计算机研究与发展
Journal of Computer Research and Development
ISSN 1000-1239,CN 11-1777/TP

《计算机研究与发展》网络首发论文

题目：域名滥用行为检测技术综述
作者：樊昭杉，王青，刘俊荣，崔泽林，刘玉岭，刘松
收稿日期：2021-02-05
网络首发日期：2022-02-11
引用格式：樊昭杉，王青，刘俊荣，崔泽林，刘玉岭，刘松．域名滥用行为检测技术综述[J/OL]．计算机研究与发展.
https://kns.cnki.net/kcms/detail/11.1777.TP.20220211.1037.002.html

计 算 机 研 究 与 发 展
Journal of Computer Research and Development
DOI:10.7544/issn1000-1239.20210121

域名滥用行为检测技术综述

樊昭杉[1,2] 王青[1,2] 刘俊荣[1] 崔泽林[1] 刘玉岭[1,2] 刘松[1]

[1]（中国科学院信息工程研究所 北京 100093）

[2]（中国科学院大学网络空间安全学院 北京 100049）

（fanzhaoshan@iie.ac.cn）

Survey on Domain Name Abuse Detection Technology

Fan Zhaoshan[1,2], Wang Qing[1,2], Liu Junrong[1], Cui Zelin[1], Liu Yuling[1,2], and Liu Song[1]

[1]（*Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100093）

[2]（*School of Cyber Security, University of Chinese Academy of Sciences, Beijing* 100049）

Abstract Domain name system is one of the most critical components of the global internet infrastructure in the network and information age. But it is also being abused by various types of cyber attacks, such as botnet command and control, spam delivery, and phishing, which are emerging as the most serious threat against cyber-security. This paper comprehensively reviews the existing domain name abuse detection technologies from the perspective of typical detection scenarios. First, the background knowledge of domain name abuse detection is introduced. By investigating the existing domain name abuse detection schemes, a taxonomy of detection scenarios is put forward. Moreover, the typical features and detection methods are also summarized. Second, the evolution process of attack and defense technologies for domain name abuse in five typical detection scenarios, including malware, phishing, cybersquatting, spam, and unrestricted abuse behavior, are respectively elaborated. Furthermore, an comprehensive summary of domain name abuse detection methods is given from multiple dimensions such as technical solutions, typical features, and detection algorithms. And a systematic overview of existing domain name abuse detection methods is conducted. Finally, the challenges faced by domain name abuse detection technology and future research directions are discussed, with a view to further improving the ecological environment of domain name system.

Key words domain name system; domain name abuse; malware; phishing; cybersquatting; spam

摘要 域名系统是网络和信息时代互联网基础结构的重要组成部分，同时也被多种严重威胁网络安全的攻击活动滥用，例如僵尸网络命令和控制、垃圾邮件分发以及网络钓鱼. 从典型检测场景的角度，全面回顾现有的域名滥用检测技术. 首先，介绍域名滥用行为检测的背景知识，并通过调研现有域名滥用检测方案，提出域名滥用检测场景分类体系、总结典型检测特征及方法. 其次，分别阐述了恶意软件、网络钓鱼、域名抢注、垃圾邮件，以及不限定滥用行为5种典型检测场景下，域名滥用攻防技术演进的过程. 并从技术方案、典型特征、检测算法等多个维度进一步全面梳理域名滥用检测工作，对现有的域名滥用检测方法进行系统概述. 最后，讨论域名滥用检测技术面临的挑战和未来研究方向，以期改善域名系统的生态环境.

关键词 域名系统；域名滥用；恶意软件；网络钓鱼；域名抢注；垃圾邮件

收稿日期：2021-02-05；**修回日期**：2021-12-06

基金项目：国家重点研发计划项目（2021YFF0307203, 2019QY1300, 2018YFB0803602）；中国科学院青年创新促进会项目（2021156）；中国科学院战略性先导科技专项（C类）（XDC02040100）；国家自然科学青年科学基金项目（61802404）；中国科学院网络评估技术重点实验室资助；北京市网络安全与保护技术重点实验室资助

This work was supported by the National Key Research and Development Program of China (2021YFF0307203, 2019QY1300, 2018YFB0803602), the Youth Innovation Promotion Association of Chinese Academy of Sciences (2021156), the Strategic Priority Research Program of Chinese Academy of Sciences (XDC02040100), the National Natural Science Foundation of China for Young Scientists (61802404), the Project of CAS Key Laboratory of Network Assessment Technology, and the Project of Beijing Key Laboratory of Network Security and Protection Technology.

通信作者：刘松（liusong1106@iie.ac.cn）

中图法分类号 TP391

随着信息技术的发展，互联网已经成为人们生活中不可或缺的重要组成部分．域名系统（domain name system, DNS）是互联网的重要技术支撑，可以为用户提供便捷和灵活的网络服务，但同时也是网络攻击者的攻击目标和恶意活动的支撑资源，被广泛地滥用在多种网络攻击中．例如，恶意软件利用域名来定位其命令和控制（command and control, C&C）服务器，垃圾邮件包含的恶意链接通过 DNS 将用户重定向到漏洞利用或网络钓鱼页面[1-2]．

近年来，域名滥用日趋严重，对网络空间安全产生巨大威胁，大量检测工作随之涌现．学术界将散布恶意软件、操控僵尸网络、分发垃圾邮件、网络钓鱼和欺诈做为典型域名滥用方式进行研究[1, 3-4]．其中不乏一些总结性工作，文献[5-8]分别从数据来源、检测特征、检测算法、评估策略以及检测对抗技术等角度对现有工作进行梳理，分析域名滥用检测的纵向技术演进．然而，尚未有工作针对具体的域名滥用行为总结现有检测方法，提供全面的横向和纵向对比分析．

本文从典型的滥用行为出发，定义并分类不同的域名滥用检测场景，提供现有域名滥用行为检测方法的系统概述，有助于更好地理解当前域名滥用检测技术的应用场景、局限性和发展方向．

1 背景介绍

1. 1 DNS 简介

DNS 提供将域名映射到 IP 地址空间的服务，由域名空间和资源记录、名称服务器和解析器构成．

域名空间和资源记录：DNS 具有层次化的树状名称空间，称为域名空间，域名空间中的每个结点具有标签和一组资源记录．域名是当前结点到根结点路径上标签的组合，如图 1 所示：

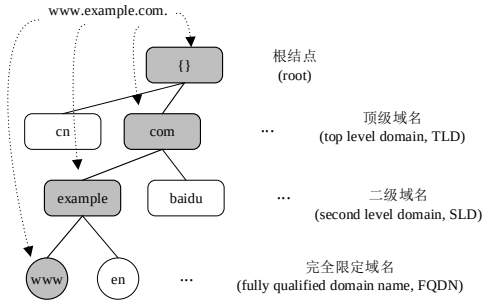


Fig. 1 DNS namespace

图 1 DNS 域名空间

名称服务器：DNS 分布式数据库由多台名称服务器构成．每台名称服务器负责保存域名空间部分结点的信息，同时提供域名解析服务．根据解析范围可划分为 4 类：根、顶级、权威和本地名称服务器．

解析器：解析器是请求资源记录的 DNS 客户端程序．DNS 支持 2 种解析方式：递归解析和迭代解析．递归解析发生在客户机向本地名称服务器发起解析请求阶段，迭代解析发生在本地名称服务器与其他名称服务器通信的过程中，图 2 展示 DNS 解析过程：

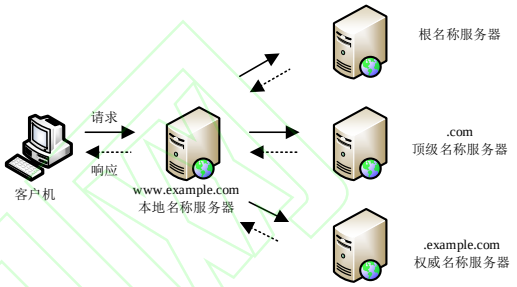


Fig. 2 DNS resolution process

图 2 DNS 解析过程

DNS 提高了网络的易用性，推动了互联网的发展进程，以 DNS 服务为基础的技术也不断涌现，例如泛域名解析和内容分发网络（content delivery network, CDN）．这些技术在为互联网提供高可用性和高性能的同时，也被攻击者广泛应用于恶意网络活动中，以提高滥用域名的可用性和对黑名单的抵抗能力，给域名滥用的甄别和检测带来极大的挑战[5, 9]．

1. 2 域名滥用检测场景分类

学术界和工业界尚未给出域名滥用的分类标准．本节通过调研现有域名滥用检测方案，对现存多种域名滥用行为进行汇总和归并，提出本文的域名滥用检测场景分类体系，如图 3 所示：

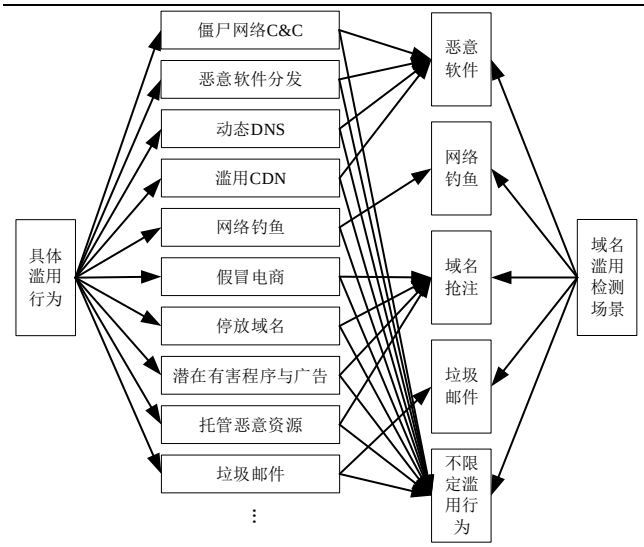


Fig. 3 Taxonomy of domain name abuse detection scenarios based on specific behaviors

图 3 基于具体行为的域名滥用检测场景分类体系

1.2.1 域名滥用行为调研分析

本文首先调研学术界和工业界有关发现、解释或缓解 DNS 生态系统中域名滥用情况的工作，抽取典型域名滥用行为，汇总得到表 1．值得注意的是，本文仅关注 2 类域名滥用行为：解析到受控恶意资源、服务于恶意活动．典型例子有恶意软件 C&C 通信、网络钓鱼．而诸如 DNS 隧道、DNS 放大和拒绝服务攻击、DNS 劫持等对 DNS 通讯协议或 DNS 基础架构发动攻击的 DNS 威胁，均不在本文的讨论范围内，在总结表 1 的过程中有针对性地予以剔除．

Table 1 Summary of Work Related to Domain Name Abuse

表 1 域名滥用行为相关工作汇总表

文献	研究目标	数据来源	域名滥用行为
[2]	分析新通用顶级域（gTLD）中的 DNS 滥用情况	新 gTLD 的区域文件、WHOIS 信息、DNS 流量数据、公共黑名单	网络钓鱼、恶意软件、垃圾邮件
[4]	构建 DNS 滥用框架，以解决 DNS 中的滥用问题	域名注册机构和注册服务商有关域名滥用行为的专业知识和经验	恶意软件，僵尸网络，网络钓鱼，垃圾邮件
[10]	分析 DNS 滥用模式、货币化形式及缓解方案	荷兰 TLD(.nl)的注册数据、权威服务器查询、资源记录	网络钓鱼、影子域名、假冒电商、恶意软件分发、僵尸网络 C&C、恶意搜索引擎优化
[11]	构建滥用域名分类体系及其威胁情报分析系统	25 种商业和公共黑名单	滥用 CDN、托管恶意资源、DGA、过期域名重新注册、Sinkhole 域名、影子域名、误植抢注、动态 DNS 域名
[12]	分析新注册域（newly registered domain, NRD）的滥用情况	1530 个 TLD 下的 NRD	C&C 域名、恶意软件分发、网络钓鱼、误植抢注、DGA、潜在有害程序和广告、欺诈域、垃圾邮件

进一步根据表 1 中列举的域名滥用行为关键词，从 Web of Science 核心数据库中筛选出近 10 年（2010—2020 年）领域相关文献 446 篇．使用 Citespace 工具对域名滥用行为检测领域的高频术语进行共现分析，得到图 4 共现关键词聚类时间轴图．

时间轴图以关键词出现年份为 *X* 轴（顶部年份）、关键词类簇标签为 *Y* 轴（右侧标签），反应该领域的热点主题．每个类簇的关键词共享一条时间轴线，轴线上的同心圆表示一个关键词，同心圆的大小代表出现频率、颜色深浅表示时间跨度，同心圆间的连线代表关键词共现情况．

从图 4 中可以看到，去除掉安全领域、通用领域关键词类簇标签，可以得到本领域的研究重点包括：恶意软件、僵尸网络、网络钓鱼、域名抢注、垃圾邮件．从同心圆连线的紧密程度分析，恶意软件和僵尸网络的共现关系较为紧密，典型共现关键词包括“Domain-Flux”“Fast-Flux”等，可将面向僵尸网络的域名滥用作为恶意软件的一个子类别进行概述．

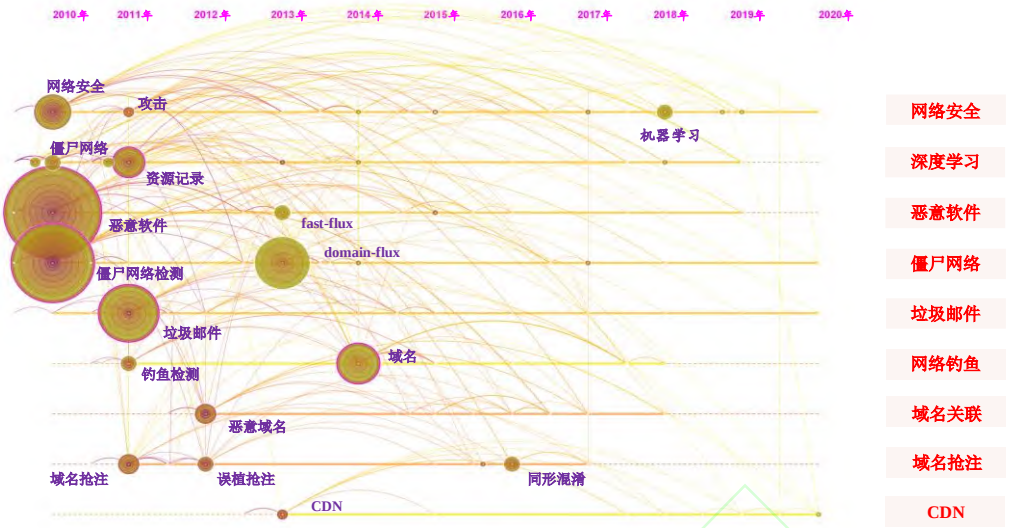


Fig. 4 Timeline view of domain name abuse detection co-occurrence keyword clustering

图 4 域名滥用行为检测领域的共现关键词聚类时间轴图

1.2.2 构建域名滥用检测场景分类体系

本文以图 4 的类簇标签为指导，对表 1 中存在重叠检测方案的域名滥用行为，面向检测场景进行合并：例如，僵尸网络 C&C、恶意软件分发、动态 DNS 和滥用 CDN 都依托于 Fast-Flux 或 Domain-Flux 等恶意软件常用的 DNS 动态解析技术实现，可归并为恶意软件场景下的检测目标；假冒电商、影子域名、托管恶意资源、欺诈域等滥用行为，常常通过抢注与流行域相似的域名诱导受害者访问，可归并为域名抢注检测场景下的检测目标；此外，本文增加不限定滥用行为检测场景，广泛捕获多种域名滥用行为，作为其他场景检测方案的补充．最终确定恶意软件、网络钓鱼、域名抢注、垃圾邮件、不限定滥用行为 5 类典型检测场景，作为本文的主要研究对象，如图 3 所示．

值得注意的是，在本文的分类体系中域名滥用行为存在交叉，例如，垃圾邮件可能包含钓鱼链接或恶意软件附件；且域名滥用检测方法往往融合多领域知识，例如，恶意软件域的检测可通过逆向工程、监听 DNS 通信流量等技术手段实现．本文在梳理域名滥用检测方法时，考虑仅适用于当前场景的检测方法，且仅关注使用 DNS 相关特征的域名滥用检测技术．

1.3 域名滥用检测特征及检测方法概述

当前阶段，域名滥用检测的主流技术方案大多采用机器学习方法，关键在于特征工程，即提取区分滥用域和良性域的检测特征．本文将滥用域检测特征划归为 6 个大类，下面分别进行简要介绍．

域名字符级特征：从域名字符级组成的角度区分良性域名与滥用域名．结构特征关注域名的结构属性；语言特征捕获语言模式的偏差；统计特征分析字符构成的随机性；相似度特征考察字符级的相似度．

Table 2 Character Level Features of Domain Name

表 2 域名字符级特征

类别	典型特征	滥用域名可能特点
结构特征	域名长度	域名长度较长
	子域个数	子域个数较多
语言特征	特殊字符/数字占比	特殊字符/数字占比较大
	数字、字母转换频率	数字、字母转换频率较高
统计特征	N-gram 频率分布指标(均值、方差、5 数概括)	N-gram 频率分布较为离散
	域名期望	域名期望较小
相似度特征	域名相似性指标(K-L 距离、马氏距离等)	字符级构成相似

域名解析特征：从 DNS 通信流量及应答记录中捕获异常．记录特征关注滥用域资源记录频繁改变和资源重用的特点；空间特征考虑滥用域的地理空间分布较为松散的特点；时间特征分析域名的活跃时段、生存时间（time to live, TTL）等与时间相关的特征．

Table 3 Domain Name Resolution Features

表 3 域名解析特征

类别	典型特征	滥用域名可能特点

记录特征	A/AAAA/NS/CNAME 记录个数	记录个数较多
	A/AAAA/NS/CNAME/MX 记录	记录重叠程度较高
空间特征	A/AAAA/NS 记录映射国家数目	国家数目较多
	A/AAAA/NS 记录映射 AS 数目	AS 数目较多
时间特征	TTL 值统计指标	TTL 值较小
	各时段内域名解析次数	突发、集中解析现象
页面特征	重定向页面（refresh 属性）	存在页面重定向
	页面质量评分	页面质量评分低
相似度特征	URL 相似性指标(编辑距离、Q-gram 距离等)	与合法知名域名的相似度较高

域名关联特征：基于与已知滥用域有强关联的域很可能是滥用域的假设，通过分析主机、域名和 IP 之间的查询、解析关系，捕获潜在滥用域．主要包括主机-域名的查询关联和域名-IP 的解析关联特征．

统一资源定位符（uniform resource locator, URL）特征：常用于网络钓鱼域名的检测．字符级特征分析 URL 的字符组成；访问特征关注用户访问 URL 的特点；页面特征考察域名托管页面的特点；相似度特征关注钓鱼链接与官方链接的相似性．

Table 4 URL Features

表 4 URL 特征

类别	典型特征	滥用域名可能特点
字符级特征	URL 中连接符/点的个数	连接符/点的个数较多
	长域名是否包含有效域名	长域名包含有效域名
访问特征	URL 是否在浏览历史记录中	不在浏览历史记录中
	用户在网站的停留时间	停留时间较短

域名抢注特征：域名抢注特征主要关注 5 类抢注域的典型构造方式，包括误植抢注构造特征（替换、增加、删减流行域名的字符），字符比特翻转特征，同音字符替换特征，同形字符替换特征，组合抢注构造特征（拼接流行域名和常见短语）．

辅助信息特征：部分检测方案借助辅助信息判别滥用域，常用的辅助信息包括：公共情报数据、whois 信息、域名区域文件信息等，挖掘这些信息可以辅助域名判别．

基于前述的特征分类，表 5 汇总 5 类典型域名滥用检测场景涉及的主要检测方法及其检测特征．以恶意软件检测场景为例进行介绍，主要检测方法可划分为基于 IP-Flux 和基于 Domain-Flux 这 2 类，其中，基于 IP-Flux 的具体检测方案，使用的典型特征包括域名字符级特征、域名解析特征、域名关联特征和辅助信息特征 4 大类别，并列出相应的参考文献．本文将在第 2 节进一步梳理各域名滥用检测场景下的典型检测方案．

Table 5 Summary of Detection Methods and Typical Features of Domain Name Abuse Detection Scenarios

表 5 域名滥用检测场景的检测方法及典型特征汇总

检测场景	检测方法	典型特征
恶意软件	基于 IP-Flux	域名字符级特征[13-15]、域名解析特征[13-23]、域名关联特征[14]、辅助信息特征[14-15]
	基于 Domain-Flux	域名字符级特征[24-39]、域名解析特征[40]、域名关联特征[26-28]、辅助信息特征[26-28, 31]
网络钓鱼	基于黑白名单	URL 特征[41-45]、辅助信息特征[43, 46]
	基于 URL	域名字符级特征[47]、域名解析特征[48-49]、URL 特征[48-58]、辅助信息特征[48-50]
域名抢注	误植抢注检测	域名解析特征[59]、域名抢注特征[59-63]、辅助信息特征[61]
	比特抢注检测	域名抢注特征[64-65]
	同音抢注检测	域名抢注特征[66]、辅助信息特征[66]
	同形抢注检测	域名抢注特征[67-71]
	组合抢注检测	域名抢注特征[72]
	多类抢注检测	域名抢注特征[73-75]
垃圾邮件	基于 spam botnet	域名解析特征[76-81]、辅助信息特征[82]
	基于注册信息	域名字符级特征[83]、辅助信息特征[83-84]
不限定滥用行为	基于信誉评分	域名解析特征[9, 85]、域名字符级特征[9, 85]、辅助信息特征[85]
	基于图推理	域名字符级特征[86-89]、域名解析特征[86-88]、域名关联特征[86-91]

2 不同场景下的域名滥用检测

本节分别针对恶意软件、网络钓鱼、域名抢注、垃圾邮件以及不限定滥用行为 5 类场景下的域名滥用检测工作进行梳理. 其中，每一场景下的检测工作都依次按照检测方法、具体检测方案进行归并. 本节聚焦检测技术的演进过程，提供一个以检测场景为导向的域名滥用行为检测工作概述.

2.1 面向恶意软件的域名滥用检测

恶意软件包括蠕虫，木马，僵尸程序或其他具有恶意意图的程序，旨在破坏计算机系统的运行，窃取专有信息或获得访问控制权限. 据 AV-Test 发布的最新数据显示，截至 2020 年底，全球范围内统计到的恶意软件总数已超 11 亿，约为 2011 年的 17 倍[92].

恶意软件大多滥用 DNS 协议，利用动态 DNS 解析技术（IP-Flux 或 Domain-Flux）实现受感染主机与 C&C 服务器之间的通信. 考虑到逆向工程成本较高，许多工作通过分析 DNS 流量，挖掘恶意通信域名，实现感染主机的检测以及 C&C 通道阻断. 下面将恶意软件检测场景的研究划分为 2 大类：基于 IP-Flux 的检测和基于 Domain-Flux 的检测. 表 6 列举典型检测工作在面向恶意软件的域名滥用检测领域的贡献及其局限性.

2.1.1 基于 IP-Flux 的检测

互联网名称与数字地址分配机构（Internet Corporation for Assigned Names and Numbers, ICANN）将 IP-Flux（Fast-Flux）描述为“对主机和/或名称服务器资源记录快速且反复地更改，使得域名解析为动态变化的 IP 地址”. 使用 IP-Flux 技术构建的网络，被称为速变服务网络（fast flux service networks, FFSN）. 长期以来，IP-Flux 技术被多种恶意软件滥用，同时也是滥用域名的有力检测指征.

基于 IP-Flux 的检测技术演进过程概述：早期采用主动 DNS 探测方案，提取 DNS 解析特征进行检测，随后发展为被动 DNS 分析技术，研究重点在于区分 FFSN 和 CDN. 在持续的攻防博弈中，攻击者不断提高 IP-Flux 技术的隐蔽性，检测算法也相应地引入多类检测特征. 现阶段，考虑到传统检测方案的局限性，学术界将深度学习算法应用于 IP-Flux 域名检测.

1）主动 DNS 探测方案

主动 DNS 探测方案通过在固定时间窗口内向待检测域名发起多次 DNS 解析请求，提取 DNS 响应特征，构造分类器，实现滥用域名的判别.

Table 6 Comparison of Typical Domain Abuse Detection Works Oriented to Malware

表 6 面向恶意软件的典型域名滥用检测工作的比较

检测方法	检测方案	文献	检测特征	检测算法	优点	缺点
基于 IP-Flux 的检测	主动 DNS 探测方案	[16]	域名解析记录特征	线性判别函数	可以区分 CDN 和 FFSN，误报率较低	无法自适应更改参数，存在检测时延
		[17]	域名解析记录特征，域名解析时间特征	贝叶斯网络	自动化参数学习；压缩时间窗口降低检测时延	主动发送 DNS 解析请求，增加网络负担
	被动 DNS 分析方案	[14]	域名解析记录特征，域名解析时间特征，域名结构特征，域名-IP 关联特征，辅助信息特征	随机森林	综合提取多维检测特征，检测精确率高	检测特征设计和提取成本较高，且易被攻击者规避
		[20]	域名解析记录特征	DenseNet、BiLSTM	隐式提取空间特征和时序特征，提升检测精度	模型较为复杂，训练和预测时间较长
基于 Domain-Flux 的检测	传统检测方案	[25]	域名结构特征，域名统计特征	J48 决策树	自动化参数学习，可解释性强，检测速度快	检测特征易被规避，仅支持二分类
		[28]	域名统计特征，域名结构特征，主机-域名关联特征，辅助信息特征	X-means 聚类、交替决策树、隐马尔科夫模型	仅分析 NXDomain 流量有效压缩检测数据，支持 DGA 家族多分类	多分类检测效果不佳，容易误分类新的滥用域家族
	基于深度学习的检测方案	[32]	域名字符级特征，域名关联特征	WordGraph、随机森林、CNN	可以检测基于字典生成的 AGD，同时获得字典	域名较多时 WordGraph 无法有效划分连通分量
		[34]	域名字符级特征	成本敏感 LSTM.MI	解决训练样本类别不平衡问题	成本项固定，不支持自动调参，泛化能力较差

[39]	域名字符级特征	成本敏感的 BiLSTM	可检测对抗 DGA 样本，对常见规避方案鲁棒	只支持检测英文字符构成的域名

最早一篇具有指导性意义的工作是 Salusky 等人[93]发表的 Honeynet 项目论文．他们给出基于 IP-Flux 的滥用域名检测问题的实验建议，提出在固定时间窗口内，对域名的 DNS 解析结果集进行评分的方案，为后续主动 DNS 探测方案奠定基础．2008 年 Holz 等人[16]发表有关 FFSN 的首个实证研究，他们对所有域执行 2 次 DNS 查询，从域解析结果集中提取检测特征，计算待检测域的 flux 分数．该方案借助阈值判定滥用域，无法自适应检测不断变化的 FFSN 架构．

主动探测方案的时间窗口大小对检测效果影响较大：过小的时间窗口无法获得足够的信息，过大的时间窗口则会造成检测延迟，这种延迟可能会导致攻击扩散．针对这一问题，Caglayan 等人[17]同时使用主动 DNS 探测和被动 DNS 监听技术，利用历史解析数据有效压缩时间窗口，实现分钟级别的 FFSN 检测．该篇工作引入被动 DNS 数据，具有开创性意义，但本质仍属于主动 DNS 探测方案．

主动 DNS 探测方案的优点是数据收集的灵活性和易用性，可以有针对性地获取待检测域名的解析信息；局限性在于会产生过多的网络传输流量，增加网络负担，密集的 DNS 请求易被攻击者发现．

2）被动 DNS 分析方案

被动 DNS 分析方案主要通过在网络中部署被动 DNS 传感器或访问 DNS 服务器日志以获取真实的 DNS 查询和响应，被动地收集 DNS 数据，相较主动 DNS 探测方案检测成本低，且较为隐蔽．目前大多数研究都采用被动分析方案，具体的研究方向包括检测特征的设计和检测算法的改进：

① 检测特征设计：许多工作根据 IP-Flux 技术的典型特点进行检测特征构造．文献[18]关注到 FFSN 具有地理分布更均匀和空间服务关系更松散的特点，使用空间快照机制取代基于时间的特征，提出一种无延迟的检测系统 SSFD，图 5 展示该系统的工作流程：

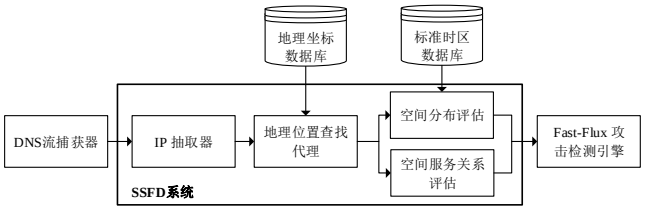


Fig. 5 System workflow of spatial snapshot Fast-Flux detection

图 5 SSFD 系统的工作流程

恶意软件逐渐引入 Domain-Flux 技术，Stalmans 等人[13]将域名字符级特征融入 IP-Flux 检测方案，分析域名的字母数字分布频率，计算总变异距离和概率分布来检测滥用域．此外，随着 IP-Flux 隐匿技术不断发展，常用检测特征被有针对性地规避，文献[14]关注到攻击者难以隐藏域名和 IP 的解析关联，从多样性、时间性、增长性和相关性等角度提取多个鲁棒检测特征，在真实 DNS 流量上检测准确率达 90%．

考虑到传统检测方法的检测成本较高，Yang 等人[19]在特征设计上抛弃依赖外部资源和不稳定的特征，从 DNS 响应中提取检测特征，训练轻量级检测模型．并加入自动更新模块，实时反馈当前流量情况，调整模型参数适应流量变化，从而提高检测精度．

② 检测算法改进：随着 IP-Flux 技术的发展，一些攻击者通过精心设计恶意网络通信架构，可以对一些传统检测特征进行有效规避，例如 TTL 值、域名字符级组成特点．此外，传统检测方案对检测特征的精确性要求很高，需要领域知识进行指导．针对上述问题，研究人员将深度学习算法应用到滥用域名检测中，神经网络模型可以自动实现隐式检测特征提取并捕获深层次的异常模式，有效识别滥用域．

文献[15]直接将域名字符，经验信息，地理和时间相关特征进行编码，输入到长短期记忆（long short-term memory, LSTM）模型中，有效检测 IP-Flux 域名．2020 年，牛伟纳等人[20]则结合 DenseNet 模型和 BiLSTM 模型，同时捕获 IP-Flux 域名的空间特征和时序特征，进一步提升检测精度．此外，Almomani 等人[21]提出一种基于自适应演化模糊神经算法（evolving fuzzy neural network, EFUNN）的 Fast-Flux 追踪系统，采用有监督和无监督的混合在线学习方案，支持实时在线检测 IP-Flux 域名．

深度学习方案的性能受限于训练样本集的大小，在样本量不够庞大，人工提取特征已经具有较强表征能力的情况下，深度学习方案的检测能力可能弱于传统机器学习方案．文献[22-23]对比了常见传统机器学习算法以及不同层数的深度神经网络（deep neural network, DNN），实验证明 DNN 检测精度并未随着隐层层数增加而提升，且弱于传统机器学习方案的检测性能．上述工作说明对于 IP-Flux 检测算法的选取，需要依据检测特征、样本集规模进行灵活选择．

被动 DNS 分析方案解决主动探测方案中大量发送 DNS 请求导致的网络负担加重和易被攻击者发现的问题；此外，被动收集数据可以获取时间跨度较大的历史数据，支持长期分析．但由于数据收集的被动性，无法检测尚未使用的潜在滥用域．

2.1.2 基于 Domain-Flux 的检测

Domain-Flux 使用域生成算法（domain generation

algorithm, DGA）定期生成大量伪随机的算法生成域（algorithmically generated domains, AGD），并将多个 AGD 映射到单个 IP 地址，有效规避静态黑名单的检测．Domain-Flux 策略常被滥用在分发垃圾邮件或实施网络钓鱼的僵尸网络基础架构中，用于标识受害者，或对反垃圾邮件技术进行绕过．

基于 Domain-Flux 的检测技术演进过程：早期滥用域名的检测主要使用传统检测方案，利用域名字符级特征，以及典型流量特征对滥用域名加以识别；现阶段的研究方案主要使用面向无特征的深度学习算法，针对检测模型的输入和结构进行改进，并关注到攻击者精心设计的 AGD 的针对性检测方案．

1）传统检测方案

早期 Domain-Flux 域名的检测主要依赖于统计算法、传统机器学习算法，根据检测特征可划分为基于域名字符级特征和基于典型流量特征 2 种方案．

① 基于域名字符级特征的检测方案：考虑到 AGD 通常较长且无语义信息等特点，抽取 DNS 流量中的域名信息，构造字符级特征．实现滥用域的检测．

文献[24]计算待检测域名在 N-gram 空间上与良性域和滥用域的相似性，实现滥用域的检测．该方案需要收集足够数量的域名，以准确估算 N-gram 分布，检测成本较高且存在较大的检测延迟．此外，该方案仅能检测已知恶意软件家族的滥用域名．文献[25]则主要使用域名的长度和域名的期望值 2 个特征，支持未知 AGD 家族滥用域的挖掘．

除上述二分类 AGD 检测方案外，部分工作支持对 AGD 域按照具体 DGA 算法进行多分类．文献[26]提出的 Phoenix 系统，使用有意义的字符比率和 N-gram 正态分数 2 类特征，根据马氏距离对待检测域进行分类，进一步构建 DGA 域-IP 二部图，利用 DBSCAN 算法实现多分类 AGD 域的检测．文献[27]对 Phoenix 的工作进行了扩展，使用了信息熵、N-gram 出现频率等域名字符级特征，并使用改进的马氏距离实现域的多分类，图 6 给出该检测系统的整体架构：

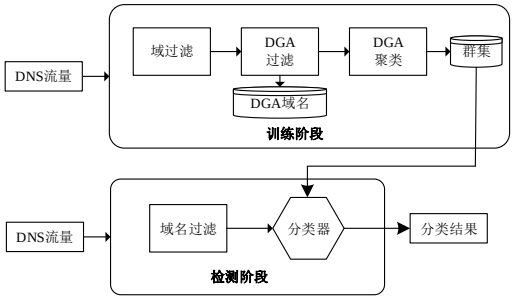


Fig. 6 Structure of the DGA botnet detection system

图 6 DGA 僵尸网络检测系统的结构

基于域名字符级特征的检测方案易于被攻击者绕过：由于 DGA 的实现方案各有不同，AGD 的字符级特征表现也不尽相同，基于 Domain-Flux 的滥用域通过特征变化可以逃避有针对性的检测．

② 基于典型流量特征的检测方案：部分工作通过分析 DNS 流中的不成功解析域名（NXDomain）响应，引入典型流量特征，检测基于 Domain-Flux 的恶意软件类域名．这是由于攻击者只会注册 DGA 算法生成的一小部分域名，用于与被感染机器通信，因此当恶意软件发起 DNS 查询时，会产生大量 NXDomain 响应，以最终定位 C&C 服务器．

文献[28]通过从 NXDomain 流中提取域名字符级特征以及关联特征，聚类待检测域名，进一步训练隐马尔可夫（hidden Markov models, HMM）模型和多分类器，识别滥用域及其所属家族．该工作首次采用无需特征工程的 HMM 模型进行检测，然而后续工作[29]证实该方案在多分类检测中表现不佳．

此外，Zhou 等人[40]观察到使用 Domain-Flux 技术的域在子域数量、域的生存周期以及域的访问模式方面与合法域存在显著差异，进一步提出一个仅依赖于流量特征的检测系统，如图 7 所示．该检测方案不使用域名字符级特征，可以提高检测方法的鲁棒性，但需要更长的时间来确定可疑域列表．

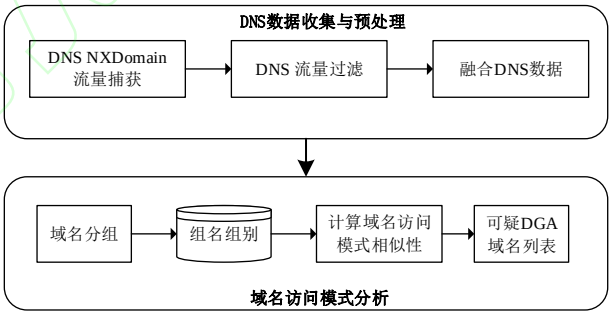


Fig. 7 Structure of detection system based on traffic features

图 7 基于典型流量特征的检测系统架构

基于典型流量特征的方案通过检测 NXDomain 流，可以压缩待分析数据、有效提高检测准确率；该方案的局限性在于提取流量特征会增加检测时延，此外，多数方法仍需构造字符级特征辅助检测．

2）深度学习检测方案

随着 Domain-Flux 技术的不断发展，许多 DGA 算法通过模拟合法域的字符级组成规避检测，使得人工设计的域名字符级检测特征失效．此外，特征提取和表示工程通常十分繁琐，无法自适应检测多种 DGA 算法生成的滥用域名．面对这一挑战，研究学者将深度学习的方法应用到滥用域名检测中，实现无人工特征提取的检测方案．这是由于 DGA 域名的检

测主要依托于对域名字符组成进行分析，可以作为自然语言处理（natural language process, NLP）领域的一个子问题，目前多数 NLP 问题都适用深度学习的解决方案．同时互联网中存在海量的域名，庞大的数据集为深度学习模型的训练提供有力的支撑．通过深层神经网络，可以有效捕获域名字符的语言模式异常，从而有效判别 DGA 域名．

最早应用深度学习的检测工作是 Woodbridge 等人[29]利用 LSTM 模型检测 DGA 及其所属家族，将域名的每个字符作为模型输入，经过嵌入层、LSTM 层和逻辑回归 3 层处理，得到分类结果，基于循环神经网络（recurrent neural network, RNN）的 DGA 检测方案大多采用类似的网络结构．该方案无需人工提取特征，对比以往的工作具有较好的检测效果．文献[30]对比随机森林、LSTM、卷积神经网络（recurrent neural network, CNN）这 3 种算法的检测效果，实验表明深度学习算法相对传统机器学习算法在检测准确率上有较大提升，可以更好地处理数据中的噪声．

后续的检测工作主要针对模型输入及模型结构进行相应改进，并关注到攻击者精心设计的 AGD．

① 模型输入改进：仅将域名字符序列作为深度学习模型的输入，对基于字典生成的 AGD 检测效果不佳，例如，Suppobox 家族使用基于字典的伪随机域名，通过拼接 2 个字典的单词创建 AGD，这使得域名在字符级别与合法域的差别很小，不足以支撑检测．针对这一问题，Curtin 等人[31]提出了一种名为 smashword 评分的机制，用于评估域名与英语单词的相似程度．进一步引入 WHOIS 信息辅助检测，可以有效捕获基于字典的 AGD 域．此外，Pereira 等人[32]使用 WordGraph 方法获取 DGA 的检测词典．该方案首先对域名进行分词预处理，构建 WordGraph 抽取连通分量，捕获域名字符级关联信息，得到检测词典，进一步训练随机森林和 CNN 分类器检测 AGD 域．

此外，考虑到感染恶意软件的主机会查询大量 NXDomain 的行为特点，Tong 等人[33]将 NXDomain 作为模型输入，提出基于 CNN 模型的检测系统 D3N．在二分类的 DGA 检测中，D3N 各项指标的性能相对传统机器学习方案都有明显能提升．

② 模型结构改进：受限于样本标注的成本，数据集中良性域以及各 DGA 家族样本的数目相差很大．考虑到原始 LSTM 模型均等对待所有样本，对样本类别失衡敏感的问题，文献[34]提出一种改进的成本敏感算法 LSTM.MI，将成本项引入反向传播机制，给样本数多的类别赋予较低的权值，提升检测精度．

此外，考虑到现有检测特征提取思路的单一性，文献[35-36]均采用 CNN 和 LSTM 的混合结构，支持从不同维度提取域名的字符级特征：CNN 提取域名的 N-gram 空间特征，LSTM 学习 N-gram 的序列上下文特征．实验证明混合结构优于单一结构的检测精度，图 8 给出文献[35]的网络架构示意图．

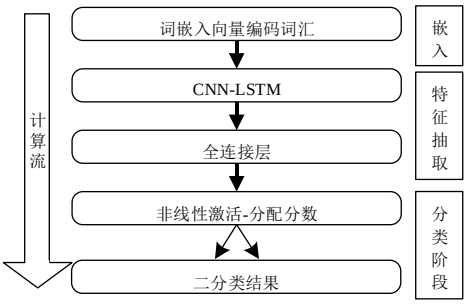


Fig. 8 CNN-LSTM based DGA detection network architecture

图 8 基于 CNN-LSTM 的 DGA 检测网络架构

③ 检测攻击者精心设计的 AGD：在长期的攻防博弈中，攻击者通过精心设计 DGA 算法生成高隐蔽性、高对抗性的 AGD，绕过普通检测方案，典型的有隐身域生成算法（stealthy domain generation algorithms, SDGA）和对抗 DGA 样本．

SDGA 最早由 Fu 等人[94]提出，他们基于 HMM 和概率上下文无关文法（probabilistic context-free grammars, PCFG）从字典中选择字符，生成模仿良性域名字符分布的 SDGA．大多数传统的基于熵的方法，或基础的 LSTM，CNN 等深度学习方案不能有效地检测 SDGA．Yang 等人[37]提出一种异构深度神经网络框架（heterogeneous deep neural network, HDNN）．HDNN 采用具有多尺寸卷积核的并行 CNN 架构，以及基于自注意力机制的双向 LSTM 结构，同时支持提取多尺度局部特征和双向全局特征，挖掘待检测域名的深层语义信息，可以有效捕获 SDGA．

一些攻击者采用对抗性机器学习（adversarial machine learning, AML）思想，基于生成对抗网络（generative adversarial network, GAN）算法生成对抗 DGA 样本．GAN 框架包括生成模型和判别模型，生成模型接收随机噪声用于生成对抗样本，判别模型判别给定样本来自真实数据还是对抗样本，构成一个动态博弈过程．训练收敛时，生成模型可以生成和真实数据高度相似的对抗样本．Anderson 等人[38]最早提出 DeepDGA 框架，将 GAN 引入 DGA 算法．该框架首先在 Alexa Top 1M 域名上预训练自动编码器（编码器+解码器），随后将编码器和解码器在 GAN 中竞争性重新组装，最后使用 GAN 生成器优化域名的伪随机生成，实验证明 DeepDGA 得到的对抗 DGA 样本可以有效绕过深度学习和传统机器学习检测方案．CharBot[95]，MaskDGA[96]，ShadowDGA[97]等方

案也相继被提出，这些方案生成的对抗 DGA 样本既可以被攻击者作为绕过现有检测方案的 AGD，也可以被防守方作为数据增强加入到检测模型的训练集中，提高对抗未知 DGA 样本的检测能力和模型的泛化能力．2021 年，Ravi 等人[39]提出一种对抗性防御方案，采用集成深度学习模型对 NXDomain 流量进行检测，可以有效捕获 DeepDGA，CharBot，MaskDGA 这 3 类对抗 DGA 样本，检测准确率可达 99%以上．

2.2 面向网络钓鱼的域名滥用检测

美国计算机应急准备小组将网络钓鱼定义为一种社会工程形式，通过伪装成可信赖的组织或实体，使用电子邮件或伪造的网站收集敏感信息并开展恶意活动．根据反网络钓鱼工作组（Anti-Phishing Working Group, APWG）2020 年第 3 季度报告[98]，全球范围内的网络钓鱼站点数量共计 199133 个，仍是网络攻击和诈骗的重要手段之一．

域名滥用与钓鱼攻击联系较为紧密，网络钓鱼者通常会注册与官方网站相似的域名，或将钓鱼页面挂载在权威域名下，增加钓鱼成功率；此外，DNS 也是网络钓鱼者的攻击目标之一，例如 DNS 缓存投毒攻击、DNS 劫持等，通过攻击 DNS 将用户引导至攻击者布置好的钓鱼界面．因此，从域名处着手是防范和阻止钓鱼攻击的有效方案．

网络钓鱼网站的检测按照技术手段可划分为基于黑白名单的检测技术、基于 URL 的检测、基于内容的检测、以及基于视觉相似性的检测．由于本篇文章仅关注网络钓鱼类域名的检测方案，因此着重对涉及域名检测的黑白名单技术和 URL 检测技术进行介绍，诸如基于内容、视觉相似性等检测方案，则主要涉及了网页内容的比对，不在本篇文章的讨论范围内．表 7 列举典型检测工作在面向网络钓鱼的域名滥用检测领域的贡献及其局限性．

2.2.1 基于黑白名单的检测

黑白名单技术是钓鱼类域名检测的早期方案，同时也是部署与应用最广泛的反钓鱼技术．

1）基于黑名单的检测方案

基于黑名单的检测方案普遍被各大安全厂商、浏览器厂商使用，根据已知的黑名单对请求的 URL 进行检查，如果存在条目匹配，则将该 URL 标注为网络钓鱼站点，进而限制访问或生成警告．黑名单一般通过手动报告、链接分析或网络爬虫等手段构建，因此不可避免的存在遗漏的情况，同时缺乏检测新钓鱼网站的能力．黑名单维护者仅在钓鱼网站变为活动状态时，才能将域名加入黑名单中，因此会存在一个攻击窗口，在这个窗口内滥用域名尚未被黑名单记录，用户可能会遭受钓鱼攻击．为了缓解黑名单的滞后检测问题，Prakash 等人[41]于 2010 年提出了 PhishNet，该方案可以基于现有黑名单预测新的恶意 URL，并实现给定 URL 与黑名单中条目的近似匹配．尽管 PhishNet 无法实时检测网络钓鱼站点，但是在大型数

Table 7 Comparison of Typical Domain Abuse Detection Works Oriented to Phishing

表 7 面向网络钓鱼的典型域名滥用检测工作的比较

检测方法	检测方案	文献	检测特征	检测算法	优点	缺点
基于黑/白名单的检测	基于黑名单的检测方案	[41]	URL 字符级特征	线性决策函数	可以标记原始黑名单外的滥用域名，误报率较低	无法实时检测钓鱼域名
		[42]	页面特征	基于规则的检测	基于少量样本，主动探测未知 URL，实时性较强	对钓鱼域名种子依赖性较强，易被规避
	基于白名单的检测方案	[44]	URL 访问特征	朴素贝叶斯	基于用户登录历史个性化定制白名单	敏感信息记录，设备依赖性强
		[45]	页面特征	线性决策函数	基于页面超链接特征检测，不访问用户敏感信息	检测方案易被规避，无法动态捕获新型攻击
基于 URL 的检测	传统机器学习检测方案	[48]	URL 字符级特征，辅助信息特征，域名解析特征，	朴素贝叶斯、支持向量机、logistic 回归分类器	扩充特征集并进行特征筛选，提升检测准确率	检测特征构造借助外部资源，检测成本较高
		[47]	域名字符级特征	随机森林	特征计算复杂度低，检测精确率高，检测速度快	检测特征较为单一，人工设计特征，易被绕过
	深度学习检测方案	[55]	URL 字符级特征	基于敏感词分词 CNN-BiLSTM	支持获取 URL 敏感词、特殊字符，提升信息利用率	模型输入较为单一，仅支持字符嵌入输入
		[58]	URL 字符级特征	GAN AutoEncoder	对抗 URL 样本扩展数据集，提高模型的泛化能力	GAN 架构复杂，训练较为困难，且输入定长

据集上，它的误报率较低，并且在标记不属于原始黑名单的新 URL 方面非常有效．除检测速度外，覆盖范围也是黑名单性能的重要评估指标，有助于跟踪短时间内网络钓鱼威胁的变化轨迹．2014 年，Lee 等人[42]提出一种主动网络钓鱼检测框架（PhishTrack），包括重定向跟踪和表单跟踪 2 个组件，主动更新网络钓鱼黑名单，以提高黑名单覆盖率，是反网络钓鱼黑名单技术的有效补充．

除反网络钓鱼黑名单更新方案外，还有一些工作对黑名单的防护能力进行了定量的评估，旨在发掘反网络钓鱼黑名单技术的现存问题，为未来的研究方向提供指导．2020 年，Bell 等人[99]调查了谷歌安全浏览（Google safe browsing, GSB）、OpenPhish、PhishTank 这 3 个网络钓鱼黑名单的钓鱼链接添加、删除、持续时间和重叠情况，并统计每个黑名单中涉及的域名总数．测量实验发现，钓鱼链接平均持续时间非常短暂，且频繁出现删除 1 天后重新添加的情况，表明钓鱼链接被过早删除或存在重新上线的行为，可以作为未来钓鱼黑名单构筑的一个研究点．Oest 等人[100]提出一个识别复杂网络钓鱼攻击的检测框架 PhishTime．PhishTime 检测使用复杂规避技术钓鱼站点，并在受控环境中大量复制它们作为测试样本，报告给 GSB，SmartScreen，Opera 这 3 个黑名单，测量黑名单对测试样本的平均响应时间，分析黑名单的检测速度和覆盖范围．PhishTime 进一步讨论黑名单对复用域名进行持续钓鱼攻击的检测能力．实验证实，现有黑名单允许网络钓鱼者重用域名进行多次攻击，可以作为优化黑名单防御能力的突破点．

2）基于白名单的检测方案

考虑到黑名单的检测延迟导致的攻击窗口问题，部分工作使用白名单对钓鱼网站进行检测，维护受信任站点的白名单列表，对未知站点请求弹出警示．

文献[43]提出了一种基于白名单的钓鱼网站防护方案，可以阻止用户访问已知的钓鱼站点，并通过执行 URL 相似性检查，警示用户疑似钓鱼站点的访问．该方案是典型的基于静态、通用白名单的钓鱼检测方案，局限性在于对外部信任站点列表具有较强依赖性，外部信任站点的准确性和完整性将会极大影响该类检测方案的性能．考虑到上述方案的局限性，2008 年，Cao 等人[44]提出了一种动态、定制化的白名单方法，根据用户的历史登录信息创建用户个人白名单，当用户尝试提交机密信息到白名单外的网站时发出警示．相似的，Dong 等人[46]提出基于用户行为的网络钓鱼检测系统，将用户访问过 3 遍以上的网站加入个人白名单，并存储域和用户凭证的对应关系，进一步使用线性决策函数判别钓鱼域．

考虑到上述自动化白名单构建方案存在用户敏感信息访问的问题，Azeez 等人[45]提出一个基于页面超链接校验的白名单自动更新方案，当用户访问未知站点时，抽取未知站点的页面超链接，通过分析是否存在空链接、超链接总数以及外部链接个数，判断站点的性质，并将良性站点加入白名单中．

白名单方案提供严格、全面的钓鱼网站防护能力，但在获取合法网站方面有很大的局限性．此外，定制化白名单需要与数据库同步更新、集中维护．

基于黑/白名单的检测方案存在一个共性缺点，即它们都需要时间来更新列表，存在检测延迟，同时难以全面覆盖互联网中每日新增的海量链接．

2.2.2 基于 URL 的检测

基于黑/白名单检测方案是静态的，难以甄别海量新增域名．专家学者引入机器学习的方案，动态识别潜在钓鱼域名：从经验数据中总结规律，构建网络钓鱼域的检测模型，实现对未知样本的有效检测，典型方案是基于 URL 的检测方案．

基于 URL 的检测主要通过分析 URL 的结构以及词汇特征，检测网络钓鱼站点．URL 用于标识因特网上资源的全球地址，由 2 个主要部分组成[50]：1）协议标识符，指示要使用的协议；2）资源名称，指定资源所在的域名、文件路径和查询字符串．如图 9 所示：

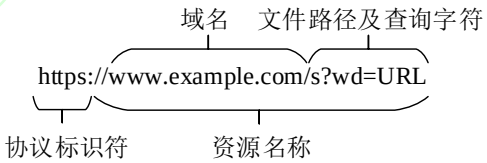


Fig. 9 Example of URL

图 9 URL 示例

1）传统机器学习检测方案

基于 URL 的传统机器学习检测方案，主要是在特征工程上进行创新，通过分析钓鱼 URL 的特性，人工提取鉴别钓鱼域的典型特征，提升检测精度．

早期基于 URL 的钓鱼域检测方案沿用黑/白名单信息作为证据特征，并加入有关域名注册和 DNS 解析信息等相关特征．典型工作有 2007 年 Garera 等人[50]使用 URL 的域名是否属于白名单，作为鉴别钓鱼域的一个显著特征．文献[48]则主要使用了 URL 词汇特征、域名辅助信息特征和 DNS 解析特征对钓鱼域进行检测．他们在后续工作[49]中采用在线学习方法，支持检测模型的自动更新．

上述检测特征的抽取都需要借助外部资源信息，例如黑/白名单或域名的注册及解析信息，检测成本

较高且存在检测时延. 一些专家学者仿照域名字符级特征的构造方式，直接分析 URL 字符级组成，常见的检测特征则包括有意义字符的长度、URL 中点的数量、URL 的长度、URL 域名中抽取品牌名、多子域、字符熵等特征. 典型工作有 Zouina 等人[51]提出的基于支持向量机和相似度索引的轻量型 URL 检测系统. 他们主要关注 URL 中的域名部分，提取出 6 种特征. 检测精度可达到 95.80%，证实 URL 的字符级特征可以有效鉴别钓鱼域. 2019 年，王雨琪等人[47]定义基元和敏感度描述钓鱼 URL 的语言特征，通过计算主级域名基元的相似性和利用随机森林算法学习子域名的语言特征，对 URL 进行检测，在降低误判率的同时，有效减小时间开销.

为进一步提升检测精度，HTML 特征、文本特征被引入到基于 URL 的检测方案中，在保证实时检测能力和较低检测成本的同时，提供更精准的钓鱼域鉴别能力. 2019 年，杨鹏等人[52]提出一种基于 Logistic 回归和 XGBoost 的钓鱼域名检测方案. 除 URL 的字符级特征，他们从站点源码中提取 24 个 HTML 特征，并利用 Logisitc 回归训练基于 TF-IDF 的网页文本特征，极大压缩融合特征的维度，最终使用 XGBoost 算法对钓鱼 URL 进行分类. 通过对比不同的特征融合方案和分类算法，证实该检测方案的优越性，在高准确率的同时兼顾检测速度.

2）深度学习检测方案

考虑 URL 的构成比域名更为复杂，且文件路径和查询字符部分完全受控于攻击者，具有高度动态变化、生命周期较短和指数爆炸式增长的特点. 传统机器学习方案对检测特征的依赖性较大，人工构造特征的成本较高，且容易被攻击者绕过，深度学习方案被引入到 URL 检测领域中. 迄今为止，深度学习方案在海量数据集上表现出强大的学习能力，并在中多分类问题中取得了最先进的结果. 通过端到端的深度学习方案，模型可以自动抽取 URL 深层次的语义特征，攻击者难以有效绕过. 现阶段基于 URL 的钓鱼域检测方案，主要是在检测算法上进行改进.

文献[53]分别在 2 个数据集上测试了基于 bigram 特征的 logistic 回归、CNN 和 CNN-LSTM 这 3 种模型的检测效果，实验证明深度学习能够自动提取较好的特征表示，检测效果优于传统机器学习方案. 2020 年，Wei 等人[54]提出一个基于 CNN 的钓鱼 URL 检测方案，将 URL 进行独热编码，通过嵌入层映射为低维稠密向量并输入到 CNN 中进行分类，检测精度高达 99.68%. 相较于 LSTM，基于 CNN 的方案训练时长压缩近 50%，检测模型体积小速度快，支持应用于内存和计算能力有限的移动设备. 对 URL 直接进行独热编码处理会得到高维稀疏向量，不利于高层语义信息的获取. 一些方案尝试对待检测 URL 进行分词处理，典型方案有 2021 年卜佑军等人[55]提出一种基于敏感词的 URL 分词方案，成分利用 URL 数据信息，并提出一种基于 CNN-BiLSTM 的检测方案，可以同时获取 URL 的空间特征和长距离依赖特征，捕获更全面的语义信息. 此外，Ozcan 等人[56]提出一种新颖的混合深度学习模型，结合基于 URL 词法分析的 DNN 模型和 LSTM 模型的强大功能，同时支持人工设计和字符嵌入 2 类特征的输入.

近年来，部分工作研究钓鱼 URL 的自动生成方案，用于改善现有检测方案的性能. Anand 等人[57]将基于字符的 LSTM 作为 GAN 的基础结构，生成对抗 URL 样本，用于解决典型数据集中存在的类不平衡问题. 此外，Burns 等人[58]在 OpenPhish，PhishTank，DNS-BH 等黑名单上训练 GAN，使用对抗 URL 样本扩展训练集，增强模型的泛化能力. 实验证实数据增强模型始终比原始分类器的检测精度高.

2.3 面向域名抢注的域名滥用检测

域名抢注（cybersquatting）是域名滥用的一种常见方式，攻击者通过注册与权威域名相似的抢注域名，骗取点击流量或开展网络诈骗活动. 根据 Proofpoint 公布的域名欺诈报告[101]显示，2018 年的第 1 季度和第 4 季度之间，抢注域名的季度注册量增加 11%. 此外，在 Proofpoint 提供数字风险保护的客户中有 76%的公司存在抢注域名. 本节首先简介抢注域名的分类，随后梳理各类抢注域名的检测方案.

2.3.1 抢注域名的分类

抢注域名可划分为 5 个大类，包括误植域名、比特抢注、同音抢注、同形抢注和组合抢注[73]. 下表以‘www.example.com’为例，给出抢注域名的示例.

Table 8 Examples of Cybersquatting Domain Name

表 8 抢注域名示例

抢注域名	类型
www(.)example.com	误植抢注——域名分隔符缺失
www.ex(a)mple.com	误植抢注——字符缺失
www.ex**ma**ple.com	误植抢注——字符排列错误
www.ex**s**mple.com	误植抢注——字符替换
www.exa**a**mple.com	误植抢注——字符插入
www.exa**o**ple.com	比特抢注
www.idle(idol).com	同音抢注
www.exampIe.com	同形抢注

www.examplelogin.com	组合抢注

1) 误植抢注（typosquatting），指攻击者注册常见键盘误操作导致的错误域名. 误植域名可以借助原域名的影响力获取大量的偶然流量，攻击者通过发布广告、网络钓鱼获取不法收益. Wang 等人[102]给出较为全面的误植域名构造方式，如表 8 所示. 当前误植抢注现象仍然较普遍，也是当前研究的热点.

2) 比特抢注（bitsquatting），如表 8 所示（将‘m’最后 1b 翻转，可得到‘o’），利用环境或制造缺陷导致的计算机内存随机单比特位翻转错误. 最早于 2011 年，安全研究人员 Dinaburg[64]在黑帽大会上介绍了比特抢注攻击，他针对 8 个合法域的 31 个比特抢注域进行为期 8 个月的监测，统计到 12949 个唯一 IP 地址的共 52317 次比特抢注域请求. 实验结果证明比特抢注域需要引起安全人员重视.

3) 同音抢注（soundsquatting），用同音字符构造的抢注域名，依赖于发音相似的字母或字符串可能彼此混淆的假设，典型例子如表 8 中示意的同音词组[idle，idol]. 最早于 2014 年由 Nikiforakis 等人[66]提出，他们将同音抢注定义为字典词的同音替换，从而与误植抢注区分开来. 概括来说，同音抢注不依赖于键入错误，而且并非所有域都可以被同音抢注.

4) 同形抢注（homograph），用同形字符构造的抢注域名，依赖于视觉相似的字母或字符串可能彼此混淆的假设，典型如表 8 中示意的小写字母'l'和大写字母'I'，在 san-serif 字体下，2 个字母看起来非常相似. 此外，随着国际化域名（internationalized domain names, IDN）的广泛使用，不同文字体系下的字母存在视觉一致性，例如，西里尔字母和拉丁字母中的“o”虽然是同形的，却具有不同的字符编码，比一般的近形字符更具迷惑性[67].

5) 组合抢注（combosquatting）通过向域名添加几个短语构造抢注域名（表 8 中示例，在‘example’后追加‘login’）. 从理论上分析，组合抢注域的构造空间是无限的，这是由于可拼接的短语是不可枚举的. 此外，组合抢注域最大程度地利用了原始域名的声誉：首先，组合抢注域名不会破坏被抢注商标在字符结构层面的完整性；其次，一些企业会主动注册组合域名，以扩大其服务范围，使得良性的组合域和滥用的组合抢注域难以区分. 上述 2 个原因使得组合抢注域较其他 4 种抢注域名更加难以进行判定[72].

2.3.2 抢注域名的检测方案

当前学术界对于抢注域名的检测，主要是将 Alexa 排名靠前的站点作为种子，根据不同类型的抢注域名构造方案，生成抢注域名的候选集合. 进一步通过主动探测或被动追踪的方案，确定是否为抢注域名. 表 9 总结典型检测工作在面向域名抢注的域名滥用检测领域的贡献及其局限性.

Table 9 Comparison of Typical Domain Abuse Detection Works Oriented to Cybersquatting

表 9 面向域名抢注的典型域名滥用检测工作的比较

检测方法	检测方案	文献	检测特征	检测算法	优点	缺点
误植抢注检测	主动探测方案	[61]	误植域名构造特征，辅助信息特征	基于规则的检测、fastcluster 聚类	对候选域名进行合法/滥用判定，并进一步划分滥用类型	对初始聚类进行人工分析，检测成本较高
	被动检测方案	[63]	误植域名构造特征	决策树、KNN	特征集较小且计算复杂度低	人工设计特征易被攻击者规避
比特抢注检测	主动探测方案	[65]	字符比特翻转特征	基于规则的检测	实验周期长，有效分析比特抢注域的增长性	未对探测到的域名进行合法/滥用判定
同音抢注检测	主动探测方案	[66]	同音字符替换特征，辅助信息特征	基于规则的检测	对同音抢注进行系统研究，并分析同音抢注域的滥用情况	需要人工辅助分类域名，检测成本较高
同形抢注检测	主动与被动检测方案	[67]	同形字符替换特征	基于规则的检测	首次混合使用多种检测方案对同形抢注进行系统研究	未对探测到的域名进行合法/滥用判定
	被动检测方案	[69]	同形字符替换特征	基于规则的检测	依据托管页面内容，对滥用域进一步划分滥用类型	采用半自动化方案检测抢注域，成本较高
组合抢注检测	主动与被动检测方案	[72]	组合抢注构造特征	基于规则的检测	对组合抢注攻击进行系统研究	人工抽样域名托管页面进行滥用行为划分，

					检测成本高	
多类抢注检测	主动与被动检测方案	[74]	误植域名构造特征，组合抢注构造特征	LSTM，Skipgram	采用表示学习方案生成置信度高的抢注域名，减少检测成本	模型结构较为复杂，可解释性较差

1）误植抢注的检测

误植抢注的检测方案可划分为两大类：一类是基于编辑距离的主动探测方案；另一类是基于编辑距离、时间相关性以及词汇相似性的被动检测方案.

① 主动探测方案：除前文 Wang 等人[102]提出的 5 种误植域名构造方式，文献[60]提出了相似的方案，他们主要关注基于编辑距离的 3 种典型构造方案：1B 替换、1B 增加和 1B 删减. 该篇工作选取 3 个顶级域名（.com，.org，.biz）下的 900 个知名站点作为原始域名集合，进行误植域构造，最终得到 300 万个原始域名变体的候选集. 通过主动探测确定原始集合中超过半数的域名，有 35%以上的变体存在于网络中，表明现存误植抢注攻击的规模较大.

此外，文献[61]采用 Wang 等人[102]的构造方案针对 500 个流行站点，共生成大小为 28179 的误植域名候选集，通过聚类算法结合人工分析的方案，对候选集中的域名进行合法和滥用性质的判别，将品牌商主动注册的域名标注为合法，将存在投放广告、实施钓鱼等滥用行为的域名标注为误植域名. 实验数据显示超过 79%的候选域名存在滥用行为，其中投放广告的误植域名比例最高.

② 被动检测方案：基于编辑距离的主动探测方案存在误报率较高的问题，例如 nhl.com 和 nfl.com 都是良性域，且内容都与体育相关，在基于编辑距离的主动探测中，非常有可能会被划分为彼此的误植域名. 考虑到上述方案的局限性，Khan 等人[59]提出基于条件概率模型的被动检测方案，使用被动收集的 DNS 和 HTTP 流量数据挖掘误植域名，主要关注用户访问某些域名后短时间内跳出，随后访问名称相似（域名的编辑距离为 1）的更具知名度的域名的流量，从而精准获取符合误植抢注定义的域名，过滤掉良性的相似域名. 此外，考虑到基于编辑距离的方案不能充分关联域名的上下文信息的问题，Ya 等人[62]提出了 TypoEval，利用暹罗神经网络来学习每个域的词嵌入，进一步使用 RNN 充分利用域名的上下文信息，并通过计算欧几里得空间中向量之间的距离来评估误植域，取得较好的分类效果. 考虑到 TypoEval 模型的复杂性，Moubayed 等人[63]提出一种集成的特征选择和分类模型来有效鉴别误植抢注域名. 实验结果证明，该方案在保证较高检测精度的同时，特征集大小缩减了 50%，且具有较低的计算复杂度.

2）比特抢注的检测

除 Dinaburg[64]的工作外，Nikiforakis 等人[65]也对比特抢注域进行实验测量，具体针对 Alexa 排名前 500 的站点执行任一比特的翻转，生成比特抢注域的候选集，通过爬虫程序尝试解析候选域名集中的 IP 地址. 在 9 个月的实验中，共记录了 5366 个比特抢注域名，相比于实验第 1 天的解析记录增加了 46%，表明比特抢注已经逐渐成为域名抢注者的攻击手段.

3）同音抢注的检测

最经典的检测方案是 Nikiforakis 等人[66]针对 Alexa 排名前 10000 个站点生成 8476 个同音域名，通过主动探测和 whois 信息查询确定已注册 1823 个同音域（占总域的 21.5%），并对已注册同音域进行分类，整体划分为权威域和抢注域，进一步按照滥用行为细分同音抢注域. 实验结果表明，用于投放广告的滥用域占现有同音抢注域的最大部分，共有 954 个（占已注册同音域的 52.3%）.

4）同形抢注的检测

早期主要是浏览器供应商和域名注册商主导研究，重点探讨缓解同形抢注的方案. 2006 年 Holgers 等人[67]首次采用被动网络跟踪和主动 DNS 探测的方案，量化同形抢注攻击的危害程度. 他们追踪科研机构的网站访问记录，在为期 9 天的实验中没有实际检测到同形抢注域的访问，一定程度上表明同形抢注在 2006 年的危害程度较低. 进一步通过 DNS 探测发现 Alexa 排名前 500 个站点的 399 个同形注册域名，近 60%的站点具有一个或多个同形抢注域，表明同形抢注域仍有较大的研究价值.

随着 IDN 的广泛使用，后续的工作主要针对同形 IDN 域名进行研究，Sawabe 等人[68]利用光学字符识别（optical character recognition, OCR）技术进行同形 IDN 的检测，使用目标 IDN 和流行域名列表作为输入，通过将目标 IDN 转换为图像，应用 OCR 以检测该 IDN 与流行域名的相似程度，从而实现同形 IDN 域名的判定. 文献[69]主要针对技术公司和金融机构的域名进行了为期 8 个月的追踪分析，检测到 2984 个 IDN 同形候选域，并通过半自动化的方案对候选域进一步标注，图 10 给出生成候选域流程示意：

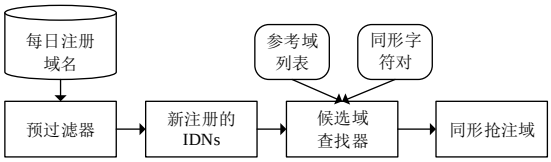


Fig. 10 Schematic diagram of the process of generating IDN homograph candidate domains

图 10 生成 IDN 同形候选域流程示意

考虑到基于编辑距离、OCR 技术的同形抢注域名的检测方案会出现较多的误报，Yu 等人[70]提出一种具有最小散列（MinHash）和局部敏感散列（locality-sensitive hashing, LSH）检索算法的双通道 CNN 分类器，进一步应用在实际的 DNS 数据中，实验证明 MinHash 和 LSH 算法在减少数据量方面表现优异．Thao 等人[71]则通过融合单字符结构相似性特征以及从 N-gram 模型中提取的 199 个语言特征，并采用集成学习算法，提升同形抢注域名的检测精度．

5）组合抢注的检测

一篇经典的工作是 Kintis 等人[72]发表的组合抢注域名的测量研究，首次针对组合抢注域进行系统性研究，通过分析从被动和主动 DNS 数据源收集的超过 4680 亿条 DNS 记录，检测到 270 万个针对 268 个流行商标的组合抢注域．并在此基础上，进一步分析组合抢注域在现实世界中的滥用情况，实验结果表明，60%的组合抢注域生命周期超过 1000 天，与组合抢注域相关的活恶意动逐年增加，需要引起商标拥有者和域名注册商的广泛重视．

6）多类抢注的检测

除了针对特定类型的抢注域名进行研究，Zeng 等人[73]提出了一个全面的抢注域名研究，他们选取 786 个流行域名，在 ISP 级的 DNS 流量中检测 5 类抢注域名．实验结果显示尽管误植抢注占比较大，但组合抢注却更容易获得访问流量．2020 年，Loyola 等人[74]提出一种面向误植抢注和组合抢注域名的主动防御方案．首先从检测到的抢注域名样本中学习其分布规律，然后自动生成相似的抢注域名候选域，通过主动探测进一步验证候选域的真实性，发掘尚未检测到的恶意抢注域名．此外，Hu 等人[75]将域名抢注的研究扩展到移动应用生态系统中，关注移动客户端市场中的抢注行为．

2.4 面向垃圾邮件的域名滥用检测

垃圾邮件泛指未经请求而发送的邮件，支撑的典型恶意行为包括：垃圾广告、网络钓鱼以及分发恶意附件等．根据 Cisco 的数据[103]，2020 年 9 月全球每日平均垃圾邮件数量达 2916.7 亿，占全部电子邮件的 84.61%．此外，Verizon 的《2019 年数据泄露调查报告》显示垃圾邮件是恶意软件分发的第一大媒介．

近年来，由于僵尸程序的成本低，相对容易传播且难以检测，垃圾邮件的分发逐渐迁移到僵尸网络上，将该类僵尸网络统称为垃圾邮件僵尸网络(spam botnet)．通过算法对现有垃圾邮件域和处于注册阶段的新域进行检测，可以破坏 spam botnet 的基础结构，有效阻断垃圾邮件的传播．

本文主要关注基于 DNS 的垃圾邮件域检测工作，通过监测异常的僵尸网络 DNS 流量或计算邮件域的信誉值等方案捕获垃圾邮件域．按照是否需要利用僵尸网络的 DNS 活动特征对现有检测工作进行划分，得到基于 spam botnet 的检测和基于注册信息的检测 2 类方案．表 10 总结典型检测工作在面向垃圾邮件的域名滥用检测领域的贡献及其局限性．

Table 10 Comparison of Typical Domain Abuse Detection Works Oriented to Spam

表 10 面向垃圾邮件的典型域名滥用检测工作的比较

检测方法	检测方案	文献	检测特征	检测算法	优点	缺点
基于 spam botnet 的检测	被动 DNS 分析方案	[82]	辅助信息特征	基于图推理	引入 DNSBL 查询特征检查僵尸网络	检测限制条件较高、无法确定僵尸主机所属网络架构
		[79]	域名解析时间特征，域名解析记录特征	多层感知机	特征提取较为简单，模型推理能力强，检测精度高	在训练样本集规模较小的情况下，难以保证检测精度
	主动 DNS 探测方案	[80]	域名解析记录特征	Adaboost 分类器	可以超前检测尚未投入使用的垃圾邮件域	攻击者可以规避检测特征，检测精确率较低
		[81]	域名解析记录特征，域名解析时间特征，域名解析空间特征	SVM	混合使用主被动数据，检测准确率、精确率较高	实验数据较少，未测试多种分类器检测效果
基于注册信息的检测	主动 DNS 探测方案	[84]	辅助信息特征	基于规则的检测	引入区域文件特征，具有超前检测的能力	检测数据敏感，影响检测算法可用性；存在检测时延
		[83]	辅助信息特征，域名语言特征，域名结构特征，	凸多面体机	注册时检测垃圾邮件域，无检测时延，误报率较低	检测特征计算复杂度高，需要多数据源信息，检测成本较高

域名相似度特征，

2.4.1 基于 spam botnet 的检测

基于 spam botnet 的检测方案主要通过分析网络流量以及攻击者行为等特征，实现 spam botnet 的僵尸主机以及 C&C 服务器域名的检测．该类方案的技术演进过程：早期主要使用被动 DNS 分析技术，实现 DNS 黑名单日志或监控网络等有限范围内的垃圾邮件域的检测；后期逐渐引入主动 DNS 查询数据，有效扩大垃圾邮件域的检测范围并缩小检测时延．

1）被动 DNS 分析方案

早期的被动 DNS 分析方案主要从 DNS 黑名单、被动 DNS 数据库及其他被动侦听到的数据中，抽取攻击者行为特征或 spam botnet 的异常 DNS 特征，识别有限范围内的垃圾邮件域．

较早的一篇工作是 2006 年 Ramachandran 等人[82]通过分析 DNS 黑名单（DNSBL）的查询记录，挖掘僵尸网络成员的方案．攻击者通常会执行 DNSBL 查找以确定僵尸主机是否被列入黑名单，通过分析 DNSBL 的日志记录，可以确定分发垃圾邮件的域及其规模．该方案的缺点在于不能确定发现的僵尸主机是否属于同一僵尸网络，此外限制条件较高，将攻击者执行 DNSBL 查找作为检测前提．

考虑到上述方案的局限性，文献[76]提出一种使用网络流数据和 DNS 元数据来检测垃圾邮件域的方法．他们使用被动 DNS 数据库来分析可疑垃圾邮件主机的流记录以及可疑控制器的 DNS 元数据，提取多种域名解析特征．经实验证明 DNS 元数据分析可以大幅提升检测精度．该方案的局限性在于攻击者可通过多种方案，诸如合法化信誉和流量统计信息，或更改通信协议来规避检测特征．此外，该检测方案需要综合构造多类检测特征，检测成本较高．

Vlaszaty 等人[77]提出一种简化方案：仅通过分析 DNS MX（mail exchange）请求，实现监控网络中垃圾邮件僵尸主机的识别．他们分别通过分析 DNS MX 的请求频率、请求周期性、请求熵、以及宿主机活跃时段内 MX 流量 4 种方案捕获垃圾邮件域．实验结果证明仅通过分析 DNS MX 请求可以识别感染垃圾邮件恶意软件的主机．然而这项研究的数据集中，仅有一个已确认的 spam botnet 主机，有必要在更大的数据集上进一步评估该方案的有效性．2020 年，Yoshida 等人[78]关注到垃圾邮件受害者的异常 DNS 查询模式：一组 FQDN 经常被一个公共客户端查询，以及客户端出现重复查询失败的情况．他们进一步基于基数分析，提取客户端可疑查询特征，识别多个客户端对不同域的访问模式，构建决策树分类模型，实现较低的漏报率和较高的检测精度．

考虑到前述传统机器学习方案对检测特征的依赖性较强，需要人工设计精确有效的检测特征．此外，传统机器学习方案，例如贝叶斯分类器和决策树，只能进行浅层次的模型推理，难以拟合现实情况中复杂多变的 spam botnet 的行为特点．2020 年，Sharma 等人[79]首次引入深度学习方案，训练基于前馈多层感知机的垃圾邮件域名检测模型．他们基于电子邮件日志和权威 DNS 记录抽取检测特征，输入到多层感知机中训练分类器模型，最终检测准确率高达 97%．

被动分析方案的优点在于可以获取监控网络内部真实的 DNS 流量，支持监控网络内部的垃圾邮件域的挖掘；其局限性在于检测范围有限，且需要借助 DNS 活动特征进行检测，存在检测时延．

2）主动 DNS 探测方案

spam botnet 的规模一般较大，被动 DNS 分析的有限检测范围，无法满足大型垃圾邮件域的检测需要，部分工作逐渐引入主动 DNS 探测方案，有针对性地发起 DNS 查询，实现垃圾邮件域的检测．

2018 年 vanderToorn 等人[80]首次使用主动 DNS 测量来检测垃圾邮件域．整体检测框架如图 11 所示，他们从 OpenINTEL 平台获取全球 60%以上已注册域名的资源记录，通过长尾分析提取候选域，进一步使用分类器对候选域进行二分类，将垃圾邮件域添加到实时黑名单列表（RBL）中．实验证明该检测方案具有超前检测未知垃圾邮件域的能力．该方案的缺点在于长尾分析会导致部分垃圾邮件域的漏检，同时也为攻击者规避检测制造机会．

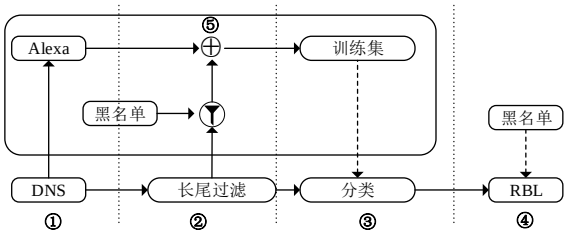


Fig. 11 Block diagram of spam domain detection

图 11 垃圾邮件域检测框图

Dan 等人[81]考虑到仅使用主动 DNS 数据造成的检测精度不高的问题，提出一种混合使用主被动数据的检测方案．利用从电子邮件接收日志和主动 DNS 数据中提取的特征，使用 SVM 分类器对候选域进行二分类，实验可达到 88.09%的准确率和 97.11%精确率，高于前述方法．此外，他们对重要性排名前 10 的特征进行分析，发现有 5 个是从电子邮件接收日志中提取的特征，进一步证明混合使用主被动数据对垃圾邮件域检测具有较高的研究价值．

主动 DNS 探测方案的优点在于可以有针对性地获取特定域的动态解析或资源记录信息，支持大规模垃圾邮件域的检测，可以提前发现尚未进行垃圾邮件活动的滥用域；其缺点在于无法获得完整的通信流量，造成检测精确度的损失，此外，需要预先构造候选域集合进行检测，容易导致漏检.

2.4.2 基于注册信息的检测

与前述方案不同，基于注册信息的检测不使用通信流量特征，而是通过主动 DNS 探测的手段，获取域名的 whois 信息或区域文件中有关域名注册的信息，以及与已知滥用域的关联，构造多维判别特征，实现垃圾邮件域的检测.

Felegyhazi 等人[84]使用来自注册服务商和注册管理机构的信息以及公共黑名单中的滥用域，来推断未被列入黑名单的垃圾邮件域. 实验表明该检测方案可以比研究中使用的黑名单更早地标记垃圾邮件域. 其局限性在于区域文件和 whois 数据库并不完全开放，或不支持批量访问，使得检测方案的可用性受到较大影响. 此外，该检测方案需要预先观察到一些相关域的滥用行为，这意味着该方案仅能减小域名滥用的时间窗口，仍存在检测时延.

考虑到上述方案的局限性，文献[83]提出了一种名为 PREDATOR 的无时延的垃圾邮件域检测器，可以利用域名注册时的特征建立域名信誉，实现潜在垃圾邮件域的检测. 他们具体关注滥用域名的异常的注册行为（例如，突发注册、文本相似的名称），从域名注册信息中抽取并编码 22 种有助于区分垃圾邮件域名与合法域名注册行为的特征，使用凸多面体机分类待检测域. 实验证明 PREDATOR 可以达到 70%的精确率，而误报率仅为 0.35%，可以有效阻止 DNS 域滥用. 该方案的局限性在于特征计算较为复杂，需要多种信息源构造检测特征.

2.5 面向不限定滥用行为的域名滥用检测

在实际的工作中，只有极少数黑名单报告特定类型的滥用域，大部分黑名单不区分具体滥用行为；此外，大部分滥用域的基础架构相似，滥用行为也不唯一，例如僵尸网络可同时用于分发垃圾邮件或布置钓鱼站点. 基于上述原因考虑，部分工作使用与滥用行为无关的方法检测滥用域名，基于域之间不同类型的关联，实现滥用域的检测，这种技术也被称为“关联有罪（guilty by association）”，按照具体的关联方案可划分为信誉评分和图推理 2 类检测方法. 表 11 总结典型检测工作在面向不限定恶意行为的域名滥用检测领域的贡献及其局限性.

2.5.1 基于信誉评分的检测

基于信誉评分的检测方案通过为参与恶意活动（例如，恶意软件传播、网络钓鱼和垃圾邮件活动）的域分配低信誉分数，以实现滥用域的检测.

最早于 2005 年 Weimer[104]发表了一篇基础性工作，提出使用被动 DNS 记录来检测多种滥用域的观点. 文献[85]率先使用被动 DNS 数据，构建名为 Notos 的 DNS 动态信誉系统，从被动数据中提取域名字符级特征和解析特征，并通过已有威胁情报提取证据特征，采用离线训练-在线检测的模型架构，离线训练信誉评分系统，在线计算新域名的信誉得分，实现滥用域检测. 文献[9]进一步提出 EXPOSURE 系统，从大规模被动 DNS 记录中提取域名解析和字符级特征，构建 J48 决策树检测滥用域名. 实验结果表明，

Table 11 Comparison of Typical Domain Abuse Detection Works Oriented to Unrestricted Abuse Behaviors

表 11 面向不限定滥用行为的典型域名滥用检测工作的比较

检测方法	检测方案	文献	检测特征	检测算法	优点	缺点
基于信誉评分	被动 DNS 分析方案	[85]	域名解析记录特征，域名解析空间特征，域名统计特征，域名结构特征，辅助信息特征	J48 决策树	引入动态 DNS 信誉评分机制	需要大量历史记录为域名分配信誉分数，离线训练时间较长
		[9]	域名解析时间特征，域名解析记录特征，域名语言特征	J48 决策树	所需的训练时间和训练数据更少	检测特征可被规避，无法检测复杂攻击中查询历史少的域名
基于图推理	主机-域二部图	[90]	主机-域名关联特征	信念传播	无需显示特征计算	使用敏感的主机访问域信息，处理脱敏被动 DNS 数据准确率不佳
	域-IP 二部图	[91]	域名-IP 关联特征	基于路径的推理	使用被动 DNS 数据构建，剔除敏感信息	使用公共 IP 池托管的不相关域易被误分类为滥用域，误报率较高
	主机-域-IP 异构图	[88]	域名相似度特征，域名解析记录特征，主机-域名关联特征，	直推分类	仅需少量标注数据，可达到较高检测精度	消耗大量计算资源，检测成本较高

域名-IP 关联特征

EXPOSURE 可以自动识别未知的滥用域，比 Notos 所需的训练时间和训练数据更少，具有更高的检测精度.

2.5.2 基于图推理的检测

信誉评分的可靠性严重依赖于检测特征提取的有效性，需要专家信息作为指导，同时需要不断动态调整检测特征以适应于复杂多变的域名滥用技术. 考虑到上述局限性，部分专家学者从图结构中受到启发，尝试通过对域名进行关联，根据已知结点推导出未知结点的性质，被称为基于图推理的检测方案. 这类检测方案主要关注域名的结构关联信息，通过主机查询域名或域名解析 IP 构建有效的域名关联. 具体而言，基于图推理的检测倾向于从 DNS 流量中提取图模型，利用全局关联挖掘潜在的滥用域，按照图模型中使用的结点可划分为：主机-域二部图、域-IP 二部图以及主机-域-IP 异构图 3 大类.

1）主机-域二部图

文献[90]将滥用域名检测问题建模为图推理问题，通过分析 DNS 查询日志，构建主机-域的二部图（如图 12 所示），进一步应用信念传播评估图中未知域为滥用域的边际概率，实现滥用域的检测. 该方案需要使用主机查询域的访问请求信息，涉及到用户隐私无法大规模部署.

2）域-IP 二部图

考虑资源有限情况下，攻击者会进行资源重用，Khalil 等人[91]从被动 DNS 数据中得提取域名-IP 解析关联特征，构建无向域解析二部图表示域和 IP 的解析关系，随后他们依据二部图构建一个无向加权域

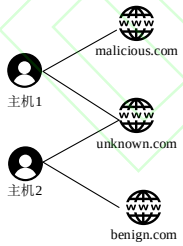


Fig. 12 Example of host-domain graph

图 12 主机-域二部图示例

图，结点是二部图中的域，边的权重由域解析到同一 IP 地址的次数决定，并基于与已知滥用域具有强关联的域很可能是恶意的假设，在关联域上使用基于图推理技术，从而实现滥用域的检测，如图 13 所示：

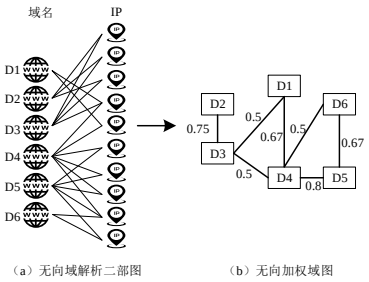


Fig. 13 Example of domain resolution graph and its corre-sponding domain graph

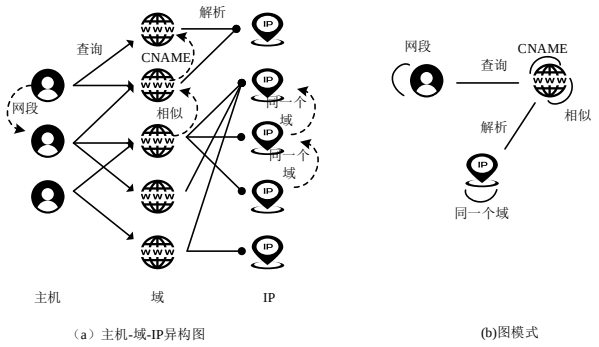
图 13 域-IP 二部图及相应域图示例

域解析为相同的 IP（co-IP 关联）是一种弱关联，没有考虑互联网中域名部署的复杂方式，尤其是公共网络托管和代理服务可能会导致毫不相关的域名被托管在在同一个 IP 池中，导致检测精度较低. 为了克服 co-IP 关联的弱点，2020 年 Nabeel 等人[86]通过构建分类器区分专用托管 IP 和公共托管 IP，并提出一种基于 IP 的域名强关联方案：1）共享一个专用 IP；或 2）共享来自不同托管服务提供商的多个公共 IP. 实验证明，该检测方案可以有效提升检测精度.

此外，Liang 等人[87]考虑到现有域图构建方案无法处理孤立域名结点，提出一种结合单个域名特征（域名字符级特征、域名解析记录特征）和域名-IP 关联特征的检测方案 MalPortrait. MalPortrait 相较于普通方案增加全局关联信息，弥补数据缺乏导致域名信息量不足的缺点，且对被动数据时间跨度的依赖性较小. 对比其他基于域图的方案，MalPortrait 可以有效解决孤立域名结点难以分类的问题.

3）主机-域-IP 异构图

Sun 等人[88]于 2019 年提出的 HinDom 系统，将 DNS 场景建模为一个异构信息网络（HIN），该网络由主机、域、IP 地址及他们之间的 6 种关系组成，HinDom 采用基于元路径的直推分类，能够仅使用一小部分标记样本来检测滥用域，图 16 给出异构图的示例（图 14（a））以及图模式说明（图 14（b））. 将 HinDom 系统与朴素贝叶斯、支持向量机和随机森林 3 种归纳分类方法进行比较，实验证明直推分类在初始标签信息的比例降低时，仍能保持较为稳定的检测精度，较归纳分类方案性能优异.



（a）主机-域-IP异构图 (b)图模式

Fig. 14 Example of host-domain-IP HIN and graph schema

图 14 主机-域-IP 异构图示例及图模式

考虑到 HinDom 系统仅使用全局关联信息，而忽略对单个域名结点的属性特征提取的问题，Liu 等人[89]提出一种基于动态图卷积网络（dynamic graph convolutional network，DGCN）的检测方案 Ringer. Ringer 充分利用主机-域名的查询关联和域名-IP 的解析关联构建域图，并采用 DGCN 方案在域图上学习集成结点特征和图结构信息的结点表示，送入全连接神经网络中进行域名分类. 实验证实，在服务提供商的真实 DNS 数据上，Ringer 在检测精度和检测速度上都有优异的表现.

3 讨论与展望

近年来涌现出大量工作旨在阻止或消除域名滥用行为，对改善 DNS 系统的生态环境做出了明显的贡献，但当前的检测能力与互联网上的域名滥用情况之间仍存在差距. 本节将进一步讨论多检测场景下域名滥用行为检测研究面临的挑战与未来的研究方向：

1）研究高动态强对抗场景下的检测技术，应对滥用技术的演进.

攻击者为提高滥用域名的可用性和对黑名单的抵抗能力，各检测场景下的域名滥用技术均呈现出高度动态发展的趋势，旨在规避现有检测技术. 攻防博弈日趋激烈，呈现出强对抗的形态. 当前各场景下的域名滥用检测难点主要包括：面向恶意软件的域名滥用检测场景下，攻击者提出基于字典词构造低随机性、高可读性域名的方案，并引入对抗样本生成技术规避检测[31-32, 38-39, 95-97]；在面向网络钓鱼的域名滥用检测场景下，攻击者尝试多种高对抗性的 URL 混淆技术规避检测[57-58]，或通过 Blackhat SEO 技术提高钓鱼链接在搜索引擎结果中的页面排名[10]；在面向域名抢注的域名滥用检测场景下，攻击者针对国际化域名进行同形抢注攻击，浏览器防御规则或一些 OCR 技术无法抵御这类抢注攻击[68-70]；在面向垃圾邮件的域名滥用检测场景下，攻击者使用不同的注册模式，并通过匿名注册服务提供商模糊其注册信息，这类混淆方案会对基于注册信息的检测造成较大影响[83-84]；在面向不限定滥用行为的域名滥用检测场景下，对于在大量通信流量中潜伏期较长、少通信关联的滥用域尚缺乏检测能力[86-88, 91]. 此外，考虑到未来 DNS 通信技术的发展，恶意攻击者的攻击目标和手法不断变化，对传统检测和防护方案提出新挑战，域名滥用行为检测需要引入新的技术方案.

2）研究跨检测场景的域名滥用检测技术，应对检测场景的融合.

随着网络恶意攻击的日趋体系化，域名作为重要的支撑资源，通常会贯穿整个恶意活动的始终，呈现出域名滥用行为检测场景融合的趋势. 在载荷投递、安装植入和命令与控制阶段，都可以通过多类域名滥用，达成实施攻击或建立通信的恶意目标. 例如，一个鱼叉式钓鱼攻击活动，可能涉及垃圾邮件域、网络钓鱼域、恶意资源下载域，以及恶意软件 C&C 通信域的协同. 面对跨检测场景的域名滥用行为，如何有效针对多种滥用行为建立相对统一的检测模型，或根据一些捕获的滥用域，迅速定位其他潜在域名滥用行为，提出行之有效的复合性检测方案，是学术界和工业界亟待解决的问题.

3）研究新基建场景的域名滥用检测技术，应对滥用模式的迁移.

随着国家对于新基建的日益重视，域名 DNS 作为一类基础设施已经部署到各种新基建场景中，同时也面临恶意软件、钓鱼站点、域名抢注、垃圾邮件等各类域名滥用行为的威胁. 域名滥用行为检测场景应全方位覆盖移动端、物联网终端、云平台等基础设施. 同时，研究人员需要针对新基建下的滥用行为特点、不同网络环境中的 DNS 通信流量特征、计算资源分配情况有针对性地调整现有域名检测技术. 目前已有检测方案通过捕获 DNS 流量异常检测基于物联网的僵尸网络[105]、移动终端的恶意软件[106]；此外，还有一些解决方案关注软件定义网络（software defined network, SDN）架构下基于 DNS 的拒绝服务攻击防御、检测和缓解技术[107-108].

4 总结

随着互联网上恶意活动的兴起，网络攻击者将 DNS 系统纳入其攻击活动的支撑资源，域名滥用行为已经得到学术界和工业界的广泛关注. 本文对域名系统的工作方式及其扩展技术进行概述，并从具体行为出发，构建本文的域名滥用检测场景分类体系，梳

理恶意软件、网络钓鱼、抢注域名、垃圾邮件以及不限定滥用行为 5 种典型场景下的域名滥用检测工作，重点阐述了各类域名滥用检测方案的演进过程，为域名滥用行为检测提供了一个以检测场景为导向的总结性工作.

参 考 文 献

[1] Stevanovic M, Pedersen J M, D'Alconzo A, et al. On the ground truth problem of malicious DNS traffic analysis [J]. Computers & Security, 2015, 55: 142-158

[2] Korczynski M, Wullink M, Tajalizadehkhoob S, et al. Cybercrime after the sunrise: A statistical analysis of DNS abuse in new gTLDs [C] // Proc of the 13th Asia Conf on Computer and Communications Security. New York: ACM, 2018: 609-623

[3] Internet Corporation for Assigned Names and Numbers. Domain abuse activity reporting [EB/OL]. [2021-02-03]. https://www.icann.org/octo-ssr/daar

[4] Public Interest Registry. Framework to address abuse [EB/OL]. (2019-12-06)[2021-02-03]. http://dnsabuseframework.org/media/files/2020-05-29_DNSAbuseFramework.pdf

[5] Zhauniarovich Y, Khalil I, Yu T, et al. A survey on malicious domains detection through DNS data analysis [J]. ACM Computing Surveys, 2018, 51(4): 1-36

[6] Yan Yida, Liu Zhenyan, Zhong Junwei, et al. Malicious domain detection based on machine learning [C/OL] // Proc of the 2017 Int Conf on Electronic and Information Technology. Lancaster, PA: DEStech, 2017 [2021-02-03]. http://dx.doi.org/10.12783/dtcse/iceit2017/19866

[7] Wang Yuanyuan, Wu Chunjiang, Liu Qihe, et al. Overview of malicious domain name detection and application [J]. Computer Applications and Software, 2019, 36(9): 310-316(in Chinese)
（王媛媛，吴春江，刘启和，等. 恶意域名检测研究与应用综述[J]. 计算机应用与软件, 2019, 36(9): 310-316）

[8] Cheng Du, Liu Zhenyan, Zhang Pengfei, et al. Profiling malicious domain by multidimensional features [C] // Proc of the 2018 Int Conf on Robots & Intelligent System. Piscataway, NJ: IEEE, 2018: 489-495

[9] Bilge L, Kirda E, Kruegel C, et al. EXPOSURE: Finding malicious domains using passive DNS analysis [C/OL] // Proc of the 18th Annual Network and Distributed System Security Symp (NDSS). Reston, VA: The Internet Society, 2011 [2021-02-03]. https://www.ndss-symposium.org/wp-content/uploads/2017/09/bilg.pdf

[10] Moura G C M, Müller M, Davids M, et al. Domain names abuse and tlds: From monetization towards mitigation [C] // Proc of the 15th IFIP/IEEE Symp on Integrated Network and Service Management (IM). Piscataway, NJ: IEEE, 2017: 1077-1082

[11] Chiba D, Akiyama M, Yagi T, et al. DomainChroma: Building actionable threat intelligence from malicious domain names [J]. Computers & Security, 2018, 77: 138-161

[12] Chen Zhenhao. Newly registered domains: malicious abuse by bad actors [EB/OL]. (2019-08-20)[2021-02-03]. https://unit42.paloaltonetworks.com/newly-registered-domains-malicious-abuse-by-bad-actors/

[13] Stalmans E, Irwin B. A framework for DNS based detection and mitigation of malware infections on a network [C/OL] // Proc of the 2011 Information Security for South Africa. Piscataway, NJ: IEEE, 2011 [2021-02-03]. https://ieeexplore.ieee.org/document/6027531

[14] Zhou Changling, Chen Kai, Gong Xiaoxu, et al. Detection of fast-flux domains based on passive DNS analysis [J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2016, 52(3): 396-402(in Chinese)
（周昌令，陈恺，公绪晓，等. 基于 Passive DNS 的速变域名检测[J]. 北京大学学报：自然科学版, 2016, 52(3): 396-402）

[15] Chen Xunxun, Li Gaochao, Zhang Yongzheng, et al. A deep learning based fast-flux and CDN domain names recognition method [C] // Proc of the 2nd Int Conf on Information Science and Systems. New York: ACM, 2019: 54-59

[16] Holz T, Gorecki C, Rieck K, et al. Measuring and detecting fast-flux service networks [C/OL] // Proc of the 16th Annual Network and Distributed System Security Symp (NDSS). Reston, VA: The Internet Society, 2008 [2021-02-03]. https://www.ndss-symposium.org/ndss2008/measuring-and-detecting-fast-flux-service-networks/

[17] Caglayan A, Toothaker M, Drapeau D, et al. Real-time detection of fast flux service networks [C] // Proc of the 2009 Cybersecurity Applications & Technology Conf for Homeland Security. Piscataway, NJ: IEEE, 2009: 285-292

[18] Huang Siyu, Mao Chinghao, Lee Hahnming. Fast-flux service network detection based on spatial snapshot mechanism for delay-free detection [C] // Proc of the 5th ACM Symp on Information, Computer and Communications Security. New York: ACM, 2010: 101-111

[19] Yang Lu, Gan Gang. Research and detection of fast-flux botnet [J]. IOP Conference Series: Earth and Environmental Science, 2021, 693: 012031

[20] Niu Weina, Jiang Tianyu, Zhang Xiaosong. Fast-flux botnet detection method based on spatiotemporal feature of network traffic [J]. Journal of Electronics & Information Technology, 2020, 42(8):1872-1880(in Chinese)
（牛伟纳，蒋天宇，张小松，等. 基于流量时空特征的 fast-flux 僵尸网络检测方法[J]. 电子与信息学报, 2020, 42(8): 1872-1880）

[21] Almomani A. Fast-flux hunter: A system for filtering online fast-flux botnet [J]. Neural Computing and Applications, 2018, 29(7): 483-493

[22] Han Chunyu, Zhang Yongzheng, Zhang Yu. Fast-flucos: Malicious domain name detection method for fast-flux based on DNS traffic [J].

Journal on Communications, 2020, 41(5): 37-47(in Chinese)
（韩春雨, 张永铮, 张玉. Fast-flucos:基于 DNS 流量的 fast-flux 恶意域名检测方法[J]. 通信学报, 2020, 41(5): 37-47）

[23] Al-Duwairi B, Jarrah M, Shatnawi A S. PASSVM: A highly accurate fast flux detection system [J]. Computers & Security, 2021, 110: 102431

[24] Yadav S, Reddy A K K, Reddy A L N, et al. Detecting algorithmically generated malicious domain names [C] // Proc of the 10th ACM SIGCOMM Conf on Internet measurement. New York: ACM, 2010: 48-61

[25] Truong D T, Cheng Guang. Detecting domain-flux botnet based on DNS traffic features in managed network [J]. Security and Communication Networks, 2016, 9(14): 2338-2347

[26] Schiavoni S, Maggi F, Cavallaro L, et al. Phoenix: DGA-based botnet tracking and intelligence [C] // Proc of the 11th Int Conf on Detection of Intrusions and Malware, and Vulnerability Assessment. Berlin: Springer, 2014: 192-211

[27] Tong V, Nguyen G. A method for detecting DGA botnet based on semantic and cluster analysis [C] // Proc of the 7th Symp on Information and Communication Technology. New York: ACM, 2016: 272-277

[28] Antonakakis M, Perdisci R, Nadji Y, et al. From throw-away traffic to bots: Detecting the rise of DGA-based malware [C] // Proc of the 21st Usenix Security Symp. Berkeley, CA: USENIX Association, 2012: 491-506

[29] Woodbridge J, Anderson H S, Ahuja A, et al. Predicting domain generation algorithms with long short-term memory networks [J]. arXiv preprint, arXiv:1611.00791, 2016

[30] Yu Bin, Gray D L, Pan Jie, et al. Inline DGA detection with deep networks [C] // Proc of the 2017 Int Conf on Data Mining Workshops (ICDMW). Piscataway, NJ: IEEE, 2017: 683-692 （无届数）

[31] Curtin R R, Gardner A B, Grzonkowski S, et al. Detecting DGA domains with recurrent neural networks and side information [C/OL] // Proc of the 14th Int Conf on Availability, Reliability and Security. New York: ACM, 2019 [2021-02-03]. https://doi.org/10.1145/3339252.3339258

[32] Pereira M, Coleman S, Yu B, et al. Dictionary extraction and detection of algorithmically generated domain names in passive DNS traffic [C] // Proc of the 21st Int Symp on Research in Attacks, Intrusions, and Defenses. Berlin: Springer, 2018: 295-314

[33] Tong Mingkai, Sun Xiaoqing, Yang Jiahai, et al. D3N: Dga detection with deep-learning through nxdomain [C] // Proc of the 12th Int Conf on Knowledge Science, Engineering and Management. Berlin: Springer, 2019: 464-471

[34] Tran D, Mac H, Tong V, et al. A LSTM based framework for handling multiclass imbalance in DGA botnet detection [J]. Neurocomputing, 2018, 275: 2401-2413

[35] Mohan V S, Vinayakumar R, Soman K P, et al. Spoofnet: Syntactic patterns for identification of ominous online factors [C] // Proc of the 2018 Security and Privacy Workshops (SPW). Piscataway, NJ: IEEE, 2018: 258-263

[36] Zhang Yongbin, Chang Wenxin, Sun Lianshan, et al. Detection method of domains generated by dictionary-based domain generation algorithm [J]. Journal of Computer Applications, 2021, 41(9): 2609-2614(in Chinese)
（张永斌, 常文欣, 孙连山, 等. 基于字典的域名生成算法生成的域名检测方法[J]. 计算机应用, 2021, 41(9): 2609-2614）

[37] Yang Luhui, Liu Guangjie, Dai Yuewei, et al. Detecting stealthy domain generation algorithms using heterogeneous deep neural network framework [J]. IEEE Access, 2020, 8: 82876-82889

[38] Anderson H S, Woodbridge J, Filar B. DeepDGA: Adversarially-tuned domain generation and detection [C] // Proc of the 9th ACM Workshop on Artificial Intelligence and Security. New York: ACM, 2016: 13-21

[39] Ravi V, Alazab M, Srinivasan S, et al. Adversarial defense: DGA-based botnets and DNS homographs detection through integrated deep learning [J/OL]. IEEE Transactions on Engineering Management, 2021 [2021-09-03]. https://ieeexplore.ieee.org/abstract/document/9377310

[40] Zhou Yonglin, Li Qingshan, Miao Qidi, et al. DGA-based botnet detection using DNS traffic [J]. Journal of Internet Services and Information Security, 2013, 3(3/4): 116-123

[41] Prakash P, Kumar M, Kompella R R, et al. Phishnet: predictive blacklisting to detect phishing attacks [C/OL] // Proc of the 29th IEEE INFOCOM. Piscataway, NJ: IEEE, 2010 [2021-02-03]. https://ieeexplore.ieee.org/document/5462216

[42] Lee Lunghao, Lee Kueiching, Chen Hsinhsi, et al. Poster: Proactive blacklist update for anti-phishing [C] // Proc of the 21st ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2014: 1448-1450

[43] Kang J M, Lee D H. Advanced white list approach for preventing access to phishing sites [C] // Proc of the 2007 Int Conf on Convergence Information Technology (ICCIT). Piscataway, NJ: IEEE, 2007: 491-496

[44] Cao Ye, Han Weili, Le Yueran. Anti-phishing based on automated individual white-list [C] // Proc of the 4th ACM Workshop on Digital Identity Management. New York: ACM, 2008: 51-60

[45] Azeez N, Misra S, Margaret I A, et al. Adopting automated whitelist approach for detecting phishing attacks [J]. Computers & Security, 2021, 108: 102328

[46] Dong Xun, Clark J A, Jacob J L. Defending the weakest link: Phishing websites detection by analysing user behaviours [J]. Telecommunication Systems, 2010, 45(2): 215-226

[47] Wang Yuqi, Liu Bowen, Lin Guoyuan. Phishing detection algorithm based on language features of URL [J]. Computer Engineering and Applications, 2019, 55(24): 84-90(in Chinese)

（王雨琪, 刘博文, 林果园. 基于 URL 语言特征的钓鱼网站检测算法[J]. 计算机工程与应用, 2019, 55(24): 84-90）

[48] Ma J, Saul L K, Savage S, et al. Beyond blacklists: Learning to detect malicious web sites from suspicious URLs [C] // Proc of the 15th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2009: 1245-1254

[49] Ma J, Saul L K, Savage S, et al. Identifying suspicious URLs: An application of large-scale online learning [C] // Proc of the 26th Annual Int Conf on Machine Learning. New York: ACM, 2009: 681-688

[50] Garera S, Provos N, Chew M, et al. A framework for detection and measurement of phishing attacks [C/OL] // Proc of the 5th ACM Workshop on Recurring Malcode. New York: ACM, 2007 [2021-02-03]. https://dl.acm.org/doi/10.1145/1314389.1314391

[51] Zouina M, Outtaj B. A novel lightweight URL phishing detection system using SVM and similarity index [J]. Human-Centric Computing and Information Sciences, 2017, 7(1): 1-13

[52] Yang Peng, Zeng Peng, Zhao Guangzhen, et al. Phishing website detection method basedon logistic regression and XGBoost [J]. Journal of Southeast University: Natural Science Edition, 2019, 49(2): 207-212(in Chinese)
（杨鹏, 曾朋, 赵广振, 等. 基于 Logistic 回归和 XGBoost 的钓鱼网站检测方法[J]. 东南大学学报：自然科学版, 2019, 49(2): 207-212）

[53] Vazhayil A, Vinayakumar R, Soman K P. Comparative study of the detection of malicious URLs using shallow and deep networks [C/OL] // Proc of the 9th Int Conf on Computing, Communication and Networking Technologies (ICCCNT). Piscataway, NJ: IEEE, 2018 [2021-02-03]. https://ieeexplore.ieee.org/document/8494159

[54] Wei Wei, Ke Qiao, Nowak J, et al. Accurate and fast URL phishing detector: a convolutional neural network approach [J]. Computer Networks, 2020, 178: 107275

[55] Bu Youjun, Zhang Qiao, Chen Bo, et al. Research on phishing URL detection technology based on CNN and BiLSTM [J/OL]. Journal of Zhengzhou University: Engineering Science, 2021 [2021-09-14]. https://doi.org/10.13705/j.issn.1671-6833.2021.04.022(in Chinese)
(卜佑军, 张桥, 陈博, 等. 基于 CNN 和 BiLSTM 的钓鱼 URL 检测技术研究[J/OL]. 郑州大学学报：工学版, 2021 [2021-09-14]. https://doi.org/10.13705/j.issn.1671-6833.2021.04.022)

[56] Ozcan A, Catal C, Donmez E, et al. A hybrid DNN-LSTM model for detecting phishing URLs [J/OL]. Neural Computing and Applications, 2021 [2021-09-03]. https://doi.org/10.1007/s00521-021-06401-z

[57] Anand A, Gorde K, Moniz J R A, et al. Phishing URL detection with oversampling based on text generative adversarial networks [C] // Proc of the 2018 IEEE Int Conf on Big Data (Big Data). Piscataway, NJ: IEEE, 2018: 1168-1177

[58] Burns J, Heath E. Using generative adversarial networks to harden phishing classifiers [EB/OL]. New Orleans, LA: FloCon, 2019 [2021-02-14]. https://flocon2019.sched.com/event/GXW1

[59] Khan M T, Huo Xiang, Li Zhou, et al. Every second counts: Quantifying the negative externalities of cybercrime via typosquatting [C] // Proc of the 36th IEEE Symp on Security and Privacy. Piscataway, NJ: IEEE, 2015: 135-150

[60] Banerjee A, Barman D, Faloutsos M, et al. Cyber-fraud is one typo away [C] // Proc of the 27th IEEE Int Conf on Computer Communications. Piscataway, NJ: IEEE, 2008: 1939-1947

[61] Agten P, Joosen W, Piessens F, et al. Seven months' worth of mistakes: A longitudinal study of typosquatting abuse [C/OL] // Proc of the 22nd Network and Distributed System Security Symp (NDSS). Reston, VA: The Internet Society, 2015 [2021-02-03]. https://www.ndss-symposium.org/wp-content/uploads/2017/09/01_3_1.pdf

[62] Ya Jing, Liu Tingwen, Li Quangang, et al. Fast and accurate typosquatting domains evaluation with Siamese networks [C] // Proc of the 2018 IEEE Military Communications Conf (MILCOM). Piscataway, NJ: IEEE, 2018: 58-63

[63] Moubayed A, Aqeeli E, Shami A . Ensemble-based feature selection and classification model for DNS typo-squatting detection [C/OL] // Proc of the 2020 IEEE Canadian Conf on Electrical and Computer Engineering (CCECE). Piscataway, NJ: IEEE, 2020 [2021-02-03]. https://ieeexplore.ieee.org/document/9255697

[64] Dinaburg A. Bitsquatting: DNS hijacking without exploitation [J/OL]. Proceedings of BlackHat Security, 2011: 136-148 [2021-03-02]. https://web.archive.org/web/20180713212603/http://media.blackhat.com/bh-us-11/Dinaburg/BH_US_11_Dinaburg_Bitsquatting_WP.pdf

[65] Nikiforakis N, Van Acker S, Meert W, et al. Bitsquatting: Exploiting bit-flips for fun, or profit [C] // Proc of the 22nd Int Conf on World Wide Web. New York: ACM, 2013: 989-998

[66] Nikiforakis N, Balduzzi M, Desmet L, et al. Soundsquatting: Uncovering the use of homophones in domain squatting [C] // Proc of the 17th Int Conf on Information Security. Berlin: Springer, 2014: 291-308

[67] Holgers T, Watson D E, Gribble S D. Cutting through the confusion: A measurement study of homograph attacks [C] // Proc of the 2006 USENIX Annual Technical Conf. (ATC). Berkeley, CA: USENIX Association, 2006: 261-266

[68] Sawabe Y, Chiba D, Akiyama M, et al. Detecting homograph IDNs using OCR [J]. Proceedings of the Asia-Pacific Advanced Network, 2018, 46: 56-64

[69] Quinkert F, Lauinger T, Robertson W, et al. It's not what it looks like: Measuring attacks and defensive registrations of homograph domains [C] // Proc of the 7th IEEE Conf on Communications and Network Security (CNS). Piscataway, NJ: IEEE, 2019: 259-267

[70] Yu Guangxi, Yang Xinghua, Zhang Yan, et al. Towards

homograph-confusable domain name detection using dual-channel CNN [C] // Proc of the 21st Int Conf on Information and Communications Security. Berlin: Springer, 2019: 555-568

[71] Thao T P, Nguyen-Son H Q, Yamaguchi R S, et al. Boosting homograph attack classification using ensemble learning and n-gram model [C] // Proc of the 19th IEEE Int Conf on Trust, Security and Privacy in Computing and Communications (TrustCom). Piscataway, NJ: IEEE, 2020: 1983-1988

[72] Kintis P, Miramirkhani N, Lever C, et al. Hiding in plain sight: A longitudinal study of combosquatting abuse [C] // Proc of the 24th ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2017: 569-586

[73] Zeng Yuwei, Zang Tianning, Zhang Yongzheng, et al. A comprehensive measurement study of domain-squatting abuse [C/OL] // Proc of the 53rd IEEE Int Conf on Communications (ICC). Piscataway, NJ: IEEE, 2019 [2021-03-02]. https://ieeexplore.ieee.org/document/8761388

[74] Loyola P, Gajananan K, Kitahara H, et al. Automating domain squatting detection using representation learning [C] // Proc of the 2020 IEEE Int Conf on Big Data (Big Data). Piscataway, NJ: IEEE, 2020: 1021-1030

[75] Hu Yangyu, Wang Haoyu, He Ren, et al. Mobile app squatting [C] // Proc of the Web Conf 2020. New York: ACM, 2020: 1727-1738

[76] Ehrlich W K, Karasaridis A, Hoeflin D A, et al. Detection of spam hosts and spam bots using network flow traffic modeling [C/OL] // Proc of the 3rd USENIX Conf on Large-Scale Exploits and Emergent Threats (LEET). Berkeley, CA: USENIX Association, 2010 [2021-02-03]. https://www.usenix.org/legacy/event/leet10/tech/full_papers/Ehrlich.pdf

[77] Vlaszaty B, Eyckelhof C J, Quarantainenet B V. Identifying spam malware infections, using DNS MX request analysis [EB/OL]. [2021-02-03]. https://www.os3.nl/_media/2013-2014/courses/rp2/p30_report.pdf

[78] Yoshida K, Fujiwara K, Sato A, et al. Cardinality analysis to classify malicious domain names [C] // Proc of the 44th IEEE Annual Computers, Software, and Applications Conf (COMPSAC). Piscataway, NJ: IEEE, 2020: 826-832

[79] Sharma C. Feed forward MLP spam domain detection using authoritative DNS records and email log [D]. Dublin: National College of Ireland, 2020

[80] vanderToorn O, van Rijswijk-Deij R, Geesink B, et al. Melting the snow: Using active DNS measurements to detect snowshoe spam domains [C/OL] // Proc of the 2018 IEEE/IFIP Network Operations and Management Symp. Piscataway, NJ: IEEE, 2018 [2021-02-03]. https://ieeexplore.ieee.org/document/8406222

[81] Dan K, Kitagawa N, Sakuraba S, et al. Spam domain detection method using active DNS data and e-mail reception log [C] // Proc of the 43rd IEEE Annual Computer Software and Applications Conf (COMPSAC). Piscataway, NJ: IEEE, 2019: 896-899

[82] Ramachandran A, Feamster N, Dagon D. Revealing botnet membership using dnsbl counter-intelligence [C] // Proc of the 2nd Workshop on Steps to Reducing Unwanted Traffic on the Internet (SRUTI). Berkeley, CA: USENIX Association, 2006: 49-54

[83] Hao S, Kantchelian A, Miller B, et al. PREDATOR: Proactive recognition and elimination of domain abuse at time-of-registration [C] // Proc of the 23rd ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2016: 1568-1579

[84] Felegyhazi M, Kreibich C, Paxson V. On the potential of proactive domain blacklisting [C/OL] // Proc of the 3rd USENIX Conf on Large-Scale Exploits and Emergent Threats (LEET). Berkeley, CA: USENIX Association, 2010 [2021-02-03]. http://usenix.org/event/leet10/tech/full_papers/Felegyhazi.pdf

[85] Antonakakis M, Perdisci R, Dagon D, et al. Building a dynamic reputation system for DNS [C] // Proc of the 19th USENIX Security Symp. Berkeley, CA: USENIX Association, 2010: 273-290

[86] Nabeel M, Khalil I M, Guan B, et al. Following passive DNS traces to detect stealthy malicious domains via graph inference [J]. ACM Transactions on Privacy and Security (TOPS), 2020, 23(4): 17

[87] Liang Zhizhou, Zang Tianning, Zeng Yuwei. MalPortrait: Sketch malicious domain portraits based on passive DNS data [C/OL] // Proc of the 2020 IEEE Wireless Communications and Networking Conf (WCNC). Piscataway, NJ: IEEE, 2020 [2021-02-03]. https://ieeexplore.ieee.org/document/9120488

[88] Sun Xiaoqing, Tong Mingkai, Yang Jiahai, et al. Hindom: A robust malicious domain detection system based on heterogeneous information network with transductive classification [C] // Proc of the 22nd Int Symp on Research in Attacks, Intrusions and Defenses (RAID). Berlin: Springer, 2019: 399-412

[89] Liu Zhicheng, Li Shuhao, Zhang Yongzheng, et al. Ringer: systematic mining of malicious domains by dynamic graph convolutional network [C] // Proc of the 20th Int Conf on Computational Science. Berlin: Springer, 2020: 379-398

[90] Manadhata P K, Yadav S, Rao P, et al. Detecting malicious domains via graph inference [C/OL] // Proc of the 19th European Symp on Research in Computer Security. Berlin: Springer, 2014 [2021-02-03]. https://link.springer.com/chapter/10.1007/978-3-319-11203-9_1

[91] Khalil I, Yu Ting, Guan Bei. Discovering malicious domains through passive DNS data graph analysis [C] // Proc of the 11th ACM on Asia Conf on Computer and Communications Security. New York: ACM, 2016: 663-674

[92] AV-TEST. Malware [EB/OL]. [2021-02-03]. https://www.av-test.org/en/statistics/malware/

[93] Salusky W, Danford R. Know your enemy: Fast-flux service networks [J/OL]. The Honeynet Project, 2007 [2021-02-03]. https://www.researchgate.net/publication/328781281_Know_Your_Ene

my_Fast-Flux_Service_Networks

[94] Fu Yu, Yu Lu, Hambolu O, et al. Stealthy domain generation algorithms [J]. IEEE Transactions on Information Forensics and Security, 2017, 12(6): 1430-1443

[95] Peck J, Nie C, Sivaguru R, et al. CharBot: A simple and effective method for evading DGA classifiers [J]. IEEE Access, 2019, 7: 91759-91771

[96] Sidi L, Nadler A, Shabtai A. MaskDGA: A black-box evasion technique against DGA classifiers and adversarial defenses [J]. arXiv preprint, arXiv:1902.08909, 2019

[97] Zheng Yu, Yang Chao, Yang Yanzhou, et al. ShadowDGA: Toward evading DGA detectors with GANs [C/OL] // Proc of the 30th Int Conf on Computer Communications and Networks (ICCCN). Piscataway, NJ: IEEE, 2021 [2021-09-03]. https://ieeexplore.ieee.org/document/9522282

[98] Anti-Phishing Working Group. Phishing activity trends report [EB/OL]. (2020-11-24)[2021-02-03]. https://docs.apwg.org/reports/apwg_trends_report_q3_2020.pdf

[99] Bell S, Komisarczuk P. An analysis of phishing blacklists: Google safe browsing, openphish, and phishtank [C/OL] // Proc of the 2020 Australasian Computer Science Week MultiConf. New York: ACM, 2020 [2021-02-03]. https://dl.acm.org/doi/fullHtml/10.1145/3373017.3373020

[100] Oest A, Safaei Y, Zhang Penghui, et al. PhishTime: Continuous longitudinal measurement of the effectiveness of anti-phishing blacklists [C] // Proc of the 29th USENIX Security Symp. Berkeley, CA: USENIX Association, 2020: 379-396

[101] Proofpoint. Proofpoint domain fraud report [EB/OL]. [2021-02-03]. https://www.proofpoint.com/sites/default/files/gtd-pfpt-us-wp-domain-fraud-report-2019.pdf

[102] Wang Yimin, Beck D, Wang J, et al. Strider typo-patrol: Discovery and analysis of systematic typo-squatting [C] // Proc of the 2nd Workshop on Steps to Reducing Unwanted Traffic on the Internet (SRUTI). Berkeley, CA: USENIX Association, 2006: 31-36

[103] Cisco. Total global email and spam volume for september 2020 [EB/OL]. [2021-02-03]. https://talosintelligence.com/reputation_center/email_rep

[104] Weimer F. Passive DNS replication [C/OL] // Proc of the 17th Annual First Conf on Computer Security Incident Handling. 2005 [2021-02-03]. https://static.enyo.de/fw/volatile/pdr-draft-11.pdf

[105] Li Wanting, Jin Jian, Lee J H. Analysis of botnet domain names for IoT cybersecurity [J]. IEEE Access, 2019, 7: 94658-94665

[106] Somarriba O, Zurutuza U. A collaborative framework for android malware detection using DNS & dynamic analysis [C/OL] // Proc of the 37th IEEE Central America and Panama Convention. Piscataway, NJ: IEEE, 2017 [2021-02-03]. https://ieeexplore.ieee.org/document/8278529

[107] Aizuddin A A, Atan M, Norulazmi M, et al. DNS amplification attack detection and mitigation via sFlow with security-centric SDN [C] // Proc of the 11th Int Conf on Ubiquitous Information Management and Communication. New York: ACM, 2017: 3

[108] Saharan S, Gupta V. Prevention and mitigation of DNS based DDoS attacks in SDN environment [C] // Proc of the 11th Int Conf on Communication Systems & Networks (COMSNETS). Piscataway, NJ: IEEE, 2019: 571-573

Fan Zhaoshan, born in 1997. Master candidate. Her main research interests include network security situational awareness and malicious domain detection.

樊昭杉，1997 年生．硕士研究生．主要研究方向为网络安全态势感知和恶意域名检测．

Wang Qing, born in 1995. PhD candidate. Her main research interests include network security situational awareness and malicious domain detection.（wangqing@iie.ac.cn）

王　青，1995 年生．博士研究生．主要研究方向为网络安全态势感知和恶意域名检测．

Liu Junrong, born in 1984. Master, senior engineer and master supervisor. Her main research interests include network security situational awareness and network security visualization.（liujunrong@iie.ac.cn）

刘俊荣，1984 年生．硕士，高级工程师，硕士生导师．主要研究方向为网络安全态势感知和网络安全可视化．

Cui Zelin, born in 1990. Master, engineer. His main research interests include situational awareness analysis and micro-services architecture analysis.（cuizelin@iie.ac.cn）

崔泽林，1990 年生．硕士，工程师．主要研究方向为态势感知分析和微服务架构分析．

Liu Yuling, born in 1982. PhD, senior engineer and master supervisor. His main research interests include network security situational awareness, network security big data analysis, evaluation and certification of information security.（liuyuling@iie.ac.cn）

刘玉岭，1982 年生．博士，高级工程师，硕

士生导师．主要研究方向为网络安全态势感知、网安大数据分析、安全测评认证等．

Liu Song, born in 1992. Master, engineer. His main research interests include data storage technology and network security situational awareness.

刘　松，1992 年生．硕士，工程师．主要研究方向为数据存储技术和网络安全态势感知．