DOI:10.16749/j.cnki.jecjtu.2014.02.019

文章编号:1005-0523(2014)02-0105-07

基于语义Web的关系型虚拟社区服务发现模型分析

孙玲芳[1],王成文[2],徐 会[2]

(1.泰州学院商学院,江苏 泰州 225300; 2.江苏科技大学经济管理学院,江苏 镇江 212003)

摘要:近几年来关系型虚拟社区服务创新滞后,没能紧随网络用户偏好,发展势头由强转弱。在前人关于服务发现和语义Web应用研究的基础上,构建关系型社区服务发现模型。通过会员信息和第三方服务数据的语义本体化、推理匹配来进行模型仿真,实现了服务集与会员的快速、准确关联,为进行针对关系型社区服务开发和内容创新做准备。

关键词:关系型虚拟社区; 会员兴趣; 语义Web; 服务发现模型

中图分类号:TP391 **文献标志码**:A

随着Web技术的蓬勃发展和日趋成熟,社交网站(SNS)的生存环境日益得到改善,用户之间的交互平台更为广阔。Rheingold(1993)[1]在这种背景下提出了虚拟社区的概念,他指出虚拟社区是社会的集合体,它的发生来自于网络人群通过不断交互而在网络空间里所形成的人际关系网络。Armstrong和Hagel(1997)[2]在此基础上按照人们追求自身四种需求的满足将虚拟社区分为兴趣型,关系型、娱乐型和事务型虚拟社区。其中关系型网络社区作为社会化媒体的主体,发展迅速,像Facebook、Twitter、人人网、开心网、朋友网等关系型网站弥漫网络世界,关系型社区的商业潜力得到持续提升。

然而,虚拟社区不得不应对新形势下用户需求的潜在变化[3]。分析一线社会化媒体(如新浪微博、豆瓣网)的发展现状不难发现其正不断向糅合兴趣、娱乐、学习共享、职场等信息服务的综合性社区迈进[4];其中微博在近三年来发展日益强劲,而关系型社区增速趋缓,用户规模缩减。与微博在服务内容上进行对比发现,这在一定程度上归因于用户偏好在转变,而关系型社区服务内容及其针对的目标群体显得日趋狭窄,服务缺乏创新,在用户个性化需求上把握不到位。面向虚拟社区的研究主要集中在电子商务和知识管理领域,涉及关系型社区的研究也主要集中在社会学方面,针对会员需求的研究在近两年才渐渐出现。本文从培养会员忠诚、促进会员持续参与和加强对新用户的吸引力的目标出发,将结合以前学者针对服务发现的研究和关系型社区的研究及发展现状,借助语义Web技术,构建服务发现模型,来进行会员群体与服务信息的组合,以期达到服务发现的目的。

1 理论依据

1.1 语义Web的概念

将WEB设计成不仅是人与人交互的网络空间,而且能被计算机理解和直接处理的信息网的远景便是语义Web[5]。语义网技术能够对领域知识提供概念化、形式化、明确化的描述,弥补了传统描述语言和工具缺乏语义表达能力的缺点[6]。借助语义Web技术,能够使计算机全面地识别会员信息,并将与其相关联的所有实体都准确地关联起来。

收稿日期:2014-02-25

基金项目:教育部人文社科基金项目(10YJAZH069);江苏省"六大人才高峰"高层次人才基金项目(XXRJ-013)

作者简介:孙玲芳(1963—),男,教授,博士,研究方向为电子商务与信息管理。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

网络本体语言OWL是语义Web的重要组成部分,它通过提供更多具有形式语义的词汇,拥有更多的机制来表达语义[7]。OWL本体包含类元素、属性元素(ObjectProperty和DatatypeProperty)、实例及本体元数据(如本体作者和版本)等信息,各种构造子和公理用于描述类、属性和个体及它们之间的逻辑关系。Protege是构建本体的关键工具,Jena是基于Java语言的开发本体流行平台。

1.2 模型基础

虽然只是一个网络社交平台,但关系型虚拟社区的用户群体日益多元化,用户偏好多样化、网络用户交互要求的实时型越来越强,关系型社区基于传统的线上、围绕社交这个主体的运营模式在新时期亟待调整。

针对Web服务开发,胡朗(2010)[8]构建了一个基于语义的Web服务发现模型,如图1。其模型分为用户接口、业务逻辑和数据存储三层,外层为用户层,业务逻辑层将外层用户请求与语义化的服务提供数据经过解析推理后匹配,然后将结果存入目录服务库中以待查询。研究通过仿真最终实现了Web服务和用户请求的准确、快速匹配。同为互联网应用,这对关系型社区进行服务发现有许多借鉴之处。

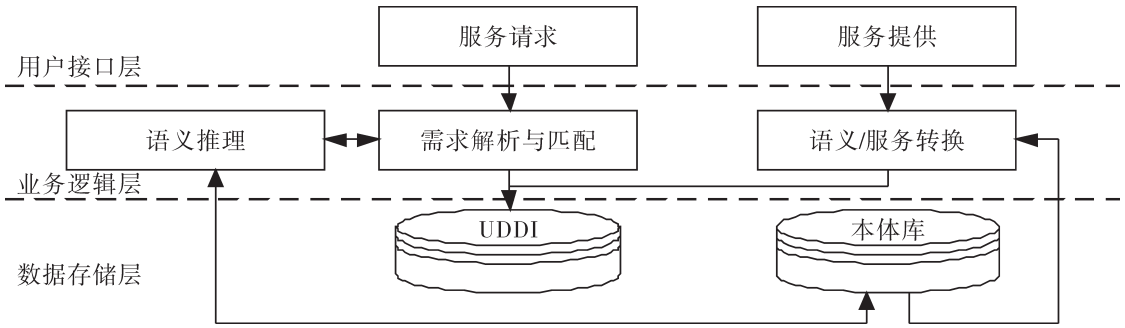


图1 Web服务发现模型

Fig.1 Web service discovery model

易手邦是中微通公司在2010年开发的一款生活信息服务软件,将语义注入网络数据中,并对自然语言进行深层次处理,从而理解用户的真正语言意图,精准快捷地反馈用户所需结果,并推出了独特的分类信息服务,成为国内首款实现生活服务信息即时撮合的平台。其服务机制如图2。

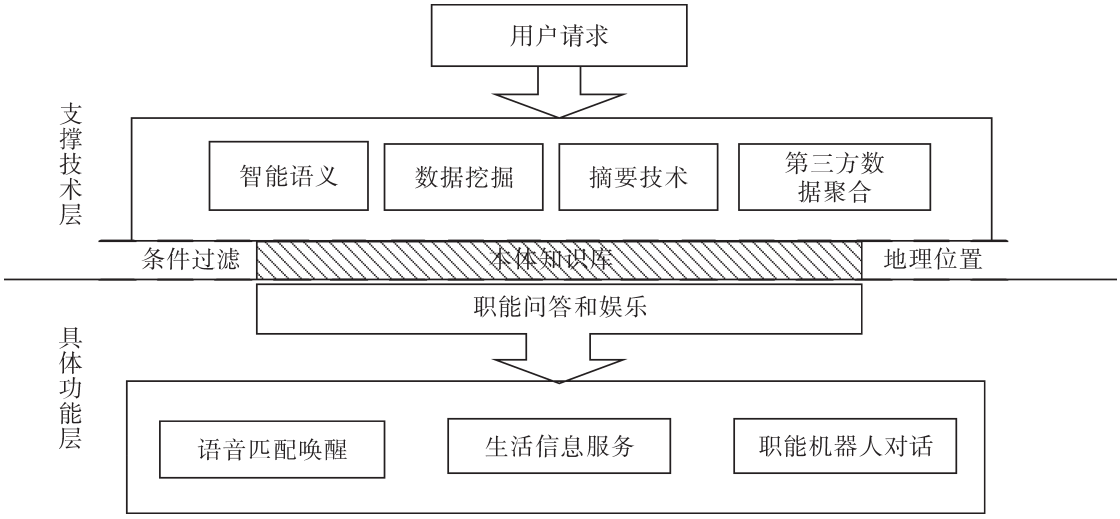


图2 易手邦电子商务一站式服务机制

Fig.2 One-stop service mechanism of Palmdeal E-commerce

易手邦正是借助语义Web技术,将用户请求与自己强大的语义本体词汇库进行检索、语义推理匹配,

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

从而快速准确地将第三方聚合的数据与用户请求相匹配，实现服务发现的。

对SNS客户各自偏好和价值进行分析来获取不同的目标群体，来实现其作为企业实施网络营销和客户关系管理新渠道的功能，进行相关服务的创新，实现与电子商务的联姻，才更能体现新时期SNS的价值[9]。作为SNS服务创新与推进持续发展的践行者，关系型社区和微博一样扮演者重要角色。根据以上分析，本研究将根据几位学者的模型，结合关系型虚拟社区的现状和特征，进行关系型社区服务发现的研究。

2　基于会员兴趣主题的服务发现模型分析

前面的研究主要是根据用户提供的信息和服务数据经过中间件的挖掘推理，最终实现服务发现的。本模型也将从数据整合、语义化，借助语义Web的推理关联，来实现会员群体与服务集的匹配。

2.1　各类实体集之间关系的组织

1）兴趣集的组织。兴趣词汇收集分析与聚类是服务挖掘的基础。兴趣词汇集来源于各领域文献的关键词筛取和一些大型网络社区。他们是原有的领域词汇给会员的推荐，或者是会员自己创造的，有标签的性质。式（1）兴趣实例涉及*Sports*、*Leisure*、*Art*等各领域。

$$\begin{aligned} Hobby=\{h_1,h_2,\cdots,h_n\}\subseteq Category=\{&Sports\cup Leisure\cup Art\cup Literature\cup\\ &Handicraft\cup Instrument\cup PublicInterest\cup Social\cup Chess\cup Science\cup Else\}\end{aligned} \tag{1}$$

其中：*Hobby*为兴趣集，h_i为其实例，*Category*为兴趣类别集。

2）服务集的组织。运用网络爬虫及网络检索工具来搜集来自第三方平台的服务集信息并整理。在*belongTo*，*relatedTo*属性下进行兴趣集和服务集在类别下的关联。服务集组合为。

$$\left.\begin{matrix} belongTo(Hobby,Category)\\ relatedTo(Service,Category)\end{matrix}\right\}\Rightarrow ossociatedWi\,th(H,S)=\{(h,s)|(h,s_{A1}),(h,s_{A2}),\cdots,(h,s_{An}),h\subseteq H,s_i\in S\} \tag{2}$$

其中：H和S分别为被关联的兴趣集和服务源集。

3）目标会员群的组织。依据某一兴趣发现会员兴趣聚类是服务对象群发现的关键。当有两个以上用户同时选择了某项兴趣，这样依据兴趣主题的一个兴趣聚类就形成了。聚集在这个兴趣周围的用户数量越多，这个群落就越壮大，越具有开发潜力。

$$h(M)=\{m|m_i\ hasHobby\ h,m_i\in M,h\in H\}. \tag{3}$$

其中：M和m_i为被聚合会员群体和会员个体，H为兴趣集*Hobby*。

2.2　以兴趣为主题的服务群组合的模型构建

以兴趣为主题，进行会员群体与服务源集的组合是研究的核心。当会员群体、兴趣、服务源集组合在一起时，这样的服务群便形成了。基本关系语义推理公理如下

$$\left.\begin{matrix} hasHobby(Member,Hobby)\\ ossociatedWi\,th(Hobby,Service)\\ location(Member,City)\\ locatedIn(Service,City)\end{matrix}\right\}\Rightarrow HSM(Hobby,Service,Member) \tag{4}$$

$$HSM=(h_i,S_i,M_i)=\{(h,s,m)|(h_i,(s_{i1},s_{i2},\cdots,s_{in}),(m_{i1},m_{i2},\cdots,m_{im}),h_i\in h\subseteq H,S_i\in S,M_i\in M\}$$

其中：h_i为与会员关联的兴趣个体，S_i为与兴趣关联的服务源集，M_i为会员集群。

本研究模型也可以看做是以会员为主导，兴趣为核心，将兴趣与同领域内服务相关联，会员依据兴趣主题进行聚类。概念模型如图3。

本模型首先从会员信息组织和语义化着手，并对第三方的服务数据进行聚合、语义化，一并存入本体库之中。其次，进行（兴趣）主题下的兴趣——服务的语义匹配，然后借助会员信息，实现（兴趣）主题下的会员群与服务集组合。最后，对形成的服务集群进行特定条件下的分群，以实现目标市场的细分，进而对各小群体进行个性化服务开发提供依据。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

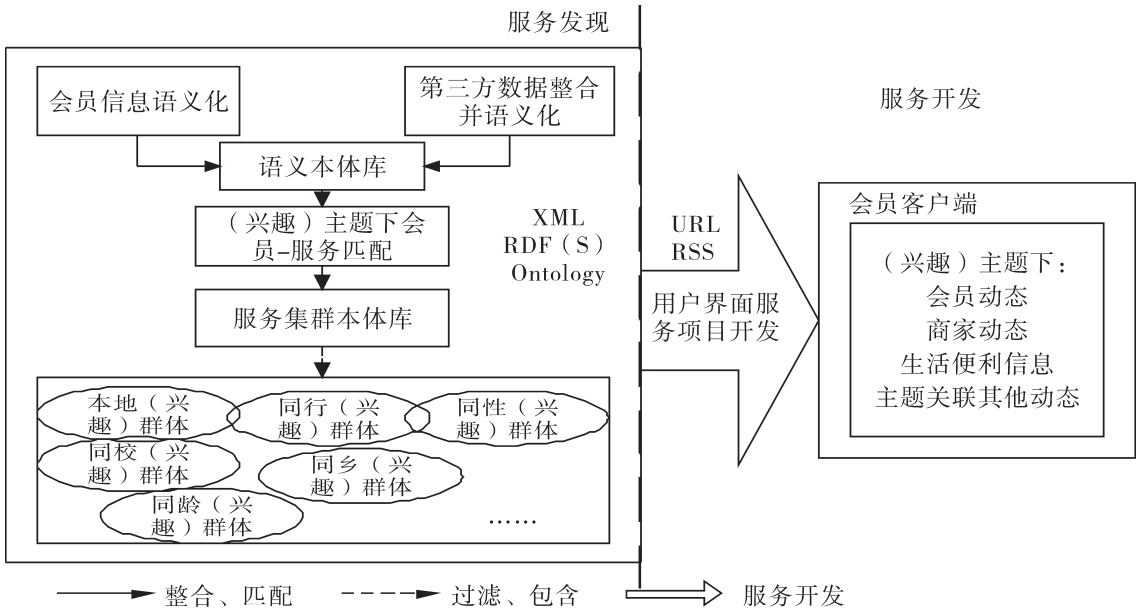


图3 （兴趣）主题下的关系型社区服务集发现模型

Fig.3 Service set discovery model of relational virtual community under the (hobby) theme

信息服务发现的趋势是进行个性化主动信息服务[10]。以兴趣为主题，结合会员的地理地址，借助语义技术，使围绕其兴趣的近在身边的实体服务群体的地址、服务产品最新动态、兴趣领域内动态都以RSS、URL链接或直接地理地址的形式展现在会员主页上，并同时反映兴趣主题下会员群体内的动态，这样的服务组合将使虚拟社区更加紧密地切入会员的现实生活，密切会员关系、加强会员参与频度与凝聚力，正体现了这一机制。

3 模型仿真

3.1 会员和服务集数据的语义化

依据关系型社区中涉及的主体和谓词，在本体编辑工具Protege中分别构建兴趣类Hobby，会员类Member，类别类Category及服务集类Service、所在城市类City，对象属性hasHobby，relatedTo、belongTo、associatedWith、locatedIn、location等，会员数据属性name，age、gender、hometown、occupation等，最后生成会员本体数据文件memOnt.owl，部分本体数据反映到RDF(S)图，如图4。

3.2 基于Jena的仿真实现

Jena是一个Java开发工具包，是开发语义网的应用系统，它已经全面支持SPARQL语言（RDF查询语言），借助SPARQL语言就能够运用Jena对本体进行推理查询，组织本体数据进行语义开发[11]。

针对语义数据的推理查询，SPARQL是查询基于RDF的信息和表示这些结果的一个标准，其核心是以简单图模式为形式的简单查询[12]。本系统涉及到SPARQL的SELECT、WHERE、Filter、OPTIONAL等关键字，他们的功能为：SELECT和WHERE，依据条件决定结果大体的格式，生成一个表作为查询的输出；Filter为过滤器，能够在一个查询中进一步限制匹配的结果；OPTIONAL为可选模式，结果中没有值则置空，便于信息分组[13]。下面进行相关推理查询，4个关键查询的具体实现方法为

```
public void queryMachine(String key)
{           String sql=qHobby/qMember/qCity/qHobby2;
    Query query=QueryFactory.create(sql);
```

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

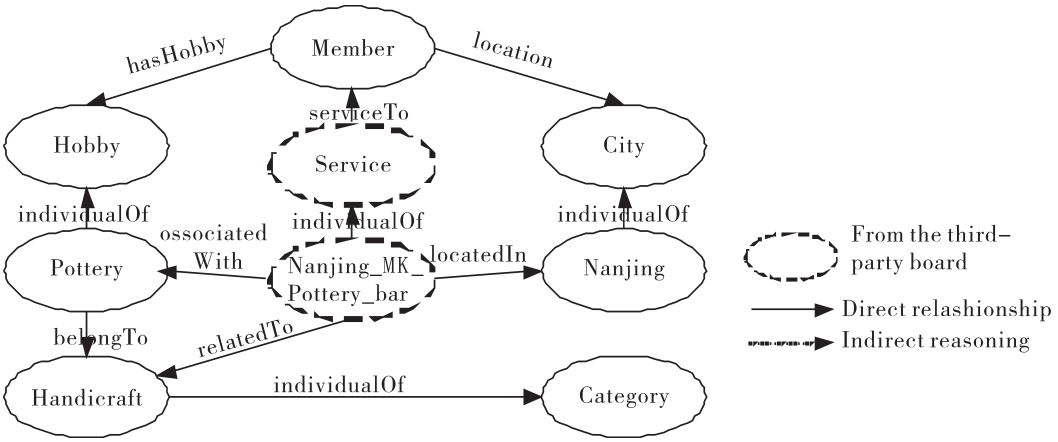


图4　会员语义化信息部分本体数据

Fig.4　Ontology data about member semantic information

```
QueryExecution qe=QueryExecutionFactory.create(query,model);
ResultSet results=qe.execSelect();
System.out.println(key);
ResultSetFormatter.out(System.out,results,query);}
```

其中:qHobby、qCity、qHobby2分别代表实现不同功能的SPARQL查询语句:

查询本体:Static final String base=

“<http://www.semanticweb.org/ontologies/2013/2/Ontology1362645251765.owl#>”;

1)根据兴趣查询会员群体:static final String qHobby=

“PREFIX :”+base+“SELECT ? member WHERE {? member :hasHobby :”+key+“ .}”;

2)根据会员编号查询会员信息:static final String qMember=

“PREFIX :”+base+“SELECT ?”name ? city ? hobby WHERE {? member :name ? name.

? member :location ? city. ? name :hasHobby ? hobby. Filter(? member=:"+key+")}";

3)根据城市查询本地会员群:static final String qCity=

“PREFIX :”+base+“SELECT ? member WHERE{? member :location :”+key+“.}”;

4)根据兴趣查询会员及关联服务的具体信息:static final String qHobby2=“FREFIX :”+base+“SELECT ? member ? serivice ? city ? serviceURL ? serviceAddress ? serviceTel WHERE { ? member :hasHobby :”+key+“. :”+key+“ : associatedWith ? service.” +“? member :location ? city. ? service :locatedIn ? city.” + “OPTIONAL{? service :URL ? serviceURL.}” +“OPTIONAL{? service :specificAddress ? serviceAddress.}” +“OPTIONAL{? service :telephone ? serviceTel.}}”;

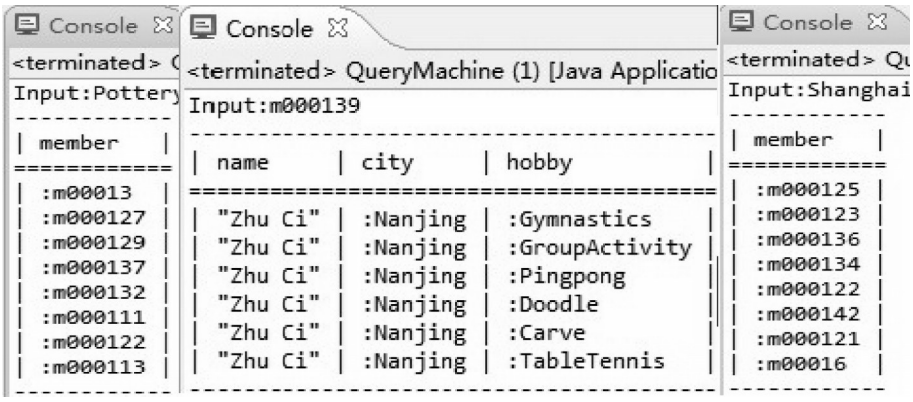
分别进行检测:QueryMachine q=new QueryMachine()、q.queryMachine(“Pottery”)、q.queryMachine(“m000139”)、q.queryMachine(“Shanghai”)、q.queryMachine(“Volunteer”)。

结果快速准确地实现了服务信息与会员的匹配,如图5所示。这个系统的原始数据不用建立很多张表,只需依对象属性添加实例,扩充语义本体数据即可,当数据规模在一定范围内使,只需内存读取数据文件,运行速度快,数据量过大时,可借助关系型数据库进行数据文件的存储。

4　结论

在以传统的现实关系为纽带的关系型社区向综合性社区发展的背景下,研究基于偏好的服务发现对开发对路的网络应用服务具有重要意义。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net



（1） （2） （3）

（4）

图5 基于Jena的语义查询结果

Fig.5 Semantic search results based on Jena

本研究的研究成果在于以用户兴趣偏好为核心的模式，扩展了社区关系网络，同时，借助语义本体技术使得概念更明确，信息个体之间的联系更清晰，对会员隐含信息的把握和分析也比较准确。

其次，语义技术在会员信息中的运用发现了会员间围绕某一主题的潜在关系网，同时将会员群体在主题下与其喜好相关的现实服务联系起来，来为其提供兴趣相关事物的动态信息服务，达到关系型社区服务创新的目的。

参考文献：

[1] HOWARD RHEINGOLD. The virtual community: homesteading on the elctronic frontier[M]. Cambridge: Addison Wesley,1993:7.

[2] HAGEL J,ARMSTRONG A. Net gain:expanding markets through virtual communities[J]. Journal of Interactive Marketing,1999,13(1):55-56.

[3] 李星玥.SNS网站的危机——以开心网为例[J].甘肃科技,2012,28(18):27-29.

[4] C WRZUS,M HÄNEL,J WAGNER,et al. Social network changes and life events across the life span:A meta-analysis[J].Psychological Bulletin,2013,139(1):53-80.

[5] 高志强,潘越,马力.语义Web原理及应用[M].北京:机械工业出版社,2009:3-4.

[6] 刁雅静,王志英,葛世伦.基于OWL的制造业知识本体构建研究[J].江苏科技大学学报:自然科学版,2009,23(6):538-542.

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

[7] GRIGORIS ANTONIOU,FRANK VAN HARMELEN. A Semantic Web Primer[M]. Cambridge: MIT Press,2008:113-134.
[8] 胡朗.基于本体的Web服务发现模型及算法研究[D].长沙:湖南大学,2011:21-22.
[9] 徐翔斌,涂欢,王佳强.基于复杂网络的社交网站用户评价模型研究[J].华东交通大学学报,2012,29(5):38-45.
[10] 谢梓琪,谢颖华.一种基于数据挖掘的SNS用户兴趣模型[J].计算机与现代化,2014,2:24-31.
[11] LIYANG YU.A developer's guide to the semantic web[M]. Heidelberg:Springer,2011:491-495.
[12] ARENAS M, PÉREZ J. Querying semantic web data with SPARQL[C]//Proceedings of the thirtieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems,Athens,2011:305-316.
[13] HITZLER P, KROTSZSCH M, SURE Y. 语义Web技术基础[M].俞勇,译. 北京:清华大学出版社,2012:202-204.

Service Discovery Model of Relational Virtual Community Based on Semantic Web

Sun Lingfang[1],Wang Chengwen[2], Xu Hui[2]

(1. Business College of Taizhou University, Taizhou 225300, China; 2. School of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract: In recent years, relational virtual community has been lagged in service innovation, and has lost in keeping pace with network users' preferences, which results in the weakening development momentum. This paper based on previous scholars' study about service discovery and semantic web application, proposes the service discovery model for relational community. And it makes the model simulation by the semantization of member information and the third-party service data, thus establishing rapid and accurate associations between service set and members and providing preparations for the development and content innovation of relational community service.

Key words: relational virtual community; member hobby; semantic web; service discovery model

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net