

# 大数据时代的人才推荐系统

## *Talent recommendation system in big data era*



洪文兴(1980-),男,厦门大学自动化系副教授,厦门信息产业与信息化研究院执行院长。2010年获得工学博士学位,为国家公派(厦门大学与新加坡南洋理工大学)联合培养博士。2013年先后任厦门大学信息科学与技术学院院长助理、厦门市信息化局软件园服务处副处长、厦门大学发展规划办公室副主任,是中国系统工程学会监事、福建省系统工程学会秘书长、中国计算机学会 YOCSEF 厦门分论坛副主席(2015—2016年)。主要研究方向为数据挖掘、大数据分析、推荐系统、软件和信息服务业产业研究、系统工程。



王宁(1979-),男,厦门华夏学院副教授、高级工程师。2001年起从事信息管理系统及其数学模型分析等领域的研究工作,主要研究方向为数据挖掘及其在人才、教育、商务等领域的应用。



陈毅伟(1990-),男,就职于厦门航空有限公司信息部软件研发中心,主要研究方向为人才推荐系统、数据处理。



周绮凤(1976-),女,博士,厦门大学自动化系副教授。2002年起从事数据挖掘及智能系统方面的研究工作,2014—2015年在美国佛罗里达国际大学访学,主要研究方向为机器学习、数据挖掘及其在可持续发展等领域的应用。



李涛(1975-),男,博士,南京邮电大学计算机学院院长,南京邮电大学大数据研究院院长,美国佛罗里达国际大学正教授。由于在数据挖掘及应用领域成效显著的研究工作,曾多次获得各种荣誉和奖励,其中包括2006年美国国家自然科学基金委颁发的杰出青年教授奖,2010年IBM大规模数据分析创新奖,并于2009年获得佛罗里达国际大学最高学术研究奖。

中图分类号:TP391

文献标识码:A

doi: 10.11959/j.issn.2096-0271.2017024

2017024-1

## 1 人才推荐系统的应用背景

随着Web 2.0技术的发展,越来越多的求职者在互联网上寻找就业机会,企业选择在招聘网站发布职位网罗人才。目前流行的招聘网站按照其应用的技术大致可以划分为5类(如图1所示):传统信息门户类、职业社交类、推荐类、垂直行业类以及人工辅助类。传统信息门户类又可以划分为专业型门户和综合型门户,该类网站以静态信息展示、发布为主,典型的代表是51JOB和58同城,前者专注于全行业全国各地的求职招聘,而后者则作为一个综合的生活分类信息网站,以兼职和短期工作招聘为主。传统信息门户类招聘网站由于缺乏个性化服务,面临着技术转型难题。LinkedIn则是职业社交类招聘网站的典型代表,它为每一个用户构建专属的职业人脉,然后利用这些人脉接触有意向的岗位或者求职候选人。Monster是一个以推荐技术为核心的求职招聘网站,通过为每一

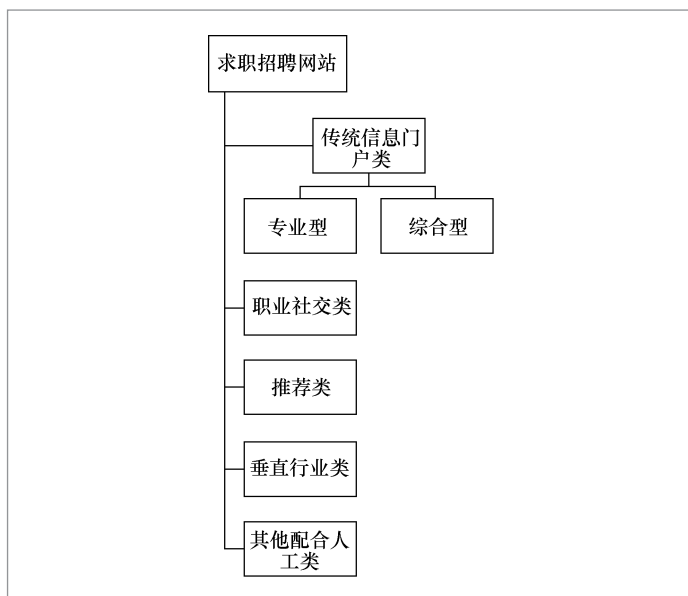


图1 求职招聘网站分类

个求职者和企业招聘人员提供个性化推荐,实现信息对接。与前三者不同,拉勾网定位于专业的互联网招聘平台,专注互联网职位机会,充分利用行业领域特点和知识。此外还有内推网等配合人工运作的信息网站。在上述各类网站中,推荐功能是众人关注的热点<sup>[1,2]</sup>。

## 2 大数据时代的人才招聘问题

各类求职招聘网站的普遍应用为人才招聘的研究提供了数据基础。与此同时,在大数据环境下,新一代人才推荐系统对服务个性化、精准化要求越来越强烈,其研究面临着新的机遇和挑战<sup>[3]</sup>。

从求职者(job seeker)的角度看,更加智能化、个性化的简历系统可以引导他们更快速、更准确地完成一份充分表现自身优势的简历。更重要的是,求职者希望应用系统可以根据求职者自身个性化的特点与偏好,从冗长的职位列表中更加精准地锁定、推荐感兴趣的职位。此外,系统应提供更简便的人性化服务。

从企业人力资源(HR)经理的角度出发,希望应用系统能够从海量的简历中快速筛选出最合适的候选人集合,并提供求职者尽可能详细的信息,以便全方位地了解求职者。与此同时,利用招聘网站的人才评价体系,可以更好地对接企业内部人力资源管理系统<sup>[4]</sup>,为企业制定切合单位发展需求的人才储备、培养以及考核等相关制度提供决策参考。

对于应用系统的经营者而言,有效地协助用户完成对接,提高系统求职与招聘对接的成功率,加强系统人性化服务、增强用户体验感,吸引各类用户、聚合用户是其基本诉求。此外,发布人才年度统计报告、人才预测报告等,从行业、地域、时间

等维度进行劳动力市场供求关系分析,为企业、政府经济规划提供参考,为系统运营赢得声誉,产生更多的社会效益。

经过多年的行业观察和实践分析,笔者发现大数据时代的人才推荐系统需要解决如下关键问题:

- 不同行业领域下个性化简历模板和职位模板的构建方法;
- 基于用户简历、系统内操作行为、用户间交互等多维异构数据的用户特征提取与用户建模方法;
- 基于求职用户职位偏好与企业用户招聘偏好相互匹配的双向推荐算法;
- 个性化、人性化的信息收集策略,数据统计、发布的多元数据可视化展示问题。

### 3 智能人才推荐系统实例: iHR<sup>+</sup>

#### 3.1 iHR<sup>+</sup>简介

自2010年以来,厦门大学与厦门人才网联合开展人才推荐研究。厦门人才网服务厦门本地及其周边城市的7万家企业,每天在线有效职位储量超过23万个,人才库拥有170多万份个人简历,日均访问量达70万页面浏览量(page view, PV)。在厦门人才网多年数据积累的基础上,研究分为3期完成<sup>[5,6]</sup>,具体内容如图2所示。

项目通过对用户求职意向的分析,结合个性化表单收集结构化、非结构化数据,实现个人简历自动创建的电子简历系统。结合用户的系统操作行为等特征完成用户建模。在此基础上,对企业职位需求进行分析,设计推荐指数、推荐热度<sup>[7]</sup>等指标,采用信息融合模型初步实现双向混合推荐算法<sup>[8]</sup>。此外,对用户搜索各类信息需求的分析、整合,实现了智能搜索以及地图搜索、信息订阅与推送、界面自适应

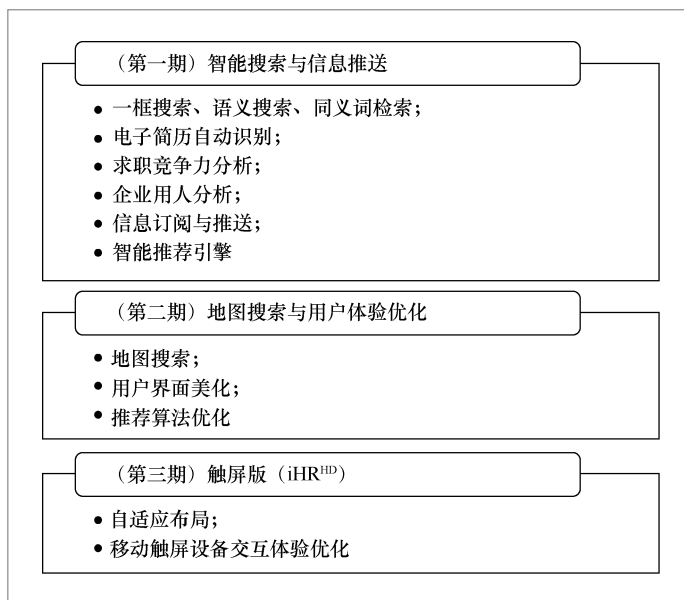


图2 厦门人才网人才推荐项目

(iHR<sup>HD</sup>)与美化等人性化服务,方便用户使用。

在理解用户需求的基础上,笔者提出若干关键问题,进行算法研究、论证评估与改进,形成了针对人才推荐的智慧解决方案——iHR<sup>+</sup>。图3为方案的整体架构。

#### 3.2 iHR<sup>+</sup>人才推荐系统关键问题的解决

在人才推荐系统中,求职者与职位提供方之间的关系是一种二元关系,即只存在二者之间的交互,而不存在求职者之间或者职位提供方之间的交互。这种强二元关系也是互惠推荐的一大特性。为此,采用二部图的方式建立人才推荐模型,如图4所示。

##### (1) 基于多维特征空间的用户建模方法

用户特征是推荐系统的数据基础,传递着求职用户的信息和兴趣偏好<sup>[9]</sup>。笔者根据用户的特征来源,将用户特征分为用户自然人信息的基本特征、个性化简历中提取的提取特征、用户在系统中各类操作的操作特征以及用户间互动的标签特征四

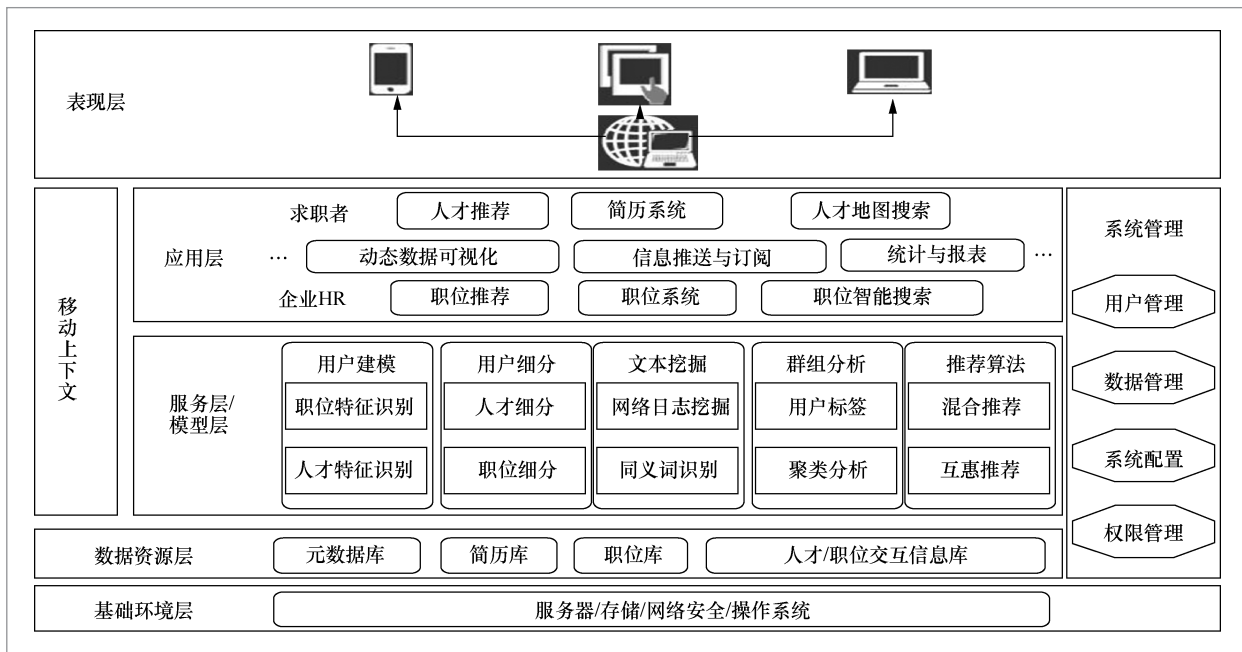


图3 iHR+ 总体架构

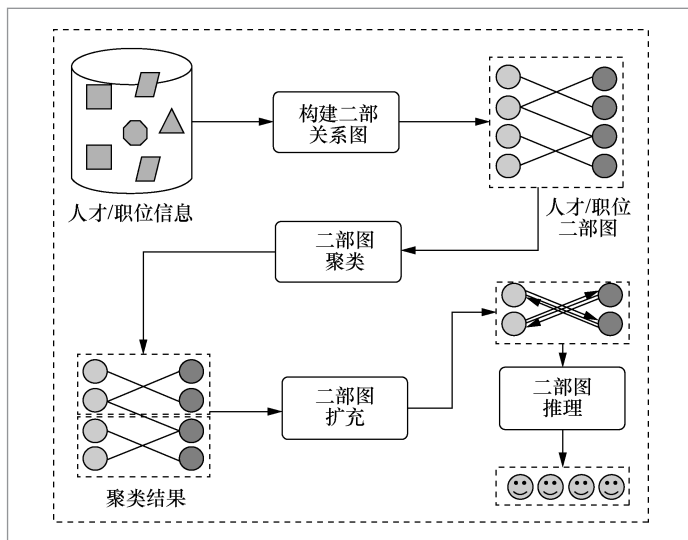


图4 二部图人才推荐模型

大类。依据每项特征的特点进行特征表示与存储, 创建了多维用户特征空间。

(2) 面向求职招聘领域的双向推荐模式  
2010年Pizzato L等人<sup>[10]</sup>首次提出双向推荐, 双向推荐是个性化推荐系统中的一个独特的分支, 该推荐算法适用于推荐双方用户之间有偏好的系统, 如在线约会

系统<sup>[11]</sup>、人才推荐系统等, 它考虑了推荐双方用户之间的相互偏好关系, 使用户双方达到双赢的效果, 也提高了推荐系统的准确性。在人才推荐系统中, Malinowski J等人<sup>[12]</sup>根据双方用户的偏好信息分别构建面向个人用户和企业用户的推荐系统, 综合考虑双方的推荐结果。Yu H等人<sup>[13]</sup>提出一种融合了双方用户的显性与隐性偏好信息的相似度计算方法, 进一步探讨双方用户偏好程度之间彼此的关联度对于推荐结果的影响, 即人才推荐中的双向推荐问题。

(3) 基于上下文信息的推荐模式

除了求职者与企业单位双方的特征信息, 潜在的上下文信息也是推荐系统可以考虑的重要信息来源。例如, 分析某求职者的工作经历, 如果发现其频繁地更换工作且每份工作的时间不长, 可推测他缺乏自身定位、能力评价或职位所需品质等可能性, 这就是潜在的上下文信息。因此, 在系统推荐中, 需以恰当的形式进行判定, 并调整向企业单位推荐该求职者的推荐度。

#### (4) 基于领域专家知识的推荐模式

在人才推荐领域中,常见的专家知识有专业的分类、职位部门的分类以及技能的分类等。专家知识具有很高的可信度与参考价值,构建一个完整的专家知识体系有利于为企业单位推荐相关专业的人才,或者为求职者推荐相关职位甚至更高一级的职位等<sup>[14,15]</sup>。例如,某求职者熟练掌握C++,通过专家知识体系中计算机方面技能分类的分析<sup>[16]</sup>,可知Java、C等语言与C++的相关性很高,则可以向该求职者推荐Java、C方面的职位。

#### (5) 基于动态时间效应的推荐模式

人才市场职位的招聘具有时效性。招聘职位在一定的时间段内有效,超过该时间段后则失效,因此在职位推荐中应考虑职位的有效期,防止将已失效的职位推荐给求职者,同理,对于求职者也应考虑活性<sup>[17]</sup>,将注册并使用推荐系统已顺利找到当前满意的工作,从此不再登录推荐系统的不活跃求职者推荐给企业单位是没有实际意义的。综上,职位的时间有效性、求职者的时间动态性(活跃度)都是推荐系统应该考虑的因素,都会影响推荐结果的质量和用户满意度。

### 3.3 人才推荐研究应用的成功拓展

基于在人才推荐领域的工作成果,笔者拓展了问题研究领域,针对专家型人才不主动提供信息且隐性信息明显的特点,设计并实现了专家信息搜寻、专家推荐等功能平台,取名“科技驿栈”。起始页如图5所示。

“科技驿栈”是一个面向中小企业科技创新的服务平台,旨在为广大产业技术工作者提供一站式创新服务,为产业研发人员之间、科研院所研究人员之间、产—学—研之间的大规模协作提供协同创新平台。它收集分布在全国各高校、科研机构



图5 大规模专家协作创新平台(科技驿栈)

的各领域专家资料信息,包括专家论文、专利、活动、企业新闻等文献以及专家活动、项目、成果等,通过建立搜索引擎,开展数据挖掘、系统推荐,提供领域专家及相关信息的智能搜索与在线推荐,方便企业寻找所需领域的专家。

### 参考文献:

- [1] 洪文兴. 基于互联网的决策理论与方法的若干问题研究[D]. 厦门: 厦门大学, 2010.  
HONG W X. Research on web-based decision-making theory and method[D]. Xiamen: Xiamen University, 2010.
- [2] 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.
- [3] 曾春, 邢春晓, 周立柱. 个性化服务技术综述[J]. 软件学报, 2002(10): 1952-1961.  
ZENG C, XING C X, ZHOU L Z. A survey of personalization technology[J]. Journal of Software, 2002(10): 1952-1961.
- [4] GONZALEZ T, SANTOS P, OROZCO F, et al. Adaptive employee profile classification for resource planning tool [C]// 2012 Annual SRII Global Conference, July 24-27, 2012, San Jose, CA, USA. New Jersey: IEEE Press, 2012: 544-553.

- [5] HONG W X, LI L, LI T, et al. iHR: an online recruiting system for xiamen talent service center[C]//The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 11-14, 2013, Chicago, IL, USA. New York: ACM Press, 2013: 1177-1185.
- [6] 李涛. 数据挖掘的应用与实践——大数据时代的案例分析[M]. 厦门: 厦门大学出版社, 2013: 123-162.
- LI T. Data mining: where theory meets practice[M]. Xiamen: Xiamen University Press, 2013: 123-162.
- [7] MELVILLE P, MOONEY R J, NAGARAJAN R. Content-boosted collaborative filtering for improved recommendations [C]//The 18th National Conference on Artificial Intelligence, July 28 - August 1, 2002, Edmonton, Canada. New York: ACM Press, 2002: 187-192.
- [8] LI L, LI T. MEET: a generalized framework for reciprocal recommender systems [C]//The 21st ACM International Conference on Information and Knowledge Management (CIKM 2012), October 29 - November 2, 2012, Maui, Hawaii, USA. New York: ACM Press, 2012: 35-44.
- [9] GAUCH S, SPERETTA M, CHANDRAMOULI A, et al. User profiles for personalized information access [M]. Berlin: Springer Berlin Heidelberg, 2007: 54-89.
- [10] PIZZATO L, REJ T, CHUNG T, et al. Reciprocal recommenders[C]//The 8th Workshop on Intelligent Techniques for Web Personalization and Recommender Systems(UMAP' 2010), June 20-24, 2010, Hawaii, USA. [S.l.:s.n.], 2010: 1-12.
- [11] PIZZATO L, REJ T, CHUNG T, et al. Recon: a reciprocal recommender for online dating[C]//The 4th ACM Conference on Recommender Systems, September 26-30, 2010, Barcelona, Spain. New York: ACM Press, 2010: 207-214.
- [12] MALINOWSKI J, KEIM T, WENDT O, et al. Matching people and jobs: a bilateral recommendation approach[C]//The 39th Annual Hawaii International Conference on System Sciences, January 4-7, 2006, Kauai, Hawaii, USA. New Jersey: IEEE Press, 2006: 137c.
- [13] YU H, LIU C, ZHANG F. Reciprocal recommendation algorithm for the filed of recruitment[J]. Journal of Information & Computational Science, 2011, 8(16): 4061-4068.
- [14] BURKE R. Knowledge-based recommender systems[J]. Encyclopedia of Library and Information Science, 2000, 69(Supplement): 32.
- [15] VELASQUEZ J D, PALADE V. Building a knowledge base for implementing a web-based computerized recommendation system[J]. International Journal on Artificial Intelligence Tools, 2007, 16(5): 793-828.
- [16] CHEN C F, CHEN L F. Data mining to improve personnel selection and enhance human capital: a case study in high-technology industry [J]. Expert Systems with Applications, 2008, 34(1): 280-290.
- [17] LI L, ZHENG L, LI T. LOGO: a long-short user interest integration in personalized news recommendation [C]//The 5th ACM Conference on Recommender Systems (RecSys 2011), October 23 - 27, 2011, Chicago, Illinois, USA. New York: ACM Press, 2011: 317-320. □