
단일 이동 객체 궤적에 대한 효율적인 분할 알고리즘에 관한 연구

박주현* · 조우현**

A Study on Efficient Split Algorithms for Single Moving Object Trajectory

Ju-Hyun Park* · Woo-Hyun Cho**

요 약

무선 네트워크 기술의 발달로, 시공간 오브젝트의 위치 정보를 저장하는 것은 아주 필수적인 일이 되었다. 하지만, 시공간 오브젝트의 움직임은 필요하지 않은 너무 많은 위치 정보를 포함하기 때문에 모든 위치 정보를 저장하는 것은 검색에 있어서 아주 비효율적이다. 따라서 본 논문에서는 시공간 오브젝트에서 필요하지 하지 않는 정보를 제거하여 검색의 효율을 높일 수 있는 효율적인 궤적을 분할하는 방법을 제시한다. 이 선형병합 분할 알고리즘은 EMBR을 이용하여 MBR들의 면적을 최소화 하는 궤적을 분할 알고리즘이다. 실험의 결과로 제시하는 분할 방법이 다른 알고리즘보다 더 효율적인 것을 알 수 있다.

ABSTRACT

With the development of wireless network technology, Storing the location information of a spatiotemporal object was very necessary. Each spatiotemporal object has many unnecessary location information, hence it is inefficient to search all trajectory information of spatiotemporal objects. In this paper, we propose an efficient method which increase searching efficiency. Using EMBR(Extend Minimum Bounding Rectangle), an LinearMarge split algorithm that minimizes the volume of MBRs is designed and simulated.

Our experimental evaluation confirms the effectiveness and efficiency of our proposed splitting policy

키워드

이동 객체, 궤적 분할, 시공간 색인

Key word

Moving Object, Trajectory Splitting, Spatiotemporal Indexing

* 정회원 : 부경대학교 (주저자, holy4001@empal.com)

** 정회원 : 부경대학교 (교신저자)

접수일자 : 2011. 07. 29

심사완료일자 : 2011. 09. 14

I. 서 론

최근 이동통신, 무선 네트워크, GPS(Global Position System)의 발전으로 인해 공간상에서 시간에 따라 다양하게 변화하는 이동 객체들의 위치 정보를 수집하고 저장하여 다양한 어플리케이션에서 활용하는 기술이 늘어나고 있는 추세이다. 그러므로 이런 정보를 저장하고 활용하는 이동 객체의 정보 저장을 하기 위한 데이터베이스의 기술 또한 관심을 받고 연구되고 있다.

이동 객체란 시간의 변화에 따라 공간적인 위치가 연속적으로 변화하는 시공간 데이터를 의미한다. 이러한 이동 객체의 색인을 하는 방법에는 여러 가지가 있는데 본 논문에서는 이동 객체의 궤적을 검색하기 위한 방법에서 검색 영역을 효율적으로 분할하여 불필요한 검색 횟수를 줄이는 방법에 초점을 맞추고 있다.

이동 객체의 궤적을 기반으로 하는 데이터를 검색하기 위한 질의 방법에는 두 가지가 있는데 좌표 기반 질의와 궤적 기반 질의가 있다.[1] 첫 번째 좌표 기반 질의는 주어진 시간 간격 동안 어떠한 공간을 지나가는 궤적의 객체 식별자와 궤적의 객체의 포인트들을 반환한다. 이렇게 반환된 궤적의 객체의 포인트들을 최소 경계 사각형으로 변환하여 입력된 공간 질의에 포함이 되는지를 찾게 된다. 두 번째 궤적 기반 질의는 궤적들이 서로 복잡하게 교차하는 위상적인 관계 혹은 주어진 시간 간격 동안 어느 공간 안에서 속도변화와 이동방향등의 정보를 반환하게 된다. 이 역시 위상적인 관계나 속도변화와 이동방향등의 정보를 이용하여 최소 경계 사각형으로 구성하여 공간 질의에 포함되는지를 찾는 방법이다. 이처럼 궤적 기반의 데이터를 검색하기 위한 방법에서는 공간 질의를 처리하기 위해서 궤적을 분할하여 최소경계사각형(MBR : Mininum Bounding Rectangle)들을 생성하고 이를 이용하여 색인을 생성하는 것이 공간 질의를 처리하기 위해서 아주 중요하다. 하지만 어떤 시공간상의 궤적들과 그 궤적이 갖는 객체의 포인트들은 무수히 많다. 이러한 궤적의 모든 정보를 찾는 것은 많은 시간이 소요된다. 따라서 본 논문에서는 검색 시간을 단축하기 위한 효율적인 색인을 생성하기 위해 저장 공간상의 궤적의 분할의 개수를 제한하는 알고리즘을 제안한다.

제안하는 방법은 질의범위의 크기를 고려한 확장된

최소경계사각형(EMBR : Extended MRB)의 면적을 최소화하는 알고리즘인 선형병합 분할(LinearMerge Split) 알고리즘을 적용하여 EMBR의 총면적이 가장 최소에 가까운 분할을 만들어 내어 색인 구성 후 질의 수행 과정에서 불필요한 탐색 공간을 감소시킬 수 있도록 제안한다. 성능 평가 결과로 병합 분할 방법(MergeSplit)[16]보다 EMBR의 총합에 있어서 더 최적의 분할이 된다는 것을 알 수 있었다.

본 논문의 구성은 다음과 같다. 2절에서는 알고리즘에 필요한 관련 연구를 설명하고, 3절에서는 효율적인 분할을 만드는 선형병합 분할 알고리즘에 대해 제안한다. 4절에서는 제안하는 알고리즘과 기존의 알고리즘과의 성능 평가를 통해 제안하는 알고리즘의 우수성을 비교 분석한다. 5절에서는 결론 및 향후 연구 과제에 대해 기술한다.

II. 관련 연구

이동 객체의 위치를 검색하기 위해 쓰이는 방법에는 공간 분할 색인과 데이터 분할 색인으로 크게 두 가지로 나눈다. 공간 분할 색인에는 격자 파일(Grid File)[8]과 사분 트리(Quad Tree)[10]와 같은 방법이 있고 데이터 분할 색인에는 R-tree[9] 방법이 있다.

이동객체의 색인을 위해 제안된 방법들은 이동 객체의 이력정보 탐색을 위한 색인 방법[2, 5-7]과 이동 객체의 궤적 탐색을 위한 색인 방법[1, 4,, 9, 12, 15]이동 객체의 미래 위치를 예측하여 탐색하기 위한 색인 방법[1, 4,, 9, 12, 15]이 있다.

여기서, 이동 객체의 궤적 탐색을 위한 색인 방법에는 공간적으로 처리 할 것인지 시간적 차원을 처리할 것인지에 따라 크게 세가지 방법으로 나누어 진다.

첫 번째 방법은 공간적 차원에 시간을 추가하는 방법인데 여기에는 기존의 R-tree에 시간을 추가하여 공간적인 차원으로 확장하여 만드는 TB-Tree[3] 방법이 있다. 두 번째 방법은 SERI-tree[7]와 SEB-tree[9]를 이용하여 시간과 공간이 결합된 스키마 안에서 서로 다르게 처리를 하는 방법이 있다. 그렇지만 이 방법들은 MBR을 이용하지 않는다. 세 번째 방법은 OP-tree[14]를 이용하는 방법인데 이 방법은 MBR을 이용하는 방법과 유사하나

이동객체의 궤적의 근사치 영역을 만들기 위해 MBR의 각 모서리를 제거해서 팔각형 형태의 영역을 만들어 사용한다. 하지만 MBR을 사용하는 방법보다 더 좋은 성능을 보장하지는 못한다.

III. 선형병합 분할 알고리즘

3.1. EMBR

본 연구에서는 3차원인 이동 객체 궤적의 위치 정보를 간단하게 나타내기 위해서 궤적을 위치 정보 x 와 시간 t 의 2차원 공간으로 나타낸다.

EMBR면적의 계산은 (그림 1)과 같이 주어지는 범위 질의 qx, qt 를 MBR의 범위에서 $qx/2, qt/2$ 만큼 MBR에서 확장된 사각형 면적을 계산한다. EMBR은 범위 질의가 주어졌을 때 질의가 MBR과 중첩하는지 여부를 찾아내기 위하여 비교 검색해야 하는 부분이다.

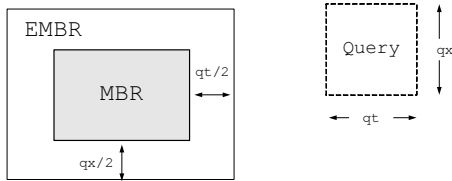


그림 1. 범위질의 크기를 고려하여 확장된 MBR
Fig 1. Extended MBR considered Size of range query

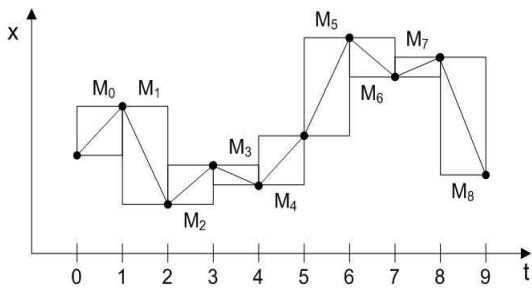


그림 2. 연속적인 두 개의 포인트를 이용하여 만든 새로운 MBR

Fig 2. New MBR using consecutive two points

3.2. 선형병합 분할 알고리즘

선형병합 분할 알고리즘은 (그림 4)와 같다. 이 방법은 첫 번째 과정 1) 서 (그림 2)와 같이 포인트 $P_0 \sim P_n$ 까지 MBR을 계산한 $M_0 \sim M_{n-1}$ 을 생성한 다음 두 번째 과정 2)에서 이웃한 두 EMBR의 값과 이 EMBR의 전체 값을 계산해서 $KEY_{(i,j)}$ 를 만든 후 차례대로 우선순위 큐에 삽입한다. 다음 세 번째 과정 3)에서 정렬된 우선순위 큐에서 첫 번째로 반환되는 키 값 $KEY_{(i,j)}$ 을 가지고 다음 이웃한 EMBR 값과 연산을 한 후 $LKEY_{(i,k)}$ 값이 더 이상 음의 수를 가지는 키 값이 더 이상 존재하지 않을 때까지 병합한 후 3)의 과정을 반복 수행한 후 (그림 3)과 같이 최종적인 EMBR 집합을 만들어 낸다.

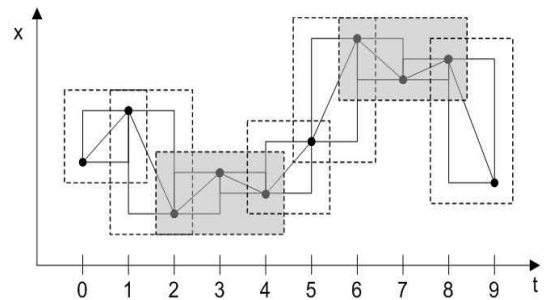


그림 3. 선형병합 분할 방법으로 구성된 최종 EMBR의 집합

Fig 3. Set of EMBR constituted the LinearMerge split algorithm

[정리] Algorithm-1의 시간 복잡도는 $O(n^2 \log n)$ 이다.

증명.

1)에서 궤적의 MBR 집합을 만들기 위한 연산 횟수는 n 이고, 2)에서 하나의 키 값을 계산하여 우선순위 큐에 삽입하는 연산 횟수는 $n + n \log n$ 이다. 3)에서 EMBR 집합의 병합과 새로운 우선순위 큐의 재구성을 위한 연산 횟수는 $n^2 \log n$ 이다. 그래서 알고리즘의 총 연산 횟수는 $n + n + n \log n + n^2 \log n$ 이다. 그러므로 이 알고리즘의 시간 복잡도는 $O(n^2 \log n)$ 이다.

Algorithm-1 : LinearMerge Split for a Trajectory

Input: A set of Trajectory's Points(P_0, P_1, \dots, P_n),
Query size.

Output: A set of MBRs that cover the Trajectory.

- 1) Construct full split MBRs $M_n[o, n]$
- 2) **For each** i ($0 \leq i \leq n-1$)
 Compute EMBRs volume for merging $M[i, i+1]$
 with $M[i+1, i+2]$
 store $KEY_{(i, i+2)}$ in a priority queue such
 that fallow holds:
 $KEY_{(i, i+2)} = E[i, i+2] - (E[i, i+1] + E[i+1, i+2])$
- 3) **While** ($KEY_{(i, j)}$ in the root node of the priority
 queue is negative)
While ($LKEY_{(i, k)}$ is negative)
 compute $LKEY_{(i, j+1)}$ that fallow holds:
 $LKEY_{(i, j+1)} = LE[i, j+1] -$
 $(LE[i, j] + LE[j, j+1])$
 Merge the pair of consecutive $M[i, k]$
 Update the priority queue with new merged MBRs;

그림 4. 선형병합 분할 알고리즘
Fig 4. LinearMerge split algorithm

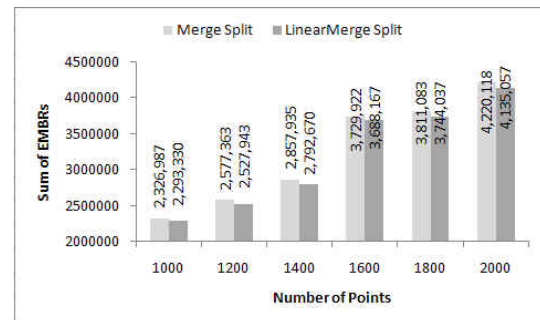
IV. 실험 및 성능 비교

본 논문에서 제안하는 이동 객체 궤적들에 대한 선형 병합 분할 방법의 성능실험은 인텔 i5 2.8GHz 프로세서와 메모리 1.9Gbyte, Windows XP 운영체제를 사용하는 시스템상에서 수행했으며, 알고리즘의 구현을 위해서 C++ 언어를 사용하였다. 그리고 성능 비교 평가를 위하여 기존의 병합분할 방법과 선형병합 분할 알고리즘을 구현하여 실험을 하였다. 본 실험에서 위치 축인 x 축의 총 범위는 0에서 1500으로 제한하고 실험에 사용되는 포인트의 개수는 1000에서 2000으로 제한하였다. 각 실험은 객관성을 위해 같은 조건상에서 임의의 생성하여 실험에 적용하였다.

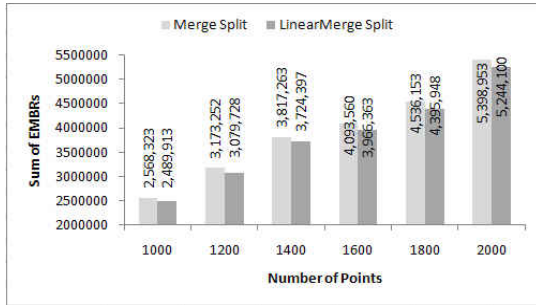
첫 번째 실험은 시간 구간을 30으로 규칙적인 시간의 간격에 따라 위치 변화의 정도는 각각 15, 30, 45로 한계를 설정하고, 평균 범위 질의의 크기를 30으로 고정하고

이동 객체의 포인트를 1000개에서 2000개까지 200개 단위로 생성하였다. 위의 조건을 통해 각 분할 방법으로 줄어드는 MBR집합의 EMBR 총면적(Sum of EMBRs)을 계산하여 비교하였다.

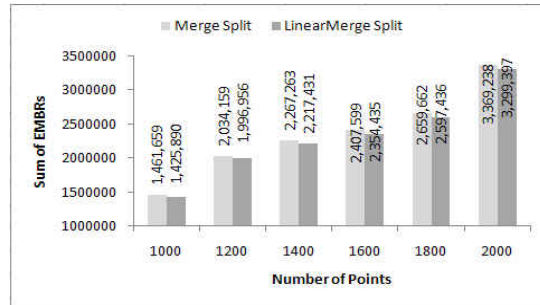
(그림 5-7)은 이동 객체의 위치 변화를 다양하게 변화하게 하여 병합 분할 알고리즘과 선형병합 분할 알고리즘을 적용하여 최종 EMBR의 총면적의 값을 도식화한 것이다. (그림 5)의 도표는 이동 객체의 좌표 (x, t) 에서 t 의 좌표는 35로 일정하게 변화시키고 이동 객체의 현 위치에서 다음 위치로의 x 좌표의 변화를 상하 15 이내로 제한하여 좌표를 생성시킨후 알고리즘을 적용하여 EMBR의 총면적을 계산 후 도식화시킨 그림이다. (그림 6)은 x 좌표의 변화를 30으로, (그림 7)은 x 좌표의 변화를 45로 각각 제한하여 좌표를 생성시킨 후 계산한 결과를 도식화한 것이다. 도표에서 보는 것 같이 두 알고리즘의 성능 비교 결과를 보면 이동 객체의 위치 변화가 클수록 본 논문에서 제안하는 선형병합 분할 방법을 이용해서 만들어지는 EMBR 총면적이 병합 분할 방법을 통해 만들어지는 EMBR 총면적보다 더 효율적으로 줄어드는 것을 알 수 있다. 이것은 인덱스 생성한 후 필요 정보를 탐색시 탐색하는 공간을 더욱 효율적으로 줄여 줄 수 있음을 의미한다.



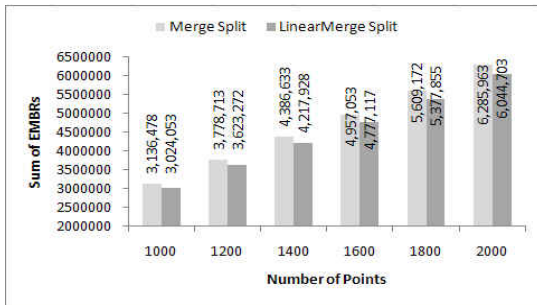
(위치변화 정도 15이내, 측정시간 구간 35)
그림 5. 위치정보에 대한 측정시간이 규칙적인 경우
(location change within 15, time interval 35)
Fig 5. The regular case of measurement time of location information



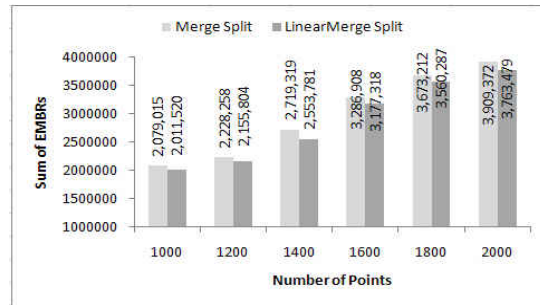
(위치변화 정도 30이내, 측정시간 구간 35)
그림 6. 위치정보에 대한 측정시간이 규칙적인 경우
(location change within 30, time interval 35)
Fig 6. The regular case of measurement time of location information



(위치변화 정도 15이내, 측정시간 구간 10-35)
그림 8. 위치정보에 대한 측정시간이 불규칙적인 경우
(location change within 15, time interval 10-35)
Fig 8. The irregular case of measurement time of location information



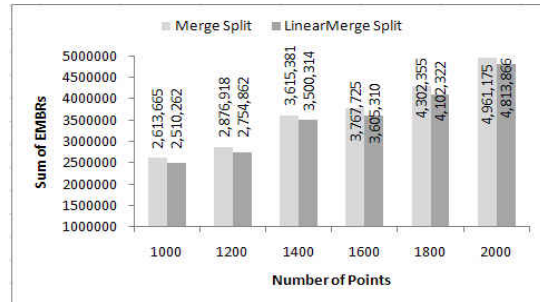
(위치변화 정도 45이내, 측정시간 구간 35)
그림 7. 위치정보에 대한 측정시간이 규칙적인 경우
(location change within 45, time interval 35)
Fig 7. The regular case of measurement time of location information



(위치변화 정도 30이내, 측정시간 구간 10-35)
그림 9. 위치정보에 대한 측정시간이 불규칙적인 경우
(location change within 30, time interval 10-35)
Fig 9. The irregular case of measurement time of location information

두 번째 실험은 첫 번째 실험과 달리 이동 객체의 좌표 (x, t) 에서 시간 구간의 좌표 이동 객체의 좌표 t 를 10-35 사이에서 불규칙적으로 생성하고 x 의 좌표를 위치 변화를 각각 상하 15, 30, 45로 제한하여 이동 객체의 위치 정보를 생성하여 실험을 했다. 그 외의 다른 조건은 첫 번째 실험과 동일한 조건을 주었다.

(그림 8-10)의 도표는 두 번째 실험을 통한 각 궤적 분할 방법의 성능 비교 결과이다. 첫 번째 실험과 마찬가지로 제안하는 선형병합 분할 방법을 통해 만들어지는 EMBR 총면적이 병합 분할 방법의 EMBR의 총면적보다 더 효율적으로 줄어들음을 알 수 있다.



(위치변화 정도 45이내, 측정시간 구간 10-35)
그림 10. 위치정보에 대한 측정시간이 불규칙적인 경우
(location change within 45, time interval 10-35)
Fig 10. The irregular case of measurement time of location information

V. 결 론

본 논문에서는 검색 범위와 시간을 줄이기 위해 쿼리 정보를 효율적으로 줄이는 알고리즘을 제안하였다. 성능 평가의 결과를 보면 제안하는 쿼리 분할 알고리즘은 평균적인 질의의 크기를 이용한 EMBR 면적을 우선 고려하여 쿼리의 적절한 분할 위치를 찾아서 분할함으로써 공간 질의에 대한 탐색 영역을 줄임으로서 효율적인 데이터 처리를 할 수 있음을 알 수 있다. 그리고 이를 바탕으로 색인을 구성하면 색인 구성을 더 효율적으로 할 수 있고 다양한 공간 질의에 대해서 효율적인 질의 처리를 할 수 있음을 알 수 있다.

향후 논문 과제는 더욱 다양한 쿼리들에 대하여 검색 공간을 줄이고 색인 구성 과정에서 이동 객체 쿼리를 정보를 효율적으로 분할할 수 있는 알고리즘에 대한 실험을 수행할 예정이다.

참고문헌

- [1] D. Pfoser, C. S. Jensen and Y. Theodoridis, "Novel Approaches to the Indexing of Moving Object Trajectories", Proc of The 26th International Conference on Very Large Data Bases, pp. 395-406, 2000
- [2] X. Xu, J. Han and W. Lu, "RT-Tree: An Improved R-Tree Indexing Structure for Temporal Spatial Databases", Proc of The International Symposium on Spatial Data Handling, pp. 1040-1049, 1990
- [3] D. Pfoser, C. S. Jensen and Y. Theodoridis, "Novel Approaches to the Indexing of Moving Object Trajectories" Proc of The 26th International Conference on Very Large Data Bases. pp. 395-406, 2000
- [4] Y. Theodoridis, M. Vazirgiannis and T. Sellis, "Spatio-Temporal Indexing for Large Multimedia Application", Proc of The 3rd IEEE International Conference on Multimedia Computing and Systems, pp. 441-448, 1996
- [5] O. Wolfson, S. Chamberlain, S. Dao and L. Jiang, "Location Management in Moving Objects Databases", Proc of The International Workshop in Satellite-Based Information Services, pp. 7-14, 1997
- [6] T. Brinkhoff, "Generating Network-Based Moving Objects", Proc of The International Conference on Scientific and Statistical Database Management, pp. 253-255, 2000
- [7] R. Benetis, C. S. Jensen, G. Karciuskas and S. Saltenis, "Nearest Neighbor and Reverse Nearest Neighbor Queries of Moving Objects", Proc of The International Symposium on Database Technology, pp. 44-53, 2002
- [8] J. Nievergelt, H. Hinterberger and K. C. Sevcik, "The Grid File: An Adaptable, Symmetric Multikey File Structure", ACM Transactions on Database System, pp. 38-71, 1984.
- [9] Z. Song and N. Roussopoulos, "SEB-tree: An Approach to Index Continuously Moving Objects", Proc of The 4th International Conference on Mobile Data Management, pp. 340-344, 2003
- [10] T. Tzouramanis, M. Vassilakopoulos, "Overlapping linear quater: a spatio-temporal access method", Proc of the 6th International Symposium on Advances in Geographic Information System, pp. 1-7, 1998.
- [11] Guttam, "R-tree: A Dynamic Index Structure for Spatial Searching", Proc of the ACM International Conference on Management of Data, pp. 47-57, 1984.
- [12] V. P. Chakka, A. C. Everspaugh and J. M. Patel, "Indexing Large Trajectory Data Sets With SETI", Proc of The 1st Biennial Conference on Innovative Data Systems Research, 2003
- [13] S. Prabhakar, Y. Xia, D. V. Kalashnikov, W. G. Aref and S. E. Hanbrusch, "Query Indexing and Velocity Constrained Indexing: Scalable Techniques for Continuous on Computers", IEEE Transactions on Computers, pp. 1124-1140, 2002
- [14] Y. Tao, D. Papadias and J. Sun, "The TPR*-Tree: An Optimized Spatio-Temporal Access Method for Predicative Queries", Proc of the 29th International Conference on Very Large Databases, pp. 790-801, 2003
- [15] M. Hadjieleftheriou, G. Kollios, V. J. Tsotras, D.

Gunopulos, Efficient Indexing of Spatiotemporal Objects, Lecture Notes in Computer Science, pp. 251-268

- [16] H. Zhu, J. Su, O. Ibarra, "Trajectory Queries and Octagouse in Moving Object Databases", Proc of The ACM International Conference on Information and Knowledge Management, pp. 413-421, 2002

저자소개



박주현(Ju-Hyun Park)

2000년 경남대학교 컴퓨터공학과
(공학사)

2002년 경남대학교 대학원
컴퓨터공학과 (공학석사)

2007년 부경대학교 컴퓨터공학과(박사수료)

※ 관심분야: 공간 데이터베이스, 객체지향 데이터
베이스



조우현(Woo-Hyun Cho)

1985년 경북대학교 전자공학과
전산 공학전공(공학사)

1988년 경북대학교 대학원
전자공학과 (공학석사)

1998년 경북대학교 대학원 전자공학과
전산공학전공(공학박사)

1989년- 현재 부경대학교 공과대학교 컴퓨터공학부
교수

※ 관심분야: 지능형 데이터베이스, 멀티미디어
인덱싱, 객체 데이터베이스 관리 기술