Dealing with False Positives in Intrusion Detection

Klaus Julisch, ZRL

kju@zurich.ibm.com
Presentation Overview

- Introduction & Problem Statement
- Characteristics of IDS alarms
- Filtering False Positives
- Problems with Filtering
- Benefits of Filtering -- Looking at What's Left Over
- Conclusion & Open Problems
Introduction & Problem Statement

Problems:
- 3 alarms/minute
- >99% false positives
Presentation Overview

- Introduction & Problem Statement
- Characteristics of IDS alarms
- Filtering False Positives
- Problems with Filtering
- Benefits of Filtering -- Looking at What's Left Over
- Conclusion & Open Problems
## Characteristics of IDS Alarms

- **Key observation: 5 / 90 Rule**

### Alarm frequencies for five different sensors in April

<table>
<thead>
<tr>
<th></th>
<th>big.gov</th>
<th>small.gov</th>
<th>bank</th>
<th>transport</th>
<th>insurance</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 (81.1%)</td>
<td>153 (83.9%)</td>
<td>47 (43.4%)</td>
<td>15 (50.6%)</td>
<td>298 (41.1%)</td>
<td></td>
</tr>
<tr>
<td>66 (7.3%)</td>
<td>47 (7.5%)</td>
<td>56 (15.0%)</td>
<td>80 (20.1%)</td>
<td>55 (18.6%)</td>
<td></td>
</tr>
<tr>
<td>9 (6.7%)</td>
<td>80 (3.8%)</td>
<td>80 (12.8%)</td>
<td>47 (12.7%)</td>
<td>119 (13.5%)</td>
<td></td>
</tr>
<tr>
<td>47 (1.2%)</td>
<td>15 (3.2%)</td>
<td>15 (8.1%)</td>
<td>29 (6.2%)</td>
<td>83 (8.8%)</td>
<td></td>
</tr>
<tr>
<td>29 (1.2%)</td>
<td>35 (0.5%)</td>
<td>66 (7.1%)</td>
<td>321 (1.8%)</td>
<td>349 (5.4%)</td>
<td></td>
</tr>
</tbody>
</table>
Presentation Overview

- Introduction & Problem Statement
- Characteristics of IDS alarms
- Filtering False Positives
- Problems with Filtering
- Benefits of Filtering -- Looking at What's Left Over
- Conclusion & Open Problems
Filtering False Positives (1/4)

Filtering by Means of NetRanger's Context Field

NetRanger’s Alarm Format

<table>
<thead>
<tr>
<th>RecId</th>
<th>Time</th>
<th>SrcIp</th>
<th>SrcPrt</th>
<th>Context</th>
</tr>
</thead>
</table>

- Idea: Use the context field to verify NetRanger's analysis
Filtering False Positives (2/4)

Filtering by Means of NetRanger's Context Field:

Alarm 80

IIS WEB Server

Problem
big.gov has 126860 type-80 alarms in April!

Black Hat  Get /script.asp%2E?
Filtering False Positives (3/4)

Filtering by Means of Alarm Patterns

- **Key idea:**
  - Find alarm patterns,
  - understand their root cause, and
  - if non-malicious, use the alarm patterns for filtering

- **Examples**
  - Host triggers alarm 120 every 30 minutes
  - Host triggers alarm 65 every \( n \times 11 \) minutes
  - Fragmented IP originating weekdays from WEB
Filtering False Positives (4/4)

Fragmented IP Originating Weekdays from WEB Server
Presentation Overview

- Introduction & Problem Statement
- Characteristics of IDS alarms
- Filtering False Positives
- Problems with Filtering
- Benefits of Filtering -- Looking at What's Left Over
- Conclusion & Open Problems
Problems with Filtering (1/4)

Overview

- Finding filtering rules
- Time constancy of filtering rules

- Other problems:
  - Genericity of filtering rules
  - Risk of filtering out true positives
Problems with Filtering (2/4)

Finding Filtering Rules

- Not all IDSs provide a "context field"
- NetRanger doesn't provide the context for all alarms
- Finally, how do you find usable patterns
  - either in the context field, or
  - in the alarm stream?
Problems with Filtering (3/4)

Time Constancy of Filtering Rules

- Two aspects of time constancy
  - Effectiveness of filtering rules over time
    - Alarm 80: 100% - 100% - 100% - 100%
    - Alarm 9: 90% - 79% - 72% - 73%
  - Justifiability
    - The most frequent alarms of one month might become insignificant in following months!
# Problems with Filtering (4/4)

**Time Constancy of Filtering Rules:**  
Justifiability

<table>
<thead>
<tr>
<th>Month</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>153 (98.9%)</td>
<td>153 (83.9%)</td>
<td>47 (35.6%)</td>
<td>15 (41.4%)</td>
</tr>
<tr>
<td></td>
<td>15 (0.4%)</td>
<td>47 (7.5%)</td>
<td>153 (27.2%)</td>
<td>47 (33.6%)</td>
</tr>
<tr>
<td></td>
<td>47 (0.4%)</td>
<td>80 (3.8%)</td>
<td>15 (11.5 %)</td>
<td>85 (14.9%)</td>
</tr>
<tr>
<td></td>
<td>80 (0.2%)</td>
<td>15 (3.2%)</td>
<td>85 (11.4 %)</td>
<td>334 (3.1%)</td>
</tr>
<tr>
<td></td>
<td>75 (&lt;0.1%)</td>
<td>35 (0.5%)</td>
<td>138 (4.8 %)</td>
<td>29 (1.1 %)</td>
</tr>
</tbody>
</table>

Development of alarm frequency for small.gov
Presentation Overview

- Introduction & Problem Statement
- Characteristics of IDS alarms
- Filtering False Positives
- Problems with Filtering
- Benefits of Filtering -- Looking at What's Left Over
- Conclusion & Open Problems
Benefits of Filtering (1/3)

The Trace of an Attack Tool

36
burst of \{20, 34, 110, 111, 115\}
burst of \{121, 123, 124\}
burst of 133
burst of 138
...
Pause
Benefits of Filtering (2/3)

Coordinated Attack: Multiple Sources Attacking one Target

Start time: April 13, 00:35

Source 1
Source 2
Source 3

synchronized alarms
Benefits of Filtering (3/3)

Opportunistic Attack: Wide Scan

Time Period: 3 days

evil.com → One probe/target → A.net1 → B.net2 → ...
Presentation Overview

- Introduction & Problem Statement
- Characteristics of IDS alarms
- Filtering False Positives
- Problems with Filtering
- Benefits of Filtering -- Looking at What's Left Over
- Conclusion & Open Problems
Conclusion & Open Problems

- Filtering is very powerful, but comes at a cost
- Based on 7 sensors over 4 months (almost 1 GB):
  - Removing 60% -- 80% of alarms is possible
  - at the cost of developing 1 -- 2 new filters per month

- Main problem: Finding filters efficiently
  - First experiments with visualization and data mining look very promising