About Me

• Senior Security Analyst at mnemonic
• Project Manager «Semi-Automated Cyber Threat Intelligence (ACT)»
• Project Manager «Threat Ontologies for Cyber Security Analytics (TOCSA)»
• Member of the Europol EC3 Advisory Group on Internet Security
Motivation – mnemonic statistics from 2014

- 150 critical security incidents
- 14300 security incidents
- 1 trillion events

ACT

Today

Threat Intelligence (TI)

Incident Response (IR)

Raw Data
ACT, TOCSA and Oslo Analytics

• Semi-Automated Cyber Threat Intelligence (ACT)
  - Open Source Threat Intelligence Platform

• Threat Ontologies for Cyber Security Analytics (TOCSA)
  - Ontologies
  - PhD Project
  - https://www.mnemonic.no/no/research-and-development/threat-ontologies-for-cybersecurity-analytics/
  - http://www.mn.uio.no/ifi/english/research/projects/tocsa/

• Operable Subjective Logic Analysis Technology for Intelligence in Cybersecurity (Oslo Analytics)
  - Analytics
  - Subjective Logic (quantifying uncertainty)
  - Trust Networks
  - Academic
  - http://www.mn.uio.no/ifi/english/research/projects/oslo-analytics/
Extended abstract presented at the Semantic Technology for Intelligence, Defense, and Security (STIDS) 2016 conference

- [http://stids.c4i.gmu.edu/](http://stids.c4i.gmu.edu/)

Collaborative work:
- Threat Ontologies in Cyber Security Analytics (TOCSA)
- Operable Subjective Logic Analysis Technology for Intelligence in Cybersecurity (Oslo Analytics)
- Semi-Automated Cyber Threat Intelligence (ACT)
THREAT INTELLIGENCE
What is Threat Intelligence?

Threat intelligence is *evidence-based knowledge*, including context, mechanisms, indicators, implications and actionable advice, about an existing or emerging *menace or hazard* to assets that can be used to *inform decisions* regarding the subject's response to that menace or hazard.

- Gartner (2013)
Evidence-Based Knowledge

[Image of a computer screen showing a ransomware message]

**What Happened to My Computer?**
Your important files are encrypted. Many of your documents, photos, videos, and other files are no longer accessible because they have been encrypted. Maybe you are busy looking for a way to recover your files, but do not waste your time. Nobody can recover your files without your decryption service.

**Can I Recover My Files?**
Sure. We guarantee that you can recover all your files safely and easily. But you have not so enough time. You can decrypt some of your files for free. Try now by clicking <Decrypt>.

But if you want to decrypt all your files, you need to pay. You only have 3 days to submit the payment. After that the price will be doubled. Also, if you don’t pay in 7 days, you won’t be able to recover your files forever. We will have free events for users who are so poor that they couldn’t pay in 6 months.

**How Do I Pay?**
Payment is accepted in Bitcoin only. For more information, click <About Bitcoin>. Please check the current price of Bitcoin and buy some bitcoins. For more information, click <How to buy bitcoins>. And send the correct amount to the address specified in this window. After your payment, click <Check Payment>. Best time to check: 9:00 am - 11:00 am.
Evidence-Based Knowledge

www[.]iuqerfsodp9ifjaposdfjhgposurijfaewrwegwea[.]com
Threat Intelligence Categories

- **Tools**
- **Artifacts**
- **Indicators**
- **Campaigns**
- **Tactics**
- **Techniques**
- **Procedures**
- **Attribution**
- **Goals**
- **Strategy**
- **Campaigns**
- **Tools**
- **Artifacts**
- **Indicators**

Strategic

Tactical

Operational

Technical

Long Term

Short Term

Less Detailed

More Detailed
Threat Information vs Threat Intelligence
THREAT INTELLIGENCE PLATFORMS
Evaluation of existing platforms

Threat Intelligence Sharing Platforms: An Exploratory Study of Software Vendors and Research Perspectives

Clemens Sauerwein¹, Christian Sillaber¹, Andrea Mussmann¹, and Ruth Breu¹

¹ University of Innsbruck, Department of Computer Science, Innsbruck, Austria

Abstract. In the last couple of years, organizations have demonstrated an increased willingness to exchange information and knowledge regarding vulnerabilities, threats, incidents and mitigation strategies in order to collectively protect against today’s sophisticated cyberattacks. As a reaction to this trend, software vendors started to create offerings that facilitate this exchange and appear under the umbrella term “Threat Intelligence Sharing Platforms”. To which extent these platforms provide the needed means for exchange and information sharing remains unclear as they lack a common definition, innovation in this area is mostly driven by vendors and empirical research is rare. To close this gap, we examine the state-of-the-art software vendor landscape of these platforms, identify gaps and present arising research perspectives. Therefore, we conducted a systematic study of 22 threat intelligence sharing platforms and compared them. We derived eight key findings and discuss how existing gaps should be addressed by future research.
Key findings

1. There is no common definition of threat intelligence sharing platforms
2. STIX is the de-facto standard for describing threat intelligence
3. Platforms primarily focus on sharing of indicators of compromise
4. The majority of platforms is closed source
5. Most platforms focus on data collection instead of analysis
6. Trust issues between users and platform providers are mostly neglected
7. Academic and commercial interest in threat intelligence sharing increases
8. Many manual tasks make the user the bottleneck
EXAMPLE: APT REPORT
| Report Contents |

<table>
<thead>
<tr>
<th>Sharewarecore</th>
<th>Sharewarecore header DLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name(s)</td>
<td>elogger-dll</td>
</tr>
<tr>
<td>Related campaign</td>
<td>&quot;Myanmar&quot;, see section 5</td>
</tr>
<tr>
<td>C2 URL</td>
<td><a href="http://unregister.rmlearn.com/rogue.php">http://unregister.rmlearn.com/rogue.php</a></td>
</tr>
<tr>
<td>MD5</td>
<td>3e1b3e4fd6f318c5b21b0f874cdd75bd67a9aefa8</td>
</tr>
<tr>
<td>SHA1</td>
<td>ab0f8baf312c13558640e8d7c708a8e6a7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sharewarecore</th>
<th>Sharewarecore header DLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name(s)</td>
<td>vmware-vmware.exe</td>
</tr>
<tr>
<td>Related campaign</td>
<td>&quot;Global campaign&quot;, see section 6.4</td>
</tr>
<tr>
<td>MD5</td>
<td>7f09a30b8a424606967819d9a7d89746345271be6</td>
</tr>
<tr>
<td>SHA1</td>
<td>8e524a7f3a0a0a9819b9fe3a7544a637</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sharewarecore</th>
<th>Sharewarecore header DLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name(s)</td>
<td>-</td>
</tr>
<tr>
<td>Related campaign</td>
<td>-</td>
</tr>
<tr>
<td>C2 URL</td>
<td>-</td>
</tr>
<tr>
<td>MD5</td>
<td>-</td>
</tr>
<tr>
<td>SHA1</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sharewarecore</th>
<th>Sharewarecore header DLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name(s)</td>
<td><a href="http://www.faw.com/proxysequence/wp/plugins/sample5.php">http://www.faw.com/proxysequence/wp/plugins/sample5.php</a></td>
</tr>
<tr>
<td>Related campaign</td>
<td>-</td>
</tr>
<tr>
<td>C2 URL</td>
<td>-</td>
</tr>
<tr>
<td>MD5</td>
<td>59be52cfc5f2b5a5d5038130d31359c6d5</td>
</tr>
<tr>
<td>SHA1</td>
<td>a65c4a8baf0a6510c37d9ef</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sharewarecore</th>
<th>Sharewarecore header DLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name(s)</td>
<td>-</td>
</tr>
<tr>
<td>Related campaign</td>
<td>-</td>
</tr>
<tr>
<td>C2 URL</td>
<td>-</td>
</tr>
<tr>
<td>MD5</td>
<td>-</td>
</tr>
<tr>
<td>SHA1</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sharewarecore</th>
<th>Sharewarecore header DLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name(s)</td>
<td>-</td>
</tr>
<tr>
<td>Related campaign</td>
<td>-</td>
</tr>
<tr>
<td>C2 URL</td>
<td>-</td>
</tr>
<tr>
<td>MD5</td>
<td>-</td>
</tr>
<tr>
<td>SHA1</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sharewarecore</th>
<th>Sharewarecore header DLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name(s)</td>
<td>-</td>
</tr>
<tr>
<td>Related campaign</td>
<td>-</td>
</tr>
<tr>
<td>C2 URL</td>
<td>-</td>
</tr>
<tr>
<td>MD5</td>
<td>-</td>
</tr>
<tr>
<td>SHA1</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sharewarecore</th>
<th>Sharewarecore header DLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name(s)</td>
<td>-</td>
</tr>
<tr>
<td>Related campaign</td>
<td>-</td>
</tr>
<tr>
<td>C2 URL</td>
<td>-</td>
</tr>
<tr>
<td>MD5</td>
<td>-</td>
</tr>
<tr>
<td>SHA1</td>
<td>-</td>
</tr>
</tbody>
</table>
Approach

- Manually create csv files
- Design simple graph structure
- Transform csv files to graph DB using Python
Graph Structure
Knowledge Graph
MATCH p=(n:Campaign)<-(:Sample)-->(o)<-(:Sample)-->(m:Campaign)
WHERE NOT o:Malware AND m <> n
RETURN p
IP addresses with multiple domains

MATCH p=(n:Domain)-[r]->(o:IP)<-[s]-(m:Domain) RETURN p
MATCH p=(m)<--[]-*1..3>-(n:IP {name: "178.209.52.72"}) WHERE NOT m:Malware AND NOT m:Filename AND NOT m:Path AND NOT m:IP RETURN p
Knowledge Graph from STIX
SEMI-AUTOMATED CYBER THREAT INTELLIGENCE (ACT)
Semi-Automated Cyber Threat Intelligence (ACT)

The main objective of the research project is to develop a platform for cyber threat intelligence to uncover cyberattacks, cyber espionage and sabotage.

The project will result in new methods for data enrichment and data analysis to enable identification of threat agents, their motives, resources and attack methodologies.

In addition, the project will develop new methods, work processes and mechanisms for the generation and distribution of threat intelligence and countermeasures, to stop ongoing and prevent future attacks.
ACT Project Goals

• Holistic workspace for analysts
• Automation
  - Repetitive tasks
  - Processing of unstructured data
  - Sharing
    • Threat information
    • Countermeasures
• Advanced automated analysis
• Advanced enrichment
• Manual analysis
  - Efficiency
  - Accuracy
• Improve our knowledge of threat agents
Data Model

• Objects
  - Global
  - Example: IP address

• Facts
  - Connected to a single object or multiple objects (relation)
  - Immutable
  - Timestamped
  - Owner
  - Role-based and explicit access control
  - Backed by evidence and comments
The Detection Maturity Level (DML) Model

- **Attacker identity**
  - DML-9: Identity

- **Attacker goals and strategy**
  - DML-8: Goals
  - DML-7: Strategy
  - DML-6: Tactics
  - DML-5: Techniques
  - DML-4: Procedures
  - DML-3: Tools

- **Attack execution plan and methods**
  - DML-2: Host & Network Artifacts
  - DML-1: Atomic Indicators
  - DML-0: None or Unknown

- **Traces of attack execution**

---

Causality and Classifiers for the DML Model

- External intelligence
- Classifiers
- Attacker goals, strategy and identity
- Attack execution plan and methods
- Traces of attack execution

Causality

Classifiers
Semantic Feature Extraction

• Formal definitions of
  - Goals
  - Strategy
  - Tactics
  - Techniques
  - Procedures

• Relevant initiatives
  - MITRE CAPEC
    • https://capec.mitre.org
  - MITRE ATT&CK
    • https://attack.mitre.org
  - MITRE CAR
    • https://car.mitre.org
Network hopping and exfiltration

Once APT10 have a foothold in victim networks, using either legitimate MSP or local domain credentials, or their sustained malware such as PlugX, RedLeaves or Quasar RAT, they will begin to identify systems of interest.

The operator will either access these systems over RDP, or browse folders using Remote Access Trojan (RAT) functionality, to identify data of interest. This data is then staged for exfiltration in multi-part archives, often placed in the Recycle Bin, using either RAR or TAR. The compression tools are often launched via a remote command execution script which is regularly named ‘t.vbs’ and is a customised version of an open source WMI command executor which pipes the command output back to the operator.
Example Procedure: Authentication with stolen credentials

Environment: Windows cmd.exe command line

1. ping -n 1 HOSTNAME
2. net use \\HOSTNAME\ipc$ "PASSWORD" /user:"DOMAIN\USERNAME"
Example Procedure Detection

Prerequisite: logging of cmd.exe command line (e.g. Sysmon)

For each COMMANDLINE in cmd.exe process:
  if COMMANDLINE matches ‘ping -n 1 HOSTNAME’:
    if next COMMANDLINE starts with ‘net use \HOSTNAME\ipc$’:
      Trigger alarm
Traces Knowledge Graph
Unstructured Data – Natural Language Processing

- No corpus for the cyber security domain
- *Snowball: Extracting Relations from Large Plain-Text Collections* ¹
- Test case: APTNotes ([https://github.com/aptnotes/data](https://github.com/aptnotes/data))

APTNotes NLP processing
Machine Learning

This is your machine learning system?

Yup! You pour the data into this big pile of linear algebra, then collect the answers on the other side.

What if the answers are wrong?

Just stir the pile until they start looking right.
Triplets and semantic reasoning

Subject

Predicate

Object

«Things»

«Relationship»

Triplet
Triplets and semantic reasoning

- Martin marriedTo Jing

Triplet
Triplets and semantic reasoning

- Martin marriedTo Jing
- Dag brotherOf Martin
- Dag brotherOf Richard
- Jing marriedTo Richard
- Richard sonOf Martin
- Martin brotherOf Dag
- Dag nephewOf Richard
- Richard nephewOf Martin
- Martin motherOf Dag
- Dag uncleOf Richard
- Richard uncleOf Martin
Triplets and semantic reasoning

Sad Panda operatorOf Operation Bulldozer

USA locatedIn Sunny Hospital

Sunny Hospital sectorMember Medical

Operation Bulldozer targetArea North America
Threat Intelligence Platform

- Data model and architecture done
  - Objects and immutable facts (relations/predicates)
  - ACL on facts
  - Queues and workers
- Platform core, API and GUI under development and testing
- Github project
  - https://github.com/mnemonic-no
- Ongoing research:
  - Threat ontologies
  - Analysis techniques
  - Enrichment techniques
  - Sharing and Countermeasures
  - Workflow orchestration
Feedback and ideas

- Useful, formal definitions of TTPs
- Examples of predicates («marriedTo») for Threat Intelligence
- Experiences, use cases
- Any other clever ideas