

基于用户情境的POI个性化推荐模型

邓鹏¹ 李霖¹ 陈功¹ 李游¹

(1 武汉大学资源与环境科学学院, 武汉市珞喻路129号, 430079)

Context-Based Personalized POI Recommendation Model

DENG Peng¹ LI Lin¹ CHEN Gong¹ LI You¹

(1 School of Resource and Environmental Sciences, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

摘要:提出了一种基于用户情境的POI个性化推荐模型,结合模糊数学中的隶属度概念,从空间、时间及属性3个方面分析了用户情境与POI之间的关联度,并利用关联度对推荐结果进行排名。实验表明,推荐结果具有较好的精度,时间在较大程度上影响POI的推荐结果。

关键词:个性化推荐;情境;POI;隶属度;LBS

中图法分类号: P208

文献标识码: A

Abstract: This paper proposes a context-based personalized POI recommendation model, by using membership degree of fuzzy mathematics to analyze the correlation between user context and POI in three aspects: space, time and other properties, and the results of recommendation are ranked by their correlations. Experiments show that the proposed method has good accuracy, and time can have a great influence on the results of recommendation.

Key words: personalized recommendation; context; POI; membership degree; LBS

随着移动定位技术和互联网技术的发展,移动终端设备的普及,地图逐渐从技术产品向公众服务过渡,产生了众多新兴概念,如智慧地图^[1]、全息位置地图^[2]等。物联网、云计算等技术的兴起所引爆的第三代信息革命也被视为现代地图学发展的新契机^[1]。在众多研究方向中,以用户位置为中心,如何为用户提供全方位的地理信息服务成为当前的研究热点^[3]。以位置为参考,基于各种定位系统、传感网、互联网、通信网等泛在网络,实时动态地获取位置的坐标、属性、关系与移动特征等多源异构信息,并实现这些相关信息的准确融合,最终构建语义一致、时空地理关联统一的地图服务平台,是地理信息服务智能化、大众化、个性化的重要核心技术。结合用户

情境,建立自适应的位置地图,为用户实时、动态地提供满足自身需求的全方位位置服务。

首先,本文从空间、时间及属性3个方面,利用隶属度函数,将用户位置、时间以及用户偏好相同的兴趣点(point of interest, POI)的位置、营业时间及其他自身属性进行关联度分析,并按与用户情境的关联程度对结果POI进行排名,建立POI个性化推荐模型;然后,以基于个性化推荐模型为用户推送合适的用餐地点为案例,重点研究用户请求时间对于POI推荐结果的影响;最后,设计一个调查问卷网站,收集大众就餐偏好数据,分析并验证模型的合理性与有效性。

1 基于用户情境的POI个性化推荐模型

目前,基于位置的各类应用服务可以根据用户当前位置为移动用户推送其附近的POI信息^[4-6]。Yu^[7]提出了一种个性化基于位置的服务(location-based service, LBS)推荐系统的框架,结合位置、时间及用户偏好对用户旅游出行方案进行推荐。Bader^[8]利用多准则决策方法,将当前位置、时间、汽车油量等作为评分指标,研究驾驶途中进行POI推荐的方法。Yuan等^[9]将协同过滤算法用于社交网络中,分析用户决策行为受时间影响的情形,例如,某个POI在某个特定时间段签到的人数较多,其在这个时间段被推荐给用户的概率也会较高。这些研究另外考虑了时间情境对于POI选取的影响,Yu和Yuan都将时间作为用户需求的一种反映,但属于情境建模的范畴;Bader计算绕道至POI所增加的额外时间,但并没有将POI自身的时间属性结合起来。本文认为,POI的营业时间作为与用户情境相关的一个重要因

基金项目: 国家863计划资助项目(2013AA12A202);中央高校基本科研业务费专项资金资助项目(2012205020211)。

素,在很大程度上影响用户对于POI的决策,如不同的餐厅营业时间不同,用户在不同的时间被推送的餐厅信息可能会不一样。结合用户位置,用户可以查询当前位置一定范围内的POI进行就近选择;通过当前时间,用户可以排除当用户到达POI时已经停止营业的POI,减少因信息不足导致的成本增加;利用用户偏好信息,地图可以对用户进行个性化的推荐,满足不同用户口味、价位上的需求。

为了计算用户情境与POI之间的关联度,本文引入隶属度函数。隶属度属于模糊评价函数里的概念,模糊综合评价是对受多种因素影响的事物做出全面评价的一种十分有效的多因素决策方法^[10],其特点是评价结果不是绝对的肯定或否定,而是以一个模糊集合来表示。隶属度函数具体解释为:若对论域(研究的范围) U 中的任一元素 x ,都有一个数 $A(x) \in \{0, 1\}$ 与之对应,则称 A 为 U 上的模糊集, $A(x)$ 称为 x 对 A 的隶属度。当 x 在 U 中变动时, $A(x)$ 就是一个函数,称为 A 的隶属度函数。用取值于区间 $[0, 1]$ 的隶属度函数 $A(x)$ 表征 x 属于 A 的程度高低。隶属度 $A(x)$ 越接近于1,表示 x 属于 A 的程度越高; $A(x)$ 越接近于0,表示 x 属于 A 的程度越低。

在本文中,用户综合多种属性对POI进行决策,由于不同属性的优先程度不一样,用户可以对它们设置不同的权重,将计算出来的隶属度进行加权平均,得到该POI最后的得分,即关联度,这个得分影响POI的排名。

用户情境和POI之间共同的基准与属性结构是进行关联度分析的基础,根据属性项性质和类型的不同,本文定义了空间关联、时间关联和属性关联3种关联方式。①利用用户当前的位置和POI的位置,实现空间关联;②利用当前的时间和POI的营业时间,实现时间关联;③利用用户的偏好和POI的其他自身属性,实现属性关联。POI个性化推荐模型处理流程如图1所示,情境模型分析出用户的情境,推荐模型根据用户情境计算POI与用户情境的关联度,将POI按照关联度降序排序后推荐给用户。

1.1 空间关联

利用GPS和其他定位技术,用户可以获取自身所处位置的坐标,通过计算用户到候选POI的距离,计算出用户选取该POI的隶属度。为了提高运算效率,本文首先通过设定缓冲半径,排除指定阈值距离外的POI,缓冲半径可由用户自己设定。用户 U 的缓冲区定义如下^[11]:

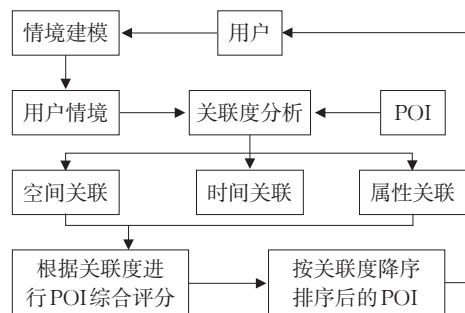


图1 POI个性化推荐模型处理流程图

$$\{X:D(U,P_i) \leq r_t\} \quad (1)$$

式中, P_i 代表第 i 个POI; $D(U,P_i)$ 代表用户 U 到第 i 个POI的距离,本文的空间距离为欧几里得距离; r_t 表示缓冲半径。筛选出缓冲半径内的POI后,计算空间隶属度,公式如下:

$$M_s = \begin{cases} \frac{r_t - D(U, P_i)}{r_t - D_{\min}}, & D_{\min} \leq D(U, P_i) < r_t \\ 0, & D(U, P_i) \geq r_t \end{cases} \quad (2)$$

式中, D_{\min} 表示到用户最近的POI的距离。POI到用户的距离越近,用户选取该POI的隶属度就越大。

1.2 时间关联

用户可以获取当前时间,不同的POI营业时间也不同,用户如果在到达POI位置时,POI已经停止营业了,则相当于用户做了失败的决策。用户如果在POI开始营业之前一段时间内到达POI,只要时间不是很长,用户通常会考虑等待到POI开始营业。因此,考虑到上述情况,本文将时间隶属度公式定义如下:

$$M_t = \begin{cases} 1, & T_o - t_b \leq T_a < T_c \\ 0, & T_a \geq T_c \cup T_a < T_o - t_b \end{cases} \quad (3)$$

式中, T_a 表示用户到达POI的时间点; T_o 表示POI的开始营业时间; T_c 表示POI的关门时间; t_b 表示POI开始营业前用户可以等待的时间长度。各参数示意图如图2所示。需要注意的是, T_a 是由地图导航功能预测出来的,由当前时间加上用户消耗在路上的时间得到,它综合了道路等级、长度、拥堵情况等各方面因素^[12]。

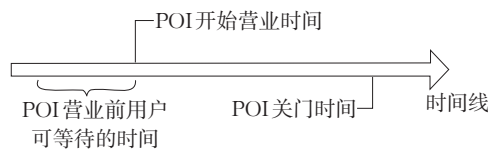


图2 时间关联参数示意图

1.3 属性关联

相对空间关联和时间关联来说,属性关联较为

复杂,不同类型的POI相应的属性也不一样。餐厅有经营菜系类别、平均消费价格等,酒店有星级、房型、房间价格等。因此,对于不同类型的属性,需定义不同的关联方法。本文将属性类型分为两类,分别为数值型(如平均消费价格)和文字枚举型(如餐厅经营菜系类别、酒店星级、房型等)。针对不同的属性类型,定义不同的隶属度函数。

1)数值型属性是可以直接量化的属性,隶属度函数定义如下:

$$M_N = \begin{cases} \frac{U_{\max} - P_{ij}}{U_{\max} - U_l}, & U_l \leq P_{ij} < U_{\max} \\ 0, & P_{ij} \leq U_{\min} \cup P_{ij} \geq U_{\max} \\ \frac{P_{ij} - U_{\min}}{U_l - U_{\min}}, & U_{\min} < P_{ij} < U_l \end{cases} \quad (4)$$

式中, P_{ij} 表示第 i 个POI的第 j 个属性项的属性值; U_l 表示对应的属性项用户所期望的理想值; U_{\min} 表示对应的属性项用户可接受的最小值; U_{\max} 表示对应的属性项用户可接受的最大值。

2)文字枚举型属性不能直接量化,属于离散变量,而且通常会存在多个用户偏好的值,如用户喜欢川菜,但偶尔也会去吃西餐;用户在入住酒店时通常会选择三星级酒店,有时也会选择经济型酒店。我们认为,如果一个用户访问某个属性类型的POI次数较多,则用户对该类型的POI的兴趣较高。因此,本文将隶属度函数定义如下:

$$M_T = \begin{cases} \frac{N_j}{N}, & P_{ij} \in U_j \\ 0, & P_{ij} \notin U_j \end{cases} \quad (5)$$

式中, P_{ij} 表示第 i 个POI的第 j 个属性项的属性值; U_j 代表用户对应属性项所偏好的属性值集合; N_j 表示用户访问该属性类型值POI的次数; N 表示用户访问该POI所属大类POI的次数总和。例如,用户去餐馆共签到20次,其中,川菜馆占了15次,则川菜馆的属性项的隶属度为0.75。

POI属性是影响用户决策行为、进行个性化推荐的关键因素。对于餐厅而言,口味是用户首要考虑的因素,其次是价格;对于酒店而言,用户可能会首先考虑酒店的档次、星级。因此,在进行推荐时,首先筛选出大致符合用户自身偏好的文字枚举型属性(如餐厅口味、酒店档次等),再对数值型属性(如价格)的隶属度进行计算,不同属性的权重可由用户进行设置,以满足不同用户的偏好。

计算出每种属性项的隶属度后,通过加权平均,计算出POI与用户情境的关联度,公式如下:

$$R_i = \sum_{j=1}^n \omega_j \times M_{ij} \quad (6)$$

式中, M_{ij} 表示第 i 个POI第 j 个属性项计算出来的隶属度; ω_j 表示用户对于该属性项的权重, $\sum_{j=1}^n \omega_j = 1$; n 表示属性项的个数。由于所有隶属度函数的值域都是 $[0, 1]$,计算出的关联度值域也为 $[0, 1]$ 。通过关联度对POI进行排序,从而得到与用户情境最为相关的POI,进而推送给用户。另外,如何对属性项的权重进行分配,以获得较好的推荐结果,将在下文实验部分进行分析和讨论。

2 案例分析

为了更清晰地对本文提出的模型进行说明,并对模型的实用性进行验证,本文以用户利用地图推送用餐地点为例进行实验。系统通过情境模型分析用户存在用餐需求,并根据用户所处情境为其推送满足其喜好要求的餐厅。通过比较不同时间、不同权重分配组合情况下系统为用户推荐POI的结果精度,研究用户请求时间对于属性权重分配的影响,说明地图对于不同属性的权重分配具有良好的自适应性,其中,用户请求时间在较大程度上影响了POI的推荐结果。

2.1 案例介绍

用户情境信息包括用户位置、当前时间和用户偏好信息。用户位置如(114°21'36" E, 30°31'32" N),当前时间信息如2014-08-30 T20:50:32。偏好信息如表1所示,包含了偏好菜系、可接受的最低和最高消费价格、期望消费价格。餐厅的相关信息示例如表2所示,包含了餐厅的名称、坐标、营业时间段、经营菜系、人均消费等相关属性^[13]。

表1 用户对餐饮的偏好

| 偏爱菜系 | 最低消费/元 | 期望消费/元 | 最高消费/元 |
|------|--------|--------|--------|
| 西餐 | 30 | 45 | 60 |
| 川菜 | 20 | 40 | 55 |

表2 餐厅相关属性信息

| 名称 | 坐标 | 营业时间段 | 经营菜系 | 人均消费/元 |
|-----------------|------------------------------|-------------------|------|--------|
| 秀玉红茶坊 (乐天城店) | 114°21'17" E, 30°31'33" N | 10:30:00~21:00:00 | 西餐 | 51 |
| 必胜客 (新群光店) | 114°21'19" E, 30°31'33" N | 10:30:00~21:30:00 | 西餐 | 53 |
| 锦庐川菜 | 114°21'19" E, 30°31'33" N | 10:00:00~21:30:00 | 川菜 | 63 |
| 成都往事 艺术餐厅 | 114°21'15" E, 30°31'30" N | 11:00:00~22:00:00 | 川菜 | 44 |

为了验证用户请求时间对于属性权重分配的影响,进而影响POI推荐结果,本文按照用户请求时间的不同,分别对武汉市广埠屯(114°21'36" E, 30°31'32" N)附近2 km半径范围以内的川菜馆(39家)和西餐厅(36家)的推荐进行了实验。通过设定基本筛选条件,如表3所示,筛选出满足该条件的 N 个餐厅。然后,使用本文提出的模型对广埠屯2 km范围内所有的餐厅按照不同的属性权重组合进行评分排序,分析不同属性权重组合下这 N 个餐厅落在前 N 中的个数 M 。最后,按个数进行降序排序,取 M 最大值的权重组合集(即最优权重组合集),对不同属性所占的权重进行求和汇总,分析请求时间对于属性权重分配的影响。

表3 川菜和西餐厅基本筛选条件

| 用户请求时间 | 空间距离 | 用户价位偏好/元 | | 时间 |
|----------|-----------|----------|-------|---------|
| | | 川菜 | 西餐 | |
| 19:00:00 | 2 km半径范围内 | 20~55 | 30~60 | 餐厅关门前到达 |
| 20:00:00 | | 20~55 | 30~60 | |
| 20:50:00 | | 20~55 | 30~60 | |
| 21:12:00 | | 20~35 | 30~60 | |
| 21:30:00 | | 20~35 | 30~60 | |

2.2 结果分析

通过对表3中5种不同请求时间下不同权重组合计算得到的POI推荐结果进行分析,本文分别统计出川菜馆和西餐厅推荐时最优权重组合集权值汇总表,如表4和表5所示。

表4 川菜馆推荐最优权重组合权值汇总表

| 请求时间 | 最优权重组合集权值汇总 | | | 最优权重组合集对应 M 值/个 | 满足基本筛选条件的POI个数 N /个 |
|----------|-------------|------|------|-------------------|-----------------------|
| | 空间 | 时间 | 价格 | | |
| 19:00:00 | 3.3 | 12.3 | 13.4 | 35 | 35 |
| 20:00:00 | 3.8 | 12.3 | 13.9 | 35 | 35 |
| 20:50:00 | 1.7 | 8.4 | 6.9 | 33 | 33 |
| 21:12:00 | 0.7 | 7.2 | 3.1 | 20 | 20 |
| 21:30:00 | 6.8 | 17.7 | 6.5 | 11 | 11 |

表5 西餐厅推荐最优权重组合权值汇总表

| 请求时间 | 最优权重组合集权值汇总 | | | 最优权重组合集对应 M 值/个 | 满足基本筛选条件的POI个数 N /个 |
|----------|-------------|-----|------|-------------------|-----------------------|
| | 空间 | 时间 | 价格 | | |
| 19:00:00 | 4.0 | 9.5 | 16.5 | 25 | 27 |
| 20:00:00 | 4.0 | 9.5 | 16.5 | 25 | 27 |
| 20:50:00 | 3.6 | 9.5 | 15.9 | 24 | 26 |
| 21:12:00 | 0.7 | 6.2 | 3.1 | 17 | 18 |
| 21:30:00 | 1.3 | 8.0 | 3.7 | 16 | 17 |

可以看出,不同请求时间下,最优权重组合集对应的 M 值与满足基本筛选条件的POI个数 N 大致

相同,说明通过这样的权重分配,系统推荐的POI较多地落在了筛选集合中,推荐结果具有较高的精度。同时,属性权重的分配组合影响推荐结果的精度,不同情境下最优权重分配组合也不同,当用户请求时间离大多数POI关门时间较长时(就餐厅而言,多数餐厅在21:00:00~22:00:00停止营业),用户多从价位、口味或自身偏好角度选择就餐地点,此时属性隶属度权重较高,空间和时间隶属度权重较低;而当用户请求时间较晚时(接近大多数POI关门时间),用户多从时间角度进行考虑,此时时间隶属度权重较高。因此,我们根据当天时间早晚程度对不同属性进行权重分配时,若时间较早,则赋予时间隶属度较小权重,属性隶属度较大权重;若时间较晚,则为时间隶属度赋予较大的权重,从而为用户推荐更加符合当前时间情境的POI。

以川菜馆为例,对最优权重组合进行归一化处理,分别计算19:00:00、20:50:00、21:12:00三个请求时间川菜馆的推荐结果,在地图上显示出来,如图3所示。可以看出,用户请求时间不同,餐厅的推荐结果也随之不同。

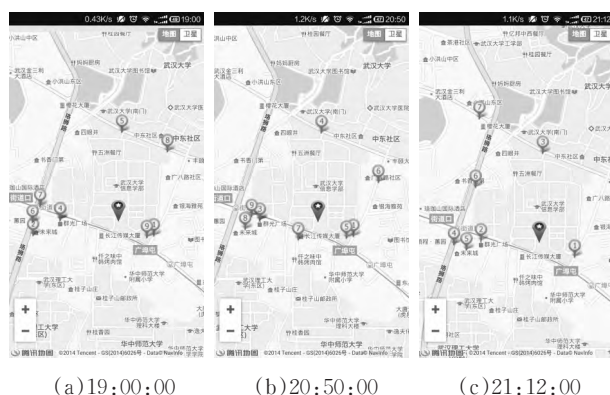


图3 不同请求时间下的川菜馆推荐结果图

3 调查问卷分析

为了进一步验证模型的合理性,针对本文的研究内容,本文设计了一个在线调查问卷网站,对大众的就餐偏好信息进行了收集,用户根据自身偏好,在不同空间(广埠屯武汉大学南门、光谷步行街入口)和时间(19:00:00、21:10:00)情境下选择若干自己倾向消费的餐厅。网站共回收37份问卷,74份餐厅选择结果,如表6所示。餐厅推荐成功的情境数统计情况如表7所示,将用户选择的结果与本文模型推荐的结果进行对比,如果用户选择的餐厅落在模型推荐餐厅前 N 个中,则推荐成功。分别统计前1、前3和前5推荐成功的情境个数,并计算各自在对应情境

总份数中所占的比例,得到餐厅推荐成功的精度统计情况,如表8所示。可以看出,广埠屯武汉大学南门前3和前5推荐成功的总精度分别为73.53%和82.35%;光谷步行街入口前3和前5推荐成功的总精度分别达到了82.50%和87.50%。总的来看,模型推荐结果具有较高的精度,说明提出的模型较为合理。

表6 不同情境下餐厅选择结果统计

| 请求时间 | 用户位置/个 | | 总计/个 |
|----------|-----------|---------|------|
| | 广埠屯武汉大学南门 | 光谷步行街入口 | |
| 19:00:00 | 17 | 20 | 37 |
| 21:10:00 | 17 | 20 | 37 |
| 总计 | 34 | 40 | 74 |

表7 餐厅推荐成功的情境统计结果

| 请求时间 | 用户位置/个 | | | | | |
|----------|-----------|----|----|---------|----|----|
| | 广埠屯武汉大学南门 | | | 光谷步行街入口 | | |
| | 前1 | 前3 | 前5 | 前1 | 前3 | 前5 |
| 19:00:00 | 9 | 13 | 14 | 10 | 17 | 18 |
| 21:10:00 | 8 | 12 | 14 | 13 | 16 | 17 |
| 总计 | 17 | 25 | 28 | 23 | 33 | 35 |

表8 餐厅推荐成功的精度统计结果

| 请求时间 | 用户位置/% | | | | | |
|----------|-----------|-------|-------|---------|-------|-------|
| | 广埠屯武汉大学南门 | | | 光谷步行街入口 | | |
| | 前1 | 前3 | 前5 | 前1 | 前3 | 前5 |
| 19:00:00 | 52.94 | 76.47 | 82.35 | 50.00 | 85.00 | 90.00 |
| 21:10:00 | 47.06 | 70.59 | 82.35 | 65.00 | 80.00 | 85.00 |
| 总计 | 50.00 | 73.53 | 82.35 | 57.50 | 82.50 | 87.50 |

4 结束语

部分学者已对利用情境模型进行个性化信息推荐服务进行了研究,但对于如何将情境融入LBS中进行POI的关联,还缺少足够的关注。本文基于用户情境,结合隶属度函数,同时从空间、时间、属性3个方面研究用户情境与POI的关联方法,建立POI个性化推荐模型,尤其关注了用户的请求时间对于POI推荐结果的影响。但本文的研究还存在很多不足,如属性权重仍需要用户根据自身需要进行设定。深入研究用户偏好的获取方法,挖掘出更为准确的用户偏好信息,为用户提供更加智能化的推荐服务,这也是作者下一步需要研究的内容。

参考文献

- [1] 刘锐,谢涛,孙世友,等. “智慧地图”体系构建研究[J]. 地理信息世界,2013,20(2):24-29
- [2] 周成虎,朱欣焰,王蒙,等. 全息位置地图研究[J]. 地理科学进展,2011,30(11):1 331-1 335
- [3] 齐晓飞,王光霞,薛志伟,等. 位置地图情境分类分级与切换研究[J]. 地理信息世界,2013,20(6):13-18
- [4] Barnes S J. Location-Based Services: The State of the Art[J]. E-Service Journal,2003,2(3):59-70
- [5] Mennecke B E, Strader T J. Mobile Commerce: Technology, Theory and Applications [M]. Hershey, Pennsylvania: IGI Global,2003
- [6] Sadoun B, Al-Bayari O. Location Based Services Using Geographical Information Systems [J]. Computer Communications,2007,30(16):3 154-3 160
- [7] Yu C C, Chang H P. Personalized Location-Based Recommendation Services for Tour Planning in Mobile Tourism Applications[C]. 10th International Conference, EC-Web 2009, Linz, Austria, 2009
- [8] Bader R, Neufeld E, Woerndl W, et al. Context-Aware POI Recommendations in an Automotive Scenario Using Multi-Criteria Decision Making Methods [C]. 2011 Workshop on Context-Awareness in Retrieval and Recommendation, New York, USA, 2011
- [9] Yuan Quan, Cong Gao, Ma Zongyang, et al. Time-aware Point-of-Interest Recommendation[C]. The 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, New York, USA, 2013
- [10] 陈水利,李敬功,王向公. 模糊集理论及其应用[M]. 北京:科学出版社,2005
- [11] 郭仁忠. 空间分析[M]. 2版. 北京:高等教育出版社,2001
- [12] 祝诗蓓,程琳. 基于最短路径的等时缓冲区分析及其应用[J]. 交通运输工程与信息学报,2011,9(1):107-113
- [13] 大众点评网. 大众点评开发者平台[EB/OL]. <http://developer.dianping.com/>, 2014

收稿日期:2014-09-12。

第一作者简介:邓鹏,硕士生,研究方向为3S集成及地图可视化技术。

E-mail: victordeng@whu.edu.cn