

《数据驱动安全》

图书基本信息

书名：《数据驱动安全》

13位ISBN编号：9787111512677

出版时间：2015-9

作者：杰·雅克布,鲍布·鲁迪斯

页数：291

译者：薛杰,王占一,张卓

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：www.tushu000.com

《数据驱动安全》

内容概要

本书由世界顶级安全专家亲笔撰写，深入剖析了安全领域中的数据分析及可视化方法，包含大量真实案例和数据。从安全数据收集、整理、分析、可视化过程，详细讲解如何设计有效的安全数据可视化，并走向数据驱动的安全研究。主要内容包括：第1章展示信息安全领域数据分析与可视化的基础知识，以及安全数据科学工作者需要掌握的技能概览。第2、3、4章分别介绍一些安全数据科学工作者需要掌握的软件工具、技术知识、使用技巧，涉及Python语言、R语言为主的实用分析方法。第5章介绍创建图表的技术以及一些核心的统计学概念。第6章讲解数据可视化的基础知识，以及有效展示的技巧。第7章介绍如何对安全漏洞进行分析和可视化，包含大量安全事件的真实数据。第8章涵盖现代数据库的概念，包含在传统数据库基础上新增的数据展示技巧以及NoSQL解决方案。第9章将带你进入机器学习领域，包括机器学习的核心概念，探索机器学习实现技术等。第10章及第11章介绍创建有效的可视化产品技巧，以及如何让这些信息的展示得更加丰富有形。在第12章呈现如何将所学的知识应用到实际的安全环境中。

《数据驱动安全》

作者简介

Jay Jacobs在IT以及信息安全领域拥有超过15年的经验，主要致力于密码学、风险学以及数据分析方面的研究。作为Versizon RISK团队的一名高级数据分析师，他参与编纂年度《Data Breach Investigation Report》，并投入大量精力进行安全相关数据的分析与可视化。Jay也是Society of Information Risk Analysts的创立人之一，现在是该组织董事会的成员。他是一名活跃的博客撰稿人与演讲者，他还是Risk Science播客的主持人并且曾经是2014 Metricon安全指标/分析大会的联席主席。可以通过@jayjacobs在推特上找到他。他拥有美国康卡迪亚大学科技管理的学士学位以及美国宾夕法尼亚州立大学的应用统计学毕业证书。

Bob Rudis拥有超过20年的利用数据来帮助维护全球财富100强企业的经验。作为Liberty Mutual Insurance的企业信息安全及IT风险管理部的主管，他负责协调与管理Advanced Cyber Security Center的多部门大范围安全分析计划。Bob是一名高级推特撰写人 (@hrbrmster)、活跃的博主 (rud.is)、作家、演讲者以及开源社区的投稿人 (github.com/hrbrmstr)。他当前正任职于Society of Information Risk Analysts (SIRA) 的董事会，是SANS Securing The Human方案的编委，同时，还是2014年Metricon安全指标/分析会议的联合主席。他拥有斯克兰顿大学的学士学位。

技术编辑Russell Thomas是一名Zions Bancorporation的安全数据科学家，还是一名乔治梅森大学社会计算科学的在读博士研究生。他拥有在计算机行业超过30年的技术、管理以及咨询方面的经验。Thomas先生是Securitymetrics.org的长期社区会员和Society of Information Risk Analysts (SIRA) 的创始成员之一。

书籍目录

目录 Contents

译者序

前言

作者介绍

第1章 通向数据驱动安全的旅程 1

1.1 数据分析简史 2

1.1.1 19世纪的数据分析 2

1.1.2 20世纪的数据分析 3

1.1.3 21世纪的数据分析 4

1.2 获取数据分析技能 5

1.2.1 领域专业知识 6

1.2.2 编程技能 8

1.2.3 数据管理 11

1.2.4 统计学 12

1.2.5 可视化 14

1.2.6 将这些技能组合起来 16

1.3 以问题为中心 16

1.3.1 创建一个好的研究问题 17

1.3.2 探索性数据分析 18

1.4 本章小结 19

推荐阅读 19

第2章 打造自己的分析工具箱 20

2.1 为什么选Python？为什么选R？为什么两者都要？ 21

2.2 用Canopy快速开始Python分析 23

2.2.1 理解Python数据分析和

可视化生态系统 24

2.2.2 设置R语言环境 27

2.3 数据帧介绍 30

2.4 组织结构 33

2.5 本章小结 34

推荐阅读 35

第3章 学习安全数据分析的“Hello World” 36

3.1 解决一个问题 37

3.2 获取数据 37

3.3 读入数据 40

3.4 探索数据 43

3.5 回到具体问题 54

3.6 本章小结 64

推荐阅读 65

第4章 进行探索性的安全数据分析 66

4.1 IP地址的剖析 67

4.1.1 IP地址的表示 67

4.1.2 IP地址的分段和分组 69

4.1.3 定位IP地址 71

4.2 IP地址数据的扩充 74

4.3 跨区域绘图 83

4.3.1 宙斯僵尸网络的可视化 85

- 4.3.2 防火墙数据的可视化 91
- 4.4 本章小结 93
- 推荐阅读 94
- 第5章 从地图到回归分析 95
- 5.1 简化地图 96
- 5.1.1 每个国家的ZeroAccess木马感染量是多少 99
- 5.1.2 改变数据范围 102
- 5.1.3 Potwin效应 104
- 5.1.4 结果奇怪吗？ 107
- 5.1.5 郡计数 111
- 5.1.6 郡级 112
- 5.2 线性回归介绍 115
- 5.2.1 回归分析中的常见陷阱 120
- 5.2.2 ZeroAccess木马感染的回归分析 121
- 5.3 本章小结 125
- 推荐阅读 125
- 第6章 将安全数据可视化 126
- 6.1 为什么要可视化 127
- 6.2 理解视觉交流的组件 133
- 6.2.1 避免第三维 133
- 6.2.2 使用颜色 135
- 6.2.3 拼在一起 137
- 6.2.4 描述分布信息 143
- 6.2.5 可视化时间序列 146
- 6.2.6 亲自实践 147
- 6.3 将数据变成电影明星 147
- 6.4 本章小结 148
- 推荐阅读 148
- 第7章 从安全失陷中进行学习 150
- 7.1 建立研究项目 151
- 7.2 数据收集框架的思考 152
- 7.2.1 瞄准目标答案 152
- 7.2.2 限制可能的答案 153
- 7.2.3 允许“其他”和“未知”选项 153
- 7.2.4 避免混淆并且合并细节 154
- 7.3 VERIS概述 155
- 7.3.1 事件追踪 156
- 7.3.2 威胁角色 157
- 7.3.3 威胁行为 158
- 7.3.4 信息资产 160
- 7.3.5 属性 162
- 7.3.6 发现/响应 163
- 7.3.7 影响 164
- 7.3.8 受害者 164
- 7.3.9 指标 166
- 7.3.10 用附加扩展VERIS 166
- 7.4 从行为中看VERIS 166
- 7.5 使用VCDB数据 168
- 7.6 本章小结 175

推荐阅读 176

第8章 离开关系数据库 177

8.1 实现有约束的存储器 180

8.1.1 架构方面的约束 181

8.1.2 存储方面的约束 183

8.1.3 RAM方面的约束 184

8.1.4 数据方面的约束 185

8.2 探索替代性的数据库 185

8.2.1 BerkeleyDB 186

8.2.2 Redis 188

8.2.3 HIVE 192

8.2.4 MongoDB 194

8.2.5 特殊目的的数据库 199

8.3 本章小结 200

推荐阅读 200

第9章 解密机器学习 201

9.1 检测恶意软件 202

9.1.1 开发机器学习算法 204

9.1.2 验证算法 205

9.1.3 实现机器学习算法 206

9.2 从机器学习中获益 209

9.2.1 用机器学习回答问题 210

9.2.2 评测良好的性能 211

9.2.3 选择特征 211

9.2.4 验证你的模型 213

9.3 具体的机器学习方法 213

9.3.1 有监督学习方法 214

9.3.2 无监督学习方法 217

9.4 实验：攻击数据聚类 218

9.4.1 受害行业的多维尺度分析 220

9.4.2 受害行业的层次聚类分析 222

9.5 本章小结 225

推荐阅读 225

第10章 设计有效的安全仪表盘 226

10.1 什么是仪表盘 226

10.1.1 仪表盘不是汽车 227

10.1.2 仪表盘不是报告 229

10.1.3 仪表盘不是搬运车 231

10.1.4 仪表盘不是艺术展 233

10.2 通过仪表盘表达及管理“安全” 237

10.2.1 帮负责人一个忙 237

10.2.2 提升仪表盘的意识 239

10.2.3 难题在细节中 241

10.2.4 突出“安全” 243

10.3 本章小结 245

推荐阅读 245

第11章 交互式安全可视化 247

11.1 从静态到交互式 248

11.1.1 用于增强的交互 248

- 11.1.2 用于探索的交互 251
- 11.1.3 用于启发的交互 254
- 11.2 开发交互式可视化 259
 - 11.2.1 使用Tableau创建交互式仪表盘 259
 - 11.2.2 使用D3创建基于浏览器的可视化 261
- 11.3 本章小结 271
- 推荐阅读 271
- 第12章 走向数据驱动的安全 273
 - 12.1 让自己走向数据驱动的安全 273
 - 12.1.1 黑客 274
 - 12.1.2 统计学 277
 - 12.1.3 安全领域专家 278
 - 12.1.4 危险区域 278
 - 12.2 带领团队走向数据驱动的安全研究 279
 - 12.2.1 对具有客观答案的事情提问 279
 - 12.2.2 查找并收集相关数据 280
 - 12.2.3 从迭代中学习 280
 - 12.2.4 寻找统计人才 281
 - 12.3 本章小结 283
- 推荐阅读 283
- 附录A 资料及工具 284
- 附录B 参考资源 287

《数据驱动安全》

精彩短评

1、烂翻译

2、本书从数据分析和数据可视化的角度来定位、分析安全问题，主要内容包括数据驱动安全的基础原理，常用的数据分析编程语言Python和R，数据分析代表性的软件工具，数据可视化和机器学习的基础算法，以及如何有效的设计交互式安全仪表盘。本书从数据的角度来分析安全，这是网络安全的一个热门领域，如何采用人工智能、机器学习从海量的网络数据中发现潜在的威胁，已经发生的安全事件，攻击踪迹，甚至威胁情报信息，这是非常值得关注的问题。

3、360团队翻译的，并且把数据驱动安全作为公司宣传语，但这本书讲的真没意思，翻译的不好，对不起这个名字

章节试读

1、《数据驱动安全》的笔记-第12页

询问我们身边人们的看法会导致错误地肯定自身的观点，因为我们很自然地 and 志同道合的人聚集在一起，且想法趋同一致。

2、《数据驱动安全》的笔记-第17页

分析工作的背景以及目的是根据研究问题来制定的，而不是源于我们可获取的数据。

3、《数据驱动安全》的笔记-第9页

Excel作为临时的解决方案，其能很好地快速处理一次性任务。但是如果你有一个需要重复分析的任务或者反复使用的模型的话，最好用某种结构化编程语言来处理。

作为一种数据清理工具，使用电子表格初看起来是一个不错的解决方案（尤其是对一些熟悉这方面技能的人来讲），但是电子表格是事件驱动的，意味着它们需要通过点击、打字、拖拽来工作。如果想用来转行一行数据，你就不得不点击表格，选中该行数据，然后再转换数据。这适合一些小的数据集或者快速的任務，但是相信我，你将会（比预期的还频繁）不得不回溯原始数据然后重新清理它。某一天，也许你有一些新的日志文件需要处理，也许你会意识到应该再从原始数据中提取另外的数据关系，也许（累得喘息）你在数据清理过程中发现了一个错误。也许不止一次地，某个点、某个处理细节会导致你重新回溯原始数据，然后重复数据清理以及转换的过程，利用电子表格的话，意味着你需要更多的无数次点击。然而，写一个脚本来运行的话，就可以很轻易、灵活以及一致地执行数据清理过程。

《数据驱动安全》

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:www.tushu000.com