



THE UNIVERSITY
OF BRITISH COLUMBIA

MANDOLINE: Dynamic Slicing of Android Applications with Trace-Based Alias Analysis

ICST, 2021



Khaled Ahmed

khaled@ece.ubc.ca



Mieszko Lis

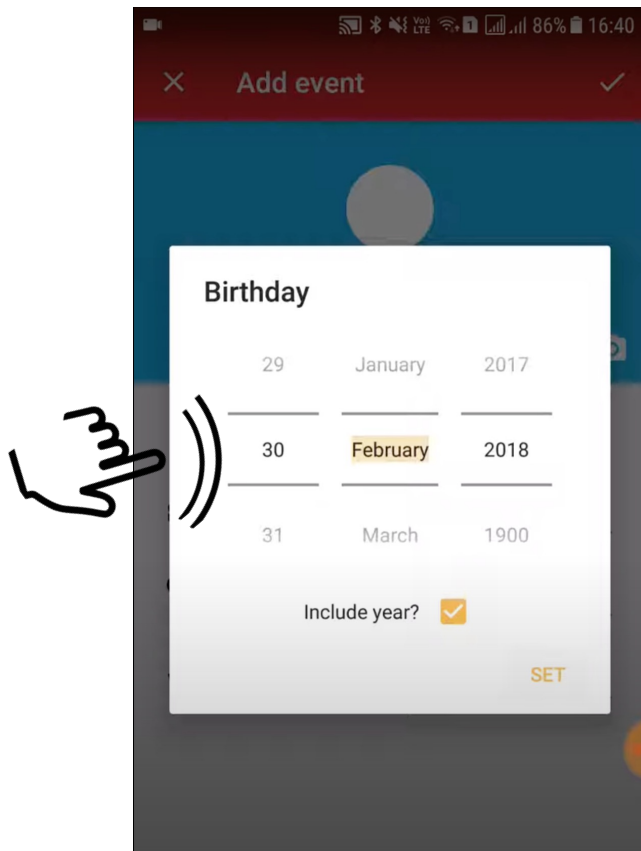
mieszko@ece.ubc.ca



Julia Rubin

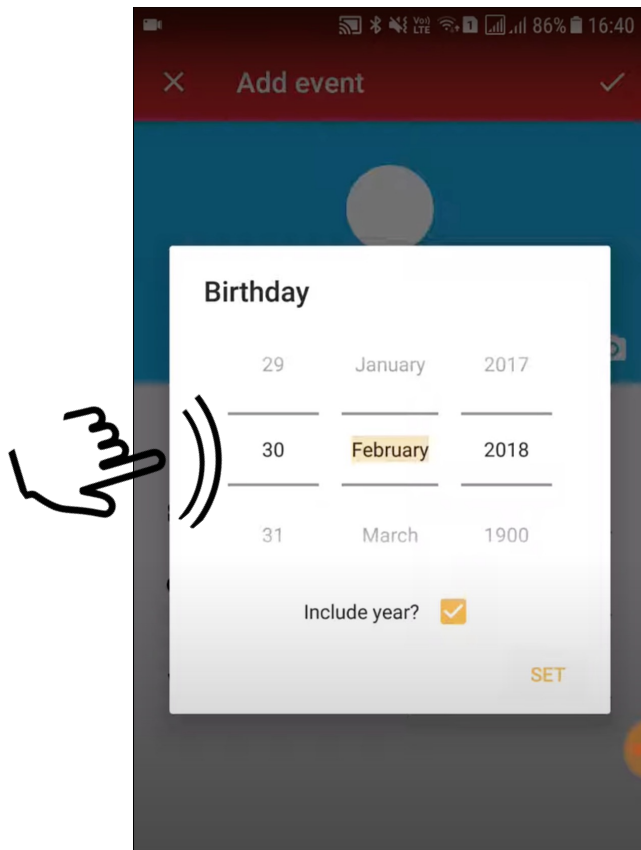
mjulia@ece.ubc.ca

Buggy Android App

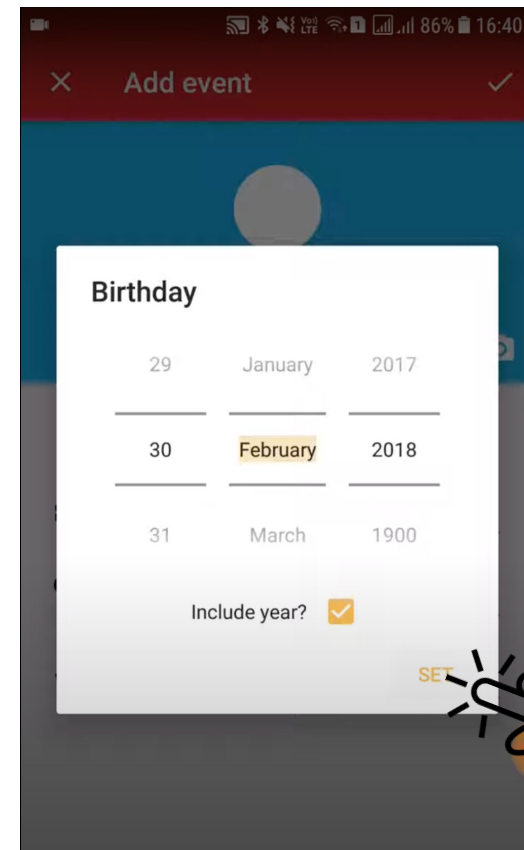


User selects monthly recurring event day

Buggy Android App



User selects monthly recurring event day



Crash! There's no February 30th!

Debugging



```
Date date = new Date(day,month,year);
```

Debugging

```
int day = userInput.d;
```


```
Date date = new Date(day,month,year);
```

A blue curved arrow originates from the variable 'day' in the second line of code and points back to the variable 'day' in the first line of code, illustrating the flow of data from the input to the date object creation.

Debugging

```
bar() {
```

```
    foo(selected);  
}  
foo(Object userInput) {  
    int day = userInput.d;  
    Date date = new Date(day,month,year);  
}
```

A blue curved arrow originates from the variable 'day' in the line 'int day = userInput.d;' and points to the 'day' parameter in the constructor call 'new Date(day,month,year);' in the line below it.

Debugging

```
bar() {
```

```
    selected.d = picker.getValue();
```

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

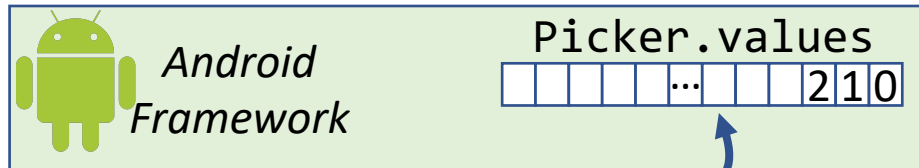
```
    int day = userInput.d;
```

```
    Date date = new Date(day, month, year);
```

```
}
```

Debugging

```
bar() {
```



```
    selected.d = picker.getValue();
```

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

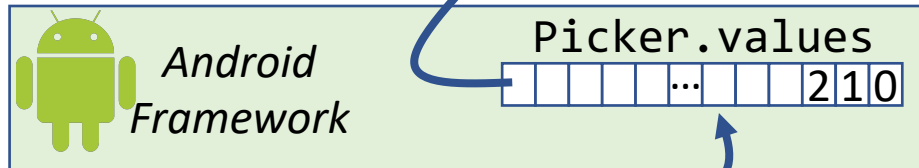
```
    Date date = new Date(day,month,year);
```

```
}
```


Debugging

```
bar() {
```

```
    picker.setMaxValue(31);
```



```
    selected.d = picker.getValue();
```

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

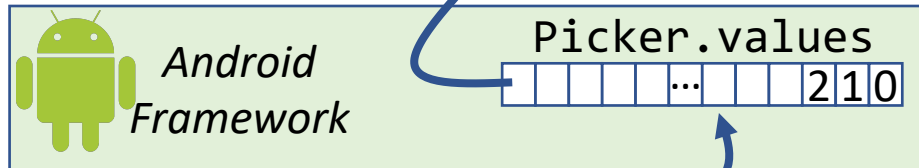
```
    Date date = new Date(day,month,year);
```

```
}
```

Debugging

```
bar() {
```

```
    picker.setMaxValue(31);
```



```
    selected.d = picker.getValue();
```

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

```
    Date date = new Date(day,month,year);
```

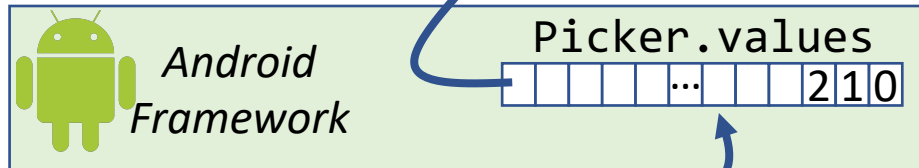
```
}
```

Automated by slicing
slice << code size

Debugging with Slicing

```
bar() {
```

```
    picker.setMaxValue(31);
```



```
    selected.d = picker.getValue();
```

```
    foo(selected);
```

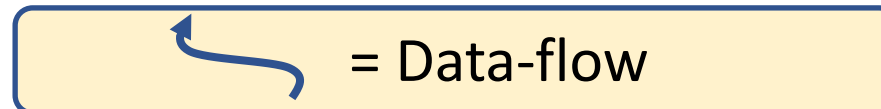
```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

```
    Date date = new Date(day,month,year);
```

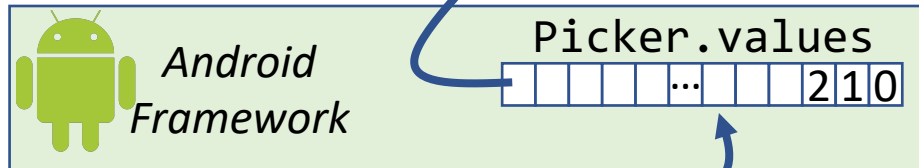
```
}
```



Challenge1: Fields

```
bar() {
```

```
    picker.setMaxValue(31);
```



```
    selected.d = picker.getValue();
```

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

```
    Date date = new Date(day,month,year);
```

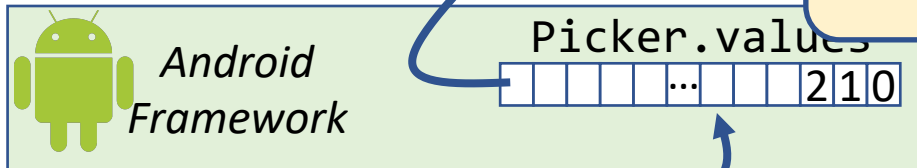
```
}
```

Challenge1: Fields

```
bar() {
```

```
    picker.setMaxValue(31);
```

Same field, different names



```
    selected.d = picker.getValue();
```

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

```
    Date date = new Date(day, month, year);
```

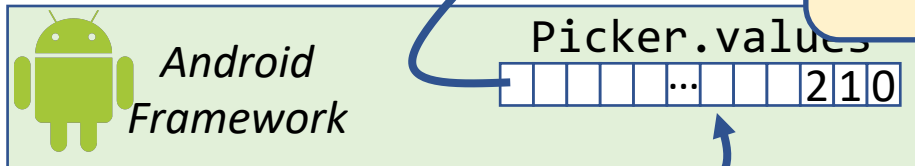
```
}
```

Challenge1: Fields

```
bar() {
```

```
    picker.setMaxValue(31);
```

Same field, different names



```
    selected.d = picker.getValue();
```

Simple solution:
record memory address
= slow for Android

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

```
    Date date = new Date(day,month,year);
```

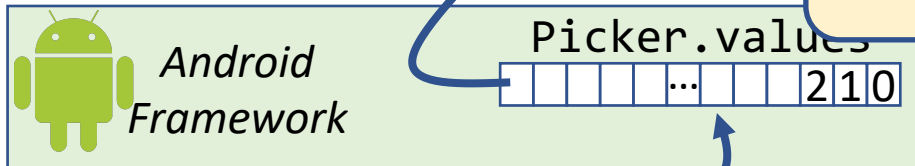
```
}
```

Challenge1: Fields

```
bar() {
```

```
    picker.setMaxValue(31);
```

Same field, different names



```
    selected.d = picker.getValue();
```

Simple solution:
record memory address
= slow for Android

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

Android kills slow apps

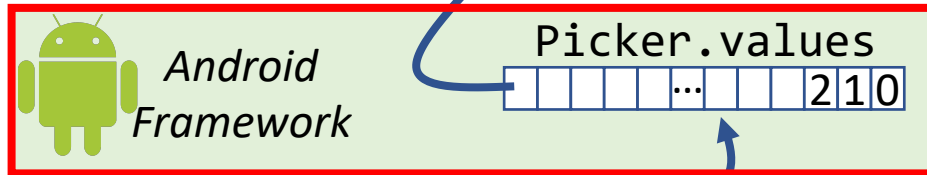
```
    Date date = new Date(day, month, year);
```

```
}
```

Challenge2: Framework

```
bar() {
```

```
    picker.setMaxValue(31);
```



```
    selected.d = picker.getValue();
```

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

```
    Date date = new Date(day,month,year);
```

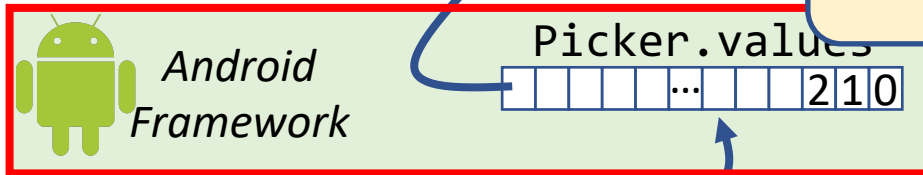
```
}
```


Challenge2: Framework

```
bar() {
```

```
    picker.setMaxValue(31);
```

Rely on framework methods



```
    selected.d = picker.getValue();
```

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

```
    Date date = new Date(day,month,year);
```

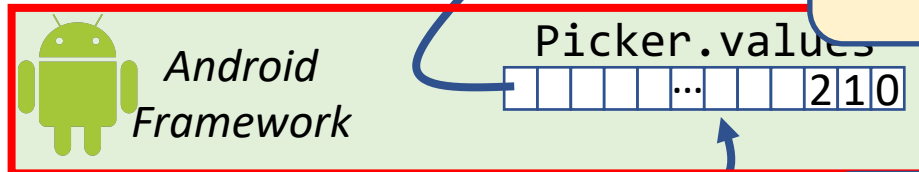
```
}
```

Challenge2: Framework

```
bar() {
```

```
    picker.setMaxValue(31);
```

Rely on framework methods



```
    selected.d = picker.getValue();
```

Simple solution:
Instrument entire framework

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

```
    Date date = new Date(day,month,year);
```

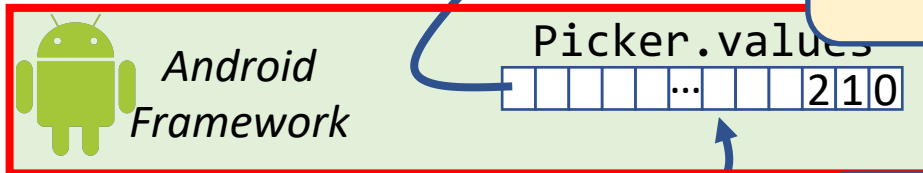
```
}
```

Challenge2: Framework

```
bar() {
```

```
    picker.setMaxValue(31);
```

Rely on framework methods



```
    selected.d = picker.getValue();
```

Simple solution:
Instrument entire framework

```
    foo(selected);
```

```
}
```

```
foo(Object userInput) {
```

```
    int day = userInput.d;
```

Not practical

```
    Date date = new Date(day,month,year);
```

```
}
```



Our Insight





Our Insight



selected.d
↑
userInput.d



Light-weight Instrumentation on-device
+
field data-flow analysis on the trace



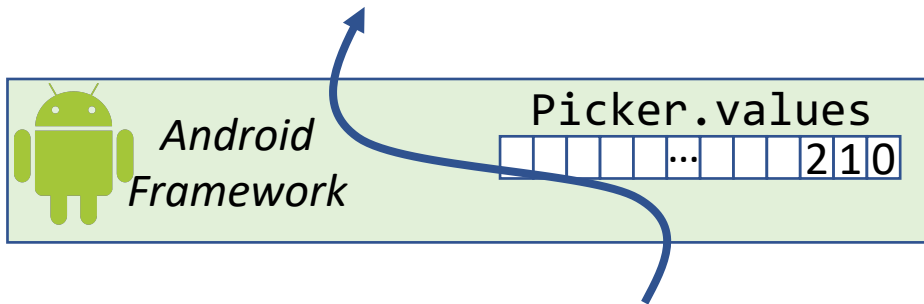
Our Insight



selected.d
↑
userInput.d



Light-weight Instrumentation on-device
+
field data-flow analysis on the trace



framework modeling

Trace-Based Field Analysis



Goal: Field data-flows w/o object addresses

selected.d



userInput.d

Trace-Based Field Analysis



Goal: Field data-flows w/o object addresses

- How to produce a trace with light-weight instrumentation?

selected.d
↑
userInput.d

Trace-Based Field Analysis



Goal: Field data-flows w/o object addresses

selected.d



userInput.d

- How to produce a trace with light-weight instrumentation?
- What information to use in the trace instead of object addresses?

Light-Weight Instrumentation



```
1  jane.age = 15
2  john      = jane
3  john.age  = 25
4  res = jane.age
```

Light-Weight Instrumentation



```
1  jane.age = 15
2  john      = jane
3  john.age  = 25
4  res = jane.age
```

Record executed statement IDs

Field Data-Flow Analysis



```
1  jane.age = 15
2  john      = jane
3  john.age  = 25
4  res = jane.age
```

Field Data-Flow Analysis




```
1  jane.age = 15
2  john      = jane
3  john.age  = 25
4  res = jane.age      use
```

Field Data-Flow Analysis



```
1  jane.age = 15
2  john      = jane
3  john.age  = 25
4  res = jane.age
```



use

Field Data-Flow Analysis



1 jane.age = 15 *definition?*

2 john = jane

3 john.age = 25

4 res = jane.age *use*

A blue arrow points upwards from the word 'use' next to line 4 to the word 'definition?' next to line 1, indicating a data flow from a use to a definition.

Field Data-Flow Analysis

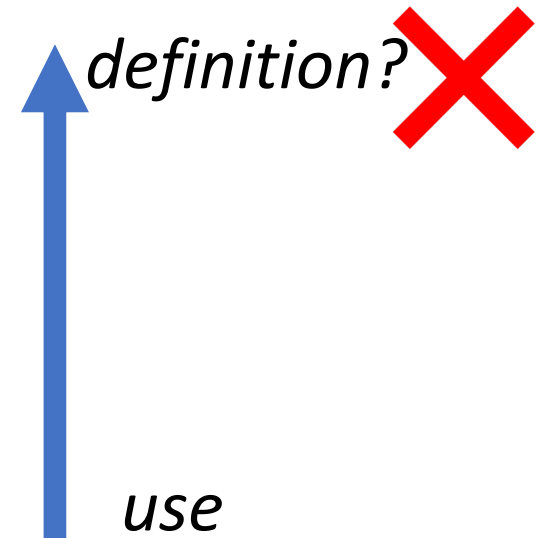


1 jane.age = 15

2 john = jane

3 john.age = 25

4 res = jane.age



Field Data-Flow Analysis



john aliases jane
john.age aliases jane.age

1 age = 15
2 john = jane

3 john.age = 25

4 res = jane.age

definition?



use

Field Data-Flow Analysis



john aliases jane
john.age aliases jane.age

age = 15

2 john = jane

definition



3 john.age = 25

4 res = jane.age

definition?



use

Field Alias Analysis

1 `jane.age = 15`

2 `john = jane`

3 `john.age = 25`

4 `res = jane.age`

Field Alias Analysis

```
1  jane.age = 15
2  john      = jane
3  john.age  = 25
4  res = jane.age      use
```

Field Alias Analysis

```

1  jane.age = 15
2  john     = jane
3  john.age = 25
4  res = jane.age

```

↑
use

Field Alias Analysis

1 jane.age = 15

2 john = jane

3 john.age = 25

4 res = jane.age



use

Field Alias Analysis

Search for `john.age` too

1 `jane.age = 15`

2 `john = jane`

3 `john.age = 25`

4 `res = jane.age`

use

Field Alias Analysis

Search for `john.age` too

```
1  jane.age = 15
```

```
2  john      = jane
```

```
3  john.age = 25
```

```
4  res = jane.age
```



use

Field Alias Analysis

Search for `john.age` too

```
1 jane.age = 15
```

```
2 john = jane
```

```
3 john.age = 25
```

```
4 res = jane.age
```

use

Field Alias Analysis

Search for `john.age` too

```
1 jane.age = 15
```

```
2 john = jane
```

```
3 john.age = 25
```

```
4 res = jane.age
```

definition ✓

use

Field Alias Analysis

Search for `john.age` too

1 `jane.age = 15`

2 `john = jane`

3 `john.age = 25`

4 `res = jane.age`

definition ✓

use

Backward + forward analysis on trace

Field Alias Analysis

```

1  Input: ICDG,  $t$ ,  $v$ 
   Output:  $T'$ 
2  begin
3     $V \leftarrow \{v\}$                                  $\triangleright$  Aliases of the variable  $v$ 
4     $S \leftarrow \text{BackwardAnalysis}(\text{ICDG}, V, t)$        $\triangleright$  Alias statements
5    return  $\text{LastDefined}(S, t)$                      $\triangleright$  Definitions of  $v$  and its fields

6  Procedure  $\text{BackwardAnalysis}(\text{ICDG}, V, t)$ 
7    begin
8       $S \leftarrow \emptyset$ 
9       $\triangleright$  Traverse the ICDG in inverse execution order
10     foreach  $t' \text{ s.t. } (t', t) \in \text{ICDG}$  do
11       if  $\text{method}(t) \neq \text{method}(t')$  then
12          $V \leftarrow \text{ChangeScope}(V, \text{method}(t), \text{method}(t'))$ 
13       if  $\exists v \in V \text{ s.t. } v \text{ and LHS}(t') \text{ have a common prefix}$  then
14          $\triangleright t'$  is a definition of a variable in  $V$ 
15          $S \leftarrow S \cup \{t'\}$                      $\triangleright$  Add the definition statement
16        $\langle V_b, V_f \rangle \leftarrow \text{AliasAnalysis}(\uparrow, V, t')$ 
17        $S \leftarrow S \cup \text{BackwardAnalysis}(\text{ICDG}, V_b, t')$ 
18        $S \leftarrow S \cup \text{ForwardAnalysis}(\text{ICDG}, V_f, t')$ 
19     return  $S$ 

18 Procedure  $\text{ForwardAnalysis}(\text{ICDG}, V, t)$ 
19   begin
20      $S \leftarrow \emptyset$ 
21      $\triangleright$  Traverse the ICDG in execution order
22     foreach  $t' \text{ s.t. } (t, t') \in \text{ICDG}$  do
23       if  $\text{method}(t) \neq \text{method}(t')$  then
24          $V \leftarrow \text{ChangeScope}(V, \text{method}(t), \text{method}(t'))$ 
25       if  $\exists v \in V \text{ s.t. } v \text{ is a prefix of LHS}(t')$  then
26          $\triangleright t'$  is a definition of a field of a variable in  $V$ 
27          $S \leftarrow S \cup \{t'\}$                      $\triangleright$  Add the definition statement
28        $\langle V_b, V_f \rangle \leftarrow \text{AliasAnalysis}(\downarrow, V, t')$ 
29        $S \leftarrow S \cup \text{BackwardAnalysis}(\text{ICDG}, V_b, t')$ 
30        $S \leftarrow S \cup \text{ForwardAnalysis}(\text{ICDG}, V_f, t')$ 
31     return  $S$ 

30 Procedure  $\text{AliasAnalysis}(d, V, t')$ 
31    $\triangleright$  Initialize with the original set of variables
32   if  $d = \uparrow$  then
33      $V_b \leftarrow V; V_f \leftarrow \emptyset$ 
34   else if  $d = \downarrow$  then
35      $V_b \leftarrow \emptyset; V_f \leftarrow V$ 
36   foreach  $v \in V \text{ s.t. } v \text{ and RHS}(t') \text{ have a common prefix}$  do
37      $\triangleright$  LHS( $t'$ ) is a new alias for  $v$ 
38      $V_f \leftarrow V_f \cup \text{ExtendFields}(\text{LHS}(t'))$      $\triangleright$  Follow it forward
39   foreach  $v \in V \text{ s.t. } v \text{ is a prefix of LHS}(t')$  do
40      $\triangleright t'$  is a re-definition of a field of  $v$ 
41      $\triangleright$  Follow the assigned variable both backward and forward
42      $V_b \leftarrow V_b \cup \text{RHS}(t')$ 
43      $V_f \leftarrow V_f \cup \text{RHS}(t')$ 
44   foreach  $v \in V \text{ s.t. LHS}(t') \text{ is a prefix or equal to } v$  do
45      $\triangleright t'$  is a full re-definition of  $v$ 
46     if  $d = \uparrow$  then
47        $V_b \leftarrow V_b \setminus \{v\}$                  $\triangleright$  Do not search before the definition
48        $\triangleright$  Follow the assigned variable both backward and forward
49        $V_b \leftarrow V_b \cup \text{ExtendFields}(\text{RHS}(t'))$ 
50        $V_f \leftarrow V_f \cup \text{ExtendFields}(\text{RHS}(t'))$ 
51     else if  $d = \downarrow$  then
52        $V_f \leftarrow V_f \setminus \{v\}$                  $\triangleright$  Do not search for new variables
53      $V_b \leftarrow V_b \cup \text{ExtendFields}(\text{RHS}(t'))$ 
54      $V_f \leftarrow V_f \cup \text{ExtendFields}(\text{RHS}(t'))$ 
55   return  $\langle V_b, V_f \rangle$ 

```

Field Alias Analysis

```

1  Input: ICDG,  $t$ ,  $v$ 
   Output:  $T'$ 
2  begin
3       $V \leftarrow \{v\}$                                  $\triangleright$  Aliases of the variable  $v$ 
4       $S \leftarrow \text{BackwardAnalysis}(\text{ICDG}, V, t)$      $\triangleright$  Alias statements
5      return  $\text{LastDefined}(S, t)$                      $\triangleright$  Definitions of  $v$  and its fields

6  Procedure  $\text{BackwardAnalysis}(\text{ICDG}, V, t)$ 
7      begin
8           $S \leftarrow \emptyset$ 
9           $\triangleright$  Traverse the ICDG in inverse execution order
10         foreach  $t' \text{ s.t. } (t', t) \in \text{ICDG}$  do
11             if  $\text{method}(t) \neq \text{method}(t')$  then
12                  $V \leftarrow \text{ChangeScope}(V, \text{method}(t), \text{method}(t'))$ 
13             if  $\exists v \in V \text{ s.t. } v \text{ and LHS}(t') \text{ have a common prefix}$  then
14                  $\triangleright t'$  is a definition of a variable in  $V$ 
15                  $S \leftarrow S \cup \{t'\}$                  $\triangleright$  Add the definition statement
16              $\langle V_b, V_f \rangle \leftarrow \text{AliasAnalysis}(\uparrow, V, t')$ 
17              $S \leftarrow S \cup \text{BackwardAnalysis}(\text{ICDG}, V_b, t')$ 
18              $S \leftarrow S \cup \text{ForwardAnalysis}(\text{ICDG}, V_f, t')$ 
19         return  $S$ 

18 Procedure  $\text{ForwardAnalysis}(\text{ICDG}, V, t)$ 
19     begin
20          $S \leftarrow \emptyset$ 
21          $\triangleright$  Traverse the ICDG in execution order
22         foreach  $t' \text{ s.t. } (t, t') \in \text{ICDG}$  do
23             if  $\text{method}(t) \neq \text{method}(t')$  then
24                  $V \leftarrow \text{ChangeScope}(V, \text{method}(t), \text{method}(t'))$ 
25             if  $\exists v \in V \text{ s.t. } v \text{ is a prefix of LHS}(t')$  then
26                  $\triangleright t'$  is a definition of a field of a variable in  $V$ 
27                  $S \leftarrow S \cup \{t'\}$                  $\triangleright$  Add the definition statement
28              $\langle V_b, V_f \rangle \leftarrow \text{AliasAnalysis}(\downarrow, V, t')$ 
29              $S \leftarrow S \cup \text{BackwardAnalysis}(\text{ICDG}, V_b, t')$ 
30              $S \leftarrow S \cup \text{ForwardAnalysis}(\text{ICDG}, V_f, t')$ 
31         return  $S$ 

```

```

30 Procedure  $\text{AliasAnalysis}(d, V, t')$ 
31      $\triangleright$  Initialize with the original set of variables
32     if  $d = \uparrow$  then
33          $V_b \leftarrow V; V_f \leftarrow \emptyset$ 
34     else if  $d = \downarrow$  then
35          $V_b \leftarrow \emptyset; V_f \leftarrow V$ 
36     foreach  $v \in V \text{ s.t. } v \text{ and RHS}(t') \text{ have a common prefix}$  do
37          $\triangleright$  LHS( $t'$ ) is a new alias for  $v$ 
38          $V_f \leftarrow V_f \cup \text{ExtendFields}(\text{LHS}(t'))$      $\triangleright$  Follow it forward
39     foreach  $v \in V \text{ s.t. } v \text{ is a prefix of LHS}(t')$  do
40          $\triangleright t'$  is a re-definition of a field of  $v$ 
41          $\triangleright$  Follow the assigned variable both backward and forward
42          $V_b \leftarrow V_b \cup \text{RHS}(t')$ 
43          $V_f \leftarrow V_f \cup \text{RHS}(t')$ 
44     foreach  $v \in V \text{ s.t. LHS}(t') \text{ is a prefix or equal to } v$  do
45          $\triangleright t'$  is a full re-definition of  $v$ 
46         if  $d = \uparrow$  then
47              $V_b \leftarrow V_b \setminus \{v\}$                  $\triangleright$  Do not search before the definition
48              $\triangleright$  Follow the assigned variable both backward and forward
49              $V_b \leftarrow V_b \cup \text{ExtendFields}(\text{RHS}(t'))$ 
50              $V_f \leftarrow V_f \cup \text{ExtendFields}(\text{RHS}(t'))$ 
51         else if  $d = \downarrow$  then
52              $V_f \leftarrow V_f \setminus \{v\}$                  $\triangleright$  Do not search for new variables
53     return  $\langle V_b, V_f \rangle$ 

```

Search for field
definitions backward

Field Alias Analysis

```

1  Input: ICDG,  $t$ ,  $v$ 
   Output:  $T'$ 
2  begin
3     $V \leftarrow \{v\}$                                  $\triangleright$  Aliases of the variable  $v$ 
4     $S \leftarrow \text{BackwardAnalysis}(\text{ICDG}, V, t)$        $\triangleright$  Alias statements
5    return  $\text{LastDefined}(S, t)$                      $\triangleright$  Definitions of  $v$  and its fields
6  Procedure  $\text{BackwardAnalysis}(\text{ICDG}, V, t)$ 
7    begin
8       $S \leftarrow \emptyset$ 
9       $\triangleright$  Traverse the ICDG in inverse execution order
10     foreach  $t' \text{ s.t. } (t', t) \in \text{ICDG}$  do
11       if  $\text{method}(t) \neq \text{method}(t')$  then
12          $V \leftarrow \text{ChangeScope}(V, \text{method}(t), \text{method}(t'))$ 
13       if  $\exists v \in V \text{ s.t. } v \text{ and LHS}(t') \text{ have a common prefix}$  then
14          $\triangleright t'$  is a definition of a variable in  $V$ 
15          $S \leftarrow S \cup \{t'\}$                      $\triangleright$  Add the definition statement
16        $\langle V_b, V_f \rangle \leftarrow \text{AliasAnalysis}(\uparrow, V, t')$ 
17        $S \leftarrow S \cup \text{BackwardAnalysis}(\text{ICDG}, V_b, t')$ 
18        $S \leftarrow S \cup \text{ForwardAnalysis}(\text{ICDG}, V_f, t')$ 
19     return  $S$ 
18 Procedure  $\text{ForwardAnalysis}(\text{ICDG}, V, t)$ 
19   begin
20      $S \leftarrow \emptyset$ 
21      $\triangleright$  Traverse the ICDG in execution order
22     foreach  $t' \text{ s.t. } (t, t') \in \text{ICDG}$  do
23       if  $\text{method}(t) \neq \text{method}(t')$  then
24          $V \leftarrow \text{ChangeScope}(V, \text{method}(t), \text{method}(t'))$ 
25       if  $\exists v \in V \text{ s.t. } v \text{ is a prefix of LHS}(t')$  then
26          $\triangleright t'$  is a definition of a field of a variable in  $V$ 
27          $S \leftarrow S \cup \{t'\}$                      $\triangleright$  Add the definition statement
28        $\langle V_b, V_f \rangle \leftarrow \text{AliasAnalysis}(\downarrow, V, t')$ 
29        $S \leftarrow S \cup \text{BackwardAnalysis}(\text{ICDG}, V_b, t')$ 
30        $S \leftarrow S \cup \text{ForwardAnalysis}(\text{ICDG}, V_f, t')$ 
31     return  $S$ 

```

```

30 Procedure  $\text{AliasAnalysis}(d, V, t')$ 
31    $\triangleright$  Initialize with the original set of variables
32   if  $d = \uparrow$  then
33      $V_b \leftarrow V; V_f \leftarrow \emptyset$ 
34   else if  $d = \downarrow$  then
35      $V_b \leftarrow \emptyset; V_f \leftarrow V$ 
36   foreach  $v \in V \text{ s.t. } v \text{ and RHS}(t') \text{ have a common prefix}$  do
37      $\triangleright$  LHS( $t'$ ) is a new alias for  $v$ 
38      $V_f \leftarrow V_f \cup \text{ExtendFields}(\text{LHS}(t'))$        $\triangleright$  Follow it forward
39   foreach  $v \in V \text{ s.t. } v \text{ is a prefix of LHS}(t')$  do
40      $\triangleright t'$  is a re-definition of a field of  $v$ 
41      $\triangleright$  Follow the assigned variable both backward and forward
42      $V_b \leftarrow V_b \cup \text{RHS}(t')$ 
43      $V_f \leftarrow V_f \cup \text{RHS}(t')$ 
44   foreach  $v \in V \text{ s.t. LHS}(t') \text{ is a prefix or equal to } v$  do
45      $\triangleright t'$  is a full re-definition of  $v$ 
46     if  $d = \uparrow$  then
47        $V_b \leftarrow V_b \setminus \{v\}$                  $\triangleright$  Do not search before the definition
48        $\triangleright$  Follow the assigned variable both backward and forward
49        $V_b \leftarrow V_b \cup \text{ExtendFields}(\text{RHS}(t'))$ 
50        $V_f \leftarrow V_f \cup \text{ExtendFields}(\text{RHS}(t'))$ 
51     else if  $d = \downarrow$  then
52        $V_f \leftarrow V_f \setminus \{v\}$                  $\triangleright$  Do not search for new variables
53   return  $\langle V_b, V_f \rangle$ 

```

Search for field
definitions forward

Field Alias Analysis

```

1  Input: ICDG,  $t$ ,  $v$ 
   Output:  $T'$ 
2  begin
3     $V \leftarrow \{v\}$                                  $\triangleright$  Aliases of the variable  $v$ 
4     $S \leftarrow \text{BackwardAnalysis}(\text{ICDG}, V, t)$        $\triangleright$  Alias statements
5    return  $\text{LastDefined}(S, t)$                      $\triangleright$  Definitions of  $v$  and its fields
6  Procedure  $\text{BackwardAnalysis}(\text{ICDG}, V, t)$ 
7    begin
8       $S \leftarrow \emptyset$ 
9       $\triangleright$  Traverse the ICDG in inverse execution order
10     foreach  $t' \text{ s.t. } (t', t) \in \text{ICDG}$  do
11       if  $\text{method}(t) \neq \text{method}(t')$  then
12          $V \leftarrow \text{ChangeScope}(V, \text{method}(t), \text{method}(t'))$ 
13       if  $\exists v \in V \text{ s.t. } v \text{ and LHS}(t') \text{ have a common prefix}$  then
14          $\triangleright t'$  is a definition of a variable in  $V$ 
15          $S \leftarrow S \cup \{t'\}$                      $\triangleright$  Add the definition statement
16        $\langle V_b, V_f \rangle \leftarrow \text{AliasAnalysis}(\uparrow, V, t')$ 
17        $S \leftarrow S \cup \text{BackwardAnalysis}(\text{ICDG}, V_b, t')$ 
18        $S \leftarrow S \cup \text{ForwardAnalysis}(\text{ICDG}, V_f, t')$ 
19     return  $S$ 
20  Procedure  $\text{ForwardAnalysis}(\text{ICDG}, V, t)$ 
21    begin
22       $S \leftarrow \emptyset$ 
23       $\triangleright$  Traverse the ICDG in execution order
24      foreach  $t' \text{ s.t. } (t, t') \in \text{ICDG}$  do
25        if  $\text{method}(t) \neq \text{method}(t')$  then
26           $V \leftarrow \text{ChangeScope}(V, \text{method}(t), \text{method}(t'))$ 
27        if  $\exists v \in V \text{ s.t. } v \text{ is a prefix of LHS}(t')$  then
28           $\triangleright t'$  is a definition of a field of a variable in  $V$ 
29           $S \leftarrow S \cup \{t'\}$                      $\triangleright$  Add the definition statement
30         $\langle V_b, V_f \rangle \leftarrow \text{AliasAnalysis}(\downarrow, V, t')$ 
31         $S \leftarrow S \cup \text{BackwardAnalysis}(\text{ICDG}, V_b, t')$ 
32         $S \leftarrow S \cup \text{ForwardAnalysis}(\text{ICDG}, V_f, t')$ 
33      return  $S$ 

```

```

30 Procedure  $\text{AliasAnalysis}(d, V, t')$ 
    $\triangleright$  Initialize with the original set of variables
31 if  $d = \uparrow$  then
32    $V_b \leftarrow V; V_f \leftarrow \emptyset$ 
33 else if  $d = \downarrow$  then
34    $V_b \leftarrow \emptyset; V_f \leftarrow V$ 
35 foreach  $v \in V \text{ s.t. } v \text{ and RHS}(t') \text{ have a common prefix}$  do
36    $\triangleright$  LHS( $t'$ ) is a new alias for  $v$ 
37    $V_f \leftarrow V_f \cup \text{ExtendFields}(\text{LHS}(t'))$        $\triangleright$  Follow it forward
38 foreach  $v \in V \text{ s.t. } v \text{ is a prefix of LHS}(t')$  do
39    $\triangleright t'$  is a re-definition of a field of  $v$ 
40    $\triangleright$  Follow the assigned variable both backward and forward
41    $V_b \leftarrow V_b \cup \text{RHS}(t')$ 
42    $V_f \leftarrow V_f \cup \text{RHS}(t')$ 
43 foreach  $v \in V \text{ s.t. LHS}(t') \text{ is a prefix or equal to } v$  do
44    $\triangleright t'$  is a full re-definition of  $v$ 
45   if  $d = \uparrow$  then
46      $V_b \leftarrow V_b \setminus \{v\}$                      $\triangleright$  Do not search before the definition
47      $\triangleright$  Follow the assigned variable both backward and forward
48      $V_b \leftarrow V_b \cup \text{ExtendFields}(\text{RHS}(t'))$ 
49      $V_f \leftarrow V_f \cup \text{ExtendFields}(\text{RHS}(t'))$ 
50   else if  $d = \downarrow$  then
51      $V_f \leftarrow V_f \setminus \{v\}$                      $\triangleright$  Do not search for new variables
52   return  $\langle V_b, V_f \rangle$ 

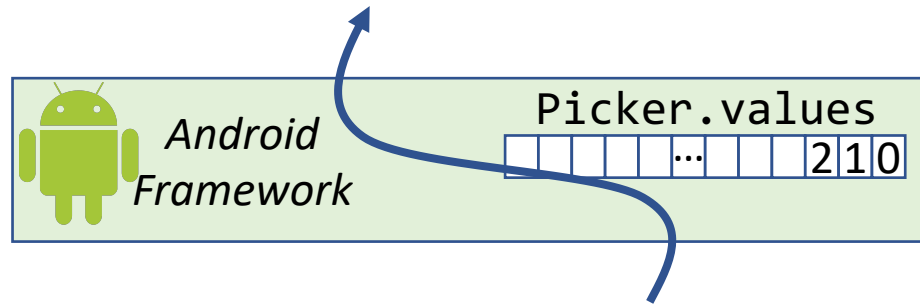
```

Decide on variables to search for and the direction

Framework Data-Flows



Goal: Data-flows inside the Framework



How to find data-flows inside framework w/o instrumentation?

Taint-Propagation Models

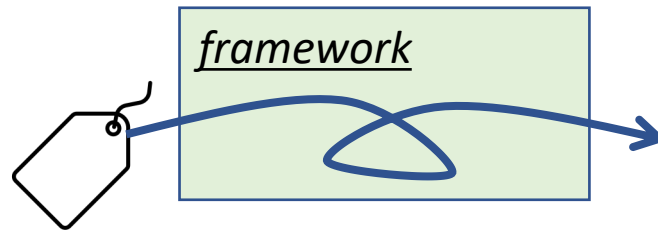


StubDroid [1] & FlowDroid [2]

Taint-Propagation Models



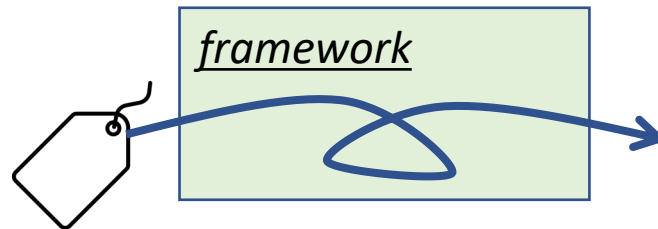
StubDroid [1] & FlowDroid [2]



Taint-Propagation Models



StubDroid [1] & FlowDroid [2]

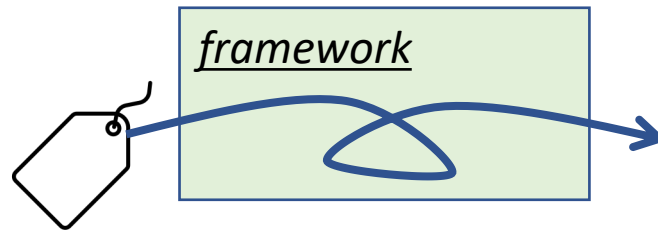


Express how taint propagates

Taint-Propagation Models



StubDroid [1] & FlowDroid [2]



Express how taint propagates

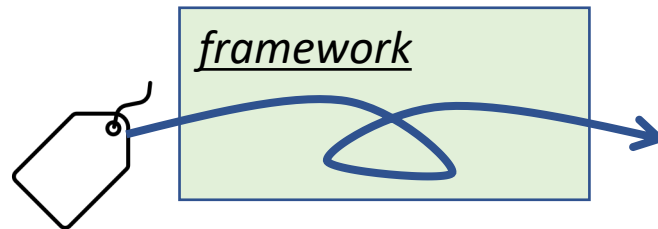
From:

- Parameters
 - Receiver
- } + Fields

Taint-Propagation Models



StubDroid [1] & FlowDroid [2]



Express how taint propagates

From:

- Parameters
 - Receiver
- } + Fields

To:

- Parameters
 - Receiver
 - Return
- } + Fields

Taint-Propagation Models



```
Array::set(index, value) {  
    this.inA[index] = value;  
}
```

Taint-Propagation Models



Taint-propagation:

```
Array::set(index, value) {  
  this.inA[index] = value;  
}
```

value  this.inA

Taint-Propagation Models



Taint-propagation:

```
Array::set(index, value) {  
  this.inA[index] = value;  
}
```

value  this.inA

index 

this 

Taint-Propagation Models



Taint-propagation:

```
Array::set(index, value) {  
  this.inA[index] = value;  
}
```

value  this.inA

index 

this 

```
Log::info(value) {  
  print(value);  
}
```

Taint-Propagation Models



Taint-propagation:

```
Array::set(index, value) {  
  this.inA[index] = value;  
}
```

value  this.inA

index 

this 

```
Log::info(value) {  
  print(value);  
}
```

value 

Taint-Propagation Models



Taint-propagation:

```
Array::set(index, value) {  
  this.inA[index] = value;  
}
```

value  this.inA

index 

this 

```
Log::info(value) {  
  print(value);  
}
```

value 

**Accurate taint propagation rules
for Java and Android methods**

Modeling w/ Implicit Assignments



value  this.inA

Modeling w/ Implicit Assignemnts



value  this.inA

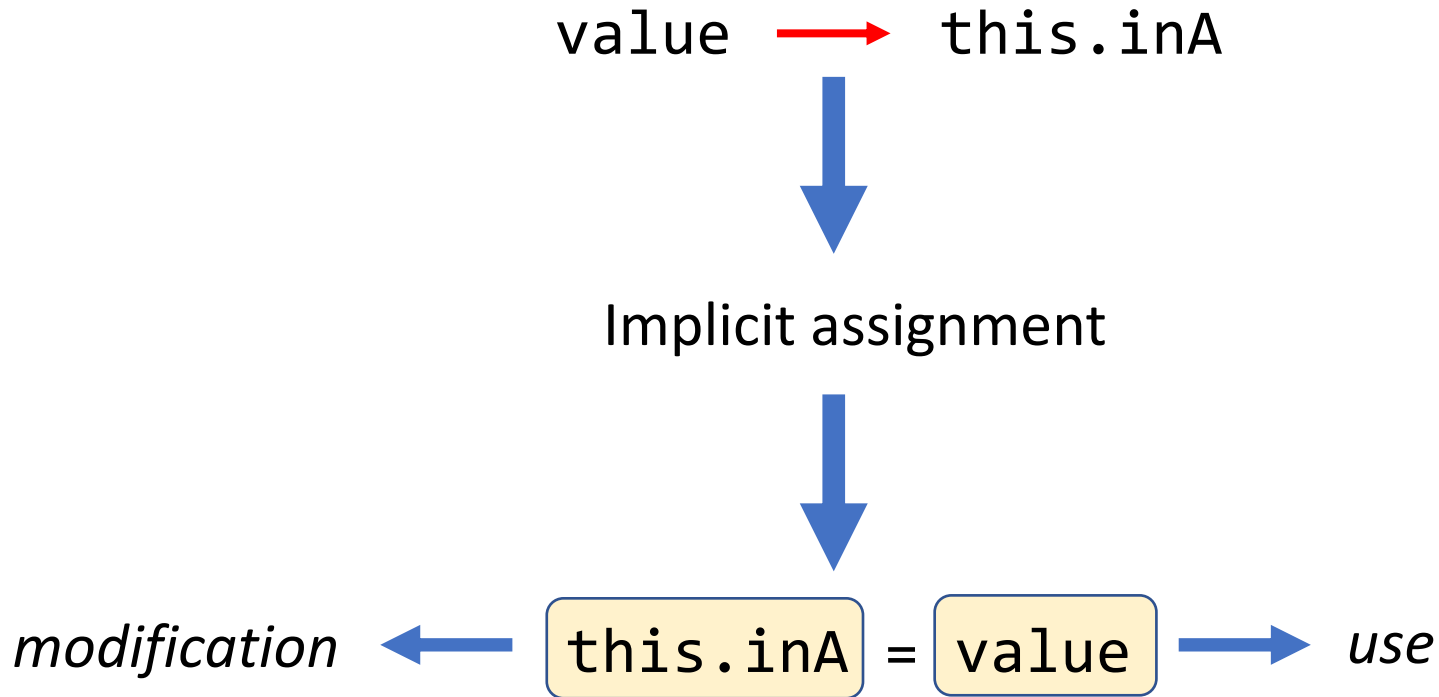


Implicit assignment



this.inA = value

Modeling w/ Implicit Assignemnts



Framework Modeling: Example



Code

```
x = 5
```

```
array.set(1, x)
```

```
y = array.get(1)
```

Model

```
y = array.inA
```

Framework Modeling: Example



Code

```
x = 5
```

```
array.set(1, x)
```

```
y = array.get(1)
```



Model

```
y = array.inA
```


Framework Modeling: Example



Code

```
x = 5
```

```
array.set(1, x)
```

```
y = array.get(1)
```



Model

```
array.inA = x
```

```
y = array.inA
```

Framework Modeling: Example



Code

```
x = 5
```

```
array.set(1, x)
```

```
y = array.get(1)
```



Model

```
array.inA = x
```

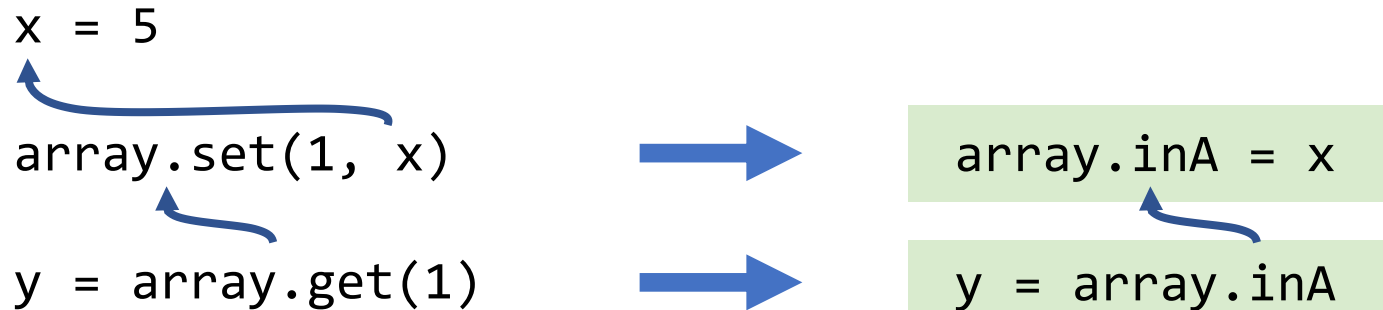
```
y = array.inA
```

Framework Modeling: Example



Code

```
x = 5  
array.set(1, x)  
y = array.get(1)
```

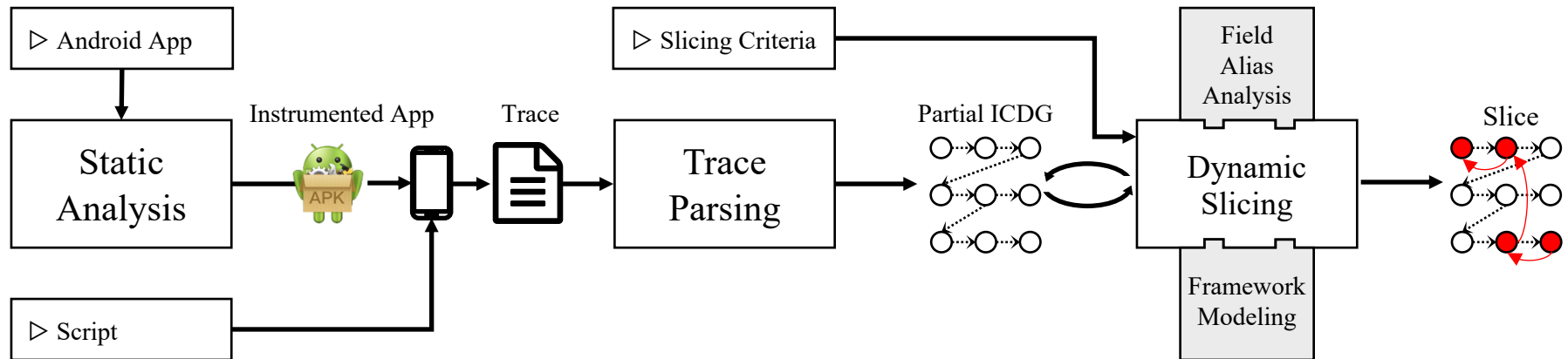


Model

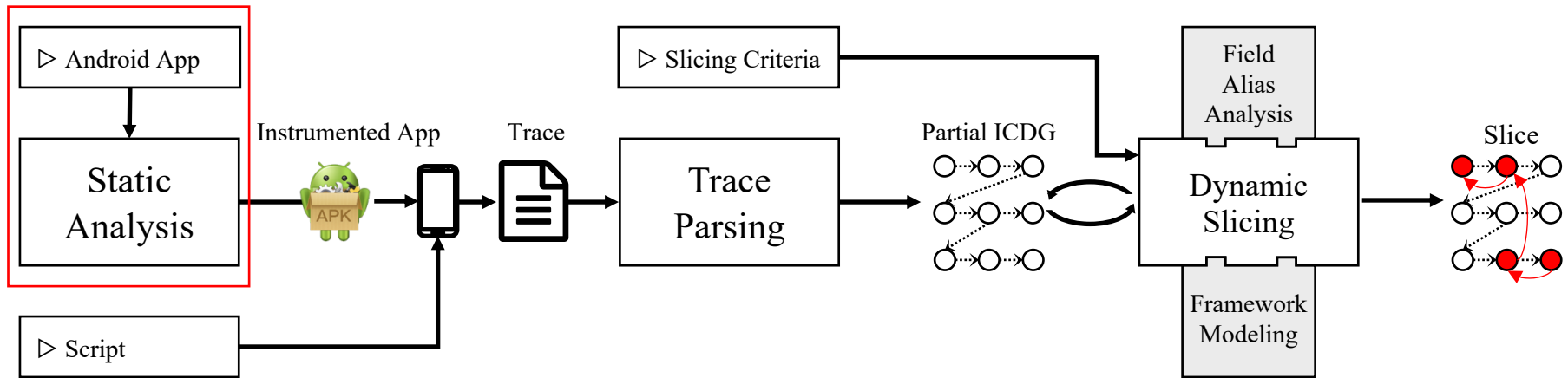
```
array.inA = x
```

```
y = array.inA
```

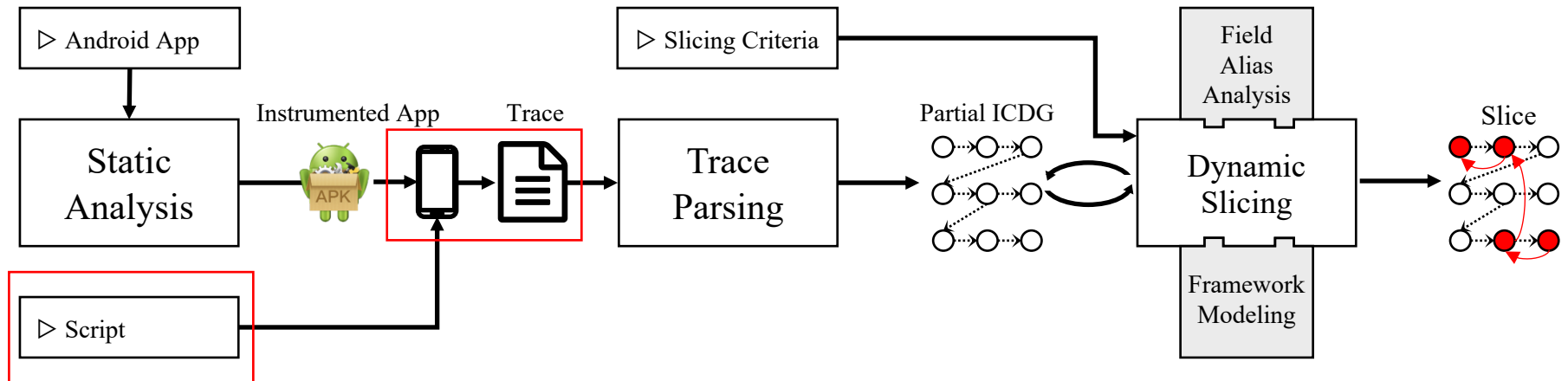
Our Solution: MANDOLINE



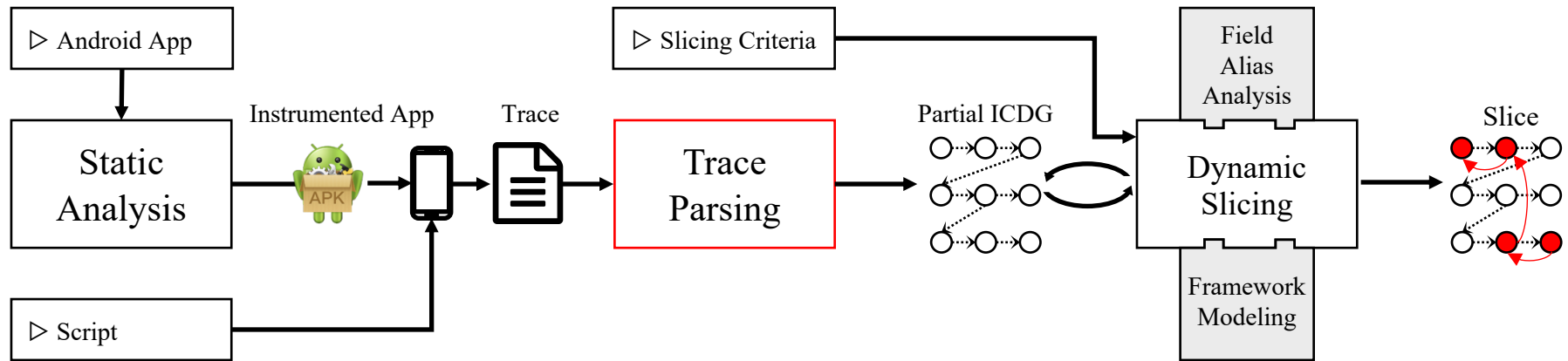
Our Solution: MANDOLINE



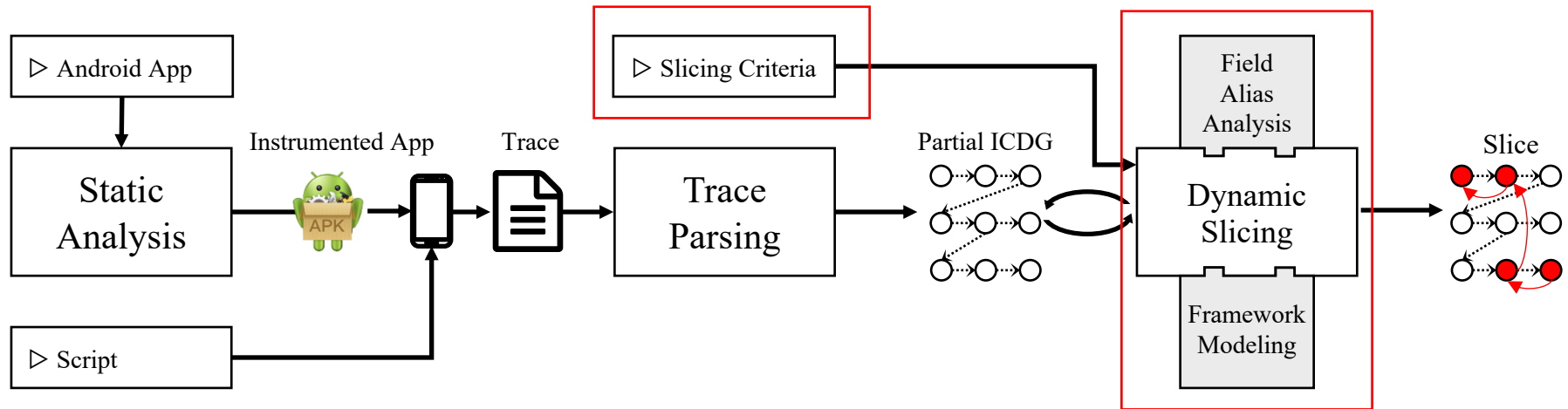
Our Solution: MANDOLINE



Our Solution: MANDOLINE



Our Solution: MANDOLINE



Evaluation



RQ1: Accuracy

vs. state-of-the-art

RQ2: Performance

Subjects



12 Apps from DroixBench [3] and ReCDroid [4]

- Faulty apps w/ crashes
- Reproducible crashes
- Crashes = slicing criteria

Baseline



AndroidSlicer [5] (state-of-the-art)

- First slicer for Android apps
- Inter-callback data-flows
- Misses data-flows in fields and framework

AndroidSlicer++:

- Fixed implementation issues
- Improved instrumentation
- Allows fair comparison

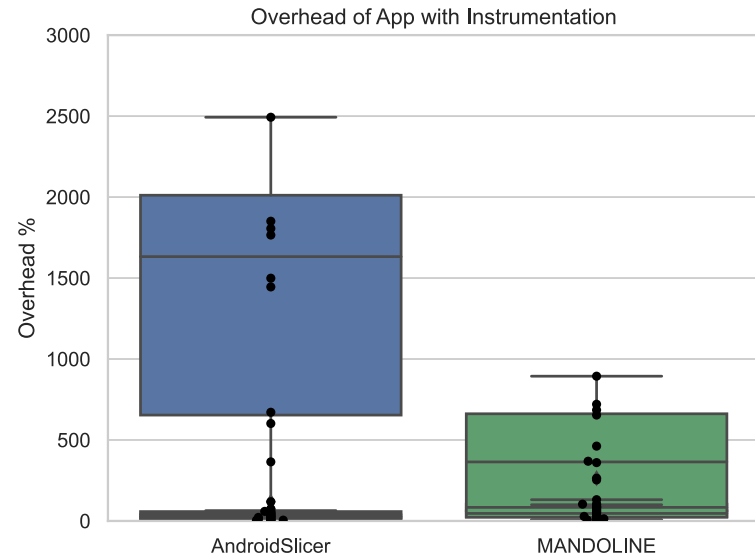
Benchmark



Manually produced slices

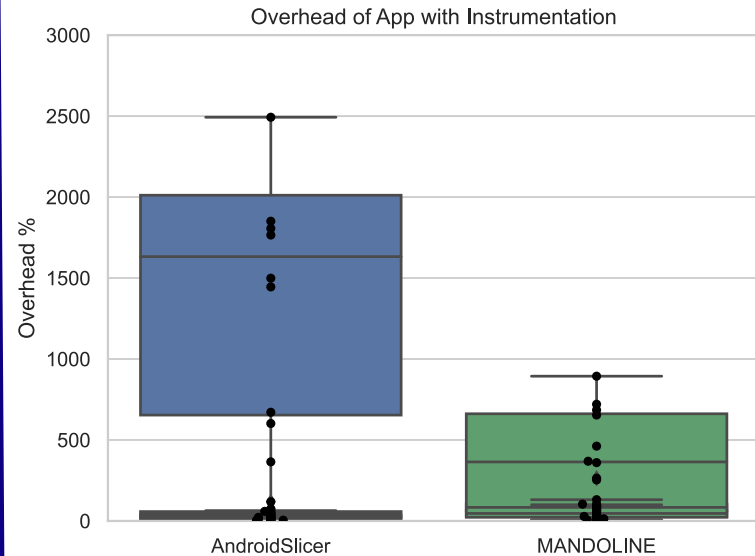
- Cross validated by two researchers
- First Android slices benchmark
- State-of-the-art [5] only compared slice to trace sizes

Accuracy and Performance



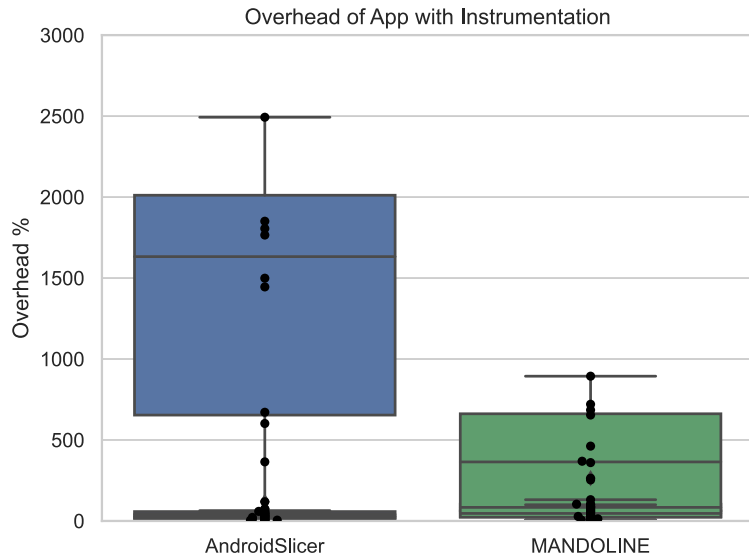
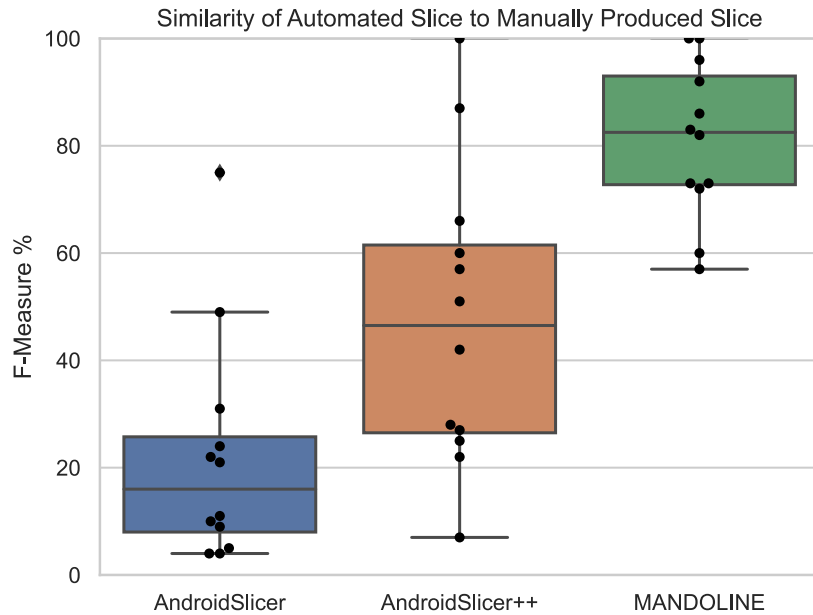
Accuracy and Performance

RQ1



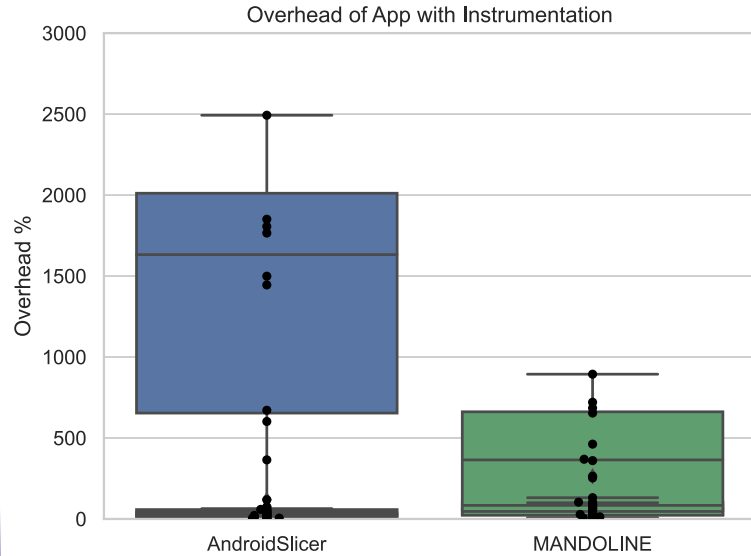
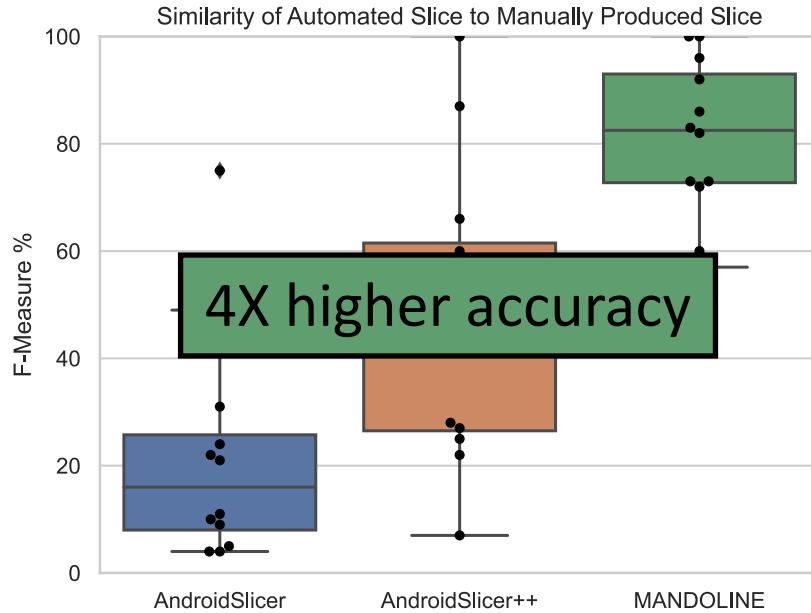
Accuracy and Performance

RQ1



Accuracy and Performance

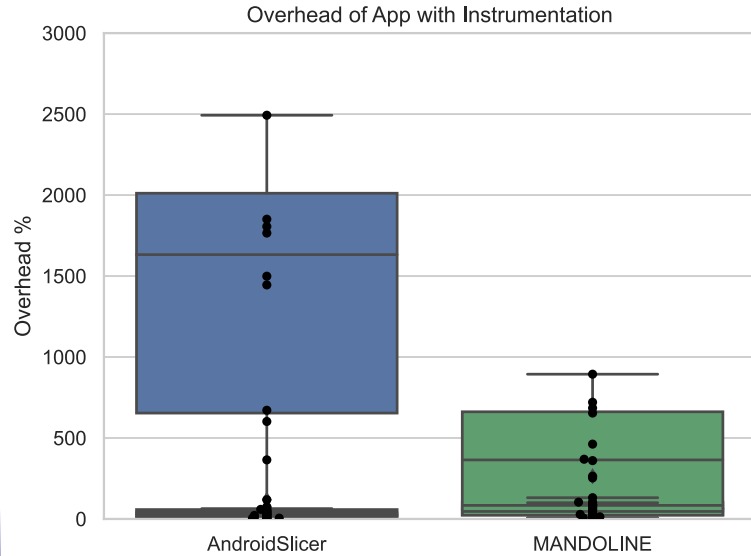
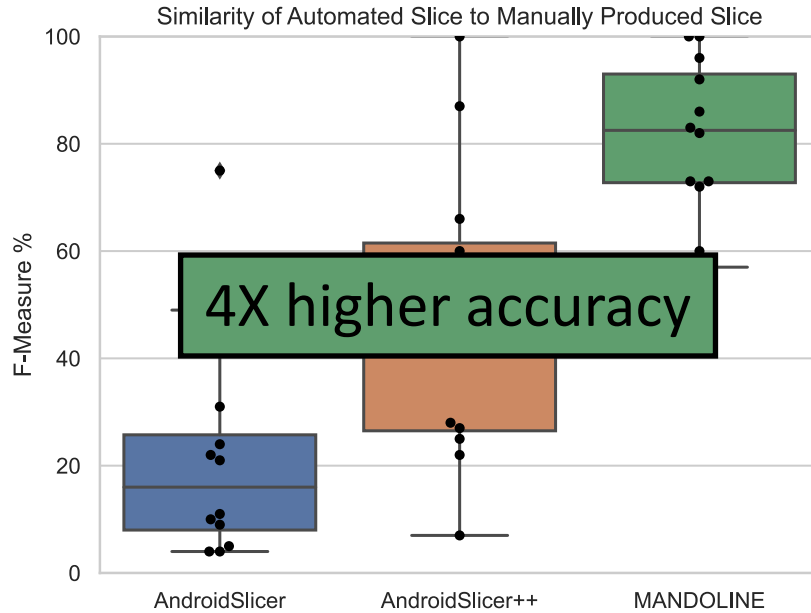
RQ1



Accuracy and Performance

RQ1

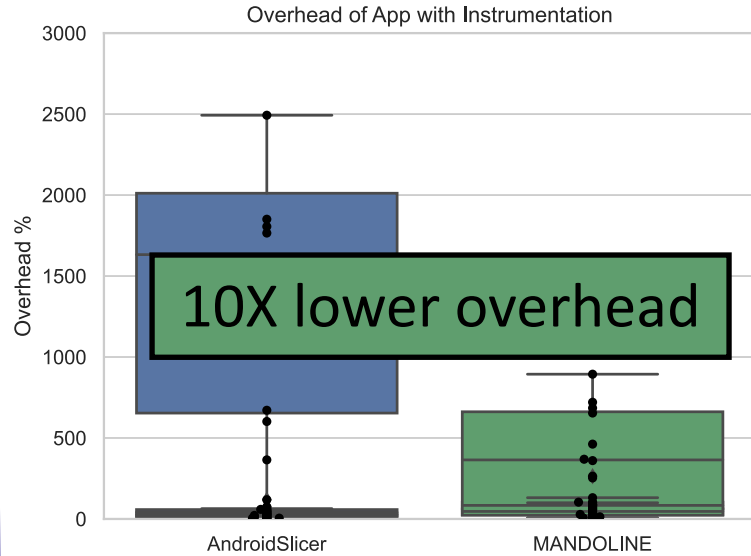
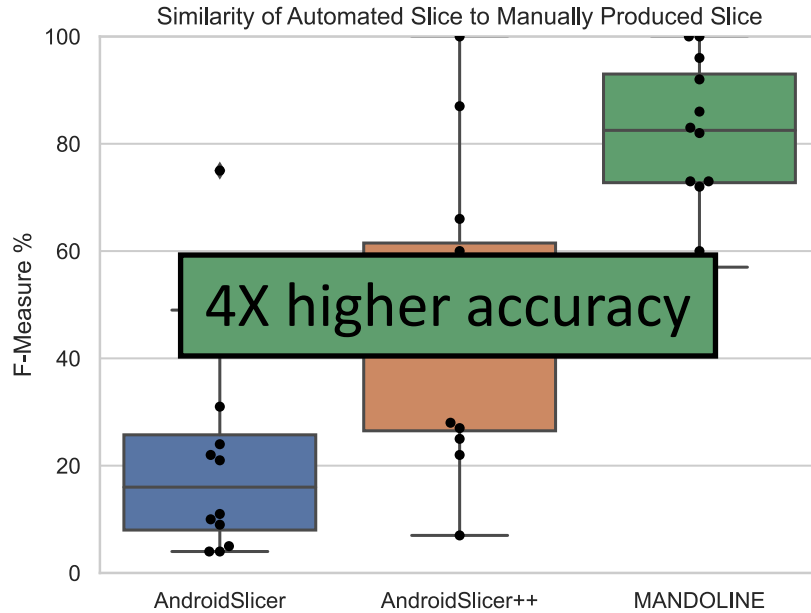
RQ2



Accuracy and Performance

RQ1

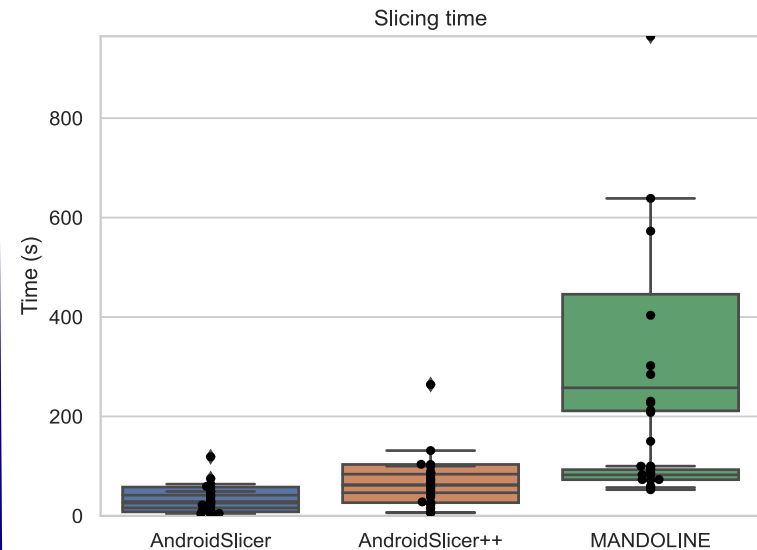
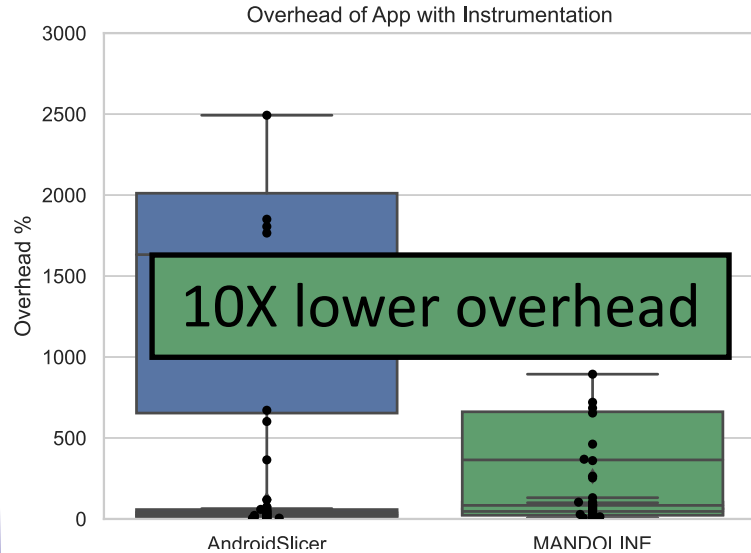
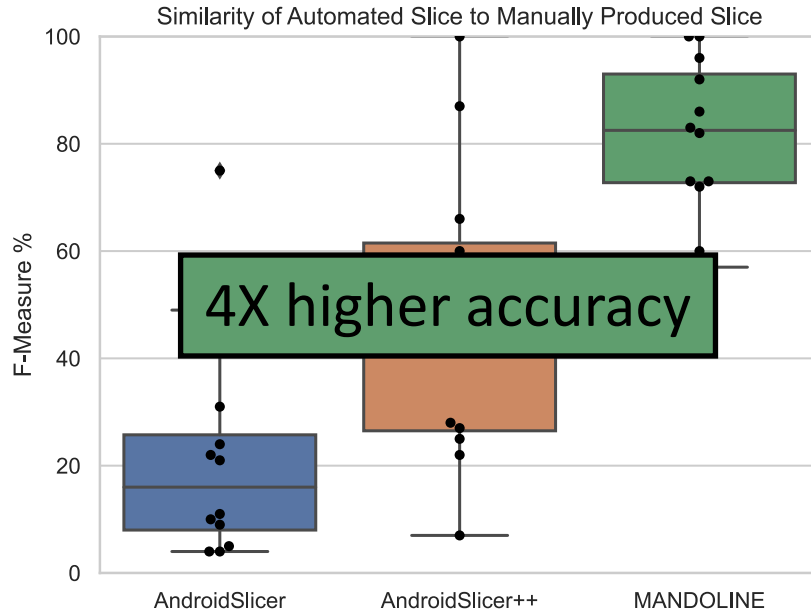
RQ2



Accuracy and Performance

RQ1

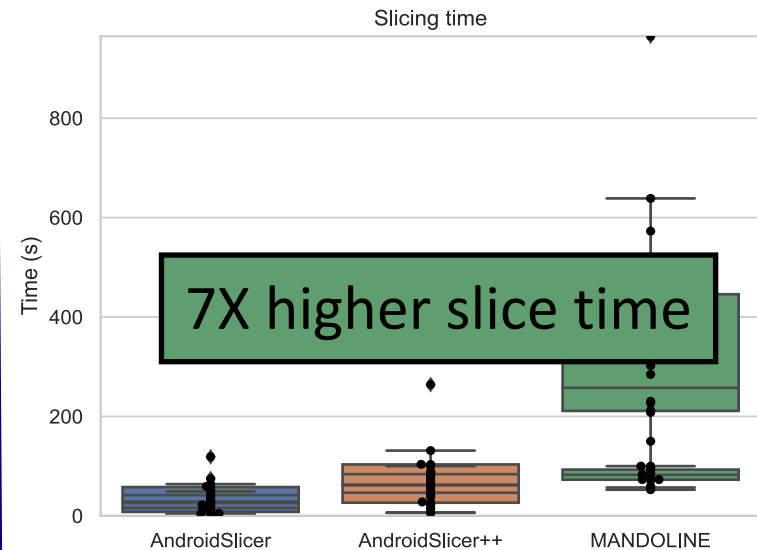
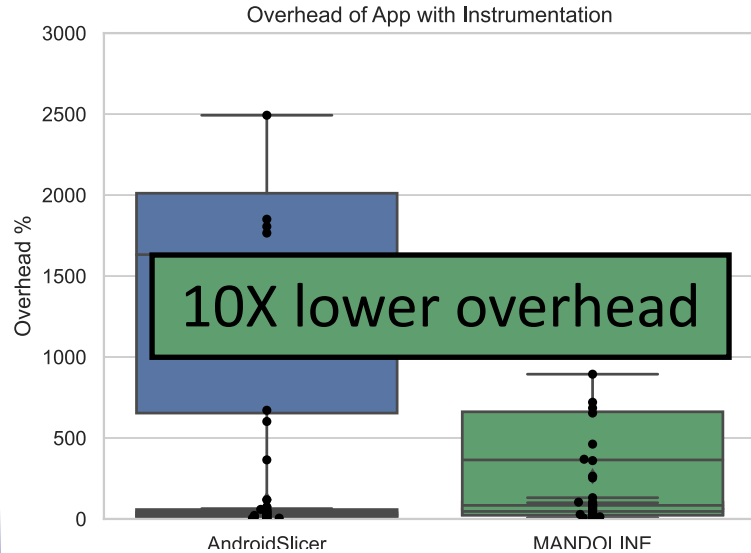
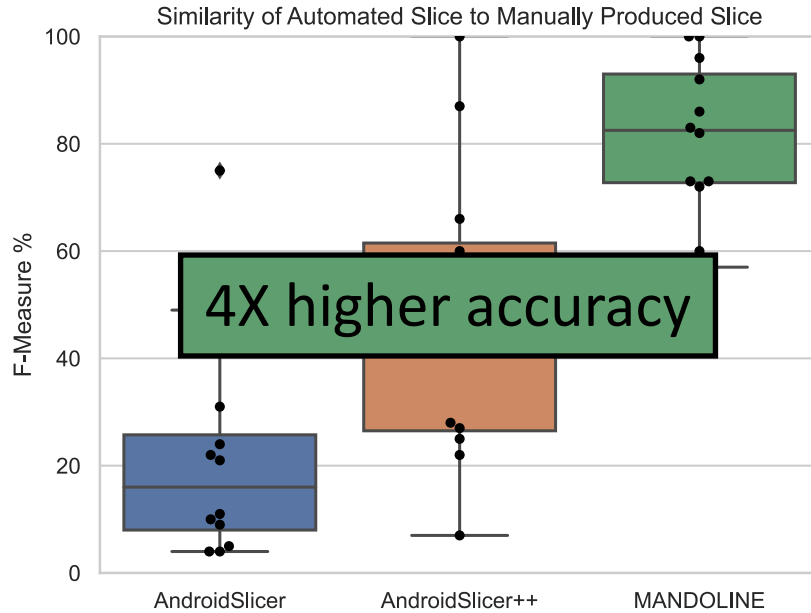
RQ2



Accuracy and Performance

RQ1

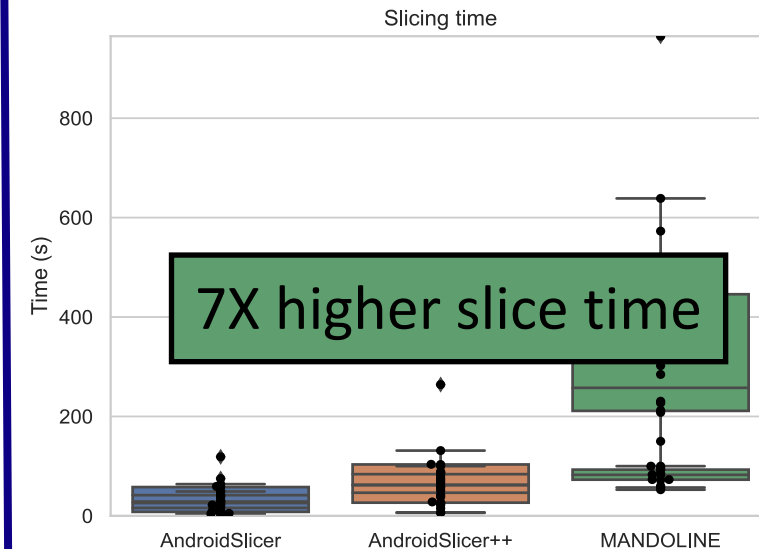
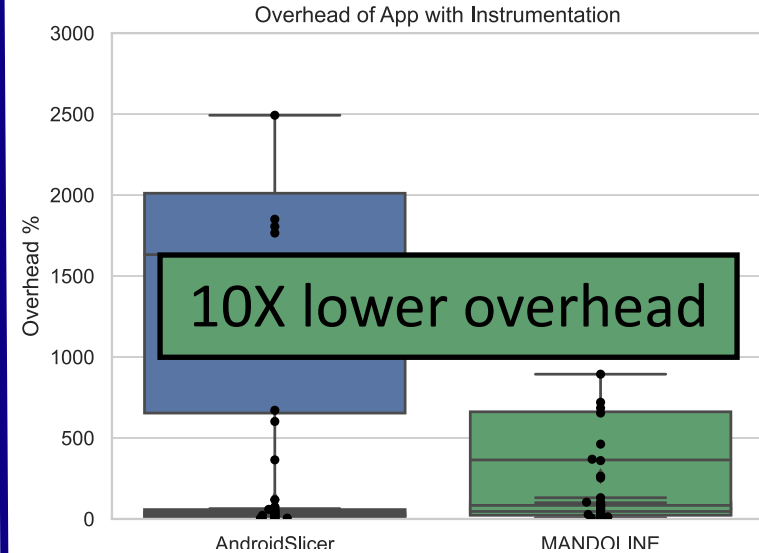
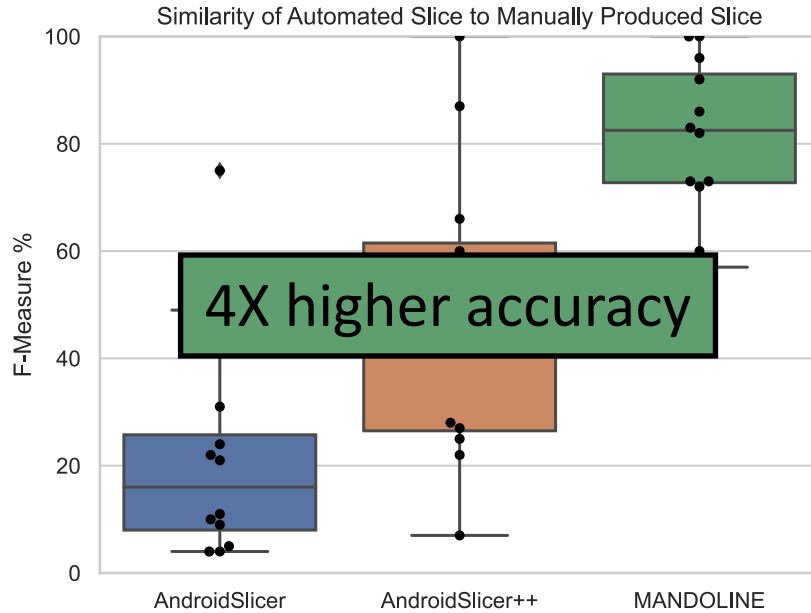
RQ2



Accuracy and Performance

RQ1

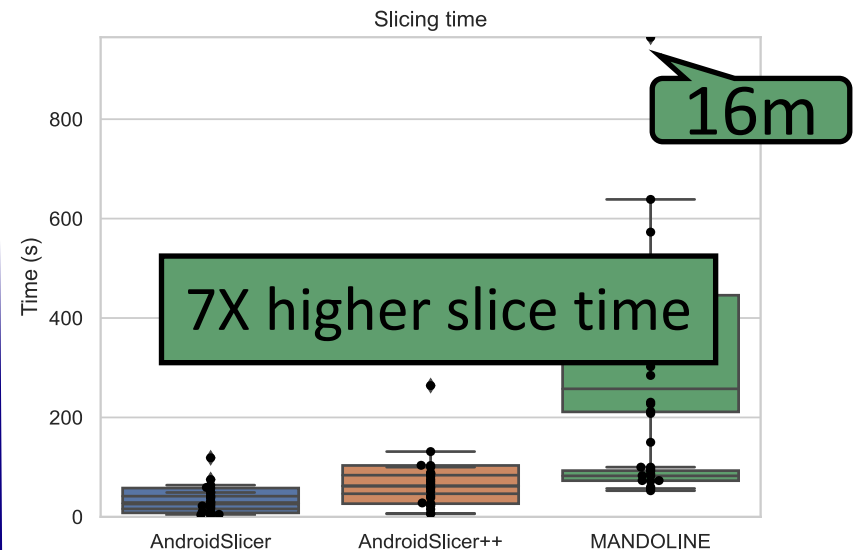
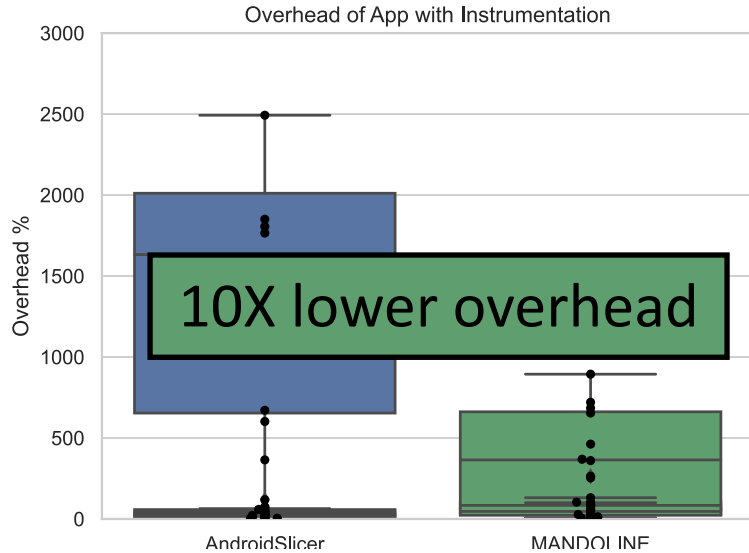
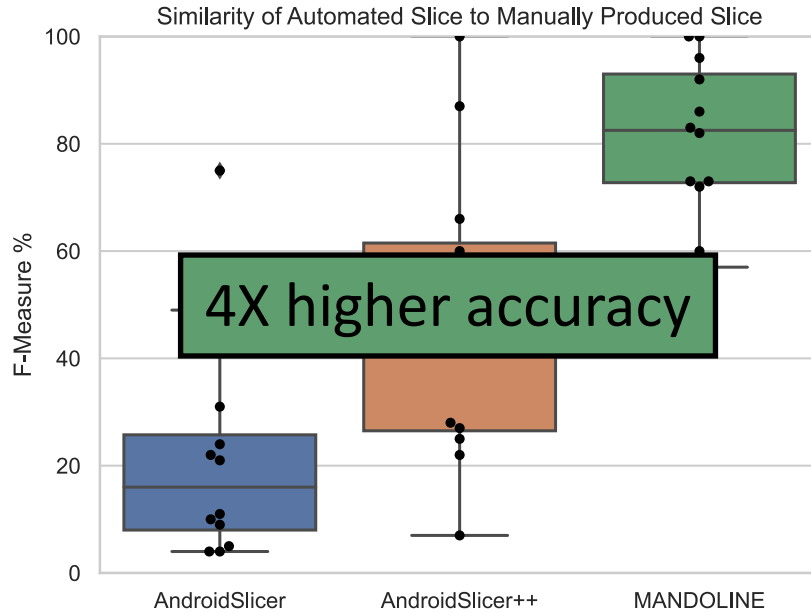
RQ2



Acceptable slice time given huge accuracy improvement

Accuracy and Performance

RQ1



Acceptable slice time given huge accuracy improvement

Summary

Insights

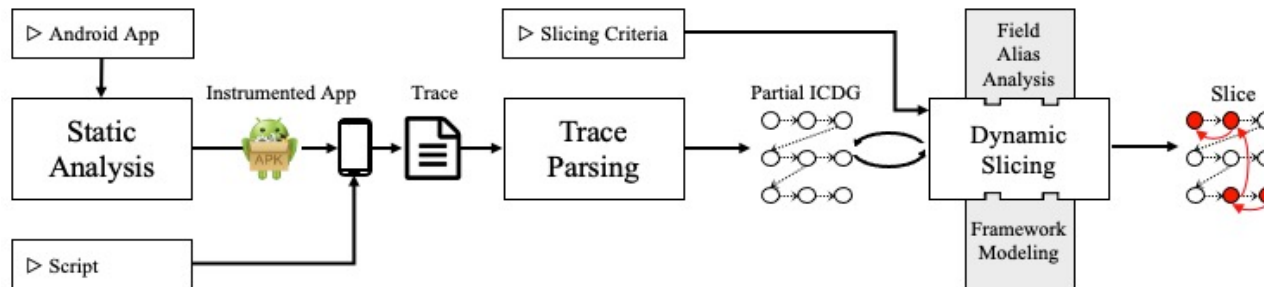


Extract field data flows from the trace using alias analysis

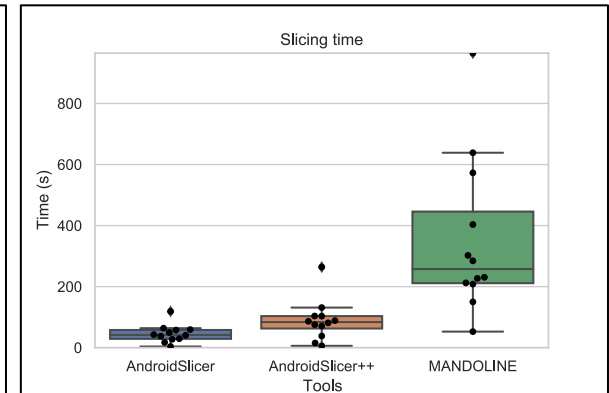
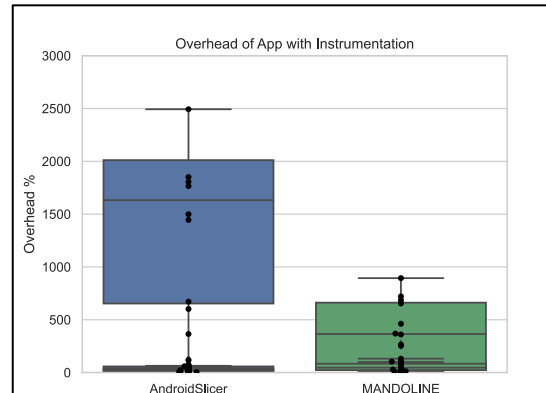
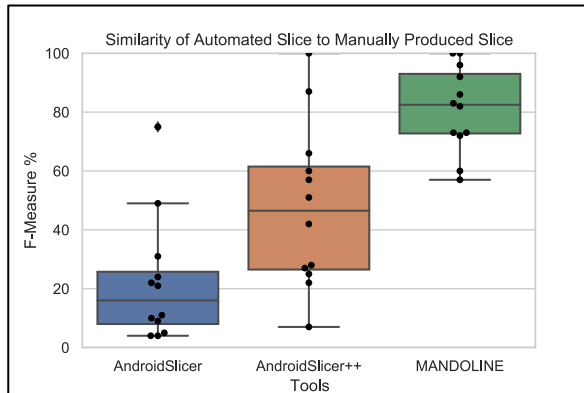


Use framework models from static analysis

Approach



Results



Code + Benchmark suite: <https://resess.github.io/PaperAppendices/Mandoline/>

Contact: Khaled Ahmed, UBC, khaled@ece.ubc.ca