

Testing and Certification of Trustworthy Systems Introduction to Minitrack

Alan R. Hevner
*Information Systems &
Decision Sciences Dept.
Univ. of South Florida
Tampa, FL 33620
ahevner@coba.usf.edu*

Richard C. Linger
*CERT Research Center
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213
rlinger@sei.cmu.edu*

Gwendolyn H. Walton
*Dept. of Mathematics &
Computer Science
Florida Southern College
Lakeland, FL 33801
gwalton@flsouthern.edu*

Modern society depends on complex information systems to control key infrastructures, with increasingly severe consequences of failure. Certifying the trustworthiness of these systems presents significant challenges. The Testing and Certification of Trustworthy Systems Minitrack focuses on research that will drive widespread use of rigorous testing and certification technologies. Specific topic areas include:

- New techniques for trustworthiness certification
- Testing and certification metrics and measures
- Testing attributes such as security and survivability
- Engineering practices and tools for certification
- Testing in system maintenance and evolution
- Specification methods to support system certification
- Role of correctness verification in system certification
- Industrial case studies in testing and certification

The Minitrack papers are summarized below.

In *Software Implemented Fault Injection for Safety-Critical Distributed Systems by Means of Mobile Agents*, authors Thomas Galla, Karin Hummel, and Roman Pallierer propose a fault injection method for testing fault-tolerant distributed real-time systems such as found in safety-critical automotive systems. An agent implementation language is used to model fault injection and system resources, agent migration, and logging of fault injection experiments.

In their paper *A Multi-layered Approach to Security in High Assurance System Development*, authors Jim Alves-Foss, Carol Taylor, and Paul Oman present a framework for design and verification of embedded trustworthy systems consisting of a partitioning real-time kernel, secure middleware, and I/O-restricted applications. Multiple levels of safety and security are supported based on separate layers of responsibility and control, with each layer enforcing its own security policy.

While survivability has emerged as a key property of information systems, author Vickie Westmark points out in the paper *A Definition of System Survivability* that there is little uniformity in how survivability is defined, measured, and computed. The paper establishes a reference baseline for understanding trends in computational survivability, and

provides a template for defining survivability to facilitate research in computational quality attributes based on standard definitions.

In their paper *What Makes a Code Review Trustworthy?* authors Stacy Nelson and Johann Schumann examine how code review can help ensure certification of trustworthy code. Technical leaders were surveyed to develop a view-based classification of important code review properties. The authors present the classification and argue that a uniform view of code review properties and tool capabilities can result in increased trust for safety-critical software.

In *Towards the Verification and Validation of Online Learning Systems: Application to RBF Neural Nets*, authors Ali Mili, Guangjie Jiang, Bojan Cukic, Yan Liu, and Rahma Ben Ayed argue that it is not possible to certify online adaptive systems, and in particular, online learning neural nets, with traditional testing and proving methods, because they rely on assumptions that do not hold for such systems. The paper introduces a framework for reasoning about and certifying online adaptive systems.

In *Breeding Software Test Cases for Complex Systems*, authors A. Watkins, D. Berndt, K. Aebischer, J. Fisher, and L. Johnson explore the use of genetic algorithms for testing complex distributed systems. They introduce a framework of environmental attributes that characterize complex systems failures and discuss visualization techniques to help understand and uncover failures.

Determining when to stop a statistical testing process is an important management decision. In *A Cost-Benefit Stopping Criteria for Statistical Testing*, author Stacy Prowell proposes a new stopping criterion based on the expected reliability of a system. The expected reliability is used, along with other factors such as units deployed and expected use, to anticipate the number of failures in the field and the resulting anticipated cost of failures.