

群智感知计算

Crowd and Participatory Sensing

杨铮



清華大學
Tsinghua University

个人简介

- 杨铮，清华大学讲师
- 在清华大学计算机系获得学士学位，在香港科技大学获得博士学位
- 发表论文40余篇，英文专著1部
- 主持自然科学基金面上项目1项，参与973、863、自然科学基金重大、重点、面上项目多项
- 2011年获得国家自然科学奖二等奖、2009年获得ACM学生研究竞赛提名奖



个人简介

- 自2009年至2013年，连续5年在计算机通信领域会议IEEE INFOCOM发表论文10篇。
- 在无线及移动计算领域的会议ACM MobiCom 2012发表论文1篇，是该会议历史上第一篇全部由中国大陆高校学者完成的论文，本人为第一作者。
- 合著关于网络定位技术的学术专著《Location, Localization, and Localizability》，由著名学术和教育出版社（Springer）出版。
- 目前担任ICDCS (2012, 2013)、INFOCOM (2013, 2014) 等国际学术会议的程序委员以及SCI索引期刊AHSWN的编辑。



提纲

- ▣ 基本思想
- ▣ 应用
 - ▣ Internet Crowdsourcing
 - ▣ Mobile Crowdsourcing
- ▣ 案例分析：城市交通感知与室内定位
- ▣ 未来研究方向

基本思想

- 在群智感知中，大量**普通用户**使用**移动设备**作为基本感知单元，通过**移动互联网进行协作**，实现感知任务分发与感知数据收集利用，最终完成大规模的、复杂的社会感知任务。



交通感知



热点发现



场景重现

相似概念

- 与群智感知相近的概念包括：



相似概念

- ▣ 与群智感知相近的概念包括：
 - ▣ 群体计算（Crowd Computing）
 - ▣ 参与式感知（Participatory Sensing）
 - ▣ 社群感知（Social Sensing）
 - ▣ 众包（Crowdsourcing）等等
- ▣ 它们都以**大量用户参与**为基础，可以说在理念上是一致的，但分别强调不同的**层次**和**方面**。

群智感知成为研究热点

- 无线网络与移动计算的几个著名学术会议中，许多论文都采纳了群智感知的思想。
 - ACM MobiCom
 - ACM MobiSys
 - IEEE INFOCOM
 - ACM SenSys等
- MobiSys 2011专门开辟一个会议单元收录相关论文，MobiCom 2012中至少有4篇论文涉及群智感知。

核心理念

- ▣ 有意识部署到无意识协作的转变
 - ▣ 有意识部署：传感器网络
 - ▣ 有意识部署与无意识协作结合：群智感知



用户参与方式

- 完成复杂感知任务的参与者无需是拥有**专业技能**的人士。
- 大量**草根用户**成为中坚力量，通过合理的协作来完成他们单独不可能完成的任务。



用户参与方式

▣ 主动参与

- ▣ 假设我们想知道某一时刻城市道路的拥堵情况
- ▣ 那么各地**广播电台交通台**的工作模式就是用户主动参与，依赖于行人、驾驶员通过**发短信**汇报路况

▣ 被动参与

- ▣ 利用手机中**自动的监测**人群的移动情况并汇报，就是用户被动参与。
- ▣ 用户无需为完成感知任务付出额外的精力和负担，用户参与热情高。



社会感知任务

- ▣ 群智感知中的感知任务是复杂的
 - ▣ 其感知对象也不仅仅是一般传感器所感知的温度、湿度、位置等物理属性，更不限于图像、声音等多媒体数据。
 - ▣ 如上文所说的交通感知等，以及其他社会事件、人群活动等，具有强烈社会属性的。



社会感知任务

- ▣ 社会感知任务的特点
 - ▣ 范围广
 - ▣ 规模大
 - ▣ 任务重
- ▣ 群智感知将大量草根用户拧成一股绳，形成随时随地、无孔不入、与人们生活密切相关的感知系统。

感知任务的分发与收集

- 要做到分工合理、负担合适、激励合情，让用户愿意参与，有所付出有所获得。
- 三个特点：



感知任务的分发与收集

▣ 众包（Crowdsourcing）

- ▣ 许多群智感知任务通过众包的方式完成感知任务的分发与收集利用。



感知任务的分发与收集

■ 众包（Crowdsourcing）

- 众包是《连线》（Wired）杂志2006年发明的一个专业术语，用来描述一种新的生产组织形式
- 即企业利用互联网将工作分配出去、发现创意或解决技术问题。通过互联网，这些企业可以利用大量用户的创意和能力。



提纲

- ▣ 基本思想
- ▣ 应用
 - ▣ Internet Crowdsourcing
 - ▣ Mobile Crowdsourcing
- ▣ 案例分析：城市交通感知与室内定位
- ▣ 未来研究方向

WIKIPEDIA

English

The Free Encyclopedia

3 816 000+ articles

Español

La enciclopedia libre

849 000+ artículos

日本語

フリー百科事典

781 000+ 記事

Русский

Свободная энциклопедия

798 000+ статей

Italiano

L'enciclopedia libera

867 000+ voci



Deutsch

Die freie Enzyklopädie

1 324 000+ Artikel

Français

L'encyclopédie libre

1 181 000+ articles

Polski

Wolna encyklopedia

848 000+ haseł

Português

A enciclopédia livre

706 000+ artigos

中文

自由的百科全書

388 000+ 條目



Internet-based public volunteer computing project for the **Search for Extra-Terrestrial Intelligence**, by leveraging **idle time** of computers.



- WHAT IS reCAPTCHA
- GET reCAPTCHA
- PROTECT YOUR EMAIL
- MY ACCOUNT
- RESOURCES: DOCS & PLUGINS

**reCAPTCHA IS A FREE
ANTI-BOT SERVICE THAT
HELPS DIGITIZE BOOKS.**

steamboat train, from New
this **morning** ran off the track
New-London. Four cars plunge

A screenshot of the reCAPTCHA interface. It shows a distorted image of the word 'morning' and a distorted image of the word 'upon'. Below the images is a text input field with the placeholder text 'Type the two words:'. To the right of the input field are icons for a refresh button, a volume icon, and a help icon. The reCAPTCHA logo is also visible in the bottom right corner of the interface.

→ **LEARN HOW reCAPTCHA WORKS**

USE reCAPTCHA ON YOUR SITE

- STRONG SECURITY**
- ACCESSIBLE TO BLIND USERS**
- 30+ MILLION SERVED DAILY**

NEW See how accurate reCAPTCHA is at
digitizing content!

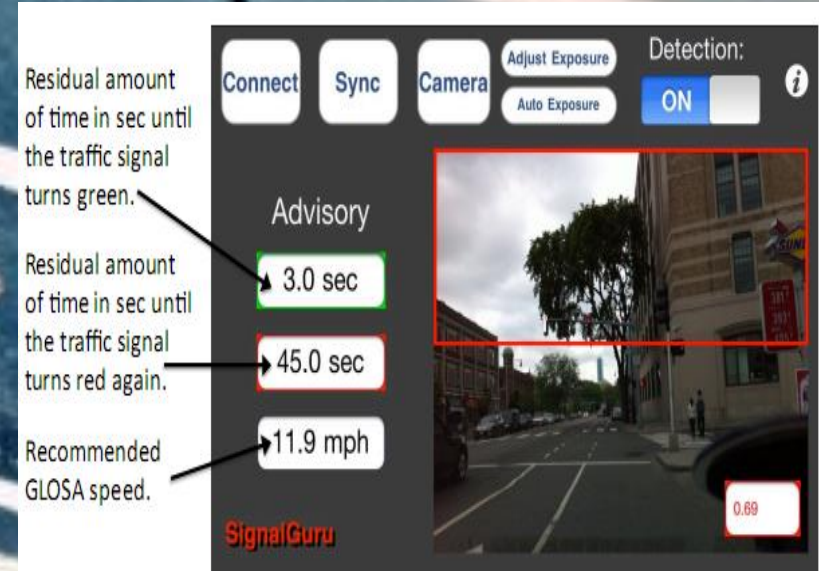
所谓巧妙的众包机制。。。。

巧妙的众包机制有种“**于无声处听惊雷**”的感觉，不知不觉中樯櫓灰飞烟灭大功告成。谁能想到探寻外星智慧以及数字化《纽约时报》类似的大工程竟能通过这样的方式进行？

巧妙的众包机制还有种“**明修栈道，暗度陈仓**”的意味，让用户在休闲、娱乐、享受、获得服务的过程中，完成具有其他目的的任务。

提纲

- ▣ 基本思想
- ▣ 应用
 - ▣ Internet Crowdsourcing
 - ▣ Mobile Crowdsourcing
- ▣ 案例分析：城市交通感知与室内定位
- ▣ 未来研究方向



SignalGuru:

Leveraging Mobile Phones for Collaborative Traffic Signal Schedule Advisory

- Emmanouil Koukoumidis, Li-Shiuan Peh, and Margaret Rose Martonosi, SignalGuru: leveraging mobile phones for collaborative traffic signal schedule advisory. MobiSys '11.

EasyTracker:

Leveraging Mobile Phones for arrival time prediction for buses



- James Biagioni, Tomas Gerlich, Timothy Merrifield, and Jakob Eriksson. 2011. EasyTracker: automatic transit tracking, mapping, and arrival time prediction using smartphones. SenSys '11.



Which App is widely accepted?

APPJOY for APPs
Rating

- Bo Yan and Guanling Chen, Appjoy: personalized mobile application discovery. MobiSys '11.



LiFS: Cutting off site survey
By using **new embedded sensors in a smartphone**:
accelerometer, gyroscope,
camera, GPS...



- Zheng Yang, Chenshu Wu, and Yunhao Liu, “Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention”, ACM MobiCom 2012, Istanbul, Turkey, August 22-26, 2012.



Building Tomography

Explore **building indoor structure** through user daily **activities** collected by smartphones



Replay of disaster scenes

Users witness an event in different angles. The collection of such **single-perspective** observation recovers the truth happenings of disasters.

提纲

- ▣ 基本思想
- ▣ 应用
 - ▣ Internet Crowdsourcing
 - ▣ Mobile Crowdsourcing
- ▣ 案例分析：城市交通感知与室内定位
- ▣ 未来研究方向

城市交通感知

- ▣ 公交车到站时间预测对于出行的乘客来说是很有价值的信息。



智能公交站牌



智能公交站牌

公交车到站时间预测

▣ 传统方法

- ▣ 公交公司提供了公交运行时刻表，但这些信息往往太过粗略，再加上当时当地交通状况的影响，人们很难得到准确的公交车到达时间。
- ▣ **EasyTracker**：2011年，美国伊利诺伊大学开发出一套基于智能手机的公交车到站预测系统，通过在公交车上嵌入智能手机，收集公交车运行的相关数据。然而不可避免的，需要公交公司的许可与配合，也会带来更高的部署成本。

公交车到站时间预测

▣ 群智感知思想的体现

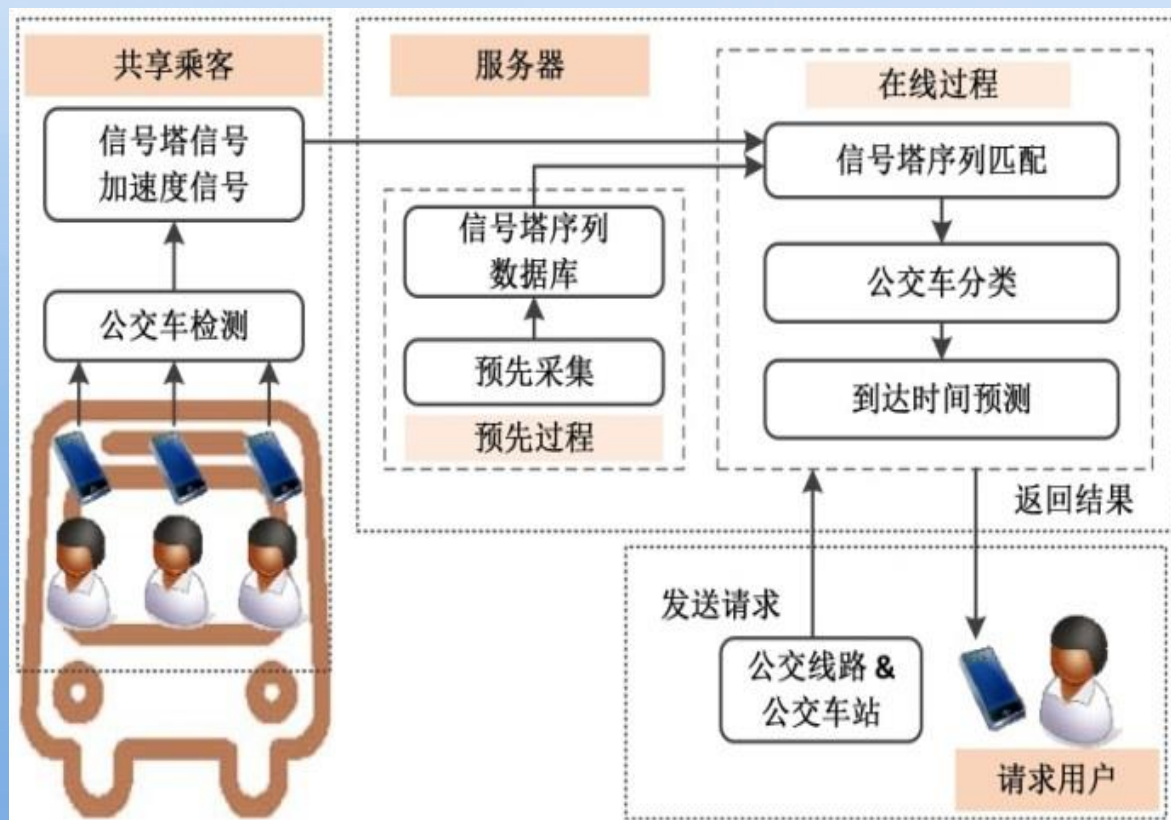
- ▣ 2012年，新加坡南洋理工大学利用群智感知思想，预测系统的**参与者正是经常乘坐公交车的乘客们**，整个系统不需要借助于公交公司的配合，乘客群体自己就可以合作地实现公交车到站时间预测。

乘客提供公交车信息

- 乘客是否在公交车上？
 - 检测刷卡的声音讯号
- 乘客在哪路公交车上？
 - 路线匹配
- 乘客的位置？
 - 利用GSM信号来定位公交车。相对于GPS定位，利用GSM几乎不会给手机带来额外的能量消耗



公交预测系统框架



- Pengfei Zhou, Yuanqing Zheng, and Mo Li. How Long to Wait?: Predicting Bus Arrival Time with Mobile Phone based Participatory Sensing. In Proceedings of ACM MobiSys, 2012.

城市交通感知

- 感知道路交通情况是一项复杂的任务。群智感知以其鲜明的特点契合城市交通感知的需求。

规模宏大

- 特别是大城市的交通状况，空间上连绵上百公里，牵扯几百万辆机动车和更多数目的行人。

时空关联

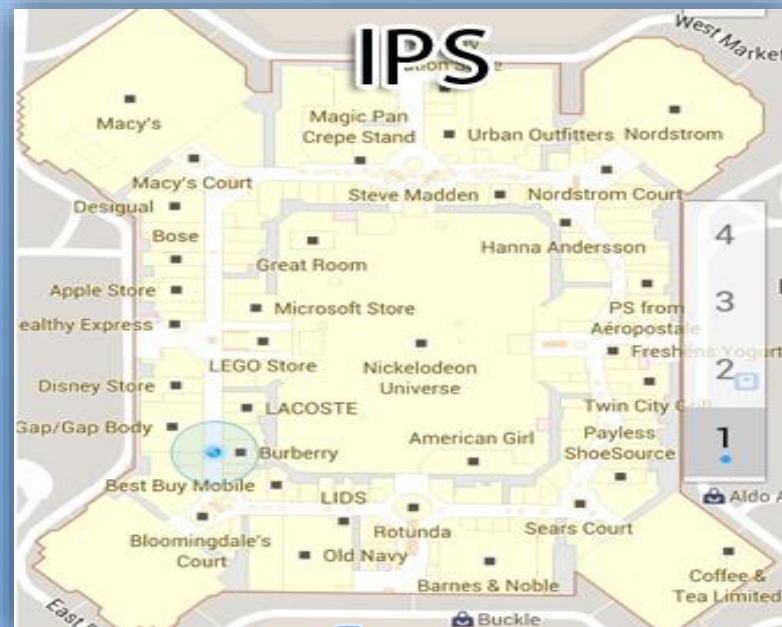
- 交通具有强烈的空间与时间关联性，牵一发而动全身。

因素众多

- 工作日，体育赛事，发改委提高油价的指令，更不用提北京“7.21”暴雨

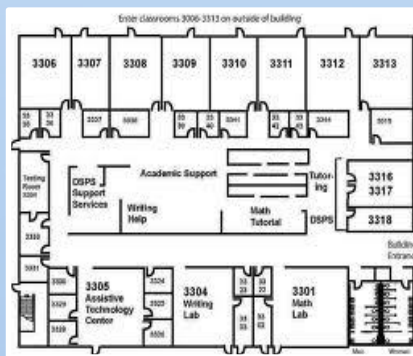
无线室内定位

- 室外定位：GPS处于领导地位
- 室内定位：基于信号指纹的定位是最普遍的。



信号指纹现场勘测

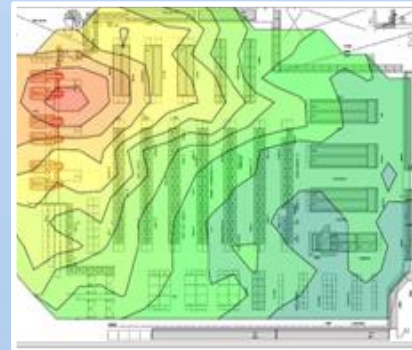
- 工程人员记录未知建筑物内部中每个位置上的无线信号特征。
- 弊端：耗费巨大人力和时间，不能适应环境动态性，限制应用范围。



Floor plan



Surveying



Radio map

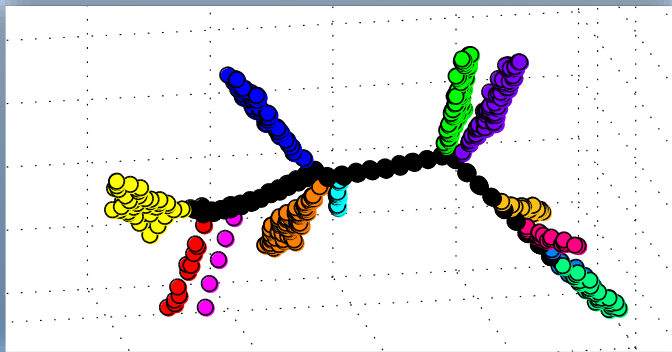
用户的移动性

让普通的用户完成现场勘测的任务

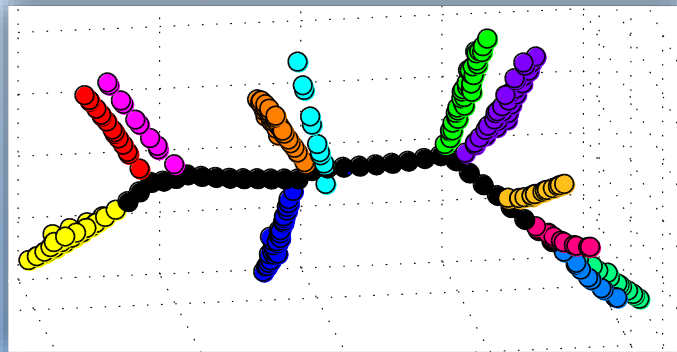
- 通过手机收集用户日常工作生活中的移动信息，利用用户移动路径在空间上关联起原本孤立的指纹，形成指纹空间。



避免现场勘测



利用手机采集的用户移动性信息构建**信号指纹空间**.



将普通的建筑室内结构图转化为空间**无压力的结构图**.

两者之间**空间上的相似性**使得他们能够完美的匹配在一起，因此每个**指纹都获得了对应的真实位置**。

这项工作以前只能通过现场勘测来完成。

普适的室内定位

- 基于群智感知的思想，实现了一个**无需依赖人工勘测、无特殊网络基础设施、无网络先验知识、无需用户主动参与**的普适室内定位系统
- 先期成果发表于INFOCOM 2012
- 后续成果发表于MobiCom 2012，是第一篇全部由中国大陆高校学者完成的MobiCom论文

- Chenshu Wu, Zheng Yang, Yunhao Liu, and Wei Xi, “WILL: Wireless Indoor Localization without Site Survey”, IEEE Transactions on Parallel and Distributed Systems (TPDS), Volume 24, Issue 4, Pages 839-848, April 2013.
- Zheng Yang, Chenshu Wu, and Yunhao Liu, “Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention”, ACM MobiCom 2012, Istanbul, Turkey, August 22-26, 2012.

提纲

- ▣ 基本思想
- ▣ 应用
 - ▣ Internet Crowdsourcing
 - ▣ Mobile Crowdsourcing
- ▣ 案例分析：城市交通感知与室内定位
- ▣ 未来研究方向

研究方向I

▣ 数据传输：“弱”联网

- ▣ 从**联网时间**来看，感知数据从产生到传输至目的地可能要经历很大的延迟，从几秒钟到几天都有可能，形成容迟网络（delay-tolerant network）。
- ▣ 从**网络接入形式**来说，也可以由直接变为间接，在没有网络基础设施的情况下，实现设备之间的多跳数据传输，形成自组织网络（ad-hoc network）。
- ▣ **联网通讯手段**由单一变为多样，移动设备可以包含多套无线数据通讯接口，例如GSM、WiFi、蓝牙等，研究人员除了应考虑可靠性、带宽、延迟等传统因素，还应考虑计费、流量、能耗、在多套通讯手段间进行选择、切换、或者并行利用等新情况

研究方向2

▣ 数据处理：大数据

- ▣ 多大的数据算是大数据？**TB**、**PB**已经不能满足胃口，**EB**正当年，**ZB**、**YB**也不是终点
- ▣ 据IDC报告，2011年全球信息总量达到**1.8ZB**
- ▣ 有些具有用户标注、有些没有，有些是结构化的（比如数值、符号）有些是非结构化的（比如图片、声音），有些时效性强有些时效性弱，有些价值密度高有些价值密度低，这都将给大数据处理技术带来关注的新焦点

研究方向3

▣ 数据质量

- ▣ 大量未经训练的用户作为基本感知单元会带来感知数据**不精确、不完整、不一致、不及时**等质量问题
- ▣ 对感知数据**去粗取精、去伪存真**，需要根据感知数据的质量设计利用方法，不同质量的感知数据都能**尽量利用又不致误读**
- ▣ Skyhook 的全球WiFi数据库

研究方向4

▣ 数据挖掘

- ▣ 原始感知数据在不同维度上刻画被感知的对象，需要经过不同层次的加工和精炼才能**形成从数据到信息再到知识的飞跃**。
- ▣ **多模态数据发掘**（multimodal data mining）：手机上的传感器包括加速度计、陀螺仪、指南针、GPS、麦克风、摄像头、以及各种无线信号如GSM、WiFi和蓝牙等。
- ▣ 通过传感器数据来分析用户的行为（行动坐卧走）或者用户所处的环境（办公室/会议室、室内/室外等）已成为**行为识别**（activity recognition）和**体域传感网**（body-area sensor network）的热点

研究方向5

▣ 让眼光更宽阔一些

- ▣ 群智感知为研究人类的社会行为提供了新的途径。比如邻近手机的发现可以记录人的社交活动，大量数据在匿名化之后可以作为社会学研究的基本资料。
- ▣ 其他还包括个体和人群的的运动信息、行为信息、社交信息、社会事件等等。
- ▣ 群智感知的数据具有广泛的社会学意义。“**有意识瞄准，无意识击发**”，谁知道这些数据将来能干什么呢！

总结

▣ 需求、现状、与未来

- ▣ 物联网发展到今天，对透彻感知的需求越来越强烈。
- ▣ 随着无线通信和传感器技术以及无线移动终端设备的爆炸式普及，群智感知计算走上移动计算的核心舞台是水到渠成。
- ▣ 在大力发展了十几年如何利用特定的有意识部署的传感器提供感知服务之后，物联网下一个成功的关键在于如何把无意识的提供“做好事”的模式融合进来——这正是“群智感知”。

参考文献

- Pengfei Zhou, Yuanqing Zheng, Mo Li, "How Long to Wait?: Predicting Bus Arrival Time with Mobile Phone based Participatory Sensing", In ACM MobiSys, Pages 379-392, Low Wood Bay, United Kingdom, June 2012.
- Chenshu Wu, Zheng Yang, Yunhao Liu, and Wei Xi, "WILL: Wireless Indoor Localization without Site Survey", IEEE Transactions on Parallel and Distributed Systems (TPDS), Volume 24, Issue 4, Pages 839-848, April 2013.
- Zheng Yang, Chenshu Wu, and Yunhao Liu, "Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention", ACM MobiCom 2012, Istanbul, Turkey, August 22-26, 2012.
- Yunhao Liu, Zheng Yang, Xiaoping Wang, and Lirong Jian, "Location, Localization, and Localizability". Journal of Computer Science and Technology, Volume 25, Issue 2, Pages 274-297, March 2010.
- 《群智感知计算》，刘云浩，计算机学会通讯

Any Questions?



杨铮
清华大学

hmilyyz@gmail.com