Back-Propagating System Dependency Impact for Attack Investigation

Pengcheng Fang, *Case Western Reserve University;* Peng Gao, *Virginia Tech;* Changlin Liu and Erman Ayday, *Case Western Reserve University;* Kangkook Jee, *University of Texas at Dallas;* Ting Wang, *Penn State University;* Yanfang (Fanny) Ye, *Case Western Reserve University;* Zhuotao Liu, *Tsinghua University;* Xusheng Xiao, *Case Western Reserve University*

https://www.usenix.org/conference/usenixsecurity22/presentation/fang

This artifact appendix is included in the Artifact Appendices to the Proceedings of the 31st USENIX Security Symposium and appends to the paper of the same name that appears in the Proceedings of the 31st USENIX Security Symposium.

August 10–12, 2022 • Boston, MA, USA

978-1-939133-31-1

Open access to the Artifact Appendices to the Proceedings of the 31st USENIX Security Symposium is sponsored by USENIX.

A Artifact Appendix

A.1 Abstract

This artifact contains a functional version of DepImpact and necessary data for the evaluation. The execution needs a virtual machine. The host machine may at least have 16GB memory and 64GB hard disk space. To facilitate the usage of this artifact, we prepare a linux virtual machine with necessary component to execute the artifact and visualize the result. Artifact users cam compare the result with our paper draft.

A.2 Artifact check-list (meta-information)

- **Algorithm:** No
- **Program:** Yes
- **Compilation:** No
- **Transformations:** No
- **Binary:** No
- **Model:** No
- **Data set:** Yes, contained in virtual machine
- **Run-time environment:** Ubuntu
- **Hardware:** No
- **Execution:** No
- **Metrics:** Please refer our paper draft
- **Output:** The graph and other necessary data
- **Experiments:** Artifact contains data for experiments
- **How much disk space required (approximately)?:** 30 GB
- **How much time is needed to prepare workflow (approximately)?:** 2 - 3 hours
- **How much time is needed to complete experiments (approximately)?:** 6 - 8 hours
- **Publicly available?:** Yes
- **Code licenses (if publicly available)?:** None
- **Data licenses (if publicly available)?:** None
- **Workflow framework used?:** None
- **Archived (provide DOI)?:** 10.5281/zenodo.5559214

A.3 Description

A.3.1 How to access

https://zenodo.org/record/5559214.YWYJT2LMKUk

A.3.2 Hardware dependencies

To effectively run the artifact, the host machine may at least need 16GB memory and 64GB hard disk spaces.

A.3.3 Software dependencies

No specific software dependencies for this artifact

A.3.4 Data sets

Virtual machine contains evaluation data. The DARPA TC raw data can be downloaded from its website.

A.4 Installation

Download Image file and import by the virtual machine.

A.5 Experiment workflow

In the virtual machine, there is a folder named **DepImpact-artifact** in the **home** directory, which contains two jar packages and a zip file.

- **DepImpact.jar** is used to generate the dependency graph from the log file and to filter out un-relative part for the POI event.
- **CalculateMissing.jar** is used to calculate false positive/negative rate based on the defined critical edges for each attack.
- **allcases.zip** contains logs and property files which are needed for the DepImpact as the input.

A.5.1 Command

java DepImpact.jar pathToRes pathToLog host logname1 logname2 ...

- **pathToRes:** the folder path of the output of DepImpact
- **pathToLog:** the folder path of the input of DepImpact
- **host:** **true** or **false** depends on the case that DepImpact needs to work with

A.6 Concrete Steps

1. Unzip allcase.zip folder
2. Create a folder for the output of DepImact (i.e. pathToRes)
3. Run listed commands:

- java -jar /home/artifact/DepImpact-artifact/DepImpact.jar pathToRes pathToLog false fileName.txt

  If the pathToLog is the path of the unzipped file from the first5cases.zip, the res folder case1 is for the attack Wget executable, the res folder case2 is for the attack Illegal Storage, the case3 is attack Illegal Storage2, the case4 is Hide File, the case5 is Steal information. If the pathToLog is the path of the unzipped file from the case67.zip, the res folder case6 is for the attack Backdoor Download, case7 is for the attack Annoying Server User. Information.
- java -jar /home/artifact/DepImpact-artifact/DepImpact.jar pathToRes pathToLog true logName.txt This command is for attack shellshock, Dataleak, and VPN Filter mentioned in our paper draft.
- java -jar /home/artifact/DepImpact-artifact/DepImpact.jar pathToRes pathToLog false fileName.dot This command is for the attack done by DARPA (Five Dir, Theia, and Trace).

Some cases may require huge memory, it may be suitable to run these cases on a powerful server. For the quick verification and try, we suggest reviewers run DepImpact on some small cases like Five Dir case1 or case3. Reviewers can take Table4 in our submission as a reference for the scale of different cases.

A.7 Evaluation and expected results

- The statistical information of dependency graph like node number and edge number will be in a file whose name ends with “json_log". In this log, it contains the number of node and edge after backtrack POI(Point of Interest), EdgeMerge, and time cost for each component of DepImpact.
- DepImpact will do forward analysis from top-ranked nodes. The filter result is under a folder whose name is DepImpact. The parent folder is set by **pathToRes**.
- The dependency graph is saved as a dot file. To show it, you may use the command:dot -Tsvg dotFilePath > svgFilePath

To calculate the false negative/positive rate for the attack, we need to provide identified critical edges for each attack. The critical edges for the attack used in the evaluation are defined in the corresponding property file.Users need to execute **calculateMissing.jar** to calculate the false positive/negative rate of DepImpact when using different numbers of top-ranked entry nodes. The command should follow this format: java -jar calculateMissing.jar path_of_the_DepImpact_outputs path_of_th*logtopN*

A.7.1 Concrete Examples

Example1:

1. run java -jar /home/artifact/DepImpact-artifact/DepImpact.jar /home/artifact/outputs /home/artifact/DepImpact-artifact/allcases/dataleak1 true dataleakhost1.txt

After this step, there will be a folder dataleak1-case1 created in folder /home/artifact/outputs.

The first thing we should do is to rename the folder "DepImpact" in folder dataleak1-case1 to "sysrep".

For the property file’s name we can ignore the part “-backward.property_", the folder name in the path_to_res(e.g. /home/aftifact/outputs) should be equal to the part befor "-backward.property_" plus "-" plus the part after "-backward.property_".

For this case, because the folder name dataleak1-case1 is not the same as the property file name dataleakhost1-backward.property_case1, we should change the folder name dataleak1-case1 to dataleakhost1-case1.

After this modification, run command:

java -jar /home/artifact/DepImpact-artifact/calculateMissing-1.0-SNAPSHOT-jar-with-dependencies.jar /home/artifact/outputs/dataleakhost1-case1 /home/artifact/DepImpact-artifact/allcases/dataleak1 2

then you will see some output in the terminal, at the same time, there will be a new json file created in the folder “sysrep".

Example2

run command: java -jar /home/artifact/DepImpact-artifact/DepImpact.jar /home/artifact/outputs /home/artifact/DepImpact-artifact/allcases/shellshock1 true shellshockhost1.txt

Modify the folder "DepImpact" in the folder shellshock1-case1 as "sysrep"

According to the property file, we need to modify the folder shellshock1-case1 to shellshockhost1-case1.

run command: java -jar /home/artifact/DepImpact-artifact/calculateMissing-1.0-SNAPSHOT-jar-with-dependencies.jar /home/artifact/outputs/shellshockhost1-case1 /home/artifact/DepImpact-artifact/allcases/shellshock1 2

then you will see some output in the terminal, at the same time, there will be a new json file created in the folder "sysrep".

A.7.2 Results explaintation

In file (eg “case-backward_json_log.json"), the key “BackTrackVertexNumber&EdgeNumbe" shows the number listed as Causality Anylysis #V & #E in Table 4 of our paper draft. The key “CPRVertexNumber& EdgeNumver " shows the number listed as Edge Merge #V & #E in Table 4 of our paper draft.