Dangerous Driving Behavior Recognition and Prevention Using an Autoregressive Time-Series Model

Hongxin Chen, Shuo Feng, Xin Pei*, Zuo Zhang, and Danya Yao

Abstract: Time headway is an important index used in characterizing dangerous driving behaviors. This research focuses on the decreasing tendency of time headway and investigates its association with crash occurrence. An autoregressive (AR) time-series model is improved and adopted to describe the dynamic variations of average daily time headway. Based on the model, a simple approach for dangerous driving behavior recognition is proposed with the aim of significantly decreasing headway. The effectivity of the proposed approach is validated by means of empirical data collected from a medium-sized city in northern China. Finally, a practical early-warning strategy focused on both the remaining life and low headway is proposed to remind drivers to pay attention to their driving behaviors and the possible occurrence of crash-related risks.

Key words: time headway; driving behavior; traffic safety; autoregressive time-series model; remaining life; driving warning strategy

1 Introduction

Traffic safety is the foundation of smooth road traffic system operation and management. Past studies have analyzed various contributory factors to traffic crashes and presented tailored solutions and measures, all with the aim of reducing traffic crashes and improving the safety performance of an urban traffic system. Driving behavior-related factors, such as speed, time headway, and so on, are the primary factors associated with traffic crashes. Speed is a determining factor influencing crash frequency and severity[1, 2], and average speed and speed dispersion have both been found to influence crash occurrence and crash severity[3–5].

Time headway has also been associated with traffic safety[6–8]. In practice, an appropriate time headway

● Hongxin Chen, Shuo Feng, Xin Pei, Zuo Zhang, and Danya Yao are with Department of Automation, Tsinghua University, Beijing 100084, China. E-mail: chenhongxin@tsinghua.edu.cn; s-feng14@mails.tsinghua.edu.cn; peixin@tsinghua.edu.cn; zhangzuo@tsinghua.edu.cn; yaody@tsinghua.edu.cn.
∗ To whom correspondence should be addressed.
  Manuscript received: 2016-07-21; revised: 2016-09-28; accepted: 2016-10-03

is usually recommended by polices or engineers so that drivers can maintain a safe distance from the leading car. For instance, driver training programs in the US suggest a time headway of not less than 2 s. In Germany, driving with time headway less than 0.9 s is penalized[6]. In relation to this, many previous studies have focused on the relationship between headway and crash occurrence[9–11]. Evans and Wasielewski[12] reported that crash-involved drivers are more likely to drive at short time headways (less than 1 s) than crash-free drivers; a similar result has been found upon comparing drivers with and without traffic violations. Vogel[6] compared time headway with time to collision, and concluded that small headways, which may potentially induce risk of rear-end crashes, can be used as an indicator for tailgating. In our preliminary research[13], we reached a similar conclusion: crash-involved drivers tended to keep shorter headways than non-crash-involved drivers in daily driving.

However, many previous studies that collected headway data mainly relied on driving simulators or on-board record units to capture the continuous driving data of a certain number of sampled drivers during a limited duration[8, 14]. Therefore, revealing the explicit relationship between time headway and crash risk for

general populations is difficult. In the current study, using traffic surveillance and traffic crash data from a city in northern China, we attempt to extract the time headway feature of each vehicle passing through the surveillance intersections, which are also the significant nodes connecting the entire road network. Thus, the time-series headway information for different types of drivers can be obtained and further analyzed.

Maintaining a shorter headway between cars is more dangerous during driving. Therefore, in this research, we focus on the association of crash risk and decreasing tendency of headway along the time-series headways of each driver, in an attempt to automatically recognize the decreasing tendency. An improved autoregressive (AR) time-series model is thus developed to detect and characterize the abnormal variation of time headways. This model is validated by comparing general and the selected crash rates. By using this model, the significant decreasing tendency of headway correlated with higher crash risk can be automatically recognized. Furthermore, based on the model parameter of remaining useful life, a practical early-warning approach is proposed to remind the drivers to maintain a safe headway, thus improving overall safety performance.

The layout of this paper is organized into sections. Section 2 presents the data preparation and preliminary research. Section 3 introduces the AR model and the associated parameter identification method. A dangerous driving behavior detection approach is further established based on the AR model. Section 4 represents the results and discussion by using the AR model in the case study. In Section 5, a practical early-warning approach for dangerous driving behavior is developed based on the concept of remaining useful life. The concluding remarks are presented in the final section.

2 Data Preparation and Preliminary Research

Traffic data of various types are paving a new way for the development of intelligent transportation systems[15]. The urban traffic surveillance systems, which are deployed at urban intersections or trunk road segments with loop or camera detectors, can capture the passing time, license plate, vehicle type, speed, traveling lane, and some driving violation behaviors that occur when vehicles pass by. Urban traffic surveillance data are based on massive vehicles' long-term travel records, which may represent routine driving behaviors. In this study, we collected traffic surveillance data from 79 monitored intersections in a city in northern China with a population of about 4.5 million. Every day, 1.5 million data are generated. For the one-year research period of this study, the database became so huge that distributed computation techniques for such big data were adopted. Owing to the continuous records of vehicles, we can thus calculate and analyze time headway, which is of great interest in terms of driving behavior.

Moreover, this study also obtained the crash information of the city from police crash records during the same one-year period. The crash records of the whole city within the entire year were recorded and analyzed. Crash information contains crash time, location, injury severity, license plates of the involved vehicles, as well as the corresponding drivers gender, age, driving age, and responsibility in the crash, to name a few.

In this research, we combined the traffic surveillance data and crash data to recognize the different driving behaviors between crash-involved and non-crash-involved drivers. First, we extracted the appropriate vehicle database using some filter criteria to ensure the consistency of vehicles and drivers. Next, we classified the drivers into different groups based on their roles or responsibilities in the crashes that occurred.

The vehicle extraction process was conducted in four steps. First, drivers of vehicles with different lengths may have significant variations in their respective driving behaviors. For example, a driver of a large vehicle tends to drive more carefully than the one driving a light-duty vehicle, resulting in diverse distributions of time headways. In this study, private vehicles were extracted from the database. Second, nonlocal vehicles were excluded, because of limited data and limited experience in driving in the city. Third, we also excluded vehicles with specialized plates for special purposes, such as learner-driven vehicles, police vehicles, etc. Finally, vehicles captured less than 20 days were also removed due to limited records. A total of 169 478 vehicles were eventually selected. Considering the recorded responsibility of the vehicles involved in a crash, we divided those into two groups: crash-involved ones, which have fault in crashes, and non-crash-involved vehicles, which have not yet experienced a crash or has just crashed in an accident. Consequently, among all of the extracted
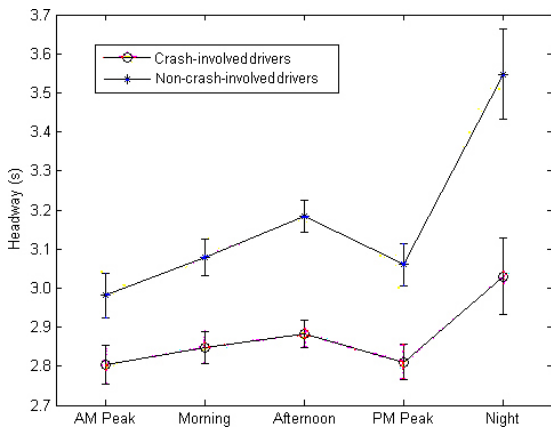
vehicles, 422 vehicles are involved in crashes, with an overall accident rate of 0.249%.

In our preliminary research presented in Ref. [13], we compared the headway difference between two adjacent vehicles passing by the monitor system in the same road environment, one of which is the crash-involved vehicle; we found that the crash-involved drivers have significantly shorter time headways than non-crash-involved drivers. Figure 1 plots the average values of the respective time headways of crash-involved and non-crash-involved drivers, at different times of a day. It seems that shorter time headway for drivers would be more dangerous, thus the inappropriate headway detection is necessary for each driver according to his own time series headways.
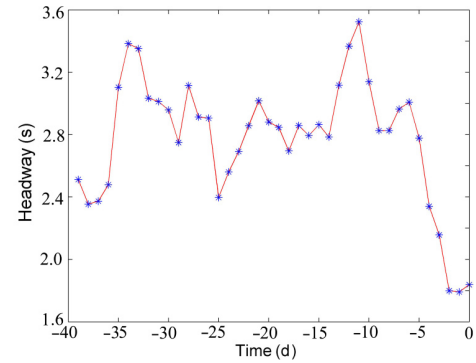
In our database, the daily time headway for each driver or vehicle can be calculated if the records are sufficient. The time-series headways along days of four sampled drivers can be illustrated, as shown in Fig. 2. As can be seen, for the crash-involved drivers with crashes that occurred on Day 0, the time headways significantly decreased before each crash. This result indicates that the continuous reduction in time headway, rather than the occasionally short headway, may be more potentially responsible for traffic accidents. This is because the driver fails to be aware of his/her slowly changing driving behavior, which may be dangerous in this kind of situation.

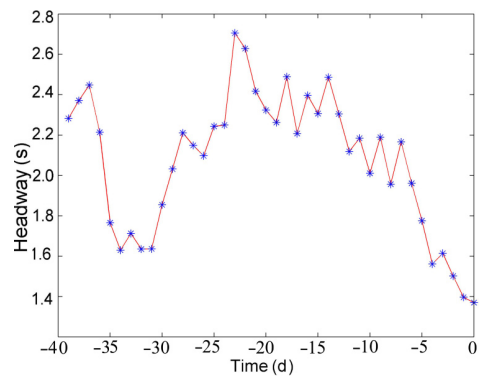# 3 AR Model for Dangerous Driving Behavior Recognition

As mentioned in the last section, accurately detecting a decreasing headway is of great importance in this study. Based on the time-series headway data, we developed an AR time-series model to describe the
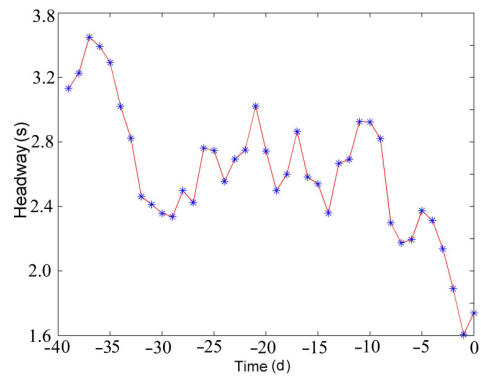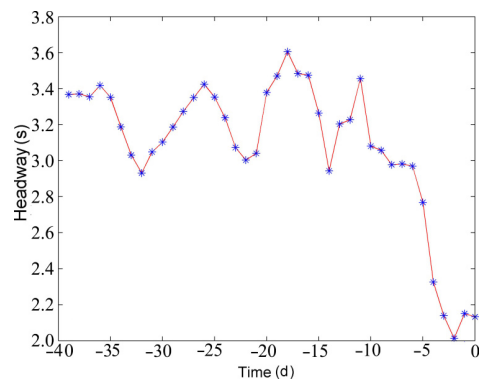

(a) Driver 1

(b) Driver 2

(c) Driver 3

(d) Driver 4

Fig. 2 **Time-series curve of headways of four sampled drivers before a crash happened on Day 0.**

Fig. 1 **Average time headway values of crash-involved drivers and non-crash-involved drivers.**

dynamic variations of daily headways. We then adopted this model to recognize dangerous driving behaviors in terms of significant decreasing headway.

3.1 Fundamental model

AR models have been extensively employed in different studies to characterize diverse dynamic phenomena in nature, economics, engineering, and so on. Particularly, in the field of engineering, AR models are considered powerful tools by which to analyze time-series data. For example, in process control community, routine variations of industrial process can be modeled by using vector AR models, from which effective process monitoring and product quality prediction strategies can be developed[16, 17]. For the purpose of conducting reliability analysis, AR models are typically used to describe the degeneration of industrial equipment, and based on these, predictions of remaining useful life can be derived[18–20].

In the present study, the average daily time headway of vehicles is regarded as an index that characterizes the daily driving behavior of drivers. Based on the assumption that the average daily time headway of a driver, denoted as $H(t)$, is linearly auto-correlated, we can perform the approximation using a first-order AR model. Its mathematical definition is given by

$$H(t) = \beta H(t-1) + e(t) \tag{1}$$

where $e(t)$ stands for zero-mean Gaussian noise causing random variations of time headway, and β is the model parameter explicitly governing the Markovian property of $H(t)$. In the case of $0 \leqslant \beta < 1$, $H(t)$ is a stationary stochastic process: when β equals unity, $H(t)$ reduces to the Brownian motion, and when β exceeds one, the trend of $H(t)$ can be considered to have exponential growth.

3.2 Recursive parameter estimation

Note in Eq. (1) that the model parameter β is time-invariant. However, in practical scenarios, the driving behaviors of drivers often manifest time-varying characteristics to varying extents. In other words, the temporal correlation of $H(t)$ may vary in different time scales and time periods. Accordingly, we can further assume that the model parameter β is time-varying, as shown below.

$$H(t) = \beta(t)H(t-1) + e(t) \tag{2}$$

The real-time value of $\beta(t)$ indicates the short-term variation tendency of time headways. When the value of $\beta(t)$ is small ($\beta(t) < 1$), the time headway decreases temporarily, which generates useful information for dangerous driving behavior recognition. Therefore, deriving real-time estimations to $\beta(t)$ is important. To estimate $\beta(t)$ in this article, we propose to adopt the recursive least square algorithm with forgetting factors. When applied to AR models, the algorithmic details are given by Ref. [21]:

$$\hat{H}(t) = \hat{\beta}(t-1)H(t-1) \tag{3}$$

$$\hat{\beta}(t) = \hat{\beta}(t-1) + K(t)[H(t) - \hat{H}(t)] \tag{4}$$

$$K(t) = P(t)H(t-1) \tag{5}$$

$$P(t) = \frac{P(t-1)}{\lambda + H(t-1)^2 P(t-1)} \tag{6}$$

where $\hat{\beta}(t)$ denotes the estimated value of $\beta(t)$ at time t, $K(t)$ represents the gain matrix in the recursive least square algorithm, and $P(t)$ refers to the variance of $\hat{\beta}(t)$, as a measure of estimation uncertainty. The initial parameters can be specified as $\hat{\beta}(0) = 1$, $P(0) = \alpha \gg 1$. Here $0 < \lambda < 1$ is the forgetting factor that alleviates the impact of historical data on the estimation and facilitates the effective tracking of short-term variation trends of time headways. Moreover, the forgetting speed of the algorithm becomes lower with an increasing λ. In such cases, more attention must be given to long-term trends instead of short-term ones.

3.3 Dangerous driving behavior recognition based on the AR Model

As mentioned above, the parameter $\beta(t)$ quantitatively delineates the variation trends of time headways. A small $\beta(t)$ value indicates an obvious reduction of time headways and a potentially increased danger in driving behaviors. Nonetheless, the inaccurate estimation of $\beta(t)$ may influence the discrimination of the declining trend of time headways. To help resolve this, the parameter $P(t)$ characterizes the uncertainty of estimation results. Therefore, we propose to synthesize the information within both $\beta(t)$ and $P(t)$, in order to select the time periods with obvious decreasing trends and high reliability. A simple yet effective discrimination policy is given by

$$\begin{cases} \beta(t) < a, \\ P(t) < b \end{cases} \tag{7}$$

where $a \leqslant 1$ and b are user-specified thresholds.

However, the normal value of $H(t)$ that a driver is accustomed to may vary. From Eq. (6) we can observe that drivers may have different $P(t)$ values because of their different driving behaviors. If the time headway $H(t)$ of a driver remains at a high value, the value of

variance $P(t)$ may be small. Thus, it is unreasonable to specify a unified threshold for $P(t)$ of different drivers, thereby rendering the normalization of $P(t)$ necessary. Assuming that the recursive algorithm finally stabilizes and $P(t) \approx P(t-1)$ holds, we could thus arrive at the following relationship:

$$P(t)H(t)^2 \approx 1 - \lambda \qquad (8)$$

In other words, the quantity $P(t)H(t)^2$ of different drivers approximately has the same scale, which depends solely on the forgetting factor λ. Therefore, we propose to use $L(t) = P(t)H(t)^2$ as the normalized variance estimation. Accordingly, Formula (7) should be modified as

$$\begin{cases} \beta(t) < a, \\ L(t) < b \end{cases} \qquad (9)$$

Due to the imperfection of practical traffic data, Formula (7) might be satisfied once in a while, when the value of $H(t)$ fluctuates over time but no obvious decline appears. To further improve the practicability of the recognition results, a simple and effective strategy is presented in Formula (9) in n consecutive days, thus leading to the modified recognition criterion given by

$$\begin{cases} \beta(t) < a, \beta(t+1) < a, ..., \beta(t+n) < a, \\ L(t) < b, L(t+1) < b, ..., L(t+n) < b \end{cases} \qquad (10)$$

4　Case Study

Based on the database introduced in Section 2, the time-series headways for a whole year of each selected vehicle's driver in the target city were identified and arranged for the AR detection model. We conducted the dangerous driving behavior detection for all 169 478 vehicles. First, we carried out the recursive parameter estimation using time headway data of each vehicle, and then referred to the policy shown in Formula (10). The threshold a for the model parameter $\beta(t)$ was set as 0.98, and n was specified as three. The determination of these two parameters was relatively easy because of their explicit physical interpretations. The choice of other parameters, including the forgetting factor λ and the threshold b for the normalized variance estimation $L(t)$, will be investigated comprehensively in the sequel.

Next, we verify our premise that a decrease in time headways serves as a primary cause of crashes and an essential factor of dangerous driving behavior. To do so, we calculated the percentage of vehicles with time headway decline in different groups of vehicles, after which we made further comparative investigations.

Three groups of vehicles were defined, and their detailed settings can be explained as follows. Groups A and B correspond to 169 478 and 422 vehicles involved in crashes, respectively. Group C contains 278 vehicles involved in crashes and has obviously low time headways. The selection criteria for vehicles in Group C include the following: the headways of the vehicles before the crashes must be lower than those of the following vehicles, and the difference must be larger than 0.5 s. In a nutshell, Group A contains Group B, and Group B contains Group C, as illustrated by Fig. 3. By varying the forgetting factor λ and the threshold b, we attained the percentage of drivers with time headway declines in each group, as plotted in Fig. 4, in which each sub-figure corresponded to a fixed forgetting factor λ. When the value of threshold b is small, most vehicles are excluded from time headway declines; when the value of threshold b is about to increase, more vehicle drivers are considered as having dangerous driving behaviors.

As shown in Fig. 4, under different parameter settings, the percentage of time headway decline in Group B is always significantly higher than that in Group A. In other words, drivers involved in crashes tend to have more time headway declines in a statistical sense, indicating that time headway decline is, indeed, a primary cause of car crashes. This assertion can be further confirmed by accident rate curves plotted in Fig. 5. As shown in the figure, when $0.16 \leqslant L \leqslant 0.20$, the overall accident rate of vehicles with obvious decrease in time headways is over twice the overall accident rate of this city.

Moreover, the role of drivers in crashes may be either active or passive. In some occasions, other possible reasons, such as dangerous driving behaviors, are mainly responsible for traffic accidents, whereas in other cases, drivers are involved in crashes in a passive manner (i.e., they do not exhibit dangerous driving behaviors). Therefore, all crash-involved drivers can
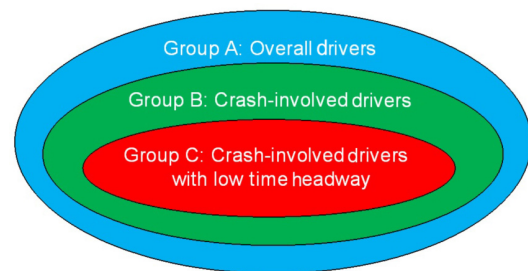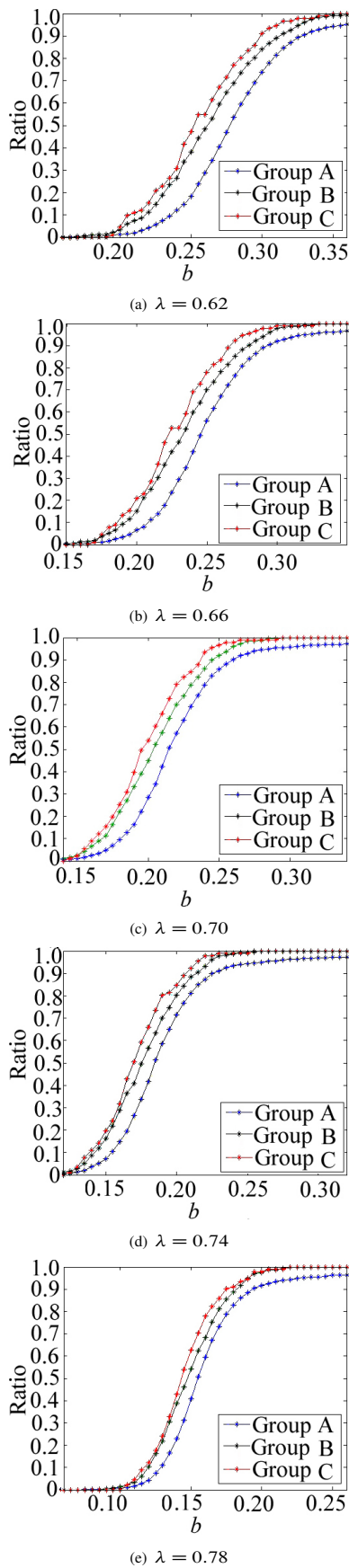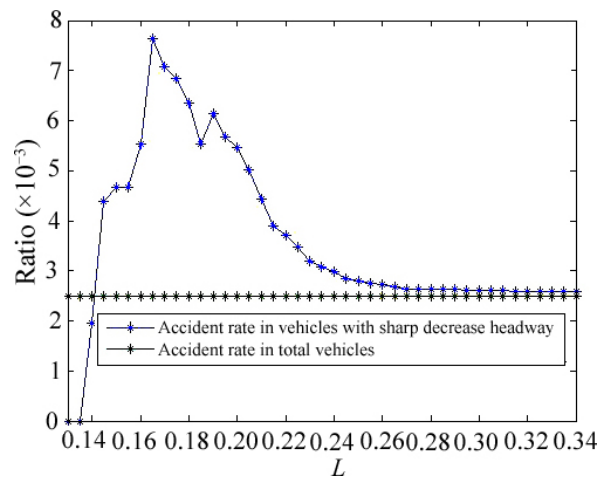

Fig. 3　Illustration of three different groups of vehicles in the case study.

(a) $\lambda = 0.62$

(b) $\lambda = 0.66$

(c) $\lambda = 0.70$

(d) $\lambda = 0.74$

(e) $\lambda = 0.78$

Fig. 4   Three different groups of vehicles in the case study under different forgetting factor λ and threshold b.

Fig. 5   Accident rate of vehicles with evident headway declines compared with other vehicles.

be loosely categorized into two categories according to their different time headways; of the two, the first category is directly related to dangerous driving behaviors (Group C in Fig. 3), which is of significant research interest. Next, to further test and verify our assumption that time headway decline is an important indicator of dangerous driving behaviors characterized by low time headways, we compared the percentages of vehicles with time headway declines in Group B and Group C. As shown in Fig. 4, vehicles in both groups that are involved in crashes and manifested low time headways have one thing in common: they have more time headway declines regardless of the parameter values. This finding suggests that a great proportion of vehicles with low time headways undergo a decline tendency of time headways. Therefore, even if the absolute value of time headway is still acceptable, raising the alarm for those drivers is necessary once the decrease of time headway is detected so as to reduce the risk of traffic accidents in a timely manner.

5   A   Practical   Warning   Strategy   for Individual Drivers

The case study in the previous section demonstrates the need to issue timely warnings to drivers regarding time headway declines. However, the judgment result of Formula (10) is difficult to understand for users. Thus, an indicator to represent the different severities of dangerous driving behavior is proposed. The indicator expected remaining life can be easily interpreted and is thus widely used in industrial process monitoring. This is measured by the remaining time before a headway falls below a pre-specified dangerous threshold as the decreasing tendency is observed. The one-step ahead

predictor given in Eq. (3), which is easy to understand, can also be used as basis through which the multi-step prediction of time headway H at time t can be established. This is given by

$$\hat{H}(t+q|t) = \hat{\beta}(t)^q H(t) \tag{11}$$

When $\hat{\beta}(t) < 1$, the prediction value of time headway exponentially decays over time. An alert limit can be pre-specified as H_{tol}. When the time headway is still tolerable, we thus have

$$\hat{H}(t+q|t) \geqslant H_{\text{tol}} \tag{12}$$

Then the remaining life at time t can be defined as

$$q_{\text{remain}} = \min_q\{q|\hat{H}(t+q|t) \geqslant H_{\text{tol}}\} = \log_{\hat{\beta}(t)}\left(\frac{H_{\text{tol}}}{H(t)}\right) \tag{13}$$

Next, a typical case of crash-involved driver is presented to illustrate the rationality and practicality of the proposed early-warning approach. As shown in Fig. 6, the red ellipse indicates day 0 at which a crash occurs. Before the crash, the value of time headway has evident decreasing tendency, especially from the 3rd to the 1st days. During these days, the value of $\hat{\beta}(t)$ stays significantly below 1 with high confidence level, as indicated by the values of $L(t)$. Hence, time headway decline is successfully recognized. Afterwards, the remaining useful life becomes positive and gradually approaches zero, indicating an increasing risk of traffic accident. This finding is in line with the physical truth of time headway data and suggests that the decreasing headways can reveal the potential risk of crashes.

Notice that the estimation value of remaining useful life may be negative in some occasions, and only becomes valid when $\hat{\beta}(t) < 1$ and $H(t) > H_{\text{tol}}$. However, the warning information in other cases should be considered. We determine the different combinations of two groups of inequalities $\hat{\beta}(t) < (\geqslant)1$ and $H(t) > (\leqslant)H_{\text{tol}}$, and design the corresponding

Fig. 6   Illustration of the early-warning strategy for a crash-involved driver.

alert strategies in Table 1.

When $H(t) \leqslant H_{\text{tol}}$ holds, the value of time headway is extremely low, and the condition is particularly dangerous. Under these circumstances, although the remaining useful life prediction can still be calculated according to Eq. (13), an explicit warning of extremely low time headway should be provided. However, when $\hat{\beta}(t) \geqslant 1$ and $H(t) > H_{\text{tol}}$, the value of time headway tends to increase, and there is no need to raise a warning, thus rendering the remaining useful life meaningless. We should also consider the information provided by $L(t)$. As mentioned in Section 3, the parameter $P(t)$ characterizes the uncertainty of estimation results, and $L(t)$ is proposed as normalized variance estimation. Hence, to avoid confusing drivers, the alert should be withdrawn when the value of $L(t)$ is large and when the result becomes untrustworthy.

Although the on-board equipment can record driving behavior-related data via Controller Area Network-BUS or other sensors, the public-use penetration of such novel equipment is still quite low, especially for time headway measurements that require an expensive forward radar. An urban traffic surveillance system, which has been widely deployed in most modern cities as a basic component of intelligent transportation systems, can facilitate accurate headway and speed monitoring for individual vehicles/drivers. Such a system allows traffic administrations to fully utilize such urban traffic surveillance data that, in turn, can be analyzed to help provide timely suggestions or alerts for individual drivers. As mentioned earlier, meeting this goal is of considerable importance in ensuring the safety of urban traffic road systems. The proposed early-warning strategy can be used in conducting preliminary attempts in mining useful and abundant information from traffic-generated big data. Hence, future investigations are required.
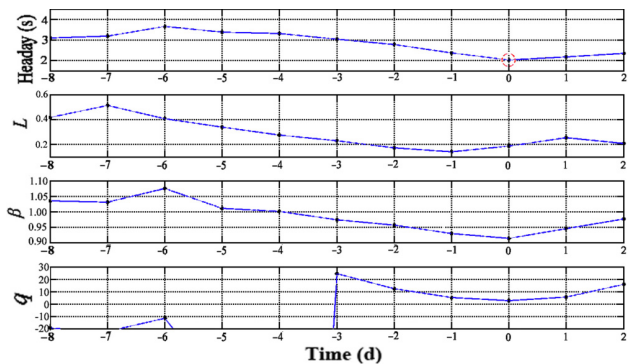
6   Conclusion Remarks

In this article, we developed an AR time-series model to detect dynamic variations of time headways. Using this model, the decline trend of time headways, as

Table 1   Warning strategies under different conditions.

	$H(t) > H_{\text{tol}}$	$H(t) \leqslant H_{\text{tol}}$
$\hat{\beta}(t) < 1$	Alert of remaining useful life	Alert of extremely low time headway
$\hat{\beta}(t) \geqslant 1$	No alert	Alert of extremely low time headway

a key indicator of dangerous driving behaviors, can be effectively detected. Case studies based on real traffic data in a city in northern China indicate that the percentage of time headway decline in crash-involved drivers is significantly higher than that in non-crash-involved drivers. Furthermore, the percentage of time headway decline in crash-involved drivers with low time headway is higher than that among all crash-involved drivers. This finding indicates that the decline of time headway should be considered an important indicator of dangerous driving behaviors as they can induce crashes. Finally, a practical warning strategy for crash prevention is developed based on the notion of remaining useful life predictions.

Future research opportunities lie in the following aspects. First, time headway is not only determined by driving behaviors but also by the external environment. Relative values of time headway in different cross-sections are commonly adopted in existing research to eliminate the impacts of external environment factors, especially traffic conditions in different cross-sections. Applying AR models to relative values of time headways may be helpful in further improving the performance of the proposed recognition method. Second, an accurate estimation of daily average time headway can be obtained from massive data generated by urban surveillance cameras. Hence, how to maintain the performance of the proposed dangerous driving behavior recognition approach and early-warning strategy method in the presence of a limited number of surveillance cameras is worthy of future investigations.

References

[1] B. E. Sabey and H. Taylor, The known risks we run: The highway, in *Societal Risk Assessment*, R. C. Schwing and W. A. Jr. Albers, eds. Boston, MA, USA: Springer, 1980, pp. 43–70.

[2] American Association of State Highway and Transportation Officials, Highway safety manual, National Research Council, Washington, DC, USA, 2010.

[3] I. Lewin, Driver training: A perceptual-motor skill approach, *Ergonomics*, vol. 25, no. 10, pp. 917–924, 1982.

[4] M. C. Taylor, D. C. Lynam, and A. Baruya, The effects of drivers' speed on the frequency of road accidents, TRL report 421, Transport Research Laboratory, UK, 2000.

[5] R. Elvik, P. Christensen, and A. H. Amundsen, Speed and road accidents: An evaluation of the power model, Nordic road and transport research, Oslo: Institute of Transport Economics, 2004.

[6] K. Vogel, A comparison of headway and time to collision as safety indicators, *Accid. Anal. Prev.*, vol. 35, no. 3, pp. 427–433, 2003.

[7] J. Paefgen, T. Staake, and E. Fleisch, Multivariate exposure modeling of accident risk: Insights from pay-as-you-drive insurance data, *Transp. Res. Part A Policy Pract.*, vol. 61, pp. 27–40, 2014.

[8] S. Husnjak, D. Perakovi, I. Forenbacher, and M. Mumdziev, Telematics system in usage based motor insurance, *Procedia Eng.*, vol. 100, pp. 816–825, 2015.

[9] A. Heino, H. H. van der Molen, and G. J. S. Wilde, Differences in risk experience between sensation avoiders and sensation seekers, *Personal. Individ. Differ.*, vol. 20, no. 1, pp. 71–79, 1996.

[10] A. Heino, H. H. van der Molen, and G. J. S. Wilde, Risk perception, risk taking, accident involvement and the need for stimulation, *Saf. Sci.*, vol. 22, nos. 1–3, pp. 35–48, 1996.

[11] W. van Winsum and A. Heino, Choice of time-headway in car-following and the role of time-to-collision information in braking, *Ergonomics*, vol. 39, no. 4, pp. 579–592, 1996.

[12] L. Evans and P. Wasielewski, Do accident-involved drivers exhibit riskier everyday driving behavior? *Accid. Anal. Prev.*, vol. 14, no. 1, pp. 57–64, 1982.

[13] H. X. Chen, X. Pei, Z. Zhang, D. Y. Yao, Q. J. Feng, and Z. Z. Wang, Driving behavior differences between crash-involved and crash-not-involved drivers using urban traffic surveillance data, presented at the IEEE International Conference on Service Operations and Logistics, and Informatics, Beijing, China, 2016.

[14] P. Händel, J. Ohlsson, M. Ohlsson, I. Skog, and E. Nygren, Smartphone-based measurement systems for road vehicle traffic monitoring and usage-based insurance, *IEEE Syst. J.*, vol. 8, no. 4, pp. 1238–1248, 2014.

[15] W. Guo, Y. Zhang, and L. Li, The integration of CPS, CPSS, and ITS: A focus on data, *Tsinghua Sci. Technol.*, vol. 20, no. 4, pp. 327–335, 2015.

[16] C. Shang, F. Yang, X. Q. Gao, X, L. Huang, J. A. K. Suykens, and D. X. Huang, Concurrent monitoring of operating condition deviations and process dynamics anomalies with slow feature analysis, *AIChE J.*, vol. 61, no. 11, pp. 3666–3682, 2015.

[17] C. Shang, B. Huang, F. Yang, and D. X. Huang, Probabilistic slow feature analysis based-representation learning from massive process data for soft sensor modeling, *AIChE J.*, vol. 61, no. 12, pp. 4126–4139, 2015.

[18] H. T. Liao, W. B. Zhao, and H. R. Guo, Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model, presented at RAMS '06 Annual Reliability and Maintainability Symposium, Newport Beach, CA, USA, 2006.

[19] X. S. Si, W. B. Wang, C. H. Hu, and D. H. Zhou, Remaining useful life estimation-a review on the statistical data driven approaches, *Eur. J. Oper. Res.*, vol. 213, no. 1, pp. 1–14, 2011.

[20] G. Li, J. Qin, Y. D. Ji, and D. H. Zhou, Reconstruction based fault prognosis for continuous processes, *Control Eng. Pract.*, vol. 18, no. 10, pp. 1211–1219, 2010.

[21] Z. L. Deng, *Self-Tuning Filtering Theory with Applications*: *Modern Time Series Analysis Method*, (in Chinese). Harbin, China: Press of Harbin Institute of Technology, 2003.
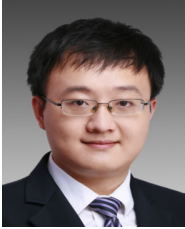
Hongxin Chen receievd the BS degree in 2010 from Tsinghua University, China. He is a PhD candidate at Tsinghua University with major research interests on traffic safety, driving behavior analysis, and intelligent transportation system.

Shuo Feng receievd the BS degree in 2014 from Tsinghua University, China. He is a PhD candidate at Tsinghua University with major research interests on traffic flow analysis, connected vehicles systems, and control theory.

Zuo Zhang received the BEng, MEng, and PhD degrees from Tsinghua University, China, in 1989, 1991, and 1995, respectively. She has been a faculty member with Department of Automation, Tsinghua University, after graduation in 1995, with one-year visit to Harvard University from September 1999 to July 2000. Since late 2010, she served as the Deputy Dean of School of Information Science and Technology in Tsinghua University. She is currently a professor with Department of Automation, Tsinghua University. Her current research interests include multi-source data acquisition, fusion and analysis, system control, and optimization techniques for intelligent transportation systems.

Xin Pei received the BS and MS degrees from Tsinghua University, China, in 2005 and 2007, respectively, and the PhD degree from the University of Hong Kong in 2011. She is currently a lecturer with Department of Automation, Tsinghua University. Dr. Pei serves as a journal reviewer for several international journals and conference including *Accident Analysis and Prevention*, *Transportmetrica*, and ITSC conference, etc. Her current research interests include road safety evaluation and driving behavior analysis.
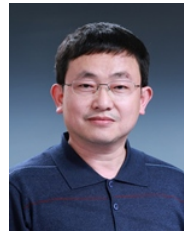
Danya Yao is currently a professor at Tsinghua University. He received the BEng, MEng, and PhD degrees from Tsinghua Unversity in 1988, 1990, and 1994, respectively. His active research areas include vehicle-infrastructure cooperation systems, advanced detection technology, and systems engineering. Prof. Yao has published more than 90 papers, including 20 SCI/EI indexed articles. Prof. Yao is PI/co-PI of more than 5 national research programs including 863 program, 973 program, and NSFC project. He is currently the Chief Expert of the National High-Tech Research and Development Program (863 Program) "Research on Key Technologies of Intelligent Vehicle-Infrastructure Co-operation System".