Actas de las segundas Jornadas Nacionales de Investigación en Ciberseguridad

Granada, 15, 16 y 17 de junio de 2016

Universidad de Granada
Abstract—Any type of software, from desktop to mobile applications, is prone to contain defects that can lead to vulnerabilities. These defects, when exploited, may put in risk the integrity, confidentiality and availability of the software. Security auditing methodologies help to reduce at some level of confidence these risks. With the explosion of mobile applications for daily activities like checking email, news, social networks, or even managing bank accounts, guaranteeing an acceptable level of application security becomes critical for the usage and trust of mobile services. In this paper, we review and classify OWASP 2014 Top Ten mobile risks in analysis blocks. Based on the blocks classification, we propose a methodology to security audit mobile software applications. We demonstrate the effectiveness of the proposed methodology by auditing the same mobile application in Google’s Android and Apple’s iOS platforms. We found several application design flaws and vulnerabilities. These vulnerabilities, when exploited, may lead to software vulnerabilities. We understand a software vulnerability as software flaws that might be accessible to an attacker who can wittingly exploit them [3]. Examples of common software vulnerabilities are buffer overflows, cross-site scripting (XSS) and SQL injection, among others [4].

Developing vulnerability-free software might be impossible to achieve due to time and economical restrictions. However, there exist different methods to surface software vulnerabilities, such as data flow analysis, taint analysis or fuzzy testing, for naming a few [5]. The more vulnerability detection techniques are used, the higher the probability that the software has less remaining vulnerabilities. However, the application of these techniques are usually expensive and burdensome.

Due to the exponential growth of mobile applications ecosystem, there are thousands of applications available to conduct any type of daily activity, from checking email, news, or social networks to managing bank accounts. Therefore, users rely on these applications sensitive data. Hence, any vulnerability in mobile applications can put in risk the user’s privacy. Besides, mobile applications ecosystem promotes the first one-of-a-kind application released. So, mobile applications are usually released to the market as soon as possible, castigating critical features like security and reliability [1], [2].

In order to cope with vulnerabilities and security threats in software products, security auditing methodologies have been proposed to certify certain security level of confidence. To the best of our knowledge, there are neither recognised consensus nor standards about how to perform a security audit in mobile applications. However, there are general best practices like the IEEE standard for software reviews and audits [6]. However, these general best practices are not considering the idiosyncrasy of mobile applications.

The most closed approach to develop some best practices or guidelines to develop and maintain secure mobile applications is managed by Open Web Application Security Project (OWASP). OWASP offers developers and security teams resources to build and maintain secure mobile applications [7], but no formal guideline is provided. However, it is important to recognise the effort conducted by OWASP in order to list the most representative risks during mobile application development.

In this paper, we introduce an auditing methodology for mobile applications covering the risks pointed out by OWASP. We firstly review, analyse and classify the OWASP 2014 Top Ten Mobile Risks [7] under different analysis blocks. Then, we propose a methodology that uses these analysis blocks to audit the security of mobile applications. The methodology is evaluated in practice by auditing the same functional application developed for two different mobile platforms: Google’s Android and Apple’s iOS. We focus on these platforms since they lead the global smartphone market [8]. Our methodology revealed several application design flaws related to implementation and handling of data, as well as information leakage. We found that sensitive application user data may be easily reachable by an attacker or third-party applications, even remotely.

The outline of this paper is as follows. Section II visits the related work. Section III summarises the OWASP Top Ten Mobile Risks, and the analysis blocks identified. The methodology proposed is introduced in Section IV. A case of use is presented and discussed in Section V. Section VI
concludes the paper.

II. RELATED WORK

As noted earlier, to the best of our knowledge, there is not a methodology for security auditing on mobile applications. Recent published works are mainly focused on the application usability [9] or rapid development [10], [11], rather than security aspects of the application itself. Nevertheless, it is worth to mention [12], where a methodology to empirical analysis of permission-based security models and enhancements to Android is introduced.

Although, several tools are available to analyse concrete security aspects – mainly related with the confidentiality – in mobile platforms, they have specific usages. However, they can be used together to conduct a security audit in mobile applications. For instance, ProtectMyPrivacy [13] detects access to private data and substitutes it by anonymised data under user requests in Apple’s iOS. Similarly, other tools, such as PiOS [14], PSiOS [15], DiOS [16], perform static or dynamic analysis into applications to detect sensitive data leakage.

Likewise, TaintDroid [17] monitors privacy data using an information flow tracking system. SCanDroid [18] allows to check whether data flow through an application is consistent with its security specifications. Other tools, such as DroidScope [19], Andrubis [20], and Google’s Bouncer [21], focus on analyzing mobile applications to detect malicious behaviors (i.e., malware).

III. ANALYSIS BLOCKS TO IDENTIFY MOBILE RISKS

In this section, we summarize the OWASP Top Ten Mobile Risks for 2014 [7] to identify where the common defects are found and how they can be surfaced during an application security auditing. The OWASP Top Ten Mobile Risks list the most common risks according to different factors such as threatening agents, attack vectors, weaknesses, technical impact, and business impact. OWASP classifies common mobile security risks and provide assurance controls to reduce its impact or likelihood of exploitation. The OWASP risks are briefly introduced:

(M1) **Weak Server Side Controls** considers the risks from third-party components like backend servers required by most mobile applications. This risk includes insecure server configuration, authentication errors, session management weakness, access control vulnerabilities, just to mention few.

(M2) **Insecure Data Storage** considers the potential risks caused by vulnerabilities on the data storage in the mobile device. These risks can cause information leakage.

(M3) **Insufficient Transport Layer Protection** identifies vulnerabilities like non-encrypted transport layer communications, the usage of weak cryptographic algorithms or the acceptance of unauthorized certifications.

(M4) **Unintended Data Leakage** risk considers the unknown potential vulnerabilities in the data management by the application or operating system.

(M5) **Poor Authorization and Authentication** considers risks associated with unacceptable authentication assumptions like that only authenticated users can send requests to the server without further validation of the user or weak authentication protocols.

(M6) **Broken Cryptography** includes the incorrect usage of the encryption/decryption process or the usage of weak in nature cryptographic algorithms.

(M7) **Client Side Injection** gathers the risks of not validating user input data, avoiding code injection.

(M8) **Security Decisions via Untrusted Inputs** brings together the risk associated with accepting any type of input source. This is specially relevant in the case of Inter Process Communication (IPC) mechanisms.

(M9) **Improper Session Handling** collects the defects that can lead in vulnerabilities in handling user session.

(M10) **Lack of Binary Protection** identified as risk using untrusted source repositories to host the mobile application code. Any source repository out of physical control by the organization has to be considered insecure.

Considering the target of the attack vector, these risks can be grouped into those which objectives are the server and communications, and those targeting the application itself.

Risks affecting server and communications include M1, M3, M5, M6 and M9. All these risks have in common that they are considering risks in the communication between the mobile application and the backend servers or lack of poor authorization/authentication in the server side when it is receiving client requests. For instance, self-signed or expired certificates, vulnerable SSL implementations, or encrypted communications weak and even unencrypted.

Those targeting the application are M2, M4, M7, M8 and M10. These risks refer to vulnerabilities like data application files containing sensitive data (such as password or username) stored in plain text and accessible by third-party applications, unknown OS vulnerabilities that may unintendedly compromise an application, and the deployment of executable code without additional protections, such as obfuscation or packing.

We have identified five analysis blocks which can detect a risk area according to application logic. Namely, the blocks are: **Environment Analysis**, **Connections**, **Sensitive Data**, **Application Own Data**, and **Application Structure**. Each block clearly delimits the aspects to be analysed in a mobile application to detect likely vulnerabilities. The risks of OWASP can belong to several blocks of analysis as OWASP classifies risks in light of the threats beyond that analysis purposes. For instance, the M5 risk accounts for authentication not only on the application side (i.e., a user must be logged in to access to concrete backend services), which belongs to **Application Structure**; but also on the server side (i.e., a backend service must not be anonymously reached), which is covered by **Environment Analysis**. Similarly, M5 is also covered by **Application Own Data** since an application may ask for a weak login password to execute.

**Environment Analysis**: This analysis block covers the OWASP Top Ten Mobile Risks M1, M5, M7, and M10. The environment analysis block encompasses the study of the mobile device (i.e., if the device is using a custom or modified firmware, platform versions, etc.), the application backend server, and the information related to the application (i.e., the market from which it was downloaded, who developed it, whether it can be decrypted/reversed, etc.).

Application backend servers are used by mobile applications for different purposes: Either to provide extra func-
tionalities, to be customized/updated, or to properly execute. Although remote servers might be considered to be out of scope when auditing an application, note that application data can be compromised whether these servers are compromised somehow. Thus, this analysis block verifies whether confidentiality and integrity properties are held.

**Connections:** This block relates to OWASP Top Ten Mobile Risks M1, M3, M4, and M7. It encompasses the analysis of all types of communications available in the mobile device, and thus likely to be used by an application. They can be from device to device, device to the Internet or device to computer. Communications can be performed using different technologies, such as mobile data, Wi-Fi, IRDA, Bluetooth, or NFC.

Four main questions are addressed by this analysis block:

- **Where is the application connecting to?** The destination of any connection performed by the application has to be located and logged. For instance, the IP address or URL when connecting through the Internet, or the target device name and MAC when connecting through a Bluetooth, IRDA, or NFC connection.

- **What is the protocol being used on the connection?** The protocol involved in any connection has to be known to detect whether is vulnerable or otherwise secure. For instance, an Internet connection using HTTP instead of HTTPS protocol enables unencrypted instead of encrypted communication.

- **What data are being transmitted in the connection?** Whether user data are being sent, we need to know what specific files are being sent, what commands are being used, etc.

- **Does the application perform certificate pinning?** Otherwise, even though the application could perform a presumably secure connection, it could be compromised using self-signed certificates. An analysis of applications potentially vulnerable to Man-in-the-middle attacks on SSL/TLS connections is introduced in [22].

**Sensitive Data:** This block only relates to OWASP Top Ten Mobile Risks M4. Sensitive data includes data managed by the operating system, such as data related to device hardware Identification (UDID, MAC, IMEI, ...), its status (Wi-Fi and Bluetooth on/off, 2G/3G/4G connection, etc.), or other data related to native OS applications (contacts, calendar events, messaging, etc.).

It is common that mobile applications request access to (or even send) sensitive user data. For instance, *Path* [23], *Hipster* [24] or *Twitter* [25]) applications.

A relevant question arises: **What is the motivation behind these data accesses?** Normally it is just for legitimate or commercial purposes, although malicious applications get this kind of data for spurious reasons. A study performed over Android applications [20] shows that, in general, applications requesting a high number of permissions are likely malicious software (i.e., malware).

**Application Own Data:** This analysis block fits to OWASP Top Ten Mobile Risks M2, M4, M5, M6, and M9. Application own data refers to data generated and managed by the application itself. For instance, format files such as XML, PList, SQLite are commonly used to store the application own data. These data are generally saved in files in the device, and may include credentials, images, documents, logs, or any other kind of data. So, any person having physical access to the device, a third-party application or a malware application installed on the device, could access this data. Thus, any data generated/handled by an application must be securely saved to prevent unauthorised access.

**Application Structure:** This analysis block covers all risks on the OWASP Top Ten except M1. This point covers everything related to the design and implementation of the application. In fact, this analysis block uses the binary code disassembled or source code of the application retrieved in the *Environment Analysis* block. Here, an analysis of the source/binary code is required.

Any vulnerability/operability analysis performed on the application is also covered by this block. This is indeed one of the most important blocks, since it provides information related to the application. For instance, the application must work properly and according to what it is expected to do. This analysis block detects when an application shows an unexpected behaviour, drains the battery or crashes during execution due to an exception. Apart from a dynamic application analysis, a static analysis is also carried out. Hard-coded strings must be studied to check whether they contain relevant information, such as developer credentials, URLs, or databases passwords.

Table I summarises the relationships between the analysis blocks and the OWASP risks.

<table>
<thead>
<tr>
<th>Environment Analysis</th>
<th>Connections</th>
<th>Sensitive Data</th>
<th>Application Own Data</th>
<th>Application Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M4</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M5</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M6</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M7</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M8</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M9</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M10</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**IV. A SECURITY ASSESSMENT METHODOLOGY**

In this section, we present a methodology to conduct security audit of mobile applications covering the analysis blocks presented in Section III, and thus, the OWASP Top Ten Mobile Risks. This methodology allows finding vulnerabilities, to detect malicious behaviours, and even to analyse application performance. The methodology also aims at covering a broad number of aspects when auditing a mobile application, such as detecting coding flaws or insecure design.

The methodology consists of three main phases subdivided into several tasks: **Pre-runtime**, which encompasses the analysis of the application before executing it; **Runtime**, which describes the analysis performed to an application under execution; and **Post-runtime**, which encompasses the analysis tasks after an application execution. Fig. 1 depicts a flowchart of the proposed methodology.

**A. Pre-RunTime Phase**

This phase is performed before executing the application. It can be divided into two tasks: **Preliminary Analysis**, and
**Static Analysis.** The former is related to Environment Analysis block, while the latter applies to Application Structure.

The Preliminary Analysis task describes the process of information gathering, decrypting, and reversing the application. Information gathering is the process of obtaining information about the application. Information can be passive, when it refers to the information collected non-intrusively from the device or remote servers (e.g., developer’s, application description, or server information), or active, when the information is collected intrusively. For instance, scanning tools are used in the remote servers to obtain the IP, OS, public services and potential vulnerabilities. Decrypting and reversing are the processes to fully obtain the binary code in a (somehow) human-readable form. For instance, Java classes and related content are obtained when analysing a Google’s Android application. Similarly, assembly code is obtained when analysing an Apple’s iOS application.

The Static Analysis is decomposed into multiple tasks, such as analysing all strings (readable text) retrieved from the source code, binary code, or configuration files, studying the source/binary code to find vulnerabilities (e.g., code coverage analysis, authorization logic schemes, or quality test), analysing the design of the application to check its abstraction, modularity, or extensible considerations, or to analyse the sensitive data accessed by the application. Note that the source code or the disassembled binary code is needed.

**B. Runtime Phase**

This phase is performed during the execution of the application. It can be divided into several tasks: File Handling, Data Sent, Access to Sensitive Data, Network Behaviour, and Running Code. In this phase, Connections, Application Own Data, and Sensitive Data analysis blocks are involved.

**File Handling** covers the process of writing, reading and handling (i.e., creation, deletion or movement) files within the application using either third-party applications or native functions of the OS. For instance, whether an Android application stores a file on the external memory, this file could be read by any other application. However, when the file is stored within the application sandbox, it is only accessible by the own application. Unlike Google’s Android, Apple’s iOS always stores files in the application sandbox, which restricts unauthorised file access.

**Data Sent** refers to the network packets being transmitted by the application. Third-party applications such as WireShark, Burp, or tcptrace can be used in this step, independently from the OS in which the application is running on. To check whether the connection is ciphered, is using certificate pinning and credentials are also sent ciphered are usually performed in this phase. An auditor also must check whether the data being sent is sensitive or private, and verify this sending is done with user consent.

**Sensitive Data** focuses on verifying to what data the application is accessing at runtime, since these data could differ from the results obtained in the Pre-Runtime phase. For instance, an application can create new code during execution, and make use of a buffer overflow to eventually access to this new code and access to some sensitive data [26]. An auditor can use specific tools (e.g., TaintDroid [17] or DiOS [16]) to monitor system calls related to data handling.

**Network Behaviour** is related to the analysis of Internet connections, i.e., the final destinations of any connection established by the application. These connections might be legitimate when it is related to the service provided by the application, or illegitimate when referring to exfiltrating user behaviour or other sensitive user data.

**Running Code** covers the entire process of dynamically analysing an application. To check whether an application is performing the role it should do is a part of these analysis phase. Messages, activities, services, events, etc. must be monitored by an auditor, and keep tracking of methods executed, parameters, and return values.

**C. Post-Runtime Phase**

The last phase of the methodology deals with forensics data, as it is performed after the application has been executed. Namely, we propose to analyse keychain, cookies, property lists, databases, logs and files downloaded during the execution of the application. This phase includes only the **Data analysis block**.

Keychain is the internal credential storage managed by the OS, usually very secure unless the device has a non-native OS firmware (i.e., it is rooted or jailbroken). An application with high privileges can read system memory, and thus read the keychain. Therefore, it is recommended to store credentials in ciphertext instead of plaintext.

Cookies are used by the OS to connect to services without further login intentions. For instance, in Apple’s iOS, the cookies are stored in the backup files, therefore an attacker can access to different services using a stolen authenticated session cookie.

Property List and Configuration Files correspond to generated files which store any kind of sensitive data, private information or even credentials. These files are sometimes stored in plain text in insecure locations and in the backup files. Thus, they can be usually accessed without privileges, and thus, stolen.

Databases, normally SQLite, refer to files generated during execution and used to store internal application data. These data can be very diverse, such as private data, sensitive data, or credentials. An auditor must review all these files to verify what they contain, how it is stored and handled.

Logs describes the files that contain information of the execution, usually used for debugging purposes. These files are useful to analyse the reasons of application failures, or the status of its execution.
Finally, Files Downloaded relate to files that are downloaded via any connection (e.g., JavaScript files, binaries, configuration files, etc.). These files must be studied to verify its type and what for they are being used.

V. CASE STUDY: eCOMMERCE APPLICATION

In order to evaluate the effectiveness of the methodology proposed, we audit the same application on Google’s Android and Apple’s iOS platforms. We have chosen an application (for anonymity called eCommerce app) from a Spanish online outlet, which is freely available at Google Play and Apple Store. eCommerce app is a multi-platform application used by more than 15 million of world-wide users, containing several functionalities where the methodology applies: Authorisation, authentication, user session handling, sensitive data handling, user private data storage, backend data handling, purchases, etc.

For the sake of space, we only reveal the most significant vulnerabilities found during the auditing process. A full report can be found at [27].

A. On Pre-Runtime Phase

The first step is to gather information about eCommerce app. Passive information matches for both platforms: The application author is a known developer named Anonimised; servers and user databases used by eCommerce app coincide those used by the eCommerce website utilizes; this also happens to other services such as databases of users and others. Active information shows a CVE-2004-0230 [28] vulnerability on the server es.ecommerce.com: A Denial-of-Service attack on this server would totally leave the application out of service.

A static analysis on eCommerce app reveals the application permissions. In the case of Google’s Android version, the declared permissions are related to the use of the Internet and the Google Cloud Messaging [29], to vibrate, to keep the device on, to access to user’s Google account registered on the device, and to read phone contacts and phone state. In our opinion, the last permissions could be considered non relevant for this kind of application. Permissions declared in Apple’s iOS version clearly differ: Apart from Internet access, it has access to the MAC address, location information, pasteboard and the iPhone’s Unique Identifier (UDID).

B. On Runtime Phase

In this phase we deeply analyse the network behaviour and data sent by eCommerce app in both platforms. Both applications are communicating with the same servers where vulnerabilities aforementioned in login and shopping requests.

Login requests are performed using an unencrypted HTTP connection and sending plain-text username and password encrypted with MD5 out. Note that MD5 has been repeatedly demonstrated to be vulnerable [30], [31]. Besides, a large number MD5 rainbow tables can be found on the Internet, so an attacker has a high chance of getting back the password in plain text. Every time a user opens the application and logs in, even when the login credentials are remembered, the login request is sent. Figure 2(a) depicts the HTTP POST login request generated by eCommerce app in Google’s Android platform.

eCommerce app uses mobile retargeting, i.e., it collects information on the mobile user to later send to the user personalised advertising. This feature is offered by several third-party companies. In this case, it uses a TapCommerce service for personalised advertising. This feature is offered by several third-party companies. In this case, it uses a TapCommerce service for personalised advertising. This feature is offered by several third-party companies. In this case, it uses a TapCommerce service for personalised advertising.

Purchase requests are done using encrypted HTTPS connections but without certificate pinning. This means that the application is vulnerable to self-signed certificates, which allow to decrypt data being sent. An Man-in-the-Middle (MitM) attack using a transparent proxy such as Mallory could be used towards this end.

Similarly, payment requests use an encrypted HTTPS connections but without certificate pinning. The data being sent out are related with credit card used for payment: Card number, card issuer, card holder, card expiration, and CVV code. Note that an attacker can retrieve these sensitive data using a similar scenario to the purchase request attack. Figure 2(b) depicts the HTTPS POST payment request generated in Apple’s iOS platform.

File handling analysis shows that files are written to the application sandbox, which prevent access by third-party applications.

During the execution of the application, the most-used permissions in Apple’s iOS platform are the access to the MAC address, pasteboard and the iPhone’s Unique Identifier (UDID).
C. On Post-Runtime Phase

In the case of eCommerce app, the main issue is to verify how the data are stored since it is a benign application. Credentials data are stored outside the OS keychain in both platforms. However, files written into the application sandbox are plain text and contain these credentials. In the case of the Google’s Android version, the user email is stored in plain text while the user’s password is stored in MD5. In the case of Apple’s iOS version of eCommerce app stores the user email and login password in plain text and this file is saved when backuping. Thus, this information can be read by any attacker with access to the backup files. This problem not only endangers the confidentiality of user credentials but also other services used by the user, since people often use the same credentials for different services (e.g., email, social networks, cloud storage, etc.).

Lastly, it is worth mentioning that a cookie of eCommerce website is stored on Google’s Android platform containing session information.

Findings Summary

In brief, we found several vulnerabilities in each phase of the proposed methodology: During the Pre-Runtime phase, we found a vulnerability on the remote servers used by the application that can lead to unavailability; On Runtime phase, we found that user credentials are transmitted in plain text but password, which is weak ciphered using MD5. Moreover, purchase requests containing credit card user data are transmitted using an encrypted connection but likely to be spoofed by a MitM attack; On Post-Runtime we found files generated by the application are totally insecure stored, being easily accessed by third-party applications, by malware, or by physically accessing to the device or a device backup file. We have notified all vulnerabilities the spanish CERT to be fixed. This methodology allows finding vulnerabilities, and developing methodologies to guarantee the security in mobile applications. Therefore, a security auditing methodology for developing device-independent mobile applications, “ is doing what is expected to do, and no hidden behaviours are detected. In this paper, we consider the OWASP Top Ten Mobile Risks, and based on them, we define five main analysis blocks that cover these risks (Environment Analysis, Connections, Application Own Data, Sensitive data, and Structure of the application). Using these blocks, we propose a methodology to guide an auditor to perform the audit of a mobile application. By applying it, the auditor can certify that the mobile application covers the top risks identified by OWASP as critical.

This methodology allows finding vulnerabilities, and detecting suspicious behaviours in a mobile application. We validate our methodology analysing a widely used spanish online outlet application available for Google’s Android and Apple’s iOS platforms. In both platforms several vulnerabilities have surfaced. These vulnerabilities may compromise the confidentiality, integrity of the user data and the availability of the service provided by the application.


