**软件设计文档**

# 1.技术选型

系统前端采用html、css、javascript以及bootstrap框架等，服务器采用python的tornado框架，数据库使用mysql。这些都是组员熟悉的技术。

爬虫使用python，并利用beautifulsoup解析html。基于html的解析，健壮性更好。

推荐算法使用FP-Growth。这是基于关联规则的推荐。由于爬虫只获取了豆列（类似购物单），不适合使用协同过滤等方法。

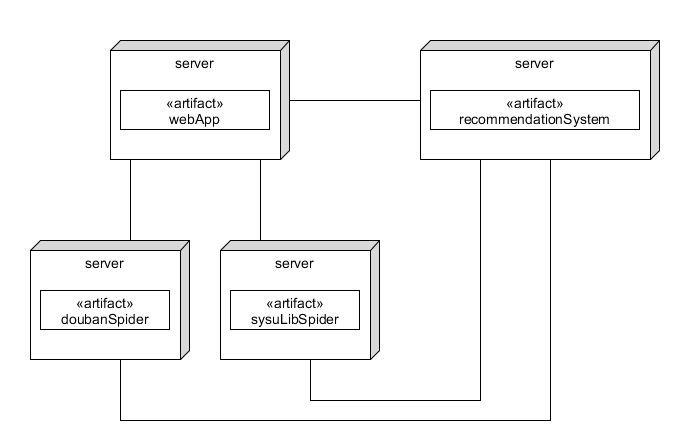
# 2.架构设计

系统由4部分组成，分别为web应用、豆瓣爬虫、图书馆爬虫、推荐算法组成。

爬虫为推荐系统提供训练数据，为web应用提供图书基本信息。

推荐系统为web应用提供推荐书籍列表。

Web应用提供访问服务。



# 3.模块划分

## 3.1服务器

## 3.2数据库

## 3.3前端

## 3.4豆瓣爬虫

## 3.5图书馆爬虫