SocietyByte
Wissenschaftsmagazin des BFH-Zentrums Digital Society
https://www.societybyte.swiss

Mit Netzwerkdaten Angriffe erkennen und analysieren

Autor : Reto Inversini

Datum : 21. März 2019

Angreifer hinterlassen bei den meisten Attacken Spuren im Netzwerk. Diese können mit Hilfe einer Vielzahl von Technologien und Werkzeugen analysiert werden. Unser Autor arbeitet im Bereich Incident Response und schildert in diesem Beitrag konkret, wie er Angriffe analysiert.

Einleitung

Systeme werden immer stärker miteinander verbunden, sowohl auf einer globalen Ebene durch das Internet wie auch auf einer lokalen Ebene, innerhalb einer Firma oder eines Haushaltes. Durch die stärkere Vernetzung sind viele neue Angriffsvektoren entstanden und Angreifer können sich in den legitimen Datenströmen verstecken. Die Kunst der Detektion auf Netzwerkebene ist es, die Spuren der Angreifer zu entdecken und genügend Beweise zu sichern, um zu verstehen, wie der Eindringling vorgegangen ist.

Netzwerkdetektion und -forensik sind eng verwandt und unterscheiden sich primär durch die

Ziele und den Zeitpunkt, an dem sie gemacht werden, aber weniger im Bereich der eingesetzten Technologien. Die Detektion ist das tägliche Geschäft von CERTs (Computer Emergency Response Teams). Das Ziel ist es, Angriffe möglichst früh zu erkennen und entsprechende Gegenmassnahmen einzuleiten. Daraus ergibt sich, dass die Geschwindigkeit, aber auch die Automatisierbarkeit der Suche ein entscheidender Faktor ist. Forensik ist "post-mortem" und versucht den Angriff möglichst lückenlos zu dokumentieren, sei dies im Rahmen eines Strafverfahrens, sei dies für die interne Aufarbeitung innerhalb einer Firma. Ein besonderes Augenmerk muss dabei auf die Beweiskette gelegt werden, damit keine Zweifel an der Gültigkeit der Beweismittel aufkommen können.

Netzwerkforensik ergänzt die klassische Disk- und Memoryforensik sehr gut. Es ist möglich, Netzwerkforensik zu betreiben, ohne dem Täter einen Hinweis zu geben, dass er überwacht wird. Da Memory- und Diskforensik oft direkten Zugang zum System des Angreifers benötigen, kann Netzwerkforensik bereits Daten sicherstellen ohne den Täter zu alarmieren. Da die allermeisten Geräte in irgendeiner Form mit einem Netzwerk verbunden sind, ist Netzwerkforensik häufig am besten geeignet, das Ausmass einer Infiltration festzustellen und die Kommunikation des Angreifers zu erkennen.

Ablauf

Analysen auf Netzwerkebene bestehen meist aus folgenden Prozessschritten:

Traffic Capturing: Dies beinhaltet das – möglichst verlustfreie – Aufzeichnen des Netzwerkverkehrs. Die grosse Herausforderung dabei sind die immer höheren Bandbreiten. So fallen beim Sniffen einer gut ausgelasteten 1GB Leitung in einem Tag ca. 10 TB an Daten an. Capturing kann dabei mit verschiedenen Methoden realisiert werden. In der Regel wird dazu entweder ein Netzwerktap eingesetzt oder der Traffic wird auf Switch Ebene von einem Switch Port auf einen anderen gespiegelt (Mirroring-Port). Ein Netzwerktap ist ein Gerät, welches inline zwischen zwei Geräten eingehängt wird und den Traffic dupliziert. Idealerweise ist das Tap fähig, transmit (Tx) und receive (Rx) Verkehr auf einen Port zu aggregieren.

Preprocessing / Normalisierung: Computernetzwerke werden als "Packet Switched" bezeichnet, d.h., dass Nachrichten meist in mehrere Datenpakete aufgeteilt werden und vom Zielrechner wieder zusammengesetzt werden müssen. Wird der Traffic aufgezeichnet, muss dieselbe Aufgabe ebenfalls wahrgenommen werden. Die Pakete müssen zu den ursprünglichen Nachrichten zusammengesetzt werden. Um möglichst einfach und effizient eine Auswertung oder eine Suche nach bestimmten Elementen zu machen, hilft es, wenn die Auswertungssoftware auch die verwendeten Protokolle (z.B. HTTP) kennt und aufbereiten

kann. Dieser Teil ist das Preprocessing.

Analyse: Dies ist die eigentliche Kernaufgabe und damit auch am Aufwändigsten. Die Analyse von Netzwerkverkehr beinhaltet die Suche nach Spuren des Angreifers oder des Täters. Je nach Fragestellung kommen dabei unterschiedliche Ansätze zum Zuge:

- Extraktion von Dateien (z.B. von Bildern) aus dem Datenstrom
- Suche nach Angriffsmustern (z.B. SQL Injections)
- Suche nach Kontrollverbindungen von Malware
- Statistische Zusammenstellung der Kommunikationspartner, Suche nach Anomalien

Die Analyse der Datenverbindungen bietet einige Herausforderungen, einerseits ist oft eine sehr grosse Datenmenge zu durchsuchen, andererseits ist immer mehr Netzwerkverkehr verschlüsselt.

Daten

Netzwerkaufzeichnungen enthalten eine Vielzahl spannender Informationen. Die folgende Liste gibt eine kurze Übersicht:

- MAC-Adresse / IP Kombination: Dies kann helfen, eine Kommunikation einem bestimmten Gerät zuzuordnen, da die MAC Adresse eindeutig ist. Geht es um eine strafrechtliche Auswertung, ist dies jedoch mit Vorsicht zu betrachten, da diese Information von einem Angreifer einfach gefälscht werden kann.
- Protokoll spezifische Daten
- Angriffsmuster (z.B. Portscans, SQL Injections in Datenströmen, Shellcode, etc.).
- Payload Daten (z.B. heruntergeladene Dateien, POST Requests, verschickte und empfangene Emails, etc.).
- Malware Kommunikation

Oft unterschätzt, aber dennoch sehr spannend sind Metadaten. Diese stehen (zumindest teilweise) auch bei verschlüsselten Verbindungen zur Verfügung und können im Firmenumfeld so gespeichert werden, dass eine Analyse über längere Zeit möglich ist:

- passive DNS bezeichnet das Sniffing und Speichern von DNS Requests. Meist werden die Daten so reduziert, dass ein First Seen und Last Seen gespeichert wird, so dass die zu untersuchende Datenmenge sehr gering bleibt.
- passive SSL bezeichnet das Sammeln von Zertifikatsinformationen, wie z.B. dem CommonName oder dem Issuer.
- Netflow Daten sind Metadaten, welche Flows identifizieren. Ein Flow besteht in der Regel aus der Quell-, der Zieladresse, des Quell- und des Zielports sowie zusätzlichen Informationen wie z.B. die ASN (Autonomous System Number).

Diese Daten sind insbesondere auch deshalb interessant, weil sie aus Sicht Datenschutz relativ unbedenklich sind, insbesondere, wenn sie so anfallen, dass der Analytiker keine Rückschlüsse

auf eine Person (oder einen Anschluss machen kann).

Übersicht Werkzeuge

Im Folgenden eine kurze, tabellarische Übersicht über hilfreiche Werkzeuge im Bereich Netzwerk Forensik (mit dem Fokus auf frei verfügbare Werkzeuge, ohne Anspruch auf Vollständigkeit):

Werkzeug	Beschreibung
chaosreader[1]	Generiert eine Übersicht über verschiedene Protokolle in einem pcap und exp Dateien aus dem pcap.
foremost[2]	Eigentlich ein Diskforensik Werkzeug, kann foremost auch dafür verwendet w Dateien aus pcaps zu carven.
Log Indexer wie ELK[3]	Das Indexieren von Logdaten ist zentral für die Erkennung von Angriffen und
oder Graylog[4]	Bereitstellen von Beweismitteln während eines Vorfalls.
Moloch[5]	Indexer für Netzwerk Traffic. Erlaubt es grosse Mengen an pcaps aufzubewah rasch zu durchsuchen.
passiveDNS[6]	passivedns von gamelinux ist ein Werkzeug zum Aufzeichnen von allen DNS einem Netzwerkstrom.
networkminer[7]	Network Miner ist ein Werkzeug für die forensische Analyse von Netzwerk Tra
ngrep[8]	Ngrep steht für network grep und tut genau das: Es erlaubt, in Netzwerktraffic bestimmten Mustern zu suchen.
ntopng[9]	Netzwerk Probe und Analyzer, der es ermöglicht, eine schnelle Übersicht übe Vorgänge in einem Netzwerk zu erhalten.
Scalpel[10]	Kein eigentliches Netzwerk Forensik Werkzeug, kann aber auch mithelfen, be Dateitypen aus Netzwerktraffic zu extrahieren.
Scapy[11]	Mächtige Python Library zur Manipulation von Netzwerkpaketen
Snort[12]	Der «Urvater» aller Netzwerk basierten IDS Systeme (NIDS)
Suricata[13]	Vergleichbar mit Snort, jedoch mit erweiterter Syntax für das Erstellen von Ru
tcpflow[14]	Hilft dabei, TCP Pakete zu einem Flow zusammenzusetzen
tcpdump[15]	Paket Sniffer, der auf fast jedem Unix System vorhanden oder installierbar ist.
TShark[16]	Das Commandline Pendant zu Wireshark, lässt sich gut automatisieren
Wireshark[17]	Vermutlich der am Häufigsten verwendete Netzwerkprotokoll Analyzer
Zeek/Bro[18]	Zeek ist ebenfalls ein oft verwendetes NIDS, das einen besonderen Fokus au Erstellen von Events aus den einzelnen Paketflüssen legt.

Beispiel

Im Folgenden betrachten wir kurz die Kommunikation von Loki[19], einer weit verbreiteten Malware, welche sich auf das Stehlen von Zugangsdaten fokussiert.

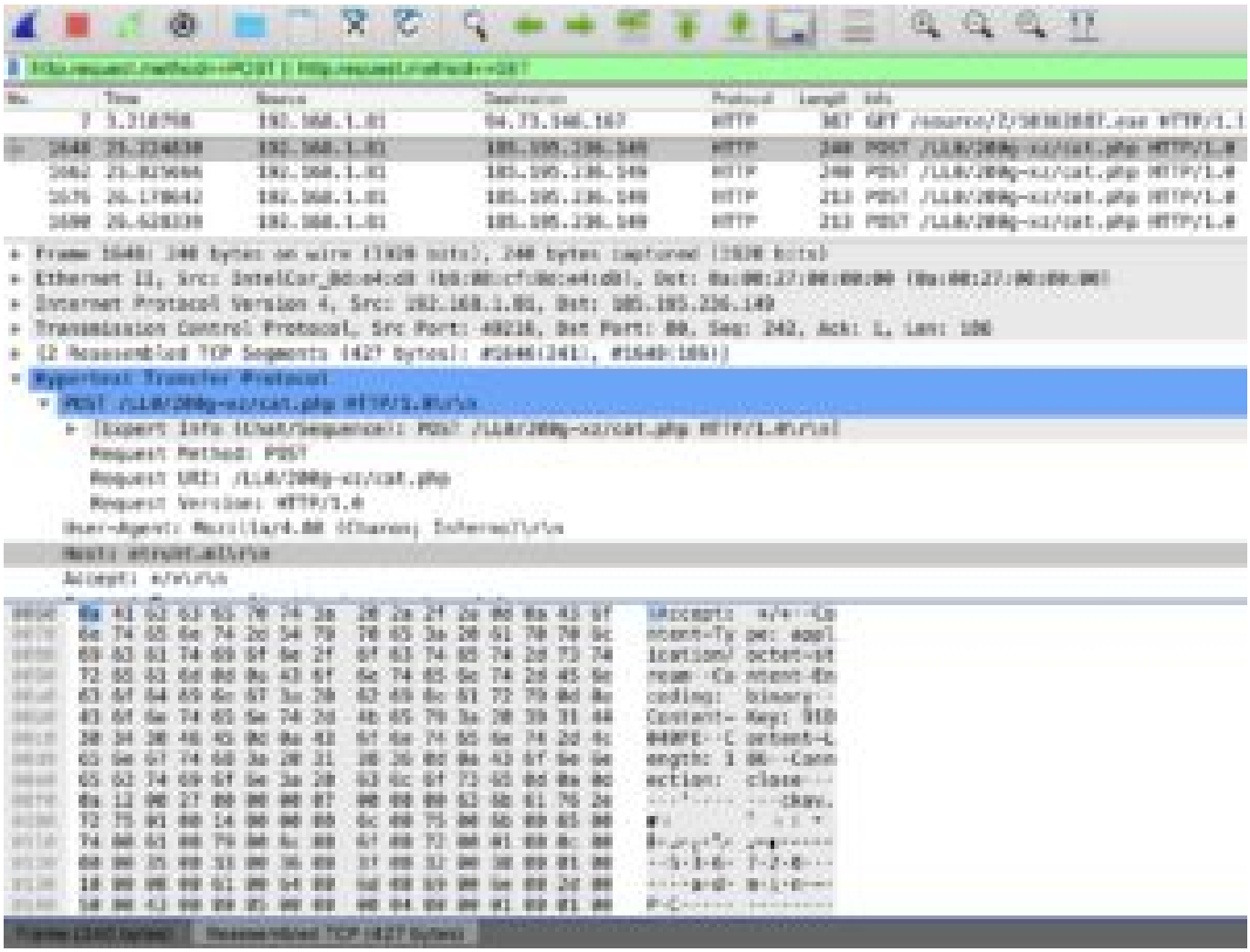
Zuerst bestimmen wir mit Hilfe von **tshark** alle erfolgreichen DNS Auflösungen:

```
tshark -n -r 752886ae315ad43235cc8a55db26f7dc.pcap -Y dns -e dns.qry.name -e dns.a
-T fields | sort | uniq

etruht[.]ml
etruht[.]ml  185.195.236[.]149
vektores[.]com
vektores[.]com      94.73.146[.]167
```

Nun untersuchen wir, ob es interessanten http Traffic gegeben hat:

```
tshark -n -r dump-752886ae315ad43235cc8a55db26f7dc.pcap -T fields -e http.host -e
http.request.method -e http.request.uri http.request.method == GET or
http.request.method == POST |sort -r |uniq

vektores[.]com      GET    /source/2/10362887.exe
etruht[.]ml  POST   /LLU/200g-xz/cat.php
```

Wir schauen dies in Wireshark grafisch aufbereitet etwas näher an:



1. Ein GET Request auf die erste Domain lädt ein executable herunter. Gleich danach folgt ein POST Request. Dies ist ein deutlicher Hinweis auf eine Infektion.
2. Der User-Agent Mozilla/4.08. (Charon; Inferno) ist aussergewöhnlich und ein guter Indicator of Compromise (IOC) für Loki.
3. Ein typisches Checkin Verhalten einer Malware. Sie übermittelt mit POST System Informationen über das infizierte Gerät

Zu den beiden IPs können wir mit wenig Aufwand folgende Meta Informationen gewinnen:

Abfrage bei whois Dienst von cymru (whois.cymru.com)

AS	IP	AS Name
209588	185.195.236[.]149	WRZ-NETWORKS-SOLUTIONS-LIMITED, GB
34619	94.73.146[.]167	CIZGI, TR

Die Länderkennzeichnung bezieht sich in dem Fall auf das Autonomous System (AS) und nicht

auf die IP selbst. Die IP 185.195.236[.]149 ist in Spamhaus als Botnet Hosting IP verzeichnet, 127.0.0.2 steht dabei als Antwort Code[20] von Spamhaus, dass die IP auf der SBL verzeichnet ist.

```
dig 149.236.195.185.zen.spamhaus.org @146.228.101.20 +short

127.0.0.2
```

Wir machen dasselbe für die Domain etruht[.]ml, die Antwort 172.0.1.6 steht dabei für eine Botnet C&C Domain.

```
dig etruht.ml.dbl.spamhaus.org @146.228.101.20 +short

127.0.1.6
```

Extrahieren wir noch das executable aus dem pcap, dies kann am Einfachsten via Wireshark/Export Objects oder mit Hilfe von Network Miner gemacht werden. Wir erhalten ein Windows PE File mit folgendem Hash (SHA-256):
181e47ab9b6c3ff0c4651492f2d401b6d3e0fa322bdd96ea9c09dee5c4a8015e

Prüfen wir diesen Hash auf Virustotal, sehen wir, dass die Datei dort bereits bekannt ist und mit den oben erwähnten Domains kommuniziert hat.

Gemäss passiveDNS Informationen, sind folgende Domains auf der IP, zu welcher die POST Requests gemacht wurden, bekannt:

Domain	firstSeen	lastSeen
etruht[.]ml	2019-02-08 17:27:07	2019-02-09 04:55:35
jatkit[.]ml	2019-01-24 12:29:59	2019-02-08 22:30:58
jatkit[.]ga	2019-01-13 05:37:36	2019-02-08 21:25:51
etruht[.]ga	2019-02-06 00:00:00	2019-02-08 04:42:27
avebx[.]ml	2019-01-25 11:46:59	2019-02-05 00:00:00

Alle Domains sind relativ neu und es würde sich sicherlich lohnen, die Logs und

Netzwerkaufzeichnungen auch nach diesen Domains abzusuchen, um allenfalls weitere, infizierte Geräte zu finden. Auf der IP Adresse, von welcher das executable heruntergeladen worden ist, finden sich 964 weitere Domains, was auf ein Shared Hosting Angebot schliessen lässt.
Möchten wir Loki Traffic unabhängig von den Domains und IPs erkennen, bietet sich der User-Agent an. Da dieser für Loki eindeutig ist, können wir daraus eine einfache Suricata Regel erstellen.

```
alert tcp $HOME_NET any -> $EXTERNAL_NET $HTTP_PORTS (msg:"Loki Useragent
Detected"; flow:established,to_server; content:"User-Agent|3A 20|"; http_header;
content:"Mozilla|2f|4.08|20 28|Charon|3B 20|Inferno|29|"; http_header;
classtype:trojan-activity; sid:50000010; rev:1;)
```

Herausforderungen für die Zukunft

Auch wenn die Datenmengen weiter zunehmen, kann dies prinzipiell gehandhabt werden, wenn auch zu teilweise beträchtlichen Kosten. Da aber immer leistungsfähigere Disksysteme zur Verfügung stehen, hält die Analysekapazität in etwa Schritt mit der Zunahme der Trafficmenge.

Durch die immer stärkere Verbreitung von Virtualisierungs- und Cloudtechnologien verändert sich die Netzwerk Forensik ebenfalls. Es ist jedoch technisch kein Problem, virtualisierten Netzwerkverkehr, der sich gar nie physisch auf einem Kabel manifestiert zu überwachen. Wichtig dabei ist aber, dass sich Firmen bewusst sind, dass ein guter Teil von Netzwerkverkehr direkt zwischen einzelnen VMs abspielt und dass sie entsprechende Vorkehrungen für das Monitoring treffen. Dasselbe gilt noch verstärkt bei der Migration von Infrastrukturen und Anwendungen in eine Cloud.

Die wohl grösste Herausforderung ist die zunehmende Verschlüsselung aller Datenströme. Die Verschlüsselung war und ist ein Kernanliegen aus Sicht der Sicherheit und hilft gegen eine Vielzahl von Angriffen und schützt die Privatsphäre. Eine Verschlüsselung macht zudem nur dann Sinn, wenn sie nicht mit Backdoors versehen ist und sie nicht "einfach ebenso mal" ausgehebelt werden kann. Dies ist umso wichtiger, je heikler die Daten sind, die übertragen werden (z.B. Patientendaten). Der steigende Einsatz von Verschlüsselungstechnologien (2017 war ca. die Hälfte des Webtraffics bereits verschlüsselt, unserer Erfahrung nach ist dieser Anteil weiter stark gestiegen[21]) verändert die Netzwerkforensik stark und beeinflusst die Fähigkeiten der Detektion und Strafverfolgung. An dieser Stelle sollen zwei Technologien betrachtet werden, die sich stark auf die Detektion und auf die Strafverfolgung auswirken werden:

- Verschlüsselung von DNS Traffic: Mit der Einführung von DNS over HTTPs[22] und DNS Crypt[23] und deren Unterstützung durch grosse DNS Operator wie z.B. OpenDNS, Cloudflare und Google, reduziert sich der Nutzen von passive DNS Datenbanken. Diese haben in der Verfolgung von Malware Akteuren jedoch bisher eine beträchtliche Rolle gespielt. Innerhalb einer Firma stehen die Daten prinzipiell nach wie vor zur Verfügung, so lange eigene Resolver eingesetzt und deren Logs entsprechend

aufbereitet werden. [24]

- TLS 1.3: Die neueste Version von Transport Layer Security erhöht – nebst weiteren Design Zielen – die Privacy von Verbindungen, u.A. durch einen neuen TLS Handshake. Dies führt einerseits dazu, dass Zertifikatsinformationen nicht mehr im Klartext übertragen werden und andererseits erhöht es die Anforderungen an TLS Interception, was für viele Firmen bedeutsam ist. Das Fehlen der TLS Handshake Informationen reduziert – ähnlich wie bei passiveDNS die Möglichkeiten der Verfolgung von kriminellen Akteuren durch das Sammeln und Ableiten von Metainformationen. Im Zusammenhang mit der TLS Interception sind zwei Punkte erwähnenswert: Es ist nach wie vor möglich im Firmenumfeld eine Interception zu machen. Keinesfalls sollte das Security Device dabei einfach die Verbindung auf TLS 1.2 herunterhandeln, es braucht jedoch einiges an Forschung und Aufwand, um TLS 1.3 so zu behandeln, dass eine Interception möglich ist, aber gewisse Seiten wie z.B.: Krankenkassenportale davon ausgenommen sind (Whitelisting von gewissen Seiten), da diese Informationen nicht mehr im Klartext zur Verfügung stehen.

Referenzen

[1] http://chaosreader.sourceforge.net

[2] http://foremost.sourceforge.net

[3] https://www.elastic.co/de/elk-stack

[4] https://www.graylog.org

[5] https://molo.ch

[6] https://github.com/gamelinux/passivedns

[7] https://www.netresec.com/?page=networkminer

[8] https://github.com/jpr5/ngrep

[9] https://www.ntop.org

[10] https://github.com/sleuthkit/scalpel

[11] https://scapy.net

[12] https://www.snort.org

[13] https://suricata-ids.org

[14] https://github.com/simsong/tcpflow

[15] man tcpdump

[16] https://www.wireshark.org

[17] https://www.wireshark.org

[18] https://www.zeek.org

[19] https://malpedia.caad.fkie.fraunhofer.de/details/win.lokipws

[20] https://www.spamhaus.org/faq/section/DNSBL%20Usage#200

[21] EFF, We're Halfway to Encrypting the Entire Web: https://www.eff.org/deeplinks/2017/02/were-halfway-encrypting-entire-web

[22] IETF, RFC 8484: https://tools.ietf.org/html/rfc8484

[23] DNS Crypt Projekt: https://dnscrypt.info

[24] IETF, TLS 1.3 RFC 8446: https://tools.ietf.org/html/rfc8446