

成都区调中心专业气象系统^①

杨代才, 秦运龙, 王宏记, 龚贤创

(湖北省气象信息与技术保障中心, 武汉 430074)

摘要: 介绍了成都区调中心专业气象系统的总体设计与实现技术, 提出一个总体逻辑结构完整、信息流程合理、各子系统功能清晰、充分体现气象与水文相结合的总体设计方案. 该系统自动化程度高, 可为开发和建设同类水电行业专业气象服务业务系统提供实际经验和技術借鉴. 选用 Oracle 数据库, 采用 c#、jsp、C++ 等开发语言, 实现了成都区调中心专业气象系统的开发建设.

关键词: 气象信息数据库; 信息发布; 应用工具集; 气象预报

Specialized Meteorological Service System for Chengdu Area Cascade Dispatching Center

YANG Dai-Cai, QIN Yun-Long, WANG Hong-Ji, GONG Xian-Chuang

(Meteorological Information and Technology Support Center of Hubei Province, Wuhan 430074, China)

Abstract: This paper introduces overall design and implementation techniques for Chengdu Area Cascade Dispatching Center, Proposed a general logical structure is complete, Reasonable flow of information, the function of each subsystem is clear, and fully reflects the overall design of a combination of meteorology and hydrology. This is an automated system, it can for the development and construction of similar meteorological service of the hydropower industry provide practical experience and technical drawing. Using oracle database and c# and jsp and c++ language, Completed the development and construction of Specialized Meteorological Service System for Chengdu Area Cascade Dispatching Center.

Key words: weather information database; information dissemination; application toolset; weather forecast

在建的金沙江下游梯级成都区调中心专业气象系统, 是目前国内最大规模的依托气象部门服务水电企业的专业气象台和专业的气象服务保障系统^[1]. 系统的建成一方面将为正在抓紧建设的向家坝、溪洛渡水电站的正式运行提供专业的气象服务保障, 有针对性地服务工程建设和电力生产. 同时, 也进一步扩大了三峡集团在长江流域气象服务及保障的范围, 有效地提高了成都区调中心专业气象台计算机网络资源和气象信息资源的使用效率, 预报的准确率, 增强了气象为水库调度服务的能力, 保障了水库运行安全, 提高了水库的运行效益, 使长江流域气象预报和水文预报能力得到进一步地提升. 系统的建立, 也是对中国三峡集团未来长江流域气象预报体系的进一步补充

和完善.

1 成都区调中心专业气象系统概述

金沙江下游梯级成都区调中心专业气象系统包括气象信息采集与处理、气象信息数据库、气象预测、气象应用工具、气象信息发布、气象预报业务工作平台、计算机信息网络七个分系统. 主要用于完成成都区调中心专业气象信息网络资料的传输、处理、存储、发布及应用; 建立气象信息数据库; 完成长江上游流域面雨量预报、日常天气分析、预报、会商、情报服务等应用平台的开发; 完成整个专业台硬件、软件系统的集成等工作^[2].

气象信息采集与处理分系统由卫星云图接收处理、

① 基金项目:三峡集团公司工程项目(JS/ME-CD-201101-GY)

收稿时间:2012-03-12;收到修改稿时间:2012-04-15

天气雷达信息采集与处理、气象信息接收管理和气象信息收集处理四大子系统组成。实现对卫星云图信息、天气雷达信息、闪电探测信息和气象局下发气象信息的收集与处理。

气象信息数据库分系统遵循气象和水文部门的有关技术规范,以文件库和数据库相结合的方式,建立满足预报服务的关系数据库和文件库,实现气象信息数据库、文件库管理。采用数据库统一访问接口,为各业务系统提供数据服务^[3]。

气象预测分系统由坝区短时灾害性天气预测子系统、流域中短期降水面雨量预测子系统和流域长期降水趋势预测子系统三大子系统组成。实现对短时灾害性天气信息、流域中短期降水面雨量、流域长期降水趋势信息的收集与处理。

气象应用工具集分系统由气象情报子系统、气象报表子系统和气象预报评价子系统组成,实现对常规气象要素、流域降水情况等气象信息的统计、显示,并形成各种产品和报表,供用户查询和使用,其中气象预报评价子系统实现各种预测预报产品检验功能。

气象信息发布分系统将各种渠道所获得的气象信息及各种预报预测分系统所生成的产品通过 WEB 页面以图文、表格等多种表现形式进行发布,以提供方便直观的气象信息服务。同时,还能够通过短信的方式发布相关的气象及预警信息,提供气象信息语音发

布功能。

气象预报业务工作平台具有浏览气象信息、制作发布中短期天气预报、流域面雨量预报等功能,集成调用成都梯调气象台相关气象业务软件,是实现气象信息分析、决策与预报服务的综合人机交互式工作平台。

计算机网络采用先进、成熟、标准化的网络技术,为各业务系统之间提供高速可靠的连接与高性能的信息交换能力,提高网络接入的可靠性与能力,增强接入系统的安全控制能力,以满足信息交换、共享服务、网络安全管理与监视等需求,实现集中远程管理。

通过系统功能的开发,建立了一个全面满足专业气象服务业务工作,集气象信息采集与处理、气象信息数据库、气象预测、气象应用工具集、气象信息发布、气象预报业务工作平台等诸多功能为一体的综合业务工作平台。

2 应用软件系统总体设计与功能

2.1 总体逻辑结构

都区调中心专业气象系统应用软件包括气象信息采集与处理、气象预测、气象应用工具集、气象信息发布共五个分系统,应用软件平台各分系统及子系统逻辑关系如图 1 所示:

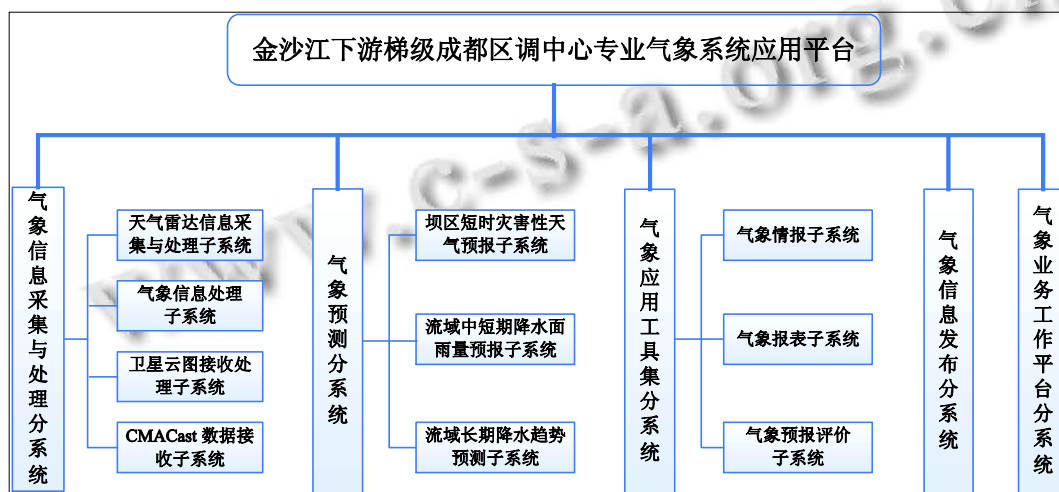


图 1 应用软件系统总体逻辑结构图

(1)气象信息采集与处理分系统能够有效监测向家坝、溪洛渡坝区库区及长江中上游流域内的灾害性和突发性天气。全部监测资料信息经加工处理后进行存

储分发,为区调中心专业气象台其它业务系统提供信息来源。

(2)气象预测分系统能提供所辖坝区灾害性天气预

警报、流域中短期降水雨量预报和流域长期降水趋势分析预测产品。

(3)气象应用工具集分系统能提供气象情报资料检索查询、气象报表制作平台,并能对业务应用中的各类预报产品进行评价。

(4)气象信息发布分系统能通过信息网络,以图、文等形式,向水情系统(水文预报系统)和三峡开发总公司系统提供气象决策预报产品、监测产品等。

(5)气象预报业务工作平台分系统能够浏览气象水文信息,制作发布中短期天气预报、流域面雨量预报,集成调用成都区调气象台相关气象业务软件,是实现气象水文信息分析、决策与预报服务的综合人机交互式工作平台。

2.2 总体信息流程

应用软件系统总体信息流程图如图 2:

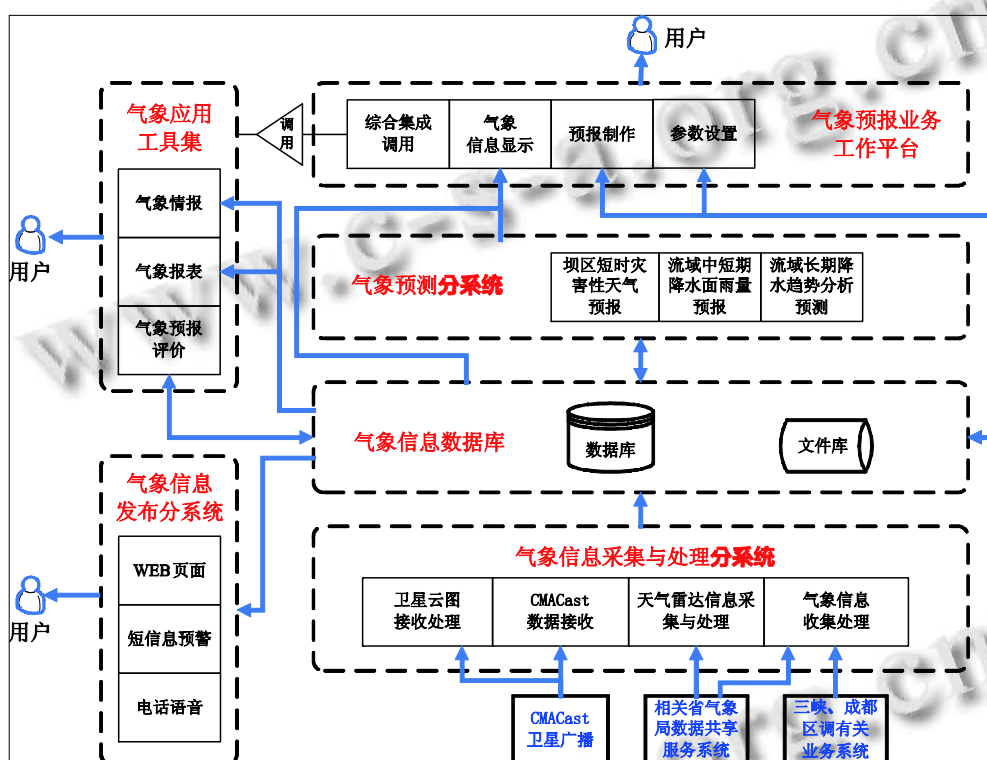


图 2 应用软件系统总体信息流程图

(1)信息的采集与处理

输入系统的气象信息包括常规地面、高空资料、自动气象站观测资料、气象卫星云图、天气雷达探测数据、闪电监测数据和数值天气预报模式产品等。其中,除了天气雷达基数据、气候长期预报所需的部分资料是由专线自有省气象局信息共享系统获取的,其他信息均由 CMACast 数据接收子系统接收,信息经过必要的处理后,存储到气象信息数据库系统中进行管理。

(2)信息产品的加工

信息产品包括灾害性天气预警报、流域中短期降水雨量预报、流域长期降水趋势分析预测产品和各

类气象预报产品评价结果,主要由气象预测系统、气象预报业务工作平台和气象应用工具集产生。其中,客观灾害性天气预警报、流域中短期降水雨量预报、流域长期降水趋势分析预测产品,是由气象预测系统直接获取气象信息数据库保存的原始信息进行加工而产生;主观天气预报、流域中短期降水雨量预报产品,由预报人员通过气象预报业务工作平台交互产生,此时,预报人员需要利用数据库中保存的天气背景、常规气象资料、数值预报产品和同类客观产品等信息;各类气象预报产品评价结果由气象应用工具集对数据库中保存的各类气象预报产品进行评价处理后产生。加工产生的信息产品将送入数据库保存。

(3)信息产品的使用与发布

预报员主要通过气象预报业务工作平台和气象应用工具集,利用气象信息数据库保存的相关信息,来进行气象水文信息查询、生成各类情报产品、进行预报制作和制作报表。气象信息产品、预报服务产品、不定期发布的专项预报服务产品、情报服务信息和预报服务产品等,通过气象信息发布系统,以 WEB 网站^[4]、手机短信、语音信箱等方式进行发布,供各类用户使用。

3 共享平台系统关键技术实现

3.1 高可用性和高弹性的双以太网结构逻辑单网技术

成都区调中心专业气象系统网络采用双以太网结构的逻辑单网,逻辑单网工作方式是将服务器或工作站的 2 块网卡分别连接到 2 个不同的交换机上(NIC Teaming 的 NFT 工作模式),且 2 块网卡捆绑在一起共享一个 IP。应用交换机 StackWise 的堆叠技术,使多个交换机实现物理堆叠,实现数据流全方位统一管理和规划的逻辑堆叠。这种逻辑堆叠使所有堆叠交换机从里到外形成统一堆叠,可作为一个单元使用,从而使整个网络构成逻辑上的单网结构。

交换机 StackWise 的堆叠体系结构,创建一个聚合带宽的通道,使高需求服务器获得更高的带宽,在网络使用高峰时最大程度的避免网络阻塞的能力。干道技术提供了更好的扩展性,保证了网络中高需求服务器有更高的可用带宽。弹性是堆叠架构与生俱来的特性,如果一条双向线缆断开了,交换机将自动使用第二条线缆传递信息,这等于对主交换机上做了 1: N 的冗余。热交换,管理员能在不关闭网络的情况下移除失效的交换机,用一个新的交换机代替,配置会在堆叠组中自动设好。低延时的高速网络,寻找所有端口都能自动适应最高速连接的解决方案。

通过以上技术实现了成都区调中心专业气象系统建设成统一的、高速有效的信息网络,并为信息资源共享提供良好的网络环境,达到各系统间共享和交换资料的目的,实现综合信息高速传输、灵活应用、安全存储和稳定运行。

3.2 雨量色斑图、等值线的显示应用

监测长江流域雨情信息并为电厂优化调度和实时防洪调度决策提供可靠的基础依据。雨量等值线图的绘制是该气象应用工具集分系统的一个研究开发子系

统。实现降水色斑图显示主要通过长江流域区域的地理信息地图上叠加显示由降水数据生成的色斑图层来实现,从而使得整个系统具有地理信息系统所具有的地图无限放大、缩小、平移等 GIS 操作功能^[5]。

首先,完成雨量值的显示,通过 ADO.Net 和 SQL 查询语句从气象信息数据库获取站点雨量数据表和站点数据表中的经度、纬度坐标以及其它必要的相关信息并生成 ADO. Net 数据集;在 ARCGIS Engine 中创建 IXEvent2 Fields Properties 对象,打开点要素数据表并根据经纬度定位,将 IXEventSource 转换为 IfeatureClass 并赋值给新建的站点要素图层,将获取的点要素数据转化为 GIS 图层数据,添加图层到地图控件;通过 ARCGIS Engine 中各种 Renderer 和 Symbol 接口来实现 GIS 图层中每个站点标注显示的内容和样式。

其次进行色斑图渲染,根据站点的降雨量值,利用 IDW 内插生成 raster 栅格,对生成的栅格图按照从小到大值进行分级渲染,生成等值线时再根据刚才 IDW 内插生成的栅格图层设定间距生成 contour 并标注。关键代码如下:

//栅格图分级渲染

```
RasterLayer SetRsLayerClassifiedColor(IRaster
pRaster, int span)
{
    IRasterClassifyColorRampRenderer pClassRen =
new RasterClassifyColorRampRendererClass(); //分级渲染对象
    IRasterRenderer pRasRen = pClassRen as
RasterClassifyColorRampRendererClass;
    //Set raster for the render and update
    pRasRen.Raster = pRaster;
    IFillSymbol pFSymbol = new
SimpleFillSymbolClass(); //背景填充标记
    IRgbColor ic = new RgbColorClass();
    Color tempcolor;
    int j = 0;
    pClassRen.set_Break(0, ca[0].start); //设置中断点
    数值
    for (int i = 0; i <= (max - min); i++)
    {
        tempcolor = Color.FromArgb(ca[min + i].color);
```

```
//从色标数组中取得色标颜色
ic.Red = tempcolor.R;
ic.Green = tempcolor.G;
ic.Blue = tempcolor.B;
pFSymbol.Color = ic as IColor; //背景标记颜色赋值
pClassRen.set_Symbol(i, pFSymbol as ISymbol);
//分级渲染对象设置背景标记
pClassRen.set_Label(i, ca[min + i].start + "~" + ca[min + i].end);//分级渲染对象设置标题
pClassRen.set_Break(i + 1, ca[min + i].end);//设置中断点数值
```

```
j = i;
}
pClassRen.set_Break(j + 1, 10000);//设置中断点数值防止出现过大数值导致异常
}
}
```

最后进行一个等值线绘制, AE 提供直接从栅格提取等值线的接口 ISurfaceOp2,该接口下的 Contour 方法包括四个参数, pGeoDataSet 为栅格面, dInterval 为等值线间距, 后面两个参数一般填 0, 设置好各个参数后即可获得满足要求的等值线.

经过上述步骤, 所显示的长江流域自动气象站降水色斑图、等值线如下如图 3 所示.

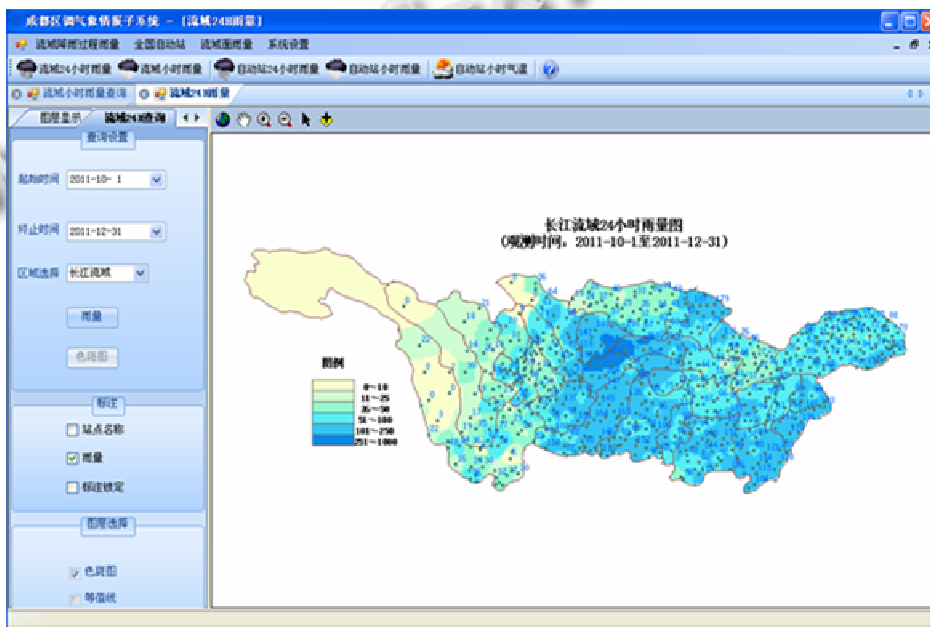


图 3 长江流域降水色斑图等值线显示

3.3 气象信息 web 发布实时无刷新自动显示最新数据功能的实现

气象信息发布分系统将各种渠道所获得的气象信息及各种预报预测分系统所生成的产品通过 WEB 页面以图文、表格等多种表现形式进行发布, 以提供方便直观的气象信息服务. 传统上, 我们浏览页面, 如果加入最新的数据. 只能是等我们重新向服务器端请求时才能显示出来. 但是, 对于气象上预警信息的发布显示、自动站实况、雷达图形等更新频率很密集、时效性很强的网站. 传统的这种做法是不能满足的. 我们让程序自动刷新定时向服务器请求数据. 30 秒取一次数据, 60 秒取一次数据. 利用 XMLHTTP 发出请求并取得

数据^[6]. 传到客户端, 客户端重新组织并显示数据. 提升用户体验, 达到很好的显示效果.

4 结语

该系统, 从信息的采集与处理、预报制作、信息发布、系统运行监控等均实现了系统自动处理, 人工干预少、维护量小, 适合预报人员少的特定环境. 在实践业务工作中充分发挥了集约化、自动化、运行稳定、实用方便等优势. 可为开发和建设同类水电行业专业气象服务业务系统提供实际经验和借鉴.

(下转第 136 页)

了大量口语化的信息特征, msg_length 在 30 左右及 $msg_length > 40$ 包含了较多非个人化的信息特征. 所以, 经过过滤的信息很大程度上能代表原信息的信息分布特征.

3.2.2 过滤算法对词频分布的影响

表 6 统计了过滤前后微博信息中词语频度分布的变化, 分析获得以下结论:

表 6 过滤前后词频分布变化

log(词频)区间	过滤前词语数	过滤后词语数	相同词语比率	减少词语数
$\geq 10e-2$	5	5	100%	0
$\geq 10e-3$	90	115	100%	-25
$\geq 10e-4$	1241	1303	99.11%	-62
$\geq 10e-5$	6898	5669	93.95%	1229
$\geq 10e-6$	17669	12463	78.23%	5206
$\geq 10e-7$	25036	16378	60.75%	8658
$\geq 10e-99$	32342	16379	47.24%	15963

(1) 过滤后词语数为 16379, 较过滤前的 32342 减少了 15963 个, 接近 50% 左右.

(2) 删除的词语中主要为低频词语, 过滤前在 $\log(\text{词频}) < 10e-5$ 的区间中有 25444 个词语, 过滤后仅有 10710 个词语, 减少的词语占总词语总删除量的 92.3%, 尤其是 $\log(\text{词频}) < 10e-7$ 的极低频区间中, 过滤前有 7306 个词语, 过滤后仅有 1 个词语. 删除低频词语可以减少待处理的数据规模, 提高算法效率.

(3) 在 $\log(\text{词频}) \geq 10e-5$ 的较高频区间, 过滤前与过滤后的词语数变化不大, 相同词语比率达到 93.95%, 这样对整体热词频度计算影响不大. 尤其由于删除了大量低频无关词语, 从而提高了少量热门词语的频度, 所以在 $\log(\text{词频}) \geq 10e-4$ 的区间内的词语数反而有所增加, 改善了热词抽取算法的结果.

算法能够较好的过滤海量微博信息, 在筛除大量无关信息的同时, 保留了大量重要信息, 从中有效获取了当前热词.

4 结语

微博信息存在着数据量巨大、内容形式多变、存在大量未登录词等特点, 本文提出的过滤及热词抽取算法能在一定程度上过滤海量信息, 并抽取微博热词, 具有速度快, 算法开销小的特点. 在未来工作中, 可以考虑引入未登录词发现的方法获取网络新词, 使用语义理解的方法挖掘微博信息, 更深入的发觉用户间信息传播模型等. 在大规模实时抽取中, 还可以采取分布式的方法提高算法效率.

参考文献

- Barabási AL. The origin of bursts and heavy tails in human dynamics. *Nature*, 2005, 435: 207.
- Dezs Z, Almaas E, Lukács A, et al. Dynamics of information access on the web. *Phys Rev E*, 2006, 73(6): 066132.
- Zhou T, Kiet HAT, Kim BJ, et al. Role of activity in human dynamics. *Europhys Lett*, 2008, 82(2): 28002.
- Grabowski A, Kruszezwska N, Kosiński RA. Dynamic phenomena and human activity in an artificial society. *Phys Rev E*, 2008, 78(6): 066110.
- 曹鹏, 李静远, 满彤. Twitter 中近似重复消息的判定方法研究. *中文信息学报*, 2011, 25(1): 20-27.
- 刘志明, 刘鲁. 微博网络舆情中的意见领袖识别及分析. *系统工程*, 2011, 29(6): 8-16.
- 许晓东, 肖银涛, 朱士瑞. 微博社区的谣言传播仿真研究. *计算机工程*, 2011, 37(10): 272-274.
- 张晨逸, 孙建伶, 丁轶群. 基于 MB-LDA 模型的微博主题挖掘. *计算机研究与发展*, 2011, 48(10): 1795-1802.

(上接第 110 页)

参考文献

- 何小明, 蒋永东. 气象信息共享服务业务平台用户认证系统的设计. *现代电子技术*, 2007, 21(2): 12-16.
- 李集明, 沈文海, 王国复. 气象信息共享平台及其关键技术研究. *应用气象学报*, 2006, 17(5): 34-36.
- 琚玲, 赵芳. ORACLE 数据库连接配置浅析及故障排除. *气象科技*, 2009, 37(4): 448-451.

综合 3.1 与 3.2, 从定性及定量两个方面证实本文

- 何险峰, 蒋丽娟, 等. 公共气象服务网站数据的及时发布. *气象科技*, 2011, (4): 483-488.
- 夏正龙, 尹新怀, 欧阳计跃等. SharpMap 在实现湖南省降水色斑图显示中的应用. *计算机系统应用*, 2011, 20(2): 134-136.
- 苏厚雄, 王婉茹. 基于 Ajax 和 PHP 数据分页的实现. *计算机系统应用*, 2012, 21(2): 218-220.