DNS 检测研究现状综述

张乐

2015.10

汇报提纲

- DNS简介及检测技术介绍
- DNS检测分类及研究现状
- DNS检测方法及其局限性

DNS简介

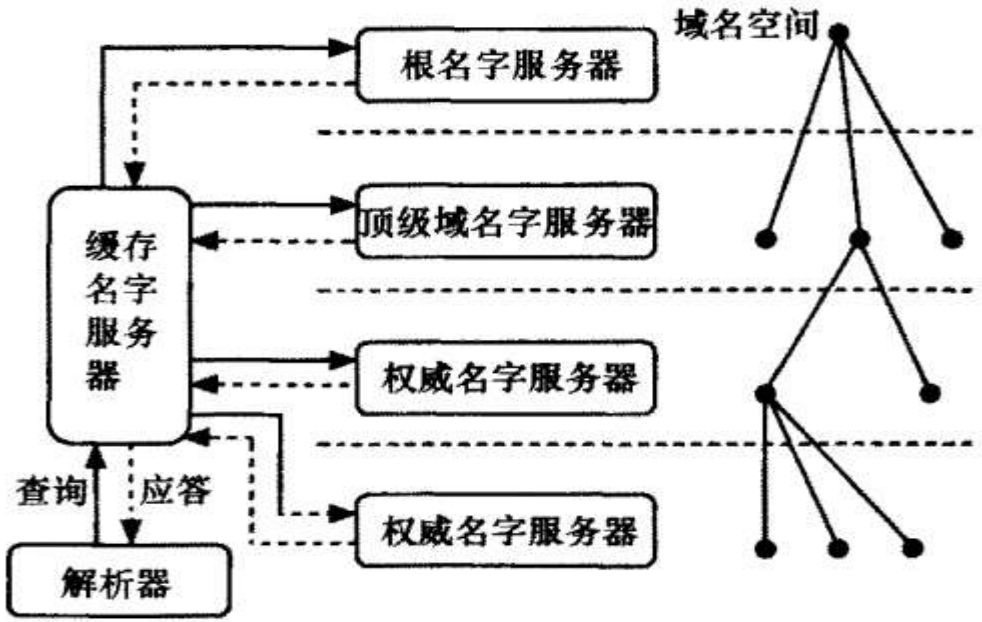
定义：DNS(domain name system，域名系统)是分布式数据库系统，其主要作用是将易于记忆的主机名称映射为枯燥难记的IP地址。

DNS体系结构：

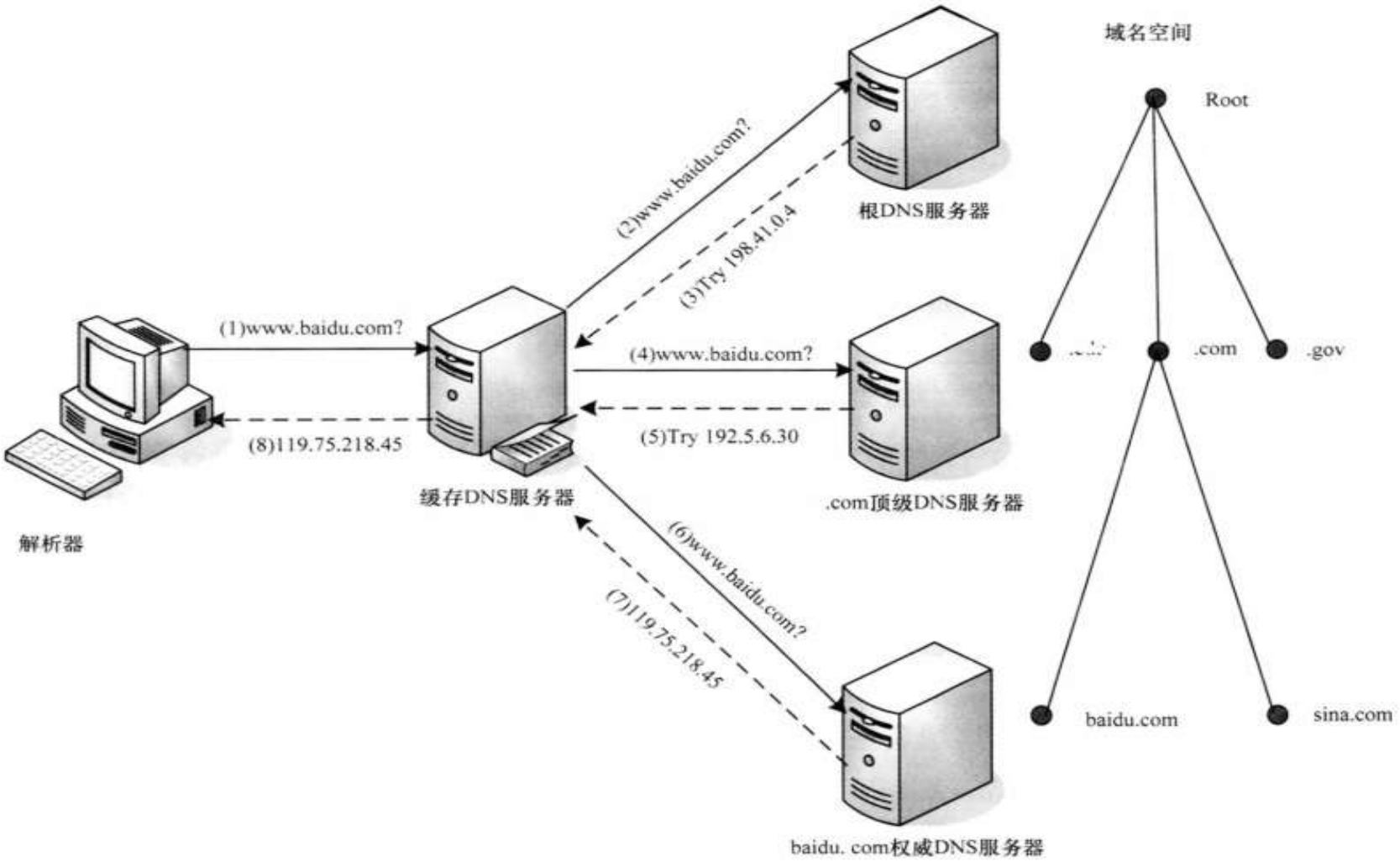
域名空间和资源记录

名字服务器

解析器



DNS解析



DNS 递归解析过程

技术介绍

DNS检测意义：域名解析服务应用范围广泛，域名解析服务也成为各类具有代表性的互联网安全威胁的重要工具，基于DNS检测研究具有重要的理论意义和实际价值。

DNS检测技术优势：

1）域名请求往往先于网络攻击行为，因而DNS检测更加及时；2）众多安全事件大多都与域名相关，更利于进行多事件关联挖掘和深层次事件分析； 3）域名信息相对较简单，易获取，便于应用。

检测分类

- DNS攻击：

  DNS隐蔽信道、DNS的DDoS、DNS缓存投毒

- DNS滥用：

  邮件蠕虫、僵尸网络、网络钓鱼、

  Fast-flux网络、恶意域名

- 域名依赖性：

DNS攻击-DNS隐蔽信道

- **定义**：隐蔽信道是指允许进程以危害系统安全策略的方式传输信息的通信信道。DNS隐蔽信道就是利用DNS通信数据包作载体进行秘密通信。
- **危害**：可作为远程控制通道，用于控制被攻破的内部主机，可以用来发布命令、传输文件。
- **特点**：具有强隐蔽性。

DNS攻击-DNS隐蔽信道

技术思路	利用特征	检测特点
基于DNS 有效载荷分析【1，2】	一个或多个 DNS 请求和响应的有效载荷大小、域名数字特征、统计最长有意义的子串长度%等	易于部署，用于检测特定的 DNS 信道，但此种检测只适用于基于域名递归解析的通道，不能解决隐藏UDP /53流量中客户端与服务器直接通信的通道
基于DNS流量分析【3】	对多个 DNS 报文进行分析，统计并分析 DNS 报文的数量、频率及其他属性。	易于部署，具有通用性，可用于检测多种 DNS 信道

DNS攻击-DNS的DDoS

- **定义**：通过主控机控制网络上的多个傀儡机同时向DNS服务器发动拒绝服务攻击。
- **危害**：耗费系统资源或者带宽，造成服务器崩溃或网络拥堵，从而无法为正常用户提供域名解析服务。
- **特点**：易发起、难追踪、流量大和隐蔽性强。

DNS攻击-DNS的DDoS

技术思路	利用特征	检测特点
基于统计分析的DDoS检测 【4】	当发生DDoS攻击时，网络流量会突然增大，与正常情况下流量的阈值进行比较判断	检测率高，但不能区分正常 的大流量和 DDoS攻击流量，易产生误判
基于神经网络技术的DDoS检测【5】	利用神经网络算法建立了一个DoS攻击检测模型，通过提取若干流量特征信息来对DoS攻击数据进行检测	检测率高，系统开销较小，能够处理噪音数据，并不依赖于对所处理 的数据 统计，易产生漏判

DNS攻击-DNS缓存投毒

- **定义**：通过污染DNS缓存，用虚假的IP地址信息替换缓存中主机记录的真实IP地址信息来造成破坏。
- **危害**：重定向用户至伪造的恶意网站，同时，缓存中的虚假记录会扩散到其他服务器，从而导致大范围的缓存投毒攻击，且互相感染，很难彻底根治。
- **特点**：易实施、隐蔽性强和防御难度大。

DNS攻击-DNS缓存投毒

- 传统DNS缓存投毒：传统 DNS 缓存投毒攻击存在的攻击所需时间长、成功率很低，污染的目标是应答数据包中带有查询结果 IP 地址的回答资源记录部分。

- Kaminsky 投毒：攻击上升了一个层次，污染的目标是应答数据包中 Authority Records（授权资源记录）因而破坏程度更高，但所需攻击时间更短，成功率也更高。

DNS攻击DNS缓存投毒

技术思路	利用特征	防御特点
增加DNS报文的信息熵提高投毒难度【6】	随机化的源端口来增加查询报文的信息熵	不能抵御持续Kaminsky投毒
修改当前的DNS体系【7】	通过部署DNS扩展协议DNSSEC，DNSSEC从概念、协议设计、报文格式、加密算法及密钥管理等方面完善了原有DNS体系的不足，为DNS解析服务提供数据源身份认证和数据完整性验证，	DNSSEC尚未全面部署，不能提供应有的安全防护

DNS滥用-邮件蠕虫

- **定义**：僵尸网络以邮件方式传播的恶意代码称之为邮件蠕虫。
- **危害**：邮件蠕虫从被感染主机的硬盘、cache或者用户的邮件列表获取感染目标，发送含有恶意代码的邮件达到扩散目的。
- **特点**：邮件蠕虫的目标性更强，感染率更高。

DNS滥用-邮件蠕虫

技术思路	利用特征	检测特点
基于阈值的检测【8】	Bot主机发送DNS请求频次高于正常主机	检测率中，需获取DNS服务器请求， 易于部署，但阈值不易确定
基于黑白名单的检测【9】	邮件服务器白名单，结合DNS MX请求数量	检测率高，需要提供大量邮件服务器地址，不易实现，容易产生漏判
基于请求流量行为分析【10】	Bot主机的DNS请求行为具有相似性	检测率中，计算复杂度较高，不宜进行在线检测

DNS滥用-僵尸网络

- **定义**：僵尸网络是指被攻击者控制的可执行协同任务的计算机集群。
- **危害**：攻击者通过僵尸网络可进行多种攻击，如垃圾邮件、DDoS 攻击、信息窃密等。
- **特点**：(a)Bot 与 Command & Control 服务器通信具有行为高度统一性 (b)僵尸网络规模相对稳定，被感染主机数量无大幅度波动。

DNS滥用-僵尸网络

技术思路	利用特征	检测特点
基于异常DNS流量的检测【11】	存在大量的请求域名失效	检测率中，需对DNS请求进行被动获取，易于部署
基于僵尸主机域名请求相似性的检测【12】	DNS请求集中，且请求行为高度相似	检测率高，需对DNS请求进行被动获取，易于部署，易误判
基于DNS 黑名单的对抗检测【13】	Bot主机会对黑名单进行探测，了解僵尸网络被封堵情况	检测率中，需对黑名单服务器请求进行被动获取，黑名单需要及时维护，易漏判

DNS滥用-网络钓鱼

- **定义**：指攻击者通过伪造的网站、 邮件等欺骗行为，诱导接受者泄露有价值或者敏感信息，如银行账号、密码等。
- **危害**：通过对钓鱼网站的分析发现， 绝大部分钓鱼网站伪造了银行或支付系统，进行经济欺诈。
- **特点**：多样性、欺骗性和针对性。

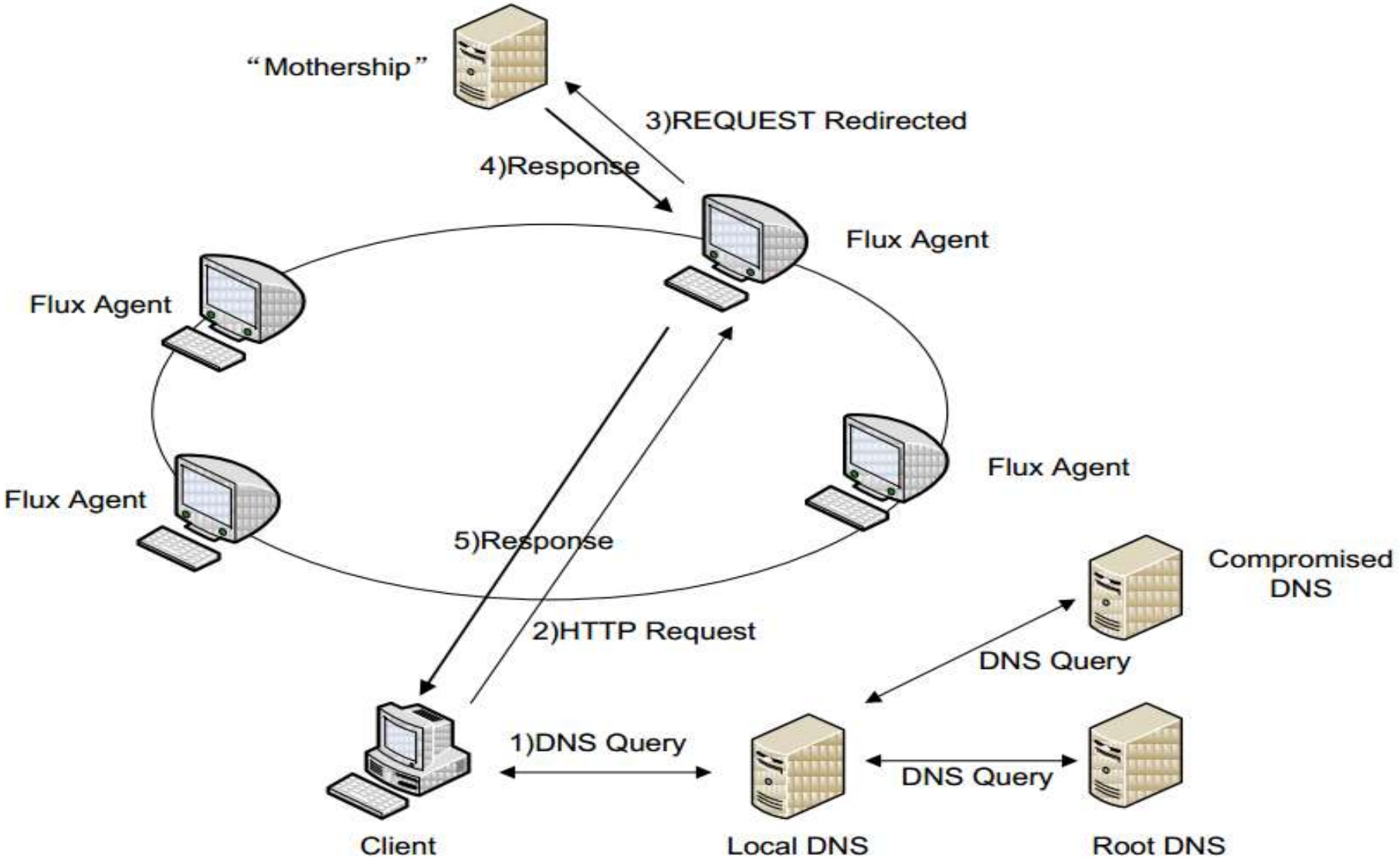
DNS滥用-网络钓鱼

技术思路	利用特征	检测特点
基于黑白名单机制的检测【14】	网络钓鱼通常意图为金融欺诈， 钓鱼网站的IP地址不属于银行或支付机构	检测率中，需要构建精确的白名单，不易实现，漏判率会比较高
基于域名字符串统计相似度的检测【15】	恶意域名 通常为自动生成，字符特征与合法域名存在较大统计差异	检测率中，需对DNS请求进行被动获取，易于部署

DNS滥用-Fast-flux网络

- **定义**：指通过不断变更僵尸代理主机，隐藏钓鱼服务站点或恶意程序分发点的DNS 技术。
- **危害**： mothership（真正的服务提供者）除承担提供服务的任务外，还可承担传统的Command & Control 服务器的角色。
- **特点**：通过多个 代理主机，实现了对mothership 的有效隐藏，极大提高了 僵尸网络的生存性。

DNS滥用-Fast-flux网络



DNS滥用-Fast-flux网络

技术思路	利用特征	检测特点
基于网络结构特性的检测【16】	域名信息、域名解析信息和域名解析得到的IP地址分布特征	检测率高，需进行主动探测，加重DNS服务器的负担，易漏判
基于网络流量相似性的检测【17】	bps（流平均每秒位数）、pps（流平均每秒包数）、bytes（流交换的总字节数）等网络流量特性	检测率中，需进行主动探测，容易受噪音干扰，易产生误判
基于行为与域名查询协同关联的检测【18】	相同僵尸网络中的各个僵尸活动相互之间具有时间、空间的行为相关性和相似性	检测率高，监测未知botnet，监测过程与botnet采用的协议结构无关，计算复杂度高

DNS滥用-恶意域名

- **定义**：指由攻击者提供的，可用于网络钓鱼、垃圾邮件传播、僵尸网络扩散等具有恶意行为的域名。
- **危害**：攻击者通过各种技术途径进行恶意域名的传播，危害网络安全。
- **特点**：隐蔽性强、追踪难、安全隐患更持久。

DNS滥用-恶意域名

技术思路	利用的特征	实用性
基于决策树的检测【19】	时间特性、DNS应答特性、TTL值和域名相关特性	检测率高，需采取主动探测机制获取域名相关信息，但总体仍然易于部署
基于域名字符串统计相似度的检测【15】	恶意域名字符特征与合法域名存在较大统计差异	检测率中，需对DNS请求进行被动获取，易于部署
基于域名评分机制的检测【20】	网络特性、基于区域特性以及基于证据特性	检测率中，需采取主动探测机制获取域名相关信息，但总体仍然易于部署

域名依赖性

- **定义**：指域名对于用户的重要程度，主要用于描述网络内部用户对于外部域名的重要性以及网络的外部用户对于内部域名的重要性。
- **方法**：选取域名的相关的特征测度，设计合理的算法，通过聚类算法对域名的依赖性进行划分。
- **意义**：便于查询域名承载的服务类型，为域名的异常检测及角色分析提供数据依据。

域名依赖性

- 刘茜从时间和空间两个维度来研究域名的依赖性问题。计算域名的依赖性等级，考察域名活动是否具有周期性，以及判定域名依赖性的稳定性，还依据域名空间位置信息分析的域名的依赖性。
- **不足：**i)域名数据库结构设计需进一步优化，ii）衡量域名的依赖性的测度选取需要进一步优化，iii）活动规律的统计实时性需要进一步提升。

检测方法

- 按分析数据及检测的方法：

特征检测：

用来检测已知攻击。

异常检测：

用来检测偏离正常模式的行为，可以检测到新的未知的攻击行为。

研究局限

DNS检测进行实时检测面临**三个问题**：

- 1、DNS 流量大， 特别是请求流量巨大，对处理和存储能力要求很高；
- 2、 为实现对 DNS 流量的有效监测， 需要覆盖一定数量的 DNS 服务器， 不易实现；
- 3、通过 DNS 流量进行安全事件检测是一个无固定目标的检测问题，计算复杂度本身就比较高， 因此检测难度较大。