

文章编号: 1001-0920(2012)03-0369-05

## 分组选择聚类融合算法

孔志周, 蔡自兴

(中南大学 信息科学与工程学院, 长沙 410083)

**摘要:** 针对聚类融合算法可能出现信息失真等问题, 提出一种新的聚类融合算法. 该算法兼顾聚类质量与成员多样性, 采用一种新的相似性度量, 并依据度量结果先对聚类成员进行剪辑操作, 再分组、选择, 最后根据每个聚类成员对于各类别的贡献设计了一种新的加权函数. 与其他方法相比, 该方法具有较好的稳定性和精确性.

**关键词:** 聚类融合; 剪辑; 分组选择; 加权函数

**中图分类号:** TP273

**文献标识码:** A

### Sub-grouping and selecting method of cluster fusion

KONG Zhi-zhou, CAI Zi-xing

(College of Information Science and Engineer, Central South University, Changsha 410083, China. Correspondent: KONG Zhi-zhou, E-mail: zhizk@yahoo.com.cn)

**Abstract:** For information distortion of cluster fusion algorithm, a new cluster fusion algorithm is proposed, which gives consideration to both the cluster quality and the diversity of cluster members, and adopts a new method of similarity measurement. According to the measure result, cluster members are edited, subgrouped and selected, finally a new weighting function is designed based on the contribution of each cluster member to each category. The practice results show that the algorithm has better stability and precise compared with other methods.

**Key words:** cluster fusion; edit; sub-group and select; weighting function

### 1 引言

聚类融合困难在于如何从不同输出结果中求得共识, 目前已被证明是NP完全型难题. Topchy等人<sup>[1]</sup>总结出其关键问题是求解共识函数, 即怎样解决聚类结果的匹配问题? 怎样保证所有划分分量对称和无偏差地一致? 怎样组合不同的聚类结果?

现有的聚类融合算法根据聚类输出结果分为“硬划分”(输出结果是类标签)和“软划分”(输出结果是属于某类的隶属度). 但客观世界中大量存在着界限不明确的聚类问题, “硬划分”聚类融合将对象分类为不相交的簇是不可接受的. 目前, 大部分研究主要集中在“硬划分”聚类融合方面, 鲜有“软划分”聚类融合成果. 文献[2]比较了聚类成员采用“硬划分”与“软划分”的不同. [3]提出了一种累积投票的方法. [4]将3种模糊相似度计算方法结合起来设计一种模糊共识函数. [5]指出未来“软划分”的聚类融合是值得研究的.

文献[6]主要论述了适度的差异性能组合出更好

的聚类结果, 并认为在组合方式多样化的同时, 如何度量多分类器系统中各聚类成员之间的关系、预测它们之间的相互结合能力将逐渐成为一项重要的研究课题. [7]提出“选择性集成(selective ensemble)”的概念, 并证明了通过选择部分个体学习器来构建集成要优于使用所有个体学习器构建的集成, 这意味着利用中小规模的选择性集成即可获得较好的性能. [8]研究了聚类成员相似性与融合精度的关系, 给出了基于该分析的3种融合方式(基于联合目标函数方法、聚类选择方法、基于kappa统计的凸壳方法), 通过实验表明聚类选择方法是最优的, 并指出用其他相似度量来进行聚类选择是未来的研究方向. [9]总结了差异性度量, 并提出了一个基于权的共识函数. [10]提出了一种自适应的投票方法. [11]设计了一种共识矩阵, 并根据该矩阵计算聚类成员的差异动态来选择聚类成员.

本文在文献[6-13]的基础上, 兼顾聚类质量与成

收稿日期: 2010-07-14; 修回日期: 2010-12-14.

基金项目: 国家基础研究项目(A1420060159); 国家博士点基金项目(200805330005); NSFC重大专项基金项目(90820302).

作者简介: 孔志周(1974—), 男, 讲师, 博士, 从事信息融合、数据挖掘的研究; 蔡自兴(1938—), 男, 教授, 博士生导师, 从事人工智能、机器人学等研究.

员多样性,采用一种新的相似性度量,并依据度量结果先对聚类成员进行剪辑操作,再分组、选择,最后借鉴文献[6]的成果提出了一种新的加权函数.与其他方法相比,该方法具有较好的稳定性与精确性.

## 2 算法组成

### 2.1 过量聚类成员产生和聚类成员匹配

#### 2.1.1 过量聚类成员产生方法

选择“软划分”聚类方法过量生成聚类成员,如 *fuzzy c means*. 在生成过程中可采用多种方法来提高聚类成员的差异性和子空间的覆盖,如不同的聚类算法、不同聚类参数、不同数据子集、不同特征子集等.

#### 2.1.2 聚类成员匹配的分析

设有3个聚类成员,聚类结果分别是3,3,4类,每类有多个样本对象装在以  $C_1, C_2, C_3, C_4$  为中心的箱体里,如图1所示.

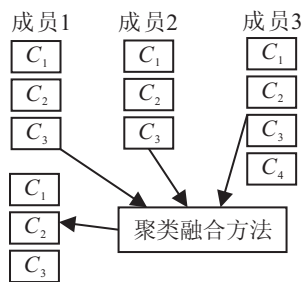


图1 聚类成员产生与融合

相比通常的基于共协矩阵的匹配方法,本文方法由于只选取各聚类成员的中心输出向量来代替共协矩阵中的标记输出向量,可有效减少运算次数,且降低了噪声数据的干扰,同时也有利于处理各聚类成员输出类别数目不等的情况. 设聚类成员1输出为3类别,中心分别是  $C_{11}, C_{12}, C_{13}$ ; 聚类成员2输出为3类别,中心分别是  $C_{21}, C_{22}, C_{23}$ ; 聚类成员3输出为4类别,中心分别是  $C_{31}, C_{32}, C_{33}, C_{34}$ . 将这些聚类成员间各类别中心相互匹配的所有可能用线连接起来,如图2所示. 如果再在左端和右端分别加上起点和终点,则可以发现,聚类成员匹配问题可以近似于TSP (traveling salesman problem) 问题. 其实质如下: 设有  $r$  个聚类成员,匹配的目标是找到诸如  $(c_1, c_2, \dots, c_i, \dots, c_r)$  若干个链,使得各聚类成员的某一个类别能一一对应. 设多数聚类类别数为  $k$ ,则匹配问题就是要找到  $k$  条这样的链. 要找到全局最优的聚类成员匹配,其实质是找到下式表达的目标函数的最优解:

$$\varphi^* = \arg \min \left( \sum_{l=1}^k \sum_{i=1}^r \sum_{j=1}^r d(c_i^l, c_j^l) \right), \quad (1)$$

其中  $d(c_i^l, c_j^l)$  为两个类别中心的距离.

在可以求解如式(1)所示的约束优化问题的算法中,蚁群算法是一较有发展前景的计算智能方法,

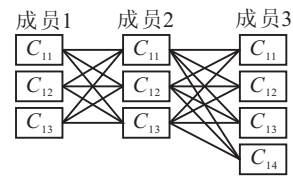


图2 聚类成员匹配问题

且是一种鲁棒算法<sup>[14]</sup>. Dorigo<sup>[15]</sup>给出了3种模型 (ant-cycle system, ant-quantity system, ant-density system), 通过实验和比较分析,认为 ant-cycle system 最好. 从目前的TSP问题研究来看,该方法仍然是最好的方法之一,因此本文采用 ant-cycle system 算法.

#### 2.1.3 基于蚁群算法的聚类成员匹配方法

通常的基于共协矩阵的匹配方法不能处理“软划分”聚类方法的输出向量,本文采用基于蚁群算法的匹配方法解决该问题. 两者的区别主要在于: 该匹配过程是针对聚类中心而不是聚类成员对于各样本的输出向量. 本文的 ant-cycle system 算法主要解决每个聚类成员输出类别相同的情况,根据图论知识和聚类成员匹配问题的特点,将聚类成员匹配模型描述为一种赋权偶图. 令  $c_x = (c_1, c_2, \dots, c_i, \dots, c_n)$  为聚类成员  $X$  的全体中心集合,  $c_y = (c_1, c_2, \dots, c_j, \dots, c_n)$  为聚类成员  $Y$  的全体中心集合,连接边  $(c_i, c_j)$  的权值  $w_{ij} = d(c_i, c_j)$  表示两聚类中心的距离,则可以得到1个赋权偶图,如图3所示.

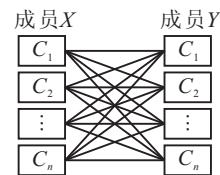


图3 聚类成员匹配模型

在图3中,至少存在一个约束条件: 保证一个聚类成员中的一个聚类中心只能与另一个聚类成员的一个聚类中心相匹配. 假设有  $r$  只蚂蚁,蚂蚁每走一步代表一次聚类成员作用于另一聚类成员的过程. 当所有蚂蚁完成一次旅行后视为算法循环一次. 为了能够利用蚁群算法求解赋权偶图,必须对 ant-cycle system 算法进行局部改进.

**Step 1:** 蚂蚁的行进策略. 首先随机生成聚类成员1的聚类中心顺序,蚂蚁由聚类中心  $c_i$  出发访问聚类成员2的聚类中心  $c_j$ ,在返回聚类成员1的聚类中心  $c_{i'}$  后再访问聚类成员2的聚类中心  $c_{j'}$ . 按此循环,直到遍历完所有聚类成员1的聚类中心. 重复此操作,直至遍历完所有的聚类成员.

**Step 2:** 设置禁忌表来实现匹配的两个约束条件.

**Step 3:** 计算蚂蚁  $k$  由一个聚类成员的中心  $c_i$  到另一个聚类成员的中心  $c_j$  的转移概率

$$p_{ij}^k(t) = \begin{cases} \frac{(\tau_{ij}(t))^\alpha (\eta_{ij}(t))^\beta}{\sum_{k \in S} (\tau_{ij}(t))^\alpha (\eta_{ij}(t))^\beta}, & j \in S; \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

其中:  $S$  为允许选择中心的可行解集;  $\tau_{ij}(t)$  为  $t$  时刻路径  $(c_i, c_j)$  上的信息量;  $\alpha$  为残留信息的相对重要程度;  $\beta$  为启发信息的相对重要程度;  $\eta_{ij}(t)$  为由  $c_i$  转移到  $c_j$  的期望程度, 可以根据某种启发式算法具体确定, 定义为  $\eta_{ij}(t) = 1/d(c_i, c_j)$ .

**Step 4:** 随着时间的推移, 前期留下的信息逐渐消逝, 用参数  $(1 - \rho)$  表示信息消逝程度,  $\rho \in [0, 1)$ . 所有蚂蚁完成一次循环后, 根据信息素全局更新规则进行调整, 即

$$\tau_{ij}(t+n) = (1 - \rho)\tau_{ij}(t) + \Delta\tau_{ij}(t), \quad (3)$$

$$\Delta\tau_{ij}(t) = \sum_{k=1}^r \Delta\tau_{ij}^k(t). \quad (4)$$

其中:  $\Delta\tau_{ij}(t)$  为本次循环中路径  $(c_i, c_j)$  上的信息素增量;  $\Delta\tau_{ij}^k(t)$  为第  $k$  只蚂蚁在本次循环中留在路径  $(c_i, c_j)$  上的信息素增量, 有

$$\Delta\tau_{ij}^k(t) = \begin{cases} Q/L_k, & L_k \text{ include}(c_i, c_j); \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

$Q$  为常数,  $L_k$  为第  $k$  只蚂蚁在循环中所走路程的长度.

## 2.2 兼顾聚类质量与聚类成员相似性度量

文献[8]研究了聚类成员相似性与融合精度的关系, 指出用其他相似性度量来进行聚类选择是未来的研究方向. 受[4]启发, 选择[12-13]的研究成果——扩展的直觉模糊集算术平均最小法. 具体算法如下: 设  $A$  和  $B$  分别表示 2 个聚类成员,  $x_i \in U, i = 1, 2, \dots, n, U$  为数据集. 为了描述方便, 假设输出类别为 2, 则  $A$  和  $B$  的输出向量分别为

$$A = \sum_{i=1}^n \langle \mu_A(x_i), \gamma_A(x_i) \rangle / x_i, \quad (6)$$

$$B = \sum_{i=1}^n \langle \mu_B(x_i), \gamma_B(x_i) \rangle / x_i. \quad (7)$$

$A$  与  $B$  之间的相似度为

$$r(A, B) = r(\mu_A, \mu_B) + r(\gamma_A, \gamma_B), \quad (8)$$

$$r(\mu_A, \mu_B) = \frac{\sum_{i=1}^n \min(\mu_A(x_i), \mu_B(x_i))}{\sum_{i=1}^n (\mu_A(x_i) + \mu_B(x_i))}, \quad (9)$$

$$r(\gamma_A, \gamma_B) = \frac{\sum_{i=1}^n \min(\gamma_A(x_i), \gamma_B(x_i))}{\sum_{i=1}^n (\gamma_A(x_i) + \gamma_B(x_i))}. \quad (10)$$

可计算聚类成员  $A$  与其他  $L - 1$  成员的相似性度量为

$$\text{Sr}(A) = \sum_{i=1}^L r(A, A_i). \quad (11)$$

直观上, 使得式(11)最大化的  $A$  表明其与其他聚类成员共享信息最多, 也意味着  $A$  是最好的信息携带者.

## 2.3 聚类成员的剪辑与分组选择

过量产生聚类成员的目的在于选择较优的聚类成员参与到聚类融合中. 在此过程中, 不可避免地产生一些相对于整个系统几乎没有作用的聚类成员, 认为其是噪声聚类成员, 可以通过剪辑操作将这些噪声聚类成员剔除. Hadjitodorov<sup>[6]</sup>认为, 只有适度的差异性才能组合出更好的聚类结果, 根据前文提出的相似性度量, 采用剪辑操作去除具有最高和最低相似性的聚类成员. 但是即使这样, 仍然可能存在过于相似的聚类成员. 假设两个相似度很高的聚类成员  $A$  与  $B$ ,  $A$  已经包含在融合系统中, 显然, 若再将  $B$  包含进来, 则系统会出现冗余. 也可能出现这种情况, 系统中的很多聚类成员与  $B$  相似度低, 而与  $A$  相似度高, 这时, 也可能会将  $B$  选入.

鉴于该情况, 采取分组选择的方法. 分组可以采用分箱技术或聚类方法, 操作的对象是各聚类成员的两两相似度. 在每组中挑选与其他  $L - 1$  成员相似性最高的那个聚类成员参与融合. 当然, 也可以从每组中选取最高的两个或更多个成员, 这取决于总聚类成员数、组数的关系. 例如, 总聚类成员数为  $2K$ , 现分成  $K$  组, 每组取最高的两个参与.

## 2.4 加权函数设计

不同的聚类成员在融合系统中所占的地位也是不同的. 文献[5-6,8-11]均认为差异性应该适度地体现到相应聚类成员的权上. 文献[9]的分析与实验表明, 当聚类成员在子空间有不同覆盖时, 其携带的信息具有互补性, 最好的信息携带者应该在系统中扮演重要的角色. 借鉴有监督的分类器融合方法, 本文采用聚类成员相对于系统各类别的重要性度量作为对该类别判定的权重, 则聚类成员  $A$  对于类别的权重为

$$\text{Sr}(\mu_A) = k \sum_{i=1}^L r(\mu_A, \mu_i), \quad (12)$$

其中  $k = 1 / \sum_{i=1}^L \text{Sr}(\mu_i)$  为标准化系数. 因此, 选择融合结果大的类别作为最后聚类融合的判定结果.

## 3 实验结果与分析

### 3.1 实验数据与性能评价方法

#### 3.1.1 测试数据的选取

实验数据使用 UCI 库数据, 为了得到可靠的实验结果, 实验根据以下准则选取数据: 1) 数据集的数据元素不能太少. 若太少, 则每次从训练数据集抽取的子集可能无法体现原数据的概率分布, 从而影响整个模型的稳定性. 2) 数据集的可分性不是很高. 若很高,

则学习得到的聚类性能会很高, 分类器间的差异性就会很小, 这样, 数据集引起的高准确率就会掩盖融合方法的效果. 依据上述原则, 选择 UCI 机器学习库中有代表性的 6 个数据集进行聚类融合的实验测试. 表 1 给出了所用数据的构成信息.

表 1 实验数据的构成信息

数据集	指标数	样本数	类别数
Wine	13	178	3
Liver disorder	6	345	2
Iris	4	150	3
Ionosphere	34	351	2
Breast	9	478	2
Glass	9	214	6

### 3.1.2 聚类性能评价方法

假设大小为  $n$  的数据集  $X$  中存在的已知类别标记为  $C$ , 用  $\{C_1 C_2 \cdots C_k\}$  表示, 通过某个聚类器对该数据集进行聚类得到的结果为具有  $k$  个聚类标记的标记向量  $\lambda$ , 用  $\{\lambda_1 \lambda_2 \cdots \lambda_k\}$  表示. 通过上述的匹配过程, 聚类标记向量中的每一个  $\lambda_i$  对应类别标记中的某一个  $C_j$ . 假设  $\nu_i$  为  $\lambda_i$  中被正确分类为对应类别  $C_j$  的个数, 那么聚类标记向量的结果可用  $\delta$  进行衡量, 具体定义为

$$\delta = \frac{1}{n} \sum_{i=1}^k \nu_i. \quad (13)$$

$\delta$  值越大, 表明聚类的准确度越高. 本文实验采用该方法来评价产生固定聚类个数的聚类器的准确度.

### 3.2 实验结果和分析实验数据与性能评价方法

由于 Tsaipei 等人<sup>[2,5]</sup>多次比较了“硬划分”与“软划分”, 本文不再对此进行比较. 为了更好地评估本文算法的性能, 将本文算法与 MM 方法<sup>[2]</sup>、SMLC 方法<sup>[5]</sup>进行比较分析.

实验采用的聚类器聚类时所设定的类别数仍然是固定的, 采用 fuzzy  $c$ -means 聚类方法, 用重采样方法生成聚类成员. 设置最大迭代步数为 150, 误差停止阈值为  $1e-5$ , 模糊数为 1.5. 蚁群算法参数如下:  $Q = 1$ ,  $\alpha = 4$ ,  $\beta = 1$ ,  $\rho = 0.8$ , 蚂蚁数  $r = m$  (总的中心数),  $NCMax = 50$ . 本文算法的分组选择策略如下: 每组有 4 个成员, 抽出 2 个参与融合. 对于每个数据集, 用上述的 3 种聚类融合算法分别构造大小为 8, 16, 24 和 32 规模的融合, 并对每一种融合重复 10 次实验, 计算其几何平均性能<sup>[5]</sup>.

为了更清楚地表明 3 种方法的性能, 根据它们在各数据集上的性能几何均数, 给出聚类融合的平均性能随成员规模变化的关系如表 2 所示.

从总体上看, 本文方法与其他两种方法相比均具有一定优势, 表明本文方法采用各聚类成员的中心

表 2 聚类性能  $\delta$ 

成员规模	MM方法	SMLC方法	本文方法
8	0.728±0.043	0.689±0.067	0.759±0.040
16	0.765±0.032	0.717±0.043	0.783±0.034
24	0.826±0.040	0.736±0.075	0.821±0.036
32	0.868±0.027	0.758±0.62	0.874±0.032

来代替常用的标记输出向量, 并采用蚁群算法进行成员匹配是可行的. 由于标记输出向量是针对每个样本对象的, 直观上可有效减少运算次数, 降低噪声数据的干扰, 也便于处理各成员输出类别数目不等的情况. 从规模为 8, 16, 24, 32 个成员的结果看, 大多数情况下本文算法优于 MM 方法, 至少不低于 MM 方法; 从 8 到 32 个成员规模上看, 本文方法均大幅优于 SMLC 方法. 因此, 本文方法具有较好的稳定性与精确性.

由表 2 还可以看出, 随着聚类成员规模的增加, 本文方法的规模优势得到了体现, 尤其是规模达到 32, 即参与的成员为 16 个时, 性能有较大的提高, 表明该方法采用剪辑、分组选择的策略是有效的. 同时, 也验证了文献 [7] 提出的“选择性集成”的概念, 表明通过选择性聚类形成的中小规模融合即可获得较好的性能.

## 4 结 论

本文提出了一种新的聚类融合算法. 该算法兼顾聚类质量与成员多样性, 采用一种新的相似性度量, 依据度量结果先对聚类成员进行剪辑操作, 再进行分组、选择, 最后根据每个聚类成员对于各类别的贡献设计了一种新的加权函数. 与其他方法相比, 该方法具有较好的稳定性与精确性.

### 参考文献(References)

- [1] Topchy A, Jain A K, Punch W F. Combining multiple weak clusterings[C]. Proc of the 3rd IEEE Int Conf on Data Mining. Florida: IEEE Computer Society Press, 2003: 331-338.
- [2] Tsaipei Wang. Comparing hard and fuzzy  $C$ -means for evidence-accumulation clustering[C]. Proc of the 18th Int Conf on Fuzzy Systems. Korea: IEEE Press, 2009: 468-473.
- [3] Ayad H G, Kamel M S. On voting-based consensus of cluster ensembles[J]. Pattern Recognition, 2010, 43: 1943-1953.
- [4] Yang Lin-yun, Lv Hai-rong, Wang Wen-yun. Soft cluster ensemble based on fuzzy similarity measure[C]. Proc of the IMACS Multiconference on Computational Engineering in Systems Applications. Beijing: IEEE Press, 2006: 1994-1997.

- [5] Punera K, Ghosh J. Soft cluster ensembles[C]. *Advances in Fuzzy Clustering and its Applications*. England: John Wiley and Sons Ltd, 2007: 69-90.
- [6] Hadjitodorov S T, Kuncheva L I, Todorova L P. Moderate diversity for better cluster ensembles[J]. *Information Fusion*, 2006, 7(3): 264-275.
- [7] Zhou Z H, Wu J, Tang W, et al. Ensembling neural networks: Many could be better than all[J]. *Artificial Intelligence*, 2002, 137(1/2): 239-263.
- [8] Xiaoli Z F, Wei L. Cluster ensemble selection[J]. *Statistical Analysis and Data Mining*, 2008, 1(3): 128-141.
- [9] Carlotta D, Muna A R. Weighted cluster ensembles: Methods and analysis[J]. *ACM Trans on Knowledge Discovery from Data*, 2009, 2(4): 1-42.
- [10] Kagan T, Adrian K A. Ensemble clustering with voting active clusters[J]. *Pattern Recognition Letters*, 2008, 29(14): 1947-1953.
- [11] Javad A, Xiaoli F. Adaptive cluster ensemble selection[C]. *Proc of the 21st Int Joint Conf on Artificial Intelligence*. California: Morgan Kaufmann Publishers Inc, 2009: 992-997
- [12] 雷英杰, 赵晔, 王涛, 等. 直觉模糊语义匹配的相似性度量[J]. *空军工程大学学报*, 2005, 6(2): 83-87.  
(Lei Y J, Zhao Y, Wang T, et al. On the measurement of similarity on semantic match for intuitionistic fuzzy[J]. *J of Air Force Engineering University*, 2005, 6(2): 83-87.)
- [13] 路艳丽, 雷英杰, 李兆渊. 直觉模糊相似关系的构造方法[J]. *计算机应用*, 2008, 28(2): 311-314.  
(Lu Y L, Lei Y J, Li Z Y, et al. Construction of intuitionistic fuzzy similarity relation[J]. *Computer Applications*, 2008, 28(2): 311-314.)
- [14] 蔡自兴, 徐光佑. *人工智能及其应用*[M]. 北京: 清华大学出版社, 2004.  
(Cai Z X, Xu G Y. *Artificial intelligence and application*[M]. Beijing: Tsinghua University Press, 2004.)
- [15] Dorigo M, Luca Maria Gambardella. Ant colony System: A cooperative learning approach to the traveling salesman problem[J]. *IEEE Trans on Evolutionary Computation*, 1997, 1(1): 53-66.

(上接第368页)

- [7] Au M H, Liu J K, Yuen T H, et al. ID-based ring signature scheme secure in the standard model[C]. *Lecture Notes in Computer Science*. Berlin: Springer-Verlag, 2006, 4266: 1-16.
- [8] 王玲玲, 张国印, 马春光. 一种签名长度固定的基于身份的环签名方案[J]. *电子与信息学报*, 2007, 29(11): 2645-2648.  
(Wang L L, Zhang G Y, Ma C G. An identity-based ring signature scheme with constant-size signature[J]. *J of Electronics & Information Technology*, 2007, 29(11): 2465-2468.)
- [9] 王化群, 于红, 李显强, 等. 两个标准模型中可证安全的环签名方案的安全性分析及其改进[J]. *通信学报*, 2008, 29(8): 57-62.  
(Wang H Q, Yu H, Li X Q, et al. Cryptanalysis and improvement of two ring signature schemes provably secure in standard model[J]. *J on Communications*, 2008, 29(8): 57-62.)
- [10] 张跃宇, 李晖, 王育民. 标准模型下基于身份的环签名方案[J]. *通信学报*, 2008, 29(4): 40-44.  
(Zhang Y Y, Li H, Wang Y M. Identity-based ring signature scheme under standard model[J]. *J of Communications*, 2008, 29(4): 40-44.)
- [11] Zhang F G, Naini R S, Lin C Y. New proxy signature, proxy blind signature and proxy ring signature schemes from bilinear pairings. *Cryptology Eprint Archive*[Z]. 2003/104.
- [12] Lang W M, Yang Z K, Cheng W Q, et al. A new improved ID-based proxy ring signature scheme from bilinear pairings[J]. *J of Harbin Institute of Technology*, 2006, 13(6): 688-691.
- [13] Amit K Awasthi, Sunder Lal. ID-based ring signature and proxy ring signature schemes from bilinear pairings[J]. *Int J of Network Security*, 2007, 4(2): 187-192.
- [14] 周晓燕, 杜伟章. 基于身份和多线性映射的代理环签名方案[J]. *计算机工程*, 2010, 36(7): 165-167.  
(Zhou X Y, Du W Z. Proxy ring signature scheme based on identity and multilinear mapping[J]. *Computer Engineering*, 2010, 36(7): 165-167.)