

연속적인 스카이라인 질의 처리를 위한 정적 유효 영역

이종혁^o 김진호 박영배

영지대학교 컴퓨터공학과 데이터베이스연구실
 mslzh@zaigen.co.kr, solbong@mju.ac.kr, parkyb@mju.ac.kr

Static Validity Region for Continuous Skyline Queries

ZhongHe Li^o JinHo Kim YoungBae Park
 Dept. of Computer Engineering, University of Myongji

요약

이동 객체는 시간마다 위치 좌표를 변경하면서 자신의 위치를 기준으로 질의를 요청한다. 최근에 이동 객체에 대한 스카이라인 질의를 처리하기 위하여 이동 객체와 대상 객체사이의 거리인 동적 속성과 대상 객체의 정적 속성 관계를 이용한 최적화된 스카이라인 영역 결정 기법이 연구되고 있다. 그러나 영역 내의 이동 객체의 위치가 가변적이어서 유효 영역을 효율적으로 결정하지 못하는 문제점이 있다. 이 논문에서는 스카이라인 영역을 형성하는 다각형의 최대 내부 원을 이용하여 정적 유효 영역을 결정함으로써 서버에 보낼 연속적인 스카이라인 질의 발생 빈도를 감소시키는 기법을 제안한다.

1. 서론

이동 객체에 대한 스카이라인 질의는 대상 객체가 지니고 있는 정적 속성과 연속적으로 위치를 변경하는 이동 객체의 동적 속성 모두를 고려해야 한다[7]. 예를 들어 “현재 위치에서 가깝고, 숙박료가 싸며, 해변과의 거리가 가까운 호텔을 검색하라”는 질의를 할 수 있다. 이러한 질의의 특징은 스카이라인 질의 결과가 이동 객체의 위치에 따라 유효하지 않을 수 있으므로 연속적인 질의 처리를 요구한다.

연속적인 스카이라인 질의에 위한 효율적인 영역 결정 기법 [7]에서는 이동 객체의 속도와 방향과는 무관한 최적화된 스카이라인 영역(*OSR*: Optimal Skyline Region)을 미리 계산하여 질의를 처리한다. 또한 이동 객체의 위치를 중심으로 하고 가장 가까운 영역 변화까지의 거리(*MINDIST*)를 반경으로 하는 원(*Vcircle*)을 유효 영역으로 결정하여 질의 발생 빈도를 감소하는 기법을 제안하고 있다.

연속적인 스카이라인 질의에서 질의 발생 빈도를 결정하는 요소는 유효 영역의 면적이다. 즉 질의 발생 빈도는 유효 영역의 면적에 반비례하므로 [7]에서는 유효 영역의 면적이 *OSR*을 얼마나 많이 차지하느냐에 따라 효율성이 결정된다. 원(*Vcircle*)은 최초 질의가 발생한 시점의 이동 객체 위치에 따라 면적이 가변적이므로 질의 발생 빈도도 가변적이고 객체가 *OSR* 내에서 이동하는 경우에도 재 질의가 빈번하게 발생하는 문제점이 있다.

이 논문에서는 항상 불록 다각형을 형성하는 *OSR*의 특징을 이용하여 스카이라인 영역의 최대 내부 원(*IVcircle*)을 정적 유효 영역으로 결정하는 기법을 제안한다. 최대 내부 원(*IVcircle*)은 질의 발생 시점이 아닌 *OSR* 결정 시 미리 계산할 수 있고, 상대적으로 원(*Vcircle*)보다 *OSR*을 차지하는 비율이 높아 유효 영역으로 결정하는 것이 효율적임을 실험을 통하여 증명한다.

2. 관련연구

2.1 최적화된 스카이라인 영역

스카이라인 영역(*SR*: Skyline Region)이란 질의의 위치에 따라 결과에 포함될 수 있는 객체가(*Skyline dynamic*) 정적 속성에

대하여 p_i 를 지배하는 객체들보다 질의의 위치와 가까운 영역으로 대상객체 p_i 가 스카이라인 질의 결과에 포함될 수 있는 영역이다.

$$SR_i = \{q \mid dist(q, p_i) \leq dist(q, p_j), \forall p_j \prec p_i\}$$

그림 2.1에서처럼 p_i 의 스카이라인 영역 SR_i 은 p_i 의 위치좌표와 정적인 속성에 대해서 p_i 를 지배하는 모든 p_j 의 위치좌표가 생성하는 수직 이동분선들이 형성하는 영역으로 계산할 수 있다. 대상객체 p_i 의 스카이라인 영역(SR_i)의 정보는 영역을 형성하는 수직이동분선을 계산하고, 수직이동분선에 대한 p_i 의 위치를 판별하여 부등식의 형태로 표현한다[7].

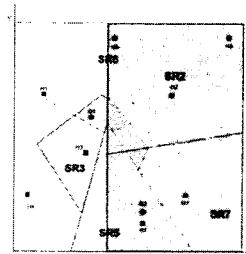


그림 2.1 대상 객체의 스카이라인 영역

최적화된 스카이라인 영역(*OSR*; Optimal Skyline Region)이란 대상 객체 p_i 가 정적 속성에 대한 지배 객체 p_j 와 피지배 객체 p_k 들보다 이동 객체(질의의 점)와 가까운 영역이다. 즉, p_i 의 *OSR*은 p_i 의 *SR*을 결정하는 선분들과 정적 속성에 대한 p_i 의 피지배 객체 p_k 들이 생성하는 수직이동분선들이 형성하는 영역이다.

$$OSR_i = \{q \mid (dist(q, p_i) \leq dist(q, p_j), \forall p_j \prec p_i) \wedge (dist(q, p_i) \leq dist(q, p_k), \forall p_k \prec p_i)\}$$

이동 객체가 *OSR*에 위치하는 경우, p_i 는 p_j 보다 이동 객체에 가깝기 때문에 항상 질의 결과에 포함되지만, p_k 는 p_i 보다 가깝지 않기 때문에 질의 결과에 포함되지 않는다[7].

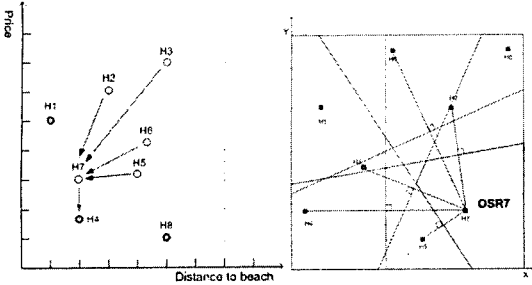


그림 2.2 최적화된 스카이라인 영역(OSR)

그림 2.2는 정적 속성에 대한 H7의 피지배 객체{H2, H3, H5, H6}과 지배 객체 H4에 대해서 H7이 생성하는 수직이등분선들에 의해 형성하는 OSR7을 나타낸다. 이 영역에 포함된 이동 객체는 H7을 항상 스카이라인 질의 결과에 포함하며, 피지배 객체{H2, H3, H5, H6}을 결과에 포함하지 않는다. OSR을 이용하여 정적 속성과 동적 속성에 대해 지배하는 객체들의 스카이라인 영역을 제외하고, 미리 계산된 나머지 스카이라인 영역과의 겹침 관계를 이용하면 검색되어질 스카이라인 영역을 최소화할 수 있다.

2.2 OSR의 동적 유효 영역(Vcircle)

기존의 이동객체의 연속적인 스카이라인 질의는 다음과 같은 과정으로 처리한다. 이동 객체는 질의 요청 시 위치좌표(x1,y1)를 서버에 전송한다. 서버에서는 이동 객체의 좌표 값이 어느 영역에 위치하는지를 판단하여 해당 영역의 정보에 의하여 영역 변까지의 가장 가까운 거리 R1을 계산하여 스카이라인 결과 집합과 함께 R1(Vcircle)을 이동 객체에 전송한다. 이동 객체는 위치가 변경될 때마다 처음 질의 시 위치에서 이동한 거리(S1)를 계산하여 결과의 유효성을 판단한다. S1 < R1이면 결과는 계속 유효하고, S1 > R1인 경우는 결과가 유효하지 않으므로 이동 위치에서 새로운 질의가 발생한다.

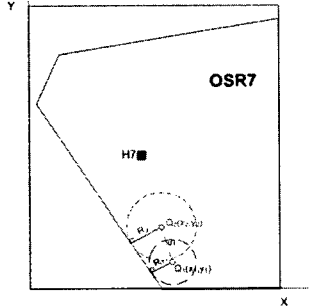


그림 2.3 Vcircle 유효 영역

3. 연속적인 스카이라인 질의 처리를 위한 정적 유효 영역

3.1 OSR의 특성 및 정적 유효 영역(IVcircle)

다각형의 최대 내부 원을 이용하여 OSR의 정적 유효 영역을 결정하는 방법은 유효 영역의 면적을 최대한 크게 하여 영역 내에서의 불필요한 질의를 줄이기 위함이다. 대상 객체의 수가 많을수록 지배관계가 복잡하기 때문에 OSR을 형성하는 다각형의 각의 개수는 상대적으로 많아진다. 또한 OSR은 대상 객체의 정적 속성에 의하여 생성된 수직이등분선으로 형성되므로 항상 볼록 다각형이다. 볼록 다각형의

특성은 각의 개수가 많을수록 원에 근접하므로 최대 내부 원(IVcircle)의 면적이 다각형에서 차지하는 비율은 다각형의 각의 개수에 정비례한다.

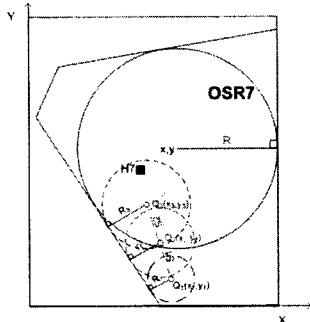


그림 3.1 최대 내부 원을 이용한 OSR의 유효 영역

다각형의 최대 내부 원을 이용한 연속적인 스카이라인 질의 처리 순서는 다음과 같다.

이동 객체는 질의 요청 시 위치좌표(x1,y1)를 서버에 전송한다. 서버에서 이동 객체의 좌표 값을 측정하여 포함된 영역을 추출하고 미리 계산된 해당 영역 정보에 의하여 좌표 값과 영역 변까지의 가장 가까운 거리 R1을 계산한다. 서버는 이동객체가 위치한 영역에 해당되는 스카이라인 결과 집합과 최대 내부 원(IVcircle)의 중심좌표(x,y), 내부원의 반경(R), 질의 위치에서 가장 가까운 변까지의 거리(R1: Vcircle)을 이동객체에 전송한다. 이동 객체는 일정시간 간격으로 이전 질의 시 위치(x1,y1)에서 이동한 거리(S1)를 계산한다. S1이 R1보다 작으면 이전 질의의 결과는 여전히 유효하고 S1이 R1보다 크면 현재 위치(x2,y2)에서 내심(x,y)까지 거리 L을 계산한다. L이 R보다 작으면 처음 질의의 결과는 여전히 유효하고 L이 R보다 크면 영역 이탈로 판단하고 서버에 질의를 요청한다.

그림 3.1에서 이동 객체가 Q1에서 Q2로 이동할 때 이동한 거리 S1이 R1보다 크지만 Q2에서 내부원의 중심까지의 거리 L이 내부원의 반경 R보다 작으므로 처음 질의의 결과는 유효하다. Q2에서 Q3로 이동할 때도 마찬가지로 이동한 거리가 영역 변까지의 최단거리(R2)보다는 크지만 최대 내부 원을 이탈하지 않았으므로 질의를 요청할 필요가 없다.

3.2 볼록 다각형의 최대 내부 원의 추출

그림 3.2-(a,b,c,d)는 볼록 다각형의 가장 큰 내부 원의 추출 순서이다.

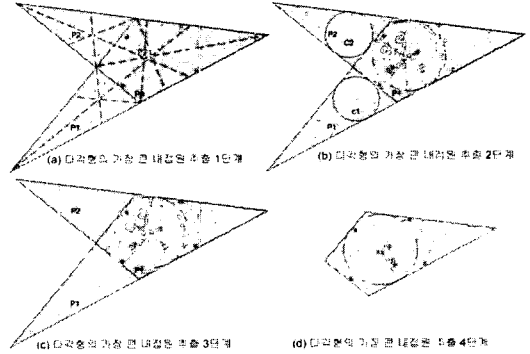


그림 3.2 볼록 다각형의 최대 내부 원 추출

1단계: 다각형의 변(a,b,c,d)을 연장하여 임의의 3개변으로 삼각형(P1, P2, P3, P4)을 만든다.
 2단계: 삼각형의 각의 이등분선을 그려 삼각형의 내심을 찾고 내접원(c1, c2, c3, c4)을 그린다.
 3단계: 영역을 형성하는 선분들의 연립부등식을 이용하여 삼각형의 내심이 다각형의 밖에 있는 내접원(c1, c2)은 제거한다.
 4단계: 내접원의 반경(R3,R4)이 내심에서 다른 변까지의 거리(S3,S4)보다 큰 내접원(C3)을 제거하고 나머지 내접원을 중에서 반경이 가장 큰 내접원을 다각형의 최대 내부 원으로 한다.

3.3 선분의 기울기를 이용한 각의 이등분선 추출

다각형의 최대 내부 원을 추출하는 1단계에서, 선분의 연장선으로 이루어진 삼각형의 각의 이등분선의 교차점을 계산하는 방법으로 내부원의 중심을 찾아냈다. OSR을 결정할 때 계산한 선분의 기울기와 곱점(꼭지점)의 좌표들을 이용하여 각의 이등분선을 찾을 수 있다.

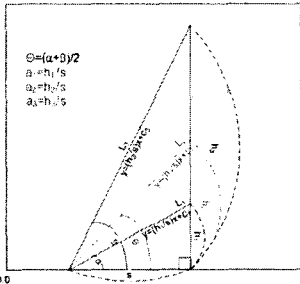


그림 3.3 선분의 기울기와 각의 이등분선

그림 3.3에서 선분 L_1 은 직선의 방정식 $y = a_1 \cdot x + c_1$ 으로 선분 L_2 는 $y = a_2 \cdot x + c_2$ 으로 표시할 수 있다. 선분 L_1 과 L_2 로 형성된 각의 이등분선인 선분 L_3 의 기울기와 y축의 절편은 다음과 같다.

$$a_2 = \tan\left(\frac{\alpha + \beta}{2}\right)$$

$$c_2 = \frac{a_1(c_2 - c_1) + c_1(a_1 - a_2) - a_2(c_3 - c_1)}{(a_1 - a_2)} \quad (a_1 \neq a_2)$$

4. 실험 및 평가

표 4.1 이동 객체의 속도 변화에 따른 질의 발생 빈도

속력 (m/s)	구분	평균 질의 발생 빈도	평균 감소	감소율
2	Vcircle	33.79	1.32	3.89%
	Vcircle + IVcircle	32.47		
5	Vcircle	76.42	4.84	6.34%
	Vcircle + IVcircle	71.58		
10	Vcircle	160.2	17.63	11.01%
	Vcircle + IVcircle	142.6		
15	Vcircle	244.2	32	13.10%
	Vcircle + IVcircle	212.2		
20	Vcircle	318.7	52.32	16.41%
	Vcircle + IVcircle	266.4		

최적화된 스카이라인 영역(OSR)의 최대 내부 원(IVcircle)을 이용한 정적 유효 영역의 효율성을 평가하기 위하여 5000M x 5000M의 영역에 균등 분포된 100개의 대상 객체를 대상으로 실험하였다. 표 4.1은 평균 속력 2, 5, 10, 15, 20m/s 인 이동 객체 1,000개의 20초 동안의 평균 질의 발생 빈도를 보여준다.

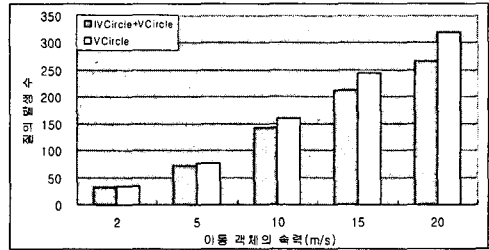


그림 4.1 이동 객체의 속력에 따른 질의 발생 빈도

그림 4.1은 속력이 2, 5, 10, 15, 20m/s 인 1,000개의 이동 객체가 20초 동안에 발생한 평균 질의 빈도를 그래프로 보여준다. 실험 결과, 연속적인 스카이라인 질의에서 Vcircle만을 고려한 경우보다 IVcircle을 같이 고려한 경우가 질의 발생 빈도는 최소 3.89%에서 최대 16.41%까지 감소하였다. IVcircle과 Vcircle을 함께 사용했을 때 이동 객체의 속력이 빠를수록 더욱 효율적임을 알 수 있다.

5. 결론

이동객체에 대한 스카이라인 질의는 다양한 위치 기반 서비스를 제공할 수 있다. 이 논문에서는 모든 대상객체에 대해서 결과 값이 변하지 않는 최적화된 스카이라인 영역(OSR)을 미리 계산하고 영역의 최대 내부 원을 이용하여 이동객체의 질의 발생 빈도를 감소시키는 기법을 제안하였다. 실험을 통해 비교 분석한 결과 평균 질의 발생 빈도를 기존의 동적 유효 영역 결정 기법보다 최대 16.41% 감소하였다.

참고문헌

- [1] Tan, K., Eng, P. Ooi, B. "Efficient Progressive Skyline Computation" In VLDB, p.301-310, 2001.
- [2] J. Zhang, M. Zhu, D. Papadias, Y. Tao, and D. L. Lee. Location-based spatial queries. In SIGMOD, p.443-454, 2003
- [3] Borzsonyi, S. Kossmann, D., Stocker, K. "The Skyline Operator" In ICDE, p.421-430, 2001.
- [4] Zheng B., Lee, D. "Semantic Caching in Location-Dependent Query Processing" SSTD, p.97-116, 2001
- [5] D. Kossmann, F. Ramsak, S. Rost, "Shooting Stars in the Sky: an Online Algorithm for Skyline Queries." In VLDB, p.275-286, 2002
- [6] Papadias, D., Tao, Y., Fu, G., Seeger, B. "An Optimal and Progressive Algorithm for Skyline Queries." In SIGMOD, 2003.
- [7] 나경석, 박영배 "연속적인 스카이라인 질의를 위한 효율적인 영역 결정기법" 명지대학교 석사학위논문. 2006년2월.