Differential Privacy for Coverage Analysis of Software Traces (Artifact)

Yu Hao ✉
Ohio State University, Columbus, OH, USA

Sufian Latif ✉
Ohio State University, Columbus, OH, USA

Hailong Zhang ✉
Fordham University, New York, NY, USA

Raef Bassily ✉
Ohio State University, Columbus, OH, USA

Atanas Rountev ✉
Ohio State University, Columbus, OH, USA

Abstract

We propose a differentially private coverage analysis for software traces. To demonstrate that it achieves low error and high precision while preserving privacy, we evaluate the analysis on simulated traces for 15 Android apps. The open source implementation of the analysis, which is in Java, and the dataset used in the experiments are released as an artifact. We also provide specific guidance on reproducing the experimental results.

2012 ACM Subject Classification Software and its engineering → Dynamic analysis; Security and privacy → Privacy-preserving protocols

Keywords and phrases Trace Profiling, Differential Privacy, Program Analysis

Digital Object Identifier 10.4230/DARTS.7.2.7

Funding This material is based upon work supported by the National Science Foundation under Grant No. CCF-1907715.

Acknowledgements We thank the ECOOP artifact reviewers for their valuable feedback and the artifact co-chairs for their help.

Related Article Yu Hao, Sufian Latif, Hailong Zhang, Raef Bassily, and Atanas Rountev, "Differential Privacy for Coverage Analysis of Software Traces", in 35th European Conference on Object-Oriented Programming (ECOOP 2021), LIPIcs, Vol. 194, pp. 8:1–8:25, 2021.
`https://doi.org/10.4230/LIPIcs.ECOOP.2021.8`

Related Conference 35th European Conference on Object-Oriented Programming (ECOOP 2021), July 12–16, 2021, Aarhus, Denmark (Virtual Conference)

1 Scope

We provide the implementation of the randomization algorithms as described in the research paper. Our experimental evaluation was conducted based on the implementation and the input data, which is simulated software traces for 15 Android apps. The artifact includes both the implementation and the input data.

In the evaluation section of the research paper, we show the experimental results of the proposed differentially private trace coverage analysis for each app in terms of the following measurements:

- Error for all covered traces.
- Recall and precision of the identified hot traces.
- Comparison of the relaxed version and strict version of the hot trace identification algorithm, in terms of their recall and precision.
- Error for identified hot traces.

Dagstuhl Artifacts Series, Vol. 7, Issue 2, Artifact No. 7, pp. 7:1–7:3
DAGSTUHL ARTIFACTS SERIES
Dagstuhl Artifacts Series
Schloss Dagstuhl – Leibniz-Zentrum für Informatik,
Dagstuhl Publishing, Germany

- Comparison of three different privacy budget choices (ln(3), ln(9), and ln(49)) in terms of error for all covered traces.

Following the documentation of the artifact, one should be able to fully reproduce these results.

2 Content

The artifact package includes:

- A self-contained Docker image that includes the source code and data for reproducing the experimental results described in the paper.
- A detailed documentation (in PDF format) that provides guidance on how to use the artifact and how to reproduce the experimental results.

3 Getting the artifact

The artifact endorsed by the Artifact Evaluation Committee is available free of charge on the Dagstuhl Research Online Publication Server (DROPS). In addition, the artifact is also available at: `https://presto-osu.github.io/ecoop21/`.

4 Tested platforms

The artifact has been tested on two following platforms:

- Mac OS X 10.15.6
- Ubuntu 18.04.5 LTS (Bionic Beaver)

5 License

6 MD5 sum of the artifact

083c4e7d3d84ac562c141799e05fd330

7 Size of the artifact

706 MiB