

基于本体的旅游资源二次推荐方法研究

徐守坤, 马 慰, 马正华

(常州大学 信息科学与工程学院, 江苏 常州 213164)

摘 要: 为降低用户寻找旅游资源时的负担,更好地处理用户的需求和偏好,从用户角度出发,提出了一种基于本体的二次推荐方法。该方法引入本体来描述旅游资源,实现了用户的直接需求及偏好兴趣信息与推荐中过滤条件的关联,用户可对推荐结果进行评价。理论分析和实验表明,该方法可行,推荐结果能够覆盖绝大部分用户的选择,具有较好的效果。

关键词: 本体; 个性化推荐; 用户偏好

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2012)11-4180-05

doi:10.3969/j.issn.1001-3695.2012.11.046

Research on double recommendation method for tourism resources based on ontology

XU Shou-kun, MA Wei, MA Zheng-hua

(School of Information Science & Engineering, Changzhou University, Changzhou Jiangsu 213164, China)

Abstract: In order to reduce the burden on users of seeking tourism resources, handle the needs and preferences of user better, this paper presented an ontology-based double recommendation method from user's perspective. It adopted ontology to describe tourism resources, achieved the goal of associating users' direct needs and their preferences or interests with filter conditions in recommendation. Moreover, it permitted evaluations. Theoretical analysis and experiments show that the proposed approach is feasible, the results of recommendation involve most of user's choices, and the effect is good.

Key words: ontology; personalized recommendation; user preference

目前,Web上信息膨胀使得信息使用效率不理想。研究发现旅行者的潜在需求和旅游资源的一些特性有很强相关性,在这样的情况下,推荐服务正越来越多地涌现。推荐系统是一种应用程序,它能够电子商务网站所开发利用,向用户推荐产品或服务,并提供必要的信息以方便用户的决策过程^[1]。主流的推荐方法包括以下几种^[2]:a)最常用的协同过滤(collaborative filtering),这种方式注重于用户对被推荐内容的行为;b)基于内容的过滤(content-based filtering),这是一种分类系统的具体类型,允许广泛的有效信息整合进入系统,这种方式关注被推荐内容的本质;c)基于知识的过滤(knowledge-based filtering),这些系统使用与用户和产品有关的知识来推测可能的建议,此类方式提供了高度的灵活性,适合大多数任务和应用,非常注重于内容或社会语义;d)将上述的方式相结合以提供推荐服务的混合过滤方法(hybrid-filtering)。推荐服务使得用户由被动的信息浏览者转变成为主动参与者^[3]。

许多学者在旅游业的推荐系统中做出过贡献,其中大多数应用了人工智能技术^[4~6]。Niemann等人^[7]研究了如何使用语义Web技术提高酒店搜索效率。LA_DMS项目^[8]通过结合语义Web技术和P2P模式提供了基于语义的旅游目的地信息。García-Crespo等人^[9]提出了能为旅客规划线路的富含语义的推荐平台,之后扩展成了目的地管理组织(destination management organizations)^[10]。Huang等人^[6]整合了贝叶斯网

络和语义,在Internet上对旅游景点提供了个性化的推荐。这些学者的工作都在试图解决个性化推荐中用户偏好与被推荐内容之间的联系问题。根据Huang等人^[6]的研究,在旅游产业中开发一个个性化推荐系统将面临两个挑战:a)是在线的异构旅游信息的整合;b)根据旅行者的偏好进行和景点的语义匹配。可以发现,在推荐服务中如何更好地处理用户的需求与偏好是当前一个比较关键的问题。许康华等人^[11]建立了一个基于用户偏好模型的标签推荐系统;García-Crespo等人^[12]运用模糊逻辑的方法找到了用户偏好与酒店特性之间的联系,这些研究着重于解决基于用户偏好的个性化匹配。考虑到用户可能参与到推荐的过程中提出一些直接的需求,这种需求和个性化匹配是独立的,本文提出了一种二次推荐方法,在推荐过程中,通过使用词汇表和语义技术分别处理用户的直接需求与个性化匹配,充分考虑用户的需求,发掘用户偏好中隐含概念,并允许用户进行反馈。

1 概念描述与建模

1.1 旅游资源建模

旅游资源是推荐的内容,以酒店作为旅游资源,引入语义网的关键技术——本体^[13]对旅游资源进行描述。本体是描述概念及概念间关系的一种基本手段,它提供了良好的数据整合

收稿日期:2012-04-20;修回日期:2012-05-21

作者简介:徐守坤(1972-),男,吉林蛟河人,教授,博士,主要研究方向为数据库与信息系统、普适计算理论与应用(jpxsk@163.com);马慰(1987-),男,江苏南京人,硕士研究生,主要研究方向为管理信息系统、本体及信息检索;马正华(1962-),男,教授,学士,主要研究方向为嵌入式系统应用、计算机控制技术。

与互操作性。酒店领域内容丰富,所建本体根据不同的研究目的是有区别的。酒店本体是本文推荐方法的数据源,因此在设计过程中,着重考虑并加入有利于推荐方面的概念。如此,本体的主要概念有:酒店类型、酒店特性(包括位置、环境在内的一些属性)等。建立酒店本体主要解决了以下问题:如何对酒店统一描述;用户可以选择哪些酒店。本体的描述语言使用了 OWL-DL^[14],因为它拥有较好的表达能力,同时又不乏可计算性。建立本体的工具选用了斯坦福大学的 Protégé 4.2。一个酒店实例局部的描述如下:

```
<SubClassOf>
  <Class IRI = "#Luxury" />
  <ObjectMinCardinality cardinality = "15" />
  <ObjectProperty IRI = "#hasRoom" />
  <Class IRI = "#LuxuryRoom" />
</ObjectMinCardinality>
</SubClassOf>
<ClassAssertion>
  <Class IRI = "#ResortHotel" />
  <NamedIndividual IRI = "#SurfPlazaResortHotel" />
</ClassAssertion>
<ObjectPropertyAssertion>
  <ObjectProperty IRI = "#hasCreditCardAccept" />
  <NamedIndividual IRI = "#SurfPlazaResortHotel" />
  <NamedIndividual IRI = "#UnionPayCard" />
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
  <ObjectProperty IRI = "#hasObjectFacility" />
  <NamedIndividual IRI = "#SurfPlazaResortHotel" />
  <NamedIndividual IRI = "#TennisCourt" />
</ObjectPropertyAssertion>
<DataPropertyAssertion>
  <DataProperty IRI = "#hasAddress" />
  <NamedIndividual IRI = "#SurfPlazaResortHotel" />
  <Literal datatypeIRI = "#&rdf;PlainLiteral" /> 山东省青岛市崂山区香港东路 316 号</Literal>
</DataPropertyAssertion>
<DataPropertyAssertion>
  <DataProperty IRI = "#hasTel" />
  <NamedIndividual IRI = "#SurfPlazaResortHotel" />
  <Literal datatypeIRI = "#&rdf;PlainLiteral" /> 0532-88890394</Literal>
</DataPropertyAssertion>
<AnnotationAssertion>
  <AnnotationProperty abbreviatedIRI = "#rdfs:comment" />
  <IRI /> #SurfPlazaResortHotel</IRI>
  <Literal datatypeIRI = "#&rdf;PlainLiteral" /> 中国山东省青岛弄海园酒店</Literal>
</AnnotationAssertion>
```

1.2 需求词汇表

通常用户在表达他们的需求时,描述都是模糊的,如“附近的”“可口的”。为了快速准确识别出用户的需求目的,需要一种方法把用户的查询需求转变成对推荐有效的信息,考虑到这些原因,本文提出了需求词汇表。词汇表的部分如表 1 所示,左边第一栏是需求类别,第二栏对应该需求类别下的概念。根据用户提出的查询(直接需求)锁定需求的类别,再根据具体要求得到检索条件。需求词汇表的主要作用是在推荐方法中处理用户的直接需求。

1.3 用户建模

用户是推荐的对象,本文采用基于知识推理的方式处理推荐过程中的个性化匹配问题。这种方式需要建立推理规则,而推理规则需要有效数据源才能实现推理,为此,建立了一个用户资料本体。设计的用户本体应有利于使用推理规则,在建立过程中,尤其是用户偏好的描述部分,考虑了酒店本体可能与

之关联的概念。用户在 Web 页面上注册个人资料,将这些信息映射到本体中,主要的建模概念有用户(类)、身份、联系信息、个人统计资料(如年龄、性别、收入等)、偏好与兴趣等。建好用户本体是应用推理规则的前提和基础,主要解决了以下问题,即如何对用户进行统一描述,推荐需要用到哪些用户特征。用户本体的建模工具与描述语言同酒店本体。一个用户实例局部的描述如下:

```
<EquivalentClasses>
  <Class IRI = "#User" />
  <ObjectIntersectionOf>
    <ObjectSomeValuesFrom>
      <ObjectProperty IRI = "#hasPreference" />
      <Class IRI = "#Preference" />
    </ObjectSomeValuesFrom>
    <DataAllValuesFrom>
      <DataProperty IRI = "#hasAge" />
      <Datatype abbreviatedIRI = "#xsd:int" />
    </DataAllValuesFrom>
    <DataMinCardinality cardinality = "1" />
    <DataProperty IRI = "#hasEmail" />
    </DataMinCardinality>
    <DataMinCardinality cardinality = "1" />
    <DataProperty IRI = "#hasTel" />
    </DataMinCardinality>
  </ObjectIntersectionOf>
</EquivalentClasses>
<ClassAssertion>
  <Class IRI = "#User" />
  <NamedIndividual IRI = "#User001" />
</ClassAssertion>
<DataPropertyAssertion>
  <DataProperty IRI = "#hasUserName" />
  <NamedIndividual IRI = "#User001" />
  <Literal datatypeIRI = "#&rdf;PlainLiteral" /> John Steven</Literal>
</DataPropertyAssertion>
<DataPropertyAssertion>
  <DataProperty IRI = "#hasEmail" />
  <NamedIndividual IRI = "#User001" />
  <Literal datatypeIRI = "#&rdf;PlainLiteral" /> JohnS@ hotmail. com</Literal>
</DataPropertyAssertion>
<ObjectPropertyAssertion>
  <ObjectProperty IRI = "#hasPreference" />
  <NamedIndividual IRI = "#User001" />
  <NamedIndividual IRI = "#Tennis" />
</ObjectPropertyAssertion>
```

表 1 需求识别词汇表

需求类别	相关概念
位置	靠海边、在北京、崂山景区、附近有地铁站…
规格/类型	会议、工作、宴会、疗养、度假…
服务	洗衣、无线网络、订机票、租车、叫醒…
环境	干净、空气好、气温适宜…
舒适程度	有沙发、按摩浴缸、有 HDTV…
⋮	

1.4 推理规则

本体只能描述一件事情,如果要做更进一步的推论,需要借助另外的补充。通过建立推理规则表达用户的偏好兴趣与酒店特性的联系,使之产生语义关联。规则集针对用户本体建立和使用,以用户偏好为核心,取出用户本体中表示用户偏好的属性(兼有 ObjectProperty 和 DataProperty),根据具体的偏好建立具体的规则,推理得到的结果是酒店的某些属性。例如,

用户的描述拥有某种运动的偏好,这就意味着那些提供了这种运动设施的酒店能够符合该用户的这一偏好,这个规则的描述如下:

@ prefix u: <http://http://www.semanticweb.org/ontologies/2012/4/TourUser.owl#>.

@ prefix htl: <http://www.semanticweb.org/ontologies/2012/5/TourHotel.owl#>.

@ include <RDF>.

[rule1: (? a u:hasPreference ? b) (? b rdf:type u:Sport) (? a u:suitFacility ? c) → (? c htl:hasFacilityFunction ? b)]

建立推理规则主要解决了个性化推荐中用户的个性特征与酒店的具体属性相匹配的问题,实现用户偏好与酒店特性之间的概念转化,使推荐的过程“有法可依”。

本文使用 Jena 框架^[15]来建立和操作推理规则。Jena 是一个为构建语义网应用程序的 Java 框架。它提供了一个可对 RDF、RDFS、OWL 编程化的环境,同时提供了 SPARQL^[16]查询语言和基于规则的推理引擎。

2 推荐方法

2.1 第一次推荐

第一次推荐工作是对用户直接需求的处理。用户先对旅行计划提出自己的要求,称为直接需求,这些需求应当认为是必须的条件。根据这些条件在酒店本体中检索,得到符合条件的酒店实例,并将这些酒店作为第二次推荐的数据源。

算法 1 筛选出满足用户直接需求的酒店过滤算法

a) 获取直接需求。提取用户直接需求中的需求元素 (query element) $qe_i, i \in N^+$ 。

b) 需求分类。将 qe_i 与需求识别词汇表中的内容进行匹配,找到 qe_i 对应的类别 cat_{e_i} 。

c) 获取第一次过滤的条件。在类别 cat_{e_i} 对应的概念中找到与 qe_i 相同的全部概念 $k_{ij} (i, j \in N^+)$ 并取出,表示为 (cat_{e_i}, k_{ij}) 的形式,将这组集合作为过滤条件,即酒店必须具备的特性集合 $HC = \{hc_1, hc_2, hc_3, \dots, hc_n\}$ 。

d) 检索符合条件内容。通过 (cat_{e_i}, k_{ij}) 的集合在酒店本体中筛选出满足条件的酒店实例集合 $H = \{h_1, h_2, h_3, \dots, h_n\}$ 。

2.2 第二次推荐

第二次推荐工作是对第一次推荐得到的结果进行个性化匹配。利用推理规则将用户偏好转换成对应的酒店特性,作为新的过滤条件。筛选出既满足用户提出的需求,又符合用户个性偏好的酒店。

算法 2 筛选出符合用户偏好兴趣的酒店过滤算法

a) 获取用户偏好信息。在用户本体中找到当前用户实例 $I_{current}$ 对应的偏好信息集合 $UP = \{up_1, up_2, up_3, \dots, up_n\}$ 。

b) 获取第二次过滤的条件。用户本体载入规则集 R , 将 UP 转换成一组新的酒店特性集合 $HC' = \{hc'_1, hc'_2, hc'_3, \dots, hc'_n\}$ 。

c) 检索符合条件的内容。以 HC' 作为过滤条件,对第一次过滤的结果进行再次筛选,得到酒店实例集合 $H' = \{h'_1, h'_2, h'_3, \dots, h'_n\}$ 。

两次推荐流程主要步骤如图 1 所示。

2.3 结果排序方法

为了衡量推荐结果对用户的适用程度,把每个具体的酒店特性对具体的用户偏好的契合度用权重表示。计算第二次过

滤结果 H' 中每一个酒店的权重,作为适用度的值。

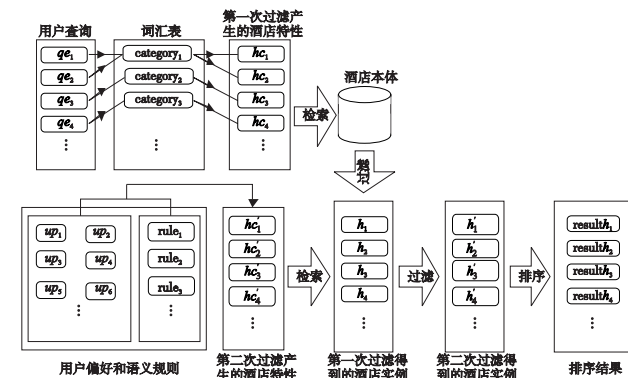


图1 过滤过程

算法 3 推荐结果的排序算法

a) 获取第二次推荐结果集 H' 中的酒店实例 h'_i , 和该酒店对应的特性集 HC'_i 。

b) 利用式(1)计算酒店 h'_i 的总契合度:

$$D_{hotel} = \sum_{i=1}^n \sum_{j=1}^m w_{ij} \quad (1)$$

$$w_{ij}: up_i \rightarrow hc_j$$

其中,权重 w_{ij} 是代表用户偏好 up_i 对应酒店特性 hc_j 的契合度,初值为 1 (1 为最大值)。

c) 通过计算结果集中每个酒店的总权重,将结果集中的酒店实例按照总权重从大到小排列。

d) 将得到的最终结果集 H_{result} 推荐给用户。

上述排序算法中, w_{ij} 的值由以下因素决定:

a) 概念对应层级。因为酒店使用本体表示,作为一种知识表示方式,本体中的概念是分层级的,若酒店特性与用户偏好恰好对应,即在同一知识层级上对应,则赋予契合度初值 1。例如,某个用户喜好运动类型中有游泳,而某个酒店的运动设施中正好有游泳池。若酒店的特性包含有用户的偏好,即在不同知识层级上 (通常不会超过两级),则赋予契合度初值由 1 逐渐向下调整。例如,某个用户资料只描述喜好运动,某个酒店提供运动设施中有网球场和游泳池,但不清楚用户是否喜欢这两种运动。

b) 社区网络用户评价。这些契合度值是动态的,需要根据实际情况调整。为此,本文从目前比较权威的旅游社区网站“去哪儿网”(qunar.com)收集了一些用户对酒店的具体特性的评价,根据评价的好坏来对契合度的初值进行调整。例如,虽然某酒店有游泳池,但是社区网络中的用户对其服务质量评价较低,就要对初值进行下调。这样做可以使酒店的信息尽可能真实,并更符合用户需求。

2.4 获取用户反馈信息

本文认为很有必要考察用户对推荐的解决方案的认同程度,作为以后推荐工作的参考。另外,用户对结果的评价往往能反映出推荐流程的问题。因此采用根据用户的反馈来定期调整知识库的策略,图 2 表明了反馈过程。本文引入 Russell 等人^[17]提出的情感网格 (affect grid) 来表达用户对推荐结果的感受。情感网格是一种快速的量化评估手段,可以用来衡量用户主观上是否满意、高兴和感兴趣,在考察用户的情感方面具有良好的效果。当某个推荐结果一个月内用户评价超过 50 个,统计用户总体感受决定是否有必要做出调整,修改词汇表、

规则集、酒店和用户的本体等。这是一个循环的过程。



图2 用户反馈过程

3 实验验证

3.1 推荐实施系统建模

图3描述了一个三层结构方案来表示实施推荐方法的 Demo 系统模型。用户在推荐界面输入直接需求,经过需求分类模块,利用需求词汇表在酒店本体中检索,得到第一次推荐结果。推理模块使用 Jena 框架组建,将规则结合用户本体进行推理,得到个性化匹配的过滤条件,在第一次推荐结果中检索,得到结果经排序后返回给用户。用户通过评价模块对推荐结果进行评价,反馈信息保存在数据库中。管理员使用管理维护界面面对系统进行设置与调整。

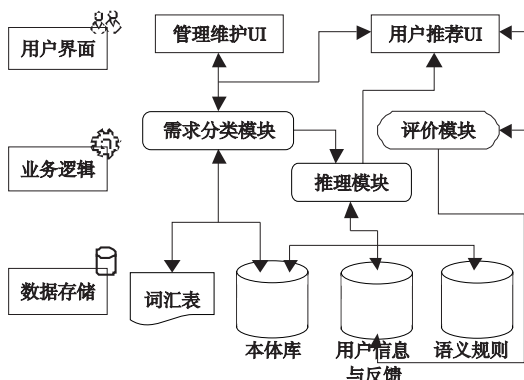


图3 Demo系统架构

3.2 实验方案

为了验证本文提出的推荐方法在一个可控的场景下是否具有令人满意的表现,分别设计了两组实验。第一组实验是 Demo 系统、专家组和用户之间两两对比,分三次测试评价该推荐方法;第二组实验是本文提出的方法与 Slope One 推荐算法^[18]的比较。根据 García-Crespo 等人的研究^[19],评价推荐方法性能的指标选择查准率(precision)、查全率(recall)以及 F1,定义如下:

$$\text{precision} = \frac{\text{CorrectHotelsFound}}{\text{TotalHotelsFound}} \quad (2)$$

$$\text{recall} = \frac{\text{CorrectHotelsFound}}{\text{TotalCorrectHotels}} \quad (3)$$

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

其中,满足条件的酒店(correct hotels)在不同实验情况下由专家或用户的选择为标准。

3.3 实验及结果分析

3.3.1 第一组实验

实验样本选择了 10 个大学在校学生使用 Demo 系统来为暑假旅游出行寻找合适的酒店。为确保实验稳定,酒店样本取中国江苏省的南京和常州两地共 20 个酒店,这些酒店的信息

来自目前比较流行的旅游网站“去哪儿网”。另外,5 个旅游领域的专业人员构成专家组将为学生提供推荐或帮助。

在第一次测试中,满足条件的酒店以专家推荐为标准。在第二次和第三次测试中,满足条件的酒店以用户的选择为标准,并且专家组把全部 20 个酒店的信息提供给每一个学生,让学生自己选择需要的酒店。实验要求学生至少选择两个酒店,因为考虑到选择的酒店可能出现客满的情况,所以在条件允许的情况下,必须要选一个备用的酒店。表 2 是三次测试的结果,其中的数据是取 10 个学生对应结果的平均值。

表 2 三次实验结果

对比项	precision	recall	F1
系统推荐对比专家推荐	0.594	0.880	0.698
专家推荐对比用户选择	0.712	0.892	0.789
系统推荐对比用户选择	0.488	0.917	0.627

从第一次测试结果来看,系统推荐的查准率不到 60%,也就是说系统推荐的结果有 59.4% 与专家推荐的相同,系统推荐的查全率则达到了 88%。通过第一次实验可以发现系统的基本性能尚佳。

在第二次与第三次测试中,以用户的选择作为基准,从专家推荐与用户选择的比较来看,查准率为 71.2%,说明专家的推荐有 71.2% 被用户接受,从侧面反映了专家提供的建议比较符合实际情况;而查全率为 89.2%,说明专家的推荐结果覆盖了用户选择的近 90%。从系统推荐与用户选择比较来看,此时系统的查准率只有 48.8%。其有两个原因:a)用户与专家的观点本身就不是 100% 相同;b)专家的推荐可以理解成一种“精炼”,因此不会为用户推荐较多结果,系统的计算能力比专家强,只要满足条件就都可以进入推荐列表,那么返回的结果肯定比专家多(从这里也可以反映出第一次实验查准率低的原因),在用户的选择不变的情况下,查准率下降。然而,正因为系统返回的结果多于专家,促使系统查全率达到了 91.7%,这是一个不错的结果,说明系统推荐的结果覆盖了用户选择的 90% 以上。

3.3.2 第二组实验

本组实验是与 Lemire 教授在 2005 年提出的 Slope One 推荐算法^[19]进行测试比较。Slope One 推荐算法是一种易于实施的基于项目(item-based)的协同过滤算法,其主要思路是利用用户对不同项目打分(评价)的平均差别预测某个用户对某个产品可能的评分。

为了将本文提出的推荐方法与 Slope One 推荐算法进行比较,实验样本又加入了 5 名学生和 10 个酒店,分别是与之前的 10 名学生和 20 个酒店同来源、同类型的不同样本。具体实验方法如下: S 和 H 分别代表实验学生集和酒店集; RS 和 RH 分别代表后加入的学生集和酒店集, RS 和 RH 作为参考样本预测实验学生集 S 对实验酒店集 H 中酒店的评分。表 3 表示了应用 Slope One 算法通过学生 rs_1 和 rs_2 分别对酒店 rh_1 、 rh_2 和 h_1 的评分(满分 5 分),预测出实验学生 s_1 对实验酒店 h_1 评分的一个示例。

表 3 中 $rh_i, i \in [1, 10]$ 是 RH 中的酒店实例; $rs_i, i \in [1, 5]$ 是 RS 中的学生实例; $h_i, i \in [1, 20]$ 是 H 中的酒店实例; $s_i, i \in [1, 10]$ 是 S 中的学生实例; x 代表 S 中的实验学生 s_1 对 H 中实验酒店 h_1 可能的评分,为预测值。计算预测值公式如下:

$$P(u)_j = \bar{u} + \frac{1}{\text{card}(R_j)} \sum_{i \in R_j} \text{dev}_{j,i} \quad (5)$$

其中: u 代表一个学生对酒店的评价集(u 中必须同时包括对 h_i 和 h_j 的评价); $P(u)_j$ 是对酒店 h_j 的预测分值; $\bar{u}$ 是该学生对一组酒店评分的均值; $\text{card}(R_j)$ 是 R_j 的数目; R_j 是所有与 u 中评价关联的酒店集合; $\text{dev}_{j,i}$ 的计算式为

$$\text{dev}_{j,i} = \frac{u_j - u_i}{\sum_{u \in S_{j,i}(\chi)} \text{card}(S_{j,i}(\chi))} \quad (6)$$

其中, $\text{card}(S_{j,i}(\chi))$ 是学生同时对酒店 h_i 和 h_j 评价的数目。经计算, 表 3 中预测值 $x = 4.25$ 。

表 3 应用 Slope One 推荐算法的一个示例

	rh_1	rh_2	h_1
rs_1	3	4	4
rs_2	4	4	5
s_1	3	4	x

通过扩展表 3 纳入参考学生集 RS 与酒店集 RH 中全部实例, 预测出实验学生集 S 中每一个学生酒店集 H 中 20 个酒店的全部 200 个评分, 将 s_i 对 H 中酒店全部评分预测值按从大到小排列, 取前 5 个推荐给 s_i 。此时, 将第一组实验中用户选择的酒店作为式 (2) (3) 中的 correct hotels 分别计算 precision、recall 和 $F1$ 的值(数据取 10 个实验学生对应结果平均值)并与第一组实验中系统推荐对比用户选择部分的结果进行比较。比较结果如图 4 所示, 可以发现, 本文提出的方法在查准率、查全率和 $F1$ 方面均优于 Slope One 推荐算法。使用 Slope One 推荐算法效果不理想的原因可能有: Slope One 算法是一种协同过滤算法, 它并不关心酒店本身具有的特性, 只考虑用户对酒店的评分, 而且该评分是唯一的衡量标准, 导致某些硬性条件不符合或几乎不符合某个用户需求的酒店得到该用户的评分, 并且受到其他用户的影响, 其预测评分值很可能与该用户本意完全不一样。在实验中有一名学生用 Slope One 算法得到的推荐结果与其自身的选择没有交集。除此以外, 用户之间的相似度也是一个影响因素。

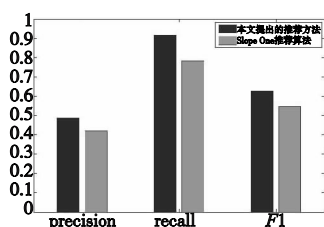


图4 本文方法与Slope One算法的比较实验结果数据

4 结束语

本文描述了一种基于本体的二次推荐方法, 兼顾用户直接需求和个性化匹配。总体来看, 研究特点在于解决用户需求及偏好和酒店特性之间的转化方面。在推荐过程中, 将用户的直接需求和偏好匹配分步处理, 从理论上保证了结果满足条件不遗漏。经实验验证, 该方法效果较好, 能够为用户接受。下一步的工作是扩大该方法应用范围, 如针对更多的旅游资源和服务; 另外, 继续研究改进推荐算法, 以提高推荐的效率。

参考文献:

[1] NIININEN O, BUHALIS D, MARCH R. Customer empowerment in tourism through consumer centric marketing (CCM) [J]. *Qualitative Market Research: An International Journal*, 2007, 10(3): 265-

281.

- [2] 徐海玲, 吴潇, 李晓东, 等. 互联网推荐系统比较研究[J]. *软件学报*, 2009, 20(2): 350-362.
- [3] 牟向伟, 陈燕. 基于模糊描述逻辑的个性化推荐系统建模[J]. *计算机应用研究*, 2011, 28(4): 1429-1433.
- [4] ACIAR S V, SERAROLS-TARRES C, ROYO-VELA M, et al. Increasing effectiveness in e-commerce: recommendations applying intelligent agents [J]. *International Journal of Business and Systems Research*, 2007, 1(1): 81-97.
- [5] LENAR M, SOBECKI J. Using recommendation to improve negotiations in agent-based systems [J]. *Journal of Universal Computer Science*, 2007, 13(2): 267-286.
- [6] HUANG Yu-xia, BIAN Ling. A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet [J]. *Expert Systems with Applications*, 2009, 36(1): 933-943.
- [7] NIEMANN M, MOCHOL M, TOLKSDORF R. Enhancing hotel search with semantic Web technologies [J]. *Journal of Theoretical and Applied Electronic Commerce Research*, 2008, 3(2): 82-96.
- [8] KANELLOPOULOS D N. An ontology-based system for intelligent matching of travellers' needs for group package tours [J]. *International Journal of Digital Culture and Electronic Tourism*, 2008, 1(1): 76-99.
- [9] GARCIA-CRESPO A, CHAMIZO J, RIVERA I, et al. SPETA: social pervasive e-tourism advisor [J]. *Telematics and Informatics*, 2009, 26(3): 306-315.
- [10] GARCIA-CRESPO A, COLOMO-PALACIOS R, GOMEZ-BERBIS J M, et al. Intelligent decision-support systems for e-tourism; using SPETA II as a knowledge management platform for DMOs and e-tourism service providers [J]. *International Journal of Decision Support System Technology*, 2010, 2(1): 35-47.
- [11] 许捷华, 王志坚, 林巧民, 等. 一种基于偏好的个性化标签推荐系统[J]. *计算机应用研究*, 2011, 28(7): 2573-2575.
- [12] GARCIA-CRESPO A, LOPEZ-CUADRADO J L, COLOMO-PALACIOS R, et al. Sem-Fit: a semantic based expert system to provide recommendations in the tourism domain [J]. *Expert Systems with Applications*, 2011, 38(10): 13310-13319.
- [13] GRUBER T R. A translation approach to portable ontology specifications [J]. *Knowledge Acquisition*, 1993, 5(2): 199-220.
- [14] SMITH M K, WELTY C, MCGUINNESS D L. OWL Web ontology language guide [EB/OL]. (2004) [2012-03-31]. <http://www.w3.org/TR/owl-guide/>.
- [15] The Apache Software Foundation. What is Jena? [EB/OL]. (2011) [2012-03-31]. http://incubator.apache.org/jena/about_jena/about.html.
- [16] PRUD' HOMMEAUX E, SEABORNE A. SPARQL query language for RDF [EB/OL]. (2008) [2012-03-31]. <http://www.w3.org/TR/rdf-sparql-query/>.
- [17] RUSSELL J A, WEISS A, MENDELSON G A. Affect grid: a single-item scale of pleasure and arousal [J]. *Personality and Social Psychology*, 1989, 57(3): 493-502.
- [18] LEMIRE D, MACLACHLAN A. Slope one predictors for online rating-based collaborative filtering [C] // Proc of SIAM Data Mining Conference. 2005: 471-475.
- [19] GARCIA-CRESPO A, RODRIGUEZ A, MENCKE M, et al. ODD-IN: ontology-driven differential diagnosis based on logical inference and probabilistic refinements [J]. *Expert Systems with Applications*, 2010, 37(3): 2621-2628.