

一种面向服务的领域特征模型

吴映波¹ 王 旭²

(重庆大学软件学院 重庆 400044)¹ (重庆大学机械工程学院 重庆 400044)²

摘 要 针对面向服务的领域需求表达与规约化组织,提出一种扩展的服务特征概念与服务特征模型,以及基于 OWL 本体描述语言的服务特征本体元模型。扩展的服务特征概念可更为完整地表征面向服务领域的服务需求语义,而扩展的服务特征模型则为面向服务领域需求提供了一种易于扩展的规约化组织模型。模型通过特征关系映射,既克服了传统特征模型的特征组织形式的局限性,也丰富了服务特征间的语义关系,并可更好地支持基于特征的领域需求配置管理与定制。最后,通过给出的一个原型辅助建模工具实现,验证了模型的适用性。

关键词 特征模型,领域分析,面向服务

中图法分类号 TP311.5 **文献标识码** A

Service-oriented Domain Feature Model

WU Ying-bo¹ WANG Xu²

(School of Software Engineering, Chongqing University, Chongqing 400047, China)¹

(School of Mechanical Engineering, Chongqing University, Chongqing 400044, China)²

Abstract To solve the problems of service-oriented domain requirements expression and specification organization, an extended concept of service feature and its model were proposed, and an ontology-meta model of service feature based on OWL (Ontology Web Language) was also presented. The extended concept of service feature can more fully describe the semantic information of service requirements, and the proposed model provides a flexible and extensible organization structure for service-oriented domain requirements. By defining the mapping relationship of service features in the model, the model overcomes the disadvantages of traditional feature model, enriches the semantic relationship of service features, and provides a more realization mechanism for configuration management and refinement of domain requirements based on feature. Lastly, a computer-aided modeling prototype tool was introduced. By developing the prototype tool, the proposed model is proved to be adaptable for service-oriented domain analysis & modeling.

Keywords Feature model, Domain analysis, Service-oriented

1 引言

面向服务不仅是一种新的分布式计算范式,也是一种区别于传统面向结构与面向对象的新的软件开发模式^[1]。面向服务的架构通过一套支持互操作的标准规范,既支持全新面向服务软件系统的构建,也广泛适用于多个遗留分布式系统间的集成应用开发。关于面向服务系统的构建,目前研究主要关注单个服务系统的分析、设计与实现,以及面向服务在特定应用领域的应用问题。在面向服务的系统中,“服务”定义了一个与业务功能或业务数据相关的接口,以及约束这个接口的契约,如服务质量要求、业务规则、安全性要求、法律法规的遵循、关键业绩指标(Key Performance Indicator, KPI)等^[2]。服务具有相对于对象更高的抽象层次、鼓励粗粒度服务接口、强调服务之间的松耦合、服务被设计为尽可能保持无状态、更灵活的服务实现与动态绑定等特性。虽然服务最终实现仍是

以对象的形式,但在面向服务的领域分析中,领域需求的分析均是以服务为中心来开展的,服务分析是获取领域需求的出发点和主要手段。

领域工程作为一种支持系统级复用的方法,关注于一类相似系统或一个特定领域内的软件系统的重用问题^[3]。领域模型是领域工程中领域分析的主要产物,也是领域工程中领域设计与领域实现的基础。现有的领域模型如 FODA^[4], FORM^[5], FeatuRSEB^[6], PuLSE^[7]等主要采用基于特征的领域需求描述与组织方法。特征模型基本上都采用层次的方式来组织特征,各层特征之间以整体一部分关系(Whole-Part Association, WPA)联系在一起^[8]。对于面向对象软件的领域分析,采用 WPA 关系的领域特征模型可以表现对象之间存在的丰富语义,便于领域模型的映射与转化。但对于面向服务的领域分析,这种特征组织结构存在一定的局限性。而且应用领域的多样性决定了每个领域都可能存在各自特有的

到稿日期:2010-07-18 返修日期:2010-10-25 本文受国家高技术研究发展计划(863 计划)(2006AA04A123),重庆市科技攻关计划项目(CSTC, 2010AC2071)和重庆市自然科学基金项目(CSTC, 2008BB2173)资助。

吴映波(1978—),男,博士生,讲师,主要研究方向为软件工程与企业信息化, E-mail: wyb@cqu.edu.cn; 王旭(1964—),女,博士,教授,主要研究方向为企业信息化与现代物流管理。

特征关系,并具有不确定性。现有的特征模型虽然对特征间常见的结构关系、配置依赖和运行依赖关系进行了总结和分析,却没有也无法完整地描述所有领域潜在的多样化特征关系。

为解决传统的领域特征概念无法完整表达面向服务的领域分析中所需表达更为丰富的服务语义问题,以及克服传统领域特征模型的特征组织方式的局限性,本文提出了一种面向服务的领域特征模型。模型提出的服务特征概念,细化扩展了传统的特征概念,可用来充分表达面向服务的领域分析中服务需求语义。通过服务特征的静态依赖与动态依赖关系,更好地表达了服务间更为丰富的语义关系。而通过定义的服务特征映射关系,更利于在领域设计与实现阶段实现领域模型的配置管理与定制。本文也给出了基于 OWL(Ontology Web Language)^[9]的服务特征本体元模型,使之既具有可扩展性,也为进一步支持基于本体推理的模型验证奠定了一定的基础。

2 面向服务的领域特征概念

FODA 方法最早将特征的概念引入到领域工程的研究中,并提出使用特征模型作为组织领域需求的一种方式。特征作为一种领域需求规约的组织方式,是从用户角度对系统进行感知,用特征对领域需求规约进行结构化组织是非常自然的一种手段。后来的众多领域工程方法都受到了 FODA 的影响,使用特征模型来组织描述领域需求。

由于服务作为领域需求特别是业务需求的表达方式,因此基于特征的领域模型需准确使用特征的概念来刻画出服务的功能特性与非功能特性。根据特征概念和面向服务计算中服务的含义,我们给出服务特征的概念的定义。

定义 1(服务特征, Service Feature, SF) 服务特征是面向服务领域需求规约的组织方式,是面向服务系统提供给用户的业务功能服务或业务数据服务,以及需遵循的服务契约。服务特征可由六元组来表示,即 $SF = (id, Tasks, Category, Type, BindingTime, Constraint)$ 。其中,

id: 服务特征的唯一标识符。每个服务特征必须特别唯一地命名,并且相应应该在领域术语词典被定义。

Tasks: 服务特征包含的服务,即说明服务特征中所包含的服务功能,它刻画服务需要“做什么”。对于一个 Web 服务,可以通过 OWL-S^[10] 建立服务特征功能的语义描述,这时 Tasks 属性值可以是一个 OWL-S 规范文档的 URI 引用。

Category: 服务特征的类别。服务特征可分为原子服务特征和组合服务特征。原子服务特征是指不可再进一步分解的服务特征;组合服务特征是两个或两个以上的服务特征通过聚合或编排形成具有一定时序逻辑关系的服务特征。

Type: 服务特征类型。服务特征有 3 种类型:强制类型(mandatory)、必选类型(alternative)、可选类型(optional)。服务特征类型描述的是同一层次的服务特征之间的选择关系,强制类型服务特征反映了领域的共性特征,必选和可选型特征则反映的是领域的变化性特征。服务特征的类型必须明确指定。

BindingTime: 变化性服务特征的绑定时间。根据服务特征在面向服务领域工程生命周期各阶段的绑定时间,典型的变化性特征绑定时间可分为复用时(Reuse-Time)、编译时

(Compile-Time)、装载时(Load-Time)和运行时(Run-time)绑定等。

Artifact: 领域制品,是服务特征在领域设计与领域实现活动中服务特征的具体化体现。在应用不同约束和设计实现策略下,服务特征可体现为不同的领域制品形式,如一个服务构件、某个数据规范定义甚至是一段代码。领域制品形式取决于服务特征绑定时间。

Constraint: 服务特征的定制与配置约束。服务特征定制和配置约束可表示为形式化的逻辑规则形式,以实现自动的约束规则逻辑分析与配置求解。

3 面向服务的领域特征模型

3.1 服务特征模型定义

特征建模是根据特征确认领域内产品的公共性和可变性并将其组织成特征模型的活动,其输出结果是产品线所有潜在产品的紧凑表达,可表示为一个图形化的特征层次结构,即特征模型。特征模型捕获了特征之间的结构关系和概念关系。传统的领域特征模型表示为二元组结构,即 $FM = (F, R)$ 。其中 F 表示特征模型中的特征集合, R 表示特征之间的依赖关系集合。为支持特征模型更好地表达刻画领域需求模型,许多学者对传统特征模型进行了一定的扩展。但目前还存在着缺乏统一形式化描述、不支持针对特定领域的可扩展,以及不能满足在领域工程全生命周期中的领域模型定制需求等问题。

我们提出的扩展的服务特征模型,在静态依赖即结构关系基础上引入服务特征之间的动态依赖关系,并通过服务特征集合与服务特征关系、约束、软件制品集合的幂集之间的映射,表达服务特征在领域工程其它阶段的配置与定制需求,并为建立基于映射关系的服务特征导航、定位与选取等特征管理,以及在特定产品需求场景中服务特征的裁剪与组装提供了实现基础。

定义 2(服务特征模型, Service Feature Model, SFM) 面向服务领域需求的规约化组织模型,可由六元组表示,即 $SFM = (F, R, C, A, Mr, Mc, Ma)$ 。其中,

F(Features): 服务特征模型中的特征集合。

R(Relationship): 服务特征模型中特征之间的依赖关系集合。服务特征间的依赖关系可分为静态依赖和动态依赖。静态依赖反映服务特征之间固有的关系,例如特征之间整体部分的层次关系;动态依赖关系反映特征之间的操作关系。常见的领域特征动态依赖关系可分为顺序依赖、并行依赖、改变依赖等。顺序依赖关系说明了一个特征必须在另一个特征之后激活,并行依赖则说明两个或者两个以上特征必须同时处于激活状态。一个特征可能被另一个特征改变,则归为改变依赖关系。

C(Constraints): 服务特征模型中的非功能性约束规则集合,是服务特征模型在领域工程各阶段中的定制约束规则。约束规则可以针对单个特征进行定义,也可以把多个特征作为约束对象建立总体约束规则。

A(Artifacts): 服务特征模型中特征对应的领域制品集合。在领域工程中,领域制品集合构成了领域资产库。

Mr: F 到 R 的幂集的映射,即 $Mr: F \rightarrow P(R)$,其中 $P(R)$ 表示集合 R 的幂集,且满足

$$\forall r \in R, \exists F' \subset F, |F'| \geq 2 \text{ 和 } \forall f \in F, \exists r \in Md(f)$$

任何一个服务特征依赖关系必然是建立在至少两个或两个以上的服务特征之间。对于任意一个服务特征,其与其它服务特征之间存在着某种或几种特征依赖关系。

$M_c: F$ 到 C 的幂集的映射,即 $M_r: F \rightarrow P(C)$, 其中 $P(C)$ 表示集合 C 的幂集,且满足

$$\forall c \in C, \exists F' \subset F, |F'| \geq 1 \text{ 和 } \forall f \in F, \exists c \in M_c(f)$$

任何一个服务特征约束必然是针对某一个服务特征或是一组服务特征。对于任意一个特征,可以存在着一项或多项的约束定义。

$M_a: F$ 到 A 的幂集的映射,即 $M_a: F \rightarrow P(A)$, 其中 $P(A)$ 表示集合 A 的幂集,且满足

$$\forall f, f' \in F, f \neq f', \text{ 则 } M_a(f) \cap M_a(f') = \emptyset, \text{ 且 } \cup M_a(f) = A$$

不同的服务特征设计实现为不同领域制品,且所有服务特征的设计与实现构成了领域资产库。在领域特征模型的定制中, M_a 反映了服务特征模型中的服务特征在具体产品需求场景下与领域资产库的对应关联关系。

3.2 服务特征模型的本体元模型

严格定义的服务特征模型必须提供相应的机制,将领域工程不同阶段的模型元素联系起来,这是实现特征驱动的领域工程的关键。本体论(Ontology)是实现这一目标的有效途径。本体具有良好的概念层次结构,支持逻辑推理,能从语义和知识层次上描述领域服务特征模型的概念,并通过概念实体之间的关联关系对特征关系进行语义信息表达。OWL 是 W3C 推荐的语义 Web 中的本体描述语言标准,我们利用其中的本体类元素、对象属性元素、数据属性元素及数据类型元素来分别表示服务特征概念中的语义主体、本体类元素之间的关联关系、本体类元素的属性信息(其定义域为本体类元素,值域为数据类型元素)。基于 OWL 本体描述语言标准,定义 2 所定义的服务特征模型的本体元模型如图 1 所示。

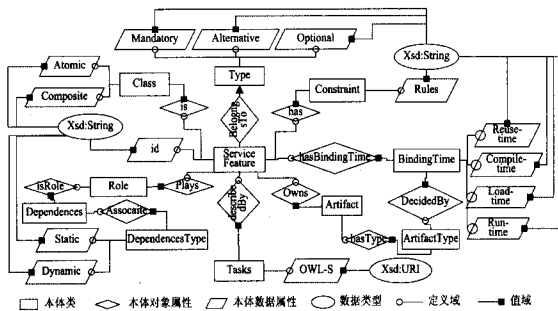


图1 服务特征模型的本体元模型

对图 2 所示的本体元模型主要组成元素简要描述如下。

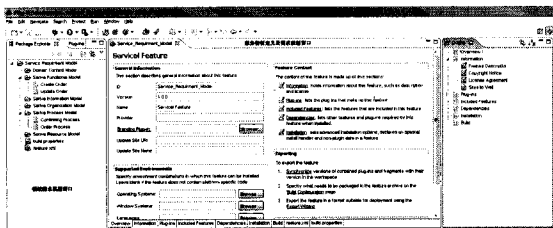


图2 面向服务的领域特征建模原型工具

ServiceFeature:服务特征模型中服务特征概念的本体表达,使用全局唯一的 id 本体数据属性标识特征本体实例。并通过 belongsTo 本体对象属性与 Type 服务特征类型本体类

关联,Type 本体类具有 Mandatory, Alternative 和 Optional 3 种本体数据属性。

Category:服务特征类别本体类。其数据属性分为原子服务特征和组合服务特征。

BindingTime:服务特征绑定时间本体类,包括 Reuse-Time, Compile-Time, Load-Time 和 Run-Time 本体数据属性。

Tasks:服务特征功能描述本体类。其数据属性为 URI 数据类型,值为 Web 服务的 OWL-S 规范文档的 URI 引用。

Artifact:领域制品本体类。通过 hasType 与 Artifact-Type 本体类关联,描述服务特征的领域制品类型。

Constraints:服务特征约束规则本体类。其本体数据属性 Rule 的数据类型可以是字符串类型的形式化规则语句。

Dependence:服务特征间依赖关系本体类。通过 Role 本体类描述服务特征依赖关系的方向,通过 DependencesType 本体类指明依赖关系的类型为静态依赖或是动态依赖。

通过 OWL 建立服务特征模型的本体元模型,从而可以依据本体推理进行特征模型的分析验证,并可对特征模型的定制裁剪提供指导规则。关于如何基于本体推理进行此相关的工作,本文由于篇幅有限暂不做进一步的探讨。

4 面向服务的领域特征建模原型工具

基于本文的服务特征模型,我们开发实现了如图 2 所示的原型工具,辅助进行领域的服务特征建模。本原型工具基于 Eclipse3.2 平台,并通过 Eclipse 插件(Plug-in)开发技术扩展实现了包括特征管理与配置、领域服务需求-服务特征-领域制品间的相互关联等功能,从而可方便实现基于服务特征的领域服务需求分类导航定位、浏览检索与编辑修改,以及服务特征模型的裁剪与定制。经应用于实际工程项目,取得了较好的实际效果。

结束语 领域分析是领域工程的首要关键任务。在传统的领域分析基础上,面向服务的领域分析需考虑面向服务的领域需求分析和建模特点与要求。基于服务特征与服务特征模型的领域需求组织方式不仅可统一和规约化领域的多样化服务需求概念与需求模型,也为在领域范围内系统级的需求复用、领域软件产品开发的需求定制和需求配置管理提供技术实现基础。未来工作可进一步研究在领域工程其它阶段(如领域设计与领域实现阶段),如何面向领域特定应用开发进行模型的定制与裁剪,以及本体理论与形式化方法如何更深入地应用于模型的分析验证。

参考文献

- [1] Papazoglou M P. Service-oriented computing: Concepts, characteristics and directions[C]//Proceedings of the 4th International Conference on Web Information Systems Engineering. Roma, Italy, 2003; 3212
- [2] 毛新生. SOA:原理方法实践[M]. 北京:电子工业出版社, 2007
- [3] 李克勤,陈兆良,梅宏,等. 领域工程概述[J]. 电子学报, 1999, 26(5): 21-25
- [4] Kang K C, Sholom G, Cohen J, et al. A Feature-Oriented Domain (FODA) Feasibility Analysis Study[R]. CMU/SEI-90-TR-21, ESD-90-TR-222. Software Engineering Institute, Carnegie Mellon University. 1990

(下转第 194 页)

BL, PNLH 进行测试, 结果如图 2 所示。图 2(a)~图 2(c) 依次是上述 5 个数据集中在不同 $\lambda\%$ 情况下各方法对于所有测试类别的 $accuracy_{neg}$, $accuracy_{pos}$ 和 $G\text{-means}$ 值的平均变化趋势。

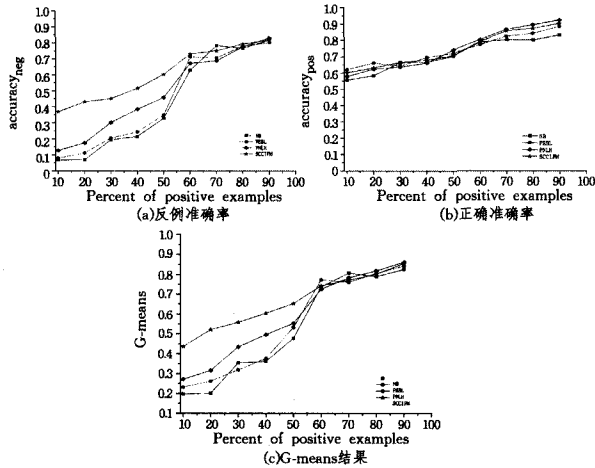


图 2 正例规模对精度的影响

由此可见, 随着正例占整个样本的百分比不断增加, 所采用的 3 个评价指标均在不同程度地增加。通过增加数据集中正例样本的数量, 改变训练数据集所给出的信息, 试图有效提高反例样本分类精度, 由此提高分类器的整体性能。

通过图 2(a) 可以看到, 在反例的分类精度方面, SCCIPM 在正例样本占整个样本的百分比越小时, 较其他方法越显示出较强的分类能力, 从而达到了较理想的分类效果。这一发现对解决“提高分类器精度”这一单分类的瓶颈问题有着重要的意义。

图 2(b) 表明增加正例样本数量对正例样本分类精度影响不是特别显著, 但仍是增加的趋势, 而且一直保持在较高值上。无论正例占整个样本的比例有多大, 正例的分类精度都高于反例的分类精度。在正例样本分类精度方面, SCCIPM 仍然在正例少的情况下分类性能仍优于其它分类器。

进而从图 2(c) 得出, SCCIPM 对样本的整体分类性能最优; 且与其它分类器相比, 不同数据集中整体的样本分布状况、样本规模、特征维度数以及正例占整个样本的百分比等因素对 SCCIPM 影响最小。当正例占整个样本的 50% 以下时, 优于其它分类器; 在正例占整个样本的 50% 以上时, 优势不是很明显, 和其它方法取得了相似的效果, 但 NB 和 PEHL 都不大稳定。这也表明了 SCCIPM 分类性能更稳定, 鲁棒性更强。这充分验证了本文提出的 SCCIPM 模型在单分类问题中具有强的普适性, 从而对从根本上解决单分类问题具有十分重要的意义。

结束语 为了更好地发现单分类问题的本质, 解决“提高分类器精度”这一单分类中的瓶颈问题, 本文提出了基于单分类环境下集成金字塔模型。该模型特点是不需要人工收集反例集合, 只使用标记的正例集合和一个全局的未标记样例集合来构造分类器, 存在着容易收集训练样例集合的优越性, 有着广阔的应用前景。对比实验结果表明, SCCIPM 在单分类上具有较强的分类效果。但模型还需要在更多的真实数据集上进行相关实验, 以充分检验其健壮性和有效性。

参考文献

- [1] Yu H. Single-class classification with mapping convergence[J]. Machine Learning, 2005, 61(1-3): 49-69
- [2] Liu B, Dai Y, Li X L, et al. Building Text Classifiers Using Positive and Unlabeled Examples[C]// ICDM-03. Melbourne, Florida, November 2003: 19-22
- [3] Yu H, Han J, Chang K C-C. PEBL: Positive Example Based Learning for Web Page Classification Using SVM[C]// Proc. Ninth Int'l Conf. Knowledge Discovery and Data Mining, 2003
- [4] Fung G P C, Lu H J. Text Classification Without Negative Examples Revisited[J]. IEEE Transactions on Knowledge and Data Engineering, 2006, 18(1): 6-20
- [5] Zhang Bang-zuo, Zuo Wan-li. Learning from Positive and Unlabeled Examples: A Survey[C]// International Symposiums on Information Processing and 2008 International Pacific Workshop on Web Mining and Web-based Application. Moscow, Russia, May 23-25, 2008: 640-644
- [6] 李和平, 胡占义, 吴毅红, 等. 基于半监督学习的行为建模和异常检测[J]. 软件学报, 2004, 18(3): 527-537
- [7] Figueiredo M A T, Jain A K. Unsupervised learning of finite mixture models[J]. IEEE Trans Patt Anal Mach Intell, 2002, 24(3): 381-396
- [8] Zhang B Z, Zuo W L. A Novel Reliable Negative Method Based on Clustering for Learning from Positive and Unlabeled Examples[C]// Proceedings of the Asia Information Retrieval Symposium 2008(AIRS 2008). Harbin, China, January 2008: 393-400
- [9] Baily T, Jain A K. A note on distance-weighted k-nearest neighbor rules[J]. IEEE Trans. Syst. Man Cybern., 1978, SMC-8(4): 311-313
- [10] 于海龙. 面向 PU 问题的文本分类的研究与实现[D]. 长春: 吉林大学, 2005
- [11] Cortes C, Vapnik V. Support-Vector-Networks [J]. Machine Learning, 1995, 20(3): 273-297
- [12] Han J-W, Kamber M. Data mining: concepts and technique(2nd ed)[M]. Morgan Kaufmann, 2007

(上接第 182 页)

- [5] Kang K, Kim S, Lee J, et al. FORM: A Feature-oriented Reuse Method with Domain-Specific Reference Architectures[J]. Annals of Software Engineering, 1998, 5: 143-168
- [6] Griss M L, Favaro J, D'Alessandro M. Integrating Feature Modeling with the RSEB[C]// ICSR98. Victoria, BC, IEEE, June 1998: 36-45
- [7] Eriksson M, Borstler J, Borg K. The PLUSS Approach-Domain Modeling with Features, Use Cased and Use Case Realizations [C]// Proceedings of the 4th International Conference on Software Product Lines(SPLC'05). LNCS 3714, Springer-Verlag, 2005: 33-44
- [8] 张伟, 梅宏. 一种面向特征的领域模型及其建模过程[J]. 软件学报, 2003, 14(8): 1345-1356
- [9] van Harmelen F A G. Web ontology language, OWL[C]// Staab S, Studer R, eds. Handbook on Ontologies in Information Systems. Springer-Verlag, 2003: 67-92
- [10] Martin D, Burstein M, et al. OWL-S: Semantic markup for Web services[OL]. <http://www.daml.org/services/owl-s/1.1/overview/>, 2003
- [11] 赵毅, 胡丹. FODA: 一种面向特征的领域分析方法[J]. 重庆大学学报学报: 自然科学版, 2004, 18(5): 45-47