جامعة خليفة
Khalifa University

C2PS
CENTER FOR CYBER-PHYSICAL SYSTEMS

AI and Cyber Warfare

Ernesto Damiani

12-09-2019

KU Cyber-Physical Systems Center (C2PS)

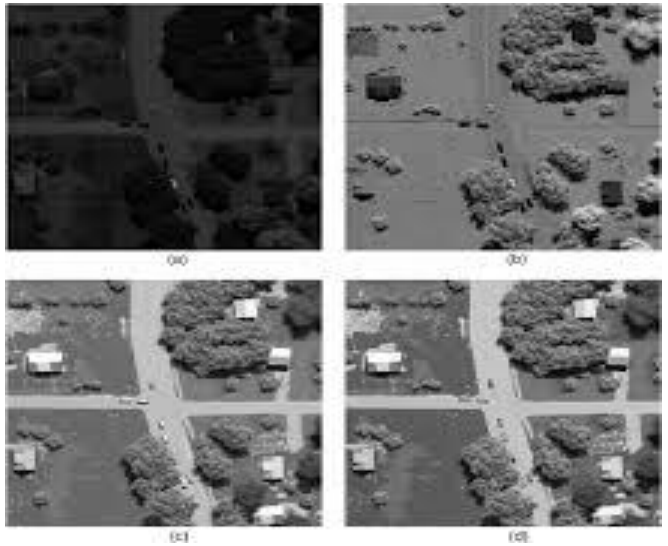
In Partnership with KHALIFA UNIVERSITY

- **SECURITY OF THE GLOBAL ICT INFRASTRUCTURE**
  - Network and Communications Security
  - Business Process Security and Privacy
  - Security and Privacy of Big Data Platforms
- **SECURITY ASSURANCE**
  - Security Risk Assessment and Metrics
  - Continuous Security Monitoring and Testing
- **DATA PROTECTION AND ENCRYPTION**
  - High Performance Homomorphic Encryption
  - Lightweight Cryptography and Mutual Authentication

CONCORDIA
Cyber security cOmpeteNCe fOr Research anD InnovAtion

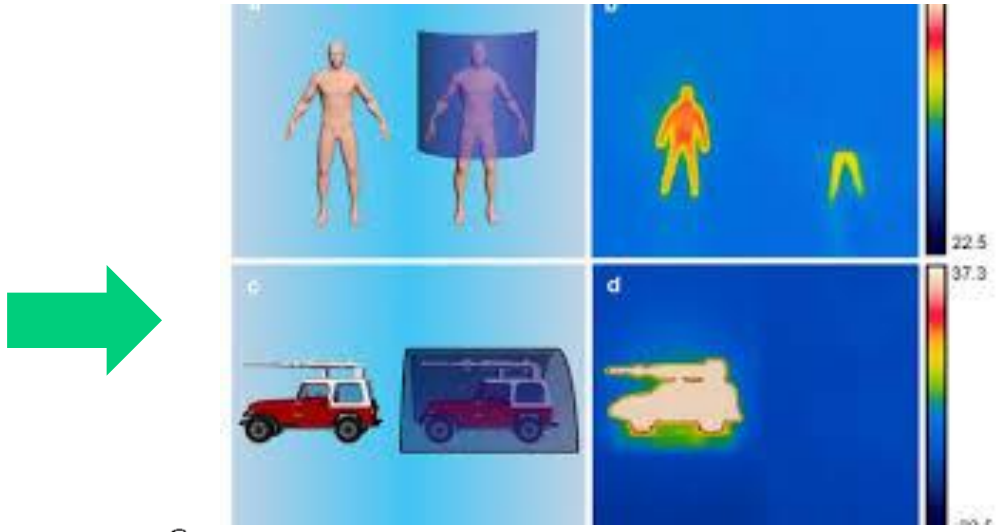
AI in cyberwarfare: the first generation

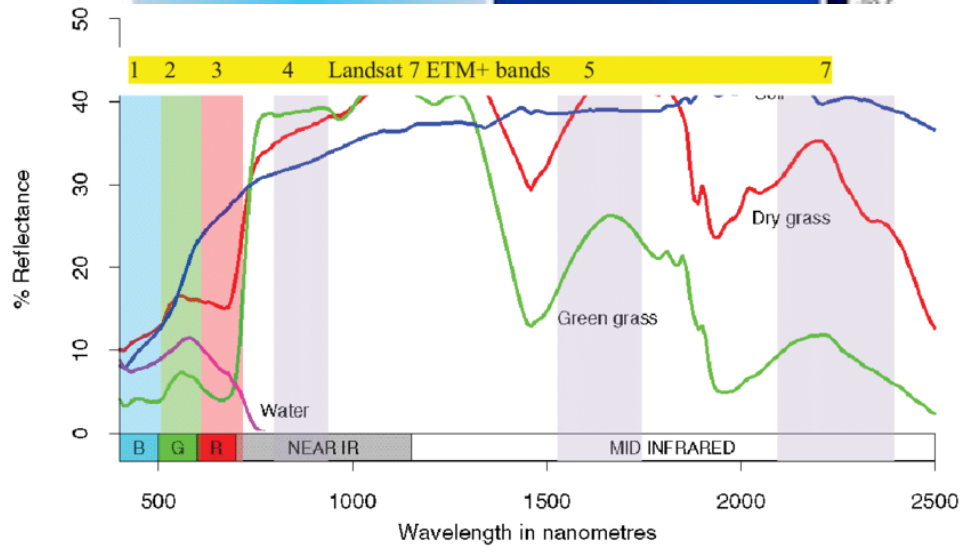
- The first generation of IA systems, coupled with encrypted videoconferencing systems for consultations between humans, has already demonstrated its potential in various field operations since the Second Gulf War
- *Multi-spectral computer-vision AI* vertically integrates support for local tactical decisions (for example, **the choice of which compound of a compound to inspect / occupy first**) with those of sector (for example, to which inspection allocate the support of drones or helicopters).
  - Extended to voice and text processing
- Military personnel have become mobile sources of information (landmarks) as well as users



Classic sensitive data identification

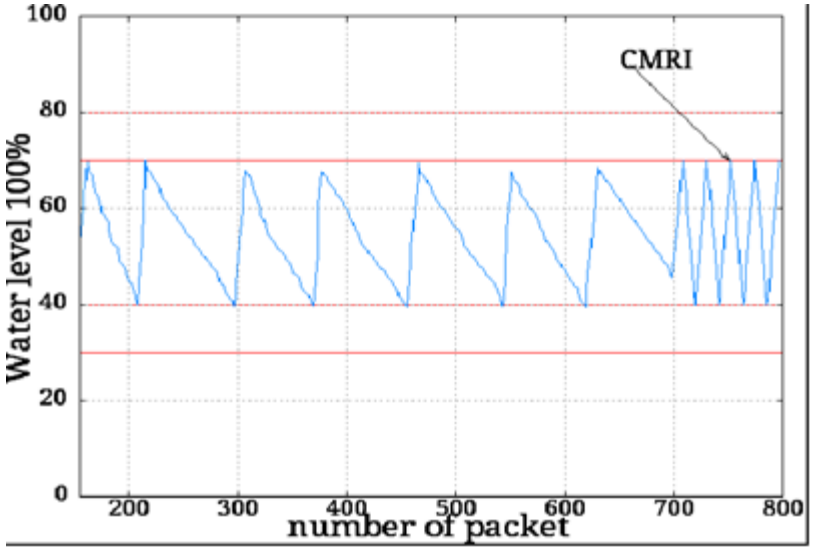
- **Spectral Fingerprints**
  - The light interacts with the bonds in the molecules, which resonate at frequencies, giving each molecule a “spectral fingerprint.” Many molecules and materials more strongly resonate in the IR end of the spectrum, which has very long wavelengths of light – often larger than the molecules themselves.
- **Regular expressions**
  - social security numbers, telephone numbers, addresses, and other data that has a significant amount of structure.
- **Keywords**
  - small number of words that can identify private data, e.g., medical or financial records





Signatures

- **Atomic signatures**
  - A single element, activity, or event is examined to determine if the signature should trigger a signature action.
  - The entire inspection can be accomplished in an atomic operation that does not require any knowledge of other activities.
- **Stateful signatures**
  - Stateful signatures trigger on a sequence of events
  - Require the analytics device to maintain state for a duration known as the *event horizon*.
  - Configuring the length of the event horizon is a tradeoff between consuming system resources and being able to detect an attack that occurs over a long period of time.
- ***Slow attacks* exploit the fact that an IPS cannot maintain state information indefinitely without eventually running out of resources.**

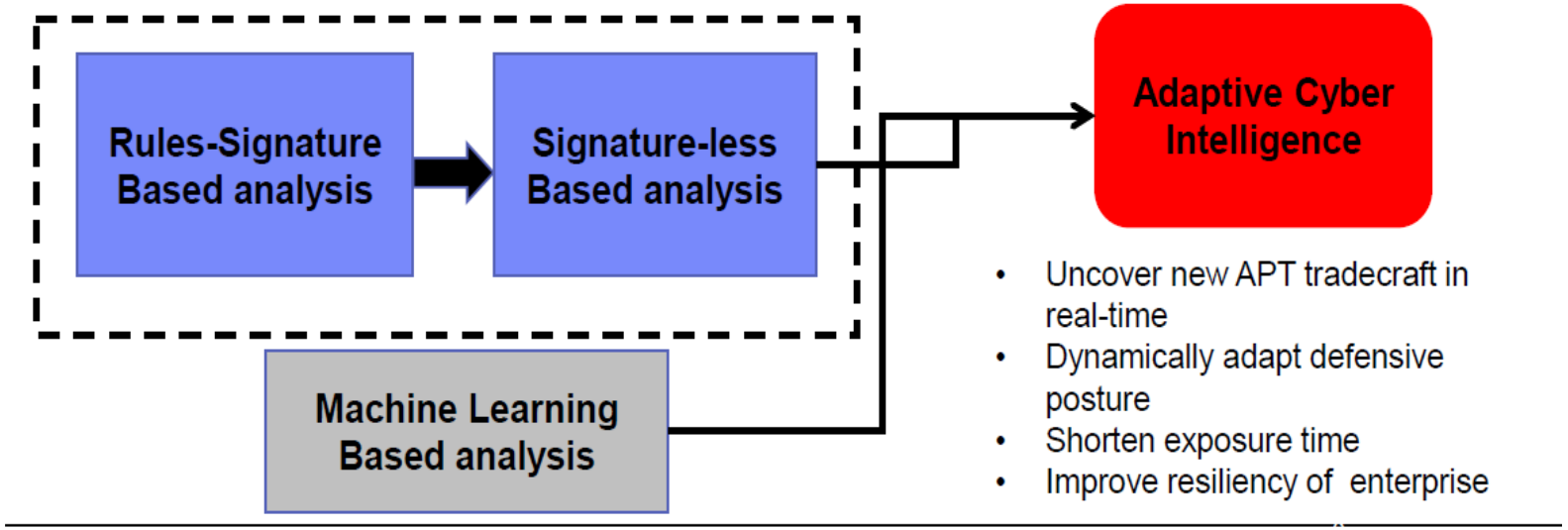


From: Wei Gao, Thomas H. Morri
ON CYBER ATTACKS AND SIGNATURE BASED
INTRUSION DETECTION FOR MODBUS
BASED INDUSTRIAL CONTROL SYSTEMS

- ***Signature-based analytics* can only detect attacks for which a signature has previously been created**
- **Machine Learning Techniques use *patterns* to detect behavior that falls outside of normal system operation**

Complementary, not alternative

- **ML improves signature-based responsiveness and increases precision**



What AI and Big Data Analytics can deliver today

Predict Discrete Attributes

Algorithms	Missions
▪ Collaborative Filtering ▪ K-Means ▪ Principal Component Analysis ▪ Belief Propagation	▪ Determine which entry ports are of most interest to a given threat ▪ Determine type of threat based on specific activity ▪ Infer an individual's tendencies based on those of his friends and family

Predict Continuous Attributes

Algorithms	Missions
▪ Collaborative Filtering	▪ Predict site visitors given historical trends ▪ Predict how an insider threat might value certain risk factors ▪ Predict likelihood that a packet might contain malware items

Determine Groups

Algorithms	Missions
▪ Community Detection ▪ K-Means ▪ Belief Propagation	▪ Analyze individuals by patterns ▪ Identify servers with similar usage characteristics ▪ Determine groups persuaded by similar interests

Predicting Influencers

Algorithms	Missions
▪ Page Rank ▪ Community Detection	▪ Determine group dynamics based on link analysis ▪ Determine the most efficient message dissemination

Limits of First Generation AI

- However high their impact may be, first-generation systems cannot be considered revolutionary
- They have improved the *speed* with which operational decisions are taken and the *quality* of their results, but not their *nature*.
- The military decision-making process remains human-centered and *climbs a chain of command* to make decisions that relate to a certain area based on information acquired in another.
  - Simply put, a company commander will make decisions about the deployment of a platoon based on video streams from another platoon in the field.

Second generation: weaponized AI

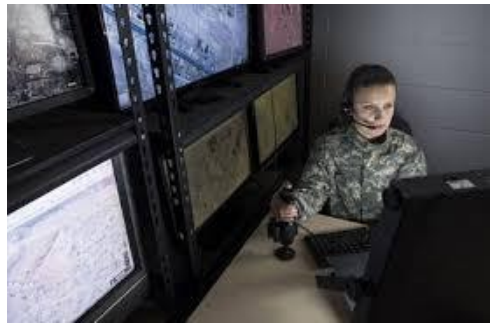
- Artificial intelligence makes it possible to conceive a "generalized battlefield" composed of three areas: *geospace* (the Earth), *space* (*satellite and airborne detectors*) and *cyberspace* where
  - Humans may **not** be involved in tactical decisions
    - Information acquired in an area is used to make automatic decisions (i.e., without going back up a chain of command) in any other area.
  - AI weaponized systems use the information flows made available by first-generation tools and its own integrated sensors to feed the inference of a Machine Learning model that can select and engage human and non-human goals without further intervention by a human operator

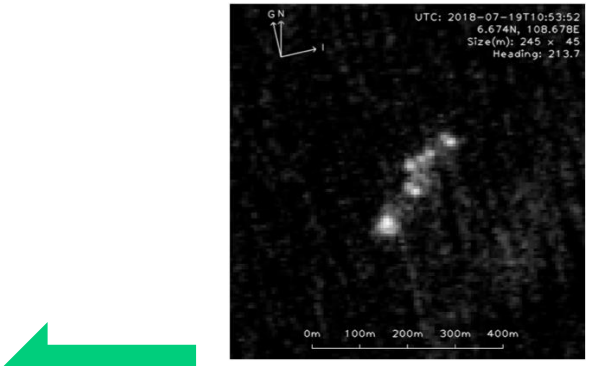
Second generation: weaponized AI

- We consider an unmanned patrol vessel which, upon encountering a cargo ship flying the flag of a country under embargo, receives from a first-generation AI system the warming of a trans-shipment in violation of the embargo.

More in detail:

- the Machine Learning model of the first-generation system is a Deep Learner implemented as a software in the cloud
- It examines medium-resolution multispectral satellite images that reveal the tonnage of the cargo ship
- Classifies as "highly probable" that the current cargo of the ship comes from a cargo ship of a third country, whose estimated route is compatible with the transshipment.



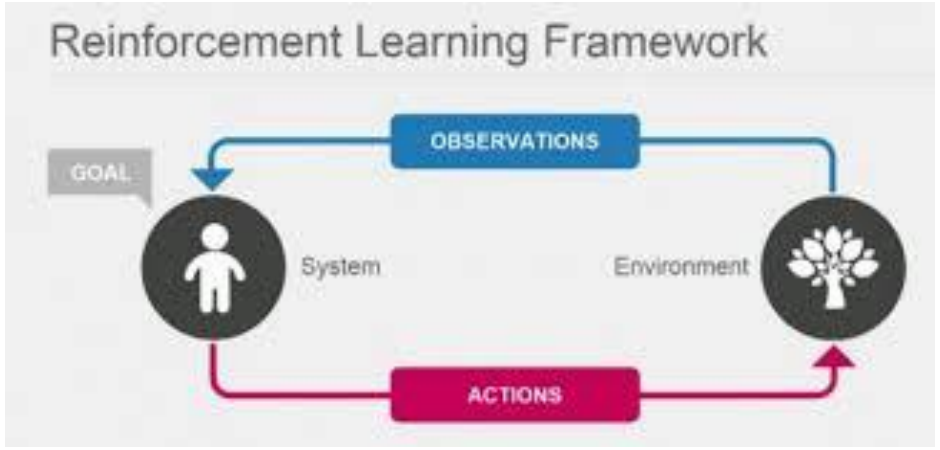


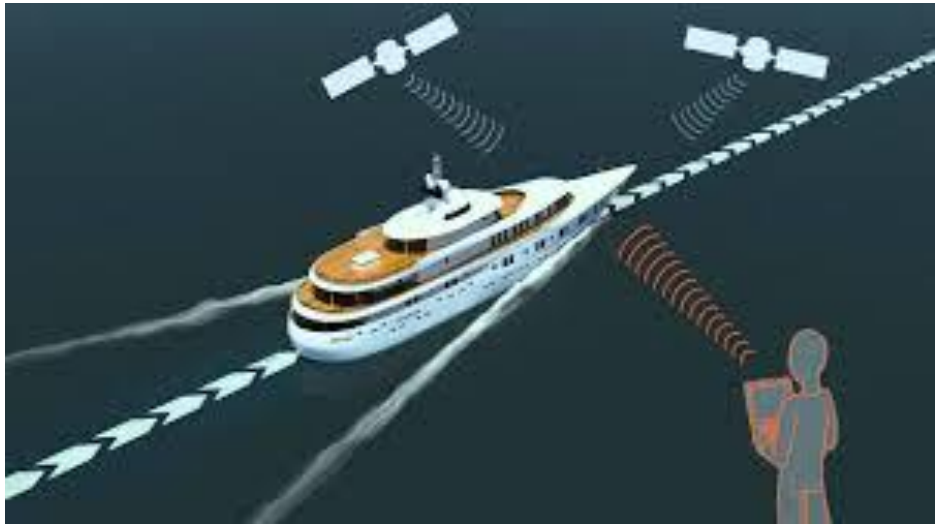
Role of adversarial training

- The ML model for ship tonnage estimate has had an adversarial training that considers disturbances and concealments
- Therefore the classification takes place correctly even if the captain of the ship has promptly embarked ballast water to conceal the tonnage decrease
- The International Convention on Tonnage Measurement of Ships formula for calculating gross tonnage of a vessel, says GT = K * V. Here, K = .2 + .02 * $\log_{10}(V)$, and V = interior volume of the vessel in cubic meters.

ML models integration

- The local intelligence of the patrol vessel is ML unsupervised model based on reinforcement learning, designed to maximize an objective function and implemented on an on-board microcontroller
- Based on the accuracy of the DL transshipment detection and the distance of the ship intercepted by the territorial waters of his country, the unsupervised model will choose the action to be requested:
  - Start a reconnaissance drone that adds close-ups to the satellite detection of the load profile
  - Activate a cyber-attack Global Navigation Satellite System (GNSS) spoofing, to take the target ship off course.



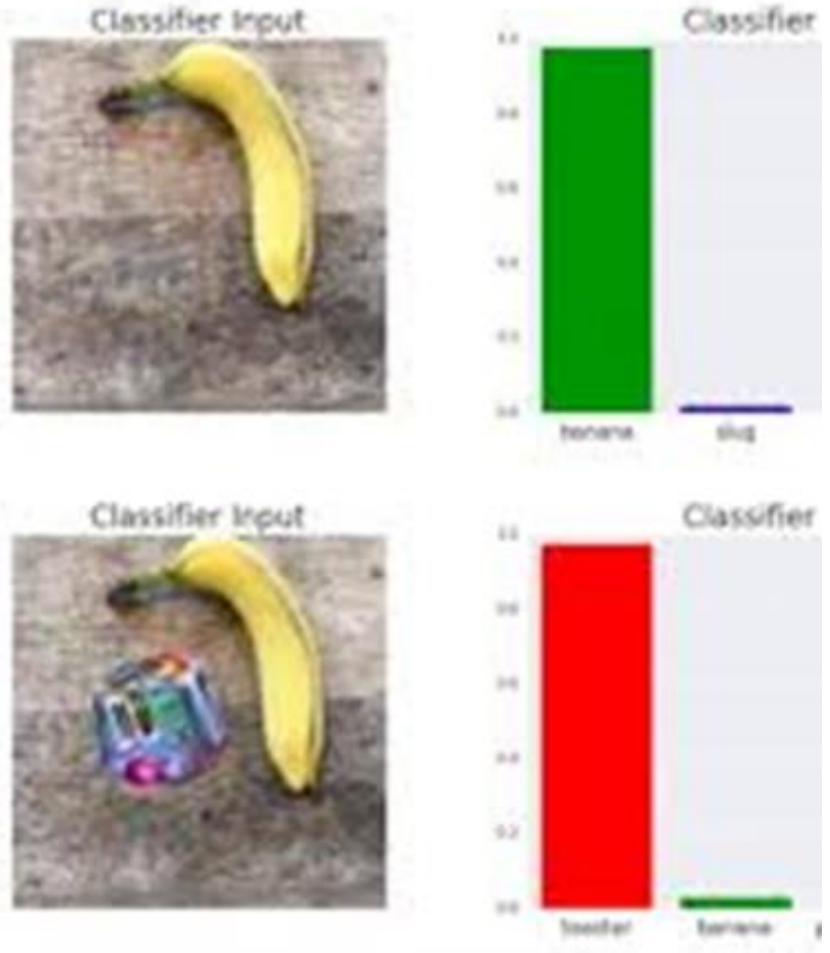


Comments

- The decision did **not** climb a command and control line: the supervised classifier for the transshipment provided input (**not** orders) to the reinforcement learning controller, who acted autonomously.
- Information collected in a domain (space) of the virtual battlefield was used to make a decision in another domain (cyberspace)
- All subsystems are already available: unmanned vessels, air and land vehicles, automatic estimates of collateral damage, and systems for automating the deployment of surveillance drones are all products already offered at trade shows.

Comments

- The use of autonomous AI systems offers clear opportunities to achieve greater accuracy and better coordination on the battlefield, although it is less certain that it can reduce the operating costs of the weapon systems. The literature speaks of a more efficient use of human resources, but complex legal, economic, social and security issues remain to be evaluated.
- Cyber security of AI plays an important role: ML models are nothing more than software or firmware and are not immune to code manipulation, and do not escape pollution of training examples.
  - Scenarios composed of several connected subsystems such as our example increase the risk of manipulation of models during training or production.
- There is also the risk that we will choose to fight a "war between AI" using generative models to create perceptive "anti-patterns" able to deceive the enemy's ML models.



جامعة خليفة
Khalifa University

Thank You

ku.ac.ae