Cyber Early Warning & the Commonality of Cyber Warfare and Electronic Warfare

SESSION ID: TRM-T08

Dr. Nitzan Barkay  & Mariana Gafni

Deputy Director, Research & Technology
IAI – Israel Aerospace Industries
## Cyber Early Warning Key Difficulty

<table>
<thead>
<tr>
<th>The Goal:</th>
<th>Early warning of cyber attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currently:</td>
<td>Many tools &amp; techniques to detect “non-legitimate” activity or “abnormal” behavior</td>
</tr>
<tr>
<td>Suggested:</td>
<td>A new layer to handle complex &amp; sophisticated attacks</td>
</tr>
<tr>
<td>The Challenge:</td>
<td>Too many False Alarms that cannot be handled; OR: Reduce false alerts by stricter criteria, while unfortunately masking out subtle events, typical to APT attacks</td>
</tr>
</tbody>
</table>

![Graph showing the relationship between probability of detection and false alarm rate.]
Towards Cyber Warfare

Electronic Warfare (EW)
- Integrated SIGINT (ELINT&COMINT) Solutions
- Electronic Protection (EP) & Electronic Attack (EA)

Cyber Warfare
- Intelligence & Situation Awareness

Communication & C4I
#Cyber Warfare vs. Electronic Warfare

<table>
<thead>
<tr>
<th></th>
<th><strong>Electronic Warfare</strong></th>
<th><strong>Cyber Warfare</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mission</strong></td>
<td>- Air-situation picture (surveillance)</td>
<td>- IT</td>
</tr>
<tr>
<td></td>
<td>- Guiding missiles</td>
<td>- SCADA</td>
</tr>
<tr>
<td></td>
<td>- Navigation</td>
<td>- Business</td>
</tr>
<tr>
<td></td>
<td>- C&amp;C/data networks</td>
<td>- Government services</td>
</tr>
<tr>
<td><strong>Intelligence</strong></td>
<td>- SIGINT (ELINT, COMINT)</td>
<td>- Hacking</td>
</tr>
<tr>
<td></td>
<td>- IMINT (Opt., Radar)</td>
<td>- Accessibility tools</td>
</tr>
<tr>
<td><strong>Attack</strong></td>
<td>- Electronic Attack (EA)</td>
<td>- Cyber attacks</td>
</tr>
<tr>
<td></td>
<td>- ECM (Victim: radars)</td>
<td>- (Victim: network services &amp; resources)</td>
</tr>
<tr>
<td></td>
<td>- ComJam (Victim: comm. links)</td>
<td></td>
</tr>
<tr>
<td><strong>Attack type</strong></td>
<td>- Jamming</td>
<td>- Jamming</td>
</tr>
<tr>
<td></td>
<td>- Spoofing, noise</td>
<td>- DoS, DDoS</td>
</tr>
<tr>
<td></td>
<td>- Deception</td>
<td>- Deception</td>
</tr>
<tr>
<td></td>
<td>- False targets, missile stealing</td>
<td>- Identity theft, MITM, phishing, Trojan horses</td>
</tr>
<tr>
<td><strong>Counter-measures</strong></td>
<td>- ECCM:</td>
<td>- Counter-measures</td>
</tr>
<tr>
<td></td>
<td>- Filters, guards, SLB&amp;SLC,…</td>
<td>- FW, IPS,…</td>
</tr>
<tr>
<td></td>
<td>- Decoys,…</td>
<td>- Honeypots,…</td>
</tr>
<tr>
<td></td>
<td>- Immunity</td>
<td>- Immunity</td>
</tr>
<tr>
<td></td>
<td>- LPI: waveform, agility,…</td>
<td>- Encryption, virtualization</td>
</tr>
</tbody>
</table>
Cyber & EW Integration in Battlefield

- An example:
  **The US Army has published(*) the ICE (Integrated Cyber & Electronic Warfare) program**
  - Define common data contexts & mechanisms to allow Cyber & EW frameworks to communicate and combat the threats in an integrated fashion

(*) http://www.army.mil/article/113678
Multi-Entity
Multi-Sensor
Scenario
&
Multi-Hypothesis
Tracking
Cyber Early Warning Challenges

- **Huge amount of activity**
  - Data availability, especially in real time
    - Technical & regulatory difficulty to maintain effective coverage of everything
  - Derive insight from the mass of data
  - Data diversity
  - Data dynamics

- **Attacker/defender asymmetry**
  - Proliferation of attack types
  - Difficulty of "attribution" to actual actors

- **Attacks that involve subtle activities**
  - Eliminating false alarms: Discrimination between legitimate activity and cyber incidents

- **Attacks that involve multiple assets**
  - Identification based on the aggregated picture
Persistent Surveillance Challenges

- A multitude of entities, of various types
- Dynamic scenario
- Integration of different sensors
  - Each interprets the situation picture in its manner
  - Some get only a partial situation picture; Some overlap
- Discrimination between “innocent” entities (false) and “malicious” targets (real threats)
  - Threats attempt to avoid interception by hiding or behaving like legitimate entities

Challenges are similar to Cyber situation awareness; Solutions can be similar, too...
Multi-hypothesis Tracking for Cyber Situation Awareness

- **Multi-Hypothesis Tracking (MHT)** is a powerful means towards achieving **Cyber Situation Awareness**
- **Situation Awareness** is a broader & better concept than Alert
  - more information & comprehension
  - more threat assessment
  - more reliable & informative alerts
Cyber Multi-Hypothesis

- **Multi-Hypothesis Analysis** is a method to **handle the uncertainty**

- An algorithmic methodology to handle complex & dynamic data
  - Collected with various sources/sensors,
  - Involving many entities,
  - Information is partial and/or ambiguous,
  - Information is streaming & dynamically changing

- **For example:**
  - Physical situation awareness (e.g., air situation picture)
  - SIGINT-based order of battle (EOB)

- **Applicable to Cyber Situation Awareness**
  - Integrating the various security tools & techniques
  - Handling the uncertainty and supporting decision making
Multi-Sensor Multi-Entity Tracking

**Tracking is** the logical process of **associating data of activity** (including past data) of **various entities** into disjoint sets - **tracks**

- **Examples:**
  - Geographical data of **platform entities** into physical movement tracks
  - EW & SIGINT data of **electromagnetic entities** into threat interception tracks

- **Logical tracks of data enable**
  - Verification of data consistency
  - Identifying the past origin of the track
  - Predicting the future evolution of the track
Cyber Tracking

Tracking is the logical process of associating data of activity (including past data) of various entities into disjoint sets - tracks

Examples:

- Geographical data of platform entities into physical movement tracks
- EW & SIGINT data of electromagnetic entities into threat interception tracks

Logical tracks of data enable

- Verification of data consistency
- Identifying the past origin of the track
- Predicting the future evolution of the track

Cyber events data

Cyber entities

Cyber Incident tracks!

Eliminate False!

Attribution!

Threat Alert!
Cyber Multi-Hypothesis Tracking (MHT)

- MHT associates distinct cyber events to a single cyber incident
  - When a new message from any sensor or information source is received, to which incident track does that message correspond?
- Events may initially be distinct
  - by “time”: evolution in time
  - by “location”: events detected at different items/hosts/etc.
  - by “sensor”: events detected by various sensor & security tools
  - by “type”: events of different type (a malicious file, illegitimate login, etc.)
Highlights of the MHT Algorithm
MHT Algorithm Main Modules

- **Hypothesis Management engine**
  - A generic module
    - Applicable to physical entities, electronic warfare signals or cyber events
    - Maintaining ambiguities, tracks, pictures, and history

- **Correlation & Scoring**
  - Specific modules
    - Depend on the application sensor characteristics
    - Correlating observations to system states and to previous data based on specific models
Tracks & Pictures

- A **Track** consists of set of data that may be associated with a single platform/system/incident
  - There can be alternative tracks to the same data
- A **Picture** includes a set of alternative tracks that are consistent with each other
  - There can be alternative pictures to the same data

- The "best" picture in any moment is the one which is selected for report, but many are maintained
Tracks & Pictures Schematic Example

1

$\text{picture}$

2

3

track
Observations & Hidden States

- **States** describe the status, behavior or properties of a system
  - There are transition probabilities between system states

- **Observations** include the data streaming from the sensors

- Observations relate to the system states, but the relationship may be ambiguous
  - There may be several possible different observations (with some probabilities) to each state; an observation may point to several possible alternative hidden states

- **Example in Electronic Warfare**
  - Looking at the electronic order of battle (EOB) picture: The hidden states are the emitter/system type and the platform carrying it, while the observations are the intercepted electronic parameters
  - Looking at the geographic situation picture: The hidden states are the position & velocity of the platforms that should be estimated from observed bearings
Correlation & Tracking Models

- **Models** are used to correlate between the observations and the hidden-states of the system, based on the knowledge of expected processes & behavior

- **Model types**
  - **Kinematic** for continuous dynamics
    - e.g., platform trajectory based on direction observations
  - **Rule-based** for simple logic correlation
    - e.g., emitter type based on electronic parameters
  - **Discrete Markov chain** for discrete states
  - **Hidden Markov model (HMM)** when the states are not directly observable
  - **Ontology-based** analytics of related entities using patterns

- **In Cyber**
  - The **hidden states** can be individual host states (trusted, compromised, etc.)
  - The **observations** derive from firewalls, IDS sensors, server/network logs, etc. as well as context & intelligence
  - Relevant **models** are HMM and ontology-based with adaptation to attack types (worm, virus, DDOS, etc., and combinations)
More Aspects of MHT Algorithms

◆ **Observability**
  - A state cannot always be estimated from a sequence of observations; necessary and sufficient conditions for observability should be evaluated.
  - In Cyber: The state of a host or network may not be identifiable from the reported events; the conditions can be estimated using attack models.

◆ **Hypotheses management**
  - The number of hypotheses may increase exponentially as observations arrive; consequently, the computational complexity of maintaining the hypotheses and finding the optimal solution may grow too much.
  - Algorithms of clustering & pruning are employed to overcome the complexity of growing number of hypotheses.
    - Deleting tracks, which have not been updated during a "purge time", which depends on estimated progress rate.
    - Pruning the unlikely (low-score) hypotheses, with the risk of eliminating the future optimal hypothesis.
  - Clustering the tracks into independent sets and using scalability in the algorithms enable distributing the computational load.
MHT & Cyber Early Warning
Cyber MHT Process (1)

- The **hidden states** can be individual states (trusted, compromised, etc.) of a host, a network, or a service.
- The **observations** derive from firewalls, IDS sensors, server/network logs, etc., as well as context & intelligence.
- Relevant **models** are HMM and ontology-based; Models are adapted to attack types (worm, virus, DDOS, etc., and their combinations).
- Each **picture** hypothesis represents a set of events associated to an independent cyber incident.
Context & Intelligence

- **Context** refers to internal (organization) information
  - Structure, procedures, etc.
- **Intelligence** refers to relevant external data collected using WEBInt & accessibility tools
  - Hints to expected attackers, their behavior and their targets

**Context & Intelligence are key factors for decision making**

They add an important dimension to MHT scoring by allowing to judge events using adapted criteria.
Cyber MHT Process (2)

- **Hypothesis score** (track & picture) depends on
  - Information **quality** & consistency
  - Likelihood as estimated by the tracking model
  - Intelligence & relevant context
  - Impact assessment

- The hypothesis with the highest score is reported
- Many of the other hypotheses & tracks are maintained

- Each new event is checked against many hypotheses (not just the previously best)
  - An updated set of hypotheses is formed with updated scoring

- **MHT keeps some history, in a special way of tracks & pictures, which is more efficient to utilize, when data is streaming and early response is required**
  - (All data is logged for later forensic research)
Multi-Hypothesis Tracking on Time

- MHT associates **events on a time scale** to follow cyber incident evolution backward & forward.

**Backward**
- Confirm information consistency
- Investigate overlooked or dropped events
- Look for actor trend & behavior

**Forward**
- Predict “expected” information
- Guide efforts of monitoring events to suspicious paths
- Project situation evolution to assess impact

---

**ALERT**
Track Triggering

- **In Electronic Warfare MHT**
  - Usually, any new intercept data is a possible trigger for tracking
  - For example: each plot of Radar, every signal in EW
  - All entities (including legitimate) are tracked, to distinguish the hostile ones

- **At a Cyber Warfare scenario**
  - The amount of data is enormous
  - Tracking all legitimate activity is impossible
  - However, analysis of known APT attacks, demonstrate that (eventually) suspect activity has been overlooked, resulting in miss of detection
Track before Detect

- The approach is to track the events even before making the decision
  
  **IF** an hypothesis track is consistent and has a high scoring, if its backward evolution suggests a threat actor and/or its forward projection indicates possible impact

  **THEN** Report “Detection”

- And immediately get Incident track details
  - All associated events
  - Possible actor attribution
  - Estimated future impact
Research Challenges

- Ongoing research to improve performance
  - Flexible data modeling to handle all types of information, structured & unformatted, activity & intelligence
  - Best tracking models of events and attacks
  - Analytic engines & optimal hypothesis scoring
  - Efficient pruning & clustering
  - etc.
New Generation Cyber Situation Awareness

Standard Operational Procedures
Incident Response Workflow

Cyber Situation Awareness

MHT: Multi-Hypothesis Tracking
Decision making & Alert

Analytics
Various Analysis & Processing Engines

Ontology
based on Cyber Defense Models
Common language to all Data, revealing relationship & tracks

Heterogeneous Information & Data Sources

Activity
Context
Intelligence
Customer
New Generation Cyber Situation Awareness

Cyber Situation Awareness

**MHT: Multi-Hypothesis Tracking**
Decision making & Alert

Various Analysis & Processing Engines

**Ontology**
Based on Cyber Defense Models
Common language to all Data, revealing relationship & tracks

Heterogeneous Information & Data Sources

- Activity
- Context
- Intelligence
- Customer

Standard Operational Procedures
Incident Response Workflow

Execution
Actionable Intelligence

Predictive Analysis
Contextual Information

Data Normalization
Data, Structured & unstructured

Data - Information - Analytics - Intelligence

Evolution

IAI
THANK YOU