

# 数字图书馆系统中基于 Ontology 的用户偏好模型\*

杨艳<sup>1,2+</sup>, 李建中<sup>1,2</sup>, 高宏<sup>1,2</sup>

<sup>1</sup>(哈尔滨工业大学 计算机科学与技术学院,黑龙江 哈尔滨 150001)

<sup>2</sup>(黑龙江大学 计算机科学技术学院,黑龙江 哈尔滨 150080)

## Ontology-Based Preference Model in Digital Library

YANG Yan<sup>1,2+</sup>, LI Jian-Zhong<sup>1,2</sup>, GAO Hong<sup>1,2</sup>

<sup>1</sup>(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

<sup>2</sup>(School of Computer Science and Technology, Heilongjiang University, Harbin 150080, China)

+ Corresponding author: Phn: +86-451-88611102, E-mail: yangyan@mail.banner.com.cn, http://www.hit.edu.cn

Received 2004-10-04; Accepted 2005-01-07

Yang Y, Li JZ, Gao H. Ontology-Based preference model in digital library. *Journal of Software*, 2005,16(12): 2080-2088. DOI: 10.1360/jos162080

**Abstract:** Personalization poses new challenges to digital library. How to describe users' preferences and how to support preference queries in digital libraries are among the top demanding tasks. A strict partial order preference model has been proposed, and a series of preference generation methods have already been developed for relational data. However, the semi-structured data stored in digital libraries make it more complex to model users' preferences than its counterpart in relational databases. The partial order preference model cannot suffice to support such complex preferences. A new ontology based preference model is proposed in this paper to overcome this difficulty. This model describes the documents along with the preferences regarding the documents in the digital libraries using ontology, which allows both the structure and the semantic of a user's preference to be fully represented. A set of complex preference operations have also been provided in the model to support the personalized query and recommendation efficiently.

**Key words:** digital library; ontology; personalization; preference; model

**摘要:** 个性化服务技术为数字图书馆的研究带来一些新的挑战.如何描述用户的偏好以及如何使数字图书馆支持偏好查询是有待研究的一个问题.人们已经提出了基于偏序关系的用户偏好模型,并针对关系数据提出了一系列偏好构造方法.数字图书馆中的数据是半结构数据.半结构数据上用户偏好的描述比关系数据复杂得多.偏序模型无法有效地表达数字图书馆中的用户偏好.提出基于 ontology 的新的用户偏好模型,用 ontology 来描述数字图书馆中的文本和文本上的偏好.该模型能够充分表达用户偏好的结构和语义,并提供了复杂的偏好操作,能够有效地支持数字图书馆中的个性化检索和推荐操作.

\* Supported by the National Natural Science Foundation of China under Grant No.60273082 (国家自然科学基金); the Key Natural Science Foundation of Heilongjiang Province of China under Grant No.zjg03-05 (黑龙江省自然科学基金重点项目)

作者简介: 杨艳(1973 - ),女,黑龙江哈尔滨人,博士生,讲师,主要研究领域为数据库,数字图书馆;李建中(1950 - ),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为数据库,并行计算技术;高宏(1966 - ),女,博士,副教授,主要研究领域为数据库,数据仓库.

关键词: 数字图书馆;ontology;个性化;偏好;模型

中图法分类号: TP311 文献标识码: A

个性化技术的出现引入了用户偏好的概念.用户偏好表达了不同用户的不同意愿,是一种非严格约束.例如,与“personalization”相关的文章.用户希望得到与个性化服务相关的文献,这包括用户感兴趣的个性化检索、个性化推荐等很多内容,而不仅仅是出现“personalization”一词的文章.

传统的数字图书馆中用户通过查询来表达自己的意愿,只能根据查询条件的严格匹配与否为用户提供资源,无法表达上述偏好.随着个性化检索技术的发展,人们开始注意到用户偏好的描述问题.用户偏好的表示主要分为两大类.一是使用用户个性化信息来表示用户偏好.二是直接给出用户偏好的表示模型,使用户能够直接表达自己的偏好.用户个性化信息大致可分为用户统计信息、用户操作历史、用户访问过的资源特征等<sup>[1]</sup>.Susan Gauch 等人在文献[2]中总结了 50 个个性化系统,讨论了不同的个性化系统中用户个性化信息的表示方法,如布尔模型、加权关键字向量、语义网络、n-grams 等,并提出用 ontology 来描述用户的个性.这些系统在支持个性化检索操作时,主要是对传统检索的结果按照用户的个性化信息进行过滤,从而区分出用户感兴趣和不感兴趣的检索结果.它们都不能真正实现偏好查询,也没有讨论偏好模型问题.Werner KieBling 等人首次提出用偏序关系来表示用户的偏好<sup>[3]</sup>.他们指出,现实世界中用户经常用“喜欢 A 胜于喜欢 B”来表达自己的偏好,这很容易用偏序关系描述.并针对关系数据库的特点,提出偏好的构造及偏好模型上的代数操作<sup>[4]</sup>,提出关系数据库中偏序偏好的实现技术<sup>[5]</sup>.文献[6]把基于偏序的偏好用于文本检索,把用户提交的检索词划分成若干个不同偏好级别的子空间,从而实现了文本数据上的偏好查询.

文本数据是半结构数据,比关系数据有着更为丰富的语义信息.因此,数字图书馆中的偏好查询与关系数据上的偏好查询有着显著的不同.在关系数据中,用户可以根据元组的各个属性表达更喜欢哪个元组.对于文本数据来说,无法根据用户提交的检索词,如“personalization”,来判断结果文本中用户更喜欢文本 A 还是文本 B.因此,Werner KieBling 等人提出的偏好模型及相应的偏好代数都不能很好地适用于数字图书馆中的文本数据.

本文提出一个新的基于ontology的用户偏好模型,它比偏序偏好模型有更强的表达能力.Ontology已经在很多领域得到广泛应用<sup>[7,8]</sup>,但它的概念仍然模糊,而且在不同的领域有不同的定义,针对不同的应用也有不同的结构.Alexander Maedche在《Ontology Learning for the Semantic Web》一书中对ontology的结构以及ontology的词汇进行了如下的形式化定义.

**定义1(ontology结构).** 一个ontology结构是一个5元组 $O := \{C, R, H^C, rel, A^O\}$ .其中, $C$ 为概念集合,其中的一个元素称为一个概念. $R$ 为关系集合,其中的一个元素称为一个关系. $H^C$ 为概念层次或分类层次, $H^C \subseteq C \times C$ 是一种有向关系, $H^C(C_1, C_2)$ 表示 $C_1$ 是 $C_2$ 的子概念. $rel: R \rightarrow C \times C$ 是一个函数,表示概念之间的非分类关系.定义函数 $dom: R \rightarrow C, dom(R) := \Pi_1(rel(R))$ 给出 $R$ 的定义域,定义函数 $range: R \rightarrow C, range(R) := \Pi_2(rel(R))$ 给出 $R$ 的值域. $rel(R) = (C_1, C_2)$ 亦可表示为 $R(C_1, C_2)$ . $A^O$ 为使用某种逻辑语言表达的ontology的公理集.

本文首先用 ontology 来描述文本,把文本的属性、结构和内容用一个统一的概念层次表示.文本上的偏好为文本的属性、结构或内容上的一些约束条件,也是一个概念层次,因而也很容易用 ontology 描述.

本文第 1 节给出基于 ontology 的文本表示.第 2 节介绍基于 ontology 的偏好模型.第 3 节与现有偏好模型进行比较.最后是文章的结论.

## 1 基于 ontology 的文本表示

文本数据不同于关系数据,文本有文本的元数据,有章节等结构信息,有内容信息.本文把文本看作一个概念层次,用一个文本 ontology 描述,记为  $O_T$ .

**定义2(文本结构).** 文本以及文本的卷、篇、章、节、段等组成部分共同构成文本结构.文本结构中的每一个组成部分称为文本的一个结构概念.文本的各结构概念之间存在包含关系,这种包含关系构成一种层次结构.最上层为文本本身.一个文本的文本结构中包括哪些结构概念是根据文本的作者对文本结构的定义以及用户

对文本结构上粒度的要求而确定的。

**定义 3(结构概念层次).** 设  $C$  为文本中结构概念的集合,  $R$  为  $C$  中概念之间存在的关系的集合,  $rel:R \rightarrow C \times C$  是一个函数, 表示两个结构概念之间的特定关系, 则称  $H=(C,R,rel)$  为一个结构概念层次. 设  $C_1$  和  $C_2$  是  $C$  中两个特定的概念, 二者满足  $R$  中特定的关系  $R_1, rel(R_1)=(C_1, C_2)$  亦可表示为  $R_1(C_1, C_2)$ .

**定义 4(属性).** 文本的每一个结构概念都有相应的属性. 文本本身的属性又称为文本的元数据, 如文本的标题、作者、作者单位、关键字、发表时间、页数、出版社、格式、语言、价格等. 文本的其他结构概念也有相应的属性, 如某章的标题、作者、页数、起止页码等. 我们把文本中所有结构概念的属性统称为属性. 一个文本的属性集中包括哪些内容是根据文本作者对属性的定义以及用户对属性的要求而确定的. 每一个属性也称为一个属性概念.

**定义 5(属性概念层次).** 设  $C$  为某一个文本概念及其属性概念的集合,  $R$  为  $C$  中概念之间存在的关系的集合,  $rel:R \rightarrow C \times C$  是一个函数, 表示  $C$  中两个概念之间的特定关系, 则称  $H=(C,R,rel)$  为一个属性概念层次. 设  $C_1$  和  $C_2$  是  $C$  中两个特定的概念, 二者满足  $R$  中特定的关系  $R_1, rel(R_1)=(C_1, C_2)$  亦可表示为  $R_1(C_1, C_2)$ .

**定义 6(文本概念).** 文本中所有属性概念和结构概念统称为文本概念.

**定义 7(值).** 每一个文本概念都有相应的值. 结构概念的值定义为该结构概念的内容, 是一个长字符串. 如果一个结构概念包含其他结构概念, 则它的值是由它所包含的所有下层结构概念的值的串连接形成的. 属性概念的值即为该属性的值. 根据属性的取值类型把属性分为数值属性和字符属性两类.

**定义 8(文本).** 我们将文本  $O_T$  定义为一个五元组  $(C_T, V_T, R_C, rel_C, map_V)$ . 其中, (1)  $C_T$  为文本概念的集合; (2)  $V_T$  为  $C_T$  中文本概念的值的集合; (3)  $R_C$  为  $C_T$  中各概念之间关系的集合; (4)  $rel_C:R_C \rightarrow C_T \times C_T$  表示两个概念之间的某种关系,  $DOM(R_C)=\{attribute\_of, composed\_of, next\_to\}$ . 设  $C_1, C_2 \in C_T$  是两个概念:  $attribute\_of(C_1, C_2)$  表示  $C_1$  和  $C_2$  之间存在  $attribute\_of$  关系, 即概念之间的属性关系, 表示  $C_2$  是  $C_1$  的属性;  $composed\_of(C_1, C_2)$  表示  $C_1$  和  $C_2$  之间存在  $composed\_of$  关系, 即概念之间的组成关系, 表示  $C_2$  是  $C_1$  的组成部分;  $next\_to(C_1, C_2)$  表示  $C_1$  和  $C_2$  之间存在  $next\_to$  关系, 即概念之间的后继关系, 表示  $C_2$  是与  $C_1$  同一层次的下一个概念. 对于任意两个概念  $C_i, C_j \in C_T$ , 若存在  $next\_to(C_i, C_j)$ , 则必存在  $C_k \in C_T$ , 使得  $composed\_of(C_k, C_i)$  和  $composed\_of(C_k, C_j)$  同时成立.  $C_T$  中各文本概念及它们之间的关系构成结构概念层次或属性概念层次. (5) 映射  $map_V$  把文本概念与其值对应起来. 设  $V_1 \in V_T, C_1 \in C_T, map_V:C_1 \rightarrow V_1$  表示  $V_1$  是  $C_1$  的值, 也表示为  $content\_of(C_1, V_1)$ .

## 2 基于 ontology 的偏好模型

### 2.1 基本概念

文本的属性概念、结构概念及它们的值可以用一个统一的文本 ontology 描述. 我们将用户对文本的偏好看作是对文本中指定的文本概念的值得上的约束, 因此可以用与文本类似的 ontology 来定义偏好模型. 一般来说, 偏好中涉及的文本概念是文本中所有文本概念的一个子集. 为了定义用户偏好, 先介绍语义概念和语义概念层次.

**定义 9(语义概念).**  $V_T$  为文本概念的值的集合,  $STR$  表示所有字符串的集合. 对于  $\forall v_i \in \{v_j | v_j \in STR \wedge v_j \in V_T\}$ , 记为  $v_i = a_1 a_2 a_3 \dots a_n, a_i (1 \leq i \leq n)$  为一个字符串,  $v_{i_{km}} = a_k a_{k+1} \dots a_{k+m-1} (k \geq 1, k+m-1 \leq n)$  称为  $v_i$  的一个子串. 令  $V'_T = \{v_{i_{km}} | v_{i_{km}}$  是单词或短语,  $\wedge v_{i_{km}}$  是  $v_i$  的子串,  $\wedge v_i$  是一个字符型文本概念的值的值}, 这里,  $v_{i_{km}}$  称为一个 term. 令  $V''_T = \{v_{i_x} | v_{i_x}$  是单词或短语,  $\wedge v_{i_x}$  和  $v_{i_{km}}$  语义相关,  $\wedge v_{i_{km}}$  是  $v_i$  的子串,  $\wedge v_i$  是一个字符型文本概念的值的值}, 这里,  $v_{i_x}$  也称为一个 term. 令  $C_V = V_T \cup V'_T \cup V''_T, v_i \in C_V$  称为一个语义概念.

**定义 10(语义概念层次).** 设  $C$  为语义概念的集合,  $R$  为  $C$  中语义概念之间的语义相关关系 (correlation 关系),  $rel:R \rightarrow C \times C \times VAL$  是一个函数, 表示两个概念之间存在相关性,  $VAL$  是相关度的值. 称  $H=(C,R,rel)$  为一个语义概念层次. 设  $C_1$  和  $C_2$  是  $C$  中两个特定的概念,  $V_1$  是二者的相关度,  $rel(R)=(C_1, C_2, V_1)$  亦可表示为  $correlation(C_1, C_2)=V_1$ .

**定义 11(用户偏好概念层次).** 在语义概念层次  $H=(C,R,rel)$  中, 如果  $C=\{v_{i_x} | v_{i_x}$  是单词或短语,  $\wedge v_{i_x}$  是  $v_i$  的子串或与  $v_i$  的子串语义相关,  $\wedge v_i$  是一个用户指定的字符型文本概念的约束值}, 则称  $H=(C,R,rel)$  为用户偏好概念层次.

用户对文本的偏好即对指定文本概念的值上的约束.根据约束条件的满足情况,把约束条件划分为严格约束和非严格约束两类.数字图书馆上传统的查询条件属于严格约束条件,即用户给出的约束条件必须严格满足.在非严格约束条件中,用户表达的是更喜欢什么,约束条件不一定严格满足.

**定义 12(偏好约束).** 偏好约束  $pcon=(c,v,cond,L)$  为如下定义的四元组:

(1)  $c$  为一个文本概念, $v$  为对  $c$  进行约束的值, $cond \in \{<, \leq, =, >, \geq, \neq, contain, max, min, around, semantic\_contain\}$  为  $c$  和  $v$  之间满足的约束条件.如果  $cond$  为单目运算符,则相应的值  $v$  为空值. $L \in \{0,1\}$  表示  $cond$  是严格约束条件(0)或非严格约束条件(1),则四元组  $pcon=(c,v,cond,L)$  为一个偏好约束.

(2) 若  $pcon$  是偏好约束, $real$  是实数,则函数  $fun: pcon \rightarrow real$  是一个偏好约束.

(3)  $c=pcon_1$  为一个偏好约束, $v=pcon_2$  为另一个偏好约束, $cond \in \{\wedge, +\}$  表示  $pcon_1$  和  $pcon_2$  之间的关系, $L \in \{0,1\}$  表示  $cond$  是严格约束条件(0)或非严格约束条件(1),则四元组  $pcon=(c,v,cond,L)$  也是一个偏好约束.

如果一个偏好约束中不涉及  $\wedge$  和  $+$  运算,则该偏好约束称为基本偏好约束,否则称为复合偏好约束.如果一个偏好约束的结果是一个布尔值,则称该偏好约束为逻辑偏好约束.如果一个偏好约束的结果是一个实数值,则称该偏好约束为数值偏好约束.约束条件“ $\wedge$ ”只能作用于两个逻辑偏好约束,表示两个约束条件同时满足;约束条件“ $+$ ”只能作用于两个数值偏好约束,表示两个数值偏好约束结果的代数和.一个复合偏好约束  $pcon$  必然由两个基本偏好约束或复合偏好约束直接复合而成.直接和间接构成偏好约束  $pcon$  的所有基本和复合偏好约束构成一个集合,称为以  $pcon$  为终结约束的一个偏好约束组.称  $pcon$  为该偏好约束组的终结约束.如果一个偏好约束组中只有一个基本偏好约束,则该基本偏好约束为该组的终结约束.偏好约束也表示为  $cond(c,v,L)$ (如果  $cond$  是单目运算符,则表示为  $cond(c,L)$ ).

**定义 13(偏好).** 偏好  $P=(C_p, R_{pc}, rel_{pc}, V_p, PCON)$  是一个五元组.其中,  $C_p$  为偏好中涉及的文本概念的集合.  $R_{pc}$  为  $C_p$  中各文本概念之间关系的集合,  $R_{pc} \subseteq \{attribute\_of, composed\_of, next\_to\}$ .  $rel_{pc}: R_{pc} \rightarrow C_p \times C_p$  是一组映射,表示某些文本概念之间存在的特定关系.  $V_p$  为对  $C_p$  中概念进行约束的值的集合,  $V_p \subseteq STR \cup NUM \cup CONCEPT$ , 其中  $STR, NUM$  和  $CONCEPT$  分别表示字符型值、数值型值和用户偏好概念层次.  $PCON$  为一个偏好约束组.如果偏好约束组中的终结约束为基本偏好约束,则称该偏好为基本偏好.用户偏好是由文本的属性概念层次、结构概念层次以及语义概念层次构成的一个复杂的概念层次,是一个 ontology.

## 2.2 基本偏好

偏好约束是描述用户偏好的关键因素.根据偏好约束中的约束条件是否严格满足,把基本偏好划分为严格约束偏好和非严格约束偏好.非严格约束偏好根据偏好约束作用的文本概念不同,又分为属性约束偏好和内容约束偏好.

- 严格约束偏好:偏好约束中的约束条件必须严格满足.对于每一个文本,若满足条件则为用户偏好的文本,否则用户不喜欢.如果没有满足条件的文本则用户偏好的文本集合为空.严格约束偏好中,偏好约束是属性概念上的约束,表示为  $cond(A, val, 0)$ . 其中,  $A$  为属性,  $val$  为对  $A$  进行约束的值,  $0$  表示  $cond$  为  $A$  和  $val$  之间的严格约束条件,  $cond \in \{<, \leq, =, >, \geq, \neq, contain\}$ .

- 属性约束偏好:在文本的属性概念上表达的非严格约束偏好.对于属性概念上的逻辑偏好约束,用户更喜欢满足条件的文本,如果没有满足条件的文本则考虑其他文本;对于属性概念上的数值偏好约束,根据偏好约束的结果值判断用户更喜欢哪个文本.属性约束偏好中的偏好约束表示为  $cond(A, val, 1)$ ,  $A$  为属性,  $val$  为对  $A$  进行约束的值,  $1$  表示  $cond$  为  $A$  和  $val$  之间满足的非严格约束条件.

- 内容约束偏好:在文本的结构概念上表达的非严格约束偏好.内容约束偏好中的偏好约束判断文本的结构概念与用户指定的  $term$  或与该  $term$  相关的用户偏好概念层次的相关程度.相关程度越高用户越喜欢.内容约束偏好中的偏好约束表示为  $cond(E, val, 1)$ ,  $E$  为结构概念,  $val$  为对  $E$  进行约束的值,  $1$  表示  $cond$  为  $E$  和  $val$  之间满足的非严格约束条件.

严格约束偏好的处理与传统的查询处理方法相同.如果严格约束偏好与非严格约束偏好组合,则把严格约束偏好作为过滤器,在经过严格约束条件过滤后的文本集上应用非严格约束条件.因此,下面只考虑非严格约束

偏好.如果已知偏好约束中的约束条件为非严格约束条件,偏好约束也可表示为 $(c,v,cond)$ 或 $cond(c,v)$ ,如果 $cond$ 为单目运算符,则偏好约束表示为 $(c,cond)$ 或 $cond(c)$ .

为便于说明,下文中引入如下记号: $x,y$ 表示文本, $P$ 表示偏好.根据偏好 $P,x<_P y$ 表示在文本 $x$ 和文本 $y$ 之间更喜欢 $y$ .! $(x<_P y)$ 表示喜欢文本 $y$ 的程度不超过喜欢文本 $x$ 的程度,包括3种情况:在文本 $x$ 和文本 $y$ 之间更喜欢 $x$ 或者喜欢 $x$ 和 $y$ 的程度相同或者 $x$ 和 $y$ 不可比(即无法说明更喜欢 $x$ 还是更喜欢 $y$ ).下面分属性约束偏好和内容约束偏好两类来介绍几种基本偏好.各种基本偏好的定义只与偏好约束相关,因此,下文对基本偏好的定义只考虑偏好约束,其参数见偏好定义.

### 2.2.1 属性约束偏好

属性约束偏好是指偏好约束中只涉及属性概念,根据属性概念的数据类型及运算符的不同分为如下几种.

**定义14(数值比较约束偏好).**  $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ 是一个数值比较约束偏好,其中 $PCON=\{(c,v,cond,L)\}$ , $c \in C_p$ 是一个数值型属性, $v \in V_p$ 为对 $c$ 进行约束的数值, $cond \in \{<,\leq,=,>,\geq,\neq\}$ 为 $c$ 和 $v$ 之间满足的约束条件, $L=1$ 表示 $cond$ 是非严格约束.

在数值比较约束偏好 $P$ 的作用下, $x<_P y$  iff  $!cond(x.c,v) \wedge cond(y.c,v)$ .其中 $x.c,y.c$ 分别为文本 $x$ 和 $y$ 的 $c$ 属性, $cond(y.c,v)$ 表示 $y.c$ 和 $v$ 之间满足 $cond$ 约束条件, $!cond(x.c,v)$ 表示 $x.c$ 和 $v$ 之间不满足 $cond$ 约束条件.

例: $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ , $C_p=\{\text{文本,页数}\}$ , $R_{pc}=\{\text{attribute\_of},\text{attribute\_of}(\text{文本,页数})\}$ , $V_p=\{8\}$ , $PCON=\{(\text{页数},8,\geq,1)\}$ . $P$ 是一个数值比较约束偏好,表示用户更喜欢大于等于8页的文章.

**定义15(最大值约束偏好).**  $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ 是一个最大值约束偏好,其中 $PCON=\{(c,v,cond,L)\}$ , $c \in C_p$ 是一个数值型属性, $v=\emptyset$ , $cond=\max$ 为 $c$ 上的约束条件, $L=1$ 表示 $cond$ 是非严格约束.

在最大值约束偏好 $P$ 的作用下, $x<_P y$  iff  $x.c < y.c$ .其中 $x.c,y.c$ 分别为文本 $x$ 和 $y$ 的 $c$ 属性.在文本集合中, $c$ 属性值越大的文本用户越喜欢.

例: $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ , $C_p=\{\text{文本,发表时间}\}$ , $R_{pc}=\{\text{attribute\_of},\text{attribute\_of}(\text{文本,发表时间})\}$ , $V_p=\emptyset$ , $PCON=\{(\text{发表时间},\emptyset,\max,1)\}$ . $P$ 是一个最大值约束偏好,表示用户喜欢最新的文本,即文本的发表时间离当前时间越近则用户越喜欢.

**定义16(最小值约束偏好).**  $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ 是一个最小值约束偏好,其中 $PCON=\{(c,v,cond,L)\}$ , $c \in C_p$ 是一个数值型属性, $v=\emptyset$ , $cond=\min$ 为 $c$ 上的约束条件, $L=1$ 表示 $cond$ 是非严格约束.

在最小值约束偏好 $P$ 的作用下, $x<_P y$  iff  $x.c > y.c$ .其中 $x.c,y.c$ 分别为文本 $x$ 和 $y$ 的 $c$ 属性.在文本集合中, $c$ 属性值越小的文本用户越喜欢.

例: $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ , $C_p=\{\text{文本,价格}\}$ , $R_{pc}=\{\text{attribute\_of},\text{attribute\_of}(\text{文本,价格})\}$ , $V_p=\emptyset$ , $PCON=\{(\text{价格},\emptyset,\min,1)\}$ . $P$ 是一个最小值约束偏好,表示用户喜欢最便宜的文本,即文本的价格越低则用户越喜欢.

**定义17(邻近值约束偏好).**  $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ 是一个邻近值约束偏好,其中 $PCON=\{(c,v,cond,L)\}$ , $c \in C_p$ 是一个数值型属性, $v \in V_p$ 为对 $c$ 进行约束的数值, $cond=\text{around}$ 为 $c$ 上的约束条件, $L=1$ 表示 $cond$ 是非严格约束.

在邻近值约束偏好 $P$ 的作用下, $x<_P y$  iff  $distance(x.c,v) > distance(y.c,v)$ .其中 $x.c,y.c$ 分别为文本 $x$ 和 $y$ 的 $c$ 属性.对于给定的 $v \in \text{dom}(c)$ ,对于任意的 $z \in \text{dom}(c)$ ,定义 $distance(z,v)=|z-v|$ .邻近值约束偏好的理想结果文本是 $c$ 属性等于 $v$ 的文本.如果没有满足这一条件的文本,则属性 $c$ 的值越接近 $v$ ,用户越喜欢该文本.

例: $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ , $C_p=\{\text{文本,页数}\}$ , $R_{pc}=\{\text{attribute\_of},\text{attribute\_of}(\text{文本,页数})\}$ , $V_p=\{10\}$ , $PCON=\{(\text{页数},10,\text{around},1)\}$ . $P$ 是一个邻近值约束偏好,表示文本的页数越接近于10则用户越喜欢.

**定义18(串相等约束偏好).**  $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ 是一个串相等约束偏好,其中 $PCON=\{(c,v,cond,L)\}$ , $c \in C_p$ 是一个字符型属性, $v \in V_p$ 为对 $c$ 进行约束的字符串, $cond$ 为 $c$ 和 $v$ 之间满足的串相等(=)约束, $L=1$ 表示 $cond$ 是非严格约束.

在串相等约束偏好 $P$ 的作用下, $x<_P y$  iff  $x.c \neq v \wedge y.c = v$ .其中 $x.c,y.c$ 分别为文本 $x$ 和 $y$ 的 $c$ 属性.满足 $c$ 属性的值等于字符串 $v$ 的文本是用户喜欢的文本.如果满足这个条件的文本集合为空,则考虑不满足条件的文本.

例: $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ , $C_p=\{\text{文本,标题}\}$ , $R_{pc}=\{\text{attribute\_of},\text{attribute\_of}(\text{文本,标题})\}$ , $V_p=\{\text{"个性化服务综"}\}$

述”, $PCON=\{(标题,“个性化服务综述”,=,1)\}$ . $P$ 是串相等约束偏好,表示用户喜欢标题为“个性化服务综述”的文本.

定义19(串包含约束偏好).  $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ 是一个串包含约束偏好,其中 $PCON=\{(c,v,cond,L)\}$ , $c \in C_p$ 是一个字符型属性, $v \in V_p$ 为对 $c$ 进行约束的字符串, $cond=contain$ 为 $c$ 和 $v$ 之间满足的串包含约束条件, $L=1$ 表示 $cond$ 是非严格约束.

在串包含约束偏好 $P$ 的作用下, $x <_P y$  iff  $\forall z(x.c \wedge v \subset y.c)$ .其中 $x.c,y.c$ 分别为文本 $x$ 和 $y$ 的 $c$ 属性.类似于串相等约束偏好,满足 $c$ 属性的值包含字符串 $v$ 的文本是用户喜欢的文本.

例: $P=(C_p,R_{pc},rel_{pc},V_p,PCON),C_p=\{文本,标题\},R_{pc}=\{attribute\_of,attribute\_of(文本,标题)\},V_p=\{“personalization”\},PCON=\{(标题,“personalization”,contain,1)\}$ . $P$ 为串包含约束偏好,表示用户喜欢标题中含有“personalization”的文本.

### 2.2.2 内容约束偏好

内容约束偏好是指偏好约束中只涉及文本的结构概念,根据运算符的不同分为如下两种:

定义20(内容包含约束偏好).  $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ 是一个内容包含约束偏好,其中 $PCON=\{(c,v,cond,L)\}$ , $c \in C_p$ 是文本的一个结构概念, $v \in V_p$ 为对 $c$ 进行约束的字符串, $cond=contain$ 为 $c$ 和 $v$ 之间满足的约束条件, $L=1$ 表示 $cond$ 是非严格约束.

在内容包含约束偏好 $P$ 的作用下, $x <_P y$  iff  $(v \subset x.c \wedge v \subset y.c) \vee (correlation(x.c,v) < correlation(y.c,v))$ .其中 $x.c,y.c$ 分别为文本 $x$ 和 $y$ 的结构概念. $correlation(x.c,v)$ 为结构概念 $c$ 与字符串 $v$ 的相关度.在出现字符串 $v$ 与没有出现过字符串 $v$ 的结构概念中,用户更喜欢出现 $v$ 的结构概念.在出现 $v$ 的众多结构概念中,相关度越高用户越喜欢.

例: $P=(C_p,R_{pc},rel_{pc},V_p,PCON),C_p=\{文本\},R_{pc}=\emptyset,V_p=\{“personalization”\},PCON=\{(文本,“personalization”,contain,1)\}$ . $P$ 是一个内容包含约束偏好,表示用户喜欢包含“personalization”一词的文本.

定义21(语义包含约束偏好).  $P=(C_p,R_{pc},rel_{pc},V_p,PCON)$ 是一个语义包含约束偏好,其中 $PCON=\{(c,v,cond,L)\}$ , $c \in C_p$ 是文本的一个结构概念, $v \in V_p$ 为对 $c$ 进行约束的字符串, $cond=semantic\_contain$ 为 $c$ 和 $v$ 之间满足的约束条件, $L=1$ 表示 $cond$ 是非严格约束.

在语义包含约束偏好 $P$ 的作用下, $x <_P y$  iff  $correlation(x.c,concept(v)) < correlation(y.c,concept(v))$ .其中 $x.c,y.c$ 分别为文本 $x$ 和 $y$ 的结构概念; $concept(v)$ 表示字符串 $v$ 对应的用户偏好概念层次:

$$correlation(c,concept(v)) = \sum_{v_i \in concept(v)} correlation(c,v_i)$$

为结构概念 $c$ 与字符串 $v$ 对应的用户偏好概念层次的相关度,用 $c$ 与字符串 $v$ 对应的用户偏好概念层次中所有语义概念的相关度之和描述.此偏好表明,相关度越高的文本用户越喜欢.

例: $P=(C_p,R_{pc},rel_{pc},V_p,PCON),C_p=\{文本\},R_{pc}=\emptyset,V_p=\{“personalization”\},PCON=\{(文本,“personalization”,semantic\_contain,1)\}$ . $P$ 是一个语义包含约束偏好,表示用户喜欢与“personalization”语义相关的文本.

## 2.3 偏好操作及代数性质

基本偏好不足以表达数字图书馆中用户的复杂偏好.我们把在基本偏好上进行的复合操作称为偏好操作,基本偏好经过一系列偏好操作的结果仍然是偏好,称为复合偏好.下面定义偏好操作以及偏好操作的代数性质.

### 2.3.1 偏好操作

为了便于说明,我们为偏好约束  $pcon=(c,v,cond,L)$  定义相应的数值函数,称为 $\gamma$ 函数.

•  $pcon$  为属性约束

a)  $cond \in \{max,min,around\} : \gamma(pcon) = 1/(1+distance(c,v))$ .

其中, $distance(c,v) \geq 0$ ,表示数值型属性概念  $c$  与其值  $v$  之间的距离, $v$  为属性  $c$  的最大值、最小值或邻近值.

b)  $cond \in \{<,\leq,=,>,\geq,\neq,contain\}$ :

$$\gamma(pcon) = \begin{cases} 0.5, & \text{如果 } pcon \text{ 为真} \\ 0, & \text{否则} \end{cases}$$

•  $pcon$  为内容约束

a)  $cond \in \{contain\}$ :  $\gamma(pcon) = correlation(c, v)$ .

其中,  $0 \leq correlation(c, v) \leq 1$  为结构概念  $c$  与值  $v$  之间的相关度, 用 IFIDF 模型定义.

b)  $cond \in \{semantic\_contain\}$ :  $\gamma(pcon) = correlation(c, concept(v))$ .

其中,  $concept(v)$  为字符串  $v$  对应的用户偏好概念层次;  $correlation(c, concept(v))$  为结构概念  $c$  与  $concept(v)$  之间的相关度, 用  $c$  与  $concept(v)$  中所有语义概念的相关度之和描述.

•  $pcon$  为复合偏好约束

a)  $cond \in \{\wedge\}$ :  $\gamma(pcon) = \gamma(pcon_1) + \gamma(pcon_2)$ .

其中,  $\gamma(pcon_1)$  和  $\gamma(pcon_2)$  分别为两个组合偏好的终结约束的  $\gamma$  函数.

b)  $cond \in \{+\}$ :  $\gamma(pcon) = pcon$ .

定义 22(与操作). 设  $P_1 = (C_{p1}, R_{pc1}, rel_{pc1}, V_{p1}, PCON_1)$  和  $P_2 = (C_{p2}, R_{pc2}, rel_{pc2}, V_{p2}, PCON_2)$  是两个非严格约束偏好,  $PCON_1$  中的终结约束  $pcon_1 = (c_1, v_1, cond_1, L_1)$ ,  $PCON_2$  中的终结约束  $pcon_2 = (c_2, v_2, cond_2, L_2)$ .  $P_1$  和  $P_2$  的与操作的结果  $P = P_1 \wedge P_2$  也是一个非严格约束偏好, 记为  $P = (C_p, R_{pc}, rel_{pc}, V_p, PCON)$ . 其中,  $C_p = C_{p1} \cup C_{p2}$ ,  $R_{pc} = R_{pc1} \cup R_{pc2}$ ,  $rel_{pc} = rel_{pc1} \cup rel_{pc2}$ ,  $V_p = V_{p1} \cup V_{p2}$ ,  $PCON = PCON_1 \cup PCON_2 \cup \{(pcon_1, pcon_2, \wedge, 1)\}$ .  $P = P_1 \wedge P_2$  称为  $P_1$  和  $P_2$  的与偏好.

在与偏好  $P$  的作用下, 对于任意两个文本  $x$  和  $y$ ,  $x <_P y$  iff  $x <_{p1} y \wedge x <_{p2} y$ . 在与偏好中, 用户在两个文本  $x$  和  $y$  之间更喜欢文本  $x$  当且仅当在两个偏好作用下  $x$  都好于  $y$ .

定义 23(同等重要操作). 设  $P_1 = (C_{p1}, R_{pc1}, rel_{pc1}, V_{p1}, PCON_1)$  和  $P_2 = (C_{p2}, R_{pc2}, rel_{pc2}, V_{p2}, PCON_2)$  是两个非严格约束偏好,  $PCON_1$  中的终结约束  $pcon_1 = (c_1, v_1, cond_1, L_1)$ ,  $PCON_2$  中的终结约束  $pcon_2 = (c_2, v_2, cond_2, L_2)$ .  $P_1$  和  $P_2$  的同等重要操作的结果  $P = P_1 \cdot P_2$  也是一个非严格约束偏好, 记为  $P = (C_p, R_{pc}, rel_{pc}, V_p, PCON)$ . 其中,  $C_p = C_{p1} \cup C_{p2}$ ,  $R_{pc} = R_{pc1} \cup R_{pc2}$ ,  $rel_{pc} = rel_{pc1} \cup rel_{pc2}$ ,  $V_p = V_{p1} \cup V_{p2}$ ,  $PCON = PCON_1 \cup PCON_2 \cup \{\gamma(pcon_1) + \gamma(pcon_2)\}$ .  $P = P_1 \cdot P_2$  称为  $P_1$  和  $P_2$  的同等重要偏好.

同等重要操作是指组合的两个偏好同样重要. 在同等重要偏好  $P$  的作用下, 对于任意两个文本  $x$  和  $y$ ,  $x <_P y$  iff  $\gamma(x.c_1, v_1, cond_1, L_1) + \gamma(x.c_2, v_2, cond_2, L_2) < \gamma(y.c_1, v_1, cond_1, L_1) + \gamma(y.c_2, v_2, cond_2, L_2)$ , 即两个偏好的终结约束的  $\gamma$  函数之和的大小决定了用户偏好程度的高低.

定义 24(优于操作). 设  $P_1 = (C_{p1}, R_{pc1}, rel_{pc1}, V_{p1}, PCON_1)$  和  $P_2 = (C_{p2}, R_{pc2}, rel_{pc2}, V_{p2}, PCON_2)$  是两个非严格约束偏好,  $PCON_1$  中的终结约束  $pcon_1 = (c_1, v_1, cond_1, L_1)$ ,  $PCON_2$  中的终结约束  $pcon_2 = (c_2, v_2, cond_2, L_2)$ .  $P_1$  和  $P_2$  的优于操作的结果  $P = P_1 \& P_2$  也是一个非严格偏好, 表示  $P_1$  优于  $P_2$ , 记为  $P = (C_p, R_{pc}, rel_{pc}, V_p, PCON)$ . 其中,  $C_p = C_{p1} \cup C_{p2}$ ,  $R_{pc} = R_{pc1} \cup R_{pc2}$ ,  $rel_{pc} = rel_{pc1} \cup rel_{pc2}$ ,  $V_p = V_{p1} \cup V_{p2}$ ,  $PCON = PCON_1 \cup PCON_2 \cup \{\delta_1 \gamma(pcon_1) + \delta_2 \gamma(pcon_2)\}$ , 这里  $\delta_1$  和  $\delta_2$  为重要性因子, 满足条件  $0 < \delta_2 < \delta_1 < 1$ .  $P = P_1 \& P_2$  称为  $P_1$  和  $P_2$  的优于偏好.

优于操作是指组合的两个偏好中的某一个比另一个重要. 在优于偏好  $P$  的作用下, 对于任意两个文本  $x$  和  $y$ ,  $x <_P y$  iff  $\delta_1 \gamma(x.c_1, v_1, cond_1, L_1) + \delta_2 \gamma(x.c_2, v_2, cond_2, L_2) < \delta_1 \gamma(y.c_1, v_1, cond_1, L_1) + \delta_2 \gamma(y.c_2, v_2, cond_2, L_2)$ . 在优于偏好中, 通过为组合的两个偏好的终结约束条件的  $\gamma$  函数分配不同的重要性因子使得两个偏好拥有不同的地位.

### 2.3.2 代数性质

$P, P_1, P_2, P_3$  表示偏好, “ $\wedge$ ”, “ $\cdot$ ”, “ $\&$ ” 分别表示与、同等重要和优于操作. 上述偏好具有如下代数性质.

自反性: 对于任意一个偏好  $P$ , 有  $P \wedge P = P$ ,  $P \cdot P = P$  且  $P \& P = P$ .

对称性:  $P_1 \wedge P_2 = P_2 \wedge P_1$ ;  $P_1 \cdot P_2 = P_2 \cdot P_1$ .

传递性:  $P_1 \& P_2$  且  $P_2 \& P_3$ , 则有  $P_1 \& P_3$ ;  $P_1 \cdot P_2$  且  $P_2 \cdot P_3$ , 则有  $P_1 \cdot P_3$ .

结合律:  $(P_1 \wedge P_2) \wedge P_3 = P_1 \wedge (P_2 \wedge P_3)$ ;  $(P_1 \cdot P_2) \cdot P_3 = P_1 \cdot (P_2 \cdot P_3)$ ;  $(P_1 \& P_2) \& P_3 = P_1 \& (P_2 \& P_3)$ .

## 3 与现有偏好模型的比较

到目前为止, 人们只提出一种偏好模型, 即基于偏序关系的偏好模型. 偏序偏好模型是针对关系数据提出的. 关系数据具有结构化的特点, 每一个元组用若干个属性值来标识. 用户能够方便地通过指明属性的取值来表达自己的偏好. 文本数据包含复杂的结构和语义信息, 用户无法简单说明更喜欢满足什么条件的文本. 正如关系

模型无法描述文本的复杂结构,基于关系的偏序模型也不能描述复杂的文本偏好.例如,“文本的一个作者的姓为“Alexander”并且标题中含有“personalization”一词并且前两章的内容与个性化相关”.该偏好包括 4 个基本偏好:

$P_1=(C_{p1},R_{pc1},rel_{pc1},V_{p1},PCON_1)$ . 其中,  $C_{p1}=\{\text{文本, 作者, 姓名, 姓}\}$ ,  $R_{pc1}=\{\text{attribute\_of, composed\_of}\}$ ,  $rel_{pc1}=\{\text{attribute\_of(文本, 作者), attribute\_of(作者, 姓名), composed\_of(姓名, 姓)}\}$ ,  $V_{p1}=\{\text{“Alexander”}\}$ ,  $PCON_1=\{\text{(姓, “Alexander”,=,1)}\}$ .

$P_2=(C_{p2},R_{pc2},rel_{pc2},V_{p2},PCON_2)$ . 其中,  $C_{p2}=\{\text{文本, 标题}\}$ ,  $R_{pc2}=\{\text{attribute\_of}\}$ ,  $rel_{pc2}=\{\text{attribute\_of(文本, 标题)}\}$ ,  $V_{p2}=\{\text{“personalization”}\}$ ,  $PCON_2=\{\text{(标题, “personalization”, contain, 1)}\}$ .

$P_3=(C_{p3},R_{pc3},rel_{pc3},V_{p3},PCON_3)$ . 其中,  $C_{p3}=\{\text{文本, 第 1 章}\}$ ,  $R_{pc3}=\{\text{composed\_of}\}$ ,  $rel_{pc3}=\{\text{composed\_of(文本, 第 1 章)}\}$ ,  $V_{p3}=\{\text{“personalization”}\}$ ,  $PCON_3=\{\text{(第 1 章, “personalization”, semantic\_contain, 1)}\}$ .

$P_4=(C_{p4},R_{pc4},rel_{pc4},V_{p4},PCON_4)$ . 其中,  $C_{p4}=\{\text{文本, 第 2 章}\}$ ,  $R_{pc4}=\{\text{composed\_of}\}$ ,  $rel_{pc4}=\{\text{composed\_of(文本, 第 2 章)}\}$ ,  $V_{p4}=\{\text{“personalization”}\}$ ,  $PCON_4=\{\text{(第 2 章, “personalization”, semantic\_contain, 1)}\}$ . 该偏好表示为 4 个基本偏好的复合  $P=P_2\&(P_3.P_4)\&P_1$ . 用 ontology 可以容易地描述该偏好. Ontology 偏好表达的是用户的一种意愿, 需要进一步判断才能确定结果文本中用户更喜欢哪一个. 对偏好查询进行处理后, 得到的结果文本根据用户的偏好程度构成一种偏序关系, 但使用偏序关系却无法表达用户在文本数据上的上述偏好. 图 1 给出了上述偏好用 ontology 表示的结果. 图中概念之间的属性关系用实边表示, 组成关系用实双线边表示, 后继关系用虚边表示, 虚双线边表示基本偏好约束中的约束条件, 用户偏好概念层次中涉及的语义概念之间也用虚双线边连接, 边上的值为概念之间的相关度, 虚波浪边表示两个基本偏好约束之间的组合关系.

综上, 基于 Ontology 偏好模型有以下优点:

能够描述复杂的文本结构. 每个文本都有很多的文本概念, 文本的结构概念构成文本的结构概念层次, 文本的属性概念构成文本的属性概念层次. 每个文本概念都有相应的值. Ontology 能够表示如此复杂的文本结构.

能够描述偏好的语义. 用户偏好中的一个 term 往往包含丰富的语义. Ontology 模型能够描述一个 term 对应的用户偏好概念层次, 使偏好带有丰富的语义.

可扩展性. Ontology 定义了文本结构上的偏好, 不同的文本涉及的概念不是固定的, ontology 偏好能够随着文本结构的扩展而扩展. 自适应性. 随着现实世界知识体系的变化以及用户偏好的变化, 基于 ontology 的偏好模型能够反映当前的语义概念层次以及用户偏好概念层次, 从而自动适应这些变化, 准确表达用户的当前偏好.

### 4 结 论

本文针对数字图书馆中大量的文本资源, 提出一种新的用户偏好模型. 该模型用 ontology 描述用户的偏好, 把用户偏好与语义概念层次相结合, 比原有的基于偏序关系的偏好模型有更强的表达能力, 能够表达数字图书馆中文本数据上的复杂用户偏好. 随着数字图书馆中知识的更新, 知识体系会随之变化, 用户的兴趣也可能发生变化. 本文提出的基于 ontology 的偏好模型使用户的偏好自动随之变化, 具有很好的扩展性和自适应性.

### References:

[1] Zeng C, Xing CX, Zhou LZ. A survey of personalization technology. Journal of Software, 2002,13(10):1952-1961 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/13/1952.htm>

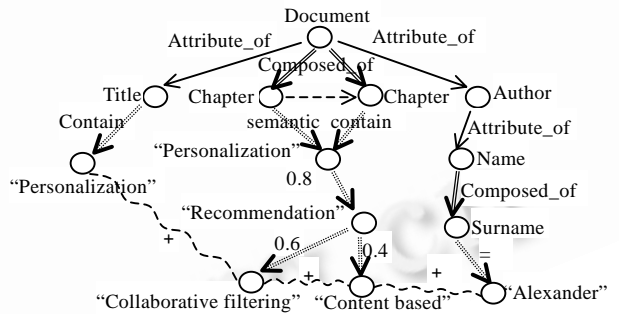


Fig.1 Preference ontology

图 1 偏好 ontology



- [2] Pretschner A, Gauch S. Personalization on the Web. Technical Report, ITTC-FY2000-TR-13591-01, Kansas: University of Kansas, 1999.
- [3] Kießling W. Foundations of preferences in database systems. In: Lochovsky F, Wang S, eds. Proc. of the 28th VLDB Conf. Hong Kong: Morgan Kaufmann Publishers, 2002. 311–322.
- [4] Kießling W, Hafenrichter B. Algebraic optimization of relational preference queries. Technical Report, D-86135, Augsburg: University of Augsburg, 2003.
- [5] Kießling W, Köstler G. Preference SQL-Design, implementation, experiences. In: Lochovsky F, Wang S, eds. Proc. of the 28th VLDB Conf. Hong Kong: Morgan Kaufmann Publishers, 2002. 990–1001.
- [6] Leubner A, Kießling W. Personalized keyword search with partial-order preferences. In: Salgado AC, Edelweiss N, eds. Proc. of the 17th Brazilian Symp. on Database Systems in Cooperation with ACM SIGMOD. Gramado: UFRGS, 2002. 181–193.
- [7] Pretschner A, Gauch S. Ontology based personalized search. In: Bastani F, ed. Proc. of the 11th IEEE Int'l Conf. on Tools with Artificial Intelligence. Washington: IEEE Computer Society, 1999. 391–398.
- [8] Guarino N, Masolo C, Vetere G. OntoSeek: Content-Based access to the Web. IEEE Intelligent Systems, 1999,14(3):70–80.

#### 附中文参考文献:

- [1] 曾春,邢春晓,周立柱.个性化服务技术综述.软件学报,2002,13(10):1952–1961. <http://www.jos.org.cn/1000-9825/13/1952.htm>

\*\*\*\*\*

## 《软件学报》网格计算专刊 征文通知

本专刊瞄准与网格技术及其应用相关的基础性、前瞻性、战略性的重大理论和关键技术问题,以及网格技术与其他学科领域的交叉性,同时为了加强网格界和各应用领域专家之间的合作与交流,共同促进技术进步,推动网格技术在国民经济重大领域的应用普及,提高该领域信息化程度。本专刊面向全国征集论文,欢迎广大学者、专家、工程技术人员积极投稿,现将专刊论文征集的有关事项通知如下:

专刊题目:网格计算

特约编辑:金海(华中科技大学)、郑纬民(清华大学)

#### 一、征文范围

##### 1. 网格基础理论研究

主要包括网格核心协议和核心服务的研究,如服务管理、网格执行管理、元数据管理、网络安全、数据管理等。

##### 2. 网格应用理论研究

针对面向科学活动和信息服务的典型应用领域,用于研究网格对这些领域提供关键技术的支持及改造,以及问题解决环境和使用门户。

##### 3. 网格技术与其他前沿技术的交叉研究

主要研究网格技术与前沿信息技术,如:语义网、对等计算、普适计算、自主计算、服务计算等的交叉结合。

论文内容并不仅限于此,所有与网格计算相关的内容均可投稿。

#### 二、投稿要求

1. 投稿方式:采用“软件学报在线投稿系统”(http://www.jos.org.cn)投稿。请在投稿时,在备注栏中注明“网格计算专刊投稿”字样。

2. 稿件格式:参照《软件学报》论文格式(学报网站上提供了论文模版,可下载)。

3. 投稿文章未在正式出版物上发表过,也不在其他刊物或会议的审稿过程中,且不存在一稿多投现象;保证投稿文章的合法性(无抄袭、剽窃、侵权等不良行为)。

4. 其他投稿须知请参阅《软件学报》投稿指南 <http://www.jos.org.cn/directory.htm>

5. 投稿作者需提交投稿声明;专刊投稿文章不收审理费。录用刊发文章将收取软件学报标准版面费。发表之后,将按软件学报标准支付稿酬,并赠送样刊及单行本。

#### 三、重要时间

截稿日期:2006年6月10日

录用通知发出时间:2006年8月10日

录用修改稿提交日期:2006年9月10日

出版日期:2006年第11期