

텔레매틱스를 위한 불확실성 기반의 이동체 색인

진희규[○] 김동현* 임덕성** 조대수* 홍봉희[○]

부산대학교 컴퓨터공학과[○], 동서대학교 소프트웨어전문대학원*, 영진전문대학교 컴퓨터정보기술계열**, 한국전자통신연구원 텔레매틱스연구단 LBS연구팀*

{heegyu, bhong}@pusan.ac.kr[○], pusrover@dongseo.ac.kr*, junsung@yjc.ac.kr**, junest@etri.re.kr*

Indexing of Moving Objects Based on Uncertainty for Telematics

Heegyu Jin[○] Donghyun Kim*, Duksung Lim**, Daesoo Cho*, Bonghee Hong[○]

Dept. of Computer Engineering, Pusan National University[○]

Graduate School of Software, Dongseo University*

Division of Computer Information Technology, Yeungjin College**

LBS Research Team, Telematics Research Division, ETRI*

요약

속도와 방향이 바뀔 때 마다 이동체의 위치를 보고하는 TPR-tree는 이동체의 위치를 예측하는 오차가 적다. 그러나 긴 시간 간격으로 이동체의 위치를 보고하면 위치 예측의 불확실성이 높아져서 위치 예측의 오차 값이 증가한다. 불확실성이 높은 이동체를 TPR-tree에 적용할 때 이동체의 위치 정보를 갱신하기 위한 색인 검색 비용이 증가하고, 질의 결과의 정확도가 낮아지는 문제가 발생한다.

이 논문에서는 긴 시간 간격으로 이동체 위치를 보고할 때 발생하는 이동체 위치의 불확실성을 고려하기 위해서 불확실성 영역(uncertainty region)을 이용한 확장 TPR-tree를 제시한다. 불확실성이 높은 이동체의 위치 데이터를 처리하기 위해서 이동체의 이동 가능한 영역을 위치 예측의 오차 값을 이용하여 계산한 불확실성 영역을 설정하고, 검색을 위하여 노드의 BR을 계산할 때 불확실성 영역을 이용하여 BR을 확장한다.

1. 서론

최근 GPS기술을 이용한 위치 기반 서비스(LBS: Location Based Service)의 요구가 증가하고 있다. LBS 응용분야로 친구 찾기, 네비게이션 시스템 등이 있다. 이러한 시스템 응용을 위하여 이동체의 위치 정보를 관리하고 현재 및 미래 위치 질의를 효과적으로 처리하는 방법이 필요하다. 예를 들어 “현재부터 10분 이내에 부산대 앞을 지나갈 수 있는 차량은?”과 같은 형태의 질의를 처리할 수 있어야 한다.

이동체의 위치보고는 일정한 주기를 가지고 이루어진다. 이동체의 위치보고가 긴 시간 간격을 주기로 이루어질 때 이동체는 위치 예측의 불확실성이 증가하는 특징을 가진다. 이러한 경우에 TPR-tree에서는 질의 결과의 정확도가 낮아지고, 이동체 위치정보를 갱신하기 위해서 전체 색인을 검색하는 문제점이 발생한다.

대표적인 이동체 현재 및 미래 위치에 대한 색인 기법으로 TPR-tree 계열의 색인[1,2,3,4]들이 있다. TPR-tree 계열의 색인에서는 이동체의 위치를 시간에 대한 선형함수로 예측한다. 그러나 이동체가 특정한 임계영역내에 있어야 하는 전제 조건이 있기 때문에 위치 예측이 부정확한 이동체를 색인에 적용할 수 있는 방법을 고려하지 않았다. 이동체 위치의 불확실성에 관한 연구들은 보고되지 않은 위치를 추정하는 방법[5,6]과 속도와 방향으로 미래 위치를 예측하는 방법[7]이 있다. 그러나 위의 불확실성 처리방법의 경우에도 이동체 이동 영역에 제한을 두지 않고 처리할 수 있는 방법은 제시되지 않고 있다.

이 논문에서는 이동체 위치 예측의 불확실성을 반영한 불확실성 기반의 확장 TPR-tree를 제시한다. 이동체 위치 예측의 불확실성을 처리하기 위해서 이동체의 불확실성 영역을 설정한다. 불확실성 영역은 현재의 속도와 방향에 위치 예측의 오차를

반영한 이동체의 이동 가능한 영역으로써, TPR-tree의 BR을 계산할 때 불확실성 영역을 이용하여 BR을 확장하는 방법을 사용한다. 그리고 이동체 정보 저장 테이블을 사용해서 이동체 위치정보의 갱신을 가능하게 한다. 이렇게 함으로써 불확실성이 높은 이동체 위치 데이터에 대한 질의 결과의 정확도를 향상시킬 수 있다.

이 논문의 구성은 다음과 같다. 먼저 2장에서는 관련 연구를 소개하고 3장에서는 대상 환경 및 문제 정의를 기술한다. 4장에서는 이동체의 불확실성 처리 방법과 색인 적용 방법을 설명하고 5장에서는 결론 및 향후 연구를 기술한다.

2. 관련 연구

이동체의 현재 및 미래 위치에 대한 질의를 지원하는 기존 연구는 이동체의 위치를 시간에 대한 선형함수로 표현한 기법이 활발히 연구되고 있다. TPR-tree 계열의 색인[1,2,3,4]에서는 이동체의 위치를 시간에 대한 선형함수로 표현하여 현재 및 미래 위치에 대한 질의를 가능하게 한 구조이다. 이러한 색인들은 이동체의 속도와 방향을 시간에 대한 함수의 매개변수로 사용하여 R*-tree 구조로 색인화하였다. 그리고 이동체의 속도와 방향이 특정 임계값 이상 변화하는 경우에만 갱신함으로써 갱신 횟수를 줄이고, 이동체의 위치가 선형함수로 예측한 위치에서 임계 영역내에 있음을 보장하고 있다. 그러나 TPR-tree 계열의 색인 기법들은 이동체의 위치 보고가 속도와 방향의 임계값을 벗어나서 이루어지는 경우를 고려하지 않았다.

시간에 따라 연속적으로 변하는 이동체의 위치를 모두 데이터베이스에 저장할 수가 없기 때문에 이산적인 위치 데이터를 저장하게 된다. 이러한 이산적인 이동체의 위치를 토대로 하여 보고되지 않은 이동체의 위치를 추정하는 연구들이 있다. 선형 보간법으로 보고되지 않은 이동체의 과거 위치를 처리하는 방

법[5]과 이동체의 이동간 변화에 대해 확률과 퍼지이론을 이용해서 처리하는 방법[6]이 있다. 이동체의 속도와 방향으로 이동체의 향후 위치를 알아내는 방법으로 [7]은 MOST 데이터모델을 소개하고 있다. 이 방법 역시 이동체의 속도와 방향이 이동체의 이동경계 내부에 존재한다는 가정으로 이동체의 위치를 예측하고 있다.

3. 대상 환경 및 문제 정의

3.1. 대상 환경

이동체의 위치보고는 무선 이동 통신으로 이루어지기 때문에 통신 비용문제가 발생한다. 이동체의 위치보고는 속도와 방향이 미리 설정한 범위를 벗어나는 경우와 일정 시간 간격으로 보고하는 두 가지 방법이 있다. 전자의 경우는 일반 도로에 적용할 경우 빈번한 변화로 인한 통신비용이 많이 발생하게 된다. 그러므로 이 논문에서는 적절한 통신비용을 유지하기 위해서 30분 이상의 긴 시간 간격으로 이동체의 위치 보고가 이루어지는 환경을 대상으로 한다. 그리고 이동체는 이동체의 식별자, 이동속도와 방향 정보를 보고한다.

3.2. 문제 정의

이동체는 현재의 이동 속도와 방향을 보고하고 위치를 시간에 대한 선형함수로 표현한다. 그러나 이 정보는 시간이 지남에 따라서 변화하는 이동체의 속도와 방향정보를 반영하지 못하기 때문에 이동체의 위치 예측에 있어서 큰 오차가 발생한다. 이러한 오차는 TPR tree에서 다음과 같은 문제점을 발생시킨다.

첫째, 이동체의 위치정보 갱신시에 이동체의 위치를 찾기 위한 색인 갱신 비용이 증가하는 문제점이 있다. 이동체의 위치를 찾기 위한 검색영역 설정이 어렵고, 최악의 경우에 전체 색인을 검색해야 하는 문제가 있다. 이는 이동체의 갱신 비용을 증가시키는 중요한 요인이 된다. 그림1은 t_0 시간에 이동체 a_1, b_1 의 위치와 BR을 나타낸다. t_1 시간에 a_2, b_2 의 위치로 갱신된다고 했을 때, 이동체 a의 경우는 a_2 의 위치가 t_1 시간의 확대된 BR 영역 내부에 해당되어서 갱신이 가능하지만, 이동체 b의 경우는 BR 영역 바깥에 있어서 갱신을 할 수 없다. 그러므로 이동체 b의 갱신을 위해서 이동체 b를 검색하는 비용이 매우 커지게 된다.

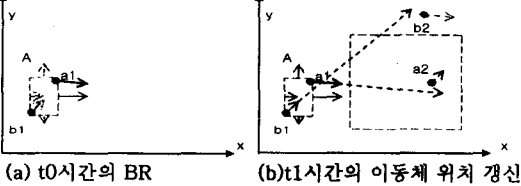


그림 1 TPR-tree의 update 문제점

둘째, 이동체의 검색의 정확도가 낮아진다. 이는 검색 영역에 해당되는 객체의 정보를 검색하지 못하는 경우가 많이 발생하는 것이다.



그림 2 TPR-tree의 검색 문제점

그림2에서 이동체 a가 t_1 시간에 a_1 , t_2 시간에 a_2 를 보고하면 t_3 시간에 a의 위치는 시간에 대한 선형함수로 결과값으로 a_2 의 위치에 있는 것으로 예측한다. 그러나 이동체 a는 t_3 시간에 X로 표시된 모든 위치에 존재할 수 있지만, 색인에서는 a_2 가 t_3 시간에 a_2 의 위치에 있는 것으로만 예측한다. t_3 시간에 점선영역에 대한 영역검색을 수행하면 이동체의 a를 검색할 가능성이 낮아

지게 된다.

4. 이동체 위치의 불확실성 처리

3장에서 언급한 문제를 해결하기 위해 이동체 위치 예측의 오차를 고려한 불확실성 영역 설정과 이동체 정보 저장 테이블의 구성을 제안한다.

제안 방법의 설명을 위해서 다음과 같은 용어를 정의한다.

[정의1] 불확실성(uncertainty): 이동체의 위치 예측의 불확실성으로써, 시간에 대한 선형함수로 예측한 위치와 실제 이동체의 위치의 차이이다.

[정의2] 불확실성 영역(uncertainty region): 이동체 움직임의 불확실성을 반영한 영역으로서, 이동체의 이동 가능한 영역을 시간에 대한 선형함수에 이동체의 위치 예측 오차를 반영한 영역이다.

4.1. 불확실성 영역 설정

문제정의에서 언급했듯이, TPR-tree에서 이동체의 현재 이동 방향과 속도만으로 위치를 예측하는 것은 오차가 크게 발생하기 때문에, 이 논문에서는 이동체의 이동이 가능한 영역을 이동체 위치 예측의 오차를 고려하여 설정하는 방법을 제안한다. 이동체 위치 예측의 오차는 TPR-tree에서 갱신시에 이동체의 선형함수로 예측된 위치와 실제 보고된 위치 사이의 차이를 말한다. (그림3 참조)

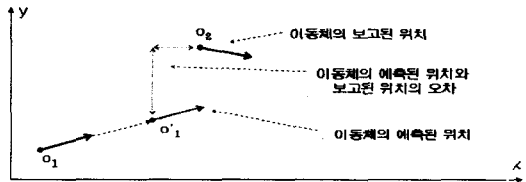
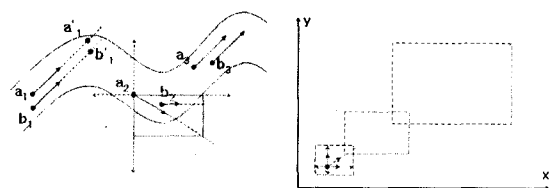


그림 3 위치 예측의 오차

불확실성 영역은 현재 이동체의 이동 속도와 방향으로 만들어진 선형함수에 공간의 각 축으로 이동체의 위치 예측의 오차 범위를 반영해서 설정한다. 공간의 각 축으로 이동체 위치 예측 오차를 현재 이동체가 이동할 수 있는 영역으로 설정함으로써 이동체 위치의 불확실성을 반영한다.

그림 4의 (a)와 같이 이동체 a, b가 t_1 시간의 위치가 a_1, b_1 이고 t_2 시간에 a_2, b_2 로 이동하는 경우에 선형함수로 예측한 이동체의 위치는 a_1', b_1' 가 된다. 불확실성 영역은 a_2, b_2 의 BR을 설정할 때 이동체의 예측한 위치와 실제 보고된 위치 사이의 오차를 반영해서 확대한 영역이다. t_3 시간에 a_3, b_3 의 위치를 보고했을 때, TPR-tree의 BR은 사각형의 영역이 되지만 불확실성 영역은 회색영역과 같이 확대되어서 a_3, b_3 를 검색할 수 있다.



(a) 불확실성 영역 설정 (b)시간에 따른 변화

그림 4 불확실성 영역

그림 5의 (b)의 회색영역으로 설정된 불확실성 영역은 시간에 따라서 선형적으로 증가해서 영역이 더 크게 확대되는 모습을 보이고 있다. 이는 시간이 지날수록 이동체 위치의 불확실성이 증가하는 점을 반영하고 있다.

이동체의 움직임을 살펴보면 먼 과거의 움직임보다는 최근의 움직임에 더 많은 영향을 받음을 알 수 있다. 이동체의 이동지역이 시내지역에서 시외지역으로 이동하고 다시 시내지역으로 이동을 한다면, 이동체는 각각 해당 지역이 있는 동안의 움직임

에 더 큰 영향을 받게 된다. 그러므로 이동체 위치 예측의 오차를 반영함에 있어서 최근의 오차에 더 큰 가중치를 둔다. 따라서 이동체의 불확실성 영역은 현재의 속도와 방향에 공간의 각 축으로 과거 위치 예측의 오차를 시간에 따라서 가중치를 반영해서 설정한다.

불확실성 영역을 설정하기 위해서 표1과 같은 표기가 사용된다.

표 1 표기 정리

- t_{upd} : 이동체의 갱신 시간
- t_{ref} : 색인의 참조 시간 (미래의 임의의 특정 시간)
- $t_i = t_u - t_{ref}$
- $V = (V_x, V_y)$: 이동체의 속도(벡터)
- 이동체의 i 번째 보고 위치 (x_i, y_i) 속도 (V_{x_i}, V_{y_i})
- $Dv_i = (Dx_i, Dy_i)$: 이동체의 $i-1$ 번째 보고로 예측된 위치와 i 번째 보고된 위치와의 차이 (벡터)
- $Dv(i)$: i 번째까지 보고한 이동체의 예측된 위치와 보고된 위치의 차이의 누적
- $Dv(i)$: 이동체의 현재 속도에 위치 예측 오차를 반영한 속도
- F : 과거 정보의 반영 가중치 ($0 < F < 1$)

공간 축을 x, y 축이라고 할 때 이동체 위치 예측의 오차를 구하기 위해서 수식 1,2를 이용한다.

$$Dx_i = |x_i - (x_{i-1} + V_{x_i}(t_i - t_{i-1}))| \quad (\text{수식 1})$$

$$Dy_i = |y_i - (y_{i-1} + V_{y_i}(t_i - t_{i-1}))| \quad (\text{수식 2})$$

이동체 위치 예측의 오차 Dx_i, Dy_i 는 수식 3과 같은 Dv_i 의 벡터성분이 된다.

$$Dv_i = (Dx_i, Dy_i) \quad (\text{수식 3})$$

최근 보고의 오차에 가중치를 준 이동체의 위치 예측 오차의 누적을 구하기 위해서 수식 4를 이용한다.

$$Dx(i) = Dx_i * F + Dx_{i-1} * F^2 + Dx_{i-2} * F^3 + \dots + Dx_1 * F^n$$

$$= (Dx_i + Dx_{(i-1)}) * F \quad (\text{수식 4})$$

이동체의 위치 예측의 오차를 반영한 속도는 공간 축 단위로 최소, 최대의 속도를 계산할 수 있다. (수식 5,6)

$$Vx^t = \min(-Dx(i), V_{x_i} - Dx(i)) \quad (\text{수식 5})$$

$$Vx^b = \max(Dx(i), V_{x_i} + Dx(i)) \quad (\text{수식 6})$$

이동체의 위치 예측의 오차를 반영한 속도를 바탕으로 수식 7, 8의 방법으로 불확실성 영역을 계산한다.

$$Ux^t = x_i(t) + Vx^t * (t) \quad (\text{수식 7})$$

$$Ux^b = x_i(t) + Vx^b * (t) \quad (\text{수식 8})$$

공간의 다른 축에 대해서도 같은 방법으로 Uy^t, Uy^b 를 설정할 수 있다. 그러므로 불확실성 영역은 (Ux^t, Ux^b, Uy^t, Uy^b) 로 표현이 된다.

4.2. 이동체 정보 저장 테이블

이동체 정보 저장 테이블은 불확실성 영역 설정을 위한 정보를 저장한 테이블이다. 테이블에는 이동체의 식별자, 시간에 대한 선형함수와 이동체의 위치 예측 오차의 누적($Dx(i), Dy(i)$)을 저장한다. 불확실성 영역을 계산할 때 이동체의 위치 예측 오차의 누적이 필요한데, 이 정보를 테이블에 저장해서 불확실성 영역을 계산한다. 그리고 테이블에 저장된 이동체의 선형함수를 이용해서 이동체의 위치 정보를 갱신한다. 새로운 보고가 들어왔을 때 테이블에 있는 선형함수를 이용하면, 이동체의 위치 정보가 색인상에 어느 노드에 있는가를 알 수 있다. 이 위치 정보로 이동체의 위치 정보를 색인에 갱신할 수 있다. 그림 3에서 이동체 o_1 가 t_2 시간에 o_1 위치에서 o_2 위치로 이동한 정보를 보고하면 테이블에서 선형함수로 t_2 시간에 o_1 이 위치해 있을 $o_{(t_2)}$ 의 위치를 갖는다. 색인에서 이동체 o 는 이동체의 실제 위치와 관계 없이 $o_{(t)}$ 위치에 존재하므로 이 위치 정보를 이용해서 이동체 o

의 정보를 갱신한다.

4.3. TPR-tree 적용

TPR-tree에서 불확실성 영역을 이용해서 TPBR(time parameterized bounding rectangle)을 설정함으로써 TPR-tree의 구조를 변경하지 않고 일정한 시간 간격으로 보고하는 불확실성이 높은 이동체를 처리할 수 있다. 공간의 한 축을 x 라고 했을 때 x 축에 대한 TPBR은 (x^t, x^b, v^t, v^b) 와 같이 표현된다.

$$x^t = x^t(t_i) = \min(o_i.Ux^t(t_i))$$

$$x^b = x^b(t_i) = \max(o_i.Ux^b(t_i))$$

$$v^t = v^t(t_i) = \min(o_i.Vx^t)$$

$$v^b = v^b(t_i) = \max(o_i.Vx^b)$$

t_{upd} 시간에 이동체의 갱신이 이루어질 때 TPBR의 경계영역을 다시 설정하는 수식은 다음과 같다.

$$x^t = \min(o_i.Ux^t(t_{upd})) - Vx^t(t_i) \quad (\text{수식 9})$$

$$x^b = \min(o_i.Ux^b(t_{upd})) - Vx^b(t_i) \quad (\text{수식 10})$$

공간의 다른 한 축에 대해서도 같은 방법으로 설정할 수 있다. 위에서 언급한 바와 같이 TPR-tree에서 TPBR의 설정 방법만 변경하고 삽입,삭제는 TPR-tree의 방법을 그대로 사용함으로써 TPR-tree의 구조변화 없이 이동체의 불확실성을 처리할 수 있다.

5. 결론

이동체의 위치보고가 긴 시간 간격으로 이루어지는 경우에 TPR-tree에서는 질의 결과의 정확도가 낮아지고, 이동체 위치 정보를 갱신하기 위해서 전체 색인을 검색하는 문제점이 발생한다.

이 논문에서는 일정한 시간 간격으로 보고되는 이동체에 대해 대표적인 현재 및 미래 위치 색인인 TPR-tree에 적용할 수 있는 방법을 제시하였다. 이동체 위치의 불확실성을 반영한 불확실성 영역을 설정하고, 이것을 색인의 BR에 반영하는 방법으로 TPR-tree에 적용하였다. 이렇게 함으로써 불확실성이 높은 이동체에 대한 색인 적용이 가능해졌고, 질의의 정확도를 향상시킬 수 있다. 또한, 이동체 정보 저장 테이블을 사용함으로써 TPR-tree의 위치정보 갱신을 가능하게 하였다.

향후 연구로 이 논문에서 제시한 불확실성 영역 설정에 대한 검증과 불확실성 영역을 적용한 색인에 대한 성능 향상을 위한 연구가 있어야 한다. 그리고 불확실성 영역을 적용한 질의처리 방안에 대한 연구가 필요하다.

6. 참고 문헌

- [1] S.Saltenis, C.S.Jensen, S.T.Leutenegger, M.A.Lopez, "Indexing the Positions of Continuously Moving Objects", In Proc. ACM SIGMOD on Management of data, p331-342, 2000.
- [2] Y.Tao, D.Papadias, J.Sun, "The tpr*-tree: an optimized spatiotemporal access method for predictive queries", VLDE, p178-193, 2003.
- [3] Cecilia Magdalena Procopiuc, Pankaj K. Agarwal, Sarel Har-Peled, STAR-Tree: An Efficient Self-Adjusting Index for Moving Objects", ALENEX p178-193, 2002
- [4] Simonas Saltenis, C.S. Jensen: Indexing of Moving Objects for Location-Based Services. ICDE, p463-472, 2002
- [5] D.Pfoser, C.S. Jensen: Capturing the Uncertainty of Moving-Object Representations. SSD, p111-132, 1999
- [6] D.Pfoser, Nectaria Tryfona: Capturing Fuzziness and Uncertainty of Spatiotemporal Objects. ADBIS, p:12-126, 2001
- [7] A.P.Sistla, O.Wolfson, S.Chamberlain, S.Dao, "Querying the Uncertain Position of Moving Objects", Temporal Databases, 310-337, 1997