FlowPic: Encrypted Internet Traffic Classification is as Easy as Image Recognition

School of Electrical Engineering

TAL SHAPIRA & YUVAL SHAVITT
Why Classification?

- Traffic Classification
  - Traffic engineering
  - Law enforcement
  - Decryption for MITM attacks
- Most of today's traffic is encrypted
  - Looking at the payload does not help
  - Port numbers are no longer indicative
Previous Approaches

- Port based methods
- Handcrafted features extraction
  - Identify many ‘important’ features
  - Feed a feature vector to some tool
  - Usually apply 'classical' supervised learning techniques
- Cross correlation between payload size distributions (PSDs)
- Payload based traffic classification methods (DPI)
  - Apply neural networks over packet payload content
Our goal is to propose a generic approach for Internet traffic classification

- We use the exact same architecture for all the experiments
- Dealing with all kinds of Internet traffic classification problems

Using all time and size related information available in a network flow, instead of just using information from manually extracted features

Dealing with only a time window of a unidirectional flow instead of the entire bidirectional session

Do not rely on the packet payload content

- Do not breach privacy
- Minimalizing storage requirement
- Classifying VPN and Tor traffic
- Capturing an intrinsic characteristic of a category behavior, regardless of the encryption technique and a specific user
Our Approach

FlowPic: Encrypted Internet Traffic Classification is as Easy as Image Recognition

Class is Video

195.154.82.180_443_10.0.2.15_56113_TCP: [(40, 0.0), (178, 0.0011), (512, 0.1427), (1012,0.1472), ...]
FlowPic Construction

- Extract records from each flow, which comprised of a list of pairs, \{IP packet size, time of arrival\} for each packet in the flow
- Split each unidirectional flow to equal blocks (15/60 seconds)
- Generate 2D-histogram. For simplicity, we set the 2D-histogram to be a square image
  - We disregard all packets with size greater than the MTU (1500)
  - For the X-axis, first we normalize all time of arrival values to be between 0 and 1500

```
195.154.82.180_443_10.0.2.15_56113_TCP:
[(40, 0.0), (178, 0.0011), (512, 0.1427), (1012, 0.1472), ...]
```
FlowPic: Encrypted Internet Traffic Classification is as Easy as Image Recognition
FlowPic: Encrypted Internet Traffic Classification is as Easy as Image Recognition
Our CNN Architecture

- A LeNet-5\(^*\) style architecture
  - \#params=309,094 + 65m

- Dropout

- Softmax final layer \[ \sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \text{ for } j = 1, \ldots, K. \]

*\[\text{Lecun et. al 1998}] \text{ “Gradient-based learning applied to document recognition”}\]
Specifications

- Categorical cross entropy loss function
- Using the Adam gradient-based optimizer with default hyper-parameters
- We build and run our networks using Keras and Tensorflow frameworks
- We set our batch size to 128, and run our network for 40 epochs of 10 batches each
Dataset

- Our dataset consists of the two datasets from Uni. of New Brunswick (UNB):
  - "ISCX VPN-nonVPN traffic dataset" (ISCX-VPN)*
  - "ISCX Tor-nonTor dataset" (ISCX-Tor) **
- In addition we generate our own small packet capture

** [Lashkari et al. 2017] “Characterization of Tor Traffic using Time based Features”
### Dataset

<table>
<thead>
<tr>
<th></th>
<th>Non-VPN</th>
<th>VPN</th>
<th>Tor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VoIP</strong></td>
<td>Google Hangouts, Facebook, VoipBuster and Skype</td>
<td>Google Hangouts, VoipBuster and Skype</td>
<td>Google Hangouts, Facebook and Skype</td>
</tr>
<tr>
<td><strong>Video</strong></td>
<td>Google Hangouts, Facebook, Netflix, Vimeo, YouTube and Skype</td>
<td>Netflix, Vimeo and YouTube</td>
<td>Vimeo and YouTube</td>
</tr>
<tr>
<td><strong>File Transfer</strong></td>
<td>FTPS, SCP, SFTP and Skype</td>
<td>FTPS, SFTP and Skype</td>
<td>FTP, SFTP and Skype</td>
</tr>
<tr>
<td><strong>Chat</strong></td>
<td>Google Hangouts, Facebook, AIM Chat, Skype, ICQ and WhatsApp Web</td>
<td>Google Hangouts, Facebook, AIM Chat, Skype and ICQ</td>
<td>Google Hangouts, Facebook, AIM Chat, Skype and ICQ</td>
</tr>
<tr>
<td><strong>Browsing</strong></td>
<td>Firefox and Chrome</td>
<td>-</td>
<td>Firefox and Chrome</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Non-VPN</th>
<th>VPN</th>
<th>Tor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VoIP</strong></td>
<td>3304</td>
<td>2872</td>
<td>2978</td>
</tr>
<tr>
<td><strong>Video</strong></td>
<td>1553</td>
<td>302</td>
<td>754</td>
</tr>
<tr>
<td><strong>File Transfer</strong></td>
<td>1174</td>
<td>242</td>
<td>1126</td>
</tr>
<tr>
<td><strong>Chat</strong></td>
<td>635</td>
<td>1061</td>
<td>422</td>
</tr>
<tr>
<td><strong>Browsing</strong></td>
<td>3191</td>
<td>-</td>
<td>2026</td>
</tr>
</tbody>
</table>
Experiments

- **Multiclass classification experiments:**
  - **Traffic categories:** VoIP, Video, Chat, File Transfer, Browsing over 4 datasets: non-VPN, VPN, Tor and merged dataset
  - **Encryption techniques:** non-VPN, VPN, Tor
  - **Application identification:** 10 classes of VoIP and Video applications

- **Class vs. all classification experiments:**
  - 5 classes: VoIP, Video, Chat, File Transfer, Browsing
  - 4 types of datasets: non-VPN, VPN, Tor and merged

- Classification of an Unknown Application
Summary of Traffic categorization Results

Former results:
- **The UNB group** gained a best average precision of:
  - 84.0% for non-VPN traffic categorization
  - 89.0% for VPN traffic categorization
  - 84.3% for Tor traffic categorization
- **Wang et al.** achieved a best accuracy of:
  - 83.0% for non-VPN traffic categorization
  - 98.6% for VPN traffic categorization

Furthermore, we test the network that was trained over the merged set over non-VPN, VPN and Tor set experiment achieves an accuracy of 88.2%, 98.4% and 67.8%, respectively.

<table>
<thead>
<tr>
<th>Multiclass Dataset</th>
<th>Top-1 Acc. (%)</th>
<th>Top-2 Acc. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Categories over Non-VPN</td>
<td>85.0</td>
<td>99.4</td>
</tr>
<tr>
<td>Traffic Categories over VPN</td>
<td>98.4</td>
<td>100.0</td>
</tr>
<tr>
<td>Traffic Categories over Tor</td>
<td>67.8</td>
<td>81.0</td>
</tr>
<tr>
<td>Traffic Categories over Merged Dataset</td>
<td>83.0</td>
<td>95.1</td>
</tr>
</tbody>
</table>
We get an average accuracy of:
- **97.0%** for non-VPN
- **99.7%** for VPN
- **85.7%** for Tor

There are no former results.
Classification of an Unknown Application

- Video vs. all dataset, where we exclude all Facebook’s videos
  - We achieve an accuracy of **99.9%**

- VoIP vs. all dataset, where we exclude all Facebook’s VoIPs
  - We achieve an accuracy of **96.3%**

- There are no former results
Our network achieves an accuracy of 99.7% classifying 10 VoIP and Video applications.

Former Results:

- Yamansavascilar et al. constructed 111 flow features and used k-NN algorithm to achieve an overall accuracy of 93.9% classifying 14 classes of applications.
- For example, they achieved an accuracy of 45.1%, 80.10% and 86.30% classifying Skype-VoIP, Vimeo and YouTube, respectively.
Future Research

- Optimize our CNN architecture or examine other well known architectures

- Reduce run-time and memory consumption by:
  - Using binary images
  - Reducing FlowPic size
  - Using blocks with a shorter time duration
  - Relying only on a sub-sampling of the flow
A generic approach for Internet traffic classification

- We use the exact same CNN architecture for all the experiments
- Dealing with all kinds of Internet traffic classification problems; traffic categorization, application identification, encryption techniques

We are the first to show:

- Identifying traffic category of an unfamiliar application, by learning samples of other applications of the same traffic category
- Classifying different Internet traffic categories that pass through different encryption techniques by learning traffic that pass through other encryption techniques

Fast & simple: Classification is done after the first 15 seconds of a unidirectional flow

- No need for manual extraction of features
- Good results for traffic over VPN and Tor traffic
Thank you for your attention.

Questions?