

基于多准则决策方法的 WSNs 不等簇数据收集算法

宋晓莹, 温涛, 孙伟, 张启龙

(东北大学 软件中心, 辽宁 沈阳 110819)

摘 要: 针对多对一的无线传感器网络“热点”问题, 提出了一种基于多准则决策方法的不等簇数据收集算法 (unequal clustering data gathering algorithm based on multiple criteria decision, UCDGAMCD). 采用直觉模糊层次分析法和层次模糊积分的多准则决策方法来竞选簇首, 提出了一个新的簇首竞争半径, 使其能够适应节点能量异构及节点非均匀分布的网络环境. 根据邻居簇首的剩余能量和传输能耗, 提出了簇首间按比例分配传输数据的路由方式, 使其能量消耗更加均衡. 仿真结果表明 UCDGAMCD 在节点均匀和非均匀分布的两种实验场景中都获得了较长的网络寿命.

关 键 词: 无线传感器网络; 分簇算法; 直觉模糊层次分析法; 多准则决策方法; 网络寿命

中图分类号: TP 393

文献标志码: A

文章编号: 1005-3026(2017)04-0476-05

Unequal Clustering Data Gathering Algorithm Based on Multiple Criteria Decision Making for WSNs

SONG Xiao-ying, WEN Tao, SUN Wei, ZHANG Qi-long

(Software Center, Northeastern University, Shenyang 110819, China. Corresponding author: SONG Xiao-ying, E-mail: songxiaoying@neusoft.edu.cn)

Abstract: Focusing on the hot spot problem in many to one wireless sensor networks, an unequal clustering data gathering algorithm based on multiple criteria decision (UCDGAMCD) was proposed. The intuitionistic fuzzy analytic hierarchy process and multiple criteria decision making of hierarchical fuzzy integral were used to select cluster heads (CHs), and a new cluster head competition radius was proposed to make it fit for the network scenarios with energy heterogeneous and nodes non-uniform distribution. To balance energy consumption, a routing algorithm with data ratio distribution was proposed based on residual energy of CHs and transmission energy consumption. The simulation results demonstrated that UCDGAMCD can obtain higher network lifetime in two scenarios—node uniform and non-uniform deployment.

Key words: wireless sensor networks; clustering algorithm; intuitionistic fuzzy analytic hierarchy process; multiple criteria decision making; network lifetime

目前,随着无线传感器节点的集成化和微型化,节点的电源大多采用能量受限的电池^[1]. 由于节点通常位于无人区域或危险环境中,难以进行能源补给,而可再生能源技术目前尚不成熟. 因此,通过优化系统能耗,最大限度地延长网络寿命成为无线传感器网络面临的首要挑战^[2].

近年来,分簇的数据收集算法是降低能耗延长网络寿命的有效方法之一^[3]. 目前,已有大量的文献研究了不等簇分簇算法. 从采用的方法划

分:一类是基于精确数学的不等簇分簇数据收集算法;另一类是基于模糊数学的不等簇分簇数据收集算法. 无论是哪种算法,它们大多适用于节点均匀分布和初始能量同构的网络场景. 鲜有同时考虑节点随机非均匀分布和能量异构的更贴近现实的网络场景.

在基于精确数学的不等簇分簇数据收集算法中,DCPVP^[4]是一种基于投票和优先级思想的不等簇分簇数据收集算法. 在 DCPVP 中,簇规模由

节点到基站的距离和局部节点的密度决定. 而簇首的选择基于邻居节点间的平均距离、剩余能量和被选择的时间. 实验表明,DCPVP 不仅延长了网络寿命而且大大降低了网络的消息复杂度. 然而,该数据收集算法仅适用于节点同构和均匀分布的网络场景. 虽然, IDUC^[3] 同时考虑的节点有非均匀分布和能量异构的情况,但是,在竞争簇首的过程中,仅考虑了候选簇首的剩余能量,簇首选择因素考虑得不够全面.

在基于模糊数学的不等簇分簇数据收集算法中,文献[5]提出了一种基于多准则决策方法的分簇算法 FAHP. 然而,该分簇算法运用梯形模糊层次分析法确定准则重要性程度时,存在一定的局限性. 因为模糊层次分析法中的模糊集能在一定程度上表示决策者对研究对象的主观判断,但不能精确表达弃权或犹豫不决的情况^[6]. 直觉模糊层次分析法是模糊层次分析法的扩充和发展,在处理上述不确定性方面更具灵活性和实用性.

综上所述,本文提出了一种基于多准则决策方法的分布式不等簇分簇数据收集算法 UCDGAMCD. 首先,采用直觉模糊层次分析法和层次模糊积分的多准则决策方法来竞选簇首. UCDGAMCD 在簇首竞争的过程中,综合考虑多种因素之间的关联关系,避免了使用 OWA 算子时的假设条件,即每个属性都是相互价值独立的. 其次,提出了一个新的簇首竞争半径,使其能够适应节点能量异构及节点非均匀分布的网络环境. 然后,根据邻居簇首的剩余能量和传输能耗,提出了按比例分配传输数据的路由方式,使其能量消耗更加均衡. 最后,仿真结果表明 UCDGAMCD 在两种实验场景中都获得了较长的网络寿命.

1 UCDGAMCD 算法

UCDGAMCD 算法的整个操作被划分为以轮^[7]为时间单位,每一轮分为簇拓扑建立阶段和数据传输阶段.

1.1 簇首选择

在获得节点成为簇首的综合得分过程中,首先,需建立综合评价簇首选择层次结构. 然后,分别求得层次模糊积分模型所需的属性评估值(h)和属性重要性程度(g).

为了提高能量有效性同时考虑到服务质量可靠性的要求,将簇首选择的主准则分为能量状况、服务质量状况及节点位置状况. 除此之外,根据影响簇首选择的多个因素,可建立层次结构.

1.1.1 簇首综合评价的属性评估值(h)

在初始化阶段完成后,每个节点都计算出邻节点的最大和最小剩余能量、邻节点数目及网络中节点与基站的最大最小距离等信息. 通过网络初始化过程,每个节点可以获得计算评估值所需要的参数值. 然后,开始计算这些属性的评估值(h).

1) 剩余能量(E):在无法自由替换电池节点构成的无线传感器网络中,能量是最重要和稀缺的资源. 第 i 个节点剩余能量评估值定义为

$$E_i = \frac{E_r - E_{\min}}{E_{\max} - E_{\min}}. \quad (1)$$

式中: E_{\max} 和 E_{\min} 分别是节点 i 的邻居节点中最大和最小的剩余能量; E_r 是节点 i 的剩余能量.

2) 消息代价(M):消息代价是在簇拓扑建立过程中,由节点获得接收和发送消息的数目来确定. 消息代价直接影响节点能量的消耗,所以,消息代价的评估值定义为

$$M_i = \frac{M_T - M_C}{M_T}. \quad (2)$$

式中: M_C 为节点 i 接收和发送消息的总数目; M_T 为其所有邻居节点接收和发送消息的总数目. M_C 值越大,消息代价 M_i 越低.

3) 覆盖因子(CI):覆盖保持是保证服务质量(QoS)最基本的问题之一. 选择覆盖保持较好的节点作为簇首能够有效延长网络的功能时间^[8],并节省节点的剩余能量. 因此,覆盖因子是判断一个节点服务质量的主要因素,其评估值定义为

$$CI_i = \frac{|C(i)| - |O(i)|}{|\cup_{i=1}^n C(i)|}. \quad (3)$$

式中: $|C(i)|$ 表示节点 i 覆盖邻居节点个数; $|O(i)|$ 表示节点 i 被其他节点覆盖的邻居节点个数; $|\cup_{i=1}^n C(i)|$ 表示覆盖节点 i 邻居节点的总节点数. CI_i 的值越高,节点 i 单独覆盖的节点数越多,该节点的覆盖保持性越好.

4) 链接可靠性(R):在数据传输的过程中,任意节点失效或是包丢失都会导致大量的包丢失. 因此,服务质量考虑可靠的数据递交是十分必要的. 簇首节点主要负责接收成员节点和其他簇首的数据,所以,可靠性高的节点作为簇首的概率应该更大. 节点可靠性的评估值定义为

$$R_i = \frac{B_{\text{ava}}(i)}{B_{\text{total}}(i)}. \quad (4)$$

式中: $B_{\text{ava}}(i)$ 为节点 i 可用的缓冲区空间; $B_{\text{total}}(i)$ 为节点 i 缓存区的总空间. 节点的缓存区可用空间越大,节点 i 的可靠性越高.

5) 邻居节点数(N):在文献[7]中,通过网络中节点的初始信息,可以求得在网络能量消耗最小时的最优簇个数,从而进一步获得最优的簇成员个数. 当一个节点的邻居节点数量越接近最佳邻居节点数时,该节点成为簇首的概率就越大. 邻居节点数评估值定义为

$$N = \frac{|N_i - N_o|}{N_o}.$$

(5)

式中: N_i 为节点 i 的邻居节点数目; N_o 为最佳簇成员数目.

6) 与基站距离(D):节点与基站的距离也是成为簇首需要考虑的关键因素,因为距离与能耗具有直接的关系. 节点与簇首的距离越大,成为簇首的概率相对越小. 因此,与基站距离的评估值定义为

$$D_i = \frac{d_{\max} - d_{(i, \text{Sink})}}{d_{\max} - d_{\min}}.$$

(6)

式中: d_{\max} 为网络中的节点与基站的距离; d_{\min} 为网络中的节点与基站的最小距离; $d_{(i, \text{Sink})}$ 为节点 i 与基站的距离. 当节点 i 与基站的距离越大, D_i 值越小.

1.1.2 确定簇首综合评价的属性重要性程度

首先,请多位专家通过对准则层和子准则各指标对上一目标层的重要性进行两两比较. 综合各位专家的意见,建立直觉模糊互补判断矩阵. 其中,第二层所有属性对目标层的直觉模糊互补判断矩阵为

$$L_1 - L_2 = \begin{bmatrix} 0.5 & (0.8, 0.9) & (0.7, 0.8) \\ (0.1, 0.2) & 0.5 & (0.4, 0.6) \\ (0.2, 0.3) & (0.4, 0.6) & 0.5 \end{bmatrix}.$$

(7)

第三层子准则层各属性对第二层准则层属性的直觉模糊互补判断矩阵为

$$L_{21} - L_3 = \begin{bmatrix} 0.5 & (0.8, 0.9) & (0.5, 0.7) & (0.6, 0.7) \\ (0.1, 0.2) & 0.5 & (0.1, 0.2) & (0.2, 0.3) \\ (0.3, 0.5) & (0.8, 0.9) & 0.5 & (0.5, 0.8) \\ (0.3, 0.4) & (0.7, 0.8) & (0.2, 0.5) & 0.5 \end{bmatrix},$$

(8)

$$L_{22} - L_3 = \begin{bmatrix} 0.5 & (0.4, 0.6) & (0.4, 0.7) \\ (0.4, 0.6) & 0.5 & (0.5, 0.7) \\ (0.3, 0.6) & (0.3, 0.5) & 0.5 \end{bmatrix},$$

(9)

$$L_{23} - L_3 = \begin{bmatrix} 0.5 & (0.2, 0.3) & (0.4, 0.5) \\ (0.7, 0.8) & 0.5 & (0.3, 0.5) \\ (0.5, 0.6) & (0.5, 0.7) & 0.5 \end{bmatrix}.$$

(10)

根据文献[9]中的算法 1,用 Matlab R2013 (b)实现了该算法,并对 4 个模糊互补判断矩阵进行一致性检验,其中, $\lambda = 0.4$, $d = 0.1$. 经过对这些矩阵进行一致性检验及调整,得到上述 4 个矩阵的满意一致性矩阵:

$$F'_{L_{11}} = \begin{bmatrix} 0.5000 & 0.5591 & 0.5659 \\ 0.4819 & 0.5000 & 0.4655 \\ 0.4925 & 0.4669 & 0.5000 \end{bmatrix},$$

(11)

$$F'_{L_{21}} = \begin{bmatrix} 0.5000 & 0.5244 & 0.5333 & 0.5282 \\ 0.4527 & 0.5000 & 0.4475 & 0.4453 \\ 0.5276 & 0.5177 & 0.5000 & 0.5223 \\ 0.5011 & 0.4824 & 0.4988 & 0.5000 \end{bmatrix},$$

(12)

$$F'_{L_{22}} = \begin{bmatrix} 0.5000 & 0.5116 & 0.5025 \\ 0.5156 & 0.5000 & 0.5206 \\ 0.4649 & 0.4586 & 0.5000 \end{bmatrix},$$

(13)

$$F'_{L_{23}} = \begin{bmatrix} 0.5000 & 0.4523 & 0.4563 \\ 0.5104 & 0.5000 & 0.5144 \\ 0.4980 & 0.5353 & 0.5000 \end{bmatrix}.$$

(14)

最后,计算权重. 获得满意一致性矩阵后,根据文献[9]中的公式(1)即可求得第二层中所有属性对目标层的权重 $W = (0.375, 0.3158, 0.3198)$. 同理可得第三层子准则层各属性相对于上一层的各权重及子准则层各属性相对于目标层元素的最终组合权重,如表 1 所示.

表 1 基于直觉模糊层次分析的权重

Table 1 Weights based on intuitionistic fuzzy analytic hierarchy process

属性	L_{21}	L_{22}	L_{23}	对目标层的权重值
	0.375	0.3158	0.3198	
L_{31}	0.2643	0.3380	0.3029	0.3027
L_{32}	0.2242	0.0000	0.0000	0.0841
L_{33}	0.2613	0.3454	0.0000	0.2071
L_{34}	0.0000	0.3078	0.0000	0.0972
L_{35}	0.0000	0.0000	0.3416	0.1092
L_{36}	0.2470	0.0000	0.3444	0.2028

1.2 簇拓扑的建立

1.2.1 综合评价与簇首广播时间的映射

簇首广播采用与文献[5]相同的计时广播机制. 根据节点成为簇首的综合评价得分映射到触发竞选簇首消息的时间轴上. 其目的是使综合评价得分高的节点较快地以簇首竞争半径 R_c 广播竞选簇首的消息 HEAD_MESSAGE. 发布竞争簇首消息的时间:

$$T_i = \delta \times (1 - CS_i) \times T_c.$$

(15)

式中： δ 为 0.9 到 1 的随机数，为了避免消息碰撞； CS_i 为通过层次模糊积分计算的节点成为簇首的综合评价得分； T_c 为预先设置的成簇时间。

1.2.2 簇首竞争半径

在异构的网络环境中，节点具有异构的初始能量。当网络中的节点消耗相同的能量时，具有较低初始能量的节点将过早死亡，从而降低了网络寿命。为了完全体现高能量节点的优势，让其承担更多的任务，在建立不等簇拓扑时，不仅考虑节点与基站的距离，还要考虑节点具有的能量高低。然而，当这些节点与基站的距离相似，而且剩余能量也接近时，除了这两个因素以外，还需要考虑节点非均匀分布的因素。因此，提出了一个新的竞争簇首半径，如式 (16) 所示。使其能够在节点能量异构且呈非均匀分布的网络场景中，高效地构建不等簇拓扑，从而避免能量消耗不均衡的现象。

$$R_{ic} = [1 - \alpha \frac{d_{max} - d_{(i, Sink)}}{d_{max} - d_{min}} - \beta (\frac{E_{max} - E_{ir}}{E_{max}}) - \gamma \frac{n_i}{N}] R_0.$$
 (16)

式中： d_{max} 和 d_{min} 为网络中节点距离基站的最大和最小距离； $d_{(i, Sink)}$ 为节点 i 与基站的距离； E_{max} 为节点 i 的邻居节点中最大的剩余能量； E_{ir} 为节点的当前剩余能量； n_i 为节点 i 的邻居节点数； N 为网络中的节点总数； R_0 为初始簇首竞争半径。

1.3 数据传输

在簇间传输阶段，提出一种按比例分配的路由传输方式使簇间数据传输的能量消耗更加均衡。如果簇首与基站的距离没有超过门限值，则采用直传的方式直接将数据传送到基站。如果超过了门限值，则采用提出的按比例分配的路由方式传输数据。具体过程如下：首先，每个簇首 CH_i ($i = 1, 2, \dots, K$) 以 $2R_{ic}$ 广播半径广播一条 $CH_RELAY_MESSAGE$ 消息，其目的是能够覆盖到其他簇首。这条消息主要包括簇首 ID、当前剩余能量及该簇首与基站的距离。接收到消息后，如果邻居簇首到基站的距离小于簇首 CH_i 与基站的距离，则簇首 CH_i 计算和这些邻居簇首 CH_j ($j = 1, 2, \dots, L, L$ 为 CH_i 邻居簇首个数) 的距离，并建立一个邻居簇首信息表。

簇首 CH_i 运用式 (17) 计算代价函数 (DRC)，将数据包按比例传输到其邻居簇首 CH_j 上。

$$DRC(i, j) = 1 - \frac{E_{t(i, j)}}{E_{jr}}. \tag{17}$$

式中： $E_{t(i, j)}$ 为从 CH_i 到簇首 CH_j 传输数据的能耗； E_{jr} 为邻居簇首 CH_j 当前剩余能量。

假设 CH_i 的数据量为 M_i ，则需要按比例分配

公式，如 (18) 所示，向邻居簇首传输数据量。也就是说向邻居簇首 1 传输的数据量为 $M_i \times DRC(I, 1)$ ，向邻居簇首 2 传输的数据量为 $M_i \times DRC(I, 2)$ ，依次类推。

$$DRC(I, 1) : DRC(I, 2) : DRC(I, 3) : \dots : DRC(I, L).$$
 (18)

2 仿真实验

仿真实验的运行环境为 Intel Pentium 双核 (2.2 GHz) 处理器，2 GB 内存，采用 Matlab R2013 (b) 实验平台。与 EEUC，IDUC，FAHP 在两种网络场景内进行仿真实验与对比，仿真环境参数如表 2 所示。

实验的两个场景描述如下：

- I. 100 个节点均匀分布在 200 m × 200 m 的区域内。
- II. 100 个节点非均匀分布在 200 m × 200 m 的区域内。

表 2 仿真参数
Table 2 Simulation parameters

参数	值
基站坐标	(250, 100)
节点初始能量/J	1 ~ 3
数据包大小/b	4 000
通信半径/m	25
$E_{elec}/(nJ \cdot b^{-1})$	50
$\epsilon_{fs}/(pJ \cdot b^{-1} \cdot m^{-2})$	10
$E_{SEN}/(J \cdot b^{-1})$	0.5
$E_{COM}/(nJ \cdot b^{-1})$	5
d_0/m	80

在两个实验场景内分别比较四个算法的网络寿命。设置 UCDGAMCD 中的关键参数： $R_0 = 140$ m， $\alpha = 0.3$ ， $\beta = 0.3$ ， $\gamma = 0.4$ 。而 IDUC 也是关于异构传感器网络不等分簇，它的关键参数： $\alpha = 0.3$ ， $\beta = 0.3$ ， $\gamma = 0.4$ ， $\omega = 0.5$ ， $R_0 = 140$ m。图 1 为四种算法的网络寿命在场景 I 中的比较。从图 1 中可以观察出 UCDGAMCD 和 IDUC 算法的网络寿命比 EEUC 和 FAHP 的要长。这主要是因为 UCDGAMCD 和 IDUC 两种算法在簇拓扑构建的过程中，考虑了邻居节点的数量，从而使分簇更加合理。而 UCDGAMCD 的网络寿命比 IDUC 长些，这主要是因为 UCDGAMCD 不仅采用多准则决策综合评价作出簇首选择，而且簇间数据传输还采用按比例传输的路由方式，使能量消耗更加

均匀. 虽然, FAHP 在簇拓扑构建的过程中, 没有采用不等簇的分簇方式, 但是在场景 I 中, 其网络寿命与不等簇的 IDUC 十分接近. 这主要是因为 FAHP 在竞选簇首的过程中也采用了综合评价节点来竞选簇首.

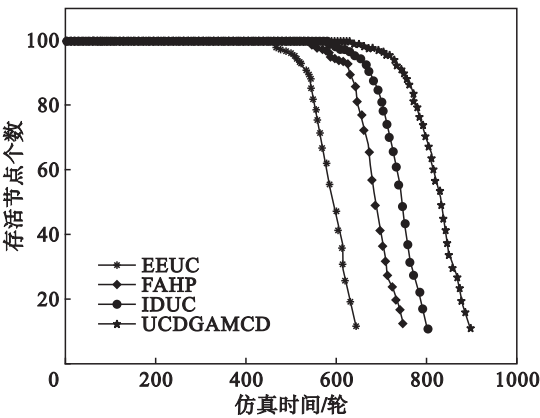


图 1 四种算法的网络寿命在场景 I 中的比较
Fig. 1 Comparison of network lifetime of four algorithms in scenario I

图 2 为四种算法的网络寿命在场景 II 中的比较. 在图 2 可以观察到四种算法在出现第一个死亡节点之后, 随着仿真时间的持续增长, 其存活节点个数都是呈线性下降的, 并没有造成网络分裂. 另外, 从图 2 中还可以观察到 UCDGAMCD 的网络寿命明显长于其他三种算法. 这主要是因为: 在簇首竞争的过程中, 采用层次模糊积分和模糊层次分析的多准则决策来综合评价节点作为簇首; 而且 UCDGAMCD 算法还考虑了节点非均匀分布对簇首竞争半径的影响, 从而构建了更加合理的不等分簇拓扑; 同时, 簇间数据传输采用按比例分配的路由方式, 进一步均衡了网络的能量消耗. 因此, UCDGAMCD 的网络寿命明显比其他三种算法长.

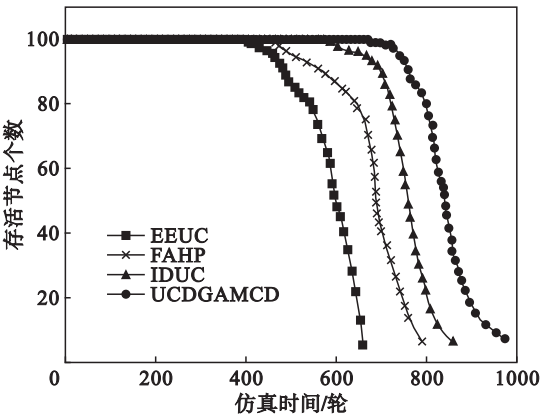


图 2 四种算法的网络寿命在场景 II 中的比较
Fig. 2 Comparison of network lifetime of four algorithms in scenario II

3 结 论

本文提出了一种基于多准则决策方法的不等簇数据收集算法 UCDGAMCD. 首先, 采用直觉模糊层次分析法和层次模糊积分的多准则决策方法来竞选簇首. 其次, 根据簇首与基站的距离、剩余能量和邻居节点个数, 提出了一个新的簇首竞争半径, 使其能够构建更加合理的不等簇拓扑. 然后, 为了均衡簇首间能耗, 提出了按比例分配传输数据的路由方式. 最后, 仿真结果表明 UCDGAMCD 在两种实验场景中都获得了较长的网络寿命.

虽然, UCDGAMCD 大大降低了网络能耗, 但是不能适应数据包重发与丢失的情况. 下一步工作将针对该问题展开讨论.

参考文献:

[1] Afsar M M, Tayarani N M H. Clustering in sensor networks: a literature survey [J]. *Journal of Network and Computer Applications*, 2014, 46: 198 – 226.

[2] Zhang D, Wang X, Song X, et al. A new clustering routing method based on PECE for WSN[J]. *EURASIP Journal on Wireless Communications and Networking*, 2015, 2015 (1): 1 – 13.

[3] Chen C, Gu X, Yu J, et al. IDUC: an improved distributed unequal clustering protocol for wireless sensor networks[J]. *International Journal of Distributed Sensor Networks*, 2015 (4): 682 – 693.

[4] Hematkah H, Kavian Y S. DCPVP: distributed clustering protocol using voting and priority for wireless sensor networks [J]. *Sensors*, 2015, 15 (3): 5763 – 5782.

[5] Gao T, Jin R C, Song J Y, et al. Energy-efficient cluster head selection scheme based on multiple criteria decision making for wireless sensor networks [J]. *Wireless Personal Communications*, 2012, 63 (4): 871 – 894.

[6] Atanassov K T. Intuitionistic fuzzy sets [J]. *Fuzzy Sets and Systems*, 1986, 20 (1): 87 – 96.

[7] Heinzelman W B, Chandrakasan A P, Balakrishnan H. An application-specific protocol architecture for wireless microsensor networks [J]. *IEEE Transactions on Wireless Communications*, 2002, 1 (4): 660 – 670.

[8] Gu X, Yu J, Yu D, et al. ECDC: an energy and coverage-aware distributed clustering protocol for wireless sensor networks [J]. *Computers & Electrical Engineering*, 2014, 40 (2): 384 – 398.

[9] Zhang C, Li W, Wang L. AHP under the intuitionistic fuzzy environment. [C] // *Proceedings of 8th International Conference on Fuzzy Systems and Knowledge Discovery*. Shanghai, 2011: 583 – 587.