人工智能与安全概述(AISec ' 13)

波茨坦大学 4 号楼，
办公室 0.20
August-Bebel-Str。89
14482 波茨坦，德国
bnelson@cs.uni-
Potsdam .de

Christos
Dimitrakakis 计算机
科学与工程
查尔姆斯大学 SE-
4172
chrdimi@
chalmers.s
e

伊莱恩·史
计算机科学系
马里兰大学
A.V.威廉姆斯大厦学院
公园，马里兰州 20742
runting@cs.cmu.edu

摘要

人工智能与安全研讨会(AISec)侧重于人工智能(AI)和机器学习在诸如安全和隐私应用等对抗性环境中的理论和应用，反过来，通过使用大规模人工智能方法产生的安全和隐私影响。该研讨会是应用程序、算法和理论融合的主要场所，并继续吸引来自不同领域的研究人员的提交，他们在这个不断增长的领域中解决新出现的问题。 AISec 为安全、隐私、人工智能和学习社区的研究人员提供了一个论坛，讨论智能技术在安全和隐私应用中的作用，并向人工智能和学习社区提出这些问题的独特需求。

类别和主题描述符

C.2.0[计算机通信网络]: 一般 - *安全和保护(例如防火墙);D.4.6[操作系统]*: 安全与保护;I.2.6[人工智能]: 学习;I.2.7【人工智能】: 自然语言处理;I.2.8[人工智能]: 问题解决、控制方法和搜索;K.4.1[电脑与社会]: 公共政策问题——*隐私*;
K.6.5[电脑及资讯系统管理]: 安全与保护

一般条款

算法、安全理论

关键字

人工智能，计算机安全，机器学习，计算机隐私，安全学习

AISec ' 13, 2013 年 11 月 4 日，德国柏林。Acm
978-1-4503-2488-5/13/11…
http://dx.doi.org/10.1145/2508859.2509031

为了解决问题和 AI/ML 研究人员开发大数据分析技术，*这*些社区只能将有限的注意力集中在另一个社区：在安全研究中，AI/ML 组件通常被视为黑盒解决方案。相反，学习社区很少考虑应用他们的算法所带来的安全/隐私影响。虽然这两个社区通常关注不同的方向，但当这两个领域相遇时，就会出现有趣的新问题。 这些已经为两个社区提出了许多新颖的问题，并创造了一个新的研究分支，即安全学习。在这个交汇处，AISec 工作坊已经成为这个独特的研究融合的主要场所。

1. 背景与动机

人工智能(AI)、机器学习(ML)和数据挖掘在安全和隐私问题上的潜在应用正在不断扩大。这些技术提供的分析工具和智能行为使得人工智能和学习在具有丰富数据或需要对不断变化的情况做出快速反应的领域的自主实时决策中变得越来越重要。特别是，这些智能技术为涉及通过云计算扩展的大数据分析*的*安全问题提供了新的解决方案。在安全敏感领域中使用学习方法为安全研究创造了新的前沿，在这些领域中，对手可能试图误导或逃避智能机器。AISec 研讨会提供了一个展示和讨论安全/隐私与人工智能和机器学习融合的新发展的场所。

AISec 连续第六年与 CCS 合作，是对安全、隐私、人工智能和机器学习结合感兴趣的研究人员的首选会议场所。过去的一年见证了 AISec 社区内的激烈活动-首先是 Dagstuhl 讲习班，然后是第五届 AISec 讲习班。这种激增有几个原因。首先，机器学习、数据挖掘和其他人工智能技术在从大数据中提取知识、态势感知和安全情报方面发挥着关*键作*用。其次，谷歌和亚马逊等以数据为中心的公司正在越来越多地探索和部署学习技术，*为*客户解决大数据问题。最后，这些趋势使公司及其客户越来越多地接触到智能技术。 因此，研究人员正在探索这些学习技术，作为安全/隐私问题的潜在解决方案，但也作为需要保护的新的潜在隐私/安全漏洞。

2. 研讨会的目标

AISec 研讨会是一个将实际安全问题与人工智能和机器学习的进步相结合的场所。 在此过程中，研究人员也在发展这一领域独特的理论和分析，并正在探索不同的主题，包括在博弈论对抗环境中的学习，隐私保护学习以及垃圾邮件和入侵检测的应用。人工智能和机器学习提供了一套有用的分析和决策技术，这些技术正在被不断增长的从业者社区所利用，包括在具有安全敏感元素的应用程序中。 然而，当安全研究人员经常使用这些技术时，情况又如何呢

3. 感兴趣的话题

在作者受众方面，我们征集了以下(但不限于)研究主题的论文提交:

学习理论与安全相关的主题

- 敌对的学习
- 健壮的统计数据
- 在线学习
- 随机博弈中的学习

安全应用

- 计算机取证
- 垃圾邮件检测
- 网络钓鱼检测和预防
- 僵尸网络检测
- 入侵检测与响应
- 恶意软件识别
- 身份识别
- 安全大数据分析

与安全相关的人工智能问题

- 安全的分布式推理和决策
- 安全多方计算和加密方法
- 保护隐私的数据挖掘
- 自适应旁信道攻击
- 验证码的设计与分析
- 人工智能接近信任和声誉
- 通过智能探测(例如，模糊测试)进*行*漏洞测试
- 内容驱动的安全策略管理和访问控制
- 生成训练集和测试集的技术和方法
- 异常行为检测(*例*如，为了防止欺诈、身份验证)

4. 项目委员会

我们感谢我们项目委员会的成员:

- 巴蒂斯塔·比吉奥，意大利卡利亚里大学
- Ulf Brefeld，德国达姆施塔特工业大学
- Michael bru¨ckner，亚马逊公司，德国
- Mike Burmester，佛罗里达州立大学，美国
- Alvaro A. C´ardenas，德克萨斯大学达拉斯分校，美国
- 马里奥·弗兰克，加州大学伯克利分校，美国
- Rachel Greenstadt，美国德雷塞尔大学
- 顾国飞，美国德州农工大学
- 黄凌，英特尔实验室，美国
- 安东尼·约瑟夫，加州大学伯克利分校，美国
- Ari Juels, RSA 实验室，美国
- 帕维尔·拉斯科夫，德国宾根大学
- 丹尼尔·洛德，俄勒冈大学，美国
- Pratyusa Manadhata，惠普实验室，美国
- Aikaterini Mitrokotsa，查尔姆斯理工大学，瑞典
- Roberto Perdisci，美国佐治亚大学
- Vasyl Pihur，谷歌公司，美国
- Konrad Rieck，德国哥廷根大学
- 法比奥·罗利，意大利卡利亚里大学
- Benjamin I. P. Rubinstein, IBM 研究院，澳大利亚
- Robin Sommer, ICSI 和 LBNL，美国
- 尼娜·塔夫脱，特艺彩色，美国
- J. D.泰格，加州大学伯克利分校，美国
- Shobha Venkataraman，美国 at&t 研究公司
- 颜廷芳，RSA 实验室

5. 研讨会的组织者

布莱恩·尼尔森是波茨坦大学的博士后研究员。他曾在宾根大学担任博士后研究员，并获得博士学位。 来自加州大学伯克利分校。 他是 2012 年 AISec 的联合主席，也是 2012 年 Dagstuhl 研讨会“机器学习方法-计算机安全方法”的共同组织者。他的研究重点是学习算法，特别是在安全敏感的应用领域。 他调查了学习对安全威胁的脆弱性，以及如何使用弹性学习技术来减轻这种安全威胁。

Christos Dimitrakakis 是查尔姆斯理工大学的一名研究员。他的主要研究兴趣是决策理论，包括强化学习和安全应用中的问题。 他获得了博士学位。 2006 年从 EPFL 毕业，并在莱奥本大学、阿姆斯特丹大学和法兰克福大学担任研究员。最近，他是 EPFL 的玛丽·居里研究员。他与 ECML 2010 合作组织了一个关于隐私、安全、数据挖掘和机器学习(PS- DML)的研讨会。

Elaine Shi 是马里兰大学帕克分校计算机科学系的助理教授。她的研究结合了系统安全、密码学和数据挖掘来设计安全且保护隐私的新计算系统。 她于 2008 年获得 Carnegie Mellon University 计算机科学系博士学位。 在加入 UMD 之前，她是 Xerox PARC 的研究人员，也是加州大学伯克利分校的研究科学家。Elaine Shi 曾在超过 25 个会议和研讨会的项目委员会任职，目前也是 AsiaCCS 的云安全研讨会的联合主席。她帮助组织了一个 NSF/Intel 赞助的安全课程研讨会，以及 2012 年 SaTC PI 会议的跨学科对话会议。