登录

AI for Security：智能化安全对抗的困境

4o4notfound / 2018-12-06 06:54:00 / 浏览数 6808 安全技术 技术讨论 顶(3) 踩(0)

本篇文章主要结合自己的亲身经历和体会表达一下对AI for Security的看法，整体态度偏消极。

0x01 背景

本段将由大到小从国家、网上公开资料、项目三个方面说明智能化安全对抗时代的来临。
正如有时候觉得学得好不如嫁得好，在网络安全方面有时也是学得好不如运气好跟上了历史的进程。网络战早已被美军纳入作战体系，现如今中美关系对抗愈加激烈，随之而

0x02 智能化安全对抗

AI和安全之间的关系可分为四点：

- 应用安全
  - 用AI来做应用安全防护
  - 用AI来攻击应用
- AI本身安全
  - (用AI)攻击AI(模型、系统、框架)
  - AI模型防护

AI是手段，目标是安全，包括应用安全和AI安全。所以是AI for Application and AI Security，即AI for Security。

从应用安全的角度来说，用AI来做应用安全防护，比如做WAF或入侵检测，属于基础安全，做反欺诈检测，属于业务安全；用AI来攻击应用，比如用AI来自动化漏洞挖掘，

四者相辅相成，假如我们在防守位，正遭遇外界的攻击，包括传统攻击和AI攻击，如果我们用AI来做应用安全防护，对手可能采用对抗样本来攻击我们的AI，那么我们就要对

0x03 学术界和工业界前沿智能化安全对抗

学术界研究者近几年比较关注对抗样本等智能化对抗技术，在四大顶会上也不断发表了系列关于机器学习和安全对抗的文章，有传统的机器学习安全检测文章，也有机器学习
Security 2018和IEEE S&P
2019，代表着学术界最新的研究进展。可以看出学术界的研究点开始转向了关于机器学习的对抗性技术研究。前几天去参加的XDef2018会议，对阿里巴巴集团安全部资深算
for
Security的理解和做法，可能由于需求导致出发点不同，该议题重点在于防：AI安全检测和抵抗机器学习流量。AI作为一种手段，可以降维打击传统攻击，抵抗同级的AI攻击

我们现在容易获取的资料主要集中在是AI安全检测方面，大部分文章都在自娱自乐，反反复复换汤不换药，都是同样的套路，未加以优化，质量低，只适合做做实验，很少在

0x04 我的AI for Security 学习笔记

针对上面分析到的问题，我筛选整理了一份安全场景、基于AI的安全算法和安全数据分析学习资料，希望对AI for Security感兴趣的朋友们有所帮助。

AI-for-Security-Learning

安全场景、基于AI的安全算法和安全数据分析学习笔记（工程类学习笔记，不包含论文、书籍、视频等，不花里胡哨嘻嘻哈哈）

项目地址：https://github.com/404notf0und/AI-for-Security-Learning

最近更新日期为：2018/12/2

同步更新于：404 Not Found：AI for Security

目录：

    - Webshell检测
    - Other

    杂项
    - WindowsRDP检测
    - PowerShell检测
    - 用户行为(UBA)检测
    - 弱口令检测
    - 安全运营
- (使用AI)保护AI(框架、数据、模型、系统)

  对抗篇
- 使用AI攻击应用

  (使用AI)攻击AI(框架、数据、模型、系统)
    - 攻击AI框架
    - 攻击AI模型

防护篇

使用AI保护应用

恶意软件和代码

- 深度学习在恶意软件检测中的应用
- 恶意软件与数据分析
- 利用机器学习进行恶意代码分类
- 用机器学习检测Android恶意代码
- Malware Detection in Executables Using Neural Networks
- 基于深度学习的恶意样本行为检测(含源码)

恶意流量检测

- 利用机器学习检测HTTP恶意外连流量

域名安全检测

- 使用fasttext进行DGA检测
- 使用CNN检测DNS隧道
- 机器学习与威胁情报的融合：一种基于AI检测恶意域名的方法

业务安全检测

- 基于设备指纹的风控建模以及机器学习的尝试
- 如何在安全风控中评估和量化机器学习有效性
- 人工智能反欺诈三部曲——特征工程

Web安全检测

Web安全之URL异常检测

- 基于机器学习的web异常检测
- 基于大数据和机器学习的Web异常参数检测系统Demo实现
- 基于机器学习的web应用防火墙
- LSTM识别恶意HTTP请求
- 基于URL异常检测的机器学习模型mini部署
- 我的AI安全检测学习笔记（一）

Web安全之XSS检测

- 机器学习识别XSS实践
- 使用深度学习检测XSS
- 使用深度学习检测XSS(续)

Web安全之攻击多分类检测

- 基于机器学习的WEB攻击分类检测模型

- 基于机器学习的攻击检测系统

Web安全之Webshell检测

- 基于机器学习的分布式webshell检测系统-特征工程（1）
- 深度学习PHP webshell查杀引擎demo
- 使用机器学习识别WebShell
- 基于机器学习的分布式Webshell检测系统
- GitChat · 安全 | 基于机器学习的 Webshell 发现技术探索
- 刘焱： Webshell 发现技术实战解析
- 安普诺张涛：再谈webshell检测
- 新开始:webshell的检测
- 基于机器学习的WebShell检测方法与实现(上)

Web安全之其他

- Web安全检测中机器学习的经验之谈

杂项

- 机器学习在WindowsRDP版本和后门检测上的应用
- 用机器学习检测恶意PowerShell
- 机器学习算法在用户行为检测(UBA)领域的应用
- 利用机器学习和规则实现弱口令检测
- 解决机器学习和安全运营之间的最后一公里问题

保护AI

- 如何利用AI对抗"数据污染"和"数据中毒"？
- 对抗数据中毒--机器学习在阿里巴巴网络安全的应用

对抗篇

使用AI攻击应用

- AI与Android漏洞挖掘的那些事儿
- AI与安全的恩怨情仇五部曲「1」Misuse AI
- 一种基于机器学习的自动化鱼叉式网络钓鱼思路

攻击AI

攻击AI框架

- 深度学习框架中的魔鬼——探究人工智能系统中的安全问题

攻击AI模型

- 安全领域中机器学习的对抗和博弈
- 基础攻防场景下的AI对抗样本初探
- 机器学习在安全攻防场景的应用与分析
- 使用生成对抗网络(GAN)生成DGA
- 详解如何使用Keras实现Wassertein GAN
- Is attacking machine learning easier than defending it?
- 对深度学习的逃逸攻击 ——探究人工智能系统中的安全盲区

0x05 对现状的个人理解

从上面我收集的资料来看，我觉得智能化安全对抗技术的第一点困境是目前易工程化的还主要集中在AI安全检测方面这种弱智能化安全对抗技术，但是大规模应用还是较难，

0x06 Reference

Machine Learning Blog:

https://plushunter.github.io

http://phoebepan.cn

http://scarletpan.github.io

http://www.jeyzhang.com

Machine Learning+Security Blog:

http://webber.tech/

http://bindog.github.io

https://www.cdxy.me

https://iami.xyz

https://www.zuozuovera.com/

https://www.cnblogs.com/LittleHann

上一篇：RWCTF-Magic Tunne... 下一篇：输入长度受限情况下的 XSS 攻击

1. 0 条回复
   - 动动手指，沙发就是你的了！