

基于粗糙集的决策树在医疗诊断中的应用

黄锦静,陈岱,李梦天

(中国矿业大学 计算机科学与计算学院,江苏 徐州 221116)

摘要:网上医疗诊断越来越受欢迎,电子病例的数据也越来越多。如何从众多的医疗数据中降低医疗数据的冗余度,快速提取有用的医疗价值,提高医疗诊断的速度和精度,成了一个大家研究的热点问题。针对这一系列问题,研究了医疗系统关于肺癌诊断的一些数据,建立了基于属性依赖改进的可分辨矩阵属性约简的C4.5算法,并用随机森林进行算法改进。属性约简算法降低了医疗数据的冗余度,决策树算法提取了肺癌诊断的一些规则,随机森林提高了医疗诊断的准确性。文中对肺癌诊断场景进行了仿真实验与应用,并将单纯的C4.5算法,属性约简与单棵C4.5决策树,属性约简和C4.5决策树随机森林进行性能比较。实验结果表明,该方法加快了计算速度,提高了医疗诊断的精度。

关键词:粗糙集;属性约简;可分辨矩阵;C4.5算法;决策树

中图分类号:TP39

文献标识码:A

文章编号:1673-629X(2017)12-0148-05

doi:10.3969/j.issn.1673-629X.2017.12.032

Application of Decision Tree Based on Rough Set in Medical Diagnosis

HUANG Jin-jing, CHEN Dai, LI Meng-tian

(School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China)

Abstract: Online medical diagnosis is becoming more and more popular, so more and more data are in electronic records. How to reduce the redundancy of medical data, extract useful medical value rapidly from a large number of medical data, and improve the speed and accuracy of medical diagnosis has become a hot issue. In view of it, some data of diagnosis of lung cancer in medical system are researched, and the C4.5 algorithm of attribute reduction based on attribute-dependent improved discernibility matrix is established and improved by stochastic forest. Attribute reduction algorithm reduces the redundancy of medical data, the decision tree algorithm extracts some rules of lung cancer diagnosis, and the stochastic forest raises the accuracy of diagnosis. In this paper, simulation and application are carried out under the scenario of lung cancer diagnosis. The simple C4.5 algorithm is made a comparison with the attribute reduction and the single C4.5 decision tree, and attribute reduction and random forests of C4.5 decision tree. The experiment shows that the proposed method accelerates the computing and improves the accuracy of medical diagnosis.

Key words: rough set; attribute reduction; discernibility matrix; C4.5 algorithm; decision tree

1 概述

随着计算机行业的快速发展,医院也开始走向电子化时代,网上挂号、网上诊断、电子病例等,积累了海量的医疗数据。如何利用这些数据挖掘出医疗价值,即基于医疗的数据挖掘,成为一个非常热门的研究领域。医疗数据的挖掘在医疗领域应用的主要内容包括医疗图像数据挖掘、医疗管理检索系统和电子病例的分析等。例如,在医疗图像方面,文献[1]提出了一种基于成像的神经退行性疾病的分类方法,以提取突出的脑模式。在电子病例这种文本数据方面,文献[2]提出了一种可变精度粗糙集模型的属性约简算法。

特别是电子病例这类的文本数据,数据量庞大,储存了大量信息,是数据挖掘的一个热点对象。病例隐含了众多信息,记录了很多病情症状,即很多属性特征,针对不同的病情不同属性的重要度也不同,但是电子病例里面往往包含很多属性,而且有很多是冗余的。大家去医院检查都要做各种繁复的检查,其实很多检查项是冗余的,记录了很多冗余的属性,同时还增加了病人的费用。针对这些问题,如何在海量的电子病例中删除冗余症状,高效和准确地建立决策树,挖掘有重要应用价值的特征数据,是一个十分重要的课题。

粗糙集理论是一种研究模糊性和不确定性问题的

收稿日期:2017-01-21

修回日期:2017-05-25

网络出版时间:2017-09-27

基金项目:国家级大学生创新项目(201510290002)

作者简介:黄锦静(1994-),女,研究方向为计算机科学与技术;陈岱,副教授,研究方向为计算机应用。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20170927.1000.074.html>

数学分析和数据挖掘工具。粗糙集的主要内容是属性约简、规则提取,在很多领域应用广泛。例如,文献[3]在邮件过滤领域提出了基于变精度粗糙集的决策树分类算法,提高了邮件的正确分类率。文献[4]提出了一种在多准则排序方面的算法,即基于粗糙集的多准则排序方法,非常高效。文献[5]将粗糙集应用到食品安全检测领域,并且得到了很好的食品检查效果,提高了食品检查的速度,文献[6]将粗糙集和 BP 神经网络相结合,提高了粮食产量预测的精确度。文献[7]为了提高图像插值方法的插值效率,改善放大图像边缘的模糊现象,提出了一种基于粗糙集约简的支持向量机图像插值方法。文献[8]为解决危险品风险分析的一些问题,应用粗糙集理论,加快了计算速度。文献[9]为了能够有效、快速地评价电网运行的优质性水平,提出一种粗糙集与证据理论相结合的方法等。

文中将粗糙集的属性约简算法应用到肺癌诊断上。属性约简算法有很多,如通过去除某属性后,判断不可区分关系是否改变来决定是否应删除该属性,但是这种算法得到的约简集不完备。自从 SKOWRON 提出可分辨矩阵这一理论后^[10],就出现了可分辨矩阵和粗糙集理论相结合的属性约简算法^[11]。这种算法可以得到比较完备的约简集,但是引入矩阵自然增加了运算量,所以算法的复杂度高、耗时长。为了得到完备的最小约简集,文中提出了一种基于属性依赖改进的可分辨矩阵的属性约简算法,减少了计算量,提高了运算速度。

经过属性约简后的病例集属性间的冗余度有所降低,将该病例集作为训练集来构建决策树,进行规则提取。这样相比直接构建决策树进行规则提取,不仅减少了计算量,也减少了树的深度,使规则更加准确。决策树是从一组无次序且无规则的元组中推理出一组以决策树为表现形式的分类规则。它是一种自顶向下的递归方式,在树的内部节点进行属性值之间的比较,根据属性值的不同从当前节点向下进行分支,而叶节点是最终要划分的类,即决策属性。从根节点到叶节点的一条路径对应着一条合取规则,而整个决策树对应着一组析取表达式规则。决策树算法有很多,如基于粗糙集的变精度算法^[12]和 ID3 算法^[13],但是它们不能很好地处理连续值;C4.5 算法是 ID3 算法的改进,能很好地处理连续值。因此文中采用 C4.5 算法通过信息熵增益率来构建决策树,同时为了提高准确率引入了决策树的随机森林。

2 问题定义

就肺癌的医疗诊断数据进行分析,定义病例样本

集 $S = \{U, C, D, V, f\}$ 。其中 U 为病例样本的实例集合 \emptyset ,即论域; C 为条件属性; D 为决策属性; V_i 为属性 i 的值域; f 为信息映射函数, $f: C(C \cup D) \rightarrow V$ 。

根据给定的训练集建立决策树,找到肺癌诊断的推理规则,进行智能诊断并预测病人的情况。因此如何建立快速且准确的诊断规则,即如何快速准确地建立决策树是关键。

2.1 粗糙集约简算法模型

定义 1:已知论域 U ,对于每一个属性子集 $P \subseteq (C \cup D)$, $P \neq \emptyset$,定义不可分辨关系 $IND(P)$;则有: $\forall x \in U, [x]_{IND(P)} = [x]_P = \cap_{R \in P} [x]_R$ 。即论域 U 上的一个划分为 $U/IND(P)$, $IND(P)$ 为论域 U 的等价关系,即有 $U/IND(P) = \{[x]_{IND(P)} \mid \forall x \in U\}$ 。

定义 2:对于 $\forall x \in U, R \in C$,属性 R 的下近似 \underline{R} 和上近似 \bar{R} 分别为:

$$R(X) = \{X \mid (\forall x \in U) \wedge ([x]_R \subseteq X)\}$$
$$\bar{R}(X) = \{X \mid (\forall x \subseteq U) \wedge ([x]_R \cap X \neq \emptyset)\}$$

(2)

$pos(R) = \underline{R}(X)$ 称为 R 的正域。

定义 3:对于 $R \in C$,对于决策条件 D ,如果 $pos_{IND(C)}(IND(D)) = pos_{IND(C - \{R\})}(IND(D))$ 则 R 是条件属性 C 上 D 的可约分关系,否则是不可约分关系。

定义 4:在论域 U 上,所有的条件属性 C 上 D 的不可约分关系构成的集合,称为 C 的核心集,记为 $Core(C)$ 。核心集为条件属性的所有约简集的交集,所以约简集一定包含核心集,且对于每个病例样本集核心集是唯一的。

定义 5:对于病例集 $S = \{U, C, D, V, f\}$,对于所有的属性 P, Q 属于 C ,有 Q 对 P 的依赖度为

$$\gamma_P(Q) = k = \frac{|\underline{pos}_P(Q)|}{|U|} = \frac{|\bigcup_{x \in U/Q} \underline{P}(X)|}{|U|} \Leftrightarrow P \Rightarrow Q_k (0 \leq k \leq 1)$$

(4)

若有属性 $k = 1$,则 P 和 Q 等价。所以为了节省计算量, P 和 Q 不应该出现在同一个最小约简集。

文中采用基于属性依赖度的可分辨矩阵约简算法,对于经典的属性频率的约简算法而言,减小了很多不必要的计算,并且考虑了属性间的依赖性,使属性约简集更加完备^[14]。

对于病例集 S ,构建其二进制可分辨矩阵:

$$m((i, j), k) = \begin{cases} 1, & (D(x_i) \neq D(x_j)) \wedge (c_k(x_i) \neq c_k(x_j)) \\ 0, & \text{others} \end{cases}$$

(5)

当 $m((i, j), k) = 1$ 时, 表示属性 c_k 在 $(x_i, x_j), x_i, x_j \in U$ 下可区分, 否则不可区分。

文中基于核心集 $\text{Core}(C)$, 首先求最小约简集 $R = \text{Core}(C)$, 然后求核心集对应的等价属性, 并在可分辨矩阵中去掉其等价属性对应的值, 然后选择 $|M(c_k)|$ 最大值的属性 c_k 加入最小约简集。

2.2 决策树算法模型—C4.5 算法

由问题定义可知, 在病例集 $S = \{U, C, D, V, f\}$ 中, 假设它的一个划分为 $\{X_1, X_2, \dots, X_n\}$, 其中 $p_i = P(x_i)$, 则有:

$$\text{Entropy}(S) = - \sum_{i=1}^m p_i \log_2(p_i) \quad (6)$$

其中, Entropy 为病例集 S 以 D 中某一属性 D_j 为决策属性时的信息熵; p_i 为 U 中任意元组属于分组 X_i 的概率, 即 $p_i = P(x_i) = \frac{|X_i|}{|U|}$ 。

信息熵最早由香农提出, 表示物理学上信息的不确定度, 而在数学上则表示为信息冗余度和概率之间的关系。信息熵增益表示病例集 S 按照某条件属性 C 划分时造成熵减少的期望, 如下:

$$\text{Gain}(S, C) = \text{Entropy}(S) - \sum_{v \in V_c} \frac{|S_v|}{|S|} \text{Entropy}(S, C) \quad (7)$$

其中, V_c 表示某一条件属性 c 的值域; S_v 表示病例集 S 中属性 C 划分时, 属性 C 取值为 v 的子集合的信息熵。

ID3 算法以信息熵增益作为条件属性选择的依据, 更倾向于选择属性取值多的属性作为决策树的节点, 而某些情况下这些节点并不是特别有价值。而 C4.5 算法用信息熵增益率代替了信息熵增益作为选择的标准, 克服了 ID3 算法的不足。信息熵增益率计算方法如下:

$$\text{GainRatio}(S, C) = \frac{\text{Gain}(S, C)}{\text{SplitInformation}(S, C)} \quad (8)$$

其中, $\text{SplitInformation}(S, C)$ 表示按照条件属性 C 划分病例集的广度和均匀性, 称为分裂因子, 其公式为:

$$\text{SplitInformation}(S, C) = - \sum_{v \in V_c} \frac{|S_v|}{|S|} \log_2 \frac{|S_v|}{|S|} \quad (9)$$

决策树算法是基于属性频率的, 所以 ID3 算法和基于粗糙集决策树无法处理好连续的实数型参数, 因此要对连续实数型属性进行离散化。假设条件属性 C_1 是连续实数型的, 且样本总数为 n , 则设它的取值序列为 $V_{C_1} = \{v_1, v_2, \dots, v_n\}$ 。C4.5 算法将对它做如下处理: 万方数据

步骤 1: 将 $V_{C_1} = \{v_1, v_2, \dots, v_n\}$ 进行排序。

步骤 2: 对已经排序好的取值序列, 以相邻两个元素的和作为断点, 如下:

$$v = (v_i + v_{i+1})/2 \quad (1 \leq i \leq n) \quad (10)$$

步骤 3: 总共取了 $n-1$ 个断点, 对决策系统而言, 单个断点将病例集分成条件属性 $C_1 \leq v$ 和 $C_1 > v$ 两部分, 对每个断点计算 $\text{GainRatio}(S, v)$ 的值, 选 $\text{GainRatio}(S, v)$ 最大的那个作为区分属性 C_1 的分割阈值, 从而将属性 C_1 离散化。

为了避免由于训练集 S 过多, 加大树的深度, 树过饱和, 而导致树的准确率下降, 采用悲观剪枝的方法对其进行剪枝。

3 算法设计

3.1 算法描述

基于可分辨矩阵的粗糙集属性约简的 C4.5 智能诊断算法包括两部分: 基于属性依赖改进的可分辨矩阵属性约简算法, 给定病例集 S , 对 S 进行属性约简, 得到最小属性约简集; 然后将 S 进行简化, 即只留下最小约简集的属性 and 决策属性, 得到简化后的病例集 S , 然后用 C4.5 算法构建决策树, 提取关于 S 的规则。

算法 1: 基于属性依赖改进的可分辨矩阵的属性约简算法。

输入: U , 原始病例样本集, 待属性约简;

输出: U , 属性约简后的样本实例集。

步骤 1: 求基于论域 U , 条件属性 C 的正域。通过式[3]找出所有不可约分关系, 并求不可约分关系的交集即为核心集 $\text{Core}(C)$ 。设最小约简集 $R = \text{Core}(C)$;

步骤 2: 求 $\text{Core}(C)$ 的等价关系, 并从条件属性和样本集 U 中去掉相关的值;

步骤 3: 构建二进制可分辨矩阵, 选择 $|M(c_k)|$ 最大值的条件 c_k 加入最小约简集;

步骤 4: 如果 R 满足知识完备性, 则得到最小约简集, 去掉不是约简集中数据的样本实例集 U , 不然返回步骤 3。

算法 2: 基于 C4.5 的决策树构造算法。

输入: U , 属性约简后待训练的病例样本实例集;

输出: Tree , 构建好的基于 C4.5 算法的决策树 $\text{C4.5_DTree}(U)$ 。

步骤 1: 计算病例集的每个条件属性的取值范围 V_{C_i} ;

步骤 2: 如果当前的样本集 U 对于决策属性的取值全部相同, 则将叶子节点赋值为该决策属性的取值, 或者样本集的条件属性集为空, 则叶子取值为决策属性取值比较多的值, 并返回, 不然转入步骤 3;

步骤 3:计算每个条件属性 C_i 的信息熵增益,首先判断是否是连续性取值,如果是,按照 2.2 节中的算法求其基于决策属性 D_j 的信息增益率,如果条件属性为离散值,直接求该属性对于决策属性 D_j 的信息熵增益率;

步骤 4:取信息熵增益率最大的条件属性 C_{\max} 作为决策树的节点,并从条件属性集中删除属性 C_{\max} ;

步骤 5:根据条件属性 C_{\max} 的取值对样本集 U 进行划分,并返回步骤 2;

步骤 6:根据样本集对决策树进行悲观剪枝。

3.2 算法改进

决策树建立在已知病例集上,一课决策树的预测和分析可能会不太准确,为了提高准确率,文中构建了决策树森林,即将属性约简之后的病例集随机划分成多个病例集,然后每个病例集生成多棵决策树,多棵决策树构成森林。由于构建每棵决策树的病例集是独立的,所以决策树之间没有关联,当有一条新的病例数据产生时,让森林里的每一棵决策树分别进行判断,以概率最大的结果作为决策结果,可以提高准确率。

4 性能评价与系统仿真

为验证文中算法对医疗诊断推理的正确性和有效性,取某医疗系统关于肺癌诊断的一些电子病例集,对基于属性依赖的可分辨矩阵属性约简算法与可分辨矩阵属性约简算法进行性能比较,将单纯的 C4.5 算法、属性约简与单棵 C4.5 决策树、属性约简和 C4.5 决策树随机森林进行性能比较。

4.1 仿真数据与仿真指标

采用某医疗系统关于肺癌手术后是否康复或者寿命只为 1 年的医疗数据进行分析,该数据总共有 870 条记录,有 16 个条件属性,1 个决策属性,如表 1 所示。

表 1 属性表			
属性名	含义	属性名	含义
DGN	肿瘤类型	FCV	强制肺活量
FCV1	强制呼气 1 s 结束时的音量	PS	性能状态
PBF	术前是否疼痛	HBS	术前是否出血
DBS	术前是否呼吸困难	CBS	术前是否咳嗽
WBS	术前是否有弱点	SOT	肿瘤的大小等级
2DM	是否是 2DM 糖尿病	MI	MI 是否高达 6 个月
PAD	是否有外周动脉疾病	Smoking	是否吸烟
Asthma	是否哮喘	AGE	手术年龄

以编写的 C++ 程序作为测试工具,主要从分类的正确性和时间消耗这两个方面对肺癌数据的分析进行评价。

4.2 仿真结果

在仿真中,为验证文中提出算法的时间性能,将其

与可分辨矩阵属性约简算法进行性能比较,如表 2 所示。

表 2 属性约简算法的实验结果比较		
算法	约简后的属性个数	时间消耗/s
基于可分辨矩阵	8	1.382
属性依赖改进	8	0.847

这个数据集的核属性有 5 个。从表 2 可知,两种算法在肺癌的诊断数据上得到的最佳简约集相同,但是基于属性依赖的可分辨矩阵属性约简算法比可分辨矩阵属性约简算法的时间消耗小,前者大约是后者的 0.6 倍左右。

表 3 列举了约简后的部分数据集。

表 3 属性约简之后的部分数据集							
DGN	FCV	PS	CBS	WBS	SOT	2DM	Smoking
DGN2	2.88	PRZ1	T	T	OC14	F	T
DGN5	2.68	PRZ0	T	F	OC12	F	T
DGN3	2.76	PRZ1	T	F	OC11	F	T
DGN6	3.96	PRZ0	F	F	OC11	F	T
DGN3	2.44	PRZ2	T	T	OC11	F	T
DGN4	3.32	PRZ0	F	F	OC12	F	T
DGN8	4.32	PRZ0	F	F	OC11	F	F

为验证属性约简之后用 C4.5 算法以及 C4.5 加随机森林的准确性,进行了规则提取的仿真。从 870 条记录中随机抽取 120 条作为检验正确性的测试集,然后用剩下的 750 条做训练集。对于单纯的 C4.5 算法,未经过属性约简,直接将原始的病例集作为训练集,构建决策树,构建好的决策树的节点有 236 个,规则有 91 条,准确率为 80.8%,比较低,耗时为 2.326 s。

属性约简后的训练集,用 C4.5 算法构建单棵决策树,得到 146 个节点,78 条规则,正确率为 91.7%,有比较大的提升,耗时减少为 1.843 s。

属性约简和单棵 C4.5 决策树直接将这 750 条测试数据当作训练集即可,而对于属性约简和 C4.5 决策树随机森林,该算法要先进行属性约简,然后将约简后的 750 条数据随机抽样出 300 条记录,抽样成 5 份训练集,然后测试集进行测试时走每棵决策树的结果,将以占比大的推理结果作为结果,时间消耗比单棵决策树有所提高,为 2.048 s,但正确率提升为 94.2%。

表 4 决策树算法的实验结果比较		
算法	时间消耗/s	准确率/%
C4.5 算法	2.326	80.8
属性约简+C4.5	1.843	91.7
属性约简+C4.5+随机森林	2.048	94.2

实验结果表明,属性约简之后再 用 C4.5 算法,减

少了时间消耗,提高了肺癌诊断的正确性,而且耗时短。说明去掉那些不必要的因素可以很好地提高推理的正确性,而且也防止了决策树因训练集太多,导致过饱和,而使正确性下降。从表 4 可以看出,加入随机森林之后,医疗诊断的准确度有所提升,这很好地避免了由于训练集过多而使决策树节点过多,从而导致不准确的情况。该算法使肺癌诊断的准确性上升了一个台阶。

5 结束语

对肺癌诊断数据进行研究与分析,着重于医疗数据属性冗余度大的问题,提出了一种基于属性依赖的可分辨矩阵的属性约简算法,找出最佳约简集,减少了很多不必要的属性。仿真结果表明,与基于可分辨矩阵的属性约简算法比较,在保证信息量的前提下,减少了算法时间开销。对快速提出准确规则问题,文中将属性约简后的训练集用 C4.5 算法进行规则提出,相比直接进行 C4.5 算法,准确率提高很多,时间消耗也降低了。为了提高分类的准确性,提出了基础 C4.5 算法的随机森林,实验结果表明,该算法提高了肺癌诊断结果的准确性。该方法不只可以用在肺癌数据的分类推理中,也可以用在其他病医疗数据中;但是文中本质上是根据医疗记录对肺癌数据进行的分类训练,对其他未知的情况无法给出明确的解答,这个问题还要在后期研究工作继续探索。

参考文献:

[1] RUEDA A, GONZÁLEZ F A, ROMERO E. Extracting salient brain patterns for imaging-based classification of neurodegenerative diseases[J]. IEEE Transactions on Medical Imaging, 2014, 33(6):1262-1274.

+++++

(上接第 147 页)

al ACM conference on embedded networked sensor systems. New York: ACM Press, 2004:39-49.

[6] PIERCE A D, SMITH P W. Acoustics: an introduction to its physical principles and applications[J]. Physics Today, 1981, 34(12):1826-1827.

[7] 丁振华, 李锦涛, 冯波. 基于 Hash 函数的 RFID 安全认证协议研究[J]. 计算机研究与发展, 2009, 46(4):583-592.

[8] OHKUBO M, SUZUKI K, KINOSHITA S. Cryptographic approach to “privacy-friendly” tags [EB/OL]. [2012-03-22]. <http://www.rfidprivacy.org/papers/ohkubo.pdf>.

[9] 雷辉, 姜卫东, 郭勇. 能量高效的水声传感器网络多跳非均匀分簇算法[J]. 计算机应用, 2013, 33(1):124-126.

[10] KOPETZ H, OCHSENREITER W. Clock synchronization in

[2] INUIGUCHI M. Attribute reduction in variable precision rough set model[J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2011, 14(4):461-479.

[3] 王靖, 王兴伟, 赵悦. 基于变精度粗糙集决策树垃圾邮件过滤[J]. 系统仿真学报, 2016, 28(3):705-710.

[4] SZELAG M, GRECO S, SŁOWIŃSKI R. Variable consistency dominance-based rough set approach to preference learning in multicriteria ranking [J]. Information Sciences, 2014, 277(2):525-552.

[5] 鄂旭, 任骏原, 毕嘉娜, 等. 基于粗糙变精度的食品安全决策树研究[J]. 计算机技术与发展, 2014, 24(1):242-245.

[6] 徐兴梅, 曹丽英. 基于粗糙集和 BP 神经网络的粮食产量预测研究[J]. 东北农业大学学报, 2014, 45(10):95-100.

[7] 贾晓芬, 赵佰周, 周孟然, 等. 基于粗糙集约简的图像插值方法[J]. 计算机应用研究, 2015, 32(2):623-626.

[8] 高举红, 赵天一. 基于粗糙集理论的危险品运输风险分析[J]. 安全与环境学报, 2015, 15(1):40-43.

[9] 蒋亚坤, 李文云, 赵莹, 等. 粗糙集与证据理论结合的电网运行优质性综合评价[J]. 电力系统保护与控制, 2015, 43(13):1-7.

[10] SKOWRON A, RAUSZER C. The discernibility matrices and functions in information system[J]. Theory and Decision Library, 2012, 11:331-362.

[11] 常梨云, 王国胤, 吴渝. 一种基于 Rough Set 理论的属性约简及规则提取方法[J]. 软件学报, 1999, 10(11):1206-1211.

[12] 常志玲, 周庆敏. 基于变精度粗糙集的决策树优化算法研究[J]. 计算机工程与设计, 2006, 27(17):3175-3177.

[13] ZHAO Guoying, HUANG Xiaohua, MATTI T, et al. Facial expression recognition from near-infrared videos[J]. Image and Vision Computing, 2011, 29(9):607-619.

[14] 王珏, 王任, 苗夺谦, 等. 基于 Rough Set 理论的“数据浓缩”[J]. 计算机学报, 1998, 21(5):393-400.

+++++

distributed real-time systems[J]. IEEE Transactions on Computers, 1987, 36(8):933-940.

[11] LI B, ZHOU S, STOJANOVIC M, et al. Multicarrier communication over underwater acoustic channels with nonuniform doppler shifts [J]. IEEE Journal of Oceanic Engineering, 2008, 33(2):198-209.

[12] 邓淼磊, 马建峰, 周利华. RFID 匿名认证协议的设计[J]. 通信学报, 2009, 30(7):20-26.

[13] CHEN Y, CHOU J S, SUN H M. A novel mutual authentication scheme based on quadratic residues for RFID systems [J]. Computer Networks, 2007, 52(12):2373-2380.

[14] 邢芙宁. 水下传感器网络时钟同步算法的研究[D]. 哈尔滨: 哈尔滨工程大学, 2013.