

基于 VSTO 的智能文档在企业环保信息调查系统中的应用^①

顾 晟 (泰州职业技术学院 江苏 泰州 225300)

陈 波 (南京师范大学 数学与计算机科学学院 江苏 南京 210097)

摘要: VSTO 2005 是内建于 Microsoft Visual Studio 中用于开发基于 Office 解决方案的一种工具,提出了一种基于 VSTO 的智能文档技术,实现了 Office 办公软件与管理信息系统的融合,并成功地应用在企业环保信息调查中。各企业用户可在熟悉的 Office 界面中填写信息,避免填写纸质表格的麻烦;利用文档变量实现的文档状态设置和文档“签名”的功能,可确保入库企业信息的准确、合法与安全,也为在此基础上提供的打印和分析提供了保证;本系统还能减少网络信息采集中频繁占用服务器资源的耗费。实际应用表明本系统具有良好的实用价值。

关键词: VSTO 智能文档 Office Visual studio .NET 系统安全

Application of Intelligent Document Based on VSTO to Enterprise Environmental Information Survey System

GU Cheng (School of Geography Science, Nanjing Normal University, Nanjing 210097, China)

CHEN Bo (School of Computer Science, Nanjing Normal University, Nanjing 210097, China)

Abstract: VSTO 2005(Microsoft Visual Studio 2005 Tools for the Microsoft Office System) are built-in tools, which are used to develop the solution of Office. This paper proposes an intelligent documents technology based on VSTO, and implement the integration of the Office and the management information system. It was successfully applied to the investigation of the state of environmental protection of the enterprises. Various enterprises users may fill the information on the familiar Office interface, avoiding filling the information on the papers. With the documents variable we set documents condition and implement documents "signature". So the presented technology may guarantee the accurateness, validity and security of the stored enterprise information. Also it can promise the dependability of the print and analysis processes. This system can reduce the cost of the server resource during collecting information in network. The running results of this application system show it has important practical significance and application value.

Keywords: VSTO; intelligent document; office; visual studio .NET; system security

在对企业进行诸如环保等信息的调查中,传统方法是将分布在多个地理位置上的企业相关数据,通过纸质或电子文件传输到汇总点,再由工作人员手工将数据录入汇总。这种传统操作模式,缺乏电子文档信息到数据库表的存入、提取的自动化;缺乏对所

收集数据格式的规范统一;缺乏对数据在传输过程中的安全保障;同时也增加了录入人员的工作负荷;延长了信息调查的工作周期。近年来,虽然采用 B/S 方式可以一定程度上避免人工收集信息的繁琐,也具有标准化、开放性、开发维护成本低等优点,但是人们

^① 收稿时间:2009-04-23

仍然期望能够将传统的电子文档与管理信息系统融合起来,构建以人为中心、以文档为基础结构、安全性更高的应用系统。

本文提出了一种基于 Microsoft Visual Studio 2005 Tools for the Microsoft Office System (VSTO 2005)^[1,2]的智能文档技术,以非结构化文档处理为核心,实现了 Office 办公软件与管理信息系统相融合的企业环保信息调查系统。智能文档不仅仅是用户输入信息的储存库,还可以包含编程逻辑,以实现数据的分布式录入、集中式自动汇总、基于模板的打印,与外部数据库的交互以及对数据的多层次安全保护等功能。

1 智能文档与VSTO基本概念

电子文档作为信息交换最主要的载体,在社会工作和生活中有着广泛的应用。而智能文档是结构化数据和非结构化数据的综合体,通过嵌入业务逻辑使文档具有一定的智能。智能文档能够将传统的电子文档与管理信息系统相结合,除了具备纸张和电子文档所具有的特性外,它还包含:

(1) 表示逻辑。提供与纸张文档和表单相同的外观,用户能够快速熟悉电子流程。

(2) 业务逻辑。有高度结构化的 XML 信息和编程逻辑,可以根据符合条件的文本字符串来触发相应的操作,可实现文档与外部数据库的交互以及对数据的多层次安全保护等功能。使制作者能够管理文档访问权限,并在文档中嵌入业务逻辑,如计算和验证。

(3) 数据传输。使企业以开放标准格式获取数据,并根据业务准则将其传递给相应的人和程序。

作为一种使用最广泛的办公软件, Microsoft Office 2003 已经不再是单一的电子文档编辑软件,它扩展了网络方面的支持,整合了最新的 XML 技术和 CRM/ERP 概念,为办公软件与管理信息系统的融合提供了一个简单而高效的平台。目前对于这类 Office 应用系统的开发,主要是通过 VBA 项目来实现^[3,4],但是它在用户界面、部署以及安全性和灵活性方面存在不足^[5]。VSTO 是一种实现 Office 智能文档,实现该融合的高效工具^[2-6]。

(1) 开发人员不但可以使用各种控件,包括 WinForm 托管控件和 VSTO 宿主控件(如 Book mark,

XMLNode 等),而且可以利用 Visual Studio 为这些控件提供的丰富的设计时支持来提高工作效率。

(2) 支持各种托管代码,如 Visual Basic .NET、Visual C++ 或 Visual C#, 这样开发人员就可以充分利用现有 VB 或 C# 的经验进行开发。

(3) .NET Framework 类库和公共语言运行库 (CLR) 对 VSTO 提供了全面的支持,例如可以使用 ADO.NET 访问数据库、支持 XML Web Services 多层结构解决方案的开发等。

(4) 支持将解决方案部署到本地计算机或网络位置,并且部署模型是基于清单的(清单在解决方案的自动升级过程中具有重要作用)。

(5) 集成了 .NET Framework 提供的安全功能,另外,在 VSTO 环境中编程时,其代码是包含在该项目所生成的程序集中的,且编写的是托管代码,与托管代码一起提供的公共语言运行库会对代码进行验证,避免了执行非法操作。另外还增加了受限制的应用程序域级策略,这使得代码在没有被显式授权之前是不能运行的。

2 企业环保信息调查系统结构及功能

本文实现的企业环保信息调查系统的基本功能是:企业用户在 Office 客户端通过数据处理界面,完成相关信息的录入,系统自动生成 Word 文档并上报;环保部门在 Web 服务器端完成对分布式上报的文档进行数据的批量自动入库以及自动汇总分析等;数据回填功能模块完成对最终汇总文档的打印确认。系统总体结构如图 1 所示,下面分别介绍。

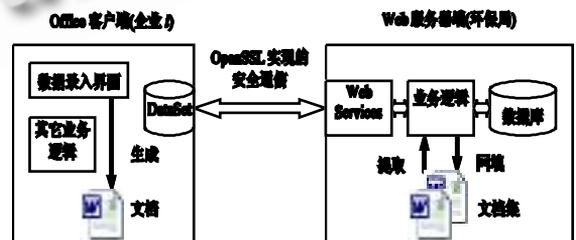


图 1 系统总体结构

2.1 Office 客户端

主要功能及特点包括:

(1) 企业用户打开来自可信服务器的文档或电子邮件中的附件的时候,智能文档和相关的组件就会被安装。它们可以自动从可信的服务器上更新自身,极

大的简化了更新的布署。智能文档遵循 Office 安全性设置,要求它们下载自可信服务器并有可信源的签名。可以为用户提供一个控件以决定是否打开不符合执行安全性标准智能文档。

(2) 客户端是用户访问智能文档的窗口,用户通过客户端查询、编辑、审核智能文档。它既是用户操作智能文档的工具,又是执行业务逻辑的容器,能根据用户在编辑过程中触发的事件执行相应的业务操作。使用用户熟悉的 Office 环境,用户不需要培训就可使用系统,Office 本身还具有强大的文字处理功能。

(3) 客户端软件要求: .Net Framework 2.0, VSTO 2005 Runtime, Word 2003, Office Primary Interop Assemblies。

2.2 Web 服务器端

主要功能模块及特点包括:

(1) Web Services 运行在环保局的服务器上,是系统的核心部份,负责处理客户端的各种数据请求以及和数据库交互。客户端可以通过各种网络(Internet、企业网)访问 Web Services。数据库(SQL Server 2005)保存着企业提交的环保调查信息等数据,它对外的接口是存储过程和视图,由 Web Services 访问。

(2) 智能文档设计组件,提供智能文档模板的动态创建、组装、定制化功能。

(3) 智能文档服务器组件,提供智能文档及其模板的管理和控制功能。通过数字签名和细粒度的权限控制,保证文档的真实性,保护敏感信息。

(4) 工作流组件,提供柔性引擎,支持协同数据管理服务,提高工作效率和质量。

(5) 数据集成与数据分析组件,将智能文档所包含的数据提取到数据库或数据仓库,进行集成、联机数据分析和数据挖掘。

(6) 开发工具: Visual Studio 2005, VSTO 2005, Office Professional 2003, SQL Server 2005。

3 基于VSTO的智能文档实现关键技术

3.1 数据录入界面的实现

(1) 界面设计

在 Word 中设计、制作企业环保信息调查表,并保存为 .doc 格式的文档。打开 Microsoft Visual Studio 2005,在新建项目对话框中,选择位于 Visual C#语言节点下的“Office”子文件夹中的 Word 文档,启动 VSTO 的项目向导,并为此应用程序选择先前制

作的 Word 文档。

VSTO 会在项目内自动创建一个 Document 宿主项。可以使用 ThisDocument 引用该 Document 宿主项。ThisDocument(对于 Visual C#为 .cs 文件)提供设计界面和代码文件的文档。下面进行录入界面设计(如图 2)。



图 2 录入界面设计图

在 ThisDocument 的设计界面,拖放若干 label、Bookmark 控件,其中 Bookmark 控件的命名规则为: bm+ 填写字段名,同时创建企业基本信息窗体。Bookmark 控件提供了一种便利的方式在 Word 文档中插入不连续数据片段。

(2) 设置文档状态

文档共设有三个状态,分别是初始状态、保护状态以及非保护状态。第一次运行,文档处于初始状态。当填写完企业信息数据,并用密码将填写内容保护后,文档处于保护状态。倘若需要对已保护的文档进行修改,输入密码解密后便可再度进行编辑,此时文档处于非保护状态。

实现时,文档的状态由文档变量表示的,文档变量是存储在文档中的变量,它的值不会随文档的关闭和打开发生变化。可通过文档状态的值判断文档当前的状态值。如文档的初始化状态,定义一个名为 initial 的文档变量,根据其值可判断文档当前是否处于初始化状态。关键代码如下:

```
private string _初始化状态属性 =
"09CDCECA-30D7-469e-8B7E-15C43C24018B";//
随机的一个 GUID 字符串

public bool initial
{get
{object objName = this._初始化状态属性;
try
{string temp=this.Variables.get_Item (ref
objName).Value;// 获取 objName 的值
return bool.Parse(temp);
```

```

    }
    catch
    {object objValue = false.ToString();
    this.Variables.Add((string)objName, ref
objValue);//在文档的 Variables 集合中添加一个名为
objName 的文档变量, 初始值为 false
        string temp=this.Variables.get_Item
(ref objName).Value;//获取 objName 的值
        return bool.Parse(temp);
    }
}
}

```

(3) 主要功能实现

信息的录入包括两部分: 企业基本情况和指标信息。如图 3, 企业基本信息须通过企业基本情况窗体来填写, 经后台处理后自动填入文档中。如, 从企业基本情况窗体中获得的年份值为 2005, 使用 Bookmark 的 Characters[i].ModifyEnclosure()方法给每个字符加圈, 最终使填写在文档中的年份为 ②①⑤。



图 3 设计信息录入界面

指标信息则直接填写在文档的 Bookmark 中, 可用 Bookmark 控件的 Editors.Add()方法对指标信息授予可写权限, 用 this.Protect()语句使文档除指标信息外处于只读状态。除指标信息部分, 文档其余部分是不可以编辑的。

3.2 数据批量提取模块的实现

(1) 界面设计

在窗体上拖放“选择 Word 存放路径”按钮和“批量提取”按钮, 如图 4 所示。



图 4 设计数据批量提取界面

(2) 主要功能实现

引入类库 Microsoft.Office.Interop.Word, 选择 MICROSOFT jet 4.0 OLE DB Provider 引擎连接 SQL Server 数据库, 通过自动化的方式操纵文档, 并提取信息。

批量提取功能代码如下:

```

private void btn 提取 _Click(object sender,
EventArgs e)
{ //获得文档所在的路径
    string FilePathdirectory = directory;
    //将 FilePathdirectory 路径下后缀名为 .doc 的文件路
径放在 dirs 数组中
    string[] dirs=Directory.GetFiles (FilePath
directory, "*.doc");
    // 循环处理 dirs 路径数组下的文档
    for each (string dir in dirs)
    {……
    //判断是否是由数据录入子系统生成的文档
    if (Check 是否验证(dir) == 标识)
    {FilePath = dir;
    _Doc = _App.Documents.Add(ref FilePath, ref
_Obj, ref _Obj, ref _Obj);
    object obj年份 = this._年份;
    int hj51 =Convert.ToInt32(_Doc.
Variables.get_Item(ref obj年份).Value. ToString());
    object v_hj52 = "bm 填报单位详细名称";
    string hj52=_Doc.Bookmarks.get_Item (ref
v_hj52).Range.Text;
    ……
    string sql = "insert into hj5(HJ1001,HJ1002,…)
values(" + hj51 + "," + hj52 + ",…)"
    ……
    }
}
}
}

```

3.3 智能文档安全性设计

本系统的安全性主要是从防录入错误, 防篡改、防伪造以及安全通信这四个方考虑, 另外由于 VSTO 解决方案集成了 Microsoft .NET Framework 提供的的安全功能, 并增加了受限制的应用程序域级策略, 使得没有被显式授予权限代码不能运行, 只有在部署系统时进行授权才能正常使用, 这也使得系统的

安全性得到了进一步的增强。

(1) 防录入错误。在数据录入子系统中, 设置了指标间关系的校验, 只有文档中填写的指标数据满足规则, 文档才能被保存, 保证最终入库数据的合理性。检查指标关系的部分代码如下:

```
private bool CheckDocument()
{.....
```

Tips = Tips + GxFrameWork.WordCheck. Check 大于等于(this.bm 污水处理能力.Text, this.bm 当年新增处理能力.Text, "2>=3"); //污水处理能力的值是否大于当年新增处理能力的值, 是则返回一个空串, 否则返回一串提示信息(如: 不满足指标关系 2>3)

Tips = Tips + GxFrameWork.WordCheck. Check 大于等于(this.bm 污水处理量.Text, this.bm 污水再用量.Text, "4>=5");

```
.....
}
```

(2) 防篡改。在 Word 录入子系统中, 放置了一个“检查保护”按钮。通过设定密码来保护文档, 防止信息被未经授权的人篡改。密码保护文档功能实现如下:

```
.....
```

密码设置 Dialog = new 密码设置(); // 创建名为密码设置的窗体

```
if (Dialog.ShowDialog()== DialogResult.OK)
{object objKey = this._默认文档保护口令;
this.Unprotect(ref objKey);
this.RestoreDocument();
objKey = Dialog.密码;
```

```
this.Protect(Microsoft.Office.Interop.Word.WdProtectionType.wdAllowOnlyReading,
ref missing, ref objKey, ref missing, ref missing); //保护文档, 使文档处于只读状态
}
```

用设置的密码解开文档保护, 授予文档编辑权限。此功能代码如下:

```
private bool ResetDocument()
{ 输入保护密码 Dialog = new 输入保护密码();
try
{ if (Dialog.ShowDialog() == DialogResult.OK)
{ object objKey = Dialog.密码;
if (this.ProtectionType !=
```

```
Word.WdProtectionType.wdNoProtection) //判断文档是否处于非保护状态下
```

```
{ this.Unprotect(ref objKey); //解开被保护的文档, 使其处于可编辑的状态
}
this.InitializationDocument();
return true;
}
}
.....
}
```

(3) 防伪造。由数据录入子系统生成的文档, 都通过文档变量定义了唯一标识, 完成对文档的“签名”。在数据批量提取子系统中, 只对具有该种标识的文档进行信息提取, 即不是由数据录入子系统生成的文档, 提取子系统不会进行处理, 避免了非法数据的干扰。

在数据批量提取子系统中识别有效文档代码如下:

```
private bool Check是否通过验证(string dir) //dir
为文档所在路径
{App = new Word.ApplicationClass();
vFilePath = dir;
_Doc = _App.Documents.Add(ref vFilePath, ref
_Obj, ref _Obj, ref _Obj);
object obj文档标识 = this._文档标识;
try
{ if (_Doc.Variables.GetItem(ref obj 文档标识).Value == "标识 1")
{ return true;
}
}
}
.....
}
```

(4) 安全通信。SSL 协议应用在传输层和应用层之间, 采用对称加密或非对称加密技术、消息摘要技术、数字证书等密码技术实现通信安全的三个方面, 即通过认证连接两端, 传输经过加密的数据, 并且能确保数据是完整没有被篡改的。OpenSSL 是一个非常优秀的实现 SSL/TLS 的著名开源软件包, 本系统基于该开源包实现了 Office 客户端与服务器端的加密通信^[7]。

(下转第 175 页)

(上接第 118 页)

4 结论

本文提出了一种基于 VSTO 2005 的智能文档技术,以非结构化文档处理为核心,实现了 Office 办公软件与管理信息系统相融合的企业环保信息调查系统。本系统在 JS 省环保局、ZJ 省环保局的实际应用中,运行良好,避免了传统环保信息收集的弊端,达到了信息的格式规范化、电子文档中数据处理的自动化、数据安全等应用要求。本系统对于企业、院校中的分布式数据收集、集中数据处理的实际应用有着重要的借鉴意义。

参考文献

1 Bruney A. Visual Studio 2005 Tools for Office. Wiley Publishing, Inc., 2006.

- 2 Schultz K, Schuler K. Creating an Estimate Solution Using Visual Studio 2005 Tools for Office. http://msdn.microsoft.com/library/default.asp?url=/library/en-us/odc_vsto2005_to/html/odc_vstoEstimateSampleSolution.asp, 2004.
- 3 魏红君,阎俊.基于对象的 VBA 技术在试卷自动生成系统中的应用.科技资讯, 2006,8:105 - 109.
- 4 李璟.利用 VB 和 VBA 实现 Word 的随机抽题和自动评分.现代计算机, 2006,7:103 - 105.
- 5 Kelly M.陶刚编译.使用 C++和 XML 建立智能文档. 2004. <http://tech.ccidnet.com/art/1081>
- 6 关巍.利用 VSTO 实现文档的智能化.电脑与电信, 2006,6:56 - 58.
- 7 陈波,张碧云.一个基于 OpenSSL 的 C/S 加密通信程序.电脑编程技巧与维护, 2007,8.