

基于 WCF 的商业智能系统的应用研究

林 昕,周永刚,严 华

(安徽邮电职业技术学院 计算机系,安徽 合肥 230031)

摘 要:在信息化过程中,企业需要实时准确地寻找信息模型,不同企业之间知识信息和智能分析能力的共享和交互的需求变得越来越迫切。针对传统商业智能存在移植性、集成性和可扩展性差等弱点,文中研究基于 WCF 的商业智能平台。在分析商业智能当前应用状况和商业智能体系结构的基础上,给出了整个系统比较完善的解决方案,构建了基于 WCF 服务的商业智能的应用平台。该平台可以将商业智能组件无缝集成到企业现有的系统中,增强了可理解性和可扩展性。

关键词:商业智能;WCF;数据仓库

中图分类号:TP311.5

文献标识码:A

文章编号:1673-629X(2010)07-0223-04

Research and Application of Business Intelligence System Based on WCF

LIN Xin, ZHOU Yong-gang, YAN Hua

(Department of Computer, Anhui Vocational College of Posts and Telecommunications, Hefei 230031, China)

Abstract: In the process of information, enterprise demands to search for a real time information model. The sharing and inner operations between different applications have become increasing demand. Focusing at the drawback of traditional business intelligence system including poor transplantation, integration, and expansibility, a platform of business intelligence based on WCF is researched in this paper. A solution for business intelligence system is proposed with the analysis of the current application status and architecture of business intelligence system. WCF-based business intelligence platform is designed and integrated seamlessly into the current application systems, and then comprehensibility and extensibility are improved.

Key words: business intelligence; WCF; DW

0 引 言

商业智能是当前中国的企业信息化应用中最热门的话题之一,它是衡量一个企业能否盈利的一个非常重要的指标。当前,商业智能系统存在的问题主要表现在集成性差和移植性差这两方面,从而使系统缺乏一定的灵活性和可扩展性。Windows 通讯基础(Windows Communication Foundation, WCF)是微软力推的一个架构产品,它是 .NET 平台下基于 SOA(Service Oriented Architecture)技术的产品,代表了下一代软件架构设计与开发的发展方向,它是构建面向服务的分布式系统的技术基础,具有开发性、可扩展性、安全性、并发性等特点^[1]。如何将 WCF 技术和商业智能技术进行整合,从而达到对数据进行高效的处理和分析,为企业决策支持提供服务,更好解决企业信息应用面临

的问题,具有十分重要的意义。

1 商业智能相关概念

商业智能是将数据仓库(DW)、联机分析处理(OLAP)和数据挖掘等多项技术整合在一起应用于企业的信息化建设中,从多数据源中提取相关的数据,然后对数据进行清洗、转换、载入(ETL)数据仓库或数据集市;然后利用 OLAP 分析工具对数据进行分析,这时数据就变成企业决策信息;最后将知识展现给用户,为企业决策提供支持^[2]。可见商业智能并不是新的技术,而是一个解决方案^[3]。

商业智能(BI)的主要技术包括三个部分:数据仓库(DW)、联机分析处理(OLAP)、数据挖掘(DM),其中数据仓库是 BI 的基础,而 OLAP 是对数据仓库中的信息进行初步分析和处理的工具,数据挖掘是对数据仓库中的信息进一步进行知识发现的工具。

数据仓库(Data Warehouse)是一个面向主题的(Subject Oriented)、稳定的(Non-Volatile)、集成的(In-

收稿日期:2009-11-11;修回日期:2010-02-23

基金项目:安徽省自然科学基金项目(KJ2009B192)

作者简介:林 昕(1975-),男,安徽合肥人,讲师,硕士,研究方向为软件工程、数据仓库、OLAP。

egrated)、反映历史变化(Time Variant)的数据集合,用于支持管理决策(Decision Making Support)^[4]。联机分析处理(OLAP)的主要技术是进行多角度的数据分析,它对分析所需要的数据进行有效集成,按多维模型进行组织,以便从多角度、多层次进行分析,并发现趋势。

数据挖掘就是从海量的、随机的、模糊的、实际应用数据中,获取隐含在其中知识的过程^[5]。数据挖掘实质是一种深层次的信息分析方法,它能够在数据仓库中自己寻找规律,本质上数据挖掘是一个知识发现的过程^[6]。

2 WCF 技术

在设计一个分布式系统时主要考虑的因素包括开放性、异构性、可扩展性、安全性、透明性、故障处理和并发性等。从功能上来看,WCF 整合了 ASP.NET Web Services (ASMX)、.NET Remoting、Enterprise Service、Web Services Enhancements (WSE)以及 Microsoft Message Queue (MSMQ)等现有技术的优点^[7]。它提供了构建面向服务的分布式系统的一种架构模型,使软件开发人员可以更加方便地进行分布式应用系统的开发。

2.1 WCF 的基本架构

WCF 支持的技术主要包括并发管理、事务性、可靠性、安全性等技术^[8],WCF 的基本架构如图 1 所示。

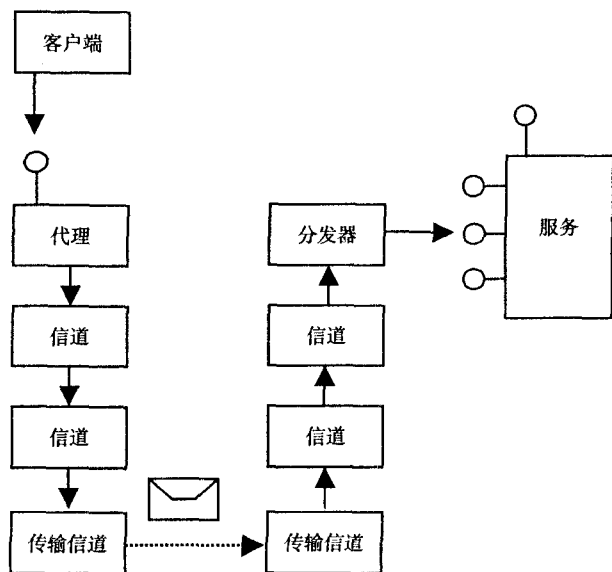


图 1 WCF 构架示意图

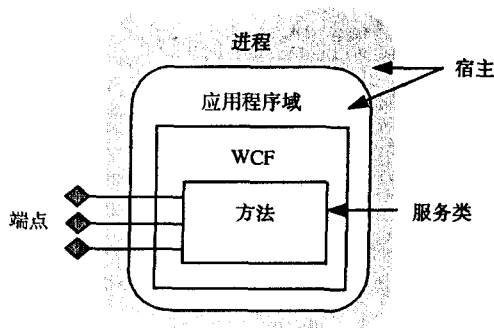
客户端,分布式应用程序通过代理来接收宿主端提供的服务,而客户端的代理拥有和宿主端服务相同的接口^[9]。此外还有附加的一些代理管理方法。这就表示即使这个服务已经存在于客户端机器的内存当

中,客户端也永远不会直接从宿主端调用服务。当客户端的代理接收到应用程序的调用的请求时,它将调用宿主端服务的消息通过信道向下传递。每个信道都会执行相关的调用前预处理,如消息的编码、消息的加密、进行可靠会话等^[10]。客户端最后一个信道负责进行传输的信道,根据配置文件的相关配置约定发送消息的传送方式。

宿主端,消息也是通过信道进行传送。与客户端相关信道一一对应,宿主端信道也要对消息执行相应的宿主端的调用前预处理,如消息的编码、消息的加密、进行可靠会话等。宿主端的最后一个信道负责将客户端的消息发送给消息分配器(Dispatcher),由分配器负责调用服务的实例。

2.2 WCF 技术要点

WCF 是基于 SOA 的一个架构产品,其最大的好处就是能够方便地建立一个服务^[11]。如图 2 所示,一个 WCF 服务主要由三部分构成。



(1)宿主(Host):WCF 服务运行的环境,它可以是进程如 Windows 服务,也可以是应用程序。

(2)服务类(Service Class):标记了[serviceContract]属性的类,在这个类中包含多个方法。除了标记了 WCF 的属性外,这个类与一般的类没什么区别。

(3)端点(Endpoints):端点可以包含一个,也可以是多个,它是 WCF 进行通信的核心部分。

一个 Endpoint 包含三部分:地址(Address),绑定(Binding),契约(Contract)。一般称为是端点的 ABC。

地址(Address):端点的地址指定端点的所在的 IP 位置,如果通过端点与 WCF 进行通信,必须指定端点的网络地址。

绑定(Binding):指定该端点与外界通信的方式,即为端点指定某一种通信协议。

契约(Contract):消息所携带的内容即是契约,同时还描述了消息的操作和组织的形式。

通过以上描述可以理解到端点中的 ABC 代表的含义分别是:what, how, where^[12]。在 WCF 发送消息

的时候,通过地址知道消息传送的位置,通过绑定知道它传送的方式,通过契约则知道传送的消息的内容是什么。

在 WCF 中, ServiceEndpoint 类表示一个端点,在这个类中包含了 Endpoint Binding, Address, Contract Description 类型分别对应端点的绑定、地址和契约。

3 基于 WCF 的商业智能平台设计

在企业内部商业智能作为企业数据仓库的解决方案,发挥了极其重要的作用,但是企业内部的数据仓库信息往往无法满足企业决策者进行分析决策的需求,需要网络分析获得的知识来共同完成企业的辅助决策。基于 WCF 的商业智能就成为解决上述弊端的最佳途径。

3.1 平台的系统结构

文中提出了一种基于 WCF 的商业智能的平台,此平台的结构模式是 B/S 架构。用户界面采用 Web 页面形式,用户通过 Web 页面在浏览器端向服务器端进行服务请求的提交,请求当中包含了数据 ETL, OLAP 数据的查询、数据挖掘等,服务器端负责对浏览器的请求进行处理,随后将处理的结果返回到浏览器端。一个浏览器可以同时访问多个服务器,形成多点到多点、一点到多点的结构模型。基于 WCF 的商业智能平台结构如图 3 所示。

各组成部分说明如下:

(1) 数据源层:也称为操作数据层,是整个数据仓库的基石,提供整个系统原始的数据。通常为企业的业务数据库和相关外围数据,其来源为企业的业务系统,包括关系数据库 (Oracle、DB2、SQL Server、Sybase 等) 以及文本中的业务信息数据。

(2) 数据 ETL 层:由 WCF 端点和数据库的 ETL 服务器组成,主要是将业务数据库中的数据信息经过抽取 (Extract)、转换 (Transform) 和加载 (Load) 到数据仓库或者数据集中。文中所设计的平台,在数据 ETL 服务器中首先按要求定义相关的 ETL 操作,明确数据的来源和目标数据仓库,然后将所定义的相关的 ETL 作业发布为 WCF 的端点来提供 ETL 的服务。数据 ETL 服务也可以连接数据 ETL 服务器进行查询和管理 ETL 作业。

(3) 物理层:采用数据仓库或者数据集市来保存经过处理的相对稳定的业务数据,本平台支持多种数据仓库或者数据集市数据的关系型数据库和 OLAP 服务器。

(4) 逻辑应用层:由基于 WCF 的商业智能服务器以及前端展现服务组成。商业智能服务器具有管理元

数据功能、分析功能、安全功能,是用于连接后台的数据仓库或者数据集市,为前端展现查询、报表、分析、数据整合、绩效管理等相关服务。前端展现服务包含查询 (报表、服务、分析等)、管理 (报表、权限和语义层等)、连接 (分析服务器等) 等功能,发布给企业内部系统用户和企业外部系统用户进行调用。

(5) 表示层:主要负责数据的展现,为方便用户的使用本平台采用了现在流行的 B/S 结构,以便实现各个方位的数据展示,配合数据挖掘技术,使用户关注的信息充分的多角度的进行展现。报表、查询、分析以及图表生成等工作简单方便,创造了一个多方位多角度的数据分析展现环境,使数据的价值得到进一步的提升。用户包含企业内部系统、企业外部系统和个人用户。

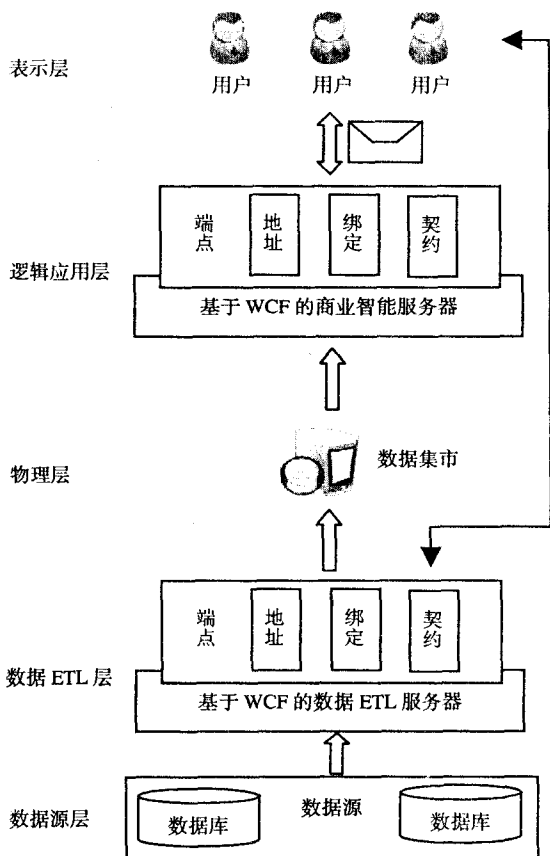


图3 基于 WCF 的商业智能平台

3.2 基于 WCF 的商业智能平台系统的实现

基于 WCF 在商业智能的平台的应用主要表现在数据的抽取、转换、装载过程和数据的前端展现过程,在该平台中,数据的 ETL 和数据的前端展现功能是平台的最为关键的部分。在 WCF 的服务模型中一个服务具有一个或多个端点,通信会被指向到这些提供服务的端点。端点则是由地址、绑定和契约三部分组成。使用 WCF 的服务模型,一般首先定义契约,契约定义

这个服务端点所提供的操作功能,也就是具体的服务内容。定义契约后则是要实现这个契约。在实现契约后是选择契约的承载服务,可以承载的服务包括以下几种:自承载、IIS 和 Windows 进程激活服务(WAS),使用何种方式则根据具体的应用需求。一般说用 IIS 来承载的服务对应 Web 应用。对于需要明确客户端与服务器端之间的进程时则使用自承载、Windows 进程激活服务承载服务(WAS)则是适用于 vista 系统,它支持 WCF 所支持的所有可用的传输协议、队列和端口,IIS 则是只能使用 HTTP 协议。最后指定绑定和地址。

下面给出一段使用自承载方式提供的按名称获取销售量数据服务的实现代码:

(1) 契约定义与实现。

a. 契约的定义:

```
namespace SaleManage{
[ServiceContract]//通过设定此特性说明该接口
为 WCF 一个契约
```

```
public interface ISaleDataManage{
[OperationContract]//通过设定此特性来说明
该接口为 WCF 契约的一个操作
```

```
double GetSaleDataById(string name)
{
}
```

b. 契约的实现:

```
namespace SaleManage{
public class SaleDataManage: ISaleDataManage {
GetSaleDataById (string name){
double data;
...
return data; } } }
```

(2) 服务的自承载。

```
namespace SaleManage{
Class Program{
Public static void main(string[] args){
Type serviceType= typeof( SaleManageType);
Using (ServiceHost host= new ServiceHost(service-
Type)){
host.open();....
host.close();} } }
```

(3) 指定端点的地址和绑定,配置 app.config 文件来实现。

```
<configuration>
<system.ServiceModel>
<services>
```

```
<service name = "SaleManage. SaleDataManage"
behaviorConfiguration = "SaleManageService" >
<host>
<baseAddresses>
<add baseAddress = "http://localhost:8080/Sale-
Manage" />
</baseAddresses>
</host>
<endpoint address = "GetSaleDataById" binding =
"basicHttpBinding" contract = "SaleManage. SaleData-
Manage" >
</service>
</services>
<behaviors>
<serviceBehaviors>
<behaviors name = "SaleManageService" >
</behavior>
</serviceBehaviors>
</behaviors>
</system.ServiceModel>
</configuration>
```

4 结束语

当前商业智能率先在政府、金融、制造等行业大规模地应用,未来经济的发展越来越离不开商业智能预测分析系统与互联网的快速性、便捷性的结合,WCF 是分布式领域的一个新技术。关注 WCF 技术的发展并在实际项目中合理运用是十分必要的。介绍了 WCF 的基本思想以及使用 WCF 构建商业智能系统的体系结构与技术要素,这个平台的应用必将带动企业决策分析能力达到一个更高的程度,快速提升企业的综合实力和竞争力。

参考文献:

- [1] Lowy J. WCF Service 编程[M]. 英文影印版. 南京:东南大学出版社,2007.
- [2] 刘桂峰,崔志明,陈建明. 智能管理决策门户体系结构研究[J]. 计算机技术与发展,2007,17(10):195-197.
- [3] 韩清池,陈世权. 商业智能及其应用的研究与发展[J]. 现代管理科学,2006(3):68-69.
- [4] 林 宇. 数据库原理与实践[M]. 北京:人民邮电出版社,2003.
- [5] Thomsen E. OLAP 解决方案:创建多维信息系统[M]. 第2版. 北京:电子工业出版社,2004.
- [6] 朱 明,王丽珍,周丽华. 数据挖掘[M]. 合肥:中国科学技术

- (1) 将区域 A 分为 $r \times r$ 的子区;
- (2) 计算边长为 r 的方柱面中包含边界的立方体个数: $n(r) = (\max I_i - \min I_i) / r + 1$;
- (3) 求 $N(r) = \sum_A n(r)$;
- (4) 通过改变 r 的大小, 得到一系列 $N(r)$;
- (5) 在双对数坐标系中, 求 $\lg N(r)$ 对 $\lg(r)$ 的斜率, 即为此曲面的分形维 D 。对图 2(a) 切削磨粒纹理分形维数应用该盒维数法进行计算 $\lg N / \lg(r)$, 得到其纹理分形维数 $D = 2.315$ 。

3 实验结果与分析

用上述方法对大量典型磨粒进行两种分形维数计算, 得到了各种典型磨粒两种分形维数范围, 见表 1。

表 1 边缘轮廓和表面纹理分形维数计算结果

磨粒类型	轮廓分形维数	表面纹理分形维数
正常磨损磨粒	1.180~1.250	2.180~2.250
切削磨粒	1.050~1.176	2.290~2.350
疲劳剥落磨粒	1.132~1.190	2.220~2.290
球状磨粒	1.010~1.042	2.120~2.170
层状磨粒	1.196~1.350	2.320~2.410
严重滑动磨粒	1.060~1.165	2.360~2.430

(1) 不同类型磨粒的边缘轮廓分形维数有其特征范围, 同种类型的磨粒轮廓分形维数比较集中。其中, 由于球状磨粒的边缘轮廓相对最为光滑, 因此计算所得的维数最小; 切削磨粒、疲劳剥落磨粒和严重滑动磨粒三种类型磨粒的轮廓分形维数比较接近, 原因为切削磨粒、疲劳剥落磨粒和严重滑动磨粒边缘轮廓比较相似, 都比较光滑。疲劳剥落磨粒轮廓分形维数相对稍大, 说明其有明显曲折的轮廓; 层状磨粒的轮廓分形维数最大, 说明层状磨粒轮廓最不规则。而分形维数的物理意义是反映该分形图形的复杂程度, 这和实验结果分析是一致的。

(2) 从实验数据可以看出, 球状磨粒表面纹理光滑, 纹理分形维数最小, 且变化范围也小; 严重滑动磨粒纹理分形维数最大; 切削磨粒和层状磨粒纹理

分形维数较大; 正常磨损磨粒分形维数较小。这样的计算结果和磨粒形成机理及理论分析是一致的。

4 结束语

针对铁谱磨粒的特点, 通过研究并应用两种分形维数的计算方法, 对大量典型磨粒分形维数进行计算, 根据数理统计原理, 研究了 6 种典型铁谱磨粒边缘轮廓及纹理分形维, 得出了相应的分布范围, 为磨粒识别提供了新的特征参数。

参考文献:

- [1] 王汉功, 陈桂明. 铁谱图像处理与识别技术[M]. 北京: 科学出版社, 2005.
- [2] Kirk T B, Stachowiak G W. Fractal parameters and computer image analysis applied to wear particles isolated by ferrography[J]. Wear, 1991(145): 347-365.
- [3] Podsiadlo P W, Stachowiak G W. Evaluation of boundary fractal methods for the characterization of wear particles[J]. Wear, 1998(217): 24-34.
- [4] Yuan C Q, Li J, Yan X P, et al. The use of the fractal description to characterize engineering surfaces and wear particles[J]. Wear, 2003(255): 315-326.
- [5] 张济忠. 分形[M]. 北京: 清华大学出版社, 1997.
- [6] 葛世荣, 朱 华. 摩擦学的分形[M]. 北京: 机械工业出版社, 2005: 283-296.
- [7] 葛世荣, 张晓云. 铜-铜磨损过程中磨粒积聚的分形研究[J]. 摩擦学学报, 2001, 21(3): 123-125.
- [8] Falconer K J. Fractal Geometry: Mathematical Foundation and Application[M]. New York: John Wiley and Sons, 1990.
- [9] 韩 炎. 分形理论及信号检测技术研究[D]. 南京: 南京理工大学, 2004.
- [10] 陆永耕. 磨粒图像纹理分形特征的研究[J]. 微型电脑应用, 2000, 16(12): 40-41.
- [11] 王 备, 王继承. 图像分割中模糊聚类数目的确定[J]. 计算机技术与发展, 2007, 17(10): 162-164.
- [12] 王 伟, 高 亮, 吴 涛. 一种基于模糊聚类的离散化方法[J]. 计算机技术与发展, 2008, 18(3): 53-55.
- [13] 焦春林, 高满屯, 史仪凯. 基于改进型聚类神经网络的图像分割[J]. 计算机工程与应用, 2007, 43(20): 93-95.
- [10] 严 商, 樟 灿. WCF: Windows 平台新一代通讯基础研究与分析[J]. 计算机与数字工程, 2008(4): 86-89.
- [11] 年福丰, 刘秋让. WCF 的异构数据源集成系统的研究和实现[J]. 科学技术与工程, 2009(11): 3316-3319.
- [12] Klein S. Professional WCF Programming. NET Development with the Windows Communication Foundation[M]. [s. l.]: Wiley Publishing, 2007.

(上接第 226 页)

术大学出版社, 2002.

- [7] Bustamante M L. 学习 WCF[M]. 英文影印版. 南京: 东南大学出版社, 2007.
- [8] 闫 冰. 基于 WCF 的分布式应用开发[J]. 电脑知识与技术, 2008(19): 68-70.
- [9] 吴良清, 黄 黎, 陈建明. 基于 Web Services 的商务智能研究[J]. 计算机与现代化, 2008(1): 33-35.