

## Analysis for problem 3: Entropy for Network Anomaly Detection

### ATTACK DETAILS:

The first attack occurs at 08:12:16 am. It seems to be a DDoS attack where the attacker(s) are using several machines to bombard the same machine in the network with many communication requests so that it can't respond to other requests. In this case the destination IP is 172.016.114.050 . As a result communications seem to be down for the next 35s.

The second attack occurs around 11:55:15 am. It seems like a brute force telnet attack where the attacker is trying to gain access to the machine using telnet requests. The attacker also seems to be using IP spoofing to protect his identity. In this case the target machine seems to be 172.016.112.050 .

### ENTROPY ANALYSIS:

We initially tried to analyze the entropy of the network at a point by maintaining a count of all events throughout the day. This however resulted in the counts being saturated so that the 2<sup>nd</sup> attack remained undetected in the entropy counts for all columns and pairs of columns.

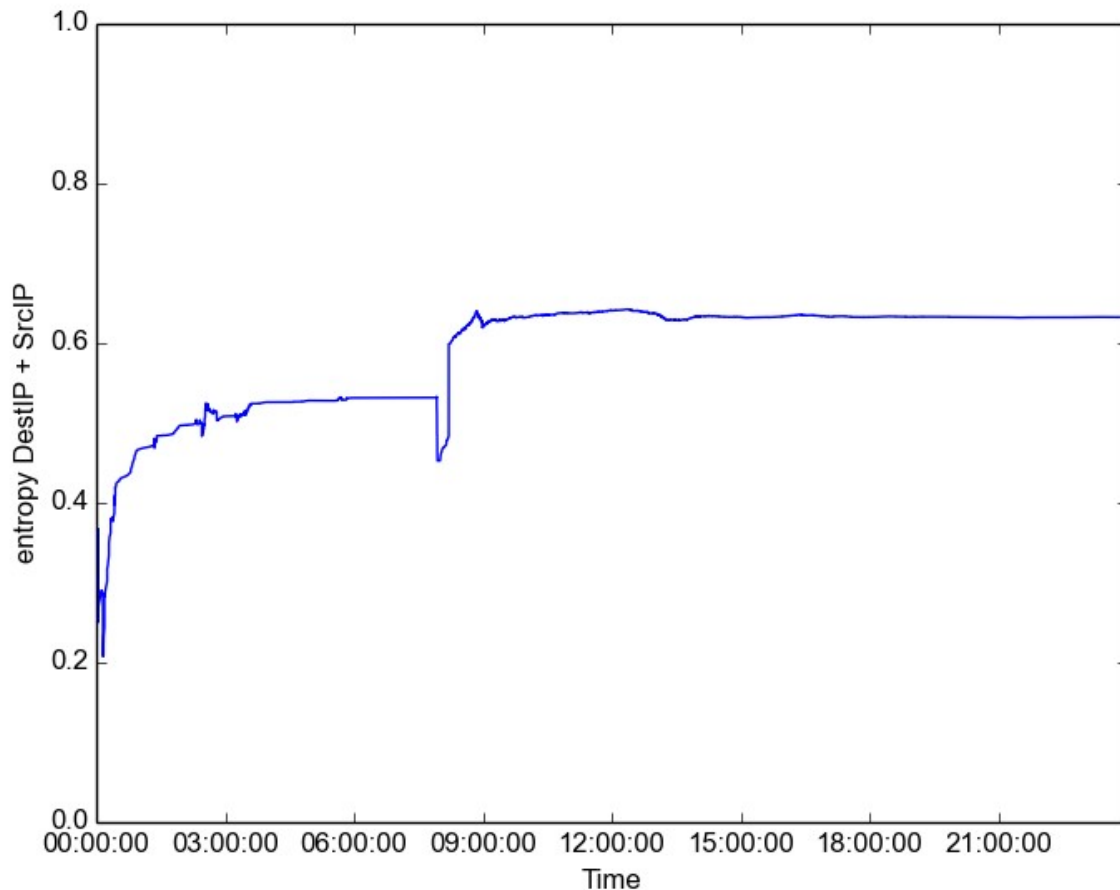
As a result we decided to use a sliding window entropy evaluation where we evaluated the entropy of the system over a window of the past 2000 requests after every 200 new requests. This resulted in more noisy entropy plots, but it was easy to spot both the attacks for certain pairs of columns.

We have included plots for some of the column pairs where the attacks can be inferred. Rest of the plots are included in the entropy\_plot folder. Files starting with w\_ are the windowed entropies while the other files are streaming entropies.

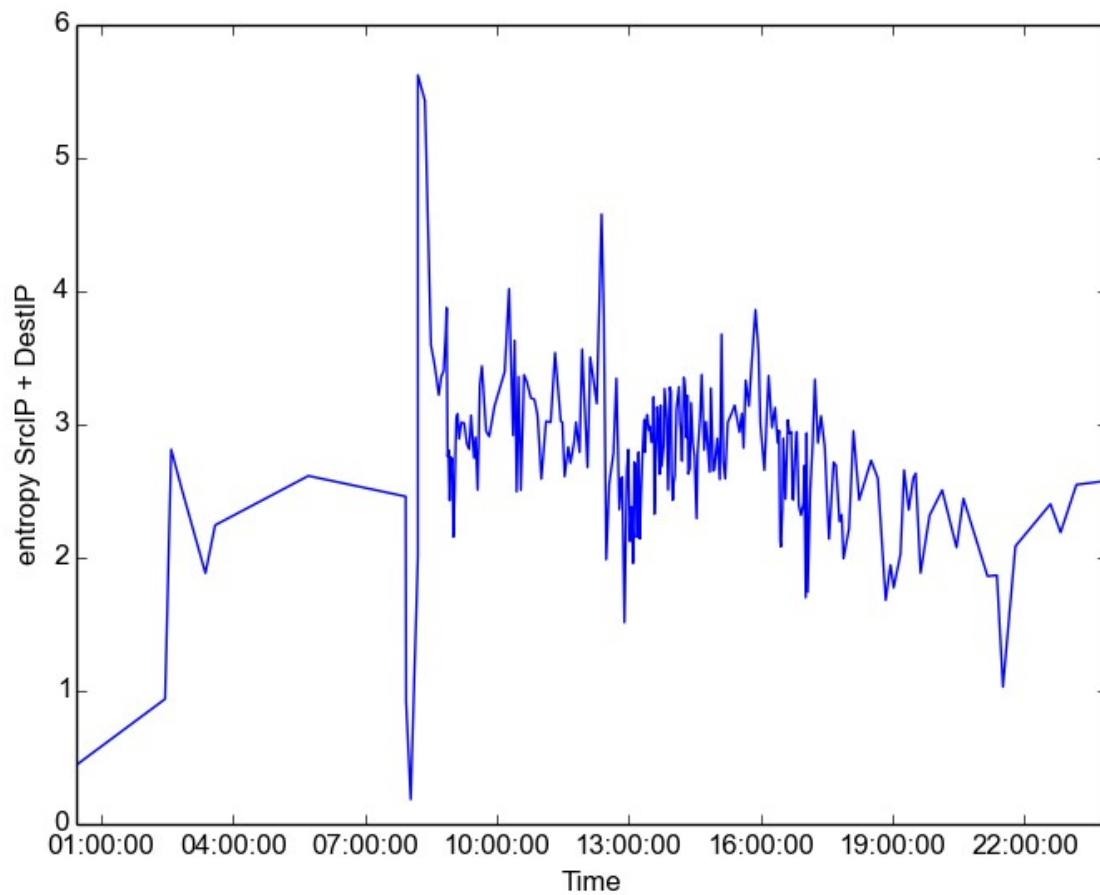
\*\*In the windowed entropy plots, events appear sometime after they have occurred since they are accounted for after the window they appear in.

### DST IP and SRC IP:

Streaming entropy

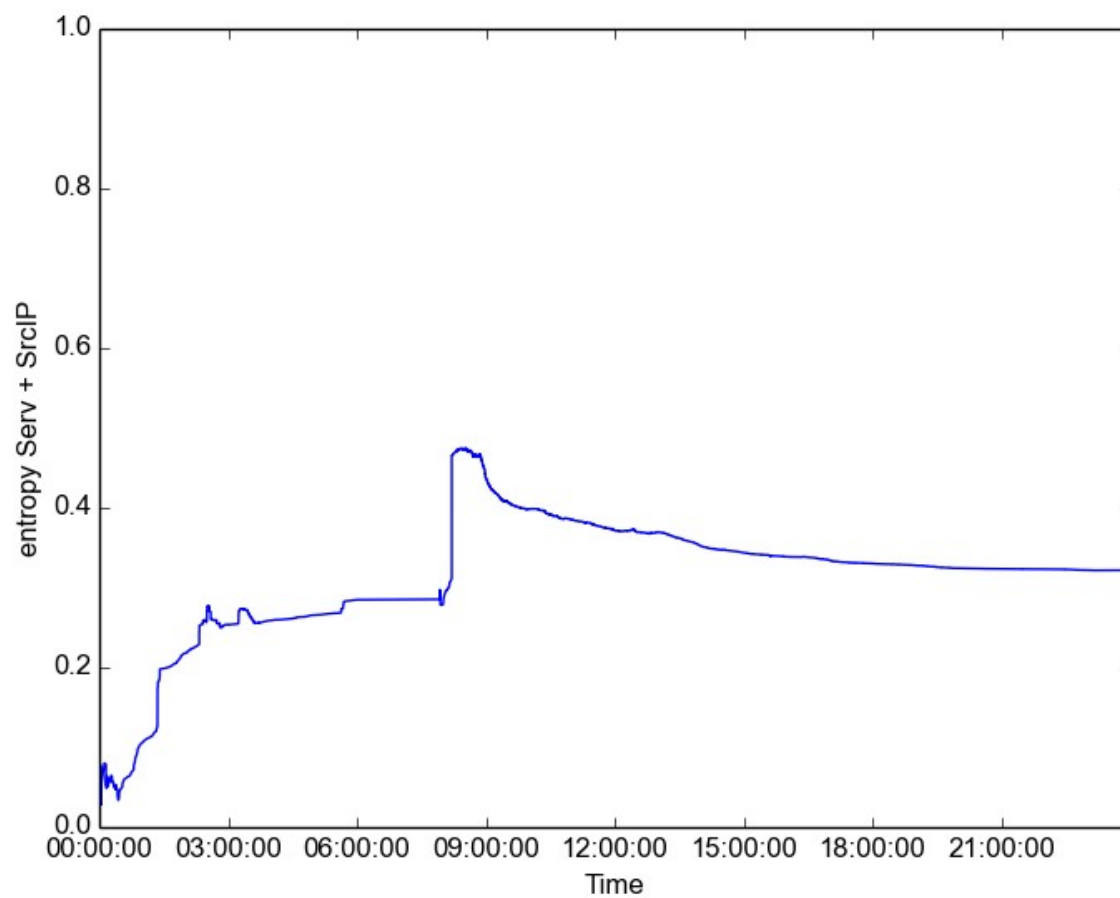


Windowed entropy

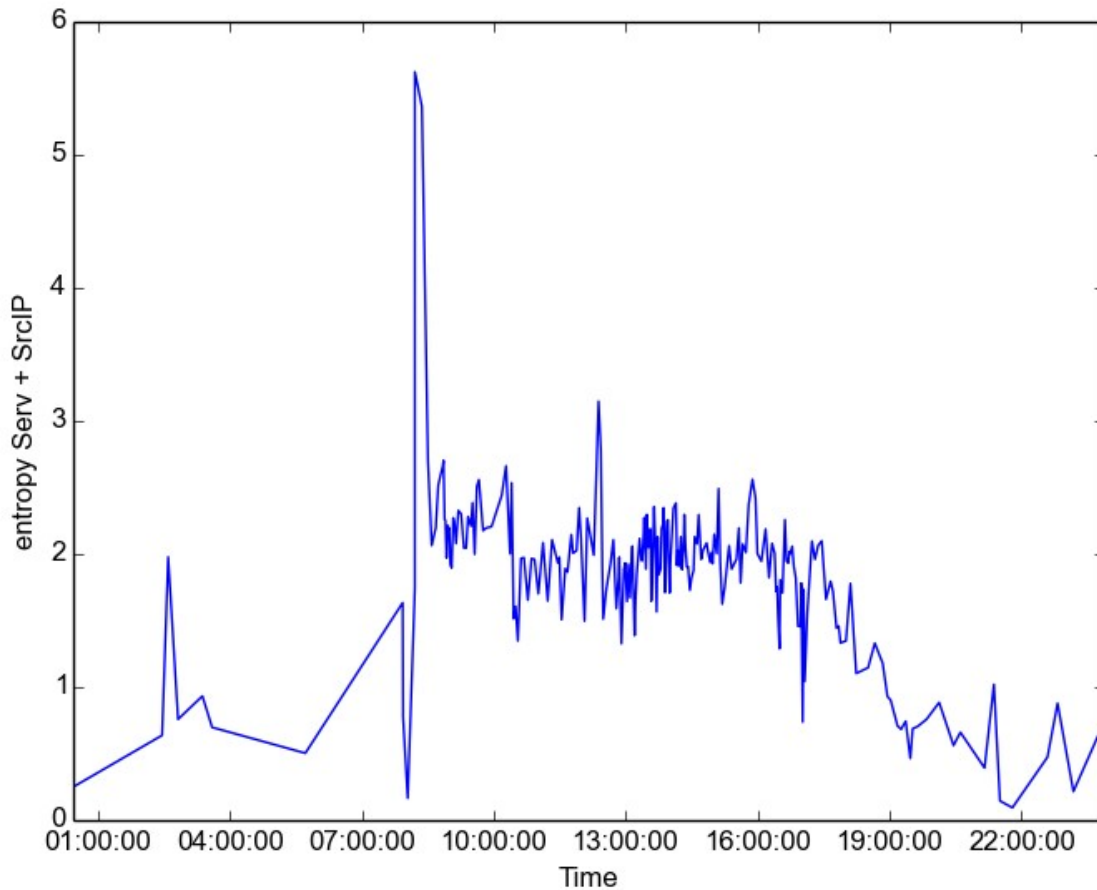


As we can see, in the streaming entropy plot, the 1<sup>st</sup> attack manifests in the form of a drop and then a steep increase in entropy. However, the second attack fails to manifest due to count saturation and the attack being smaller in duration. In the second plot, the 2 attacks clearly appear as the 2 largest maxima. However, since the entropy is calculated over a smaller period of time, it results in a noisy plot where a few other network anomalies are also seen, even though they might not be actual attacks.

**SERV and SRC IP:**  
Streaming entropy



## Windowed entropy



Similar patterns are observed in the plot of serv + src IP entropy vs time. For streaming entropy the 1<sup>st</sup> attack is clearly visible, but the second attack remains unnoticed. For windowed entropy we can see both the attacks as the 2 highest maxima but the plot is noisy and using automated methods for anomaly detection might result in false positives.

Most of the single column entropies (DestIP, Duration, Serv, Src Port, SrcIP) can be used to detect attack1 but don't detect attack 2. However, a few pairs of windowed entropies (Duration - SrcIP, Serv - SrcIP, SrcIP - DestIP) can be used to detect the second attack, although they are slightly noisy and might result in false detections. This could probably be removed by increasing the sliding window size to smoothen the estimates, but that might also result in the attacks being less visible from the entropies (the false positive - false negative trade-off).

**SUMMARY:**

From the results it seems like entropy could be useful in identifying network traffic anomalies. Using pairwise entropy seems like a better approach since it is more successful in identifying the attacks since the attacks leave signals over combinations of variables. However, pairwise entropy plots are pretty jagged even for normal network traffic and might result in misdetections.