HZ BOOKS
ELSEVIER

Google核心技术丛书

Google Hacking 技术手册

Google Hacking for Penetration Testers

（美） Johnny Long 等著
李 静 等译

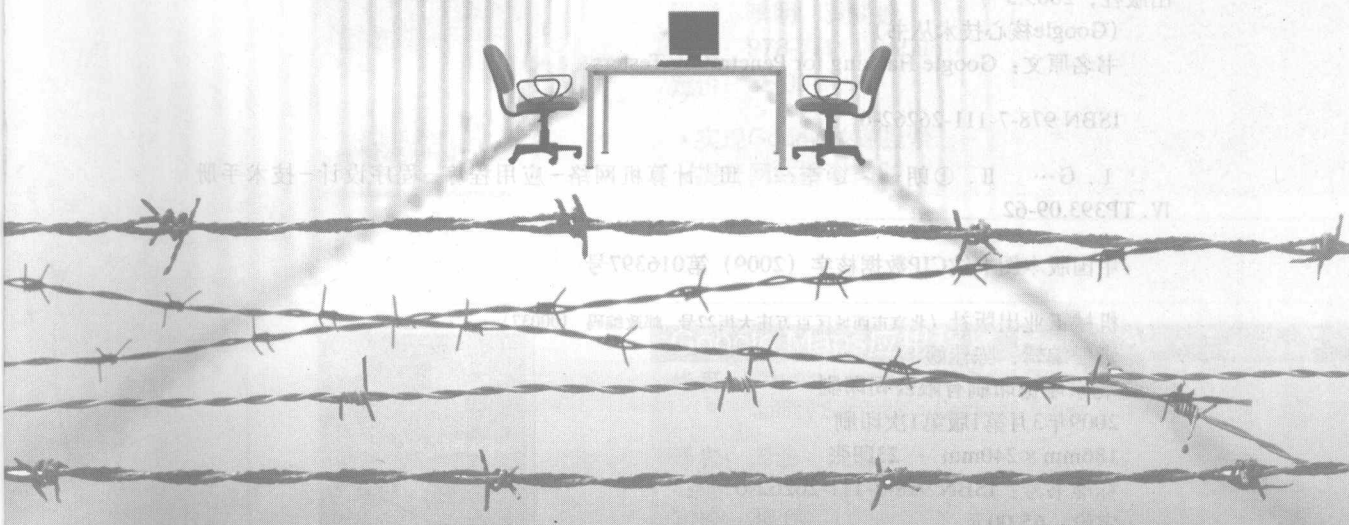
机械工业出版社
China Machine Press

Google核心技术丛书

Google Hacking 技术手册

Google Hacking for Penetration Testers

（美） Johnny Long 等著
李 静 等译



机械工业出版社
China Machine Press

当人们的信息来源渠道日益扩大时，搜索、过滤信息已成为网民必不可少的日常操作，因此Google自然而然地成为了人们常用的工具之一。本书采纳了详尽的Google语法与工具。在发挥Google的最大搜索效用后，你又将如何反其道，让自己的信息无法被别人搜索到呢？只要你细细品读书中的讲解，顺着作者的逆向思维提示，便可轻松地找到甩掉中低级黑客的方法。

本书内容详实，通俗易懂，可作为技术人员的参考用书。

Google Hacking for Penetration Testers
Johnny Long
ISBN:978-1-59749-176-1

ISBN:978-981-272-105-1

图书在版编目（CIP）数据

Google Hacking技术手册/（美）朗格（Long, J.）等著；李静等译. —北京：机械工业出版社，2009.3
（Google核心技术丛书）
书名原文：Google Hacking for Penetration Testers

ISBN 978-7-111-26262-6

Ⅰ. G…　Ⅱ. ①朗…　②李…　Ⅲ. 计算机网络-应用程序-程序设计-技术手册
Ⅳ. TP393.09-62

中国版本图书馆CIP数据核字（2009）第016397号

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：陈佳媛
北京瑞德印刷有限公司印刷
2009年3月第1版第1次印刷
186mm×240mm　·　23印张
标准书号：ISBN 978-7-111-26262-6
定价：65.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换
本社购书热线：（010）68326294

译 者 序

人类的每一个进步都悄然而巨大地改变了人们的生活，计算机的出现使得信息化办公成为可能，网络的出现使得地球成为了“村庄”，搜索引擎更是让网络无秘密可言。查找网页，我们用搜索；查找信息，我们用搜索；查找朋友，我们用搜索。离不开计算机、离不开网络的人们都无一例外地爱上了搜索。一时间，网络拉近了全世界各地的人们之间的距离，搜索改变了人们的生活，让人不禁感慨：“夜夜思君不见君，共用Google搜索。”

然而任何事物都有两面性，你可以使用它来美化生活，也可以用它来破坏世界。Google也不例外，它的强大功能可以为黑客中的“黑帽”所用，也可以为“白帽”工作。在黑客日益增长且黑客技术日益强劲的网络时代，你是否已经察觉到了身边的危机？在网络如此盛行的时代，你不可能脱离网络生活，也正是因为离不开网络，才使得你时刻身陷网络危机之中。你的重要个人信息不但用于电子邮箱注册、论坛注册，还经常在聊天时被你不经意地传送给朋友，你开通了网上购物功能，开通了银行卡的网上购物功能，甚至把家人的敏感信息也登在了自己的博客之中，你可知道这些信息对于黑客来说是多么大的馈赠吗？难道你一点也不担心黑客冒用你的名义做坏事，用不正当的方式刷你的银行卡吗？

现在是潜心了解Google Hacking的时候了！在这么恶劣的网络环境中，维护个人正当权益不能仅仅依靠法律、道德、自觉，更重要的是自己要有防患意识，知道黑客们如何“黑”到你的信息。

本书是一本告诉你黑客是如何通过Google这一强大的引擎在线精确搜索敏感信息并据为己有。阅读本书后，不仅可以大大强化你的搜索知识和能力，还可以将了解到的黑客知识用于防范。书中讲解了很多搜索的技巧和案例，相信这些技巧和案例会使黑客技术爱好者、专业的安全测试或渗透测试人员，以及网络管理人员获益匪浅。总而言之，Hacking的最高境界就是强化搜索别人的功能并避免泄露自己的敏感信息。想要一试究竟的读者不妨仔细地学习并实践一下，跟随本书的步骤，你一定会搜索到让你惊讶的东西。不信？那你就试试看吧！

参与本书翻译的全体人员有：李静、贺倩、李凌燕、贺强、梁晓琴、陈平锋、吴启文、卢祖英、幸慧、陈爱萍、马睿倩、翁子扬、苏建忠和穆陟晅。译者水平有限，书在难免存在瑕疵，敬请广大读者朋友不吝赐教。

译 者

2008年12月

前言

我是Johnny。我爱好黑客技术。

你是否曾有一种会改变你的生活的业余爱好？Google Hacking就成了我的一种爱好，不过在2004年，它成了一份意想不到的礼物。那一年，我参加了Defcon会议并发表了演讲，这是我事业的顶峰。那一年我处于世界之巅，并且有些飘飘然——我实际上是一个微不足道的小人物。我发表了有关Google Hacking的演讲，并在演讲的时候模仿了我的偶像。演讲进行得很顺利，确实引起了诸多评论，并预示着我辉煌演说生涯的开始。前景一片光明，但是到了周末，我却感到空虚。

在两天的时间里，一连串不幸随之而来，把我从成功的顶峰无情地拽到了绝望低谷的峭壁上。过分？有点，不过这是我的真实感受，而且我甚至没有从中得到任何收获。我不确定是什么原因导致我这样，但是我向上张开双臂，把我工作中所有的烦恼，我的事业，我的500位网站用户以及我刚刚开始的演讲生涯，都交给了上帝。

那时，我并不是很清楚那意味着什么，不过我是认真的，我渴望能实现巨大的改变，并且莫名地希望为了更高的目标而生活。我有生以来第一次看到了我生命的浅薄和自私，而这让我感到了震惊。我希望得到更多，而且是希望真的得到。可笑的是，我得到的要比我要求的多。

Syngress出版公司找到了我并问我是否愿意写一本Google Hacking方面的书，也就是你现在看到的这本书的第1版。我极其希望我能掩饰自己经验的缺乏以及对写作的厌恶，因此我接受了这份我后来视为“原创的礼物（original gift）”的邀请。如今，Google Hacking已经成为了畅销书之一。

我的网站用户从500名上升至近8万名。Google图书出版计划衍生了10个左右的额外图书出版计划。媒体如潮汐般涌来，给我留下了极深的印象，首先是Slashdot，紧随其后的是在线媒体、出版界、电视台和有线电台。随着邀请我参加的会议逐渐增多，我得以很快进行了全球旅行。我特别希望成为Hacking社区中的一员，很庆幸他们无条件地接收了我，尽管我最近持有很多保守观点。他们从我的网站购买图书，并将由此产生的收入用于公益事业，他们甚至全额资助我和我妻子的非洲乌干达之旅。这一系列的事件改变了我的生活，并且为ihackcharities.com做好了铺垫，这是一个旨在Hacking社区技术与需要这些技术的公益事业间搭建连接桥梁的组织。我“真正的”生活也得到了改变，我与妻子和孩子的关系比任何时候都要好。

如你所见，对我来说它不仅是一本书。它实际上是一份原创的礼物，我非常认真地进行了本书的升级改版工作。我亲自审核了每一句话和每一幅图片（特别是我编写的那一部分）以保证它的正确性。我以本书的第2版为荣，我很感激各位读者，感谢你支持那些为本书倾注了很

多心血的人们。谢谢你！

欢迎你访问我们的网站http://johnny.ihackstuff.com，感谢你阅读本书。欢迎你链接Google Hacking Database，通过单击我们的Amazon链接来资助公益事业。感谢你为我们提供一个影响现实变化的平台——不只是在安全社区，而是在更大的全球范围。我真诚感谢你的支持。

—Johnny Long

2007年10月

撰稿作者介绍

Roelof Temmingh出生于南非，就读于比勒陀利亚大学，并于1995年获得电子工程技术学位。从那时起，他就一直对计算机安全充满了激情。他曾经是一名开发人员，随后在1995年到2000年之间，在一个信息安全工程公司做一名系统架构师。2000年初，他与一些安全评估和咨询领域中的引领者一同创建了安全评估和咨询公司SensePost。在SensePost工作的期间，他出任评估团队技术主管，之后，他又负责公司的技术革新中心的管理。Roelof曾经在多个国际会议上发表演说，例如Blackhat、Defcon、Cansecwest、RSA、Ruxcon和FIRST会议。本书中由他撰写的内容包括：窃取网络、如何拥有一块陆地，渗透测试者的开放源代码工具箱等，同时他还是“通过数字攻击”训练课程的首席培训师之一。Roelof编写了几个非常有名的安全测试应用程序，如Wikto、Crowbar、BiDiBLAH和Suru。2007年初，他创建了Paterva以便能自主从事研发工作。Roelof在Paterva期间开发了一个称为Evolution（现在称为Maltego）的应用程序，该程序在信息收集和信息关联领域展现了巨大的潜力。

Petko “pdp” D. Petkov是英国伦敦的一名高级IT安全咨询师。他每天的工作就是识别漏洞，构建攻击策略，创建攻击工具和渗透测试基础结构。私底下，人们都知道Petko是pdp或者架构设计师，但是在IT安全行业，他的名字却是因为他强大的技术背景和富有创造性的思维而广为人知。

他最近的一个项目，GNUCITIZEN（gnucitizen.org）是一个领先的在线Web应用程序安全资源，为公众利益而公开了其部分工作。Petko将他自己定位为安全圈内的“酷”猎人。

他与他心爱的女友Ivana一起居住，如果没有他的女友，他将不可能参与本书的撰写。

CP是GHDB和位于http://johnny.ihackstuff.com的论坛的仲裁员，他是Advanced Dork:等许多开源的工具的研发者，是Google站点索引编著者，是http://tankedgenius.com的合作创立者，是自由安全咨询顾问，是DC949（http://dc949.org）的积极分子——在DC中，他参与了一年一度的名为“Amateur/Open Capture the Flag”的hacking竞赛以及各种搜索项目。

“我的身份众多，但是最重要的身份是黑客。”——CP

Jeff Stewart（Jeffball55）正在East Stroudsburg大学主修计算机科学、计算机安全和应用数学。他是johnny.ihackstuff.com论坛的活跃分子，经常在这个论坛中编写与Google服务相关的程序以及Firefox扩展程序。他当前从事的所有项目都可以在http://www.tankedgenius.com上找到。最近，他接了一个关于FD软件计划（FD Sofware Enterprise）的工作：帮助几家医院创建

事故管理系统（Incident Management System）。

Ryan Langley是加利福尼亚人，目前居住在洛杉矶。身为兼职程序员以及安全评估员的Ryan一直都在坚持不懈地研究和学习有关IT安全的知识以及新的评估技术。Ryan拥有五年的系统维修和管理员经验。他经常与CP和Jeffball合作项目。

致谢

此时此刻，我要感谢很多人，不过我不会全部列出他们的名字，但是我会尽最大的努力。

感谢我的妻子和三个可爱的孩子。无法用言语来表达你们对于我的意义。感谢你们包容“真实的”j0hnny。

感谢这本书的团队：CP、Seth Fogie、Jeffball55、L0om、pdp、Roelof Temmingh、Rar、Zanthas。感谢我的朋友Nathan、Mike “Corn” Chaney、Seth Fogie、Arun、@tlas和Apu。感谢Shmoo组中我众多的知己和支持者，ihackcharities志愿者以及支持者，Malcolm Mead和Pat、Predestined（David、Em、Isaac、Josh、Steve、Vanessa）、Tushabe家庭、Dennis以及所有AOET家庭成员。

我还要借此机会感谢Google Hacking社区的成员。正是他们的付出，才让此书以及Google Hacking的运转如期进行。以下列出了这些成员的名字，根据他们对GHDB的发贴数量排序。

Jimmy Neutron（107）、rgod（104）、murfie（74）、golfo（54）、Klouw（52）、CP（48）、L0om（32）、stonersavant（32）、cybercide（27）、jeffball55（23）、Fr0zen（22）、wolveso（22）、yeseins（22）、Rar（21）、ThePsyko（20）、MacUk（18）、crash_monkey（17）、MILKMAN（17）、zoro25（15）、digital.revolution（15）、Cesar（15）、sfd（14）、hermes（13）、mlynch（13）、Renegade334（12）、urban（12）、deadlink（11）、Butt-Pipe（11）、FiZiX（10）、webby_guy（10）、jeffball55+CP（8）、James（7）、Z!nCh（7）、xlockex（6）、ShadowSpoof（6）、noAcces（5）、vipsta（5）、injection33（5）、Fr0zen+MacUK（5）、john（5）、Peefy（4）、sac（4）、sylex（4）、dtire（4）、Deakster（4）、jorokin（4）、Fr0zen rgod（4）、zurik6am（4）、brasileiro（4）、miss.Handle（4）、golfo42（3）、romosapien（3）、klouw（3）、MERLiiN（3）、Darksun（3）、Deeper（3）、jeffball55+klouw（3）、ComSec（3）、Wasabi（3）、THX（3）、putsCTO（3）。

以下所列主体为GHDB提供了两个附件：HaVoC88、ToFu、Digital_Spirit、CP and golfo、ceasar2、namenone、youmolo、MacUK / CP / Klouw、242、golfo、CP and jeff、golfo and CP、Solereaper cp、nuc、bigwreck_3705、ericf、ximum、/iachilles、MacUK/CP、golfo and jeffball55、hevnsnt、PiG_DoG、GIGO、Tox1cFaith、strace、dave@cirt.net、murk、klouw & sylex、NRoberts、X-Ravin、ZyMoTiCo、dc0、Fr0zen jeffball55、Rar CP、rgod jeffball55、vs1400、pitt2k、John Farr、Kartik、QuadsteR、server1、rar klouw、Steve Campbell。

以下主体为GHDB提供了一个附件：Richie Wolk、baxter_jb、D3ADLiN3、accesspwd1、darkwalk、bungerScorpio、Liqdfire、pmedinua、WarriorClown、murfie & webbyguy、stonersavant、klouw、thereallinuxinit、arrested、Milkman & Vipsta、Jamuse and Wolveso、

FiZiX and c0wz、spreafd、blaqueworm、HackerBlaster、FiZiX and klouw、Capboy118、Mac & CP、philY、CP and MacUK、rye、jeffball55 MacUK CP9、rgod + CP、maveric、rar、CP、rgod + jeffball55、norocosul_alex R00t、Solereaper、Daniel Bates、Kevin LAcroix、ThrowedOff、Apoc、mastakillah、juventini、plaztic、Abder、hevensnt、yeseins & klouw、bsdman & klouw & mil、digital.ronin、harry-aac、none90810、donjoe145、toxic-snipe、shadowsliv、golfo and klouw、MacUK / Klouw、Carnage、pulverized、Demogorgo、guardian、golfo、macuk、klouw、Cylos、nihil2006、anonymous、murfie and rgod、D. Garcia、offset、average joe、sebastian、mikem、Andrew A. Vladimirov、bullmoose、effexca、kammo、burhansk、cybercide cybercide、Meohaw、ponds、blackasinc、mr.smoot、digital_revolution、freeeak、zawa、rolf、cykyc、golfo wolveso、sfd wolveso、shellcoder、Jether、jochem、MacUK / df、tikbalang、mysteryman0122、irn-bru、blue_matrix、dopefish、muts、filbert、adsl3000、FiNaLBeTa、draino、bARDO、Z!nCh & vs1400、abinidi、klouw & murfie、wwooww、stonersavant、jimmyn、linuxinit、url、dragg、pedro#、jon335、sfd cseven、russ、kg1、greenflame、vyom、EviL_Phreak、golfo、CP、klouw、rar murfie、Golem、rgod +murfie、Madness!、de Mephisteau、gEnTi、murfie & wolveso、DxM、l0om wolveso、olviTar、digitus、stamhaney、serenh、NaAcces、Kai、goodvirus、barabas、fasullo、ghooli、digitalanimal、Ophidian、MacUK / CP / Jeffb、NightHacker、BinaryGenius、Mindframe、TechStep、rgod +jeffball55 +cp、Fusion、Phil Carmody、johnny、laughing_clown、joenorris、peefy & joenorris、bugged、xxC0BRAxx、Klouw & Renegade334、Front242、Klouw & digital.revo、yomero、Siress、wolves、DonnyC、toadflax、mojo.jojo、cseven、mamba n*p、mynewuser、Ringo，Mac / CP，MacUK / golfo，trinkett，jazzy786，paulfaz，Ronald MacDonald，.-DioXin-.，jerry c、robertserr、norbert.schuler、zoro25 / golfo、cyber_、PhatKahr4u2c、hyp3r、offtopic、jJimmyNeutron、Counterhack、ziggy1621、Demonic_Angel、XTCA2S、m00d、marcomedia、codehunter007、AnArmyOfNone、MegaHz、Maerim、xyberpix、D-jump Fizix、D-jump、Flight Lieutenant Co、windsor_rob、Mac、TPSMC、Navaho Gunleg、EviL Phreak、sfusion、paulfaz、Jeffball55、rgod + cp clean +、stokaz、Revan-th、Don、xewan、Blackdata、wifimuthafucka、chadom、ujen、bunker、Klouw & Jimmy Neutro、JimmyNeutron & murfi、amafui、battletux、lester、rippa、hexsus、jounin、Stealth05、WarChylde、demonio、plazmo、golfo42 & deeper、jeffball55 with cle、MacUK / CP / Klou、Staplerkid、firefalconx、ffenix、hypetech、ARollingStone、kicktd、Solereaper Rar、rgod + webby_guy、googler。

最后，我想重申一下我对第1版中所提到的所有人的谢意，这些人仍与我有关：

感谢我的母亲和父亲，感谢你们允许我对数字生活的废寝忘食。感谢这本书的团队：Alrik“Murf”van Eijkelenborg、James Foster、Steve、Matt、Pete和Roelof。Cooper先生、Elliott夫人、Athy C、Vince Ritts、Jim Chapple、Topher H、Mike Schiffman、Dominique Brezinski和rain.forest.puppy都停下手头的工作来帮助我成就未来。没有我的那些亲密的朋友们的帮助，就没有这本书，他们是Nathan B、Sujay S、Stephen S。感谢Mark Norman让一直保持生活在现实中。来自Google Hacking论坛的Google大师们对论坛和GHDB做了许多的贡献，我很荣幸能在

这里把他们列出来（以发帖数降序排列）：murfie、jimmyneutron、klouw、l0om、ThePsyko、MILKMAN、cybercide、stonersavant、Deadlink、crash_monkey、zoro25、Renegade334、wasabi、urban、mlynch、digital.revolution、Peefy、brasileiro、john、Z!nCh、ComSec、yeseins、sfd、sylex、wolveso、xlockex、injection33、Murk。特别感谢Murf在我写作此书时维护网站的运行，同时也要感谢版主团队：ThePsyko、l0om、wasabi和jimmyneutron。

StrikeForce总是难以描述，但是它占据了我生活的大部分，所以即使我只能玩一小部分也是非常感谢：Jason A、Brian A、Jim C、Roger C、Carter、Carey、Czup、Ross D、Fritz、Jeff G、Kevin H、Micha H、Troy H、Patrick J、Kristy、Dave Klug、Logan L、Laura、Don M、Chris Mclelland、Murray、Deb N、Paige、Roberta、Ron S、Matty T、Chuck T、Katie W、Tim W、Mike W。

感谢CSC和许多我曾经的帅气的老板。你们很牛："FunkSoul"、Chris S、Matt B、Jason E和Al E。感谢'TIP成员让生活充满乐趣。'TIP成员有许多人，但是我只记得一些与我合作比较多的：Anthony、Brian、Chris、Christy、Don、Heidi、Joe、Kevan、The 'Mikes'、"O"、Preston、Richard、Rob、Ron H、Ron D、Steve、Torpedo、Thane。

在写作本书的过程中，我听了许多音乐来掩盖那些噪声。感谢P.O.D（感谢Sonny）、Pillar、Project 86、Avalon O2 remix、D.J. Lex、Yoshinori Sunahara、Hashim and SubSeven（很棒的名字！）（第2版的更新：Green Sector、Pat C.、Andy Hunter、Matisyahu、Bono和U2）。Shouts to securitytribe、Joe Grand、Russ Rogers、Roelof Temmingh、Seth Fogie、Chris Hurley、Bruce Potter、Jeff、Ping、Eli、Blackhat的Grifter，以及作者的整个Syngress家族。我非常荣幸能成为这一团队的一员，尽管在你们面前我很自惭形秽！感谢Andrew和Jaime。你们很牛！

感谢苹果计算机公司制造出如此帅气的笔记本（和操作系统）。

—Johnny Long

目　录

译者序
前言

第1章　Google搜索基础知识 ……1
1.1　简介 ……1
1.2　探索Google的Web界面 ……1
1.2.1　Google的搜索页面 ……1
1.2.2　Google的查询结果页面 ……3
1.2.3　Google Groups ……4
1.2.4　Google图片搜索 ……5
1.2.5　Google使用偏好 ……6
1.2.6　语言工具 ……8
1.3　建立Google查询 ……9
1.3.1　Google搜索的黄金法则 ……10
1.3.2　基本搜索 ……11
1.3.3　使用布尔操作符和特殊字符 ……12
1.3.4　搜索缩简 ……13
1.4　使用Google URL ……16
1.4.1　URL语法 ……17
1.4.2　特殊字符 ……17
1.4.3　组合各个部分 ……18
1.5　总结 ……26
1.6　快速查找解决方案 ……27
1.7　网站链接 ……27
1.8　常见问题 ……28
第2章　高级操作符 ……29
2.1　简介 ……29
2.2　操作符语法 ……29
2.3　Google高级操作符 ……31
2.3.1　Intitle与Allintitle：在页面标题中搜索 ……31
2.3.2　Allintext：在网页内容里查找字符串 ……34
2.3.3　Inurl与Allinurl：在URL中查找文本 ……34
2.3.4　Site：把搜索精确到特定的站点 ……35
2.3.5　Filetype：搜索指定类型的文件 ……37
2.3.6　Link：搜索与当前网页存在链接的网页 ……40
2.3.7　Inanchor：在链接文本中查找文本 ……42
2.3.8　Cache：显示网页的缓存版本 ……43
2.3.9　Numrange：搜索数字 ……43
2.3.10　Daterange：查找在某个特定日期范围内发布的网页 ……44
2.3.11　Info：显示Google的摘要信息 ……44
2.3.12　Related：显示相关站点 ……45
2.3.13　Author：搜索Groups中新闻组帖子的作者 ……46
2.3.14　Group：搜索Group标题 ……47
2.3.15　Insubject：搜索Google Group主题行 ……48
2.3.16　Msgid：通过消息ID来查找Group帖子 ……48
2.3.17　Stocks：搜索股票信息 ……49
2.3.18　Define：显示某个术语的定义 ……50
2.3.19　Phonebook：搜索电话列表 ……50
2.4　操作符冲突与糟糕的Search-Fu ……52
2.5　总结 ……55
2.6　快速查找解决方案 ……55
2.7　网站链接 ……58
2.8　常见问题 ……58
第3章　Google Hacking基础 ……60
3.1　简介 ……60
3.2　使用缓存进行匿名浏览 ……60

3.3 目录列表 ……65
3.3.1 查找目录列表 ……65
3.3.2 查找特定的目录 ……66
3.3.3 查找特定的文件 ……67
3.3.4 服务器的版本 ……67
3.4 险境：遍历技术 ……72
3.4.1 目录遍历 ……72
3.4.2 递增置换 ……73
3.4.3 拓展遍历 ……74
3.5 总结 ……76
3.6 快速查找解决方案 ……76
3.7 网站链接 ……78
3.8 常见问题 ……78
第4章 文档加工与数据库挖掘 ……79
4.1 简介 ……79
4.2 配置文件 ……79
4.3 日志文件 ……84
4.4 数据库挖掘 ……87
4.4.1 登录入口 ……88
4.4.2 帮助文件 ……89
4.4.3 错误消息 ……90
4.4.4 数据库转储 ……95
4.4.5 实际的数据库文件 ……96
4.5 自动加工 ……97
4.6 Google桌面搜索 ……100
4.7 总结 ……101
4.8 快速查找解决方案 ……101
4.9 网站链接 ……102
4.10 常见问题 ……102
第5章 Google在信息收集框架中扮演的角色 ……104
5.1 简介 ……104
5.2 自动搜索原则 ……104
5.2.1 原始搜索关键字 ……106
5.2.2 扩展搜索关键字 ……107
5.2.3 从数据源获取数据 ……112
5.2.4 解析数据 ……123
5.2.5 后期处理 ……129
5.2.6 数据挖掘的应用 ……131
5.3 收集搜索关键字 ……144
5.3.1 在Web上收集 ……144
5.3.2 自行收集 ……145
5.3.3 甜言蜜语 ……149
5.3.4 引用者 ……150
5.4 总结 ……151
第6章 搜索漏洞利用与查找目标 ……152
6.1 简介 ……152
6.2 搜索漏洞利用代码 ……152
6.3 通过常见代码字符串搜索漏洞利用 ……153
6.4 使用Google代码搜索查找代码 ……155
6.5 搜索恶意软件和可执行文件 ……156
6.6 搜索易受攻击的目标 ……160
6.6.1 利用演示页面搜索目标 ……160
6.6.2 利用源代码搜索目标 ……162
6.6.3 利用CGI扫描搜索目标 ……175
6.7 总结 ……177
6.8 快速查找解决方案 ……177
6.9 网站链接 ……178
6.10 常见问题 ……178
第7章 简单有效的安全性搜索 ……180
7.1 简介 ……180
7.1.1 site ……180
7.1.2 intitle:index.of ……181
7.1.3 error | warning ……181
7.1.4 login | logon ……182
7.1.5 username | userid | employee.ID | “your username is” ……183
7.1.6 password | passcode | “your password is” ……184
7.1.7 admin | administrator ……184
7.1.8 -ext:html -ext:htm -ext:shtml -ext:asp -ext:php ……186

7.1.9 inurl:temp | inurl:tmp | inurl:backup | inurl:bak ……188
7.1.10 intranet | help.desk ……189
7.2 总结 ……189
7.3 快速查找解决方案 ……190
7.4 常见问题 ……191
第8章 跟踪搜索Web服务器、登录入口和网络硬件 ……192
8.1 简介 ……192
8.2 定位并剖析Web服务器 ……192
8.2.1 目录列表 ……193
8.2.2 Web服务器软件的错误消息 ……194
8.2.3 应用软件错误消息 ……203
8.2.4 默认页面 ……205
8.2.5 默认文档 ……209
8.2.6 示例程序 ……211
8.3 定位登录入口 ……212
8.4 瞄准使用Web的网络设备 ……225
8.5 查找各种网络报告 ……226
8.6 查找网络硬件 ……227
8.7 总结 ……235
8.8 快速查找解决方案 ……235
8.9 常见问题 ……236
第9章 用户名、口令和其他秘密信息 …239
9.1 简介 ……239
9.2 搜索用户名 ……239
9.3 搜索口令 ……242
9.4 搜索信用卡账号和社保号码等 ……249
9.4.1 社保号码 ……250
9.4.2 个人财务数据 ……251
9.5 搜索其他有利可图的信息 ……251
9.6 总结 ……254
9.7 快速查找解决方案 ……254
9.8 常见问题 ……255
第10章 Hacking Google服务 ……256
10.1 AJAX Search API ……256
10.1.1 嵌入式Google AJAX Search API …257
10.1.2 深入了解AJAX Search ……260
10.1.3 攻击AJAX Search Engine ……263
10.2 Calendar ……267
10.3 Blogger和Google的Blog Search ……269
10.4 信号警报 ……277
10.5 Google Co-op ……278
10.6 Google Code ……283
10.6.1 SVN简介 ……283
10.6.2 在线获取文件 ……284
10.6.3 查找代码 ……286
第11章 Google Hacking案例 ……289
11.1 简介 ……289
11.2 低级信息 ……289
11.2.1 工具 ……290
11.2.2 开放的网络设备 ……292
11.2.3 开放的应用程序 ……298
11.3 摄像头 ……302
11.4 电话设备 ……307
11.5 电源 ……310
11.6 敏感信息 ……312
11.7 社保号码 ……319
11.8 Google之外的信息 ……324
11.9 总结 ……326
第12章 防卫Google黑客 ……327
12.1 简介 ……327
12.2 完善且坚固的安全策略 ……327
12.3 Web服务器安全防护 ……327
12.3.1 目录列表和缺失的索引文件 ……328
12.3.2 利用Robots.txt阻止Crawler ……329
12.3.3 NOARCHIVE：缓存“杀手” ……330
12.3.4 NOSNIPPET：去除摘要 ……331
12.3.5 口令保护机制 ……331
12.3.6 软件默认设置和程序 ……332
12.4 攻击你自己的站点 ……333
12.4.1 用Site操作符搜索自己的站点 ……334
12.4.2 Gooscan ……334

12.4.3 Windows平台下的工具和.NET框架 ……340
12.4.4 Athena ……340
12.4.5 Wikto ……344
12.4.6 Google Rower ……346
12.4.7 Google Site Indexer ……347
12.4.8 Advanced Dork ……349
12.5 从Google获取帮助 ……351
12.6 总结 ……352
12.7 快速查找解决方案 ……352
12.8 网站链接 ……353
12.9 常见问题 ……353

第1章 Google搜索基础知识

1.1 简介

Google的Web界面很清爽。它的“观感效果（look and feel）”[⊖]是受版权保护的。其界面清新而简约。然而大多数人都没有意识到这个界面的功能也是十分强大的。在本书中，我们将了解到怎样利用Google来展示那些真正让人感到惊奇的事。但是，就像日常生活一样，在开始跑之前，要先学会走。

本章简要介绍Google搜索的基础知识。Google功能强大的基于Web的界面，让它成为一个家喻户晓的词，我们就从探索它开始。甚至许多高级Google用户仍在使用基于Web的界面来完成他们的日常查询工作。在掌握了怎样浏览和解释各种界面生成的结果之后，我们将开始探索基本的搜索技巧。

掌握基本的搜索技巧是学习高级查询技巧的基础。你将学到怎样合理地使用Boolean（布尔）操作符（AND、NOT以及OR），同时探索强大和灵活的群搜索。我们也将学到Google几种独特的通配符实现。

最后，我们将了解Google的URL（Uniform Resource Locator，统一资源定位符）结构的语法。学习Google URL的输入和输出能够让你以更快的速度和更大的灵活性来提交一系列相关的Google搜索。Google URL结构为朋友和同事之间交流有趣的搜索提供了一种极棒的简约形式。

1.2 探索Google的Web界面

1.2.1 Google的搜索页面

Google的主页面可以在www.google.com上看到，如图1-1所示。这个界面以其简洁的线条、令人愉悦的整洁感觉和友好的接口而广为人知。尽管这个界面初看起来似乎功能有些少，但是我们将看到许多不同的搜索功能正是从此页开始执行的。

如图1-1所示，用户只可以在一个地方输入内容。这是搜索域（search field）。如果想问Google一个问题或者进行一次查询，你只需简单地把要查找的东西输入进去，然后敲一下Enter（回车键）（前提是你的浏览器支持这种操作），或者点击“Google搜索”（Google Search）按钮，Google就可以给出你想要的查询结果了。

⊖ look and feel（外观和感觉）通常是指程序的图形用户界面的形式与功能，它包含的元素有颜色、形状、布局以及字体等，同时也包含按钮、输入框以及菜单等动态元素的行为。这个术语可用于软件以及网站。由于人们逐渐认识到了look and feel的重要性，所以一些国家制定了相关的版权保护法律。——译者注

图1-1 Google的主页面

在界面顶部的那些链接包括：Web（网页），Images（图片），Video（视频）等。用于打开表1-1所示的其他搜索区域。每一部分的基本搜索功能均相同。我们将在第2章中看到，Google网页界面的每个搜索区域都有不同的功能，并能接受各种不同的查询操作符。例如，author操作符专门设计用于Google Groups搜索领域。表1-1描述了Google主页面的各种不同搜索区域的功能。

表1-1 Google主页面的链接及其功能

界面部分	描　述
Google工具栏	我正在使用的浏览器安装了一个Google“工具栏”，并把它放在了地址栏的后面。下一节将介绍各种不同的Google工具栏
网页，图片，视频，新闻，地图，Gmail和更多标签	这些标签支持用户分别搜索网页，照片，群组发布的消息，Google地图，Google邮件。如果用户初次使用Google，则要知道这些标签并不总能代替“提交搜索”（Submit Search）按钮。这些标签只是简单地切换到其他Google搜索应用
iGoogle	该链接将返回到用户的个性化Google主页
（Signin）登录	该链接支持用户通过登录到个人的Google账户（Google Account）来注册访问额外的功能
搜索项输入域	该文本域位于那些备用的搜索标签的正下方，允许用户输入一个Google搜索项。我们将在整本书中讨论Google搜索的语法
Google Search（Google搜索）按钮	这个按钮是用来提交用户搜索项的。对于大多数浏览器而言，在输入搜索项之后简单地敲击Enter/Return键（回车键）就可以激活这个按钮
I'm Feeling Lucky（手气不错）按钮	与列出搜索结果列表不同的是，这个按钮会挑选出针对输入的搜索项而言最佳的页面。通常这个页面是和输入的搜索项最相关的页面

（续）

界面部分	描　述
Advanced Search（高级搜索）	这个链接可转到高级搜索页面。第2章将介绍这些高级选项
Preferences（使用偏好）	这个链接允许用户设置某些选项（这些设置会保存在用户机器里的cookies中，以便下次访问时使用）。可选的选项包括语言选择，父辈过滤器，每页列出的结果数量以及结果视窗选项
Language tools（语言工具）	这个链接允许你设置许多不同的语言选项以及各种语言之间的翻译

1.2.2 Google的查询结果页面

在处理完一次搜索查询后，Google显示出一个结果页面。如图1-2所示，这个结果页面列出查询结果并且提供包含查询内容的网页。

图1-2 典型的网页查询结果页面

查询结果页面的顶部与主搜索页面相同。注意该页面顶部的Image（图片）、Video（视频）、News（新闻）、Maps（地图）和Gmail（邮件）链接。如果点击搜索页面的这些链接，那么就可以自动地将搜索作为另一种搜索类型重新提交，而不用重新输入一遍。

结果行显示出当前页列出了哪些结果（图中是1—10），约有多少项符合（这里是800万以上），搜索查询自身（包括可以在字典中查找每个单词的链接）以及执行搜索所用的时间。人们通常不会关注查询速度，不过它确实非常迅速。即使有上百万项结果的大量查询也能够在不到一秒的时间内返回！

对于结果页面上的每一项，Google都列出了网站的名称，网站的介绍（通常是该网站内容的前几行），符合查询的页面URL，页面的大小和上次抓取⊖该页面的日期，一个显示Google上次抓取的时候页面内容的经过缓存的链接，以及一个到类似网页的链接。如果查询结果页面未采用你的本地语言，那么Google还支持把该页面翻译成该语言（可以在使用偏好中设置），这时将会出现一个标题为“翻译此页”（Translate this page）的链接，它允许用户能够使用自己的语言来近似阅读该页面（如图1-3所示）。

图1-3 Google翻译

Google搜索背景知识

翻译代理

通过翻译服务，可以把Google当作一个透明的代理服务器来使用。当你点击“翻译此页”（translate this page）链接时，会转到一个位于Google的服务器上的该页面的翻译副本。它可以替你获取相关页面，这可以看作是一种代理服务器。即使你想要查看的页面并不需要翻译，通过修改URL中的hl变量值来匹配该页的本地语言这种办法，你仍然可以把翻译服务作为一种代理服务器来使用。注意，图片并不能以这种方式作为代理。

1.2.3 Google Groups

由于基于Web的论坛、博客、邮件列表和实时消息技术的风靡，作为最原始的公众论坛讨论形式的USENET新闻组已经被人们忽视。但是仍然有许多用户每天都在USENET上发布消息。关于USENET所包含内容的完整介绍可以在www.faqs.org/faqs/usenet/what-is/part1/中找到。DejaNews（www.deja.com）过去曾被认为是收集过去和现在的新闻组消息的权威站点，而它在2001年2月被Google收购（参见www.google.com/press/pressrel/pressrelease48.html）。这个收购使得用户能够通过简单明了的Google搜索界面来查询自1995年以来发布的全部USENET消息存档。Google把USENET新闻组称作Google Groups。今天，全球范围内的互联网用户都转向Google Groups进行讨论和解决某些问题。IT从业人员从Google Groups获得各种技术相关问题的答案已经是非常普遍的了。在Google Groups搜索引擎漂亮的界面背后，古老的USENET交流方式仍然散发着其旺盛的生命力。

Goolge Groups搜索可以通过点击主页面中的Groups标签或者通过浏览http://groups.google.com来访问。虽然它的查询界面（如图1-4所示）看起来和其他的Google查询页面有点儿不同，但是搜索操作方法都是一样的。网页搜索页面和Groups搜索页面的主要不同之处在于新闻组浏览链接。

⊖ 抓取（crawl），又译作爬行，搜索引擎术语，是对搜索引擎访问网站的一个形象说法。是指搜索引擎派出机器人程序（又称为spider，智能代理等）对网页进行抓取并分析，建立索引的过程。——译者注