



手机恶意软件网络行为分析

杜雪涛,常玲,刘佳,张晨

(中国移动通信集团设计院有限公司 北京 100080)

摘要: 分析并提出了手机恶意软件网络行为分析计算方法以及网络侧的实现方案。该方法综合考虑了手机恶意软件网络行为指数以及病毒体行为特征双因素,从而提高了手机恶意软件研判的及时性以及准确性,对现网手机恶意软件实时封堵起到了重要的指导作用。

关键词: 手机恶意软件;网络行为;分析算法

doi: 10.11959/j.issn.1000-0801.2015200

Behaviour Analysis of Mobile Malicious Software in Network

Du Xuetao, Chang Ling, Liu Jia, Zhang Chen

(China Mobile Group Design Institute Co., Ltd., Beijing 100080, China)

Abstract: A behavior analysis method of mobile malicious softwares and a corresponding solution on the network-side were proposed. The method combined network behaviors and virus-inner behaviors of a mobile malicious software to improve the accuracy and the timeliness. The method was a useful guidance for communication operators to block mobile malware in real time.

Key words: mobile malicious software, network behavior, analysis algorithm

1 引言

随着智能手机的普及,手机恶意软件日益泛滥。某厂商安全报告报道,该厂商 2014 年全年累计截获新增恶意程序样本高达 326 万个,约为 2013 年的 4 倍,平均每天截获新增恶意程序样本 8 932 个。2014 全年用户感染恶意程序 3.19 亿人次,下半年恶意程序感染量是上半年的 3 倍左右。手机恶意软件对用户危害巨大,其产生的恶意定购不仅造成用户经济损失,还造成了用户的异常网络流量及短/彩信发送量。此外,手机恶意软件还造成了包括通话记

录、短信内容、位置信息等的用户个人信息泄露,使用户手机出现无故关机、电话无法接通等情况。对于运营商而言,手机恶意软件造成了网络资源被大量占用的后果,严重地影响了业务质量,对用户的网络感知影响巨大。

2 网络侧的技术解决方案

为了有效地防治手机恶意软件,可针对手机恶意软件的生命周期开展针对性的监测、封堵及查杀工作。手机恶意软件的传播一般经过潜伏和大规模发作两个阶段。在第一阶段,手机恶意软件通过用户下载软件、短/彩信推送、

收稿日期:2015-03-19;修回日期:2015-07-29

论文引用格式:杜雪涛,常玲,刘佳等. 手机恶意软件网络行为分析. 电信科学, 2015200

Du X T, Chang L, Liu J, et al. Behaviour analysis of mobile malicious software in network. Telecommunications Science, 2015200



蓝牙等方式进入网络散播。经过一段时间的潜伏,手机恶意软件进入大网内大规模传播及发作阶段。在这一阶段,手机恶意软件将产生用户不知情定购、自动发送短/彩信、自动上网等被感染行为。这两个阶段中,通过对手机恶意软件的软件行为和特征行为进行联合分析,达到封堵、查杀恶意软件的目的。

结合以上治理思路,在现网手机恶意软件监控系统中,可以采用如图1所示的3层架构:数据采集层从五大数据来源采集数据,并对手机恶意软件共性特征进行过滤;数据分析层对用户的网络行为进行扫描、分析,确认受感染的手机号码和疑似案例;系统管理层实现对侦测结果的展现和分类统计、查询、系统管理等。

其中,数据分析功能层是本文讨论的重点,该层采用用户网络行为分析法与病毒体扫描分析法相结合的模式,完成对手机恶意软件的研判。通过网络行为分析法,收集用户行为数据,与已有病毒行为特征库对比,如特征符合,则确认用户中毒。通过病毒体扫描分析法,收集用户数据流,与已有病毒体特征库对比,如符合特征,则确认用户下载了病毒文件。

2.1 数据模型

数据分析层从数据采集层获取并维护以下资源数据:全网手机恶意软件特征数据、疑似恶意软件网络行为特征数据、非恶意软件文件样本 MD5 特征数据、已入库未研判疑似样本 MD5 特征数据、待封堵 URL 地址数据、全网恶意软件下载 URL 数据、全网恶意软件主控 URL 数据、染毒用户日志数据等。

2.2 分析流程

(1)病毒体扫描分析流程

手机恶意软件监控系统对于采集的样本文件,利用全

网手机恶意软件特征库和非恶意软件文件样本 MD5 特征库进行过滤,对上报的疑似恶意软件样本及网络日志进行汇总去重,并利用已入库未研判疑似样本 MD5 特征库,对疑似恶意软件样本进行过滤。

(2)用户网络行为分析流程

手机恶意软件监控系统利用疑似恶意软件网络行为筛选、发现疑似恶意软件相关的网络日志信息,对疑似恶意软件网络日志进行汇总去重后,对需要下载疑似样本文件的疑似恶意软件网络行为对应的 URL 进行爬取下载。对疑似恶意软件样本进行自动研判,若经过计算的疑似恶意软件网络行为指数达到了筛选门限,送至人工审核平台。当该行为确定为新型手机恶意软件时,则提取特征码信息,将其增加到全网已知手机恶意软件特征库,将其对应的下载 URL 和主控端 URL 新增到全网恶意软件下载 URL 库、全网恶意软件主控 URL 库和待封堵 URL 地址库。若疑似样本确定为正常软件,则将其增加到非恶意软件文件样本 MD5 特征库。

3 手机恶意软件的网络行为分析算法

在手机恶意软件监控流程中,疑似手机恶意软件网络行为的指数计算是研判的关键。若该行为指数定义精准,会对恶意软件研判的准确性、高效性起到至关重要的作用。在对恶意软件进行网络行为参数梳理时,可以对恶意软件的网络行为数据进行以下分类。

(1)垃圾短/彩信识别参数

包含 URL 的垃圾短信 URL、短信内容、主叫号码、被叫号码、时间、疑似 URL 等信息。

(2)终端设备类型参数

包含 IMEI、终端型号、终端厂商、软件版本号、操作系

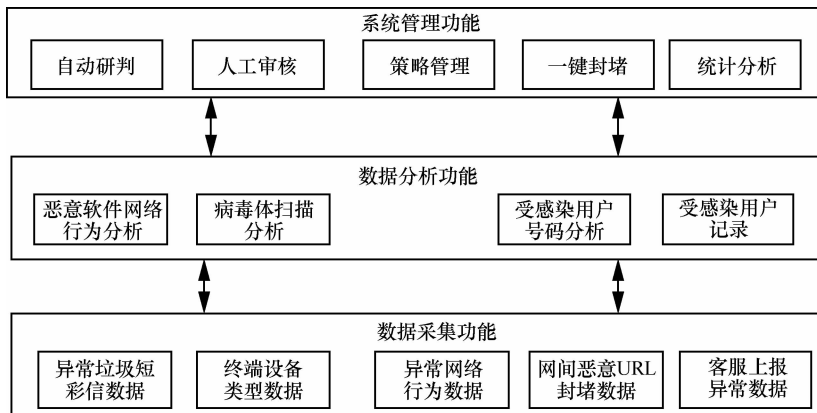


图1 恶意软件监控系统功能

统平台等信息。

(3) 异常网络行为参数

包含 URL、IMSI、IMEI、user agent (用户代理)、MSISDN、APN、时间、源 IP 地址、源端口、目的 IP 地址、目的端口、文件样本、样本的文件名、样本 MD5、样本大小、染毒用户 MSISDN、事件类型、用户连接主控或者下载恶意软件的 URL、恶意软件名称、主控 URL 地址等信息。

(4) CRM 上报异常数据

包含恶意软件类型、危害类型、恶意软件影响的操作系统及版本、恶意软件的描述、用户发生网络事件的时间等信息。

根据前面的分析可知,恶意软件的网络行为构成具有多个影响因素,网络行为指数应该由一个多维数组进行描述。利用主成份分析的方法对多维数组进行降维,去掉数组中的冗余信息,获取描述恶意软件网络行为的主要信息。具体计算方式如下。

(1) 数据量化

- 垃圾短/彩信识别参数一般包括但不限于:短信内容关键词中标次数、主叫号码人工审核次数、被叫号码人工审核次数、短信内包含的 URL 中标此数等参数。
- 终端设备类型参数一般包括但不限于:设备终端类型的安全分级或第三方安全评分。
- 异常网络行为参数一般包括但不限于:URL 涉黄指数、URL 访问热度、文件大小等。
- CRM 上报异常数据一般包括但不限于:恶意软件类型评级或第三方危害指数评分、危害类型评级或评分等。

(2) 数据采集

假设恶意软件网络行为的相关参数个数为 m , 参数分别为 x_i ($1 \leq i \leq m$), 数据采集次数为 n , 采集时间点是 t_j ($1 \leq j \leq n$), 则恶意软件网络行为指数数组 e 表示如下:

$$e = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix} \quad (1)$$

(3) 数组 e 的标准化

数组 e 中各参数描述的属性不同,其量纲也不同,因此在计算恶意软件网络行为危害指数前,需要对数组 e 进行标准化,得到标准化数组:

$$E = \begin{pmatrix} X_{11} & \cdots & X_{1n} \\ \vdots & \ddots & \vdots \\ X_{m1} & \cdots & X_{mn} \end{pmatrix} \quad (2)$$

(4) 协方差矩阵 R 的特征值 λ 和对应特征向量 L 的计算。

(5) 特征值选择

选择最大的 M 个特征值 $\lambda_{M_1}, \lambda_{M_2}, \dots, \lambda_{M_M}$, 使这 M 个特征值的方差贡献率大于 90% (主成分分析的方差贡献率一般认为大于 85% 即可, 参见参考文献[1]), 即满足:

$$\frac{\sum_{i=1}^M \lambda_{M_i}}{\sum_{i=1}^n \lambda_n} \geq 90\% \quad (3)$$

(6) 恶意软件网络行为指数的计算

设 λ_i 对应的特征向量为 $(l_{i1}, l_{i2}, \dots, l_{in})$, 其中, l_{M_j} 描述了第 M_i 个主成分中含有的第 j ($1 \leq j \leq n$) 个参数信息的权重系数, $l_{M_i,1}^2 + l_{M_i,2}^2 + \dots + l_{M_i,n}^2$ 描述了第 M_i 个主成分含有的信息量, 则恶意软件网络行为指数计算式为:

$$P = \sqrt{\sum_{i=1}^M \sum_{j=1}^n l_{M_i,j}^2} \quad (4)$$

参考文献[2,3]已经证明,仅用最大特征值即能够检测出网络中的隐藏信息,已经设定选取的主成分方差贡献值大于 90% 的特征信息,因此只要数据采集足够有效,就能够检测出较隐蔽的恶意软件信息。由以上分析可知,恶意软件网络行为指数 P 包含了多维数组 e 的 90% 的信息,只需要设定一定的门限 P_T , 则只需满足 $P \geq P_T$, 即可推送至人工审核平台。

此外,由于手机恶意软件的变种频繁,监控系统自动研判对恶意软件仍存在误判的可能性。在人工审核为正常软件后,正常软件会再次被误判而推送至人工审核,会造成人工的大量重复审核;但如果软件一次被审核为正常之后就加入正常规则库,有可能会造成漏判。因此,可考虑对曾经审核为正常的软件加入“信誉因子”,并结合历史审核记录对软件进行“平均危害指数”评价。具体方法如下。

(1) 在人工审核结束后,对于审核为正常的软件,系统将该软件的相关参数传送到危害指数计算模块,软件的相关参数包括:软件被连续审核为正常的次数 n ; 第 n 次审核为正常的审核时间 T_n ; 第 n 次审核为正常的危害指数为 P_n 。

(2) 当 $n \geq 1$ 时,即软件曾经多次被审核为正常,则网站第 n 次被审核为正常的危害指数可采用平均危害指数



进行评价计算。平均危害指数代表了该软件累加的平均危害程度,其计算方式如下。

- 软件第 n 次被监控到的危害指数为 P_n ;
- 软件曾经 $n-1$ 次审核为正常,则第 n 次软件的奖励因子为 $e^{-(n-1)}$,软件第 n 次被系统监控的危害指数为 $P_n e^{-(n-1)}$;

- 软件第 n 次被系统监控到,则其平均危害指数 $\bar{P} =$

$$\sum_{i=1}^n P_i e^{-(i-1)};$$

- 当 $\sum_{i=1}^n P_i e^{-(i-1)} > P_T$ 时,则软件再次被推送至人工审核。

(3) 参数清零条件即为约束软件重复送审的条件,其目的是减少人工对历史信誉高的软件的重复审核量。但是历史信誉应具有有一定的有效条件,在有效条件之外,信誉数据将清零。数据的清零条件如下。

- 软件的两次连续中标时间间隔过长,即 $T_n - T_{n-1}$ 超出一定的门限,则软件的信誉机制清零,即 $n=0$,系统认为软件是一个新的软件来计算其危害指数。
- 软件的再次被监测时间与首次被监测时间的间隔过长,即 $T_n - T_1$ 超出一定的门限(如一个月),则软件的信誉机制清零,即 $n=0$,系统认为软件是一个新的恶意软件来计算其危害指数。
- 软件被审核为恶意软件,则提交并且相关数据清零。

通过以上恶意软件危害指数定义及再计算,使得恶意软件的网络行为判断更为精准,减少了误判及漏判,提高了人工审核的效率。

4 结束语

本文分析并提出了手机恶意软件网络行为指数的计算方法以及网络侧的实现方式。该方式综合考虑了手机恶意软件网络行为指数以及病毒体行为特征双因素,提高了手机恶意软件研判的及时性以及准确性,对现网恶意软件实时封堵起到了重要的指导作用。

参考文献

- 1 张新荣,熊伟丽,徐保国.基于PCA的发酵过程监控模型主元数的确定.计算机测量与控制,2009,17(6):1120~1122
Zhang X R, Xiong W L, Xu B G. Determination of principal component of performance monitoring model in fermentation

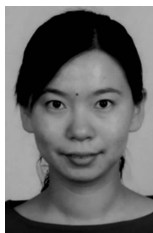
process based on PCA. Computer Measurement & Control, 2009, 17(6): 1120~1122

- 2 Plerou V, Gopikrishnan P, Rosenow B, et al. Universal and nonuniversal properties of cross correlaitons in financial time series. Physics Review Letter, 1999, 83(7):1471~1474
- 3 Yuan J, Mills K. Monitoring the macroscopic effect of DDoS flooding attacks. IEEE Transactions on Dependable and Secure Computing, 2005, 2(4): 324~335
- 4 常玲,吴兴耀,杜雪涛.手机恶意软件网络侧封堵效果验证方法研究.电信工程技术与标准化,2013,26(187):62~64
Chang L, Wu X Y, Du X T. Research of the mobile phone viruses plugging effect verification method in the network side. Telecom Engineering Technics and Standardization, 2013, 26(187): 62~64

[作者简介]



杜雪涛,女,中国移动通信集团设计院有限公司教授级高级工程师、研发技术总监,主要研究方向为移动通信网络及信息安全技术。



常玲,女,中国移动通信集团设计院有限公司高级工程师,主要研究方向为移动通信信息安全、移动智能终端安全技术研究等。



刘佳,女,博士,中国移动通信集团设计院有限公司高级工程师,主要研究方向为不良信息治理关键技术、运营商网络的客户信息保护技术。



张晨,男,中国移动通信集团设计院有限公司高级工程师、研究员,主要研究方向为信息安全、内容安全。