

AndroEvolve: Automated Update for Android Deprecated-API Usages

Stefanus A. Haryono*, Ferdian Thung*, David Lo*, Lingxiao Jiang*, Julia Lawall[‡], Hong Jin Kang*, Lucas Serrano[†], and Gilles Muller[‡]

*School of Information Systems, Singapore Management University, Singapore
{stefanusah,ferdianthung,davidlo,hjkang.2018,lxjiang}@smu.edu.sg

[†]Sorbonne University/Inria/LIP6, France

Lucas.Serrano@lip6.fr

[‡]Inria, France

{Gilles.Muller,Julia.Lawall}@inria.fr

Abstract—The Android operating system (OS) is often updated, where each new version may involve API deprecation. Usages of deprecated APIs in Android apps need to be updated to ensure the apps’ compatibility with the old and new versions of the Android OS. In this work, we propose AndroEvolve, an automated tool to update usages of deprecated Android APIs, that addresses the limitations of the state-of-the-art tool, CocciEvolve. AndroEvolve utilizes data flow analysis to solve the problem of out-of-method-boundary variables, and variable denormalization to remove the temporary variables introduced by CocciEvolve. We evaluated the accuracy of AndroEvolve using a dataset of 360 target files and 20 deprecated Android APIs, where AndroEvolve is able to produce 319 correct updates, compared to CocciEvolve which only produces 249 correct updates. We also evaluated the readability of AndroEvolve’s update results using a manual and an automatic evaluation. Both evaluations demonstrated that the code produced by AndroEvolve has higher readability than CocciEvolve’s. A video demonstration of AndroEvolve is available at <https://youtu.be/siU0tuMITXI>.

I. INTRODUCTION

Android is one of the most widely used operating systems (OS) in the recent years due to the popularity of smartphones. It is frequently updated to add new features and fix bugs. Following each OS update, changes and modifications in the Android APIs are common, which often results in deprecation of the older version of an API and makes the deprecated APIs unusable for the newer version of the OS. This problem, named Android fragmentation [1], [2], is a common occurrence.

Due to the above issue, updating usages of deprecated Android APIs should be prioritized. Several studies have promoted automated approaches to updating deprecated Android API usages. A4 [3] automatically assists Android API migrations by learning API migration patterns from the differences between the deprecated and updated code examples. AppEvolve [4] uses both before- and after- update code examples to automatically learn the update. However, a replication study by Thung et al. [5] found several weaknesses in AppEvolve, where the target file must be structured the same way as the code examples.

Following Thung et al.’s findings, Haryono et al. [6] presented CocciEvolve, an automated deprecated Android API

usage update tool built on top of Coccinelle4J [7]. CocciEvolve offers several improvements from AppEvolve: (1) a readable update script expressed in the Semantic Patch Language (SmPL) [8]; (2) only requiring a single after-update example; and (3) the ability to update multiple invocations of a deprecated API within a single file. Code example used in CocciEvolve must be in the form of an `if` statement, containing the updated API in the `then` statement and the old API in the `else` statement, or vice versa. An example of such a code example is given in Fig. 1, where the deprecated API `getCurrentMinute` is shown at line 5 and the updated API `getMinute` at line 3.

```
1 if (android.os.Build.VERSION.SDK_INT >=
2     android.os.Build.VERSION_CODES.M) {
3     minutes = picker.getMinute();
4 } else {
5     minutes = picker.getCurrentMinute();
6 }
```

Fig. 1. An example of after-update code for `getCurrentMinute()` API

While CocciEvolve significantly improves on AppEvolve, it has several shortcomings. CocciEvolve is only able to resolve (i.e., locate and identify the definition for) the value of a variable within the same method, meaning that if an expression used in the updated API invocation is defined outside of the currently processed method, CocciEvolve will not be able to resolve its value. Furthermore, during the update process, CocciEvolve introduces temporary variables that refer to other existing variables in the target file. These temporary variables are not deleted in the update result, making the update result less readable.

In this paper, we present AndroEvolve, an automated Android deprecated API usage update tool that addresses the limitations of CocciEvolve. This paper is a tool demo paper that supplements our full paper [9]. AndroEvolve addresses the limitations through two features: data flow analysis and variable denormalization. Data flow analysis is used to resolve the values used as the updated API invocation arguments, including values of all out-of-method variables. Using the data flow analysis, AndroEvolve automatically locates the definitions of those variables, and uses their values to replace

the variables used in the API invocation arguments. Variable name denormalization refers to the process that reverts the normalization done by CocciEvolve that introduces the temporary variables in the update result. This process reverts the temporary variables back to their original variable names within the code, removing all unnecessary changes and improving the overall readability of the update result.

We evaluated AndroEvolve on a dataset of 360 target files containing 20 deprecated Android APIs. Based on this evaluation, AndroEvolve produces 316 correct updates while CocciEvolve only produces 249 correct updates, showing that AndroEvolve outperforms CocciEvolve by 26.90%. We also measure the readability of the update results of both tools using manual and automatic evaluation methods. For the manual evaluation, we asked the opinions of two experienced Android engineers. For the automated evaluation, we use a popular readability scoring tool [10]. We found that AndroEvolve’s update results are 83% and 50% more readable compared to CocciEvolve’s according to the manual and automated readability evaluation, respectively.

II. DESIGN AND IMPLEMENTATION

A. AndroEvolve Architecture

The overall architecture of AndroEvolve is shown in Fig. 2. The workflow of AndroEvolve is comprised of *update-script creation* followed by *update-script application*.

Update-script creation takes as input the API update mapping and an after-update example of the API. The API update mapping consists of the API signatures of the deprecated API and its corresponding updated API. This mapping is used to identify the deprecated and the updated API from the code example and target code. Next, data flow analysis is employed to resolve the values of out-of-method variables in the update example that are used as the arguments of the updated API invocation but are not used in the deprecated API invocation. Variable normalization is then applied to the part of the code containing the deprecated and updated API invocations, introducing temporary variables to name all complex subexpressions to facilitate the update. The goal is to minimize the syntactic differences between the update example and the target code to be updated. The update script, expressed using the Semantic Patch Language (SmPL), is created from the normalized update example.

Update-script application takes as input the target code to be updated, the update script from the update-script creation process, and the API update mapping. The update-script application is also comprised of several steps. First, variable normalization is applied into the target code. Then, the update script is applied to the normalized target code using Coccinelle4J [7], resulting in the updated code. Following the update script application, we copy method and class definitions that are being used in the updated API arguments to the updated code. Finally, variable denormalization is applied to this updated code, replacing the temporary variables with their original expressions.

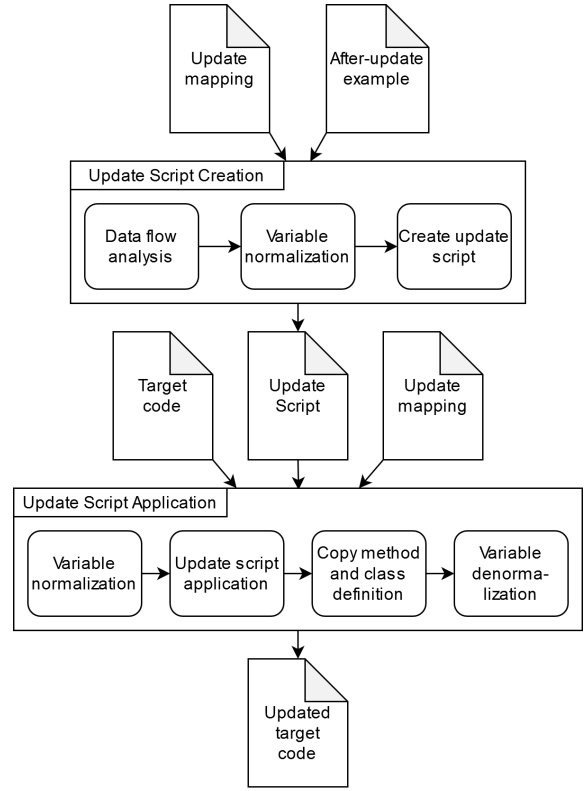


Fig. 2. Summary of AndroEvolve workflow

The following subsections further describe these components. The complete details of each AndroEvolve component and process are available in our full paper [9].

B. Data Flow Analysis

AndroEvolve utilizes a data flow analysis (DFA) to resolve the values of expressions used as arguments in an API method invocation. The data flow analysis is able to gather and predict a set of possible values of an object/variable at any given point inside the code, thus it is able to predict and resolve the correct replacement values for any expressions used in the updated API invocation arguments.

We build the DFA using the symbol resolver from Java Symbol Solver that is available in Javaparser.¹ The DFA performs a bottom-up search from the variables or expressions used as the API invocation arguments and expands the search scope until it finds the value or methods referred to by the expressions. Values and method invocations that are found by this DFA are then used to replace the original expressions used as API invocation arguments to ensure that the API arguments are in the form of either a literal expression, a static class member, a method invocation, or an object creation.

If the found values are in the form of method invocations or object creations that are not available in the target code, it is necessary to copy the method or class definition from the after-update example to the target file to ensure that it is accessible in the updated code. For this purpose, if the resolved API

¹<https://javaparser.org/>

invocation arguments are in the form of method invocations or object creations, AndroEvolve will extract their definitions and copy them to the target file during the update. AndroEvolve only copies the methods and classes that are related to the resolved API arguments.

C. Variable Denormalization

Variable normalization introduces temporary variables for each API invocation argument which clutter the update result and reduce its readability. Using variable denormalization, AndroEvolve aims to remove these temporary variables and replace them with their values or referenced variables. For each temporary variable, AndroEvolve locates its definition and replaces it with the resolved values or expressions. Then, AndroEvolve deletes the declarations and definitions of the temporary variables as they are no longer needed.

III. EVALUATION

A. Dataset and Experiment Setup

Our dataset is comprised of three components: after-update examples used for the update-script creation, one-to-one API mappings from the deprecated APIs to the replacement APIs, and the target files to be updated. We utilize the after-update examples that were used in the evaluation of AppEvolve [4]. For the target files to be updated, we extend the target file dataset from CocciEvolve [6]. We randomly selected public GitHub projects obtained using AUSearch [11], a tool to search GitHub repositories for API usages. We collected 360 target files containing 20 deprecated Android APIs for our target files dataset.

We evaluated the performance of AndroEvolve by comparing its update accuracy (i.e., the percentage of correct updates) against CocciEvolve's. We also compare the readability of the update results of the two tools, using an automated and a manual approach. In the automated approach, we utilized a state-of-the-art automated readability measuring tool from Scalabrino et al. [10] which gives a score of 0.0-1.0 where higher score indicates better readability. In the manual approach, we asked two experienced Android developers to score 60 successful updated code fragments, with 30 updated code fragments from each of CocciEvolve and AndroEvolve. We choose 30 as the sample size to represent the variation in the syntax of the updated code. The developers were not told which tool was used to produce each updated code fragment. For each updated code fragment, the developers were asked to give score in the Likert scale of 1-5 for the readability and the naturalness of the code. A higher score indicates higher readability and higher confidence that the code resembles code produced by a human.

B. Experiment Results

We manually check the correctness of the updated code produced by both AndroEvolve and CocciEvolve. On the dataset, AndroEvolve and CocciEvolve are able to provide correct updates for 316 and 249 target files, respectively. Our analysis shows that data flow analysis helps to improve the

number of correct updates produced by AndroEvolve. It works particularly well in cases where the after-update example uses out-of-method variables.

In the readability evaluation, using the automatic approach, we measure the readability scores of the updated code and calculate the average for each API. We found that AndroEvolve's updated code achieves higher readability than CocciEvolve's for all of the evaluated deprecated API migrations. This finding is strengthened by our manual readability measurement where the average readability score given by the developers for AndroEvolve achieves a score of 4.817, while CocciEvolve only achieves a score of 2.633. Similarly, the code naturalness measure for AndroEvolve achieves an average score 4.917, while CocciEvolve achieves an average score of 2.433. These evaluations highlight that AndroEvolve offers a significant improvement in code readability as compared to CocciEvolve.

IV. USAGE EXAMPLE

AndroEvolve can be used via a command line interface (CLI). Two main functionalities are provided:

1) **Update Patch Creation** Update patch creation takes as input the API update mapping and the after-update example. AndroEvolve's CLI takes the command line argument `--generate-patch` to specify the update patch creation, `--input` to indicate the path for the update example file, and `--output` to indicate the path for the generated patch file. The format of this command is as follows:

```
java -jar AndroEvolve.jar --generate-patch
    <deprecated_api_signature> <updated_api_signature>
    --input <update_example_path>
    --output <output_path>
```

2) **Update Patch Application** Update patch application takes as input the API update mapping, the target file to be updated, and the previously created update patch file. AndroEvolve's CLI takes the command line argument `--apply-patch` to specify the update patch application process, `--input` to indicate the path to the target file, `--patch` to indicate the path to the update patch file, and `--output` to indicate the path for the produced updated target file. The format of this command is as follows:

```
java -jar AndroEvolve.jar --apply-patch
    <deprecated_api_signature> <updated_api_signature>
    --input <target_filepath> --patch <patch_filepath>
    --output <output_path>
```

```
1 private static final int DURATION = 50;
2 private static final int AMPLITUDE = 175;
3 public static void itemActivated(Context context) {
4     long milliseconds = DURATION;
5     if (android.os.Build.VERSION.SDK_INT >=
6         android.os.Build.VERSION_CODES.O) {
7         int amplitude = AMPLITUDE;
8         VibrationEffect effect = VibrationEffect.
9             createOneShot(milliseconds, amplitude);
10        vibrator.vibrate(effect);
11    } else {
12        vibrator.vibrate(milliseconds);
13    }
14 }
```

Fig. 3. After-update example for `android.os.Vibrator#vibrate(long)` deprecated API

As an example of the patch creation and update application process provided by AndroEvolve, consider the `android.os.Vibrator#vibrate(long)` deprecated API. This deprecated API needs to be updated into `android.os.Vibrator#vibrate(android.os.VibrationEffect)`, meaning that a change to the API parameter is necessary. An after-update example is provided in Fig. 3. Line 10 contains the updated API usage, while line 12 contain the deprecated API usage. Using this update example, we use the following command to run the update patch creation process:

```
java -jar AndroEvolve.jar --generate-patch
  android.os.Vibrator#vibrate(long)
  android.os.Vibrator#vibrate(android.os.VibrationEffect)
  --input example.java
  --output vibrate_update.cocci
```

In line 10, we can see that the updated API invocation uses the variable `effect` as its argument. The data flow analysis conducted a bottom-up search in the example for this variable, finding its definition at lines 8-9. The search is continued as the `effect` variable definition uses other variables, i.e, `milliseconds` and `amplitude`. This search is done recursively until all variables related to the API invocation arguments are resolved. Following this data flow analysis, variable normalization is conducted, and the update patch is created. The created update patch is shown in Fig. 4.

```
1 @bottomupper_classname@
2 expression exp0, exp1;
3 identifier iden0, classIden;
4 @@
5 ...
6 + if (android.os.Build.VERSION.SDK_INT >=
7 +   android.os.Build.VERSION_CODES.O) {
8 +   VibrationEffect newParameterVariable0 =
9 +     VibrationEffect.createOneShot(50, 175);
10 + classIden.vibrate(newParameterVariable0);
11 + } else {
12 classIden.vibrate(iden0);
13 + }
```

Fig. 4. Update patch based on the after-update example in Fig. 3 for the `android.os.Vibrator#vibrate(long)` deprecated API

Using this created update patch, we can apply the update to a target file containing the deprecated API usage. The following command runs the update application process:

```
java -jar AndroEvolve.jar --apply-patch
  android.os.Vibrator#vibrate(long)
  android.os.Vibrator#vibrate(android.os.VibrationEffect)
  --input main.java --patch vibrate_update.cocci
  --output main_updated.java
```

The target file is first normalized to remove syntactic differences from the update patch. Then, the update patch is applied, resulting in the updated code. Finally, variable denormalization is applied to remove the temporary variables that are introduced in the normalization. Fig. 5 highlights the code differences resulting from the application of this update. An `if` statement that checks the Android version is added at line 3. Lines 5-7 contain the updated `android.os.Vibrator#vibrate(android.os.VibrationEffect)` API usage, along with the new parameter required for its invocation.

V. CONCLUSION

We have proposed AndroEvolve as the next step in the automated update of usages of Android deprecated APIs. AndroEvolve provides a significant improvement compared to the state-of-the-art tool, CocciEvolve, through the addition of data flow analysis to handle out-of-method variables, and variable denormalization to increase the readability of the update result by removing temporary variables previously introduced by CocciEvolve. We evaluated the accuracy and readability of AndroEvolve and demonstrated its higher accuracy and readability compared to CocciEvolve. AndroEvolve is available at <https://github.com/soarsmu/AndroEvolve> **Acknowledgement.** This research is supported by the Singapore NRF (award number: NRF2016-NRF-ANR003) and the ANR ITrans project.

```
1 public void Once(long milliseconds) {
2     if (MyVibrator.hasVibrator()) {
3         + if (Build.VERSION.SDK_INT >=
4         +   Build.VERSION_CODES.O) {
5         +   MyVibrator.vibrate(VibrationEffect.
6         +     createOneShot(500, VibrationEffect.
7         +       DEFAULT_AMPLITUDE));
8         +   } else {
9             MyVibrator.vibrate(milliseconds);
10        +   }
11    }
12 }
```

Fig. 5. Update result for a target file using the update patch in Fig. 4

REFERENCES

- [1] D. Han, C. Zhang, X. Fan, A. Hindle, K. Wong, and E. Stroulia, "Understanding Android fragmentation with topic analysis of vendor-specific bugs," in *2012 19th Working Conference on Reverse Engineering*. IEEE, 2012, pp. 83–92.
- [2] L. Wei, Y. Liu, and S.-C. Cheung, "Taming Android fragmentation: Characterizing and detecting compatibility issues for Android apps," in *ASE*. IEEE, 2016, pp. 226–237.
- [3] M. Lamothe, W. Shang, and T.-H. Chen, "A4: Automatically assisting Android API migrations using code examples," *ArXiv*, vol. abs/1812.04894, 2018.
- [4] M. Fazzini, Q. Xin, and A. Orso, "Automated API-usage update for Android apps," in *ISSTA*. ACM, 2019, pp. 204–215.
- [5] F. Thung, H. Kang, L. Jiang, and D. Lo, "Towards generating transformation rules without examples for Android API replacement," in *2019 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 09 2019, pp. 213–217.
- [6] S. A. Haryono, F. Thung, H. J. Kang, L. Serrano, G. Muller, J. Lawall, D. Lo, and L. Jiang, "Automatic Android deprecated-API usage update by learning from single updated example," in *IEEE International Conference on Program Comprehension*, 2020.
- [7] H. J. Kang, F. Thung, J. Lawall, G. Muller, L. Jiang, and D. Lo, "Automating program transformation for Java using semantic patches," in *European Conference on Object-Oriented Programming (ECOOP)*, 2019.
- [8] J. Lawall and G. Muller, "Coccielle: 10 years of automated evolution in the Linux kernel," in *USENIX Annual Technical Conference*, 2018, pp. 601–614.
- [9] S. A. Haryono, F. Thung, D. Lo, L. Jiang, J. Lawall, H. J. Kang, L. Serrano, and G. Muller, "AndroEvolve: Automated Android API update with data flow analysis and variable denormalization," 2020.
- [10] S. Scalabrino, M. Linares-Vásquez, R. Oliveto, and D. Poshyvanyk, "A comprehensive model for code readability," *Journal of Software: Evolution and Process*, vol. 30, 06 2018.
- [11] M. H. Asyrofi, F. Thung, D. Lo, and L. Jiang, "AUSearch: Accurate API usage search in GitHub repositories with type resolution," in *IEEE International Conference on Software Analysis, Evolution and Reengineering*, 2020.