

# 도로 네트워크에서 이동 객체의 과거 궤적 분석을 통한 미래 경로 예측<sup>†</sup>

## Path Prediction of Moving Objects on Road Networks through Analyzing Past Trajectories

김종대\* / Jong-Dae Kim

김상욱\*\* / Sang-Wook Kim

원정임\*\*\* / Jung-Im Won

### 요약

본 논문에서는 도로 네트워크에서 이동하는 객체들의 미래 경로를 예측하는 방안에 대하여 다룬다. 기존의 대부분 미래 예측 기법들은 유클리드 공간에서 이동하는 객체들을 대상으로 한다. 그러나 텔레매틱스 등 다양한 응용에서 객체들은 도로 네트워크 상에서 이동하는 경우가 많으므로 이를 위한 미래 예측 방법이 요구된다. 본 연구에서는 질의 객체의 현재까지의 이동 궤적과 유사한 경향을 가지는 과거 궤적들을 분석함으로써 이 객체의 미래 경로를 예측하는 기법을 제안한다. 우선, 도로 네트워크의 특성을 반영하여 궤적들의 유사도를 측정하는 새로운 함수를 제안한다. 이 함수를 이용하여 주어진 이동 객체의 미래 경로를 다음과 같이 예측한다. 먼저, 이동 객체 데이터베이스 내의 과거 궤적들을 대상으로 주어진 질의 궤적과 유사한 부분 궤적을 갖는 후보 궤적들을 검색한다. 그 다음, 검색된 후보 궤적들의 현재 위치 이후부터 목적지까지의 이동 경로를 분석함으로써 객체의 미래 이동 경로를 예측한다. 작은 차이를 갖는 이동 경로들을 같은 그룹으로 간주함으로써 경로 예측의 정확성을 높이는 방안을 제안한다.

### Abstract

This paper addresses techniques for predicting a future path of an object moving on a road network. Most prior methods for future prediction mainly focus their attention on objects moving in Euclidean space. A variety of applications such as telematics, however, deal with objects that move only over road networks in most cases, thereby requiring an effective method of future prediction of moving objects on road networks. In this paper, we propose a novel method for predicting a future path of an object by analyzing past trajectories whose changing pattern is similar to that of a current trajectory of a query object. We devise a new function that measures a similarity between trajectories by reflecting the characteristics of road networks. By using this function, we predict a future path of a given moving object as

† 본 논문은 제주대학교를 통한 정보통신부 및 정보통신진흥원의 대학 IT 연구센터 지원사업의 지원을 받았습니다. (IITA-2005-C1090-0502-0009)

■ 논문접수 : 2006.8.21      ■ 심사완료 : 2006.9.22

\* 한양대학교 정보통신학부 연구원(jdkim@agape.hanyang.ac.kr)

\*\* 교신저자 한양대학교 정보통신학부 교수(wook@hanyang.ac.kr)

\*\*\* 한양대학교 정보통신학부 연구교수(jiwon@hanyang.ac.kr)

follows: First, we search for candidate trajectories that contain subtrajectories similar to a given query trajectory by accessing past trajectories stored in moving object databases. Then, we predict a future path of a query object by analyzing the moving paths along with a current position to a destination of candidate trajectories thus retrieved. Also, we suggest a method that improves the accuracy of path prediction by regarding moving paths that have just small differences as the same group.

**주요어 :** 이동 객체 데이터베이스, 도로 네트워크, 미래 예측, 유사 검색

**Keyword :** moving object databases, road networks, future prediction, similarity search

## 1. 서론

최근, 휴대폰 및 PDA 등의 휴대용 이동 단말기의 확산과 통신 기술의 발달로 인하여 이동하는 객체의 시간 흐름에 따른 공간적 위치 정보를 활용하는 다양한 서비스들이 제공되고 있다[1]. 시공간 데이터(spatio-temporal data)란 공간(x, y)과 시간(t)의 특성을 함께 가지고 있는 3차원 공간(x, y, t)상의 데이터를 의미한다. 객체들의 이동 경로는 3차원의 시공간상에서 라인 세그먼트들의 집합으로 표현할 수 있으며, 이를 궤적(trajecory)이라 부른다[1, 2, 3, 4, 5]. 이동 객체의 궤적 정보를 잘 분석하면, 도로 정보, 차량 정보 등과 연계하여 단순한 위치 추적 뿐만 아니라 이동 객체와 관련된 사람의 행동 패턴 등을 파악할 수 있으며, 이를 다양한 비즈니스에 활용할 수 있다.

이동 객체에 대한 사용자 질의는 크게 이동 객체의 과거에 움직인 위치를 검색하는 과거 이력 질의(historical query)와 미래 위치를 예측하는 미래 예측 질의(future query)로 구분된다. 과거 이력 질의는 영역 질의(range query), 궤적 질의(trajecory query), 복합 질의(complex query)로 구분된다[4]. 영역 질의는 주어진 질의 영역 내에 존재하는 이동 객체를 검색하는 질의이며, 궤적 질의는 주어진 시간 간격 동안에 이동 객체가 움직인 경로를 검색하는 질의

이다. 복합 질의는 영역 질의와 궤적 질의를 결합한 형태로서 주어진 시간 간격 동안에 특정 영역에 있었던 이동 객체의 궤적을 검색하는 질의이다. 미래 예측 질의의 경우 이동 객체의 현재 위치, 이동 속도, 이동 방향 등을 이용하여 미래 위치를 예측하는 것이 일반적이다[1, 6, 7]. 본 논문에서는 이들 중 미래 예측 질의를 연구의 대상으로 한다.

미래 예측 질의를 위한 인덱스 구조로는 VCI-트리[6], TPR-트리[1], TPR\*-트리[7] 등이 제안된 바 있다. 이들 연구에서는 이동 객체의 움직임을 유클리디안(Euclidean) 공간 상에서 이동 가능한 모든 2차원의 공간 좌표(x, y)를 대상으로 하며, 이동 객체의 이동 방향과 속도가 일정 시간 내에는 유지된다는 가정 하에 이동 객체의 미래 위치를 예측한다. 이 중 가장 널리 이용되고 있는 TPR\*-트리는 이동 객체들의 현재 위치들에 대한 MBR(minimum bound rectangle)과 이동 방향 및 속도를 모두 포함하는 CBR(conservative bounding rectangle)을 R\*-트리[8] 내에 저장함으로써 이동 객체의 미래 위치를 빠르게 검색할 수 있도록 한다.

그러나 실제 응용에서 대부분의 이동 객체들은 도로 네트워크(road network) 공간 상에 존재하며, 이동 객체의 공간 정보를 파악하기 위하여는 1차원의 도로 정보가 사용자에게 보다 직관적이고, 유용한 정보를 제시

할 수 있다[13, 14, 15]. 또한, 도로 네트워크 상에서의 이동 객체의 이동 방향과 속도는 도로 네트워크의 사정에 의하여 빈번하게 달라질 수 있다[16, 17]. 따라서 이동 객체의 현재 위치, 이동 방향, 속도를 이용하여 시간에 대하여 선형적으로 변화한다는 가정 하에 이동 객체의 위치를 예측하는 기존의 방법을 도로 네트워크 공간 상에 그대로 적용할 수 없다.

본 연구에서는 도로 네트워크 공간 상에서 이동하는 객체들을 위한 미래 경로 예측 방법을 제안한다. 도로 네트워크 공간에 존재하는 이동 객체를 관리하는 데이터베이스 기술로는 트래킹, 유사 궤적 검색, 데이터 생성, 도로 네트워크 인덱싱 등이 제안된 바가 있다[12, 17, 18, 19, 20, 21]. 그러나 도로 네트워크 공간 상에서 이동 객체의 미래 경로 예측 방법은 본 저자가 아는 한 아직까지 제안된 바가 없다. 제안된 기법에서의 미래 경로 예측은 다음과 같이 수행된다. 먼저 현재 이동 중인 객체의 출발지부터 현재 위치까지의 궤적이 질의로 주어지면, 이동 객체 데이터베이스 내의 과거 궤적들을 대상으로 주어진 질의 궤적과 유사한 부분 궤적을 갖는 후보 궤적들을 검색한다. 그 다음, 검색된 후보 궤적들의 현재 위치 이후부터 목적지까지의 이동 경로를 분석하여, 빈도를 기반으로 확률을 부여함으로써 객체의 미래 이동 경로로 예측한다. 이러한 예측 방법을 이용하여 다음과 같은 형태의 질의의 처리가 가능하다. “원자력 병원에서 출발하여 동부간

선로를 지나 현재 용비교에 도착한 사용자 A는 이후 어떤 경로를 이용하여 목적지인 한양대로 이동할 것인가?” 또한, 본 연구에서는 경로 예측의 정확성을 높이기 위하여 작은 차이를 갖는 이동 경로들을 그룹핑하여 유사 경로로 간주함으로써 이를 예측 과정에 활용하는 방안을 제안한다.

본 논문의 구성은 다음과 같다. 제 2장에서는 본 논문에서 연구 동기와 해결하고자 하는 문제를 정의하고, 제 3장에서는 도로 네트워크 상에서 이동 객체의 과거 궤적에 대한 유사 궤적 검색과 이를 이용한 미래 경로 예측 기법을 제안한다. 제 4장에서는 본 논문을 요약하고, 결론을 내린다.

## 2. 연구 동기 및 문제 정의

본 연구에서는 도로 네트워크 상에서 이동하는 객체의 궤적  $T_i$ 를 (userId, mold, tId, <rseg<sub>1</sub>, rseg<sub>2</sub>, ..., rseg<sub>k</sub>>)로 구성한다. 여기서, userId는 사용자의 식별자이다. 동일 목적지로 움직이는 이동 객체라도 사용자의 운전 성향에 따라 경로 선택이 달라질 수 있으므로, 본 연구에서는 사용자별로 궤적을 분리하여 관리한다. mold는 이동 객체의 식별자이며, tId는 궤적 식별자이다. 또한, rseg<sub>j</sub>( $1 \leq j \leq k$ )는 궤적  $T_i$ 를 구성하는 도로 세그먼트이며, (rsId<sub>j</sub>, rsLen<sub>j</sub>)로 구성된다. rsId<sub>j</sub>는 도로 세그먼트의 식별자이며, rsLen<sub>j</sub>은 도로 세그먼트의 길이이다. 본 연구에서는 교차로와 교차로 사이의 도로들의 시퀀스를 하나의

<표 1> 도로 네트워크 상에서 생성된 궤적의 예

출발지	이동경로	목적지	출발시간
원자력병원	동부간선로/용비교/성동교/한양대	한양대	9:00
원자력병원	동부간선로/서울외대/성동교/한양대	한양대	10:00
서울대	한강대교/강변북로/용비교/한양대	한양대	8:20
강남역	한남대교/올림픽대로/상도터널/신림역	신림역	7:30
⋮	⋮	⋮	⋮

도로 세그먼트로 간주한다[9]. 이는 사용자의 경로 선택이 교차로 지점에서만 발생하기 때문이다.

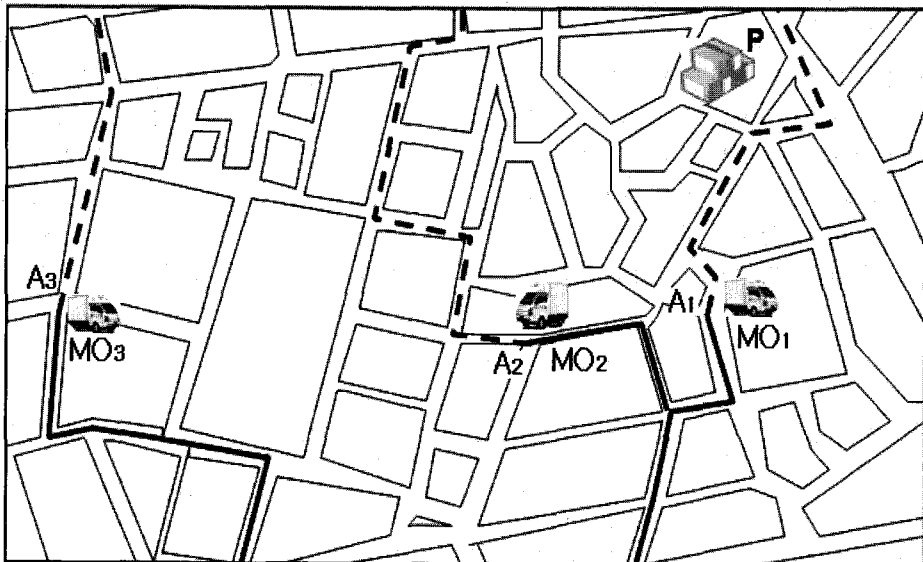
도로 네트워크 상에서 시간 흐름에 따라 이동 객체가 지나간 도로 세그먼트들은 이동 객체 데이터베이스에 궤적의 형태로 축적되며, 임의의 출발지에서 목적지까지 움직인 이동 경로들은 대부분 사용자의 운전 성향을 반영하여 반복적인 패턴이나 유사 패턴을 갖는 경향이 있다. 도로 네트워크 상에서 임의의 사용자 A에 의해서 생성된 궤적의 예를 표 1에 보인다.

이와 같은 상황에서 이동 객체들의 미래 경로를 예측하기 위한 다음과 같은 질의 형태가 가능하다. “원자력 병원에서 출발하여 동부간선로를 지나 현재 용비교에 도착한 사용자 A는 이후 어떤 경로를 이용하여 목적지인 한양대로 이동할 것인가?” 본 연구에서는 대규모 이동 객체 데이터베이스를 대상으로 하여 위와 같은 미래 경로 예측 질의를 지원하기 위한 방법을 제안한다. 여기서 입력 값으로 운전자와 출발지, 이동 객체의 현

재 위치까지의 이동 경로, 목적지가 사용자에 의하여 주어진다 가정한다.

실제 응용 예로 택배 회사의 물류 수송을 들 수 있다. 택배 회사의 택배 수송 차량은 현재 위치에서 출발하여 목적지인 택배물이 있는 장소로 이동하며, 각 차량 운전자는 자신이 선호하는 경로를 따라 목적지까지 이동한다. 또한, 택배물은 택배 수송 차량이 이동하는 중에도 빈번하게 발생할 수 있다. 이런 환경에서 택배 차량의 이동 경로를 미리 예측하여 새로운 택배물이 있는 장소로 이동할 가능성이 높은 차량을 검색하여 배차할 수 있다면, 택배 차량의 이동 경로를 최소화하여 효율적인 물류 수송 시스템을 구축할 수 있을 것이다.

그림 1은 이와 관련된 예를 나타낸 것이다. A<sub>1</sub>, A<sub>2</sub>, A<sub>3</sub>는 현재 도로 네트워크 상에서 이동 중인 택배 차량 MO<sub>1</sub>, MO<sub>2</sub>, MO<sub>3</sub>의 현재 위치를 나타내며, 검은 실선은 각 택배 차량이 지금까지 이동하였던 경로를 나타낸다. 이때, P 위치에서 택배물의 수거 요청이 발생하였다면, 그림에서 점선으로 표시된 이



<그림 1> 택배 차량의 이동 경로 예측의 예

들 택배 차량의 현재 위치  $A_1, A_2, A_3$  이후의 이동 경로를 예측한다. 이 결과 P 위치로 이동할 가능성이 가장 높은 택배 차량  $MO_1$ 에게 P 위치의 택배물을 수거하도록 배차함으로써 물류 수송 효율을 높일 수 있다.

### 3. 제안하는 기법

본 장에서는 이동 객체의 과거 궤적을 이용한 미래 경로 예측 방안에 대하여 논한다. 제 3.1절에서는 제안하는 기법의 기본 전략을 요약하고, 제 3.2절에서 과거 궤적을 대상으로 한 유사 궤적 검색 방법에 대하여 설명한다. 제 3.3절에서는 미래 경로 예측 방법에 대하여 설명한다.

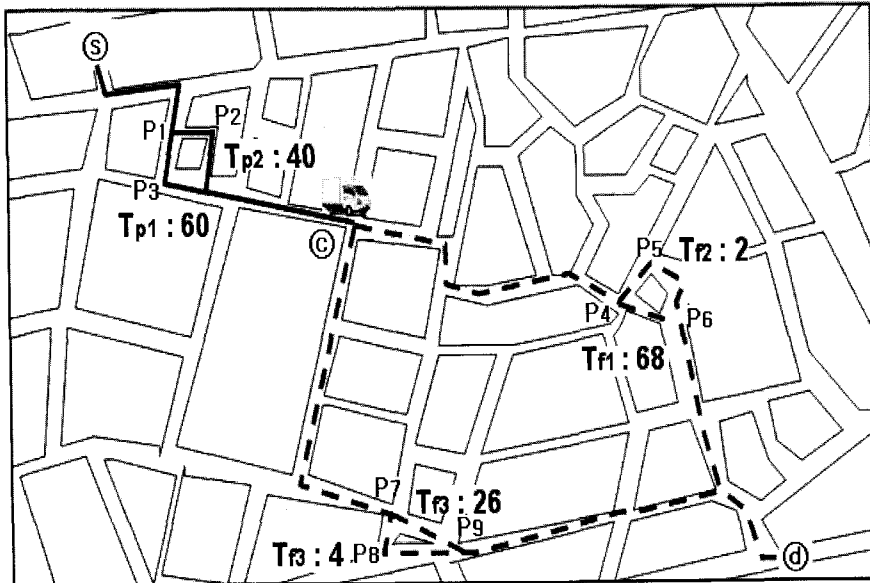
#### 3.1 기본 전략

본 연구에서의 미래 경로 예측 과정은 유사 궤적 검색과 미래 경로 예측의 두 단계로 이루어진다. 도로 네트워크 상에서 시간 흐름에 따라 움직인 이동 객체의 출발지부

터 현재 위치까지의 이동 경로가 질의 궤적으로 서버에 전달되면, 우선 서버에서는 질의 궤적과 이동 패턴이 유사한 후보 궤적을 데이터베이스 내에 저장된 해당 이동 객체의 과거 궤적들로부터 검색한다. 그 다음, 이동 객체의 현재 위치 이후의 이동 경로를 예측하기 위하여 검색된 후보 궤적으로부터 목적지까지의 예측 이동 경로들에 대하여 이용 빈도를 기반으로 확률을 부여하여 사용자에게 전달한다. 이때, 작은 경로 차이를 가지는 궤적들을 유사 궤적으로 그룹핑하여 각 이동 경로들에 대한 개별적인 이용 빈도를 계산하는 대신에 구성된 그룹내에 속한 이동 경로들에 대한 총괄적인 이용 빈도를 계산함으로써 경로 예측에 대한 정확성을 높인다.

직관적인 이해를 돕기 위하여 그림 2의 예를 이용하여 제안하는 미래 경로 예측 과정을 설명하면 다음과 같다.

이동 객체가 ㉓ 위치에서 출발하여  $P_1, P_3$ 를 경유하여 현재 ㉑ 위치에 도착하였고, 최종 목적지는 ㉔ 위치라고 가정한다. 이동 객



<그림 2> 미래 경로 예측의 예

체가 ㉔ 위치 이후에 목적지 ㉕로 어떤 경로를 선택하여 이동하게 될 지를 이동 객체가 움직인 과거 궤적을 이용하여 예측한다. 먼저 이동 객체 데이터베이스 내의 과거 궤적들을 대상으로 ㉓ 위치에서 ㉔ 위치까지의 궤적과 유사한 궤적을 검색한다. 그 결과로 그림 내에서 ㉓에서 출발하여  $P_1$ ,  $P_3$ 을 경유하여 ㉔ 위치까지 완전히 일치하는 60개의 과거 궤적  $T_{p1}$ 과  $P_1$ 과  $P_2$ 를 경유하는 유사한 40개의 과거 궤적  $T_{p2}$ 가 검색되었음을 알 수 있다.

이후 ㉔ 위치부터 ㉕까지의 경로 예측은, 검색된  $T_{p1}$ ,  $T_{p2}$ 의 궤적에 대하여 ㉔ 이후의 경로를 살펴보면  $P_4$ ,  $P_6$ 를 경유하는  $T_{f1}$  궤적으로는 68번,  $P_4$ ,  $P_5$ 를 경유하는  $T_{f2}$  궤적으로는 2번,  $P_7$ ,  $P_9$ 를 경유하는  $T_{f3}$  궤적으로는 26번,  $P_7$ ,  $P_9$ 를 경유하는  $T_{f4}$  궤적으로 4번의 이동 빈도로 목적지 ㉕로 이동하였음을 알 수 있다. 따라서 이동 객체의 현재 위치 ㉔ 이후의 목적지 ㉕로 미래 이동 경로는 이들 궤적 중 이동 빈도가 가장 높은  $T_{f1}$  궤적의 이동 경로를 따라  $P_4$ 와  $P_6$  지점을 경유하여 목적지 ㉕로 이동할 것이라고 예측할 수 있다.

그러나 그림 내의  $T_{f1}$ 과  $T_{f2}$ 의 궤적, 그리고  $T_{f3}$ 와  $T_{f4}$ 의 궤적은 서로 작은 경로 차이만을 갖는 유사 궤적임을 알 수 있다. 따라서 궤적  $T_{f1}$ ,  $T_{f2}$ ,  $T_{f3}$ ,  $T_{f4}$  각각에 대하여 이동 빈도를 계산한 미래 경로 예측 방법 대신에 본 연구에서는 이들 궤적을 유사 궤적으로 묶어 두 개의 그룹  $G_1 = \{T_{f1}, T_{f2}\}$ 과  $G_2 = \{T_{f3}, T_{f4}\}$ 를 구성하고,  $G_1$ 의 이동 빈도( $T_{f1} + T_{f2} = 70$ )와  $G_2$ 의 이동 빈도( $T_{f3} + T_{f4} = 30$ )를 합하여 그룹별로 예측 경로 확률을 제시한다.

### 3.2 유사 궤적 검색

본 연구에서는 이동 객체 데이터베이스를

대상으로 유사 궤적 검색을 수행하기 위하여 먼저 이동 객체의 출발지와 목적지를 이용하여 이와 일치하는 궤적만을 검색 대상에 포함시킨다. 그 다음, 이들 궤적을 대상으로 질의 궤적과 이동 패턴이 유사한 부분 궤적을 갖는 과거 궤적을 검색한다. 도로 네트워크에서의 유사한 부분 궤적 정의는 과거 유틸리티 공간에서의 궤적간 거리 차이를 기반으로 하는 유사 궤적 모델 [10, 11]이나 시공간 네트워크 제약 특성을 나타낸 모델 기법 [12]과 다른 새로운 유사 궤적 판정 모델이 필요하다. 이를 위하여 본 연구에서는 주어진 질의 궤적  $Q$ 와 과거 궤적  $T$ 간의 유사 정도를 측정하기 위해서 아래와 같은 유사 궤적 모델을 정의한다.

본 연구에서는 도로 네트워크 내의 이동 객체 궤적들 간의 유사도를 측정해야 하므로, 제 2장에서 정의된 궤적  $T$ 를 구성하는 도로 세그먼트의 개수와 도로 세그먼트의 길이의 합을 이용한 식(1) DSN(Dissimilarity based on Segment Number)과 식(2) DSL(Dissimilarity based on Segment Length)를 유사도 측정 함수로 사용한다. 이 방식은 두 궤적 간에 일치하는 도로 세그먼트의 개수가 많을수록, 도로 세그먼트의 길이의 차가 작을수록 두 궤적이 유사하다고 판별하는 것이다.

$$DSN(Q, T) = \frac{Q \text{와 } T \text{의 비공통 세그먼트의 개수의 합}}{Q \text{ 세그먼트 개수의 합} + T \text{ 세그먼트 개수의 합}} \quad (1)$$

$$DSL(Q, T) = \frac{Q \text{와 } T \text{의 비공통 세그먼트의 길이의 합}}{Q \text{ 세그먼트 길이의 합} + T \text{ 세그먼트 길이의 합}} \quad (2)$$

본 논문에서는 이와 같은 유사도 측정 함수를 이용한 이동 객체의 유사 궤적 검색 문제를 다음과 같이 정의한다. 질의 궤적  $Q$ 와 유사 허용치  $\epsilon_1$ 와  $\epsilon_2$ 가 주어지면 먼저  $Q$ 의 출발지와 목적지를 이용하여 검색 대상이 되는 궤적만을 이동 객체 데이터베이스에서 여과하고, 여과된 궤적을 대상으로 DSN( $Q$ ,

X)의 값이  $\epsilon_1$  이하이고, DSL(Q, X)의 값이  $\epsilon_2$  이하인 모든 부분 궤적 X를 찾고, X를 포함하는 궤적 T를 반환하는 것이다.

이를 위한 사용자 질의 궤적은  $Q = (\text{userId}, \text{mold}, \text{start}, \text{dest}, \langle \text{qrseg}_1, \text{qrseg}_2, \dots, \text{qrseg}_k \rangle, \epsilon_1, \epsilon_2)$ 의 형태로 표현된다. 여기서,  $\text{userId}$ 는 사용자의 식별자,  $\text{mold}$ 는 이동 객체의 식별자,  $\text{start}$ 는 이동 객체의 출발지,  $\text{dest}$ 는 이동 객체의 목적지,  $\text{qrseg}_i (1 \leq i \leq k)$ 는 출발지에서 현재 위치까지 움직인 이동 객체의 이동 경로로 ( $\text{rsld}_i, \text{rsLen}_i$ )으로 구성된다.  $\epsilon_1$ 와  $\epsilon_2$ 는 유사 궤적 검색과 그룹핑을 위한 유사 허용치로,  $\epsilon_1$ 은 도로 세그먼트의 개수 차이에 대한 비율이며,  $\epsilon_2$ 는 도로 세그먼트의 길이 차이에 대한 비율이다.

이와 같이 비율을 사용하는 이유는 이들 유사 허용치의 적용 대상이 서로 다른 특성을 갖기 때문이다. 즉,  $k$ 개의 도로 세그먼트로 이루어진 출발지부터 현재 위치까지의 질의 궤적 Q와 과거 궤적간의 유사 궤적 검색을 위한 유사도 측정 뿐만 아니라,  $l$ 개의 도로 세그먼트로 이루어진 현재 위치 이후의 목적지까지의 궤적간 그룹핑을 위한 유사도 측정에 적용되어야 하기 때문이다. 예를 들어, 질의 Q의 출발지부터 현재 위치까지의 도로 세그먼트 개수가 20개로 많지 않다면 허용치  $\epsilon_1$ 으로서 1의 값을 사용하여 유사 궤적을 검색하는 것은 유용하다. 그러나 이후 미래 경로를 예측하는 과정에서 현재 위치부터 목적지까지의 도로 세그먼트의 개수가 평균 100개로 비교적 많았다면 허용치  $\epsilon_1$ 으로서 1의 값을 동일하게 사용하여 그룹핑을 하는 것은 무의미하다.

본 연구에서는 구현을 단순하게 하기 위해서 유사 궤적 질의를 처리할 때, 순차 검색 기법을 사용한다. 물론, 별도의 인덱싱 기법을 사용하여 검색 성능을 향상 시킬 수 있으나, 본 연구에서는 다음과 같은 이유로 순차

검색을 그대로 채택한다. 첫째, 각 사용자별로 별도의 데이터베이스를 구성하여 궤적 데이터를 관리하며, 또한 동일 사용자는 동일 이동 경로를 반복적으로 운행하는 성향이 많으므로, 검색 대상이 되는 궤적 데이터의 양이 많지 않기 때문이다. 둘째, 유사 허용치  $\epsilon_1$ 와  $\epsilon_2$ 에 의하여 비교 대상이 되는 궤적의 부분 길이가 길지 않기 때문이다. 셋째, 주어진 유사 허용치  $\epsilon_1$ 와  $\epsilon_2$ 이 크더라도 교차로 지점까지의 도로들의 집합을 하나의 도로 세그먼트로 간주하므로 데이터베이스내에 저장된 원본 궤적의 길이가 길지 않다. 따라서 일반적인 순차 검색 비용은  $O(n)$ 이나, 위에서 언급된 이유로 인하여 검색 대상이 되는 궤적의 수  $n$ 이 많지 않으며, 궤적의 길이  $l$ 이 길지 않기 때문에 비교적 구현이 간단한 순차 검색 방식을 그대로 적용하는 것이 가능하다.

### 3.3 미래 경로 예측

주어진 질의 Q의 현재 위치 이후의 목적지까지의 이동 경로를 예측하기 위한 단순한 방법은 유사 궤적 검색 과정을 통하여 얻어진  $n$ 개의 모든 후보 궤적들을 대상으로 현재 위치 이후부터 목적지까지의 이동 경로들에 대한 이동 빈도를 계산한 후, 빈도가 가장 높은 이동 경로를 이동 객체의 미래 경로로 예측하는 방법을 사용할 수 있다. 그러나 그림 2에 나타난 바와 같이 제안된 유사 궤적 측정방법을 이용하여 허용치 이하의 값을 갖는 궤적을 서로 다른 개별 경로로 취급하여 이동 빈도를 계산하는 것은 예측 결과의 정확도를 떨어뜨릴 수 있다. 따라서 본 연구에서는 이러한 문제점을 해결하기 위하여 검색된 유사 궤적에 대한 그룹핑 기법을 제안한다.

이를 위하여 후보 궤적들을 대상으로 제 3.2절에서 언급한 유사도 측정 함수 DSN과

<표 2> 검색된 후보 궤적의 이동 경로 예

궤적	이동 경로
T <sub>f1</sub>	<(R01, 5), (R02, 2), (R03, 7), (R04, 1), (R05, 1), (R06, 1), (R09, 3), (R10, 2)>
T <sub>f2</sub>	<(R01, 5), (R02, 2), (R03, 7), (R04, 1), (R05, 1), (R07, 2), (R08, 2), (R09, 3), (R10, 2)>
T <sub>f3</sub>	<(R11, 3), (R12, 3), (R13, 2), (R14, 4), (R15, 4), (R18, 5), (R19, 7)>
T <sub>f4</sub>	<(R11, 3), (R12, 3), (R13, 2), (R14, 4), (R16, 1), (R17, 2), (R18, 5), (R19, 7)>

DSL를 이용하여 질의 Q의 현재 위치 이후부터 목적지까지에 해당하는 이동 경로들간의 유사도를 측정한다. 그 다음, 얻어진 DSN 및 DSL 값이 주어진 유사 허용치  $\epsilon_1$ 와  $\epsilon_2$ 를 만족하는 궤적들을 그룹핑한다. 본 연구에서는 이동 객체의 미래 경로 예측을 위한 이동 빈도로 각 이동 경로의 개별적인 이동 빈도 대신에 각 그룹내에 속한 궤적들의 이동 빈도를 합한 그룹의 이동 빈도를 이용한다. 이 과정에서 하나의 궤적이 여러 개의 그룹에 동시에 속하는 경우가 발생할 수 있다. 이 경

우 이동 경로의 빈도가 n이고, 이동 경로가 속한 그룹의 개수가 m이라면 해당 이동 경로의 이동 빈도는  $\lfloor n/m \rfloor$ 가 되며, 이 값을 이용하여 그룹의 이동 빈도를 계산한다.

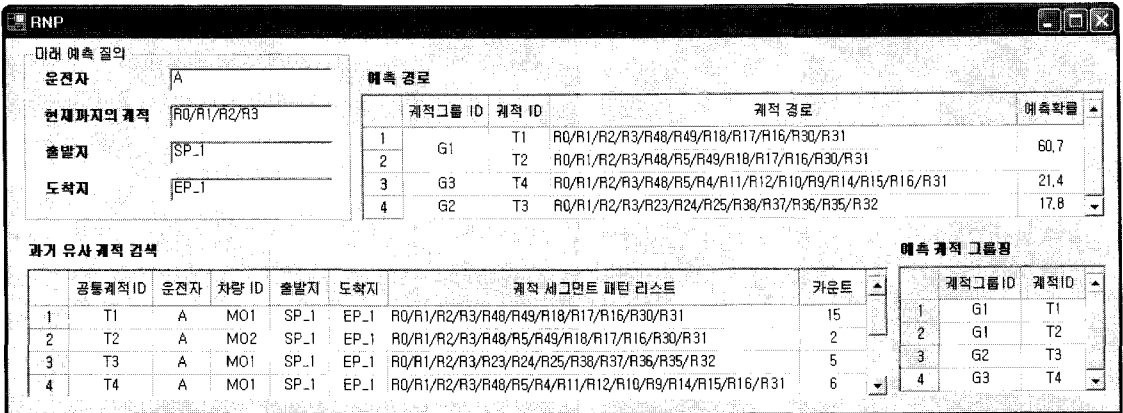
표 2에 그림 2의 궤적 T<sub>f1</sub>, T<sub>f2</sub>, T<sub>f3</sub>, T<sub>f4</sub>에 대하여 현재 위치 ㉠부터 목적지 ㉡까지의 이동 경로를 보인다. 다음 표 3에 그룹핑을 위하여 이들 이동 경로들을 대상으로 계산된 유사도 값을 보인다. 표 3에 의하여 우리는 그룹 G<sub>1</sub> = {T<sub>f1</sub>, T<sub>f2</sub>}와 G<sub>2</sub> = {T<sub>f3</sub>, T<sub>f4</sub>}를 얻을 수 있다.

Algorithm PathPredict(User\_Id, P<sub>start</sub>, P<sub>dest</sub>, P<sub>current</sub>, Traj<sub>query</sub>)

User\_Id: 사용자 Id  
 P<sub>start</sub>: 출발지  
 P<sub>dest</sub>: 목적지  
 P<sub>current</sub>: 현재 위치  
 Traj<sub>query</sub>: 출발지부터 현재 위치까지의 질의 궤적

1. 데이터베이스 내 과거 궤적들 중 P<sub>current</sub>를 포함하며, 사용자 Id, 출발지, 목적지가 각각 User\_id, P<sub>start</sub>, P<sub>dest</sub>인 궤적들을 검색하여 TrajSet에 넣어 둔다.
2. FOR TrajSet 내의 각 궤적 Traj<sub>data</sub>의 P<sub>start</sub>로부터 P<sub>current</sub>까지의 부분궤적 TrajPrefix에 대하여 IF (DSN(TrajPrefix, Traj<sub>query</sub>) <  $\epsilon_1$ ) AND (DSL(TrajPrefix, Traj<sub>query</sub>) <  $\epsilon_2$ ) Traj<sub>data</sub>를 TargetTrajSet에 넣는다.
3. TargetTrajSet내의 궤적들을 SimTrajGroup<sub>i</sub>(1 ≤ i < k)로 그룹핑한다. SimTrajGroup<sub>i</sub>내에서는 임의의 두 궤적 Traj<sub>i,x</sub>와 Traj<sub>i,y</sub>의 P<sub>current</sub>부터 P<sub>dest</sub>까지의 각 부분 궤적 TrajPostfix<sub>i,x</sub>과 TrajPostfix<sub>i,y</sub>에 대하여 (DSN(TrajPostfix<sub>i,x</sub>, TrajPostfix<sub>i,y</sub>) <  $\epsilon_1$ ) AND (DSL(TrajPostfix<sub>i,x</sub>, TrajPostfix<sub>i,y</sub>) <  $\epsilon_2$ )가 성립한다.
4. 각 궤적 그룹 SimTrajGroup<sub>i</sub>(1 ≤ i < k)에 대해서 예측 확률 |SimTrajGroup<sub>i</sub>|/|TargetTrajSet| 을 부여한다.





<그림 3> 유사 궤적 검색의 예

<표 3> 후보 궤적간의 유사도

궤적	DSN	DSL
T <sub>11</sub> , T <sub>12</sub>	0.176	0.106
T <sub>11</sub> , T <sub>13</sub>	1	1
T <sub>11</sub> , T <sub>14</sub>	1	1
T <sub>12</sub> , T <sub>13</sub>	1	1
T <sub>12</sub> , T <sub>14</sub>	1	1
T <sub>13</sub> , T <sub>14</sub>	0.2	0.127

출발지부터 목적지까지의 전체 궤적을 대상으로 그룹핑을 수행하지 않는 이유는 질의 Q내의 출발지와 현재 위치가 사용자마다 다를 수 있기 때문이다. 이를 처리하기 위하여는 궤적의 모든 부분 궤적을 추출한 후, 이에 대하여 그룹핑을 수행하여야 하므로 많은 오버헤드가 발생한다. 따라서 본 연구에서는 출발지부터 현재 위치까지는 순차 검색에 의한 유사 궤적 검색을 수행하고, 현재 위치 이후부터 목적지까지의 경로에 대해서만 클러스터링을 수행한다. 다음 알고리즘 1은 미래 경로 예측을 위한 제안된 기법의 전체적인 처리 과정을 나타낸다.

그림 3에 유사 궤적 검색의 예를 보인다. 운전자 A의 출발지는 SP\_1이고, 도착지는 EP\_1이고, 출발지부터 현재 위치까지의 이동 경로는 (R0-R1-R2-R3)이다. 이들 값

을 이용하여, 먼저 운전자 A의 모든 과거 궤적 중에서 출발지가 SP\_1이고 목적지가 EP\_1인 후보 궤적만을 검색한다. 그림에 나타난 바와 같이 검색된 후보 궤적을 대상으로 주어진 경로 (R0-R1-R2-R3)와 유사한 부분 궤적을 갖는 T1, T2, T3, T4의 궤적을 얻었으며, 각 궤적의 이동 빈도는 15, 2, 5, 6임을 알 수 있다. 또한, 얻어진 궤적 T1, T2, T3, T4의 궤적들에 대하여 그룹핑을 수행하여 G1={T1, T2}, G2={T3}, G3={T4}의 그룹을 구성하고, 이들 그룹에 대한 이동 빈도를 구한다. 그림에서 얻어진 그룹 G1, G2, G3의 이동 빈도는 각각 17, 5와 6이 된다. 즉, T1, T2로 이동할 확률은 60.7%, T3로 이동할 확률은 17.8%, T4로 이동할 확률은 21.4%이다. 최종 결과로 사용자에게 이동 확률이 높은 순서대로 목적지까지의 예측된 이동 경로들의 리스트를 제시한다.

#### IV. 결론

본 논문에서는 도로 네트워크 상에서 이동 객체의 미래 경로를 예측하는 기법에 대하여 논하였다.

제안된 기법에서는 미래 경로 예측을 위하여 이동 객체 데이터베이스내에 축적되어 있는 이동 객체의 과거 궤적을 대상으로 주어진 질의 궤적과 이동 패턴이 유사한 궤적을 검색하는 방식을 사용한다. 이를 위하여 본 연구에서는 궤적을 도로 네트워크의 특성을 반영한 도로 세그먼트의 연속으로 표현하고, 도로 세그먼트의 개수와 길이를 이용한 새로운 유사도 측정 방식을 제안하고, 이를 기반으로 하는 미래 경로 예측 기법을 제안하였다. 또한, 미래 경로 예측의 정확성을 높이기 위하여 작은 경로 차이를 갖는 이동 경로들을 그룹핑하고, 이를 경로 예측 과정에 활용하는 방안을 제안하였다.

향후 연구로서 실험을 통하여 제안된 유사 궤적 검색 기법과 미래 경로 예측 기법에 대한 정확성 및 효율성을 검증하기 위한 성능 평가를 수행할 예정이다. 또한, 현재 제안된 기법은 운전자와 이동 객체의 현재 위치까지의 이동 경로, 출발지, 목적지 등이 입력 값으로 주어진다든 특수 상황을 대상으로 한다. 향후 이를 보다 일반화된 상황에 적용하기 위한 연구를 추가로 수행할 예정이다.

## 참고문헌

1. S. Saltenis, C. Jensen, S. Leutenegger, and M. A. Lopez, "Indexing the Positions of Continuously Moving Objects," In *Proc. Int'l. Conf. on Management of Data*, ACM SIGMOD, 2000, pp. 331-342.
2. R. Guting et al., "A Foundation for Representing and Querying Moving Objects," In *Proc. ACM Trans. on Database Systems*, Vol. 25, No. 1, 2000, pp. 1-42.
3. G. Kollios, D. Gunopulos, and V. Tsotras, "On Indexing Mobile Objects," In *Proc. Int'l Symp. on Principles of Database Systems*, ACM PODS, 1999, pp. 261-272.
4. D. Pfoser, C. Jensen, and Y. Theodoridis, "Novel Approaches in Query Processing for Moving Object Trajectories," In *Proc. Int'l. Conf. on Very Large Data Base*, VLDB, 2000, pp. 395-406.
5. E. Pitoura, and G. Samaras, "Locating Objects in Mobile Computing," *IEEE Trans. on Knowledge and Data Engineering*, Vol. 13, No. 4, 2000, pp. 571-592.
6. M. Nascimento and J. Silva, "Towards Historical R-trees", In *Proc. ACM Symp. on Applied Computing*, ACM SAC, 1998, pp. 235-240.
7. Y. Tao, D. Papadisas and J. Sun, "The TPR\*-tree: An Optimized Spatio-Temporal Access Method for Predictive Queries," In *Proc. Int'l. Conf. Very Large Data Base*, VLDB, 2003, pp. 790-801.
8. N. Beckmann, H. Kriegel, R. Schneider, and B. Seeger, "The R\*-tree: An Efficient and Robust Access Method for Points and Rectangles," In *Proc. Int'l. Conf. on Management of Data*, ACM SIGMOD, 1990, pp. 322-331.
9. K.-S. Kim et al., "Fast Indexing and Updating Method for Moving Objects on Road Networks," In *Proc. IEEE Int'l.*

- Conf. on Web Information Systems Engineering*, 2003, pp. 34-42.
10. L. Chen and R. Ng, "On the Marriage of Lp-norms and Edit Distance," In *Proc. Int'l. Conf. on Very Large Data Bases, VLDB*, 2004, pp. 1040-1049.
  11. Y. Yanagisawa, J. Akahani, and T. Satoh, "Shape-Based Similarity Query for Trajectory of Mobile Objects," In *Proc. Int'l. Conf. on Mobile Data Management*, 2003, pp. 63-77.
  12. 황 정래, 강 혜영, 이 기준, "시공간 유사성을 이용 도로 네트워크 상의 유사한 궤적 검색," *정보처리학회 논문지*, Vol. 13-D, No. 3, 2006, pp. 337-346.
  13. V. Almeida and R. Güting, "Indexing the Trajectories of Moving Objects in Networks," *Geoinformatica*, Vol. 9, No. 1, 2005, pp. 33-60.
  14. C. Shahabi, M. Kolahdouzan, and M. Sharifzadeh, "A Road Network Embedding Technique for K-Nearest Neighbor Search in Moving Object Databases," In *Proc. Int'l. Symp. on Advances in Geographic Information Systems, ACM GIS*, 2002, pp. 94 - 100.
  15. R. Benetis et al., "Nearest Neighbor and Reverse Nearest Neighbor Queries for Moving Objects," In *Proc. Int'l. Conf. on Database Engineering Applications Symposium, IDEAS*, 2002, pp. 44-53.
  16. N. V. de Weqhe et al., "Representation of Moving Objects along a Road Network," In *Proc. Int'l. Conf. on Geoinformatics, Geospatial Information Research*, 2004.
  17. M. Vazirgiannis, and O. Wolfson, "A Spatiotemporal Model and Language for Moving Objects on Road Networks," In *Proc. Int'l. Symp. on Spatial and Temporal Databases, SSTD*, 2001, pp. 20-35.
  18. Thomas Brinkhoff, "A Framework for Generating Network-based Moving Objects," *GeoInformatica*, Vol. 6, No. 2, 2002, pp. 153-180.
  19. A. Civilis et al., "Techniques for Efficient Road-Network-Based Tracking of Moving Objects," In *Proc. IEEE Trans. on Knowledge and Data Engineering*, Vol. 17, No. 5, 2005, pp. 698-712.
  20. H. Hu, D. Lee, and V. Lee, "Distance Indexing on Road Networks," In *Proc. Int'l. Conf. Very Large Data Base, VLDB*, 2006, pp. 894-905.
  21. L. Speicys, C. Jensen, and A. Kligys, "Computational Data Modeling for Network-Constrained Moving Objects," In *Proc. Int'l. Symp. on Advances in Geographic Information Systems, ACM GIS*, 2003, pp 118-125.

**김종대**

2003년 2월 숭실대학교 컴퓨터학부 졸업(학사)  
2003년 3월~2006년 1월: 에넥스 Information System 부서 개발팀원  
2006년 3월~현재: 한양대학교 정보통신대학 정보통신학부 연구원(석사)  
관심분야: 데이터베이스 시스템, 데이터 마이닝, 시공간 인덱싱, 플래쉬 메모리 DBMS

**김상욱**

1989년 2월 서울대학교 컴퓨터공학과 졸업(학사)  
1991년 2월 한국과학기술원 전산학과 졸업(석사)  
1994년 2월 한국과학기술원 전산학과 졸업(박사)  
1991년 7월~8월 미국 Stanford University, Computer Science Department 방문 연구원  
1994년 2월~1995년 2월 KAIST 정보전자연구소 전문 연구원  
1999년 8월~2000년 8월 미국 IBM T.J. Watson Research Center Post-Doc.  
1995년 3월~2000년 8월 강원대학교 컴퓨터정보통신공학부 부교수  
2003년 3월~현재 한양대학교 정보통신대학 정보통신학부 교수  
관심분야: 데이터베이스 시스템, 저장 시스템, 트랜잭션 관리, 데이터 마이닝, 멀티미디어 정보 검색, 공간 데이터베이스/GIS, 주기억장치 데이터베이스, 이동 객체 데이터베이스/텔레매틱스, 사회 연결망 분석, 웹 데이터 분석

**원정임**

1992년 2월 한림대학교 전자계산학과 졸업(학사)  
1997년 8월 한림대학교 전자계산학과 졸업(석사)  
2004년 2월 한림대학교 전자계산학과 졸업(박사)  
2000년 3월~2004년 2월 한림대학교 교양교육부 강의전담교수  
2004년 3월~2006년 2월 연세대학교 컴퓨터과학과 연구교수  
2006년 3월~6월 서울대학교 유전자이식연구소 선임연구원  
2006년 7월~현재 한양대학교 정보통신대학 정보통신학부 연구교수  
관심분야: 데이터베이스 시스템, 데이터 마이닝, XML 응용, 바이오 정보공학, 데이터베이스 보안, 이동객체 데이터베이스, 텔레매틱스