

무선 센서 네트워크에서 네트워크 트래픽 감소를 위한 데이터 중심 클러스터링 알고리즘

(A Data-Centric Clustering Algorithm for Reducing Network Traffic in Wireless Sensor Networks)

여명호[†] 이미숙[†] 박종국[†] 이석재^{**} 유재수^{***}
 (Myung Ho Yeo) (Mi Sook Lee) (Jong Guk Park) (Seok Jae Lee) (Jae Soo Yoo)

요약 센서 네트워크를 사용하는 응용분야에 따라 보다 고차원적인 데이터 처리를 필요로 하는 경우 모든 센서 노드의 수집 데이터를 싱크 노드로 전송한다. 수집된 데이터는 일반적으로 센서 네트워크의 환경적인 특성상 시간적으로 혹은 공간적으로 연관성을 지닌다. 이러한 연관성은 싱크 노드가 일부의 데이터만 수집하고도 모든 데이터를 복원할 수 있는 기회를 제공한다. 센서 네트워크에서는 데이터 수집을 위한 기법으로 클러스터링 기법을 널리 사용한다. 하지만 기존의 클러스터링 기법의 경우 수집한 데이터의 연관성을 고려하지 않고, 센서 노드의 지역성(locality)만을 고려하여 클러스터를 생성하기 때문에 이러한 기회를 활용하기에 비효율적이다. 본 논문에서는 수집된 데이터를 중심으로 클러스터를 생성하고, 싱크 노드로 전송되는 데이터의 크기를 획기적으로 줄일 수 있는 클러스터링 기법을 제안한다. 제안하는 클러스터링 기법의 우수함을 보이기 위해 시뮬레이션을 통한 성능 평가를 수행하였으며, 그 결과 기존 기법들에 비해 네트워크 트래픽이 약 4~40% 감소하고, 네트워크의 수명이 약 20~30% 연장되었다.

키워드 : 센서 네트워크, 클러스터링, 네트워크 생명주기, 에너지 효율

Abstract Many types of sensor data exhibit strong correlation in both space and time. Suppression, both temporal and spatial, provides opportunities for reducing the energy cost of sensor data collection. Unfortunately, existing clustering algorithms are difficult to utilize the spatial or temporal opportunities, because they just organize clusters based on the distribution of sensor nodes or the network topology but not correlation of sensor data. In this paper, we propose a novel clustering algorithm with suppression techniques. To guarantee independent communication among clusters, we allocate multiple channels based on sensor data. Also, we propose a spatio-temporal suppression technique to reduce the network traffic. In order to show the superiority of our clustering algorithm, we compare it with the existing suppression algorithms in terms of the lifetime of the sensor network and the size of data which have been collected in the base-station. As a result, our experimental results show that the size of data was reduced by 4~40%, and whole network lifetime was prolonged by 20~30%.

Key words : sensor networks, clustering, network lifetime, energy efficiency

• 이 논문은 2007년도 정부(교육인적자원부)의 재원으로 한국학술진흥재단 지원
 원(지방연구중심대학육성사업/충북BIT연구중심대학육성사업단)과 과학기술
 부·한국과학재단의 지원(No.R01-2006-000-10809-0)을 받아 수행된 연구임
 • 이 논문은 2007 한국컴퓨터종합학술대회에서 '무선 센서 네트워크에서 네트
 워크 트래픽 감소를 위한 데이터 중심 클러스터링 알고리즘'의 제목으로 발
 표된 논문을 확장한 것임

[†] 학생회원 : 충북대학교 정보통신공학과
 mhyeo@netdb.cbnu.ac.kr
 mslee@netdb.cbnu.ac.kr
 jgpark@netdb.cbnu.ac.kr
^{**} 정회원 : 충북대학교 정보통신공학과
 sjlee@netdb.cbnu.ac.kr

^{***} 종신회원 : 충북대학교 정보통신공학과 교수
 yjs@chungbuk.ac.kr

(Corresponding author)

논문접수 : 2007년 9월 27일

심사완료 : 2008년 1월 4일

Copyright@2008 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작
 물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다.
 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처
 를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든
 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야
 합니다.

정보과학회논문지: 정보통신 제35권 제2호(2008.4)

1. 서 론

최근 무선 센서 네트워크는 환경 감시, 스마트 공간, 의료 기술, 농업 기술 분야와 같은 다양한 분야의 응용으로 각광받고 있다[1-3]. 이러한 센서 네트워크는 온도, 습도, 진도 등과 같은 주변 환경의 유용한 정보를 수집하고, 보다 다양한 처리를 위해서 수집한 정보를 싱크 노드로 전송한다. 센서 네트워크에서 사용되는 센서 노드는 에너지 사용의 제약을 가지고 있기 때문에 데이터를 수집하거나 전송하는데 있어서 에너지를 효율적으로 사용하기 위한 방법이 필요하다.

무선 통신의 발전과 다기능, 저가의 스마트 센서(smart sensor)들의 개발로 보통 한 번에 배포되는 센서 노드들의 수가 수백~수만 개에 이르고 센서 노드들이 배포되는 환경이 매우 가변적이므로 센서 노드를 일일이 재충전하거나 교체를 하는 것은 쉽지 않다. 이에 제한된 에너지를 효율적으로 사용하여 센서 네트워크의 수명(lifetime)을 최대화 할 수 있는 기법과 함께 확장성(scability), 부하 균등(load balancing) 이 센서 네트워크의 응용에 중요한 요구사항이 되었다[4-6].

클러스터링 기법은 이러한 요구사항을 효과적으로 처리할 수 있는 기법 중 하나이다. 클러스터링 기법은 여러 가지 장점을 지닌다. 실제 네트워크는 센서 노드가 추가로 배포되거나 배터리를 모두 소모하여 센서 노드가 동작을 하지 않는 경우와 같이 동적인 위상 변화가 발생했을 때 상위 계층은 모든 변화를 인식해야 한다. 또한, 센서 네트워크처럼 많은 수의 노드로 이루어진 네트워크의 경우, 이러한 변화는 상위 계층에게 상당한 부담으로 작용하게 된다. 하지만 이러한 부담을 클러스터링을 통해 효과적으로 줄일 수 있다. 즉, 클러스터 내부의 노드와는 상관없이 응용 계층에서 내려다보는 시각에서는 클러스터를 하나의 단위로 보고 통신하기 때문에 라우팅 테이블 등의 간접에 대한 비용을 보다 줄일 수 있다. 이와 함께 응용 계층에서는 클러스터 헤드에서 클러스터 멤버 노드의 데이터를 압축하거나 병합하여 좀 더 큰 범위의 확장성을 제공한다[5].

응용 분야에 따라 데이터의 전송 비용을 줄이기 위해서 싱크 노드는 병합된 데이터(aggregated data)를 수집하기도 하고, 싱크 노드에서 고차원적인 데이터 처리를 위해서 개별적인 센서의 데이터를 모두 수집하는 것을 요구하기도 한다. 센서에서 수집한 데이터는 환경적인 특성상 시간적으로 혹은 공간적 연관성을 지닌다. 데이터 제한 기법은 이러한 특성을 활용하여 싱크 노드로 전송해야 할 데이터의 크기를 획기적으로 줄여 통신에 따른 에너지 소모를 줄인다[7]. 데이터 제한 기법은 시간적인 제한(temporal suppression) 기법과 공간적인

제한(spatial suppression) 기법으로 나눌 수 있다. 시간적인 제한 기법이란 센서 노드에서 마지막으로 보고된 데이터를 기준으로 현재 데이터가 변화가 없을 때 싱크 노드로 해당 데이터를 보고하지 않는 것을 의미한다. 이 때, 싱크 노드는 보고되어지지 않은 데이터의 경우 이전과 변화가 없다고 가정한다. 데이터 제한 기법의 다른 한 종류인 공간적인 제한 기법은 센서 노드의 새로운 수집 데이터가 이웃 노드들의 데이터와 동일하면 제한을 수행하는 것이다.

일반적으로 클러스터 멤버로부터 데이터를 수집한 클러스터 헤드는 병합된 데이터를 상대적으로 먼 거리에 위치한 싱크 노드로 전송하기 때문에 더 많은 에너지를 소모한다. 이는 전송하는 데이터 크기와 비례하며, 시/공간적인 데이터 제한 기법을 클러스터링 기법에 적용하여 전송하는 데이터 크기를 줄일 경우, 통신비용의 감소와 함께 네트워크 수명의 연장 효과를 기대할 수 있다. 하지만 기존의 클러스터링 기법의 경우 센서 분포의 지역성(locality)만 고려하여 클러스터를 생성하기 때문에 공간적인 데이터 제한(data suppression) 기법을 적용할 경우 효과적이지 못하다.

본 논문에서는 먼저 시/공간적인 데이터 제한 기법의 적용에 따른 효과와 문제점을 분석하고, 데이터 제한 기법의 효율을 높일 수 있는 데이터 기반 클러스터링 기법(data-centric clustering algorithm)을 제안한다. 본 논문의 구성은 다음과 같다. 제2장에서는 관련연구로 본 논문에서 사용하고 있는 에너지 소모 모델과 기존에 제안된 클러스터링 기법, 그리고 시/공간 특성에 따른 데이터 제한 기법을 살펴본다. 제3장에서는 시/공간적인 데이터 제한 기법을 클러스터링에 적용하는 방법과 그 문제점에 대해 분석하고, 제4장에서는 본 논문에서 제안하는 데이터 중심 클러스터링 알고리즘의 특징과 상세 과정을 기술한다. 제5장에서는 성능평가와 분석을 통해 제안하는 기법의 우수성을 보이고, 제6장에서 결론과 향후 연구 방향을 제시한다.

2. 관련연구

2.1 데이터 제한 기법

응용 분야에 따라 데이터의 전송 비용을 줄이기 위해서 싱크 노드는 병합된 데이터(aggregated data)를 수집하기도 하고, 싱크 노드에서 고차원적인 데이터 처리를 위해서 개별적인 센서의 데이터를 모두 수집하는 것을 요구하기도 한다. 그림 1은 센서를 통해 수집한 데이터의 특성을 나타낸다. 센서에서 수집한 데이터는 환경적인 특성상 시간적 혹은 공간적 연관성을 지닌다. 그림 1(a)는 자동차의 디스크, 페드, 브레이크 액에 센서를 부착하여 언덕길 주행 중 브레이크 페달을 밟는 순간 각

부위의 온도 변화를 나타낸다. 이때, 온도는 시간이 변함에 따라 연속적으로 변화한다. 이와 같이 센서를 통해 측정하는 대부분의 물리량은 급격하게 변화하기 보다는 시간적으로 연속적인 값을 가진다. 센서 데이터들은 그림 1(b)와 같이 공간적으로도 연관성이 있다. 그림 1(b)는 기압을 측정하는 센서들을 배포하여 수집된 데이터를 기반으로 작성한 우리나라 주변의 등압선이다. 인접한 지역인 A 지역과 B 지역의 수집된 기압을 분석해보면 실제로 두 지역의 기압이 연관성을 가지고 있다. 이는 공간적으로 인접한 지역에 배치되어 있는 센서들이 유사한 값을 수집했다는 것을 나타낸다.

데이터 제한 기법은 이러한 센서 데이터의 특성을 활용하여 싱크 노드로 전송해야 할 데이터의 크기를 줄이는 방법이다. 데이터 제한 기법은 크게 시간적인 제한(temporal suppression) 기법과 공간적인 제한(spatial suppression) 기법으로 나누어진다.

시간적인 제한 기법이란 센서 데이터의 시간적인 유사도를 이용하여 센서 노드에서 마지막으로 보고된 데이터를 기준으로 현재 데이터의 변화가 발생한 경우 싱크 노드로 데이터를 전송하는 기법이다. 이때, 싱크 노드는 보고되어지지 않은 데이터의 경우 이전과 변화가 없다고 가정한다. [8]에서 제안하는 시간적인 제한 기법은 센서 데이터의 오차 허용치를 미리 정하고, 수집한 데이터의 변화가 허용치보다 큰 경우만 데이터를 싱크 노드로 전송한다.

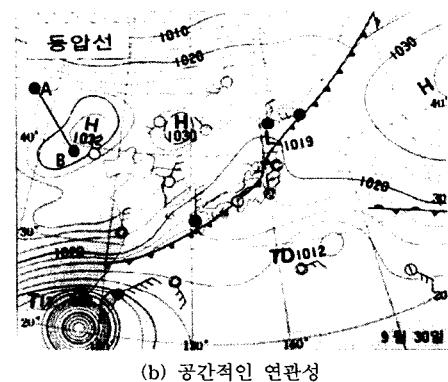
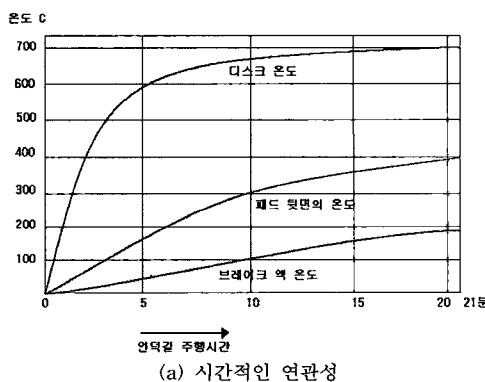
데이터 제한 기법의 다른 한 종류인 공간적인 제한 기법은 이웃 노드들의 데이터와의 연관성을 고려한다. [9]에서는 평균 연산자를 이용하여 공간적인 제한을 수행한다. 모든 노드들은 서로 다른 시간 슬롯(slot)을 할당 받고, 시간 슬롯에 따라 자신의 데이터를 싱크 노드에게 전송한다. 센서 노드들은 자신의 시간 슬롯을 기다리는 동안 이웃 노드들이 싱크 노드로 전송하는 데이터를 도청한다. 이때, 도청한 모든 값에 대한 평균값을 계

산하고, 자신의 값과 이 평균값을 비교하여 자신의 값과 동일한 경우 싱크 노드에게 데이터를 전송하지 않는다. 데이터 제한 기법의 효율을 높이기 위해서 클러스터링과 트리 구조와 같은 인-네트워크 처리(in-network processing)를 이용하는 방법이 제안되었다[10,11]. 공간적으로 인접한 센서들의 데이터가 서로 연관성이 높은 경우, 각각의 센서들이 싱크 노드에게 데이터를 전송하는 것보다 네트워크 내부에서 데이터를 병합한 후 전송하는 것이 더 효율적이다. [10]은 센서 노드들이 클러스터를 형성하고, 각 클러스터의 멤버 노드들은 자신의 클러스터 헤드 노드에게 자신의 데이터를 전송한다. 이때 클러스터 헤드 노드는 수집된 멤버 노드들의 데이터에서 중복된 데이터를 제거함으로써 공간적인 데이터 제한을 수행한다. 그리고 [11]은 네트워크 내에 특정 영역을 선언하고 그 영역 범위 안의 노드들 중 싱크 노드에게 가장 가까운 가장자리 노드를 대표 노드로 선정하여 영역 범위 내의 노드들에 대한 데이터를 수집한다.

최근에는 시간적인 데이터 제한 기법과 공간적인 데이터 제한 기법의 장점을 결합하기 위한 연구로 시/공간적인 데이터 제한 기법이 많이 연구되고 있다. 두 데이터 제한 기법은 “OR” 혹은 “AND” 형태로 결합할 수 있으나, “OR” 형태로 결합하는 경우, 싱크 노드는 보고되어 지지 않은 데이터가 시간적 데이터 제한으로 보고되지 않았는지 공간적 데이터 제한으로 보고되지 않았는지를 알 수 없는 문제점이 있다. “AND” 형태로 결합하는 경우, 하나의 데이터 제한 기법을 사용하는 것보다 데이터의 효율이 떨어지는 문제점이 있다.

2.2 대표적인 클러스터링 기법

일반적으로 클러스터를 구성하게 되면, 클러스터 헤드는 데이터를 수집하고 수집된 데이터를 싱크 노드에게 전송하기 위해서 다른 센서 노드들에 비해 더 많은 에너지를 소모하게 된다. 이때, 클러스터 헤드의 에너지 소모비용은 클러스터의 크기와 노드의 수에 따라 달라



진다. 따라서 센서 네트워크의 수명을 연장하기 위해서 클러스터 헤드의 에너지 소모를 균형 있게 만드는 것이 중요하다.

LEACH는 클러스터를 이용하여 센싱된 데이터를 싱크 노드까지 전달하는 가장 널리 알려진 클러스터링 기법 중 하나이다[12]. 이 기법이 가정하고 있는 대상은 다음과 같다. 먼저 사용자는 센서 필드와 노드의 수를 알고 있어야 하고 가장 적절한 클러스터 헤드의 수를 지정할 수 있어야 한다. 모든 노드는 자신의 신호 세기를 조정할 수 있으며, 클러스터 헤드는 클러스터 내의 노드로부터 센싱된 데이터를 정확히 수집하며 수집된 정보를 싱크 노드로 한 번에 전송한다. 즉, 필드 내의 모든 노드는 최대 세기로 신호를 보낼 경우 싱크 노드까지 정보를 전달할 수 있다. 클러스터 헤드의 에너지 소모를 분산하기 위한 방법으로 LEACH 알고리즘은 매 주기(round)마다 클러스터 내부에서 클러스터 헤드를 랜덤하게 선출한다. LEACH에서 클러스터 헤드를 결정하는 방법은 다음과 같다. 사용자는 미리 알고 있는 노드의 수와 필드에 따라 확률 P를 결정하는데, 이것은 매주기마다 각각의 노드가 클러스터 헤드가 될 수 있는 확률을 뜻한다. LEACH는 시간을 사이클로 나누고, 그것을 각 라운드로 다시 나누는데, 모든 노드는 하나의 사이클 내에 단 한 번씩 클러스터 헤드가 될 수 있다. 하지만 이는 주변 노드의 상태를 고려하지 않기 때문에 불균형적인 클러스터가 형성이 될 수 있으며, 클러스터에 포함된 센서 노드의 수에 따라 불균형적인 에너지 소모를 초래한다. 또한, 클러스터 구성 시 에너지에 대한 고려가 전혀 없다는 것도 문제가 될 수 있다. 이러한 문제를 해결하기 위해 LEACH-C 알고리즘에서는 잔여 에너지 정보를 이용하여 더 많은 에너지를 가지고 있는 센서 노드를 클러스터 헤드로 선출하는 방법을 제안했다[13]. 하지만 여전히 확률적인 방법을 이용하여 클러스터를 생성하기 때문에 상황에 따라 불균형적인 네트워크 위상을 만들 수 있다는 문제점을 가지고 있다.

HEED는 LEACH의 문제점을 해결하기 위해서 주기적으로 노드의 에너지 혹은 그 이외의 여러 요소를 고려하여 클러스터 헤드를 선출한다[14]. 이 기법은 잔여 에너지뿐 아니라 이웃 노드에 대한 근접도와 같은 다양한 파라미터를 고려하여 적절한 노드의 밀도, 클러스터 내/외의 전송 거리를 결정함으로써 균등한 분산 클러스터를 구성한다.

ANTCLUST는 클러스터 노드간의 지역 통신(local communication)을 통해서 클러스터의 정보를 교환하고, 클러스터 헤드까지의 거리, 잔여 에너지와 같은 추가적인 정보를 통해 동적으로 클러스터를 재구성하는 방법을 제안했다[15]. 그 결과 네트워크의 생존주기를 LEACH

와 HEED보다 30% 연장시킨다.

3. 분석: 클러스터링 기법을 위한 데이터 제한 기법 적용

본 장에서는 시/공간 데이터 제한 기법을 클러스터링 기법에 적용하는 방법을 알아보고, 적용에 따른 문제점을 분석한다.

3.1 시간적인 데이터 제한 기법

시간적인 데이터 제한 기법을 클러스터링 기법에 적용하는 경우, 별다른 수정 없이 알고리즘을 적용할 수 있다. 즉, 그림 2와 같이 클러스터에 포함된 각 센서 노드들은 클러스터 헤드로 데이터를 보내는 시점에서 최근에 전송한 데이터와 비교하여 동일한 경우 데이터를 전송하지 않는다. 예를 들면, $T=T_2$ 인 경우, 이전 시점인 $T=T_1$ 일 때 수집한 데이터와 동일하기 때문에 클러스터 헤드로 데이터를 전송하지 않는다. 클러스터 헤드를 통한 데이터 수집은 TDMA 스케줄에 의해 이루어지기 때문에 데이터를 수신하지 않은 클러스터 헤드는 해당 시간 슬롯의 센서 노드가 이전 데이터와 동일한 값을 수집했다고 판단한다. 최종적으로 클러스터 헤드가 싱크 노드로 데이터를 전송하는 경우, 싱크 노드 역시 수집되지 않은 센서 노드는 이전에 수집된 데이터와 동일하다고 판단한다.

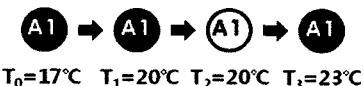


그림 2 시간적인 제한 기법

3.2 공간적인 데이터 제한 기법

클러스터링 기법을 이용한 공간적인 데이터 제한은 클러스터 헤드에 의해 제어된다. [11]은 인접한 노드를 클러스터로 구성하고, 클러스터 헤드는 수집된 데이터의 중복성을 제거하여 싱크 노드로 데이터를 전송한다. 그림 3은 [15]에서 사용된 클러스터 기반의 데이터 제한 기법의 예를 나타낸다. 두 개의 클러스터로 구성된 센서 네트워크가 있다고 할 때, 각 클러스터의 헤드 노드는 클러스터 멤버로부터 수집한 데이터의 중복성을 파악하여 중복된 데이터를 제거한다. 그림 3(a)와 같이 클러스터를 구성한 경우, 클러스터 A는 {A1~A7}의 데이터 중 {A1, A2, A4, A6}, {A3, A5}의 중복되는 센서 데이터를 제거하여 {A1, A3, A7}의 데이터만을 싱크 노드에게 전송한다. 이것은 클러스터 내에 동일한 값을 수집한 센서 노드들이 많을수록 효율이 높아진다는 것을 의미한다. 동일한 데이터를 수집한 상황이라 할지라도 클러스터를 어떻게 구성하느냐에 따라 압축 효율은 크게

달라진다. 하지만, 기존의 클러스터링 기법의 경우, 수집한 데이터의 값과 무관하게 센서 노드의 위치를 기반으로 클러스터를 구성하기 때문에 압축 효율이 낮아지는 경우가 빈번히 발생한다. 그림 3(b)는 클러스터 구성에 따라 압축 효율이 낮아지는 대표적인 예이다. 노드 B2, B3 과 노드 A5, A6은 다른 클러스터 멤버 노드들과 비교해 센서 데이터의 유사도가 낮다. 다시 말해, 클러스터 A의 데이터 발생은 {A1, A2, A4}의 동일한 데이터 발생 집단과 {A3, A7}, {B2}, {B3}의 다른 클러스터 멤버 노드들과 데이터 유사도가 낮은 센서들로 구성되어 있다. 따라서 A3, A7, B2, B3 노드는 각자 자신의 데이터를 싱크 노드로 전송해야 하므로 데이터 제한 발생 확률이 낮아진다. 또한, 그림 3(a)와 같이 유사한 데이터를 가진 센서 노드들끼리 클러스터를 형성하더라도 실제 응용에서는 수집되는 센서 데이터가 잡음을 포함하기 때문에 중복 제거 방법의 효율이 그리 높지 않다.

4. 제안하는 클러스터링 기법

4.1 특징

본 논문에서는 공간적인 데이터 제한 기법의 효율을 높일 수 있는 “데이터 기반 클러스터링 알고리즘(Data-Centric Clustering Algorithm)”을 제안한다. 앞장에서 언급했듯이 공간적인 데이터 제한 기법의 효율을 높이기 위해서 동일한 데이터를 수집한 센서 노드를 클러스터로 구성하는 것이 중요하다. 또한 보다 압축 효율을 높이기 위해서는 중복된 값을 제거하는 기존의 방식보

다 정교한 압축 기법의 제안이 필요하다. 제안하는 기법은 크게 데이터 기반 클러스터링 기법과 공간적인 데이터 제한 기법으로 구성된다.

4.2 데이터 기반 클러스터 생성 과정

그림 4(a)는 기존 클러스터링 기법의 클러스터 생성 과정을 나타낸다. 기존 클러스터 생성 과정은 크게 (1) *Advertisement*, (2) *Join*, (3) TDMA 스케줄 배포, (4) 데이터 수집 과정으로 구성된다. *Advertisement* 과정은 센서 노드가 클러스터 헤드로 선출되기 위해서 주위로 *Advertisement* 메시지를 방송(broadcast) 하는 것을 의미하며, 메시지를 수신한 센서 노드들은 메시지를 전송한 센서 노드를 클러스터 헤드로 선정하기 위해서 *Join* 메시지를 전송하게 된다. *Join* 메시지를 수신한 클러스터 헤드는 TDMA 방식을 이용한 데이터 수집을 하기 위해서 클러스터를 구성하는 노드의 수를 기반으로 TDMA 스케줄 테이블을 작성하고 클러스터 내부로 방송(broadcast) 한다. 따라서 클러스터 헤드는 배포한 TDMA 스케줄에 따라 자신의 클러스터 멤버 노드들의 모든 데이터를 수집한다. 그림 4(b)는 제안하는 클러스터링 기법의 클러스터 생성과정을 나타낸다. 수집된 데이터를 기반으로 클러스터를 생성하기 위해서 클러스터 생성 이전에 각 센서 노드들은 자신의 데이터 수집과정을 먼저 수행한다. 그 다음, *Advertisement* 과정과 *Join* 과정을 거쳐 클러스터를 구성하는 노드의 수를 기반으로 TDMA 스케줄을 작성하고, 작성된 스케줄에 따라 클러스터 헤드가 자신의 클러스터 멤버 노드들의 데

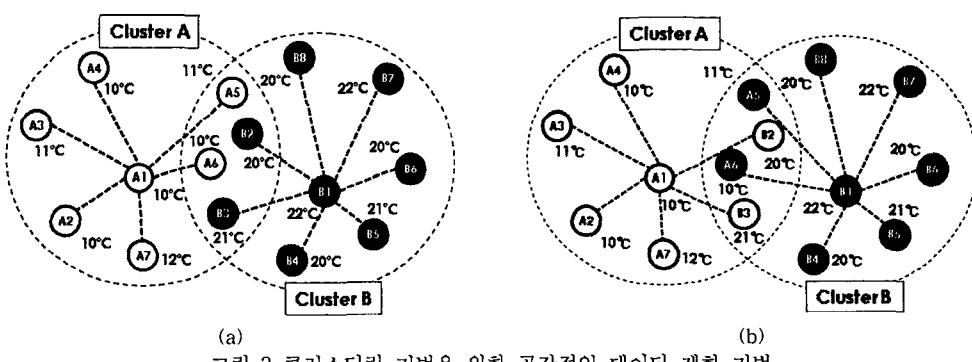


그림 3 클러스터링 기법을 위한 공간적인 데이터 제한 기법

지연시간(전여에너지) + Advertisement	Join	TDMA 스케줄 배포	데이터 센싱	데이터 수집
--------------------------------	------	-------------	--------	--------

(a) 기존 클러스터링 기법

데이터 센싱	지연시간(수집데이터+전여에너지) + Advertisement	Join	TDMA 스케줄 배포	데이터 수집
--------	--------------------------------------	------	-------------	--------

(b) 제안하는 클러스터링 기법

그림 4 클러스터링 생성 과정

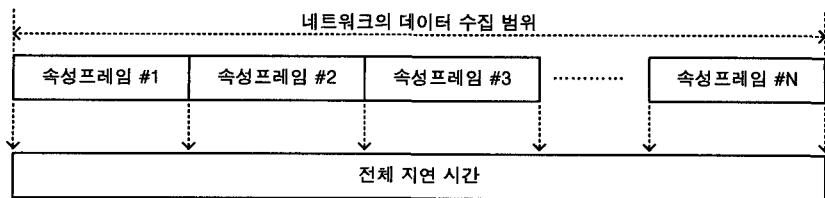


그림 5 수집된 데이터의 범위에 따른 지연 시간 결정

이터를 수집한다. 이때, *Advertisement* 과정에서 유사한 데이터를 수집한 센서 노드끼리 클러스터를 생성하기 위해서 일부 과정을 수정하였다.

제안하는 기법은 수집된 데이터가 속하는 범위에 따라 별도의 클러스터 구성성을 목표로 한다. 따라서 센서 노드들은 자신이 수집한 데이터에 따라 클러스터의 *Advertisement* 메시지 방송 시점을 조정한다. 그림 5는 수집된 데이터의 범위에 따라 각 센서 노드가 자신의 지연시간을 결정하는 방법을 나타낸다. 전체 네트워크의 데이터 수집 범위는 센서 노드를 배포할 때, 응용 프로그램의 특성을 고려하여 미리 정의된다. 따라서 각 센서 노드는 네트워크의 수집 데이터 범위를 N개의 속성 프레임(Attribute Frame)으로 나누고, 각각의 속성 프레임을 전체 지연 시간으로 매핑시킨다. 이때, 중간의 어떠한 계산도 필요하지 않으며, 센서 노드는 오직 자신이 생성한 데이터에 의해 자신이 속하는 속성 프레임을 선택할 수 있고, 스스로 자신의 지연시간을 결정한다.

기존의 클러스터링 기법의 경우, 잔여에너지가 많은 센서 노드를 클러스터 헤드의 선출 확률 CH_{prob} 을 계산하기 위해서식 (1)을 이용한다[8,10]. 이때, CH_{prob} 는 초기 클러스터 헤드의 선출을 제한하기 위해서 사용되는 값이고, E_{max} 는 최대 에너지, $E_{residual}$ 은 센서 노드의 잔여 에너지를 의미한다.

$$CH_{prob} = C_{prob} \times \frac{E_{residual}}{E_{max}} \quad (1)$$

제안하는 기법 역시식 (1)을 이용하여 잔여에너지가 많은 센서 노드를 클러스터 헤드로 선출한다. 모든 센서

노드는 수집한 데이터에 따라 자신의 속성 프레임을 결정하고, 자신의 속성 프레임이 동작하는 시기에 송/수신 기를 활성화 시켜 *Advertisement* 메시지를 송/수신하게 된다. *Advertisement* 메시지를 수신한 센서 노드는 클러스터 헤드로 Join 메시지를 송신하게 된다. 결국 추가적인 통신 비용 없이 동일한 속성 프레임, 즉 유사한 데이터 범위를 가진 센서 노드간의 클러스터 생성이 이루어진다.

4.3 데이터 기반 TDMA 스케줄링 기법

클러스터링 기법은 클러스터 내부의 센서 노드간의 필요 없는 데이터 수신 및 통신 충돌을 방지하기 위해서 TDMA 스케줄에 따라 데이터를 수집한다. 제안하는 기법 역시 TDMA 스케줄을 이용하여 데이터를 수집하게 된다. 이때, 데이터 범위에 따라 형성된 클러스터의 독립적인 통신을 보장하고, 통신의 효율성을 높이기 위해서 그림 6과 같이 전체 TDMA 스케줄을 조정하였다. 제안하는 기법은 2단계의 스케줄 구조를 이용한다. 먼저 전체 데이터 수집 시간을 각 데이터 범위와 매핑되는 슈퍼 프레임(Super Frame) 단위로 나눈다. 그리고, 각 슈퍼 프레임은 같은 데이터 범위를 가진 클러스터들로 구성된다. 이때, 각 클러스터를 슈퍼 프레임에 속해 있는 각각의 프레임(Frame)으로 간주한다. 모든 센서들은 자신이 포함된 프레임에서 상태를 활성화(Active)시키고 데이터 송/수신을 수행함으로 데이터 범위에 따른 독립적인 통신을 보장한다.

4.4 수정된 공간적인 데이터 제한 기법

제안하는 공간적인 데이터 제한 기법은 클러스터 해

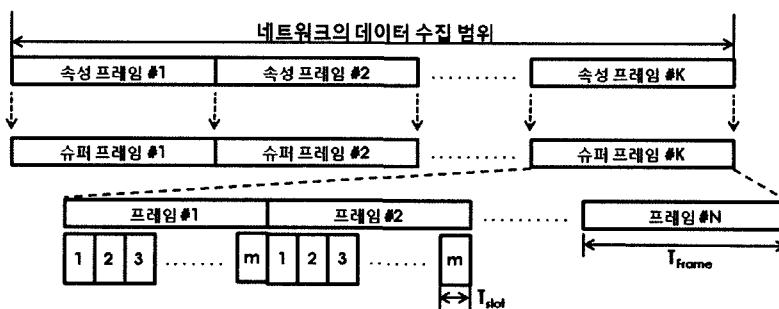


그림 6 데이터 기반 TDMA 스케줄 기법

드에서 이루어지며, 기존의 공간적인 데이터 제한 기법에 비해 싱크 노드에게 전송되는 데이터 크기를 더 많이 줄인다. 클러스터를 생성하는 단계에서 클러스터를 이미 동일한 범위를 가진 센서 노드들로 구성하였기 때문에 공간적인 데이터 제한을 적용할 때 효율이 높아진다. 하지만 보다 높은 압축 효율을 위해서 몇 가지 단계를 추가하였다. 그림 7(a)는 제안하는 공간적인 데이터 제한 기법의 수행 과정을 나타낸다. 먼저 속성 범위의 가장 작은 값을 “대표값”으로 설정하고, 수집된 데이터와 차이를 구한다. 그 다음, 이 값을 양자화하여 데이터 크기를 줄인다. 그리고 기존 방법과 동일하게 수집 데이터 중 중복된 정보를 제거함으로써 압축의 효율을 높인다. 예를 들면, 그림 7(b)와 같이 온도를 측정하는 센서 네트워크의 10도~20도 범위의 프레임에 할당된 센서 노드들이 있다고 가정할 때, 제안하는 공간적인 데이터 제한 기법에 따라 데이터를 수집한 클러스터 헤드는 속성 범위에 해당하는 대표 값과 수집 데이터의 차를 계산한 후 그 값을 양자화한다. 이때 N0, N3, N5는 동일한 데이터를 가지기 때문에 데이터 크기를 줄이기 위해 중복된 정보를 제거하여 (0x0000, {N0, N3, N5})를 싱크 노드로 전송한다.

5. 성능 평가

제안하는 데이터 기반 클러스터링 기법의 우수성을 보이기 위해서 대표적인 클러스터링 기법인 LEACH[12] 기법을 기반으로 본 논문에서 제안하는 시/공간적인 데이터 제한 기법, 데이터 기반 TDMA 스케줄링 기법을 적용하여 네트워크의 수명과 싱크 노드에서 수집된 데이터의 크기를 비교하였다. 또한, 클러스터링 기법에 영향을 끼칠 수 있는 파라미터를 변화시키면서 다양한 실

표 1 성능 평가 환경

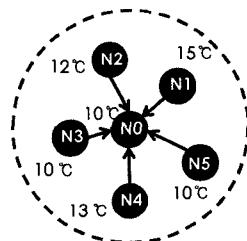
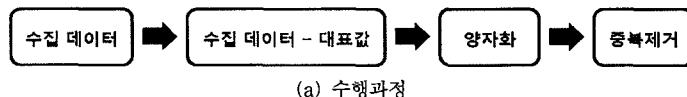
파라미터	값
센서 네트워크의 크기	100m × 100m
싱크 노드의 위치	(50, 175)
센서 노드의 개수	50~400 (기본값=100)
수집 데이터의 크기	4byte
Join/Advertisement 메시지의 크기	16byte
센서 ID의 크기	12byte
슈퍼 프레임의 개수	2~60 (기본값=10)
데이터 변경율	25% / 1라운드
데이터 유사도	25% / 1라운드
클러스터의 크기	0.1~0.8배 (기본값=0.25배)

험을 수행하였다. 시뮬레이션 환경은 표 1과 같다.

5.1 기존 기법과의 비교

그림 8은 각 노드에서 전송하는 평균 데이터 크기를 나타낸다. 기존 기법 중 LEACH는 단순 클러스터링 기법으로 데이터 제한이 일어나지 않기 때문에 다른 기법들에 비해 전송되는 데이터가 많다. 시간/시공간 데이터 제한 기법의 경우에는 센서 노드에서 데이터 제한이 일어나기 때문에 LEACH보다 전송되는 데이터의 크기가 감소한다. 제안하는 기법 역시 시공간 데이터 제한 기법을 사용하기 때문에 LEACH보다 데이터 전송량이 감소하지만, 시공간 제한 기법과 비교하여 상대적으로 클러스터의 수가 증가하기 때문에 평균 데이터 크기가 오히려 증가한다.

그림 9는 네트워크의 수명(lifetime)을 비교한 것이다. LEACH 기법을 사용한 경우와 시간/시공간 제한 기법을 적용한 경우에 비해 제안하는 클러스터링 기법의 네트워크 수명이 연장되었다. 기존의 클러스터링 기법을 기반으로 할 경우, 클러스터를 구성하는 단계에서 이웃 노드간의 도청(overhearing)이 발생하여 전체 네트워크



	N0	N1	N2	N3	N4	N5
수집데이터	10°C	15°C	12°C	10°C	13°C	10°C
차이값 대표값: 10°C	+0°C	+5°C	+2°C	+0°C	+3°C	+0°C
양자화	0000	0101	0010	0000	0011	0000
중복제거	(0x0000, {N0, N3, N5}), (0x0101, {N1}), (0x0010, {N2}), (0x0011, {N4})					

(b) 적용의 예

그림 7 제안하는 공간적인 제한 기법

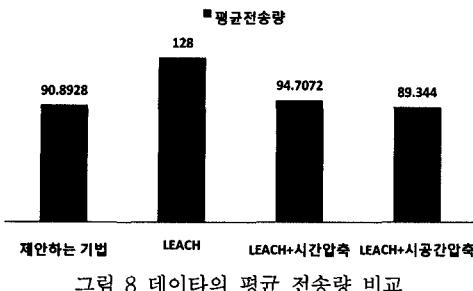


그림 8 데이터의 평균 전송량 비교

수명이 단축된다. 하지만 제안하는 클러스터링 기법의 경우, 데이터 기반 TDMA 스케줄링 기법을 통해 독립적인 통신을 보장하였으며 그 결과 전체적인 네트워크의 수명이 연장되는 결과를 가져왔다.

5.2 슈퍼 프레임의 개수

그림 10과 그림 11은 각각 제안하는 기법의 슈퍼 프레임의 개수 변화에 대한 데이터의 평균 전송량과 네트워크의 수명을 비교한 것이다. 슈퍼 프레임의 개수는 2개~60개로 변화시키면서 실험하였다. 슈퍼 프레임의 개수는 센서 노드를 배포하기 이전에 설정하는 속성 값을 나누는 범위의 수를 의미하며 슈퍼 프레임의 개수가 많을수록 각 슈퍼 프레임에 할당되는 속성 값의 범위는 좁아진다. 수집 데이터의 크기의 경우, 슈퍼 프레임의 수가 증가하는 경향을 가지고 있다. 이것은 슈퍼 프레임의 수가 많아질수록 클러스터의 수가 많아지고, 싱크 노드와 통신하는 클러스터 헤드의 수가 증가하기 때문이다. 그리고, 증가한 클러스터 헤드의 수만큼 싱크 노드

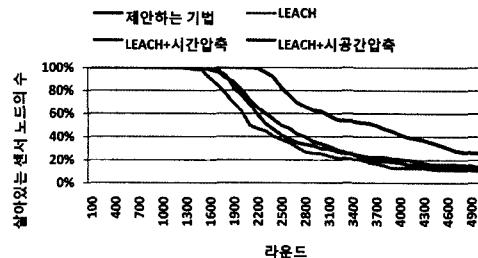


그림 9 네트워크 수명(lifetime) 비교

와의 직접적인 통신이 증가하기 때문에 그림 11과 같이 네트워크의 수명이 감소한다. 슈퍼 프레임의 수가 10개 일 때, 가장 좋은 성능을 보인다.

5.3 센서 밀집도

그림 12와 그림 13은 각각 제안하는 기법의 센서 밀집도 변화에 대한 데이터의 평균 전송량과 네트워크 수명(lifetime)을 나타낸다. 센서의 밀집도는 전체 네트워크 면적에 센서 50개(0.005)~400개(0.04)를 균일하게 배포한 상황을 가정으로 변화시키면서 실험하였다. 센서의 밀집도가 높을수록 같은 반경 내에 수집 대상이 되는 노드의 수는 증가한다. 즉, 클러스터 당 클러스터 멤버의 수가 증가하게 된다. 따라서 전체 네트워크의 동일한 데이터 유사도에 대해 센서 밀집도가 높은 네트워크 일수록 데이터 수신에 따른 클러스터 헤드의 에너지 소모가 증가하기 때문에 전체 네트워크의 수명이 감소된다. 센서의 밀집도를 100개(0.01)로 설정한 경우, 가장 좋은 성능을 보인다.

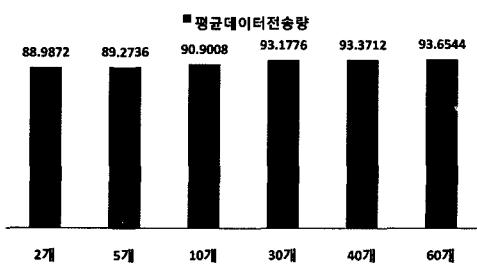


그림 10 데이터의 평균 전송량 비교

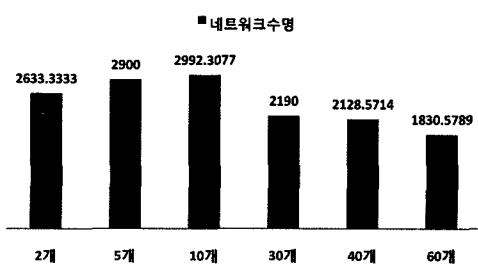


그림 11 네트워크 수명(lifetime) 비교

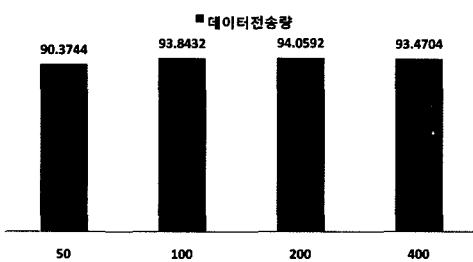


그림 12 데이터의 평균 전송량 비교

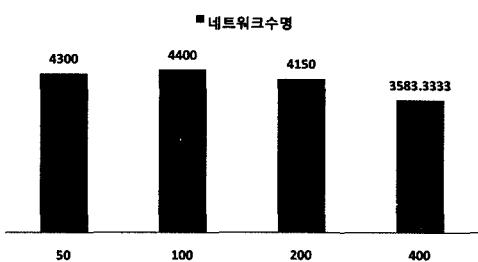


그림 13 네트워크 수명(lifetime) 비교

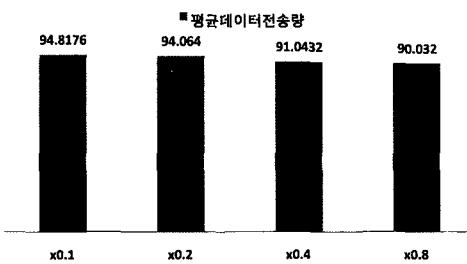


그림 14 데이터의 평균 전송량 비교

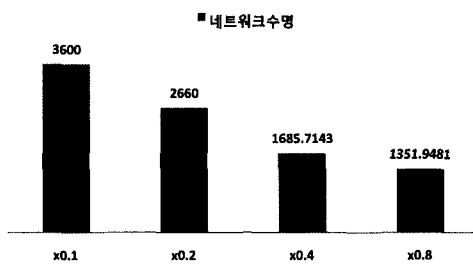


그림 15 네트워크 수명(lifetime) 비교

5.4 클러스터 크기

그림 14와 그림 15는 각각 제안하는 기법의 클러스터 크기 변화에 대한 데이터의 평균 전송량과 네트워크의 수명을 나타낸다. 슈퍼 프레임의 수와 밀집도를 고정하고, 클러스터의 크기는 전체 네트워크의 0.1배~0.8배로 조정한다. 데이터의 평균 전송량의 경우, 동일한 밀집도에서 클러스터의 크기가 커질수록 포함하는 멤버의 수가 많아지는 대신 클러스터의 수가 줄어들고, 공간 제한의 효율이 좋아지기 때문에 평균 데이터의 전송량은 약간 감소한다. 하지만, 클러스터 크기가 커지게 되면 클러스터 멤버 수가 증가하게 되고, 상대적으로 클러스터 헤드와 클러스터 멤버간의 거리가 멀어지게 된다. 따라서 클러스터 크기가 커질수록 전체 네트워크 수명이 단축되는 결과를 가져온다. 클러스터의 크기가 0.1배일 때, 클러스터의 수는 증가하지만 클러스터 헤드와 멤버간의 평균 거리가 가장 짧기 때문에 가장 좋은 성능을 보인다.

6. 결 론

본 논문에서는 무선 센서 네트워크에서 네트워크 트래픽을 감소시키기 위한 데이터 기반 클러스터링 기법을 제안했다. 기존의 클러스터링 기법의 경우 센서 분포의 지역성(locality)만 고려하여 클러스터를 생성하기 때문에 데이터 제한(data suppression) 기법을 적용할 경우 효과적이지 못하다. 또한 단순히 유사한 데이터를 가진 노드들을 클러스터로 생성한 경우, 클러스터 간 겹침이 크게 발생하고, 불필요한 데이터 수신으로 인해 네트워크 수명이 단축되는 문제점을 가지고 있다. 본 논문은 수집한 데이터를 기반으로 클러스터를 생성한다. 이 과정에서 클러스터의 특성을 이용한 TDMA 스케줄링 기법을 통해 클러스터간 겹침 문제를 해결하고, 공간적인 데이터 제한 기법의 효율을 높임으로써 전체적인 네트워크 수명이 연장되었다. 실험 결과, 기존 클러스터링 기법에 비해 최종 싱크 노드에서 수집하는 데이터의 크기가 약 4%~40% 정도 줄어들었고, 네트워크의 수명이 약 20%~30% 연장되었다. 향후 연구는 클러스터 헤드 간의 협력을 통한 보다 효율적인 클러스터 기반의 네이

타 압축 기법을 제안하는 것이다.

참 고 문 헌

- [1] D. Estrin, L. Girod, G. Pottie and M. Srivastava, "Instrumenting the World with Wireless Sensor Networks," In Proceedings of International Conference Acoustics, Speech, and Signal Processing, 2001.
- [2] G. J. Pottie and W. J. Kaiser, "Wireless Integrated Network Sensors," In Proceedings of Comm. ACM, pp. 51-58, May 2000.
- [3] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A Survey on Sensor Networks," In Proceedings of IEEE Communications Magazine, 2002.
- [4] D. Estrin, R. Govindan, J. Heidemann and S. Kumar, "Next Century Challenges: Scalable Coordination in Sensor Networks," In Proceedings of the Mobile Computing and Networking, Seattle, WA., pp. 263-270, Aug. 1999.
- [5] N. Bulusu, D. Estrin, L. Girod and J. Heidemann, "Scalable Coordination for Wireless Sensor Networks: Self-Configuring Localization Systems," In Proceedings of the Sixth International Symposium on Communication Theory and Applications, Ambleside, Lake District, UK, Jul. 2001.
- [6] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," In Proceedings of IEEE Transactions on Information Theory, vol. IT-46, no. 2, pp. 388-404, Mar. 2000.
- [7] A. Silberstein, R. Braynard and J. Yang, "Constraint Chaining: On Energy-Efficient Continuous Monitoring in Sensor Networks," In Proceedings of the ACM SIGMOD International Conference on Management of Data, pp. 157-168, Jun. 2006.
- [8] M. Sharaf, J. Beaver, A. Labrinidis and P. Chryanthis, "Tina: A scheme for temporal coherency-aware in-network aggregation," In Proceedings of the 2003 ACM Workshop on Data Engineering for Wireless and mobile Access, Sept. 2003.
- [9] X. Meng, L. Li, T. Nandagopal and S. Lu, "Event contour: An efficient and robust mechanism for tasks in sensor networks," In Proceedings of Technical report, 2004.

- [10] S. Pattem, B. Krishnamachari and R. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," In Proceedings of International Conference on Information Processing in Sensor Networks, 2004.
- [11] D. Petrovic, R. Shah, K. Ramchandran and J. Rabaey, "Data funneling: Routing with aggregation and compression for wireless sensor networks," In Proceedings of the 2003 IEEE Sensor Network Protocols and Applications, May 2003.
- [12] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," In Proceedings of the Hawaii International Conference on System Sciences, pp. 3005-3014, Jan. 2000.
- [13] W. R. Heinzelman, A. P. Chandrakasan and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," In Proceedings of IEEE Transactions on Wireless Communications, pp. 660-670, Oct. 2002.
- [14] O. Younis and S. Fahmy, "Distributed clustering in adhoc sensor networks: A hybrid, energy-efficient approach," In Proceedings of IEEE INFOCOM, pp. 366-379, Mar. 2004.
- [15] J. Kamimura, N. Wakamiya and M. Murata, "Distributed Clustering Method for Energy-Efficient Data Gathering in Sensor Networks," In Proceedings of the 1st IEEE Communications Society Conference (SECON 2004), Oct. 2004.



여명호

2004년 2월 충북대학교 정보통신공과 공학사. 2006년 2월 충북대학교 정보통신공학과 공학석사. 2006년~현재 충북대학교 정보통신공학과 박사과정. 관심분야는 메인 메모리 기반 데이터베이스 시스템, 시공간 데이터베이스 시스템, 무선 센서 네트워크 등



이미숙

2007년 2월 충북대학교 정보통신공학과 공학사. 2007년~현재 충북대학교 정보통신공학과 석사과정. 관심분야는 무선 센서 네트워크, 데이터베이스 시스템 등



박종국

2007년 2월 충북대학교 정보통신공학과 공학사. 2007년~현재 충북대학교 정보통신공학과 석사과정. 관심분야는 데이터베이스 시스템, 무선 센서 네트워크 등



이석재

2000년 충북대학교 정보통신공학과(공학사). 2002년 충북대학교 정보통신공학과(공학석사). 2006년 충북대학교 정보통신공학과(공학박사). 2006년~2007년 7월 충북대학교 BK21 Post-Doc. 2007년 08 월~현재 애플테크(주) 기술연구소장. 관심분야는 DBMS, 저장 시스템, 주기억장치 DBMS, 센서 네트워크 등



유재수

1989년 2월 전북대학교 컴퓨터공학과 공학사. 1991년 2월 한국과학기술원 전산학과 공학석사. 1995년 2월 한국과학기술원 전산학과 공학박사. 1995년 2월~1996년 8월 목포대학교 전산통계학과 전임강사. 1996년 8월~현재 충북대학교 전기전자컴퓨터공학부 및 컴퓨터정보통신연구소 정교수. 관심분야는 데이터베이스 시스템, 센서네트워크 데이터 관리, 멀티미디어 데이터베이스, 바이오 인포매티스 등