Summary of Session 2

Sy-Yen Kuo

- Takashi Nanya: FAQs on Dependability
- Roy Maxion: Dependable Design of Experiments: Theory and Exemplars
- Marco Vieira: Perspectives on Dependability and Security Benchmarking


HAPPY NEW YEAR
2017
THE YEAR OF ROOSTER

FAQs on Dependability - for the DSN community to answer

- Correct service vs. intended service vs. expected service
  - Deterministic vs. probabilistic
- FAQs on dependability for higher-level layers
  - Higher-level dependability matters more than lower-level
  - Need a "bottom-up" approach to define system boundary so that dependability efforts work as well at a higher-level layer
  - Raise a lot of questions and try to give answers
  - Are those valid answers?
- FAQs on Economics of Dependability
  - Engineers' question: how much more dependability is achieved by the deployment from user's point of view.
  - Project Managers' question: What should we do to achieve high dependability within a given budget and delivery time ?

FAQs on Dependability - for the DSN community to answer

- CTO/CIOs' question: How much should we invest in it? How much return can we get ?
- Customers' question: How much more should I pay for it?
- Little work on mapping of dependability measures to economic value

- What we need

Quantifying relations between:

Dependability Efforts	Dependability Measures	Economic value

Dependable Design of Experiments: Theory and Exemplars

- Goal
  - Provide examples of how real experiments can (and do) go wrong (Context: using typing rhythms to detect stress)
  - Explain each of several potential design flaws.
  - Explain how to address each one.
- Needs
  - Validation !!
  - Everything must work as claimed
- What questions will (should) readers/reviewers ask about the work?
  - Apparatus/instruments, materials, subjects/objects being measured, instructions to subjects, data, design, procedure, Instructions to experimenters, analysis, inclusion/exclusion criteria
- What are your science-based needs?
  - How reliable do you need to be?
  - How accurate?
  - How consistent?

Dependable Design of Experiments: Theory and Exemplars

- Reproducibility
  - All materials and instruments are vetted, in the literature, documented, peer reviewed
  - Thorough and transparent documentation
  - All equipment identified and calibrated; calibration procedures are provided.
  - Experiment procedures are clearly laid out.
  - Code available.
- Avoiding experimental-design flaws
  - Make (up front) a clear and concise statement of the problem being solved.
  - Put it in the form of a claim.
  - Ask what kind of credible evidence would be needed to support the claim.
  - Back up every assertion or claim/subclaim with credible evidence
- Relation to dependability - Competent experiment design is a way of thinking ... just like dependability design

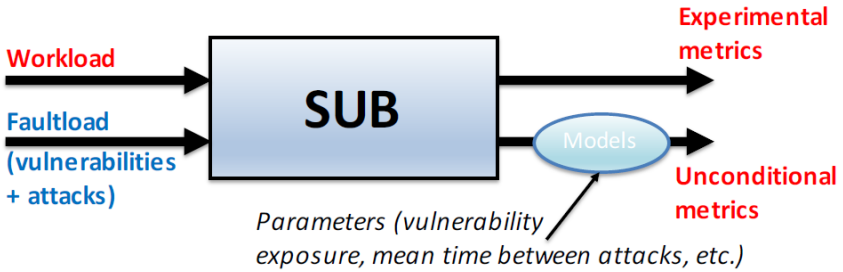
Dependable Design of Experiments: Theory and Exemplars

Summary

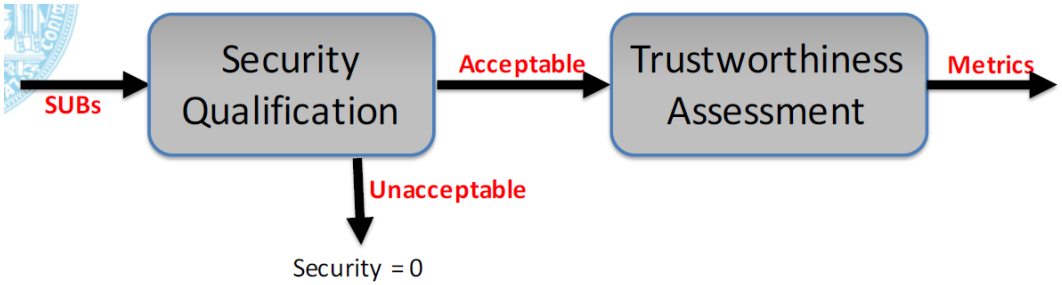
- Idea - Stress detection at the keyboard, using typing rhythms – it works
- Main points
  - How to make the experiment dependable
  - How to ensure confidence in the outcome
  - Rigorous experimental procedure
  - Reproducible
  - Good planning is essential
- What if less were done than was suggested?
  - Still valid? Still confident? Still dependable?

PERSPECTIVES ON DEPENDABILITY AND SECURITY BENCHMARKING

- TO BENCHMARKOR NOT TO BENCHMARK
  - Performance benchmarking: mostly for marketing
  - Dependability benchmarking: no endorsement from the industry
  - Security benchmarking: no common approach available yet
  - Towards the future: resilience and trustworthiness benchmarking
- Security benchmarking



  - Does not work if one wants to benchmark how secure different systems are
- A different approach



PERSPECTIVES ON DEPENDABILITY AND SECURITY BENCHMARKING

- Metrics
  - Portray trust from a human perspective
  - Dynamic: may change over time
  - Depend on the type of evidences gathered
  - Different metrics for different attack vectors
- What is wrong
  - Established benchmarks are mostly for marketing
  - Strict benchmarking conditions
  - Workload & faultload
  - Metrics
- A potential approach
  - Benchmarking conditions adaptable to the user needs
  - Include multiple usage scenarios
  - Use quality models instead of independent metrics
- Future: Resilience benchmarking and Trustworthiness Benchmarking
- Acceptance by “big” industry and acceptance by users?