SIEMENS

Schlussbericht

Effiziente Verhaltensanalyse von modernem Schadcode

siemens.com

This page intentionally left blank

Projektname

Effiziente Verhaltensanalyse von modernem Schadcode

Akronym

VAMOS

Förderkennzeichen

16KIS0535

Projektlaufzeit

06/2016 - 10/2020

Teilvorhaben

Verbesserung der Analysefähigkeiten von Angriffsindikatoren und Anbindung in die Detektion bei Großunternehmen

Eingereicht im Rahmen der Bekanntmachung „Erkennung und Aufklarung¨ von IT-Sicherheitsvorfällen”

Gefördert vom

Bundesministerium
für Bildung
und Forschung

This page intentionally left blank

Berichtsblatt

1. ISBN oder ISSN ---	2. Berichtsart (Schlussbericht oder Veröffentlichung) *Schlussbericht*
3. Titel Abschlussbericht des Teilvorhabens „Effiziente Verhaltensanalyse von modernem Schadcode“	
4. Autor(en) [Name(n), Vorname(n)] *Caselli, Marco*	5. Abschlussdatum des Vorhabens *31.10.2020*
	6. Veröffentlichungsdatum *Geplant*
	7. Form der Publikation *Bericht*
8. Durchführende Institution(en) (Name, Adresse) *Siemens Aktiengesellschaft* *Werner-von-Siemens-Straße 1* *80333 Munich* *Germany*	9. Ber. Nr. Durchführende Institution ---
	10. Förderkennzeichen *16KIS0535*
	11. Seitenzahl 33
12. Fördernde Institution (Name, Adresse) Bundesministerium für Bildung und Forschung (BMBF) 53170 Bonn	13. Literaturangaben 16
	14. Tabellen 1
	15. Abbildungen 11
16. Zusätzliche Angaben ---	
17. Vorgelegt bei (Titel, Ort, Datum) ---	

18. Kurzfassung

Heutzutage ist die Sicherheit von IT-Systemen durch eine immer größere Anzahl von Cyberangriffen massiv gefährdet. Täglich dringt immer ausgefeiltere Malware in die Anwendungen von Einzelpersonen und Unternehmen ein, die Daten stehlen, den Betrieb behindern und Geschäfte beeinträchtigen. Computer Emergency Response Teams (CERTs) und allgemeiner Security Operation Center (SOCs) bilden die Verteidigungslinie gegen diese Bedrohungen. Die Aufgabe der Sicherheitsexperten, die diese Teams bilden, erfordert jedoch Technologien, die ihnen helfen können, nicht nur Cyberangriffe rechtzeitig zu erkennen, sondern auch diese hochentwickelte Malware effektiv zu untersuchen, ihre Mechanismen zu verstehen und schließlich Wissen innerhalb der gesamten Sicherheitsgemeinschaft auszutauschen.
Das VAMOS-Projekt zielte genau auf diese Ziele ab und konzentrierte sich auf die Entwicklung neuartiger Techniken zur Analyse und Erkennung hochkomplexen Schadcodes. In diesem Zusammenhang brachte Siemens AG seine langjährige Erfahrung in der Cybersicherheit und insbesondere in der Behandlung von Sicherheitsvorfällen (innerhalb des Siemens CERT) ein. Die Hauptaufgabe von Siemens Technology konzentrierte sich speziell auf die Entwicklung von Methoden zur automatischen Ableitung und Anreicherung von Kompromitierungsindikatoren aus dem beobachteten Verhalten von Schadcode. Um dies zu erreichen, hat Siemens Technologies bekannte und weit verbreitete Technologien (z. B. MISP, YARA) und kundenspezifische Sicherheitslösungen, die intern entwickelt wurden (z. B. CMAP) verwendet. Zusätzlich zu dieser Technologie wurde eine Reihe neuartiger Ansätze entwickelt, um von VAMOS-Partnern (VMRay) generierte Kompromitierungsindikatoren zu erfassen und ihre semantische Bedeutung zu verbessern. Die entwickelten Lösungen wurden schließlich in der Produktion getestet und in die am Siemens CERT verwendeten Sicherheitstoolketten integriert.

19. Schlagwörter Cybersicherheit, Schadcode, Computer-Notfallteam, Kompromitierungsindikatoren	
20. Verlag ---	21. Preis ---

This page intentionally left blank

Document Control Sheet

1. ISBN or ISSN ---	2. type of document (e.g. report, publication) Final report
3. title Abschlussbericht des Teilvorhabens „Effiziente Verhaltensanalyse von modernem Schadcode“	
4. author(s) (family name, first name(s)) *Caselli, Marco*	5. end of project *31.10.2020* 6. publication date *Planned* 7. form of publication Report
8. performing organization(s) (name, address) *Siemens Aktiengesellschaft* *Werner-von-Siemens-Straße 1* *80333 Munich* *Germany*	9. originator’s report no. --- 10. reference no. *16KIS0535* 11. no. of pages 33
12. sponsoring agency (name, address) Bundesministerium für Bildung und Forschung (BMBF) 53170 Bonn	13. no. of references 16 14. no. of tables 1 15. no. of figures 11
16. supplementary notes ---	
17. presented at (title, place, date) ---	
18. abstract Nowadays, the security of IT systems is massively endangered by an always increasing number of cyberattacks. More and more sophisticated malware daily breaks into individuals’ and companies’ applications stealing data, hindering operations, and impairing businesses. Computer emergency response teams (CERTs) and, more in general, Security Operation Centers (SOCs) represent the line of defense against these threats. However, the job of the security experts making those teams requires technologies that can help them not just by timely detecting cyberattacks but also by effectively investigating this sophisticated malware, understanding its mechanisms, and finally sharing knowledge within the whole security community. The VAMOS project aimed exactly at these objectives and focus on the development of novel techniques for the analysis and detection of highly complex malicious code. In this context, Siemens Technology brought its years of experience in cybersecurity and, especially, in the handling of security incidents (operated within the Siemens CERT). Siemens AG main task specifically focused on the development of methods for the automatic derivation and enrichment of indicators of compromise from observed malicious code behavior. To achieve this, Siemens Technologies has employed well-know and widely used technologies (e.g., MISP, YARA), custom security solutions developed in-house (e.g., CMAP) and a set of novel approaches to ingest indicators of compromise generated by VAMOS partners (VMRay) and enhance their semantical meaning. The developed solutions have been eventually tested in production and integrated within the security tool chains in use at the Siemens CERT.	
19. keywords Cybersecurity, Malware, Computer Emergency Response Team (CERT), Indicator of Compromise (IoC)	
20. publisher ---	21. price ---

This page intentionally left blank

Table of Contents

Ziele des Teilvorhabens

Das Ziel dieses Teilvorhabens ist die Verbesserung der Analysefähigkeiten von Indikatoren, die aus der Schadsoftwareanalyse gewonnen werden. In Sicherheitsteams fallen durch die tägliche Auswertung der diversen Quellen, eine sehr große Anzahl von potenziellen Indikatoren an. Diese müssen von Analysten teilweise manuell gesichtet und auf Qualität geprüft werden.

Des weiteren ist ein wesentlicher Punkt bei der Aufarbeitung der Analysen, die effiziente Speicherung der Indikatoren wichtig. Zu einem ist es notwendig, dass man schnell die notwendigen Informationen in die Detektion von Angriffen bringt, zum anderen müssen Analysten schnell nach Indikatoren und deren Kontext suchen können.

Derzeitige Ansätze in der IT-Sicherheit können diese Anforderungen nicht erfüllen. Mit dem hier erarbeitenden Ansatz wollen wir durch die Kombinationen der hier entwickelten Analyseverfahren und der Anreicherung dieser Informationen mit externen Quellen, ein Verfahren entwickeln, welches Sicherheitsteams unterstützt schnell und adäquat die richtigen Schritte durchzuführen.

Die Ziele dieses Teilvorhabens sind:

Verbesserung der Analysefähigkeiten

Ein essenzielles Ziel dieses Teilvorhabens ist die Verbesserung der derzeitigen Situation, wie Analysten Angreifer Daten analysieren. Es ist ein fast ausschließlicher manueller Prozess. Dies führt dazu, dass sehr viel Zeit verloren geht, bis Angriffe detektiert werden können und auch die Qualität der Indikatoren leidet darunter. Durch die im Rahmen von VAMOS entwickelten Verfahren, erhoffen wir eine wesentliche Verbesserung dieser Situation.

Integration des Ansatzes in Unternehmensnetzwerken

Unternehmensnetzwerke sind eine komplexe Infrastruktur und die Integration von neu-en Erkennungs- und Analyseverfahren dauert oft sehr lange, weil viele Aspekte nicht betrachtet werden. Durch die Zusammenarbeit der Projektpartner ist es möglich diese Anforderungen direkt bei der Entwicklung einfließen zu lassen.

Datenschutzrechtliche Bewertung der erhobenen Daten

Bei der Erkennung von Angriffen mittels Indikatoren wird oft vergessen, dass hier Daten verarbeitet werden, welche möglicherweise unter den Datenschutz fallen. Aus diesem Grund wird eine datenschutzrechtliche Bewertung der verarbeiteten Daten, sowie der Ver-fahren in diesem Teilvorhaben durchgeführt und wenn notwendig, Losungen¨ integriert, um den Datenschutz zu verbessern.

Arbeitsplan

Das Vorhaben hat eine Laufzeit von 36 Monaten und ist in fünf Teilprojekte gegliedert, die jeweils zusammengehörige Arbeitspakete bündeln. Der Arbeitsaufwand im Teilvorhaben kann Tabelle 1 entnommen werden.

Teilprojekte		**Siemens AG**
TP.1	Anforderungsanalyse und Datenerhebung	2 PM
TP.2	Ableitung und Filterung von Indikatoren	8 PM
TP.3	Linearzeit-Anomalieerkennung und -Clustering	0 PM
TP.4	Korrelation und Visualisierung von Verhalten	0 PM
TP.5	Evaluation und Testbetrieb	2 PM
Gesamt		12 PM

Tabelle 1 - Teilprojekte des Vorhabens und Aufwand in Personenmonaten

TP.1 Anforderungsanalyse und Datenerhebung

Dauer: 2PM

Im ersten Teilprojekt soll zur Entwicklung und Evaluierung der neuen Methoden und Ver-fahren ein umfangreicher Datensatz von Schadcode erhoben werden. Weiterhin sollen für die Praxis relevante Anforderungen, wie standardisierte Formate für den Datenaustausch und unternehmenstypische Arbeitsprozesse analysiert und spezifiziert werden.

AP.1.1 Erhebung und Aufbereitung eines umfangreichen Schadcode-Datensatzes

Dauer: 1PM

Die Entwicklung und Evaluierung neuer Methoden im Rahmen dieses Projekts erfordert einen umfangreichen Datensatz aus verschiedenartigem Schadcode. In der Praxis erhalt¨ man solche Schadcode-Samples über verschiedene Wege: unter akademischen und industriellen Partnern existieren diverse Sample-Sharing-Programme, es gibt diverse öffentliche Datenbanken mit unsortierter, aber auch bereits vorklassifizierter Malware, und in vielen Fällen ist es zusätzlich sinnvoll, synthetische Selbsterzeugnisse zu verwendet.

Die Dateien selbst sollen möglichst unterschiedliche Eigenarten von typischem Malware-Verhalten aufweisen und zusammengenommen eine optimalerweise vollständige Abdeckung aller denkbaren Verhaltensweisen mit sich bringen. So müssen zum Beispiel alle möglichen Formen von ausfuhrbaren¨ Dateien berücksichtigt werden (32-Bit und 64-Bit; Anwendungen, Bibliotheken und Systemtreiber; Skriptsprachen), diverse Formen von Schaddokumenten (PDF-Dateien, Office-Dokumente, usw.) und zusätzlich auch noch bösartige Webseiten mit den unterschiedlichsten verwendeten Technologien (z.B. Java- oder Flash-Exploits). Insbesondere die Erhebung von den für dieses Vorhaben relevanten APT-Schadcodes ist eine

komplizierte Angelegenheit, da diese naturgemäß nicht in groß er Zahl existieren und erst recht nicht häufig an die Öffentlichkeit geraten.

Neben den eigentlichen Dateien werden weiterhin zusätzliche Metainformationen benötigt, insbesondere eine Form sogenannter Ground Truth, also eine Vorklassifizierung und Einteilung in die unterschiedlichen Malware-Typen und -Familien. Diese werden vor allem benötigt, um die Wirksamkeit und Korrektheit der zu entwickelnden Methoden zu verifizieren.

AP.1.3 Aufbau eines effizienten Datenbanksystems für Indikatoren (IoCs)

Dauer: 1PM

Die effiziente Abspeicherung der Indikatoren essenziell für die spätere Weiterverarbeitung der Daten. In diesem Arbeitspaket sollen derzeitige Ansätze weiterentwickelt werden, damit mit den zu erwartenden Datenmengen umgegangen werden kann. Dabei werden neue Datenmodel-le, wie z.B. die semantische Darstellung der Indikatoren, evaluiert und neue Datenbanksysteme verwendet. Ziel soll es sein, am Ende ein System zu haben, welches die Angreifer Daten semantisch darstellen kann, aber auch eine schnelle Suche innerhalb der Daten ermöglicht.

TP.2 Ableitung und Filterung von Indikatoren

Dauer: 8 PM

In diesem Teilprojekt sollen Methoden entwickelt werden, um Verhaltensindikatoren aus relevanten Analysedaten — zum Beispiel aus einer manuellen oder automatischen Analyse — abzuleiten und in einem geeigneten Standard zu beschreiben (vgl. AP.1.2).

AP.2.1 Methoden zur Ableitung von Indikatoren aus Verhaltensdaten

Dauer: 2 PM

Die verhaltensbasierte Analyse von Schadcode bietet eine Vielzahl von möglichen Indikatoren, die aus den Ergebnissen abgeleitet werden können. Die von VMRay eingesetzte Technologie liefert zudem eine noch weitaus größere Menge an detaillierten Analysedaten als konkurrierende Verfahren. Um die später zur Verfügung stehenden Ressourcen möglichst effizient einzusetzen, ist es notwendig, die vielversprechendsten Indikatoren zu identifizieren. In diesem Arbeitspaket sollen verschiedene Verfahren entwickelt und getestet werden, um an diese Indikatoren zu gelangen. Dabei wird von bereits vorhandenen Standard-Formaten und den darin berücksichtigen Verhaltensmustern ausgegangen und dann sukzessive nach weiteren charakteristischen und im Feld leicht zu messenden Verhaltenscharakteristika gesucht. Ziel dabei ist, eine möglichst kompakte Menge an Indikatoren zu bestimmen, die Ressourcenschonendend zur Erkennung eingesetzt werden können und dabei eine möglichst gute Trefferrate mit sich bringen.

AP.2.2 Anbindung und Korrelation von externen Datenquellen

Dauer: 2 PM

Neben den generierten Analysedaten gibt es eine hohe Anzahl von externen Datenquellen, welche die Indikatoren verbessern können und die Qualität der Bewertung erheblich steigern. In diesem Arbeitspaket sollen mögliche externe Datenquellen recherchiert und bewertet werden. Darauf aufbauend sollen diese Datenquellen in unser System integriert werden. Die Daten sollen so angereichert werden, dass es immer noch für darauf aufbauende Verfahren, sowie den Analysten möglich ist, den Ursprung festzustellen.

Durch die Anreicherung der Daten sollen auch die Korrelationen zwischen den einzelnen Indikatoren verbessert werden. Angreifer verwenden sehr oft, gleiche Vorgehensweisen, aber unterschiedliche Programme dies zu erreichen. Durch die Korrelationen sollen Ähnlichkeiten dieser Angriffe erkannt werden und möglicherweise dem Analysten bei der Attribution geholfen werden.

AP.2.3 Regelbasierte Filterung und Verfeinerung der Indikatoren

Dauer: 2 PM

Auf Grund der Art und Weise, wie die Indikatoren gewonnen werden - nämlich durch Schadsoftware-Analyse - sind alle resultierenden Indikatoren auf eine konkrete Schadsoftware beziehbar. Mit zunehmend besser funktionierender Verteilung und Austausch von Indikatoren stehen Sicherheitsabteilungen häufig vor der Frage ob ihr Unternehmen von der Bedrohung, die hinter den vorliegenden Indikatoren steckt, betroffen sind. Aus diesem Grund ist es wichtig effiziente Möglichkeiten zur Filterung und zur regelbasierten Alarmierung bereitzustellen. Zu diesem Zweck wir ein Rahmenwerk für die Hinterlegung von entsprechenden Regeln implementiert. Dabei müssen Kontext abhängige Verallgemeinerungen (z.B. Finden ähnlicher Treffer) berücksichtigt werden und auch das Ableiten von gemeinsamen Indikatoren (inkl. konfigurierbarer Unscharfe¨ Hinsichtlich der Ähnlichkeit) für eine vor ausgewählte Menge an Schadsoftware soll unterstutzt¨ werden. Auf diese Weise wird es möglich von beobachteten Indikatoren auf konkrete Schadsoftware zurück zu schließen und Indikatoren zu finden, die als Hinweis für eine ganze Gruppe von Schadsoftware dienen.

AP.2.4 Methoden zur effizienten Abfrage der erstellten Indikatoren

Dauer: 2 PM

Die erzeugten Indikatoren sollen in einem Datenbanksystem verwaltet, aufbereitet und effizient für die Suche nach Schadcode abgefragt werden können. Um den unterschiedlichsten Konsumenten diese Daten in der jeweils benötigten Form zur Verfügung stellen zu können, muss das System nicht nur effizient, sondern vor allem flexibel sein. So können im einfachsten Fall die ermittelten Indikatoren sofort in eine statische Exportdatei mit festem Format geschrieben werden. In anspruchsvolleren Szenarien kann es notwendig sein, dass die weiterverarbeitende Instanz das Format und den Umfang der benötigten Indikatoren selbst

bestimmt. Außerdem sollen Ansätze verfolgt werden, die inkrementelle Updates und nachjustierende Verfahren unterstutzen¨.

TP.5 Evaluation und Testbetrieb

Dauer: 2 PM

Die entwickelten Techniken zur Verhaltensanalyse von Schadcode sollen in diesem Teilprojekt entsprechend verschiedener Verwertungsszenarien implementiert und evaluiert werden. Hierbei soll die Erkennungsleistung der generierten Indikatoren, die Genauigkeit der Anomalieerkennung und des Clusterings sowie die Gute¨ der Visualisierung unter realen Bedingungen untersucht werden. In enger Zusammenarbeit mit den drei CERT-Verbundpartnern soll hier das Potential der entwickelten Technologie bei der Bekämpfung von modernem Schadcode erprobt werden.

AP.5.3 Evaluation und Testbetrieb mit den abgeleiteten Indikatoren

Dauer: 1 PM

Die Qualität der abgeleiteten Indikatoren wird durch deren Anwendbarkeit im Echtbetrieb bestimmt. Wichtig ist dabei, ob und wie weit die Indikatoren von bestehenden Sicherheits-produkten verarbeitet werden können, wie aufwendig deren Integration, und wie effektiv der spätere Einsatz ist. Des Weiteren muss die Anwendung in einer sehr geringen Fehlerrate resultieren, so dass maximal viele Schadcodes erkannt und abgewehrt und nur minimal viele gutartige Dateien fälschlicherweise als bösartig erkannt und blockiert werden bzw. zu Fehlalarmen fuhren¨. Neben der engen Zusammenarbeit mit den CERT-Teams im Feld, werden in diesem Arbeitspaket umfangreiche Testreihen mit gutartigen und bösartigen Dateien im Labor durchgeführt werden.

AP.5.4 Datenschutzrechtliche Prüfung im Feld

Dauer: 1 PM

In diesem Arbeitspaket werden die Verfahren und erhobenen Daten datenschutzrechtlich bewertet. Innerhalb von VAMOS werden möglicherweise personenbezogenen Daten verarbeitet. Dies ist möglicherweise nicht zu verhindern, weil sonst die Qualität der Erkennung durch die Indikatoren nicht gewährleistet ist. Aus diesem Grund ist diese Bewertung wichtig und es sollen Losungen¨ entwickelt werden, wie es dennoch möglich ist, die Ansätze Datenschutzkonform zu verwenden.

Eingehende Darstellung

Der folgende Abschnitt beschreibt alle Ergebnisse im Zusammenhang mit dem Beitrag von Siemens. Wie im Arbeitsplan dargestellt, beziehen sich die wichtigsten Erfolge auf die Verwaltung der Kompromitierungsindikatoren und insbesondere auf deren Verarbeitung und Filterung, bevor diese unter den Partnern geteilt werden können.

TP.1 Anforderungsanalyse und Datenerhebung

Die zu Beginn des Projekts durchgeführte Anforderungsanalyse konzentrierte sich auf die Gestaltung der Architektur, die später darauf abzielte, Kompromitierungsindikatoren (IoCs) zu erstellen und zu verarbeiten. Diese Architektur wurde kontinuierlich verbessert und vor allem in die bereits vorhandenen Siemens-Toolchains für die Datenerfassung und -speicherung integriert. Die endgültige Architektur, die Siemens für das VAMOS-Projekt eingerichtet hat, umfasst zwei Hauptkomponenten:

- Eine MISP-Instanz - MISP ist eine Open-Source-Plattform, die vom Computer Incident Response Center Luxemburg (CIRCL) entwickelt und unterstützt wird [1]0. Die MISP-Instanz (exklusiv für das VAMOS-Projekt bereitgestellt) ist die Kernkomponente der Architektur und für das Speichern, Korrelieren und Freigeben von Indikatoren verantwortlich.
- CMAP - Das „Corporate Malware Analysis Portal" ist eine von Siemens entwickelte Plattform, die von Siemens CERT zur Analyse von Malware-Beispielen verwendet wird. Ursprünglich basierend auf der Cuckoo-Sandbox [3], hat sich CMAP im Laufe der Zeit weiterentwickelt und ermöglicht es nun, heterogene Informationen von externen Plattformen (z. B. VMRay) zu analysieren und Teile des Schadcodes gründlich zu analysieren, um nützliche Erkenntnisse (z. B. mögliche Indikatoren für Malware) und Klassifizierungen (z. Zu welchen Malware-Familien eine böswillige ausführbare Datei gehören könnte).

Eine grundlegende Übersicht über die wichtigsten Architekturkomponenten ist in Abbildung 1 dargestellt.

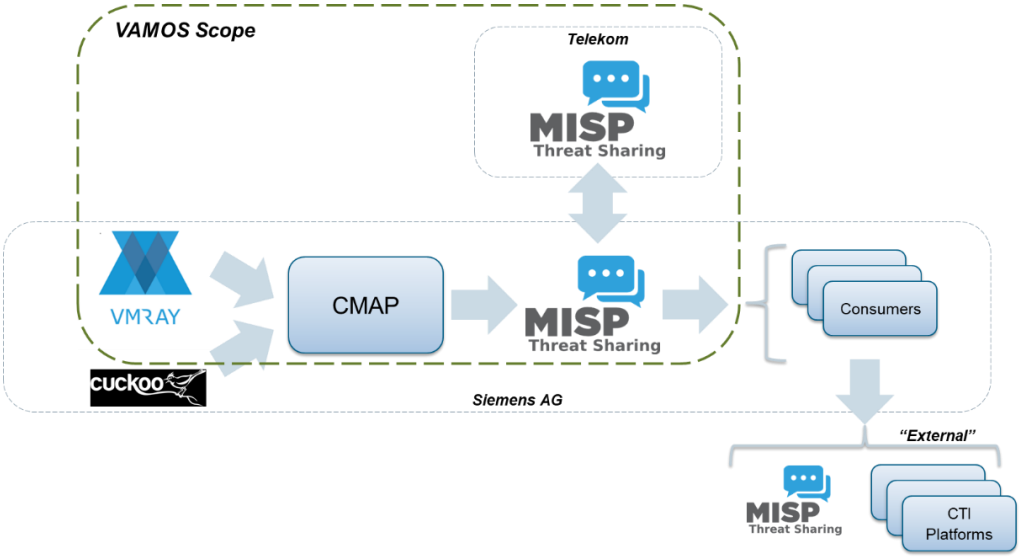

Abbildung 1 - Architekturübersicht

Die von CMAP bereitgestellten Ergebnisse werden über benutzerdefinierten Code, der auf den verfügbaren MISP-APIs entwickelt wurde, an MISP gesendet. Dieser Prozess ist vollständig automatisiert und hat sich im Laufe des VAMOS-Projekts zu einem integralen Bestandteil des gesamten CTI-Handhabungsprozesses bei Siemens CERT entwickelt.

In der Siemens MISP-Instanz gespeicherte Daten wurden offen gestellt und standen allen Projektpartnern zum Abrufen zur Verfügung. Die Verbindung zwischen dieser MISP-Instanz und der bei der Telekom laufenden Haupt-MISP-Plattform wurde im letzten Projektjahr getestet, um die Richtigkeit der Push- und Pull-Mechanismen für die Übertragung von CTI-Informationen sicherzustellen.

TP.2 Ableitung und Filterung von Indikatoren

Ein Hauptziel des VAMOS-Projekts war es, automatisch wertvolle Erkenntnisse aus der Analyse bösartiger Codebeispiele zu gewinnen. Dieser Prozess, der auf der oben genannten CMAP-Plattform stattfindet, umfasst mehrere Aufgaben. Beispielsweise spielt ein vorläufiger „Filterungsschritt" eine wichtige Rolle, um das Vorhandensein von Indikatoren zu vermeiden, die sich auf bekannte harmlose Dienste beziehen. Da mehrere Malware-Verbindungen möglicherweise Verbindungsprüfungen mit weltweit verfügbaren Internetressourcen (z. B. google.com) durchführen, ist eine Reinigungstechnik erforderlich, mit der die zugehörigen Indikatoren als "falsch positiv" gekennzeichnet werden, bevor sie MISP erreichen. In dieser Hinsicht stützt sich die Implementierung von Siemens auf die Alexa Top 1Mio-Liste von Websites [4] (weitgehend in der Literatur verwendet, wie in [5] und [6] diskutiert), um bekannte URLs zu vermeiden, sowie auf eine Liste privater URLs, um eine Unterbrechung der internen Dienste zu verhindern. Eine andere Aufgabe überprüft die Neuheit der zu analysierenden Indikatoren und stellt sicher, dass alle in der Vergangenheit analysierten Schadcodes, die dieselben (oder manchmal ähnliche) Indikatoren enthalten, korreliert / referenziert werden. Schließlich konzentriert sich die wichtigste Aufgabe auf die Klassifizierung und „Kennzeichnung" von Schadcodebeispielen auf der Grundlage einer Reihe von „genau definierten Regeln". Diese Aufgabe wurde mit YARA umgesetzt.

YARA [7] ist ein weit verbreitetes Open-Source-Tool zur Identifizierung und Beschreibung von Malware-Typen basierend auf erkennbaren Text- und Binärmustern [8]. Das Tool arbeitet mit einem regelbasierten Ansatz, bei dem eine Reihe von Mustern sowie bestimmte Bedingungen und Boolesche Ausdrücke die Merkmale einer bekannten Malware eindeutig identifizieren. Durch die Nutzung von YARA-Funktionen (z. B. Tags, Metadaten) wird diese Identifikation mit Informationen angereichert, die verwendet werden können, um später eine Landschaft von Malware-Familien zu erstellen, die hinsichtlich ihrer Eigenschaften und ihres allgemeinen Verhaltens gruppiert sind.

Abbildung 2 zeigt, wie sich YARA in die vorhandene Gesamtarchitektur integriert, um Kompromitierungsindikatoren zu verarbeiten und auszutauschen.

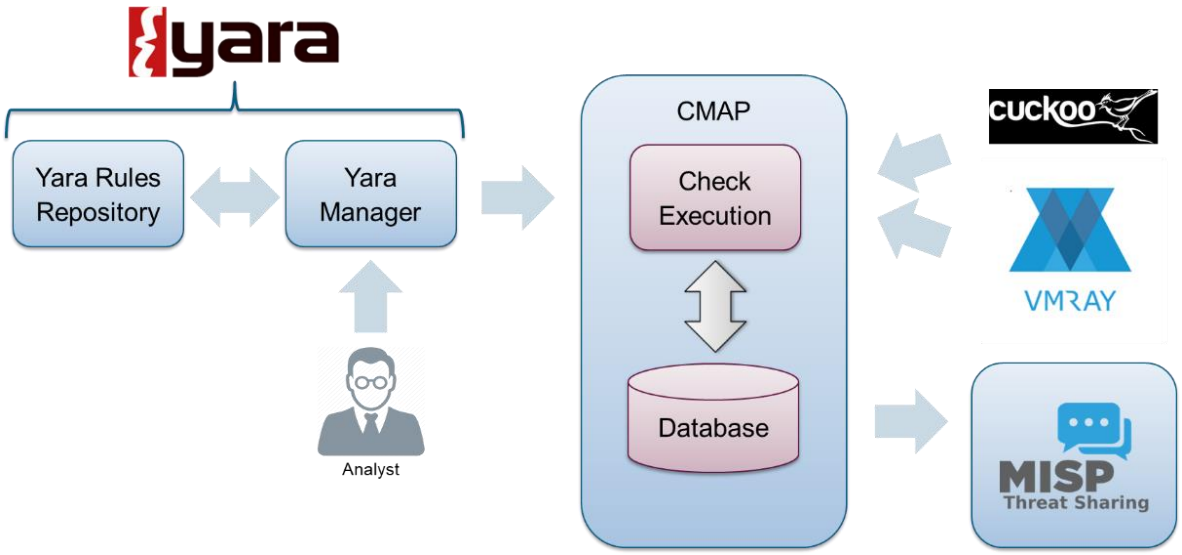

Abbildung 2 - Komplette Architekturübersicht

Nach Abschluss des Analyseprozesses werden alle extrahierten Erkenntnisse sowie das ursprüngliche Malware-Beispiel in CMAP gespeichert. Die Art und Weise, wie alle Daten gespeichert werden, wurde mit allen Projektpartnern besprochen und vereinbart, um ein Informationsschema zu definieren, das eine effiziente Abfrage und Datenabfrage ermöglicht. Der unternommene Ansatz sieht zwei interne Hauptverzeichnisverzeichnisse vor. Die erste dient als Sammelstelle für alle Endergebnisse der Analyse (z. B. Indikatoren, Etiketten), während die zweite die tatsächlichen Malware-Beispiele enthält und eine effizientere Möglichkeit zum Umgang mit größeren Dateien (z. B. Systemabbildern) bietet.

Die Analyseergebnisse sind nach drei verschiedenen Informationsgruppen gegliedert:

- Die „Sample" -Gruppe enthält Hashes des ursprünglichen Schadcodes (z. B. MD5, SHA1, SHA256 usw.) sowie eine Liste der aus der Analyse stammenden Bezeichnungen (z. B. Ergebnisse der auf YARA-Regeln basierenden Klassifizierung).
- Die Gruppe „Dropped-file" sammelt Informationen zu allen Dateien, die von dem Schadcode verarbeitet werden. Dies schließt alle Dateien ein, die während der Ausführung von Malware geöffnet, geschrieben oder geändert wurden

- Schließlich sammelt die Gruppe „Network“ Informationen zu allen eingehenden und ausgehenden Verbindungen, die während der Ausführung von Malware beobachtet wurden. Dies umfasst Hostnamen und IP-Adressen sowie detailliertere Informationen wie Netzwerkpaket-Header und Nutzdaten (z. B. innerhalb einer HTTP-Verbindung, die verwendeten Methoden, Benutzeragenten usw.).

Der Zugriff auf alle in CMAP gespeicherten Informationen erfolgt über eine benutzerdefinierte Schnittstelle. Die benutzerdefinierte Benutzeroberfläche implementiert die Standardbenutzerverwaltung sowie die Authentifizierung und ermöglicht ein effizientes Durchsuchen von Daten.

Die Ergebnisse einer Suche in CMAP werden in einem JSON-Format (wie in Abbildung 3 dargestellt) bereitgestellt und können exportiert werden.

Abbildung 3 - Beispiel für eine Suche in CMAP (nach den Schlüsselwörtern "Crypto" und "Packers")

Darüber hinaus überwacht die Schnittstelle den Code, der die Interaktionen zwischen CMAP und MISP implementiert, und überwacht so die Erstellung spezifischer MISP-Ereignisse für jede analysierte Malware. Dieser Prozess ermöglicht es schließlich, den Partnern alle Ergebnisse und Informationen über die von MISP bereitgestellten Standardfreigabeschemata zur Verfügung zu stellen.

Der Synchronisationsprozess zwischen CMAP und MISP ist vollständig automatisiert und erstellt eine Reihe von Querverweisen zwischen den auf den beiden Plattformen gespeicherten Informationen. Aufgrund der unterschiedlichen Darstellung der Informationen auf den Plattformen mussten jedoch einige Techniken entwickelt werden, um eine ordnungsgemäße Ausrichtung sicherzustellen. Ein wichtiger Aspekt dieses Prozesses betrifft die Umwandlung von YARA-Labels in MISP-Tags.

YARA bietet zwei Möglichkeiten, um Malware zu kennzeichnen, die einer bestimmten Regel entspricht: Tags und Metadaten. Tags werden häufig verwendet, um Klassen von YARA-Regeln zu identifizieren (z. B. alle Regeln, die mit "Ransomware" übereinstimmen). Metadaten bereichern stattdessen jede Regel mit zusätzlichen Informationen in Form von Schlüssel-Wert-Paaren (z. B. einem Namespace für Sensitivitätsmessungen, der verwendet wird, um die Vertraulichkeit einer Regel auszudrücken, sowie mit übereinstimmender Malware wie TLP und dem zugehörigen Wert als "rot"). Wenn eine Malware mit einer YARA-Regel übereinstimmt, müssen Informationen, die in den Tags und Metadaten der Regel enthalten sind, an MISP weitergeleitet werden. Dies kann auf verschiedene Weise erreicht werden, indem unterschiedliche Strategien verfolgt werden. Innerhalb des VAMOS-Projekts bestand die Strategie der Wahl darin, das MISP-Tagging-System zu nutzen und alle YARA-Tags und Metadaten benutzerdefinierten MISP-Tags zuzuordnen.

Jedes MISP-Tag wird durch drei Elemente eindeutig identifiziert: ein "Namensschema", ein "Prädikat" und einen "Wert". Die oben erwähnte Zuordnungsstrategie muss alle von einer YARA-Regel übertragenen Informationen den drei Elementen zuordnen, aus denen ein MISP-Tag besteht. Innerhalb von VAMOS wurden vier Techniken implementiert und evaluiert, um diese Zuordnung zu implementieren: "regelbasiert", "codebasiert", "in sich geschlossen" und "hybrid". Jede Technik hat ihre Vor- und Nachteile (insbesondere in Bezug auf Benutzerfreundlichkeit und Datenabruf). Alle Regeln haben ein gemeinsames Namensschema, nämlich „vamos“, da wir gemäß den Partnern entschieden haben, dass ein eindeutiges Schema zum Sammeln aller MISP-Tags unter dem Dach des Projekts die beste Wahl ist, um das Tagging-System schließlich für die Community öffentlich zu machen von MISP-Benutzern. Aufgrund dieser Auswahl waren die verfügbaren Elemente zum Speichern der Informationen aus der YARA-Klassifizierung und -Kennzeichnung nur „Prädikate“ und „Werte“. In den folgenden Abschnitten beschreiben wir jede Mapping-Technik einzeln.

Der regelbasierter Ansatz

Ziel der regelbasierten Technik ist es, die Aufmerksamkeit eines MISP-Benutzers auf YARA-Regeln zu lenken. Für jedes MISP-Ereignis, das als Antwort auf eine Analyse erstellt wurde, stellt diese Technik sicher, dass alle übereinstimmenden YARA-Regeln auf den MISP-Tags explizit sichtbar sind. Aus diesem Grund wird der Name der Regel als Prädikat der Tags an MISP übertragen. Alle anderen Informationen werden in den Werten der Tags unter Verwendung eines Schemas gespeichert, das die Initialen der YARA-Metadaten und die zugehörigen Werte verkettet (z. B. würden sich die YARA-Metadaten "category = ransomware" in "c_ransomware" verwandeln). Besonders wichtige Metadaten wie "TLP" bleiben "wie sie sind", ohne Initialen zur Verbesserung der Lesbarkeit zu verwenden (z. B. würden sich YARA-Metadaten "tlp = red" in "tlp_red" verwandeln). Schließlich werden alle MISP-Tags unter dem Schema "t_ <Tag>" ausgedrückt.

Fields	Rule name	Rule Filename	Category	TLP	Source	Tags
Example	invalid_trailer_structure	Maldoc_PDF.yar_compiled	Malicious_Documents	White	yara_rules_project	“PDF”, “raw”

Template → vamos:”rule”=“attr.”

Results →
vamos:invalid_trailer_structure =“f_Maldoc_PDF.yar_compiled”
vamos:invalid_trailer_structure =“c_malicious_document”
vamos:invalid_trailer_structure =“tlp_white”
vamos:invalid_trailer_structure =“s_yara_rules_project”
vamos:invalid_trailer_structure =“t_PDF”
vamos:invalid_trailer_structure =“t_raw”

Abbildung 4 - Ergebnisse des regelbasierten Ansatzes

Wie bereits erwähnt, konzentriert sich die „regelbasierte“ Technik auf die Lesbarkeit (ein MISP-Benutzer ist sich immer sofort der YARA-Regeln bewusst, die mit der im MISP-Ereignis referenzierten Malware übereinstimmen), kann jedoch unter Namenskollisionen leiden. Tatsächlich sind YARA-Regelnamen nicht eindeutig, und zeitgemäße Übereinstimmungen von Regeln mit denselben Namen können zu Unklarheiten und Verwirrung darüber führen, welche Tags zu welcher Regel gehören.

Der codebasierte Ansatz

Die "codebasierte" Technik verwendet den gleichen Ansatz wie die "regelbasierte", versucht jedoch, die durch Regelnamen verursachten Mehrdeutigkeiten zu lösen. Bei dieser Technik wird der Name der YARA-Regel, der für die Prädikate der MISP-Tags verwendet wird, durch einen Code ersetzt, der jede in CMAP verwendete (und gespeicherte) YARA-Regel eindeutig identifiziert.

Fields	Rule name	Rule Filename	Category	TLP	Source	Tags
Example	invalid_trailer_structure	Maldoc_PDF.yar_compiled	Malicious_Documents	White	yara_rules_project	“PDF”, “raw”

Template → vamos:”rule id”=“attr.”

Results →
vamos:rule1432 = “r_invalid_trailer_structure”
vamos:rule1432 = “f_Maldoc_PDF.yar_compiled”
vamos:rule1432 = “c_malicious_document”
vamos:rule1432 = “tlp_white”
vamos:rule1432 = “s_yara_rules_project”
vamos:rule1432 = “t_PDF”
vamos:rule1432 = “t_raw”

Abbildung 5 - Ergebnisse des codebasierten Ansatzes

Obwohl die Technik das Problem löst, das durch die Verwendung expliziter Regelnamen entsteht, wurde der Nachteil des Speicherns nutzloser Informationen wie der CMAP-YARA-Codes

von MISP-Benutzern während der Tests im Allgemeinen negativ bewertet. Einer der Vorteile der Verwendung von MISP liegt in der übersichtlichen Oberfläche, über die alle wichtigen Informationen schnell auf die Benutzer aufmerksam gemacht werden. Das Füllen von MISP-Ereignissen mit unbrauchbaren Daten würde sich definitiv auf dieses Prinzip auswirken und die allgemeine Benutzerfreundlichkeit verringern.

Der „in sich geschlossen" Ansatz

Mit der "in sich geschlossenen" Technik versuchen wir, das Problem der Mehrdeutigkeiten von Regelnamen zu lösen, ohne bedeutungslose Codes zu verwenden. Der Ansatz funktioniert dank des Prädikats eines komplexeren Tags. Wenn für die beiden Techniken vor jedem Prädikat ein einzelner Begriff war (entweder der YARA-Regelname oder der Regelcode), werden Prädikate in der "in sich geschlossenen" Technik zu einer Folge von YARA-Tags und Metadaten. Insbesondere verknüpft das Prädikat eines MISP-Tags den YARA-Regelnamen (z. B. "invalid_trailer_structure"), den Dateinamen, in dem die Regel gespeichert ist (z. B. "Maldoc_PDF.yar_compiled"), die Regelkategorie (z. B. "Malicious_Documents") und die TLP-Level (z. B. „Weiß"). Diese vier Elemente werden einfach unter Verwendung des Zeichens "/" verkettet (z. B. "invalid_tralier_structure / Maldoc_PDF.yar_compiled / Malicious_Documents / White"). Schließlich werden alle anderen YARA-Tags und Metadaten einfach als MISP-Tag-Werte verwendet, um die MISP-Tags zu vervollständigen.

Fields	Rule name	Rule Filename	Category	TLP	Source	Tags
Example	invalid_trailer_structure	Maldoc_PDF.yar_compiled	Malicious_Documents	White	yara_rules_project	"PDF", "raw"

Template → vamos:"rule + attr."="attr."

Results →
vamos:invalid_trailer_structure/Maldoc_PDF.yar_compiled/Malicious_Documents/White ="yara_rules_project"
vamos:invalid_trailer_structure/Maldoc_PDF.yar_compiled/Malicious_Documents/White ="PDF"
vamos:invalid_trailer_structure/Maldoc_PDF.yar_compiled/Malicious_Documents/White ="raw"

Abbildung 6 - Ergebnisse des "in sich geschlossen" Ansatzes

Als Kompromiss zwischen den "regelbasierten" und den "codebasierten" Techniken arbeitet die "in sich geschlossene" Technik daran, Mehrdeutigkeiten zu lösen, ohne dass nutzlose Informationen hinzugefügt werden. Die Lesbarkeit der Prädikate von MISP-Tags nimmt jedoch drastisch ab, sodass MISP-Benutzer mit langen Folgen von Begriffen umgehen müssen, die bei der Analyse von MISP-Ereignissen berücksichtigt werden müssen. Es ist auch erwähnenswert, dass mit diesem Ansatz die Anzahl der Tags pro Regel abnimmt (da ein Teil der Informationen an anderer Stelle im Tag übertragen wird), aber mehrere Übereinstimmungsregeln mit überlappenden Tags eine insgesamt zunehmende Anzahl von MISP-Tags verursachen würden (z. Zwei mit „PDF" gekennzeichnete YARA-Regeln, die mit derselben Malware übereinstimmen, würden zu zwei nahezu identischen Tags führen, die sich möglicherweise nur in einem kleinen Abschnitt des Prädikats des MISP-Tags unterscheiden. Dieses letzte Problem kann die Nützlichkeit von MISP-Tags im Laufe der Zeit drastisch verringern.

Der hybride Ansatz

Schließlich implementiert die „Hybrid" -Technik einen völlig anderen Ansatz als die vorherigen drei, wobei die semantische Bedeutung der Prädikate der MISP-Tags grundlegend geändert wird. Anstatt MISP-Tags in den YARA-Regeln (z. B. Namen oder Codes) zu "schwenken", wird der Fokus auf die tatsächlichen Tags und Metadaten verschoben. Tatsächlich verwenden wir die Prädikate, um die Schlüssel der YARA-Metadaten (z. B. "Kategorie", "TLP") und die entsprechenden Werte als Werte für MISP-Tags (z. B. "Malicious_Documents", "Weiß") zu speichern. Zusätzliche Informationen wie Regelnamen und Regeldateien sind in MISP-Tags mit identischer Struktur enthalten (z. B. "rule =" invalid_trailer_structure "). Gleiches gilt für YARA-Tags, die das feste Prädikat "Tag" verwenden (z. B. "Tag =" PDF ").

Fields	Rule name	Rule Filename	Category	TLP	Source	Tags
Example	invalid_trailer_structure	Maldoc_PDF.yar_compiled	Malicious_Documents	White	yara_rules_project	“PDF”, “raw”

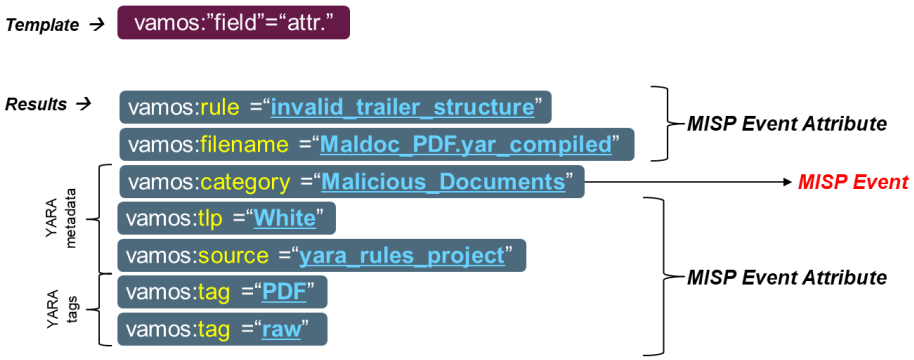

Abbildung 7 - Ergebnisse des hybriden Ansatzes

Mit der "Hybrid" -Technik erhöht sich die Lesbarkeit der MISP-Tags in Bezug auf die "Code-basierten" und "in sich geschlossenen" Ansätze (was die "Regel-basierten" betrifft, ist dies vergleichbar, was möglicherweise davon abhängt, was ein MISP-Benutzer tut schaut nach). Was die Anzahl der Tags betrifft, verbessert die Technik die Situation drastisch, indem einfache Überlappungen zugelassen werden (z. B. würden zwei mit "PDF" gekennzeichnete YARA-Regeln nur ein Tag in Form von "Tag =" PDF "erzeugen).

Schlussfolgerungen

In Übereinstimmung mit den anderen Partnern war der Hybridansatz aufgrund der im vorherigen Abschnitt erörterten Vorteile letztendlich die Technik der Wahl. Neben der Tag-Struktur zeigte eine spätere Studie zum Anwenden von Tags auf MISP-Ereignisse den Vorteil der Verteilung der Informationen, sodass die wichtigsten Tags für einen Benutzer sofort sichtbar waren. Wie in Abbildung 7 dargestellt, wurden Tags, die sich auf YARA-Regelkategorien beziehen, als Haupt- und aussagekräftigste Information identifiziert und somit direkt auf das gesamte MISP-Ereignis angewendet (Abbildung 8).

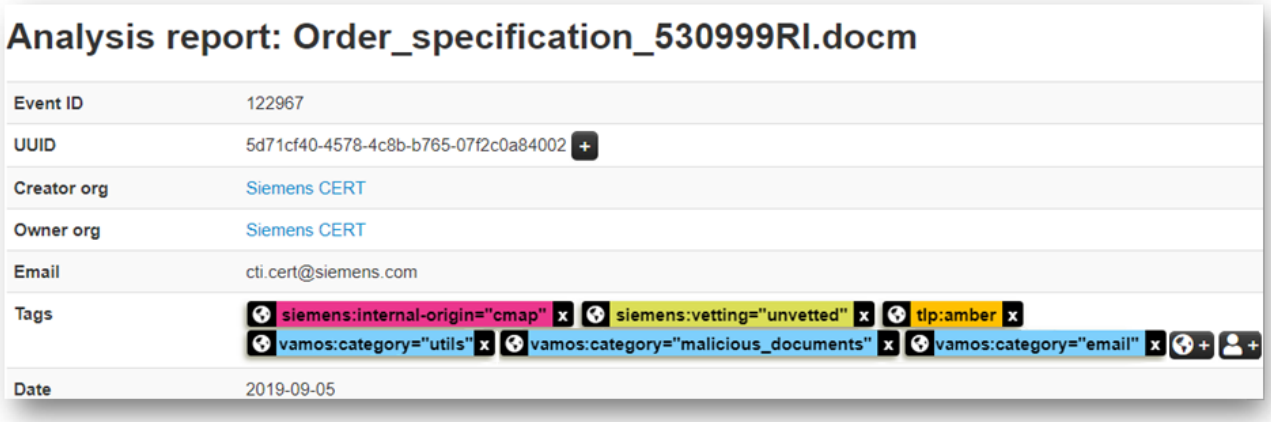

Abbildung 8 - MISP-Ereignis mit benutzerdefinierten Tags (1/2)

Alle anderen Tags wurden gruppiert und auf das benutzerdefinierte Attribut "YARA Results " angewendet, das im MISP-Ereignis selbst enthalten ist (Abbildung 9).

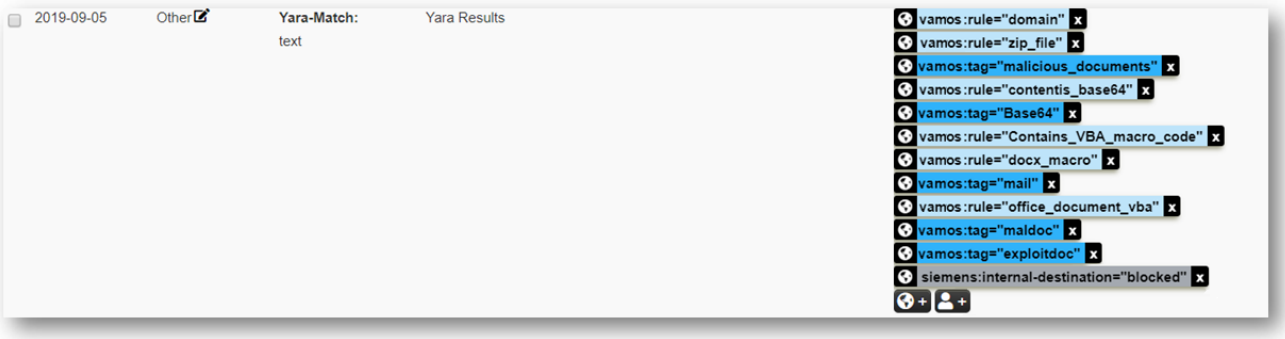

Abbildung 9 - MISP-Ereignis mit benutzerdefinierten Tags (1/2)

TP.5 Evaluation und Testbetrieb

Wie in TP.1 erläutert, wurde die gesamte Werkzeugkette während der Projektlaufzeit getestet. Die Kerntests umfassten die folgenden vier Aktivitäten:

1. Die Erstellung neuer YARA-Regeln und deren Bereitstellung in CMAP

2. Kennzeichnung neuer Malware-Beispiele und Auswertung der Ergebnisse (z. B. Analyse von falsch positiven und falsch negativen Ergebnissen)

3. Senden von gekennzeichneten Informationen an MISP und Anwenden der zugehörigen MISP-Tags (gemäß den Strategien der Wahl)

4. Weiterleitung dieser Informationen an die Hauptinstanz (zentral) der MISP bei der Telekom und damit Weitergabe der Informationen an alle VAMOS-Partner

Parallel dazu trug Siemens zur Gesamtsammlung von Schadcodebeispielen bei, die von VMRay und der Technischen Universität Braunschweig zur Analyse verwendet wurden, indem Tausende von Malware zur Verfügung gestellt wurden, die zuvor in den internen CERT-Datensätzen von Siemens gespeichert waren.

Die Testergebnisse wurden mit den Partnern besprochen und der Gesamtprozess des Austauschs von Bedrohungsinformationen wurde bewertet und mit deren Feedback verbessert.

Zusätzlich zu den Kerntests haben wir die Toolkette in die interne Infrastruktur von Siemens integriert und deren Einsatz in breiteren Szenarien wie den Siemens-Prozessen zur Reaktion auf Vorfälle weiter getestet. Diese zusätzlichen Tests halfen bei der Bewertung der Verwendbarkeit der generierten Ergebnisse sowie ihres Mehrwerts in realen betrieblichen Anwendungsfällen. In dieser Hinsicht wurde die Werkzeugkette erfolgreich eingesetzt mit:

- Siemens Vetting Platform - wird verwendet, um verfügbare Informationen zu Bedrohungsinformationen weiter anzureichern und die Gesamtinformationen halbautomatisch auszuwerten (z. B. Sicherheitsexperte, der die Korrelationsergebnisse überwacht und zusätzliche Erkenntnisse liefert

- Siemens Security Monitoring & Analytics Platform - wird verwendet, um Korrelationen und Beziehungen von Sicherheitsereignissen über grafische Darstellungen und Analysen (basierend auf Open Source-Technologien und -Standards wie Neo4j [9] [10] und STIX [11]) hervorzuheben

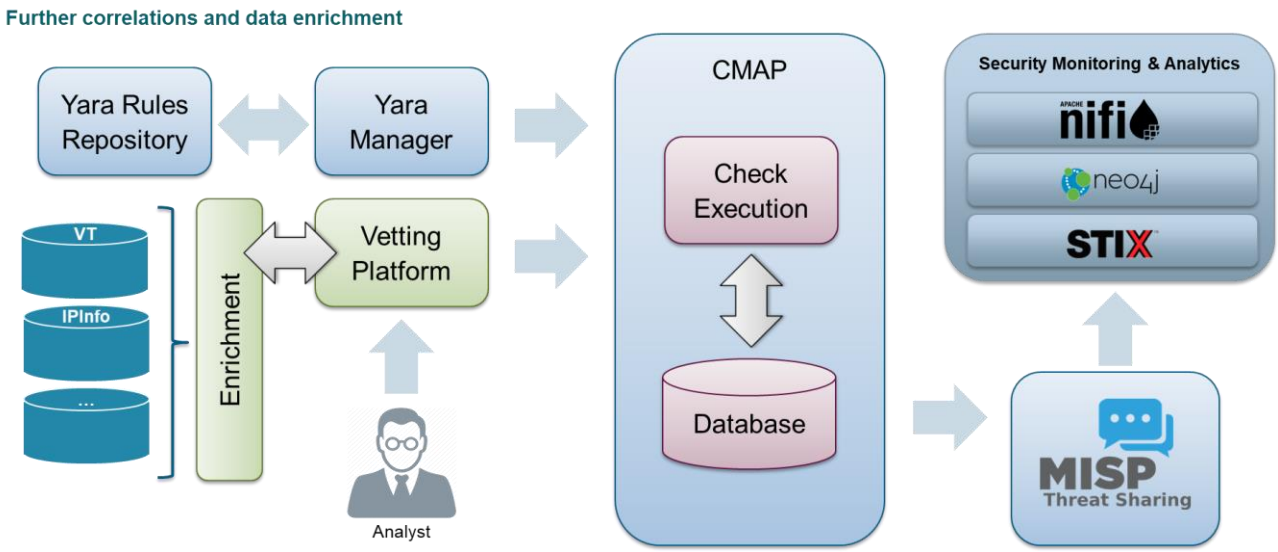

Abbildung 10 - Vollständige Pipeline mit VAMOS-Technologien und -Ergebnissen

Diese Plattformen nahmen Daten auf, die direkt von der VAMOS MISP-Instanz stammen (und somit nach der Verarbeitung in CMAP), und haben die Ergebnisse zur weiteren Freigabe direkt an MISP zurückgesendet. In diesem Fall wurde die Infrastruktur so instrumentiert, dass die Informationen aufgrund des Vorhandenseins vertraulicher Daten (z. B. laufende Analyse

von Cyber-Vorfällen, Siemens-Kundendaten usw.) nicht außerhalb von Siemens weitergeleitet werden.

Um die Benutzerfreundlichkeit weiter zu verbessern (und die Projektergebnisse außerhalb des VAMOS-Konsortiums zu fördern), definierte Sie-mens eine offene Vorlage für Vorschläge (im MISP-Standardformat) für den Austausch von Ergebnissen aus der Analyse von VMRay und der Technischen Universität Braunschweig. Sobald weitere Informationen zu einem Malware-Beispiel verfügbar sind (z. B. Ergebnisse der von der Technischen Universität Braunschweig durchgeführten Clusteranalyse), erfordert eine effektive Freigabe eine strukturierte Darstellung aller Analysedetails. Um dies zu erreichen, haben wir in Übereinstimmung mit allen Partnern eine neue JSON-Vorlage definiert, um diese Informationen über MISP in Form eines neuen „MISP-Ereignisses" zu übertragen. Dieses neue Format wurde der gesamten Sicherheitsgemeinschaft öffentlich zugänglich gemacht und kann unter [12] abgerufen werden. Abbildung 11 zeigt einen Auszug der im benutzerdefinierten MISP-Ereignis verfügbaren Informationen.


```
{
  "required": [
    "sandbox-type"
  ],
  "requiredOneOf": [
    "web-sandbox",
    "on-premise-sandbox",
    "saas-sandbox"
  ],
  "attributes": {
              ...
    },
    "score": {
              ...
    },
    "results": {
              ...
    },
    "raw-report": {
              ...
    },
    "sandbox-file": {
              ...
    },
    "sandbox-type": {
              ...
    },
    "on-premise-sandbox": {
              ...
    },
    "web-sandbox": {
              ...
    },
    "saas-sandbox": {
              ...
    }
  },
  "version": 2,
  "description": "Sandbox report",
  "meta-category": "misc",
  "uuid": "4d3fffd2-cd07-4357-96e0-a51c988faaef",
  "name": "sandbox-report"
}
```


Abbildung 11 - Auszug aus dem benutzerdefinierten MISP-Ereignis

Schließlich haben wir uns im vergangenen Jahr auch auf den Datenschutz konzentriert. In diesem Zusammenhang haben wir die Herausforderungen und Kritikpunkte beim Umgang mit und beim Austausch von Informationen über Malware untersucht, die an Cybersicherheitsvorfällen beteiligt sind. Diese Aktivität war hilfreich, um die internen Prozesse zum Schutz von Unternehmensdaten und deren Übereinstimmung mit den geltenden Gesetzen und Vorschriften zu bewerten.

Ausgangspunkt dieser Untersuchung war die Analyse der 2018 veröffentlichten Europäischen Allgemeinen Datenschutzverordnung (GDPR) [13]. Die im Dokument dargestellten allgemeinen und umfassenden Perspektiven sowie alle damit verbundenen Richtlinien mussten jedoch für die VAMOS-Anwendungsfälle instanziiert werden. In diesem Zusammenhang war der Hauptaspekt, der im Rahmen des Projekts erörtert wurde, die Rolle der Privatsphäre und des Datenschutzes beim Austausch von Bedrohungsinformationen. Dieses Thema stellt keinen umfassenden Stand der Technik dar, aber einige Arbeiten haben kürzlich Herausforderungen bei SOC- und CERT-Operationen erörtert, die sich aus Datenschutzanforderungen ergeben.

Im Allgemeinen beginnt die Analyse von Anforderungen und Best Practices zum Schutz personenbezogener Daten in einem bestimmten Anwendungsfall mit einer grundlegenden Beschreibung der personenbezogenen Daten selbst. Im Zusammenhang mit dem Austausch von Bedrohungsinformationen haben wir zwei Arten von persönlichen Informationen identifiziert, die möglicherweise Teil eines Informationspakets für Sicherheitsvorfälle werden können. Die erste Art stellt Informationen dar, die direkt mit einem bestimmten Sicherheitsvorfall verknüpft sind (Informationen als „Teil des Vorfalls"). Dies ist der Fall bei Daten, die an der Dynamik eines Cyberangriffs beteiligt sind (z. B. gestohlene Anmeldeinformationen, die möglicherweise mit einer physischen Person verknüpft sind) und wahrscheinlich erforderlich sind, um die Situation zu beschreiben, in der dieser Angriff durchgeführt wurde, und um einige seiner Eigenschaften eindeutig zu identifizieren (z. B. seine Zuordnung zu einem bestimmten Gegner oder einer bestimmten gegnerischen Gruppe). Die zweite Art stellt Informationen dar, die indirekt mit einem Sicherheitsvorfall verknüpft sind (Informationen, die als „mit dem Vorfall geteilt" bezeichnet werden). In diesem Fall beziehen sich die Informationen auf den umgebenden Kontext, in dem der Angriff ausgeführt wurde, anstatt Teil des Angriffs selbst zu sein. Ein prominentes Beispiel für diese Gruppe waren Daten über eine physische Person, die einen Sicherheitsvorfall gemeldet hat. Während in diesem letzten Fall die Notwendigkeit der Weitergabe personenbezogener Daten fraglich ist und von mehreren Faktoren abhängen kann, ist dies im ersten Fall meist unvermeidbar und muss entsprechend behandelt werden.

Basierend auf dieser Unterscheidung haben wir im letzten Jahr des VAMOS-Projekts das Thema interviewt und mit dem operativen Personal von Siemens diskutiert, um Szenarien zu beschreiben, in denen diese beiden Arten von Informationen in den Austausch von Bedrohungsinformationen involviert waren (oder gewesen sein könnten). Wir haben diese Interviews verwendet, um einige grundlegende Aspekte des Problems zu extrapolieren und schließlich einen Prozess zum Sammeln aller Informationen zu beschreiben, die für den Austausch von Bedrohungsinformationen von Siemens erforderlich sind, um GDPR-konform zu

werden. Dieser Prozess sieht detaillierte Beschreibungen von mindestens den folgenden Elementen vor:

- Zwecke, für die personenbezogene Daten gesammelt oder weitergegeben werden Siemens
- Kategorien aller gespeicherten Daten
- Beteiligte und Dienstleister
- Technische und organisatorische Maßnahmen zum Schutz dieser Daten
- Eine Folgenabschätzung zum Datenschutz
- Lösch- und Aufbewahrungsrichtlinien vorhanden

Für jede Aktivität zum Austausch von Bedrohungsinformationen werden die oben genannten Beschreibungen vereinbart und unter den beteiligten Stakeholdern geteilt.

Abschließende Überlegungen

Am Ende des Projekts wurden alle Ziele erfolgreich erreicht und die damit verbundenen Technologien und Ansätze auf einem zufriedenstellenden Niveau entwickelt. Wie im vorigen Kapitel angedeutet, wurden die Projektergebnisse nicht nur in mehreren Siemens-internen Veranstaltungen und Werkstätten vorgestellt und diskutiert, sondern auch in bereits vorhandene Toolchains und Prozesse integriert und aktiv verwendet. Im Folgenden beschreiben wir kurz die beobachteten Vorteile der in VAMOS entwickelten Kooperationen und Lösungen.

Verwertungsplan

Das Vorhaben leistet einen wichtigen Beitrag zur Bekämpfung von modernem Schadcode und gezielten Angriffen. Nur durch eine effiziente und vor allem automatisierbare Analyse von komplexen und spezialisierten Angriffen wird es möglich, Unternehmen, Staaten und deren Bürger zuverlässig zu schützen. Die in diesem Vorhaben entwickelten Verfahren und Methoden bringen der Firma Siemens AG ganz konkrete wirtschaftliche und technische Vorteile und können darüber hinaus der Anbahnung weiterer Kooperationen und Anschluss Projekte dienen.

Wirtschaftliche Verwertung

Gezielte Angriffe auf IT-Systeme von Unternehmen haben häufig das Ziel Daten aus dem Unternehmen zu entwenden. Die Anzahl der Falle¨ diese Art der Industriespionage steigt im letzten Jahrzehnt sehr stark und hat wirtschaftliche Auswirkungen auf die Industrie. Aus diesem Grund ist es wichtig, dass sich Unternehmen mit neuen Erkennungsmöglichkeiten für diese Angriffe beschäftigen und so schnell wie möglich integrieren.

Die in diesem Vorhaben entwickelnden Verfahren helfen der Firma Siemens AG bei der effizienteren Auswertung von Schadsoftware, der schnelleren Bewertung von Indikatoren, und zuletzt bei der Erkennung von neuen Angriffen gegen das Unternehmen. Unmittelbar können dadurch Angriffe schneller gestoppt werden und bei frühzeitiger Erkennung auch das Abfließen der Daten verhindert werden.

Weiterhin können von dem Vorhaben VAMOS viele andere Unternehmen in Deutschland profitieren, da das Siemens CERT sich aktiv in den Informationsaustausch mit diesen einbringt und dort gezielter gewonnene Informationen zu aktuellen Angriffen austauschen kann.

Technische Verwertung

Neben der Verbesserung der Schadsoftwareanalyse innerhalb von Siemens AG, wird eine effiziente Indikatoren Datenbank entwickelt. Die derzeitigen Lösungen erlauben nicht, eine Speicherung von großen Datenmengen und führen dazu, dass nur ein Bruchteil von Indikatoren gespeichert und verwendet wird.

Das hier entwickelte System ist daher ein wesentliches System, dass Sicherheitsteams in Zukunft mit diesen Daten umgehen können und auch bei der automatischen Erkennung und Analyse verwenden können. Durch diese Grundlagen können weitere Verfahren in der

Erkennung und Analyse entwickelt werden. Somit schafft das Projekt VAMOS den notwendigen Grundstein für weitere Forschungsarbeiten auf diesem Gebiet.

Anschluss Fähigkeit

Durch die Zusammenarbeit mit VMRay und den akademischen Experten erhofft sich die Siemens AG die entwickelten Systeme in ihre Infrastruktur zu integrieren und darüber hinaus weitere Anschluss Projekte durchzuführen. Durch die bereits vorhandene enge Zusammenarbeit im deutschen CERT-Verbund erhoffen wir uns zusätzlich, die entstandenen Erfahrungen auch anderen deutschen Unternehmen weiterzugeben.

Eine Fortführung der gemeinsamen Forschung durch den aktuellen Verbund ist sehr wahrscheinlich und im Kontext der Planungen des VAMOS-Projektes sind bereits jetzt schon diverse neue Ideen entstanden.

Literaturverzeichnis

[1] Wagner, C., Dulaunoy, A., Wagener, G., & Iklody, A. (2016, October). Misp: The design and implementation of a collaborative threat intelligence sharing platform. In Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security (pp. 49-56).

[2] "MISP - Open Source Threat Intelligence Platform & Open Standards For Threat Information Sharing" (Online). Verfügbar unter: https://www.misp-project.org/

[3] "Cuckoo Sandbox - Automated Malware Analysis" (Online). Verfügbar unter: https://cuckoosandbox.org/

[4] "Alexa top sites list" (Online) Verfügbar unter: https://www.alexa.com/topsites

[5] Englehardt, S., & Narayanan, A. (2016, October). Online tracking: A 1-million-site measurement and analysis. In Proceedings of the 2016 ACM SIGSAC conference on computer and communications security (pp. 1388-1401).

[6] Scheitle, Q., Hohlfeld, O., Gamba, J., Jelten, J., Zimmermann, T., Strowes, S. D., & Vallina-Rodriguez, N. (2018, October). A long way to the top: Significance, structure, and stability of internet top lists. In Proceedings of the Internet Measurement Conference 2018 (pp. 478-493).

[7] "YARA Documentation" (Online) Verfügbar unter: https://yara.readthedocs.io/en/v3.5.0/index.html

[8] Naik, N., Jenkins, P., Cooke, R., Gillett, J., & Jin, Y. (2020, December). Evaluating automatically generated YARA rules and enhancing their effectiveness. In 2020 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 1146-1153). IEEE.

[9] "Neo4j Graph Platform – The Leader in Graph Databases" (Online) https://neo4j.com/

[10] Guia, J., Soares, V. G., & Bernardino, J. (2017, January). Graph Databases: Neo4j Analysis. In ICEIS (1) (pp. 351-356).

[11] Barnum, S. (2014). STIX—Structured Threat Information Expression. MITRE

[12] "MISP Objects – VAMOS custom object" (Online) Verfügbar unter: https://github.com/MISP/misp-objects/blob/master/objects/sandbox-report/definition.json

[13] Regulation (EU) 2016/679 (General Data Protection Regulation) OJ L 127, 23.5.2018

[14] Borden, R., Mooney, J., Taylor, M., & Sharkey, M. (2019). Threat information sharing under GDPR. Scitech Lawyer, 15(3), 30-35.

[15] Hellwig, O., Quirchmayr, G., Hötzendorfer, W., Tschohl, C., Huber, E., Vock, F., ... & Langner, G. (2018, August). A GDPR compliance module for supporting the exchange of information between CERTs. In Proceedings of the 13th international conference on availability, reliability and security (pp. 1-7).

[16] Albakri, A., Boiten, E., & De Lemos, R. (2019, June). Sharing cyber threat intelligence under the general data protection regulation. In Annual Privacy Forum (pp. 28-41). Springer, Cham.

This page intentionally left blank

SIEMENS

Siemens Aktiengesellschaft
Werner-von-Siemens-Straße 1
80333 Munich
Germany

siemens.com