

文章编号:1007-6735(2011)05-0444-08

社会化标签系统的结构、演化和功能

张子柯^{1,2,3}(1. 杭州师范大学 信息经济研究所, 杭州 310036; 2. 杭州师范大学 阿里巴巴商学院, 杭州 310036;
3. 电子科技大学 互联网科学中心, 成都 611731)

摘要: 从复杂性科学角度总结了社会化标签系统的结构、演化和功能问题. 包括多类异质性节点和超图模型的网络结构; 基于标注行为的网络演化模型; 基于标签的个性化推荐系统. 系统地总结和比较了当前几种代表性的模型和推荐算法, 并指出各方法的优缺点和适用范围. 有助于更深层次理解和解决社会化标签系统中的理论 and 应用问题.

关键词: 社会化标签; 推荐系统; 模型; 超图

中图分类号: F 293.3

文献标志码: A

Social tagging systems: structure, dynamics and function

ZHANG Zi-ke^{1,2,3}(1. *Institute of Information Economy, Hangzhou Normal University, Hangzhou 310036, China;*2. *Alibaba Business College, Hangzhou Normal University, Hangzhou 310036, China;*3. *Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 611731, China*)

Abstract: The recent research progress on understanding the structure, dynamics and function of social tagging systems was summarized. In particular, the following was emphasized: network structures of heterogeneous nodes and hypergraph; tagging behavior based evolving models; tag-aware recommender systems. The representative models and algorithms were systematically studied and compared, and advantages and disadvantages of each, as well as their applicable areas, were pointed out respectively. This survey could shed some lights on deep understanding on both theoretical and application problems of social tagging systems.

Key words: *social tag; recommender system; model; hypergraph*

社会化标签系统研究是以“用户-对象-标签”(user-object-tag)为基本研究对象, 用定量和定性的方法研究其结构关系、演化动力学机制和功能的一个研究方向^[1]. 作为 Web2.0 的一个核心构件^[2], 社会化标签自产生后就得到了科学界和工业界的广泛

关注, 被普遍认为是解决信息爆炸时代数据组织、管理和挖掘的有效工具^[3-4]. 同时, 个性化推荐则被认为是目前解决信息过载问题的有效手段之一^[5-6]. 由于标签中包含着丰富的用户个性化信息, 因此很多学者都尝试将标签功能应用到推荐系统中来进行

收稿日期: 2011-10-03

基金项目: 国家自然科学基金资助项目(11105025, 61103109, 60973069); 四川省科技厅国际合作项目(2010HH0002)

作者简介: 张子柯(1982-), 男, 教授. 研究方向: 复杂网络、社会标签、推荐系统等. E-mail: zhangzike@gmail.com

个性化推荐工作^[7-9]. 其中大多数研究都是从应用层面出发,以设计推荐算法为主. 然而同一推荐算法在不同标签系统的数据上表现迥异,其根本原因在于对不同标签系统的特性不尽相同,而“社会化标签系统的统计结构、演化及其功能应用的关联关系”这个基本问题尚没有理论上清晰和系统的认识.

社会化标签系统的研究工作涉及信息科学、物理学、社会科学、管理科学等多门学科,是一个典型的交叉研究领域. 其研究对象——“用户-对象-标签”三元关系,是一个复杂动态系统. 受互联网和社会经济发展的影响,社会化标签系统在不同时期不同的统计特性有很大不同. 因此,更好地理解标签系统所特有的多节点结构和演化机制,将有利于设计具有高精度性和更为个性化的推荐算法. 从而系统地认识和理解社会化标签系统的结构、演化和功能.

社会化标签系统中的“用户-对象-标签”三元关系网络是典型的复杂动态网络. 而分析这类网络的拓扑结构、统计性质和演化机制,及其对个性化推荐的影响,正是统计物理和非线性动力学的强项,也是当前复杂性科学研究的前沿方向之一. 因此,相关研究不仅可以拓展复杂网络研究所涵盖的对象,而且可以从理论层面一定程度上弥补信息科学、社会科学和管理科学等其他科学在研究社会化标签系统中的理论不足,为交叉科学的研究提供一定的理论支持,并切实有效地解决个性化推荐系统中的重大理论和实践问题.

1 国内外研究现状分析

随着社会化标签系统在互联网、电子商务和图书管理等领域的重要性日益凸显,深入系统地研究其结构、演化和功能应用的研究已成为目前的研究热点. 以下将从社会化标签系统的结构和演化模型,以及其在个性化推荐中的功能应用两个方面介绍相关的研究背景.

1.1 社会化标签系统的结构和演化模型

复杂网络理论被普遍认为是认识和分析复杂动态系统的一个强大而有效的工具^[10-12]. 在一个典型的复杂动态系统中,复杂网络用节点代表个体(如用户),连边表示了个体之间的相互作用. 复杂网络理论就是为了理解和解释这一类复杂动态系统的结构、演化和功能等问题而提出来的^[13-17]. 一般而言,可以用图论的相关理论来描述和刻画复杂网络,其

中最为普遍的就是一部分图^[18]和二部分图理论^[19-20]. 所谓一部分图,是指网络中的所有节点都是同一类节点,如图1所示. 典型的一部分图如社会网络、WWW网络和航空网络等. 而二部分图是指网络中存在两类节点,连边只存在于一类节点到另一类节点之间,同类节点之间没有连边,如图2所示. 如用户购买商品,用户给电影打分等,可以统一抽象为“用户-对象”二部分图. 社会化标签系统除了用户-对象二元关系之外,还存在第三类属性的节点:标签. 用户利用标签管理自己的对象或者寻找自己感兴趣的对象. 社会化标签系统是一类典型的复杂动态系统,它包含了用户、对象和标签3类不同质的信息,构成了其独特的“用户-对象-标签”三元结构^[1],而传统的一部分图和二部分图理论都无法完整的描述和刻画这种结构. 因此很多学者试图寻找新的理论来刻画这种三元结构,其中最受关注的就是三部分图^[21]和超图理论^[22]. 三部分图是对二部分图的延伸,一般来说,是将用户、对象、标签看作成3类不同的节点,连边只存在于不同类别的节点之间,而同类别的节点之间不存在连边. 而超图理论则打破了传统的一条边有且仅有两个节点的约束,允许一条边连接一个及以上数目的节点^[22],如图3所示(见下页).

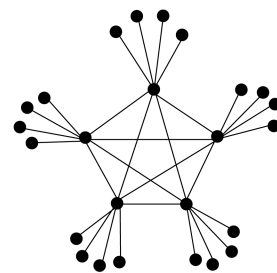


图1 一部分图

Fig.1 Unipartite graph

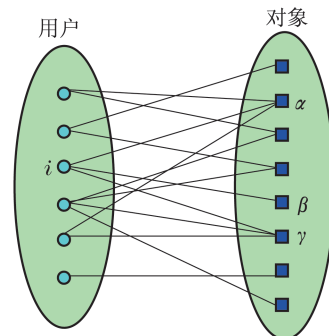


图2 二部分图

Fig.2 Bipartite graph

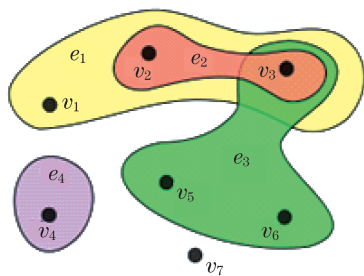


图3 超图

Fig.3 Hyper graph

由于超图理论具有一般性,因此在社会科学^[23]、生物学^[24]和信息科学^[25]等领域都有广泛的应用.而在社会化标签系统中,如果将用户的每次标注行为都看做一条边的话,由于每次标注行为所给的标签数目不同,因此这样的一条边会包含不同数目的节点,这与超图对边的定义是一致的.因此,利用超图来研究标签系统中的结构和统计特性成为了目前该领域的热点课题^[26]. Estrada 等^[27]探索了基于超图的子图中心性和聚类问题. Ghoshal 等^[28]提出了用随机超图模型来表示这种三元结构,其核心思想是定义每条边有且仅有 3 个节点:一个用户节点、一个对象节点和一个标签节点.在此基础上, Zlatic^[29]定义和分析了基于这种三角形模型的统计特性,包括超度分布、超度-簇系数相关性和节点之间的最短距离,并在两个数据集 (Delicious 和 Flickr) 上分别予以展示,都与实际数据的统计特性比较吻合. Neubauer 等^[30]利用超图理论给出了一种在标签网络中寻找最大连通子图的方法,并半定量地指出这种最大连通子图的不同构成是由恶意用户的标注行为导致的,并进一步提出了一种检测恶意用户行为的方法. Palla 等^[31]从自相似的角度研究了标签网络的统计特性.

另外,很多学者利用复杂网络动态建模的思路来研究标签网络的演化规律. Golder 等^[32]较为系统地研究了标签系统的结构和用户使用模式,并且定义了标签的 7 种功能. Cattuto 等^[1]在经典的 Yu-Simon 模型^[33]基础之上,加入了记忆效应,考察了记忆效应对用户标注行为的影响,从一定程度上揭示了标签系统的增长规律.进一步地, Cattuto 等^[34]利用网络中的随机游走思想揭示了标签系统中蕴含的群集动力学机制^[35],该模型能较好的重现标签系统中标签规模分布服从的 Zipf 定律^[36]、标签增长服从次线性的 Heaps 定律^[37]. Halpin 等^[38]提出了标签具有“信息价值”这一理念,定义标签所包含的信息价值是利用它能找到指定资源的概率,并利用用户标

注行为之间的模仿作用设计了标签系统的演化模型,该模型能较好地解释标签系统中的“共现”现象和标签规模分布服从 Zipf 定律,但不能解释标签增长服从次线性的 Heaps 定律.在此基础上, Dellschaft 等^[39]除了考虑用户模仿标注行为外,还考虑了用户自身的背景知识对标签系统演化的影响.该模型能较好地弥补文献^[38]不能解释标签增长服从次线性的 Heaps 定律这一缺陷.

国内在社会化标签系统的结构和建模方面的研究起步较晚,但进展较快,近年来取得了丰硕的成果.王志平、王众托^[40]和何大韧等^[41-43]分别从系统工程科学和复杂网络的角度提出利用“超网络”和交联网理论来刻画和解决复杂动态系统中的结构和应用问题,为复杂系统的建模提供了新思路^[44]. 王建伟等^[45]采用了超边增长和超边优先连接两种增长机制建立了基于超网络的增长模型,并解析给出了标签系统中的超度分布,同时给出了该模型的簇系数和平均最短距离的数值分析结果. 张子柯等^[46]从语义动力学的角度出发,建立了基于用户背景知识和对象、标签双重优先连接机制的超图增长模型,该模型不仅可以得到标签系统中的超度分布的解析解,还能得到标签系统中的超度-簇系数分布的理论解,同时该模型能较好地重现实际数据的统计特性. 刘闯等^[47]提出了用户标注行为存在趋同性,从而最终导致了标签系统中用户标注行为自组织现象的涌现. 裴伟东等^[48]研究了一类三角形结构的动态网络结构的演化模型,利用平均场方法得到了这一类网络结构的理论解,证明了该类演化模型具有和很多真实网络相似的无标度性^[10]和小世界现象^[9].此外,国内学者还利用复杂网络理论对标签系统的统计特性进行了广泛而深入的研究^[49-52].进一步地,刘丹^[53]总结了近年来社会化标签系统的发展和改进方面的研究进展. 杨青云等^[54]总结了当前国际上对用户标注行为研究的现状和最新进展.当然,国内的研究成果远不止于此,由于篇幅限制这里就不做过多介绍.

1.2 基于社会化标签的个性化推荐系统

互联网和 Web2.0 技术的迅速发展将我们带入了信息爆炸的时代.海量数据信息的涌现,在给人们带来更多选择的同时,也导致人们很难快速有效地定位自己所需要的信息.个性化推荐则被认为是解决这一问题的有效工具之一^[5-6],也被认为是复杂网络理论的发展方向之一^[55].由于标签包含了丰富的个性化信息,因此将社会化标签应用到个性化推

荐中也是当前推荐系统领域的一个重要研究课题^[7-8],属于社会计算^[56]的范畴. Hotho 等^[57]使用改进后的 PageRank^[58]算法进行推荐标签的研究,为使用标签进行推荐提供了切实有效的思路. 到目前为止,利用标签进行个性化推荐主要有以下几个方向:

a. 基于三部分图的推荐算法

最近,很多学者从网络角度,利用物质扩散和热传导等物理学方法在“用户-对象”二部分图上来进行个性化推荐研究^[19,59-64]. 同时将标签作为单独一类节点,在“用户-对象-标签”三部分图上利用物理方法进行推荐,也成为该领域的热点课题. 张子柯等^[65]将物质扩散方法分别应用到“用户-对象”和“对象-标签”两个二部分图中,得到对象分别基于用户和标签的得分值,最后通过两个得分混合得到对象的最终得分. 该算法在最优参数下会得到比二部分图更高的推荐准确性、多样性和新颖性. 尚明生等^[66]将该方法应用到以用户为中心的“对象-用户-标签”的三部分图上,进一步提高了推荐效果. 张子柯等^[67]提出标签拥有两种功能:管理和检索功能,并将其应用到推荐系统中,得到了更高的推荐精确性和多样性. 进一步的统计分析表明,该算法尤其对相对冷门的对象有效,从而在一定程度上解决了困扰个性化推荐的“冷启动”难题. Liang 等^[68]指出简单的混合方式会导致信息的丢失,并提出了利用标签所蕴含的语义信息来衡量用户对标签和对象的喜好程度,在此基础上利用混合算法得到更好的推荐效果.

基于三部分图算法存在的不足是将用户、对象、标签看作成3类不同的节点,连边只存在于不同类别的节点之间,同类别的节点之间不存在连边. 这种人为的划分方式,割裂了3类节点相互之间的“共现”关系,从而不可避免地会造成信息丢失. 为了更为完善地解决这个问题,需要从分析社会化标签系统的结构入手,从理论上建立较为完整的理论模型. 只有建立合理的理论模型,这个问题才能更好地得到解决.

b. 基于概率模型的推荐算法

基于概率模型的方法最早出现在信息检索领域中,如潜层语义分析(latent semantic analysis, LSA)^[69],基于概率的潜层语义分析(probabilistic latent semantic analysis, PLSA)^[70]和潜层狄利克雷分配(latent dirichlet allocation, LDA)^[71]. 近年来,很多学者将基于概率模型的方法应用到个性化推荐

中去,期望获得更好的推荐效果. Said 等^[72-73]提出了基于 PLSA 的混合推荐算法. 该算法利用潜层主题变量来衡量“用户-对象”和“对象-标签”的共现概率,再将两者混合起来进行推荐. Krestel 等^[74-75]利用 LDA 模型从对象的标签库里抽取出现层主题,然后在这些主题之上向用户推荐个性化标签. Bundschuh 等^[76-77]则使用 LDA 模型混合利用用户信息和标签信息,并提出用“主题-标签”模型和“用户-主题-标签”模型来获得潜层主题,然后再使用两步迭代来进行个性化标签推荐. Li 等^[78]则将 LDA 模型和 GN 社团发现算法^[79]结合在一起,发现同一个社区的用户倾向于喜好类似的主题,该算法为设计基于社团的推荐算法提供了新思路. Li 等^[80]进一步将概率模型和网络结构模型结合起来,获得了更好的推荐效果.

基于概率模型算法存在的不足是以机器学习为基础. 这种方法一般采用 Gibbs 抽样^[81]或者期望最大化方法^[82]来迭代获得最优的推荐结果,因此对计算机的计算能力要求较高. 特别是当数据量规模非常大的时候,使用基于概率模型的算法将会消耗较长的计算时间. 目前,如何快速有效地设计基于概率模型的推荐算法还是一个难题.

c. 基于协同过滤的推荐算法

基于协同过滤的推荐算法是目前为止应用最为广泛的个性化推荐算法^[83-84]. 其原理是根据具有相似兴趣的用户以往的购买记录来预测目标用户未来可能喜好的物品^[6]. 类似地,用户对相似标签的喜好程度也一定程度上反映了用户之间相似度. Tso-sutter 等^[85]分别计算了基于标签的用户相似度和对象相似度,分别归一化后使用混合方法进行个性化推荐. Liang 等^[86]指出由于标注行为的自由性所导致的噪音问题会影响到推荐性能,因此提出使用“流行”标签来衡量用户相似性,再在协同过滤框架下进行推荐,实验结果表明取得了比文献^[68]和^[85]更高的精确性. 与以往单纯的推荐一类节点不同的是, Peng 等^[87]提出了利用协同过滤算法推荐“对象-标签”对的想法,为个性化推荐拓展了思路. Kim 等^[88]使用标签来过滤用户偏好,该算法对较为稀疏的数据集和“冷启动”用户有一定效果. Durao 等^[89]利用标签之间的语义关联来衡量对象之间的相似度,并在协同过滤框架下进行个性化推荐. 陈超等^[90]则将协同过滤和三部分图结合起来,提出了一种基于含权三部图的协同过滤算法. 该算法除了能提高推荐效果外,还在一定程度上解决了用户“冷启

动”问题.最近国内学者对应用社会化标签进行个性化推荐的研究进展做了详细的综述^[7,91].

基于协同过滤的推荐算法存在的不足是以用户过去的行为来计算相似度的,因此需要大量的历史数据才能较好地衡量相似性.同时,在海量数据的情况下计算相似度也是非常消耗时间的.此外,很多在线用户的行为并不十分显著(比如购买过程中的浏览和比较行为),如何衡量基于这一类用户行为的相似度还是一个研究难题.

2 总结与讨论

从复杂性科学角度总结了社会化标签系统的结构、演化和功能问题研究进展.包括基于三节点结构和超图结构的网络模型、基于个体和群体效应的网络演化建模机制、以及基于社会化标签的个性化推荐系统.作为一类典型的复杂系统,除了一般复杂系统的共性以外,社会化标签系统的特性及其引发的一系列科学问题也引起越来越多的科学界的广泛关注.

目前,社会化标签系统的结构、演化和功能问题的研究已经取得了一些进展.然而,也可以看到,当前的研究仍然存在诸多问题:a. 尽管社会化标签已经在 Web 2.0 网站中得到了广泛应用,然而对社会化标签系统的结构和演化机制缺乏足够清晰的认识.社会化标签系统中的3类节点不是各自孤立的,而是彼此之间有着很紧密的关联关系.因此无法直接用经典的复杂网络理论进行建模分析.需要综合分析用户、对象和标签这3类节点的微观联系,进而更为深刻地理解社会化标签系统的宏观结构和统计特性.而基于超图模型的结构^[22,29,46]将有望处理类似这种异质性节点系统的一种有效理论工具.b. 对用户标注行为习惯缺乏足够的了解.目前对社会化标签系统建模的研究都是假设用户每次只标注一个标签.然而,在实际的社会化标签系统中,每次用户标注行为往往包含了多个标签.因此,需要从大量实际数据中分析用户的同一个标注行为中多个标签的关系,在此基础上,进行用户标注行为的模拟和建模.目前涉及这方面的内容的研究工作还较少.c. 无法定量分析社会化标签系统的结构和演化规律对推荐算法的影响.现有的基于标签的个性化推荐算法都没有考虑标签系统的结构特性,以及不同系统的用户标注行为习惯.因此,正确利用标注行为设计准确性和多样性都较高的推荐算法,将有望成为解决

困扰推荐系统已久的“稀疏性”、“冷启动”等问题的一种有效途径.

参考文献:

- [1] CATTUO C, LORETO V, PIETRONERO L. Semiotic dynamics and collaborative tagging[J]. Proceedings of the National Academy of Sciences, 2007, 104(5): 1461-1464.
- [2] 翟爽,宋文. 社会标签进展研究概述[J]. 图书情报工作, 2010, 54(20): 41-44.
- [3] TRANT J. Studying social tagging and folksonomy: a review and framework[J]. Journal of Digital Information, 2009, 10(1): 1-42.
- [4] DING Y, JACOB E, ZHANG Z, et al. Perspectives on social tagging[J]. Journal of the American Society for Information Science and Technology, 2009, 60(12): 2388-2401.
- [5] ADOMAVICIUS G, TUZHILIN A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions [J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(6): 734-749.
- [6] 刘建国,周涛,汪秉宏. 个性化推荐系统的研究进展[J]. 自然科学进展, 2009, 19(1): 1-15.
- [7] 魏建良,朱庆华. 基于社会化标注的个性化推荐研究进展[J]. 情报学报, 2010, 29(4): 625-633.
- [8] ZHANG Z K, ZHOU T, ZHANG Y C. Tag-aware recommender systems: A state-of-the-art survey[J]. Journal of Computer Science and Technology, 2011, 26(5): 767-777.
- [9] MILICEVIC A K, NANOPOULOS A, IVANOVIC M. Social tagging in recommender systems: A survey of the state-of-the-art and possible extensions[J]. Artificial Intelligence Review, 2010, 33(3): 189-204.
- [10] WATTS D J, STROGATZ S H. Collective dynamics of “small-world” networks[J]. Nature, 1998, 393(6684): 440-442.
- [11] BARABÁSI A L, ALBERT R. Emergence of scaling in random networks [J]. Science, 1999, 286(5439): 509-512.
- [12] 汪小帆,李翔,陈关荣. 复杂网络理论及其应用[M]. 北京:清华大学出版社, 2006.
- [13] ALBERT R, BARABÁSI A L. Statistical mechanics of complex networks[J]. Review of Model Physics, 2002, 74(1): 47-97.
- [14] DOROGOVTSSEV S N, MENDES J F F. Evolution of networks[J]. Advances in Physics, 2002, 51(4): 1079-1187.

- [15] NEWMAN M E J, The structure and function of complex networks [J]. SIAM Review, 2003, 45 (2): 167 - 256.
- [16] BOCCALETTI S, LATORA V, MORENO Y, et al. Complex networks: Structure and dynamics [J]. Physics Reports, 2006, 424(4-5): 175 - 308.
- [17] COSTA L F, RODRIGUES F A, TRAVIESO G, et al. Characterization of complex networks: A survey of measurements [J]. Advances in Physics, 2007, 56(1): 167 - 242.
- [18] NEWMAN M E J, WATTS D J, STROGATZ S H. Random graph models of social networks [J]. Proceedings of the National Academy of Sciences, 2002, 99 (3): 2566 - 2572.
- [19] ZHOU T, REN J, MEDO M, et al. Bipartite network projection and personal recommendation [J]. Physical Review E, 2007, 76(4): 046115.
- [20] SHANG M S, LÜ L, ZENG W, et al. Empirical analysis of web-based user-object bipartite networks [J]. Europhysics Letters, 2009, 90(4): 48006.
- [21] LAMBIOTTE R, AUSLOOS M. Collaborative tagging as a tripartite network [J]. Lecture Notes in Computer Science, 2006(3993): 1114 - 1117.
- [22] VITALY I V. Introduction to Graph and Hypergraph Theory [M]. Hauppauge: Nova Science Publishers, 2009.
- [23] VAZQUEZ A. Population stratification using a statistical model on hypergraphs [J]. Physical Review E, 2008, 77(6): 066106.
- [24] KLAMT S, HAUS U U, THEIS F. Hypergraphs and cellular networks [J]. PLoS Computational Biology, 2009, 5(5): e1000385.
- [25] ZHOU D, HUANG J, SCHOLKOPF B. Learning with hypergraphs: clustering, classification, and embedding [J]. Advances in Neural Information Processing Systems, 2007, 19: 1601 - 1608.
- [26] CATTUTO C, SCHMITZ C, BALDASSARRI A, et al. Network properties of folksonomies [J]. AI Communications, 2007, 20(4): 245 - 262.
- [27] ESTRADA E, RODRIGUEZ-VELÁZQUEZ J A. Subgraph centrality and clustering in complex hyper-networks [J]. Physica A, 2006, 364: 581 - 594.
- [28] GHOSHAL G, ZLATIĆ V, CALDARELLI G, et al. Random hypergraphs and their applications [J]. Physical Review E, 2009, 79(6): 066118.
- [29] ZLATIĆ V, GHOSHAL G, CALDARELLI G. Hypergraph topological quantities for tagged social networks [J]. Physical Review E, 2009, 80(3): 036118.
- [30] NEUBAUER N, OBERMAYER K. Hyperincident connected components of tagging networks [C] // Proceeding of the 20th ACM Conference on Hypertext and Hypermedia. New York: ACM press, 2009: 229 - 238.
- [31] PALLA G, FARKAS I J, POLLNER P, et al. Fundamental statistical features and self-similar properties of tagged networks [J]. New Journal of Physics, 2008, 10 (12): 123026.
- [32] GOLDER S A, HUBERMAN B A. Usage patterns of collaborative tagging systems [J]. Journal of Information Science, 2006, 32(2): 198 - 208.
- [33] YULE G U. A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FRS [J]. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character, 1925, 213: 21 - 87.
- [34] CATTUTO C, BARRAT A, BALDASSARRI A, et al. Collective dynamics of social annotation [J]. Proceedings of the National Academy of Sciences, 2009, 106 (26): 10511 - 10515.
- [35] CASTELLANO C, FORTUNATO S, LORETO V. Statistical physics of social dynamics [J]. Review of Model Physics, 2009, 81(2): 591 - 646.
- [36] ZIPF G K. Human Behaviour and the Principle of Least Effort; An Introduction to Human Ecology [M]. Cambridge: Addison-Wesley Press, 1949.
- [37] HEAPS H S. Information Retrieval: Computational and Theoretical Aspects [M]. New York: Academic Press, 1978.
- [38] HALPIN H, ROBU V, SHEPHERD H. The complex dynamics of collaborative tagging [C] // Proceeding of the 16th International WWW Conference. New York: ACM Press, 2009: 211 - 220.
- [39] DELLSCHAFT K, STAAB S. An epistemic dynamic model for tagging systems [C] // Proceeding of the 9th ACM Conference on Hypertext Hypermedia. New York: ACM press, 2008: 71 - 80.
- [40] 王志平, 王众托. 超网络理论及其应用 [M]. 北京: 科学出版社, 2008.
- [41] 张培培, 侯威, 何阅, 等. 淮扬菜系的网络描述 [J]. 复杂系统与复杂性科学, 2005, 2(2): 49 - 53.
- [42] 张培培, 何阅, 周涛, 等. 一个描述合作网络顶点度分布的模型 [J]. 物理学报, 2006, 55(1): 59 - 66.
- [43] 何阅, 张培培, 许田, 等. 一个科研合作网的双粒子图自适应演化模型 [J]. 物理学报, 2004, 53 (6): 1710 - 1715.
- [44] 席运江, 党延忠. 基于加权超网络模型的知识网络鲁棒性分析方法 [J]. 系统工程理论与实践, 2007, 27

- (4):134-140.
- [45] WANG J W, RONG L L, DENG Q H. Evolving hyper-network model[J]. *European Journal of Physics B*, 2010, 77(4):493-498.
- [46] ZHANG Z K, LIU C. A hypergraph model of social tagging networks[J]. *Journal of Statistical Mechanics*, 2010, (10):P10005.
- [47] LIU C, YEUNG C H, ZHANG Z K. Self-Organization in Social Tagging Systems[J]. *Physical Review E*, 2011, 83(6):066104.
- [48] 裴伟东, 夏玮, 王全来, 等. 一类三角形结构动态复杂网络演化模型分析[J]. *中国科学技术大学学报*, 2010, 40(11):1186-1190.
- [49] 鲁晓明. Web 2.0 中社会性标签系统的复杂网络特性研究[J]. *现代情报*, 2007, 27(12):64-66.
- [50] 李静. Folksonomy 的网络性质分析[J]. *现代情报*, 2009, 29(7):61-63.
- [51] 吴超, 周波. 基于复杂网络的社会化标签分析[J]. *浙江大学学报(工学版)*, 2010, 44(11):2194-2197.
- [52] 刘知远, 司宪策, 郑亚斌, 等. 中文博客标签的若干统计性质[C]//中国计算技术与语言问题研究——第七届中国信息处理国际会议论文集. 北京:电子工业出版社, 2007:533-539.
- [53] 刘丹. 近年来国外 Folksonomy 的改进研究进展[J]. *图书情报工作*, 2010, 54(8):55-59.
- [54] 杨青云, 裴雷, 吴克文. 国外社会化标注系统中标注行为研究现状[J]. *资源管理*, 2009, 28(11):185-188.
- [55] NEWMAN M E J. Complex systems: A survey [J]. 2011, to be published.
- [56] 王飞跃. 社会计算:科学技术人文[J]. *中国科学院院*, 2005, 20(5):370-376.
- [57] HOTH O A, JASCHKE R J, SCHMITZ C, et al. Information retrieval in folksonomies: Search and ranking[J]. *Lecture Notes in Computer Science*, 2006, 4011:411-426.
- [58] BRIN S, PAGE L. The anatomy of a large-scale hypertextual Web search engine[J]. *Computer Networks ISDN System*, 1998, 30(1):107-117.
- [59] ZHANG Y C, BLATTNER M, YU Y K. Heat conduction process on community networks as a recommendation model [J]. *Physical Review Letters*, 2007, 99(15):154301.
- [60] ZHOU T, KUSCSIK Z, LIU J G, et al. Solving the apparent diversity-accuracy dilemma of recommender systems[J]. *Proceedings of the National Academy of Sciences*, 2010, 107(10):4511-451.
- [61] ZHOU T, SU R Q, LIU R R, et al. Accurate and diverse recommendations via eliminating redundant correlations [J]. *New Journal of Physics*, 2009, 11(12):123008.
- [62] SUN D, ZHOU T, LIU J G, et al. Information filtering based on transferring similarity[J]. *Physical Review E*, 2009, 80(1):17101.
- [63] ZHOU T, JIANG L L, SU R Q, et al. Effect of initial configuration on network-based recommendation[J]. *Europhysics Letters*, 2008, 81(5):58004.
- [64] LIU J G, ZHOU T, CHE H A, et al. Effects of high-order correlations on personalized recommendations for bipartite networks [J]. *Physica A*, 2010, 389(4):881-886.
- [65] ZHANG Z K, ZHOU T, ZHANG Y C. Personalized recommendation via integrated diffusion on user-item-tag tripartite graphs [J]. *Physica A*, 2010, 389(1):179-186.
- [66] SHANG M S, ZHANG Z K, ZHOU T, et al. Collaborative filtering with diffusion-based similarity on tripartite graphs[J]. *Physica A*, 2010, 389(6):1259-1264.
- [67] ZHANG Z K, LIU C, ZHANG Y C, et al. Solving the cold-start problem in recommender systems with social tags[J]. *Europhysics Letters*, 2010, 92(2):28002.
- [68] LIANG H, XU Y, LI Y, et al. Connecting users and items with weighted tags for personalized Item recommendations[C]//Proceeding of the 21st ACM Conference on Hypertext and Hypermedia. New York: ACM press, 2010:51-60.
- [69] DEERWESTER S, DUMAIS S T, FURNAS G W, et al. Indexing by latent semantic analysis[J]. *Journal of the American Society for Information Science and Technology*, 1990, 41(6):391-407.
- [70] HOFMANN T. Probabilistic latent semantic indexing [C]//Proceeding of the 22nd Ann. Intl. ACM SIGIR Conference on Research and Development Information Retrieval. New York: ACM press, 1999:50-57.
- [71] BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation[J]. *Journal of Machine Learning Research*, 2003, 3(3):993-1022.
- [72] SAID A, WETZKER R, UMBRATH W, et al. A hybrid PLSA approach for warmer cold start in folksonomy recommendation[C]//Proceeding of ACM Recsys 2009 Workshop on Recommender Systems & the Social Web. New York: ACM Press, 2009:87-90.
- [73] WETZKER R, UMBRATH W, SAID A. A hybrid approach to item recommendation in folksonomies[C]//Proceedings of the Workshop Exploiting Semantic Annotation Information Retrieval. 2009:25-29.
- [74] KRESTEL R, FANKHAUSER P. Tag recommendation

- using probabilistic topic models[C] // Proceeding of ECML PKDD Discovery Challenge. 2009:131 - 141.
- [75] KRESTEL R, FANKHAUSER P, NEJDL W. Latent dirichlet allocation for tag recommendation[C] // Proceedings of the 3rd ACM Conference on Recommender Systems. New York:ACM press,2009:61 - 68.
- [76] BUNDSCHUS M, YU S, TRESP V, et al. Hierarchical bayesian models for collaborative tagging systems[C] // The 9th IEEE International Conference on Data Mining. Los Alamitos: IEEE Computer Society Press, 2009:728 - 733.
- [77] BUNDSCHUS M, TRESP V, KRIEGEL H P. Topic models for semantically annotated document collections[R]. Whistler;NIPS workshop,2009.
- [78] LI D, HE B, DING Y, et al. Community-based Topic Modeling for Social Tagging[C] // Proceeding of the 9th ACM International Conference on Information Knowledge Management. New York:ACM press,2010: 54 - 63.
- [79] GIRVAN M, NEWMAN M E J. Community structure in social and biological networks[J]. Proceedings of the National Academy of Sciences, 2002, 99 (12): 7821 - 7826.
- [80] LI H Q, XIA F, ZENG D, et al. Exploring Social Annotations with the Application to Web Page Recommendation[J]. Journal of Computer Science Technology, 2009,24(6):1028 - 1034.
- [81] GELFAND A E, SMITH A F M. Sampling-based approaches to calculating marginal densities[J]. Journal of the American Statistical Association, 1990, 85 (410):398 - 409.
- [82] DEMPSTER A P, LAIRD N M, RUBIN D B. Maximum likelihood from incomplete data via the EM algorithm [J]. Journal of the Royal Statistical Society: Series B, 1977,39(1):1 - 38.
- [83] KONSTAN J A, MILLER B N, MALTZ D, et al. GroupLens:applying collaborative filtering to use net news [J]. Communications of ACM,1997,40(3):77 - 87.
- [84] LINDEN G, SMITH B, YORK J. Amazon. com recommendations: Item-to-item collaborative filtering [J]. IEEE Internet Computing,2003,7(1):76 - 80.
- [85] TSO-SUTTER K H L, MARINHO L B, SCHMIDT-THIEME L. Tag-aware recommender systems by fusion of collaborative filtering algorithms[C] // Proceedings of the 2008 ACM Symposium on Applied Computing. New York:ACM press,2008:1995 - 1999.
- [86] LIANG H, XU Y, LI Y. Collaborative filtering recommender systems based on popular tags[C] // Proceedings of the 4th Australasian Document Computing Symposium. 2009:3 - 10.
- [87] PENG J, ZENG D D, ZHAO H, et al. Collaborative filtering in social tagging systems based on joint item-tag recommendations[C] // Proceedings of the 19th ACM International Conference on Information Knowledge Management. New York:ACM press,2010:809 - 818.
- [88] KIM H N, JI A T, HA I, et al. Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation[J]. Electronic Commerce Research and Applications,2010,9(1):73 - 83.
- [89] DURAO F, DOLOG P. Extending a hybrid tag-based recommender system with personalization[C] // Proceedings of the 2010 ACM Symposium on Applied Computing. New York:ACM press,2010:1723 - 1727.
- [90] 陈超,张颖超,缪进.一种基于三部图网络的协同过滤算法[J].南京信息工程大学学报(自然科学版),2010,2(4):337 - 339.
- [91] 吴思竹.社会标注系统中标签推荐方法研究进展[J].图书馆杂志,2010(3):48 - 52.