韓國經營科學會誌
第 14 券 第 1 號
1989年 6月

-研究論文-

다차원 개체를 위한 차이등급 clustering
(The Difference Order Clustering for Multi-dimensional Entities)

이 철*
강 석 호*

ABSTRACT

The clustering problem for multi-dimensional entities is investigated. A heuristic method, which is named as Difference Order Clustering (DOC), is developed for the grouping of multi-dimensional entities. DOC method has an advantage of identifying the bottle-neck entities. Comparisions among the proposed DOC method, modified rank order clustering (MODROC) method, and lexicographical rank order clustering using minimum spanning tree (lexico-MMSTROC) are illustrated by a part type selection problem.

1. 서론

clustering은 다량의 개체(entity)들을 비교적 적은 수의 동질적인 그룹으로 형성시키는 것으로서 많은 응용분야에서 이용되고 있다. 그러나 'cluster'의 수학적 정의의 부재, 문제정의의 불명확성, 해에 대한 평가기준의 미비 등 clustering 문제에는 불확실한 요소들이 존재하여 문제해결의 난점이 되고있다. 문제의 성격을 명확히 하기 위하여 clustering 문제들을 분류성격, 개체차원(the dimension of entities), 병목개체 등의 관점을 통해 구체화하면 다음과 같다.

Jambu와 Lebeaux[3]는 clustering 문제를 classification 문제와 classing 문제로 양분하였다. "classification" 문제란 주어진 개체들을 묶어 class들의 hierarchy나 partition을 형성시키는 문제이며 "classing"은 기존의 이미 정의되어 있는 분류시스템에 특정 개체를 편입시키는 문제이다. 본 연구의 범위는 classification 문제로 한정한다. Jambu와 Lebeaux의 classification 문제의 대상은 균질적인 개체들로 통계학적 clustering 기법의 대부분이 이 범주의 문제를 위해 개발된 것이다.

group technology의 도입과 더불어 발생하는 문제 중의 하나인 MPGF(Machine-Part Group Formation) 문제를 고려해 보자. 이 문제는 classification 문제 유형에 속하지만 개체들의 특성이 Jambu와 Lebeaux의 개체특성과 상이하다. MPGF의 개체들은 기계와 부품으로 나뉘어 있으며 기계-부품간의 관계만 존재한다. 또한 clustering의 복

* 서울대학교 산업공학과

적은 기계그룹과 부품그룹을 형성시키며 동시에 부품그룹과 기계그룹간에 1 : 1 대응관계를 가지도록 하는데 있다. 이런 clustering 문제는 2차원 개체 clustering 문제로 생각할 수 있다. 2차원 개체 clustering 문제의 해법으로 기존의 통계학적 clustering 기법은 적절치 못하다. MPGF의 해법으로서 개발된 대표적인 clustering 기법인 rank order clustering(ROC)[4], modified rank order clustering(MODROC)[1], machine-component cell formation(MACE)[8] 등은 모두 2차원 개체 clustering 문제에 대한 해를 제공한다.

분산처리 시스템이 요구되는 사업체의 전산화 계획 수립시 발생하는 업무－자료 분석(process-data class analysis) 문제를 고려해 보자. 분산처리를 고려하지 않는다면 이 문제는 2차원 개체 clustering 문제이다. 그러나 지역이 고려대상에 포함되어야 하므로 업무, 자료, 지역을 함께 clustering 해야 한다. 그리고 각 개체들간의 관계는 서로 이질적인 개체들 사이에서만 존재한다. 이러한 문제는 3차원 개체 clustering 문제로 분류될 수 있다. 이러한 개체차원의 관점에서 볼 때 classification 문제는 n차원 개체 clustering 문제로 유형화됨을 알 수 있다.

상기한 논의들과는 다른 관점에서 Vannelli와 Kumar[7]는 classification 문제를 tearing 문제로 간주하였다. 이 문제는 해로서 cluster 뿐만 아니라 병목개체(bottle-neck entity)의 탐색을 요구한다. 병목개체의 판별은 해의 개선을 위한 선택이 가능한 경우 가치있는 정보를 제공한다. 본 연구의 목적은 n차원 개체 clustering 문제의 cluster 해 및 병목개체 해를 제공하는 해법의 개발에 있다.

〈표-1〉 clustering 문제에의 접근방법

문 제	입력정보	접근방법	출력정보
classification	1차원 개체들과 개체관계들	hierarchical clustering	개체들의 hierarchy 또는 주어진 기준하의 cluster 해
classification	2차원 개체들과 개체관계들	mathematical programming, ROC	clusters
tearing	1 또는 2차원 개체들과 개체관계들	bottle-neck finding	cluster와 병목 개체
다차원 개체 clustering	다차원 개체들과 개체관계	DOC	cluster와 병목 개체

2. 해법의 개발

2.1. 부호 및 용어

본 연구에서 편의상 사용된 부호 및 정의된 용어는 다음과 같다.

a) n : 개체들의 차원

b) e_i : i 차원의 개체

c) $|e_i|$: i 차원의 개체들의 cardinality

d) $e_i[k]$: i 차원에서의 k번째 개체

e) $r(e_1, e_2, \cdots, e_n)$: $e_1, e_2, \cdots, e_n$ 개체들 사이의 관계

f) 관계공간 R : r의 집합

g) 부분관계공간 $S(e_k=j)$: k 차원에 j 개체만을 가지는 개체관계들의 집합

h) $r_{S(e_k=j)}(e_1, e_2, \cdots, e_{k-1}, e_{k+1}, \cdots, e_n)$: 부분관계공간 $S(e_k=j)$의 k 차원을 배제한 개체

관계

i) 기준개체관계 : 특정 cluster의 성격을 대표하는 개체관계

j) 기준부분관계공간 $S(e_k=j)$: j가 기준개체관계의 개체인 부분관계공간

k) 기준관계선 : n−1개 차원의 개체가 기준개체관계와 동일한 개체관계의 집합

l) 블록 : 예비 cluster(a candidate for cluster)

m) 예외 개체관계 : 어느 블록에도 속하지 않는 '1'의 값의 개체관계

n) 통합기준 : 복수의 블록이 주어졌을 때 블록과 블록을 통합할 것인가 분리할 것인가의 처리기준

2.2. 가정

a) 각 개체의 차원은 알려져 있다.

b) 개체관계는 1 또는 0의 값을 갖는다.

c) 동일 차원의 개체들간에는 관계가 존재하지 않는다.

d) 모든 개체들은 다른 차원의 개체들을 통하여 관계를 맺고 있다(connectivity).

2.3. 제약

cluster 해는 각 차원에 대해 적어도 하나 이상의 개체를 포함하여야 한다.

2.4. 목적(objective)

cluster내의 개체들의 동질성과 cluster 간의 이질성을 증가시키도록 개체들을 그룹으로 분할하고 분할의 장애가 되는 병복개체들을 판별한다.

2.5. 해법의 절차

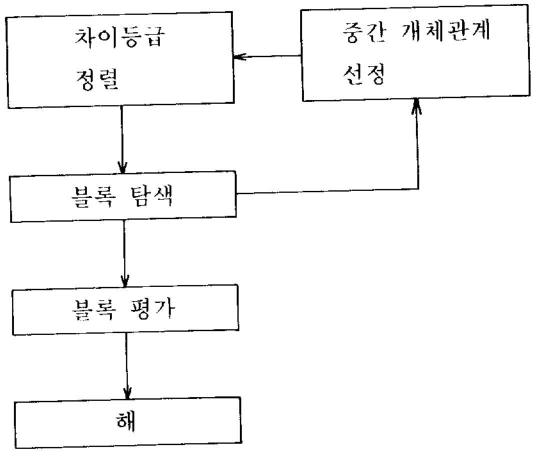
본 연구에서 제시한 DOC(Difference Order Clustering) 기법은 다음의 4가지 절차로 구성되어 있다.

1) 차이등급정렬(Difference Order Sorting) : 기준개체와 밀접한 순으로 각 차원의 개체들을 배열한다.

2) 블록 탐색(Block Search) : 기준개체관계와 가까운 블록과 먼 블록을 분리한다.

3) 중간 개체관계 선정(Median Cell Selection) : 탐색된 블록으로부터 기준 개체관계에 대한 정보를 이용하여 적정 블록을 획득한다.

4) 블록 평가(Block Evaluation) : 부속블록과 병복개체관계 등을 판별한다.



(그림-1) DOC의 절차

2.5.1. 차이등급정렬절차

하나의 적절한 기준개체관계가 선정되었다고 하자. 기준개체관계가 선정되면 이에 따라 n개의 기준부분관계공간이 결정된다. 차이등급정렬 절차의 개념은 부분관계공간들을 '차이'의 관점에서 기준개체관계에 가까운 순으로 배열하고자 하는 것이다. 두 부분관계공간 사이의 '차이'는 다음과 같이 정의된다.

정의 : 차이부분공간

두 부분관계공간 $S(e_k=i)$, $S(e_k=j)$가 존재할 때 이 두 부분관계공간과 동일한 개체들을 가지는 부분관계공간 d를 구성할 수 있다. 이때 다음의

관계가 성립한다.

$$\{e_l \mid e_l \in S(e_k=i)\} = \{e_l \mid e_l \in S(e_k=j)\}$$
$$= \{e_l \mid e_l \in d\}$$
where $l=1, 2, \cdots, k-1, k+1, \cdots, n$

이때 세 부분관계공간의 개체관계들 사이에 다음의 관계가 성립하면 d를 '차이부분공간'이라 부르고 $D(S(e_k=i), S(e_k=j))$로 표시한다.

$$r_d(e_1,e_2,\cdots,e_n)=1, \text{if} \quad r_{S(e_k=i)}(e_1,e_2,\cdots,e_n)$$
$$= r_{S(e_k=j)}(e_1,e_2,\cdots,e_n)$$
$$=0, \text{ otherwise}$$

예를 들어 (그림-2) (b)의 첫번째 행은 $S(e_1=2)$이며 벡터 (1,1,1,0,0)으로 표현될 수 있다. 이와 마찬가지로 4번째 행도 (1,0,1,1,0)의 벡터로 표현되어지는 $S(e_1=1)$이다. 이때 $D(S(e_1=2), S(e_1=1))$는 (0,1,0,1,0)이 된다. DOC의 경우 $D((S(e_k=i), S(e_k=j))$에서 $S(e_k=i)$가 항상 기준부분관계공간이므로 $D(e_k=j)$로 축약하여 표현할 수 있다.

	1	2	3	4	5
1	1		1		1
2	1	1			1
3			1	1	1
4	1	1			
5	1	1			
6			1	1	
7			1	1	

(a)

	2	1	5	3	4
2	1	1	1		
4	1	1			
5	1	1			
1	1		1	1	
3			1	1	1
6				1	1
7				1	1

(b)

(그림-2) 2차원 개체 clustering

정의 : 등급벡터

차이부분공간 $D(e_k=j)$의 등급벡터 $V(e_k=j)=(v1, v2, \cdots, vm)'$는 다음과 같이 정의된다.

$$v1 = \sum_{[ki]\in\Sigma[ki]=n-1} r_{S(e_k=j)}(e[k1], e[k2], \cdots, e[kn])$$

$$v2 = \sum_{[ki]\in\Sigma[ki]=n} r_{S(e_k=j)}(e[k1], e[k2], \cdots, e[kn])$$

$$\vdots$$

$$vm = \sum_{[ki]\in\Sigma[ki]=\Sigma|ei|} r_{S(e_k=j)}(e[k1], e[k2], \cdots, e[kn]) \quad \cdots (1)$$

차이등급정렬은 어느 차원에서도 (2)식이 만족되도록 개체들의 순서를 바꾸는 것이다.

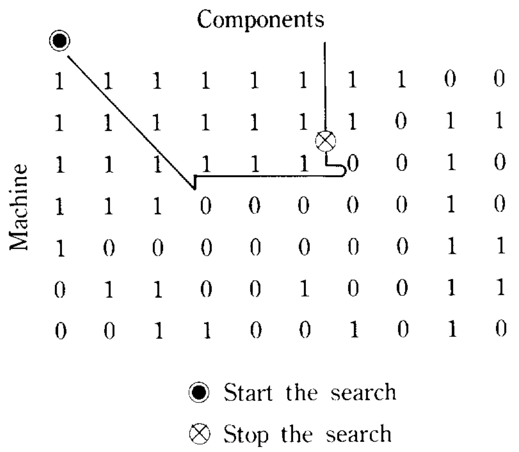
$$V(e_i[1]) \preceq V(e_i[2]) \preceq \cdots\cdots\cdots \preceq V(e_i[\,|e_i|\,])$$
for all i $\cdots(2)$

이때 $\preceq$는 'lexicographically not greater than'을 의미한다. 차이등급정렬의 결과로서 기준개체관계와 관련이 많은 개체일 수록 순서가 앞설 것이 기대된다.

2.5.2. 블록탐색절차

차이등급 정렬이 완료되면 기준개체관계와 밀접한 개체들과 관계가 없는 개체들은 2개의 블록을 형성하게 된다. 블록의 탐색에 있어 Chandrasekharan과 Rajagopalan[1]은 MPGF의 해법으로 개발한 MODROC에서 블록 탐색방법을 제시하였다(그림-3). 이 방법은 2차원 개체의 관계행렬(incidence matrix)이 정렬(rank order) 되었을 때의 다음 성질을 이용한 것이다.

"어느 행이나 열의 최초에 k개의 연속된 '1'이 존재한다면 그 이전의 행(또는 열)에는 적어도 k개의 '1'이 존재한다."



(그림-3) **MODROC**의 블록 탐색

상기한 가정 d와 함께 이 성질은 n차원 개체들의 관계에도 확장 응용될 수 있다. 즉, $r(e_1[1], e_2[1], \cdots, e_n[1])$에서 시작하여 최초로 '1'의 값이 아닌 $r(e_1[k], e_2[k], \cdots, e_n[k])$를 탐색하여 기준 관계와 연관성이 높은 블록, Bo 를 분리해낼 수 있다. Chandrasekharan과 Rajagopalan의 방법은 하나의 특정 차원의 개체들에 편향되어 있으므로 본 연구에서는 Bo를 근간으로 통합기준을 각 차원에서 반복적으로 적용함으로써 적절한 블록을 구한다. 통합기준으로는 예외 개체관계의 최소화를 사용하였다.

블록 탐색은 다음과 같이 이루어진다.

step 1.

Bo에 $r(e_1[1], e_2[1], \cdots, e_n[1])$를 부가한다.

Eo에 $e_1[1], e_2[1], \cdots, e_n[1]$를 부가한다.

step 2.

$r(e_1[k], e_2[k], \cdots, e_n[k]) = 1$인 동안

Eo에 $e_1[k], e_2[k], \cdots, e_n[k]$를 부가한다.

$k = k + 1$

Bo에 $e_i \in$ Eo인 개체들로 구성된 개체관계들을 부가한다.

step 3.

Er에 $e_i \notin$ Eo인 개체관계들을 부가한다.

step 4.

예외 개체관계를 감소시키는 개체 e_i 가 존재하는 동안

Eo에 e_i 부가

Bo에 $e_i \in$ Eo인 개체관계를 부가

Br에 $e_i \notin$ Eo인 개체관계를 부가

2.5.3. 중간 개체관계 선정절차

적절치 못한 기준 개체관계의 선정은 다음의 두 가지 결과를 초래할 수 있다.

(1) 병목 기준 부분관계공간의 경우
(bottle-neck index subspace case)

한 블록을 가정하자. 이 블록을 어떻게 분리(partition)하여도 그 결과인 분리된 하부블록들에 통합기준을 적용할 경우 다시 원래의 블록으로 통합된다면 이 블록은 "최소 블록(minimal block)"이라 할 수 있다. 최소블록의 부분집합은 다시 최소블록을 구성할 수 있으므로 어느 최소블록이 존재할 때 자신을 포함하는 최소블록이 자기자신인 경우를 최대 최소 블록(maximum minimal block)이라 부르자.

복수의 최대 최소블록이 존재할 때 어느 특정 부분관계공간이 이들 최대 최소 블록들과 관계를 가지고 있어 통합기준적용시 이들 모두가 하나의 블록으로 통합된다면 이 부분관계공간을 '병목 부분관계공간'이라 부를 수 있다. 왜냐하면 기준 부분관계공간으로서 병목 부분관계공간이 선정되는

것을 피한다면 각각의 최대 최소 블록들은 독립되어 보다 개선된 해를 제공할 수 있기 때문이다. 이때의 부분관계공간의 선정된 개체는 '병목개체'가 된다. (그림-4)에 예를 보였다. 첫번째 행이 병목 부분관계공간으로 작용하고 있다. 따라서 잘못된 기준 개체관계의 선정은 병목 기준 부분관계공간을 형성하게 되어 해의 품질에 영향을 끼침을 알 수 있다.

	1	2	3	4	5
1	1		1		1
2			1	1	1
3	1	1			1
4			1	1	
5			1	1	
6	1	1			
7	1	1			

(a) 문제

	5	1	3	2	4
1	1	1	1		
3	1	1		1	
2	1		1		1
6		1		1	
7		1		1	
4			1		1
5			1		1

(b) 병목 기준 부분관계공간의 결과

(그림-4) 병목 기준 개체 예제

(2) 부속 기준 부분관계공간의 경우 (annexed index subspace case)

만일 어느 한 개체가 특정 cluster에 소속되어 있으나 그 개체의 성격이 동일 cluster내의 다른 개체들을 대표하지 못하고 오히려 단편적 성격만을 가지는 경우가 있을 수 있다. 이 경우 기준 개체관계가 이 개체를 포함하게 되면 차이등급정렬은 부적절한 결과를 초래할 수 있다. 예를 들어 부속기준개체 예제의 (a)는 어느 주어진 2차원 개체들에 대하여 차이등급정렬 및 블록 탐색을 수행한 결과이다. (그림-5) 이 해는 만족스럽다고 할 수 없다. 이는 '1'행과 같이 대표성이 없는 부속 부분관계공간이 기준 관계공간으로 선정되었기 때문이다.

	1	2	3	4	5	6	7	8	9	10
1	1	1								
2	1		1	1						
3	1		1	1						
4	1		1	1	1					
5		1		1	1				1	1
6						1	1	1		
7					1	1	1	1		
8					1	1	1	1		

(a)

	1	2	3	4	5	6	7	8	9	10
4	1	1	1	1						
2	1	1	1							
3	1	1	1							
1	1									
5			1	1	1				1	1
7				1		1	1	1		
8				1		1	1	1		
6						1	1	1		

(b)

(그림-5) 부속 기준 개체 예제

상기한 두 가지의 부적절한 기준 개체관계의 선정이 일어났을 경우 결과로서 나타나는 블록들은 다음과 같은 성질을 지니게 된다.

병목 기준 부분관계공간의 경우 :

상대적으로 블록의 밀도는 낮고 크기는 크다.
부속 기준 부분관계공간의 경우 :
상대적으로 블록의 밀도는 높고 크기는 작다

이러한 성질을 이용하면 병목 기준 부분관계공간의 경우를 단순하고 효과적으로 회피할 수 있다. 이 방법이 중간 개체관계 선정으로서 선행된 블록으로부터 상기한 성질을 파악하여 이용하는 것이다. 구체적인 방법은 다음과 같다.

step 1.
각 차원에서 $\lceil e_i \rceil /2$ 번째가 되는 개체관계를 찾는다.

step 2.
이 개체관계의 값이 '0'이면 값이 '1'인 가장 가까운 개체관계를 선정한다.

step 3.
차이등급정렬을 수행한다.

step 4.
블록 탐색으로 새로운 블록을 구한 후 기존 블록과 다음의 규칙들을 사용하여 비교 및 선정을 수행한다.

규칙 1 : 블록의 밀도가 높은 블록을 선택한다.

규칙 2 : 블록의 밀도가 같은 경우에는 개체의 수가 많은 블록을 선택한다.

규칙 3 : 규칙 1과 2가 모두 일치하면 예외 개체관계 수가 적게 되는 블록을 선택한다.

규칙 4 : 규칙 1, 2, 3이 모두 일치하면 기존의 블록을 선택한다.

중간 개체관계 선정 절차를 (그림-4)의 (a)와 (그림-5)의 (a)에 적용한 결과를 (그림-6)에 보였다. 중간 개체관계 선정은 성공적으로 병목 개체를 피할 것으로 기대되는데 이는 블록의 크기가 상대적으로 크고 병목 개체의 블록에서의 비중이 상대적으로 적기 때문이다. 이러한 병목 개체의 판별은 minimal cut node 문제로서 NP-complete로 알려져 있다[7].

	1	5	2	3	4
3	1	1	1		
1	1	1		1	
6	1		1		
7	1		1		
2		1		1	1
4				1	1
5				1	1

(a)

	2	1	9	10	5	4	6	7	8	3
1	1	1								
5	1		1	1	1	1				
2		1				1				1
3		1				1				1
4		1			1	1				1
6							1	1	1	
7					1		1	1	1	
8					1		1	1	1	

(b)

(그림-6) 중간 개체관계 선정예

부속 개체에 대하여 상기한 방법은 개선된 해를 가져다 줄 기회를 제공하기는 하나 높은 신뢰도를 기대하기 어렵다. 부속 개체의 경우에는 블록이 상대적으로 적은 개체만을 포함하게 될 것이므로 블록내의 다른 개체관계를 선정하는 방식으로 문제를 해결하기 보다는 적정한 다른 블록이 형성되었을 때 해당 블록에 통합하는 접근 방법이 보다 적절하다. 따라서 우리가 획득한 블록이 부속 개체에서 비롯된 것인가의 판정과 적정 블록에의 통합에 대한 절차가 구성되어야 한다. 이 절차는 블록 평가에서 다루게 된다.

2.5.4. 블록 평가

블록의 획득과 평가에는 여러가지 정책의 채택이 가능하다. 채택되는 정책은 기준 개체관계의 선정 및 블록의 구성, 블록의 평가에 대한 절차를 제공하여야 한다. 가장 품질이 우수한 블록을 얻기 위하여서는 모든 개체관계를 기준으로 삼아 블록을 구할 수도 있으며 일부 개체관계만을 대상으로 삼을 수도 있다. 보다 효과적인 정책의 수립을 위하여 다음과 같이 블록 평가의 중간 단계의 상황을 상정하여 평가에 대한 논의를 전개시켜 보자.

블록 평가의 어느 단계에서 이미 탐색된 블록의 집합 $C=\{C_1, C_2, \cdots, C_k\}$가 존재하고 이들 블록들은 서로 배타적(mutually exclusive)이라 가정하자. 어느 기준 개체관계를 통하여 블록 B가 탐색되었다면 B와 C 간에는 다음과 같은 관계와 그에 상응하는 해석이 가능하다.

Case 1 : $B \cap C = \phi$ for all $C_i \in C$.

B가 C와는 유리된 별도의 블록임을 알 수 있다.

Case 2 : $B \subseteq C$ for one C_i .

집합 $C_i \setminus B$는 이전 평가에서는 B와 결합되지 않았으나 현재의 평가에서는 B와 결합되었다. 이는 $C_i \setminus B$가 병목 개체관계들이거나 B가 부속 개체관계들임을 의미한다.

Case 3 : $B \supseteq C$ for some C_i 's

집합 $B \setminus C_i$ 는 이전 평가에서는 C_i 와 결합되지 않았으나 현재의 평가에서는 C_i 와 결합되었다. 이는 $B \setminus C_i$ 가 병목 개체관계들이거나 C_i 가 부속 개체관계들임을 의미한다.

Case 4 : $B \cap C \neq \phi$ for some C_i 's and neither $B \subseteq C_i$ nor $B \supseteq C_i$

$B \setminus \cup C_i$ 가 병목 개체관계들이거나 각각의 $C_i \setminus B$가 병목 개체관계들임을 의미한다.

Case 5 : $B \cap C \neq \phi$ for some C_i 's and $B \supseteq C_i$ for some C_j

case 3과 case 4가 결합된 경우이다.

상기한 해석들에 따라 타당성있는 블록들과 병목 개체관계들을 분리함으로써 서로 겹치는 cluster(overlapping natural clusters)에 대한 해를 구할 수 있다. 모든 개체관계들을 기준으로 삼아 블록을 구성해내어 평가하는 것은 계산상의 비효율성 뿐만 아니라 개개의 블록들을 지나치게 세분하여 의미없는 해를 제공할 위험이 있다. 중간 개체관계 선정을 이용하는 경우 병목 부분관계공간들을 성공적으로 피하고 또 판별할 수 있으므로 기준 개체관계의 대상에서 기존의 블록들내의 관계들을 배제할 수 있다. 또한 이미 탐색된 블록들에 대하여는 병목 개체관계로 판정하거나 타당한 블록으로 판정되어 있으므로 새로운 블록의 도입시 고려해야 할 범위로 축소된다. 즉, case 2와 case 5의 경우는 일어나기 어렵다. 따라서 새로운 블록과 기존 블록과의 관계 및 그에 따른 평가는 다음과 같다.

Case 1 : B를 새로운 블록으로서 C에 추가한다.

Case 3 : C_i 가 복수이거나 C_i 의 개체 수가 B보다 많으면 $B \setminus \cup C_i$ 를 C_i 에 추가하고 $B \setminus \cup C_i$ 를 병목 개체관계들로 판정한다.

Case 5 : 각각의 C_i 를 $C_i \setminus B$로 바꾸고 $B \setminus \cup C_i$ 를 C에 추가한다. 그리고 $B \setminus \cup C_i$ 를 병목 개체관계들로 판정한다.

C가 모든 개체들을 포함하게 되면 부속 블록들을 통합기준에 의거하여 적당한 블록에 결합시킬 수 있다.

2.6. Algorithm과 복잡도

2.6.1. Algorithm

상기한 논의는 동시에 여러 블록들의 탐색을 수행해야 할 필요가 없음을 시사하고 있다. 따라서 기준 개체관계의 선정 방법을 구체화함으로써 algorithm이 확정된다. 본 연구에서는 초기 기준 개체관계의 선정은 첫번째 개체관계를 사용하며 이후의 기준 개체관계 선정은 직전의 기준 개체단계에서 가장 먼 C에 속하지 않는 '1'의 값의 개체관계를 선정한다. 이에 따른 DOC algorithm은 다음과 같다.

step 1.
초기 기준 개체관계 선정 및 $C=\phi$, $Br=R$,
$Er=\{e_i\}$ where $e_i \in R$.

step 2.
Br에 대하여 차이 등급정렬을 수행한다.

step 3.
블록 탐색을 수행한다.

step 4.
중간 개체관계 선정을 수행한다.

step 5.
블록 평가를 수행한다. $C \leftarrow C \cup B$

step 6.
$Er \leftarrow Er \setminus \{e_i \mid e_i \in C\}$
$Br \leftarrow r(e_1, e_2, \cdots, e_n)$ where $e_i \in Er$
If $Er=\phi$, goto step 7
Else, $r(e_1, [\,|e_1|\,], e_2[\,|e_2|\,], \cdots, e_n[\,|e_n|\,])$에서 가장 가까우며 C와 중복되지 않고 1의 값을 갖는 개체관계를 기준관계로 선정한 후 step 2로 간다.

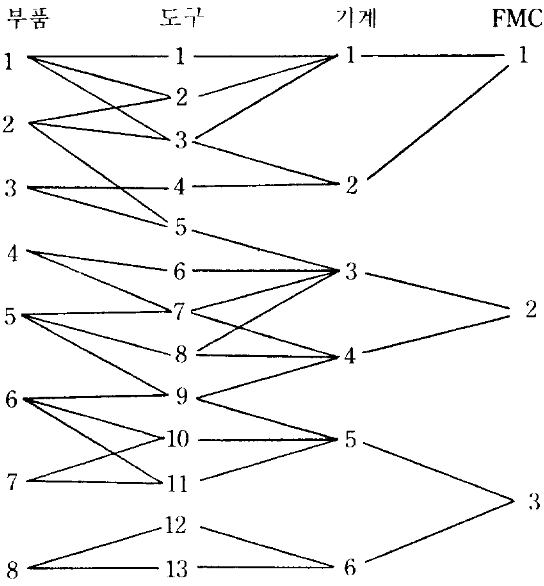
step 7.
C에 대하여 통합기준을 적용한다.

2.6.2. 복잡도

algorithm DOC는 차이정렬에 대하여 polynomial이며 차이정렬은 식 (2)에 대하여 polynomial이다. (1) 식의 등급 벡터를 확장하여 R 전체에 대한 등급 벡터 $V=(v_1, v_2, \cdots, v_{m+1})^t$를 상정하여 보자. 한 차원에서 정렬을 할 때마다 V는 이전 단계보다 최소 lexicographically $(0,0,\cdots,0,1)^t$ 만큼 감소하며 최소치에 다달았을 때 정렬은 완료된다. 따라서 DOC의 복잡도는 $O(\Sigma|e_i|\cdot|R|^2\cdot \text{sorting})$이며 sorting에는 polynomial algorithm이 존재하므로 DOC는 polynomial algorithm이다.

2.7. 예제

FMC 방식으로 운영 중인 생산공장에서의 부품 선택 문제(Part Type Selection Problem)를 고려하여 보자. 각 FMC에는 복수의 기계가 있으며 각 기계에는 설치 가능한 도구(tool)의 집합이 존재한다. 부품 주문이 도래한 경우 주어진 도구와 기계로 모두를 동시에 가공할 수 없을 때 부품 선택을 하게되며 이때의 목적은 도구의 준비작업(set-up)을 최소화하는 데 두게 된다.



(그림-7) 예제

주어진 FMC (C1, C2, C3), 기계 (M1, M2, …, M6), 도구 (T1, T2, …,T13) 등은 tool magazine의 용량제한 (5)으로 주어진 부품 주문 (P1, P2, …, P8)을 동시에 처리할 수 없다.

부품 선택 문제는 복잡도가 NP-hard이며 해법으로 여러가지 기법이 제안되고 있다[2]. 본 예제의 경우 그 목적이 DOC의 평가에 있으므로 clustering 기법인 MODROC와 lexico-MMSTROC[9]를 비교 대상으로 선정하였다. MODROC는 2차원 개체를 대상으로 하므로 부품 주문과 도구들을 clustering 대상으로 하였고 lexcio-MMSTROC는 FMC를 제외한 부품 주문, 도구, 기계들을 alternative route로 처리하였다.

〈표-2〉

	MODROC	lexico-MMSTROC	DOC
해	P1 P2 P3 T1 T2 T3 T4 P4 P5 P6 P7 T6 T7 T8 T9 T10 T11 P8 T12 T13	P1 P2 T1 T2 T3 T5 M1 P3 P4 T4 T5 T6 T7 M2 M3 P5 T7 T8 T9 M4 P6 P7 T9 T10 T11 M5 P8 T12 T13 M6	P1 P2 P3 T1 T2 T3 T4 M1 M2 C1 B{T5} P4 P5 T6 T7 T8 M3 M4 C2 B{T9} P6 P7 T10 T11 M5 B{C3} P8 T12 T13 M6
set-up	3	4	3
over lapping tools	—	T5 T7 T9	T5 T9
기계도구 mismatch	T5 — M3	T5 — M1	—
기계그룹간 부품이동	P2 P3	P2	P3 P5
기계FMC mismatch	있음	있음	없음

3가지 기법의 적용 결과를 〈표-2〉에서 보였다. 이 경우 DOC가 기계그룹간의 부품이동 측면을 제외하면 모두 MODROC나 lexico-MMSTROC보다 우수함을 알 수 있다. 또한 기계그룹간의 부품이동의 경우에도 DOC의 해는 T5, T9의 도입으로 이동을 제거할 수 있는데 비해 lexico-MMSTROC의 해는 그렇지 못하다.

3. 결론

다차원 개체에 대한 clustering 기법으로서 DOC가 개발되었다. DOC는 cluster해 뿐만 아니라 병목 개체들을 찾아주므로 문제 해결자에게 귀중한 정보를 제공한다. 부품 선택 문제의 한 예에서 DOC는 다른 두 가지 방법에 비하여 우수한 해를 제공하였다. 그러나 이 결과를 일반적인 다차원 개체의 clustering 문제에까지 일반화하기는 어려우며 따라서 다차원 개체 clustering 문제에 있어서의 DOC의 일반적 수행도에 대한 평가가 필요하다. 대부분의 기존 연구에 있어 clustering 기법의 평가에 상호 비교가 사용되어온 데 비해 DOC의 경우 비교 대상이 없으므로 추후 연구과제로서 절대적 척도에 의한 평가가 요망된다.

REFERENCES

1. M. P. Chandrasekharan an R. Rajagopalan, "MODROC : and extension of rank order clustering for group technology", IJPR, Vol. 24, No. 5, pp. 1221-1233, 1986.
2. S. S. Hwang, "An FMS production planning system and evaluation of part selection approaches", Working Paper No. MS-43, Berkeley Business School, 1987, September.
3. M. Jambu and M-O. Lebeaux, "Cluster analysis and data analysis", North-Holland, 1983.
4. J. R. King, "Machine-component grouping in production flow analysis : an approach using a rank order clustering algorithm", IJPR, Vol. 18, No. 2, pp. 213-232, 1980.
5. J. R. King and V. Nakornchai, "Machine-component group formation in group technology : review and extension", IJPR, Vol. 20, No. 2, pp. 117-133, 1982.
6. W. T. McCormick, Jr., P. J. Schweitzer and T. W. White, "Problem decomposition and data reorganization by a clustering technique", OR, Vol. 20, pp. 993-1009, 1972.
7. A. Vannelli and K. R. Kumar, "A method for finding minimal bottle-neck cells for grouping part-machine families", IJPR, Vol. 24, No. 2, pp. 387-400, 1986.
8. P. H. Waghodekar and S. Sahu, "Machine-component cell formation in group technology : MACE", IJPR, Vol. 22, No. 6, pp. 937-948, 1984.
9. 이 철, "Alternative route와 복수동일기계를 고려한 MPGF의 발견적 해법", 한국경영과학회/대한산업공학회 '89 춘계공동학술대회논문집, pp. 151-156, 1989.