Cyber Threat Intelligence mit Open Source Intelligence

Ein Vorteil für das Gemeinwohl

sarah.julia.kriesch@accenture.com

Agenda

Open Source Intelligence und OSINT Daten versus Nicht-OSINT Daten

Cyber Threat Intelligence und Cyber Threat Attribute

Structured Threat Expression (STIX)

Der Threat Score

Beispiel einer Energy Cloud

Edge Intelligence

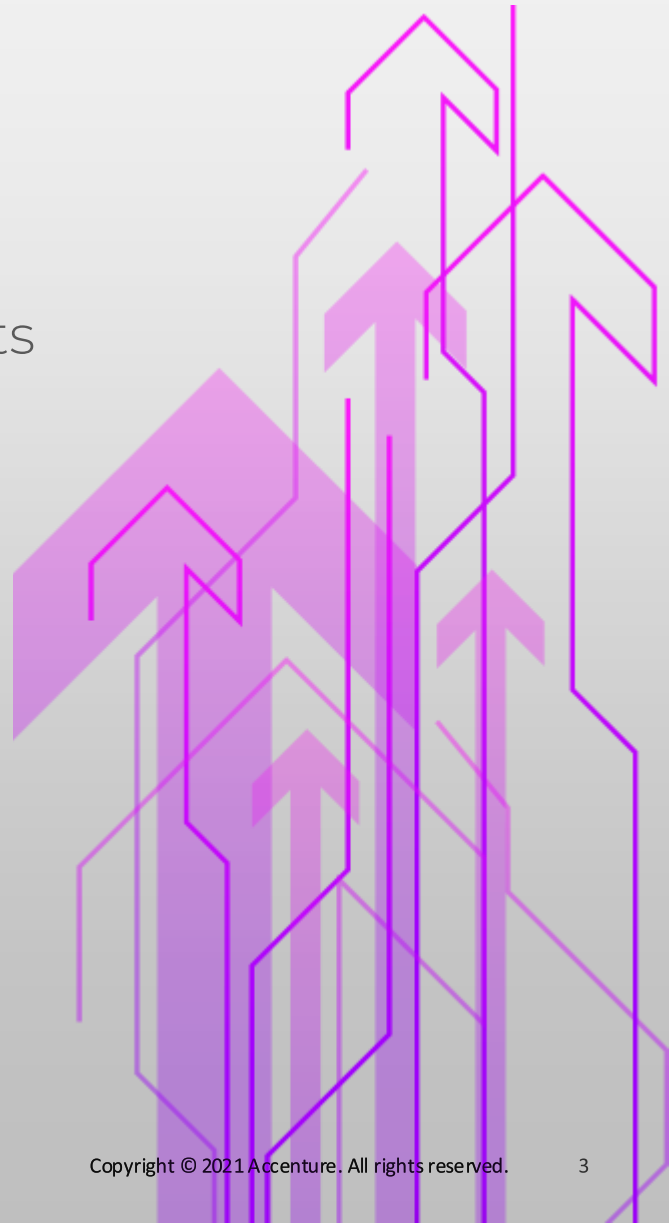
Federated Learning

Homomorphe Verschlüsselung

Zukunft eines zentralen Servers (Quantum Computing)

Open Source Intelligence

- Open Source Intelligence ist die Grundlage anderer Intelligence Bereiche
- Bietet die Möglichkeit mit der Perspektive von außen relevante Threats zu identifizieren
- Bekannt als “Hardware Hacking” (besonders Embedded Systeme)
- Das Sammeln, Analysieren und Verwenden von Open Source Intelligence Data (öffentlich zugänglich)
- Design von OSINT-Tools sind auf Basis des OSINT Cycles



OSINT Cycle

Prozess & Design Best Practices für OSINT Tools

- 4 Phasen:

  - **Data Collection:** Scannen von Quellen und Sammlung von Daten

  - Data Processing: Normalisierung, Vereinigung, Daten Indexe, Zuordnung, Berechtigungen anpassen, Bereitstellung von OSINF

  - **Data Analysis:** Validierung der Quellen und OSINF, Verhinderung von Daten-Manipulation

  - **Intelligence Dissemination:** Daten - mit intelligenter Indexierung, Verteilung, OSINT Speicherung und Feedback - in Produktion bringen

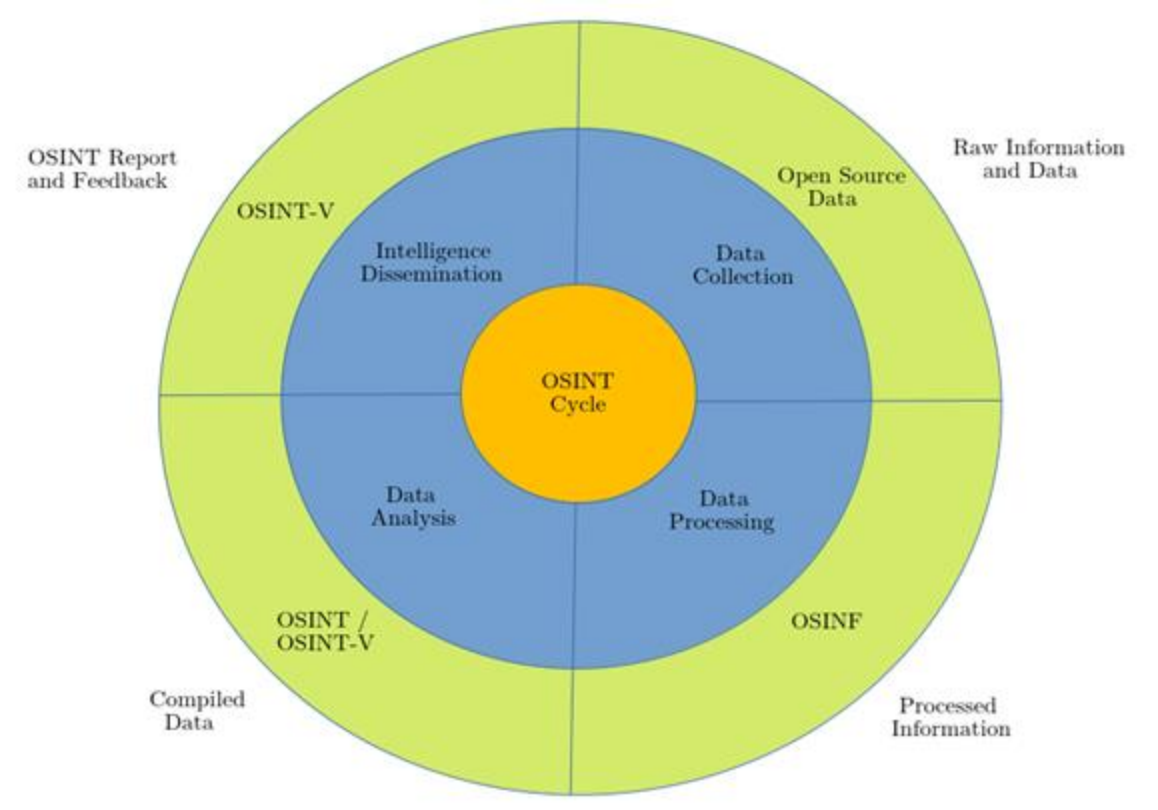


Bild: Der OSINT Cycle (von Kriesch, basierend auf Ungureanu)

OSINT-Daten versus Nicht-OSINT-Daten

OSINT-Daten

Öffentlich zugängliche Daten

Darauf zugreifbar ohne Authorisierung

Strukturierte und unstrukturierte Daten

Unterschiedliche öffentliche Security Datenbanken

Beispiel: CVE-Datenbank

Nicht-OSINT-Daten

Vertrauenswürdige Daten

Nicht öffentlich verfügbar

Meistens gespeichert in Unternehmen/Organisationen

Persönliche Daten (z.B. Mitarbeiter-Nummern)

Cyber Threat Intelligence

- Bietet wertvollen Einblick in potentielle Cyber Threats und Angriffe
- Der Prozess zur Verwendung verfügbarer Informationen um Threats zu verstehen, die auf Organisationen ausgerichtet sind
- Ziel: Frühe Erkennung von (möglichen) Incidents und Bestärkung/Bereitstellung der Umgebung

Definition von Lee:

"The process and product resulting from the interpretation of raw data into information that meets a requirement as it relates to the adversaries that have the intent, opportunity and capability to do harm"

Cyber Threat Attribute

- Information über “wo, warum, und wie” ein Angriff passiert ist

→ Identifikation, was ein Angreifer erreichen wollte (Ausrichtung, Zeitstempel, Ziele)

- Verwendung zur Verhinderung weiterer Angriffe

- Konvertierung unstrukturierter/strukturierter Daten in Intelligenz

- **High-level IoCs:** Tactics, Techniques und Malware

- **Low-level IoCs**: IP-Adressen, URLs, Hash, Domain Names, Timestamp, …

- 4 Attribut-Gruppen: “*network address*”, “*file hash*”, “*file name*” und “*other info*”

Structured Threat Intelligence Expression (STIX)

Etabliertes und standardisiertes Format für CTI-Daten entwickelt vom OASIS Consortium

- Kann gespeichert werden als Maschinen- und Menschen-lesliches Format (JSON)
- Verwendbar zur Automatisierung mit Maschinen (als Python libraries verfügbar)
- Erweiterbar für CTI-Plattformen (Beispiel ETIP)
- Unterstützt 4 Cyber Threat Use-Cases:
  - Analyse von Cyber Threats
  - Spezifikation von Indicator Patterns
  - Verwaltung von Response Aktivitäten
  - Teilen der CTI
- OSINF mit Threat Attributen werden als STIX Domain Objects (SDOs) kategorisiert
- Die Plattform Trusted Automation Exchange of Intelligence Information (TAXII) kann Threat Feeds basierend auf STIX-Informationen extrahieren
- Beispiel-Plattform, die STIX und TAXII verwendet: MISP (Malware Information Sharing Platform)

STIX Domain Objects

- **ThreatActor:** Ein Threat-Actor, der bekannt ist
- **campaign:** Eine Gruppe an Aktivitäten vom Threat-Actor
- **CourseOfAction:** Empfehlungen für nächste Aktivitäten
- **ExploitTarget:** Schwachstellen in Netzwerken, Software, Systeme, und andere Ziele
- **Incident:** Ein STIX-Incident
- **Indicator:** Erkennung Indikator-basierend auf Cyber Attack Patterns
- **Observable:** STIX Cyber-observable Objects (SCOs) um Information zu teilen
- **TTP:** Taktiken, Techniken und Prozeduren
- **Further SDOs:** "attack pattern", "grouping", "identity", "infrastructure", "intrusion set", "location", "malware analysis", "malware", "note", "opinion", "report", "tool", "vulnerability"

Der Threat Score

- Berechnet basierend auf heurischem Analyse-Prozess zur Priorisierung der Incidents
- Hinzugefügt als Attribut im gesendeten IoC innerhalb der CTI
- Prozess:
  - Evaluation der Threat-Daten:
    - Quellen-Identifikation (Log-Dateien, Datenbanken, …)
    - Heuristische Identifikation (IP-Adressen, Ports, …)
    - Threshold Definition (CVE oder nicht)
  - Berechnung des Weighted Mean/TS:
    - X_i: Heuristischer Wert; P_i: Gewicht-Faktor (0-5)
    - C_p: Nicht leere Features/alle Features

Threat Score Berechnung von Gonzalez:

$$TS = C_p \times \left(\sum_{i=1}^{i} X_i \times P_i \right)$$

Low priority (0-1), medium priority (2-3), high priority (3-4), critical priority (5)

CTI für eine Energy Cloud

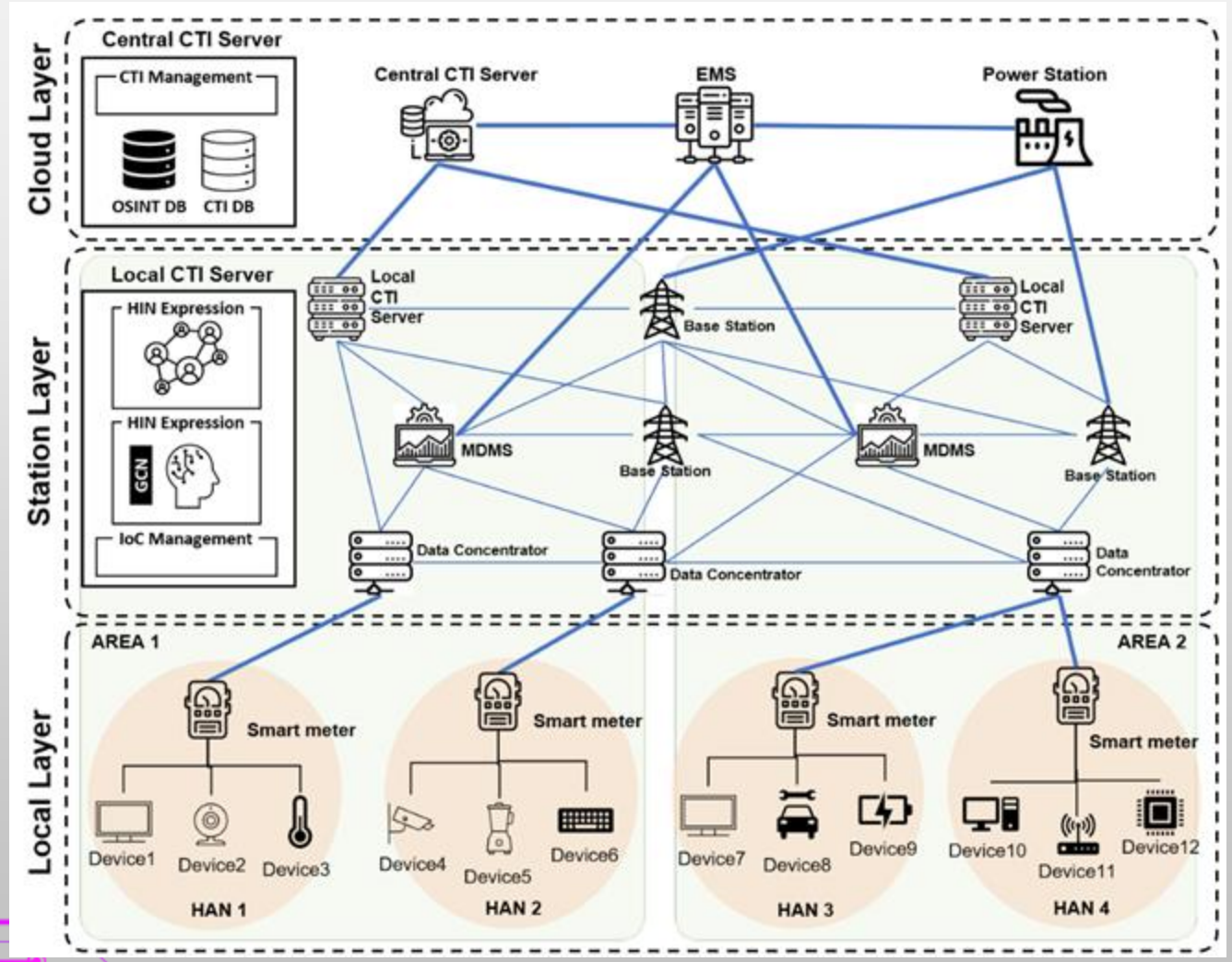


Bild: Energy Cloud Architektur für CTI von Gong

Edge Intelligence

- Verwendung von KI zusammen mit Edge Computing wird Edge Intelligence / Edge AI genannt
- Haupt-Ziel für Cybersecurity ist CIA (Confidentiality, Integrity, Availability)
- Beinhaltet auch Informationssicherheit angewendet auf KI-Algorithmen oder Malfunctions in Software
- Vorteil mit KI zur automatisierte Erkennung von Cyber Risiken
- (Quanten-)Kryptographie kann angewendet werden um Vertrauenswürdigkeit zu garantieren

Federated Learning für CTI

- Federated Learning für Distributed Learning

  (Machine Learning als Teil von KI)

- Online/Incremental Learning um von geänderten und angepassten Daten zu lernen

  → Nützlich für CTI und Edge Intelligence (siehe Beispiel mit Energy Cloud)

- Training ist verteilt mit einem Neuronalen Netzwerk über mehrere Teilnehmer hinweg (CTI Server und Geräte)

- Hao et al. hat das Privacy-Enhanced Federated Learning für Industrial Artificial Intelligence eingeführt

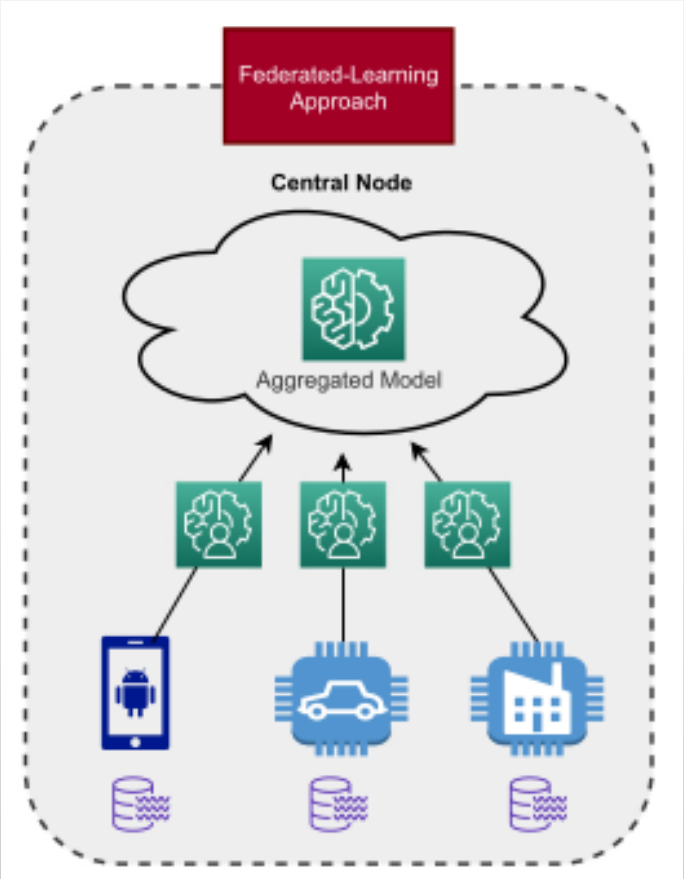
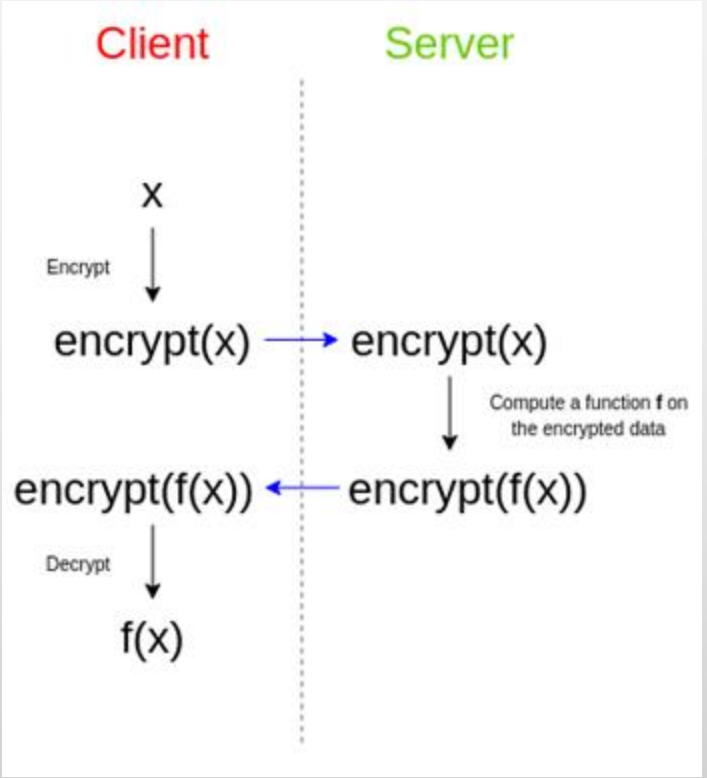


Bild: Federated-Learning Appoach von Campos

Homomorphe Verschlüsselung



- Alle Geräte nutzen den gleichen Secret Key und können verschlüsselte Daten und vorherige Entschlüsselung lesen
- KI benötigt “Full Homomorphic Encryption” (nicht nur adding, sondern auch XOR und Multiplikationen)
- Schutz von sensitiven Informationen
- Erweiterung von Datenschutz
- Sichere Abstimmung von Cyber Security Professionals möglich
- Anwendung von verschlüsselten Daten innerhalb des eigenen Netzwerks

Quelle:
https://github.com/redhat-et/homomorphic-learning/blob/main/docs/homomorphic_learning.md

Idee mit zentralem CTI Server

Quantum Computing für CTI

- Sun et al. (2021) schlug open-source threat intelligence publishing platforms (OSTIPs) vor

→ Neu gefundene Issues können veröffentlicht und gepushed warden (vergleichbar mit CVE-Datenbanken)

- Neue Findings können zu maschinen-lesbaren CTI-Records konvertiert werden

→ Sammlung vieler CTI-Daten

- Dalvi et al. (2023) nannte Quantum Computer als leistungsstarke und sichere Systeme zur Lösung komplexer Probleme für Cyber Threat Intelligence

Benefit: Schnelle Verarbeitung von CTI-Daten (inkl. Strukturierung) und Bereitstellung zum Pulling

Sarah Julia Kriesch
Consultant
Accenture GmbH

sarah.julia.kriesch@accenture.com

Let‘s gain your edge together.