

유사구간 클러스터링을 이용한 에너지 효율적인 데이터 전송방법*

홍창기^o 김창화 이승재 김상경
강릉대학교 컴퓨터 공학과

{doctorband^o, kch, silveree, skim98}@kangnung.ac.kr

An Energy-Effective Data Transmission Method using Clustering Similar Ranges

Changgi Hong^o, Changhwa Kim, Seungjae Lee, Sangkyeung Kim

Department of Computer Engineering, Kangnung National University

요 약

센서 네트워크는 제한된 에너지를 가지는 작은 노드들로 구성이 된다. 이 센서 네트워크에서 가장 큰 에너지 손실을 가져오는 부분은 RF통신 부분이라 할 수 있다. 해양 센서 네트워크는 통신 매체로 음파를 사용하기 때문에 RF를 사용하는 센서 네트워크보다 통신하는데 더 많은 에너지를 소모한다. 센서 네트워크에서 통신 횟수를 줄여 에너지 효율을 높이는 방법으로 네트워크 내 집계 연산이나 필터링 등이다. 해양환경에서 데이터 값들이 유사한 층을 가지고 있다. 이 유사층에서 네트워크 내 집계 연산과 필터링의 의미를 살펴보겠다. 해양 센서 네트워크는 기존의 토폴로지와 다른 구조를 가지고 있다. 새로 제안하는 구조에 어떠한 개념과 기능이 있는지를 살펴본 후 센서 노드들 임계값을 사용하여 센싱된 데이터 값이 유사한 구간을 클러스터로 묶고 묶여진 클러스터 내에서 어떻게 데이터를 전송할 방법을 제안한다.

1. 서 론

센서 네트워크(sensor network)는 제한된 에너지를 가지는 작은 무선 센서 노드들로 구성 된다. 구성된 센서 네트워크가 동작하기 위해서는 작은 무선 센서 노드들이 서로 연결하여 특정 토폴로지를 형성하여야 한다. 노드들 간에 가까운 거리는 직접 연결하고, 통신 거리가 먼 경우 멀티홉을 이루어 통신을 한다. 연결된 노드들은 사용자가 질의를 요청하면, 특정 센서 혹은 모든 센서에게 요청을 전달한다. 요청을 받은 센서 노드들은 요청에 의해 수집한 데이터를 직접 또는 멀티홉을 통해서 질의를 요청한 노드에게 전달하게 된다.

RF센서네트워크 통신과 해양센서네트워크 통신의 가장 큰 차이는 통신 매체이다. 일반적으로 RF를 사용하는 센서 네트워크는 수중에서 전파가 물에 흡수가 되기 때문에 사용할 수 없다. 음파는 수중에서도 흡수가 되지 않고 데이터를 전달 할 수 있기 때문에 해양센서네트워크에서 사용 가능하다. 음파를 이용하여 데이터 전송을 하게 되면 RF통신보다 많은 에너지가 소모된다.

센서 네트워크는 수심에서 수백 개의 노드가 함께 네트워크를 이루어 작동한다. 에너지가 소진되면 모든 노드들의 전력을 교체해주는 것이 현실적으로 불가능하며 가능하더라도 비효율적이다. 어떤 센서 네트워크의 유효 기간이 끝

났는지 여부를 결정하는 것은 여러 가지 기준으로 정할 수 있다.[1] 예를 들어, 전체 센서 네트워크 중에서 주요 센서 하나라도 동작하지 않거나, 일정 부분 센서가 동작하지 않을 때 생명(lifetime)이 끝난 것으로 본다. 이러한 이유로 센서 네트워크에서 핵심 이슈가 되는 것이 저전력, 에너지 효율이다. 센서 네트워크에서 중요시하는 것이 노드들의 전송 횟수를 최소화 하여 센서 노드의 수명을 증가시키면서 더불어 센서 네트워크 전체의 수명을 증가시키는 연구이다. 가장 대표적인 것이 집계 내 연산(in-network aggregation)과 필터링이다. 집계 내 연산은 노드들을 경유 해가는 메시지를 집계 가능한 정보(SUM, AVG, COUNT, MAX, MIN)을 효과적으로 처리 하는 것이다. 필터링은 센서 노드들이 센싱하는 값이 허용되는 오차가 문제가 되지 않는 경우 메시지 전송을 줄이는 방법이다. 초음파를 사용하는 해양센서 네트워크에서는 송신과 수신에 필요한 에너지가 크기 때문에 데이터 전송 횟수를 줄여 에너지 소모를 최대한 줄이는 더욱 효과적이라 할 수 있다.

본 논문은 해양 환경이 데이터 값이 유사 층을 이루는 특성에서 시작된다. 해양은 수심에 따라 수온 및 염도, 용존 산소량 등의 정보들이 층을 이루고 있다. 유사층에 노드들이 센싱하는 정보들은 유사한 값을 가진다. 이 범위는 작게는 수 미터에서 수십 km까지 존재한다.

이 영역에서의 집계 내 연산이나 필터링의 효율을 살펴본 후, 해양 환경에 맞는 클러스터링을 통한 데이터의 효율적인 전송 방법을 제안한다. 제안한 방법은 무선 네트워크 구성이 완료가 되면 싱크 노드는 하위 노드의 센싱 값을 수집하고, 그 센싱 값에 따라 그룹을 만든다. 만들어진 그룹에서 영역을 구분하여 클러스터를 만들고 영역 안의 클러

* 본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (IITA-2005-C1090-0501-0010)

스터 헤더를 선정한다. 선정된 클러스터 헤더는 그룹에서 전송 노드를 선정한다. 선정된 노드는 직접 또는 멀티 홉을 통해서 싱크 노드로 데이터를 전송한다.

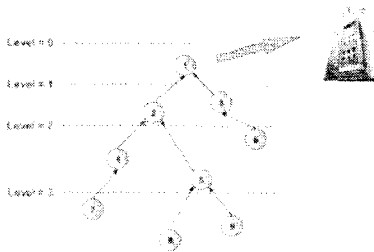
논문의 구성은 다음과 같다. 2장에서 센서 네트워크에서 쓰이는 데이터 횡수를 줄이는 전송 방법을 알아본다. 3장 해양환경을 설명하고 해양센서네트워크 클러스터링방법을 소개한다. 4장에서 본 논문의 결론을 내리고 향후 과제를 제시하겠다.

2. 관련연구

센서 네트워크에서 데이터 전송을 줄여 에너지를 효율적으로 사용하는 방법 있는데, 해양센서 네트워크 클러스터링기법을 살펴보기 전에 알아보도록 하겠다.

첫 째, 네트워크 내 집계 연산이다. 센서 데이터를 집계 하기 위한 방법으로 쉽게 생각할 수 있는 것은, 일단 센서가 읽은 모든 집계 처리를 수행하는 것이다.[1] 중앙 처리 방식(centralized, serverbased approach)은 메시지를 매우 많은 노드들을 경유하여 전송한다. 이때 전송시간과 처리시간이 오래 걸리게 된다. 이러한 단점을 극복하기 위해 네트워크 내 집계를 사용하면 중앙 처리 방식보다 메시지 전송시간을 줄일 수 있으며, 각 노드마다 전송하는 데이터 횡수를 줄여 에너지를 절약할 수 있다. 우선 네트워크 내 집계 연산을 수행하기 위해서는 토폴로지를 형성하여야 한다.

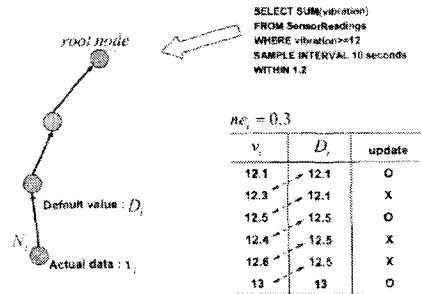
[그림 1]은 센서 네트워크에서 네트워크 내 집계 연산을 수행하기 위해 노드들이 기지국(Base Station)까지의 경로를 트리 구조로 표현 한 것이다. 기지국에서 질의를 요청하게 되면 센서 노드에게 전달이 되고 센서 노드들은 부모노드에게 데이터를 전송한다. 이때 부모노드에게 전송을 할 때는 모든 자식 노드가 데이터를 정해진 시간까지 기다렸다가 보내게 된다. 부모노드는 정해진 시간 내에 데이터를 받으면, 부모노드는 부분 집계 결과(partial aggregate)를 계산해서, 그 부분 집계 결과를 부모 노드에게 전송을 한다. 이 과정은 [그림 1]에서 level 3 이하부터 level0을 거치게 된다. level0을 거치고 나면, 루트 노드는 집계 연산된 결과를 기지국에 전달하게 된다. 이때 level3의 데이터는 기지국까지 전달되는 것이 아니라 상위 레벨 즉 level2까지 전송되어 메시지 전송을 효율적으로 할 수 있게 된다.



[그림 1] 네트워크 내 집계 연산[1]

둘 째, 데이터 필터링이다. 센서 데이터에 기반을 두는 많은 응용들은, 근사 결과를 어느 정도의 부정확도 안에서 허용할 수 있다. 다시 말해, 센서 데이터는 그 값에 어느 정도의 오차가 있어도 큰 문제가 되지 않는 경우가 많다는 것이다.[1] 이것은 데이터의 정확도가 크게 중요하지 않은 응용에서 사용될 수 있다. 센서가 측정하는 온도, 습도, 염도, 용존 산소량 등의 데이터를 정확하게 측정하는 것은 쉽지 않다. 또 데이터 값이 100%정확할 필요가 없을 수도 있다. 실제로 해양에서 수온이 9.0℃와 9.1℃도를 크게 차이가 없는데, 실제 해양 수온에서 0.1℃도 차이는 크게 문제가 되지 않는다. 이와 같은 특징을 이용하여 센서 노드에서 센싱한 값이 이전의 데이터와 오차 허용 범위 내 문제가 없으면 메시지 전송을 줄이는 방법을 데이터 필터링이라고 한다.

[그림 2]와 같은 구조로 노드들이 구성이 되면 Actual data는 실제 데이터를 센싱하는 부분이고 Default Value 는 Actual data를 저장하는 곳이다. 이때 허용되는 오차를 0.5라고 가정을 한다. Actual data가 이전의 데이터와 비교하여 오차 허용 범위 내에 있게 되면 데이터 값을 전송하지 않고 오차 허용 범위 밖의 데이터이면 부모 노드로 전송을 하게 된다. 이 과정은 [그림 2]에서 Actual Data가 12.1일 때 첫 번째 데이터로 Default Value에 저장되지만 12.3은 오차 허용 범위 내에 있기 때문에 갱신되지 않고 12.5일 때 오차 허용 범위 밖에 있기 때문에 갱신이 일어난다. 이처럼 필터링은 일정 시간 동안 자식노드가 데이터를 보내지 않으면 캐싱하고 있던 데이터를 사용하여 오차 범위 내의 데이터 전송을 줄여 효율적으로 할 수 있게 된다.

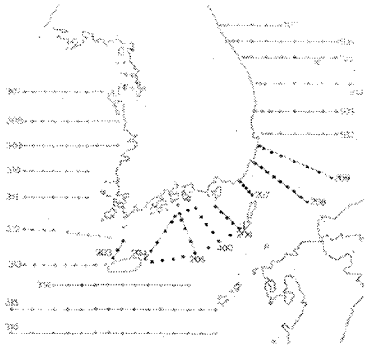


[그림 2] 필터링 [1]

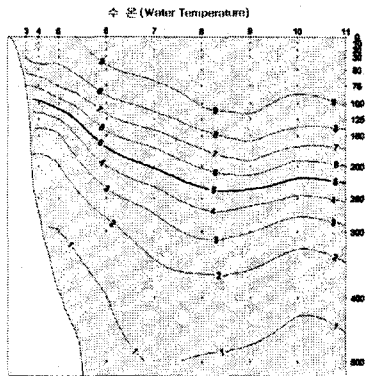
3. 해양센서 네트워크 클러스터링 기법

3.1 해양환경의 특성

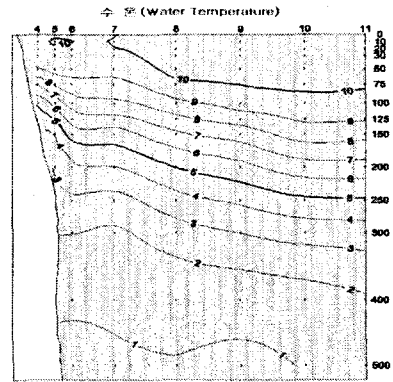
해양환경에서 온도에 따라 표수층, 변수층, 심수층으로 구분을 한다. 이는 해양환경에서 온도 데이터가 수심에 따라 층을 이루고 있는 것을 보여준다. [그림 3]은 정선 관측지점 위치도로 우리나라의 해양 정보를 수집할 장소를 표시한 것이다.



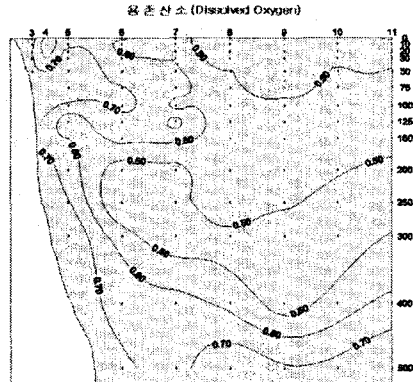
[그림 3] 정선관측지점 위치도[3]



[그림 4] 105 정지선 수온[3]



[그림 5] 104 정지선 수온[3]



[그림 6] 105 정지선 용존산소량[3]

[그림 4]는 [그림 3]의 105 관측지점 단면도이다. [그림 3]의 가로축은 [그림 3]에서 보는 105 관측지점의 정점이 다. 첫 번째 정점은 3이고 마지막 정점은 11이다. 세로축은 수심을 나타낸다. [그림 4]를 볼 때 수심에 따라 온도층이 넓은 구간에서 형성된다.

[그림 4]와 [그림 5]는 [그림 3]의 105 관측지점과 104 관측지점의 단면도이다 105와 104의 실제 거리는 수심 km 이지만 [그림 4]와 [그림 5]는 보이는 것처럼 온도의 변화가 비슷하고 유사한 층을 이루고 있다.

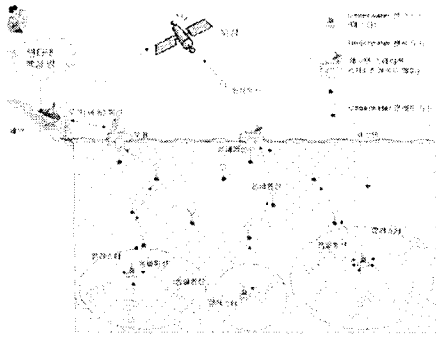
또 [그림 5]와 [그림 6]은 [그림 3]의 105 관측지점 단면도의 수온과 용존산소량을 나타낸다. 이때 용존산소량과 온도가 유사한 층이 있다. 즉 가로 구간 8에서 11까지 세로구간 0m에서 30m까지 용존산소량과 온도는 유사한 층을 이루고 있다. [그림 5]와 [그림 6]에서 8지점에서 11지점까지의 수심 50m 지점에서 센싱 한다고 가정하자. 이 지점에서 온도차이가 1°C이하, 용존산소량이 0.1이하를 같은 값이라 생각하는 것이 무방하다면 이 범위 내의 모든 노드들은 같은 값을 가진다고 볼 수 있다. 이처럼 해양에서는 유사한 측정값을 가지는 구간이 넓게 존재한다.

해양환경의 유사층에서 얻어지는 유사한 데이터의 전송 횟수를 줄이게 되면 배터리 소모를 줄일 수 있다. 데이터 전송 횟수를 줄이기 위한 방법으로는 앞에서 살펴본 바와 같이 집계 내 연산, 필터링 등이 있다. 유사층에 10개의 노드가 있다고 가정하자. 이때 싱크 노드가 질의를 하면 센서

노드들은직접 또는 멀티 홉을 통해서 싱크 노드로 데이터를 전송한다. 네트워크 내 연산을 사용할 경우 10개의 노드에서 전송되는 데이터 정보는 유사 하며 데이터 전송 횟수는 최소 10일 것이다. 다음으로 데이터 필터링을 생각해보자 이 유사층에서 데이터의 변화가 없다면 데이터 재전송하지 않고 이전에 캐싱해 놓은 데이터를 사용하여 질의응답을 할 수 있을 것이다. 하지만 유사 층의 정보가 변화가 되면 필터링을 통한 데이터 전송의 횟수는 10이 될 것이다. 본 논문에서는 이와 같은 유사 층 혹은 구간이 존재하는 해양 환경에 맞는 클러스터링 방법을 제안하여 데이터 전송 횟수를 최소로 줄이는 방법을 제안하고자 한다. 즉 유사층의 노드를 각각의 클러스터로 묶어 이들을 대표하는 노드를 정하고 데이터를 전송하는 것이다.

3.2 해양 센서 네트워크 구성

해양 센서 네트워크는 초음파를 이용하여 물속에서 구성되는 무선 센서 네트워크로 일반 센서 네트워크와 마찬가지로 저전력과 고신뢰도, 설치의 용이성 같은 특징을 가져야 하며, 물속에서 초음파를 이용하여 발생하는 낮은 전송율과 높은 에러율을 극복할 수 있도록 설계되어야 한다.



[그림 7] 해양 센서 네트워크 구조 [2]

해양 센서 네트워크는 [그림 7]과 같이 기지국, 인공위성, 게이트웨이, 릴레이 노드, 싱크노드, 센서 노드로 구성된다. 각 구성요소를 알아보겠다.

○ 기지국

해양 센서 네트워크의 게이트웨이가 기지국과 충분히 가까운 거리에 있어 전파를 이용하여 직접 통신이 가능한 경우는 기지국과 게이트웨이가 직접 통신을 하게 되며, 직접 통신을 수행하지 못할 정도로 멀리 떨어져 있는 경우 인공위성을 통하여 통신을 한다.

○ 인공위성

기지국과 게이트웨이가 모두 해안 가까이 위치할 경우는 전파를 이용하여 직접 통신을 할 수 있다. 그러나 기지국이 내륙 깊숙한 곳에 위치하거나 게이트웨이가 먼 바다에 떠 있는 경우는 직접적으로 전파를 이용하여 통신을 하기 어렵다. 이러한 경우 인공위성을 이용하여 위성통신으로 기지국과 게이트웨이 간에 데이터를 전달한다.

○ 게이트웨이

게이트웨이는 수중에서 수집한 데이터의 값을 기지국 또는 인공위성으로 보내거나 인공위성에서 수중 센서 노드로 명령 및 데이터를 보내기 위한 장치로, 서로 다른 프로토콜을 연결 시켜주는 역할을 하는 장치이다.

○ 릴레이노드

릴레이노드는 게이트웨이와 싱크노드 간의 거리가 너무 멀어 데이터를 한 번에 전송할 수 없을 경우 중간에서 받은 신호를 재전송하고자 할 때 사용하는 장치

○ 싱크노드

싱크노드는 하나의 클러스터를 관리하는 클러스터헤드 역할을 하는 노드로 센서노드로부터 온 데이터를 취합하여 릴레이노드로 전달하거나 릴레이노드로부터 온 명령 등을 센서노드로 전달하며 스스로 센서노드의 기능도 수행한다.

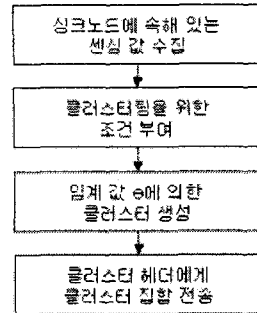
○ 센서노드

센서노드는 실질적으로 필요한 데이터를 센싱하는 역할을 수행하는 노드이다. 모든 센서노드는 라우팅 기능을 수행할 수 있으며, 필요에 따라 싱크노드로 전환도 가능하다.

3.3. 센서 값 기반 클러스터링

기존의 데이터 전송 횟수를 줄이기 위한 방법은 우선적으로 모든 센서들이 센싱을 한 후 그 값에 따라 데이터를 전송할 것인지 아닌지를 판별한다. 또 전송되는 데이터의 횟수를 줄이기 위해서 싱크노드를 경유하는 동안 데이터의 양을 줄이는 방법을 사용하였다. 전자는 필터링이고 후자는 네트워크 내 집계 연산이다.

본 논문은 비슷한 환경에 위치한 노드들을 클러스터링하여 데이터 전송을 최소화 하는 방법을 제안 하고자 한다. 비슷한 환경에서는 모든 노드가 센싱할 필요가 없으며, 이 환경에 대한 대표 값을 전송하게 된다면 센서노드의 기능을 충분히 수행한다고 볼 수 있다. 이때 센싱되는 값을 비교하기 위해 임계값을 두어 센서노드간의 데이터 값이 유사한지를 판별한다. 임계값에 의해 센서노드가 서로 유사한 값을 가진다면 같은 클러스터로 묶는다. 진행 순서는 [그림 8]과 같이 전개 될 것이다.

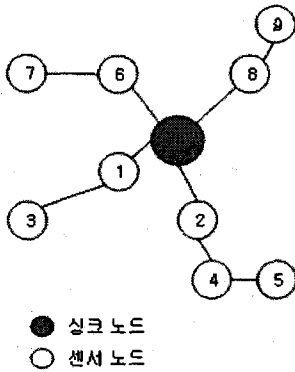


[그림 8] 센서 값 기반 클러스터링 순서

본 논문에서 고려하는 네트워크는 한 개의 싱크 노드를 가지며, 각각의 센서 노드는 싱크 노드와 직접 또는 멀티홉을 통해서 데이터를 전송하며 네트워크 토폴로지는 재구성 없는 환경으로 가정한다.

네트워크 시작시 [그림 9]와 같은 토폴로지가 완성하게 되면 싱크노드는 센서 값 기반 클러스터링을 하기위해 관리하고 있는 모든 센서노드들에게 센싱 질의를 전달한다. 질의를 받는 센서노드들은 센싱한 데이터를 싱크노드들에게 전달한다. 전달된 데이터들은 싱크 노드에 저장 이 된다. 하나의 노드가 총 n 개의 센서를 가지고 있을 때 i 번째 센서가 센싱한 값을 v_i 이라 한다면 노드 S 가 센싱한 값은 다음과 같은 센싱 벡터 V 로 표현할 수 있다.

$$V = (v_1, v_2, v_3, \dots, v_n)$$



[그림 9] 해양센서 토폴로지

이 때 총 노드의 개수를 m 이라 하면 노드 S_j ($1 \leq j \leq m$)가 센싱한 값의 벡터를 v_j 이라 하고 각각의 v_j 는 다음과 같이 정의한다.

노드	센싱벡터	벡터값
S_1	V_1	$(v_{11}, v_{12}, v_{13}, \dots, v_{1n})$
S_2	V_2	$(v_{21}, v_{22}, v_{23}, \dots, v_{2n})$
\vdots	\vdots	\vdots
S_m	V_m	$(v_{m1}, v_{m2}, v_{m3}, \dots, v_{mn})$

[표 1] 센서 노드별 데이터 벡터

센서값을 비교하기 위한 임계값은 C 로 표현하고 다음과 같이 벡터로 표현할 수 있다. C 는 노드 S_i 와 노드 S_j 유사 관계를 비교할 때 쓰인다.

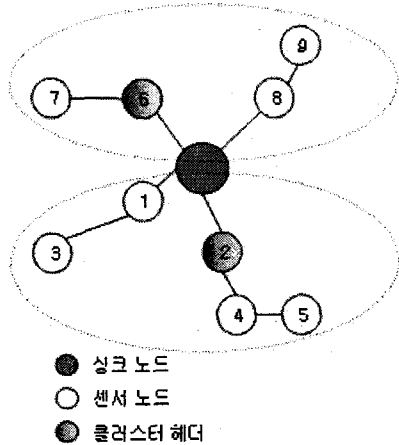
$$C = (c_1, c_2, c_3, \dots, c_n)$$

싱크노드가 유사한 구간 내의 센싱값을 가지는 노드들끼리 클러스터링을 하기 위하여 먼저 임의의 노드 S_i 를 선택한다. 이후 S_i 를 제외한 모든 노드 S_j ($1 \leq j \leq m, i \neq j$)를 대상으로 $1 \leq j \leq m$ 인 모든 k 에 대하여 $|v_{ik} - v_{jk}| \leq c_k$ 을 만족하는 S_j 를 찾아 S_i 와 같은 클러스터로 분류한다. 이와 같은 과정을 통하여 하나의 클러스터가 완성된다.

하나의 클러스터가 완성되면 클러스터에 포함되어 있지 않은 임의의 노드를 선택하여 동일한 과정을 반복하여, 더 이상 클러스터에 포함되어 있지 않은 노드 없으면 클러스터링을 완료 한다. [그림 10]과 같이 클러스터링이 형성되었다면 싱크노드는 각 클러스터 마다 클러스터 헤더를 선정하고 클러스터 헤더에게 같은 클러스터에 포함된 센서 노드들의 정보를 제공한다.

3.4 클러스터링 전송 프로토콜

클러스터 헤더의 역할은 싱크노드가 질의가 있을 때 클러



[그림 10] 클러스터링 형성

스터 내의 어느 노드가 전송할 것인지 정해준다. 이때 싱크노드의 질의가 있을 때 마다 전송 노드를 선정하게 되면 오버헤드가 발생하므로 한번 지정된 노드는 에너지 레벨이 E_i 이하가 될 때까지 전송을 담당하게 된다. 에너지 레벨이 E_i 이하가 되면 노드는 클러스터 헤더에게 전송 노드를 교체 요청을 한다. 클러스터 헤더는 클러스터 내 에너지 레벨에 따라 다른 노드를 지정하여 싱크노드로 데이터를 전송하게 하여 균형적으로 에너지소비를 한다.

4. 결론 및 향후 과제

본 논문에서는 해양 환경에서 유사한 정보 구간을 클러스터링하여 데이터 전송 횟수를 줄이는 방법을 제안 했다. 센서 노드가 싱크 노드로 정보를 보내게 되면 싱크노드는 센서 노드를 유사 정보구간으로 나눈다. 나눠진 구간들 각각의 클러스터로 묶고 클러스터 헤더를 정한다. 클러스터 헤더에게 같은 클러스터에 포함되어 있는 노드들의 정보를 전송한다. 싱크노드가 클러스터 헤더에게 질의 요청을 하게 되면 클러스터 헤더는 자신이 속한 클러스터 중 한 개의 노드를 선정한다. 선정된 노드는 싱크노드가 질의 요청을 하게 될 때 클러스터를 대표하여 정보를 전송하게 된다.

이번 연구에서는 토폴로지 변화의 가능성, 클러스터링 방법을 선택적으로 가정하였으며 정량적인 방법을 통한 효율성 입증은 하지 못하였다. 따라서, 이러한 주제에 대한 연구를 다음 연구로 남겨며 본 논문을 마치도록 한다.

[참고 문헌]

- [1] 박노준, 현동준, 김영호, "센서 네트워크에서 집계 연산을 위한 적응적 필터링", 정보과학학회논문지: 데이터베이스 제 32 권 제 4호, pp.372-382, 2005.8
- [2] 조용만, 김상경, 김창화, "해양 음파 센서 네트워크 구조 연구", 한국컴퓨터종합학술대회 논문집 Vol. 33, No 1(D), pp.58-60,2006
- [3] <http://research.nfrdi.re.kr/index.html?PageNo=3>