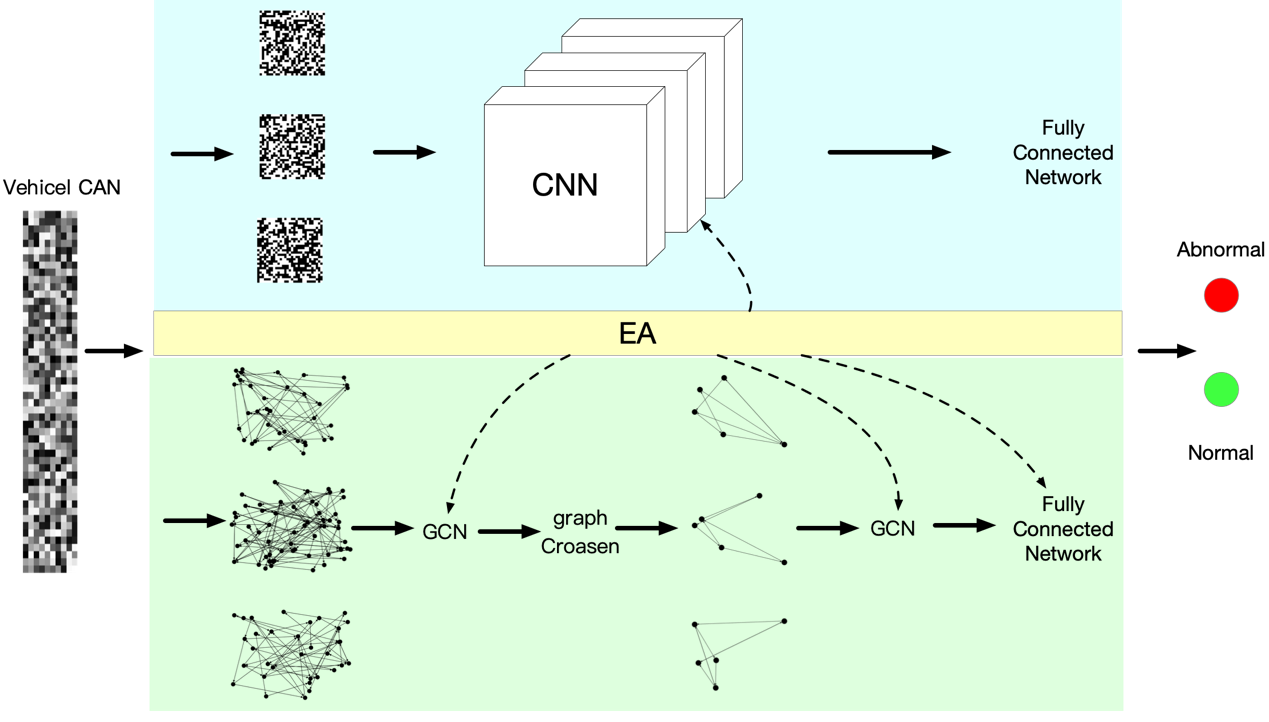
提出的方法



我们提出的架构如图所示，可以分为三大部分，检测样本采集部分，图网络部分和卷积网络部分。样本采集部分由强化学习网络负责，负责动态的获取一定长度的CAN报文数据送入检测神经网络。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。被遗传算法优化的部分在图中用虚线标了出来，它们分别是卷积网络的卷积层，图神经网络的图卷积层以及全连接层。同时利用两种不同神经网络的优势，图神经网络具有提取逻辑特征的优势，卷积神经网络具有提取空间特征的优势，最后整合两个网络各自的输出结果，得到最后的输出，具体实行见下文。

1. 数据采集部分

在每次入侵检测之前先动态采集一定长度的CAN报文数据作为一个输入到检测网络的样本。强化学习的任务是获得适当的每次检测需要采集的CAN报文长度。车辆CAN报文中，不同功能的报文数据，或者是同一功能的报文也可能需要不同的帧数去完成，所以使用强化学习网络解决动态采集的问题。在训练强化学习网络之前首先先训练出一个具有入侵检测能力的图神经网络，把图神经网络的卷积层的输出转换成一维向量。我们的工作中有两个卷积层，所以得到两个一维向量，把得到的两个一维向量合并作为强化学习的状态输入。得到图网络的识别结果设计奖励函数，如果图网络识别正确获得正向奖励，如果识别错误获得负值的惩罚。强化学习的输出是一个具有多维的向量，维度如下所示。

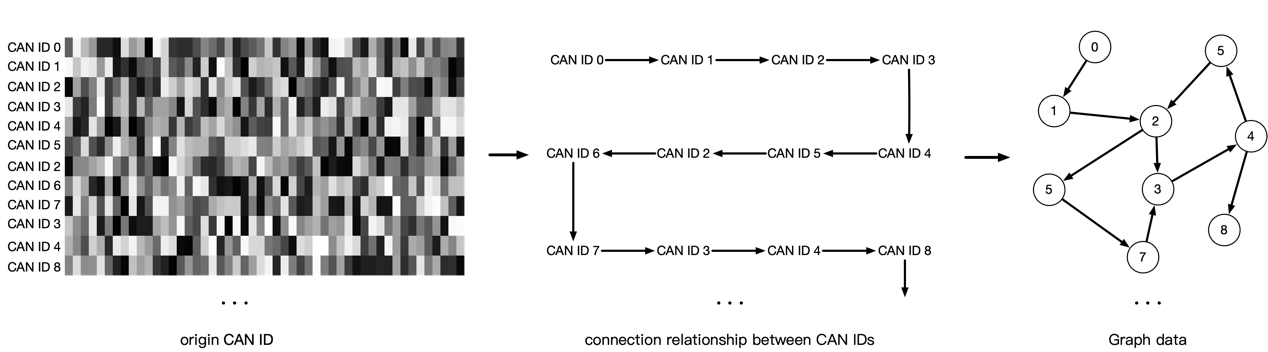
入侵检测的最大报文长度与最小报文长度之差。本工作的强化学习网络架构设计参考TD3。

2. 图网络部分

1.1构建图数据

Our work is based on a publicly available datasets which is ATTACK & DEFENSE CHALLENGE 2020 DATASET [27] which is issues on IEEE-DATAPORT. the dataset contains two states of vehicle driving and stationary, and each part has two types of dataset files: intrusion and normal. In the intrusion file, several intrusion CAN messages are interspersed with normal messages. Directly convert these labeled messages into a format that can be input to the network for training and testing. Our task is to quickly and accurately identify the intrusion CAN messages from the dataset. Each data file has millions of CAN data frames, which can better extract various features in the message.

First, we set the previous frame of message to point to the next frame of message. Because the message ID is limited, not all messages will point to a new message, it may point to a message frame that has appeared before. The direction relationship of the converted graph data is related to the time sequence of the message frame sequence. A sequence of CAN IDs always indicates different functions in the vehicle, and whether the logic of the execution sequence of the functions is reasonable or not can be learned by the graph neural network. For example, a large acceleration signal appears when braking, or a vehicle ignition signal appears without a car key signal. These are unreasonable logic. It is very likely that an intrusion message has appeared on the CAN bus, which has practical significance. For the graph neural network (GNN) to learn more abstract features, we input a partial graph of sequence data to determine whether the message is an intrusion message, instead of inputting one or two frames of messages. Our proposed approach is shown in Fig. 3.



1.2 整体图网络架构

本工作的图神经网络是图层面的分类任务，图分类任务的网络架构参考工作GraphCoarsen。 是将构建好的图数据使用聚类的方法分成若干个簇，每个簇可以看作是图数据的子图，得到一系列子图的特征矩阵，使用特征向量构建池化矩阵，把每个子图池化为一个超级节点。如公式所示：

给定k个子图连接的图G，**c**是图G的一部分。表示子图的节点数量，是子图的节点列表。每个子图可以看作是图G的超级节点。定义采样算子如下：

当且仅当

表示矩阵中第i行第j列的元素，表示节点列表中第j个节点。该操作表明节点在子图和在原始图G的对应关系。

因为傅立叶变换可以将图信号转换到频域，可以将信号信息和图数据的结构信息都考虑在内，所以我们参考了【图坍缩】使用傅立叶转换设计池化操作。将构建的图信号G池化为

。池化操作基于了每个子图的傅立叶变换。子图的拉普拉斯矩阵为。,…表示第k个子图的拉普拉斯矩阵的特征向量。然后使用上采样操作上采样特征向量到整个图G。

集合所有上采样的特征向量构成一个矩阵，表示包含所有子图的第l个特征向量的池化矩阵。

每个子图不一定有相同的节点数，也就是说每个子图的特征向量的个数是不一定相等的。表示所有子图中最大的节点个数。对于拥有个节点的子图的第l次池化操作表示为：

表示第l次池化操作的结果，的第k行包含的第k个子图的信息，即第k个超级节点。

基于以上结构我们可以构造一系列次的池化操作，结合所有池化操作的结果构成矩阵。

在一些其他采用池化方法的图分类任务是使用求和或者求平均的方法，无差别对待子图中的每个节点，无法提取每个子图中节点的结构信息。我们的工作对图数据进行了一次池化，池化前和池化后都做了一次图卷积操作提取特征。最后把两次图卷积的输出都通过求和或者求平均的方法化为一维向量，把两个一维向量拼接成一个输入到全连接层做分类。每次图卷积的输出化成一维向量的方法根据遗传算法搜索得到。

1.3 使用遗传算法优化图网络

如何找到识别准确并且需要的运算量小是我们设计的遗传算法想要达到的。

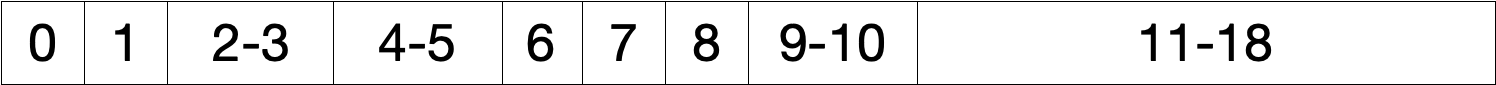
## 讲解图网络部分的架构，并使用公式表达出来

基因第一部分，图数据可以根据相邻两帧的顺序关系确定有向图的边的方向，具体说是上一帧报文指向下一帧。是否使用方向信息使用遗传算法搜索。基因第二部分，在进行图坍缩过程中，根据上一节的公式可知，需要得到每个子图的拉普拉斯矩阵计算每个子图的特征向量。在求每个子图的特征向量之前是否对拉普拉斯矩阵使用正则化。图网络正则化的拉普拉斯矩阵和非正则化的L的公式表示为：

其中A表示图数据的邻接矩阵，D表示图数据的度矩阵。

基因第三部分和第四部分分别是两个图卷积块的深度和预测层中每一层的神经元的数目，网络结构的设计参考slim，合适的图卷积的层数和预测层的神经元个数可以在复杂度不高的的同时达到效果很好的检测准确性。基因的第6位到最后都是网络结构中的超参数，使用遗传算法进行优化。其中第9位到第10位两位分别表示每个卷积块输出卷积结果之后如何降维成一维向量。其中有对每个维度求和，取平均，取最大三种选择。

在图网络中的结果设计比如是否使用有向图的信息，构造池化矩阵时是否正则化，每次图卷积的层数，预测层的层数以及预测层中每层的神经元数等等。还有超参数我们也进行了搜索，比如Dropout rate，学习率。具体的基因和相应的对应关系如下图所示。



|  |  |  |
| --- | --- | --- |
| Gene Position | meaning | Search space |
| 0 | directed/undirected graph | 0: directed 1: undirected |
| 1 | Whether Normalization of subgraph | 0: not normalization  1: normalization |
| 2-3 | Layers of GCN | 0: one 1: two 2: three |
| 4-5 | Proportion of neurons in prediction layer | 0: 0.25 1: 0.75 2: 1.00 |
| 6 | DropOut | 0: 0.05 1: 0.1 2: 0.2 3: 0.3  4: 0.4 5: 0.5 |
| 7 | Weight Decay Rate | 0: 5e-4 1: 8e-4 2: 1e-3 3: 4e-3 |
| 8 | Learning Rate | 0: 5e-4 1: 1e-3 2: 5e-3 3: 1e-2 |
| 9-10 | way to merge the vectors of every GCN layer | 1: sum 2: average 3: max |
| 11-18 | Activation function | 0: sigmoid 1: tanh 2: relu  3: leaky\_relu 4: relu6 |

虽然使用构建图数据的方向，会更多的利用图数据的信息，实验结果表明使用更多的图信息并不一定可以提高识别精度。具体实验结果分析见下一节。

2. 卷积部分

2.1 数据获取

数据获取参考卷积网络识别CAN入侵报文论文，首先把十六进制CAN IDs 转化为二进制，构成与图像相似的样本。每个像素点分别是0或者1，参考论文，车辆CAN扩展帧的ID位数是29位，采集29帧构成2929的样本。

2.2 卷积网络架构

搜索空间对于架构设计质量和搜索效率都有很大的影响，在本节介绍搜索空间的基本设计思路和属性。

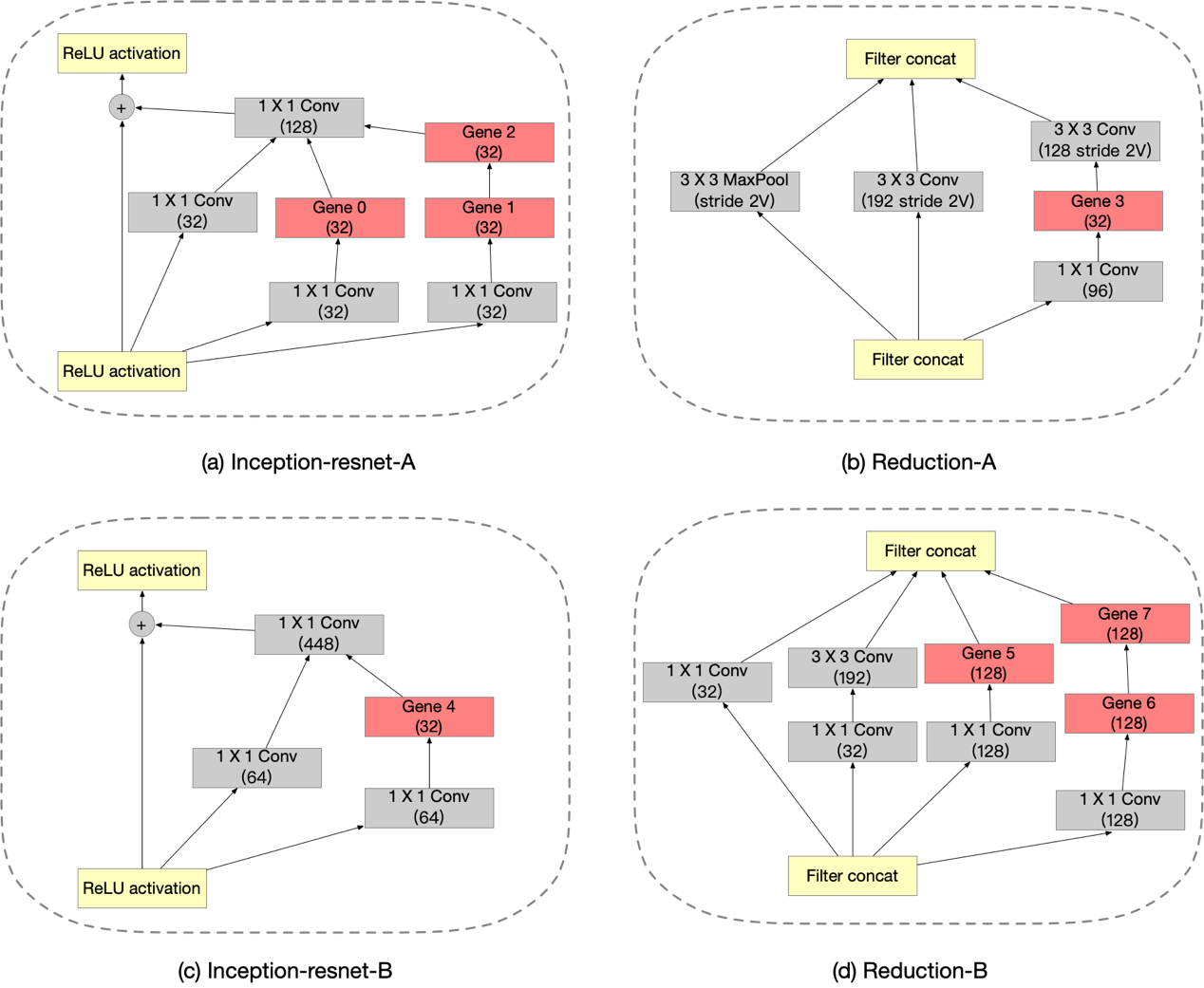
深度卷积网络例如Inception和Resnet的变种都是通过堆叠多个块来设计构建神经网络。网络结构设计包括深度（层数）、宽度（通道数）和空间分辨率变化（池层数）的确定，而块结构设计则规定了分层连接和局部计算。通过这种分块设计方法，生成的模型不仅可以实现高性能，而且可以推广到不同的数据集和任务。因此，我们遵循与[8,31,59]中相同的块级设计方法，这将在下一小节中详细阐述。

块是一个小的卷积网络，为了在前向传播中更有效地处理不同的中间信息，根据特征映射的不同网格大小设计了四种种卷积块。同时，reduction块的设计目的是增加深层的感受野，并通过以2的步幅应用所有操作，将特征图的网格大小减半。按照现代CNN架构[5,60,61]的惯例，当特征图的网格大小减半时，我们将块的通道（过滤器）数量增加一倍，以保持大致恒定的隐藏状态维度。

O是一个操作空间，由一组可能的网络架构基本组件、由人类专家设计的已知成功模块组成。本研究中使用的5种O操作和相应的基因型-表型映射如表一所示。在表中，空间可分离卷积(SP)和深度可分离卷积（DW）能够在不牺牲网络性能的情况下减少网络参数。这里我们使用两个DW操作和两个SP操作，内核大小分别为3×3和5×5，简称SP3、DW3和SP5、DW5。

|  |  |  |  |
| --- | --- | --- | --- |
| Operation type | Kernel size | Short name | Code |
| Spatial Separable Convolutions | 3 | SP3 | 0 |
| Spatial SeparableConvolutions | 5 | SP5 | 1 |
| Depthwise separable convolution | 3 | DW3 | 2 |
| Depthwise separable convolution | 5 | DW5 | 3 |
| Normal convolution | 3 | 3\*3 | 4 |

我们定义遗传算法的交叉操作是在两个亲代染色体上分别随机选择一个点断开，相互交换染色体基因。定义遗传算法的变异操作是在一个染色体上随机选取两个基因，然后进行交换。因为每次都重新训练生成的个体，所需要的计算量是非常巨大的。为了提高遗传算法的评估效率，在交叉操作和变异操作的过程中，把相应神经网络位置的权重参数也进行交换。通过四个不同的卷积块的示意图可以得知，每个卷积块的基因的通道数都是不同的，每个卷积块内的基因的通道数都是相同的，这对于交叉操作是没有影响的，但是会影响变异操作。在一个染色体的两个基因交换位置，并且需要交换神经网络的相应的权重，所以两个基因的通道数需要相同。redA和resB卷积块都只有一个基因，resA和redB卷积块各自分别有三个基因。所以变异操作只能在resA块或者redB块内进行。



Inception-ResNet是深度卷积模型的一种，为了图像分类领域把图片分为1000个类而设计并且表现出了非常出色的性能。神经网络的整体架构如图所示。输入的尺寸是29X29X1经过Stem模块把输入数据尺寸转换为13\*13\*28。经过我们设计的可优化的四个模块后得到的数据尺寸为2\*2\*896。最终经过Softmax模块把数据转化为维度为2的二分类。

