



CRC Press  
Taylor & Francis Group

★★★★★  
大数据行业实践与应用译丛

BIG DATA ANALYTICS  
STRATEGIES FOR  
THE SMART GRID

# 智能电网 大数据分析

[美] 卡罗尔·L. 斯蒂米尔 (Carol L. Stimmel) ◎著

张荣◎译

基于电网全业务数据，应用大数据分析，  
实现电网诊断、需求预测和智能规划

**Amit Narayan**  
AutoGrid公司首席执行官

**Adrian Tuck**  
Tendril公司首席执行官

**Ivo Steklac**  
SunPower公司电气住宅和商业能源解决方案总经理

**Ron Gerrans**  
Genus Zero公司首席执行官、E Source前首席执行官

—— 联袂推荐 ——



中国工信出版集团



人民邮电出版社  
POSTS & TELECOM PRESS

# 智能电网 大数据分析

[美]卡罗尔·L. 斯蒂米尔 (Carol L. Stimmel) ◎著  
张荣◎译

基于电网全业务数据，应用大数据分析，  
实现电网诊断、需求预测和智能规划

BIG DATA ANALYTICS  
STRATEGIES FOR  
THE SMART GRID

人民邮电出版社  
北 京

## 图书在版编目 (C I P) 数据

智能电网大数据分析 / (美) 卡罗尔·L. 斯蒂米尔  
(Carol L. Stimmel) 著; 张荣 译. — 北京: 人民邮  
电出版社, 2018.4

(大数据行业实践与应用译丛)

ISBN 978-7-115-47518-3

I. ①智… II. ①卡… ②张… III. ①数据处理—应  
用—智能控制—电网 IV. ①TM76-39

中国版本图书馆CIP数据核字(2017)第316728号

## 版权声明

Carol L. Stimmel

Big Data Analytics Strategies for the Smart Grid

Copyright © 2015 Carol L. Stimmel

Authorized translation from English language edition published

by CRC Press, part of Taylor & Francis Group LLC. All Right Reserved.

本书原版由 Taylor & Francis 出版集团旗下 CRC 出版社出版, 并经其授权翻译出版。版权所有, 侵权必究。

Posts and Telecommunications Press is authorized to publish

and distribute exclusively the Chinese (Simplified Characters) language edition. This edition is authorized for sale throughout Mainland of China. No part of the publication may be reproduced or distributed by any means, or stored in a database or retrieval, without the prior written permission of the publisher.

本书中文简体字翻译版授权由人民邮电出版社独家出版并限在中国大陆地区销售, 未经出版者书面许可, 不得以任何方式复制或发行本书的任何部分。

Copies of this book sold without a Taylor & Francis Sticker on the cover are unauthorized and illegal.

本书封面贴有 Taylor & Francis 公司防伪标签, 无标签者不得销售。

---

◆ 著 [美]卡罗尔·L. 斯蒂米尔 (Carol L. Stimmel)

译 张 荣

责任编辑 李 强

责任印制 彭志环

◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号

邮编 100164 电子邮件 315@ptpress.com.cn

网址 <http://www.ptpress.com.cn>

大厂聚鑫印刷有限责任公司印刷

◆ 开本: 700×1000 1/16

印张: 14.5

2018 年 4 月第 1 版

字数: 236 千字

2018 年 4 月河北第 1 次印刷

著作权合同登记号 图字: 01-2015-3539 号

---

定价: 88.00 元

读者服务热线: (010)81055488 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147 号

---

## 致 谢

大多数人仅对本书的这一段匆匆一瞥就注意到我忘了感谢他们，他们可能会有些恼怒或伤心。因此，对于所有允许我缺席一些重要事件、容忍我忽略了他们的短信、并接受了我为写这本书而给他们布置的任务的人，感谢你们的体谅。对于一直给予我鼓励和安慰的忠实朋友，我希望你们能明白我对你们有多感激。在这里，我要感谢的是我亲爱的朋友 Argot，她读了这本书的每一个字，她放弃了晚上和周末的休息时间来帮助我，让我开心，使我变得更好——请接受我永久和衷心的感谢之情。

就写到这里吧！

---

## 献 词

本书献给我心爱的儿子杰克——为了你极富感染力的热情、不断迸发出的创意、怀疑一切的态度以及满腹的幽默感。我爱你，宝贝。

---

# 序

## 能源数据时代

这本书提供了一个深入的数据分析，这将有助于电力公司高管以及监管机构、投资者、大型电力用户和企业家了解行业的一些结构性变化，而这些改变是一个局外人不能看透的。对于一个混乱的未来，Carol 绘制了一条任何人都可以从中受益的行业发展路径。

——Amit Narayan 博士

预测分析和数据系统将对电力行业产生变革性影响。大数据技术的整合不仅能使电网运营更有效率，也从根本上改变了销售电力的人群、定价方式以及监管者、电力公司、电网运营商和终端用户之间的交互方式。

为了充分了解软件和硬件带来的全部影响，我们必须先退后一步，观察一下电力行业与大多数其他行业的差距有多大。电力公司通常不必通过市场份额竞争来增加收入和利润。相反，它们通过合法垄断的方式生存，享受稳定的价格，并且可以以可预测的方式收回在固定资产上的投资。电力公司最关心的是服务质量、稳定性和可靠性，而不是收入。它们是具有受托责任的公共或准公共机构，为服务领域内的每个人提供基本的服务，对社会安全和福祉产生重要的影响。如果 Facebook 关闭了两个小时，Twitter 就会闹得沸沸扬扬。但如果一个电力公司停电两个小时，高管们必须向消费者群体和管理电力事务的官员解释发生了什么事情。能源部估计，停电和电力质量问题使美国企业每年花费超过 1 000 亿美元。

为了实现这些不同水平的运营要求，电力公司把注意力集中在如何进行控制及消除不同水平的冗余上。峰值发电厂耗资数亿美元，制造一种每年只被使用 50 小时的工具，建造它们是因为它们是一种用来抵消需求的时间的经过验证

（如果效率不高）的工具。一些人反对利用太阳能和风能等可再生能源，这是因为这些能源的变异性太大。为了控制性和一致性，电力公司通过缓冲、暴力（破解）工程和有意限制可选项来弥补可见性和可预测性的缺乏。

能源互联网的搭建改变了上述措施，它能够为电力公司一次性提供实时反馈和洞察。简言之，通过能源互联网，电力公司终于知道它们的客户在做什么，他们想要什么。如今，在电力公司内部，预测是在系统层面完成的，这是一个基本的操作，但实际上它驱动电力公司所有的运营和规划决策。预测每个电表、变压器、馈线和职责范围的能力使我们能够提高这些决策的质量，并削减数十亿美元的运营费用。如果一个中型的欧洲电力公司仅在预测能力方面获得 0.1% 的提升，那么将有助于减少约 300 万美元不平衡市场的运营成本。

在实践中，是什么推动电力公司这样做的？主要影响之一是对需求响应的快速接受。联邦能源管理委员会（FERC）估计，需求响应系统基本上是用动态控制电力消费的云计算平台，这个系统可以替代 188GW 的需求，并且仅在美国就可以避免 4 000 亿美元的峰值电厂的投资。然而，这套系统在很大程度上是针对大型电力公司和大客户的，因为它所要求的硬件成本对于大多数公司而言是负担不起的。

基于软件的需求响应使实现需求响应的成本降低了 90%。更重要的是，它首次引入了需求响应的可见性概念。电力公司，或者更准确地说那些采用云平台的电力公司，可以一次性查看数百万用户的消费模式，并迅速确定哪些客户愿意参与 DR（需求响应）活动、这些客户为参与活动支付了多少费用以及实际上节省了多少费用。

需求响应已经从每年仅能使用几次的、成本昂贵的系统转换为一个电力公司可以日常使用的基本工具，以帮助消费者节省资金，满足社区排放标准，并最大限度地提高固定资产（如发电厂）的回报。

与硬件不同，基于软件的系统会随着时间的推移而改善。想想你自己在 Google、Amazon 或 Netflix 上的经验。这些网络平台随着不断吸收和分析更多的数据将会不断改进。同样，基于软件的需求响应系统将在电力服务方面变得更加专业和有效。可以每隔几分钟向数百万客户发布预测，以对整个地区、特定地区或特定分销部门的用户的耗电量预测进行微调。这些影响最终可能不是很明显。消费者和企业直到收到使他们惊喜的账单时才会知道他们正在节省电力。



类似地，软件可以代替传统的硬件系统进行频率调节和空转备用，在提高性能和加速采用的同时降低成本。就像我们在互联网上看到的那样，可能有无数个可以提供粒度可预测性的系统应用。

电力消耗、传输的控制性和灵活性反过来将为增加太阳能、风电、电动汽车和存储铺平道路。这些技术可以在有数字控制系统辅助的情况下被安全且更容易地集成，并为业主提供更快的投资回报。峰值和非峰值电力之间的价格差距将逐渐消失。

因此，人们将可以看到基础业务如何变化。事实上，它一直在改变。在北美，大约有 44% 的电力在放松管制的市场上销售。Comcast 和其他公司正在进军家庭能源和零售领域。欧洲、澳大利亚和新西兰的电力零售商（这些地方的市场已经放松管制）通过与其他公司竞争，以获取和留住客户，并正在使用一些新方法对客户习惯和负荷曲线进行细分。

使用动态数据分析和控制，让更多的电力供应商能够和电网连接起来，既可作为供方也可作为需方，他们之间的联系也会变得更加紧密。同样，消费者和企业也将更加擅长通过消费模式获利。电力零售商——在得克萨斯和英国已经成为一种特色——随着技术的放松管制将扩大到其他市场。

随着控制电源成为可能，效率也将开始明显好转，并且衡量和监测这些措施的实施结果也更加容易。同时，消费者和企业之间也会变得相互掣肘。据美国能源与环境保护局（Department of Energy and Environmental Protection Agency）表示，商业建筑消耗了美国全部能源的 18%，但接近总消耗 30% 的电力由于浪费或低效率被白白损失掉。在全世界范围内，照明消耗所有电力的 19%——超过核电站和水电站的组合发电——但自动调节亮度来节约电能的系统仍然很少。当你在晚上看着闪闪发光的城市轮廓时，你看到的不止是风景如画的景观，而是一笔巨大、长久以及看似不可阻挡的浪费。

同样，工业用户将开始采用基于云的系统来帮助控制需求费用。需求费用可占大型电力用户账单的 30%。通过采用智能自动化，大型电力用户可以减少基本功耗（如白天照明），同时维持生产运行，并且避免过高的电力峰值。如果没有数据，那么也许大型电力用户只能猜测他们的电力需求是什么。数据通过严格定义可能的结果，有效地消除了风险。

数据也可用于阻挡电力盗窃。世界银行估计，每年有价值 850 亿美元的



电力被盗。在新崛起的国家，这个问题是一场永无止境的危机：在印度大约有30%的电被盗，导致那里长期停电、生产率下降和费率上涨。但在美国这也是一个问题，有价值50亿美元的电力是被非法运营“吞掉”。

在新兴国家中，这些影响甚至更令人瞠目。国际能源机构估计，全球有超过12亿人无法获得电力，超过26亿人无法获得清洁的烹饪设施。为了缓解这个问题，这些国家的许多电网一直在提供肮脏、昂贵和低效率的柴油发电机组。这种情况的根本原因是以供应为中心的电网架构的局限性。由太阳能电池板、电池组和智能数据系统赋予活力的微电网将弥补这些不足。

的确，能源与数据的相互关联已经开始了。最初推出的智能电表——数据时代的基础元素——引起了许多危机。太平洋天然气电力公司的客户对智能电表安装使用表示抗议，声称智能电表对他们的健康构成威胁，侵犯了隐私。

但是，当你去了解一些有争议的头条新闻时，你会看到不同的画面。过去5年来，俄克拉荷马州天然气和电力公司（OG&E）已经开展了成功的项目之一，使用数据来控制能源成本、消耗和排放。该电力公司使用来自Silver Spring Networks、AutoGrid Systems和其他公司的技术，向消费者提供有关高峰定价的信息、管理用户时间计划和其他举措。在早期对6500名客户进行的测试中，得到了一致的积极回应。客户表示，他们不知道高峰定价如何降低他们的账单费用，后来他们改变了使用空调和洗衣机的方法和时间。

OG&E已经将该计划扩展到7万个客户，并在2014年底增加到12万人次。OG&E也赢得了J.D.电力客户服务奖，这对电力公司来说是少有的，OG&E相信，在2020年前，在数据系统的帮助下，它们不需要新建任何发电厂。

智能设备的全球扩散将最终产生名副其实的数字化信息浪潮。一个典型的智能电表每月提供的数据相当于2880只普通电表的总和。到2020年，世界各地的9.8亿智能电表每年将产生431000PB的数据。办公大楼的楼宇管理系统每年将产生约100GB的信息。

实施和整合数据系统将需要时间。我们仍然必须要谨慎任何变革并考虑安全性。但是，变革是不可避免的。消费者将如何精确地与数据进行互动仍有待观察，但我认为电力行业不会走回头路。

数据是新的电力。

——Amit Narayan 博士

---

# 前言

可以肯定地说，这是一本实用的书，也是一本能够带来希望和积极改变的书。电力供应应该深深植根于普遍使用的原则。当所有人都能用上电时，清洁可靠的能源有助于减轻贫困、改善社会现状、加强经济发展。在发达国家，我们知道这是真实的。我们生活的数字化时代证明了能源安全的重要性。在全球范围内，我们看到电气化为经济发展和提高生活质量带来的重要贡献。尽管如此，要实现这个至高无上的工程成就，我们仍然面临着很大的挑战。

现代电气化系统正在蜕变，并且在众多方面是低效率的，但复杂和困难的能源业务经营条件已经在缓慢适应和提升，从而改善这些情况。然而，随着信息化双路电网的到来，我们可以直面这些挑战。这就是本书的主题，通过应用大数据分析以及对全球数百万公里电网的情境感知进行改进，我们将能够整合可再生能源系统，引入经济效益和运营效率，并将能源服务带给全球超过 10 亿不能使用电力的人们。从商业角度来看，电力公司在利用这些关键技术实现快速发展时的确面临着资金困难的问题——这是本书的一个观点。技术专家、电力公司利益相关者、政治机构和能源消费者要坚决采取措施来保护和改善电网的运营，以及做出必要的改变以保护我们的经济、捍卫我们的环境。

我希望凭借一个完全实现的大数据分析策略，阐明大数据分析为电力系统做出的巨大贡献。如今，电力系统能够被史无前例、容量巨大且日益增长的数据所刻画，并且这些数据是可被访问的，而且凭借这些数据，电力系统的消费者们能够瞬时提供一些强有力的洞察力，它们不仅提高了优化日常运营的能力，而且在供不应求时，是有效决策和关键沟通的核心推动力。当面临一些不确定状况时，如极端天气或其他灾害，也能够保证电力输送的安全性和连续性。

今天我们几乎不会看到电网发生故障的情况，但在过去的几十年里，系统的可靠性和效率很低，而且变革一直来得很慢。对于全社会而言，我们必须设

法使电网更有弹性、安全、高效、可靠，并且能够与消费者的生活相融合。智能电网大数据分析固有的技术创新是实现上述目标的第一步并且是踏向美好未来的第一步。

——Carol L. Stimmel

---

## 作者简介

Carol L. Stimmel 于 1991 年在为气候研究写代码和为 3D 系统建模时开始使用“大数据分析”一词——很多年后，这个词已经变成流行词。在这 23 年中，她花了 7 年的时间关注能源行业，包括智能电网数据分析、微电网、家用自动化、数据安全和隐私、智能电网标准和可再生发电。她在职业生涯中的大部分时间参与了新兴技术市场研究，包括做工程、设计新产品以及向电力公司和其他能源行业的利益相关者提供市场情报和数据分析。

Carol 拥有和经营一家数字取证公司，曾与尖端的创业团队合作，共同撰写了关于组织管理的权威作品：*The Manager Pool* (《经理池》)，并在 Gartner、E. Source、Tendril 和 Navigant Research 担任领导。她是一家研究和咨询可持续发展公司——Manifest Mind 有限责任公司的创始人兼首席执行官，该公司为先进的技术项目提供了严格的基于行动的洞察，为人类和环境创造和维护健康的生态系统。Carol 拥有兰道尔夫—麦肯女子学院的哲学学士学位。

---

# 目 录

## 第一部分 数据分析的变革力量

### 第1章 将智能引入电网 3

- 1.1 章节目标 4
- 1.2 建立数据驱动型电力公司的必要性 5
- 1.3 大数据：当我们看到它时，我们了解了 7
- 1.4 什么是数据分析 9
- 1.5 从头开始 11
  - 1.5.1 注意差距 11
  - 1.5.2 文化转型 12
  - 1.5.3 个人案例研究 13
  - 1.5.4 “灵应盘”经济学 14
  - 1.5.5 一如既往的业务对电力公司是致命的 16
  - 1.5.6 生存与灭亡 16
- 1.6 通过智能电网数据分析发现机会 18

### 第2章 构建数据分析基础 19

- 2.1 章节目标 20
- 2.2 毅力是最重要的工具 21
- 2.3 构建数据分析架构 22
  - 2.3.1 数据管理的艺术 24
  - 2.3.2 管理大数据是一个大问题 24
  - 2.3.3 真相不会给你自由 24
  - 2.3.4 每个办法不能“包打天下” 27
  - 2.3.5 解决“特定情境”的难题 27

- 2.3.6 自主构建与外包之争如火如荼地进行着 29
- 2.3.7 当“云”有意义时 31
- 2.3.8 变革既是危险也是机遇 32

### 第3章 让大数据为高价值行动服务 35

- 3.1 章节目标 36
- 3.2 数据型的电力公司 37
- 3.3 算法 38
  - 3.3.1 算法业务 39
  - 3.3.2 数据类别 40
  - 3.3.3 及时性 40
- 3.4 看得见的智能 42
  - 3.4.1 记住人类 44
  - 3.4.2 客户的问题 44
  - 3.4.3 电力公司的变革 46
  - 3.4.4 越大未必越好 47
- 3.5 评估业务问题 48

## 第二部分 智能电网数据分析的优势

### 第4章 在电力公司中应用数据分析模型 53

- 4.1 章节目标 54
- 4.2 了解数据分析模型 55
  - 4.2.1 到底什么是模型 57
  - 4.2.2 警告：相互关系并不意味着因果关系 58
- 4.3 使用描述性模型进行数据分析 59
- 4.4 使用诊断性模型进行分析 60
- 4.5 预测性分析 61
- 4.6 规范性分析 63
- 4.7 电力公司的优化模型 64
- 4.8 面向情境智能 65

## 第 5 章 企业数据分析 67

- 5.1 章节目标 68
- 5.2 超越商业智能 69
  - 5.2.1 电力预测 70
  - 5.2.2 资产管理 70
  - 5.2.3 需求响应和能源分析 72
  - 5.2.4 动态定价分析 78
  - 5.2.5 收入保护分析 80
  - 5.2.6 打破部门间壁垒 81

## 第 6 章 运营分析 83

- 6.1 章节目标 84
- 6.2 调整力量以改善决策 85
- 6.3 洞察的机会 86
- 6.4 关注有效性 87
- 6.5 分布式发电运营：管理混乱 90
- 6.6 电网管理 91
- 6.7 弹性分析 94
- 6.8 从运营数据分析中提取价值 95

## 第 7 章 客户运营和参与分析 97

- 7.1 章节目标 99
- 7.2 提升客户价值 99
  - 7.2.1 客户服务 99
  - 7.2.2 高级客户细分 100
  - 7.2.3 情绪分析 101
  - 7.2.4 收入追缴 102
  - 7.2.5 呼叫中心运营 103
  - 7.2.6 客户沟通 104
- 7.3 为了客户需要具备什么 107



7.3.1 提升账单的价值和面向客户的 Web 门户 108

7.3.2 家庭能源管理 110

7.3.3 战略价值 111

## 第 8 章 网络安全分析 113

8.1 章节目标 114

8.2 电力行业的网络安全 115

8.2.1 对关键基础设施的威胁 115

8.2.2 智能电网是如何增加风险的 116

8.2.3 智能电网是阻止黑夜灾祸的机会 117

8.3 大数据网络安全分析的作用 119

8.3.1 预测和保护 120

8.3.2 网络安全应用 122

8.3.3 主动方法 123

8.3.4 协调网络安全的全球行动 123

8.3.5 风险变化的格局 124

## 第三部分 实施持续变化的数据分析程序

## 第 9 章 寻源数据 129

9.1 章节目标 130

9.2 了解寻源数据 131

9.2.1 智能电表 132

9.2.2 传感器 134

9.2.3 控制设备 135

9.2.4 智能电子设备 136

9.2.5 分布式能源 136

9.2.6 消费者设备 137

9.2.7 历史数据 138

9.2.8 第三方数据 139

9.3 如何处理大量的数据源 140

## 第10章 大数据集成、框架和数据库 143

- 10.1 章节目标 145
- 10.2 这是要花成本的 145
- 10.3 存储方式 146
  - 10.3.1 超大规模存储 146
  - 10.3.2 网络连接存储 146
  - 10.3.3 对象存储 147
- 10.4 数据集成 147
- 10.5 低风险方法的成本 148
- 10.6 让数据流动起来 149
  - 10.6.1 Hadoop 150
  - 10.6.2 MapReduce 151
  - 10.6.3 Hadoop 分布式文件系统 152
  - 10.6.4 如何帮助电力公司 153
- 10.7 其他大数据库 154
  - 10.7.1 NoSQL 154
  - 10.7.2 内存或主内存数据库 155
  - 10.7.3 面向对象的数据库管理系统 156
  - 10.7.4 时间序列数据库服务器 156
  - 10.7.5 空间和 GIS 数据库 156
- 10.8 丰富并非好事 157

## 第11章 提取价值 159

- 11.1 章节目标 160
- 11.2 我们需要明确的答案 161
- 11.3 从数据中挖掘信息和知识 164
- 11.4 数据提取过程 166
  - 11.4.1 当更多不总是更好的时候 168
  - 11.4.2 提升性能 169
  - 11.4.3 Hadoop: 专门为批量数据服务的平台 169
- 11.5 流处理 171