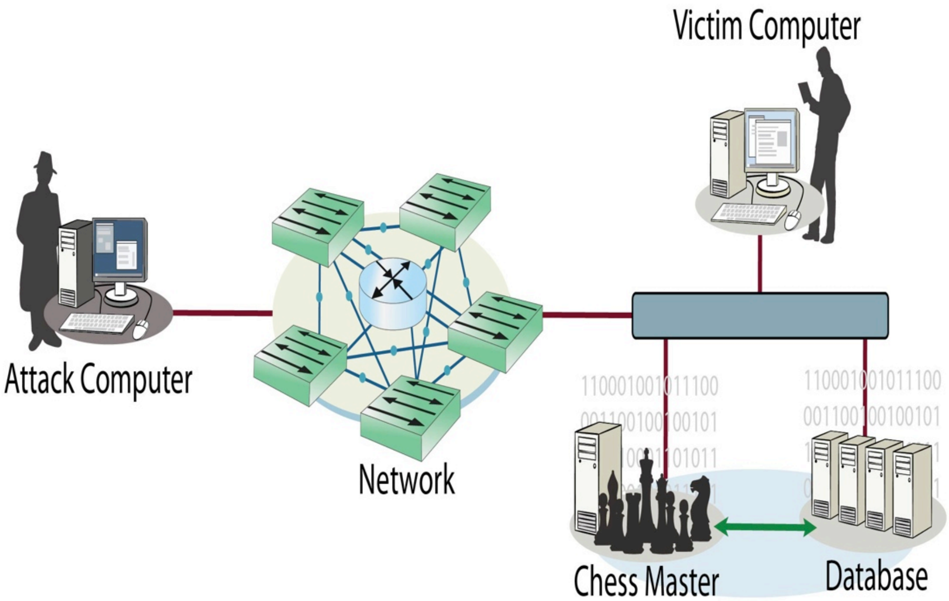
Adrian R. Chavez
Sandia National Laboratories

Dynamic Defense

Cybersecurity for Energy Delivery Systems Peer Review
August 5-6, 2014

Summary: Dynamic Defense

- **Objective**
  - Identifying and actively defending against past, present and future attacks within an ICS setting
- **Schedule**
  - February 2013 – December 2014
  - Develop Proof-of-Concept Machine Learning Algorithms to detect attacks and actively respond
  - Automatically trigger specific responses for specific attacks



- **Total Value of Award:** $500K
- **% Funds expended to date:** 71%
- **Performer:** Sandia National Laboratories
- **Partners:** Tennessee Valley Authority

Advancing the State of the Art (SOA)

- Current approaches have developed a framework for automatic response and deception within an IT setting
- Our approach is based on machine learning algorithms to classify traffic and host measurements and respond accordingly
- R&D is driven by TVA input and control system specific datasets
- Logging and alerting for interactive responses
- We are focused on ICS based systems and meeting the unique environmental constraints inherent to these systems

Challenges to Success

- **Detecting attack vs. benign measurements**
  - Using an ensemble of machine learning algorithms with results that match or improve upon existing classifiers
- **Respond to an attack with an appropriate response**
  - Initially focused on responding to know attacks with pre-determined responses
  - Future implementations will dynamically choose response strategy
- **Classify traffic while meeting ICS unique constraints**
  - Leverage training data and feature sets to quickly classify traffic

Progress to Date

- **Implemented proof-of-concept prototype for detection**
  - Leveraging Kyoto 2006 dataset
  - University of Mississippi State Datasets
    - Water Pump
    - Gas Pipeline
    - Powersystem
- **Preliminary results of classifying datasets**
- **Implemented a framework for appropriate response strategies**
  - Cocooning
  - Network Randomization

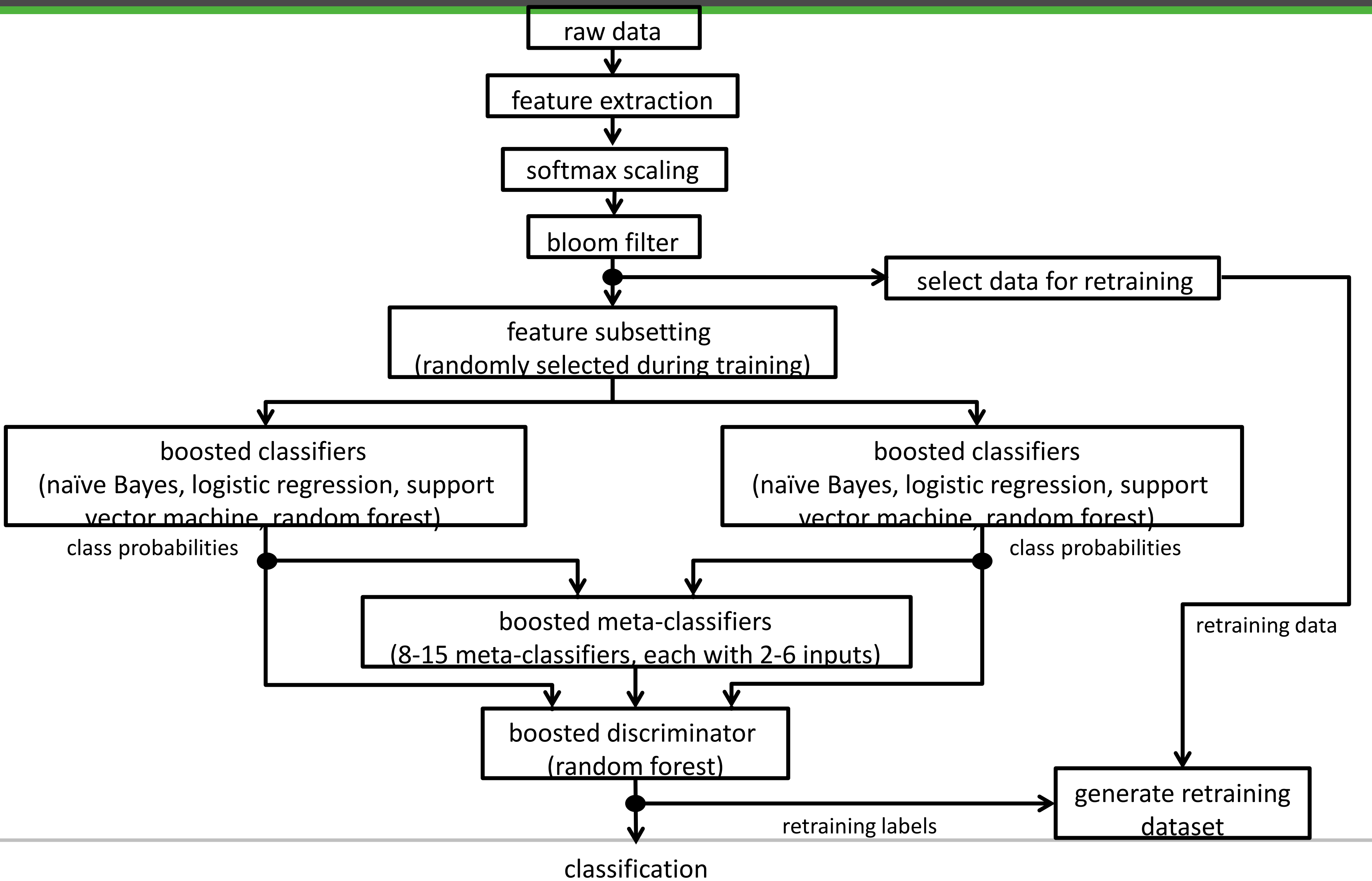
Collaboration/Technology Transfer

- **TVA providing requirements and input throughout R&D**
- **Accepted into Department of Homeland Security (DHS) Transition To Practice (TTP) Program**
  – Seeking additional partners to pilot technology in a representative environment
  – Continue to engage industry in use-case/applications of our solution
- **Transition technology into OPSAID reference implementation**
  – Lemnos is going through IEEE standardization process
    - Vehicle to harness our solution

Next Steps for this Project

- **Test algorithms using additional data sets**
  - Utilize internal data sets with SCADA traffic
  - Working with TVA
- **Integrate network based detection methods**
  - Currently using sequences of System call analysis + system info
- **Continue to gather performance metrics**

Framework



raw data
feature extraction
softmax scaling
bloom filter
select data for retraining
feature subsetting
(randomly selected during training)
boosted classifiers
(naïve Bayes, logistic regression, support vector machine, random forest)
boosted classifiers
(naïve Bayes, logistic regression, support vector machine, random forest)
class probabilities
class probabilities
boosted meta-classifiers
(8-15 meta-classifiers, each with 2-6 inputs)
retraining data
boosted discriminator
(random forest)
generate retraining dataset
retraining labels
classification

Results (1)

- **MCC is Matthew's Correlation Coefficient (http://en.wikipedia.org/wiki/Matthews_correlation_coefficient).**
- **AUC is the area under the receiver operating characteristic curve (http://en.wikipedia.org/wiki/Receiver_operating_characteristic#Area_under_curve)**
- **Recall is TP / (TP + FP)**
- **Accuracy is (TP + TN) / (P + N)**
  - MCC: 0.89532, recall: 0.91616, FPR: 0.027601, accuracy: 0.95338, TP: 27285209, TN: 56694092, FP: 1609213, FN: 2496904

Results (2)

	Recall	FPR	AUC
Signature IDS	0.09	0.016	N/A
Anomaly Detection	0.809	0.05	N/A
Max Entropy	0.773	0.02	0.72
Linear SVM	0.9895	0.035	0.963
Laplacian Eigenmap	0.64	0.087	0.759
Laplacian RLS	0.89	0.027	0.987
Ours (same test data)	0.9837	0.012	0.967

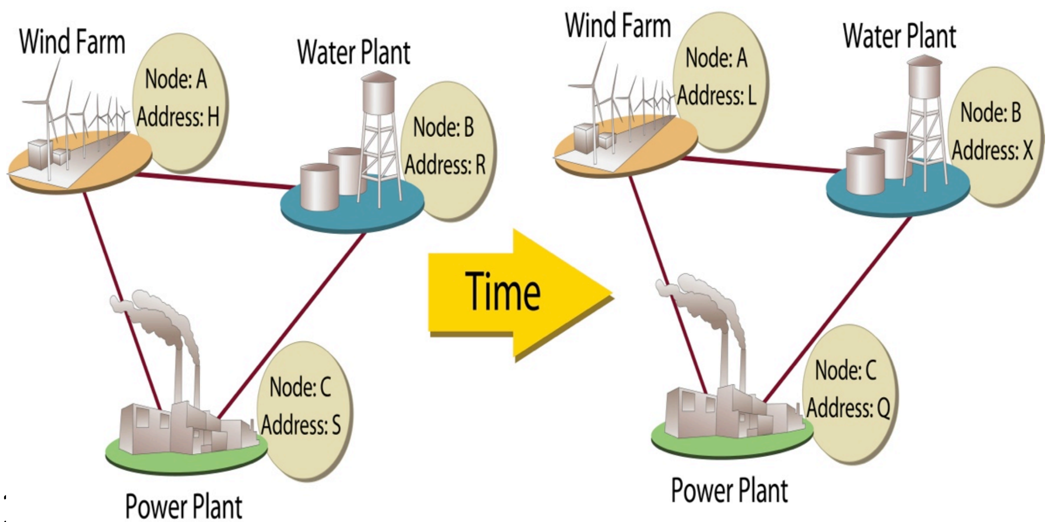
Adrian R. Chavez
Sandia National Laboratories

Network Randomization

Cybersecurity for Energy Delivery Systems Peer Review
August 5-6, 2014

Summary: Network Randomization

- **Objective**
  - Convert statically configured control system networks into dynamic moving targets
    - Create uncertainty
    - Eliminate targeted attacks
- **Schedule**
  - February 2013 - December 20:
  - We have Randomized:
    - IP Addresses (Dec 2013)
    - Port Numbers (Feb 2014)
    - Applications (Aug 2014)
    - Tested in a laboratory environment (300 nodes – Apr 2014)
  - Proof-of-concept implementation built-in OPSAID



- **Total Value of Award:** $250K
- **% Funds expended to date:** 65% (Through June)
- **Performer:** Sandia National Laboratories
- **Partners:** Tennessee Valley Authority

Advancing the State of the Art (SOA)

- **IP and port hopping implemented in traditional IT networks**
  - We consider combining the two within an ICS setting
- **Our approach leverages SDN technologies**
  - Open source - OpenFlow
  - Transparent to end devices
- **Randomization can be retrofitted into existing systems with OpenFlow capable hardware/software**
  - Increased difficulty in launching targeted attacks and gaining reconnaissance information

Challenges to Success (1)

- **Maintain network connectivity before, during and after randomization**
  - Allow configurable overlapping time windows when re-randomization occurs
- **Designing a scalable solution that can be applied on a large number of nodes and diverse set of end devices**
  - Randomization resides at the network level
    - Transparent to end devices
    - Network layer nodes < end device nodes
  - Tested in 300 node environment
- **Managing randomization across different networks**
  - Controller(s) communicate across network subnets

Challenges to Success (2)

- **IP Address Exhaustion**
  - Multiple subnets constrain IP address space
- **Lack of separate control network**
  - Receiving router needs to accept gratuitous ARPs to associate endpoints with overlay network
    - Separate control/data networks do not have this issue

Progress to Date

- **Port Randomization**
  - Leverage Linux iptables to manipulate port numbers entering/leaving network (host-based)
- **IP Randomization**
  - OpenFlow implementation that is transparent to end devices (host- or network-based)
    - Port Randomization can also be done here
- **Path Randomization**
  - Randomize path packets take through network
- **Application Randomization**
  - Compiler modifications to randomize instruction set

Collaboration/Technology Transfer

- **TVA providing requirements & input throughout R&D**
- **Accepted into Department of Homeland Security (DHS) Transition To Practice (TTP) Program**
  - Seeking additional partners to pilot technology in a representative environment
  - Continue to engage industry in use-case/applications of our solution
- **Transition technology into OPSAID reference implementation**
  - Lemnos is going through IEEE standardization process
    - Vehicle to harness our solution

Next Steps for this Project

- **Combine randomization schemes into a single solution**
  - Test and validate that each independent scheme does not interfere with one another
- **Collecting metrics for impact/effectiveness**
  - Red team assessment
  - Performance
- **Continue documenting results**
  - Aid the development of a new Interoperable Configuration Profile (ICP) for Lemnos IEEE efforts
  - Gather performance metrics

Questions?