



# 南京大學

## 研究生畢業論文 (申請碩士專業學位)

論文題目 基于矢量瓦片和优先点树的相似路径检索和可视化服务的设计和实现

作者姓名 韩淳

学科、专业 软件工程

研究方向 软件工程

指导教师 刘海涛 讲师

年 月 日

学 号 : MF1732038

论文答辩日期 : 年 月 日

指 导 教 师 : (签字)



# **The Design and Implementation of Similar Path Search and Visualization Service based on vector-tile and vp-tree**

By

**Han Chun**

Supervised by

Advisor Title **Si Li**

A Thesis

Submitted to the Software Institute

and the Graduate School

of Nanjing University

in Partial Fulfillment of the Requirements

for the Degree of

**Master of Engineering**

Software Institute

May 2019



# 南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目：基于矢量瓦片和优先点树的相似路径检索和可视化服务的设计和实

软件工程 专业 2017 级硕士生姓名：韩淳

指导教师（姓名、职称）：刘海涛 讲师

## 摘 要

地理数据，是直接或间接关联着相对于地球的某个地点的数据，是表示地理位置、分布特点的自然现象和社会现象的诸要素数据。地理数据包括轨迹数据和瓦片数据。其中轨迹数据指的是一个物体的运动轨迹在空间中经过的点的集合。瓦片数据指的是用于展示的地图数据。在地理数据用户的业务中，一个非常有价值的场景就是犯罪同伙分析，通过框选选定目标嫌疑人轨迹，再通过检索与目标嫌疑人活动轨迹相似的其他人员的轨迹，从而找到目标嫌疑人的潜在同伙。针对以上这样的使用场景，本文构建了一个能够提供地图路径展示和相似路径检索的服务。本文主要关注两部分，第一部分是如何建立高效的轨迹索引结构，以提高良好的相似性检索的性能。第二部分是如何构建一个通用的瓦片数据服务，用来高效地提供地图数据。

该数据服务的主体思路是分别构建轨迹数据和地图数据的存储检索结构，然后分别从轨迹数据服务和地图数据服务中提取相关数据，再利用前端库整合两份数据共同显示在浏览器上，从而让用户方便快捷地看到轨迹检索结果。

文章的主要工作就是介绍以上两个数据服务的设计和实现。首先，本文介绍了目前在地理数据服务中主要流行的技术，包括通用的技术标准和本文的关键数据结构优先点树和矢量瓦片，并分析了这两种结构在地理数据服务的作用和优势。在后续章节，通过整个系统的需求以及检索端的流程来详细介绍这两种结构在系统中的角色。接着介绍了整个服务的设计架构，各个模块的职责划分和各个模块的详细设计和实现。在该过程中还讨论并分析了系统中出现的一些性能以及模型上的问题，提供了更好的解决方案，并且通过具体的性能数据来说明解决方案的提升点。最后，详细分析了整个服务对地理轨迹数据检索方面的效果提升。

地理轨迹数据服务，（Geographic trajectory data service, 以下简称GTDS），通过构建独立的索引结构和通用的瓦片数据服务，实现了相似轨迹的检索和展示功能，解决了用户应用中一些难题，提高了系统的可用性。文章的最后部分，通过总结和展望，对该技术以及应用前景进行了一些分析。

**关键词：** 地理数据，轨迹数据，瓦片数据，优先点树,矢量瓦片，轨迹索引结构，轨迹相似检索，可视化

## 南京大学研究生毕业论文英文摘要首页用纸

THESIS: The Design and Implementation of Similar Path Search and Visualization Service based on vector-tile and vp-tree

SPECIALIZATION: Software Engineering

POSTGRADUATE: Han Chun

MENTOR: Advisor TitleSi Li

### **Abstract**

Geographic data is data that is directly or indirectly related to a certain location on the earth. It is the data of natural phenomena and social phenomena that represent geographic location and distribution characteristics. Geographic data includes trajectory data and tile data. Where trajectory data refers to a collection of points through which an object's motion trajectory passes in space. Tile data refers to the map data used for display. In the business of geographic data users, a very valuable scenario is the analysis of criminal accomplices. By selecting the target suspect's trajectory, and by searching the trajectory of other people similar to the target suspect's trajectory, the target suspect is found. Potential associates. In view of the above usage scenarios, this paper constructs a service that can provide map path display and similar path retrieval. This article focuses on two parts. The first part is how to build an efficient trajectory index structure to improve the performance of good similarity retrieval. The second part is how to build a generic tile data service to efficiently provide map data. The main idea of the data service is to separately construct the storage retrieval structure of the trajectory data and the map data, and then extract the relevant data from the trajectory data service and the map data service respectively, and then integrate the two pieces of data into the browser by using the front-end library, thereby Let users see the track search results quickly and easily. The main work of the article is to introduce the design and implementation of the above two data services. First of all, this paper introduces the current popular technologies in geographic data services, including common technical standards and the key data structures of this paper, vp-tree and vector tile, and analyzes the roles and advantages of these two structures in geographic data services. In the following chapters, the roles of the two systems in the system are described in

detail through the requirements of the entire system and the processes at the search end. Then it introduces the design structure of the whole service, the division of duties of each module and the detailed design and implementation of each module. In the process, some performance and model problems appearing in the system are also discussed and analyzed, which provides a better solution and shows the solution's lifting point through specific performance data. Finally, the effect of the whole service on the retrieval of geographic trajectory data is analyzed in detail. The geographic trajectory data service realizes the retrieval and display function of similar trajectories by constructing an independent index structure and a common tile data service, solving some problems in the user application and improving the availability of the system. In the last part of the article, through the summary and outlook, some analysis of the technology and application prospects.

**Keywords:** English, Geographic data, trajectory data, tile data, vantage point tree, vector tile, trajectory index structure, trajectory similarity retrieval, visualization



# 目 录

目录 .....	v
第一章 引言 .....	1
1.1 这是节标题 .....	1
1.1.1 这是小节标题 .....	1
1.1.2 这是小小节标题 .....	1
第二章 相关技术综述 .....	3
2.1 这是节标题 .....	3
2.1.1 这是小节标题 .....	3
2.1.2 这是小小节标题 .....	3
第三章 系统需求分析与概要设计 .....	5
3.1 GTDS系统概述 .....	5
3.2 轨迹数据服务需求分析 .....	6
3.2.1 轨迹数据服务的功能需求 .....	6
3.2.2 轨迹数据服务的非功能需求 .....	6
3.3 瓦片数据服务需求分析 .....	6
3.3.1 瓦片数据服务的需求综述 .....	6
3.3.2 瓦片数据服务功能需求 .....	7
3.3.3 瓦片数据服务非功能需求 .....	7
3.3.4 瓦片数据服务用例图 .....	7
3.3.5 瓦片数据服务用例描述 .....	8
第四章 系统详细设计与实现 .....	11
4.1 这是节标题 .....	11
4.1.1 这是小节标题 .....	11
4.1.2 这是小小节标题 .....	11

第五章 总结和展望 .....	13
5.1 这是节标题 .....	13
5.1.1 这是小节标题 .....	13
5.1.2 这是小小节标题 .....	13
简历与科研成果 .....	15
致谢 .....	17

## 表 格

3.1	轨迹数据服务功能需求列表.....	6
3.2	轨迹数据服务非功能需求列表 .....	6
3.3	瓦片数据服务功能需求列表.....	7
3.4	瓦片数据服务非功能需求列表 .....	7
3.5	bounding-box更新地图瓦片数据用例描述表 .....	9



## 插图

3.1 GTDS的总体架构 .....	5
3.2 瓦片数据服务的功能需求 .....	8



## 第一章 引言

### 1.1 这是节标题

这是地理信息系统的引言

#### 1.1.1 这是小节标题

#### 1.1.2 这是小小节标题





## 第二章 相关技术综述

### 2.1 这是节标题

这是地理信息系统的相关技术综述

#### 2.1.1 这是小节标题

#### 2.1.2 这是小小节标题



## 第三章 系统需求分析与概要设计

### 3.1 GTDS系统概述

地理轨迹数据服务总体上分为三个部分，轨迹数据服务，瓦片数据服务和业务展示服务。其中业务展示服务是用户直接操作的前端，完成轨迹数据和瓦片数据的可视化功能。**注意：由于业务展示服务只有前端库的调用，没有有价值的实现，因此本文不做介绍。**瓦片数据服务主要面向外部数据源，支持瓦片数据源的配置和瓦片的增，删，改，查等功能。轨迹数据服务主要面向数据库管理员，主要功能是负责轨迹存储和相似性检索。如图 3.1所示，可视化服务

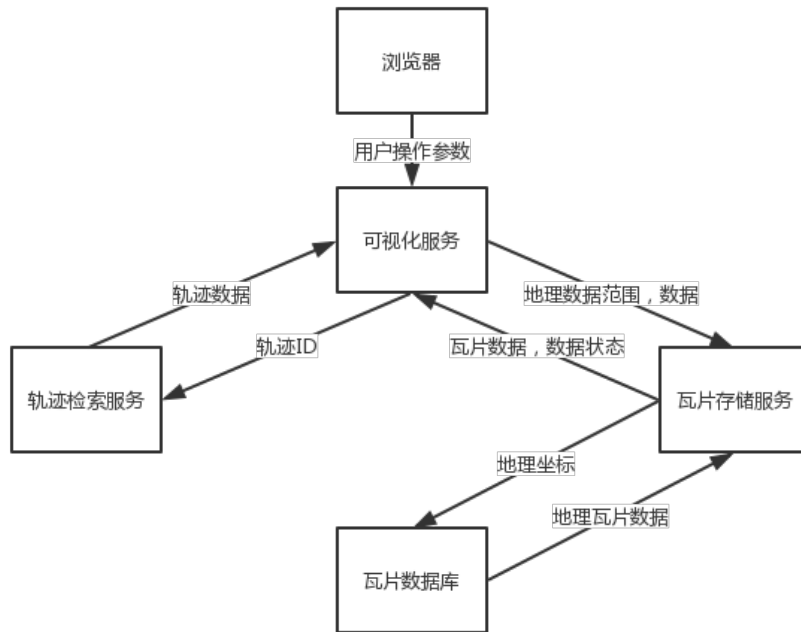


图 3.1: GTDS的总体架构

直接接收用户的界面操作，将其转化为轨迹ID和瓦片数据范围，作为参数，分别传递给轨迹搜索服务和瓦片存储服务。再使用前端库汇总这两部分的数据共同完成可视化功能。而瓦片存储服务本身不存储瓦片数据，只负责地图范围的解析，地理坐标的转换以及数据格式的转换，瓦片数据的存放位置是在具体的

瓦片数据库中。

## 3.2 轨迹数据服务需求分析

### 3.2.1 轨迹数据服务的功能需求

整个地理轨迹数据服务的核心功能就是对相似路径的检索，因此轨迹数据服务的功能需求实际上就是一个Trajectory KNN功能。即根据目标轨迹的ID，执行检索算法，找出与目标轨迹最相近的K条轨迹。

表 3.1: 轨迹数据服务功能需求列表

需求ID	需求名称	需求描述
R1	相似轨迹knn检索	服务调用方能够通过传递目标轨迹ID和检索量K，在规定时间内返回K个与目标轨迹最相似的K个轨迹

### 3.2.2 轨迹数据服务的非功能需求

Trajectory KNN问题是一个计算密集型问题，用户对响应时间一定会有要求。同时，并发量也是必须考虑的问题

表 3.2: 轨迹数据服务非功能需求列表

时间特性	对于百万级别的轨迹数据量，服务应该在2s之内返回检索结果
负载特性	服务应该能应对10w以上的并发访问

## 3.3 瓦片数据服务需求分析

### 3.3.1 瓦片数据服务的需求综述

在轨迹检索服务中，对于地图瓦片数据的要求有增加，删除，修改，查询，数据源配置。其中以查询和更新这两部分功能的细分功能最多。对于查询而言，可能被用户需要的有整张地图的全量查询，局部bounding-box渲染查询，单个瓦片数据的查询。对于更新而言，可能需要局部地图的更新，bounding-box更新。注意：在GTDS的功能需求中，对于瓦片的增加和删除都是以整张地图为单位的，局部的增加和删除这种需求并不存在，所以此处不列为功能需求。而对于非功能需求，瓦片数据服务涉及到的计算主要是坐标转换，请求解析，缓存处理。计算量不大，所以并不是计算密集型应用，而是IO密集型应用，高并发和快速响应是其必须满足的非功能特性。除此之外，由于瓦片服务本身不存

储瓦片数据，用户的瓦片数据可能是存在各种不同的数据库中的，因此服务还应该独立于不同的数据库，做到高可用性，高扩展性。

### 3.3.2 瓦片数据服务功能需求

表 3.3: 瓦片数据服务功能需求列表

需求ID	需求名称	需求描述
R1	新增地图瓦片数据	服务调用方能够通过瓦片服务，参数为地图名称和图瓦片数据，增加一个地区的完整地图瓦片数据
R2	删除地图瓦片数据	服务调用方能够通过瓦片服务，参数为地图ID，删除一个地区的全部地图瓦片数据
R3	局部更新地图瓦片数据	服务调用方能够通过瓦片服务，参数为地图ID和地图数据，更新一张大地图中某一个地区的地图瓦片数据
R4	bounding-box更新地图瓦片数据	服务调用方能够通过瓦片服务，参数为地图ID，经纬度范围，地图数据，更新一张大地图中某一个地区某一经纬度矩形范围内的地图瓦片数据
R5	局部更新地图瓦片数据	服务调用方能够通过瓦片服务，参数为地图ID和地图数据，更新一张大地图中某一个地区的地图瓦片数据
R6	地图瓦片数据全量查询	服务调用方能够通过瓦片服务，通过获取全量数据的JSON文件，作为输入的数据源，获取整张地图的全量数据
R7	单个地图瓦片数据查询	服务调用方能够通过瓦片服务，参数为地图ID和栅格坐标zxy，获取指定的瓦片
R8	局部bounding-box的渲染查询	服务调用方能够通过瓦片服务，参数为地图ID和经纬度矩形范围，获取这一部分的地图渲染结果

### 3.3.3 瓦片数据服务非功能需求

表 3.4: 瓦片数据服务非功能需求列表

时间特性	在正常负载情况下，服务的平均响应时间应在1s之内
负载特性	服务能稳定应对百万级别的并发访问，不会出现延迟超过10s或服务崩溃的情况
高可用性	服务能通过设置中间件的方式，便捷地对接到各种不同的数据库，并保证运行正常

### 3.3.4 瓦片数据服务用例图

注意：以上用例中的功能并不是逻辑完备的，某些逻辑功能，比如局部地图瓦片非渲染查询，这种需求在实际应用中并没有使用场景，这里就没有列举。

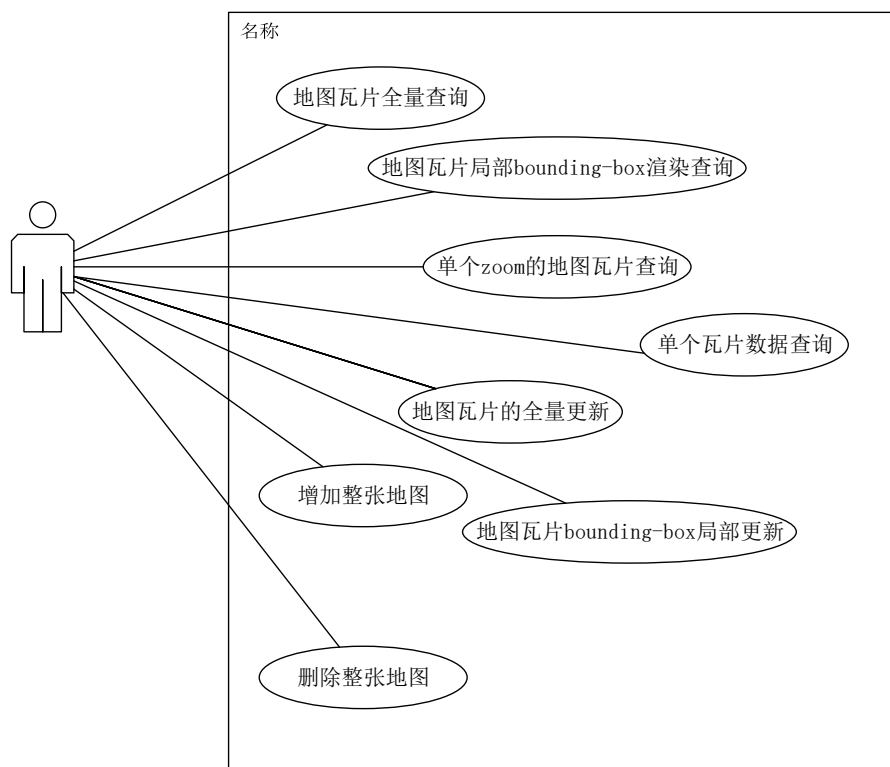


图 3.2: 瓦片数据服务的功能需求

### 3.3.5 瓦片数据服务用例描述

瓦片数据服务的部分功能操作简单明确，无需使用用例描述。本文只对操作比较复杂的“bounding-box更新地图瓦片数据”进行了具体描述。

表 3.5: bounding-box更新地图瓦片数据用例描述表

<b>ID</b>	<b>UC1</b>
<b>名称</b>	<b>bounding-box更新地图瓦片数据</b>
<b>参与者</b>	普通用户
<b>目的</b>	更新某张地图中一个矩形范围内的瓦片数据，以改变轨迹展示背景
<b>描述</b>	用户通过浏览器发送矩形参数和瓦片数据，以实现对瓦片数据库中数据的修改，进而改变业务展示的结果
<b>优先级</b>	高
<b>触发条件</b>	某一地图中某一部分数据发生变更，需要更改
<b>前置条件</b>	服务正常运行，用户进入服务界面
<b>后置条件</b>	地图瓦片数据完成更新，业务展示结果经过刷新可以看到地图变化
<b>正常流程</b>	1.进入瓦片数据更新界面 2.设置north,south,west,east,4个经纬度值 3.设置zoom区间 4.设置是否进行精度模糊 5.点击浏览本地文件并上传文件数据 6. 点击更新瓦片数据
<b>异常流程</b>	2a.经纬度大小值错误，前端应自动检测并警告 3a.指定矩形范围太小，在较小的zoom下无法更新瓦片,服务端应发现错误并返回告知用户 5a.用户选择的文件格式错误，则前端对用户发出警告 5b.用户选择的文件中并没有指定bounding-box中的数据，服务端应返回错误报告并提示用户更改文件





## 第四章 系统详细设计与实现

### 4.1 这是节标题

这是地理信息系统的系统详细设计与实现

#### 4.1.1 这是小节标题

#### 4.1.2 这是小小节标题



## **第五章 总结和展望**

### **5.1 这是节标题**

这是地理信息系统的总结和展望

#### **5.1.1 这是小节标题**

#### **5.1.2 这是小小节标题**



## 简历与科研成果

**基本情况** 韩淳，男，汉族，1994 年 8 月出生，吉林省松原市人。

### 教育背景

**2017.9～2019.6**    南京大学软件学院    硕士

**2013.9～2017.6**    南京大学软件学院    本科

这里是读研期间的成果（实例为受理的专利）

1. 刘海涛，**韩淳**，“基于矢量瓦片和优先点树的相似路径检索和可视化服务的设计和实现”，申请号：20xx1018xywz.a，已受理。



## 致 谢

这里是致谢。一般的感谢顺序：导师，其他指导老师，师兄姐妹、同学，父母和伴侣。