

Chicker

A Personalized Restaurant Recommender

Yongfeng Zhang Tsinghua University
zhangyf07@gmail.com <http://yongfeng.me>

Outline

- Chicker : WHAT ?
- Chicker : WHY ?
 - Demo
- Chicker : HOW ?
 - Crawler & Data & Indexing
 - Recommendation Strategies
 - Client & User interaction

Chicker: WHAT ?

- A **Personalized** Restaurant Recommender
 - Intelligently recommend restaurants according to your tastes.
 - Different persons get different recommendations.
 - Recommend according to crowd scores, your history preferences as well as current states.
 - Location based service.

Chicker: WHY?

- Not My Taste
- Sham Grade by Zombies
- Which is Nearby



第一次使用？

您可以立刻通过一个简单的步骤表达自己的个性化需求，从而得到更加准确的搜索结果

个性化

也可以暂时跳过，在使用过程中逐渐积累个性化信息

跳过

后退

Chicker

北大农园餐厅

家常菜，¥10，北京，距离：800

地址：海淀区颐和园路5号北京大学南门（中关村大街）

推荐理由：颐和园的家常菜餐馆中排名第1，喜欢北大博雅国际会议中心西餐厅的人也喜欢该餐馆

您的评价：很好



想去

没去过

Chicker



1 公里内

不限价格

您的评价：很好

北大农园餐厅

家常菜，¥10，北京，距离：800

推荐理由：颐和园的家常菜餐馆中排名第1，喜欢北大博雅国际会议中心西餐厅的人也喜欢该餐馆

可能喜欢：90%

黔之味

贵州菜，¥45，北京，距离：940

推荐理由：喜欢清华大学甲所餐厅的人也喜欢该餐馆

可能喜欢：87%

北京大学燕南美食

家常菜，¥6，北京，距离：700

推荐理由：喜欢北大农园餐厅的人也喜欢该餐馆

可能喜欢：87%

畅春园餐厅

家常菜，¥16，北京，距离：1000

推荐理由：颐和园的家常菜餐馆中排名第3，喜欢北大农园餐厅的人也喜欢该餐馆

您的评价：一般

竹鱼坊(北大总店)

烤鱼，¥58，北京，距离：170

推荐理由：北京很受欢迎的烤鱼餐馆，喜欢北大农园餐厅的人也喜欢该餐馆

当前位置: Qinghua West Rd, Haidian, Beijing, China

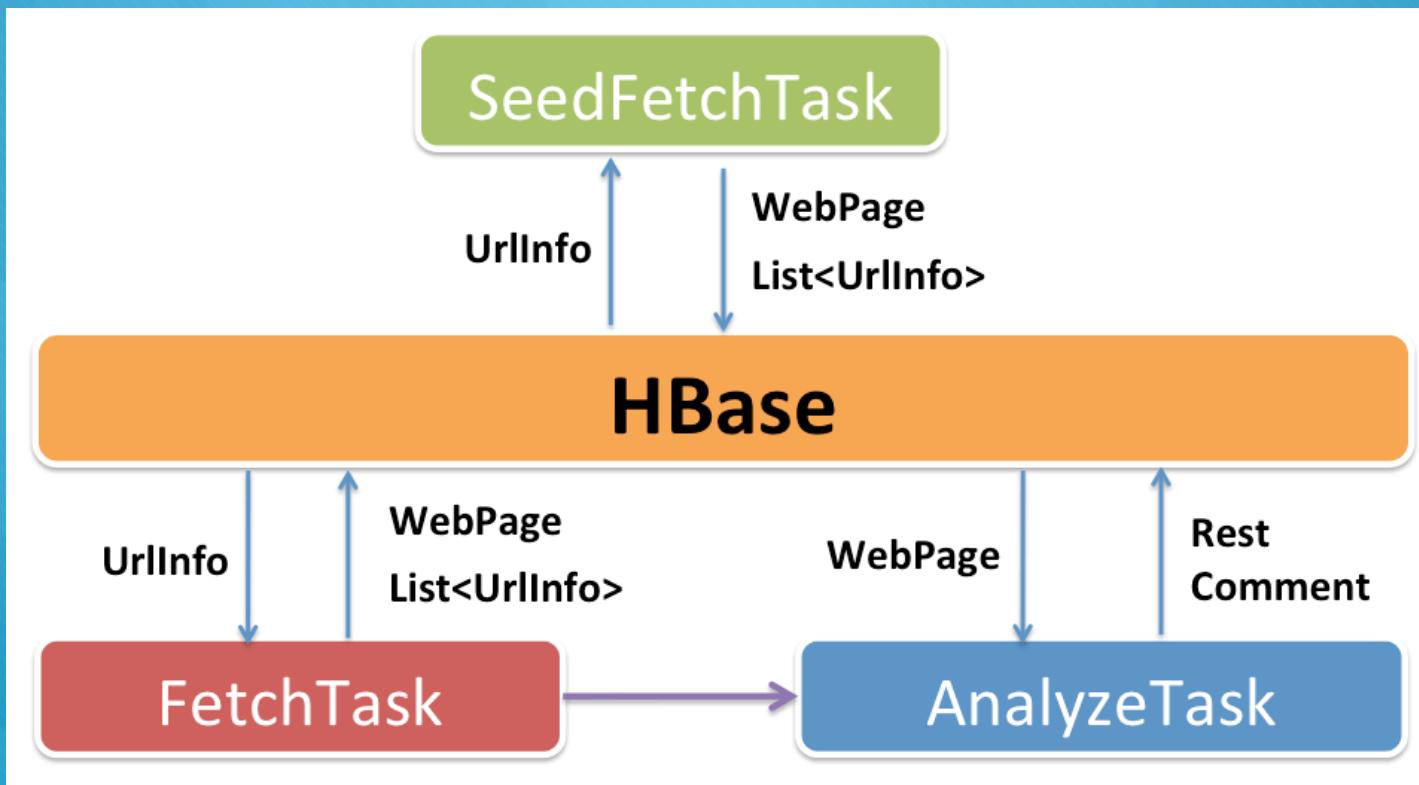
Demo

Chicker: HOW ?

Crawler & Data & Indexing

- Crawler
 - Distributed crawler based on Hadoop MapReduce
 - Storage by Hadoop hbase distributed file system
 - Fetch document by Httpclient
 - Parse webpages by Nekohtml + xpath

Crawler



Data

○ Restaurant information & User reviews

北京餐厅 » 海淀区 » 五道口 » 北京菜 » 烤鸭 » 全聚德烤鸭店(清华园店)

全聚德烤鸭店(清华园店) 其他25家分店 ▾

★★★★★ 1251封点评 | 人均 ¥128

口味: 25
环境: 25
服务: 22

地址: 海淀区中关村东路1号院清华科技园科技大厦A座1楼(清华大学南门东) 公交/驾车

电话: 010-82150018 82151015

标签: 烤鸭 婚宴酒店 五道口

修改/报错

 小雷弹 ★★★

★★★★★ 口味: 3(很好) 环境: 3(很好) 服务: 2(好) 人均: ¥100

最喜欢吃的居然是炸鲜奶! 应该算是我吃过的最好吃的炸鲜奶了
烤鸭我觉得哪家吃着都一样, 因为都是裹成一卷就这样吞下去了, 到底脆不脆, 酥不酥, 真的没法评论
火燎鸭心不错, 有人说这是用茅台做的, 哈哈, 我真的没吃出酒味来, 但是还是很好吃的
鸭汤很浓很鲜
不好吃的有木瓜牛腩, 蒜香排骨, 前者味道怪怪的, 后者又炸得过焦, 吃着费力

推荐菜: 火燎鸭心 烤鸭 炸鲜奶 鸭汤

等位时间: 15-30分钟

12-05-25 15:27 全聚德烤鸭店 | 详情 | 送鲜花 | 收藏 | 不当内容

Data

推荐菜

烤鸭(743) 芥末鸭掌(278) 盐水鸭肝(251) 精品烤鸭(197) 鸭汤(146) 火燎鸭心(146) 榴莲酥(99) 大拌菜(79) 炸鲜奶(61)
干锅鸭杂(54) 红烧肉墨鱼(28) 豌豆黄(24)

详细信息

商户描述: “驰名中外”的京城“饮食”名片, 提到烤鸭, 必连着他家的名号。挂炉烤鸭确实做得“不错”, “入口即化”的“酥脆”外皮, “肥而不腻”的鸭肉, 搭配“薄嫩”的春饼, “香甜”的面酱, ... [更多](#)

餐厅氛围: 商务宴请 朋友聚餐 家庭聚会 [更多](#)

餐厅特色: 可以刷卡 是老字号 [更多](#)

营业时间: 周一 至 周日 11:30-14:00 17:30-21:00 [修改](#)

公交信息: 307路、319路、331路、355路、375路、438路、562路、628路、656路、731路、... [更多](#) [修改](#)

贡献榜: 蓝梅奶... 添加商户 / 丁旨 发布第一个点评 / 系统在10-12-07最后更新 | [查看贡献榜»](#)

Data

- Location (longitude & latitude) of restaurants
- Cracked location encryption code of Dianping



HESRHBZVVIJIGK



Latitude: 39.91472
Longitude: 116.41302

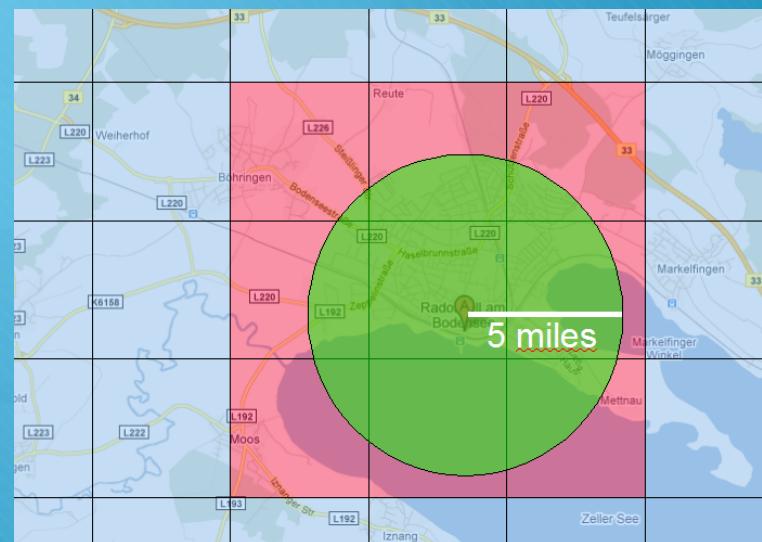
Data

- We selected users who rated 20 times or more.
- Statistics of crawled restaurants

# Users	# Restaurant	# User Reviews
~ 60w	~ 30w	~ 200w

Indexing

- Spatial indexing by Lucene
 - *Given center and radius, retrieve all the restaurants located in the circle.*

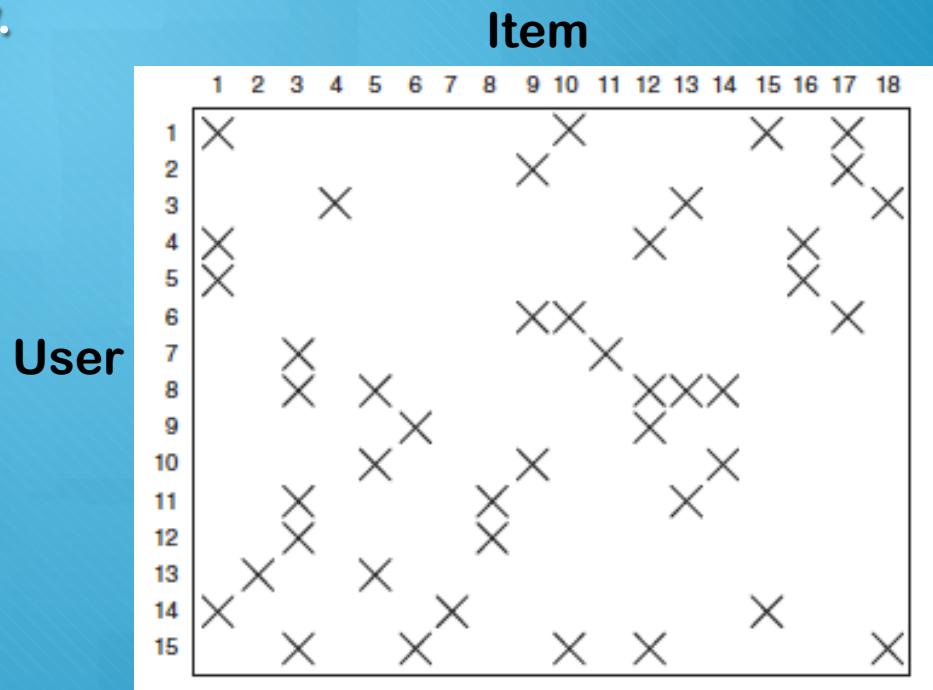


Chicker: HOW ?

Recommendation Strategies

- Data that we have by now.

- User-Item rating matrix
- Restaurant meta data
 - Name
 - Description
 - Special dishes
 - Special services
 - Location
 - ...

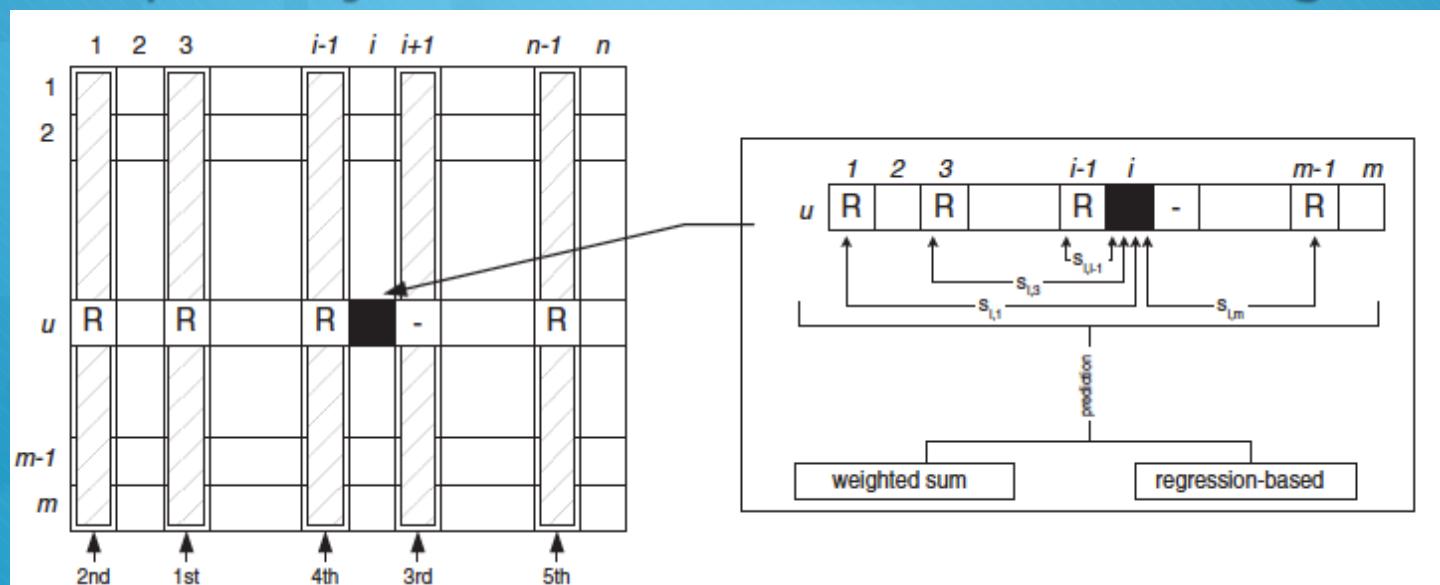


Preliminary definitions

- Prediction
 - Dynamic similarity
 - Static similarity (text similarity)
- Correlation
- Hotness (per month)
- Geographical similarity

Preliminary definitions

- Dynamic similarity
- Inspired by item-based collaborative filtering



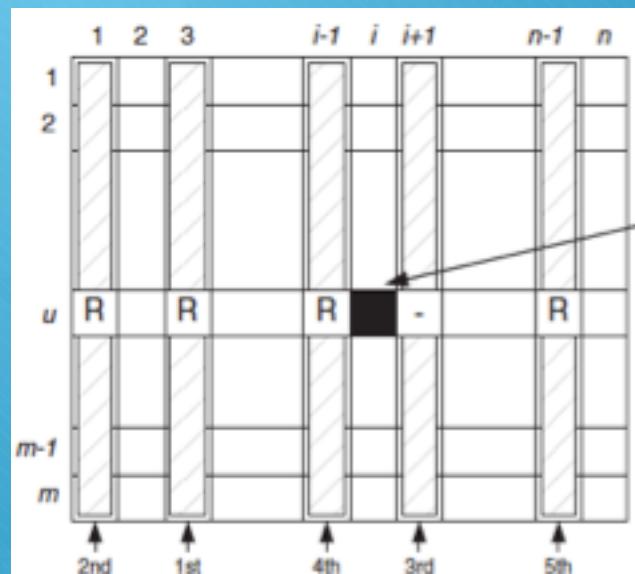
Preliminary definitions

- Static similarity
- text similarity
 - Restaurant descriptions + Special dishes + Special services
 - Cosine similarity

Preliminary definitions

- Correlation
 - Jaccard similarity of two restaurants
 - Average Correlation
- Hotness
- Number of ratings per month

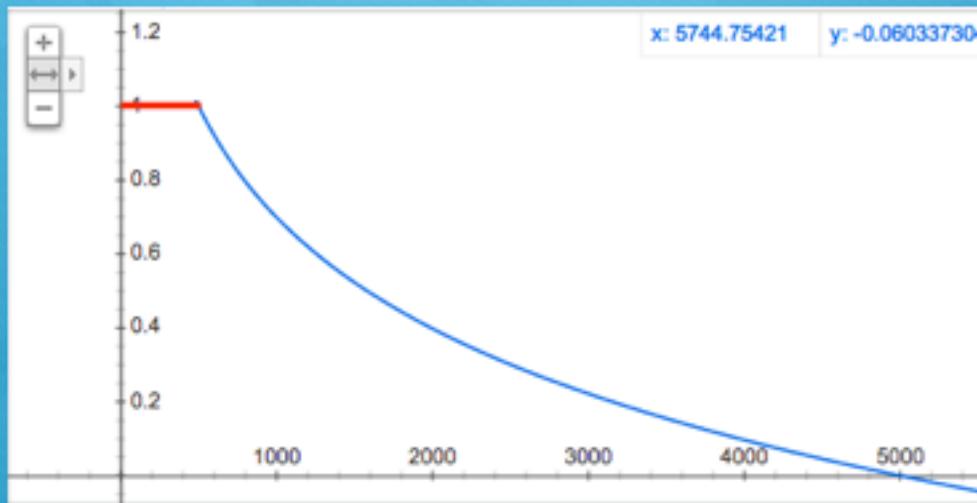
$$Jaccard = \frac{A \cap B}{A \cup B}$$



Calculation

- Geographical similarity
- $\text{MIN_DIST} = 500$ $\text{MAX_DIST} = \text{user specified}$

$$GeoSim = 1.0 - \frac{\ln(1 + dist) - \ln(1 + MIN_DIST)}{\ln(1 + MAX_DIST) - \ln(1 + MIN_DIST)}$$



Calculation

- Prediction
- Utilize both dynamic and static similarity
- Weighted sum predictor

$$R_D = \frac{Ds(I_1, I_x) \cdot R_1 + Ds(I_2, I_x) \cdot R_2 + \dots + Ds(I_t, I_x) \cdot R_t}{Ds(I_1, I_x) + Ds(I_2, I_x) + \dots + Ds(I_t, I_x)}$$

$$R_S = \frac{Ss(I_1, I_x) \cdot R_1 + Ss(I_2, I_x) \cdot R_2 + \dots + Ss(I_t, I_x) \cdot R_t}{Ss(I_1, I_x) + Ss(I_2, I_x) + \dots + Ss(I_t, I_x)}$$

$$R = \frac{R_D + R_S}{2}$$

Calculation

- Prediction
 - For those restaurant that can not be predicted:
 $R_D \text{ or } R_S = \text{Average Rating of the Item}$
 - This happens when:
 - The user didn't make any rating
 - All the similarities are 0

Calculation

- How to do RANKING ?
- Normalization
 - Normalize Prediction, GeoSim, Average Correlation, Hotness to range 0~5
 - Prediction : no need to normalize
 - GeoSim, Average Correlation: ×5.0
 - Hotness

$$H' = 5.0 \cdot \frac{H - H_{min}}{H_{max} - H_{min}}$$

Calculation

- How to do RANKING ?
 - Calculate weighted rank

$$Rank = 0.4 \cdot P + 0.3 \cdot G + 0.2 \cdot C + 0.1 \cdot H$$

- Rank inversely

Calculation

- How to COLD START ?
- Will not consider prediction and average correlation
- Aim to find restaurants user may have visited

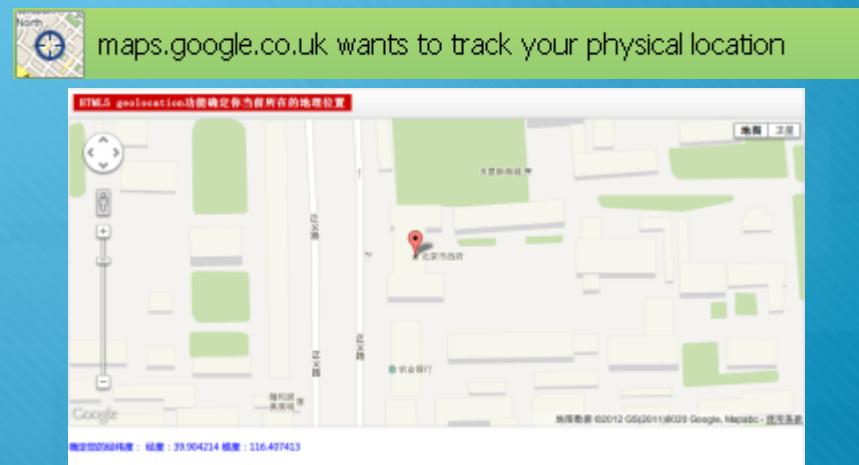
$$Rank = \frac{G + H}{2}$$

Client & User Interaction

- HTML5
- Available on PC, mac, tablet, and smart phone.
- Perfectly function with Chrome, Safari, Maxthon, Android WB, WB for S60

HTML5 Geolocation

- LBS Spirit
- navigator.geolocation.getCurrentPosition



User Identity

- Hash session info
- Cookie
- Initialization Score

Recommendation Reasons

- Post explaining by 3 kinds of reasons
 - [X位置]的[X类]餐馆中排名第[X]
 - 城市 > 行政区 > 商区 (每类/每区域按**Hotness**排序存好)
 - 喜欢[X餐厅]的人也喜欢该餐厅
 - **Correlation >= 0.5**
 - [X位置]很受欢迎的[X类]餐厅
 - 商区 > 行政区 > 城市 (**Hotness>=100 & Average Rate >= 4**)

Third-party libs

- Phonegap
- Sencha Touch
- Ext



Future Work

- More detailed recommend reasons
 - price
- Custom Query
 - Location, breakfast
- Connect to social network
 - Get and push recommendations from/to friends

第一次使用?

您可以立刻通过一个简单的步骤表达自己的个性化需求,从而得到更加准确的搜索结果

个性化

也可以暂时跳过,在使用过程中逐渐积累个性化信息

跳过

后退

Chicker

北大农园餐厅

家常菜, ¥10, 北京, 距离: 800

地址: 海淀区颐和园路5号北京大学南门(中关村)

推荐理由: 颐和园的家常菜餐馆中排名第1, 喜欢北大
博雅国际会议中心西餐厅的人也喜欢该餐馆

您的评价: 很好



想去

没去过



Chicker



1 公里内

不限价格

您的评价: 很好

北大农园餐厅

家常菜, ¥10, 北京, 距离: 800

推荐理由: 颐和园的家常菜餐馆中排名第1, 喜欢北大博
雅国际会议中心西餐厅的人也喜欢该餐馆

可能喜欢: 90%

黔之味

贵州菜, ¥45, 北京, 距离: 940

推荐理由: 喜欢清华大学甲所餐厅的人也喜欢该餐馆

可能喜欢: 87%

北京大学燕南美食

家常菜, ¥6, 北京, 距离: 700

推荐理由: 喜欢北大农园餐厅的人也喜欢该餐馆

可能喜欢: 87%

畅春园餐厅

家常菜, ¥16, 北京, 距离: 1000

推荐理由: 颐和园的家常菜餐馆中排名第3, 喜欢北大农
园餐厅的人也喜欢该餐馆

竹鱼坊(北大总店)

烤鱼, ¥58, 北京, 距离: 170

推荐理由: 北京很受欢迎的烤鱼餐馆, 喜欢北大农园餐
厅的人也喜欢该餐馆

您的评价: 一般

当前位置: Qinghua West Rd, Haidian, Beijing, China

Q & A

Related Works

Ness



食神摇摇





Recommender vs Searcher

- Information \leftrightarrow Human

HTML5 vs App

- Develop Cost
- Cross platforms