

## 权利要求书

1、一种缓解交通拥堵驾驶员控制方法，其特征在于，所述缓解交通拥堵驾驶员控制方法采用集中学习但分散执行的设定，每个目标车辆节点在每个时刻均做一个决策，为所有节点实现相同的给定目标，有序通过从而解决交通拥堵问题；节点间的通信和信息传播采用图神经网络 GNN 进行建模，决策处理器采用 Deep Q learning, 形成的决策信息以建议指令的形式下发给每一位环境内的驾驶员。

2、如权利要求 1 所述的缓解交通拥堵驾驶员控制方法，其特征在于，所述缓解交通拥堵驾驶员控制方法，还包括环境建模，对于环境的建模定义为信息拓扑结构；所述环境建模，包括：根据车辆的空间位置和相对关系，将该模型分解为局部和全局两层，所述局部网络是一个星形图，包括目标车辆以及其周围的所有其他车辆个体；所述全局网络由当前环境内的所有车辆个体组成；目标车辆既通过车载传感器从附近的其他车辆获取局部信息，又通过云端连接通道从其他环境内的车辆获取全局信息。

3、如权利要求 2 所述的缓解交通拥堵驾驶员控制方法，其特征在于，在局部星形网络中，信息从周围车辆传递到目标车辆；所有的目标车辆共享其他车辆的局部感知信息。

4、如权利要求 1 所述的缓解交通拥堵驾驶员控制方法，其特征在于，所述缓解交通拥堵驾驶员控制方法，还包括强化学习模型构建，所述强化学习模型结构，包括：在每一个时间步  $t$ ，在目标车辆周围的  $N$  个其他车辆均可被检测到，故每个时间步  $t$  对应的模型空间的输入设置为状态  $s$ ，该状态  $s$  是一个由三个信息模块组成的元组，包括：节点特性  $x_t$ 、关联矩阵  $A_t$ ，记录车辆的索引矩阵  $M_t$ ；其中节点特性  $x_t$  包括速度  $v_t$ 、纵向位置  $p_t$ 、横向车道位置  $l_t$  以及行驶意图  $I_t$ ，关联矩阵  $A_t$  表示目标车辆及其周围车辆之间的交互关系，索引矩阵  $M_t$  用于将目标车辆从所有节点中筛选出来。

5、如权利要求 4 所述的缓解交通拥堵驾驶员控制方法，其特征在于，所述强化学习模型结构，还包括：节点节点特性  $x_t$  输入到全连接 FCN 层，FCN 的输

## 权利要求书

出与关联矩阵  $A_i$  同时输入图神经网络 GCN 层并行计算，将输出结果与索引矩阵  $M_i$  进行点积运算进行车辆节点的筛选，最后通过 Q 网络计算输出 Q 值供参数进化迭代使用；

在正式训练开始前，设置  $T$  个时间步作为“热身阶段”；从  $T+1$  时间步开始，按照奖励最大化以及损失最小化的原则对模型进行训练，用于保证模型训练的稳定性，使得车辆节点对环境进行探索。

6、一种实施权利要求 1~5 任意一项所述缓解交通拥堵驾驶员控制方法的缓解交通拥堵驾驶员控制系统，其特征在于，所述缓解交通拥堵驾驶员控制系统，包括：

环境模型构建模块，用于根据车辆的空间位置和相对关系，将环境模型分解为局部和全局两层，并将环境的建模定义为信息拓扑结构；

强化学习模型构建模块，用于将节点节点特性  $x_i$  输入到全连接 FCN 层，FCN 的输出与关联矩阵  $A_i$  同时输入图神经网络 GCN 层进行并行计算，将输出结果与索引矩阵  $M_i$  进行点积运算进行车辆节点的筛选，最后通过 Q 网络计算输出 Q 值供参数进化迭代使用；

交通拥堵缓解模块，用于采用集中学习但分散执行的设定，每个目标车辆节点在每个时刻均做一个决策，为所有节点实现相同的给定目标，即有序通过从而解决交通拥堵问题；

决策信息建立下发模块，用于将节点间的通信和信息传播采用图神经网络 GNN 进行建模，决策处理器采用 Deep Q learning，形成的决策信息以建议指令的形式下发给每一位环境内的驾驶员。

7、一种计算机设备，其特征在于，所述计算机设备包括存储器和处理器，所述存储器存储有计算机程序，所述计算机程序被所述处理器执行时，使得所述处理器执行如下步骤：采用集中学习但分散执行的设定，每个目标车辆节点在每个时刻均做一个决策，为所有节点实现相同的给定目标，即有序通过从而解决交通拥堵问题；节点间的通信和信息传播采用图神经网络 GNN 进行建模，

## 权 利 要 求 书

---

决策处理器采用 Deep Q learning, 形成的决策信息以建议指令的形式下发给每一位环境内的驾驶员。

8、一种计算机可读存储介质, 存储有计算机程序, 所述计算机程序被处理器执行时, 使得所述处理器执行如下步骤: 采用集中学习但分散执行的设定, 每个目标车辆节点在每个时刻均做一个决策, 为所有节点实现相同的给定目标, 即有序通过从而解决交通拥堵问题; 节点间的通信和信息传播采用图神经网络 GNN 进行建模, 决策处理器采用 Deep Q learning, 形成的决策信息以建议指令的形式下发给每一位环境内的驾驶员。

9、一种信息数据处理终端, 其特征在于, 所述信息数据处理终端用于实现如权利要求6所述缓解交通拥堵驾驶员控制系统。

10、一种应用如权利要求 1~5 任意一项所述缓解交通拥堵驾驶员控制方法的应用, 其特征在于, 所述应用包括: 将感知系统获得的局部信息以及全局信息进行初步处理可以获得符合网络输入要求的数据类型, 将此数据元组输入到训练成熟的网络获得全局的最优决策输出, 再将决策结果通过驾驶员建议系统下发到每个车辆节点的实际操作人, 由实际操作人完成最后的车辆行驶控制任务。