Anomaly Detection

Egemen K. Çetinkaya

Department of Electrical & Computer Engineering
Missouri University of Science and Technology

cetinkayae@mst.edu

http://web.mst.edu/~cetinkayae/teaching/CPE5420Fall2016
Anomaly Detection in Networks

Outline

• Introduction and motivation
• Anomaly detection taxonomy
• Detection in wireless networks
• Signal analysis of anomalies
• Conclusions and summary
Anomaly Detection in Networks

Introduction and Motivation

• Introduction and motivation
• Anomaly detection taxonomy
• Detection in wireless networks
• Signal analysis of anomalies
• Conclusions and summary
ResiliNets Strategy
$D^2R^2 + DR$

- **Real time control loop:** $D^2R^2$
  - defend against challenges to normal ops
    - passive
    - active
    - *detect* when defenses fail
  - remediate to do the best possible
  - recover to original state

- **Background loop:** DR
  - diagnose fault that lead to failure
  - refine future $D^2R^2$ behavior

[SHÇ+2010]
Anomaly Detection in Networks

Introduction and Motivation

- Detection is *required* once the defenses fail
- Threat model is essential for successful detection
- Primarily work in literature is about security
  - flash crowds don’t follow this
- Different domains need different detection algorithms
  - wired vs. wireless
- Several types of detection mechanisms exist
- Historical progress:
  - console monitoring for user activity
  - reviewing logs: tedious, takes long time, after the attack?
  - real-time systems: new attack recognition is not easy
Challenges to Normal Operation

Perturbations

- Unintentional misconfiguration, operational mistakes
  - random node or random link failures
- Large scale natural disasters (geo-correlated failures)
  - natural: hurricanes, tsunami, floods, earthquakes, etc.
  - man-made: fire, explosions, etc.
- Attacks from an intelligent adversary
- Environmental challenges
  - primarily wireless environments
- Unusual but legitimate traffic (e.g. flash crowd)
- Dependent failures
- Social, political, economical, and business factors
T1 Model
Survivability Attributes

\[ P(t) \]

\[ S_a + S_r \]

\[ S_u \] fraction unservable after failure at \( t_0 \)
  - \( S_a \) fraction available at \( t_0 \)

\[ S_r \] fraction restored at \( t_1 \)

full restoration at \( t_2 \)
• Correct operation
  – all events occur within expectations
  – $\forall e_i \in E \ \exists o_i \in O : \text{val}(o_i) = (t_i, t'_i)$
Anomaly Detection in Networks

Anomaly Detection Taxonomy

- Introduction and motivation
- Anomaly detection taxonomy
- Detection in wireless networks
- Signal analysis of anomalies
- Conclusions and summary
Anomaly Detection in Networks

**Detection Model**

- $\lambda = \langle M, D \rangle$
- Anomaly detection system: $\lambda$
- Model of normal behavior: $M$
- Degree of deviation: $D$

[TTV2004]
Anomaly Detection in Networks

Classification Model

- Learn a model – *classifier*
- From a set of labeled data instances – *training*
- Classify using learned model – *testing*
Anomaly Detection in Networks

Classification Boundaries

- Objects might be classified based on attributes: $x_1$ & $x_2$
- Linear classification finds a line between classes
- Classification boundaries may be non-linear

[BBK2014]
Anomaly Detection in Networks

Correct Detection of Events

Detection

Anomalous

Normal

Event Nature

Harmless event labeled normal

Attacks not detected, similar to normal behavior

Also known as false alarm

Successful detection

false positive

true positive

true negative

false negative

Also known as false alarm

Successful detection

Harmless event labeled normal

Attacks not detected, similar to normal behavior

false positive

true positive

true negative

false negative

[TTV2004]
Anomaly Detection Taxonomy

Overview

- Network feature analyzed
  - network traffic
    - flow analysis
    - protocol analysis
  - network elements and topology

- Behavior model
  - learnt models
  - specification-based models

- Analysis scale
  - microscale
  - mesoscale
  - macroscale
Anomaly Detection Taxonomy

Network Feature Analyzed

- Network traffic
  - flow analysis: utilizes temporal evolution of traffic flow
    - stochastic process and signal analysis are main methods
  - protocol analysis
    - data link
    - network
    - transport
    - application

- Network elements and topology
Anomaly Detection Taxonomy

Behavior Models

• Constructing the normal behavior of the system
  – by machine learning techniques
  – manually providing specifications

• Learnt models
  – statistical (e.g. estimators, Markov chains, etc.)
  – rule-based
  – artificial neural networks
  – high false alarm rate since data is trained

• Specification-based models (signature or misuse)
  – difficult, time consuming, only done once
  – low false alarm rate
Anomaly Detection Taxonomy

Analysis Scale

- Several dimensions of analysis exist
  - microscale
  - mesoscale
  - macroscale
- Functional perspective:
  - service, host, network
- Time dimension:
  - hourly, daily/weekly, seasonal
- Protocol analysis:
  - packets, streams, simultaneous analysis of connections
Anomaly Detection in Networks
Detection in Wireless Networks

- Introduction and motivation
- Anomaly detection taxonomy
- Detection in wireless networks
- Signal analysis of anomalies
- Conclusions and summary
IDS in Wireless Networks

Overview

• Mobile and wireless networks differ than wired nets
  – links susceptible to attacks
  – nodes do not have enough physical protection
  – dynamic topology

• IDS agents in each mobile host

• IDS agents can detect and decide locally

• Collaborative among the nodes
  – if node detects anomaly with weak evidence
  – detection state information is propagated to neighbor nodes
IDS in Wireless Networks
Detection State Information

• Use of level-of-confidence value: $p\%$
• Node A concludes from local data there is intrusion
• Node A concludes from local and neighbor states
• Node A, B, C, ... collectively concludes about intrusion
• Weights can be included in computation
  – nearby nodes can have more weight than distant nodes
IDS in Wireless Networks
Detection Procedure

• Node sends *intrusion state request* to neighbor
• Each node propagates state information
  – likelihood of intrusion to its immediate neighbors
• Each node determines
  – whether *majority* of reports indicate intrusion
  – if yes; node concludes network is under attack
• Any node detecting intrusion can initiate response
  – e.g. initiating re-authentication to exclude malicious nodes
IDS in Wireless Networks
Anomalies of MANET Routing Protocols

- Training of networks required
  - for simulation this entails aggregation of trace files

- Routing intrusion metrics:
  - percentage of changed routes
  - it can also include:
    - bad routes
    - stale routes
    - updated routes
  - percentage of changes in the sum of hops of all routes

- Traffic flow, routing activities, topological patterns
  - preferred in correlating change of information
Introduction and motivation
Anomaly detection taxonomy
Detection in wireless networks
Signal analysis of anomalies
Conclusions and summary
Signal Analysis of Anomalies

Overview

• Wavelets describes time series in frequency and time
  – useful for characterizing data with spikes and discontinues
  – Fourier transform is good for frequency analysis

• Following results show IP flow and SNMP data
  – flow is end-to-end association: by src/dst address, port
    • defines a set of MIB (Management Information Base)

• From Univ. of Wisconsin – Madison gateway router
  – data collected over 6 months
  – analysis includes high, medium, low band data
Signal Analysis of Anomalies

Ambient IP Flow Traffic

- Baseline IP flow traffic

[See BKPR2002 for details.]
Signal Analysis of Anomalies

Ambient SNMP Traffic

- Baseline SNMP traffic

[BKPR2002]
Signal Analysis of Anomalies

Byte Traffic for Flash Crowd

- Long-lived events can be captured by:
  - low-band and mid-band filters

[BKPR2002]
Signal Analysis of Anomalies
Average Packet Size for Flash Crowd

- Long-lived events can be captured by:
  - low-band and mid-band filters

[PKR2002]
Signal Analysis of Anomalies
Flow Traffic During DoS Attacks

- Short-lived events can be captured by:
  - high-band and mid-band filters

[BKPR2002]
• Short-lived events can be captured by:
  – high-band and mid-band filters
Anomaly Detection in Networks

Conclusions and Summary

• Introduction and motivation
• Anomaly detection taxonomy
• Detection in wireless networks
• Signal analysis of anomalies
• Conclusions and summary
Anomaly Detection Taxonomy
Open Challenges [PP2007]

• Mechanisms to keep up with the high-speed nets
• Reduction of false alarm rate: \( 1/10^5 \)
• Evaluation of detection mechanisms/algorithms
• Defending detection systems from attacks
  – attacks generating false alarms
• Better understanding of anomalies
  – taxonomy of challenges
• IDS and encryption does not work
• IDS for internal attacks
  – different access requirements for different users
Anomaly Detection Taxonomy

Conclusions and Summary

- Detection is essential part of a resilient network
- Anomaly detection primarily applied in security area
- Wired/wireless domains have different mechanisms
References and Further Reading


References and Further Reading


References and Further Reading

End of Foils