

상황변화에 따른 엔트로피 기반의 클러스터 구성 알고리즘

최윤정^{1*}

¹서일대학 정보통신과

Efficient Clustering Algorithm based on Data Entropy for Changing Environment

Yun-Jeong Choi^{1*}

¹Department of Information Communication, Seoil University

요약 무선 센서네트워크를 효율적으로 운영하고 관리하기 위해서는 센서노드들이 에너지를 효율적으로 사용하도록 프로세스를 설계하고 운영하는 것이 매우 중요하다. 최소한의 자원으로 무인 동작되는 센서 시스템에 결함이 발생하거나 디바이스의 전력이 소진된다면 전체 네트워크에 치명적인 영향을 미치므로, 변화하는 환경을 감지하여 상황에 따라 처리하여 신뢰도를 높이기 위한 연구가 활발하다. 본 논문에서는 센서데이터들의 변화에 따라 환경변화를 인지하고 능동적으로 클러스터링을 재 수행 하게 함으로써 에너지 효율을 높일 수 있는 알고리즘을 제안하고 있다. 상황변화에 따라 변화하는 클러스터의 안정도를 물리량의 무질서 정도를 나타내는 엔트로피의 개념을 이용하여 기존의 확률적 접근방식을 보완하여 클러스터링에 활용하고 있다. 또한, 클러스터링 기반의 LEACH 알고리즘과 엔트로피를 이용하는 제안방법과의 비교를 위해 상황에 따른 조건을 정하여 실험하고 전체 센서노드의 생존율을 체크하여 실험 결과를 비교분석 하였다.

Abstract One of the most important factors in the lifetime of WSN(Wireless Sensor Network) is the limited resources and static control problem of the sensor nodes. In order to achieve energy efficiency and network utilities, sensor nodes can be well organized into one cluster and selected head node and normal node by dynamic conditions. Various clustering algorithms have been proposed as an efficient way to organize method based on LEACH algorithm. In this paper, we propose an efficient clustering algorithm using information entropy theory based on LEACH algorithm, which is able to recognize environmental differences according to changes from data of sensor nodes. To measure and analyze the changes of clusters, we simply compute the entropy of sensor data and applied it to probability based clustering algorithm. In experiments, we simulate the proposed method and LEACH algorithm. We have shown that our data balanced and energy efficient scheme, has high energy efficiency and network lifetime in two conditions.

Key Words : Sensor Network, Routing Algorithms, Clustering Algorithms, Sensor Data, Entropy

1. 서론

유비쿼터스 환경에서 USN(Ubiquitous Sensor Network)은 국내외적으로 무선 센서네트워크(Wireless Sensor Networks)라는 이름으로 오래 전부터 많은 연구가 진행되어 온 분야 중 하나이다. 이는 기존의 유선으로 구축된 센서네트워크를 무선 네트워크로 대체하는 기술

로서, 기존의 단순한 무선 네트워크 체제와는 차이가 있다. USN은 다수의 센서노드로 구성된 무선 네트워크로 다양한 위치에 설치된 센서노드로부터 환경 정보를 수집하고 이를 자유롭게 이용할 수 있는 정보서비스의 인프라를 뜻한다[1,2]. 센서노드들은 응용분야의 성격과 환경에 따라 고정위치에 혹은 임의의 위치에 무작위로 배치된다. 환경적인 특성상 배터리 교환이나 충전이 어렵고

*교신저자 : 최윤정(cris@seoil.ac.kr)

접수일 09년 09월 29일

수정일 (1차 09년 10월 30일, 2차 09년 11월 30일)

게재확정일 09년 12월 16일

한정된 자원으로 노드간 멀티 홉 통신을 하게 되므로 최대한으로 오래 살아있으면서 정보수집기능을 유지하는 것이 관건이다.

센서노드의 에너지 소모를 최소화하기 위한 기법으로 노드의 계층적인 구조를 이용한 클러스터링 기법이 제안되었고 많은 연구가 이루어져 왔다[3,4,6]. 클러스터링 이후 다수의 노드들로 그룹핑 된 각 클러스터 내에는 일반적으로 하나의 클러스터 헤더가 존재하게 되며, 이들은 하나의 일정 영역인 해당 클러스터를 대표하는 노드로서 동작한다. 따라서 클러스터내의 나머지 멤버 노드는 일정 기간 휴지상태(Sleep)를 유지하며 에너지 소모를 줄일 수 있는 것이다. 이 때, 클러스터 헤더는 멤버 노드와의 주기적인 메시지 전송을 통해 동기화를 유지하고 멤버 노드의 스케줄 즉, 활동/휴지기간을 조절하는 역할을 해야 한다. 또한 에너지균형화 (Load Balancing)와 데이터 통합(Data Aggregation)을 통해 넓은 공간에 높은 밀도로 분포된 다수의 노드들이 안정적인 네트워크 토폴로지를 유지하게 하며, 에너지 소비를 최소화 하면서 한정된 네트워크 자원을 효율적으로 활용할 수 있도록 해야 한다. 따라서 클러스터 헤더는 멤버 노드들에 비해 에너지 소모가 커서 생명주기가 짧아지므로 교체가 필요하며 클러스터 헤더의 효율적인 변경 및 선정을 위한 다양한 기법들이 연구되고 있다. 계층구조망에 대한 연구로서 LEACH(Low-Energy Adaptive Clustering Hierarchy)가 대표적이고, LEACH-C, EACHS, HEED, TEEN(Threshold Sensitive Energy Efficient Sensor Network), PEGASIS 프로토콜에 이르는 여러 알고리즘들이 제안되었으며 이들의 알고리즘을 개선한 연구들이 보고되어 있다[4,5,7,8]. 이들은 센서노드들의 잔존 에너지량을 검사하여 클러스터 헤더를 변경함으로써 수명연장을 기대하고 있다. 그러나 초기배치가 이루어지고 난 이후상황과 환경변수들은 클러스터가 생성된 이후에도 얼마든지 변경될 가능성이 높는데 비해서 센서노드가 배치된 위치나 분포, 센서데이터들의 변화량에 따른 문제점에 대해서 근본적인 해결방법을 제시하고 있지는 않다. 대부분 노드들이 초기 배치된 이후의 유동적인 상황 변화를 인지하여 동적으로 대처하기 어려운 구조를 띄고 있기 때문이다. 이는 결국 클러스터의 불안전성을 높이고 센서노드에 심각한 전력손실을 일으키는 요인이 되며 나아가 시스템의 가용성을 떨어뜨리는 원인으로 지적된다. 따라서 노드들의 잔존 에너지 뿐 만 아니라, 노드의 공간상의 위치정보, 센서가 수집한 데이터의 변화에 따라 클러스터링을 재구성하여 환경변화에 유연하게 대처할 수 있는 고효율적인 알고리즘들에 대한 연구의 필요성이 높다[9,11].

본 논문에서는 센서데이터들의 변화에 따라 환경변화

를 동적으로 인지하고 능동적으로 클러스터링을 재 수행하게 함으로써 에너지 효율을 높일 수 있는 알고리즘을 제안하고 있다. 상황변화에 따라 변화하는 클러스터의 안정도를 물리량의 무질서 정도를 나타내는 엔트로피의 개념을 이용하여 기존의 확률적 접근방식을 보완하는 것이다. 이는 센서노드의 위상이나 센서데이터의 변화의 양상을 예측할 수 없는 네트워크 상황에서 효율성을 증대시키기 위해 보다 더 적합한 방법을 찾아내기 위한 방법으로 의미를 가진다.

본 논문은 구성은 다음과 같다. 2장에서는 관련연구로서 기존의 대표적인 클러스터링 알고리즘과 제안내용에 요구되는 기반기술에 대해 설명한다. 3장에서는 엔트로피 개념을 이용하여 센서노드로부터 수집된 데이터들의 변화 상태에 따라 초기의 클러스터 상태를 재구성할 수 있는 EOC(Entropy-based Optimal Clustering Algorithms) 알고리즘을 제안한다. 4장에서는 기존의 LEACH알고리즘과 제안방법으로 보완한 방법의 생존을 비교실험으로 성능을 평가하며 5장에서 결론을 맺는다.

2. 관련 연구

2.1 클러스터 헤더 선정방법

클러스터 기반의 계층적 라우팅 프로토콜에서 클러스터내의 센서노드들은 센싱 기능을 담당하는 멤버이거나 멤버들의 데이터를 수집하고 통합하여 싱크나 다른 클러스터 헤더로 데이터를 전송하는 역할을 하는 클러스터 헤더로 나뉜다. 이때, 클러스터 헤더를 선정하는 방식에 따라 중앙형과 분산형으로 구분되는데, 중앙형 방식은 센서필드의 센서들로부터 수집된 데이터를 바탕으로 싱크가 클러스터 헤더를 선정하고, 선정된 헤더를 중심으로 클러스터가 형성되는 방식이며, 분산형은 정해진 확률에 의해 클러스터 헤더가 선출되며 멤버들을 모아 클러스터가 형성되는 방식을 말한다[3,6,8]. 비교적 간단한 방법으로는 지역적인 위치를 고려하여 클러스터를 구성하기도 하는데, 클러스터 헤더는 해당 클러스터의 중앙에 위치하도록 선정하며 이때 에너지의 균형을 위해 각 클러스터내의 노드 수는 거의 일정하게 제한한다. 어느 방법이든 일단 클러스터헤더로 선정되면 클러스터의 크기와 노드의 수에 따라 차이는 있지만 다른 멤버에 비해 훨씬 더 많은 에너지를 소모하게 된다. 따라서 시스템의 생존기간을 늘리기 위해서는 클러스터 헤더의 에너지 소모를 균형있게 조절해 주기 위해 교체해주는 재선정작업이 필요하다.

2.2 클러스터 구성 방법

클러스터를 구성하는 프로토콜은 대부분의 센서 네트워크와 마찬가지로 구분되는데, 센서노드들이 대상 환경을 센싱하고 전송하는 시기가 주기적으로 이루어지는지 혹은 데이터 발생형태에 따라 이루어지는지에 따라 전체 조건을 정하고 프로토콜을 설계한다. 센서노드들이 일정한 시간에 따라 데이터를 주기적으로 센싱하고 동시에 전송하는 과정에서 불필요한 에너지 소모가 발생할 수 있다. 따라서 침입 및 폭발 탐지와 같이 어느 시점에 센싱된 데이터들의 상이한 차이가 나타나는 경우에는 수집된 데이터의 변화에 따라 전송하는 것이 더 효율적일 것이다.

2.3 클러스터링 알고리즘

계층망의 대표적인 프로토콜인 LEACH 에서는 네트워크에 있는 모든 센서노드들의 에너지를 공정하게 분산시키기 위해 주기적으로 일정한 확률에 의해 클러스터 헤더를 임의로 순환하여 선택한다. 시간을 사이클로 나누고 그것을 각 라운드로 다시 나누는데 이때 클러스터 헤더는 특정 라운드에서 확률함수에 의해 계산된 임계값보다 작을 때 선정된다. 여기서 선정된 클러스터 헤더를 기반으로 클러스터가 형성되며 헤더들은 멤버들에게 타임슬롯을 할당하여 데이터를 전송받고, 전송받은 데이터를 융합(Fusion)하여 싱크에게 전달한다. LEACH 알고리즘에 의해 계산된 확률 식에 의해 모든 노드가 한번 씩 클러스터 헤더가 될 수 있으나 주변노드의 상태를 고려하지 않고 있기 때문에 불균형적인 클러스터가 형성될 수 있으며, 클러스터에 포함된 센서노드의 수에 따라 불균형적인 에너지 소모를 초래하게 되는 문제가 있다.

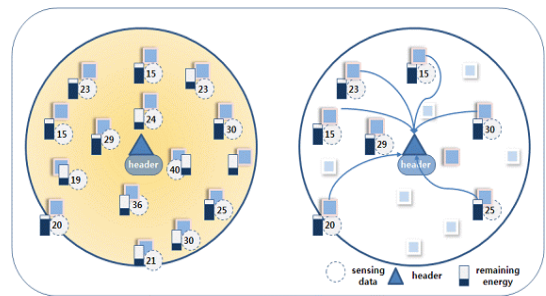
LEACH-C는 클러스터 헤더를 싱크가 직접 결정해주는 중앙형 방식으로써, 기존의 LEACH의 방법에서 잔여 에너지 정보를 이용하여 더 많은 에너지를 가지고 있는 노드의 우선순위를 높여주는 방법을 제안했다. 센서노드들은 자신의 위치와 에너지상태를 싱크에게 전송하고 싱크는 이를 바탕으로 필드 전체의 평균에너지를 계산한다. 평균 이하의 에너지를 가진 노드는 클러스터 헤더후보에서 제외시키고, 센싱 노드들의 전송에너지가 최소가 되도록 일정한 수의 클러스터 헤더를 선정한다. 매 라운드마다 일정한 수의 클러스터를 재구성하는 분산형 방법인 LEACH에서는 클러스터 헤더가 고르게 분포될 때 좋은 성능이 발휘될 수 있으나 언제나 보장되는 상황은 아니다. LEACH-C 또한 상황에 따라 불균형적인 네트워크를 형성할 수 있으며 둘 다 매 라운드마다 클러스터 선출이 일어나고 클러스터가 재형성되므로 불필요한 에너지 소

모가 있을 수 있다는 문제점을 지니고 있다.

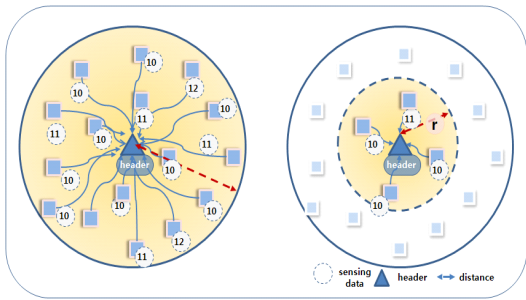
HEED는 LEACH의 문제점을 해결하기 위해서 노드의 에너지 이외의 여러 추가적인 요소를 고려하여 클러스터 헤더를 선출하고 있다[4]. 이 기법은 잔여에너지뿐 아니라 이웃노드에 대한 근접도 같은 다양한 파라미터를 고려하여 적절한 노드의 밀도, 클러스터 내/외의 전송거리를 결정함으로써 등한 분산클러스터를 구성한다.

3. 클러스터 구성 알고리즘

LEACH가 전제하고 있는 조건상으로 모든 노드가 동일하게 에너지를 소비한다는 가정 하에서는 우수하게 동작하지만, 실제 적용된 환경에서는 데이터의 전송상태나 거리 등에 의한 잔존에너지의 상태 등 변수들이 많으므로 단순히 클러스터헤더 선출요소만을 고려한 방법으로는 부족하다. 이후 LEACH-C, EACHS 와 같이 선출 시점에서의 전체 노드의 에너지와 잔존에너지, 평균에너지를 고려하는 방법이 제안되었으나 추가적인 오버헤더가 크고 또 다른 가정조건이 필요하다는 면에서 단점이 있다. 따라서 에너지효율을 높이기 위하여 적용분야와 목표, 노드의 이동성, 그리고 데이터의 발생형태에 따라서 다른 방법이 강구되어야 한다. 예를 들면, 센서노드가 자연환경에 노출되어 설치되는 경우 데이터 및 센서노드 자체가 유실될 위험이 크므로 토폴로지 변화와 주위환경에 대한 강인성을 고려하여 설계해야 할 것이다. 또한 노드가 이동성이 있는 경우 잦은 토폴로지 변화에 따른 전력의 소모가 증가하므로 전체 시스템의 생존기간에 미치는 영향을 최소화 하도록 하는 처리방법을 고려해야 할 것이다. 따라서 클러스터를 형성과정에서는 데이터의 발생형태에 따라 동일한 데이터를 수집한 센서노드들은 같은 클러스터로 구성하는 것이 효율적이며, 클러스터 헤더에서 중복된 값을 제거하는 방식보다 센서노드 차원에서 전송하지 않는 방법을 고려해야 한다.



(a) 클러스터의 불안전성(entropy)이 클 경우



(b) 클러스터의 불안정성(entropy)이 낮은 경우

[그림 1] 클러스터의 불안정성에 따른 데이터의 전송처리

3.1 최적화된 클러스터 구성을 위한 알고리즘

최적화된 클러스터란 같은 클러스터를 이루는 요소들 간의 유사도는 높고 각 개별적인 클러스터들 간의 유사도는 낮게 구성된다. 동일한 데이터들을 수집하는 센서노드들은 서로 같이 그룹핑 되게 하여 정보처리의 효율을 높이는 것이다. 이러한 이상적인 클러스터를 형성하기 위하여 클러스터의 내부 유사도와 외부 유사도의 개념을 적용할 수 있다. 식(1)과 식(2)는 각각 내부 유사도와 외부 유사도 검사를 위한 기준함수를 나타낸다. 센서노드들을 개별적인 클러스터로 초기화하고 각 노드들을 일정 기준에 따라 서로 그룹핑 해가면서 최적의 클러스터 $S = \{ S_1, \dots, S_k \}$ 를 구성하는 것이다.

$$\text{maximize } I_Sim = \sum_{r=1}^k n_r \left(\frac{1}{n_r} \sum_{d_i, d_j \in S_r} sim(d_i, d_j) \right) \quad (1)$$

내부 유사도는 각 클러스터들에 속하는 노드들이 수집한 데이터들 간의 유사도를 최대화 하도록 구성하는 방법이다. 여기서 S_r 은 r 번째 클러스터, 그리고 k 는 총 생성된 클러스터의 수를 나타낸다. n_r 은 r 번째 클러스터를 이루는 센서노드의 수를 말하며 해당 클러스터의 크기를 의미한다.

$$\text{minimize } E_Sim = \sum_{r=1}^k n_r sim(Cr_r, Cr) \quad (2)$$

외부 유사도는 내부 유사도와 다른 개념으로 서로 다른 다양한 클러스터들을 구성하기 위한 방안으로 사용된다. 식 (2)에서 Cr 은 전체 센서노드들의 중심값(Centroid Vector), 그리고 Cr_r 는 r 번째 클러스터의 중심값을 의미한다. n_r 은 r 번째 클러스터를 이루는 노드의 수를 나타낸다.

$$\text{maximize } \frac{I_Sim(\text{Internal Similarity})}{E_Sim(\text{External Similarity})} \quad (3)$$

즉, 식 (2)에서는 전체노드 집합의 중심 값과 각 클러스터의 중심 값의 유사도가 최소화되도록 하여 다양한 클러스터들을 생성할 수 있도록 하는 것이다. 따라서 클러스터의 내부 유사도는 최대가 되고 각 클러스터간의 외부 유사도는 최소가 되는 이상적인 클러스터를 생성하기 위해 식 (3)과 같은 기준함수를 이용하여 최적화된 k 개의 클러스터를 구성한다.

3.2 클러스터의 엔트로피에 의한 전송처리

클러스터 헤더의 역할로 인해 통합 과정에서 인접 노드간 중복된 데이터를 처리하여 싱크에게 전송함으로써 낭비되는 에너지는 감소될 수 있다. 그러나 그 이전 과정인 노드가 클러스터 헤더에게 데이터를 전송하는 과정에서 전송될 데이터를 비교함으로써 불필요한 전송으로 낭비되는 자원을 최소화 할 수 있을 것이다.

그림 1은 데이터가 발생하는 형태에 따라 다른 처리가 필요하다는 개념을 보인다. 그림 1의 (a)의 경우에는 클러스터내의 데이터들의 편차범위가 큰 상태를 나타낸다. 반면, 그림 1의 (b)의 경우에는 주변데이터들의 편차가 거의 적은 상태이며 정보화 관점에서 볼 때 클러스터내의 데이터 유사성이 높은 상태이다. 주변 데이터의 편차가 큰 (a)의 경우에는, 클러스터내의 데이터를 대표할 수 있는 데이터를 전송하기 위해 몇 개의 소수의 노드를 선정하는 일이 거의 불가능하며, 최악의 경우 전체 노드에서 센싱 된 정보를 클러스터에게 빠짐없이 알려야 경우도 생길 것이다. 반면, (b)의 경우에는 클러스터 내에서 수집된 데이터들의 유사성이 높은 형태이므로, 거리 r 내에 있는 소수의 노드를 선정하여 전송 처리를 하고, 나머지 노드는 휴지상태로 전환시켜 에너지를 절약할 수 있도록 할 수 있다.

한 클러스터를 이루는 데이터들이 얼마나 유사한지 가늠하는 척도로서 가장 보편적으로 사용되는 엔트로피(Entropy)의 개념을 정리하면 다음과 같다. 엔트로피는 원래 물질계의 열적상태를 나타내는 물리량으로 무질서 정도를 나타내는 지표로 사용되었다. 정보이론에서는 이런 엔트로피의 개념을 도입하여 상호정보량(Mutual Information)이라는 개념으로 기존의 확률적 접근방식을 보완하고 있다[10]. Shannon이 제안한 엔트로피는 식 (4)와 같이 정의된다. 여기서 S 는 s 를 확률 변수로 갖는 집합을 의미하고 $H(S)$ 는 이 집합의 엔트로피, 그리고 $P(s)$ 는 s 가 발생할 확률 값을 의미한다.

```

// the process of the cluster re-formation
If ( Node N is Cluster Header in Cluster Ci) then
{
    Analyze the Entropy of Ci;
    If (Entropy of Ci is low) then Broadcast Advertisement_Message with threshold r;
    Else Broadcast Advertisement_Message using sensing data + residual energy of N
    Wait for Join_Cluster_Messages;
}

If ( Node N is Not Cluster Header in Cluster Ci) then
{
    Receive all Advertisement_Message;
    Analyze the Entropy of Ci for each Cluster : residual energy and sensing data ;
    Broadcast the Join_Cluster_Message;
}
    
```

[그림 2] 클러스터의 재구성 프로세스

$$H(S) = - \sum_{s \in S} p(s) \log_2 p(s) \quad (4)$$

엔트로피는 확률 변수에 대해 확률이 균등하게 분배되어 있을수록 큰 값을 가지며, 일반적으로 임의의 클러스터의 엔트로피가 낮을수록 내부 자질들의 유사도가 높은 좋은 클러스터임을 나타낸다. 내부 복잡(Internal Complexity)의 개념인 엔트로피는 반대로 클러스터내의 복잡도가 이전보다 높아졌다면 그 클러스터 구성을 변동시킬 수 있는 상황의 변화가 있음을 의미한다. 터의 품질 평가에 유용하게 적용되며, 클러스터의 크기가 n_r 인 임의의 클러스터 S_r 에 대한 일반적인 엔트로피는 다음의 식 (5)와 같이 계산된다.

$$Entropy(S_r) = - \frac{1}{\log q} \sum_{i=1}^q \frac{n_r^i}{n_r} \log \frac{n_r^i}{n_r} \quad (5)$$

여기서 q 는 클러스터의 개수를 나타내며, n_r^i 는 r 번째 클러스터로 그룹핑 된 센서노드들 중 클러스터 i 에 해당하는 노드들의 개수를 의미한다. 이 때, 전체 클러스터 S 에 대한 엔트로피는 개별 클러스터들의 크기에 가중치를 두어 다음과 같이 합산된다. 클러스터내의 복잡도가 낮으면 엔트로피는 1에 가깝게 될 것이고 센싱 된 데이터 상태의 변화가 거의 없는 것으로 판단할 수 있다. 반대로 클러스터내의 복잡도가 이전보다 높아졌다면 그 클러스터 구성을 변동시킬 수 있는 상황의 변화가 있음을 의미한다.

4. 실험결과 및 평가

본 장에서는 앞서 제안한 방법의 성능측정을 위해 기존의 LEACH 알고리즘과의 비교실험을 수행하였다.

이를 위해, 초기 구성된 클러스터내의 데이터 값을 검사하여 엔트로피가 높을 경우와 높지 않을 경우를 가정하였고, 클러스터 재구성이 필요한 시기 및 재구성 방법을 고려하여 시뮬레이션을 수행하였다. 초기 클러스터링 수행방법은 잘 알려져 있는 LEACH 알고리즘의 조건을 따르며 노드의 생존 비율을 비교하였다. 차원영역은 100m*100m로 정하고 노드의 수는 100개로 설정하였다. 초기에너지로 각 총 2J의 자원을 주었으며 500byte의 메시지를 전송한다. LEACH에서 적용한 확률값 P 는 0.05로 고정하고, 잔존에너지는 1-(현재에너지/초기에너지)로 구하였다. 그림 2는 앞장에서 제안한 클러스터의 재구성 프로세스를 나타낸다. 노드 N 이 클러스터 헤더인지 검사하고, 해당 클러스터의 엔트로피를 분석하여 주변 노드에게 어떠한 광고 메시지를 전송할 것인지 혹은 수신할 것인지를 결정한다. 그림 1의 (b)와 같이 초기 클러스터를 구성하는 단계에서, 엔트로피가 높지 않을 경우에는 거리 임계값을 이용하여 일정거리내의 영역에 있는 노드의 전력 상태에 따라 노드를 활성화시키고 휴지상태로 전환한다. 또한, (a)와 같이 엔트로피가 클 경우에는 거리 임계값이 아닌 클러스터의 잔존에너지와 센싱 데이터를 이용하여 노드의 활성/휴지 상태를 정하도록 한다. 각 클러스터의 엔트로피 증가는 노드가 수집한 데이터와 이전데이터의 오차범위를 이용할 수 있는데, 본 실험에서는 클러스터의 복잡도가 0.5 이상일 때 높다고 간주하였고, 이를 복잡도의 임계값으로 사용하였다. 클러스터의 복잡도는 5회마다 반복 계산하여 반영하였다.

Round별 노드의 생존율

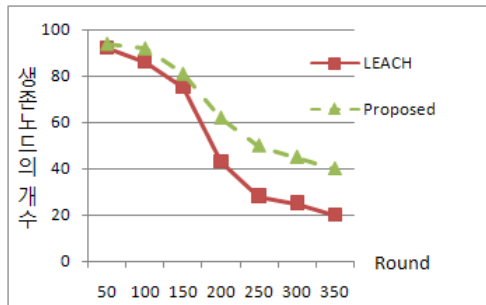
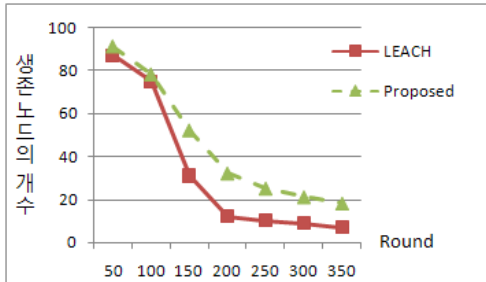
(초기엔트로피>0.5, re-clustering 1회)

Round	LEACH(A)	Proposed(B)	증가율(%)
50	87	91	5%
100	75	78	4%
150	31	52	68%
200	12	32	167%
250	10	25	150%
300	9	21	133%
350	7	18	157%

Round별 노드의 생존율

(초기엔트로피<0.3, re-clustering 1회)

Round	LEACH	Proposed	증가율(%)
50	92	94	2%
100	86	92	7%
150	75	81	8%
200	43	62	44%
250	28	50	79%
300	25	45	80%
350	20	40	100%



- (a) 복잡도가 높은 경우(엔트로피 > 0.5)
- (b) 복잡도가 낮은 경우(엔트로피 < 0.3)

[그림 3] 라운드별 dead node의 개수

그림 3은 기존방법과 제안방법을 적용했을 때의 생존한 노드들의 개수와 단순 증가율을 나타낸다. 그림 3의 (a)와 같이 클러스터의 복잡도가 높은 상황에서 라운드 100이하에서의 생존 노드는 두 방법에서 모두 약간의 증가율을 보인다. 이 때의 에너지 소모량 역시 비슷하게 감

소하고 있음을 알 수 있다. 그러나 클러스터내의 복잡도가 높은 경우 라운드 150 이후에서는 횡수가 증가할수록 노드 생존율이 이전보다 최고 157% 증가하고 있음을 보인다. 이 경우 회를 거듭해갈 수록 해당 클러스터 내/외의 센서데이터의 변화량을 포함한 환경의 변화가 발생했고 에너지가 급격히 소모되고 있음을 알 수 있다. 이러한 동적인 변화에 능동적으로 대처함으로써 기존보다 약 3배에 가까이 증가한 노드 생존율을 보인다. 그림 3의 (b)와 같이 클러스터의 복잡도가 낮은 상황에서는 생존 노드의 수가 최고 100% 증가함을 보이고 있는데 불필요한 비교와 재계산이 수행되지 않으므로 자원소모가 최소화되었음을 나타낸다.

에너지 소모율을 최소한으로 하여 네트워크의 생존기간을 늘릴 수 있는 연구들이 활발한 가운데, [11]은 LEACH를 개선한 PEGASIS 프로토콜의 약점을 보완한 연구로서 잔존 에너지 상태를 기준으로 서브클러스터마다 클러스터 헤드를 선출하고 있다. Greedy 알고리즘을 적용한 이 논문의 실험결과 생존노드에 따른 라운드 수가 12% 연장된 결과를 얻을 수 있었다. 본 연구는 클러스터 내의 센서 노드들이 수집한 데이터 형태를 함께 고려하고 있으므로 잔존 에너지를 변수로 한 위 연구와 직접적인 수치비교에 의한 성능평가는 어렵지만, 재클러스터링이 이루어질 때 클러스터 헤드를 선정하는 개선된 알고리즘을 반영함으로써 계산에 드는 오버헤드를 줄여나가는 연구가 필요할 것으로 생각된다.

5. 결론 및 향후연구

본 논문에서는 센서데이터들의 변화에 따라 환경변화를 동적으로 인지하고 능동적으로 클러스터링을 재 수행하게 함으로써 에너지 효율을 높일 수 있는 알고리즘을 제안하였다. 본 연구에서는 클러스터 헤터를 중심으로 형성된 클러스터 내의 노드에서 수집된 데이터들의 엔트로피를 분석하고 각 노드의 정보를 이용하여 에너지 효율적인 클러스터를 구성한다. 초기 설정하는 방법은 최대한 비용이 낮은 기존의 알고리즘을 택하여 수행한다고 할 때, 상황변화에 따라 변화하는 클러스터의 안정도를 물리량의 무질서 정도를 나타내는 엔트로피의 개념을 이용하여 기존의 LEACH 알고리즘이 갖는 장점인 확률적 접근방식을 보완하였고, 잔여에너지와 센싱 데이터의 정보에 따라서 노드들이 활성/휴지상태가 됨으로써 초기 구성된 클러스터가 상황에 따라 재구성되도록 하고 있다. 이는 변화의 양상을 예측할 수 없는 네트워크 상황에서 효율성을 증대시키기 위해 보다 더 적합한 방법을 찾아

내기 위한 방법으로 의미를 가진다.

추후, 조정이 필요한 시기를 정하여 재 클러스터링을 수행하는 문제에 대해 데이터간의 유사도와 센서노드들의 공간적인 토폴로지를 이용하여 추론할 수 있는 부분과 싱크가 수집한 데이터 수준에서 다시 고려할 수 있는 문제에 대해 정의하여 접근하는 연구가 필요하다.

참 고 문 헌

[1] 최경진, 윤명준, 심인보, 이재용, "무선 센서 네트워크에서의 에너지 효율적인 클러스터 헤드 선출 알고리즘", 한국통신학회논문지, 제 32권, 제 6호, pp. 342~349, 6월, 2007.

[2] W.B. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture For Wireless Microsensor Networks," *Ieee Trans. Wireless Comm.*, Vol.1, No.4, pp.660-670, Oct. 2002.

[3] 김재현, 이재용, 김석규, 도윤미, 박노성, "무선 센서 네트워크에서 동적 예비 클러스터헤드를 이용한 효율적인 토폴로지 관리 방안에 관한 연구," 한국통신학회 논문지, 제31권, 제 6호, pp. 534-543, 6월, 2006.

[4] Ossama Younis and Sonia Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for AdHoc Sensor Networks" *IEEE Transactions on Mobile Computing*, Vol. 3, No. 4, pp.366-379, Oct., 2004.

[5] I. Stojmenovic, A. Nayak, and J. Kuruvila, "Design Guidelines for Routing Protocols in Ad Hoc and Sensor Networks with a Realistic Physical Layer," *IEEE Communication Magazine*, Vol. 43, Issue 3, pp. 101-110, Mar. 2005.

[6] A. Manjeshwar and D.P. Agrawal, "APTEEN: A Hybrid Protocol For Efficient Routing And Comprehensive Information Retrieval In Wireless Sensor Networks", *Int'l Parallel and Distributed Process. Symp. (IPDPS'02) Workshops*, pp.0195b, April 2002.

[7] V. Tran Quang and T. Miyoshi, "ARPEES: Adaptive Routing Protocol For Large Scale Wireless Sensor Networks", *Proc. Commun. Conf. IEICE 2007*, BS-12-3, pp.S-154-S-155, Sept. 2007.

[8] A. Visvanathan, J.-H. Youn, and J. Deogun, "Hierarchical Data Dissemination Scheme For Large Scale Sensor Networks", *IEEE Int'l Conf. Commun. (ICC 2005)*, Vol.5, pp.3030-3036, May

2005.

[9] 남도현, 민흥기, "센서 네트워크에서 클러스터 헤드의 load-balancing을 통한 에너지 효율적인 클러스터링", 정보처리학회논문지, 제14권, 제3호, pp.277~283, 2007

[10] C.E. Shannon, "A Mathematical Theory of Communication", *Bell System Technical Journal*, Vol 27, pp.379-423, 623-656, 1948.

[11] 정수형; 유해영, "서브 클러스터링 기법을 이용한 효과적인 USN 라우팅 프로토콜", 한국산학기술학회논문지, 제9권, 제3호, pp.1647-1652, 2008년 12월

[12] 이창열, "그룹 개념 기반 지능형 USN 미들웨어 플랫폼 연구", 한국산학기술학회논문지, 제9권, 제6호, pp.1666-1672, 2008년 12월

최 윤 정 (Yun-Jeong Choi)

[정회원]



- 2001년 9월 : 이화여자대학교 대학원 컴퓨터학과 (공학석사)
- 2007년 2월 : 이화여자대학교 대학원 컴퓨터학과 (공학박사)
- 2007년 8월 ~ 2008년 2월 : 서강대학교 컴퓨터학과 Post.Doc
- 2009년 3월 ~ 현재 : 서일대학교 정보통신과 강의전담 교수

<관심분야>

인공지능, 기계학습, 온톨로지, 유비쿼터스 센서네트워크