GATOR Documentation

PRESTO

July 25, 2017

1 Overview

GATOR is a Program Analysis Toolkit For Android. It requires a Unix-like operating system to run, and has been tested on Ubuntu 16.04 and Mac OS X 10.10. The toolkit takes as input either Java bytecode or APK files, which are processed with the Soot program analysis framework (http://www.sable.mcgill.ca/soot/). GATOR's analyses are built on top of Soot.

This release (version 3.3) includes the source code for the static analyses described in our CGO'14[1], ICSE'15[3], ASE'15[4] and CC'16[2] papers:

- GUI structural analysis [CGO'14] with extensions and modifications;
- Callback control flow-analysis [ICSE'15] with minor extensions.
- Analysis for constructing the window transition graph (WTG) [ASE'15] with minor extensions.
- The analysis from [ICSE'15] is included as a building block of WTG construction and is not intended for independent use.
- The analysis from [CC'16] is included as a client based on the WTG.
- The Android programs used in the experiments for these papers are also included.

Compared to the previous release (version 3.2), the current release adds following features:

- Handling of GUI event handlers defined in XML layout files
- Minor enhancements and bug fixes

1.1 Prerequisite Setup

1.1.1 JDK

JDK 1.6+ is required to run Gator. Please refer to http://www.oracle.com/technetwork/java for details of how to obtain a copy of the JDK.

1.1.2 Android SDK

Android SDK is required to run GATOR. It can be downloaded from http://developer.android.com/sdk. After the SDK is installed, support files for individual API levels should be installed as well. For example, if you want to analyze an Android application developed for API level 17, support files for android-4.2.2 must be installed. To do that, you can use the Android SDK manager, whose documentation can be found at http://developer.android.com/tools/help/sdk-manager.html. In order to run Gator on example apps included in this package, at least following API levels and Google APIs should be installed:

- android-8
- android-10
- android-14
- android-15
- android-16
- android-17

Note: Some API levels and their Google APIs are only visible after clicking "obsolete" in Andriod SDK Manager.

2 Usage

2.1 Build Gator

Before running GATOR analysis on applications the GATOR needs to be built. There are two ways to build GATOR: (1) you can import the project rooted at SootAndroid directory into Eclipse or IntelliJ IDE and built it; or (2) you can use the Apache ant (http://ant.apache.org) to compile the project, just run following command under the SootAndroid directory:

\$ ant

2.2 Use the Analyses

Before running any analyses using Gator, there are 2 environment variables need to be defined. The GatorRoot should be assigned the path which contains the AndroidBench and SootAndroid folders. The ADK should be assigned the path to the Android SDK. Under bash, this can be done using following command:

```
export GatorRoot=PATH_TO_ROOT_DIRECTORY_OF_GATOR
export ADK=PATH_TO_ANDROID_SDK
```

2.2.1 Perform Analysis on Demo Applications

If this is the first time you use GATOR, we provide a script AndroidBench/guiAnalysis.sh, which allows you to run the GATOR on applications we used in the [CGO'14], [ICSE'15] and [CC'16] paper.

You should change the current working directory to the \$GatorRoot/AndroidBench using

```
cd $GatorRoot/AndroidBench
```

There are several options to run guiAnalysis.sh script. The easiest one is use:

```
./quiAnalysis.sh runAll
```

It will perform analysis on applications we used in [CGO'14] and [CC'16] with default clients. If you only want to perform analysis on applications in [CGO'14], you can replace the option "runAll" with "runAllCGO". If you only want to perform analysis on applications in [CC'16], you can replace the option with "runAllEnergy". It is also possible to perform analysis on a single application. e.g. To perform analysis on apv, you can use:

```
./guiAnalysis.sh apv
```

The applications available are: apv, astrid, barcodescanner, beem, connectbot, fbreader, k9, keepassdroid, mileage, mytracks, notepad, npr, openmanager, opensudoku, sipdroid, supergenpass, tippytipper, vlc, vudroid, xbmc, osmdroid, osmdroid-fixed, recycle-locator, recycle-locator-fixed, sofia, sofia-fixed, ushahidi, ushahidi-fixed, droidar, droidar-fixed, speedometer, heregps, whereami, locdemo, wigle.

If you need to change the analysis client used during the analysis, you need to modify the configuration information in the \$GatorRoot/AndroidBench/cgo.json and \$GatorRoot/AndroidBench/ccl6.json. The detail information of these two files will be introduced in the following sections.

Please note, due to the oversize issue, only apv and recycle-locator is provided in the gator-3.3.tar.gz package. If you want to analysis all example applications, you need to download the bench-3.3.tar.gz from http://web.cse.ohio-state.edu/presto/software/gator

Note on the CC'16 analysis results for ushahidi and ushahidi-fixed: the detected energy defects reported by GATOR are for an activity named LocationMap. This activity does have an energy defect if it can be opened. However, the activity is not possible to open by the user; therefore, we removed it from the published paper.

2.2.2 Using new Python script runGator.py

Besides the AndroidBench/guiAnalysis.sh, we provide another python script to initiate GATOR to perform analysis on applications. This method consist of two parts: (1) the runGator.py python script; (2) the JSON configuration file.

Let us explain the JSON configuration file first. In this release, we attached 3 JSON files. The cgo.json, which contains the configuration for the applications from the [CGO'14] paper. The cc16.json, which contains the configuration for the applications from the [CC'16] paper. And the example.json-example, which shows a brief example of a JSON configure for the Gator. Here is a snippet from the cc16.json file:

```
"BASE_DIR" : "",
  "BASE_PARAM" : "",
  "BASE_CLIENT" : "EnergyAnalysisClient",
  "BASE_CLIENT_PARAM" : "-clientParam WTPK5",
  "osmdroid" : {
    "relative-path" : "osmdroid/osmdroid-pack/OpenStreetMapViewer",
   "abs-path" : "",
   "api-level" : "android-8",
    "zip-file" : "osmdroid/osmdroid-pack.zip",
   "abs-zip-file" : "",
   "extra-lib" : "",
   "append-param" : "",
   "append-client-param" : "",
   "override-client" : "",
   "override-param" : "",
   "override-client-param" : ""
 }
}
```

The BASE_DIR, BASE_PARAM, BASE_CLIENT and BASE_CLIENT_PARAM are four global keywords. The osmdroid is the name of the application. It can be changed to anything except the keywords defined above. Though this example file shows only 1 application, you can put any number of applications inside a JSON file.

Inside the osmdroid (the application name) variable, there are several fields need to be defined. The relative-path field will be concatenated with BASE_DIR keyword to pinpoint the directory of the application project or the APK file, which is convenient if you have multiple applications in a common location. If you do not want to use the relative path representative, you can leave the relative-path field as blank and put the absolute path of the application source code inside the abs-path field. The api-level field defines the Android API level used by this application, If the application is using the Google API, please change the "android" in android-8 into "google". The zip-file and abs-zip-file are used only for our demo applications, you should leave them as blank when you define your own configuration

files. The extra-lib field defines the location of third party library used by an application. It is rarely used. The append-param and the override-param fields control the option parameters passed into the GATOR. If you have multiple applications that share the same GATOR parameters, it can be defined inside the BASE_PARAM keyword. However if one of the applications need some additional parameters, those parameters can be placed in the append-param. As it will be appended to the BASE_PARAM. If in rare cases, one of the applications in the configure file need to use a set of completely different GATOR option parameters, you can put these options in the override-param. In this case, the options in the BASE_PARAM will be discarded. The BASE_CLIENT keyword and override-client field controls which subclass of GUIAnalysisClient will be used during the analysis. If an application defines its unique client in the override-client field, it will be used during the GATOR analysis, otherwise the client defined in the BASE_CLIENT will be used. append-client-param and the override-client-param controls the client parameters passed to the analysis client class. If you have multiple application that share the same client parameters, it can be defined in the BASE_CLIENT_PARAM. However, if one of the application need additional client parameters, it can be defined in the append-client-param field, as it will be concatenated with BASE_CLIENT_PARAM and passed to client. If the application require a unique client parameter, it can be defined in override-client-param field. In this case, the parameters in the BASE CLIENT PARAM will be discarded.

You can use the runSoot.py script to perform analysis using the configuration in a JSON configuration file. This script is compatible with Python 2.7 and Python 3. The basic options for this script is:

If we just want to run all applications from the [CGO'14] paper, we can simply use:

```
python runGator.py -j cgo.json --base_dir $GatorRoot/AndroidBench
```

If we want to override the client defined in the cgo.json file, e.g. the WTGDemoClient. Instead of putting WTGDemoClient in the BASE_CLIENT in the cgo.json file, we can use following command:

If we only want to run analysis on one or two applications from a JSON configuration file, e.g. apv and connectbot from [CGO'14], we can use following command;

2.3 Perform Analysis on APK files

In this release, we add full support for APK files. If you prefer to use the JSON configuration file explained in the last section, you can simply put the path to the apk file you want to analyze the in the relative-path or abs-path section, it will allow Gator to perform analysis on APK.

We also provide another python script, which is AndroidBench/runGatorOnApk.py, specially designed for apk files. The basic options for this script is:

For example, if we want to perform analysis on an apk located at /tmp/example.apk using WTGDemoClient, we can use following command:

```
python runGatorOnApk.py /tmp/example.apk -client WTGDemoClient
```

If we want to perform analysis on the same apk using EnergyAnalysisClient, we can use following command:

Please note, for some obfustcated apps, the apktool, which we used to extract the apk package, may failed to decode correct tag names in the layout xml files. Unless apktool fix this issue, Gator may crash when performing analysis on these apps.

2.3.1 GATOR Option Parameters

In GATOR, we provide several options to control the way GATOR analyze the application. There are two categories of options, one is PARAM, the other is CLIENT_PARAM. Currently we only provide 1 PARAM option, which is <code>-worker</code> <code>NUM_OF_THREAD</code>. In default configuration, GATOR will analyze the application using 16 threads. However, in rare cases, it may experience concurrency issues as some of part of Soot framework is not thread-safe. In this case, you can put <code>-worker</code> 1 in the PARAM. For example, in the JSON configuration file:

```
{
  "BASE_DIR" : "",
  "BASE_PARAM" : "-worker 1",
  "BASE_CLIENT" : "WTGDemoClient",
```

```
"BASE_CLIENT_PARAM" : "",
  "osmdroid" : {
    "relative-path" : "osmdroid/osmdroid-pack/OpenStreetMapViewer",
    "api-level" : "android-8",
    "zip-file" : "osmdroid/osmdroid-pack.zip"
  }
}
Or:
  "BASE_DIR" : "",
  "BASE_PARAM" : "",
  "BASE_CLIENT" : "WTGDemoClient",
  "BASE_CLIENT_PARAM" : "",
  "osmdroid" : {
    "relative-path" : "osmdroid/osmdroid-pack/OpenStreetMapViewer",
    "api-level" : "android-8",
    "zip-file" : "osmdroid/osmdroid-pack.zip",
    "append-param" : "-worker 1",
  }
}
```

The CLIENT_PARAM, on the other hand, can be used to transfer parameters to the analysis client. For example, in the cc16.json file, we provide an option -clientParam WTPK5 for the EnergyAnalysisClient to define the maximum length of WTG path it should generate. If you want to change this limit to 3, you can replace this option to -clientParam WTPK3. For example:

```
"BASE_DIR": "",
"BASE_PARAM": "",
"BASE_CLIENT": "EnergyAnalysisClient",
"BASE_CLIENT_PARAM": "-clientParam WTPK3",

"osmdroid": {
    "relative-path": "osmdroid/osmdroid-pack/OpenStreetMapViewer",
    "api-level": "android-8",
    "zip-file": "osmdroid/osmdroid-pack.zip"
}
```

If you define your own analysis client, in your source code, you can access all CLIENT_PARAM by accessing following global variable:

3 Customize GUIAnalysisClient

In this section, we will show the way to create a customize GUIAnalysisClient from scratch.

3.1 GUIAnalysisClient

In order to implement a customized GUIAnalysisClient, user needs to add his own class which implements the GUIAnalysisClient interface in presto.android.gui.clients package. The declaration of GUIAnalysisClient interface is:

```
public interface GUIAnalysisClient {
  public void run(GUIAnalysisOutput output);
}
```

When the GUI analysis of Gator is finished, if the user specified the name of user implemented GUIAnalysisClient in the Gator options mentioned in the previous section, the run method in this interface would be called. The parameter output of the run method provides the results from the GUI analysis ([CGO'14]), which can be further used to build the Window Transition Graph.

3.2 Build the Window Transition Graph

The window transition graph (WTG) can be build inside a GUIAnalysisClient. A basic example is like this:

The example code shown above will create a WTG from the result saved in the output parameter. All WTG nodes and WTG edges are stored in the WTG wtg variable. And the number of WTG nodes and edges will be printed on the screen.

3.2.1 WTG related APIs

We provide several APIs to access these nodes. As shown in the example above, API WTG.getEdges() will return all available edges in the WTG and API WTG.getNodes() will return all available nodes in the WTG.

Every application has a launcher node which stands for starting the application from the launcher. This node can be accessed by using:

```
public WTGNode WTG.getLauncherNode();
```

For each WTG node. The window (activity/dialog/menu) it represents can be accessed through:

```
public NObjectNode WTGNode.getWindow();
```

Any inbound WTG edges of a WTG node can be accessed by:

```
public Collection<WTGEdge> WTGNode.getInEdges();
```

Any outbound WTG edges of a WTG node can be accessed by:

```
public Collection<WTGEdge> WTGNode.getOutEdges()
```

For each WTG edge, its source and target window can be accessed through following APIs:

```
public WTGNode WTGEdge.getSourceNode();
public WTGNode WTGEdge.getTargetNode();
```

Each WTG edge is associate with an EventType, for example, it can be clicking on a button, or pressing the **BACK** button. This information can be accessed through:

```
public EventType WTGEdge.getEventType();
```

The GUI event handler triggered in this edge can be accessed through:

```
public Set<SootMethod> WTGEdge.getEventHandlers();
```

In some cases several possible GUI event handlers may be associated with the same event; thus, this method returns a set.

If the edge triggers window life cycle callbacks, these callback methods can be accessed by:

```
public List<EventHandler> WTGEdge.getCallbacks();
```

The sequence of lifecycle callbacks is provided as a list (i.e., the callbacks are ordered). These lifecycle callbacks will occur after the GUI event handlers returned by method getEventHandlers() described earlier. For historic reasons, this methods returns a helper EventHandler object. This EventHandler object is a wrapper for a SootMethod object, which can be accessed via EventHandler.getEventHandler().

Each WTG edge is annotated with window stack operations, which can be **push** a window or **pop** out a window. This information can be accessed by:

```
public List<StackOperation> WTGEdge.getStackOps();
```

And the declaration of the StackOperation class is:

```
public class StackOperation {
  public boolean isPushOp();
  public NObjectNode getWindow();
}
```

The isPushOp() method will return whether current window stack operation is **push**. It will return false if the window stack operation is **pop**. The getWindow() method will return the window this stack operation is pushing or popping.

3.2.2 WTG usage example

Here is a demo of the APIs introduced above:

```
public class WTGDemoClient implements GUIAnalysisClient {
   @Override
   public void run(GUIAnalysisOutput output) {
```

```
VarUtil.v().guiOutput = output;
  WTGBuilder wtgBuilder = new WTGBuilder();
  wtgBuilder.build(output);
  WTGAnalysisOutput wtgAO = new WTGAnalysisOutput(output, wtgBuilder);
  WTG wtg = wtgAO.getWTG();
  Collection<WTGEdge> edges = wtg.getEdges();
  Collection<WTGNode> nodes = wtg.getNodes();
  Logger.verb("DEMO", "Application: " + Configs.benchmarkName);
  Logger.verb("DEMO", "Launcher Node: " + wtq.getLauncherNode());
  for (WTGNode n : nodes) {
    Logger.verb("DEMO", "Current Node: " + n.getWindow().toString());
    Logger.verb("DEMO", "Number of in edges: "
            + Integer.toString(n.getInEdges().size()));
    Logger.verb("DEMO", "Number of out edges: "
            + Integer.toString(n.getOutEdges().size()) + "\n");
  }
  for (WTGEdge e : edges) {
    Logger.verb("DEMO", "Current Edge ID: " + e.hashCode());
    Logger.verb("DEMO", "Source Window: "
            + e.getSourceNode().getWindow().toString());
    Logger.verb("DEMO", "Target Window: "
            + e.getTargetNode().getWindow().toString());
    Logger.verb("DEMO", "EventType: " + e.getEventType().toString());
    Logger.verb("DEMO", "Event Callbacks: ");
    for (SootMethod m : e.getEventHandlers()) {
      Logger.verb("DEMO", "\t"+ m.toString());
    Logger.verb("DEMO", "Lifecycle Callbacks: ");
    for (EventHandler eh : e.getCallbacks()) {
      Logger.verb("DEMO", "\t"+ eh.getEventHandler().toString());
    Logger.verb("DEMO", "Stack Operations: ");
    for (StackOperation s : e.getStackOps()){
      if (s.isPushOp())
        Logger.verb("DEMO", "PUSH " + s.getWindow().toString());
      else
        Logger.verb("DEMO", "POP " + s.getWindow().toString());
  }
}
```

}

This example, will print out details information in the WTG nodes and WTG edges with the APIs introduced in the previous section. There is another example client which is presto.android.gui.clients. ASE15Client in the GATOR's source code. It provides more advanced usage of the WTG.

4 Path Generation

We provide a generic class to perform WTG Path generation. The name of the class is DFSGenericPathGenerator. As the name suggests. It performs depth-first traversal on the Window Transition Graph and it will record the path when it satisfies users' requirements. One of its factory methods of this class is:

There are 2 interface objects required by this class. The first one is the interface IPathFilter, which determines if the path traversed by DFSGenericPathGenerator satisfies the requirement of the user. The declaration of this interface is:

```
public interface IPathFilter {
    /***
    * Specify the stop rule for the DFS traversal
    * @param P
    * @param S
    * @return
    */
    boolean match(List<WTGEdge> P, Stack<NObjectNode> S);

    /***
    * Return the name of the filter
    * @return the name of the filter.
    */
    String getFilterName();
}
```

Every time when the DFSGenericPathGenerator generates a path, it will call the match method in the IPathFilter, if the match method returns true, it means that the path is matched by the pattern defined in this IPathFilter. The matched path will be recorded in matchedPath passed in the factory method.

Another interface, IEdgeFilter is used to determine if a WTG edge should be added to the generated WTG path during the path expansion. The declaration of this interface is:

Every time when the DFSGenericPathGenerator adds a new WTG edge into the current temporary path, it will call the discard method in the IEdgeFilter. If the discard returns true, it means the edge does not satisfy the requirement defined in this IEdgeFilter. The WTG Edge will be discarded.

The List<WTGEdge> initEdges parameter defines the start point of the path generation. Every WTG edge inside this list will be put in the first place in the generated path. This parameter should not be empty.

The boolean parameter stopAtMatch defines the behavior of the DFS traversal when the match method in IEdgeFilter returns true. When this boolean flag is set to true, which is its default value, the DFS traversal will stop at this depth when all IPathFilter have been evaluated. The DFS traversal will return the previous depth. If this boolean flag is set to false, the DFS traversal will continue no matter what is returned by the match method.

The boolean parameter allowRepeatedEdge defines whether repeated edge is allowed in the generated path. If it is set to true, the generated path might contain the same edge for multiple times, which will cause a loop.

The integer parameter K defines the maximum length of the path. In our energy analysis, this value is set to 5.

Here is an example which generates WTGPath from any activity with maximum length of 3:

```
public class PathGenerationDemoClient implements GUIAnalysisClient {
    @Override
    public void run(GUIAnalysisOutput output) {
        //Perform WTG Construction
        WTGBuilder wtgBuilder = new WTGBuilder();
        wtgBuilder.build(output);
        WTGAnalysisOutput wtgAO = new WTGAnalysisOutput(output, wtgBuilder);
```

```
WTG wtg = wtgAO.getWTG();
//Create a placeholder filter class
IPathFilter ph = new IPathFilter() {
  @Override
 public boolean match(List<WTGEdge> P, Stack<NObjectNode> S) {
    return true;
  @Override
 public String getFilterName() {
    return "PlaceHolder";
  }
};
List<IPathFilter> pathFilterList = Lists.newArrayList();
pathFilterList.add(ph);
//Create Initial Edges.
//The path generation will begin from these
//Initial Edges
List<WTGEdge> initEdges = Lists.newArrayList();
for (WTGNode n : wtg.getNodes()){
  if(!(n.getWindow() instanceof NActivityNode)){
    //Ignore any window that is not Activity
   continue;
  }
 List<WTGEdge> validInboundEdges = Lists.newArrayList();
  for (WTGEdge curEdge : n.getInEdges()){
    switch (curEdge.getEventType()) {
      case implicit_back_event:
     case implicit_home_event:
      case implicit_rotate_event:
      case implicit power event:
        continue;
    List<StackOperation> curStack = curEdge.getStackOps();
    if (curStack != null && !curStack.isEmpty()) {
      StackOperation curOp = curStack.get(curStack.size() - 1);
      //If last op of this inbound edge is push
      if (curOp.isPushOp()) {
        NObjectNode pushedWindow = curOp.getWindow();
        WTGNode pushedNode = wtq.getNode(pushedWindow);
        if (pushedNode == n) {
          validInboundEdges.add(curEdge);
        }
      }
```

This example code will use the DFSGenericPathGenerator class to generate any path from any Activity node with its length less or equal to 3. It implements a IPathFilter that will always return true, as the only requirement for the generated path is its length, which is already defined in the parameter K.

The DFSGenericPathGenerator.doPathGeneration() method will start the DFS path generation. After the execution of this method, the recorded path can be accessed from parameter matchedPath passed in the factory method. The key of this map is the filter name defined in IPathFilter.getFilterName() method.

References

- [1] A. Rountev and D. Yan. Static reference analysis for GUI objects in Android software. In *International Symposium on Code Generation and Optimization*, pages 143–153, 2014.
- [2] H. Wu, S. Yang, and A. Rountev. Static detection of energy defect patterns in Android applications. In *International Conference on Compiler Construction*, pages 185–195, 2016.
- [3] S. Yang, D. Yan, H. Wu, Y. Wang, and A. Rountev. Static control-flow analysis of user-driven callbacks in Android applications. In *International Conference on Software Engineering*, pages 89–99, 2015.
- [4] S. Yang, H. Zhang, H. Wu, Y. Wang, D. Yan, and A. Rountev. Static window transition graphs for Android. In *IEEE/ACM International Conference on Automated Software Engineering*, pages 658–668, 2015.