

**CHENNAI – PONDICHERRY**

**Reviving Sequential Program Birthmarking for Multithreaded Software Plagiarism Detection**

Abstract

As multithreaded programs become increasingly popular, plagiarism of multithreaded programs starts to plague the software industry. Although there has been tremendous progress on software plagiarism detection technology, existing dynamic birthmark approaches are applicable only to sequential programs, due to the fact that thread scheduling nondeterminism severely perturbs birthmark generation and comparison. We propose a framework called TOB (Thread-oblivious dynamic Birthmark) that revives existing techniques so they can be applied to detect plagiarism of multithreaded programs. This is achieved by thread-oblivious algorithms that shield the influence of thread schedules on executions. We have implemented a set of tools collectively called TOB-PD (TOB based Plagiarism Detection tool) by applying TOB to three existing representative dynamic birthmarks, including SCSSB (System Call Short Sequence Birthmark), DYKIS (DYnamic Key Instruction Sequence birthmark) and JB (an API based birthmark for Java). Our experiments conducted on large number of binary programs show that our approach exhibits strong resilience against state-of-the-art semantics-preserving code obfuscation techniques. Comparisons against the three existing tools SCSSB, DYKIS and JB show that the new framework is effective for plagiarism detection of multithreaded programs. The tools, the benchmarks and the experimental results are all publicly available.