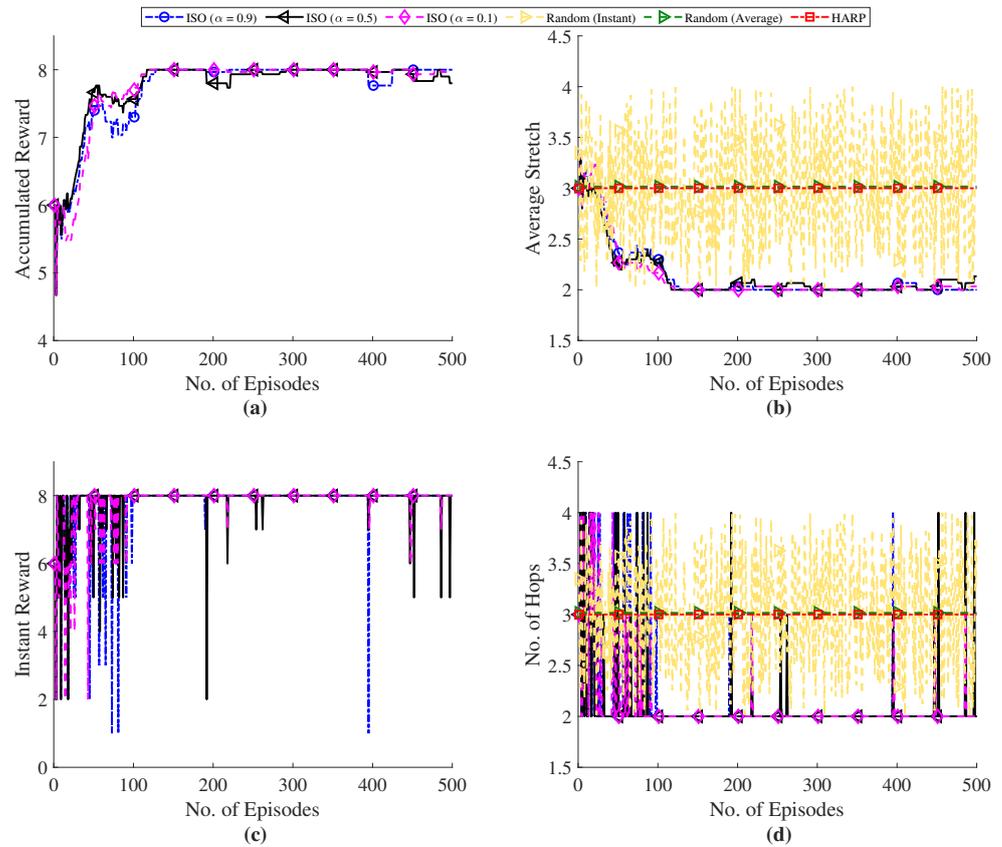


Figure 6. Stretch reduction and routing optimization progress of Q-learning agent at $\gamma = 0.1$: (a) Accumulated reward gain; (b) Average stretch gain; (c) Reward earned per episode; (d) Total hops/routers covered per episode.

Furthermore, the average stretch has been evaluated and illustrated in Figure 6b. We have illustrated the random routing strategy and HARP and compared them with our proposed strategy. It describes the number of hops traveled between the consumer and the producer. It manifests that all the initial stages of α value at 0.1, 0.5, and 0.9, the average number of hops count between consumer and producer varies between 3 and 4. Therefore, it means that at the initial stage, the agent is learning. After some experiences, because of the exploration/exploitation factor, the average number of hops count decreases gradually between consumer and producer. Moreover, at the alpha value of 0.5, the system is much more converged and optimal as compared to the rest of alpha 0.1 and 0.9. Moreover, The results clearly shows that the proposed strategy clearly converges towards the optimal stretch values and lesser number of hops as compared to random and HARP strategy. Thus our proposed strategy outperforms in comparison to both routing strategies i.e., random and HARP.

Moreover, Figure 6c illustrates the instant reward incurred in each episode the agent has taken. We have illustrated the reward gained by the agent or the cost value reward attained by the agent when the agent has followed a certain routing path. It is observed that even after the agent has convergence to the optimal path, there is still some point where the agent has shown the reward of 2, 5, and 7. This is due to the fact that we have kept the ϵ value at 0.5, which leads the agent to 50% explore and 50% exploit. Figure 6d illustrates the total hops covered by an agent from consumer to producer in one episode by an agent. Thus, Figure 6b illustrates the average stretch incurred while the path has been opted by the agent from consumer to producer. Figure 6d shows the hop count incurred in one episode. Similar to Figure 6c, it can be seen that even after the convergence to

optimality, there are still some spikes at 4, 3 because of the reinforcement learning nature of exploration and exploitation.

Figure 7a–d has been evaluated and illustrated by changing γ value from 0.1 to 0.5. The performance of the agent is also evaluated γ value of 0.5 with different variants of learning rate i.e., 0.1, 0.5 and 0.9. Figure 7a computes and illustrates the behavior of agent with performance metric of the accumulated reward of routing path between agent and final content router. The γ value is increased to 0.5 and accumulated reward is evaluated at α values 0.1, 0.5, and 0.9. in Figure 7a more optimal and better trend at α value 0.5 is observed. Similarly, Figure 7b illustrates the average stretch between the consumer and publisher when the content has been identified at content router.

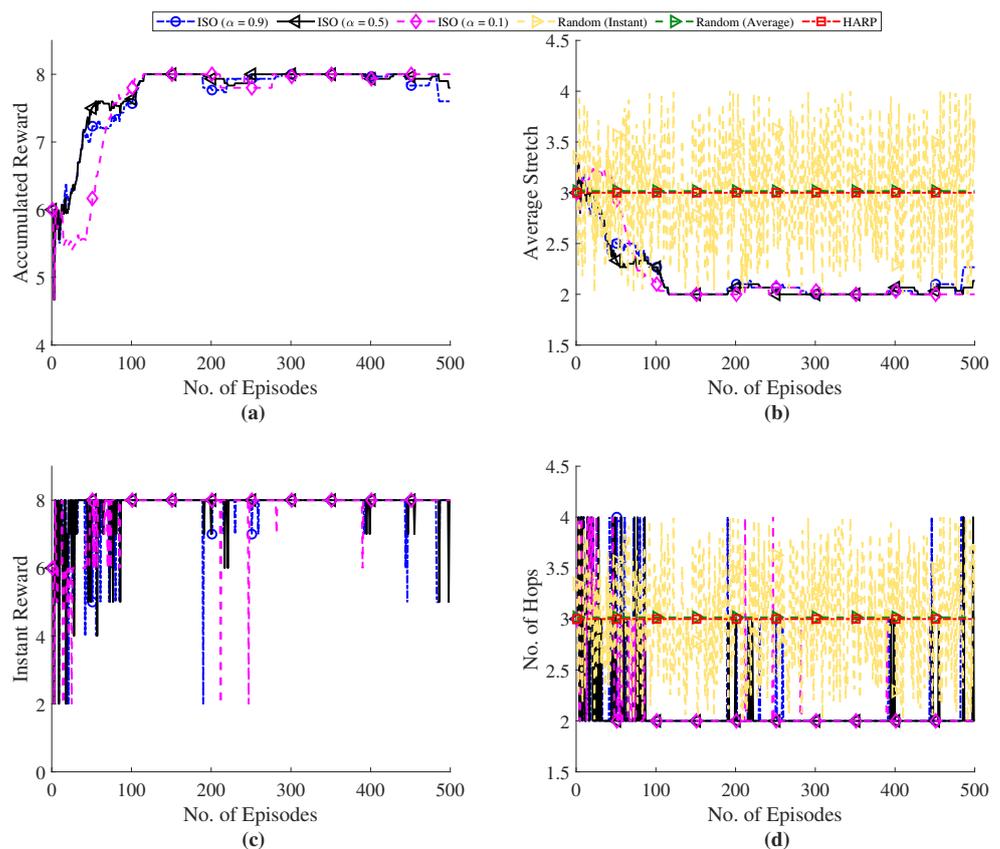


Figure 7. Stretch reduction and routing optimization progress of ISO at $\gamma = 0.5$: (a) Accumulated reward gain; (b) Average stretch gain; (c) Reward earned per episode; (d) Total hops/routers covered per episode.

Similarly, we have computed and evaluated the performance of our model with other routing strategies i.e., random and HARP. In our proposed algorithm the stretch incurred at one episode when the agent has completed one episode is illustrated in Figure 7c. It is also illustrated at different α values of 0.1, 0.5, and 0.9 when γ value is kept constant. The average steps show the corresponding hop count while st in Figure 7a, it illustrates the average stretch. Therefore, Figure 7a,b are describing each other for example, when the average reward is eight at episode 200, then the number of hops taken by the agent from consumer to the producer is 2. Therefore, the agent has been passed two routers while choosing the optimal path. Similarly, Figure 7c illustrates instant reward incurred at each episode the consumer has requested the content and Figure 7d illustrates the total hops the interest packet has traveled. At the same time, the discounted factor is increased to 0.5 and observing the agent’s behavior at different α values.

Thus we have observed that while changing the γ value to 0.5 and at α value of 0.5 we have gained the optimality with minimum stretch described and illustrated in Figure 7a–d

with less overall average stretch, average hop count, stretch per episode, and hop count at each step, respectively.

In Figure 8a–d we have observed the results while the γ is 0.9 with α varying to 0.1, 0.5 and 0.9. Figure 8a evaluates and illustrates the behavior of the agent when the γ value is increased to 0.5. It illustrates that the average stretch at α value of 0.5 outperforms in comparison to other α values of 0.1 and 0.9 and provides optimal and better stretch.

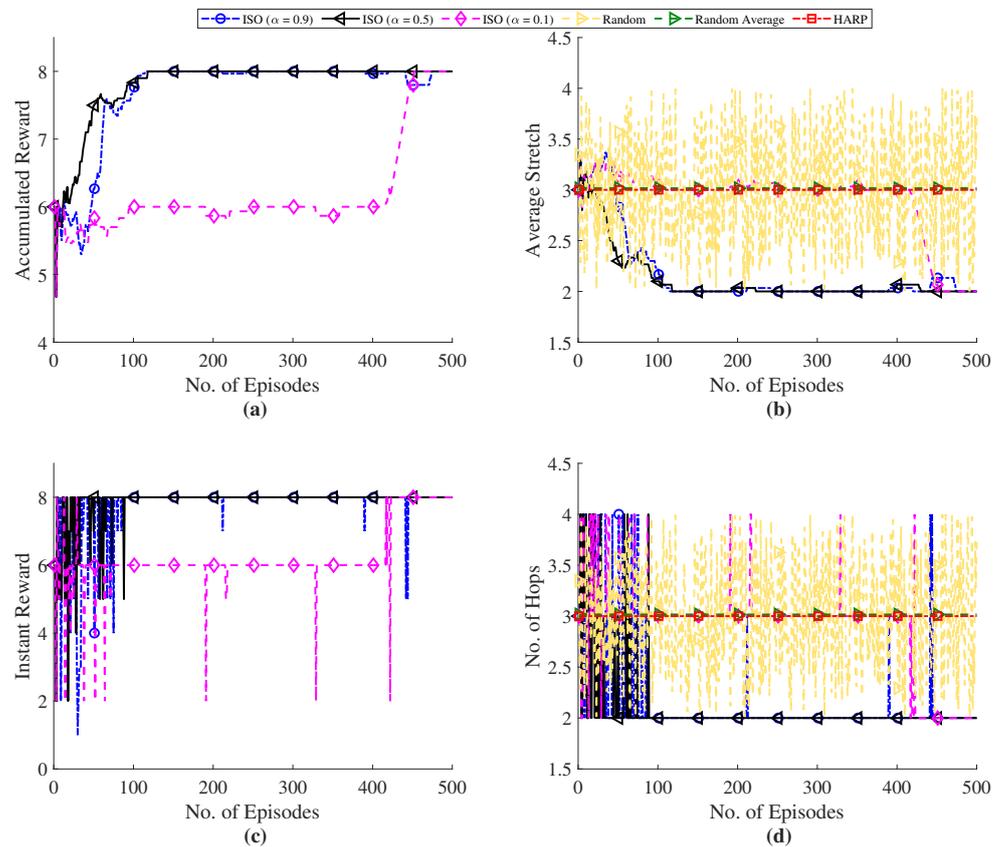


Figure 8. Stretch reduction and routing optimization progress of Q-learning agent at $\gamma = 0.9$: (a) Accumulated reward gain; (b) Average stretch gain; (c) Reward earned per episode; (d) Total hops/routers covered per episode.

Moreover, the average hop count between the AR/VR user and content router where the content has been identified is illustrated in Figure 8b. In Figure 8c we have illustrated the stretch incurred at one episode when the agent has completed one episode. We have also illustrated at different α values of 0.1, 0.5, and 0.9 when γ value is constant at 0.9. The average steps shows the corresponding hop count while Figure 8a illustrates the average stretch.

Therefore, Figure 8a,b are describing each other for example, when the average reward is eight at the 150th iteration, then the number of hops taken by the agent from consumer to the producer is 2. Therefore, the agent has been passed two routers while choosing the optimal path. In Figure 8c, we have illustrated that instant reward incurred at each episode the consumer has requested the content, and Figure 8d represents the total hops the interest packet has traveled. At the same time, the discounted factor is increased to 0.5 and observing the agent’s behavior at different α values. Thus we have observed that while changing the γ value to 0.5. At α value of 0.5, we have obtained the optimality i.e., the minimum stretch described as illustrated in Figure 8a–d with less overall average stretch, average hop count, stretch per episode, and hop count at each step, respectively. We have compared our results while changing the γ value to 0.9.

The whole performance of the RL agent has been summarized in Table 3. The convergence and optimal routing path attained at a specific episode given the corresponding values of α and γ of an agent are recorded. It can be observed that when $\gamma = 0.5$, we have achieved a better and faster optimal routing path.

Table 3. Optimal stretch comparison of learning agent.

Discounted Factor (γ)	Learning Rate (α)	Convergence Attained at Respective Episode
0.1	0.1	121
	0.5	117
	0.9	128
0.5	0.1	115
	0.5	116
	0.9	115
0.9	0.1	452
	0.5	118
	0.9	115

We have also illustrated the link delay of our proposed strategy in comparison with the random routing and HARP in Figure 9. We have kept the link delay between two content routers of 10 ms. Therefore, as our proposed algorithm converges towards optimal stretch in comparison to random routing and HARP. The average of random routing evaluated to 30.163 ms. Similarly, the HARP is converging at 30 ms. Furthermore, the proposed strategy converges to 20 ms. Therefore, our proposed strategy is 33.69% efficient and better in performance than that of random routing. Similarly, the proposed strategy is 33.33% good in performance in comparison to HARP.

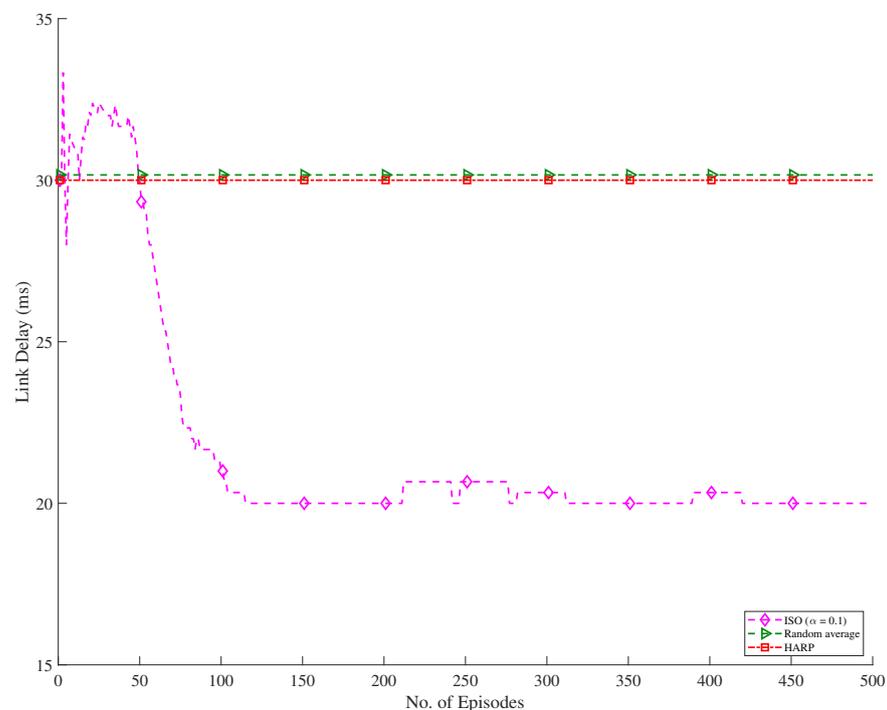


Figure 9. ISO link delay at $\gamma = 0.5$ and comparison with random routing and HARP.

6. Conclusions

In this paper, we have proposed and evaluated the use of ISO as a routing path optimization technique for the stretch reduction in the ICN framework in 5G and next-generation TI. We tested our simulation using Q-learning. We evaluated our results for different reinforcement learning parameters like α , γ , and ϵ . We have compared the average

and instantaneous results with three experiments of γ , and hence the optimal routing path selection in ICN using Q-learning algorithm proves to be efficient. The path with the lesser stretch gives the higher reward. Therefore, the performance and efficiency of the system is greatly improved. We designed the MDP model for our problem statement and solved that MDP with the help of Q-learning. The algorithm decides the optimal path between consumer and publisher based on previous experiences. We have evaluated our algorithm to illustrate the efficiency of our proposed algorithm towards stretch reduction. The proposed algorithm discusses the identification of requested content for AR/VR users at the certain content router in RCR and retrieves the content from the certain content router back to the AR/VR user. Moreover, we evaluated our performance and compared our proposed ISO strategy with random routing algorithm and HARP. The best optimal path lead to better QoS, QoT, and QoE for the proposed infrastructure, for haptic-driven applications of TI. Our Results show that our ISO algorithm converges to the optimal stretch within a lesser number of episodes with less link delay and is 33.69% better and 33.33% better than random routing and HARP, respectively.

Author Contributions: Methodology, H.A., R.A. and A.H.; Software and coding, H.A. and M.Z.I.; Experimentation and formal analysis, H.A., A.H. and R.A.; Writing—original draft preparation, H.A. and A.H.; Writing—review and editing, H.A., A.H., M.Z.I. and R.A.; Supervision, H.K.; Funding acquisition, A.H. and H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (Nos. 2019R1A4A1023746, 2019R1F1A1060799) and the Strengthening R&D Capability Program of Sejong University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: Authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

RL	Reinforcement Learning
MDP	Markov Decision Process
ICN	Information centric networking
MEC	Multi-Access Edge Computing
TI	Tactile Internet
AR	Augmented Reality
VR	Virtual Reality
URLLC	Ultra Reliable Low Latency Communication
eMBB	enhanced Mobile Broadband
QoS	Quality of Service
QoE	Quality of Experience
QoT	Quality of Tasks
5G	Fifth-Generation
RAN	Radio Access Network
BS	base station
RCR	RAN-based Content Router
PIT	pending interest table
FIB	forward interest base
CS	Content Store
CHR	Cache hit ratio

References

1. Bhalla, A.V.; Bhalla, M.R. Article: Generations of Mobile Wireless Technology: A Survey. *Int. J. Comput. Appl.* **2010**, *5*, 26–32.
2. Gui, G.; Liu, M.; Tang, F.; Kato, N.; Adachi, F. 6G: Opening New Horizons for Integration of Comfort, Security, and Intelligence. *IEEE Wirel. Commun.* **2020**, *27*, 126–132. [[CrossRef](#)]
3. Shafique, K.; Khawaja, B.A.; Sabir, F.; Qazi, S.; Mustaqim, M. Internet of Things (IoT) for Next-Generation Smart Systems: A Review of Current Challenges, Future Trends and Prospects for Emerging 5G-IoT Scenarios. *IEEE Access* **2020**, *8*, 23022–23040. [[CrossRef](#)]
4. Li, Q.C.; Niu, H.; Papathanassiou, A.T.; Wu, G. 5G Network Capacity: Key Elements and Technologies. *IEEE Veh. Technol. Mag.* **2014**, *9*, 71–78. [[CrossRef](#)]
5. Tullberg, H.; Popovski, P.; Li, Z.; Uusitalo, M.A.; Høglund, A.; Bulakci, O.; Fallgren, M.; Monserrat, J.F. The METIS 5G system concept: Meeting the 5G requirements. *IEEE Commun. Mag.* **2016**, *54*, 132–139. [[CrossRef](#)]
6. Lee, W.; Suh, E.S.; Kwak, W.Y.; Han, H. Comparative Analysis of 5G Mobile Communication Network Architectures. *Appl. Sci.* **2020**, *10*, 2478. [[CrossRef](#)]
7. Holland, O.; Steinbach, E.; Prasad, R.V.; Liu, Q.; Dawy, Z.; Aijaz, A.; Pappas, N.; Chandra, K.; Rao, V.S.; Oteafy, S.; et al. The IEEE 1918.1 “Tactile Internet” Standards Working Group and its Standards. *Proc. IEEE* **2019**, *107*, 256–279. [[CrossRef](#)]
8. Nasrallah, A.; Thyagaturu, A.S.; Alharbi, Z.; Wang, C.; Shao, X.; Reisslein, M.; ElBakoury, H. Ultra-Low Latency (ULL) Networks: The IEEE TSN and IETF DetNet Standards and Related 5G ULL Research. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 88–145. [[CrossRef](#)]
9. Ali, R.; Zikria, Y.B.; Bashir, A.K.; Garg, S.; Kim, H.S. URLLC for 5G and Beyond: Requirements, Enabling Incumbent Technologies and Network Intelligence. *IEEE Access* **2021**, *9*, 67064–67095. [[CrossRef](#)]
10. Marshall, A.; Yap, K.M.; Yu, W. Providing QoS for networked peers in distributed haptic virtual environments. *Adv. Multimed.* **2008**, *2008*, 841590. [[CrossRef](#)]
11. Haider, A.; Rahman, M.; Khan, T.; Niaz, M.T.; Kim, H.S. Multi-Gigabit CO-OFDM System over SMF and MMF Links for 5G URLLC Backhaul Network. *CMC-Comput. Mater. Contin.* **2021**, *67*, 1747–1758. [[CrossRef](#)]
12. Ahmad, H.; Islam, M.Z.; Haider, A.; Ali, R.; Kim, H.S. Intelligent Stretch Reduction in Information-Centric Networking towards 5G-Tactile Internet realization. In Proceedings of the 2021 International Conference on Green and Human Information Technology (ICGHIT), Online, 13–15 January 2021; pp. 1–6.
13. McClellan, M.; Cervelló-Pastor, C.; Sallent, S. Deep Learning at the Mobile Edge: Opportunities for 5G Networks. *Appl. Sci.* **2020**, *10*, 4735. [[CrossRef](#)]
14. Shu, Y.; Zhu, F. An edge computing offloading mechanism for mobile peer sensing and network load weak balancing in 5G network. *J. Ambient Intell. Humaniz. Comput.* **2020**, *11*, 503–510. [[CrossRef](#)]
15. Parvez, I.; Rahmati, A.; Guvenc, I.; Sarwat, A.I.; Dai, H. A Survey on Low Latency Towards 5G: RAN, Core Network and Caching Solutions. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 3098–3130. [[CrossRef](#)]
16. Sevgican, S.; Turan, M.; Gökarslan, K.; Yilmaz, H.B.; Tugcu, T. Intelligent network data analytics function in 5G cellular networks using machine learning. *J. Commun. Netw.* **2020**, *22*, 269–280. [[CrossRef](#)]
17. Rehman, M.A.U.; Ullah, R.; Kim, B.S.; Nour, B.; Mastorakis, S. CCIC-WSN: An Architecture for Single Channel Cluster-based Information-Centric Wireless Sensor Networks. *IEEE Internet Things J.* **2021**, *8*, 7661–7675. [[CrossRef](#)]
18. Erol-Kantarci, M.; Sukhmani, S. Caching and computing at the edge for mobile augmented reality and virtual reality (AR/VR) in 5G. *Ad Hoc Netw.* **2018**, *223*, 169–177.
19. Feng, L.; Li, W.; Lin, Y.; Zhu, L.; Guo, S.; Zhen, Z. Joint Computation Offloading and URLLC Resource Allocation for Collaborative MEC Assisted Cellular-V2X Networks. *IEEE Access* **2020**, *8*, 24914–24926. [[CrossRef](#)]
20. Simsek, M.; Aijaz, A.; Dohler, M.; Sachs, J.; Fettweis, G. 5G-Enabled Tactile Internet. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 460–473. [[CrossRef](#)]
21. Bastug, E.; Bennis, M.; Medard, M.; Debbah, M. Toward Interconnected Virtual Reality: Opportunities, Challenges, and Enablers. *IEEE Commun. Mag.* **2017**, *55*, 110–117. [[CrossRef](#)]
22. Alshahrani, A.; Elgendy, I.A.; Muthanna, A.; Alghamdi, A.M.; Alshamrani, A. Efficient multi-player computation offloading for VR edge-cloud computing systems. *Appl. Sci.* **2020**, *10*, 5515. [[CrossRef](#)]
23. Meng, Y.; Naeem, M.A.; Almagrabi, A.O.; Ali, R.; Kim, H.S. Advancing the state of the fog computing to enable 5G network technologies. *Sensors* **2020**, *20*, 1754. [[CrossRef](#)] [[PubMed](#)]
24. Golrezaei, N.; Molisch, A.F.; Dimakis, A.G.; Caire, G. Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution. *IEEE Commun. Mag.* **2013**, *51*, 142–149. [[CrossRef](#)]
25. Sukhmani, S.; Sadeghi, M.; Erol-Kantarci, M.; El Saddik, A. Edge Caching and Computing in 5G for Mobile AR/VR and Tactile Internet. *IEEE Multimed.* **2019**, *26*, 21–30. [[CrossRef](#)]
26. Ebrahimzadeh, A.; Chowdhury, M.; Maier, M. Human-Agent-Robot Task Coordination in FiWi-Based Tactile Internet Infrastructures Using Context- and Self-Awareness. *IEEE Trans. Netw. Serv. Manag.* **2019**, *16*, 1127–1142. [[CrossRef](#)]
27. Erol-Kantarci, M. Cache-at-relay: Energy-efficient content placement for next-generation wireless relays. *Int. J. Netw. Manag.* **2015**, *25*, 454–470. [[CrossRef](#)]
28. Gregori, M.; Gómez-Vilardebó, J.; Matamoros, J.; Gündüz, D. Wireless Content Caching for Small Cell and D2D Networks. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 1222–1234. [[CrossRef](#)]

29. Meng, Y.; Naeem, M.A.; Sohail, M.; Bashir, A.K.; Ali, R.; Zikria, Y.B. Elastic caching solutions for content dissemination services of ip-based internet technologies prospective. *Multimed. Tools Appl.* **2020**, 1–26. [[CrossRef](#)]
30. Din, I.U.; Hassan, S.; Almogren, A.; Ayub, F.; Guizani, M. PUC: Packet update caching for energy efficient IoT-based information-centric networking. *Future Gener. Comput. Syst.* **2020**, *111*, 634–643. [[CrossRef](#)]
31. Ruan, J.; Xie, D. Networked VR: State of the Art, Solutions, and Challenges. *Electronics* **2021**, *10*, 166. [[CrossRef](#)]
32. Hasan, K.; Jeong, S.H. Efficient caching for data-driven IoT applications and fast content delivery with low latency in ICN. *Appl. Sci.* **2019**, *9*, 4730. [[CrossRef](#)]
33. Serhane, O.; Yahyaoui, K.; Nour, B.; Mounghla, H. A Survey of ICN Content Naming and In-network Caching in 5G and Beyond Networks. *IEEE Internet Things J.* **2020**, *8*, 4081–4104. [[CrossRef](#)]
34. Elayoubi, S.E.; Fallgren, M.; Spapis, P.; Zimmermann, G.; Martín-Sacristán, D.; Yang, C.; Jeux, S.; Agyapong, P.; Campoy, L.; Qi, Y.; et al. 5G service requirements and operational use cases: Analysis and METIS II vision. In Proceedings of the 2016 European Conference on Networks and Communications (EuCNC), Athens, Greece, 27–30 June 2016; pp. 158–162. [[CrossRef](#)]
35. Naeem, M.A.; Ali, R.; Kim, B.S.; Nor, S.A.; Hassan, S. A periodic caching strategy solution for the smart city in information-centric Internet of Things. *Sustainability* **2018**, *10*, 2576. [[CrossRef](#)]
36. Zafar, W.U.I.; Rehman, M.A.U.; Jabeen, F.; Kim, B. Context-Aware Naming and Forwarding in NDN-Based VANETs. *Sensors* **2021**, *14*, 4629. [[CrossRef](#)] [[PubMed](#)]
37. Ullah, R.; Rehman, M.A.U.; Naeem, M.A.; Kim, B.S.; Mastorakis, S. ICN with edge for 5G: Exploiting in-network caching in ICN-based edge computing for 5G networks. *Future Gener. Comput. Syst.* **2020**, *111*, 159–174. [[CrossRef](#)]
38. Chakareski, J. VR/AR immersive communication: Caching, edge computing, and transmission trade-offs. In Proceedings of the Workshop on Virtual Reality and Augmented Reality Network, Los Angeles, CA, USA, 25 August 2017; pp. 36–41.
39. Qiao, J.; He, Y.; Shen, X.S. Proactive Caching for Mobile Video Streaming in Millimeter Wave 5G Networks. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 7187–7198. [[CrossRef](#)]
40. Mun, J.H.; Lim, H. On Sharing an FIB Table in Named Data Networking. *Appl. Sci.* **2019**, *9*, 3178. [[CrossRef](#)]
41. Jung, K.; Nguyen, V.T.; Lee, J. BlocklyXR: An Interactive Extended Reality Toolkit for Digital Storytelling. *Appl. Sci.* **2021**, *11*, 1073. [[CrossRef](#)]
42. Kang, M.W.; Seo, D.Y.; Chung, Y.W. An Efficient Delay Tolerant Networks Routing Protocol for Information-Centric Networking. *Electronics* **2020**, *9*, 839. [[CrossRef](#)]
43. Naeem, M.A.; Ali, R.; Alazab, M.; Meng, Y.; Zikria, Y.B. Enabling the content dissemination through caching in the state-of-the-art sustainable information and communication technologies. *Sustain. Cities Soc.* **2020**, *61*, 102291. [[CrossRef](#)]
44. Naeem, M.A.; Rehman, M.A.U.; Ullah, R.; Kim, B.S. A comparative performance analysis of popularity-based caching strategies in named data networking. *IEEE Access* **2020**, *8*, 50057–50077. [[CrossRef](#)]
45. Meng, Y.; Naeem, M.A.; Ali, R.; Zikria, Y.B.; Kim, S.W. DCS: Distributed caching strategy at the edge of vehicular sensor networks in information-centric networking. *Sensors* **2019**, *19*, 4407. [[CrossRef](#)] [[PubMed](#)]
46. Naeem, M.A.; Nor, S.A.; Hassan, S.; Kim, B.S. Compound popular content caching strategy in named data networking. *Electronics* **2019**, *8*, 771. [[CrossRef](#)]
47. Hussaini, M.; Naeem, M.A.; Kim, B.S.; Majjama'a, I.S. Efficient producer mobility management model in Information-Centric Networking. *IEEE Access* **2019**, *7*, 42032–42051. [[CrossRef](#)]
48. Hu, H.; Jin, Y.; Wen, Y.; Westphal, C. Orchestrating caching, transcoding and request routing for adaptive video streaming over icn. *ACM Trans. Multimed. Comput. Commun. Appl. TOMM* **2019**, *15*, 1–23. [[CrossRef](#)]
49. Chuan, J.; Bai, B.; Wu, X.; Zhang, H. Optimizing Content Placement and Delivery in Wireless Distributed Cache Systems Through Belief Propagation. *IEEE Access* **2020**, *8*, 100684–100701. [[CrossRef](#)]
50. Ahmad, F.; Ahmad, A.; Hussain, I.; Uthansakul, P.; Khan, S. Cooperation Based Proactive Caching in Multi-Tier Cellular Networks. *Appl. Sci.* **2020**, *10*, 6145. [[CrossRef](#)]
51. Chen, M.; Challita, U.; Saad, W.; Yin, C.; Debbah, M. Artificial neural networks-based machine learning for wireless networks: A tutorial. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 3039–3071. [[CrossRef](#)]
52. Doan, K.N.; Van Nguyen, T.; Quek, T.Q.; Shin, H. Content-aware proactive caching for backhaul offloading in cellular network. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 3128–3140. [[CrossRef](#)]
53. Thar, K.; Tran, N.H.; Oo, T.Z.; Hong, C.S. DeepMEC: Mobile edge caching using deep learning. *IEEE Access* **2018**, *6*, 78260–78275. [[CrossRef](#)]
54. Ozcan, A.; Oguducu, S.G. Multivariate time series link prediction for evolving heterogeneous network. *Int. J. Inf. Technol. Decis. Mak.* **2019**, *18*, 241–286. [[CrossRef](#)]
55. Jiang, F.; Yuan, Z.; Sun, C.; Wang, J. Deep Q-learning-based content caching with update strategy for fog radio access networks. *IEEE Access* **2019**, *7*, 97505–97514. [[CrossRef](#)]
56. Zhang, C.; Ren, P.; Du, Q. Learning-to-rank based strategy for caching in wireless small cell networks. In Proceedings of the International Conference on Internet of Things as a Service, Xi'an, China, 17–18 November 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 111–119.

-
57. Lei, F.; Dai, Q.; Cai, J.; Zhao, H.; Liu, X.; Liu, Y. A proactive caching strategy based on deep Learning in EPC of 5G. In Proceedings of the International Conference on Brain Inspired Cognitive Systems, Xi'an, China, 7–8 July 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 738–747.
 58. Li, L.; Xu, Y.; Yin, J.; Liang, W.; Li, X.; Chen, W.; Han, Z. Deep reinforcement learning approaches for content caching in cache-enabled D2D networks. *IEEE Internet Things J.* **2019**, *7*, 544–557. [[CrossRef](#)]