Feedback perimeter control with online estimation of maximum throughput for an incident-affected road network

Jiawen Wang, Xiaozheng He, Srinivas Peeta & Xiaoguang Yang

To cite this article: Jiawen Wang, Xiaozheng He, Srinivas Peeta & Xiaoguang Yang (2020): Feedback perimeter control with online estimation of maximum throughput for an incident-affected road network, Journal of Intelligent Transportation Systems, DOI: 10.1080/15472450.2020.1797501

To link to this article: https://doi.org/10.1080/15472450.2020.1797501

Published online: 27 Jul 2020.

Submit your article to this journal

Article views: 47

View related articles

View Crossmark data
Feedback perimeter control with online estimation of maximum throughput for an incident-affected road network

Jiawen Wang, Xiaozheng He, Srinivas Peeta, and Xiaoguang Yang

ABSTRACT

This study develops a feedback perimeter control strategy to maximize the throughput of an incident-affected network. The proposed perimeter control strategy is innovative in two aspects. First, the control variables, i.e., the inflow rates to the controlled subnetwork within the incident-affected network, are adjusted based on the online estimation of maximum network throughput that is updated dynamically using real-time traffic data and road vulnerability. The incident-dependent network throughput provides the perimeter control a more legitimate control target. Second, the proposed perimeter control strategy applies the proportional-integral-derivative controller, which enhances control stability given the dynamically-updated control target. The results of simulation experiments demonstrate that the proposed strategy can enhance the average speed and reduce the total delay of the incident-affected traffic.

ARTICLE HISTORY

Received 2 February 2019
Revised 9 July 2020
Accepted 10 July 2020

KEYWORDS

maximum network throughput; online estimation; perimeter control; proportional-integral-derivative controller

Introduction

Because of the high spatial correlation between high-density intersections in the urban network, traffic incidents, referred to as “all events which affect (or may affect) the capacity of the road and hinder the smooth flow of traffic” (Steenbruggen et al., 2014), can cause severe congestion that may propagate to a large area of the transportation network and result in gridlock for a network with heavy traffic demand (Qi et al., 2018). Influenced by sudden traffic incidents, the interrupted flow of urban network presents complex dynamics, randomness and instability. It is urgent to provide real-time control strategies to cope with the short-term mismatch of traffic supply and demand caused by the incidents, thereby improving travel reliability and overall network throughput.

To mitigate the congestion caused by traffic incidents, efficient use of existing infrastructure through route guidance and traffic control is preferred for a sustainable urban transportation system (Ding et al., 2017; Sirmatel & Geroliminis, 2018). Route guidance seeks to manage the demand to leverage the existing infrastructure; however, its efficiency is sensitive to travelers’ response (Paz & Peeta, 2009a, 2009b). Thereby, the resulting traffic pattern may deviate from the expected one. Furthermore, route guidance strategies may not be able to mitigate traffic congestion caused by incidents for traffic networks with particular topological structures, e.g., a central business district where travelers do not have acceptable alternative routes to their destinations.

Traffic control strategies, including traffic signals at urban streets, are preferable to mitigate congestion after an incident. Real-time traffic control strategies are deemed to be an efficient and cost effective way to ameliorate traffic conditions and prevent gridlock phenomena in cities (Kouvelas et al., 2017; Xu et al., 2019, 2020). Although many methodologies have been developed for real-time signal control over the last decades, see Papageorgiou et al. (2003) for a good review, the design of efficient control strategies for large-scale urban networks affected by traffic incidents remains a significant challenge. To enhance the throughput of urban networks affected by incidents, this study proposes an approach to update the control target dynamically and develops a perimeter control strategy.
Traffic signal control strategies can be classified into data-driven and model-based approaches. Data-driven approaches adjust the traffic signal parameters using traffic and signal data. Widely-used data-driven traffic signal control systems include SCOOT (Hunt et al., 1982) and SCATS (Lowe, 1992). These systems become less efficient under saturated traffic conditions induced by situational events (Keyvan-Ekbatani et al., 2012; L. Zhang et al., 2013). By contrast, model-based approaches can better manage the saturated traffic, e.g., model predictive control (Aboudolas et al., 2010; Ma et al., 2020; Gartner, 1983; Henry et al., 1984; Li & Sun, 2019; Lin et al., 2012; Mirchandani & Head, 2001), linear quadratic controller/integrator (Aboudolas & Geroliminis, 2013; Diakaki et al., 2002), and others (Khattak et al., 2020; Yao et al., 2020).

However, because of the large number of signalized intersections and urban links, modeling the traffic flow dynamics of an urban network is a complex task (Geroliminis et al., 2013). Given the high computational requirement, the model predictive control methods might have difficulty in ensuring real-time feasibility when applied to a large network (Aboudolas et al., 2010; Lo et al., 2001). Hence, instead of micromodeling approach, the macroscopic fundamental diagram (MFD) aims at simplifying the micromodeling task of the urban network (e.g., Aboudolas & Geroliminis, 2013; Geroliminis et al., 2013; Haddad, 2017; Keyvan-Ekbatani et al., 2012, 2013, 2015; Kouvelas et al., 2017; Wan et al., 2020; Wu et al., 2018, and others).

MFD-based control strategies can be utilized to introduce effective control strategies to improve mobility and decrease delays in large urban networks by considering network performance holistically. MFD-based control strategies are built upon the relationship between vehicular accumulation and exit flow of a traffic network (Ardekani & Herman, 1987; Daganzo, 2007; Dinopoulou et al., 2005; Geroliminis & Sun, 2011b; Godfrey, 1969; Herman & Prigogine, 1979; Leclercq et al., 2014; Wardrop, 1968). Based on the notion of MFD, Daganzo (2007) first proposed the network exit function to develop a control rule to maximize network exit flow, and Geroliminis et al. (2013) developed further for two regions. Then, the notion of MFD was applied to develop perimeter control strategies, with the goal of maximizing network throughput that is estimated using MFD.

The MFD-based perimeter control includes two categories: (i) nonlinear models using the model predictive control (e.g., Geroliminis et al., 2013; Haddad et al., 2013; Hajiahmadi et al., 2015), and (ii) linearized models using the feedback control (e.g., Aboudolas & Geroliminis, 2013; Keyvan-Ekbatani et al., 2012). Since it is a difficult task to conduct the stability analysis and control synthesis of a nonlinear system (Haddad & Geroliminis, 2012), the stability of feedback controller for a nonlinear system, even for a simple class of nonlinear system such as piecewise affine systems, is hard to be guaranteed (Bemporad et al., 2000). In this paper, we aim at designing a stable feedback controller for the incident-affected region based on the linearized models following the classical approach.

MFD-based feedback control methods aim at maximizing the network throughput using stable and fast convergence linear system control methods. Zhang et al. (2010) proposed a bang-bang type control strategy to optimize the aggregate vehicular accumulation of a controlled region. Li et al. (2012) proposed a fixed-time signal perimeter control to relieve congestion in a traffic network. Keyvan-Ekbatani et al. (2012) proposed a real-time feedback-based perimeter control strategy based on the MFD notion, which applies a static control target. Keyvan-Ekbatani et al. (2013) developed a perimeter control approach whose efficiency preserves even less traffic data are used. Some linearized feedback control strategies consider the heterogeneity of a large-scale urban network, including the concentric-boundary gating strategy (Haddad & Mirkin, 2017; Keyvan-Ekbatani et al., 2013; Kouvelas et al., 2017; Wu et al., 2018), sub-network approach (Dong et al., 2019; Geroliminis et al., 2013), three-dimensional MFD method (Geroliminis et al., 2014), approaches with time-delay (Haddad & Zheng, 2020; Wan et al., 2020), multi-region robust control (Ampountolas et al., 2017; Haddad, 2017; Haddad & Mirkin, 2017), and approaches combined with adaptive signal systems (Keyvan-Ekbatani et al., 2019). Haddad and Shraiber (2014) proposed a robust perimeter controller, which is robust under various traffic conditions from free flow to gridlock. Zhong et al. (2018) deal with the scattering in the MFD caused by stochastic and dynamic travel demand.

MFD-based feedback control methods determine the maximum network throughput (MNT) as the control target based on a fixed MFD curve no matter how the network topological structure, signal plans, travelers’ route choice, or road characteristics change. However, the MNT may change caused by disturbances on road network (e.g., road blockage due to an incident). Mariotte et al. (2017) measured the inaccuracy of the MFD estimation approach when the demand varies rapidly. Nguyen et al. (2016) stated that the MFD will present a variation over time, and introduced a model for a continuous updating of the
MFD in traffic control applications, which inspired the proposed MNT estimation approach for incident-affected networks. Mohebifard et al. (2019) proposed that properties of an MFD may change as signal timing parameters vary in the network and filled this gap by developing an optimization program and a solution technique for cooperative traffic signal and perimeter control in semi-connected urban-street networks.

When travelers routing behavior changes in response to incidents, a traffic jam can occur at a lower vehicle accumulation (Zhang et al., 2013). As the vehicle accumulation increases to an over-saturated condition, the MNT reduces forming a metastable hysteresis loop in the MFD (Gayah & Daganzo, 2011; Geroliminis & Sun, 2011a, 2011b; Mariotte et al., 2017; Saberi & Mahmassani, 2013). As the MNT is metastable (Krauss et al., 1997; Nagatani, 2002; Nguyen et al., 2016), perimeter control strategies using a predetermined control target will become ineffective under incidents when the incident-affected MNT differs significantly from the MNT under normal conditions. This observation motivates the need to develop an online estimation method for the MNT-based on the limited real-time traffic data in the subnetwork under control (i.e., controlled subnetwork within the incident-affected network). The estimated MNT, based on an MNT deformation mechanism and a notion of link vulnerability, provides a more tangible control target to improve the effectiveness of the perimeter control strategy.

This paper applies the notion of adaptive MNT, which varies with incidents, to facilitate the design of a feedback controller for subnetworks. The proposed control strategy follows the model-based approach, in which the reference model is updated in real time using available traffic data. The control strategy applies a proportional-integral-derivative (PID) feedback controller, which factors the impacts of incident through the dynamically-updated MNT. This PID feedback controller has the advantages of fast response time (Desborough & Miller, 2002). The proposed perimeter control is applied to the urban network constructed in a microscopic simulation environment. The simulation results illustrate that, after traffic incidents occur, the proposed perimeter control strategy can enhance the average traffic speed and reduce the average delay of the whole incident-affected network that including the subnetwork under control.

The contributions of the paper are as follows. First, the study proposes an online approach to update the MNT dynamically, providing control targets that are responsive to the unfolding traffic conditions. The approach estimates the MNT, instead of the whole MFD curve using only a small amount of real-time traffic data by leveraging the notion of link vulnerability. Second, the study develops an innovative perimeter control strategy based on the PID controller with a well-tuned derivative term, which is suitable for online implementation with fast response time and stable convergence. The stability of PID controllers (Kiam Heong et al., 2005) enables the proposed online approach to avoid sharp fluctuations of inflow at each gated link and the regulation of controlled intersections when the control target varies. Third, the study investigates the performance of the incident-affected network in terms of total traffic delay and average speed using the proposed control strategy. It provides insights on effective signal control design based on dynamically-updated MNT. Using the adaptive MNT, the dynamically-updated control target improves the realized network throughput when traffic is moderately or severely affected by incidents. It leads to a reduction in average delay and an increase in traffic speed. The simulation results demonstrate that the proposed approach performs better than the fixed-MNT-based perimeter control strategies for the whole incident-affected network, including the subnetwork under control.

The remainder of this paper is organized as follows. Section 2 provides background information for the development of the perimeter control strategy. Section 3 presents the feedback mechanism applied to the perimeter control strategy. Section 4 describes the simulation settings and discusses the numerical experiment results. Section 5 provides concluding comments.

**Preliminaries**

**Maximum network throughput**

In general, the throughput of a traffic network is the sum of traffic volumes leaving the network per unit time (i.e., including vehicles arriving at the destination within network and vehicles driving out of the incident-affected region). This definition has been applied in many other studies (Keyvan-Ekbatani et al., 2012; Mohebifard et al., 2019; Papageorgiou et al., 2003).

The motivation to estimate incident-affected MNT is threefold. (i) Throughput of a traffic network might be measured directly when there were enough detectors (including point detector and floating car). However, the number of detectors is insufficient and the distribution is uneven in current real-world traffic networks. (ii) Based on the estimated real-time throughput, obtaining the MNT at the normal state
(i.e., the state of the network without the incidents and the gridlock links) of the road network requires time-consuming data analyses. (iii) The MNT affected by incidents is difficult to measure. Therefore, this paper develops an online approach for MNT estimation, rather than the whole MFD, to provide effective control target for the subnetwork.

In this study, the MNT is based on the definition of Keyvan-Ekbatani et al. (2012). The x-axis represents the Total Time Spent (TTS) per hour by vehicles traveling in the subnetwork (in passenger car unit h per h, or pcu h per h), which can be viewed as a proxy for vehicular accumulation. The y-axis represents the Total Traveled Distance (TTD) per hour of vehicles in the subnetwork (in pcu km per h). Following the research of Keyvan-Ekbatani et al. (2012), this study assumes that the sum of traffic volumes leaving the network of the incident-affected region (i.e., the throughput of the incident-affected network) is proportional to the TTD of the network. Therefore, TTD per hour of vehicles in the network (in pcu km per h), is the proxy for vehicle throughput in this paper.

This study divides the perimeter control time horizon into $U$ time intervals of equal length $T$. Parameter $i = 0, 1, \ldots, I$ is the index of time interval. Denote $Z$ as the set of subnetwork links and $M \subseteq Z$ as the set of links with traffic sensors. A critical task is to update the MNT dynamically. To obtain MNT, the relationship between network throughput and the number of vehicles traveling inside the subnetwork is dynamically updated using the latest link length and network flow information on TTS and TTD defined by the following equations:

$$TTS(i) = \sum_{z \in M} \frac{T \cdot \hat{N}_z(i)}{T} = \sum_{z \in M} \hat{N}_z(i) = \hat{N}(i) \quad (1)$$

$$TTD(i) = \sum_{z \in M} \frac{T \cdot \hat{q}_z(i) \cdot L_z}{T} = \sum_{z \in M} \hat{q}_z(i) \cdot L_z \quad (2)$$

where $z$ is link index. Notation $\hat{q}_z(i)$ represents the measured flow on link $z$ during time interval $i$. Notation $L_z$ represents the length of link $z$; $\hat{N}_z(i)$ represents the estimated number of vehicles on link $z$ during time interval $i$; $\hat{N}(i)$ represents the total number of vehicles traveling on links belong to set $M$. $TTS(i)$ is the estimated total vehicle-time spent per hour during interval $i$. $TTD(i)$ is the estimated total traveled distance per hour during time interval $i$. When the detector is located in the middle of a link, ref. Vigos and Papageorgiou (2010) proposes the following equation to estimate $\hat{N}_z(i)$ using occupancy data from loop detectors:

$$\hat{N}_z(i) = \frac{L_z}{100 \cdot \lambda} \cdot \alpha_z(i) \quad (3)$$

where $\alpha_z(i)$ is the measured time-occupancy (in %) on link $z$ during interval $i$; $\mu_z$ is the number of lanes on link $z$, and $\lambda$ is the average effective vehicle length. The solid curve in Figure 1 represents the relationship between network throughput ($TTD$) and the number of vehicles ($TTS$) of the network under normal traffic conditions, which is characterized based on Equations (1)–(3). The dotted line in Figure 1 represents a deformed relationship between the throughput and the vehicle accumulation. Our research is focused on how to estimate the affected MNT.

Figure 1. Relationship between normal MNT and the affected MNT.
The maximum proxy value of the throughput, e.g., represented by total network flow (Aboudolas & Geroliminis, 2013), total traveled distance (Keyvan-Ekbatani et al., 2012), network outflow (Ramezani et al., 2015), is commonly considered as the control target in existing perimeter control strategies. However, the perimeter control strategies may overestimate the MNT of the subnetwork under control (Geroliminis & Boyaci, 2012). It signifies the need to dynamically update the MNT to ensure the effectiveness of perimeter control strategies for subnetworks.

To an urban road network, the MFD may be quite stable from day to day, particularly if the traffic load is homogeneously distributed in network links (Geroliminis & Sun, 2011a, 2011b). Under recurrent traffic conditions, the normal MFD of the subnetwork (regarded as a homogeneous network) can be estimated through long-term historical field data to obtain the normal MNT (see Figure 1). For large-scale networks, clustering methods should be used to divide the network into multiple homogeneous subnetworks (Saeedmanesh & Geroliminis, 2016). These clustering methods, including the unconstrained clustering method (Ji & Geroliminis, 2012) and the hierarchical clustering method (Guo, 2008), ensure the feasibility when applying MFD-based control.

When an incident occurs, the traffic load may not be distributed homogeneously. Under this inhomogeneous condition, the MNT cannot maintain at the highest value (Ramezani et al., 2015). Meanwhile, a small amount of real-time traffic data is insufficient to reconstruct the whole affected MFD with possible transient phenomena such as hysteresis. Nevertheless, it is critical to estimate the incident-affected MNT to facilitate the design of effective congestion alleviation operations. To address this difficulty, this paper proposes a viable approach to estimate the incident-affected MNT by incorporating a transient relationship between network throughput and the accumulation under incident condition.

**Control scheme**

This study seeks to mitigate the traffic congestion caused by incidents via perimeter control. Our approach seeks to maximize the traffic throughput by limiting the number of vehicles accessing the subnetwork.

In perimeter control, traffic signals are modified at the gated links of the subnetwork. The gated inflows lead to a decrease in vehicular accumulation, and thereby balance the network supply and demand. Perimeter control seeks to maintain the vehicular accumulation of the incident-affected region close to the optimum value. The potential gated links for perimeter control are those close to the border of the subnetwork. If all boundary links of the incident-affected region are under control, perimeter control can achieve the best performance. However, the perimeter control strategy must choose gated links to satisfy the queue length constraint. This study selects the gated links using the following two principles. First, the lengths of the storage bays of the gated links are long enough to accommodate vehicle queues. Second, the gated links that satisfy the first principle are selected on the boundary of the subnetwork.

Figure 2 shows the schematic of the proposed control strategy. A signal control strategy can be deployed to reduce the total inflow of the subnetwork to mitigate congestion. Part (i) in Figure 2 represents the incident-affected network in the proposed feedback control approach. Notations $q_{in1}, q_{in2}, \ldots, q_{inn}$ in part (i) represent the incident inflows into the subnetwork. Part (ii) shows the detector data, which includes the network density estimation and incident information. Part (iii) illustrates the PID regulator, which compares the desired density estimation with the current network density to determine the affected MNT estimation. The selected gated links based on the detected incident information are indicated in part (i).
(i) denote the inflows at the gated links. Notations $g_1$, $g_2$, $g_3$ in part (i) of Figure 2 denote green phases of the gated links. The gated inflows are determined by the feedback controller in Part (iii).

Part (ii) in Figure 2 represents the traffic state monitoring module to archive traffic density. For decades, research has been dedicated to establishing traffic incident detection systems to identify the time, locations, and types of traffic incidents in real time (Gu et al., 2016, Hashemi & Abdelghany, 2018, Singh & Mohan, 2019). Based on the incident detection systems, incident information can be inferred using detector data or observed by surveillance cameras. Using the surveillance data, the MNT can be estimated. The information on the targeted critical density $\hat{k}$ (i.e., the vehicular accumulation of the subnetwork) and the current subnetwork density $k$ is sent to part (iii) to determine the green phases at the gated links.

Part (iii) in Figure 2 represents a feedback-based perimeter controller. In this module, a PID controller outputs the optimal inflow rate to each gated link based on the current network density $k$ and the targeted critical density $\hat{k}$. The perimeter controller can improve the exit flow rate of the incident-affected network, with the cost of additional delays imposed on the vehicles approaching the gated links (Keyvan-Ekbatani et al., 2012).

Part (iii) in Figure 2 represents a feedback-based perimeter controller. In this module, a PID controller outputs the optimal inflow rate to each gated link based on the current subnetwork density $k$ and the targeted critical density $\hat{k}$. This perimeter control strategy can improve the exit flow rate of the subnetwork, at the cost of additional delays imposed on the vehicles approaching the gated links (Keyvan-Ekbatani et al., 2012).

Feedback perimeter control strategy

Feedback control approach

The feedback control approach is a simple, efficient and robust method to adjust control variables. Several MFD-based perimeter control strategies in the literature apply real-time traffic data without relying on traffic prediction (Keyvan-Ekbatani et al., 2012, 2013, 2015; Ramezani et al., 2015).

This study designs a feedback control strategy that factors the impacts of incidents on traffic flow. Figure 3 illustrates the proposed feedback control model. The input of subnetwork traffic model is the gated total inflow rate $q_{in}$. The model output is the TTS of the subnetwork under control. The external disturbances are the inflows at the ungated links, denoted by $q_d$. We adopt the nonlinear continuous-time control design model considering the normal state and the incident-affected state.

As the flow conservation equation list in the work of Keyvan-Ekbatani et al. (2012), flow conservation equation of the incident-affected region can be formulated as:

$$\dot{N}(t) = q_{in}(t) + q_d(t) - q_{out}(t)$$

In an ideal case, we have $N(t) = \sim TTS(t)$, where $\sim TTS$ is the ideal value of TTS, and $N$ is the actual number of vehicles traveling on the subnetwork. However, the measured value of TTS may differ from the ideal value because: (i) sensors are not available at every link, so the measured TTS could be proportional to the actual value with factor $a \leq 1$; and (ii) the occupancy measurement and the estimation Equation (3) may be inaccurate, causing an estimation error $e_0$. See Keyvan-Ekbatani et al. (2012). These two conditions yield:

$$TTS(t) = a \cdot N(t) + e_1(t).$$

Moreover, the normal MFD function, denoted by $F_0(\cdot)$, may change to an affected function, denoted by $F_1(\cdot)$, due to disturbances from incidents in the subnetwork. Based on the work of Keyvan-Ekbatani et al. (2012), the measured TTD is as follows:

$$TTD(t) = \begin{cases} F_0[TTS(t)] + e_2(t), & \text{Normal State} \\ F_1[TTS(t)] + e_2'(t), & \text{Incident - affected State} \end{cases}$$

where $e_2(t)$ and $e_2'(t)$ denote fitting errors. Note that the form of $F_1(\cdot)$ is incident-dependent.

The value of $TTD$ can be estimated in real-time using data collected from links in set $M$ to update the

Figure 3. Block diagram of the feedback controller.
form of $F_k(\cdot)$ dynamically. Denote $\tilde{TTD}(t)$ as the ideal value of $TTD$, which includes the measurements of all links in the subnetwork. The relationship between the measured $TTD$ and $\tilde{TTD}$ is formulated as:

$$TTD(t) = b \cdot \tilde{TTD}(t)$$  \hfill (7)

where $b \leq 1$ is a proportion parameter.

In this paper, $TTD$ of the network is calculated as Equation (1), which can represent the link length weighted total flow of the network. In a traffic network, the proportion of detouring flow is relatively small. Other than the detouring flow, the other flow is driving to a destination within the network or leaving the road network. The sum of traffic volumes leaving the network should be increased with the increase of link length weighted total flow. Namely, the outflow of the incident-affected region increases when the value of $TTD$ becomes larger. Then, following other studies, e.g., Keyvan-Ekbatani et al. (2012), this study assumes that the outflow of the incident-affected region $q_{out}$ is proportional to $TTD$, i.e.:

$$q_{out}(t) = \frac{\Gamma}{L} \tilde{TTD}(t)$$  \hfill (8)

where $0 \leq \Gamma \leq 1$ is the exit flow rate and $L$ is the average link length of the subnetwork. Substituting (5)–(8) into (4) yields:

$$\frac{d}{dt} \text{TTS}(t) = \left\{ \begin{array}{ll} (q_u(t) + q_d(t) - \frac{\Gamma}{BL} F_0(\text{TTS}(t))) \cdot a + \epsilon_3(t) \\ (q_u(t) + q_d(t) - \frac{\Gamma}{BL} F_1(\text{TTS}(t))) \cdot a + \epsilon'_3(t) \end{array} \right.$$  \hfill (9)

where $\epsilon_3(t)$ and $\epsilon'_3(t)$ are fitting errors with zero mean.

When incidents occur, the $y$-axis of the MNT (i.e., the maximum of the TTD) will decrease because of the degraded capacity at the affected links. The free flow speed of the network will not drop significantly if no variable speed limit system is implemented. In such cases, the slop of the MFD in the uncongested regime will not change significantly. With the decrease of the maximum of the TTD, the $x$-axis of the MNT, i.e., the position of the critical density/TTS, can only move to the left. Hence, we assume that the values of MNT and its corresponding accumulation reduce after an incident occurs. In particular, the reductions of MNT and accumulation follow the equations below:

$$\tilde{TTS}' = \omega \tilde{TTS}$$  \hfill (10)

$$\tilde{TTD}' = \theta \tilde{TTD} + \varepsilon_4(t)$$  \hfill (11)

where $0 \leq \theta \leq 1$ is a throughput reduction factor, $0 \leq \omega \leq 1$ is an accumulation reduction factor, and $\varepsilon_4$ indicates the reduction error. Denote $(\tilde{TTS}, \tilde{TTD})$ as the MNT at the normal state, and $(\tilde{TTS}', \tilde{TTD}')$ as the incident-affected MNT. Based on Equation (11), $(\omega \tilde{TTS}, \theta \tilde{TTD} + \varepsilon_4(t))$ indicates the MNT at which the $TTD$ reaches the maximum value after the incident.

The values of $\theta \tilde{TTD} + \varepsilon_4$ and $\tilde{TTD}$ can be regarded as estimates of MNT under the incident-affected state and normal state, respectively. Then,

$$C_0 = \tilde{TTD}$$

$$C_1 = \tilde{TTD} + \varepsilon_4(t)$$

It yields that:

$$\theta = \frac{C_1 - \varepsilon_4(t)}{C_0}$$  \hfill (14)

where $C_0$ is the MNT under the normal state and $C_1$ is the MNT under the incident-affected state. Suppose the incident-affected MNT follows an approximated relationship:

$$C_1 = C_0 - \sum_{z \in M_I} q_z'(t) \cdot L_z$$  \hfill (15)

where $t_1$ is the occurrence time of incident $I$, $q_z(t_1)$ is the measured flow on link $z$ at time $t_1$, $L_z$ is the length of link $z$, and $M_I$ indicates the set of links where incidents occur.

Note that the incidents lead the degraded capacity at the affected links, resulting in the reduction of the MNT (i.e., the network capacity) at the normal state. Not only does the incident reduce the capacity of the link where the incident occurs, but also reduces the capacity of other links in the affected region due to the congestion or gridlock. As a result, the MNT can be smaller than the sum of capacity for each single link (Geroliminis & Boyaci, 2012). Here, we introduce vulnerability parameter $x'_z$ to indicate the overall impact of incident $I$ on link $z$ on the vehicle accumulation.

The vulnerability of the road transportation system is referred to as “a susceptibility to incidents that can result in considerable reductions in road network serviceability. These incidents may then be more or less predictable, caused voluntarily or involuntarily, by man or nature” (Berdica, 2002). This parameter can be estimated by examining the removal of incident-affected links on the performance of the subnetwork. Vulnerability parameter $x'_z$ can be calculated as:

$$x'_z = \frac{P(G_0)}{P(G_z)} \cdot \frac{C_z^I}{C_z^0}$$  \hfill (16)

where $P(G_0)$ measures the performance (e.g., mean velocity or $TTD$) of traffic network $G_0(V,E)$ ($V$ is the
set of nodes and $E$ is the set of directed edges in the network. $P(G_z)$ measures the performance of traffic network $G_z(V,E)$, which has link $z \in E$ removed. The ratio $C_z/C_z^0$ is the capacity discount rate of link $z$, in which $C_z^0$ is the capacity of link $z$ affected by incident $I$ and $C_z^0$ is the capacity of link $z$ under normal state. Note that $C_z^0$ is estimated based on information on incidents and data collected in real time. Depending on the severity of the incident, the value of capacity discount rate varies between 0 and 1. For each link, $P(G_0)/P(G_z)$ can be estimated offline using traffic assignment or simulation (Wang, 2017). The vulnerability parameter $x_z^d$ is then derived to estimate the MNT online.

Substituting Equation (15) into Equation (14) yields an estimation of $\theta$:

$$ \theta = \frac{C_0 - \sum_{z \in E_z} x_z^d \cdot q_z(t) \cdot L_z - e_4(t)}{C_0} \tag{17} $$

We denote parameter $c$ as the relationship between throughput reduction and accumulation reduction of the incident-affected MNT:

$$ e = \frac{c}{\theta} \tag{18} $$

Parameter $c$, whose value is close to 1 based on our simulation experiments, can be calibrated using field data. It significantly affects the control target to maximize network throughput.

The goal of the feedback controller is to maintain the system state around the optimal steady state. Around the MNT, according to Equation (11), the following equation can be derived approximately:

$$ F_i(TTS(t)) = \theta \cdot F_0 \left( \frac{TTS(t)}{\theta} \right) + e_4(t) \tag{19} $$

Substituting (17), (18), and (19) into (10) yields a first-order nonlinear system around the MNT:

$$ \frac{d}{dt} (TTS(t)) = \left( q_{in}(t) + q_d(t) - \frac{\Gamma \theta}{bL} F_0 \left[ \frac{TTS(t)}{c\theta} \right] \right) \cdot a + e(t) \tag{20} $$

where $e(t)$ integrates errors $e_3(t)$ and $e_4(t)$. In Equation (19), $q_{in}(t)$ is the control variable, $q_d(t)$ is disturbance term and $TTS(t)$ is the output of the first-order nonlinear system. Note that feedback control laws for nonlinear systems, such as Equation (20), are complex and difficult to implement. Therefore, we introduce a linearization technique to facilitate the design of effective feedback control for nonlinear system (20).

In control theory, using an optimal steady state as the control target helps to derive a linearized model and design a linear feedback control. Thereby, we linearize Equation (20) around the MNT where the slope of the MFD is positive on the left side and negative on the right side (Daganzo, 2007). This assumption does not impact on the proposed PID controller, whose operation is only governed by the controller error $\omega TTS - TTS(i)$ (Keyvan-Ekbatani et al., 2012).

The relationship between the steady-state variables is:

$$ \frac{\bar{q}_{in} + \bar{q}_d}{\bar{q}_{out}} = \frac{\Gamma}{b \cdot L} TTD \tag{21} $$

Define $\Delta x = x - \bar{x}$ for all variables. The linearized system is as follows:

$$ \frac{d}{dt} (\Delta TTS) = \begin{cases} \left( \Delta q_{in} + \Delta q_d - \frac{\Gamma}{bc_L} \Delta TTS \right) \cdot a + e, & TTS \leq \omega TTS \\ \left( \Delta q_{in} + \Delta q_d - \frac{\Gamma}{bc_R} \Delta TTS \right) \cdot a + e, & TTS > \omega TTS \end{cases} \tag{23} $$

where $\overline{TTS}$ is the slope of the assumed MFD in the left side of the MNT $\left( \omega TTS, \omega TTD + e_4(t) \right)$, i.e. $\overline{TTS} > 0$, and $\overline{TTS}$ is the slope of the assumed MFD in the right side of the MNT $\left( \omega TTS, \omega TTD + e_4(t) \right)$, i.e. $\overline{TTS} < 0$. The continuous-time Equation (23) can be transformed into a discrete-time formula for the variation of total number of vehicles from time interval $i$ to $i + 1$, i.e., $\Delta TTS(i+1)$ as follows:

$$ \Delta TTS(i+1) = \begin{cases} \mu_L \cdot \Delta TTS(i) + \xi_L \cdot (\Delta q_{in}(i) + \Delta q_d(i)) + e(i), & TTS \leq \omega TTS \\ \mu_R \cdot \Delta TTS(i) + \xi_R \cdot (\Delta q_{in}(i) + \Delta q_d(i)) + e(i), & TTS > \omega TTS \end{cases} \tag{24} $$

where $\mu_L = \exp \left( -\Gamma F'_L Ta/bcL \right)$, $\mu_R = \exp \left( -\Gamma F'_R Ta/bcL \right)$ and $\xi_L = (1 - \mu)bcL/\Gamma F'_L$, $\Delta q_{in}(i)$ is the control variable, and $\Delta q_d(i)$ is a disturbance term at time interval $i$. The stability of the linearized system can be proved using the Jury stability criterion (Seborg et al., 2010). The main purpose of model linearization around MNT is to reduce the complexity of the modeling structure of the underlying dynamics of the network, which is essential for a proper choice of the feedback controller.

**Controller design**

The effectiveness of the linearized model relates to the control target as well as the stability and convergence of the controller. The model contains two steps. First, the control target is determined adaptively in each time interval $i$ (i.e., the measurement time intervals...
A PID-type feedback controller is constructed as:

\[
\text{PID controller} = K_P \frac{1}{s} + K_I s + K_D s^2
\]

around the targeted value. The problem of this study applies a PID controller to maintain the TTS or the target state while increasing the system state convergence speed of the incident-affected region. Therefore, this study applies a PID controller to increase the stability of the system (Åström & Hågglund, 1995). This helps to ensure system stability while increasing the system state convergence speed of the incident-affected region. Therefore, this study applies a PID controller to increase the stability of the system while increasing the proportional gain (i.e., \( K_P \)) and PID controller (Åström & Murray, 2010). PID controller is popular for solving industrial control problems (Desborough & Miller, 2002).

Unlike the feedback control at normal state, the traffic state and MNT of the network change rapidly when incidents occur. To help the TTS quickly converge to the time-dependent control target, feedback control requires better dynamic performance. Increasing the proportional gain (i.e., \( K_P \)) is a method to increase the convergence speed of the PID controller, yet reduce the control stability. The derivative term (i.e., \( D \)) of the PID controller can increase the stability of the system while increasing the proportional gain (i.e., \( K_P \)) (Åström & Hågglund, 1995). This helps to ensure system stability while increasing the system state convergence speed of the incident-affected region. Therefore, this study applies a PID controller to maintain the TTS around the targeted value. The TTD of the subnetwork is maximized based on the dynamically updated MNT.

Based on the discrete-time system (24), a standard PID-type feedback controller is constructed as:

\[
q_{in}(i) = q_{in}(i - 1) + K_P \left[ \omega \bar{\text{TTS}} - \text{TTS}(i) \right] + K_I \sum \left[ \omega \bar{\text{TTS}} - \text{TTS}(i) \right] + K_D \left[ \text{TTS}(i - 1) - \text{TTS}(i) \right]
\]

where \( K_P \) is the proportional gain, \( K_I \) is the integral gain, and \( K_D \) is the derivative gain. Their functionalities are highlighted as follows. First, the proportional term provides an overall control action proportional to the error signal. Second, the integral term reduces steady-state errors through low-frequency compensation using an integrator. Third, the derivative term improves the transient response through high-frequency compensation using a differentiator. Here, the control stability is referred to as the bounded-input, bounded-output stability, namely, the output of a system will be bounded for every input to the system that is bounded (Camacho & Alba, 2013). When a system is unstable, the output of the system may be infinite though the input is finite.

In this study, two tuning approaches are applied. (i) The linear model is established to describe the dynamic of the network. Then we applied the Ziegler-Nichols method (Åström & Hågglund, 1995) to get initial value for each parameter. (ii) We applied the simulation with the initial value of three parameters and further fine-tuning these parameters to satisfy the constraints of setting time and the amplitude. The calibration of the second approach contains the following three steps for each subnetwork:

i. Establish a simulation model of the subnetwork. The demand of the network increases with simulation time from 0 pcu/h to the saturation state. Use Equation (1) to estimate the subnetwork TTS based on the simulation data. Sum up the current gated links’ inflow rates to obtain the subnetwork gated inflow rate \( q_{in} \). Use the subnetwork TTS and gated inflow rates \( q_{in} \) as inputs to establish the control plan of the subnetwork.

ii. Apply the PID Controller dialog in MATLAB to calibrate its three parameters. This study chooses parameters that satisfy the following two constraints: (a) the settling time (i.e. the time required for the current TTS to reach and stay within a range of certain percentage of the target value (Tay et al., 2012)) should be greater than the time interval \( i \) (e.g., in the following simulation, the setting time is set to 3 times of the time interval \( i \), 300 seconds), and (b) the amplitude is the smallest. Note that as the settling time decreases, the convergence rate of the controller increases. Yet, the stability of the controller will become difficult to guarantee. Therefore, it is necessary to take as short settling time as possible while ensuring that the controller is stable. For the conducted simulation experiments, it was observed that the controller resulted in stable closed-loop behavior.

iii. Test the stability of the PID controller in the simulation subnetwork with the proposed perimeter control method and observe whether the TTS converges to the control target TTS.

In practice, the three control parameters can be calibrated based on subnetwork data under the no-incident condition. When incidents occur, the calibrated control parameters can be used as the initial values.

The gating procedure includes six steps:

i. The controller estimates TTS using Equation (1) continuously, which is based on real-time measured occupancy and flow of each link in the subnetwork over the previous time interval. The system also detects incidents based on the collected sensor data.
ii. The controller targets the incident-affected MNT \(\left(\omega \hat{TTS}, \theta \hat{TTD} + e_4(t)\right)\). The value of \(\hat{TTS}\) is the \(x\)-coordinate value of the MNT under normal state. Parameter \(\omega\) is calculated using Equations (17) and (18) in each time interval \(i\). Then, the controller outputs the optimal gated inflow rate \(q_{in}\) to be applied in the next time interval.

iii. A switch-on/off logic decides whether or not the controller implements the optimal inflow rate \(q_{in}\). Gating is switched on when the measured \(TTS\) is larger than the maximum throughput \(\omega \hat{TTS}\).

iv. During the switch-on period, the optimal inflow rate \(q_{in}\) is distributed among the gated links \((q_{in1}, ..., q_{inn})\) proportional to the real-time inflow rate of each link, where \(n\) is the number of gated links. Then, the distributed inflows are converted into green phases \((g_{qin1}, ..., g_{qinn})\), where \(g_{qinj}\) is the gated green split at gated link \(j\). Here, the cycle length is not a control variable. For each gated link, the cycle length remains unchanged, while the green time is reduced proportional to the value of \(q_{inj}\), \(j = 1, ..., n\) until it reaches the minimum green, which factors the queue length.

v. During the switch-off period, the controller implements the ordinary signal timing plan \(g_{qj}, j = 1, ..., n\).

vi. The traffic state in the incident-affected region (i.e., \(TTS\), \(TTD\), and \(C_i\)) is estimated in real time as an online process during the whole procedure.

In step (ii), note that the time interval \(i\) is the time interval of \(TTS/TTD\) measurement, MNT estimation and the feedback control. This time interval should be long enough to cover the variations of traffic signal cycles of the network (e.g., the cycle length of each intersection is between 60 seconds and 90 seconds). Based on this premise, the computation burden and control accuracy of the controller will continue to increase as the interval is shortened. Therefore, in practical applications, the time interval \(i\) should be determined according to the size of the control road network, the complexity of the road network, and the computing power.

In step (iv), green times for the inbound traffic of the gated link decrease proportional to the flow rates. The lower bound is the minimum green time at each gated link. The objective of the proposed approach can be altered to other performance measures, e.g., minimizing the average delay. Thereby, we can develop different signal optimization models.

**Numerical experiments**

**Study network**

The study network uses a part of Nanjing urban road network, where congestion rarely occurs. Nanjing Olympic Center Stadium is on the south side of the network. During major sports events, the traffic state of this area is significantly impacted by the spectators entering or leaving the stadium. This part of Nanjing urban road network is illustrated in Figure 4, and modeled in VISSIM traffic simulator. The incident-affected network consists of 92 links and 23 intersections. A loop detector is installed in the middle of every link in the network. Real-time measurements are processed every 100 seconds in the proposed approach. This measurement time interval is long enough to cover the variations of traffic signal cycles.

![Figure 4. Locations of incidents and gated links in VISSIM.](image-url)
in the study network. The gated links have been chosen to minimize the negative impacts on traffic outside of the incident-affected region. In the simulation, all the non-gated intersections in the incident-affected network are set as pre-timed intersections, while drivers do not change their route choice.

Note that in a small network, the primary incidents may lead to more than one secondary accident. Secondary accidents are caused by the creation and existence of the traffic queue (Moore et al., 2004; Park & Haghani, 2016). Even in a small network, it is possible to have multiple incidents simultaneously (Karlaftis et al., 1999). The simulation network established in this study is relatively small due to the limited computing power, which is used to verify the validity of the proposed method. In reality, the multiple adjacent incidents will affect a larger traffic network, which is large enough to apply the proposed method.

We investigate the proposed control strategy using two simulation experiments. The first experiment based on the whole test network focuses on verifying the MNT estimation method using scenarios with incidents occurring at locations 1, 2, and 3 shown in Figure 4. In this experiment, the network demand increases with simulation time from 0 pcu/h. The second experiment compares the proposed control strategy to the fixed-MNT feedback control strategy and pre-timed control strategy, using scenarios with incidents occurring at locations 1, 4, and 5 shown in Figure 5. Note that in the second experiment the MNT of the controlled subnetwork is estimated dynamically. In the second experiment, the VISSIM simulation model is calibrated by the RFID (Radio Frequency Identification) field detector data, which includes volume, the average speed of the vehicles per 5 minutes. The time interval of the MNT estimation and the feedback control is the same of measurement time interval, which is 100 seconds as well. Note that the expected vehicle speeds in both experiments are calibrated based on the RFID data.

**MNT estimation**

As discussed in Section 2, the MFD for the normal state is obtained using long-term historical field data, and then the MNT under the normal state is estimated. This study estimates the MNT using traffic measurements from VISSIM. They include occupancy (in %) and flow rate (in pcu/h) for every 100-second time interval. When incidents occur, the transient relationship between network throughput and the accumulation can be constructed to approximate an MNT using the network demand loading phase of the affected network. The MNTs of the network under normal state and incident-affected states are shown in Figure 5.

The simulation period is 3.5 hours, with a warm-up period of 30 min. It is divided into eighteen 10-min demand intervals, except for the 30-min warm-up period. The network demand increases with simulation time from 0 pcu/h. As VISSIM is a stochastic simulator, simulation runs with different random seeds lead to different results. We executed 10 simulation runs to obtain occupancy and flow for each link. Three incidents were simulated to investigate the online MNT estimation method. Here, the MNT is obtained by taking the average of TTS and X-percentile TTD (the selected percentile depends on the dispersion degree of the original MFD data). In the experiments, we use the 80th percentile TTD for each scenario. At location 1, a small traffic accident was simulated, which affected one of the three lanes. At location 2, a lane was closed at 50 meters from the link beginning to simulate a temporary lane closure. At location 3, a serious traffic accident was simulated, which affected two of the three lanes. As mentioned in the introduction section, the MNT varies with incidents.

To verify the MNT estimation method, we performed 10 runs of 3-hour simulation for each of four scenarios. Scenario 1 is a normal scenario. Scenario 2 contains an incident at location 1. Scenario 3 contains incidents at locations 1 and 2. Scenario 4 contains incidents at locations 1, 2, and 3. In the incident-affected scenarios, the link capacity reduction caused by incidents persists during the loading phase. Figure 5(a) shows the MNT under the normal state. Figure 5(b)–(d) show the MNTs and the estimation of MNTs (labeled as MNTE in Figure 5) under the three scenarios, respectively.

In each scenario, the MNT (\(\hat{\text{TTS}}, \hat{\text{TTD}}\)) can be identified directly based on data from the 10 runs, as shown by the red squares in Figure 5. As shown in Figure 5, the TTD decreases up to 13% (12,466 to 10,833 pcu.km per h) and TTS reduces up to 6% (396 to 374 pcu.h per h). Moreover, the proposed perimeter control approach will only be applied when incidents occur and cause congestion, i.e., perimeter control approach is applied when measured TTS is in the TTS domain where TTD is close to the maximum value. Such a reduction is significant in this TTS domain (e.g., if TTS domain is between 300 and 450 pcu.h per h, the TTS decrease can be up to 14.7%) where the control approach is applied. If the reductions are ignored, the control efficiency for incident-
affected scenarios will be significantly affected, as illustrated in Table 2.

When the reduction of TTD and TTS is ignored, the control target will be made a bit larger than the MNT, which will fall into the congestion regime. This will cause the feedback controller pointing to an unstable state. Moreover, the larger control target will delay the start of the perimeter control and will miss the optimal timing for the perimeter control. In the following section of comparison of control strategies, the second experiment (i.e., experiment to compare the proposed control strategy to the fixed-MNT feedback control strategy and pre-timed control strategy) was conducted to verify the benefits of proposed MNT estimation method when the control target is variable.

The incident-affected MNT is unknown in practical applications, i.e., the controller does not have the information on real MNT (red squares) in Figure 5. Instead, the proposed online MNT estimation method can be used to determine the control target (i.e., the green circles in Figure 5). This dynamically estimated MNT is not far from the ground truth as demonstrated by the results summarized in Table 1.

Specifically, this study proposes Equations (17) and (18) to estimate the MNT. Firstly, parameter $c = 1.02$ was calibrated using the simulation data. The simulation data also estimates the vulnerability parameter $\alpha^T$ in Equation (16) using incident severity in different scenarios. Parameter $c$ was used to estimate the MNTE based on Equations (17) and (18). The values of $\{TTS, TTD\}$ are summarized in the fourth and fifth columns in Table 1, showing that the errors between estimated MNTs and simulated MNTs are less than 5.0% in these scenarios. The small errors demonstrate the feasibility of the proposed method, defined.

**Figure 5.** (a) MNT of the incident-affected region under the normal state for 3 h, (b) MNT of the incident-affected region for 3 h with one incident, (c) MNT of the incident-affected region for 3 h with two incidents, and (d) MNT of the incident-affected region for 3 h with three incidents.
by Equations (11)–(18), for estimating the MNT under the incident scenarios.

Comparison of control strategies

Four scenarios are designed to compare the proposed control strategy with the pre-timed and fixed-MNT feedback control strategies.

i. Scenario 1, labeled as the base scenario, represents the normal state without incidents.

ii. In scenario 2, labeled as the small-effect scenario, a small traffic accident occurs at 8:15 a.m. at location 4 in Figure 4. The impact of the accident lasts 60 minutes.

iii. In scenario 3, labeled as the moderate-effect scenario, one of the three lanes is closed for temporary road construction from 8:00 a.m. to 9:30 a.m. at location 5, in addition to the incident in scenario 2.

iv. Scenario 4, labeled as the severe-effect scenario, is constructed based on scenario 3. At 8:15 a.m. an additional incident (e.g., a serious traffic accident) occurs at location 1. The impact of the serious traffic accident lasts 90 minutes. To analyze the impact, the demand profile is scaled upwards based on the detector data collected from 7:30 a.m. to 10:30 a.m.

The effectiveness of the three control strategies is investigated using three performance metrics: (i) the total delay in hours, (ii) the average speed in km/h, and (iii) the total number of vehicles that have left the network during the 3-hour period. The statistics for the above performance metrics were computed using the full network, including the subnetwork (i.e., network within the gated intersections) and the non-controlled network (i.e., network outside the gated intersections). We compare the performance of the study network under different control strategies.

Pre-timed control strategy

Without gating control, the fixed-time signal plans used in the real world are applied to the study network. The averaged values of performance metrics based on 10 replications are compared across different scenarios.

Since VISSIM is a stochastic simulator, link spillovers and partial gridlocks may occur in some simulation runs, which lead to higher delays and lower speeds. The results show that the network is not empty at the end of the simulation period, implying poor network performance without gating control.

Fixed-MNT control strategy

Six gated links are placed on the edge of the subnetwork. Based on the controlled system (24) and (25), the discrete-time linearized system is specified by the vectors $q_{in}$ and $TTS$ around the MNT. After the linearized discrete-time system is specified, stable controller parameters were estimated as $K_p = 77.015 \text{ h}^{-1}$, $K_I = 0.083 \text{ h}^{-1}$, and $K_D = 0.083 \text{ h}^{-1}$. The fixed-MNT feedback control strategy applies $TTS = 396 \text{ pcu.h per h}$ in each scenario.

Proposed feedback control strategy

Here, we apply the proposed feedback-based perimeter control strategy with incident-dependent MNT. The locations of the gated links and the control parameters are the same as those in the fixed-MNT control strategy. The control target $\omega TT$S is dynamically updated based on Equations (17) and (18) in each scenario.
Figure 6. (a) TTS variation in Scenario 1, (b) TTS variation in Scenario 2, (c) TTS variation in Scenario 3, and (d) TTS variation in Scenario 4.
Comparison of strategies

Table 2 compares the performance metrics under different control strategies. It shows that the overall network delay reduces significantly under the proposed control strategy compared to the pre-timed strategy for each scenario. In scenarios 2, 3, and 4, the overall network delay reduces under the proposed strategy compared to fixed-MNT strategy. As the incident severity increases, the performance increases. This demonstrates the effectiveness of the proposed strategy.

To highlight the differences between the proposed strategy and the pre-timed and fixed-MNT strategies, we present the detailed results of the first simulation run in all scenarios. The comparison focuses on the TTS of the subnetwork.

In Figure 6, the red dashed lines show the control target in the fixed-MNT control strategy and the blue dashed curves show the dynamic control targets of the proposed strategy. The solid green curves show the evolution of TTS without gating control, the solid red curves show the evolution of TTS under the fixed-MNT control strategy, and the solid blue curves show the evolution of TTS under the proposed strategy.

Under the pre-timed control, the TTS gradually increase in each scenario due to the increase in traffic demand from 0 to 5,400 seconds of the simulation. From 1,800 to 5,400 seconds in simulation, congestion leads to an increase in TTS. During this period, the control target of the proposed control strategy varies, as shown by the blue dash curves. In the base scenario (i.e., scenario 1), the control target of the proposed control strategy is the same as that of the fixed-MNT control strategy. Hence, the TTS values are the same under these two control strategies, as shown in Figure 6(a). In scenario 2, the control target of the proposed strategy is lower than that of the fixed-MNT control strategy, as illustrated in Figure 6(b). In the moderate- and severe-incident scenarios (i.e., scenarios 3 and 4), the control target of the proposed strategy is updated dynamically, leading to a much lower TTS value than that in the fixed-MNT control strategy, as illustrated by Figure 6(c). Also, the proposed strategy leads to a lower TTS value as illustrated in Figure 6(d).

Figure 6 also illustrates that in the simulation conditions, the proposed PID controller is stable. As shown by the relationship between TTS and the control target in Figure 6, the TTS value steadily shifts toward the control target based on the well-tuned PID controller parameters described in Controller Design Section.

In summary, through the incident-dependent MNT, the dynamically updated control target can improve network throughput of the whole incident-affected network when traffic is moderately or severely affected by incidents. The improved network throughput leads to a reduction in average delay and an increase in traffic speed when the proposed control strategy is applied to networks affected by moderate or severe incidents.

Concluding remarks

This study develops a perimeter control strategy to maximize the throughput of the controlled subnetwork within an incident-affected network. In the control strategy, the incident-dependent maximize network throughput (MNT) is dynamically determined by an online estimation method, which applies a data-driven approach to improve the accuracy of control targets and prevent the overestimation of network throughput. Following a feedback control mechanism, the proposed control strategy employs a PID controller to prevent sharp fluctuations of gated inflows when updating the MNT. Results from simulation experiments illustrate that the proposed online perimeter control strategy is superior to the fixed-MNT feedback control strategy for an incident-affected network. The results also highlight the benefits of the dynamically-updated control target in reducing average delay and increasing network throughput.

This study provides several future research directions. First, the proposed approach can be improved by looking into the dynamic mechanism of incident-affected MFD. The shape and scatter level of affected MFD and the hysteresis loops phenomenon are valuable knowledge in analyzing the network performance and estimating accurate control targets for non-recurring congestion. This will help in designing more effective methods to mitigate the degradation caused by incidents. Second, developing approaches to estimate the vulnerability parameter $\alpha^*_2$ will enhance the control efficiency. Since $\alpha^*_2$ characterizes the relationship between $\omega$ and $\theta$ for different incident-affected networks, a reliable estimation approach can enhance the transferability of the proposed control strategy. Third, the impact of the incidents on the larger-scale network is worth further studying. Due to the inherent heterogeneity of large-scale networks, a multi-region feedback control approach based on adaptive partitioning could be further integrated into perimeter control.

Moreover, the proposed approach that regulates the control target around the dynamic MNT is a simple version of trajectory-tracking control (Aguiar & Hespanha, 2007; Reyhanoglu et al., 1999; Wu et al., 2019). Trajectory-tracking control can track the MNT of the incident-affected region when the change in
demand is fast (Mariotte et al., 2017), which maximizes the throughput of incident affected network. Moreover, robust control approaches can improve the performance of the controller under different traffic scenarios (Ampountolas et al., 2017; Haddad & Shraiber, 2014; Wu et al., 2019). Therefore, it will be valuable to extend the proposed approach by integrating with other control concepts, e.g., trajectory-tracking control and robust control in the future.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was financially supported by the Shanghai Sailing Program (Grant #19YF1435100). Any errors or omissions remain the sole responsibility of the authors.

**ORCID**

Jiawen Wang [http://orcid.org/0000-0003-2317-1390](http://orcid.org/0000-0003-2317-1390)

Xiaozheng He [http://orcid.org/0000-0002-1633-4748](http://orcid.org/0000-0002-1633-4748)

**References**


Berdica, K. (2002). An introduction to road vulnerability: What has been done, is done and should be done. *Transport Policy, 9*(2), 117–127. [https://doi.org/10.1016/S0967-070X(02)00011-2](https://doi.org/10.1016/S0967-070X(02)00011-2)


Li, X., & Sun, J.-Q. (2019). Multi-objective optimal predictive control of signals in urban traffic network. Journal of...


