Packet Data Analysis

DTS: 360 - Grady Blair

Data Collection:

PWNAGOTCHI

What is it?

A packet data collection tool, primarily for collection handshakes.

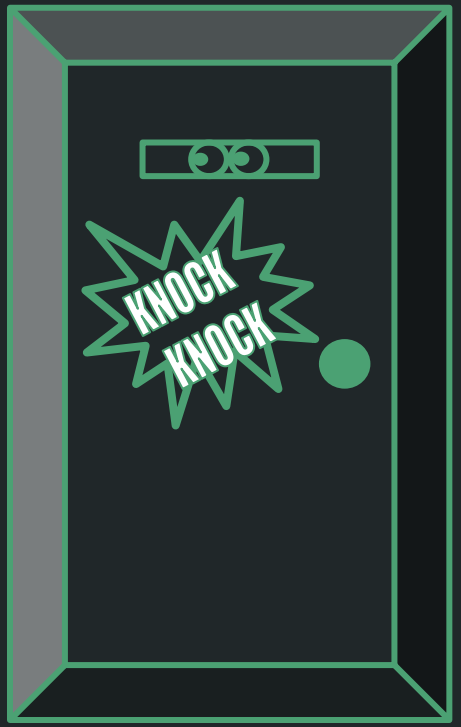
How does it work?

Tries to connect to wireless networks and records the interaction.

What is a HANDSHAKE?

Specifically a 4-way handshake, is a conversation between a device and a network:

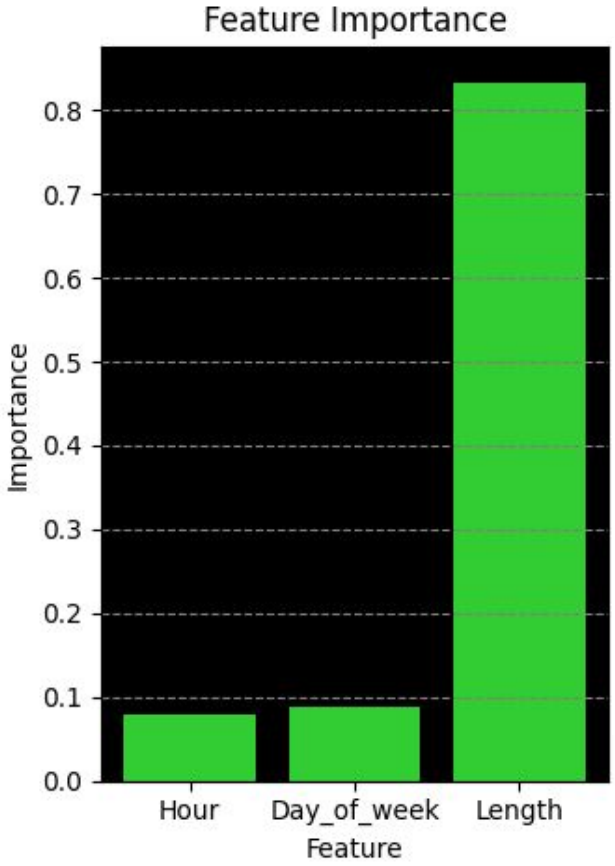
1. The client ask to connect.
2. The AP asks who is connecting
3. The client sends its identity to the AP
4. The AP verifies the identity of the client and sends the necessary information to connect securely.

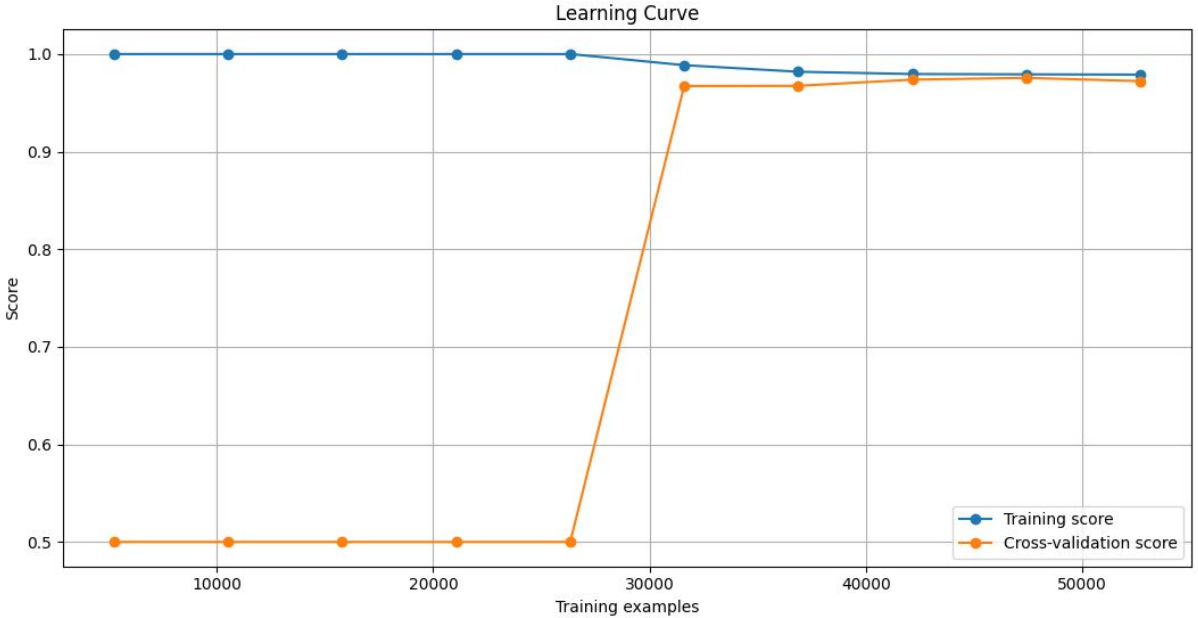


My Model -> Predicting Handshakes

- Trained on multiple variables of packet data such as packet length (bytes) and capture time
- The model is not super complex so it is fairly accurate
- The goal of the model is to predict handshakes, but it could serve a purpose in helping protect networks or attack them.

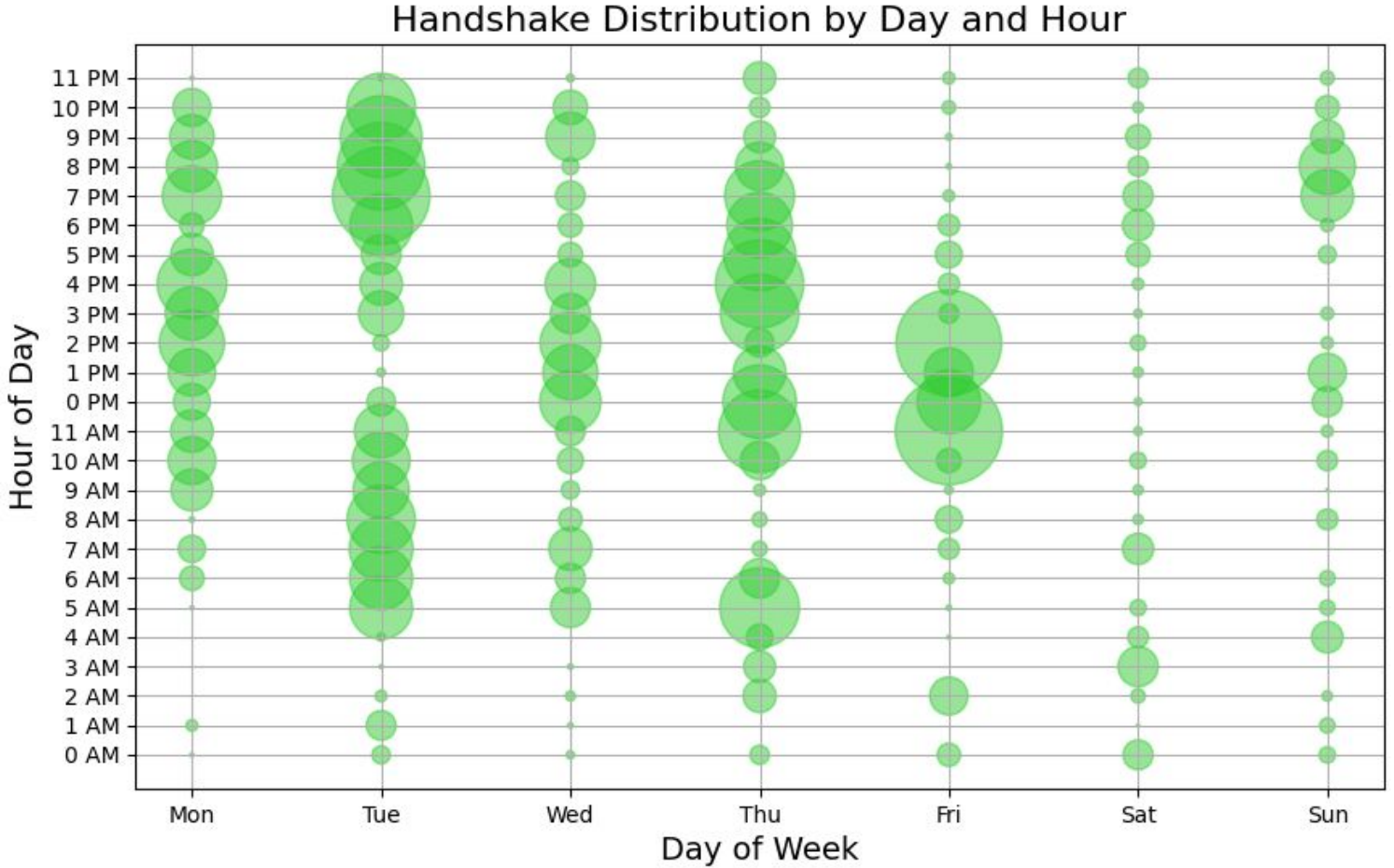
No.	Time	Source	Destination	Protocol	Length	Info
1	2024-04-11 22:36:18.776996	32:92:5e:d5:f0:87	Broadcast	802.11	285	Beacon frame, SN=1047, FN=0, Flags=........C, BI=100, SSID
2	2024-04-11 22:41:09.085700	32:92:5e:d5:f0:87	Microsoft_35:51:a7	EAPOL	161	Key (Message 1 of 4)
3	2024-04-11 22:41:09.084185	32:92:5e:d5:f0:87	Broadcast	802.11	285	Beacon frame, SN=10, FN=0, Flags=........C, BI=100, SSID='
4	2024-04-11 22:41:09.085700	32:92:5e:d5:f0:87	Microsoft_35:51:a7	EAPOL	161	Key (Message 1 of 4)


Feature Importance
Importance
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Hour
Day_of_week
Length
Feature


Learning Curve
Score
1.0
0.9
0.8
0.7
0.6
0.5
10000
20000
30000
40000
50000
Training examples
Training score
Cross-validation score

How/Use Case

- WPA/WPA2 are the most common security standards used for wireless networks
- However they are vulnerable.
- The handshakes captured contain a hash or an encrypted version of the networks password which can be cracked.
- When networks are busy they are more likely to produce handshakes for deauthentication attacks like those used by the pwnagotchi.


Handshake Distribution by Day and Hour
Hour of Day
11 PM
10 PM
9 PM
8 PM
7 PM
6 PM
5 PM
4 PM
3 PM
2 PM
1 PM
0 PM
11 AM
10 AM
9 AM
8 AM
7 AM
6 AM
5 AM
4 AM
3 AM
2 AM
1 AM
0 AM
Mon
Tue
Wed
Thu
Fri
Sat
Sun
Day of Week

Future Development

With more collected data and the ability to process that data better I believe the model could be used to help security analyst to:

- Identify deauth attacks
- Recognize suspicious devices
- Find vulnerabilities in their network security

Overall create an Intrusion Detection System to detect abnormal behavior on a network

Questions?