

**北 京 科 技 大 学**

本科生毕业设计(论文)选题报告

基于用户地理位置和偏好的

题　　目： \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

移动推荐系统的设计与实现

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

计算机与通信工程学院

学　　院： \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

信息安全

专　　业： \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

芦珊

姓　　名： \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

41259046

学　　号： \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

指导教师签字： \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

2016 年 03 月 27 日

目 录

[1 课题的研究背景及意义 1](#_Toc446872798)

[1.1 课题背景 1](#_Toc446872799)

[1.2 研究意义 2](#_Toc446872800)

[2 文献综述 4](#_Toc446872801)

[2.1 Skyline查询的经典算法 4](#_Toc446872802)

[2.1.1 经典算法 4](#_Toc446872803)

[2.1.2 研究中的问题 6](#_Toc446872804)

[2.2 基于位置的Skyline查询研究 6](#_Toc446872805)

[2.2.1 LBS服务的概述 7](#_Toc446872806)

[2.2.2 研究中的问题 8](#_Toc446872807)

[3 研究内容、预期目标及研究方法 9](#_Toc446872808)

[3.1 研究内容 9](#_Toc446872809)

[3.1.1 Skyline查询相关定义 9](#_Toc446872810)

[3.1.2 算法流程 9](#_Toc446872811)

[3.1.3 算法示例 9](#_Toc446872812)

[3.2 预期目标 9](#_Toc446872813)

[3.3 研究方法 10](#_Toc446872814)

[3.3.1 系统平台开发与工具介绍 10](#_Toc446872815)

[3.3.2 系统功能模块设计 11](#_Toc446872816)

[3.3.3 系统数据的来源 11](#_Toc446872817)

[4 研究进度安排 12](#_Toc446872818)

[参考文献 13](#_Toc446872819)

1. 课题的研究背景及意义
   1. 课题背景

随着移动设备技术、通讯技术和GPS系统的迅速发展，人们收集数据的方法和获取的数据更加丰富，并且可以把这些丰富的信息提供给那些位置和环境都不断变化的用户。如何在海量的数据中找到合适的信息推荐给移动用户，帮助人们做出有效的决策，是数据库以及信息查询领域的重要课题之一。

目前，移动设备的存储空间和计算能力都非常可观，因此我们很容易在移动设备端嵌入一个小型的数据库，用于给用户提供有用的信息和服务。我们可以从数据库中选出基于用户情况和喜好的相应合适的信息做一个推荐。虽然可处理的数据量在增加，我们依然很难使用移动设备去浏览和查询大量的数据。这是因为移动设备的显示尺寸有限，用户移动时缺乏时间在大量的数据中查询。所以，仔细筛选的信息往往是有利的更容易被用户理解和接收的，挑选一定数量的项目推荐给用户变得越来越重要。

信息推荐是在计算机科学研究领域的一个重要的课题，它包括网站设计，电子商务和数字图书馆等。给移动设备用户做信息推荐一个重要的子课题，特别是在基于地理位置的服务中。一种最常见方法的就是只考虑用户到目标项目的距离，例如餐厅，选择距离最近的项目为最优的。然而，最近的项目可能无法满足用户的喜好，目标项目如餐厅可能有不同的特征，如食物的类型、价格的范围。当我们推荐项目给移动用户时，我们考虑的应该不仅仅是距离，还有用户对这些特征的喜好。因此，我们提出了一种移动推荐方法同时基于距离和用户对其他属性的喜好。我们扩大了skyline查询的概念，从数据库中选择少量合适的数据推荐给用户，完成个性化的推荐。

Skyline查询是从一个数据库中抽取不被其他任何数据对象支配的数据对象集合[1]。因其在数据挖掘和可视化、多目标决策、用户偏好查询等领域具有潜在的应用，研究者们对skyline查询越来越重视。Skyline查询已经被集成到了商业数据库系统中，目前，如淘宝，大众点评，eBay，亚马逊等几乎所有大型的电子商务系统都在不同程度上使用了skyline查询。其中亚马逊研究电子商务的推荐系统长达十年的时间，目前亚马逊35%的销售量直接来自其推荐系统[2]。并且skyline查询还在用户优先选择查询、信息检索以及Web交互式系统等实际应用中发挥重要的作用。

位置服务（Location-Based Service，LBS）又称适地性服务、行动定位服务、位置服务、基于位置的服务，它是通过移动运营商的无线电通讯网络（如GSM网、CDMA网）或外部定位方式（如GPS）获取移动终端用户的位置消息（地理坐标）[3]。在GIS（地理信息系统）的支持下，为用户提供相应服务的一种增值业务。位置服务可以被应用在不同的领域，例如：健康、工作、个人生活等。此服务可以用来辨认一个人或物的位置，例如发现最近的提款机或朋友同事的目前的位置，也能通过客户目前所在的位置提供直接的手机广告，并包括个人化的天气消息提供，甚至提供本地化的游戏。早期的 LBS系统主要用于在紧急情况下快速定位求助者的位置，以实施救援，如美国的E911系统和欧洲的E112系统[4]。

空间位置是人们日常生活中考虑的一个重要因素，如何在skyline 查询中考虑位置信息，为用户提供基于位置的skyline查询服务成为skyline研究领域的一个新的研究热点。2008年，基于位置的skyline查询被正式提出[5]。随着查询需求的日益复杂，出现了基于位置的多源skyline查询，它是一种特殊的基于位置的skyline查询，在传统的skyline查询上考虑数据点与多个查询点之间的位置关系，查询点是查询用户的当前位置或者用户给定的其它空间对象的位置。传统的基于位置的空间查询包括范围查询，例如对于用户所在的城市查出所有的日本饭馆，或者对于用户所在位置查找出最近的加油站（最近邻查找）。

* 1. 研究意义

随着无线通信技术及智能移动终端的成熟，人们已经越来越离不开便捷的移动生活，而空间位置和用户偏好是人们日常考虑的重要因素。

本文设计的基于Android平台的餐厅推荐系统将有以下特点：

1.将用户位置与用户偏好相结合进行skyline查询：不再是单纯的依靠位置来提供推荐，而是提供离用户近并且用户喜欢的餐厅，。

2. 将查找餐厅和分享美食相结合：系统在查找餐厅的基础上为用户提供该美食所在餐厅的评价，可根据其他用户对美食的评价来选择，也可根据餐厅的菜式和位置和价格来选择，给用户提供更高的可选择性。

3.基于Android手机操作系统：以目前用户规模最大的Android作为平台，不但可利用Google Map提供位置服务，而且界面友好方便，可提高系统的普遍性。

4.可在移动过程中搜罗各种餐厅：系统只需在Android客户端上便能访问，用户可随时随地搜索餐厅、分享美食、收藏餐厅，发现各种有特色的餐厅。

5.为用户和餐厅商家提供便利：用户可以使用系统节省寻找餐厅的时间，同时，系统也为商家宣传餐厅提供了一个平台，提高营业额。

总之，开发一套具有较高灵活性、可扩展性、易管理、易操作餐厅推荐系统，不但可为消费者提供生活上的便利，也对餐厅商家都的效益提高有直接推进作用，从而具有较大的研究意义和价值。

1. 文献综述
   1. Skyline查询的经典算法

早在20世纪70年代，skyline就以最大向量问题的形式出现[6]。随后，2001年Börzönyi等人第一次在数据库研究领域提出了skyline查询的概念[7]，从此以后针对skyline查询技术的研究不断展开。直观上来看，Skyline计算最简单的办法就是将数据点与数据集中的其他点依次进行比较，看是否被支配，可判断这个点是否为Skyline点。然而，当数据集的规模很大时，Skyline查询的响应时间很长，这样做效率就会低。如何最大程度减少不必要的支配比较，执行尽可能少的计算来完成Skyline查询是我们优化以及加快Skyline查询的重点。

根据计算的过程中是否需要建立索引，主要有两大类的算法：不借助索引的算法和索引算法。不借助索引的经典算法有分治Skyline查询方法D&C(Divide-and-Conquer)[7]、嵌套循环查询方法BNL(Block Nested Loop)[7]、位图Skyline查询方法(Bitmap)[8]、排序过滤SFS算法(Sort Filter Skyline)[9]；借助索引的经典算法有索引Skyline查询方法(Index)[8]、NN (Nearest Neighbor) [10]、分支限界Skyline查询方法BBS(Branch and Bound Skyline)[11]。

* + 1. 经典算法

Borzonyi等人提出了BNL算法[7]和D&C算法[7]：BNL算法即循环嵌套算法，在内存中保留一个窗口队列用于存储互不支配的点。一个新到达的点与窗口队列进行比较，初始时队列为空，第一个点放入窗口队列，依次取出下一个点与队列中的点进行比较。如果被队列中的点支配直接将该点删除，如果窗口队列中的点被该点支配，直接将被支配点删除。如果窗口未满则将该点存入窗口，否则写入本地磁盘临时文件中。依次遍历完数据集中的所有点，窗口队列中的点可以直接作为skyline点输出。

D&C算法的思想是按照某一维度的中值进行分组，把维度上大于中值的点存储在𝑃1分组，把维度上取值小于中值的点存储在𝑃2分组。按照同样的划分方法，递归的将P1，P2划分直到划分得到的点集只包含一个点或为空，最后再递归地合并两部分的Skyline结果集。当数据集大到无法一次完全放入内存中时，D&C算法需要从硬盘中读取数据，还至少要将数据写入硬盘一次，导致I/O开销较大，因此，D&C算法在数据集较大时扩展性并不很好，由于D&C需要在所有划分工作完成以后才能计算Skyline点，因此它的渐进性较差。

Tan等人提出了两种渐进式算法：Bitmap算法[8]和Index算法[8]。Bitmap算法也即位图算法，将数据集中的每个点映射为一个位串。要判断点*c*是否为Skyline点时，不需要与其他点进行比较，通过按位操作来确定一个点是否是Skyline点。由于位运算的计算速度快，所以可很快判断这个点是否为Skyline点。然而，当数据集点较多且各维度上不同取值数较多时，位串较长，进行位运算时不仅需要耗费更多的空间，时间效率也随之降低。

Index算法即索引算法，采用B+树对每个列表进行索引，它通过将数据集分成多个列表，如果点在第*i*维上取值最小，那么这个点将被放到第*i*个列表。每个列表又根据索引值分成多个组，在每个组中，组内Skyline点与全局Skyline点中的每一个点进行比较合并到全局Skyline点中。Index算法能够渐进的返回Skyline点，但偏向于先返回那些在某个维度上取值特别好的点，无法按照用户规定的顺序返回结果。

Kossmann等人提出了针对skyline查询的在线查询最近邻居算法(NN)[10]。用户给出自定义的距离公式，通过对数据集利用𝑅-树进行索引。找到离原点最近的点，以这个点为基准对数据空间进行递归划分，原点与最近邻居之间的那个区域中的点一定是skyline点。该算法在确定Skyline点后可以直接输出，不需要读取全部数据集并等待全部处理结束之后才能够给出查询结果，可以实时输出skyline点。因此NN算法适用于实时交互式应用中，具有用户友好性，能够根据用户规定的顺序返回 Skyline 点，但算法执行过程中很多子空间被重复计算，导致算法性能降低。

Chomicki等人提出了SFS算法[9]。在BNL算法基础上按照一个单调函数对数据集进行预排序过滤，排在前面的点不可能被排在后面的点支配。在算法执行的过程中，放在窗口队列的点可以直接输出，避免了将它们与排在它们后面的点进行不必要的比较。它比BNL算法更加高效，提高了BNL算法的时间效率和渐进性。

Padadias等人根据最近邻居的搜索方法提出了渐进式的分值界限算法(BBS)[11]。BBS算法主要采用数据结构中经典的R-树实现，BBS算法依次扩展R-树的节点，逐一返回Skyline点，具有良好的渐进性。它只访问那些最可能包含Skyline数据项的结点，并且不会对同一个结点进行超过两次的访问。当维度较小时，BBS 算法在空间、时间和 I/O 性能上表现的比其它算法优越，能够很好地适应用户需求。但是由于R-树在高维空间上的性能降低，当数据库维数大于6维时，BBS算法性能将受到影响而降低。从综合性能等各方面的考虑，BBS被认为是当前解决集中式Skyline 计算的最优算法。

* + 1. 研究中的问题

1. 结果集过多

很多情况下，返回给用户的Skyline数据集相对于用户期望的过多，不加以选择的将计算所得的全部结果返回给用户。如何选择合适的点返回数据集，是算法是否实用的一个标准。合适的点是符合用户偏好、具有代表性、方便用户的下一步查询，选择合适的点可以大大的提高算法的性能，并且使用户对整个结果集的全貌有一个了解。这个处理过程要求是瞬时的，所以算法需要有相当高的时间性能。

1. 数据不完整

现阶段我们队skyline查询算法的研究大多基于数据完整性这一前提条件，然而现实中我们搜集到的各种类型的数据集往往都是不完整的，即存在属性值缺失的问题。传统处理属性值空缺的方法主要是利用统计学方法或者机器学习方法进行各种预处理填充和修复，然后在填充和修复完整的数据集上做查询，所以要花费大量的时间在预处理的过程中，这样会延长整个系统的响应速度。

1. 数据实时性

实际上，真实数据集往往是动态更新的，几乎所有的应用都期望具有实时的特点。动态数据流上的skyline查询往往很困难，因为数据流模型下的数据规模大、速率快，所以对数据进行快速的处理，连续的查询要求算法有实时在线特性。但是由于数据流的不断到达，时间也在不断地累计，会造成算法的服务质量降低。并且，流上数据速率也不是稳定的,而是随外部环境动态变化。综上，在数据流环境下计算是我们面临的很困难的问题。

* 1. 基于位置的Skyline查询研究

近年来，随着无线通信技术和智能移动终端的广泛应用，基于位置的服务得到飞速发展与普及。用户触发查询并将实时的位置信息传给LBS服务器，服务器根据用户所在的位置返回用户感兴趣的信息以支持用户决策。传统的查询主要考虑用户与数据点之间的位置关系，然而当用户既关注与数据点的距离同时又在数据集的某些属性上有偏好时，传统的基于位置的查询便无法满足用户这种复杂的查询需求。基于位置的Skyline查询可用于解决这个问题[5]，这种特殊的Skyline查询融合了基于位置的查询与Skyline查询的优点，同时考虑数据点的自身属性及数据点与查询用户的位置关系，能够提供更加有效和精致的服务。

* + 1. LBS服务的概述

基于地理位置的服务（Location-Based Services，LBS）起源于上个世纪70年代的美国，它并不是一项新技术。早期的LBS系统主要用于在紧急情况下快速定位求助者的位置，以便实施救援，如美国的E911系统和欧洲的E112系统[4]。当前LBS已经广泛应用在军事、交通、物流、医疗等领域中。

LBS是利用现代计算机技术、移动通信技术、移动互联网技术、移动定位技术、全球定位系统、地理信息系统以及遥感等相关技术集成提供与位置相关的信息及规律的特殊服务。LBS移动应用能够满足当前手机用户对信息服务四个方面的要求，即在任何时间、任何位置、任意设备和所有信息内容上享有合理的自由度。一些位置服务，包括提供天气信息，提供公交信息的服务在内，尽管与用户当前的位置无关，但也可以归于LBS服务之内[19]。

LBS系统由计算机网络（Internet）、移动通信网络（GPRS或HSDPA）组成、移动终端（智能手机、平板电脑等）以及LBS服务平台组成[12]。LBS 系统通过移动设备本身的全球定位功能（Global Positioning System，GPS）或者通过运营商提供的无线通讯网络（如GSM，CDMA等）获取用户的具体位置以及随时间变化规律，并通过移动通讯网络向LBS应用运营商的服务平台发送当前位置信息及其他请求，LBS服务平台利用推荐算法跟据用户当前位置、情境信息对用户请求进行处理，将结果通过移动通讯网络发送给用户。

经过十几年的发展，LBS应用已具备定位快捷准确、计算功能强大、种类丰富多样、使用简单方便等一系列特点。当前的LBS应用以不仅仅满足于向用户提供导航信息、天气信息等传统服务模式，而是向着个性化、电子商务化、社交化的模式不断发展。当下常用的各类移动应用，包括通讯软件—微信，电子商务平台—大众点评网，社交软件—微博等都纷纷加入LBS系统。传统LBS移动应用与其他类型移动应用的交叉与融合已经成为移动应用开发的重要方向，这也使得LBS系统不断深入人们日常生活的各个方面，从而不断改善人们的生活方式。

随着LBS系统功能的不断丰富和扩展，LBS系统需要处理大量信息，LBS系统的通信速度与计算能力已成为影响其成功与否的重要因素。另外，随着LBS技术的不断发展，如今的LBS系统已不仅仅根据用户的地理位置提供服务，而是综合考虑用户当前所处情境与上下文信息。用户的情景信息是能够影响用户需求表达的环境因素，既包括主观因素也包括客观因素。其中，主观因素主要指和用户自身的特性如个人偏好，生活习惯等; 客观因素主要指当前的时间、当前地理位置、当前天气等。情境信息会很大程度的影响用户的短期偏好。为了给用户提供个性化服务和信息，LBS系统所使用的推荐算法需要综合分析用户的情境信息。

* + 1. 研究中的问题

1. 隐私保护

基于位置服务中的隐私内容涉及两个方面[13]，即位置隐私和查询隐私。位置隐私中的位置指用户过去或现在的位置，查询隐私指涉及敏感信息的查询内容。目前也有一些技术来保护隐私，如假位置[14]、时空匿名[15]、空间加密[16]等。然而位置隐私研究中依然面临着几个方面的挑战：隐私保护与位置服务是一对矛盾，提供位置服务的同时必然会获取你的位置；基于位置服务的请求需要进行在线处理，所以位置匿名具有实时性要求;不同用户的隐私要求大相径庭，所以隐私保护需要满足个性化的需求。

1. 个性化需求

基于位置服务的个性化需求越来越突出，因此这类服务需要符合用户个人爱好以及用户上下文信息，实现高效精准的服务。通常情况下，用户的需求不仅仅是距离，还会附加一些属性，所以在实现算法时需要考虑一些因素。

1. 研究内容、预期目标及研究方法
   1. 研究内容

假设用户*A*在寻找一家餐厅吃饭，希望通过请求Skyline查询寻找感兴趣的餐馆，他们不仅关注餐馆的价格和类型同时希望餐馆离他们越近越好。一个餐厅的数据集存在用户移动客户端的RDBMS中，如表3.1所示。表中存有十个餐厅，除了location这个属性外，还有type和price这两个属性。

**表3-1餐厅数据库**

|  |  |  |  |
| --- | --- | --- | --- |
| Restaurant | Location | Type | Price |
| A | (3,9) | Chinese | Low |
| B | (7,5) | French | Medium |
| C | (7,7) | Japanese | Low |
| D | (5,1) | Italian | High |
| E | (3,4) | Japanese | Medium |
| F | (4,8) | French | Low |
| G | (5,6) | Chinese | High |
| H | (1,3) | Italy | Low |
| I | (5,2) | Japanese | Medium |
| J | (9,3) | Chinese | High |

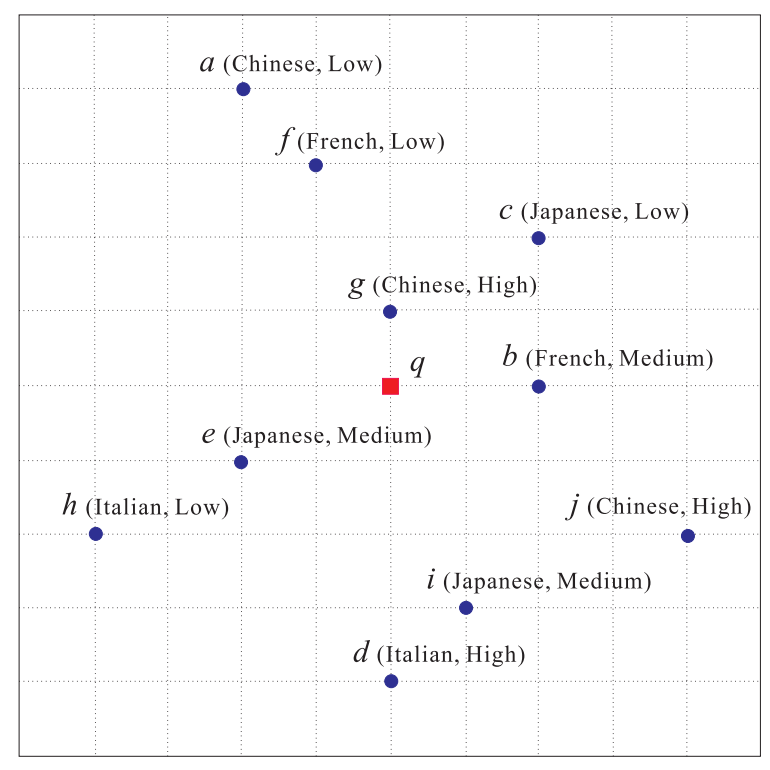
假设用户*A*的位置是(5,5)，用户提供的偏好文件信息是：

Type这个属性的排序是： ，符号’’表明互不支配。

Price这个属性的排序是：。

用户的位置以及10家餐厅的位置如图3.1显示，*g*表示用户所在位置（5,5）：

**图3-1 餐厅位置**



* + 1. Skyline查询相关定义

对于*λ*-维空间*A*={*a*1 ,*a*2,…*ad*}的数据集*D*，任意一点*o*∈*D*，*o*在维度*ai*上的取值记为*o*.*ai*，我们假设*D*在各个维度*ai*上的属性值是可比较的。具体用＞和＜比较，我们假设属性值越小越优，对于数据集中的任意两点o和o’，若o.ai＜o’.ai，表明ai维度上点o的取值要优于点o’；若o.ai＞o’.ai，则说明在维度ai上o的取值要差于o’。基于这些假设，我们有如下的定义：

**定义3.1 支配**：

对于D中的任意两点o和o’，如果o在所有属性上的取值不比o’差且至少在一个属性上的取值优于o’，那么o支配o’，记为。

在上面的例子中，e餐厅在location，type，price所有属性上的值都是优于a餐厅的，所以我们表示为。

**定义3.2 Skyline点：**

任意一点，o是数据集D在空间S上的Skyline点当且仅当不存在任何点满足。

定义3.3 **Skyline查询**：

Skyline查询返回数据集D在在空间A上的所有Skyline点。所有的Skyline点都不被其他点所支配。

在上个例子中，餐厅b，d，g不被任何点所支配，所以查询返回的结果就是S(O)={b,d,g}。

* + 1. 算法流程

这个算法应用于每个维度由有限个离散数据组成的数据集。算法将用户的偏好存储在本地文件中，根据用户偏好文件，将数据集映射成一个模式表，利用最近邻搜索，查找出离用户当前位置最近的数据点，查看模式表，检查该数据点的模式是否在模式表中，如果在，说明该数据点是Skyline点，之后将这个数据点对应的模式及这个模式能支配的所有模式从模式表中删除。算法进行下一轮查找，利用最近邻搜索，找出下一个离用户最近的点，重复上面的步骤，直到输出数据集中的所有Skyline点。算法执行的过程本质上是按数据点与用户的距离从小到大的排序依次查找Skyline点。利用最近邻搜索找到的离用户最近的点为第一个Skyline点，可直接输出，因此算法能快速响应用户，同时具有良好的渐进性。

根据描述，我们写出这个算法来解决用户既需要距离近又满足偏好的需求。

|  |
| --- |
| 算法1 基于距离和偏好的Skyline查询 |
| 1. **function** SSP(p, q) p:用户的偏好文件，q:用户的位置 2. S←∅ 数据集非空 3. 把用户偏好文件p映射成一个模式表T 4. 初始化最近邻搜索，用户的初始位置为q 5. **repeat** 6. 找到最近的点o 7. **if** o在模式表T中，那么o是一个skyline点 8. 把o放入S中 9. 将o点所对应的模式从T中删除 10. 将o点对应的模式所能支配的所以模式从T中删除 11. **end if** 12. **until** T为空 13. **return** S 14. **end function** |
|  |

* + 1. 算法示例

根据上面的算法和例子，我们对算法做一个演示。

首先先根据用户偏好文件映射成一个模式表T，如表3-2所示。

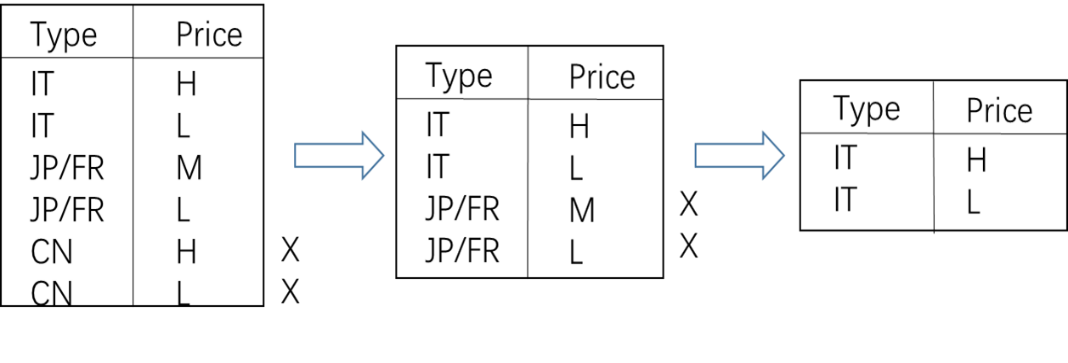
**表3-2 模式表T**

|  |  |
| --- | --- |
| Type | Price |
| Italy | High |
| Italy | Low |
| Japanese/French | Medium |
| Japanese/French | Low |
| Chinese | High |
| Chinese | Low |

我们根据最近邻搜索搜索到餐厅g（Chinese，High），在距离这个属性上，它不被任何一点所支配。所以它属于skyline点，我们把它加入数据集S中，S={g}。这时我们可以把g对应的模式（Chinese，High）和这个模式能支配的模式（Chinese，Low）从模式表T中删除。现在还剩四个模式在表T中，其他的餐厅如果对应的模式不在模式表中，那它就不可能是skyline点。然后我们重复之前的步骤，找到餐厅b（French，Medium），我们发现b没有被其他的skyline点所支配，所以我们把b也加入S中，S={b，g}。继续删除（Japanese/French，Medium）和（Japanese/French，Low）。此时，模式表T中只剩下两个模式。

这个过程如图3-2所示：

**图3-2 模式表T的变化**



然后我们又继续搜索到e（Japanese，Medium），发现模式表T中没有符合它的模式，所以它不是skyline点，我们依然采用最近邻搜索，直到搜索到餐厅d（Italy，High），才发现模式表T中有所对应的模式，并且之前的skyline点不能支配d，我们把d放入S中，S={b，d，g}。然后模式表被删空，算法结束，我们把S返回给用户。

* 1. 预期目标

本设计将实现在Android平台下的基于位置和用户喜好的skyline查询及推荐系统，系统将餐厅搜索、餐厅点评和餐厅导航于一身，为用户提供随时随地的查询服务。首先对推荐系统的总体进行设计，包括系统服务器端和客户端的架构设计。接着，对推荐系统各个功能模块的设计，包括Web 端网站管理、餐厅搜索、餐厅导航和餐厅评价的设计。最后，对系统的后台数据库进行设计。

本设计需要在设计完成后，能实现以下功能：

1.用户可以在Android客户端实现用户登录及用户注册功能，并进入系统Android客户端主界面。

2.在客户端界面可以实现查看附近的餐厅、查看推荐、进行餐厅评分，并分别实现各个功能。

3.在客户端实现查看餐厅详情，访问餐厅路径，最后实现推荐功能。

4.管理员可数据库中添加餐厅的信息，如餐厅的位置，电话等等。

另外，本设计在性能上要达到安全性、可维护性和稳定性三项要求。安全性是要求对网站管理员和一般客户实现对用户操作和数据库访问权限进行管理与控制，对用户的合法权进行保护。可维护性就是要求在系统开发过程中，系统文档、指示代码体系遵循软件开发的规范要求。稳定性就是要求系统必须保持性能稳定，在用户访问时能保证较好的响应速度。

* 1. 研究方法
     1. 系统平台开发与工具介绍

基于用户位置和偏好的skyline查询与推荐系统是基于Android平台开发的，包括Android客户端和Web端，

采用Java、JSP等编程语言，具体实现平台如表3-3所示。

表3-3系统开发平台

|  |  |
| --- | --- |
| 操作系统 | Windows 10 |
| 开发工具 | eclipse+ADT23.0.7+Android SDK |
| 运行环境 | JDK1.7+Android4.0+Google API |
| Web服务器 | Tomcat6.0 |
| 虚拟设备 | Genymotion |
| 数据库 | SQLite |
| 地图 | Google Map |

Eclipse是一个开放源代码的，基于Java的整合型可扩展开发平台，利用它可以在数据库和JAVA EE的开发、发布以及应用程序服务器的整合方面极大地提高工作效率。Eclipse的本身只是一个框架平台，但是众多插件的支持使得Eclipse拥有其他功能相对固定的IDE软件很难具有的灵活性。本设计将采用eclipse作为代码主要开发工具。

Android开发环境也需要在eclipse平台上搭建，包括安装ADT（Android Develop Tools）插件和Android SDK（Software Development Kit）。ADT是一个提供了与eclipse IDE集成工具套件的eclipse插件，它集成到 eclipse中的许多开发工作流任务，使开发者可以快速开发和测试 Android 应用程序。Android SDK为开发者提供了库文件以及其它开发所用到的工具。另外，我们在eclipse环境下使用Genymotion模拟器可以创建虚拟的安卓设备，它的运行速度和流畅度都非常快，方便开发人员在电脑上直接测试应用程序。

Tomcat 是一个小型的轻量级 Web 应用服务器，在中小型系统和并发访问用户不是很多的场合下被普遍使用，是开发和调试JSP 程序的首选，因为Tomcat 技术先进、性能稳定，成为目前比较流行的Web 应用服务器。本系统设计的Android客户端和Web端都需要通过Web服务器与后台数据交互，因此采用Tomcat服务器。

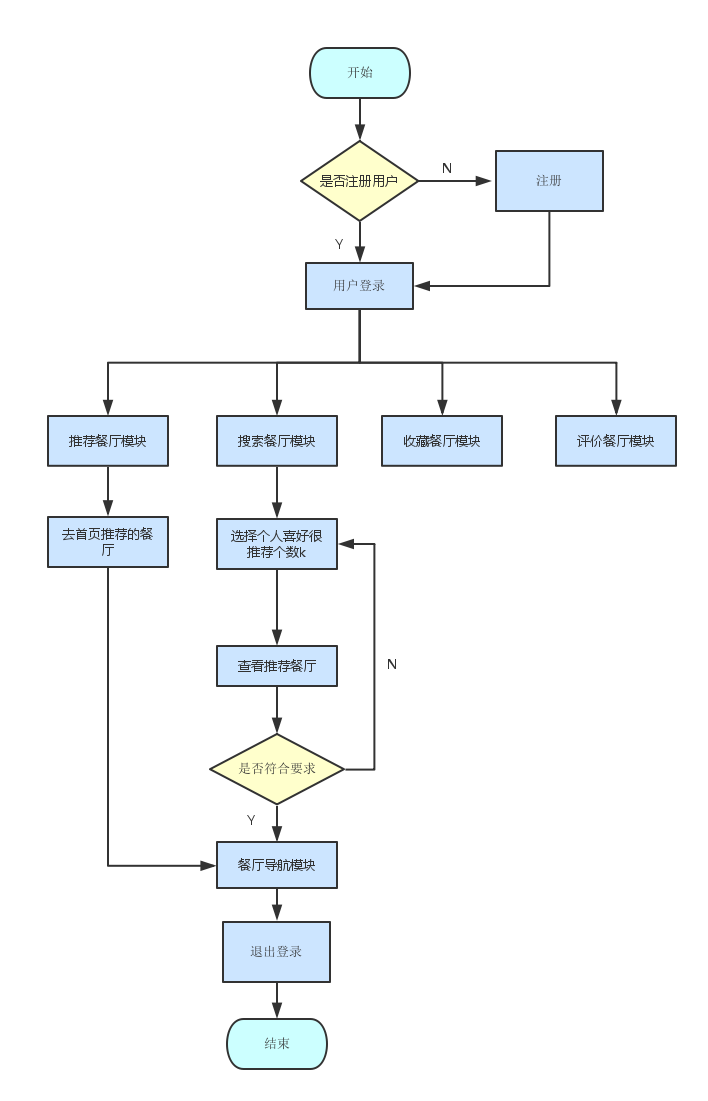
SQLite 一个非常流行的嵌入式数据库，它支持SQL语言，并且只利用很少的内存就有很好的性能。此外它还是开源的，任何人都可以使用它。许多开源项目（(Mozilla, PHP, Python）都使用了SQLite。所以每个Android应用程序都可以使用SQLite数据库。对数熟悉SQL的开发人员来时，使用 SQLite相当简单。我们开发这个轻量级的系统使用SQLite数据库非常的方便实用。

* + 1. 系统功能模块设计

在基于位置和用户喜好的餐厅推荐系统中。客户可通过Android终端访问系统Android客户端，通过Web浏览器访问系统Web客户端；网站管理员可通过Web浏览器访问系统的Web管理端，并进行相应的管理和维护。Android客户端和Web端都是通过同一台服务器来操作后台的数据库服务器，实现两端获取的数据的一致性。

基于位置和用户偏好的skyline查询与推荐系统是在移动过程中进行skyline查询，查出好的餐厅来推荐给用户。因此系统的Android客户端的主要工作流程如图3-3所示。从图中可以看出，用户首先登录系统，或者注册后登录系统后，分别可以进行查看餐厅推荐、查看收藏的餐厅，以及搜索餐厅操作，在搜索餐厅操作后，我们让顾客先选择价格区间以及喜爱的菜式，通过skyline算法推荐给用户*k*个餐厅。如果用户对某个餐厅感兴趣，可以查看餐厅的地图位置，当这些符合用户的要求，就可以进入餐厅导航功能模块，完成推荐过程。用餐完成后中还可给餐厅作出评论并上传相关信息。

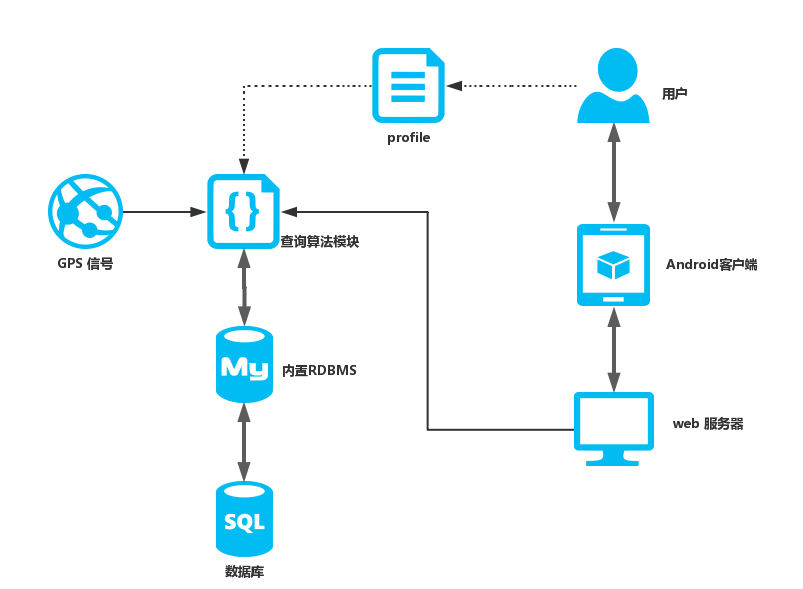
**图3-3 系统流程**



* + 1. 系统结构

该系统的结构如图3-4所示。我们的实用了Windows系统的PC端作为系统中的服务器端，用户最初设置他的偏好和参数*k*（查询结果的最小值）存储在查询处理模块里。查询处理模块是在这个系统中的重要的组成部分，执行skyline查询处理。当用户开始移动时，GPS信号定期获取和查询用户的当前位置。查询处理过程的一些部分是在数据库系统中执行，当计算出skyline点的合集时，通过socket通信传输给用户的Android客户端，并且在客户端上接了一个Google Map的API实现交互功能，可以使用户使用Google map去导航。

**图3-4 系统结构**



### 3.3.4 系统数据来源

1. 把采集的数据或者现成的数据集存入数据库中对应的数据表中。

2. 把获取来的数据存入SQLite中，开发轻量级的应用使用这个数据库十分便捷。

3. 把存在数据表里的地址信息读取出来，通过调用Google Map API把地址转化成经纬度。

4. 按行更新GIS信息。

1. 研究进度安排

1月—2月 收集资料，完成开题报告，确定使用的编程语言和平台。

3月—4月 实现系统的代码编写，并且测试系统的各个功能模块，完成中期报告。

5月—6月 撰写毕业论文。

学生本人签字：

年 月 日

参考文献

1. 朱琳, 关佶红, 周水庚. Skyline计算研究综述,计算机工程与应用, 2008, 44(6):160-165.
2. 蒋国瑞，青海，黄梯云。一种柔性的电子商务推荐系统[J].计算机应用研究，2009,03:930-932,976.
3. 维基百科。地理位置服务相关概念[EB/01]. <https://wikipedia.kfd.me/zh-cn/>，2016-03-29.
4. 周傲英，杨彬，金澈清，马强。基于位置的服务：架构和进展[J].计算机学报，2011,34(7):1155-1171.
5. Zheng B.H, Ken C. K. Lee ,and Wang-Chien Lee. Location-Dependent Skyline Query[A].The 9th IEEE International Conference on Mobile Data Management[C].2008:148-155.
6. Kung H T, Luccio F, Preparata F P. On finding the maxima of a set of vectors[J].J ACM,1975,22(4):469-476.
7. Borzsonyi S, Kossmann D, Stocker K. The Skyline Operator [A] Proceedings of the 17th International Conference on Data Engineering[C]. Heidelberg, Germany,2001:421-430.
8. Tan K, Eng P, Ooi B. Efficient Progressive Skyline Computation[A].

The 27th International Conference on Very Large Data Bases[C]. Rome,Italy,2001:301-310.

1. Chomicki J, Godfrey P, Gryz J, et al. Skyline with Presorting[A]. The 19th International Conference on Data Engineering[C]. Bangalore,India,2003: 717-719.
2. Kossmann D, Ramsak F, Rost S. Shooting stars in the sky:An online algorithm for skyline queries[A]. In: Proc. of the Int’l Conf. on Very Large Data Bases[C]. 2002:275-286.
3. Papadias D ,Tao Y.F, Fu G,et al. Progressive Skyline Computation in Database Systems[J].ACM Transaction on Database Systems,2005,30(1):41-82.
4. Ana M. Bernardos, José R. Casar, Paula Tarrio. Building a framework to characterize location-based services[C].The 2007 International Conference on Next Generation Mobile Applications, Services and Technologies, Proceedings,2007:110-115.
5. GEDIK B, LIU Ling. Location privacy in mobile systems:a personalized anonymization model[C]. Proc of the 25th IEEE International Conference on Distributed Computing Systems.2005:620-629．
6. MOKBEL M F, CHOW C Y, AREF W G. The new casper: query

processing for location services without compromising privacy[C]//Proc of the 32nd International Conference on Very Large Data Bases.2006:763-774.

1. KHOSHGOZARAN A, SHAHABI C. Blind evaluation of nearest neighbor queries using space transformation to preserve location privacy[C]//Proc of the 10th International Conference on Advances in Spatial and Temporal Databases.2007:239-257．
2. KIDO H, YANAGISAWA Y, SATOH T. Protection of location privacy using dummies for location-based services[C]//Proc of the 25th International Conference on Distributed Computing Systems.2005．

指导教师意见

该生通过与指导教师充分讨论，参考了许多文献，确定了基于地理位置和用户偏好的移动推荐系统这个课题。移动推荐系统是目前学术界和工业界的研究热点，并且具有一定的实用性。该课题基于学术界已经提出的基于位置和用户偏好的Skyline查询技术，设计思路合理明确。通过设计和开发该系统，可以帮助学生巩固专业知识，提高编程技能和研究能力。研究方法和研究计划基本合理，难度适中，学生能够在预定时间内完成该课题。

指导教师签字：

年 月 日