网站业务安全(BotGuard)防护
服务白皮书V1.0

网世界 心归宿

网宿科技 全球领先的互联网基础设施平台

关注公众号，及时了解网宿产品与服务动态

400-010-0617 | www.wangsu.com

Content 目录

网宿“业务安全（BotGuard）服务”为“网宿网盾”品牌旗下一款强大的网站业务安全防护产品，依托于大数据平台和智能行为分析技术，在云端形成 Bot 管理网络，然后基于平台配置的情报库、访问控制、指纹识别、布设陷阱、机器识别等技术，实时对网站流量进行检测和分析，识别出真假用户的流量，并针对不同 Bot 流量采用合理的管理策略。

“网宿网盾”是网宿科技发布的云安全全新品牌，其业务全景主要覆盖网络安全、业务安全、内容安全、DNS 安全、源站可用性、传输安全、安全管理和安全评估等八大领域，“网宿网盾”依托高容量节点和自主研发的安全技术，构建出一套云端智能安全体系。

行业现状和挑战

1.1. Bot 访问量越来越多

随着网络的迅速发展，互联网已成为大量信息的载体，如何有效地提取并利用这些信息成为一个巨大的挑战，因此 Bot 程序便应运而生。Bot 程序是指按照一定的规则自动的抓取互联网信息的程序或者脚本，已被广泛应用于互联网领域，对于多数企业网站来说，超过 50%的网站总流量来自于 Bot 程序而非正常用户。

1.2. 恶意 Bot 的常见应用场景

并不是所有的 Bot 流量都是企业所期望或者所排斥的，部分 Bot 如搜索引擎爬虫有利于网站的推广属于善意 Bot，而有些 Bot 程序被用于窃取商业竞争对手的敏感信息等，属于恶意 Bot。恶意 Bot 在企业业务中的常见应用场景主要体现在：

- 恶意注册

非法分子通过批量注册工具获取大量账号，并以此进行非法谋利。同时大量的恶意注册会导致网站运营者无法正确统计出正常用户量、注册转化率等信息，破坏网站的正常运营。

此外，有些企业网站注册页面需进行手机短信验证，非法者可能会通过频繁恶意注册来攻击企业注册的短信接口，耗尽短信接口资源，使用户无法注册，给企业造成损失。

- 非法登录

几乎每个网站的登录页面都面临着攻击威胁，非法者通过暴力破解工具或者撞库工具绕过网站的权限控制，获取用户账号权限访问更多敏感数据，并以此为基础导致一系列欺诈性行为。

- 活动作弊

非法者对互联网上各种活动进行作弊，例如：

1. 薅羊毛，例如：电商、网游行业经常做“0 元购”、“优惠券”、“红包”等活动来吸引用户，而这种方式在获取用户的同时了也催生了“羊毛党”。“羊毛党”通过 Bot 程序大量爬取优惠资源，并转卖二次获利。
2. 投票作弊，目前网络投票的范围也越来越广，内容涉及文艺评选、企业满意度、人物评选等社会各个领域，而非法者通过使用投票软件恶意刷票来获取相关利益，不仅造成投票结果远离公正公平，同时还损害了平台和用户的利益。
3. 其他活动作弊场景，例如点击/评论作弊、刷信誉、刷好评等。

- 恶意刷票/虚占座位

黄牛党通过各种抢票软件在互联网购票平台上进行抢票囤票、虚占座位，导致正常用户买不到票只能以高价向黄牛购买。

1.3. 恶意 Bot 带来的安全隐患

恶意 Bot 会给企业带来诸多影响：

- 影响企业活动效果

Bot 模拟正常用户行为参与注册、登陆和提交订单等业务过程，占用限量资源（如：商品、优惠、返现等），再进行高价二次转卖获利，不仅影响了活动的效果，也极大损害了商家的利益。

- 增加企业运营成本

大量恶意的 Bot 访问并不会给网站带来收益，反而增加了带宽、计算资源等成本。此外，大量自动登录、注册过程会频繁调用短信发送接口，浪费企业短信资源，增加了企业运营成本。

- 降低网站性能

大量的恶意 Bot 流量会给企业网站增加服务器负载，降低网站性能，无法为正常用户提供服务。

- 降低企业竞争力

商品定价、库存数据、知识产权、财务信息等信息被竞争对手利用 Bot 技术抓取，导致企业核心数据泄露。

- 用户流失、经济损失

非法者通过 Bot 程序盗取用户敏感信息，损害用户利益，导致企业的用户流失，进一步造成企业经济损失。

1.4. 传统 Bot 管理方式及不足

随着 Bot 流量的爆发，各种检测防护技术也随之应运而生，包括：请求限速、图片验证码、异常特征检测等方法，这些方法虽然在一定程度上可以缓解 Bot 流量，但是存在着以下不足：

✓ 误杀严重

传统的 Bot 防护方式对所有 Bot 流量采用相同的拦截策略，无论 Bot 程序是否对业务有益，导致善意 Bot 被严重误杀。

✓ 缓解能力不足

1. 请求限速需要设置个合理的阈值，然而由于 Web 网关、HTTP 反向代理、NAT 等技术会造成来自相同 IP 的访客数不确定，因此会影响阈值的设置，如阈值设置过高或过低，都将导致漏过攻击影响正常业务。
2. 部分 Bot 发送的请求头部与正常浏览器发出的请求差异较大（如：User-Agent 的内容不同），因此可以根据请求头部的异常特征来识别出 Bot；然而目前很多 Bot 工具可以伪造请求头部进而模拟浏览器请求，影响了识别准确度。
3. 由于图像识别及人工智能技术的发展及普及，计算机程序对图片验证码识别的准确性大幅增加，而为了对抗 Bot 使用的图像识别及人工智能技术，图片验证码增加了图片内容的识别难度，也同时增加了用户识别图片验证码的难度，影响了用户体验。

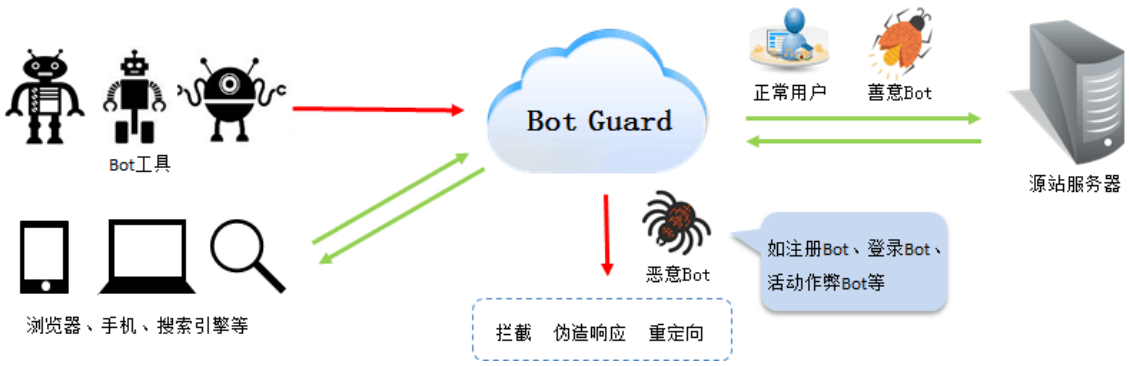
2.产品介绍

2.1. 产品简介

网宿业务安全（Bot Guard）产品，依托于大数据分析平台形成 Bot 管理网络。基于情报库、客户端速率控制、客户端指纹采集、布设陷阱、机器识别等技术，实时对网站流量进行检测和分析，识别真实用户流量、善意 Bot、恶意 Bot 流量，并针对不同 Bot 流量采用合理的管理策略，防止网站敏感信息被抓取，同时避免 Bot 流量占用大量服务器资源，保障企业业务稳定运行，保持竞争优势。业务安全平台内置的威胁评估模型会自动预测攻击趋势和网站风险，对于高风险攻击事件能够全网快速部署防御策略，防患于未然，保障网站安全稳定的运行。

2.2. 产品技术架构

网宿业务安全（Bot Guard）产品架构如下图所示：



2.3. 产品应用场景

网宿业务安全（BotGuard）产品应用的场景包括但不限于：

1、 注册场景

针对注册场景中，非法者使用恶意注册工具进行机器人注册、垃圾注册、短信滥刷等，BotGuard 能智能识别正常用户与恶意 Bot，降低恶意注册给企业带来的业务风险。

2、 登录场景

针对登录场景中，非法者使用恶意注册工具进行刷库撞库、暴力破解等账号盗取，BotGuard 能智能识别正常用户与恶意 Bot，避免用户账号信息而导致用户个人及企业经济损失。

3、 营销活动场景

针对互联网开展的活动场景中，非法者使用自动化程序进行“薅羊毛”、“刷单/刷信誉”、投票作弊等，BotGuard 能智能识别正常用户与恶意 Bot，保障活动效果。

4、 订票场景

针对订票场景中，非法者使用订票程序恶意刷票、抢票、虚占座位等，BotGuard 能智能识别正常用户与恶意 Bot，使用户能正常订票，保障用户和平台的利益。

5、 其他场景

此外，非法者还可通过 Bot 程序进行其它行为，例如：抓取竞争对手商品定价、库存/、知识产权信息、财务信息等业务数据和恶意点击广告等，BotGuard 均能智能识别正常用户与恶意 Bot。

2.4. 产品适用行业

网宿业务安全（BotGuard）产品面向的行业包括但不限于：

1、电商行业

近年来电商行业通过发放优惠券、红包等方式来获取用户、培养用户的消费习惯，而这种方式在获取用户的同时了也催生了“羊毛党”。“羊毛党”通过恶意 Bot 工具大量爬取优惠资源，通过转卖二次获利，不仅影响活动效果，也给商家造成了经济损失。

2、P2P 金融行业

线上 P2P 金融通常对新注册用户有一定的理财优惠，如年收益率翻倍、发现金红包等，容易被非法者觊觎爬取，使得企业旨在扩展自身业务的现金、优惠被黑产截获，无法发放到商家的正常目标客户，极大损害了商家的利益。

3、旅游服务行业

旅游服务行业包含旅游、交通、餐饮等一系列服务，其中诸多环节已成为黑客觊觎的对象，如在交通方面，航空售票、铁路售票等网络售票平台频繁遭受恶意 Bot 刷单、占座、恶意查询等；在景区网上订票系统中，一些“黄牛党”通过恶意 Bot 工具大量购买真票，并把价格炒高数倍获利；这些不仅危害了旅客的个人利益，也给各平台造成了经济损失。

4、政企行业

门户网站、信息公示系统作为政府、企业提供给互联网用户获取信息服务的重要渠道，频繁遭受非法者大量爬取，导致网站敏感信息泄露、服务器性能下降网站响应缓慢，严重损害了政企形象。

5、地产行业

地产、二手房交易类门户网站，往往依靠真实、独家的优质房源信息吸引客户，而这些独有资源亦会被竞争对手通过 Bot 工具所获取，从而丧失了竞争优势。

6、信息资讯行业

如招聘网站、文学博客、论坛、新闻网站等。这类网站以内容为王，恶意 Bot 将导致网站的核心文本在几小时甚至几分钟内就被抓取并复制到别的网站，极大影响网站在搜索引擎上的排名，而低排名会导致访问量降低和销量、广告收益降低等。

3.产品功能

网宿业务安全（BotGuard）产品，基于大数据分析，提供 Bot 定义、Bot 智能识别、Bot 管理、Bot 可视化等功能，能够对 Bot 流量进行智能分类与管理，保障企业业务安全稳定运行。

3.1. Bot 定义

网宿业务安全（BotGuard）将 Bot 分为善意 Bot、恶意 Bot 两类进行管理：

- 善意 Bot

善意 Bot 主要通过抓取网站的各项内容（如：网页内容、流量、载入时间等），用于企业网站的优化和推广，主要包括如下范围： ：

- ✓ 搜索引擎类 Bot，如：谷歌爬虫、百度爬虫、搜狗爬虫等；
- ✓ 网站流量监测和排名类 Bot，如：Alexa 公司用于网站流量监测和发布网站世界排名的工具 Archive Bot 等；
- ✓ 网站在线监控服务类 Bot，如：用于网页速度监测的 Pingdom Bot 等；

- ✓ 图片搜索引擎类 Bot，如：用于反向图片搜索的 Tineye Bot 等；
- ✓ ……

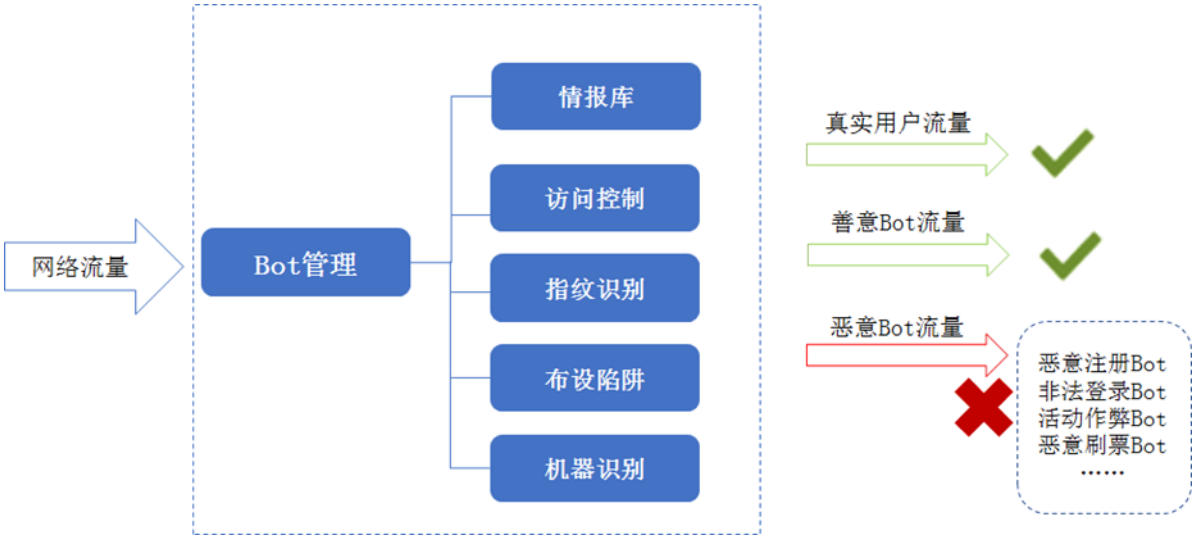
- ● 恶意 Bot

恶意 Bot 通过修改自己的请求方式模拟真实用户，可能潜伏在网站的每一个角落，文档、图片、价格、评论、接口等都有可能被爬取，会给企业网站带来诸多影响：如：核心资源被抓取、活动被“薅羊毛”等。根据恶意 Bot 在企业业务的应用场景，大概定义为以下几类：

- ✓ 恶意注册 Bot，如：批量注册工具等；
- ✓ 非法登录 Bot，如：暴力破解工具、撞库工具等；
- ✓ 活动作弊 Bot，如：投票软件、自动领红包工具、自动秒杀工具等；
- ✓ 黄牛刷票/虚拟占座 Bot，如：刷票/订票软件等；
- ✓ 其他 Bot，如：价格爬虫工具、恶意点击工具等。

3.2. Bot 智能识别

网宿业务安全（BotGuard）依托于大数据分析技术，对网站的请求进行检测与分析，智能识别出真实用户、善意 Bot 以及恶意 Bot 流量。



1、业务安全情报库

业务安全情报库包括威胁情报库和善意 Bot 白名单库，最大限度提高 Bot 识别效率。

1) 威胁情报库

实时汇总分析历史攻击事件的日志，提取攻击特征（如：IP、UA 等），并基于内置的威胁评分机制对攻击威胁等级进行评定。

2) 善意 Bot 库

善意 Bot(如：搜索引擎爬虫)一般会在 HTTP 头的 UA 中标识自己的身份，便于网站识别，但这类特征也较易模仿伪造，因此不能将此类特征作为识别善意 Bot 的唯一依据。BotGuard 通过反向域名解析技术形成善意 Bot 库，快速区分真实/伪造善意 Bot 程序。

同时，网宿通过与多家 SE 厂商、网站流量监测和排名厂商、网站在线监控服务厂商等合作，收录了 SE、网站流量监测和排名、网站在线监控等各类别的爬虫程序相关信息（包括：UA、IP 等），不断完善善意 Bot 库。

2、访问控制

由于反向代理与 NAT 技术的广泛使用，导致同一个 IP 下用户数量不确定，无法设置合理的 IP 限速阈值，影响了对 Bot 检测防护的精度。

业务安全平台为每一个首次访问网站的客户端添加“唯一标识信息”，针对每个客户端的访问进行区分跟踪，通过设定单客户端的访问频率阈值，防止大量恶意 Bot 流量占用服务器资源，同时速率阈值设置能够根据自学习引擎的学习结果动态自动调整。

3、指纹识别

业务安全平台可在响应页面中添加检测脚本，对客户端的各种特性进行校验（如：是否支持 JS、H5、Cookie 等属性），采集每个客户端的指纹信息，进而识别客户端为正常用户或者 Bot 工具。

4、布设陷阱

Bot 程序为了尽可能多抓取目标网站的信息，在抓取网页的过程中会不断从当前页面中识别抓取新的 URL 放入爬取目标中。业务安全平台利用该特性布设“陷阱”诱导 Bot 访问（如：设置隐藏的 URL，正常用户不会访问），进而识别出恶意 Bot 的请求。

5、机器识别

由于图片验证码识别难度的增加，采用该方式进行 Bot 检测会对用户体验产生极大影响。网宿业务安全（BotGuard）采用独创的 WIMC 技术，能够在不影响用户体验的基础上，智能识别 Bot 程序与正常用户。

网宿业务安全（Bot Guard）自动学习用户网站浏览习惯，通过行为分析（如：鼠标移动轨迹、页面访问逻辑等）建立用户行为模型，并为每个客户端添加事先设计好的交互场景，诱导用户下意识地进行简单的操作(如：鼠标移动、键盘敲击等)，通过客户端用户的行为数据进行监测分析是否为正常的用户反应，从而识别出 Bot 程序与正常用户。

3.3. Bot 管理

- **实时监控报警**

网宿业务安全（BotGuard）能够对 Bot 流量进行实时监控，以便第一时间发现异常流量并报警。

- **善意 Bot 管理**

善意 Bot 虽然对企业的业务推广有利，但是有些 Bot 程序对服务器负担较大(如：搜狗爬虫算法恶劣会对页面进行大量反复而无实际意义的扫描，增加服务器计算压力，抓取压力大；)，网站管理者可在业务安全（Bot Guard）平台上选择放行、限速或拒绝某类善意 Bot。

支持自定义善意 Bot 特征码（如：IP、UA 信息），以便识别正常与网站交互的 Bot 程序，以免误杀影响企业业务。

- **恶意 Bot 管理**

可通过 BotGuard 平台配置恶意 Bot 检测维度以及处理机制。

1) 检测维度：

支持自定义恶意 Bot 检测维度，如是否进行 JS 检测、HTML5 检测、用户行为检测等。

2) 处理机制：

针对识别出来的恶意 Bot，可采用拦截、限速、伪造响应、重定向等处理机制，满足网站不同的场景需求。

✓ 拦截

针对识别出来的 Bot 流量，直接阻断，避免恶意 Bot 流量影响企业业务。

✓ 伪造响应

支持网站管理者自定义响应内容，避免恶意 Bot 流量影响企业业务。

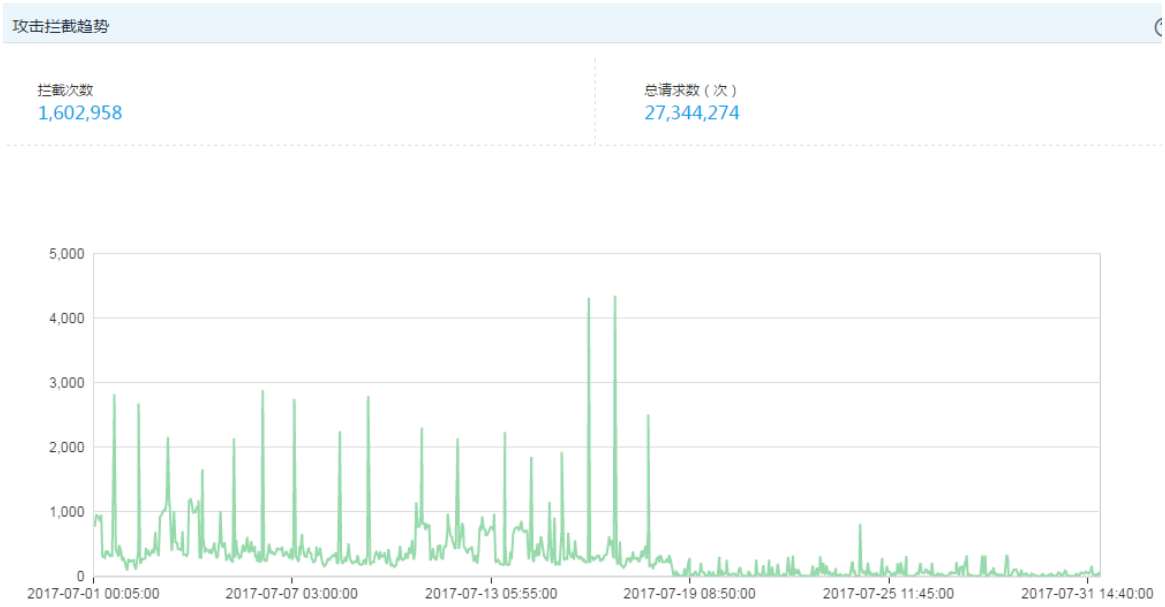
✓ 重定向

支持重定向到网站管理者指定的已有页面，避免恶意 Bot 流量影响企业业务。

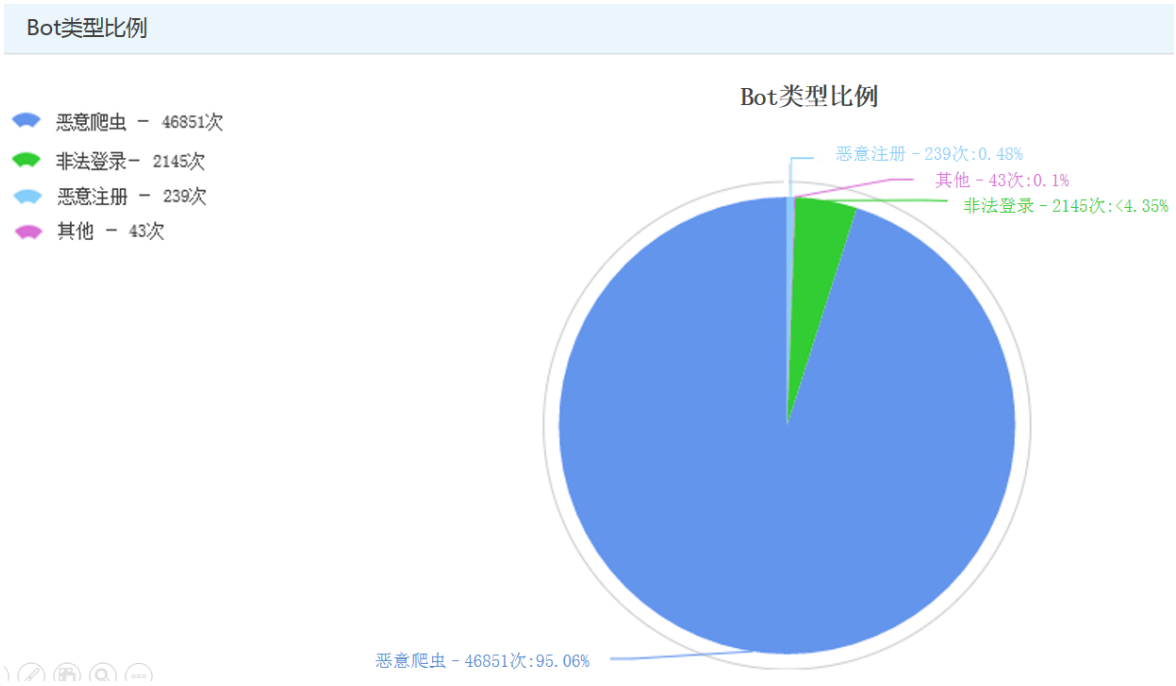
3.4. Bot 可视化

网宿业务安全（BotGuard）可实时展示 Bot 流量趋势、Bot 类型分布、Bot 流量来源和拦截情况等。

(1) 实时展示流量拦截趋势，网站管理者可以直观地了解各个时段的安全状况。



(2) 实时展示 Bot 流量类型分布，帮助网站管理者了解各业务系统的安全状况。



(3) 展示发起恶意 Bot 的 IP 详细信息，包括：IP 所在地、访问次数、拦截次数等，以利于网站管理者后续对非法者的处理。

攻击IP	所属区域	总访问次数	攻击类型	攻击次数
218.200.128.69	中国大陆	997	恶意注册	182
106.120.128.249	中国大陆	27,713	恶意爬虫	156
211.137.49.165	中国大陆	249	恶意注册	127
110.247.153.90	中国大陆	360	非法登录	107
103.225.196.217	香港	224	恶意爬虫 非法登陆 恶意注册	47 43 16
223.198.216.80	中国大陆	349	恶意注册	74
220.167.246.91	中国大陆	307	非法登录	69
211.137.52.135	中国大陆	150	恶意注册	69
1.25.82.184	中国大陆	229	恶意注册	64
60.164.204.140	中国大陆	126	恶意注册	54

(4) 展示恶意 Bot 所爬取的 URL 信息。

序号	攻击时间	URL
1	2017-07-18 09:40:00	http:/[redacted]gid=1000855960&r=goods/index
2	2017-07-18 09:40:00	http:/[redacted]d=1000514109&name=%E5%B7%A6%E6%89%8B%E5%80%92%E5%BD%B1%E5%8F%B3%E6%89
3	2017-07-18 09:30:00	http:/[redacted]gid=1002951565&r=goods/index
4	2017-07-18 09:00:00	http:/[redacted]id=1002172707&name=CROSS%20TOP%20%E5%A4%9A%E5%8A%9F%E8%83%BD%E6%8C%8E9
5	2017-07-18 05:10:00	http:/[redacted]gid=1000655121&r=goods/index
6	2017-07-18 05:10:00	http:/[redacted]gid=1002077940&r=goods/index

4.产品价值

4.1. 大数据安全分析，全网联动

网宿业务安全（Bot Guard）产品依托大数据分析平台，能够在云端进行攻击事件特征分析及关联分析，内置的威胁评估模型自动预测攻击趋势及风险，对于高风险攻击事件能够全网快速部署防御策略，防患于未然，保障网站正常运行。

4.2. Bot 合理管理，保障业务正常开展

网宿业务安全（BotGuard）产品能够采用合理的措施管理善意 Bot 和恶意 Bot 流量，避免 Bot 流量占用大量服务器、带宽等资源，降低企业运营成本，保障网站业务稳定运行。

4.3. 智能人机识别，用户无感知

网宿业务安全（Bot Guard）产品通过机器识别技术智能区分正常请求与 Bot 请求，不需要网站访问者去识别复杂的图片验证码等，在保证企业业务安全的同时，不影响用户体验。

4.4. 私有内容防窃取，保持竞争优势

网宿业务安全（Bot Guard）产品能够防止企业私有内容（如：价格、库存等信息）被竞争对手通过 Bot 程序窃取，保持企业竞争优势。

关于网宿

网宿科技 始创于 2000 年 1 月，主要提供互联网内容分发与加速（CDN）、云计算、云安全、全球分布式数据中心(IDC) 等服务。

2009 年 10 月，网宿科技在深交所上市，股票代码 300017。

网宿科技拥有遍布全球的 1000+ CDN 加速节点，在北京、上海、广州、深圳等地设有分公司，在美国、香港、印度、爱尔兰、马来西亚、济南、南京、杭州等地建有多家全资子公司，并在厦门及美国硅谷设立了研发中心。现有员工 3000 多名，研发以及技术人员占总人数 60% 左右。客户群覆盖各类互联网门户网站、视音频网站、网络游戏公司、电子商务网站、政府网站、企业网站以及运营商等，公司服务的客户超过 3000 家，是市场同类公司中拥有客户数量较多、行业覆盖面较广的公司。