Algorithm Design Document

1 Introduction

The algorithm is the core part of a recommendation system. It needs to analyze the data of the user's imported song list and the song list of the database to get a recommended song list or song.

Our recommendation algorithm mainly uses Collaborative Filtering(CF). CF is composed of 3 parts: User-based CF(UserCF), Item-based CF(ItemCF) and Latent Factor Model(LFM). UserCF and ItemCF are similar, and we will introduce them together.

2 User-based CF and Item-based CF

	Description	Input	Expected output
1.lst_dist(lst1, lst2)	calculate distance between 2 music lists	lst_dist(lst2music['伤感翻唱版集合'], lst2music['又是一个睡不着的夜晚'])	5.0
2.music_dist(music1, music2)	calculate the distance between 2 musics	music_dist('星河清梦', '缭星')	5.0
3.most_similar_lst(lst)	get the most similar music list	most_similar_lst(lst2music['刷题 看书 学习 工作 冥想'])	({'夏日喝汽水 lofi', '我去宇宙偷星星，放在夜里等你??', '浮生若梦', '银河赴约 - (网易云音乐助力高考自制曲目)', '你能来保護我的世界嘛', '等', '风的小径', '夜里失联', '这个深海是你的眼眸', '奇怪吗'}, 5.0)
4.most_similar_music(music)	get the most similar music	most_similar_music('无人之岛（翻自 任然）')	('好想好想（翻自 群星）', 5.0)

5.recommend(list)	recommend according to music list	recommend({'呼吸','无人之岛（翻自任然）', '水星记', '像鱼', '大鱼 - (动画电影《大鱼海棠》印象曲)', '千千阙歌(Live)', '心はいつもあなたのそばに Piano'})	{'永不失联的爱', '多远都要在一起','你的样子', '想い', '是想你的声音啊 - (你快听滴答滴)', '古诗中国', '心领神会', '你还好吗', '不爱我', '玻璃之情', '或是一首歌', '那个她', '秋海棠', '看见你的声音 - (电视剧《想见你》插曲)','好想好想（翻自 群星）'}

3 Latent Factor Model

	Description	Input	Expected output
1.getRatingMatrix(music_list)	music_list is a List composed of input songs	getRatingMatrix(music_list)	a rating matrix
2.matrixDecomposition(matrix, k)	decompose one matrix into two	matrixDecomposition(matrix, 50)	two matrix
3.recommend(music_list)	recommend according to music list	recommend({'呼吸','无人之岛（翻自任然）', '水星记', '像鱼', '大鱼 - (动画电影《大鱼海棠》印象曲)', '千千阙歌(Live)', '心はいつもあなたのそばに Piano'})	{'摸月光 - (独家国风自制企划《一期一会》)', '1107', 'Seven Days', 'My Apologies', '相信', 'Bloom (Bonus Track)', 'No Love', '?????? - (Siren)', '幻听（片段版）（翻自 许嵩）','拔刀出鞘'}