

在线音乐平台项目规格说明书

1. 概况

项目名称	在线音乐平台
行业	网络媒体
架构类型	B/S 结构
开发技术	<p>1. B/S 开发环境和技术</p> <p> (1) Eclipse+Tomcat+Java</p> <p> 采用技术: JSP、Servlet、SSH 框架、Tomcat。</p> <p> (2) .NET+IIS+C#</p> <p> 采用技术: ASP.NET、ADO.NET、IIS。</p> <p>2. 数据库技术</p> <p>MySQL、SQL Server 或 Oracle 数据库。</p> <p>3. 其它技术</p> <p>jspSmartUpload 组件使用、网络爬虫、IO 流操作、多媒体相关技术。</p>
规模	系统总体规模: 22.5 人月

2. 项目概述

2009 年全球网民数量已突破 10 亿, 中国网民达 3.84 亿, 其中经常使用音乐服务的用户占互联网用户总数的近 80%。与之相对应的全球网络音乐销售额也从 2005 年的 15 亿美元攀升至 2010 年的 107 亿美元, 网络音乐正成为全球音乐产业最重要的组成部分。

文化部在《2009 年中国网络音乐市场年度报告》中指出, 2009 年中国网络音乐市场保持了较好的运行态势, 在线音乐和无线音乐两大领域都实现了用户规模的稳步增长和市场规模的迅速发展。从市场规模来看, 2009 年我国网络音乐总体市场规模已经达到 20.1 亿元人民币。

近年来, 网络音乐在中国保持了快速的发展势头。文化部文化市场司网络文化处处长刘强介绍, 网络音乐的发展优势主要体现在三个方面, 即网络音乐的使用率在整个互联网运用环节里始终排在第一位; 网络音乐应用方式的多元化使更多网民参与网络音乐创作和网络音乐社区活动; 网络音乐已进入 3G 时代, 移动网络音乐的快速发展使音乐的应用更加广泛、便利。

作为一种网络音乐服务, 用户可以选择自己喜爱的流派、收听、创建或修改, 并可以给收听到的音乐评论。

本系统致力于成为中国最大的互联网音乐资讯共享交流平台,中国最好的音乐爱好者交流交友平台,中国最好的网络音乐人的展示平台!

本平台使命: 娱乐天下, 音乐你我!

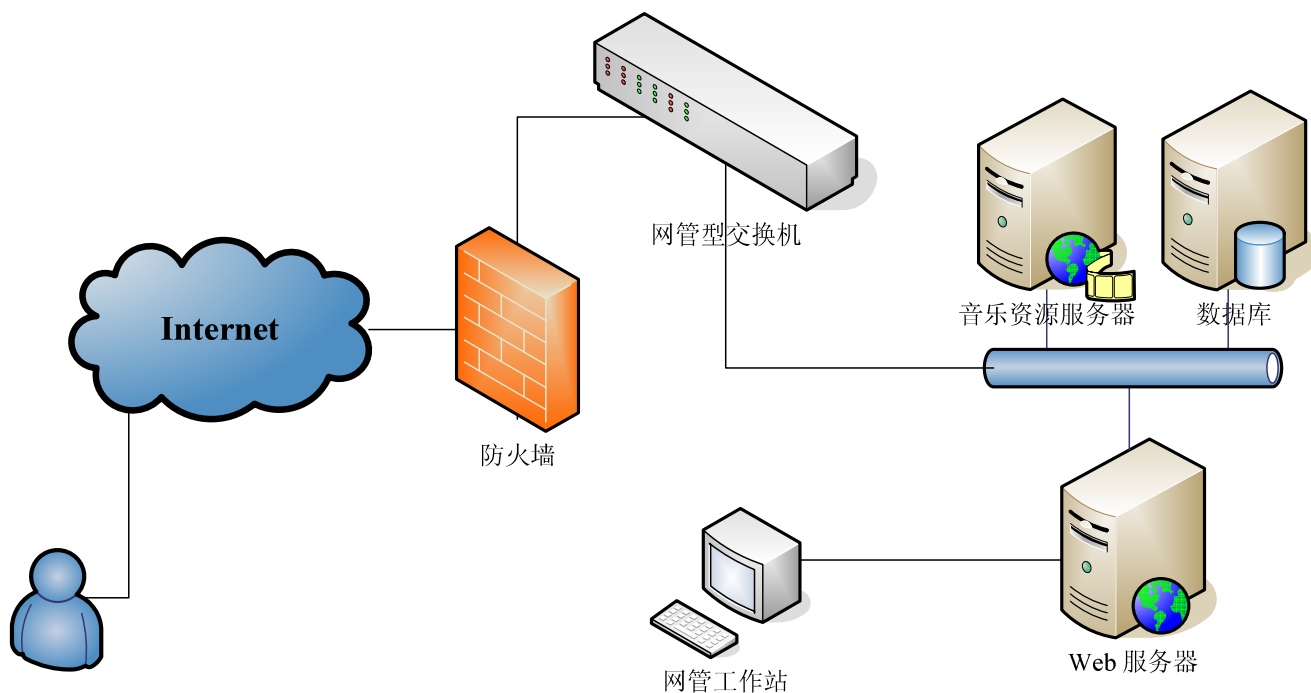


图 1 系统拓扑结构图

3. 总体业务流程介绍

用户的前台业务流程主要包括: 浏览最新上传歌曲、按类别浏览歌曲列表、搜索歌曲、在线试听、下载歌曲、歌曲连播、歌曲评论。

管理员的后台业务流程主要包括: 浏览歌曲信息、多条件搜索歌曲信息、添加歌曲信息、上传歌曲文件及歌词、保存歌曲信息、删除歌曲信息及歌曲文件、资源信息采集。

其处理流程如下图所示:

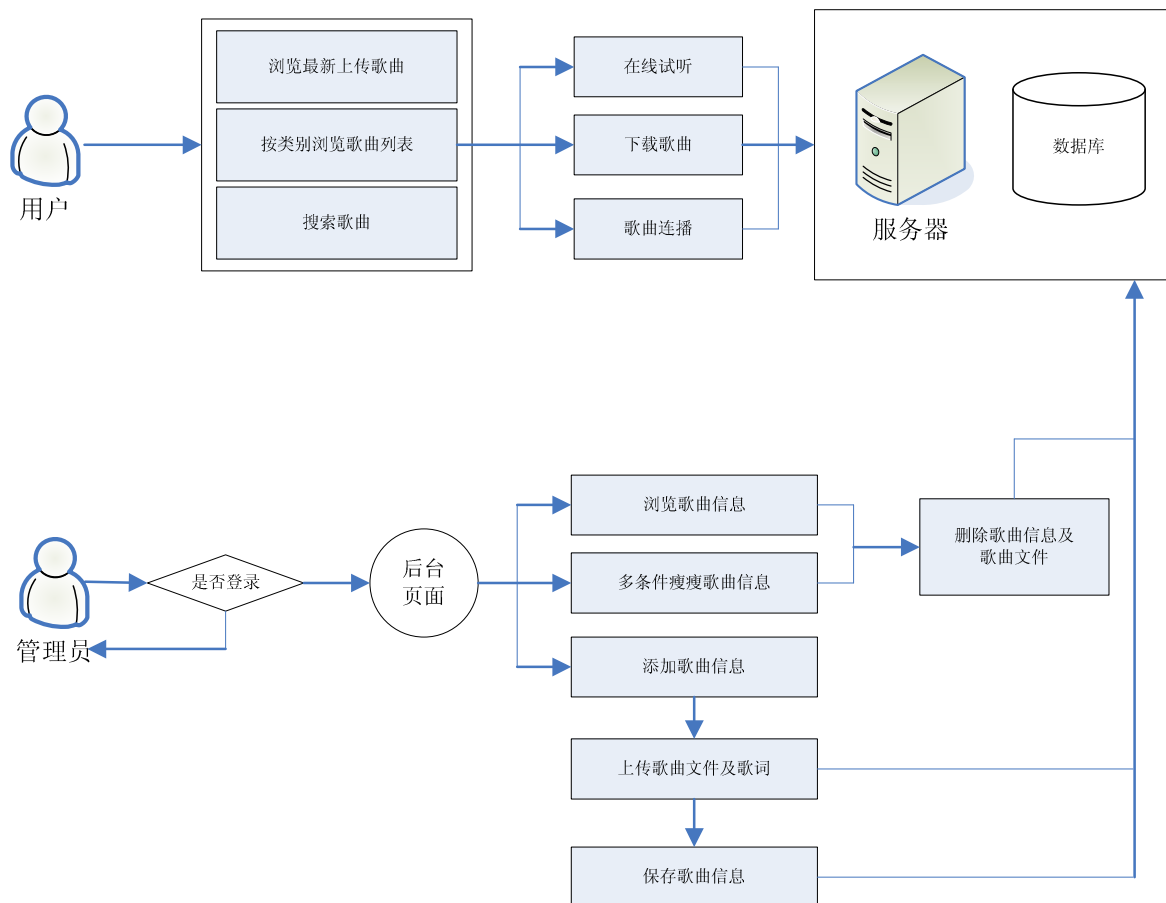


图 2 系统总体业务流程图

(1) 前台业务：

主要针对普通用户这一角色进行相关的业务处理，完成普通用户这一个角色在音乐平台上的一些基本需求。

(2) 后台业务：

主要针对管理员这一角色进行相关的业务处理，完成本音乐平台的一些资源配置和操作管理。

4. 系统功能

4.1 系统功能结构

系统采用 B/S 结构进行开发。主要功能包括前台：歌曲显示：分类显示最新的 10 首歌曲、按类别分页显示各类歌曲、试听歌曲并同步显示歌词、连续播放歌曲、下载歌曲、搜索歌曲、排行榜：试听排行、下载排行、歌曲评论；后台：分页显示歌曲列表、多条件查询歌曲、添加歌曲：检测歌曲是否添加、上传歌曲文件、上传歌词文件、保存歌曲信息、删除歌曲及歌曲文件、资源信息采集等 17 个方面，功能结构图如下所示：

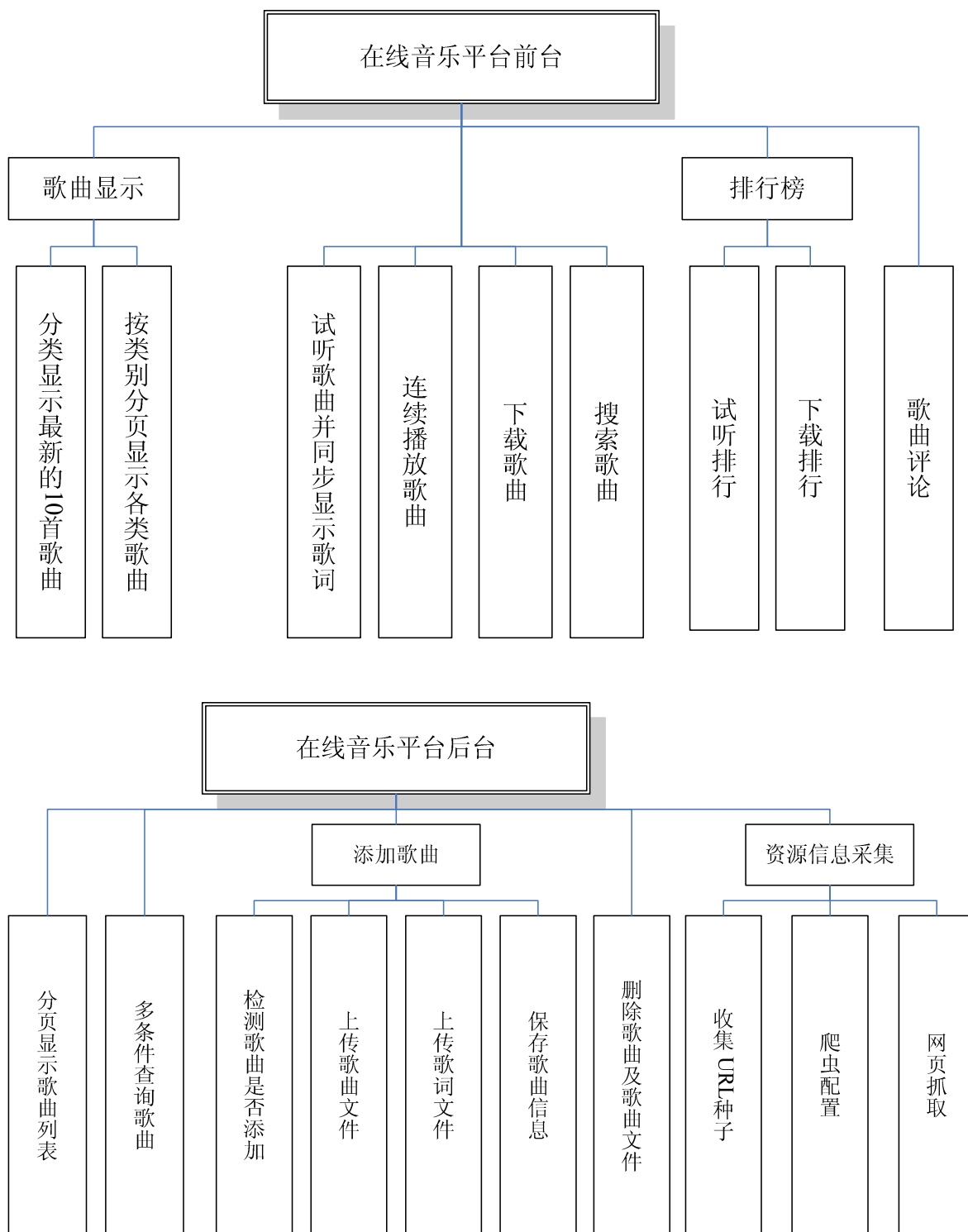


图 3 系统功能结构图

4.2 系统功能需求

4.2.1 功能需求概述

- 本系统主要由前台模块和后台模块两部分组成。前台模块：歌曲显示：分类显示最新的 10

首歌曲、按类别分页显示各类歌曲、试听歌曲并同步显示歌词、连续播放歌曲、下载歌曲、搜索歌曲、排行榜：试听排行、下载排行、歌曲评论；后台模块：分页显示歌曲列表、多条件查询歌曲、添加歌曲：检测歌曲是否添加、上传歌曲文件、上传歌词文件、保存歌曲信息、删除歌曲及歌曲文件、资源信息采集等 17 个方面。

系统前台

- **歌曲显示**

用户进入本音乐平台后，可以一目了然的在固定板块看到：

- 1.分类显示的最新 10 首歌曲；
- 2.按照类别分页显示各类歌曲。

- **试听歌曲并同步显示歌词**

实现了用户的试听功能，有候选歌曲列表选择，并有歌词同步伴随，歌曲播放的语句和歌词同步，并有高亮显示。

- **连续播放歌曲**

可以选择多首歌曲，进行勾选后连续播放。

- **下载歌曲**

当用户查看到自己心仪的歌曲时，可以点击下载按钮进行下载。

- **搜索歌曲**

用户可以根据歌曲名字进行模糊查询，方便快捷的定位自己想找的歌曲。

- **排行榜**

点击排行榜页面，可以查阅下载最多的歌曲排行和试听排行。

- **歌曲评论**

对歌曲进行留言和评论。

系统后台

- **分页显示歌曲列表**

管理员可以进入后台功能模块，分页检索到所有歌曲列表，方便管理和查阅。

- **多条件查询歌曲**

本查询不同于用户查询，该多条件查询可以通过歌手名、歌曲名、专辑名等多条件进行查询和组合查询，方便管理员精确定位歌曲资源。

- **添加歌曲**

本模块包括：检测歌曲是否添加、上传歌曲文件、上传歌词文件、保存歌曲信息。

- **删除歌曲及歌词文件**

管理员对歌曲、歌词进行删除操作。

- **资源信息采集**

利用时下流行的网络爬虫技术爬取其他网站的资源信息，为我所用。

对指定的资源网站 URL 进行解析，提取指定资源的歌手信息、歌曲信息以及可播放歌曲资源的网络地址。提取后，按照本音乐平台的规则写入到数据库中，进行存储。

4.2.2 歌曲显示

主要包括分类显示最新的 10 首歌曲和按类别分页显示各类歌曲。本子系统采用的是 B/S 架构实现。

(1) 分类显示最新的 10 首歌曲

- 可便捷、灵活地在音乐平台的某个版块浏览到最新的 10 首歌曲；
- 具有实时更新的特点；
- 本模块是在主体页面的一个部分，显示排名在前 10 名的最新歌曲做为新歌推荐；

(2) 按类别分页显示各类歌曲

按照华语、欧美、日语、韩语、影视、合辑、轻音乐等分类要求进行歌曲的板块或者选项卡页面的分类显示；

4.2.3 试听歌曲并同步显示歌词

主要包括：调用 Window Media Player 或者 Real Player 进行歌曲的播放。勾选歌曲后点击播放按钮弹出播放器的页面，页面左侧是播放器和播放列表，右侧为歌词显示区域，播放歌曲的同时，歌词应同步移动，并高亮显示当前的同步歌词语句；

歌词同步采用读取 txt 文本格式的 LRC 歌词，对其进行解析和同步以及高亮显示。在实现音乐试听中的同步显示歌词时需要获取歌词的行数。在 LRC 歌词中，通过：

[MM:SS.MS]指定时间；

[ar: 演唱者名]；

[ti:歌曲名]指定歌曲名；

[al:专辑名]指定专辑名；

[by:歌词编辑者]指定歌词编辑者；

[Offset: MS]用于调整整个歌词文件的时间标签值，单位是毫秒，可以是负值，它也是 LRC 歌词格式中，唯一可以使用负值的时间标签；在 LRC 歌词格式中，每一个括号对应一行歌词。

3.2.4 连续播放歌曲

主要包括：用户选择多首歌曲，进行勾选后连续播放。本连续播放仅限同板块的列表歌曲；

4.2.5 下载歌曲

本模块主要功能是对用户喜爱的歌曲提供方便的下载服务，点击歌曲右方的下载按钮，歌曲可以保存到用户本机。

4.2.6 搜索歌曲

该搜索主要是针对用户进行的，用户可以输入任意的歌名信息或者关键字，便可在本站快捷

的搜索到歌曲的相关信息，可以进行试听、下载等后续操作。

4.2.7 排行榜

主要包括：试听排行和下载排行；

-- **试听排行：**罗列了近期试听频率最高的一些歌曲，用户可以查阅该排行榜进行选择试听，走在时尚前沿的前列；

-- **下载排行：**罗列了近期本站下载次数较高的一些歌曲，用户也同样可以进行查阅，并了解歌曲进行试听等操作；

对于排行榜，本系统采用了图形和报表相结合的方式，方便用户直观的看到各种排行情况。

4.2.8 歌曲评论

本模块是方便普通用户对歌曲进行评论。

用户登录后，在自己喜爱的歌曲中可以进行歌曲评论，方便其他用户对本歌曲的好听与否进行直观的参考查阅。

4.2.9 分页显示歌曲列表

本模块为管理员操作的模块，管理员可以进入后台功能模块，分页检索到所有歌曲列表，方便管理和查阅。

4.2.10 多条件查询歌曲

本模块为管理员操作的模块，本查询不同于用户查询。

该多条件查询可以通过歌手名、歌曲名、专辑名等多条件进行查询和组合查询，方便管理员精确定位歌曲资源。

4.2.11 添加歌曲

本模块为管理员操作的模块。

管理员可以添加歌曲的相关信息，如：歌曲名、歌手名、专辑名称、上传歌曲文件，上传歌词文件。

- (1) 检测歌曲是否添加：管理员在添加歌曲资源前，本模块有程序判断的功能，判断其歌曲是否被添加过，以免造成冲突和磁盘资源的浪费；
- (2) 上传歌曲文件：上传歌曲文件到服务器，做为本音乐平台的主要资源来源；
- (3) 上传歌词文件：对歌曲进行歌词的上传工作，为歌曲播放时的歌词同步功能做资源支持；
- (4) 保存歌曲信息：管理员可以后台添加歌曲的信息，对该歌曲的歌手名、专辑名等进行录入；

特别要求，本模块的上传歌曲文件和上传歌词文件，要求应用 jspSmartUpload 组件实现文件重命名上传。为了实现歌词同步现象，需要保证歌词文件和歌曲文件同名，这样上传歌词文件时，就需要将歌词文件重命名。在应用 jspSmartUpload 组件上传文件时，可以通过文件上传

组件的 `getFiles()` 方法获取 `Files` 类的对象，然后通过 `Files` 类的 `getFile()` 方法获取文件集合中指定的文件对象，再通过该文件对象的 `saveAs()` 方法将文件重命名上传即可。

4.2.12 删除歌曲及歌曲文件

本模块为管理员操作的模块。

管理员可以删除歌曲以及歌曲对应的相关文件，本删除属于级联删除，即：删除歌手后，歌手对应的专辑信息，歌曲信息，歌曲对于的歌词信息一并删除。

4.2.13 资源信息采集

本模块属于系统的附加功能，利用时下流行的网络爬虫技术爬取其他网站的资源信息，为我所用。

网络爬虫是一个自动提取网页的程序，它为搜索引擎从万维网上下载网页，是搜索引擎的重要组成部分。传统爬虫从一个或若干初始网页的 URL 开始，获得初始网页上的 URL，在抓取网页的过程中，不断从当前页面上抽取新的 URL 放入队列，直到满足系统的一定停止条件。

本系统中，利用能够从互联网上自动收集数码产品网站网页的网络蜘蛛程序，自动访问互联网，并沿着数码产品网站任何网页中的所有 URL 爬到其它网页，重复这过程，并把爬过的所有数码产品信息网页收集到“本地网页库”中。网络爬虫从数码产品网站上抓取网页，把网页送入“本地网页库”，从网页中“提取 URL”，把 URL 送入“URL 种子队列”，爬虫控制得到网页的 URL，控制网络爬虫抓取其它网页，反复循环直到把所有的网页抓取完成。

--收集 URL 种子：分析音乐网站的网页地址结构，URL 规则，选取有代表性 URL，通过此 URL 能够尽量达到更多其它页面，将这些 URL 作为种子。从首页或者给定网页中提取种子 URL；或者直接由用户手动输入种子 URL。

--爬虫配置：爬虫配置为网页抓取提供配置信息，包括 URL 种子队列、爬虫目标网站页面 DOM 结构、URL 规则、页面信息提取的正则表达式等。这些信息通过配置保存于 XML 文件中，爬虫程序根据该 XML 文件提供的信息规则来执行页面信息抓取任务。

当爬取多个网站时，爬虫配置控制爬虫不需手动输入起始地址，爬虫按照配置策略对网页信息进行抓取，爬虫自动读取配置文件，自动获得相关配置信息，自动爬取。

爬虫配置同时可以配置爬虫运行条件，用以控制爬虫何时开始爬取，何时停止运行等策略。

--网页抓取：根据网络爬虫配置规则，从大量页面中提取与数码信息相关的 URL，按照某特定的算法决定访问页面顺序，沿着网页中的所有 URL 爬到其他网页，并把爬过的所有网页内容下载到本地服务器中，即本地网页库。

4.2.14 非功能需求

- ◇ **可用性：**系统具备傻瓜式操作特点，好记易学、实用高效、令人满意；
- ◇ **可靠性：**系统应经过完善的设计和充分的测试运行，具备在较长时间内连续无故障的运行能力；

- ◇ **健壮性：**系统应具备强大的容错、数据恢复与稳定运行的能力；
- ◇ **性能：**系统在响应时间、数据吞吐量和持续高速性等多方面提供较高性能的数据处理和查询服务。
- ◇ **安全性：**系统应提供全面、有效的系统安全机制，能有效防止病毒感染、非法用户或恶意程序对应用系统或数据的入侵；
- ◇ **扩展性：**系统应易于扩展和升级，能够根据具体需求快速、方便地定制、扩展原系统的功能，以更好地满足档案管理的新增和变更的需求；
- ◇ **开放性：**系统应具备开放的标准化体系结构，可方便地与其它业务系统衔接，实现与其它业务系统间的无缝集成；
- ◇ **先进性：**系统应采用业界先进、主流的档案数据管理、网络管理及信息安全技术，具备较强的可用性、可靠性、健壮性、性能、安全性、扩展性和开放性。

4.3 系统逻辑结构

系统的核心组件包括歌曲显示组件、试听歌曲并同步歌词组件、连续播放歌曲组件、下载歌曲组件、搜索歌曲组件、排行榜组件、歌曲评论组件、分页显示歌曲组件、多条件查询歌曲组件、添加歌曲上传组件、删除歌曲组件、资源信息采集组件等。系统逻辑分层组件结构如下图所示：

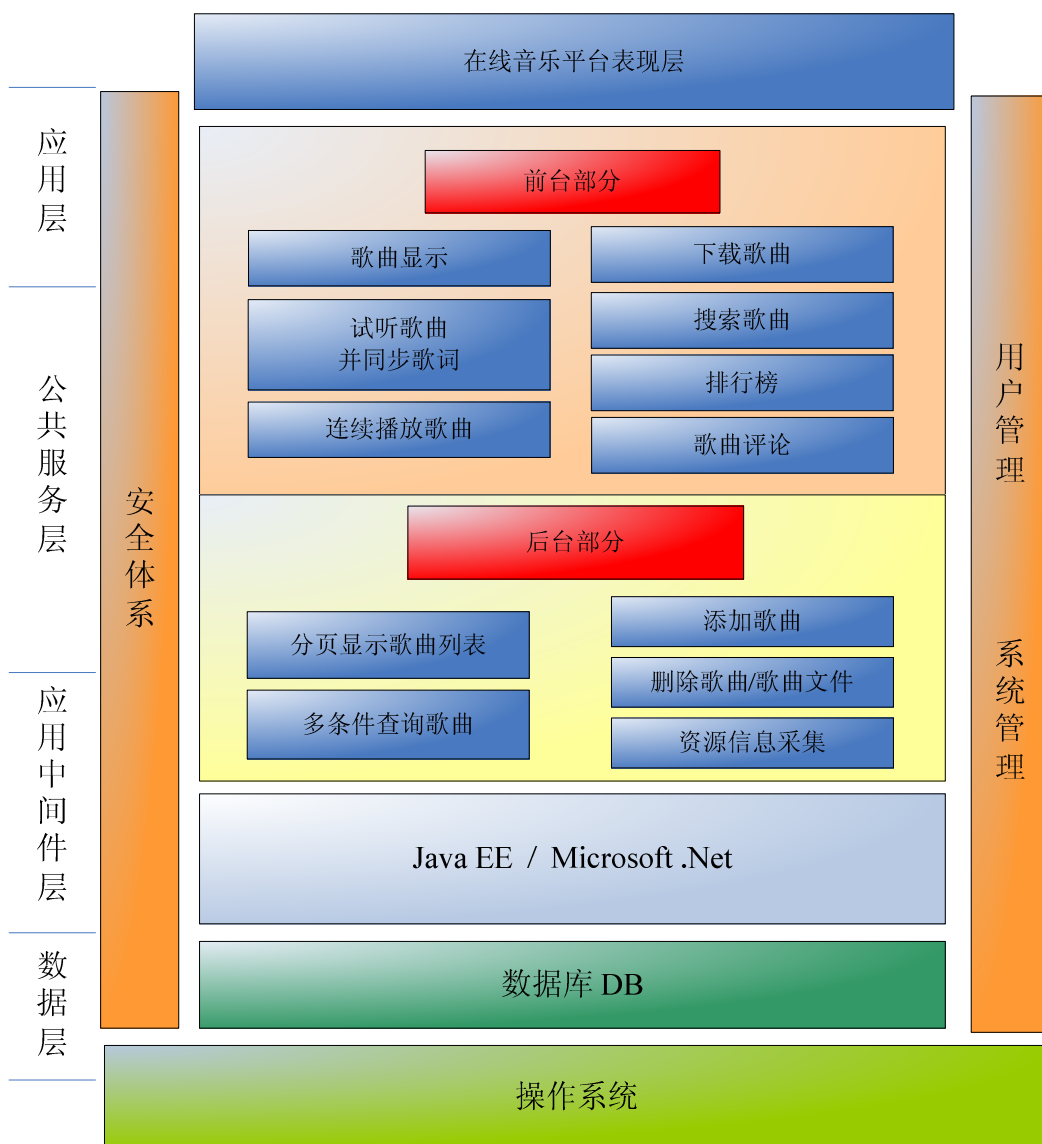


图 4 系统逻辑结构

5. 关键技术与难点

- B/S 架构。
- 信息管理平台：
 - 1) J2EE 平台：Eclipse+Tomcat+SSH 框架
 - 2) .NET 平台：VS.NET 2008+IIS+ADO.NET+MVC 框架
- 数据报表统计生成技术（报表格式可自定义）
JFreechart 或 JFreeReport 或 Crystal Reports 等
- 支持多种数据库平台（MySQL、SQL Server、Oracle 等）
- IO 流操作

- Ajax 技术的使用

- 数据交换平台接口规范

- 验证客户端是否安装 Windows Media Player 和 Real Player 播放器的方法

要想通过网页在客户端播放音乐或视频，客户端必须安装指定的播放器，如果客户端没有安装指定的播放器，将不能播放音乐或视频。这时为了当客户端机器上没有安装指定的播放器时给予相关提示，可以在程序中添加检测客户端是否安装播放器的功能。

- 调用 Windows Media Player 和 Real Player 播放器播放歌曲

通过 html 提供的<object>标记可以调用指定媒体播放器播放多媒体文件。

调用 Window Media Player 时 classid 属性值为：

```
<OBJECT id=MediaPlayer height=64 width=546
```

```
classid=CLSID:6BF52A52-394A-11d3-B153-00C04F79FAA6 name=' wghMediaPlayer' >
```

- 获取 LRC 歌词的行数

在实现音乐试听中的同步显示歌词时需要获取歌词的行数。

在 LRC 歌词中，通过[MM:SS.MS]指定时间；

[ar: 演唱者名]；

[ti:歌曲名]指定歌曲名；

[al:专辑名]指定专辑名；

[by:歌词编辑者]指定歌词编辑者；

[Offset: MS]用于调整整个歌词文件的时间标签值，单位是毫秒，

可以是负值，它也是 LRC 歌词格式中，唯一可以使用负值的时间标签；在 LRC 歌词格式中，每一个括号对应一行歌词。

- 应用 jspSmartUpload 组件实现文件重命名上传

为了实现歌词同步现象，需要保证歌词文件和歌曲文件同名，这样上传歌词文件时，就需要将歌词文件重命名。在应用 jspSmartUpload 组件上传文件时，可以通过文件上传组件的 getFiles()方法获取 Files 类的对象，然后通过 Files 类的 getFile()方法获取文件集合中指定的文件对象，再通过该文件对象的 saveAs()方法将文件重命名上传即可。

- 防止被盗链下载歌曲文件

防止网站中的歌曲被盗链下载可以应用请求头中的 referer 字段，该字段记录了当前请求的上一次访问地址，在需要被保护的页面或请求响应中，通过判断请求头中的 referer 字段的内容来得知当前的请求是不是从其他网站中触发的，从而达到防止网站中的歌曲被盗链下载的目的。

- 爬虫技术的研究

网络爬虫是一个自动提取网页的程序，它为搜索引擎从万维网上下载网页，是搜索引擎的重要组成部分。传统爬虫从一个或若干初始网页的 URL 开始，获得初始网页上的 URL，在抓取网页的过程中，不断从当前页面上抽取新的 URL 放入队列，直到满足系统的一定停止条件。

聚焦爬虫（主题爬虫）的工作流程较为复杂，需要根据一定的网页分析算法过滤与主题无关的链接，保留有用的链接并将其放入等待抓取的 URL 队列。然后，它将根据一定的搜索策略从队列中选择下一步要抓取的网页 URL，并重复上述过程，直到达到系统的某一条件时停止。另外，所有被爬虫抓取的网页将会被系统存贮，进行一定的分析、过滤并建立索引，以便之后的查询和检索。对于聚焦爬虫来说，这一过程所得到的分析结果还可能对以后的抓取过程给出反馈和指导。

本系统中，利用能够从互联网上自动收集数码产品网站网页的网络蜘蛛程序，自动访问互联网，并沿着数码产品网站任何网页中的所有 URL 爬到其它网页，重复这过程，并把爬过所有数码产品信息网页收集到“本地网页库”中。网络爬虫从数码产品网站上抓取网页，把网页送入“本地网页库”，从网页中“提取 URL”，把 URL 送入“URL 种子队列”，爬虫控制得到网页的 URL，控制网络爬虫抓取其它网页，反复循环直到把所有的网页抓取完成。

6. 行业知识

1、网络信息化

涉及计算机网络、信息管理系统、多媒体技术。

2、各种规范

《互联网文化管理暂行规定》；

《关于规范网络音乐市场秩序、整治网络音乐网站违规行为的通告》；

7. 系统工作量估计

本项目总的实施工作量为 495 人天（22.5 人月），具体工作量分析如下表所示：

模块	功能	工作量 (人天)	合计	备注
歌曲显示	分类显示最新 10 首歌曲	15	35	
	按类别分页显示各类歌曲	20		本模块涉及分类较多，如：歌手、专辑、华语、欧美、日韩等。
试听歌曲并同步显示歌词	歌曲播放	20	60	调用播放器进行播放。
	候选列表控制	20		候选列表中，罗列待播放的歌曲。
	歌词同步显示	20		歌曲播放的同时歌词同步高亮显示。
连续播放歌曲		30	30	涉及各个分类板块的歌曲

				连播。
下载歌曲		20	20	
搜索歌曲		20	20	提供用户级别的歌曲搜索功能。
排行榜	试听排行	10	10	
	下载排行	10	10	
分页显示歌曲列表		25	25	分页显示歌曲列表，以满足后台管理员的操作。
多条件查询歌曲		30	30	本查询为多条件综合查询，方便管理员做一系列操作。
添加歌曲	检测歌曲是否添加	10	60	
	上传歌曲文件	25		
	上传歌词文件	15		
	保存歌曲信息	10		
删除歌曲及歌曲文件		20	20	
资源信息采集	收集 URL 种子	20	80	
	爬虫配置	20		
	网页爬取	40		
其他	界面美化	25	25	
	开发文档	50	50	
	性能优化	20	20	
合计	495 人天 (22.5 人月)			

注：以上工作量估算是按照人员专业水平至少达到软件工程师水平，并且有 2~3 年实际开发经验的人员。