
공간객체의 영향력을 고려한 클러스터링 알고리즘의 설계와 구현

Design and Development of Clustering Algorithm Considering Influences of Spatial Objects

김병철
남서울대학교 지리정보공학과
Byung-Cheol Kim(bckim@nsu.ac.kr)

요약

본 논문은 공간객체의 영향력을 고려한 클러스터링을 위한 알고리즘인 DBSCAN-SI를 제안한다. DBSCAN-SI는 기존의 DBSCAN과 DBSCAN-W를 확장한 것으로 공간클러스터링 시 비공간 속성들을 영향력으로 변환한다. DBSCAN-SI는 클러스터링에 사용되는 속성에 의한 영향력이 클수록 클러스터에 포함될 확률을 높여주어, 단지 공간적인 거리뿐만이 아니라 영향력의 크기를 반영하여 군집화를 수행하기 위한 알고리즘이다.

이 논문에서 제안한 클러스터링 기법은 주변에 있는 객체들이 특정 속성 중심으로 보았을 때, 영향력이 큰 객체임에도 불구하고 주변에 객체가 드물게 있으므로 인하여 클러스터에서 배제되게 되는 기존 알고리즘의 단점을 보완해 줄 수 있다.

■ 중심어 : 클러스터링, 공간객체, 영향력, 밀도 기반

Abstract

This paper proposes DBSCAN-SI that is an algorithm for clustering with influences of spatial objects. DBSCAN-SI that is extended from existing DBSCAN and DBSCAN-W converts from non-spatial properties to the influences of spatial objects during the spatial clustering.

It increases probability of inclusion to the cluster according to the higher the influences that is affected by the properties used in clustering and executes the clustering not only respect the spatial distances, but also volume of influences.

For the perspective of specific property-centered, the clustering technique proposed in this paper can makeup the disadvantage of existing algorithms that exclude the objects in spite of high influences from cluster by means of being scarcely close objects around the cluster.

■ keyword : Clustering, Spatial Object, Influence, Density-Based

1. 서론

클러스터링이란 실제적 또는 추상적인 객체의 집합을 클러스터라는 서로 유사성을 갖는 객체들의 분류로 그룹

화 하는 과정으로, 하나의 클러스터에 속하는 데이터 점들 간에는 서로 다른 클러스터내의 점들과는 구분되는 유사성을 갖게 된다[1].

클러스터링 방법은 크게 분할(partitioning) 방법[2]과

접수번호 : #081026-002
접수일자 : 2008년 10월 26일

심사완료일 : 2008년 11월 28일
교신저자 : 김병철, e-mail : bckim@nsu.ac.kr

계층적(hierarchical) 방법[3][4], 밀도기반(density-based) 방법[5][6], 격자기반(grid-based) 방법[7]과 모델기반(model-based) 방법으로 나눌 수 있다[8]. 그리고 많은 클러스터링 알고리즘은 관련된 응용시스템의 목적에 따라 가중치를 고려한다.

이중에서 밀도 기반 공간 클러스터링은 공간 데이터베이스 내에서 객체간의 거리, 연결성, 밀도를 기반으로 유사한 객체들을 그룹화 하는 것이다. 밀도기반 공간 클러스터링 알고리즘에 가중치를 고려하기 위해서 기존의 DBSCAN(Density-Based Spatial Clustering of Applications with Noise)을 확장한 DBSCAN-W(a DBSCAN algorithm using region expressed as Weight)[9]도 제안되었다.

그러나 기존의 알고리즘들은 어느 객체를 중심으로 하여 일정 영역안에 들어오는 이웃객체의 수, 즉 양에 따라서만 중심객체 여부를 판단하였다.

그러므로 이웃객체의 크기, 즉 질까지 동시에 고려하지 못했다. 본 논문에서는 한 객체가 중심객체가 되기 위한 조건으로 이웃객체의 수와 크기를 동시에 고려할 수 있는 밀도기반 클러스터링 알고리즘으로 DBSCAN-SI(a DBSCAN algorithm expanding the core object conditions as Sum of Influences)를 제안한다. DBSCAN-SI는 DBSCAN과 DBSCAN-W를 확장한 것

으로, 객체가 갖는 속성들에 의해 가중치를 산출하여 클러스터링 시 주변에 있는 이웃들의 수 뿐 만이 아니라 주변 이웃들의 가중치들의 합도 함께 고려할 수 있는 알고리즘이다.

II. 관련연구

1. DBSCAN

밀도 기반 군집화 방식인 DBSCAN은 Ester, Kriegel, Sander 과 Xu[10]에 의해서 제안되었다.

DBSCAN은 데이터 집합이 주어지면 클러스터와 어느 클러스터에도 속하지 않은 데이터 점들인 잡음을 식별한다. DBSCAN 알고리즘은 다음과 같이 정의하고 있다.

정의 1. 한 점 p 의 Eps -neighborhood는 p 로부터 반경 Eps 내에 있는 이웃의 집합이다.

여기에서 Eps 는 점 p 를 중심으로 그려지는 원의 반지름(r)이 된다.

정의 2. MinPts(minimum number of points)는 최소 이웃수로 한 점의 Eps -neighborhood의 수가 최소 이웃수 이상인 경우 이 점을 중심객체라 한다.

주어진 점들의 집합 D 에서 한 점 p 가 q 로부터 직접 밀도 도달 가능(directly density-reachable)하다는 의미는

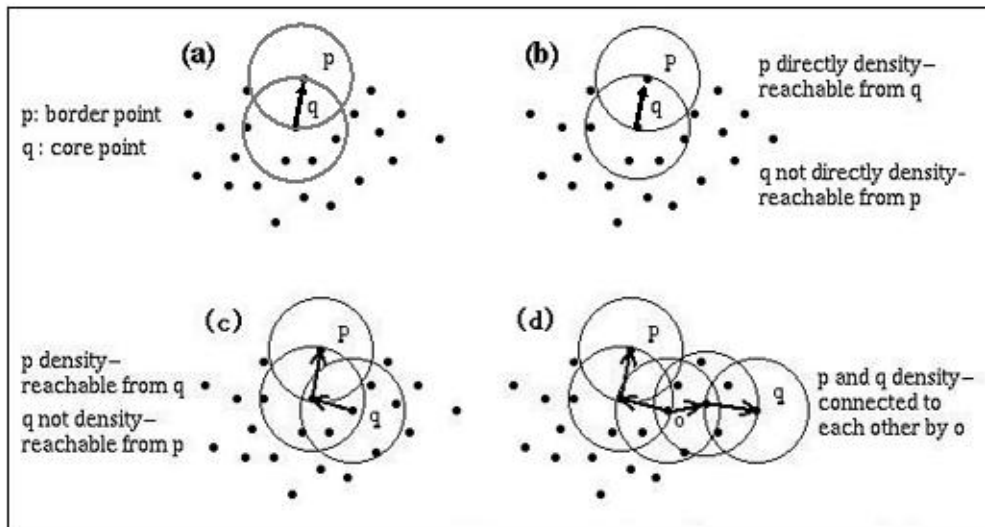


그림 1. DBSCAN 관련용어

p 가 점 q 의 Eps-neighborhood 내에 있고 q 가 중심객체인 경우이다.

정의 3. 한 점 p 가 점 q 로부터 밀도 도달 가능(density-reachable)하다는 의미는 p 에서 시작하고 q 에서 끝나는 직접 밀도 도달 가능한 연결(chain)이 존재한다는 의미이다.

정의 4. 한 점 p 가 q 로부터 밀도 연결(density-connected)하다는 의미는 p 와 q 로부터 밀도 도달 가능한 점 o 가 존재한다는 의미이다. 클러스터는 밀집 연결된 점들의 최대 집합이다.

2. DBSCAN-W

밀도 기반 공간 클러스터링 알고리즘에 가중치를 고려하기 위해서 기존의 DBSCAN을 확장한 DBSCAN-W(a DBSCAN algorithm using region expressed as Weight)을 제안했다[9]. DBSCAN-W에서 재 정의한 개념은 다음과 같다.

모든 객체는 해당 응용시스템에서 그 객체가 갖는 중요도에 따라서 서로 다른 크기의 원으로 표현되는 영역을 갖는다. 이때 원의 중심 좌표값은 공간상에 객체의 위치 좌표이고, 반지름은 가중치로 선택된 비공간 속성에 의해 결정된다.

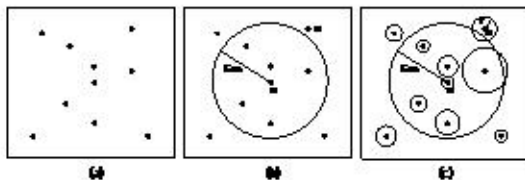


그림 2. DBSCAN과 DBSCAN-W 비교

한 객체 p 의 Eps-neighborhood는 p 의 중심점으로부터 반경 Eps 내에 각 객체를 표현하는 영역이 겹쳐지는 이웃들의 집합이다.

[그림 2]는 기존의 DBSCAN 알고리즘과 DBSCAN-W에서 객체의 Eps-neighborhood를 결정하는 과정을 비교한 것이다. [그림 2(a)]는 점으로 표시된 객체의 분포를 나타낸 것이다. [그림 2(b)]는 DBSCAN에서 객체 p 의 Eps-neighborhood를 구하는 과정이다. 객체 p 의

Eps-neighborhood는 객체 p 를 중심으로 반경 Eps 내에 포함된 5개의 객체들이다.

[그림 2(c)]는 DBSCAN-W에서 Eps-neighborhood를 결정하는 과정으로 이때 각 객체들은 서로 다른 크기의 원으로 표현된다. 즉 중심점은 객체의 공간 속성인 위치 값(x, y)이고, 반지름 r 은 응용 프로그램의 목적에 따라 사용자가 가중치의 기준으로 선택한 비공간 속성에 의해서 결정된 값이다. 객체 p 의 Eps-neighborhood는 객체 p 를 중심으로 반경 Eps 내에 포함된 5개의 객체들과 영역이 겹치는 q 를 포함하여 6개 객체가 된다.

III. 가중치의 합을 고려한 밀도 기반 클러스터링 알고리즘

DBSCAN-SI는 각 대상물의 위치 속성뿐 아니라 그 점으로 표현된 객체가 갖는 다양한 비공간 값들을 클러스터링에 고려한다. 객체가 갖는 많은 속성들의 값과 가중치를 고려해 계산한 영향력을 원의 크기로 부여하여 영향력이 큰 경우 클러스터에 소속될 확률을 높이거나 영향력의 합이 큰 경우 클러스터를 구성할 수 있도록 알고리즘을 설계하였다.

DBSCAN-SI는 이웃의 개수를 정하는데 기준이 되는 거리를 확장하여 영향력이 큰 값들은 주변 객체들의 이웃이 될 확률을 높여주는 방법과 이웃들의 영향력의 합에 의해 중심객체를 결정하는 방법을 병행하는 알고리즘이다. 이를 DBSCAN-SI(1)과 DBSCAN-SI(2)로 구분한다.

1. DBSCAN-SI(1)

DBSCAN-SI(1)은 객체가 갖는 여러 속성들을 영향력 값으로 변환하여, 영향력이 큰 객체일수록 주변 객체를 이웃으로 삼을 확률(흡입률)을 높여주고, 또한 주변에 있는 객체는 영향력이 클수록 다른 객체의 이웃이 될 확률(흡수율)을 높여주도록 설계되었다. 여기에서 영향력이란 객체의 크기를 의미한다. DBSCAN에서 객체는 하나의 점으로 표현되어 지지만 DBSCAN-SI에서 점 p 는 크기가 서로 다른 동심원이 되어진다. 그러므로 가중치가 큰 객체 p 는 함수에 의해 계산되어진 영향력 값인 자신

의 반지름만큼 Eps를 확장하여 주변 객체를 더 많이 Eps-neighborhood가 되도록 하고, 또한 이웃으로 고려되는 객체 q는 영향력에 대응하는 만큼 원의 크기를 크게 하여 주변 객체의 이웃이 될 확률을 높여준다.

이때 원의 중심 좌표값은 공간상의 객체의 위치 좌표이고 반지름은 영향력 값으로, 위치속성외의 다른 속성(비공간 속성 포함) 값에 의해 계산된다.

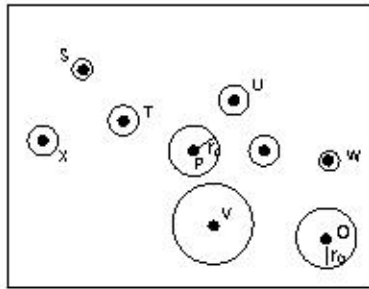


그림 3. 영향력 반영

정의 5. 한 점의 이웃을 정하는 Eps'의 길이는 DBSCAN과 DBSCAN-W 알고리즘에서의 Eps에 중심점의 반지름을 더한 값으로 한다.

$$Eps' = Eps + r_p \quad (1)$$

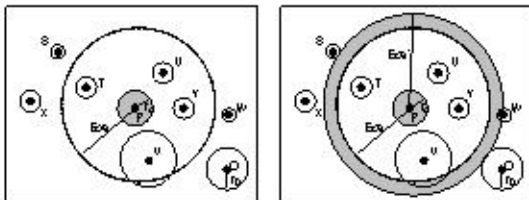


그림 4. Eps 확장

[그림 4(a)]는 DBSCAN 및 DBSCAN-W에서의 Eps로 객체 p로부터의 거리이다. [그림 4(b)]는 DBSCAN-SI에서의 Eps'로 기존 알고리즘보다 객체 p의 반지름만큼 더 길다. 객체 p가 갖는 영향력에 비례하여 주변 객체를 이웃으로 삼는 흡입력을 증가시키기 위한 것이다.

정의 6. 한 점 p의 Eps'-neighborhood는 p로부터 반경

Eps'내의 영역과 각 점들의 영향력으로 표현된 원들이 겹쳐지는 점들의 집합이다. 여기에서 dist()함수는 공간상의 유클리드 거리를 구하는 함수이다.

$$NEps(p) = \{q \in D \mid dist(p,q) \leq (Eps + r_p + r_q)\}$$

$$- Eps\text{기준: } dist(p, q) \leq Eps + r_p + r_q$$

$$- Eps'\text{기준: } dist(p, q) \leq Eps' + r_q$$

[그림 4(a)]에서 점 p의 이웃은 T, U, V, Y이지만 [그림 4(b)]에서 점 p의 이웃은 T, U, V, Y와 Q, S, W를 포함한다. 즉 객체 Q, S, W가 Eps' 범위 밖에 있지만, 영향력이 클수록 그만큼의 거리를 삭감해 주는 효과를 통해 이웃이 될 확률을 높여주는 것이다.

2. DBSCAN-SI(2)

DBSCAN-SI(2)는 객체가 갖는 여러 속성들을 영향력 값으로 변환하여, 어느 한 점이 이웃의 수에 의해 중심객체 여부가 결정되는 DBSCAN과 DBSCAN-W 알고리즘을 보완하여 이웃 객체의 수는 적더라도 이웃들의 영향력의 합이 인자로 주어지는 최소 영향력 이상이라면 그 점을 중심 객체가 되게 한다.

정의 B. MinPts는 최소이웃수를 의미하고, MinInf는 최소이웃의 영향력의 합을 의미한다. 이때 한 점의 Eps'-neighborhood의 수가 MinPts 이상인 경우이거나, 최소이웃의 영향력의 합이 MinInf 이상인 경우 이점을 중심객체라 한다. 여기에 자신의 영향력 값도 포함한다.

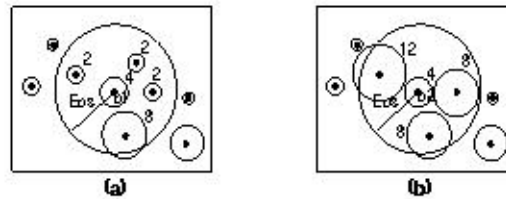


그림 5. DBSCAN-SI

[그림 5]에서 (a)는 점 p를 기준으로 Eps 반경내의 이웃 객체의 수가 4이고, 이웃들의 영향력의 합은 자신을 포함하여 18이다. 그리고 (b)의 경우는 이웃 객체의 수가 3이고, 이웃들의 영향력의 합은 자신을 포함하여 32이다.

여기에서 중심 객체의 조건으로 이웃의 수가 4이상이거나 이웃의 영향력 값의 합이 30 이상인 경우로 설정하게 되면 그림 5의 (b)도 중심객체가 된다. 이는 주변에 객체의 수는 비록 적더라도, 객체들이 비중 있는(영향력 값이 큰) 경우라면 중심객체가 되게 하여 클러스터에 포함되도록 하는 효과가 있다. 또한 MinPts를 최대화 할 경우는 이웃의 영향력의 합에 의해서만 중심객체를 선정하게 된다.

DBSCAN-SI(1,2)를 위한 알고리즘은 다음과 같다.

Input:
 $D = \{ \delta_1, \delta_2, \dots, \delta_n \}$ // 객체들의 집합(n 은 객체의 수)
 MinPts // 중심객체가 되기 위한 최소 이웃의 수.
 MinInf // 중심 객체가 되기 위한 이웃의 최소 영향력의 합.
 Eps' // 이웃을 정하는 DBSCAN-SI의 임계값.
Output:
 $K = \{ K_1, K_2, \dots, K_k \}$ // 클러스터의 집합(k 는 클러스터의 수, $k \leq n$).

그림 6. DBSCAN-SI 입출력 파라미터

추가된 입력 파라미터는 MinInf로 이는 객체 p 가 중심 객체가 되도록 하는 조건으로 Eps' 반경내의 객체들이 갖는 영향력의 합이 MinInf 이상인 경우 중심 객체가 된다.

IV. 구현결과

1. 실험데이터

실험데이터 집합은 제안하는 알고리즘의 특성을 실험하기 위하여 VB6.0으로 구현된 실험데이터 생성기를 이용하여 데이터 집합을 생성하였다.

클러스터링을 위한 데이터의 구성은 어느 도시에 거주하는 고객들의 공간적 클러스터를 찾고자 하는데, 여기에 가중치로 각 고객들의 거래금액을 고려하는 것을 가

정한다. 데이터 생성을 위한 가정은 산과 같은 장애물에 의해 일부 위치가 구분되고, 도로 등의 여건에 의해 밀집되는 현상을 반영하였다. 일반적으로 고객별 거래금액은 랜덤하다는 기본 전제에 고객의 공간적 밀집도는 다소 낮지만 고객별 거래금액의 규모가 큰 지역이 포함된 경우를 포함하였다. 주어진 데이터는 고객의 공간적 분포 형태와 가중치인 거래금액을 고려한 것으로 [그림 7]과 같이 표현되어 질 수 있다.

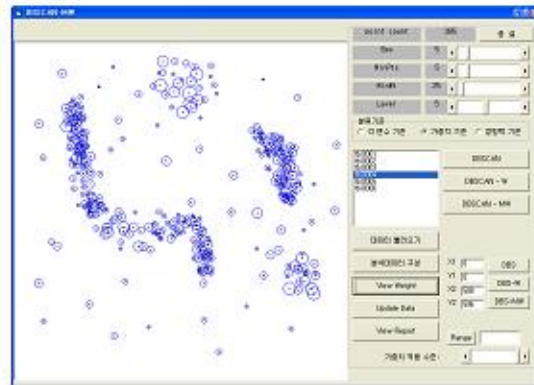


그림 7. 영향력 척도

[그림 7]은 모든 속성의 영향력의 합에 의해 영향력을 표현한 것이다. DBSCAN-SI는 이 영향력 값을 반영하여 클러스터링을 수행한다.

2. 클러스터링 결과

클러스터링 조건으로 Eps', MinPts, MinInf의 값을 부여하게되는데, 여기서 Eps'는 이웃으로 인정되는 거리를 말하고, MinPts는 어느 한 객체가 중심객체가 되기 위한 최소 이웃객체의 수이며, MinInf는 어느 한 객체가 중심 객체가 되기 위한 이웃객체들의 영향력의 합의 최소값이다.

다음은 Eps', MinPts, MinInf의 서로 다른 조건하에서 DBSCAN과 DBSCAN-W, DBSCAN-SI 각각의 알고리즘으로 클러스터링을 수행한 결과이다. 입력 파라미터는 DBSCAN와 DBSCAN-W의 경우는 Eps, MinPts이고, DBSCAN-SI는 Eps', MinPts, MinInf이다.

그리고 주어지는 데이터로는 분석 대상인 공간 데이터와 함께 DBSCAN-W의 경우에는 가중치를 의미하는 하

나의 속성 값이 추가적으로 사용되고, DBSCAN-SI는 위치 속성외의 여러 속성의 각 가중치를 고려하여 도출된 복수의 영향력들의 합이 사용된다.

2.1 DBSCAN 수행 결과

Eps = 5, MinPts = 5의 조건하에서 [그림 8(a)]과 같이 4개의 클러스터가 생성되었다.

2.2 DBSCAN-W 수행 결과

[그림 8(b)]는 DBSCAN-W의 수행결과로 Eps = 5, MinPts = 5의 조건하에서 클러스터링을 수행하였고, 비공간 속성인 Weight를 고려하여 수행하였다. DBSCAN에서와 같이 4개의 클러스터를 발견하였지만 클러스터의 위치나 크기가 다르다. DBSCAN에서 두 개의 클러스터로 나뉘었던 부분이 하나의 클러스터로 합쳐졌고, DBSCAN에서 발견하지 못했던 영역을 클러스터로 발견하였다.

DBSCAN-W는 Eps 범위 밖에 위치한 객체라도 그 객체의 가중치가 큰 경우 이웃 객체가 될 수 있도록 하는 알고리즘이다.

2.3 DBSCAN-SI(1) 수행 결과

[그림 8(c)]은 Eps = 5, MinPts = 5, MinInf = 999의 조건에서 DBSCAN-SI(1)로 클러스터링을 수행한 결과로, MinInf를 아주 크게 부여하여 Eps와 MinPts에 의해서만 클러스터링이 이루어지도록 한 것이다.

클러스터링 결과 클러스터의 수는 4개이지만 DBSCAN-W에서의 서로 다른 두 개의 클러스터가 병합

되었고, 상단 부분에 새로운 클러스터를 발견하였다. 상단 부분은 객체간 거리는 다소 멀지만 각 객체들이 갖는 영향력이 큰 부분으로 DBSCAN-SI(1)은 이런 객체들도 클러스터 안에 포함되도록 하는 알고리즘이다.

2.4 DBSCAN-SI(2) 수행 결과

DBSCAN-SI(2)는 Eps = 5, MinPts = 5, MinInf = 15인 조건에서 클러스터링을 수행한 것으로 MinPts와 MinInf를 동시에 고려한다.

[그림 8(d)]는 DBSCAN-SI(2)에서 MinInf를 적용하여 클러스터링을 수행한 결과로 클러스터의 수는 마찬가지로 4개이다. DBSCAN-SI(1)에 비해 클러스터의 영역이 다소 확장되었다.

주목할 부분은 데이터의 상단 부분으로 DBSCAN-SI(1)에서는 주변으로 갈수록 인접한 이웃객체의 수가 적어서 중심객체가 되지 못하므로 주변의 객체들은 영향력이 크어도 불구하고 클러스터에 포함되지 못하는 결과가 되었다.

그러나 DBSCAN-SI(2)는 인접한 이웃객체의 수는 적더라도 이웃객체들의 영향력의 합이 일정 수준(MinInf) 이상이 되면 중심객체가 될 수 있도록 한다. DBSCAN-SI(2)는 최소 이웃객체의 수(MinPts)와 최소 이웃객체의 영향력의 합(MinInf)을 동시에 고려한다.

3. 알고리즘별 특징 비교

[그림 9]는 각 알고리즘에 의해 클러스터링 된 객체들을 각 객체가 갖는 영향력의 값에 의해 분류한 것을 그래프화 한 것이다.

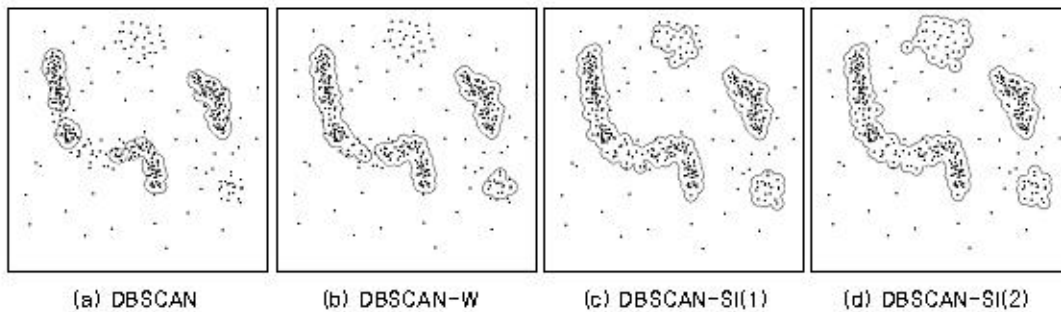


그림 8. 실험 결과

각 객체의 영향력 값을 백분율로 변환하여 클러스터링 한 경우, 각 알고리즘별로 영향력 값의 구간에 따라 클러스터를 구성하는 점들의 수를 누적하여 그린 것이다. 이는 영향력 값이 큰 객체 일수록 DBSCAN-SI(1)과 (2)에 의해 클러스터 안에 포함될 확률이 높아진다는 알고리즘의 특징을 확인하기 위한 것이다. 즉 영향력이 작은 객체들은 각 알고리즘별 차이가 두드러지지 않을 것이고, 영향력이 큰 객체들일수록 각 알고리즘별 차이가 나타날 것이라는 가설을 확인하기 위한 것이다. 그러나 이는 데이터 셋의 구성에 따라 그 결과는 알고리즘 별 차이가 거의 없을 수도 있다. 즉, 객체들의 영향력이 거의 같아서 서로 차별성이 없을 경우에 해당된다. [그림 9]에서는 알고리즘별 차이가 나타나고 있으며, 50%이상에서 DBSCAN-SI(1)과 (2)의 객체수가 상대적으로 증가함을 볼 수 있다.

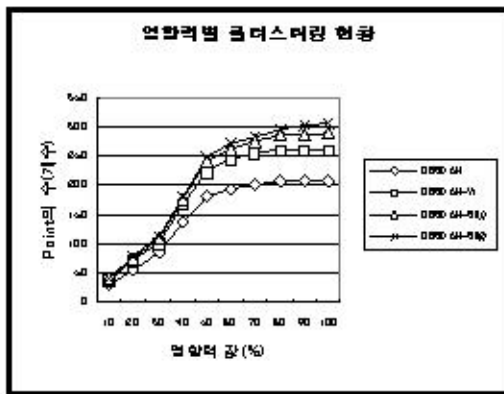


그림 9. 영향력별 클러스터링 현황

V. 결론

클러스터링이란 실제적 또는 추상적인 객체의 집합을 클러스터라는 서로 유사성을 갖는 객체들의 분류로 그룹화 하는 과정으로, 데이터 마이닝 뿐만 아니라 통계학, 생물학, 기계학습, 공간분석 등에 많이 사용되고 있으며, 최근 효과적으로 대량의 데이터를 취급할 수 있는 많은 방법들이 개발되었다.

본 논문은 각 객체가 갖는 영향력을 고려한 밀도기반

알고리즘을 설계하고 구현한 것으로 대상 객체들이 갖는 속성값들을 영향력으로 변환하여, 클러스터의 조건을 이웃의 개수에만 두는 것이 아니라, 이웃의 크기도 함께 반영하도록 하였다. 그리고 실험을 통해 각 알고리즘별 차이를 살펴보았다. 본 논문에서 제안한 DBSCAN-SI 알고리즘은 인구수, 교통, 지가, 환경, 균형발전 등 매우 많은 비공간 속성을 고려하여 입지선정이나 상권분석 등을 수행하는 경우 현실적인 문제를 해결하는데 사용되어질 수 있다.

그러나 객관적인 영향력 값의 도출이 쉽지 않으므로 향후 대상 오브젝트가 갖는 많은 속성들로부터 영향력을 도출해 내는 객관적이고 가장 합리적인 방법에 대한 연구로 확장할 것이다.

참고문헌

- [1] J. Mchae, A. Berry, and Gordon Linoff, *Data Mining Techniques: For Marketing, Sales, and Customer Support*, John Wiley & Sons, Inc., 1997.
- [2] R. T. Ng and J. Han, "Efficient and Effective Clustering Method for Spatial Data Mining", In Proc. of the VLDB Conf. on Santiago, Chile, pp.144-155, Sep. 1994.
- [3] T. Zhang, R. Ramakrishnan, and M. Livny, "BIRCH: An Efficient Data Clustering Method for Very Large Database," Proc. of the ACM SIGMOD Conf. on Management of Data, Montreal, Canada, pp.103-114, June 1996.
- [4] S. Guha, R. Rastogi, and K. Shim, "CURE: An Efficient Clustering Algorithm for Large Database," Proc. of the ACM SIGMOD Conf. on Management of Data, Seattle, Washington, USA, pp.73-84, May 1998.
- [5] M. Ester, H. P. Kriegel, J. Sander, and X. W. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," Proc. of ACM SIGMOD 3rd Int. Conf. on

- Knowledge Discovery and Data Mining, pp.226-231, 1996.
- [6] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, 2001.
- [7] M. H. Huh and Y. G. Lee, "Input Variable Importance in Supervised Learning Models," *The Korean Journal of Applied Statistics*, Vol.10, pp.239-246, 2003.
- [8] K. P. Sankar, "Unsupervised Feature Evaluation: A Neuro-Fuzzy Approach," *IEEE Transaction on Neural Networks*, Vol.11, No.2, pp.366-376, 2000.
- [9] 김호숙, 임현숙, 용환승, "공간 데이터 마이닝에서 가중치를 고려한 클러스터링 알고리즘의 설계와 구현", *한국지능정보시스템학회 논문지*, 제8권 제2호, pp.177-187, 2002.
- [10] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Database with Noise," *Proc. of Int. Conf. on Knowledge Discovery and Data Mining*, 1996.

저자 소개

김 병 철(Byung-Cheol Kim)

정회원



- 1992년 2월 : 청주대학교 응용통계학과 (경제학사)
- 1997년 8월 : 청주대학교 전자계산학과 (공학석사)
- 2005년 8월 : 충북대학교 전자계산학과 (이학박사)

•2006년 8월 ~ 현재 : 남서울대학교 지리정보공학과 겸임교수

<관심분야> : 데이터마이닝, 시공간데이터베이스, GIS, 공간분석, 기업정보화