(顶级期刊TOPS)Efficient and Accurate Behavior-Based Tracking of Malware-Control Domains in Large ISP Networks

本文工作：

1. 创新性的提出了一种基于行为的系统Segugio，segugio在大的ISP网络中通过追踪被恶意软件感染主机的DNS请求行为来高效的发现新增的malware-control域名。
2. 实现了Segugio原型系统，并将其部署在三个大型的ISP网络（和咱们的应用场景比较像），能在十几分钟内分类一整天的ISP层面的DNS流数据，实现94%TP / 0.1%FP
3. 其他结果
   1. Segugio能检测出unseen malware家族相关的malware-control域名，85%TP / 0.1% FP。
   2. 在一个ISP网络训练出的Segugio可以部署在其他ISP网络，依然保持high accuracy
   3. 可以在商业黑名单出来之前的数天甚至数周检测出新的malware-control域名
   4. 在ISP网络中新的、unknown的被感染主机可以被检测出来
   5. Segugio性能优于Notos以及其他基于loopy belief propagation算法的系统

注：belief propagation是一种动态编程方法，解决图模型中的条件概率请求；给定一些已知节点（图的子集），计算图剩余部分的条件概率。在这个文章中，我粗略的理解为知道了二部图的一部分连接求，剩余的未知连接。

相比前人工作的不同之处：

1. Notos（信誉系统）和EXPOSURE都是建立domain-IP映射关系模型（使用域名字符串的特征、域名承载的恶意内容等信息）；他们没有利用本地DNS服务器下游的主机请求行为，Segugio通过监测ISP网络用户的DNS请求行为，重在精确的追踪新增的“malware-only”域名。
2. Kopis[Antonakakis et al. 2011]的做法和本文有相似性（Kopis用请求者散度、请求者画像等信息）, 但Kopis利用权威或TLD服务器的数据，这种数据不好获得（需要与大的DNS区域运营商紧密合作）。Segugio不用关心顶级域名，通过监测本地ISP流量（在ISP使用者和他们的本地DNS解析器之间的DNS流量）。因此Segugio可以依靠ISP网络权限独立开发，不需要与外部的DNS运营商合作。

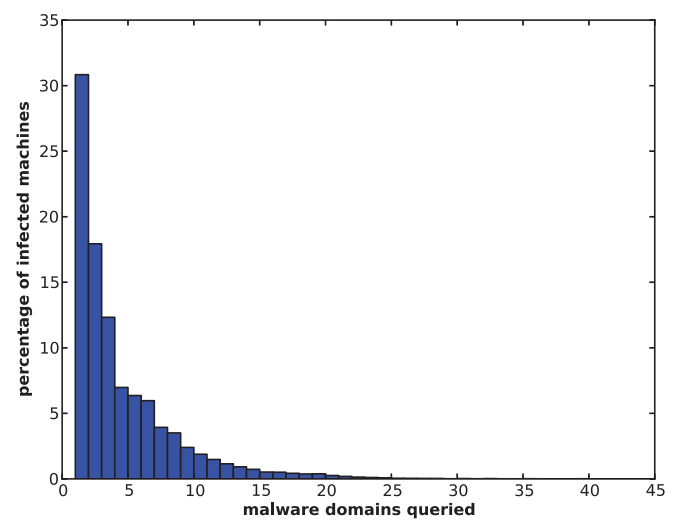
|  |  |  |  |
| --- | --- | --- | --- |
| 与Segugio比较 | 共同点 | 老方法缺点 | Segugio特点 |
| Manadhata2014 | 基于loopy belief propagation用图模型检测恶意域名 | 仅使用了machine-domain关系 | 充分利用了访问行为，结合被请求域名的特性（lifetime，解析IP等） |

Segugio的几个出发点：

1）在时间特性上，被感染主机倾向于请求新的恶意域名（或新的IPs）；

2）被相同的恶意软件家族感染的主机访问的恶意域名集合有重叠；

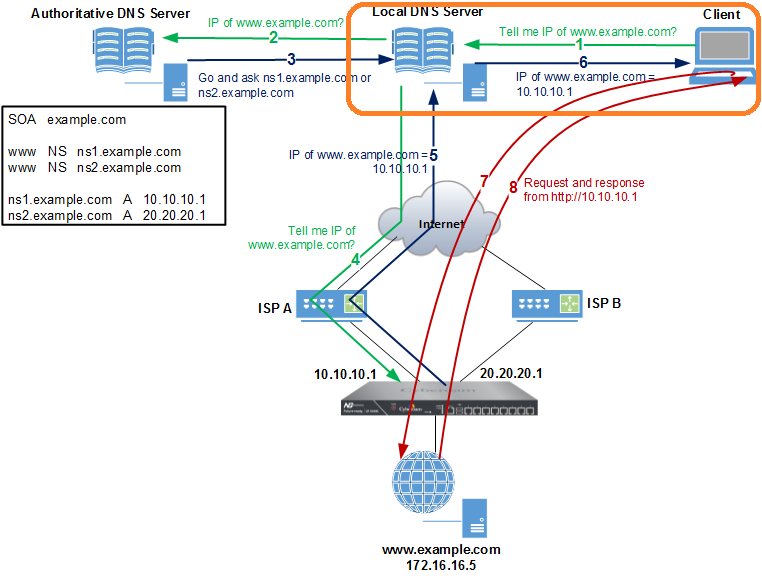
3）健康的主机没有理由访问纯粹是为了恶意行为的域名



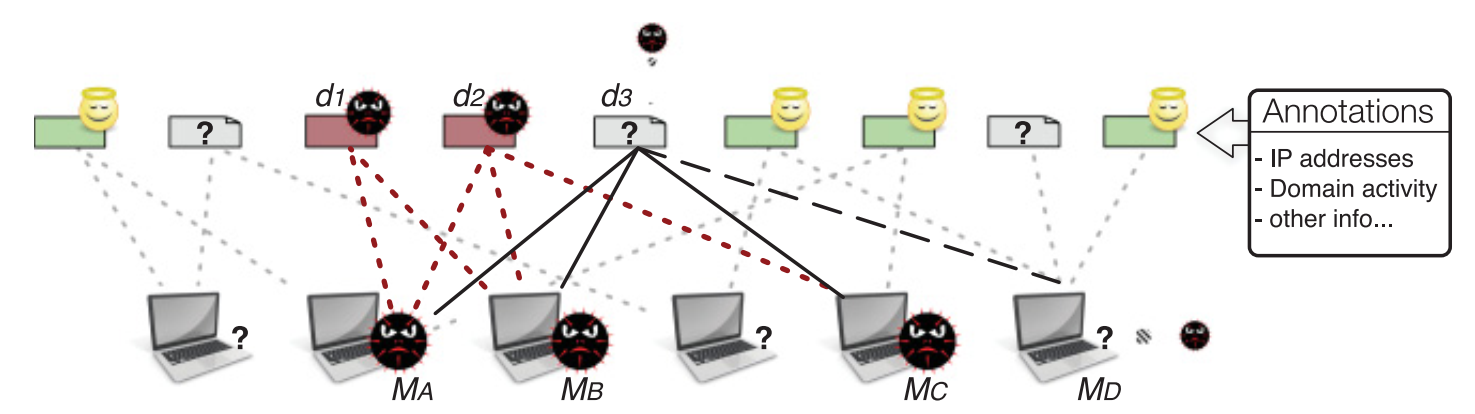
一天的DNS观测数据显示， 被感染主机访问malware-control域名的分布情况，70%左右的感染主机访问一个以上的恶意域名

Segugio基本的原理:

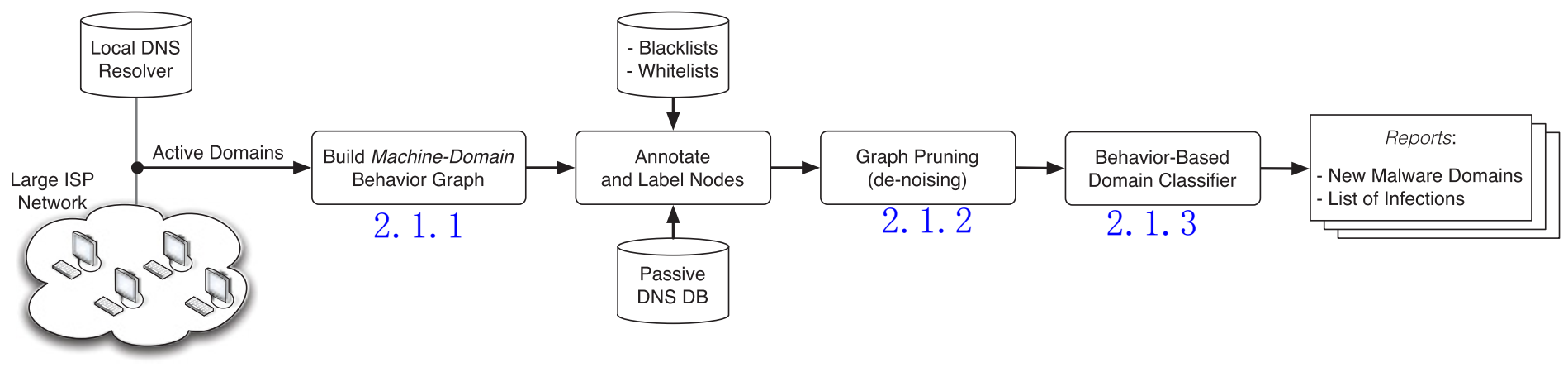
Segugio使用的数据是用户主机与ISP本地DNS解析器之间的DNS流量



通过观察谁请求了谁,我们能推断d3很有可能是恶意域名, 而M\_D很有可能是被感染的主机.



### Segugio系统结构图:



#### 1主机-域名行为图(Machine-Domain Behavior Graph)

通过观察大规模ISP网络, 观察一段时间后构建 机器-域名图 来描述谁请求了什么. Segugio构建无向二部图G=(M,D,E), M表示主机构成的node集合, D表示域名构成的node集合. 在一个给定的窗口期T内（例如1天）, 如果某个主机访问了某个域名, 我们就用将和连接起来.



**节点标记：**

1. 增大节点, 通过记录T时间段内此域名解析到的IP地址集合的壮大;
2. 对节点标记malware(利用公开或私有的C&C域名blacklist), benign(top 1M alexa.com的2LD), unknown(其他标记为unknown);
3. 标记请求Mal-ctrl域名的machine为infected, 标记只请求benign的machine为benign.

#### 2图修剪

上千万的machine，几亿个不同的域名，潜在的数十亿的边都需要，需要修剪图来提高性能降低噪音。

修剪规则如下：

R1）识别并丢弃不活跃的请求主机（去掉请求域名数<5的）;

R2）去掉domain请求数 > 的主机，设置为一个主机可能访问域名数的99.99%，这种主机可能是大的代理或企业网络的DNS承载者。

R3）去掉只被一个主机访问过一次的域名

R4）如果一个域名被整个网络中的1/3主机都访问，那么抛弃掉这个域名（不可能是恶意域名）。

**附加条款**： 如果主机被标记为malware-control，它不会被R1剔除；被标记为mal-ctrl的域名不会被R3剔除（例如NXdomain）；

### 基于访问行为的分类器(Behavior-Based Domain Classifier)

此分类器需要三组特征

F1）Machine Behavior：



S：代表所有请求域名d的主机构成的集合

I：被感染的主机构成的集合，是S的子集

U：不确定是否被感染，是S的子集

三个特征: m=|I|/|S|, u=|U|/|S|, t=|S|

m越大说明d是恶意域名的可能性就越大。

F2）Domain Activity：

如果一个新发现的域名被一些感染主机请求，那么它很有可能是恶意域名，但是靠查注册信息不靠谱，因为有些域名注册了但没使用处于休眠状态。本文侧重捕捉域名活跃度，提取一些特征（2LD域名的特征）

1. 二向图建立时间
2. 观察窗口宽度n天，从往前推的n天的时间窗[]
3. 在观察窗口内域名d的有效活跃天数(d活跃是指d被访问)
4. 域名d的连续活跃天数

F3）IP Abuse:

1. A: 在窗口期T内域名d解析到的IPs集合
2. A中有多少IP在过去曾经指向过恶意软件控制的域名
3. 在一个窗口期W(eg. W=5 months)测量A中与恶意域名关联的比例
4. 在W周期内, /24前缀中指向恶意域名的比例
5. 在W周期内, 指向未知域名的IP和/24的数量.

注：在Notos和exposure中使用了一些与F3）相似的信息，可能对本文方法有用但不是至关重要的。

**训练数据集**：这里提到了一个训练数据集的方式，选取一个标记好的主机-连接-域名的图数据集，去掉其中的域名d的标记，统计d的F1～F3的特征，然后再把d原来的label打到这个特征向量上。不断重复这个过程，可以让数据集更完备。这个步骤就是训练数据集的过程还没有涉及到classifier。本文并没有给出这种数据训练方式的参考文献。

实验条件的设定：

1. 数据来源：

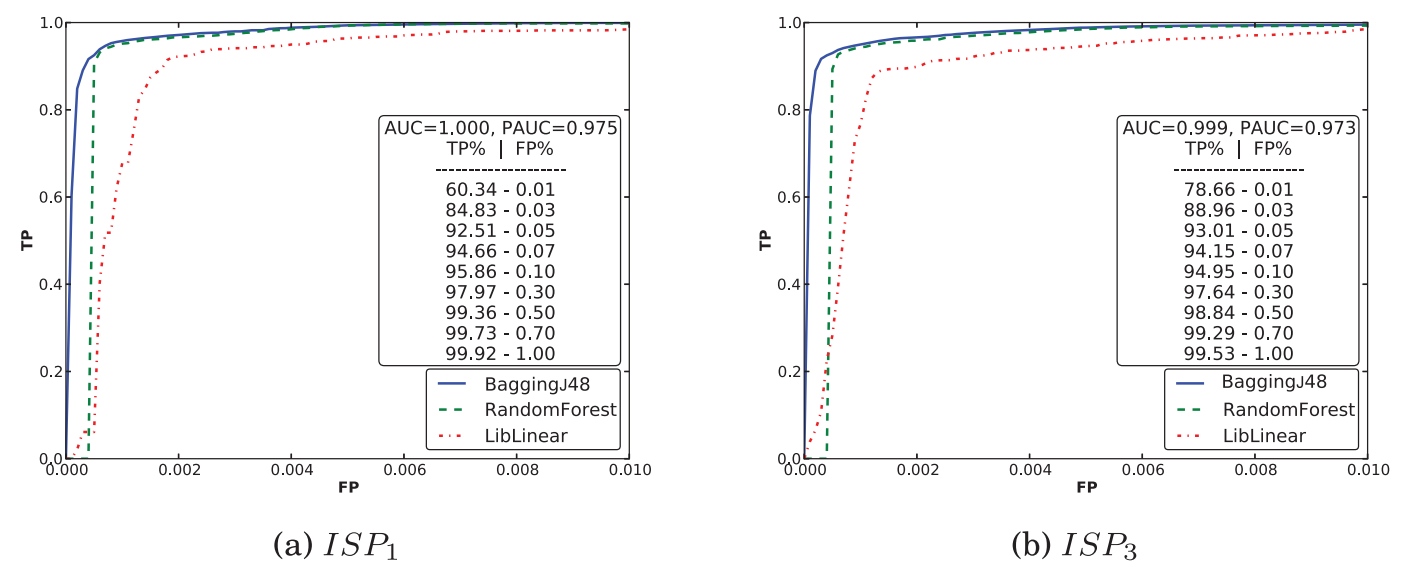
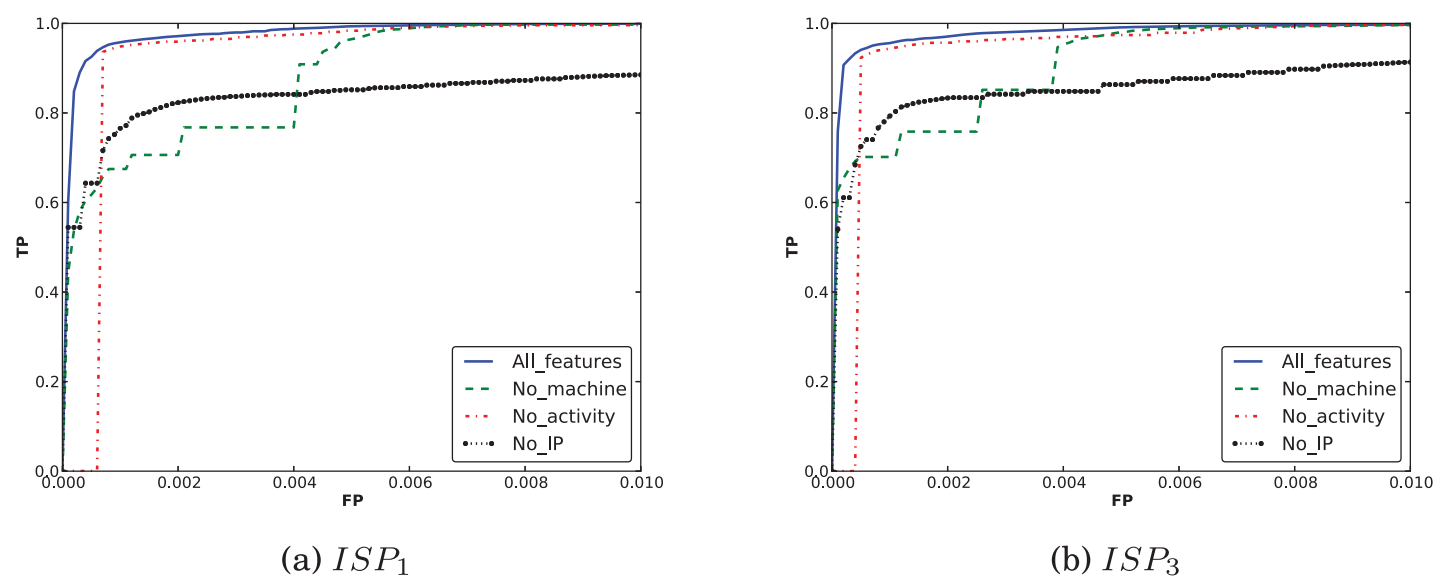
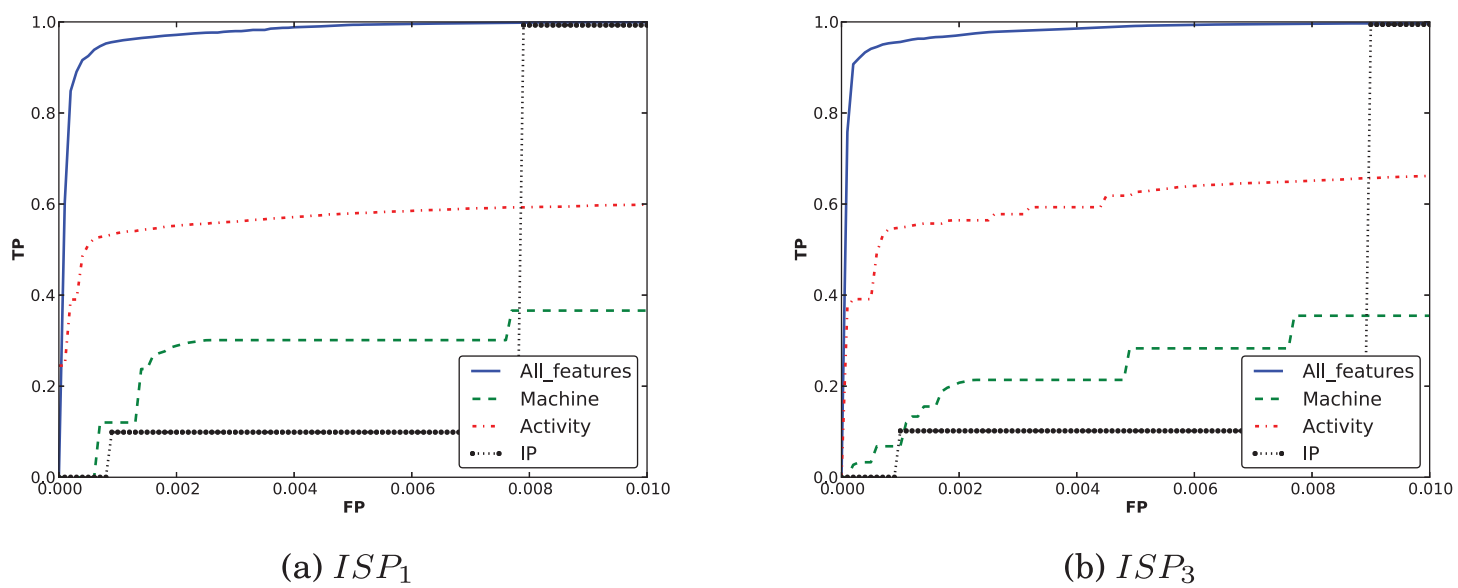
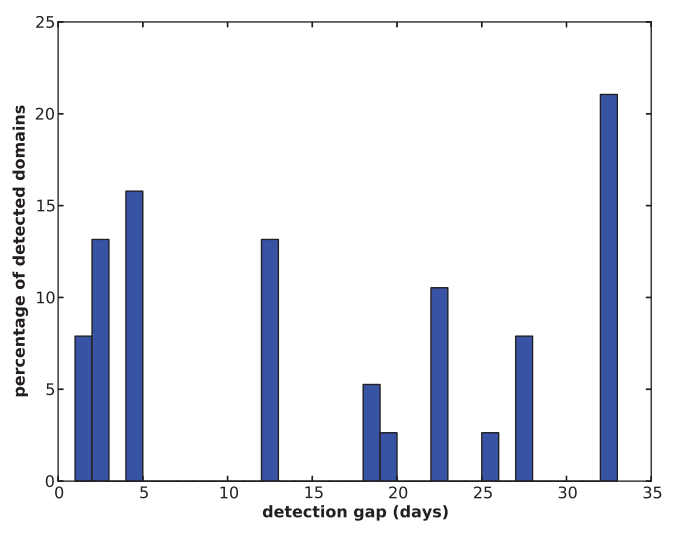
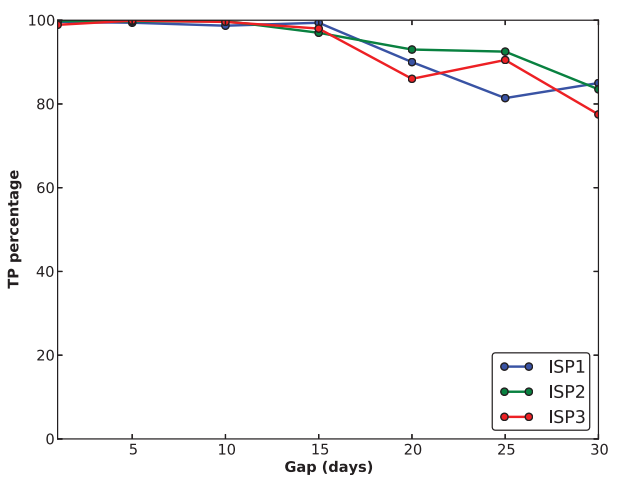
3个ISP：一个在美国东南部ISP\_1，一个在美国西南海岸ISP\_2，一个在美国西部ISP\_3。每天粗略的观测到1M~4M的机器识别码(这里使用机器识别码,不会受到DHCP的影响).

2）域名-机器 打标签:

Malware：商业黑名单;

benign: alexa.com一年的top1M的名单,包括458564二级域名.

实验结果：

1. 交叉验证测试  
   使用10重交叉验证，TP能实现95%，FP能实现0.1%  
   
2. 不同特征对实验结果的影响  
     
   
3. 恶意域名的提前发现  
     
   恶意域名被Segugio发现的时间与它出现在blacklist的时间间隔所占的比例。
4. 检测精度  
     
   对于一个训练好的Segugio（后续不再训练它），它的检测精度随着时间推移的变化情况。