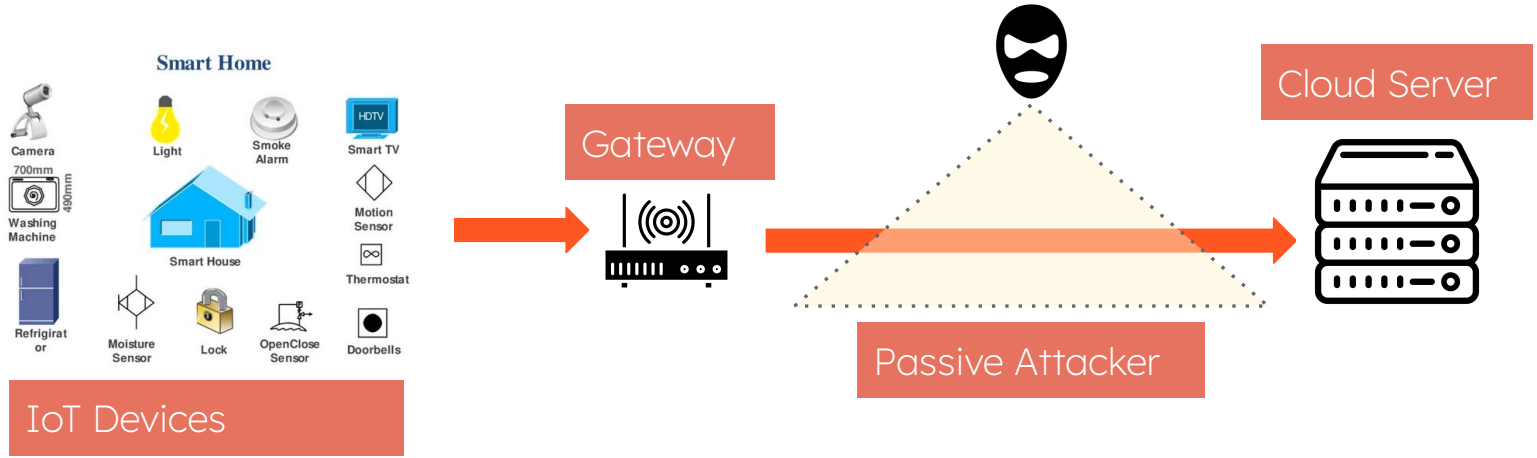
IoT Device Identification Against Shaped Traffic

Praneeth Damarla

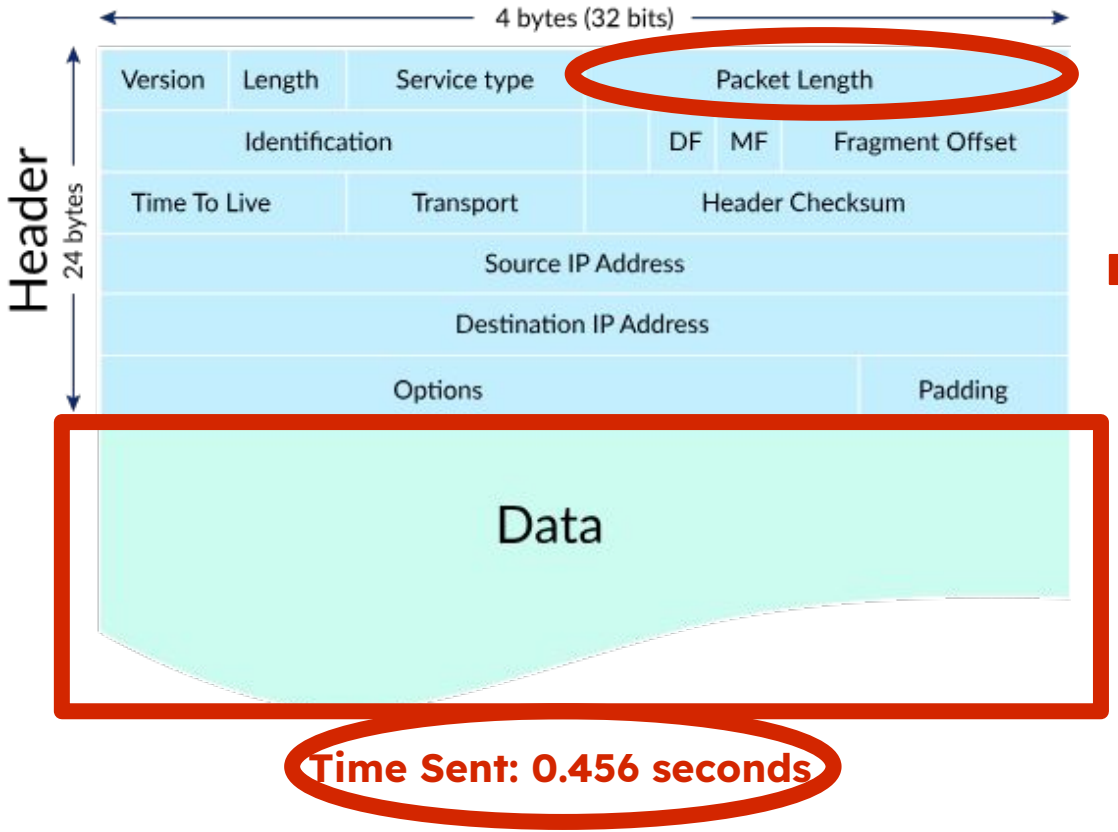
Advisor: Prof. Anand Sarwate

Motivation



DEVICE FINGERPRINTING	USER ACTIVITY
Amazon Echo	“Hey Alexa, what’s the weather like today?”
Canary Camera	“Person detected in living room”
Withings Sleep Sensor	“58 BPM, currently resting”

Network Packet Information

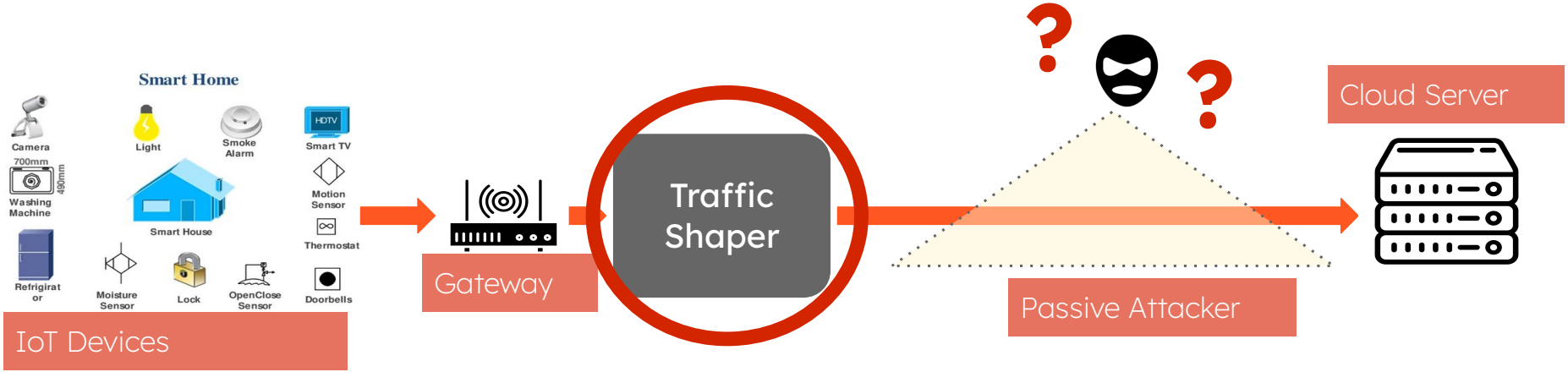


Captured as .PCAP binary file by Wireshark

Encrypted

Time Sent: 0.456 seconds

Privacy Enhancement



Traffic Shaper: Masks network header information by altering packet size and timing

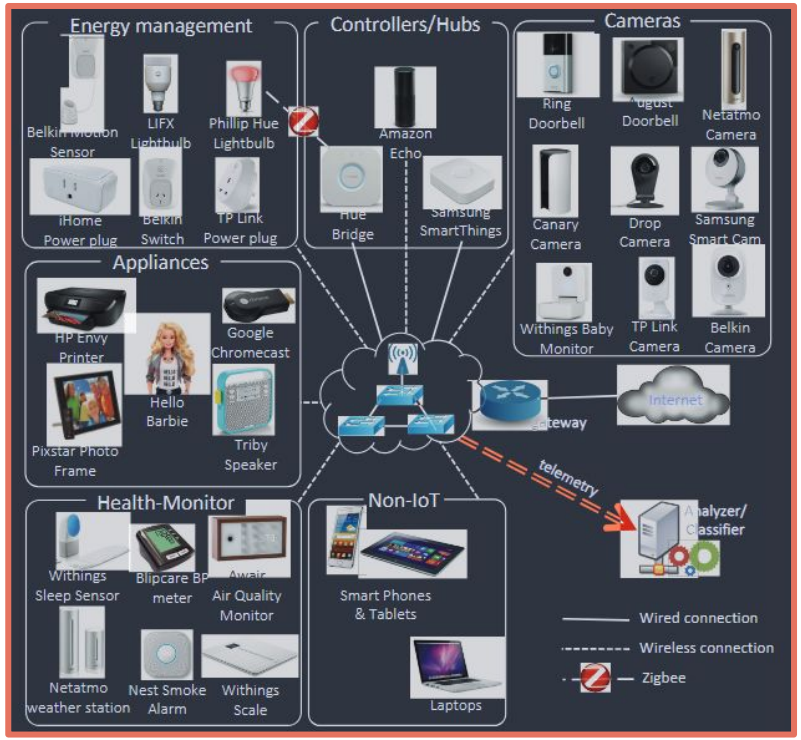
- Does so by reducing significant latency and resource consumption

Goals of Project

1. Develop a passive attack model capable of performing device fingerprinting

2. Evaluate traffic shaping effectiveness against model

Data Set



University of New South Wales (UNSW) IoTraffic

27 various IoT devices

200 days of continuous data captured on wireless testbed

Sequential Model: CNN

1D Convolutional Neural Network (CNN)

9 AM — 1 Day of Data — 11:59 PM

3 second size filter

10 second sample:
- 100 bytes total packet size
- 0.45 seconds average inter arrival timing

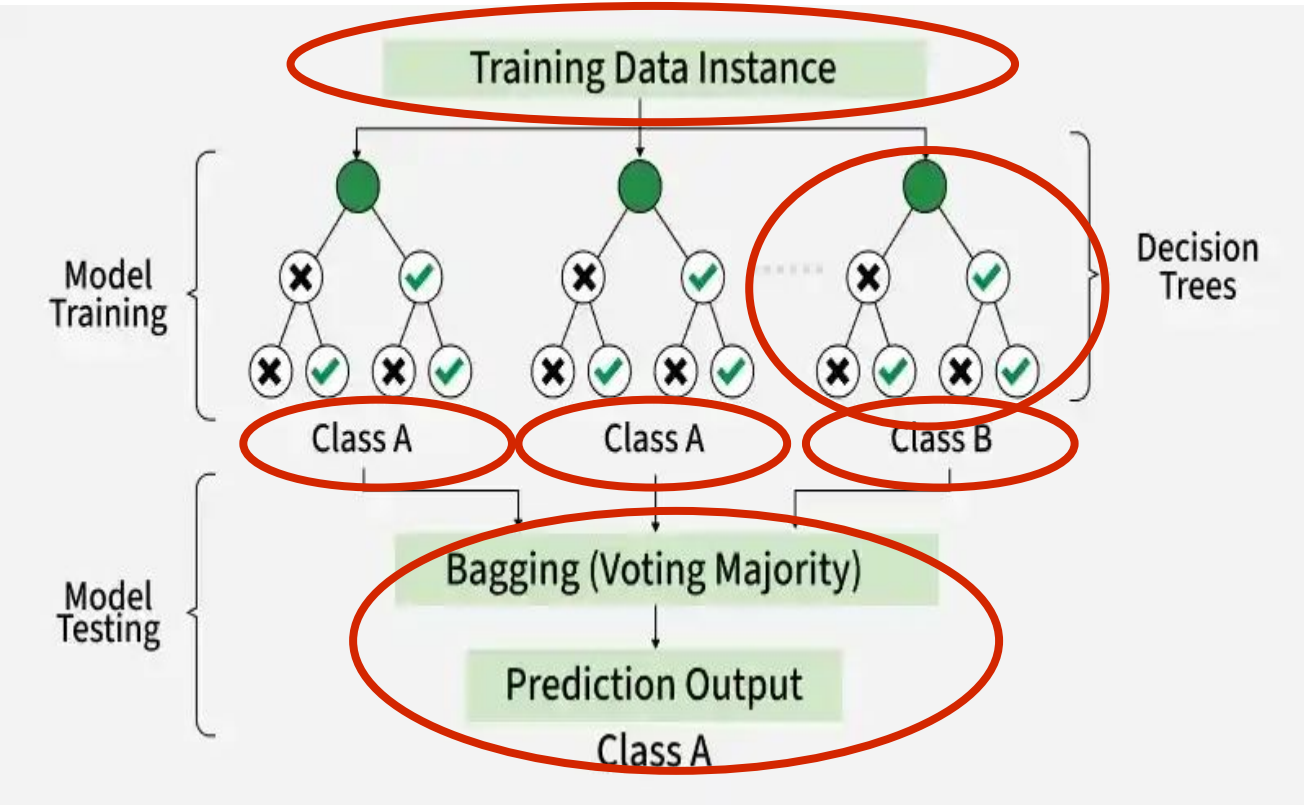
- Extracts complex patterns learned from each sequential time sample in the model
- Chooses most likely device based on patterns

Sequential Model: LSTM

Long Short Term Memory (LSTM)

- Type of Recurrent Neural Network (RNN)
- Cell structure with multiple gates choosing what information to retain long term
- Looks at data sequentially as it comes in; Good for time series

Random Forest



Each tree gets a subset of the data sampled with replacement

Each node of tree chooses subset of features without replacement to split on.

Each tree submits its individual vote on what they think the device is.

Device with most votes is decided as winner.

Optimal Split decided by minimum Gini Impurity

Preferred over Deep Learning in this case due to training speed and overall superiority.

E.g. "AvgPacketSize 200 bytes", "MinIAT .02 ms"

Feature Engineering

45 s Window 1

Total Packet size

Number of Packets

Packet size statistics
- Average
- Min/Max
- Standard Deviation
- ……

Inter-arrival Time (IAT) statistics

Duration of Transmission

• • •

45 s Window 10

Total Packet size

Number of Packets

Packet size statistics
- Average
- Min/Max
- Standard Deviation
- ……

Inter-arrival Time (IAT) statistics

Duration of Transmission

• • •

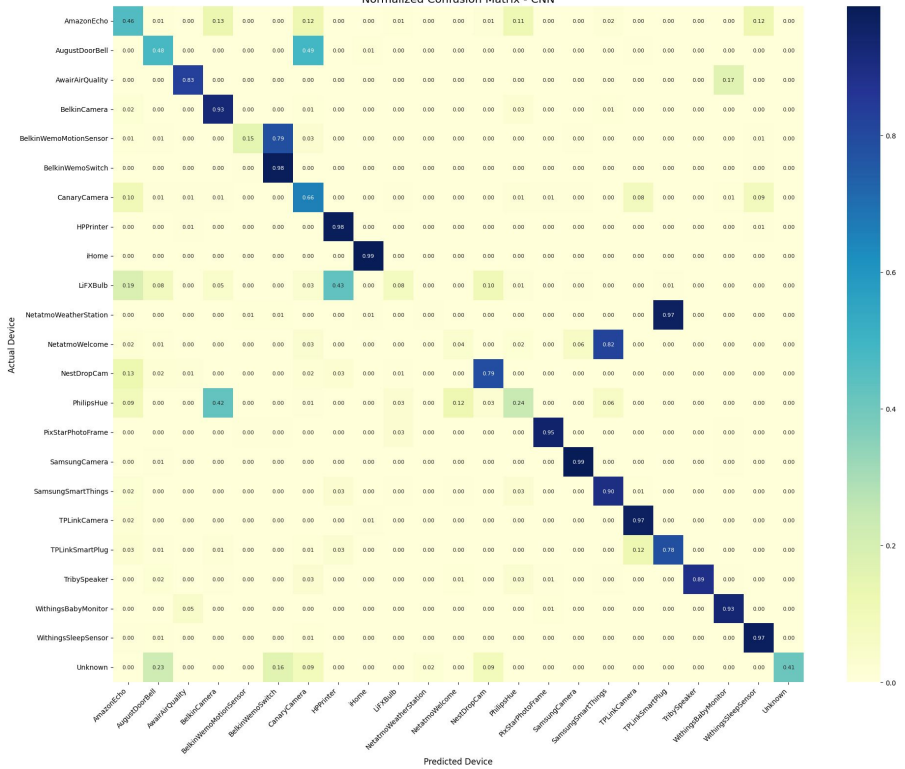
Results

Table 1: Classification performance metrics by Machine Learning Model

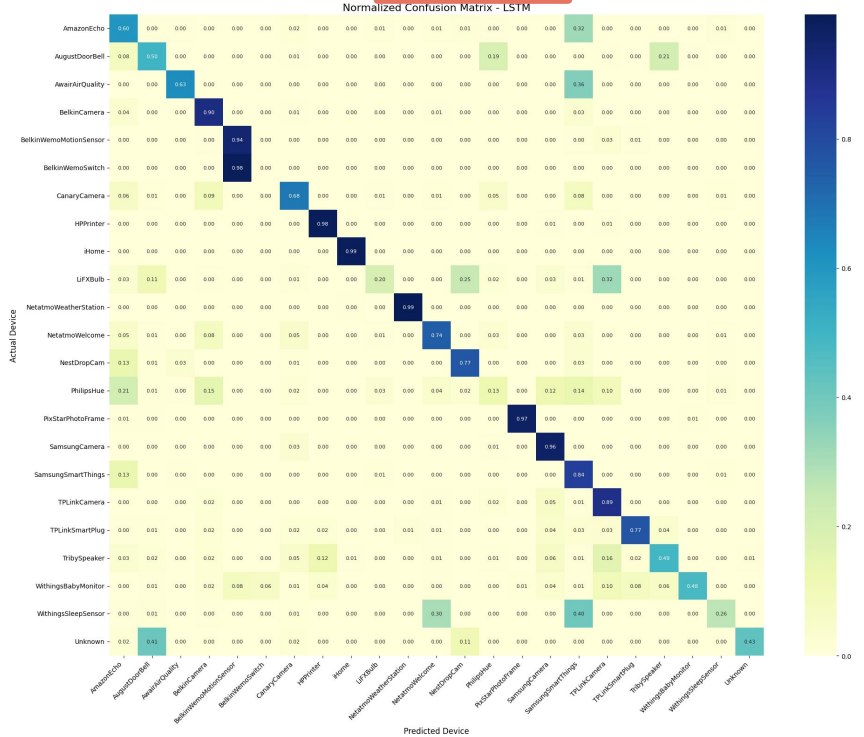
Model	Accuracy	Weighted F1 Score
CNN	69%	0.66
LSTM	66%	0.64
Random Forest	88%	0.89

Confusion Matrices

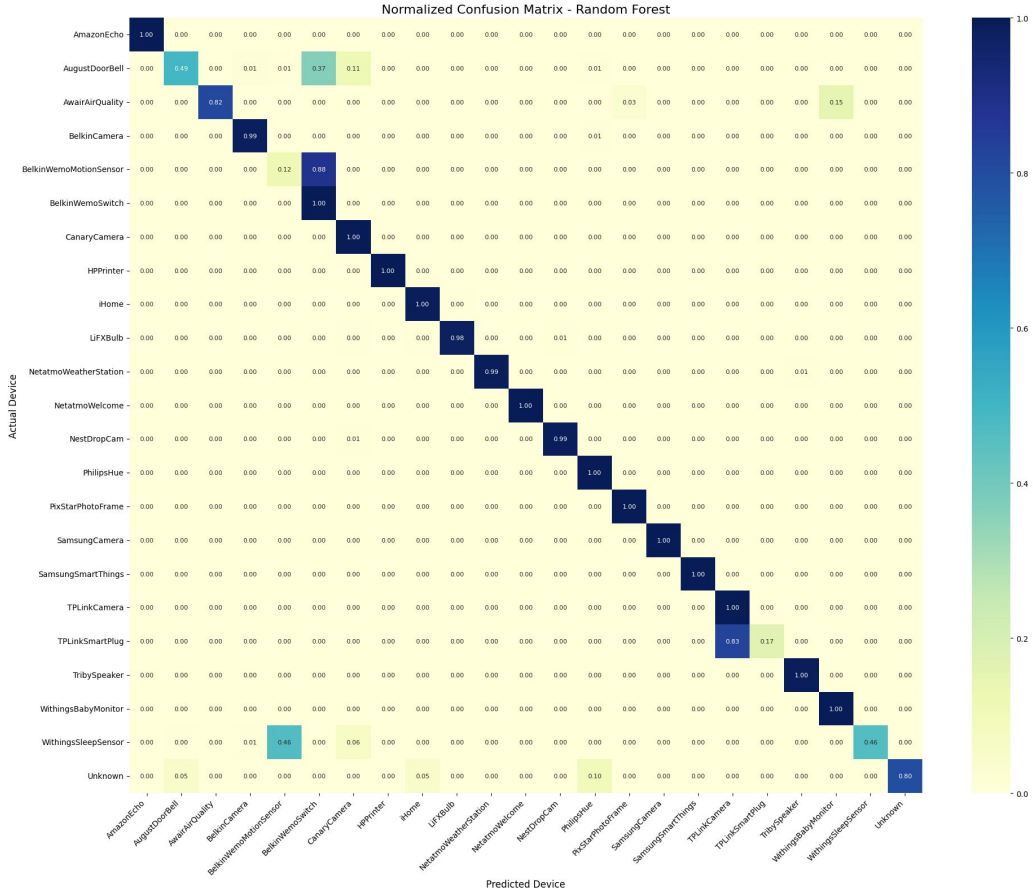
CNN



LSTM

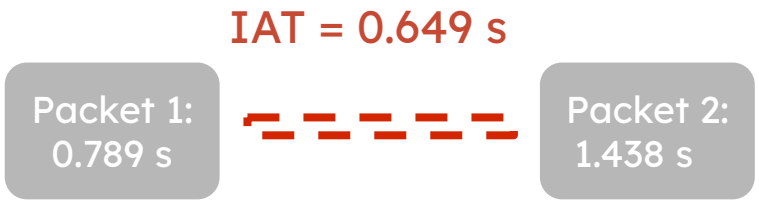


Confusion Matrices



Shapers

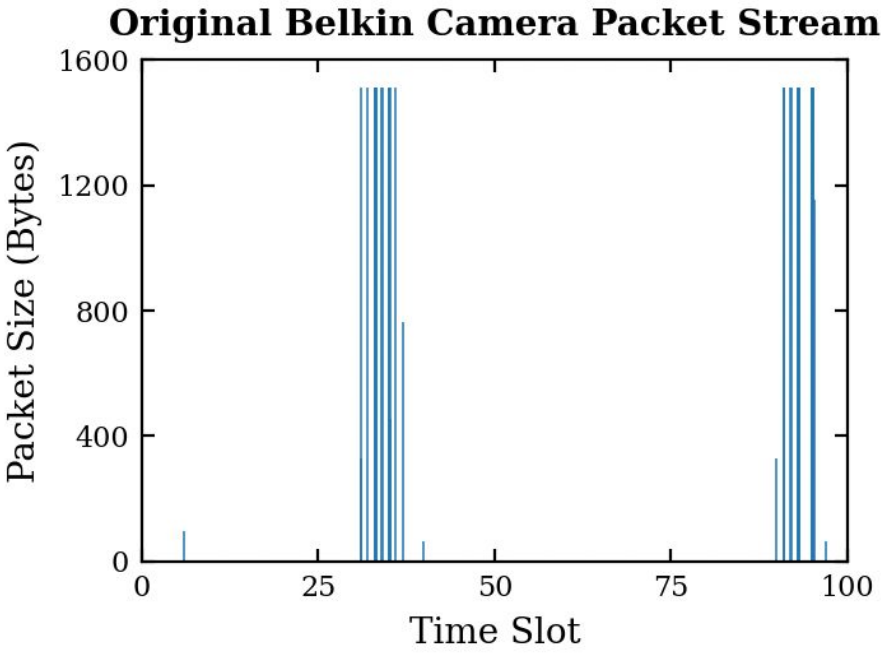
Shapers tested do not change packet size only inter-arrival timing



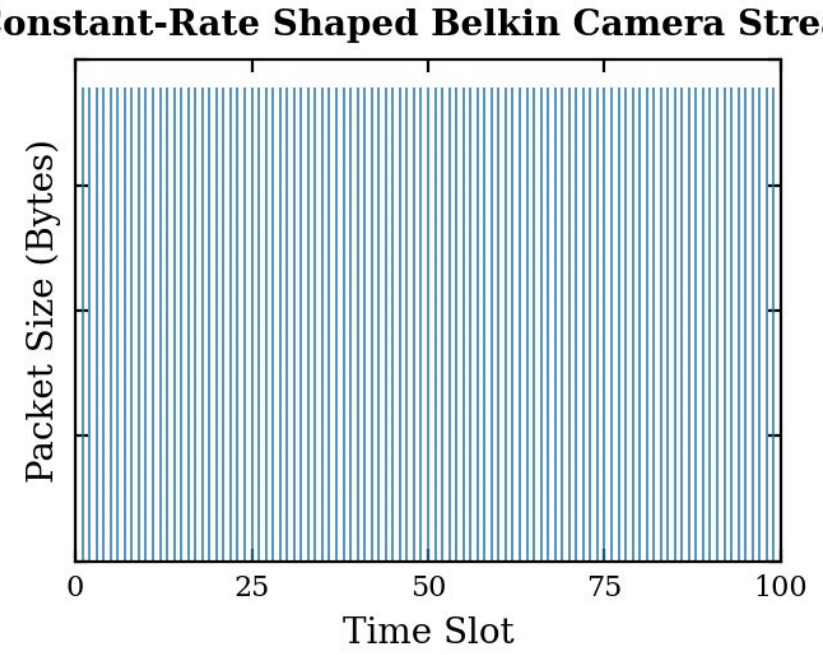
Constant Rate: Alters outgoing packets to all have a constant inter-arrival time.

Exponential Rate: Alters outgoing packets to all have a constant inter-arrival time.

Constant Rate



(a) Original Belkin Camera traffic.



(b) Constant-rate shaped Belkin Camera traffic.

Shaper Results

Table 2: Fingerprinting model confidence before and after viewing shaped traffic

Device Name	Unshaped Confidence	Shaper	Confidence
Belkin Camera	97.34%	Constant Rate	79.17%
Belkin Camera	97.34%	Exponential	76.71%
Amazon Echo	99.82%	Constant Rate	92.79%

Conclusions

- Shaping reduced confidence overall.
- Relying only on packet timing traffic shaping is not enough to deter attackers.
- Future work should work on more complex shaping that involves both **timing** and **size**.

Questions?