

# 데이터 중심 저장 기법을 위한 효율적인 센서 데이터 압축 기법

## Efficient Sensor Data Compression Algorithm for Data-Centric Storage

노규종\*, 여명호\*\*, 성동욱\*\*\*, 복경수\*\*\*\*, 신재룡\*\*\*\*\*, 유재수\*\*\*  
 삼성SDS 전자LCD인프라그룹\*, 국방과학연구소 제2기술연구본부\*\*,  
 충북대학교 정보통신공학과\*\*\*, 가인정보기술(주)\*\*\*\*, 광주보건대학\*\*\*\*\*

Kyu-Jong Roh(kyujong.roh@samsung.com)\*, Myung-Ho Yeo(myungho.yeo@gmail.com)\*\*,  
 Dong-Ook Seong(seong.do@gmail.com)\*\*\*, Kyoung-Soo Bok(ksbok@gainit.co.kr)\*\*\*\*,  
 Jae-Ryong Shin(shr@ghc.ac.kr)\*\*\*\*\*, Jae-Soo Yoo(yjs@chungbuk.ac.kr)\*\*\*

### 요약

데이터 중심 저장 기법은 센서 네트워크에서 측정값을 효율적으로 저장하고 관리하기 위한 대표적인 연구 중 하나이다. 이 기법은 센서가 센싱한 데이터를 특정 센서 노드로 보내 저장한다. 하지만, 측정값을 네트워크 내부의 특정 노드로 전송하는 과정에서 많은 에너지를 소모하는 문제점을 가지고 있다. 본 논문에서는 데이터 중심 저장 기법에서 데이터의 전송 비용을 줄이기 위한 새로운 센서 데이터 압축 기법을 제안한다. 제안하는 기법은 측정값의 안전영역(Safe Region) 범위를 설정하고, 측정값이 안전영역을 벗어나는 경우만 데이터를 전송한다. 이를 통해 데이터의 전송을 줄이고 네트워크의 수명을 연장 시킨다. 성능평가 결과, 기존 기법에 비해 제안하는 기법의 에너지 소모가 60% 감소하였다.

■ 중심어 : | 센서 네트워크 | 데이터 압축 | 데이터 중심 저장구조 |

### Abstract

Data-centric storage schemes(DCS) are one of representative researches that efficiently store and manage sensor readings in sensor nodes in the sensor networks. In DCS, a sensor sends the sensed data to a specific node in order to store them. However, it has a problem that sensor nodes consume a lot of energy for transmitting their readings to remote sensor node. In this paper, we propose a novel sensor data compression algorithm to reduce communication costs for DCS. The proposed algorithm establishes a safe region and transmits the sensed data only when current measurement is out of the safe region. As a result, the proposed algorithm extends network life time and reduces data transmission. It is shown through performance evaluation that our proposed algorithm reduces energy consumption by about 60% over the conventional algorithm.

■ keyword : | Sensor Network | Data Compression | Data Centric Storage |

## 1. 서론

최근 프로세서, 메모리 그리고 무선 통신 기술의 발

전으로 인해 낮은 가격, 저전력 소모, 다양한 기능을 갖춘 스마트 센서 노드들이 개발되었다. 센서 노드는 주변의 온도, 습도 등과 같은 유용한 정보를 측정하기 위

\* 이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No. 2009-0080279)과 교육과학기술부와 한국산업기술재단의 지역혁신인력양성사업으로 수행된 연구결과임

접수번호 : #100907-002

접수일자 : 2010년 09월 07일

심사완료일 : 2010년 10월 13일

교신저자 : 유재수, e-mail : yjs@chungbuk.ac.kr

한 센서와 무선 통신 장치로 구성된다. 센서 네트워크는 수 백 ~ 수 천 개의 센서 노드로 구성되며, 다양한 응용분야에서 사용된다. 예를 들어, 야생 동물의 거주지와 같은 사람이 쉽게 접근할 수 없는 지역에 배포하여 야생 동물의 생태와 습성 등을 모니터링 하는 목적으로 사용된다. 또는, 군용 목적으로 분쟁 지역에 배포하여 적군의 이동 경로 등을 파악하는 용도로도 사용된다 [1][2]. 일반적으로 센서 네트워크의 응용은 센서에서 수집한 데이터의 가공 처리를 기반으로 한다. 따라서 센서에서 수집된 데이터에 대한 효율적인 저장, 관리는 데이터 기반 서비스의 QoS(Quality of Service)와 효율적인 운용에 직결된다.

데이터 저장 기법은 데이터 저장 위치와 방식에 따라 크게 외부 저장 방식(ES: External Storage), 내부 저장 방식(LS: Local Storage) 그리고 데이터 중심 저장 방식(DCS: Data Centric Storage)으로 구분된다[3-6]. 외부 저장 방식은 발생한 센싱 데이터를 싱크노드에 전송하여 저장하는 방식이다. 이 방식은 네트워크에서 발생하는 센싱 데이터를 기지국이나 싱크노드에 저장하여 이를 바탕으로 질의를 처리한다. 하지만 이 방식은 모든 수집 데이터를 전송해야 하므로 데이터 수집에 많은 에너지를 소모한다. 반면에 내부 저장 방식은 센서 노드 자체의 저장 공간을 활용하여 센싱 데이터를 저장하는 방식이다. 이 방식은 외부 저장 방식에서 발생하는 높은 데이터 전송비용을 크게 줄인다. 하지만 이 방식은 질의의 대상이 되는 데이터가 저장된 위치를 특정화할 수 없다. 따라서 질의를 전체 네트워크에 배포해야 하며, 이로 인해 추가적인 통신비용이 발생한다. DCS 방식은 데이터 값의 범위에 따라 저장을 담당할 센서를 할당하고, 특정 범위에 포함되는 센싱 데이터가 발생하면 해당 범위를 담당하는 센서에 데이터를 저장하는 방식이다. 이 방식은 외부 저장 방식에서의 과도한 데이터 전송 비용 발생 문제를 해결하고, 동시에 질의 처리에 필요한 데이터가 저장되어 있는 위치를 특정화할 수 있어 질의 처리에 필요한 센서 노드에게 정확히 라우팅하여 질의의 불필요한 확산을 막을 수 있다. 하지만 데이터 중심 저장 기법의 경우 데이터를 수집한 센서 노드에서 데이터를 저장하기 위한 센서 노드로 전송

해야 하는 부가적인 전송 비용이 발생한다.

본 논문에서는 데이터 중심 저장 기법에서 데이터 저장에 필요한 전송 비용을 줄이기 위한 새로운 접근의 데이터 압축 기법을 제안한다. 제안하는 기법은 데이터 전송 시 데이터를 전부 전송하는 것이 아니고 안전영역을 설정하고, 측정값이 안전영역을 벗어나는 경우에만 데이터를 전송함으로써 데이터의 전송을 줄이고 이를 통해 네트워크의 수명을 연장 시킨다.

본 논문의 구성은 다음과 같다. 2장에서 관련 연구에 관해 기술하고 3장에서 제안하는 데이터 저장 및 전송 기법과 질의 처리과정에 대해 설명하고, 4장에서 기존 연구와의 시뮬레이션을 통해 제안하는 기법의 우수성을 증명한다. 5장에서는 결론 및 향후연구를 기술한다.

## II. 관련연구

### 1. 데이터 중심 저장구조

대표적인 DCS 기법으로는 GHT[3], DIM[4], KDDCS[5], EDCS[6]가 있다. GHT기법은 공간 해상 기법을 이용하여 데이터가 저장될 지리적인 위치를 랜덤하게 선정하고, 데이터의 저장, 탐색, 질의를 위해 해당 위치로 패킷을 GPSR[7]기법으로 최단 경로 라우팅을 하는 기법이다.

DIM (Distributed Index for Multi-dimensional data) [4]은 센서 노드들이 온도, 습도, 조도 등 다차원의 속성을 갖는 데이터를 저장하고 질의를 처리하는 것을 효과적으로 지원하기 위한 색인구조이다. DIM은 [그림 1]처럼 2차원 평면상에 배치되는 센서의 위치에 따라서 센서 네트워크 영역을 1차원과 2차원으로 번갈아 가면서 균등하게 분할하여 하나의 영역에 하나의 센서만 남도록 분할한다. 센서 노드가 수집한 데이터는 GPSR 방식에 따라서 소유노드로 전송된다. 이 방식은 데이터가 많이 발생했을 때 외부 저장 공간(싱크 노드, 기지국 등)에 가까운 센서들에게 부하가 집중되거나, 질의를 수행할 때 필요치 않은 센서가 질의 처리에 관여되는 것을 방지할 수 있다.

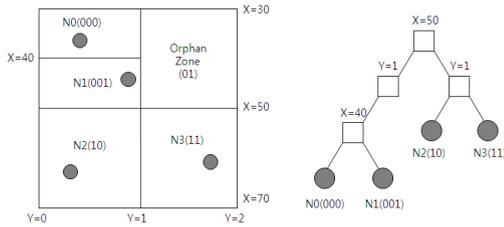


그림 1. DIM 의 구조

KDDCS (K-D tree based Data-Centric Storage) [5] 는 DIM에서 발생할 수 있는 핫스팟(hotspot) 문제와 센서 노드가 존재하지 않는 영역(Orphan Zone) 문제를 해결하기 위해 제안되었다. KDDCS는 DIM과 같이 K-D 트리를 기반으로 하고 있지만, [그림 2]와 같이 영역을 분할할 때 단순하게 2 등분하지 않고 센서 노드가 존재하지 않는 영역이 발생하지 않도록 임의의 지점에서 분할을 한다. 이를 통해 모든 영역에 센서 노드가 포함되도록 할 수 있지만, 라우팅시 DIM과는 다르게 K-D 트리의 하위노드에 대한 분할정보를 가지고 있는 상위 노드들을 방문해야 다음 노드로 이동할 수 있다.

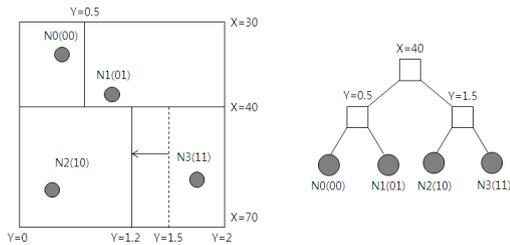


그림 2. KDDCS 의 구조

EDCS[6]는 GPSR 라우팅 기법[7]을 그대로 사용하면 데이터 핫스팟을 해결하는 데이터 중심 저장 기법을 제안한다. 센서 네트워크는 설치된 환경의 영향으로 오류를 발생시키기 쉽기 때문에 일정 연결성(Connectivity)를 보장하기 위해서 단위 면적당 많은 수의 센서 노드를 설치한다. EDCS는 이러한 센서 네트워크의 특성을 이용하여 [그림 3]과 같이 각 공간이 하나 이상, 특정 임계치 이하의 센서 노드를 포함하도록 데이터 공간을 분할한다. DIM의 경우, 각 공간이 하나의 센서 노드를 포함하

는데 비해 제안하는 기법은 다수의 센서 노드를 포함한다. 이때, 각 그리드 셀에 포함된 센서 노드 중 하나를 대표 노드(RN:Representative Node)로 선정하고, 해당 셀의 다른 노드를 대표 노드의 멤버 노드(MN:Member Node)로 정의한다. 초기, 대표 노드를 제외한 나머지 멤버 노드는 데이터 저장에 참여하지 않는다. 대표 노드는 측정값을 효율적으로 저장하기 위해서 수신된 측정값을 인코딩하여 멤버 노드로 전송하며 멤버 노드는 인코딩된 데이터를 저장하는 역할을 수행한다.

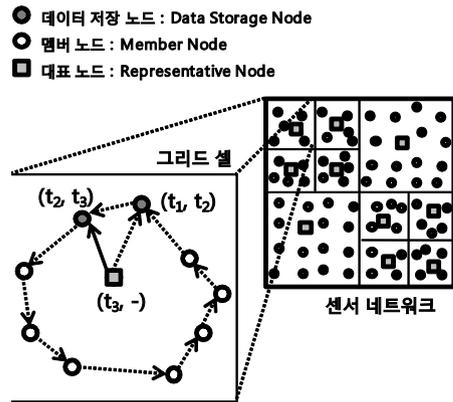


그림 3. EDCS 의 구조

기존에 제안된 DCS 기법은 데이터가 발생할 때 마다 데이터를 저장 영역으로 모두 전송하기 때문에 데이터 전송에 많은 에너지를 소모한다. 따라서 비슷한 데이터가 빈번하게 발생한다면 데이터를 모두 전송해야 하기 때문에 비효율적이다. 따라서 데이터 전송비용을 줄이고 네트워크의 수명을 연장하기 위한 에너지 효율적인 기법에 대한 연구가 필요하다.

## 2. 데이터 압축 기법

현재 많은 응용에서 센서가 수집한 데이터를 전송하거나 모니터링 하는 비용을 줄이는 연구를 다양하게 진행하고 있다[8]. 비용을 줄이기 위해 공간적인 압축이나 시간적인 압축이 필요하다. 첫 번째로 센서에서 수집한 데이터는 공간적으로 연관되어 있다. 생체 모니터링의 예에서, 만약 한 센서가 동물의 소리를 감지했다고 할

때 그 주변의 센서들도 같은 소리를 감지한다. 또한 만약 센서가 토양 오염을 감지했다면 주변의 센서들도 마찬가지로 토양오염을 감지한다. 그렇기 때문에 여러 센서에서 수집된 값을 전송하기 보다는 비슷한 값을 수집한 센서의 그룹이 데이터를 전송한다면 더 효율적이다. 두 번째로 센서에서 수집된 데이터는 시간에 따라 천천히 변화하고, 예상치를 벗어나지 않는다는 특성을 가지고 있기 때문에 센서에서 수집된 데이터는 시간적인 압축이 가능하다. 만약에 마지막으로 수집된 값이 이전 값과의 변화의 정도가 적다면 전송하지 않는다. 싱크노드는 데이터를 수집해야 할 시점에 데이터 전송이 없다면 이전 값으로 데이터를 유지 시킨다. 이러한 기법은 데이터의 변화가 거의 발생하지 않을 때 더욱 효율적이다. DCS는 값이 발생할 때 마다 수집된 값을 저장할 노드로 전송한다. 그래서 많은 에너지 소모가 발생한다. 지금까지는 이러한 압축기법이 DCS에 적용되지 않았다. 그러므로 이러한 불필요한 값의 전송비용을 줄이고 센서네트워크의 수명을 늘리기 위한 DCS를 위한 압축 기법이 필요하다.

### III. 데이터 중심 저장 기법을 위한 제안하는 센서데이터 압축 기법

본 장에서는 제안하는 데이터 중심 저장 구조를 위한 압축 기법을 기술한다. 먼저, 센서데이터의 근사성을 이용하기 위한 안전영역 개념을 설명하고, 데이터의 전송 및 저장 과정을 상세히 기술한다.

#### 1. 특징

제안하는 기법은 [그림 4]와 같이 안전영역을 설정하고 이전 측정값을 기준으로 현재 측정값이 안전영역을 벗어나는 경우만 데이터를 전송한다. 일반적인 DCS 기법은 데이터가 발생하면 지정된 센서로 데이터를 전송하여 저장한다. 하지만 데이터가 발생할 때마다 해당 데이터가 저장 될 영역으로 계속 전송을 하기 때문에 불필요한 전송 비용이 소모된다. 센서 네트워크는 응용 분야에 따라 정확한 데이터를 수집하기 보다는 근사 데

이터를 수집한다. 이것은 두 가지 관점으로 설명할 수 있다. 첫째, 센서가 위치한 지역은 지리적으로 인접해 있기 때문에 비슷한 데이터가 발생할 확률이 높다. 그렇기 때문에 센서네트워크 상에서 데이터가 발생될 때마다 중복 데이터를 전송하는 것은 불필요하다. 둘째, 센서 자체의 오차 범위로 인해 측정값은 어느 정도의 오차를 가지며, 무선으로 데이터가 전송이 되기 때문에 환경적인 제약에 따라 오류가 발생하기 쉽다. 예를 들면, 0.5도 단위로 수집하는 센서는 20.2도 같은 데이터는 20도 또는 20.5도로 수집하기 때문에 정확한 값이 아니다. 그렇기 때문에 센서 데이터에 대한 근사적인 접근을 활용하여 오차 허용범위( $\epsilon$ )를 설정할 수 있다. 오차 허용 범위란 데이터의 오차를 허용할 수 있는 범위이다. 오차 허용 범위 내에서 데이터의 전송은 발생하지 않는다. 예를 들면, 오차 허용 범위를 1도로 설정하고 처음 발생한 값이 20도라고 가정하자. 그러면 20도와 1도 차이가 나기 전까지의 데이터는 전송하지 않고, 1도 차이 이상인 값만 전송한다. 이와 같이, 제안하는 기법은 측정값이 해당 범위를 벗어나는 경우만 측정값을 전송함으로써 데이터 전송 비용을 줄일 수 있다. 제안하는 기법은 측정값이 오차 허용범위 이내에 존재하는 경우, 데이터를 전송하지 않는다. 데이터를 전송하지 않는 오차 허용 범위를 안전 영역으로 정의한다. 안전 영역은 전역 안전 영역(Global Safe Region)과 지역 안전 영역(Local Safe Region)으로 구분된다.

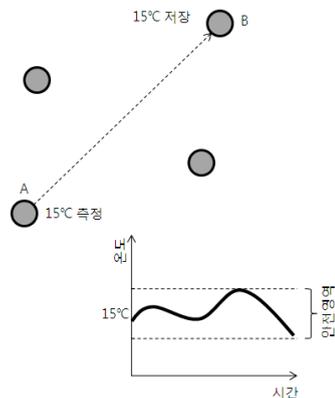


그림 4. 노드A에서의 안전영역 설정

## 2. 전역 안전 영역

제안하는 기법은 측정값이 오차 허용범위 이내에 존재하는 경우, 데이터를 전송하지 않는다. 데이터를 전송하지 않는 오차 허용 범위를 안전 영역으로 정의한다. 안전 영역은 전역 안전 영역과 지역 안전 영역으로 구분된다.

전역 안전 영역은 한 개의 센서가 담당하는 영역을 의미한다. [그림 5]와 같이 온도 15도~20도, 습도 15%~20%의 값이 센싱되면 해당 데이터는 모두 B 노드로 전송된다. 처음으로 데이터가 발생하면 해당 데이터를 저장하는 B 노드에 보고하고 전역 안전 영역을 설정한 후 데이터를 전송한다. 데이터의 전송은 해당 전역 안전 영역을 벗어나기 전까지 수행되며, 센싱 값이 전역 안전 영역을 벗어나면 B 노드에 해당 영역을 벗어났다고 보고한다. B 노드는 A 노드에서 전송한 데이터를 시간 값과 함께 데이터 저장 테이블에 저장한다.

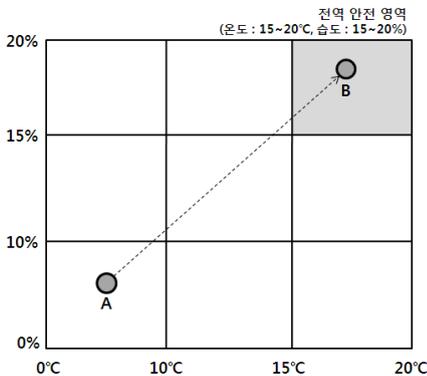


그림 5. 전역 안전 영역 설정

[그림 6]는 전역 안전 영역의 변경과 데이터 저장 테이블의 구조를 나타낸다. 처음에 A 노드는 T0의 시점에 20도~24도의 값을 센싱 한다. 이 값을 전송하기 위해 B 노드를 전역 안전 영역으로 설정한 후 데이터를 전송한다. 그 후 T1의 시점에 25도~30도의 값이 발생하면 해당 전역 안전 영역을 벗어나는 값이기 때문에 A 노드는 B 노드에 범위를 벗어났다는 메시지를 보내고 C 노드를 전역 안전 영역으로 설정한다. 그 후 센싱된 데이터는 모두 C 노드로 전송한다. 데이터 저장 테이블에는 고

유한 센서의 아이디와 데이터 그리고 시간이 저장된다. B 노드의 저장테이블에는 T0 시점에 A 노드가 센싱한 20도의 값이 저장된다. T1 시점에서 A 노드는 25도를 센싱하고 이 값은 C 노드의 저장 테이블에 저장된다.

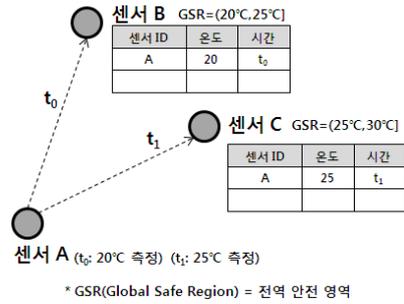


그림 6. 전역 안전 영역의 변경 및 데이터 저장 테이블의 구조

## 3. 지역 안전 영역

지역 안전 영역은 오차 허용 범위내에서 데이터가 전송되지 않는 영역이다. 일반적으로 센서 네트워크상의 데이터는 시간의 흐름에 따라 크게 변화하지 않고 유사한 값이 지속적으로 발생한다는 특성이 있다. 제안하는 기법은 일단 데이터가 센싱되면 처음 센싱된 값을 가지고 전역 안전 영역을 설정한다. 다음으로 오차 허용 범위를 설정하고 오차 허용 범위 내의 영역을 지역 안전 영역으로 선정한다. 데이터의 전송 여부는 지역 안전 영역 내에서 설정한 오차 허용 범위에 따라 결정된다.

[그림 7]은 지역 안전 영역의 설정에 따른 센싱값 전송을 나타낸다. [그림 7(a)]와 같이 오차 허용 범위를 설정하지 않고 주기적으로 센싱값을 전송하는 경우 총 28회 전송이 발생한다. [그림 7(b)]와 [그림 7(c)]와 같이 오차 허용 범위를 각각 1°C, 2°C로 설정한 경우, 지역 안전 영역 또는 전역 안전 영역이 변경되는 시점에서 센싱값을 전송하기 때문에 각각 13회, 8회의 데이터 전송이 발생한다. 오차 허용 범위의 설정은 응용에 따라 다르게 할 수 있다. 만약 정확한 값을 요구하는 응용에서는 오차 허용 범위를 작게 설정해서 더욱 정확한 값을 전송해야 하고 그렇지 않다면 오차 허용 범위를 크게 설정해서 근사 데이터 값을 전송한다. 제안하는 기법은 이처럼 전역 안전 영역과 지역 안전 영역을 이용하여

데이터 압축의 효과를 내며 데이터의 전송 빈도를 줄여 에너지 소모를 줄인다.

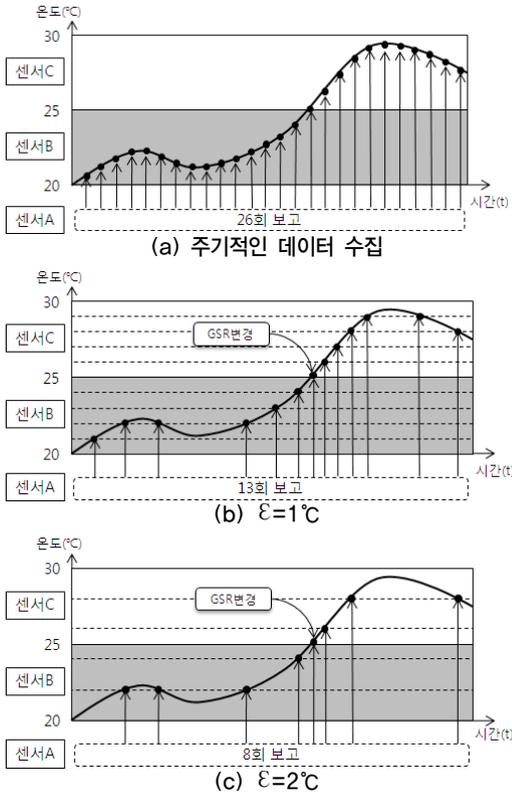


그림 7. 지역 안전 영역 설정에 따른 센싱값 전송

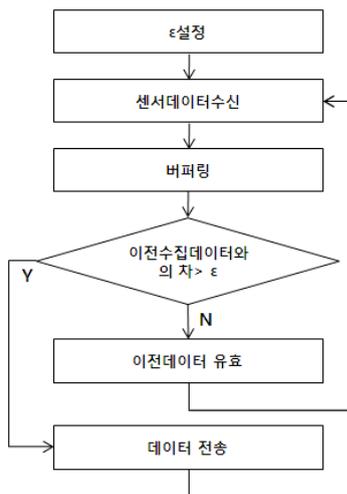


그림 8. 데이터 전송과정

[그림 8]은 제안하는 기법의 데이터 전송과정을 보여 준다. 우선 처음에 오차 허용 범위를 설정한 후, 센서 데이터를 수신한다. 값을 수신하면 처음의 데이터를 저장하고, 다음 데이터가 수집될 때까지 대기 한다. 다음 값이 수집되면 이전에 저장된 값과 현재 수집된 값의 차이가 처음에 설정한 오차 허용 범위보다 크지 비교한다. 그 후, 만약 값의 차이가 오차 허용 범위보다 작다면 이전 값을 유지 시키고 크다면 현재 수집된 값을 전송한다.

#### 4. 오차 허용 시간(E) 설정

DCS 기법은 질의 발생 시 해당 값을 저장하고 있는 노드로 질의를 배포하고 질의에 대한 값을 수신한다. 또한 각 노드에서 수집한 값을 계속해서 전송하기 때문에 저장된 값의 최신성을 보장할 수 있다. 하지만 제안하는 기법은 오차 허용 범위 이상이 변할 때 전송을 하기 때문에 저장된 값의 최신 성을 보장할 수 없고 노드의 fail이나 오류를 알아보기 위해 오차 허용 범위를 설정한다. 예를 들어 ‘현재 기온이 20도인 지역을 알려줘’라는 질의가 20도를 저장하는 노드로 내려졌을 때 저장된 데이터가 1시간 전에 수집된 데이터라면 데이터의 최신 성을 보장할 수 없다. 제안하는 기법은 오차 허용 시간을 설정하여 오차 허용 시간 이내의 데이터는 직접 전송하고 오차 허용 시간을 벗어나면 데이터를 수집한 센서 노드로 최신 데이터 값을 요청하여 질의 결과를 전송한다. 오차 허용 시간의 설정은 응용에 따라 다르게 설정할 수 있다. 만약 정확한 데이터를 요구하는 응용에서는 오차 허용 시간을 짧게 설정하고, 그렇지 않다면 길게 설정한다. [그림 9]는 오차 허용 시간이 설정된 상태에서 질의에 대한 응답의 예를 나타낸다. 센서 데이터가 저장된 시점에서 오차 허용 시간은 갱신되며 질의 Q1, Q2와 같이 오차 허용 시간 내에 요청된 질의에 대해서 센서 데이터를 저장하고 있는 센서 노드 B, C가 각각 질의 결과를 반환한다. 하지만, 질의 Q3, Q4와 같이 오차 허용 시간을 벗어나 요청된 경우 데이터의 최신성을 보장하기 위해서 센서 노드 C는 센서 노드 A로 질의 처리를 요청하며 센서 노드 A는 질의 Q3, Q4에 대해 센서 데이터를 반환한다.

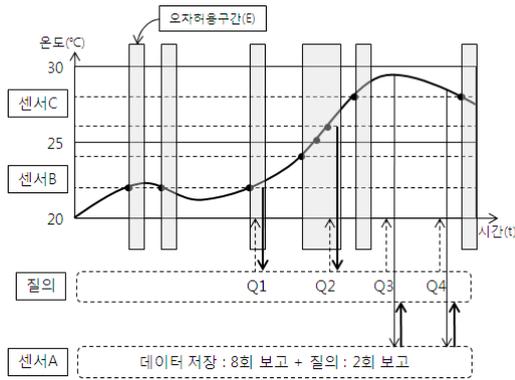


그림 9. 오차 허용 시간을 고려한 질의 처리

#### IV. 성능평가

##### 1. 실험 환경

제안하는 기법의 우수성을 보이기 위해서 기존에 제안된 히스토그램 기반 기법에서 수행한 실험과 동일한 방법으로 성능 평가를 수행하였다. 실험에 사용한 데이터는 인텔 연구실[9]에서 1개월 동안 54개의 센서로부터 수집한 실제 환경 데이터이다. 중앙에 기지국이 위치한 100m X 100m의 영역 공간을 가정하였으며, 임의로 54개의 센서 노드를 위치시키고 센서 노드는 다중-홉으로 통신한다고 가정하였다. 임의의 네트워크 토폴로지에서 연결성을 보장하기 위해서 통신 반경은 18로 설정하였다. 제안하는 압축기법의 유무에 따른 DCS 기법의 성능을 평가하기 위해서 대표적인 DCS 기법인 DIM을 기반으로 하였다. [표 1]은 임의로 구성된 1,000개의 센서 네트워크로부터 측정된 평균 중간 노드의 수, 평균 전송 거리이다. 센서 노드의 통신 반경은 18m

표 1. 성능 파라미터

센서노드의수(N)	54
중간노드의평균수(N <sub>m</sub> )	25.7
통신반경	18m
평균통신홉수	4.26
측정값의 범위	10 ~ 40도
측정값의 수	94156

로 설정한다. 센서 노드의 메시지 전송에 소모되는 에너지 모델은 {메시지 크기} x {전송 비용} + {증폭 비용} x {거리}²이며, 전송 비용은 50nJ/b, 증폭 비용은 100pJ/b/m²으로 설정한다. 메시지 수신에 소모되는 에너지 모델은 {메시지 크기} x {수신 비용}이며, 수신 비용은 50nJ/b으로 설정한다.

##### 2. 통신 비용

시뮬레이션은 54개의 센서 노드를 균등하게 배포한 환경에서 수행하였다. 제안하는 기법의 환경변수인 전역 안전 영역은 10~40도 범위를 54개로 균등하게 분할 설정하였고, 지역 안전 영역을 0.1~0.5의 범위 내에서 0.1단위로 변화시켜가며 테스트하였다. [그림 10]은 통신비용 측면에서 제안하는 기법과 DIM을 비교한 결과이다. 비교 결과, DIM의 경우 네트워크에서 발생하는 모든 데이터를 아무런 가공 처리 없이 전송하여 높은 통신비용을 소모한다. 하지만 제안하는 기법의 경우 발생하는 데이터의 유사한 정도를 판단하여 설정된 오차 허용범위 이내의 데이터가 발생하였을 때, 이러한 데이터의 전송을 불필요한 데이터 전송으로 간주하고, 이를 제거함으로써 많은 통신비용적 측면의 이득을 가진다. [그림 10]과 같이 실험결과 제안하는 기법에서 오차 허용범위가 증가함에 따라 발생하는 통신비용의 이득이 증가하는 것을 확인할 수 있으며, 0.3을 기점으로 수렴하는 것을 확인할 수 있다.

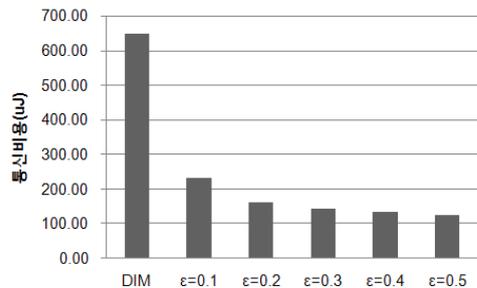


그림 10. 오차 허용 범위에 따른 통신비용 비교

##### 3. 정확도

[그림 11]은 데이터의 정확도 관점에서 제안하는 기

법과 DIM기법을 비교한 결과이다. DIM 기법에서는 전체 데이터를 모두 전송하기 때문에 100%의 정확도를 보인다. 하지만 제안하는 기법은 지역 안정 영역을 설정하여, 설정한 오차 허용범위를 넘어설 경우만 전송하기 때문에 DIM에 비해서 정확도는 다소 떨어진다. 그러나 데이터의 전송이득을 고려해볼 때, 제안하는 기법의 데이터에 대한 정확도의 차이는 0.5~1% 정도이기 때문에 큰 차이를 보이지 않는다.

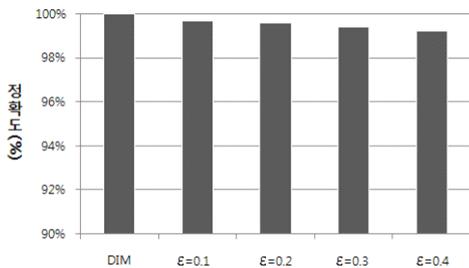


그림 11. 오차 허용 범위에 따른 정확도 비교

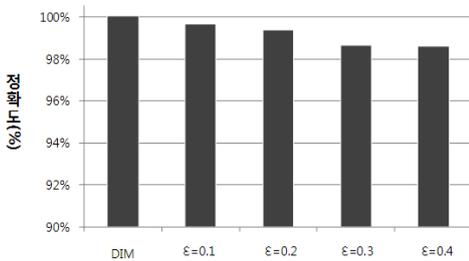


그림 12. 질의 처리 시 정확도 비교

[그림 12]는 한 노드에 질의가 내렸을 때 데이터의 정확도 관점에서 제안하는 기법과 DIM기법을 비교한 결과이다. DIM기법은 계속 해서 데이터를 전송하기 때문에 노드로 질의를 내렸을 때 데이터의 최신성을 보장할 수 있고 100%의 정확도를 보인다. 하지만 제안하는 기법은 오차 허용범위를 넘어서면 전송하고, 전송되지 않았을 때는 이전에 전송된 데이터 값을 유지하므로 데이터에 대한 정확도는 다소 떨어진다. 그러나 데이터의 전송이득을 고려해볼 때, 제안하는 기법의 질의에 대한 데이터 정확도의 차이는 1~2% 정도이기 때문에 큰 차이를 보이지 않는다. 또한 오차 허용범위는 0.4를 기점으로 수렴하는 것을 알 수 있다. 이러한 정확도에 대한

결과는 근사데이터를 수집하는 센서 네트워크의 응용에 적합하다는 것을 보여준다.

## V. 결론 및 향후 연구

본 논문에서는 데이터 중심 저장 기법에서 데이터의 전송 비용을 줄이기 위한 새로운 접근의 데이터 압축 기법을 제안하였다. 제안하는 기법은 발생한 데이터를 전부 전송하는 것이 아니라 안전영역을 설정하고 오차 허용 범위를 사용하여 오차 허용 범위를 넘었을 때만 데이터를 전송함으로써 데이터의 압축의 효과와 함께 데이터 전송량을 줄이는 기법이다. 성능평가 결과 기존 기법에 비해 데이터 전송 비용이 60% 감소하였다. 향후 연구는 시간의 근사 속성을 이용하여 질의 처리 비용을 줄이는 것이다.

### 참고 문헌

- [1] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks", In Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, pp.56-67, 2000.
- [2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: A Survey", Computer Networks, Vol.38, No.4, pp.393-422, 2002.
- [3] S. Ratnasamy, B. Karp, S. Shenker, D. Estrin, R. Govindan, L. Yin and F. Yu, "Data-Centric Storage in Sensornets with GHT, a geographic hash table", Mobile Network and Applications, Vol.8, No.4, pp.427-442, 2003.
- [4] X. Li, Y. J. Kim, R. Govindan and W. Hong, "Multi-Dimensional Range Queries in Sensor Networks.", In Proceedings of ACM Conference

on Embedded Networked Sensor Systems, pp.63-75, 2003.

- [5] M. Aly, K.Pruhs, and P. K. Chrysanthis, "KDDCS:A LoadBalanced In Network Data Centric Storage Scheme for Sensor Networks", In Proceedings of International Conference on Information and Knowledge Management, pp.317-326, 2006.
- [6] 여명호, 성동욱, 노규중, 송석일, 유재수, "센서 네트워크에서 지리적인 라우팅과 데이터 핫스팟을 고려한 인코딩 기반 데이터 중심 저장 기법", 2009년 한국컴퓨터종합학술대회, 한국정보과학회, pp.54-55, 2009.
- [7] B. Karp and H. T. Kung, "GPSR: Greedy Perimeter Stateless Routing for Wireless Sensor Networks", In Proceedings of International Conference on Mobile Computing and Networking, pp.243-254, 2000.
- [8] M. Yeo, D. Seo, and J. Yoo, "Data Correlation-Based Clustering Algorithm in Wireless Sensor Networks", KSII Transactions on Internet and Information Systems, Vol.3, No.3, pp.331-343, 2009.
- [9] <http://berkeley.intel-research.net/labdata/>

저 자 소 개

노 규 중(Kyu-Jong Roh)

정회원



- 2008년 2월 : 충북대학교 정보통신공학과(공학사)
- 2010년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2010년 2월 ~ 현재 : 삼성SDS 전자LCD 인프라그룹

<관심분야> : 센서네트워크, RFID 시스템, 플래시 메모리, DB 시스템

여 명 호(Myung-Ho Yeo)

정회원



- 2004년 2월 : 충북대학교 정보통신공학과(공학사)
- 2006년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2010년 2월 : 충북대학교 정보통신공학과(공학박사)

▪ 2010년 2월 ~ 현재 : 국방과학연구소 제2기술연구본부 (선임연구원)

<관심분야> : 데이터베이스시스템, 정보검색, 센서네트워크 및 RFID, 멀티미디어데이터베이스, 분산객체 컴퓨팅

성 동 욱(Dong-Ook Seong)

정회원



- 2005년 2월 : 충북대학교 정보통신공학과(공학사)
- 2007년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2007년 2월 ~ 현재 : 충북대학교 정보통신공학과 박사과정 재학중

<관심분야> : LMS, LCMS, 데이터베이스 시스템, 센서 네트워크, 저장 시스템

복 경 수(Kyoung-Soo Bok)

정회원



- 1998년 2월 : 충북대학교 수학과(이학사)
- 2000년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2005년 2월 : 충북대학교 정보통신공학과(공학박사)

▪ 2005년 3월 ~ 2008년 2월 : 한국과학기술원 정보전자연구소 Postdoc

▪ 2008년 3월 ~ 현재 : 가인정보기술 연구소 차장  
<관심분야> : 데이터베이스 시스템, 위치기반서비스, 이동 P2P, 센서네트워크 및 RFID

신 재 룡(Jae-Ryong Shin)

정회원



- 1996년 2월 : 충북대학교 정보통신공학과(공학사)
- 1998년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2002년 8월 : 충북대학교 정보통신공학과(공학박사)

- 2003년 3월 ~ 현재 : 광주보건대학 의약정보관리과 교수

<관심분야> : 실시간데이터베이스, 내용기반검색

유 재 수(Jae-Soo Yoo)

종신회원



- 1989년 2월 : 전북대학교컴퓨터공학과(공학사)
- 1991년 2월 : 한국과학기술원 전산학과(공학석사)
- 1995년 2월 : 한국과학기술원 전산학과(공학박사)

- 1995년 3월 ~ 1996년 8월 : 목포대학교 전산통계학과(전임강사)

- 1996년 8월 ~ 현재 : 충북대학교 전기전자컴퓨터공학부 및 컴퓨터정보통신연구소 교수

<관심분야> : 데이터베이스시스템, 정보검색, 센서네트워크 및 RFID, 멀티미디어데이터베이스, 분산객체컴퓨팅