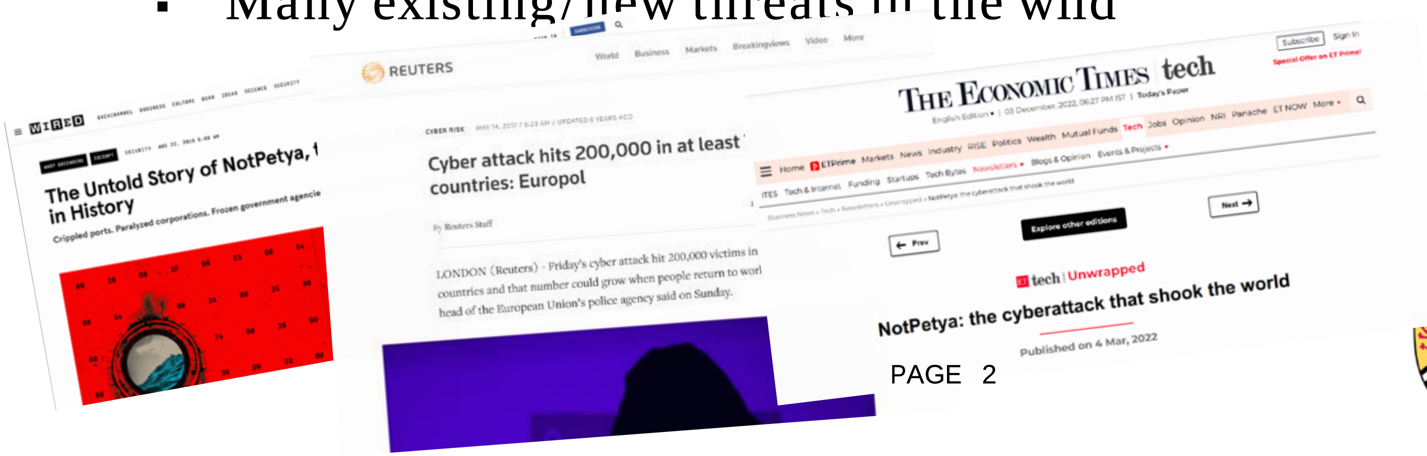
Soteria: An Approach for Detecting Multi-Institution Attacks

Saif Zabarah, Omar Naman, Mohammad A. Salahuddin, Raouf Boutaba, Samer Al-Kiswany

UNIVERSITY OF WATERLOO

Multi-institution Attacks (MIA)

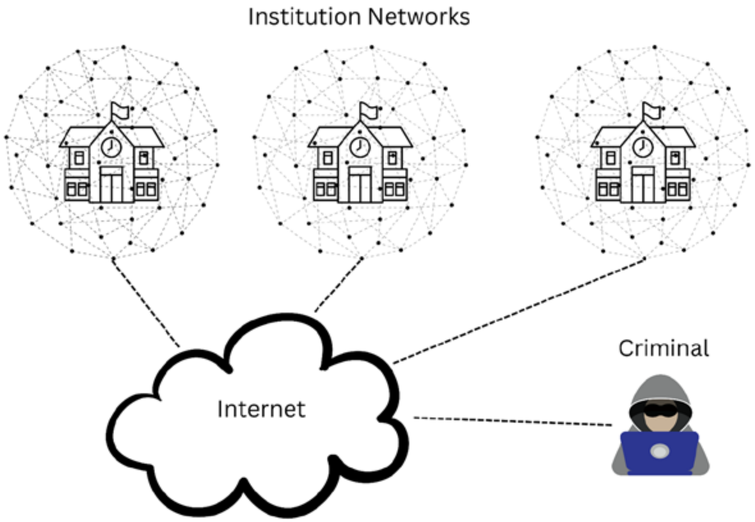
- An Attack targeting multiple institutions in a short time period
- Examples:
  - WannaCry affected 200,000 computers in 150 countries (2017)
  - NotPetya, estimated loss is $10 billion (2017)
- Challenging to defend:
  - Vulnerabilities change quickly
  - Attacks happen quickly
  - Many existing/new threats in the wild



UNIVERSITY OF WATERLOO

MIAs are Challenging in the Education Sector

- Large and constantly changing networks
- Low budget and understaffed teams
- Prime targets for MIAs
  - Cybercrime cost institutions an average of $9.25 million in 2019
  - 46% of institutions reported attacks in 2017



Related Works

- Reconnaissance works:
  - Limited to detecting port scans
- Heavy Hitter detection:
  - Detecting hosts that communicate with large number of hosts
  - Limited to predicting the number of hosts
- Current approach relies on sharing intel (e.g. Virus Total)
  - Threat sharing delays
  - It requires cybersecurity experts time
  - Privacy constraints

Requirement of an MIA detection tool

- Accurately predict attacks
- Severity estimation
- Predicting the next victims of an attack

UNIVERSITY OF WATERLOO

Soteria - the contribution

- A data analysis pipeline for detecting MIAs
- Uses graph analysis and ML
- Deployed as part of CANARIE IDS
- Overview of the results,
  - Able to predict MIAs
  - Predict future attacks with 95% recall rate
  - Estimates the severity of the attack with high accuracy
  - Predict the next targets of an attack with 95% recall rate
  - Detect attacks in the first 20% of their life span

Outline

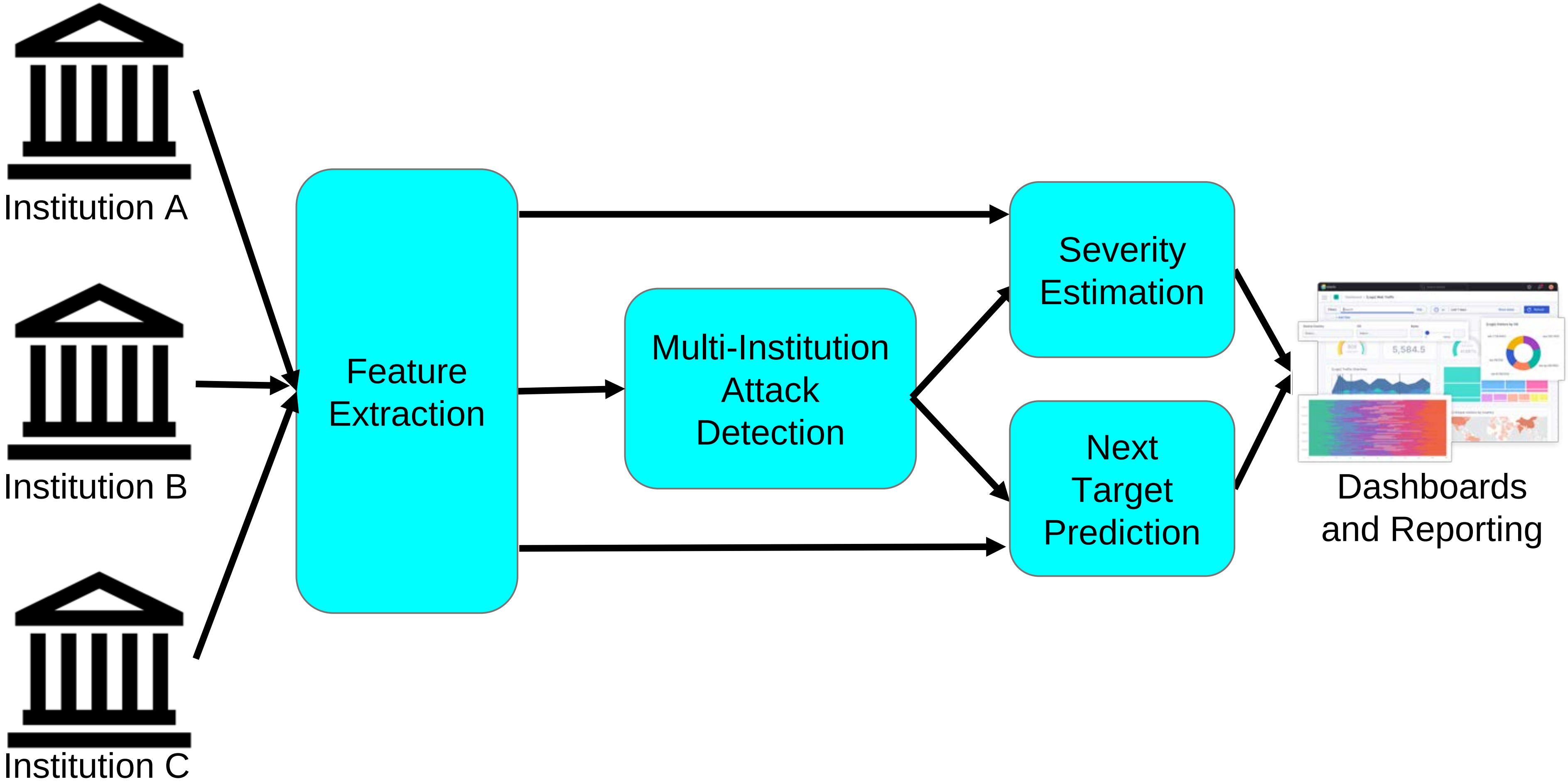
- Motivation and Introduction
- Soteria design
  - Feature Extraction
    - Static Metrics
    - Dynamic Metrics
  - Attack detection
  - Severity Estimation
  - Next Target Prediction
- Evaluation
- Conclusion

Outline

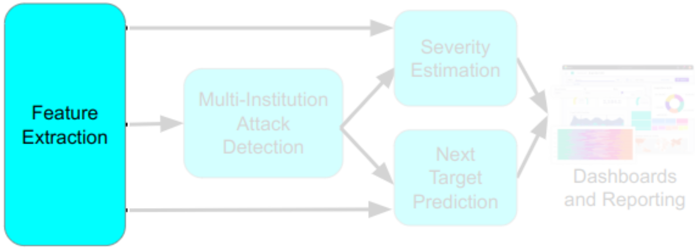
- Motivation and Introduction
- Soteria design
    - Feature Extraction
        - Static Metrics
        - Dynamic Metrics
    - Attack detection
    - Severity Estimation
    - Next Target Prediction
- Evaluation
- Conclusion

UNIVERSITY OF WATERLOO

Soteria design

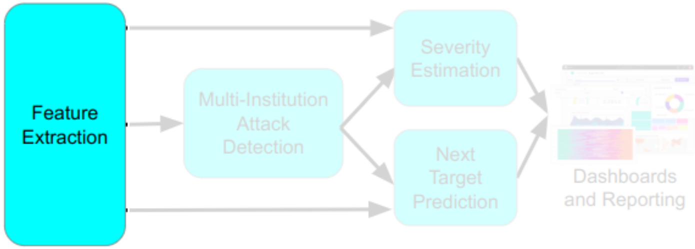


Feature Extraction



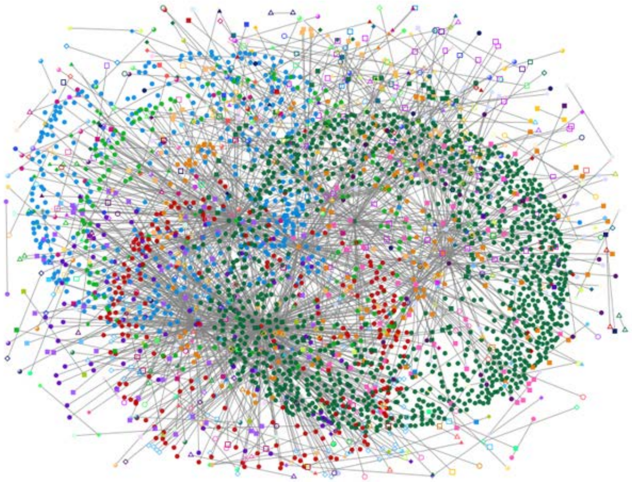
- Institution share zeek logs
- Input dataset from connection logs
  - id.orig_h: Source ip
  - id.resp_h : Destination ip
  - ts: Timestamp
  - local_orig: is the orig ip local
- Topological graphs are a natural representation for the dataset and the attack

Feature Extraction

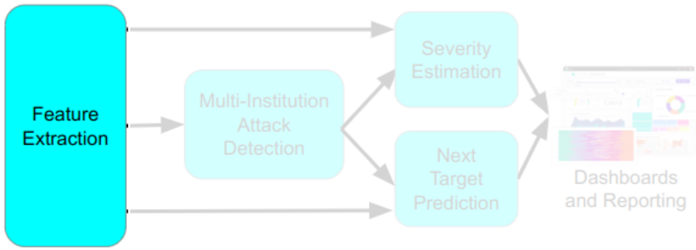


Challenge: Generating a graph from the dataset does not scale

- Graphs are massive
- Processing metrics is slow

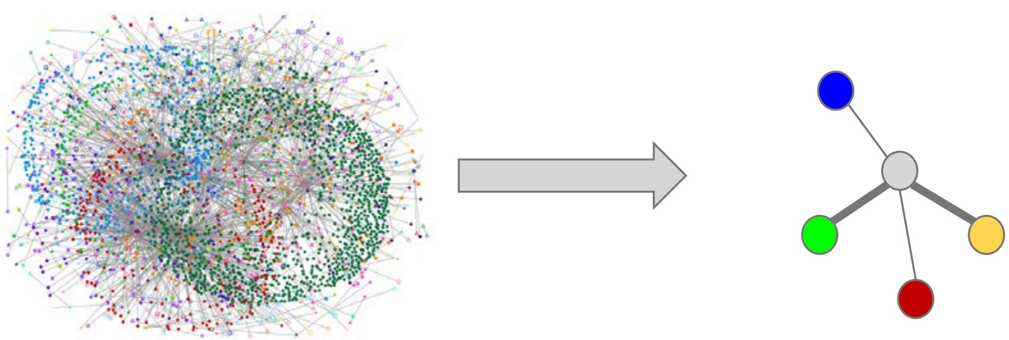


Feature Extraction

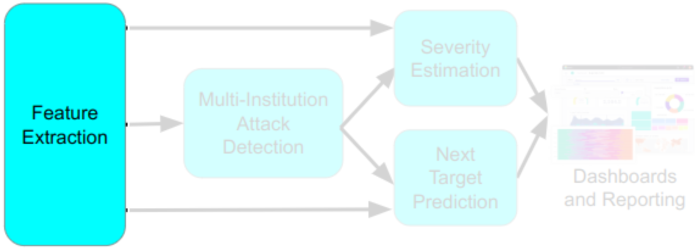


Solution: Graph compression without losing relevant information

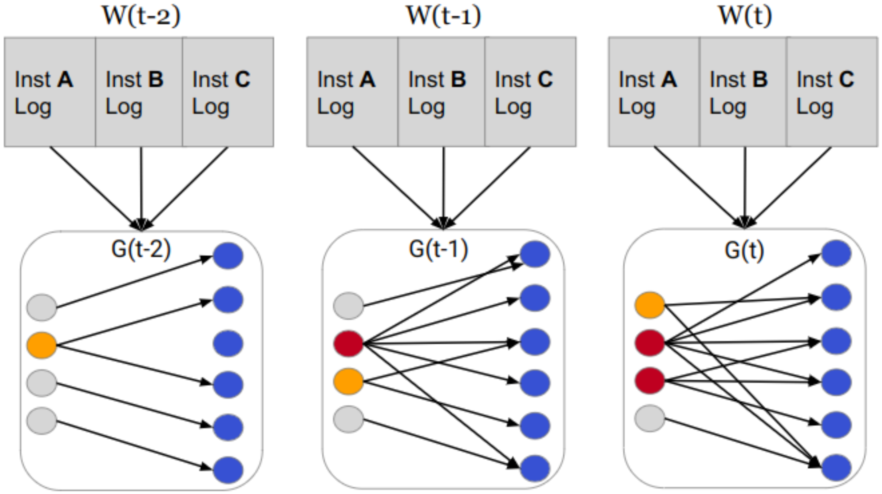
- Removing connections initiated internally
- Each educational institute's IPs clustered into a single vertices
- Directed aggregate edge
  - With weights



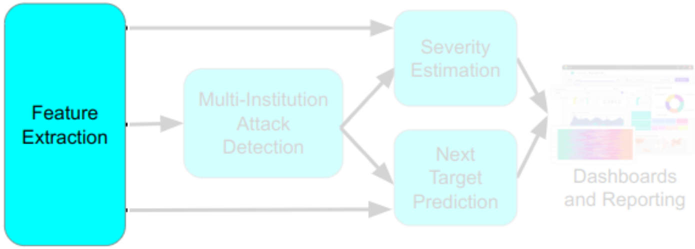
Graph Creation



- Collect logs by time windows
- Windows have two variables:
  - Window length
  - Number of windows
- Create graph for each window

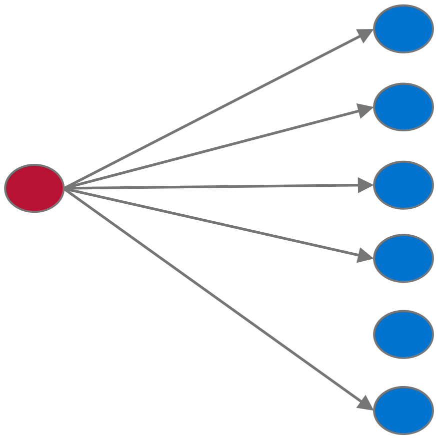


External IP Metrics

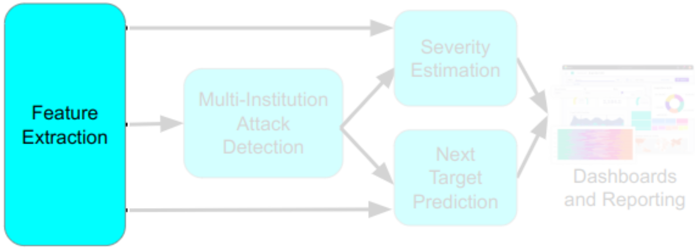


For each external IP in a time window:

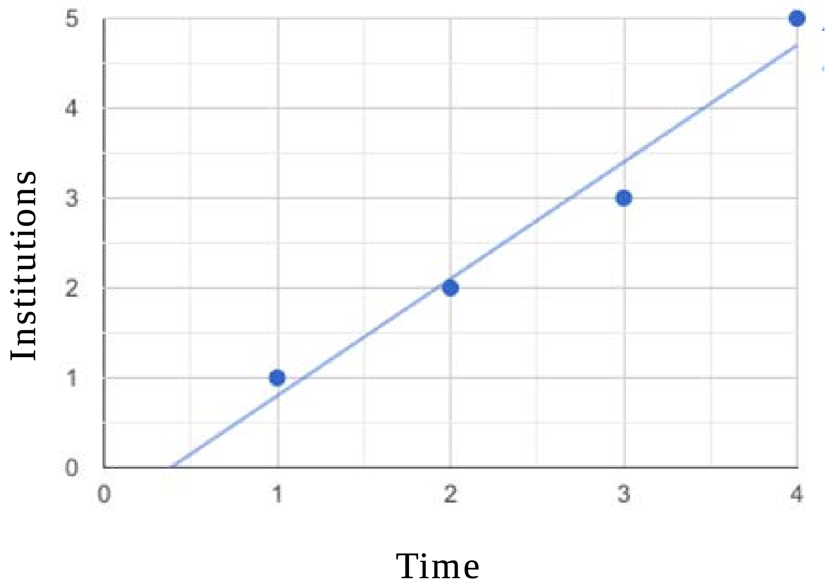
- inst_count : Number of institutes targeted
- ip_count: Number of IPs targeted
- conn_count: Number of connections attempted
- total_count: Total number of institutes targeted
- V(Adj): List of the institutions targeted in current window
- V(cumltv): Cumulative list of all institutions targeted until now



Dynamic Feature Extraction

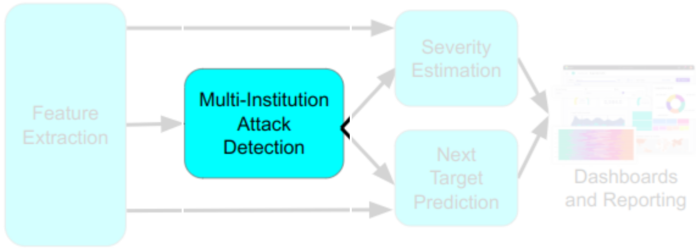


- For each metric
  - Capture growth across windows
- Use linear regression:
  - Predict an attack
  - Get growth metric

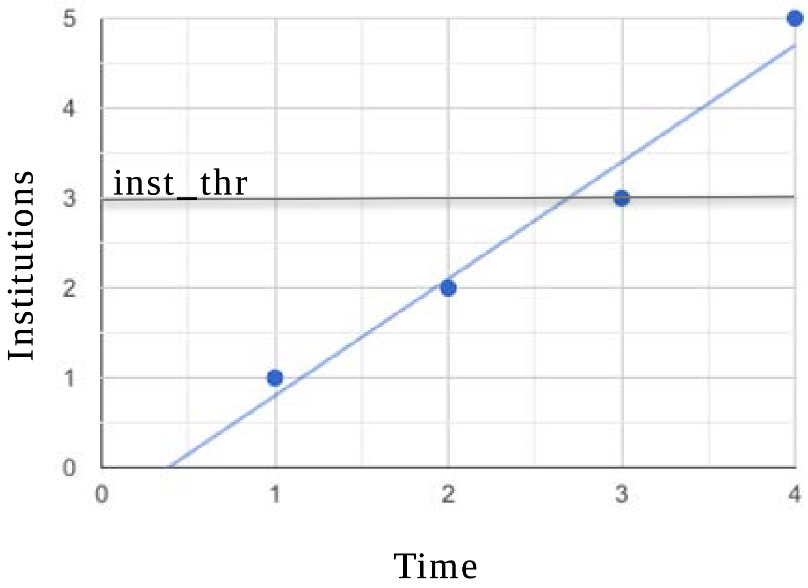


UNIVERSITY OF WATERLOO

Attack Detection

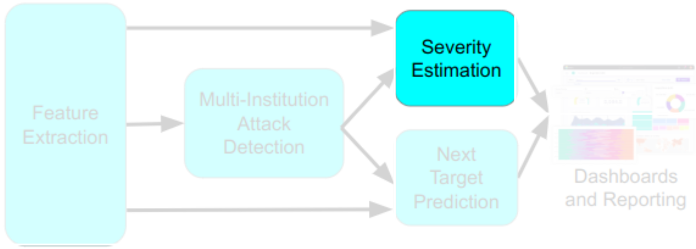


- An IP is identified as an attacker if its total_count exceeds a threshold (inst_thr)
- Predict an Attack:
  - If the Linear Regression line of total_count exceeds the inst_thr

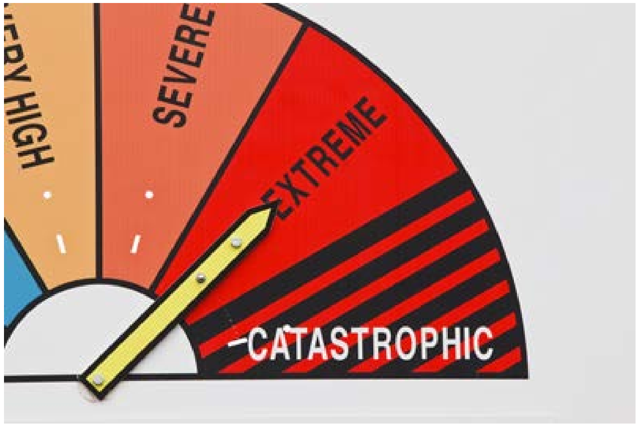


UNIVERSITY OF WATERLOO

Severity Estimation

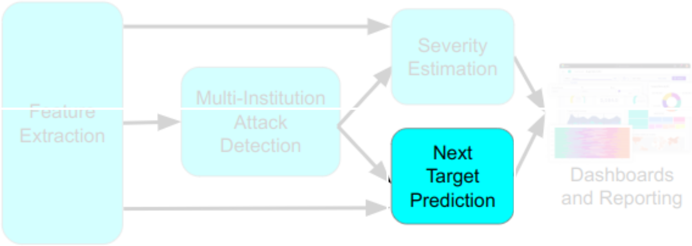


- Calculate a severity indicator in the range of [0,1]
  - Normalizing each feature in the range [0,1]
  - Robust scaling: to mitigate outliers stretching boundaries
- Sort these threats using severity indicator

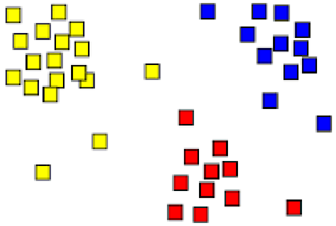


UNIVERSITY OF WATERLOO

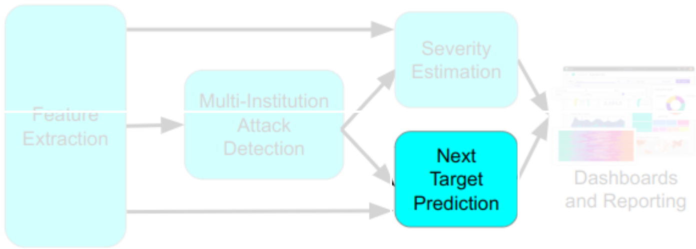
Next Target Prediction



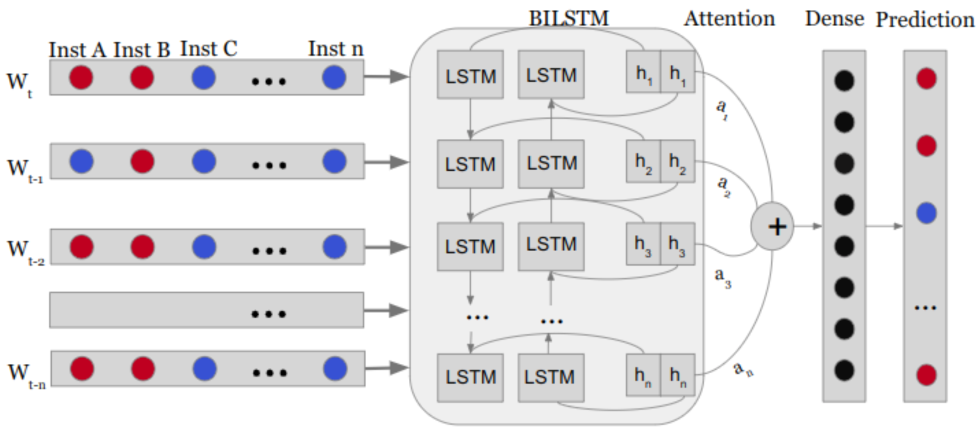
- Can we predict their path?
- Hypothesis:
  - Attackers follow a pattern in their movement.
    - Institute types are targeted together due to:
      - Service types
      - Security standards
      - Size of networks
      - etc...



Next Target Prediction



- Bidirectional LSTM with Attention
- Benefits:
  - Learns relationship between institutions
  - Arranges windows in sequence and learns attack sequence
    - In both directions
  - Captures growth or decline of attack



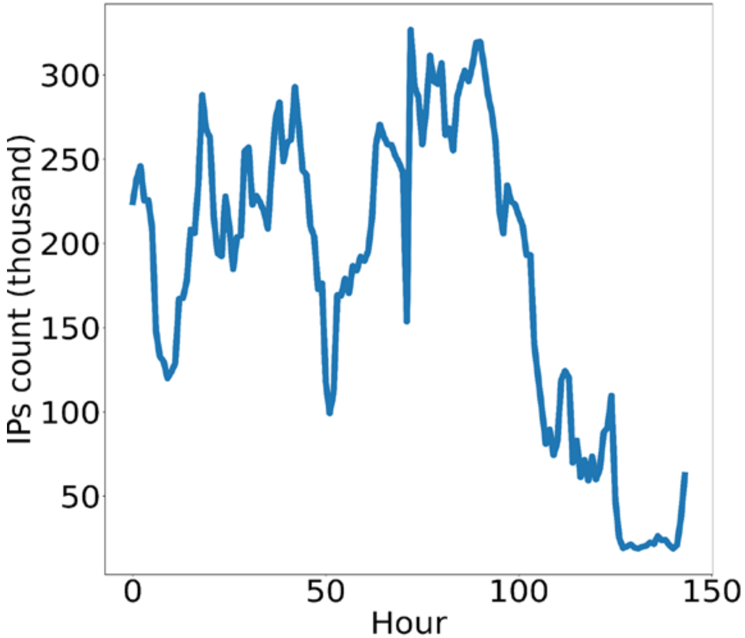
Outline

- Motivation and Introduction
- Soteria design
    - Feature Extraction
        - Static Metrics
        - Dynamic Metrics
    - Attack detection
    - Severity Estimation
    - Next Target Prediction
- Evaluation
- Conclusion

UNIVERSITY OF WATERLOO

Evaluation (Data Used)

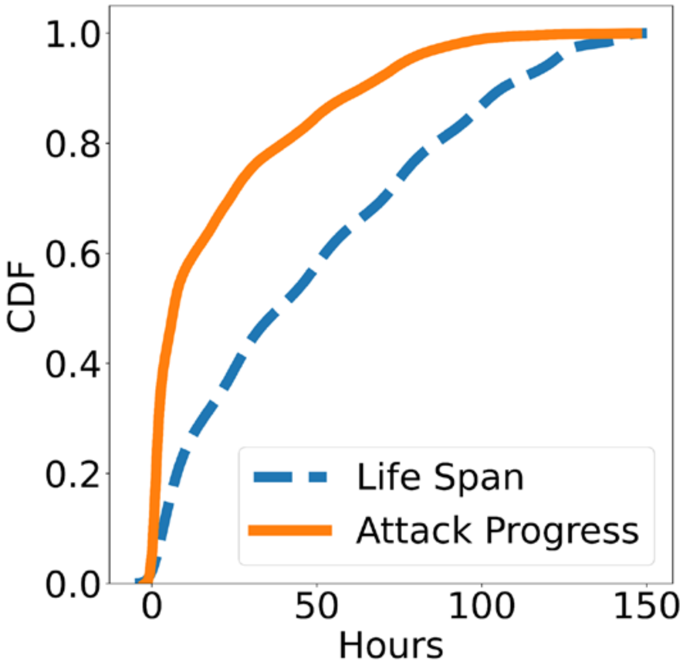
- 6 days of data
- 25th - 30th of Jan 2022
- 52 institutions
- 12 million external IPs
  - 2.7 million are attackers
- External IPs count



UNIVERSITY OF WATERLOO

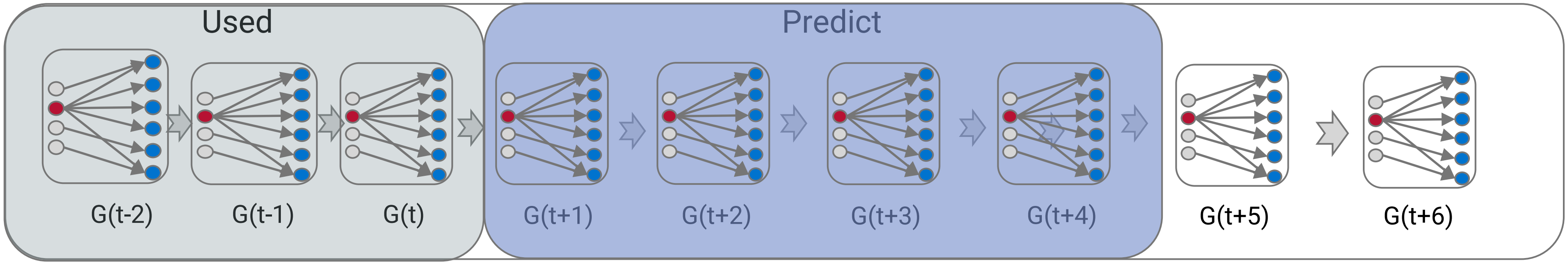
Evaluation (MIA Lifespan and attack progress)

- CDF of the MIA lifespan
  - 70% of attackers live 3 days or less
  - 50% live a day or less
- CDF of attack progress
  - When are they first targeted
    - Attacker contacts 70% of targets within the 1st day



UNIVERSITY OF WATERLOO

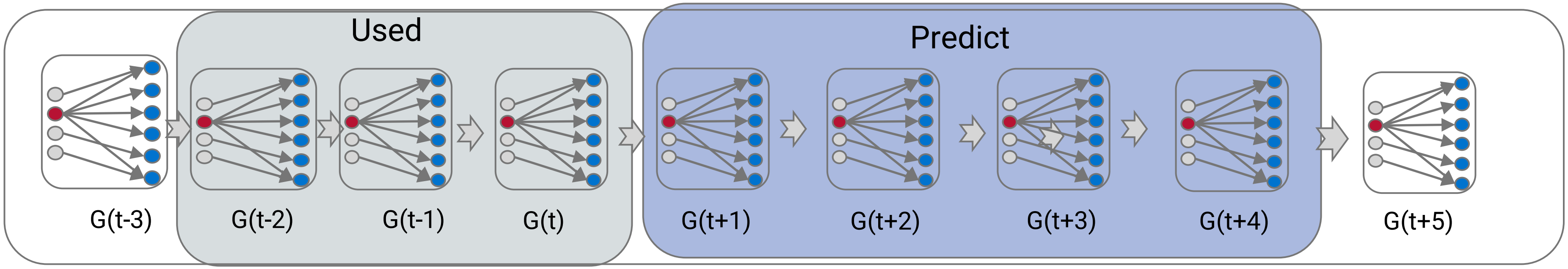
Evaluation (Life cycle of experiments)



- We divide time into windows example:
  - Window size is 6 hours

- We use 3 windows to predict attacks in the next 4 windows

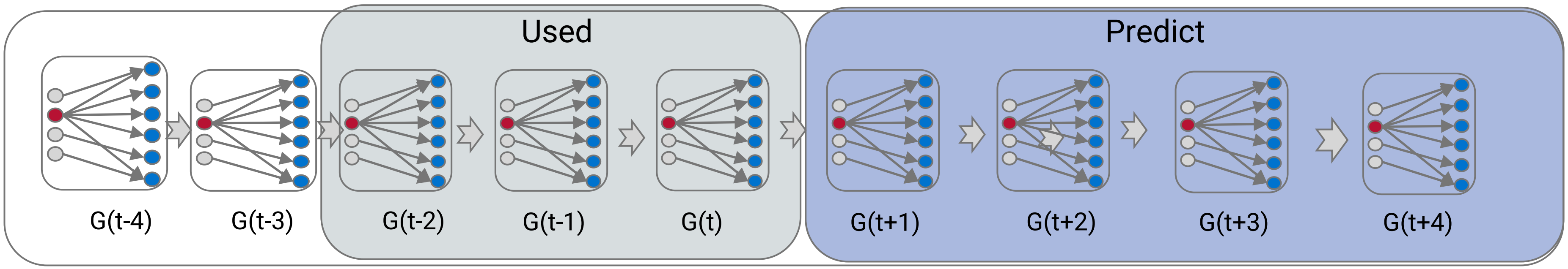
Evaluation (Life cycle of experiments)



- We divide time into windows example:
  - Window size is 6 hours

- We use 3 windows to predict attacks in the next 4 windows

Evaluation (Life cycle of experiments)



- We divide time into windows example:
  - Window size is 6 hours

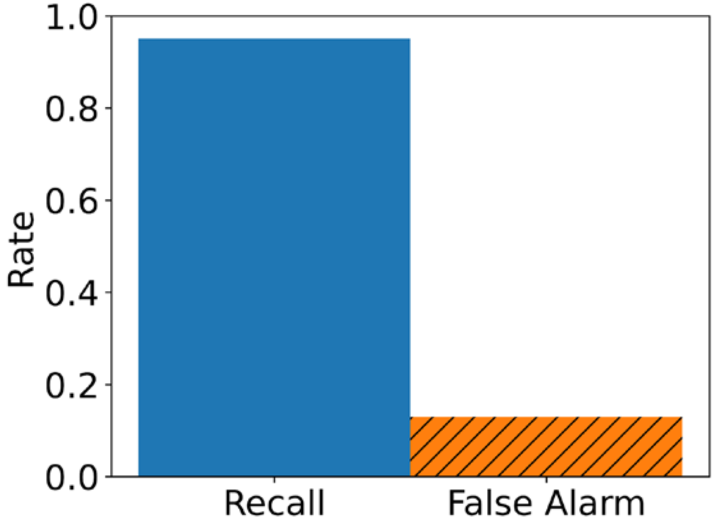
- We use 3 windows to predict attacks in the next 4 windows

Evaluation (Metrics)

- Metrics used:
  - Recall = True Positives / (True Positives + False Negatives)
  - False Alarm = False Positive /(True Negatives + False Positives)
- Aggregated results
  - We take the cumulative results of all the runs
  - In all runs, has the model been able to predict that an institution will be reached.

Evaluation (can it detect future Multi-institution attacks?)

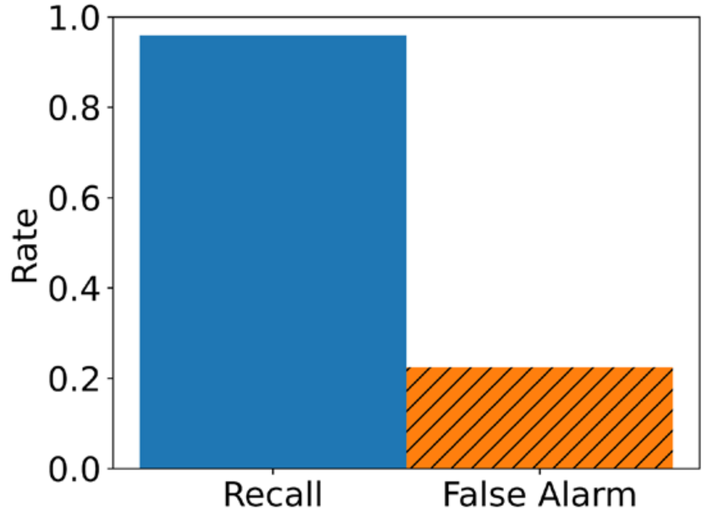
- 3 hour windows with 3 windows
- Soteria predict future attacks well:
  - Recall 95%
  - False alarm 15%



UNIVERSITY OF WATERLOO

Evaluation (can it find next target?)

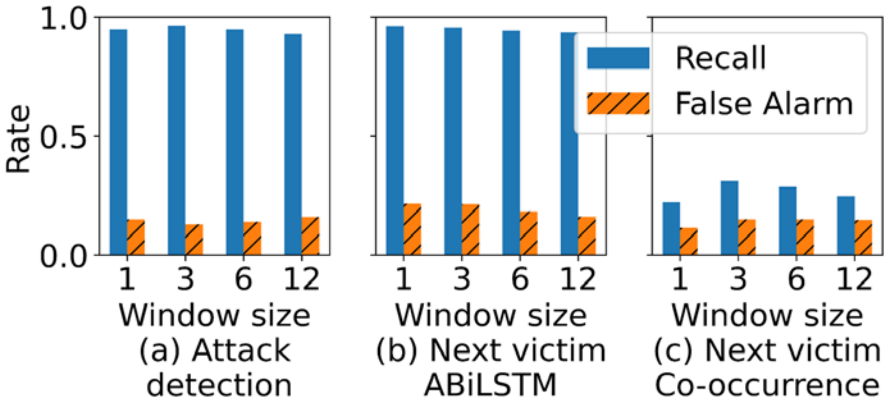
- 3 hour windows with 3 windows
- Soteria predicts effectively the next target:
  - Recall of 97%
  - False Alarm of 20%

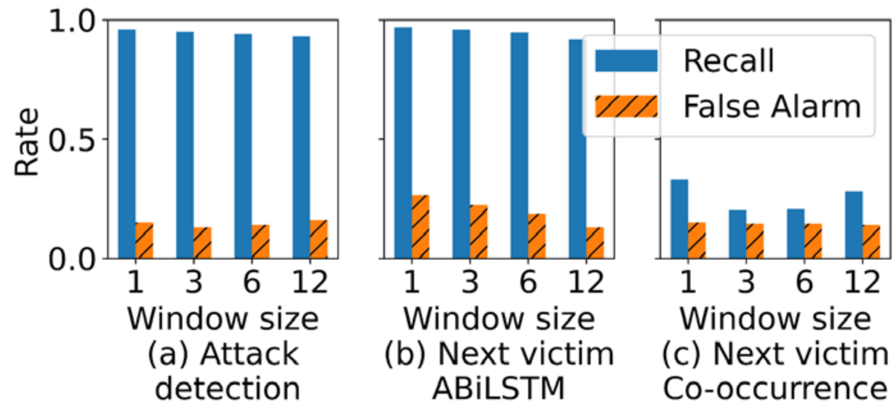


UNIVERSITY OF WATERLOO

Evaluation (Which window size and count is best?)

- Evaluated multiple window sizes
  - Fixed window count (3 windows)
  - Fixed the lookback time (24 hours)
- Slightly better with:
  - Smaller windows
  - Smaller number of windows

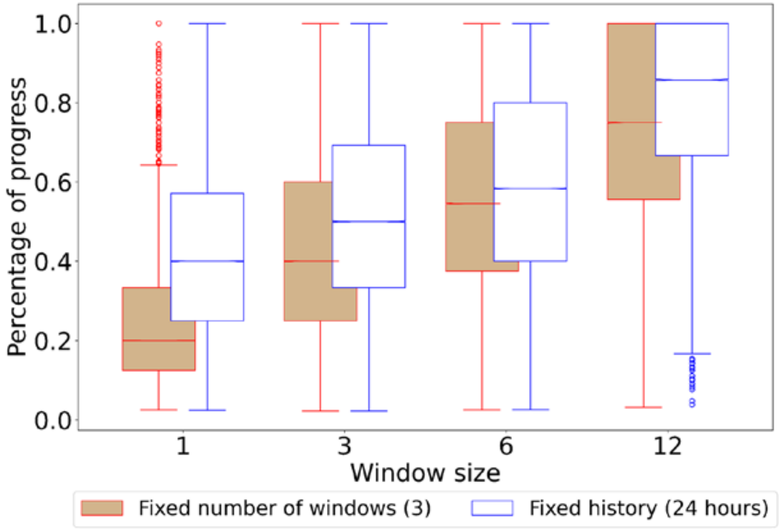




UNIVERSITY OF WATERLOO

Evaluation (How soon can Soteria predict an attack?)

- Evaluated using all the window size and count combinations used previously
- Soteria can predict an attack is happening at 20% progress
- Smaller windows with less windows predicts faster.
  - 1 hour windows provide up to 4x earlier detection



Insights

- External IPs contacting more than 2 institutions are most likely involved in an attack
- A simple linear regression model is highly effective in predicting future attack
- To accurately predict the next target of an attack we need to learn:
  - The relationships between institutions
  - The sequence of the attack
  - The level of activity of an attacker

UNIVERSITY OF WATERLOO

Conclusion

- Educational Institutions have huge networks and inadequate cyber security resources.
  - Attackers take advantage of this.
- Proposed model is able to:
  - Detect multi-institutional attacks
    - Current and future
    - Recall 95%
    - False Alarm 15%
  - Able to predict institutions targeted
    - Recall 97%
    - False Alarm 20%
- Currently deployed in the CANARIE IDS

UNIVERSITY OF WATERLOO