

基于社会化群体作用的信息聚合服务

胡昌平 胡吉明 邓胜利

摘 要 信息聚合服务在一定程度上提高了用户获取信息的效率和质量,但越来越多且越来越“碎片”化的信息分散了人们的注意力,纯粹的信息聚合服务已不能有效地进行信息的质量控制、无法满足用户的社会化需求。在分析当前社会化信息推荐聚合的基础上,结合社会化的群体作用,构建基于社会化群体作用的信息推荐聚合服务模型,并以“玩聚 SR”为例阐述了其实现的策略。研究表明,社会化推荐聚合服务可以有效地提高信息推荐的准确度,将成为互联网环境下挖掘有价值信息的有效方式。图 2。参考文献 17。

关键词 信息聚合服务 社会化群体作用 推荐聚合

分类号 G250

ABSTRACT Information aggregation service improves the efficiency and quality for user to access information to a certain extent, but there are some bottlenecks for development. By analyzing the role of user's socialization groups, the users' socialization needs would be met more effectively and the services would be further improved. The paper first analyzes the principle and development bottlenecks of information aggregation services, and then discusses the socialization groups and its impact on information aggregation service. Finally, the paper analyzes the current status of social recommendation and aggregation, builds information recommendation and aggregation service model based on the role of socialization groups, and takes "WanJu-SR" as an example to illustrate its implementation strategies. 2 figs. 17 refs.

KEY WORDS Information aggregation service. Role of socialization groups. Recommendation and aggregation.

CLASS NUMBER G250

Web2.0 环境下,越来越多且越来越“碎片”化的信息分散了人们的注意力,人们迫切需要将“信息碎片”和分散的注意力聚合起来^[1],利用信息聚合寻找有用的知识信息,利用聚合的用户关系分享知识,让知识信息在个人和群体间更加自由地流动。信息聚合服务使信息获取和利用变得高效、快捷和准确,成为当前重要的信息传播渠道和主流信息服务模式之一;但其服务方式使用用户的信息行为大多只和自己有关,忽视了用户关系对信息传播的巨大作用^[2],也抑制了用户的沟通、分享和情感需求。用户聚合形成的社会化有机群体所产生的巨大作用不仅使信息通过“群体关系”快速乃至以几何级数传播,加速了知识信息的分享和创造,也满足了用户的情感认知需求。

1 信息聚合服务的原理与发展瓶颈

Web2.0 环境下,零碎且不断创造出来的信息根据用户的兴趣逐渐地被聚合起来,大大减少了用户寻找分散信息所花的精力和时间,提高了信息获取和利用的效率和质量。信息聚合服务正在成为互联网的最新动向,并逐步走向社会化。

1.1 信息聚合服务的原理

目前,基于 RSS/Atom 的信息聚合服务已成为互联网上一种重要的服务模式,广泛应用于网络新闻、博客等网络信息的聚合组织和传播,解决了网络信息资源的有序组织、有效传播问题^[3]。它将用户所关心的 Web 信息内容,通过对标题、链接和摘要等基本项目甚至剥离了布局结构的完整内容进行描述,同时利用 XML 命

名空间扩展描述对象,按照标准的格式生成 XML 内容描述文件,最终达到信息聚合的目的。聚合以后,通过阅读器推送给用户,并随着信息内容的更新而及时更新^[4]。一方面,应用它们可以在站点之间使用相同(或兼容)的规范达到信息共享、内容聚合的效果;另一方面,可以帮助用户把有用的网络信息“拉”(pull)回来^[5]。

信息聚合服务具有以下优势^[6]:首先,根据用户需求有选择地聚合信息,从庞杂的信息源中获取有价值的信息,具有个性化优势;其次,服务器数据库中信息的更新会立即聚合到阅读器中,使得网络信息的发布和共享变得更加方便和高效,极大地提高了信息的时效性和价值;再次,充分体现了控制信息的能力,结合用户参与,网络信息资源的组织变得更加合理;最后,随着信息聚合数量、密度以及层次的递增,信息的价值递增,出现了信息聚合的马太效应,服务的质量得到了质的提升。

2003 年以来,互联网上出现了较多的信息聚合服务商,国外主要有 Twitter、delicious、Google Reader,国内则有叽歪、玩聚 SR、鲜果等,它们通过各种功能让用户参与到服务的流程中,如订阅跟踪服务、“多对一”的信息源格式转化、多个信息源的“合烧”、Tag 标注、评论、共享等,向用户推荐感兴趣或相关的信息。

1.2 信息聚合服务的发展瓶颈

纯粹的信息聚合只是信息的简单收集和拼贴,服务的发展出现了瓶颈,具体有以下两点:

(1)不能有效地进行信息内容的质量控制。之前,信息聚合服务之所以能够帮助用户获取信息,高效、准确且个性化,究其原因信息聚合技术本身并未普及,支持规范格式的内容并不多,相对而言用户通过信息聚合服务获得的内容的质量足够高。但是,当呈爆炸式增长的网络信息都能以信息聚合格式迅速地呈现在阅读器上的时候,信息聚合服务系统却很少对内容和用户进行智能处理,内容推送的先后顺序以及重要程度无从得知,用户所获得的内容的质量不能得到有效控制,因而再次陷入“信息过载”的局面。

(2)用户的社会化需求无法得到有效满足。

目前的信息聚合服务主要是基于对信息共享数据的统计分析来为用户提供服务。通过数据挖掘统计用户的信息共享行为数据,如访问时间、Tag 引用的频繁程度等,对用户喜爱的内容进行排序并基于 Tag 的引用进行类别排序,以此来向用户推荐^[7]。不可否认,这种模式为用户提供了新的体验,对服务质量有很大的改善,但是它对用户一视同仁,对所有的用户进行相同的分析,没有将内容和用户之间的关系考虑进去。因此用户获得的仅仅是信息的内容,而不是知识;通过服务所获得的体验仅限于与自己有关系,很难通过获取信息的过程向其他用户贡献自己的经验,不能找到自己所认同的群体和体现自我价值。目前纯粹的信息聚合技术对信息的获取和利用很难继续改良和优化,不能帮助用户接触所感兴趣领域的更多更好的信息源,进而改善自己的知识结构以及结识一些志同道合的朋友;单纯的信息聚合服务已不能很好地满足用户的社会化需求。

2 社会化群体作用及其对信息聚合服务的影响

当网络逐渐成熟和用户对技术的新鲜感降低后,用户对服务的期望由“技术”逐渐转向“认知”。基于 Web2.0 的互联网,本质上也是社会化的互联网,所要实现的是服务的亲和力、可操作性、用户体验以及可用性,让用户获取信息的同时,建构自己的社会群体关系网络,由此所产生的社会化群体作用对信息聚合服务的拓展具有深远的影响。

2.1 信息聚合服务中的社会化群体作用

物以类聚,人以群分,人们总是希望加入所认同的群体,远离所不认同的群体;作为社会中的一员,其决策必然会受到周遭环境的影响^[8]。网络信息的泛滥让用户基于信息分析进行决策的难度越来越大,越来越受其他用户的影响,越来越依赖于所信任的某个领域“专家”。而且,用户对内容互动性的要求越来越高^[9],在获取所需信息的同时,还存在着强烈地表达观点、与其他用户分享经验以及实现自我价值的需求;越来越重视自己所认同的

领域群体,希望自己的经验得到群体的认同,与群体其他用户持续有效地交流。

在信息聚合服务中,用户的行为具有社会群体化的趋势,不仅希望将所喜好的信息聚合,更希望将志趣相投的用户联系起来,增强沟通和交流,聚合形成群体。一方面可以在用户之间形成相互影响、相互促进的人际关系;另一方面获得学术性、社会性认知和情感上的支持帮助^[10]。具体表现为:众多用户总是簇拥在少数的专业信息来源(其发布者可称为“专家”,在领域内有特殊的专注研究,对群体中用户的决策具有相当大的影响力)周围,以“专家”为中心形成特定的交流群体,如豆瓣网的兴趣小组,用户群组 and 话题背后的群体将会散发出巨大的影响力,对用户知识交流和知识构建具有非常大的帮助。这即为用户行为的社会群体化,其产生的社会化群体作用可以有效地解决“信息过载”和改善服务。

2.2 社会化群体作用对信息聚合服务的影响

随着群体内用户的交流,“专家”的影响力在群体内扩散,社会化的群体作用逐渐凝聚、放大并开始影响服务的质量。首先,“专家”推荐所产生的影响和用户之间的交流,将拓展用户信息来源的渠道,让用户得知更多有益的信息源。其次,通过社会化的群体作用,按照用户的认同与否来进行信息的聚合推荐,而且聚合的信息在群体中被反复接触、吸收、理解和整合后转化为知识,将信息的简单收集和拼贴上升到知识的挖掘与分享。最后,在信息转化为知识的过程中,志趣相投的用户建立联系、彼此关注,增强了用户的群体认同感和凝聚力,激励用户持续分享知识和经验,获得情感上的满足,提升用户对服务的粘度。用户只有在认同的群体中,通过积极的用户关系才能充分地分享信息和体验,运用群体智慧将信息转化为知识^[11]。上述过程通过社会化的群体作用无缝地衔接在一起,相互促进,循环上升,对信息聚合服务的开展具有重要的推进作用。

现在 Web2.0 的核心就是社交,或者说是“用户关系”,纯粹的信息聚合已经逐渐走向消亡。如,Google News 当时被认为极大地提升了

新闻的价值,而现在 Google 已经宣布关闭这个功能;曾经小而精彩的聚合网站 Techmeme 已经远离了科技人士关注的重心,2008 年宣布引入了首位人工编辑,对新闻进行调整,Techmeme 纯聚合模式终结;Digg 则引入了用户的参与因素,利用参与的用户民主选择和推荐新闻,充当更好的、更有效的分享工具。社交网 Facebook 和迷你博客服务 Twitter 出现之后,基于用户关系的信息流获取机制彻底击败了 Digg;Facebook 的用户已突破 2 亿,基于用户关系的社会化群体作用使其信息流的获取早已超过了当年信息聚合的成就;Twitter 中的社会化群体作用更是将信息的获取和传播发挥到了极致^[12]。因此,用户在信息传播和获取活动中越来越承担起一个社会化的媒体角色,纯粹的机器语义信息聚合已经逐渐走向没落,基于社会化的用户群体交往关系的聚合推荐才是未来的方向。

3 基于社会化群体作用的信息推荐聚合服务模型

社会化推荐聚合的产生有效地提高了信息推荐的准确度,越来越成为海量互联网下挖掘有价值信息的有效方式。本文在分析当前社会化信息推荐聚合的基础上,结合社会化的群体作用,构建基于社会化群体作用的信息推荐聚合模型。

3.1 社会化推荐聚合服务模型的不足

社会化推荐/聚合 (Social Recommendation/Aggregation) 是基于社会化媒体分享源的再聚合,从推荐中发现“推荐”,为用户发现、推荐和共享知识,提高了“推荐”的精确度,并实现了内容之间的互通^[13](模型见图 1)。首先,社会化推荐聚合系统追踪并实时扫描各种社会化聚合媒体,如 Google Reader、Twitter、delicious、FriendFeed、鲜果等,将每个 RSS 内容进行处理,提取里面所有的 URL,并将 Tinuurl 等缩短网址解析出来;其次,汇总推荐分享的链接得到一个推荐链接库,在去掉推荐链接库中的不相关内容后,将 URL 聚合;最后,根据 $R(t, n)$ 算法 (R 是与时间 t 和推荐数 n 相关的函数),计算出 URL 的不

同 R 值,根据 R 值将文章排序,得出既新鲜又好看的文章,按照顺序推荐给用户^[14]。如上述国外的 Techmeme,通过 link 追踪发现 tech blogger 讨论热点的平台,在形成固定阅读群体后,就反

向实现了制造热点的能力。国内比较典型的则是“玩聚 SR”,利用智能语义聚合技术帮助用户更准确地发现新知识、结识志趣相投的用户。

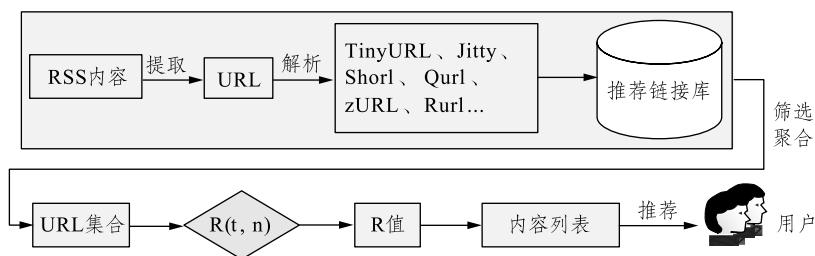


图1 社会化推荐/聚合服务模型

社会化推荐聚合服务可以让用户及时地获得最新的热点、精华内容,但不一定就是用户所喜欢的。社会化推荐聚合的核心还是内容的聚合和推荐,对信息共享背后的用户关联的挖掘不尽理想,没有通过内容的互动建立牢固的用户关系;用户的关联分散,没有形成志趣相投的群体,基于社会化群体关系的分享与推荐作用尚浅,用户无法得到充分满足自己个性化需求的推荐内容。只有运用群体智慧和志趣相投的

用户间的分享和推荐,用户才能获得最符合自己个性化需求的推荐内容。

3.2 基于社会化群体作用的信息推荐聚合服务模型设计

在综合分析社会化群体作用对信息聚合服务的影响和目前信息推荐聚合服务的基础上,构建出基于社会化群体作用的信息推荐聚合服务模型(图2)。

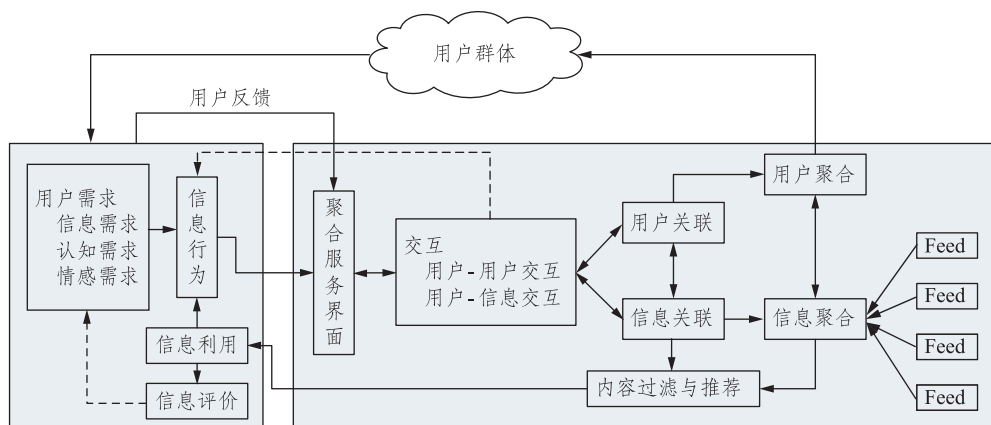


图2 基于社会化群体作用的信息推荐聚合服务模型

在此模型中,信息聚合服务系统基于聚合交互界面,推动实现信息聚合和用户聚合。一方面,系统为用户提供个性化的信息聚合服务;另一方面,用户聚合为群体后,随着用户认同感和

粘性的增强,激励着用户无私地分享信息和贡献自己的经验,用户群就可以源源不断地向系统提交其交流的信息和创新的知识成果。

(1)信息聚合。通过系统提供的聚合交互

界面,用户对服务系统有了基本认识,根据已有的知识结构,应用知识与经验以及感知、思维、判断等向系统提交感兴趣的信息源,获取聚合推荐信息并进行相应的操作,完成信息利用过程。系统在与用户的交互过程中,通过观察用户以及用户的操作,了解用户的反应特征,推导出用户的个体特征;然后在此基础上,为用户组织适应性的资源,利用智能化的资源匹配和用户知识供需匹配机制在恰当的时间、以恰当的方式提供给用户。此外,系统将定期或不定期提供新的刺激,以维持用户的注意力,包括:为不同的用户需求制定不同等级的知识转化策略;提供简单易懂、生动活泼的使用指南和模拟演示;通过信息查询、知识地图、知识导航、信息检索等服务功能,使用户能够顺利地利用系统中各种资源。同时,通过制作友好和简洁的用户界面,建立用户交互的反馈机制^[15]。

(2)用户聚合。基于信息关联以及用户的交流和兴趣将用户聚合为群体,实现社会化的群体交流机制,不仅可以有效地过滤信息,而且可以提高用户的服务体验,满足情感需求。系统通过分析信息交互记录,对用户和信息进行聚类,然后根据聚类得到的用户相关性和信息相关性实现用户之间的协同,让具有相近需求的用户能借鉴彼此的经验,完成信息内容的选择、过滤和推荐,从而激发新的信息需求和行为。提供有利于用户交流的各种环境和情境,营造良好的交流氛围,开展灵活多样的用户交流服务,包括与服务提供者、专家及用户群中其他用户的交流,搭建系统和用户群的桥梁,建立基于用户需求内容和方法的学习和交流方式,帮助用户聚合为用户群。基于用户的使用习惯,提供用户自组织功能和协作工具,让每个用户可以自主建立特定的兴趣小组,或自主选择加入别人创建的小组,根据用户参与情况,依据聚类算法给不同用户推荐其最可能感兴趣的信息和志趣相投的用户。让用户动态地选择、接触和认同群体,集成用户智慧,帮助用户协同分享经验,减轻用户的智力负担,提升服务体验。

该模型将信息聚合上升为用户聚合,强调信息交互与用户交互过程中对用户动态的反

映,注重认知和非认知因素的结合,旨在构建支持对用户自适应的群体交流环境,利用群体中催化和激发的社会化机制,过滤推荐信息,“理解”并在必要的时候“激发”用户,最大化地激励用户在所认同的群体内分享和交流信息。

3.3 基于社会化群体作用的信息推荐聚合服务实现

本文以社会化推荐聚合服务“玩聚 SR”为例,进行基于社会化群体作用的信息推荐聚合服务的改进。

目前,国内社会化推荐聚合的典型代表“玩聚 SR”以简单、丰富、智能、准确的特点吸引了众多用户。它就像一个自动化的 digg 平台,用于追踪中国大陆 IT 业界人士使用的各种社会化聚合媒体分享源,如微博客(twitter、叽歪、饭否)、阅读器分享(Google Reader、鲜果)、网摘收藏(del.icio.us)等,从中聚合用户推荐、分享以及讨论的文章或话题,并伴有用户推荐和分享的评论^[16]。

“玩聚 SR”的服务理念就是基于社会化群体作用,先从某个领域内挑选一些有影响力的意见领袖,再延伸出去,综合分享推荐的次数、推荐者权重和信息源类型等多种因素,从有限的信息源中分析出重点和热点,使推选的内容更加具有“热度”(经过用户的推荐而来,更加反映出用户的喜欢程度,也保证了高质量);并将引入智能语义技术,进化到协同过滤和语义关联的 Social Web Explorer,自动化地为用户发现、推荐和共享知识^[17]。“玩聚 SR”的二次聚合推荐,大大降低了用户搜寻信息的时间成本,也加强了各社会化媒体用户之间的沟通;相比之前的内容聚合的确有一些创新,是一种更加优秀的聚合服务方式。

目前,“玩聚 SR”注重聚合的思想和呈现逻辑,核心还是聚合,推荐依然较“大众化”,用户大多数只能被动地接受,没有很好地考虑到用户行为的群体特性,无法有效保证推荐最多的内容就是用户一定感兴趣的;每个用户难以通过网络上的社会化关系进行信息推荐与被推荐,“大众化”与“个性化”难以平衡。应在继续

拓展信息聚合源的基础上,提高用户参与聚合推荐与分享的程度,聚合用户的网络社会关系,挖掘用户的群体关系和特性;将选择权还给用户,让用户自己选择心目中志同道合的意见领袖和感兴趣的信息源,并分别给予不同的权重;根据“玩聚 SR”的语义聚合算法,构建机器筛选和用户好友推荐相结合的个性化聚合推荐平台,通过社会化的群体作用机制替用户筛选信息,尽可能地减少无用干扰,在“大众化”和“个性化”两者之间找到平衡点;保证既能将最新热点推荐给用户,又能够契合用户的兴趣。

4 结语

信息服务的本质是提供给用户控制信息的能力,而不是可获得信息的总量。信息聚合服务暂时缓解了信息爆炸的危机,但并没有从根本上解决信息过载的问题和有效地满足用户的社会化需求。解决信息过载、改善用户体验的根本途径则是服务的社会化。当技术成熟和用户对网络的新鲜感退却之后,用户的网络行为必然回归现实,进行网络上的社会交往和激发交流表达的深层需求,寻找自己所认同的群体。用户信息行为的社会化必然形成社会化的群体,只有在用户所认同的群体中,用户才能积极地分享知识和经验,才能获得最符合个性化需求的信息。信息聚合表明了用户在信息需求层面依赖于领域内的少数“专家”,走向信息的“中心化”;用户聚合则会将用户聚合为群体,依据群体中少数专家的群体影响力推动用户的决策,走向了用户的“中心化”。信息聚合服务必将从信息聚合走向信息背后用户的聚合,利用社会化群体的作用为用户更好地过滤推荐信息,通过群体的交流机制改善用户的服务体验,提高服务质量。

参考文献:

- [1] 信息琐碎时代 聚合谁的生活[EB/OL]. [2009-10-14]. <http://forum.techweb.com.cn/viewthread.php?tid=295752&page=1&authorid=158645>.
- [2] 胡昌平. 信息资源管理原理[M]. 武汉: 武汉大学出版社, 2008: 235.
- [3] 姜恩波. 基于信息聚合的服务与技术[J]. 现代图书情报技术, 2007(4): 32-34.
- [4] JamesSnell. RSS and Atom [EB/OL]. [2008-05-04]. <http://www.intertwingly.net/wiki/pie/Rss20AndAtom10Compared>.
- [5] 伍玉伟. RSS:网络信息“聚合”利器[J]. 现代情报, 2006(2): 221-222.
- [6] 陈峰,熊励. 基于RSS信息服务联盟的内容聚合技术研究[J]. 计算机技术与发展, 2009(1): 9-12.
- [7] 黄晓斌. 社会书签与网络信息推荐服务[J]. 情报理论与实践, 2006(1): 121-124.
- [8] 鲍尔. 预知社会:群体行为的内在法则[M]. 北京: 当代中国出版社, 2007: 186.
- [9] 邓胜利. 基于用户体验的交互式信息服务[M]. 武汉: 武汉大学出版社, 2008: 10.
- [10] Habib M C. Toward academic library 2.0: Development and application of a library 2.0 methodology. A Master's Paper for the M. S. in L. S degree [D]. University of North Carolina at Chapel Hill, November, 2006: 48.
- [11] 甘永成. 虚拟学习社区中的知识建构和集体智慧研究[D]. 上海: 华东师范大学, 2004: 136.
- [12] 曹增辉. 机器聚合之殇[EB/OL]. [2009-10-14]. <http://www.caozenghui.cn/archives/569.html>.
- [13] Anatomy of a social recommendation engine [EB/OL]. [2009-11-04]. <http://blog.recommenders06.com/wp-content/uploads/2006/09/shur1.pdf>.
- [14] 社会化推荐聚合[EB/OL]. [2009-10-16]. <http://www.kenengba.com/post/648.html>.
- [15] 邓胜利,张敏. 基于用户体验的交互式信息服务模型构建[J]. 中国图书馆学报, 2009(1): 65-70.
- [16] 玩聚 SR [EB/OL]. [2009-10-20]. <http://www.ju690.com/>.
- [17] 玩聚 SR 社会化推荐聚合[EB/OL]. [2009-10-16]. http://blog.csdn.net/zhengyun_ustc/archive/2008/12/21/3575578.aspx.

胡昌平 武汉大学信息资源研究中心教授,博士生导师。通讯地址: 武汉大学信息管理学院。邮编: 430072。

胡吉明 武汉大学信息管理学院博士研究生。通讯地址同上。

邓胜利 武汉大学信息管理学院讲师。通讯地址同上。

(收稿日期: 2009-11-05; 修回日期: 2009-12-10)