

O'REILLY®



Web开发的 身份和数据安全

Identity & Data Security for Web Development

Jonathan LeBlanc
Tim Messerschmidt 著
安道 译

出版社

TP393.292.2
201812

Web开发的身份和数据安全

Jonathan LeBlanc 和 Tim Messerschmidt 著
安道 译



Beijing • Boston • Farnham • Sebastopol • Tokyo

O'REILLY®

O'Reilly Media, Inc. 授权中国电力出版社出版

中国电力出版社

Copyright © 2016 Jonathan LeBlanc, Tim Messerschmidt. All rights reserved.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and China Electric Power Press, 2018.
Authorized translation of the English edition, 2016 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由 O'Reilly Media, Inc. 出版 2016。

简体中文版由中国电力出版社出版 2018。英文原版的翻译得到 O'Reilly Media, Inc. 的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc. 的许可。

版权所有，未得书面许可，本书的任何部分和全部不得以任何形式重制。

图书在版编目 (CIP) 数据

Web开发的身份和数据安全 / (美) 乔纳森·勒布朗 (Jonathan LeBlanc), (美) 蒂姆·梅塞施米特 (Tim Messerschmidt) 著; 安道译. — 北京: 中国电力出版社, 2018.1

书名原文: Identity and Data Security for Web Development

ISBN 978-7-5198-1420-5

I. ①W… II. ①乔… ②蒂… ③安… III. ①网页制作工具 IV. ①TP393.092.2

中国版本图书馆CIP数据核字(2017)第293147号

北京市版权局著作权合同登记 图字: 01-2017-7622号

出版发行: 中国电力出版社

地 址: 北京市东城区北京站西街19号 (邮政编码100005)

网 址: <http://www.cepp.sgcc.com.cn>

责任编辑: 刘 焯 (liuchi1030@163.com)

责任校对: 闫秀英

装帧设计: Karen Montgomery, 张 健

责任印制: 蔺义舟

印 刷: 三河市百盛印装有限公司

版 次: 2018年1月第一版

印 次: 2018年1月北京第一次印刷

开 本: 787毫米×980毫米 16开本

印 张: 12.75印张

字 数: 238千字

印 数: 0001—3000册

定 价: 48.00元

版权专有 侵权必究

本书如有印装质量问题, 我社发行部负责退换

O'Reilly Media, Inc. 介绍

O'Reilly Media通过图书、杂志、在线服务、调查研究和会议等方式传播创新知识。自1978年开始，O'Reilly一直都是前沿发展的见证者和推动者。超级极客们正在开创着未来，而我们关注真正重要的技术趋势——通过放大那些“细微的信号”来刺激社会对新科技的应用。作为技术社区中活跃的参与者，O'Reilly的发展充满了对创新的倡导、创造和发扬光大。

O'Reilly为软件开发人员带来革命性的“动物书”；创建第一个商业网站（GNN）；组织了影响深远的开放源代码峰会，以至于开源软件运动以此命名；创立了Make杂志，从而成为DIY革命的主要先锋；公司一如既往地通过多种形式缔结信息与人的纽带。O'Reilly的会议和峰会集聚了众多超级极客和高瞻远瞩的商业领袖，共同描绘出开创新产业的革命性思想。作为技术人士获取信息的选择，O'Reilly现在还将先锋专家的知识传递给普通的计算机用户。无论是通过书籍出版，在线服务或者面授课程，每一项O'Reilly的产品都反映了公司不可动摇的理念——信息是激发创新的力量。

业界评论

“O'Reilly Radar博客有口皆碑。”

——Wired

“O'Reilly凭借一系列（真希望当初我也想到了）非凡想法建立了数百万美元的业务。”

——Business 2.0

“O'Reilly Conference是聚集关键思想领袖的绝对典范。”

——CRN

“一本O'Reilly的书就代表一个有用、有前途、需要学习的主题。”

——Irish Times

“Tim是位特立独行的商人，他不光放眼于最长远、最广阔的视野并且切实地按照Yogi Berra的建议去做了：‘如果你在路上遇到岔路口，走小路（岔路）。’回顾过去Tim似乎每一次都选择了小路，而且有几次都是一闪即逝的机会，尽管大路也不错。”

——Linux Journal

前言

“黑客每年导致企业损失 4 千亿美元。”^{注 1}

—— Inc. 杂志

2015 年第四季度，Cybersecurity Ventures 发布的一份网络安全市场报告指出，网络攻击一年导致企业损失 4 千亿到 5 千亿美元。^{注 2} 为了应对这一问题，2015 年在 IT 安全上的花费增长了 4.7%，达到 754 亿美元，而且预计到明年全球的花费将达 1010 亿美元，2020 年则增长到 1700 亿美元。据此估计，到后年，网络安全从业人员的短缺将达 150 万人，而整体需求将增长到 600 万人。

作为 Web 和应用开发者、设计师、工程师和创作者，我们责无旁贷，必须掌握身份和数据安全方面的知识。如果 Web 开发者不知道如何在传输过程中正确隐蔽数据，不经意间就会敞开网站的安全大门。如果项目经理不知道原本安全的密码算法现已包含瑕疵，不把再次处理数据库中的用户记录当做要事，会导致应用暴露严重的攻击媒介。如今，在一个系统中工作的每个人都要担起责任，确保用户和数据得到保护。

尽管都知道安全第一，但是每周似乎都能见到有公司被攻击了，从初创公司到大型企业，无一幸免。遭泄露的有重要的用户信息、信用卡数据、病例等，而这些都是用户托付给那些公司保护的。总是有很多这样的组织从不把数据加密妥当，所有信息都明文存储，好像等待着黑客去攫取。

注 1: <http://www.inc.com/will-yakowicz/cyberattacks-cost-companies-400-billion-each-year.html>。

注 2: <http://cybersecurityventures.com/cybersecurity-market-report>。

真正的问题是，如今黑客不再单独行事，证明他们可以攻破系统，而是一种有组织的行为，为的是获取金钱或破坏别人的生意。

这就是笔者写作这本书的目的。每一章都会讲解一些概念，告诉你如何封堵现有系统中的漏洞，如何防护可行的攻击媒介，如何在某些自身并不安全的环境中工作。我们将讲解的概念有：

- 了解 Web 和应用安全的现状。
- 构建安全的加密方式，以及与各种密码攻击媒介斗争。
- 创建数字指纹，在浏览器、设备和配对设备中识别用户。
- 通过 OAuth 和 OpenID Connect 构建安全的数据传输系统。
- 使用其他的识别方法提供第二种身份验证方式。
- 加固 Web 应用，防止攻击。
- 使用 SSL/TLS 及同步和异步加密创建安全的数据传输系统。

读完本书后，你将全面了解身份和数据安全的现状，知道如何保护自己，免受潜在的攻击，保护用户托付给你存储的数据，以防泄露。

排版约定

本书使用下述排版约定。

斜体 (*Italic*)

表示新术语、URL、电子邮件地址、文件名和扩展名。

等宽字体 (`Constant Width`)

表示代码清单，在段落中出现则表示程序元素，例如变量、函数名、数据库、数据类型、环境变量、语句和关键字。

粗体等宽字体 (`Constant width bold`)

表示应该由用户输入的命令或其他文本。

斜体等宽字体 (*Constant Width Italic*)

表示应该替换成用户提供的值，或者由上下文决定的值。



这个图标表示提示或建议。



这个图标表示一般性说明。



这个图标表示警告或提醒。

Safari® Books Online



Safari Books Online (<http://safaribooksonline.com>) 是应需而变的数字图书馆，它同时以图书和视频的形式出版世界顶级技术和商务作家的专业作品。

Safari Books Online 是技术专家、软件开发人员、Web 设计师、商务人士和创意人士开展调研、解决问题、学习和认证培训的第一手资料来源。

Safari Books Online 为企业、政府部门、教育机构和个人提供了多种套餐和价格。

订阅者可以在一个完全可搜索的全文数据库中访问上千种图书、培训视频和正式出版之前的书稿。这些内容由以下出版社提供：O'Reilly Media, Prentice Hall Professional, Addison-Wesley Professional, Microsoft Press, Sams, Que, Peachpit Press, Focal Press, Cisco Press, John Wiley & Sons, Syngress, Morgan Kaufmann, IBM Redbooks, Packt, Adobe Press, FT Press, Apress, Manning, New Riders, McGraw-Hill, Jones & Bartlett,

Course Technology 等。关于 Safari Books Online 的更多信息，请访问我们的网站。

联系方式

请把你对本书的意见和疑问发给出版社：

美国：

O'Reilly Media, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472

中国：

北京市西城区西直门南大街 2 号成铭大厦 C 座 807 室 (100035)
奥莱利技术咨询 (北京) 有限公司

这本书有专属网页，你可以在那儿找到本书的勘误、示例和其他信息。这个网页的地址是 <http://shop.oreilly.com/product/0636920044376.do>。

如果你对本书有一些评论或技术上的建议，请发送电子邮件到 bookquestions@oreilly.com。

若想了解 O'Reilly 图书、培训课程、会议和新闻的更多信息，请访问我们的网站，地址是 <http://www.oreilly.com>。

我们在 Facebook 的地址：<http://facebook.com/oreilly>。

请关注我们的 Twitter 动态：<http://twitter.com/oreillymedia>。

我们的 YouTube 视频地址：<http://www.youtube.com/oreillymedia>。

致谢

首先，感谢 O'Reilly 的全体工作人员，感谢他们出版这本书，让我们与世界各地的人们分享我们的知识、想法和观点。特别感谢编辑 Meg Foley，她在写作的过程中充满耐心，给我们支持和帮助。

还要感谢 Lenny Markus、Allen Tom 和 Aaron Parecki，他们细致审阅了本书的草稿，极大地提升了书稿的质量。

我们还要感谢开发者同事，他们帮助校对、为我们提建议，给我们充足的时间写作本书。

最后，我们要对你们——亲爱的读者表示感谢，谢谢你们购买这本书。祝你们阅读愉快！

Jonathan

首先，感谢与我合著本书的 Tim，他是一个优秀的伙伴。我们积极交流，不断冒出想法又推倒重来，最终让本书呈现在你面前。你的想法、鼓励和幽默让写作这本书成为我最珍贵的体验之一。

感谢我的妻子 Heather，大约五年前，我决定写第一本书时，你让我保持理智。尽管我早已忘记写书的辛劳，但是当我决定再写一本的时候，你依然给我支持。没有你，我无法保持头脑清醒，支撑我走完整个过程。你始终陪伴在我身旁，鼓励我追求自己的梦想。我做任何事，你都是我最大的拥护者。我爱你！

感谢我的女儿 Scarlett，做你的父亲真是一件乐事，你为我的生活带来了安宁。世界充满混乱，但是你让我认识到情况并没有我认为的那样严重，这让我受用终生。

感谢我的团队、我的朋友。我们走上不同的道路，在世界各地不同的公司工作，但是你们始终是我最近的朋友之一。我们在一起经历了很多，也做了诸多牺牲。尽管如此，你们是我做任何事的支持者，为我打气，让我向着成功迈进。感谢你们！

Tim

我要感谢 Jonathan，他不仅是难得的同事和朋友，还是本书优秀的合著者。与你不断讨论意见和想法真是让人怀念，我相信，如果没有你的影响、支持和付出，这本书将失色不少。

由衷感谢我的妻子 Karin，或许应该送上一大束鲜花，她给我充足的时间，让我写完本书。

Joe Nash、Alan Wong、Steven Cooper 和 Cristiano Betta 组成了一个优秀团队，为本书的写作贡献颇多，必须把你们列在这里。

感谢鼓励我写这本书的每一个人，感谢在各种场合跟我一起畅谈安全和可用性概念的每一个人。

特别感谢 PayPal 开源部门的主管 Danese Cooper，他强烈建议我在博客之外的媒介上写下我的想法。

最后，感谢 John Lunn 和 Taylor Nguyen，他们在我写作本书时给予了大力支持，而且在我的职业发展中给予建议。

目录

前言	1
第 1 章 导论	7
现有安全模型的问题	7
弱密码	8
为了安全牺牲可用性	9
不当的数据加密	10
最薄弱的环节：人类	11
单点登录	12
理解密码安全中的熵	13
随机选择的密码熵	14
人为选定的密码熵	15
区分用户名和密码在系统中的作用	17
保护身份的当前标准	18
好的和不好的安全算法	18
应该保护哪些数据？	20
账户恢复机制和社会工程	20
安全问题带来的问题	20
下一步	21
第 2 章 密码加密、哈希和加盐	22
静态数据和动态数据	22
静态数据	22
动态数据	24

密码攻击媒介	24
暴力攻击	26
使用 reCAPTCHA 添加验证码	27
字典攻击	33
反向查询表	34
彩虹表	35
加盐	37
生成随机盐值	38
重用盐值	39
盐值的长度	39
把盐值存储在哪里	39
撒胡椒	40
选择正确的密码哈希函数	41
bcrypt	41
PBKDF2	43
scrypt	44
对比哈希值，验证密码	46
密钥延伸	47
重新计算哈希值	48
下一步	48
第 3 章 身份安全基础知识	49
理解不同的身份类型	49
社会身份	50
实际身份	50
弱身份	51
利用身份提升用户体验	51
信任区简介	52
浏览器指纹识别	53
阻碍浏览器指纹识别的配置	54
可识别的浏览器信息	55
获取浏览器细节	56

位置追踪	58
设备指纹识别 (手机 / 平板)	61
设备指纹识别 (蓝牙配对设备)	62
实现身份	63

第 4 章 通过 OAuth 2 和 OpenID Connect 实现安全的

登录	64
身份验证和授权之间的区别	64
身份验证	64
授权	65
OAuth 和 OpenID Connect 是什么	65
OAuth 2.0 简介	68
使用 OAuth 2.0 处理授权	70
OAuth 2.0 权限核发类型	72
使用 Bearer Token	73
使用 OpenID Connect 授权和验证身份	73
OAuth 2 和 OAuth 1.0a 之间的安全注意事项	75
构建一个 OAuth 2.0 服务器	75
创建 Express 应用	76
设置服务器的数据库	76
生成授权码和令牌	77
ES5 中 Math.random() 函数的官方文档	79
授权端点	80
处理令牌的存活期	83
处理资源请求	87
使用刷新令牌	89
处理错误	90
为服务器添加 OpenID Connect 功能	94
ID 令牌模式	95
修改授权端点	96
调整令牌端点	97

userinfo 端点	99
使用 OpenID Connect 管理会话	99
构建 OAuth 2 客户端	100
使用授权码	100
使用资源属主凭据或客户端凭据授权	103
为客户端添加 OpenID Connect 功能	105
OpenID Connect 基本流程	105
OAuth 2.0 和 OpenID Connect 之外	107
第 5 章 身份认证的其他方法	108
设备和浏览器指纹识别	108
双因素身份验证和 n 因素身份验证	109
n 因素身份验证	109
一次性密码	110
使用 Authy 实现双因素身份验证	113
使用生物特征代替密码	120
如何评价生物特征的效果	121
面部识别	121
视网膜和虹膜扫描	122
静脉识别	123
新出现的标准	123
FIDO Alliance	123
Oz	126
区块链	126
小结	127
第 6 章 增强 Web 应用的安全	128
保护会话	128
会话的种类	129
Express 处理会话的方式	130
使用 SHA-2 保护密码	131
处理 XSS	134
XSS 攻击的三种类型	134

测试 XSS 保护机制	135
小结	140
CSRF 攻击	140
使用 csrf 处理 CSRF	140
有用的 Node 资源	141
Lusca	142
Helmet	142
Node 安全项目	143
其他减轻危害的技术	144
小结	145
第 7 章 数据传输安全	147
SSL/TLS	147
证书验证类型和权威机构	149
创建供测试的自签名证书	151
异步加密	159
用例	159
具体示例	161
异步加密的优缺点和用途	167
同步加密	168
初始向量	169
填充	170
分组加密的操作模式	172
使用 CTR 加密模式的 AES	174
使用 GCM 验证加密模式的 AES	177
同步加密的优缺点和用途	179
附录 A GitHub 仓库	181
附录 B 技术前提条件和要求	183
词汇表	191

前言

“黑客每年导致企业损失 4 千亿美元。”^{注 1}

—— Inc. 杂志

2015 年第四季度，Cybersecurity Ventures 发布的一份网络安全市场报告指出，网络攻击一年导致企业损失 4 千亿到 5 千亿美元。^{注 2} 为了应对这一问题，2015 年在 IT 安全上的花费增长了 4.7%，达到 754 亿美元，而且预计到明年全球的花费将达 1010 亿美元，2020 年则增长到 1700 亿美元。据此估计，到后年，网络安全从业人员的短缺将达 150 万人，而整体需求将增长到 600 万人。

作为 Web 和应用开发者、设计师、工程师和创作者，我们责无旁贷，必须掌握身份和数据安全方面的知识。如果 Web 开发者不知道如何在传输过程中正确隐蔽数据，不经意间就会敞开网站的安全大门。如果项目经理不知道原本安全的密码算法现已包含瑕疵，不把再次处理数据库中的用户记录当做要事，会导致应用暴露严重的攻击媒介。如今，在一个系统中工作的每个人都要担起责任，确保用户和数据得到保护。

尽管都知道安全第一，但是每周似乎都能见到有公司被攻击了，从初创公司到大型企业，无一幸免。遭泄露的有重要的用户信息、信用卡数据、病例等，而这些都是用户托付给那些公司保护的。总是有很多这样的组织从不把数据加密妥当，所有信息都明文存储，好像等待着黑客去攫取。

注 1: <http://www.inc.com/will-yakowicz/cyberattacks-cost-companies-400-billion-each-year.html>。

注 2: <http://cybersecurityventures.com/cybersecurity-market-report>。

真正的问题是，如今黑客不再单独行事，证明他们可以攻破系统，而是一种有组织的行为，为的是获取金钱或破坏别人的生意。

这就是笔者写作这本书的目的。每一章都会讲解一些概念，告诉你如何封堵现有系统中的漏洞，如何防护可行的攻击媒介，如何在某些自身并不安全的环境中工作。我们将讲解的概念有：

- 了解 Web 和应用安全的现状。
- 构建安全的加密方式，以及与各种密码攻击媒介斗争。
- 创建数字指纹，在浏览器、设备和配对设备中识别用户。
- 通过 OAuth 和 OpenID Connect 构建安全的数据传输系统。
- 使用其他的识别方法提供第二种身份验证方式。
- 加固 Web 应用，防止攻击。
- 使用 SSL/TLS 及同步和异步加密创建安全的数据传输系统。

读完本书后，你将全面了解身份和数据安全的现状，知道如何保护自己，免受潜在的攻击，保护用户托付给你存储的数据，以防泄露。

排版约定

本书使用下述排版约定。

斜体 (*Italic*)

表示新术语、URL、电子邮件地址、文件名和扩展名。

等宽字体 (`Constant Width`)

表示代码清单，在段落中出现则表示程序元素，例如变量、函数名、数据库、数据类型、环境变量、语句和关键字。

粗体等宽字体 (**Constant width bold**)

表示应该由用户输入的命令或其他文本。