Applications of Conformal Prediction in Information Security Problems

Giovanni Cherubin

12 April 2016
Classification

- Attacks detection: malware, network traffic.
- Pre-analysis malware filtering.
- Attacks: traffic analysis, side channel.
Machine Learning and Information Security

Classification

- Attacks detection: malware, network traffic.
- Pre-analysis malware filtering.
- Attacks: traffic analysis, side channel.

Anomaly detection

- Discover new threats.
- NIDS.
Machine Learning and Information Security

Classification

• Attacks detection: malware, network traffic.
• Pre-analysis malware filtering.
• Attacks: traffic analysis, side channel.

Anomaly detection

• Discover new threats.
• NIDS.

Clustering

• Monitoring tools.
• NIDS.
Conformal Prediction
[VGS05]

• A statistical framework to make predictions.

• Controls the number of error.

• Applications that require confidence.

For classification, anomaly detection, clustering, …
Conformal Prediction

“non-conformity measure”

\[ A(\text{;}) = 0.2 \]
Conformal Prediction

The error committed is smaller or equal to $\varepsilon$
Classification of Malware

Conficker

Torpig

Zeus

New object
Classification of Malware

Conficker

Torpig

Zeus

New object
Classification of Malware

Conficker

Torpig

Zeus

New object
Classification of Malware

- Conficker
- Torpig
- Zeus

New object
Classification of Malware

Conficker

Torpig

Zeus

New object
Classification of Malware

$Y = \{\text{Zeus, Torpig}\}$

New object
Classification of Malware

\[ Y = \{\text{Zeus, Torpig}\} \]

Validity: \( Pr(y \notin Y) \leq \varepsilon \)
Anomaly Detection

[SNCOG14]

Anomaly detection of maritime trajectories
Bots clustering

[CNNG+15]

Forming clusters of similar bots network traces
An attack...

hello darling...

tic tic tic tic tic tic tic...
Hidden Markov Models

tic .... tic

t_{12} \quad t_{23} \quad t_{34}

he \quad el \quad ll

...
Confident Prediction for HMM

[CN16]
Confident Prediction for HMM

[CN16]

t_{12} → CP
  da, he, ge, ho, ...

t_{23} → CP
  el, ol, ad, ...

t_{34} → CP
  ll, dd, ...

t_{34} → CP
  dy, lo, ly, ...

daddy, hello, fully, gello, ...

Confident Prediction for HMM [CN16]

daddy, hello, fully, gello, ...

da, he, ge, ho, ...
el, ol, ad, ...
l, dd, ...
dy, lo, ly, ...

t_{12} \rightarrow \text{CP}
t_{23} \rightarrow \text{CP}
t_{34} \rightarrow \text{CP}
t_{34} \rightarrow \text{CP}
Confident Prediction for HMM

[CN16]
Conclusions

- Conformal Prediction has interesting properties for Information security.

- It can be used for anomaly detection, classification, clustering, and beyond.

- Applications in Information Security exist, and much more can be done.
References


Pictures from https://openclipart.org.
Applications of Conformal Prediction in Information Security Problems

Giovanni Cherubin

12 April 2016