

**CHENNAI – PONDICHERRY**

**An Empirical Evaluation of Mutation Testing for Improving the Test Quality of Safety-Critical Software**

**Abstract**

Testing provides a primary means for assuring software in safety-critical systems. To demonstrate, particularly to a certification authority, that sufficient testing has been performed, it is necessary to achieve the test coverage levels recommended or mandated by safety standards and industry guidelines. Mutation testing provides an alternative or complementary method of measuring test sufficiency, but has not been widely adopted in the safety-critical industry. In this study, we provide an empirical evaluation of the application of mutation testing to airborne software systems which have already satisfied the coverage requirements for certification. Specifically, we apply mutation testing to safety-critical software developed using high-integrity subsets of C and Ada, identify the most effective mutant types, and analyze the root causes of failures in test cases. Our findings show how mutation testing could be effective where traditional structural coverage analysis and manual peer review have failed. They also show that several testing issues have origins beyond the test activity, and this suggests improvements to the requirements definition and coding process. Our study also examines the relationship between program characteristics and mutation survival and considers how program size can provide a means for targeting test areas most likely to have dormant faults. Industry feedback is also provided, particularly on how mutation testing can be integrated into a typical verification life cycle of airborne software.