基于行为异常的 Symbian 蠕虫病毒检测方法

A Methodology of Detecting the Symbian Worm Virus Based on Abnormal Behavior

陈雅娴　袁津生　郭敏哲　(北京林业大学 计算机应用技术系 北京 100083)

摘　要: 随着智能手机持有者大幅度增多,以手机终端传播,短信、彩信形式为载体的蠕虫病毒对用户已造成多方面的危害。基于行为异常的 Symbian 蠕虫病毒检测方法以 Symbian OS60 操作系统为平台,通过频繁行为序列算法分析特定的系统函数调用行为,并结合对键盘按键操作行为的分析,检测出智能手机蠕虫病毒。通过实验,该方法可以有效的检测到当前流行的手机蠕虫病毒。

关键词: 智能手机蠕虫病毒 行为异常 SymbianOS60 频繁行为序列算法 按键操作行为

1　引言

手机蠕虫病毒是通过无线网络自我传播、利用手机操作系统、应用程序的功能和漏洞主动攻击,具有自身复制性和快速传播性,使其可以在很短时间内通过蓝牙或彩信等手段蔓延整个网络,造成了用户手机费用的损失和手机系统资源的消耗[1,2]。

当前提出的手机反病毒技术大多是通过提取手机蠕虫病毒的特征码建成特征码病毒库,这些技术虽能成功的检测已知攻击,但攻击者稍稍改变病毒特征,便可绕过特征码的匹配检测,对未知的手机蠕虫病毒检测效果不太明显[3]。

本文提出了一种关注于应用程序和用户操作的异常行为与正常行为之间差异的检测方法,即利用频繁行为序列匹配算法对应用程序对接口函数的调用行为分析,再进一步监控键盘按键行为,通过两者的结合,检测出手机蠕虫未知或变种攻击。

2　手机蠕虫病毒异常行为分析

手机蠕虫病毒主要的传播方式分为四类:①通信方式传播,主要通过彩信(MMS)、蓝牙等;②手机 WAP 上网方式,主要是用户下载的安装程序;③与其他通信设备相连;④其存储设备等。但目前所发现的蠕虫病毒大多是通过 MMS 方式传播[2]。本文主要以通过 MMS 和蓝牙传播感染的蠕虫病毒为研究对象。手机蠕虫主要以 MMS 方式伪装成安装程序传送到手机。传送的 MMS 中包含附件,附件形式本质是病毒安装程序 sis 文件,目前附件约有 60% 的文件名以. sis 为扩展名,其他以. jpg、. rm、. mp3三种形式为扩展名。

手机蠕虫病毒的发作方式分为两大类:一类是在安装行为发生后,病毒程序潜伏在系统内,待用户重启手机后自动加载到系统中运行。另一类是在安装行为发生后,病毒程序马上运行[4,5]。运行特征分为:

(1)病毒程序自动查找被感染手机中的电话本,在一定时间内或短时间内不断复制自身含有 sis 安装程序的文件转发给其他联系人。或通过寻找附近的蓝牙设备传播自身。

(2)病毒程序自动找到最近接收到的 MMS,并自动回复自身或删除接收到的 MMS。

(3)有些蠕虫会使手机键盘失灵。

为便于对手机应用程序监控,我们需要定义行为异常的概念:当应用程序访问的资源是已知的、固定的,那么其行为是可以预测的,根据程序自身行为的目的性,设定访问的资源范围,若一个未知的程序超越访问的资源范围,那么它的行为便是异常的。

经过对多种手机蠕虫的观察与分析,我们总结出如下异常行为:

(1)消息传送的一般数据项操作异常:手机中的各

种消息都是以数据项形式供程序操作的,一般数据项的操作主要是索引、定位、查找、更改和打开。正常的消息传送行为对数据项的操作只根据用户行为进行操作,不会在短时间内频繁对数据项操作。而蠕虫的异常行为有创建带自身病毒的 MMS、调用打开或删除 MMS 函数接口的行为频繁发生。

(2)频繁访问敏感数据文件:根据用户操作手机的频率,正常程序对这些操作不会是持续进行的(持续一小时内有规律操作),或短时间内(2 – 3 秒)频繁进行。而蠕虫会不断有调用访问电话薄数据库函数查找、复制联系人等函数接口的行为。

(3)蓝牙通信设备异常:只要用户打开蓝牙设备,正常情况下,蓝牙设备是不会自动搜索附近的蓝牙手机,或经过用户许可后搜索频率也是有限。而蠕虫会频繁调用 RSocket 函数接口,不断扫描附近的蓝牙设备并进行通信。

(4)按键行为异常:当用户无操作键盘行为时,即无按键事件传递给应用程序,却有以上异常行为的发生。

蠕虫病毒一般是同时出现多个异常行为,所以如果一个应用程序行为满足以上条件中的两点以上,该应用程序就极可能为蠕虫病毒。

3 基于行为异常的 Symbian OS 智能手机蠕虫检测

由于 Symbian 操作系统提供了一系列核心的应用程序接口(APIs)和被所有基于 SymbianOS 的手机共享使用的技术,所以当第三方应用程序运行时,一定会调用系统的接口函数来实现特定的功能,只要分析其对 API 函数调用行为,就可大概了解该程序的功能和行为[6,7]。键盘异常行为检测是当手机键盘无操作时,应用程序却具有蠕虫病毒的行为特征,通过监控键盘行为可知应用程序是否为蠕虫病毒。

检测方法的主要思想是将特定 API 函数接口调用行为监控和键盘操作行为的监控相结合,通过异常分析,检测手机蠕虫病毒,然后对检测出的蠕虫病毒进行报警,阻断等处理。检测方法流程如图 1 所示:

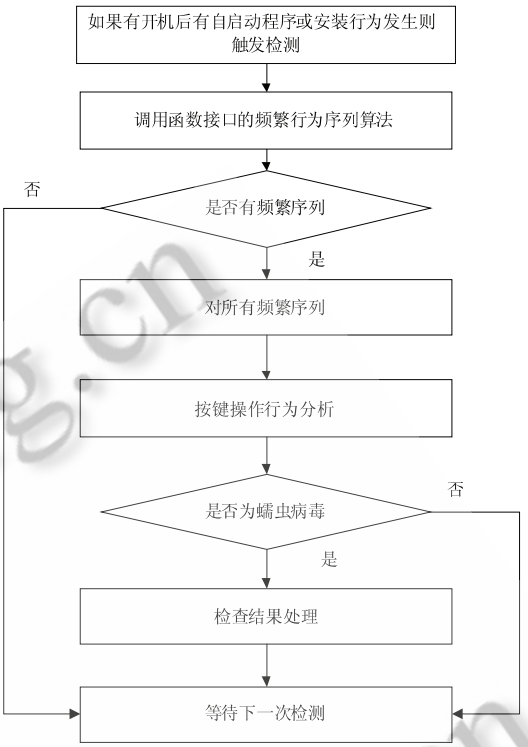

图 1　检测流程图

3.1 监听特定函数接口算法设计

3.1.1 本研究开启监听功能触发方法

为了获取异常行为,需要进行 API 函数接口监控,但是盲目、持续的监控又会浪费系统资源,因此需要寻找一个合适的监控触发时机。根据手机蠕虫病毒异常行为特征,本研究在下列两种行为发生之一开启监控功能,一是开机后有自启动程序,二是有应用程序安装行为的发生。

1. 在 Symbian S60 系统中,应用程序实现开机自启动主要有两种方法:

(1)利用 Recognizer 提供的功能,编译完成后形成 MDL 文件。蠕虫病毒程序都会在开机后自启动,其程序所用的 MDL 文件,一定会调用下列 API 函数:

CApaCommandLine::SetLibraryNameL()由 APPARC.dll 导入

CApaCommandLine::SetCommandL()由 APPARC.dll 导入

CApaDataType::DoRecognizeL()由 APMIME.dll 导入

(2)利用 Symbian Start – On – Boot API,虽然此机

制一般也通过 Recognizer 来实现，但也有特定的接口函数，病毒程序也会调用：

RServiceStartOnBoot::CreateL()

RServiceStartOnBoot::AddServiceL()

由于 Symbian 的函数导入、导出根据序号，而不是函数名，若要判断应用程序在开机自启动时调用了哪些 dll 中的函数，可根据下列方法：

①如果存在 Symbian 可执行程序(E32 文件)的导入函数，读取导入节表的地址；

②通过定位导入节，读取导入节头部信息，取得第一个导入块的地址；

③由导入节数目 iDllRefTableCount 循环读取每一个导入 的 DLL 和导入函数，其中，取得导入的库文件名字和该库导入的函数。

2. 应用程序安装行为是否发生可以通过 C 盘下 Applications. dat 文件的变化可知。

以上两种行为发生任一种，便运行调用函数接口的频繁行为序列算法。

3.1.2 函数接口调用的频繁行为序列算法

此算法目的是从当前可执行程序正在调用的函数接口中找出调用最频繁的函数接口集合。为了更好的描述算法，我们给出如下定义：

定义 1　设 Symbian 操作系统上的任意一个 API 函数接口为 p，所有 API 函数接口为全集 I，而在一个时间段内(例如 3 分钟)收集到的所有应用程序调用 API 函数接口 p 的集合称为 Q，$Q=\{p_1,p_2\ldots p_n\}$。

定义 2　每个异常行为的功能会调用某些 API 函数，设每一个异常行为所调用的函数集合为 S，$S=\{p_1,p_2\ldots p_k\}$，而所有异常行为的全集是 T，则 $T=\{S_1,S_2\ldots S_r\}$。

定义 3　对于某一个序列 S_i，定义 $Sup(S_i)$ 为序列 S_i 在集合 Q 中的支持度，而 $MinSup(S_i)$ 为此支持度的最小阈值，如果 $Sup(S_i)\geqslant MinSup(S_i)$ 成立，则称 S_i 为 Q 中的频繁序列。

说明①定义 1 中的时间段长度是不固定的，若集合 Q 中元素个数达到某特定值，则停止收集。根据多次对病毒发作时函数行为的实验，此特定值分布在 50 左右，为更好满足不同使用环境下的安全设定，用户可以在 10～100 的范围内自行选择。

②定义 2 中，将某一异常行为序列集合 S_i 的元素在集合 Q 中进行匹配，若集合 Q 中出现序列 S_i 的所有元素则 $S_i\subseteq Q$。

则在集合 Q 中，序列 S_i 的支持度如(1)式所示：

$$\sup(S_i)=\frac{Npr(Q,S_i)}{N_t},i=1,2,\ldots,6 \tag{1}$$

式(1)中，$Npr(Q,S_i)$ 是序列 S_i 在集合 Q 中出现的次数，Nt 是集合 Q 的元素个数取值，通过多次实验分析，异常行为调用的函数序列 S_i 主要是下列 6 项：(1) MTM API、(2) 创建新 MMS 条目、(3) 处理条目一般项[1]、(4) 发送信息、(5) 访问电话薄数据库、(6) 拨打电话。这 6 项已经涵盖了目前病毒发作的大部分异常行为，所以我们将主要以这 6 项为研究对象。

频繁行为序列算法的输入为集合 Q、异常行为调用函数集合 S_i 和最小支持度阈值 MinSup，输出为一个频繁行为序列二元组的集合：$\{(S_i,Sup(S_i))\}$，$0\leqslant i\leqslant 6$，其中 S_i 是 T 中的第 i 个序列，$Sup(S_i)$ 是第 i 个序列的支持度；如果 i 为 0，则表示 Q 中没有频繁序列，而 $S_1\sim S_6$ 分别代表前述 6 种主要异常行为所调用的函数序列。

以一种 MMS 蠕虫病毒为例子，可更直观地阐述上述算法。此类蠕虫病毒通过 MMS 传播，并每隔三秒钟查找被感染手机中电话薄联系人，然后复制自身感染文件，并将其伪装成 MMS 附件形式发送到其它手机。

假设病毒样本只频繁发送自身感染文件，所以调用函数接口有限。由于病毒每三秒钟发作一次，无论病毒发作和未发作期间，设集合 Q 的个数取值 $N_t=90$，病毒发作 1 分钟之内，根据此病毒样本设最小支持度阈值 MinSup 为 0.3，由频繁行为序列算法可得输出结果，$Npr(Q,S_1)=0$，$Sup(S_1)=0$；$Npr(Q,S_2)=60$，$Sup(S_2)=0.6$；$Npr(Q,S_3)=3$，$Sup(S_3)=0.03$；$Npr(Q,S_4)=60$，$Sup(S_4)=0.6$；$Npr(Q,S_5)=70$，$Sup(S_5)=0.7$；$Npr(Q,S_6)=0$，$Sup(S_6)=0$，S_2、S_4、S_5 支持度均大于最小支持度阈值 MinSup，则 S_2、S_4、S_5 是频繁序列。

S_2 通过调用 IMPORT_C::CreateMessageL(TmsvIDaID)、CMmsClientMtm::CreateAttachmentL() 等 Symbian OS60 中的 接口来创建新 MMS 条目；S_4 通过调用：CMmsClientMtm::SendL()、CActiveScheduler::Start() 等接口来发送信息；S_5 则调用 CpbkContactEngine::ReadContactL() 接口等来实现访问电话薄数据的操作。对比 MMS 蠕虫发作症状的描述，可以看出算法所

得到的频繁序列所执行的功能与蠕虫发作时的异常行为相符。因此针对上述例子,使用频繁行为序列算法就可准确的表现出应用程序行为的异常。

上述算法在非操作系统程序调用特定函数接口情况下,病毒程序调用函数接口频繁序列的置信度为0.8 。下面我们将结合键盘操作行为的分析来进一步判断蠕虫病毒。

3.2 键盘按键行为分析

经过研究我们发现,在正常应用程序运行情况下,对于集合 T 中的每个序列 Si 在实现相应的功能时都会产生对应的键盘操作行为。如果每个序列为频繁序列,则该序列所包含的函数接口就会被频繁调用,如果此时键盘的操作行为与函数调用行为不一致,那么就认为键盘操作行为是异常的。如果一个应用程序的所有频繁序列所定义的键盘操作行为都是异常,那么我们就可判定该应用程序为蠕虫病毒。

当有频繁序列行为发生时,有两个步骤分析键盘按键行为:

1. 通过扫描控件栈中有无按键事件传递,判断有无按键行为。控件栈是一个结构,当有按键事件发生时,这个按键事件从控件栈的顶部依次传递到最底端,直至得到某个控件的处理。将按键事件传递到该控件是通过调用函数 offerkeyEventL( ) 来实现的,那么可以监控此函数的调用可知控件栈中有无按键事件。

2. 若控件栈中有按键事件,判断按键事件发生的次数。经过多次实验得出:左选择键B1、确认键B2、拨打电话键B3,是蠕虫病毒发作时操作有明显异常行为的三个键。

在正常行为下,序列 S_i 执行一次对应的实际按键次数分别为:左选择键 B1i,确认键 B2i,拨打电话键 B3i,i = 1,2,...,6。

当有频繁序列行为发生时,根据频繁行为序列算法输出的二元组集合(Si,Sup(Si)),可知 Npr(Q,S_i)是序列 S_i 发生的次数(i 不能取 0),则此时序列 Si 对应的按键次数分别为左选择键 Bfr1i = B1i * Npr(Q,S_i),确认键 Bfr2 i = B2i * Npr(Q,S_i),拨打电话键 Bfr3i = B3i * Npr(Q,S_i),i = 1,2,...,6。

而对于键盘按键行为,当3.1.1 中提到的两种触发行为发生时,便开始捕获序列 S_i 对应的三个按键的按键次数,将捕获次数设定一个固定值,当捕获次数达到这个固定值时停止捕获,可得序列 Si 所对应的实际按键次数分别为左选择键 Bth1i,确认键 Bth2i,拨打电话键 Bth3i,i = 1,2,...,6。

将序列 S_i 所对应的按键次数与键盘实际按键次数比较,对于 B1、B2、B3 三键,若 Bfri > Bthi, i = 1, 2, ..., 6,则可判断按键行为异常。

至此,根据上述函数接口调用的频繁行为和键盘按键行为的结合分析,可比较准确的判定目前运行的应用程序为蠕虫病毒。

4 实验结果及分析

为验证本检测方法,我们编写一个模拟手机蠕虫病毒的恶意程序 worm. sis,另下载 Commwarroir 病毒安装在手机上各自验证。并下载一款瑞星杀毒软件手机版 V2.0,同本检测方法比较。

Worm. sis 是一个模拟蠕虫程序。此病毒特征是:安装完成重启手机后自动运行并隐藏自身进程,频繁搜索电话薄号码,发送 MMS。Commwarrior. B 病毒特征是:安装后自动搜索电话薄里的电话号码,然后向这些号码发送彩信,彩信中包含有该病毒副本,发送的频率可以达到每1~2 秒一条。

首先在 Nokia N72 手机上安装并运行这两个病毒和本检测程序 wormdetect,其次再用瑞星手机杀毒软件检测,比较结果如下:

手机安装完两个病毒后,本检测程序自启动监控特定函数接口调用功能,约 10 秒后检测到 Commwarrior. B,1 分钟内截获模拟蠕虫病毒 Worm. sis。无需扫描所有文件,并将截获结果提醒用户。

由于瑞星杀毒软件手机版 V2.0 支持实时监控功能不太理想,必须手动执行扫描病毒,查杀功能才会运行。安装完两个病毒后,瑞星手机杀毒软件都没有相关病毒警报提示,通过执行病毒扫描以后,所有的病毒才被查出来。扫描到本测试手机有文件 3781 个,用时 8 分钟。如表1 所示:

表1　测试结果

检测程序 / 功　能	wormdetect	瑞星杀毒软件手机版 V2.0
实时监控	有	无
查杀速度	约1分钟	扫描文件个数而定

(下转第 31 页)

(上接第 52 页)

由表 1 的测试结果可知,本检测程序能在短时间(最长约 1 分钟,最短 10 秒钟)成功截获恶意程序,与此相比的瑞星手机杀毒软件并未能很有效、快捷的检测到病毒。

5 结束语

本论文研究并实现了一种基于异常行为的手机蠕虫病毒检测方法。实验结果表明,该方法内存占用小,检测速度快,并对大部分手机蠕虫病毒的检测具有较高的准确率和较低的误报率。

此外,该方法主要针对目前具有一定特征的手机蠕虫病毒的检测,如果未来的变种蠕虫病毒无自启动或具备拒绝服务等特征,该方法还需进一步研究和发展。

参考文献

1 连一峰,戴英侠,王航. 基于模式挖掘的用户行为异常检测. 计算机学报,2002,(3):331 – 336.

2 李志, 王延巍, 朱林. 手机病毒的现状与未来 . 电信技术,2006,(3):87 – 90 .

3 Eskin E. Anomaly Detection over Noisy Data Using Learned Probabilit. Distributions. Proceedings of the 17th International Conferenceon Machine Learning(ICML2000),2000. 237 – 243.

4 Traynor,P. Enck,W. McDaniel,P. and Porta. T. L. Mitigating Attacks on Open Functionality in SMS – Capable-cellular Networks. In ACM MobiCom 06.

5 吴建刚, 鲁士文. 针对恶意代码的行为阻断方法研究. 微电子学与计算机,2004,21(2) .

6 Harrison R 著,周良忠译. Symbian OS 手机应用开发. 北京:人民邮电出版社,2004. 9. 10 – 221.

7 Cheng J, Starsky H. Y. Wong, Hao Yang and Songwu Lu. SmartSiren: Virus Detection and Alert for Smartphones. Dept. of Computer Science, UCLA, 4732 Boelter Hall, Los Angeles, CA 90025 1IBM T. J. Watson Research,19 Skyline Drive, Hawthorne, NY 10532 2.