

# 上下文感知的移动社交网络服务选择机制研究

王玉祥 乔秀全 李晓峰 孟洛明  
(北京邮电大学网络与交换技术国家重点实验室 北京 100876)

**摘 要** 服务选择机制作为为用户提供智能性、主动性应用服务的一项关键支撑技术,是移动社交网络中重要的研究内容之一.文中提出了基于上下文、信任网络和协作过滤算法的移动社交网络服务选择机制,将上下文相似度引入到服务选择的过程中,并且和信任度相结合,构成“用户-服务-上下文”三维协作过滤服务选择模型.该方法提高了服务选择的准确性和可靠性,避免了服务选择的盲目性和随意性.仿真实验表明,与传统的协作过滤推荐算法相比,该服务选择算法具有更高的准确率.

**关键词** 移动社交网络服务;服务选择;上下文感知;协作过滤算法;平均绝对误差(MAE)

中图法分类号 TP311 DOI号: 10.3724/SP.J.1016.2010.02126

## Research on Context-Awareness Mobile SNS Service Selection Mechanism

WANG Yu-Xiang QIAO Xiu-Quan LI Xiao-Feng MENG Luo-Ming

(State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876)

**Abstract** The service selection mechanism, as a key enabling technology to provide users with intelligent and proactive services, is one of the important research topics in Mobile SNS (Social Networking Service). This paper presents the service selection mechanism for mobile SNS based on user context, trust network and collaborative filtering algorithm. By importing the context similarity into the service selection process and also considering users' trust value, this paper proposes a “user-service-context” three-dimensional collaborative filtering model. Based on the user context information, the accuracy and reliability of service selection are improved, and the blindness and arbitrary are also avoided. Evaluation results show that the service selection approach based on context, trust relationships and collaborative filtering algorithm outperforms the traditional collaborative filtering algorithm in the accuracy aspect.

**Keywords** mobile social networking service; service selection; context-awareness; collaborative filtering algorithm; Mean Absolute Error(MAE)

## 1 引 言

随着社交网络服务(Social Networking Service, SNS)的出现,尤其是移动社交空间 MOSS(Mobile Social Spaces)<sup>[1]</sup>的兴起,新兴社交网络中大规模涌

现出各种丰富多彩的服务和内容.如何从纷繁复杂的相似或者相近的服务中为用户选择最佳服务,做到真正的以用户为中心,提供更好的应用和服务体验,服务选择机制逐渐成为移动社交网络中的一项关键支撑技术.服务选择是为用户提供智能性、主动性应用服务的前提基础.

收稿日期:2010-06-08;最终修改稿收到日期:2010-08-31.本课题得到国家自然科学基金(60802034,60672122)、高等学校博士学科点专项科研基金(20070013026)、北京市科技新星计划(2008B50)以及“新一代宽带无线移动通信网科技重大专项课题(2011ZX03002-002-01)”资助.王玉祥,男,1971年生,博士,主要研究方向为网络管理与通信软件.乔秀全(通信作者),男,1978年生,副教授,主要研究方向为网络服务智能化理论及技术.E-mail: qiaoxq@bupt.edu.cn.李晓峰,女,1950年生,教授,主要研究领域为网络智能与通信软件.孟洛明,男,1955年生,教授,博士生导师,主要研究领域为通信网与网络管理.

服务选择就是要根据用户的移动性、其他服务使用者的服务评价信息、服务的功能性以及非功能属性等综合因素, 从众多相同或者相似的服务中为用户提供最合适、最相关的服务。服务的使用者(用户)的上下文信息也是影响服务选择的重要因素(例如用户的兴趣偏好、位置等上下文信息), 对服务选择起到了至关重要的作用, 尤其是用户的动态上下文信息。传统社交网络通常依据的是服务的静态属性信息, 基本上没有考虑用户的动态变化的上下文环境。与传统的面向互联网的 SNS 相比, 移动社交网络无缝地将移动计算(mobile computing)和社会计算(social computing)结合起来, 极大地增强了用户的真实性、地域性和交互的实时性。为了向用户提供适时适地的合适服务, 服务选择除了要满足用户的需求之外, 还要适应环境的动态变化。在移动社交网络系统中, 用户的上下文信息得到了极大的丰富, 如手机终端能力信息、用户的位置信息(通过手机 GPS 或者移动网络定位获得)、用户的状态(Presence)信息(如在线、离线、开会等)、手机日历/日程信息、用户的偏好信息、用户的社会关系网络和好友的评价信息等。因此, 如何基于用户的社会关系网络, 利用其他用户的评价信息, 并根据用户本身的动态上下文信息来为用户提供随时、随地的智慧型主动服务, 是移动社交网络服务进一步发展所面临的挑战。

不同的领域和不同的专业对服务有不同的理解, 目前服务计算领域对服务的一个通用理解是: 服务表示一种至少包含一个服务提供者和一个服务消费者之间的交互(活动)以取得一定的商业目标或解决问题的目的。本文所指的移动社交网络服务是指利用移动社交网络为用户提供的各种服务, 如用户根据当前的位置和偏好、好友的评价信息等选择兴趣点(如饭馆、酒店等)的服务, 具有一般意义的服务。现有的基于位置的移动社交网络服务提供系统能够根据用户的当前位置来把附近的某些兴趣点(如饭店、商场、咖啡店等)显示出来, 用户可以选择签到(CheckIn), 从而获取积分、打折券等, 其不足之处就是: 没有做到根据用户本人的偏好以及其社会关系网络中好友的评价信息来向用户推荐服务的功能, 用户社会关系网络中熟人的评价信息没有发挥实质性的导向作用。而在现实生活中, 熟人推荐往往是促成用户发生消费行为的一个非常重要的手段。目前, 国内外研究者针对不同的问题域对服务选择问题进行了大量的研究, 出现了许多解决办法。比如在电子商务领域的商品选择/推荐中也涉及到类似

的研究<sup>[2-3]</sup>, 用户购买商品之后可以对产品或者服务进行评价, 后来的用户可以根据这些评价信息作出是否购买的判断, 很多文献研究利用协作过滤算法来分析、比对用户之间的消费行为, 从而产生相关的推荐, 供用户进行选择。然而其存在的明显缺陷是: 一方面, 目前的电子商务系统中用户与用户之间缺乏实质性联系, 没有建立起相关的社会关系网络, 导致用户对有些用户评价信息的可信度产生怀疑; 另一方面, 现有的电子商务系统基本上是基于用户静态的上下文信息, 如注册的个人资料(偏好或者已发生的消费历史记录), 没有与用户实时实地的动态上下文信息相关联。另外, 服务选择的研究还体现在 Web 服务领域, 如 Maximilien 等<sup>[4]</sup>认为服务选择应基于用户偏好、商业策略以及考虑服务的信任关系, 他们利用决策理论和本体, 开发了一个基于本体和信任模型的多 Agent 框架。Casati 等<sup>[5]</sup>假定服务会话日志作为每个业务过程执行和服务消费者指定的质量目标, 然后利用数据挖掘技术构造作为分类模型的决策树, 使所选择的服务最大可能满足用户定义的目标。Kalepu 等<sup>[6]</sup>归纳了现已提出的 Web 服务选择服务质量 QoS 需要包含的属性, 例如可用性、可靠性、性能和代价等, 并提出还需要增加一个真实性属性。Sreenath 等<sup>[7]</sup>认为服务选择必须考虑用户观点, 不同于基于传统推荐方法的产品选择, 他们对服务选择问题进行了形式化, 提出了一种利用 Agent 协作评估服务提供者的方法。东北大学的林涛<sup>[8]</sup>等人分析了人类社会的信任关系, 并以人们之间的信任关系模型为基础, 对普适计算环境进行建模, 将传统服务选择模型中的服务定义扩充, 提出了一种基于信任关系的服务选择算法。文献[9-10]提出了一种基于语义 Web 的上下文感知动态服务选择机制, 这种机制可以根据待匹配的服务的动态上下文属性来对其进行过滤和排序, 从而为用户提供相关的服务。文献[11-12]还提出了基于 QoS 的 Web Service 服务选择经验算法。虽然 Web 服务领域已有的服务选择方法对服务选择问题中的语义描述、服务发现、系统架构、服务绑定等都提供了一些很好的解决方法, 但还远远没有解决服务选择问题面临的用户动态上下文的挑战, 用户动态上下文的获取一直缺乏相应的实际环境支撑, 而且信任关系在实际的 Web 服务应用中存在着严重缺失, 在工业界缺乏生存的土壤和发展的动力。

而移动社交网络能够很好地解决基于用户评价信息的可信度、用户动态上下文信息以及用户偏好

和消费行为的服务选择问题,从而为用户适时适地提供合适的个性化服务.用户评价信息的可信度可以基于社会关系网络利用信任网络理论来解决;用户的动态上下文信息可以通过智能手机终端和应用服务器来获取,解决了用户动态上下文的获取问题;通过社会关系网络中偏好相似的用户行为习惯来向用户推荐服务,从而提供服务选择的候选目标.因此,本文提出了一种“用户-服务-上下文”三维协作过滤的服务选择模型,将动态上下文信息、信任网络和协作过滤算法巧妙地结合起来研究移动社交网络服务的选择机制.本文引入上下文相似度,在上下文环境相似的情况下,结合用户的社会关系即信任网络,将用户信任度和相似度结合起来,能够较全面准确地进行服务选择,比较传统的服务选择方法,基于上下文、信任网络和协作过滤服务选择机制既避免了传统协作过滤算法的稀疏性和冷启动性,同时又考虑了上下文的动态性,更能接近现实世界和具有更高的服务选择准确率.

本文第2节首先介绍传统协作过滤算法的基本流程;第3节提出基于上下文、信任网络和协作过滤的服务选择机制,并详细阐述服务选择方法的整体流程和具体执行过程;第4节设计移动社交网络中兴趣点服务选择应用场景来验证提出的基于上下文、信任网络和协作过滤的服务选择机制.仿真实验验证了该方法的有效性和合理性;最后给出结论和未来的工作.

## 2 协作过滤算法基础知识介绍

本文主要是基于协作过滤算法,并结合信任网络和上下文感知技术来研究移动社交网络系统中的服务选择问题.因此,首先简要介绍协作过滤算法相关的基础知识<sup>[13-14]</sup>.

推荐算法分为基于内容的推荐算法和协作过滤推荐算法,基于内容的推荐算法需要内容与用户模型严格地“关键词”匹配,只能推荐结构化对象,因此常用的推荐算法是协作过滤推荐算法,主要是用在电子商务的商品推荐上.协作过滤推荐是目前推荐系统中使用最多的个性化推荐技术,它基于邻居用户的评分数据得到目标用户的推荐,其推荐的个性化程度高.协作过滤算法的最大优点是对推荐对象没有特殊要求,能处理非结构化的复杂对象,如书店、音乐、餐馆等.

传统协作过滤算法流程大体分为4步:首先输

入用户-服务评分矩阵;然后选择用户相似性度量方法;再形成最近用户邻居集合;最后产生推荐数据.传统协作过滤算法的详细推荐过程描述如下.

### 2.1 用户模型的建立

在一个采用协作过滤技术的推荐系统中,用户的评分数据可以用一个  $n \times m$  阶用户-服务评分矩阵  $R$  来表示.  $n$  行代表  $n$  个用户,  $m$  列代表  $m$  个服务,第  $i$  行第  $s$  列的元素  $R_{i,s}$  代表用户  $i$  对服务  $s$  的评分.评分可以用 2 进制的 0 和 1 来表示用户偏好(喜欢/不喜欢)或购买状态(已购买/未购买),也可以用分级的用户对服务的喜好值表示(例如:MovieLens 中用户对电影的评分是用 0~5 之间的整数来代表用户的喜好,0 表示没有评分,1~5 表示用户喜欢程度).

### 2.2 选取用户相似性度量公式

常用的有如下几种相似性度量公式可供选择:

#### (1) 余弦相似性度量公式

余弦相似度,即余弦相关系数度量,用户评分可看作  $m$  维服务空间上的向量,用户  $i$  和用户  $j$  之间的相似度通过向量间的余弦夹角度量,余弦值越大表明两个用户的相似程度越高.设用户  $i$  和用户  $j$  在  $m$  维服务空间上的评分数据分别表示为向量  $i$  和向量  $j$ ,则用户  $i$  和用户  $j$  的相似度  $sim(i, j)$  为

$$sim(i, j) = \cos(\vec{i}, \vec{j}) = \frac{\vec{i} \cdot \vec{j}}{\|\vec{i}\| \cdot \|\vec{j}\|} \quad (1)$$

其中,分子为两个用户评分向量的内积,分母为两个用户向量模的乘积.

#### (2) 修正的余弦相似性度量公式

修正的余弦相似度:即 Spear 相关系数度量,余弦相似性度量方法中并未考虑不同用户的评分尺度问题,为了改善上述缺陷,修正的余弦相似性度量方法减去用户对服务的平均评分,设用户  $i$  和用户  $j$  共同评分过的服务集合用  $I_{i,j}$  表示,  $I_i$  和  $I_j$  分别表示用户  $i$  和用户  $j$  评分过的服务集合.  $R_{i,s}$  表示用户  $i$  对服务  $s$  的评分,  $R_i$  和  $R_j$  分别表示用户  $i$  和用户  $j$  的平均评分,则用户  $i$  和用户  $j$  相似度  $sim(i, j)$  为

$$sim(i, j) = \frac{\sum_{s \in I_{i,j}} (R_{i,s} - R_i)(R_{j,s} - R_j)}{\sqrt{\sum_{s \in I_i} (R_{i,s} - R_i)^2} \sqrt{\sum_{s \in I_j} (R_{j,s} - R_j)^2}} \quad (2)$$

#### (3) 相关相似性度量公式

相关相似性度量公式,即 Pearson 相关系数度量:设用户  $i$  和用户  $j$  共同评分过的服务集合为  $I_{i,j}$ ,则用户  $i$  和用户  $j$  的相似度  $sim(i, j)$  为

$$sim(i, j) = \frac{\sum_{s \in I_{ij}} (R_{i,s} - R_i)(R_{j,s} - R_j)}{\sqrt{\sum_{s \in I_{ij}} (R_{i,s} - R_i)^2} \sqrt{\sum_{s \in I_{ij}} (R_{j,s} - R_j)^2}} \quad (3)$$

其中,  $R_{i,s}$  表示用户  $i$  对服务  $s$  的评分,  $R_i$  和  $R_j$  分别表示用户  $i$  和用户  $j$  对服务的平均评分. 注意, 平均评分是指两个用户有共同评分服务的平均值. 求得和相关相似性系数的范围是  $[-1, 1]$ , 相关系数越大, 则表示这两个用户的兴趣爱好越接近, 其推荐的准确率就越高.

### 2.3 邻居的形成

一般地说, 邻居用户的确定有两种方法, (1) 根据预先确定的相似度阈值, 选择相似度大于该阈值的用户作为该用户的邻居; (2) 根据预先确定的邻居数  $N$ , 选择相似度最大的前  $N$  个用户作为邻居用户即 TOP- $N$  方法.

### 2.4 评分值的预测

预测方法主要采用如下几种统计的方法:

#### (1) 简单平均值

最简单的就是平均所有邻居用户的对预测服务的评分值.

#### (2) 加权平均值

考虑到当前用户及其邻居用户之间相似性, 对于与当前用户更为相似的用户在预测中给予更大的推荐权重, 否则给予较小的权重.

$$P_{ui} = R_u + \frac{\sum_{p \in SU} (R_p(i) - R_p) \times sim(u, p)}{\sum_{p \in SU} |sim(u, p)|} \quad (4)$$

其中,  $u$  为目标用户,  $SU$  为最近邻居集合,  $p$  为最近邻居集合中的任一用户.

#### (3) 中心加权平均值

其为考虑到不同用户的评分值, 加权评分值与该用户平均评分值间的差值, 而得到的统计方法. 这种预测评分值的方法也很少使用.

## 3 基于上下文、信任网络和协作过滤算法的服务选择机制

### 3.1 基于上下文、信任网络和协作过滤算法的服务选择框架

针对目前还没有成熟的针对移动社交网络特点的服务选择模型, 本文提出了基于上下文、信任网络和协作过滤算法的服务选择机制. 模型包括服务注册中心、服务提供者、服务用户、信任评价、服务选择

等模块, 如图 1 所示.

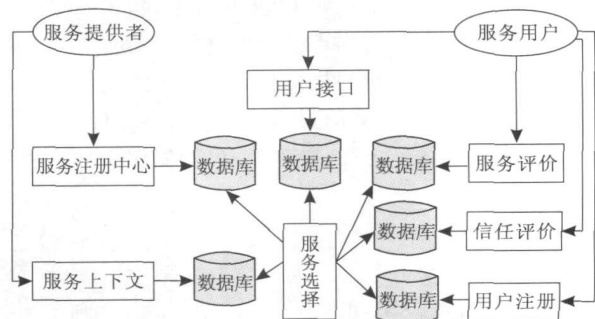


图 1 基于上下文、信任网络和协作过滤算法的服务选择机制

各实体的功能描述如下:

**服务注册中心.** 服务注册中心的角色类似于 Web Service 中的 UDDI, 只不过针对的是一般意义上的服务, 服务提供者可以通过服务注册中心去注册其所提供的服务, 提供维护服务、查询服务的功能, 这是系统的关键部分.

**用户注册.** 用户注册是服务选择系统的数据来源, 只有注册用户才能对使用过的服务进行评价, 从而形成评价数据. 也是个性化推荐的基础.

**服务上下文.** 存储了多维(本文用的是三维模型)推荐的上下文信息, 其中包括 QoS 服务质量非功能性的属性信息, 例如服务的可靠性、稳定性、安全性和响应时间等等. 服务上下文是和服务注册中心紧密配合的, 当服务提供者在服务注册中心注册其服务时服务上下文会主动分配其上下文属性, 当然, 服务提供者可以注册新的上下文属性在服务上下文模块中.

**服务评价.** 服务评价存储了所有用户的对服务的评价信息, 包括用户对服务的一些偏好, 这也是该系统的关键部分.

**信任评价.** 信任评价存储了所有用户的对服务的信任关系信息, 这也是该系统的关键部分.

**服务选择.** 通过降维的方法, 运用协作过滤算法, 进行服务选择.

### 3.2 协作过滤与上下文的结合

协作过滤算法往往被应用在电子商务的环境中来为用户推荐商品信息, 在此应用环境中没有对用户上下文进行定义, 而且商品信息的内容相对固定, 用户的推荐不会随环境的变化改变. 而在移动社交网络服务环境中, 上下文环境是动态变化的, 用户的消费行为也会因环境中上下文信息的变化而发生改变. 在普适计算环境中用户对于服务的评价往往是以隐式的方式做出的. 当用户对于事物做出评价时, 如果其所处的上下文环境同其他用户所处的上下文

环境具有很强的相似性, 则相似上下文中其他用户做出的评价对当前用户的评价的影响度相对较高. 为了能有效地找出和当前上下文相关联的评价, 我们需要对上下文之间的相似性进行度量.

首先, 给出上下文信息形式化的定义. 对于系统中  $n$  种不同种类的上下文信息, 我们定义如下的向量来予以表示:

$$C=(C_1, C_2, C_3, \cdots, C_n) \quad (5)$$

向量中的分量  $C_t(t=1, 2, \cdots, n)$  表示一种类型的上下文, 例如温度、时间或位置信息等. 系统所采集的上下文信息都可以用这里定义的向量来表示. 假设  $X, Y$  属于  $C$ , 那么我们可以用  $sim_t(X, Y)$  来表示  $X, Y$  在上下文类型  $t$  上的相似度. 对于同一种类型的上下文信息, 可以用上下文信息之间的关系来衡量它们之间的相似程度.

我们采用 Pearson 相关系数来计算上下文相似度. 假设上下文信息  $X$  和  $Y$ , 则上下文信息  $X$  和  $Y$  在关于服务  $s$  的相似度可以通过以下公式得到.

$$rel_t(X, Y, s)=\frac{\sum_{u=1}^n(r_{usx_t}-\bar{r}_s)(r_{usy_t}-\bar{r}_s)}{\sigma_{x_t}\sigma_{y_t}} \quad (6)$$

但是, 该方法的使用具有一定的限制. 只有当用户对于服务的评价不会因上下文的不同而发生巨大变化时, 该方法才适用.

上下文引入到传统协作过滤算法中, 原来的用户-服务矩阵得到了扩充, 成为用户-服务-上下文三维协作过滤模型, 使服务选择更加复杂, 评分矩阵更加稀疏, 当然推荐更加准确, 如图 2 所示. 对于引入上下文相似度后三维协作过滤推荐模型处理方法是采用两种方式: 一种是将上下文的信息建立在初始用户模型中, 通过用户模型的初始向量计算用户的相似度, 也就是首先基于上下文选择相同兴趣的用户, 将三维模型降为二维, 然后再根据传统协作过滤

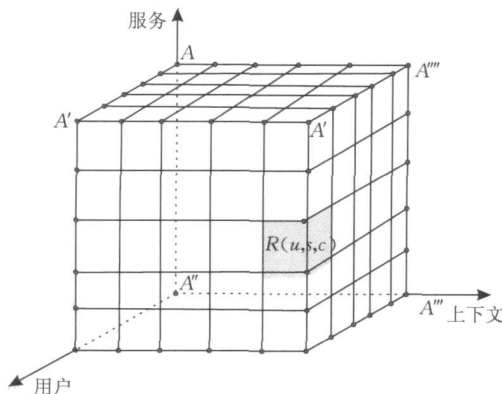
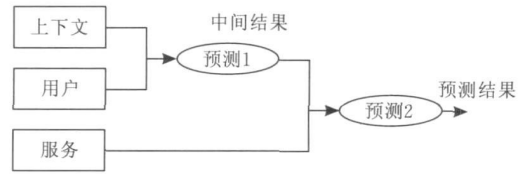
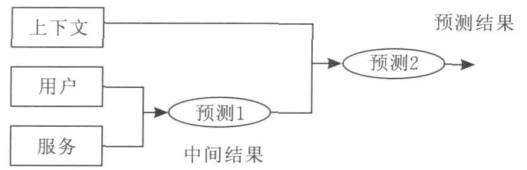


图 2 用户-服务-上下文三维矩阵模型

算法进行推荐, 如图 3(a); 另一种是为用户上下文单独建立模型, 将它看作是基于用户-服务之上的另外一维, 建立三维协作过滤模型. 也就是先按照传统的协作过滤算法计算相似度, 生成用户最近邻居, 再采用上下文信息进行推荐, 如图 3(b). 考虑第二种方法的简单性和可行性, 本文采用后者来进行上下文服务推荐.



(a) 三维矩阵模型降维处理方法之一



(b) 三维矩阵模型降维处理方法之二

图 3

### 3.3 基于上下文、信任网络和协作过滤算法的服务选择方法

#### 3.3.1 基于上下文、信任网络和协作过滤算法的服务选择总体流程

本文提出了基于信任网络和协作过滤的服务选择方法, 将用户相似度和信任度结合起来, 并引入上下文的相似度, 其基本处理流程如图 4.

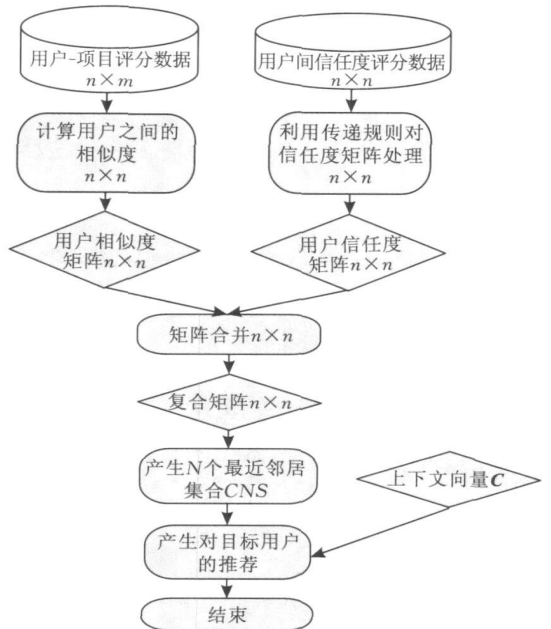


图 4 基于上下文、信任网络和协作过滤算法的服务选择算法处理流程

其中数据库部分表示用户的评分数据, 矩形部分表示该方法的推荐基本流程, 平行四边形表示各种推荐过程中的中间结果矩阵或者向量。

### 3.3.2 基于上下文、信任网络和协作过滤算法的服务选择具体步骤

根据基于信任网络和协作过滤的服务选择算法流程, 该服务选择方法的具体步骤如下。

#### 1) 建立用户模型 (User Profile)。

用户模型是推荐算法的输入部分, 它包括用户对服务的评分值以及用户对用户的信任值。这两种用户的评分数据可以分别用一个矩阵来表示, 矩阵的每一个元素表示用户对某个服务的评分值, 或者是用户对其他用户的信任度的评分值。这一部分主要有以下两方面的工作。

首先, 将用户的隐式评分信息转换成显式的评分数据。用户对服务的评分有两种形式: 显式的评分和隐式的评分, 对于显式的评分数据可以在计算中直接运用, 而隐式的评分, 如非常好、好、较好、一般等, 则必须根据关键词将其转换成显式评分。

其次, 另一种形式的评分——信任评分也可以表示成一个矩阵。矩阵中的每一个元素表示一对用户的信任度值。由于在该模型中将相似度与信任度合并为一个复合值, 所以信任度矩阵的初始化过程中必须考虑到相似度与信任度的取值范围的一致性。传统协作过滤算法利用相关相似性度量方法计算得到的相似度范围是 $[-1, 1]$ , 所以信任评分值也应满足该范围。信任评分为正数时表示信任的程度, 为负数时则表示不信任的程度。本文的实验中所选择的信任评分值取值范围是 $[0, 1]$ , 其中, 1 表示完全信任, 0 表示不信任, 而没有考虑信任度为负的情况, 这也是本文未来工作的一个方向。因此与 Pearson 相关性系数相符合, 保证了下文信任度与相似度相结合的可行性。

#### 2) 计算用户相似度和用户信任度。

(1) 计算用户相似度。对于系统中所有的用户, 分别计算其与其他用户间的相似度。用户的相似度计算方法很多, 考虑到相关相似性度量方法较高的准确性和简单性, 本文所选用的计算方法是相关相似性度量方法。将计算得到的相似度用一个矩阵来表示, 矩阵中的每个元素值表示两个用户间的相似度值。由于在评分数据稀疏情况下, 某些用户间无法计算相关相似性, 因此, 所得到的相似度矩阵和用户评分数据一样, 也是非常稀疏的。

(2) 计算用户信任度。初始化得到的信任度矩

阵非常稀疏, 许多用户间还不存在信任关系, 也就没有信任度。因此, 本文中利用信任关系的传递特性, 根据一定的信任度传递规则, 计算用户间的间接信任度, 使原来稀疏的信任度矩阵变得相对稠密。

一种计算用户信任度的方法是:

① 若信任网络中节点  $A$  与节点  $B$  之间存在唯一路径  $D(A, N_1, N_2, \dots, N_i, B)$ , 则节点  $A$  对节点  $B$  的间接信任度为在信任网络中节点  $A$  到节点  $B$  所有直接信任度的最小值。

② 若信任网络中节点  $A$  与节点  $B$  之间存在多条路径  $(D_1, D_2, \dots, D_i)$ , 则节点  $A$  对节点  $B$  的间接信任度值为在信任网络中  $A$  到  $B$  经过每一条路径所得到的间接信任度的平均值。

上述方法的优点是在计算  $A, B$  两点之间信任值时, 使用了包括  $A, B$  间直接信任和能使  $A, B$  建立起信任联系的所有它点的推荐信任, 参考了所有因素, 是对  $A, B$  间信任关系的最全面评价, 结果也是最可信的; 缺点是该方法过于复杂, 在计算  $A, B$  间推荐信任关系时, 要查找图中  $A, B$  两点间所有的路径, 特别是对于复杂的信任关系图, 这种方法是难以接受的。

本文采用了一种改进的方法如下:

#### ① 计算直接信任值 $dt(A, B)$ 。

查找信任关系图, 看  $A$  点与  $B$  点是否有直线相连。

if(有), then 直接信任值  $dt(A, B) = \text{权值 } w(A, B)$ ;

if(无), then 直接信任值  $dt(A, B) = 0$ 。

#### ② 计算推荐信任值 $ct(A, B)$ 。

查找信任关系图中这样的路径, 路径的起点是  $A$ , 路径的终点是  $B$ , 对于该路径的子路径  $A'B'$ ,  $A'B'$  两点是相连的点, 并且对于任何非  $B'$  的与  $A'$  相连的点  $X$ , 存在边  $A'X$ , 边  $A'B'$  的权值  $w(A', B') > w(A', X)$ 。找到这样的路径  $AB, U$  是路径上所有点的集合, 有

$$ct(A, B) = \prod_{X, Y \in U} w(X, Y) \quad (7)$$

#### ③ 求信任值 $tr(A, B)$ 。

$$tr(A, B) = \alpha \times dt(A, B) + (1 - \alpha) \times ct(A, B) \quad (8)$$

$\alpha$  是参数, 是一个取值范围为  $0 \sim 1$  的实数, 表示直接信任值  $dt$  在信任值  $tr$  中所占的比重。选择与  $A$  点直接信任值最大的点  $X$ ,  $X$  点是  $A$  最信任(直接信任)的点, 该点给出的推荐信任值要比其它的点更可信; 然后再找与  $X$  点直接信任值最大的点  $Y$ , 以此类推, 直到找到  $B$  点。根据现实中的经验, 为了确

定一个没有直接接触过的陌生人是否值得信任, 我们总是选择最令自己信任的人所作的对陌生人的评价, 无论我们最信任的人对陌生人所作的评价是高还是低. 该算法是对以上算法的改进, 显然要比上面算法简单, 对于复杂的系统, 改进后的算法更有实用价值.

### (3) 合并矩阵.

由于相似度矩阵与信任度矩阵都非常稀疏, 许多用户间无法计算相似度, 同时也不存在信任关系. 另外, 信任度也可以作为一个可选的推荐权重. 因此, 本文中考考虑将相似度矩阵与信任度矩阵进行合并, 将相似度和信任度综合进行考虑, 得到相似度与信任度的一个综合取值, 并用它作为最终的推荐权重. 结合相似度与信任度可以更加精确的实现推荐.

假设  $Similarity(i, j)$  为用户  $i$  与用户  $j$  的相似度,  $Trust(i, j)$  为用户  $i$  与用户  $j$  的信任度,  $Weight(i, j)$  为合并后相似度与信任度的一个综合值. 若  $Similarity(i, j)$  用  $A$  表示,  $Trust(i, j)$  用  $B$  表示, 则

$$Weight(i, j) = \begin{cases} \frac{2 \times A \times B}{A + B}, & A > 0 \text{ and } B > 0; \\ B, & A = 0 \text{ and } B > 0; \\ 0, & B = 0 \end{cases} \quad (9)$$

### 3) 生成邻居集合.

邻居是向目标用户推荐的用户集合, 是与目标用户兴趣最近的用户, 邻居集合的选取与最终的推荐结果有非常紧密的关系. 传统的协作过滤推荐算法中, 邻居的选取通常有两种策略, 即与目标用户相似度最高的  $N$  个用户; 与目标用户的相似度大于某个固定值的用户集合.

在本文的推荐模型中, 将相似度与信任度相结合, 得到一个复合值, 因此, 在该模型中, 最近邻居的选取的参照因素为相似度与信任度的复合值, 即在复合矩阵中与目标用户的复合值最高的  $N$  个用户作为其邻居集合, 或者是复合值大于某个固定值的用户集合, 我们将得到最近邻居集  $CNS$  (Closest Neighbor Set).

### 4) 产生推荐.

接下来我们要预测用户对于任意项的兴趣度. 设活动用户  $a$  对于服务  $s$  在上下文为  $C$  时的评价用  $P_{a, s, c}$  来表示.

$$P_{a, s, c} = \bar{r}_a + k \sum_{u \in CNS_a} (R_{u, s, c} - \bar{r}_u) \times Weight(a, u) \quad (10)$$

其中  $k$  为线性回归系数,  $k = \frac{1}{\sum_{u \in CNS_a} (|Weight(a, u)|)}$ ,  $\bar{r}_a$  为用户  $a$  的平均评价,  $\bar{r}_u$  为最近邻集合  $CNS_a$  中用户  $u$  的平均评价,  $Weight(a, u)$  则表示用户  $u$  对用户  $a$  的推荐权重, 我们采用式 (5) 来计算.

和传统的计算方法相比, 该方法只需要计算当前用户和最近邻集合中的其他用户的相似度, 而不需要计算当前用户和所有其他用户的相似度. 在上下文感知的协作过滤系统中, 每个评价都有一个与之相关的上下文, 表示评价做出时所处的上下文. 当前用户所处的上下文和其他用户在做出评价时所处的上下文之间的相似性决定了评价之间的相关程度. 定义  $R_{u, s, c}$  为用户  $u$  在上下文环境为  $C$  时对服务  $s$  做出的评价关于上下文相似度的加权平均.

$$R_{u, s, c} = k \sum_{x \in C} \sum_{t=1}^z r_{u, s, x} \times sim_t(c, x) \quad (11)$$

其中  $k = \frac{1}{\sum_{x \in C} \sum_{t=1}^z |sim_t(c, x)|}$ ,  $sim_t(c, x)$  则表示上下文  $c$  和  $x$  在上下文类型  $t$  上的相似程度. 上下文信息  $C$  是由多个上下文分量组成, 上下文之间的相似度则由  $R_{u, s, c}$  体现出来.

## 4 仿真实验及结果分析

### 4.1 实验场景

为了验证提出的基于上下文、信任网络和协作过滤算法的移动社交网络服务选择机制, 我们设计了一个用户兴趣点 (如旅游景点、餐馆、酒吧等) 服务选择的应用场景. 移动社交网络系统用户到达某地, 通过移动社交网络, 根据好友的推荐, 依据好友的信任度的不同, 而且按照自己的兴趣偏好的不同, 搜索出与自己相似度比较接近的邻居集合, 结合用户当前的实际位置等上下文信息, 从而选择出最适合自己的兴趣点. 因此, 兴趣点服务的选择既参考移动社交网络中好友相似度和信任度, 同时又结合了用户当前的位置等上下文信息, 能够比较好地验证本文提出的基于上下文、信任网络和协作过滤算法的移动社交网络服务选择算法.

### 4.2 实验结果评价准则

本文采用绝对平均偏差  $MAE$  来衡量服务选择算法的准确率.  $MAE$  的定义为

$$MAE = \frac{\sum_{i=1}^N |p_i - q_i|}{N} \quad (12)$$

其中,  $N$  表示推荐服务的个数, 推荐的用户评价为  $p_i$ , 而用户的实际评价为  $q_i$ .  $MAE$  体现了算法预测值和用户实际评分间的差异.  $MAE$  值越小就表明服务选择算法的准确度越高.

4.3 实验环境

从实验环境来说, 作者所在实验室已经搭建了一个基于手机通讯录的移动社交网络服务系统, 新型手机通讯录具有即时通信功能和社交服务能力. 因此, 基于手机通讯录形成了一个天然的社会性关系网络. 并且基于该网络用户可以方便地实现信息的分享和服务的使用. 用户的位置信息可以通过手机客户端软件来获取到, 通过调用智能手机的 GPS 功能来实现. 在服务器端将一个个的手机通讯录进行关联, 从而形成了一个社会性关系网络, 并且基于该网络用户可以进行兴趣点的评价和选择. 这样就解决了文章中所提出的移动社交网络服务的必备的实验环境.



图 5 基于手机通讯录的移动社交网络服务原型系统

系统硬件配置如下:

CPU: IntelPentium 2 8GHz;

内存: DDR 4GB;

软件环境如下:

操作系统: Window s Server 2003;

数据库: MySQL Server 5 0;

智能手机: Nokia N97 mini;

手机操作系统: Symbian S60 5 0;

实验环境编程语言: Java, C++, Qt;

数据分析处理语言: Matlab 7.

4.4 实验结果及分析

从实验室数据来说, 本文从作者实验室内建立的移动社交网络实验系统平台中采集数据进行了模拟仿真实验, 包括 1000 个兴趣点, 其中对 800 个兴趣点进行了评分, 15 个上下文, 150 个系统用户, 服务评价信息采集及呈现系统如图 6 所示. 数据集分成两部分, 一部分是训练集, 另一部分是测试集, 以

$F$  表示两组数据的分割比例, 例如  $F=0.8$  表示 80% 为训练集, 20% 为测试集.

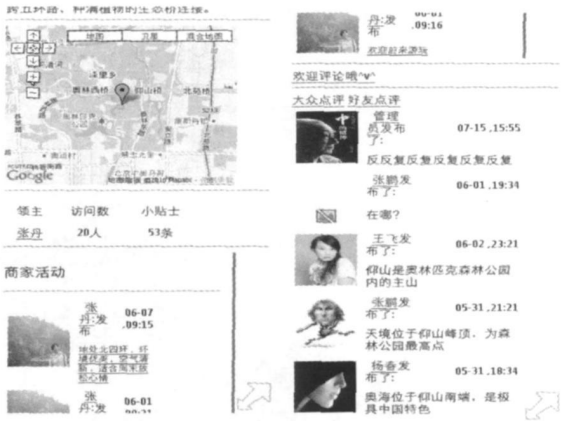


图 6 移动社交网络服务评价信息

实验 1. 测试 3 个相似性度量方法的准确性: 在图 7 中, 取  $F=0.8$ ,  $N$  为 [25, 150], 在图 8 中,  $N=150$ , 而  $F$  值为 [0.5, 0.9], 从中可以看出调整的余弦相似性度量方法更准确. 此外,  $F=0.8$  时准确率最高.

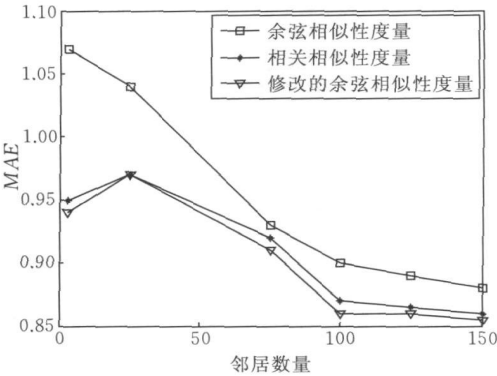


图 7 3 种相似性度量方法比较

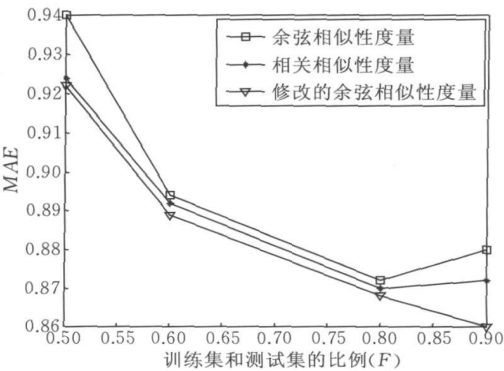


图 8 训练集和测试集比例准确率比较

实验 2. 传统的协作过滤算法和信任关系的推荐准确性的比较, 如图 9, 可以看出加入信任关系后推荐准确性得到了提高, 说明用户的信任度也是



服务选择中需要考虑的重要因素。

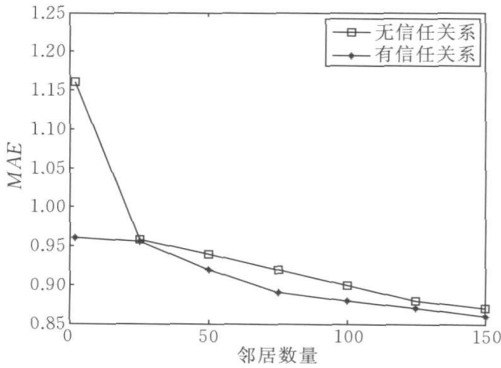


图 9 信任关系对预测准确性的影响

实验 3. 训练集和测试集比例不同的推荐准确性不同, 训练集容量较大的情况下能得到更高的推荐准确性, 如图 10.

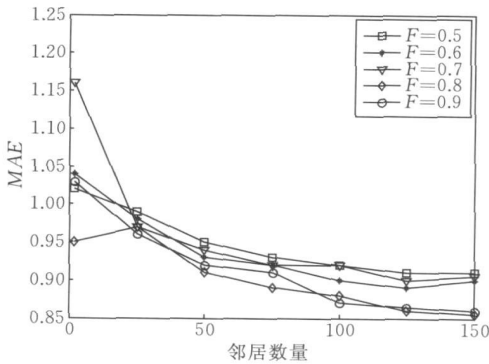


图 10 不同训练集和测试集比例的准确度

实验 4. 基于传统协作过滤的服务选择算法 TCF、基于上下文的协作过滤算法 CCF、在上下文环境下基于信任网络和协作过滤算法(TNCF)的推荐数据的准确性. 用户的评价数据量分别为 200、400、600 和 800 时 MAE 的值, 如图 11.

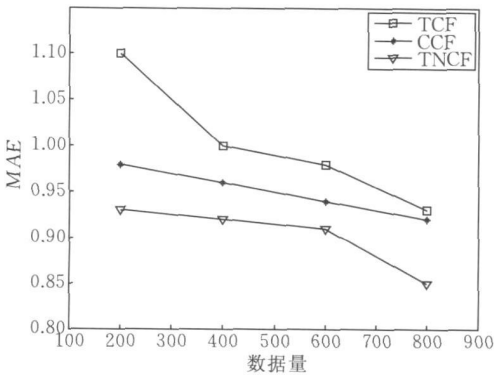


图 11 MAE 值随数据量的变化曲线

从图中可以看出, TCF 的 MAE 值不仅在数据稀疏时偏高, 而且在数据较稠密时也无法达到理想

的值. 这其中主要的原因是, 传统的协作过滤技术没有考虑到环境中上下文的相似度. 而 CCF 在数据稀疏的情况下, MAE 值很高, 由于加入了对上下文的考虑, 进一步稀释了用户-评价矩阵. 使得算法在启动阶段很难找到类似的用户, 因此造成了在数据稀疏条件下算法准确性的下降. TNCF 在启动时和 TCF 以及 CCF 相比具有较好的性能, 由此可见, 该算法有效地改进了在数据稀疏时推荐的准确性.

## 5 总结与展望

本文提出了基于上下文、信任网络和协作过滤算法的移动社交网络服务选择机制, 将用户相似度和信任度结合起来, 提高了服务选择的准确性和可靠性. 引入上下文相似度构成用户-服务-上下文三维协作过滤模型来进行服务选择, 并详细描述了该方法的基本流程. 最后通过仿真实验验证了此方法的服务选择准确率比传统协作过滤算法更高.

另外, 用户上下文信息还包括用户的社会上下文, 特别是用户本身的社会属性, 例如年龄、性别、职业等, 也是将来服务选择机制中需要考虑的重要因素, 此外, 本文考虑用户的信任度取值区间仅限于  $[0, 1]$  之内, 在  $[-1, 1]$  区间内服务选择的研究还需要进一步探讨.

## 参 考 文 献

- [1] Hdanova A V, Davies M, Jorns O et al. MOSS: Mobile social spaces//Proceedings of the 1st International Workshop on Blending Physical and Digital Spaces on the Internet (OneSpace2008) at the 1st Future Internet Symposium (FIS'08). Vienna, Austria, 2008; 94-100
- [2] Cui Yazhou, Duan Gang. Collaborative filtering recommendation system based on Web log and commodities classification. Journal of UESTC, Chengdu, 2006, 8(3): 39-42
- [3] Cheng Guangyao. Research on the recommending method used in C2C online trading//Proceedings of the IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology-Workshops. Silicon Valley, California, USA, 2007; 103-106
- [4] Maximilien E M, Singh M P. Multiagent system for dynamic Web services selection//Proceedings of the Workshop on Service-Oriented Computing and Agent-Based Engineering (SOCABE 2005). Utrecht, The Netherlands, 2005; 25-29
- [5] Casati F, Castellanos M, Dayal U et al. Probabilistic context-sensitive and goal-oriented service selection//Proceedings of the 2nd International Conference on Service Oriented Computing (ICSOC04). New York, 2004; 316-321

- [ 6 ] Kalepu S, Krishnaswamy S, Verity S L. A QoS metric for selecting Web services and providers//Proceedings of the 4th International Conference on Web Information Systems Engineering Workshops(WISE W'03). Roma, Italy, 2003: 131-139
- [ 7 ] Sreenath R M, Singh M P. Agent-based service selection. *Journal of Web Semantics*, 2004, (3): 261-279
- [ 8 ] Lin Tao. Service selection in ubiquitous computing[ Ph. D. dissertation]. Northeastern University, Shenyang, 2004(in Chinese)  
(林涛. 普适计算中的服务选择问题[ 博士学位论文]. 东北大学, 沈阳, 2004)
- [ 9 ] Huang Run-Cai, Zhuang Yi-Wen, Zhou Ji-Liang. Semantic Web-based context-aware service selection in task-computing. *Journal of Computer Applications*, 2009, 29(3): 892-895(in Chinese)  
(黄润才, 庄怡雯, 周集良. 任务计算中基于语义 Web 的上下文感知服务选择. *计算机应用*, 2009, 29(3): 892-895)
- [ 10 ] Wang X, Vitvar T, Kerrigan M et al. A QoS-aware selection model for semantic Web services//Proceedings of the Service-Oriented Computing (ICSOC 2006). Chicago, USA, 2006: 390-401
- [ 11 ] Tsoumetsis D, Roussaki I, Sykas E. QoS-aware service evaluation and selection. *European Journal of Operational Research*, 2008, 191(3): 1101-1112
- [ 12 ] Ma Y, Zhang C. Quick convergence of genetic algorithm for QoS-driven Web service selection. *Computer Networks*, 2008, 52: 1093-1104
- [ 13 ] Resnick P, Iacovou N, Suehak M et al. Grouplens: An open architecture for collaborative filtering of netnews//Proceedings of the ACM Conference on Computer Supported Cooperative Work. North Carolina, USA, 1994: 175-186
- [ 14 ] John S B, David H, Carl K. Empirical analysis of predictive algorithms for collaborative filtering//Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence. Madison, WI, 1998: 43-52



**WANG Yu-Xiang** born in 1971, Ph. D.. His main research interests include network management and communication software.

**QIAO Xiu-Quan** born in 1978, associate professor. His

main research interests include the intelligent theory and technology of network services.

**LI Xiao-Feng** born in 1950 professor, Ph. D. supervisor. Her main research interests include intelligent network and communication software.

**MENG Luo-Ming** born in 1955 professor, Ph. D. supervisor. His main research interests include network management and communication software.

## Background

The context-awareness service provisioning is the pursuing goal of pervasive computing. There are a lot of researches on context-awareness service providing in different fields such as mobile network, cognitive network and Web service. However, the acquisition of real context information is always a problem for most of researchers. The mobile social networking services seamlessly integrate mobile computing with social computing based on smart mobile phone, and are able to provide rich context information. So the context-awareness service providing has become a hot topic in mobile SNS, especially location-based service.

This group's research interests are currently focused on the context-awareness mobile SNS provisioning. The long-

term goal is to build a user-centric intelligent convergent service environment, which can seamlessly integrate the traditional telecommunication services, instant messaging service, social networking services and personal information management service. This work is supported by National Natural Science Foundation of China under grant Nos. 60802034 and 60672122, Specialized Research Fund for the Doctoral Program of Higher Education under grant No. 20070013026, Beijing Nova Program under grant No. 2008B50, and "New generation broadband wireless mobile communication network" Key Projects for Science and Technology Development under grant No. 2011ZX03002-002-01.