HZ BOOKS
华章IT

· 网络空间安全技术丛书 ·

DATA SECURITY ARCHITECTURE
DESIGN AND PRACTICE

数据安全架构设计与实战

郑云文 编著

- 资深数据安全专家十年磨一剑的成果，多位专家联袂推荐。
- 本书以数据安全为线索，透视整个安全体系，将安全架构理念融入产品开发、安全体系建设中。

机械工业出版社
China Machine Press

数据安全架构设计与实战

DATA SECURITY ARCHITECTURE DESIGN AND PRACTICE

郑云文 编著

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

数据安全架构设计与实战 / 郑云文编著 . —北京：机械工业出版社，2019.9（2020.5 重印）
（网络空间安全技术丛书）

ISBN 978-7-111-63787-5

I. 数… II. 郑… III. 数据处理 – 安全技术 IV. TP274

中国版本图书馆 CIP 数据核字（2019）第 213686 号

数据安全架构设计与实战

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）
责任编辑：吴 怡 责任校对：李秋荣
印 刷：三河市宏图印务有限公司 版 次：2020 年 5 月第 1 版第 3 次印刷
开 本：186mm × 240mm 1/16 印 张：22.75
书 号：ISBN 978-7-111-63787-5 定 价：119.00 元

客服电话：（010）88361066 88379833 68326294 投稿热线：（010）88379604
华章网站：www.hzbook.com 读者信箱：hzit@hzbook.com

Praise 对本书的赞誉

数据安全问题其实一直存在，只是在大数据、基于大数据的人工智能时代变得更加重要。郑云文的这本书在覆盖信息安全、网络安全基础知识与最佳实践的基础上，对数据安全相关问题做了更深入的探讨。如同书中的观点，安全的系统是设计出来、开发出来的，没有一招见效的“安全银弹”。这本书非常适合软件开发型企业的开发主管、信息安全主管与开发工程师阅读，也适合高校信息安全专业的同学作为了解业界网络安全实践的参考书使用。

——谭晓生，北京赛博英杰科技有限公司创始人

数据安全是企业安全建设的重点与难点，相对应用安全和内网安全，大多数企业在数据安全的投入要少很多，但数据安全的重要性却要高很多。少数金融机构设置了专门的数据安全团队，投入大量人财物，未雨绸缪。但多数企业安全建设在数据安全领域还是被动的，其难处在于数据安全缺乏体系性解决方案和安全前置措施。这本书将为读者提供目前急缺的这部分内容，是数据安全领域不错的专业好书。

——聂君，奇安信首席安全官

这本书以数据安全实践为基础，结合网络与信息安全相关的理论、技术、方法、案例，系统全面地介绍了数据安全保护技术设计与实现中的知识和经验，并对数据安全相关的法律、法规、标准等合规性要求进行了梳理，是一本很好的数据安全架构设计与实现的参考书。

——王绍斌，亚马逊 AWS 大中华区首席信息安全官

这本书系统地总结了作者在互联网巨头公司安全部门长期工作的最新成果，体系性、实战性及可操作性都很强，对安全产业人员非常有参考价值！

——蔡一兵，恒安嘉新副总裁

这本书的作者云文是安全领域的一名老兵，他对于安全的热情和善于总结提炼的特点一

直令我印象深刻。随着信息化社会的高速发展，安全问题越来越多地得到大家的关注。但安全是一个庞大的系统工程，而数据安全往往是贯穿整个工程的核心焦点。数据安全的难点不仅在于复杂多样的对抗技术，更在于如何进行工程落地，整合成适配的解决方案，并在业务成长的过程中同样完善发展。这一切都深深地困扰着安全从业者。这本书作者结合在大型互联网企业的工作经验，针对上述难点，系统地讲述数据安全建设方法，并总结出一套保障数据安全的方法论。推荐从事安全工作的朋友们阅读此书。

——杨勇（Coolc），腾讯安全平台部负责人，腾讯安全学院副院长

这本书站在企业的角度对安全架构设计给出了详尽的指导，从理论到实现，都沉淀自作者多年的工作实践经验，有着重要的参考意义。这本书对数据安全建设给出了非常完整的框架，是当前企业接触互联网与大数据后所急需的工作手册。

——吴翰清，阿里云首席安全科学家

如何保障数据安全是目前各行业广泛面临的挑战，而保障好数据安全要涉及基础设施、系统架构、业务应用甚至生态链条的方方面面，做好、做扎实尤其不易。郑云文的这本书非常契合当前行业的需求，特别对于数据安全保障的关键环节有着翔实的实践经验分享，能够给安全从业者良好的借鉴。

——韦韬（Lenx），百度首席安全科学家，北大客座教授

市面上的书在企业安全领域分散的知识点很多，系统性的方法比较少，讲攻防的居多，讲数据安全的极少。企业安全、信息安全本质上还是要保护数据安全，但数据安全的问题大部分不是因为攻防对抗的缺漏引起的，而是发生在整个企业价值链和全生命周期，甚至在泛生态、产业链上也有衍生问题。对于这些问题的解决不能只靠单点技术对抗，而是需要有贯穿全局的视野和系统性风险防范的意识。这本书是难得的系统性讲述数据安全建设方法的书，对于广大安全从业者是不可多得的必备读物，对于从事安全数据分析（态势感知）的同学可以增加态势的全局能力，对于从事 SDL 的同学可以提升安全设计能力，对于从事应急响应的同学可以提升数据视角应急和溯源的能力。对红蓝对抗的攻方和防守方以及业务安全的风控，这本书补全了数据视角下需要的闭环运营工作的视野。强烈推荐。

——赵彦（ayazero），美团点评集团安全部总经理

如今已进入“数据为王”的时代，谁掌握了数据，谁就拥有了最宝贵的资源和最强大的业务潜能，同时必然面临着重大的安全威胁与责任。

业务系统快速迭代，攻防博弈不断升级，数据安全治理政策逐渐出台，这些因素合并产

生的压力与日俱增。为此，这本书针对数据安全保护这一核心问题，从安全架构基础、产品安全架构、安全技术体系架构、数据安全与隐私保护治理等多个角度进行了全面探讨和循序渐进的系统化梳理，并给出了在各阶段保障数据安全的有效解决思路和业界最佳实践，有利于读者快速了解数据安全的全貌，构建数据安全技术体系以及架构性思维模式，有助于相关单位从系统设计阶段就开始系统规划和引入安全策略、部署安全技术，并遵循数据安全治理要求，防患于未然。

这本书适合于信息系统设计、开发和运维人员，以及安全从业人员，同时也非常适合于网络空间安全与计算机学科的在校大学生。

——彭国军，武汉大学国家网络安全学院教授

这是围绕数据安全来考虑安全架构设计的书籍，作者从认证授权和数据资产保护的角度，对安全防御体系做出了诠释，将多年工作经验沉淀其中，值得一读。

——董志强，腾讯安全云鼎实验室负责人

随着互联网时代的发展，数据已经逐渐成为企业的核心资产，对数据资产的保护成为新的课题。这本书脱离传统的网络安全视角，以数据安全为中心展开讨论，在数据生命周期的各个流转阶段引入安全措施，观点新颖，既有丰富的理论知识也有最佳实践，是一本不错的信息安全专业书籍。推荐。

——胡珀（lake2），资深网络安全专家、腾讯安全平台部总监

随着欧盟《通用数据保护条例 GDPR》生效和各国监管法案的出台，数据安全和隐私保护已经成为企业安全建设的关注点。数据安全可以归类为信息安全或网络安全众多安全领域中的一个，也可以视为与信息安全和网络安全并驾齐驱的独立安全体系。数据安全在信息安全 CIA 三性基础上增加了数据主体的权利，如何平衡数据价值的合规利用和有效保护？需要数据安全的方法论和最佳实践，供企业在实践中借鉴和指导其落地。

郑云文的这本著作不仅包括数据保护的方法论和框架，同时对数据安全和隐私保护的核心技术做了详细阐述；这本书不仅适合产品经理、开发工程师理解数据安全的方法论，同时适合安全合规人员、律师了解相关技术措施在隐私保护上起到的作用。

——宋文宽，小米安全与隐私合规总监

数据安全是安全工程建设和运营的重要结果。这本书围绕数据安全，从技术、管理、合规等维度，与读者分享相关思考、方法论和实践，是该领域不可多得的一本好书。

——方勇，腾讯云安全首席架构师

数据安全是安全线人员和业务线人员都可以理解的为数不多的领域，业务线的人员很可能不懂 SQL 注入，也很可能不懂 WebShell，但是他们可以理解业务数据库中数据的价值以及数据丢失后对公司的损失。数据安全从广义来理解，技术视角是数据安全，2C 视角是个人隐私，这些都是目前安全领域的热门话题。相关的大部头文章其实很多，各种规范也层出不穷，但是如何落地到业务中呢？这本书的作者有着多年安全领域的从业经历，既有互联网公司的安全经验，也有运营商和传统制造领域的安全经验。作者把自己对数据安全的理解和实践，浓缩在这本书的 20 章中，强烈推荐安全从业人员研读这本书。

——兜哥，百度安全资深研究员，《企业安全建设入门》作者

很高兴能看到这本书出版，这本书内容丰富翔实，令人赞叹。曾经和郑兄在同一互联网巨头公司的安全平台部门共事过两年，打开这本书仿佛使我又回到了当年一起对抗黑产入侵、一起推动数据保护项目的日子。数据资产是互联网企业的核心资产，关系到企业的生命线。作为多年奋战数据安全保护战线的一名互联网安全老兵，郑兄从实战出发，很好地总结了数据安全体系的建设思路，将实践中的经验教训升级为朴实的方法论，非常具有借鉴价值。

——马传雷（Flyh4t），同盾科技反欺诈研究院负责人

我曾经跟郑云文共同工作过一段时间。在数据安全团队的职责上，不同的公司目前还是有一些区别的。普遍认为数据安全很难独立于传统的基础安全 / 应用运维安全而存在，因此，站在保护数据的视角来看，传统的很多做法本质上是为了保护公司的信息资产——数据。书中既有站在甲方安全运营视角下的传统建设思路和释疑，又有站在数据视角下的方法论介绍和实战经验。因此，本书的架构、内容，其实非常适合那些想要对数据进行保护，却不知道如何下手的企业和工程师，也很适合学生时期就计划投身于安全领域的新人。

——赵弼政（职业欠钱），美团点评基础安全负责人

Foreword 序 一

在互联网和新兴技术高速发展的今天，数据信息充斥在各行各业中，并发挥着重要的作用。然而，在享受信息化时代带来便利的同时，数据安全问题也成为大家关注的焦点。无论是从 toG、toB、toC 的各业务场景来看，还是从网络安全（Cyber Security）的架构来看，数据安全（Data Security）都是一个主要的组成部分，而且在新兴技术日新月异的数据时代变得越来越重要，范围也越来越大。

中国云安全与新兴技术安全创新联盟已经把数据安全，包括大数据安全作为云安全之后的一个重要研究方向。

中央网信办赵泽良总工程师多次表示："数据安全已经成为网络安全的当务之急。"今年，国家互联网信息办公室就《数据安全管理办法》向社会公开征求意见。此外，还密集公布了《网络安全审查办法（征求意见稿）》和《儿童个人信息网络保护规定（征求意见稿）》。《数据安全管理办法》以网络运营者为主要规制对象，重点围绕个人信息和重要数据安全，在数据收集、数据处理使用、数据安全监督管理等方面进行了系统的管理规定，重点明确了使用范围、监管主体、个人信息收集和处理、问题处置等内容。

然而，在数据安全管理的要求下，达成数据安全的技术保障措施需要有一大批具有安全意识的业务研发人员和具有数据安全专业能力的安全人员。

这本书是云安全联盟（CSA）技术专家郑云文先生在数据安全领域实践多年的心得，适合广大架构师、工程师、信息技术人员、安全专家阅读，是一本实践性很强的安全架构设计书籍，被列为 CSA "注册数据安全专家"（Certified Data Security Professional）认证的学习参考资料。

安全防护的重心已经从"以 Network（网络）为中心"向"以 Data（数据）为中心"转

移，不论你是甲方企业还是乙方供应商，也不论你是进行单位的业务保障还是想提升个人能力，相信广大读者一定能从这本书受益。

李雨航（Yale Li）
CSA 大中华区主席
中国科学院云安全首席科学家
2019 年 6 月 26 日

Forword 序 二

初识云文是看到他写的一篇讲 SDL 的文章，觉得这个作者在安全领域很有研究，后来几经辗转联系到他，一聊下来大家非常投缘，于是就邀请云文加入腾讯数据安全团队一起开展数据安全工作。云文在数据安全团队工作期间做出了重要贡献，是多个重要安全系统的主要架构设计者，也是数据安全合规标准的主要制定者。

随着互联网时代的发展，越来越多的在线业务会产生大量数据，这些数据已经成为企业的核心资产。不同于过去的静态信息资产，现在的数据资产是流动的，对数据资产的动态保护已成为安全行业新的课题。数据安全已成为企业安全的重中之重，从过去的无数案例可以看到，许多企业因为数据泄露事件导致品牌受损、用户流失、高层辞职甚至业务停摆。

时势造英雄。越来越多的安全从业者开始关注和研究数据安全，也在实践过程中摸索出一些经验。这本书即是云文多年的数据安全研究和实践经验的总结。在这本书中，云文以精湛的文笔系统地阐述了数据安全体系，将数据安全架构设计、数据安全治理与数据全生命周期的预防性设计引入安全体系中。这本书与众不同之处在于它脱离传统的网络安全视角，而是从防御者的角度出发，将安全建设从“以产品为中心”逐步过渡到“以数据为中心”，并围绕数据生命周期的各个阶段引入安全措施进行保护，既有丰富的理论知识，又有能落地的最佳实践，可作为安全人员的案头必备书籍。

此外，还应当认识到，网络安全是一个整体，安全体系的建设也是一个漫长的过程，时代变化很快，唯有紧跟形势，不断学习、不断迭代优化，方能立于不败之地。

胡珀（lake2）
资深网络安全专家
腾讯安全平台部总监
2019 年 6 月 17 日

前　言 *Preface*

你一定听说过非常厉害的黑客，各种奇技淫巧，分分钟拖走大量数据！或入侵到目标内网，Get Shell、提升权限、拖走数据库！抑或根本不用进入内网，直接远程操作一番，就能窃取到大量数据，犹如探囊取物一般容易。

可是，站在黑客的对立面，作为防御的一方，公司频频遭遇入侵、网络攻击或数据泄露事件，一方面会面临巨大的业务损失，另一方面也会面临来自用户、媒体、监管层面的重重压力。

数据安全这是一个非常严峻的问题。数据泄露事件层出不穷，就算是安全建设得比较好的企业，也不能保证自己不出问题，况且在日常安全工作中，还面临着三大困境——资源有限、时间不够、能力不足，使得我们距离数据安全的目标还有不小的差距。

“资源有限”体现在企业在安全方面的投入往往不足，特别是在预防性安全建设、从源头开始安全建设的投入方面，更加缺乏。在有的产品团队，人力几乎全部投在业务方面，没有人对安全负责，产品发布上线后，也缺乏统一的安全增强基础设施（例如在统一的接入网关上实施强制身份认证），导致产品基本没有安全性可言。

“时间不够”是因为业务开发忙得不可开交，完成业务功能的时间都不够，哪里还有时间考虑安全呢？这也是为什么我们经常会发现有的 JSON API 接口根本就没有身份认证、授权、访问控制等机制，只要请求过来就返回数据。

“能力不足”体现在具备良好安全设计能力和良好开发能力的人员太少，基层开发人员普遍缺乏良好的安全实践和意识，写出来的应用频频出现高危漏洞。就算能够事先意识到安全问题，在实现上，安全解决方案也是五花八门，重复造轮子，且互不通用，往往问题多多，效率低下；就算发现了安全问题，然而牵一发而动全身，修改了问题还担心业务服务是否正常运转。

在几大困境面前，各产品团队往往寄希望于企业内安全团队的事后防御。殊不知，事后

解决问题，也有诸多局限：

- 时间不等人，险情就是命令！当漏洞或事件报告过来的时候，无论是节假日，还是半夜时分，都需要立即启动应急响应，“三更起四更眠”屡见不鲜。数量不多时还可以承受，但长此以往，负责应急的同学身体也吃不消，需要不断招聘新人及启用岗位轮换机制。
- 依赖各种安全防御系统，没有从根本上解决问题，属于治标不治本，黑客经常能找到绕过安全防御系统的方法，就如同羸弱的身体失去了铠甲的保护。
- 事后修复很可能会影响业务连续性，即便产品团队已经知道问题出在哪里了，但是由于业务不能停，风险迟迟得不到修复，因此还可能引发更大的问题。

安全不是喊口号就能做好的。实际上，安全是一项系统性工程，需要方法论的指导，也需要实践的参考。

我们如何才能克服上述三大困境，更好地保护业务，防止数据泄露呢？本书尝试通过一套组合拳逐一化解：

- 通过安全架构方法论的引入，探讨如何从源头开始设计产品自身的安全架构，快速提升产品自身的安全能力，让产品（网络服务等）天然就具有免疫力，构建安全能力的第一道防线。
- 梳理安全技术体系架构，建立并完善安全领域的基础设施及各种支撑系统，让产品与安全基础设施分工协作，并对协作进行疏导（即“哪些应该交给产品自身来实现，哪些交给安全基础设施进行落地”），减少各业务在安全上的重复性建设和资源投入，避免重复造轮子，让业务聚焦到业务上去，节省业务团队在安全方面投入的时间。产品外部的安全能力，构成了第二道防线。
- 以数据安全的视角，一览企业数据安全治理的全貌，协助提升大家的架构性思维，站在全局看问题，了解数据安全与隐私保护治理实践。

总的来说，这是一本有关数据安全架构的技术性书籍，但也会涉猎数据安全治理的内容，目的在于让大家了解数据安全的全局，培养架构性思维模式，希望能给企业安全建设团队或有志于从事安全体系建设的读者一些建设性的参考。

内容简介

随着数据时代的到来，安全体系架构逐步由之前的“以网络为中心”（称之为网络安全）过渡到“以数据为中心”（称之为数据安全）。本书将使用数据安全这一概念，并以数据的安全收集或生成、安全使用、安全传输、安全存储、安全披露、安全流转与跟踪、安全销毁为

目标，透视整个安全体系，进而将安全架构理念融入产品开发过程、安全技术体系及流程体系中，更好地为企业的安全目标服务。

我们将站在黑客的对立面，以防御的视角，系统性地介绍安全架构实践，共包含四个部分。

第一部分为安全架构的基础知识，为后续章节打好基础。

第二部分为产品安全架构，从源头开始设计产品自身的安全架构，提升产品的安全能力，内容包括：

- 安全架构 5A 方法论（即安全架构的 5 个核心要素，身份认证、授权、访问控制、审计、资产保护）。
- 产品（或应用系统）如何从源头设计数据安全（Security by Design）和隐私安全（Privacy by Design）的保障体系，防患于未然。

第三部分为安全技术体系架构，通过构建各种安全基础设施，增强产品的安全能力，内容包括：

- 建立和完善安全技术体系（包括安全防御基础设施、安全运维基础设施、安全工具与技术、安全组件与支持系统）。
- 安全架构设计的最佳实践案例。

第四部分为数据安全与隐私保护领域的体系化介绍，供读者了解数据安全与隐私保护的治理实践，内容包括：

- 数据安全治理，包括如何设定战略，组织、建立数据安全文件体系，以及安全运营、合规与风险管理实践等。
- 隐私保护治理，包括隐私保护基础、隐私保护技术、隐私保护治理实践等。

本书使用的源代码发布在 https://github.com/zhyale/book1，欢迎读者在此提交问题或反馈意见。

安全理念

本书将使用如下安全理念：

- “主动预防”胜于“事后补救”。
- 默认就需要安全，安全贯穿并融入产品的生命周期，尽可能地从源头改善安全（架构设计、开发、部署配置）；对数据的保护，也不再是保护静态存储的数据，而是全生命周期的数据安全与隐私保护（包括数据的安全收集或生成、安全使用、安全传输、安全存储、安全披露、安全流转与跟踪、安全销毁等）；在安全设计上，不依赖于广大员工的自觉性，而是尽量让大家不犯错误。

- 数据安全与隐私保护可以和业务双赢，数据安全与隐私保护不是妨碍业务的绊脚石，也可以成为助力业务腾飞的核心竞争力。只有真正从用户的立场出发，充分重视数据安全，尊重用户隐私，才能赢得市场的尊重。

在安全架构实践中，我们将采用基于身份的信任思维：默认不信任企业内部和外部的任何人、设备、系统，需基于身份认证和授权，执行以身份为中心的访问控制和资产保护。在涉及算法或理论细节时，我们将基于工程化及建设性思维：不纠缠产品或技术的理论细节，只考虑是否属于业界最佳实践，是否可以更好地用于安全建设，做建设性安全。

读者对象

本书主要面向安全领域的从业者、爱好者，特别是：

- 网络安全、数据安全从业人员；
- 希望提升产品安全性的应用开发人员；
- 各领域架构师；
- 有意进入安全行业、隐私保护行业的爱好者、学生。

致谢

感谢我所任职过的公司，在工作中让我有了练兵、成长、积累的机会，也感谢各位领导、同事、安全圈同行与各位朋友的帮助，他们包括但不限于：谭晓生 @ 赛博英杰、李雨航 @CSA、聂君（君哥的体历）@ 奇安信、王绍斌 @ 亚马逊、蔡一兵 @ 恒安嘉新、杨勇（coolc）@ 腾讯、赵彦（ayazero）@ 美团、吴翰清 @ 阿里巴巴、韦韬（Lenx）@ 百度、彭国军 @WHU、董志强（killer）@ 腾讯、胡珀（lake2）@ 腾讯、郑斌（天明）@ 阿里巴巴、宋文宽 @ 小米、方勇（包子）@ 腾讯、刘焱（兜哥）@ 百度、赵弼政（职业欠钱）@ 美团、马传雷（Flyh4t）@ 同盾、王珂、郑兴（召唤）@ 腾讯、刘宁 @ 腾讯、郭铁涛 @ 腾讯、胡享梅（梅子）@ 腾讯（排名不分先后）以及一大批曾经一起奋战过的同事们（应公司要求不能列出名字和单位）。

感谢云安全联盟 CSA 大中华区主席、中国科学院云安全首席科学家李雨航为本书作序。

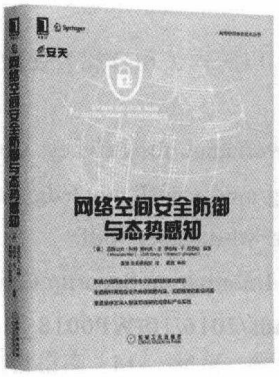
感谢资深网络安全专家、腾讯安全平台部总监胡珀为本书作序。

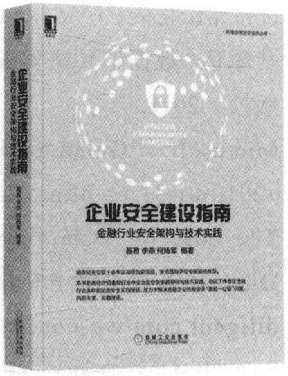
感谢编辑吴怡，为本书提出大量改进意见和建设性建议。感谢机械工业出版社的各位编辑、排版、设计人员。

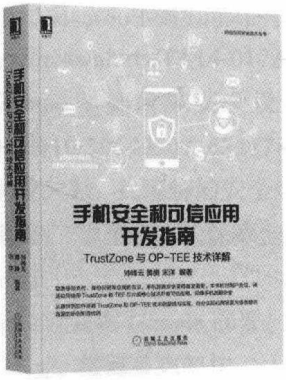
感谢我的家人，是你们的支持，我才得以完成此书。

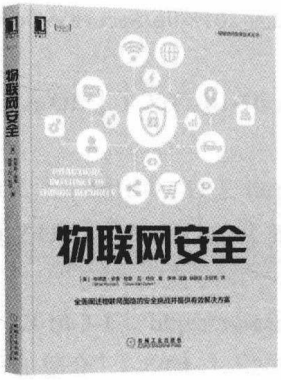
推荐阅读

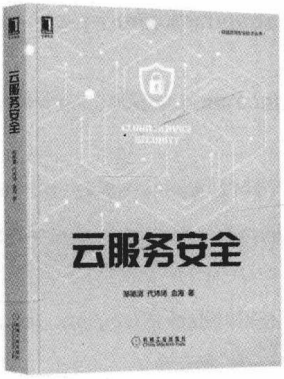

计算机安全
原理与实践

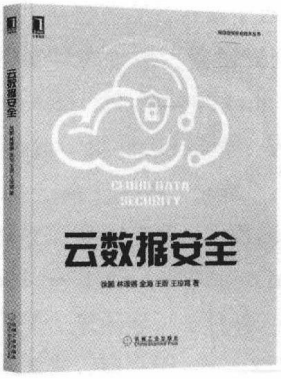

网络空间安全防御与态势感知

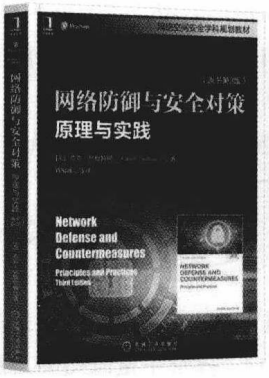

企业安全建设指南
金融行业安全架构与技术实践

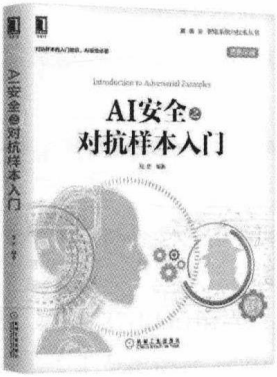

手机安全和可信应用开发指南
TrustZone 与 OP-TEE 技术详解


物联网安全

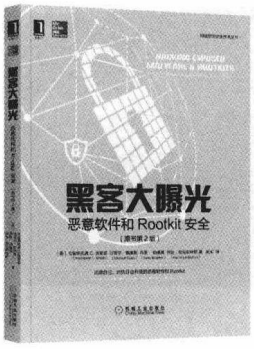

云服务安全

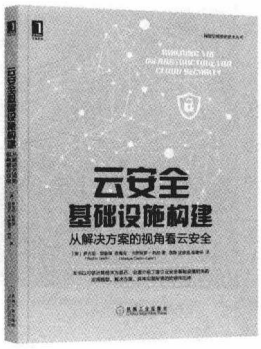

云数据安全

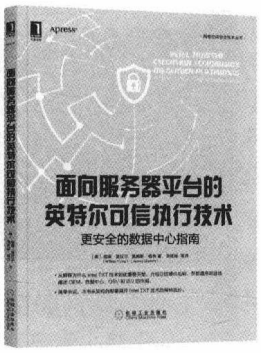

网络防御与安全对策
原理与实践
Network Defense and Countermeasures

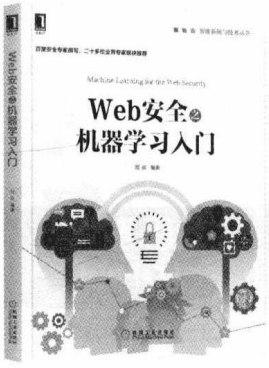

AI安全之对抗样本入门

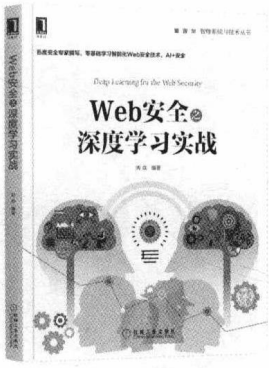
推荐阅读

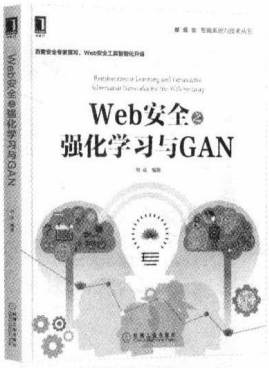












- **黑客大曝光：恶意软件和Rootkit安全(原书第2版)**
  作者：克里斯托弗 C. 埃里森
  ISBN：978-7-111-58054-6
  定价：79.00元

- **云安全基础设施构建：从解决方案的视角看云安全**
  作者：罗古胡. 耶鲁瑞
  ISBN：978-7-111-57696-9
  定价：49.00元

- **面向服务器平台的英特尔可信执行技术：更安全的数据中心指南**
  作者：威廉.普拉尔
  ISBN：978-7-111-57937-3
  定价：49.00元

- **Web安全之机器学习入门**
  作者：刘焱
  ISBN：978-7-111-57642-6
  定价：79.00元

- **Web安全之深度学习实战**
  作者：刘焱
  ISBN：978-7-111-58447-6
  定价：79.00元

- **Web安全之强化学习与GAN**
  作者：刘焱
  ISBN：978-7-111-59345-4
  定价：79.00元

推荐阅读