**蓝盾前沿技术分享会会议记要**

会议时间:

* 2018年8月9日 14:00 – 16:00

会议主题：

* 基于人工智能的恶意软件检测与对抗性学习，以及数据服务平台建设

## 会议流程：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 时间 | 演讲主题 | 演讲者 |
| 1 | 14:00 - 14:20 | 从恶意软件检测到对抗性学习 | 江纬 |
| 2 | 14:20 - 14:35 | 数据服务平台 | 王木梯 |
| 3 | 14:35 - 14:50 | 深度学习模型的对抗 | 陈俊航 |
| 4 | 14:50 - 15:05 | 基于人工智能的恶意文档检测与对抗性学习 | 王凤娇 |
| 5 | 15:05 – 15:20 | 双语文档及语料库管理——标准化&内容管理 | 魏舒敏 |
| 6 | 15:20 – 15:40 | 画像的得失 | 刘峰 |
| 7 | 15:40 – 16:00 | 技术讨论与交流 | 全体 |

## 会议内容和讨论：

1. 从恶意软件检测到对抗性学习—江纬 ：
   1. 主要内容
      * 1. 恶意软件检测(难点，研究进展)
        2. AI模型的安全 —— 对抗性学习（未来研究的重点）
        3. 数据科学团队建设
   2. 讨论
      * 1. 对抗性学习怎样与产品应用结合？

在一些安防领域，图片识别领域等，以提高模型的安全性。

* + - 1. 现有的研究成果（如PE模型，APK模型，PDF模型）怎样结合到防火墙，终端安全产品中？

公司十月份后会对现有的模型进行技术评估，以决定是否将现有研究成果结合到产品当中。

* + - 1. 深度学习能否／怎样运用于终端产品中？

目前深度学习对计算资源要求较高，有两种解决思路：1）模型的裁剪与简化，2）加入GPU提高终端计算力

1. 数据服务平台—王木梯：
   1. 主要内容
      * 1. 数据基础平台（CPU:42 Cores；内存：106G； 存储：34T）
        2. 病毒数据样本来源渠道与共享（https://admin.qy.com）
        3. 病毒数据的解析与分类处理（病毒家族分类）
        4. 病毒样本数据服务扩展（沙箱分析报告，病毒危害等级分析，逆向分析报告，数据可视化，cert年度报告等）
        5. 安全四大顶级会议的相关学习资料与视频共享平台搭建
   2. 讨论
      * 1. 都同意数据收集的重要性，并鼓励与支持继续做好数据收集的工作
        2. 能否收集关于CVE的漏洞信息解决方案（建立漏洞信息库）？

审计系统里面希望提供 哪些病毒运用的CVE，以及对应的 CVE和补丁信息，目前还没有关于CVE漏洞信息收集，后续增加相关人员进行辅助会加入漏洞信息库。目前主要的研究方向还是漏洞检测。

* + - 1. 在数据收集的同时，能否做到数据（如情报威胁）的关联分析，并且做出报告，提出解决方案？

理论方法是可行的，可是鉴于目前人员配备还不够完善，扩展更多的数据源和数据，我们也正在实践当中。

* + - 1. 数据处理结果提供API对其他开发人员调用。

1. 深度学习模型的对抗－陈俊航
   1. 主要内容
      * 1. 黑盒攻击与白盒攻击
        2. 提高模型健壮性的防守策略（对抗学习，扰乱梯度，降维）
        3. 对抗样本的检测方法（Sample Statistic, Training a Detector, Prediction Inconsistency）
   2. 讨论
      * 1. 对抗性学习有哪些应用场景？

在人脸识别，自动驾驶领域等，以提高模型的安全性。可以应用于子公司安防业务当中。

1. 恶意文件检测与对抗性学习 – 王凤娇
   1. 主要内容
      * 1. 传统AI模型的建立
        2. 对抗性学习逃逸AI模型
        3. AI模型如何抗逃逸
        4. 模型的应用场景
   2. 讨论
      * 1. 从产品的出发点来看将PDF检测模型放入到出入口检测器
        2. 针对于有恶意性质的，并且具有逃逸性的恶意文件，就标识为有重大危险的标签，这种文件的危害会更大，并且不容易被检测到（本质是检测对抗样本，PDF能否像软件基因一样抽取到不易攻击的内容。）关于恶意文件找到原始变异样本的问题目前正在研究当中，之前一直的研究重点是在AI模型的对抗。
2. 双语文档及语料库管理(标准化&内容管理)– 魏舒敏
   1. 主要内容
      * 1. 技术用语标准化管理（产品界面中文用于一致性，建立各产品预料库，语言包）
        2. 英文技术文档阅读支持
   2. 讨论
      * 1. 建立语料库是一个很大又重要，同时需要各产品线的同事相互配合，同时改善工作效率
3. 画像的得失 – 刘峰
   1. 主要内容
      * 1. 基于标签体系实现用户画像
        2. 安全领域画像的问题
        3. 深度学习与经典机器学习（深度学习-必然趋势）
   2. 讨论
      * 1. 结合各自业务场景，选择合适的方法

## 会议总结：

1. 建立数据科学团队的重要性（举例：前沿组的人员构成）
2. 数据架构体系构建的重要性（加强数据收集、存储，结构与非结构化存储，机器学习算法及调优等部分）
3. 对抗性学习在安防领域提前布局的重要性（基于图片的研究）
4. 对抗性学习对于AI模型安全的重要性
5. 在全公司范围内建立语料库的重要性（举例：防火墙UI的不统一）
6. 研究与工程紧密结合的重要性（举例：态势感知）
7. 数据中心的建立，现在各种方法都在兴起，可是所有的方法都依赖于数据，大家也表示，数据中心的建立与数据处理，是一个长久的战略
8. 前沿技术与产品紧密交互，引导与支持产品开发
9. 威胁情报云的建立，加强云的输出能力