Deep Reinforcement Learning for Handover-Aware MPTCP Congestion Control in Space-Ground Integrated Network of Railways

Jianpeng Xu and Bo Ai

Abstract

Space-ground integrated networks (SGINs) have been recently regarded as a promising way to provide resilient, dependable as well as efficient data transmission in the high-speed railway (HSR) scenario. Applying multipath transmission control protocol (MPTCP) to SGIN can realize data transmission simultaneously via terrestrial and satellite networks. However, since the existing congestion control (CC) mechanisms of MPTCP fail to distinguish between adverse influences (such as packet loss and/or round-trip time increase) caused by congestion and those caused by handovers, it suffers severe performance degradation in the HSR scenario where handover frequently occurs. In this article, we first present the SGIN oriented HSR (SGIN-HSR) with MPTCP. Then leveraging cross-layer information (i.e., reference signal received power), we design a novel cross-layer aided MPTCP CC mechanism targeted at SGIN-HSR based on deep reinforcement learning, which is referred to as HSR-CC, to alleviate performance degradation problems induced by handover. The experimental results show that HSR-CC significantly enhances the goodput and outperforms state-of-the-art MPTCP CC algorithms in SGIN-HSR environments where handover frequently occurs.

Introduction

Thanks to the emergence of many advanced technologies, such as artificial intelligence (AI), edge computing, and Internet of Things, smart railways are under high speed development. Nowadays, more and more countries are trying to build smart railways, where more and more passengers, goods, and high-speed trains (HSTs) are anticipated to be interconnected. In smart railway systems, operation mobility, safety, and dependability will be enhanced. In addition, passengers can also enjoy more cost efficient, eco-friendly as well as comfortable services [1, 2].

To accelerate construction of smart railways, it is significant to provide high data transmission rate and wide coverage in the high-speed railway (HSR) scenario. With the development of railways, geometric environments of railway transportation lines appear diverse and complicated. However, owing to the lack of terrestrial wireless network infrastructure in remote regions, it presents a big challenge to the construction of smart railways. Space-ground integrated networks (SGINs) are one mode expected to accede to sixth generation (6G), which provides an efficient way to settle communication issues in remote regions [1]. The authors in [3] have investigated that SGIN can improve transmission performance in the HSR. However, considering the restriction of single-path transmission control protocol (TCP), almost all works related to SGIN assume data can only be transferred over a single network at any given time; as a result, the capacity of multiple interfaces cannot be fully utilized [4]. Moreover, according to the measurement in HSR networks [5, 6], TCP will suffer from significant performance degradation, mainly due to frequent handover. To alleviate performance issues caused by frequent handovers in HSR networks, Hd-TCP [5], a congestion control (CC) strategy based on deep Q-network (DQN), was investigated with the help of observing the reference signal received power (RSRP) from the physical layer. Nevertheless, since HST experiences one handover approximately every 10 seconds, single TCP struggles to maintain service continuity as well as performance in HSR environments, causing a frustrating user experience. Facing these problems, multipath TCP (MPTCP) has appeared, which is a possible solution to ensure that data connections are resilient, dependable, and efficient in HSR networks [2], [4]. This is because it is scarcely possible that all the paths experience handovers simultaneously, which helps to maintain service continuity and performance. What's more, the authors in [4] have presented that the architecture of SGIN oriented HSR (SGIN-HSR) can better satisfy railway services quality requirements. Applying MPTCP to SGIN can realize multi-bearer communications, as a result, data can be transmitted simultaneously over the terrestrial and satellite networks [4].

Unfortunately, the authors of [6] measured the performance of MPTCP in HSR networks, concluding that though MPTCP is more robust to handover than TCP, frequent handovers also lead to MPTCP performance degradation. MPTCP CC

Digital Object Identifier:
10.1109/MWC.001.2100116

Jianpeng Xu and Bo Ai (corresponding author) are with Beijing Jiaotong University, and also with Frontiers Science Center for Smart High-speed Railway System. Bo Ai is also with Peng Cheng Laboratory and Zhengzhou University.

algorithms, such as LIA and OLIA [7], fail to distinguish between adverse influences (such as packet loss and/or round-trip time (RTT) increase) caused by congestion and those caused by handovers, and always refer to packet loss as congestion, therefore executing extremely aggressive CC when high packet loss events induced by handover happen. To differentiate between congestion and handover, cross-layer assisted MPTCP (CLA-MPTCP) [8], a ruled-based CC strategy, was investigated to improve service continuity during handover with the cross-layer information (i.e., RSRP) assistance. However, the dynamic properties of the HSR networks are so complex that achieving an effective control of the congestion window (CWND) adjustment by relying on a set of rules is completely unrealistic. In addition, CWND decisions of rule-based CC rely on precise network environment models, which are extremely challenging in the HSR networks. As a result, the rule-based MPTCP CC schemes, including RVeno [2], OLIA [7], and CLA-MPTCP [8], cannot adapt fast enough and perform well in the HSR networks.

In recent years, machine learning (ML), particularly deep reinforcement learning (DRL), hasl shown the advantages in adapting to complicated and highly-dynamic network environments [9]. Past research has demonstrated that the CC based on DRL can achieve much better performance than the rule-based CC mechanisms [9–11]. That is because the equipment with DRL-based CC mechanisms is able to learn the sound CC policy based on the past experiences and network environment observations, which is beneficial for adapting to the time-varying network. Owing to the fact that the railway lines are fixed, it is likely that the same train or other trains travelling on the same lines can use the previously acquired experiences. Therefore, applying DRL-based CC mechanism in HSR networks is beneficial for learning to adapt CC strategies directly from past experience through trial and error, instead of precise network models, showing better potential than the rule-based CC mechanism.

Nevertheless, there is a shortcoming in existing DRL-based CC mechanisms. The performance of the DRL model is largely decided by the network condition representation quality. In general, packet loss and/or RTT increment is the sign of congestion [5–7]. Existing mechanisms represent the network environment by employing a few characteristics from the perspective of the transport layer, such as RTT, which are not adequate to represent the network's state, especially the HSR networks with frequent handovers. It is undesirable to observe the results of interruptions in service continuity induced by handover in HSR networks. When each handover occurs, it will cause the fluctuations of HSR network environments, such as the sharp increase of RTT and packet loss rate (PLR) [5, 6, 12]. Misinterpretation of the fluctuations due to the handover in the HSR scenario as a sign of network congestion will cause poor performance. Ideally, we should investigate a more efficient DRL-based CC of the MPTCP algorithm. Based on the measurements in [5, 6, 12], it can be inferred that network conditions (i.e., RSRP, throughput, PLR and RTT) are highly related in anticipating when a handover is about to occur. Combining these metrics can accurately represent the state of network conditions. In this context, to the best of our knowledge, this article is the first work to investigate a novel cross-layer aided MPTCP CC mechanism targeted at SGIN-HSR by leveraging DRL, which is named as HSR-CC and can alleviate performance degradation problems induced by handover. A simple comparison between HSR-CC and recent CC mechanisms for MPTCP is summarized in Table 1. As the Table 1 shows, rule-based CC mechanisms are often intuitive with reasonable interpretations. In contrast, ML-based mechanisms are usually non-intuitive, which is hard to interpret how to make policy decisions. However, ML-based mechanisms are much more promising than the rule-based mechanisms in highly-dynamic network environments [9–11]. Compared with existing ML-based CC mechanisms of MPTCP, HSR-CC has the following advantages:

Algorithm	Year	Category	Cross-layer	Interpretation	Optimization objective
CLA-MPTCP [8]	2016	Rule-based	Yes	Intuitive	High throughput
DRL-CC [9]	2019	ML-based	No	Non-intuitive	High goodput
SmartCC [10]	2019	ML-based	No	Non-intuitive	High throughput, low jitter, low RTT
DQN-MPTCP [11]	2020	ML-based	No	Non-intuitive	High throughput
HSR-CC	2021	ML-based	Yes	Non-intuitive	High goodput

TABLE 1. A simple comparison between HSR-CC and recent CC mechanisms for MPTCP.

- Integrating the network characteristics of the transport layer and that of the physical layer is capable of representing network state more sufficiently.
- HSR-CC is handover-aware. The knowledge learned in RSRP fluctuations is beneficial for differentiating between congestion and handover.

The remainder of this article is structured as follows. We first present the MPTCP-empowered SGIN-HSR architecture in detail. Next, the necessity of a cross-layer aided CC mechanism is analyzed and we exploit DRL to develop a cross-layer aided CC method. After that, numerical results demonstrate that HSR-CC outperforms other state-of-the-art CC methods in goodput. We conclude the article and provide future research opportunities.

MPTCP-Empowered SGIN-HSR Architecture

Figure 1 shows the MPTCP-empowered SGIN-HSR architecture, which includes the principle architecture and protocol architecture. With the purpose of reducing the penetration losses, the onboard multi-homed router (MHR) is connected to the in-cabin wireless access point (AP) that is deployed on top of each carriage. The MHR can simultaneously communicate with the terrestrial and satellite networks, playing the role of forwarding data between the user equipments (UEs) on the train and the broadband wireless networks. Note that the MHRs are equipped with satellite and terrestrial network interfaces, such as third generation (3G), long term evolution (LTE), fifth generation (5G), and so on. The UEs in HST can communicate with each other directly instead of via base stations to improve service quality by the aid of a device-to-device (D2D) technique [1].

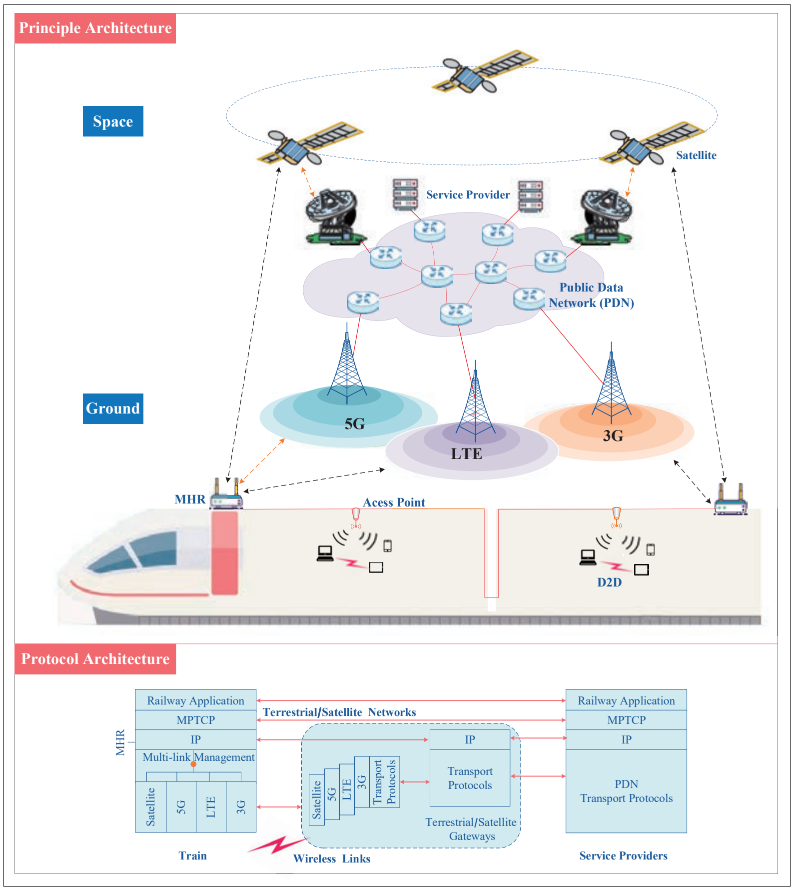

FIGURE 1. MPTCP-empowered SGIN-HSR Architecture.

Current railway communication is based on the Internet protocol (IP) technique. MPTCP does not make changes to the railway application layer by extending IP to multiple link environments. As shown in the Fig. 1, the MPTCP protocol layer resides on top of the multi-link management layer that supports all the available radio bearers. The principle protocol architecture shown in Fig. 1 is suitable for being applied in SGIN-HSR. In a word, compared to current railway communication systems, MPTCP-empowered SGIN-HSR architecture has the following advantages.

Bandwidth Aggregation: With the help of MPTCP, the MHR can simultaneously communicate with terrestrial and satellite networks, significantly increasing resource utilization as well as bandwidth utilization.

Resilience and Reliability: The MHR can transmit data via multiple paths, which is good for realizing seamless super-connectivity. Even if one of the paths fails, it can transmit data by choosing another path, increasing HSR network resilience and reliability.

Wide Coverage: Due to the satellite network's global coverage, MPTCP-empowered SGIN-HSR can achieve wide coverage.

Integration of Multiple Techniques: Multiple advanced techniques such as D2D, AI and so on can be integrated into the architecture.

Intelligent Collaboration: This architecture is good for the intelligent cooperation of various network service providers.

However, there is no doubt that existing MPTCP CC schemes suffer from severe performance degradation induced by frequent handovers. Next, we will analyze the challenges for CC in the MPTCP-empowered SGIN-HSR architecture.

Why is Cross-Layer Information Necessary

There is not much work focused on investigating MPTCP challenges during handover. In HSR networks, the HST experiences one handover approximately every 10 seconds [1]. Due to the frequent handovers, it is extremely difficult to maintain service continuity and consistency in HSR networks, resulting in bad influence on the passenger experience. For simplicity, we consider that the MHR can simultaneously transmit data via the satellite and LTE networks. The MHR keeps the connection to the satellite network, and experiences frequent handovers in the LTE networks. Consequently, we mainly analyze the fluctuations of LTE network environments when the handover occurs.

Handover Decision

Figure 2a illustrates the process of the event A3 handover decision algorithm. Note that due to its most stable performance, the event A3 handover decision algorithm is the most popular handover decision algorithm [12]. It was observed from Fig. 2a that LTE performs the event A3 handover process through monitoring the RSRP values of serving evolved Node B (SeNB) as well as target eNB (TeNB). The UE measures regularly the RSRP of SeNB and TeNB. Naturally, when the train is travelling away from the SeNB, the RSRP of SeNB decreases gradually. In contrast, the RSRP of TeNB gradually increases. When the RSRP of TeNB is higher SeNB and is above a certain threshold, handover entering condition (event A3's trigger point) is met.

To prevent the ping-pong handover phenomenon from happening, handover lasts from the time to trigger (TTT). Within the TTT period, once the measured RSRP value meets the handover cancelling criterion, handover will not happen. Nevertheless, in the case that the handover cancelling criterion is not met, the SeNB will be in the operation of handover based on transmitting a *Handover Request* signal to TeNB after TTT expires. The handover operating point represents the time point when the SeNB transmits a Handover Request signal. When it comes to handover operating point, handover is performed by exchanging control messages among the UE, SeNB and TeNB. Note that the RSRP that is served as an indispensable parameter of handover decision is transmitted through the HSR channel. Since the HSR channel is extremely time-varying, the variations of RSRP from the SeNB is rapidly dynamic, which is measured by [12]. As a result, rapid variations may be beyond the range of LTE systems, causing link failure associated with the SeNB before performing handover, which will lead to handover failure or delay.

Handover Operation

When it comes to handover operating point, the UE's connection to the SeNB is disconnected and the connection to the TeNB is established. During the HO operation stage, delay can be divided into three categories, T1, T2 and T3, which respectively express the pre-handover delay, handover interruption time and post-handover delay.

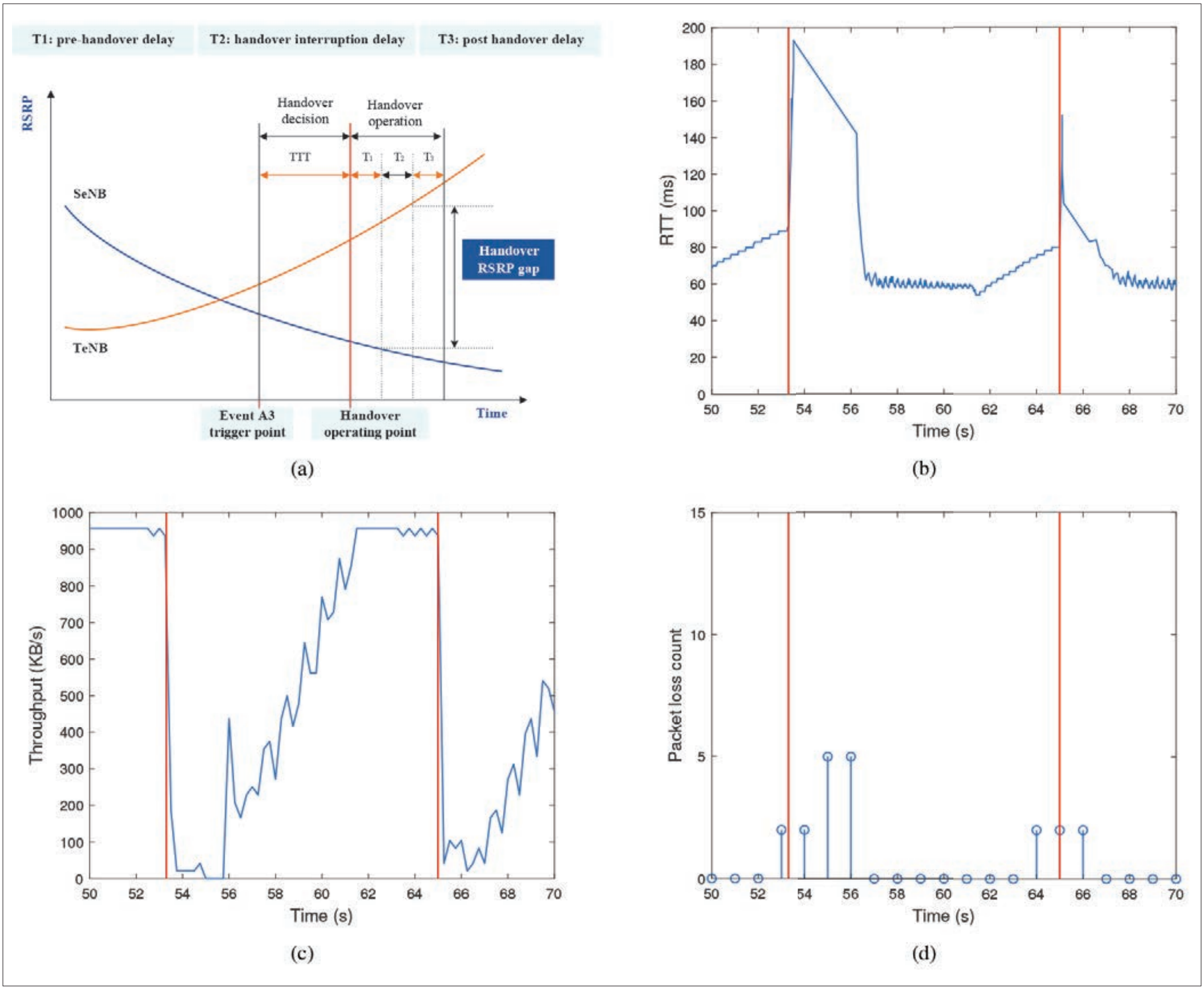

FIGURE 2. Performance metrics relevant to handovers: a) RSRP relevant to handovers; b) RTT relevant to handovers; c) Throughput relevant to handovers; b) Packet loss count relevant to handovers [5, 12].

T1 represents the time gap between the handover operating point and the SeNB disconnection. T2 represents the time gap between the SeNB disconnection and TeNB connection. T3 represents the signal processing time to complete the handover.

Handover Challenges for MPTCP

According to the measurement of [12], if the train is going at the rate of 50 km/h, the last measured RSRP value before performing handover is −58.25 dBm. Nevertheless, if the train is going at the rate of 300 km/h, the last measured RSRP value before performing handover is −77.38 dBm. The RSRP value from the SeNB is decreased with the increase of the train's speed, causing not only the high PLR, but also handover failure or delay. As a result, service continuity cannot be maintained. In addition, the rapid changes of RTT, PLR and throughput also are present in [5] and [6] during the handover process, which are similar to the rapid changes of RSRP.

In a word, when the handover happens, these phenomena including the sharp increase of RTT and PLR, the sharp decrease of RSRP and throughput will appear, shown in Fig. 2 [5, 12]. That is to say, MPTCP performance is unstable and unsatisfactory in HSR networks. The main reason is that MPTCP fails to distinguish between adverse influences (such as packet loss and/or RTT increase) caused by congestion and those caused by handovers. Note that among the subflows of MPTCP, the subflow with the lowest RTT (or least congestion level) is preferred to balance the load. Under the certain bandwidth, a larger CWND usually means a larger RTT. Generally speaking, when the congestion level of a chosen subflow increases, MPTCP will choose another subflow with less congestion [7]. Nevertheless, RTT increase and packet loss alone do not always represent an advisable indication of link congestion, because the fluctuations in the HSR scenario may be merely due to an imminent handover. Consequently, MPTCP may treat mistakenly depleting connection caused by handover as network congestion, leading to significant performance degradation. In other words, it is not conducive to the experience of the user in the way that a congested link is indicated by the time requested for updating RTT or packet loss. With the aid of cross-layer information, the problem can

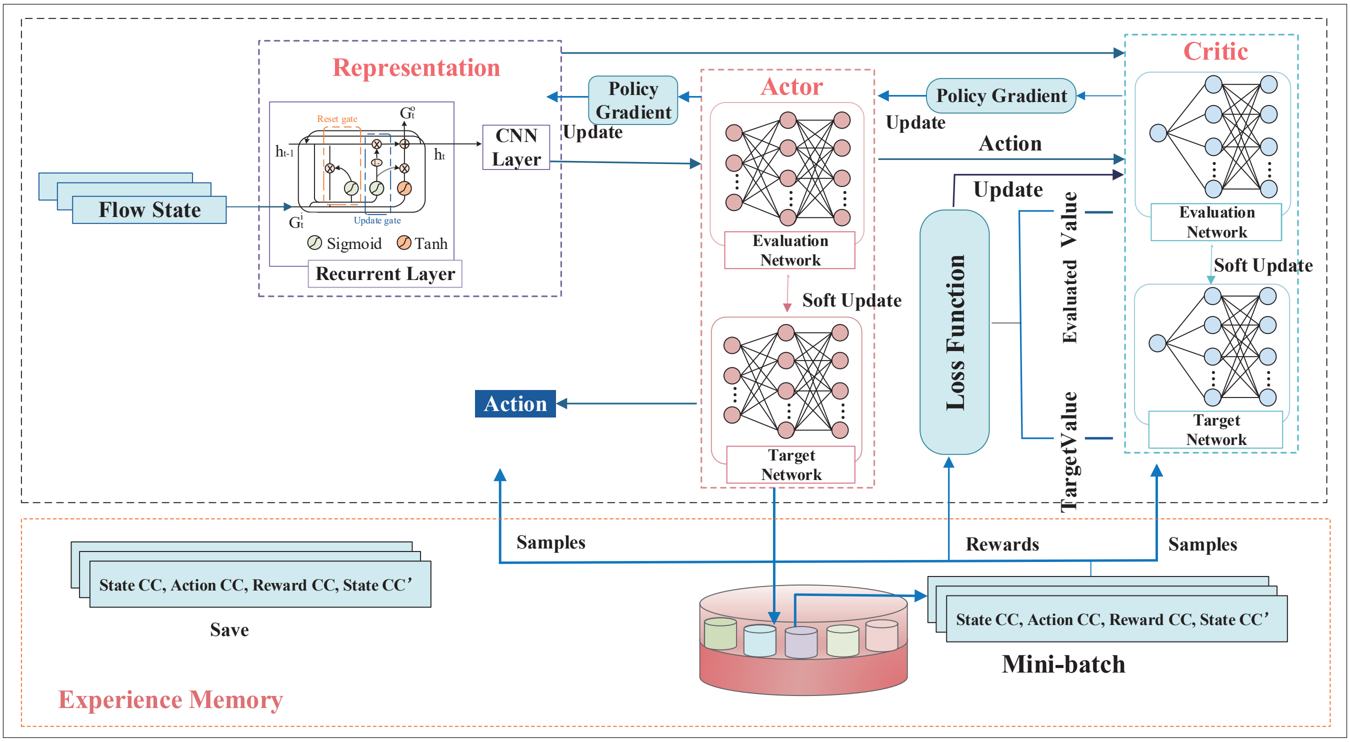

FIGURE 3. Overview of HSR-CC.

be alleviated by considering physical network factors, such as RSRP. In other words, if we can adjust the CWND through reducing the dependence on packet loss and/or RTT, MPTCP has a chance to be more responsive to the depleting connection due to handover.

HSR-CC: Cross-Layer Aided MPTCP

As discussed above, in the HSR scenario, it is very necessary to distinguish between the fluctuations induced by congestion and those induced by handovers. If a MPTCP connection just relies on transport layer conditions, such as packet loss and RTT, it is hard to achieve prompt reaction to handover. Therefore, it is significant to investigate a CC algorithm that can balance between the imminent handover and network congestion. Combining real-time network monitoring and cross-layer aid into MPTCP is a potential scheme.

In this context, we design a novel cross-layer aided MPTCP CC mechanism with leveraging DRL. We name it HSR-CC. The MHR is taken as an agent who can generate the CC action based on observed state. Note that the authors in [9] have proposed DRL-CC and implemented it in practice. They regarded a regular laptop, which is equipped with an Intel i7-3630QM CPU and 4GB memory, as a DRL-CC agent, and they found the DRL-CC algorithm can easily be run and trained. Therefore, it is suggested that the MHR is equipped with higher hardware configuration than the laptop used in [9], guaranteeing that the MHR equipment has powerful computing capability to finish the HSR-CC algorithm.

Overview of HSR-CC

Since DQN can only deal with discrete control, it is not efficient to solve the CC problem. Therefore, HSR-CC is investigated in this work, which is a continuous CC scheme with DRL. To the best of our knowledge, this work is the first to alleviate performance issues caused by frequent handover and investigate a novel CC mechanism in MPTCP with cross-layer information (i.e., RSRP) assistance and DRL to differentiate between congestion and handover. The model that integrates gated recurrent unit (GRU), convolutional neural network (CNN), and deep deterministic policy gradient (DDPG) [13] is innovative, which has not been employed in the context of DRL. Fig. 3 illustrates the detailed structure of HSR-CC, which consists of four components as follows.

Representation Network: The representation network plays the role of generating a representation of the network state, which is then input to the actor and critic networks for designing CC actions. When the HST is travelling, the networks that can be accessed by the MHR may change over time. For example, in some transportation lines, the MHR can simultaneously connect to the satellite, LTE and 3G networks. But in some transportation lines, the MHR can only simultaneously connect to the satellite and LTE networks. In this case the number of MPTCP subflows will change from time to time. However, the input size of the great majority of deep neural networks is constant. In order to make the input size variable, the recurrent layer (i.e., GRU) is adopted. The key advantage of GRU is the gate mechanism, which can decide what information to forget or remember. Different from long short-term memory (LSTM), which is composed of a cell and three gates, that is, the forget, input and output gates, GRU shown in Fig. 3 is composed of two gates, that is, the reset and update gates. It has been shown that in contrast to LSTM, GRU has comparable performance with much more efficient computing ability [14]. On the other hand, since the HSR network is extremely dynamic and complex [1], it will generate huge state space. In the case that the state is input to the actor and critic networks for designing CC actions without further treatment, the model will inevitably generate too many parameters. It is helpful to decrease the number of parameters with convectional layers. In brief, the CC state is first processed by the representation network, which contains one GRU layer and one CNN layer, and then input to the actor and critic networks for designing CC actions.

Actor Network: The actor network, which is made up of an evaluation network as well as a target network, plays the role of deriving CC actions based on observed state. The evaluation network's structure is the same as the target network's, but they have different parameters. The parameters of the evaluation network, which are periodically cloned to the target network, are updated in real time.

Critic Network: The critic network plays the role of assessing the performance of CC actions generated by the actor networks, which helps the actor to optimize the policy gradient and update CC actions. Finally, the CC action can be converged by continuously interacting among the representation network, actor network and critic network.

Experience Memory: The experience memory saves experience sequences, including CC state, CC action, CC reward, and next CC state. The representation network, critic network and actor network are trained by random sampling from the saved experience sequences that can efficiently reduce the dependence of data.

Elements of HSR-CC

As shown in Fig. 3, the CC mechanism of HSR-CC can be characterized as a DRL task to find sagacious CC strategies under SGIN-HSR circumstances. Next, we will introduce the critical elements of the HSR-CC algorithm in detail.

State: For a MPTCP flow including multiple subflows, the state of each subflow at each epoch includes its RSRP, loss count, queuing delay, average RTT, goodput as well as CWND size. Here, RSRP has a powerful influence on predicting handover process. The loss count denotes the number of lost packets in each epoch. The queuing delay represents the gap between average RTT and minimal RTT. As mentioned above, packet loss and/or RTT increment is usually the sign of congestion. Observing Fig. 2, we can find that RSRP, throughput, PLR and RTT are highly related in anticipating when a handover is about to occur. What's more, the authors in [5] input these parameters as the state of the DQN agent, and proposed Hd-TCP algorithm, which indicates that combining these metrics can really help to distinguish between adverse influences (such as packet loss and/or RTT increase) caused by congestion and those caused by handovers. That is to say, in the case that the agent observes that average RTT, packet loss count and RSRP present large fluctuations, the packet loss events are usually as a result of handovers, instead of congestion. Consequently, the agent is capable of distinguishing between adverse influences caused by congestion and those caused by handovers, and then generate sound CC strategies through trial and error.

Action: In the CC for MPTCP, the action plays the role of adjusting the CWND of the subflow. In this article, the action at each epoch means how much change in the CWND of MPTCP subflow will be required. The zero value denotes that the CWND size will not be changed. The amount of CWND size increment and reduction separately.

Reward: The reward plays the role of evaluating the quality of CC actions. In this article, the reward function adopted is identical to that in [9], which can strike a balance between fairness and goodput.

Training Process

Based on Fig. 3, the training process of HSR-CC performs as follows. The state information of MPTCP sub-flows collected by the agent is input into the representation network, then the representation of the state is generated. After that, based on the representation of the state as well as experience memory, the agent designs the CC action for the target subflow by continuously interacting between the actor network and critic network. Actually, the critic network can be taken as a DQN. The critic network updates its parameters according to minimizing the squared error loss. Both the actor network and representation network update their parameters according to policy gradient. Note that the representation, actor and critic networks are trained in an end-to-end way.

During the training process, it may take a lot of time to train the agent to converge, which can be performed offline. Once the agent is trained to converge, the agent can rapidly generate sound CC strategies. In our tests, it takes about 0.6 ms to make the online inference, which is advantageous to real-time decision making in HSR networks. Moreover, the HSR transportation lines are fixed within the strictly planned time frame, which makes the HSR networks obey certain space-time rules. According to environment dynamics [15], the agent does not need to be retrained in a short time. In other words, it is likely that the agent is retrained once a week or even every two weeks.

The HSR transportation lines are fixed within the strictly planned time frame, which makes the HSR networks obey certain space-time rules. According to environment dynamics, the agent does not need to be retrained in a short time. In other words, it is likely that the agent is retrained once a week or even every two weeks.

Performance Evaluation

To assess the HSR-CC performance, we configure the SGIN-HSR architecture in Fig. 1 based on NS3-DCE. The MHR is equipped with satellite and LTE interfaces, and utilizes the HSR-CC protocol to simultaneously transmit data via satellite and LTE links. In the simulation, as mentioned above, the MHR experiences handover on the LTE link while keeping the connection to the satellite network. The LTE network handover is triggered by the event A3 algorithm [12]. HSR-CC modifications were implemented based on MPTCP Linux kernel v0.92. According to measurements of [4], we set the LTE network bandwidth and RTT as 5 Mb/s and 100 ms, separately. The satellite network bandwidth and RTT are set as 10 Mb/s [3] and 260 ms, separately. The MHR transmits data to the destination at 11 Mb/s. We configure that there are eight eNBs. That is to say, the HST will perform handover seven times. The distance between eNBs ranges from 1000 m to 1600 m. The HST's speed is 300 km/h. The entire simulation time is 120 seconds. On the other hand, before testing, the HSR-CC agent, that is, the MHR, was trained for more than 30,000 epochs with the off-line way. In the training phase, the representation network contains a single-layer GRU unit and a single-layer CNN unit. The actor

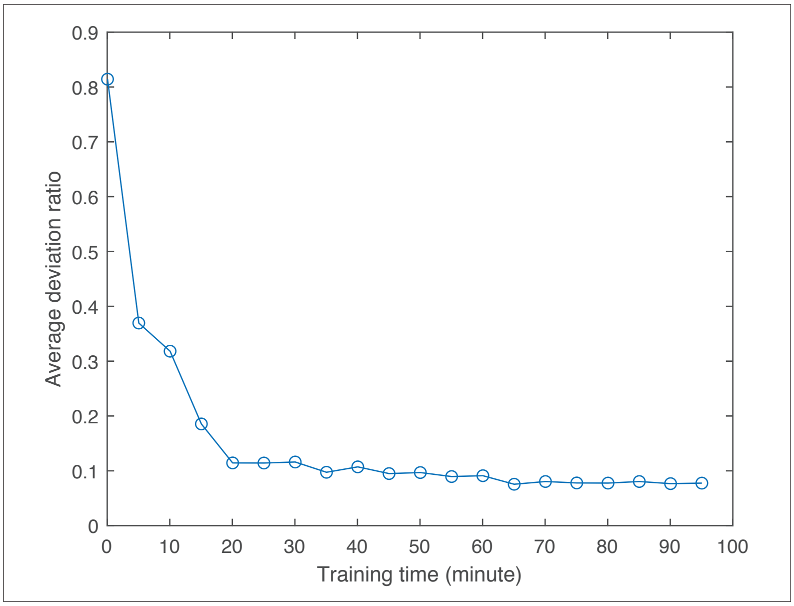

FIGURE 4. Average deviation ratio of HSR-CC versus training time.

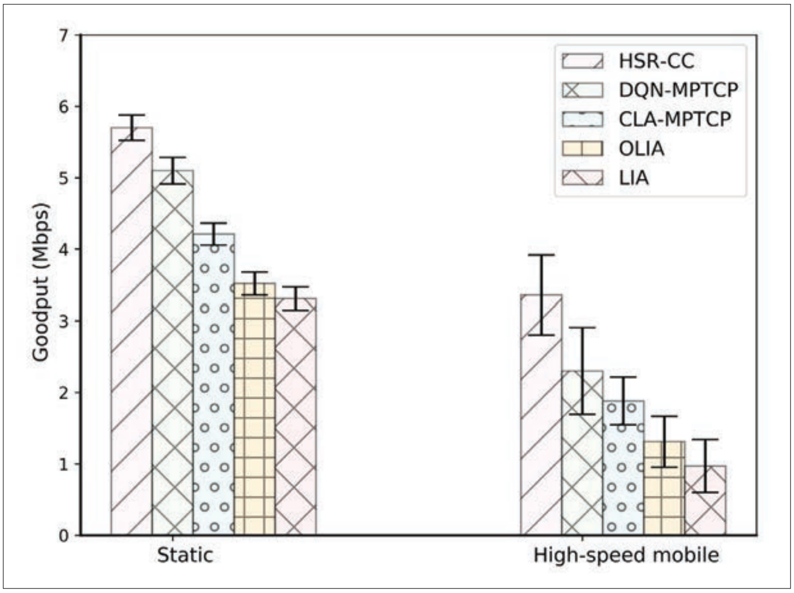

FIGURE 5. Goodput comparison of different MPTCP CC algorithms.

and critic networks both contain a three-layer fully connected neural network with the learning rate of 0.001. After training HSR-CC, we compared it with three rule-based CC mechanisms (i.e., CLA-MPTCP [8], LIA, OLIA [7]) and one DRL-based CC mechanism (i.e., DQN-MPTCP [11]). LIA is the default MPTCP CC mechanism. OLIA is a revised version of LIA, which realizes Pareto optimization. CLA-MPTCP is a rule-based CC mechanism with cross-layer information (i.e., RSRP) assistance. DQN-MPTCP is a DQN-based MPTCP CC mechanism. The four benchmark schemes did not consider cross-layer information assistance.

We employ the deviation ratio between the real reward and the anticipated value [10] to show the convergence of our model. Fig. 4 shows the average deviation ratio variation of HSR-CC versus the training time. We can observe that as the training time goes on, the average deviation ratio declines rapidly. It takes about 60 minutes to dip below 0.1, which denotes HSR-CC can obtain the convergence after 60 minutes.

Figure 5 depicts the goodput performance of all algorithms in static and high-speed mobile environments. Note that goodput is different from throughput, which indicates the amount of data successfully received by the receiver within the specified deadline. As Fig. 5 shows, after the MHR experiences the handover, the goodput of all algorithms decreases. The DRL-based CC algorithms can achieve better performance than the rule-based algorithms in terms of goodput, but they have poorer performance than rule-based CC algorithms in terms of the standard deviation of the goodput values. This is because the DRL-based CC algorithms can employ the neural networks to generate numerous CC rules being automatically designed in adaption to the affluent surroundings, which is impossible for the rule-based CC algorithms. It is also important to note that in the two kinds of ML-based CC algorithms, DQN-MPTCP's action space is discrete, which cannot perform well under highly-dynamic HSR networks. Nevertheless, compared to the DRL-based CC algorithms, the rule-based CC algorithms also have two advantages. One is steady and the other is interpretable. In addition, HSR-CC obtains the highest throughput among all algorithms. For example, in a high-speed mobile environment, HSR-CC outperforms DQN-MPTCP by 63 percent. The reason is that LIA, OLIA and DQN-MPTCP only rely on the path's RTT as well as packet losses and ignore the fluctuations of RSRP, resulting in that they react slowly and irrationally to abrupt handover across the LTE link. CLA-MPTCP considers the cross-layer information (i.e., RSRP), which shows better performance than LIA and OLIA. But CLA-MPTCP is a rule-based CC mechanism, which cannot take rapid action in the time-varying networks. In contrast, HSR-CC pays close attention to the fluctuations of RTT, packet losses and RSRP in SGIN-HSR, and uses the cognitive ability to distinguish between negative effects due to congestion and those due to handover. This proves that an efficient ML-based CC mechanism should pay more attention to cross-layer information besides RTT as well as packet losses. Last but not least, in the test, we observe that HSR-CC's Jain's fairness index is close to 1, which indicates that it can obtain good fairness. That is due to the reward that we adopted, which is able to strike a balance between fairness and goodput.

Conclusion and Future Work

In this article, we first presented the MPTCP-empowered SGIN-HSR architecture for meeting future railway service requirements. Then to distinguish between adverse influences caused by congestion and those caused by handover, we investigated a novel cross-layer information (i.e., RSRP) aided CC scheme with DRL, namely HSR-CC based on the MPTCP-empowered SGIN-HSR architecture, which is capable of intelligently adjusting the CWND from past experience rather than pre-defined rules. Numerical results demonstrated that HSR-CC is more robust to the handover than existing MPTCP CC schemes in the SGIN-HSR.

Though HSR-CC has showed the performance advantages of employing cross-layer aided information, there is still room for improvement. First, HSR-CC performance was only evaluated based on an emulated platform. The performance of HSR-CC, such as goodput and RTT, in real HSR environments should be measured. Second, compared to the rule-based CC mechanisms, HSR-CC is uninterpretable. It is necessary to improve the interpretability for the CC strategies derived by ML. Third, if HSR-CC is directly applied to a new HSR environment, it needs to be retrained for generating corresponding CC state-action pairs, which will require long training time and heavy calculation work. Transfer learning is capable of shortening training time with the knowledge gained from another relevant environment. In the future, we will integrate HSR-CC with transfer learning, which is beneficial to improve the transferability of CC mechanisms in terms of rapidly fitting into the new scenarios and ensuring good efficient performance.

Acknowledgment

This work was supported in part by the Royal Society Newton Advanced Fellowship under Grant NA191006; in part by the National Key Research and Development Program under Grant 2016YFE0200900; in part by the NSFC under Grant 6196113039, 61725101 and U1834210; in part by the Major Projects of the Beijing Municipal Science and Technology Commission under Grant Z181100003218010; in part by the Scholarship from the China Scholarship Council under Grant 202007090172; in part by the State Key Lab of Rail Traffic Control and Safety under Grant RCS2020ZT010, RCS2019ZZ007, and RCS2018ZZ007; in part by the Beijing Natural Haidian Joint Fund under Grant L172020; and in part by the Open Research Fund from the Shenzhen Research Institute of Big Data under Grant 2019ORF01006.

References

[1] B. Ai *et al*., "5G Key Technologies for Smart Railways," *Proc. IEEE*, vol. 108, no. 6, June 2020, pp. 856–93.

[2] J. Xu *et al*., "When High-Speed Railway Networks Meet Multipath TCP: Supporting Dependable Communications," *IEEE Wireless Commun. Lett*., vol. 9, no. 2, Feb. 2020, pp. 202–05.

[3] L. Yan *et al*., "Safety-Oriented Resource Allocation for Space-Ground Integrated Cloud Networks of High-Speed Railways," *IEEE JSAC*, vol. 38, no. 12, Dec. 2020, pp. 2747–59.

[4] F. Mazzenga *et al*., "Integrated Public Mobile Radio Networks/Satellite for Future Railway Communications," *IEEE Wireless Commun*., vol. 24, no. 2, Apr. 2017, pp. 90–97.

[5] L. Cui *et al*., "Improving the Congestion Control Performance for Mobile Networks in High-Speed Railway via Deep Reinforcement Learning," *IEEE Trans. Vehic. Tech*., vol. 69, no. 6, June 2020, pp. 5864–75.

[6] L. Li *et al*., "A Measurement Study on Multi-Path TCP with Multiple Cellular Carriers on High Speed Rails," *Proc. ACM SIGCOMM*, Budapest, Hungary, Aug. 2018, pp. 161–75.

[7] R. Khalili *et al*., "MPTCP Is Not Pareto-Optimal: Performance Issues and a Possible Solution, " *IEEE/ACM Trans. Netw*., vol. 21, no. 5, Oct. 2013, pp. 1651–65.

[8] H. Sinky, B. Hamdaoui, and M. Guizani, "Proactive Multipath TCP for Seamless Handoff in Heterogeneous Wireless Access Networks," *IEEE Trans. Wireless Commun*., vol. 15, no. 7, 2016, pp. 4754–64.

[9] Z. Xu *et al*., "Experience-Driven Congestion Control: When Multi-Path TCP Meets Deep Reinforcement Learning," *IEEE JSAC*, vol. 37, no. 6, June 2019, pp. 1325–36.

[10] W. Li *et al*., "SmartCC: A Reinforcement Learning Approach for Multipath TCP Congestion Control in Heterogeneous Networks," *IEEE JSAC*, vol. 37, no. 11, Nov. 2019, pp. 2621–33.

[11] S. R. Pokhrel and S. Garg, "Multipath Communication with Deep QNetwork for Industry 4.0 Automation and Orchestration," *IEEE Trans. Ind. Informat*., vol. 17, no. 4, Apr. 2021, pp. 2852–59.

[12] H. Cho *et al*., "LTE-R Handover Point Control Scheme for High-Speed Railways," *IEEE Wireless Commun*., vol. 24, no. 6, Dec. 2017, pp. 112–19.

[13] T. P. Lillicrap *et al*., "Continuous Control with Deep Reinforcement Learning," *Proc. 4th ICIR*, San Juan, Puerto Rico, May 2016, pp. 1–14.

[14] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An Empirical Exploration of Recurrent Network Architectures," *Proc. 32nd ICML*, vol. 37, July 2015, pp. 2342–50.

[15] J. Xu and B. Ai, "Artificial Intelligence Empowered Power Allocation for Smart Railway," *IEEE Commun. Mag*., vol. 59, no. 2, Feb. 2021, pp. 28–33.

Biographies

Jianpeng Xu [S'20] (18111037@bjtu.edu.cn) is currently pursuing a Ph.D. degree with the State Key Laboratory of Rail Traffic Control and Safety, School of Electronic and Information Engineering, Beijing Jiaotong University, China. He is also with the Frontiers Science Center for Smart High-speed Railway System. His research interests include mobility management and deep reinforcement learning.

Bo Ai [SM'10] (boai@bjtu.edu.cn) received his Ph.D. from Xidian University in 2004. He is a professor at Beijing Jiaotong University, China, where he is the Deputy Director of the State Key Lab of Rail Traffic Control and Safety. He is also with the Research Center of Networks and Communications, Peng Cheng Laboratory, Shenzhen 518055, China, and the Frontiers Science Center for Smart High-speed Railway System. He is also with the Henan Joint International Research Laboratory of Intelligent Networking and Data Analysis, Zhengzhou University, China. His research interests are rail traffic mobile communications and channel modeling. He is an IET Fellow.

Transfer learning is capable of shortening training time with the knowledge gained from another relevant environment. In the future, we will integrate HSR-CC with transfer learning, which is beneficial to improve the transferability of CC mechanisms in terms of rapidly fitting into the new scenarios and ensuring good efficient performance.