基于能量分析技术的芯片后门指令分析方法

马向亮[1,2,3],王 宏[3],李 冰[3],方进社[3],严 妍[4],白学文[5],王 安[6,7]
(1. 中国科学院软件研究所可信计算与信息保障实验室, 北京 100190;2. 中国科学院大学,北京 100049;
3. 国家信息技术安全研究中心,北京 100084;4. 中国网络安全审查技术与认证中心,北京 100020;
5. 北京航天发射技术研究所,北京 100076; 6. 北京理工大学计算机学院,北京 100081;
7. 中国科学院信息工程研究所,中国科学院网络测评技术重点实验室,北京 100093)

摘 要: 芯片后门指令是激活硬件木马的典型方式之一,其安全风险高,影响范围广,且难于检测.本文提出了一种基于能量分析的后门指令检测方法,通过对指令分段穷举、并分别采集其能量信息,可有效区分常规指令和后门指令.实验表明,通过简单能量分析即可从能量迹中直接判定出后门指令.进一步,本文提出了一种自动化识别后门指令的相关能量分析方法,通过判断其相关系数与系数均值之间的关系,可高效、自动地完成后门指令分析.

关键词: 芯片; 简单能量分析; 差分能量分析; 相关能量分析; 后门指令; 智能卡
中图分类号: TP309.1 **文献标识码:** A **文章编号:** 0372-2112 (2019)03-0686-06
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2019.03.024

A Power Analysis Method Against Backdoor Instruction in Chips

MA Xiang-liang[1,2,3], WANG Hong[3], LI Bing[3], FANG Jin-she[3], YAN Yan[4], BAI Xue-wen[5], WANG An[6,7]
(1. *Trusted Computing and Information Assurance Laboratory, Institute of Software, Chinese Academy of Sciences, Beijing* 100190, *China*;
2. *University of Chinese Academy of Sciences, Beijing* 100049, *China*;
3. *National Research Center for Information Technology Security, Beijing* 100084, *China*;
4. *China Cybersecurity Review Technology and Certification Center, Beijing* 100020, *China*;
5. *Beijing Institute of Space Launch Technology, Beijing* 100076, *China*;
6. *School of Computer Science, Beijing Institute of Technology, Beijing* 100081, *China*;
7. *Key Laboratory of Network Assessment Technology & Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100093, *China*)

Abstract: The backdoor instruction of chip is one of the typical ways to activate hardware Trojan, which has high security risk and a wide range of impact besides being difficult to be detected. In this paper, we propose a detection method of the backdoor instruction based on power analysis technology. By utilizing the segmented exhausting process and some power traces, the backdoor instruction can be distinguished from the conventional instruction effectively. The experiments show that the backdoor instruction can be analyzed successfully from the power traces by simple power analysis (SPA). Moreover, we also present an automatic detection method for the backdoor instruction based on correlation power analysis (CPA). By comparing the correlation coefficient with the mean value of the coefficient, backdoor instruction can be analyzed efficiently and automatically.

Key words: chip; SPA; differential power analysis (DPA); CPA; backdoor; smart card

1 引言

随着我国信息技术的发展,带有微处理器的智能芯片技术发展已非常迅速,应用范围也日益广泛,涵盖了从日常生活使用的公交卡、社保卡、银行金融卡到涉及国防安全的各种芯片.然而,芯片后门[1]的存在可能使非法访问者获得最高权限,甚至访问内部的所有敏感信息,控制所有资源,给最终用户的设备和应用系统

收稿日期: 2018-04-12; 修回日期: 2018-08-13; 责任编辑: 覃怀银
基金项目: 国家重点研发计划(No. 2018YFB0904901, No. 2016YFF0204005, No. 2016YFF0204003, No. 2016YFF0204002); "十三五"装备预研领域基金(No. 6140002020115); CCF-启明星辰"鸿雁"科研计划(No. 2017003); "十三五"国家密码发展基金密码理论课题(No. MMJJ20170201); 信息保障技术重点实验室开放基金(No. KJ-17-009)

带来了极大的安全风险.

本文研究的后门指令[2]是由一串二进制数字固定序列组成的电子信号,是芯片后门的激活方法中最常用的一种方式,存在于芯片的 COS 或直接存储于硬件中. 如何对后门指令进行有效搜寻探测[3~6],一直是困扰分析者的难题,本文提出一种基于能量分析技术[7~9]的有效后门指令分析方法. 该方法通过逐字节分段穷举,采集芯片解析 APDU 指令产生的能量迹,并对能量迹进行去噪、滤波、差分等信号处理[10],然后使用简单能量分析(SPA)直接分析该指令的有效性. 进一步,我们通过实验选择皮尔森相关系数[10]找到合适阈值,使用相关能量分析(CPA)自动识别后门指令.

2　基于能量分析技术的后门指令分析

2.1　分析原理

芯片内的后门指令通常由一个固定的字符串表示. 字符串可能是几个字节,几十个字节. 在黑盒的条件下,即使一个字符串仅 8 个字节,则穷尽此 8 字节字符串的次数,难度仍大于破解 DES 算法密钥难度. 但我们利用分而治之的思想逐字节穷举,可以将穷举的总次数缩短为 $2^8 \times 8$.

实现分段穷举的关键在于如何分辨出为“真”的那个字节. 当整个指令不是有效指令即芯片中不存在该指令时,可能其中的某字节指令段存在,通过芯片返回响应数据无法得知. 我们猜测芯片在收到 APDU 指令时,在处理器内部逐字节比较解析处理该指令. 比较过程中,对于“真”字节和“伪”字节的处理具有不同的解析处理流程,这两种流程差异可能通过某种侧信道方式泄露给攻击者,如时间、能耗或电磁辐射. 因此,处理器在进行解析指令逐字节比较时,无论比较失败或成功,比较能量迹都会产生,而且正确和失败之间的能量迹是有差别的,将“真”能量迹与“伪”能量迹之间进行差分,会出现明显的差异信息即尖峰. 如两个“伪”能量迹之间进行差分,则不会出现明显的差异尖峰信息.

2.2　分析算法

通过单字节穷举方式,采集芯片解析处理每条指令产生的能量迹. 通过初步判断选择一条“伪”指令产生的能量迹作为基准,其余指令逐一与该指令进行差分计算. 在差分能量迹中,有明显差分信息的能量迹对应的单字节指令有效. 通过重复以上操作,依次逐字节逼近完整的被隐藏的指令字段. 如算法 1.

算法 1　SPA 单字节穷举后门指令算法

输入:所有可能的指令集 *ABCDEF*…

输出:芯片存在的指令集 *ABCDEF*…

步骤 1:输入指令集 *ABCDEF*…,其中 *ABCDEF*…都是单字节,*A* 是 0x00,0x01,…,0xFF256 个字节中的一个字节,其余 4 字节均先设为 0xFF.

步骤 2:采集芯片解析、判断比较上述首字节指令段对应的 256 条能量迹(T_0, T_1,…,T_{255}).

步骤 3:对上述能量迹进行滤波等预处理,选取最后一个首字节为 FF 的指令产生的能量迹为基准,剩余的单字节对应的能量迹与该能量迹进行差分,得到 255 个差分能量迹集合(N_0, N_1,…,N_{254}).

步骤 4:判断差分能量迹集合中每条能量迹是否存在明显可利用的尖峰信息,若存在则说明该指令首字节有效,保留该字节,得到新的指令集 A = XX…YY,B = 0xFF,C = 0xFF…

步骤 5:遍历步骤 4 中的指令 *A* 集合,并更改 *B* 集合中的元素为 0x00,0x01,…,0xFF ,256 个字节中的一个字节.

步骤 6:重复步骤 2 -5,并逐步向后更改一个字节集合为 0x00,0x01,…,0xFF,得到新的指令集 *ABCDEF*…

步骤 7:直到更改单字节指令集后,尖峰信息没有明显变化,则可提取出该指令集 *ABCDEF*…

步骤 8:排列组合指令集 *ABCDEF*…即得到该芯片所有的指令集合.

通过单字节穷举方式,多次采集芯片解析处理每条指令产生的能量迹. 对能量迹进行均值和滤波等预处理,通过计算皮尔森相关系数,小于系数均值的为有效指令段,否则为无效指令段,若整个子图中未出现明显尖峰,说明该字节为非判断字节,可取任意值. 通过重复以上操作,依次逐字节逼近完整的被隐藏的指令字段. 如算法 2.

算法 2　CPA 单字节穷举后门指令算法

输入:所有可能的指令集 *ABCDEF*…

输出:芯片存在的指令集 *ABCDEF*…

步骤 1:输入指令集 *ABCDEF*…,其中 *ABCDEF*…都是单字节,*A* 是 0x00,0x01,…,0xFF,256 个字节中的一个字节,其余 4 字节均先设为 0xFF.

步骤 2:至少重复采集 5 次芯片解析同一首字节指令段对应的能量迹,共 256 个能量迹集(T_0,T_1,…,T_{255}).

步骤 3:对上述能量迹进行均值、滤波等预处理,得到新的能量迹 256 个能量迹集(N_0,N_1,…,N_{255}).

步骤 4:对上述 256 条能量迹计算相关系数,小于系数均值的为有效指令段,否则为无效指令段,若整个子图中未出现明显尖峰,可取任意值. 得到新的指令集 A = XX…YY,B = 0xFF,C = 0xFF…

步骤 5:遍历步骤 4 中的指令 *A* 集合,并更改 *B* 集合中的元素为 0x00,0x01,…,0xFF 256 个字节中的一个字节.

步骤 6:重复步骤 2 -5,并逐步向后更改一个字节集合为 0x00,0x01,…,0xFF,得到新的指令集 *ABCDEF*…

步骤 7: 排列组合指令集 *ABCDEF*…即得到该芯片所有的指令集合.

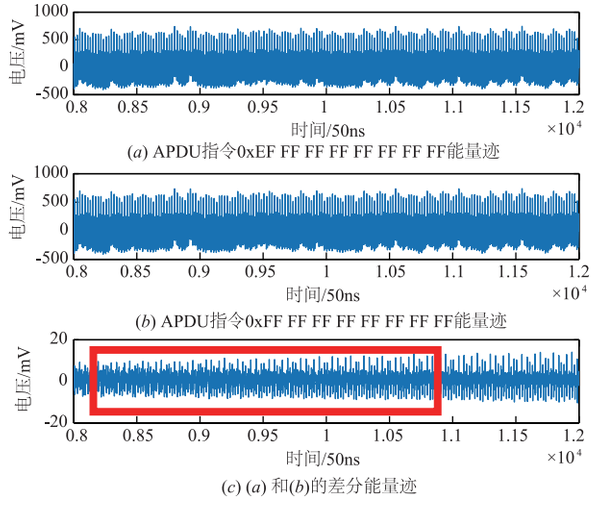
通过算法 1 和 2 可以得到被测芯片的所有可能指令集合,在此基础上再逐一判断哪些指令是公开已有的,哪些指令是未公开的,未公开的指令即后门指令. 验证其后门指令有效性,并不能通过查看返回数据或响应进行判断,而是尝试是否拿到一些特殊权限(比如可以读出密钥等)进行判断.

3 实验结果及分析

3.1 基于 SPA 后门指令分析实验

本实验由已知该智能卡指令 A0 04 00 00 08 作为隐蔽后门指令去验证,即假设该指令存在但未知. 通过本文提出的分段穷举能量分析法,采集该智能卡收到 APDU 指令后解析处理该指令的能量迹,并对预处理后的能量迹进行 DPA、SPA 分析和判断,最后恢复出该 APDU 指令.

根据算法 1 SPA 单字节穷举后门指令算法,本实验首先发送指令“FF FF FF FF FF FF FF FF”,然后将首字节以“FF”~“00”进行穷举,其余指令段保持不变,并逐一采集其对应的能量迹,本实验以首条能量迹作为基准,进行 DPA、SPA 分析和判断,得出该指令中的首字节指令段. 依次逐字节穷举重复进行实验,直到恢复出该指令的完整指令段. 本文选出具有代表性的能量迹说明该方法在恢复指令时如何进行 DPA、SPA 分析和判断. 第一组能量迹是该智能卡解析指令“EF FF FF FF FF FF FF FF”和指令“FF FF FF FF FF FF FF FF”迹预处理后的能量迹以及它们之间进行差分产生的能量迹,如图 1 所示.



(*a*) APDU指令0xEF FF FF FF FF FF FF FF能量迹

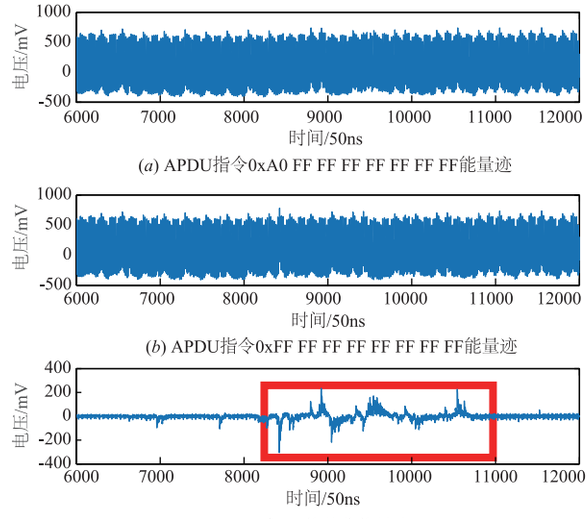
(*b*) APDU指令0xFF FF FF FF FF FF FF FF能量迹

(*c*) (*a*) 和(*b*)的差分能量迹

图1 第一组能量迹

由图 1 可得,第一组能量迹由三条能量迹组成,该能量迹横坐标轴是采样率,每 50ns 采集一个点,纵坐标是电压值,单位 mV. 从图 1(*c*)可以得出图 1(*a*)中的能量迹与图 1(*b*)中的能量迹作差分后,差分波形平稳,上下波动介于正负 20mV 之间,而原始波形的振幅介于正负 700mV 之间,可以认为两条波形近似相同,没有明显的差异信息可以利用. 猜测该智能卡在收到该指令后,在其 CPU 中进行了 if 比较判断,逐字节进行解析处理. 由于该卡的指令集中,任何指令首字节指令段没有 EF 和 FF 两个字段,即有效指令集中不存在这两条指令,CPU 在处理无效指令时会进行相同的操作,因此这两个指令比较产生的能量迹进行差分运算后,没有发现明显的可利用差异信息. 利用这一点,我们可以在实验中快速锁定无效指令的能量波形特征,而在区别出无效指令的波形之后,有效指令的能量波形也就可以很简单的区分出来了.

第二组能量迹是该智能卡解析指令“A0 FF FF FF FF FF FF FF”和指令“FF FF FF FF FF FF FF FF”迹预处理后的能量迹以及它们之间进行差分产生的能量迹,如图 2 所示.



(*a*) APDU指令0xA0 FF FF FF FF FF FF FF能量迹

(*b*) APDU指令0xFF FF FF FF FF FF FF FF能量迹

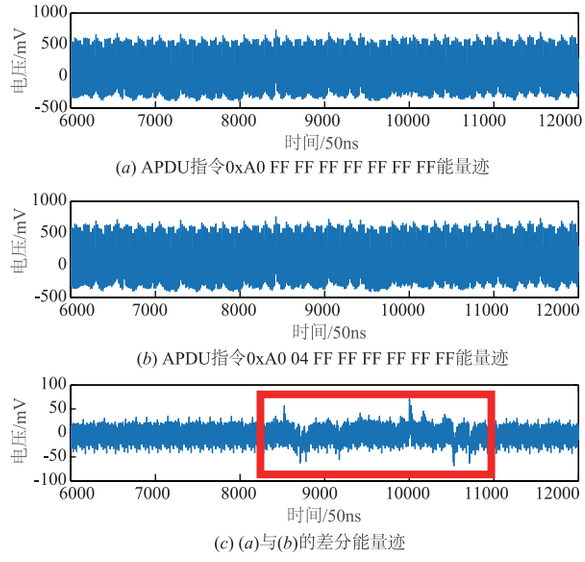
(*c*) (*a*)与(*b*)的差分能量迹

图2 第二组能量迹

由图 2(*c*)中的可以得出这两个指令产生的图2(*a*)中的能量迹与图 2(*b*)中的能量迹差分后存在明显的尖峰差异信息,如图 2(*c*)矩形框内所示. 这些尖峰差异信息是可以被利用的. 由图 1 结果猜测,可以进行推测这些差异信息是由于在该卡的 APDU 指令集中这两个 APDU 指令执行的是不同的 CPU 操作而导致的. 按此逻辑,由于 CPU 在处理无效指令的操作上是一样的,因此下发至卡中的指令中的字节指令段每多一个正确字节段应该会产生更多的差异信息. 另外,结合图 1 中的能量迹,甚至可以猜测该尖峰信息的具体来源是智能卡在解析 APDU 指令进行了多条 *if* 语句判断,或者是在执行其他的 CPU 级别的指令操作,但在本文研究的后门指令是否存在分析中不需要该级别的猜测. 在本实验中,我们的目的可以归约为只要将有效 APDU 指令和无效 APDU 指令区分开,即可算是成功的找到后门 APDU 指令.

第三组能量迹是该智能卡解析指令“A0 04 FF FF FF FF FF FF”和指令“FF FF FF FF FF FF FF FF”迹预

处理后的能量迹以及它们之间进行差分产生的能量迹,如图 3 所示.



(*a*) APDU指令0xA0 FF FF FF FF FF FF FF能量迹

(*b*) APDU指令0xA0 04 FF FF FF FF FF FF能量迹

(*c*) (*a*)与(*b*)的差分能量迹

图3　第三组能量迹

由图 3(*c*)中的差分能量曲线与图 2(*c*)中的曲线比较后可以得出,在矩形框内尖峰信息中,这两个指令产生的能量迹存在明显的可利用差异信息,而且这些差异信息明显多于图 2(*c*)中的差分曲线,验证了上面的猜测.

通过进行实验,该智能卡在收到 APDU 指令后,解析处理正确的(该智能卡已有的)指令段和错误的(该智能卡未有的)指令段时产生的能量信息差异性交大,并且这些差异信息可以通过它们的能量迹进行差分获得. 因此差分能量迹作为判断该字节指令段是否存在的依据是可行的. 重复以上实验,直到差分能量曲线没有出现明显的可利用差分信息,即最后一个字节指令段产生的差分能量迹信息没有变化或变化很微小,至此可得出该智能卡中存在的指令集. 对比已公开的指令集,如存在一些未公开的指令集,那么可以得出该芯片指令存在隐蔽后门指令.

3.2　基于 CPA 后门指令分析实验

本实验由已知该智能卡指令 00 b2 00 00 02 和 80 74 xx xx xx 等作为隐蔽后门指令去验证,其中 xx 为任意数值. 通过本文提出的分段穷举能量分析法,多次采集芯片解析处理每条指令产生的能量迹. 对能量迹进行均值和滤波等预处理,通过计算皮尔森相关系数,进行 CPA 恢复出该 APDU 指令.

根据算法 2 CPA 单字节穷举后门指令算法,本实验首先发送指令“FF FF FF FF FF FF FF FF”,然后将首字节以“FF”~“00”进行穷举,其余指令段保持不变,并逐一重复采集其对应的能量迹 5 次. 对能量迹进行均值和滤波处理,计算相关系数,进行 CPA 自动识别恢复全部所有指令,本文选出具有代表性的实验数据图进行说明.

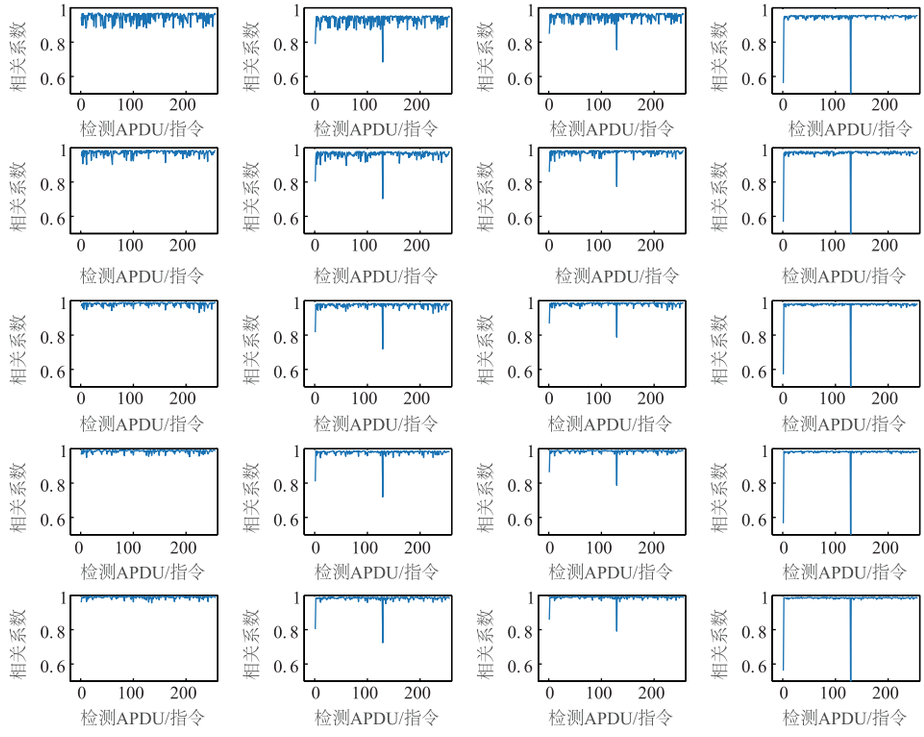


图4　能量迹均值和滤波

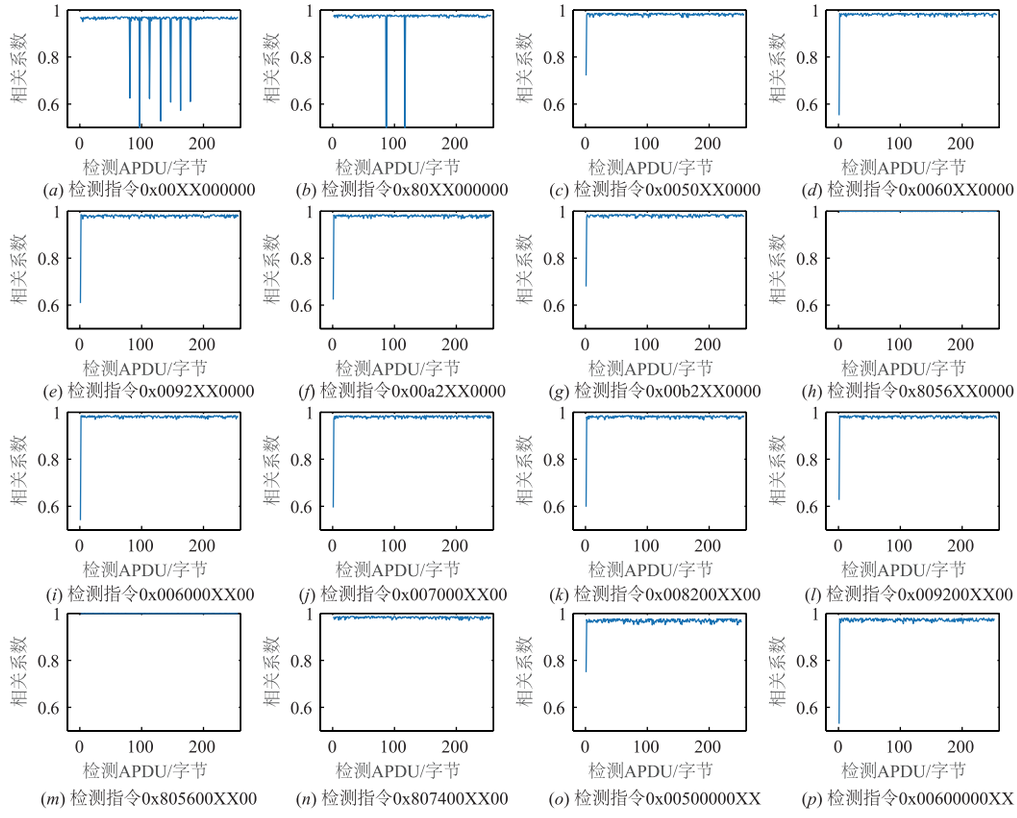


图5 CPA后门指令分析结果

如图4检测系数分析图所示,第一列至第四列分别为5点均值滤波,10点均值滤波,15点均值滤波,20点均值滤波,第一行至第五行分别为单指令采集一次波形,单指令采集二次波形,单指令采集三次波形,单指令采集四次波形,单指令采集五次波形. 由图4可以得出,随着采集波形数的增加,会提高相关系数的区分度. 而在滤波方面,与未滤波效果相比,滤波效果明显增强. 但随着滤波系数的增加,效果不一定增强,在系数为20时,滤波的效果达到峰值,而在之后,滤波的效果逐渐下降,直至无法区分. 在非可控平台或未知平台上,利用该方法可以迅速的得到检测所需的各项系数,避免由于采集能量波形时设置的系数的不合理造成额外的采集资源开销,甚至由于滤波系数设置过大或者过小而无法成功检测出有效APDU指令的问题.

由图5实验结果可知,在芯片使用逐字节比较判断机制来解析APDU指令的系统当中,使用本文所述的CPA方法,通过计算皮尔森相关系数,小于系数均值的为有效指令段,否则为无效指令段,若整个子图中未出现明显尖峰,说明该字节为非判断字节,可取任意值. 通过重复以上操作,有效区分出了芯片当前字节为有效还是无效. 根据树形结构遍历方法,逐字节向下自动穷举搜索,直到找到所有的有效指令集. 同样对比已公开的指令集,如存在一些未公开的指令集,那么可以得出该芯片指令存在隐蔽后门指令.

4 结论

由于技术敏感性,如何有效探测搜寻后门指令,目前尚无公开的具体技术和实验方案. 本文首次提出分段穷举的能量分析技术进行后门指令分析,通过猜测芯片已有的指令段和未有的指令段产生的能量信息差异性,通过SPA直接从差分曲线分析后门指令. 或将采集到同指令段的多条能量迹做均值和滤波处理后,计算相关性系数,通过CPA自动识别分析后门指令. 另外,本文提出的方法可以从智能卡产品扩展到其它类型的安全芯片产品,如FPGA.

总之,芯片产品后门的存在严重威胁了用户的关键信息和安全资产. 而且,运行在各关键系统中的电子密码芯片,如果存在后门,将是我国我国基础设施和国防体系的巨大威胁. 本文提出的基于能量技术的后门指令分析方法代价低、可行性强,可以作为国家网络安全审查的一个重要手段.

参考文献

[1] Yang K, Hicks M, Dong Q, et al. A2: Analog malicious

hardware[A]. IEEE/S&P Security and Privacy[C]. California,USA: IEEE,2016. 18 - 37.
[2] Sergei Skorobogatov,Christopher Woods. Breakthrough silicon scanning discovers backdoor in military chip[A]. Workshop on Cryptographic Hardware and Embedded Systems (CHES)[C]. Germany: Springer,2012. 23 - 40.
[3] 忽朝俭,薛一波,赵粮,等. 无文件系统嵌入式固件后门检测[J]. 通信学报,2013,34(8):140 - 145.
Hu Chao-jian,Xue Yi-bo,Zhao Liang,et al. Backdoor detection in embedded system firmware without file system [J]. Journal on Communications,2013,34(8):140 - 145. (in Chinese).
[4] R Torrance,D James. The state-of-the-art in IC reverse engineering[A]. Workshop on Cryptographic Hardware and Embedded Systems (CHES)[C]. Germany: Springer, 2009. 363 - 381.
[5] S Jha,S K Jha. Randomization based probabilistic approach to detect trojan circuits[A]. IEEE High Assurance System Engineering Symp[C]. California,USA: IEEE,2008. 117 - 124.
[6] M Banga,M Hslao. A region based approach for the identification of hardware trojans[A]. Workshop on Hardware-Oriented Security and Trust (HOST)[C]. California, USA: IEEE,2008. 40 - 47.
[7] Paul C Kocher,Joshua Jaffe,Benjamin Jun. Differential power analysis[A]. Annual International Cryptology Conference[C]. Germany:Springer,1999. 388 - 397.
[8] Clavier C,Reynaud L. Improved blind side-channel analysis by exploitation of joint distributions of leakages[A]. Workshop on Cryptographic Hardware and Embedded Systems[C]. Germany: Springer,2017. 24 - 44.
[9] 杜之波,吴震,王敏,等. 基于 SM3 的动态令牌的能量分析攻击方法[J]. 通信学报,2017,38(3):65 - 72.
Du Zhi-bo,Wu Zhen,Wang Min,et al. Power analysis attack of dynamic password token based on SM3[J]. Journal on Communications,2017,38(3):65 - 72. (in Chinese)
[10] Stefan Mangard,Elisabe Thoswald,Thomas Popp. 能量分析攻击[M]. 冯登国,周永彬,刘继业,等,译. 北京:科学出版社,2010. 100 - 111.

作者简介

马向亮(通信作者)　男,1986 年 3 月生于山西省临汾市,博士研究生,主要研究方向为信息安全、密码工程与旁路攻防技术.
E-mail: maxiangliang@163.com

王　宏　男,1972 年 9 月生于江西省玉山县,博士,高级工程师,主要研究方向为密码学、网络安全、信息安全测评.
E-mail: wh@nitsc.cn

李　冰　男,1962 年 8 月生于河北省张家口市,高级工程师,主要研究方向为网络安全、密码应用.
E-mail:lb0682@126.com

方进社　男,1956 年生于北京市,高级工程师,主要研究方向为智能卡安全、电子设备安全性检测.
E-mail:fangjinshe@163.com

严　妍　女,1979 年生于辽宁省沈阳市,硕士,主要研究方向为信息安全.
E-mail: yany@isccc.gov.cn

白学文　男,1988 年生于河南省新乡市,硕士,主要研究方向为检测技术与自动化装置.
E-mail: 412106777@qq.com

王　安　男,1983 年生于山东省莱州市,博士,硕士生导师,主要研究方向为密码工程与侧信道攻防技术.
E-mail:wanganl@bit.edu.cn