INF FAKULTÄT FÜR INFORMATIK

Phishing Infrastructure Knowledge

GITS
06.07.22

Gina Bartenwerfer
Bernhard Birnbaum
Vincent Donat
Felix Gretschel
Abdalla Khalil
Michelle Kirst

Inhaltsverzeichnis

Organisation

- 20 Treffen (Stand 04.07)
- HedgeDoc für Protokoll und wöchentliche Aufgabenverteilung
- OVGU Cloud für Dateien
- Dynamische Organisation
- Keine festen Rollen und Aufgaben, sondern Einteilung je nach Bedarf

Motivation & Stand der Technik

- Zentrale Fragestellung: welche Daten im Zusammenhang mit Phishing-Attacken stehen den Dienstanbietern zur Verfügung
- In welchem Umfang können diese Dienstanbieter solche Angriffe erkennen bzw. unterbinden
- Sicherheitsaspekte werden verletzt
  - Sicherheitsaspekte, die unsere Fragestellung betreffen sind:
    - Authentizität
    - Vertraulichkeit
    - Nichtabstreitbarkeit

Motivation & Stand der Technik

- Werkzeuge
  - Mitmproxy
  - Wireshark
  - Thunderbird
  - Evolution



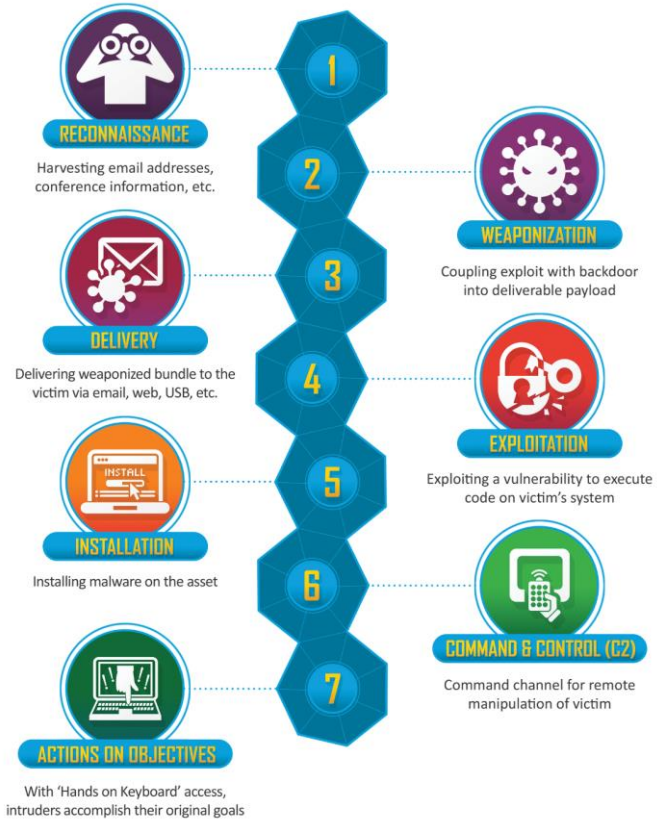

Motivation & Stand der Technik

- CERT-Taxonomie

Angreifer	Werkzeuge	Schwachstelle	Aktion	Ziel	Resultat	Absicht
Kriminelle Beauftragte	Malware Websites Nachrichten	Mensch Adressobfuskation	Masquerade Stehlen Auslesen	Account Daten Informationen	Zugriff auf persönliche Daten Unerlaubter Zugriff auf Informationen	Finanziell motiviert

Motivation & Stand der Technik

- Kill-Chain
  - Basiert auf militärischem Konzept
  - Beschreibt einen Angriff aus der Sicht des Angreifers
  - Unterteilt sich in 7 Ebenen
    - Reconnaissance
    - Weaponization
    - Delivery
    - Exploitation
    - Installation
    - Command and Control
    - Actions on objective



Konzept

Vorbereitung der Testumgebung

In der Testumgebung müssen folgende Anforderungen und Eigenschaften erfüllt sein:

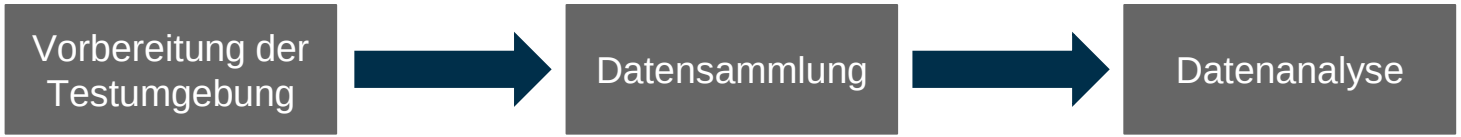
- Rauscharme Umgebung
  - Virtual Machine mit Linux OS
- Netzwerkverkehr aufzeichnen
  - mittels Mitmproxy und Wireshark
- E-Mails öffnen
  - durch Thunderbird und Evolution
- Arbeitsspeicher festhalten
  - Funktion der VM
- Massenspeicher sichern
  - kopieren der wichtigen Dateien

Konzept

In der Datensammlung werden folgende Daten erhoben:

- Metadaten
  - Header der Mail und Sender der Mail
- Netzwerkverkehr
  - Öffnen der Mail in TB/ Evol.
  - Öffnung der verlinkten Website in Firefox und die dortige Eingabe von Daten
- Arbeitsspeicher und Massenspeicher
  - vor Schließung aller relevanten Programme

Konzept



Die Analyse betrachtet folgende aufgezeichneten Daten:

- Metadaten werden auf auffällige Daten untersucht
- Analyse der Netzwerkverbindungen
  - Untersuchung der Verbindungen in Wireshark/ Mitmproxy
- Arbeitsspeicher wird mittels String vorsortiert und anschließend untersucht
- Massenspeicher
  - Cookies von TB und Firefox mit SQLite analysiert
  - Formulardaten die in Firefox eingegeben wurden ebenfalls in SQLite

Implementierung

- Nutzung des Testersticks der Uni (rauscharme Umgebung) mit Hilfe von VirtualBox
- E-Mail-System mit synthetischen Daten für die Phishing-Mails
- Tools zur Datenakquise (Firefox ESR, Wireshark, mitmproxy, Thunderbird und Evolution)
- Bei der Untersuchung von Phishing-Mails ein eigenes Skript erstellt
- Anmelden auf vielen Seiten an Foren/Netzwerke, um Phishing-Mails zu gelangen (auf Newsletter, Dating-Seiten und tempr.email)
- Wireshark und MITMProxy (Eingriff in Datenstrom) für Netzwerk-Traffic
- Main-Storage mit SQLite-Viewer und vmdump mit Strings konvertiert

Evaluierung

- Hauptproblem:
  - Fehlende Phishing E-Mails
    - Eher normaler Spam
  - Wir konnten jede Art von Phishing E-Mail untersuchen (URL-Obfuscation, HTML-Attachment, Open-Redirect)
    - Aber auch von denen nur wenige

    → wir konnten nicht das volle Spektrum analysieren

- Warum ist das so?
  - “Published” Leute haben den Vorteil, dass ihre E-Mail-Adresse im Internet verbreitet ist
  - Unsere benutzte E-Mail Adresse ist nicht öffentlich genug
    - Dadurch nicht genug authentische Phishing E-Mails

Evaluierung

- Tracker in E-Mails
  - Dienstanbieter haben Mails auf Servern zu liegen
    - Existenz/Status müssen ihnen bekannt sein
  - Anhand der E-Mail fast unmöglich herauszufinden, ob es sich um Phishing oder legitime E-Mail handelt
- Tracker auf Phishing Seiten
  - Große Dienstanbieter haben zwar die Möglichkeit Links in E-Mails zu verfolgen
    - Aufgrund von Kosten aber unwahrscheinlich
  - Automatisierte Phishing Erkennung theoretisch durch Abgleich und Markierung möglich
    - Dazu müsste Phishing Adresse aber bereits bekannt sein
    - Unterscheidung ansonsten schwierig (Unsere Interaktion mit Website müsste beobachtet werden)

Zusammenfassung & Ausblick

- Projekt mehr oder weniger erfolgreich
  - Untersuchungen konnten zwar durchgeführt werden, allerdings kann die geringe Menge der Analysen nicht das vollständige Spektrum aller Phishing-Angriffe abdecken
  - Große Dienstanbieter können durchaus Phishing-Angriffe wahrnehmen:
    - Tracker in E-Mails (JavaScript, Schriftarten, Stylesheets)
    - Redirects/URL-Shortener (z.B. goo.gl bzw. Firebase Dynamic Links, …)
  - Viel Raum für weiterführende Untersuchungen
    - Was könnten Dienstanbieter tun, um Phishing-Angriffe zu verhindern?
    - Gibt es weitere Mechanismen, wodurch Dienstanbieter von solchen Angriffen wissen können? (mit Hilfe eines größeren Testsets)
    - Welche weiteren Tools könnten hilfreich sein, um die Analyse effizienter durchzuführen (z.B. Volatility, …)

Fragen?