Slicer4J: A Dynamic Slicer for Java

Khaled Ahmed
Univ. of British Columbia, Canada
khaledea@ece.ubc.ca

Mieszko Lis
Univ. of British Columbia, Canada
mieszko@ece.ubc.ca

Julia Rubin
Univ. of British Columbia, Canada
mjulia@ece.ubc.ca

ABSTRACT
Dynamic program slicing is used in a variety of tasks, including program debugging and security analysis. Despite being extensively studied in the literature, the only dynamic slicing solution for Java programs that is publicly available today is a tool named JavaSlicer. Unfortunately, JavaSlicer only supports programs written in Java 6 or below and does not support multithreading. To address these limitations, this paper contributes a new dynamic slicing tool for Java, named Slicer4J. Slicer4J uses low-overhead instrumentation to collect a runtime execution trace; it then constructs a thread-aware, inter-procedural dynamic control-flow graph and uses the graph to compute the slice. To support slicing through Java framework methods and native code, Slicer4J relies on a set of pre-constructed data-flow summaries of the main framework methods. It also allows the users to further customize this set, adding user-defined methods when needed. We demonstrate the applicability of Slicer4J on ten benchmark and open-source Java programs, comparing it with JavaSlicer, and discuss how to use and extend the tool.

Tool package and demo: https://github.com/resess/Slicer4J

CCS CONCEPTS
• Theory of computation → Program analysis; • Software and its engineering → Maintaining software.

KEYWORDS
Program analysis, dynamic slicing, Java

ACM Reference Format:

1 INTRODUCTION
Program slicing [22] computes the set of statements that affect a particular variable or statement of interest, often referred to as a slicing criterion. Slicing techniques are used in a variety of tasks, e.g., program debugging, to help locate the origin of an error. Slicing can be performed either statically or dynamically [14]. While static slicing considers all possible program paths leading to the slicing criterion, dynamic slicing focuses on one concrete execution.

The main idea behind a dynamic slicing tool is to first collect an execution trace of a program, i.e., the set of all executed program statements. Then, the tool inspects control and data dependencies of the trace statements, identifying statements that affect the slicing criterion and omitting the rest. The produced dynamic slices are more compact than static ones, making them particularly suitable for debugging activities [1, 2].

While there is a large number of papers proposing and optimizing dynamic slicing techniques, e.g., [6, 16, 17, 20], only one dynamic slicing solution of Java programs is publicly available at the time of writing: JavaSlicer [10, 11]. JavaSlicer is an efficient tool that is optimized for low instrumentation overhead with the capability of producing compressed traces for space optimization. However, the tool has several limitations: it does not support newer language constructs from Java versions 6, e.g., lambda expressions; it does not support multithreading, producing instead one separate slice per thread; and it cannot handle certain Java framework methods, as well as framework and custom methods implemented in native code, e.g., java.lang.System.arraycopy(...).

To address these issues, we contribute a dynamic slicing tool, named Slicer4J, which supports Java version 9, multithreading, and modeling of Java framework methods. Slicer4J relies on the infrastructure we created for slicing Android applications [3], adapting it to support stand-alone Java programs. The implementation and the documentation of Slicer4J is available online [4]. We also demonstrate the applicability of Slicer4J and compare it with JavaSlicer on ten Java programs.

Researchers and tool developers can use Slicer4J as a part of their dynamic analysis systems in a wide variety of projects, such as fault localization [6], regression analysis [21], and malware detection [15]. Software developers can also utilize Slicer4J for fault localization, e.g., to narrow down the scope of statements to inspect given failed test assertions or crash sites.

2 DYNAMIC SLICING
We now define slicing for Java programs on an example in Figure 1. This program calculates the lengths of two arrays, given as the program command line arguments. Its main method defines and spawns two threads of type PThread, which will read and parse the first and the second argument, respectively (lines 2-6). The parsed arrays are stored in the arr field of PThread (line 13). After the threads terminate (the check is omitted in line 7), the main method reads, sums up, and outputs the length of these arrays (lines 8-9).

Each thread is initialized with the array of input arguments args and an index within this array index (lines 15-16). When the thread starts executing, it checks if the index is within the range of the input parameters and, if so, populates the array using a helper function parse, which we omit here for brevity (lines 19-20).
Figure 1: SliceMe: a faulty app.

The bug in this program occurs when one or both input parameters are omitted. For example, if the second parameter is omitted, p2.arr is not initialized, leading to an exception when calculating p2.arr.length in line 8. Slicing from the statement in line 8 produces the subset of executed program statements which contribute to the failure. These are the statements highlighted in the figure. We define the slice more formally next.

For simplicity of presentation, we refer to a statement in line \(i\) of our examples as \(s_i\), e.g., the if statement in line 19 is denoted by \(s_{19}\). When a program runs, each statement can be triggered multiple times during the execution, e.g., in multiple iterations of a loop or in different instances of a thread. We refer to each individual execution of a statement as a statement instance and denote the \(k\)th execution of a statement \(s_i\) as \(s_{i,k}\). For example, assuming \(p1\) is executed before \(p2\), the trace will contain statement instances \(s_{1,1}, s_{1,2}, s_{1,3}, s_{1,4}, s_{1,5}, s_{1,6}, s_{1,7}\). We refer to the full sequence of statement instances from a particular app run as an execution trace.

Dynamic control-flow dependencies represent a concrete transfer of control recorded during the program execution. We say that a statement instance \(s_{m,k}\) is control-flow-dependent on \(s_{j,k}\) if \(s_{m,k}\) is executed immediately after \(s_{j,k}\) in the same execution trace [5]. For the example in Figure 1, \(s_{1,1}\) is control-flow-dependent on \(s_{1,15}\).

A sequence of statement instances without jump statement (i.e., conditional or method call), except at the end of the sequence, is referred to as a basic block. For example, \(s_{1,15}\) and \(s_{1,16}\) are in the same basic block, while \(s_{1,19}\) and \(s_{2,0}\) are not.

A statement instance \(s_{m,k}\) is control-dependent on \(s_{j,k}\) if \(s_{j,k}\) can alter the program’s control and it determines whether \(s_{m,k}\) executes [9]. In Figure 1, \(s_{1,20}\) is control-dependent on \(s_{1,19}\), as the outcome of the if determines whether the control reaches \(s_{1,19}\) or not.

A statement instance \(s_{m,k}\) is a data-flow-dependent on \(s_{i,k}\) w.r.t. the variable \(v\) used in \(s_{m,k}\) if and only if \(s_{i,k}\) defines \(v\) and no other statement redefines \(v\) between \(s_{i,k}\) and \(s_{m,k}\) in the trace [1]. For instance, \(s_{1,5}\) is data-flow-dependent on \(s_{1,20}\) w.r.t. \(p1.arr\), as the statement in line 20 defines the \(p1.arr\) variable used in line 8.

A slicing criterion for an execution trace is a statement instance and all variables of interest used in this statement instance [14]. For the example in Figure 1, the app throws a NullPointerException in line 8, making line 8 with all variables that it uses a control- or data-flow-dependent slice for the developer interested in finding the cause of the crash.

A backward dynamic slice [14] is the set of all statement instances whose execution affects the slicing criterion, i.e., the set of instances on which the slicing criterion is control- or data-flow-dependent, either directly or transitively. For the example in Figure 1, the backward slice from \(s_{1,6}\) w.r.t. variable \(p1.arr\) consists of \(s_{1,1}, s_{1,2}, s_{1,3}, s_{1,4}, s_{1,5}, s_{1,6}\). The backward slice from \(s_{2,0}\) w.r.t. variable \(p2.arr\) consist of \(s_{1,6}\) and \(s_{2,1}\). The source code lines corresponding to these statement instances are highlighted in the figure.

3 SLICE4J DESIGN

We now describe the main design decisions behind SLICE4J. Figure 2 gives its high-level overview. Our tool receives the following inputs: (1) a Java Archive (JAR) file of the program to slice, (2) a slicing criterion, e.g., a statement that throws an exception, (3) a test that triggers this criterion, and (4) an optional list of user-defined method summaries, as described below. It produces as output a backward dynamic slice for the provided slicing criterion.

Instrumentation. As the first step, SLICE4J runs a lightweight static analysis to instrument the JAR. Our instrumentation relies on Soot [18] and works on the Jimple intermediate code representation [19]. It records a unique identifier of each execution thread (using Thread.currentThread()) and for each statement within the thread. SLICE4J also includes an option to perform basic-block-level instrumentation, which records a unique identifier of each basic block rather than each statement within the thread [13]. Such implementation relies on the assumption that control is unlikely to switch from one thread to another within a basic block; it thus should not be used in heavily multi-threaded code.

While every local variable can be uniquely identified within its declaring method, shared variables (e.g., object fields) may have different names in the same/different methods. We thus follow Agrawal et al. [1] and record unique identifiers for fields (with Java. lang. System.identityHashCode method), which we use to identify fields when calculating data-flow dependencies.

As Java framework code is not part of the input JAR, it cannot be instrumented by SLICE4J. To still enable slicing through framework methods, SLICE4J relies on statically-generated Java framework method summaries, which we borrow from FlowDroid [8] and StubDroid [7]. These summaries, originally designed for taint analysis, define how a taint originates from each method input (i.e., fields and parameters) and propagates to its output (i.e., fields, parameters, and return value). We recast the taint-flow summaries as assignment statements, which our analysis uses to find data-flows.
For example, the StubDroid summary for the method java.lang.System.arraycopy(a, srcPos, b, ...) will state that a taint from the first array reaches the second array b. We use this information to represent a method as an assignment b = a, expressing the fact that the values of the first array are copied to the second array (directly or indirectly). Slicer4J also provides the option to augment this initial framework method summaries with a user-defined list of additional summaries, e.g., to model user-defined native methods.

An alternative solution, employed by JavaSlicer, is to instrument Java classes when they are loaded by the Java Virtual Machine (JVM). The main advantage of this solution is that it can handle framework methods and dynamically generated code. Its main disadvantage is that it cannot handle Java methods used as part of its instrumentation logic, e.g., java.lang.String: instrumenting these classes would cause an infinite loop when producing the trace. Moreover, such instrumentation has to be repeated every time a trace is generated, which slows things down when a program needs to be sliced multiple times and for different slicing criteria. Because of these limitations, we opt to statically instrument JARs and model framework/native methods. We plan to add handling of dynamically generated code as part of future work.

**Trace Parsing.** The instrumented JAR is executed (from command line or with a test script) to produce a trace. Slicer4J parses the trace and builds a Thread-aware, Inter-procedural Dynamic Control-flow Graph (TDCG), whose nodes are statement instances of each executed basic block. The graph has three types of unidirectional edges. Control-flow edges connect statement instances with the same thread, in their order of execution. Thread-control-flow edges connect statement instances from different threads, if they were executed immediately after each other in the trace. Finally, control-dependence edges connect statement instance $s_j^m$ to the instance $s_i^k$ if $s_j$ can alter the execution of $s_i$. We rely on a static control dependence graph produced by Soot and create a control-dependence edge between $s_j^m$ and $s_i^k$ if $s_i$ is control-dependent on $s_j$ in the static control dependency graph.

As the last step, we replace each statement instance that invokes a framework method (or any other method in the user defined list of methods summaries) with its corresponding summary, which we treat as a set of assignments for slicing purposes.

**Dynamic Slicing.** The TDCG is given as the input to the dynamic slicing step. We follow the backward slicing algorithm by Agrawal et al. [1]. The main objective of the algorithm is to identify control and data-flow dependencies of all variables used in the slicing criterion. In a nutshell, it works in an iterative manner, initializing the working set $W$ to include variables used in the slicing criterion. Then, it transitive adds all variables used in their control and data dependencies to $W$. The algorithm terminates when dependencies of all variables in $W$ are identified.

For identifying the control dependency of a statement instance, we simply follow the control-dependence edges in TDCG. For data-flow dependencies, we separate between local and shared variables (i.e., object fields). The scope of local variables is within a single method/thread. Thus, to identify the definition of a local variable, we simply traverse control-flow edges until the definition of a variable with the same name is found. Shared variables can be read/written in different threads. We thus traverse both control-flow and thread-control-flow edges, to find the definition of a variable with the same memory address; when a statement instance has outgoing edges of both kinds, a thread-control-flow edge takes preference over control-flow edges, to accurately represent the sequence of statement execution in the trace. For completeness, we add the identified data-flow-dependency edges to TDCG.

**Outputs.** At the end of this process, Slicer4J maps each statement instance in the produced slice back to its corresponding line of source code and outputs the list of source code lines. It also outputs the raw slice representation, where each line corresponds to a statement instance and has three fields: the name of the source file and the line number within the file for the statement instance; the thread of the instance; and the statement instance itself, in simple format (see Figure 3). Finally, Slicer4J also makes TDCG available for further inspection of the relevant control and data-flow dependencies.

Slicer4J is launched from the command line and provides several slicing options and configuration parameters, which are described in our tool package [4]. Information about hardware and software specification is also available online.

## 4 EVALUATION

We demonstrate the applicability of Slicer4J by running it on a number of subject applications.

**Benchmark Programs.** First, we borrowed three benchmark applications created by the JavaSlicer authors to evaluate their tool. These applications check the ability of a slicing solution to handle data-flows within a method (intra-procedural), data-flows across methods (inter-procedural), and control-dependence of exception blocks on the statements that throw exceptions.

We created four additional benchmark applications to verify capabilities introduced in Slicer4J and not supported by JavaSlicer: tracking data flows through multiple threads; Java 9 constructs, e.g., lambda expressions; native methods; and framework methods used by the JavaSlicer instrumentation. The list of the benchmarks, together with the size of each benchmark in terms of source lines of code (LoC), is given in the first two columns of Table 1. The benchmarks are small, ranging from 9 to 44 LoC.

As slices produced by JavaSlicer are at the Java bytecode level and contain both the program and framework code, for fair comparison between the tools, we map both slices to the source-code level and disregard statements within the Java framework code. The last three columns of Table 1 show the size of the expected slice, as well as the size of slices produced by JavaSlicer and Slicer4J.

The table shows that both tools produce the expected slice for the three JavaSlicer benchmarks, but JavaSlicer fails to produce the expected slice on the last four cases handled by Slicer4J due to its limitations outlined in Section 1. While Slicer4J finds the correct slice in all cases, it has an extra statement in the slice of the Native methods and Instrumentation classes benchmarks. In both cases, this is caused by an overapproximation in the framework method modeling; as this model is created by static analysis of the framework methods, it assumes that all possible execution paths...
within the framework may be taken at runtime, which is not the case for concrete executions.

**Defects4j Programs.** To further demonstrate the applicability of Slicer4j and compare it with JavaSlicer on bigger examples, we selected the top three open-source projects from the Defects4j dataset [12], ordering the projects in this dataset by the number of project usages on GitHub. The list of selected projects, together with their size, is shown in the first two columns of Table 2.

Each project in the Defects4j dataset contains a list of bugs, a test case that triggers each of these bugs, and an annotated fix for each bug. We examined the bugs one-by-one, selecting the first bug on the list where at least one buggy line is executed in the trace, i.e., where fixes involve modifying or replacing a line of code. We then used the failing test assertion statement as a slicing criterion and checked whether the buggy lines indeed appears in the slice produced by both Slicer4j and JavaSlicer. The number of buggy lines in the trace per project is shown in the third column of Table 2. The size of the execution trace for the test triggering the bug is in the fourth column, while the time to produce the trace, in seconds (without any instrumentation) is in the fifth column.

Next, we show the number of buggy lines found by each tool and the size of the slice that each tool produces. While in all three cases both JavaSlicer and Slicer4j found the expected buggy lines (these are older programs, which are written in Java 6 and contain no threads), the slices produced by Slicer4j are substantially smaller in the first two cases. That is mostly because of the extra unnecessary data-flow dependencies that JavaSlicer adds due to inaccurate modeling of bytecode instructions. For example, when looking for data-flow dependencies of y.f in the code y.f = x, the variable x should be added to the working set W. Yet, JavaSlicer adds y to W as well (and further collects all its transitive dependencies).

For the last subject, JavaSlicer and Slicer4j have identical slices besides one statement missing for JavaSlicer. This statement is a definition of a static field inside a static constructor, which is correctly identified by Slicer4j. JavaSlicer does not handle static fields in static constructors correctly, which we confirm with a mini-benchmark that we made available online [4].

Finally, we measured instrumentation (Inst.), execution (Exec.), and slicing times for Slicer4j for each subject. Unlike Slicer4j, which instruments the JAR file, JavaSlicer attaches to the JVM and instruments each class as it is loaded. We thus measure Inst.+Exec. and slicing times for JavaSlicer for each subject. All measured times, for both of the tools are shown in Table 2.

The combined instrumentation and execution time for Slicer4j is lower than that of JavaSlicer in the last two subject programs but is higher in the first one. This program is relatively large (more than 46,000 LoC) but the execution trace is small (only 3,927 lines), meaning that only a small portion of the program is executed. As JavaSlicer instruments on-demand, it saves a substantial portion of instrumentation, leading to a better performance in this case.

For the other two cases, the trace is substantially larger than the code size. In such cases, execution time "dominates" the instrumentation time and thus the efficient instrumentation capabilities of Slicer4j help to bring the overall processing time down. We also do not slow down framework code by instrumenting it.

The higher slicing time of Slicer4j is the cost of the Slicer4j ability to handle data-flows across threads. Unlike JavaSlicer, Slicer4j processes the entire trace to untangle threads and create the TDCG, which is used for more accurate slicing. The overall slicing time for Slicer4j is less than five minutes, which we still find acceptable.

**Limitations.** Slicer4j is unable to slice classes that are dynamically generated at runtime. We plan to address this limitation by adding a JVM instrumentation component that instruments such classes on-demand. Moreover, Slicer4j’s slicing time can be reduced by generating TDCG on-demand instead of pre-processing the entire trace before slicing starts. Slicer4j can map the Jimple slice to the source code only if the JAR file is compiled with debug information (preserving line numbers); otherwise, Slicer4j outputs the slice in the raw Jimple format. The interleaved writes to the same trace by multiple threads reduce parallelism as threads have to synchronize when writing to the trace, which can slow applications down. This synchronization can also "hide" concurrency bugs that occurred in the uninstrumented version of the program. Finally, Slicer4j relies on Soot to instrument the program and thus is limited to the Java 9 constructs supported by Soot.

5 CONCLUSIONS

This paper presents Slicer4j – an accurate and efficient tool for dynamic slicing of Java 9 programs. Slicer4j uses low-overhead instrumentation to collect a runtime execution trace. It then constructs a thread-aware, inter-procedural dynamic control-flow graph and uses the graph to compute the slice, while relying on pre-defined data-flow summaries of the main framework methods. We demonstrate the applicability of Slicer4j on ten benchmark and open-source Java programs, comparing it with JavaSlicer – the only other openly available slicing solution for Java 6 programs.

**Table 1: Accuracy: Benchmarks.**

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Slice size (LoC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>LoC</td>
</tr>
<tr>
<td>Intra-procedural</td>
<td>9</td>
</tr>
<tr>
<td>Inter-procedural</td>
<td>12</td>
</tr>
<tr>
<td>Exceptions</td>
<td>30</td>
</tr>
<tr>
<td>Multiple threads</td>
<td>44</td>
</tr>
<tr>
<td>Native methods</td>
<td>10</td>
</tr>
<tr>
<td>Java 9 constructs</td>
<td>17</td>
</tr>
<tr>
<td>Instrumentation classes</td>
<td>12</td>
</tr>
</tbody>
</table>

**Table 2: Defects4j Programs.**

<table>
<thead>
<tr>
<th>Name: Bug ID</th>
<th>LoC</th>
<th># Buggy lines</th>
<th>Trace (LoC)</th>
<th>Exec. time (s)</th>
<th># Buggy lines</th>
<th>Slice (LoC)</th>
<th>Time (s)</th>
<th># Buggy lines</th>
<th>Slice (LoC)</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JacksonDatabind3</td>
<td>46,091</td>
<td>1</td>
<td>3,927</td>
<td>0.23</td>
<td>1</td>
<td>367</td>
<td>18.6</td>
<td>3.7</td>
<td>1</td>
<td>38</td>
</tr>
<tr>
<td>Gson4</td>
<td>7,639</td>
<td>2</td>
<td>380,143</td>
<td>0.23</td>
<td>2</td>
<td>311</td>
<td>29.4</td>
<td>42.2</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>JacksonCore4</td>
<td>15,663</td>
<td>1</td>
<td>809,814</td>
<td>0.21</td>
<td>1</td>
<td>32</td>
<td>16.8</td>
<td>38.7</td>
<td>1</td>
<td>35</td>
</tr>
</tbody>
</table>
REFERENCES


