**张翔宇工作汇报（2022.11.16）**

1. **论文阅读**

**1.**

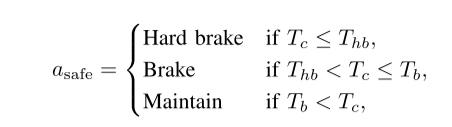
**题目：**

**Deep Reinforcement Learning with Enhanced Safety for Autonomous Highway Driving**

**摘要：**

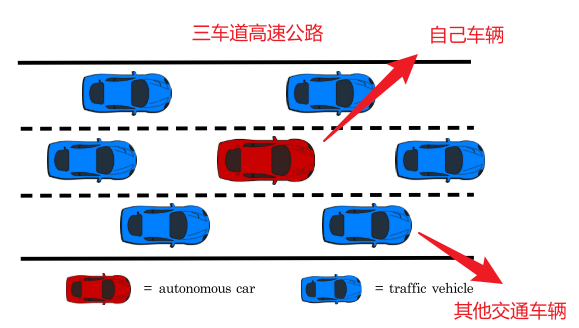
提出了一种用于自动驾驶的安全深度强化学习系统。提出的框架利用基于规则和基于学习的方法的优点来确保安全。我们的**安全系统由两个模块组成，即手工安全和动态学习安全。**手工制作的安全模块是基于常见驾驶实践的启发式安全规则，可确保与交通车辆的相对间隙最小。另一方面，动态学习的安全模块是数据驱动的安全规则，它从驾驶数据中学习安全模式。具体来说，如果将来的状态之一导致接近失误或碰撞，那么将为奖励函数分配负奖励，以避免碰撞并加速学习过程。我们在不同流量密度的模拟环境中演示了所提出框架的功能。结果表明，通过动态学习的安全模块增强了该策略的性能。

**论文整体思路：**

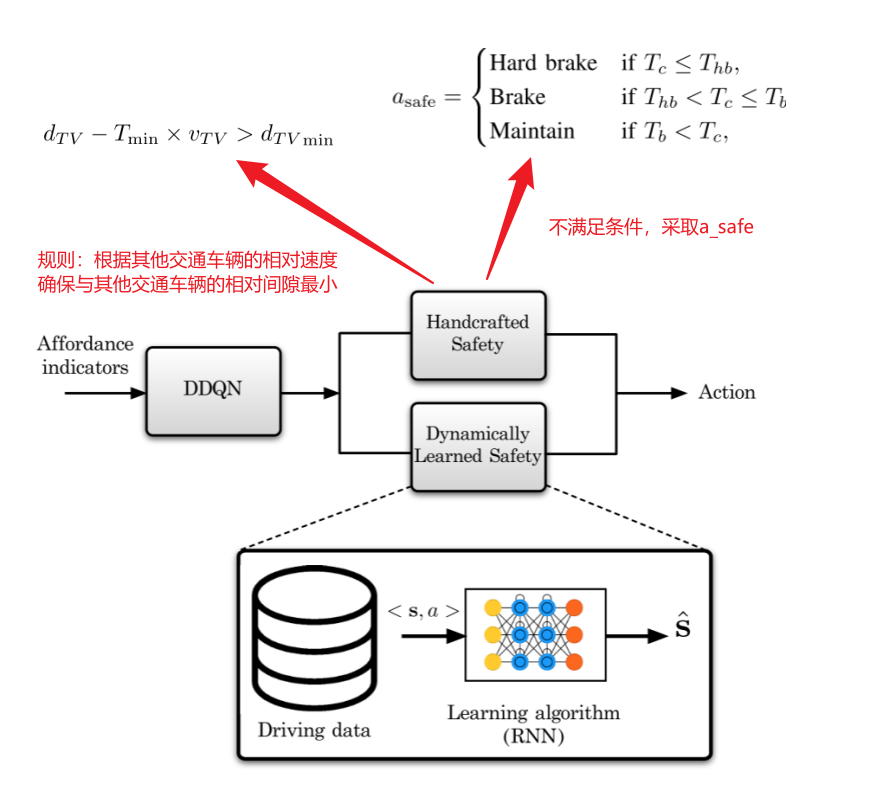
这篇论文主要使用了这个启发式规则去判断当前状态是否安全。不满足这个条件就采取。但是当外部环境变化很快或遇到其他意外情况时，这个规则很难进行处理。为了解决这个问题，这篇文章引入了学习的机制，使用RNN预测有限范围内的未来状态，然后确定未来状态之一是否违反安全规则。如果违反了这个规则，就设定负的reward，以避免碰撞并加速学习过程。

通过与什么机制都不使用、和使用启发式规则，进行对比实验证明，启发式规则+学习机制，累计收益最高。相比于，只用启发式规则，加入学习机制，碰撞次数明显减少了。最终得出结论，加入学习机制，可以减少碰撞次数，提高安全性。

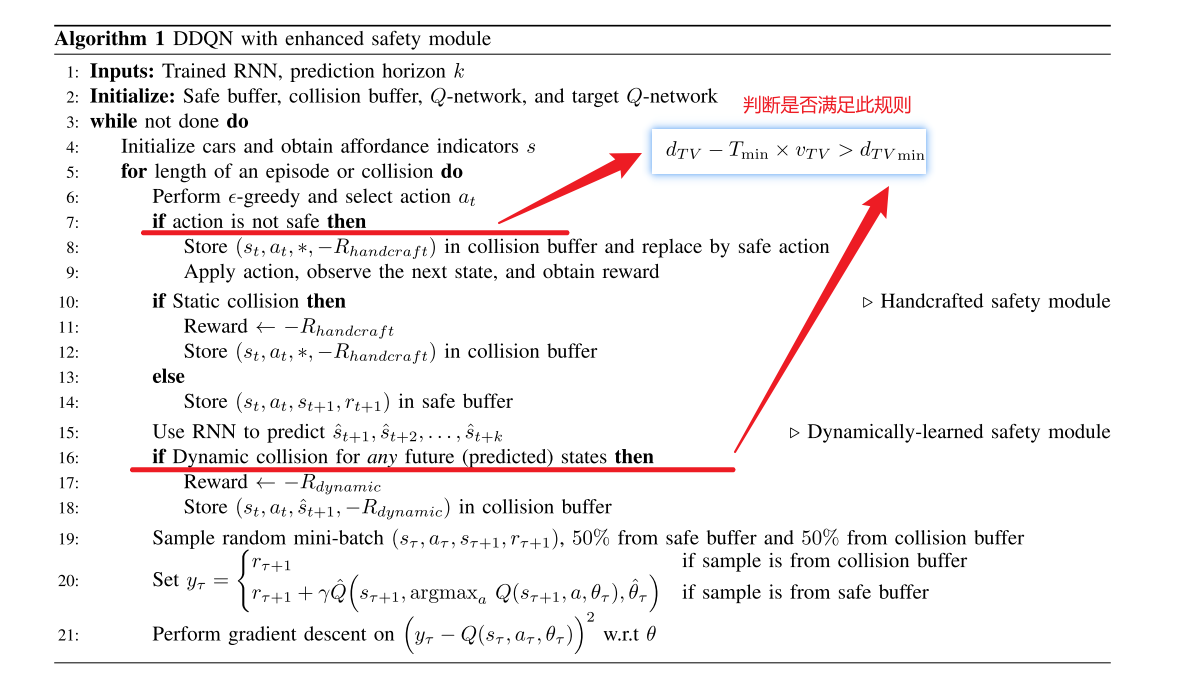
**本文场景：**



**整体架构：**



**算法伪代码：**



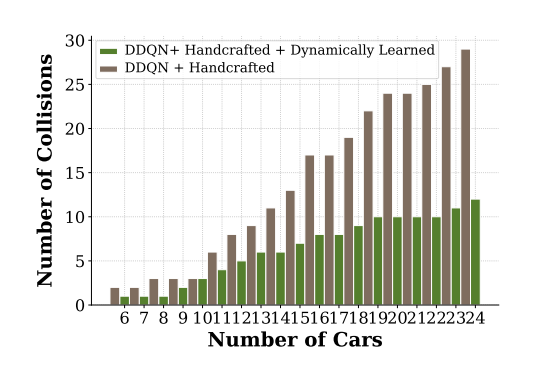
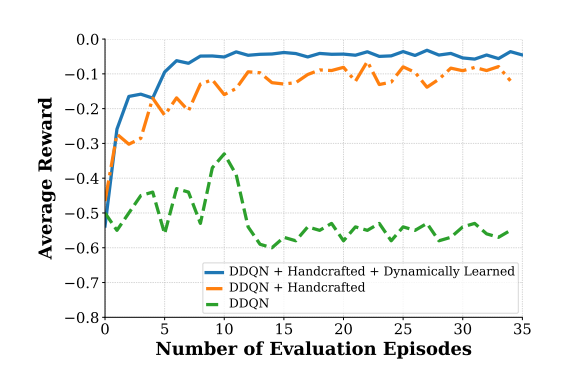
**Handcrafted safety module：**

用规则判定策略生成的action是否安全，不安全则进行负的reward，存入collision buffer；如果判定该行为是安全的，则存入Safe buffer。

**Dynamically-learned safety module：**

输入状态和行为使用RNN进行下一时刻状态预测——>判断预测出来的状态是否安全（用什么进行的判定? ）——>如果预测出来的状态会导致碰撞，将会对其进行施加负的reward，存储到collision buffer中。

**对比实验的结果：**



**结论：**

左图展示出：使用动态学习的安全模块增强的策略优于仅使用手工制作的安全模块的策略，优于无任何安全模块的策略。

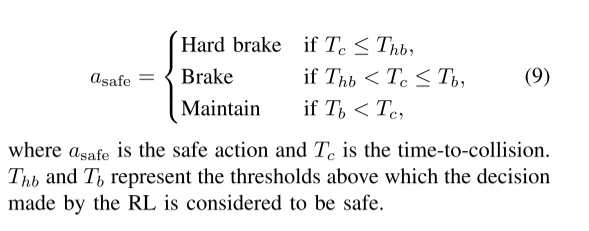
右图展示出：相比于只基于手工模块，通过动态学习的安全模块增强的策略导致的碰撞明显减少。整体趋势是随着车辆数量的增加，碰撞次数增加。

**文章中的heuristic safety rule：**

判断是否符合此不等式：



如果不满足此不等式，则采取



**2.**

**题目：**

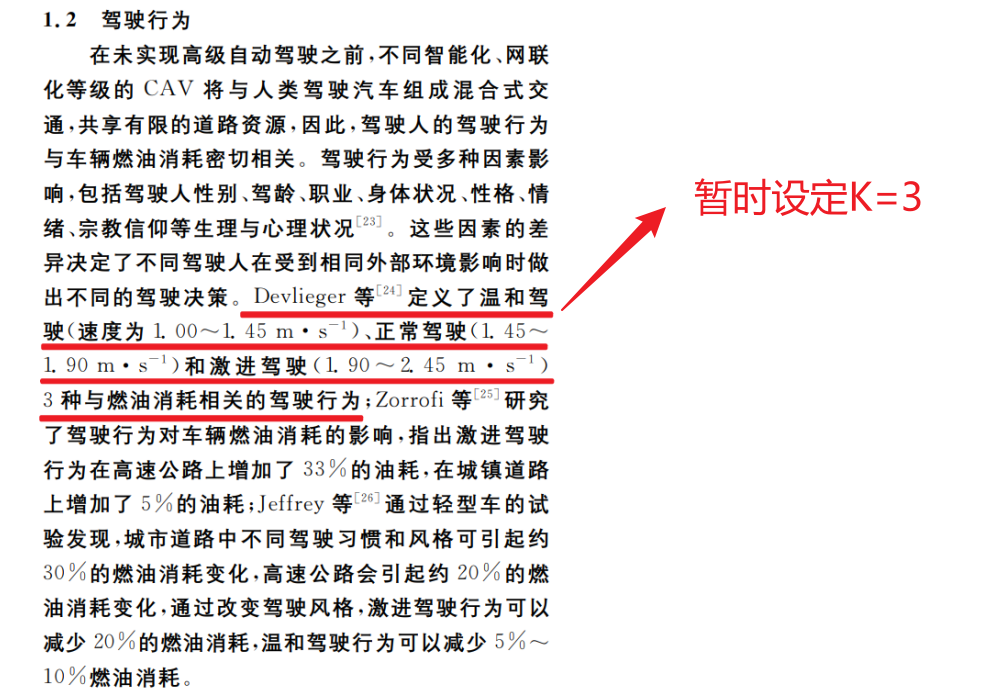
**Incorporating driver preferences into eco-driving assistance systems using optimal control**

近年来，由于对气候变化的担忧，减少燃料消耗，从而减少运输造成的二氧化碳排放，已成为一个热门问题[1]。

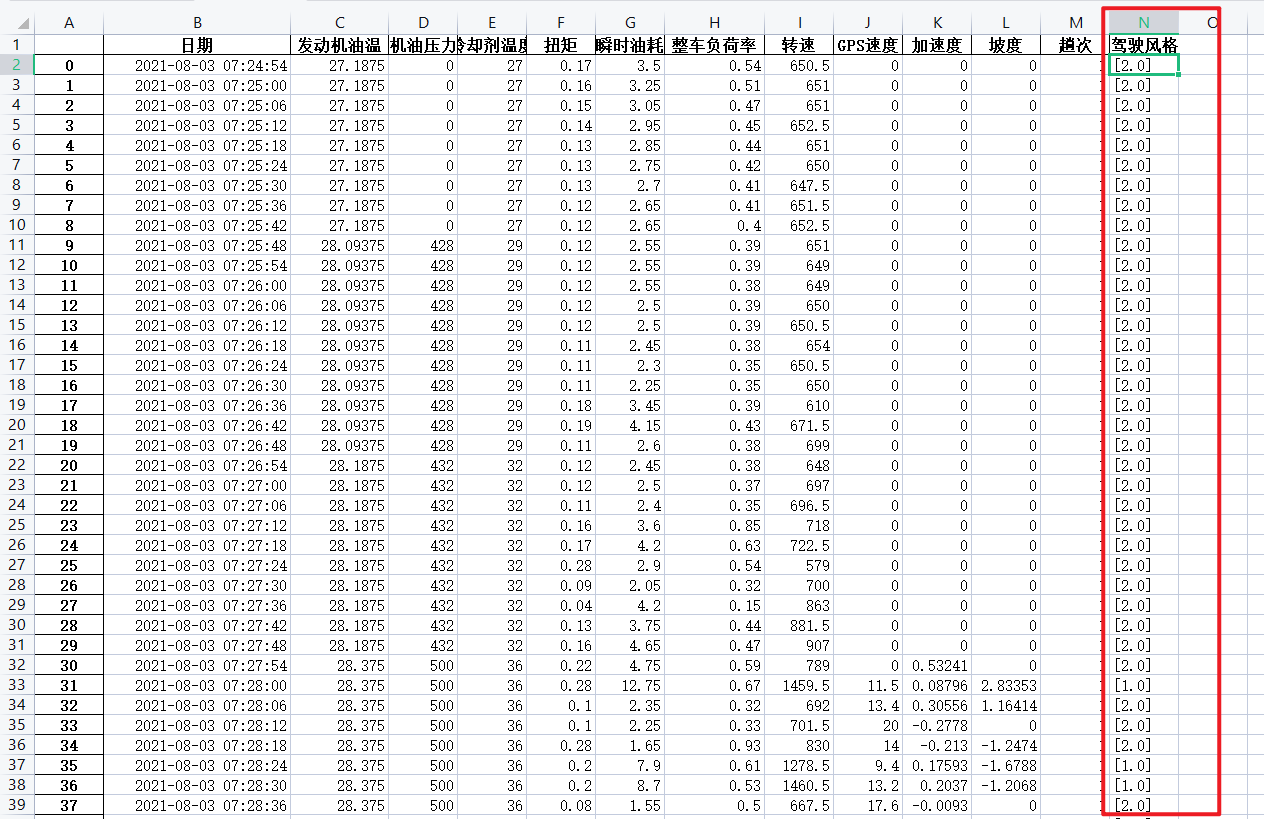
经济驾驶行为，或 “生态驾驶”，已被建议作为一种可以通过鼓励驾驶员缓慢加速，预测信号和交通流量，以避免停车，保持平稳的速度，并避免空转的技术，减少道路车辆10%的二氧化碳排放[2]。

针对这两个问题的潜在补救措施是向驾驶员提供有关其行为的主动反馈，例如通过在车辆内使用听觉，视觉或触觉人机界面 (HMIs) [8]，[9]。使用V2X通信 [10] 或机器学习 [11] 等方法改进这些接口有相当大的研究兴趣。

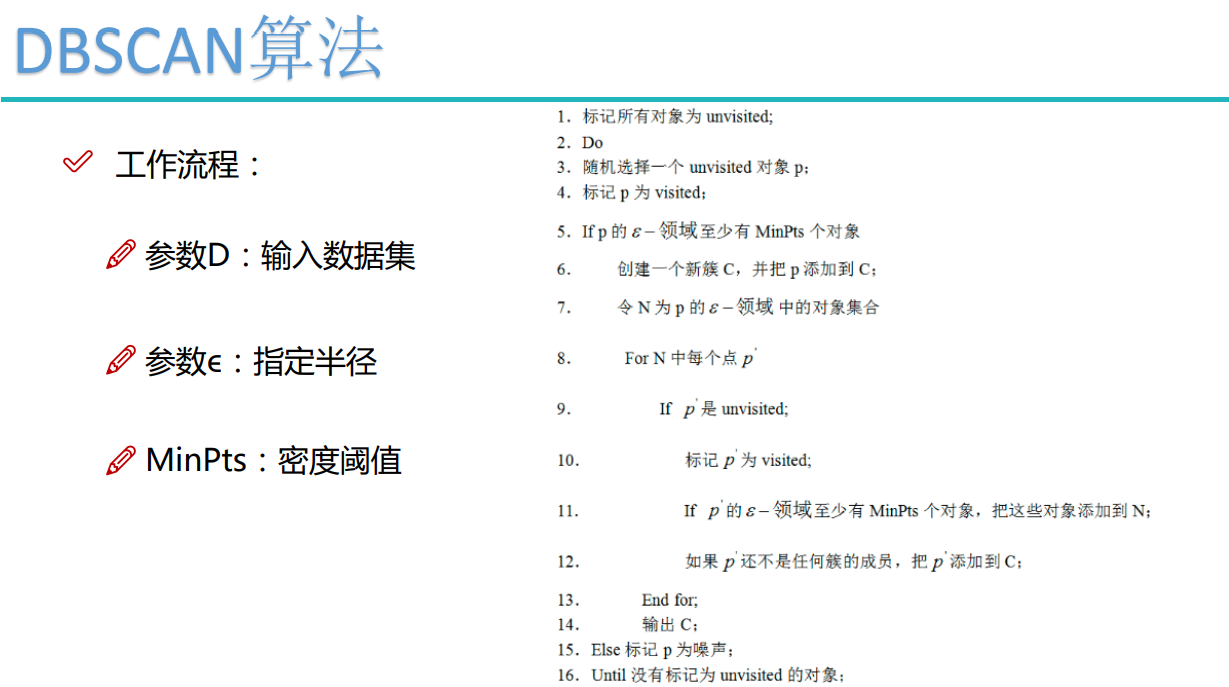
1. M. P. Vandenbergh, J. Barkenbus, and J. Gilligan, “Individual carbon emissions: The low-hanging fruit,” UCLA L. Rev., vol. 55, p. 1701, 2007.
2. J. N. Barkenbus, “Eco-driving: An overlooked climate change initiative,”Energy Policy, vol. 38, no. 2, pp. 762–769, 2010.
3. **实践**
4. 使用K-Means算法对之前暑假的数据进行聚类。



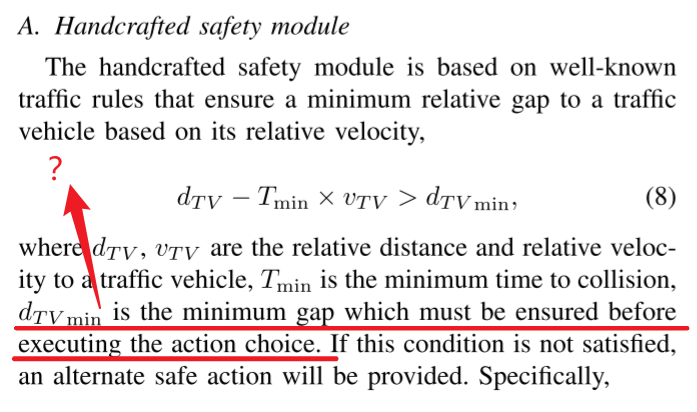
使用K-Means算法之后，最终聚类结果：



1. 了解DBSCAN算法流程



1. 和师兄讨论，订正了论文整体框架。
2. **遇到的问题**
3. 对驾驶行为识别方面，整体实现思路不是很明确。使用聚类进行驾驶行为识别，很难细致到具体行为。正常来说，使用聚类可以划分类别（激进、温和、正常），这样的话能称为驾驶行为识别吗？
4. 整体论文最后落点到生态驾驶上，是想通过对细致的、具体的行为进行油耗预测分析，从而通过建议等落点到生态驾驶上。如果使用聚类只划分了三种不同的驾驶风格的话，最后落点到生态驾驶上，分析风格太过于笼统？
5. 安全性驾驶，论文中的



**四、未来计划**

1. 和老师、师兄师姐讨论，解决上述问题。

2. 对新的数据进行整理和处理。