

基于协同过滤算法的嘻哈之家平台设计

杨 灿¹, 郭 龙^{1,2}, 杨秀涛¹, 杨 峰¹, 袁 洪¹, 翟德坤¹

(1. 贵州师范学院 数学与大数据学院, 贵州 贵阳 550018;

2. 贵州师范学院 大数据科学与智能工程研究院, 贵州 贵阳 550018)

摘 要: 基于协同过滤算法的嘻哈之家平台专为国内众多嘻哈文化喜好者而开发。平台的研发主旨是为这一群体提供一个可靠途径以实现整合市场、文化交流、商演推广等。该平台采用协同过滤算法对用户进行个性化推荐, 同时还涵盖 BBS 论坛、海量数据存储技术、实时网络分享技术。嘻哈之家平台采用 MVC 设计模式、B/S 架构、tomcat 服务器, 以 Java 作为主要开发语言, 打造一个整合性平台, 最终促进国内嘻哈文化发展。

关键词: 嘻哈之家; 文化推广; 协同过滤算法; MVC; 个性化推荐; BBS

中图分类号: TP391

文献标识码: A

文章编号: 2095-1302 (2020) 07-0085-03

0 引言

近年来国内音乐被嘻哈浪潮席卷, 嘻哈文化爱好者数量激增, 从而为嘻哈文化的发展奠定了一个庞大的基数。但这一群体缺乏整合, 分散在国内众多群体中。大多数采用的国内嘻哈文化推广方式都是线下演出, 推广力度较小。如今国内环境缺少一个专属的、规范的、全面的嘻哈文化平台, 导致国内嘻哈文化发展一直处于停滞阶段。嘻哈之家平台是国内外嘻哈文化浪潮孕育下产生的一种新媒介平台^[1-2], 其优势便在于使用更加现代化的网络方式进行推广, 有效扩大嘻哈文化的影响范围, 同时为这一人群提供出整合的平台^[3]。

1 项目背景

1.1 总体背景

在高度发展的媒体时代, 大量新潮文化相互渗透、相互碰撞。这使得嘻哈音乐的发展受到限制, 无法进入更多人的眼球。因此需要一个助推器加速成长。嘻哈之家平台的推出的意义就在于突破局限, 展现出嘻哈文化所蕴含的价值, 并促使新潮流风向的产生。这一平台将学习交流、嘻哈知识普

及、音乐作品的分享等方面集为一体, 使得大众可以更加全面并且系统地进行学习与实践^[4]。并且能帮助学习者更快速、更精确地解决学习中的盲点和交流方面的障碍。嘻哈之家作为多元素综合的一个新媒介平台, 将嘻哈文化向着流行文化般的传播度进行转变, 让更多人熟知。

1.2 国内外现状分析

嘻哈文化于 20 世纪 70 年代诞生于美国, 现已在当地进入成熟阶段; 20 世纪 80 年代左右传入日本与韩国, 而后发展至盛行状态; 20 世纪 90 年代传入中国, 在国内属于小众文化范畴^[5]。国内嘻哈文化发展缓慢的原因在于缺少有效的传播媒介, 嘻哈之家平台提出的必要性便在于此。

2 系统意义

2.1 文化交流

国内嘻哈文化市场一直缺少一个代表本土特色嘻哈文化的平台。嘻哈之家平台为广大爱好者提供了一个交流中枢, 可以自由地进行交流以及相互学习, 相互提高。将本土特色嘻哈文化推出, 由小众走向大众。

2.2 商演推广

平台会发布专属音乐节活动以进一步推广, 让大家也能感受到线下嘻哈的氛围。线上与线下的结合将使嘻哈文化得到最大程度的推广, 扩大喜爱这一文化的人群数量, 并逐步增强在大众心中的影响力。

2.3 整合市场

由于国内嘻哈市场一直处于离散状态, 市场无法得到有效的集中, 无法形成一定的影响力。平台将对分布于嘻哈文化行业中的人群进行整合, 使得嘻哈文化影响力得到提高, 进一步提高国内嘻哈文化的整体性, 让嘻哈文化在国内得到

收稿日期: 2020-02-11 修回日期: 2020-03-10

基金项目: 贵州省 2018 年大学生创新创业训练计划项目 (2018520558); 贵州省省级重点学科“计算机科学与技术”(ZDXK[2018]007号); 2016 年贵州省省级重点支持学科“计算机应用技术”(黔学位合字 ZDXK[2016]20号); 贵州师范学院大学生互联网+创新创业训练中心(黔教高发[2015]337号, 黔教高发[2017]158号); 贵州省教育厅创新群体重大研究项目(黔教合 KY 字[2016]040); 贵州省普通高等学校工程研究中心(黔教合 KY 字[2016]015)

有效发展。

2.4 促进国内嘻哈文化发展

国内嘻哈文化发展缓慢的主要因素在于推广方式一直是传统线下演出,要发展国内嘻哈文化必然要将其网络化,使其得到最大程度的推广。将具有国内特色的嘻哈文化推向大众视野,让大众得到更大程度了解。

3 功能及设计要素

3.1 设计目标

嘻哈之家是一个嘻哈文化的媒介平台。用户可以集学习、娱乐、交流为一体,并且陪伴用户完成学习期间的一切困扰,降低以往大众接触嘻哈文化的门槛,让更多人群认识嘻哈文化。

3.2 功能介绍

系统功能模块如图1所示。

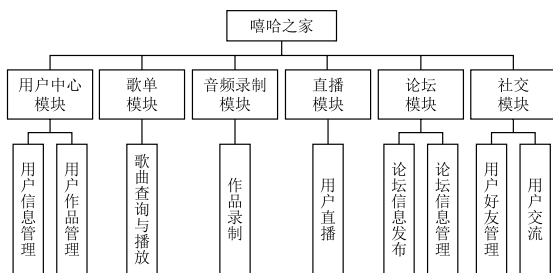


图1 系统功能模块图

(1) 用户中心模块:用于用户进行信息修改和空间建设。

(2) 歌单模块:对所有的嘻哈音乐进行分类,比如 jazz, trap, oldschool 等。

(3) 音频录制模块:用户可在平台上进行录制与分享,让其余用户可以欣赏与点评自己的作品。

(4) 直播模块:每个用户可以在直播间进行直播,分享自己的技巧等^[6]。

(5) 论坛模块:一个自由发表论坛的模块,用户可以自由发表自己的文章,比如求助说唱技巧方面的经验等。

(6) 社交模块:用户可以自由地和感兴趣的人进行交流并成为好友。

3.3 程序设计

3.3.1 设计模式

如图2所示,使用MVC(Model View Controller)设计模式。按照代码功能分为三层,即模型层(Model)、显示层(View)、控制层(Controller)。这样会使维护更加方便,代码也不会显得繁琐^[7]。

3.3.2 系统设计

系统的前端方面采用Bootstrap框架做响应式设计,保证系统能自适应于移动设备。各功能子系统分别为用户中心

子系统、歌单子系统、音频录制与发布子系统、直播子系统、论坛子系统、社交子系统。

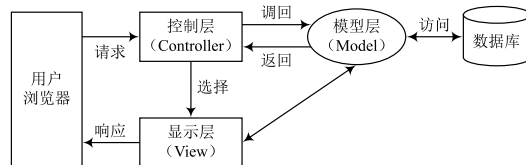


图2 MVC模式图

4 技术特点

4.1 个性化推荐

在系统中通过用户的历史行为来构建数据集,再通过协同过滤算法对用户的喜好进行预测,从而为用户计算出符合用户喜好的音乐以及视频形成个性化推荐。

4.1.1 协同过滤推荐^[8-9]

系统采用基于物品的协同过滤推荐,是指为用户推荐出与之前所感兴趣物品相似的物品。就本系统而言,通过用户对不同作品的评分记录来构建模型,根据作品间的相似度来进行推荐^[10]。

4.1.2 余弦相似度

在相似度函数方面选用余弦相似度(Cosine similarity)进行相似度计算,如下:

$$\cos \theta = \frac{\sum_{i=1}^n (A_i + B_i)}{\sqrt{\sum_{i=1}^n A_i^2} + \sqrt{\sum_{i=1}^n B_i^2}}$$

式中, A_i 与 B_i 即代表不同用户对作品A与作品B的不同评分,带入公式得出余弦值。余弦值越接近1则代表相似度越高即值得推荐,余弦值越接近0则相似度越低即不值得推荐。

4.1.3 具体实现

首先对获取的数据进行处理,将数据处理完毕后。通过Mahout调用基于物品的协同过滤推荐算法,并且选用余弦相似度作为相似度函数,从而实现个性化推荐效果,部分程序如下:

```
public class MusicRecommend {
    public static void main (String[] args) throws Exception {
        DataModel model = new FileDataModel (new File ("D : \\hadoop
数据 \\ratingsout.csv"));

        // 余弦相似度
        ItemSimilarity similarity = new UncenteredCosineSimilarity
(model);

        // 构造推荐
        Recommender recommender = new
GenericItemBasedRecommender (model, similarity);
```

```
List<RecommendedItem>items = recommender.recommend (10,5);  
LongPrimitiveIterator iter = model.getUserIDs ();  
while (iter.hasNext ()) {  
    longuserid = iter.nextLong ();  
    List<RecommendedItem>list = recommender.recommend (userid,  
3);  
    System.out.println ("userid="+userid+" : ");  
    for (RecommendedItem item : list) {  
        System.out.println ("["+item.getItemID () + " : "+item.getValue ()  
+"]");  
    }  
    System.out.println ();  
}  
}
```

上述部分代码是个性化推荐的部分代码，通过调用对用户感兴趣的作品进行预测，便于对用户推送。

4.2 BBS 模块

在这一平台中推出了 BBS 模块^[11]。用户可在论坛中发起自己感兴趣的帖子，其余用户可以进行评论。贴主具有增删查改的权限，使得信息的合理性得到保证。同时后台管理员可监控所发布的信息，保证信息的合法性和合理性。

4.3 海量数据存储

在存储日志以及用于个性化推荐计算的数据时，采用 hive 进行存储。hive 具有对数据集的扩展能力与计算能力，这样一来在应对庞大数据时计算效率将得到提高。同时，将很少发生数据变化的页面缓存起来，每次需要时，就不必再重新加载，从而腾出更多的 CPU 资源^[12]，减小建立数据库的压力，也能有利于加快网站的运行速度。在建立数据库时，会对所有数据库表结构进行优化，使得 SQL 语言更加严谨精炼，索引也更快速更准确。

4.4 实时网络分享

平台会把网络上搜索次数较多的作品放到热搜榜单中，每天定时刷新数据，再统计该榜单中访问次数最多的作品，将最好作品放到显而易见的页面上供大众欣赏。系统所设置的分享功能，只要用户注册账号，就能拥有自己的“空间”，可以在空间分享自己喜欢的作品，也可以上传自己的作品进行分享，可添加好友、实时聊天分享。

4.5 系统架构 (B/S 架构)

如图 3 所示，在系统架构方面选择了 B/S 架构^[13]。现如今软件更新频率不断上升，增加了工作人员的负担，假设工作人员对数量为上千台电脑中的软件进行维护，工作效率将

非常低下。但如果是采用 B/S 架构的软件，用户使用的只是浏览器端，并不需要做维护，不管用户规模多大，所有操作只需要对服务器进行操作，只需对服务器进行管理就可以，大大减少工作负担。

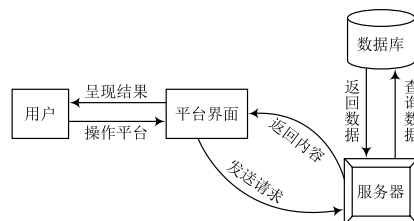


图 3 B/S 架构图

5 结 语

嘻哈之家平台线上加线下的推广模式在嘻哈文化的推广中独树一帜。其线上和线下的各个模块囊括了当今人群文化交流的大部分途径，所以其完全可以满足大众输入和输出嘻哈文化的需求。而便捷的平台操作毫无疑问能大大增加用户的使用兴趣，实现快速入门到精通。精致的内容和包装终将促进嘻哈之家平台的迅速传播。

参 考 文 献

- [1] 陈敏. 中国嘻哈文化的本土化及其启示 [J]. 中华文化论坛, 2013 (7): 162-165.
- [2] 李凡卓. Hip-Hop 文化在我国青少年中流行的社会学分析 [J]. 青少年研究, 2004 (2): 7-9.
- [3] 李姗. 关于“互联网+”音乐的思考与探索: 基于音乐视角的考察 [J]. 北京联合大学学报 (人文社会科学版), 2016 (1): 73-78.
- [4] 王俊霞. 浅谈网络音乐对音乐教育的影响 [J]. 艺术教育, 2016 (2): 82-83.
- [5] 郭佳珺. 嘻哈乐在中国的发展探析 [J]. 当代音乐, 2019 (9): 177-178.
- [6] 朱莹. 基于 Web 的视频直播系统的应用研究 [J]. 电脑知识与技术, 2009 (4): 968-970.
- [7] 李展飞. Web 软件系统开发框架设计在 MVC 模式的实现 [J]. 电子技术与软件工程, 2017 (8): 61.
- [8] 马宏伟, 张光卫, 李鹏. 协同过滤推荐算法综述 [J]. 小型微型计算机系统, 2009 (7): 1282-1288.
- [9] 蒲鲜霖. 智能推荐系统中协同过滤算法综述 [J]. 中国新通信, 2018 (23): 31-32.
- [10] 王炳祥. 基于协同过滤的歌曲推荐算法研究 [J]. 数字技术与应用, 2019 (10): 126-127.
- [11] 何博文, 郑剑. BBS 论坛系统设计与实现 [J]. 电脑知识与技术, 2018 (18): 42-44.
- [12] 王成, 李少元, 郑黎晓, 等. Web 前端性能优化方案与实践 [J]. 计算机应用与软件, 2014 (12): 89-95.
- [13] 史瑞刚, 周亮, 秦琴琴, 等. 基于 B/S 的网络教育管理系统的设计与实现 [J]. 信息技术与信息化, 2019 (5): 177-180.