Evaluating State-of-the-Art Free and Open Source Static Analysis Tools against Buffer Errors in Android Apps

Bushra Aloraini
David R. Cheriton School of Computer Science
University of Waterloo
Waterloo, ON, Canada
baloraini@uwaterloo.ca

Meiyappan Nagappan
David R. Cheriton School of Computer Science
University of Waterloo
Waterloo, ON, Canada
mei.nagappan@uwaterloo.ca

Abstract*—Modern mobile apps incorporate rich and complex features, opening the doors for different security concerns. Android is the dominant platform in mobile app markets, and enhancing its apps security is a considerable area of research. Android malware (introduced intentionally by developers) has been well studied and many tools are available to detect them. However, little attention has been directed to address vulnerabilities caused unintentionally by developers in Android apps. Static analysis has been one way to detect such vulnerabilities in traditional desktop and server side desktop. Therefore, our research aims at assessing static analysis tools that could be used by Android developers. Our preliminary analysis revealed that `Buffer Errors` are the most frequent type of vulnerabilities that threaten Android apps. Also, we found that `Buffer Errors` in Android apps have the highest risk on Android that affects data integrity, confidentiality, and availability. Our main study therefore tested whether state-of-the-art static analysis tools could detect `Buffer Errors` in Android apps. We investigated 6 static analysis tools that are designed to detect `Buffer Errors`. The study shows that the free and open source state-of-the-art static analysis tools do not efficiently discover `Buffer Error` vulnerabilities in Android apps. We analyzed the tools carefully to see why they could not discover `Buffer Errors` and found that the lack of semantic analysis capabilities, inapplicability to Android apps, and the gap between native code and other contexts were some of the reasons. Thus, we concluded that there is a need to build better free and open source static analysis tools for detecting `Buffer Errors` in Android apps.

I. Introduction

Android dominates the mobile market as it has gained tremendous popularity recently [1]. The main Android market, Google Play Store, included 3 million apps and at least 65 billion downloads as of June 2017 [2][3]. Users assume that app markets guarantee the security of offered apps [4]. Yet, app markets typically do not fully ensure the security of the apps they offer [5]. In fact, apps hosted on these markets may possess malware or vulnerabilities (that were introduced unintentionally by developers). A study by Symantec shows that 17% of all Android apps were actually malware in disguise [6]. Yet, every app could contain an unintentional vulnerability. Such vulnerabilities could be exploited by attackers to leak private data, modify the software or data, or deny availability of systems and data causing substantial economic loss.

Studies indicate that the number of vulnerabilities are rapidly rising [7][8]. According to Risk Based Security [7], in 2015 the number of vulnerabilities increased by 77% compared to the vulnerabilities reported in 2011. In 2017 Q1, the number of vulnerabilities increased by 29.2%, over the same period in 2016 [9]. Another study showed that nearly 75% of tested mobile apps showed at least one critical or high-severity security vulnerability [10].

Studying Android vulnerabilities is therefore a very important area of research. Beside the security concerns, Android apps that have vulnerabilities are likely to get negative reviews and being abandoned by users [11]. Thus, Android developers need to ensure that their apps are secure. In the desktop and server side software, static analysis tools are a common proactive method to find security vulnerabilities in source code, early during the coding phase [12]. This leads to cost savings (a key benefit of static analysis tools), as the earlier a vulnerability is detected, the cheaper it is to fix [13]. In fact, applying a static analysis approach to Android apps to detect vulnerabilities could potentially ensure quality and reliability of the apps and hence the app market. Thus, static analysis tools are common in both maintenance and evolution of software.

However, we do not know how effective static analysis tools are in detecting vulnerabilities in Android apps. *This research, thus, aims to study state-of-the-art static analysis tools that can be used to detect Android vulnerabilities during the coding process.* In our study we ask the following RQs.

Motivational RQ: What are the most common vulnerabilities in Android apps? From data in the National Vulnerability Database (NVD) [14], we found that `Buffer Errors` are the most frequent vulnerability that happens in Android apps. They also have the highest risk that compromises integrity, confidentiality, and availability.

Case study RQ: Are state-of-the-art static analysis

978-1-5386-0992-7/17 $31.00 © 2017 IEEE
DOI 10.1109/ICSME.2017.77

tools for `Buffer Errors` able to detect vulnerabilities reported in the wild for Android apps? In this study, we investigated 17 static analysis tools that could discover `Buffer Errors`. Out of the 17, we tested 6 free and open source static analysis tools on 9 real world `Buffer Errors`, and it was found that none of the studied tools could efficiently detect the reported `Buffer Errors`.

All our empirical data (including the vulnerabilities from NVD and the Android apps with the vulnerabilities) are available for download [15].

II. Motivational RQ: What are the most common vulnerabilities in Android apps?

A. Motivation

Since examining the efficiency of static analysis tools on all types of vulnerabilities is beyond the scope of a single study, we wanted to find the most frequent type of vulnerability that occur in Android apps. Therefore, in this RQ we identify Android vulnerability trends by examining the published vulnerabilities in the wild. Software security vulnerability records are maintained by multiple security vulnerability databases, such as NVD, Open Source Vulnerability Database (OSVDB) [16], CERT [17], and Bugtraq (BID) [18]. These databases utilize the Common Vulnerabilities and Exposures Identifier (CVE-ID) [19], which is a unique number for a publicly recognized security vulnerability, as an identifier of a vulnerability record. NVD is synchronized with CVE-ID and its records are mainly based on CVE information. Thus, NVD is preferred by researchers as it is the most comprehensive database. Therefore, NVD was selected to construct the dataset for this RQ.

B. Methodology

1) Data Gathering: In this phase, the relevant vulnerability records of Android from 2008 (when Android was first released) to 2015 have been extracted from NVD using an automated web-scraping tool. The "Android" keyword was used to filter the NVD database and extract relevant Android vulnerability reports. For each vulnerability record, the CVE-ID, original release date, last revised date, description, CVSS v2 base score, impact score, exploitability score, access vector, access complexity, authentication, impact type, and vulnerability type were retrieved. 2,089 records were extracted in the initial data gathering.

2) Removing Irrelevant Data: The dataset has been cleaned in order to get more accurate results. All 2,089 entries have been examined manually to check whether they are accurately related to Android; if not, the record is excluded from the dataset. Four records were found unrelated to Android: CVE-2015-3906, CVE-2015-3815, CVE-2012-1344, and CVE-2011-1001.

3) Data Processing: In this phase, we traced each vulnerability record manually in order to collect more information, such as whether the vulnerability has been confirmed or patched, and how it was discovered. This kind of information was obtained from the software vendor websites and other resources, such as BID, OSVDB, CERT, and Japan Vulnerability Note (JVN) [20]. Based on the obtained information, we further sanitized the dataset. We classified the vulnerability records into multiple categories. Vulnerability record categories with hyperlinks of the acquired information could be found in our dataset in columns "Record Category" and "Record Category URL". The categories are as follow:

- **Confirmed and patched:** An advisory has been published by the vendor that explains the vulnerability and provides patching information. We found that 556 vulnerability records fall into this category.
- **Reported and patched:** The vulnerability has been reported to the vendor and patches were released. We found 67 vulnerability records fall into this category.
- **Proof of concept and patched:** A proof of concept has been demonstrated by the reporter, and it is indicated that the vendor has released patches. We found 5 vulnerability records belong to this category.
- **Confirmed but not patched:** The vulnerability has been confirmed by the vendor, however no patches were provided. We found 5 vulnerability records belong to this category.
- **Proof of concept but not patched:** A proof of concept has been demonstrated by the discoverer, and it is indicated that the vendor has not released patches. We found 30 vulnerability records fit into this category.

While categories that are excluded from the study are:

- **Not enough information:** In case that the vulnerability report does not include enough information, such as a proof of concept, patching information, or confirming from the vendor, then the record is excluded from the dataset. 33 records were excluded as they do not include enough information.
- **Large-scale experiment:** An automated large scale experiment has been conducted by Dormann [21] to test whether Android apps properly validate SSL certificates provided by HTTPS connections. The study was conducted on 23668 Android apps and 13 Android libraries, as a result 5.9% (1379 apps and 10 libraries) have been reported to be vulnerable. 23.2% of the tested apps were found vulnerable because of a vulnerability in the libraries. These reports have been found in NVD from 9/8/2014 to 10/29/2014 as `Cryptographic Issue` vulnerability type. Those records were excluded in order to produce more balanced and meaningful results. 1389 vulnerability records belong to this category.

4) Data Classification: After cleaning the dataset, 663 records remained. Next, the vulnerability records were examined manually in order to categorize them into two groups: Android platform vulnerabilities, and Android app vulnerabilities. Android app vulnerabilities are any vulnerability that could be triggered within Android apps. All other vulnerabilities are considered as Android platform vulnerabilities, for example,

vulnerabilities that reside in the Linux kernel of the Android system.

C. Results

Android platform related vulnerabilities are 187, and Android app vulnerabilities are 476, representing 72% of the all vulnerabilities in our cleaned dataset (663).

1) Vulnerability Type Trend: We aim to tackle vulnerabilities at app level, but not at the OS level. Thus, an important question of this empirical study is "what is the most frequent type of vulnerability that occurs historically in Android apps". To reveal the trend of Android vulnerabilities, NVD classification of vulnerability type was utilized, which is based on the Common Weakness Enumeration (CWE) [22]. CWE is a list of software flaw types that is maintained by the MITRE Corporation and used by security organizations and researchers. In the studied dataset, there are 24 different type of flaws that occurred in the Android platform, and 20 types of them occurred in the Android apps. The percentage of each vulnerability type has been calculated in order to uncover the trending vulnerability in Android apps. As shown in Table I, `Buffer Error` is the dominant vulnerability in Android apps – 28.6% of all discovered Android app vulnerabilities.

`Permissions`, `Privileges`, and `Access Control` (18.1%) and `Information Leak/Disclosure` (11.3%) vulnerabilities are the second and third most frequent vulnerabilities in Android ecosystem respectively. `Insufficient Information` (12.4%) type indicates that there is insufficient information about the vulnerability to categorize it. Such case usually happens when vendors confirm a vulnerability but decline to release certain details about the vulnerability.

Although, `Permissions`, `Privileges`, and `Access Control` and `Information Leak/Disclosure` are dominating the Android security research community [23], rarely are `Buffer Error` vulnerabilities related to Android apps being discussed (See Section VII-B for more details). Since `Buffer Error` is the most frequent vulnerability that occurred historically in Android apps, further analysis is needed to recognize how severe and dangerous `Buffer Error` vulnerabilities are.

2) Buffer Error Severity Trend: In order to analyze how severe discovered `Buffer Error` vulnerabilities are, the Common Vulnerability Scoring System (CVSS) [24] was employed. CVSS is an open standard used to assess the severity of security vulnerabilities, and it is platform and technology independent. We used CVSS Base score version 2 standards. We found during our analysis that 97% of the `Buffer Error` vulnerabilities have high risk. These results show that `Buffer Errors` have serious implications on the security of Android apps and its end users. Also, 99.26% of `Buffer Errors` in Android apps are remotely exploitable and no authentication is required at all. Thus, it could be concluded that the `Buffer Errors` in Android apps are easy to exploit. Finally, we found that 95.6% of `Buffer Error` vulnerabilities completely impacted the target app in terms of confidentiality, integrity, and the availability.

Vulnerability Type	CWE-ID	Total	%
Buffer Errors	CWE-119	136	28.6%
Permissions, Privileges, and Access Control	CWE-264	86	18.1%
Insufficient Information	NVD-CWE-noinfo	59	12.4%
Information Leak / Disclosure	CWE-200	54	11.3%
Cryptographic Issues	CWE-310	28	5.9%
Input Validation	CWE-20	23	4.8%
Cross-Site Scripting (XSS)	CWE-79	16	3.4%
Numeric Errors	CWE-189	15	3.2%
Path Traversal	CWE-22	15	3.2%
Other	NVD-CWE-Other	11	2.3%
Resource Management Errors	CWE-399	8	1.7%
Code Injection	CWE-94	7	1.5%
Authentication Issues	CWE-287	4	0.8%
Cross-Site Request Forgery (CSRF)	CWE-352	3	0.6%
Improper Access Control	CWE-284	3	0.6%
Security Features	CWE-254	3	0.6%
Credentials Management	CWE-255	3	0.6%
Code	CWE-17	2	0.4%
Data Handling	CWE-19	1	0.2%
OS Command Injections	CWE-78	1	0.2%

TABLE I: Vulnerability types distribution in Android apps

Also, our data shows that even though Android becomes more mature, `Buffer Error` still appear in the latest versions of Android. Due to space constraints, we omitted some aspects of `Buffer Errors` analysis. However, all details could be found here [15].

> Buffer Errors are the most common vulnerability in Android apps. They also have the highest risk that compromises integrity, confidentiality, and availability. As a result, we concluded that we need to focus our main study on `Buffer Errors`.

Determining the most impactful vulnerability affecting Android apps (which involved hundreds of hours of manual work), is not the focus of the study. However, we believe that the empirical evidence shown above not only motivates the rest of our study, but is a useful contribution to the research community that looks to solve the issue of `Buffer Errors` in Android apps.

III. Background

Before we examine state-of-the-art static analysis tools, we want to give some background about the Android app architecture and `Buffer Errors`.

A. Android Application Architecture

Android apps are unlike standard applications in two important ways. First, Android apps run in a security sandbox to manage application resources. Hence, each app represents a different process with a unique UID, and it is executed in isolation from other apps with restricted permissions. Second,

Android apps are framework-based and event-driven. Thus, Android apps and Android OS communicate through callbacks. Android apps have no single entry point, the so called main method, though there are multiple entry points. These entry points represent components that can be used by other apps if needed or called by the Android OS. There are four types of components in Android apps: activities are a single focused user interface, services run background tasks, content providers act as a database storage, and broadcast receivers listen for framework events. The components of an application could be executed in any order, and each component has a complete lifecycle [25].

Android apps are typically written in the Java programming language. However, the Android Native Development Kit (NDK) that was provided by Google, allows developers to implement Android apps using native programming languages, such as C and C++. In fact, native code allows developers to use existing third party libraries, and allows hardware specific optimization of performance critical code. Indeed, Zhou et al. [26] in 2012 reported that 4.52% of Android applications use native code. In 2014, Qian et al. indicted that 16.46% of tested Android apps uses native code [27]. NDK enables Android developers to combine native code with an Android app using Java Native Interface (JNI). JNI is a foreign function interface (FFI) by which a program written in Java can call routines or make use of services written in native code, such as C/C++ and assembly; yet, JNI can be also utilized to invoke Java objects from native code.

Both the Java code and native code of an Android app run within the same process. Thus, the native code still adheres to the entire application permissions set in the manifest file. Java code is managed and executed by the Dalvik Virtual Machine (DVM)/Android RunTime (ART) [1], while native code is not restricted to DVM/ART and manages itself throughout the lifetime of the application. This requires further responsibilities on developers, such as memory management tasks, hence `Buffer Errors` occur in native code due to improper memory management. To execute Android apps effectively, the native components are also expected to carefully interact with Java components. If this interaction is not appropriately managed, the native components can cause errors that could crash the entire application [28].

B. Buffer Errors

`Buffer Errors` usually occur on the stack or on the heap. Stack is adjacent blocks of memory that is controlled by OS, and it is used for storage of local variables, and return addresses, and passing extra arguments to subroutines when there are inadequate argument registers available. Stack-based buffer overflow happens when an application writes extra data outside an intended memory address on the program's call stack. This type of vulnerability is a serious one as the stack contains the return addresses for all function calls. If the affected application is running with higher privileges, or receives data from untrusted sources then the defect is a potential security vulnerability. Stack-based buffer overflow could result in corrupting local variables, crashing the application, or executing arbitrary code.

In Android, the heap is managed by particular APIs in Bionic library [2] using the dynamic memory allocator jemalloc. Heap memory is used to create dynamic data objects, and to store objects that must live longer than one function's lifetime. Buffers allocated on the heap are subject to the same boundary checking issues as those located on the stack. Heap-based buffer overflow usually happens when an application writes extra data to a heap chunk buffer, which leads to corruption of the control sections of the chunk. This leads the memory allocator to an undefined state, and potentially, crashing the application or even executing arbitrary code.

In fact, `Buffer Errors` in Android do not differ much from other platforms as vulnerability characteristics. What is distinct is that static analysis tools that support other native projects might not be applied well in Android, as `Buffer Errors` in Android apps could involve communication from Java to native code through JNI. Yet, most current static analysis tools focus on one language per analysis, thus they would not contain the whole picture when analyzing a portion of the program. Another issue is that Android apps do not have a main method, but they compromise multiple entry points which could involve different call graphs than traditional apps.

IV. Case Study

A. Selection Criteria for Buffer Error Vulnerabilities

In our case study we want to test state-of-the-art static analysis tools against `Buffer Error` vulnerabilities in Android apps. While we could write our own toy Android apps and inject `Buffer Error` vulnerabilities, we feel that it would be a biased experiment. Hence, we mined the vulnerability records from NVD to identify example vulnerabilities in open source Android apps that we can use as case study subjects. Having open source Android apps was necessary so that we could run static analysis tools to analyze the source code.

We found 9 `Buffer Error` vulnerability records in 3 open source Android apps (Google Chrome, Android browser, and Mozilla Firefox). Thus by examining the static analysis tools against existing popular Android apps with real word vulnerabilities makes our experiments, results and conclusions stronger. In this section, we describe each of the vulnerabilities categorized within corresponding reason in Table II and follow up with a discussion on some of the common attributes among all of them.

1) Buffer Size Miscalculation:

- **CVE-2008-0985**
  A remote attacker could cause a heap-based buffer overflow by persuading a victim to visit a malicious web site that contains GIF components. The vulnerability occurs at the GIF library in the WebKit framework in Android web

[1] In recent Android versions, Java code is managed and executed by Android runtime (ART).

[2] https://android.googlesource.com/platform/bionic/

CVE ID	Reason	C/C++	Affected app	Source (input)	Sink (Container)	Data flow
CVE-2008-0985	Buffer size miscalculation	C++	Android Browser	User (Gif image)	Class	Inter-procedural
CVE-2017-5014	Buffer size miscalculation	C++	Google Chrome	User (Image)	uint32_t pointer	Inter-procedural
CVE-2016-5182	Lack of boundary checking	C++	Google Chrome	User (Bitmap image)	Smart pointer	Inter-procedural
CVE-2014-1705	Lack of boundary checking	C++	Google Chrome	User (JS code manipulation)	Function template	Inter-procedural
CVE-2014-3201	Lack of boundary checking	C++	Google Chrome	User (Scroll size)	Smart pointer class template	Inter-procedural
CVE-2014-1710	Lack of boundary checking	C++	Google Chrome	User (GPU Command Buffer)	Class	Inter-procedural
CVE-2012-4190	Null pointer dereference	C	Mozilla Firefox	Within the app	int pointer	Inter-procedural
CVE-2016-5200	Incorrectly applied type rules	C++	Google Chrome	User (JS code manipulation)	Class	Inter-procedural
CVE-2016-5199	Off by one error	C	Google Chrome	User (video file)	Struct	Inter-procedural

TABLE II: Studied Buffer Error vulnerabilities in Android apps

browser. It fails to properly sanitize input which is a .gif file before copying it to an inadequately sized memory buffer. The problem occurs due to allocating buffer size based on the logical screen width and height field of the GIF header. However, the buffer is filled in with bytes based on the real width and height of the GIF image.

- **CVE-2017-5014**
  A remote attacker could cause a heap-based buffer overflow in Google Chrom by persuading a victim to visit a malicious web site that contains crafted image components. The overflow happens during image processing in Skia [3] that miscalculates the buffer size of the image.

2) Lack of Boundary Checking:

- **CVE-2016-5182**
  A remote attacker could cause a heap-based buffer overflow by persuading a victim to visit a malicious web site that include a crafted bitmap. The Google Chrome rendering engine Blink fails to render that particular size causing `Buffer Error`.
- **CVE-2014-3201**
  A remote attacker could cause a buffer overflow by persuading a victim to visit a malicious web site that embeds another document using iframe. The embedded page specifies large dimensions for ::webkit-scrollbar and embeds an image with ::-webkit-scrollbar-corner. The Google Chrome rendering engine Blink fails to render that particular size causing `Buffer Error`.
- **CVE-2014-1705**
  A heap-based buffer overflow vulnerability was found in the Google V8 JavaScript engine, which is an open source JavaScript engine written in C++. It exists within handling of TypedArray objects. The vulnerability occurs due to missing bounds checking for the length of ArrayBuffer when manipulated using js defineGetter method, which is then fed to the TypedArray object during initializing. This may allow an attacker to read and write data to any memory address which could be leveraged to arbitrary code execution in the Google Chrome sandbox process.
- **CVE-2014-1710**
  In this vulnerability, Google Chrome does not validate whether a certain location is within the bounds of a shared-memory segment. This allows remote attackers to cause GPU command-buffer memory corruption and a denial of service. The GPU command-buffer is the way in which Chrome communicates to the GPU either OpenGL or OpenGL ES, which are APIs for rendering 2D and 3D vector graphics. User interaction is required to exploit this vulnerability and cause GPU process to crash; in such that the user should open a carefully malicious a crafted page that has scripts to dynamically modify the structure of the web page after load time.

3) NULL Pointer Dereferences:

- **CVE-2012-4190**
  This vulnerability was reported by a user that was complaining about Mozilla Firefox app crashing in Android in CyanogenMod kernel. The developers took a month to figure out the exact problem. The issue was triggered because of the Cairo library, written in C, which was calling the FreeType library from the system path instead of calling it from in-tree causing memory corruption. So when FreeType library is initially created, it has non-NULL module pointers. However, at some later point, one of the pointers has become NULL. The vulnerability was patched by forcing Cairo library to use Mozilla in-tree setlcdfilter of FreeType. Thus, calling a system function instead of using local function led to NULL pointer dereference.

4) Incorrectly Applied Rules:

- **CVE-2016-5200**

[3]https://skia.googlesource.com/skia/

A heap-based buffer overflow vulnerability was found in Google V8 JavaScript engine, that allowed a remote attacker to it by persuading a victim to visit a crafted HTML page. The Typer in V8 incorrectly applied type rules when using asm optimizer that could cause an out of bound read/write.

5) Off by One Error:

- **CVE-2016-5199**
  A remote attacker could cause a heap corruption via a crafted video file. The vulnerability occurs in Google Chrome due to an off by one error that leads to an allocation of zero size in FFmpeg MP4 decoder which results in corrupting a number of pointers.

B. Common Attributes of Buffer Errors in Android Apps

It was found that all the studied `Buffer Error` vulnerabilities are in client-side apps, such as web browsers. Also, 7 out of 9 of the studied `Buffer Error` vulnerabilities are C++ based and they have some common characteristics. For instance, input is read from untrusted sources, untrusted input is inadequately validated, or lack of boundary checking. In addition, most of the studied vulnerabilities involve pointer indirection. Also, `Buffer Errors` in our case study are inter-file/inter-procedural, as the buffer is allocated in one function in one file, and overflow in another function in a different file. In fact, the inter-file/inter-procedural communication occurs in cross-language manner, meaning that the input come from Java context while buffer errors occur in native context.

In addition, it was found that these vulnerabilities occur through common attack surfaces in web browser, such as JavaScript, Hypertext Transfer Protocol (HTTP), Hypertext Markup Language (HTML), Document Object Model (DOM), and Cascading Style Sheets (CSS).

C. Examined Static Analysis Tools

Static analysis tools examine the program statically without executing it. They can analyze either the source or binary code of the program. In our study, we only focus on tools that support source code analysis. As `Buffer Errors` is a well-known problem, multiple static analysis tools and methods have been already proposed. We investigated 17 popular static analysis tools to detect `Buffer Errors`. We gathered some tools by reviewing research work that evaluated methods that detect `Buffer Errors` [29] [30] [31] [32] [33]. Also, we studied some other static analysis tools that are popular in the wild, such as Clang Static Analyzer [34] and Frama-C [35]. When we collected and studied tools and methods that discover `Buffer Errors`, we found that none of these tools was implemented specifically for Android. However, they are general tools for multiple platforms.

We classified the studied static analysis methods based on classification by Shahriar and Zulkernine [32]. Shahriar and Zulkernine [32] classified static analysis methods that target `Buffer Errors` based on several features, such as inference algorithm, analysis sensitivity, analysis granularity, and target language [32]. The underlying inference algorithm refers to how does the static analysis method infer potential vulnerabilities methodically by analyzing program code. Inference methods are categorized into four types: string pattern matching, tainted data flow, constraint, and annotation. String pattern matching method is one of the simplest methods of static analysis tools (e.g., RATS, and Flawfinder). This technique tokenizes program source code to identify a well-known set of tokens or library function calls that could cause a `Buffer Error`. The other three types; tainted data flow, constraint, and annotation involve more advanced analysis to understand the semantic of the source code [32]. Analysis sensitivity indicates how the static method uses pre-computed information based on program code before running the inference algorithm. On the other hand, analysis granularity refers to the granularity level of program code at which an inference is carried out. Table III shows the studied `Buffer Error` static analysis tools classified based on [32]. Also, we categorized the static analysis methods into two categories: Open source and Commercial.

In this study, we want to determine which of the state-of-the-art static analysis tools could potentially detect the nine Android `Buffer Error` vulnerabilities described in Table II. As our study revealed that most of Android `Buffer Errors` occurred in C++ language, thus all static analysis tools that target C language could not detect these kind of vulnerabilities. So we excluded 5 out of 17 static analysis tools that target C language, such as Splint, BOON, ARCHER, and UNO. We found that there are 12 tools could analyze C++ and could potentially detect studied vulnerabilities.

Six out of the 12 static analysis tools are commercial tools that we did not test (We tested CodeSonar, see Section VI for more details). However, in this paper we only focus on free and open source static analysis tools, therefore we only tested the six free and open source static analysis tools.

V. RQ: Are state-of-the-art static analysis tools for Buffer Errors able to detect vulnerabilities reported in the wild for Android apps?

Static analysis tools are a typically adopted solution by developers to keep project costs down. In this section, we evaluate the efficiency of the six free and open source static analysis tools.

A. Methodology

1) Collecting open source static analysis tools that detect `Buffer Error`:

We gathered the six open source static analysis tools that discover `Buffer Error`. Our study tests the following tools that support C++: IKOS [37], Frama-C [38], Clang Static Analyzer [34], Cppcheck [31], Flawfinder [31], and RATS [31]. Table IV summarizes all tested tools and their versions.

2) Collecting the source code of the vulnerable apps: We then gathered the source code of open source Android apps that have `Buffer Error` vulnerabilities in our study.

The following source code versions of Android apps were collected and checked:

Method Name	Type	Language	Inference Algorithm	Sensitivity	Granularity
Polyspace Bug Finder [29] [35] [36]	Commercial	C/C++	Constraint: abstract interpretation	Flow	Inter-procedural
Parasoft C/C++test	Commercial	C/C++	String pattern matching, Constraint: symbolic execution	Flow	Inter-procedural
Klocwork Insight [36]	Commercial	C/C++	Unpublished	Flow, path	Inter-procedural
Coverity [36]	Commercial	C/C++	Unpublished	Flow, path, context	Inter-procedural
PVS-Studio [35]	Commercial	C/C++	Constraint: symbolic execution, annotations	Flow, path, value range	Inter-procedural
CodeSonar [35]	Commercial	C/C++	Constraint: symbolic execution, taint data flow	Flow, path	Inter-procedural
ASTREE [30][35]	Commercial	C	Constraint: abstract interpretation	Context	Inter-procedural
ARCHER [29]	Open source	C	Constraint: symbolic execution	Flow, path, context, alias	Inter-procedural
BOON [29][30]	Open source	C	Constraint: integer range	N/A	Inter-procedural
Splint [29][31][30]	Open source	C	Annotation	Flow	Intra-procedural, inter-procedural
UNO [29][35][31]	Open source	C	Annotations	Flow, path	Inter-procedural
Flawfinder [31][30]	Open source	C/C++	String pattern matching	N/A	Token
RATS [31][30]	Open source	C/C++	String pattern matching	N/A	Token
Cppcheck [31][30]	Open source	C/C++	Constraint: integer range	Flow, context	Inter-procedural
Clang Static Analyzer [34]	Open source	C/C++	Constraint: symbolic execution, annotation	Flow, path	Inter-procedural
Farma-C [35]	Open source	C/C++	Constraint: abstract interpretation, annotations	Flow, value range, point-to	Inter-procedural, System dependence graph
IKOS [37]	Open source	C/C++	Constraint: abstract interpretation	Flow, path, point-to	Inter-procedural

TABLE III: State-of-the-art static analysis tools to detect Buffer Errors

Tool Name	Version Number
RATS	2.4
Flawfinder	1.31
Cppcheck	1.72
Clang Static Analyzer	279.1
IKOS	1.2
Frama-C with Frama-Clang plugin	Aluminium-20160502

TABLE IV: The studied free and open source static analysis tools that detect Buffer Errors in C++

- CVE-2008-0985: Android web browser's webkit rendering engine webkit-522-android-m3-rc20
- CVE-2012-4190: Mozilla Firefox web browser 16.0
- CVE-2014-1705: Google Chrome web browser 33.0.1750.165 V8 JavaScript engine
- CVE-2014-1710: Google Chrome web browser 33.0.1750.15
- CVE-2014-3201: Google Chrome web browser 37.0.2062.94 Blink rendering engine.
- CVE-2016-5182: Google Chrome web browser 54.0.2840.85 Blink rendering engine.
- CVE-2016-5199: Google Chrome web browser 55.0.2883.83
- CVE-2016-5200: Google Chrome web browser 55.0.2883.83 V8 JavaScript engine
- CVE-2017-5014: Google Chrome web browser 56.0.2924.86

3) Running static analysis tools against the source code of the vulnerable apps: All source code of each app was tested through the static analyzers. Each tool was run several times using different options and flags. Then, automated scripts have been built to run tests automatically and save results to files.

4) Analyzing the results: in this step, we analyzed the results that were stored to files. Since most of the tools could report errors with the source code file path, error type, and error line number, we opened the source code files and traced and analyzed all reported errors. All collected source code and scripts could be found here [15]

B. Results

It was found that Flawfinder, RATS, and Cppcheck static analysis tools are easy to use. They could be executed through a command line interface, and accepting a list of files or projects directories to test with a set of options as parameters. Also, these tools show their results by default in the system's command line. The results contain the files path, the lines of code that are suspected to have vulnerabilities, and descriptions of the potential issues.

Clang Static Analyzer needs to be integrated into the build process. Frama-C and IKOS were built on top of LLVM/Clang, they accept source files and they are not meant to be integrated into a build process. To analyze source files with Frama-C, source files in our case study need to be preprocessed first, then preprocessed C++ files could be fed to Frama-C using Frama-Clang plugin. IKOS was developed by NASA to analyze flight systems, and it could convert C++ source files to LLVM bitcode first which could be analyzed by the tool then. Similar to Frama-C the files need a lot of preprocessing settings when analyzing source files in complex apps. However, setting preprocessing for source files manually is hard to be achieved with large and complex apps such as Google Chrome and Mozilla Firefox.

Our analysis reveals that none of the tools were able to detect any of the studied vulnerabilities. Table V shows the results of our analysis. In Table V, we include a number to refer to the reason why a tool cannot determine the specific vulnerability.

C. Discussion

In this subsection we want to discuss possible reasons for why the studied free and open source static analysis tools could not detect the nine vulnerabilities. In order to do that we look at characteristics of the techniques underlying the tools (in the context of `Buffer Error` vulnerabilities). In total we present six such characteristics (the numbers in the enumeration below corresponds to the numbers in Table V):

CVE ID	RATS	Flaw finder	Cpp check	Clang Static Analyzer	IKOS	Frama-C
CVE-2008-0985	1,2,3	1,2,3	1,2,4	1,2,5	1,6	1,6
CVE-2012-4190	2,3	2,3	2,4	2,5	6	6
CVE-2014-1705	1,2,3	1,2,3	1,2,4	1,2,5	1,6	1,6
CVE-2014-1710	1,2,3	1,2,3	1,2,4	1,2,5	1,6	1,6
CVE-2014-3201	1,2,3	1,2,3	1,2,4	1,2,5	1,6	1,6
CVE-2016-5182	1,2,3	1,2,3	1,2,4	1,2,5	1,6	1,6
CVE-2016-5199	1,2,3	1,2,3	1,2,4	1,2,5	1,6	1,6
CVE-2016-5200	1,2,3	1,2,3	1,2,4	1,2,5	1,6	1,6
CVE-2017-5014	1,2,3	1,2,3	1,2,4	1,2,5	1,6	1,6

TABLE V: Tested static analysis tools results: The numbers indicates the reason in the enumerated list below which explains why the tool could not find the vulnerability.

1) Unable to examine tainted data from code written in another language: In our case study, almost always the tainted data (data from an untrusted source) comes from Java (except CVE-2012-4190), while the static analysis tools only examine the C++ native code. None of the six tools examined are able to determine tainted data that comes from code written in Java.
2) Unable to keep track of pointer operations: The tool is unable to keep track of data when a buffer is manipulated using non trivial pointer operations (CVE-2017-5014), or when a null-pointer is dereferenced (CVE-2012-4190). Almost all of the studied nine vulnerabilities involves some sort of pointer operation.
3) Simple lexical analysis only instead of semantic analysis: RATS and Flawfinder only perform simple pattern matching, which is basically tokenizing the source code and looking for tokens that are well-known to cause `Buffer Errors`. They mostly focus on tokens that could potentially be a source or sink to buffer overflow such as dangerous functions like the **strcpy** function. However, the sink has more compound settings than the well-known set of tokens. For example, in CVE-2008-0985, the buffer is copied to the sink inside a loop that involves pointer indirection. Although, string pattern matching could be a powerful technique that could be enhanced to catch this kind of complicated sink by combining high level semantic analysis methods, it should not be used solely. In general, Flawfinder and RATS, have high false positive rate, since they always report vulnerabilities based on simple lexical analysis without semantically analyzing the code.
4) Inadequate semantic analysis: Cppcheck is an example of a static analysis tool that uses inadequate semantic analysis. Unlike RATS or Flawfinder, it uses semantic analysis based on Abstract Syntax Trees (AST) and it also utilizes control-flow analysis. However, it does not always perform control flow analysis on all situations and sometimes it assumes that all statements are reachable. Having inadequate semantic analysis may lead to miss problems related to buffer allocation and validation or pointer dereferencing. For the CVE-2012-4190 vulnerability, Cppcheck discovered two NULL pointer dereference issues as they are passed as parameter and used without checking if they are NULL (we checked and found that they are true positives, but not reported to NVD). However, it did not discover the NULL pointer dereference that was discovered in CVE-2012-4190 and reported to NVD.
5) Limited scope of analysis: Clang Static Analyzer performs a high level AST analysis. However, since it uses the AST generated by Clang, the analysis is performed at the compilation unit level. That is, it uses inter-procedural analysis, however it does not support inter-procedural analysis for cross-translation-unit. In fact, almost all of the vulnerabilities in our case study are inter-file/inter-procedural.
6) Tied to a specific compiler: IKOS and Frama-C directly invoke Clang/LLVM compiler framework that supports both C and C++. Therefore, they cannot analyze Android apps that uses specific toolchains. In addition, these tools are difficult to be integrated into build systems. These tools, however, perform sophisticated semantic, data flow, control flow, pointer operation, and inter-procedural analysis in a robust fashion. However, because they are unable to analyze the native code in Android apps, they are unable to accurately determine the `Buffer Errors` in our case study.

From the reasoning above we can see that an ideal static analysis tool to detect `Buffer Errors` in Android apps will:

- performs cross-language analysis to be able to understand both Java and native code contexts.
- utilizes tainted data flow as inference algorithm.
- employs sophisticated semantic, data flow, control flow and pointer operation analysis
- employs Inter-procedural/Inter-file analysis

VI. Threats to Validity

One possible threat to external validity is that we did not test many commercial tools that use their own parser which might be more accurate. Though, we tried to analyze studied vulnerabilities using CodeSonar. We got an academic license of CodeSonar, which only allows to analyze one million lines

of code. However, the analyzed apps in our study includes several million lines of codes.

In addition, knowing that most of the studied vulnerabilities involve pointer indirection, we know from past research that it could be hard to discover them by most of the commercial tools [39]. For instance, the Heartbleed is a pointer indirection vulnerability in OpenSSL [40] that was not discovered by commercial tools, such as Coverity Code Advisor, Code Sonar, Klocwork, and Veracode [39]. Commercial tools were not able to find Heartbleed as they are unsound. In fact, Heartbleed vulnerability was hard to be found due to using multiple level of pointer indirection and the complexity of the execution path from the buffer allocation to buffer misuse which is similar to our case. Additionally, we qualify that our experiments are done only on free and open source tools, by clearly stating "free and open source" in the title, abstract and conclusion to avoid any misunderstanding.

Another threat to validity is that some apps, such as Google Chrome, contains some closed source parts that might include the vulnerabilty. However in our study , we ensured that all vulnerabilities reside in the open source component (e.g. V8 and webkit). We browsed the source code as we know where the vulnerabilities exist from NVD website.

There is a threat that the vulnerability records we analyzed may not be related to Android. Hence. we manually inspected and studied all collected records from NVD to ensure that they related to the Android ecosystem and removed all biased records. Also, we manually reviewed static analysis tool results to guarantee the correctness. In order to address any threat to internal validity (i.e., mistakes we could have made) and for the ability for anyone to replicate our experiments, we provide all the data in our experiments [15].

Another limitation of our study is the sample size is small, yet we have tested all available vulnerabilities that occurred in open source apps. There is a possibility that the tested vulnerabilities may not be representative of other `Buffer Error` in general. However, when observing the descriptions of other `Buffer Error` vulnerabilities in our dataset that occurred in closed source apps, we found some similarities with our sample, such as apps are written in C++ and `Buffer Errors` happened due to input that was read from untrusted sources.

VII. Related Work

A. Android Vulnerabilities Analysis

To support the development of vulnerability detection models, several studies have been conducted to understand the patterns of the vulnerability in Android software. Huang et al. [41] studied the mobile vulnerability market to reveal the unique vulnerability patterns of mobile software. The study shows that the vulnerabilities in the Android market are more exploitable than in the entire market; and the exploitation impact is higher based on CVSS metrics. Our research took further steps to reveal the pattern of these vulnerabilities in Android apps.

B. Android Static Analysis Tools

Although, extensive work has been conducted to introduce static analysis tools for Android vulnerabilities and malicious behavior, most of the studies focus on permissions and information leakage issues and do not address `Buffer Error` [23]. For instance, the SCanDroid tool [42] performs a data-flow analysis of installed Android apps, to track inter-component communication through intents in order to detect the potential violation of permission through a coalition of applications. Similarly, Quire [43] tracks permissions through the IPC call to prevent privilege attacks among applications. In addition, number of tools have been developed to detect private data leakage. TaintDroid [44] preforms dynamic taint analysis, while FLOWDROID [45], LeakMiner [46], and AndroidLeaks [47] are static analysis approaches to detect data leakage. VulHunter [48] is a static analysis framework to support vulnerability detection for Android apps by extracting information from applications. It aims to detect five types of vulnerabilities that are related to information leakage and permissions violations. In fact, there is a lack of research on static analysis tools that target native code in Android apps. Such mapping between the native and Java contexts will improve the discovery of Android vulnerabilities by constructing a better call graph. The issue of the lack of static analysis methods that link Java code to native code has been recently highlighted [49] [50]. Our research looks at `Buffer Error` vulnerabilities. The above tools do not address `Buffer Error` vulnerabilities, and we showed that no open source static analysis tools could statically determine `Buffer Error` vulnerabilities in Android apps.

C. Dynamic Analysis for Buffer Errors in Android

`Buffer Error` is a well-known problem that still resurfaces. Several run-time approaches have been proposed to mitigate `Buffer Error` vulnerabilities in Android. Hardware-based No eXecute (NX) technique has been added to Android 2.3 Gingerbread to prevent code execution on the stack and heap. Another approach is using compiler extensions, such as ProPolice, which is improved over StackGuard [51], to protect against stack-based buffer overflow. ProPolice has been introduced in Android 1.5 CupCake [52]. In Android Jelly Beans version 4.2 and later, a new feature has been added that protects against `Buffer Errors`, as all applications and system libraries during compile time are checked with FORTIFY_SOURCE feature [52], which detects and stops a certain types of `Buffer Errors`. Defense side obfuscation is another approach to mitigate `Buffer Errors`. An example of this approach is Address Space Layout Randomization (ASLR) that randomizes memory addresses of stack and heap each time the memory is allocated for a process, such that finding executable code becomes unreliable. ASLR has been introduced in Android 4.0 Ice Cream Sandwich [52]. However, overcoming the above mentioned techniques is possible and this was discussed in [53] [54]. Real exploitation and code execution have already been proven in the wild [55]. *In fact, the dynamic approaches that prevent* `Buffer Error` *attacks*

do not eliminate the vulnerabilities from the source code. While we studied static analysis tools that could potentially identify the issue earlier. One advantage of static analysis is that security vulnerabilities can be removed before code is deployed, which reduces the cost of the risk.

D. Static Analysis for Buffer Errors in C/C++

Several static analysis prototypes have been introduced in academia to detect `Buffer Errors`, yet the majority of the proposed methods target C programming language. For instance, Mjolnir [56], MACKE [57], MESCs [58], Wagner et al. [59], Kim et al. [60], Hackett et al. [61], CSSV [32], Vulncheck [62] [32], Avots et al. [63], and Livshits and Lam [64] are all academic prototype that only focus in C. However, it is noticed that little attention has been directed lately to study and propose methods to target C++ which could be able to address the OOP model, such as Marple [65] and Li et al. [66].

E. Inter-language Static Analysis

In fact, studying inter-language analysis tools has been investigated in the past. Su and Wassermann [67] introduced an algorithm for checking type safety across a foreign function interface, such as between OCaml and C. The system prevents foreign function calls in C from introducing type and memory bugs into an safe language such as OCaml. Siefers et al. [68] presented static analysis techniques to detect bugs in programs using JNI. Their analysis detects bugs such as exception handling, memory leaks, and invalid local references. Li and Tan [69] proposed a static analysis framework to examine exceptions and report errors in JNI programs. Their framework can be applied to other foreign function interfaces, including the Python/C interface and the OCaml/C interface. Tan and Croft [70] have conducted an empirical security study on the native code portion of Suns JDK 1.6. They used ITS4, Flawfinder, and Splint to carry out the analyses. It was mentioned the importance of building inter-language analysis across Java and C, as most existing tools are limited to code written in a single language. However, our study focuses on `Buffer Errors` particularly which have not been considered before in such tools. Also, our study suggests that unlike Java, when building a tool for Android apps, the multiple entry points should be considered as well.

F. Evaluation of Static Analysis Tools for Buffer Errors

Although, a few studies have evaluated the use of static analysis tools for `Buffer Error` detection [29] [30] [31] [71], to the best of our knowledge, this is the first empirical study that evaluated these tools against known `Buffer Error` in the Android apps domain. Our results expose the need for more advanced static analysis tools to detect `Buffer Errors` in Android apps taking into consideration Android app nature.

VIII. Conclusions

Our study found that `Buffer Errors` were the most frequent type of vulnerability occurring in Android apps, and they are easy to exploit. Therefore, we decided to study the effectiveness of state-of-the-art free and open source static analysis tools for detecting `Buffer Error` vulnerabilities in Android apps. In addition, our goal was to understand how `Buffer Errors` happen in Android context.Our findings indicate that there is a lack of free and open source static analysis tools that target Android, particularly to detect `Buffer Errors`. To the best of our knowledge, there is no free and open source static analysis tool for Android that analyzes both native code (which may introduce `Buffer Errors`), and Java code (which may originate the unsanitized input). Thus, general static analysis tools for native code are usually used by Android developers. However, current free and open source static analysis tools for C++ could not detect all `Buffer Errors` in Android.

Also, our study found some pattern of characteristics for `Buffer Errors` that occurred in Android apps, such as occur within C++, due to pointers indirection, untrusted input travels from Java to C++ where buffer is misused. Thus, by utilizing these patterns, static analysis tools could be built to work more efficiently. The experimental results show that such an evaluation brings an important contribution characterizing an effective static analysis tool to detect `Buffer Errors` in Android apps. Therefore, we conclude that an efficient static analysis technique for detecting `Buffer Errors` in Android (1) should perform a taint analysis that traces inputs to a program from outside, (2) should involve inter-language analysis (this could effect other modern apps in other platforms that include native C++ as library), (3) has better understanding of the code semantics and involve pointer operation analysis (this is a general problem that effect apps in all kinds of platforms.), and (4) should perform inter-procedural/inter-file analysis to analyze data travel cross procedures.

REFERENCES

[1] IDC Research Inc, "Smartphone OS market share, 2017 Q1," May 2017. [Online]. Available: http://www.idc.com/prodserv/smartphone-os-market-share.jsp

[2] Statista, "Number of available applications in the Google Play Store from December 2009 to June 2017," Jun. 2017. [Online]. Available: https://www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play-store/

[3] Statista, "Cumulative number of apps downloaded from the Google Play as of May 2016 (in billions)," May 2016. [Online]. Available: https://www.statista.com/statistics/281106/number-of-android-app-downloads-from-google-play/

[4] Kantar Media, "Apps environment: Research report," *Ofcom*, 2014. [Online]. Available: https://www.ofcom.org.uk/__data/assets/pdf_file/0015/40290/apps_environment.pdf

[5] N. J. Percoco and S. Schulte, "Adventures in bouncerland: failures of automated malware detection with in mobile application markets," *Black Hat USA*, 2012.

[6] Symantec, "2015 Internet security threat report," *Internet Security Threat Report*, vol. 20, p. 119, April 2015. [Online]. Available: https://www4.symantec.com/mktginfo/whitepaper/ISTR/21347932_GA-internet-security-threat-report-volume-20-2015-social_v2.pdf

[7] Risk Based Security, "Vulndb quickview 2015 vulnerability trends," 2015. [Online]. Available: https://www.riskbasedsecurity.com/vulndb-quickview-2015-vulnerability-trends/

[8] Y. Younan, "25 years of vulnerabilities: 1988–2012," *Sourcefire Vulnerability Research Team*, 2013.

[9] Risk Based Security, "29% increase in vulnerabilities already disclosed In 2017," May 2017. [Online]. Available: https://www.riskbasedsecurity.com/2017/05/29-increase-in-vulnerabilities-already-disclosed-in-2017/

[10] Hewlett Packard Enterprise, "HPE security research cyber risk report 2016," 2016.

[11] Dimensional Research, "Mobile app usage and abandonment survey," Jan 2015.

[12] N. Nagappan and T. Ball, "Static analysis tools as early indicators of pre-release defect density," in *Proceedings. 27th International Conference on Software Engineering, 2005. ICSE 2005.*, May 2005, pp. 580–586.

[13] M. Soni, "Defect prevention: Reducing costs and enhancing quality," *iSixSigma. com*, vol. 19, 2006.

[14] National Vulnerability Database (NVD). [Online]. Available: https://nvd.nist.gov/

[15] B. Aloraini and M. Nagappan, "Evaluating state-of-the-art free and open source static analysis tools against buffer errors in Android apps," 2017. [Online]. Available: https://github.com/BushraAloraini/Android-Vulnerabilities

[16] OSVDB. [Online]. Available: https://blog.osvdb.org/

[17] CERT. [Online]. Available: http://www.cert.org/

[18] Security Focus. [Online]. Available: http://www.securityfocus.com/

[19] Common Vulnerabilities and Exposures (CVE). [Online]. Available: https://cve.mitre.org/

[20] Japan Vulnerability Note (JVN). [Online]. Available: https://jvn.jp/en/

[21] W. Dormann, "Finding Android SSL vulnerabilities with CERT Tapioca," Sept 2014. [Online]. Available: https://insights.sei.cmu.edu/cert/2014/09/-finding-android-ssl-vulnerabilities-with-cert-tapioca.html

[22] Common Weakness Enumeration (CWE). [Online]. Available: https://cwe.mitre.org/

[23] A. Sadeghi, H. Bagheri, J. Garcia, and S. Malek, "A taxonomy and qualitative comparison of program analysis techniques for security assessment of Android software," *IEEE Transactions on Software Engineering*, vol. 43, no. 6, pp. 492–530, June 2017.

[24] Common Vulnerability Scoring System (CVSS). [Online]. Available: https://www.first.org/cvss/v2/guide

[25] Android, "Application fundamentals," 2016. [Online]. Available: https://developer.android.com/guide/components/fundamentals.html

[26] Y. Zhou, Z. Wang, W. Zhou, and X. Jiang, "Hey, you, get off of my market: detecting malicious apps in official and alternative Android markets." in *NDSS*, vol. 25, no. 4, 2012, pp. 50–52.

[27] C. Qian, X. Luo, Y. Shao, and A. T. S. Chan, "On tracking information flows through JNI in Android applications," in *Proceedings of the 2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, ser. DSN '14. Washington, DC, USA: IEEE Computer Society, 2014, pp. 180–191. [Online]. Available: http://dx.doi.org/10.1109/DSN.2014.30

[28] O. Cinar and G. Allen, *Pro Android C++ with the NDK*. Springer, 2012.

[29] M. Zitser, R. Lippmann, and T. Leek, "Testing static analysis tools using exploitable buffer overflows from open source code," in *Proceedings of the 12th ACM SIGSOFT Twelfth International Symposium on Foundations of Software Engineering*, ser. SIGSOFT '04/FSE-12. New York, NY, USA: ACM, 2004, pp. 97–106.

[30] T. Hofer, "Evaluating static source code analysis tools," Master's thesis, School of Computer and Communications Science, Switzerland, 2010.

[31] L. Torri, G. Fachini, L. Steinfeld, V. Camara, L. Carro, É. Cota, P. Box, and P. Alegre, "An evaluation of free/open source static analysis tools applied to embedded software," in *2010 11th Latin American Test Workshop*, March 2010, pp. 1–6.

[32] H. Shahriar and M. Zulkernine, "Classification of static analysis-based buffer overflow detectors," in *Secure Software Integration and Reliability Improvement Companion (SSIRI-C), 2010 Fourth International Conference on*, June 2010, pp. 94–101.

[33] K. Vorobyov and P. Krishna, "Comparing model checking and static program analysis: A case study in error detection approaches," *Proc. SSV*, pp. 1–7, 2010.

[34] C. Cifuentes, C. Hoermann, N. Keynes, L. Li, S. Long, E. Mealy, M. Mounteney, and B. Scholz, "BegBunch: Benchmarking for C bug detection tools," in *Proceedings of the 2Nd International Workshop on Defects in Large Software Systems: Held in Conjunction with the ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2009)*, ser. DEFECTS '09. New York, NY, USA: ACM, 2009, pp. 16–20.

[35] Spinroot, "Static source code analysis tools for C," 2014. [Online]. Available: https://spinroot.com/static/

[36] P. Emanuelsson and U. Nilsson, "A comparative study of industrial static analysis tools," *Electron. Notes Theor. Comput. Sci.*, vol. 217, pp. 5–21, Jul. 2008. [Online]. Available: http://dx.doi.org/10.1016/j.entcs.2008.06.039

[37] G. Brat, J. A. Navas, N. Shi, and A. Venet, "IKOS: A framework for static analysis based on abstract interpretation," in *International Conference on Software Engineering and Formal Methods*. Springer, 2014, pp. 271–277.

[38] Frama-C. [Online]. Available: http://frama-c.com/index.html

[39] J. A. Kupsch, , and B. P. Miller, "Why do software assurance tools have problems finding bugs like Heartbleed?" 2014.

[40] Codenomicon, "The Heartbleed bug," 2014. [Online]. Available: http://heartbleed.com/

[41] K. Huang, J. Zhang, W. Tan, and Z. Feng, "An empirical analysis of contemporary Android mobile vulnerability market," in *2015 IEEE International Conference on Mobile Services*, June 2015, pp. 182–189.

[42] A. P. Fuchs, A. Chaudhuri, and J. Foster, "SCanDroid : Automated security certification of Android applications," *Technical report, University of Maryland*, vol. 10, no. November, p. 328, 2009.

[43] M. Dietz, S. Shekhar, Y. Pisetsky, A. Shu, and D. S. Wallach, "Quire: lightweight provenance for smart phone operating systems," vol. 271, no. 2012, 2011, p. 23.

[44] W. Enck, P. Gilbert, B.-G. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. N. Sheth, "TaintDroid: An information-flow tracking system for realtime privacy monitoring on smartphones," in *Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation*, ser. OSDI'10. Berkeley, CA, USA: USENIX Association, 2010, pp. 393–407. [Online]. Available: http://dl.acm.org/citation.cfm?id=1924943.1924971

[45] S. Arzt, S. Rasthofer, C. Fritz, E. Bodden, A. Bartel, J. Klein, Y. Le Traon, D. Octeau, and P. McDaniel, "FlowDroid: Precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for Android apps," *SIGPLAN Not.*, vol. 49, no. 6, pp. 259–269, Jun. 2014. [Online]. Available: http://doi.acm.org/10.1145/2666356.2594299

[46] Z. Yang and M. Yang, "LeakMiner: Detect information leakage on Android with static taint analysis," in *Software Engineering (WCSE), 2012 Third World Congress on*, Nov 2012, pp. 101–104.

[47] C. Gibler, J. Crussell, J. Erickson, and H. Chen, "AndroidLeaks: Automatically detecting potential privacy leaks in Android applications on a large scale," in *Proceedings of the 5th International Conference on Trust and Trustworthy Computing*, ser. TRUST'12. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 291–307.

[48] C. Qian, X. Luo, Y. Le, and G. Gu, "VulHunter: Toward discovering vulnerabilities in Android applications," *IEEE Micro*, vol. 35, no. 1, pp. 44–53, Jan 2015.

[49] P. Lantz and B. Johansson, "Towards bridging the gap between Dalvik bytecode and native code during static analysis of Android applications," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2015 International*, Aug 2015, pp. 587–593.

[50] V. Afonso, A. Bianchi, Y. Fratantonio, A. Doupe, M. Polino, P. de Geus, C. Kruegel, and G. Vigna, "Going native: Using a large-scale analysis of Android apps to create a practical native-code sandboxing policy," in *Proceedings of the ISOC Network and Distributed System Security Symposium (NDSS)*, San Diego, CA, February 2016.

[51] C. Cowan, C. Pu, D. Maier, H. Hintony, J. Walpole, P. Bakke, S. Beattie, A. Grier, P. Wagle, and Q. Zhang, "StackGuard: Automatic adaptive detection and prevention of buffer-overflow attacks," in *Proceedings of the 7th Conference on USENIX Security Symposium - Volume 7*, ser. SSYM'98. Berkeley, CA, USA: USENIX Association, 1998, pp. 5–5. [Online]. Available: http://dl.acm.org/citation.cfm?id=1267549.1267554

[52] Source Android, "Security enhancements in Android 1.5 through 4.1." [Online]. Available: https://source.android.com/security/enhancements/enhancements41

[53] J. J. Drake, Z. Lanier, C. Mulliner, P. O. Fora, S. A. Ridley, and G. Wicherski, *Android Hacker's Handbook*. John Wiley & Sons, 2014.

[54] H. Marco-Gisbert and I. Ripoll, "On the effectiveness of NX, SSP, RenewSSP, and ASLR against stack buffer overflows," in *2014 IEEE 13th International Symposium on Network Computing and Applications*, Aug 2014, pp. 145–152.

[55] G. Gong, "Exploiting heap corruption due to integer overflow in Android libcutils," *Black Hat USA*, 2015.

[56] M. Weber, V. Shah, and C. Ren, "A case study in detecting software security vulnerabilities using constraint optimization," in *Proceedings First IEEE International Workshop on Source Code Analysis and Manipulation*, 2001, pp. 1–11.

[57] S. Ognawala, M. Ochoa, A. Pretschner, and T. Limmer, "MACKE: Compositional analysis of low-level vulnerabilities with symbolic execution," in *2016 31st IEEE/ACM International Conference on Automated Software Engineering (ASE)*, Sept 2016, pp. 780–785.

[58] R. Gjomemo, P. H. Phung, E. Ballou, K. S. Namjoshi, V. N. Venkatakrishnan, and L. Zuck, "Leveraging static analysis tools for improving usability of memory error sanitization compilers," in *2016 IEEE International Conference on Software Quality, Reliability and Security (QRS)*, Aug 2016, pp. 323–334.

[59] D. Wagner, J. S. Foster, E. A. Brewer, and A. Aiken, "A first step towards automated detection of buffer overrun vulnerabilities." in *Network and Distributed System Security Symposium (NDSS)*, 2000, pp. 2000–02.

[60] Y. Kim, J. Lee, H. Han, and K. M. Choe, "Filtering false alarms of buffer overflow analysis using SMT solvers," *Information and Software Technology*, vol. 52, no. 2, pp. 210–219, 2010.

[61] B. Hackett, M. Das, D. Wang, and Z. Yang, "Modular checking for buffer overflows in the large," in *Proceedings of the 28th International Conference on Software Engineering*, ser. ICSE '06. New York, NY, USA: ACM, 2006, pp. 232–241. [Online]. Available: http://doi.acm.org/10.1145/1134285.1134319

[62] A. I. Sotirov, "Automatic vulnerability detection using static source code analysis," Master's thesis, The University of Alabama, May 2005.

[63] D. Avots, M. Dalton, V. B. Livshits, and M. S. Lam, "Improving software security with a C pointer analysis," in *Proceedings. 27th International Conference on Software Engineering, 2005. ICSE 2005.*, May 2005, pp. 332–341.

[64] V. B. Livshits and M. S. Lam, "Tracking pointers with path and context sensitivity for bug detection in C programs," *SIGSOFT Softw. Eng. Notes*, vol. 28, no. 5, pp. 317–326, Sep. 2003. [Online]. Available: http://doi.acm.org/10.1145/949952.940114

[65] W. Le and M. L. Soffa, "Marple: A demand-driven path-sensitive buffer overflow detector," in *Proceedings of the 16th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, ser. SIGSOFT '08/FSE-16. New York, NY, USA: ACM, 2008, pp. 272–282. [Online]. Available: http://doi.acm.org/10.1145/1453101.1453137

[66] L. Li, C. Cifuentes, and N. Keynes, "Practical and effective symbolic analysis for buffer overflow detection," in *Proceedings of the Eighteenth ACM SIGSOFT International Symposium on Foundations of Software Engineering*, ser. FSE '10. New York, NY, USA: ACM, 2010, pp. 317–326. [Online]. Available: http://doi.acm.org/10.1145/1882291.1882338

[67] Z. Su and G. Wassermann, "The essence of command injection attacks in web applications," in *Conference Record of the 33rd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, ser. POPL '06. New York, NY, USA: ACM, 2006, pp. 372–382. [Online]. Available: http://doi.acm.org/10.1145/1111037.1111070

[68] J. Siefers, G. Tan, and G. Morrisett, "Robusta: Taming the native beast of the JVM," in *Proceedings of the 17th ACM Conference on Computer and Communications Security*, ser. CCS '10. New York, NY, USA: ACM, 2010, pp. 201–211. [Online]. Available: http://doi.acm.org/10.1145/1866307.1866331

[69] G. Tan and J. Croft, "An empirical security study of the native code in the JDK," in *Proceedings of the 17th Conference on Security Symposium*, ser. SS'08. Berkeley, CA, USA: USENIX Association, 2008, pp. 365–377. [Online]. Available: http://dl.acm.org/citation.cfm?id=1496711.1496736

[70] L. Davi, A. Dmitrienko, A.-R. Sadeghi, and M. Winandy, "Privilege escalation attacks on Android," in *Proceedings of the 13th International Conference on Information Security*, ser. ISC'10. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 346–360. [Online]. Available: http://dl.acm.org/citation.cfm?id=1949317.1949356

[71] D. Pozza, R. Sisto, L. Durante, and A. Valenzano, "Comparing lexical analysis tools for buffer overflow detection in network software," in *Communication System Software and Middleware, 2006. Comsware 2006. First International Conference on*, 2006, pp. 1–7.