

# 항만 경쟁력 평가를 위한 유사도 기반의 이산형 평균 알고리즘

추봉성\* · 이철영\*\*

\*한국해양대학교 대학원, \*\*한국해양대학교 물류시스템공학과 교수

## A Dispersion Mean Algorithm based on Similarity Measure for Evaluation of Port Competitiveness

Bong-Sung Chu\* · Cheol-Yeong Lee\*\*

\*Graduate school of National Korea Maritime University, Busan 606-791, Korea

\*\*Dept. of Logistics Engineering, National Korea Maritime University, Busan 606-791, Korea

**요약** : 평균법과 클러스터링은 다속성 평가문제에서 널리 쓰이고 있는 중요한 데이터 마이닝 기법들이다. 그러나, 다양한 다속성 평가 문제에서 데이터 마이닝을 할 때, 데이터들의 특징은 그 중요성이 달라질 수 있기 때문에 이러한 데이터의 중요도 차이를 고려해야 할 필요가 있다. 따라서, 이러한 기법들은 데이터의 선택 및 중요도 등과 같이 그 특징을 얼마나 잘 반영하는 지가 중요하다. 게다가, 산술평균법의 경우에는 우선순위 및 가중치로 정의되는 평가구조에서 적합한 결과를 산출하기에는 한계가 있을뿐 만 아니라, 평가자 그룹별 특징을 반영하기 곤란하다. 따라서, 본 연구에서는 기하학적 도형을 바탕으로 유사도를 평가하여, 평가자 그룹별로 특징지어지는 이산적인 환경에서의 평균을 산출하는 알고리즘을 제안하였다. 본 알고리즘의 핵심사항 중 하나는, 항목별 우선순위의 혼돈없이 유사도를 평가할 수 있다는 점이다.

**핵심용어** : 유사도, 정규화 배분법, 위치비교 매트릭스, 수량화 이론, 이산형 평균

**Abstract** : The mean and Clustering are important methods of data mining, which is now widely applied to various multi-attributes problem. However, feature weighting and feature selection are important in those methods because features may differ in importance and such differences need to be considered in data mining with various multif-attributes problem. In addition, in the event of arithmetic mean, which is inadequate to figure out the most fitted result for structure of evaluation with attributes that there are weighted and ranked. Moreover, it is hard to catch hold of a specific character for assume the form of user's group.

In this paper, we propose a dispersion mean algorithm for evaluation of similarity measure based on the geometrical figure. In addition, it is applied to mean classified by user's group. One of the key issues to be considered in evaluation of the similarity measure is how to achieve objectiveness that it is not change over an item ranking in evaluation process.

**Key Words** : Similarity Measure, Regularity Distribution, Location Comparison Matrix, Quantification Theory, Dispersion Mean

## 1. 서 론

다속성(Multi-attributes) 평가항목을 가지는 일반적인 평가구조에서 데이터 전체의 특징이나 경향을 수치로 나타내는 기술적 척도, 즉, 각 항목에 대한 물리량  $x$ 를 대표할 수 있는 최적 결과값을 얻기 위해서 목적에 따라 다양한 중심화 경향의 척도가 이용되고 있다.

일반적인 중심화 경향의 척도로 주어지는 대표적인 방법으로는, 크게 데이터 집합에서 가장 큰 빈도수를 가지는 분포의 최고점(Peak)  $x$ 를 구하는 최빈치(Mode)와 중위수(Median), 그리고, 계량적 데이터 집합에 대한 중심화 척도로서 가장 잘 알려져 있는 산술평균법(Arithmetic mean)이 있다. 하지만, 가장 널리 이용되고 있는 산술평균법의 경우, 각 평가항목에 대해서 평가자 혹은 평가자그룹 형태의 데이

터가 특정한 외적기준이 없이 상호 독립적인 회답패턴으로 주어지는 경우나 중요도 혹은 우선순위의 형태로 주어질 때는, 다음과 같은 문제점들이 존재한다.

첫째, 평가자 그룹별 대표속성 및 특징, 즉, 기존 그룹별 평가에서 주어진 우선순위의 반영이 곤란해 진다.

둘째, 평가항목  $A_i$ 에 대해서 상반된 회답패턴이 동등한 수치로 주어질 경우,  $\sum x_i / n = 0$  이 된다.

따라서, 본 연구에서는 평가자 그룹으로 주어지는 다속성 평가문제에서 산술평균이 갖는 문제점을 해결하고, 평가자 그룹별 우선순위와 특징을 반영할 수 있는 이산형 평균 알고리즘을 제안하고자 한다.

우선, 평가자의 부담을 줄이기 위해서 각 평가항목에 대해 주어진 중요도를 수직선상에 위치로 표현하여 우선순위를 매기고, 상대적인 중요도 관계를 고려해 가는 방법인 상

\* 대표저자 : 추봉성(정회원), super-bs@hanmail.net 051)410-4911

\*\* 중신회원 yjici@mail.hhu.ac.kr 051)410-4911

대위치 평가법을 이용하여 가중치를 산출한다. 그리고, n개의 평가항목(A<sub>i</sub>, i=1,2,...,n)을, n개의 축으로 이루어진 평면상에 배치하여 각각의 축상에 평가자별 가중치를 표시한 후, 축상의 가중치를 직선으로 연결했을 때 얻어지는 도형의 각으로 유사도를 평가하여 유사한 성질을 가진 개체나 거의 유사한 반응을 나타내는 개체별로 그룹핑한다. 그리고, 그룹별 특징을 보다 정확하게 반영하기 위해서 적절한 못한 데이터들, 즉, 그룹핑 되지 못한 개체들을 제외한 후, 그룹별로 평균을 산출하고 통합하여 평가한다.

## 2. 평가구조 및 가중치 동정

### 2.1 다속성 평가구조

Fig. 1과 같이 n개의 평가항목 A<sub>i</sub>를 가지는 다속성 단일 계층구조 K<sub>a</sub>에 대해 각 항목(A<sub>i</sub>, i=1,2,...,n)는 평가자 N<sub>n</sub>에 따라 가변적으로 주어지는 중요도를 W<sub>an</sub>, 평가자 그룹 G<sub>n</sub>으로 나타내고, 각 항목은 상호독립적이라 가정한다.

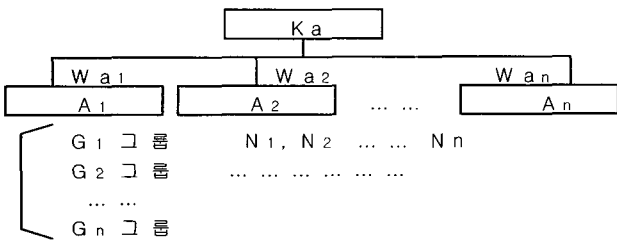


Fig. 1 Structure of evaluation with multi-attributes

또한, 각 항목에 대해서, 특정한 외적기준이 주어지지 않고, 다양한 분야별로 n개의 평가자 그룹(G<sub>j</sub>, j=1,2,...,n)이 참여하는 평가구조로서 평가자 혹은 평가자 그룹은 중요도 및 우선순위의 형태로 평가하게 된다.

### 2.2 상대위치 평가법에 의한 가중치 계산

AHP와 같이 평가항목간 가중치를 일대비교(Pair comparison)할 경우, 평가항목수를 n이라 하면, n(n-1)/2회의 일대비교가 필요하게 되어, 평가자의 부담이 증가하게 된다. 따라서 본 연구에서는, 평가항목의 가중치를 수직선상의 「위치」로 평가하고, 그 위치 데이터를 바탕으로 각 평가항목의 가중치를 산출함으로써, 평가자의 부담을 경감시킬 수 있는 AHP에 의한 「상대위치 평가법」을 이용한다.

#### 1) 정규화 배분법

특정 평가항목 x(순위는 y위)에 관해서, 원점부터의 위치 데이터 d<sub>x<sup>y</sup></sub>를 측정 한 후, Fig. 2와 같이 각 위치 데이터를 위치데이터의 총합으로 나눈다[7].

$$w = d_x^y / \sum_{x=1}^n d_x^y \quad (1)$$

#### 2) 위치비교 매트릭스법

Fig. 2와 같은 위치데이터로부터, 일대비교 매트릭스에 대응하는 “위치비교 매트릭스”를 구축한다. 여기서, 순위a의 평가항목 A<sub>i<sup>a</sup></sub>와 순위b의 평가항목 A<sub>j<sup>b</sup></sub>와의 위치비교평가치, D<sub>ij<sup>ab</sup></sub>,는 위치 데이터의 차로 정의한다. 즉,

$$D_{ij}^{ab} = (d_i^a - d_j^b) \quad (2)$$

( i, j = 1, 2, ..., x, ..., n )  
 ( a, b = 1위, 2위, ..., y위 ..., m위)

이 되고, Fig. 2와 같은 형식으로 표현할 수 있다.

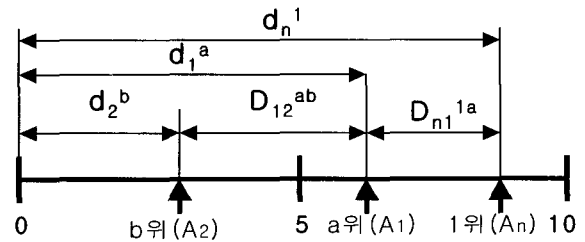


Fig. 2 Measurement of location data

위치비교 평가치 P<sub>ij</sub>는 a와 b의 순위관계에 의해 아래 식(3)과 같이 정의한다. 단, 동순위의 평가항목이 존재할 경우, D<sub>ij<sup>ab</sup></sub>=0 가 되기 때문에 위치비교 매트릭스 구축시 동순위의 위치비교 평가치는 0이 되고 만다. 따라서, 모든 위치비교 평가치는 1을 더한 수치로 해서, 동순위의 경우는 1이 되게 한다.

- a>b (D<sub>ij<sup>ab</sup></sub> >0)일때, P<sub>ij</sub> = D<sub>ij<sup>ab</sup></sub> + 1
- a<b (D<sub>ij<sup>ab</sup></sub> <0)일때, P<sub>ij</sub> = 1 / -D<sub>ij<sup>ab</sup></sub> + 1
- a=b (D<sub>ij<sup>ab</sup></sub> =0)일때, P<sub>ij</sub> = 1

위치비교 평가치를 바탕으로 위치비교 매트릭스를 구축하면 아래 식(4)와 같다.

$$P = [ P_{ij} ] = \begin{matrix} & A_1 & A_2 & \dots & A_m \\ \begin{matrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{matrix} & \begin{bmatrix} 1 & P_{21} & \dots & P_{m1} \\ 1/P_{21} & 1 & \dots & P_{m2} \\ \dots & \dots & \dots & \dots \\ 1/P_{m1} & 1/P_{m2} & \dots & 1 \end{bmatrix} \end{matrix} \quad (4)$$

따라서, 식(4)의 최대고유치에 대한 고유벡터가 각 평가항목 A<sub>i</sub>의 가중치(Weight)가 된다[7].

## 3. 유사도 평가 및 군집화

### 3.1 거리 척도에 의한 유사도 평가

클러스터링을 위한 유사도 평가는 주어진 데이터 집합을

특징이 유사한 클러스터들(Homogeneous clusters)로 나누는 것을 의미한다. 또한, 데이터 군집을 정의하기 위해서는 특징 벡터를 특징한 군집 중심의 영역으로 배정하기 위한 유사도(Similarity measure)를 정의해야 한다.

일반적으로 유사도 평가시, 널리 이용되고 있는 방법으로는 거리척도를 들 수 있는데, 두 특징 벡터 x와 y간의 유클리드 평방거리(Euclidian distance)는 아래 식(5)와 같이 정의된다.

$$\Delta_{ij}^2 = \| x - y \|^2 = (x_i - y_j)^2 \quad (5)$$

그러나, 거리 척도로 유사도를 평가하여 클러스터링 할 경우에는, 각 평가항목에 대해, 평가자 혹은 평가자 그룹별로 주어진 우선순위가 혼돈될 우려가 있다. 다시말해서, 평가항목  $A_i$ 에 대한 우선순위를 순서대로  $r_1, r_2, \dots, r_m$ 이라고 하고, N개의 대상항목에 대해 i번째 항목과 j번째 항목의 유사도를,

$$E = (e_{ij}) = e_{ij}(x_i - y_j)^2 \quad (6)$$

라 할 때, Table 1의 a, b의 경우, 평가자별로 주어진 우선순위가 동일함에도 불구하고, 우선순위와는 상관없이 유클리드 거리가  $D_{ab}=31$ 이 되어, 비유사성이 높게 나타나는 문제점이 있다.

Table 1 Example of weighted data

	$A_1$	$A_2$	$A_3$	$A_4$
a	8(r2)	9(r1)	5(r4)	7(r3)
b	5(r2)	6(r1)	3(r4)	4(r3)
c	7(r1)	6(r2)	5(r3)	4(r4)
d	8(r1)	8(r1)	7(r3)	6(r4)

### 3.2 기하학적 접근에 의한 유사도 평가

특정한 외적기준이 없는 다속성 평가구조에서 평가자그룹으로 분류가능한 평가자간의 유사도 평가시 우선순위의 혼돈을 방지하고, 동일한 우선순위를 가지는 항목 또는 유사한 개체들끼리 클러스터링을 하기 위해 본 연구에서는 기하학적인 접근을 통해 유사도를 평가하였다.

평가자에 의해 주어진 평가 데이터를 n개의 평가항목을 축으로 하는 평면상에 배치하고, 상대위치 평가법에 의해 구해진 가중치를 축상에 표현한다. 그리고, 각 평가항목에 대한 가중치를 직선으로 연결함으로써 생성되는 삼각형(이하, 도형이라 함)의 내각으로 유사도를 평가하여 우선순위의 혼돈없이 우선순위가 동일한 항목과, 또는 이에 유사한 반응을 나타내는 개체들을 평가한다.

다시말해서, Table 1과 같이 4개의 평가항목을 가지는 평가문제에서 각 평가항목  $A_i$ 에 대한 가중치를  $w_{ai}$ 라 하고, 상대위치 평가법으로 산출한 평가자 a와 b의 가중치를 결

과를 4개의 축상에 표현하여 가중치값을 직선으로 연결하면 다음과 같이 나타낼 수 있다.

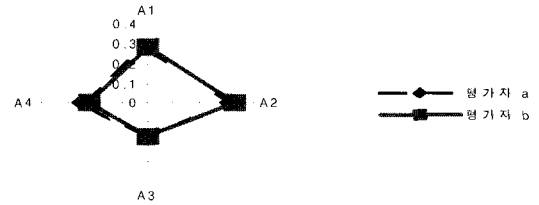


Fig. 3 Similarity measure for a geometrical figure

각 축에 대한 도형의 내각은 다음과 같이 정의할 수 있다. 즉, n개의 평가항목  $A_i$ 에 대해, 평가자별로 주어진 가중치  $W_i$ 가 좌우로 인접한 평가항목과 이루는 도형의 내각을  $A_{ij}\theta$ 이라 하고,  $A_i$ 축 좌우 두 내각의 합을  $A_i\theta$ 라 한다.

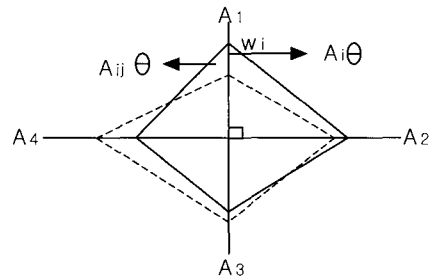


Fig. 4 Concept of an interior angle

그리고, n개의 평가항목은 상호독립적이므로 원점을 중심으로 n개의 축이 이루는 각은  $2\pi/n$ 이 되고, 내각  $A_{ij}\theta$ 는 아래 식(7)과 같이 삼각함수 제2 코사인 법칙을 통해 산출할 수 있다.

$$\begin{aligned} \cos A &= -a^2 + b^2 + c^2 / 2bc \\ \cos B &= -b^2 + c^2 + a^2 / 2ca \\ \cos C &= -c^2 + a^2 + b^2 / 2ab \end{aligned} \quad (7)$$

기하학적인 접근을 통한 유사도 평가는 다음과 같이 정의한다. 대상 i와 j간의 유사도를 나타내는 유사도 행렬을  $S = (s_{ij})$ 라 하고, 특정 평가 항목  $A_i (i=1, 2, \dots, n)$ 에 대해 대상 i에  $c_i$ , 대상 j에  $c_j$ 라는 수치를 아래의 식(8)과 같이 도형의 각으로 할당한다.

$$\begin{aligned} |S_{ij}| &= \sum_{i,j=1}^n (c_i - c_j) \\ (\text{단, } c_i &= A_i\theta) \end{aligned} \quad (8)$$

즉, Table 2와 같이,  $n \times n$  정방행렬의 행과 열을 기준으로

평가항목을 할당하여 i행과 j열을 요소로 하는 유사도  $s_{ij}$ 의 수치를 기입한다.

Table 2 Matrix of similarity measure

평가항목	1	2	...	j	...	R
1		$s_{12}$	...	$s_{1j}$	...	$s_{1R}$
2						
...						
i	$s_{i1}$	$s_{i2}$	...			
...						
R	$s_{R1}$	$s_{R2}$	.....			

이 때, 유사도는 수량화 이론 제4류의 성질을 이용하여, 아래 식 (9)와 같이 계수 Q로 정의할 수 있다[5].

$$Q = - \sum_{i=1}^n \sum_{j=1}^n S_{ij}(c_i - c_j) \quad (\text{단, } S_{ij} = S_{ji}) \quad (9)$$

그리고,  $c_i$ 는 계수 Q가 최대가 되도록 결정하면 되지만, 계수 Q의 수치가 무한히 증가하지 않도록 하기위해, 아래 식(10)과 같이 c의 분산은 일정하다고 가정한다.

$$1/R \sum_{i=1}^n c_i^2 - (mc)^2 = a^2 \quad (10)$$

(단, a는 정수이고,  $mc = \sum_{i=1}^n c_i / R$  임.)

따라서, 식(9)에서 계수 Q를 최대로 하는 해  $c = [c_1, c_2, \dots, c_n]$ 는, 고유치가 최대가 되는  $H=[h_{ij}]$ 의 제2고유벡터가 된다(제1고유벡터의 고유치는 0임). 다시 말해서, 평가항목  $A_i$ 에 대해, 우선순위가 가장 높은 항목의 각  $A_i$ 가 가장 작게 되고, 우선순위가 가장 낮은 항목의 각이 가장 크게 나타난다.

또한, 도형간 각의 차가 작을 수록, 개체간의 유사성이 높아지고, 각의 차가 클수록 유사성은 낮아지게 되므로, 우선순위의 혼돈없이 우선순위가 같거나 유사한 반응을 나타내는 개체끼리 군집을 형성하게 된다.

### 3.3 배치 알고리즘에 의한 클러스터링

대상간의 계량적 관계를 공간상에 표현하는 방법으로는 크게 다차원척도법과 수량화 이론을 들 수가 있고, 구체적인 클러스터링 기법들로는 크게, 인접 데이터가 유사한 것(거리가 가까운 것)끼리 묶고, 그 평균값을 사용하는 형태로 Hierarchical cluster 정보를 제공하는 UPGMA와 k개 만큼의 초기 포지션을 설정한 후, 각 포지션에서 가까운 거리의 노드들을 묶고, 이들로부터 다시 중점계산을 반복하여 더

이상의 멤버들이 바뀌지 않을때까지 수행하는 K-means법, 그리고, 다차원의 데이터를 2차원으로 사상시켜주는 Neural network의 한 종류인 SOM이 있다.

본 연구에서는 외적기준이 주어지지 않을 때, 거리척도가 아닌 기하학적인 접근을 통해서, 대상간의 유사도를 요소로 하는 행렬을 2차원 평면상에 배치하는 방법인 수량화 이론의 성질을 이용하여, 유사한 반응을 나타내는 평가자그룹을 2차원 평면상에 클러스터링 하였다.

즉, 식(9)에서 고유치가 최대가 되는  $H=[h_{ij}]$ 의 고유벡터의 제2 고유벡터 및 제3 고유벡터를 각각 x, y축의 값으로 하여 2차원 평면위에 배치하여 표현하였다.

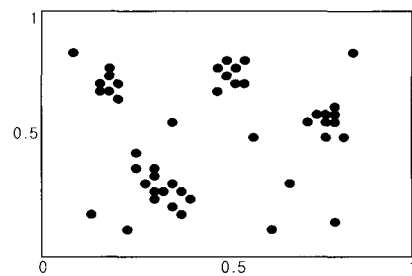


Fig. 5 Example data of two-dimensional space

## 4. 이산형 평균 알고리즘

### 4.1 이산형 평균 알고리즘

본 연구에서 제안하는 이산형 평균법이란 평가자그룹별로 구성되어지는 평가문제에서 우선순위의 혼돈없이 개개의 그룹별 특성을 파악해서 반영하고자, 서로 다른 포지션별로 군집화된 개체들을 각각의 평가항목에 대해서 평균을 산출한 후, 이를 평준화해서 통합평균치를 산출하는 방식을 말한다. 다시 말하면, 각 평가항목별로 군집을 형성하지 못하는 개체들, 즉, 평가자 그룹의 의견과는 동떨어진 부적합한 개체들은 제외하고, 우선순위가 같거나 혹은 유사한 개체들로 형성된 군집의 평균을 산출하여 통합한다는 것이다.

군집별 평가항목  $A_i$ 에 대해서 중요도를  $w_{ai}$ , 개개의 군집에 소속된 n개의 개체들의 평균치를  $C_{mi}$ , 통합평균치를  $T_{mi}$ 이라 하고, 군집별로 소속된 n개의 개체들의 개수의 영향을 받지않게 하기위해 아래의 식(11)과 같이 평준화를 실시한다.

$$C_{mi} = \sum_{i=1}^n w_{ai} / n$$

$$T_{mi} = \sum_{i=1}^n C_{mi} \quad (\text{단, } i=1,2,\dots,n) \quad (11)$$

지금까지 설명한 내용을 중심으로 다속성 및 평가자그룹으로 정의 가능한 평가문제에서의 평균치 산출 알고리즘을 정리하면 아래와 같다.

- Step 1. 평가자로부터 주어진 중요도 및 우선순위를 상대 위치 평가법으로 가중치를 산출한다.
- Step 2. n개의 평가항목, 즉, n개의 축상에 가중치를 표현하여 생성되는 도형의 각도에 의해 유사도를 평가한다.
- Step 3. 최대고유치에 대한 고유벡터의 값을 x, y좌표로 해서 클러스터링한다.
- Step 4. 그룹별로 각 항목 A<sub>i</sub>에 대한 평균치를 산출하여 평균화한 후, 통합한다.

4.2 적용 및 비교평가

1) 유사도 평가법 비교

아래 그림 Fig. 6과 같이, 세 종류의 평가 항목(a<sub>i</sub>)에 대해, 10점 척도로 주어지는 A, B, C 세명의 평가결과를 바탕으로, 거리에 의한 유사도 평가법과 본 연구에서 제안하는 유사도 평가법을 비교해 보았다.

Table 3 Example of weighted data

	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>
A	10(1)	8(2)	7(3)
B	9(1)	5(2)	3(3)
C	7(3)	8(2)	9(1)

1-1) 거리척도에 의한 유사도

유클리드 평방거리  $\Delta_{ij}^2 = \|x - y\|^2 = (x_i - y_i)^2$  를 통해 유사도를 평가하고 2차원 공간상에 각 개체들을 배치하면 다음과 같다.

Table 4 Eigen value and eigen vector

(H.Matrix)	1축	2축	3축
고유치	0.00000	-0.16906	-0.37094
	0.57735	-0.29229	0.76239
	0.57735	0.80639	-0.12807
	0.57735	-0.51410	-0.63432

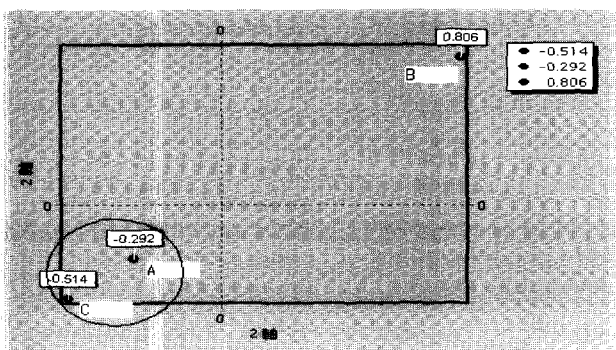


Fig. 6 A cluster by distance

그 결과, 우선순위가 정반대인 A개체와 C개체가 서로 군집을 형성하게 되어 각 항목에 대한 군집별 평균 산출시 적당치 않음을 알 수 있다.

1-2) 도형의 각에 의한 유사도

우선, 아래와 같이 각 항목에 대한 평가자별 가중치를 정규화 배분법으로 산출하였다.

Table 5 Example of weighted data

	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>
A	0.4	0.32	0.28
B	0.53	0.29	0.18
C	0.29	0.33	0.38

그리고, 산출된 가중치를 세항목(a<sub>i</sub>, i=1,2,3), 즉, 3개의 축상에 표시하여 직선으로 연결했을때 얻어지는 도형의 각으로 유사도를 평가하였다.

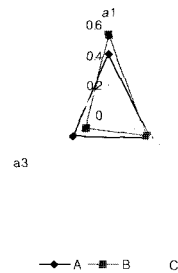


Fig. 7 Similarity measure for a geometrical figure

Table 6 Angle to represent from each attribute point of view

	Cos(a <sub>1</sub> )	Cos(a <sub>2</sub> )	Cos(a <sub>3</sub> )
A	93	125	142
B	72	125	163
C	140	120	100

각 축에 대해 평가자별로 형성되는 도형의 각으로 유사도를 평가하여 2차원 공간상에 배치한 결과는 아래 Fig. 6과 같다.

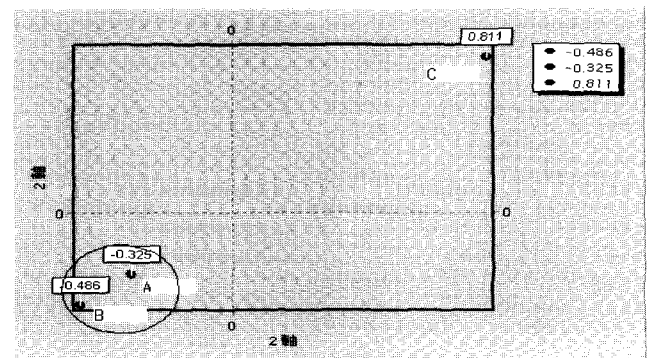


Fig. 6 A cluster by angle

즉, 위 그림 Fig. 6에 나타난 것과 같이, 거리척도가 아닌 각에 의한 유사도를 평가했을때, 우선순위가 동일한 A개체와 B개체가 군집을 형성하였다. 따라서, 본 연구에서 제안한 기하학적 도형에 의한 유사도 평가법에서는 우선순위가 동일한 개체 혹은 유사한 반응을 나타내는 개체들을 클러스터링 해 줌으로써, 각 항목별 군집의 평균 산출시 각 평가자 그룹의 성향을 충분히 반영할 수 있게 되는 셈이다.

2) 실제 적용을 통한 평균 산출

실측치 산출을 통한 비교검증을 행하기 위해 Table 7과 같이 항만 경쟁력에 대해 5가지 평가항목을 가지는 일반적인 다속성 평가구조의 간단한 예제를 들어 보기로 한다. 평가과정상에서 평가자 그룹별 성향이 혼돈되고 그룹별 우선순위 파악이 곤란한 유클리드거리 척도와는 달리, 본 알고리즘의 적용에 있어서는 각 평가자 그룹별로 중요시하는 항목의 우선순위를 파악하고, 우선순위별로 군집화되는 경향을 살펴보기 위해 임의의 두 평가자 그룹을 설정하고 적용해 보았다. 우선, 항만관련 운송업자와 화주 20명으로 설정한 두 평가자 그룹에서 각 평가자별로 평가항목에 대한 중요도 및 우선순위를 n점 척도를 통해 평가한 후 각 축상에 표현되는 평가자별 도형의 각을 산출하여 유사도를 평가하였다.

Table 7 Example of weighted data

	입지 (a <sub>1</sub> )	시설 (a <sub>2</sub> )	물동량 (a <sub>3</sub> )	비용 (a <sub>4</sub> )	서비스 (a <sub>5</sub> )
1	0.154	0.205	0.154	0.256	0.231
2	0.179	0.179	0.179	0.256	0.205
...	...	...	...	...	...
20	0.179	0.205	0.154	0.231	0.231

Table 7에서 평가자들에 의해 산출된 가중치를 축상에 표현하면 아래 Fig. 7과 같다.

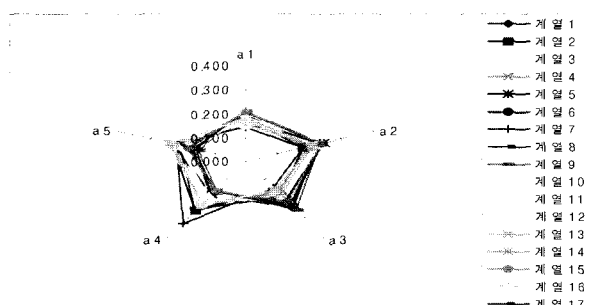


Fig. 7 Similarity measure for a geometrical figure

각 축에 대한 평가자별 도형의 각을 Table 8과 같이 산출하고, 유사도를 평가하여 2차원 평면상에 나타내면 아래 Fig. 8과 같다.

Table 8 Angle to represent from each attribute point of view

	Cos(a <sub>1</sub> )	Cos(a <sub>2</sub> )	Cos(a <sub>3</sub> )	Cos(a <sub>4</sub> )	Cos(a <sub>5</sub> )
1	135	86	138	85	96
2	113	108	122	85	112
...	...	...	...	...	...
20	123	92	134	93	98

Table 9 Eigen value and eigen vector

(H.Matrix)	X(1)	Y(2)
E.V	-3.5655	-4.1990
C(1)	-0.2043	-0.0065
C(2)	-0.113	0.0289
C(3)	-0.1588	-0.1070
.....	.....	.....
C(19)	0.2920	0.0528
C(20)	-0.1294	-0.0301

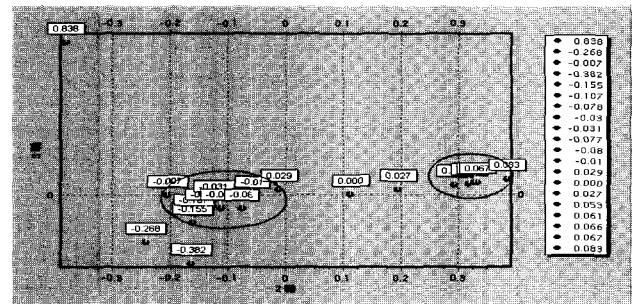


Fig. 8 A cluster by angle

적용 결과, 군집은 크게 운송업자 10명으로 이뤄진 A그룹과 화주4명 운송업자1명으로 이뤄진 B그룹, 그리고 군집이라 보기 어려운 나머지 5개체로 나누어졌다.

결과적으로, Table 10에서 나타난 결과와 같이, 기존의 산술평균(A.M)으로 계산했을때는 화주 및 운송업자 그룹별 우선순위 파악이 곤란할뿐더러, 전체적인 순위, 즉, 시설, 서비스, 그리고 순차적으로 비용, 물동량, 입지순이었다. 그러나 본 연구에서 제안한 알고리즘(D.M)을 적용했을 때는 화주그룹이 서비스 및 비용을 가장 중시하고, 운송업자는 물동량과 시설을 가장 중요시 여기는 것으로 나타났으며, 각 평가항목별 화주그룹과 운송업자 그룹의 통합 평균치는 순차적으로 시설, 물동량, 서비스, 비용, 입지순으로 나타났다.

Table 10 Arithmetic mean and dispersion mean

	a1	a2	a3	a4	a5
A.M	7.05	8.55	7.45	7.95	8.10
D.M	7.50	8.70	8.05	7.60	7.90

## 5. 결론 및 추후과제

다수의 평가자그룹으로 분류가능한 다속성 구조문제에서 각 항목별 물리량의 대표치 산출, 즉, 평균 산출에서는 각각의 평가자그룹별 우선순위 파악이 무엇보다 중요할뿐만 아니라, 전체적인 통합평균치면에서도 평가자그룹의 특징을 잘 반영할 수 있어야 한다.

그러한 측면에서 볼 때, 본 연구에서 제안한 알고리즘은 기하학적 도형의 각으로 유사도를 평가하여, 우선순위가 같은 개체 및 이와 유사한 반응을 나타내는 개체들을 클러스터링함으로써, 기존의 거리척도에 의한 유사도 측정법이 가진 문제점을 해결하였다. 또한, 개개의 평가자 그룹별 대표치 파악이 용이해졌을뿐만 아니라, 그룹핑되지 못한 개체들은 평균치 산출시 제외함으로써, 통합평균치를 산출할때, 평가자그룹별로 보다 정확한 데이터 및 그룹별 특징을 반영할 수 있게 되었다. 그러나, 본 연구에서 이용했던 수량화 이론은 상이도가 큰 개체끼리 가깝게 배치되는 경향이 있으며, 유사도 평가시 일대비교를 통해 값을 산출함으로써 데이터의 양이 방대해질수록 수행속도가 떨어진다는 단점이 있다. 또한, 개체간 거리치가 일정 범위를이상으로 커져버리면 우선순위가 같은 항목이라 할지라도 군집화되지 않는 문제점이 존재하며, 평가항목간 평가기준이 다를시에 이를 동일한 기준으로 평가할수 있도록 재조정이 필요하다. 따라서, 향후, 평가 범위의 제한이나 보정이 필요할 것으로 판단되며, 이와 더불어, 보다 다양한 분석 방법들과의 비교 검증을 통한 보완이 병행되어야 할 것으로 사료된다.

## 참 고 문 헌

- [1] 김장욱, 박춘일(1995), “선형대수학”, 한국해양대학교 출판부.
- [2] 안상형, 이명호(1998), “현대통계학”, 학현사.
- [3] 이철영, 이석태((1993), “상호연관성을 지닌 계층구조형 문제의 평가 알고리즘”, 한국항만학회지, 제7권, 제1호, pp.6~11.
- [4] 임봉택, 양원, 이철영(1998), “대규모 다계층 MADM 문제의 퍼지평가 알고리즘에 관한 연구 -퍼지측도의 동정을 중심으로-”, 한국항만학회지, 제12권, 제1호, pp.12.
- [5] 寺野壽郎(1985), “システム工学入門 -あいまい問題への挑戦-”, 共立出版株式會社, pp.192~194.
- [6] 井上光平, 堀山政二, 浦浜喜一(1998), “ファジークラスターリングと數量化4類によるデータの階層表示”, 電子情報通信學會論文誌, D-2 Vol. J82-D-2 No.12, pp.2425
- [7] 盛亞他子, 鈴木聰士(2000), “AHPにおける位置評価法に關する研究”, 電子情報通信學會論文誌.
- [8] 竹内啓, 柳井晴木(昭和48年), “多變量解析の基礎”, 東洋經濟.

---

원고접수일 : 2003년 12월 8일

원고채택일 : 2004년 2월 7일