

# 一种基于背景声音识别的移动社会活动推荐系统

杨 曜<sup>1</sup> 郭 斌<sup>2</sup> 於志文<sup>2</sup>

(西北工业大学软件与微电子学院 西安 710072)<sup>1</sup> (西北工业大学计算机学院 西安 710072)<sup>2</sup>

**摘 要** 随着智能手机和移动互联网的快速发展,人们的生活方式正在发生着变革。目前,智能手机终端集成了不同种类的传感器,如 GPS、Wi-Fi、摄像头、麦克风等。通过内置传感器,可以搜集用户的相关信息,利用这些信息能有效地理解和识别用户行为,并为用户提供更好的个性化服务。文中主要面向大型工作或生活社区(如大学校园),提出基于背景声音识别的社会活动推荐系统 MSSAR(Mobile Sound Sensing and Activity Recommender)。该系统通过音频传感器采集用户的背景声音,实时感知用户所处环境,并对用户实时活动(如在咖啡厅、在开会等)进行识别。MSSAR 还能结合用户在线交互历史数据计算其偏好及社会关系亲密程度,并据此为用户进行活动推荐。本系统基于以人为本的计算理念,对增强用户间社会联系、促进社区成员间活动交流具有积极意义。

**关键词** 智能手机,移动互联网,背景声音识别,在线线下关联,社会活动推荐

**中图分类号** TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.10.014

## Mobile Social Activity Recommendation System Based on Background Sound Recognition

YANG Yao<sup>1</sup> GUO Bin<sup>2</sup> YU Zhi-wen<sup>2</sup>

(School of Software and Microelectronics, Northwestern Polytechnical University, Xi'an 710072, China)<sup>1</sup>

(School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)<sup>2</sup>

**Abstract** With the rapid development of smart phone and mobile Internet, the life style of people is changing. At present, the smart phone terminal integrates different sensors, such as GPS, Wi-Fi, camera, microphone, and so on. By analyzing the information collected from mobile phone sensing, it can help understand and identify users' activities and provide personalized activity recommendation service. This paper mainly focused on improving the life in big communities (such as in a university campus). We put forward a social activity recommendation system based on background sound recognition using mobile phones. The system, named Mobile Sound Sensing and Activity Recommender (MSSAR), can gather the background sounds through the embedded microphone of mobile phones, recognizing the ongoing user activities (such as in the coffee shop, in a meeting). Furthermore, based on the online interactive historical data, MSSAR can calculate the intimacy among friends and suggest activities accordingly. This system has positive effects to enhance social connection among users and promote communication in the community.

**Keywords** Smart phone, Mobile internet, Background recognition, Association with online and offline, Social activity recommendation

## 1 引言

情境感知技术最早由 Schilit 于 1994 年提出<sup>[1]</sup>。其目标是通过传感器及其相关技术使计算设备能够“感知”用户当前的情境。随着传感器的发展,移动智能设备集成了不同种类的传感器,利用传感器获取信息,可以更好地理解用户行为。而利用背景声音感知用户情境、识别用户行为活动的方法还处于初始阶段,目前研究者一般采用语音识别的相关技术解决背景声音识别的问题,包括语音特征提取和识别技术。对于背景声音识别而言,由于存在很多不确定因素,使得识别率

达不到应用要求,而使应用往往失去了价值。另一方面,随着社交网站的发展,越来越多的年轻用户开始利用社交网站发布、分享自己的所见所闻,例如 Facebook、Twitter、人人网等,甚至可以通过在线交互发起兴趣活动等,例如豆瓣网。目前的社交网络往往基于线上交互,与物理世界的活动及交互行为没有进行有效关联,例如用户不能及时获知周围正在发生的有趣活动,或者线上朋友在线下正进行什么活动。

由于社交网站的使用者往往是年轻人,大学生在社交网站上的活跃程度更高。大学校园每时每刻都发生着各种各样的活动,但是学生们往往因不能及时了解周围朋友正进行的

到稿日期:2013-06-11 返修日期:2013-09-12 本文受国家重点基础研究发展计划(973 计划)(2012CB316400),国家自然科学基金(61373119, 61222209, 61103063),教育部“新世纪优秀人才支持计划”(NCET-12-0466),教育部高等学校博士学科点专项科研基金(20126102110043),陕西省自然科学基金项目(2012JQ8028),西北工业大学基础研究基金(JC20110267)资助。

杨 曜(1988-),男,硕士生,主要研究方向为移动计算, E-mail: yangyao308@gmail.com; 郭 斌(1980-),男,博士,副教授,主要研究方向为普适计算、社会智能、移动社会网络等; 於志文(1977-),男,教授,博士生导师,主要研究方向为普适计算、移动互联网、人机交互和智能信息技术等。

活动而错过。例如这样一种情景:学生小张在某大学上学,一次周末,小张在宿舍,想看看周围的朋友都在干什么,他打开手机应用,系统检测到周围小张的朋友正在发生的活动:小李正在操场打篮球,小王正在图书馆看书。一般来说,由于周围活动比较多,需要根据用户需求进行推荐。这里针对两种需求,1)和老朋友交互,2)和最近线上联系多的朋友交互,小张选择了最近线上联系多的推荐方案。由于小张和小李刚认识不久,且最近常通过社交网站相互交流,最终系统为小张推荐去操场找小李打篮球。

基于以上情境,本文提出了一种基于背景声音识别的校园活动推荐系统,该系统通过获取用户的背景声音,采用分层次的算法识别用户正在进行的活动,用户活动以校园为背景,识别的活动包括说话类、音乐类、其它活动场景 3 大类,分别代表开会/讨论、休闲等不同校园活动场景。其它活动类场景又细分为图书馆、篮球场、街道 3 小类。本文的创新性体现在以下两个方面:

(1)提出了基于背景声音的分层次用户活动识别算法:识别算法不是单一算法,首先依靠决策树对说话类、音乐类和其它活动场景类进行识别;其它活动类根据不同的背景声音再使用 DTW 算法对活动进行识别分类。采用分层活动识别算法可以弥补单一算法的不足,达到更好的识别效果。

(2)结合在线及线下数据分析进行活动推荐:用户参加活动时,常常考虑该活动是否值得参加和该活动中有哪些人参加,而用户与好友的在线交互数据可以反映两人的亲密程度,线下活动数据可以作为用户是否值得去参加该活动的依据,结合两种不同数据可以有效地进行活动推荐。

## 2 相关工作

在个性化推荐和服务中,推荐类型包括朋友、内容、商品、地点等<sup>[2]</sup>。在文献[3]中,作者通过对 Facebook 上 1200 个用户进行分析,采用 GA 算法进行好友推荐;在文献[4]中,作者通过用户的社交关系图对朋友进行聚类分析,根据用户与好友的聚类关系匹配最适合的活动;在文献[5]中,作者从用户思考角度出发,对影响用户是否参加活动的因素进行了分析;而在文献[6]中,作者在活动推荐中考虑了活动群体对用户的影响。目前研究者主要根据在线交互情况进行推荐,而本文强调物理世界活动的推荐,并结合在线和线下信息,如用户亲密密度、距离、活动类型等进行推荐。

通过背景声音识别活动是近年来出现的一个研究领域。作为情境感知的一部分,具有背景声音识别功能的计算终端设备更加具有人性化。通过背景声音识别活动一般沿用语音识别相关技术。文献[7]中,作者对不同的声音特征提取技术进行了比较,包括 MFCC、HCC、短时傅里叶转换等,对这些特征提取技术运用不同的识别算法进行了验证比较,证明 MFCC 特征参数在不同的识别算法中均有比较好的识别率。文献[8]中,作者通过对声音、GPS 等不同传感器收集到的用户信息进行分析学习,识别用户所在场景(例如书店、服装店、酒吧等)。本文与这些工作的区别在于结合在线社交数据,并根据物理活动属性推荐用户最佳活动;此外,在基于背景声音的活动识别方面,本文主要以大学生校园活动为识别对象,具有特定的应用背景和较高的区分度。另外,我们之前的工作<sup>[9]</sup>提出了基于 DTW 的背景声音活动识别方法。与该工作相比,本文有两点改进,1)提出了在线和线下结合的社会活动

推荐方法;2)在背景声音识别方面,提出了分层次的识别方法。

## 3 系统架构

本系统是基于背景声音识别的校园活动推荐系统,系统架构如图 1 所示。该系统分 2 个部分,第 1 部分是校园活动的背景声音识别,包括背景声音采样预处理,最后通过多层次的识别算法对该学生用户的活动进行识别。第 2 部分是结合用户在社交网站上的交互信息,推断该用户在现实生活中的社交好友关系的紧密程度,并结合 GPS 等信息,为该用户推荐最佳活动。

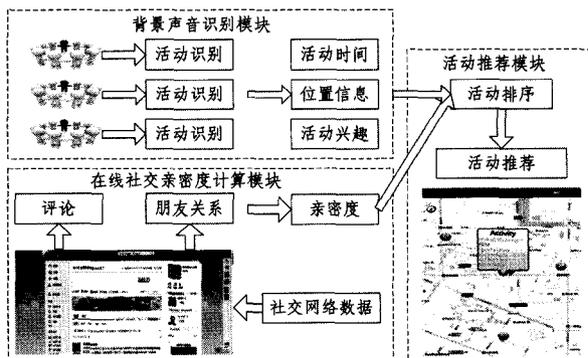


图 1 MSSAR 系统架构

背景声音识别模块:该模块对用户正在发生的行为活动进行识别,识别过程中不需要用户进行干预而自动完成。本文主要识别在校学生的校园活动,识别结果分为说话类、音乐类、环境声音类,其中环境声音类又分为图书馆、篮球场、街道,总共 5 类场景。说话类可以反映学生用户正在交谈、开会或上课等情境,音乐类可以反映学生用户正处于休闲娱乐等状态。由于采用分层次的识别方法,第一层识别结果分为 3 类,包括说话类、音乐类、环境声音类,第二层对环境声音类识别,识别结果分为图书馆、篮球场、街道。

活动推荐模块:该模块根据用户搜索到的周围朋友正在发生活动的相关属性(例如活动位置、用户对该活动的喜爱程度),结合用户与该好友之间的亲密密度计算活动推荐值。亲密密度又根据引言部分提出的用户需求分为动态亲密密度和静态亲密密度两种,其中动态亲密密度指用户和好友在社交网络中的状态评论的交互频繁程度(即最近联系多的朋友),而静态亲密密度指用户和好友的共同好友之间的联系紧密程度(即老朋友关系)。

## 4 背景声音识别算法

在背景声音识别中,识别算法采用传统的语音识别技术(例如决策树、DTW、HMM 等),不同的算法、模型有各自的优点,结合 MSSAR 系统需求,本文提出了分层次的识别策略,对不同的识别类型采取不同的识别算法。

在第一层识别中,识别结果有较大区分度,利用不同的声学特征训练决策树,而决策树是树型选择结构,具有识别效率高、快等优点。第二层识别中构建不同背景声音类的 MFCC 特征,采用路径约束的 DTW 算法作为识别器。

### 4.1 特征选择

背景声音识别分 2 个阶段完成,在第一个阶段中采用决策树算法识别说话类、音乐类、其它活动场景类。在使用决策树之前,需要对声音特征进行提取和对决策树进行构建。在

第二个阶段中,识别算法采用改进后的 DTW 算法<sup>[9]</sup>,在识别之前,需要对每类背景声音训练生成 MFCC 模板。决策树识别过程如图 2 所示。

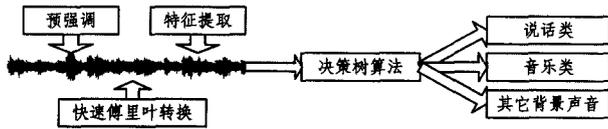


图 2 决策树识别过程

首先,声音信号需要经过预处理阶段,预处理阶段分为预强调(pre-emphasis)和快速傅里叶转换(Fast Fourier Transform, FFT),其中,预强调公式如式(1)所示:

$$H(z) = 1 - \alpha * z^{-1} \quad (1)$$

式中,  $\alpha$  介于 0.9~1.0 之间,作用是为了突出高频部分的共振峰以及使信号变得平坦。实验中设置  $\alpha = 0.95$ 。

训练决策树时,选取了 4 个声音特征属性构建决策树。其中,说话类具有较低的 ZCR 值,而非说话类具有较高的 ZCR 值。音乐类具有较低的 SF 值,而其它活动背景声音具有较低的 BW 值。

ZCR 的计算公式如式(2)所示,其中  $\text{sign}()$  为符号函数。

$$ZCR = \frac{\sum_{i=0}^n |\text{sign}(s_i) - \text{sign}(s_{i-1})|}{2} \quad (2)$$

SF 的计算公式如式(3)所示,其中  $p_t(i)$  和  $p_{t-1}(i)$  是相邻 2 帧的声音信号, SF 可以衡量相邻 2 帧之间声音信号变化的程度。

$$SF_t = \sum_{i=1}^n (p_t(i) - p_{t-1}(i))^2 \quad (3)$$

SC 的特征并没有在构建决策树的时候使用,但是在 BW 特征计算过程中会使用 SC 特征参数的值,其计算公式如式(4)所示。

$$SC_f = \frac{\sum_{i=1}^n i \cdot p(i)^2}{\sum_{i=1}^n p(i)^2} \quad (4)$$

BW 是信号频率的范围宽度,其计算公式如式(5)所示。

$$BW_f = \frac{\sum_{i=1}^n (i - SC_f)^2 \cdot p(i)^2}{\sum_{i=1}^n p(i)^2} \quad (5)$$

## 4.2 识别算法

通过 J48 算法构造决策树可以识别说话类、音乐类和其它活动的背景声音,由于说话和音乐可以较好反映用户的状态,但对说话类和音乐类进行进一步识别必然会带来更大的系统开销,因此说话类、音乐类声音不需要再进一步识别,而其它活动的背景声音类更具体的活动情境需要采用改进后的 DTW 算法来进一步识别<sup>[9]</sup>。

### (1) 传统的 DTW

传统的 DTW 算法基于动态规划思想,运算量较大,但具有技术上容易实现、正识率也较高等特点。为了描述传统的 DTW 算法,我们用  $R$  和  $T$  分别代表参考模板和测试模板。其中:

$$R = \{R(1), R(2), \dots, R(m), \dots, R(M)\}$$

$$T = \{T(1), T(2), \dots, T(n), \dots, T(N)\}$$

$m$  为训练声音帧的时序标号,  $n$  为测试声音帧的时序标号,这样可以形成一个  $M \times N$  的网格, DP 算法就是寻找一条从  $(1, 1)$  到  $(M, N)$  的最短路径。

由于 DTW 算法基于动态规划思想,其最优解为累积距离  $D(M, N)$ , 其由  $R$  和  $T$  对应的帧距离  $d(m, n)$  计算得出。累积距离的递归求解如式(6)所示。其中,初始条件  $D(1, 1) = d(1, 1)$ 。

$$\begin{cases} D(m, n) = d(m, n) + \text{Min} \\ \text{Min} = \min\{D(m, n-1), D(m-1, n-1), D(m-1, n)\} \end{cases} \quad (6)$$

### (2) 改进的 DTW

动态规整过程中,该方法计算量大且需要存储每一个帧距离  $d(n, m)$ , 为了减少计算量和存储空间而采取约束搜索路径方法,将搜索路径限制在两边斜率分别为 0.5 和 2 的平行四边形区域内,以减少搜索范围而提高计算效率。如果搜索路径通过当前坐标  $(n, m)$ , 则下一个坐标只可能是  $(n+1, m+2)$ ,  $(n+1, m+1)$ ,  $(n+1, m)$ 。

如图 3 所示,由于减少了搜索路径,最终结果不一定是最优解,识别率会有所下降。由于平行四边形的两边斜率分别为 0.5 和 2,  $y$  的上界和下界随  $x$  的变化而变化,为了计算  $y$  的边界值,必须把搜索路径计算过程动态地分为 3 个阶段:  $(1, x_a)$ ,  $(x_a+1, x_b)$ ,  $(x_b+1, N)$ , 其中每个阶段对应不同的  $y$  的上边界和下边界,  $x$  的划分必须满足式(7):

$$\begin{cases} x_a = \frac{1}{3}(2M - N) \\ x_b = \frac{2}{3}(2N - M) \end{cases} \quad (7)$$

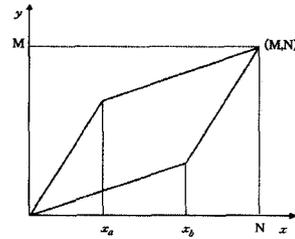


图 3 DP 算法示意图

由于搜索路径限制为一个平行四边形,对于平行四边形以外的帧距离不需要计算,只需要计算平行四边形的上界和下界之间的帧距离,平行四边形边界计算公式如式(8)、式(9)所示。

$$y_{\min} = \begin{cases} \frac{1}{2}x, 0 \leq x \leq x_b \\ 2x + (M - 2N), x_b < x \leq N \end{cases} \quad (8)$$

$$y_{\max} = \begin{cases} 2x, 0 \leq x \leq x_a \\ \frac{1}{2}x + (M - \frac{1}{2}N), x_a < x \leq N \end{cases} \quad (9)$$

由于 DTW 算法把搜索路径限制在如图 3 所示的平行四边形中,在搜索路径中用来计算最短路径的格点都是落在平行四边形内,因此平行四边形外的所有点不会对结果产生影响,计算量会降低很多。其次,累积距离在  $x$  轴每前进一格,都只和前一格的 3 个点的累积距离有关,所以不需要一次性存储所有平行四边形内的格点,这样可以大大减少算法的运行时间和提高其效率。但是由于减少了搜索路径,导致识别率会略微下降。

## 5 活动推荐算法

在校活动推荐根据学生对活动的喜爱程度、活动发生地点和不同的好友亲密度计算策略进行推荐。

• 活动兴趣度:活动兴趣度指用户对该活动感兴趣的程度,所有活动兴趣度初始化为相同值,兴趣度会随着用户多次参与活动而变化。用  $I(u_i, a_j)$  表示该用户对活动的兴趣度。

• 好友亲密度分为动态亲密度和静态亲密度:

动态亲密度依靠用户和好友在社交网站的交互信息计算用户线下的好友亲密度,用户的在线交互信息只考虑最近一个月,可以动态地反映用户的线上好友亲密度的变化。

静态亲密度依靠用户和好友在社交网站中的共同好友数量、共同好友联系的紧密程度来反映该用户和好友的线下亲密度。由于只考虑用户与好友之间的社交拓扑关系,这种关系相对比较稳定。

对于好友的每条状态,考虑其近因效应<sup>[10]</sup>,即用户评论时间越近,亲密度越高,计算公式如式(10)所示,其中  $\Delta t$  表示评论时间和当前时间的相差天数,CR 表示用户在该状态的评论条数中占有评论条数的比例,该比例反映用户在好友的每条状态评论中的积极程度。

$$C_{(u_i, u_j)} = \sum_{k=1}^n \left(\frac{1}{2}\right)^{\Delta t} \cdot CR \quad (10)$$

在计算动态亲密度时,必须考虑用户与好友之间的评论互动,即用户对好友的评论和好友对用户的评论,以减少只有用户对好友的状态评论而没有好友对该用户的评论的影响,并且考虑用户对好友的评论和好友对用户的评论权重关系。计算公式如式(11)所示。

$$DI_{(u_i, u_j)} = w_1 \cdot C_{(u_i, u_j)} + w_2 \cdot C_{(u_j, u_i)} \quad (11)$$

式中,  $C_{(u_i, u_j)}$  表示用户  $u_i$  对好友  $u_j$  的亲密度,  $C_{(u_j, u_i)}$  表示好友  $u_j$  对用户  $u_i$  的亲密度,我们认为用户对好友的评论更能反映其兴趣度,对推荐活动作用更大,故而给其更大的权重。目前系统中的  $w_1$  值设置为 1,  $w_2$  为 0.5。

在计算静态亲密度时,为了描述静态亲密度的计算,提出如下定义:  $E$  代表两个用户的共同好友可能存在边数的最大值,  $e$  代表两个用户的共同好友实际存在的边数,即两个用户存在边的关系。用户与好友之间的共同好友联系的紧密程度的计算公式如式(12)所示。

$$D = \frac{e}{E} \quad (12)$$

式中,如果好友之间联系越紧密,即边的关系越多,该值越大。仅仅计算共同好友之间的联系紧密程度是不够的,如果用户只有 2 个共同好友,且共同好友存在边的关系,那么  $D$  会得到最大值 1,所以还必须计算共同好友占有所有好友的比例,计算公式如式(13)所示。

$$R_{U_c} = \frac{CF(u_i, u_j)}{AF(u_i)} \quad (13)$$

式中,  $CF(u_i, u_j)$  代表用户  $u_i$  与好友  $u_j$  之间的共同好友数量,  $AF(u_i)$  代表用户  $u_i$  所有好友的数量。最后,静态亲密度计算公式如式(14)所示。

$$SI_{(u_i, u_j)} = R_{U_c} \cdot D \quad (14)$$

通过计算好友之间的亲密度,并结合正在发生的活动的其它属性,活动推荐分数计算公式如式(15)所示。

$$Rank = \frac{I(u_i, a_j) \cdot Intimacy = \{DI, SI\}}{Distance} \quad (15)$$

式中,  $Distance$  表示距离因子,由于用户在选择活动时,活动地点离用户有较长距离时才会考虑,而在一定距离范围内,距离不会成为用户考虑活动选择的重要因素,因此在式(16)中  $t$

是距离阈值,小于该阈值时,距离因子变化趋势很小,而大于该阈值时,距离因子变化趋势明显。计算公式如式(16)所示,  $t$  在测试中取值为 3 公里。

$$Distance = \begin{cases} (\sqrt{t}), & d < t \\ (\sqrt{t})^d \cdot (\sqrt{2})^{d-t}, & d \geq t \end{cases} \quad (16)$$

## 6 实验结果

在介绍 MSSAR 算法的基础上,本节将对该算法的性能进行分析和比较。

### 6.1 基于背景声音的活动识别

在 MSSAR 系统中,由于系统运行在资源受限的移动终端,为了降低识别过程中的系统开销和电池能耗,背景声音采集格式为 8000Hz, 16bit, 单声道。目前提出的原型系统主要针对校园内的活动进行推荐,故而实验数据的采集(如背景声音的采集、在线交互数据的采集)主要在校园内进行。但本文工作不限于校园内活动推荐,相关技术和方法可推广到更大的范围。背景声音共采集 600 个文件,每个文件声音长度 1 分钟。其中,200 个背景声音文件(包括 80 个图书馆、60 个街道、60 个篮球场)、200 个音乐类文件、200 个说话类文件,用于训练的文件占 75%,其余 25% 用于测试。

在背景声音第一层识别中,首先构造决策树,由于不同长度的帧对训练决策树结果有影响,实验中选取了 4 种不同长度的帧进行测试,识别结果如表 1 所列。

表 1 不同帧长度的识别结果

	0.032s	0.064s
背景声音类型	识别率/查全率	识别率/查全率
其它活动类	72.00%/75.78%	79.14%/76.73%
说话类	75.93%/62.99%	78.00%/72.41%
音乐类	65.13%/77.11%	70.42%/80.16%
	0.128s	0.256s
背景声音类型	识别率/查全率	识别率/查全率
其它活动类	76.14%/78.71%	76.96%/73.36%
说话类	77.77%/70.62%	76.96%/70.40%
音乐类	67.97%/72.98%	64.70%/75.42%

从 3 种类别识别结果可以看到,帧长度在 64ms 时,识别率都较高,最终在系统中选取 64ms 作为帧长度。而在第二层识别中,采用 MFCC 特征提取,基于路径约束的 DTW 算法作为识别器。第二层识别结果如表 2 所列。

表 2 其它类活动识别结果

	篮球场	图书馆	街道	其它
篮球场	65.89%	19.22%	2.94%	11.95%
图书馆	1.44%	91.50%	0.00%	7.06%
街道	10.89%	13.89%	58.11%	17.11%

从表 2 中可以看到,由于街道的背景声音具有复杂多变等特性,使得识别率只有 58.11%,而背景声音特征比较明显如图书馆之类的安静环境中,识别率相对较高。

### 6.2 推荐结果比较

由于亲密度计算有不同的策略,不同的策略计算结果也不一样,因此推荐也存在不同的方案。其中,通过 Sina API 采集用户和好友最近 30 天内 Sina 微博发布的微博信息计算动态亲密度,包括西北工业大学软件学院 6 位学生用户的 198 个好友的 4852 条评论;通过采集用户和好友之间的 1148 个共同好友计算静态亲密度。不同亲密度计算结果如表 3 第 2 列所示。

表3 亲密度推荐结果

好友	动态亲密度/ 静态亲密度	距离 (单位:km)	推荐值 (动态/静态)
A	0.5625/0.44	0.68	46.07/4.38
B	0.0694/0.75	0.11	4.57/6.49
C	0.8653/0.28	1.32	66.87/2.45
D	0.1016/0.26	0.57	7.65/2.99
E	1.0416/0.66	4.13	44.68/4.24
F	0.0878/0.47	0.98	5.49/4.08

从表3中可以看到,好友A和好友D活动距离相差不多,但是好友D的动态亲密度要低于好友A,因此推荐值好友A较高,而虽然好友C的动态亲密度低于好友E,但是由于好友C的活动距离远低于好友E,因此推荐值C更高。在基于静态亲密度推荐中,好友A和好友D的活动距离相差不多,但是好友A静态亲密度要高于好友D,所以推荐值好友A高于好友D。虽然好友B的动态亲密度(即最近在线交互)较低,但是其静态亲密度(即用户与好友B的共同好友之间联系紧密程度)却最高,推荐值从原来6位好友中的最低变成最高,说明两种亲密度从不同角度对用户进行活动推荐。从表3中可以看到,距离越远,亲密度越低,活动推荐值也越低,推荐结果和实际情况基本一致。由于实验中用户的所有活动兴趣度初始值相同,因此表3中没有列出兴趣度关系。

### 6.3 系统应用

MSSAR系统运行在Android平台,该系统包括界面模块和后台运行模块。界面模块包括功能选择、用户个人足迹和推荐结果显示。其中推荐结果又根据用户选择可以分为3类:基于静态亲密度推荐、基于动态亲密度推荐和多人活动推荐,多人活动推荐通过MSSAR将用户周围多个好友在相同地点进行的相同活动划分为朋友组,并按朋友组对用户进行活动推荐。后台运行模块包括背景声音识别、相关推荐算法等计算过程。

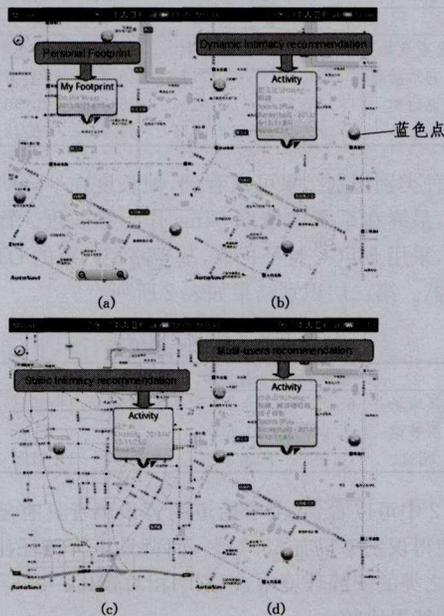


图4 MSSAR系统原型用户界面

MSSAR系统每隔一定时间将通过背景声音识别记录用户的所处场景,如图4(a)所示,系统记录下用户发生的活动,并已可视化形式呈现给用户。图4(a)中每个蓝色记录点是用户在该地点发生的活动,包括活动类型、时间等信息。而图4(b)(c)分别是基于动态亲密度和基于静态亲密度的推荐结果。图4(b)(c)中蓝色点为用户周围好友正在发生的活动,

该活动通过背景声音识别进行分类,并记录好友的活动发生的位置信息、时间信息等,最后为用户进行推荐。而图4(d)是多人推荐结果,通过MSSAR感知到用户的多个好友在同一地点参加同一活动,并将这类好友作为朋友组对用户进行活动推荐。

**结束语** 本文提出了基于背景声音识别的社会活动推荐系统,它通过声音识别物理活动类型,结合在线社交信息为大学生的校园活动进行推荐。通过背景声音活动识别及基于在线交互数据的活动推荐等实验,表明了本文方法的有效性。目前在MSSAR系统中,只考虑了单人活动的识别和推荐,在今后的工作中需要考虑多个好友进行同一活动(即群组活动)的识别,并根据群组活动成员与个体用户间的关系进行亲密度计算和推荐算法的改进。此外,活动识别需要通过智能手机麦克风进行声音采集,这会造成较高的系统能耗,因此应结合用户活动情境(如移动距离、活动地点等)等进行自适应的数据采集间隔设置,以减少系统能耗。由于能耗是智能手机应用开发的重要影响因素,在未来工作中将继续考虑其它耗能相对较低的感知方法来提高活动的识别精度。如通过加速度传感器来识别用户的运动状态(如走路、静止、跑步等),可以增强对用户活动的识别。

### 参考文献

- [1] Schilit B, Adams N, Want R. Context-aware computing applications[C]//IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'94), Santa Cruz, CA, US, IEEE Press, 1994:85-90
- [2] 刘建国,周涛,汪秉宏. 个性化推荐系统的研究进展[J]. 自然科学进展, 2009, 19(1):1-15
- [3] Naruchitparames J, Giine M H, Louis S J. Friend Recommendations in Social Networks using Genetic Algorithms and Network Topology[C]//The 2011 IEEE Congress on Evolutionary Computation, New Orleans, LA, IEEE Press, 2011:2207-2214
- [4] Zanda A, Menasalvas E, Eibe S. A Social Network Activity Recommender System for Ubiquitous Devices[C]//The 11th International Conference on Intelligent Systems Design and Applications, Cordoba, IEEE Press, 2011:493-497
- [5] Quijano-Sanchez L, Recio-Garcia J A, Diaz-Agudo B, et al. Social Factors in Group Recommender Systems[J]. ACM Transactions on Intelligent Systems and Technology, 2013, 4(1):1-30
- [6] Chen Y L, Cheng L C, Chuang C N. A group recommendation system with consideration of interactions among group members [J]. Expert System with Application, 2008, 34:2082-2090
- [7] Cowling M, Sitte R. Comparison of techniques for environmental sound recognition [J]. Pattern Recognition Letters, 2003, 24:2895-2907
- [8] Azizyan M, Constandache I, Choudhury R R. SurroundSense: Mobile Phone Localization via Ambiance Fingerprinting[C]//Proceedings of the 15th annual international conference on mobile computing and networking, New York, USA, ACM Press, 2009:261-272
- [9] 杨曜,郭斌,於志文. 一种基于背景声音识别的社会情境感知方法[C]//第八届和谐人机环境联合学术会议, 2012. 广州, 中国, 清华大学出版社, 2012
- [10] Guo B, He H L, Yu Z W, et al. GroupMe: Supporting Group Formation with Mobile Sense and Social Graph Mining[C]//The 9th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous' 12). Beijing, China, 2012