

불확실한 경계를 가진 시공간 객체의 모델 설계[†]

김대중, 지정희, 류근호
충북대학교 데이터베이스 연구실

e-mail:{kimdj, jhchi, khryu}@dblab.chungbuk.ac.kr

A Spatio-temporal model for simple objects with a broad boundary

Dae Jung Kim, Jung Hee Chi, Keun Ho Ryu
Database Laboratory, Chungbuk National University

요 약

최근 지리 정보 시스템이나 공간 데이터 베이스 시스템이 발전하면서 실세계를 효율적으로 모델링하려는 많은 연구가 수행되고 있다. 기존의 단순한 point, line, polygon으로 모델링했던 방법들을 multipoint, multiline, multipolygon과 같은 복잡한 객체를 표현하기 위한 연구로 확장되었다. 또한 이러한 공간 객체의 이력을 효율적으로 관리하기 위해 공간 객체를 시간으로 확장하는 연구가 수행되고 있다. 이러한 모델들은 명확한 경계를 가진 시공간 객체를 대상으로 한다. 자연, 사회 문화 현상, 생태 같은 경계가 명확하지 않은 객체를 단순한 경계를 가진 공간 객체로 모델링하기에는 그 표현력이 부족하다. 따라서 이 논문에서는 이런 문제점을 해결하기 위해서 불확실한 경계를 가진 시공간 데이터 모델을 제시한다.

이 논문에서 제시한 불확실한 경계를 가진 시공간 모델의 설계는 상호 운용을 증진하기 위하여 OGC(OpenGIS Consortium)에서 제시한 2차원 기하 객체 모델을 기반으로 불확실한 경계를 가지는 공간 데이터 모델로 확장한다. 또한, 이 불확실한 경계를 가지는 단순한 공간 모델을 시간 영역으로 확장하여 유효시간을 갖는 불확실한 시공간 데이터 모델을 설계한다. 또한, 이 논문에서 제시한 불확실한 경계를 가진 시공간 객체 모델을 바탕으로 실세계의 다양하고 복잡한 객체에 대한 효율적인 질의가 가능한 시공간 연산자를 제시한다.

1. 서 론

최근 몇 년 동안 공간 데이터 베이스에 대한 관심이 고조되면서 효율적으로 시공간 데이터를 모델링 하기 위한 많은 연구들이 수행되어져 왔다[10, 11]. 이러한 연구들의 대부분은 기존의 관계형 데이터베이스를 확장하여 시공간 데이터를 모델링하거나 객체 지향 데이터베이스를 기반으로 시공간 데이터를 모델링하기 위한 연구들이 [11]였다. 이러한 연구들의 대상이 되는 공간 객체는 길, 가옥, 필지와 같은 명확한 경계를 갖는 객체들이며, [3]에서는 이러한 개체들을 determinate 공간 객체라고 정의하고 있다.

그러나, 명확한 경계를 갖는 공간 객체만으로는 복잡하고 다양한 실세계를 표현하는 데에는 한계가 있다. 예를 들어, 태풍을 표

현하다고 할 경우 기존의 모델링 방식을 사용하면, 태풍은 단순한 polygon이나 region으로 표현된다. 따라서 단순한 선으로 표현한 태풍의 경계 밖의 인접 지역은 태풍의 영향권에서 벗어났다고 정의한다. 그러나 실세계의 자연 현상을 명확한 경계로 표현하는 것은 문제가 있다. 왜냐하면, 인근 지역 역시 태풍의 피해를 받으며, 다만, 정의된 태풍의 경계를 벗어났을 뿐이다. 이렇게 공간 객체의 경계는 그 객체에 대한 속성, 관찰 방법 등에 따라 달리 정의 되어질 수 있다. 이런 문제를 해결하기 위해서 경계에 대한 buffering 기법이나 Minimum bounding Rectangle(MBR)기법 등 불확실한 경계로 표현되는 공간 객체를 모델링하기 위한 많은 연구들이 수행되었다[3, 4, 5]. 이들은 불확실한 경계로 표현되는 공간 객체를 vague, indeterminate 공간 객체 또는 Broad Boundary 공간 객체로 정의하고 있으며, 이러한 객체를 모델링하기 위한 연구를 수행하였다. 이러한 연구를 기반으로, 토지 속성의 변화(인구 밀도, 토질)와 같은 자연, 사회, 또는 문화 현상, 생태 등을 보다 효율적으로 모델링 할 수 있다.

[†] 이 연구는 과학재단 2001년 목적 기초(R01-1999-00243)의 연구비 지원으로 수행되었음.

그러나, 이러한 모델링 방식들은 불확실한 경계를 가진 공간 객체들을 잘 표현하고 있지만, 이러한 객체들의 시간에 따른 변화를 관리하기 위한 연구들은 아직 수행되지 않았다. 따라서 이 논문에서는 불확실한 경계를 가진 공간 객체의 이력에 관한 질의를 효율적으로 처리하기 위한 시공간 모델과 연산을 제시한다.

효율적인 전개를 위해서 2장에서는 불확실한 경계를 가진 공간 객체에 대한 관련 연구를 시작으로 3장에서는 기존의 시공간 데이터 모델의 공간 모델을 불확실한 경계를 가지는 공간 모델로 확장하여 시공간 데이터 모델을 설계한다. 4장에서는 이 논문에서 제시한 모델을 기반으로 시공간 연산자를 정의 및 분류하며, 끝으로 5장에서는 이 연구의 결과와 향후 연구 방향을 제시한다.

2. 관련 연구

2.1 불확실한 공간 데이터베이스 모델

불확실성은 임의의 공간 객체나 현상을 분류하거나 정의할 때 발생하게 되며, 이러한 불확실성은 퍼지 집합론이나 퍼지 모델분야에서 많은 연구가 진행되었다[3]. 불확실성에 대한 정의는 학자마다 서로 다른 견해 차이를 보이지만, 일반적으로 공간 객체의 불확실성은 크게 Uncertainty와 Fuzziness로 구분되어진다. 먼저, Uncertainty는 객체의 정확한 측정이 불가능하거나 그 경계의 정확한 위치와 모양이 확실하지 않은 경우이며, Fuzziness는 임의의 영역을 가지고는 있으나 명확한 경계를 가지지 않는 객체의 모호성을 나타낸다[10]. 이러한 불확실한 경계를 갖는 공간 객체를 모델링하기 위해 공간의 축소, 단순한 경계로 공간 기하 객체의 표현(point, line, region)등의 추상화 과정을 거치게 된다. 일반적으로 불확실한 경계를 갖는 객체의 모델링은 3가지 방법으로 설계된다.

- ① fuzzy model : 퍼지 집합론과 모델을 기반으로 한다.
- ② probabilistic model : 개연성 이론과 위치, 측정의 불확실성 모델에 기반을 둔다.
- ③ exact model : 불확실한 경계를 갖는 공간 객체에 대해서 명확한 경계를 갖는 공간 객체 모델로 전환된다[3].

위에 기술된 모델 중 fuzzy와 probability 모델을 공간 데이터를 설계하는데 사용되지 못하고 있다. 왜냐하면, 현재의 컴퓨터 기술이 uncertain과 fuzzy개념으로 공간 데이터의 효율적인 처리 과정을 수행하지 못하기 때문이다. 따라서 이 논문에서의 모델링 방식은 exact 모델을 기반으로 불확실한 경계를 가진 객체를 모델링하는 방식을 사용한다.

2.2 불확실한 경계를 가진 공간 객체의 정의

[7]에서는 불확실한 경계를 가진 region, line, point에 대해서 정의하고 있으며, 이 논문에서는 [7]에서 수행된 불확실한 경계를 가진 객체에 대한 정의와 표기법을 수용한다. 크게 두 가지로 구분이 되는데 불확실한 경계를 가진 region과 불확실한 경계를 가진 line으로 구분된다.

- ① 불확실한 경계를 가진 region

불확실한 경계를 가진 region 을 A라고 가정할 때, A는 내부 경계(∂A_i)와 외부 경계(∂A_o)를 가지며, 두 경계가 이루는 닫혀진 경계 즉, 불확실한 경계를 ΔA 라 정의한다, A의 내부와 외부는 open set이 되고 불확실한 경계인 ΔA 는 closed set이 된다. 단순한 영역은 내부 경계와 외부 경계가 일치하는 즉, $\Delta A = \partial A$ 인 경우이며, 불확실한 경계를 가지는 region의 특수한 경우이다.

- ② 불확실한 경계를 가진 line

2차원 평면에서 line에 대한 불확실한 경계의 종류는 두 가지로 나누어진다. 첫째, 전체 line의 위치를 모델링 하는 관점과 line의 양 끝점의 위치를 모델링 하는 관점이 있다.

전자의 불확실한 경계는 전체 line을 둘러싸고 있는 region이 되고, 이것을 broad line라 정의한다. 또한 후자 경우, 불확실한 경계는 각각의 끝점에 대한 region으로 구성되며, 불확실한 경계를 가진 line으로 정의한다. broad point는 broad line의 특수한 경우로 정의된다.

3. 불확실한 경계를 가진 시공간 데이터 모델의 설계

이 논문에서 제시하는 불확실한 경계를 가진 시공간 데이터 모델을 위해서는 불확실한 경계를 가지는 공간 데이터의 모델이 필수적이다. 불확실한 경계를 가진 공간 데이터의 모델은 그 객체에 대한 불확실성 기준의 선정이 그 객체의 경계에 대한 정의와 해석에 영향을 끼친다. 이 불확실성의 기준이 되는 것은 크게 객체의 특성, 객체에 대한 관찰 방법, 사용자의 목적에 의해 영향을 받는다. 객체의 특성은 객체가 연속, 불연속 객체인지에 따라 또는 고정객체인지 이동 객체인지에 따라 영향을 받는다. 또한 측정이나 해상도와 같은 관찰방법에 따라 불확실성의 기준은 달라진다.

3.1 설계 원칙

- 1) 객체 지향 개념을 사용한다.

객체 지향 데이터 모델은 객체, 클래스, 캡슐화, 상속, 그리고 다형성 등의 객체 지향 패러다임에 기반 한다. 이 개념은 동일한 객체의 이력들을 하나의 단일 엔티티에 내장시키는 것을 가능하게 한다. 또한 추상 클래스의 속성과 메소드는 하위 클래스에 상속된다.

- 2) 기존의 시공간 객체 모델을 불확실한 경계를 가지는 시공간 모델로 확장한다.

기존의 연구된 시공간 데이터 모델을 최소한으로 확장하기 위하여 이 논문에서는 기존의 연구된 명확한 경계를 갖는 시공간 객체 모델을 확장하여 불확실한 경계를 갖는 객체도 모델링 할 수 있도록 한다.

- 3) 불확실한 경계를 가진 객체의 이력 질의를 지원한다

시공간 객체를 이력 객체로 관리함으로써 시공간 객체에 대한 보다 정확한 접근이 가능하다.

이 논문은 이런 설계 원칙에 따라 [그림 1]과 같이 유효시간을 지원하는 불확실한 경계를 가진 데이터 모델을 제시한다. 이 모델

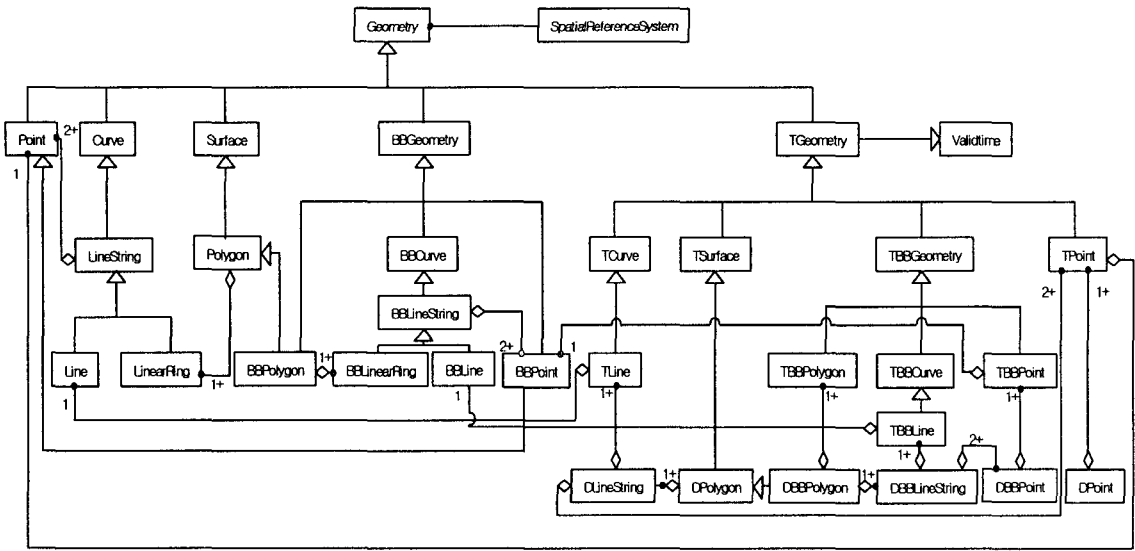


그림 1. 불확실한 경계를 가지는 시공간 기하 객체의 클래스 계층도

은 공간 참조모델로는 OpenGIS의 명세서에서 나타나 있는 단순한 2차원 공간 객체 모델을 기초로 불확실한 경계를 가진 시공간 모델을 설계한다. 이 모델에서는 시공간 객체 클래스를 TGeometry라 정의하고 Geometry 클래스와 Validtime 클래스로부터 각각의 공간과 시간의 속성을 상속받는다. BBGeometry 클래스는 불확실한 경계를 갖는 객체에 대한 클래스로 Geometry 클래스를 상속받고, BBPoint, BBCurve, BBPolygon를 포함한다. 또한, 각각 공간 객체의 유효시간을 갖는 클래스를 TCurve, TSurface, TBBGeometry, TPoint로 하며, 각 시공간 객체의 유효 시간에 대한 버전들의 집합을 DLineString, DPolygon, DBBPolygon, DBBLineString, DBBPoint, DPoint로 한다. BBPoint는 BBCurve의 특수한 경우이다[8, 9, 12].

4. 불확실한 경계에 대한 시공간 연산자

이 장에서는 3장에서 설계한 불확실한 경계를 가진 시공간 데이터 모델에서 수행될 수 있는 연산자를 제시한다.

기존의 위상관계 연산자는 객체간의 내부, 경계, 외부간의 교차 결과 값을 이진값(0,1)으로 나타내고 각각의 위상관계를 3x3의 행렬로 표현하였다. 명확한 경계를 가지는 두개의 region간에는 기본적으로 8개의 위상관계가 disjoint, meet, overlap, coveredBy, inside, covers, contains, equal 이라 정의된다[5, 7].

하지만 9-IM을 이 논문에서 제시한 모델에 적용시키기 위해서는 기존의 단순한 경계를 불확실한 경계를 가진 경계(Δ)로 확장시킨 방법을 사용한다. 이런 방법으로 공간 모델에 적용시킨 [7]에서는 기하 조건에 따라 불확실한 경계를 가지는 단순한 region에 대한 경우의 수가 44가지임을 보여준다. [9]에서는 이렇게 많은 경우의 수를 유사한 기하 속성과 기하 거리 함수에 의해 클러스터링하여 위상

관계를 명명하였다. 이 논문은 [9]에서 제시한 공간 연산자를 유효시간을 지원하는 불확실한 경계를 가진 시공간 연산자를 설계한다. 단, 여기에서 α , β 는 각각 TGeometry형의 시공간 객체이며, α , β 의 내부경계 : α_1 , β_1 이다. 또한 α , β 의 외부경계 : α_2 , β_2 이고 α , β 의 불확실한 경계 : $\Delta\alpha$, $\Delta\beta$ 라고 표현한다.

a.BVnearlyEqual(β) - α 의 유효시간은 β 의 유효시간에 대해서 $a.tEqual(\beta)$ 의 관계가 성립하고, α 의 내부경계(α_1)가 β 의 내부가 Equal 관계가 성립하면서 α 의 외부경계(α_2)가 β 와 적어도 한번은 교차하는 경우 참을 반환한다.

a.BVstronglyOverlap(β) - α 의 유효시간은 β 의 유효시간에 대해서 $a.tOverlap(\beta)$ 의 관계가 성립하고, α , β 의 내부가 서로 overlapping 되면서, 또한 α 와 β 의 경계에 포함되어지면 참을 반환한다

a.BVboundaryOverlap(β) - α 의 유효시간은 β 의 유효시간에 대해서 $a.tOverlap(\beta)$ 의 관계가 성립하고, α , β 의 $\Delta\alpha$, $\Delta\beta$ 가 서로 overlapping 하지만, 두 개의 α , β 의 내부가 교차하지 않고 두 개의 모든 경계에 포함되어지면 참을 반환한다.

a.BVnearlyFill(β) - α 의 유효시간은 β 의 유효시간에 대해서 $a.tEqual(\beta)$ 의 관계가 성립하고, α 의 내부 경계가 β 의 내부에 포함 되어지면서, α 의 $\Delta\alpha$ 는 β 의 모든 부분과 교차가 일어나는 경우 참을 반환한다.

a.BVcoveredByBoundary(β) - α 의 유효시간은 β 의 유효시간에 대해서 $a.tduring(\beta)$ 의 관계가 성립하고, α 의 내부가 β 의 $\Delta\beta$ 에 포함되어지면 참을 반환한다.

a.BVcoverWithBoundary(β) - α 의 유효시간은 β 의 유효시간에 대해서 $a.tduring(\beta)$ 의 관계가 성립하고 β 의 내부가 α 의 $\Delta\alpha$ 에 포함 되어지면 참을 반환한다.

또한, 일반적으로 시공간 데이터베이스에서 지원하는 기본적인 시공간 연산자를 기반으로 이 논문에서 설계한 불확실한 경계를 가지는 시공간 연산자를 분류하면 [표.1]과 같다.

표 1. 불확실한 경계를 가지는 시공간 기하 객체의 연산자 분류

분류	연산자 이름		
	시공간 객체 연산	InstanceAt	InstancePeriod
Validtime			
시공간 위상관계 연산	Disjoint	Meet	Overlap
	CoveredBy	Inside	Covers
	Contains	Equal	BVcover- WithBoundary
	BVnearlyEqual	BVstrongly- Overlap	BVboundary- Overlap
	BVnearlyFill	BVcovered- ByBoundary	
시공간 기하 연산	Boundary	Distance	Intersection
	Union	Difference	

5. 결론

기존의 시공간 데이터베이스는 단순한 공간 데이터 타입으로 실세계를 표현하였다. 이 과정에서 공간 객체의 단순화, 추상화는 실세계의 복잡, 다양한 객체를 표현하기에는 부적절하였다. 또한 이미 데이터화 된 자료들의 변환이나 이기종의 자료간의 통합이나 변환 등을 통해서 공간 데이터의 불확실성은 발생한다. 이런 문제점을 해결하기 위해 이 논문은 불확실한 경계를 가진 시공간 데이터 모델을 제시하였다. 이 모델은 기존의 시공간 데이터 모델보다 실세계의 다양한 표현이 가능하고 그에 따른 다양한 질의가 가능하다. 이 논문에서 제시한 불확실한 경계를 가진 시공간 모델의 설계는 상호 운용을 증진하기 위하여 OGC(OpenGIS Consortium)에서 제시한 2차원 기하 객체 모델을 기반으로 불확실한 경계를 가지는 공간 데이터 모델로 확장한다. 불확실한 경계를 가지는 단순한 공간 모델을 시간 영역으로 확장하여 유효시간을 갖는 불확실한 시공간 데이터 모델이 설계된다. 이 모델에서 BBGeometry는 불확실한 경계를 가지는 단순한 공간 객체이며, 시공간 기하 객체인 TGeometry는 Validtime 클래스를 상속을 받아 각각의 공간 기하에 대해 유효시간을 갖는다. 이 모델에서는 TCurve, TSurface, TBBGeometry, TPoint로 정의하였고, 각 시공간 객체의 유효시간에 대한 버전들의 집합을 DLineString, DPolygon, DBBPolygon, DBBLinestring, DBBPoint, DPoint로 하였다.

또한, 이 논문에서 제시한 불확실한 경계를 가진 시공간

객체 모델을 바탕으로 실세계의 다양하고 복잡한 객체에 대한 효율적인 질의가 가능한 시공간 연산자를 정의 및 분류하였다.

앞으로 연구되어야할 분야는 불확실하고 복잡한 기하 객체에 대한 모델링 설계와 그에 따른 수많은 위상 관계에 대한 클러스터링 기술 및 혼합된 데이터 타입간의 가능한 관계를 연구되어야 할 것이다. 또한, 불확실한 영역을 갖는 단순한 시공간 및 복잡한 시공간 객체간의 위상 관계 연산자에 대한 정의 및 설계가 요구된다.

Reference

- [1] H.Couclelis. towards an Operational Topology of Geographic Entities with Ill-defined Boundaries. Geographic Objects with Indeterminate Boundaries, GISDATA Series, vol. 3, Taylor & Francis, pp.45-55, 1996.
- [2] R.HGutting & M. Schneider. "Realm-Based Spatial Data Types : the ROSE Algebra." VLDB Journal, vol.4, pp 100-143, 1995
- [3] Martin Erwig & Markus Schneider, "Vague region " 5th Int. Symp.on Advances in Spatial Databases (SSD'97), LNCS 1262, 298-320
- [4] Eliseo Clementini & Paolino Di Felice "A spatial model for complex objects with a broad boundary supporting queries on uncertain data" DKE 37(30: 285-305 (2001)
- [5] Eliseo Clementini & Paolino Di Felice, "spatial operator" SIGMOD Rexord 29(3): 31-38 (2000)
- [6] M.F. Worboys, "A Unified Model for Spatial and Temporal Information," The Computer Journal, Vol.37, No1, pp.26-34, 1994
- [7] Eliseo Clementini & Paolino Di Felice, "Approximate topological relations," International Journal of Approximate Reasoning, 16:173--204, 1997.
- [8] OpenGIS Consortium. OpenGIS Simple Features Specification for SQL.(1998) undetermined boundaries using the Realm/ROSE"
- [10] F.J. Allen, "Maintaining Knowledge About Temporal Intervals," Communications of the ACM 26, pp832-843, 1983.
- [11] 김봉호, 남광우, 류근호, "시공간 데이터 모델". 1996년 가을 학술 발표 논문집, 제 23권, 제 2호, 한국 정보과학회 논문지, 1996년 10월
- [12] 이현아, 남광우, 류근호 "2차원 지리 객체를 위한 시공간 객체 모델 설계" 한국 정보처리학회, 제 9-권 제 1호, 2002년