

실세계 도로 네트워크 환경에서의 이동객체 패턴기반 분산 예측 프레임워크 설계

정재화*

요약

최근 모바일 스마트 기기의 보급으로 스마트 기기에 탑재된 다양한 센서에서 수집되는 대량의 데이터를 분석하여 처리하는 빅 데이터의 시대는 위치기반 서비스(LBSs: Location-Based Services)에 까지 확대되고 있다. 이동객체 위치 예측 기술은 차세대 시공간 정보 서비스에서 요구하는 필수적인 기술로 많은 연구가 진행되어 왔다. 그러나 대부분의 기존연구는 특정 어플리케이션에 종속적인 이동객체 위치 예측 질의처리 연구로서 증가하는 차세대 시공간 정보 서비스의 기술적 요구를 반영하기가 어렵다. 따라서 본 논문은 실세계 공간 네트워크에서 이동객체들의 초대용량 시공간적 데이터를 토대로 필수적으로 예측에 필요한 기본적이고 다양한 질의를 효율적으로 처리할 수 있는 범용적 분산 이동객체 예측 질의처리 프레임워크 설계를 제안한다.

키워드 : 시공간 데이터베이스, 분산 데이터 색인, 예측 질의처리 프레임워크

Design of Moving Object Pattern-based Distributed Prediction Framework in Real-World Road Networks

Jaehwa Chung*

Abstract

Recently, due to the proliferation of mobile smart devices, the innovation of bigdata, which analyzes and processes massive data collected from various sensors implanted in smart devices, expands to LBSs. Many location prediction techniques for moving objects have been studied in literature. However, as the majority of studies perform location prediction which depends on specific applications, they hardly reflect the technical requirements of next-generation spatio-temporal information services. Therefore, this paper proposes the design of general-purpose distributed moving object prediction query processing framework that is capable of performing primitive and various types of queries effectively based on massive spatio-temporal data of moving objects in real-world space networks.

Keywords : Spatio-temporal database, Distributed indexing, Prediction query processing framework

1. 서론

스마트기기, 스마트카의 대중화는 지능형 교통

체계(ITS), 물류관리 체계(LMS), 위치기반 서비스(LBS), 위치기반 광고(LBA) 등으로 대표되는 시공간 정보서비스 시장을 창출하였으며 이러한 양상은 향후 지속적으로 증대될 것으로 예상되고 있다. 이에 따라 시공간 정보서비스를 위해 주요 기반 기술로서 이동 객체(MOs: Moving Objects)의 위치 예측과 이동 객체 인덱싱 기술 등의 이동객체 데이터베이스에 대한 여러 연구가 진행되고 있다[1, 2].

기존의 지능형 시공간 정보 서비스를 지원하기 위한 도로 네트워크 기반의 예측 질의처리

※ 교신저자(Corresponding Author): Jaehwa Chung
접수일:2014년 08월 07일, 수정일:2014년 08월 27일
완료일:2014년 08월 31일

* 한국방송통신대학교 컴퓨터과학과

jaehwachung@knou.ac.kr

■ 이 논문은 2013학년도 후기 한국방송통신대학교 학술연구비 지원을 받아 작성된 것임.

프레임워크에 관한 연구들은 중앙집중식 환경에서 실세계 환경을 명확하게 반영하지 못하였으며 예측의 정확성을 보장하기 위해서는 고비용적인 예측 처리가 불가피하다.

정교하고 지능적인 시공간 정보 서비스를 지원하기 위해서는 실세계 도로 네트워크에 대한 고려가 필요하며 각각의 이동 객체들의 이동상황 데이터를 토대로 기존의 단순 거리 예측에서 벗어나 집합을 이루는 객체들이 이동하는 양상, 즉 객체 이동패턴 정보를 기반으로 예측 기능을 제공하는 범용적 이동객체 예측 질의처리 프레임워크에 대한 설계가 절실히 요구된다.

이를 위해, 본 논문에서 제안하는 이동객체 데이터상의 분산 예측 질의처리 프레임워크 설계는 첫째, 실세계 도로 네트워크 모델 설계, 둘째, 이동객체 데이터의 분산관리 모듈 설계, 셋째, 객체 이동상황 패턴 기반의 분산 예측 질의처리 모듈 설계, 넷째, 클라이언트 모듈 설계, 마지막으로 이들을 종합한 최종적인 프레임워크 설계이다.

본 논문의 구성은 다음과 같다.

2장은 기존 연구를 살펴보고 3장은 분산 예측 프레임워크 설계를 위해 다양한 모듈 설계에 대해 다루며, 마지막으로 연구 결과를 요약하고 향후 연구 방향을 설명한다.

2. 관련연구

도로 네트워크 모델에 기반한 연구로 Ralf Hartmut Guting 외[3]는 ADT를 확장하여 고정적인 도로 네트워크를 모델링하고 그 상에서의 이동객체의 위치를 상대적으로 표현하는 기법을 제시하였다. 또한 도로 네트워크의 동적 속성을 고려하였다. 그러나 실세계 도로 네트워크의 불균형적인 특성을 고려하지 않았으며, 이동객체를 도로 네트워크상의 상대적 위치로 표현하여, 실세계 도로에서 이동객체가 위치하는 차선 정보를 파악할 수 없어 예측의 정확도가 제한되는 단점이 있다. Lin Qi 외[4]는 도로 네트워크상의 이동객체를 모델링하고, 이동객체 데이터 접근을 위한 질의 언어를 개발하였다. 그러나 단일계층 그래프 도로 네트워크 모델로 동적으로 변하는 속성과 현실 세계의 도로 네트워크의 불균형적

인 구조를 고려하지 못했고, 차선과 같은 상세도로 정보를 표현하지 못한다.

이동객체 위치 예측 질의처리 기술로서 Jong-Dae Kim 외[5]는 도로 네트워크를 지리상의 좌표로 표현하고, 객체의 궤적을 세그먼트들의 집합으로 표현하는 라인 세그먼트 모델을 제안하였다. 이 연구는 출발지-현재 위치까지의 경로를 토대로 순차 검색하여 탐색된 유사 궤적을 분석하여 이동 경로를 예측한다. 그러나 데이터의 효과적 탐색하기 위한 색인구조가 없어 원하는 데이터를 찾기 위한 탐색 시간과 비용이 크기 때문에, 대규모 실세계 도로 네트워크로부터 생성된 대용량 데이터를 관리하기 어렵다. Hoyoung Jeung 외[6]는 단일 계층의 그래프 모델로 정의하여 교차로, 도로, 가중치 값 및 도로의 부가적인 속성정보를 표현 모델을 제안하였다. 객체의 위치는 도로 상의 상대적 거리와 시간으로 표현하며, 각 객체의 이동 패턴정보를 통계로 관리하고 그리드 공간상에서 각 셀 간의 이동에 걸리는 시간을 계산하여 인덱스를 구축하고 예측이 이루어진다. 그러나 도로 네트워크를 그리드로 분할시, 현실의 도로는 불균등하게 분포되어 대용량 데이터를 처리과정에 부하 불균형을 야기하여 분산처리에 부적합한 문제점이 있다. 또한 Gyozo Gidofalvi 외[7]는 그리드 공간상에서 객체의 과거 궤적을 분석하여 FP-tree로 관리하고 유사 궤적을 찾아 객체의 위치를 예측한다. 이것은 빈번하게 생성된 궤적을 기반으로 이동객체 위치 예측을 하여, 궤적이 밀집된 구간에만 질의처리가 유효하고 소규모 도로 네트워크에서의 범위 예측 질의에 한정된 연구로 다양한 어플리케이션의 지원에 어려움이 존재하는 문제점을 갖는다.

지금까지의 도로 네트워크에 대한 기존연구는 실세계 대규모 도로 네트워크에 부적합한 단순 구조 모델로 현실 세계에서의 불균형적인 도로 네트워크 구조에 대한 고려가 미흡하다. 또한 이동객체의 위치를 단순 도로 네트워크 모델 상의 상대적 위치로 표현함으로써 다각도적인 정보 활용도가 저하되어 이동객체 데이터 분석을 통한 다양한 예측 지원에 한계를 지닌다. 뿐만 아니라 중앙 집중식 예측 질의 처리를 수행하여 정확한 예측정보를 제공하기 위한 성능을 확보하는 것이 매우 어렵다는 단점이 있다.

따라서 본 논문에서는 지능형 시공간 정보 서비스는 실세계 도로 네트워크에서 이동객체 데이터 분석을 지원하기 위해 어플리케이션에 독립적이고 객체 이동상황 패턴 기반 분산된 예측 질의처리 프레임워크 설계를 제안한다.

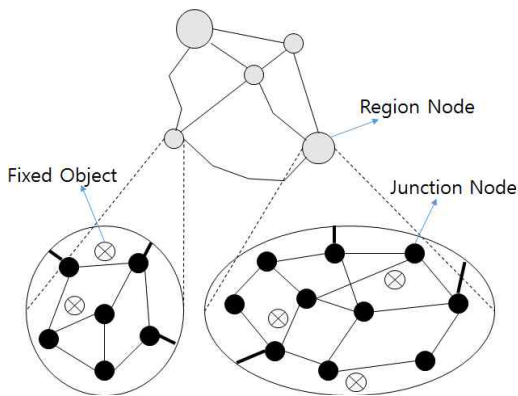
3. 시스템 설계

기존연구의 경우 실세계 대규모 도로 네트워크에 부적합한 모델 사용, 대용량 이동객체 처리가 어려운 중앙 집중식 예측 질의처리, 객체 이동상황을 반영하지 못한 예측 질의, 어플리케이션 종속적 예측 질의라는 다수의 문제점을 가지고 있기에 다음과 같은 세부 모듈 설계를 진행하여 본 연구의 목표인 분산 예측 프레임워크 설계를 진행한다.

3.1 실세계 도로 네트워크 모델 설계

다양한 시공간 정보서비스에서 다루는 도로 네트워크 및 이동객체 데이터를 각각 통합적으로 표현하기 위하여 현실 세계의 대규모 도로 네트워크 특성들을 분석하고 해당 특성들을 효과적으로 표현할 수 있는 동적 도로 네트워크 모델을 설계해야 한다.

(그림 1) 도로 네트워크 모델



(Figure 1) Road Networks Model

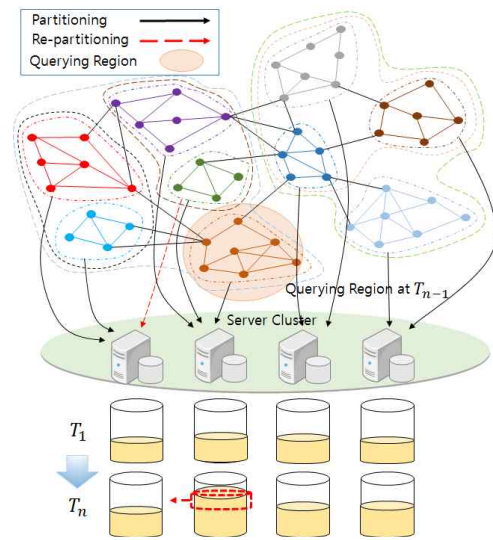
그래프 구조 기반의 도로 네트워크는 노드(node)와 간선(edge)이 공간 속성을 가지기 때문에 공간적 불균형 특성과 도로의 상황에 표현을 위한 동적 속성을 고려한 그래프 구조를 설계

한다. 이를 위해 현실 세계의 대규모 도로 네트워크 특징 및 속성에 대한 분석을 진행하여 차후 데이터 분산적 관리에 용이한 네트워크 표현을 위해 노드(region node)와 간선의 유형을 계층적 구조 기반으로 정의하며 그 구조는 (그림 1)과 같다.

3.2 대용량 데이터 분산관리 모듈 설계

기존 대부분의 분산 데이터 관리 시스템은 대용량 데이터에 대해 동일한 데이터 크기로 분할하여 다수의 서버 노드들에서 관리한다.

(그림 2) 분산관리 기법



(Figure 2) Distributed Management Technique

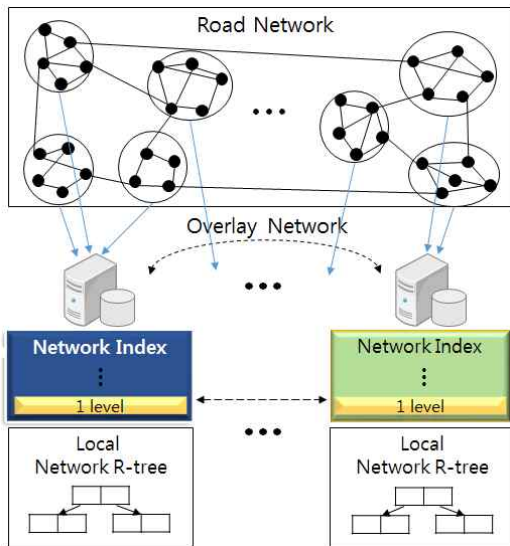
이러한 데이터 크기 관점의 단순 분할 기법은 구현이 쉽지만 시공간적 속성을 가진 이동·고정 객체 데이터에 대해 지역성을 보장하지 못하여 시공간적 특성을 기반으로 하는 데이터 처리 시 서버 노드들 간의 네트워크 비용을 크게 증가시켜 전체 분산 처리 성능을 저하시킨다. 또한 실세계의 불균형적인 데이터 분포로 인해 데이터 집중 발생 지역을 담당하는 서버 노드는 데이터 양 증가 속도가 다른 노드들에 비해 크기 때문에 노드들 간의 불균형 문제가 야기되어 전체 분산 처리 성능을 저하시킨다.

제안하는 분산 관리 모듈에서는 데이터양 및 지역성을 고려한 분할 기법 연구를 위해 상향식 접근 전략을 적용한 (그림 2)와 같이 계층적 도

로 네트워크의 분할 및 재분할을 진행한다.

또한, 이것을 바탕으로 대용량 이동객체 데이터 상에서 예측 질의처리의 성능 향상을 위해 서버 클러스터 환경을 위한 도로 네트워크 기반의 분산 인덱스를 설계한다. 기존 이동객체 예측에 대한 인덱싱 관련 연구들은 도로 네트워크를 균일 분포의 자유공간에 적합한 그리드(grid)에 매핑하는 인덱싱 연구가 대부분으로 불균형적 특성을 갖는 현실 세계의 도로 네트워크 환경에 부적합하다. 따라서 본 연구는 대용량 이동·고정 객체 데이터에 대한 관리 및 처리를 위해 (그림 3)과 같이 다수의 서버 노드 환경에 적합한 계층적 도로 네트워크 기반의 분산 인덱스 구조를 설계한다.

(그림 3) 분산색인



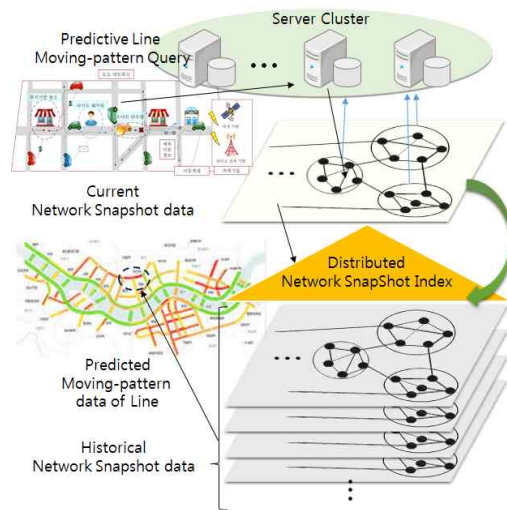
(Figure 3) Distributed Indexes

본 연구에서 제안하는 분산 인덱스의 설계는 분할된 각 서버 노드의 데이터를 관리하는 로컬 인덱스와 서버들 간의 연결 관계를 관리하는 네트워크 인덱스로 구분하고 이를 위해 각 서버 노드들의 분할 및 클러스터된 그래프의 연결 관계를 분석하여 오버레이 네트워크 기반의 인덱스 구조를 갖는다.

3.3 객체 이동상황 패턴 기반의 분산 예측 질의처리 모듈 설계

도로 네트워크에서 이동객체 예측 질의는 질의(예측 시간, 선분, 범위 등)로 주어지는 미래의 한 시점에 선분이나 범위 안에 포함될 이동객체들을 검색하는 것이 목적이다. 기존 대부분의 이동객체 예측 질의처리 연구는 주어진 선분이나 범위 안에 포함될 이동객체를 예측하기 위해 각각의 이동객체의 현재 이동정보(e.g. 속도 및 방향 등) 또는 이동궤적에만 의존한 예측하기 때문에 먼 미래 시점에 대한 예측이 어렵거나 예측의 정확도를 높이기 위해 각각의 모든 이동객체 데이터에 대한 이동경로 예측을 계산해야하기 때문에 고비용적 질의처리가 불가피하다.

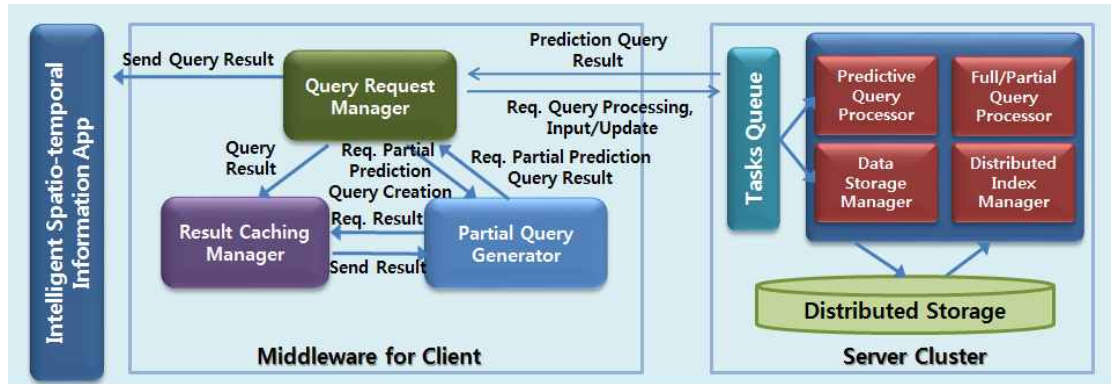
(그림 4) 분산질의 처리 기법



(Figure 4) Distributed Query Processing

본 연구는 도로의 객체 이동상황 예측을 위해 (그림 4)에서와 같이 해당 예측 질의의 공간 속성(i.e. 그림의 예로는 선분)을 기반으로 네트워크 거리(upper bound moving-pattern distance)를 계산하고 이를 기반으로 대용량 네트워크 스냅샷 데이터의 검색 공간(search space) 축소를 위한 시공간적 필터링 기법을 이용하여 질의처리가 효율적으로 진행할 수 있는 모듈을 제안한다.

(그림 5) 클라이언트 모듈의 프로세스



(Figure 5) Processes in Client Modules

또한, 선분 및 범위 객체 이동상황 패턴 예측 질의에 대해 후보 네트워크 스냅샷 데이터 집합에서 질의의 예측 시간을 바탕으로 현재 도로 네트워크 스냅샷 데이터와 유사 객체 이동상황 패턴 데이터를 검색하는 모듈을 제안한다.

3.4 클라이언트 모듈 설계

실시간 연속적 지능형 시공간 정보 서비스 지원을 위한 예측 프레임워크는 서버 측 모듈 구성만으로는 어플리케이션 지원에 한계가 있기 때문에 프레임워크와 모바일 시공간 정보 어플리케이션의 중간 모듈이 필요하다. 따라서 비교적 모바일 플랫폼 중에서 개방적 특성을 가진 모바일 OS 플랫폼을 타겟으로 하여 클라이언트 측에 위치하는 모듈을 설계 한다. (그림 5)은 이동객체 예측 프레임워크에서 고려하고 있는 환경의 서버/클라이언트 아키텍처를 나타내고 있으며 클라이언트에 위치하는 세부 모듈과 그 관계를 나타낸다.

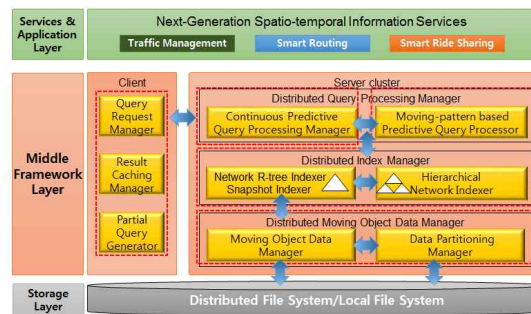
- **Query Request Manager:** 서버에 질의처리 및 데이터의 삽입 갱신 연산을 수행한다.
- **Result Caching Manager:** 서버로부터 전송 받은 이동객체 위치 예측 질의 결과 및 해당 질의 결과의 메타(meta) 정보를 클라이언트에 캐싱한다.
- **Partial Query Generator:** 전체(full) 또는 부분 이동객체 위치 예측 질의의 캐싱된 정

보를 바탕으로 갱신이 요구되는 이동객체들에 대한 부분 질의를 생성하는 역할을 담당한다.

3.5 이동 객체를 예측하는 프레임워크

본 연구는 다양한 시공간 정보 서비스를 지원하기 위한 이동객체 데이터베이스 기반 기술 연구로서, 구체적으로는 실세계 도로 네트워크에서 이동객체들의 위치 데이터를 토대로 다양한 시공간 정보 서비스 영역에서 활용 가능한 기본 예측 정보를 제공하는 범용적 이동객체 예측 질의처리 프레임워크는 (그림 6)과 같다.

(그림 6) 이동객체 예측 프레임워크



(Figure 6) Moving Object Prediction Framework

- **Distributed MOD Manager:** 분산 스토리지에서의 이동객체 데이터의 저장 관리 모듈

- **Distributed Index Manager:** 네트워크 기반의 비용-최적화된 데이터 관리 및 처리를 위한 분산 인덱스 구조 및 생성 모듈
- **Moving-pattern based Predictive Query Processor:** 객체 이동상황 패턴 기반의 이동 객체 위치 예측 질의처리 모듈
- **Continuous Predictive Query Processing Manager:** 연속적 예측 질의처리 모듈
- **Query Request Manager:** 클라이언트에서 데이터 삽입·갱신 및 질의처리 요청 관리 모듈
- **Result Caching Manager:** 예측 질의처리 결과 및 결과 메타데이터 캐싱 관리 모듈
- **Partial Query Manager:** 실시간 지속적인 지능형 서비스 지원을 위한 부분 질의 생성 모듈

4. 결론

본 논문은 지능형 시공간 서비스를 효율적으로 지원하기 위한 분산 예측 프레임워크 설계를 제안하였다. 복잡한 예측 질의처리에서 저비용/고성능의 목표를 달성하기 위해 데이터 모델, 색인, 질의처리 모듈로 나누어 시스템 설계를 진행하였다. 또한 부가적으로 클라이언트에서의 모듈들을 설계함으로써 전체 시스템에 실시간적인 성능 향상을 할 수 있게 설계를 진행하였다. 향후 본 시스템 설계를 오픈소스 기반의 하둡 시스템으로 구현하고 다양한 이동객체 데이터 환경에서의 비교 실험을 통해 본 시스템의 우수성을 증명하고자 한다.

References

[1] Jaehwa Chung, "Design of Trajectory Data Indexing and Query Processing for Real-Time LBS in MapReduce Environments", Journal of Digital Contents Society, 14(3), 313-321, 2013

[2] Jong-Ho Kim, "Finding Critical Success Factors for Spatial Data Industry by Comparing Strategies of Digital Earth Enterprises", Journal of Digital Contents Society, 13(3), 318-329, 2013

[3] Ralf Hartmut guting et al, "Modeling and querying

moving objects in networks", VLDB Journal pp. 165-190, 2006

[4] Lin Qi et al, "MONET: Modeling and Querying Moving Objects in Spatial Networks", ACM SIGSPATIAL IWGS, 2012

[5] Jong-Dae Kim et al, "Path Prediction of Moving Objects on Road Networks", Proceedings of the Korea Information Processing Society Conference 2006 Nov. 10 pp.437-440

[6] Hoyoung Jeung et al, "Path prediction and predictive range querying in road network databases", VLDB Journal pp. 585-602, 2010

[7] Gyozo Gidofalvi et al, "Frequent Route Based Continuous Moving Object Location and Density Prediction on Road Networks", ACM SIGSPATIAL GIS, 2011.



정재화

1999년 : 고려대학교 대학원 (이학 석사)

2011년 : 고려대학교 대학원 (이학 박사-공간질의처리)

2011년~2012년: 고려대학교 정보창의연구소 연구교수

2012년~현재: 한국방송통신대학교 컴퓨터학과 조교수

관심분야 : 공간질의처리 및 인덱싱(Spatial Query Processing and Indexing), 모바일 데이터 관리(MDM), RFID, 무선 센서 네트워크(WSNs)