# 前沿研究方向规划

目录

[前沿研究方向规划 1](#_Toc509850396)

[AI与文件检测引擎 1](#_Toc509850397)

[背景 1](#_Toc509850398)

[1、PE恶意文件检测引擎 1](#_Toc509850399)

[2、深度神经网络与恶意文件分析 1](#_Toc509850400)

[3、APK检测 1](#_Toc509850401)

[4、HSO 1](#_Toc509850402)

[5、恶意文档检测 1](#_Toc509850403)

[6、社会工程学 1](#_Toc509850404)

[安全沙箱 1](#_Toc509850405)

[背景 1](#_Toc509850406)

[1、沙箱逃逸的对抗分析 1](#_Toc509850407)

[蓝盾数据中心 1](#_Toc509850408)

[大数据分析平台 1](#_Toc509850409)

AI与文件检测引擎

## 背景

SDN全称是Software Defined Network,即软件定义网络，是一种网络设计理念，或者一种推倒重来的设计思想。只要网络硬件可以集中式软件管理，可编程化，控制转发层面分开，则可以认为这个网络是一个SDN网络。NFV全称是Network Function Virtualization，即网络功能虚拟化，关注网络转发功能的虚拟化和通用化，将网络功能从专用设备(硬件架构不一定是X86)迁移到通用X86服务器上运行的虚拟机或者容器中。

NFV和SDN如果相互结合，又可以是互补的存在。借助SDN，不仅传统的NF连接方式都能支持，SDN还能提供更高效的NFV实现方式。毕竟SDN提供的管理层和转发层的分离，使得网络变得极其灵活。反过来，NFV也能够提供SDN的运行环境，帮助SDN的实现。

## 1、PE恶意文件检测引擎

P4是一门编写协议无关的包处理器的高级语言，是一种声明式编程语言，主要用于编程程序以下达指令给数据转发平面的设备(如交换机、网卡、防火墙、过滤器等)如何处理数据包。通过P4我们可实现：定制数据解析流程、可执行并行和串行的匹配+动作操作、支持协议无关的转发。通过这些特性，我们可以更好的控制数据流向，并在高性能的保障下为公司的业务提供更强的功能。

## 2、深度神经网络与恶意文件分析

当前SDN控制器种类繁多，主流的控制器有14种。其中最为出名的是由服务商共同参与的开源OpenDaylight（ODL）项目及运营商主要关注的ONOS项目。我们将基于现有项目的接口或开源的代码，快速实现公司的产品功能。由于控制器的更新较快，我们将持续关注新功能的推出，以求更好的利用到产品中来，同时降低开发成本。

## 3、APK检测

## 4、HSO

## 5、恶意文档检测

### 1. 背景

网络攻击者正在转向基于文件的恶意软件，因为越来越多的攻击者使用邮件附件和网络链接。用户通常会被浏览器，电子邮件代理或AV产品更多地警告可执行文件的危险，而像PDF这样的文档由于受到静态文件的影响而受到的关注和审查要少得多，并且几乎没有什么危害。

但是，随着时间的推移，PDF规格已经改变。增加的脚本功能可以使文档以与可执行文件几乎相同的方式工作，包括连接到Internet的能力，运行进程以及与其他文件/程序进行交互。内容复杂性的增长为攻击者提供了更多的武器来发动强大的攻击，并且更灵活地隐藏恶意有效载荷（例如，加密，隐藏为图像，字体或Flash内容）并逃避检测。

### 2. 静态分析

#### ● 基于签名的

基于签名的检测技术主要基于模式匹配的思想，为每种已知恶意代码产生一个唯一的签名特征标记来创建恶意代码库。这些签名包括很多不同的属性，如文件名、内容字符串或字节等，并且也从排除这些恶意代码所产生安全漏洞的角度来探讨保护系统安全。将未知代码的签名特征与恶意代码库进行对比，搜索恶意代码库查找时候存在相匹配的恶意代码签名，若有吻合，则判定为恶意代码；反之则判断为正常代码。

这种方法是最方便应用最广的检测方法，商业的杀毒产品都是采用这种技术。其优点是检测速度快，凡是病毒库中已有的恶意代码，全都能够准确检测出来，误报率较低。缺点是对于未出现过的病毒无能为力，面对不断出现的新病毒，必须不断更新版本，否则检测工具变回老去，逐渐失去实用价值，因此在将来某一种恶意代码的签名加入恶意代码签名库之前，计算机系统可能会遭到此类恶意代码的攻击。

#### ● 机器学习检测

机器学习(Machine Learning, ML)是一门多领域交叉学科，它是[人工智能](https://baike.baidu.com/item/%E4%BA%BA%E5%B7%A5%E6%99%BA%E8%83%BD)的核心，是使计算机具有智能的根本途径，Tom Mitchell的机器学习(1997)对[信息论](https://baike.baidu.com/item/%E4%BF%A1%E6%81%AF%E8%AE%BA)中的一些概念有详细的解释,其中定义机器学习时提到，“机器学习是对能通过[经验](https://baike.baidu.com/item/%E7%BB%8F%E9%AA%8C)自动改进的计算机算法的研究”。（Machine Learning is the study of computer algorithms that improve automatically through experience.）

综合考虑各种学习方法出现的历史渊源、知识表示、推理策略、结果评估的相似性、研究人员交流的相对集中性以及应用领域等诸因素。将机器学习方法 区分为以下六类

1. **经验性归纳学习 (empirical inductive learning)**
2. **分析学习（analytic learning）**
3. **类比学习**
4. **遗传算法（genetic algorithm）**
5. **联接学习**
6. **增强学习（reinforcement learning）**

### 3. 动态分析

全面的动态分析。值得注意的是，只有一种动态检测方法对整个文件进行分析，相反，其余的方法只对提取的JavaScript代码执行分析。这与传统的沙盒式恶意软件分析（如Cuckoo或CWSandbox）形成鲜明对比，后者执行恶意软件，并在运行时检查其行为和对主机操作系统的影响。一个原因可能是因为maldoc运行在AAR之上，AAR本身是一个复杂的软件，在主机系统上留下了很大的空间。

针对对整个文档的动态分析，监测起调用的函数与返回值，进而来对整个恶意文件的检测提供更有利的证据，动态分析在调用外部链接时起到了很好的检测效果，同时可以增加恶意文档的检测率。Cuckoo在安装有PDF阅读器的虚拟机中运行提交的示例，并报告示例的行为，包括调用的网络API及其参数。通过对动态行为的检测来提取一些动态行为进行分析取证。

## 6、社会工程学

社会工程学是一种通过人际交流的方式获得信息的**非技术渗透手段**。不幸的是， 这种手段有效， 而且效率很高。 事实上，社会工程学已是企业安全最大的威胁之一。对hacker来说，通过网络远程渗透破解获得数据，可能是最为麻烦的方法。一种无需电脑网络，更注重研究人性弱点的黑客手法正在兴起，这就是社会工程学攻击。

### ●　APT　攻击

APT（Advanced Persistent Threat）是指高级持续性威胁。 利用先进的攻击手段对特定目标进行长期持续性网络攻击的攻击形式，APT攻击的原理相对于其他攻击形式更为高级和先进，其高级性主要体现在APT在发动攻击之前需要对攻击对象的业务流程和目标系统进行精确的收集。在此收集的过程中，此攻击会主动挖掘被攻击对象受信系统和应用程序的漏洞，利用这些漏洞组建攻击者所需的网络，并利用0day漏洞进行攻击。

APT攻击是近几年来出现的一种高级攻击，具有难检测、持续时间长和攻击目标明确等特征。2016年6月在DNC电脑系统中曾发现APT28的渗透痕迹，同年2月孟加拉国央行被黑客攻击导致8100万美元被窃取的事件被曝光后，如越南先锋银行、厄瓜多尔银行等，针对银行SWIFT系统的其他网络攻击事件逐一被公开。

# ●　用户实体行为分析(UEBA)

**用户与实体行为分析(UEBA)**已经在过去的几年得到了迅速的发展。最开始它只是用户行为分析，关注与捕捉内部恶意威胁。 今天UEBA已经不再仅仅是一个单点解决方案，它已经被视为从内部威胁，到安全信息事件管理工具，再到网络风险分析和终端保护的重要因素。

**通过把UEBA技术同其他数据资源和分析方法结合起来，机构越来越方便在更大的尺度上解决网络风险挑战。**比如，他们使用UEBA分析异常用户和机器行为的重叠作为攻击指标，以识别有问题的账户；或者用UEBA确定基于供应商的内部威胁，并且这方面信息和其他的风险智能结合起来，以获得一个360度的第三方风险视角。这一视角可以被安全和供应商风险管理参与方用来减少侵入公司网络和信息的外部威胁。

伴随2018年5月生效的《通用数据保护规则》(GDPR)，保护雇员、顾客和股东的私人数据的关注点也会更新。**UEBA将被用于检测处理不当的敏感数据**，这些没有被妥善处理的数据会导致公司在合规性方面违反通用数据保护条例。这包括了解人们在访问什么、异常访问、异常数据分类水平处理、异常解密和异常电子邮件和云上传的行为模式。

安全沙箱

## 背景

区块链是对过往技术的重新组合，是现阶段信息科技创新的高潮。作为分布式账本的一种，区块链兼具去中心化和环环相扣的特点。因为区块链的特性，很多企业或者说机构都看到了区块链在未来发展中的效用。而且从近期国家频频出现的动作来看，并不是只有银行及支付行业被区块链科技影响到，很多领域都在慢慢涉及。

区块链本质上是一个去中心化的分布式账本数据库。其本身是一串使用密码学相关联所产生的数据块，每一个数据块中包含了多次比特币网络交易有效确认的信息。

## 1、沙箱逃逸的对抗分析

区块链已经成为了当前的热点所在。如何寻找到该技术和行业的密切结合点，推出先进的产品，在市场上产生巨大的影响力。

蓝盾数据中心—木梯

大数据分析平台—待定