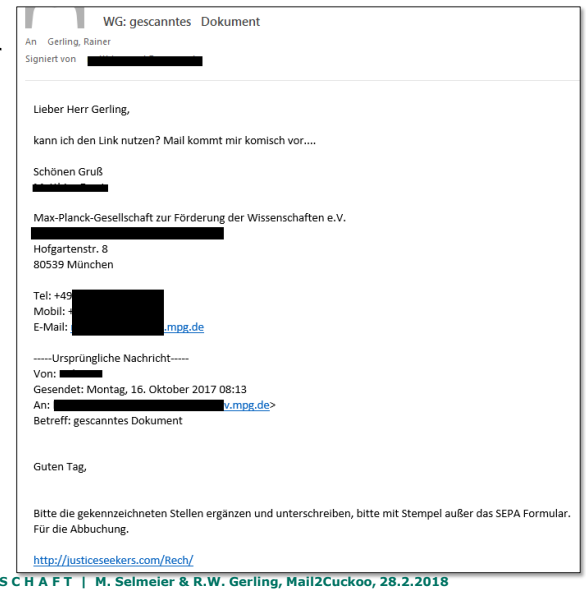
E-Mail-Analyse auf Basis von Cuckoo Sandbox

Manuel Selmeier, Rainer W. Gerling
Generalverwaltung
Max-Planck-Gesellschaft

Motivation

- Fast jeder in der IT oder IT-Sicherheit bekommt derartige E-Mails.
- Aus dem Bauch beantworten?
- Jedes Mal Recherche Aufwand?
- Da schreit nach einer Bachelor-Arbeit!



WG: gescanntes Dokument

An Gerling, Rainer
Signiert von [redacted]

Lieber Herr Gerling,

kann ich den Link nutzen? Mail kommt mir komisch vor....

Schönen Gruß
[redacted]

Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.
[redacted]
Hofgartenstr. 8
80539 München

Tel: +49[redacted]
Mobil: +[redacted]
E-Mail: [redacted].mpg.de

-----Ursprüngliche Nachricht-----
Von: [redacted]
Gesendet: Montag, 16. Oktober 2017 08:13
An: [redacted]v.mpg.de>
Betreff: gescanntes Dokument

Guten Tag,

Bitte die gekennzeichneten Stellen ergänzen und unterschreiben, bitte mit Stempel außer das SEPA Formular. Für die Abbuchung.

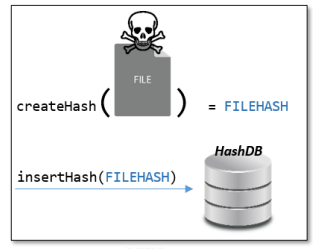
http://justiceseekers.com/Rech/

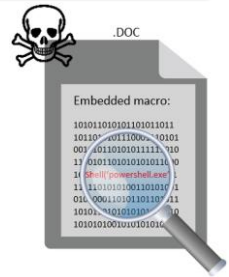
Randbedingungen

- Der Nutzer solle möglichst ohne Kontakt mit der Malware den Dienst nutzen können ➔ E-Mail-Weiterleitung
  - Weiterleitung an dedizierte E-Mail-Adresse
- Datenschutz und Mitbestimmung einhalten
  - Alle Dateien bleiben in der Max-Planck-Gesellschaft.
  - Kein Upload nach Virustotal! Prüfung nur über Hashwert.
- Vorgegebenes System für die Bachelor-Arbeit war Cuckoo Sandbox
  - Cuckoo möglichst nicht patchen
- Erstellung der Bachelor-Arbeit von Sept. – Dez. 2016

Malware Erkennungsansätze

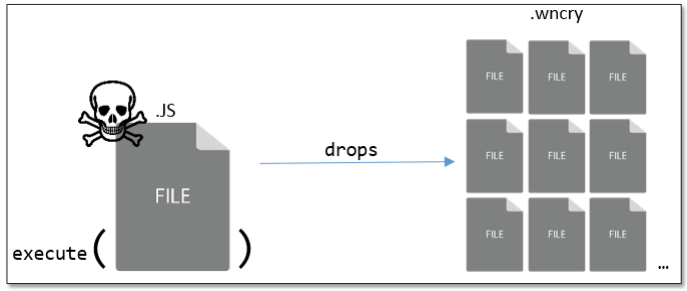
- Hashbasiert
  - Lookup in jeweilige HashDB
  - Schnell
  - Zuverlässig
  - Nicht up to date
- Signatur-/Patternbasiert
  - Untersuchung des Speichers/Binarys in Echtzeit nach maliziösen Mustern
  - Relativ schnell
  - False-positives möglich





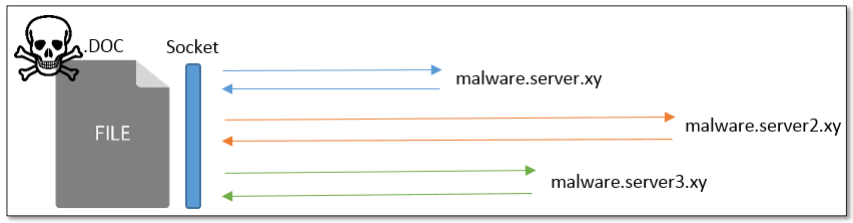
Malware Erkennungsansätze II

- **Verhaltensbasiert**
  - Verhaltensanalyse der ausführenden Datei **in Echtzeit**
  - Beispiel 1: File Input/Output
    - JavaScript verursacht ungewöhnliche Dateiströme mit verdächtigen Dateiendungen



Malware Erkennungsansätze III

- Beispiel 2: Socket-Streams
  - Ausführung der .DOC-Datei verursacht ungewöhnliche Netzwerkverbindungen



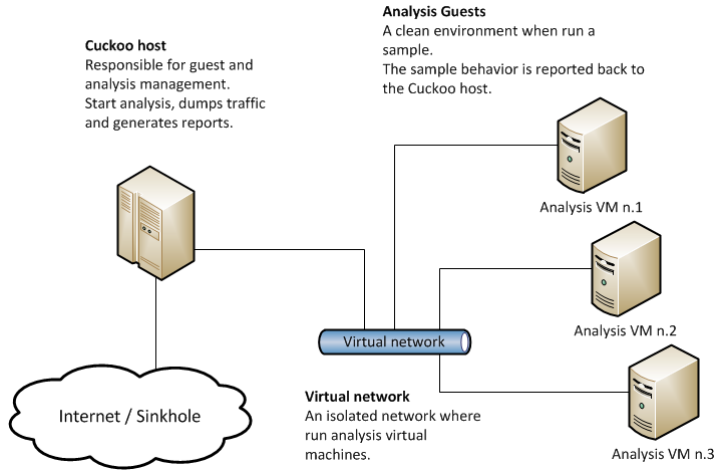
Cuckoo Sandbox

"Cuckoo Sandbox is an advanced, extremely modular, and 100% open source automated malware analysis system with infinite application opportunities."

- Cuckoo Sandbox …
  - kombiniert alle gängigen Erkennungsansätze
  - analysiert verbreitete Dateitypen (u.a. .exe, Office-Dateien, .pdf, .com, .js, .py)
  - führt vollautomatische Echtzeitanalysen auf VM und bare-metal-Maschinen durch
  - erstellt detaillierte und sehr umfangreiche Report-Summarys (JSON)

Quelle: https://cuckoosandbox.org/

Cuckoo Sandbox: Architektur



Quelle: http://docs.cuckoosandbox.org/en/latest/introduction/what/

Cuckoo Sandbox: Was wird geloggt

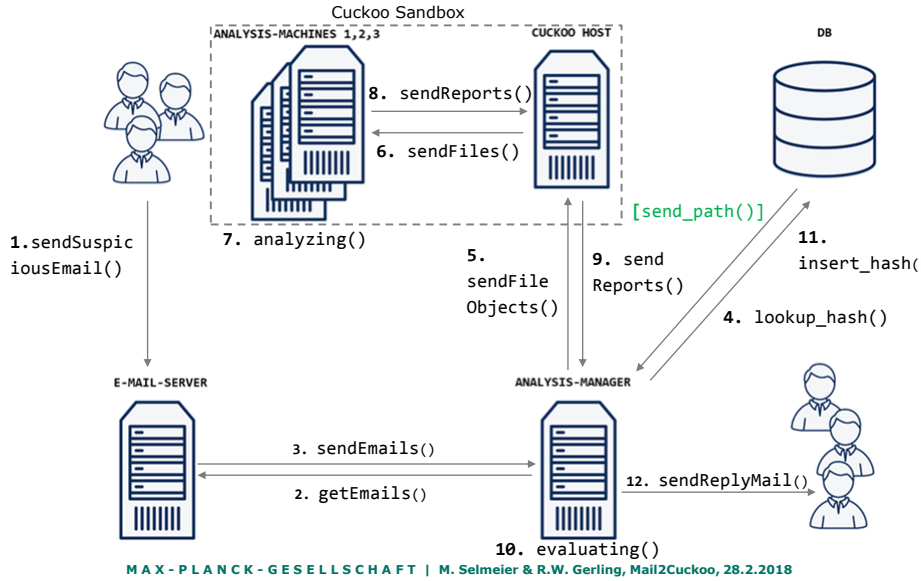
- Die Analyse von Cuckoo Sandbox erfasst
  - Prozessmanagement
  - Threadmanagement
  - Registry-Modifikationen
  - File Input /Output
  - Socket-Aktivitäten

  - Signaturen
  - Weitere beliebige Reporting-Module

Cuckoo Erweiterung: Mail2Cuckoo

- E-Mail-basierte Analyseerweiterung
- Implementiert Cuckoo Sandbox für Dateianalysen
- Erweitert Cuckoo Sandbox in den Punkten
  - Bewertung: Erstellung einer aussagekräftigen Reply-E-Mail
  - Handhabung: E-Mail-Weiterleitung an Cuckoo-Postfach genügt
  - URL-Analyse: Erkennt blacklisted URLs und referenzierte Dateien
  - Benutzer-Sicherheit: Kein direkter Malwarekontakt

Mail2Cuckoo: Analyseablauf

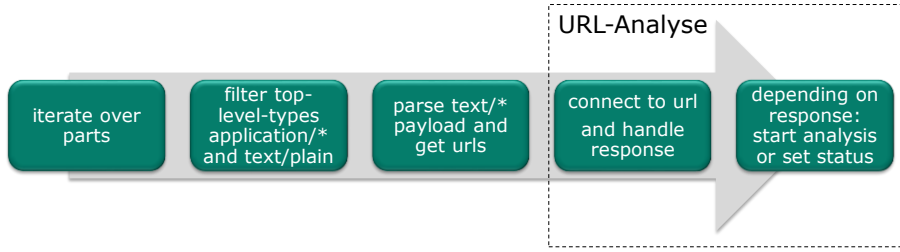


Mail2Cuckoo: Weitere Funktionen

- Analyse von Dateianhängen
- Analyse von Internetlinks in E-Mails auf:
  - Referenz auf Datei
  - Blacklisting (falls Information darüber verfügbar)
  - Phishing (in Planung)
- Vollautomatisiert nach E-Mail-Weiterleitung
- Auto-E-Mail-Reply mit Bewertung
- Hash-Caching: Ergebnisabruf bereits analysierter Dateien

E-Mail-Analyse

- Die E-Mail wird nach folgendem Schema untersucht:



- Die Ermittlung von Dateien aus dem Mailanhang/hinter URLs steht im Vordergrund

URL-Analyse

1. Request wird mit gespooften Firefox-Browser als User-Agent erzeugt
   - `request = urllib2.Request(url, headers={"User-Agent" : "Mozilla/5.0 (Windows NT 10.0; Win64; x64;rv:56.0) Gecko/20100101 Firefox/56.0", …})`
2. Request wird ausgeführt
   - `response = urllib2.urlopen(request)`
3. Überprüfung des Response:
   - `response_code == 200`: Überprüfung des „Content-Type"- Felds im Response
   - `response_code != 200`: Reporting

URL-Analyse 2

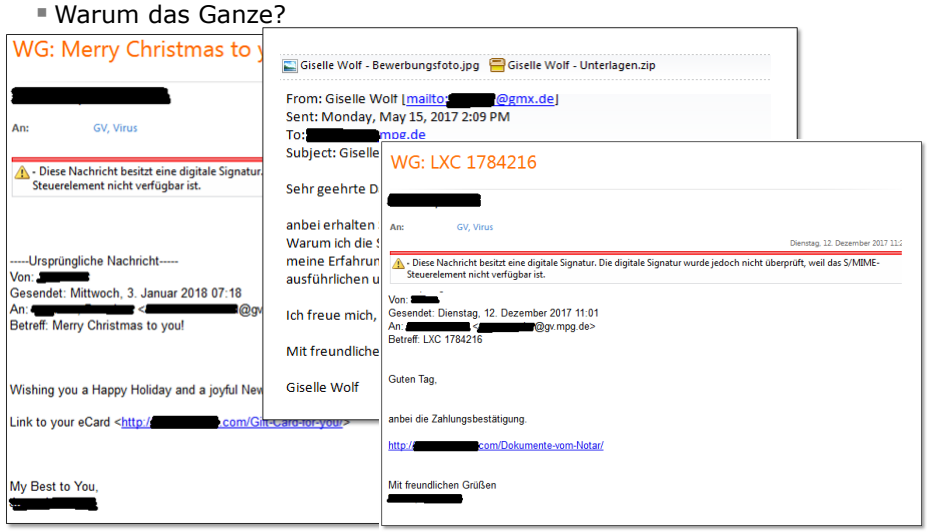
Beispiel-Response einer ZIP-Datei

Content-Type: application/zip
Content-Length: 3867203
Connection: close
Server: Apache
Accept-Ranges: bytes
Last-Modified: Web, 1 Feb 2018 11:05:19
[…]

- Response enthält „Content-Disposition“-Feld:
  Datei gefunden! -> cuckoo
- Content-Type enthält MIME-Type application/*:
  Datei gefunden! -> cuckoo
- Content-Type enthält MIME-Type text/*:
  Keine Datei gefunden! -> Check blacklisted status

URL-Analyse 3

- Warum das Ganze?



Bewertung

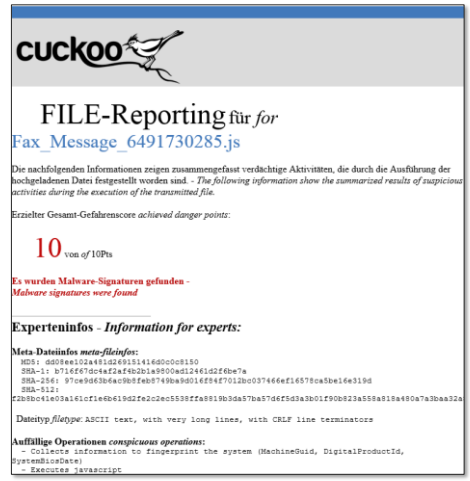
- Grundlage für die Bewertung von Mail2Cuckoo ist der Cuckoo Report
- Kriterien für die Einstufung als Malware sind:
  - Auffälligkeiten bei Socket Streams
  - Auffälligkeiten im File Input/Output
  - Erfüllung typischer Malwaresignaturen
  - Treffer in AV-Datenbanken
  - Ausführung ungewöhnlicher Prozesse
  - Statische Analyse (bei Office-Dateien und PDFs)

Bewertung: Scoring

- Signatur: $hit * 0{,}5\ pts$
- Netzwerk: $hit * 0{,}4\ pts$
- Prozesse und File I/O:
  - Prozesse: $hit * 2\ pts$
  - Änderungen Dateistruktur: $hit * 0{,}3\ pts$
- VirusTotal
  - Bekannte AV-Engine: $hit * 2\ pts$
  - Sonstige AV-Engine: $hit * 1\ pt$
- Strings: $hit * 0{,}5\ pts$
- **Score = Summe aller Module**

Bewertung: Auto-Mail-Reply

- Durch die Vergabe eines Scores ist die Gefährlichkeit der Datei leicht beurteilbar
- IOC`s und weitere score-beeinflussende Faktoren werden unter Experteninfos aufgeführt
- Das Reportlayout ist (wie alles andere) beliebig veränderbar


cuckoo

FILE-Reporting für *for*
Fax_Message_6491730285.js

Die nachfolgenden Informationen zeigen zusammengefasst verdächtige Aktivitäten, die durch die Ausführung der hochgeladenen Datei festgestellt worden sind. - *The following information show the summarized results of suspicious activities during the execution of the transmitted file.*

Erzielter Gesamt-Gefahrenscore *achieved danger points:*

10 von *of* 10Pts

Es wurden Malware-Signaturen gefunden -
Malware signatures were found

Experteninfos - ***Information for experts:***

Meta-Dateiinfos ***meta-fileinfos*:**
MD5: dd08ee102a481d26915141 6d0c0c8150
SHA-1: b716f67dc4af2af4b2b1a9800ad12461d2f6be7a
SHA-256: 97ce9d63b6ac9b8feb8749ba8d16f84f7012bc037466ef1657 8ca5be16e319d
SHA-512:
f2b8bc41e03a161cf1e6b619d2fe2c2ec5538ffa8819b3da57ba57d6f5d3a3b01f90b823a558a818a480a7a3baa32a

Dateityp *filetype*: ASCII text, with very long lines, with CRLF line terminators

Auffällige Operationen ***conspicuous operations*:**
- Collects information to fingerprint the system (MachineGuid, DigitalProductId, SystemBiosDate)
- Executes javascript

Mail2Cuckoo

GitHub

- Der Sourcecode ist öffentlich zugänglich und abrufbar unter

https://gitlab.mpcdf.mpg.de/mselm/mailAtCuckoo


ANY
QUESTIONS
?