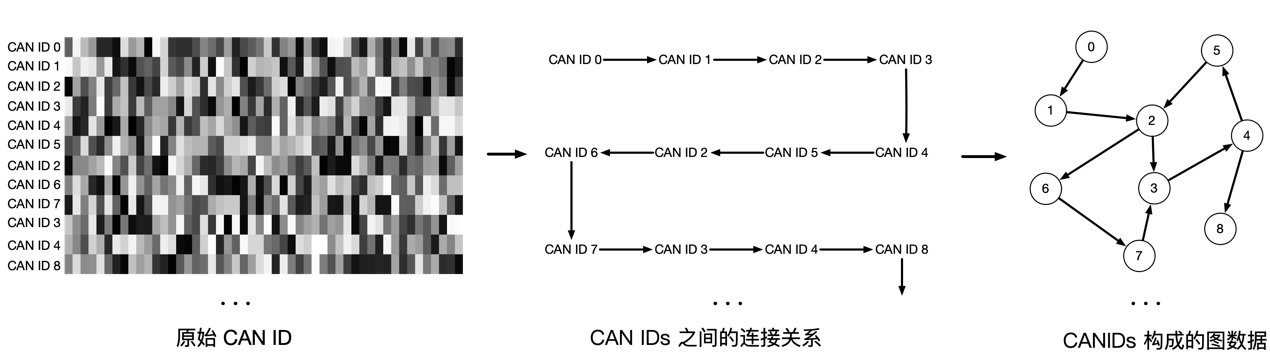
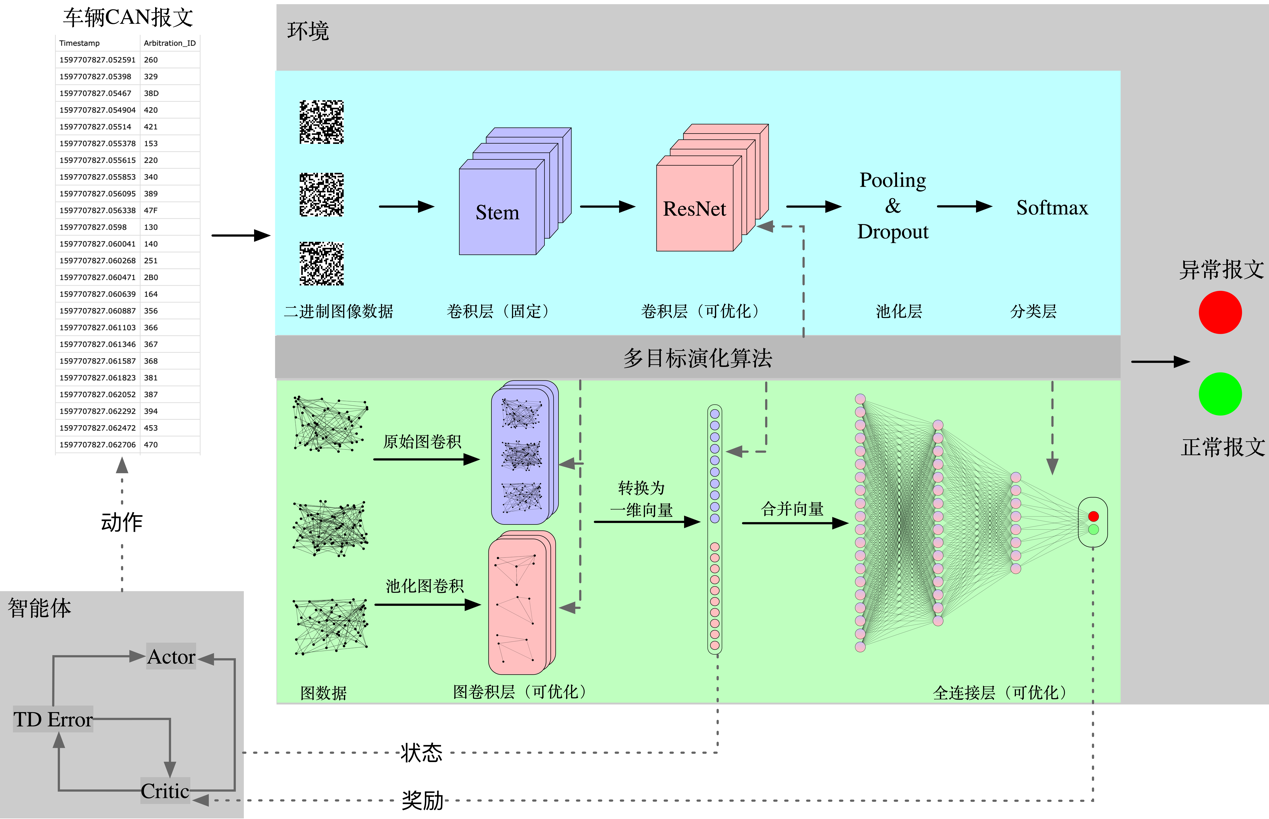


典型的消息注入攻击场景和CAN的格式。攻击者可以通过目标车辆的车载诊断端口或提供无线通信信道的远程通信设备，将伪造的消息注入目标车辆的can总线。



首先，CAN报文中前一时刻报文帧CANID指向下一帧CANID。CANID的种类有限，并非所有CANID都指向新的CANID，它可能指向以前出现过CANID。转换的图形数据的方向关系与消息帧序列的时间序列相关。不同的CANID对应车辆中的不同功能，可以通过图形神经网络学习车辆功能执行的逻辑是否合理。例如，制动时会出现较大的加速信号，或者在没有车钥匙信号的情况下会出现车辆点火信号等不合理的逻辑。学习一定帧数的CAN数据包中的抽象判断消息是否是入侵报文。



CANID检测样本的示意图在图的右上角。报文入侵检测系统的检测部分是强化学习的环境部分，在右侧最大的灰色框内，强化学习智能体部分是由多个全连接网络组成，在左下方的灰色框内。检测部分由CNN和GNN两种网络组成，位于上面蓝色背景的是CNN，位于下面绿色背景的是GNN。两种网络都使用MOEA算法优化，图中使用虚线由中间的深灰色框指向被优化的部分。优化的部分包括CNN的ResNet部分以及CNN中的超参数如学习率，Drop out率等等。优化的部分包括GNN的图卷积层数，全连接网络每层的神经元数还有其他的超参数，学习率，Dorpout等。RL、CNN、GNN三个网络的设计细节见图。

我们提出的架构如图所示，可以分为检测样本采集部分和检测部分两个大部分。样本采集部分强化学习网络根据图神经网络学习到的车辆状态来动态的获取一定长度的CAN报文数据。图网络和卷积网络的网络架构都使用了遗传算法进行了优化。遗传算法优化了卷积网络的卷积层，图神经网络的图卷积层以及全连接层，这些部分在图中用虚线标出。模型同时利用图神经网络提取逻辑特征的优势，卷积神经网络提取空间特征的优势，最后整合两个网络各自的输出结果得出模型的检测结果, GNN和CNN都是二分类网络，最后输出检测结果之前判断每个网络的输出一行两列向量，结果输出为两个网络输出向量中两个列数据相差较大的网络结果(即置信度更大的结果)。