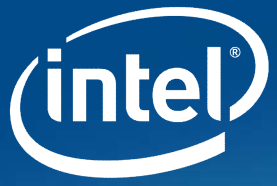


Interactions between Artificial Intelligence and Security

Sohrab Aftabjahani, PhD, SIM, SAM
Staff Security Researcher, Pre- and Post- Silicon Security Validation Lead,
Artificial Intelligence Product Group, Intel Corporation

26th Annual Electronic Design Process Symposium (EDPS) on Efficient Design and Manufacturing

SEMI, Milpitas, California

Oct 3-4, 2019

My Definitions

Artificial Intelligence and Machine/Deep Learning

- Artificial Intelligence (AI): A program that can sense, reason, act, and adapt
- Machine Learning (ML): A subset of AI with algorithms & statistical models whose performance improve as they are exposed to more data over time. Features are selected manually. E.g. Decision Trees, Support Vector Machines, and Ensemble methods.
- Deep Learning (DL): A subset of machine learning in which multi-layered neural networks (similar to neural pathways of the human brain) learn from vast amounts of data. Features are automatically learned. E.g. Convolutional neural networks (CCN's), Recurrent Neural Networks (such as Long Short-term memory), and Deep Q networks.
- Model: The trained program that predicts outputs given a set of inputs.
- Algorithm: set of rules/instructions that will train the model what to do.

My Definitions

X Neural Networks (xNN's)

- Artificial Neural Network (ANN): A framework for many ML algorithms to work together and process complex data inputs.
- Deep Neural Network (DNN): An ANN with multiple layers between the inputs and outputs.
- Convolutional Neural Network (CNN): A class of Deep Neural Networks – mostly used to analyze visual data.

High Level Machine Learning Classification

Machine Learning (ML)

- Artificial Neural Networks and Deep Learning: Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Generative Adversarial Networks (GAN), Autoencoders
- Supervised Learning: Classification, Regression
- Unsupervised Learning: Transformation (Dimensionality Reduction), Clustering, Outlier Detection
- Ensemble Learning: Stacking, Bagging, Boosting
- Reinforcement Learning

How is a ML model created?

Build Mathematical Models

- Neural Nets: e.g. ANN, CNN
- Probabilistic Classifiers: e.g. naïve-Bayes Classifier

Training: How to program the Mathematical Models?

- Training Dataset (input vector, output vector) pairs

ML Model Validation: How accurately a predictive model perform?

- Testing/Validation Dataset: Find the prediction errors.

Uses of ML/DL

General usage of ML/DL

- Classification
- Prediction

Example applications of ML/DL

- Autonomous driving
- Abnormality detection
- Surveillance

Extending the usage of ML/DL to

- Safety
- Security

Interactions between Security and AI

How AI affects security of systems?

- Use AI to facilitate security attacks
- Use AI to mitigate/defend against security attacks

How security affects the AI systems?

- Create secure AI Training engines to develop trustworthy AI Models
- Evaluate the trustworthiness of AI Models
- Create secure AI Inference engines to guarantee trustworthiness of the AI Model Usage

How can one improve the other one?

- Use Design for Security to create trustworthy and secure AI systems and provide Security Assurance for AI systems.
- Use AI to improve Design for Security and Security Assurance of systems (including AI systems).

Current uses of AI/ML in Security

Defense/Mitigation

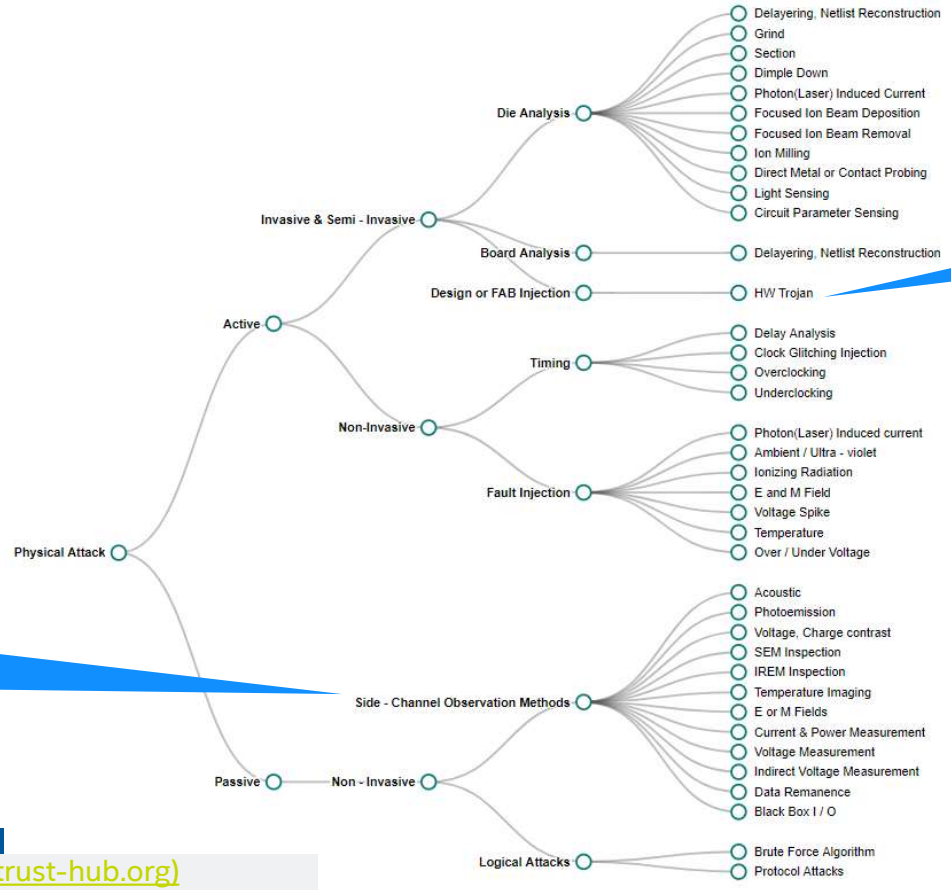
- Abnormality Detection
  - Malware Detection (Network Security, Operating System Security, Application Security)
  - Hardware Trojan Detection (Hardware Security)

Attacks

- Facilitating Side Channel Analysis (Weakening the security guarantee of Cryptography)
- Creating ML models to bypass the foundations for security (Modeling Physical Uncloneable Functions PUF's)

Physical Attacks Taxonomy & Vulnerability Database

Mission Statement	Physical Vulnerabilities	SOC Vulnerabilities (Coming Soon)	CAD Solutions for Security

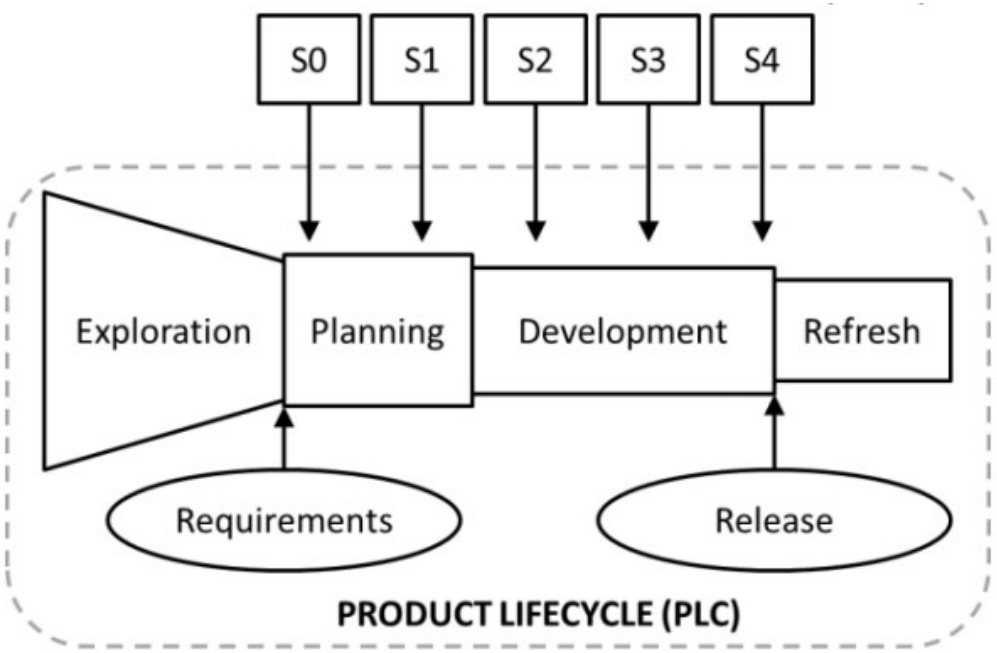


HW Trojan

Side channel Analysis

Source: Vulnerability Database (trust-hub.org)

Product Lifecycle (PLC), Security Development Lifecycle (SDL)



S0 – Security Risk Assessment
S1 – Security Architecture Review
S2 – Security Design Review
S3 – Pre-Silicon Security Validation
S4 – Post-Silicon Security Validation

Source IEEE Publication: http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6224330

Security Validation of ML/DL systems?

(a) In general, SDL can be used for high quality security validation of HW/SW (as a system) although it is very difficult problem by itself.

- Validation of security requirements
- Validation of Threat Models
- Validation of Security Architecture
- Validation of Design of Security
- Validation of Implementation of Security

(b) There are specific validation problems not very well solvable by methods used for (a)

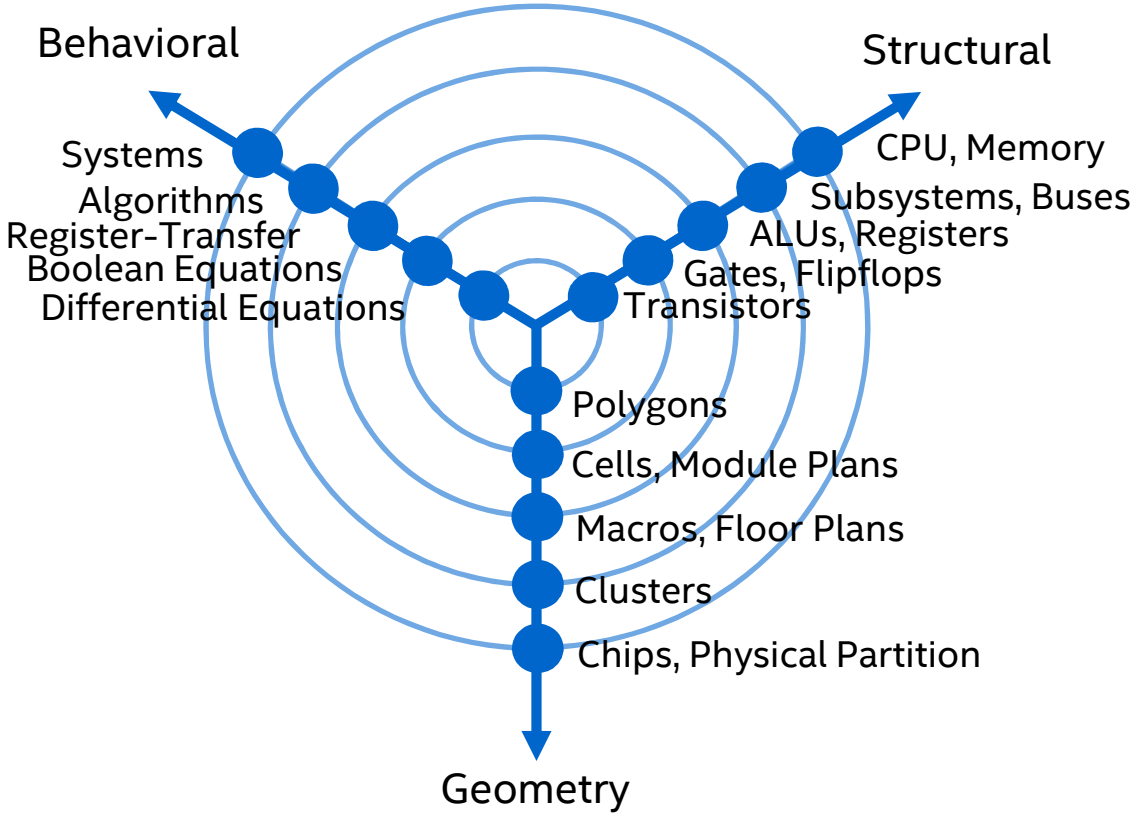
- Validation of training data is still a challenging problem.
- Validation of the ML/DL models to guaranteed they always give correct answers is a big challenge.

Typical Security Requirements of ML/DL Systems

Type	Usage	Training Dataset	Input Data	Model Graph Type	Model Weights	HW/SW Training	HW/SW Inference	Output
ML	Training	C?,I	-	-	-	I,A?	-	-
ML	Inference	-	C?,I	C?,I	C?,I	-	I,A?	I,C?
DL	Training	C?,I	-	-	-	I,A?	-	-
DL	Inference	-	C?,I	C?,I	C?,I	-	I,A?	I,C?

* Classical Security Properties: C: Confidentiality, I: Integrity, A: Availability

Design Database: Y Chart – Design Domains & Levels



Source IEEE Publication: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1654264

Opportunities to use AI/ML for PLC

PLC involves a lot of Data including the following:

- Product Requirements -> Business Data
- Architecture and Design -> Design Data
- Pre-Silicon (Design Automation) -> Simulation Data
- Post-Si (Manufacturing Automation) -> Test Data
- In-Field (System/on-line testing) -> Customer Data (Bugs)

There are many opportunities to use AI/ML for optimizing products and PLC.

Opportunities to use AI/ML for SDL

SDL involves a lot of Data including the following:

- Product Security Requirements -> Business Security-related Data
- Security Architecture and Design for Security-> Design Security-related Data
- Pre-Si Security Validation (Design Automation) -> Simulation Security-related Data
- Post-Si Security Validation (Manufacturing Automation) -> Test Security-related Data
- In-Field (System/on-line testing) -> Product Security Vulnerability Data (Bugs)

There are many opportunities to use AI/ML for increasing the security robustness of a product and optimizing executing SDL.

Potential Uses/Challenges of AI/ML in Security

Potential Uses

- Guiding for Design for Security.
- Better guiding and optimizing SDL steps to free up engineering time to target the areas with less coverage (corner cases).
- Optimizing Security-Aware EDA CAD tools.

Challenges

- Continuous Learning ML Models become capable to of detection of malicious inputs
- Overcoming the large effort and time required for training the models
- Improving the accuracy of models to 100% (Without overfitting is it really possible?)

Challenges for using ML/DL models in Security

ML/DL models can be used for solving specific set of problem.

- Can ML models be used to provide 100% accurate categorization, interpolation, extrapolation? Not really, but they can speed up finding an approximate solution for a problem.
- Is it possible to keep the quality of ML good enough as the number of variables in the input vectors are increased? No always very easy.
- Does training data represent the testing data? Ideally, it should be, but not always true.
- Are there curve fitting errors? Yes (In general) and No (Over-fitted – Not really a good model).
- Is it possible to have a large amount of reliable data (to avoid creating an adversarial model) for training a model? Not really.
- Are models portable from one Company/Group/Technology to another? Not really, as their quality depends on the sampling space for the training data. It can create biases for the models not to make them usable for a more general dataset.
- How can the ML model be validated (always providing a correct answer)? Very difficult problem.

Takeaways

AI and security have a strong interaction.

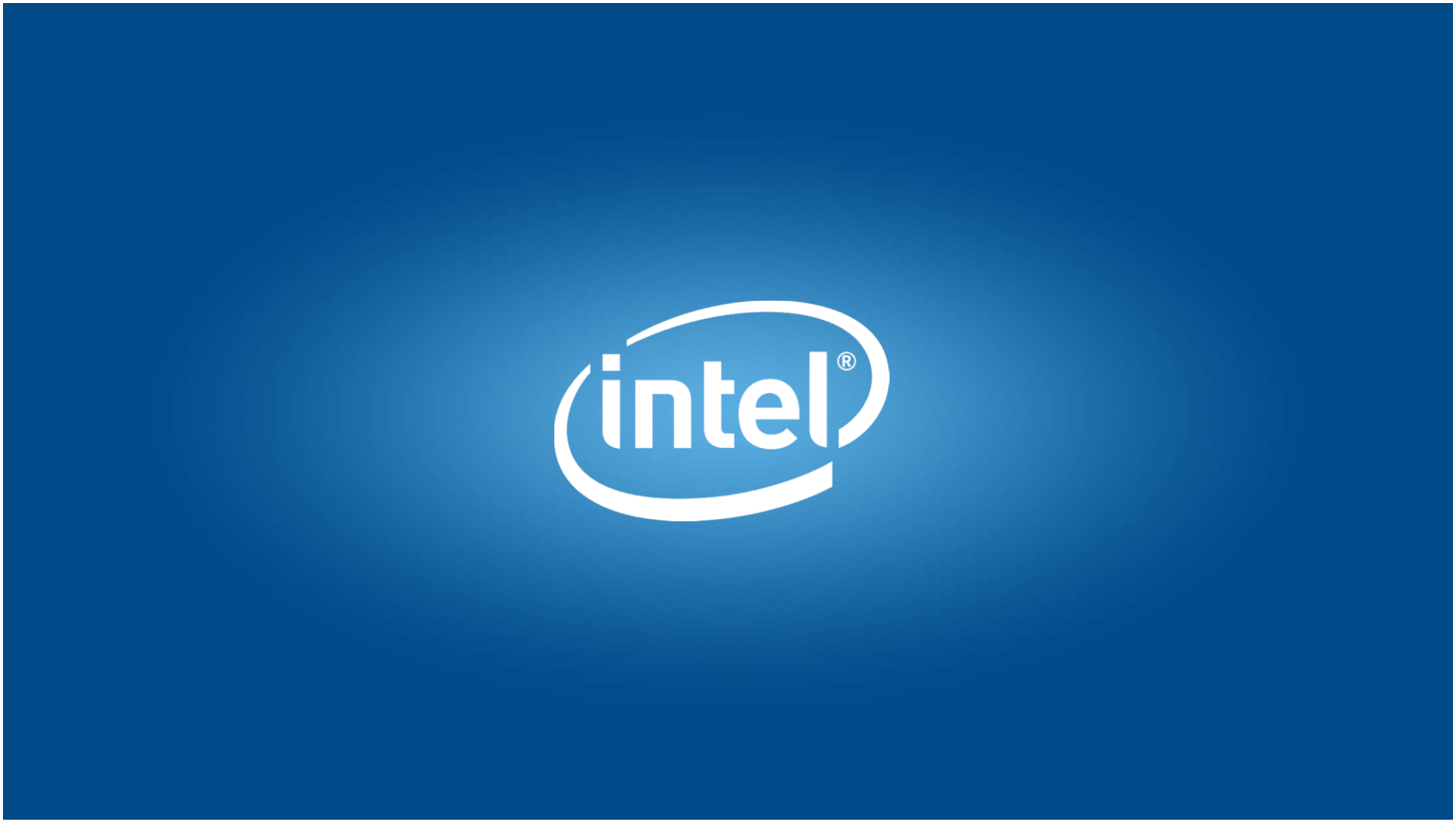
- AI can facilitate security attacks and coming up with mitigations to them with a potential for improvement of the Design for Security and Security Assurance practices.
- Design for Security and Security Assurance practices can potentially see improvements using AI.

Although using AI for solving problems in security is promising, there are challenges.

- Typical accuracy of the AI models in categorization and prediction (close to 98%) has surpassed the capability of humans; however, security assurance problems still require 100% accuracy.
- The accuracy of the models depends on quality of dataset, training, and validation as well as time and cost factors.

AI Systems for Training and/or Inference require rigorous Design for Security and Security Assurance.

- For mission critical AI applications, this requirement is stringer as lack of security can be disastrous.


intel®