各位老师你们好，我是兰雨晴老师的学生王文博，我的题目名称是恶意域名检测技术研究

以下是我开题ppt主要内容

首先通过介绍研究背景来说明我的题目意义，说明我为什么要做这个题目

第二介绍国内外的研究现状

第三是针对第一部分提出的问题而制定的研究内容与方案

最后我将简要说明进度安排与其他工作

首先，什么是恶意域名，

百度百科上查到的是故意在计算机系统上执行恶意任务的病毒、蠕虫和特洛伊木马的非法网站的域名这个解释很不全面，例如僵尸网络控制服务器的域名就不在此范畴之内

加州大学圣巴巴拉分校的Bilge给出的解释是涉及恶意活动的域名这个范围又太大了，受害主机域名或者很久之前被恶意使用过的域名也包含了进去

因此我这里定义恶意域名为当前存在恶意行为或者被恶意使用的域名

接下来我将从攻击和对抗两个方面来细致说明我为什么要做这个题目

首先是种类繁多的攻击手段

域名阴影是指利用合法域名下的大量子域名进行恶意行为

僵尸网络一种感染大量主机的传播行为

钓鱼网站伪装并窃取信息的一种恶意活动

DNS放大DDoS攻击是利用DNS协议的DDoS攻击

域名停放服务本身是合法的，但是往往这些网站存在商标滥用行为或者跳转到其他恶意网站的链接

相关对抗手段有信誉系统、逆向工程、机器学习

这里是两个恶意域名应用事例

第一个是Locky，一种勒索木马，主机感染之后将使用域名生成算法，从控制服务器上下载公钥对主机内指定文件进行加密处理，危害极广，实验室的韩涛学姐电脑就受到了感染

第二个是DNS放大DDoS攻击，放大效果可以达到100倍

这一页是主要对抗系统的建设，所使用方法就是之前提到的信誉系统、机器学习和逆向工程

接下来我简要说明这些方法的优缺点

利用恶意域名和DNS的攻击如此多样，对抗手段也存在着诸多不完备的地方，因此我要做的研究工作是意义极大的

这里根据我之前的一些研究，点出三个我认为需要解决问题

然后是第二部分，国内外研究现状

首先是国外研究现状

上一部分提及的经典系统中有近一半是他们的作品

商业机构中需要重点说明的是Nominum

构建了现在最完备、最准确的域名监测系统

这一页是国内研究现状

研究机构主要有中科院信工所、CNCERT

第三部分是研究内容与方案

针对之前提出的三个问题，制订了以下三个研究内容

其中原始数据的收集主要包括被动DNS数据和Whois数据收集，

被动DNS数据来源各省递归DNS服务器，

Whois数据主要来源自网络上的爬取和数据交换。

利用被动DNS和已有黑名单实时地对DNS异常进行检测，发现活跃的异常数据，并进行分析。

同时原始数据经过简单过滤、提取存储起来。域名打分系统利用历史数据进一步加工，例如提取一个时间片段内的域名访问量、访问IP数、网络数量等，按照一定规则进行计分，按照结果来更新黑白名单。

最后综合利用历史数据、域名打分结果以及黑白名单进行操作，实现恶意域名准确检测。

接下来将针对这三个部分重点说明其关键技术难点

这个猜测的来源是，我之前尝试过利用pleiades中的方法来检测山西省恶意域名，准确率和召回率比论文结果要差很多

而第一部分提到的经典系统在美国和欧洲很多大型ISP检测结果都比较好，所以我排除了过拟合的问题，至少在国外的网络环境中不是过拟合的问题

所以我有了这样的猜测，希望能利用打分系统来构建具有中国地方特色、汉语语言特点的白名单

第二是尝试深度学习在恶意域名检测中的应用

之前论文中并没有深度学习在恶意域名中的应用，但是Nominum公司的系统已经使用了神经网络，我认为这也是一个很好的探索方向。

这是最终的评价指标

最后一部分是