Data Management Plan

Following the NSF Data Management Guidelines, we will administrate, archive, and maintain the data and the intellectual properties generated during this project. In this document, we describe the expected data that we will produce and collect, and a detail plan to openly manage them without violating user's privacy over the course of the project.

I. Expected Data. This proposed work is expected to generate new techniques, develop new systems and software tools, and have new curriculum materials and hands-on labs. In general, the data to be produced by the proposed research falls into the following categories:

1. **Software Prototype** includes source code of SGX infrastructure, toolchains and kernel extension, as well as the tools to secure SGX program and defend SGX malware.
2. **Experimental Data and Result** that characterizes the behavior of our prototype, and includes the details of experiment settings and analysis results, and the security benchmark suite to evaluate SGX programs.
3. **Academic Papers and Technical Reports** that describe the work proposed in this research.
4. **Curriculum Materials** including lecture notes, presentation slides, homework and solutions, projects, in systems security and hands-on labs.

II. Management Plan. As our goal is to make an tangible influence to our society, all the developed techniques and implementation, will be maintained publicly accessible, and tech-transfer to the open source team (e.g., Apache organization) for future maintenance. More specifically, in accordance with NSF data sharing policy for scientific research to ensure its usability, accessibility, and preservation, we use the following plan to manage each specific category of data above:

1. **Experimental Data**. The input and output data conducted to evaluate our prototype will be archived by the PI for the duration of the project and made public available. In addition, the source code and experimental data will also be permanently hosted as a digital archive operated by Georgia Tech Information Security Center (GTISC) and backed up by the D-Space system at University of Texas at Dallas. The operational team at GTISC is a group of permanent system administrators who are specialized in maintaining our group's infrastructure. The central repository (e.g., gitolite, code review system, monitoring, and backup services) is protected with RAID5 in local storage, and is backed up daily to a dedicated archiving server.
2. **Software and the Source Code**. For prototype that is built on top of open source project (e.g., GPL) will be correspondingly licensed using the same policy. If the software is entirely developed from scratch without using any other proprietary software, PIs will make it open source with detailed documentation using Apache license. In particular, we plan to make an early release of our research prototype via our group's GitHub repository at https://github.com/sslab-gatech/, and https://github.com/utds3lab. For instance, our preliminary system OpenSGX is already publicly available at https://github.com/sslab-gatech/opensgx.
3. **Academic papers**. Academic conference and journal papers, technical reports, presentation slides, posters, and other materials will be released through PIs' web site (e.g., https://sslab.gtisc.gatech.edu/. While published papers are subject to the discretion of copyright owners (e.g., ACM, IEEE, USENIX), the technical reports will be made freely available without any restriction.
4. **Curriculum Materials**. Curriculum materials will be openly shared with anyone in the world through PIs' web site. Except homework solutions and hands-on labs require permissions from instructors, all other materials is freely available.