分类号 密级

UDC

学 位 论 文

基于深度学习的惯性导航室内定位算法研究

|  |  |  |  |
| --- | --- | --- | --- |
| 作者姓名： | 付萌 | | |
| 指导教师： | 邓庆绪 教授 | | |
|  | 东北大学计算机科学与工程学院 | | |
| 申请学位级别： | 硕士 | 学科类别： | 工学 |
| 学科专业名称： | 计算机软件与理论 | | |
| 论文提交日期： | 2018年12月 | 论文答辩日期： | 2018年12月 |
| 学位授予日期： | 2019年1月 | 答辩委员会主席： |  |
| 评阅人： |  | | |

东 北 大 学

2016年12月

##### A Thesis in Computer Software and Theory

**The Research on Redestrian Dead Reckoning for Indoor Localization based on Deep Learning**

By Fu Meng

Supervisor: Professor Deng Qingxu

**Northeastern University**

**Decemeber 2018**

独创性声明

本人声明，所呈交的学位论文是在导师的指导下完成的。论文中取得的研究成果除加以标注和致谢的地方外，不包含其他人己经发表或撰写过的研究成果，也不包括本人为获得其他学位而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：

日 期：

学位论文版权使用授权书

本学位论文作者和指导教师完全了解东北大学有关保留、使用学位论文的规定：即学校有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人同意东北大学可以将学位论文的全部或部分内容编入有关数据库进行检索、交流。

作者和导师同意网上交流的时间为作者获得学位后：

半年 □ 一年□ 一年半□ 两年□

学位论文作者签名： 导师签名：

签字日期： 签字日期：

摘要

目前因GPS(Global Positioning System)、北斗等室外定位技术因为室内信号较弱、室内空间复杂等问题，在室内定位效果偏差很大，针对室内定位的技术例红外线定位、超声波定位、射频识别定位等对基础设施的依赖很强，灵活性均不足，很难应对极端条件下的室内定位需求。行人航迹推算是一种基于惯性传感器信息对行人的行动轨迹进行推算的技术，对基础设施的依赖很小，能够在极端环境下体现出更好的鲁棒性。

本文对行人航迹推算技术的各个步骤进行了研究，细化该技术的流程，包括数据收集、数据滤波、主轴选择、判步、滤步、步长估计、方向估计、高度移动、位置矫正以及额外针对小车移动的位置计算，并对流程中每一个部分进行了研究，均尝试使用了不同的方法和参数，并对这些结果进行对比。其中在滤步、步长估计、高度移动中给出可使用深度学习进行判定的接口。本文尝试结合流程中各个部分，寻找定位效果、计算开销综合最佳的室内空间定位策略。

本文中提到的所有方法均已经实现，其中数据收集和数据标记实现在手机客户端，数据处理和存储、参数设定、定位计算、多种方法的对比实验等内容在实现在服务端，能够实时进行室内定位并对流程中各个方法进行对比，定位准确性良好。

**关键词：**室内空间定位；手机惯性传感器；步行者航位推算；深度学习

Abstract

In recent years,

**Keywords:** Indoor positioning;Smartphone inertial sensor;Pedestrian dead reckoning;Deep learning

目录

[独创性声明 I](#_Toc517961744)

[摘要 II](#_Toc517961745)

[Abstract III](#_Toc517961746)

[第 1 章 绪论 1](#_Toc517961747)

[1.1 研究背景与意义 1](#_Toc517961748)

[1.2 国内外研究现状 1](#_Toc517961749)

[1.3 本文主要研究内容 1](#_Toc517961750)

[第 2 章 数据收集与前期处理 3](#_Toc517961751)

[2.1 数据收集 3](#_Toc517961752)

[2.1.1 数据内容 3](#_Toc517961753)

[2.1.2 采集与传输 3](#_Toc517961754)

[2.2 前期处理 3](#_Toc517961755)

[2.2.1 数据滤波 3](#_Toc517961756)

[2.2.2 数据储存 4](#_Toc517961757)

[2.3 本章小结 4](#_Toc517961758)

[第 3 章 步数估计 5](#_Toc517961759)

[3.1 主轴选择 5](#_Toc517961760)

[3.1.1 主轴选择的意义 5](#_Toc517961761)

[3.1.2 主轴选择方法 5](#_Toc517961762)

[3.1.3 方法对比 5](#_Toc517961763)

[3.2 步数估计 5](#_Toc517961764)

[3.2.1 步数估计方法 5](#_Toc517961765)

[3.2.2 方法对比 5](#_Toc517961766)

[3.3 错步过滤 6](#_Toc517961767)

[3.3.1 错步过滤方法 6](#_Toc517961768)

[3.3.2 方法对比 6](#_Toc517961769)

[3.4 步数估计与错步过滤的整合 6](#_Toc517961770)

[第 4 章 步长估计 7](#_Toc517961771)

[4.1 步长估计算法分类 7](#_Toc517961772)

[4.2 固定数值步长估计 7](#_Toc517961773)

[4.3 基于人体规律的步长估计 7](#_Toc517961774)

[4.4 基于加速度的步长估计 7](#_Toc517961775)

[4.5 一般公式与公式族 8](#_Toc517961776)

[4.5.1 一般公式 8](#_Toc517961777)

[4.5.2 公式族 8](#_Toc517961778)

[4.5.3 分类方法 8](#_Toc517961779)

[4.6 步长估计方法对比 8](#_Toc517961780)

[4.7 本章小结 8](#_Toc517961781)

[第 5 章 方向判定 9](#_Toc517961782)

[5.1 方向判定方法 9](#_Toc517961783)

[5.2 方法对比 9](#_Toc517961784)

[5.3 本章小结 9](#_Toc517961785)

[第 6 章 纵向移动与位置校正 10](#_Toc517961786)

[6.1 纵向移动判定方法 10](#_Toc517961787)

[6.2 方法对比 10](#_Toc517961788)

[6.3 位置校正 10](#_Toc517961789)

[6.4 针对车辆的位置校正 10](#_Toc517961790)

[第 7 章 总结与展望 11](#_Toc517961791)

[参考文献 12](#_Toc517961792)

[致谢 13](#_Toc517961793)

[攻读硕士期间发表论文和参与项目 14](#_Toc517961794)

# 绪论

## 研究背景与意义

随着时代的飞速变迁，科学技术的快速发展，人们对信息服务的质量和效率的需求不断地提高，鲁棒性良好，精确性高的室内空间定位技术逐渐在科学研究中占有了一席重要之地。在室外，目前随着GPS和北斗定位系统的普及，人们可以享受到较为精确的定位服务。然而这些服务在室内却很难达到让用户满意的程度。其原因在于，室内空间中移动终端能够获得的信号微弱，室内空间的环境较为复杂，由此产生的误差值相对于室内空间而言过大，定位的精确性很低。

为了解决这些问题，目前有一些专门针对室内空间定位的技术，如红外线定位技术、超声波定位技术、射频设别定位技术、蓝牙室内定位技术、Wifi室内定位技术、ZigBee室内定位技术、超宽带室内定位技术等。这些技术能够在室内为用户提供较为精确的定位服务。然而这些技术都需要事先在室内空间建设基础设施或者采集存储相关的位置信息以用于定位。也正因为这样的机制，这些定位系统相对较为脆弱，一旦基础设施受到干扰甚至毁坏，定位的精度会大幅下降甚至难以定位，而这样得情况又往往正是需要有定位系统支持的时候。如在一些极端条件或灾害现场，依赖基础设施的定位技术很难为用户提供准确而稳定的定位服务。

行人航迹推算技术，是利用惯性传感器的信息对行人的移动进行推断的技术，其思想在于惯性导航，其核心定位功能几乎不会依赖基础设施的搭建，鲁棒性姮好，并且也可以满足一些极端环境下的定位需求。目前，手机中已经集成诸如加速计、陀螺仪、磁力计等传感器，可以利用这些传感器的数据来进行行人航迹推算，从而达到室内空间定位的目的。

## 国内外研究现状

图是一种现实生活中最普遍的数据模型。显著性子图挖掘是对以图模式存储的数据进行进一步处理的基础。这项研究成果可以广泛用于各个领域。例如，在药物研发过程中，化学家们在

## 本文主要研究内容

本文针对论文组织结构

本文后序章节采用如下的组织结构方式：

1. 为“绪论”部分。该部分首先对文章的研究背景进行了简单的描述。介绍了研究目的。

## 论文章节结构

1. 为本文的“相关工作”部分。本章主要介绍了子图模式挖掘的概念，包括频繁子图挖掘，区分子图挖掘以及统计显著子图挖掘。介绍了这些子图特征提取方法的研究现状和意义，引出了如何高效地进行统计显著子图挖掘这一主要研究问题。然后介绍了多重假设检验校正方法的研究现状。以及对蚁群优化算法目前研究情况做了简单的介绍。
2. 为“基本概念和问题定义”部分。本章主要介绍了在本算法中提到的基本概念，主要包括费舍尔精确检验和置换检验的概念。然后用形式化的语言对本文要解决问题进行了定义。
3. 为“改进蚁群算法用于统计显著子图挖掘”部分。本章对算法的框架进行了详细的介绍。首先介绍了算法的整体流程，包括多重假设检验校正过程和子图挖掘过程。然后分别介绍了多重假设检验校正的主要过程和蚁群优化的数学模型以及具体的挖掘过程。其中详细分析了统计显著性与子图支持度的关系。将该性质利用在了挖掘过程中作为削减策略。同时在本章中提出了一系列的策略，通过在挖掘过程中调节参数控制蚁群优化方向和对挖掘结果进行后处理两类方法来保证挖掘结果的具有较少的信息冗余和较高的覆盖度。
4. 为“实验测试和结果分析”部分。本章通过在三种真实数据集，包括大规模数据和较小规模数据以及平衡数据和不平衡数据集上进行了实验。也利用了人工合成的大规模图数据进行了实验验证。通过大量实验证明了算法的准确性和高效性。通过对不同参数下算法的运行效率得到了不同参数对于算法效率的影响。通过与其他算法的时间空间效率比较验证了算法性能的提升。保证了挖掘结果的统计效力。并保证了分类的准确率。
5. 为“结束语”部分。对本文进行了总结与展望。

# 数据收集与前期处理

在使用惯性导航的方法进行室内空间定位时，判断何时前进了一步是一个非常重要的过程。步数的估计和判断的准确性对坐标估计的影响较大，与数据的储存、方向估计、步长估计和高度移动的估计都有重要影响，是室内空间定位系统中较为核心的一个步骤

## 数据收集

在数据挖掘领域图挖掘是一个重要的研究领域。这个领域的研究主要集中在从图数据集中挖掘特征子图。研究者们提出了一系列的特征子图度量标准，从频繁度，到区分度再到统计显著性。这些研究都有两个主要问题：（1）生成候选子图的有效机制；（2）怎么更好的处理候选子图，从候选子图中获得目标子模式。接下来详细介绍了这些方面的研究现状以及方法。

### 数据内容

数据挖掘的最基

掘的过程中去，一些研究者们将目光转向了区分子图挖掘过程中。

### 采集与传输

区分子图是指那些频繁子图与区分子图关系

区分子图是指那些频繁子图与区分子图关系

## 前期处理

### 数据滤波

那些频繁子图那些频繁子图那些频繁子图那些频繁子图那些频繁子图

### 数据储存

是指那些频繁

## 本章小结

本章主要对于论文相关的其他研究做了简要的介绍。首先介绍了子图挖掘的两个主要方向，并同时介绍了他们之间的关系。接下来，介绍了多重假设检验校正的错误度量标准以及校正的方法。最后，介绍了蚁群优化算法的基本原理和算法主要流程，并对其适用范围进行了分析。

# 步数估计

本章详细介绍了第四章算标。

## 主轴选择

### 主轴选择的意义

费舍尔精确检验[26]是一种用来分析列联表统计显著性的检验方法。在使用中通常将其用于较小。

### 主轴选择方法

置换检验[27]因其对总

计算过程中相当于*n*次挖掘统计显著子图，而挖掘过程又是一种非常耗时的操作。因此，这是一种时间复杂度很高的方法，于是本文采用了一系列的技术来减少其时间消耗。

### 方法对比

在实本发鬼地方第三方士大夫算法算法是打发啥

## 步数估计

### 步数估计方法

计算过程中相当于*n*次挖掘统计显著子图，而挖掘过程又是一种非常耗时的操作。因此，这是一种时间复杂度很高的方法，于是本文采用了一系列的技术来减少其时间消耗

### 方法对比

计算过程中相当于*n*次挖掘统计显著子图，而挖掘过程又是一种非常耗时的操作。因此，这是一种时间复杂度很高的方法，于是本文采用了一系列的技术来减少其时间消耗

## 错步过滤

计算过程中相当于*n*次挖掘统计显著子图，而挖掘过程又是一种非常耗时的操作。因此，这是一种时间复杂度很高的方法，于是本文采用了一系列的技术来减少其时间消耗

### 错步过滤方法

程中相当于*n*次挖掘统计显著子图，而挖掘过程又是一种非常耗时的操作。因此，这是一种时间复杂度很高的方法，于是本文采用了一系列的技术来减少其时间消耗

### 方法对比

计算过程中相当于*n*次挖掘统计显著子图，而挖掘过程又是一种非常耗时的操作。因此，这是一种时间复杂度很高的方法，于是本文采用

## 步数估计与错步过滤的整合

# 步长估计

在本章中介绍了利用改进的蚁群算法从大规模图数据中挖掘统计显著子图的算法框架。4.1节介绍了算法的总体结构，然后详细讨论各部分的具体内容。4.2节介绍了基于多重假设检验过程中族错误率的控制方法，对已有的方法的计算效率进行了改进。4.4节提出了针对于图数据库的蚁群优化搜索方法，通过演化计算的方法加快对子图的搜索过程。4.5节对本章内容进行了总结。

## 步长估计算法分类

本文提出了基于p值过滤和改进蚁群算法的统计显著子图挖掘算法PABSMiner。算法主要包括两部分。第一部分为多重假设检验校正部分。该部分中利用用户给定的族错误率阈值进控制FWER的统计显著性阈值校正

## 固定数值步长估计

在本节中，对蚁群优化模型进行了扩展。蚁群优化算法是一种模拟蚂蚁群体智慧来求解组合优化问题的一种方法。子图挖掘也可以看做是一种组合优化问题。一个

## 基于人体规律的步长估计

样性也是研究人员们重点着力的一个

为了更直观地描述算法的执行过程，在本节中给出了一个完整的例子来描述利用蚁群进行挖掘的整个过程。

## 基于加速度的步长估计

然后利用显著性值计算公式，同时能够计算得出该子图的正负样本支持度。根据该子图的正负样本支持度，得到该子图统计显著性。若其显著性小于预设统计显著性，则将其加入候选结果列表。

## 一般公式与公式族

本章对大规模图数据中统计显著子图挖掘问题进行了深入的研究与分析，提出了一种基于假设检验和演化计算的统计显著子图挖掘算法PABSMiner。该算法首先利用置换检验模拟出数据的零分布，获得了在一定族错误率下统计显著子图的显著性阈值，以及对应的支持度

### 一般公式

章对大规模图数据中统计显著子图挖掘问题进行了深入的研究与分析，提出了一种基于假设检验和演化计算的统计显著子图挖掘算法PABSMiner。该算法首先利用置换检验模拟出

### 公式族

基于假设检验和演化计算的统计显著子图挖掘算法PABSMiner。该算法首先利用置换检验模拟出

### 分类方法

基于假设检验和演化计算的统计显著子图挖掘算法PABSMiner。该算法首先利用置换检验模拟出

## 步长估计方法对比

图数据中统计显著子图挖掘问题进行了深入的研究与分析，提出了一种基于假设检验和演化计算的统计显著子图挖掘算法PABSMiner。该算法首先利用置换检验模拟出数据的零分布，

## 本章小结

本章对大规模图数据中统计显著子图挖掘问题进行了深入的研究与分析，提出了一种基于假设检验和演化计算的统计显著子图挖掘算法PABSMiner。该算法首先利用置换检验模拟出数据的零分布，获得

# 方向判定

本章主要从算法的效率，子图的显著性和图分类的准确性对提出的基于多重假设检验校正的统计显著子图挖掘算法进行分析和验证，并对比了该算法与过去算法性能的比较。对比其在准确率和时间效率方面的区别。通过对人工合成数据集和真实数据集的实验，证明了用于大规模图数据分类的统计显著子图挖掘算法的高效性和有效性，并且进行了分类准确性的验证。

## 方向判定方法

时间增长。随着图模式尺寸的增加，算法运行时间增长。同时，空间使用效率基本与现有算法持平。并且由于减少了搜索空间，避免了生成候选子图过程，节省了较多的

## 方法对比

在本节中从两个角度研究了子图挖掘结果的统计效力。一是从挖掘结果的经验

## 本章小结

本章通过实验验证了基于蚁群优化算法的统计显著子图挖掘算法的高效性，验证了基于多重假设检验校正的统计显著子图挖掘算法的有效性，并检验了算法的挖掘结果的可靠性和对于分类的准确性。通过在真实数据集和人工数据集上运行算法所得到的结果，表明将演化计算引入统计显著子图挖掘过程中，可以在提高算法效率的同时保证挖掘结果依然很好。

# 纵向移动与位置校正

## 纵向移动判定方法

在图分类研究中，图分类依据从频繁子图到区分子图，然后逐渐发展到统计显著子图。这些方法不断的改进对子图区分能力的评价标准，从而为用户决策提供具有价值的信息。近年来，随着大规模数据的不断发展，

## 方法对比

复杂数据结构的不断产生，对于图数据挖掘的需求也在日益旺盛，关于图数据挖掘的区分模式的研究热度逐渐增加。然而，由于数据量的急剧增加，图数据结构的复杂性导致现有算法的时间性能不足。其数目庞大的子结构使得其上的挖掘工作变得异常复杂。因此，对于图数据上区分模式挖掘的研究还亟待发展。

## 位置校正

本文就大规模图数据中统计显著子图挖掘过程中的时间效率问题和挖掘结果的假阳性比例问题进行了深入的研究。提出了基于蚁群优化算法的统计显著子图挖掘算法。与以往统计显著子图挖掘的算法不同，本文采用了基于演化计算的搜索框架，有效地减少了候选子图生成过多，搜索冗余过的等问题。此外，为了进一步避免挖掘过程中遇到的计算瓶颈，在计算显著性的时候运用了两种p值索引，以及一系列的削减策略，有效的减少了重复计算。

## 针对车辆的位置校正

本文就大规模图数据中统计显著子图挖掘过程中的时间效率问题和挖掘结果的假阳性比例问题进行了深入的研究。提出了基于蚁群优化算法的统计显著子图挖掘算法。与以往统计显著子图挖掘的算法不同，本文采用了基于演化计算的搜索框架，有效地减少了候选子图生成过多，搜索冗余过的等问题。此外，为了进一步避免挖掘过程中遇到的计算瓶颈，在计算显著性的时候运用了两种p值索引，以及一系列的削减策略，有效的减少了

# 总结与展望

放虎归山东方闪电存在可供福超能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不能力者开发就好像不

参考文献

1. Inokuchi A, Washio T, Motoda H. An Apriori-Based Algorithm for Mining Frequent Substructures from Graph Data[A].Proceedings of the 4th European Conference on Principles of Data Mining and Knowledge Discovery[C]. Springer-Verlag, 2010. 13-23.
2. He H, Singh A K. GraphRank: Statistical Modeling and Mining of Significant Subgraphs in the Feature Space[A]. international conference on data mining[C]. 2006. 885-890.
3. Chen M, Gao X, Li H, et al. An efficient parallel FP-Growth algorithm[A]. cyber-enabled distributed computing and knowledge discovery[C].2009. 283-286.
4. Thoma M, Cheng H, Gretton A, et al. Discriminative frequent subgraph mining with optimality guarantees[J]. Statistical Analysis and Data Mining, 2010, 3(5): 302-318.
5. (美)Anand Rajaraman, Jeffrey David Ullman著.大数据:互联网大规模数据挖掘与分布式处理[M].王斌译.人民邮电出版社, 2012.
6. 裴艳波. 多重假设检验问题中关于三种错误测度-FWER,FDR和pFDR的讨论[D]. 长春：东北师范大学,2005.

致谢

感谢大吸神教，让我长生不老！

攻读硕士期间发表论文和参与项目

攻读硕士期间参与项目：

1.国家自然科学基金重点项目: 集群环境下基于内存的高性能数据管理与分析(61332006)