Introduction to Anomaly Detection
Standard CERT Disclaimer

NO WARRANTY

THIS MATERIAL OF CARNEGIE MELLON UNIVERSITY AND ITS SOFTWARE ENGINEERING INSTITUTE IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

Use of any trademarks in this presentation is not intended in any way to infringe on the rights of the trademark holder.

This Presentation may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

This work was created in the performance of Federal Government Contract Number FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center. The Government of the United States has a royalty-free government-purpose license to use, duplicate, or disclose the work, in whole or in part and in any manner, and to have or permit others to do so, for government purposes pursuant to the copyright license under the clause at 252.227-7013.
Where We Are Going

• Introduction, Definitions, and Usage
• Anomaly Collection and Classifications
• Anomaly Detection: Profiles & Attention Focusing
• Conclusion
Introduction

Assumption “Attacks exhibit characteristics that are different than those of normal traffic” (Denning, 1987).

Assumption validity
Introduction

Why do we care?

• In spite of following “best practices” vulnerabilities are still being discovered and exposed.

• Signature based solutions are failing miserably - new malware has < 10% detection rate by certain signature products.

• Fuzzing technologies make it easier for attackers to create their own 0 day attack. Fuzzing technologies work by automating the process of creative inputs, this in turn makes it easier for hackers to create their own 0 day attack.
Introduction

Why do we care?

• Anomaly detection provides an alternate approach than that of traditional intrusion detection systems. Jung et al., suggests modeling both normal and malicious behavior. (Jung, Paxson, Berger and Balakrishnan, 2004).
• Not all anomalies are malicious acts. (Sommer & Paxson, 2010)
• Most compelling reason: Anomalies have the potential to translate into significant critical and actionable information. (Chandola, Banerjee, & Kumar, 2009)
• AD is gaining popularity, this introductory presentation provides information and insight for deciphering the terms.
Introduction

The value of AD?

- AD represents an opportunity to see everything.
  - **Good:**
    - Capture 0 day attacks.
    - Define new analytics.
    - Gain a greater understanding of the network environment.
    - Proactive security posture.
    - Ability to better understand own environment.
    - Ability to complement existing solutions.
  - **Bad:**
    - Information overload.
    - Potential for improper use of models.
    - False positives are costly and incident handling is not easy nor automated.
    - Intrusion Detection has been shown to have fundamental differences from other areas where machine learning has been applied (Sommer & Paxson, 2010).
Definitions

Anomaly Definition

- A deviation from the norm; strange condition, situation or quality; an incongruity or inconsistency.
- Examples of network traffic anomalies:
  - IP address changes – New IP addresses appearing on sources and/or destinations found in logs.
  - Destination port changes – New destination ports showing up, especially combined with new destination addresses.
  - Command changes – sudden use of rarely used commands (e.g. Debug command, in HTTP or any other service).
  - Volume changes – sudden increases in service volume, destination volume.
  - Protocol anomalies – ssh over port 80, odd TCP flags, etc.
Definitions

Operational Profile

The operational profile of a system is defined as the set of operations that the software can execute along with the probability with which they will occur. An operation is a group of runs that typically involve similar processing (Lyu, 2002).

The Role of Profiles

Profiles are used to determine the norm, usual or expected behavior. They represent “baseline behavior”. More on how we obtain profiles when we discuss collection and classifiers.
Anomaly Detection Usage

Uses for Anomaly Detection

• Detect precedent attack behavior. (CERT 2010)
  — APT assistance.

• Zero day attack detection.
• Intrusion detection.
• Insider threat detection
• Situational awareness.
• Validate and assist with signature data.

Anomaly detection can be considered the thoughtful process of determining what is normal and what is not.
Where We Are Going

• Introduction, Definitions & Usage
• **Anomaly Collection and Classifications**
• Anomaly Detection: Profiles & Attention Focusing
• Conclusion
Anomaly Collection

Machine Learning

• Un-Supervised learning
  • Gather information on the network passively, determine normal, build profile, then set decision boundaries.
  • Collects and builds.
  • Fast collection increase time spent on categorization.

• Supervised learning
  • Uses training data in order to learn the environment.
  • Provides groupings of learned categories.

Regardless of the learning method, the operational profile is the result of this step.
Anomaly Classification

Classifiers (decision support for uncertainty)

- Classifiers provide ways to organize the data.
- Commonly referenced models in anomaly classification:
  - Decision Tree
  - Bayes
  - Fuzzy
  - Certain types of clusters*

* Note: The term “Certain types of clusters” is marked with an asterisk, indicating it might be a specific type or context relevant to anomaly classification.
Where We Are Going

• Introduction. Definitions & Usage
• Anomaly Detection Usage
• Anomaly Collection and Classifications
• **Anomaly Detection: Profiles & Attention Focusing**
• Conclusion
Operational Profile Candidates

Here are a few candidates for operational profiles:

- **Netflow (using SiLK names for fields)**
  - sIP, dIP, sPort, dPort, pro, packets, bytes, flags, sTime, dur, eTime, sen, in, out, nhIP, scc, dcc, cla, type, sTime+msec, eTime+msec, dur+msec, iTy, iCo, initialF, sessionF, attribut, appli

- **External Data Sources**
  - DNS, ASN, WHOIS, GeoIP, blacklists, reputation

- **Full Packet Data and Logs**
  - IDS alerts, extracted URLs, extracted DNS responses, authentication logs, email headers, AV data…

- **Application data, User behavior, Policy Violations**

- **Combinations of any of the above**
Example Operational Profile

Outbound Bytes per Port

Operational Profile

- Statistical breakdown of outbound calls by service (proto+port)
- First data below shows top 13 services, %99.87 of all bytes
- Second data below shows bottom services. The interesting things are often in the noise.
- Graph shows the first data set.
Example Operational Profile
Drilling down on one service: 443

Operational Profile
We know what services are normal, now we must find what is normal for the services.

- Drill down on outbound port 443
- Look at total bytes to destinations
- First data below shows top dests
- Second data shows bottom dests
- Graphic shows first data.
- Caveat: this data is cooked for the slides. There are inconsistencies.

<table>
<thead>
<tr>
<th>port</th>
<th>dIP</th>
<th>dPort</th>
<th>Bytes</th>
<th>%Bytes</th>
<th>cumul_%</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>74.125.228.105</td>
<td>443</td>
<td>99184312</td>
<td>20.571027</td>
<td>20.571027</td>
</tr>
<tr>
<td>6</td>
<td>74.125.228.117</td>
<td>443</td>
<td>60136869</td>
<td>12.472508</td>
<td>33.043535</td>
</tr>
<tr>
<td>6</td>
<td>74.125.228.118</td>
<td>443</td>
<td>41171738</td>
<td>8.539102</td>
<td>41.582637</td>
</tr>
<tr>
<td>17</td>
<td>128.237.28.12</td>
<td>443</td>
<td>28945131</td>
<td>6.003279</td>
<td>47.585916</td>
</tr>
<tr>
<td>6</td>
<td>69.171.248.16</td>
<td>443</td>
<td>16700996</td>
<td>3.463802</td>
<td>51.049718</td>
</tr>
<tr>
<td>6</td>
<td>69.171.229.16</td>
<td>443</td>
<td>14444553</td>
<td>2.995830</td>
<td>54.045547</td>
</tr>
<tr>
<td>6</td>
<td>74.125.228.101</td>
<td>443</td>
<td>14141718</td>
<td>2.933021</td>
<td>56.978568</td>
</tr>
<tr>
<td>6</td>
<td>69.171.224.32</td>
<td>443</td>
<td>10705577</td>
<td>2.220358</td>
<td>59.198926</td>
</tr>
<tr>
<td>6</td>
<td>69.171.237.16</td>
<td>443</td>
<td>10206146</td>
<td>2.116774</td>
<td>61.315781</td>
</tr>
<tr>
<td>6</td>
<td>69.171.237.32</td>
<td>443</td>
<td>9759443</td>
<td>2.024128</td>
<td>63.339829</td>
</tr>
</tbody>
</table>
Where We Are Going

- Introduction. Definitions & Usage
- Anomaly Detection Usage
- Anomaly Collection and Classifications
- Anomaly Detection: Profiles & Attention Focusing
- Conclusion
Conclusion

AD is gaining in popularity.
There are many different components of AD and the ones discussed represent only a portion, not a complete picture.
Understanding how the profile is built and what it represents is vital to understanding how the results were obtained.
It is important to how attention focusing is being directed.
References


References


Thank you!

Char Sample & George Jones
CERT/NetSA
October 2012