A Case-Based Approach for Modelling the Risk of Driver Fatigue

Qiaoting Zhong, Guangnan Zhang

A Case-based Approach for Modelling the Risk of Driver Fatigue

Qiaoting Zhong and Guangnan Zhang*

Center for Studies of Hong Kong, Macao and Pearl River Delta, Institute of Guangdong Hong Kong and Macao Development Studies, Sun Yat-Sen University, Guangzhou, 510275 China.

Abstract. Fatigue-related crashes are one of the major threats to road safety worldwide. Despite the substantial work in the domain of transportation science by both the industry and academia, there are few studies in applying case-based reasoning (CBR) approach to modelling the risk of driver fatigue. This research explores the potential for fatigued driving using a database of 16,459 traffic crashes reported from 21 cities in Guangdong province, China from 2006 to 2010. The CBR system under development differentiates between fatigued-driving and non-fatigued-driving cases based on various personal and environmental traffic characteristics. The advantage of using CBR in modelling fatigued driving has been demonstrated through empirical evaluation.

Keywords: intelligent information processing, intelligent decision making, case-based reasoning, traffic safety management, driver fatigue

1 Introduction

Driver fatigue has already become a leading factor contributing to traffic crashes around the world [48]. In China, fatigued driving caused 887 (9.26%) of all highway crashes in 2011, resulting in 520 (8.1%) deaths and over RMB 37 million (10.82%) property losses [43]. To decrease the occurrences of traffic crashes and promote road safety, studying drivers' decision on driving under fatigue condition is urgent. Despite the substantial work in the domain of transportation science by both the industry and academia, there are few studies in applying case-based reasoning approach to modelling the risk of driver fatigue.

Case-based reasoning (CBR) has been used in both cognitive science and artificial intelligence [32], it makes use of the most similar previous cases to solve new problems [32, 30, 27]. Since the early 1980s, CBR has been successfully applied in various fields such as legal reasoning [42, 5, 6, 45], planning [28, 14, 25, 15, 34], E-commerce [41, 18], medical diagnosis [4, 19, 9, 40], incident management [49, 24], and risk analysis [35, 23]. CBR is the best fit for modelling fatigued driving because of its ability to simultaneously process a large number of highly interrelated variables to arrive at a decision [20, 13]. Moreover, CBR models have significant merits as compared to statistical ones (*e.g.*, the widely employed matched case-control logistic regression [2, 31])

* Author to whom all correspondence should be addressed. E-mail: zsuzgn@hotmail.com; Tel.: 020-84110641

and other artificial intelligence models (*e.g.*, neural network) regarding the comprehensibility of output [13].

The novelty of the research work presented lies in the following aspects. First of all, this work explores the potential for using CBR in dealing with driver fatigue management. Secondly, this research examines a highly comprehensive set of risk factors that relate to the driver fatigue. A database of 16,459 real traffic crashes reported from 21 cities in Guangdong province, China from 2006 to 2010 is used, and a total of 50 variables in 21 categories are examined in this study. Compared with the logistic regression model that uses the same variables, our proposed CBR model had better performance in modelling fatigued driving on AUC measures in all simulations. The outcome of our proposed CBR system can help targeting the group of drivers that may intend to drive under fatigue condition, which constitutes the crucial part of driver safety management in an intelligent transportation system.

This paper is organised as follows. Section 2 introduces previous relevant work from the literature. Section 3 describes the methodology. Section 4 presents the case study and empirical evaluation of our approach. Section 5 concludes the study.

2 Related Work

In recent years, there had been an increase in uptake of case-based reasoning (CBR) concepts in the research area of traffic control and management. For instance, Jagannathan et al. [22] developed a CBR prediction system that is capable of differentiating between the accident and non-accident cases. Mounce et al. [36] designed a CBR system to help selection of signal timing plans. Sadek et al. [39] developed a prototype CBR routing system. Li and Zhao [33] applied CBR to intersection control. Kofod-Petersen et al. [26] presented a prototype implementation of a CBR system that predicts traffic flow and calculates signal plans for urban intersections. CBR methods have also been used to model the vehicle control behaviour (*i.e.*, steer, throttle, and brake) of teen drivers [37]. Different from the previous ones, this work aims to explore the potential for using CBR in dealing with driver fatigue management.

The risk of driver fatigue is related to a combination of situational and individual factors. The increased risk may result from a mix of biological, personal, road and environmental related factors [29]. Previous studies on detecting driver's state of fatigue have used various biological and personal indicators, such as visual indicators (*e.g.*, face images and eye state [12, 17, 10]), physiological sensor signals (*e.g.*, EEG [46] and ECG [7]) and driving states (*e.g.*, accelerate, brake, shift and steer [21]). However, none of them (including the CBR approach of identifying the drivers' stress state [7]) considered road and environmental features that are crucial towards drivers' intention to drive under fatigue condition. By using CBR in this research, the risk of driver fatigue is modelled based on past similar circumstances considering not only personal but also road and environmental characteristics. Moreover, most previous studies verify their models through experiments with driving simulators, whereas we validate ours by using the reliable official source of traffic crash data.

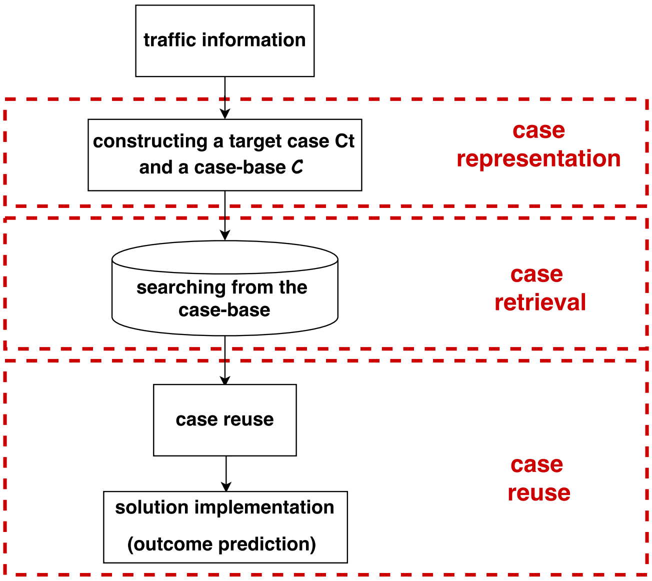

Fig. 1. Schematic Diagram Showing Steps Involved In Modelling Fatigued Driving

3 System Design

Following the work of [27, 33], we propose a case-based system of fatigued driving that consists of three main components: representation module, retrieve module, and reuse module. Figure 1 shows a schematic overview of the system architecture. The proposed system receives traffic information, through the *representation module*, a target case C_t and a case-base $\mathcal{C}$ are constructed based on such information. The *retrieve module* retrieves the most similar previous case from the case-base, and *reuse module* further manipulates the retrieved cases for the best outcome of the target case.

3.1 Case Representation

A *case* in the system represents a previous traffic situation [33], and it is composed of a case number, a case description, and a case outcome. Formally, a case can be represented as follows:

- Case $C = \langle N, D, O \rangle$, where,
  - N is the case number;
  - D is the case description;
  - O is the case outcome (*i.e.*, fatigued-driving or non-fatigued-driving);

$\mathcal{C}$ is used to denote the case base that contains all possible cases.

The description of a case depicts the situation when the case appears [33, 26]. Four categories of attributes are needed to describe the case. The first category contains the demographic information of the driver. The second one describes the vehicle information. The third category consists of road situation factors. The fourth category contains some environmental factors such as the weather situation. Formally, we have:

- Case description $D = \langle DD, VI, R, E\rangle$, where,
  - DD is the demographic information of the driver such as gender;
  - VI is the vehicle information such as vehicle type;
  - R is the road situation such as road type;
  - E is the environmental factors such as bad weather.

For example, a simple case (No. 0) that a crash occurred on the expressway at about 8 am on Saturday was caused mainly because the fatigued driving of the male motorcyclist would be represented as $\langle$0, $\langle\{$male$\}$, $\{$motorcycle$\}$, $\{$expressway$\}$, $\{$moring, weekend$\}\rangle$, fatigued-driving$\rangle$.

3.2 Case Retrieval and Reuse

The crucial part of the retrieval module is the similarity measure, which calculates the similarity between the target case and the cases in the case-base [22]. For each target case, the similarity between the target case and the cases in the case-base is measured by calculating the Euclidean distance (the most widely used distance metric in CBR) between them. Given a target case t and a case c in the case-base, the similarity between both cases is:

$$S(t,c) = \sqrt{\sum_a (D_{t,a} - D_{c,a})^2}$$

where $D_{t,a}$ is the normalised value of attribute a $\in DD, VI, R, E$ in the target case t and $D_{c,a}$ is the normalised value of attribute a in the case c from the case base. Note that all the attributes are of equal importance in this paper for simplicity reasons.

With the above similarity measure, the most similar cases will be retrieved by using the K-nearest neighbour (KNN) algorithm [11], where K refers to the number of neighbours. The KNN algorithm, widely used in classification problems to assign objects to classes, matches each attribute in the target case to its corresponding attribute in the retrieved case [22]. The predicted outcome of the target case is then obtained by the majority vote of its K nearest neighbours, *i.e.*, *the most common outcome of the K retrieved cases are reused and assigned to the target case directly*.

4 Case Study: Based on the Traffic Crash Data in China

In this study, we use the traffic crash data for the period 2006-2010 in Guangdong Province, China. These data are extracted from the Traffic Management Sector-Specific Incident Case Data Report, the Road Traffic Accident Database of China's Public Security Department. They are the only officially available source of traffic crash data in China. Data are recorded and reported by the traffic police on-scene who conducted assessments and provided feedback immediately to the headquarters of the Traffic Management Department. These reports include characteristics of drivers, vehicle features,

Table 1. Case Structure

Case description		
Category	**Feature**	**(Binary) Attributes**
demography	gender	*male*
	age	*0-24, 25-44, 45 and above*
	residence	urban
	driving experience	*0-2 years, 3-5 years, above 5 years*
	occupation	*clerk, migrant worker, farmer, self-employed, other*
vehicle	license condition	*valid*
	safety condition	*poor*
	insured or not	*insured*
	overloaded or not	overloaded
	vehicle type	*passenger car, truck, motorcycle*
	commercial or not	commercial
road	road type	*express, first-class highway, second-class or below highway, urban expressway, urban ordinary highway, other*
	traffic lane	isolated
	road surface	*dry, wet, other*
	traffic control device	*no device, signal, sign, other*
environment	light condition	*daytime, with street lighting in the night, without street lighting in the night*
	weather	*rainy*
	weekend or not	*weekend*
	holiday or not	holiday
	time	*00:00-06:59 (midnight to dawn), 07:00-08:59 (morning rush hours),12:00-13:59 (noon),17:00-19:59 (afternoon rush hours), other*
	season	*spring, summer, fall, winter*
Case outcome: *fatigued-driving, non-fatigued-driving*		

road conditions, the time of crashes, the environmental context for each crash and the cause of crashes such as traffic violations like driving under fatigue condition [47]. To experiment with the proposed approach, we set up a case-based system in which each of the samples in our datasets was a case. Table 1 shows the case structure used in the experiment where relevant binary attributes are selected based on [48]. According to working time patterns and peoples' lifestyles in China, we classify five groups of time.

4.1 Evaluation Settings

To evaluate our proposed case-based system of drivers' risk of fatigued driving, the data set was partitioned into a training set and a test set. To be aligned with the literature (*e.g.*, [3]), we varied the percentage of the training and test data to study the possible variations of performance based on different partitions. In particular, ratios 10:90, 30:70, 50:50, 70:30 and 90:10 were used as the portions of training and test data. Each case in the test set was consecutively made the target case with the cases in the training set constituting the case-base. For each portion of training and test data, KNN algorithms

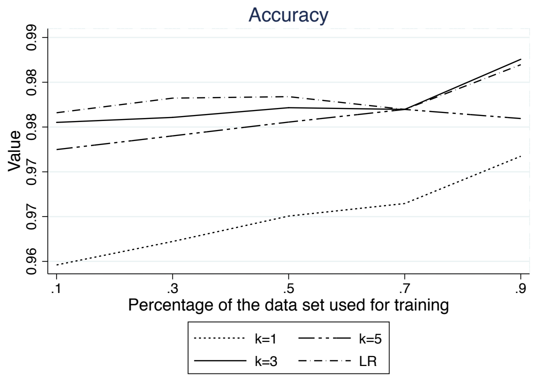

Fig. 2. Accuracy of the different models

with different numbers of K were run. Ties may occur in a binary classification problem if $K > 1$, to avoid ties, the most similar case ($K = 1$), the three most similar cases ($K = 3$), and the five most similar cases ($K = 5$) were retrieved in the current study. In comparison, the widely employed logistic regression was also conducted for each partition. The dependent variable indicated the occurrence of driving under fatigue condition in a crash, while the independent variables under consideration were the same as the CBR setting.

The performance of a computational intelligent system can be measured in many different ways. Typically, the models' predictive performances are measured regarding *overall accuracy*, *e.g.*, the percentage of fatigued-driving and non-fatigued-driving classified accurately in the current research [16]. However, other measures like the *true positive rate* (TPR, also called recall or sensitivity) tends to be the key concern when infrequent events (as fatigued driving in the current study) are to be predicted [44]. Complementary to TPR that measures the proportion of positives that are correctly identified as such, another primary evaluator is the *false positive rate* (FPR, also known as fall-out or false alarm ratio) that calculates the ratio of negatives wrongly categorised as positive. Based on pairs of TPR and FPR, a receiver operating characteristic (ROC) curve could be obtained to compare the performance of different models. We calculate the area under the ROC curve based on the work of Cantor and Kattan [8], and use IBM SPSS Statistics 23 to perform all the statistical analysis.

4.2 Results and Discussion

The available data set contained 16,459 traffic crashes, of which 384 (2.33%) were fatigue-related. Among all the fatigue-related crashes, 99% of drivers were male, 88% of them had more than two years of driving experience, and 72.4% came from urban areas. Drivers that were involved in fatigue-related crashes included self-employed workers (23.7%), migrant workers (18.5%), farmers (17%), clerks (8%), and employees from other professions (32.8%). Concerning vehicle information, trucks and motorcycles constituted 56% and 16.7%, respectively. For road situation, 65.9% of these fatigue-related crashes occurred in separate lanes, and 38% occurred in expressway.

Table 2. Evaluation Results (TPR for true positive rate and FPR for false positive rate)

	10:90				30:70				50:50			
	k=1	k=3	k=5	LR	k=1	k=3	k=5	LR	k=1	k=3	k=5	LR
TPR	0.12	0.01	0	0	0.16	0.04	0.02	0	0.22	0.06	0.02	0
FPR	0.02	0	0	0	0.02	0.01	0	0	0.02	0	0	0

	70:30				90:10			
	k=1	k=3	k=5	LR	k=1	k=3	k=5	LR
TPR	0.26	0.11	0.04	0	0.23	0.17	0.1	0
FPR	0.02	0	0	0	0.01	0.01	0	0

Considering environmental features, 51.8% of the fatigue-related crashes occurred during 00:00-06:59, and 38% took place at daytime. Seasonal distribution of the fatigue-related crashes: 32.3%, 26.3%, 24.7%, and 16.7% occurred in the summer, winter, spring, and fall, respectively.

Figure 2 shows the comparison of the overall accuracy for all simulations. In overall, all simulations received very high accuracy in modelling fatigued driving, *i.e.*, more than 96 percent. Logistic regression (LR) had a better performance in comparison to the proposed CBR method (no matter what the value of K) for most training/test partitions. When 70% or above of the dataset was used for training, the CBR method with K=3 achieved slightly higher accuracy than LR.

Table 2 summarises the results of true positive rate (TPR) and false positive rate (FPR) from all simulations. In all scenarios, we observed that the FPR is nearly zero, indicating that both LR and the proposed CBR method have excellent performance in correctly identifying non-fatigued-driving cases. However, LR received zero TPR for all the training/test partitions, which reflects that LR is relatively biased towards the majority class in our dataset (*i.e.*, the non-fatigued-driving case). Such bias may be caused by the imbalance situation of our dataset (*i.e.*, only 2.33% cases were fatigue-related). Recall that TPR reflects the hits of fatigued driving reality, in such circumstances, although LR got very high accuracy and very low FPR, it is meaningless [44]. In contrast, our proposed CBR models with K=1 or K=3 had shown their merits by receiving non-zero TPR in all the training/test partitions.

Figure 3 illustrates the trade-off between the TPR and the FPR of each simulation. Our proposed CBR models (no matter what the value of K) had better performance than LR in modelling fatigued driving on AUC measures. Simulation of K=1 outperformed the rest as the percentage ratio of the training and test data were varied.

To sum up, we can see that the proposed CBR model with K=1 and that with K=3 have their strengths in modelling fatigued driving. To further differentiate them, we calculated the *precision* values in various partitions of training and test data. Precision measures the probability that a driver with a positive screening test indeed conducted fatigued driving. As shown in Fig. 4, when 10% of the data set were used for training, the K=1 model had higher precision value, whereas the K=3 model had better performance when more than 10 percent were used for training. Further empirical validation is required to study the optimal value of K for modelling fatigued driving.

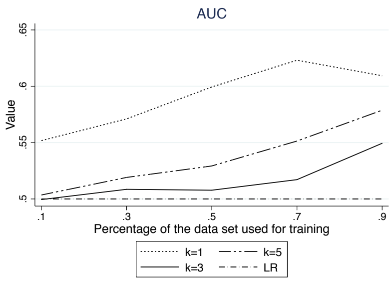

Fig. 3. AUC of the different models

Fig. 4. Precision of the K=1 and K=3 models

4.3 Further Implications

Fatigue-related crashes are one of the major threats to road safety worldwide. The important implication of our research is: if driving under fatigue condition could be reduced or controlled successfully through the early detection, the occurrences of a crash would be reduced accordingly. When designing road traffic interventions to decrease incidences, it is well established that a change in the driver's attitude is of utmost importance. The outcome of our proposed CBR model can help targeting the group of drivers that may have an intention to drive under fatigue condition. In particular, our proposed CBR model can highlight the important features of such group of drivers in the aspects of various personal, vehicle, road, and environmental conditions. Specific countermeasures by integrating these highlighted features into a driver fatigue detection component of the intelligent transportation systems will be effective in improving traffic safety [38]. The results obtained in this study may likely be generalised to other provinces in China because the crash data in Guangdong Province and features considered are rather representative and comprehensive. The fact that Guangdong has a high percentage of residents migrating from other provinces, and is having one of the largest numbers of vehicles among all provinces in China reinforces the generalisability [48]. Furthermore, the findings of this study can also contribute as a reference to future road safety research for other countries.

5 Limitations and Future Work

A live case-based reasoning system has a "4 REs" cycle (*i.e.*, retrieve, reuse, revise, and retain) [1]. Since our proposed case-based system is not operating on live data, there is no component designed in the current study for the standard revise and retain process. Our proposed system could be extended to deal with live data by adding a new part of revise and retain in the future.

The advantage of using case-based reasoning in modelling fatigued driving has been demonstrated through comparing with logistic regression. In getting higher values of recall, precision and AUC, an important consideration is case attribute selection and weighting. The attributes we have chosen are based on previous work. However, it is likely that fatigued driving also occurs due to other conditions that have not yet been studied. Moreover, the contribution of each attribute to the similarity between the target case and the cases in case-base could differ as well. Thus, it is worth testing out different attributes and weighted assignments for attributes considered.

The empirical experiments show that the CBR system under development is capable of differentiating between fatigued-driving and non-fatigued-driving cases to some extent based on certain personal and environmental traffic characteristics. The evaluation of our proposed method is not exhaustive and could be expanded in several directions, for instance, to compare with other artificial intelligence techniques (*e.g.*, neural network). It would also be worth evaluating whether or not our proposed approach could be fruitfully applied to other risky behaviours rather than the kind we have considered.

Acknowledgements

This research was supported in part by the National Natural Science Foundation of China grant 71573286 and Ministry of Education Project for Humanities and Social Sciences Research (16JJDGAT006).

References

1. Aamodt, A., Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI communications 7(1), 39–59 (1994)
2. Abdel-Aty, M., Pande, A., Lee, C., Santos, V.G.C.D.: Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways. Journal of Intelligent Transportation Systems 11, 107–120 (2007)
3. Adedoyin, A., Kapetanakis, S., Petridis, M., Panaousis, E.: Evaluating case-based reasoning knowledge discovery in fraud detection. In: Workshops Proceedings of the 24th International Conference on Case-Based Reasoning. pp. 182–191 (2016)
4. Althoff, K.D., Bergmann, R., Wess, S., Manago, M., Auriol, E., Larichev, O.I., Gurov, S.I.: Case-based reasoning for medical decision support tasks: the inreca approach. Artificial Intelligence in Medicine 12(1), 25–41 (1998)
5. Ashley, K.D.: Modeling Legal Arguments: Reasoning with Cases and Hypotheticals. MIT press (1991)
6. Athakravi, D., Satoh, K., Broda, K., Russo, A.: Generating legal reasoning structure by answer set programming. In: Proceedings of the 8th International Workshop on Juris-informatics. pp. 24–37 (2014)
7. Begum, S., Ahmed, M.U., Funk, P., Filla, R.: Mental state monitoring system for the professional drivers based on heart rate variability analysis and case-based reasoning. In: Proceedings of Federated Conference on Computer Science and Information Systems. pp. 35–42 (2012)
8. Cantor, S.B., Kattan, M.W.: Determining the area under the roc curve for a binary diagnostic test. Medical Decision Making 20(4), 468–470 (2000)
9. Chattopadhyay, S., Banerjee, S., Rabhi, F.A., Acharya, U.R.: A casebased reasoning system for complex medical diagnosis. Expert Systems 30(1), 12–20 (2013)
10. Chen, D., Wang, Z., Zhuo, Q., Wu, F.: A new algorithm for fatigue detection in driving. In: Foundations of Intelligent Systems. pp. 1041–1049 (2014)
11. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE transactions on information theory 13(1), 21–27 (1967)

12. Devi, M.S., Bajaj, P.R.: Driver fatigue detection based on eye tracking. In: Proceedings of the First International Conference on Emerging Trends in Engineering and Technology. pp. 649–652 (2008)
13. Goh, Y.M., Chua, D.K.H.: Case-based reasoning for construction hazard identification: case representation and retrieval. Journal of Construction Engineering and Management 135(11), 1181–1189 (2009)
14. Hammond, K.: CHEF: A model of case-based planning. In: Proceedings of the Fifth National Conference on Artificial Intelligence. pp. 267–271 (1986)
15. Hammond, K.J.: Case-based planning: Viewing planning as a memory task. Elsevier (2012)
16. Han, J., Kamber, M., Pei, J.: Data Mining: Concepts and Techniques (third ed.). Morgan Kaufmann (2011)
17. He, J., Roberson, S., Fields, B., Peng, J., Cielocha, S., Coltea, J.: Fatigue detection using smartphones. Journal of Ergonomics 3(3), 1–7 (2013)
18. Heras, S., Jordn, J., Botti, V., Julian, V.: Argue to agree: A case-based argumentation approach. International Journal of Approximate Reasoning 54(1), 82–108 (2013)
19. Holt, A., Bichindaritz, I., Schmidt, R., Perner, P.: Medical applications in case-based reasoning. The Knowledge Engineering Review 20(3), 289–292 (2005)
20. Hossain, M., Muromachi, Y.: A bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways. Accident Analysis and Prevention 45, 373–381 (2012)
21. H.Wang, Liu, H., Song, Z.: Fatigue driving detection system design based on driving behavior. In: Proceedings of the 2010 International Conference on Optoelectronics and Image Processing. pp. 549–552 (2010)
22. Jagannathan, R., Petrovic, S., Powell, G., Roberts, M.: Predicting road accidents based on current and historical spatio-temporal traffic flow data. In: Proceedings of 2013 International Conference on Computational Logistics. pp. 83–97 (2013)
23. Jin, R., Han, S., Hyun, C., Cha, Y.: Application of case-based reasoning for estimating preliminary duration of building projects. Journal of Construction Engineering and Management 142(2), `DOI:http://dx.doi.org/10.1061/(ASCE)CO.1943--7862.0001072` (2016)
24. Johnson, C.: Using case-based reasoning to support the indexing and retrieval of incident reports. In: Proceedings of European Safety and Reliability Conference: Foresight and Precaution. pp. 1387–1394 (2000)
25. Khattak, A., Kanafani, A.: Case-based reasoning: a planning tool for intelligent transportation systems. Transportation Research Part C: Emerging Technologies 4(5), 267–288 (1996)
26. Kofod-Petersen, A., Andersen, O.J., Aamodt, A.: Case-based reasoning for improving traffic flow in urban intersections. In: Proceedings of the 22nd International Conference on Case-Based Reasoning. pp. 215–229 (2014)
27. Kolodner, J.: Case-Based Reasoning. Morgan Kaufmann (1993)
28. Kolodner, J.L.: Retrieval and organisation strategies in conceptual memory: A computer model. Erlbaum, NJ (1984)
29. Lal, S.K., Craig, A.: A critical review of the psychophysiology of driver fatigue. Biological psychology 55(3), 173–194 (2001)
30. Leake, D.B.: CBR in context: the present and future. In: Leake DB (ed) Case-based reasoning, experiences, lessons, & future directions. p. 330 (1996)
31. Lee, C., Abdel-Aty, M., Hsia, L.: Potential real-time indicators of sideswipe crashes on freeways. Transportation Research Record: Journal of the Transportation Research Board 1953, 41–49 (2006)
32. Li, K., Waters, N.M.: Transportation networks, case-based reasoning and traffic collision analysis: A methodology for the 21st century. In: Methods and Models in Transport and Telecommunications: Cross Atlantic Perspectives. pp. 63–92 (2005)

33. Li, Z., Zhao, X.: A case-based reasoning approach to urban intersection control. In: Proceedings of the 7th World Congress on Intelligent Control and Automation. pp. 7113–7118 (2008)
34. Liu, W., Hu, G., Li, J.: Emergency resources demand prediction using case-based reasoning. Safety Science 50, 530–534 (2012)
35. Lu, Y., Li, Q., Xiao, W.: Case-based reasoning for automated safety risk analysis on subway operation: Case representation and retrieval. Safety Science 57, 75–81 (2013)
36. Mounce, R., Hollier, G., Smith, M., Hodge, V.J., Jackson, T., Austin, J.: A metric for pattern-matching applications to traffic management. Transportation Research Part C: Emerging Technologies 29, 148–155 (2013)
37. Ontañón, S., Lee, Y.C., Snodgrass, S., Bonfiglio, D., Winston, F.K., McDonald, C., Gonzalez, A.J.: Case-based prediction of teen driver behavior and skill. In: Proceedings of the 22nd International Conference on Case-Based Reasoning. pp. 375–389 (2014)
38. Paaver, M., Eensoo, D., Kaasik, K., Vaht, M., Mäestu, J., Harro, J.: Preventing risky driving: a novel and efficient brief intervention focusing on acknowledgement of personal risk factors. Accident Analysis and Prevention 50, 430–437 (2013)
39. Sadek, A.W., Demetsky, M.J., Smith, B.L.: Case-based reasoning for real-time traffic flow management. Computer-Aided Civil and Infrastructure Engineering 14(5), 347–356 (1999)
40. Sharaf-El-Deen, D.A., Moawad, I.F., Khalifa, M.E.: A new hybrid case-based reasoning approach for medical diagnosis systems. Journal of medical systems 38(2), 1–11 (2014)
41. Shimazu, H.: Expertclerk: A conversational case-based reasoning tool for developing salesclerk agents in e-commerce webshops. Artificial Intelligence Review 18(3-4), 223–244 (2002)
42. Simpson, R.L.: A computer model of case-based reasoning in problem solving: An investigation in the domain of dispute mediation. In: Georgia Institute of Technology, School of Information and Computer Science Technical Report no. GIT-ICS-85/18 (1985)
43. Traffic Management Bureau, Ministry of Public Security, PRC, China: China Road Traffic Accidents Annual Statistical Report 2011 (2011)
44. Weiss, G.M., Hirsh, H.: Learning to predict extremely rare events. In: Proceedings of the AAAI'2000 Workshop on Learning from Imbalanced Data Sets. pp. 64–68 (2000)
45. Yeow, W.L., Mahmud, R., Raj, R.G.: An application of case-based reasoning with machine learning for forensic autopsy. Expert Systems with Applications 41(7), 3497–3505 (2014)
46. Zhang, C., Wang, H., Fu, R.: Automated detection of driver fatigue based on entropy and complexity measures. IEEE Transactions on Intelligent Transportation Systems 15(1), 168–177 (2014)
47. Zhang, G., Yau, K.K., Chen, G.: Risk factors associated with traffic violations and accident severity in china. Accident Analysis and Prevention 59, 18–25 (2013)
48. Zhang, G., Yau, K.K., Zhang, X., Li, Y.: Traffic accidents involving fatigue driving and their extent of casualties. Accident Analysis and Prevention 87, 34–42 (2016)
49. Zubair, M., Khan, M.J., Awais, M.M.: Prediction and analysis of air incidents and accidents using case-based reasoning. In: Proceedings of the Third Global Congress on Intelligent Systems. pp. 315–318 (2012)