(Boyce and Geller 2002)驾驶模式：driving pattern

5的文献综述

Reliable accident prediction and proactive prevention are undoubtedly of great benefit and necessity. Accident occurrence is a complex mechanism, with many contributing factors (Mannering et al., 2016). Generally, driver-centric factors could be found in most crash accidents, and driving behaviour assessment is an important aspect to enhance safety and reduce crashes. There is a perennial quest about assessing driving behaviour and predicting crash risk potentials in driving.

Many studies have been conducted to evaluate driving behaviour. Typical approaches include self-reported questionnaires, simulatorbased experiments, and naturalistic driving studies (NDS) (Hong et al., 2014). In NDS, various characteristics and variables about both driving and drivers’ behaviours are investigated, based on detailed information recorded using sensors, in-vehicle devices, and even smartphones (Eftekhari and Ghatee, 2018). A range of features has been developed to describe unsafe behaviours, such as speeding, abrupt braking or jerk, tailgating, frequent and intense lane change, yaw, among others (Bagdadi, 2013; Wahlström et al., 2017). Generally, statistical profiles of movement-related variables are used as behaviour features. For reliable driving assessment and risk prediction, in-depth and multi-view mining of features are necessary, especially features with predictability. Data labelling of risk levels is another challenging but valuable work. As a complement to accident data, surrogate measures of vehicle conflicts are well accepted for safety evaluation (Zheng et al., 2014; Chai and Wong, 2015a). Many indicators have been developed to reflect certain views of risk potentials, such as Time to Collision (TTC). Shi et al. (2018) retrieved two real-world accident cases, and developed hybrid indicators to identify pre-accident risk signals. Besides, kinematic characteristics are also commonly used for risk evaluation (Wu and Jovanis, 2013). For example, a rapid evasive manoeuvre is used to flag a near-crash event (Perez et al., 2017). However, detailed risk levels are inherently problematic to determine, since accurate classification and well-fitting thresholds are difficult to establish. A clear and synthetical assessment of risk levels is therefore of great interest, but is still lacking.

Furthermore, risk assessment is a distinctly imbalanced problem. The class imbalance issue has been discussed in various concepts, such as the safety pyramid model suggested by Hydén (1987). Algorithms often display bias in favour of the numerical majority class (i.e., safe and low-risk cases), and ignore or wrongly discard the minority instances (i.e., events of higher risk levels), which might be treated as noise or outliers (Díez-Pastor et al., 2015). Besides, there are five main observed challenges in class imbalance, related to intrinsic data characteristics (López et al., 2013; Beyan and Fisher, 2015). Class imbalance problems make the behaviour assessment and risk prediction much more complicated. The focus of this study is to extract and select behaviour features for driving assessment and risk prediction using machine learning.

The methodology is introduced in Section 2. Section 3 elaborates on extracting driving behaviour features and risk indicator features. The learning-based feature selection and risk prediction are described in Section 4. The final two sections cover the discussion and conclusions.

动态的行为导向驾驶风险评估模型

（+为什么要评估驾驶风险？）由于As驾驶风险预测和评估能够为社会和个人带来的巨大潜在好处，关于驾驶风险，不同领域的研究者已经使用过许多传统的办法尝试进行预测和评估。

但是他们并没有具体情况具体分析，有的结果是考虑不全面的，甚至是矛盾的

1通过量化问卷调查结果，发现饮酒较多的人，发生事故的风险也较高，并将醉驾者的personality按聚类为五类（并非为了预测，仅仅是临床诊断和分析，且长问卷、主观性）

2 将driving risk定义为In this paper, driving risk is defined as a potential threat that may cause vehicle crashes or other accidents. /Hence, the driving risk level can be represented by the braking process characteristics./ K-means cluster is used to cluster different driving risk levels involved in a near-crash case using the braking process features./ the velocity when braking and triggering factors have the largest influence on the driving risk level, which, to some extent, are in accordance with some previous studies.

3

1、（Personality Subtypes Among Driving-While-Intoxicated Offenders: Relationship to Drinking Behavior and Driving Risk）本研究探讨了临床和理论上的经验推导。醉酒驾驶被捕的男性中，有意义的子类型。通过对驾驶态度的聚类分析，确定了五个子类型。性格和敌意测量。有两个亚型被发现有特别的 高水平的风险增强特征。第一种特征是最高的。抑郁和怨恨的程度，以及最低水平的自信。情绪调整和感知控制。第二个特点是 与驾驶有关的攻击性、竞争速度、寻求感觉的水平最高。攻击性、易怒性、间接性和言语敌意。随后的分析 表示，这两个亚型的社会地位也较低，在社会上的地位也较低。

饮酒较多的人，发生事故的风险也较高。

Five distinct groups were derived based upon differential levels of driving-related attitudes, personality functioning, and hostility. Three of these groups (Clusters 1, 3, and 4) appeared to be more deviant, having particularly high levels of risk-enhancing characteristics relative to the overall DWI sample norms; the remaining two groups (Clusters 2 and 5) had relatively low levels on these characteristics, evidencing considerably lower levels across the affective and behavioral dimensions assessed.

问题：问卷主观性、长问卷保真度问题、

2、是一个**experiment**

**Using acquired real-world driving data, a near-crash database is built, which contains vehicle status, potential crash object, driving environment and road type, and weather condition. K-means cluster analysis is applied to classify the near-crash cases into different driving risk levels using braking process features, namely maximum deceleration, average deceleration and percentage reduction in the vehicle kinetic energy. The results indicate that the velocity when braking and triggering factors have strong relationship with the driving risk level involved in near-crash cases.**

K-means cluster is used to cluster different driving risk levels involved in a near-crash case using the braking process features.

the velocity when braking and triggering factors have the largest influence on the driving risk level, which, to some extent, are in accordance with some previous studies.

3、跟2基本一样，只是2用了聚类，3用的是CART，结论也一样

4、啊这，就是卞一洋老师这文章2333

5、Based on extracted risk indicator features, vehicles are clustered into various groups labelled with graded risk levels. Data under-sampling of the safe group is performed to reduce the risk-safe class imbalance. Afterwards, the linkages between behaviour features and corresponding risk levels are built using XGBoost, and key features are identified according to feature importance ranking and recursive elimination. The risk levels of vehicles in driving are predicted based on key features selected.

Findings show that this approach is effective and reliable to identify important features for driving assessment, and achieve an accurate prediction of risk levels.

5.4. Limitations

For clustering-based risk grading, the major challenge is the lack of crash instances, which makes it hard to verify the linkages between the highest risk level and actual crash occurrence. A potential way of validation is to examine the accident records of the drivers/vehicles which are clustered as higher risk levels.

The feature extraction is far from being exhaustive. In-depth feature extraction is recommended to further improve modelling, which should cover a broader range of driving behaviours and risk conditions, such as lane-changing, conflicts between motorcycles and vehicles. The interests of feature extraction are mainly twofold, namely, making risk assessment more reliable, and providing early signals for risk-based crash prediction.

6、基于驾驶员追尾碰撞避免过程的风险评估模型--受手机使用和性别影响

- 基于驾驶模拟器的研究

为了预警驾驶员，提示他不要做一些事儿而做的模拟实验——使用手机撞击风险明显更高（不管是不是handsfree，驾驶需要注意力）

As for the gender effect, the results indicated that female drivers had longer reaction time than male drivers in critical situation, but they were more quickly in braking with larger maximum deceleration rate, and thus kept larger safety margin with the leading vehicle compared to male drivers.

the present study provided an alternative and probable paradigm for measuring the effect of intervention factors and driver characteristics on on-road collision avoidance

performances.

局限性：p47

7、An evaluation of relationships between age, personality and driving style revealed that driver age and type. A personality characteristics were significant predictors of vehicle speed and following distance, P0.05. However, contrary to

the earlier research, which relies heavily on a self-reported driving criterion, no significant gender differences were obtained.

年轻人比老年人反应更快，更容易熟悉路况

Proponents of personality psychology argue that some people are more prone than others to take risks. Specifically, more than a decade ago, theoretical formulations and research findings suggested that at-risk driving (e.g. non-use of safety belts, speeding, and driving while intoxicated (DWI)) are components of a larger risky driving syndrome (Jessor, 1987; Beirness and Simpson, 1988; Donovan et al., 1988; Wilson and Jonah, 1988). Furthermore,

it was suggested that certain social, psychological and behavioral factors (e.g. health compromising behaviors such as smoking, drug use and heavy drinking) distinguished young drivers (ages 12–19), who became involved in vehicle crashes from those who did not (Beirness and Simpson, 1988), and were significant predictors of DWI, driving while using marijuana, and willingness to ride with alcohol-impaired drivers.

8、distraction and secondary tasks

9、一本书？？太长了再见

10、The paper analyses risk behaviour as described by a group of convicted drink-drivers.

This paper analyses risk assessment and risk behaviour in relation to drink-driving.

We interviewed 25 participants—20 men and five women—of five courses in three different

locations. Contact with the interviewees was established in the following way.

11、Driving risk assessment using driving behavior data under continuous tunnel environment

12、Individual driver risk assessment using naturalistic driving data

CNC risk for individual drivers is associated with CIE rate, age, and personality characteristics. Furthermore, the CIE rate is an effective predictor for high-risk drivers.

13、我们根据以下信息推测风险评估将对应一个基于经验的技能领域，具有隐性知识的特征：数据显示风险评估随着经验的提升而提升。

The driving knowledge tests required individuals to provide safe-speed and crash-risk judgments across a variety of conditions. The correlations and odds ratios obtained for these scales demonstrate the importance and relevance of this knowledge to driver safety.

14、MobiDriveScore – A System for Mobile Sensor Based Driving Analysis A Risk Assessment Model for Improving One’s Driving

通过危险驾驶行为，评估分数，（we propose a novel method named MobiDriveScore for routine

ventures using which the consumer can assess his/her own driving pattern. Thus the consumer can consciously reduce the risk associated with his/her driving by using MobiDriveScore.）

15、雨天高速路驾驶风险评估

developed based on drivers’ subjective questionnaire and its performance is validated by using actual crash data.

to estimate the relationship between drivers’ perceived risk and factors, including vehicle type, rain intensity, traffic volume, and location.

16、Probabilistic Analysis of Dynamic Scenes and Collision Risk Assessment to Improve Driving Safety

主要deals with urban driving

直接预测，并不是基于驾驶员历史行为，而是直接预测可能性。基于车而非基于人

17、也是主要基于环境和车而非基于人

我们提出了一种在人工场理论框架内评估道路车辆所承担的风险的方法，设想用于安全分析和设计驾驶支持/自动化应用。在这里，任何障碍物（道路上的相邻实体）都被视为一个有限标量风险场，该风险场是在主体车辆的预测配置空间中制定的。驾驶风险估计是标的车辆未来位置的风险场的强度。这个风险域是两个因素的乘积：碰撞概率和。预期碰撞能量。与相邻车辆的碰撞概率是根据以下几点来估计的 概率运动预测。风险可以评估单个时间步骤或多个未来的时间步骤，取决于所需的估计的时间分辨率。我们验证了在自然数据集的三种近似碰撞情况下，以及在切入和插入的情况下，采用单步法。仿真的硬制动场景，并展示了多步法在的应用。选择最安全的路径，在一个车道段。拟议方法的风险描述 定性地反映了情况的叙述，与《时间》基本一致。碰撞。与目前的代用安全措施相比，拟议的风险估计提供了： 考虑到不确定性，为评估个体车辆的驾驶安全提供了更好的依据。对未来环境交通状况和预期碰撞后果的严重度。拟议的 驾驶风险模型可以作为智能汽车安全应用的一个组成部分。并作为评估交通安全的综合代用措施。

18、In general, safety can be defined as the absence of unreasonable risk.

**文献综述**

**pedestrian-collision risk could be predicted based on driving behavior upon commencement of steering when making an across-traffic turn.**

**In addition, we plan to examine the influence of other elements of the traffic environment (e.g., road geometry, existence of other road users) on the driving-behavior features to extend the range of applications of the results of the present study.**

**（只考虑了一个因素，什么时候开始拐弯）**

**A limitation of this study is the relatively small sample size.**

**Another limitation is the limited experimental conditions of the driving scenario. The experiment was conducted in an environment where no other traffic participant exists and therefore findings**

**of this research are limited to the scope of this elementary right-turn driving scenario.**

19、不是实景的，是想象环境，然后打分

Incidents of angry and aggressive driving, often referred to as “road rage,” are becoming more and more commonplace in everyday driving. Many people might benefit from counseling interventions to help manage driving anger and aggression. This article provides a review of research on road rage risk factors, a description of inventories for assessing angry and aggressive driving tendencies, and a discussion of intervention strategies to reduce angry driving behavior.

Given the risk factors associated with particular demographically

identified groups, it may make sense to target certain drivers, particularly younger, inexperienced drivers, for road rage prevention efforts.

路怒的特征—减少和避免，尤其是经验少的和年轻的，说年轻人更容易犯问题。但是时代背景不同，文化程度和社会环境不一样，我们还能持有这种想法吗

但是仅靠驾驶经验和年龄并不能认定一个人的驾驶行为，

drivers, younger drivers

(ages 16–25) tend to display a more risky driving style, drive

faster, accept narrower gaps when pulling into traffic, leave

shorter distances between cars, and are more likely to violate

traffic lights (Summala, 1987)。

我们可以把年龄的原因归结为经验的多少

有的人做研究说男人比女人更激进，但还有人并没有做的出来，发现男女其实很像，也没有什么所谓的差异

Factor in aggressive and risky driving. Men have been found to commit more dangerous traffic violations and engage in more risky driving behavior than do women (Ellison-Potter et al.,

2001; Lawton, Parker, Manstead, & Stradling, 1997; Parker, Manstead, Stradling, & Reason, 1992). Male drivers have been observed to react more aggressively than do women to

congestion (Shinar, 1998) as well as to reckless, inconsiderate, and impatient driving of others (Parker et al., 2002). In addition, male drivers may be more prone than women to engage in revengeful and physically aggressive thinking (Deffenbacher, Petrilli, Lynch, Oetting, & Swaim, 2003) and physical and verbal aggression (Hennessy & Wiesenthal,

2001), particularly men who exhibit high trait driving anger (Deffenbacher et al., 2000).

Several studies (Novaco, Stokols, Campbell, & Stokols, 1979; Novaco, Stokols, & Milanesi, 1990; Stokols, Novaco, Stokols, & Campbell, 1978) have demonstrated that traffic congestion

and travel impedance can negatively affect mood, behavior, and health of commuters. Simply stated, daily driving, particularly in conditions of high traffic congestion, can be a

source of annoyance and stress. But do the stress and irritation associated with traffic congestion lead to angry and aggressive driving? The findings have been mixed. Some researchers

(Hennessy & Wiesenthal, 1999; Shinar, 1998) have found driver anger and aggression to be reported more often in high-congestion conditions than in low-congestion conditions,

whereas others (Lajunen, Parker, & Summala, 1999; Underwood, Chapman, Wright, & Crundell, 1999) have not

found any relationship between congestion and reports of

driver anger and aggression. Perhaps there is an interaction

effect such that some individuals are more predisposed to

respond with anger under conditions of congestion or impedance

while driving. For example, drivers who are generally

prone to getting angry while driving may be particularly

anger-prone and aggressive under high impedance conditions

(Deffenbacher, Deffenbacher, Lynch, & Richards, 2003). There

may also be different reactions of individuals depending on

whether they perceive being impeded as a result of another

driver (e.g., someone driving too slow) versus other situational

conditions (e.g., road construction).

Similar to other forms of aggression, aggressive

环境易使人路怒，但是他没有具体到驾驶员，没有给驾驶员画像，不能具体到个人去减少损失

后面又讲了personality，但是我们是从驾驶行为，预测人的行为，且在现实场景中

20、Transportation systems comprise various elements including drivers, vehicles, and roads. However, a generalized instability of the driver-vehicle-road closed-loop system can lead to a driving risk owing to these potential dynamic factors, such as driver error, vehicle failure, road condition, and environment state (Laugier et al., 2011; Rolison et al., 2018; Wang et al., 2016). Previous studies have suggested that

drivers’ perceptions and reactions can change in different conditions (Martensen and Dupont, 2013; Morgan and Mannering, 2011; Saifuzzaman and Zheng, 2014).

In the past decades, several studies (Goerlandt and Reniers, 2016; Ni et al., 2010) have assessed various influencing factors that lead to driving risks, which can be divided into objective risk analysis, subjective risk analysis, and combined ones.

光从环境和问卷的主观评估并不准确，我们将从综合的数据

(+Difficulty to identify driving risk, For instance, safe speed in some place or time may be dangerous in others.)（常开路段常开时间有没有考虑——如果仅仅考虑了年龄、开车经验知识、性别等等，也许评估结果看起来较为平稳，但是当他常常进入特定路段如拥堵路段，或者。。。危险性其实就完全不一样——考虑动态环境变化，结合实际实事求是）

事实上，已经有许多人做过driving risk assessment。In the past decades, several studies (Goerlandt and Reniers, 2016; Ni et al., 2010) have assessed various influencing factors that lead to driving risks, which can be divided into objective risk analysis, subjective risk analysis,

and combined ones.

其中获取数据主要包括三种传统的方法，让驾驶者填写问卷，场景模拟实验以及数学概率模型的预测评估（直接收集GPS and on-board diagnostic (OBD II) data 的 the detection and aggregation of high-risk maneuvers by the driver, which typically constitutes of hard bump, sharp cornering, hard stop and speed limit violations**.** ）

Boyce, T. E. and E. S. Geller (2002). "An instrumented vehicle assessment of problem behavior and driving style:: Do younger males really take more risks?" Accident Analysis & Prevention **34**(1): 51-64.