**北京航空航天大学计算机学院**

**硕士学位论文中期检查报告**

**论文题目**：恶意域名检测技术研究

**专 业**：计算机技术

**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**

2017 年 08 月 16 日

目 录

1 论文工作计划 1

1.1 论文研究背景 1

1.2 论文研究目标 2

1.3 论文研究进度 2

2 已经完成的工作 2

2.1 可疑域名提取 2

2.2 流量监控 4

2.3 恶意域名检测研究 4

3 关键技术或难点 6

3.1 关键技术或难点一 6

3.2 关键技术或难点二 6

4 下一阶段工作计划 6

4.1 存在的问题 6

4.2 尚未完成的工作 6

4.3 解决问题的技术思路或措施 7

4.4 下一阶段计划 7

5 主要参考文献 7

**恶意域名检测技术研究**

# 论文工作计划

## 论文研究背景

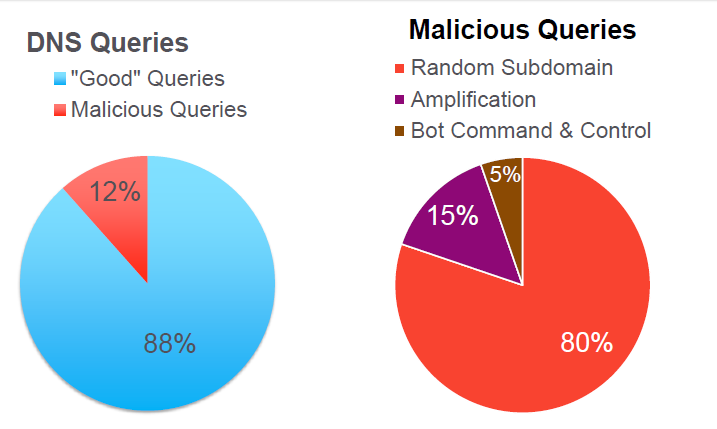
IP地址是由IP协议提供的数字型统一地址标识，作为一种逻辑地址来定义一台设备在网络之中的位置，网络设备逐渐增多IP地址的记忆困难显现出来，保罗·莫卡派乔斯 (Paul Mockapetris) 在1983年的第882和在南加州大学里资讯科学研究院所提出的883号因特网标准草案中提出DNS的架构，提议将其改进为分布式和动态的数据库域名系统，也就是我们今天所用的域名系统的雏形。从1985年Symbolics公司注册的第一个.com域名到如今仅中国域名总数增长至3698万个，域名产业飞速发展，随之而来的安全问题也越来越多的暴露出来。Bilge认为恶意域名是涉及恶意活动的域名，我认为这样说是不准确的，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。

图1-1 2015年某日流量分布

如图1-1所示[5]，恶意的DNS查询约占12%，在这些恶意查询中，主要有三个来源。一是随机子域名攻击的流量；二是放大攻击的流量；三是僵尸网络的C&C流量。我的研究也基于此，树立了三个研究对象，对于这三个研究对象又分别构建了针对性的阶段检测手段。接下来我将对已完成工作进行详实的叙述，并阐明下一个阶段的工作存在的问题。

## 论文研究目标

对于现如今恶意利用DNS和域名的问题，本研究课题主要针对随机子域名攻击、DDOS攻击和DGA所涉及到的域名进行研究，并实现相应原型系统，该系统实现以下功能：

1. 流量监控子系统，对各种请求类型的流量进行监控；
2. 可疑域名提取子系统，分别提取三类可疑域名及其流量；
3. 域名检测子系统，对可疑域名进行检测，找到恶意域名。

## 论文研究进度

表1-1 论文研究进度表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **工作内容** | | **已完成** | **进行中** | **未展开** |
| 系统总体框架设计 | | √ |  |  |
| 基础研究 | 域名与DNS系统基础知识研究 | √ |  |  |
| 各类攻击行为研究 | √ |  |  |
| 算法设计 | 可疑域名提取可行性研究 | √ |  |  |
| 恶意域名检测技术研究 | √ |  |  |
| 机器学习算法及其应用研究 |  | √ |  |
| 系统实现 | 流量监控子系统实现 | √ |  |  |
| 可疑域名提取子系统实现 | √ |  |  |
| 域名检测子系统实现 |  | √ |  |

# 已经完成的工作

## 可疑域名提取

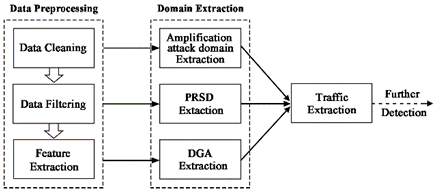
在本节中，我将介绍可疑域名提取部分的工作情况。如图一所示，我们将系统分为两个部分：被动DNS预处理模块和流量提取模块。我们将对每个模块功能进行表述，并讨论如何共同达到目的并且尽可能提高召回率与效率。

图2-1 可疑域名提取系统总览

首先预处理的过程主要达到三个目标，一是剔除包含错误的域名的DNS记录，即进行数据清洗，例如请求的域名中出现了‘？’，‘！’等符号；二是剔除无关流量，一方面将良性的流量和恶意的流量中都不会出现的流量剔除，例如逆向解析流量。另一方面我们在针对某一类型流量提取过程中，也会尝试剔除不会出现在恶意流量中的域名，例如DGA流量的提取可以剔除白名单中的域名。三是为流量提取模块计算特征值，例如每个域名的ANY查询的比例等。

域名提取的部分，其主要目的是要将放大攻击中作为跳板的那部分域名找出来。攻击者想要利用这些域名，必然是利用其TXT记录或者ANY查询返回该域名所有资源记录。我们获取ANY类型请求比例qar和TXT类型请求比例qtr，带入公式1中，β为我们设定的一个参数，当qar+qtr<=β时，结果为0，当qar+qtr>β时qar+qtr与s1成正相关，同时设定阈值α，其中s1>α，我们认定为疑似放大攻击的流量。

对于随机子域名攻击我使用类似的方法，攻击者会在二级域名下伪随机的生成大量子域名，这些域名都是不存在的。因此我们用sdc与nxdr相乘来表示被恶意使用的可能性大小，而这个值范围较大，我们使用公式2将结果变到0到1之间。

对于僵尸网络的C&C域名，我主要考虑DGA所产生的域名。这里的模型使用黑白名单训练，分类器我们选择随机森林，训练每棵树时，从全部训练样本（样本数为N）中选取一个可能有重复的大小同样为N的数据集进行训练（即bootstrap取样）。随机森林的训练速度快，并且可以平衡误差。

## 流量监控

图片包含 电子产品, 陈列

已生成极高可信度的说明在获取了原始的被动DNS之后，需要一个切实有效的可视化方案。这里我选择使用Influxdb + Grafana构建一个流量监控系统。Influxdb是一个开源的分布式时序、时间和指标数据库，使用go语言编写，无需外部依赖。它有三大特性：一是时间序列，因此与时间相关的函数使用非常的灵活（诸如最大、最小、求和等）；二是度量，可以对实时大量数据进行计算；三是事件，它支持任意的事件数据。同时Influxdb还具有无结构（无模式）、扩展性强、支持大量统计函数、原生的HTTP支持、强大的类SQL语法等一系列优点。我选择Influxdb作为数据库也正是基于此。

图2-2 流量监控系统

Grafana 是一个开箱即用的可视化工具，具有功能齐全的度量仪表盘和图形编辑器，有灵活丰富的图形化选项，可以混合多种风格，支持多个数据源特点。

## 恶意域名检测研究

首先是基于特征的机器学习检测方法，这里特征的选取分为两类，一是字符特征，例如n-gram特征，基于熵的特征，域名结构的特征。在计算时通常将连续的DNS记录分割成一个固定长度的子序列，这样的一个子序列在计算这些字符特征后，又会产生中位数、平均数、方差等一系列特征，选取合适的聚类算法。另一部分是利用访问关系来检测，即利用一个二部图来检测。感染了同一种DGA的恶意软件的主机有着很大的可能生成互相重叠的NXDomains集合。另一方面，其他“非DGA”的NXDomains则不会被多个主机查询。例如，在一段时间内多个用户同时犯同样的拼写错误是不太可能的。这促使我们考虑那些被一些相同的主机查询的NXDomains的相似性，而通过利用这些相似性我们可以将可能由同种DGA产生的NXDomains聚类起来，算法如下。

输入：稀疏矩阵,行代表*l*个主机，列代表*k*个NXDomains

[1]：归一化M：

[2]：依据M计算相似矩阵S=

[3]：通过特征值分解依据S计算个特征向量

令为从S特征值分解得来的包含k个维度为的向量的矩阵(向量代表第i个NXDomain压缩成维后的向量)

[4]：使用X-means算法聚类向量(也就是NXDomains)

输出：NXDomains簇

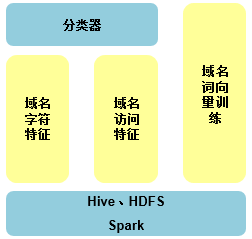
相比传统检测方法,基于深度学习的域名相关性分析不仅简单高效,而且使用范围更广,在未来网络安全中必将发挥着重要作用。如图所示，以数据流图的形式，详细描述了本系统中数据的处理过程与流向。原始DNS流量数据从DNS服务器获得，首先，格式化解析成为可识别的DNS数据；其次，经过一系列的数据预处理，也就是数据清洗，获得符合训练要求的数据；然后，再组织成可供深度学习训练的域名语料(以域名为词的句子)，这些语料经过训练会形成关联模型保存下来；之后，通过对关联模型的一系列分析处理（诸如：关联分析，向量降维，域名聚类，以及可视化等），最终将实验结果呈现给安全人员进一步分析研究。这部分我们采用word2vec算法，它是由google推出的用于词向量计算的算法。该方法可以针对当前的词来计算前后文词语出现的条件概率，或者由前后文的词语计算当前的词出现的条件概率，进而建立有效的预测模型，使所有词语的条件概率乘积取最大值。通过取对数的方法简化运算,转乘积为求和,并采用Negative Sampling 提高计算速度,使用随机梯度下降的方法来更新概率, 在不断更新概率与建立模型的过程中,将目标数据集转换为高维空间向量。对于这部分向量我们有多种处理方式：一是进行降维，实现可视化便于直接分析；二是直接计算两个向量之间的相关性；三是与传统检测方法结合，将向量视为一类特征输入到机器学习的模型之中。

图2-3 域名检测子系统

# 关键技术或难点

## 流量监控系统的展示

原始数据位于hive上，通过构建定时脚本将hdfs上的数据导入指定服务器上的Influxdb中，Grafana将同步更新。这一部分的关键在于数据库和展示平台的选择与搭建。

## 可疑域名提取阈值计算

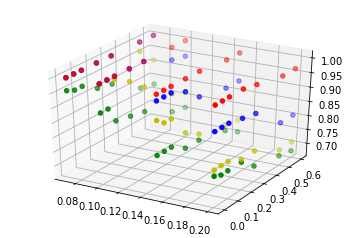
这一部分中出现了大量的阈值和参数，首先是公式2-1中涉及到的三个，分别是时间间隔、判断域名是否可疑的阈值α、公式2-1中的参数β。其中时间间隔取值为10、30、60、120，α取值为0、0.05、0.1、0.2、0.3、0.5、0.6，β的取值为0.07、0.1、0.15、0.2，如图所示，z轴代表召回率，当时间间隔为10时，召回率始终在0.98以上，当时间间隔为120时，召回率最高为0.94，都无法满意的对α与β取值。图3-1展示了在不同参数的取值与可疑域名数量之间的关系。可以看到时间间隔选为30和时间间隔为60所能达到的效果差别不大，而前者的执行次数为后者的两倍，因此我们将时间间隔定为60分钟，为了获得尽可能小的域名数量，我们将α设定为0.1，β设定为0.05。

图3-1 不同阈值下的召回率

另一个需要讨论的参数是2.1节中公式2-2中的θ，θ的值越小，这个函数的曲线越平滑。如图3-2所示，描述了θ的取值和提取出来的域名的数量之间的关系。其中当θ为0.1时，召回率为50%。当θ取值大于0.2时，召回率达到100%。从图中可以看出θ的取值域名数量保持正相关，为避免过拟合，我们将θ取值为0.3。

图3-2 流量监控系统

最后讨论一下我们在DGA域名的处理过程中的特征选择，一方面我们不能选择过多的特征，以保证模型的训练时间和分类时所花费的时间不能过长，同时也不能选择过少的特征，不然难以达到满意的召回率。最后我们选取了二级域名的长度、熵值以及bigram、trigram、fourgram，这五个特征。使用随机森林作为分类器。如图7所示，我们通过混淆矩阵来观察结果，召回率达到96以上，并且进行了十折交叉验证，召回率都在85%以上，最高可以达到98%以上。

## 恶意域名检测

现阶段研究主要集中在技术的应用上，因此关键技术主要集中在算法的实现之上。我对域名根据以下两个相似准则分别被聚类：(1)域名字符串所具有的相似的统计特征(例如相似的长度、随机性和有关字符的频率分布等)；(2)域名被一组相同或相近的IP访问。对NXDomain聚类的主要目的是将那些使用同一种域名生成算法的主机和域名聚类出来。理所当然地，由于这个聚类过程是无监督的，所以聚类出来的某些NXDomain簇里可能包含那些偶然错误的域名(例如由于拼写或者配置错误导致的NXDomains)。因此，我们需要对这些数据进行过滤。我们使用有监督的DGA分类器去修整这些聚类簇，修正的对象包括我们已经发现并建模的DGA产生的域名和那些与合法的域名相近的域名。DGA发现模块最终的输出是NXDomains的聚簇集，其中的每一个集合都代表了已知或者未知DGA产生的域名。

为了进一步优化模型，因此引入了word2vec。选择每个IP所访问的域名序列作为一句话来构建语料库，使用skip-gram或者cbow模式训练出词向量，这个词向量可以直接利用其相关性进行检测或者作为特征使用。

# 下一阶段工作计划

## 存在的问题

1. 通过选择更加合理的特征提高恶意域名检测效果，或者制定特征评价策略；
2. 尝试多种模型来完善实验；

## 后续进度安排

|  |  |
| --- | --- |
| **时间** | **任务安排** |
| **2017年9月—2017年10月** | 检测算法的完善，效果测试 |
| **2017年10月—2017年11月** | 完成系统成果的展示部分 |
| **2017年11月—2017年11月** | 对数据、材料进行整理 |
| **2017年12月—2018年12月** | 撰写毕业论文，毕业答辩。 |

# 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. Bilge, L., Kirda, E., Kruegel, C., Balduzzi, M., & Antipolis, S. (2011). EXPOSURE : Finding Malicious Domains Using Passive DNS Analysis. Ndss, 1–17.
5. Thomas Orthbandt. (2015). Nominum. <https://nominum.com/press_item/sharp-rise-in-dns-based-ddos-last-year-signals-larger-more-frequent-attacks-in-2015/>. DNS-based DDoS rise signals more attacks in 2015
6. Shadow Puppets – Domain Shadowing 101. http://defintel.com/blog/index.php/tag/domain-shadowing. 2016.
7. 2015DDoS威胁报告.绿盟科技.2015
8. Plohmann Daniel,Fkie Fraunhofer,Yakdan Khaled,Klatt, Michael.A Comprehensive Measurement Study of Domain Generating Malware, USENIX Security Symposium,2016
9. 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
10. Riden J. Know your enemy: fast-flux service net- works [EB/OL]. 2008.http://www.honeynet.org/papers/ff
11. T. Holz, C. Gorecki, K. Rieck, and F. Freiling. Measuring and detecting fast-flux service networks. In Proceed- ings of NDSS, 2008.
12. 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
13. Vissers, T., Joosen, W., & Nikiforakis, N. (2015). Parking Sensors-PPT. Proceedings 2015 Network and Distributed System Security Symposium, (February), 8–11.
14. 中国网，勒索者病毒“Locky”每小时感染德国5300台计算机，2012.
15. 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.
16. URIBL. Real time URI blacklist. http://uribl. com.
17. DSBL. http://www.dsbl.org/.
18. RBLS. one-stop source of checking ip-numbers and domain names in multiple blocklists. http://rbls.org/.
19. F. Weimer. Passive DNS replication. In Proceedings of FIRST Conference on Computer Security Incident, Hand ling, Singapore, 2005.
20. D. Plonka and P. Barford. Context-aware clustering of DNS query traffic. In Proceedings of the 8th IMC, Vou- liagmeni, Greece, 2008. ACM.
21. Bilge, L., Kirda, E., Kruegel, C., Balduzzi, M., & Antipolis, S. (2011). EXPOSURE : Finding Malicious Domains Using Passive DNS Analysis. Ndss, 1–17.
22. 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
23. 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.
24. 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.
25. Rahbarinia, B. (2016). Segugio:Efficient and Accurate Behavior-Based Tracking of Malware-Control Domains in Large ISP Networks, 19(2).
26. 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.