

基于协同过滤算法的自动化隐式评分音乐双重推荐系统

李涛, 符丁

(黔南民族师范学院, 贵州 都匀 558000)

摘要: 提出了基于协同过滤算法的自动化隐式评分音乐双重推荐系统; 在异构普适环境推荐框架下, 对系统总体结构进行设计; 其中硬件部分采用四元件组成方式, 使用 W900710 型号芯片作为播放器核心板, 并将隐式评分提取器与推荐引擎结合起来, 可避免噪声干扰; 而软件部分设计场景模拟衰减现象, 采用协同过滤算法描述衰减过程, 根据描述结果, 设立双重推荐机制来实现抗人为影响的音乐双重推荐系统; 由实验结果可知, 对于大规模音乐数据推荐具有良好可扩展性。

关键词: 协同过滤; 自动化; 隐式评分; 音乐; 双重推荐; 普适环境

An Automated Implicit Scoring Music Dual Recommendation System Based on Collaborative Filtering Algorithm

Li Tao, Fu Ding

(Qiannan National Minorities Normal College, Duyun 558000, China)

Abstract: According to the influence of the display scoring mechanism adopted by the traditional system, which is influenced by the external interference and the precision of the recommended results is low, a dual recommendation system based on collaborative filtering algorithm is proposed. Under the recommendation framework of heterogeneous ubiquitous environment, the overall structure of the system is designed. The hardware part is composed of four components, using W900710 model chip as the core of the player, and combining the implicit score extractor with the recommendation engine, the noise interference can be avoided. The software part design scene simulation attenuation phenomenon, use collaborative filtering algorithm to describe the attenuation process, according to the description results, set up a double recommendation mechanism to achieve the anti human impact of the music dual recommendation system. The experimental results show that the maximum precision of the proposed system can reach 90%, and it has good scalability for large-scale music data recommendation.

Keywords: collaborative filtering; automation; implicit rating; music; dual recommendation; universal environment.

0 引言

由于信息不断增加与更新, 使网络结构中信息量以指数规律迅速扩展, 因此形成了信息爆炸显式社会网络结构, 对于音乐爱好者来说, 网络上信息过载成为了严重问题, 爱好者在网络复杂结构中迷失航向, 无法明确所要寻找信息位于信息空间中的具体位置, 也无法返回到某个节点, 再加上网络信息的复杂性, 限制爱好者的认知能力, 造成用户无法获取满意音乐信息^[1]。为了解决这些问题, 需将网络从被动浏览转化为主动浏览, 实现网络结构对用户提供主动的信息服务, 为此, 构建了基于协同过滤算法的自动化隐式评分音乐双重推荐系统。采用协同过滤算法的自

动化隐式评分结构能够反映出不同用户之间对音乐兴趣以及内容之间关联性, 为具有共同兴趣用户推荐相似内容的可能性大大增加。对于同一用户对音乐喜欢的具体内容也是具有相似特征关联的, 其音乐种类也具有风格和品质。基于用户自身、用户与用户之间的联系, 可通过内容推荐方式, 引导用户对音乐内容访问行为, 进而适应网络能力, 实现网络带宽利用效率的优化, 根据用户兴趣爱好, 推荐满足用户需求的信息^[2]。

1 异构普适环境的推荐框架

在普适环境下, 人们能够任意使用设备来获取相关信息, 基于普适环境背景下构建异构普适环境推荐框架, 为自动化隐式评分音乐双重推荐系统研究奠定基础^[3]。异构普适环境推荐框架设计如图 1 所示。

普适环境结合了网络技术与移动技术, 设计了面向客户的自适应推荐结构。普适环境是由网络设备与服务共同组成的, 其中网络设备包括计算机、手机和各种网络连接家电等, 网络服务包括计算、管理与控制等。在该环境下, 网络可收集来自用户、管理员的查询、配置与管理信息,

收稿日期: 2018-05-05; 修回日期: 2018-05-30。

基金项目: 黔科合 J 字 LKQS[2013]23 号; 黔科合 LH 字 [2014]7439 号。

作者简介: 李涛(1974-), 男, 安徽芜湖人, 硕士, 讲师, 主要从事软件工程及机器学习方向的研究。

符丁(1964-), 男, 湖南岳阳人, 硕士, 副教授, 主要从事软件工程方向的研究。

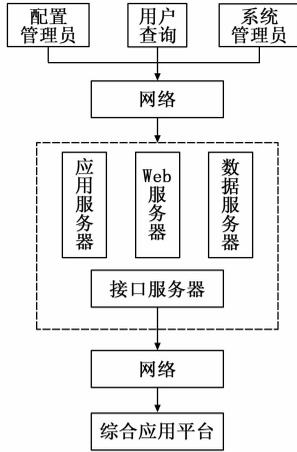


图 1 异构普适环境推荐框架

将这些信息传递到各个服务器端口，再次通过网络应用到综合平台，为推荐系统设计提供基础^[4]。

2 自动化隐式评分音乐双重推荐系统

基于异构普适环境推荐结构的自动化隐式评分音乐双重推荐系统设计如图 2 所示。

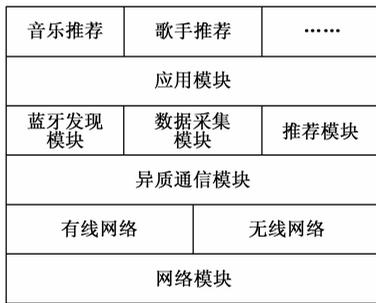


图 2 自动化隐式评分音乐双重推荐系统

由图 2 可知：双重推荐系统是由应用、异构通信和网络模块构成的，其中核心部分主要位于异构通信模块之中。采用协同过滤算法，从一个节点中提取信息，随机发送给另外一个节点，重复该动作，使其他节点接收到信息^[5]。基于协同过滤算法的信息传送方式完全模拟人们通过网络的日常交流方式，保证用户信息的高效率传输。其中蓝牙模块是连接外围设备的统一桥梁，采用无线连接方式可实现用户自然享受信息推荐服务；数据采集模块是收集可推荐给用户的主要信息，也可根据用户偏好作出个性化推荐服务。对于不同推荐系统，需要收集的内容也是不同的，音乐推荐需要收集用户音乐播放器中的数据，而歌手推荐需要收集音乐播放器中的歌手信息^[6]。根据推荐内容分别将信息存储到服务器上，用户可将自己感兴趣的数据保存到播放器中，通过设备之间的交互实现数据共享；推荐模块是通过数据采集到的用户偏好信息来推荐内容，采用协同过滤算法获取合适的推荐列表给用户。

2.1 硬件设计

如图 3 所示，音乐双重推荐系统硬件结构是由 4 个元件

组成的。

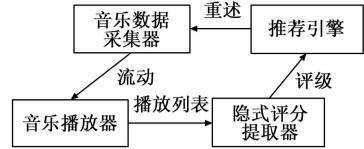


图 3 音乐双重推荐系统硬件结构

根据图 3 分别对音乐播放器和隐式评分提取器进行设计。

2.1.1 音乐播放器

在异构普适环境下，用户界面友好性展示的更加明显，根据用户接入的电子设备以及接入方式，对用户端界面做出适当调整^[7]。具体设计如图 4 所示。

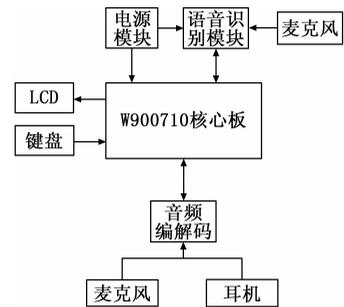


图 4 音乐播放器

由图 4 可知：采用 W900710 型号的核心板与电源模块、语音识别模块、LCD、键盘和音频相连。利用麦克风向播放器中输入语音数据，通过语音识别模块对数据进行识别；利用麦克风和耳机向音频编解码器输入语音数据，通过音频编解码器对数据进行处理^[8]。最终将所有数据集中在 W900710 型号的核心板上，通过 LCD 显示在屏幕上。

2.1.2 隐式评分提取器

根据音乐数据采集器采集到的用户信息，设计一种隐式评分提取器，专门对用户隐式信息进行评分。在普适环境中分别提取歌手与歌曲这两种项目评分，并在移动环境中充分考虑资源分布情况，通过推荐引擎，将用户浏览过的多种浏览器和自定义播放器信息进行集中处理，采用协同过滤算法使网络拓扑结构中的每个节点都部署一个智能代理节点。结合普适环境，将经过音乐数据采集器上的信息分散在各个服务器之中^[9]。当用户进入到普适推荐环境之中，服务器会根据用户以往搜索记录，将可感兴趣的歌曲或歌手推荐到用户设备上，将这些数据在客户端与服务端中进行交互，采用协同过滤算法不会受到噪声干扰影响得以共享^[10]。隐式评分提取内部电路图如图 5 所示。

由图 5 可知：当引脚为高电平时，Q1 三极管截止，电路不工作；当引脚为低电平时，Q1 三极管导通，电路工作，向外发射信号，经过三极管将电流放大，发射载波所占的比例增加，保证采集器能够容纳更多信息，使评分结果更准确。

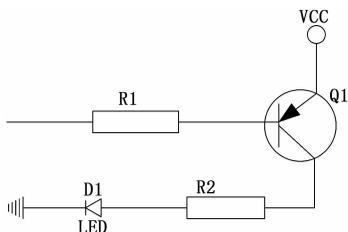


图 5 隐式评分提取电路连接

音乐双重推荐系统是由音乐数据采集器、音乐播放器和隐式评分提取器和推荐引擎组成的, 其中音乐数据采集器可为系统提供海量数据; 音乐播放器可为系统提供清晰推荐结果; 将隐式评分提取器与推荐引擎结合起来, 可为推荐系统提供准确评分结果。采用协同过滤算法, 尽可能利用集中式推荐效果, 以便达到异构环境下的推荐服务一致性^[1]。

2.2 软件设计

针对硬件中的隐式评分提取器, 对推荐系统软件功能进行设计。在构建协同过滤自动化隐式评分音乐双重推荐系统时, 需先统计用户听歌历史记录, 从该记录中提取用户的音乐偏好信息, 根据提取的信息来决定该用户对歌曲以及歌手喜爱情况。采用双重提取策略, 首先对所有歌曲进行隐式评分, 根据评分结果提取出得分相对较高的歌曲; 然后分析这些歌曲的类型, 定位到某种唱这种风格歌曲的歌手, 再次将这些歌曲与歌手发行的歌曲进行整合处理; 最后将整合结果做成推荐列表, 推荐给用户。在此过程中, 利用协同过滤算法, 为用户提供合理推荐建议。

自动化隐式评分是指在用户寻找感兴趣信息过程中, 系统追随用户行为来分析用户兴趣, 该过程只需通过用户浏览某个网站时, 就可对某个阅读时间较长的页面保存下来, 以此说明该网页涉及用户感兴趣内容。区别于传统系统, 在该推荐系统中采用双重推荐机制, 综合考虑歌曲和歌手推荐, 使每个歌手很自然地融合为一个曲风载体。在实际生活中, 经常使用歌手名来表达用户对音乐偏好程度。如果用户音乐播放器播放的是“林俊杰”的歌, 那么可推断出该用户比较喜欢抒情类的歌曲。但是, 如果用户音乐播放器播放的是《一千年以后》—林俊杰某个专辑中的一首歌, 那么很难分辨出该用户的音乐偏好。为此, 采用协同过滤算法能够更精准度量出用户音乐偏好。具体流程如图 6。

对于特定用户来说, 假设用户能多听一次该歌手的歌曲, 那么相应隐式评分会大大增加, 因此, 每听一次歌曲的行为就要被赋予新的权重。设立如表 1 所示的场景来模拟衰减现象。

表 1 衰减模拟场景

用户	歌曲 1	歌曲 2
1	1 次	50 次
2	10 次	500 次

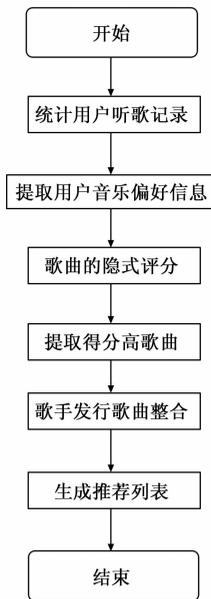


图 6 自动化隐式评分系统流程

由表 1 可推断出用户 1 不喜欢歌曲 1, 而用户 2 喜欢歌曲 1; 用户 1 和 2 都喜欢歌曲 2。用户 2 听的次数是用户 1 的 10 倍, 虽然获取的偏好相似度与实际完全相反, 但是后续听歌行为可增加隐形评分, 其权值大小是小于最初用户听歌时的权值大小。在推荐系统之中, 采用协同过滤算法来设计衰减函数来描述上述具体衰减现象。

假设听歌行为为次数为 n , 那么第 n 次听歌行为的权值为:

$$weight(n) = \alpha \times e^{-\beta \times n} \quad (1)$$

公式 (1) 中: β 为衰减速度; α 为控制歌曲的隐式评分程度; e 为常数。

假设用户 X 听歌曲 Y 的次数为 $n'(X, Y)$, 那么根据公式 (1) 的衰减公式, 可获取用户 X 对歌曲 Y 的双重推荐机制:

$$f = \sum_{n=1}^{n'(X, Y)} weight(n) = \alpha \times \sum_{n=1}^{n'(X, Y)} e^{-\beta \times n} \quad (2)$$

根据上述双重推荐机制, 将所有用户对于某个歌手或歌曲偏好程度作为向量, 计算两者之间的相似度, 获取歌手与歌曲的相似内容之后, 根据用户历史偏好预测当前用户还没有确定的歌曲种类, 获取一个排序的歌曲列表作为推荐项目。具体推荐内容如图 7 所示。

由图 7 可知: 对于歌曲 A, 根据用户历史偏好, 喜欢歌曲 A 的用户都喜欢歌手 C, 得出歌曲 A 与歌手 C 所唱的歌曲风格相似, 而用户 3 喜欢歌曲 A, 那么可推断出用户 3 也可能喜欢歌手 C。

根据上述推荐内容的需求用例设计隐式评分推荐功能, 给出自动化隐式评分音乐双重推荐功能代码如下:

```
import javax.sound.midi.*; // The midi package is necessary
public class MiniMusicApp {
```

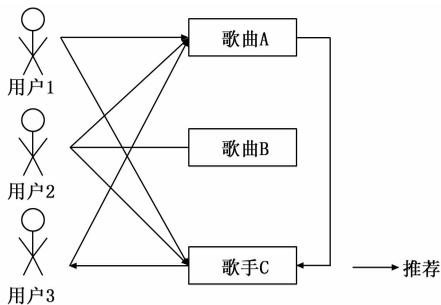


图 7 推荐内容

```
public static void main(String[] args) {
    MiniMusicApp mini = new MiniMusicApp();
    mini.play();
}
public void play() {
    try {
        // Get a sequencer and open it
        Sequencer player = MidiSystem.getSequencer();
        player.open();
        Sequence seq = new Sequence(Sequence.PPQ, 4); // Treat the
arguments as Ready-bake arguments
        Track track = seq.createTrack(); // Ask the sequencer for
a track
        // Put some MidiEvents into the Track, the setMessage()
method is what we should really care
        ShortMessage a = new ShortMessage();
        a.setMessage(144, 1, 44, 100);
        MidiEvent noteOn = new MidiEvent(a, 1);
        track.add(noteOn);
        ShortMessage b = new ShortMessage();
        b.setMessage(128, 1, 44, 100);
        MidiEvent noteOff = new MidiEvent(b, 16);
        track.add(noteOff);
        player.setSequence(seq); // Give the sequence to the Sequen-
cer
        // like pushing a CD to a CD player
        player.start(); // Start the sequencer like pushing PLAY
    }
    catch (Exception ex) {
        ex.printStackTrace();
    }
} // Close play
} // Close class
```

根据上述推断内容，可作为推荐依据实现软件部分的设计，保证用户获取更加准确的推荐列表。

3 实验

采用协同过滤算法设计的自动化隐式评分音乐双重推荐系统，从该系统推荐结果准确性角度出发，设计了如下

对比实验。由于协同过滤算法需要上百个节点才能体现网络复杂程度，因此，采用基于 Java 的工具包模拟离散点来构建复杂网络结构。

3.1 实验环境

实验环境设置如表 2 所示。

表 2 实验参数设置

参量	数值
节点	10 个
CPU	2
核心频率	1.9GHz
内存	8GB

与计算固定数量节点不同，采用协同过滤算法需计算相邻节点最大值的限制，将距离一定范围内的所有节点都视为当前邻居，虽然邻居个数不确定，但是相似度不会相差较大，对于孤立节点处理也较为妥当。

3.2 实验结果与分析

传统系统采用显示评分方法，为用户进行音乐推荐；而双重推荐系统采用隐式评分方法，为用户进行音乐推荐。这两种推荐系统分别在噪声干扰和人为因素影响下推荐结果精准度进行对比，具体对比情况如下所示。

3.2.1 噪声干扰

传统系统对音乐数据进行提取时，受到噪声干扰影响，使数据不能被全部统计，导致最后提取到的数据为无效数据。而协同过滤系统设计隐式评分提取电路对音乐数据进行提取时，不会受到噪声干扰影响，使数据被全部统计，提取效果较好，对于推荐结果也较为准确。将这两种系统受到噪声干扰影响程度进行对比，结果如图 8 所示。

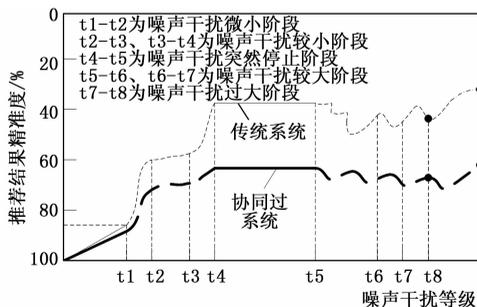


图 8 两种系统在噪声干扰影响下推荐精准度对比

由图 8 可知：当噪声干扰为 t1~t2 等级时，传统系统推荐结果精准度从 85% 下降到 60%，而协同过滤系统推荐结果精准度从 88% 下降到 70%；当噪声干扰为 t2~t4 等级时，传统系统推荐结果精准度从 60% 下降到 35%，而协同过滤系统推荐结果精准度从 70% 下降到 65%；当噪声干扰为 t4~t5 等级时，传统系统推荐结果精准度与协同过滤系统推荐结果无明显变化；当噪声干扰为 t5~t6 等级时，传统系统推荐结果精准度从 35% 升高到了 40%，而协同过滤系统推荐结果精准度从 65% 升高到了 67%；当噪声干扰为

t6~t7 等级时, 传统系统推荐结果精准度从 40% 升高到了 43%, 而协同过滤系统推荐结果精准度从 67% 升高到了 69%; 当噪声干扰为 t7~t8 等级时, 传统系统推荐结果精准度从 43% 下降到了 40%, 而协同过滤系统推荐结果精准度从 67% 下降到了 65%; 当噪声干扰大于 t8 等级时, 传统系统推荐结果精准度下降到了 30%, 而协同过滤系统推荐结果精准度下降到了 55%, 说明噪声干扰等级超过 t8 时, 两种系统都受到较大程度的干扰, 推荐结果精准度无法得到有效控制, 而在实际生活中, 超过 t8 等级的噪声会噪声机器设备的损坏, 为此, 可忽略超过该等级的噪声干扰影响。

根据上述对比内容可知: 当噪声干扰为 t1~t2 等级时, 两种系统推荐结果精准度都达到最高, 分别为 85% 和 88%, 且协同过滤系统在受到 t1~t8 等级噪声干扰影响, 推荐结果精准度始终不低于 60%。由此可知, 在噪声影响下, 采用协同过滤系统推荐结果精准度较高。

3.2.2 人为因素影响

采用传统系统会受到人为因素影响, 播放器中的音乐杂乱无章, 无法获取准确的推荐信息, 采用显示评分机制, 使推荐结果更加偏向后期用户听到的歌曲; 而采用协同过滤系统, 虽然播放器中的音乐播放种类较为混乱, 但是可通过隐式评分来获取准确推荐信息, 并采用双重推荐机制, 提高推荐结果精准度。将这两种系统受到人为因素影响进行对比, 结果如图 9 所示。

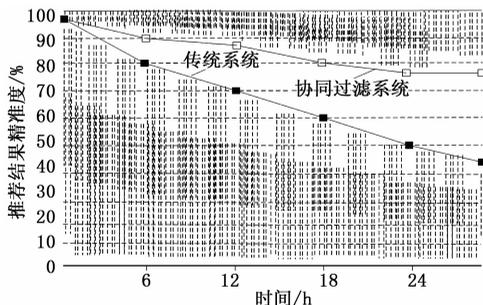


图 9 两种系统在人为因素影响下推荐精准度对比

由图 9 可知: 在人为因素影响下, 当时间为 6h 时, 采用传统系统推荐结果精准度为 80%, 而采用协同过滤系统推荐结果精准度为 90%; 当时间为 12 h 时, 采用传统系统推荐结果精准度为 70%, 而采用协同过滤系统推荐结果精准度为 88%; 当时间为 18 h 时, 采用传统系统推荐结果精准度为 60%, 而采用协同过滤系统推荐结果精准度为 80%; 当时间为 24 h 时, 采用传统系统推荐结果精准度为 50%, 而采用协同过滤系统推荐结果精准度为 78%。

根据上述对比内容可知: 当时间为 6h 时, 两种系统推荐结果精准度都达到最高, 分别为 80% 和 90%, 由此可知, 在人为因素影响下, 采用协同过滤系统推荐结果精准度较高。

3.3 实验结论

综合上述实验内容可得出结论: 采用协同过滤算法设计的自动化隐式评分音乐双重推荐系统是具有合理性的。①在噪声干扰影响下, 当干扰等级为 t1~t2 时, 传统系统与协同过滤系统推荐结果精准度都达到最高, 分别为 85% 和 88%; ②在人为因素影响下, 当时间为 6 h 时, 传统系统与协同过滤系统推荐结果精准度都达到最高, 分别为 80% 和 90%。在这两种因素干扰下, 采用协同过滤算法设计的系统推荐结果精准度都比传统系统精准度要高。

4 结束语

针对设计内容具有以下创新点: 以推荐系统内容为主, 结合普适环境, 通过收集当前用户偏好信息, 利用协同过滤处理方式预测用户感兴趣内容, 以此为基础设计推荐系统, 可有效避免噪声干扰和人为因素的影响, 保证较高的推荐结果精准度。

对于设计系统的研究展望内容有: 采用协同过滤算法设计的推荐系统是具有合理性的, 但是有关蓝牙交互可靠性还有待提高, 主要原因是蓝牙交互协议的不对称性, 导致异构通信模块工作效率较低。因此, 在以后研究过程中, 争取使用量化方法来研究蓝牙交互行为, 保障系统的实时性。

参考文献:

- [1] 朱 白. 数字图书馆推荐系统协同过滤算法改进及实证分析 [J]. 图书情报工作, 2017, 20 (9): 130-134.
- [2] 吴彦文, 齐 旻, 杨 锐. 一种基于改进型协同过滤算法的新闻推荐系统 [J]. 计算机工程与科学, 2017, 39 (6): 1179-1185.
- [3] 孙 萍, 李 镕, 关 欣, 等. 融合评分倾向度和双重预测的协同过滤推荐算法 [J]. 计算机科学与探索, 2017, 11 (10): 1642-1651.
- [4] 孙艳歌, 王志海, 黄 丹. 基于时间的局部低秩张量分解的协同过滤推荐算法 [J]. 计算机科学, 2017, 44 (7): 227-231.
- [5] 潘涛涛, 文 峰, 刘勤让. 基于矩阵填充和物品可预测性的协同过滤算法 [J]. 自动化学报, 2017, 43 (9): 1597-1606.
- [6] 许明杰, 蔚承建, 沈 航. Spark 并行化基于物品协同过滤算法 [J]. 计算机工程与设计, 2017, 38 (7): 1817-1822.
- [7] 李昆仑, 万品哲, 张德智. 基于改进用户相似性度量 and 评分预测的协同过滤推荐算法 [J]. 小型微型计算机系统, 2018, 18 (3): 17-22.
- [8] 熊聪聪, 邓 滢, 史艳翠, 等. 基于协同过滤的美食推荐算法 [J]. 计算机应用研究, 2017, 34 (7): 1985-1988.
- [9] 王建芳, 李 骁, 武文琪, 等. 一种 SVD 和信任因子相结合的协同过滤算法 [J]. 小型微型计算机系统, 2017, 38 (6): 1290-1293.
- [10] 钱晓捷, 张路一. 融合评分结构特征与偏好距离的协同过滤推荐算法 [J]. 计算机工程, 2017, 43 (5): 185-190.
- [11] 李 培. 基于 HMM 模型的网络社区推荐技术 [J]. 电子设计工程, 2018, 26 (3): 1-5.