IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 19 July 2023, accepted 31 July 2023, date of publication 4 August 2023, date of current version 10 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3301964

RESEARCH ARTICLE

Optimal Control Method for Congestion Control and Delay Reduction in Deterministic Networks

Q. GAO, J. LI, Y. LIU, AND Y. XING
6G Research Center, China Telecom Research Institute, Beijing 102209, China

Corresponding author: Q. Gao (gaoqy3@chinatelecom.cn)

This work was supported by the 2020 National Key Research and Development Program ''Broadband Communication and New Network'' Special ''6G Network Architecture and Key Technologies'' under Grant 2020YFB1806700.

ABSTRACT The proliferation of the types and number of the beyond fifth generation / the six generation (B5G/6G) network services will contribute to severe network congestion collapse, which is detrimental to deterministic service assurance for time-sensitive businesses. Faced with the congestion control problem in deterministic networks, the regulation and control effect of network macroscopic congestion state on the node various microscopic control strategies should be sufficiently considered from a global view. In this paper, a congestion control system model is constructed, which includes node-level congestion states and control strategies and network-level optimization objective. Continuous congestion state transition and real-time control strategy selection are analyzed to minimize the network delay and congestion degree with the lowest control cost. The analytical solution to optimal control strategy is derived based on the optimal control theory, and the numerical solution is analyzed by proposing a congestion control discretization algorithm (CCDA). Simulation results demonstrate that the congestion degree of each network node gradually decreases with the obtained CCDA strategies. The impacts of model parameters such as congestion deterioration probability and delay weight coefficients on the evolution of network total loss and optimal strategies are also analyzed. Moreover, the network total loss of the proposed congestion control model is lower than the baseline model, which can provide effective and referable theoretical guidance for congestion control problem in deterministic networks.

INDEX TERMS Congestion control, deterministic networks, optimal control theory.

I. INTRODUCTION

The continuously evolving fifth generation (5G) network and the future six generation (6G) network will provide more ubiquitous connectivity, more extreme performance experience, and more customized services for ''to customer'' (2C) and ''to business'' (2B) fields. At the same time, the explosion of service traffic leads to a large number of network congestion and packet delay, which is not conducive to the quality of service (QoS) guarantee of time-sensitive services. In particular, industrial control, automatic driving and other vertical industry scenarios that have high real-time and security demands require ultra-low millisecond delay, microsecond jitter and more than 99.999% reliability [1]. High-definition video surveillance scenarios have high requirements for video resolution and size, and need stable transmission with large bandwidth [2]. However, the traditional best-effort (BE) delivery mechanism can no longer satisfy the punctual and accurate delivery demands of the future services due to the long-tail effect of end-to-end delay [3], which urgently calls for a novel deterministic delivery mode provided for time-sensitive services.

As a consequence, the Internet Engineering Task Force (IETF) DetNet working group proposes the deterministic network [4] to achieve end-to-end routing data flow on the basis of IEEE 802.1 bridged networks and to expand the network scale of deterministic assurance. Typical deterministic network technologies mainly include Flexible Ethernet (FlexE) [5], [6] for Layer-2 Ethernet, Time-sensitive Networking (TSN) [7] for local area network (LAN), Deterministic Internet Protocol (DIP) [8], [9] for Layer-3 large-scale networks, Deterministic Networking

The associate editor coordinating the review of this manuscript and approving it for publication was Xujie Li.

(DetNet) [10] for wide area network (WAN), and Deterministic Mobile Edge Computing (D-MEC) [11] for MEC networks, etc. Although applicable to different networks, these technologies are commonly dedicated to providing accurate, bounded network performance. Service traffic aggregation and congestion collapse can be mitigated and time-critical services QoS can be deterministically guaranteed by various deterministic mechanisms such as resource reservation, multipath redundancy, and explicit path planning.

Existing work on deterministic networks mainly focuses on the technologies or mechanisms to reduce the network end-to-end delay, increase the number of schedulable flows and mitigate network congestion, including the traditional queue scheduling mechanisms and the novel queue deterministic enhancement to improve the network determinacy. For example, at the queue input ports for enqueue congestion control, the active queue management (AQM) actively discards packets that exceed the queue length threshold [12], whereas the explicit congestion notification (ECN) notifies the upstream nodes to reduce the sending rate to mitigate traffic aggregation [13]. At the queue output ports for rate control, traffic flows are shaped based on different criterions such as token, credit value, gate control list (GCL), and priority [14]. Among multihop nodes for packet forwarding, the path and delay are optimized deterministically through slot mapping based on time synchronization and cycle mapping based on frequency synchronization [15], [16], [17], [18].

However, existing work mainly concerns the influence of a certain technology on the partial-network performance optimization from the microscopic technical level, and lacks the overall-network multiple strategy decision-making from the macroscopic regulation level. A thorough solution to network congestion could be achieved only with the cooperation of the various control strategies, which is inseparable from the macroscopic level regulation and control. Especially in the case of a large number of congested nodes, the decision ignoring the overall-network real-time states may aggravate the burden of the whole network, e.g., extra delay and resource waste. In addition, the node types and traffic flow types will also affect the node strategy selection in deterministic networks where ordinary traffic flows and deterministic traffic flows coexist. But the existing work mainly concentrates on the QoS enhancement of deterministic services accompanied by relatively high cost consumption. Appropriately relaxing the deterministic-applicable mechanisms for ordinary traffic flows may reserve more available resources for the network especially deterministic traffic flows. This trade-off between network performance and cost budget is definitely a challenge for current research on deterministic network.

To solve the above problems, a real-time congestion control system model in deterministic network is constructed in this paper. The real-time prediction and macroscopic regulation of congestion control strategies according to the different states of nodes are carried out. The network end-to-end delay and congestion degree are jointly minimized with the lowest control strategy cost while taking distinguished node states and traffic flows into account. Simulation results are provided to demonstrate the evolution of network congestion degree with node control strategies. The network total loss is reduced compared with the baseline model, proving the effectiveness of the proposed congestion control model.

The main contributions of this paper are summarized as follows.

- To alleviate the network congestion and better guarantee the determinacy of the future network, a congestion control system model is constructed in the deterministic network scenario. The node-level congestion state evolution, real-time control strategy selection and network-level optimization objective are sufficiently considered.
- The transition of congestion state and the optimal congestion control strategy are analyzed based on the optimal control theory. In addition, a CCDA algorithm is proposed to jointly minimize the network end-to-end delay and congestion degree with the lowest control cost.
- Simulation results demonstrate the CCDA strategies can reduce the congestion degree of network nodes and the network total loss. In addition, the impacts of model parameters e.g., congestion deterioration probability and delay weight coefficients are also analyzed.

The remainder of this paper is organized as follows. Related work is provided in Section II. The congestion control system model in deterministic networks and optimization problem formulation are provided in Section III. The optimal control theory based solution and the CCDA algorithm are proposed to obtain the optimal congestion control strategies in Section IV. Simulation results are provided and discussed in Section V. Finally, Section VI summarizes this paper.

II. RELATED WORK

Reducing network delay and congestion to improve the network determinacy has been studied in recent decades, and we roughly divide them into the following four categories: enqueue congestion control, queue scheduling mechanism, dequeue shaping mechanism, multihop forwarding mechanism. The research gap analysis is provided in Subsection II-E.

A. ENQUEUE CONGESTION CONTROL

As a typical congestion control scheme to reduce the buffer occupation and end-to-end delay, AQM drops the subsequent packets when the packet length exceeds the queue length threshold. Based on this, an information compression model for AQM design is proposed to deal with the heterogeneous RTT and time-varying traffic load in [19], achieving low latency and high scalability. To deal with head of line (HOL) blocking, regional explicit congestion notification (RECN) for source deterministic routing networks and flow-based implicit congestion management (FBICM) designed for distributed deterministic routing networks are proposed in [20]. A deterministic latency congestion control system is provided in [21] that integrates rate control, multiqueue management, and per-hop delay correction mechanisms, reducing the delay jitter by 90%. A two-timescale congestion window control mechanism is proposed in [22], considering distinguished delay-aware priorities of flows and enqueue delay to reduce the average service delay compared with TCP Vegas. Reference [23] proposes an online learning Crimson to continuously perceive and adaptively satisfy the deterministic application QoS. Consistent performance can be achieved under diverse network scenarios. The above schemes mainly implement congestion control in the enqueue process by means of intra-node operation or inter-node collaboration.

B. QUEUE SCHEDULING MECHANISM

According to the granularity of queue scheduling, it can be classified into: per-port scheduling, per-user scheduling, per-service-class scheduling, per-queue scheduling, per-flow scheduling and per-packet scheduling, from coarse-grained to fine-grained. Especially in TSN and DetNet, per-service-class scheduling and per-flow scheduling are mainly deployed. For per-service-class scheduling, the priority is allocated based on the service class, and reflected in the Ethernet Priority Code Point (PCP). However, the flows with the same PCP will easily result in congestion. For per-flow scheduling, burst can be managed stream by stream and fine-grained QoS services can be provided. The cost and scalability will get worse with the network expansion. Therefore, classified stream by stream hybrid queue is proposed to integrate the advantages of per-service-class scheduling and per-flow scheduling [24], [25].

C. DEQUEUE SHAPING MECHANISM

Typical dequeue shaping mechanisms such as token bucket, credit-based shaping (CBS) of IEEE 802.1Qav, time-aware shaping (TAS) of IEEE 802.1Qbv, asynchronous traffic shaping (ATS) of IEEE 802.1Qcr, frame preemption (FP) of IEEE 802.1Qbu can provide delay guarantee with different granularity. For token bucket [26], tokens are put in at a certain rate, packets with tokens can be sent whereas packets without tokens will be dropped or stored. It guarantees determinacy in terms of bandwidth but has poor latency performance. CBS [27] provides services with different priorities, and traffic flows are sent based on the constantly changing credit value. TAS [28] uses GCL to control the periodic packet transmission at the queue exit, which can realize microsecond per-hop-per-packet scheduling but needs strict time synchronization. An enhanced TAS is proposed in [29] to guarantee determinacy for aperiodic key traffic neglected by TAS. ATS [30] relaxes time synchronization requirement and uses urgency-based scheduler. Emergency traffic can be prioritized and ATS still has high bandwidth utilization when periodic and sporadic traffic coexist. FP [31] allows high-priority eMAC traffic to interrupt low-priority pMAC traffic, based on which a frame truncation assisted FP is proposed in [32]. By discarding the least important bytes, this enhanced-FP can transmit more payload bytes of image and video without increasing delay of higher-priority services.

D. MULTIHOP FORWARDING MECHANISM

Queue forwarding mechanisms generally place a gate at the entry and exit of the queue, respectively. The time is divided into equal slots T and periodic mapping between multiple hops needs to be maintained. The enqueuing and dequeuing of packets are cyclic/rotated. Jitter can generally be controlled within $2T$. Cyclic queue forwarding (CQF) [33] requires strict time synchronization. The slot mapping is $x \rightarrow x \rightarrow x+1$ (send in time slot x, reach the next hop in x and send in $x+1$) which requires propagation delay less than T and is suitable for LAN. To deal with the shortcomings of CQF, CQF-3 introduces a third queue for caching to avoid frame arrival slot errors. And to release the time synchronization requirement of CQF, scalable deterministic forwarding (SDF) adds the explicit period identifier to the datagram [25]. However, the above forwarding mechanisms are limited to LAN. For WAN, deterministic internet protocol (DIP) requires frequency synchronization with $x \rightarrow y \rightarrow y+1$ [34], whereas cyclic specified queue forwarding (CSQF) requires frequency synchronization with $x \rightarrow y \rightarrow y+n$ [35]. Multi-CQF, an extension of CSQF is evaluated in [36], the simulated annealing solution is developed and the latency is lower than CSQF.

E. RESEARCH GAP ANALYSIS

Existing work mainly focuses on the low-level microscopic technologies to improve the network deterministic performance. For example, proposing scheduling and forwarding mechanisms to solve the congestion control problem, reducing the cost and complexity of congestion control from the hardware perspective, reducing the delay and jitter from the algorithm perspective, etc. Besides, existing work mainly concentrates on the enhancement of deterministic services accompanied by relatively high cost consumption. However, a thorough solution to network congestion control could be achieved only with the appropriate cooperation and joint optimization of the various control strategies such as packet dropping in AQM, rate limitation notification in ECN, packet queuing in CQF, etc. At the same time, it is required to take both deterministic services and important nondeterministic services into account, and balance their cost and performance.

In summary, the research for theoretical modeling and upper-level regulation is generally lacked, where node-level congestion states and control strategies and network-level optimization objective should be studied. Therefore, it is necessary to model the interplay between macroscopic network congestion states and microscopic congestion control decision-making to assist the network to predict the congestion situation and adjust the strategies in time. In addition, it has important theoretical guiding significance for the prevention and mitigation of network congestion to consider the congestion control problem from the global perspective.

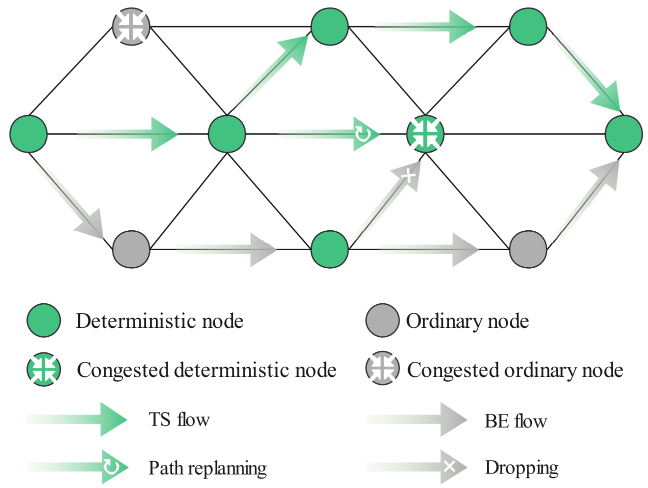

FIGURE 1. **Congestion control system model in deterministic networks.**

III. CONGESTION CONTROL SYSTEM MODEL IN DETERMINISTIC NETWORKS

As illustrated in Fig. 1, the congestion control system model in the deterministic network includes two types of nodes, the ordinary nodes and the deterministic nodes. The type of node $i \in [1, N]$ is denoted by $v_i(t)$.

$$v_i(t) = \begin{cases} 1, & \textit{deterministic node}, \\ 0, & \textit{ordinary node}. \end{cases} \tag{1}$$

The traffic flows can be transmitted from node i to node j only when the path between them is reachable, denoted by $e_{ij}(t)$.

$$e_{ij}(t) = \begin{cases} 1, & \textit{reachable}, \\ 0, & \textit{unreachable}. \end{cases} \tag{2}$$

The types of traffic flows include BE flows and time-sensitive (TS) flows, where the traffic flow from node i to node j is denoted by $p_{ij}(t)$.

$$p_{ij}(t) = \begin{cases} 1, & \textit{TS flow}, \\ 0, & \textit{BE flow}. \end{cases} \tag{3}$$

The ordinary nodes can only handle BE flows, whereas the deterministic nodes can handle both TS flows and BE flows. The node congestion states are influenced by the traffic arrival rate, the node congestion control strategies and the congestion states of other nodes. Each node adjusts the packet processing measures (e.g., path replanning, packet dropping) to alleviate the network congestion and to ensure the lowest user plane delay and control plane cost.

A. NODE CONGESTION STATE

Two node congestion states are defined in the proposed system model, congested state denoted by J and free (uncongested) state denoted by F.

- J, the forwarding capacity and buffer space of the nodes in this state are severely limited, (80% of the incoming packets can not be processed, just as an example.) The packets arriving at this kind of nodes will experience high delay.
- F, the nodes in this state can process the received packets normally.

Let the variable $C_i(t)$ denote the congestion state of node i at time t, the congestion probability of which is denoted by $c_i(t)$.

$$C_i(t) = \begin{cases} 1, & \textit{congested state } J_i(t), \\ 0, & \textit{uncongested state } F_i(t). \end{cases} \tag{4}$$

B. TRANSITION OF NODE CONGESTION STATE

BE flows and TS flows coexist in the deterministic network. When congested, network nodes perform path replanning for TS flows and dropping/path replanning for BE flows.

For congested node i, direct dropping may be performed after packet arrival or path replanning may be performed before packet arrival. Note that due to explicit path planning and resource reservation in deterministic networks, we assume that when congestion occurs, there must be reserved paths and resources for the forwarding of deterministic flows. The congestion control strategies for congested nodes are defined as

- Dropping rate $\delta_{ji}(t)$. When node j is free and node i is congested, node i can drop the BE packets received from node j, so as to alleviate the traffic aggregation at node i. After dropping, the packet sender will resend the packets at intervals, resulting in large waiting and resending delay. However, the extra control cost is small or even negligible. $e_{ji}(t) = 0, \delta_{ji}(t) = 0; C_j(t) = 1, \delta_{ji}(t) = 0;$ $C_i(t) = 0, \delta_{ji}(t) = 0$ and $v_i(t)\, v_j(t)\, p_{ji}(t) = 1,$ $\delta_{ji}(t) = 0.$
- Path replanning rate $\gamma_{ji}(t)$. When node j is free and node i is congested, the path of TS/BE packets sent from node j to node i can be replanned to alleviate the traffic aggregation at node i. The delay caused by this behavior is less than that of direct dropping, because the packets only need to be replanned from an intermediate node rather than resent from the source node. However, replanning paths and calculating routes will consume large extra control cost. $e_{ji}(t) = 0, \gamma_{ji}(t) = 0; C_j(t) = 1, \gamma_{ji}(t) = 0$ and $C_i(t) = 0, \gamma_{ji}(t) = 0.$

For uncongested node i: Assuming that the packet arrival follows a Poisson distribution with parameter λ_p, then the congestion deterioration probability λ_i of node i is defined as the probability of sufficient packet arrival, where (5) is taken as an example.

$$\lambda_i = Pr\left(X > \lambda_p\right) = 1 - \sum\nolimits_{k=0}^{\lambda_p} \frac{\lambda_p^k}{k!} \exp\left(-\lambda_p\right). \tag{5}$$

It can be easily concluded through calculation that, as the packet arrival mean λ_p increases, the congestion deterioration probability λ_i increases, which is consistent with the actual scenario.

Therefore, the evolution of the congestion probability of node i over time, i.e., the congestion state differential equation, can be expressed as (6). The congestion state

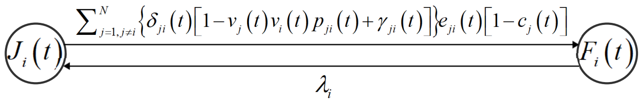

FIGURE 2. **Congestion state transition of network nodes.**

transition of network nodes is shown in Fig. 2.

$$\dot{c_i}(t) = [1 - c_i(t)]\lambda_i - c_i(t)\sum_{j=1,j\neq i}^{N}\{\delta_{ji}(t)[1 - v_j(t)v_i(t)p_{ji}(t)] + \gamma_{ji}(t)\}e_{ji}(t)[1 - c_j(t)]. \tag{6}$$

C. OPTIMIZATION PROBLEM FORMULATION

The optimization objective of the system is to minimize the additional delay and to alleviate the network congestion degree with the lowest control cost.

The additional delay includes additional queuing delay caused by network congestion, additional sending and propagation delay after being dropped, and additional processing delay for path replanning. Define $l_i(t)$ as the additional delay for node i at time t, which can be expressed as

$$l_i(t) = \left[ac_i(t) + d_1\sum_{j=1,j\neq i}^{N}\delta_{ji}(t)V_{ji}(t)E_{ji}(t)c_i(t) + d_2\sum_{j=1,j\neq i}^{N}\gamma_{ji}(t)E_{ji}(t)c_i(t)\right]. \tag{7}$$

For ease of expression, we define

$$V_{ji}(t) = 1 - v_j(t)v_i(t)p_{ji}(t). \tag{8}$$

$$E_{ji}(t) = e_{ji}(t)[1 - c_j(t)]. \tag{9}$$

a, d_1 and d_2 are the delay weight coefficients used to measure the delay caused by congestion, dropping, and path replanning, respectively.

The control cost includes dropping cost and path replanning cost, proportional to the quadratic form of the congestion control strategies. Define $f_i(t)$ as the control cost for node i at time t, which can be expressed as

$$f_i(t) = \frac{b_1}{2}\sum_{j=1,j\neq i}^{N}\delta_{ji}^2(t)V_{ji}(t)E_{ji}(t)c_i(t) + \frac{b_2}{2}\sum_{j=1,j\neq i}^{N}\gamma_{ji}^2(t)E_{ji}(t)c_i(t). \tag{10}$$

b_1 and b_2 are the cost coefficients for implementing dropping and path replanning strategies, respectively.

Therefore, the instantaneous objective loss $g(t)$ at time t for the entire network is defined as

$$\begin{aligned} g(t) &= \sum_{i=1}^{N}[l_i(t) + f_i(t)] \\ &= \sum_{i=1}^{N}c_i(t)[a + d_1\sum_{j=1,j\neq i}^{N}\delta_{ji}(t)V_{ji}(t)E_{ji}(t) \\ &\quad + d_2\sum_{j=1,j\neq i}^{N}\gamma_{ji}(t)E_{ji}(t) \\ &\quad + \frac{b_1}{2}\sum_{j=1,j\neq i}^{N}\delta_{ji}^2(t)V_{ji}(t)E_{ji}(t) \\ &\quad + \frac{b_2}{2}\sum_{j=1,j\neq i}^{N}\gamma_{ji}^2(t)E_{ji}(t)]. \end{aligned} \tag{11}$$

The total objective loss G for the entire network, i.e., network total loss, is the integral of the instantaneous objective loss over the time period $[t_0, t_f]$.

$$G = \int_{t_0}^{t_f} g(t)\,dt = \int_{t_0}^{t_f}\sum_{i=1}^{N}[l_i(t) + f_i(t)]\,dt. \tag{12}$$

Network nodes adjust their congestion control strategies to minimize the network total loss G. Therefore, the optimization problem can be expressed as P_1:

$$P_1: \min_{\{\delta_{ji}(t)\},\{\gamma_{ji}(t)\}} G, \tag{13a}$$

$$s.t.\ o_{ji}(t) \leqslant \delta_{ji}(t) + \gamma_{ji}(t), \tag{13b}$$

$$c_i(t) \in [0, 1], \tag{13c}$$

$$d_1 > d_2 > 0, \tag{13d}$$

$$0 < b_1 < b_2, \tag{13e}$$

$$\delta_{ji}(t) \geqslant 0, \tag{13f}$$

$$\gamma_{ji}(t) \geqslant 0. \tag{13g}$$

(13b) indicates that the sum of the two strategies cannot be lower than the amount of overflowed packets received from node j at node i. Since the dropping delay is larger than the path replanning delay, the delay weight coefficient for dropping d_1 is larger than that of path replanning d_2 as (13d). Similarly, since the dropping cost is smaller than the path replanning cost, the cost coefficient for dropping b_1 is smaller than that of path replanning b_2 as (13e).

IV. OPTIMAL CONTROL THEORY BASED SOLUTION AND CCDA ALGORITHM

A. OPTIMAL CONTROL THEORY BASED SOLUTION

Optimal control theory seeks a control strategy that optimizes the given system performance, subject to given constraints. Especially, Hamilton function method is widely used, which transforms the optimization of the system optimization objective to the optimization of the constructed Hamilton function [37]. Therefore, the Hamilton function with respect to the instantaneous objective loss (11) and the congestion state differential equation (6) is constructed by introducing the costate variable $\mu_i(t)$ for each node i.

$$\begin{aligned} H(t) &= g(t) + \sum_{i=1}^{N}\mu_i(t)\dot{c_i}(t) \\ &= \sum_{i=1}^{N}c_i(t)[a + d_1\sum_{j=1,j\neq i}^{N}\delta_{ji}(t)V_{ji}(t)E_{ji}(t) \\ &\quad + d_2\sum_{j=1,j\neq i}^{N}\gamma_{ji}(t)E_{ji}(t) \\ &\quad + \frac{b_1}{2}\sum_{j=1,j\neq i}^{N}\delta_{ji}^2(t)V_{ji}(t)E_{ji}(t) \\ &\quad + \frac{b_2}{2}\sum_{j=1,j\neq i}^{N}\gamma_{ji}^2(t)E_{ji}(t)] \\ &\quad + \sum_{i=1}^{N}\mu_i(t)\{[1 - c_i(t)]\lambda_i \\ &\quad - c_i(t)\sum_{j=1,j\neq i}^{N}[\delta_{ji}(t)V_{ji}(t) + \gamma_{ji}(t)]E_{ji}(t)\}. \end{aligned} \tag{14}$$

Therefore, the optimization problem P_1 is equivalent to the optimization problem P_2.

$$P_2: \min_{\{\delta_{ji}(t)\},\{\gamma_{ji}(t)\}} H(t), \tag{15a}$$
$$s.t.\ \dot{c}_i(t) = \frac{\partial H(t)}{\partial u_i(t)}, \tag{15b}$$
$$c_i(t_0) = c_{i0}, \tag{15c}$$
$$\dot{\mu}_i(t) = -\frac{\partial H(t)}{\partial c_i(t)}, \tag{15d}$$
$$\mu_i(t_f) = 0, \tag{15e}$$
$$o_{ji}(t) \leqslant \delta_{ji}(t) + \gamma_{ji}(t), \tag{15f}$$
$$c_i(t) \in [0,1], \tag{15g}$$
$$d_1 > d_2 > 0, \tag{15h}$$
$$0 < b_1 < b_2, \tag{15i}$$
$$\delta_{ji}(t) \geqslant 0, \tag{15j}$$
$$\gamma_{ji}(t) \geqslant 0. \tag{15k}$$

It is easy to prove that

$$\frac{\partial H(t)}{\partial \delta_{ij}(t)} \geqslant 0, \tag{16a}$$
$$\frac{\partial H(t)}{\partial \gamma_{ij}(t)} \geqslant 0. \tag{16b}$$

Therefore, the minimum of $H(t)$ with respect to the control strategies exists. Since the system model is initial-state-fixed and final-state-free, the state constraints and costate constraints of this kind of model are (15b)-(15c) and (15d)-(15e), respectively, which is according to the Hamilton optimal control theory. Moreover, the constraints (13b)-(13g) of P_1 should be also applied to P_2, i.e., (15f)-(15k).

The optimal control strategies $\left\{\delta_{ji}^*(t)\right\}$ and $\left\{\gamma_{ji}^*(t)\right\}$ in optimization problem P_2 should satisfy (17).

$$H\left(\left\{\delta_{ji}^*(t)\right\}, \left\{\gamma_{ji}^*(t)\right\}\right) \leqslant H\left(\left\{\delta_{ji}(t)\right\}, \left\{\gamma_{ji}(t)\right\}\right). \tag{17}$$

To solve $\left\{\delta_{ji}^*(t)\right\}$ and $\left\{\gamma_{ji}^*(t)\right\}$, the state variables $\{c_i(t)\}$ and costate variables $\{\mu_i(t)\}$ need to be solved as Lemma 1 and Lemma 2, respectively.

Lemma 1: The state variables $\{c_i(t)\}$ can be obtained by solving the congestion state differential equation (18) with boundary condition (15c).

$$\begin{aligned}\dot{c}_i(t) = {}& [1 - c_i(t)]\lambda_i - c_i(t)\sum\nolimits_{j=1,j\neq i}^{N}\left[\delta_{ji}(t)V_{ji}(t)\right.\\ &\left.+\gamma_{ji}(t)\right]E_{ji}(t).\end{aligned} \tag{18}$$

Proof: According to (15b), solve the partial derivative of (14) with respect to $\{\mu_i(t)\}$, (18) can be obtained. ■

Lemma 2: The costate variables $\{\mu_i(t)\}$ can be obtained by solving the costate differential equation (19) with boundary condition (15e).

$$\begin{aligned}\dot{\mu}_i(t) = {}& -a - \sum\nolimits_{j=1,j\neq i}^{N}\left[d_1\delta_{ji}(t)V_{ji}(t) + d_2\gamma_{ji}(t)\right]E_{ji}(t)\\
&- \sum\nolimits_{j=1,j\neq i}^{N}\left[\frac{b_1}{2}\delta_{ji}^2(t)V_{ji}(t) + \frac{b_2}{2}\gamma_{ji}^2(t)\right]E_{ji}(t)\\
&+\mu_i(t)\left\{\lambda_i + \sum\nolimits_{j=1,j\neq i}^{N}\left[\delta_{ji}(t)V_{ji}(t) + \gamma_{ji}(t)\right]E_{ji}(t)\right\}\\
&+\sum\nolimits_{j=1,j\neq i}^{N}\left[d_1\delta_{ij}(t)V_{ij}(t) + d_2\gamma_{ij}(t)\right]c_j(t)e_{ij}(t)\\
&+\sum\nolimits_{j=1,j\neq i}^{N}\left[\frac{b_1}{2}\delta_{ij}^2(t)V_{ij}(t) + \frac{b_2}{2}\gamma_{ij}^2(t)\right]c_j(t)e_{ij}(t)\\
&-\sum\nolimits_{j=1,j\neq i}^{N}\left[\delta_{ij}(t)V_{ij}(t) + \gamma_{ij}(t)\right]\mu_j(t)c_j(t)e_{ij}(t).\end{aligned} \tag{19}$$

Proof: According to (15d), solve the partial derivative of (14) with respect to $\{c_i(t)\}$, (19) can be obtained. ■

Based on Lemma 1 and Lemma 2, the optimal congestion control strategies can be obtained by Theorem 1.

Theorem 1: The optimal congestion control strategies $\left\{\delta_{ji}^*(t)\right\}$ and $\left\{\gamma_{ji}^*(t)\right\}$ are

$$\delta_{ji}^*(t) = \begin{cases} 0, & c_i(t) = 0,\ or\ V_{ji}(t) = 0,\ or\\ & E_{ji}(t) = 0,\ or\ \mu_i(t) \leqslant d_1,\\ \dfrac{\mu_i(t) - d_1}{b_1}, & c_i(t)V_{ji}(t)E_{ji}(t) \neq 0. \end{cases} \tag{20}$$

$$\gamma_{ji}^*(t) = \begin{cases} 0, & c_i(t) = 0,\ or\ E_{ji}(t) = 0,\\ & or\ \mu_i(t) \leqslant d_2,\\ \dfrac{\mu_i(t) - d_2}{b_2}, & c_i(t)E_{ji}(t) \neq 0. \end{cases} \tag{21}$$

Proof: Solve the partial derivatives of Hamilton function with respect to $\left\{\delta_{ji}(t)\right\}$ and $\left\{\gamma_{ji}(t)\right\}$ and make them 0 as

$$\begin{aligned}\frac{\partial H(t)}{\partial \delta_{ji}(t)} &= d_1V_{ji}(t)E_{ji}(t)c_i(t) + b_1\delta_{ji}(t)V_{ji}(t)E_{ji}(t)c_i(t)\\ &\quad - \mu_i(t)c_i(t)V_{ji}(t)E_{ji}(t) = 0.\end{aligned} \tag{22}$$

$$\begin{aligned}\frac{\partial H(t)}{\partial \gamma_{ji}(t)} &= d_2E_{ji}(t)c_i(t) + b_2\gamma_{ji}(t)E_{ji}(t)c_i(t)\\ &\quad - \mu_i(t)c_i(t)E_{ji}(t) = 0.\end{aligned} \tag{23}$$

Then, the optimal congestion control strategies $\left\{\delta_{ji}^*(t)\right\}$ and $\left\{\gamma_{ji}^*(t)\right\}$ can be solved. ■

From P_2 and (18)-(21), we can see that the optimization objective is the time integral function of the congestion state variables, the control strategies, and the costate variables. These three variables are mutually dependent on each other, whose close loop relationship makes it difficult to attain the explicit solution. Therefore, CCDA algorithm is proposed to obtain the optimal control strategies in Section IV-B.

B. CCDA ALGORITHM

The total continuous time interval $\left[t_0, t_f\right]$ is divided into K discrete subintervals. The length of each subinterval is σ.

$$\sigma = \frac{t_f - t_0}{K}. \tag{24}$$

Algorithm 1 Congestion Control Discretization Algorithm (CCDA)

Input: c_{i0}, $\mu_i(K)$, $\delta_{ji}(t_0)$, $\gamma_{ji}(t_0)$, σ and other network parameters

Output: Optimal congestion control strategies $\left\{\delta_{ji}^*(k+1)\right\}$, $\left\{\gamma_{ji}^*(k+1)\right\}$.

1. **For** $k = 0, k < K, k++$
2. **For** $i = 1, i \leqslant N, i++$
3. **For** $j = 1, j \neq i \,\&\, j \leqslant N, j++$
4. Calculate discrete state variables $c_i(k+1)$ via (29);
5. Calculate discrete costate variables $\mu_i(K-k-1)$ via (30);
6. Calculate optimal congestion control strategies $\delta_{ji}^*(k+1)$ and $\gamma_{ji}^*(k+1)$ according to Theorem 2;
7. **If** $\delta_{ji}^*(k+1) + \gamma_{ji}^*(k+1) < o_{ji}(k+1)$
8. $\gamma_{ji}^*(k+1) = o_{ji}(k+1) - \delta_{ji}^*(k+1)$;
9. **End**
10. **End**
11. **End**

The k-th subinterval is denoted by $k, k \in [0, K-1]$. When σ is extremely small, $\left\{\delta_{ji}^*(k)\right\}$ and $\left\{\gamma_{ji}^*(k)\right\}$ can be considered as the optimal approximate solution as Theorem 2. Based on which, the optimal congestion control strategies can be finally obtained via Algorithm 1.

Theorem 2: The discrete optimal congestion control strategies (defined as CCDA strategies) can be expressed as

$$\delta_{ji}^*(k+1) = \begin{cases} 0, \ c_i(k+1) = 0, \ or \ V_{ji}(k+1) = 0, \ or \\ E_{ji}(k+1) = 0, \ or \ \mu_i(K-k-1) \leqslant d_1, \\ \dfrac{\mu_i(K-k-1) - d_1}{b_1}, \ c_i(k+1) \neq 0, \ and \\ \qquad V_{ji}(k+1)\, E_{ji}(k+1) \neq 0. \end{cases} \tag{25}$$

$$\gamma_{ji}^*(k+1) = \begin{cases} 0, \quad c_i(k+1) = 0 \ or \ E_{ji}(k+1) = 0 \\ \qquad or \ \mu_i(K-k-1) \leqslant d_2, \\ \dfrac{\mu_i(K-k-1) - d_2}{b_2}, c_i(k+1)\, E_{ji}(k+1) \neq 0. \end{cases} \tag{26}$$

$c_i(k+1)$ and $\mu_i(K-k-1)$ are respectively shown as (29) and (30) in APPENDIX A.

Proof: Please refer to APPENDIX A. ■

V. SIMULATION RESULTS

In this section, simulation results are provided to analyze the evolution of congestion states and control strategies of nodes under the formulated model and proposed CCDA algorithm. In addition, the CCDA strategies can reduce network total loss compared with other strategies, proving the effectiveness of the formulated model and proposed CCDA algorithm in reducing delay and mitigating congestion.

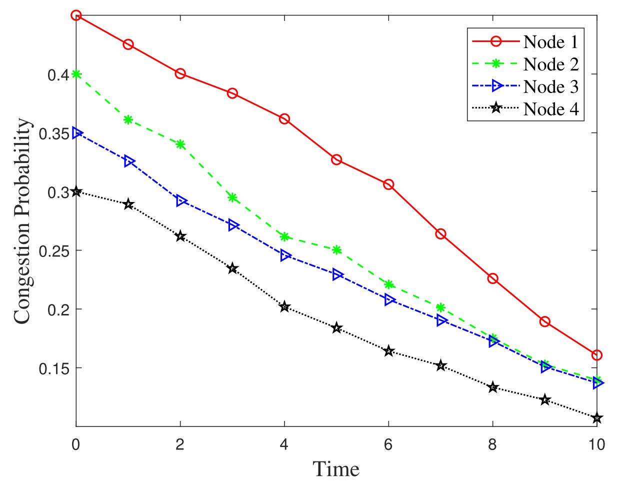

FIGURE 3. The evolution of the congestion states.

The simulation is performed via MATLAB R2019b. Simulation parameters are set as follows. The network includes four nodes, where node 2 and node 3 are deterministic nodes, and node 1 and node 4 are ordinary nodes. The nodes are completely connected to each other. The initial congestion probability of the four nodes is set as $c_{10} = 0.45$, $c_{20} = 0.4$, $c_{30} = 0.35$, $c_{40} = 0.3$, respectively. The delay weight coefficients are $a = 3$, $d_1 = 10$, $d_2 = 4$ and the cost coefficients are $b_1 = 2$, $b_2 = 10$. The congestion deterioration probability is $\lambda_i = 0.384$ with Poisson parameter $\lambda_p = 5$. The number of subintervals is $K = 10$ and the subinterval length is $\sigma = 0.01$. The type of traffic flows among nodes $p_{ji}(t)$ is randomly assigned. The amount of overflowed data among nodes $o_{ji}(t)$ and the initial control strategies are both randomly assigned in [0, 10].

A. NODE CONGESTION STATE AND STRATEGY EVOLUTION

The evolution of congestion states of each node is shown in Fig. 3. It is demonstrated that the congestion probability gradually decreases over time through congestion control. The dropping rates of deterministic node 2 and deterministic node 3 are partially shown in Fig. 4. It is shown that the dropping rate between deterministic nodes is lower than the dropping rate between ordinary nodes and deterministic nodes most of the time, e.g., in most cases, $\delta_{32} \leqslant \delta_{42}$ exists for node 2, and $\delta_{23} \leqslant \delta_{13}$ exists for node 3. However, at subinterval $k = 6$, since BE flow is transmitted from node 3 to node 2, and the amount of overflowed data is more than that from node 4 to node 2, there exists $\delta_{32} > \delta_{42}$ in this period. This is because TS flows are not allowed to be dropped in the proposed model to guarantee the determinacy of them.

In addition, the dropping rates of deterministic node 2 and ordinary node 4 are partially shown in Fig. 5. Results and conclusions similar to those in Fig. 4 can be obtained: the

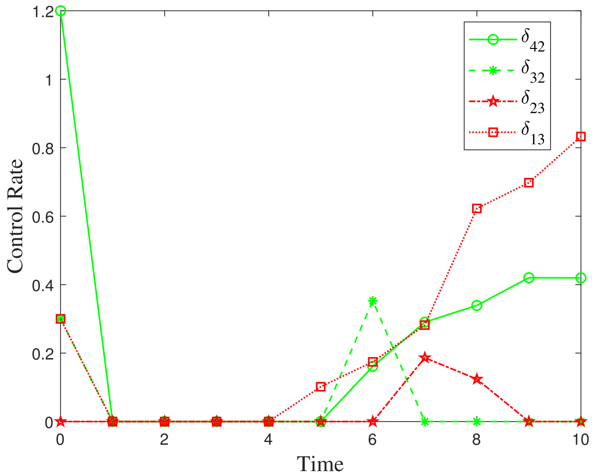

FIGURE 4. The evolution of dropping rate of deterministic nodes.

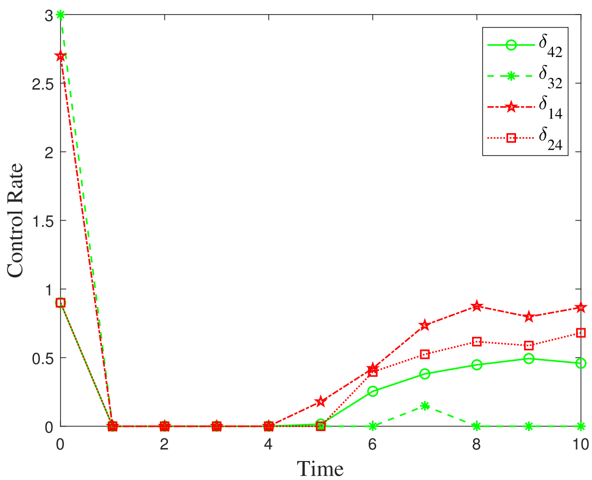

FIGURE 5. The evolution of dropping rate of deterministic node and ordinary node.

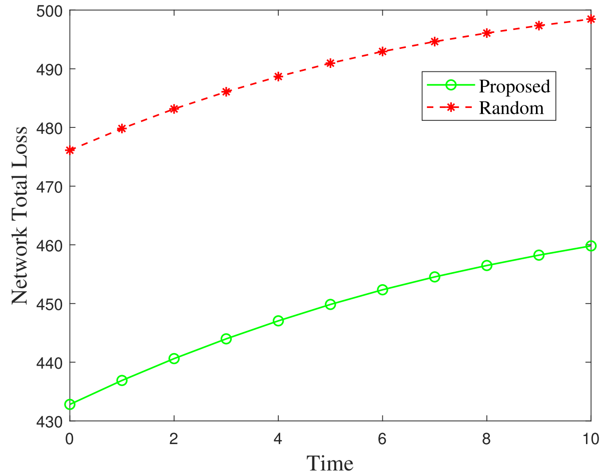

FIGURE 6. Comparison of network total loss.

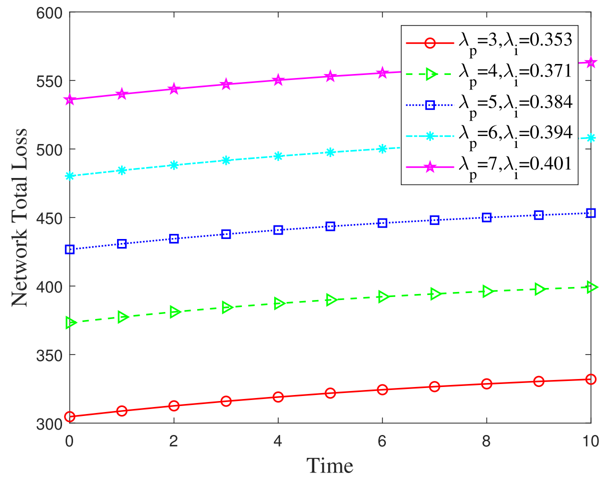

FIGURE 7. Comparison of network total loss with different congestion deterioration probability.

dropping rate between deterministic nodes is lower than the dropping rate between ordinary nodes and deterministic nodes most of the time, e.g., in most cases, $\delta_{32} \leqslant \delta_{42}$ exists for node 2, and $\delta_{24} \leqslant \delta_{14}$ exists for node 4. However, at subinterval $k = 0$, since BE flow is transmitted from node 3 to node 2, and the amount of overflowed data is more than that from node 4 to node 2, there exists $\delta_{32} > \delta_{42}$ in this period. Another remarkable phenomenon is that, the dropping rate of deterministic node 2 is lower than that of ordinary node 4 in most cases. It is revealed that the deterministic nodes and time sensitive flows are more subject to network delay than to strategy control cost. In other words, in order to ensure the deterministic transmission of TS flows, the deterministic nodes may attempt to avoid packet dropping despite having to sacrifice the strategy control cost.

B. NETWORK TOTAL LOSS EVALUATION

The comparison of network total loss with the proposed CCDA strategies and random strategies is shown in Fig. 6. Note that we aim to take a more global view to study the optimal selection of multiple congestion techniques or mechanisms rather than a certain technique like related work in Section II, so maybe no related work is suitable for comparison and we conservatively compare with the random strategies. Compared with the random strategies, the network total loss with the proposed control strategies is lower. This result indicates the proposed model can minimize delay and mitigate congestion with the lowest control cost, because the joint reduction of node congestion degree, node control strategy cost and network delay is formulated as the reduction of network total loss. In addition, the network total loss gradually rises because both the congestion control cost and the network delay in the defined network total loss function accumulate over time. However, the upward trend is more and more gentle, indicating that the proposed model and algorithm can achieve effective congestion control. In summary, the constructed congestion control model in deterministic networks and the proposed CCDA algorithm can achieve the goal of minimizing delay and congestion degree with the lowest control cost.

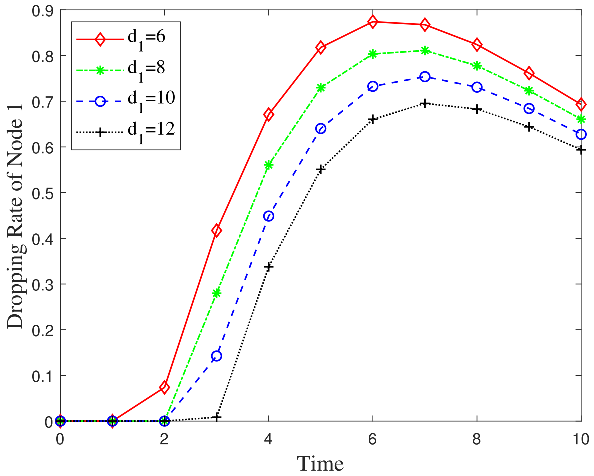

FIGURE 8. Comparison of dropping rate with different delay weight coefficient for dropping strategy.

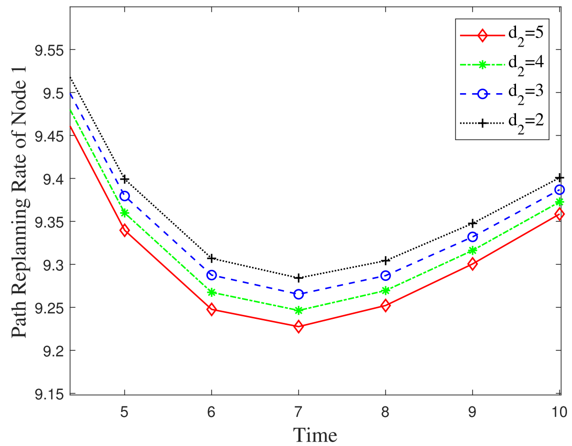

FIGURE 9. Comparison of path replanning rate with different delay weight coefficient for path replanning strategy.

C. MODEL PARAMETER EVALUATION

1) CONGESTION DETERIORATION PROBABILITY

The comparison of the network total loss with different congestion deterioration probability λ_i is shown in Fig. 7. According to (5), as the packet arrival mean λ_p increases from 3 to 7, λ_i will increase from 0.353 to 0.401, and the network total loss increases consequently. This is because as λ_i increases, both the probability and the number of nodes from the uncongested state to the congested state increase. The implementation intensity of the control strategies and network congestion degree increase accordingly, so does the network total loss. Therefore, mechanisms such as traffic shaping at the queue entrances, ECN at downstream nodes can be deployed to control the packet arrival rate of congested nodes, e.g., notifying upstream nodes to reduce the forwarding rate within an acceptable range or limiting the dequeue of unimportant packets to alleviate the network congestion.

2) DELAY WEIGHT COEFFICIENT

The comparison of dropping rate of node 1 with different delay weight coefficient for dropping strategy d_1 is shown in Fig. 8. It is demonstrated that with the increase of d_1, the dropping rate of node 1 decreases. In addition, the comparison of path replanning rate of node 1 with different delay weight coefficient for path replanning strategy d_2 is shown in Fig. 9. Note that we just select partial subintervals to more clearly show the relationship among the four almost coincident lines. It is demonstrated that with the decrease of d_2, the path replanning rate of node 1 increases. In summary, the increase of delay weight coefficient for each control strategy will lead to the decrease of the implementation rate

TABLE 1. Mathematical symbols.

Symbol	Quantity	Remark
i	node sequence	$i \in [1, N]$
t	time	continuous time $t \in [t_0, t_f]$
$v_i(t)$	node type	type of node i at time t, $v_i(t) = \begin{cases} 1, & \textit{deterministic node}, \\ 0, & \textit{ordinary node}. \end{cases}$
$e_{ij}(t)$	path reachability	path reachability from node i to j at time t, $e_{ij}(t) = \begin{cases} 1, & \textit{reachable}, \\ 0, & \textit{unreachable}. \end{cases}$
$p_{ij}(t)$	traffic flow type	traffic flow type from node i to j at time t, $p_{ij}(t) = \begin{cases} 1, & \textit{TS flow}, \\ 0, & \textit{BE flow}. \end{cases}$
$C_i(t)$	congestion state	congestion state of node i at time t, $C_i(t) = \begin{cases} 1, & \textit{congested state } J_i(t), \\ 0, & \textit{uncongested state } F_i(t). \end{cases}$
$c_i(t)$	congestion state variable	congestion probability of node i at time t
$\delta_{ji}(t)$	dropping rate	intensity of dropping strategy to be optimized
$\gamma_{ji}(t)$	path replanning rate	intensity of path replanning strategy to be optimized
λ_p	packet arrival mean	packet arrival follows a Poisson distribution with parameter λ_p
λ_i	congestion deterioration probability	probability of sufficient packet arrival, e.g., $\lambda_i = Pr(X > \lambda_p) = 1 - \sum_{k=0}^{\lambda_p} \lambda_p^k \exp(-\lambda_p)/k!$
$l_i(t)$	delay	refer to (7)
$V_{ji}(t)$	none	$V_{ji}(t) = 1 - v_j(t)\, v_i(t)\, p_{ji}(t)$, just for ease of expression
$E_{ji}(t)$	none	$E_{ji}(t) = e_{ji}(t)\,[1 - c_j(t)]$, just for ease of expression
a	delay weight coefficient	a constant to weight the delay caused by congestion
d_1	delay weight coefficient	a constant to weight the delay caused by dropping strategy
d_2	delay weight coefficient	a constant to weight the delay caused by path replanning strategy
$f_i(t)$	control cost	refer to (10)
b_1	cost coefficient	a constant to weight the control cost of dropping strategy
b_2	cost coefficient	a constant to weight the control cost of path replanning strategy
$g(t)$	instantaneous objective loss	$g(t) = \sum_{i=1}^{N} [l_i(t) + f_i(t)]$, refer to (11)
G	total objective loss	network total loss to be minimized, refer to (12)
$o_{ji}(t)$	overflowed packet	a lower bound to constrain strategies
$\mu_i(t)$	costate variable	each costate variable corresponds to a state variable in optimal control theory
$H(t)$	Hamilton function	refer to (14)
$\delta_{ji}^*(t)$	optimal dropping rate	optimized intensity of dropping strategy
$\gamma_{ji}^*(t)$	optimal path replanning rate	optimized intensity of path replanning strategy
k	subinterval sequence	discrete time $k \in [0, K-1]$
σ	subinterval length	length of each discrete subinterval $\sigma = (t_f - t_0)/K$
c_{i0}	congestion probability	congestion probability of node i at subinterval $k = 0$

of the corresponding strategy. This is because when the delay weight coefficient increases, the delay caused by the control strategy of the same intensity will increase. To maintain the ideal delay and network total loss, the network nodes have to reduce the implementation intensity of the corresponding strategy. Therefore, in practical applications, preferentially choosing the low-delay countermeasures can effectively reduce network congestion and network total loss. And it is also required to trade off the delay performance and cost of these countermeasures to make a comprehensive decision.

VI. CONCLUSION

In this paper, we investigate the network congestion alleviation problem in deterministic networks, and conduct the real-time prediction and macroscopic regulation of congestion control. Firstly, a congestion control system model is constructed, where the congestion states are affected by the node strategies. Then, an optimization problem is formulated to jointly alleviate network congestion and minimize delay with the lowest control cost. By proposing a CCDA algorithm based on the optimal control method, the CCDA strategies are derived, achieving the continuous state evolution and real-time strategy selection. Finally, simulation results show that with the formulated model and proposed CCDA algorithm, the node congestion probability gradually decreases over time, and the network total loss is lower than the baseline model. It is revealed that this solution can effectively mitigate network congestion, reduce network delay, and achieve the trade off between control cost and network performance.

$$
\begin{aligned}
\frac{c_i(k+1)-c_i(k)}{\sigma} &= \left[1-c_i(k+1)\right]\lambda_i \\
&\quad - c_i(k+1)\sum\nolimits_{j=1,j\neq i}^{N}\big[E_{ji}(k)\,\gamma_{ji}(k) \\
&\quad + E_{ji}(k)\,\delta_{ji}(k)\,V_{ji}(k)\big].
\end{aligned} \tag{27}
$$

$$
\begin{aligned}
\frac{\mu_i(K-k)-\mu_i(K-k-1)}{\sigma} &= -a+\mu_i(K-k-1)\,\lambda_i \\
&\quad - \sum\nolimits_{j=1,j\neq i}^{N}\left[d_1\delta_{ji}(k)+\frac{b_1}{2}\delta_{ji}^2(k)\right]V_{ji}(k)E_{ji}(k) \\
&\quad - \sum\nolimits_{j=1,j\neq i}^{N}\left[d_2\gamma_{ji}(k)+\frac{b_2}{2}\gamma_{ji}^2(k)\right]E_{ji}(k) \\
&\quad + \mu_i(K-k-1)\sum\nolimits_{j=1,j\neq i}^{N}\big[\delta_{ji}(k)\,V_{ji}(k)+\gamma_{ji}(k)\big]E_{ji}(k) \\
&\quad + \sum\nolimits_{j=1,j\neq i}^{N}c_j(k)\,e_{ij}(k)\big[d_1\delta_{ij}(k)\,V_{ij}(k)+d_2\gamma_{ij}(k)\big] \\
&\quad + \sum\nolimits_{j=1,j\neq i}^{N}c_j(k)\,e_{ij}(k)\left[\frac{b_1}{2}\delta_{ij}^2(k)\,V_{ij}(k)+\frac{b_2}{2}\gamma_{ij}^2(k)\right] \\
&\quad - \sum\nolimits_{j=1,j\neq i}^{N}\mu_j(K-k)\,c_j(k)\,e_{ij}(k)\big[\delta_{ij}(k)\,V_{ij}(k)+\gamma_{ij}(k)\big].
\end{aligned} \tag{28}
$$

$$
c_i(k+1) = \frac{c_i(k)+\sigma\lambda_i}{1+\sigma\left\{\lambda_i+\sum_{j=1,j\neq i}^{N}\big[\delta_{ji}(k)\,V_{ji}(k)+\gamma_{ji}(k)\big]E_{ji}(k)\right\}}. \tag{29}
$$

$$
\begin{aligned}
\mu_i(K-k-1) &= \{\mu_i(K-k)+\sigma a \\
&\quad + \sum\nolimits_{j=1,j\neq i}^{N}\sigma E_{ji}(k)\big[d_1\delta_{ji}(k)\,V_{ji}(k)+d_2\gamma_{ji}(k)\big] \\
&\quad + \sum\nolimits_{j=1,j\neq i}^{N}\frac{\sigma}{2}E_{ji}(k)\left[b_1\delta_{ji}^2(k)\,V_{ji}(k)+b_2\gamma_{ji}^2(k)\right] \\
&\quad - \sum\nolimits_{j=1,j\neq i}^{N}\sigma c_j(k)\,e_{ij}(k)\big[d_1\delta_{ij}(k)\,V_{ij}(k)+d_2\gamma_{ij}(k)\big] \\
&\quad - \sum\nolimits_{j=1,j\neq i}^{N}\frac{\sigma}{2}c_j(k)\,e_{ij}(k)\left[b_1\delta_{ij}^2(k)\,V_{ij}(k)+b_2\gamma_{ij}^2(k)\right] \\
&\quad + \sum\nolimits_{j=1,j\neq i}^{N}\sigma\mu_j(K-k)\,c_j(k)\,e_{ij}(k)\,\big[\delta_{ij}(k)\,V_{ij}(k)+\gamma_{ij}(k)\big]\} \\
&\quad / \{1+\sigma\lambda_i+\sigma\sum\nolimits_{j=1,j\neq i}^{N}\big[\delta_{ji}(k)\,V_{ji}(k)+\gamma_{ji}(k)\big]E_{ji}(k)\}.
\end{aligned} \tag{30}
$$

In summary, this research provides a general modeling thought and a reasonable mathematical optimization method for the research of deterministic networks especially the congestion state evolution analysis. In addition, it has a theoretical and practical guiding significance for the real-time selection of congestion control mechanisms and the global regulation of network congestion situation.

For the future work, multilevel congestion states for network nodes and the accurate mapping between actual network configurations and model parameters will be considered. More realistic network topology may also be included in the simulation, so as to improve the authenticity and accuracy of the congestion control model.

APPENDIX A
PROOF OF THEOREM 2

The discrete forms of (18) and (19) can be expressed as (27) and (28), shown at the bottom of the previous page, respectively.

Therefore, $c_i(k+1)$ and $\mu_i(K-k-1)$ can be derived as (29) and (30), shown at the bottom of the previous page, respectively.

APPENDIX B
MATHEMATICAL SYMBOLS

Please refer to Table 1.

REFERENCES

[1] *Deterministic Networking Use Cases*, document RFC 8578, 2019.
[2] Q. Gao, J. Li, and Y. Liu, "Research on deterministic cyber-physical systems architecture for 6G evolution," *Appl. Electron. Techn.*, vol. 48, no. 12, pp. 11–14, Dec. 2022.
[3] A. Badar, D. Z. Lou, U. Graf, C. Barth, and C. Stich, "Intelligent edge control with deterministic-IP based industrial communication in process automation," in *Proc. 15th Int. Conf. Netw. Service Manage. (CNSM)*, Halifax, NS, Canada, Oct. 2019, pp. 1–7.
[4] *Deterministic Networking Architecture*, document RFC 8655, 2019.
[5] Z. Ding, W. Li, Y. Cheng, Y. Xu, Y. Dai, Y. Wang, and P. Yu, "Slice network framework and use cases based on FlexE technology for power services," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Harbin, China, Jun. 2021, pp. 57–62.
[6] T. Huang, S. Wang, Y. Huang, Y. Zheng, J. Liu, and Y. Liu, "Survey of the deterministic network," *J. Commun.*, vol. 40, no. 6, p. 17, Jun. 2019.
[7] X. Wang, H. Yao, T. Mai, Z. Xiong, F. Wang, and Y. Liu, "Joint routing and scheduling with cyclic queuing and forwarding for time-sensitive networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 3793–3804, Mar. 2023.
[8] S. Wang, B. Wu, C. Zhang, Y. Huang, T. Huang, and Y. Liu, "Large-scale deterministic IP networks on CENI," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Vancouver, BC, Canada, May 2021, pp. 1–6.
[9] B. Liu, S. Ren, C. Wang, V. Angilella, P. Medagliani, S. Martin, and J. Leguay, "Towards large-scale deterministic IP networks," in *Proc. IFIP Netw. Conf. (IFIP Networking)*, Espoo, Helsinki, Finland, Jun. 2021, pp. 1–9.
[10] T. Li and Y. Cai, "Joint routing and scheduling for deterministic networking: A segment routing approach," in *Proc. IEEE 23rd Int. Conf. High Perform. Switching Routing (HPSR)*, Taicang, Jiangsu, China, Jun. 2022, pp. 189–194.
[11] B. Wu, J. Wang, Y. Wang, W. Tan, and Y. Huang, "Achieving deterministic service in mobile edge computing (MEC) networks," in *Proc. IEEE 29th Int. Conf. Netw. Protocols (ICNP)*, Dallas, TX, USA, Nov. 2021, pp. 1–6.
[12] C. Pan, S. Zhang, C. Zhao, H. Shi, Z. Kong, and X. Cui, "A novel active queue management algorithm based on average queue length change rate," *IEEE Access*, vol. 10, pp. 75558–75570, 2022.
[13] S. Shahzad, E.-S. Jung, J. Chung, and R. Kettimuthu, "Enhanced explicit congestion notification (EECN) in TCP with p4 programming," in *Proc. Int. Conf. Green Human Inf. Technol. (ICGHIT)*, Hanoi, Vietnam, Feb. 2020, pp. 35–40.
[14] G. Peng, S. Wang, Y. Huang, R. Huo, T. Huang, and Y. Liu, "Traffic shaping at the edge: Enabling bounded latency for large-scale deterministic networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Montreal, QC, Canada, Jun. 2021, pp. 1–6.
[15] Y. Liu, Z. Cheng, J. Ren, and D. Yang, "Joint routing and scheduling for CQF," in *Proc. 7th Int. Conf. Comput. Commun. Syst. (ICCCS)*, Wuhan, China, Apr. 2022, pp. 1–5.
[16] J. Krolikowski, S. Martin, P. Medagliani, J. Leguay, S. Chen, X. Chang, and X. Geng, "Joint routing and scheduling for large-scale deterministic IP networks," *Comput. Commun.*, vol. 165, pp. 33–42, Jan. 2021.
[17] Y. Huang, S. Wang, T. Feng, J. Wang, T. Huang, R. Huo, and Y. Liu, "Towards network-wide scheduling for cyclic traffic in IP-based deterministic networks," in *Proc. 4th Int. Conf. Hot Inf.-Centric Netw. (HotICN)*, Nanjing, China, Nov. 2021, pp. 117–122.
[18] S. Chen, J. Leguay, S. Martin, and P. Medagliani, "Load balancing for deterministic networks," in *Proc. IFIP Netw. Conf. (Networking)*, Paris, France, Jun. 2020, pp. 785–790.
[19] Z. Liu, J. Sun, S. Hu, and X. Hu, "An adaptive AQM algorithm based on a novel information compression model," *IEEE Access*, vol. 6, pp. 31180–31190, 2018.
[20] J. Escudero-Sahuquillo, P. J. Garcia, F. J. Quiles, J. Flich, and J. Duato, "Cost-effective congestion management for interconnection networks using distributed deterministic routing," in *Proc. IEEE 16th Int. Conf. Parallel Distrib. Syst.*, Shanghai, China, Dec. 2010, pp. 355–364.
[21] J. Liu, J. Huang, W. Jiang, Z. Li, Y. Li, W. Lyu, W. Jiang, J. Zhang, and J. Wang, "End-to-end congestion control to provide deterministic latency over internet," *IEEE Commun. Lett.*, vol. 26, no. 4, pp. 843–847, Apr. 2022.
[22] X. Yang, W. Wang, J. Han, K. Xue, and Z. Zhang, "Congestion control with deterministic service delay guarantee," in *Proc. IEEE 19th Int. Conf. Mobile Ad Hoc Smart Syst. (MASS)*, Denver, CO, USA, Oct. 2022, pp. 555–563.
[23] L. Zhang, K. Zhu, Y. Cui, and Y. Jiang, "QoS-aware congestion control with online learning," *China Commun.*, vol. 20, no. 2, pp. 168–178, Feb. 2023.
[24] A. Grigorjew, F. Metzger, T. Hoßfeld, J. Specht, F.-J. Götz, F. Chen, and J. Schmitt, "Constant delay switching: Asynchronous traffic shaping with jitter control," in *Proc. IFIP Netw. Conf. (IFIP Networking)*, Catania, Italy, Jun. 2022, pp. 1–9.
[25] Y. Mo, Z. Yang, H. Liu, and T. He, "Global cyclic queuing and forwarding mechanism for large-scale deterministic networks," in *Proc. IEEE 23rd Int. Conf High Perform. Comput. Commun., 7th Int. Conf Data Sci. Syst., 19th Int. Conf Smart City; 7th Int. Conf. Dependability Sensor, Cloud Big Data Syst. Appl. (HPCC/DSS/SmartCity/DependSys)*, Haikou, Hainan, China, Dec. 2021, pp. 275–282.
[26] T. Fukui, Y. Sakaue, K. Minami, and T. Taniguchi, "Delay-based shaper with dynamic token bucket algorithm for deterministic networks," *IEEE Access*, vol. 10, pp. 114424–114433, 2022.
[27] *IEEE Standard for Local and Metropolitan Area Networks—Virtual Bridged Local Area Networks Amendment 12: Forwarding and Queuing Enhancements for Time-Sensitive Streams*, IEEE Standard 802.1Qav-2009, IEEE Standard 802.1Q-2005, 2010.
[28] *IEEE Standard for Local and Metropolitan Area Networks—Bridges and Bridged Networks—Amendment 25: Enhancements for Scheduled Traffic*, IEEE Standard 802.1Q-2014/Cor 1-2015, 2016.
[29] M. Kim, D. Hyeon, and J. Paek, "ETAS: Enhanced time-aware shaper for supporting nonisochronous emergency traffic in time-sensitive networks," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 10480–10491, Jul. 2022.
[30] *IEEE Standard for Local and Metropolitan Area Networks—Bridges and Bridged Networks—Amendment 34: Asynchronous Traffic Shaping*, IEEE Standard 802.1Qcx-2020, 2020.
[31] *IEEE Standard for Local and Metropolitan Area Networks—Bridges and Bridged Networks—Amendment 26: Frame Preemption*, IEEE Standard 802.1Qbu-2016, 2016.

[32] H. Wang, J. Zhao, X. Liu, and C. Yin, "Enhanced frame preemption in image and video transmission over time sensitive networks," in *Proc. IEEE 33rd Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Kyoto, Japan, Sep. 2022, pp. 1185–1190.

[33] *IEEE Standard for Local and Metropolitan Area Networks—Bridges and Bridged Networks—Amendment 29: Cyclic Queuing and Forwarding*, IEEE Standard 802.1Qch-2017, 2017.

[34] L. Qiang, B. Liu, D. Yu, and C. Wang, "Large-scale deterministic network forwarding technology," *Telecommun. Sci.*, vol. 35, no. 9, pp. 12–19, Sep. 2019.

[35] Y. Huang, S. Wang, T. Huang, and Y. Liu, "Cycle-based time-sensitive and deterministic networks: Architecture, challenges, and open issues," *IEEE Commun. Mag.*, vol. 60, no. 6, pp. 81–87, Jun. 2022.

[36] K. Alexandris, P. Pop, and T. Wang, "Configuration and evaluation of multi-CQF shapers in IEEE 802.1 time-sensitive networking (TSN)," *IEEE Access*, vol. 10, pp. 109068–109081, 2022.

[37] B. Sun and B.-Z. Guo, "Convergence of an upwind finite-difference scheme for Hamilton–Jacobi–Bellman equation in optimal control," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 3012–3017, Nov. 2015.

Q. GAO received the master's degree from the Beijing University of Posts and Telecommunications (BUPT), in 2022. She is currently a Researcher with China Telecom Research Institute. Her current research interests include mobile communications and networks, core network architecture design, and deterministic networks, with an emphasis on deterministic networks and key technologies towards B5G/6G.

J. LI received the master's degree from the Beijing University of Posts and Telecommunications (BUPT). She is currently a Professorate Senior Engineer with China Telecom Research Institute. Her current research interests include the standardization and technique evolution of mobile networks, especially in the area of architecture evolution and service deployment.

Y. LIU received the B.S. and Ph.D. degrees from the Beijing University of Posts and Telecommunications (BUPT), in 2006 and 2011, respectively. She is currently a Senior Researcher with China Telecom Research Institute, with a focus on the RAN architecture design towards 6G. Her current research interests include virtualization and intelligence for wireless communications.

Y. XING received the master's degree from Capital Normal University, in 1999. She is currently a Senior Engineer with China Telecom Research Institute. Her research interests include mobile communication network architecture and key technologies, and network operation and management.