

学校编码: 10384

分类号\_\_\_\_\_密级\_\_\_\_\_

学号: 24320121152274

UDC\_\_\_\_\_

厦 门 大 学

硕 士 学 位 论 文

基于位置社交网络潜在好友推荐算法研究

Research on Potential Friend Recommendation Algorithm in  
Location-Based Social Network

李 翔

指导教师姓名: 林 坤 辉 教 授

专 业 名 称: 计 算 机 软 件 与 理 论

论文提交日期: 2015 年 3 月

论文答辩日期: 2015 年 5 月

学位授予日期: 年 月

指 导 教 师: \_\_\_\_\_

答 辩 委 员 会 主 席: \_\_\_\_\_

2015 年 3 月

## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为( )课题(组)的研究成果,获得( )课题(组)经费或实验室的资助,在( )实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

# 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1.经厦门大学保密委员会审查核定的保密学位论文，  
于 年 月 日解密，解密后适用上述授权。

2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

## 摘 要

伴随着移动互联网技术与地理定位技术的崛起，基于位置的服务迅速地渗透到互联网的各类网站和应用中。其中，融合基于位置的服务与传统社交网络结构的位置社交网络发展迅猛，许多结合位置服务的社交类应用受到了许多互联网用户的青睐。用户在位置社交网络上分享自己的位置信息，在自己喜欢的地点进行签到，发表包含位置信息的状态和微博，通过关注周围的人寻找与自己有相同兴趣爱好的人。

随着位置社交网络应用中用户数据和地理位置数据的迅速膨胀，用户想要快速、准确地发现自己需要的信息变得越来越困难，基于位置社交网络的推荐系统开始出现。研究者们致力于根据用户数据和地理位置数据中包含的信息，分析用户的喜好，并针对用户的喜好对用户进行个性化的推荐，包括潜在好友推荐、地点推荐和活动推荐。但是，由于数据稀疏性和敏感性等因素，位置社交网络的个性化推荐面临许多的问题。

本文针对位置社交网络中潜在好友的推荐展开研究。介绍了基于位置社交网络的基本概念和相关推荐算法。同时，针对位置社交网络好友推荐中社交影响和位置影响这两个主要因素，介绍了三种相关的潜在好友推荐算法：基于共同好友比重的好友推荐、基于用户协同过滤的好友推荐和基于习惯签到地点的好友推荐。另外，针对位置社交网络用户签到数据稀疏的问题，运用聚类方法对用户签到矩阵进行降维处理，提高用户相似性计算效果。最后，提出了一个线性框架融合了基于社交影响和位置影响的好友推荐算法，力求提高位置社交网络好友推荐的效果。

本文通过改进算法与传统算法的实验对比，验证了改进算法的可行性。实验结果表明，改进后的算法能够有效提高好友推荐的效果，并且在数据稀疏的情况下，算法的稳定性优于传统推荐算法。

**关键词：** 好友推荐；位置社交网络；聚类算法；数据稀疏

厦门大学博硕士学位论文摘要库

## Abstract

With the rise of mobile Internet technology and geograohic positioning technology, location-based services rapidly penetrate into various types of Internet sites and applications. Among them, the location-based social networks which contained location-based services and traditional social network structure developed rapidly. A lot of social networking applications combined with location-based services have attracted many Internet users. Users shared their location information on the location-based social network, checked in the location which they were interested, published state and microblog contained location information, and searched the people who have the same interests by focusing on the people around.

With the rapid expansion of user data and geographic data on the location-based social network, users want to quickly and accurately find the information which they need become more and more difficult. Location-based social network recommendation system began to emerge. The researchers committed to analyze user preferences according to the information contained in the user data and geographic data. And for users personalized recommendations through the users' preferences, including potential friend recommendation, location recommendation and events recommendation.

This dissertation studies the potential friend recommendation in the location-based social network. First of all, this dissertaion introduces the basic concepts of the location-based social network and related recommendation algorithms. Secondly, for the social influence and location influence which are the two main factors to the potential friend recommendation in the location-based social network, this dissertation introduces three friend recommendation algorithm: friend recommendation based on the proportion of common friends, friend recommendation based on user-based collaborative filtering and friend recommendation based on normal check-in location. In addition, for the problem of the user check-in sparse data in the location-based social network, this dissertation proposes the method of reduce

the check-in matrix dimension by clustering algorithms to improve the user similarity computing. Finally, this dissertation proposes a potential friend recommendation algorithm combined the social influence and the location influence, in order to improve the performance of friend recommendation.

In this dissertation, we compared the improved algorithm with the traditional algorithm through comparison experiments and verify the feasibility of the improved algorithm. The experimental results show that, the improved algorithm can effectively improve the effect of friend recommendation. And in the case of sparse data, the stability of the algorithm is superior to the traditional recommender algorithms.

**Key words:** Friend Recommendation; Location-Based Social Network; Clustering Algorithm; Sparse Data



## 目 录

<b>第一章 绪论</b> .....	<b>1</b>
1.1 研究背景与意义 .....	1
1.2 研究现状与存在的问题 .....	2
1.2.1 好友推荐算法研究现状.....	3
1.2.2 存在的问题.....	3
1.3 论文研究内容 .....	4
1.4 论文组织结构 .....	5
<b>第二章 相关理论介绍</b> .....	<b>7</b>
2.1 基于位置的社交网络 .....	7
2.1.1 在线社交网络.....	7
2.1.2 基于位置的服务.....	8
2.1.3 基于位置的社交网络.....	9
2.2 个性化推荐算法 .....	10
2.2.1 基于内容的推荐.....	10
2.2.2 协同过滤推荐.....	11
2.3 相关聚类算法 .....	12
2.3.1 K-means 算法 .....	12
2.3.2 Canopy 算法 .....	14
2.3.3 K-medoids 算法.....	15
2.4 本章小结 .....	16
<b>第三章 个性化好友推荐算法研究</b> .....	<b>17</b>
3.1 问题描述 .....	17
3.2 数据表示 .....	18
3.2.1 好友关系表示.....	18
3.2.2 签到信息表示.....	20
3.3 基于社交影响的好友推荐 .....	21
3.4 基于位置影响的好友推荐 .....	23

3.4.1 基于用户协同过滤的好友推荐.....	23
3.4.2 基于习惯签到地点的好友推荐.....	24
3.5 本章小结 .....	26
<b>第四章 改进的个性化好友推荐算法 .....</b>	<b>27</b>
4.1 数据稀疏问题 .....	27
4.2 数据稀疏问题解决策略 .....	28
4.2.1 改进的 K-means 算法 .....	29
4.2.2 Canopy 算法 .....	33
4.2.3 K-medoids 算法.....	34
4.3 改进的个性化好友推荐算法 .....	36
4.4 本章小结 .....	38
<b>第五章 实验设计与结果分析 .....</b>	<b>40</b>
5.1 实验数据集 .....	40
5.2 实验环境及实验的算法 .....	44
5.2.1 实验环境.....	44
5.2.2 实验中采用的算法.....	44
5.3 评测方法 .....	45
5.3.1 评价指标.....	45
5.3.2 验证流程.....	46
5.4 实验结果与分析 .....	47
5.4.1 参数调优.....	48
5.4.2 实验结果分析.....	48
5.5 本章小结 .....	52
<b>第六章 总结与展望 .....</b>	<b>53</b>
6.1 总结.....	53
6.2 展望.....	54
<b>参考文献.....</b>	<b>55</b>
<b>攻读硕士学位期间发表的论文 .....</b>	<b>59</b>

致 谢..... 60

厦门大学博硕士论文摘要库

厦门大学博硕士学位论文摘要库

## Contents

<b>Chapter 1 Introduction.....</b>	<b>1</b>
<b>1.1 Background and Significance of Research .....</b>	<b>1</b>
<b>1.2 Research Status and Problems.....</b>	<b>2</b>
1.2.1 Research Status of Friend Recommendation Algorithm.....	3
1.2.2 The Problems .....	3
<b>1.3 Contents of this Dissertation .....</b>	<b>4</b>
<b>1.4 Structure of this Dissertation .....</b>	<b>5</b>
<b>Chapter 2 Related Theory Introduction .....</b>	<b>7</b>
<b>2.1 Location-Based Social Network.....</b>	<b>7</b>
2.1.1 Online Social Network.....	7
2.1.2 Location-Based Services.....	8
2.1.3 Location-Based Social Network .....	9
<b>2.2 Personalized Recommendation Algorithm .....</b>	<b>10</b>
2.2.1 Content-Based Recommendation.....	10
2.2.2 Collaborative Filtering Recommendation.....	11
<b>2.3 Related Clustering Algorithm.....</b>	<b>12</b>
2.3.1 K-means Algorithm.....	12
2.3.2 Canopy Algorithm.....	14
2.3.3 K-medoids Algorithm .....	15
<b>2.4 Summary.....</b>	<b>16</b>
<b>Chapter 3 Research on Personalized Friend Recommendation</b>	
<b>Algorithm.....</b>	<b>17</b>
<b>3.1 Problem Description .....</b>	<b>17</b>
<b>3.2 Data Representation .....</b>	<b>18</b>
3.2.1 Friend Relationship Representation.....	18
3.2.2 Check-in Information Representation.....	20
<b>3.3 Friend Recommendation Based on Social Influence .....</b>	<b>21</b>

---

<b>3.4 Friend Recommendation Based on Location Influence .....</b>	<b>23</b>
3.4.1 Friend Recommendation Based on User-Based Collaborative Filtering .....	23
3.4.2 Friend Recommendation Based on Normal Check-in Location .....	24
<b>3.5 Summary.....</b>	<b>26</b>
<b>Chapter 4 Improved Personalized Friend Recommendation Algorithm.....</b>	<b>27</b>
<b>4.1 Sparse Data Problem .....</b>	<b>27</b>
<b>4.2 Solving Strategies of Sparse Data Problem .....</b>	<b>28</b>
4.2.1 Improved K-means Algorithm .....	29
4.2.2 Canopy Algorithm.....	33
4.2.3 K-medoids Algorithm .....	34
<b>4.3 Improved Personalized Friend Recommendation Algorithm.....</b>	<b>36</b>
<b>4.4 Summary.....</b>	<b>38</b>
<b>Chapter 5 Experimental Design and Result Analysis.....</b>	<b>40</b>
<b>5.1 Experimental Dataset .....</b>	<b>40</b>
<b>5.2 Experimental Environment and Algorithm.....</b>	<b>44</b>
5.2.1 Experimental Environment .....	44
5.2.2 Algorithm Used in the Experiment .....	44
<b>5.3 Evaluation Method.....</b>	<b>45</b>
5.3.1 Evaluation Measure .....	45
5.3.2 Algorithm Used in the Experiment .....	46
<b>5.4 Experimental Results and Analysis .....</b>	<b>47</b>
5.4.1 Evaluation Measure .....	48
5.4.2 Validation Process .....	48
<b>5.5 Summary.....</b>	<b>52</b>
<b>Chapter 6 Conclusions and Prospect .....</b>	<b>53</b>
<b>6.1 Conclusions.....</b>	<b>53</b>

<b>6.2 Prospect</b> .....	<b>54</b>
<b>References</b> .....	<b>55</b>
<b>Publications</b> .....	<b>59</b>
<b>Acknowledgements</b> .....	<b>60</b>

厦门大学博硕士学位论文摘要库

厦门大学博硕士学位论文摘要库



Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”.

Fulltexts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.