Memory Safety Bug Hunting with BAP

Tiffany Bao, Dominic Chen, Rijnard van Tonder

April 29, 2015

1 Introduction

The Problem

Although numerous debugging tools have been developed for performing static and dynamic software analysis, the majority of these require access to source code or utilization of special compile-time flags to embed debug information within the output binary. In contrast, little effort has been spent on developing static binary analysis tools which can directly analyze off-the-shelf executables without executing them. For this project, we plan to focus on a subset of common programming bugs that tend to introduce security vulnerabilities, known as memory safety errors. Examples of these include buffer overflows, double frees, and uninitialized variables. In particular, we define *Security Outcomes* with respect to certain memory-safety errors, and subsequently implement and test static analyses that yield these outcomes. All analyses were written against our group's Binary Analysis Platform (BAP[1]) [1, 3], and are available online [2].

A factor of this research area that appeals to many is the ability to conclude whether vulnerabilities do or do not exist by performing an analysis purely on the binary level. Due to the inherent difficulties of operating at this level, our approach to the problem starts out coarse, and continues to be incrementally refined as we become more familiar with the numerous barriers that inhibit vulnerability detection in binaries. In this respect, we have much to say towards lessons learned in §5.

1.1 Approach Overview

At the outset of this project, we planned to implement and test the static analyses with respect to our original proposal:

1. Intraprocedural data dependency resolution of instructions

2. Intraprocedural detection of unchecked memory allocations

3. Basic intraprocedural detection of buffer overflows in the stack

4. Basic intraprocedural detection of buffer overflows in the heap

---

[1] https://github.com/BinaryAnalysisPlatform/bap
[2] https://github.com/ddcc/15745-s15-project

Equipped with the knowledge of data flow dependency between instructions (1), we determined to reason about various properties of typical vulnerabilities that we could check (2-4). For example, we might like to follow how arguments to a 'strcpy' call are influenced by previous instructions as a function of the stack register.

1.2 Related Work

Engler et. al. found numerous bugs with system-specific static analyses, implemented by way of a meta-compilation language for the `gcc` compiler [5, 6]. Analyses are composed in the form of correctness rules and automatic inference of rules form source code. Memory errors that are checked include null pointer bugs, not checking allocation results, and uses of freed pointers. This relates to our goal in that we would ideally like to perform similar checks, albeit at the level of complexity of binary code. Whereas Engler's work matured into a generic platform for specifying analyses, we must start first with an ad-hoc approach to gain insight and familiarity in the problem domain presented by binaries.

The Jakstab framework [7] provides a platform for analyzing binaries. However, it is not oriented toward detecting security vulnerabilities, but rather recovering abstractions from binaries (e.g., CFG reconstruction). Frama-C is an example of a framework which allows one to check various properties, including security properties, but works on the source level [4].

1.3 Contributions

Our key technical contributions are as follows:

1. A dataflow analysis framework for BAP.

2. General plugins for BAP that deliver

    (a) reaching definitions and
    (b) data flow dependency resolution (use-def and def-use chains).
    (c) registers corresponding to function call arguments and returns, according to the ARM ABI.

3. Security-specific plugins for BAP that produce security outcomes by checking

    (a) whether memory allocation checks are performed on 'memcpy',
    (b) whether calls such as strcpy are safe from buffer overflows under specific conditions, and
    (c) whether malloc calls...

4. Results of testing the aforementioned plugins on the `GNU Coreutils`[3] software suite.

1.4 Structure

Due to the exploratory nature of this project on the binary level, we concentrate much of our efforts on looking at viable approaches of analysis, rather than optimizing properties of the analysis or testing against measurable benchmarks. For this reason, it doesn't come as a surprise that our evaluation section may be less expressive than our approach and "lessons learned".

---

[3]`https://www.gnu.org/software/coreutils/`

2 Approach

2.1 Overview

Here we discuss our approach toward ensuring security properties as it evolved over the course of the project. Initially, we determined to reason about each of the cases in §1.1 that relate to vulnerabilities (2-4). To do so, we needed some basic infrastructure (1) which we describe first. Thereafter, we describe our specific approaches to performing security analyses with respect to specific *Security Outcomes.*

2.2 The Case for Reaching Definitions

For example, consider the following code, where we desire to verify that the return value of `malloc` is checked:

Listing 1: malloc.c

```
1   #include <stdlib.h>
2   #include <stdio.h>
3
4   int main() {
5       // Uninitialized variable
6       char *array = 0;
7
8
9       array = malloc(32 * sizeof(*array));
10
11      if (!array)
12          return -1;
13
14      free(array);
15
16      return 0;
17  }
```

The disassembly of this binary might appear as follows (with two basic blocks, before and after the call):

Listing 2: Malloc disassembly

```
begin(main_ENTRY)
    000082c8:  04 e0 2d e5      str lr, [sp, #-4]!
    000082cc:  20 00 a0 e3      mov r0, #0x20
    000082d0:  49 00 00 eb      bl #0x124           \; call malloc
end(main_ENTRY)

begin(main_0xc)
    000082d4:  00 00 50 e3      cmp r0, #0x0        \; check malloc return value
    000082d8:  02 00 00 0a      beq #0x8
end(main_0xc)
```

We would like to confirm that a check takes place after the call to malloc; i.e., we would check for the presence of the `cmp` instruction and the operand `r0` in `main_0xc`.

However, it may easily be the case that the contents of `r0` are first moved to another register, `r3`, which is subsequently checked by a `cmp` instruction. In order to verify the buggy condition that the malloc return value is never checked, we must verify that any flow of this data from `r0` to other registers are also not checked. To obtain the statements where `r0` data flows to, we make use of def-use chains. For the definition of `r0` in the BIL IR, we check all uses of `r0` and transitively look up the def-use chains for further registers, such as `r3`, which the value in `r0` flows to. On the final set of instructions, we verify that the data is subject to a `cmp` instruction.

This example demonstrates the need for def-use chains; our stack and heap based analyses in turn also require use-def chains. Due to retreating edges in the CFG, it is difficult to produce def-use chains given a single instruction. For this reason we first implemented a dataflow framework and a reaching-definitions pass. An exhaustive description of the implementation details for this framework, as implemented in OCaml for BAP, may be viewed in Appendix A. From the reaching definitions, we generate def-use and use-def chains as necessary—e.g., we filter on those reaching definitions which concern a particular use or def of a register.

2.3 Data Flow Dependency

An extension to the def-use and use-def chains was to produce the data (flow) dependencies of a given statement in a forward or backward direction, respectively. Essentially, given a register in a definition (such as a Move instruction in the BIL IR), we can produce a set of all instructions that are determined to transitively affect its value. We implement this by way of a data dependency plugin for BAP. Consider the arguments to the following `strcpy` call:

Listing 3: strcpy.c

```
#include <stdio.h>
#include <string.h>

int main(int argc, char** argv) {
  char dst[10] = {0};
  strcpy(dst, argv[1]);
}
```

For the `dst` variable on the stack, encoded as an `<address,statement>` pair, we have:

```
<(0x8490:32, 0)> R0 := R2
```

The plugin will determine all preceding instructions it is dependent on:

```
<(0x8444:32, 3)> SP  := SP - 0x8:32
<(0x8448:32, 2)> R11 := SP + 0x4:32
<(0x848C:32, 2)> R2  := R11 - 0x10:32
```

Now, we can observe the role that the stack pointer `SP` plays when considering `dst`. A full listing of the BIL disassembly, the data dependencies for `argv[1]`, and a graphical output may be viewed in the Appendix B.

2.4 Malloc Checking Security Outcomes

For Malloc checking, we suppose that a security program should always check the return value of `malloc`. This is a strong yet reasonable requirement for ensuring secure code, because if the pointer fails to be assigned a value due to a lack of space for allocation, an attacker may be able to exploit a NULL pointer dereference at a later point in the program.

To demonstrate how our plugin check works, observe the example shown in 2.2. In this example, the function `malloc` is called at address `0x82d0`. Followed by the instruction, we have a comparison instruction which compare the return value `r0` to constant value `0`. This piece of code is thus a candidate for `malloc` checking.

Note that the comparison instruction is not necessarily available immediately after the malloc call instruction. Also, a more complicated case could be that the return value is passed to another register which in turn is checked instead of `r0`.

Another example that we came across is from binary coreutils_O1_[, which is shown below:

4

Listing 4: Malloc disassembly

```
d4d0:        e0800009        add     r0, r0, r9
d4d4:        e280a002        add     sl, r0, 2
d4d8:        e2800003        add     r0, r0, 3
d4dc:        ebffee80        bl      8ee4 <malloc@plt>              \;
d4e0:        e1a08000        mov     r8, r0                         \*
d4e4:        ea000006        b       d504 <locale_charset+0x208>
d4e8:        e0891000        add     r1, r9, r0
d4ec:        e081100a        add     r1, r1, sl
d4f0:        e281a002        add     sl, r1, 2
d4f4:        e1a00004        mov     r0, r4
d4f8:        e2811003        add     r1, r1, 3
d4fc:        ebffee4b        bl      8e30 <realloc@plt>
d500:        e1a08000        mov     r8, r0
d504:        e3580000        cmp     r8, 0                          \;
d508:        1a000005        bne     d524 <locale_charset+0x228>    \;
d50c:        e1a00004        mov     r0, r4
```

In this example, after calling malloc at address d4dc, the program copies the return value to register `r8`, and then to address `d504` where it compares the value of `r8` to constant 0. The reason of compiling the program in such way is mainly because of optimization. Specifically, this helps decreasing the redundant code, by reusing `d504:    cmp r8, 0` and the subsequent instructions.

To address this problem, we use the forward-data-deps plugin to help finding the instructions that depends on the return value. Here dependence means that one or more variables used in the instruction depends on register `r0`. For example, given instruction `d4e0:    mov r8, r0`, we find that instruction `d504:    cmp r8, 0` depends on `r0`, because the value of `r8` is from `r0`. Also, instruction d508 depends on r0, because the flag registers are determined by the `d504`, and in `d504` `r8` depends on `r0`.

Looking back to the example, and using forward-data-deps plugin, we get all dependence instructions once we find a move instruction with `r0` as source. For In the next step, we check if there is `cmp` instruction among all instructions we found. If there is one, we consider call is safe. If not, we consider the call unsafe.

2.5   Stack-based Security Outcomes

A number of unfruitful approaches documented in §5 lead us to opting for an approach where we identify two properties which are important from a security perspective, without inferring sizes of stack-based buffers.

src is a constant   If we can determine that the `src` buffer is a constant, (and therefore not dependent on any user input), the `strcpy` call is safe. While a good property to check, it was found that a constant value corresponding to the `src` argument didn't occur in practice. Depending on compiler optimizations `strcpy` (as well as `sprintf` calls) would be compiled to a `LDMIA.W` ARM instruction which directly copies memory. Lastly, we observed that `strcpy` and `sprintf` may be replaced by the `__strcpy_chk` and `__sprintf_chk` counterparts, which are less interesting to analyse from a security standpoint.

Nevertheless, we implemented a partial check of this case in the plugin, and assumed the standpoint that any `src` address originating from the `ro.data` section implies a constant, fixed value. However, for this check we only consider the statement which loads a value immediately before the call, and not originating previous statements.

Stack accesses within the frame   For this outcome, we confirm that all instructions which access elements on the stack, plus a fixed offset, are within the stack range of a function. We infer the size of the frame by considering the statements that manipulate SP in the function prologue.

While we didn't observe any cases where the stack range was exceeded, this type of check forms the basis for further, more sophisticated improvements.

2.6 Malloc-based Security Outcomes

Another common memory safety problem is reading or writing to memory exceeding an allocated buffer; in certain situations, these buffer overflows are remotely exploitable and can be leveraged by an attacker to take control of a given program. As a result, we decided to implement a basic heap check plugin for the Binary Analysis Platform (BAP) that is capable of identifying and warning when static program execution violates heap safety.

Another common memory safety problem is reading or writing to memory The below example program, known as `doublefree`, suffers from a number of common memory safety problems. Among these is an out of bounds array write to the `array` variable, demonstrating a common off-by-one error that exceeds the allocated buffer size by one byte. Additionally, this program also exemplifies a number of other memory safety issues that we have developed BAP plugins to detect, including a missing return value check for the call to `malloc`.

Listing 5: malloc_free.c

```
1   #include <stdlib.h>
2   #include <stdio.h>
3
4   int main() {
5     char *array;
6     printf(  Hello World!\ n   );
7
8     // Use of uninitialized memory
9     printf("Pointer:_%p\n", array);
10
11    // Missing check against NULL return value
12    array = malloc(32 * sizeof(*array));
13
14    // Out of bounds array access
15    array[32] = 0;
16
17    // Double free
18    free(array);
19    free(array);
20
21    return 0;
22  }
```

At the most basic level, the `heap_check` functions by searching for calls to external library functions responsible for performing memory allocation such as `malloc`, `xmalloc`, `calloc`, etc. Then, it queries the BAP ABI interface to determine the number of arguments passed to the library function, and the register that contains the return value. Template matching is used to identify the specific argument that contains the allocation size, since the former two functions take only one argument, whereas the latter takes two arguments, with the allocation size passed in the second argument.

At this point, the program searches backwards through the function from the callsite for the most recent definition of the register that contains the allocation size, and analyses this definition to determine if it can statically identify a constant allocation size. If this is the case, then the plugin tags this allocation callsite with the size of the allocation, and proceeds to the next allocation callsite. If it cannot statically determine the allocation size, then the program currently tags the allocation callsite with an allocation size `MAX_INT`, essentially performing widening to the supremum of the semilattice.

Next, after identifying and tagging all of the memory allocation callsites, the plugin iterates through each of the callsites, and attempts to statically identify all possible uses of the return value register that involve a memory operation (e.g. LDR and STR assembly instructions). For each of these uses, the plugin identifies the register that contains the memory address, searches backwards from the memory instruction for the most recent definition of this register, and statically attempts to determine the register offset from the allocation site. This process is similar to that performed above for determining the size of a memory allocation.

Finally, if the program was able to statically identify both the memory allocation size and the memory operation offset, the program compares the two values to ensure that the operation is less than the allocation size. If this is not the case, the program emits a warning about this unsafe memory operation, otherwise it labels the operation safe and proceeds to the next use.

Since the current BAP framework lacks a number of memory optimizations that would simplify analysis, this plugin currently only performs a basic level of memory operation checking. In particular, it currently only analyzes memory write operations, and is restricted to simple memory offset computations that only involve addition. As a result, it cannot perform analysis on loop operations or complex assembly sequences in which the pointer to the memory allocation is stored into memory or selectively modified through indirect control flow. Nevertheless, the current plugin is capable of identifying memory safety vulnerabilities for the example program shown above, as shown below.

```
0x8330:32: .init_proc ([832c, 8333]) -> call_weak_fn ([83dc, 83f3])
0x83C8:32: _start ([83a0, 83cb]) -> __libc_start_main ([837c, 8387])
0x83CC:32: _start ([83cc, 83cf]) -> abort ([8394, 839f])
0x83F4:32: call_weak_fn ([83f4, 83f7]) -> __gmon_start__ ([8388, 8393])
0x848C:32: __do_global_dtors_aux ([848c, 848f]) -> deregister_tm_clones ([8400, 8417])
0x84C8:32: frame_dummy ([84c4, 84cb]) -> register_tm_clones ([8438, 8457])
0x84E4:32: main ([84d4, 84e7]) -> puts ([8364, 836f])
0x84F0:32: main ([84e8, 84f3]) -> printf ([834c, 8357])
0x84F8:32: main ([84f4, 84fb]) -> malloc ([8370, 837b])
>> 0x84F8:32: malloc (0x20:32)
0x8518:32: main ([84fc, 851b]) -> free ([8358, 8363])
0x8520:32: main ([851c, 8523]) -> free ([8358, 8363])
>> 0x84FC:32: R3 = R0
>> 0x8508:32: R3 = R3 + 0x20:32
>> 0x8510:32: R3 + 0x0:32 <- t_1605 (UNSAFE: 0x20:32 <= 0x20:32)
0x8560:32: __libc_csu_init ([853c, 8563]) -> .init_proc ([832c, 8333])
```

The above plugin output identifies addresses at which calls were made from one function to another using the -¿ (arrow) symbol, labeling both the source and destination function with their symbol name and relevant basic blocks.Then, for each memory allocation callsite, the program attempts to statically determine the allocation size, shown above with the malloc (0x20:32) syntax. Subsequently, the program attempts to statically determine the use of each allocation site, the offset of the memory operation, and the actual memory operation, which are respectively shown above by the movement of memory allocation pointer from R0 (the return value of malloc) to R3, the increment of register R3 by 0x20:32, and the storage of temporary variable t_1605 (containing the value 0 loaded from R2) into R3. At the point, the plugin identifies that memory operation offset

exceeds the size of the allocation site, and emits a warning that the operation is unsafe (UNSAFE: 0x20L32 ¡= 0x20:32).

Note that the :32 label indicates that the respective memory address is 32-bit.

3   Experimental Setup

Our experimental setup consisted of running our plugins with BAP `0.9.6` [1] on a set of binaries, including GNU `Coreutils` and our own examples.

3.1   Unchecked Malloc

In this experiment, we checked the following property for all arm `Coreutils` binaries:

- after a call instruction with target `malloc`, the return value must be checked.

- if the return value is checked to be zero, there must be a conditional jump followed to deal with the zero case.

- the return value of `malloc` can be passed to other registers. If so, checking other register with the return value should be also valid.

3.2   Stack Check Plugin

The stack check plugin was tested against the GNU `Coreutils` suite. We primarily chose the `Coreutils` suite since a) it can be compiled for ARM, and b) it contains numerous calls to `strcpy`, `memcpy`, `malloc`, and so forth. The stack check plugin, while configurable, concretely checked for function calls to:

- strcat

- strcpy

- strncpy

- memcpy

The plugin collected a number of statistics on the binaries, including the number of argument sites (that is, statements that set up registers before a call), the number of functions which contained calls to 'dangerous' functions, and the maximum number of statements on which an argument site statement was dependent. These are illustrated in Figure 1.

Figure 3.2 presents the total number of statements in a binary which affect unique argument sites.

The figures illustrate some interesting properties with respect to argument sites and data dependencies. Consider that each potentially dangerous function takes 2 or 3 parameters–of these, our data dependency plugin determined the maximum number of statements a single argument site statement is dependent on. We observe from Figure 1 that this could number in the hundreds, whereas the total number could range in the thousands (Figure 3.2). The implication here is that any additional reasoning and refinement to static analysis methods across cases composed of sizable dependencies need to be carefully considered. For instance, if we add forms of pointer analysis, we may speculate about the scale of complexity of memory operations for argument sites.
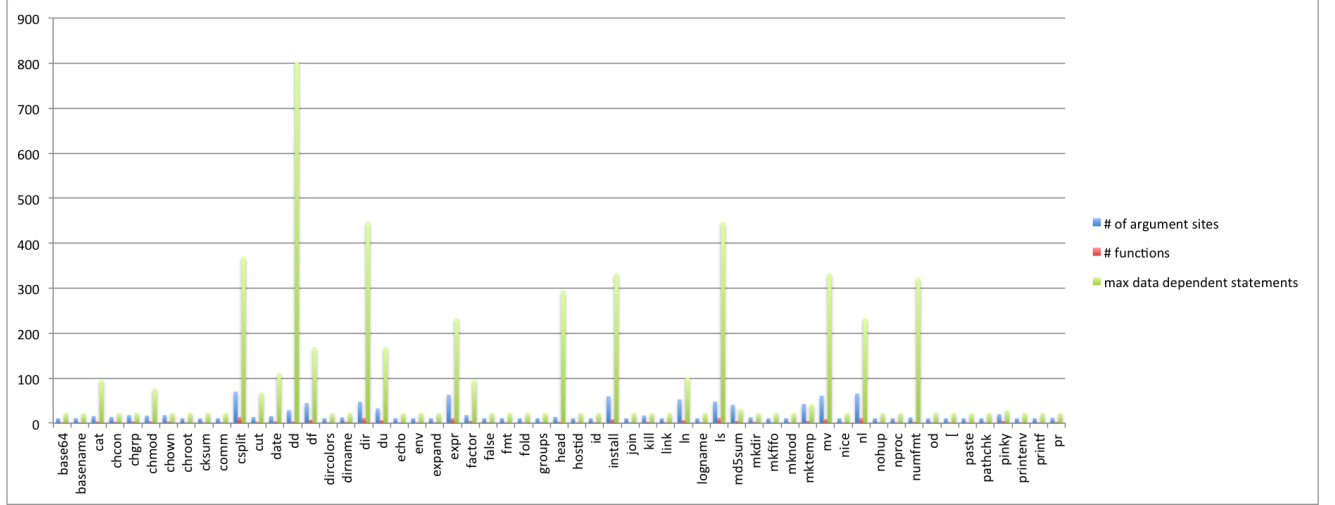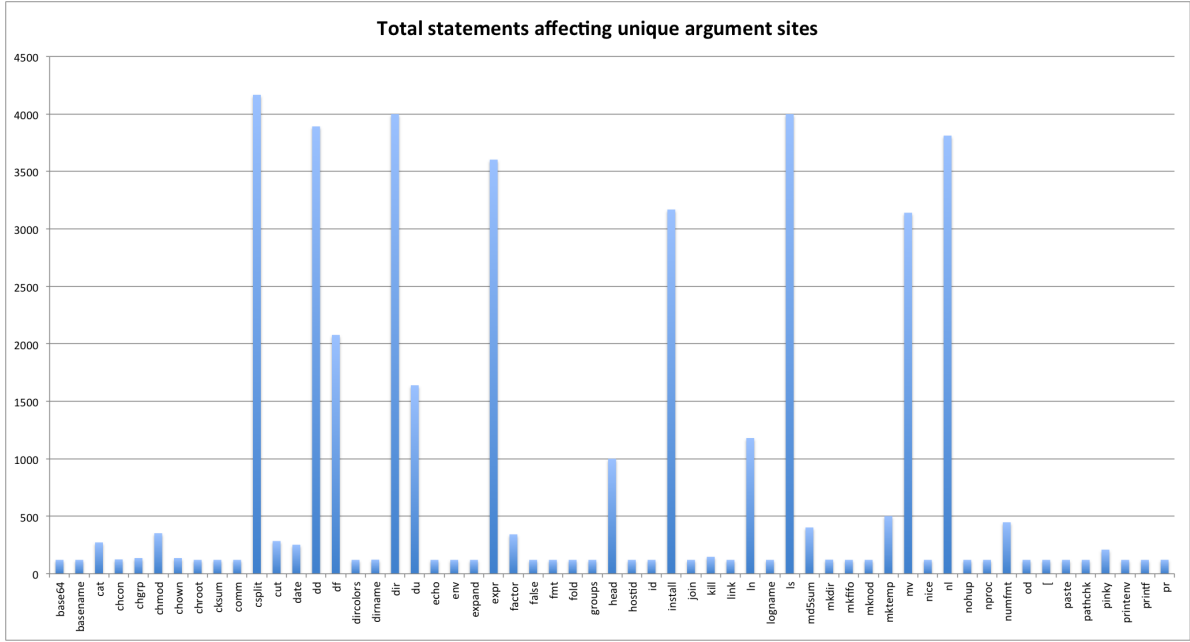
Figure 1: Coreutils statistics generated by plugin

We did not observe any violations with respect to the security outcomes in §2.5. This is perhaps unsurprising, given the ubiquitous use of the `Coreutils` suite. At the same time, it is difficult to choose a test suite which purposefully exhibits the stack properties that we check; see §5 for further discussion.

3.3 Malloc Check Plugin

For evaluation of the heap check plugin, we executed the plugin on each executable from the `coreutils` suite of system programs. This is composed of 103 binaries, such as `cat`, `wc`, `who`, etc. Additionally, we compiled four separate versions of each plugin with different optimization levels; one at `O0`, one at `O1`, one at `O2`, and one at `O3`.

Of this output, our plugin identified a total of 4303 calls to preidentified memory allocation functions, namely "xmalloc", "malloc", "xrealloc", "realloc", and "calloc". Among these, the memory allocation size was identified for 369 of these callsites, whereas 3934 of these callsites had an unknown allocation size. Drilling down further, 804 of these unknown callsites involved left arithmetic shifts (multiplication) on a temporary variable, which our plugin would likely be able to identify given some additional development work. Another 426 of these unknown callsites had an allocation size that was loaded from memory, which would be difficult resolve without pointer analysis and significant additional development effort. The remaining 2707 of these unknown callsites were due to the allocation size value being passed through more than one register, which our plugin is currently unable to resolve. By making the callsite identification routine recursive, it is likely that we would be able to resolve these unknown callsites further into one of the above subcases. Below is a sample of the program output demonstrating these different results:

```
coreutils_O3_groups.out:182:>> 0xAFFC:32: xmalloc (UNKNOWN: R1)
coreutils_O3_head.out:17:>> 0x9C8C:32: xrealloc (UNKNOWN: t_26405 << 0x2:32)
coreutils_O3_head.out:52:>> 0x9730:32: xmalloc (UNKNOWN: mem[SP + 0x64:32, el]:u32)

coreutils_O3_who.out:335:>> 0xD2D0:32: xmalloc (0x21:32)
```

Of the above callsites with memory allocation sizes that were statically identified, the memory operation offset was successfully determine and identified as safe for five of these, as shown below. None were identified as unsafe, and 301 were unchecked. It is likely that more of these unchecked uses would be resolvable given additional development effort, since the current implementation is only capable of resolving additive (and not negative, multiplicative, etc) offsets to the allocation site. Note that of the five memory operations identified as safe, all were in the O0 version of the respective `coreutils` binary, indicating that higher optimization levels are generating more complex code that currently cannot be resolved.

```
coreutils_O0_chmod.out:194:>> 0xAF14:32: R3 + 0x0:32 <- t_67547 (UNCHECKED)
coreutils_O0_chmod.out:196:>> 0xAF20:32: R3 + 0x1:32 <- t_67551 (UNCHECKED)
coreutils_O0_chmod.out:198:>> 0xAF2C:32: R3 + 0x4:32 <- R2 (UNCHECKED)
coreutils_O0_chmod.out:200:>> 0xAF38:32: R3 + 0x8:32 <- R2 (UNCHECKED)
coreutils_O0_chmod.out:202:>> 0xAF44:32: R3 + 0xC:32 <- R2 (UNCHECKED)

coreutils_O0_chmod.out:204:>> 0xAF54:32: R3 + 0x1:32 <- t_67567 (SAFE: 0x20:32 > 0x10:32)
coreutils_O0_install.out:1515:>> 0x179BC:32: R3 + 0x1:32 <- t_186264 (SAFE: 0x20:32 > 0x10:32)
coreutils_O0_mkdir.out:176:>> 0xA880:32: R3 + 0x1:32 <- t_36422 (SAFE: 0x20:32 > 0x10:32)
coreutils_O0_mkfifo.out:109:>> 0x9B14:32: R3 + 0x1:32 <- t_25151 (SAFE: 0x20:32 > 0x10:32)
coreutils_O0_mknod.out:125:>> 0xA2B0:32: R3 + 0x1:32 <- t_31175 (SAFE: 0x20:32 > 0x10:32)
```

4 Experimental Evaluation

In general, we were pleased that our plugins ran successfully on our own example binaries as well as `Coreutils`. A key aspect of evaluating the success of our project is that we were able to start with a basic API in BAP, and after implementing a non-trivial amount of 'infrastructure' code, arrive at three individual plugins which are able to produce security outcomes.

Our evaluation of these plugins are largely qualitative rather than quantitative at this stage. However, they serve as very useful data points toward developing more sophisticated analyses.

4.1 Malloc Check Plugin

So far we applied our checking on 129 binaries from `Coreutils`. The result showed that we found 15 places that `malloc` is not checked according to our rules.

Listing 6: Function re_node_set_alloc

```
0000f910 <re_node_set_alloc>:
    f910:       e92d4010        push    {r4, lr}
    f914:       e1a04000        mov     r4, r0
    f918:       e5801000        str     r1, [r0]
    f91c:       e3a03000        mov     r3, 0
    f920:       e5803004        str     r3, [r0, 4]
    f924:       e1a00101        lsl     r0, r1, 2
    f928:       ebffe6f6        bl      9508 <malloc@plt>
    f92c:       e5840008        str     r0, [r4, 8]
    f930:       e3500000        cmp     r0, 0
    f934:       03a0000c        moveq   r0, 12
    f938:       13a00000        movne   r0, 0
    f93c:       e8bd8010        pop     {r4, pc}
```

For example, we found that although function `re_node_set_alloc` compares the return value of `malloc`, it does not deal with the case when the return value is 0. Instead, the return value will be stored to the stack and overwritten to other value.

4.2 Stack Check Plugin

As mentioned, our stack check plugin did not trigger any violations when run over the `Coreutils` test suite. This result was largely anticipated: the `Coreutils` library was chosen so that we could verify that our plugins run and to rule out the possibility that binaries in `Coreutils` exhibited rudimentary flaws.

In critique of this plugin, there are a number of stack-based errors that could occur despite our checks. For example, consider a stack frame containing multiple local buffers at runtime. It is feasible that a buggy `strcpy` call could copy bytes beyond the boundary of a single buffer into an adjacent buffer, while remaining within the stack bounds. Our plugin would not detect this case. Consider a further case where memory accesses to the stack are dependent on general purpose registers—here we cannot conclude whether the access is within range of the stack frame.

A positive side-effect of the plugin is the potential of reusability, and by extension, refinement toward more sophisticated analyses. Recall that this plugin collects and makes use of statements that an argument site depends on. Thus, we have the ability to store the output of our analysis without having to rerun the plugin. For example, we may be able to couple this data with pointer-analysis, and infer stack boundaries.

4.3 Malloc Plugin

5 Surprises and Lessons Learned

There were a number of surprises and lessons learned throughout this project.

5.1 Reasoning about buffer overflows

An initial approach towards ensuring stack buffer overflow detection was to attempt to reason about sizes of buffers which are passed to "Potentially Dangerous Functions" [8] such as `strcpy`. This turned out to be an unfruitful approach for the scope of the project, since one would have to invariably have to reason about memory accesses.

A concrete example that illustrates the difficulty of performing checks on `strcpy` was found when we considered `libxkbfile.so.1.0.2`. Initially we thought that this is a candidate we could perform our check on, due to containing a `strcpy(buf,"none")` [2] call. However, this was found to compile to the `LDMIA.W` instructions mentioned before. This example principally highlights the difficulty in choosing candidates for testing.

One surprising result was the great number of statements (over 200) that some argument sites were found to depend on in `Coreutils` binaries. It was also the case that this result was exhibited inside large functions. When testing the plugin on small examples initially, the number of statements rarely numbered more than 20. As mentioned previously, we speculate that this will affect

5.2   General Things Learned

In general, this project made clear the difficulty of tackling binary problems. Specifically, we cannot advance the problem of analyzing binaries without basic "building blocks" which provide information about the security outcome we wish to address. For each such potential outcome in the future, we will have to consider critically what types of "building blocks" we need, whether pointer analysis, path-sensitivity, context-sensitivity, etc.

We learned that it is more effective in our research to develop our approaches with respect to security outcomes. By doing so, we incrementally identify the requirements for producing desired security outcomes (e.g., reaching definitions, use-def). It also seems to be a pattern that more sophisticated analyses are built on top of simpler analyses (like those produced in this project), rather than sophisticated analysis being stand-alone. This follows because we require more fundamental information before reasoning about further complexities.

6   Conclusions and Future Work

We will focus specifically on refining the sophistication of our analyses to handle broader cases as outlined in our discussions. A major benefit of identifying the limitations in our approch and evaluation is that it immediately highlights areas of refinement. The additional goals beyond 100% in our original proposal are all candidates for formulating future analyses. In summary, these are:

- Implement intraprocedural detection of null pointer dereferences

- Implement detection of dynamic memory errors e.g. dangling pointers, double frees, and invalid frees

- Implement intraprocedural detection of variables with uninitialized variables e.g. usage of variables with unassigned values

- Evaluate soundness, completeness, and performance of our tool against similar bug checking tools, e.g. Address Sanitizer, Valgrind MemCheck, Clang Static Analyzer, Coverity, CodeSonar, etc.

It has been a promising exercise to prototype and test our analyses within the context of this project. We consider the experience and insights gained throughout to be particularly valuable for performing continued research in our group.

7 Distribution of Total Credit

A   Appendix

A.1   Technical Description of the BAP Dataflow framework

Due to the preliminary status of the current BAP framework, a number of common compiler analyses have not yet been implemented, including a basic dataflow framework. In particular, BAP currently only includes a disassembly engine using the LLVM framework, a lifter from ARM/x86/x86_64 to an intermediate language known as BIL (BAP Intermediate Language), and a basic semantic analysis for identifying basic blocks and performing a graph-based traversal through predecessors and successors.

As a result, we spent the first portion of our project implementing a basic dataflow analysis framework capable of performing both forwards and backwards dataflow analysis. This was accomplished by first implementing a sparse Bitvector module capable of representing both the infimum and supremem lattice elements (e.g. the empty and universal sets). Existing Bitvector representations in the standard library and Jane Streets Core library are not capable of this, because they utilize a fixed-size bit representation for the Bitvector. As a result, this would require a preliminary pass through each basic block and each instruction to identify the number of lattice elements in the Bitvector, reducing performance and flexibility of analysis.

In particular, our custom implementation defines a Bitvector record (datatype) composed of three subvalues: (1) infinite: a boolean value that determines whether the bitvector is of infinite size, (2) value: a boolean value that determines whether the bits of the bitvector are 0 or 1, and (3) indexes: a set of integers that identifies indexes of the bitvector that are set to value. As a result, this implementation allows for the empty and universal sets to be respectively identified by the following simple Bitvector values: {infinite: true, value: false (0), indexes: {}} and {infinite: true, value: true (1), indexes: {}}. Standard Bitvector operations, such as union, intersection, and difference, are defined using these primitives. As an example, consider an arbitrary Bitvector a, and the following operations: empty - a = empty, a - empty = a, universal - a = not a, a - universal = empty. These operations are straightforward, even for the relative complement of the universal set with the arbitrary Bitvector; this can be represented by simply toggling the value attribute from true to false or vice-versa.

Specifically, this highlights the idea that each finite Bitvector has two distinct representations: one with value set to true and indexes containing a set of bit indexes that are set, and another with value set to false and indexes containing a set of bit indexes that are unset. Although this does increase the logical complexity of the implementation, it allows for greater lattice representation flexibility with decreased memory consumption.

Otherwise, the dataflow framework implementation is relatively standardized, and is built on top of this lattice representation. A pair of maps are defined for translating unique instruction addresses of BIL or assembly instructions to lattice values, and vice-versa. This is achieved through the use of a mutable counter that identifies the unique Bitvector index associated with each instruction. Then, analysis results from each sequence of BIL instructions are propagated to each assembly instruction, and subsequently to each basic block, allowing for analysis granularity at all levels. Users of the framework specify initial values for each interior and boundary point, the entry block of analysis, and the direction of the dataflow analysis. Since the BAP framework is written in a functional language, the user can also pass a user-defined meet function and transfer function

for performing the body of analysis. An internal worklist is used to keep track of basic blocks that needs to be scheduled for reanalysis; this is achieved by making the main dataflow analysis function recursive, and iterating through every basic block on the worklist before constructing a new worklist from those basic blocks with modified dataflow lattice values.

B   Appendix

BIL Dissasembly of Listing 6

```
SP := SP - 0x8:32
mem := mem          with [base_436 + 0xFFFFFFF8:32, el]:u32 <- R11
mem := mem          with [base_436 + 0xFFFFFFFC:32, el]:u32 <- LR
base_436 := SP
R11 := SP + 0x4:32
t_439 := 0x4:32
s_438 := SP
SP := SP - 0x18:32
t_442 := 0x18:32
s_441 := SP
mem := mem          with [R11 + 0xFFFFFFE8:32, el]:u32 <- R0
mem := mem          with [R11 + 0xFFFFFFE4:32, el]:u32 <- R1
R3 := R11 - 0x10:32
t_447 := 0x10:32
s_446 := R11
R2 := 0x0:32
mem := mem          with [R3 + 0x0:32, el]:u32 <- R2
R3 := R3 + 0x4:32
t_452 := 0x4:32
s_451 := R3
R2 := 0x0:32
mem := mem          with [R3 + 0x0:32, el]:u32 <- R2
R3 := R3 + 0x4:32
t_457 := 0x4:32
s_456 := R3
R2 := 0x0:32
mem := mem          with [R3 + 0x0:32, el]:u16 <- t_460
t_460 := low:16[R2]
R3 := R3 + 0x2:32
t_463 := 0x2:32
s_462 := R3
R3 := mem[R11 + 0xFFFFFFE4:32, el]:u32
R3 := R3 + 0x4:32
t_467 := 0x4:32
s_466 := R3
R3 := mem[R3 + 0x0:32, el]:u32
R2 := R11 - 0x10:32
t_471 := 0x10:32
s_470 := R11
```

```
R0 := R2
R1 := R3
jmp 0x82E0:32
LR := 0x849C:32
```

BIL data dependency instructions for `argv[1]`

```
<(0x8444:32, 0)> base_436 := SP
<(0x8444:32, 1)> mem := mem          with [base_436 + 0xFFFFFFFC:32, el]:u32 <- LR
<(0x8444:32, 2)> mem := mem          with [base_436 + 0xFFFFFFF8:32, el]:u32 <- R11
<(0x8444:32, 3)> SP := SP - 0x8:32
<(0x8448:32, 2)> R11 := SP + 0x4:32
<(0x8450:32, 0)> mem := mem          with [R11 + 0xFFFFFFE8:32, el]:u32 <- R0
<(0x8454:32, 0)> mem := mem          with [R11 + 0xFFFFFFE4:32, el]:u32 <- R1
<(0x8458:32, 2)> R3 := R11 - 0x10:32
<(0x845C:32, 0)> R2 := 0x0:32
<(0x8460:32, 0)> mem := mem          with [R3 + 0x0:32, el]:u32 <- R2
<(0x8464:32, 2)> R3 := R3 + 0x4:32
<(0x8468:32, 0)> R2 := 0x0:32
<(0x846C:32, 0)> mem := mem          with [R3 + 0x0:32, el]:u32 <- R2
<(0x8470:32, 2)> R3 := R3 + 0x4:32
<(0x8474:32, 0)> R2 := 0x0:32
<(0x8478:32, 0)> t_460 := low:16[R2]
<(0x8478:32, 1)> mem := mem          with [R3 + 0x0:32, el]:u16 <- t_460
<(0x8480:32, 0)> R3 := mem[R11 + 0xFFFFFFE4:32, el]:u32
<(0x8484:32, 2)> R3 := R3 + 0x4:32
<(0x8488:32, 0)> R3 := mem[R3 + 0x0:32, el]:u32
```
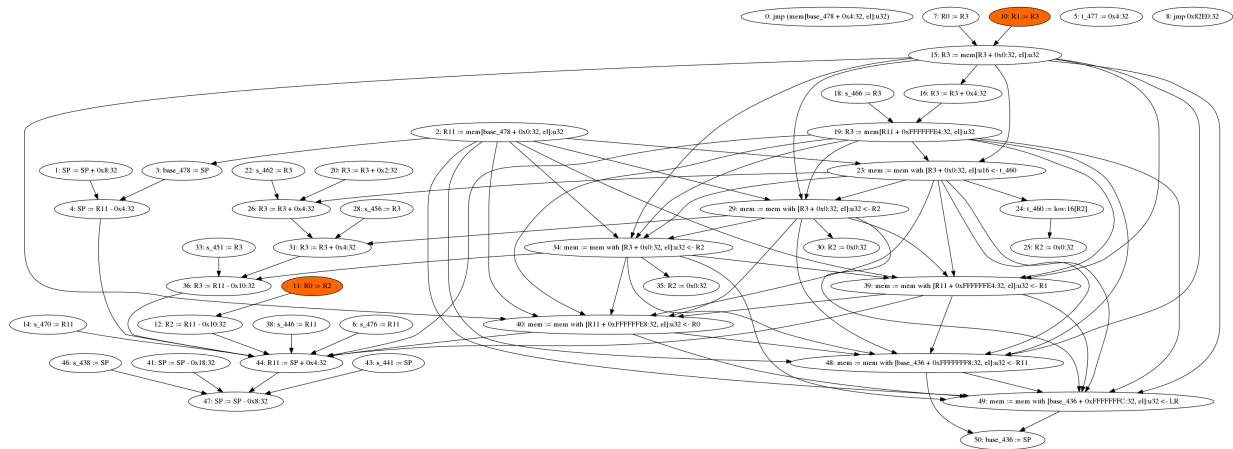

Graphical output of BIL data dependencies, with highlighted arguments

References

[1] Binary analysis platform. `https://github.com/BinaryAnalysisPlatform/bap`. Accessed 03/16/2015.

[2] xorg-server. `https://github.com/wereHamster/xorg-server/blob/server-side-xcb/xkb/xkbtext.c`.

[3] BRUMLEY, D., JAGER, I., AVGERINOS, T., AND SCHWARTZ, E. J. Bap: a binary analysis platform. In *Computer aided verification* (2011), Springer, pp. 463–469.

[4] CUOQ, P., KIRCHNER, F., KOSMATOV, N., PREVOSTO, V., SIGNOLES, J., AND YAKOBOWSKI, B. Frama-c. In *Software Engineering and Formal Methods*. Springer, 2012, pp. 233–247.

[5] ENGLER, D. Finding bugs with system-specific static analysis. `http://web.stanford.edu/~engler/paste02-talk.pdf`. Accessed 03/16/2015.

[6] HALLEM, S., CHELF, B., XIE, Y., AND ENGLER, D. *A system and language for building system-specific, static analyses*, vol. 37. ACM, 2002.

[7] KINDER, J., AND VEITH, H. Jakstab: A static analysis platform for binaries. In *Computer Aided Verification* (2008), Springer, pp. 423–427.

[8] SEACORD, R. C. *The CERT C secure coding standard*. Pearson Education, 2008.