Detecting Network Intrusions

Janet Carson

3 Million Records Stolen Every Day

- Retail
- Healthcare
- Government

Source: https://www.cybintsolutions.com/cyber-security-facts-stats/

Network logs

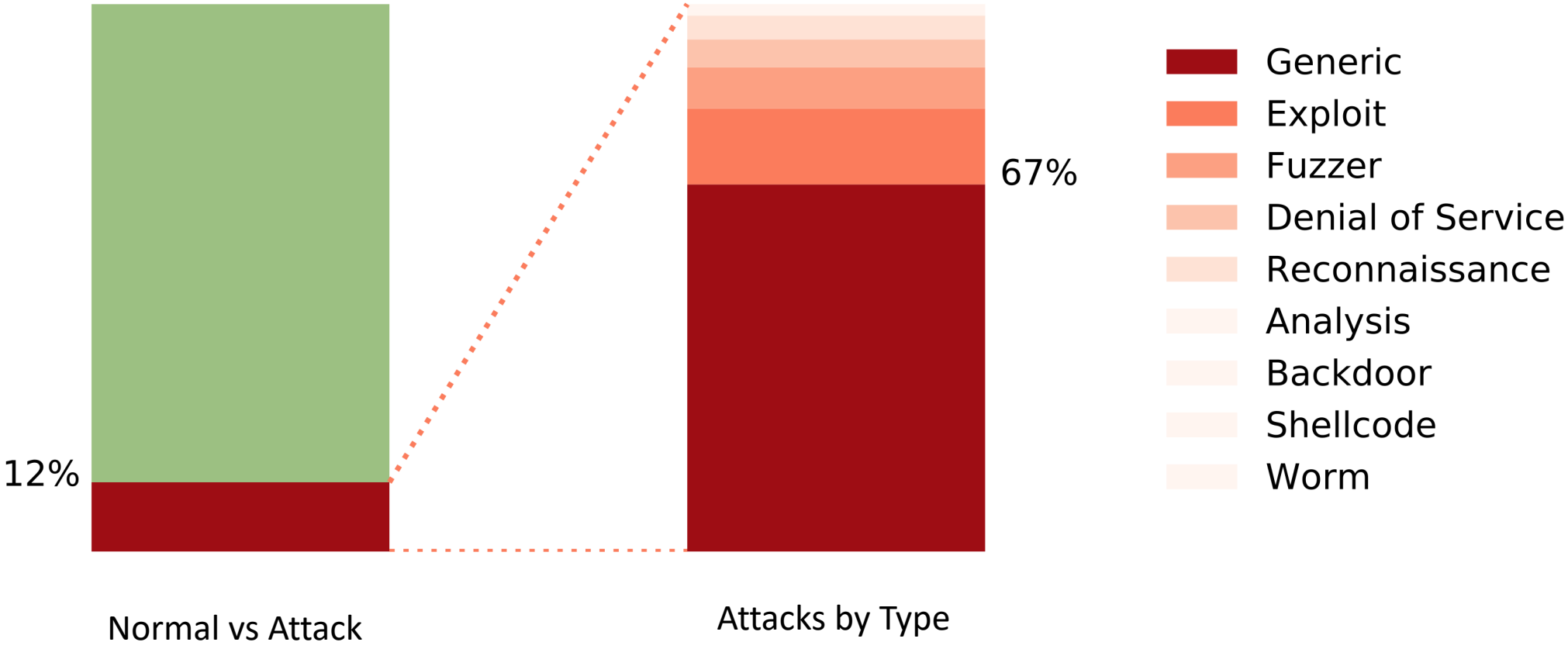
- Records of computer network traffic
- Can be used for real-time analysis or for forensic investigation

UNSW-NB15 Dataset

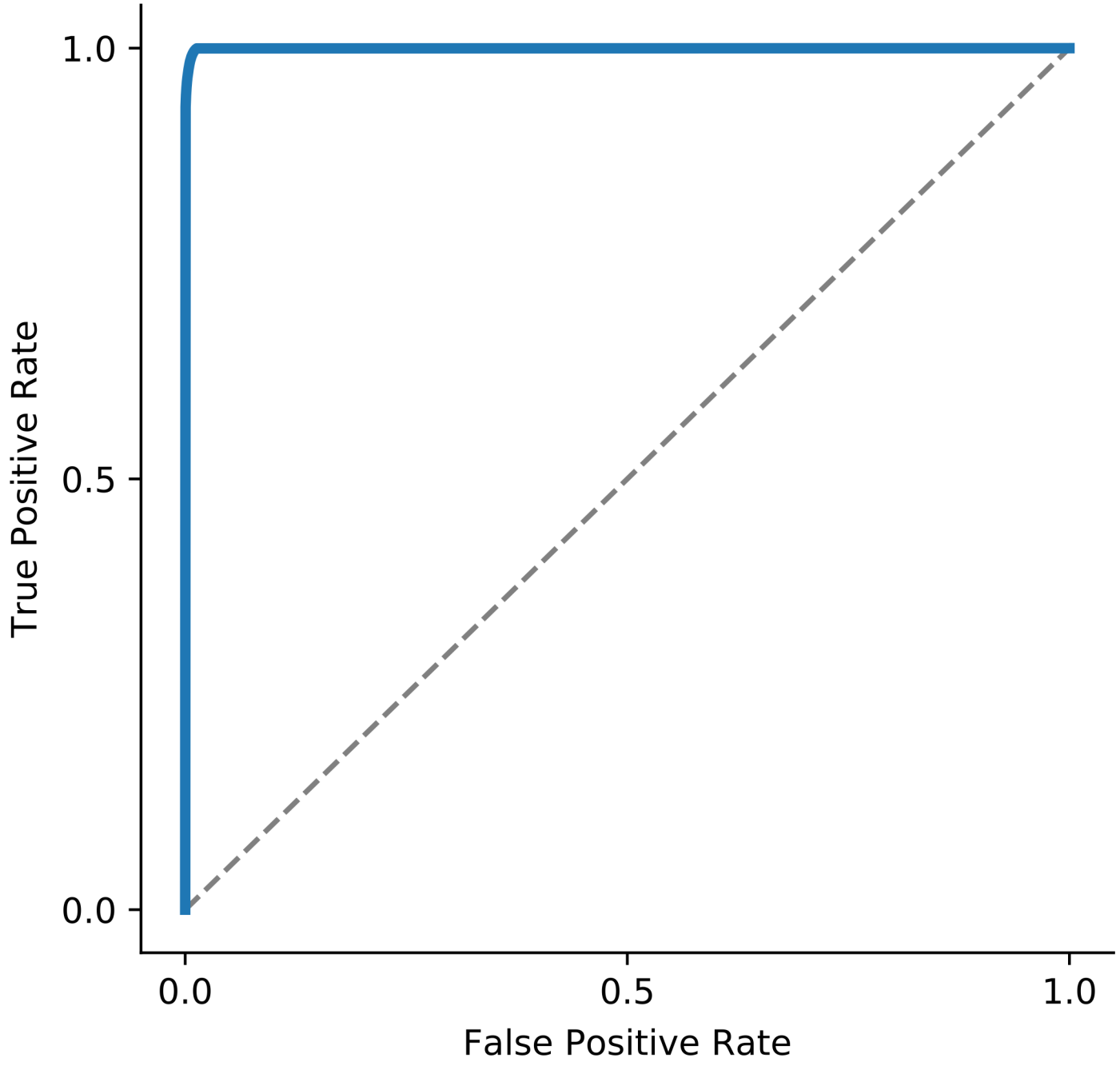
- 2,500,000 records
- 174,000 Attacks
- 3,000 Vulnerabilities

1. Normal
2. Generic
3. Exploits
4. Fuzzers
5. Denial of Service
6. Reconnaissance
7. Analysis
8. Backdoors
9. Shell code
10. Worms

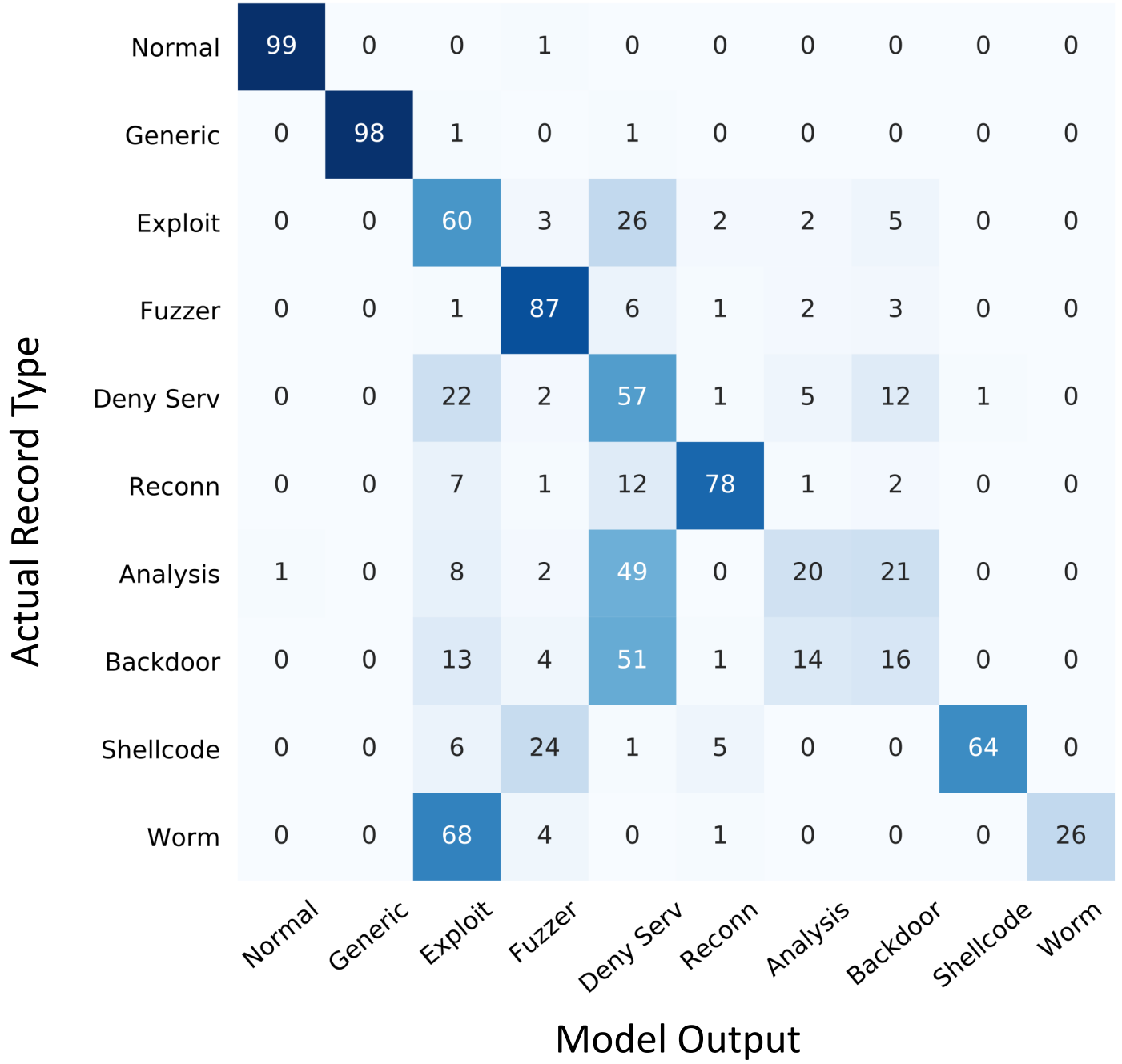
Unbalanced Data



Attack Detection

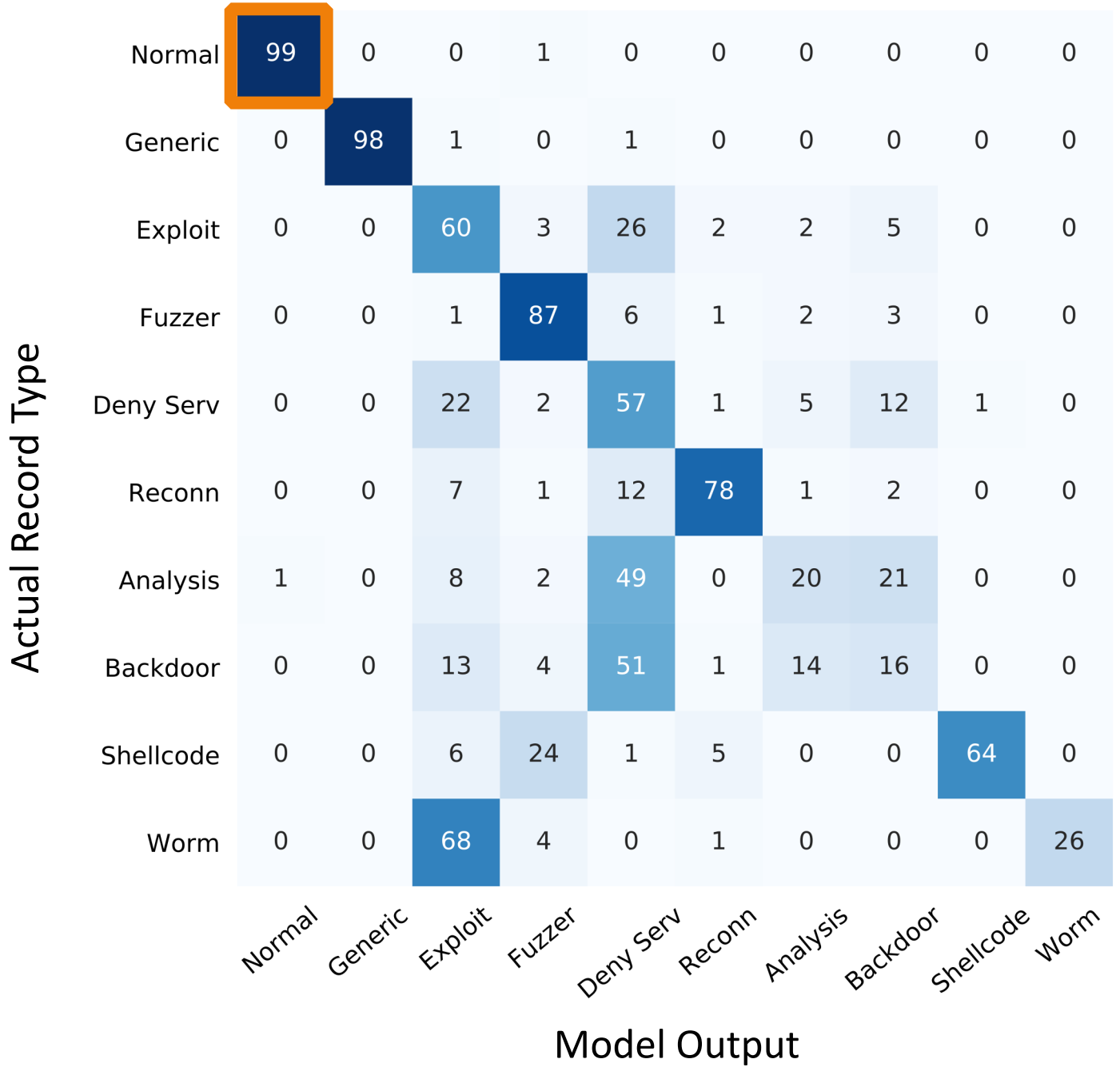


Attack Labeling



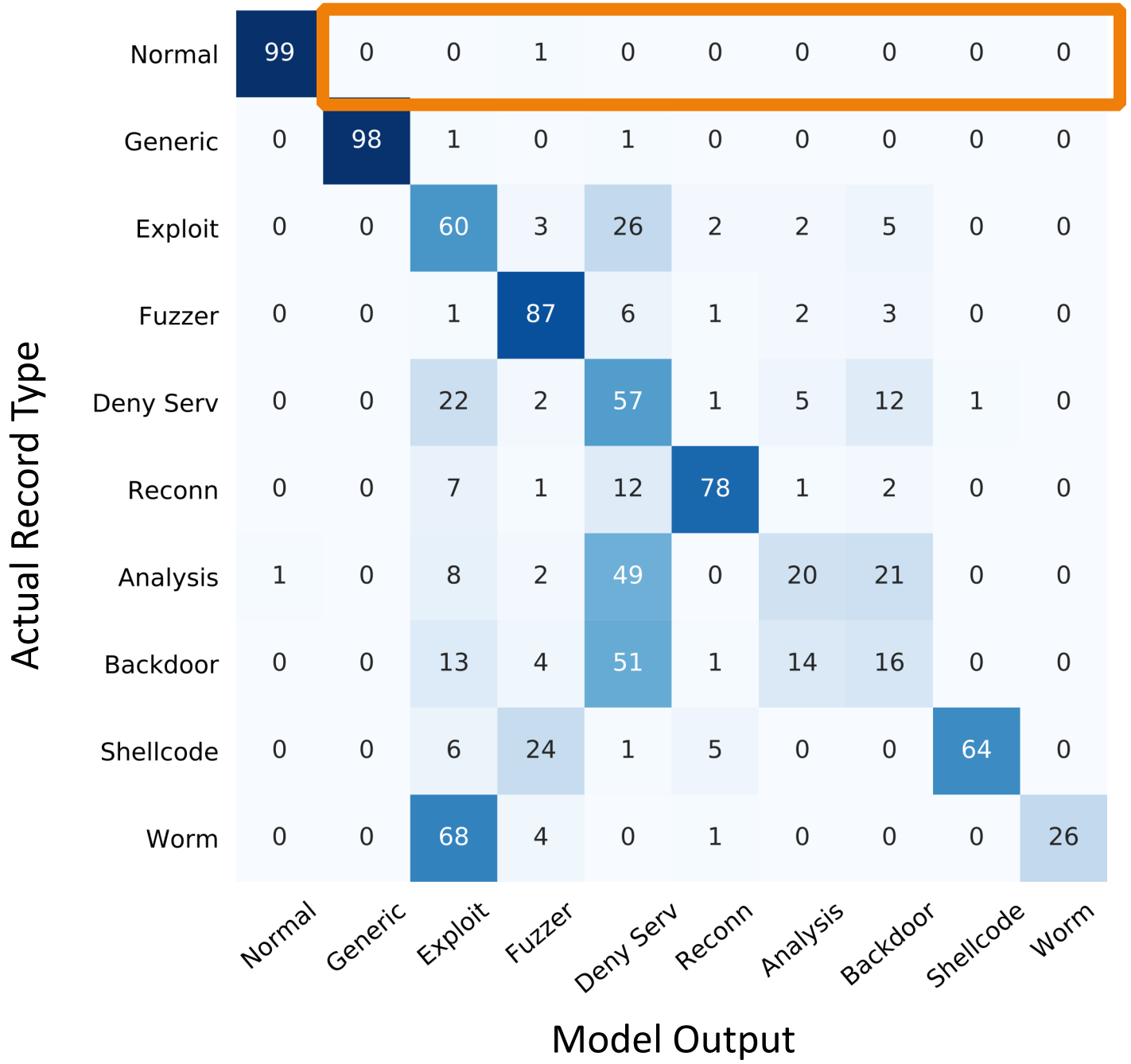
Attack Labeling

99% Normal



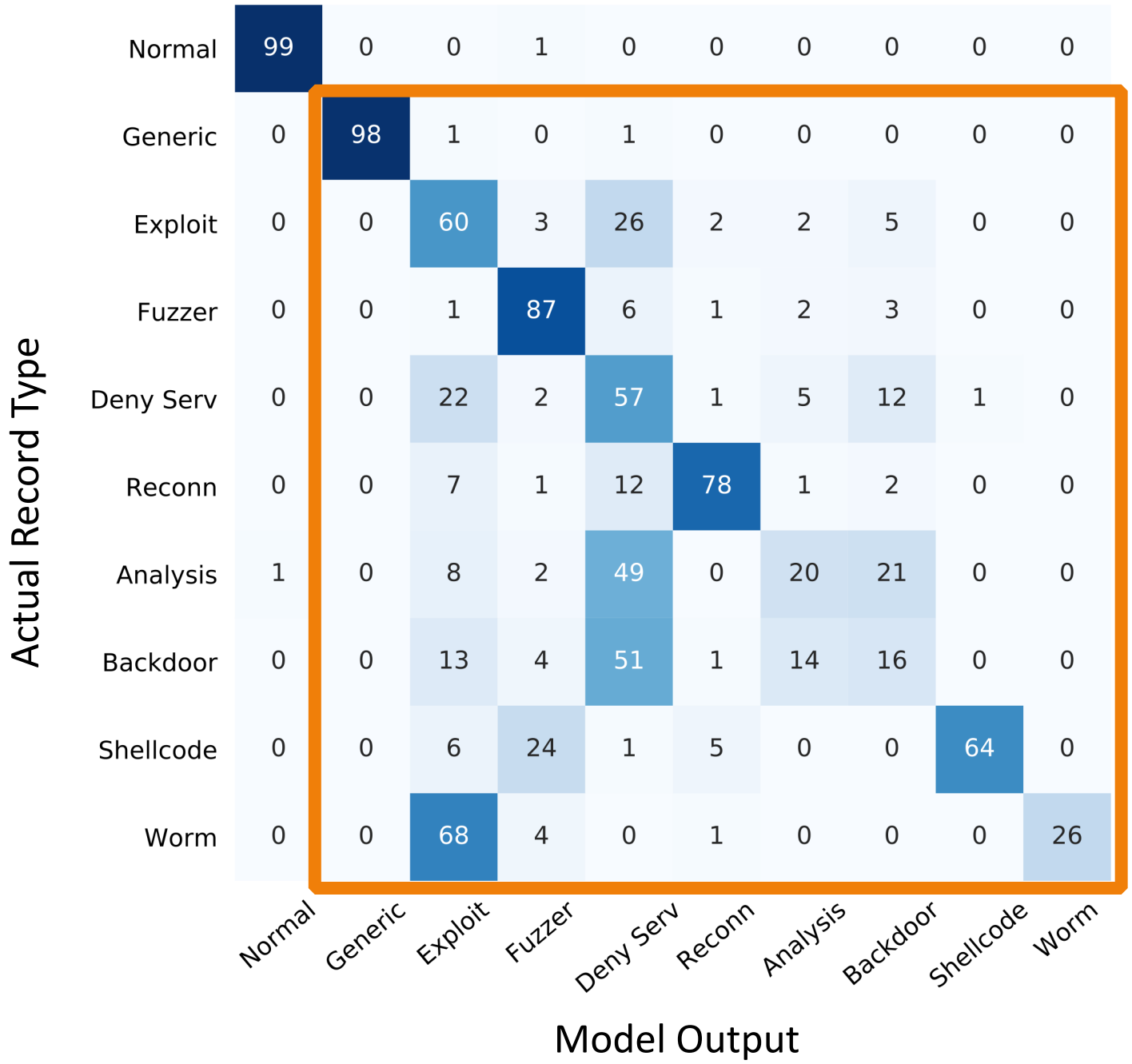
Attack Labeling

1% False alarms

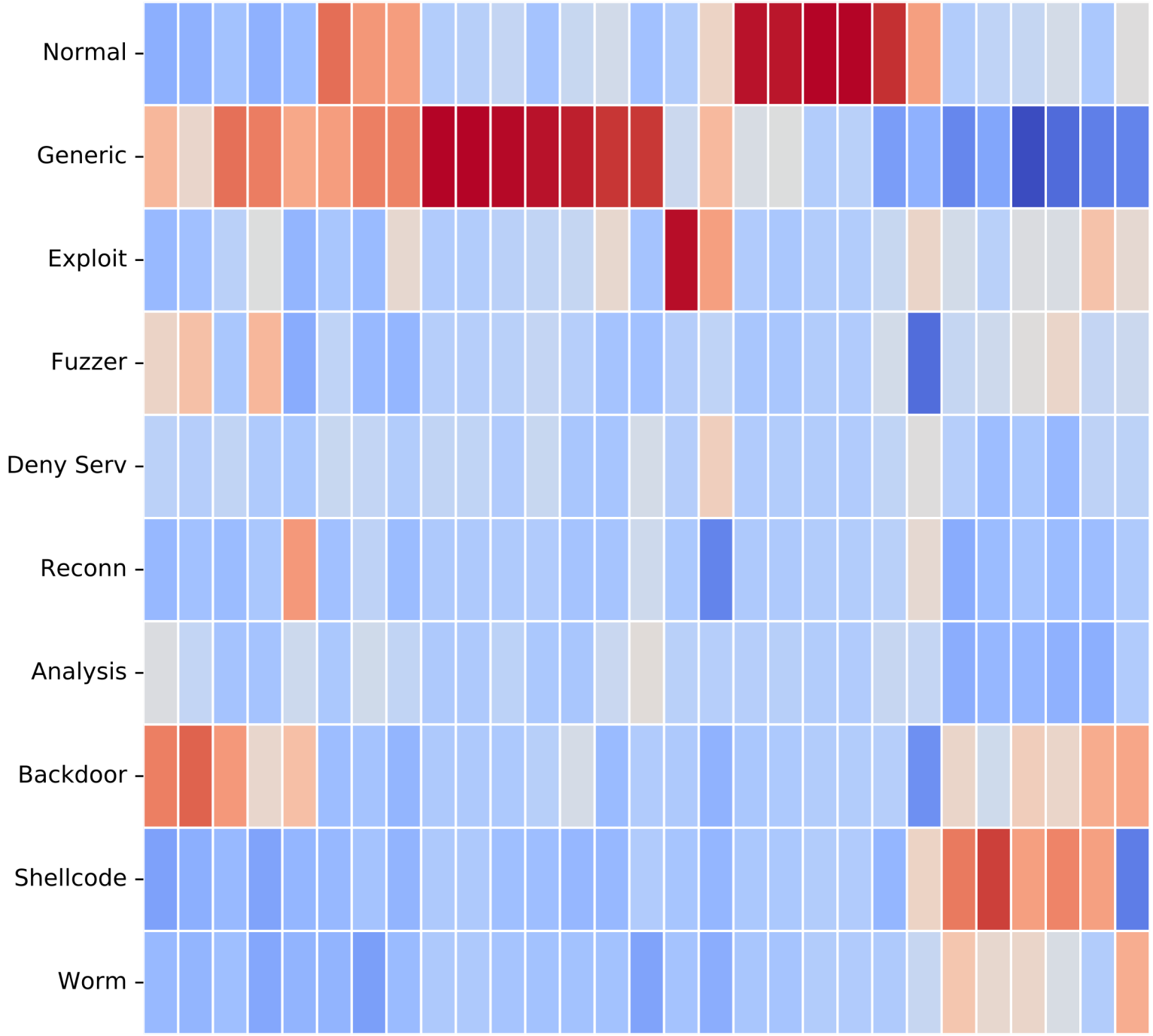


Attack Labeling

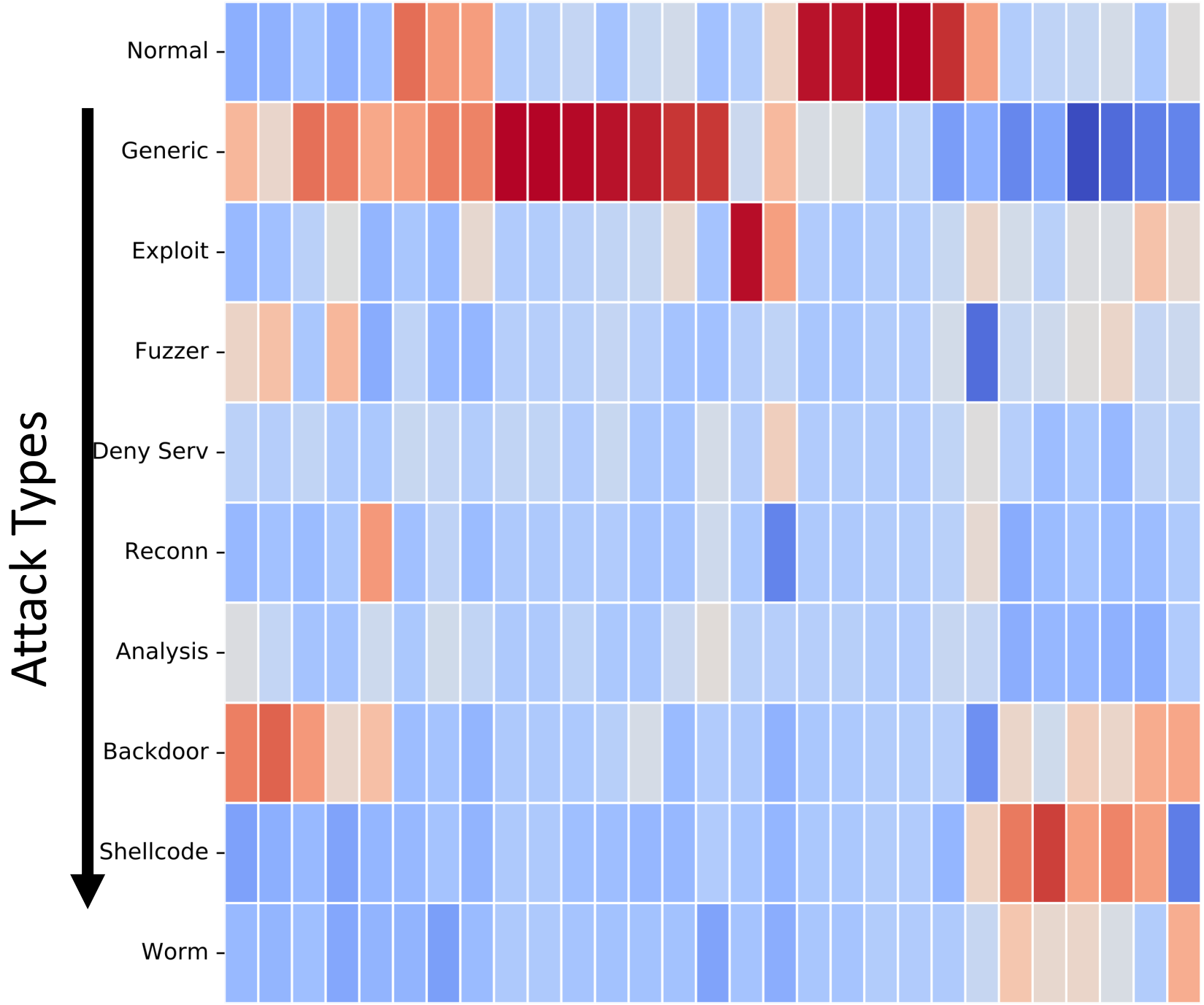
99.99% of attacks
-- 87% correct label



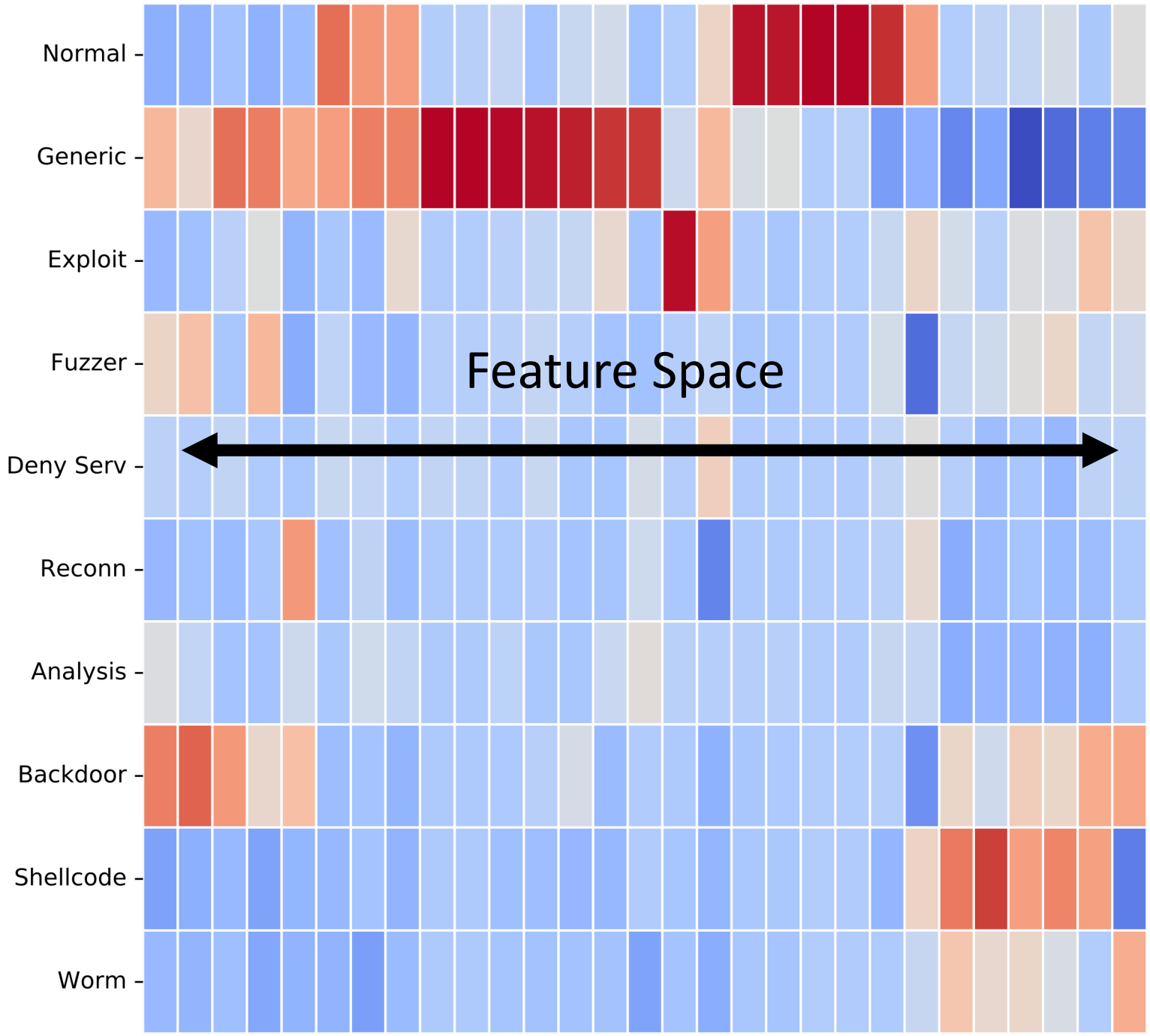
Underlying Data



Underlying Data

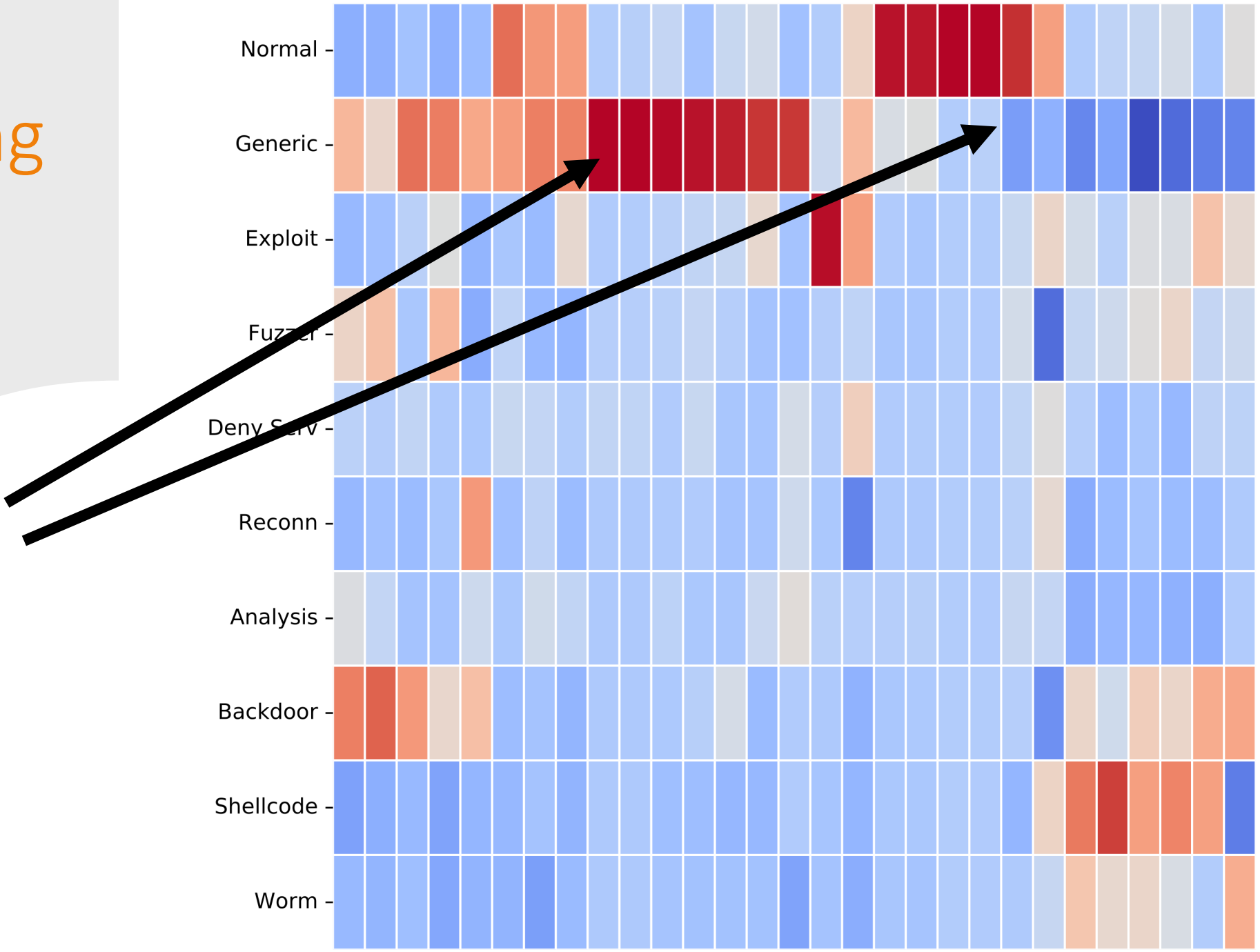


Underlying Data

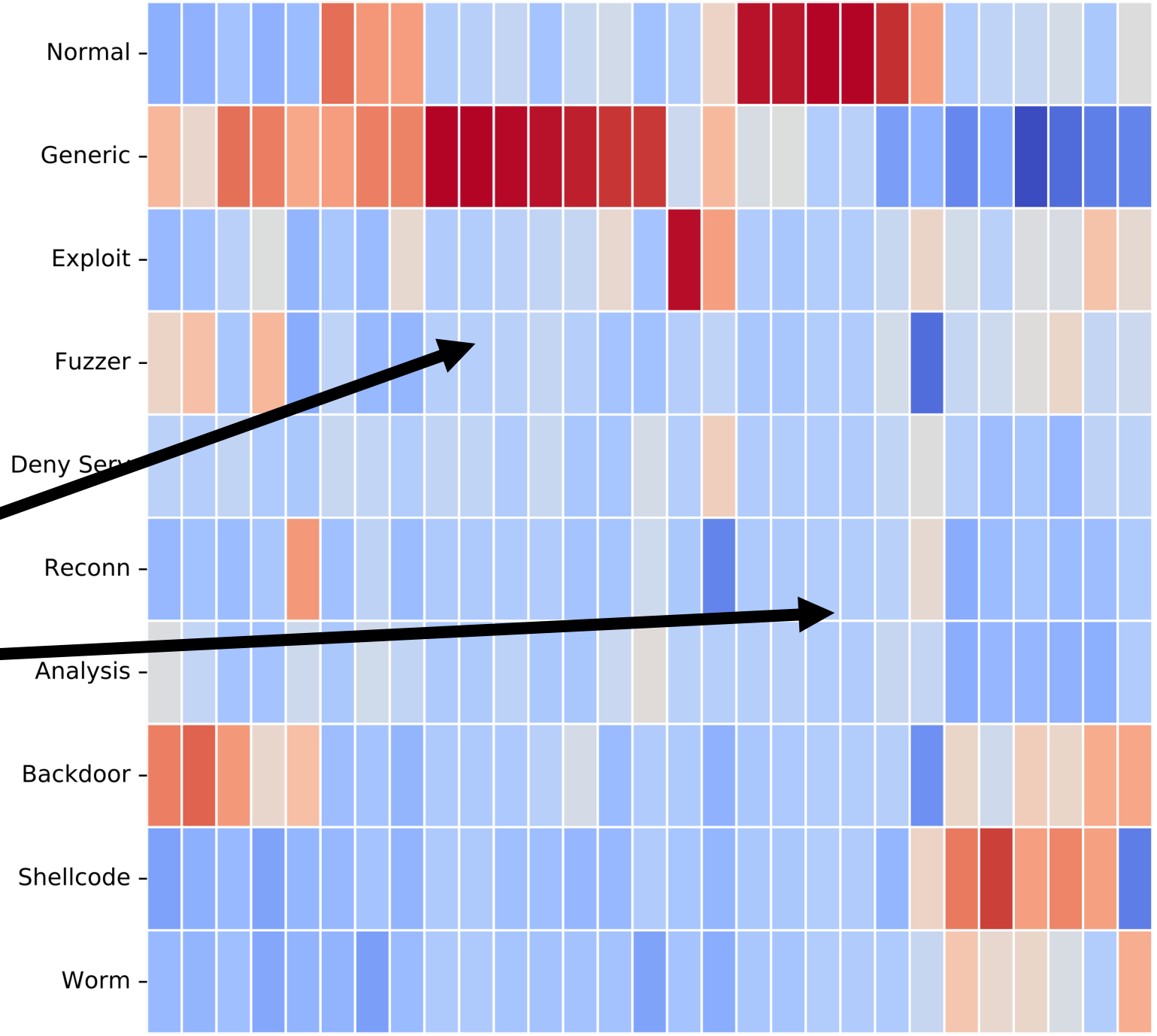


Underlying Data

Distinctive features

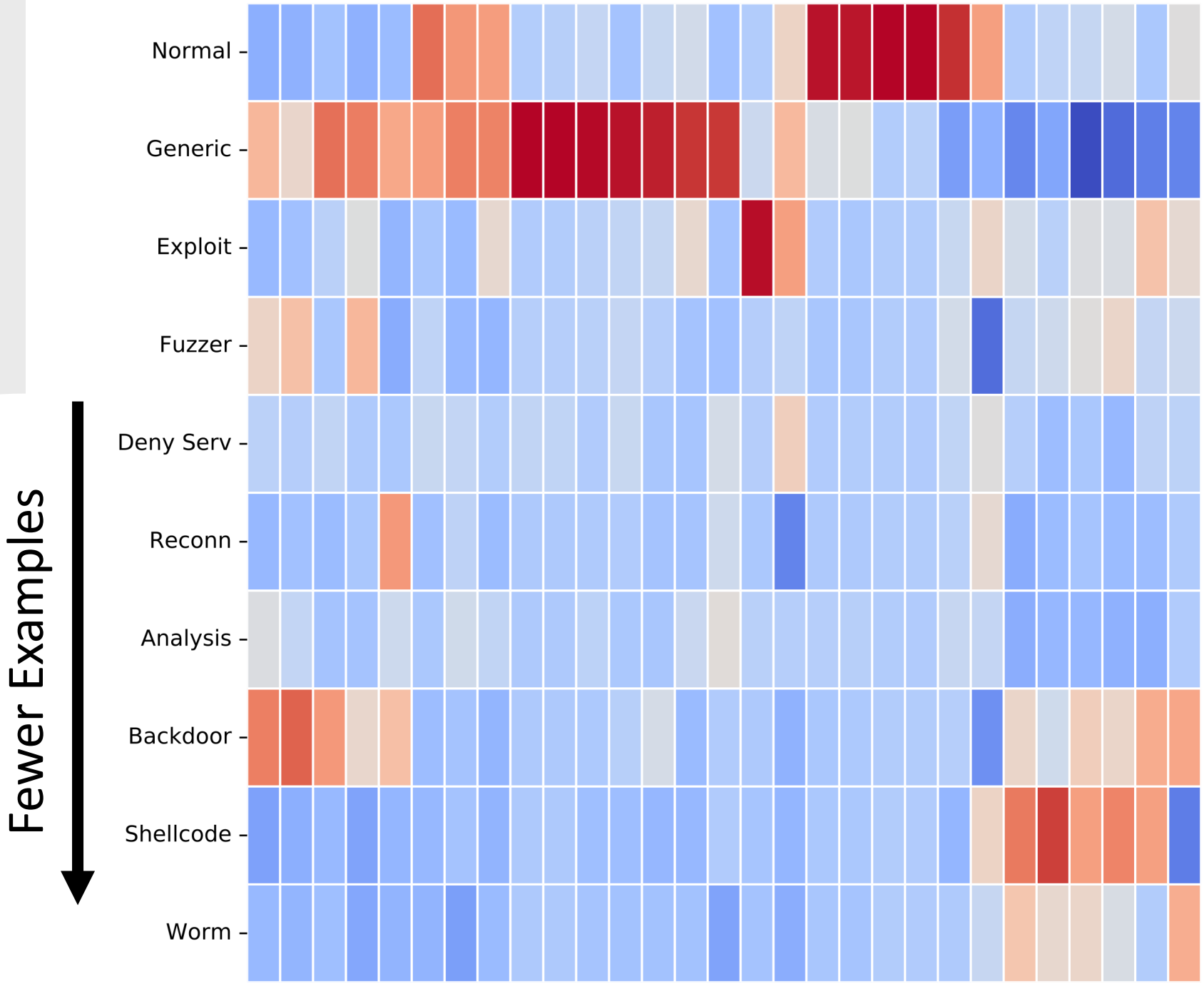


Underlying Data

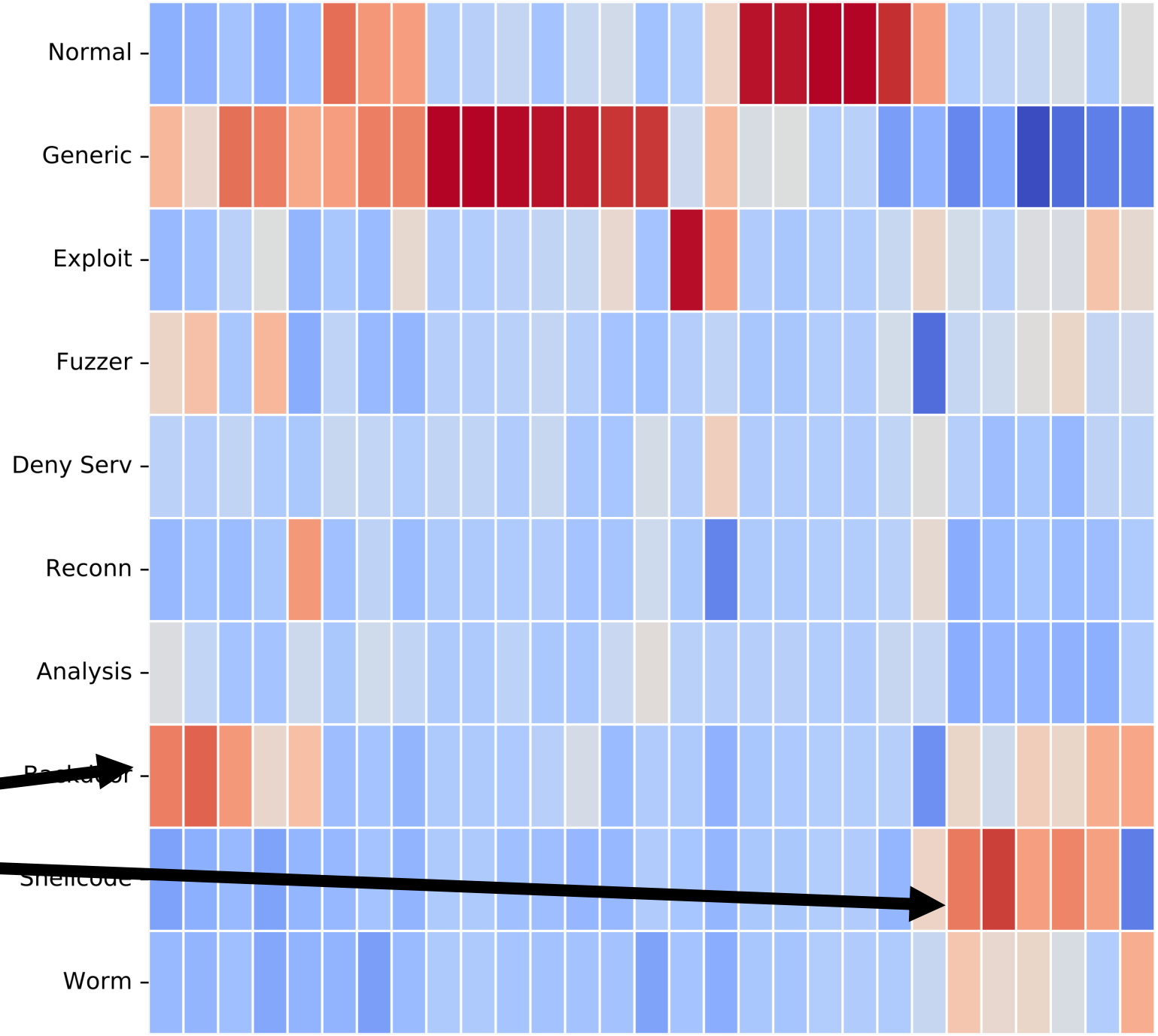


Zone of "blah"

Underlying Data



Underlying Data



Fewer strong indicators

Conclusions and Future Work

- Successfully detected attacks, often their subtype
- Better training data would help
  - More balanced training examples
  - Feature extraction

Thank You

Janet Carson

carsondata