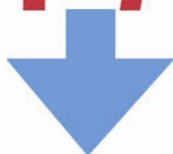


INTERNET ART



互联网艺术

——互联网初学者指南

原著 布伦丹·帕特里克·基鸥
翻译 杨福宝



中国出版集团



世界图书出版公司

互联网艺术

——互联网初学者指南

原著 布伦丹·帕特里克·基欧
翻译 杨福宝



世界图书出版公司

图书在版编目(CIP)数据

互联网艺术 / (爱尔兰)基欧(Kehoe, B. P.) 著; 杨福宝译.

—广州: 世界图书出版广东有限公司, 2011. 9

ISBN 978-7-5100-3981-2

I. ①互… II. ①基… ②杨… III. ①互联网络 IV. ①TP393.4

中国版本图书馆 CIP 数据核字(2011)第 197398 号

互联网艺术

——互联网初学者指南

原 著: (爱尔兰)基欧(Kehoe, B. P.)

翻 译: 杨福宝

责任编辑: 陈 洁

装帧设计: 高 山

出 版: 世界图书出版广东有限公司

发 行: 世界图书出版广东有限公司

(广州市新港西路大江冲 25 号 电话 020-84451969 邮编: 510300)

销 售: 各地新华书店

印 刷: 湖北万隆印务有限公司

开 本: 787×1092 毫米 1/16

印 张: 9.5

字 数: 140 千字

版 次: 2011 年 9 月第 1 版

印 次: 2011 年 9 月第 1 次印刷

ISBN 978-7-5100-3981-2/TP·0009

定价: 36.50 元

服务热线: 020-84183753

关于原著者：“一个真正的黑客和君子”



布伦丹·帕特里克·基欧 (1970. 12. 3—2011. 7. 19.)

爱尔兰的软件开发人员和作家。出生在都柏林，他在美国缅因州 china 镇度过了自己的童年。在刚刚十几岁时，他得到了一台 Commodore 64 电脑，这是他第一次接触计算机。用这台机器，他自学了计算机和计算机网络。在离开高中后，他进入怀德纳大学，在那里继续他的计算机研究。

其主要作品是两本书和数量不少的发表在专业期刊(例如,《Boardwatch 杂志》)上的专业技术文章。他的第一本书,《禅与互联网艺术:初学者指南》,由 Prentice Hall 出版社于 1992 年 7 月首次出版。被誉为第一本广泛发布的互联网用户的启蒙教程。在怀德纳时,他就与出版商协商,以确保该书的第一版将继续在互联网上供大家免费使用。在一项旨在纪念《PC 杂志》发行 20 周年的调查中,他的《禅与互联网艺术》入选了“二十年(1981—2001)来最顶级的非小说类科技书籍”。它也成为了“谢尔盖·布林最受欢迎书籍排行榜”的上榜书。他后来成为 20 世纪 90 年代初的 GNU C 库的开发员及 Cygnus Support 的主要成员。

1993 年 12 月 31 日,他和朋友斯文 Heinicke 出了严重的车

祸,为此留下了脑损伤,包括失语。所幸的是随后他完全康复了。1996年10月5日他结了婚。婚后,他和家人一直住在爱尔兰的都柏林。

2010年,他被任命为“共同教育”(注:爱尔兰教育组织)董事会董事。

不幸的是在2011年3月开始,他被确诊为急性髓系白血病。为与疾病做斗争,他接受了四个月的化疗,但不幸还是于2011年7月19日离开了这个世界,离开了他所钟爱的互联网,以及爱戴他的读者和网民们。埃里克·史蒂文·雷蒙德(注:黑客文化圈内的著名的“行话文件”的维护者和“开源运动”代言人)称他为“一个真正的黑客和君子”。

前 言

编写这本小册子的最初的动因,是由于怀德纳大学计算机科学系对“伟大而新兴的互联网链接”能力描述的迫切需要。

本书尽力使读者了解并熟悉当前互联网的许多性能。针对初学者,本册保持操作系统的“中立性”——对于 Unix、VMS 或其他操作系统涉及的信息较少。希望本册能对所有人有所帮助。

本指南中沿用了传统的印刷风格。所有概念术语如假定的文件名,用户名都用斜体表示。同样,特定的文件名和电子邮件地址用带引号的“”印刷字体表示。用户输入的内容将相对于段落的其他部分缩进,如:

提示 > 命令

结果通常显示在这里。

编写这本小册子有两个目的:首先,它作为一本小参考书,可供人在忙碌中抢读或查阅。另外,它为人们搭建了能够从中探索广袤的互联网的基础。《禅宗与互联网的艺术》不打算花费大量的篇幅针对某一点;相反,它提供了足够多的细节为人们了解其本地的系统。

这里,或许需要提醒的是:我们正在进入的这个领域可能成为一个梦幻般的时间吸收器,时间正一小时一小时地悄悄流逝,人们来来往往,而你则被锁定在网络空间。别忘了要去做你自己的工作哟!

为此,欢迎你——新用户,来上网吧!

brendan@cs.widener.edu
Chester, PA

目 录

第 1 章	网络基础知识	(1)
第 2 章	电子邮件	(7)
第 3 章	匿名文件传输协议	(16)
第 4 章	Usenet 新闻	(28)
第 5 章	远程登录	(49)
第 6 章	各类工具	(66)
第 7 章	商业服务	(92)
第 8 章	你即将了解的互联网	(97)
第 9 章	了解更多	(117)
后记	(120)
附录 A	进入其他网络	(121)
附录 B	通过电子邮件检索文件	(123)
附录 C	创建新闻组	(125)
术语表	(128)
参考文献	(140)
致谢	(144)

第 1 章 网络基础知识

我们确实生活在一个信息社会。相比以往,现在我们最迫切的需求之一就是远距离快速传输大量的信息。从小型单打独斗式企业家的努力到巨型企业,越来越多的专业人士发现,在 90 年代找到成功或取得超越唯一的途径就是意识到技术是在以一个步步紧逼的步伐前进,而他们必须想办法跟上。同样的,来自全球各个角落的研究人员发现他们的工作正在网络环境下茁壮成长。及时分享同事的工作成果,以及方便地查阅虚拟图书馆里成千上万的图书和论文,使得他们能够轻而易举地吸取知识海洋里的营养,而这些在以前是完全无法想象的。现在工作小组成员间可以举行网络交互会议,再也不用考虑地理位置的问题——网络的可能性真是永无止境啊!

短短三十分钟内,你可以用你的手指尖实时地跟身在日本的某个人交流,可以发送 2000 字的短篇故事给一群纯粹以此为乐的人去评价,可以看到加拿大苹果公司的一个实验室里的 Macintosh 电脑是否开着,可以发现澳大利亚是否有人碰巧正端坐在电脑前面(上网登录)。没有哪一家航空公司(或交通和道路驾驶综合系统)对此能够与这种互联网中的“旅行”相媲美。

人们第一次使用网络时遇到的最大的困难是如何才能掌握所有的有用知识。即使是经验丰富的用户在发现存在一个自己从不知道的新的服务或者功能时都会显得十分惊讶。但是一旦熟悉了网络术语,对偶尔出现的错误学会变得心安理得,那么学习网络的进程就会显著地加快。

1.1 域名

人们在使用网络时常遇到的最大困难之一是得知你想去访问什么地方,而那些地方命名方式的多样性最初可能会让你一筹莫

展。不过,你也别太担心,因为这种表面上的迷茫实际上总是有办法解决的。

如果有人寻找一个家庭住址,他就可能期望得到一个街道、公寓、城市、省份、邮政编号。这些是邮局合理高效的收发邮件所需要的基本信息。同样地,计算机地址里也有一个类似的内部结构,通常的形式会是:

个人电子邮件地址:用户名 @ 地址. 域名

电脑名:地址. 域名

除了例外的情形外,大多数情况下,用户的分配地址通常指的是该个人在某一系统中的帐户名。地址. 域名则告诉你该系统或局域的名称,以及组织的类型。而域名的后缀通常是如下某一类:

com

通常是某公司或其他商业机构或组织,比如 convex.com 就是指康威克斯计算机公司

edu

教育机构,如纽约大学被命名为 nyu.edu

gov

政府网站,比如美国航空航天管理局被命名为 nasa.gov

mil

军事网站,例如美国空军(“af.mil”)。

net

网关和其他行政主机对网络(这并不意味着系统内的所有主机网络)。(注:矩阵 111)“near.net”便是指的这样的一个网关。

org

是为某些不喜欢其他类型域名的私人机构所保留的域名,比如电子前线基金会的域名就是 eff.org(参见 8.3.3 [EFF], 第 66 页)

每个国家也有自己的顶级域名。例如,美国的域名包括了五十个州。另外其他国家的域名有:

au

澳大利亚

ca

加拿大

fr

法国

uk

英国。下面还包括一些子域名,比如“ac. uk”代表了英国的学术性网站,而“co. uk 指的是英国的商业性网站。

一个网站的域名的标准用语是指它的完全合格域名(包括地址. 域名加其他),即所谓 FQDN。它通常会很清楚地标明该网站所属的组织或者赞助商。例如,美国麻省理工学院的 FQDN 是“mit. edu”;同样地,苹果电脑公司的域名是“apple. com”。这些域名的所指清晰无误,然而这只是常态,的确还有另外一些例外,它们模棱两可,很容易令人产生误解。就像‘vt. edu’,一开始你会凭直觉去猜测它应该属于在美国佛蒙特州的一个教育机构,可事实并不是如此。它实际上是弗吉尼亚理工大学的域名。在大多数情况下,还是比较容易理解域名的含义的,像这种令人费解的情况毕竟还是比较少的。

1.2 因特网编号

每一个机器在互联网上有一个唯一的地址[注:至少有一个地址,有的也可能有两个甚至三个——但我们不会去探究它],被称为网络编码或 IP 地址。它们往往是由一串用点分隔的名字组成的,用于在数据传输时标识计算机的电子方位(有时也指地理位置)。DNS 是因特网的一项核心服务,它作为可以将域名和 IP 地址相互映射的一个分布式数据库,能够使人更方便地访问互联网,而不用去记住能够被机器直接读取的 IP 地址数串。

每个 IP 地址实际上是一个 32 位的数字,但通常以四个相关联的由“.”号相隔开的数字形成,就像 147. 31. 254. 130. 这个数字有时也被称为四元组;有数以千计的可能组合。ARPnet,今天的互联网之母,最初时受系统分配地址的方式所限,只能连接 256 台系统。在八十年代初期,这种的限制很快得以摆脱,因为新的 32 位的地址分配方式的诞生了,它可以容纳成千上万的主机编号。

每个互联网的地址值(比如 192)被称为一个“八位字节”,代表四个八位元组中的一个。前两个或三个代表系统所在的网络,(例如:192.55.239)称它们为子网。例如,所有卫斯理大学的电脑子网都是 129.133,它们的 IP 号可以为 129.133.10.10,或 133.230.19,像这种可能的组合多达 6.5 万种(也就意味着 6.5 万台计算机)。

另外,在域名中大小写是没有区分的。域名一般不能超过 5 级,从左到右,域的级别逐渐升高,高的级域包含低的级域。域名在整个互联网中是唯一的,当高级子域名相同时,低级子域名不允许重复。一台服务器只能有一个 IP 地址,但是却可以有多个域名。域名是网络寻址方式的“第二代”。

IP 地址和域名分配不是随机的,因为那样的话会导致难以置信的混乱现象。任何一次使用网络都必须在网络信息中心(NIC)存档,不论是通过电子邮件(发到 `hostmaster@nic.ddn.mil`)或者常规的邮件的递送。

1.3 名称和编号

好的,我们可以利用 FQDN 和网络地址来区分计算机,但是一个普通电脑用户怎么可能记住这么多呢?

实际上,他们并不需要记住这么多。互联网的设计使我们可以使用别的任何一种方法。自从人类发现在大多数情况下处理文字比处理数字编号更自然后,每个主机的 FQDN 就被映射到它的英特网编号上了。每个域名都有个域内的电脑为它做服务器,提供所有需要的从域名到 IP 地址的信息,反之亦然。例如,当某人访问 `foosun.bar.com` 服务器时,我们就知道它需要访问 `bar.com` 系统里的 `foosun.bar.com`,也就是想要知道 `foosun.bar.com` 里面的内容;如果 `foosun.bar.com` 确实是存在的,它将会反馈回对应的编号。所有的这些“魔幻”都发生在看不见的后台。

普通用户很少需要记住网站的因特网编号(即使你发现自己记住了一组明显很难理解的编号,可能也仅仅是因为你经常访问该系统而已)。然而,无论如何,你还是会记住很多的 FQDN。最

终,你一定会达到能够仅仅根据学院,大学或者公司的名字就能合理地猜测出它们的域名的程度。

1.4 英特网

互联网是一个巨大的“相互交错的网络系统”,没有一个网络像英特网这样被人们熟知。特别是像 SuraNet, PrepNet, Near-Net 等的局域网,都是内部相互连接在一起成为一个庞大的系统、通过 TCP/IP 协议以惊人的速度相互沟通。而且所有的数据交换都是实时的!

UUCP

即 UNIX 间复制协议(Unix to Unix Copy Protocol)的英文缩写,是贝尔实验室在 1977 年首先开发出来的,用于在他们的 Unix 开发站点之间的通信。它同时包括一个电脑程序以及一个协议,UUCP 允许在未连上因特网的 UNIX 主机间远程执行命令以及传送文件,email 或 netnews. UUCP 包是由多种程序组成。刚开始它是作为一个程序包,用于在串行线路上传输文件、确定这些传输的时间、并且在远程站点上启动程序的执行。

UUCP 网络把通过 UUCP 协议通讯的所有系统松散地连接在一起。它基于两个按照特定的时间间隔相互连接的系统,称为轮询,执行它们其中任何一个的任务列表。过去大多数 UUCP 都是由 Unix 设备来处理,虽然有些软件在其他平台上实现(比如 VMS)。例如,牛至系统每两小时监测一次罗勒系统,如果有任何邮件需要等待给牛至系统,罗勒即刻会传输给它。同样的,牛至系统同时也会传输给罗勒系统任何正在等待的任务。

BITNET(因时网)

因时网(是" Because It's Time Network" 的英文缩略语)是 1980 年由美国纽约州大学和耶鲁大学的研究者们建立的,一种连接世界教育单位的计算机网络,类似于因特网,但是与因特网相互独立,由点对点连结的系统组成,所有的链接都基于 NJE 协议。它是今天互联网的雏形,如 ARPA 网一样。它的功能包括:提供收发电子邮件,交换文件以及在空间内共同分享文字信息。它正

在持续增长中,但却受到了互联网连接成本下降的影响。同时大量的邮件网关准备好了与其他的网络的连接。

1.5 终端连接

多种网络之间的实际连接有很多种不同的形式。最常用的网络连接是 56k 线路(专用电话线路,可以达到 56k 每秒的连接速度)和 T1(特定电话线路,达到 1M 每秒的接连速度)线路。还有的安装了 T3 线路,高达 45M/S 的速度像人体的脊椎一样担负起了网络的主要区域间的数据传输任务。

这些线路都由各机构支付给当地的运营商的(例如,贝尔大西洋公司拥有 PrepNet,是宾西法尼亚州的主要供应商)。此外,还有 SLIP 连接,它用高速调制解调器来传输互联网流量(包)。

大部分的 UUCP 线路运用的是调制解调器,其传输的速度可以从 1200 波特率直升到到 38.4kbps。就像在前面第 1.4 节“网络”的第 8 页中提到过的,这种连接是处于多样的存储并发送状态。基于因特网的 UUCP 连接现在仍在使用(好像情况还没有完全混淆不清)。这个系统完成 TCP/IP 连接的 UUCP 流量统计,它给了 UUCP 基础网络一些飞速跳跃式的发展,导致了整个网络连接的更优化。UUCP 连接在 1970 年首次流行起来,并从此以后保持了广泛的应用。有了 UUCP 连接,乔·史密斯才可以和全国乃至全世界的人们通讯,所付费用不过是相当于本地电话的通话价格。

大部分 BITNET 连接采用 14.4Kbps 的调制解调器连接各个站点。通常一个站点需要连接 3 个或者更多的线路;然而多数站点都有一个唯一连接到“上游”的站点。

“荣耀,没有一个名字。”

拜伦,丘吉尔墓地

第 2 章 电子邮件

网络化的本质是为了交流。当缺少正常的沟通方法时,人们总是想与对方以尽可能最快的方式进行交流。电子邮件(或 email)便是这种愿望在计算机网络中最普遍的实现。它允许人们来回反复地交流而不需要费心去考虑信息实际上是如何传送出去的。随着技术的发展越来越贴近生活,最终成为的日常生活的最常见的一部分,了解它能够被利用到的一些方面以及它是如何工作的这样一种需求,至少从某种程度来说,是至关重要的。

电子邮件的信息传递是任何传统的方式也无法相比的。正是由于电子邮件的使用简易、投递迅速、收费低廉、易于保存、全球畅通无阻,使得电子邮件被广泛地应用,它使人们的交流方式得到了极大的改变。另外,电子邮件还可以进行一对多的邮件传递,同一邮件可以一次发送给许多人。最重要的是,电子邮件是整个互联网以至所有其他网络系统中直接面向人与人之间信息交流的系统,它的数据发送方和接收方都是人,所以极大地满足了大量存在的人与人之间的通信需求。

2.1 电子邮件地址

电子邮件是围绕地址的概念而定的。在引入域名时,网络基础部分提供了一些参考。您的电子邮件地址为您提供接收一条来自世界上任何一个地方的消息所需要的一切信息。一个地址不一定非要针对一个人,它有可能是存档服务器[注:见存档服务器,附录 B 的相关内容描述], 一列名单,或者甚至是某个人的口袋传呼机。这些情形与传统意义上的通信大相径庭——因为传统邮件的地址大多是由人类来阅读的。

电子邮件地址的格式由三部分组成。第一部分“USER”代表用户信箱的帐号,对于同一个邮件接收服务器来说,这个帐号必须

是唯一的；第二部分的“@”是分隔符；第三部分是用户信箱的邮件接收服务器域名，用以标志其所在的位置。

2.1.1 %@!.:符号的杂音

电子邮件地址通常以这样两种形式的其中之一出现——使用互联网格式即包含‘@’，一个“at”符号或使用 UUCP 格式，包含一个感叹号“!”，发音为“砰”。而后者，UUCP 格式的“砰”路径，有更严格的限制，但却更清晰地指示出邮件将如何传送。发送给位于 south.america.org 系统的 Jim Morrison 邮件的地址可能是这样‘jm@south.america.org’。但如果 jim 的帐户名位于一个名为 brazil 的 UUCP 地址，那么他的地址将会是‘brazil! jm’。如果可能的话（事实上，的确存在这种情况），当你尝试使用互联网形式的地址时，如果路径里一个中间的环节偶然失败了，砰路径就会失败。目前，UUCP 地址注册互联网域名表现出一种日益增长的趋势，这可以帮助减少这些路径出现的失败问题。

另外还有一个符号是“%”——它的作用是作为一个额外的“选择途径”的方法。例如，如果 UUCP 地址 dream 与 south.america.org 相连接，但是没有它自身的互联网域名，一个叫 debbie 的用户就可以通过这样一个地址收到邮件：debbie%dream@south.america.org

地址的形式是很重要的。这个地址表示当地系统应该先把邮件发送至 south.america.org。在那里，地址 debbie%dream 就会转变成 debbie@dream，这将有希望成为一个有效的地址。然后 south.america.org 处理所获得的邮件，并将它发至主机 dream，到那里以后，它将会被传送给当地叫 Debbie 的人。

所有的错综复杂的电子邮件寻址方法全部包括在由 O'Reilly 及其合著者出版的《!%@.. 电子邮件寻址和网络目录》这本书里，作为他们的《坚果壳简明手册》系列书的一部分。对于任何电子邮件的常用用户，了解这本书的内容是必不可少的。如果您想订购这本书，请致信给 nuts@ora.com 订购。

2.1.2 发送和接收邮件

我们将从这里的 OS—中间者开始,做一个快速的转换,向你们展示它将会如何在一个 Unix 系统上发送和接收一封邮件。请与您的系统管理员核对有关从您的地址发送邮件的特别指示事项。

有人发送邮件给作者可能会做些像这样的一些事:

```
% mail brendan@cs.widener.edu
```

主题:打印工作的停顿

我输入“输入“印工作 babe.gif”但是没有用!为什么??

下次作者检查他的邮件,他会在他的邮箱里看到:

```
%mail "/usr/spool/mail/brendan":1 条新的未读信息
```

```
U 1 joeuser@foo.widene Tue May 5 20:36 29/956 打  
印工作的停顿?
```

以上信息给出了电子邮件发送者的有关信息,如邮件被发送的时间,邮件的主题等。而邮件接收者可能使用 Unix 邮件的“回复”命令给出这样的回应:

```
? r
```

```
至:joeuser@foo.widener.edu:
```

主题:回复:打印工作的停顿

你不应该用打印机来打印像 GIFs 这样的二进制文件

Brendan

试着多给自己发几次邮件,来学会适应你系统的邮寄者。这样会为您和您的系统管理员省下很多不必要的阿斯匹林(解热镇痛药)。

2.1.3 邮件信头的结构分析

每一封电子邮件都有一个特定的结构,在每一种类型的计算机系统中都是很普遍的。[注:相关标准参阅 RFC—822。更多关于如何取得各种 RFCs 的复制品的信息,请参考 RFCs,第 9.2 章节]典型的电子邮件头通常是这样的:

```
发送者 bush@hq.mil Sat May 25 17:06:01 1991
```


接收者：from hq. mil by house. gov with SMTP id AA21901
(4.1/SMI for dan@house. gov)； Sat, 25 May 91 17:05:56
-0400

日期：Sat, 25 May 91 17:05:56 -0400

来自：The President

消息-ID：<9105252105. AA06631@hq. mil>

至：dan@senate. gov

主题：会议

你好，Dan！早上 9:30 将召开一个董事会的会议，请不要迟到。

第一行的“发送者”和第二行的“接收者”：通常不是大家关注的内容。它们给出邮件出自的“真实”地址（相对于你应该回复的地址，可能看上去会有很大的不同），以及邮件经过了哪些地方最后到达你这里。在因特网上，总是有至少一个‘received:’的信头，而且通常不超过四个或五个。当使用 UUCP 发送一个消息时，一个“接收者”：的信头是邮件经过时被每个系统自动添加的。这种情况经常导致出现很多“接收者”：信头。尽管它们能帮忙检查在递送邮件的过程中出现的问题，但普通用户对此却毫无兴趣。大多数电子邮件程序会自动过滤掉邮件信头中的这种“无聊的东西”。

“日期”中包含信息被传送的日期和时间。同样地，“好”的地址（相对于“真实”地址）在“发送者”：中列出。有时它不会包含人名的全称（例如上例中“总统”这种情况），可能看起来也会不一样，但它应该都含有一个某种形式的电子邮件地址。

信息中的“Message-ID:”主要是用来追踪邮件路由，普通用户很少会对此感兴趣。每一个“Message-ID:”肯定是独一无二的。

“发至”则列出电子邮件接受者的地址（或多个地址）信息。可能会有一个“Cc:”信头，用来列出额外的地址。最后，关于信息的一个简要的主题会列在“subject”里。

还有另外一个问题，那就是电子邮件的安全问题。将电子邮件地址与没有必要知道的人分享，是一个不好的做法。在未经允许的情况下，将电子邮件地址与陌生人分享也是不礼貌的。在发