

基于OSGi的移动广告平台订单系统^①

杨林, 王晶, 沈奇威, 朱晓民

(北京邮电大学网络与交换技术国家重点实验室, 北京 100876)
(东信北邮信息技术有限公司, 北京 100191)

摘要: 基于组件化构建应用已经成为一种趋势。针对模块划分缺少理论依据和指导, 提出了依靠软件数据模型进行细粒度划分基本单元, 再根据组件之间的服务交互频度和权值对基本单元的相关性进行分析, 最终, 利用模糊图的聚类方法, 将基本单元聚类成模块的模块划分方法。最后, 通过基于OSGi(Open Services Gateway Initiative's framework)的移动广告平台订单系统的应用实例表明了该方法的合理性和有效性。该模块划分方法适用于简单的面向数据模型的应用系统。

关键词: 模块划分; 数据模型; 模糊图聚类; OSGi; 订单系统

Mobile Advertising Order System Based on OSGi

YANG Lin, WANG Jing, SHEN Qi-Wei, ZHU Xiao-Min

(State Key Lab of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)
(EB Information Technology Co. Ltd., Beijing 100191, China)

Abstract: Building applications from components is the ongoing trend. Lack of the theoretical basis and guidance to separate the software modules, the paper represents the module partition approach. It divides fine basic units based on software data model, and then according to the service interaction frequency and weight between modules, It analyses the relativity among the basic units. Finally, it assembles the basic units to module using fuzzy dendrogram. An example named order system in Mobile Advertising platform based on Open Services Gateway Initiative's framework (OSGi) was used to illustrate the rationality and validity. The approach is suitable for simple applications based on data model.

Keywords: module partition; data model; fuzzy dendrogram; OSGi; order system

1 引言

软件的可靠性, 可扩展性等等一直是软件界的持续的研究热点。可靠性可以依赖多机备份等方法^[1], 可扩展性则多依赖于程序的设计过程, 因此, 出现了面向构件的软件开发。

在现代的软件工程实践过程中, 基于模块化的构建应用已经成为一种趋势。模块化的设计方法可以使应用具有可扩展, 可重用的, 可积累的特性。

模块划分是依据模块划分的原则, 将产品划分成

可复用的模块。模块划分是模块化设计中的一项重要工作。模块划分的合理性对模块化产品的功能、性能以及模块的通用程度均有很大的影响。然而, 模块化设计过程中, 模块划分主要以架构师的丰富经验和敏锐的观察力作为依据, 缺少一定的理论指导。为了避免模块划分过程中的隐藏错误和风险, 模块划分需要理论分析和指导。

本文提出一种基于分层的面向数据模型的模块划分方法。在小型面向数据模型编程的应用体系中,

^① 基金项目: 国家杰出青年科学基金(60525110); 国家 973 计划(2007CB307100, 2007CB307103); 国家自然科学基金(60902051); 中央高校基本科研业务费专项资金(BUPT2009RC0505); 电子信息产业发展基金

收稿时间: 2010-07-16; 收到修改稿时间: 2010-08-14

可以利用此方法做模块划分的指导。

基于 OSGi(Open Services Gateway Initiative's framework)的移动广告平台订单系统利用 OSGi 框架技术为基础完成系统的需求。OSGi 框架的一个重要特点就是使得产品具有很强的模块特性。因此,在此系统的设计过程中,模块划分是一个不可或缺的步骤。最终,以本文提出的模块划分方法作为指导,完成了系统的设计。

2 模块划分研究现状

由于模块化在各行各业的广泛应用,模块划分的研究比较火热。由于机械行业的大规模定制化生产的要求,他们对模块化、可复用性要求高,因此在这个领域研究的模块划分方法也层出不穷。

当前的研究中,模块划分方法主要有两种:第一种是自顶向下的,根据一定的原则,划分到合适的模块;第二种是自下向上的,先将系统划分成比较独立的单元,然后计算功能单元之间的交互度,最后根据一定的聚类算法将高耦合的单元组合成模块。

在第一种方法上,有基于公理和模糊图的划分方法。这种方法可以避免大量模块的繁杂聚类过程。但是,这种方法也不是完全的自顶向下的,有可能划分的模块不合理,需要对这些模块继续聚类的过程。

第二种方法的研究比较多样化。首先是划分的原则和面向的角度的研究。模块划分的基本原则是用最少的模块能够组合成最多的产品,就是软件工程里的可扩展和可复用性。但是,由于每个产品自己的特点,具体的划分原则有细微的变化。例如现有的基于生命周期,产品复杂度,继承性的三维模块划分原则,火炮模块划分原则,功能模块划分原则等^[2]。

自下而上的模块划分方法主要分成三步。每步都有不同的研究方法去完成目标。

第一步是划分成基本的单元。由于各个系统选择面向的角色不同,最终划分的方式也不一样。但是在划分成基本的单元的时候,需要考虑功能需求,架构层次,功能流程等多方面的因素。

第二步是计算基本独立单元之间的相关度。相关

度是基本独立单元之间的关联程度、相互影响程度,以此作为模块的聚类依据。

第三步是聚类的过程,将划分的独立单元聚类成模块。聚类的过程有两种,分别是定性的模块聚类和定量的模块聚类过程^[3]。当前定性的模块聚类有启发式算法,基于模糊树图的模块划分算法^[4],模糊聚类算法。定量的模块聚类过程有基于遗传算法的模块聚类,基于模拟退火算法的模块聚类。

作为模块化设计的基础和关键,模块划分一直是国内外学者的研究的热点。但是,由于各个领域均有自己的特色,这些方法在软件工程的模块划分中不具有实际可操作性。因此,需要考虑软件工程实践过程中的各项因素,确认模块划分过程中的影响因子。

3 模块划分方法

小型面向数据模型的应用系统的设计与开发过程中的模块划分具有面向数据模型,应用架构简单,影响因素少等特点,根据这些特点结合当前模块划分的研究方法,确认影响因子,并完善模块划分的过程。最终,本文提出一种合适的模块划分方法。

模块划分的原则是功能之间的高耦合,功能内部的低内聚。该模块划分方法主要分成三个部分。首先,根据需求和软件架构,依赖数据模型划分基本操作单元。其次,根据模块间的调用原则,将基本单元之间的调用关系用相关度计算。最后,选取模糊树图的模块聚类算法将基本单元聚集成模块。

3.1 基本单元划分

软件行业的设计中,不可或缺的一项就是数据模型。每一个应用都有自己的数据模型。应用的创建都是基于数据模型的基本操作完成的。因此,根据这个特色,在基本单元划分的时候依据应用的数据模型,综合考虑软件架构的因素,对基本单元进行划分。

由于一般的软件设计都会有自己的软件架构层次和特殊要求。因此,基本单元划分依据软件设计的数据模型进行划分后,再按照软件架构的因素考虑继续细粒度的划分。

3.2 相关性分析

基本单元的相关性分析以基本单元之间的调用服

务数和服务调用频度作为相关度的基本点。具体的参看公式 1:

$$r_{ij} = \sum_{k=1}^N W_{ij}^k * F_{ij}^k \quad (1)$$

其中, r_{ij} 表示两个基本单元 i, j 之间的相关度, 值越大, 说明关系越密切。N 表示基本单元之间的服务数目, W_{ij}^k 表示 i 单元和 j 单元之间第 k 个服务的权值。这个权值是可以根据设计者的经验来定。 F_{ij}^k 表示 i 单元和 j 单元之间第 k 个服务的频度。由于每个服务的调用次数是不一样的, 因此, 用该项表征。这些相关度可以组成相关矩阵 R 。

3.3 基本单元聚类

根据前面的调研, 可以了解到聚类算法的多样化。而且, 在数据挖掘领域的聚类算法就丰富, 有 k -means 算法、 k -medoids 算法、Clara 算法等。然而, 基于模糊图的聚类算法具有简单可操作的特点, 而且, 软件的模块划分具有易更改的特点, 因此, 在不影响聚类的效果的前提下, 选择简单可操作的算法是可取的。

基本单元之间的关系可以采用模糊关系图 $G=(V, R^*)$ 来表示, V 是一个以 N 个单元为节点的集合, R^* 是由相关矩阵 R 得到的节点之间的模糊相似关系。若 $E = \{e_k = v_i v_j / R^*(v_i, v_j) > 0; v_i, v_j \in V\}$, 则称 $G^*=(V, E)$ 是 $G=(V, R^*)$ 的基础图。边 $v_i v_j$ 赋权值 $w(v_i v_j) = r_{ij} + r_{ji} (i \neq j)$ 。

设 T 是 G 的生成树, 若对于 G 的一切生成树 T' 都有 $\sum_{e \in E(T')} w(e) \leq \sum_{e \in E(T)} w(e)$ 则称 T 是模糊图 G 的最大生成树。最大生成树描述了节点之间的模糊等价关系, 因此可用来对产品进行模块划分, 根据最大生成树中的 $w(e)$, 选取不同的阈值 λ , 对 T 中 $w(e) < \lambda$ 的边进行截割, 得到的各个子树就是 T 在 λ 水平上的一个模块划分方案^[4]。

4 订单系统的设计

基于 OSGi 的移动广告平台订单系统是移动广告平台的子系统, 主要完成了系统的交易流程^[5], 撮合广告主购买广告位的交易实现。

OSGi 是基于 Java 构建模块化和动态化系统的规范^[6]。OSGi 从物理级别支持系统的模块化。每个模块

拥有自己的生命周期, 并且是一个独立的 Java 工程。基于 OSGi 规范开发的应用程序必须完全遵循模块化设计的基本原则。可见, 模块化是 OSGi 规范的立足之本。

基于 OSGi 的移动广告平台的设计与实现依赖于 OSGi 框架技术, 因此模块化分是子系统设计不可或缺的一部分。

首先, 由于订单系统中的数据模型主要包含的数据对象有提案、订单、子订单、子订单排期, 广告内容以及任务。而且, 系统采用传统的 Spring+Hibernate 的形式开发, 主要分成数据访问层(简称 DAO)和业务逻辑层(简称 SL)。因此, 基本单元的划分考虑两个因素: 数据模型和分层的架构。

最终划分的基本单元有: 1)提案 DAO; 2)订单 DAO; 3)子订单 DAO; 4)子订单排期 DAO; 5)广告内容 DAO; 6)任务 DAO; 7)提案 SL; 8)订单 SL; 9)子订单 SL; 10)子订单排期 SL; 11)广告内容 SL; 12)任务 SL。

OSGi 的模块称为 Bundle。Bundle 之间的服务调用有其严格的规定。根据 OSGi 规范和需求的约束, 确认模块间的调用服务频度。

每个模块的基本服务有增加、删除、修改和查看的功能。以这四个服务为中心, 统计出每个服务的频度。如表 1 和表 2 所示。

表 1 数据对象调用服务频度-1

	1	2	3	4	5	6
1		0,1,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
2	1,0,1,1		2,3,3,1	0,0,0,0	0,0,0,0	0,0,1,1
3	0,0,0,0	1,1,0,1		2,3,3,1	2,2,2,1	0,0,0,0
4	0,0,0,0	0,0,0,0	1,1,0,0		0,0,0,0	0,0,0,0
5	0,0,0,0	0,0,0,0	1,1,0,1	0,0,0,0		0,0,0,0
6	0,0,0,0	0,0,0,0	1,0,0,1	0,0,0,0	0,0,0,0	
7	4,4,4,8	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
8	2,0,2,4	6,6,6,8	4,4,4,8	4,4,4,8	2,2,4,6	4,2,4,6
9	0,0,0,0	0,0,0,0	6,6,6,8	2,2,2,2	0,0,0,0	0,0,0,0
10	0,0,0,0	0,0,0,0	0,0,0,0	4,4,4,8	0,0,0,0	0,0,0,0
11	0,0,0,0	0,0,0,0	1,1,1,0	0,0,0,0	4,4,4,8	0,0,0,0
12	0,0,0,0	0,0,0,0	1,1,1,0	2,2,2,4	0,0,0,0	4,4,4,8

表 2 数据对象调用服务频度-2

	8	9	10	11	12
1	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
2	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
3	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
4	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
5	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
6	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
7	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
8		2,2,4,6	2,2,4,6	0,0,0,0	0,0,0,0
9	0,0,0,0		2,3,4,5	1,2,2,2	0,0,0,0
10	0,0,0,0	1,1,1,1		0,0,0,0	0,0,0,0
11	0,0,0,0	1,1,1,1	0,0,0,0		0,0,0,0
12	0,0,0,0	1,1,1,1	0,0,0,0	0,0,0,0	

表中的第 i 行, 第 j 列表示第 i 个基本单元完成增加、删除、修改和查看的操作需要第 j 个对象直接提供的服务频度。

公式(1)显示了如何计算基本单元之间相关度, 根据公式, 令权值均取值 1, 得到的相关矩阵 R 如图 2 所示:

∞	1	0	0	0	0	0	0	0	0	0	0
3	∞	9	0	0	2	0	0	0	0	0	0
0	3	∞	9	7	0	0	0	0	0	0	0
0	0	2	∞	0	0	0	0	0	0	0	0
0	0	3	0	∞	0	0	0	0	0	0	0
0	0	2	0	0	∞	0	0	0	0	0	0
12	0	0	0	0	0	∞	0	0	0	0	0
8	26	20	20	14	16	10	∞	14	14	0	0
0	0	26	8	0	0	0	0	∞	14	7	0
0	0	0	20	0	0	0	0	4	∞	0	0
0	0	3	0	20	0	0	0	4	0	∞	0
0	0	3	10	0	20	0	0	4	0	0	∞

图 1 相关矩阵 R

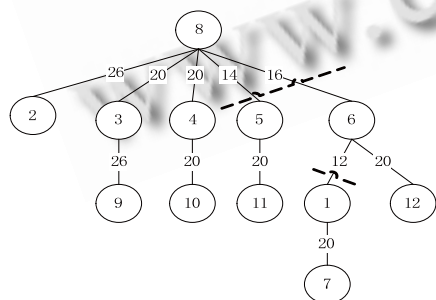


图 2 最大生成树 T

根据矩阵生成的最大生成树 T 如图 3 所示。对执行的结果进行简单的分析, 我们选择阈

值 $\theta=20$, 从而将边值小于 20 的割切, 最后得到四个组成部分 {2, 3, 4, 8, 9, 10}, {5, 11}, {6, 12}, {1, 7}。根据这个模块划分, 我们将系统分成四个大的 Bundle 组件, 提案组件包含提案 DAO, 提案 SL, 订单组件包含订单 DAO, 订单 SL, 子订单 DAO, 子订单 SL, 子订单排期 DAO, 子订单排期 SL, 广告内容组件包含广告内容 DAO 和广告内容 SL, 任务组件包含任务 DAO 和任务 SL。因此, 系统根据其耦合特性划分组件, 不会产生过多的组件, 导致系统的组件冗余, 且组件间可灵活完成任务。根据组件划分结果, 设计系统的总体架构图如图 3 所示。

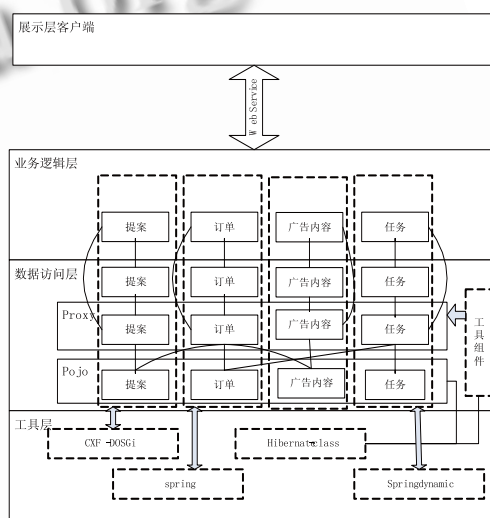


图 3 软件架构图

其中, 虚框表示一个服务组件 Bundle, 细线表示有服务调用关系, 粗箭头表示整体都调用服务组件。

5 总结

本文针对面向数据模型的分层简单应用系统提出一种模块划分方法。首先, 根据软件的数据模型和软件架构层次进行基本单元划分; 再对基本单元需要暴露的服务频度作相关度的分析; 最后, 以模糊图聚类算法将基本单元聚类成模块。文中还利用该方法根据广告平台的订单系统的数据模型, 划分模块, 并设计软件的架构。

此模块划分思路清晰, 方法简单, 适用于简单的面向数据模型的分层应用架构系统的模块划分。因此, (下转第 44 页)

③ 系统主界面

系统由 Java 语言开发, 为了能更好的分类, 数据的提取以城区为单位, 系统运行结果将保存在数据库。系统主界面如图 5 所示。

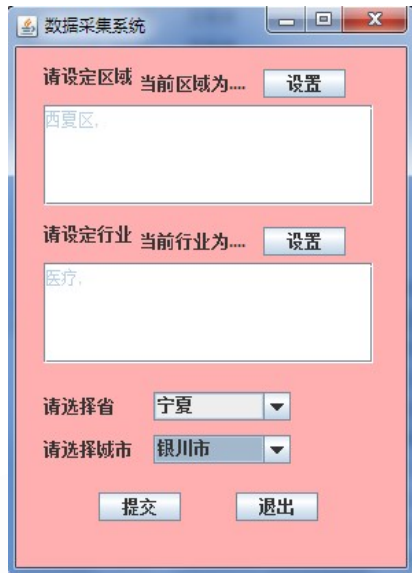


图 4 系统运行界面

④ 待完善之处

对 Google Maps 深入分析发现, 通过程序查询到的数据是以当地政府为中心的一个有限区域的企业数据, 这个区域是以当地政为中心的周边地区, 用一个距离参数来表示, 如图 1 所示。

参考文献

- 1 杨孟辉, 廖建新, 吴乃星. 下一代网络核心业务平台的可靠性分析. 通信学报, 2006, (4): 60-66.
- 2 祁卓娅, 王建正, 韩新民. 模块柔性划分方法. 机械工程学报, 2007, 43(1): 87-94.

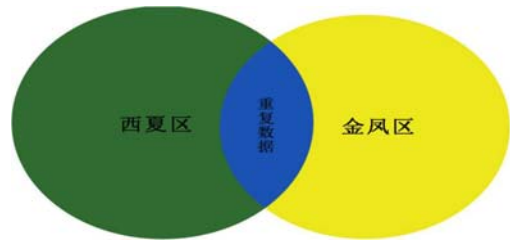


图 5 重复数据示意图

本系统虽然能比较完整的提取所需的数据, 但依然存在一些待完善的地方, 如结果中有大量重复数据存在。

图 5 形象地表示银川的两个城区的重复数据问题, 这个问题是可以解决的, 或在写入数据库时先判断当前数据是否存在, 如果存在就舍弃本次插入操作; 或者在提取的数据中本身就有个表示距离的参数, 通过限定距离来提取数据也能避免重复数据的存在, 不过这样做可能会遗漏部分数据。

4 结语

在信息化时代, 必然离不开互联网这个共享网络, 如何合理、高效地利用这些共享资源是值得研究的课题。本文以从 Google 地图中提取企业黄页信息为例, 简述了提取动态页面信息的过程, 设计和实现了相应的系统原型, 并对系统中待完善之处提出了解决方案。

参考文献

- 1 <http://htmlparser.sourceforge.net/>
- 2 <http://ditu.google.cn/maps?hl=zh-CN&tab=wl>
- 3 <http://allenj2ee.javaeye.com/blog/222454>
- 4 <https://www.ibm.com/developerworks/cn/open-source/os-cn-cr-awler/>

(上接第 35 页)

当系统架构复杂, 涉及的特征和影响因素繁杂时需要变更划分方法中的基本单元划分和相关度分析影响因素。

参考文献

- 3 史俊友, 陶庆斌, 翟红岩. 基于遗传算法的产品族模块划分. 青岛科技大学学报, 2010, (4): 188-193.
- 4 王日君, 张进生等. 基于公理设计与模糊树图的集成式模块划分方法. 农业机械学报, 2009, (4): 179-183.
- 5 廖安舟, 王纯. 移动广告系统的研究与设计. 计算机系统应用, 2009, 18(8): 15-18.
- 6 OSGi Alliance, OSGi Service Platform Core Specification Release 4, April 2007. <http://www.osgi.org>