

在线社交网络信息传播 实证分析与模型研究

吴联仁 著



南开大学出版社

在线社交网络信息传播 实证分析与模型研究

吴联仁 著

南开大学出版社
天 津

图书在版编目(CIP)数据

在线社交网络信息传播实证分析与模型研究 / 吴联仁著. —天津: 南开大学出版社, 2016.1

ISBN 978-7-310-05041-3

I. ①在… II. ①吴… III. ①计算机网络—传播媒介—研究 IV. ①G206.2

中国版本图书馆 CIP 数据核字(2015)第 312554 号

版权所有 侵权必究

南开大学出版社出版发行

出版人: 孙克强

地址: 天津市南开区卫津路 94 号 邮政编码: 300071

营销部电话: (022)23508339 23500755

营销部传真: (022)23508542 邮购部电话: (022)23502200

*

北京楠海印刷厂印刷

全国各地新华书店经销

*

2016 年 1 月第 1 版 2016 年 1 月第 1 次印刷

230×170 毫米 16 开本 13.5 印张 2 插页 253 千字

定价: 30.00 元

如遇图书印装质量问题, 请与本社营销部联系调换, 电话: (022)23507125

本书获得“北京市属高等学校高层次人才引进与培养计划项目 (The Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions)”资助。

序：专心致志，持之以恒

网络科学的研究方兴未艾，课题不断涌现。现在，我们不得不思考这样一个问题：什么是网络科学下一个最前沿的问题？全球复杂网络研究权威、“无标度网络”的创立者艾伯特-拉斯洛·巴拉巴西已经做出了比较明确的回答：“寻找或发现网络的动力学普适规律及其同网络拓扑结构、功能的关系。”正如巴拉巴西的洞察那样，“我们需要攻克的下一个前沿问题，就是理解在网络中发生的动力学过程”。

巴拉巴西在 2005 年发表于《自然》的一篇论文《The Origin of Bursts and Heavy Tails in Human Dynamics》中提出了“从记录人类活动历史的数据库中挖掘出人类行为的统计规律”，从而开创了一个名为“人类动力学”的新研究方向。受这篇开创性文章的影响，大量文章相继在《自然》《科学》和《美国科学院院刊》等期刊上发表，掀起了有关人类行为动力学研究的热潮。

就在同一时期，我们小组在例会谈论中，也开始关注上述研究热点问题。并且，相关研究工作得到了国家自然科学基金“即时通信系统中消息传播模型及其干预机制研究”和教育部新世纪优秀人才“移动互联网情境中人类行为动力学统计特征实证分析及其机理研究”等项目的支持。2010 年，当吴联仁刚到北邮来跟随我攻读博士学位时，结合他的专业背景，我把他引到了复杂网络人类行为动力学这个方向上来。同时，2010 年新浪微博推出，据当年的官方公布数据显示，新浪微博每天发布微博数超过 2500 万条。这为我们小组研究社交媒体用户行为动力学提供了丰富的数据和前所未有的机会。

吴联仁在博士论文开题时，将复杂网络和人类动力学理论应用到社交媒体用户网络和行为动力学研究上，在这一研究问题上经过四五年的研究，得到了一些研究结论和研究成果。我们开始在国内外学术期刊如《Communications in Theoretical Physics》《Physica A: Statistical Mechanics and its Application》《清华大学学报：自然科学版》、《电子科技大学学报》（复杂网络专栏）等期刊上发表一系列关于社交媒体用户网络和行为动力学的文章。这些研究成果的发表凸显了吴联仁博士论文选题的前沿性和重要性。

现在，他在博士论文的基础上补充修改出版本书，我觉得是一件很有意义的

事情。

这里简单介绍一下本书的主要内容和观点。

第一，对新浪微博用户发布微博信息的时间间隔进行了统计分析，发现具有爆发性，时间间隔服从幂律分布。这与巴拉巴西出版的专著《爆发》一书中的结论一致。随后，通过建立基于时间异质性的 SI (Susceptible-Infected) 传播模型，研究了这种爆发性评论行为所导致的时间异质性对微博信息传播的影响。仿真结果得到的结论是：与传统的用户行为时间间隔服从泊松分布的假设相比，在用户行为时间间隔服从幂律分布的情况下，传播速度要缓慢很多，传播范围也较小。

第二，为了理解社交网络的形成及演化机制，在连接最近邻 (Connecting Nearest Neighbors) 机制和强度驱动 (Strength-Driven) 机制的基础上，提出了基于强度驱动的连接最近邻 (Strength-Driven Connect Nearest Neighborhood) 模型。

模型的价值在于反映了社交媒体网络形成的实际情况：两个具有共同好友的用户成为好友的可能性比随机选取的两个用户大得多，同时又考虑了用户选择的偏好性。对模型进行数值仿真，得到的结论有：生成的网络具有服从幂律分布的节点度和节点强度，并且节点度和节点强度呈正的非线性相关关系。

该书的研究工作具有两方面的贡献：第一，丰富了社交网络用户行为动力学研究成果和学科交叉研究成果；第二，书中研究结论为媒体信息传播的引导和控制提供了理论依据。

北京邮电大学经济管理学院教授、博士生导师 闫强

2015 年 10 月

前 言

随着互联网技术和移动通信技术的日益成熟,基于互动服务的新型社交网络应用得到了前所未有的发展机会。社交网络已经成为人们传播信息的工具,以及分享、沟通和交流的重要平台。不同于传统的信息传播方式,社交网络用户的行为模式和社交网络的结构特征使得信息的传播过程变得更复杂和不确定。并且,在传统信息传播建模中过于简化的模型假设和简单的理论方法难以刻画真实社交网络中信息传播的演进过程。鉴于此,本书结合人类动力学理论、复杂网络理论和传播动力学理论等学科的思想和方法,对社交网络环境下的信息传播进行实证分析和建模研究。

对于该交叉热点学科问题,本书从以下四个方面进行了研究。

第一,研究了社交网络用户爆发性评论模式对网络信息流行度的影响。

社交网络的出现不仅改变了人们的通信方式,而且打开了获取大规模人类通信数据的大门,为深入研究人类行为模式和社会互动提供了前所未有的机会。这些研究得到的一个惊人发现是人类活动行为具有爆发性,表现为事件的密集爆发被长的静默间隙所分隔。人类行为明显地偏离了泊松过程,而服从重尾(Heavy-tail)的幂律分布。书中首先对来自新浪微博平台的数据进行统计分析。统计结果显示用户评论信息的行为具有周期性和爆发性。微博信息获得的评论数(即流行度)随时间呈现幂律消亡。信息的生命周期服从幂律分布,表明只有少部分信息可以在微博社区中传播很长时间,具有较高的流行度。并且,在群体水平评论的时间间隔服从幂指数为 2.5 的幂律分布。其次,根据用户的评论行为,对评论数消亡(即流行度消亡)进行建模,并给出理论证明。结果表明用户的评论行为模式与信息的流行度密切相关,表示为群体水平用户连续评论时间间隔分布幂指数与流行度消亡指数相关。最后,将理论证明结果和实证结果进行比较分析,两者都很好地说明了用户的评论行为模式对信息流行度具有影响。当用户评论时间间隔分布幂指数越大,评论爆发性越强,大部分评论集中在很短的时间内,因此信息的消亡速度就越快。该结论为电子商务网站在管理和引导用户评论策略方面提供理论依据。

第二,建立了基于时间异质性和有限注意力的社交网络信息传播模型。实证

研究表明人类的生活节律和活动模式会极大影响社交网络的信息传播,特别是时间异质性。这主要表现为不同个体对信息做出处理的等待时间不同。此外,盛行的在线社交网络使得网络用户每天被大量的信息所淹没,用户的有限注意力也对信息的传播产生了深远的影响。该部分首先将时间异质性和有限注意力这两个因素考虑在内,建立了社交网络信息传播模型。从理论角度对模型进行分析,提出了相关的假设、命题并给予证明。其次,对模型进行数值仿真。仿真结果显示,用户活跃时间间隔分布幂指数越小,即时间间隔异质性越大,信息传播速度越慢,感染用户比例也越小。最后,对新浪微博平台的数据进行统计分析,实证分析结果与仿真结果高度一致。该部分结论为相关政府部门引导和控制社交网络信息或舆论演化、传播提供参考。

第三,对社交网络形成及演化机制进行了建模研究。首先,通过对社交网络的实证数据进行统计分析,获得了网络的相关统计特征,如该网络度分布服从幂指数为 1.9 的幂律分布,节点强度服从幂指数为 1.6 的幂律分布,节点度和节点强度呈非线性相关关系, $s \sim k^{1.5}$ 。为了理解社交网络的网络统计特征,在连接最近邻机制和强度驱动机制的基础上,提出了基于强度驱动的连接最近邻模型。该模型不仅反映了社交网络形成的实际情况,即两个具有共同好友的用户成为好友的可能性比随机选取的两个用户大得多,同时又考虑了用户选择的偏好性。对模型进行数值仿真,仿真结果与理论预测结果和实证数据结果相互一致,验证了模型建立的合理性和准确性。

第四,建立了基于社团结构的信息传播模型。目前,有关社交网络拓扑结构对信息传播的影响的研究,主要集中在社团结构间的链接强度对信息传播的影响方面。而本书定量考虑社交网络中社团结构数量和社团结构间的链接数,详细深入地研究了具有社团结构的社交网络对信息传播速度和传播范围的影响。研究结论表明:对一条特定的信息在具有社团结构的网络上的传播,各社团间的链接数决定了信息传播的速度和范围;传播速度(即个体的平均感染时间)随网络中社团数的增加,表现出先加快后减慢的规律;传播范围(即感染个体比例)随网络中社团数的增加,基本保持在同一水平,影响不大;同时,每个社团结构的内部连接方式(如随机、小世界和无标度)对信息的传播速度和范围也存在影响。该结论具有重要的实际意义:谣言或新闻一般在规模适中的网络中传播得比较快、信息生命周期较长。

目 录

第 1 章 概述	1
1.1 社交网络概述	1
1.2 社交网络主要研究问题	7
1.3 相关概念及研究对象	9
1.4 本书思路与采用工具	10

第一篇 基于人类行为动力学的信息传播研究

第 2 章 人类行为动力学研究综述	19
2.1 人类行为动力学特征实证统计研究进展	20
2.2 人类动力学行为内在机制研究成果	26
2.3 人类动力学理论模型研究	27
2.4 人类行为动力学实证案例	36
第 3 章 传播动力学模型研究综述	47
3.1 谣言传播模型研究	47
3.2 传染病传播模型研究	53
3.3 社交网络信息传播模型研究	60
第 4 章 爆发性评论行为对微博信息传播的影响	73
4.1 引言	73
4.2 微博信息传播研究	74
4.3 爆发性评论行为建模研究	83
4.4 爆发性评论行为对信息传播影响建模	94
4.5 幂律分布检验	98
4.6 本章小结	102
第 5 章 活跃时间异质性对信息传播的影响研究	103
5.1 引言	103
5.2 实证数据分析	104
5.3 活跃时间异质性对信息传播影响建模	107

5.4 理论分析	110
5.5 数值仿真和分析	112
5.6 本章小结	118

第二篇 基于复杂网络的社交网络信息传播研究

第 6 章 社交网络形成及演化机制研究	123
6.1 引言	123
6.2 社团结构形成机制及建模研究	123
6.3 社交网络实证数据分析	127
6.4 在线社交网络建模	132
6.5 连接最近邻 (CNN) 模型	148
6.6 模型建立与数值仿真	152
6.7 本章小结	161
第 7 章 在线社团结构对信息传播的影响	162
7.1 引言	162
7.2 在线社团结构对信息传播影响研究	163
7.3 社团结构对信息传播的影响建模	165
7.4 理论分析	167
7.5 数值仿真	169
7.6 本章小结	171
第 8 章 结论与展望	173
8.1 研究成果及结论	173
8.2 研究展望	174
参考文献	177
主题索引	201
后 记	205

第1章 概述

在社会生活中，每个人都有各自的人际关系网，这些关系网共同构成了现实生活中的社交网络（Social Network）。其中，人是网络中的节点，人际关系是网络中的边。由于人际关系与人们的生活、工作和学习息息相关，因此研究和分析社交网络对提高人们的生活质量有着重要的意义。互联网的产生与发展在很大程度上改变了人们之间的交流方式。在线社交网络作为现实社交网络在互联网上的扩展具有何种特性？它与现实社交网络有何差异？在线社交网络对人们工作和生活产生了哪些影响？这些问题已经引起学术界的关注。为了更好地了解和认识上述问题，本书首先介绍在线社交网络的概念、分类和发展历史，然后探讨学术界在这个领域的研究进展。在不引起歧义的情况下，书中“在线社交网络”统一简称为“社交网络”。

1.1 社交网络概述

1.1.1 社交网络的概念

1967年，哈佛大学的心理学教授斯坦利·米尔格兰姆（Stanley Milgram）创立了“六度分隔理论”，被认为是社交网络的理论基础。六度分隔理论（Six Degrees of Separation）指的是假设世界上所有互不相识的人只需要很少中间人就能建立起联系。而在斯坦利·米尔格拉姆教授的实验中，尝试证明平均只需要5个中间人就可以联系任何两个互不相识的美国人，见图1-1^①。这种现象，并不是说任何人与人之间的联系都必须要通过6个层次才会产生，而是表达了这样一个重要的概念：任何两位素不相识的人之间，通过一定的联系方式，总能够产生必然联系或关系。显然，随着联系方式和联系能力的不同，实现个人期望的机遇将产生明显的区别。

^① 图片来自维基百科。

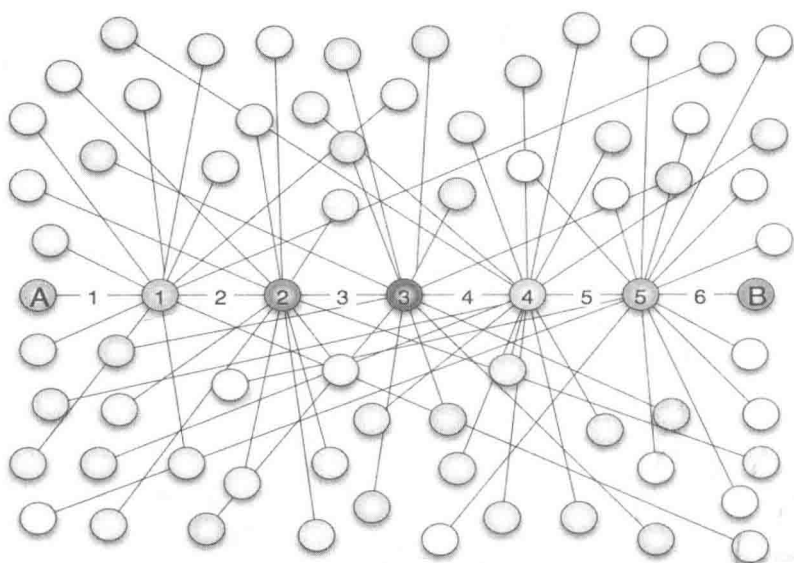


图 1-1 六度分隔理论

按照六度分隔理论，网络上每个用户的社交圈都在不断放大，最后就构成一个大型的社交网络。然而，关于什么是社交网络，目前还没有统一的定义。为便于理解社交网络的概念，本书给出一种供参考和可商榷的描述性定义。

社交网络与传统的万维网（Web）最大的不同之处在于：传统的万维网的主体是内容信息，依靠内容信息组织在一起，呈现给用户；而社交网络的主体是人，依靠人与人之间的朋友关系组织在一起。社交网络必须具备三项基本功能，即允许用户：①创建和维护朋友关系；②上传自己欲分享的内容信息；③浏览其他用户分享的内容信息。但这三项功能在不同的社交网站上的体现形式可能存在较大差异，如 Facebook 只允许用户遍历三层朋友关系，而国内的人人网则没有这个限制。

社交网络是一个系统：①系统中的主体是用户（User），用户可以公开或半公开个人信息；②用户能创建和维护与其他用户之间的连接（或朋友）关系及个人欲分享的内容信息（如日志或照片等）；③用户通过连接（或朋友）关系能浏览和评价朋友分享的信息。

根据用户使用社交网络的目的，以及各社交网络的定位，社交网络可以分为以下几类。

（1）休闲娱乐类社交网络：以满足用户的休闲娱乐需求为主，目前以人人网、腾讯朋友和 QQ 空间、微信朋友圈等为典型代表。

(2) 婚恋交友类社交网络: 以婚恋交友的需求为主, 如世纪佳缘、百合网、珍爱网等。

(3) 商务交友类社交网络: 以商务交友的需求为主, 如优士网、若邻网、经纬网等。

(4) 其他社交网络。

1.1.2 社交网络的起源

社交网络即社交网络服务, 源自英文 SNS (Social Network Service), 中文直译为社会性网络服务或社会化网络服务, 意译为社交网络服务。虽然 SNS 包括硬件、软件、服务及应用, 但由于四字构成的词组更符合中国人的构词习惯, 因此人们习惯上用社交网络来代指 SNS。

SNS 是 1954 年由 Barnes J A 首先使用(“Human Relations”, 在章节“Class and Committees in a Norwegian Island Parish”内)。一个社交网络最大规模约为 150 人左右(这个数值被称为邓巴数, 即 Dunbar's number), 平均规模约为 124 人左右(Hill and Dunbar, 2002)。

社交网络源自网络社交, 网络社交的起点是电子邮件。互联网本质上就是计算机之间的联网, 早期的电子邮件(E-mail)解决了远程的邮件传输的问题, 至今它也是互联网上最普及的应用, 同时它也是网络社交的起点。BBS 则更进了一步, 把“群发”和“转发”常态化, 理论上实现了向所有人发布信息并讨论话题的功能(疆界是 BBS 的访问者数量)。

BBS 把网络社交推进了一步, 即从单纯的点对点交流成本的降低, 推进到了点对面交流成本的降低。即时通信(IM)和博客(Blog)更像是前面两个社交工具的升级版, 前者提高了即时效果(传输速度)和同时交流能力(并行处理); 后者则开始体现社会学和心理学的理论——信息发布节点开始体现越来越强的个体意识, 因为在时间维度上的分散信息开始可以被聚合, 进而成为信息发布节点的“形象”和“性格”。比如从 RSS、Flickr 到最近的 YouTube、Digg、Mini-feed、Twitter、Fetion、Video-Mail 都解决或改进了单一功能, 是丰富网络社交的工具。

随着网络社交的悄悄演进, 一个人在网络上的形象更加趋于完整, 这时候社交网络出现了。

交友只是社交网络的一个开端。就像 Google 的开端只是每个网页的 Backlinks 那么普通一样, 社交网络的开端只是获取用户的个人资料和好友列表。社交网络大体经历了这样一个发展过程: 早期概念化阶段——SixDegrees 代表的六度分隔理论; 结交陌生人阶段——Friendster 帮你建立弱关系从而带来更高社会资本的理论; 娱乐化阶段——MySpace 创造的丰富的多媒体个性化空间吸引注意力的理论;

社交图阶段——Facebook 复制线下真实人际网络来到线上低成本管理的理论；云社交阶段——著云台分布式网际社交理论。整个 SNS 发展的过程是循着人们逐渐将线下生活的更完整的信息流转移到线上进行低成本管理，这让虚拟社交越来越与现实世界的社交出现交叉。

人类历史上，大凡重要的技术革命都伴随媒介革命，几乎人类的任何活动都包含着信息活动，信息流的传递介质、管理方式的不同将决定人们接受信息的不同，所有有关信息流媒介的变革一定是底层的变革——网络社交也是如此。从网络社交的演进历史来看，它一直在遵循“低成本替代”原则。网络社交一直在降低人们社交的时间和物质成本，或者说是降低管理和传递信息的成本。

与此同时，网络社交一直在努力通过不断丰富的手段和工具，来替代传统社交以满足人类这种社会性动物的交流需求，并且正在按照从“增量性的娱乐”到“常量性的生活”这条轨迹不断接近基本需求。

如果说在网络社交的起点——电子邮件时代，网络仅仅可以满足人们 5% 的社交需求，那么今天丰富的社交网络已经可以把这个数字至少提升 10 倍，除了“接触型”的社交行为，或者说是“接触型”信息的收集和发布之外，网络社交已经开始承担大部分传统社交的作用。实际上，“非接触型”的社交，原本就占据了人类社交的 80% 以上，这意味着网络社交对传统世界必然会带来巨大的影响。

其实，网络社交不仅仅是一些新潮的商业模式，从历史维度来看，它更有一股推动互联网向现实世界无限靠近的关键力量。社交网络涵盖以人类社交为核心的所有网络服务形式。互联网是一个能够相互交流、相互沟通、相互参与的互动平台，互联网的发展早已超越了当初 ARPANET 的军事和技术目的。社交网络使得互联网从研究部门、学校、政府、商业应用平台扩展成一个人类社会交流的工具。

网络社交更是把其范围拓展到移动手机平台领域，借助手机的普遍性和无线网络的应用，利用各种交友/即时通讯/邮件收发器等软件，使手机成为新的社交网络的载体。社交网络，也就是“网络+社交”的意思。通过网络这一载体把人们连接起来，从而形成具有某一特点的团体。

也必须注意到，人与人之间的信息沟通，只有近 7% 的信息是通过语言传递的（文字更少），语气、情感、态度、肢体语言等却占了 93%。除非网络社交的发展逼近传统接触型社交效果，否则只能是有益的补充，真正认识和了解一个人还是需要在现实中接触才行，但人们有意无意夸大了网络虚拟社交的作用。其实认识了这点也为网络社交指明了未来发展方向（类似星球大战中的远程全息“面对面”交流）。最近研究表明，依赖 Facebook 进行社交的人易发生抑郁情绪。

1.1.3 社交网络发展史

1971年,人类第一封电子邮件诞生(相应的,本书中称传统水陆路信件为“普通信件”以示区别)。其缘起就是为了方便阿帕网(ARPANET)项目的科学家们互相之间分享研究成果。直到1991年,伯纳斯·李经过多年实践和改进,创办了以“超链接”为特征的万维网(WWW)。

1994年,斯沃斯莫尔学院(Swarthmore College)学生Justin Hall建立了自己的个人站点“Justin's Links from the Underground”,开始与外部网络互联。Justin Hall把这个站点更新了11年,因此被称为“个人博客元勋”(Founding Father)。

1995年,Classmates.com成立,旨在帮助曾经的幼儿园同学、小学同学、初中同学、高中同学、大学同学重新取得联系;Classmates.com在2008年的时候还拥有5000万会员,到2010年才跌出社交网站前十名(TOP 10)。这里有一个罗曼蒂克的鲜活例子:Ray Sears在Classmates.com上找到了自己七年级的女友,他通过站内信息工具问她还记得他,她回答“我怎能忘记我的初恋”,最终两人走进了婚姻的殿堂,并生下两子。

1996年,早期搜索引擎Ask.com上线,它允许人们用自然语言提问,而非关键词(比如:“今天上映什么电影”,而不是“10月23日电影上映”)。1997年,美国在线实时交流工具AIM上线;在这一年,一位名为Jorn Barger的先锋博客作者创造了“Weblog”一词。

1998年,在线日记社区Open Diary上线,它允许人们即使不懂HTML知识也可以发布公开或私密日记。更重要的是,它首次实现人们可以在别人的日记下面进行评论回复。

1999年,博客工具Blogger和LiveJournal出现;后来Blogger在2003年被Google收购,但该产品目前仍然存在——全球科技公司之间的专利站捧红的FOSS Patent就是用Blogger建的网站。

2000年,Jimmy Wales和Larry Sanger共同创建Wikipedia,这是全球首个开源、在线、协作而成的百科全书,完全不同于《大英百科全书》的编纂方式。维基百科(Wiki)的用户在第一年就贡献了20000个在线词条。目前维基百科仍然坚持募捐的方式筹措运营资金,如2011年底他们募集2000万美元来维持2012年的运营。

2001年,Meetup.com网站成立,专注于线下交友。这个网站大家应该比较陌生,但是如果告诉你这个有着12年历史的网站,现在每月还有34万个群组举行线下活动,你应该会感到诧异了。网站的创建者是Scott Heiferman,2001年“9·11”事件以后,他成立了Meetup.com来帮助人们互相联系——而且不只是线

上的。Meetup.com 是一个兴趣交友网站,鼓励人们走出各自孤立的家门,去与志趣相投者交友、聊天。现在它每月会有 34 万个群组在当地社区进行聚会,一起吃喝玩乐、聊天、社交甚至学习。

2002 年, Friendster 上线,这是首家用户规模达到 100 万的社交网络。Friendster 开创了通过个人主页进行交友的先河。两年之后, Facebook 正式在哈佛大学寝室上线。目前在 Facebook 的攻势下, Friendster 在全球范围内基本式微,不过在印尼和菲律宾仍然受欢迎——有意思的是, Facebook 在这两个国家也极其受欢迎,看来那里的人民有强烈的社交欲望。

2003 年,面向青少年和青年群体的 MySpace 上线,它再一次刷新了社交网络的成长速度:一个月注册量突破 100 万。MySpace 发展到后来,涉黄和无谓谩骂逐渐增多而管理者不加规范,以致难以扩大用户量,最终被出售。如果说 2005 年以 5.8 亿美元卖给新闻集团象征着它是未来的新星,那么 2011 年以 3500 万美元卖给广告商则意味着这颗流星陨落。

2003 年上线的还有 WordPress,它由全球各地的几百名网友通过在线协作创建,目前在全球已经拥有数千万用户——截至 2011 年 12 月,发布刚满一年的 WordPress 3.0 获得了 6500 万次下载。与 WordPress 相关的故事不计其数,信息图中讲到它帮助一个自闭症女孩走出病症阴影的故事,这位名叫 Carly Fleishmann 的女孩通过在电脑上敲打文字的方式使自己摆脱了自闭症;随后她创办了 Carly's Voice 的 WordPress 个人博客,帮助其他人摆脱自闭症。

2004 年, Facebook 成立。根据其 2015 年第一季度的财报, Facebook 目前每月有 14.41 亿活跃用户 (Monthly Active Users),每月移动平台活跃用户数有 12.5 亿。

2004 年同年创立的还有 Flickr,现在它依然是非常活跃的图片社区,只不过东家已经变成了雅虎。

2005 年, YouTube 成立,它在成立后迅速被 Google 相中,2006 年从 Google 那里得到的收购价是 16.5 亿美元。

2006 年, Twitter 成立,由于它内容限制在 140 字以内,所以它迅速成为方便的交流工具和强大的自媒体平台。

2006 年成立的还有 Spotify,它现在是社交音乐分享型应用的典型,拥有 1500 万活跃用户和 400 万付费用户。

2007 年,轻博客平台 Tumblr 成立,目前该平台上有 7700 万个博客;根据 2011 年 7 月的数据,该网站每月的独立访问量是 1340 万。信息图中举了一个例子,一个名叫 Ana White 的女木匠在 Tumblr 上开通博客分享自己对木匠工作的喜爱,现在她的博客每月访问量有 300 万,所获得的广告收入足以养家。

2008年, Groupon 上线, 它是国际上最大的团购网站, 但 2012 年股价跌逾 80%, 一直走下坡路。

2009 年, Foursquare 上线, 它是以“签到”(Check-in) 组建基于地理位置的社交网络。Foursquare 成立于纽约市, 每年 4 月 16 日在纽约拥有一个独特的“4SQ 日”。截至 2014 年 4 月, Foursquare 拥有 2000 万注册用户。

2010 年, Google 围绕最成功的产品 G-mail 推出微博客和沟通工具 Google Buzz, 但这是一个失败的产品, 2011 年 12 月 15 日彻底被 Google 终结。

2011 年, Google Buzz 的继承者——Google+ 上线, 根据 2014 年 9 月的数据, Google+ 目前拥有 4 亿的注册用户, 每月 1 亿的活跃用户。

2012 年 Pinterest 呈现爆发式增长(很大一部分原因是它在 2011 年底被 TechCrunch 评为“年度最佳创业公司”), 它是目前网站史上最快达到 1000 万独立访客的网站^①。

1.2 社交网络主要研究问题

基于互联网的社交网络正在成为人类社会管理维系和信息传播的重要载体和渠道, 并对国家安全和社会发展产生深远的影响。在社交网络中, 首先社会个体通过各种网络链接在社交网络上形成“关系结构”; 其次, 社交网络中的大量社会个体以虚拟社区的形式, 围绕某些事件而聚合, 这些社会个体相互依赖和相互影响, 从而形成具有共同行为特征的“网络群体”; 最后, 在基于社交网络的关系结构和网络群体中, 各类“网络信息”快速发布并传播扩散。

近几年, 社交网络的研究引起学术界越来越多的关注, 研究内容包括社交网络的影响力、社交网络中的信息传播(如口碑和谣言)、社交网络的拓扑分析、社会化推荐、在线社会网络的网络结构等, 同时在这方面的研究也取得了大量的研究成果。

(1) 社交网络影响力研究(丁兆云, 2014)。互联网已演变为普遍存在的信息传播工具和平台。在线社交网站、微博、论坛等社交网络应用的出现和迅猛发展, 使得人类使用互联网的方式产生了深刻变革——由简单信息搜索和网页浏览转向网络上社会关系的构建与维护, 基于社会关系的信息创造、交流和共享。社交网络中个体间的交互形成影响力。社交网络中个体的影响力表现在以下四个方面: ①个体间的关系强度; ②个体间的网络距离; ③时间因子; ④网络特征与个体特

^① 社交网络发展史简介来自百度百科。