# Evaluation Of Different Methods

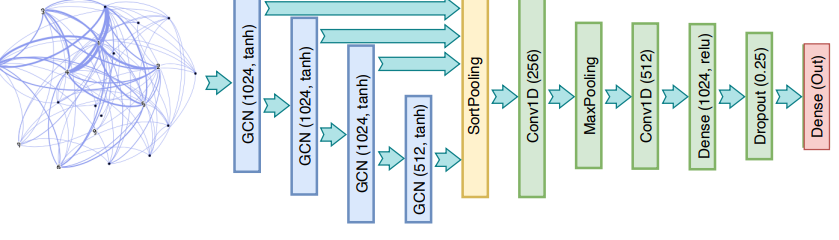
Graph building.  
Dataset.   
The model being used.

**MAppGraph: Mobile-App Classification on Encrypted Network Traffic using Deep Graph Convolution Neural Networks –**

A picture containing text, clock, watch, gauge

Description automatically generatedGraph building:  
A node is defined as (dst ip, dst port). And an edge is defined as the cross correlation between two nodes, and the weight is the cross correlation.  
 is the activation of node i in slice t of time.  
After wards they normalize the edge weights to [0,1].   
Each node has a feature vector that is constructed from different extracted features.   
(The more slices that we are going to take the less edges that there are going to be)

Dataset:   
The team collected their data since they claimed each dataset they found the time delta collection was too short. Using 10 students and 8 smartphones during 6 months. The data collection is happening at the same time and is mirrored to a different computer with enough storage, they can differ between apps using the IP source of the device. Each session is 3-4 hours.

  
The model being used:  
The used model is of 3 layers of convolutional layers, then a pooling layer, after that two convolutional layers with max pooling layer between them. Then flattening and sending to a fully connected layer, and after that sending to dropout with 25% to avoid overfitting.

**CGNN: Traffic Classification with Graph Neural Network-**

Graph building:  
The graph is a chained graph; every node is a packet and edges are the adjacency between these packets. Also, every node has a feature vector of 1500 by bytes of the corresponding packet. The edges in the graph are undirected.

Chart

Description automatically generated  
  
  
  
  
Dataset:  
They defined a time window T in which they sniffed the packet of some application. And then they combined them to pcap objects. On these pcap objects they did some processing to turn them into graphs.

The model being used:  
The model of the network is two layers of SGC to decrease the vectors sizes in half.

**GAP-WF: Graph Attention Pooling Network for Fine-grained SSL/TLS Website Fingerprinting-**

Text, letter

Description automatically generatedGraph building:  
The graph is a traffic trace graph. And each node is a vector that represent a flow through a five tuple, they also added time\_inter and packet size with +/-, as + is outgoing and – ingoing. An edge between two nodes i and j is constructed if and only if

Dataset:   
They collected their own data, since they couldn’t find any real other dataset to satisfy them.  
The data was collected using two computers one has windows the other is ubuntu, both have fire-fox and chrome and they used selenium module to simulate of people surfing the net. They sniffed the traffic using. It’s mentioning that they do 3 or 4 different scenarios.

Diagram

Description automatically generatedThe model being used:  
There are 3 layers for the actual model that’s learning,   
Multil-head graph attention

**A Network Traffic Classification Method Based on Graph Convolution and LSTM-**

Text

Description automatically generatedGraph building:  
The nodes are vector features of flows with d-dimensions. And the edges between them are Established like this:  
  
  
  
  
  
The nodes features might go through bit of normalization.

Dataset:  
they used UNSW-NB15.   
Link to the data <https://cloudstor.aarnet.edu.au/plus/index.php/s/2DhnLGDdEECo4ys>.

The model:  
the model is 3 layers of and then lstm. Sgc is for dimension reduction and lstm is for normal abnormal classification.

# Comparison

|  |  |  |
| --- | --- | --- |
|  | Pros | Cons |
| MAppGraph: Mobile-App Classification on Encrypted Network Traffic using Deep Graph Convolution Neural Networks | * Graph building is straight forward * Rich topology. | * complexity of memory and calculations are not efficient * A lot of time to collect the dataset. Each session is 3-4 hours. |
| CGNN: Traffic Classification with Graph Neural Network | * Keep the semantic and the casual relationship of the packet * Build time is O(n) | * The DL model is only two layers * The node features are the 1500 bytes of the packet, which is less efficient for the model. |
| GAP-WF: Graph Attention Pooling Network for Fine-grained SSL/TLS Website Fingerprinting | * Can find associations between multiple services which occurred on the same time. | * Depending on time features which can cause overfitting. * The complexity of building the graph is high * Graph node features are not simple and can be big and vary. * The proposed method is only for website fingerprinting |
| A Network Traffic Classification Method Based on Graph Convolution and LSTM | * The usage of the sgc layers can help extract better features for the flows. | * In comparison to other papers their results are not so high. * The graph topology is not well defined. S |