

基于 LBSN 的商业选址推荐系统的研究与实现

屈弘扬 於志文 田 苗 郭 斌

(西北工业大学计算机学院 西安 710072)

摘 要 随着智能移动设备的发展和普及,空间定位技术不断成熟,基于位置的社交网络(Location-based Social Network, LBSN)得到了广泛应用。大量用户在 LBSN 签到,以及针对签到进行的评论不仅记录了用户的时空行为轨迹,也为研究用户行为模式和特征偏好提供了巨大的机会。提出一种基于 LBSN 签到数据的商业店铺选址推荐系统,首先分析用户在 LBSN 上的签到时间、签到地点、签到商铺类型 3 个方面的特征;然后提出 4 个影响商铺选址的因素:多样性、竞争性、相关性和客流性;最后实现商业选址推荐系统,并根据选址因素生成最优候选。并以此为基础进行相关实验来验证推荐结果,结果符合相关预期。

关键词 LBSN,行为轨迹数据,商业选址

中图法分类号 TP393

文献标识码 A

DOI 10.11896/j.issn.1002-137X.2015.9.007

Research and Implementation of Commercial Site Recommendation System Based on LBSN

QU Hong-yang YU Zhi-wen TIAN Miao GUO Bin

(School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract With the development and popularization of smart mobile devices, spatial positioning technology continues to develop, and based on this, location-based social network is widely used. The majority of users check in LBSN, and comment on check-in activity, which not only record the spatial-temporal behavior track, but also provide great opportunities for us to study user behavior patterns and characteristics of preference. This paper proposed a commercial site recommendation system based on LBSN. Firstly, it analysed the characteristics about the check-in time, the check-in location and the category of check-in retail in LBSN. Then it proposed four kinds of factors that affect retail location: diversity, competitive, relevance, passenger flow. Finally the system was implemented that can provide the best candidate based on various factors. The paper used those as the basis for experiments to verify the recommendation result. The results comply with the relevant expectations.

Keywords LBSN, Behavior track data, Commercial site

1 引言

近年来,随着智能移动设备的广泛普及,各种移动社交网络快速发展,逐渐融入人们的日常中。移动社交网络不仅连接了物理和虚拟两个社会空间,融合人们线上、线下体验,形成混合社交空间^[1],而且为社会感知计算^[2]提供了更加方便快捷的途径。

基于位置的社交网络是一种基于用户位置,以智能通信网络为媒介、智能终端为主要载体的社交网络。LBSN 作为移动社交网络快速发展的代表,除了具有传统社交网络中人与人的联系外,还可以通过签到跟踪和共享人的位置信息。作为用户社交内容的重要组成部分,位置可以在某种程度上反映个体用户的兴趣与行为^[3]。以国外的 FourSquare、Google+ 及国内的街旁网、人人网为代表的 LBSN,每天都会

产生数以亿计的数据。通过对这些含有时间、地点、评论的用户签到数据的分析与研究,不仅能够得到用户的特征偏好,而且能够进一步挖掘出地点的特征。

商铺选址问题在经济、管理领域具有很大的研究价值,但它又是一项复杂的研究,具有难度大、不确定性高的特点。一些研究人员开始尝试结合社交网络,探索商铺的选址问题。Karamshuk 等人^[4]借助用户在 FourSquare 的签到数据,基于地理位置特征和用户移动特征预测商铺的最佳选址。Jensen^[5]提出在空间上对商铺的内部和外部协同因素研究最佳零售区域。Yu^[6]等人使用对比线性回归、协同过滤、关联规则 3 种方法,预测最大市场。而这些研究或没有使用 LBSN 数据或没有深入挖掘用户签到行为特征,来提出商铺选址的解决方案。并且这些研究多从地点角度考虑选址问题,并未细粒度地考量不同类型商铺选址的特点和差异性。与之不同的

到稿日期:2014-05-17 返修日期:2014-08-22 本文受国家重点基础研究发展计划(973 计划)(2012CB316400),国家自然科学基金(61332005,61373119,61222209,61103063)资助。

屈弘扬(1991—),男,硕士生,主要研究方向为普适计算,E-mail:lonicera0036@163.com;於志文(1977—),男,博士,教授,博士生导师,CCF 杰出会员,主要研究方向为普适计算、移动互联网和智能信息技术,E-mail:zhiwenyu@nwpu.edu.cn;田 苗(1992—),女,硕士生,主要研究方向为普适计算、移动社交网络;郭 斌(1980—),男,博士,教授,博士生导师,CCF 高级会员,主要研究方向为普适计算、社会智能、移动社会网络。

是,我们将更加注重对 LBSN 数据的分析与研究,深入挖掘用户签到特征。

本文通过收集和分析 LBSN 用户的签到数据,提取用户在时间、地点、商铺类型 3 个方面的签到特征,进一步对影响商铺选址的因素进行分析,提出一种商铺选址解决方案,最后实现商业选址推荐系统。

2 数据采集与分析

2.1 数据采集

利用街旁 API^[7] 采集了上海地区从 2011 年 5 月到 2012 年 10 月这 17 个月中 54054 位用户共计 964972 条签到数据。由于街旁数据的时间较为久远,部分签到情况和商铺情况可能会发生变化,我们借助百度 LBS 开放平台提供的 API^[8] 采集了上海地区共计 138876 条信点(Point of Interest, POI) 和大众点评网的 15196 条商铺信息。虽然这些数据之间存在时间差异,但通过相互对比发现,商铺的类型和数量的变化相对于这些基数是很小的。这两部分数据作为街旁数据的补充,可以有效地提高数据的可用性。

采集到的 3 种数据虽然相互补充,但是主要作用不同。街旁数据主要用于研究用户的签到情况和签到特征,计算用户客流情况。收集的百度 LBS 开放平台的 POI 信息主要用于辅助研究商铺的特征和商铺数量。大众点评网提供用户对商铺的评分,用于对系统的验证。

下面对街旁数据进行分析,以反映用户签到规律。街旁主要数据信息如表 1 所列。

表 1 街旁主要数据信息

名称	举例
用户 ID	19
记录 ID	483
签到时间	2012-01-17 08:51:47
签到位置(经纬度)	121.4665198102 31.269701600722
签到内容	“摩卡,半价半价!!!”
该位置的总用户数	207
该位置的总签到数	298
商铺地址	平型关路 108 弄 6 号
商铺名称	少山咖啡
商铺类型	咖啡馆
类型 ID	0102

2.2 数据分析

2.2.1 签到时间

采集的街旁数据是 17 个月的数据,导致部分月份是两年的签到记录,并且数据量较大,存在噪声数据,因此抽取 2102 条签到数据,并记录签到时间。这些签到数据涵盖了完整一年 12 个月的签到记录。依照一天 24 小时绘制签到时间频率统计与累计分布,如图 1 所示。

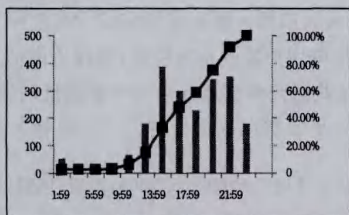


图 1 整体签到时间频率统计与累计分布

从图 1 可以发现,用户一般在上午 8 点左右开始进行签到活动,这个时间点正是上班、上学的起始时间。签到数在上

午 10 点左右开始攀升,一天之中 12 点到 14 点达到峰值,这个时间段正是午饭时刻。同上午相比,下午用户签到数都保持了一个较高的数量级,在 18 点到 22 点,用户的签到数一直保持在一个较高层次,用户多在此时间段进行休息、娱乐等活动。因此整体来看,用户的签到活动多在中午到晚上进行,并且在饭点时刻的签到数量较多。

2.2.2 签到地点

在收集的数据中用户的签到数据有签到位置的经纬度。由于经纬度数据不方便使用,为了便于观察与分析,根据这些经纬度,使用百度地图绘制了用户签到的热力图。热力图既反映了签到地点情况,又反映了签到数量。根据签到的经纬度和签到数量,对上海徐汇区进行了绘制,如图 2 所示。



图 2 徐汇区热力

图中颜色越深的区域表明此处签到数量越多,通过热力图可以发现用户签到具有分散性和区域性。用户签到数量较多的地点多在该地区的商业圈,如衡山路商圈、徐家汇商圈。这些商圈的人流量较大,在此签到的用户概率也就随之增加。

2.2.3 签到商铺类型

街旁数据中提供了用户签到地点信息的商铺类型。在研究用户的签到地点特征之后,借助上述特征进而研究用户签到的商铺类型分布规律。

通过对用户的签到数据的分类与统计,可以发现当签到记录个数超过 5000 时,用户的签到地点类型种类增加缓慢,特别是商铺类型种类几乎不再增加。因此统计了 5000 个签到记录,共收集到 84 种类型,其中商铺类型有 72 种。将这 72 种商铺类型再进行归类,共得到中餐、西餐、休闲、娱乐、零售 5 种类型。这 5000 个签到记录中商铺有 2014 家,各个类型的数量分布如图 3 所示。

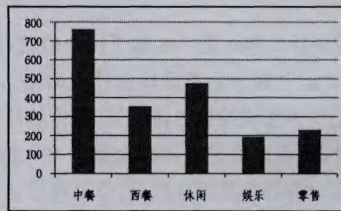


图 3 商铺数量统计

3 商铺选址因素分析

根据上述用户签到数据和签到商铺的分析与研究发现,商铺的种类之间、商铺的数量之间、商铺与商铺之间、客流之间都存在着一定关系。为研究这些关系,定义 4 种影响商铺选址的因素:多样性、竞争性、相关性和客流性。

3.1 多样性

在一个区域内,商铺类型的多样性反映出该区域的繁荣程度^[4]。通常来说,多样性越高,该区域越繁华,越能吸引人群。一个区域的多样不仅表现为商铺类型种类,还表现为各种类型的商铺的数量。多样性所表示的商铺种类和数量之间的关系可以根据香农-威纳指数确定,在区域 L 内的多样性因素 P_v 定义如式(1)所示。

$$P_v = -\sum_{\kappa} \left(\frac{N_{\kappa}}{N} * \ln \frac{N_{\kappa}}{N} \right) \quad (1)$$

其中, κ 为商铺的类型, N_{κ} 为该区域商铺类型为 κ 的商铺数量, N 为所有商铺总数量。根据定义, P_v 越大,表示多样性越大,区域内越繁荣;反之, P_v 越小,表示多样性越小,区域内越不繁荣。

3.2 竞争性

在一个区域内,同种类型的商铺在吸引客流上形成一种竞争关系,导致客流的分散^[4]。竞争性具有优胜劣汰的特征,但是对商铺选址具有负面影响。同种类型的商铺数量越多,那么这些同种类型商铺之间的竞争就越激烈。竞争性所表现出来的就是商铺数量之间的关系。商铺数量是竞争的主要因素,竞争性主要根据商铺数量来进行度量。因此在区域 L 内的竞争性因素 P_c 定义如式(2)所示。

$$P_c = e^{-\frac{N_{\tau}}{N}} \quad (2)$$

其中, N_{τ} 为该区域商铺类型为 τ 的商铺数量, N 为所有商铺总数量。根据定义, P_c 越大,表示竞争性小;反之, P_c 越小,表示竞争性大。根据指数函数特征,当 $\frac{N_{\tau}}{N}$ 很大时, P_c 很小,其变化也趋于稳定。

3.3 相关性

在一个区域内,不同类型的商铺之间能够互相影响,这种影响直观体现为提供客流,例如餐饮类商铺增加了休闲、娱乐类商铺的客流。相关性主要表示商铺之间的关系。相关性不仅与商铺种类、数量有关,还与客流也有关。相关性涉及因素较多,如果将客流分散研究,那么研究商铺之间的相关性的复杂程度就降低了。因此在区域 L 内的相关性因素 P_r 定义如式(3)所示。

$$P_r = \lg \left(\sum_{\kappa} \frac{\lg \frac{M_{\kappa}}{N_{\kappa}}}{\lg \frac{M_{\tau}}{N_{\tau}}} - 1 \right) \quad (3)$$

其中, κ 为商铺的类型, N_{κ} 为该区域商铺类型为 κ 的商铺数量, M_{κ} 为该区域商铺类型为 κ 的签到数量, N_{τ} 为该区域商铺类型为 τ 的商铺数量, M_{τ} 为该区域商铺类型为 τ 的签到数量。根据定义, P_r 越大,多样性越高,商铺之间联系越紧密;反之, P_r 越小,多样性越低,商铺之间联系越不紧密。

3.4 客流性

在一个区域内,客流会成为影响商铺选址的另一个主要因素。客流量越大,用户在此消费的几率就大。客流性表示客流情况,是客流关系的一个度量。由于人群具有流动性,传统的商铺选址研究中,很难准确确定该区域的客流情况,而借助 LBSN 中的签到服务可以直观、准确地获取到客流情况。用户签到地点具有分散性和区域性,我们通过对比签到数来确定客流情况。因此在区域 L 内客流性因素 P_f 定义如式(4)所示。

$$P_f = \frac{M_l}{M_{2r}} \quad (4)$$

其中, M_l 为在以点 A 为圆心、 r 为半径的 L 区域内总的签到数, M_{2r} 为在以点 A 为圆心、 $2r$ 为半径的区域内总的签到数。根据定义, P_f 越大,客流性越高;反之, P_f 越小,客流性越低。

4 系统设计

4.1 系统架构

商铺选址推荐系统主要包含两主模块:用户界面模块、推荐引擎模块,如图4所示。

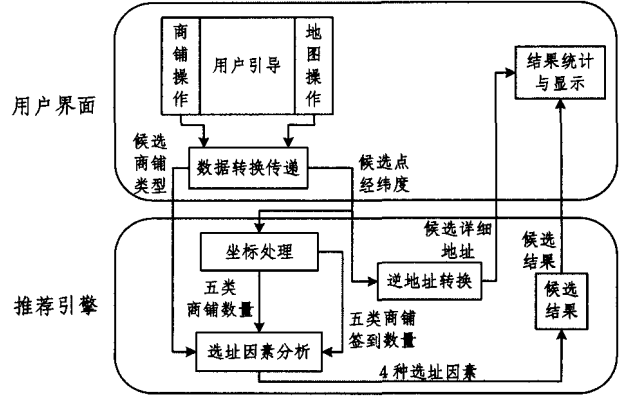


图4 商业选址推荐系统架构

用户界面模块不仅方便用户操作,也能显示推荐结果,主要包含用户引导模块、地图操作模块、商铺操作模块、数据转换传递模块、结果统计与显示模块5个子模块。

其中,用户引导模块用于指导用户操作;地图操作模块和商铺操作模块分别用于用户选址候选点和候选商铺类型;数据转换传递模块是将抽象的候选数据进行量化处理成机器可识别的数据;结果统计与显示模块用于对比结果和显示结果。

推荐引擎模块的主要功能是根据相关特征数据和候选数据集,对它们进行处理,进而返回候选推荐结果。其具体包含坐标处理模块、选址因素分析模块、候选结果模块、逆地址解析模块4个子模块。其中坐标处理模块用于获取某点形成的区域内各种商铺类型的数量和签到数量;选址因素分析模块用于分析计算4种选址因素;候选结果模块用于计算所有候选结果,并传递给用户界面;逆地址解析模块用于将经纬度转化为具体地址。

4.2 地址推荐

4.2.1 问题定义

商铺选址问题定义为:给定地点和候选商铺类型集,选择最优商铺类型。该问题可以描述为:对于地点 A ,选择商铺类型 k 。因此首先以 A 为中心,得到半径为 r 的区域 L ,对该区域内的商铺类型、数量和签到数量进行分析,通过多样性、竞争性、相关性、客流性4个因素来设计商铺选址方案,即给定地点 A 和候选商铺类型集 S ,为地点 A 选择最优商铺类型 $s(s \in S)$ 。

4.2.2 因素融合

第3节提出了4种影响商铺选址的因素:多样性、竞争性、相关性、客流性,但是每种因素产生的影响效果是不同的。我们首先确定各因素对商铺选址的影响权重,然后通过线性相加的方式将各个因素进行融合。具体如式(5)所示。

$$P = \alpha * P_v + \beta * P_c + \gamma * P_r + \delta * P_f \quad (5)$$

其中, $\alpha, \beta, \gamma, \delta$ 分别为多样性、竞争性、相关性、客流性4种因素的权重, P_v, P_c, P_r, P_f 分别表示为多样性、竞争性、相关性、客流性4种因素。 P 表示融合了4种因素之后,得到的商

铺推荐值。即在区域 L 内,某种商铺类型 s 的 P 值大于其他商铺类型值,则该区域推荐的商铺类型即为 s 。

4.2.3 推荐流程

首先,用户根据操作引导,在地图操作模块和商铺操作模块选择候选地点和候选商铺类型。数据转换传递模块将用户操作结果解析以获得量化结果,并传递给推荐引擎。

然后,推荐引擎借助坐标处理模块,获得 5 类商铺类型的数量和签到数;接着使用选址因素分析模块根据相关数据,分析计算 4 种选址因素;然后通过候选结果模块分析计算候选 P 值作为候选结果传递给结果统计与显示模块。使用逆地址解析模块将抽象的经纬度坐标转化为具体的地址,并将其传递给结果统计与显示模块。

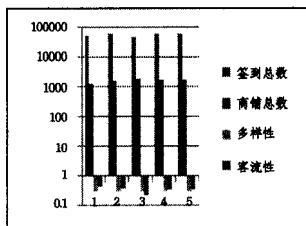
最后,用户界面的结果统计与显示模块将候选结果进行对比分析,选取 P 值最大的作为最优商铺,并显示相关统计结果。

5 实验结果

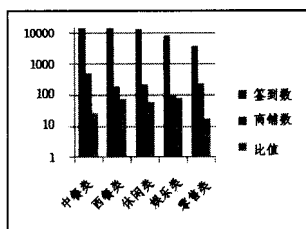
5.1 商铺选址因素分析结果

用户主要的签到商铺类型为中餐、西餐、休闲娱乐、零售这 5 种类型。因此商铺选址因素中 κ 同样也为这 5 种类型。

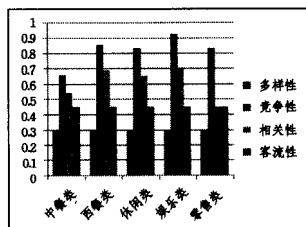
用户的签到数据多是不连贯的、碎片化的和即时性的,这会使确定区域 L 的半径 r 产生较大误差。但是根据用户主要商业圈范围内活动这一特征,选取相邻两个商业圈之间的距离作为区域 L 的半径 r 。为此我们选取了上海市区 15 组相邻的商业圈,计算相邻商业圈的距离,结果显示两个相邻商业圈的平均距离为 1.6km。因此区域 L 的半径 r 设定为 1.6km。



(a) 5 个区域中的签到总数、商铺总数、多样性因素和客流性因素



(b) 一个区域的商铺数量、签到数量、签到数量与商铺数量比



(c) 一个区域中各个商铺类型的多样性、竞争性、相关性和客流性

图 5

在确定 κ 与 r 这两个参数之后,对商铺选址因素进行分析。选取点 A ,分别统计点 A 所在区域内用户在各类商铺的签到数量、各类商铺分布数量、签到数量与商铺数量比以及多样性因素、竞争性因素、相关性因素和客流性因素。在上海市

区选取了 5 个点,统计以上信息。相关性因素如图 5(a)所示。在同一个区域内的统计结果如图 5(b)、图 5(c)所示。

图 5(a) 示出商铺数量随着多样性因素值增加而增加。从图 5(b)、(c) 中可以看出,单一类型的商铺数量越多,竞争性因素值越低,则竞争性越大;相关性因素不仅与商铺数量有关,还与签到数量有关。当区域半径增加一倍,区域面积增加 3 倍,半径为 3.2km 的区域面积足以体现人群的移动特性,通过签到比值,可以较好地体现客流性因素。因此多样性因素、竞争性因素、相关性因素、客流性因素的选取符合实际情况,是可以作为选址影响因素的。

5.2 选址推荐实验结果

5.2.1 实验设计

为了验证选址推荐实验结果,需要首先确定 4 种商铺选址因素在融合结果中所占的权重。然后根据融合因素对部分地点进行实验,计算这些地点的各个商铺类型的 P 值,对比选址结果,获得最优选址结果。最后验证推荐结果的可信度。

5.2.2 确定 4 种商铺选址因素的权重

将用户在大众点评上商铺的评分数据作为各个选址影响因素的训练数据。在上海市市区,共计搜集 200 个商铺评分,商铺评分的数据基本处于 0.70~0.90 之间。因此训练数据时将评分数据控制在不超过 5% 的波动范围内。在训练时,设定每个权重值每次以 0.05 为增幅,权重最大值为 1。通过训练,对各个影响因素取均值,则结果中各个影响因素权重值分别取 $\alpha=0.93, \beta=0.27, \gamma=0.31, \delta=0.32$ 。

5.2.3 推荐结果

根据以上各个选址影响比重,进行选址推荐实验。对 10 个随机地点的各个类型的商铺选址推荐结果进行实验,并对比结果,如图 6 所示。

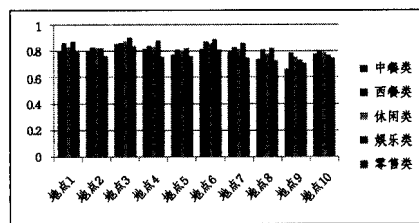


图 6 10 个地点各个商铺类型推荐结果

从图中发现,各个商铺类型的推荐结果因地点不同而不同,但最优推荐结果基本上集中于西餐类和娱乐类。系统的最优推荐结果分布不均衡,这 5 种商铺类型形成 3 个层次,第一层是西餐类和娱乐类两种,第二层次为休闲类,第三层次为中餐和零售类。

造成此结果的原因主要归为以下几个方面:

- 1) 中餐类的商铺较多,导致竞争性因素值较低;
- 2) 西餐和娱乐类商铺数量较少,但是签到数量较多,使竞争性因素和相关性因素值较高;
- 3) 零售类虽然商铺数量较少,但是签到数较少,使相关性因素值较低;
- 4) 由于仅在二维平面上进行选址设计,各个商铺的实际位置也会存在误差。

通过进一步对比发现,计算 3 个层次所得的推荐结果依次降低,层次之内的差距很小,不超过 10%;层次之间的差距比层次之内的差距大,但是差距仍然不超过 20%。

(下转第 44 页)

- [2] 韩家炜,坎伯. 数据挖掘:概念与技术[M]. 北京:机械工业出版社,2001:130-160
Han Jia-wei, Kamber M. Data Mining, Concepts and Techniques [M]. Beijing: China Machine Press, 2001:130-160
- [3] Yen S J, Wang C K, Ouyang L Y. A Search Space Reduced Algorithm for Mining Frequent Patterns[J]. Journal of Information Science and Engineering, 2012, 28(1):177-191
- [4] Park J S, Chen M S, Yu P S. An effective hash-based algorithm for mining association rules[J]. ACM SIGMOD Record, 1995, 24(2):175-186
- [5] Singh J, Ram H, Sodhi D J S. Improving Efficiency of Apriori Algorithm Using Transaction Reduction[J]. International Journal of Scientific and Research Publications, 2013, 3(1):1-4
- [6] Savasere A, Omiecinski E R, Navathe S B. An efficient algorithm for mining association rules in large databases[J]. ACM SIGMOD Record, 2000, 29(2):1-12
- [7] Toivonen H. Sampling large databases for association rules [C]//VLDB. 1996, 96:134-145
- [8] Brin S, Motwani R, Ullman J D, et al. Dynamic itemset counting and implication rules for market basket data[J]. ACM SIGMOD Record. ACM, 1997, 26(2):255-264
- [9] 孟军,王蓬,张静,等. 基于项集依赖的最小关联规则挖掘[J]. 计算机科学, 2013, 40(1):183-186, 217
Meng Jun, Wang Peng, Zhang Jing, et al. Minimal Association Rules Mining Based on Itemset Dependency[J]. Computer Science, 2013, 40(1):183-186, 217

(上接第36页)

5.2.4 推荐结果可信度的验证

基于以上推荐实验结果和原因,依旧根据大众点评网的数据来进行推荐结果的对比验证。由于大众点评网中的数据不一定是最优的结果,因此对于推荐结果的验证,采用可信度来衡量。由于商铺选址的复杂性较高,准确性较低,我们仅考虑了4个会对最终结果造成误差的影响因素,因此经过训练之后,当系统的整体可信度达到50%时,我们认为推荐结果可信。根据推荐过程中可能遇到的情况,做如下情况定义。

情况一 如果某地点的原有商铺类型为A,系统最优推荐商铺类型也为A,实际评分 P_1 与最优推荐评分 P_2 的差距小于10%,即如式(6)所示,则认为推荐结果可信。

$$\frac{|P_1 - P_2|}{\min\{P_1, P_2\}} < 10\% \quad (6)$$

情况二 令某地点的原有商铺类型为A,系统最优推荐商铺类型为B,实际评分为 P_1 ,最优推荐评分为 P_2 ,以及系统计算A类型的评分为 P_3 ,若任意两个评分差距都小于10%,即如式(7)所示,则认为推荐结果可信。

$$\begin{cases} \frac{|P_1 - P_2|}{\min\{P_1, P_2\}} < 10\% \\ \frac{|P_1 - P_3|}{\min\{P_1, P_3\}} < 10\% \\ \frac{|P_2 - P_3|}{\min\{P_2, P_3\}} < 10\% \end{cases} \quad (7)$$

根据以上定义,随机选取50个地点,统计验证结果,如表2所列。

表2 推荐结果验证

可信结果数量		不可信结果数量		可信度	
情况一	情况二	情况一	情况二	情况一	情况二
9	23	10	8	47%	74%
合计	32	18		64%	

根据表2的数据,满足情况一的只有17个地点,而这17个地点的最优推荐结果是西餐类或者娱乐类,但是其可信度只有47%;而情况二的可信度要比情况一的高,这也反映了最优推荐结果分布不均衡,并且各个推荐结果差距不是很大。将两种情况综合起来,可信度达到了64%。因此根据多样

性、竞争性、相关性、客流性4种商铺选址因素而实现的商铺选址推荐系统的推荐结果的可信度是符合设计期望的。

结束语 本文根据用户在LBSN上的签到数据,提出商铺选址的影响因素,并以此为基础实现了商铺选址推荐系统。

下一步,我们将更细粒度地挖掘用户签到特征和选址影响因素,包括签到用户的行为、情感特征以及用户的签到内容、评论等方面。

参考文献

- [1] 於志文,周兴社,郭斌. 移动社交网络中的感知计算模型、平台与实践[J]. 中国计算机学会通讯, 2012, 8(5):15-20
Yu Zhi-wen, Zhou Xing-she, Guo Bin. Aware computing model, platform and practice of Mobile social networks[J]. Communications of China Computer Federation, 2012, 8(5):15-20
- [2] 於志文,於志勇,周兴社. 社会感知计算:概念、问题及其研究进展[J]. 计算机学报, 2012, 35(1):16-26
Yu Zhi-wen, Yu Zhi-yong, Zhou Xing-she. Socially aware computing[J]. Chinese Journal of Computers, 2012, 35(1):16-26
- [3] Zheng Y. Location-based social networks: Users[M]//Computing with Spatial Trajectories. Springer, New York, 2011:243-276
- [4] Karamshuk D, Noulas A, Scellato S, et al. Geo-spotting: mining online location-based services for optimal retail store placement [C]//Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2013:793-801
- [5] Jensen P. Analyzing the localization of retail stores with complex systems tools[M]//Advances in Intelligent Data Analysis VIII. Springer Berlin Heidelberg, 2009:10-20
- [6] Yu Z, Zhang D, Yang D. Where is the Largest Market: Ranking Areas by Popularity from Location Based Social Networks[C]//Ubiquitous Intelligence and Computing, 2013 IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC). IEEE, 2013:157-162
- [7] <http://dev.jiebang.com>
- [8] <http://developer.baidu.com/map>