SAND2021-14547C

Unclassified

LDRD

Laboratory Directed Research and Development

(U)SECURE: Science and Engineering of Cyber security by Uncertainty quantification and Rigorous Experimentation

Ali Pinar
Sandia National Laboratories
Livermore, CA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

What is the return of investment for cyber?

Credit: Staff Sergeant Jason Gamble, United States Air Force

We cannot improve,
what we cannot measure

(U) SECURE's story in a nutshell

(U)Cyber experimentation should be a pillar of science of cyber security, just as computational Science and Engineering (CSE) is a pillar of science.

- (U) Cyber experimentation is commonly used to answer questions about cyber systems
  - (U) but lack of rigor limits its use in high-consequence systems
- (U) To study complex cyber systems, we need to
  - (U) answer "what if questions" with high-confidence

Emulytics

  - (U) assess confidence in our results under uncertainty

Uncertainty Quantification

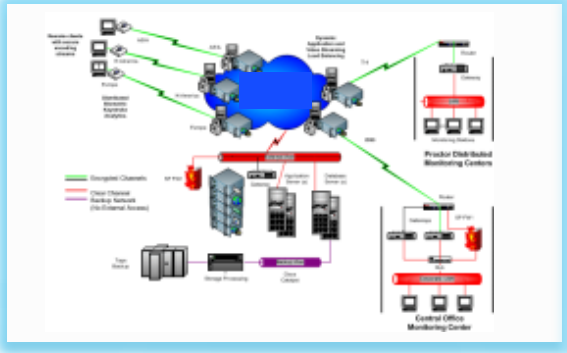
  - (U)make robust decisions under uncertainty in an adversarial environment

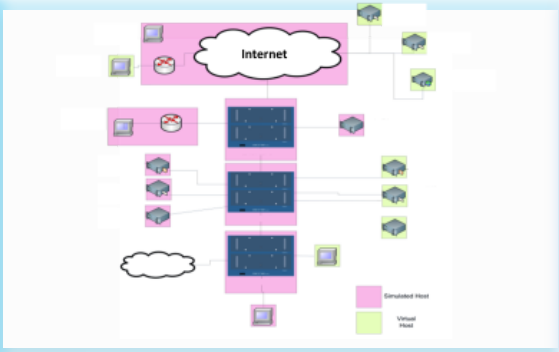
Adversarial Optimization

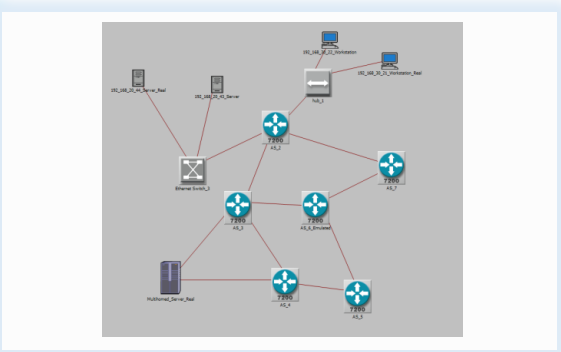
- *(U) Inspiration:* Sandia's know-how and capabilities from our nuclear stockpile stewardship
- *(U) Challenge: Cyber systems are different than physics-based systems*

(U) Cyber experimentation approaches

Figure unclassified







ACTUAL SYSTEM

EMULATION TESTBED

MATHEMATICAL MODELING

Interoperability in a single experiment

LIVE ← Increase Realism — Decrease Cost, Decrease Time → SIMULATED

REAL HARDWARE REAL SOFTWARE	ABSTRACT HARDWARE REAL SOFTWARE	ABSTRACT HARDWARE ABSTRACT SOFTWARE

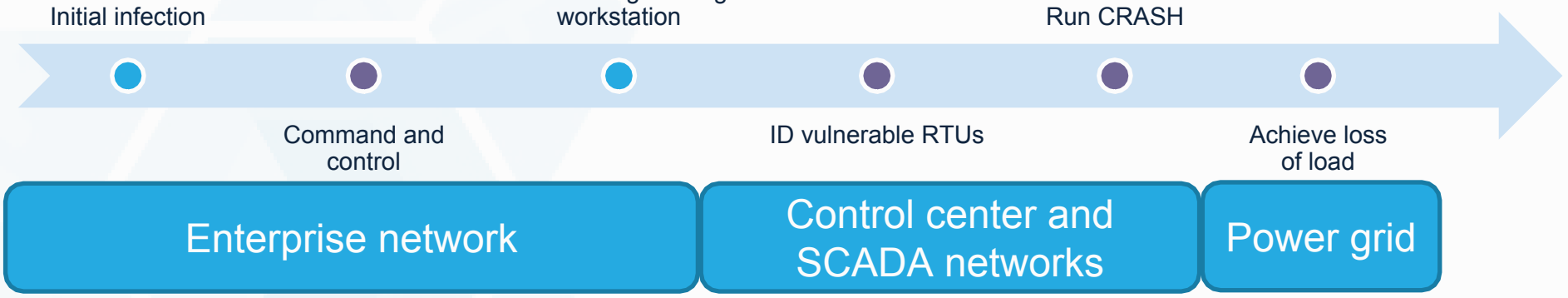
(U) SECURE's position:

- (U) Results should be independent of the platform and the tools used for the experiment

(U) Exemplar: How vulnerable is the power grid against a cyber attack?

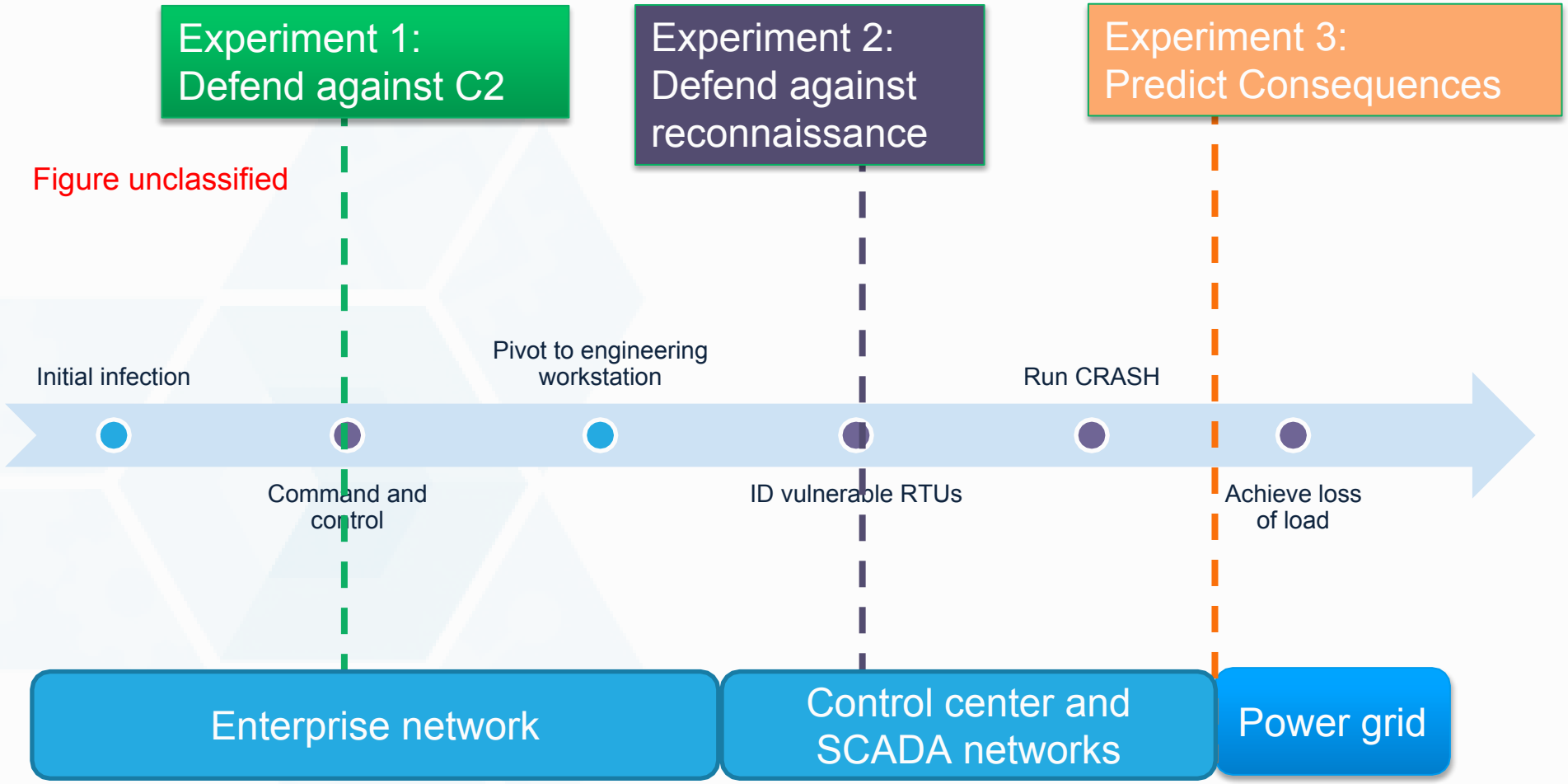
- (U) Ukraine attack was based on Crash Override Malware
- (U) The attacker gains remote access to power grid components to turn them on and off.

Figure unclassified

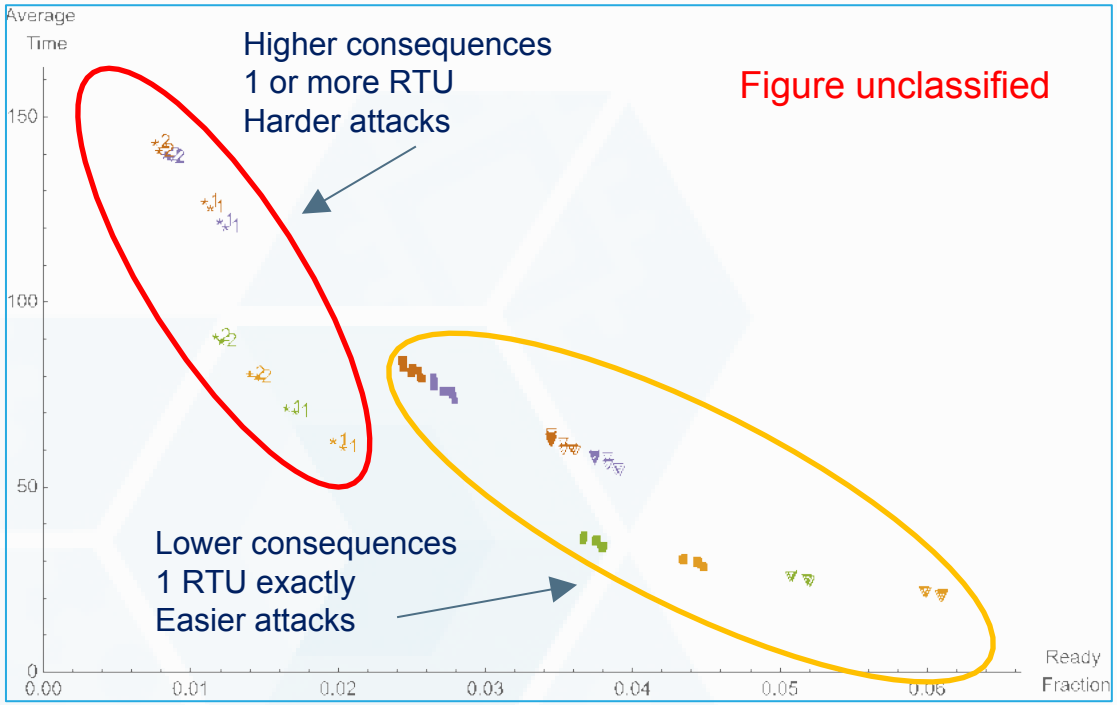


- (U) Goal: characterize loss of load resulting from malware infection in enterprise network
  - (U) Account for uncertainties in threat, network conditions
- (U) Approach: Piecewise studies to inform Markov transition probabilities and uncertainties

(U) We design an experiment for each step, and aggregate results with a Markov model



(U) Aggregated results



- (U) Plotting attack success metrics from Markov analysis: mean time to attack success vs. fraction of time in the "READY" state.
- (U)Extended our analysis framework to support **UQ in transition probabilities**, and **variations in each step's inherent timestep.**
- (U) Experiments provide range of transition probabilities (depending on scenario, attacker strategy, etc.)

(U) Defender goal: push attacker toward top-left of the plot (e.g. through better IDS)

(U) Each dot on the chart above represents a combination of C2 data, scanning/detection data, and attacker/defender strategy

(U) Markov analysis allows:
- Estimates of how secure the system is under attack
- Ranking of attacker/ defender strategies

(U) So What?

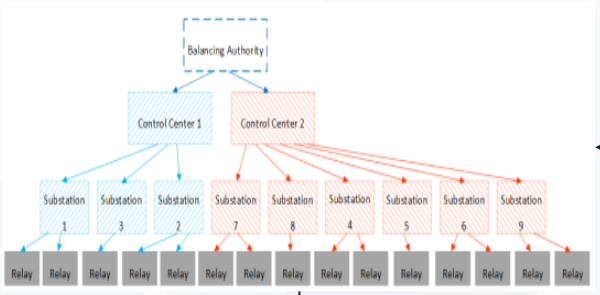
- (U) What changed?
  - (U) We produced an ***objective*** process that can ***quantify*** security.
    - (U) All assumptions are listed; all processes are repeatable;
    - (U) All experiments are verified; all models are validated;
  - (U) We have a scientific processes that can, and will be improved.
    - (U) No more disagreeing with expert opinions.
    - (U) Instead challenge assumptions; propose better algorithms/metrics.

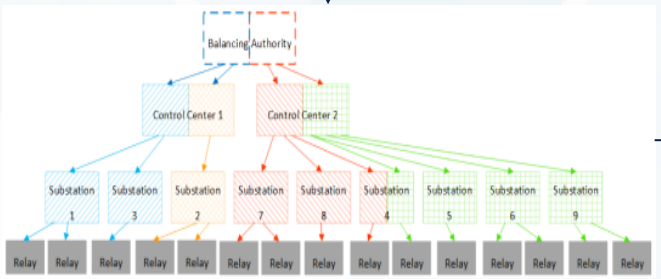
- (U) What can we do now? ***Quantifiable Security***
  - (U) Quantify return on investment for cyber security
  - (U) Identify critical components both for improving security and model fidelity
  - (U) Quantify attack consequences and enable mission-driven cyber security

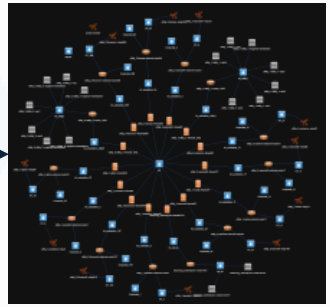
(U) Optimal Segmentation – rigorous comparison of two solutions

Figure unclassified

Optimization/Emulation Workflow

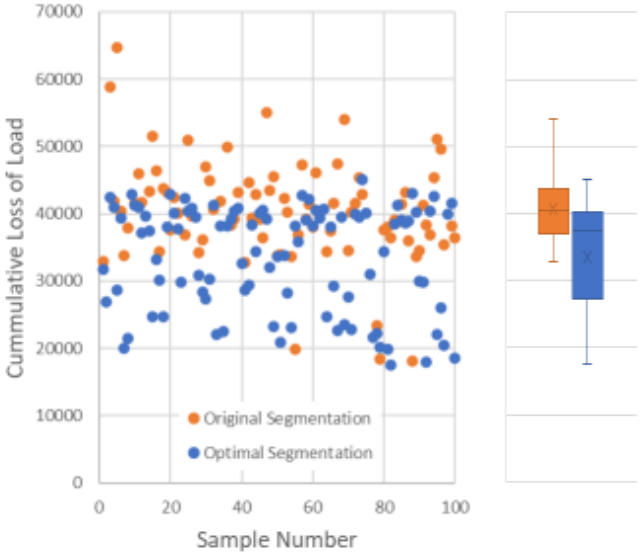






Results

Figure unclassified



(U) Takeaway: Designed a workflow that interfaces emulation with mathematical optimization to investigate network segmentation

(U) Takeaway: Mathematical optimization identifies a segmentation policy that is more robust under a CrashOverride attack

(U) Identifying extreme events is crucial

- (U) We need to identify events with low-likelihood yet high-consequence
  - (U) Solution: Multi-fidelity sampling for tail events; optimization for extreme points

Figure unclassified

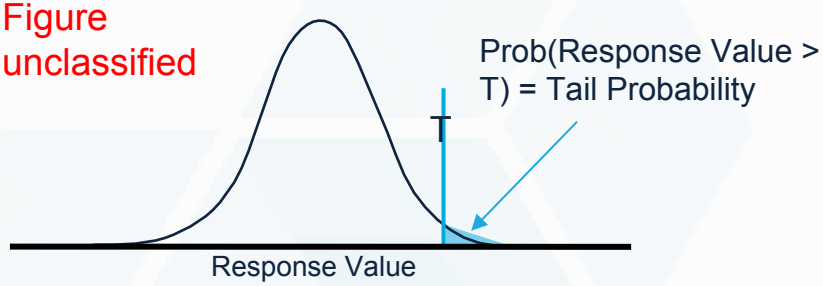
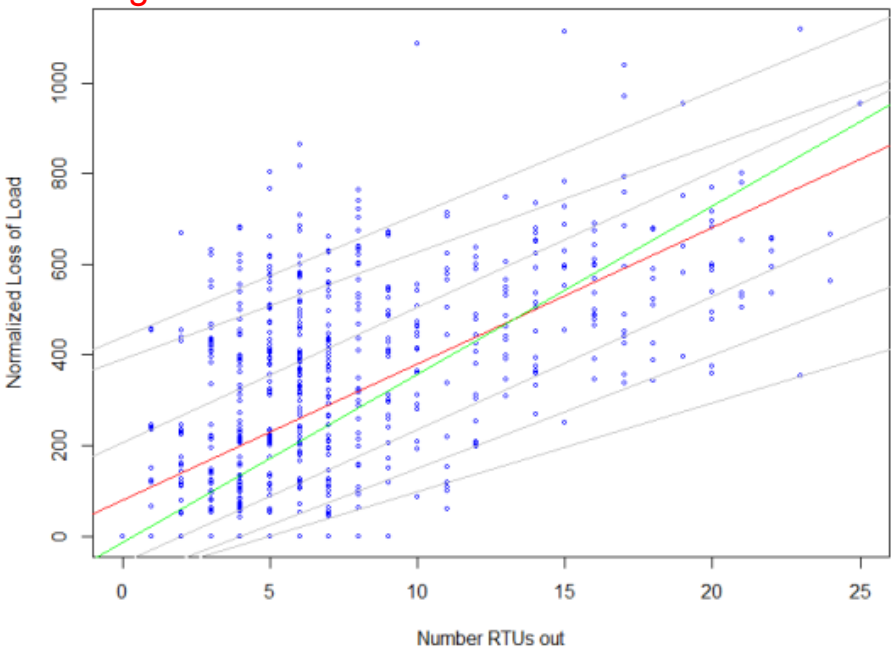


Figure unclassified



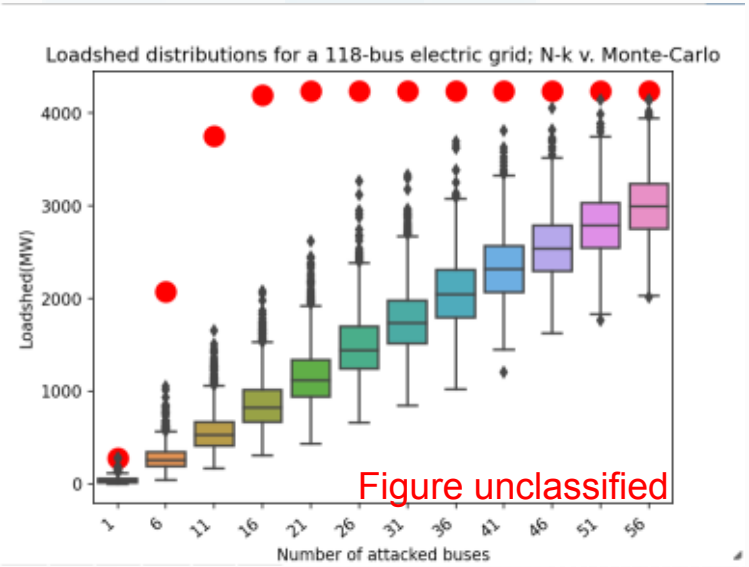


Figure unclassified

(U) Verify each experiment

- (U) Distribution of alert times shift as namespaces are added
- (U) Quantified similarity with Tukey Multiple Comparison Test
  - (U) Shows clear drop in similarity after 10 namespaces
- (U) Large p-value indicates that the null hypothesis can't be rejected
  - **(U) Larger p-value -> similar results**

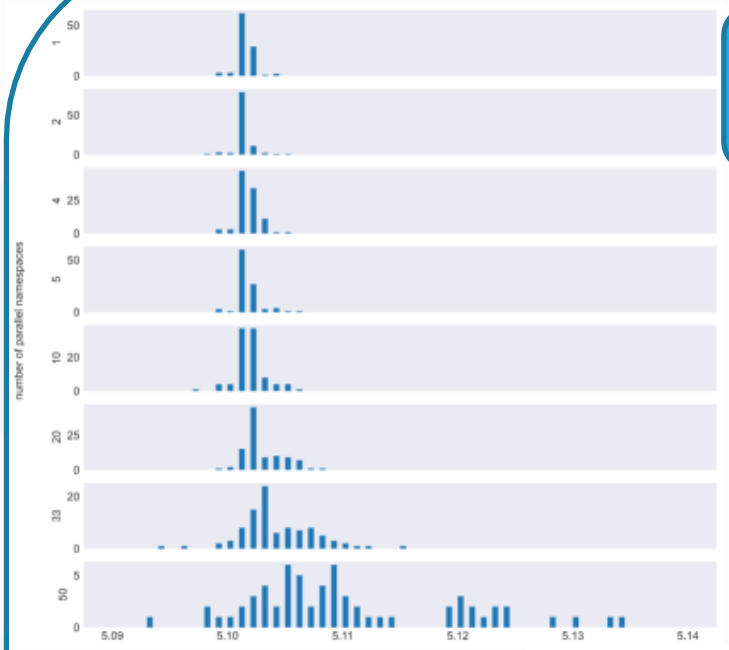
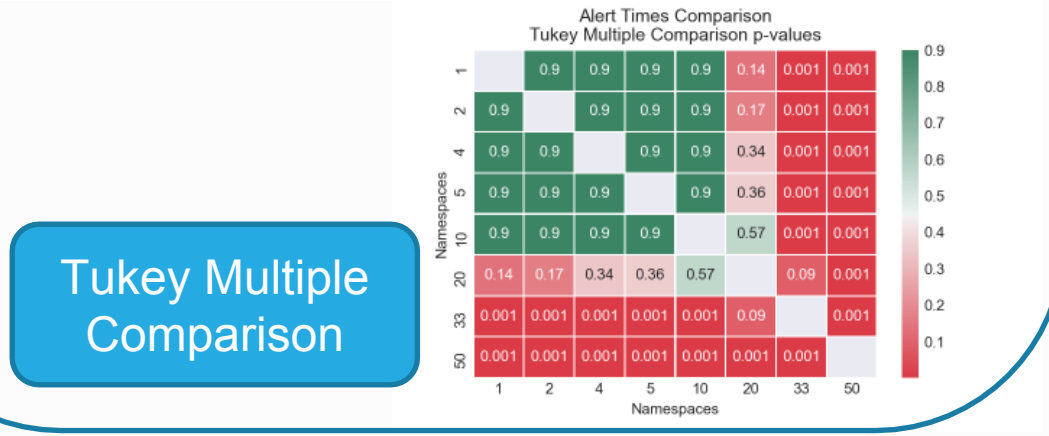


Figure unclassified



(U) Validate each model

- (U) Can we validate our models against data from real systems?
  - (U) Joint study with TAMU
- (U) Mean and median indicate good agreement. The low values of the 5th percentile between 100-120 seconds help identify times which have some realizations with less agreement.

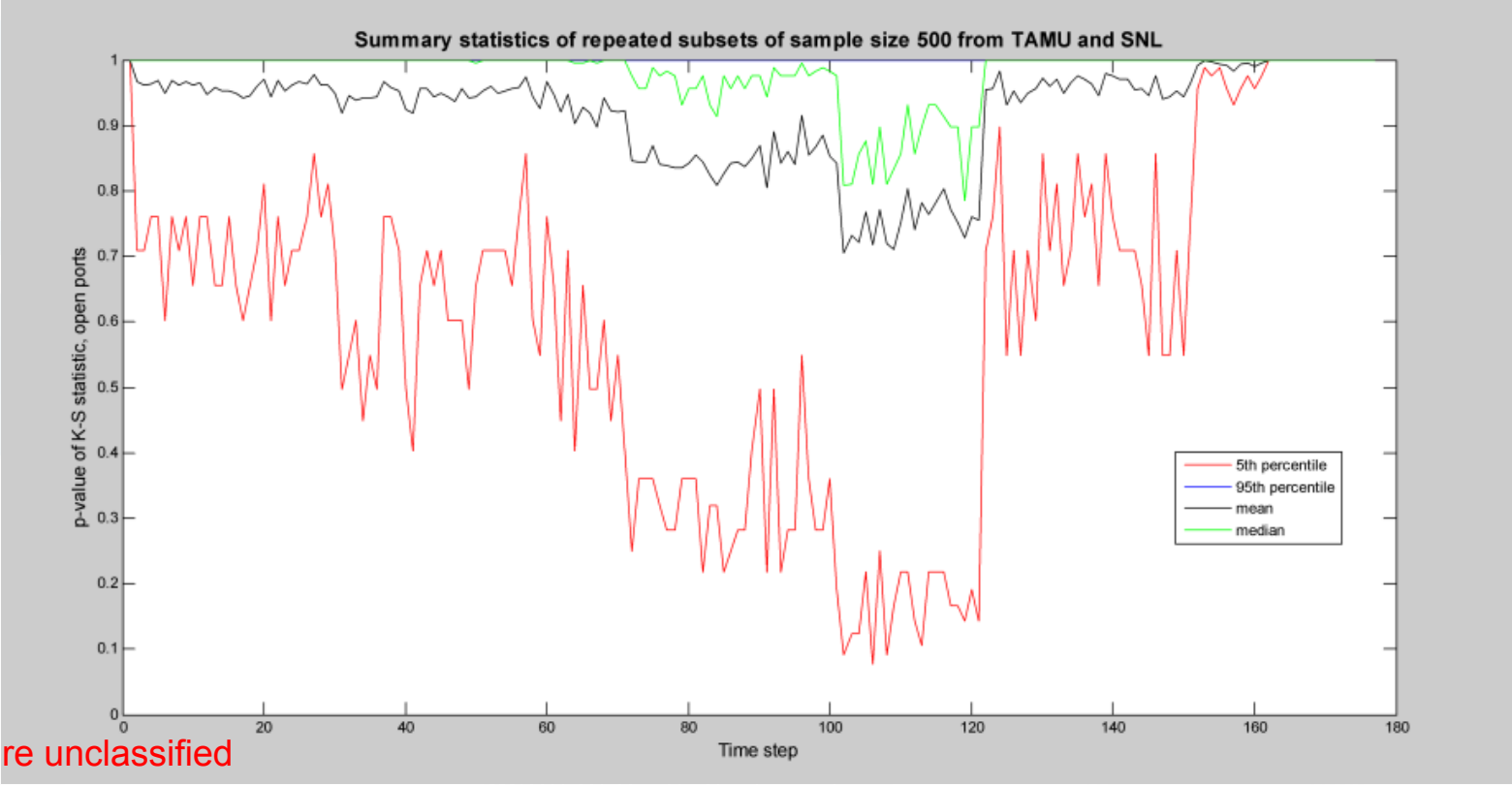
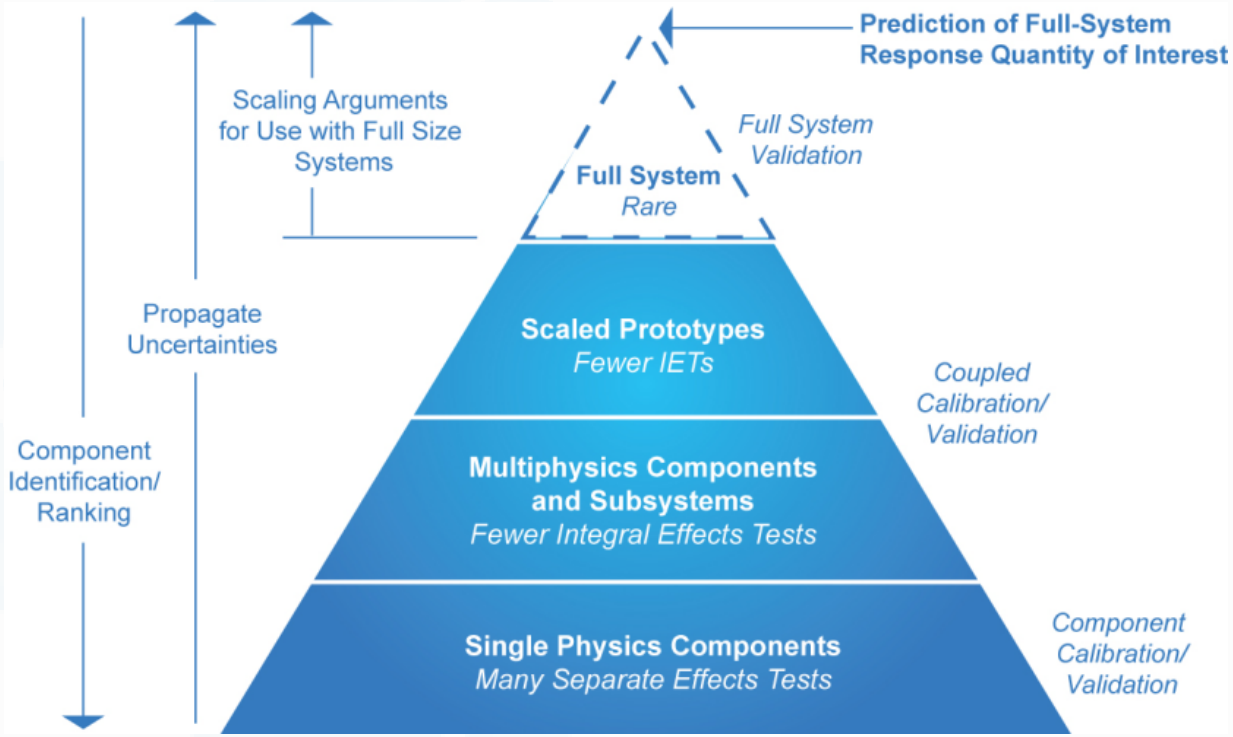


Figure unclassified

Validation Lessons from Weapons program: small scale tests → full system



Leverage information across the hierarchy

Example study for NC3 survivability/endurability

Experimental plan

- Experimental questions
- Identify inputs, outputs, topology

Topology

- Topology information from NC3
- Transfer this information to experimentation topology

Tools

- Scenario orchestration
- Fault/degradation injection (Netflix “Chaos Monkey”, but for experimental testbeds)

Efficient experiments

- Multifidelity models (emulation + math/simulation/surrogate models)
- Sampling strategies to comprehensively cover space of possible fault scenarios

Validation

- Validate low fidelity models against high fidelity models
- Start with small topology, exhaustively enumerate fault scenarios, test MF model against exhaustive results
- Where possible, compare to real world data

Rigorous Cyber Experimentation can provide NC3, what CSE provided to the nuclear weapons programs

- We cannot improve what we cannot measure
- Cyber experimentation provides measurements and is an essential tool for designing future complex systems
  - Rigor is paramount for high-consequence systems
- We need to look at the whole system and
  - build our confidence bottom up
  - tailor requirements top down
- Computation Science and Engineering (CSE) is a pillar of our nuclear weapons programs
  - Inspiration behind SECURE
- SECURE has been developing methods and tools to bring rigor into cyber experimentation. I can be used to
  - Assess a system and/or its components
  - Set justifiable requirements for components
  - Enable survivability/endurance by design
- We worked with NC3 in mind, and we are ready to face this challenge