

수직구조 기반의 빈발 이동 객체 패턴 탐색 기법

A Pattern Retrieval Method of Frequent Moving Objects Using Vertical-Based Framework

홍 성 한*, 황 병 연
Sung-Han Hong* Byung-Yeon Hwang
가톨릭대학교 컴퓨터공학과
{hongta, byhwang}@catholic.ac.kr

요 약 최근에 모바일 기기의 급속한 발전과 더불어 다양한 서비스들이 개발되고 있다. 그 중에서도 위치 기반 서비스는 사용자에게 위치와 관련된 유용한 정보를 제공하는 서비스를 말한다. 효과적인 서비스를 제공하기 위해서는 먼저 위치정보를 나타내는 이동 객체 관련 기술 연구가 선행되어야 한다. 이러한 연구의 핵심 기술로 현재 빈발한 이동 객체 탐색을 위한 마이닝 기법들에 관한 연구가 진행되고 있다. 본 연구에서는 기존의 수평적 마이닝 기법에서 문제시되었던 많은 후보 이동 객체 발생을 줄이기 위해 새로운 수직적 마이닝 기법을 적용한 방법을 제안한다.

1. 서 론

최근에 모바일 기술의 발전과 더불어, 사용자의 위치정보를 활용한 다양한 연구가 진행되고 있다. 위치기반 서비스는 개인의 위치를 파악하여 각종 마케팅이나 고객 서비스 등을 개인화하여 서비스를 제공하는 기술이다[1]. 위치 정보에 대상이 되는 PDA, 이동전화, 차량 등은 시간의 흐름에 따라 위치가 연속적으로 변경되는 특징을 가지고 있다[2, 3]. 이러한 PDA, 이동전화, 차량 등에서 수집되는 데이터를 이동 객체(Moving Object)라고 한다[1, 2, 6].

최근 시간에 따른 공간 패턴 변화에 관한 연구가 활발히 진행되고 있다. 기존의 연구들은 정적 데이터 중심의 마이닝 기법을 이동 객체의 패턴을 탐색하는데 적용하고 있다. 그

러나 시공간적 개념을 동시에 고려해야 하는 많은 공간 데이터들을 처리해야 하므로 기존의 정적 데이터를 처리하는 마이닝 기법으로는 그 한계가 있다. 따라서 수행시간과 저장공간에 대한 효율적인 탐색 기법에 관한 연구가 필요하다.

이와 관련된 연구로는 수평적 구조기반 패턴 마이닝을 직접 이동 객체 패턴을 탐색하는데 적용한 연구가 있다. 수평적 구조 기반 패턴 마이닝은 후보-빈발 구조의 2중 구조를 가지고 있어, 수행 시간 측면에서 효율성이 저하된다[1, 2]. 따라서 본 논문에서는 탐색 수행 시간에 효율성을 개선하고자 수직적 구조 기반의 빈발 탐색 방법을 제안하려고 한다.

이 논문의 구성은 다음과 같다. 2장에서는 관련연구를 기술하고, 3장에서는 이동 객체 탐색을 위한 일반화 과정을 기술한다. 4장에

서는 본 논문에서 제안하는 방법을 이용한 실험 예를 기술하고, 5장에서는 결론에 대해 기술한다.

2. 관련 연구

이동 객체들의 패턴 탐색은 우선적으로 시공간적 속성을 고려해야 한다. 여기에 데이터 마이닝 기법을 접목시키면 효율적인 탐색이 가능하다. 기존의 정적 데이터 분석 중심의 데이터 마이닝 기법은 대부분 시간과 공간을 분리하여 패턴을 탐색해 왔다[2, 6]. 시간패턴 중심의 기법으로는 순차 패턴, 주기 패턴, 시간 관계 패턴 등이 있다[3]. 대부분의 기법들은 공간적 속성을 고려하지 않았으므로 직접적으로 이동 객체 패턴을 탐색하는 것이 어렵다.

[1, 2]에서는 시간적 제약조건 기반에 공간적 속성을 고려한 탐색기법을 제안하였다. 이 기법은 시간과 공간적 속성을 이동 객체들의 패턴을 찾는데 적용하기 위해서 이동 객체들을 일반화시키고, 정적 데이터로 변환시켜 빈발한 패턴을 탐색하였다. 특히 시간 속성의 주기적 변화에 대한 공간 속성 위치를 2차원 평면상의 좌표로 나타내고 공간 연산 Contain()을 이용하여 일반화했다. 그리고 기존의 정적 데이터 패턴 분석인 순차패턴 마이닝(Sequence Pattern Mining) 기법을 적용해서 빈발한 이동객체들에 대한 패턴을 탐색하였다. 후보-빈발 항목의 2중 구조를 가진 수평적 패턴 마이닝 기법을 적용하였기 때문에 반복적인 스캔과 수행 시간의 효율성이 저하되는 문제점을 가지고 있다.

본 연구에서는 [2]에서 제안했던 수평적 구조 기반의 이동 객체 패턴 탐색 시 발생하는 초기 후보 항목 탐색 시간의 비효율성 문제를 보완하고자 수직적 구조 기반의 빈발 패턴 탐색 기법을 제안한다.

3. 이동 객체의 탐색을 위한 일반화

3.1 이동 객체의 표현

이동 객체의 위치변화는 특정 시점에 따라 연속적(continue)으로 혹은 이산적(discrete)으로 변한다. 이런 변화 때문에 객체를 표현하는 다양한 방법들이 존재한다. 이동 객체들의 공간 속성을 고려하기 위해서는 객체에 대한 고정된 시점에서 위치를 샘플링 해야 한다. 왜냐하면 컴퓨터 내에서 동적인 객체를 표현하기 어렵기 때문이다[1, 2, 3].

이동 객체가 일반성을 갖기 위해서는 특정 시점에서 위치가 고정된 것으로 가정 한다. 아래의 [정의 1]은 이동 객체 탐색 일반화에 대한 정의이다.

[정의 1]

이동 객체 A 를 유한집합이라고 가정하자. MT 는 이동 시간(Moving Time), 위치 좌표는 $P(X_A, Y_B)$, 객체의 유일한 식별자는 OID 이다.

$$A = \langle OID, MT, P(X_A, Y_B) \rangle$$

이동 객체에서 빈발한 패턴을 탐색하기 위해서는 탐색할 시점과 연속된 이동 객체들에 대한 공간 영역도 함께 정의해야 한다. 또한 시간 제약에 따른 연속된 객체의 순서를 이동 시퀀스(Moving Sequence)라고 하며, 이동 시퀀스들 간의 빈발한 패턴을 추출하기 위해서는 지지도에 대한 정의도 필요하다[1, 2, 3].

[정의 2]

이동 시퀀스에 관한 지지도(Support)는 전체 이동 시퀀스의 비율에 대한 특정 시퀀스를 비율로 정의한다.

[정의 3]

시점은 특정한 시간동안에 발생하는 이동 객체를 샘플링하는 일정한 구간으로 정의한다.

3.2 이동 객체들의 정적 데이터 변환

<표 1>은 각 이동 객체들의 시간에 따른 공간 좌표를 임의로 측정한 다음, 정적 좌표로 변환한 예이다.

<표 1> 이동 객체 변환

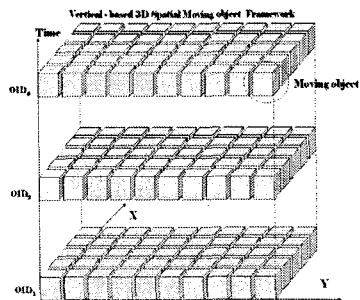
OID	Moving Time	X	Y
OID ₁	2005.11.10.09:10	0	17
	2005.11.10.09:20	44	13
	2005.11.10.09:30	23	45
	2005.11.10.09:40	35	03
OID ₂	2005.11.10.13:10	5	20
	2005.11.10.13:20	41	58
	2005.11.10.13:30	45	17
	2005.11.10.13:40	35	61
OID ₃	2005.11.10.15:10	47	11
	2005.11.10.15:20	25	38
	2005.11.10.15:30	31	67
	2005.11.10.15:40	44	06

OID는 이동객체 아이디이고, X와 Y는 좌표를 의미하며, Moving Time은 이동시간을 의미한다. 해당 시간에 이동 객체 위치를 표현한 것이다.

3.3 이동 객체의 3차원 공간구조

기존의 방법에서는 임의의 이동 객체가 어느 셀에 속하는지를 나타내기 위해서는 2차원 평면에 좌표로 나타내었고 이것을 공간 인덱싱 기법인 Contain 연산방법을 이용하여 각 이동 객체의 공간위치를 구했다[2].

그러나 본 논문에서 제안하는 수직적 구조 기반의 빈발 패턴탐색을 하기에 2차원구조는 적합하지 않다. 따라서 기존의 3차원 공간 인덱싱 기법인 3D R-Tree[4, 7]를 변형시켜 사용하려 한다. <그림 1>은 수직적 구조 기반의 빈발 패턴 탐색에 적합한 3차원 이동 객체 공간 구조를 나타낸다.



<그림 1> 3차원 이동 객체 공간 구조

<그림 1>에서 사용한 데이터는 <표 1>의 데이터를 이용하여 3차원 공간으로 나타낸 것이다. X축과 Y축은 공간 좌표를 의미하며, 시간 간격 축 T(OID)는 분단위로 구분된다. 단위 셀(Unit Cell)은 객체의 위치를 의미한다.

3.4 이동 객체의 특정 패턴 표현

<그림 1>에서는 3차원 공간상에 이동 객체들을 각 셀로 나타냈다. 이렇게 표현된 각 이동 객체들을 분석하면, 특정한 공간 위치에서 특이한 경향이나 패턴이 발견된다. 예를 들어 2005년 11월 10일 9시 10분에 측정된 이동 객체와 13시 10분에 측정된 이동 객체 위치는 유사하다.

따라서 <표 1>에서 탐색된 각 이동 객체들을 공간영역으로 변화하여 나타내면 <표 2>와 같이 나타낼 수 있다.

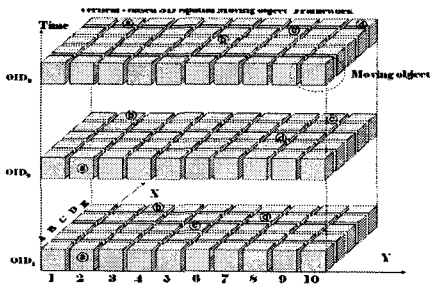
<표 2> 이동 객체 패턴

OID	Moving Time	Spatial Area
OID ₁	2005.11.10.09:10	A
	2005.11.10.09:20	B
	2005.11.10.09:30	C
	2005.11.10.09:40	D
OID ₂	2005.11.10.13:10	A
	2005.11.10.13:20	E
	2005.11.10.13:30	B
	2005.11.10.13:40	C
OID ₃	2005.11.10.15:10	B
	2005.11.10.15:20	C
	2005.11.10.15:30	D
	2005.11.10.15:40	E

4. 수직적 빈발 패턴 구조

<표 2>에서 나타낸 연속된 이동 객체들의 특정한 패턴들을 분석하여 그 경향을 알아보기 위해서는 새로운 수직적 빈발 패턴 구조가 필요하다. <그림 2>는 새로운 패턴 분석 구조를 나타낸 것이다. 각 이동 객체들이 유사한 공간 좌표 값에 따라 특정한 공간상에 분포하게 된다.

빈발한 이동 객체 패턴을 분석하기 위해서는 <그림 2>에 나타난 형태를 표현하며, 이동 객체에 대한 위치 패턴을 분석한다.



<그림 2> 수직기반 빈발 패턴 구조

4.1 빈발한 이동 객체의 추출 단계

빈발한 이동 객체를 추출하기 위해서는 <표 2>에 나타난 특정한 시간 제약 조건에 따른 이동 객체들을 패턴 마이닝 해야 한다. 일반적인 데이터 마이닝 분석 테이블로 나타내면 <표 3>과 같이 나타낸다.

<표 3>에서는 트랜잭션을 OID로 변형하였고, 각각의 아이템을 이동 객체로 변형하여, 마이닝 기법을 적용하기 편하도록 테이블 형태로 나타냈다.

<표 3> OID-Moving Object 테이블

OID	Moving Object
1	A, B, C, D
2	A, E, B, D
3	B, C, D, E

4.2 수직적 빈발 패턴 마이닝 단계

<표 3>에 나타난 내용을 중심으로 수평구조 중심의 데이터 마이닝 기법을 적용하면 빈발한 이동 객체 패턴을 추출할 수 있다. 하지만 이동 객체의 수가 많을 경우에는 빈발한 패턴을 탐색하기 위해서 많은 수행시간과 저장 공간이 필요한 단점을 가진다.

수직 구조 중심의 마이닝 기법은 빈발한 항목을 찾기 위해서 계속적으로 방대한 양의 데이터베이스를 스캔할 필요 없이 한 번의 스캔으로 빈발 항목을 쉽게 찾을 수 있다[4]. 이런 기법을 이동 객체 탐색하는데 적용한다면 기존의 수평적 구조 마이닝 기법보다 더 빠른 탐색이 가능하다.

4.2.1 수직 구조 트랜잭션 테이블

<표 3>에 나타난 식별자-이동 객체 테이블을 수직 구조 탐색을 위해 변형하면 <표 4>와 같다. 이동 객체들이 고유 ID형태로 변형되고, 각 이동 객체들은 트랜잭션 항목을 나타내는 구조로 변형된다. 아래의 <표 4>는 수직 구조 이동 객체 테이블을 나타낸 것이다.

<표 4> 수직 구조 이동 객체 테이블

Moving Object	OID
A	1, 2
B	1, 2, 3
C	1, 3
D	1, 2, 3
E	2, 3

4.2.2 수직구조 빈발 이동 패턴 마이닝

<표 4>에 나타난 데이터를 가지고 빈발한 패턴을 찾기 위해서는 최소지지도를 고려하여 가장 빈발한 객체들을 찾아야 한다. 따라서 각 이동 객체에 대한 지지도를 계산하기 위해 Bit Vector를 이용하여 표현하면 <표 5>와 같다.

그러면 <표 5>에서 bit 값이 1인 경우만을 카운트하여 지지도를 구한다. 임의의 최소지지도 값을 입력받아 최소지지도 값을 넘지 않는 객체들은 제거된다.

예들 들어 최소지지도를 3으로 가정하면 A, C, E는 제거 된다.

<표 5> 이동 객체 지지도 계산

Moving Object	O1	O2	O3	Support
A	1	1	0	2
B	1	1	1	3
C	1	0	1	2
D	1	1	1	3
E	0	1	1	2

4.3 최대 빈발 이동 객체 패턴 탐색

<참 고 문 헌>

최대로 빈발한 이동 객체 시퀀스를 찾기 위해서는 후보 항목의 수를 줄일 수 있는 Apriori 알고리즘을 이용한다[5]. 이 알고리즘은 탐색할 데이터에 대해서 여러 번의 스캔 과정을 거친다. 각 패스에서 후보항목을 생성하고 데이터베이스에서 후보 항목을 포함하는 이동 객체를 탐색한다[5].

데이터베이스에 포함된 후보항목을 여러 번 스캔하기 때문에 탐색 효율이 떨어진다. 본 연구에서 제안한 수직적 기반 패턴 방법을 적용하면 한번의 스캔 과정을 통해 빠른 후보 항목을 찾을 수 있다.

5. 결 론

본 연구에서는 위치 기반 서비스를 제공하기 위해서 필요한 이동 객체 탐색에 관한 수직적 구조 기반 패턴 탐색기법을 제안하였다. 이 기법은 기존의 수평적 구조 기반의 패턴 탐색 방법의 단점인 초기 후보 탐색 시 발생하는 여러 번의 스캔으로 인한 탐색 시간의 비효율성 문제를 보완하였다. 본 논문에서 제안한 수직적 마이닝 기법을 적용하면 보다 빠른 빈발 항목의 탐색이 가능하므로 빈발한 이동 객체를 찾는데 좀더 효율적이다.

- [1] 이준옥, 남광후, “이동 객체 위치 일반화를 이용한 시공간 이동 패턴 탐사,” 정보처리학회논문지, 10-D권, pp.1103-1114, 2003.
- [2] 이준옥, 백옥현, 류근호, “위치 기반 서비스를 위한 이동 객체의 시간 패턴 탐사 기법,” 정보과학회논문지, 제29권, 제5호, pp.335-346, 2002.
- [3] 이용준, 서성보, 류근호, 김혜규, “시간간격을 고려한 시간관계 규칙 탐사 기법,” 정보과학회논문지, 제28권, 제3호, pp.52-56, 2001.
- [4] 고윤희, 김현철, 이원규, “LSI 유사도를 이용한 효율적인 빈발항목 탐색 알고리즘,” 한국컴퓨터교육학회논문지, 제6권, 제1호, pp.1-8, 2003.
- [5] R. Agrawal and R. Srikant, "Fast Algorithm for Mining Association Rules," Proceedings of the 20th International Conference on Very Large Database, pp.487-499, Santiago, Chile, 1994.
- [6] 안윤애, 류근호, “이동 객체의 불확실한 위치 정보 관리,” 충북대학교 컴퓨터정보통신연구, 제9권, 제1호, pp.81-91, 2000.
- [7] 이양구, 이웅재, 남광우, 류근호, “과거 및 현재 정보의 통합 관리를 위한 이동체 색인 기법,” 개방형지리정보시스템학회, 추계학술대회, pp.153-158, 2003.