DiSIT, Computer Science Institute
Università del Piemonte Orientale “A. Avogadro”
Viale Teresa Michel 11, 15121 Alessandria
http://www.di.unipmn.it

UNIVERSITÀ DEL PIEMONTE ORIENTALE

The Android Forensics Automator (AnForA): a tool for the Automated Forensic Analysis of Android Applications

C. Anglano, M. Canonico, M. Guazzone (cosimo.anglano@uniupo.it, massimo.canonico@uniupo.it, marco.guazzone@uniupo.it)

TECHNICAL REPORT TR-INF-2019-06-02-UNIPMN
(June 2019)

Research Technical Reports published by DiSIT, Computer Science Institute, Università del Piemonte Orientale are available via WWW at URL `http://www.di.unipmn.it/`.
Plain-text abstracts organized by year are available in the directory

Recent Titles from the TR-INF-UNIPMN Technical Report Series

2019-01 *Deriving Symbolic and Parametric Structural Relations in Symmetric Nets: Focus on Composition Operator*, L. Capra, M. De Pierro, G. Franceschinis, March 2019.

2018-03 *Deriving Symbolic Ordinary Differential Equations from Stochastic Symmetric Nets without Unfolding*, M. Beccuti, L. Capra, M. De Pierro, G. Franceschinis, S. Pernice, July 2018.

2018-02 *Power (set) Description Logic*, L. Giordano, A. Policriti, February 2018.

2018-01 *A Game-Theoretic Approach to Coalition Formation in Fog Provider Federations (Extended Version)*, C. Anglano, M. Canonico, P. Castagno, M. Guazzone, M. Sereno, February 2018.

2017-02 *Configuration and Use of Android Virtual Devices for the Forensic Analysis of Android Applications (see below for citation details)*, C. Anglano, M. Canonico, M. Guazzone, June 2017.

2017-01 *A dynamic simulation model for comparing kidney exchange policies*, M. Beccuti, G. Franceschinis, S. Villa, March 2017.

2016-04 *Tracing sharing in an imperative pure calculus*, P. Giannini, M. Servetto, E. Zucca, December 2016.

2016-03 *SUPPORTING DATA COMMUNICATION AND PATIENT ASSESSMENT DURING EMERGENCY TRANSPORTATION*, M. Canonico, S. Montani, M. Striani, September 2016.

2016-02 *TECHNICAL NOTE TO Forensic Analysis of the ChatSecure Instant Messaging Application on Android Smartphones (see below for citation details)*, C. Anglano, M. Canonico, M. Guazzone, September 2016.

2016-01 *Reasoning in a rational extension of SROEL*, L. Giordano, D. Theseider Dupré, May 2016.

2014-02 *A Provenly Correct Compilation of Functional Languages into Scripting Languages*, P. Giannini, A. Shaqiri, December 2014.

2014-01 *An Intelligent Swarm of Markovian Agents*, A. Bobbio, D. Bruneo, D. Cerotti, M. Gribaudo, M. Scarpa, June 2014.

2013-01 *Minimum pattern length for short spaced seeds based on linear rulers (revised)*, L. Egidi, G. Manzini, July 2013.

2012-04 *An intensional approach for periodic data in relational databases*, A. Bottrighi, A. Sattar, B. Stantic, P. Terenziani, December 2012.

2012-03 *Minimum pattern length for short spaced seeds based on linear rulers*, L. Egidi, G. Manzini, April 2012.

2012-02 *Exploiting VM Migration for the Automated Power and Performance Management of Green Cloud Computing Systems*, C. Anglano, M. Canonico, M. Guazzone, April 2012.

The Android Forensics Automator (AnForA): a tool for the Automated Forensic Analysis of Android Applications

Please, cite this paper as:

Cosimo Anglano, Massimo Canonico, Marco Guazzone

"The Android Forensics Automator (AnForA): a tool for the Automated Forensic Analysis of Android Applications,"

Submitted for publication.

The Android Forensics Automator (AnForA): a tool for the Automated Forensic Analysis of Android Applications*

Cosimo Anglano[1], Massimo Canonico[1,*], Marco Guazzone[1]

[a] *Computer Science Institute, DiSIT, University of Piemonte Orientale, Viale T. Michel 11, 15121 Alessandria, Italy*

Abstract

Most of our daily activities are carried out by means of mobile applications, that typically generate and store on the device large sets of data. The forensic analysis of these data thus plays a crucial role during an investigation, as it allows to reconstruct the above activities. Manually analyzing these applications is a long, tedious, and error-prone task.

In this paper we present the design, implementation, and evaluation of AnForA, a software tool that automates most of the activities that need to be carried out to forensically analyze Android applications, and that has been designed in such a way to yield various important properties, namely fidelity, completeness, soundness, effectiveness, repeatability, and generality.

AnForA is based on a dynamic "black box" approach, in which the application to be analyzed is first installed on a virtualized Android device, and then a set of experiments are carried out, in which actions of interest are automatically performed on the application by emulating a human user that interacts with its interface. During the experiments, the filesystems of the device storage are actively monitored, so that the data created or modified by each one of these actions can be located and correlated with that action.

We have devised a proof-of-concept implementation of AnForA, that we use to assess its ability in achieving its design goals, by analyzing through it several Android applications already studied in the literature, so that we can compare AnForA's results against those reported in these papers. The results of our evaluation confirm that AnForA greatly simplifies the forensic analysis of Android applications, and exhibits all the properties mentioned above.

Keywords: Digital forensics, mobile forensics, Android applications, digital evidence, automated forensics analysis.

---

1. Introduction

Mobile devices are an integral part of our everyday lives. More often than not, they play a key role in all our activities, and not only in our interpersonal communications. Most of our daily activities (e.g., eating, sleeping, doing sport, driving, interacting with other people, etc.) are indeed carried out, at least in part, by using suitable apps installed on our mobile devices. These apps generate and store on the device large sets of data, that may be later used to reconstruct the activities carried out by the user on the device. Hence, the forensic analysis of these applications may (and usually does) play a crucial role during an investigation.

To reconstruct user activities starting from the data generated by a given application, the analyst needs to know (a) which data are generated by the application, (b) how these data are encoded, (c) where these data are stored on the device, and (d) the data generated or modified by each operation allowed by the application. In this way, it is indeed possible to establish the causal correlation between the data generated by the application and the user action that led the application generate it. Once these correlations have been established in the general case, it is possible to infer, from the presence of certain data on a given device, whether a given action may have been performed or not on that device.

Unfortunately, gaining the above knowledge is usually a rather complex affair. It indeed involves to perform a set of controlled experiments, in which (a) the analyst carries out each action of investigative interest (e.g., sending a text message), (b) the internal and external storage of the devices are inspected to determine which data are generated by these actions and where they are stored, and (c) these data are analyzed in order to decode them and to assign them their correct meaning [3]. Doing so in a systematic manner, where all the relevant actions allowed by the applications are considered, is a long, tedious, and error-prone task.

Moreover, the huge and always growing number of apps available to users, as well as the frequent updates of existing ones, places a hard-to-sustain burden on the analyst, given that a new application

⋆This research has received the financial support of the Università del Piemonte Orientale.

∗Corresponding author

Email addresses: cosimo.anglano@uniupo.it (Cosimo Anglano), massimo.canonico@uniupo.it (Massimo Canonico), marco.guazzone@uniupo.it (Marco Guazzone)

(or a new version of an existing application) must be analyzed before the reconstruction of user activity may take place. For these reasons, the idea of automating the forensic analysis of mobile applications has recently received the attention of the scientific community [5, 13, 22].

In order to be adequate, an automated solution for the forensic analysis of mobile applications should provide:

1. *fidelity*, i.e. the ability to reproduce, as faithfully as possible, the interactions that a human experimenter would have with the application under analysis in order to perform the actions of investigative interest;
2. *completeness*, i.e. the ability to identify all the data that are generated/modified by the application as effect of the above actions;
3. *soundness*, i.e. the ability to exclude data generated by other applications/services running on the same device used to carry out the experiments;
4. *effectiveness*, i.e. the ability to correlate each user action of interest with the data it modified/generated;
5. *repeatability*, i.e. the ability to provide to a third party the possibility of replicating the same set of experiments and to obtain the same results;
6. *generality*, i.e. the ability of generating results that hold for as many different devices as possible (possibly all).

Existing proposals for the automation of the forensic analysis of mobile applications [5, 13, 22] exhibit only a subset of the above features (see the related work in Sec. 2), hence they do not represent a completely satisfactory solution.

In this paper we fill this gap by proposing AnForA (acronym for *Android Forensics Automator*), a system that automates the forensic analysis of Android applications. AnForA is based on a novel analysis methodology (that extends the approach proposed in [3]), that ensures the achievement of repeatability, and generality. This methodology provides the basis for the design of a software architecture whose components interact among them to achieve fidelity, completeness, soundness, and effectiveness. Furthermore, these components fully automate the execution of the experiments required to characterize the behavior of mobile applications, the collection of the results generated in them, and the correlation of each action performed in these experiments with the data they generate.

In particular, AnForA is based on a dynamic "black box" approach, in which the application to be analyzed is first installed on a virtualized Android device, and then a set of *experiments* are carried out, in which actions of interest are automatically performed on the application by emulating a human user that interacts with its interface. During the experiments, the filesystems of the device storage are actively monitored, so that the data created or modified by each one of these actions can be located and correlated with that action.

We have devised a proof-of-concept implementation of AnForA that couples off-the-shelf software components already available in the Android ecosystem with components purposely developed by us. We experimentally evaluate the ability of AnForA in achieving the goals mentioned above, by using the above implementation to carry out the forensic analysis of several Android applications already studied in the literature, so that we can compare AnForA's results with those reported in these papers. The results of our evaluation confirm that AnForA exibits all the properties mentioned above, namely fidelity, completeness, soundness, effectiveness, repeatability, and generality.

The rest of this paper is organized as follows. In Sec. 2 we discuss related works. Then, in Sec. 3 we present the design and the implementation of AnForA, while in Sec. 4 we illustrate its practical use by using as example the Gmail app. Next, in Sec. 5 we report the results of the experimental validation of AnForA, and in Sec. 6 we conclude the paper and outline future research work.

2. Related work

The forensic analysis of mobile applications has received a considerable attention in the recent literature [1, 2, 3, 10, 11, 14, 15, 19, 20, 21], where a large set of different mobile applications have been analyzed in order to identify and decode the artifacts they store on the device where they run. In all these works, the analysis (i.e., the identification of the artifacts, their location, and their decoding) has been carried out manually. However, the high complexity of the applications makes the manual approach cumbersome, time consuming, and prone to errors. Hence, the need for an automated solution clearly emerges from these works.

The problem of automating the forensic analysis of mobile applications has recently received a significant interest in the literature, where various proposal – focusing on the automatic identification and decoding of the data that are generated by Android applications during their execution – have been published [5, 13, 22].

In particular, ForDroid [13] and EviHunter [5] combine static code analysis, to discover all the possible execution paths in the application code, with taint analysis, to track the flow of relevant information from their *sources* (i.e., the places in the code where these data are generated) to their *sinks* (i.e., the places in the code where these data are written to the file system). However, both approaches suffer from the following drawbacks, due to the fact that they rely on static code analysis:

1. Sources and sinks, that correspond to specific methods of the Android API (e.g., those that are used to obtain GPS coordinates) or to specific variable types (e.g., strings), must be known in advance. However, gaining such knowledge is not trivial, as it requires a complex analysis of the Android API, that must be repeated each time the above API changes [5].
2. They lack soundness and effectiveness: static analysis, on which they rely, merely enumerates all the possible execution paths within the code of the application, instead of considering only those corresponding to the actions chosen by the analyst. Hence, they identify also data which do not correspond to the actions chosen by the analyst (lack of soundness), and are unable to correlate each user action with the data it generates (lack of effectiveness).
3. They lack completeness, since are unable to deal with (a) code paths that cannot be determined statically (e.g., those that contain multithreading or reflection [5]), (b) applications that are highly obfuscated [16], and (c) with service requests, issued by the app, that are served by other apps or system services available on the device.

To overcome the limitations of static analysis, [22] proposes an approach based on dynamic taint analysis, which is based on the execution of the application of interest on a modified version of the ART Android runtime.

This approach, however, suffers from drawbacks similar to those characterizing static analysis. In particular, it lacks effectiveness, and requires the knowledge of the sources and sinks in the applications. Furthermore, it lacks completeness, since it is unable to deal with applications that use native code, or that exhibit implicit data flows [22].

In contrast, AnForA does not present anyone of the problems affecting the alternative solutions mentioned before. As a matter of fact, it is explicitly driven by the user actions that have been chosen by the analyst, and therefore (a) identifies only those data that are created or modified as direct consequence of these actions, disregarding instead data generated by other actions (soundness), and (b) is able to correlate each of the chosen actions with the data it generates (effectiveness). Finally, AnForA is based on a dynamic "black box" approach and not source code analysis, and therefore is not hindered by the presence, within the application, of code obfuscation techniques or of dynamic code update mechanisms (completeness).

3. The AnForA System

As anticipated in the Introduction (Sec. 1), AnForA is based on a methodology for the forensic analysis of mobile applications that has been specifically conceived in order to provide repeatability and generality. This methodology provides the basis for its architecture, that encompasses various components that interact among them to provide fidelity, completeness, soundness, and effectiveness through the full automation of the execution of analysis experiments, the identification of the location and format of all (and only) the data generated during these experiments, and their correlation with the actions that generated them.

In this section we first describe the analysis methodology (Sec. 3.1), then we describe the architecture of AnForA (Sec. 3.2), and its proof-of-concept implementation (Sec. 3.3).

3.1. The Analysis Methodology

The methodology providing the foundations of AnForA is based on the design of a set of experiments, each one focusing on one of the operations allowed by the application under analysis (e.g., sending a text message or a picture), on their systematic execution using the application on a mobile device, on the inspection of the device storage during and after each experiment (so as to identify the data generated during it), and on the analysis of the generated data to determine their meaning and context.

Generality and repeatability, that are the main goals of this methodology, are achieved through the use of virtualized mobile devices in place of physical ones. A mobile virtualization platform makes indeed simple and cost-effective to run experiments on a multitude of different virtual smartphones (featuring different hardware and software configurations), thus yielding generality. Furthermore, it allows a third-party to easily replicate experiments on the same smartphone models and configurations, as well as to enforce the same operational conditions holding during the experiments, thus yielding repeatability.

The methodology, whose workflow is schematically depicted in Fig. 1, is articulated in a sequence of steps, as discussed below.

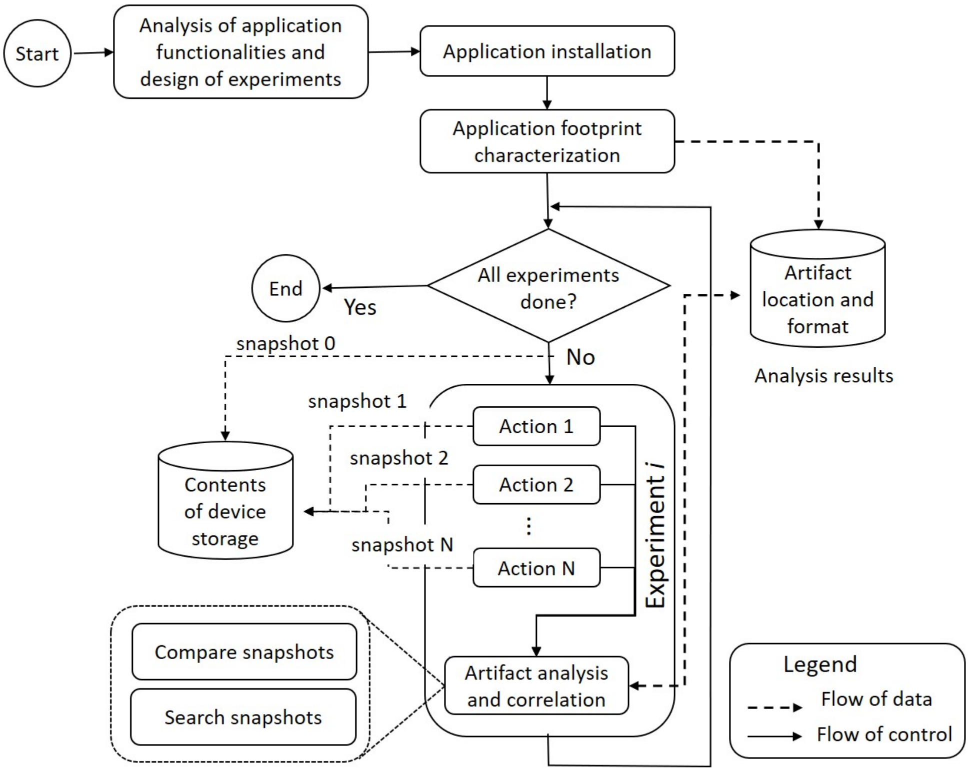

Figure 1: Workflow of the analysis methodology for mobile applications.

First, the functionalities of the application under consideration are analyzed, in order to identify its actions that have a potential investigative interest, and then suitable experiments, aiming at eliciting the generation of the data corresponding to the above actions, as well as their storage on the local memory of the device, are designed.

Next, in the *Application installation* step the application is installed on the device. Then, in the *Application footprint characterization* step, the set of directories that contain data generated either directly or indirectly by the application (the *Analysis Paths*) are identified, so that they can be monitored during the execution of the experiments in order to detect changes to the data they store. Furthermore, the data generated during the installation is also collected and examined, and the results of this analysis are stored into the *Artifact location and format* database.

After these preparatory steps, the set of experiments is carried out in a systematic way, until all of them have been completed. As shown in Fig. 1, each experiment consists in a set of actions, carried out by the analyst in a predefined order by interacting with the application user interface. At the beginning of the experiment, and also after each action of the experiment is completed, a snapshot of the contents of the *Analysis Paths* may be collected and stored for subsequent analysis.

After all the actions of a given experiment have been completed, the various snapshots are compared in order to identify which files have been created, deleted and/or updated as effect of each action. Furthermore, snapshots may be searched for known information (e.g., the text of a message that has been sent) to determine the data that have been written in, or deleted from, the above files. These findings, jointly with the association of each artifact with the (set of) action(s) that generated them, are recorded into the *Artifact location and format* database.

3.2. The Architecture

To carry out the analysis according to the above methodology, AnForA couples a mobile virtualization platform, enabling the configuration of, use of, and interaction with a virtual device where the app under analysis is installed and executed, with an *analysis machine*, where its software components run and where the data extracted from the mobile device are stored and analyzed.

The architecture of AnForA is shown in Fig. 2, while in Fig. 3 we show how its components automate the various steps of the analysis methodology, by annotating each one of them with the corresponding component automating it.

As shown in Fig. 2, AnForA consists of seven components, whose role is detailed below:

1. The App Installer, which installs the application to be analyzed on the virtual mobile device, thus automating the *Application installation* step. In particular, the App Installer takes as

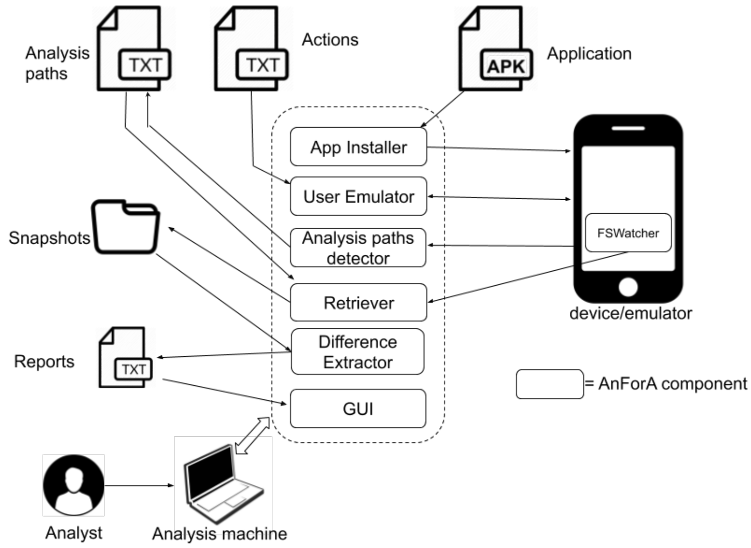

Figure 2: AnForA architecture.

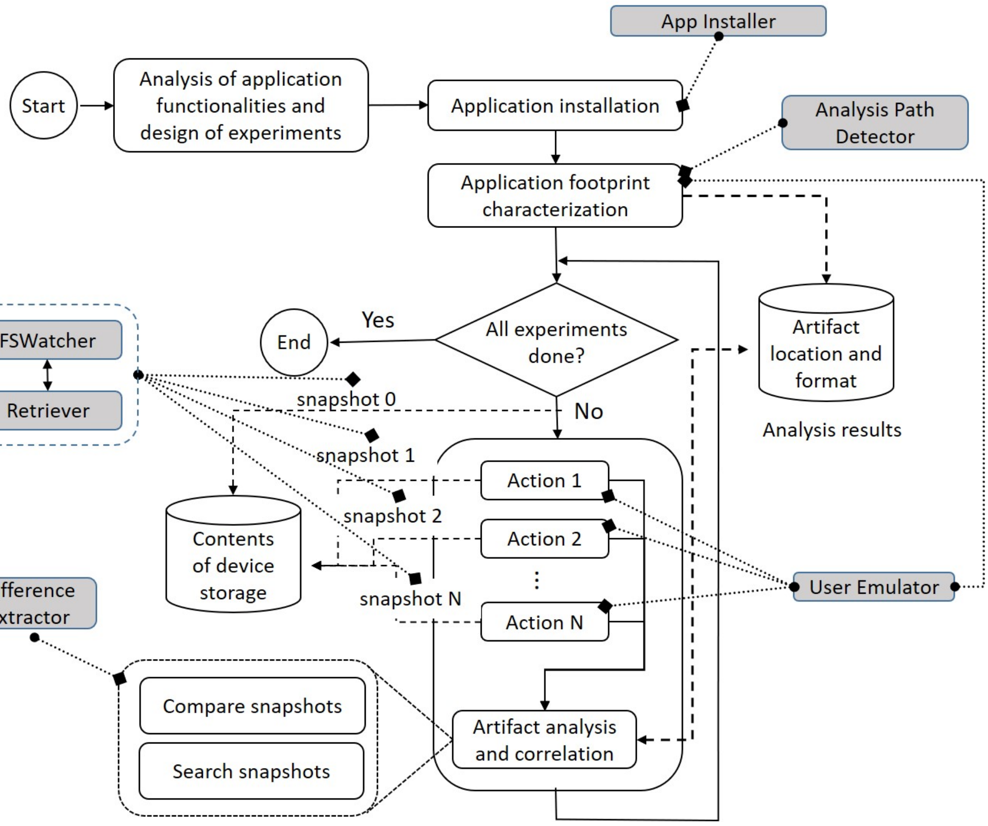

Figure 3: The workflow of AnForA. Annotations are shown as gray boxes attached via a dotted connector to the methodology steps it is used in.

input the installation file of the application (the *APK file*), extracts from it the *Manifest* file that lists the permissions the application needs to function properly, and executes suitable commands to install the application, as well as to grant the permissions it needs.

2. The User Emulator, which interacts with the application to perform sequence of actions that make up an experiment by using the application *Graphical User Interface* (or *GUI*, for brevity) exactly as a human user would do. These actions are specified by the analyst into an *action file*, using a suitable command language (see Sec. 3.3) As shown in Fig. 3, the User Emulator automates the execution of the experiments.

3. The Analysis Paths Detector, which automates the *Application Footprint Characterization*. In particular, it determines the folders, in the internal and external storage of the device, where the application can write data (the *Analysis Paths*). More specifically, the Analysis Path Detector identifies, and includes in the *Analysis Paths*, the following folders:

   - the *private data folder* of the application, i.e. the private directory where the application may write data;
   - the *additional folders* of the applications, i.e. public directories where the application may write if it is granted the corresponding write permissions;
   - the *EC private and additional folders*, i.e. those folders where third-party applications or system services (henceforth referred to as *External Component* or *EC* for brevity), write data on the behalf of the app when it requests service to them.

4. The FSWatcher, which runs on the mobile device, and monitors all the folders in the *Analysis Paths* to detect which files are created/deleted/modified during the experiments. The FSWatcher works in close interaction with the Retriever, from which it receives the start/stop monitoring commands (see below), and to whom it reports the results of the monitoring sessions.

5. the Retriever, which retrieves the set of files reported by the FSWatcher, and creates a snapshot that is stored on the analysis machine. In particular, as shown in Fig. 3, at any time during the experiment, the Retriever may get a snapshot of the contents of each directory in the *Analysis Paths* if specified in the action file, and tells the FSWatcher to start its monitoring

activity. Then, monitoring is stopped at the end of each action of a given experiments, when the Retriever gets a new snapshot.

6. The Difference Extractor, which compares two versions of the same file contained in different snapshots, in order to find which data have been added to/removed from/modified in that file, and associates these data with the corresponding action. In particular, as shown in Fig. 3, the Difference Extractor comes into play at the end of each experiment to examine all the pairs of consecutive snapshots to identify all the files that have been created/modified/deleted as effect of the action corresponding to the second snapshot of each pair. For each one of these files, the Difference Extractor locates the portions that have been modified, decodes them (if it knows their encoding scheme), and writes this information into a report. At the end of the experimental campaign, the report contains – for each individual experiment – the list of the files that have been added/created/modified/deleted by each one of its actions. In this way, the analyst can establish the correspondence between each user action, and the artifacts it generates. Furthermore, the report contains also – for each one of the above files – which data have been modified, and where this data are located within it, thus enabling the analyst to quickly focus the analysis of them.
7. The GUI, which provides the analyst with the access to the various functionalities of AnForA, and allows him/her to control the various steps of the analysis workflow by controlling the operations of its components.

Jointly, AnForA components provide fidelity, completeness, soundness, and effectiveness, as discussed below:

- *fidelity* is ensured by the App Installer, which reproduces all the steps that a human user would have to do to install the app (including the grant of the permissions it requires), and by the User Emulator, which interacts with the application via its user interface exactly as a human user would do (i.e., by issuing commands like swipe or tap on the screen, as well as by typing on the keyboard);

- *completeness* is ensured by the Analisys Paths Detector, which is able to identify all the device folders belonging to the *Analysis Paths* of the application, and by the FSWatcher, which identifies all the files, stored in these paths, that are modified by the application during the experiments;

- *soundness* is achieved by means of the FSWatcher, which identifies only the files modified by the application;
- *effectiveness* is provided by the Difference Extractor, which is able to (a) precisely locate, in the file storing them, the data generated by each action, and (b) correlate the above action with the corresponding data.

It is worth noting that AnForA supports the execution of experiments where multiple users interact among them using the application under analysis, or even different applications. To do so, it is sufficient to start as many instances of AnForA are necessary, either on the same or on different analysis machines, and provide to each instance its specific actions file.

3.3. The Implementation

We have developed a proof-of-concept implementation of AnForA, which relies on a mix of freely available tools and of software components that we developed specifically for it using the `Python` and the `C++` languages, as discussed below.

First of all, for the configuration and use of virtualized mobile devices, AnForA relies on the *Android Mobile Device Emulator* [9], a software tools running on the analysis machine that allows the creation and the execution of the so-called *Android Virtual Devices* (*AVD*s), i.e. emulated smartphones behaving exactly like real physical devices that can be customized with different hardware characteristics and Android versions.

The components of AnForA running on the analysis machine interact with those running within the virtual device, as well as with its Android operating system, by means of the Android Debug Bridge (ADB) service [6], that is part of the standard Android SDK [7].

In the following, we briefly discuss how each one of the components of AnForA has been implemented.

- The App Installer is implemented as a Python script that uses suitable ADB commands to install the app using its APK file, and to grant it the permissions it needs to function properly.

These permissions are automatically extracted from the AndroidManifest.xml file, which is contained in the APK file of the application. The App Installer also installs the FSWatcher on the AVD and suitably sets port forwarding on it to enable the FSWatcher – User Emulator network communications.

- The User Emulator is implemented in Python by means of the *UI Automator* testing framework [8], that provides a set of APIs to build UI tests that perform interactions on user apps and system apps. This APIs allows to programmatically interact with the various components of the application GUI by emulating via software typical user gestures such as tap, swipe, long tap and so on. In particular, the User Emulator reads from the actions file the sequence of actions it has to perform on the application user interface, and for each one of them calls the appropriate function of the UI Automator API. The User Emulator currently supports the actions listed in Table 1 below. To find out the identifier of a widget, or its position on the

Action name	**Description**
`Home`	Go to the device home screen
`Back`	Tap on the back button
`TapOn(widgetId)`	Tap on the widget identified by `widgetId`
`TapXY(x,y)`	Tap on point ⟨`x,y`⟩ of the screen
`SetTxt(widgetID,t)`	Insert text `t` into the widget identified by `widgetId`
`SetTxtXY(x,y,t)`	Insert text `t` into the widget placed on position ⟨`x,y`⟩ of the screen
`Dump`	stops the FSWatcher, extract from the device the list of files it reports, and then restarts it.

Table 1: Actions supported by the User Emulator

screen, the analyst uses the *UI Automator Viewer*, a tool which is part of the UI Automator framework (see Sec. 4 for the discussion on how to use it).

- the Analysis Path Detector is implemented as a Python script that performs the following two actions:

    1. it parses the APK file of the application to extract the information concerning the corresponding private and additional folders (from the *AndroidManifest.xml* file);
    2. it discovers the EC private and additional folders by performing a "dry run" (i.e., a run

in which no data are collected from the device memory) in which the *User Emulator* carries out all the experiments defined by the analyst. As a matter of fact, ECs can be determined only at run-time, since they depend on specific choices made by the user when the application runs. By performing a dry run, the application is forced to issue all the service requests to the ECs it uses. The information about these requests (called *intents* in the Android jargon) are recorded by Android into a log file that, after all the experiments have been performed, is extracted and analyzed, so that the Analysis Path Detector can identify the above ECs. At the end of the dry run, the virtualized device is brought back to a clean state, so that all the modifications it induces are wiped away.

- The FSWatcher is implemented as a `C++` program, and relies on the Linux's *inotify* mechanism for watching filesystem events under specific root paths. For each given root path to watch, the FSWatcher recursively monitors its entire subtree and can follow all symbolic links found therein. On stopping, the FSWatcher produces a report containing all the changes that took place in the monitored paths since when started. The FSWatcher is installed by he App Installer and it communicates with the User Emulator through a socket interface by means of which they exchange JSON messages.

- The Retriever is implemented as a Python script that use ADB commands to retrieve files from the Android device to the analysis machine, and it is invoked any time a `Dump` action is specified in the *action file*.

- The Difference Extractor is implemented as a Python script that uses the magic Linux utility to determine the type of the file it needs to process, and calls the appropriate diffing utility. In the current implementation, it is able to compute the differences between text and binary files (using the diff Linux utility), and SQLite databases (using the SQLDiff utility).

- The GUI is implemented in Python and uses the *PyQt* Python bindings for the Qt cross-platform framework.

4. GMail application use case

To illustrate how to use AnForA, in this section we discuss how to set up and run an experiment in which the *Gmail* email app is used to compose and send a message to a specific destination address.

First of all, we need to specify the sequence of actions, that need to be carried out to compose and send the message, using the User Emulator language. More specifically, this sequence consists in the following actions:

1. bring the device to the main screen by tapping on the "home" button of its interface, so that the icons of all the installed apps are shown;
2. start the Gmail app by tapping on the corresponding icon;
3. open the "compose" window by tapping on the `Compose` button of the application GUI;
4. fill the `to`, `subject`, and `message body` text boxes shown on the application GUI;
5. send the message by tapping on the `send` button of the application GUI.

As discussed in Sec. 3.3, for each one of these activities, it is necessary to determine either the identifier of the corresponding widget, or its position on the screen. The identifier of a widget is stored into either one of two *properties* of the widget (i.e., attributes that can store values), namely either the content-desc or the resource-id property, while its position on the screen is stored in the bounds property.

As discussed in Sec. 3.3, both information can be retrieved by loading in the *UI Automator Viewer* the window of the GUI where the widget of interest is placed, and by using it to inspect its properties.

Fig. 4 shows how to identify the icon corresponding to the Gmail app using the *UI Automator Viewer*, which is shown in the left side pane. When the analyst selects this icon, its properties – and in particular `content-desc` – are shown in the bottom-right pane. From this property it can be seen that the identifier of the Gmail icon is "`Gmail`", so the User Emulator action that launches the Gmail app is set to `TapOn(Gmail)`.

In the same way, the "compose" window is opened by tapping on the `Compose` button of the application GUI, whose identifier is stored – as shown in Fig. 5 – in the resource-id property (note

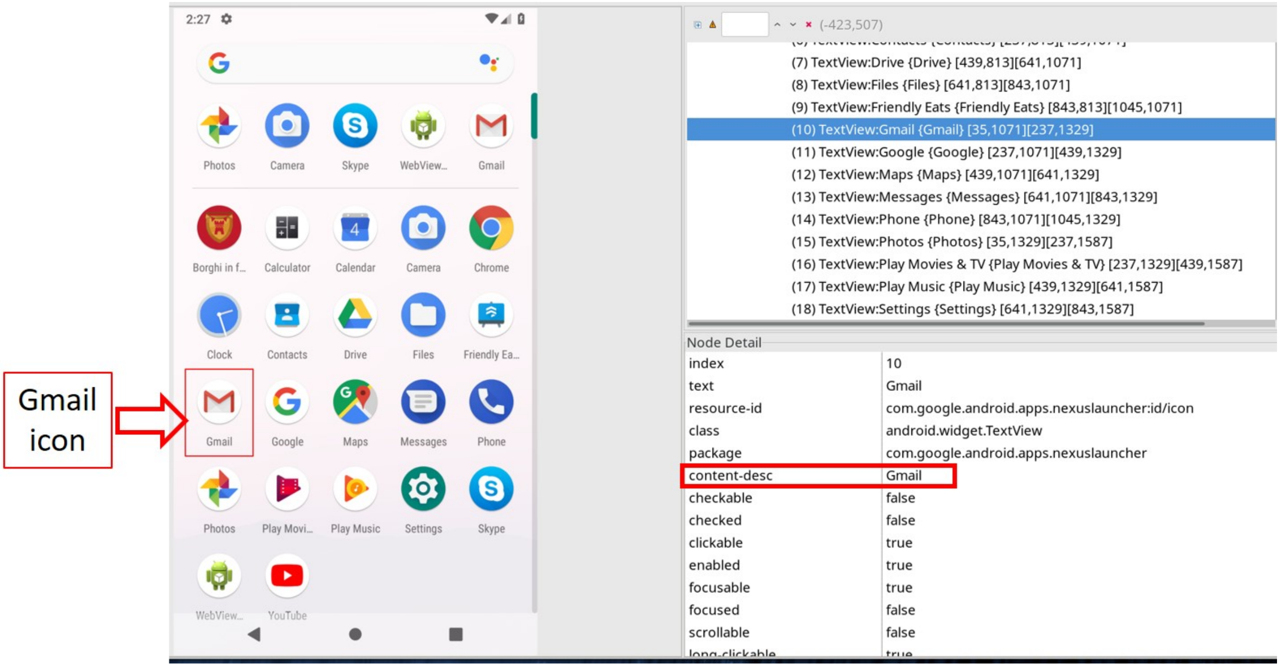

Figure 4: Identification of the Gmail icon.

that the content-desc property could have been used equivalently). Hence, the User Emulator action that opens the "compose" window is `TapOn(com.google.android.gm:id/compose_button)`.

To compose and send a message, the User Emulator needs to place suitable textual information into the destination address, the subject, and the email body text boxes; hence, the corresponding widgets need to be identified. For the first two text boxes, we proceed exactly as described above, i.e. we use either the content-desc or the resource-id properties, so we do not discuss it here again. However, as shown in Fig. 6, the position on the screen of the email body text box needs to be used, since the above two properties are empty. The screen position of this widget, which is stored in the bounds property, corresponds to the rectangle whose opposite edges correspond to points $(42, 654)$ and $(1039, 794)$. To identify the widget, any point falling inside this rectangle – e.g., $(50, 700)$ – may be used. Hence, the User Emulator action that fills the email body text box is `SetTxtXY(50,700,"Test Message #1 - body")`.

The resulting action file, that will be provided as input to the User Emulator, contains the following actions:

1. `TapOn(Home)`

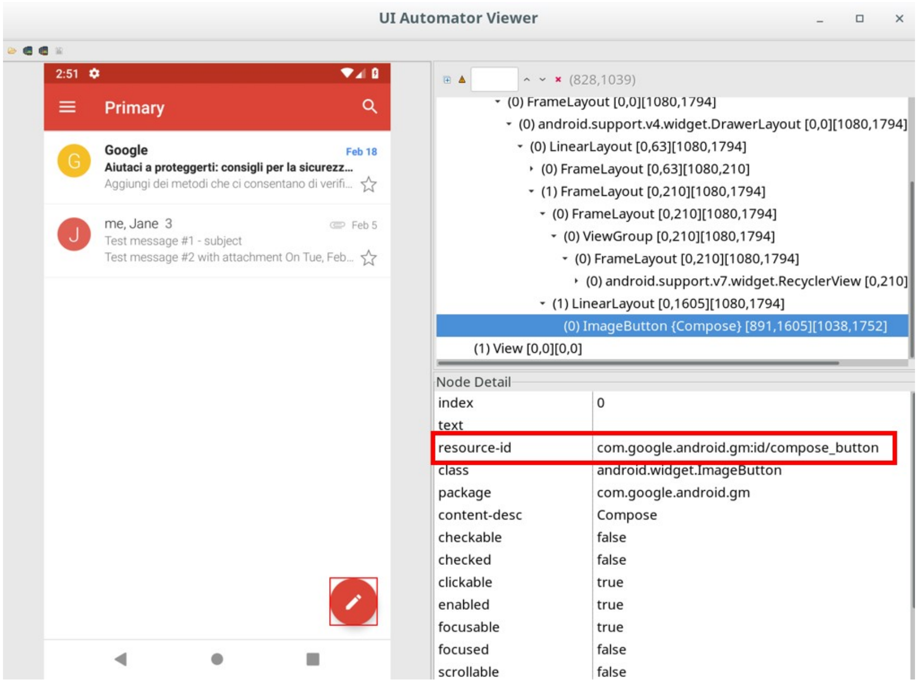

Figure 5: Identification of the Gmail Compose button.

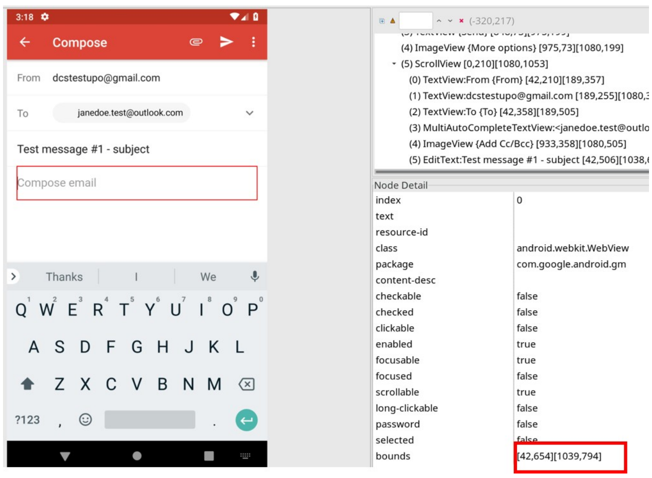

Figure 6: Identification of the email body text box.

2. `TapOn(Gmail)`
3. `TapOn(com.google.android.gm:id/compose_button)`
4. `SetTXT(com.google.android.gm:gm-to,"janedoetest@outlook.com")`
5. `SetTxt(com.google.android.gm:gm-subject,"Test message # 1 - subject")`
6. `SetTxtXY(50,700,"Test Message #1 - body")`
7. `TapOn(com.google.android.gm:id/send)`

Now, the experiment may be carried out by means of the AnForA's GUI, which is shown in Fig. 7, and provides various buttons, each one corresponding to a specific step of the analysis methodology.

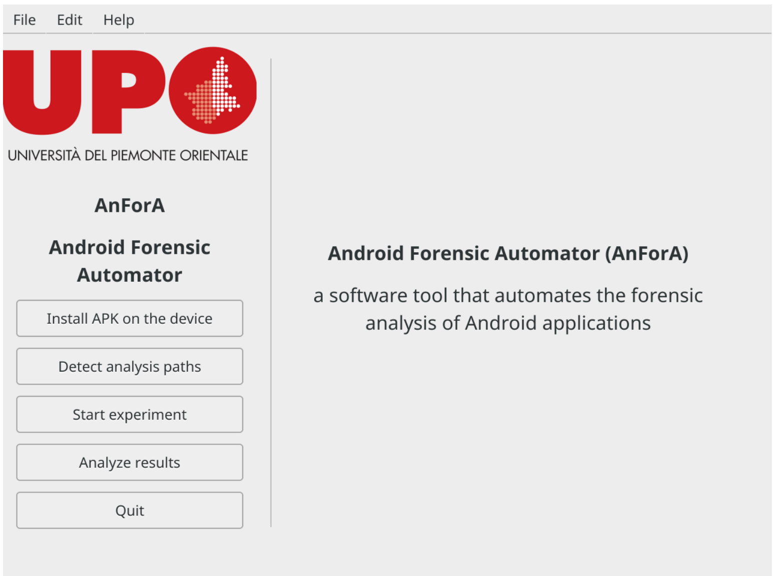

Figure 7: The AnForA main screen.

In particular, the app is installed via the *Install APK on the device* button, which prompts the analyst to specify the path name of the APK, and to click on the *Install* button (see Fig. 8). At the end of the installation, AnForA notifies the user with the outcome of the operation.

The other actions, namely the detection of the analysis paths, and the execution of the experiment, are carried out by means of the *Detect analysis paths* and *Start experiment*. When the experiment is done, AnForA notifies the analyst (see Fig. 9), and enables him/her to download the results and to analyze them using the *Download results* button.

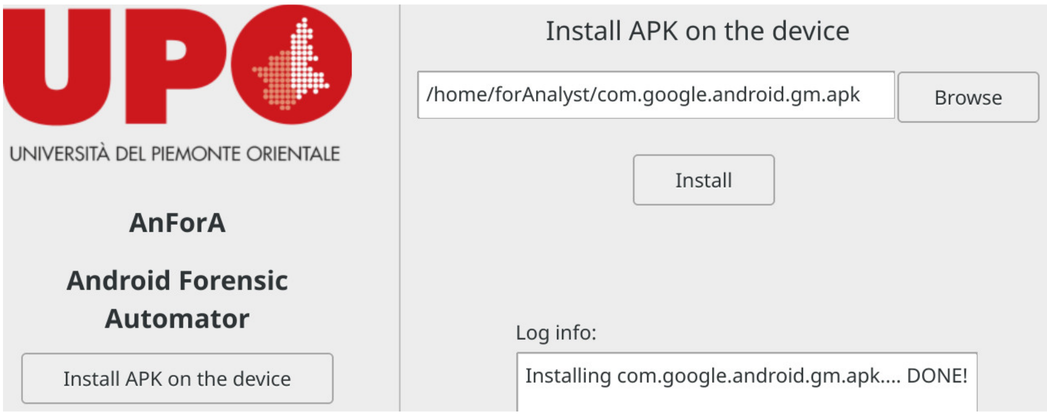

Figure 8: The APK installation menu. The figure has been cut to show only the information of interest.

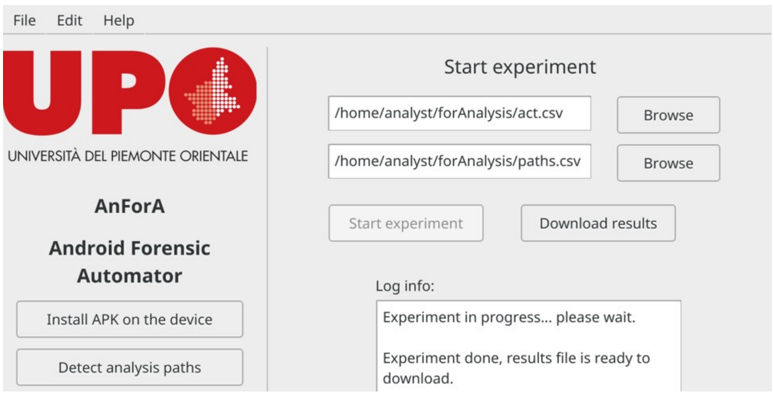

Figure 9: The experiment has been completed. The figure has been cut to show only the information of interest.

At this point, once the analyst clicks on the *Analyze results* button, (s)he is presented with the set of snapshots generated as consequence of the various Dump actions in the *action file*. By selecting two different snapshots, the analyst is presented with the list of files that have been created, modified or deleted in the second snapshot with respect to the first one (Fig. 10).

Experiment

File

Row ID	Status	Snapshot ID	Pathname	File	Type	Action
0	Created	snapshot_2	/data/data/com.google.android.gm/app_webview:	Cookies	application/x-sqlite3	Show
1	Created	snapshot_2	/data/data/com.google.android.gm/app_webview:	Cookies-journal	inode/x-empty	Show
2	Created	snapshot_2	/data/data/com.google.android.gm/app_webview:	GPUCache	inode/directory	Show
3	Created	snapshot_2	/data/data/com.google.android.gm/cache:	http	inode/directory	Show
4	Created	snapshot_2	/data/data/com.google.android.gm/cache:	org.chromium.android_webview	inode/directory	Show
5	Created	snapshot_2	/data/data/com.google.android.gm/cache:	uploader	inode/directory	Show
6	Modified	snapshot_2	/data/data/com.google.android.gm/databases/	bigTopDataDB.993253759	application/x-sqlite3	Show
7	Deleted	snapshot_1	/data/data/com.google.android.gm/databases:	bigTopDataDB.993253759-shm	application/octet-stream	Show
8	Deleted	snapshot_1	/data/data/com.google.android.gm/databases:	bigTopDataDB.993253759-wal	application/octet-stream	Show
9	Created	snapshot_2	/data/data/com.google.android.gm/databases:	downloader.db	application/x-sqlite3	Show
10	Created	snapshot_2	/data/data/com.google.android.gm/databases:	downloader.db-shm	application/octet-stream	Show
11	Created	snapshot_2	/data/data/com.google.android.gm/databases:	downloader.db-wal	application/octet-stream	Show
12	Modified	snapshot_2	/data/data/com.google.android.gm/databases/	EmailProviderBody.db	application/x-sqlite3	Show
13	Deleted	snapshot_1	/data/data/com.google.android.gm/databases:	EmailProviderBody.db-shm	application/octet-stream	Show
14	Deleted	snapshot_1	/data/data/com.google.android.gm/databases:	EmailProviderBody.db-wal	application/octet-stream	Show
15	Modified	snapshot_2	/data/data/com.google.android.gm/databases/	internal.dcstestupo@gmail.com.db	application/x-sqlite3	Show
16	Deleted	snapshot_1	/data/data/com.google.android.gm/databases:	internal.dcstestupo@gmail.com.db-shm	application/octet-stream	Show
17	Deleted	snapshot_1	/data/data/com.google.android.gm/databases:	internal.dcstestupo@gmail.com.db-wal	application/octet-stream	Show
18	Modified	snapshot_2	/data/data/com.google.android.gm/databases/	mailstore.dcstestupo@gmail.com.db-shm	application/octet-stream	Show
19	Created	snapshot_2	/data/data/com.google.android.gm/databases:	metadata.993253759.db	application/x-sqlite3	Show
20	Created	snapshot_2	/data/data/com.google.android.gm/databases:	metadata.993253759.db-shm	application/octet-stream	Show
21	Created	snapshot_2	/data/data/com.google.android.gm/databases:	metadata.993253759.db-wal	application/octet-stream	Show
22	Created	snapshot_2	/data/data/com.google.android.gm/files/downloads:	077e6f5e9112c56220b3da770ec7e0dc	inode/directory	Show
23	Modified	snapshot_2	/data/data/com.google.android.gm/files/	persisted_config	application/octet-stream	Show
24	Modified	snapshot_2	/data/data/com.google.android.gm/shared_prefs/	Account-dcstestupo@gmail.com.xml	text/xml	Show

Figure 10: List of files that differ among two snapshots.

As shown in the figure, for each file AnForA reports its status (one of created, modified, or deleted), the identifier of the snapshot containing the file, its pathname on the device file system where it is stored, its name, and its type, and allows the analyst to visualize the file (in case it has been created or deleted), or the changes occurred in this file as a consequence of the actions corresponding to the last snapshot.

For instance, from Fig. 10 we see that file in row 5 has been created in `snapshot_2` (i.e., it was not present in `snapshot_1`), while file in row 7 has been deleted (hence, it is contained in `snapshot_1` only).

By clicking on the `Action` column of each file, AnForA shows the changes occurred in that file from the first to the second snapshot. Fig. 11 shows the changes occurred in the text file

`Account-dcstestupo@gmail.com.xml` as consequence of sending an email message using the Gmail app. As can be seen from the figure, the contents of that file in the two snapshots are shown

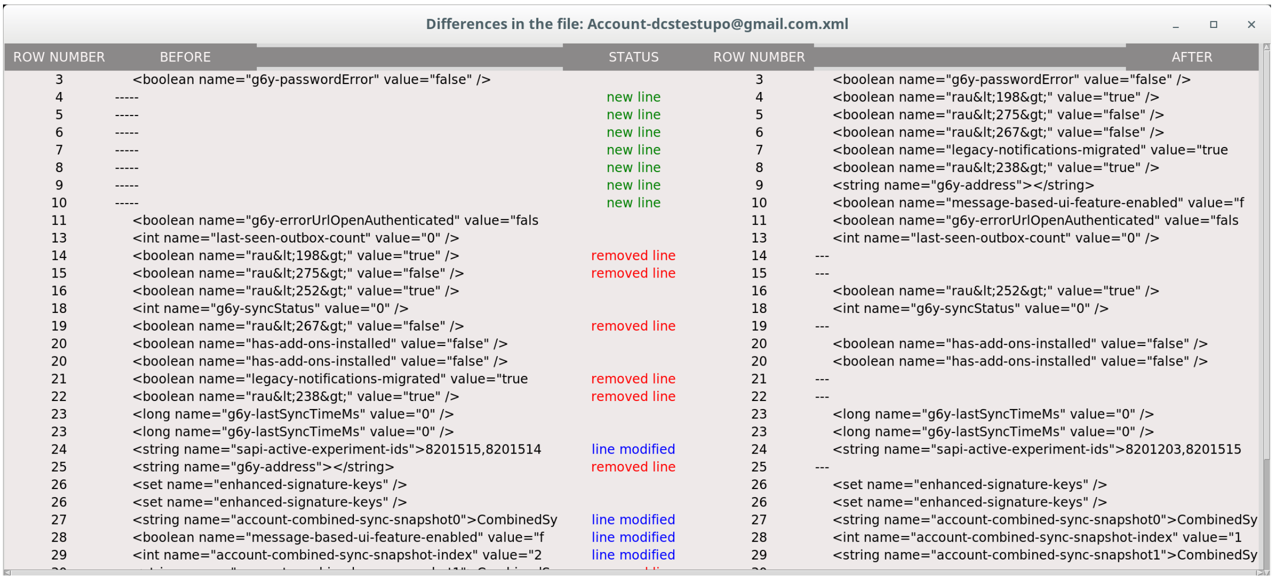

Figure 11: AnForA shows what has been changed in file Account-dcstestupo@gmail.com.xml.

side-by-side, and the lines that have been added (labeled as `new line`), removed (labeled as `removed line`), or modified (labeled as `line modified`) are suitably identified.

In addition to text files, in its current implementation AnForA is able to identify and show the differences also in SQLite databases, while for file whose encoding is not supported it only pinpoints the blocks of data that differ between two snapshots.

5. Validation

In order to validate the results generated by AnForA, and in particular its ability to provide completeness and soundness, we use it to perform the forensic analysis of Android apps already analyzed in the literature, so that the results it yields can be compared against those that have been already published.

More specifically, we consider the following apps:

- *Google Gmail*: we analyze version 8.5.6.199637500, and compare AnForA's results against those reported in [4, 18] (both referring to an unspecified version);

- *Microsoft Skype*: we analyze version 8.37.0.98, and compare AnForA's results against those reported in [4] (for version 6.31.0.709), [17] (for an unspecified version), and [18] (for an unspecified version);

- Facebook Messenger: we analyze version 113.0.0.21.70, and compare AnForA's results against those reported in [4] (for version 68.0.0.22.67), [12] (for version 86.0.0.17.70), and [23] (for version 113.0.0.21.70).

All the experiments have been performed on a virtualized mobile device running Android 9 for the Intel Atom x86_64 platform.

The results of our validation can be summarized as follows. AnForA has been able to identify and locate all the data reported in the literature for the applications we considered. In most cases, however, these data were stored in different files and/or different paths than those previously reported; we believe this is a direct consequence of the fact that in our experiments we considered later versions of these applications and of Android.

Furthermore, for the Gmail application, AnForA found artifacts that had not been previously reported in the literature. This could also be due to differences in the versions of Gmail considered in these studies with respect to the one we use in our experiments.

The results of our experiments demonstrates the ability of AnForA of achieving completeness, as it has been able to find – for each application – all (actually, even more) the data reported in previous studies. Furthermore, they also demonstrate soundness, since for each application, only the data it generates have been included by AnForA into its reports. Finally, our experiments also demonstrate the advantage of using AnForA instead of manual analysis, both in terms of time and effectiveness, as it enables to quickly and accurately repeat the analysis of new versions of already-analyzed applications when needed.

In the following, we discuss the comparison of AnForA's results against those obtained in published works for *Gmail* (Sec. 5.1), *Skype* (Sec. 5.2), and *Facebook Messenger* (Sec. 5.3).

5.1. *Google Gmail*

The results of the analysis of Gmail are reported in Table 2, in which we compare the relevant files identified by AnForA (column `AnForA`) against those reported by [4, 18] (column `Literature`).

Table 2: Location of artifacts for Google Gmail. Unless otherwise specified, all paths are relative to `/data/data/com.google.android.gm`.

Row	Artifact	Location	
		AnForA	Literature
1	*Preferences*	`shared_prefs/*.xml`	`shared_prefs/*.xml`
2	*Account details (Google account)*	`shared_prefs/MailAppProvider.xml`,	`cache/<username>@gmail.com.db`
3	*Messages (Google account)*	`databases/bigTopDataDB.<account-id>`	`databases/mailstore.<username>@gmail.com.db`
4	*Attachment metadata (Google account)*	`databases/metadata.<account-id>`	`cache/<username>@gmail.com/`
5	*Attachment files (Google account)*	`files/downloads/`	`cache/<username>@gmail.com/`
6	*Account details (IMAP account)*	`shared_prefs/MailAppProvider.xml`, `databases/EmailProvider.db`	n/a
7	*Messages (IMAP account)*	`databases/EmailProvider.db`	n/a
8	*Attachments metadata (IMAP account)*	`databases/EmailProvider.db`	n/a
9	*Attachment files (IMAP account)*	`cache/*.attachment`	n/a
10	*Third-party services*	Android's account files, Google Mobile Services' files	n/a

As in the studies already appeared in the literature, AnForA found that the installation folder of Gmail is `/data/data/com.google.android.gm`. However, it also found several important differences with respect to them, namely:

- There are differences in the location and contents of the data generated when using a Google email account with respect to using an account of another provider via the *IMAP* protocol.
- For Google email accounts, AnForA identified the same data reported in the literature, but located them in different files and folders (rows $2-5$ in Table 2).
- For generic email accounts, AnForA found that Gmail generates data not identified by previous studies (rows $6-9$ in Table 2), and in particular:
  - When a new generic account is registered in Gmail, a new record is added to table `Account` of database `databases/EmailProvider.db` to store account information, including the email address, the display name and the sender name. Furthermore, a file named `Account-<imap server>.xml` is created to store cipher settings for the IMAP server `<imap server>` associated with this account.
  - Messages are stored in table `Message` of database `databases/EmailProvider.db`.

  – Attachment metadata are stored in table `Attachment` of database `databases/EmailProvider.db`.
  – Downloaded attachment contents are saved as files, named as `<name>.attachment`, in the `cache` subdirectory, where `<name>.attachment` encodes the download timestamp (e.g., `2019-02-27-21:57:157770246422182752973.attachment`). These attachments are also stored as URIs in field `cachedFile` of table `Attachment` (so that a downloaded attachment can be easily related to its metadata).

- Unlike previous studies, AnForA has been able to identify two third-party services (namely, *Android account service* and *Google Mobile Service*) that are used by Gmail, and that may write – in their folders – data generated on the behalf of Gmail. In particular, when an account is added to Gmail, a related entry is also added to the device's accounts by updating the Android's contacts database `contacts2.db` located in the directory `/data/data/com.android.providers.contacts/databases`. Furthermore, for Google accounts, the Google Mobile Services collection is also updated accordingly by updating the corresponding databases located in the directory `/data/data/com.google.android.gms/databases`.

5.2. Microsoft Skype

The results of the analysis of Skype are reported in Table 3, in which we compare the relevant files identified by AnForA (column `AnForA`) against those reported by [17, 4, 18] (column `Literature`).

Table 3: Location of artifacts for Skype. Unless otherwise specified, all paths are relative to `/data/data/com.skype.raider`.

Artifact	Relevant paths	
	As reported by AnForA	Literature
Contacts, conversations, and call logs	`databases/s4l-live:<username>.db`	`files/live#3 <username>/main.db` [4], `files/<username>/main.db` [17, 18]
Shared media metadata	`databases/s4l-live:<username>.db`	`files/<username>/main.db` [18]
Shared media files	`cache/FileCache/`	`files/live#3 <username>/media_messaging/media_cache/` [4], `cache/` [17]

Unlike the Gmail case, the results obtained by AnForA for Skype agree with those reported in the literature, as far as the data that are generated by this app are concerned, while differences have been found with respect to the location of these data, as shown in Table 3. Hence, we do not discuss the contents of these files here (the interested reader may refer to the literature for this discussion).

5.3. Facebook Messenger

The results of the analysis of Facebook Messenger are reported in Table 4, in which we compare the relevant files identified by AnForA (column `AnForA`) against those reported by [4, 12, 23] (column `Literature`). Again, as in the Skype case, the results obtained by AnForA agree with those reported

Table 4: Location of artifacts for Facebook Messenger. Unless otherwise specified, all paths are relative to `/data/data/com.facebook.orca`.

Artifact	Relevant paths	
	As reported by AnForA	Literature
Contacts	`databases/contacts_db2`	`databases/contacts_db2` [4, 23]
Conversations	`databases/threads_db2`	`databases/threads_db2`
Call logs	`databases/call_log.sqlite`	`databases/call_log.sqlite` [4]
Shared media files	`cache/fb_temp/`	`cache/fb_temp/` [4], `cache/`, `files/` [12, 23]

in the literature as far as the data that are generated by this app, while differences have been found with respect to the location of these data, as shown in Table 4. Therefore, as for the Skype results, we do not discuss the contents of these files here (the interested reader may refer to the literature for this discussion).

6. Conclusions and future work

In this paper we have presented the design, implementation, and evaluation of AnForA, a software tool that automates most of the activities that need to be carried out to forensically analyze Android applications designed in such a way to yield various properties namely fidelity, completeness, soundness, effectiveness, repeatability, and generality. In particular, AnForA relies on the use of virtualized Android devices, on which the application is installed, and on a set of software components that (a) interact with its interface as a human user would do, (b) monitor the changes to the file systems induced by these interactions, (c) extract modified files, and (d) locate these modifications.

We have devised a proof-of-concept implementation of AnForA, and we have used it to assess its ability in achieving its design goals, by analyzing through it several Android applications already studied in the literature, so that we can compare AnForA's results against those reported in these papers. The results of our evaluation confirm that AnForA greatly simplifies the forensic analysis of Android applications, and exhibits all the properties mentioned above.

As future work, we plan to extend AnForA's capabilities in decoding files along two directions: (a) integrate in it suitable helper applications able to decode files whose encoding is known (e.g., PDF or Microsoft Office files), and (b) use machine learning techniques to enable it to automatically discover the proprietary encoding schemes that many application use for the data they store on the device (e.g., applications that serialize complex data structure containing fields of different types).

References

[1] Anglano, C., 2014. Forensic Analysis of WhatsApp Messenger on Android Smartphones. Digital Investigation 11, 201–213. doi:10.1016/j.diin.2014.04.003.

[2] Anglano, C., Canonico, M., Guazzone, M., 2016. Forensic Analysis of the ChatSecure Instant Messaging Application on Android Smartphones. Digital Investigation 19, 44–59. doi:10.1016/j.diin.2016.10.001.

[3] Anglano, C., Canonico, M., Guazzone, M., 2017. Forensic Analysis of Telegram Messenger on Android Smartphones. Digital Investigation 23, 31–49. doi:10.1016/j.diin.2017.09.002.

[4] Cahyani, N.D.W., Rahman, N.H.A., Glisson, W.B., Choo, K.K.R., 2017. The role of mobile forensics in terrorism investigations involving the use of cloud storage service and communication apps. Mobile Networks and Applications 22, 240–254. doi:10.1007/s11036-016-0791-8.

[5] Cheng, C.C.C., Shi, C., Gong, N.Z., Guan, Y., 2018. EviHunter: Identifying Digital Evidence in the Permanent Storage of Android Devices via Static Analysis, in: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, ACM, New York, NY, USA. pp. 1338–1350. doi:10.1145/3243734.3243808.

[6] Google, a. Android Debug Bridge. Available at https://developer.android.com/studio/command-line/adb.html. Last visited on May 2019.

[7] Google, b. Android SDK Tools. Available at https://developer.android.com/studio/index.html. Last visited on May 2019.

[8] Google, c. UI Automator. Available at https://developer.android.com/training/testing/ui-automator. Last visited on May 2019.

[9] Google, 2016. Run Apps on the Android Emulator. Available at https://developer.android.com/studio/run/emulator.html.

[10] Gregorio, J., Gardel, A., Alarcos, B., 2017. Forensic analysis of telegram messenger for windows phone. Digit. Investig. 22. doi:10.1016/j.diin.2017.07.004.

[11] Husain, M.I., Sridhar, R., 2010. iForensics: Forensic Analysis of Instant Messaging on Smart Phones, in: Goel, S.

(Ed.), Digital Forensics and Cyber Crime. Springer Berlin Heidelberg. volume 31 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*. doi:10.1007/978-3-642-11534-9_2.

[12] Lin, X., 2018. Android Forensics. Springer. pp. 335–371. doi:10.1007/978-3-030-00581-8_15.

[13] Lin, X., Chen, T., Zhu, T., Yang, K., Wei, F., 2018. Automated forensic analysis of mobile applications on android devices. Digital Investigation 26, S59 – S66.

[14] Mehrotra, T., Mehtre, B.M., 2013. Forensic analysis of Wickr application on android devices, in: 2013 IEEE International Conference on Computational Intelligence and Computing Research, pp. 1–6.

[15] Ovens, K.M., Morison, G., 2016. Forensic analysis of Kik messenger on iOS devices. Digital Investigation 17. doi:10.1016/j.diin.2016.04.001.

[16] Pauck, F., Bodden, E., Wehrheim, H., 2018. Do Android Taint Analysis Tools Keep Their Promises?, in: Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, ACM, New York, NY, USA. pp. 331–341.

[17] Sgaras, C., Kechadi, M.T., Le-Khac, N.A., 2015. Forensics acquisition and analysis of instant messaging and voip applications, in: Garain, U., Shafait, F. (Eds.), Computational Forensics, Springer. pp. 188–199. doi:10.1007/978-3-319-20125-2_16.

[18] Tamma, R., Skulkin, O., Mahalik, H., Bommisetty, S., 2018. Practical Mobile Forensics. 3rd ed., Packt Publishing.

[19] Tso, Y.C., Wang, S.J., Huang, C.T., Wang, W.J., 2012. iPhone Social Networking for Evidence Investigations Using iTunes Forensics, in: Proc. of the 6^{th} International Conference on Ubiquitous Information Management and Communication, ACM, New York, NY, USA. pp. 1–7. doi:10.1145/2184751.2184827.

[20] Walnycky, D., Baggili, I., A.Marrington, Moore, J., Breitinger, F., 2015. Network and device forensic analysis of Android social-messaging applications. Digital Investigation 14, Supplement 1, S77–S84. doi:10.1016/j.diin.2015.05.009. proc. of 15^{th} Annual DFRWS Conference.

[21] Wu, S., Zhang, Y., Wang, X., Xiong, X., Du, L., 2017. Forensic analysis of WeChat on Android smartphones. Digital Investigation doi:10.1016/j.diin.2016.11.002.

[22] Xu, Z., Shi, C., Cheng, C.C., Gong, N.Z., Guan, Y., 2018. A Dynamic Taint Analysis Tool for Android App Forensics, in: Proc. of the 2018 IEEE Security and Privacy Workshops (SPW).

[23] Zhang, H., Chen, L., Liu, Q., 2018. Digital forensic analysis of instant messaging applications on android smartphones, in: 2018 International Conference on Computing, Networking and Communications (ICNC), pp. 647–651. doi:10.1109/ICCNC.2018.8390330.