

## 基于社会网络面向个性化需求的可信服务推荐

张佩云<sup>1,2</sup>, 陈恩红<sup>2</sup>, 黄波<sup>3</sup>

- (1. 安徽师范大学 数学计算机科学学院, 安徽 芜湖 241003;  
2. 中国科学技术大学 计算机科学与技术学院, 安徽 合肥 230026;  
3. 南京理工大学 计算机科学与技术学院, 江苏 南京 210094)

**摘 要:** 在当前服务计算和社会计算背景下, 针对难以获取满足用户个性化需求的可信 Web 服务问题, 给出基于社会网络面向个性化需求的可信 Web 服务推荐模型; 设计用户个性化功能需求分解与匹配算法, 并利用语义词典提高功能需求语义匹配的准确性; 基于个性化功能需求、社会网络节点信任度及服务信任度, 设计了一种满足用户个性化需求的可信服务推荐算法, 通过对社会网络节点之间、节点与服务之间的信任相关性进行分析, 提高服务协同可信推荐性能。算法分析及实验结果表明该方法是有效和可行的。

**关键词:** 社会网络; 个性化需求; 信任度; 可信服务推荐

中图分类号: TP311

文献标识码: A

文章编号: 1000-436X(2013)12-0049-11

## Personalized requirements oriented trustworthy services recommendation based on social network

ZHANG Pei-yun<sup>1,2</sup>, CHEN En-hong<sup>2</sup>, HUANG Bo<sup>3</sup>

- (1. School of Mathematics and Computer Science, Anhui Normal University, Wuhu 241003, China;  
2. School of Computer Science and Technology, University of Science and Technology of China, Hefei 230026, China;  
3. School of Computer Science & Technology, Nanjing University of Science & Technology, Nanjing 210094, China)

**Abstract:** In current background of service computing and society computing, it is difficult to obtain the trustworthy Web services to meet users' personalized requirements. To solve the problem, the trustworthy services recommendation model was proposed which orients personalized requirements based on social network. The decomposing and matching algorithm was proposed for users' personalized requirements, and the algorithm utilizes the semantic dictionary to improve the semantic match of users' requirements. Based on personalized functional requirements and social trust values of network nodes and the services' direct trust values and indirect trust values, a trustworthy service recommendation algorithm was designed which meets the personalized requirements. By analyzing the correlation among the social network nodes and services' trust values, the method can improve the performance of trustworthy services collaborative recommendation. The analysis of algorithms and experiments' results show that the approach is effective and feasible.

**Key words:** social network; personalized requirement; trust values; trustworthy services recommendation

### 1 引言

当前, 面向服务的体系结构 (service-oriented architecture) 、软件即服务(SaaS)、云计算等新的软

件架构思想和运营理念<sup>[1]</sup>得到不断发展, 使得服务正日益成为互联网中重要的计算资源和软件资产。随着社会计算技术的快速发展, 社会网络已成为用户获取服务和交互信息的重要平台, 由于社会网络

收稿日期: 2013-03-27; 修回日期: 2013-09-13

基金项目: 国家自然科学基金资助项目(61201252, 61203173); 安徽省自然科学基金资助项目(1308085MF100); 中国博士后科学基金资助项目(2013M531528); 安徽省高校省级自然科学基金重点基金资助项目(KJ2011A128); 安徽省科技厅软科学研究计划基金资助项目(11020503009)

**Foundation Items:** The National Natural Science Foundation of China(61201252, 61203173); The Natural Science Foundation of Anhui Province(1308085MF100); The Postdoctoral Foundation of China (2013M531528); The Natural Science Research Key Project of Anhui Provincial Higher Education (KJ2011A128); The Soft Science Project of Anhui Province (11020503009)

的虚拟性和动态性,当服务的发现和运行从传统的相对集中、封闭和静态环境变迁到开放、多变、动态的社会网络环境中时,服务可能面临着不可用、不可靠等信任风险,导致基于社会网络获取的服务往往难以满足用户的需求。因此,如何快速获取满足用户个性化需求的可信服务日益成为当前社会计算领域中的重要研究内容。

推荐系统作为个性化服务研究领域的重要分支,可以通过挖掘用户与服务之间的二元关系向用户推荐满足其个性化需求的服务以及 Web 信息、在线商品等<sup>[2]</sup>。在推荐算法方面,主要研究有协同推荐、基于内容的推荐、聚类技术、关联规则推荐等<sup>[3]</sup>。Bobadilla<sup>[4]</sup>提出了基于用户上下文信息与相似度计算的协作过滤推荐方法,Walter<sup>[5]</sup>给出了基于知识的推荐系统,指出推荐系统中需要对用户需求引入语义技术,以提高推荐的质量,Tang<sup>[6]</sup>提出了基于位置感知的协同过滤方法,通过整合用户和服务的位置向用户推荐服务。在基于社会网络的可信推荐方面,相关研究有:Golbeck<sup>[7]</sup>提出了社会网络中一种基于平均信任值的信任推理机制,评估社会网络中 2 个不相邻的参与者之间的信任值,Liu<sup>[8]</sup>提出了一种启发式算法以发现参与者之间可信的社会路径,文献[9]研究了社会网络中上下文感知的信任问题,Yao<sup>[10]</sup>提出了一种基于声誉的方法来推荐可信服务,Zhong<sup>[11]</sup>研究了社会网络中基于个人托管的服务推荐方法,Jamali<sup>[12]</sup>提出了一种社会网络中基于内存的可信推荐方法,Massa<sup>[13]</sup>研究了基于协作过滤的可信服务推荐问题等。前人在社会网络及推荐方面做了许多有益的研究工作,对本文的研究有一定的启发,但社会网络服务环境面临着复杂多变性,仅基于用户—服务的二元信任关系难以实现可信服务推荐,因此,需要根据社会网络的动态监测以及服务质量的变化,并结合社会网络节点—节点—服务三元信任关系,来推荐满足用户个性化需求的服务。本文结合社会网络节点—节点—服务三元信任关系设计可信服务推荐算法。

## 2 基于社会网络面向个性化需求的可信服务推荐模型

### 2.1 相关定义与概念

社会网络成员节点之间因为交互和联系而形成相对稳定的成员间信任关系体系,而社会节点由于调用服务使得社会网络与服务之间形成了社会

网络服务信任关系,其定义如下。

**定义1** 社会网络服务信任关系:用有向图  $G=(V,S,E,F)$  表示,其中, $V$ 是社会网络节点集, $V=\{v_i|i \text{ 为正整数}\}$ , $E$ 是社会网络节点之间信任关系的集合, $E=\{(v_i,v_j)|v_i \in V \wedge v_j \in V\}$ , $(v_i,v_j)$ 表示社会网络节点  $v_i$  和节点  $v_j$  之间存在一条从  $v_i$  指向  $v_j$  的有向边,该有向边表示节点  $v_i$  对节点  $v_j$  有信任关系,其信任值记为  $v_{ij}$ ;用  $w_{ij}$  表示两节点交易的权重, $w_{ij}=\langle \alpha_{ij}, \beta_{ij} \rangle$ ,其中, $\alpha_{ij}$  与  $\beta_{ij}$  分别表示社会网络节点  $v_i$  对节点  $v_j$  成功交易的次数及总次数。 $S$ 是服务集, $S=\{s_k|k \text{ 为正整数}\}$ 。 $F$ 是社会网络节点与服务之间的交易关系集合, $F=\{(v_i,s_k)|v_i \in V \wedge s_k \in S\}$ ,节点对服务的直接信任值记为  $f_{ik}$ ,用  $q_{ik}$  表示节点与服务交易的权重, $q_{ik}=\langle a_{ik}, b_{ik} \rangle$ ,其中, $a_{ik}$  与  $b_{ik}$  分别表示社会网络节点  $v_i$  与服务成功交易次数及总次数。社会网络服务信任关系如图1所示。

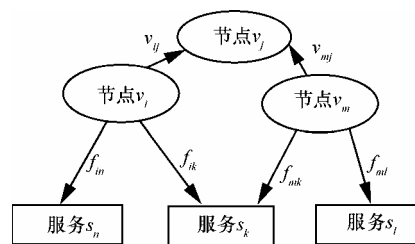


图1 社会网络服务信任关系

图1中,椭圆表示社会网络节点(如  $v_i$  和  $v_j$ , 简称  $v$ );方框表示社会网络节点交易过的服务; $v_{ij}$ 表示社会网络节点  $v_i$  对节点  $v_j$  的信任度( $v_{mj}$  含义类似),信任具有有向性,箭头线指明了信任的方向; $f_{ik}$ 表示社会网络节点  $v_i$  对服务  $s_k$  的直接信任度( $f_{in}$ 、 $f_{mk}$ 、 $f_{ml}$  含义类似)。社会网络服务关系具有非对称性。图1中,社会网络节点定义如下。

**定义2** 社会网络节点  $v$ :  $v=(ID, IP, NList, Status, SList)$ ,其中, $ID$ 表示社会网络节点  $v$  的标识符, $IP$ 表示节点  $v$  的主机IP地址, $NList$ 表示节点  $v$  的邻居集合, $SList$ 表示节点  $v$  调用过的服务集合。其中, $NList=(NID, NIP, v_{ij})$ , $NID$ 表示节点  $v$  的邻居节点标识符, $NIP$ 表示邻居节点主机IP地址, $v_{ij}$ 表示节点间的信任度。 $Status$ 表示是否允许该节点向社会网络推荐服务,取值为0或1,分别表示不允许和允许推荐。 $SList=(SID, URI, Des, f_{ik})$ , $SID$ 表示服务标识符, $URI$ 是该服务的资源标识符, $Des$ 为该服务的

功能描述,  $f_{ik}$  表示节点  $v$  对该服务的直接信任度。

定义2中社会网络节点间的信任度  $v_{ij}$ 、直接信任度  $f_{ik}$  含义分别如下。

1) 社会网络节点间的信任度 (记为  $v_{ij}$ ): 该信任度是在两者交易的过程中逐渐建立起来的, 其取值范围为  $[-1, 1]$ 。社会网络节点对自身的信任度为最大值1, 当节点  $v_i$  与节点  $v_j$  之间没有交易记录时,  $v_{ij} = 0$ , 若  $v_{ij} < 0$ , 则表明社会网络节点  $v_i$  对  $v_j$  有过信任, 但现在不再信任。信任度  $v_{ij}$  保存在矩阵  $E$  中。

2) 对服务  $s_k$  的直接信任度 (记为  $f_{ik}$ ): 是指当社会网络节点  $v_i$  直接调用过服务  $s_k$ 、基于前期与该服务的调用经历而获得的直接经验信任度。 $f_{ik}$  取值范围为  $[-1, 1]$ , 其初始值为0.5; 当社会网络节点  $v_i$  成功调用服务  $s_k$ , 则  $f_{ik} = (f_{ik} + 1)/2$ ; 当社会网络节点  $v_i$  调用过服务  $s_k$  失败, 则  $f_{ik} = f_{ik}/2$ 。当  $f_{ik} = 1$  时, 表示  $v_i$  完全信任  $s_k$ ; 当  $f_{ik} = -1$  时, 表示社会网络节点  $v_i$  将不再信任服务  $s_k$  并屏蔽该服务,  $f_{ik}$  保存在矩阵  $F$  中。

基于直接信任度  $f_{ik}$ , 可以计算节点对服务  $s_k$  的间接信任度。

3) 对服务  $s_k$  的间接信任度 (记为  $ndt_{jk}$ ): 如果社会网络节点  $v_j$  对服务  $s_k$  没有直接信任度, 则通过聚合社会网络中其他节点与服务  $s_k$  的直接信任度 (基于加权平均法计算)。基于图1, 社会网络节点  $v_j$  对服务  $s_k$  的间接信任度  $ndt_{jk}$  可以表示为

$$ndt_{jk} = \frac{w_i f_{ik} + w_m f_{mk}}{w_i + w_m} \quad (1)$$

其中, 社会网络节点  $v_j$  将通过社会网络中与  $s_k$  之间有交易的服务使用者的直接信任度  $f_{ik}$ 、 $f_{mk}$  来计算间接信任度  $ndt_{jk}$ ,  $w_i$ 、 $w_m$  分别为  $f_{ik}$  和  $f_{mk}$  的权重, 权重取值范围在 0~1 之间。由于直接信任度  $f_{ik}$ 、 $f_{mk}$  是动态变化的, 因此间接信任度  $ndt_{jk}$  也具有动态变化的特点。

## 2.2 可信服务推荐模型

基于上述概念与定义, 本文提出基于社会网络面向用户个性化需求的可信服务推荐模型, 其中, 用户个性化需求的定义如下。

定义3 个性化需求: 用二元组表示为  $Requirement = (FunctionCategory, QoS[N])$ , 其中,  $FunctionCategory$  是服务的功能需求类别,  $QoS$  表示用户对服务的性能需求约束条件,  $N$  表示  $QoS$  约束的个数, 每个  $QoS$  约束均满足一个三元组:  $QoS = (Attribute, Relation, Value)$ , 其中,  $Attribute$  是  $QoS$  属性名称,  $Relation$  是约束关系,  $Value$  是  $QoS$  属性值。

本文中侧重于面向  $FunctionCategory$  的功能需求及个性化信任需求, 如设置:  $QoS[0] = (trust, >, 0.9)$  表示用户对服务的个性化信任度要求是大于 0.9。

基于定义1~定义3, 可信服务推荐模型如图2所示。

图2所示的模型中相关模块含义如下。

1) 服务发布。服务提供者以广播的形式向社会网络发布服务, 包括服务的描述、QoS及信任信息等。

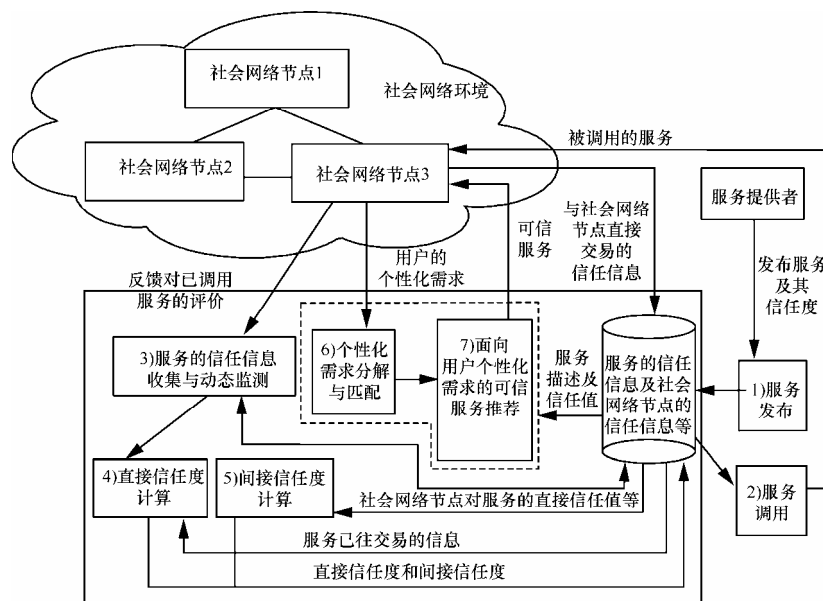


图2 基于社会网络面向个性化需求的可信服务推荐模型

2) 服务调用。调用满足用户对服务功能和性能需求的服务, 根据服务提供者发布的服务功能描述、信任值来直接调用服务或通过其他社会网络节点推荐后调用服务。

3) 服务的信任信息收集与动态监测。社会网络的每个节点维持一个邻居列表, 服务使用者一方面收集基于直接经验的服务交易信息, 并从自身数据库查询已有的服务交易信息, 包括已往交易次数及成功率、服务使用者的评价信息、服务提供者发布的服务信任信息等; 另一方面, 当需要时, 服务使用者通过社会网络向其邻居社会网络节点反馈并广播对已调用服务的评价(如服务的可信性), 而邻居社会网络节点收集该信息, 进行服务的间接信任度计算。

4) 直接信任度。基于服务调用信息的收集与动态监测结果, 形成服务的直接信任度。

5) 间接信任度计算。基于社会网络节点的信任网络关系和服务的直接信任度, 借助聚合函数方法, 计算社会网络节点对服务的间接信任度。

6) 个性化需求分解与匹配。服务使用者通过该模块分析用户的个性化需求, 进行需求分解与需求匹配。具体见第3节的算法1。

7) 满足用户个性化需求的可信服务推荐。面向用户个性化需求, 从数据库查询并分析匹配用户需求的服务功能描述、服务信任度, 通过推荐算法向用户推荐满足其个性化需求的服务。具体见第3节的算法2。

### 3 基于社会网络面向个性化需求的可信服务推荐算法

#### 3.1 基于社会网络面向用户个性化功能需求的分解与匹配算法

为了实现面向个性化需求的可信服务推荐功能, 首先需要对用户的个性化需求进行分解并进行功能需求匹配, 为此, 针对定义3中的 *Function-Category*, 设计面向个性化功能需求分解与匹配算法如算法1所示。

##### 算法1 个性化功能需求的分解与匹配算法

输入: 服务集 *servicesSet*, 用户的个性化功能需求描述 *individualDescription*

输出: 个性化功能需求与服务匹配信息 *services-Subset*

**BEGIN**

```

1)      demandVector=OpenNLP
(individualDescription); /*对用户的个性化需求进行分词*/
demandNounVector=getNoun(demandVector);
/*从用户需求分词向量中获取名词*/
2)      FOR EACH demandNoun[i] in
demandNounVector
3)          {wordnetSynonyms=getSynonyms
(demandNoun[i]); /*语句4是从WordNet中获取
demandNoun[i]的同义词*/
4)              {wordnetHypernyms=getHypernyms(demand
Noun[i]); /*语句5是从WordNet中获取demand
Noun[i]的上位词*/
5)                  demand←wordnetSynonyms & wordnet
Hypernyms & demandNounVector;
6)              }ENDFOR
7)          }ENDFOR
8)      FOR EACH service[i] in servicesSet /*
针对服务集中每个服务service[i]进行处理*/
9)          { desVector=OpenNLP (service[i].
description); /*对每个服务service[i]的功能描述进
行分词*/ service =getNoun(desVector); /*从服务功
能描述的分词向量中获取名词*/
10)         degree=match(demand, service); /*对个性
化需求与服务功能进行语义匹配, 匹配度值放入
degree中*/
11)         servicesSubset←(demand, service, degree);
/*保存满足用户个性化功能需求、servicei及degree
到servicesSubset中*/
12)     }ENDFOR
13) return servicesSubset;
14) END

```

为提高个性化功能需求的语义理解性, 首先利用 *OpenNLP* 对需求进行分词(见算法1中的1))和获取名词(见算法1中的2)), *OpenNLP* 是一个机器学习工具包, 本文用于处理自然语言文本, 包括句子切分、词性标注及名称抽取。其次, 借助词典 *WordNet* 来处理用户需求的语义模糊性, 主要使用 *Synonyms* (同义词关系, 见算法1中的4))及 *Hyponym/Hypernym* (即Is-A关系, 见算法1中的5))来扩充对需求的语义理解, 并将语义扩充理解后的需求保存在 *demand* 中(见算法1中的6))。其次, 针对服务的功能描述(*description*)进行分词和处理(分别见算法1中的8)和算法1中的9)), 最后, 在个性化功能需求分解与处理及服务的描述分解与处理的基

基础上,对个性化功能需求和服务功能描述进行语义匹配计算(见算法1中的10)和算法1中的11),这种匹配为demand和service的最大语义匹配问题。本文利用二分图实现最大匹配,从提高匹配度的角度计算用户需求与服务的功能描述之间的语义匹配度。

**定义 4 二分图的最大匹配** 给定一个二分图  $G=(V, E)$  将二分图  $G$  中的顶点集划分为  $X \cup Y = V(G)$ ,  $X=\{x_1, x_2, \dots, x_k\}$ ,  $Y=\{y_1, y_2, \dots, y_n\}$   $x_k$  与  $y_n$  均属于  $E(G)$ , 若  $M$  包含的边数是  $G$  的所有匹配中包含的边数中最大的,且仅当概念  $x_k$  与  $y_n$  的语义匹配度最大时,则  $M$  称为  $G$  的一个最大匹配。

基于二分图的最大匹配,计算用户个性化功能需求与服务功能描述之间的最大语义匹配,计算公式如式(2)所示。

$$\begin{aligned} match(demand, service) = & \\ & ((\sum_{j=1}^m (getSimilarity(demand_j, service_j)) / k + \\ & (\sum_{j=1}^m getSimilarity(demand_j, service_j)) / n) / 2 \quad (2) \end{aligned}$$

其中,  $m$  表示 demand 与 service 在执行二分图的最大匹配后的语义匹配集大小,  $k$  表示 demand 集的大小,  $n$  表示 service 集的大小,计算 demand<sub>*j*</sub> 与 service<sub>*j*</sub> 的语义匹配度时,本文采用 Java WordNet API 的 getSimilarity() 函数进行概念间的语义相似度计算。

### 3.2 基于社会网络面向个性化需求的可信服务推荐算法

在设计的面向个性化功能需求分解与匹配算法的基础上,本文设计了基于社会网络面向个性化需求的可信服务推荐算法(如算法2所示)。在基于间接信任度推荐服务时,往往存在多个满足用户个性化需求的同类服务,需要采用一定的方法来区分最优服务并推荐。作为一种度量两组变量相关程度的方法,皮尔逊相关系数可以很好地计算出社会网络节点之间的信任度与其推荐服务的直接信任度之间的相关性,在服务推荐时推荐相关性强的最佳服务。面向可信服务推荐的皮尔逊相关系数计算公式如式(3)所示。

$$r_{vf} = \frac{\sum_{j=1}^m (v_{ij} - \bar{v}_j)(f_{jk} - \bar{f}_k)}{\sqrt{\sum_{j=1}^m (v_{ij} - \bar{v}_j)^2} \sqrt{\sum_{j=1}^m (f_{jk} - \bar{f}_k)^2}} \quad (3)$$

其中,  $r_{vf}$  表示社会网络节点信任度  $v_{ij}$  与推荐的服务  $s_k$  的信任度相关的强弱程度,其取值范围为  $[-1, 1]$ 。当取值为 1、-1、0 时,分别表示完全正相关、完全负相关和无关。 $\bar{v}_j$  表示社会网络节点  $v_i$  对  $v_j$  的平均信任度,  $i$  取值为整数且为某一定值,  $j$  为整数且依次取值,取值范围为  $[1, m]$ ,  $m$  表示向节点  $v_i$  推荐服务  $s_k$  的社会网络节点的个数;  $f_{jk}$  表示社会网络节点  $v_j$  对服务  $s_k$  的直接信任度,  $\bar{f}_k$  表示  $m$  个社会网络节点  $v_j$  对服务  $s_k$  的直接信任度加权平均值。

#### 算法 2 面向个性化需求的可信服务推荐算法

输入: 服务集 servicesSubset 和用户对服务的个性化信任度阈值 trustDegree 和个性化需求匹配度阈值 threshold

输出: 被推荐的可信服务序列 recommendList

**BEGIN**

1) **IF**(servicesSubset is null)

2) return “No satisfied services with personalized requirements function”; /\*表示没有满足用户个性化功能需求的服务\*/

3) **ELSE** /\*表示有满足用户个性化功能要求的服务\*/

4) { **FOR EACH**  $s_k$  in servicesSubset /\*下面依次对 servicesSubset 中每个服务  $s_k$  的信任度进行处理\*/

5) { **IF** (degree > threshold &  $f_{jk}$

trustDegree)

6) { store degree and  $f_{jk}$  and  $s_k$  into

recommendList[0]; continue;}

7) **ELSE IF**(degree > threshold &  $ndt_{jk}$

trustDegree and  $r_{vf}$  is the maximum)

8) { store degree,  $ndt_{jk}$  and  $s_k$  into

recommendList[1]; continue;}

9) } **ENDFOR**

10) /\*下面按照 recommendList 返回服务的推荐列表\*/

11) **IF**(recommendList[0] != null)

12) return recommendList[0]; /\*先推荐具有直接信任度且满足用户个性化功能需求的服务\*/

13) **ELSE IF** (recommendList[1] != null)

14) return recommendList[1]; /\*否则,推荐社会网络环境下具有间接信任度且满足用户个

性化功能需求的服务\*/

15) ELSE IF ( $degree > threshold$ )

16) return  $servicesSubset$ ; /\*否则推荐满

足用户个性化功能需求的服务\*/

17) }

END

在算法2中, 首先判断 $servicesSubset$ 是否为空(见算法2中的1)), 若不为空, 则针对用户对服务的个性化需求(包括功能和信任需求)进一步处理, 当该服务满足信任度大于用户设定的个性化信任度 $trustDegree$ , 且个性化功能需求匹配度大于 $threshold$ 时, 根据算法2的推荐策略来推荐服务。由于服务信任度存在直接信任度、间接信任度、服务发布者提供的信任度及无信任度4种情况, 因此, 有必要针对不同的信任度进行判断。算法2在 $servicesSubset$ 中查找是否满足用户个性化需求的服务, 并保存相应的满足用户信任度需求的服务到 $recommendList$ (见算法2中的4)~算法2中的9))。其中, 5)的功能是判断对服务 $s_k$ 是否存在直接信任度且直接信任度值是否大于 $trustDegree$ , 且个性化功能需求匹配度 $degree$ 是否大于阈值 $threshold$ 。7)的功能是判断对服务 $s_k$ 的间接信任度值是否大于 $trustDegree$ 且 $r_{vj}$ 是否具有最大值, 且用户个性化功能需求匹配度是否大于 $threshold$ 。向社会网络节点推荐服务 $s_k$ 的规则依据是: 由于直接信任度是来自社会网络节点(令为 $v_i$ )自身对服务 $s_k$ 的直接经验体验, 信任值的客观性最高, 因此首先推荐具有直接信任度且满足用户个性化功能需求的 $s_k$ (见算法2中的11)和算法2中的12)); 由于间接信任度是来自社会网络中其他用户(不包括社会网络节点 $v_i$ )对某服务 $s_k$ 的直接经验信任度的体验, 间接信任度值的可信度是次高的, 因此, 在 $v_i$ 没有对服务的直接信任度情况下, 向 $v_i$ 推荐具有间接信任度且满足用户个性化功能需求的 $s_k$ (见算法2中的13)和算法2中的14)); 若 $s_k$ 无直接信任度和间接信任度, 则向 $v_i$ 推荐在功能上满足用户个性化功能需求的服务(见算法2中的15)和算法2中的16)), 以增加服务推荐的灵活性。令 $k=|servicesSubset|$ , 令 $n$ 为需要计算 $r_{vj}$ 和间接信任度的服务数目, 则算法2的时间复杂度为 $O(kn)$ 。

当社会网络节点 $v_j$ 向 $v_i$ 推荐了服务 $s_k$ 后,  $v_i$ 调用服务 $s_k$ , 根据服务调用后的信任反馈信息, 修正社会网络节点 $v_i$ 对 $v_j$ 的信任关度 $v_{ij}$ , 计算公式如式(4)所示。

$$v_{ij} = v_{ij} - \sqrt{\frac{\sum_{j=1 \wedge j \neq i}^n (f_{ik}^t - f_{jk}^t)^2}{n}} \quad (4)$$

其中,  $f_{ik}^t$ 表示 $t$ 时刻节点 $v_i$ 对服务 $s_k$ 的直接信任度( $f_{ik}^t$ 含义类似)。通过计算社会网络节点( $v_j$ )推荐的服务信任度 $f_{jk}^t$ 与 $v_i$ 对其直接信任度 $f_{ik}^t$ 的差异, 对社会网络节点间的信任度 $v_{ij}$ 进行修改。当社会网络节点 $v_j$ 推荐了不真实的服务时, 将可信服务推荐与社会网络节点间的信任关系有机关联, 社会网络节点 $v_i$ 将根据式(4)修改其对社会网络节点 $v_j$ 的信任度 $v_{ij}$ , 从而提高了社会网络节点间信任关系的可信性。

## 4 算法分析与实验测试

为了验证本文提出的服务推荐算法, 本文在真实的数据集上进行仿真实验和测试。

### 4.1 数据来源及实验相关参数

#### 1) 社会网络数据来源

目前真实的社会网络有很多, 如文献[12]提到的Flixster、Facebook、Myspace以及Flickr、Gnutella等。基于实际社会网络开展实验的研究有: 文献[3]通过新浪微博的开放平台API抓取实验数据, 以种子用户为起始, 通过逐步在新浪微博社会网络中生成兴趣图的方式来模拟基于弱关系的网络社区的形成过程; 文献[14]通过在实验室搭建一个基于手机通讯录的移动社会网络服务系统来获取实验数据。相对而言, 基于真实社会网络进行的实验研究较少, 因为获取社会网络数据的时间较长且获取的数据往往由于稀疏性而难以满足实验要求。因此, 很多研究是在真实社会网络的数据集上进行实验<sup>[15,16]</sup>。

社会网络数据集可以表现为图结构, 图数据特点主要体现为方向性和权重, 其中, 无向图和无权图所表示的社会网络数据集不适合具有信任关系算法的研究, 无向图(如ego-Facebook)忽略了用户间以及用户与服务之间信任关系的有向性, 而无权图(如wiki-Vote)难以满足节点之间以及服务应具备信任度的数据要求。目前仅有少数公认的社会网络数据集(如Epinions<sup>[15]</sup>)是具有权重(信任关系)的有向图, Epinions数据集来自现实社会网络www.epinions.com(面向用户的产品服务评论网站), 该社会网络独特地提供了“信任机制”, 使得节点之间具有信任关系。

本文基于Stanford Social Network Dataset Collection的Epinions<sup>注1</sup>数据集进行实验, 与文献

lection 的 Epinions<sup>注1</sup>数据集进行实验,与文献[15]类似(基于 Epinions 数据集),通过数据集表现具有信任关系的社会网络。Epinions 数据集包括 trust-data 和 rating\_data 两部分,其中,trust-data 是从实际的 Epinions 社会网络中获取的社会网络节点间的信任值,值为逻辑值(取值为 0 或 1),对于信任度值是实型值的算法(如本文算法及文献[16]等),其取值范围要求为[0,1],因此,为执行实验,还需要对 Epinions 的 trust-data 数据集进行改造,使用用户信任度的取值范围为[0,1],且用户信任度的赋值符合高斯分布( $\mu=0.5$ 、 $\sigma=0.25$ )。Epinions 的第二部分数据集是 ratings\_data,为社会网络节点对产品服务的定级信息,评价等级为整数值[1,5],该数据集缺少节点对服务的信任评价信息,因此,为执行实验,需要对 ratings\_data 数据集增加节点对服务的信任信息,服务的信任取值范围为[0,1],服务数共 10 000  $\alpha$  个, $\alpha$  为调节系数(即满足个性化功能需求及信任度需求的服务比例数)。实验数据集 Epinions 中 80%为训练数据,20%为测试数据。

为了适应本文实验场景,对社会网络节点进行扩充,以保存节点调用服务后的信任信息。笔者采用服务数据集 service\_dataset<sup>注2</sup>进行实验,该数据集共有 26 类服务(如 music、money、news、engineering 等),数据比较全面和充分,另外,为使数据集更为充实和真实,将注册在 WebServiceList 站点(<http://webservicelist.com/>)上的部分属于这 26 类的服务按数据集格式扩充到该服务数据集中。

## 2) 相关实验参数

为分析算法 1 和算法 2 的性能,需要设置相关实验参数。由算法及相关定义可知,每个社会网络节点拥有邻居列表( $NList$ ),邻居节点之间可以互相通信,在间接信任度计算时,通过信息回馈,基于社会网络节点推荐服务的  $SID$  和直接信任度来计算对服务的间接信任度。为避免信息无限地在社会网络中转发而造成时间等待和过多的资源耗费,令在社会网络中转发信息的跳数为 500  $\beta$ ,  $\beta$  为调节系数。

在服务功能的个性化需求分解与匹配方面,笔者以 OpenNLP 机器学习工具包处理个性化需求的切分及名称抽取,借助词典 WordNet 来处理用户需求的语义模糊性(见算法 1)。在此基础上,当实现面向个性化需求的可信服务推荐时,笔者基于直接信任度、间

接信任度和皮尔逊相关系数等计算结果进行可信服务的推荐(见算法 2)。每项测试至少包括 30 次实验,将 30 次实验结果的平均值作为一项测试结果。

表 1 仿真实验中的相关参数设置

参数名称	取值	说明
$ nodes $	75 879	Epinions 数据集中社会网络节点数
$ edges $	508 837	Epinions 数据集中社会网络节点之间的边数
$ S $	10 000	社会网络拥有的服务数
$ NList $	9	社会网络的节点的平均邻居数
threshold	0.5	算法 2 的个性化需求匹配阈值
trustDegree	0.5	算法 2 的个性化信任度阈值

## 4.2 算法的性能测试与分析

评价指标采用广泛认可的 precision(准确率)、recall(召回率)、 $F1$ -measure 方法。定义  $C$  为符合用户个性化需求的服务集, $R$  为推荐算法返回的服务集,令  $I = C \cap R$ ,即  $I$  是基于推荐算法得到的正确的服务推荐结果。基于这些前提,准确率、召回率及  $F1$ -measure 的计算公式如下

$$precision = \frac{|I|}{|R|}, recall = \frac{|I|}{|C|},$$

$$F1\text{-measure} = \frac{2precisionrecall}{precision + recall}$$

由于召回率可以通过准确率及  $F1$ -measure(简称  $F1$ )分析得到,因此,此处只对准确率及  $F1$  值进行分析。

共分成两大部分实验,其中,实验 1 的内容如下。

实验 1 主要从社会网络节点与服务之间的不同关系(包括直接信任、间接信任、服务发布者信任及无信任)及需求满足角度分析可信服务推荐的准确率和  $F1$  值。将本文提出的算法 2 中基于直接信任度满足个性化需求的可信服务推荐(简称 DTR)、基于间接信任度满足个性化需求的可信服务推荐(简称 ITR)与文献[17]及文献[18]推荐方法进行比较。

1) PTR 方法<sup>[17]</sup>:该方法实现的是基于服务提供者发布的信任度的服务推荐,该方法简称为 PTR。

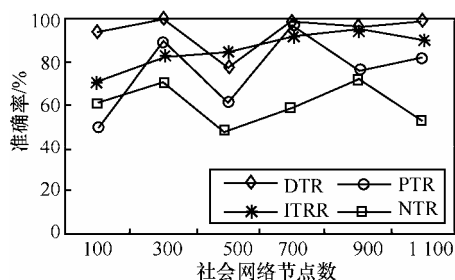
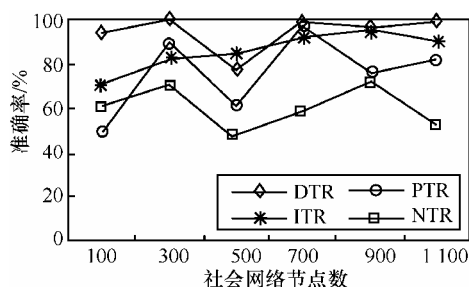
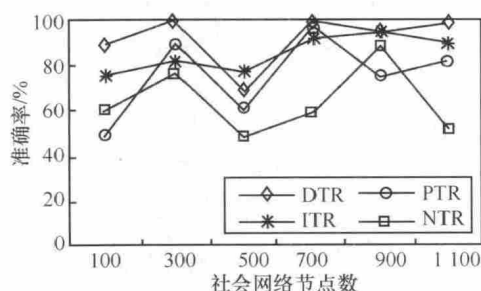
2) NTR 方法<sup>[18]</sup>:该方法实现的是基于无信任度的服务推荐,属于随机游走搜索推荐,当找到满足用户需求的服务时,推荐就终止,而不考虑服务是否可信的问题。该方法简称为 NTR。

评价指标 1 precision 分析

可信服务推荐的准确率(precision)数据如图 3~图 5 所示。

注1 <http://snap.stanford.edu/data/#socnets>。

注2 <http://andreas-hess.info/projects/annotator/index.html>。

图 3 服务推荐的准确率 ( $\alpha=0.8, \beta=0.5$ )图 4 服务推荐的准确率 ( $\alpha=0.5, \beta=0.5$ )图 5 服务推荐的准确率 ( $\alpha=0.8, \beta=1$ )

由图 3~图 5 分析可知，在 3 种配置下，基于直接信任度满足个性化需求的可信服务推荐（DTR）的准确率都很高，原因在于 DTR 是服务使用者向其自身推荐服务，但由于服务自身的变化或网络通信问题，可能导致推荐不可用或描述不一致的服务，出现推荐失误的情况，从而导致服务推荐的准确率值有时会下降。相对 DTR 而言，基于间接信任度满足个性化需求的可信服务推荐（ITR）的准确率要低些，原因在于社会网络节点推荐的服务真实度难以保证，有些社会网络中节点出于自身原因可能虚报了某服务的信任度，但该服务在实际运行后并不具备所描述和所推荐的功能，导致服务推荐的准确率开始时不高，但随着交易次数的增加，不断从社会网络中记录这些非善意的社会网络节点，通过降低对这些社会网络节点的信任度及其推荐的服务信任度，协同过滤非善意的节点和服务，提高 ITR 的服务推荐正确率。因此，随着交易次数的增加，ITR 方法的准确率也在逐渐上

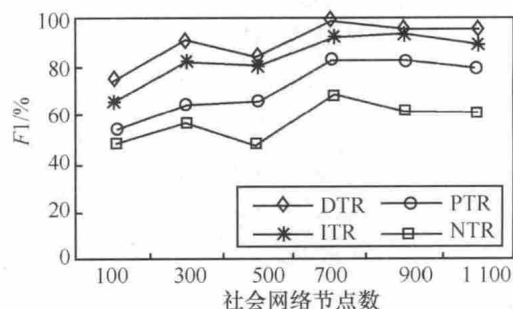
升。基于服务提供者发布的服务信任推荐（PTR），若服务提供者提供的信任度是准确的，则服务的推荐准确率高，否则准确率会受到影响。基于无信任推荐的服务推荐，其影响服务推荐的主要因素是服务提供者发布的服务质量，若服务质量高，则服务推荐的准确率高，否则推荐的准确率低。因此，PTR 与 NTR 服务推荐的准确性处于波动起伏较大的不定状态。从图 3 和图 4 可以看出，不同的调节系数  $\alpha$  对准确率影响不是很明显，基本没多大变化；从图 3~图 5 可以看出，不同的调节系数  $\beta$  对准确率有所影响，随着社会网络转发信息的跳数的增大，社会网络推荐的范围增大，推荐的准确性相对有所提高。

另外，图 3~图 5 中，由于节点数为 500 的社会网络中增加了不可信节点的比重（达 80% 以上），通过实验可知，当不可信的社会网络节点比例达 80% 以上时，对可信服务推荐的性能有很大的负面影响，导致服务推荐的性能在图 3~图 5 中出现规律性的下折效果（表现为在节点数为 300 和 500 时效果反差很大）。

#### 评价指标 2 $F1$ -measure 分析

$F1$ -measure（简称  $F1$ ）分析如图 6~图 8 所示。

从图 6~图 8 综合而言，基于直接信任度推荐的  $F1$ -measure 值是最高的，其次是基于间接信任度的推荐，基于间接信任度的可信服务推荐会随着社会网络节点间信任及时更新而排除掉不满足信任要求的社会网络节点，并且通过社会网络节点间的协同操作来不断过滤掉不满足信任要求的服务，从而提高社会网络节点的可信服务推荐的准确率和  $F1$ -measure 值。不同的调节系数  $\alpha$  和  $\beta$  对  $F1$ -measure 影响不是很明显。节点数为 500 的社会网络增加了不可信社会网络节点的比重（达 80% 以上），导致  $F1$ -measure 值相对降低较多。

图 6  $F1$ -measure 值分析 ( $\alpha=0.8, \beta=0.5$ )



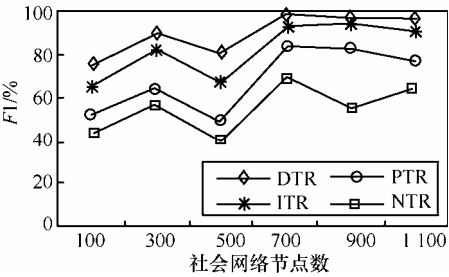


图 7 F1-measure 值分析 (  $\alpha=0.5, \beta=0.5$  )

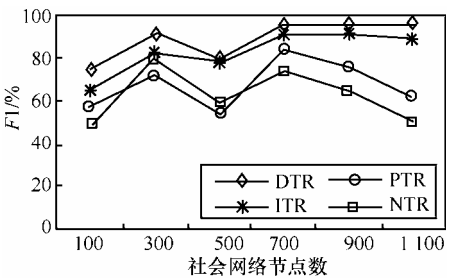


图 8 F1-measure 值分析 (  $\alpha=0.8, \beta=1$  )

综合图 3~图 8，算法 2 中提出的 DTR 和 ITR 方法及推荐策略是可行和有效的。在算法 2 中，由于基于直接信任度满足个性化需求的服务推荐局限于社会网络节点的内部，而基于间接信任度满足个性化需求的服务推荐是其他社会网络节点聚合直接信任度而成的，相对而言，基于间接信任度的服务推荐具有推广应用价值。实验表明通过皮尔逊相关系数可提高基于间接信任度服务推荐的准确率，可实现从候选服务集中推荐最相关的可信服务。

基于图 3~图 8，分析社会网络节点数对算法准确性和 F1 值的影响。由于节点数是 500 的社会网络中不可信节点比例很高，达 80% 以上，为异常点，分析时去掉该节点数所对应的数据。节点数对准确性的影响分析如表 2 所示。

表 2 社会网络节点个数的选择对算法准确值的影响

算法	社会网络节点数					平均偏差
	100	300	700	900	1 100	
DTR	0.014 8	0.067 1	0.171 5	0.048 8	0.040 8	0.068 6
ITR	0.114 5	0.016 5	0.014 5	0.070 9	0.074 5	0.058 18
PTR	0.257	0.177	0.157	0.216	0.021	0.165 6
NTR	0.034 4	0.081 9	0.187 4	0.002 3	0.137 6	0.088 72

表 2 前 5 列分别表示图 3~图 5 中相同节点数对应的平均准确率与 5 种节点数对应的总平均准确率的偏差的绝对值（即绝对偏差，记为  $\delta_n$ ）， $n$  表示节点数（取值分别为 100、300、700、900、1 100），最后一列表示 5 种节点数对应的准确率绝对偏差的

均值（记为  $\varepsilon$ ）， $\varepsilon = (\sum_{i=1}^n \delta_i) / n$ 。由表 2 可知，对 DTR、ITR 及 NTR 3 种方法而言，节点数对准确性的影响较小（体现在最后一列的数据较小），而对 PTR 方法影响相对较大。

对 F1 值的影响分析数据如表 3 所示。

表 3 社会网络节点个数的选择对算法 F1 值的影响

算法	社会网络节点数					平均偏差
	100	300	700	900	1 100	
DTR	0.157	0.005 6	0.067 1	0.054 4	0.038 4	0.064 5
ITR	0.209 5	0.016 9	0.088 5	0.098 8	0.039 1	0.090 6
PTR	0.175 1	0.057 1	0.116 9	0.094 5	0.020 9	0.092 9
NTR	0.134 5	0.058 5	0.095 1	0.005 8	0.024 9	0.063 8

表 3 前 5 列分别表示图 6~图 8 中相同节点数对应的平均 F1 值与 5 种节点数对应的总平均 F1 值的偏差的绝对值（计算类似于表 2 中的  $\delta_n$ ）。最后一列表示 5 种节点数对应的 F1 值绝对偏差的均值（计算类似于表 2 中的  $\varepsilon$ ）。从表 3 可以看出，社会网络节点数对 DTR、ITR、PTR 及 NTR 4 种方法的 F1 值影响较小（体现在最后一列的数据较小）。

**实验 2** 从社会网络节点—节点—服务之间的三元信任关系及需求满足角度出发，比较本文方法与文献[10]及文献[12]的推荐方法的性能（包括准确率、F1-measure）。

1) RS方法<sup>[10]</sup> 通过声誉来计算社会网络节点信任度，并推荐可信服务，简称RS方法。

2) SOCIALMF方法<sup>[12]</sup>：通过直接邻居节点的信任度矩阵来计算节点的信任度，并进行可信推荐，简称SOCIALMF方法。

评价指标 1 准确率分析

3 种方法的准确率比较分析如图 9 所示（取  $\alpha=0.5, \beta=0.5$ ，下同）。

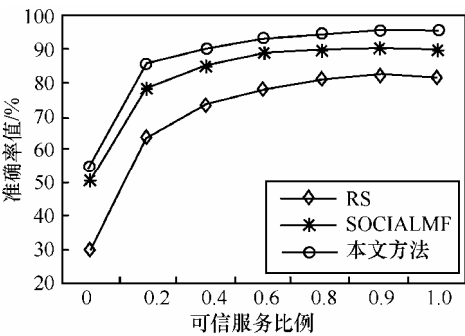


图 9 3 种方法的准确率分析

图 9 中横坐标 0.2 表示可信服务比例占 20%，而不可信服务比例占 80%。由图 9 可以看出，0.2 是一个比较明显的分界线，当可信服务比例小于 0.2 时，RS 方法的准确率值最小，本文方法的准确率与 SOCIALMF 方法的准确率比较相近，在 50% 左右。当可信服务比例大于 0.2 时，随着比例的增加，3 种方法的准确率值均上升，表示算法性能均有所提高，当比例为 0.9~1 时，准确率值达到最高值。实验表明比例太小（小于 0.2）时，可信服务推荐的准确率相对是较低的。

#### 评价指标 2 $F1-measure$ 分析

3 种方法的  $F1-measure$  比较分析如图 10 所示。

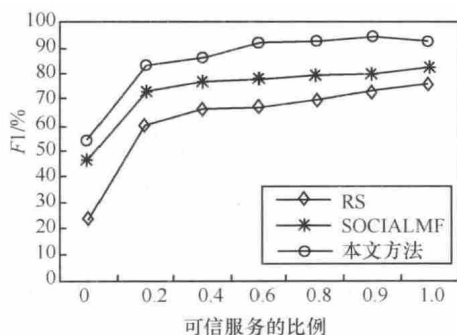


图 10 3 种方法的  $F1-measure$  分析

图 10 与图 9 类似，0.2 在图 10 中仍是一个比较明显的性能界限。由图 10 可知，本文方法的  $F1-measure$  值相对优于 SOCIALMF 方法和 RS 方法的  $F1-measure$  值。

综合图 3~图 10 可知，本文方法既考虑到社会网络节点间的信任度，也考虑到社会网络节点与服务的直接信任度和间接信任度，通过多重协作过滤方式推荐满足个性化需求的可信服务，相对而言，其  $precision$  和  $F1-measure$  性能是三者中最好的。

## 5 结束语

在社会计算环境中，基于用户需求定位可信任服务是服务发现的主要手段之一，而如何面向用户的个性化需求推荐可信任的服务是当前需要解决的重要问题。在社会计算与服务计算背景下，针对难以获取满足用户需求的可信 Web 服务问题，给出社会网络服务信任关系，提出基于社会网络面向个性化需求的可信服务推荐模型，设计针对用户个性化功能需求的分解与匹配算法；在功能需求匹配的基础上，设计可信服务推荐算法，该算法在满足用户对服务的个性化功能需求的同时，要求被推荐的

服务满足信任度需求。本文提出的方法既可降低传统的集中式控制过程中服务发现的瓶颈问题及可信服务推荐可靠度低的问题，又可通过需求匹配来推荐个性化服务。

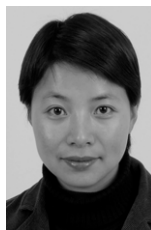
社会网络作为当前服务应用的主要场景之一，未来工作包括以下 2 个方面：1) 结合可信计算及服务计算，通过凝聚式聚类方法抽取网络的层次结构<sup>[19]</sup>，研究利用分层社会网络技术提高可信服务推荐的性能；2) 结合可信计算及服务计算，通过获取社会网络中近似的最小地点控制子集<sup>[20]</sup>，研究基于控制子集技术来提高可信服务推荐的效率。

#### 参考文献：

- [1] 韩燕波, 陈俊亮, 王千祥. 云计算和服务计算[J]. 计算机学报, 2011, 34(12): 2251-2252.  
HAN Y B, CHEN J L, WANG Q X. The preface of special issue of clouding and services computing[J]. Chinese Journal of Computers, 2011, 34(12): 2251-2252.
- [2] 王立才, 孟祥武, 张玉洁. 上下文感知推荐系统[J]. 软件学报, 2012, 23(1): 1-20.  
WANG L C, MENG X W, ZHANG Y J. Context-aware recommender systems[J]. Journal of Software, 2012, 23(1): 1-20.
- [3] 陈克寒, 韩盼盼, 吴健. 基于用户聚类的异构社交网络推荐算法[J]. 计算机学报, 2013, 36(2): 249-359.  
CHEN K H, HAN P P, WU J. User clustering based social network recommendation[J]. Chinese Journal of Computers, 2013, 36(2): 249-359.
- [4] BOBADILLA J, ORTEGA F, HERNANDO A. A collaborative filtering similarity measure based on singularities[J]. Information Processing and Management, 2012, (48): 204-217.
- [5] WALTER C N, MARIA L H A, RAFAE V G, et al. Social knowledge-based recommender system application to the movies domain[J]. Expert Systems with Applications, 2012, (39): 10990-11000.
- [6] TANG M, JIANG Y, LIU J, et al. Location-aware collaborative filtering for QoS-based service recommendation[A]. Proc of IEEE 19th International Conference on Web Services (ICWS)[C]. Washington DC, USA, 2012: 202-209.
- [7] GOLBECK J, HENDLER J. Inferring trust relationships in Web based social networks[J]. ACM Transactions on Internet Technology, 2006, 6(4): 497-529.
- [8] LIU G, WANG Y, ORGUN M A, et al. Discovering trust networks for the selection of trustworthy service providers in complex contextual social networks[A]. Proc of IEEE 19th International Conference on Web Services (ICWS), IEEE Computer Society[C]. Washington DC, USA, 2012: 384-391.
- [9] LIU G, WANG Y, ORGUN M A. Finding k optimal social trust paths for the selection of trustworthy service providers in complex social

- networks[A]. Proc of IEEE 18th International Conference on Web Services (ICWS) IEEE Computer Society[C]. Washington DC, USA, 2011.41-48.
- [10] YAO J, TAN W, NEPAL S, *et al.* ReputationNet: a reputation engine to enhance servicemap by recommending trusted services[A]. Proc of IEEE Ninth International Conference on Services Computing (SCC), IEEE Computer Society[C]. Washington DC, USA, 2012. 454-461.
- [11] ZHONG Y, ZHAO W, YANG J. Personal-hosting RESTful web services for social network based recommendation[A]. Proc of the 9th International Conference on Service-Oriented Computing(ICSOC)[C]. Paphos, Cyprus, 2011.661-668.
- [12] JAMALI M, ESTER M. A matrix factorization technique with trust propagation for recommendation in social networks[A]. Proc of ACM RecSys, ACM[C]. USA, 2010.135-142.
- [13] MASSA P, AVESANI P. Trust-aware recommender systems[A]. Proc of ACM Recommender Systems Conference[C]. Minneapolis, Minnesota, USA, 2007.17-24.
- [14] 王玉祥, 乔秀全, 李晓峰等. 上下文感知的移动社交网络服务选择机制研究[J]. 计算机学报, 2010, 33(11):2126-2135.  
WANG Y X, QIAO X Q, LI X F, *et al.* Research on context-awareness mobile SNS Service selection mechanism[J]. Chinese Journal of Computers, 2010, 33(11):2126-2135.
- [15] MA H, KING I, LYU M R. Learning to recommend with explicit and implicit social relations[J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3):1-19.
- [16] RICHARDSON M, AGRAWAL R, DOMINGOS P. Trust management for the semantic Web[A]. Proc of 2nd International Semantic Web Conference (ISWC), IEEE Computer Society[C]. White Plains, NY, USA, 2003.351-368.
- [17] HANG C, SINAGH M P. Trustworthy service selection and composition[J]. ACM Transactions on Autonomous and Adaptive Systems, 2011, 6(1):1-23.
- [18] HURLEY N, ZHANG M. Novelty and diversity in top-N recommendation analysis and evaluation[J]. ACM Transactions on Internet Technology, 2011,10(4):1-30.
- [19] 韩毅, 方滨兴, 贾焰等. 基于密度估计的社会网络特征簇挖掘方法[J]. 通信学报, 2012,33(5):38-48.  
HAN Y, FANG B X, JIA Y, *et al.* Mining characteristic clusters: a density estimation approach[J]. Journal on Communications, 2012, 33(5):38-48.
- [20] 胡玉鹏, 罗昊, 林亚平等. 社会网络中时空周期行为模式挖掘算法[J]. 通信学报, 2013, 34(1):8-18.  
HU Y P, LUO H, LIN Y P, *et al.* Spatio-temporal periodic behavior mining algorithm for social networks[J]. Journal on Communications, 2013, 34(1):8-18.

#### 作者简介：



张佩云（1974-），女，安徽安庆人，博士后，安徽师范大学副教授，主要研究方向为服务计算、数据挖掘和智能信息处理等。

陈恩红（1968-），男，安徽宁国人，博士，中国科学技术大学教授、博士生导师，主要研究方向为约束满足问题、机器学习、数据挖掘和网络信息处理等。

黄波（1980-），男，湖南醴陵人，博士后，南京理工大学副教授，主要研究方向为计算机网络技术、智能信息处理等。