

퍼지 클러스터링에 의한 항만의 분류

Classification of Foreign Trade Ports using Fuzzy Clustering

양 원 재*, 금 종 수**

Weon-Jae Yang*, J. S. Keum**

* 한국해양대학교 대학원 해사정보공학전공(E-mail : ywj07@hanmail.net)

** 목포해양대학교 해상운송시스템학부 국제물류시스템학전공

ABSTRACT

Grouping ports in certain region by their characteristics could be used as the principal informations to establish national policy for port development or investment and also to analyze the competitiveness between ports.

Currently Korean ports are divided into two groups such as the local port and the designated port containing foreign trade port and coastal port under the Korean port law. These divisions seem to be used for port administration as the matter of convenience but some qualitative grouping is needed for research of port problems.

In this paper, 28 major Korean ports were clustered by the similar characteristics using Fuzzy C-Means and found to be classified 8 qualitative groups.

I. 서론

최근 경제·사회환경이 급변함에 따라 항만에 부과된 역할도 크게 변하고 있어 효율성과 공평성의 관점에서 항만분류에 따른 역할 및 배치방안을 개선하여 항만분류를 재조정할 필요가 있다. 우리 나라의 경우 모든 지역항만이 국제거점항 수준의 항만개발을 요구하고 있으나 항만의 개발 또는 개발방향을 설정할 때 먼저 대상으로 하는 항만의 특성을 명확히 파악하여 그 지역의 특성을 최대한 살릴 수 있는 다양한 형태의 항만개발과 배치방안을 마련할 필요가 있다.

항만의 분류를 위한 기존의 연구는 주로 컨테이너 항만을 중심으로 항만의 특성과 조건들을 한 두가지 변수를 이용하여 유사한 항만군으로 분류하고 있다. 그러나 기존의 연구에서 사용한 하드 클러스터링 방법은 분명한 경계선을 가지고 항만을 몇 개의 항만군집으로 분류하지만 경계부근에 있는 항만의 경우 어느 하나의 항만군집에 완전하게 분류하는 것이 불가능할 수도 있다는 문제점이 있다[1-4].

따라서 본 연구에서는 퍼지 클러스터링 방법 중의 하나인 Fuzzy C-Means 알고리즘을 이용하여 내부적인

상황도 식별할 수 있도록 소속하는 정도를 [0,1]로 확장하고 경계부근의 항만들을 한 개 이상의 항만군집에 소속될 수 있게 허용하는 방법을 사용하여 우리 나라 항만법상의 무역항을 중심으로 각 항만의 선박입출항 현황, 연간 화물처리량 및 항만시설현황 등을 파악하여 항만의 조건이나 성격이 유사한 항만들을 8개의 군집으로 분류하고 그 특성을 분석하였다.

II. FCM법에 의한 클러스터링

항만의 분류하기 위한 클러스터링 방법은 계층적 방법, 그래프 이론적 방법, 목적함수 방법으로 나누어지고, 각각의 방법에는 크리쉬(crisp) 분류법과 퍼지 분류법이 있는데 본 연구에서는 퍼지 목적함수 방법 중 FCM(Fuzzy C-Means) 알고리즘을 이용하여 국내 무역항의 분류에 적용한다[5-7].

어떤 항만 k 가 항만군집 S_i 에 속하는지 속하지 않는 지만을 나타내는 하드 클러스터링법에 대해서 FCM법은 각 항만이 복수개의 항만군집에 서로 다른 정도로 속한다는 퍼지 이론의 특성을 포함시킨 클러스터링 방법이다[11-14].

n 개의 항만의 집합 $X = \{x_1, x_2, \dots, x_n\}$ 퍼지 부분집합 S_i 로의 분할은 소속함수 $u_{si} : X \rightarrow [0,1]$ 로 표시할 수 있다. 여기서 $u_{ik} = u_{si}(x_k)$ 는 폐구간 $[0,1]$ 의 값을 가지며 항만 k 의 항만군집 S_i 에 소속정도를 나타낸다. 항만 k 의 항만군집 S_i 에 소속정도 u_{ik} 를 이용하면 c 분할의 개념을 퍼지 c 분할로 확장할 수 있다.

n 개의 t 차원의 데이터벡터 $X_k = x_{k,p} (p=1, 2, \dots, t)$

($k=1, 2, \dots, n$) c 개의 항만군집으로 분류할 때, 각 군집의 중심벡터 $v_i (i=1, 2, \dots, c)$ 데이터 x_k 와의 비유사도(dissimilarity) d_{ik} 는 유클리드 거리(Euclidean Distance)로 식(1)을 이용하여 구한다.

$$d_{ik} = d(x_k, v_i) = \|x_k - v_i\| \quad (1)$$

$$= \left\{ \sum_{j=1}^t (x_{kj} - v_{ij})^2 \right\}^{(1/2)}$$

FCM 알고리즘은 주어진 입력력 공간상의 데이터(x_k)에 대하여 다음 식(2)와 같은 목적함수를 최적화 시키는 클러스터 중심 및 소속도 함수 값을 구하는 알고리즘이다.

$$\text{Minimize } J(U, v) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik})^m \|x_k - v_i\|^2 \quad (2)$$

이 때, 식(2)의 목적함수를 최적화 시키기 위한 클러스터 중심(v_i)과 소속도 함수 값(u_{ik})은 다음 식(3) 및 식(4)로부터 구할 수 있다.

$$v_i = \frac{1}{\sum_{k=1}^n (u_{ik})^m} \sum_{k=1}^n (u_{ik})^m x_k, \quad 1 \leq i \leq c \quad (3)$$

$$u_{ik} = 1 / \sum_{j=1}^c (d_{ik} / d_{jk})^{2/(m-1)} \quad (4)$$

한편, FCM법의 알고리즘은 기본적으로 통상의 C-Means법의 U (소속도 함수값)와 v (클러스터의 중심)를 갱신하기 위한 루틴을 추가한 것으로 FCM 알고리즘을 요약하여 정리하면, 다음 4단계의 과정으로 구성된다.

단계 1 : 클러스터의 수 $c(2 \leq c \leq n)$ 와 지수가중치 $m(1 < m < \infty)$ 값을 결정하고, 퍼지 c 분할행렬($U^{(0)}$)을 초기화한다.

단계 2 : 단계 1에서 구한 $U^{(l)}$ 과 식(3)을 이용하여 새로운 클러스터의 중심 $v_i^{(l)} (i=1, 2, \dots, c)$ 를 구한다.

단계 3 : $x_k \neq v_i^{(l)}$ 일 때 식(4)에 의해 $U^{(l)}$ 를 $U^{(l+1)}$

로 갱신하고 그 외에는 $U_{ik}^{(l+1)} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$ 된다.

단계 4 : 수렴판정치 ϵ 에 대해서

$\|U^{(l+1)} - U^{(l)}\|_G \leq \epsilon$ 를 만족하면 종료하고 그렇지 않으면 $l = l + 1$ 을 한 후 단계 2로 되돌아가서 반복 수행한다.

III. 항만의 분류 및 그 특성 분석

3.1 항만의 분류

각 항만의 조건과 특성이 비슷한 항만들을 분류하고

자 할 때 일반적으로 항만의 조건과 특성을 나타낼 수 있는 변수들로는 연간 화물처리량, 입출항 선박 척수, 입출항 선박의 평균 톤수, 항만하역능력, 항만시설 등을 들 수 있다.

따라서 우리나라의 28개 무역항을 대상으로 하여 항만의 특성과 조건이 유사한 항만들로 분류하고, 각 항만군의 특성을 파악하기 위하여 Table 1에서와 같이 구체적으로 정량화 할 수 있는 항만조건과 항만의 특성을 나타낼 수 있는 5개의 변수를 선정하였다.

X1 : 연간 선박 입출항 척수, X2 : 입출항 선박 평균톤수, X3 : 하역능력, X4 : 보관능력, X5 : 연간 화물처리량

Table 1 Characteristic data of each port

항만 \ 변수	X1	X2	X3	X4	X5
부산항	69,429	6,482	84,764,000	6,379,573	107,757,128
인천항	40,639	5,478	59,590,000	3,992,772	108,227,010
평택항	8,922	4,900	6,951,000	1,117,510	26,132,060
여수항	6,385	3,877	3,016,000	276,364	6,751,662
광양항	35,604	6,823	73,805,000	2,132,657	131,058,788
마산항	13,199	3,806	14,233,000	2,017,662	11,320,805
삼천포항	2,695	6,129	13,581,000	386,436	13,560,678
옥포항	2,153	2,513	454,000	0.000	942,516
거제항	146	35,441	76,000	32,760	1,743
진해항	1,453	1,771	1,869,000	431,000	527,766
통영항	3,340	215	172,000	20,580	145,933
고현항	2,186	1,555	901,000	10,476	976,142
울산항	45,454	6,578	24,776,000	1,423,773	151,116,765
동해항	4,370	7,889	23,035,000	960,690	17,071,613
삼척항	2,022	3,635	7,286,000	30,000	6,141,137
목포항	2,976	1,372	6,422,000	214,000	3,656,588
속초항	220	355	896,000	46,000	40,896
육계항	1,919	3,139	4,914,000	93,000	5,159,884
군산항	8,954	4,484	7,596,000	1,805,043	11,502,821
강항항	836	1,160	1,068,000	24,294	622,689
목포항	14,714	778	6,271,000	350,000	5,918,542
완도항	3,185	300	817,000	68,300	334,478
포항항	13,871	5,341	44,542,000	2,316,011	49,119,671
제주항	6,332	777	3,589,000	85,027	2,385,552
서귀포항	2,210	570	1,359,000	72,584	563,519
대산항	7,293	9,585	4,987,000	0.000	49,546,337
보령항	472	20,263	10,682,000	0.000	8,095,858
태안항	336	18,639	5,483,000	0.000	5,278,476

먼저, 항만의 조건과 특성이 유사한 항만들을 군집화하기 위하여 FCM 알고리즘을 적용할 때 고려해야 할 파라메타인 항만군집의 수(c), 지수 가중치(m), 수렴판정치(ϵ)를 설정하여야 한다[9-10].

퍼지 클러스터링에서 항만군집의 타당성을 측정하기 위하여 분할계수(pc), 분할 엔트로피(pe)를 사용하며, 분할계수 및 분할 엔트로피는 항만군집의 수가 증가함에 따라 단조 감소 및 증가한다. 따라서 최적 항만 군집의 수는 항만군집의 수 2~12개에 대하여 클러스터링을 수행하여 분할 엔트로피가 $c-1$ 에서 c 로 갈 때의 기울기 아래에 있는 c 와 분할계수가 c 로부터 $c+1$ 로 갈 때 많은 감소를 얻는 c 를 선택하여 8개로 결정하였으며, 지수 가중치는 2, 수렴판정치는 $1.0e^{-5}$ 를 사용하였다.

FCM 알고리즘으로 항만을 분류하는 경우 각 항만이 단지 하나의 항만군집에 할당되는 것이 아니라 각 항만 군집에 소속되는 정도를 나타내는 멤버십 값으로 표현된다. Table 2는 각 항만의 조건과 특성을 나타내는 5개의 변수를 이용하여 클러스터링 한 결과 28개의 항만이 8개의 각 항만군집에 속하는 소속의 정도를 나타내고 있다.

Table 2 Membership values of the ports to the individual clusters

항만 \ 군집	A	B	C	D	E	F	G	H
부산항	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
인천항	0.039	0.114	0.125	0.046	0.121	0.027	0.497	0.030
평택항	0.324	0.003	0.008	0.487	0.031	0.028	0.007	0.112
여수항	0.216	0.002	0.004	0.096	0.014	0.022	0.004	0.642
광양항	0.005	0.010	0.019	0.005	0.014	0.004	0.939	0.004
마산항	0.076	0.003	0.006	0.819	0.032	0.015	0.005	0.044
삼천포항	0.918	0.001	0.001	0.028	0.005	0.008	0.001	0.037
옥포항	0.030	0.000	0.001	0.015	0.003	0.005	0.001	0.944
거제항	0.120	0.021	0.039	0.097	0.068	0.529	0.035	0.092
진해항	0.056	0.001	0.002	0.037	0.006	0.009	0.002	0.888
통영항	0.028	0.001	0.001	0.019	0.004	0.006	0.001	0.940
고현항	0.010	0.000	0.000	0.005	0.001	0.002	0.000	0.981
울산항	0.000	0.000	0.999	0.000	0.000	0.000	0.000	0.000
동해항	0.555	0.005	0.010	0.211	0.064	0.052	0.010	0.093
삼척항	0.230	0.002	0.004	0.064	0.013	0.021	0.003	0.664
목포항	0.064	0.001	0.002	0.032	0.006	0.007	0.002	0.887
속초항	0.036	0.001	0.002	0.022	0.005	0.007	0.001	0.926
옥계항	0.103	0.001	0.002	0.037	0.007	0.012	0.002	0.836
군산항	0.031	0.001	0.002	0.934	0.007	0.005	0.001	0.019
장항항	0.020	0.000	0.001	0.011	0.002	0.004	0.001	0.961
목포항	0.237	0.005	0.012	0.232	0.037	0.037	0.010	0.429
완도항	0.021	0.000	0.001	0.014	0.003	0.004	0.001	0.956
포항항	0.000	0.000	0.000	0.000	0.999	0.000	0.000	0.000
제주항	0.053	0.001	0.002	0.033	0.006	0.008	0.002	0.896
서귀포항	0.014	0.000	0.001	0.009	0.002	0.003	0.001	0.972
대산항	0.388	0.010	0.031	0.177	0.073	0.128	0.023	0.170
보령항	0.020	0.001	0.002	0.011	0.006	0.947	0.002	0.011
태안항	0.048	0.002	0.004	0.026	0.011	0.877	0.004	0.028

3.2 항만군집의 특성분석

우리 나라의 28개 무역항을 대상으로 클러스터링 한 결과 8개의 항만군집으로 분류할 수 있었으며 각 군집의 특성을 나타내는 클러스터 중심은 Table 3에 나타나있다.

Table 3 Centers of the 8 clusters

항목 \ 군집	X1	X2	X3	X4	X5
A	4,399	6,584	13,152,935	494,293	17,491,369
B	69,014	6,482	84,349,896	6,344,593	107,706,739
C	45,258	6,606	25,220,496	1,459,752	150,072,097
D	10,383	4,503	10,068,194	1,681,731	13,800,519
E	14,110	5,502	44,081,960	2,306,498	49,479,622
F	468	21,641	7,123,021	8,553	6,245,730
G	36,646	6,560	69,912,934	2,535,678	125,854,610
H	2,795	1,468	2,291,851	102,855	1,971,375

각 항만군집의 특성을 살펴보면, 군집 A에 속하는 항만의 경우 입출항 선박척수는 4,399척, 평균톤수는 6,584톤, 하역능력은 13,153천톤, 보관능력은 494천톤, 연간 화물처리량은 17,491천톤으로 A군집에 소속의 정도가 가장 높은 항만은 삼천포항이며 입출항 선박척수는 많지 않으나 평균톤수가 크고 화물처리량이 많음을 알 수가 있다.

군집 B에 속하는 항만의 경우 입출항 선박척수는 6만 9,014척으로 타 군집에 비하여 가장 많은 선박이 입출항하고 있으며, 하역능력은 84,350천톤, 보관능력은 6,345천톤으로 다른 군집에 비하여 가장 높게 나타나고 있다. 또한 연간 화물처리량이 107,707천톤으로 세 번째로 많은 화물을 처리하고 있음을 알 수 있다. B군집에 소속의 정도가 가장 높은 항만은 부산항으로 우리 나라의 수출입 물동량을 처리하는 대표적인 항만이다.

군집 C에 속하는 항만의 경우 입출항 선박척수는 45,258척으로 타 군집에 비하여 두 번째로 많은 선박이 입출항 하고 있으며, 연간 화물처리량이 150,072천톤으로 다른 군집에 비하여 가장 많은 화물을 처리하고 있다. C군집에 소속하는 정도가 가장 높은 항만은 울산항으로 연간 화물처리량이 가장 많은 것은 배후지역에 석유화학, 조선, 자동차, 비철금속 등의 공단을 지원하는 기능을 수행하고 있기 때문이라 할 수 있다.

군집 D에 속하는 항만의 경우 입출항 선박척수는 10,383척, 평균톤수는 4,503톤, 하역능력은 10,068천톤, 보관능력은 1,682천톤, 연간 화물처리량은 13,801천톤이며, D군집에 소속의 정도가 가장 높은 항만은 군산항이다.

군집 E에 속하는 항만의 경우 하역능력은 44,082천톤, 보관능력은 2,307천톤으로 타 군집에 비하여 세 번째로 높은 하역능력과 보관능력을 가지고 있으며, 연간 화물처리량은 49,480천톤으로 다른 군집에 비하여 네 번째로 많은 화물을 처리하고 있다. E군집에 소속하는 정도가 가장 높은 항만은 포항항으로 배후지역의 재철소 및 관련공업단지를 지원하는 기능을 수행하고 있다.

군집 F에 속하는 항만의 경우 입출항 선박척수는 468척으로 타 군집에 비하여 가장 적은 수의 선박이 입출항하고 있으나, 평균톤수는 약 21,641톤으로 다른 군집에 비하여 가장 크게 나타나고 있다. 또한 보관능력은 8,553톤으로 가장 낮은 것으로 나타난다. F군집에 소속하는 정도가 가장 높은 항만은 보령항이고, F군집에 소속하는 항만 중에서 부령, 태안항은 대중화물이 유·무연탄으로 배후지역의 화력발전소의 연료수송 지원 항으로 기능을 수행하고 있다.

군집 G에 속하는 항만의 경우 입출항 선박척수는 36,646척으로 다른 군집에 비하여 세 번째로 많은 선박이 입출항 하고 있고, 하역능력은 69,913천톤, 보관능력은 2,536천톤으로 타 군집에 비하여 두 번째로 높은 하역능력과 보관능력을 보유하고 있으며, 연간

화물처리량이 125,855천톤으로 다른 군집에 비하여 두 번째로 많은 화물을 처리하고 있음을 알 수 있다. G군집에 소속하는 정도가 가장 높은 항만은 광양항으로 포항항과 배후지역의 제철소 및 관련공업단지를 지원하는 기능을 수행하고 있음을 알 수 있다. 또한 G군집에 소속하는 정도가 높은 항만은 인천항으로 수도권 지역을 배후지로 두고 있어 부산항과 함께 우리나라의 수출입화물을 처리하는 대표적인 항만이다.

군집 H에 속하는 항만의 경우 입출항 선박의 평균 톤수는 1,468톤, 하역능력은 2,292천톤, 연간 화물처리량은 1,971천톤으로 8개의 군집 중에서 최하위를 차지하고 있음을 알 수 있다.

IV. 결론

본 연구에서는 퍼지 클러스터링 방법 중의 하나인 FCM 알고리즘을 이용하여 항만의 조건이나 특성이 유사한 항만들을 분류하고 각 항만군집의 특성을 파악하였다.

본 연구의 결과를 요약하면 다음과 같다.

- 1) 연간 입출항 선박의 척수 면에서는 부산, 울산, 인천, 광양항의 순으로 나타났으며, 입출항 선박의 평균 톤수 면에서는 거제, 보령, 태안, 대산항 순으로 나타나고 있다.
- 2) 입출항 선박의 평균톤수 면에서는 통영, 속초, 목포, 완도, 제주, 서귀포항이 평균톤수 1천톤 미만인 항만으로 나타나고 있으며, 부산, 인천, 광양, 삼천포, 울산, 동해, 포항, 대산항은 평균톤수가 5천톤 이상 1만톤 미만으로 항만이며, 거제, 보령, 태안항은 평균톤수 1만톤 이상인 항만으로 나타나고 있다.
- 3) 연간 화물처리량 측면에서는 울산, 광양, 인천, 부산, 대산, 포항항이 전체 물동량의 약 82%를 차지하고 있으며, 옥포, 거제, 진해, 통영, 고현, 속초, 장항, 완도, 서귀포항은 우리나라 28개 무역항 중에서 1999년 연간 화물처리량이 1백만톤 이하인 항만으로 나타나고 있다.
- 4) 하역능력은 부산, 광양, 인천, 포항, 울산항의 하역능력이 전체의 약 70%를 차지하고 있으며, 보관능력에 있어서는 부산, 인천, 포항, 광양, 마산항이 전체 보관능력의 약 70%를 차지하고 있다. 항만의 보관능력이 전혀 없는 항만은 옥포, 대산, 보령, 태안항으로 나타나고 있다.
- 5) 각 항만의 입출항 선박의 척수, 연간 화물처리량, 항만시설 등을 고려하여 항만의 조건과 특성이 유사한 항만들을 분류한 결과 우리나라의 28개 무역항은 8개의 항만군집으로 분류됨을 알 수 있었다.
- 6) 각 항만군집의 특성을 분석한 결과 같은 군집에 속하는 항만들간에는 실제 유사한 특성을 나타내고 있으며, 서로 다른 군집간에는 매우 다른 특성을 지니고 있음을 알 수 있었다.

IV. 참고문헌

- [1] 윤명오 외2, 항만의 분류 및 그 특성 분석에 관한 연구, 한국항해학회지, 제24권, 제4호, 2000, pp.247~254.
- [2] 금종수 외2, FCM법에 의한 항만의 분류 및 그 특성 분석에 관한 연구, 한국항만학회지, 제14권, 제2호, 2000, pp.143~154.
- [3] 문성혁·이준구, 주성분분석 및 군집분석을 이용한 컨테이너항만의 분류, 한국항만학회지, 제13권, 제1호, 1999, pp. 11-25.
- [4] J. L. Tongzon, Systematizing International Benchmarking for Ports, *Maritime Policy and Management*, Vol.22, No.2, 1995, pp. 171-177.
- [5] R. L. Cannon, J. V. Dave and J. C. Bezdek, Efficient Implementation of the Fuzzy C-Means Clustering Algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.8, No.2, 1986, pp. 248-255.
- [6] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.
- [7] H. Choe, J. B. Jordan, On the Optimal Choice of Parameters in a Fuzzy C-Means Algorithm, *IEEE International Conference on Fuzzy Systems*, San Diego 1992, pp. 349-354.
- [8] M. R. Anderberg, *Cluster Analysis for Applications*, Academic Press, New York, 1973.
- [9] J. C. Dunn, Well-separated Clusters and Optimal Fuzzy Partitions, *Journal of Cybernetics*, Vol.4, 1974, pp. 95-104.
- [10] M. P. Windham, Cluster Validity for Fuzzy Clustering Algorithms, *Fuzzy Sets and Systems*, Vol.5, 1981, pp.177~185.
- [11] 坂和正敏, ファジィ理論の基礎와 應用, 森北出版社, 1989, pp. 74-83.
- [12] 石岩, 水本雅晴, 湯崎崎直哉, ファジィc-平均 クラスタリング アルゴリズムによる ファジィ規則生成法の 改善, 日本ファジィ學會誌, Vol.9, No.4, 1997, pp. 525-532.
- [13] 宮本定明, ファジィクラスタリングあれこれ, 日本ファジィ學會誌, Vol.8, No.3, 1996, pp. 423-430.
- [14] 宮腰政明, ファジィクラスタリングの畫像解析への應用, 日本ファジィ學會誌, Vol.8, No.3, 1996, pp. 440-447.