Introduction to Anomaly Detection

Char Sample <csample@cert.org>
George Jones <gmj@cert.org>
CERT/NetSA
FloCon 2013

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*

Where We Are Going

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

Operational Profile Candidates

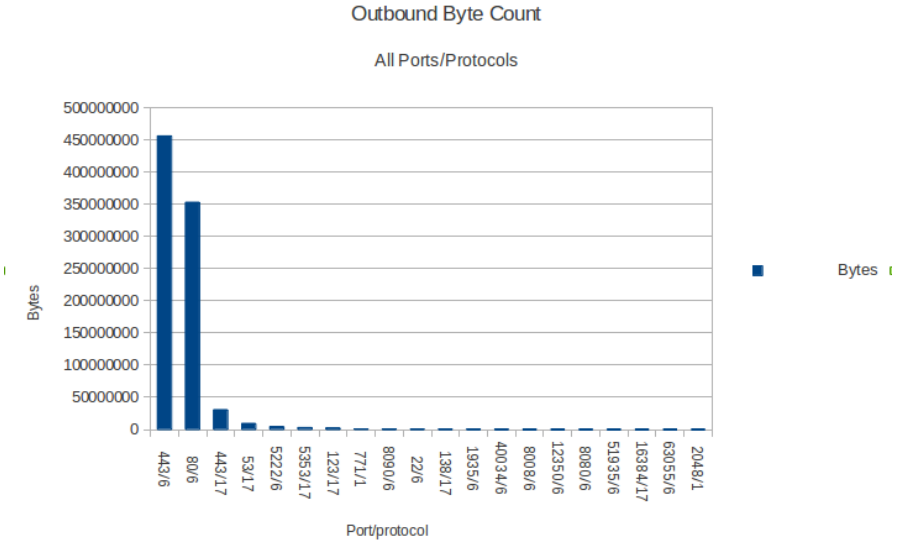
Here are a few candidates for operational profiles:

- 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.



pro	dPort	Bytes	%Bytes	cumul_%
6	443	452535897	52.718420	52.718420
6	80	353117060	41.136567	93.854987
17	443	29619475	3.450537	97.305524
17	53	9202986	1.072107	98.377631
6	5222	4347436	0.506457	98.884088
17	5353	3030322	0.353019	99.237107
17	123	2416876	0.281555	99.518662
1	771	982035	0.114403	99.633065
6	8090	642693	0.074871	99.707936
6	22	531313	0.061896	99.769831
17	138	415305	0.048381	99.818213
6	1935	303337	0.035337	99.853550
6	40034	223559	0.026044	99.879594

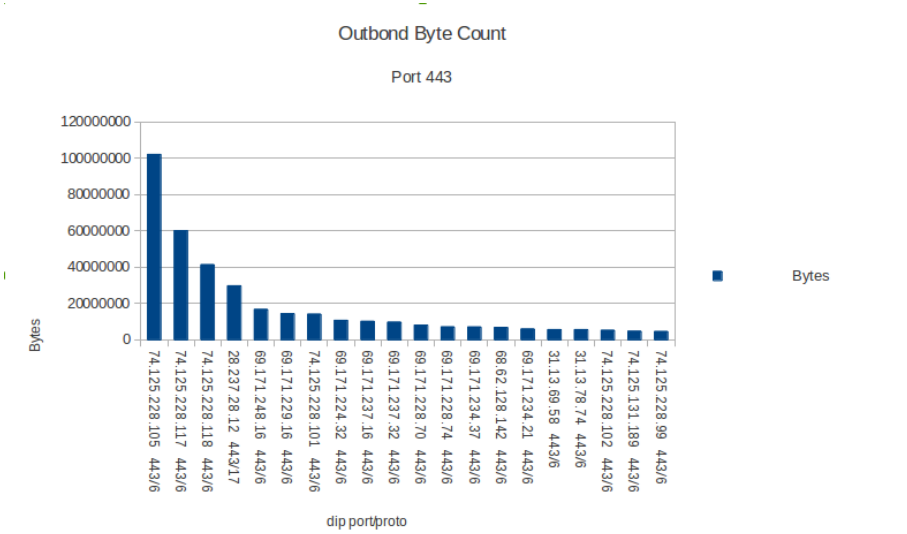
pro	dPort	Bytes	%Bytes	cumul_%
6	1935	303337	5039	99.858457
6	40034	208713	3424	99.880346
6	8080	192331	873	99.900517
6	8008	146607	853	99.915893
6	12350	110309	1830	99.927462
6	993	75351	1171	99.935365
6	51935	74724	1437	99.943202
17	16384	70248	104	99.950569
6	63055	58396	1123	99.956693

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.



pro	dIP	dPort	Bytes	%Bytes	cumul_%
6	74.125.228.105	443	99184312	20.571027	20.571027
6	74.125.228.117	443	60136869	12.472508	33.043535
6	74.125.228.118	443	41171738	8.539102	41.582637
17	128.237.28.12	443	28945131	6.003279	47.585916
6	69.171.248.16	443	16700906	3.463802	51.049718
6	69.171.229.16	443	14444553	2.995830	54.045547
6	74.125.228.101	443	14141718	2.933021	56.978568
6	69.171.224.32	443	10705577	2.220358	59.198926
6	69.171.237.16	443	10206140	2.116774	61.315701
6	69.171.237.32	443	9759443	2.024128	63.339829

pro	dIP	dPort	Bytes	%Bytes	cumul_%
17	157.55.235.146	443	46	0.000008	99.999943
17	111.221.77.154	443	46	0.000008	99.999951
17	111.221.77.162	443	46	0.000008	99.999959
17	157.55.235.148	443	46	0.000008	99.999968
17	157.55.235.161	443	46	0.000008	99.999976
17	64.4.23.140	443	46	0.000008	99.999984
17	111.221.74.16	443	46	0.000008	99.999992
17	111.221.77.152	443	46	0.000008	100.000000

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

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, *41*(3), 15.

Denning, D.E., 1987, "An intrusion-detection model", *Software Engineering IEEE Transactions on,* (2):222-232.

Jung, J., Paxson, V., Berger, A.W., and Balakrishman, H., 2004, "Fast Portscan Detection Using Sequential Hypothesis Testing", *Security and Privacy 2004. Proceedings 2004 IEEE Symposium on* (pp. 211-225). IEEE.

References

Lyu, M., 2002, “Software Reliability Theory”, *Encyclopedia of Software Engineering.*

Sommer, R., & Paxson, V., (2010), “Outside the closed world: On using machine learning for network intrusion detection.”, In *Security and Privacy (SP), 2010 IEEE Symposium on,* pp. 305-316, IEEE, 2010

Thank you!

Char Sample & George Jones
CERT/NetSA
October 2012