

다차원 스트림 데이터 환경에서의 효율적인 데이터 수집 기법

김재인*, 황부현*

*전남대학교 전자컴퓨터공학과

e-mail:sereno3@naver.com

A Method for Efficiently Collecting Data from Multiple Data Streams

Jae-In Kim*, Bu-Hyun Hwang*

*Department of Computer Science, Chonnam National University

요 약

USN 환경에서의 센서는 일반적으로 많은 제약사항을 가지고 있다. 센서의 제한된 전원의 문제는 센서의 동작 수명과 관련된 것으로 최근의 연구들에서 중요 이슈가 되고 있다. 본 논문에서는 고도화된 USN 환경에서 발생하는 다차원 스트림데이터를 수집하는데 있어서 센서의 전원 문제를 해결하고 데이터를 효율적으로 수집하기 위한 기법을 제안한다. 제안하는 기법은 센서에 이상 이벤트를 정의하고 이상 이벤트에 해당하는 데이터를 수집하는 경우에만 데이터를 전송하도록 하여 센서의 통신 빈도를 줄여 센서의 전원 문제를 해결하고 스트림 데이터를 기호화 하여 처리함으로써 스트림 데이터를 효율적으로 수집할 수 있다.

1. 서 론

최근 네트워크 기술의 발전과 하드웨어의 발전으로 언제 어디서나 컴퓨팅 환경을 제공하는 유비쿼터스 기술에 대한 연구가 진행되고 있다. 특히 유비쿼터스 기술 중 저전력, 저비용, 초소형의 센서를 무선 네트워크로 연결하여 지속적이고 연속적으로 데이터를 수집하여 유용한 서비스를 제공하는 USN(Ubiquitous Sensor Network)기술에 관한 연구가 활발하게 진행되고 있다[1]. USN의 기술은 그 응용 범위가 매우 다양하다. 예를 들어 군사 분야에서는 위험 지역의 위험 요소를 효과적으로 감지하여 아군의 손실을 최소화하는데 이용되며, 화재나 홍수 등의 재난지역에서 사상자의 위치를 파악하여 신속한 구조를 돕는데 이용된다. 또한 바이오센서를 이용하는 U-Health 시스템에서는 독거노인이나 만성질환 환자에게 실시간으로 의료 서비스를 제공하며 응급 상황 발생시 신속하게 구조할 수 있도록 한다[1].

센서의 기술 발달로 USN의 응용 분야는 더욱 고도해지고 있으며 복잡한 컴퓨팅 기술을 요구하고 있다. 이러한 요구로 센서의 제한된 전원을 효율적으로 사용하면서 응용의 목적에 맞는 동작을 실시할 수 있도록 하는 연구들이 진행 중이다[2,3,4]. 또한 고도화된 USN의 응용은 단일 센서의 스트림데이터 뿐만 아니라 각기 다른 종류의 센서로부터 발생하는 다차원 스트림데이터를 효율적으로 처리하기 위한 기법을 요구하고 있다.

본 논문에서는 다차원 스트림 데이터 환경에서 센서들

의 한정된 전원의 소모를 최소화하여 센서의 동작 수명을 최대화하고, 다차원 스트림데이터에 대한 효율적인 데이터 수집을 위하여 이상 이벤트를 정의 하고 이상 이벤트에만 해당하는 데이터만을 선별하여 베이스 노드로 전송하는 기법을 제안한다. 본 논문의 구성은 다음과 같다. 논문의 2장에서는 관련연구로 센서의 전원 소모를 줄이는 연구에 대하여 알아보고, 3장에서는 제안하는 기법을 기술한다. 4장에서는 제안한 기법을 적용한 환경 모니터링 시스템을 시뮬레이션 하여 그 적합성을 확인하며 5장에서는 결론을 맺고 향후 연구를 기술한다.

2. 관련 연구

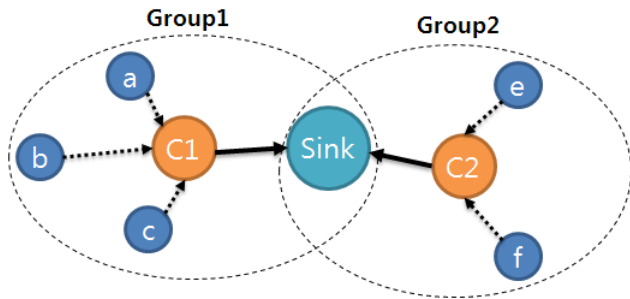
본 절에서는 센서의 전원 소모를 줄이기 위해 제안된 관련 연구에 대하여 알아본다. 전원 소모를 줄이기 위한 방법은 무선 송수신기의 전원 차단, 무선 송수신기의 출력 제어, 에너지 효율을 높인 USN 프로토콜 사용 등이 있다 [2,3,4].

송수신기의 출력을 제어하는 방법은 무선 송수신기가 출력하는 신호의 세기를 제어하는 방법으로 신호의 세기를 줄이게 되면 센서의 전원 소모가 감소하게 된다. 또한 송수신 출력의 세기를 줄이게 되면 송수신 거리가 줄어들어 노드간의 신호 간섭이 줄어들어 장점을 보인다. 이러한 출력 제어를 위한 다양한 방법은 다음과 같이 분류 할 수 있다[5].

첫째, 송수신기의 물리계층인 RSSI(Received Signal

Strength Indicator), SIR(Signal to Interference Ratio), BER(Bit Error Rate) 등을 참조하여 신호의 세기를 조절하는 방법이다. 수신측의 패킷 수신율에 따라서 신호의 세기를 조절하는 PCBL(Power Control with Black Listing) 알고리즘은 물리계층의 정보를 이용하는 예이다. 두 번째 방법은 하나의 네트워크를 구성하여 데이터를 송수신하는 센서들의 신호를 센서들 간 통신의 장애가 발생하지 않는 최소 범위로 모두 같게 조절하는 CPC(Common Power Control) 방법이다. 세 번째 방법은 수신 노드가 수신한 신호의 세기를 기준으로 자신의 송신 신호를 조절하는 방법(open loop)과 수신 노드가 수신 신호의 세기를 송신 노드에게 전송하여 송신 노드의 신호를 조절하도록 하는 방법(closed loop)이 있다. PCSMAC(Power Controlled Sensor MAC) 프로토콜은 수신 노드가 수신 데이터의 RSSI 정보를 송신 노드에게 전송하게 하여 신호의 세기를 조절하게 하는 closed loop 방식을 사용한다[6].

이러한 송수신기의 출력 제어 방법은 다양한 USN 프로토콜의 개발을 가능하게 하였는데, 대표적인 예로 센서들 간의 신호의 세기를 효율적으로 관리하여 동작 수명을 연장하는 LEACH(Low Energy Adaptive Clustering Hierarchy) 프로토콜이 있다. LEACH 프로토콜은 클러스터 기반으로 동작하는 USN 라우팅 기법으로 클러스터의 헤드가 클러스터 내에 있는 센서 노들로부터 데이터를 전송받아 병합하고 싱크노드로 전송하는 기법이다[7].



(그림 1) LEACH 프로토콜

그림 1은 LEACH 프로토콜의 동작 예를 나타낸 것으로 Group1은 센서노드인 a, b, c를 포함하는 하나의 클러스터이고 클러스터의 헤더로 C1을 포함하고 있다. 클러스터의 헤더의 전송은 랜덤하게 이루어지며 클러스터 내의 센서들이 헤더로 데이터를 전송할 때 신호의 세기를 최소화 하도록 하여 전체적인 센서의 전원 소모를 효율적으로 할 수 있다. 효율적인 전력소모를 위한 LEACH 프로토콜뿐만 아니라 위치 기반 프로토콜로 정보가 필요한 특정 지역으로 데이터를 송수신함으로써 전원의 소모를 줄이는 GEAR(Geographic and Energy Aware Routing), LAR(Location-Aware Routing)[8] 등이 있다.

이와 같은 신호의 세기를 조절하여 센서의 동작 수명을 최대화 하는 방법들은 USN 응용 시스템의 효율적인

동작을 가능하게 한다. 하지만 실세계에서 발생하는 데이터가 대부분 정상적인 상태의 데이터이고 이상 상태의 데이터를 모니터링 하고자하는 응용 분야에서는 앞서 기술한 방법들과는 다른 관점의 연구가 필요하다. 본 논문에서는 센서노드에서 정상적인 데이터와 비정상적인 데이터를 선별하여 비정상적인 데이터만을 전송하는 기법을 제안한다. 이러한 비정상적인 데이터를 선출하기 위하여 센서노드에는 이상 이벤트를 정의하고 이러한 이상 이벤트에 해당하는 경우에만 데이터를 전송함으로써 센서의 송수신 횟수를 줄여 전원 소모를 줄일 수가 있다.

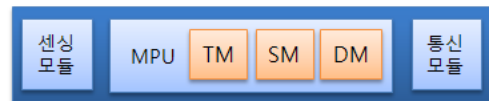
3. 효율적인 데이터 수집

본 절에서는 센서의 전원의 소모를 줄이고 효율적으로 데이터를 수집하는 기법을 제안한다. 특히 단일 센서가 아닌 다양한 센서로부터 수집되는 다차원 스트림데이터 환경에서의 데이터 수집을 고려한다.

다차원 스트림데이터는 실세계를 모니터링 하는데 있어서 더욱 정확한 모니터링을 실시한다. 예를 들어 지하 건물 내부의 화재를 감지하기 위한 USN 응용에서는 온도 센서만을 이용해서 화재를 감지하는 것보다 온도와 습도, 조도 센서를 이용하여 더욱 정확하게 화재를 감지할 수 있다. 온도가 높고 습도가 낮으며 빛이 지하 공간에서 조도가 높다면 이는 화재로 간주할 수 있고 적절한 대응을 실시 할 수 있다. 이처럼 다차원 스트림데이터는 실세계의 다양한 응용을 위한 것으로 다차원 스트림데이터를 효율적으로 수집하는 방법은 의미 있는 연구이다.

본 논문에서 제안하는 센서의 전원 소모 최소화는 센서노드에 대한 이상 이벤트 정의를 실시함으로써 해결하고자 한다. 이상 이벤트란 실세계의 데이터 중 정상적인 상태의 데이터가 아닌 모든 데이터로 볼 수 있는데, 이러한 이상 이벤트가 발생하는 경우에만 데이터를 선별하여 센서노드가 데이터를 전송함으로써 통신의 빈도를 낮춰 궁극적으로 센서의 전원 소모를 최소화 하는 것이다.

3.1 센서노드의 구조



(그림 2) 센서노드의 구조

센서노드는 일반적으로 센서와 통신모듈, 그리고 이들을 제어하는 MPU로 구성되어 있다. 본 논문에서는 이상 이벤트를 정의하고 이를 처리하기 위하여 MPU 내부에 TM(Tokenize Module), SM(Symbolize Module), DM(Detect Module)를 정의한다.

TM은 센서를 통해 수집되는 스트림데이터를 일정한 크기(시간, 데이터의 양)로 나눈다. 이때 나뉘지는 크기는 센서의 센싱 주기에 따라서 다양하게 정의 될 수 있으며 일정한 크기로 나누어진 스트림데이터의 조각은 SM으로 전달된다. SM은 미리 정의된 기호에 따라 스트림데이터를 분석하여 스트림데이터의 조각을 기호화 한다. 기호화 작업은 연속 시퀀스 특징을 가지는 스트림데이터를 요약하는 작업으로 방대하게 발생하는 데이터를 요약하기 때문에 통신모듈을 통해 전송되는 데이터의 패킷의 크기를 줄여주고 통신의 빈도를 낮출 수 있는 여건을 마련한다. SM을 거쳐 생성된 기호화된 요약 스트림데이터는 DM 모듈로 전달된다. DM 모듈에는 시스템의 목적에 맞는 이상 이벤트가 정의되어 있고 SM으로부터 생성된 요약 데이터가 이상 이벤트에 해당 하는지 여부를 판단한다. 만약 요약 데이터가 이상 이벤트에 해당되면 이 데이터를 통신 모듈을 통하여 베이스 노드로 데이터를 전송한다.

TM, SM, DM을 거친 이상 이벤트에 대한 데이터는 베이스노드로 전송되어 서버측에서 적절한 동작한 실시하게 된다. 이러한 이상 이벤트에 기반한 데이터 전송은 센서노드의 통신 빈도를 낮추어 센서 노드의 전원 소모를 최소화 하게 되고 다양한 사이트로부터 발생하는 데이터를 수집해야하는 베이스 노드의 처리비용을 줄여주게 되어 패킷 손실, 시스템 다운과 같은 문제를 해결한다.

3.2 다차원 스트림데이터 환경의 적용

그림 3은 다차원 스트림 환경에서의 효율적인 수집 기법을 적용한 예이다. 실세계를 모니터링 하는 하나의 센서 노드에 3개의 센서, s1, s2, s3가 존재한다고 할 때 센서 노드의 MPU는 각각의 센서에 대한 데이터에 대하여 이상 이벤트를 검출한다. 이러한 구조는 환경을 모니터링 하는 센서인 온도, 습도, 조도 센서를 한 번에 처리 하는 센서 노드에 적용될 수 있다.

각 센서의 스트림데이터는 센싱 주기에 기반하여 TM에 의해 토큰화 된다. 토큰화된 데이터는 각각 SM에 정의된 기호에 의해 데이터 값에 따라 기호화 된다. 예를 들어 온도 데이터를 기호화하기 위하여 SM은 표 1과 같이 구성할 수 있다.

<표 1> 온도 데이터 기호화 예

	온도	기호
1	5℃ 이하	VL
2	5℃ ~ 15℃	L
3	15℃ ~ 25℃	N1
4	25℃ ~ 30℃	N2
5	30℃ 이상	VH

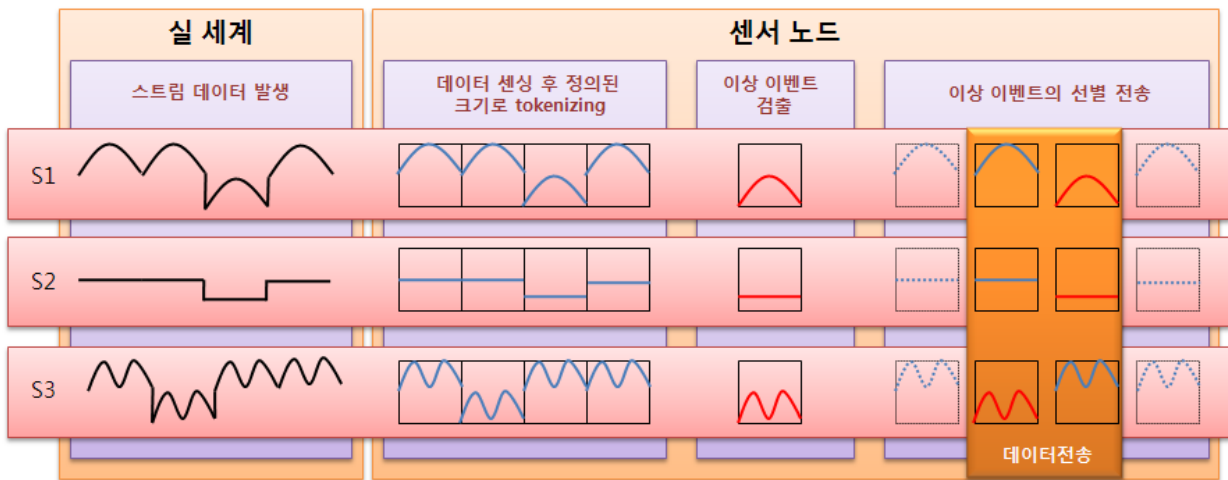
SM에 의하여 기호화된 각 센서의 스트림데이터는 통합되어 하나의 문자열로 기호화 된다. 예를 들어 온도, 습도, 조도는 센서의 ID와 발생 시간 등이 포함된 그림 4와 같이 기호화 될 수 있다.

센서ID 발생시간 온도:낮음 습도:정상 조도:정상

023 000000015 T 0L H N1 L N2

(그림 4) 기호화된 요약정보의 예

기호화된 스트림데이터는 PM에서 이상 이벤트에 해당 하는지 여부를 판단 받게 된다. PM은 정상적인 데이터를 정의 해놓고 그 데이터에 해당하지 않는 경우를 모두 이상 데이터로 정의 할 수 있다. 특히 다차원 스트림데이터 환경에서는 한 개 이상의 센서에서라도 이상 이벤트가 발생하게 되면 전체 요약 데이터를 전송한다. 그림 3의 이상 이벤트 전송처럼 정상 적인 데이터는 전송하지 않고 한 개 이상의 센서에서 이상 이벤트가 발생되면 전체 데이터를 전송한다.

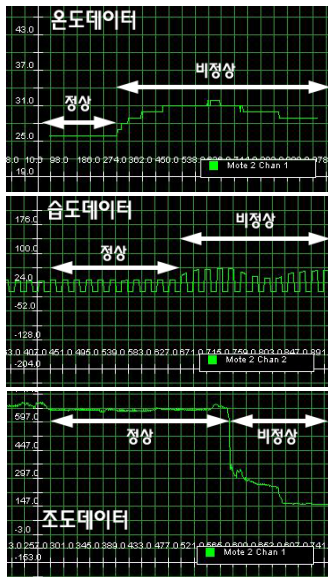


(그림 3). 다차원 스트림 환경에서의 데이터 전송

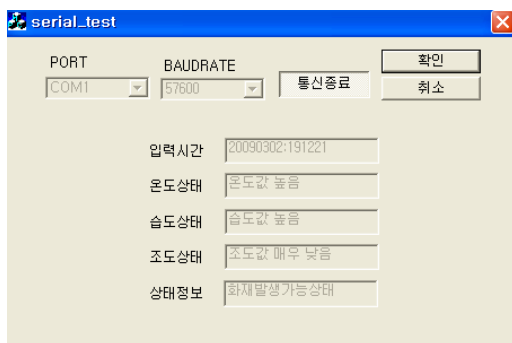
이상 이벤트가 발생한 경우에 데이터를 전송하는 기법은 센서 노드의 데이터 전송 빈도를 낮추지만 보다 효율적인 스트림데이터 수집을 위하여 정상적인 상태에서도 정의된 임계시간(ϵ)의 주기로 정상 데이터를 전송해야 한다. 이러한 주기적인 정상 데이터의 전송을 통해 센서의 올바른 동작여부를 확인 할 수 있다.

4. 시뮬레이션

본 절에서는 다차원 스트림 데이터의 효율적 수집 기법을 적용한 환경 모니터링 시스템을 시뮬레이션 한다. 사용된 센서는 한백전자의 ZigbeX 모드를 사용하였으며 TinyOS 1.0 환경에서 구현하였다. 시스템 UI는 WindowXP환경의 VisualStudio 6.0를 통해 작성하였다. 그림 5는 센서로부터 획득되는 모든 스트림데이터를 출력한 것이고 그림 6은 스트림데이터를 요약하고 이상 이벤트를 검출한 UI이다.



(그림 5) 입력 스트림 데이터



(그림 6) 이상이벤트 수집 서버 UI

5. 결론 및 향후연구

본 논문에서는 다차원 스트림 환경에서의 효율적인 스트림데이터 처리를 위하여 이상 이벤트 발생시에만 데이터를 전송하는 기법을 제안하였다. 이상 이벤트에 해당하는 데이터는 기호화된 요약정보 형태로 처리되어 베이스 노드로 전송된다. 이러한 기법은 센서노드의 통신 빈도를 낮추어 전원의 소모를 최소화 하고 다양한 사이트에서 다량의 데이터를 수집하는 베이스 노드의 처리 비용을 줄여 주는 장점을 갖는다. 그리고 제안하는 기법을 온도, 습도 센서에 적용하여 적합성을 확인 하였다.

향후 연구로는 USN의 다양한 라우팅 프로토콜과의 연계를 통해 보다 효율적이고 안정적인 센서네트워크를 구성하는 연구를 실시하고 요약된 이상 이벤트 데이터들을 수집하여 실시간 데이터 마이닝을 실시하는 시스템을 개발하고자 한다.

참고문헌

- [1] I. F. Akyiliz, W.Su, Y.Sankarasubramaniam and E.Cayirci, "A survey on sensor networks," IEEE Communications Magazine, pp.102-114, 2002. 8.
- [2] K. Akkaya and M. Younis, "A Survey on Routing Protocols for Wireless Sensor Networks," Ad Hoc Networks, Vol.3, pp.325-349, 2005.
- [3] 엄홍식, 김건욱, "전송전력 최적화를 통한 센서네트워크의 효율적인 에너지관리에 대한 연구," 전자공학회 논문지, 제44권 CI편, 제3호, pp.37-42, 2007.5
- [4] 박재홍, 류경식, 김용득, "무선 센서 네트워크에서의 S-MAC 기반의 에너지 효율적인 MAC 프로토콜," 전자공학회 논문지, 제44권 CI편 제2호, pp.19-24, 2007.3.
- [5] P. C. Nar and E. Cayirci, "PCSMAC: A Power Controlled Sensor-MAC Protocol for Wireless Sensor Networks," Proc. of 2nd European Workshop on Wireless Sensor Networks, pp.81-92, 2005.1.
- [6] 박지원, 모상만, 정일용, 배용근, "에너지 효율적인 데이터 수집을 이용한 LEACH 기반 무선 센서 네트워크 수명 연장," 전자공학회 논문지, 제45권 CI편, 제3호, 2008.
- [7] W. B. heinzelman, A. P. Chandrakasan, and H. Balakrishana, "An Application-Specific Protocol Architecture for Wireless Microsensor NetWorks," IEEE Transactions on Wireless Communications, Vol.1, Issue.4, pp660-670, 2002.10.
- [8] Y. Yu, R. Govindan, and D. Estrin, "Geographical and Energy Aware Routing: a recursive data dissemination protocol for wireless sensor networks," Tech. Rep. UCLA/CSD-TR-01-0023, Computer Science Department, University of California at Los Angeles, 2001.5.