



(12) 发明专利申请

(10) 申请公布号 CN 103281177 A

(43) 申请公布日 2013. 09. 04

(21) 申请号 201310123028. X

(22) 申请日 2013. 04. 10

(71) 申请人 广东电网公司信息中心

地址 510600 广东省广州市越秀区水均岗路
6-8 号

(72) 发明人 王甜 陈军 邹洪 崔磊 朱奕
艾解清 吴一阳 卜夏锋

(74) 专利代理机构 深圳市君胜知识产权代理事
务所 44268

代理人 王永文 刘文求

(51) Int. Cl.

H04L 9/14 (2006. 01)

G06F 21/56 (2013. 01)

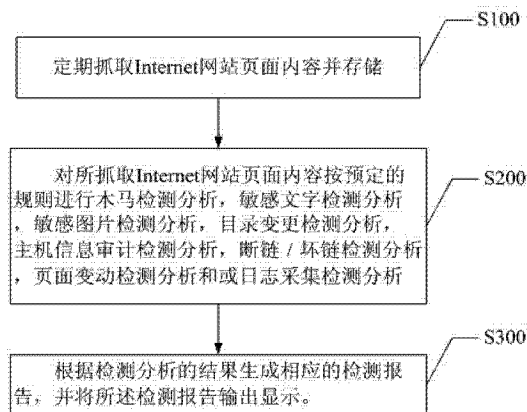
权利要求书3页 说明书10页 附图2页

(54) 发明名称

对 Internet 信息系统恶意攻击的检测方法及系统

(57) 摘要

本发明公开了对 Internet 信息系统恶意攻击的检测方法及系统, 包括步骤 :A、定期抓取 Internet 网站页面内容并存储 ;B、对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析, 敏感文字检测分析, 敏感图片检测分析, 目录变更检测分析, 主机信息审计检测分析, 断链 / 坏链检测分析, 页面变动检测分析和 / 或日志采集检测分析 ;C、根据检测分析的结果生成相应的检测报告, 并将所述检测报告输出显示。本发明的对 Internet 信息系统恶意攻击的检测方法检测全面, 提高了检测的安全, 并且减少了工作量, 降低了人力成本, 为用户提供了方便。



1. 一种对 Internet 信息系统恶意攻击的检测方法,其特征在于,包括步骤:

A、定期抓取 Internet 网站页面内容并存储;

B、对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析,敏感文字检测分析,敏感图片检测分析,目录变更检测分析,主机信息审计检测分析,断链/坏链检测分析,页面变动检测分析和/或日志采集检测分析;

C、根据检测分析的结果生成相应的检测报告,并将所述检测报告输出显示。

2. 根据权利要求1所述对 Internet 信息系统恶意攻击的检测方法,其特征在于,所述步骤A具体包括:通过 NMAP 协议和/或 SYSLOG 协议定期抓取 Internet 网站页面内容。

3. 根据权利要求1所述对 Internet 信息系统恶意攻击的检测方法,其特征在于,所述步骤B中的对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析具体包括:

B10、采用挂马链接匹配技术,通过公认的木马网站及恶意链接收集各种公认的木马信息,并将收集的木马信息建立一个恶意链接库;

B11、在对所抓取 Internet 网站页面内容的网页源码进行分析时,将所抓取 Internet 网站页面内容的网页源码与所述恶意链接库存储的木马信息进行比较,判断是否有与所述恶意链接库存储的木马信息相同的信息;

B12、当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述恶意链接库中的链接相匹配,则控制立即预警并输出相应的检测分析结果。

4. 根据权利要求3所述对 Internet 信息系统恶意攻击的检测方法,其特征在于,所述步骤B中的对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析具体还包括:

B20、采用网页源码分析技术,针对所述恶意链接库中未收集的木马链接,根据木马在浏览器中隐藏自身的通用特征,对木马链接的隐藏技术进行分析,结合不同类型的挂马形式,分别提取出各类型挂马特征,组装成挂马特征库;

B21、将所抓取 Internet 网站页面内容的网页源码与所述挂马特征库存储的木马信息进行比较,判断是否有与所述挂马特征库存储的木马信息相同的信息;

B22、当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述挂马特征库中的链接相匹配,则控制立即预警并输出相应的检测分析结果。

5. 根据权利要求3所述对 Internet 信息系统恶意攻击的检测方法,其特征在于,所述步骤B中的对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析具体还包括:

B30、采用蜜罐技术,利用虚拟技术搭建一台无任何防范的服务器,将可疑的网页链接在该服务器的浏览器中打开,监测网页打开后的系统状况;

B31、判断系统时间是否被修改、关键的系统目录下是否有文件写入、系统文件是否被篡改、注册表是否被篡改、网络流量是否异常、和/或是否连接了可疑端口或IP;

B32、当判断系统时间被修改、关键的系统目录下有文件写入、系统文件被篡改、注册表被篡改、网络流量异常、和/或连接了可疑端口或IP,则控制立即预警并输出相应的检测分析结果。

6. 根据权利要求1所述对 Internet 信息系统恶意攻击的检测方法,其特征在于,所述步骤B中的对所抓取 Internet 网站页面内容按预定的规则进行敏感文字检测分析和/或

敏感图片检测分析具体包括：

B40、预先定义指定的关键词为敏感文字和 / 或带指定内容的图片为敏感图片；

B41、对所抓取 Internet 网站页面内容进行监测过滤,判断是否包含所述敏感文字和 / 或所述敏感图片；

B42、当检测到所抓取 Internet 网站页面内容中包含所述敏感文字和 / 或所述敏感图片,则控制立即预警并输出相应的检测分析结果。

7. 根据权利要求 1 所述对 Internet 信息系统恶意攻击的检测方法,其特征在于,所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行页面变动检测分析和 / 或目录变更检测分析具体包括：

B50、将所抓取 Internet 网站页面内容,与历史数据进行匹配比较；

B51、进行页面变动检测分析和 / 或目录变更检测分析包括 js 脚本、html 脚本、a 链接、图片链接、文字内容信息,计算页面变动和 / 或目录变更百分比；

B52、当识别页面和 / 或目录被恶意篡改信息,则控制立即预警并输出相应的检测分析结果。

8. 根据权利要求 1 所述对 Internet 信息系统恶意攻击的检测方法,其特征在于,所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行主机信息审计检测分析具体包括:对主机的各种基线,进行检查和分析,将各外网应用系统按照预定的策略进行配置,对不合规内容进行预警；

所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行日志采集检测分析具体包括:通过对网络设备、主机、数据库、中间件、应用程序的日志文件进行采集和关联分析,动态识别出外网应用系统遭受攻击的情况,并检测攻击是否造成危害影响,当检测攻击是造成危害影响,则控制立即预警并输出相应的检测分析结果。

9. 一种对 Internet 信息系统恶意攻击的检测系统,其特征在于,包括：

抓取模块,用于定期抓取 Internet 网站页面内容并存储；

检测分析模块,用于对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析,敏感文字检测分析,敏感图片检测分析,目录变更检测分析,主机信息审计检测分析,断链 / 坏链检测分析,页面变动检测分析和 / 或日志采集检测分析；

生成模块,用于根据检测分析的结果生成相应的检测报告,并将所述检测报告输出显示。

10. 根据权利要求 9 所述对 Internet 信息系统恶意攻击的检测系统,其特征在于,所述检测分析模块包括：

第一检测分析单元,用于采用挂马链接匹配技术,通过公认的木马网站及恶意链接收集各种公认的木马信息,并将收集的木马信息建立一个恶意链接库;在对所抓取 Internet 网站页面内容的网页源码进行分析时,将所抓取 Internet 网站页面内容的网页源码与所述恶意链接库存储的木马信息进行比较,判断是否有与所述恶意链接库存储的木马信息相同的信息;当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述恶意链接库中的链接相匹配,则控制立即预警并输出相应的检测分析结果；

第二检测分析单元,用于采用网页源码分析技术,针对所述恶意链接库中未收集的木马链接,根据木马在浏览器中隐藏自身的通用特征,对木马链接的隐藏技术进行分析,

结合不同类型的挂马形式,分别提取出各类型挂马特征,组装成挂马特征库;将所抓取 Internet 网站页面内容的网页源码与所述挂马特征库存储的木马信息进行比较,判断是否有与所述挂马特征库存储的木马信息相同的信息;当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述挂马特征库中的链接相匹配,则控制立即预警并输出相应的检测分析结果;

第三检测分析单元,用于采用蜜罐技术,利用虚拟技术搭建一台无任何防范的服务器,将可疑的网页链接在该服务器的浏览器中打开,监测网页打开后的系统状况;判断系统时间是否被修改、关键的系统目录下是否有文件写入、系统文件是否被篡改、注册表是否被篡改、网络流量是否异常、和/或是否连接了可疑端口或 IP;当判断系统时间被修改、关键的系统目录下有文件写入、系统文件被篡改、注册表被篡改、网络流量异常、和/或连接了可疑端口或 IP,则控制立即预警并输出相应的检测分析结果;

第四检测分析单元,用于预先定义指定的关键词为敏感文字和/或带指定内容的图片为敏感图片;对所抓取 Internet 网站页面内容进行监测过滤,判断是否包含所述敏感文字和/或所述敏感图片;当检测到所抓取 Internet 网站页面内容中包含所述敏感文字和/或所述敏感图片,则控制立即预警并输出相应的检测分析结果;

第五检测分析单元,用于将所抓取 Internet 网站页面内容,与历史数据进行匹配比较,进行页面变动检测分析和/或目录变更检测分析包括 js 脚本、html 脚本、a 链接、图片链接、文字内容信息,计算页面变动和/或目录变更百分比,当识别页面和/或目录被恶意篡改信息,则控制立即预警并输出相应的检测分析结果;

第六检测分析单元,用于对所抓取 Internet 网站页面内容按预定的规则进行主机信息审计检测分析具体包括:对主机的各种基线,进行检查和分析,将各外网应用系统按照预定的策略进行配置,对不合规内容进行预警;

第七检测分析单元,用于对所抓取 Internet 网站页面内容按预定的规则进行日志采集检测分析具体包括:通过对网络设备、主机、数据库、中间件、应用程序的日志文件进行采集和关联分析,动态识别出外网应用系统遭受攻击的情况,并检测攻击是否造成危害影响,当检测攻击是造成危害影响,则控制立即预警并输出相应的检测分析结果。

对 Internet 信息系统恶意攻击的检测方法及系统

技术领域

[0001] 本发明涉及网络检测技术领域,尤其涉及的是一种对网站恶意攻击的检测方法及系统。

背景技术

[0002] 网站恶意攻击的检测通过对网站进行扫描,分析扫描数据并提取可疑恶意信息。

[0003] 国内外网站恶意攻击的检测方式有漏洞扫描、人工检查和渗透测试。漏洞扫描是采用安全漏洞扫描软件或设备,以网络扫描的方式发现漏洞,自动完成检查工作;人工检查是以系统管理员方式登陆系统,查找工具扫描无法发现的安全漏洞,了解网站配置信息和查看网站受攻击情况;渗透测试是通过模拟恶意黑客的攻击方法,来检测网站安全的一种方法。

[0004] 目前,网站恶意攻击检测工具分为网络漏洞扫描器、数据库漏洞扫描器、应用漏洞扫描器和渗透性测试工具,这类工具通常包括黑客工具、脚本文件。渗透性测试的目的是检测已发现的脆弱性是否真正会给系统或网络带来影响。通常渗透性工具与脆弱性扫描工具一起使用,并可能会对检测网站的运行带来一定影响。当网站受到攻击时,往往会安装防火墙等工具来躲避攻击,在网站性能降低的同时,也需花费大量资金。

[0005] 除此,还有通过日志查看网站是否被恶意攻击。通常,访问网站时系统会把访问日志记录起来,人工查看海量的日志信息,通过解读日志文件来了解网站的健康状态。

[0006] 现有技术对网站恶意攻击检测过程中,存在如下问题:

(1) 检测工作量大:漏洞扫描繁琐,需要人工值守和配置,扫描结果的解读也需要大量时间;

(2) 专家依赖性高:不管是漏洞扫描、人工检查还是渗透测试,都依赖经验丰富的专家,扫描结果的验证和渗透测试严重依赖个人素质和技术能力,不同人给出的结果可能差异较大;

(3) 检测不够全面:对一个新的测评对象,除了从漏洞方面检测,还需要从应用程序本身的内容、安全等角度进行评估,需要快速的收集其主机、应用服务器、数据库服务器等信息,并进行综合分析;

(4) 缺乏分析手段:在检测过程中由于使用多种测试工具,数据不集中,且缺乏综合的平台进行综合展示和分析,使得安全能力成熟度、安全风险趋势等无法进行高级分析;

(5) 内容安全重视度不够:网站最终展示给客户的是直观的网页内容。对于网页上的文字、图片、恶意代码、外部链接等,一般成为“黑客”恶意行为的主要“载体”,通过对网站内容的高频率分析,是快速识别恶意行为、弥补安全防护手段的有效方式;

(6) 安全检测频率不高:目前的恶意攻击检测,基本是静态检测;但是对外网站变更频繁,没法及时发现安全漏洞和恶意行为结果。

[0007] 因此,现有技术还有待于改进和发展。

发明内容

[0008] 本发明要解决的技术问题在于,针对现有技术的上述缺陷,提供一种对网站恶意攻击的检测方法及系统,其检测全面,提高了检测的安全,并且减少了工作量,降低了人力成本。

[0009] 本发明解决技术问题所采用的技术方案如下:

一种对 Internet 信息系统恶意攻击的检测方法,其中,包括步骤:

A、定期抓取 Internet 网站页面内容并存储;

B、对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析,敏感文字检测分析,敏感图片检测分析,目录变更检测分析,主机信息审计检测分析,断链/坏链检测分析,页面变动检测分析和/或日志采集检测分析;

C、根据检测分析的结果生成相应的检测报告,并将所述检测报告输出显示。

[0010] 所述对 Internet 信息系统恶意攻击的检测方法,其中,所述步骤A具体包括:通过 NMAP 协议和/或 SYSLOG 协议定期抓取 Internet 网站页面内容。

[0011] 所述对 Internet 信息系统恶意攻击的检测方法,其中,所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析具体包括:

B10、采用挂马链接匹配技术,通过公认的木马网站及恶意链接收集各种公认的木马信息,并将收集的木马信息建立一个恶意链接库;

B11、在对所抓取 Internet 网站页面内容的网页源码进行分析时,将所抓取 Internet 网站页面内容的网页源码与所述恶意链接库存储的木马信息进行比较,判断是否有与所述恶意链接库存储的木马信息相同的信息;

B12、当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述恶意链接库中的链接相匹配,则控制立即预警并输出相应的检测分析结果。

[0012] 所述对 Internet 信息系统恶意攻击的检测方法,其中,所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析具体还包括:

B20、采用网页源码分析技术,针对所述恶意链接库中未收集的木马链接,根据木马在浏览器中隐藏自身的通用特征,对木马链接的隐藏技术进行分析,结合不同类型的挂马形式,分别提取出各类型挂马特征,组装成挂马特征库;

B21、将所抓取 Internet 网站页面内容的网页源码与所述挂马特征库存储的木马信息进行比较,判断是否有与所述挂马特征库存储的木马信息相同的信息;

B22、当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述挂马特征库中的链接相匹配,则控制立即预警并输出相应的检测分析结果。

[0013] 所述对 Internet 信息系统恶意攻击的检测方法,其中,所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析具体还包括:

B30、采用蜜罐技术,利用虚拟技术搭建一台无任何防范的服务器,将可疑的网页链接在该服务器的浏览器中打开,监测网页打开后的系统状况;

B31、判断系统时间是否被修改、关键的系统目录下是否有文件写入、系统文件是否被篡改、注册表是否被篡改、网络流量是否异常、和/或是否连接了可疑端口或 IP;

B32、当判断系统时间被修改、关键的系统目录下有文件写入、系统文件被篡改、注册表被篡改、网络流量异常、和/或连接了可疑端口或 IP,则控制立即预警并输出相应的检测分

析结果。

[0014] 所述对 Internet 信息系统恶意攻击的检测方法,其中,所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行敏感文字检测分析和 / 或敏感图片检测分析具体包括:

B40、预先定义指定的关键词为敏感文字和 / 或带指定内容的图片为敏感图片;

B41、对所抓取 Internet 网站页面内容进行监测过滤,判断是否包含所述敏感文字和 / 或所述敏感图片;

B42、当检测到所抓取 Internet 网站页面内容中包含所述敏感文字和 / 或所述敏感图片,则控制立即预警并输出相应的检测分析结果。

[0015] 所述对 Internet 信息系统恶意攻击的检测方法,其中,所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行页面变动检测分析和 / 或目录变更检测分析具体包括:

B50、将所抓取 Internet 网站页面内容,与历史数据进行匹配比较,

B51、进行页面变动检测分析和 / 或目录变更检测分析包括 js 脚本、html 脚本、a 链接、图片链接、文字内容信息,计算页面变动和 / 或目录变更百分比,

B52、当识别页面和 / 或目录被恶意篡改信息,则控制立即预警并输出相应的检测分析结果。

[0016] 所述对 Internet 信息系统恶意攻击的检测方法,其中,所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行主机信息审计检测分析具体包括:对主机的各种基线,进行检查和分析,将各外网应用系统按照预定的策略进行配置,对不合规内容进行预警;

所述步骤 B 中的对所抓取 Internet 网站页面内容按预定的规则进行日志采集检测分析具体包括:通过对网络设备、主机、数据库、中间件、应用程序的日志文件进行采集和关联分析,动态识别出外网应用系统遭受攻击的情况,并检测攻击是否造成危害影响,当检测攻击是造成危害影响,则控制立即预警并输出相应的检测分析结果。

[0017] 一种对 Internet 信息系统恶意攻击的检测系统,其中,包括:

抓取模块,用于定期抓取 Internet 网站页面内容并存储;

检测分析模块,用于对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析,敏感文字检测分析,敏感图片检测分析,目录变更检测分析,主机信息审计检测分析,断链 / 坏链检测分析,页面变动检测分析和 / 或日志采集检测分析;

生成模块,用于根据检测分析的结果生成相应的检测报告,并将所述检测报告输出显示。

[0018] 所述对 Internet 信息系统恶意攻击的检测系统,其中,所述检测分析模块包括:

第一检测分析单元,用于采用挂马链接匹配技术,通过公认的木马网站及恶意链接收集各种公认的木马信息,并将收集的木马信息建立一个恶意链接库;在对所抓取 Internet 网站页面内容的网页源码进行分析时,将所抓取 Internet 网站页面内容的网页源码与所述恶意链接库存储的木马信息进行比较,判断是否有与所述恶意链接库存储的木马信息相同的信息;当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述恶意链接库中的链接相匹配,则控制立即预警并输出相应的检测分析结果;

第二检测分析单元,用于采用网页源码分析技术,针对所述恶意链接库中未收集的木马链接,根据木马在浏览器中隐藏自身的通用特征,对木马链接的隐藏技术进行分析,结合不同类型的挂马形式,分别提取出各类型挂马特征,组装成挂马特征库;将所抓取 Internet 网站页面内容的网页源码与所述挂马特征库存储的木马信息进行比较,判断是否有与所述挂马特征库存储的木马信息相同的信息;当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述挂马特征库中的链接相匹配,则控制立即预警并输出相应的检测分析结果;

第三检测分析单元,用于采用蜜罐技术,利用虚拟技术搭建一台无任何防范的服务器,将可疑的网页链接在该服务器的浏览器中打开,监测网页打开后的系统状况;判断系统时间是否被修改、关键的系统目录下是否有文件写入、系统文件是否被篡改、注册表是否被篡改、网络流量是否异常、和/或是否连接了可疑端口或 IP;当判断系统时间被修改、关键的系统目录下有文件写入、系统文件被篡改、注册表被篡改、网络流量异常、和/或连接了可疑端口或 IP,则控制立即预警并输出相应的检测分析结果;

第四检测分析单元,用于预先定义指定的关键词为敏感文字和/或带指定内容的图片为敏感图片;对所抓取 Internet 网站页面内容进行监测过滤,判断是否包含所述敏感文字和/或所述敏感图片;当检测到所抓取 Internet 网站页面内容中包含所述敏感文字和/或所述敏感图片,则控制立即预警并输出相应的检测分析结果;

第五检测分析单元,用于将所抓取 Internet 网站页面内容,与历史数据进行匹配比较,进行页面变动检测分析和/或目录变更检测分析包括 js 脚本、html 脚本、a 链接、图片链接、文字内容信息,计算页面变动和/或目录变更百分比,当识别页面和/或目录被恶意篡改信息,则控制立即预警并输出相应的检测分析结果;

第六检测分析单元,用于对所抓取 Internet 网站页面内容按预定的规则进行主机信息审计检测分析具体包括:对主机的各种基线,进行检查和分析,将各外网应用系统按照预定的策略进行配置,对不合规内容进行预警;

第七检测分析单元,用于对所抓取 Internet 网站页面内容按预定的规则进行日志采集检测分析具体包括:通过对网络设备、主机、数据库、中间件、应用程序的日志文件进行采集和关联分析,动态识别出外网应用系统遭受攻击的情况,并检测攻击是否造成危害影响,当检测攻击是造成危害影响,则控制立即预警并输出相应的检测分析结果。

[0019] 本发明所提供的一种对网站恶意攻击的检测方法及系统,其检测全面,提高了检测的安全,并且减少了工作量,降低了人力成本,为用户提供了方便。并具有如下优点:

1、针对网站恶意攻击检测新技术,整合不同的测评工具,所有检测工作在一个中心统一管理,对一个网站的检测策略,检测任务等统一管理和下发,减轻测评人员的工作量,保证检测工作的有序高效进行;

2、针对网站恶意攻击检测新技术,研究和建立一个优化的检测过程,从主机、应用服务器、数据库等信息的发现,到对应网站的漏洞扫描和分析,到漏洞确认和渗透测试,到在线的检测等,整个过程在平台 IT 化的支持下运行;

3、对网站开展文字、图片、外链、挂马等检测,需要借助强大的分析能力和实时更新能力。平台需要与商业化的第三方云安全厂商合作,确保平台的检测和评估能力具备国际领先水平。

[0020] 4、针对网站,新技术支持“一次性检测”和“长期运营检测”两种模式。并且具备成熟的任务申请、调度、执行、告知、报告等功能。

[0021] 5、新技术能根据网站检测数据,分别从纵向和横线分析网站的安全性,并给出打份,便于直观、客观查看网站的健壮性。

附图说明

[0022] 图1是本发明的一种对Internet信息系统恶意攻击的检测方法的较佳实施例的流程图。

[0023] 图2是本发明的一种对Internet信息系统恶意攻击的检测系统的原理框图。

[0024] 图3是本发明的一种对Internet信息系统恶意攻击的检测系统中的检测分析模块的原理框图。

具体实施方式

[0025] 为使本发明的目的、技术方案及优点更加清楚、明确,以下参照附图并举实施例对本发明进一步详细说明。应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0026] 如图1所示的是本发明的一种对Internet信息系统恶意攻击的检测方法的较佳实施例的流程图。具体包括以下步骤:

S100、定期抓取Internet网站页面内容并存储。

[0027] 在具体实施例中,主要是通过NMAP协议和/或SYSLOG协议定期抓取Internet网站页面内容。其中NMap,也就是Network Mapper,是Linux下的网络扫描和嗅探工具包。SYSLOG协议是用于系统日志或系统记录,是一种用来在互联网协定(TCP/IP)的网络中传递记录档讯息的标准。

[0028] 本实施例中对需进行检测的Internet网站页面上内容每隔一定时间进行扫描并提取,然后将提取到数据按照内容类型进行归类存储,所有提取的数据都暂存于一个预先设置的数据库中。所抓取的Internet网站页面内容不仅包括页面文字内容,图片,链接,脚本等信息,也包括主机、应用服务器、数据库服务器等信息。

[0029] S200、对所抓取Internet网站页面内容按预定的规则进行木马检测分析,敏感文字检测分析,敏感图片检测分析,目录变更检测分析,主机信息审计检测分析,断链/坏链检测分析,页面变动检测分析和/或日志采集检测分析。

[0030] 根据不同恶意代码的不同攻击方式以及目前网站漏洞被攻击和被利用的特点,制定不同的规则从不同角度的,全方位的对于所抓取的网站内容进行木马检测分析,主要包括敏感文字检测分析,敏感图片检测分析,目录变更检测分析,主机信息审计检测分析,断链/坏链检测分析,页面变动检测分析和/或日志采集检测分析,具体在如下所述。这种从主机、应用服务器、数据库等信息的发现,到对应网站的漏洞扫描和分析,到漏洞确认和渗透测试,到在线的检测等整个过程在平台IT化的支持下形成了立体化检测,保证了在检测策略上的全覆盖,不留死角。

[0031] S300、根据检测分析的结果生成相应的检测报告,并将所述检测报告输出显示。

[0032] 具体实施过程中,在检测报告输出显示同时,对报告中出现的危险信息或者可疑

信息发出预警,并对该类危险 / 可疑信息的风险级别给出评级,按照风险由高到低的顺序进行显示,同时按照预设的解决方案同步显示相应地初步处置方案报告,对于解决方案中未能识别和解决的信息,初步处置方案报告中将该类信息单独归类并突出显示。

[0033] 目前网页木马主要有两种存在形式,一种形式是将包含恶意代码的链接(木马链接)嵌入到正常网页源代码中并且在浏览器中隐藏自身,使用户在浏览该网页时不自觉打开木马链接,从而达到攻击的目的;另一种形式是将恶意代码直接嵌入到正常网页源代码中,使用户在浏览网页时即刻运行恶意代码,从而达到攻击的目的。

[0034] 因此针对以上两种存在形式,具体实施过程中,所述步骤 S200 中的对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析具体包括:

S210、采用挂马链接匹配技术,通过公认的木马网站及恶意链接收集各种公认的木马信息,并将收集的木马信息建立一个恶意链接库。该恶意链接库也是自动定期进行木马信息的收集和汇编,更新,在不断更新过程中,使得该恶意链接库的信息量不断扩展,检测效能不断提高。

[0035] S211、在对所抓取 Internet 网站页面内容的网页源码进行分析时,将所抓取 Internet 网站页面内容的网页源码与所述恶意链接库存储的木马信息进行比较,判断是否有与所述恶意链接库存储的木马信息相同的信息,若有,则证明 Internet 网站正常页源代码中嵌入并隐藏了木马信息,即需要对该相关信息作出处理。

[0036] S212、当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述恶意链接库中的链接相匹配,则控制立即预警并输出相应的检测分析结果。

[0037] 同时按照预设的解决方案同步显示相应地初步处置方案报告,对于解决方案中未能识别和解决的信息,初步处置方案报告中将该类信息单独归类并突出显示。

[0038] 在进一步的具体实施例中,所述步骤 S200 中的对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析具体还包括:

S220、采用网页源码分析技术,针对所述恶意链接库中未收集的木马链接,根据木马在浏览器中隐藏自身的通用特征,对木马链接的隐藏技术进行分析,结合不同类型的挂马形式,如 JS 调用型挂马、框架嵌入式挂马、图片伪装挂马等,分别提取出各类型挂马特征,组装成挂马特征库。该挂马特征库也是自动定期进行挂马类型信息的收集和汇编,更新,在不断更新过程中,使得该挂马特征库的信息量不断扩展,检测效能不断提高。

[0039] S221、将所抓取 Internet 网站页面内容的网页源码与所述挂马特征库存储的木马信息进行比较,判断是否有与所述挂马特征库存储的木马信息相同的信息。

[0040] S222、当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述挂马特征库中的链接相匹配,则控制立即预警并输出相应的检测分析结果。同时按照预设的解决方案同步显示相应地初步处置方案报告,对于解决方案中未能识别和解决的信息,初步处置方案报告中将该类信息单独归类并突出显示。

[0041] 在进一步的具体实施例中,所述步骤 S200 中的对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析具体还包括:

S230、针对恶意代码直接嵌入到正常网页源代码中的形式采用蜜罐技术,利用虚拟技术搭建一台无任何防范的服务器,将可疑的网页链接在该服务器的浏览器中打开,监测网页打开后的系统状况。

[0042] S231、判断系统时间是否被修改、关键的系统目录下是否有文件写入、系统文件是否被篡改、注册表是否被篡改、网络流量是否异常、和 / 或是否连接了可疑端口或 IP。除上述重要信息的检测,本发明系统也会根据木马更新类型及“发病”特点灵活改变所监控的内容。

[0043] S232、当判断系统时间被修改、关键的系统目录下有文件写入、系统文件被篡改、注册表被篡改、网络流量异常、和 / 或连接了可疑端口或 IP,则控制立即预警并输出相应的检测分析结果。同时按照预设的解决方案同步显示相应地初步处置方案报告,对于解决方案中未能识别和解决的信息,初步处置方案报告中将该类信息单独归类并突出显示。

[0044] 在进一步的具体实施例中,所述步骤 S200 中的对所抓取 Internet 网站页面内容按预定的规则进行敏感文字检测分析和 / 或敏感图片检测分析具体包括:

S240、预先定义指定的关键词为敏感文字和 / 或带指定内容的图片为敏感图片。本系统采用智能分词技术,以便有效监测关键词。分词技术属于自然语言理解技术的范畴,是语义理解的首要环节,它是能将语句中的词语正确切分开的一种技术,它是文本分类,信息检索,机器翻译,自动标引,文本的语音输入输出等领域的基础,而由于中文本身的复杂性及其书写习惯,使中文分词技术成为了分词技术中的难点。本系统采用目前主流的智能分词技术,以便敏感文字检测分析能保证检测分析的质量。

[0045] 对于图片检测,重点针对带指定内容的敏感图片的检测,其检测方法主要包括去背影、人脸识别、肤色识别、骨架分析、向量分析等技术,针对所获取的敏感图片信息建立有效的特征库,从而可进一步提高图片的识别率。

[0046] S241、对所抓取 Internet 网站页面内容进行监测过滤,判断是否包含所述敏感文字和 / 或所述敏感图片;

S242、当检测到所抓取 Internet 网站页面内容中包含所述敏感文字和 / 或所述敏感图片,则控制立即预警并输出相应的检测分析结果。同时按照预设的解决方案同步显示相应地初步处置方案报告,对于解决方案中未能识别和解决的信息,初步处置方案报告中将该类信息单独归类并突出显示。

[0047] 而本实施例中所述步骤 S200 中的对所抓取 Internet 网站页面内容按预定的规则进行页面变动检测分析和 / 或目录变更检测分析具体包括:

S250、将所抓取 Internet 网站页面内容,与历史数据进行匹配比较。

通过网络爬虫利用网页的链接地址来寻找网页,从网站某一个页面(通常是首页)开始,读取网页的内容,找到在网页中的其它链接地址,然后通过这些链接地址寻找下一个网页,这样一直循环下去,直到把这个网站所有的网页都抓取完为止,将所抓取的页面内容与数据库中存储的历史数据进行匹配比较,包括 js 脚本、html 脚本、a 链接、图片链接、文字内容信息。

[0048] S251、进行页面变动检测分析和 / 或目录变更检测分析包括 js 脚本、html 脚本、a 链接、图片链接、文字内容信息,根据预设规则计算页面变动和 / 或目录变更百分比。

[0049] S252、预先设置不同的页面变动和 / 或目录变更百分比对应不同的危险等级。变动率越高,其危险等级越高,根据所计算的页面变动和 / 或目录变更百分比迅速识别页面和 / 或目录被恶意篡改信息,立即预警并输出相应的检测分析结果。同时按照预设的解决方案同步显示相应地初步处置方案报告,对于解决方案中未能识别和解决的信息,初步处

置方案报告中将该类信息单独归类并突出显示。

[0050] 所述步骤 S200 中的对所抓取 Internet 网站页面内容按预定的规则进行主机信息审计检测分析具体包括：对主机的各种基线，进行检查和分析，将各外网应用系统按照预定的策略进行配置，对不合规内容进行预警。

[0051] 所述步骤 S200 中的对所抓取 Internet 网站页面内容按预定的规则进行日志采集检测分析具体包括：通过对网络设备、主机、数据库、中间件、应用程序的日志文件进行采集和关联分析，动态识别出外网应用系统遭受攻击的情况，并检测攻击是否造成危害影响，当检测攻击是造成危害影响，则控制立即预警并输出相应的检测分析结果。同时按照预设的解决方案同步显示相应地初步处置方案报告，对于解决方案中未能识别和解决的信息，初步处置方案报告中将该类信息单独归类并突出显示。

[0052] 另外，所述步骤 S200 中包括对所抓取 Internet 网站页面内容按预定的规则进行断链/坏链检测分析，其是检测网站页面 a 链接，记录不可访问链接，并对其进行分析统计，识别断链/坏链来源，并出具报告，该报告为保障网站页面可用、快速访问提供了决策依据。

[0053] 由上可见，本发明实施例的对网站恶意攻击的检测方法及系统，其检测全面，提高了检测的安全，并且减少了工作量，降低了人力成本。

[0054] 基于上述实施例，本发明实施例还提供了一种对 Internet 信息系统恶意攻击的检测系统，如图 2 所示，所述对 Internet 信息系统恶意攻击的检测系统，包括：

抓取模块 100，用于定期抓取 Internet 网站页面内容并存储，具体如上述 S100 步骤所述。

[0055] 检测分析模块 200，用于对所抓取 Internet 网站页面内容按预定的规则进行木马检测分析，敏感文字检测分析，敏感图片检测分析，目录变更检测分析，主机信息审计检测分析，断链/坏链检测分析，页面变动检测分析和/或日志采集检测分析，具体如上述 S200 步骤所述。

[0056] 生成模块 300，用于根据检测分析的结果生成相应的检测报告，并将所述检测报告输出显示，具体如上述 S300 步骤所述。

[0057] 如图 3 所述是本发明的一种对 Internet 信息系统恶意攻击的检测系统中的检测分析模块 200 的原理框图。所述检测分析模块 200 包括：

第一检测分析单元 210，用于采用挂马链接匹配技术，通过公认的木马网站及恶意链接收集各种公认的木马信息，并将收集的木马信息建立一个恶意链接库；在对所抓取 Internet 网站页面内容的网页源码进行分析时，将所抓取 Internet 网站页面内容的网页源码与所述恶意链接库存储的木马信息进行比较，判断是否有与所述恶意链接库存储的木马信息相同的信息；当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述恶意链接库中的链接相匹配，则控制立即预警并输出相应的检测分析结果，具体如上述步骤所述。

[0058] 第二检测分析单元 220，用于采用网页源码分析技术，针对所述恶意链接库中未收集的木马链接，根据木马在浏览器中隐藏自身的通用特征，对木马链接的隐藏技术进行分析，结合不同类型的挂马形式，分别提取出各类型挂马特征，组装成挂马特征库；将所抓取 Internet 网站页面内容的网页源码与所述挂马特征库存储的木马信息进行比较，判断是否

有与所述挂马特征库存储的木马信息相同的信息；当所抓取 Internet 网站页面内容的网页源码中包含的链接与所述挂马特征库中的链接相匹配，则控制立即预警并输出相应的检测分析结果，具体如上述步骤所述。

[0059] 第三检测分析单元 230，用于采用蜜罐技术，利用虚拟技术搭建一台无任何防范的服务器，将可疑的网页链接在该服务器的浏览器中打开，监测网页打开后的系统状况；判断系统时间是否被修改、关键的系统目录下是否有文件写入、系统文件是否被篡改、注册表是否被篡改、网络流量是否异常、和 / 或是否连接了可疑端口或 IP；当判断系统时间被修改、关键的系统目录下有文件写入、系统文件被篡改、注册表被篡改、网络流量异常、和 / 或连接了可疑端口或 IP，则控制立即预警并输出相应的检测分析结果，具体如上述步骤所述。

[0060] 第四检测分析单元 240，用于预先定义指定的关键词为敏感文字和 / 或带指定内容的图片为敏感图片；对所抓取 Internet 网站页面内容进行监测过滤，判断是否包含所述敏感文字和 / 或所述敏感图片；当检测到所抓取 Internet 网站页面内容中包含所述敏感文字和 / 或所述敏感图片，则控制立即预警并输出相应的检测分析结果，具体如上述步骤所述。

[0061] 第五检测分析单元 250，用于将所抓取 Internet 网站页面内容，与历史数据进行匹配比较，进行页面变动检测分析和 / 或目录变更检测分析包括 js 脚本、html 脚本、a 链接、图片链接、文字内容信息，计算页面变动和 / 或目录变更百分比，当识别页面和 / 或目录被恶意篡改信息，则控制立即预警并输出相应的检测分析结果，具体如上述步骤所述。

[0062] 第六检测分析单元 260，用于对所抓取 Internet 网站页面内容按预定的规则进行主机信息审计检测分析具体包括：对主机的各种基线，进行检查和分析，将各外网应用系统按照预定的策略进行配置，对不合规内容进行预警，具体如上述步骤所述。

[0063] 第七检测分析单元 270，用于对所抓取 Internet 网站页面内容按预定的规则进行日志采集检测分析具体包括：通过对网络设备、主机、数据库、中间件、应用程序的日志文件进行采集和关联分析，动态识别出外网应用系统遭受攻击的情况，并检测攻击是否造成危害影响，当检测攻击是造成危害影响，则控制立即预警并输出相应的检测分析结果，具体如上述步骤所述。

[0064] 综上所述，本发明所提供的一种对网站恶意攻击的检测方法及系统，其检测全面，提高了检测的安全，并且减少了工作量，降低了人力成本，为用户提供了方便。并具有如下优点：

1、针对网站恶意攻击检测新技术，整合不同的测评工具，所有检测工作在一个中心统一管理，对一个网站的检测策略，检测任务等统一管理和下发，减轻测评人员的工作量，保证检测工作的有序高效进行；

2、针对网站恶意攻击检测新技术，研究和建立一个优化的检测过程，从主机、应用服务器、数据库等信息的发现，到对应网站的漏洞扫描和分析，到漏洞确认和渗透测试，到在线的检测等，整个过程在平台 IT 化的支持下运行；

3、对网站开展文字、图片、外链、挂马等检测，需要借助强大的分析能力和实时更新能力。平台需要与商业化的第三方云安全厂商合作，确保平台的检测和评估能力具备国际领先水平。

[0065] 4、针对网站，新技术支持“一次性检测”和“长期运营检测”两种模式。并且具备

成熟的任务申请、调度、执行、告知、报告等功能。

[0066] 5、新技术能根据网站检测数据,分别从纵向和横线分析网站的安全性,并给出打份,便于直观、客观查看网站的健壮性。

[0067] 应当理解的是,本发明的应用不限于上述的举例,对本领域普通技术人员来说,可以根据上述说明加以改进或变换,例如,所有这些改进和变换都应属于本发明所附权利要求的保护范围。

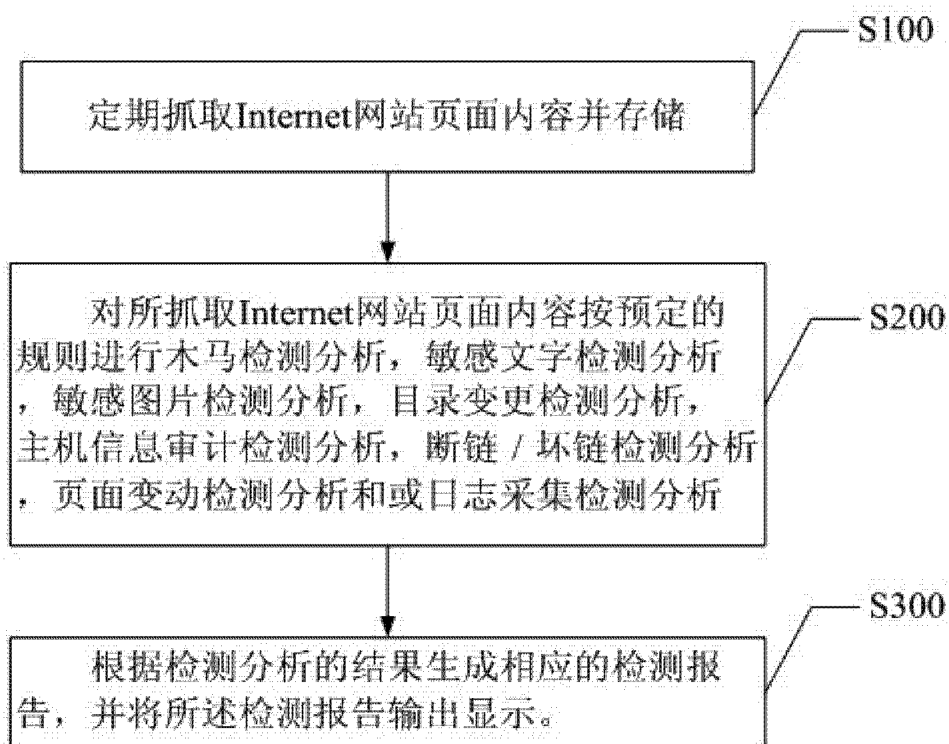


图 1

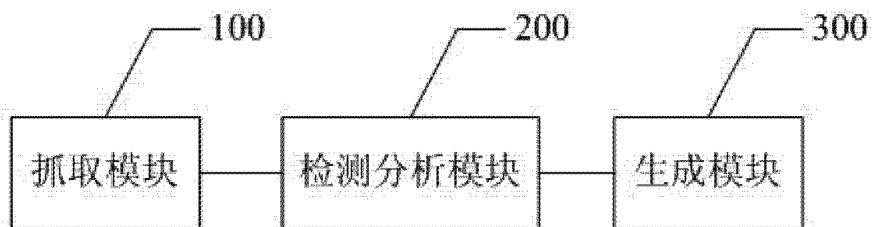


图 2

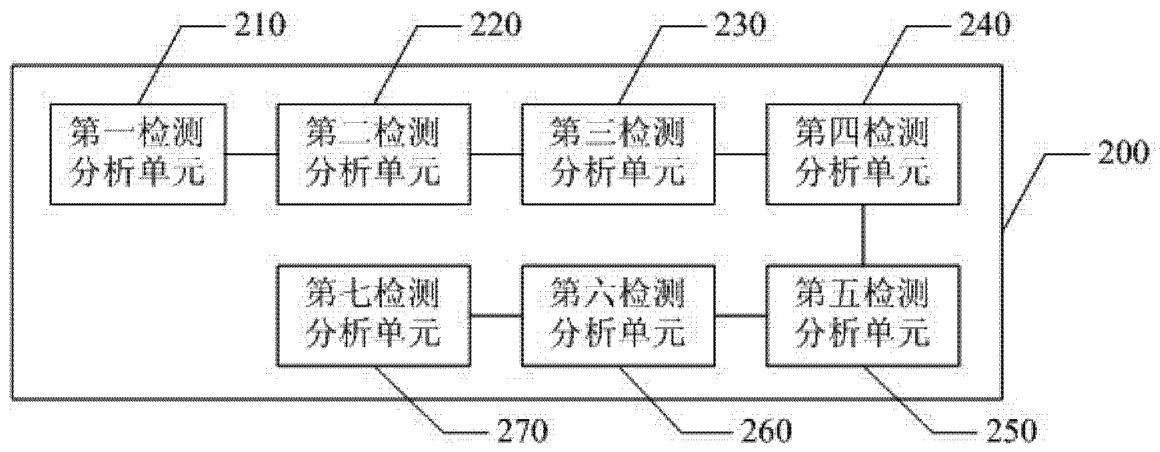


图 3