

센서 모니터링 환경에서 캐쉬 테이블을 사용한 센서 스트림 데이터 고속 처리 시스템 설계

박준용[○], 이광원^{**}, 황윤철^{**}, 오염덕^{*}

^{○**} *충주대학교 컴퓨터정보공학과, *교신저자

e-mail: pjy1418@gmail.com, {kwlee,ychwang,rdohe}@cjnu.ac.kr

Design of High-speed Sensor Stream Data Processing System to use Cache Table base on Sensor Monitoring Environment

Junyong Park[○], Kwangwon Lee^{**}, Yunchul Hwang^{**}, Ryumduck Oh^{*}

^{○**} Dept. of Computer Science and Information Engineering Chungju National University

● 요약 ●

USN(Ubiquitous Sensor Network)기술이 다양한 분야에 적용하고 응용함으로써 센서 네트워크에서 발생하는 스트림 데이터를 효율적으로 처리 하는 기술이 활발하게 연구되고 있다. 그리고 센서들에서 발생한 이기종 센서 스트림 데이터는 미들웨어의 메타데이터를 이용하여 센서 모니터링 환경의 상위 응용시스템에서 사용하기 적합한 센서 데이터로의 변환이 필요하다. 기존에 개발된 센서 데이터 처리 시스템에서는 동일한 노드나 지역에서 발생하는 센서 스트림 데이터에 대해서도 미들웨어에서의 불필요한 검색과 연산을 수행하기 때문에 본 논문에서는 캐쉬 테이블방식을 이용하여 중복적인 센서 스트림 데이터 처리의 수행을 줄일 수 있는 센서 스트림 데이터 처리 시스템의 설계를 제안한다.

키워드: USN(Ubiquitous Sensor Network), 메타데이터(Metadata), 데이터 스트림(Data Stream)

1. 서론

최근 빠르게 발전하고 있는 USN(Ubiquitous Sensor Network) 기술은 모든 사물에 컴퓨팅 서비스를 내포하여 언제 어디서나 유용한 서비스를 받을 수 있도록 하는 것이며 이러한 서비스를 얻기 위해서는 다양한 센서들을 대량으로 분산시켜서 센서 네트워크를 구성하여 일정한 지역이나 환경을 모니터링 한다. 그리고 모니터링 환경에서는 다양한 종류의 센서가 사용되며 이러한 센서에서는 센서 데이터가 발생하게 된다. 이러한 센서 데이터들은 매우 다양하고 복잡하여 이질적인 센서 환경을 정의를 위해서 메타 데이터가 사용되고 있다. 기존의 TinyDB 같은 센서 네트워크 미들웨어는 데이터 처리가 센서 노드 내에서 센싱 데이터를 처리하는 방식이다. 이러한 방식은 센서 노드에서 불필요한 데이터 처리가 발생되고 센서 네트워크 내부에서 전송되는 데이터의 양이 증가시키는 원인이 된다. 그리고 데이터의 양과 센서노드의 데이터 처리의 양이 늘어나게 되면서 센서 데이터의 수명이 감소하는 문제가 발생하기 때문에 현재 구축되고 있는 센서네트워크 시스템에서는 센서노드의 정보를 매우 적은 양만을 처리하고 추가정보는 서버 측면의 미

들웨어의 데이터베이스에 저장되어 있는 메타데이터를 이용하여 센서 데이터를 처리 하는 방식을 제안하고 있다. 하지만 센서 스트림 데이터들은 노드에 대한 정보가 최소로 적체된 패킷을 보내기 때문에 미들웨어에서는 센싱 타입과 하드웨어 스펙등에 대한 데이터를 이용하여 상위 응용시스템에서 사용하기 적합한 데이터 형태로 만들어 주는 절차가 요구된다. 최근 센서 네트워크 미들웨어 시스템에서는 센서 스트림 데이터를 처리하는 경우에는 데이터베이스 접근을 통하여 전체 메타 데이터에서 해당 센서 노드 데이터를 검색하고 해당 데이터를 이용하여 가공하는 형식과 절차로 구성된다. USN 센서 네트워크 환경에서는 특정한 공간이나 노드에서 동일한 이벤트가 주로 발생하기 때문에 불필요한 메타데이터의 검색과 처리가 발생하게 된다. 본 논문에서 제안하는 시스템은 캐쉬 테이블 방식을 이용해서 데이터 처리에 필요한 메타데이터를 구성하여 센서 데이터 처리 시스템의 성능을 개선하였다 [1,2].

본 논문에서는 센서 메타데이터와 스트림 데이터 처리 절차를 설명하고 캐쉬 테이블을 이용한 센서 스트림 데이터 처리시스템의 구성과 그에 따른 메타 데이터의 구조 그리고 센서 스트림 데이터의 처리기법을 제안하였다.

II. 관련 연구

1. 센서 메타데이터

USN 환경의 센서들은 다양한 데이터들을 수집하여 미들웨어로 전송하게 된다. 이렇게 수집된 데이터는 원시적인 스트림 형태의 데이터로 사용자가 이용할 수 있는 정보가 되기 위해서는 센서가 만들어진 제조사나 기종에 따라 다양한 형태의 메타 데이터와 결합하여 사용자가 사용할 수 있는 데이터로 변환하는 작업이 수반된다. 이러한 효과적인 결합작업을 위해서는 표준화된 데이터의 구조의 필요성이 제기되었다. 그래서 국제표준 기구인 OGC(Open Geospatial Consortium)에서는 센서의 표준화된 형태의 데이터를 제공하기 위해서 센서에서 제공하는 서비스를 정의하고 하위레벨의 미들웨어에서 상위 레벨의 사용자 어플리케이션으로 데이터를 효과적으로 전달하기 위해서 SensorML(Sens or Mode 1 Language Encoding Standard) 기술을 제안하였다.

SensorML은 측정데이터 뿐만 아니라 센서와 관련된 높은 수준의 정보를 얻을 수 있는 정보를 얻을 수 있는 명령을 포함하는 표준모델과 XML 인코딩을 제공한다. 하지만 이러한 SensorML 기술은 다양한 센서의 데이터를 결합하기 매우 어렵기 때문에 고안되었지만 이러한 SensorML은 센서가 아닌 기능적인 모델을 제공하는데 목적을 가지고 있기 때문에 센서에서 들어오는 원시적인 데이터를 처리하는 하부 미들웨어에서는 적합하지 않는 메타데이터의 형태를 가지고 있다[3].

2. 스트림 데이터 관리

스트림 데이터에 관한 연구는 현재 국제적으로 데이터베이스 분야에서는 많은 관심을 가지고 활발한 연구가 진행되는 분야이다. 이러한 관심은 기존의 비즈니스 영역에서의 DBMS의 역할의 효용성이 인터넷의 발전을 통한 다른 영역의 필요성으로까지 인식되어졌기 때문이다. 이러한 스트림 데이터의 처리 기술은 USN 환경에서 센서에서 발생 되는 정보가 스트림 형식으로 서버로 전달된다고 볼 수 있기 때문에 스트림 데이터 처리와 USN 환경은 매우 밀접한 관계가 있다. 그리고 USN 환경에서는 네트워크의 수명이 배터리로 유지되며 최대한 오랜 시간동안 센서 네트워크의 수명을 유지해야하는 문제점이 있다. 그래서 센서 스트림 데이터는 매우 소량의 원시 데이터만을 전송하여 에너지 사용의 효율성을 높이고 미들웨어에서는 센서네트워크에서 들어온 원시적인 스트림 데이터를 메타데이터를 이용하여 상위 응용프로그램이 요구하는 데이터로 변환하기 위해서 미들웨어에서는 데이터베이스의 검색과 연산의 데이터베이스 접근이 필요하다[4,5].

본 논문에서는 센서 모니터링 환경에서 원시 센서 스트림 데이터를 처리 하는데 가장 최적의 메타데이터를 이용하여 센서 데이터를 처리하기 위한 모델링 방식을 제안하였다.

III. 캐쉬 기법을 적용한 센서 스트림 데이터 처리 시스템

1. 센서 스트림 데이터 처리 시스템의 구성

USN의 모니터링 환경에서 발생하는 이기종의 센서 데이터들은 기본적인 센서 메타데이터만을 적재하여 패킷의 형태로 미들웨어로 전송하게 된다. 이렇게 수집된 데이터들은 상위 어플리케이션에서 사용할 수 없는 특정한 수치의 연속이다. 이러한 정보들은 미리 정의된 센서 네트워크의 메타데이터에 의해서 상위 응용프로그램에서 요구하는 센서 데이터로 변환하여 정보를 제공하는 센서 스트림 데이터 처리 시스템이 필요하며 이러한 시스템은 이기종의 센서 스트림 데이터를 처리하기 위해서 센서 네트워크와 연결하는 추상화된 인터페이스와 센서 네트워크의 정보를 저장하고 있는 메타데이터베이스를 가진다.

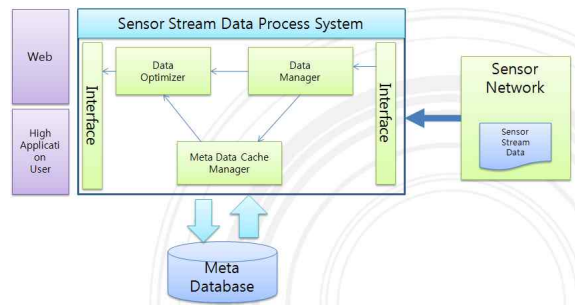


그림 1. 센서 스트림 데이터처리 시스템

본 논문에서 제안하는 센서 스트림 데이터 처리 시스템의 구조는 센서 네트워크에서 들어오는 패킷 형태의 스트림 데이터는 Data Manager에서 센서 노드에서 센싱되어진 정보들과 센서 노드ID나 네트워크 ID등과 같은 메타데이터를 분류해낸다. 그리고 분류한 센서노드 ID 메타데이터는 MDC(Meta Data Cache) Manager로 제공되어진다. 그리고 MDC Manager는 제공된 정보를 통해서 MDC(Meta Data Cache) 테이블을 구성한다. MDC 테이블은 지속적으로 삽입과 삭제를 진행하면서 센서 네트워크에서 반복적으로 사용되는 노드의 정보를 구성하게 된다. MDC 테이블에서 구성된 정보는 Data Optimizer에 전달된다. Data Optimizer의 역할은 MDC 테이블에 구성된 정보를 이용하여 센서노드에서 보내진 데이터와 조합하여 센싱 데이터를 최적화된 형태로 변환하고 메타데이터와 센싱 데이터를 종합하여 하나의 노드의 객체 정보를 구성하여 상위 어플리케이션에 보내기 위해서 인터페이스로 전송하는 역할을 수행한다.

2. 메타 데이터 구성

Data Manager에서 구성되는 메타데이터의 정보는 센서 스트림 데이터를 최적화 할 수 있는 메타데이터로 아래 표1과 같이 구성하게 된다.

표 1. 메타데이터의 구성

항목	설명
SensorNetwork ID	센서 네트워크망의 고유 번호
SensorNetwork Location	센서 네트워크 설치 위치
SensorNetwork Admin	센서 네트워크 관리자 정보
SensorNetwork BuildTime	센서 네트워크 구축 시기
SensorNetwork Protocol	센서 네트워크 구성 통신규격
SensorNodeID	센서 노드의 고유번호
SensorNode Location	센서 노드의 설치 위치
SensorNodeType	센서 노드의 센싱 데이터 타입
SensorNodeTime	센서 노드에서 데이터 측정시간
SensorNode HWSpec	센서 노드의 제조사 스펙

Data Manager에서 추출된 메타데이터는 센서 노드의 ID와 같은 유일한 항목을 통해서 캐쉬 테이블에 해당 노드가 존재 여부를 판단하는데 사용하고 추가적인 항목의 경우는 서버 미들웨어의 데이터베이스에 저장되게 하지만 본 시스템에서는 캐쉬 테이블에 기존에 사용되었던 노드의 정보를 저장하고 있는 상태에서 해당 노드 ID 정보가 들어 왔을 때 노드ID 정보를 이용하여 기존의 메타데이터와 조합하게 되면 상위 어플리케이션에 제공할 수 있는 노드의 객체의 정보를 가지게 된다. 이러한 데이터들의 메타데이터의 효율적 관리와 설계를 위해서 XML(Extensible Markup Language)을 이용하여 모델링하였다.

3. 센서 스트림 데이터 처리

본 논문에서 제안하는 MDC Manager는 센서 메타데이터의 캐쉬를 구성하는 부분으로서 테이블 형태의 힙 저장구조를 가진다. 그리고 메타데이터 캐쉬 구성 방식은 가장 일반적인 LRU(Least Recently Used) 방식을 이용하여 구성을 하게 된다. LRU 방식은 가장 최근에 사용된 데이터가 가장 상위의 테이블에 구성되어 데이터의 검색이 가장 먼저 검색되는 방식이며 LRU를 사용하였을 경우 동일한 노드에서 중복적으로 발생하는 이벤트에 대한 데이터 처리에서 높은 효율성을 나타내는 방식이다.

```

Algorithm Meta_Data_Handler_for_Sensor_Stream_Data

Data Transfer: Sensor Network로부터 센서 스트림 데이터를 Interface로 전달
Read Sensor Stream Data : Interface로부터 읽은 스트림 데이터 분류
Call MDC Manager by Data Manager: 메타데이터 전달 Cache Table에서 동일 메타 데이터 검색

if(Cache Table 에 Meta Data Set의 존재)
    Call Data Optimizer: 메타데이터 조합을 최적화
    Cache Table Reconstruction : Heap 저장구조에서 최근 사용한 메타데이터로 활용하기 위해 최상위 Data set으로 설정
else
    Request Meta Data set : 캐쉬 테이블에 해당 메타데이터 Set이 존재 하지 않기 때문에 메타 데이터베이스 에 해당 Data Set 요청
    Call Data Optimizer: 메타데이터 조합을 최적화
    Cache Table Reconstruction : Heap 저장구조에서 최근 사용한 메타데이터로 활용하기 위해 최상위 Data set으로 설정
    
```

그림 2. 캐쉬 테이블을 통한 센서 스트림데이터 처리

3.1 MDC(Meta Data Cache)의 삽입

Data Optimizer에서 요청하는 메타데이터가 캐쉬 테이블에 존재 하지 않을 경우 Meta Data를 데이터베이스에서 요청하여 캐쉬 테이블에 구성한다. 힙 형태의 저장구조 이기 때문에 가장 오랜 시간동안 사용되지 않은 메타 데이터는 삭제한다.

3.2 MDC(Meta Data Cache)의 유지

Data Optimizer에서 요청한 메타데이터가 존재 하는 경우에는 사용한 메타데이터의 테이블을 저장공간의 가장 상위 위치에 이동 저장하여 테이블에서 삭제되는 확률을 낮게 한다.

본 시스템은 센서 노드에서 센싱이 일어 날 때 이벤트 기반의 모니터링 환경의 경우는 이전에 일어났던 동일한 노드에서 이벤트 발생확률이 가장 높기 때문에 센서 메타데이터의 구성을 Cache Table로 구성할 경우 미들웨어에서 데이터베이스의 접근과 연산을 최소한으로 줄이는 결과를 주어 데이터 처리 시스템에서 높은 성능을 가질 수 있다.

표 2. 제안된 센서 스트림 데이터 처리 시스템의 특성

항목	기존시스템	제안 시스템
반복 이벤트 데이터 처리	중복된 데이터베이스 접근	캐쉬 테이블 데이터 접근
데이터 처리 시간	데이터베이스 접근 시간 증가	데이터베이스 접근 시간 감소
메타데이터 저장	데이터베이스에 직접 저장	캐쉬 테이블에 임시 저장

본 센서 데이터 처리 시스템에서는 스트림 센서 데이터 처리과정에서 발생하는 불필요한 데이터베이스의 접근과 연산의 중복 문제를 줄이도록 시스템을 설계하였다.

IV. 결론

본 논문에서는 USN 환경에서 다양한 이기종의 센서 네트워크에서 발생하는 센서 스트림데이터를 효율적으로 처리하기 위해 센서 네트워크에서 수집되는 최소한의 메타데이터를 이용하여 서버 미들웨어에서 반복적으로 사용되는 메타데이터의 테이블을 구성하였다. 메타 데이터 캐쉬 테이블 미들웨어에서 데이터베이스의 접근을 최소화 시켜 데이터 처리속도를 높이는 시스템을 설계하였다. 현재는 메타데이터를 LRU 방식을 적용하여 메타데이터 캐쉬를 구성하였다. 또한 효과 적인 획득과 저장 관리를 위한 메타데이터 관리 시스템 구현할 예정이다.

참고문헌

- [1] I.F. Akyildiz, W. Su*, Y. Sankarasubramaniam, E. Cayirci. "A survey on sensor networks". Computer Networks, 2002.
- [2] S.R. Madden, M.J. Franklin, and J.M. Hellerstein, TinyDB: An Acquisitional Query Processing System for Sensor Networks, ACM TODS, Vol.30,No.1,2005, pp.122-173
- [3] Golab, L., Ozler, M. T.: Issues in Data Stream Management. In:SIGMOD Record 32(2) (2003) 5-14
- [4] Robin, A., Havens, S., Cox, S., Ricker, J., Lake, R., Niedzwiatek, H.: OpenGIS c sensormodel language (SensorML) implementation specification. Technical report, Open Geospatial Consortium Inc (2006)
- [5] Kawashima H., Hirota Y., Satake S., and Imai M. Met: A Real World Oriented Metadata Management System for Semantic Sensor Networks. 3rd International Workshop on Data