# Recommending friends and locations based on individual location history.读书笔记

## 背景：

134. Yu Zheng, Lizhu Zhang, Zhengxin Ma, Xing Xie, and Wei-Ying Ma. Recommending friends and locations based on individual location history. ACM Transactions on the Web (TWEB), 5(1):5, 2011.



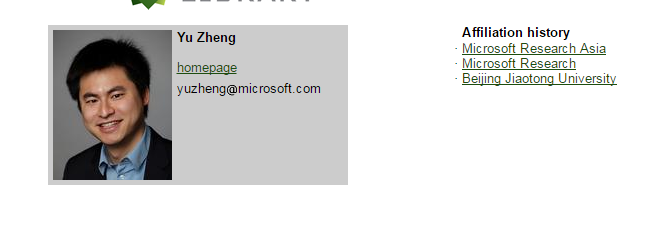
文章链接：

<http://cs.joensuu.fi/pages/franti/lami/papers/Recommending%20Friends%20and%20Locations%20Basedon%20Individual%20Location%20History.pdf>

### 作者

这位第一作者15年还在研究LBSN方向，这里面有他的一些近作。持续跟进中.

<http://dblp.uni-trier.de/pers/hd/z/Zheng:Yu>



### Recommendations in Location-based Social Networks: A Survey综述对文章的描述

进一步延伸通过考虑不同的空间粒度的位置序列的用户相似性度量框架。作者提出了一种新的序列匹配算法，该算法划分的位置序列，并认为受欢迎的访问位置分别。该框架，称为分层的基于图形的相似性度量（HGSM），提出了每一个人的位置历史和每个用户之间的相似性度量。这种相似性是基于用户位置的历史和使用的是三个因素来衡量，1）用户动作的共同序列，即不再类似的探访用户共享序列相似的用户，2）位置基准的普及，例如两个用户访问一个位置不走可能是比别人更多的相关访问一个受欢迎的位置，3）地理空间的层次，即细的地理区域由两个人共享粒度。

## 文章结构：

### 背景

略写

geographical information systems (GIS)

hierarchical-graph-based similarity measurement (HGSM),基于分层图的相似度量

### 介绍

假设一种框架HGSM考虑三因素：

用户移动序列属性

地理空间分层属性

不同地点的流行属性

使用HGSM评判用户的相似性

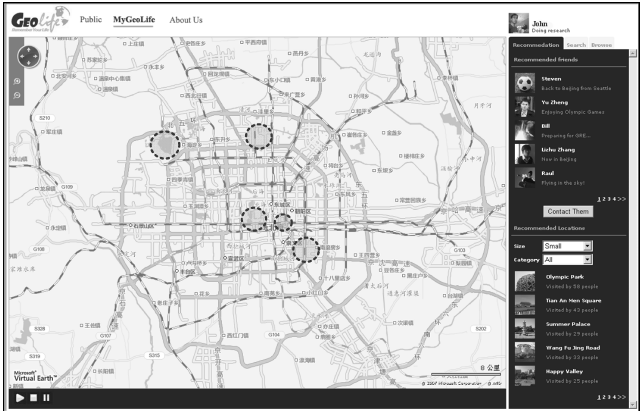
数据使用的是全球范围的75个人的一年的数据定位数据

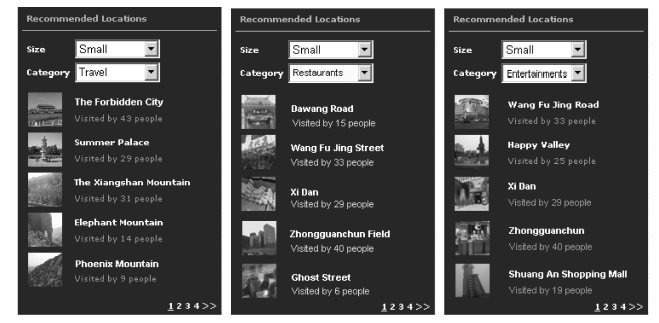
In Section 2, we first present the user interface of the system. Later, the architecture of our recommender system, which consists of three parts, location history representation, user similarity mining, and CF-based location recommendation, is introduced. In Section 3, we detail the processes of mining the similarity between users based on their location histories. Section 4 describes the CF-based location recommender, and Section 5 reports major experimental results. After giving a survey on the related works in Section 6, we draw our conclusions in Section 7, and propose the future work we attempt to conduct in Section 8.

1. 第二部分 用户系统界面，推荐系统，位置历史，用户相似性挖掘，和基于位置的推荐。
2. 第三章用户历史轨迹的挖掘相似性的进展，
3. 第四章CF位置推荐，
4. 第五章主体之言结果反馈。第六章相关的研究工作，
5. 第七章结论总结，
6. 第八章是将来的工作方向

### 回顾推荐系统

1. 用户的系统界面

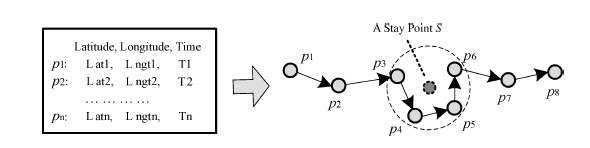




1. 以往刊物和这篇的区别

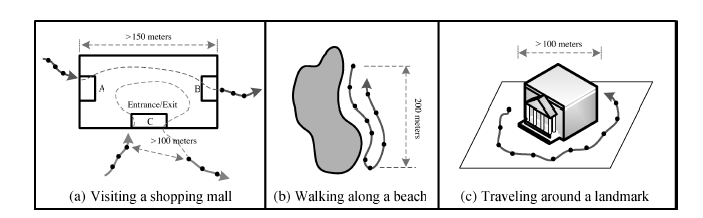
a collaborative filtering (CF) model 协同过滤比较模型

1. 初步



a stay point can be constructed by points *{p*3, *p*4, *p*5, *p*6*}*.

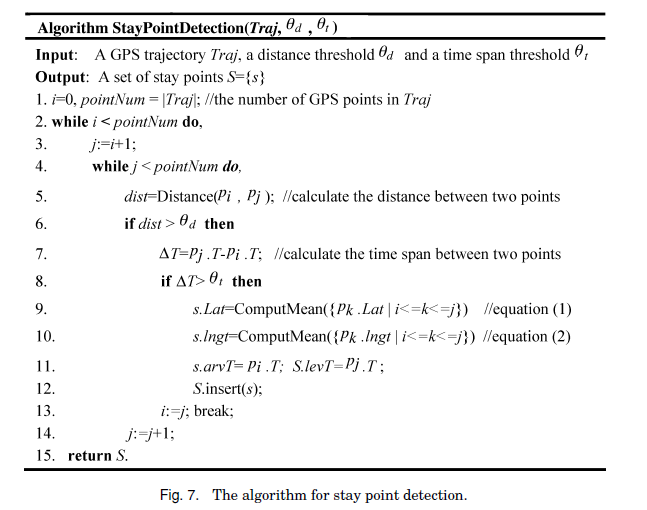
作者分析了是因为进入了建筑导致没有了信号，出来之后又重新连接上了



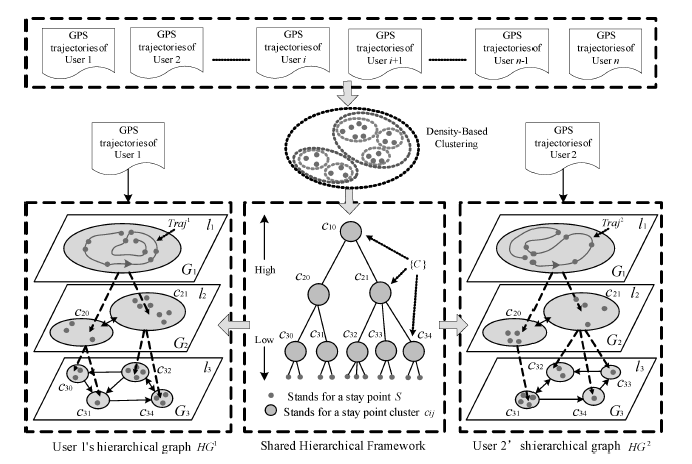
这些都是对轨迹的一些描述

访问shopping mall，地标建筑啊游玩的湖啊。

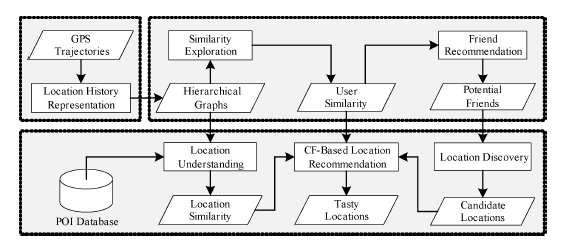
Staypoint的探测思想：



然而，不同的人的位置历史是不一致的，因为不同的人的停留点是不相同的。同时，它是主观的直接衡量类似留下的点，基于它们之间的位置距离。而且，用户相似度不是二进制值。我们的目标是要找出相关的个人是如何与其他人相比，然后为每个用户，根据相似性排名。



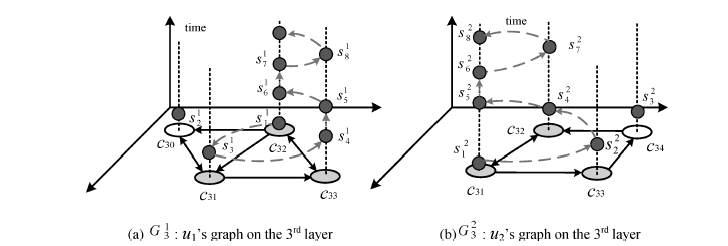
**Architecture of the System建筑系统**

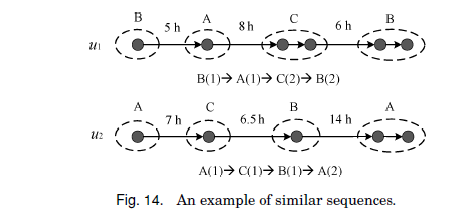


First, based on individual GPS trajectories, we build a hierarchical graph for each user using the method we proposed in Section 2.2. This hierarchical graph is capable of modeling the user’s location histories on different geospatial scales. Second, given two users’ hierarchical graphs, we are able to match the similar sequences shared by them on each layer of the hierarchy and calculate a similarity score for them. Later, a group of people, called potential friends, with relatively high scores will be retrieved for a particular individual. Third, using a POI database, we understand the profile of a geospatial region by exploring the categories of POIs located in the region. Such profiles enable us to detect the similarity between geospatial regions and recommend locations based on users’ diverse requests. At the same time, with the similarity between locations, we are able

* 1. 首先根据GPS轨迹建立分层图
  2. 然后根据两个用户的分层图来匹配相似模式，进行判分。
  3. 最后一组人，潜在用户，使用POI数据去探索定位区域，这些信息能用来探测地理区域和基于用户不同请求的推荐地理位置的相似性。同时我们能在位置之间找到相似性。

### 用户相似性探索



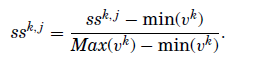


Using part of the two sequences (seq1 3 and seq2 3 ) depicted in Figure 13, Figure 14 further illustrates the definitions mentioned in this Section.

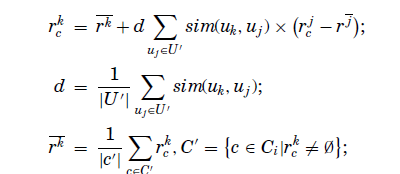
最后，对不同GPS轨迹，为用户提供一个公平的结果，我们把整体的相似性得分到他们的数据集规模的乘法。直观地说，在一个网络社区的人加入更容易积累更多的全球定位系统的轨迹比新的用户。如果我们不考虑数据的规模，更多类似的序列将从这些人的相对大的数据集检索。因此，这些高级用户可能会推荐给别人，虽然他们不是最完美的人选。

### 地理位置推荐

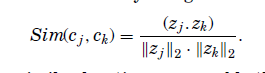
#### 地点发现



#### cf推理

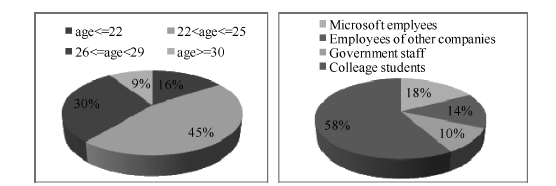


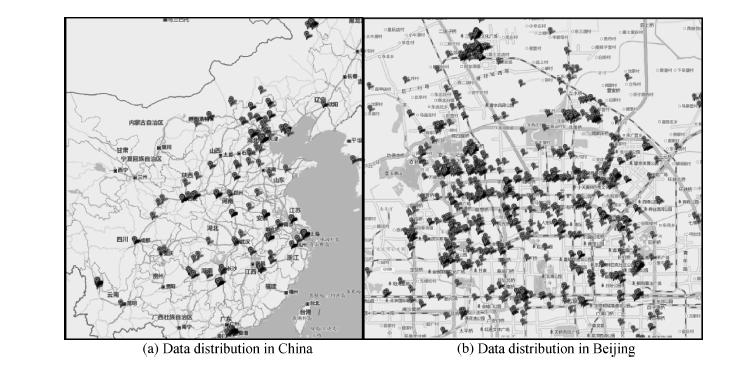
#### 位置理解

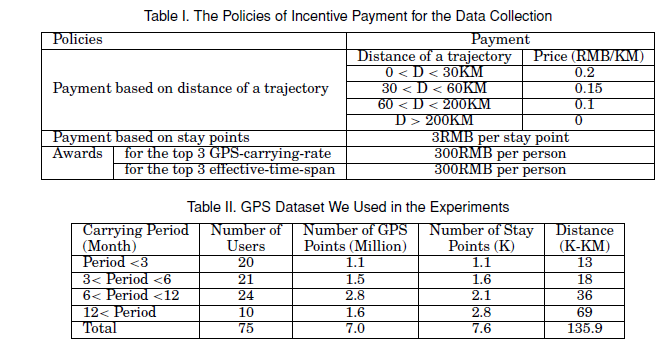


### 实验

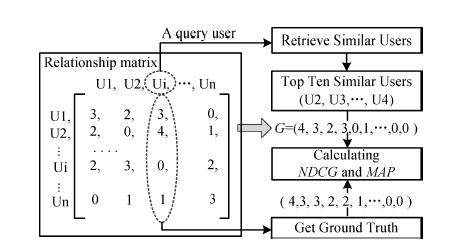
#### 实验设置





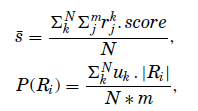


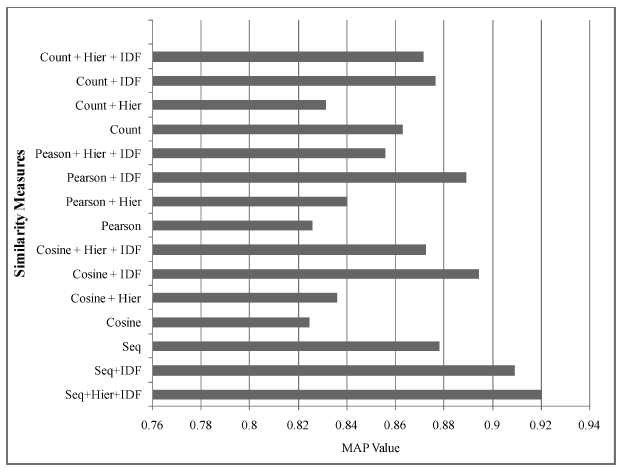
#### 评价趋近

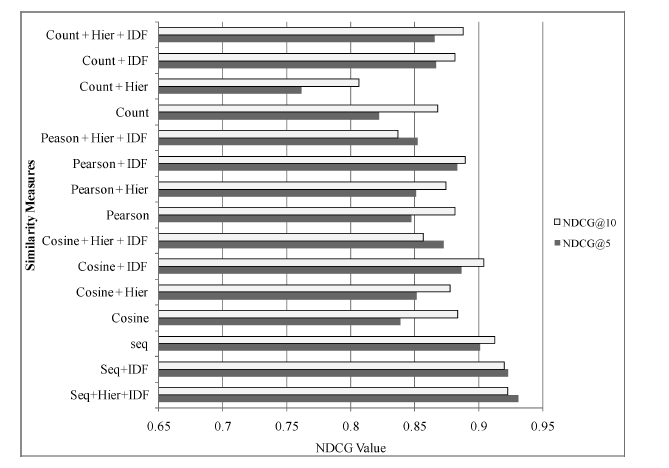


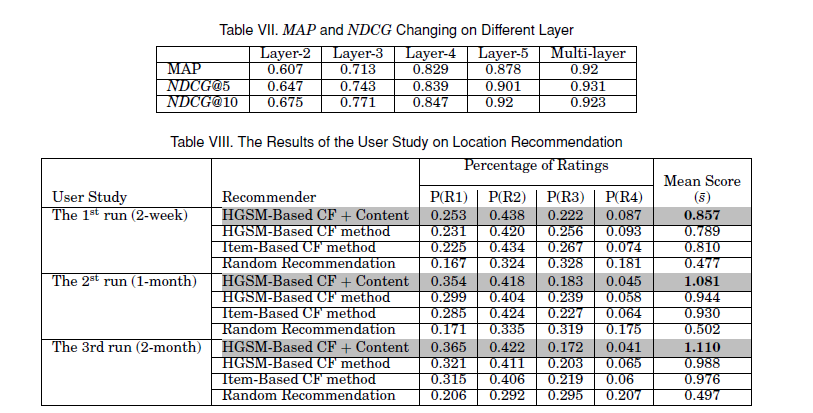
评价朋友推荐的框架

#### 结果

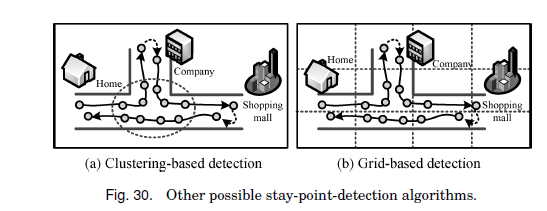




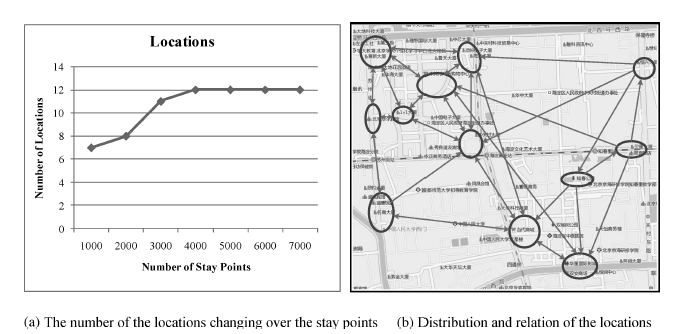




#### 讨论



如图30（b），网格划分的边界问题 方法也可能会错过重要的地方。换句话说，地理空间区域 购物中心可能被分成若干个网格，在每一个密度的全球定位系统 点不可能达到的条件，制定一个集群。



在实验中，我们观察到在一个固定的地理空间范围内发现的位置的数目不继续增加时，加入我们的系统中的用户的数目超过了一定值。如图所示，我们随机添加的停留点，从75个用户的位置记录步进的数据集，这将分层聚集成若干区域。其结果是，在3公里的地理空间范围内3公里，发现的位置的数目不增加任何更多的停留点的数目超过4000。说明了在给定的地理空间范围内的分布。同时，在用户的GPS轨迹两位置与一些指导线显示。2个位置之间的一个边缘意味着至少有5个用户在一个序列中的方向已前往这些位置。

### 相关工作

1. 挖掘定位历史
2. 通常的定位系统
3. 基于定位的推荐系统

### 结论

在这篇文章中，基于定位历史的推荐系统使用特定个人的访问在一个地理位置为位置的隐含评级和预测在他们的访问位置和其他特定用户的兴趣。在这个系统中，每一个用户都会推荐一组潜在的朋友，他们可能会有类似的旅游、体育、娱乐的爱好，以及可能与用户利益匹配的地理空间位置。因此，用户可以在一个社区组织一些社会活动，以最小的努力扩大他们的空间。了解用户和位置之间的相关性，使用用户生成的地理空间数据。相似性度量，HGSM，提出对该推荐模型的相似性。用户移动的序列属性，层次属性的地理空间和访问的位置，已被认为是在这种相似性度量。然后，结合一个基于内容的方法为基于用户的协同过滤算法，它利用HGSM用户相似性测度来判断一个用户在一个项目的评级。因此，在一定程度上，冷启动问题的推荐系统为用户提供了一个更好的位置推荐使用。对于朋友的推荐，HGSM超过相似计数基准方法，余弦相似度和皮尔森的相似性。此外，这三个功能表现出他们的优势，在测量用户之间的相似性。结合基于内容的方法基于HGSM CF取得了显着的改善。系统的有效性取决于用户的数据集的规模。更多的GPS轨迹从一个个体，我们的系统更容易推断出用户兴趣的准确。

### 将来的工作

以下三个方面：

1. 关于相似性度量，考虑更多用户的移动特性，例如类似序列的距离和用户在地理空间中的时间停留时间。同时，提高测量用户相似性的效率也是一个潜在的工作。
2. 关于朋友推荐的评价。设想一个新的方法来评估我们的相似性度量，研究用户行为后，他们建议与潜在的朋友，检查他们是否成为真正的朋友，并与更多的用户从不同的城市进行这种评价。
3. 关于位置推荐。进一步了解位置之间的相似性，提出了一种更先进的基于内容的方法。其次，尝试将基于内容的方法，在一个更复杂的方式协同过滤方法。研究CF方法的一些项目估计用户的兴趣水平。