

基于类依赖的语音情感特征选择

陈 俊¹, 王爱国¹, 王坤侠², 安 宁¹, 李 廉¹

(1 合肥工业大学 计算机与信息学院, 安徽 合肥 230009; 2 安徽建筑大学 电子与信息工程学院, 安徽 合肥 230601)

摘 要: 提出一种基于类依赖的语音情感特征选择模型, 该模型可以提高语音情感识别的准确率, 应用马尔科夫毯分别为每类情感筛选出具有较高表征能力的特征子集, 通过支持向量机用于学习情感分类. 为了更好地融合多分类器的分类决策, 将支持向量机的二分类转化为概率分类, 以解决多分类器联合使用的投票冲突问题. 在公开数据集上的实验结果表明, 与信息增益、主成分分析以及类独立的特征选择方法相比, 此方法可以较大幅度减少样例维数, 同时获得较高的情感识别率.

关键词: 语音情感识别; 特征选择; 类依赖; 马尔科夫毯; 支持向量机

中图分类号: TP391.4

文献标识码: A

文章编号: 1000-7180(2016)08-0092-05

Speech Emotion Recognition with Class-dependent Feature Selection Methods

CHEN Jun¹, WANG Ai-guo¹, WANG Kun-xia², AN Ning¹, LI Lian¹

(1 School of Computer and Information, Hefei University of Technology, Hefei 230009 China;

2 Department of Electronic Information Engineering, Anhui University of Architecture, Hefei 230601, China)

Abstract: This study proposes a novel class-dependent feature selection model to improve the performance of speech emotion recognition. Particularly, we adopt the markov blanket technique to select discriminative features for each emotion class, and use the support vector machine classifier to build the emotion recognition model. To make better decision in multi-label classification, the binary discriminative output of the support vector machine is transformed to a probability output for solving the voting conflict problem. Extensive experimental results on the a publicly available dataset show that in comparison with information gain, principal component analysis, and class-independent feature selectors, the proposed method significantly reduces the feature dimensionality and obtains better classification accuracy.

Key words: speech emotion recognition; feature selection; class-dependent; Markov blanket; support vector machine

1 引言

语音情感识别是计算机智能的一个重要组成部分, 并被广泛地应用在人机交互领域中. 在人机交互中, 语音情感识别的研究, 不仅推动计算机技术的进一步发展, 也将给包括新型人机交互在内的诸多领域带来良好的经济效益和社会效益.

在实际应用中, 研究人员从原始语音信号中提

取了大量的时域、频域语音特征并用于情感识别. 但不相关的特征会降低基于实例的学习器的性能^[1]. 为提高语音情感识别的准确率, 特征选择为我们提出了一种有效的方法. 常用的特征选择方法包括主成分分析(PCA)^[2]、序列前向选择(SFS)^[3]、序列后向选择(SBS)^[4]、优先选择算法(PFS)^[5]等. 前三种方法是基于启发式的搜索策略进行特征选择的, 能够在一定程度上缩小候选特征子集的搜索空间, 并

收稿日期: 2015-11-01; 修回日期: 2015-12-16

基金项目: 国家自然科学基金项目(61305064); 国家高等学校学科创新引智计划(“111”)(B14025); 国家科技支撑计划(2013BAH19F01); 教育部直属高校外国文教专家年度聘请计划—学校特色项目(TS2013HFGY031)

在实际应用中获得不错的效果^[4]. 例如,谢波等人^[5]应用 SFS、SBS 和 PFS 方法在普通话情感语音库上进行语音情感特征选择,并详细比较了这几种方法的有效性. 然而现有的这些语音情感特征选择方法都是针对所有的情感类别选择一个共同的特征子集,而不是为每一个情感类别筛选出各自的具有判别能力的特征子集. 不同特征对不同情感类别的分类所起的作用不同,甚至某些特征对某个分类起着决定性的作用,而对另一分类却造成了很大的干扰,所以只利用一个共同的特征子集进行分类会使情感的分类具有较低的识别率.

因此,本文提出将语音情感识别的多分类问题转化为多个二分类问题,并在此基础上建立基于类依赖的语音情感特征选择模型,然后为每类情感筛选出相应的表征能力强的特征子集. 本文采用支持向量机分类器,并且通过将支持向量机的判别输出转化为语义向量输出,解决投票冲突问题.

2 语音情感特征提取

语音情感识别技术主要包括三个过程:语音信号采集、语音情感特征提取和特征选择、语音情感识别. 首先通过麦克风等语音录制设备获取语音信号,并将这些信号传递给情感特征提取与选择模块. 此模块主要负责提取语音信号中的与说话人情感关系紧密的声学参数,然后送入语音情感识别模块进行语音情感识别. 从语音识别流程可以看出,语音之所以能够表达情感,是因为语音信号中包含能表达情感的特征参数. 因此,提取合适的语音信号参数对于语音情感的识别具有重要的作用.

目前,在语音情感识别技术中广泛用到的声学特征主要包括三种类型的特征:韵律学特征、基于谱的相关特征以及音质特征. 其中,韵律是指语音中的音高、音长、快慢和轻重等方面的变化. 韵律学特征又被称为“超音段特征”,最常用的韵律学特征有时长、基频、能量等. Luengo 等人进行了一系列的韵律特征分析研究,并通过计算一阶差分和二阶差分曲线,统计出曲线的最大值、最小值、均值、方差、峰度等作为语音的特征; Iliou 和 Wang 等人则分别将基频、能量、时长等韵律特征用于德语和汉语的语音情感识别. 基于谱的相关特征体现的是声道形状变化和发声运动之间的相关性. Nwe 等人研究发现,不同情感的语音所对应的频谱能量在各个频谱区间的分布是不一样的. 基于谱的相关特征主要包括线性谱特征和倒谱特征. 线性谱特征一般有线性预测系数(Linear Predictor Coefficient, LPC)、

对数频率功率系数(Log-frequency Power Coefficient, LEPC),以及单侧自相关线性预测系数(One-sided Autocorrelation Linear Predictor Coefficient, OSALPC)等;倒谱特征一般有线性预测倒谱系数(Linear Predictor Cepstral Coefficient, LPCC)、梅尔频率倒谱系数(Mel-frequency Cepstral Coefficient, MFCC)等. MFCC 是 Davis 和 Mermelstein 于 1980 年首次提出并广泛应用于语音情感识别中. 声学质量用于衡量语音是否纯净、清晰、容易辨识,其声学表现主要有喘息、颤音、哽咽等. 声音质量特征一般包括共振峰频率及其带宽、频率微扰和振幅微扰、声门参数. 例如, Luggner 和 Yang 等人使用共振峰频率及其带宽和部分韵律特征进行语音情感识别; A. Moshier 根据基音、能量和共振峰频率等特征进行语音情感识别; 王治平、赵力使用基于语速、振幅、基音和共振峰的语音特征参数,并提出一种基于模糊熵的特征有效性分析方法. 另外, Sun 等人通过实验比较了声门参数、基频、能量等韵律特征在情感识别中发挥的作用. 鉴于以上三种特征是从不同侧面来表达语音情感信息的, Schuller 和 Malandrakis 等人将这三种特征进行融合,并应用于语音情感识别以改善系统识别率.

3 基于类依赖的语音情感特征选择模型

虽然针对所有的语音情感样例选出一个共同的特征子集进行情感分类可以获得不错的效果,然而该特征子集并不会对每个类别都具有很好的表征能力;例如,特征 f_1 对于类别 C_1 具有很好的判别能力,特征 f_2 对于类别 C_2 具有很好的判别能力, f_3 对于类别 C_3 具有较好的判别能力,但是 f_1 可能会对类别 C_2 的分类或类别 C_3 的分类造成一定的干扰,或者 f_2 会对于类别 C_1 或类别 C_3 造成一定的干扰,这样 $\{f_1, f_2, f_3\}$ 用于多分类时,其分类的准确率不见得很高. 为此,本文提出一种基于类依赖的语音情感特征选择模型,旨在为每个情感类别筛选出最具有区分能力的特征子集;然后在此基础上为每个情感建立相应的预测模型.

3.1 基于类依赖的特征选择框架

基于类依赖的特征选择的目的是从每类情感筛选出具有判别能力的特征子集. 与基于类独立的特征选择方法不同,在类依赖的特征选择方法中,为不同类别选择的特征子集可能是不一样的. 具体来说,假设待分类的语音情感的类别数目是 C , 标签空间 $L = \{L_1, L_2, \dots, L_c\}$, 原始特征空间 $F = \{F_1, F_2, \dots, F_n\}$ 由 n 个特征组成. 在基于类依赖的特征

选择框架中, 首先将原训练集 Data 的 C 分类问题, 转化成 C 个二分类问题, 记该步骤为数据集标签的二值化. 假设, 当前关注的类标签是 $L_i (1 \leq i \leq C)$, 二值化操作将原标签值为 L_i 的标签值设置为 1, 原标签值不是 L_i 的标签值设置为 0; 然后在二值化后的数据集 $Data(L_i)$ 上进行特征选择, 筛选出与类别 L_i 相关的特征子集 S_i . 图 1 给出了本文提出的基于类依赖的特征选择框架. 根据该模型, 可获得与每个情感类别相关的特征子集.

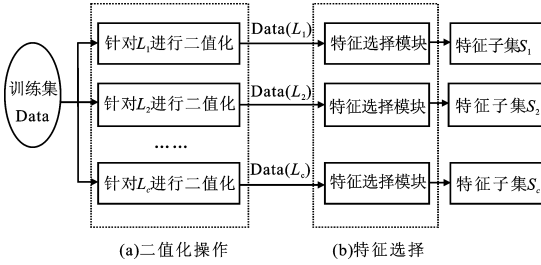


图 1 基于类依赖的特征选择框架

3.2 马尔科夫毯特征选择算法

针对每个情感类别, 本文使用马尔科夫毯技术 (Markov Blanket) 进行特征选择. 马尔科夫毯特征选择方法是通过基于约束的条件独立性来寻找目标变量的马尔科夫毯局部结构, 它不需要学习整个网络结构, 因此在实际应用中具有较低的时间复杂度, 并且通过马尔科夫毯技术获得的特征子集考虑了特征之间相关性和冗余性, 并且能够在理论上说明获得的特征子集是全局最优的.

Markov Blanket 是由 Pearl 于 1988 年提出的, 并且 Pearl 在贝叶斯网络中对马尔科夫毯的性质进行了深入研究. 在图形化的贝叶斯网络中, 某个变量的马尔科夫毯由该变量的子节点、父节点以及子节点的父节点组成. 给定某个节点的马尔科夫毯, 该节点与其他节点是条件独立的, 因此马尔科夫毯可以屏蔽网络中其他变量对该变量的影响. 马尔科夫毯特征选择方法与其他方法相比, 不仅考虑了特征与类别之间的相关性, 而且能够识别特征之间的冗余. 在马尔科夫毯的寻找上, Koller^[6] 提出根据交叉熵理论寻找目标变量的马尔科夫毯, 从而剔除与目标变量不相关的、冗余的特征. 下面给出马尔科夫毯的定义和相关定理.

定义 1 Markov Blanket. 给定一个变量 T , 变量集为 F , 假设该变量的马尔科夫毯为 $MB(T)$, $MB(T) \subseteq F (T \notin MB(T))$, $MB(T)$ 满足如下条件:

$$P(F - MB(T) - \{T\} | T, MB(T)) =$$

$$P(F - MB(T) - \{T\} | MB(T)) \quad (1)$$

即 $MB(T)$ 是 T 的马尔科夫毯, 当且仅当在给定 $MB(T)$ 的条件下 T 和 $F - MB(T) - \{T\}$ 是独立的.

定义 2 配偶节点. 给定一个变量 T , 把 T 的子节点的父节点称为其配偶节点, 并记为 $SP(T)$.

定理 1 当有向无环图 G 忠实于联合概率分布 J 时, G 中节点 T 的马尔科夫毯 $MB(T)$ 是唯一的, 也就是 T 的父节点、子节点, 以及子节点的父节点 (配偶节点).

从以上定义和定理可以得出, 一个变量 T 的马尔科夫毯是由其父节点、子节点, 以及子节点的父节点构成的集合. 例如, 图 2 中圆圈包围的阴影节点 $MB(T) = \{C, D, E, F, G\}$ 就构成了变量 T 的马尔科夫毯, 并且有

$$P(A, B | T, MB(T)) = P(A, B | MB(T)) \quad (2)$$

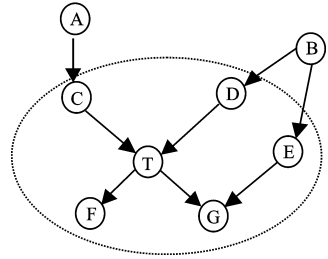


图 2 T 的马尔科夫毯 (阴影节点)

常用的马尔科夫毯算法有 Grow-shrink (GS)^[7]、Incremental Association Markov Blanket (IAMB)^[8]、Max-Min Markov Blanket (MMMB)^[9]、Hiton-Markov Blanket (Hiton-MB)^[10] 和 Parents-Children Markov Blanket (PCMB)^[11]. GS 和 IAMB 包括增长和收缩两个阶段, GS 的时间复杂度大, 效果较差, IAMB 运行时间虽然较少, 但样本量较少时, 算法的学习效果较差; MMMB、Hiton-MB、和 PCMB 都是基于拓扑结构的分治法, MMMB 和 Hiton-MB 比 IAMB 有效, 但比 IAMB 运行时间长; PCMB 则面临着时间复杂度高的问题. 为较好地平衡时间代价和获得的特征子集的质量, 本文拟采用 MMMB 算法.

最大-最小马尔科夫毯算法 (Max-min markov blanket) 是由 Tsamardinos 等人于 2006 年提出的. MMMB 算法采用分治法首先调用 MMPC 算法寻找节点 T 的父节点和子节点的集合 $PC(T)$, 然后找到 T 的配偶节点 $SP(T)$. 那么 $PC(T)$ 和 $SP(T)$ 组成了节点 T 的马尔科夫毯 $MB(T)$. 然后, MMMB 寻找节点 T 的候选马尔科夫毯 CMB, 并且应用条件独立性测试, 从 CMB 找到 T 的配偶节点, 从而最

终得到 T 的马尔科夫毯.

定义 3 条件独立和条件依赖. T 和 X 为两个变量,当且仅当 $P(T|X,S) = P(T|S)$ 时,称变量 T 和 X 在给定变量集 S 情况下条件独立,记作 $\text{Ind}(X;T|S)$;否则称变量 T 和 X 在给定 S 情况下条件依赖,记作 $\text{Dep}(X;T|S)$.

为方便理解马尔科夫毯算法,下面给出 MMPC 算法和 MMB 算法的描述,在下面的伪代码中把在给定 Z 的前提下, X 与 T 的关联性记为 $\text{assoc}(X;T|Z)$.

算法 1:MMPC 算法描述

输入:目标变量 T

输出: T 的父子节点集合 $\text{PC}(T)$

$\text{CPC} = \emptyset$

Repeat

 Foreach X find

$\text{Minassocset}(X) = \text{subset } s \text{ of } \text{CPC}$

 That minimizes $\text{assoc}(X;T|s)$;

 End for

 Find $F \in \Phi \setminus (\{T\} \cup \text{CPC})$ that

 Maximizes $\text{assoc}(F;T|\text{minassocset}(F))$;

 If $\text{Dep}(F;T|\text{minassocset}(F))$

$\text{CPC} = \text{CPC} \cup F$;

End if

Until CPC has not changed

Foreach $X \in \text{CPC}$

 If $\exists s \subseteq \text{CPC}, s.t. \text{Ind}(X;T|s)$

$\text{CPC} = \text{CPC} \setminus \{X\}$ /* 把 X 从 CPC 中移除 */

 End if

End for

Return CPC

算法 2:MMMB 算法描述

输入:目标节点 T

输出: T 的马尔科夫毯 $\text{MB}(T)$

/* 获取 $\text{MB}(T)$ 的 CMB (候选马尔科夫毯) */

$\text{PC}(T) = \text{MMPC}(T)$; /* $\text{PC}(T)$ 表示 T 的父子节点集合 */

$\text{MB}(T) = \text{PC}(T)$; /* 用 $\text{MB}(T)$ 来表示 T 的马尔科夫毯集合 */

$\text{Set} = \{\}$;

For each $C \in \text{PC}(T)$:

$\text{Set} = \text{Set} \cup \text{MMPC}(C)$;

End for

/* 找出 T 的配偶节点 */

For each $X \in \text{Set}$:

 寻找集合 S ,使得 $\text{Ind}(X;T|S)$;

 For each $Y \in \text{PC}(T)$:

 if $\text{Dep}(X;T|Y \cup S)$ /* 条件成立说明: X 就是要找的 T 的配偶节点 */

$\text{MB}(T) = \text{MB}(T) \cup \{X\}$; /* 把 X 加入到 T 的马尔科夫毯集合中 */

 End if

 End for

End for

return $\text{MB}(T)$;

3.3 语音情感识别模型

根据上述的特征选择框架,可以获得 C 个特征子集 $\{S_1, S_2, \dots, S_c\}$,分别对应于 C 个不同的情感类别.为检验最终选择的特征子集的质量,我们在每个特征子集上分别建立相应的预测模型,其中 Model_i 表示针对标签 L_i 建立的模型,用于预测测试样例的标签.图 3 给出的是所建立的语音情感训练模型.具体地,针对类别 L_i ,首先将原始数据集 Data 映射到特征子集 S_i 上,获得在 S_i 上的数据集 $\text{Data}(S_i)$,然后在该数据集上训练分类器得到预测模型 Model_i .

在对未知测试样本进行预测时,本文提出如图 4 所示的框架.具体地,给定某个测试样本 Test ,首先将其映射在不同的特征空间 S_i 中,得到维度消减后的样本 $\text{Test}(S_i)$;然后将该样本分别送到每一个二分类模型 Model_i 中.在采用 SVM 分类器时, Model_i 的输出是判别输出,即判定结果只有两种:测试样本属于类别 L_i ,输出 1,以及测试样本不属于类别 L_i ,输出 0.支持向量机在使用判别输出进行投票时,可能会出现多个二分类模型输出 1 的情况,这样便会出现投票冲突问题,导致不能确定 Test 的标签.为解决该问题,本文提出使用将支持向量机的判别输出转化为语义向量输出,即当测试样例经由每一个二分类器时,二分类模型不仅给出该测试样例的标签,而且计算出该判断的可信度 Score . Score 表示模型做出当前判断的可信值,值越大表示当前的判断越可信.因此,在出现投票冲突问题时,可以结合 C 个二分类模型的预测标签和可信值推测未知样本的标签.具体地,在 C 个二分类模型的输出中:(1) 如果只有模型 Model_i 的输出值是 1,那么测试样本的标签为 L_i ;(2) 如果有两个以上的模型输出 1,那么取具有较大可信值的模型对应的标签;(3) 如果所有的二分类模型输出均为 0,那么选择具有最低可信度的模型对应的标签.接下来本文将给出可信度 Score 的数学计算公式^[12]:

$$\text{Score} \approx \frac{1}{1 + e^{\text{Ag} + B}} \quad (3)$$

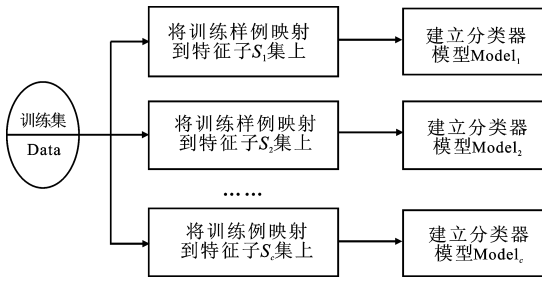


图3 语音情感训练模型

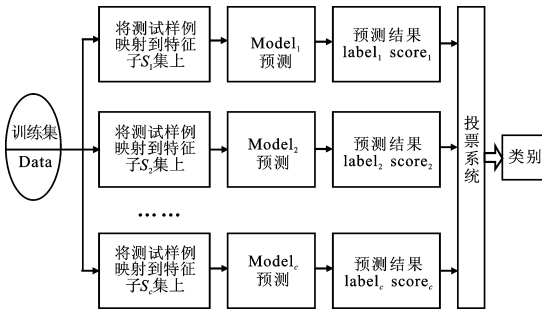


图4 测试样例的分类模型

式中, g 是样例 x 的预测标签, 另外, 通过最小化训练数据的负对数似然来估计 A 和 B 的值(用它们的真实标签和预测标签)。

4 实验结果

本文提出的方法所选的特征数比这三种方法所选的特征数要少, 同时, 识别率却比这三种方法都要高。由此验证了基于类依赖的特征选择方法无论是在所选的特征子集维数方面还是在分类性能方面, 都优于主成分分析, 信息增益和基于类独立的MMMB特征选择方法。

5 结束语

本文应用马尔科夫毯技术进行特征选择, 该方法能够有效地识别不相关的、冗余的特征, 在实际中可以获得较好的性能。针对测试样例在分类过程中出现的投票冲突问题, 本文通过将支持向量机的判别输出转化为语义向量输出, 并给出相应的解决方案。实验结果表明, 基于类依赖的语音情感特征选择方法能够获得比主成分分析、信息增益以及基于类独立的特征选择方法更好的识别率, 而且能够选出较少数目的特征。实际上, 在基于类依赖的特征选择模型中, 可以选用的特征选择方法并不仅仅局限于马尔科夫毯算法, 其他的特征选择方法也可以嵌入本文提出的模型中。在下一步的工作中, 一方面将探索本文提出的方法在其他领域中的应用, 另一方面

将在本文提出的模型中集成其他的特征选择方法。

参考文献:

- [1] Wang Ai-guo, An Ning, Chen Gui-lin, et al. Accelerating wrapper-based feature selection with K-nearest-neighbor [J]. Knowledge-Based Systems, 2015, 83 (1):81-91.
- [2] Ntalampiras S, Fakotakis N. Modeling the temporal evolution of acoustic parameters for speech emotion recognition [J]. Affective Computing, IEEE Transactions on, 2012, 3(1):116-125.
- [3] Morrison D, Wang R, De Silva L C. Ensemble methods for spoken emotion recognition in call-centres [J]. Speech communication, 2007, 49(2):98-112.
- [4] Kwon O W, Chan K, Hao J, et al. Emotion recognition by speech signals [C]// Proceedings of International Conference Eurospeech. [s. l.]:IEEE, 2003:1-8.
- [5] Xie B, Chen L, Chen G C, et al. Feature selection for emotion recognition of mandarin speech [J]. Journal of Zhejiang University Engineering Science, 2007, 41 (11): 1816-1822.
- [6] Koller D, Sahami M, T, Koller D, et al. Toward Optimal Feature Selection [J]. International Conference on Machine Learning, 1996, 20(3):284-292.
- [7] Margaritis D, Thrun S. Bayesian network induction via local neighborhoods [R]. Carnegie-Mellon University Pittsburgh PA Department of Computer Science, 1999.
- [8] Tsamardinos I, Aliferis C F, Statnikov A R, et al. Algorithms for Large Scale Markov Blanket Discovery [C] // FLAIRS Conference. Nashville, TN: IEEE, 2003: 376-381.
- [9] Tsamardinos I, Aliferis C F, Statnikov A. Time and sample efficient discovery of Markov blankets and direct causal relations [C]// Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. Washington, DC, USA: ACM, 2003: 673-678.
- [10] Aliferis C F, Tsamardinos I, Statnikov A. HITON: a novel markov blanket algorithm for optimal variable selection [C]// AMIA Annual Symposium Proceedings. American Medical Informatics Association Nashville, TN: IEEE, 2003:21-25.
- [11] Peñ J M, Nilsson R, Björkegren J, et al. Towards scalable and data efficient learning of Markov boundaries [J]. International Journal of Approximate Reasoning, 2007, 45(2):211-232.

(下转第 101 页)

5 结束语

本文针对 LEACH 协议中存在的节点簇头选举不合理、能耗不均以及监测得到的数据冗余度较大的问题,提出一种基于正六边形网格的新型路由协议 LEACH-HG. 该协议引进虚拟六边形网格来划分整个监测区域,并在各个网格中选取代表节点作为候选簇头节点,采用综合考虑节点能量和单元格内存活节点数的簇头选举方式,有效避免簇头随机选择的同时引入休眠机制使能量消耗和数据冗余现象得到改善.

参考文献:

[1] 孙利民,李建中,陈渝,等. 无线传感器网络[M]. 北京:清华大学出版社,2005.

[2] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless micro sensor networks[C]// Proceedings of the 33rd Annual Hawaii International Conference on System

Sciences, Maui; IEEE,2000:3005-3014.

[3] 沈玉龙,徐启建,裴庆祺,等. 基于栅格的无线传感器网络路由方法[J]. 通信学报,2009,30(11):96-100.

[4] 郭彬,李喆. 无线传感器网络中基于剩余能量的联合选举动态成簇路由算法[J]. 电子与信息学报,2007,29(12):3006-3010.

[5] LI B, WANG W J, YIN Q Y, et al. An energy-efficient geographic routing based on cooperative transmission in wireless sensor networks[J]. Sci China Inf Sci, 2013, 56(11): 4757-4762.

[6] 蒋畅江,石为人,唐贤伦,等. 能量均衡的无线传感器网络非均匀分簇路由协议[J]. 软件学报, 2012,23(5): 190-200.

作者简介:

严斌亨 男,(1993-),硕士. 研究方向为无线传感器网络路由协议. E-mail:one01.23@foxmail.com.

刘 军 男,(1963-),硕士,教授. 研究方向为无线传感器网络.

(上接第 96 页)

[12] Chang C C, Lin C J. LIBSVM: a Library for Support Vector Machines [J]. ACM Transactions on Intelligent Systems & Technology, 2006, 2(3):389-396.

作者简介:

陈 俊 男,(1990-),硕士. 研究方向为情感计算和机器学习.

王爱国 男,(1986-),博士. 研究方向数据挖掘及其在健康信息学中的应用.

王坤侠(通讯作者) 女,(1980-),博士. 研究方向语音信号处理、机器学习和情感计算. Email: kxwang@ahjzu.edu.cn

安 宁 男,(1971-),博士生导师,教授. 研究方向情感计算、人机交互和机器学习.

李 廉 男,(1951-),博士生导师,教授. 研究方向计算数学、网格计算和机器学习.