4.2.2 漏洞分析

本文对目前市场上占有率较高的几个 Android 版本的漏洞进行统计,结果见表 4.其中,漏洞 1~漏洞 7 分别代表本文在第 2.4 节列举出的 7 种 WebView 漏洞.从表中可以看出,Android 4.1 系统的漏洞最多,共发现 5 个已知漏洞.最新系统 Android 7.0 的漏洞最少,目前只有利用侧信道漏洞才能发起攻击.这说明,随着 Android 系统的不断升级,很多已知漏洞在新版系统中被成功地修复,也表明 Google 公司对 WebView 相关漏洞的高度重视.表 4 的最后一列也给出了每个版本的市场占有率情况[24],排名第一的是 Android Kitkat(6.0).该系统在 2017 年 5 月的占有率为 25.59%,共发现两个已知漏洞.对于最新系统 Android 7.0,在 2017 年 5 月的占有率仅为 6.33%.另外,有 3 个主流 Android 版本(4.4,5.0 和 5.1)的漏洞超过了 3 个,市场总占有率超过 55%.这主要是因为 Android 系统的碎片化问题造成的,最新系统很难及时发布到所有的 Android 手机上.这样,虽然新版的 Android 系统修复了一些已知漏洞,但是由于各大厂商更新速度不同,市场上仍然有大量的 Android 手机运行着旧版系统.随着 Android 市场整体占有率的提升,攻击者仍然可以从大量的旧版系统中成功地利用已知漏洞发起攻击.

Table 4 Distribution of loophole among different Android versions

表 4 不同 Android 版本的漏洞分布情况

Android 版本	漏洞 1	漏洞 2	漏洞 3	漏洞 4	漏洞 5	漏洞 6	漏洞 7	占有率(%)
7.0	–	–	–	–	–	–	√	6.33
6.0	–	–	–	–	–	√	√	25.59
5.1	–	–	–	√	–	√	√	22.52
5.0	–	–	√	√	√	√	√	10.44
4.4	√	–	√	√	√	√	√	22.53
4.1 以前	√	√	√	√	√	√	√	12.59

4.2.3 攻击效果

为了评估自动生成载荷的攻击效果和攻击范围,本文随机选择 1 000 个包含移动广告的应用进行实验.通过 Android 提供的 appt(Android asset packaging tool)工具,对这些应用的 APK 文件和相应的广告库 AdSDK 文件进行静态分析.其中,315 个应用嵌入的广告库中不包含可以执行 JavaScript 代码的 iframe,104 个应用通过 HTTPS 实现加密传输,剩余的 581 个应用通过 HTTP 协议向远端服务器动态获取广告内容.在这 581 个移动应用中,有一些应用处于休眠状态,没有应用开发者负责更新和维护,需要剔除出去.因此,本文进一步利用 MoneyRunner 动态执行工具对这些应用进行分析,确定哪些应用仍然可以发送和接收广告数据包.最终,本文共对 252 个应用进行实验.另外,虽然可以通过 SSLStrip 工具对 HTTPS 会话进行劫持,但是为了便于读者理解本文的攻击方案,暂不考虑通信加密的情况,而重点关注 HTTP 明文传输协议.而且,超过 58%的广告库传输使用的是 HTTP 协议,所以本文方案的攻击范围仍然很大.

图 8 表示宿主权限和广告库权限的分布情况.从图中可以看出:165(65%)个应用会申请超过 5 个权限,内部嵌入的 72(28%)个广告库 AdSDK 会申请超过 3 个权限.在宿主应用申请的权限列表中,有一些是自身需要,而有一些是为了广告库的需要.通过分析发现,广告库要求的权限一般会占到宿主应用所有权限的 20%.在这些广告库 AdSDK 申请的权限中,有很多都是与用户隐私相关的,例如 ACCESS_COARSE_LOCATION,READ_EXTERNAL_STORAGE 和 BROADCAST_SMS.为了广告的正常展示,很多宿主应用都没有遵循最小权限原则,会直接在自身的配置文件 AndroidManifest.xml 中加入广告库要求的权限.

通过人工对上述提取出的 252 个移动应用发起攻击,图 9 给出了 5 种攻击技术的实验结果.从图中可以看出:直接获取隐私的攻击效果最好,能够从 124(49%)个应用中窃取到用户的敏感数据;其次是拒绝服务和破坏系统这两种攻击方法,分别有 85(34%)和 63(25%)个应用被成功地攻击;然后是网络欺诈,能够从 31 个应用中获取到用户的个人信息;效果相对最差的是恶意扣费攻击方法,只有 5 个应用向预设的短信服务器发送了指定命令.这主要是因为该攻击方法会给用户带来直接损失,而且需要在电信运营商上注册相应业务.所以在本文的实验过程中没有将其作为主要攻击手段,只是用来验证方案的可行性.值得注意的是,在现实的攻击环境下,很多攻击者倾向于直接获取利益,所以这种攻击方法同样需要引起重视.

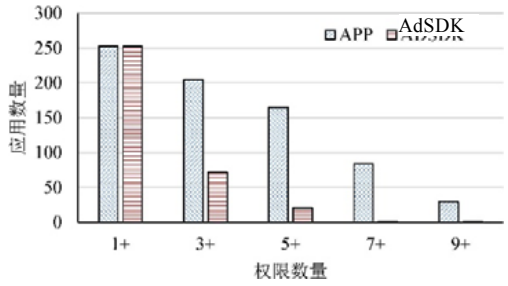


Fig.8 Distribution of permission request between app and AdSDK

图 8 宿主应用和广告库权限分布

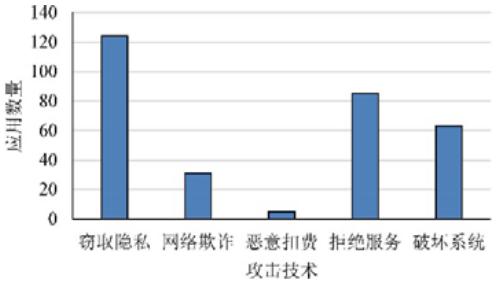


Fig.9 Attack effect

图 9 攻击效果

接下来,本文继续对攻击者能够获得的具体隐私数据进行分析,实验结果如图 10 所示.其中,获取设备信息和位置信息是攻击成功率最高的两大数据,分别能够从 103(41%)和 96(38%)个应用中获取.这与宿主应用的权限申请情况是相符的,即很多广告服务商为了提供定制化广告,需要获取这两类个性化信息,而这也恰恰被攻击者所利用.另外,对于访问联系人、获取通话记录和短信内容这样的敏感信息,本文方案也能获得不错的攻击效果,分别有 43 个、36 个和 33 个应用被攻击成功.而这主要归因于本文方案在发起攻击之前已经对宿主应用做了深入的能力分析,能够做到有的放矢,攻击也更有针对性.

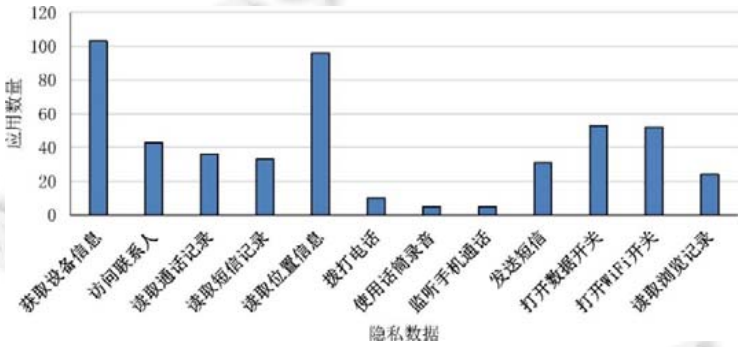


Fig.10 Attack result of different privacy data

图 10 不同隐私数据的攻击结果

4.3 攻击实例

为了更好地验证所生成攻击载荷的可行性,本节选择 4 个含有移动广告的 Android 应用进行详细介绍,包括 Angry Bird 2、盗墓笔记、nice 和 WiFi 万能密码.这 4 个应用的下载量都超过了 100 万次,各自嵌入了不同的 AdSDK,并且申请了不同的权限.本文将每个应用都分别安装到 3 个 Android 设备(Mi 5S,M8 和 S6)上,依次启动并激活攻击系统.具体的样本说明和攻击结果见表 5.

从表 5 中可以看出:在 3 个实验设备上,针对这 4 个应用的攻击均可以获得一些隐私数据.例如,Angry Bird 2 是一款免费的游戏应用,下载量超过 500 万,在其主界面包含一个广告展示框.该应用在提供娱乐功能时需要玩家的位置信息,因此在运行时申请了位置相关权限.通过利用 WebView 任意代码执行漏洞,攻击者在广告内容中植入读取位置的攻击代码,能够将设备当前位置发送到指定服务器上.盗墓笔记是一款电子书应用,下载量超过 100 万,启动时会展示一个全屏广告框.该应用将用户收藏和阅读过的文章缓存在本地 SDCard 中,在运行时需要申请存储相关权限.通过利用外部存储文件推测漏洞,攻击者在广告内容中植入访问本地存储的攻击代码,能够推测出用户感兴趣的文章类型.

在网络欺诈和恶意扣费方面,针对 nice 应用的攻击能够获得较好的攻击效果.该应用是一款社交类应用,下

载量超过 100 万,能够发布用户的最新动态,并与其他好友保持联系.在应用首页的不同位置嵌入了多个广告框,运行时需要申请读取联系人、打开摄像头和读取短信等高风险权限.通过利用 Web 层和本地层语义误差漏洞,结合数据挖掘技术,攻击者能够获得用户的身份信息,为进一步发动社会工程学攻击和网络钓鱼提供条件.另外,基于短信相关权限,攻击者通过向特定的业务号码发送短信对用户造成直接的财产损失.

Table 5 Four attack cases

表 5 4 个攻击实例

宿主应用				测试设备	攻击结果				
应用名称	分类	安装次数	广告库		窃取隐私	网络欺诈	恶意扣费	拒绝服务	破坏系统
Angry Bird 2	游戏	5 000 000+	Inmobi	Mi 5S	√	–	√	√	–
				M8	√	–	√	–	–
				S6	–	√	–	–	
盗墓笔记	电子书	1 000 000+	亿动广告	Mi 5S	√	√	–	√	√
				M8	√	–	√	–	–
				S6	√	–	–	–	
nice	社交	1 000 000+	AdColony	Mi 5S	√	√	√	–	√
				M8	–	√	√	–	
				S6	√	–	–	–	
WiFi 万能密码	工具	10 000 000+	百分通联	Mi 5S	√	–	–	√	√
				M8	√	–	–	√	√
				S6	√	–	–	–	√

在拒绝服务和破坏系统方面,针对 WiFi 万能密码应用的攻击更加有效.该应用是一款免费的系统工具应用,下载量超过 1 000 万,能够帮助用户连接和共享 WiFi 热点.除了基本的功能以外,应用开发者没有申请过多的权限,因此攻击者只能获得设备相关信息(例如设备 ID、WiFi 状态等),无法发起网络欺诈或者恶意扣费攻击.但是读取各种传感器(例如加速计、陀螺仪等)信息不需要申请任何权限,攻击者利用侧信道攻击漏洞能够妨碍设备的正常使用,并利用 WebView 域控制不严格漏洞卸载设备上已安装的其他应用.

上述分析表明,本文提出的攻击载荷自动生成方法能够根据宿主应用的权限和设备相关漏洞发起相应的攻击,攻击效果显著.另外,在攻击过程中不会出现申请过多权限的情况,从而避免了系统异常的发生,用户更加不容易察觉.

5 防范措施

本文攻击方法能够带来威胁的最主要原因是能够从网络流量中分析得到宿主应用的信息,而这些信息进一步暴露了宿主应用的能力.因此对于移动应用开发者来说,可以借鉴数据发布时隐私保护的方法来隐藏自身信息,从而阻止攻击载荷的自动生成.在广告库获取广告内容时,不要直接加入宿主应用的标识,而是发送处理过的信息.例如,利用 *k*-匿名[25]技术加入其他 *k*–1 个应用的标识信息,使得真正的宿主应用无法直接获得.利用泛化技术使用更概括、更抽象的应用数据来代替原始的信息,加大机器学习聚类分析算法识别成功的难度.值得注意的是，在对宿主应用信息进行混淆的过程中,需要考虑对最终服务质量的影响.因为宿主应用嵌入广告的目的是从广告服务商那里获得利益,所以广告服务商需要明确知道哪个应用展示了广告.如果完全去除了宿主应用的信息,那么势必会影响到开发者的利益,也会破坏整个移动广告生态系统.

对于广告服务商来说,为了防止攻击者对广告内容进行非法修改,广告库 AdSDK 可以对获取的内容进行完整性校验.在广告服务端产生广告内容之后,同时生成一段校验码,嵌入在广告内容中,并且保证该校验码无法被修改.当广告库客户端接收到广告内容之后,首先需要提取出校验码和广告内容进行比对.如果在通信过程中广告内容被攻击者修改,那么在客户端校验时会失败,这样,广告库就会知道自身受到了攻击,能够进一步采取防御措施.该方法虽然简单,但是目前大部分广告服务商都没有提供完整性校验功能.

另外,中间人攻击方法为了能够截获宿主应用和广告服务器时间传送的数据,需要依赖 ARP 欺骗或者 DNS 欺骗技术.在移动互联网环境中,ARP 欺骗很难完全防御,只要用户接收和发送 ARP 报文,就有可能会受到虚假

信息的欺骗.传统的配置静态 ARP 缓存的方法对于移动网络环境也不太有效,静态手工维护 MAC 表的方式很难实施.但是,我们仍然可以采取一些方法来降低 ARP 欺骗攻击的几率,例如使用 ARP 服务和 DHCP 服务等.无论是 ARP 欺骗还是 DNS 欺骗,都利用了协议维持信息一致性操作上的缺陷.通过指定局域网内部的一台机器作为 ARP 服务器或者在网关上建立 DHCP 服务器,能够保持网内的机器 IP/MAC 一一对应,从而防止攻击者的冒名顶替.

6 总结与展望

移动应用广告作为一种新型的商业模式,给移动互联网的发展带来了新的机遇,同时,它也给用户的隐私与安全带来了诸多挑战.通过对现有移动广告生态系统的深入分析,本文提出了一种基于宿主权限的移动广告漏洞攻击方法.该方法能够在广告主、广告平台和移动应用都是可信的前提下,通过广告网络发起中间人攻击.首先,本文对广告流量进行分析,从中提取出宿主应用的标识和设备相关信息.宿主应用的标识能够用来得到攻击者的能力上限,设备相关信息能够用来确定攻击者的攻击途径.然后,本文提出了一种基于能力描述语言(CDL)的攻击载荷自动生成方法,利用有限状态自动机生成可行的攻击代码.实验结果表明,本文提出的攻击方案能够达到很好的攻击效果,大量移动广告存在泄露宿主应用标识的问题.最后,针对本文提出的攻击方法给出了一些可行的防御方法.随着攻击者威胁的加剧和移动广告研究的不断深入,相信人们会越来越关注移动广告生态系统中的安全问题.

References:

[1] Rastogi V, Shao R, Chen Y, Pan X, Zou SH, Riley R. Are these ads safe: Detecting hidden attacks through the mobile app-Web interfaces. In: Proc. of the 23rd Network and Distributed System Security Symp. Internet Society, 2016. [doi: 10.14722/ndss.2016.23234]

[2] Demetriou S, Merrill W, Yang W, Zhang A, Gunter CA. Free for all! Assessing user data exposure to advertising libraries on Android. In: Proc. of the 23rd Network and Distributed System Security Symp. Internet Society, 2016. [doi: 10.14722/ndss.2016.23082]

[3] Meng W, Ding R, Chung SP, Han S, Lee W. Thre price of free: Privacy leakage in personalized mobile in-app ads. In: Proc. of the 23rd Network and Distributed System Security Symp. Internet Society, 2016. [doi: 10.14722/ndss.2016.23353]

[4] Book T, Wallach DS. A case of collusion: A study of the interface between ad libraries and their apps. In: Proc. of the 2013 ACM Workshop on Security and Privacy in Smartphones and Mobile Devices. Berlin: ACM Press, 2013. 79−86. [doi: 10.1145/2516760.2516762]

[5] Grace M, Zhou W, Jiang X, Sadeghi A. Unsafe exposure analysis of mobile in-app advertisements. In: Proc. of the 5th Security and Privacy in Wireless and Mobile Networks. Tucson: ACM Press, 2012. 101−112. [doi: 10.1145/2185448.2185464]

[6] Datta A, Tschantz MC, Datta A. Automated experiments on ad privacy settings. Proc. on Privacy Enhancing Technologies, 2015, 1(1):92−112. [doi: 10.1515/popets-2015-0007]

[7] Son S, Kim D, Shmatikov V. What mobile ads know about mobile users. In: Proc. of the 23rd Network and Distributed System Security Symp. Internet Society, 2016. [doi: 10.14722/ndss.2016.23407]

[8] Louw MT, Ganesh K, Venkatakrishnan V. AdJail: Practical enforcement of confidentiality and integrity policies on Web advertisements. In: Proc. of the 19th USENIX Conf. on Security Symp. Washington: USENIX Association, 2010. 371−388.

[9] Zarras A, Kapravelos A, Stringhini G, Holz T, Kruegel C, Vigna G. The dark alleys of Madison avenue: Understanding malicious advertisements. In: Proc. of the 2014 Conf. on Internet Measurement Conf. Vancouver: ACM Press, 2014. 373−380. [doi: 10.1145/2663716.2663719]

[10] Li Z, Zhang K, Xie Y, Yu F, Wang X. Knowing your enemy: Understanding and detecting malicious Web advertising. In: Proc. of the 19th ACM Conf. on Computer and Communications Security. Raleigh: ACM Press, 2012. 674−686. [doi: 10.1145/2382196.2382267]

[11] Liu B, Nath S, Govindan R, Liu J. Decaf: Detecting and characterizing ad fraud in mobile apps. In: Proc. of the 11th USENIX Symp. on Networked Systems Design and Implementation. Seattle: USENIX Association, 2014. 57−70.

[12] Crussell J, Stevens R, Chen H. Madfraud: Investigating ad fraud in android applications. In: Proc. of the 12th Annual Int'l Conf. on Mobile Systems, Applications, and Services. Bretton Woods: ACM Press, 2014. 123−134. [doi: 10.1145/2594368.2594391]

[13] Tuncay GS, Demetriou S, Gunter CA. Draco: A system for uniform and fine-grained access control for Web code on Android. In: Proc. of the 23rd ACM Conf. on Computer and Communications Security. Vienna: ACM Press, 2016. 104−115. [doi: 10.1145/2976749.2978322]

[14] Pearce P, Felt AP, Nunez G, Wagner D. AdDroid: Privilege separation for applications and advertisers in Android. In: Proc. of the 23rd ACM Conf. on ASIA Computer and Communications Security. Seoul: ACM Press, 2012. 71−72. [doi: 10.1145/2414456.2414498]

[15] Shekhar S, Dietz M, Wallach DS. AdSplit: Separating smartphone advertising from applications. In: Proc. of the 21st USENIX Conf. on Security Symp. Bellevue: USENIX Association, 2012. 553−567.

[16] Georgiev M, Jana S, Shmatikov V. Breaking and fixing origin-based access control in hybrid Web/mobile application framework. In: Proc. of the 21st Network and Distributed System Security Symp. San Diego: Internet Society, 2014. [doi: 10.14722/ndss.2014.23323]

[17] NetMate. http://f00l.de/netmate/

[18] Shahshahani B, Landgrebe D. The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon. IEEE Trans. on Geoscience and Remote Sensing, 1994,32(5):1087−1095. [doi: 10.1109/36.312897]

[19] Hartigan JA, Wong MA. Algorithm AS 136: A *K*-means clustering algorithm. Journal of the Royal Statistical Society, 1979,28(1):100−108.

[20] Viennot N, Garcia E, Nieh J. A measurement study of google play. In: Proc. of the Int'l Conf. on Measurement and Modeling of Computer Systems. Austin: ACM Press, 2014. 221−233. [doi: 10.1145/2591971.2592003]

[21] Hopcroft JE, Motwani R, Ullman JD. Introduction to Automata Theory, Language, and Computation. 2nd ed., Addison Wesley, 2003.

[22] Metasploit. http://www.metasploit.com

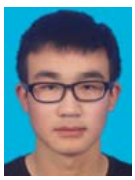
[23] Monkeyrunner for android developer. http://cs.szpt.edu.cn/android/tools/help/monkeyrunner_concepts.html

[24] Operating system market share. https://www.netmarketshare.com/operating-system-market-share.aspx?qprid=10&qpcustomd=1

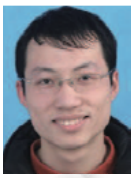
[25] Sweeney L. *k*-anonymity: A model for protecting privacy. Int'l Journal on Uncertainty, Fuzziness and Knowledge-Based Systems, 2002,10(5):557−570. [doi: 10.1142/S0218488502001648]

王持恒(1990－),男,河南漯河人,博士生,CCF 学生会员,主要研究领域为移动安全,恶意软件检测,隐私保护.

陈晶(1981－),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为网络安全,分布式系统安全.



苏涵(1994－),男,硕士生,主要研究领域为移动安全,漏洞挖掘.



何琨(1986－),男,博士,主要研究领域为网络安全,云安全.

杜瑞颖(1964－),女,博士,教授,博士生导师,主要研究领域为网络安全,密码学.