Classifying Encrypted Traffic with TLS-aware Telemetry

Blake Anderson, David McGrew, and Alison Kendler
blaander@cisco.com, mcgrew@cisco.com, alkendle@cisco.com

FloCon 2016
Problem Statement

- “I need to understand traffic even when it is encrypted”
  - Malware detection
  - Application identification
- “I need to understand how crypto is being used on my network”
  - Weak crypto algorithms and/or key sizes
  - Vulnerable cryptographic library detection
  - The ports where TLS shows up

![Non-443 Malicious TLS](chart.png)
Solution

• Our solution is to gather additional, TLS-aware telemetry.

• This solution:
  • Could be baked into a flow telemetry exporting device
  • Can be run in a VM off a SPAN port (with our open source package)

• Passive monitoring is used to gather all data.
  • Not costly or difficult to deploy (as opposed to MITM solution)
TLS-aware Telemetry Data Types

Client

- Client Hello
- Server Hello / Certificate
- Client Key Exchange / Change Cipher Spec
- Change Cipher Spec
- Application Data

Server

- TLS Version, Offered Ciphersuites, TLS Extensions
- Selected Ciphersuite
- Client Key Length
- Sequence of Record Lengths, Times, and Types
Malware Detection
Malware Detection

• Malware is making use of TLS to communicate.
  • We observed that 7-13% of malware communication is over TLS

• Traditional IPS/IDS signatures fail.
  • Malicious communication is encrypted

• We leverage TLS-aware telemetry for malware classification.
  • Increases classification accuracy
  • Reduces false positives
Ciphersuites

Offered Ciphersuites

Selected Ciphersuites

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential
TLS Extensions

Percentage of TLS Flows that Use Each Extension

TLS Extension Hex Code

DMZ
Malware
Client Key Lengths

Client Key Length

Key Size (bits)

DMZ
Malware

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential
Test Setup

- **Malware**
  - September 2015 pcaps from ThreatGRID
  - TLS (443) traffic, > 100 in and out bytes
  - 26,404 flows, Telemetry enhanced with TLS extensions, ciphersuites, and client key lengths

- **Benign**
  - traffic taken from a large enterprise DMZ
  - TLS (443) traffic, > 100 in and out bytes
  - 50,848 flows, Telemetry enhanced with TLS extensions, ciphersuites, and client key lengths

- 10-fold CV
Telemetry Data Types

- SPLT – Sequence of Packet Lengths and Arrival Times
  
- Byte Distribution
  - Relative frequency for each byte in a flow

- traditional: sp, dp, prot, ib, ip, ob, op, dur
Results

- L1-logistic regression
- SPLT + 7-tuple + BD

- L1-logistic regression
- SPLT + 7-tuple + BD + TLS
Results

- L1-logistic regression
- SPLT + 7-tuple + BD
  - 172.2 non-zero parameters
  - 0.01 FDR: 0.1%
  - Total Accuracy: 96.1%

- L1-logistic regression
- SPLT + 7-tuple + BD + TLS
  - 138.1 non-zero parameters
  - 0.01 FDR: 90.4%
  - Total Accuracy: 99.7%
Crypto Audit
Crypto Audit

• We observe what cryptography is being used in TLS (same principles can be applied to SSH, IPsec, etc.).
  • Who is using weak crypto on my network?

• We infer the version of the cryptographic library in use.
  • Initial results with OpenSSL
  • Vulnerable implementations, not active attacks

• We passively monitor traffic, no active probing.
Client Key Lengths (DMZ)

![Graph showing TLS Client Key Lengths with key sizes in bits.]
Selected Ciphersuites (DMZ)
OpenSSL Similarity Matrix
TLS Extensions

TLS Extensions by Default
Heartbleed

TLS pad extension to fix TLS hang bug
logjam

Removed the export ciphers from the DEFAULT ciphers
Benefits of TLS-awareness

• TLS-aware telemetry provides a passive monitoring approach for:
  • Improved malware classification
  • The ability to audit an enterprise network’s crypto usage

• TLS-aware telemetry is a relatively lightweight system compared to MITM solutions or full packet capture.

• joy (our open source package) currently implements the described functionality (https://github.com/davidmcgrew/joy).
Thank You