

스냅샷을 사용하는 다중 레벨 공간 DBMS 를 가지는 효율적인 센서 미들웨어 구조 설계

오은석*, 김호석*, 배해영*
*인하대학교 컴퓨터 정보공학과
e-mail : oes@dblabb.inha.ac.kr

Efficient Sensor Middleware Architecture on Multi Level Spatial DBMS with Snapshot

Eun-Seog Oh*, Ho-Seog Kim*, Hae-Young Bae*
*Dept. of Computer Science & Information Engineering, Inha University

요 약

최근 들어, 모든 장소에 컴퓨터가 있고 그것을 누구나 자유롭게 사용할 수 있는 인간 중심 유비쿼터스 환경에 대한 연구가 활발하게 진행되고 있다. 그러나 이러한 컴퓨팅 환경에서의 센서 미들웨어들은 연속적인 스트림 데이터 처리에 대한 막대한 비용을 줄이기 위해서 분석이 끝난 스트림 데이터를 삭제한다. 따라서 사용자의 서비스에 삭제된 스트림 데이터가 요구될 경우, 또는 과거 데이터들에 대한 확률 통계 정보가 요구될 경우, 이에 대한 서비스를 사용자에게 제공할 수 없다.

본 논문에서는 이러한 미들웨어에서의 스트림 데이터 재사용 문제를 해결하기 위해서 스냅샷을 가지는 다중 레벨 공간 DBMS 를 기반으로 하는 효율적인 센서 데이터 미들웨어 구조를 설계하였다. 본 시스템은 사용자에게 재사용될 가치가 높은 스트림 데이터들을 다중 레벨 공간 DBMS 의 디스크 데이터베이스에서 관리한다. 또한 빈번한 서비스를 가지는 스트림 데이터의 경우 해당 스트림 데이터를 메모리 데이터베이스에 관리함으로써 사용자에게 신속한 서비스를 제공한다.

본 시스템은 기존의 미들웨어에서 지원할 수 없었던 동일한 스트림 데이터에 대한 빈번한 요청, 또는 확률이나 통계 자료와 같은 데이터 서비스 요청 문제들을 해결하였다. 그리고 메인 메모리 데이터베이스에 스냅샷 형태로 저장되는 스트림 데이터의 높은 데이터 재사용성을 유지함으로써, 사용자에게 지속적으로 정확하고 신속한 데이터 서비스를 제공한다.

1. 서론

최근 5C(Computing, Communication, Connectivity, Contents, Calm) 또는 5Any(Any-time, Any-where, Any-network, Any-device, Any-service)를 지향하는 유비쿼터스 환경을 실현시키기 위한 많은 연구 개발이 진행되고 있고, 이를 위한 주요 기술로서 RFID 기술 및 RFID 데이터를 처리하는 소프트웨어인 RFID 미들웨어 기술이 주목을 받고 있다. RFID 미들웨어는 유비쿼터스 컴퓨팅 환경 실현을 위한 주요 시스템의 하나로서, RFID를 기반으로 유비쿼터스 환경에서 폭주하는 데이터들을 효율적으로 정제,

관리하는 시스템이다. RFID 미들웨어는 사용자가 필요로 하는 다양한 응용 서비스에 대응하기 위하여 사용자가 관심있는 정보를 적절한 위치에서 적절한 시간에 정확한 서비스를 제공하기 위한 목적을 가지며, 이를 위해서 처리해야 할 데이터의 양, 데이터 형식, 그리고 동시에 처리되는 필터링 조건의 수 등의 조건에 대응하여 다양한 필터링 과정을 실시간으로 수행한다.

그러나 기존의 이러한 미들웨어들은 스트림 데이터에 대한 재사용 문제를 고려하지 않는 이유로 다음과 같은 두 가지 문제점을 가지고 있다. 첫 번째로, 사용자가 동일한 스트림 데이터를 이용한 서비스를 빈번하게 요청할 경우, 미들웨어는 정적인 스트림 데이터 저장 장소를 가지지 않기 때문에

*본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT 연구센터 육성, 지원사업의 연구결과로 수행되었음.

동일한 트리거 연산을 반복 수행해야 하는 문제를 가진다. 두 번째로, 사용자가 과거 스트림 데이터들을 이용한 확률, 통계 자료들을 필요로 할 경우, 미들웨어는 이를 서비스하기 위한 상당한 데이터 누적 시간이 필요하거나, 스트림 데이터 처리 부하 문제로 인하여 해당 서비스를 제공하지 않는다는 문제점을 가진다. 따라서 미들웨어는 현재 처리되는 스트림 데이터 이외에, 사용자의 의사 결정이나 빈번한 서비스 요청에 대응할 수 있는 정적 저장 공간이 필요하다[1].

본 논문에서는 이러한 문제점들을 해결하기 위해서 스냅샷을 가지는 다중 레벨 공간 DBMS 를 가지는 효율적인 센서 미들웨어 구조를 설계하였다. 본 시스템에서는 미들웨어가 지정된 트리거 조건을 만족하는 스트림 데이터를 발견할 경우, 해당하는 스트림 데이터를 스냅샷을 가지는 다중 레벨 공간 DBMS 의 디스크 데이터베이스에 저장한다. 그리고 일부의 재사용 빈도가 높은 스트림 데이터의 경우, 해당하는 스트림 데이터를 스냅샷의 형태로 메인 메모리 데이터베이스에 저장한다 [2]. 또한, 주기적으로 저장된 스냅샷이 가지는 재사용 가치를 분석함으로써 재사용 빈도가 높은 스트림 데이터 정보가 메인 메모리 데이터베이스에 유지하도록 한다.

본 논문의 구성은 다음과 같다. 2 장은 관련 연구로서 기존의 미들웨어 구조와 스냅샷을 가지는 다중 레벨 공간 DBMS 에 대해 설명하고, 3 장에서는 본 시스템의 구조를 설명하고 스트림 데이터에 대한 효율적인 저장 및 관리 구조를 설명한다. 마지막으로 4 장에서는 결론 및 향후 연구 방향을 제시한다.

2. 관련 연구

본 장에서는 본 논문의 기반이 되는 RFID 센서 미들웨어와 스냅샷을 가지는 다중 레벨 공간 DBMS에 대하여 기술한다.

2.1 RFID 미들웨어

RFID 미들웨어는 다양한 응용 서비스에서 사용되는 RFID 를 적절한 장소와 적절한 시간에 응용 서비스로 전달하는 구조를 가진다. RFID 미들웨어는 RFID 기반의 유비쿼터스 환경의 응용 서비스를 지원하며 끊임없이 지속적으로 입력되는 데이터를 실시간으로 정확하게 처리하고 응용 서비스에서 요구하는 결과를 획득해서 전달한다[3, 4]. 이러한 처리 과정을 위해서 미들웨어는 RFID 센서 데이터 처리를 지원하기 위한 다양한 관리 도구, 센서로부터 들어온 데이터 서비스를 지원하기 위한 센서 데이터 관리자, 입력된 센서 데이터를 분석하고 원하는 내용을 얻기 위한 매핑 래퍼 과정을 지원하는 센서 데이터 스트림 관리자로 구성된다. 미들웨어는 이러한 RFID 플랫폼을 사용하여 백엔드 어플리케이션으로부터 요구되는 요구 사항에 효율적으로 대응한다. 그러나 이러한 기존의 미들웨어는 정적인 저장 장소가 없기 때문에 사용자의 동일한 서비스의 빈번한 요청, 또는 과거 스트림 데이터를 이용한 확률, 통계 등의

서비스 요청을 지원하지 못한다.

2.2 스냅샷을 사용하는 다중 레벨 공간 DBMS

스냅샷을 사용하는 다중 레벨 공간 DBMS 는 대용량 데이터, 그리고 일부의 빠른 응답 속도를 요구하는 데이터들을 효율적으로 관리하기 위한 DBMS 이다[5]. 스냅샷을 사용하는 다중 레벨 공간 DBMS 는 상대적으로 낮은 사용 빈도를 가지는 대용량 공간 데이터를 관리하기 위한 GMS 기반의 디스크 저장 계층을 가지며, 높은 사용빈도와 가까운 미래에 높은 재사용 가능성을 가진다고 판단되는 일부의 디스크 계층 데이터를 스냅샷의 형태로 메인 메모리 계층에서 관리한다[6, 7]. 또한 계층 간의 데이터가 가지는 일관성을 유지시키기 위해서, 갱신된 디스크 계층의 데이터 값을 메모리 계층에 저장된 스냅샷 테이블 데이터를 주기적으로 갱신한다. 그리고 메모리 계층의 스냅샷 데이터가 가지는 중요도를 평가하여 사용가치가 상대적으로 낮은 스냅샷 데이터를 삭제한다. 스냅샷을 사용하는 다중 레벨 공간 DBMS 는 스냅샷 데이터의 저장 가치를 평가하기 위해서 데이터의 접근 빈도수, 중요도, 빠른 응답 속도 등을 고려하고 데이터를 갱신하며 갱신 후에 발생하는 데이터베이스 계층 간의 데이터 일관성을 위해서 데이터 동기화 과정을 가진다.

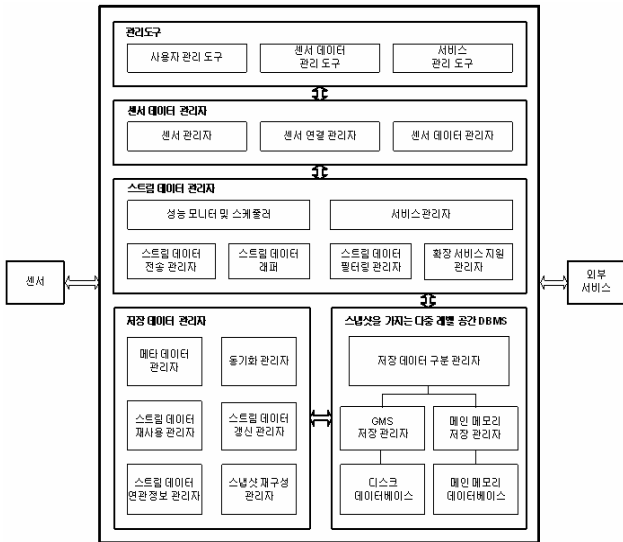
3. 스냅샷을 사용하는 다중 레벨 공간 DBMS를 가지는 효율적인 센서 미들웨어 구조

본 장에서는 스냅샷을 사용하는 다중 레벨 공간 DBMS 와 센서 미들웨어를 연동하여 확률이나 통계 정보와 같은 과거 스트림 데이터를 필요로 하는 사용자에게 DBMS 를 이용하여 해당하는 데이터 정보를 효율적으로 서비스하는 센서 미들웨어 구조를 설명한다. 본 시스템은 스냅샷을 사용하는 다중 레벨 공간 DBMS 의 디스크 데이터베이스에 기본적인 필터링 과정을 거친 스트림 데이터를 저장하여 과거 스트림 데이터에 대한 사용자 요청 문제를 해결하고, 사용자에게 대한 데이터 재사용 빈도를 가지는 스트림 데이터를 메인 메모리 데이터베이스에서 관리하여 는 시스템 구조를 가진다. 본 논문에서 제안하는 센서 미들웨어 구조는 (그림 1)과 같다.

센서 데이터 관리자는 센서의 종류와 연결 상태를 관리하는 센서 연결 관리자, 센서의 종류를 관리하는 센서 관리자, 그리고 센서로부터 입력받는 데이터를 관리하는 센서 데이터 관리자로 구성된다. 그리고 스트림 데이터 관리자는 센서로부터 얻은 데이터를 미들웨어에서 처리할 수 있는 데이터 형태로 변환하는 스트림 데이터 래퍼, 지정된 트리거 조건을 만족하지 않는 스트림 데이터를 걸러내는 스트림 데이터 필터링 관리자, 다중 레벨 공간 DBMS 에 데이터를 전송하는 스트림 데이터 전송 관리자, 그리고 확률 통계나 누적된 스트림 데이터를 필요로 하는 서비스를 위하여 다중 레벨 공간 DBMS 에 해당하는 데이터를 요청하는 확장 서비스 지원 관리자 등으로 구성된다. 본 시스템은 이러한 스트림

데이터 처리 콤포넌트를 사용하여 미들웨어에서 지정된 트리거 조건을 만족하는 스트림 데이터를 스냅샷을 가지는 다중 레벨 공간 DBMS 의 디스크 데이터베이스에 저장하고, 사용자가 해당 스트림 데이터에 관한 서비스를 요청할 경우 이에 대한 서비스를 제공한다. 그리고 디스크 데이터베이스에 저장된 스트림 데이터 중에서 사용자가 빈번하게 요청하는 스트림 데이터를 메인 메모리 데이터베이스에 저장함으로써 사용자의 요청에 대해 보다 신속한 데이터 서비스를 제공한다. 또한 본 시스템은 메인 메모리 데이터베이스에 저장된 스트림 데이터에 대한 높은 데이터 재사용성을 유지하기 위해서 주기적으로 저장된 스트림 데이터를 분석하고 더 이상 필요로 하지 않는 데이터에 대한 삭제 연산을 수행한다.

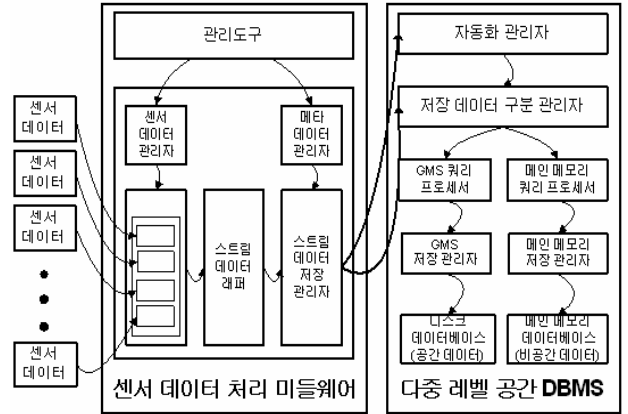
요구하는 스트림 데이터를 메모리 기반 데이터베이스 계층으로, 그리고 빠른 응답 속도보다는 대용량의 과거 데이터 정보, 또는 누적 데이터 통계 정보 등을 디스크 기반 데이터베이스 계층에 저장한다. 아래의 (그림 2)는 이러한 스트림 데이터의 저장 과정을 나타낸다.



(그림 1) 센서 미들웨어 구조

3.1 스트림 데이터의 저장

미들웨어는 외부로부터 끊임없이 연속적으로 들어오는 정보 중 응용 서비스가 관심있어 하는 스트림 데이터 만을 가공할 수 있는 기술이 필요하다. 이를 위해서 미들웨어로부터 수집된 스트림 데이터는 중복이나 위조 등의 기본적인 데이터 필터링 과정을 거치고, 백엔드 어플리케이션에서의 필요 여부가 분석된 후, 미들웨어에서 지정된 트리거를 만족한 스트림 데이터들을 디스크 레벨의 데이터베이스에 저장된다. 또한, 저장된 디스크 기반 데이터베이스의 데이터에서 상대적으로 높은 데이터 우선 순위를 가지는 일부 데이터들은 스냅샷의 형태로 메모리 데이터베이스에 저장된다. 분석 과정에 사용되는 데이터 우선 순위는 DBMS 의 메타 데이터에 정의된 기준에 따른 처리 과정을 거친다. 데이터를 전달 받은 다중 레벨 공간 DBMS 는 시스템 내에서 높은 재사용 가치를 가지는 스트림 데이터들의 저장 위치를 판단한다. 시스템은 그 판단 기준에 따라 백엔드 어플리케이션에게 높은 사용빈도, 처리 속도를



(그림 2) 센서 스트림 데이터 저장 구조

시스템은 센서로부터 얻은 스트림 데이터를 스트림 데이터 레퍼를 통해 필요로 하는 데이터를 얻는다. 그리고 변환된 데이터에 대한 저장 가치를 분석하는 과정은 데이터를 임시 저장하고 분석하는 데이터 스트림 저장 관리자에 의해서 수행되며, 기존의 저장된 메타 데이터 정보와 데이터 관리는 관리도구에서 지원한다. 분석 결과, 미들웨어에서 저장 가치가 높다고 판단된 데이터는 다중 레벨 공간 DBMS로 보내지고 저장 데이터 고유의 특성에 따라 메인 메모리와 디스크 기반의 데이터베이스로 구분되어 각각의 데이터 특성에 맞는 데이터베이스로 저장된다. 저장할 데이터 데이터에 대한 분류 과정은 저장 데이터 구분 관리자에 의해 처리되며 일부의 정해진 작업은 자동화 관리자에 의하여 관리된다.

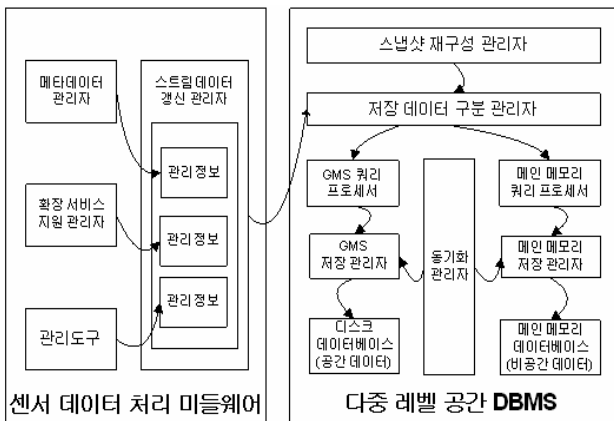
3.2 스트림 데이터의 관리

본 시스템은 사용자가 높은 빈도로 재사용하는 스트림 데이터를 관리하고 사용자가 데이터를 요청할 경우 이에 대한 신속한 스트림 데이터 서비스를 제공한다. 이러한 사용자 서비스를 제공하는 스트림 데이터에 대한 관리를 위해서 해당하는 스트림에 대한 사용자의 요청 횟수, 요청 시간, 서비스를 구성하는 데 기여하는 정도 등의 관리 정보를 수집하고, 이를 이용하여 데이터베이스에 저장된 스트림 데이터들에 대한 재사용 가치를 분석한다. 이러한 데이터의 저장 가치 판단에 사용되는 관리 정보는 확률, 또는 통계 자료와 같은 사용자의 과거 데이터 서비스 요청에 사용된다[8].

(그림 3)은 시스템의 미들웨어 계층에서 스트림 데이터의 재사용 가치를 판단하는 데 필요한 갱신 정보를 수집하고, 이를 이용하여 스냅샷을 가지는 다중 레벨 공간 DBMS 의 데이터베이스에 저장된 스트림 데이터를 재사용 가치에 따라 관리하는

시스템 구조를 나타낸다. 센서 데이터 처리 미들웨어 계층은 스트림 데이터 관리 정보 수집기, 메타 데이터 관리자, 스트림 데이터 저장 관리자, 그리고 관리 도구들로 구성된다. 메타 데이터 관리자는 사용자의 서비스 요청 빈도, 횟수, 최근 요청 시간 등과 같은 관리 정보 등을 수집하고, 확장 서비스 지원 관리자는 아직 메타 데이터에 반영되지 않았으며 데이터베이스에 저장된 스트림 데이터를 필요로 하는 서비스 정보를 수집한다. 스트림 데이터 갱신 관리자는 이러한 관리 정보들을 기반으로 스트림 데이터들에 대한 재사용 가치를 분석하고, 이를 다중 레벨 공간 DBMS 계층에 전달한다.

센서 데이터 처리 미들웨어 계층으로부터 스트림 데이터들에 대한 재사용 가치 정보를 전달받은 다중 레벨 공간 DBMS 계층에서는 전달받은 정보를 기준으로 디스크 데이터베이스에 저장된 스트림 데이터가 높은 재사용 가치를 가질 경우, 해당하는 스트림 데이터를 메인 메모리 데이터베이스에 스냅샷의 형태로 저장하고, 더 이상의 높은 재사용 가치를 가지지 않는 메인 메모리 스냅샷 데이터의 경우, 해당 스냅샷에 대한 삭제 작업을 수행한다[9]. 또한 갱신 과정에서 일어날 수 있는 데이터베이스 간의 데이터 일관성 문제를 동기화 관리자를 통하여 해결하며, 많은 수의 스냅샷에 대한 갱신 연산이나, 그 외 스냅샷 데이터 최적화, 재구축 등의 문제들은 스냅샷 재구성 관리자에서 지원한다.



(그림 3) 센서 스트림 데이터 관리 구조

4. 결론 및 향후 연구

본 논문에서는 사용자에게 서비스하기 위한 센서 스트림 데이터 처리 미들웨어에 보다 정확한 데이터 분석 및 제공을 위한 정적 저장소인 스냅샷을 사용하는 다중 레벨 공간 DBMS를 통합하기 위해, 스냅샷을 사용하는 다중 레벨 공간 DBMS를 가지는 효율적인 센서 미들웨어 구조를 설계하였다. 또한, 기존 미들웨어의 문제점이었던 별도의 시간 영역에 발생하는 데이터 정보, 또는 빈번하게 요청되는 대용량의 공간 데이터에 대한 중복 탐색 문제를 개선하고, 기존의 데이터베이스를 사용하면서 발생할 수 있는 미들웨어 내에서의 디스크 처리 비용, 제한된 메모리 공간 문제 등을 해결하였다.

본 논문의 향후 연구로는 저장 가치가 있는 스트림 데이터를 분석하기 위한 구체적인 메타 데이터, 그리고 미들웨어의 부하를 고려한 다중 레벨 공간 DBMS에서의 지능적인 갱신 알고리즘을 연구하는 것이다[10, 11].

5. 참고문헌

[1] Arvind Arasu, Brian Babcock, "STREAM: The Stanford Data Stream Management System," IEEE Data Engineering Bulletin, March 2003.

[2] Tobin J. Lehman, Eugene J. Shekita and Luis-Felipe Cabrera, "An Evaluation of Starburst's Memory Resident Storage Component", IEEE Transactions on Knowledge and Data Engineering, Vol. 4, No. 6, pp. 555 - 566, 1992.

[3] Christian Floerkemeier, Matthias Lampe, "Middleware: RFID middleware design: addressing application requirements and RFID constraints", In Proc of the 2005 joint Conf. on Smart objects and ambient intelligence: innovative context-aware services", pp. 219-224, 2005

[4] 정원일, "유비쿼터스 컴퓨팅을 위한 센서 데이터 스트림 처리 기술", 한국전자통신연구원, 2005

[5] 천종현, 어상훈, 김호석, 배해영, "스냅샷 데이터를 갖는 다중레벨 공간 DBMS 설계 및 구현", 추계학술발표회 논문집, Vol. 32, No. 02, pp. 217-219, 2005.

[6] James J. Lu, Guido Moerkotte, Joachim Schue and V. S. Subrahmanian, "Efficient maintenance of materialized mediated views", ACM SIGMOD Record, In Proc of the 1995 ACM SIGMOD international Conf. on Management of data, pp. 340 - 351, 1995.

[7] 박상근, 박순영, 정원일, "GMS: 공간 데이터베이스 관리 시스템", 2003 공동 춘계학술대회, Vol. 00, No. 00, pp. 217-224, 2003.

[8] W. Gilani, N.H. Naqvi, O. Spinczyk, "On adaptable middleware product lines", In Proc of the 3rd workshop on Adaptive and reflective middleware, pp. 207-213, 2004.

[9] 최종무, 조성제, 노삼혁, "적응력있는 블록 교체 기법을 위한 효율적인 버퍼 할당 정책", 정보과학회 논문지: 시스템 이론, Vol 27, No. 3, pp. 324-336, 2000.

[10] Manuel Roman, Nayeem Islam, "ACM International Conference Proceeding Series; Vol. 78, Proc of the 5th ACM/IFIP/USENIX international Conf. on Middleware", pp. 372-396, 2004.

[11] María Jesús Garzarán, Milos Prvulovic, José María Llabería, "Tradeoffs in buffering speculative memory state for thread-level speculation in multiprocessors", ACM Transactions on Architecture and Code Optimization (TACO), Vol 2, pp. 247-279, 2005.