

XML 解析技术在 IPTV 机顶盒中的应用^①

杜 源, 张永安, 周剑超

(昆明理工大学 理学院, 昆明 650093)

摘 要: 根据实际要求, 设计了一种可以在 Linux 嵌入式操作系统中运行的 XML 解析器, 详细说明了设计过程以及基于 C 语言的编程实现, 并通过实际检测, 证明了其可行性和有效性, 已成功应用于 IPTV 机顶盒。

关键词: XML 解析; IPTV; C 语言编程; Linux 嵌入式操作系统

Application of IPTV STB Based on XML Parser

DU Yuan, ZHANG Yong-An, ZHOU Jian-Chao

(Institute of Laser, School of Kunming University of Science and Technology, Kunming 650093, China)

Abstract: According to the actual needs, we design a XML parser based on Linux Embedded system. The design process is analyzed in detail and the program based on C Language is given in this paper. The feasibility and effectiveness are proved by actual test. Now, this XML parser has already been applied to IPTV STB successfully.

Key word: XML parser; IPTV; C program; Linux embedded operating system

1 技术概述

IPTV 即交互式网络电视(Interactive Personality TV), 是一种利用宽带有线电视网, 集互联网、多媒体、通讯等多种技术于一体, 向家庭用户提供包括数字电视在内的多种交互式服务的崭新技术。用户可以在各种终端, 对自己喜爱的节目进行点播、下载、录制等。IPTV 机顶盒是将网络数据流转换成电视机可接受信号的附加装置, 它的软件平台为嵌入式操作系统(如 WinCE、Linux、VxWorks), 由制造商自定义内核^[1]。本文根据实际要求自定义了一套基于 DOM(Document Object Model)的 XML 文本, 并利用 C 语言在 ST-Linux 环境下解析 XML 中的内容, 通过解析程序, 提取相关数据, 最后在普通电视机上播放显示内容, 而如要更改界面和流媒体内容, 只需更改 XML 文本文档的内容即可, 大大简化开发流程, 缩短了研发周期, 增加了可移植性, 适应实际应用需求。这种方式的优点是, 当需要更改节目菜单甚至是整个机顶盒界面时, 不再需要对程序代码有任何修改, 只简单更改 XML 里面的内容信息即可^[4]。

XML 意为可扩展标识语言(extensible markup language), 是标准通用标识语言标准(SGML)的一个子集, 用于描述网络上的数据内容和结构。它是一种简单的数据存储语言, 使用一系列简单的标记描述数据, 其数据容易加载到程序中进行分析, 并以 XML 格式输出结果。通过解析提取出 XML 中的数据, 调用相关函数, 可以为开发工作节省大量的研发时间。

本文采用的操作系统是由意法半导体公司提供的 ST-Linux 系统, 开发平台也是由 ST 公司提供的集成达芬奇方案的 ARM 开发板。

2 基于DOM的XML解析器详解

2.1 XML 解析器

XML 的数据可以用 DOM 来构建, DOM 是将所有信息写成树状的结构以文档形式储存起来, 必须包含详细的信息, 因此, DOM 并不适合大规模高数据量的解析工作, 本文所解析的 XML 文档, 定义单页所包含的一级子目录最多不能超过 128 项。其结构的树状结构如下所示:

① 收稿时间:2010-12-21;收到修改稿时间:2011-01-23

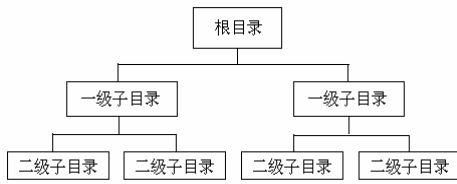


图 1 DOM 中信息树状结构

```

<?xml version="1.0" encoding="utf-8" ?>
<epg>
<list type="list" text-size="1x1">
<link cnumber="0" pre_xml="null" current_xml="vod/vod1.xml" nxt_xml="vod/vod2.xml" />
<item StreamName="Laislabonita_H264.ts" pos="65,90,150,10" NetStreamType="VOD"
IsPAL="TRUE" href="rtsp://192.168.12.68/Alizee3_1.5Mb.ts" VideoType="H264"
AudioType="MP2A" VideoID="69" AudioID="68" PcrID="69" />
</list>
</epg>
  
```

图 2 XML 解析文本样本

2.2 XML 解析函数

XML 解析函数实现的主要功能是解决多页节目菜单能按正确的顺序显示，正确解读当前页中的各个子元素包含的信息，使 IPTV 机顶盒设计中存在的多种 API 函数能够调用解析数据，读取外部 XML 文件，正确调度内存空间。具体函数设置如下：

2.2.1 头文件 axmlengine.h 函数

本文的 XML 解析函数的头文件定义了如下结构体：STGXOBJ_Rectangle_t(), BtnInfo(), VideoFrameInfo(), VideoBtnInfo()。其中 STGXOBJ_Rectangle_t() 用来存放二级子目录中流媒体信息在屏幕上显示的位置，BtnInfo() 用来存放二级子目录的个数，VideoFrameInfo() 用来存放一级子目录的信息，VideoBtnInfo() 用来存放二级子目录中流媒体的各种信息，以 VideoFrameInfo() 为例，其具体定义如下：

```

typedef struct
{
char      Type[NODE_VALUE_MAXLEN];
BYTE      TextSize[2];
char      PreXml[NODE_VALUE_MAXLEN];
//前一页 XML 文档名称
char      CurrentXml[NODE_VALUE_MAXLEN]; //
当前 XML 文档名称
char      NxtXml[NODE_VALUE_MAXLEN]; //
下一页 XML 文档名称
  
```

```

char      Cnumber;
//当前页编号
}VideoFrameInfo;
  
```

2.2.2 VideoFrame()函数

VideoFrame()函数实现的功能是解决多页节目菜单能按正确的顺序显示的问题，在 XML 文档中，每一页都加入了上一页及下一页信息，在成功解析的基础上，当机顶盒检测到红外遥控器发出按键值为“0x1A08”即遥控器向上按键时，主程序调用显示函数打开 pre_xml 中的文本，当遥控器发出按键值为“0x1A1B”即向下按键时，再调用显示函数打开 nxt_xml 中的文本，这样就完成了翻页操作，解析代码如下：

```

src=strstr(src,"pre_xml=");
src=strstr(src,"");
end=strstr(src+1,"");
strncpy(pVideoFrame->PreXml,src+1,end-src);
*(pVideoFrame->PreXml+(end-src)-1)='\0';
src=strstr(src,"nxt_xml=");
src=strstr(src,"");
end=strstr(src+1,"");
strncpy(pVideoFrame->NxtXml,src+1,end-src);
*(pVideoFrame->NxtXml+(end-src)-1)='\0';
  
```

2.2.3 VideoButton()函数

在解决了翻页问题后，VideoButton()函数地功能是正确解读当前页中的各个子元素包含的信息，在设计 XML 文档规范时，规定节目子元素包含的信息有流媒体名称 StreamName,显示的位置 pos,流媒体类型 NetStreamType,流媒体网络地址 href, 音频 ID, 视频 ID, 当规定好文本后，按照格式规范，逐行解析出相关内容，最后放入数据库保持，供其他 API 函数调用其解析。

2.2.4 GetVideoXmlInfo()函数

所以在成功解析 XML 的基础上，为了能让 IPTV 机顶盒设计中存在的多种 API 函数调用解析数据，设计 GetVideoXmlInfo()函数，用来读取外部 XML 文件，并统一分配、释放内存。

3 解析结果及流程图

3.1 解析结果

将以上函数编写成解析程序，在 Linux 的 Fedora7

版本下,将图二所示的 XML 文档解析后,其解析结果如下所示:

```

cnumber=0 pre_xml=null current_xml=vod/vod1.xml next_xml=vod/vod2.xml
PositionX=65 PositionY=90 Width=150 Height=10
StreamName=Laislabonita_H264.ts href=rtsp://192.168.12.68/alizee3_1.5Mb.ts

```

图 3 解析结果

3.2 解析流程

整个解析流程如下图所示:

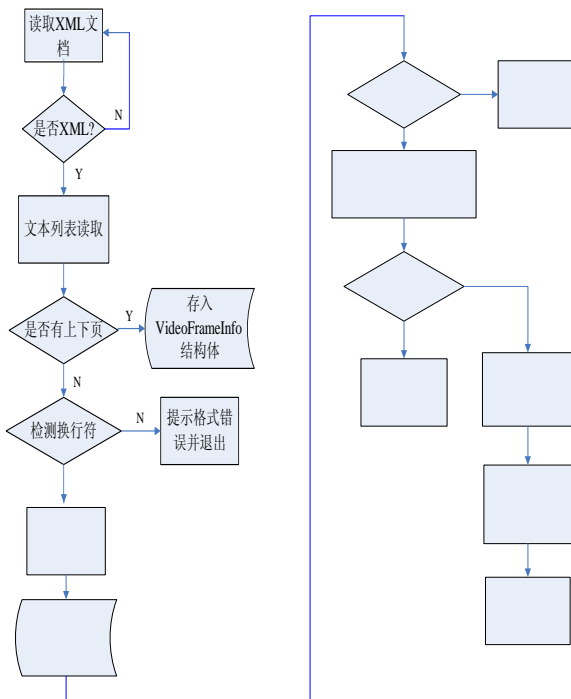


图 4 解析流程图

(上接第 247 页)

6 结语

本文中给出的思路、数据结构和算法实现,实际上就是“一种将结果记录集转换为二维数组的方法”这种方法已经在物流信息系统集成中间件 LESB 的数据交换情形中进行了应用,较好地解决了前文中提出的数据库操作的连接释放、结果记录集对象占用资源、异常处理等棘手的问题。

这种“将结果记录集转换为二维数组的方法”在各种应用系统开发的过程中具有普遍而实用的价值。

由如上结果对照图 2 可见,用此方法对按照本文所定义的 XML 文档格式进行文档解析,解析结果完全正确,解析出来的内容可供上层的 API 函数调用,如需更改流媒体内容或界面内容,只需按照 XML 文档格式更改相应内容即可,让服务器端的操作变得轻松,简单,快捷。

4 结论

本文应用 C 语言编程解决了基于 DOM 模型的 XML 解析工作,成功为上层 API 函数提供了所需要的数据,使得整体 IPTV 设计工作大大简化,从而减轻了开发的工作量,加快了研发进程。为 IPTV 的改进和开发提供了一种有效方法。

参考文献

- 1 谢质文.IPTV 产品,营及案例.北京:电子工业出版社,2008.
- 2 倪继利.Linux 内核分析及编程.北京:电子工业出版社,2005.
- 3 李善平,刘文山,王焕龙.Linux 与嵌入式系统.北京:清华大学出版社,2006.
- 4 Wall K, Watson M, Whitis M. GNU-Linux 编程指南.自由软件基金会,1991.
- 5 勤研工作室.U-boot 运行机制.http://www.kiyi.com.cn.
- 6 Salzman PJ, Burian M, Oripomerantz. The Linux kernel module programming Guide, 2005.
- 7 Prate S. C primer plus.第五版.中文版.北京:人民邮电出版社,2005.
- 8 冯进,丁博,史殿习,张囡熹,许凯.XML 解析技术研究.计算机工程与科学,2009,31(2):120-124.

检测换行符 N 提示格式错误并退出

参考文献

- 1 梁伍七.基于 JDBC 的 Web 应用程序数据库连接技术研究.合肥学院学报(自然科学版),2010,20(4):29-32.
- 2 梁清翰,沈占锋,骆剑泉,吴红,明冬平,李均力.构建 LBS 系统的数据库连接池技术研究.计算机工程,2006,32(18):39-41.
- 3 邓子云,黄友森,杨晓峰,陈玉林,罗涛.基于SOA-BPM 组合架构的第三方物流XML文档集成平台.计算机系统应用,2010,19(3):1-5.

N

Applied Technique 调用 199

下个元素

GetVXmlInfo 函数