**北京航空航天大学计算机学院**

**硕士学位论文开题报告**

**论文题目**：恶意域名检测技术研究

**专 业**：网络与信息安全

**研究方向**：网络安全

**研 究 生**：王文博

**学 号**：ZY1506221

**指导教师**：兰雨晴 周渊

**北京航空航天大学计算机学院**

2016年12月13日

目 录

1. 论文选题的背景和意义 2

1.1. 相关攻击手段 2

1.1.1. 域名阴影 2

1.1.2. DGA 3

1.1.1. Fast-Flux 4

1.1.2. DNS放大DDoS攻击 4

1.1.3. 钓鱼网站 5

1.1.4. 域名停放 5

1.2. 相关防御手段 5

1.2.1. 信誉系统 5

1.2.2. 逆向工程 6

1.2.3. 机器学习 6

2. 国内外研究现状及趋势 7

3. 论文的研究内容及拟采取的技术方案 9

3.1. 课题目标 9

3.2. 研究内容 9

3.3. 拟采取的技术方案 10

3.3.1. 总体架构 10

3.3.2. 统计学的方法 11

3.3.3. 机器学习 11

4. 关键技术或难点 12

5. 论文研究计划 13

5.1. 进度安排 13

5.2. 预期研究成果 13

6. 主要参考文献 14

**恶意域名检测技术研究**

# 论文选题的背景和意义

IP地址是由IP协议提供的数字型统一地址标识，作为一种逻辑地址来定义一台设备在网络之中的位置，网络设备逐渐增多IP地址的记忆困难显现出来，保罗·莫卡派乔斯 (Paul Mockapetris) 在1983年的第882和在南加州大学里资讯科学研究院所提出的883号因特网标准草案中提出DNS的架构，提议将其改进为分布式和动态的数据库域名系统，也就是我们今天所用的域名系统的雏形。[1]从1985年Symbolics公司注册的第一个.com域名[2]到如今仅中国域名总数增长至3698万个[3]，域名产业飞速发展，随之而来的安全问题也越来越多的暴露出来。Bilge认为恶意域名是涉及恶意活动的域名[20]，我认为这样说是不准确的，不能通过历史上存在的恶意行为就简单认为当前仍然是恶意的，而且恶意行为的受害者显然也不能视为恶意域名，所以我把恶意域名定义为当前存在恶意行为或者被恶意使用的域名。接下来我将从当前相关攻击手段和对抗手段两个层面来论述此课题的意义。

## 相关攻击手段

### 域名阴影

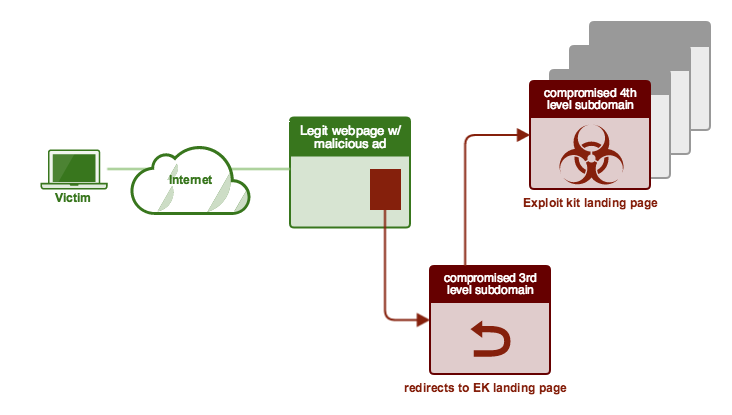
域名阴影（Domain Shadowing）是指利用劫持的合法域名账户在合法域名之下创建大量的子域名。Angler作为当前市场上最强大的钓鱼工具包，于2015年3月份首先推出了这项技术[4]，2016年年初WordPress站点被曝检测到了大量的

图1-1 域名阴影技术原理

恶意广告活动，攻击者使用的正是这项技术[5]。域名阴影技术的原理如图1-1所示，攻击者一般会使用渗透技术（钓鱼网站或者键盘记录木马等）或者社会工程方法获取合法的域名登记者账户，一次性在这些域名下创建大量的子域名来满足恶意的目的。通常这些子域名都进行重定向到攻击者准备好的恶意网站（例如钓鱼网站、广告网站），这些网站充满了挂马与陷阱。而这些危险都隐藏在一个合法的域名之下，导致难以检测和对抗。思科Talo小组在于域名阴影对抗时发现，这些被劫持的子域名被设计为两层，第一层子域多数是一个三级域名，由合法网站页面上的恶意广告跳转而来，然后通过域名的重定向，到第二层子域名。第二层子域名多数为四级域名，这一层页面上部署了扫描和攻击的程序，会自动扫描被攻击者系统的指定漏洞，并执行恶意代码，这一层的子域名数量极大并且有效时间较短以快速更新。

域名阴影技术难以控制在于以下几点：一、注册域名的账户很少被查看，通常域名拥有者只有在第一次注册以及之后添加域名的时候才会查看自己的账户，也就是说基本上一年才会查看一次或者两次，这为攻击者提供了充足的时间来达到目的；二、一次性产生大量的子域名，攻击者不用担心数量问题，因而子域名的生存时间被设计的很短，通常不到一个小时，安全工作者难以得到充分的数据与时间来分析攻击；三、域名阴影对现在的许多安全防护措施是免疫的，例如域名信誉系统，如果这些子域名建立在分数较高的域名之下，就会无从辨别。

### DGA

DGA（Domain generation algorithms）是域名生成算法，被大量的恶意软件家族所使用，用以建立受感染主机与C&C服务器（command and control servers）的联系。

传统僵尸网络使用固定IP或者域名与C&C服务器建立连接，隐蔽性很差，极易被发现。后来出现的例如Nugache，Storm，Waledac，Zeus等基于P2P的僵尸网络虽然具有较好的鲁棒性与稳定性，但实现难度和维护成本较高。如今大部分活跃的僵尸网络都采用了DGA，依赖于集中的C&C服务器，相比于前两者具有简单易行，兼顾稳定性与隐蔽性的优点。今年在网络上大肆传播的locky正是使用了这个算法[12]，造成了极大的危害，通过DGA算法找到域名，获得公钥，然后使用AES或 RSA算法批量加密上百种后缀文件，并且留下勒索信息。

DGA作为一种随机算法，输出为域名，我们将算法输入称为种子（例如数字常量、当前时间、Tiwwter动态等），按照种子产生类型以及算法类型，可以对DGA进行分类。如果这个种子与时间有关，称之为TD（time-dependent，时间相关），反之，称为TI（time-independent，时间无关），如果种子可以估计（例如日期），称为D（Determinism，可估计），反之（例如欧洲央行每天外汇参考利率），称为N（Non-determinism，不可估计）。域名产生模式分别有A（Arithmetic，算法类）、W（Wordlist，单词表类）、H（Hashing，哈希类）、P（Permutation，置换类）四种。进行排列组合理论上有16类DGA，实际上只出现了TDD-A, TID-A,TDD-W , TDD-H, TDN-A , TID-P 这六种类型。[7]

### Fast-Flux

速变服务网络(Fast-Flux Service Network,FFSN)[9]也是一种恶意软件隐藏自己恶意行为的一种常见技术。Fast-Fux不仅在僵尸网络方面有所应用，在被广泛使用于钓鱼网站和垃圾邮件。Fast-Flux为一个域名配置多个IP地址，DNS的解析结果会不断的变化，两次访问同一个域名，可能会得到不同的结果，可以有效的避免IP地址的封锁。利用此特征，僵尸网络控制者可以将其控制的肉鸡中服务能力比较强、具有公有ＩＰ的肉鸡作为代理，其他一些服务能力比较弱的主机通过与这些代理通信，代理将通信重定向到后面真正的控制者，完成控制者与肉鸡的通信过程。僵尸网络控制者通过频繁的更换域名到代理ＩＰ的映射，从代理网络中剔除不可用的以及服务能力较弱的代理，提高网络的健壮性和可用性。[10]

### DNS放大DDoS攻击

DDoS（Distributed denial-of-service）是指分布式拒绝服务攻击，在2015年，DDoS攻击峰值流量不断上升，甚至出现了1T超大流量攻击事件，全年的攻击总流量接近28万Tbytes[6] ，常用手法有UDP攻击、SYN攻击、ICMP包攻击等，这里我主要研究利用DNS的放大DDoS攻击。

在利用流量来实施的DDoS攻击中，只要有可以利用来进行放大数据量的协议，都有可能被恶意使用。正常的DNS查询是从原IP地址向DNS服务器（递归或者权威），大小相对固定（70字节左右），也就是攻击成本相对固定，DNS返回数据由于请求域名和类型的不同，数据大小从几百到几千字节（查询[www.baidu.com.的返回数据为302](http://www.baidu.com.的返回数据为302)字节），放大了四倍以上。由于利用A记录或者Cname记录放大效果并不十分明显，Akamai研究人员在2014年发现了利用TXT记录来进行的DNS放大攻击，攻击者使用名为DNS Flooder的工具，从guessinfosys.com获得TXT记录，攻击峰值高达4.3Gbps[8]。

### 钓鱼网站

钓鱼网站（phishing page）一直以来都是一个难以规避的网络问题，同时他也是一种社会工程学的攻击方式。通常攻击者对恶意页面进行伪装，使得页面看起来像是银行网站或是其他可以获取关键隐私信息的网站，受害者往往由于受到邮件或者短信的欺骗，通过其中的链接导向到钓鱼网站，或者由于误植域名导向到钓鱼网站。而受害者对于网站的真伪缺少足够的判断和辨识的能力，导致银行账户密码等隐私信息被攻击者窃取。

### 域名停放

域名停放（Domain Parking）又叫做域名停靠，是将未开发的域名用于显示除自动生成的广告banner及链接以外，没有任何实质内容的网页，并从中获取经济收入[12]。用户到达Parked Domain的方式通常有以下三种，一是误植域名，目的是访问已知网站，由于输入错误（字母顺序输错、漏输、错输等）进入停放域名；二是猜测域名，按照域名的意思，例如我想找985学校信息，于是我输入了“985.hk”，而事实上该网站使用了sedo的域名停放服务；三是从挂马网站或C&C服务器跳转，有些停靠网站依靠网站上的挂马链接或是僵尸网络产生的多余域名来获取流量。

常见的域名停放服务有SedoParking、InternetTraffic、SilverClicks、RedDns等等。这些服务提供商提供的都是合法的服务，但同时这些网站却存在着大量的恶意行为，例如误植域名或是商标的滥用，用一些错误对浏览网页的人产生误导，再比如一些恶意的重定向，页面上的一些链接会直接重定向到布满木马的网站、色情网站或是准备好扫描脚本直接执行的网站。最后这一类网站通常会获取用户敏感信息或者声称用户计算机处在高危之中，骗取服务费。

## 相关防御手段

### 信誉系统

域名信誉系统（DRS，Domain Reputation System）简要来说就是一个给域名打分的系统，域名的良性与恶性都由分数来决定，我们熟知的黑名单也可以看作是一种最简单的信誉系统，黑名单上的域名的分数统一为0分。

建立黑名单是一种切实有效的防御手段，国内外很多安全厂商和团队都长期维护着黑名单，可以说这是误报率最低同时也是检测速度最快的的一种防御措施。但黑名单也有着不可避免的缺点，将数据信息进行人工或者程序的审核，确认为恶意的数据信息加入黑名单，因此维护和更新的成本相对较高，而且黑名单的建立必然在恶意行为发生之后，相对滞后。

针对网络安全的黑名单通常有域名黑名单和IP地址黑名单，DGArchive[14]、URIBL[15]等都是针对域名的黑名单，更新维护依托于学校、研究机构，国内360、工信部也都有相应的黑名单。DSBL[16]和RBLS[16]则是针对IP的黑名单，由于互联网上广泛使用的地址转换协议难以精确到C类地址，再加上例如Fast-Flux等技术对于IP黑名单有很高的抗性，因此IP黑名单很难做到面面俱到，但对使用固定IP频繁更换域名的攻击者来说，这个黑名单至关重要。

### 逆向工程

逆向工程对恶意软件防御方面而言是一种很常规的手段，今年Khaled Yakdan、Daniel Plohmann等人的论文中正是对现在活跃的大部分DGA使用了逆向工程[14]，在实际使用中可以做到FP值为0，他们历时数年的细致工作完美体现了逆向工程精准的特质，但是逆向工程的缺点也暴露无遗，那就是太耗费人力和时间，一旦出现可能是新的DGA算法，就需要专业的逆向工程师来处理、验证，这对于如今层出不穷的新型木马而言是远远不够的。

### 机器学习

信息安全领域仍然是一个富数据，穷分析的领域。2016乌云白帽大会上，phunter做了报告《What can you get from 100 billion DNS queries, each day, in real time?》，可以想见当我们有1000亿条实时数据，如果仍然使用蜜罐、逆向工程，周期长度以及工程量都是无法承受的。如何做到又快又精确地对给定域名进行分析，那么机器学习就是不二之选。特征的选择体现了人类的经验，而把这些量化特征交给机器，相当于机器利用人类的经验来完成这些重复性的工作。

如今机器学习在域名安全领域的利用大概都有如下的流程，首先收集数据并进行过滤，然后进行特征工程，往往特征的好坏就决定了整个系统的优劣，之后是模型选择与训练，最后是评价函数和交叉验证，可以用精度和召回率评价模型的分类效果。

# 国内外研究现状及趋势

被动DNS（PassiveDNS）数据是网络安全领域最为常用的资源之一，该数据收集技术由Florian Weimer于2004年提出[18]，主要目的是将DNS流量转换为易于访问的格式。递归域名服务器会响应其接收到的来自其它域名服务器的请求，对响应进行记录并将日志数据复制到中央数据库当中，因此记录这些数据就掌握了DNS解析的历史动态。

2008年Plonka等人提出了Treetop[19]，这是早期较为经典的一个系统，它可以弹性地管理逐渐增长的被动DNS数据，并在同时关联域和网络属性。他们的聚簇区域是基于不同的种类的网络。Treetop依据是否符合各种DNS RFC标准以及解析结果来区分DNS流量并提供部分安全信息。

佐治亚理工学院的M.Antonakakis、Perdisci.Roberto等人也在那个时期开始对DNS的相关研究。在2010年他们提出了Notos[22]，是一个动态的综合性的信誉系统，首次提出了针对域名的信誉系统，利用被动DNS数据来输出信誉分数，而于此之前主要针对IP的信誉系统完全不同。Notos使用了网络和地区的特征，通过配置信息、使用情况和域名管理情况等，能够学习良性和恶性的域名分别是如何工作的，并对每一个新域名计算一个信誉得分。2011年Manos Antonakakis等人构建了一个叫做Kopis的恶意域名监测系统[25]，可以通过对DNS查询解析模式的分析得到恶意域名，与EXPOSURE和Notos这些依赖本地递归DNS服务器的系统相比，Kopis使用的是上级DNS数据。在此基础上Manos Antonakakis等人在2012年又提出了一个新颖的检测系统Pleiades[23]，首先使用了不存在域名（NXDOMAIN）响应。NXDOMAIN是指像域名服务器提出解析请求，但是无法解析得到对应IP。对于使用DGA的被控主机大多数查询都会返回一个NXDomain，而同一僵尸网络控制下(使用相同的DGA算法)的被控主机会产生有类似特征的NXDomain的流量。Pleiades聚类和分类算法的结合，充分利用了机器学习的优势，将有着相似字符特征和相似访问特征的域名聚集在一起。分类算法用来将这些生成好的聚簇分配到已知的DGA模型中去。如果一个聚簇不能被分到已知的DGA中去，那么意味着可能出现了新的DGA变种或DGA家族。

2011年L.Bilge等人建立了EXPOSURE系统[20]，一个检测恶意域名的系统。该系统从DNS数据中分别基于时间、基于DNS响应、基于生存时间值（TTL）、基于域名提取了这4类共计15种特征，使用J48决策树训练分类器。相比于Notos，Exposure有着更完善的特征选取，也弥补了无法检测一个IP地址只被恶意使用一次的恶意域名。相比于之前文章大多数只能在DNS数据中找到特定种类的恶意域名，例如Perdisci 等提出 FluxBuster 系统[21]， 将域名的 IP 变迁情况引入特征集，共 9 组 13 个特 征，采用聚类算法来识别速变域名，EXPOSURE对恶意域名的检测更加广泛。

2013年Terry Nelms，Roberto Perdisci等人构建的ExecScent系统是一个旨在从真实的企业网络流量中挖掘新的、从未出现过的C&C域名的系，同时ExecScent也是第一个使用自适应C&C流量模型的系统。ExecScent从已知的C&C通信样本中自动地学习控制协议模板(CPT)，并且这些CPT会匹配它们所部署的网络流量。ExecScent构建的这种自适应模板从部署模板的网络流量中学习，这种“自适应”的方法使得ExecScent在保持一个很高检出率的同时极大地降低了误报率。

2016年B.Rahbarinia创新性的提出了一种基于行为的系统Segugio，segugio在大的ISP网络中通过追踪被恶意软件感染主机的DNS请求行为来高效的发现新增的malware-control域名。相比于Notos和EXPOSURE，前两者建立的都是domain-IP映射关系模型（使用域名字符串的特征、域名承载的恶意内容等信息）而没有利用本地DNS服务器下游的主机请求行为，Segugio通过监测ISP网络用户的DNS请求行为，重在精确的追踪新增的“malware-only”域名。相比于Kopis，他的做法和本文有相似性（Kopis用请求者散度、请求者画像等信息），但Kopis利用权威或TLD服务器的数据，这种数据难以获得（需要与大的DNS区域运营商紧密合作）。Segugio不用关心顶级域名，通过监测本地ISP流量（在ISP使用者和他们的本地DNS解析器之间的DNS流量）。因此Segugio可以依靠ISP网络权限独立开发，不需要与外部的DNS运营商合作。

D.Plohmann，F.Fkie等人针对DGA做了大量细致的工作[14]，他们对43种DGA恶意软件家族和变种进行一个综合性的研究，针对DGA提出一个分类学方法，并用它对所研究的DGA进行分类与比较。并重现了这些算法，预先计算所有可能的AGD，覆盖了大部分已知的活跃DGA，以过去8年总计18millionDGA域名的注册状态，来证实预计算得到的域名确实是是可靠的。对于botmaster的域名注册策略也提出了见解。

# 论文的研究内容及拟采取的技术方案

## 课题目标

针对现如今恶意利用DNS和域名的问题，研究一种高效准确的鉴别技术，并实现一个恶意域名检测系统，该系统实现以下功能：

1. 针对域名的评价系统，维护一个具有中国特色的域名排名；
2. 能够监测DNS流量中的异常；
3. 检测恶意域名，并对攻击类型进行分类。

## 研究内容

如今针对网站的排名一方面是例如百度、Google一类以使用PageRank为主的排名，另一方面是Alexa一类的利用访问行为统计的排名[26]。无论是哪一种都存在着难以回避的问题，对于第一类，大量黑帽SEO会使用诸如关键字堆砌、地址重定向、域名轰炸、虚假关键字、桥页、门页、黑链来作弊，每年百度都要被动对SEO策略升级上百次[27]，也没办法提供一个综合性的网站排名。对于第二类，Alexa数据来源是Alexa工具条，如果没有安装则无法统计，而且仅仅利用了连接数和访问时间，一方面很容易被黑帽利用刷排名，另一方面小众网站可能因为用户没有使用Alexa工具条而无法进入排名，在大多数情况看来几万或者几十万排名或者不在排名中都没有区别，但事实上Alexa的Top十万、Top百万是安全领域对域名数据预处理用作白名单的必然之选，而且由于中国大量地方性网站都使用汉语拼音的缩写作为网站域名（例如北京市小客车指标管理信息系统，www.bjhjyd.gov.cn），汉语拼音与英语所属的印欧语系有着本质上的区别，当利用字符特征来进行域名处理时，如果没有中国特色的白名单或者语料库，而是直接使用Alexa Top百万会造成较高的误报。我想做的是以省为单位维护一个地方的域名评价系统，评分结果一方面可以用于构建具有地方特色的白名单，另一方面可以将计算好的历史数据作为域名安全领域最为贴合国内环境的域名访问特征。

在此基础之上，可以利用CNCERT的数据进行分析，例如对DNS流量进行实时分析，发现其中的异常，通过对DNS异常的分析可以发现针对DNS的网络攻击或者对潜在网络威胁产生预警。比如说发生了利用DNS的DDoS放大攻击，攻击者通常会利用A记录或者TXT记录，那么这些记录的激增都在可观察的异常范围内，再比如出现了大量解析带有相同字符特征的请求，并且解析rcode为3时，那么很有可能出现了一个大型僵尸网络。当DNS请求发生异常时，无法确认说一定发生了攻击，但是有针对性的将异常记录下来并进行分析绝对是必要的。

无论是对域名打分或者是对DNS异常进行分析，最终的目的都是为了找到并识别恶意域名。如今的网络环境中，我所理解的安全并不仅仅是保障一台主机或者网络设备的安全，而是全网的安全。我所要做的就是正是如此，找到一个C&C域名，可能破坏了一个僵尸网络；找到一个钓鱼域名，可能避免很多人蒙受损失；找到一个恶意子域名，可能就处理了一次潜在的攻击。当整个互联网环境都安全的时候，个人设备的安全自然不必担心。

## 拟采取的技术方案

恶意域名检测的过程就是一个数据挖掘的过程，对恶意域名的挖掘就是从众多数据中找到最有价值的内容，具体使用的技术方案，无外乎有这两类，统计学方法和机器学习方法，下面对总体思路进行介绍。

### 总体架构

图3-1 总体架构图

总体架构如图3-1所示。其中原始数据的收集主要包括被动DNS数据和Whois数据收集，被动DNS数据来源各省递归DNS服务器，Whois数据主要来源自网络上的爬取和数据交换。利用被动DNS和已有黑名单实时地对DNS异常进行检测，发现活跃的异常数据，并进行分析。同时原始数据经过简单过滤、提取存储起来。域名打分系统利用历史数据进一步加工，例如提取一个时间片段内的域名访问量、访问IP数、网络数量等，按照一定规则进行计分，按照结果来更新黑白名单。最后综合利用历史数据、域名打分结果以及黑白名单进行操作，实现恶意域名准确检测。

### 统计学的方法

现如今任何一个领域都在谈论大数据，那是因为现在数据的收集、存储变得简单起来，就数量而言是很大的，也意味着除了我们想得到信息以外存在着大量的噪声，而利用统计学的方法可以帮助我们很好的处理以一些问题。具体举例来说，当拿到原始数据之后，需要做一些初步的过滤，我们对在一个固定时间段访问某一域名的IP数计数，低于或者高于某一阈值我们认定为合理的，把他们过滤掉，留下可疑的进行下一步研究，这是一种统计方法；我们对访问规律进行研究，尝试回归的模型，对聚类结果进行处理，做一些主成分分析，这些也是统计方法。

### 机器学习

机器学习相比于统计学的方法更具有实验性质，学习的效果要通过统计学的方法来对比和提升，同时对统计量的选取提供指导，两种方法相互配合才能更好的挖掘所需信息。大部分情况下，机器学习的是人类的经验，按照人类制定的规则，完成一系列重复性的工作，比如监督学习和无监督学习。机器学习的方法在信息安全领域的已经有了广泛的应用，但是值得注意的是，随着2012年Hinton等人利用卷积神经网络在ImageNet竞赛中夺冠[29]，神经网络或者说是深度学习成为了一个十分热门的话题，在计算机视觉领域发展迅速，但是安全领域的应用还几乎是空白。今年乌云大会上，域名安全领域最为成功的公司Nominum所做报告中，已经将神经网络应用到了域名检测实时系统之中[28]，这也是我之后实验的一个重点。

# 关键技术或难点

本课题存在以下几个课题难点：

1、地区化差异带来的影响

在前期工作之中，尝试直接将已有的方法复用到从山西省收集得到的PassiveDNS数据集上，实际结果与论文效果差距很大，而论文中的系统普遍布置在众多欧美地区ISP之下，皆取得不错的结果，因此过拟合的问题在印欧语系下应该不存在，我思考可能的原因在于中国域名的独特性，或者说是汉语的独特性。而且相比于欧美国家恶意域名的泛滥，新型恶意域名生成算法与使用方法层出不穷，国内黑客显得十分简单粗暴，2008年就已经被发现的Conficker仍然是国内最主要的蠕虫病毒，这一点也是国内外现状的不同。即使在国内范围，由于地区经济与科技的发展的不同步，各类攻击的分布也与地理位置有着极大关系，例如360天眼实验室公布的《OceanLotus（海莲花）APT报告》中提到，感染者地域分布仅北京一地就占22.7%[30]。

2、域名打分系统

我希望利用DNS数据以及Whois信息对域名打分能真实的反映访问情况，访问量大可能是良性域名，但是也要界定黑帽SEO或者僵尸网络的行为；访问有周期性可能是恶意代码请求，但是也要考虑一些特需网站要求定期访问。最理想的状态就是直接通过分数来评定网站的优劣，由于需要考量的方面和具体情况太过于复杂，难以准确代表域名的特性。我的备选方案是划分小的时间段，每个时间段单独计分，通过分数的变化来定义异常的域名。

3、深度学习在恶意域名检测方面的应用

至今所阅读的论文之中尚没有将深度学习应用在恶意域名的检测之中的实例，如何合理的应用是之后学习的关键和难点。

# 论文研究计划

## 进度安排

* 2016.11－2016.12 调查背景和研究现状，撰写文献综述和开题报告
* 2017.01－2017.03 实现针对汉语特征的域名打分系统
* 2017.03－2017.04 整理方法和数据写成技术文档
* 2017.04－2017.06 编码实现，撰写小论文
* 2017.07－2017.09 实现域名检测系统，进行试验并完善技术
* 2017.10－2017.12 整理资料，撰写毕业论文

## 预期研究成果

* 技术文档： 2篇
* 工具：恶意域名检测系统原型
* 小论文：2篇以上
* 申请发明专利：2个

# 主要参考文献

1. https://zh.wikipedia.org/wiki/%E4%BF%9D%E7%BD%97%C2%B7%E8%8E%AB%E5%8D%A1%E6%B4%BE%E4%B9%94%E6%96%AF
2. 25 Years Later, First Registered Domain Name Changes Hands,https://techcrunch.com/2009/08/27/25-years-later-first-registered-domain-name-changes-hands/
3. 第38次中国互联网络发展状况统计报告. CNNIC，2016.
4. Nick Biasini.Threat Spotlight: Angler Lurking in the Domain Shadows. 2015.
5. Shadow Puppets – Domain Shadowing 101. http://defintel.com/blog/index.php/tag/domain-shadowing. 2016.
6. 2015DDoS威胁报告.绿盟科技.2015
7. Plohmann Daniel,Fkie Fraunhofer,Yakdan Khaled,Klatt, Michael.A Comprehensive Measurement Study of Domain Generating Malware, USENIX Security Symposium,2016
8. Attackers are increasingly leveraging large Domain Name System (DNS) TXT records in an effort to amplify the impact of their distributed denial-of-service (DDoS) attacks, Akamai's Prolexic Security Engineering and Research Team (PLXsert) warned on Tuesday,http://www.securityweek.com/large-dns-text-records-used-amplify-ddos-attacks-akamai
9. Riden J. Know your enemy: fast-flux service net- works [EB/OL]. 2008.http://www.honeynet.org/papers/ff
10. T. Holz, C. Gorecki, K. Rieck, and F. Freiling. Measuring and detecting fast-flux service networks. In Proceed- ings of NDSS, 2008.
11. 康乐，李东，余翔湛，基于SVM的Fast-flux僵尸网络检测技术研究，《智能计算机与应用》, 2011, 01(3):24-27
12. Vissers, T., Joosen, W., & Nikiforakis, N. (2015). Parking Sensors-PPT. Proceedings 2015 Network and Distributed System Security Symposium, (February), 8–11.
13. 中国网，勒索者病毒“Locky”每小时感染德国5300台计算机，2012.
14. Plohmann, D., Fkie, F., Yakdan, K., Klatt, M., Bader, J., Gerhards-padilla, E., … Bader, J. (2016). A Comprehensive Measurement Study of Domain Generating Malware. USENIX Security Symposium.
15. URIBL. Real time URI blacklist. http://uribl. com.
16. DSBL. http://www.dsbl.org/.
17. RBLS. one-stop source of checking ip-numbers and domain names in multiple blocklists. http://rbls.org/.
18. F. Weimer. Passive DNS replication. In Proceedings of FIRST Conference on Computer Security Incident, Hand ling, Singapore, 2005.
19. D. Plonka and P. Barford. Context-aware clustering of DNS query traffic. In Proceedings of the 8th IMC, Vou- liagmeni, Greece, 2008. ACM.
20. Bilge, L., Kirda, E., Kruegel, C., Balduzzi, M., & Antipolis, S. (2011). EXPOSURE : Finding Malicious Domains Using Passive DNS Analysis. Ndss, 1–17.
21. Perdisci R, Corona I, Giacinto G. Early detection of malicious flux networks via large-scale passive DNS traffic analysis. IEEE Transactions on Dependable and Secure Computing, 2012, 9(5): 714–726
22. Antonakakis, M., Perdisci, R., Dagon, D., Lee, W., & Feamster, N. (2010). Building a Dynamic Reputation System for DNS. USENIX Security’10: Proceedings of the 19th USENIX Conference on Security, 1–17.
23. Antonakakis, M., & Perdisci, R. (2012). From throw-away traffic to bots: detecting the rise of DGA-based malware. Proceedings of the 21st USENIX Security Symposium, 16.
24. Rahbarinia, B. (2016). Segugio:Efficient and Accurate Behavior-Based Tracking of Malware-Control Domains in Large ISP Networks, 19(2).
25. Antonakakis, M., Perdisci, R., Lee, W., Ii, N. V., & Dagon, D. (2011). Kopis:Detecting Malware Domains at the Upper DNS Hierarchy. USENIX Security Symposium., 11, 1–16.

1. <https://en.wikipedia.org/wiki/Alexa>
2. 中国互联网地下产业链分析白皮书.2016.
3. What can you get from 100 billion DNS queries , each day , in real time ? (n.d.).2016.
4. Krizhevsky, A., Sutskever, I., & Geoffrey E., H. (2012). Imagenet. Advances in Neural Information Processing Systems 25 (NIPS2012), 1–9.
5. 天眼实验室：OceanLotus（海莲花）APT报告.2015.