

클러스터 구조의 센서 네트워크에서 효율적인 데이터 모음 기법

지재경^o 하란

홍익대학교 컴퓨터공학과 실시간시스템 연구실

{jkjee^o, rhanha}@cs.hongik.ac.k

An Efficient Data Aggregation Method in Clustered Sensor Network

Jaekyoung Jee^o Rhan Ha

Department of Computer Engineering, Hong-ik University

요 약

배터리를 사용하여 정보의 처리와 센싱 작업을 수행하는 무선 센서 네트워크를 오래 동안 가동시키기 위해서 한정된 자원을 효율적으로 사용할 수 있는 여러 기술들이 제안되고 있다. 이런 기법들 중 클러스터를 구성하거나, 데이터 모음 기법을 수행하여 중복된 데이터를 하나의 패킷으로 압축하여 전송 횟수를 줄이면 에너지 절감 효과를 볼 수 있다. 본 논문에서는 클러스터 구조를 이루고 있는 센서 네트워크에서 두 개 이상의 클러스터가 중복된 지역을 센싱 할 경우 발생하는 중복 데이터 제거 기법을 제안한다. 제안하는 기법은 Meta-data를 사용한 사전 교섭으로 동일한 정보가 각각 다른 클러스터에 전송되는 것을 방지하여 에너지 절감 효과를 볼 수 있다. 또한, 클러스터 내에서 발생하는 다른 정보들을 시간 지연 기법을 사용하여 하나의 패킷으로 데이터 모음을 수행하는 기법도 제안한다. 성능 평가를 통해 제안하는 알고리즘은 기존의 기법에 비하여 지연 시간과 에너지 소모 면에서 모두 효율적인 것을 확인할 수 있다.

1. 서 론

무선 센서 네트워크는 넓은 지역이나 건물 등에 위치한 센서를 통해 주변 환경의 변화를 수집하고 그 정보를 중앙의 기지 노드(Base Station or Sink)로 보내 정보를 수집한다. 센서 네트워크는 지리적으로 사람이 쉽게 접근하기 어렵고, 크기가 넓기 때문에 관리가 어려운 지역 등의 환경을 모니터링 하거나, 또는 빌딩 안에 출입, 보안, 화재 경보 등 계속적인 관리가 필요한 작업들을 수행하는데 이용되고 있다[1, 2].

센서 네트워크를 구성하는 노드들은 에너지원으로 배터리를 사용하기 때문에 배터리 용량을 모두 소모한 노드는 더 이상 정상적인 라우팅 동작을 수행할 수 없다. 따라서 노드의 가용 시간을 보장하기 위해서는 제한된 배터리 용량으로 발생하는 에너지의 고갈을 해결해야 한다. 일반적으로 센서 노드들은 스스로 충전할 수 있는 능력을 가지고 있지 않을 뿐만 아니라, 넓게 분포되어 있는 수많은 노드들의 배터리를 교체하는 일도 매우 어렵다. 게다가 노드 크기의 제한으로 많은 용량의 배터리를 탑재하는 것도 쉽지 않기 때문에 높은 가용 시간을 보장하기 위해서는 한정된 자원을 효율적으로 사용할 수 있는 여러 기술들이 제안되어야 한다. 센서 네트워크에서도 노드들을 그룹화하여 관리를 용이하게 하고 범위성을 증가시킨 클러스터 기반의 라우팅 알고리즘[3, 4]이 제안되었는데, 클러스터 기반의 라우팅 알고리즘은 클러스터 내에서 발생한 동일한 정보를 클러스터 헤드에서 데이터 모음이 수행되어 하나의 데이터 패킷으로 전송할 수 있기 때문에 에너지를 절약할 수 있다. 여러 개의 동일한 데이터를 전송할 때보다 하나의 패킷으로 데이터를 모아서 전송하는 것이 전송 횟수와 전송 양을 줄일 수 있기 때문에 많은 에너지 절약 효과가 있다[5].

그러나 클러스터를 이용하여 데이터 모음을 하게 되면 클러스터와 클러스터 경계에 위치한 중복된 지역을 각각의 클러스터가 센싱 할 수 있다. 이 때 중복된 지역에서 발생한 이벤트가 두 개 이상의 클러스터에서 개별적으로 센싱되며, 중복된 데이터는 서로 다른 클러스터 헤드에서 데이터 모음을 하고 이 데이터들은 BS으로 각각 전송될 것이다. 따라서 중복된 데이터의 전송은 에너지 절약 측면에서는 매우 비효율적이므로 중복된 데이터를 제거하여 에너지 효율적인 센서 네트워크를 구성하도록 한다.

본 논문은 다음과 같이 구성된다. 2장에서는 본 논문과 관련된 연구들에 대하여 알아보도록 하고, 3장에서는 구체적으로 본 논

문에서 제안하는 클러스터 알고리즘과 중복 데이터를 제거하는 알고리즘에 대하여 알아본다. 이어서 4장에서는 실험을 통하여 본 논문의 우수성을 증명하고 5장에서 결과에 대하여 정리하며 결론을 맺는다.

2. 관련 연구

SPIN (Sensor Protocols for Information via Negotiation)[6]은 실제 데이터를 얻는 작업을 수행하기 전에 Meta-data를 미리 주고받아서 실제 데이터 전송의 수행 여부를 결정한다. SPIN에서는 센서 네트워크에서 일어날 수 있는 세 가지 문제점을 지적하며 그 문제점의 해결을 시도하고 있는데, SPIN에서 지적하는 문제점으로는 메시지의 중복전달로 발생할 수 있는 Implosion 문제, 센서 노드의 센싱 범위 중복으로 같은 지역을 센싱한 데이터가 중복 전달되는 Overlap 문제, 마지막으로 노드들의 에너지 상태를 알기 어려운 Resource Blindness 문제가 있다. 이 문제점들을 SPIN에서는 Meta-data를 이용하여 네트워크에서 발생하는 불필요한 전송을 제거하고 전송횟수를 줄여 에너지 효율을 높일 수 있는 것이다. SPMS (Shortest Path Minded SPIN) [7]은 SPIN-Based 라우팅 알고리즘으로서 SPIN과 같이 meta-data를 이용하는 Negotiation을 수행하여 라우팅을 진행한다. SPIN과 마찬가지로 Meta-data를 교환하여 데이터의 필요 여부를 확인 후 실제 데이터 전송을 수행한다. 그러나 SPMS는 데이터 전송을 위해서 최소 거리의 멀티 홉을 구하여 최단 거리로 데이터 전송을 시도하기 때문에 단순히 point-to-point나 broadcasting을 이용하여 데이터 전송을 하는 SPIN 보다 적은 에너지가 소비되어 좋은 성능을 보이게 된다.

데이터 모음(Data Aggregation) 기법은 실제 전송이 일어나기 전에 중복된 데이터를 압축하고 하나의 패킷으로 만들어 전송횟수를 줄임으로서 에너지 절감효과를 볼 수 있는 기법이다. Data Funneling[8]은 평면기반 네트워크에서 센싱이 필요한 지역을 지정하여 데이터 모음을 수행하는 라우팅 알고리즘이다. 각 센서 노드에서 발생한 데이터를 BS에 개별적으로 보내는 대신에 Border 노드로 지정되는 지정된 지역 안에서 BS와 가장 가까이 있는 노드로 데이터를 하나의 패킷으로 묶정한 후 BS로 보내는 방법이다.

본 논문에서 바탕을 두고 있는 클러스터 구성 알고리즘은

Energy Efficient Hierarchical Clustering Algorithm [4]으로 이 알고리즘은 Single-Level의 클러스터링 알고리즘을 제안하고 이를 확장하여 계층적인 클러스터링 알고리즘을 구성한다. 계층적인 클러스터링 알고리즘은 Single-Level 클러스터링 알고리즘을 확장해 가면서 먼저 가장 낮은 레벨의 클러스터 헤드를 선출하고, 이 헤드가 다시 한 단계 높은 클러스터 헤드를 선출하며, 이러한 반복을 통해서 계층적인 클러스터가 구성된다.

3. 중복 제거 알고리즘

3.1 클러스터 구성

본 논문은 중복 데이터 제거 알고리즘에 초점을 맞추므로 클러스터 구성은 Energy Efficient Hierarchical Clustering Algorithm의 알고리즘을 차용한다. 클러스터 헤드는 일정한 확률 p 에 의해서 선택되는데 이 값은 일정하지 않으며, 최적의 p 값은 네트워크 전체 노드의 수에 따라서 달라진다. 또한, 클러스터의 크기와 네트워크에서 생성되는 클러스터 수에 영향을 미치는 클러스터의 최대 홉 수인 k 의 값도 일정하지 않는데 이 값도 네트워크에 참여하고 있는 노드의 수에 따라서 변하게 된다. 아래의 [표 1]은 네트워크에 참여하고 있는 노드의 수에 따라 결정되는 최적의 p 값과 최대 홉 수 k 값을 나타낸다[4].

표 1 클러스터 구성 변수 값

노드 수	헤드 선출 가능성	최대 홉 수
500	0.1012	5
1000	0.0792	4
1500	0.0688	3
2000	0.0622	3

3.2 중복 데이터 제거

본 논문에서 제안하는 중복 데이터 제거 알고리즘을 전개하기 위해서 SPIN[6]에서 사용되었던 Meta-data 형식을 사용한다. SPIN에서 정의하고 사용하는 Meta-data란 다음과 같다.

- Meta-data : Meta-data란 노드가 센싱한 데이터의 간단한 요약(description)이다. 어떤 센싱 노드에서 a 를 실제 데이터인 A의 Meta-data라고 한다면 전송 이익을 얻기 위하여 a 의 크기는 A의 크기보다 작아야 한다. 또한, 두 개의 다른 실제 데이터가 있다면, 각 데이터의 Meta-data들 역시 구분할 수 있도록 달라야 한다.

클러스터 구조를 이루고 있는 센서 네트워크에서 Meta-data를 사용하여 중복 데이터를 제거하는 방법은 다음과 같이 진행된다. 클러스터에서 Leaf 노드가 이벤트를 센싱하게 되면 센싱한 데이터를 바로 클러스터 헤드로 보내지 않고 그 데이터의 Meta-data를 최대 전송 전력을 사용하여 브로드캐스팅 한다. 최대 전송 전력을 사용하는 것은 두 개 이상의 노드가 중복된 지역을 센싱할 때 그 중복 데이터를 전송하지 못할 수도 있기 때문이다. Leaf 노드에서 Meta-data를 개별적으로 브로드캐스팅 하여 다른 클러스터의 Leaf 노드도 Meta-data를 통하여 같은 데이터가 센싱된 것을 알게 되면 두 Leaf 노드 중 BS와의 거리가 가깝고 클러스터 헤드와의 거리가 가까운 Leaf 노드가 자신의 클러스터 헤드로 데이터를 전송하게 된다. 이것은 데이터의 전송 에너지는 전송거리에 비례하기 때문에 더욱 적은 전송 횟수를 보장하기 위해서이다.

본 논문에서 제안하는 중복 제거 기법은 SPIN[6]과 SPMS[7]에서 언급했던 세 가지 문제 중에서 Overlap 문제와 유사한 점이 존재한다. 실제 SPMS에서는 클러스터 구조 상태에서 효율성을 실험하였다고 언급하였다. 그러나 본 논문에서 제안하는 문제는 두 개 이상의 클러스터에서 중복된 지역이 발생했을 경우인데 이 때에는 같은 지역을 센싱한 노드의 Upstream 노드가 다르기 때문에 Overlap 문제와 차이점을 보인다. Overlap 문제는 이런 경우 문제를 해결하기 어렵지만 본 논문에서 제안하는 알고리즘은 이

용하면 간단히 해결할 수 있다.

3.3 데이터 모음

데이터를 전송하기 전에 Meta-Data를 이용하여 Negotiation을 하면 Overlap이나 Implosion이 일어난 데이터를 제거할 수 있다. 그러나 비슷한 지역에서 발생한 다른 이벤트는 중복된 데이터가 아니기 때문에 Meta-Data를 이용하여 제거가 불가능 하다. 이런 경우에는 데이터 모음을 수행하여 두 개 이상의 다른 패킷을 하나의 패킷으로 통합하여 데이터 전송 횟수를 줄일 수 있다. 하나의 패킷으로 통합하여 BS로 전송을 시도한다면 데이터를 전송하는데 사용되는 에너지를 상당부분 절약할 수 있다. 따라서 [8]에서 제안했던 각 노드에서 일정시간만큼 지연을 일으켜 데이터를 모은 후 전송하는 방법을 사용하면 추가적으로 에너지를 절약할 수 있다. 아래의 식은 각 노드에서 발생하는 지연시간이다.

$$\frac{1}{W_k} = W_{k-1} + H_p$$

여기서 k 는 홉 수를 나타내며, WT_k 는 k 홉에서의 Waiting time 이 된다. 즉, k 홉에서의 Waiting time WT_k 은 k 홉 수의 역수 값으로, 이 값에 비례상수 G_w 를 곱한 값만큼 각 노드에서 지연을 발생시키고 데이터 전송이 일어난다. 예를 들어서 G_w 가 0.1이고, 3홉으로 구성된 클러스터의 Leaf 노드에서 Event가 발생하여 데이터를 클러스터 헤드로 모으는 상황이 발생한다면 각 센서 노드에서 기다려야 하는 Waiting time은 3홉 노드에서는 33ms, 2홉 노드에서는 50ms, 1홉 노드에서는 100ms 만큼의 지연이 발생한 후 데이터 전송을 시도한다. 따라서 각 노드에서는 순차적으로 데이터를 하나로 모을 수 있게 되고, 결국 클러스터 헤드로 데이터가 전송되어 하나의 데이터 패킷으로 모아지게 된다.

3.4 이론적 평가

이번 장에서는 중복 데이터 제거 기법을 사용하였을 때 소모되는 에너지 값과 시간 지연에 대하여 이론적으로 평가한다. 사용되는 변수는 다음과 같다. D_R 과 D_M 은 실제 데이터와 Meta-data의 길이가 되며, A 는 Advertisement 메시지, R 은 Request 메시지의 길이가 된다. n_A 는 중복된 클러스터 수이고, E 은 bit당 소모되는 에너지 값으로 E_r 은 패킷을 받는데 소모되는 에너지이고, E_m 은 주변 노드에게 전송할 수 있는 최소 전력에너지가 된다.

C_{normal} 은 [4]에서 구성하는 일반적인 클러스터 모델의 전송 비용이고, C_{EDD} 는 일반적인 클러스터에 중복 제거 기법을 적용하였을 때의 전송 비용이다. C_{EDD1} 과 $C_{normal1}$ 은 Leaf 노드에서 중복된 지역이 아닌 지역에 Event가 발생할 때의 비용이고 C_{EDD2} 와 $C_{normal2}$ 는 중복된 지역 안에 Event가 발생할 때의 비용이다.

$$C_{EDD1} = k(D_R E_m + D_R E_r) + D_M E_m$$

$$C_{EDD2} = k(D_R E_m + D_R E_r) + n_A(D_M E_m + D_M E_r)$$

$$C_{normal1} = k(D_R E_m + D_R E_r)$$

$$C_{normal2} = n_A(D_R E_m + D_R E_r)$$

Leaf 노드에서 Event가 발생할 때 Event가 여러 클러스터로 이루어진 중복된 지역에서 발생할 수도 있고 아닐 수도 있기 때문에 두 경우를 모두 계산해야 한다. C_{EDD1} 은 중복된 지역 밖에 Event가 발생할 경우로서, 노드에서 Meta-Data의 오버헤드와 k 개의 멀티 홉을 따라 정보들이 클러스터 헤드로 전송되는 비용이 된다. C_{EDD2} 는 n_A 개가 중복된 지역 안에서 Event가 발생한 경우로서, n_A 개의 Meta-data 송수신과 한 번의 데이터 전송 비용이 든다. $C_{normal1}$ 은 k 개의 멀티 홉을 따라 정보들이 클러스터 헤드로 전송되는 비용과 같고, $C_{normal2}$ 는 n_A 개의 중복된 데이터가 n_A 개

의 서로 다른 클러스터 헤드로 전송되는 비용이 된다.

다음으로 시간 지연에 대하여 알아본다. 본 논문에서는 Waiting time을 이용한 데이터 모음을 시도하기 때문에 어느 정도의 지연이 발생하게 된다. 따라서 본 논문에 기반을 두고 있는 SPIN과 SPMS와의 지연을 비교하여 어느 정도 차이가 있는지 살펴해보도록 한다. T_{tz} 는 하나의 데이터를 전송하는데 소모되는 시간이고 $T_{Out_{ADV}}$ 와 $T_{Out_{DAT}}$ 는 타이머 값이 된다. T_{proc} 는 노드에서 데이터를 처리하는데 소모되는 시간이며, $T_{channel}$ 는 채널 접근 지연으로서 n^2 에 비례해서 증가한다. 여기서 n 은 전송 범위 안에 있는 노드 수이다. G 는 비례상수이고 n_M 과 n_m 는 도달할 수 있는 노드 수로서 각각 최대 전송 범위와 최소 전송 범위 일 때의 노드 수이다. 지연 비교는 Leaf 노드에서 클러스터 헤드로 도달하는데 걸리는 시간을 비교하는데, Leaf 노드에서 클러스터 헤드는 k 층만큼 떨어져 있다. 본 논문에서 제안하는 알고리즘은 먼저 Leaf 노드에서 중복 데이터를 검색한 후에 Waiting time 만큼의 지연 발생 후 Upstream 노드로 데이터를 전송하게 된다.

우선 노드에서 노드로 전송하는데 걸리는 지연 시간은 다음과 같다.

$$T = Gn_M^2 + DT_{tz} + T_{proc}$$

Leaf 노드에서는 Meta-Data를 주고받아야 하기 때문에 ADV와 REP 메시지를 서로 전송하게 된다. 따라서 Leaf 노드에서 발생하는 시간 지연은 다음을 만족한다.

$$T_{hopMAX} = Gn_M^2 + 2Gn_m^2 + (A + R + D)T_{tz} + 2T_{proc}$$

Leaf 노드를 제외한 클러스터 헤드까지 $k-1$ 개의 노드들에서 발생하는 시간 지연은 다음과 같다.

$$T_k = (Gn_m^2 + DT_{tz} + T_{proc})(k-1)$$

따라서 k 층만큼 떨어져 있는 클러스터 헤드로 전송되는데 경과되는 총 지연 시간은 다음과 같다.

$$T_{EDD} = Gn_M^2 + (k+1)Gn_m^2 + (kD + A + R)T_{tz} + (k+1)T_{proc}$$

3.3에서 제안한 Waiting time을 적용하면 추가적으로 시간 지연이 발생하며 이 값은 T_{EDD} 보다 크다.

중복 제거 기법을 적용하지 않은 일반 클러스터에서 경과되는 시간 지연은 ADV나 REQ 메시지 없이 k 층 만큼 떨어져 있는 클러스터 헤드로 데이터 전송을 시도하는데 경과되는 시간이다. 식은 다음과 같다.

$$T_{normal} = k(Gn_m^2 + DT_{tz} + T_{proc})$$

SPMS는 멀티 홉 전송을 시도하기 때문에 자연스럽게 SPIN에 비해서 시간 지연이 발생하게 되는데 그 지연 값은 다음과 같다 [7].

$$T_{SPMS} = (k-1)T_{round} + T_{Out_{ADV}} + T_{hopMAX}$$

여기서 T_{round} 는 한 번의 전송 과정이고, $T_{Out_{ADV}}$ 는 노드가 ADV를 보내기 전까지의 시간 지연이며, T_{hopMAX} 는 Leaf 노드에서 소모되는 시간 지연이다.

4. 성능 평가

성능 평가는 전체 센서 네트워크에서 소모되는 에너지 측면을 비교해보도록 한다. 그리고 기존 알고리즘과 비교하여 어느 정도의 시간 지연(Delay) 효과가 있는지를 기준으로 수행하였다. 실험에 사용된 값은 $T_{tz} = 0.05$, $T_{proc} = 0.02$, $A : R : D = 1 : 1 : 30$, $G = 0.01$, 그리고 $n_M = 20$, $n_m = 5$ 의 값을 갖는다. 에너지 측면에서 비교해 봤을 때에는 중복 데이터 제거 기법을 사용했을 경우와 적용하지 않은 경우에 대하여 에너지 소비에 대하여 실험하였다. [그림 1]에 나타나듯이 중복 데이터를 제거하여 데이터 전송 횟수를 줄이므로 더 적은 에너지가 소모됐음을 실험을 통해서 알 수

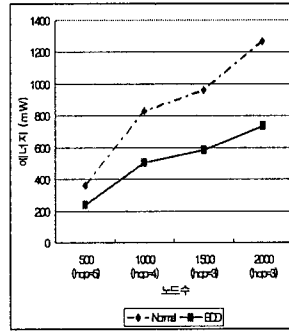


그림 1 에너지 소모 비교

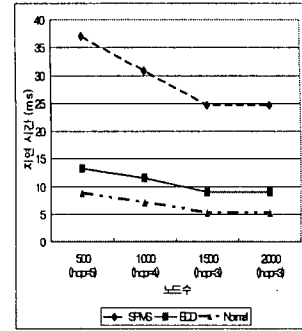


그림 2 지연시간 비교

있었다.

시간 지연에 대한 실험에서는 중복 제거 기법과 일반 클러스터 이외에 클러스터 구조로 된 SPMS에 대한 실험도 추가하였다. [그림 2]에서와 같이 본 논문에서 제안하는 알고리즘은 ADV와 REQ의 과정이 없기 때문에 SPMS와 비교하여 많은 지연 시간을 절약할 수 있었다. 그러나 중복 데이터 제거 기법에 대한 시간 지연이 발생하기 때문에 중복 데이터 제거 기법을 적용하지 않는 클러스터에 비해서 약간의 시간 지연이 발생하였다.

5. 결론

본 논문에서는 클러스터 구조를 이루고 있는 센서 네트워크에서 두 개 이상의 클러스터의 중복 센싱이 발생 가능한 지역에서 수행하는 중복 데이터 제거 기법과 클러스터 내에서 수행할 수 있는 데이터 모음 기법에 대하여 제안했다. 실험 결과 기존 알고리즘에 비하여 크지 않은 시간 지연으로 보다 적은 에너지 소모를 하는 효율적인 클러스터 구성을 이루어 정보 수집이 가능한 것을 확인할 수 있었다.

참고 문헌

- [1] Deborah Estlin, Ramesh Govindan, John Heidmann, and Satish Kumar, "Next Century Challenges : Scalable Coordination in Sensor Networks", In Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking, August 1999.
- [2] Gregory J. Pottie and William J. Kaiser, "Embedding the Internet Wireless Integrated Network Sensors", Communications of the ACM, V.43 (N.5), pages 51-58, May 2000.
- [3] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," in Proceeding of the 33rd Annual Hawaii International Conference on System Sciences, pages 3005-3014, 2000.
- [4] S. Bandyopadhyay, E.J.Coyle, "An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks", IEEE INFOCOM 2003.
- [5] K. Dasgupta, K. Kalpakis, P. Namjoshi, "An Efficient Clustering-based Heuristic for Data Gathering and Aggregation in Sensor Networks"
- [6] W. R. Heinzelman, J. Kulik, and H. Balakrishnan, "Negotiation-based Protocols for Disseminating Information in Wireless Sensor Networks," in Proc. of the 5th Annual ACM/IEEE International Conference on Communications, 2001.
- [7] Khanna, G., Bagchi, S., Yu-Sung Wu, "Fault tolerant energy aware data Dissemination protocol in sensor networks", International Conference on 2004.
- [8] Petrovic, D. Shah, R.C. Ramchandran, K. Rabaey, J. "Data funneling: Routing with Aggregation and Compression for Wireless Sensor Networks", Proceedings of the First IEEE 2003.