Task 2 – AI Integration of BlockShield Project

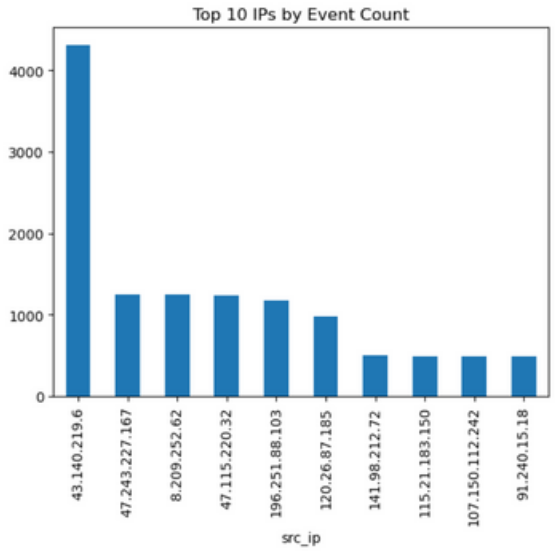
Exploring Cowire Data & Classifying IPs Using Clustering

Nashat Alfarajat – BlockShield AI Development Team

Understanding the Data

To analyze attack patterns, I explored a Cowrie honeypot log dataset containing around 29,700 rows and 36 columns, mostly consisting of text, timestamps, and network data. Some missing values were expected due to the nature of logging.

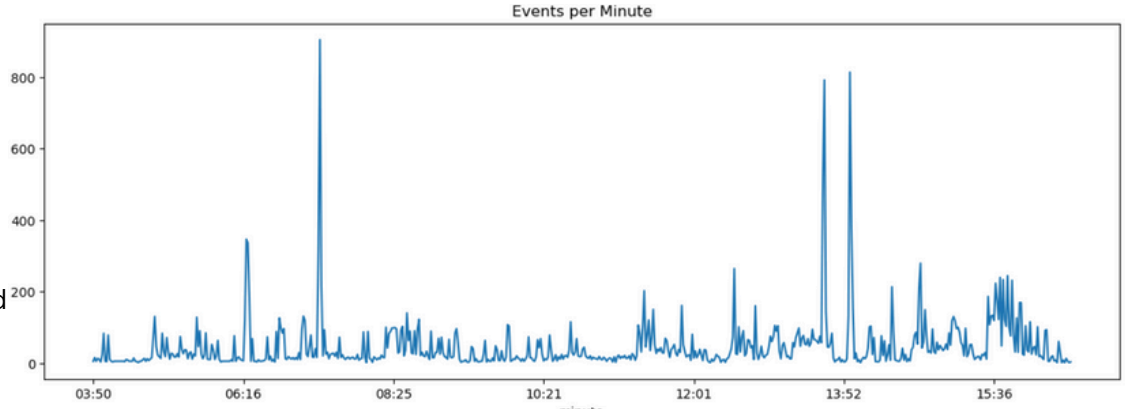
I started by reviewing key columns like protocol, src_ip, and message. Many values were missing, but patterns were clear. Some IPs appeared thousands of times, and many messages showed repeated connection attempts. I also saw common weak credentials like admin/admin and root/123456, which likely point to automated attack tools.



Time-Based Analysis

The dataset spans July 11, from 03:50 AM to 04:43 PM (UTC).
I plotted event frequency by minute. The highest spike occurred at 07:36, with 904 events in one minute. Another burst appeared around 13:50.

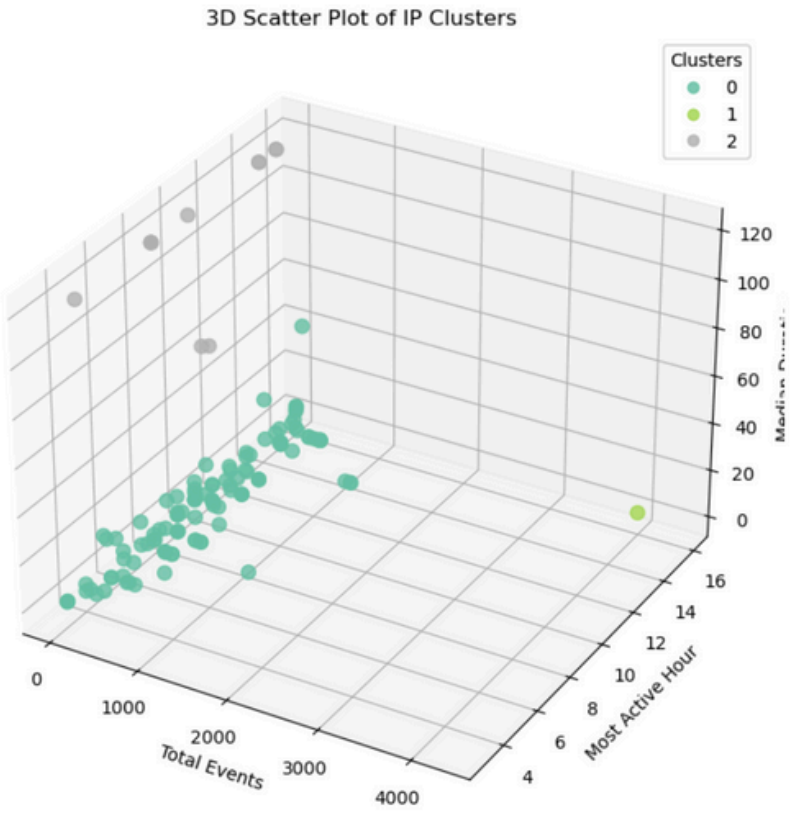
Peak IP Analysis:
During the 07:36 peak, a single IP — 47.243.227.167 — was responsible for 884 of 904 requests (99%). This is a strong signature of automated scanning or bot behavior.



Clustering for IP Classification

I applied unsupervised clustering to group IPs based on activity patterns. First I reduced data to unique IPs by aggregating events, enabling IP-level behavior clustering. The features used were all derived from the timestamp:

- **Total events per IP**
- **Most active hour**
- **Median session duration**



These features reflect how active each IP is, when it operates, and how long it stays connected — helping spot patterns like brute-force bots, scanners, or possible manual intrusions.
The clustering revealed:

- A large group with low, short activity (likely baseline noise - low risk)
- A small group with long sessions (potentially suspicious)
- A single clear outlier with extreme activity (likely malicious)

The clustering achieved a **silhouette score of 0.61**, showing the clusters are well-formed and distinct. Most IPs are more similar to others in their cluster than to those in different clusters, reflecting clear behavioral differences like activity and session duration. This indicates the clustering captures meaningful patterns and should remain effective when applied to larger datasets for deeper analysis.

Clustering gave helpful insights, but a larger dataset and expert knowledge (to understand what each group represent exactly) would improve results. Merging with datasets like CIC-IDS 2017 isn't ideal due to different data structures — Cowrie logs show events, while CIC-IDS logs network flows, also a few thousand records wouldn't add much to a 2.6M dataset. A better model could be built using supervised learning on labeled data.