

基于 Vuetify 和 LSTM 的音乐论坛

冯驰、钟睿哲、刘梓池、张峰玮、王奥迪、何梓欣、姚思悦、韩雨虹


2022 年 5 月 20 日



上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

目录

- 
- The logo of Xinjiang University is a large, faint circular watermark in the background. It features a gear-like outer ring with the university's name in Chinese and English. In the center, there is an anvil with a hammer and a stack of books, with the year '1896' below it.
- 1 项目背景
 - 2 项目概述
 - 3 项目管理
 - 4 AI 音乐生成
 - 原理简述
 - 模型训练
 - 模型推理
 - 其他网络选择
 - 5 后端架构与 api 设计
 - 后端设计思路
 - 6 前端界面与功能
 - 前端设计思路
 - 用户身份认证
 - 创作中心-AI 生成音乐
 - 创作中心-合成音乐
 - 音乐论坛

目录

- 1 项目背景
- 2 项目概述
- 3 项目管理
- 4 AI 音乐生成
 - 原理简述
 - 模型训练
 - 模型推理
 - 其他网络选择
- 5 后端架构与 api 设计
 - 后端设计思路
- 6 前端界面与功能
 - 前端设计思路
 - 用户身份认证
 - 创作中心-AI 生成音乐
 - 创作中心-合成音乐
 - 音乐论坛



项目背景

疫情期间，无聊吗 ???

那么来享受音乐合成的快乐、发掘 AI 续写音乐之美吧。

目录

- 1 项目背景
- 2 项目概述
- 3 项目管理
- 4 AI 音乐生成
 - 原理简述
 - 模型训练
 - 模型推理
 - 其他网络选择
- 5 后端架构与 api 设计
 - 后端设计思路
- 6 前端界面与功能
 - 前端设计思路
 - 用户身份认证
 - 创作中心-AI 生成音乐
 - 创作中心-合成音乐
 - 音乐论坛



项目概述与技术分工

音乐生成论坛项目主要实现了以下功能：

- 在线合成音乐
- AI 生成音乐
- 音乐分享论坛

项目分工如下：

- AI 生成音乐：钟睿哲
- 后端服务器架构：刘梓池
- 网站前端搭建：冯驰、张峰玮、王奥迪、何梓欣、姚思悦、韩雨虹

目录

- 1 项目背景
- 2 项目概述
- 3 **项目管理**
- 4 AI 音乐生成
 - 原理简述
 - 模型训练
 - 模型推理
 - 其他网络选择
- 5 后端架构与 api 设计
 - 后端设计思路
- 6 前端界面与功能
 - 前端设计思路
 - 用户身份认证
 - 创作中心-AI 生成音乐
 - 创作中心-合成音乐
 - 音乐论坛



项目管理

- 

Danielle 4月20日
...Ql07Dbyy0 #腾讯会议: 533-18...
- 

Danielle 4月20日
3点腾讯会议
- 

Danielle 4月18日
...rlsBn4Mioa #腾讯会议: 312-72...
- 

Danielle 4月16日
...Q8RMj1PN #腾讯会议: 469-36...
- 

Danielle 4月14日
...uPYIz7d8xf #腾讯会议: 840-47...



页面交互

刘梓池: okk

5月11日



钟睿哲、Danielle、刘梓池

语音通话已经结束

周一

项目管理

我们主要使用 github 进行版本管理：

daniellePie / cooperation Public

<> Code Issues Pull requests 1 Actions Projects Wiki Security Insights Settings

main 6 branches 0 tags Go to file Add file Code

File/Folder	Description	Commit Date	Commits
Auxiliary_inf	刘梓池完成音乐相关一部分服务--对应前端王奥迪	12 days ago	56
ffmpeg_instruction	刘梓池 提交前后端对应说明文档	17 days ago	
front_back_interfacing	add new interfacing	4 days ago	
music_generation	update training and inference code for music generation	5 days ago	
musicforum	add new interfacing	4 days ago	
musicforum_back	刘梓池 后端完成用户setting服务---对应前端何梓欣	10 days ago	
picture	刘梓池 搭建图片服务器, 完成头像服务, friends表相关基本服务	27 days ago	
record	upload 音乐生成器调研.md	last month	
.DS_Store	delete	23 days ago	
LowFidelityModelofFrontEnd.pdf	update the Low fidelity model of front end	28 days ago	
README.md	刘梓池 提交服务器初步部分说明	last month	

README.md

Music generation Forum—cooperation project

About
No description, website, or topics provided.
2 stars
1 watching
0 forks

Releases
No releases published
[Create a new release](#)

Packages
No packages published
[Publish your first package](#)

Contributors 9

项目管理

main ▾ cooperation / record /

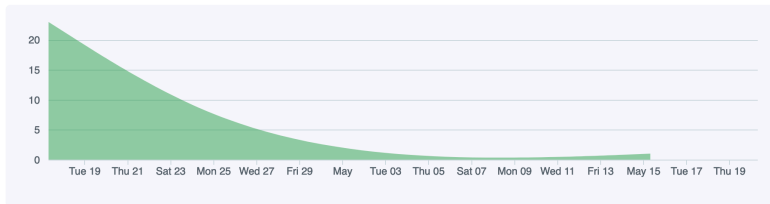
Go to file Add file ▾ ...

FengweiZhang upload 音乐生成器调研.md	9b5b738 on 18 Apr	History
..		
4.16fc_summary.md	update record	last month
4.18talk.md	update record	last month
text_to_image.md	text2image and music generation	last month
音乐生成器调研.md	upload 音乐生成器调研.md	last month


Apr 17, 2022 – May 20, 2022

Contributions: Commits ▾

Contributions to main, excluding merge commits and bot accounts



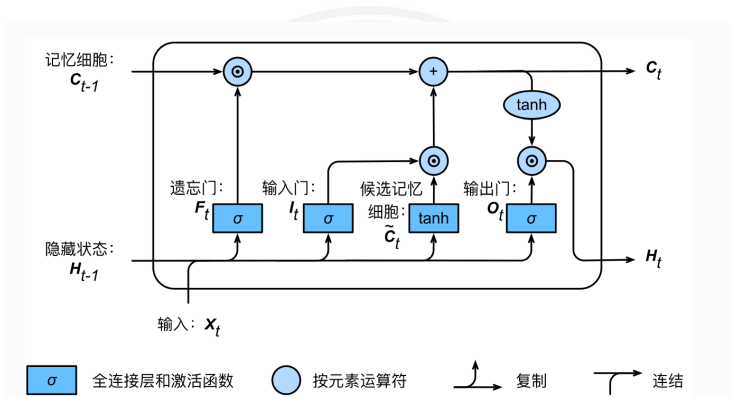
目录

- 
- 1 项目背景
 - 2 项目概述
 - 3 项目管理
 - 4 AI 音乐生成
 - 原理简述
 - 模型训练
 - 模型推理
 - 其他网络选择
 - 5 后端架构与 api 设计
 - 后端设计思路
 - 6 前端界面与功能
 - 前端设计思路
 - 用户身份认证
 - 创作中心-AI 生成音乐
 - 创作中心-合成音乐
 - 音乐论坛

基本概念

- 选择时序网络进行音乐的生成
- 通过 user 指定/上传音乐初始片段 (Begin Of Sequence, BOS), 根据这段音频进行后续的音频生成
- 时间预测角度, 将这段音频视为已知的观测值, 进行音频的预测
- 条件生成角度, 将这段音频视为 condition, 基于该 condition 进行后续的音乐生成

网络选择

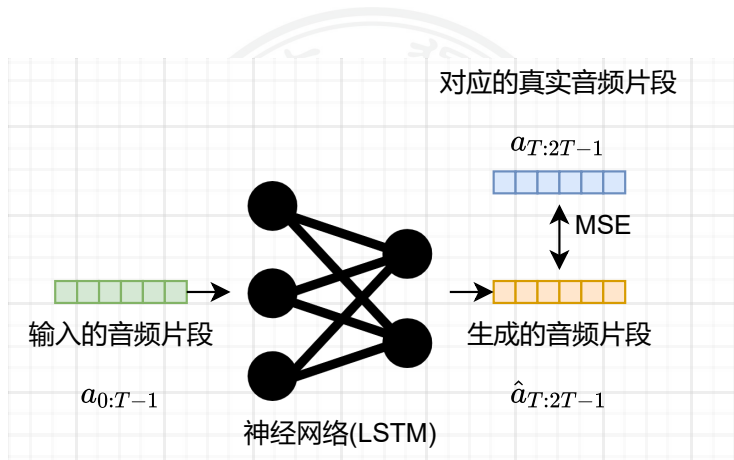


我们选择 LSTM(Long Short Term Memory) 作为我们的音乐生成网络，引入了遗忘门、输入门和输出门来更好地解决传统 RNN 的梯度衰减和遗忘的问题

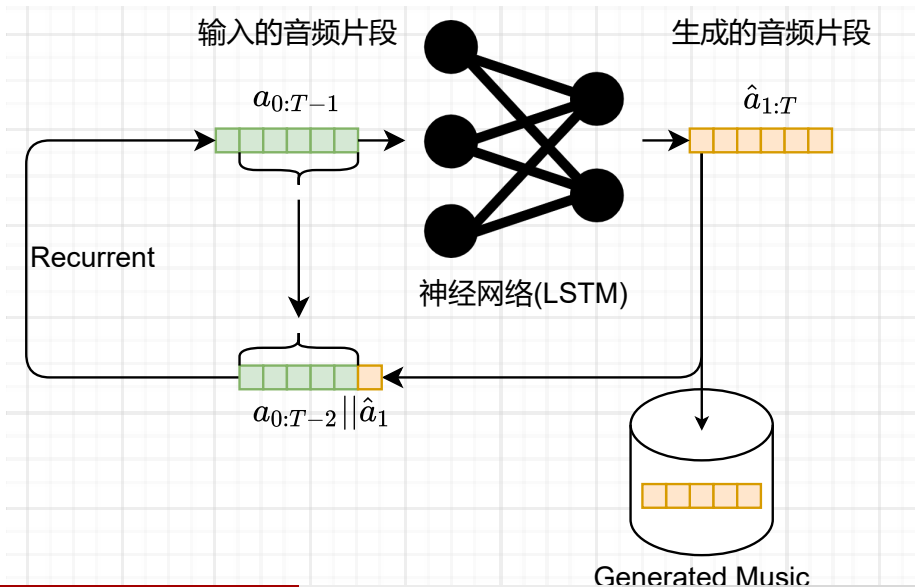
音频处理

- 为降低训练开销和网络容量，对音频进行降采样
- 选定滑窗，对音频进行 FFT
- 对网络的输出进行 IFFT，得到生成的可播放音频文件

模型训练




模型推理



其他网络选择

- 使用 LSTM 相当于将音乐生成视为预测任务进行处理
- 通过增大数据集，仍然存在过拟合的问题，加入 Dropout 之后依然难以缓解
- 将尝试使用 cGAN(condition Generative Adversarial Network)，将给定的音乐片段视为 condition，用于生成新的音乐。通过 Generator 进行音乐生成，Discriminator 进行音乐质量判别，实现音乐的生成

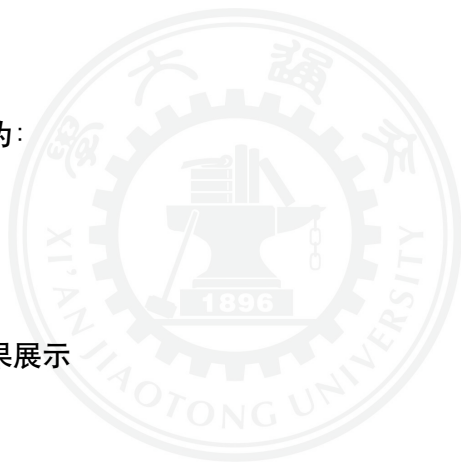
目录

- 
- The logo of Xinjiang University is a large, faint circular watermark in the background. It features a gear-like outer ring with the university's name in Chinese '新疆大学' at the top and 'XINJIANG UNIVERSITY' at the bottom. In the center is an anvil with a hammer, and the year '1896' is inscribed on a base below it.
- 1 项目背景
 - 2 项目概述
 - 3 项目管理
 - 4 AI 音乐生成
 - 原理简述
 - 模型训练
 - 模型推理
 - 其他网络选择
 - 5 后端架构与 api 设计
 - 后端设计思路
 - 6 前端界面与功能
 - 前端设计思路
 - 用户身份认证
 - 创作中心-AI 生成音乐
 - 创作中心-合成音乐
 - 音乐论坛

后端设计思路

后端工作主要分为：

- 数据库
- 服务器
- 后端架构
- 音频处理
- 部分查询结果展示



数据库

1. 根据需求设计数据库表和字段，形成并更新了数据库设计文件
2. 使用阿里云 mysql 服务器进行数据存储，数据表包括用户表，好友表，音频表，喜爱与收藏表
3. 数据库尽量保证业务的拓展，比如支持了多层评论等等
4. Sqlyong 展示

服务器使用

1. 使用了三台服务器，除了 mysql 服务器外还有一台网站服务器，一台存储图片和音频的文件服务器用于减轻网站服务器的传输负担
2. 有时间有必要的话可以添加 redis 缓存来缓存常用数据和用户登录信息，nginx 网关，服务拆分为微服务 + docker，mysql 主从复制等等

后端架构

1. 后端使用 springboot+mybatisplus 架构

2. 由于开始我对 springcloud+ 微服务并不熟练并且网站规模不大，因此采用了 springboot 的架构，使用 mybatisplus 与 mysql 数据库交互，使用 maven 包管理

3. 分多层，controllor 层匹配 url 分配 service 层获取服务，service 层调用 mybatis-mapper 层操作数据库，另外还是数据库接口的 dao 层，vo 层进行前后端数据交互（交互文档），还有配置包过滤包（用户登录过滤）等等，见代码

音频处理

1.music/musicInfo/infoForGenerate 查询所有音频信息

2.music/musicInfo/detail/1 查询 id 为 1 的音频信息


3.music/musicInfo/myUpload/11 查询用户 id 为 11 上传的所有音频 id

4. 各种服务还有待与前端调试，包括过滤器异常包的设置，还有对各种异常查询的处理等等，由于这是我本身的技术栈因此后端进度很快但由于我个人懒和时间原因懒得自己写 form 测试或者用工具测试 post，需等待与前端调试

部分查询结果展示

1. 在服务器安装了 mmfpeg, 使用 ffmpeg 进行处理
2. 包括获取音频时间, 生成空音频, 对一个 list 的音频添加空音频做 concat, 对多个音轨音频做 mix
3. Ffmpeg 代码由钟睿哲同学与我合作完成, 由我整合进后端代码
4. Java 代码中使用进程类单独开一个进程执行 ffmpeg, 其中时间还做了输出重定向之后 + 正则匹配, 见代码

目录

- 
- The logo of Xinjiang University is a large, faint watermark in the background. It is a circular emblem with a gear-like outer ring. Inside the ring, the text 'XINJIANG UNIVERSITY' is written in English, and '新疆大学' is written in Chinese. In the center of the logo is an anvil with a hammer resting on it, and the year '1896' is inscribed below the anvil.
- 1 项目背景
 - 2 项目概述
 - 3 项目管理
 - 4 AI 音乐生成
 - 原理简述
 - 模型训练
 - 模型推理
 - 其他网络选择
 - 5 后端架构与 api 设计
 - 后端设计思路
 - 6 前端界面与功能
 - 前端设计思路
 - 用户身份认证
 - 创作中心-AI 生成音乐
 - 创作中心-合成音乐
 - 音乐论坛

前端设计思路

前端工作主要分为三大部分：

1. 用户身份认证

2. 创作中心

- AI 生成音乐
- 合成音乐

3. 音乐论坛

用户身份认证

Music Forum

[登录](#)[注册](#)

用户昵称 0 / 20

请输入密码 0

请重复输入密码 0

邮箱 [获取验证码](#)

请输入验证码

[注册](#)

Music Forum

[登录](#)[注册](#)

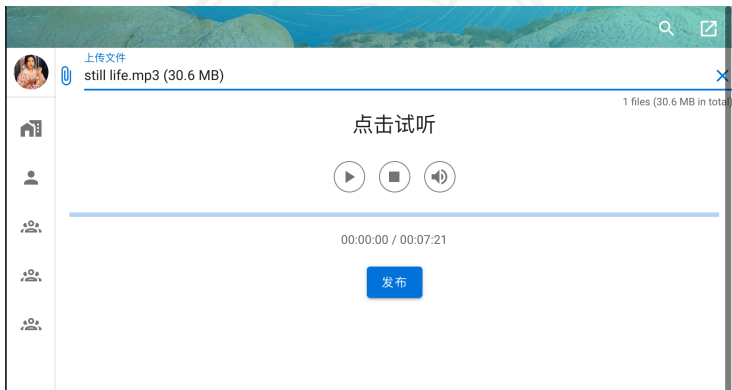
用户昵称 0 / 20

请输入密码 0

[登录](#)[忘记密码](#)

创作中心-AI 生成音乐界面

上传音乐-> 得到 AI 续写的音乐-> 试听



创作中心-AI 生成音乐界面



创作中心-合成音乐功能

在创作中心，用户可以使用后端数据库中存储的音乐片段合成新的音乐，是用户发挥自身创意并上传作品的页面。

创作中心-合成音乐主要实现的功能如下：

1. 使用现有的短音乐片段合成长音乐片段
2. 上传并保存用户生成的长音乐片段

创作中心-合成音乐 demo 界面

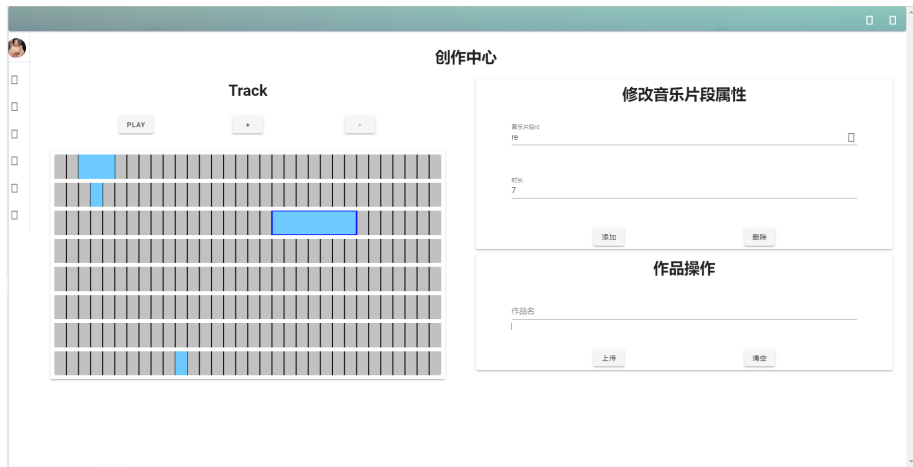


Figure: 创作中心界面 (demo)

音乐论坛实现功能

在为音乐论坛设计功能时，我们参考了新浪微博、网易云等较为成熟的应用，选取了核心社交功能进行设计。

在至今已完成的迭代过程中，我们已经完成了热门音频、个人主页、关注列表等主要页面的设计。在音乐卡片中实现了点赞、收藏、评论功能。

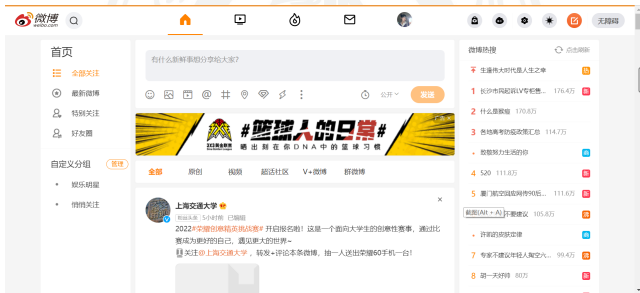


Figure: 微博主页

热门音频

热门音频界面由热门的音频卡片组成，按照点赞数从高到低依次排列

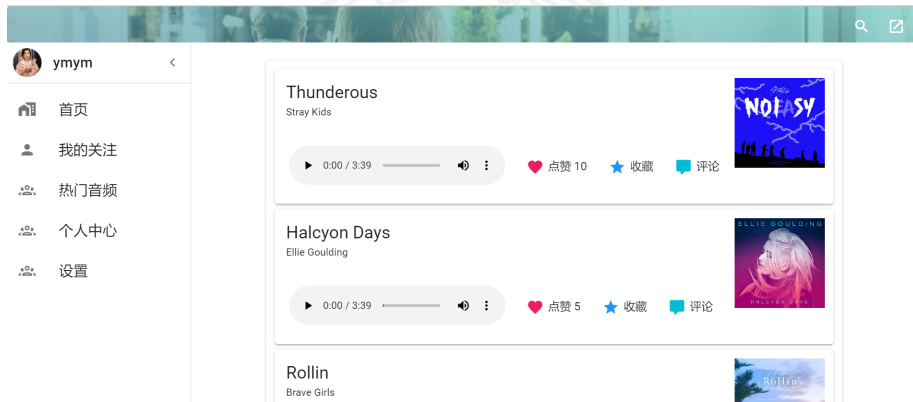


Figure: 热门音频界面

音乐卡片

由于 Vue 组件化的特点，音频卡片作为单独的子组件不仅用在热门音频，还有个人中心、首页这些页面中，降低了代码的冗余度。

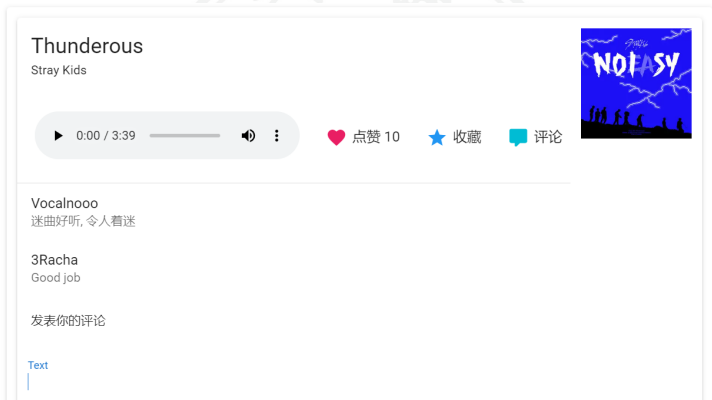


Figure: 音乐卡片

个人设置

个人设置界面分为账号信息、个人信息和权限设置三部分，可以通过填写或选择来更改用户的信息。

Figure: 个人设置界面

个人设置

在对前端页面进行设计的过程中加入了对用户名称、个人简介字数的限制和对邮箱、生日等格式的限制，使数据库存储形式更加规范；权限设置部分与个人主页部分进行交互，使得用户可以对自己个人主页显示的内容进行管理。

用户昵称
1234567890-

11 / 10

2022

May 2022

S	M	T	W	T	F	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

☒ 允许所有人观看我的作品

☐ 仅允许好友观看我的作品

☐ 仅允许我观看我的作品

Figure: 个人设置页面

个人中心

个人中心界面直接使用音频卡片子组件，从而显示“我”发布的音频。同时，从该界面还可以跳转“我的关注”、“我的点赞”、“我的评论”等界面。

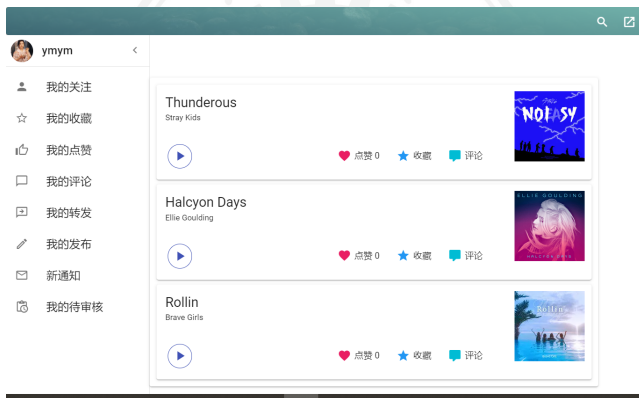


Figure: 个人中心页面

关注列表

我的关注使用列表组件显示关注的用户。列表组件易于在集合中识别特定项目，为组织一组文本和图像提供了一致的样式。

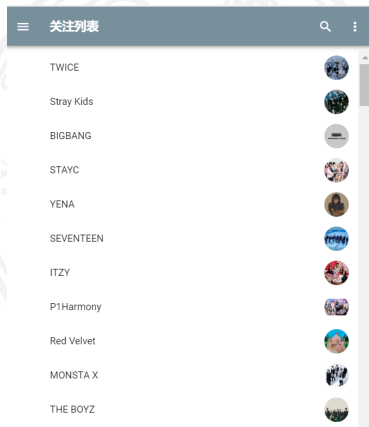


Figure: 关注列表



Thank You

