# 构建基于DNS的动态信誉系统

Manos Antonakakis, Roberto Perdisci, David Dagon, Wenke Lee, Nick Feamster

佐治亚理工学院 计算机学院

**摘要** DNS是一个十分重要的协议，它不仅被用于合法的互联网应用，还被用于网络攻击。例如，僵尸网络可以利用DNS作为它C&C基础设施的一部分。摧毁这些攻击的有效的方法就是使用恶意域名黑名单，或者在防火墙和入侵检测系统中加入新的过滤规则。攻击者通过每天产生新的域名的方法来绕过静态黑名单和防火墙的检测。在本篇论文里我们提出了一个基于DNS的动态信誉系统Notos。该系统依据恶意域名有着独特的特征这一特点，将恶意域名与合法域名区分开来。Notos使用被动DNS查询作为数据集去分析这些域名的特征。它为合法域名和恶意域名建立了一系列的模型，并通过这些模型计算一个未知域名的信誉值。我们在一个有着140万用户的大型ISP网络的DNS流量下评估了Notos。 结果显示，Notos可以以96.8%的精度和0.38%的误报率识别恶意域名，并且相比于那些公开的域名黑名单，它可以提前几周甚至几个月对这些域名进行告警。

**1** 引言

域名系统(Domain Name System, DNS)是互联网的一项核心服务，它将域名映射到对应的IP地址上。DNS在网络安全中还可被用于分发IP的信誉信息，例如可以使用基于DNS的黑名单来过滤垃圾邮件和拦截恶意网站。

互联网规模的网络攻击通常利用DNS，因为它可以被互联网规模的恶意程序所利用。例如，间谍软件可以利用匿名注册的无意义域名盗取私人信息；恶意广告软件可以利用一次性域名托管恶意或虚假的内容；僵尸网络可以灵活地使用生存周期短的域名来改变它的C&C服务器地址；Fast-Flux网络频繁地改变DNS记录来躲避黑名单的封锁。攻击者们充分利用DNS的灵活来躲避域名黑名单的封锁。最常见的例子就是Fast-Flux，同样还有很多其他的方式诸如一次性域名(为了垃圾邮件或者僵尸网络生成成千上万的随机生成的域名)、拥有很多A记录与NS记录的域名(这些记录的数目远多于RFC标准的推荐值，目的是防止被封锁)和生命周期极短的僵尸网络域名等。可能最好的例子就是Conficker.C蠕虫。当该蠕虫感染了一台主机后，它通过尝试访问每天随机产生的50000个域名来连接C&C服务器。很明显，Conficker.C的这样做的目的就是使域名黑名单失去作用。利用DNS的恶意程序还包括Sinowal(即Torpig)、Kraken和Srizbi。其他恶意行为也充分利用新注册域名的这一特点，例如垃圾邮件以及恶意网络流量。这些会使得黑名单的封堵策略失效，并且无意义的域名还会污染整个互联网空间。

在本文中，我们研究了为新的未知域名动态分配信誉值的问题。我们的主要目的是给那些涉及到恶意软件传播、网络钓鱼、垃圾邮件等恶意行为的域名一个低的信誉值。与此同时，我们希望给那些合法的域名一个高的信誉值。相比于黑名单的做法，这种为域名动态评分的机制能更有效地对抗网络攻击。例如，当有足够证据证明一个域名是恶意的并将它加入黑名单进行封堵时，它已经在恶意活动中被使用了很重要的一段时间。我们的目标是建立一个动态的域名黑名单，这样即使对于那些新的域名，也可以判断它是否用于恶意用途。为此，我们提出了Notos——一个为域名动态分配信誉值的系统。我们的工作是基于对这些恶意域名的观察，发现它们独特的特征，这些特征可以与合法的、专业配置过的DNS服务区分开来。简而言之就是那些恶意和用于欺诈的网络行为为了躲避安全检查难免会有一些特征暴露出来。通过识别和衡量这些特征，Notos可以分配合理的信誉值。

Notos依据分布在互联网上的递归DNS服务器的历史流量来建立合法网络资源分配的模型。与此同时，Notos使用垃圾邮件陷阱、蜜网和恶意软件分析中得到的恶意域名和IP地址来建立恶意的网络资源分配的模型。有了这些模型，Notos可以给一个未知的新域名一个信誉值，进而实现未知恶意域名与IP的动态黑名单。

过去对动态信誉系统的研究主要集中在IP信誉上。据我们所知，我们的系统是第一个关于域名的动态信誉系统。

概括地讲，我们的主要贡献如下：

* 我们设计了完整的域名动态系统Notos，它可以给出域名的信誉值。我们根据捕捉到的资源配置、用途以及域名管理等特点构建网络和区域特征。这些特征可以让Notos学习到合法域名和恶意域名是如何运作的，并且可以为新的域名计算准确的信誉值。
* 我们实现了该系统的自验证的版本，并将它部署在美国的亚特兰大和圣何塞的大型ISP的DNS网络下，在这个网络中我们可以观察到140万用户产生的DNS流量。我们也使用了安全信息交流项目中的被动DNS流量。在大规模的真实流量评估下， Notos可以以96.8%的精度和0.38%的误报率来对新的域名进行分类。Notos可以发现与恶意软件或垃圾邮件相关的域名并且给予它们一个低的信誉值，而这些域名往往会在数天甚至数周后才出现在公开的黑名单里。

本篇论文的第二章介绍了DNS以及相关的技术背景，对这些熟悉的读者可以跳过直接阅读第三章。在第三章我们首先描述了被动DNS流量的收集方式以及其他白名单、黑名单的选取过程。接着介绍了主要网络、域以及基于证据的三个特征的提取模块。最后我们描述这些特征是如何聚集并合并到信誉引擎中去的。为了评估Notos的输出结果，我们收集了大量的网络跟踪数据。第四章介绍了数据的收集过程，第五章详细介绍了每个模块的精度以及最终的结果。

**2** 背景以及相关工作

DNS是将域名(如www.example.com)解析成IP地址(如192.0.2.10)的协议。为了解析这个域名，主机通常需要询问当地的递归DNS服务器(RDNS)。递归DNS服务器向负责每个域的权威域名服务器发起迭代查询。这个迭代过程的最终结果是在RDNS上建立请求域名和当前IP地址的一个映射关系。

根据整合所有在递归DNS服务器上成功解析为A记录的应答，我们可以建立一个被动DNS数据库。这个被动DNS(pDNS)数据库是被监控网络的有效DNS指纹，其中包含了不重复的A记录应答。一个域名典型的RR记录如下所示：{example.com. 78366 IN A 192.0.2.10}，含义为{域名，TTL，类别(Internet类)，形式(主机地址)，描述资源的信息(长度可变)}。为了简单起见，在本文中RR仅指域名-IP这一元组。

被动DNS数据的收集最早由Florian Weimer提出。他的系统第一次在出现在DNS社区，主要目的是将DNS流量转换为易于访问的格式。Zdrnja等人在他们的“监视被动DNS异常”中研究了如何将pDNS数据用于信息安全的数据收集中。尽管他们提出可以根据被动DNS数据来开发DNS信誉系统，但是他们并没有实现这方面的功能。在我们的工作中，只将被动DNS数据作为每个成功解析了的域名的计算特征。动态域名信誉的相关函数就是建立在这些特征的基础上的。Plonka等人提出了Treetop，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。Plonka提出的方法，虽然新颖高效，但该方法只提供了有限的关于DNS安全的信息，而且并不能对这些记录赋予信誉值。

还有一些论文如Sinha等人研究了IP黑名单的有效性；Zhang等人说明了高度可预测黑名单的命中率会在一段时间后显著减少。我们解决了动态DNS黑名单的问题，让它与高度可预测黑名单有着明显的不同。更重要的是，Notos并不是要创建IP黑名单。Notos通过利用DNS协议的属性，可以判断一个域名是否是潜在恶意的。Garera等人使用了URL的属性而不是通过观察域名或IP的统计数据来检测网络钓鱼行为。Holz等人使用统计特征来检测fast-flux网络，这和我们的工作很相似。但是，Notos利用了更完整的网络统计特征而且并不局限于对fast-flux网络的检测。

研究人员试图使用恶意网络独特的特征去检测恶意行为的来源。Anderson等人第一个提出了利用七层分析的方法识别并描述垃圾邮件基础设施的系统——Spamscatter。Hao等人提出了SNARE，一个有着高精度和低误报率的检测垃圾邮件的时空信誉引擎。SNARE信誉引擎是第一个利用基于网络的统计特征来检测垃圾邮件信息的。Notos互补了SNARE和Spamscatter的优点，而且扩展为不仅可以检测垃圾邮件，而且可以识别例如网络钓鱼、恶意主机等其他的恶意行为。Qian等人使用了基于网络的聚类来检测垃圾邮件。他们的论文表明，基于网络的聚类可以提高针对垃圾邮件的黑名单的准确度。我们的工作有着更强的通用性，因为我们试图识别出各种各样的恶意域名。不过共同点是我们都用了基于网络的聚类来识别恶意活动。

Felegyhazi等人提出了基于对WHOIS的观察来实现DNS信誉黑名单的方法。我们的系统在设计上并不将WHOIS的信息作为补充方法。Sato等人提出了通过观察IP地址信息的共性来扩展现有的黑名单的方法。Notos比Sato更具有通用性，并且不局限于与僵尸网络域名相关的检测。总之，Notos主要基于被动DNS作为数据来实现信誉功能，数据来自真实的网络流量，而不是从蜜罐中的监测的。

之前没有工作试图为网络边界上的任何一个域名赋予动态域名信誉值。Notos收集域名、有效域、IP地址、IP地址所属网络、自治系统以及蜜罐分析这些数据。此外，Notos使用短期的被动DNS信息。因为很难有恶意域名可以在被动DNS上面不留痕迹。

**3** Notos:一个动态的信誉系统

Notos信誉系统的目标是为域名动态地赋予信誉值。对于一个域名*d*，如果它与恶意活动(例如僵尸网络、垃圾邮件、恶意软件传播等等)有关，我们就希望给他赋予一个低的信誉值。另一方面，我们希望为那些合法的域名赋予一个高的信誉值。

Notos的主要数据来源是被动DNS(pDNS)数据库。数据库里有域名解析成对应IP地址的历史记录。我们的pDNS数据库是依据真实的DNS流量不断更新的，这些流量来自很多不同的地方，如图1所示。 我们从位于亚特兰大和圣何塞的两个ISP的递归DNS服务器收集DNS流量。ISP节点在峰值时DNS查询可以达到30000次/秒。我们也从安全信息交流(SIE)项目中收集DNS数据，这里的数据主要来自北美和欧洲的权威域名服务器。总的来说，SIE项目大约处理每秒200Mbit的DNS数据，这是美国单一ISP的DNS流量的好几倍。

我们使用的另一个信息来源是已知的恶意域名列表。例如，我们在受控的环境中运行已知的恶意软件，我们将恶意软件尝试连接的所有不在白名单中的域名都被标记为可疑的。此外，我们还从大的垃圾邮件陷阱中提取垃圾邮件中的可疑域名。接着，我们将与白名单匹配 的域名过滤掉，将其他的域名认为是可疑的。另外我们从alexa.com收集了大量的常用合法域名(我们将在第四章详细地讨论数据收集与分析的细节)。我们的数据库由已知的恶意域名和合法域名组成，在训练我们的信誉引擎(第四章)时会用到这些数据。

直观地讲，如果一个域名*d*或者它的IP地址(在过去的几个月里)与已知的恶意行为有关就会被认为是可疑的。如果我们更找到更多的关于这方面的证据，它的信誉分就越低。反过来说，如果有证据表明(过去)*d*与互联网上的合法的服务有关联，那么我们就会赋予它较高的信誉值。

**3.1** 系统概述

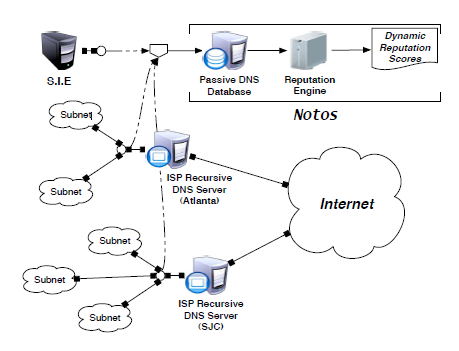


图1 系统的整体流程

在描述系统的内部实现之前，我们先介绍一些基本的术语。一个域名*d*可以看作由一些字符串组成或者看作被点“.”隔开的一组标签。最右边的是顶级域名TLD，次右边的被称作二级域名，倒数第三的标签后被称为三级域名，以此类推。例如，对于给定域名*d*=“a.b.example.com”，顶级域名为“com”，二级域名为“example.com”，三级域名为“b.example.com”。

令*s*代表一个域名(如s=“example.com”)，我们定义*Zone*(*s*)代表包含*s*且以.*s*为结尾的所有域名的集合(如以“.example.com”结尾的域名)。

令*D*={}表示一组域名的集合，让*A*(*D*)表示*D*中所有域名曾经指向过的IP地址的集合。

给定一个IP地址，定义表示的BGP前缀下的所有IP的集合，定义表示位于所在的自治系统的所有IP的集合。此外，对输入的IP的集合，我们可以扩展这些定义。给定IP集合，；同理。

为了对给定域名赋予信誉值，我们进行如下操作。首先，我们考虑域名最近指向的IP地址集合。接着我们去查询pDNS数据库来检索如下信息：

* 相关历史IP(RHIPs)：它包括，，。为了简化符号表达，我们将和分别表示为和。
* 相关历史域名(RHDNs)：它包括所有曾经解析到IP地址为的域名。换句话说，RHDNs包括满足的所有域名。

在pDNS数据库中提取以上信息后，我们构建了一些统计特征。具体来说，我们为每个域名提取三组特征，如图2所示。

* 基于网络的特征：第一组的统计特征是从RHIPs的集合中提取的。我们计算如下特征：与*d*历史相关的IP地址的总数量、它地理位置的多样性、它们所在的不同的自治系统的数量等等。
* 基于域的特征：第二组特征是从RHDNs集合提取的。我们计算在RHDNs里域名的平均长度、不同TLD的个数、不同字符出现的频率等等。
* 基于证据的特征：最后一组特征包括与*d*相关的不同恶意软件样本的数量、与*d*曾经指向的IP相关的恶意软件样本的数量等等。

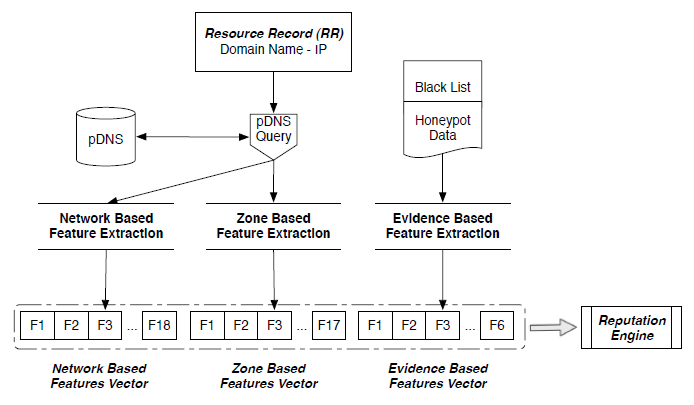


图2计算基于网络的特征、基于域的特征和基于证据的特征

这些特征一旦被提取出来，它们就会被传入信誉引擎中去。Notos的信誉引擎有两个运行模式：离线训练模式和在线分类模式。在离线模式下，Notos使用在知识库中收集到的数据来训练信誉引擎，知识库即一组已知的恶意和合法的域名及其相关的IP地址。在之后的在线模式中，对于每一个新域名*d*，Notos询问信誉引擎从而为*d*计算一个信誉值(如图3所示)。我们接下来详细解释这些统计特征，以及信誉引擎是如何在两个模式中使用这些特征为域名进行评分的。

**3.2** 统计特征

在本节中我们确定关键的统计特征以及阐述选择这些特征的原因。

3.2.1 基于网络的特征

给定一个域名*d*，我们从*d*的RHIPs集合中提取一些统计特征(正如3.1节中提到的)。我们的基于网络的特征描述了有着域名*d*以及*d*指向的那些IP地址的控制者是如何利用这些网络资源的。网络罪犯通常滥用DNS来达到灵活敏捷地操纵他们的恶意网络的目的。也就是说，那些被用于恶意目的的IP和域名通常很短命并且有着很高的流失率。这种灵活避免了简单黑名单的封锁以及被相关法律清除的危险。为了衡量一个域名*d*的灵活性，我们提取了用来描述*d*的网络轮廓的18个统计特征。我们的网络特征可分为以下三类：

* BGP特征。这一类一共有9个特征。我们计算与相关的不重复的BGP前缀的个数、这些BGP前缀分布的国家数、拥有这些BGP前缀的组织个数；在和集合中不重复的IP地址个数；与和相关的不重复的BGP前缀个数、这两个前缀集合所在的国家的个数。
* AS特征。这一类一共有3个特征。即与、和相关的不重复的自治系统的个数。
* 注册特征。这一类一共有6个特征。我们计算与在集合中的这些IP相关的不同注册人的个数；与在集合中的这些IP相关的不同注册日期；与在和集合中的这些IP相关的不同注册人的个数、不同注册日期。

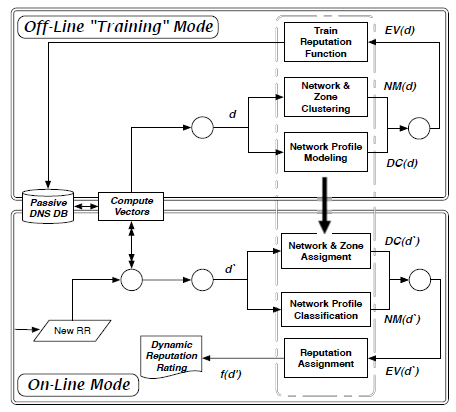


图3 Notos的离线和在线模式

大多数合法、正规运行的互联网服务有着非常稳定的网络轮廓，它们会在上述网络特征中取得一个低的数值。恶意网络配置(如 fast-flux)相对来说改变非常频繁，因此它们的网络特征将取得较高的数值。我们期望一个从合法的区域中选取的域名*d*会在它的AS特征中得到低数值，这主要是因为在RHIP中的IP应该属于相同的或者有着少数差异的组织。另一方面，如果一个域名*d*参与了恶意活动(如僵尸网络、flux网络)，它将会驻留在大量不同的网络中。恶意域名对应的RHIP中的IP列表将会在AS特征中产生较高的特征值。同样，我们也衡量良性域名的注册信息。合法域名通常指向那些有秩序地获取网络资源的组织的地址空间。这意味着，相同组织拥有的一个合法域名*d*的注册特征值会导致它的RHIP中的IP也有着较小注册特征值。如果这组IP地址有着高的注册特征值，这意味着它们很有可能位于不同的注册地并在不同日期注册的。这样的注册特征通常与欺诈域名有着联系。

3.2.2 基于域的特征

基于网络的特征计算了与一个域名*d*相关的历史IP的一系列特征。另一方面，基于域的特征计算了与一个域名*d*相关的历史域名的特征。利用基于域的特征的理由是，合法的互联网服务可能与许多不同的域名相关联，这些域名通常有着很强的相似性。例如，google.com, googlesyndication.com, googlewave.com等等都是谷歌提供的互联网服务，它们也都包含“Google”字符串。另一方面，与同一垃圾邮件活动相关的恶意域名通常看起来是随机生成的并且有很少的共同特征。因此，基于域的特征旨在衡量RHDN集合中域名的多样性。给定一个域名*d*，我们提取17个用于描述域名RHDN集合的性质的统计特征。我们把这17个特征分成两组：

* 字符串特征。这一组共有12个特征。我们计算RHDN中不同域名的数量、域名长度的平均值和标准差、在RHDN中的域名的每个字符出现频率的平均数、中位数和标准差；域名2-grams和3-grams分布的平均数、中位数和标准差。
* TLD特征。这一组共有5个特征。对于在RHDN集合中的域名，我们提取它的顶级域*TLD*()并且统计不同TLD的个数；我们计算*TLD*()=“.com”的域名个数与*TLD*()“.com”的域名总数的比值；我们计算TLD出现的频率的平均数、中位数和标准差。

值得注意的是当我们计算一个特定属性的平均数、中位数和标准差时，我们这么做是为了总结其分布的轮廓。例如，通过计算一组域名中的每个字符出现的频率的平均数、中位数和标准差，我们可以总结出字符频率是如何分布的。

3.2.3 基于证据的特征

我们利用基于证据的特征来判断一个给定的域名*d*是否与其他已知的恶意域名或IP地址相关联。正如之前提到的，Notos从公开信息中构建了一个已知可疑的、恶意的以及合法的域名和IP的知识库。例如，我们通过在受控环境中执行大量的恶意软件样本的来收集恶意软件相关的域名。同时，我们将IP地址与一些列的IP黑名单做比较。我们会在第四章详细说明了如何构建Notos的知识库的。给定一个域名*d*，我们使用知识库来计算6个统计特征。我们把这些特征分成两组：

* 蜜罐特征。我们计算3个特征：运行时试图访问*d*或*A*(*d*)中的IP的不同恶意样本的个数；连接*BGP*(*A*(*d*))中的IP的不同恶意样本个数；连接*AS*(*A*(*d*))中的IP的不同样本个数。
* 黑名单特征。我们计算3个特征：*A*(*d*)中的IP在公开IP黑名单中的个数；*BGP*(*A*(*d*))中的IP在公开IP黑名单中的个数；*AS*(*A*(*d*))中的IP在公开IP黑名单中的个数。

Notos使用从证据向量中的黑名单特征，因此它可以识别已知恶意的网络资源如IP、BGP前缀甚至AS的“复燃”活动。相比于IPv4地址，域名是非常廉价的。所以恶意用户往往会使用新的域名来重复利用地址空间。我们应该知道基于证据的特征只是我们用来计算信誉值参考信息的一部分。事实上，一个被恶意软件访问过的域名并不自动意味着它将得到一个较低的信誉值。

**3.3** 信誉引擎

Notos信誉引擎负责决定一个域名*d*是否与合法或者恶意的域名有着相似的特征。为了达到这一目的，我们首先需要训练信誉引擎去识别*d*是否属于(或接近)一类已知的域名。这种训练使用了Notos的知识库(见第四章)中收集到的历史信息，它可以在离线环境下周期性地重复。一旦引擎被训练完毕，在在线模式下它就可以为每个新的域名*d*赋予信誉值。

在本节中，我们首先介绍如何训练信誉引擎，接着我们介绍这个训练好的信誉引擎是如何(为域名)赋予信誉值的。

3.3.1 离线训练状态

在离线训练中(如图3所示)，信誉引擎建立了三个不同的模型。我们先简要介绍各个模型，接着再去详细地说明它们。

* 网络轮廓模型：一个描述已知的网络行为的模型。例如，我们对市面上的内容分发网络(如Akamai, Amazon CloudFront)和大型网站(如google.com, yahoo.com)的网络特征进行建模。在在线模式中，我们把每个域名*d*与这些已知的网络模型进行对比，并且使用这些信息去计算最终的信誉值，下面会详细说明。
* 域名聚类：我们将有相似特征的域名聚在一起。我们创建这些聚类的目的是识别包含大部分恶意域名或大部分合法域名的域名组。在在线模式中，给定一个新域名*d*，如果*d*(更准确地说应该是*d*的特征向量)落在(或接近)一个由大部分恶意域名组成的集群中，这就给我们了一个提示，我们应该赋予*d*一个低的信誉值。
* 信誉函数：对在Notos的知识库中的每一个域名,，我们用训练后的网络轮廓模型以及域名聚类去“测试”它。让和分别代表网络轮廓和域名聚类的输出。信誉函数将、、和)是合法的、可疑的还是恶意的(如果他们出现在恶意域名或IP的黑名单中)这一信息作为输入，并建立一个可以为*d*赋予一个介于0-1之间的信誉值的模型。信誉值接近0表示*d*是恶意域名，信誉值接近1表示*d*是合法域名。

现在我们详细描述各个模块。

3.3.2 网络轮廓模型

在离线训练过程中，信誉引擎为已知的网络行为建立了一个模型。建模过程的概述如图4(a)所示。在实际中我们选取了五组具有相似特征的域名并且“学习”了它们的网络轮廓。例如，我们识别一组与常见网站(如google.com, yahoo.com, amazon.com)相关的域名，并且对其中的每一个域名我们提取了它的网络特征，正如3.2.1所述。然后我们使用提取好的特征向量去训练一个统计分类器，这个分类器可以识别一个新的域名*d*是否有着与这些常用网站相似的网络特征。

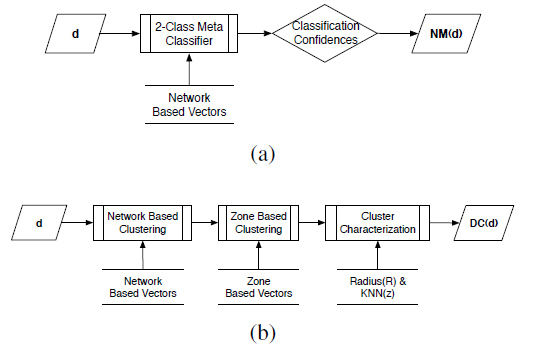


图4(a) Notos网络轮廓建模 (b)Notos中基于网络和域的聚类

在Notos的当前实现中，我们为域名建立了如下类别：

* 热门域名。这一类包括在以下DNS域的大量域名：google.com, yahoo.com, amazon.com, ebay.com,msn.com, live.com, myspace.com, and facebook.com.
* 常见的域名。这一类域名包括alexa.com中前100个域名。我们从这一组中剔除掉那些已经建模过的热门域名。
* Akamai 域名。Akamai是一个大的CDN厂商，与此CDN相关的域名有非常奇特的网络特性。为了对Akamai的域名的网络特性建模，我们收集了在以下域中的域名：akafms.net, akamai.net, akamaiedge.net, akamai.com, akadns.net, and akamai.com.
* CDN域名。在该类中我们涵盖了那些不是Akamai的CDN域名。例如，我们收集了在以下域中的域名：panthercdn.com, llnwd.net, cloudfront.net, nyud.net, nyucd.net and redcondor.net. 自从我们发现Akamai的域名有着非常不同的网络特征(第四章讨论)，我们选择不把这些CDN域名和Akamai域名归为一类。因此，为Akamai域名和CDN域名学习两个不同的模型会让我们与只学习一个模型相比得到一个更准确的分类效果(见3.3.5)。
* 动态DNS域名。这一类包括在两个最大的动态DNS服务商No-IP(no-ip.com)和DynDNS(dyndns.com)下注册的域名。

对于每一类域名，我们训练了一个统计分类器用来将它与其他类别区分开来。因此，我们训练了5个不同的分类器。例如，我们训练一个可以区分热门域名与其他种类域名的分类器。即对于给定域名*d*，分类器可以判断出*d*的网络轮廓是否与已知热门域名的网络轮廓相似。按照相同的逻辑，我们也可以识别其他域名种类的网络轮廓。

3.3.3 构建域名聚类

在这个阶段，信誉引擎将使用在“训练阶段”收集的pDNS数据库中的域名，对那些有着相似网络和域特征的域名构建聚类。我们分两步进行聚类，在第一步中，我们只使用基于网络的特征来创建粗粒度聚簇。然后在第二步中，我们利用基于域的特征将各个粗粒度的聚簇分成更细的簇，如图5所示。

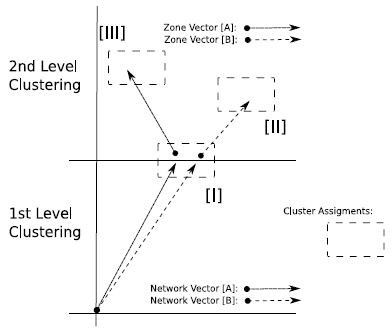


图5 Notos的基于网络和域的聚类过程，其中A代表Akamai，B代表恶意域名

* 基于网络的聚类。基于网络的聚类目的是将那些有着相同“灵活”性的域名聚在一起。这将把域名分成有着稳定和不稳定的网络特征(如CDN和恶意的flux网络)的两块。
* 基于域的聚类。在基于网络的聚类过后，我们将基于网络聚类的结果继续细分。在此步骤中，我们在同一网络聚类中将那些有着相似域特征的域名聚在一起。为了更好地理解基于域的聚类是如何工作的，考虑下面基于域聚类的例子：

聚类1：

..., **72.247.176.81 e55.g.akamaiedge.net**, 72.247.176.94 e68.g.akamaiedge.net, 72.247.176.146 e120.g.akamaiedge.net, 72.247.176.65 e39.na.akamaiedge.net, 72.247.176.242 e216.g.akamaiedge.net, 72.247.176.33 e7.g.akamaiedge.net, 72.247.176.156 e130.g.akamaiedge.net, 72.247.176.208 e182.g.akamaiedge.net, 72.247.176.198 e172.g.akamaiedge.net, 72.247.176.217 e191.g.akamaiedge.net, 72.247.176.200 e174.g.akamaiedge.net, 72.247.176.99 e73.g.akamaiedge.net, 72.247.176.103

e77.g.akamaiedge.net, 72.247.176.59 e33.c.akamaiedge.net, 72.247.176.68

e42.gb.akamaiedge.net, 72.247.176.237 e211.g.akamaiedge.net, 72.247.176.71

e45.g.akamaiedge.net, 72.247.176.239 e213.na.akamaiedge.net, 72.247.176.120 e94.g.akamaiedge.net, ...

聚类2：

..., **90.156.145.198 spzr.in**, 90.156.145.198 vwui.in, 90.156.145.198 x9e.ru, 90.156.145.50 v2802.vps.masterhost.ru, 90.156.145.167 www.inshaker.ru, 90.156.145.198 x7l.ru, 90.156.145.198 c3q.at, 90.156.145.198 ltkq.in, 90.156.145.198 x7d.ru, 90.156.145.198 zdlz.in, 90.156.145.159 www.designcollector.ru, 90.156.145.198 x7o.ru, 90.156.145.198 q5c.ru, 90.156.145.159 designtwitters.com, 90.156.145.198 u5d.ru, 90.156.145.198 x9d.ru, 90.156.145.198 xb8.ru, 90.156.145.198 xg8.ru, 90.156.145.198 x8m.ru, 90.156.145.198 shopfilmworld.cn, 90.156.145.198 bigappletopworld.cn, 90.156.145.198 uppd.in, ...

聚类中的每一个元素都是一个域名-IP对。这两组域名属于同一个网络聚类，但是根据基于域的聚类把它们分成了两组不同的簇。聚类1包括了属于Akamai的CDN的域名，而聚类2的域名全都与恶意软件访问的恶意站点相关。这两组域名簇有着相似的网络特征，但是在域特征上却有显著的不同。例如，考虑聚类1中的域名=“e55.g.akamaiedge.net”，聚类2中的域名=“spzr.in”。和被聚在同一基于网络的簇中的原因是和的RHIP集合有着相似的特征。特别地，所表现出来的“网络灵活性”让它看起来像是一个大型CDN的一部分。然而，当考虑和的RHDN集合时，我们可以注意到的域特征要远比的要稳定。换句话说，的RHDN集合中的域名有着很强的相似性(如每个域名里都包含“akamai”)并且有着较低的字符串特征值方差(见3.2.2)。强大的“域灵活性”影响了的RHDN的基于域特征的计算，并且使得看起来和有显著的不同。

Notos的一个主要的优点是它会可靠地为那些参与“敏捷”恶意活动的域名赋予低的信誉值。一些少数“敏捷”的恶意活动如Fake AVs可能将域名表示成类似CDN相关域名的样子。这样的诡计对FakeAV的活动并不会有帮助，因为像virus-scan1.com, virus-scan2.com等域名会被简单的正则表达式所阻拦住。换句话说，为了逃避简单的域名黑名单，攻击者需要同时在网络和域名层面上做得更加灵活。Notos为了训练目的只需要一些属于恶意活动的标记了的域名，然后信誉引擎接着会为属于相同的恶意活动的域名(之前未知)归纳赋予一个低的信誉值。

3.3.4 构建信誉函数

一旦我们建立了已知网络轮廓的模型(见3.3.2)和域名聚类(见3.3.3)，我们可以构建信誉函数。这个信誉函数会为域名赋予一个介于[0,1]的信誉值，0代表低信誉(也就是说可能恶意)，1代表高信誉(也就是说可能合法)。我们通过统计分类器来实现我们的信誉函数。为了训练信誉函数，我们考虑在Notos知识库中的所有域名=1,…,*n*, 我们将每个域名放入网络轮廓模块和域名聚类模块来计算两个输出向量和。我们会在3.3.5节中详细介绍和的计算。现在考虑和这两个特征向量已经足够了。对于每一个域名，我们也计算证据向量，如3.2.3节描述的那样。将、和结合成这一特征向量。我们使用标记了的数据集训练信誉函数，当是一个已知的恶意域名时，否则.

3.3.5 在线模式

训练完成后，信誉引擎可以在在线模式中(如图3)为新的域名赋予信誉值。例如，给定输入域名*d*，信誉引擎计算一个分数. *S*的值接近0意味着*d*似乎与恶意活动相关从而有一个低的信誉值。另一方面，*S*的值接近1预示着*d*似乎与良性的互联网服务相关，所以会有高的信誉值。信誉值的计算方法如下。首先，*d*作为网络轮廓模块的输入，这个模块包含了五个统计分类器，如3.3.2节讨论的一样。网络轮廓模块的输出是一个特征向量，其中是第一个分类器的输出，并且可以看作*d*属于热门域名这一类的可能性，同样可以看作*d*属于常见域名这一类的可能性等等。与此同时，*d*作为域名聚类模块的输入，这个模块计算了一个特征向量。这个向量的元素的计算方法如下。给定域名*d*，我们首先提取它的网络特征并且找到在离线模式中由域名聚类模块计算出的网络聚类中的与*d*网络特征最接近的簇(见3.3.3节)。接着，我们提取基于域的特征并且找到与*d*域名聚类最接近的簇。令这个最接近的域名簇为. 此时我们考虑所有基于域的特征向量，并且我们选择满足以下两个条件的向量的子集：1), 其中是*d*的域特征向量，是预定义的半径；2), 其中是的*k*个最近邻居的集合。

中的特征向量与Notos知识库提取出的域名相关。因此，我们可以根据计算的域名的性质为每一个向量赋予一个标签。Notos知识库中的域名分属不同的类。特别地，我们区分了8类不同的域名，即3.3.2节提到的热门域名、常见域名、Akamai、CDN以及动态DNS域名，以及垃圾邮件域名、Flux域名以及恶意软件域名。

为了计算输出向量, 我们计算以下5个统计特征：“主要类”标签*L*(例如*L*可能是恶意域名)，即在向量中出现最多的标签；标签出现频率的标准差，即给定在向量中每个标签出现的频率，我们计算它们的标准差；给定与标签*L*相关的的向量子集，我们计算与那些的向量的距离分布的平均值、中位数和标准差。

3.3.6 分配信誉值

给定一个域名*d*，我们计算了如上所述的向量和，我们也计算了如3.2.3节所说的证据向量。此时，我们将这三个特征向量整合到一个16维的特征向量中，并且我们将输入到已经训练好的信誉函数中(见3.3.4节)。信誉函数计算一个分数，其中可以被理解为域名*d*是恶意域名的可能性。介于[0,1]之间，并且的值越低，域名*d*的信誉就越低。

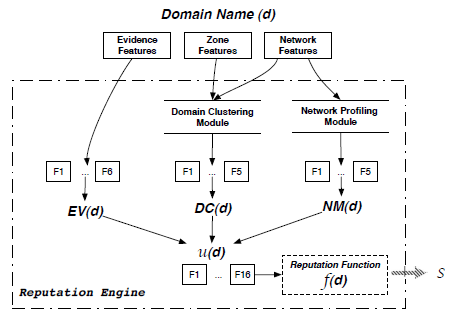


图6网络轮廓模块和域名聚类模块的输出和证据向量会帮助信誉函数为域名打分

**4** 数据的收集与分析

本节总结了对被动DNS测量的观察结果，并探讨了我们是如何将专业、合法的DNS服务与恶意服务区分开来的。这些观察结果为我们的动态域名信誉系统提供了参考标准。我们还使用了一些例如Akamai和Google这样的主要的Internet域来提供一个直观的例子来说明这些特性。

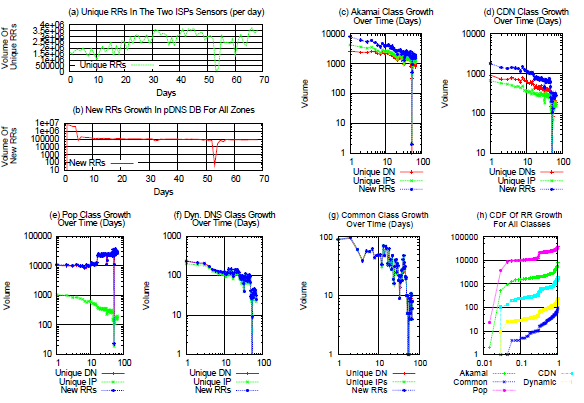


图7 Notos的基于网络和域的聚类过程，其中A代表Akamai，B代表恶意域名

**4.1** 数据收集

动态信誉评分系统的基础是从成功解析的域名A记录中获得的历史或“被动的”信息。我们使用了来自两个ISP的采集节点的DNS流量，一个位于美国东海岸(亚特兰大)，另一个位于美国西海岸(圣荷西)。另外我们也汇总了SIE项目中不同网络的DNS流量。我们的数据库收集到了从2009年7月19号至2009年9月24号这68天的解析记录共27377461条。

在这个大的数据集上通过利用我们选择的特征来进行一些简单的操作可以展现出数据的一些重要的性质。短短几天后，不重复的新pDNS条目的增长率趋于平稳。图7(b)说明了尽管每天都有非常多的RR记录(如图7(a)所示)，但是每天只有大约100000~150000的新域名产生(在第53天有一个简短的中断)。这表明大多数的RR记录是重复的，大约几天后，在测量点观察到的新的RR记录有平均94.7%已经在被动DNS数据库中有了记录。因此，Notos也可以部署在一个相对较小的pDNS数据库上。在第5节中，我们会衡量系统的精确度，该系统的流量来自规模较小的网络。

图7中的其他图片从五个不同的角度展示了我们日渐增长的被动DNS数据库。图7(c)和7(d)展示了与CDN网络(Akamai以及其他所有的CDN)相关的增长率。不重复的IP地址的数量与不重复的域名的数量几乎保持恒定(即每个新的RR记录包括一个新的IP和新的CDN的子域名)。在几个星期后，我们就知道了大部分CDN的IP地址，这意味着我们可以将CDN映射到一个合适的训练集中。这是因为CDN网络虽然很大，但它们往往只有一组固定的提供高可用的服务的IP地址。直觉上我们相信这不会是恶意“CDN”(如flux网络)出现的情况，因为它们会持续使用新的IP来进行传播和感染。

图7(e)展示了常用网站(如Google, Facebook)的新发现的RR记录中新的IP地址与域名的比例。Facebook为它们Web聊天客户端配备了独立的子域名，其他知名的互联网公司也使用了相似的策略(将信息编在域名中而不是URI中)，这也解释了为什么图7(e)中的域名会增长。这些常用网站使用了一组小范围的IP地址，并且经过几周的训练我们的pDNS数据库已经能将它们标识出来了。由于这些常用的域名占用了各种流量中一大部分，所以我们的想法就是使用简单的白名单显著降低分类器的工作量。

图7(f)展示了动态DNS服务提供商的pDNS增长率。这些服务有时候会被僵尸主控所利用，该图表明了一个近乎匹配的新IP与新域名的比例。采集的数据不包括不可路由到的应答(如指向127.0.0.1的动态的DNS域名)，因为它们不包含唯一的网络信息。直观地讲，我们可以认为动态DNS是域名对IP的几乎完整的双向映射。图7(g)展示了alexa.com中前100个域名的RR记录增长率。不同于动态的DNS域名，这些域名指向了一组少数的不重复的IP地址，并且我们可以在几周的训练中将大多数域名标识出来。

图7(h)是所有类别的一个对比，它展示了从图7(c)到7(g)的不重复的RR记录的累积分布情况。不同的变化速率说明每个类都有一套独特的RR记录的利用模式：有些IP地址数目不多但是域名的变数很大；有些几乎是每出现一个新的域名，就会出现一个新的IP。经过观察了每个区域类中大约90%的唯一RR记录后发现，我们只需要利用不同的RR记录即可。该图意味着尽管我们在研究中使用了很大的数据集，Notos也可以使用那些从相对较小的网络中收集到的数据进行工作。

**4.2** 建立参考标准

为了建立参考标准，我们用两个不同的标记过程。首先，我们在新的RR记录出现时就为他们附上一个标签。该操作为很多域名赋予了一个初始的静态标签。当然黑名单永远都是不全的和动态变化的。因此第二个标记过程发生在评估阶段，并且搜集了若干个公开的域名黑名单和白名单。

我们的标记的数据来自几个渠道。我们黑名单的主要来源是诸如malwaredomainlist.com和malwaredomains.com这样的服务。为了给我们pDNS数据库中的IP地址做标记我们也使用了来自Spamhaus的SBL(Sender Policy Block)列表。我们也从Zeus追踪项目中收集域名和IP的黑名单信息。所有这些黑名单信息在2009年8月的第一天前收集完毕(我们收集了15天的pDNS数据)。由于传统的黑名单跟不上活跃的威胁，所以在实验结束前我们会一直收集新的数据。

我们有限的白名单取自2009年8月1日的alexa.com的前500个域名。我们认为，尽管一些恶意域名很常见，但因为审查等原因它们不会长久，而且永远不会进入域名排名的前列。此外，我们使用了来自不同CDN厂商的18个最常见的二级域名列表，它们是我们标记为CDN的RR记录集的重要组成部分。最终，464个动态DNS的二级域名列表使我们可以辨认和标识那些来自动态DNS服务区域的域名和IP。我们通过汇总更新了的恶意域名和IP的黑名单用来标记待评估的数据集。

为了计算蜜罐特征(见3.2.3节)我们需要一个能处理尽可能多的新的恶意样本的分析设施。我们的蜜罐基础设施和Ether相似，并且有能力处理一列的恶意样本。它会在一个受控的环境和时间(5分钟)下分析每个恶意样本。我们在2009年7月的过去的15天内一直重复这一过程。15天过后我们得到了一组每个恶意软件查询过并成功解析了的DNS记录(包括域名和IP)。我们尝试执行恶意软件并且收集这些DNS证据，当然与此同时这些我们也会在被动DNS数据库中汇总出来。我们的虚拟机安装了五个常见的商业杀毒引擎。如果其中一个引擎判断这个可执行程序是恶意的，我们就捕捉这个恶意软件在它执行时间内访问过的所有域名和IP地址。在排除了alexa.com前500个域名之后，我们整理它们作为我们的“蜜罐数据”。我们自动地爬取和收集黑名单和蜜罐执行的信息。

读者应该注意到我们以一种尽可能透明的方式来标记我们的数据。在我们建立模型和训练信誉函数前，我们使用公开的黑名单来标记我们的训练集。接着，我们为数据赋予信誉值并且使用同样的公开的黑名单来验证结果。我们可以放心地假设私有的IP和DNS黑名单会比公开的黑名单包括更全的信息和更低的假正(FP)率。通过使用这种类型的私有黑名单，Notos的信誉函数的精度会大幅提高。

**5** 结果

在本节中，我们给出评估的实验结果。我们展示了Notos可以以96.8%的精度和0.38%的误报率发现恶意域名，并且它比黑名单方法更快。第一步我们从250000个不重复的RR记录中计算基于统计特征的向量(见3.2节)。这些对应于那些新的~以前从未见过的~一天内一个主要ISP的两个递归DNS服务器观察到的RR记录的平均值，如第4节图7(b)提到的那样。这些向量是基于从亚特兰大和圣何塞的ISP递归服务器中观察到的最近两周的DNS流量所产生的历史被动DNS数据计算而来的。

**5.1** 网络轮廓建模的准确度

网络轮廓模块中的元分类系统(图4(a))的准确度对于Notos的整体性能十分重要。这是因为在在线模式中，Notos会收到没有被标记的向量，Notos必须将这些向量进行分类并且将它们与我们知识库中存在的数据相关联。例如，如果分类器收到了新的RR记录并且非常肯定地为它赋予了Akamai标签，这说明由该RR计算出的向量与Akamai这样的网络的一部分很相似。但是这并不一定意味着它是实际的Akamai CDN的一部分。我们会在下一节看到我们是如何能在基于相同的区域簇内的标记和未标记的RR记录的相似中得出结论的。此外，我们讨论了元分类器为每个不同类别的网络轮廓建模时的准确度(网络轮廓在3.3.2节有描述)。

我们的元分类器包括五个不同的分类器，不同分类器针对不同种类的域名。我们使用元分类系统而不是传统的分类方法的原因是元分类系统比单一的统计分类器效果要好。我们在实验中也证明了这种方法的正确性。图8的ROC曲线表明元分类器可以对不同种类的网络轮廓的RR记录准确地进行分类。

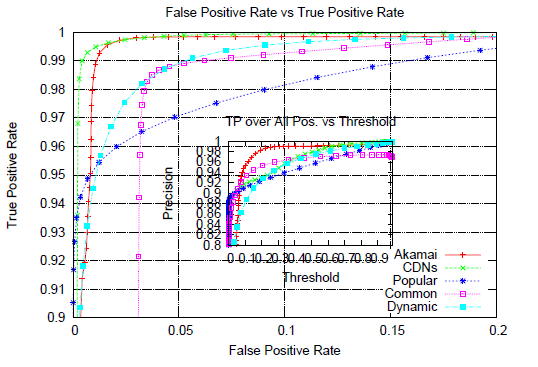


图8 展现元分类器准确度的不同种类的网络轮廓的ROC曲线

元分类器的数据集由五类网络轮廓的2000个向量组成。评估数据集由10000个向量组成，五类网络轮廓各提供2000个。Akamai，CDN，动态DNS以及热门域名的分类结果表明Notos中的监督学习的过程是准确的，只有少数关于常见域名的误报(3.8%)。通过手动地分析这些误报，我们的结论是动态DNS域名和常见域名在一定程度上仍然存在分不清的现象。但是网络轮廓中的这点分类错误并不会很大程度地影响信誉函数。这是因为动态DNS域名和常见域名在域特征上是截然不同的。域配置上的特征差异会把因基于网络配置聚成一起的动态DNS域名和常见域名的RR记录分为不同的聚簇。

尽管网络建模的过程提供了准确的结果，但这并不意味着这一步可以独立地指定一个域名是良性的还是恶意的。聚类阶段不仅会从网络轮廓上还会从域配置上来帮助Notos聚集向量。在下面的部分我们会展示网络和域配置的聚类模块是如何根据域名结构的性质更好地关联相似向量的。

**5.2** 网络和基于区域的聚类结果

在域名聚类过程中(见3.3.3节，如图4(b))我们使用一系列的X-means算法。第一次是基于网络的聚类第二次是基于域的聚类。在这两个步骤中我们分别为数据集中的向量集设置聚簇个数最大和最小值。我们使用不同数量的域和网络向量来运行这两个步骤。图11说明了在第一个100000个向量以后，网络和域的聚类个数就基本保持稳定。这意味着通过计算至少100000个网络和域的向量——使用15天的被动DNS数据库——我们可以得到一个监控网络下的基于域和网络的稳定聚类个数。我们应该注意到这个网络和集群的平衡并不意味着我们不希望看到任何在ISP的递归DNS服务器上的新类型的域名。这只说明对于我们被动DNS数据库中存在的RR记录以及ISP递归服务器的日常流量，100000个向量已经足够反应被监控网络的网络轮廓的趋势。造成这个趋势的结果应该取决于被监控网络的种类、规模以及用途。如果没有这较小组织的网络中的数据，我们很难可靠地评估出需要训练Notos的最少的RR记录的数目。

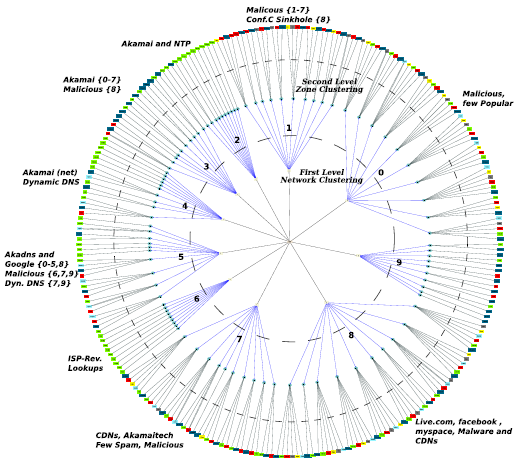


图10通过两步聚类，Notos可以对较大的DNS行为进行聚类

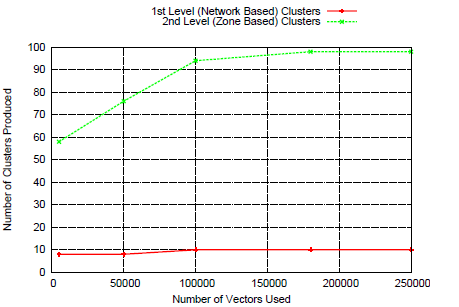


图11通过对不同数目的网络和域向量进行测试，我们发现过了第一个100000后，两个聚类阶段的结果都没有明显的变化

我们使用不重复的250000条域名和IP记录作为评价的数据集。聚类的情况如图10所示，接下来我们会讨论从这些基于域和网络的聚类中得到的一些有意思的观察结果。例如，网络聚类0和1主要是由参与欺诈活动如垃圾邮件(黄色)和恶意程序访问或C&C(红色)组成。而另一方面，网络聚类2到5包括了Akamai、动态DNS以及热门的如Google这样的域名，这些都标记为良性的(绿色)。基于信誉函数的准确度，我们还包括了未被标记了的向量(蓝色)。几乎所有的网络和域的聚类都有未被标记的向量样本。我们将看到已标记的向量是如何在眉睫之内帮助我们描述那些未标记的向量的。

在我们描述基于域的聚类的两个动态特征的样本前，我们需要讨论我们的半径*R*和*k*的选取(见3.3.5)。在3.3.5节，我们讨论了我们是如何构建域名聚类的。在那时我们介绍了动态表征过程，它使得Notos可以根据先验知识，利用已经标记的向量描述那些新得到的未标记的向量。通过观察同一域簇下的已标记和未标记向量的欧式距离的分布，我们得出了多数情况这个距离介于0~1000的这一结论。我们为K近邻算法测试了多个不同的*R*和*k*值。我们发现实验中，*R*介于50和200间时信誉评分的结果最准确，我们会在下节描述它。我们也发现如果*k*>25，信誉函数的准确度就不会受到所有介于50和200的*R*的影响。基于这些实验结果，我们决定令*k*=50,*R*=100.

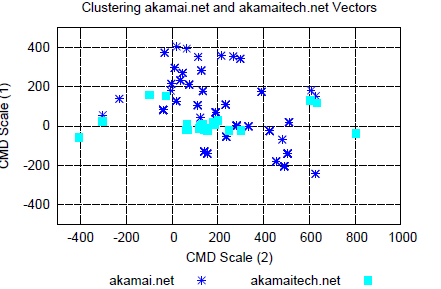


图12 根据已标记的向量(akamai.net)将akamitech.net的这些未知向量标记为良性的例子

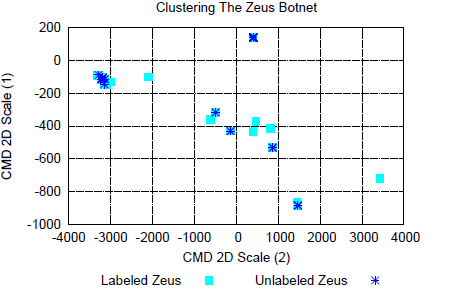


图13 Zeus僵尸网络聚类的过程。所有的向量都是基于同一网络聚类，两个不同的域聚类中。

图12和13说明了在两个不同聚类问题上半径选择的效果。在图12中，未知的RR记录akamaitech.net与已标记的向量akamai.net聚在了一起。如第4节提到的，例如Akamai的CDN有使用新域名的倾向，但是会重用它们的IP。通过训练已标记akamai.net的RR记录的一小部分，我们的分类器将未知的akamaitech.net的RR记录分到了Akamai的类别中。所以特定的IP特征将新的RR记录与已经标记了的类紧密连在一起。图12将所有维度压缩成2维的图片(为了更好地展示)，很明显未知的RR记录都在已标记的数据集的100距离内。

这些结果验证了第4节中的设计理念，即仅仅几周的已标记的数据足以用来训练了。因此，我们不必详尽地发掘所有的白名单。Notos可以弹性地应对我们选择的域类别的改变。类似CDN和主流的web站点会增加新IP或者调整域名的格式，而这些都会自动地与一个已知标记了的类别相关联。

图13再一次展示了Notos基于有限标记的输入将新的RR记录关联起来的能力。在这种情况下，标记为Zeus的域名(三个不同的与Zeus相关的黑名单的大约2900条RR记录)用来分类新的RR记录。图13描绘了标记为与Zeus有关的RR记录和新的(以前未知)的RR记录的距离，证明那些新的RR记录也与Zeus僵尸网络有关。我们可以从第4节知道，大多数的新的(未知)Zeus RR分布非常近，甚至常常与已知的Zeus RR重叠。这是个很好的结果，因为由于Zeus的灵活性，Zeus僵尸网络是出了名难以跟踪。诸如zeustracker.abuse.ch和malwaredomainlist.com这样的追踪系统只能“看见”僵尸网络的一部分，并且经常产生不相干的黑名单。Notos通过利用有限的训练数据对新的RR记录进行正确的分类解决了这一问题。在我们的评测集中，Notos成功检测出了685条新的Zeus RR记录。

**5.3** 信誉函数的准确度

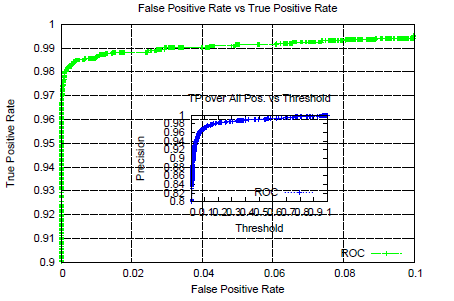


图9信誉函数的ROC曲线预示Notos具有很高的准确度

我们在本节中解决的第一件事是决定使用Logit-Boost策略的决策树(LAD)作为信誉函数。我们的决定是出于对分类器的时间复杂度、检测结果和精度的考虑。我们使用了一个典型的模型选择过程将LAD分类器与其他统计分类器进行了比较。LAD是为了在最短的训练时间内为信誉函数提供最准确的结果而诞生的。我们可以在图9的ROC曲线中看到，LAD分类器呈现出了96.8%的精度(TP%)和0.38%的误报率(FP%)。值得注意的是这是10折交叉验证得到的结果，并且检测阈值设定为0.5。用于评价的数据集包括10719条RR记录，其中有9530个已知的恶意域名。已知的良性域名列表包括Alexa的前500个热门域名。

我们也在另外两个包括大量已知良性域名的数据集上测试了信誉函数。我们将前10000个和前100000个Alexa域名分别作为实验数据。实验的检测结果如下。当使用前10000个Alexa域名时，我们得到了93.6%的精度和0.4%的误报(再一次使用了10折交叉验证和0.5的检测阈值)。正如我们所看到的，这些结果与当我们只用Alexa前500的域名没有什么不同。但是，当我们将良性域名列表扩大到Alexa的前100000个时候，我们看到精度明显下降，误报明显增多。这次我们得到了80.6%的精度以及0.6%的误报。我们认为准确度的下降主要是由于Alexa前100000个域名不仅仅包含了专业的域名以及相关的网络设施，还包括了一些不怎么好的域名，如文件共享、色情相关等等，这些域名没有按常理出牌并且有一个有争议的信誉值。

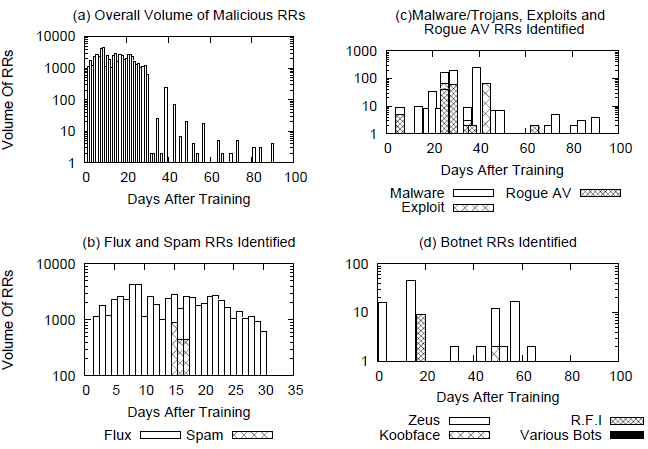


图14 8月1日Notos赋予低的信誉值后，不同种类的黑名单证实RR记录为恶意的时间。

我们也想去评判Notos相比于静态的黑名单有多好用。为此我们进行了如下的实验。给定一个Notos实例，该实例使用了截至2009年7月31日的数据作为训练集，我们将2009年8月1日DNS流量中观察到的250000条不同的RR记录作为输入。我们接着为每一条RR记录计算信誉值。首先我们设置检测阈值为0.5，在这阈值下我们鉴定54790条RR记录有一个低信誉值(低于阈值)。这些RR记录一共与10294个不同的域名有关(注意一个域名可以映射到多个IP，这就是为什么RR记录比域名要多的原因)。在这10294个域名中，有7984(77.6%)至少出现我们用来比较的其中的一个8月1日之后的60天的黑名单里，所以它们被证实是恶意的。图14(a)展示了出现在公开黑名单的被Notos判断为低信誉值的RR记录的数目和日期。其他的三个图(图14(b)、(c)、(d))展示了根据恶意域名的类型分别组成的结果。特别地，值得一提的是与公开黑名单比，Notos可以提前几天甚至几周发现与Zeus僵尸网络相关的从未出现过的域名。

对于10294中剩余的22.4%的域名，我们不能得出一个确定的结论。但是，我们认为这其中的绝大多数域名都或多或少地参与了恶意活动。我们也发现7980或7984个已经被证实是恶意的域名被赋予一个不大于0.15的信誉值，并且那些没有被证实的可疑域名得到的信誉值都比0.15高。在实际中，这意味着那些希望将Notos作为独立的动态黑名单系统的管理员可以将检测阈值微调至0.15左右，这样可以将误报率降低至一个可以忽略(甚至为0)的数额。

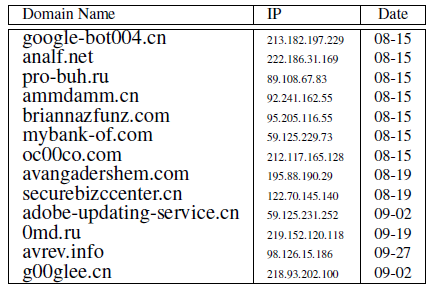


表1 Notos检测到的Zeus 域名样本以及它们出现在公开黑名单的时间。表格中的信息全部来自zeustracker.abuse.ch

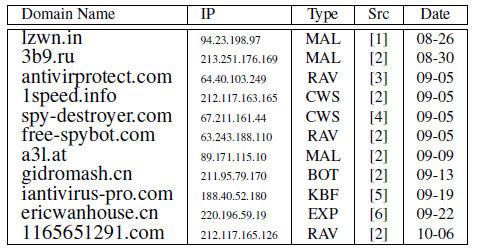


表2 Notos检测到的各种各样的恶意域名以及它们出现在公开黑名单的时间。

**5.4** 讨论

这一节讨论了Notos的局限以及在真实网络下潜在的绕过姿势。主要的局限就是Notos不能为那些仅有非常少的历史信息(被动DNS)的域名评分。创建一个准确的被动DNS数据库需要足够长的时间和相对足够多的被动DNS数据。因此，如果一个攻击者总是购买新域名和新的地址空间，并且永远不重用那些用于恶意域名的网络资源，Notos就不能精确地为新的域名赋予信誉值。在IPv4地址空间里，这不太可能发生，因为地址空间已经被用尽了。一旦IPv6成为主流的协议，在我们提取基于IP粒度的统计特征时这就是一个问题。但是，我们认为基于BGP前缀和AS号的特征依旧能捕捉到恶意DNS的灵活的行为。

只要新产生的域名与现有标记了得RR记录有着一些相同的网络特征(如IP或BGP前缀)，Notos就可以为它赋予一个精确的信誉值。特别地，因为网络资源是有限的并且更新或变更它们是非常昂贵的，所以即使域名的属性变化，Notos仍然可以辩认出一个域名是否与恶意活动有关。此外，对于一个给定的域名，我们想知道它的得分但是它并没有在被动DNS数据库里，我们可以通过寻找相关的被动DNS条目从而主动地去探索它。但是，只有当域名成功解析到不为空的IP地址时这种方法才奏效。

我们使用Alexa前10000的域名作为良性域名，使用它们进行测试得到了0.4%的误报率。尽管比例低，在Notos每日有大量的新域名输入的情况下(例如部署在大型ISP网络上)误报率的绝对值就显得非常重要。不过，我们并不想将Notos信誉系统独立地使用而是打算将它与其他的防御机制合作。例如，Notos可以与垃圾邮件过滤系统一起协作。如果一封邮件中包含Notos认为较低的信誉值的链接，垃圾邮件过滤器就可以增加它的过滤过则。但是如果其他的邮件看上去是正常的，过滤器仍然会决定接收这封邮件。

在我们对我们评估过程中的出现的(一部分)误报进行人工分析时，我们可以发现一些有趣的事情。我们发现网络中一些合法的站点(如goldsgym.com)内“寄居”了大量的恶意域名。这种情况Notos就会对该合法域名的信誉值进行“处罚”因为它们有一些“坏的邻居”。随着时间的推移，如果这个网站的管理员采取行动清理了在他们地址空间下的恶意域名，这些合法域名的信誉值也会随之改变。

**6** 结论

在本篇论文中，我们介绍了Notos，一个基于DNS的动态信誉系统。据我们所知，Notos是第一个可以为被监控网络边界下的DNS查询的任何域名赋予动态信誉值的系统。Notos从多个信息源如域名属于的DNS区域、相关IP地址、BGP前缀、AS信息以及蜜罐分析得到的数据来维护最新的关于合法和恶意域名的DNS信息。基于这些信息，Notos使用自动分类和聚类算法为合法和恶意的域名的网络和域行为进行建模，接着将这些模型用于新域名的评分上。

我们使用了真实的数据进行评估，这些数据来自大型ISP网络的流量，通过评估表明了Notos在监控网络的DNS流量下能很准确地以96.8%的精度和0.38%的误报率识别新的恶意域名。另外，Notos可以比公开的黑名单提前几周甚至几月发现这些恶意域名，从而可以对网络攻击采用更加积极的对策。

**7** 致谢

感谢我们的导师Steven Gribble帮助提高最终版本的质量，感谢为本文提出建设性意见的匿名评审。感谢Gunter Ollmann和Robert Edmonds提出的宝贵意见。此外，我们还感谢互联网安全联盟安全信息交换项目(ISC@SIE)为我们提供的部分DNS实验数据。

本课题收到了美国国家科学基金会下的第0831300号、美国国土安全部下的第FA8750-08-2-0141号、海军研究办公室的第N000140710907和第N000140911042的相关支持。在本文中表达的任何观点、研究成果、结论或者建议只与作者有关，与国家科学基金会、美国国土安全部和海军研究办公室无关。