

# Architecting Publish/Subscribe Applications: A Graph-Oriented Approach

Biao Dong<sup>1, a\*</sup>, Jinhui Chen<sup>2, b</sup>

<sup>1</sup>School of Computer and Software, Nanjing Institute of Industry Technology, Nanjing, 210023, China;

<sup>2</sup>School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, 210044, China.

<sup>a</sup>dongb@niit.edu.cn, <sup>b</sup>cjh@nuist.edu.cn

\*The corresponding author

**Abstract.** This paper presents a novel approach, called GOM, for modeling and implementing the architecture of content publish/subscribe applications using a graph-oriented model (GOM). The GOM constitutes a self-contained piece of content that indivisibly unites three of the content's aspects, namely the content aspect, the semantic aspect, as well as the functional aspect. The model provides means for the integration of ontological knowledge within a GOM's graph structure. A prototype of GOM is introduced to verify the feasibility of author's approach.

**Keywords:** Publish/Subscribe; Graph-Oriented modeling; Multimedia content; Ontology

## 一种面向图的发布/订阅应用架构

董飏<sup>1</sup>, 陈金辉<sup>2</sup>

(1. 南京工业职业技术学院计算机与软件学院, 江苏 南京 210023;

2. 南京信息工程大学计算机与软件学院, 江苏 南京 210044)

**摘要:** 提出了一种面向图的描述和实现技术(称为 GOM)来支持内容发布/订阅应用架构。GOM 由内容所涉及的不可分割的、自包含的内容、语义和功能三方面特征。提供把本体知识集成为 GOM 图结构的方法, 设计实现一个 GOM 发布框架的原型系统, 并对这个原型系统的关键性能指标进行了测试以表明该技术的可行性。

**关键词:** 发布/订阅; 面向图的建模; 内容; 本体

中图分类号: TP 311.1 文献标志码: A

## 引言

目前, 许多分布式系统采用基于发布/订阅(Publish/Subscribe, P/S)的架构。这使得 P/S 系统不再仅仅是一个信息交换的场所, 同时也是一个全球规模的分布式软件应用的基础环境。新型面向内容的 P/S 系统结构复杂, 开发代价高昂。面向内容的 P/S 应用的功能需求常常变化, 用户的使用偏好各异, 有效地共享、协同的 P/S 系统始终是一个开放性问题[1-5]。

RDF、TopicMaps[6-8]等提出了多媒体内容的语义模型, 这些方法把多媒体内容所涉及到的媒体、语义和功能等特征相分离, 不能实现多媒体内容的互换、共享、协同发布。XHTML+SMIL、HyTime、和 XMT、SMIL、SVG 等[9]是多媒体文档模型的标准, 按照时间、空间和交互关系安排基本媒体对象来建模, 主要解决了多媒体内容的媒体特征。这些模型完全依赖于外部的应用, 不能实现 P/S 系统中的提取、重用、自动合成、推荐等要求和技术框架。EJB、CORBA、DCOM 和 Web 服务等中间件提供远程调用对象的功能, 它们没

有提供对发布内容的特别支持。NMM[10]是一个专门的多媒体中间件，使用对象表示一个分布式多媒体框架中的组件，如流服务器，但不能通过对象表现 P/S 系统中的多媒体内容。RDFS、DAML+OIL 和 OWL 等<sup>[16-18]</sup>是描述本体的 web 标准，RDFS 缺少版本支持和对于本体实例的标识，DAML+OIL 没有唯一命名，并且仅仅提供对推理的部分支持，OWL 通过加入建模原语，提供满足需求的表达能力，但缺乏必要的推理支持。它们仅仅在一定程度上适合多媒体内容的 P/S 系统。

与上述系统相比，本文面向图创建一个 GOM 图，为基于内容的 P/S 应用提供共享、协同发布框架，主要特点：

- (1)在逻辑层次上描述物理数据实现，满足面向内容的 P/S 系统对存储容量、CPU 等的较高要求。
- (2)基于元对象建立一个语义网，为发布者创建复杂、上下文敏感和个性化的发布环境。
- (3)订阅者在应用环境中可以使用发布者提供的操作来处理内容。

## 1 应用架构

### 1.1 面向图的 GOM

GOM 由一组相互协作、自包含的成分组成，通过实体、操作和属性等表示 P/S 系统中发布内容。实体包括逻辑媒体、本体对象、联系和容器。逻辑媒体和属性等描述媒体特征，说明内容的组成；本体对象和联系等描述语义特征，建立对象之间的语义联系；操作描述功能特征，提供了任意定义的、与领域有关的能被订阅者的应用程序调用的操作。在 P/S 系统中，GOM 聚集成一个容器，该容器能被序列化成可发布、可交换的格式。

GOM 通过文件的嵌入或者链接，说明所包含的对象。GOM 通过一个和概念图类似的基于图的模型进一步封装了它所包含的媒体对象的语义联系，并且可以具体化联系。GOM 提供一组操作用于处理发布的内容，该操作集能被订阅者调用；发布操作提供了以不同的格式向订阅者发布内容的能力。

### 1.2 本体

本体是一个三元组  $O = \{\Theta, H^O, A^O\}$ ， $H^O \subseteq \Theta \times \Theta$  表示本体对象之间具有子类关系， $A^O$  表示本体公理的集合。

在 GOM 中，本体由本体对象表现，分为对象概念和关系概念。GOM 图中的结点、边分别标记对象概念、关系概念。本体的公理集说明了概念的属性、限制以及概念之间的关系。可以通过元本体对象来描述 GOM 中的关系。本体与 GOM 集成有两种方法：

- (1)按照本体公理，扩展 GOM 图。
- (2)使用本体知识压缩 GOM 图，即从 GOM 中删除任何能被推导的冗余的联系。

在 P/S 系统中，方法 1 提供了有效的订阅匹配和路由，方法 2 能有效的发布 GOM。选择哪一种方法依赖于如应用场景、本体、GOM 的大小以及响应时间、设备的存取空间、带宽等因素。

### 1.3 面向图的本体集成

GOM 建立了一个协同分布式的面向内容的 P/S 应用架构，能被广泛应用于知识发布、电子学习、多媒体任务管理系统等各种应用领域。对于本体集成问题，GOM 通过三方面满足各种应用要求：

- (1)提供建模原语来描述类和类的层次结构，类能被实例化。在说明类和关系时，可以使用限制条件。允许对关系属性赋值，如可逆、对称或传递关系。支持公理。

(2)形式化定义了语义。最大限度的与现存的标准兼容，提供唯一标识符系统，支持版本信息。

(3)本体被作为资源对待，拥有标识符。

GOM 使用一些预定义的元本体对象与典型的本体结构，如子概念、关系概念、逆概念和领域等，这些元本体对象被再用来分类其他本体对象。

## 2 原型实现与性能测试

### 2.1 原型实现

为了适应不同应用场景，GOM 提供两种发布模式：部分模式和实体模式。部分模式指明所发布的内容不包括引用的文件，不包括被递归包含进该 GOM 的内容等。实体模式指明所发布的实体如何被替换，分为强、弱和空三种方法。强方法是实体模式默认的发布方法，包括实体的完整信息。弱方法包括除实体外的所有 GOM 的信息。空方法仅仅包括实体的基本信息，如对象标识符。

面向内容的 P/S 应用架构是一个分布式的 GOM 管理框架，用 java 实现 GOM 图类型，实现了平台独立性，用面向对象的数据库管理系统 MySQL 实现内容的发布。使用 MySQL 作为发布平台有两个方面原因：首先，MySQL 是一个功能完备的数据库服务器，适合大规模内容发布，保证了 GOM 容器的扩展性。其次，MySQL 是一个面向对象的数据库管理系统，能较好地处理 GOM 图类型。

### 2.2 性能测试

对 GOM 原型系统作了性能测试，结果表明其 GOM 所带来的额外开销被控制在一个合理的可以接受的范围以内。为了考察 GOM 与订阅语言的匹配效率，对实体订阅操作的开销进行了测试。实体订阅操作以任意一个给定的实体作为订阅条件，在发布图中找出该实体。

在实验中配置了一台 Intel 酷睿 i5-7200U 处理器、8GDDR3 内存的笔记本，安装 Windows 7 操作系统和 JDK1.8 环境。实现了一个可创建任意大小和拓扑结构的 GOM 图的测试环境。在实验中设定每个图直接嵌套 0 到 50 个子图，每个图最多有 100 个实体。图 1-2 对两种类型的图嵌套深度的操作开销进行测试，普通、深度类型图中的最大嵌套深度分别为 5、10。

图 1 所示的结果是运行 10 次实验后的平均值。其中：低速为 100GOM/S，中速为 300 GOM/S，高速为 500 GOM/S。图 1 中，实体订阅操作的响应时间随着 GOM 的嵌套深度增加而增大。

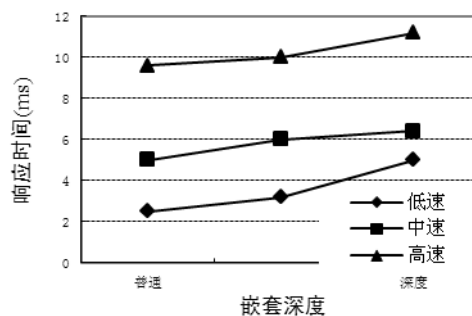


图 1 实体订阅操作的响应时间

这些基本性能测试表明我们的原型系统的实现是较为高效的，同时也表明订阅匹配算法的优化应基于一个真实的应用场景，不能与 P/S 系统的特定应用领域相分离。有很多因素影响订阅匹配效率，关于子图

嵌套深度对响应时间的影响，需要进行进一步的分析。

### 3 结论

本文从内容的表示、语义和功能出发，提出一种面向图的内容发布、本体知识描述和实现技术来支持面向内容的P/S应用架构。该应用架构已被应用于知识发布、电子学习、多媒体任务管理系统等领域。

### 4 致谢

基金项目：江苏省“青蓝工程”资助。

### Acknowledgement

This work was sponsored by Qing Lan project(Jiangsu province, china).

### 参考文献：

- [1] 杨煜尧, 杨宇威, 张扶桑, 等. 构造服务质量感知的发布/订阅系统[J]. 计算机系统应用, 2016(01):48-55.
- [2] Onus M, Richa A W. “Parameterized Maximum and Average Degree Approximation in Topic-Based Publish-Subscribe Overlay Network Design” [J]. Computer Networks, Vol.94(2016), No.C, p. 307-317.
- [3] Jergler M, Jacobsen H A, Sadoghi M, et al. “Safe distribution and parallel execution of data-centric workflows over the publish/subscribe abstraction” [C]// IEEE International Conference on Data Engineering. IEEE, 2016, p.1498-1499.
- [4] AntoniĆ A, MarjanoviĆ M, PripužiĆ K, et al. “A mobile crowd sensing ecosystem enabled by CUPUS: Cloud-based publish/subscribe middleware for the Internet of Things” [J]. Future Generation Computer Systems, Vol.56(2016), p. 607-622.
- [5] Roffia L, Morandi F, Kiljander J, et al. “A Semantic Publish-Subscribe Architecture for the Internet of Things” [J]. 2016, p.1-1.
- [6] Kim J, Shin H, Han W S, et al. “Taming subgraph isomorphism for RDF query processing” [J]. Proceedings of the Vldb Endowment, Vol.8(2015), No.11, p. 1238-1249.
- [7] Alam F, Rahman S U, Khusro S, et al. “A Roadmap for Killer Applications in Resource Description Framework (RDF) and Topic Maps” [J]. Bulletin Et Annales De La Societe Royale Belge Dentomologie, Vol. 27(2015), p. 185-190.
- [8] Berlanga R, Nebot V, Pérez M. “Tailored semantic annotation for semantic search” [J]. Web Semantics Science Services and Agents on the World Wide Web, Vol.30(2015), No.C, p. 69-81.
- [9] Neumann A. “Scalable Vector Graphics (SVG)” [M]// Encyclopedia of GIS. 2016:1013-1021.
- [10] Lohse, Marco, Florian, et al. “Network-integrated multimedia middleware (NMM)” [J]. Proceedings of Acm Multimedia, 2016, p.1081-1084.

### References

- [1] Yang Yuyao, Yang Yuwei, ZhangFusang, et al. “Building a Quality of Service Awareness Pub/Sub System”[J]. Computer Systems and Applications, 2016, No.1, p. 48-55 (in Chinese)
- [2] Onus M, Richa A W. “Parameterized Maximum and Average Degree Approximation in Topic-Based Publish-Subscribe Overlay Network Design”[J]. Computer Networks, Vol.94(2016), No.C, p. 307-317
- [3] Jergler M, Jacobsen H A, Sadoghi M, et al. “Safe distribution and parallel execution of data-centric workflows over the publish/subscribe abstraction”[C]// IEEE International Conference on Data Engineering. IEEE, 2016, p.1498-1499

- [4] Antić A, Marjanović M, Pripužić K, et al. "A mobile crowd sensing ecosystem enabled by CUPUS: Cloud-based publish/subscribe middleware for the Internet of Things"[J]. *Future Generation Computer Systems*, Vol.56(2016), p. 607-622
- [5] Roffia L, Morandi F, Kiljander J, et al. "A Semantic Publish-Subscribe Architecture for the Internet of Things"[J]. 2016, p.1-1
- [6] Kim J, Shin H, Han W S, et al. "Taming subgraph isomorphism for RDF query processing"[J]. *Proceedings of the Vldb Endowment*, Vol.8(2015), No.11, p. 1238-1249
- [7] Alam F, Rahman S U, Khusro S, et al. "A Roadmap for Killer Applications in Resource Description Framework (RDF) and Topic Maps"[J]. *Bulletin Et Annales De La Societe Royale Belge Dentomologie*, Vol.27(2015), p. 185-190
- [8] Berlanga R, Nebot V, Pérez M. "Tailored semantic annotation for semantic search"[J]. *Web Semantics Science Services and Agents on the World Wide Web*, Vol.30(2015), No.C, p. 69-81
- [9] Neumann A. "Scalable Vector Graphics (SVG)"[M]// *Encyclopedia of GIS*. 2016:1013-1021
- [10] Lohse, Marco, Florian, et al. "Network-integrated multimedia middleware (NMM)"[J]. *Proceedings of Acm Multimedia*, 2016, p.1081-1084