

무선 센서 네트워크에서 부분 속성값을 활용한 에너지 효율적인 질의처리

(Energy-efficient query processing based on partial attribute values in wireless sensor networks)

김 성 석 [†] 김 형 순 ^{**} 양 순 옥 ^{***}
(Sungsuk Kim) (Hyong-Soon Kim) (Sun Ok Yang)

요 약 유비쿼터스 컴퓨팅 환경이 일반화되면서 센서의 역할이 중요해지고 있다. 응용에 따라 단순히 주변의 환경 정보를 수집하는 기능보다는 그 자체가 계산 기능을 가지고 다양한 역할을 수행할 수 있게 되었다. 이러한 센서를 활용한 과제에서 중요한 고려사항 중 하나는 에너지의 효율성이다.

본 연구에서는 무선 센서네트워크에서 속성 질의 처리를 수행할 수 있는 알고리즘을 개발하고자 한다. 이를 위해 각 센서들은 모든 자식 노드들의 속성값에 대한 부분 정보를 유지하도록 한다. 하지만 정보의 양이 너무 커지면, 정보 유지 비용이 커지게 된다. 또한 정보의 갱신 비용 역시 무시할 수 없다. 따라서 각 노드가 수집한 속성값 자체를 전달하는 대신 그 값의 범위를 표현한 비트값 즉, AVB(Attribute-Value Bits),을 보내도록 한다. 이는 적은 공간으로 모든 자손노드들의 속성값에 대한 영역 범위를 유지할 수 있어서 질의 처리 과정동안 필요한 메시지의 수를 크게 줄일 수 있다. 이에 대한 실험을 통하여, 제안한 기법의 다양한 속성을 살펴보았다.

키워드 : 무선 센서 네트워크, AVB, 라우팅 경로 설정, 속성질의, 데이터 값의 갱신

Abstract Wireless sensors play important roles in various areas as ubiquitous computing is generalized. Depending on applications properties, each sensor can be equipped with limited computing power in addition to general function of gathering environment-related information. One of main issues in this environment is to improve energy-efficiency in sensor nodes.

In this paper, we devise a new attribute-query processing algorithm. Each sensor has to maintain partial information locally about attributes values gathered at its all descendent nodes. As the volume is higher, however, the maintenance cost also increases. And the update cost also has to be considered in the proposed algorithm. Thus, some bits, AVB(Attribute-Value Bits), are delivered instead of the value itself, where each bit represents a bound of attribute. Thus, the partial information can decrease the number of exchanged messages with a little cost during query processing. Through simulation works, the proposed algorithm is analyzed from several points of view.

Key words : wireless sensor network, AVB, routing path, attribute-query, data update

[†] 정 회 원 : 서경대학교 컴퓨터과학과 교수
sskim03@skuniv.ac.kr

^{**} 비 회 원 : 한국정보화진흥원
khs@nca.or.kr

^{***} 비 회 원 : 세종대학교 컴퓨터과학과 교수
soyang@korea.ac.kr
(Corresponding author)

논문접수 : 2009년 8월 24일

심사완료 : 2010년 3월 8일

Copyright©2010 한국정보과학회: 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 데이터베이스 제37권 제3호(2010.6)

1. 서 론

무선 센서 네트워크는 관련 기술이 발전함에 따라 전장지역 감시, 환경 모니터링, 재난지역 구조 등과 같이 응용 영역이 갈수록 넓어지고 있다. 센서 네트워크는 여러 자동화된 센서들로 구성된다. 각 센서는 임시 무선 통신 경로를 확보할 수 있으며 어떤 중앙 노드의 제어 없이 주변 지역에 대한 정보 수집기능을 담당할 수 있다. 경우에 따라 제한된 계산 능력 및 에너지 제약조건은 센서의 활용과정에서 중요한 고려사항이 되며, 특히 네트워크의 전체 수명을 늘리기 위한 연구가 다양한 영역에서 진행되고 있다.

오늘날 하나의 센서 가격을 낮춤으로써 센서 네트워크의 대중화를 이루는 것이 가능해졌다. 특히 센서의 소형화 및 부품의 저전력화를 실현함으로써 하나의 센서 내에 여러 센싱 장치를 부착하는 것도 가능해졌다. 예를 들면, 산불 감시 시스템에서 사용되는 센서는 온도 센서, 풍향 센서, 습도 센서 등을 부착하여 감시 및 진화에 필요한 정보를 수집할 수 있다. 이 경우, 사용자는 단순히 하나의 속성값을 대상으로 한 질의뿐만 아니라 여러 속성에 대한 조건을 부여하는 일반적인 질의도 요청할 수 있다.

```
select 센서location, 센서온도
from 무선네트워크
where 센서온도 > 30 and 센서습도 < 50 ;
```

위와 같은 질의에 대하여 네트워크 내의 각 센서들이 이웃노드들의 수집 값에 대한 정보를 미리 알고 있다면 질의처리 과정에서 필요한 메시지의 수를 충분히 줄일 수 있다. 하지만 이러한 정보를 교환하고 관리하기 위한 메시지의 수 역시 엄청날 수 있으며, 특히 각 센서에서 수집, 관리하는 정보가 지속적으로 변경된다면 메시지 전송으로 인하여 내부 배터리의 수명이 크게 단축될 것이다. 이와 같이 이웃노드에 대한 정보를 유지하기 위한 비용 역시 중요한 고려사항이 되어야 한다. 무선 센서 네트워크 환경은 기본적으로 제한된 자원을 기본 가정으로 하며, 따라서 지역적 저장장치의 크기 및 무선 통신 대역폭 역시 기존 시스템과 비교할 수 없을 만큼 매우 작다고 가정하기 때문이다.

에너지 효율적인 질의처리와 관련된 이전 연구는 크게 2가지 방향으로 구분할 수 있다. 먼저 센서 네트워크에서 해당 속성값을 획득하기 위한 연구가 있으며, 이 경우에는 효율적인 라우팅 경로 설정이나 스케줄링 기법을 개발하는데 초점을 맞추고 있다[1,2]. 이에 반해 센서 네트워크 전체 혹은 일부에 대한 속성값을 얻기 위한 방식으로써, 불완전한 값을 가정으로 한 근사값 질의가 있다. 저자들은 집계연산(예를 들면, SUM, MAX 등)을 효율적으로 처리하기 위한 이론적 근거나 모델을 제안하고, 이를 기반으로 한 알고리즘을 개발하였다[3,4].

본 연구에서는 기본 응용 대상을 다차원 속성 질의로 정하였다. 즉, 전체 네트워크에 질의를 전달하면, 해당 조건에 해당하는 센서 노드의 측정값을 얻어오는 방식을 가정한다. 먼저 라우팅 경로는 가장 일반적인 방식인 베이스 노드로부터 출발하여 전체 노드로 메시지를 전파하여 부모-자식 관계를 설정하도록 한다. 이때 자식노드들은 부모노드에게 자신의 수집한 측정값에 대한 부분값을 보내도록 하며, 그 크기는 너무 크지 않도록 한

다. 이를 위해 속성값 비트(Attribute-Value Bits:AVB) 방식을 제안한다. 이는 각 속성별로 최대값과 최소값을 N 등분한 후, 이를 N 비트 데이터로 만든다(초기값은 모두 0). 이후 각 센서들은 자신의 측정값에 해당하는 비트를 1로 설정한 후, 부모노드에게 전달하며, 부모노드는 자식노드들의 값 및 자신이 수집한 값을 OR 연산한 후 그 결과를 자신의 부모노드에게 전달하도록 한다. 이와 같은 방식으로 노드들은 라우팅 경로 상의 모든 자식 노드들에 대한 속성값에 대한 추정치를 저장할 수 있다. 이후 사용자 질의가 전달되면, 자신을 포함한 자식노드들이 질의 조건을 만족할 수 있는 경우에만 질의를 전파하는 방식으로 불필요한 메시지 수를 줄이려고 한다. 또한 각 센서에서 수집한 속성값이 갱신되는 경우라도, AVB 값이 변화가 없다면 부모노드에게 알려줄 필요가 없으므로 이 경우에도 메시지의 수를 크게 줄일 수 있다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 본 연구와 관련된 연구들을 정리한다. 이후 3장에서 제안하는 AVB 데이터를 기반으로 한 질의처리 알고리즘을 설명하고, 4장에서는 제안한 기법에 대한 다양한 성능 요소에 대한 실험 결과를 설명한다. 마지막으로 5장에서 본 논문의 결론을 맺는다.

2. 관련연구

본 논문에서는 센서 네트워크에서 사용자 질의처리에 관심을 가지고 있다. 센서 네트워크 환경에서 질의처리와 관련된 연구들은 크게 연속 질의(continuous query) 처리 및 근사값 질의(approximate query) 처리로 구분할 수 있다. 연속질의란 대상에 대하여 지속적으로 정보를 수집하여 전송하는 기능이다. 따라서 연속질의의 최적 스케줄링 기법 혹은 질의에 해당하는 데이터들을 에너지 효율적으로 수집하는 기법의 개발이 주요한 연구 대상이며, 이때 네트워크 위상(topology)이나 여러 다른 환경적 요소들이 고려되고 있다[1,2,5]. 근사값 질의는 기본적으로 센서 네트워크에서 각 센서들이 수집한 부정확한 값을 이용하여 질의 처리하려는 방식이다. 관련된 연구는 1) 적절한 수학적 모델링을 기반으로 하여 수집한 데이터들의 시공간적 의미 분석, 2) 비용 효율적으로 질의 처리 등이 있다[3,4,6,7]. 본 연구는 연속질의와 관련되어 있다. 즉, 라우팅 경로를 따라 하위 노드들이 수집한 부분 정보를 활용하여 질의처리의 효율성을 높이려고 한다. 하지만 부하가 많은 새로운 스케줄링 기법이나 라우팅 경로설정 방식을 개발하기 보다는 센서들끼리 부분 정보를 효율적으로 교환 및 지역적으로 관리하는 기법에 초점을 맞춘다.

센서 네트워크는 무엇보다 에너지의 효율성을 중요한

고려요소로 한다. 이 경우, 무선 통신은 센서 노드의 에너지 사용에 가장 크게 영향을 미치는 요소이며, 따라서 교환되는 메시지의 수를 줄이려는 기법이 여러 가지 제안되었다. 질의처리를 포함한 여러 응용에서 에너지 효율성을 높이기 위한 연구는 다시 크게 두 가지 방식으로 구분할 수 있다. 첫 번째는 효율적인 라우팅 프로토콜이나 전달 알고리즘을 개발하는 것이다. 예를 들면, [8]의 저자들은 다중 경로 라우팅을 사용하여 결합포용적인 라우팅 기법을 제안하였다. 먼저 멀티캐스트 트리를 만들고, 단일방향 안테나의 특성에 의한 브로드캐스트 기법을 사용하고자 하였다.

두 번째 방식은, 꼭 전송되어야 할 데이터만 선별하여 보내려는 연구이다. 예를 들면, 네트워크-내 데이터 수집(In-Network data aggregation)이 대표적인 방식이다. 이 기법들은 평균(AVG), 합(SUM) 등과 같은 집계 연산을 분산적인 방식으로 수행하도록 하는 것이다[9]. [10]의 저자들은 최대값을 수집하는 문제를 선형(linear) 프로그래밍 문제로 형식화한 후 반복적 알고리즘을 제안하였다.

이러한 연구에 비해 본 연구는 라우팅 경로 설정에 대한 특별한 가정을 하지 않으며, 또한 네트워크-내 데이터 수집과 같은 제한된 응용을 가정하지도 않는다. 즉 일반적인 라우팅 기법에서 각 센서가 하나 이상의 속성에 대한 정보를 수집하며, 이 정보를 이용하여 연속적인 질의 처리를 할 수 있는 기법을 제안한다.

3. 속성값 비트 기반 질의처리 기법

앞서 간단하게 언급하였듯이, 본 연구에서는 각 센서가 하나 이상의 속성값을 측정할 수 있다고 가정한다. 이를 위해 각 속성값마다 속성값 비트라고 하는 비트값을 저장하도록 한다. 본 장에서는 먼저 속성값 비트(Attribute-Value Bits:AVB)에 대하여 간단하게 설명한 후, 질의처리 과정에서 AVB를 활용하여 불필요한 메시지의 전송 회수를 줄이는 방법에 대하여 설명한다. 마지막으로 각 센서가 주기적으로 수집하는 속성값에 갱신이 일어나는 경우를 처리하는 방법을 설명한다.

3.1 속성값 비트

각 센서가 수집하고자 하는 속성 A_i ($i \geq 1$)에는 최대값(Max_i)과 최소값(Min_i)이 존재한다. 먼저 Max_i와 Min_i 사이의 어느 값을 표현하기 위해 N개의 동일 간격으로 나눈 후, 이를 N 개의 비트로 표현한다. 초기값은 모두 0으로 설정한다. 예를 들어 어느 대상 지역의 온도의 경우, 최대값은 40°C, 최소값은 -10°C로 가정하고 이를 일정 크기로 등분한다면(예를 들어 최대값-최소값 사이를 20 등분), 각 비트는 2.5°C간격으로 구분되며 이를 표현하기 위해 20비트가 있으면 된다. 그림 1은 설명을

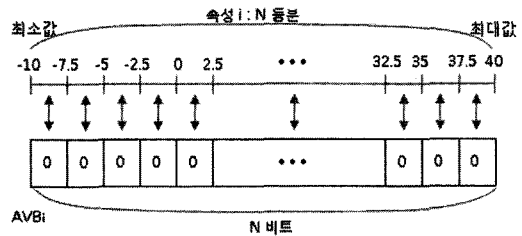


그림 1 온도에 대한 AVB 예

그림으로 표현한 것이다. 그림에서 온도에 대한 AVB는 20비트 크기로 가정하고 있다. 만약 어느 센서가 수집한 온도 값이 -4°C라면, 이는 -5°C보다 크고 -2.5°C 보다 작은 값이므로 세 번째 비트 값이 1로 설정하게 된다.

이때 N의 값이 커질수록 각 속성값 영역의 범위는 작아지므로, 실제 측정값과 유사하게 된다. 하지만 N의 값만큼의 비트가 할당되어야 하므로, AVB를 위한 저장 공간의 크기 역시 커지게 된다. 그림 1에서 온도의 최대값과 최소값을 100(=N) 등분한다면 이를 표현하기 위해 100비트가 필요하지만, 각 비트는 단지 0.5°C 범위를 가지므로 정확성은 높아지게 된다. 물론 AVB는 비트 단위로 할당되는 것이므로, 다른 기법보다 훨씬 적은 공간을 필요로 한다. 이후 3.3절에서 설명되지만, 각 등분의 크기가 작아진다면(즉, N의 값이 커진다면), 자신이 수집한 값이 갱신될 때마다 자신의 AVB값을 빈번하게 이웃노드들에게 전파하게 되는 문제가 발생할 수 있다.

전체 센서들이 이웃노드들의 AVB값을 저장하는 것은 특별한 라우팅 경로 설정 기법에 의존하지 않는다. 따라서 본 연구에서는 가장 기본적인 라우팅 경로 설정 방식인 부모-자식 관계를 이용하기로 한다[11]. 우선 각 센서는 고유한 ID 값을 가지고 있으며, 자신의 (절대적 혹은 상대적) 위치 정보를 알고 있다고 가정한다. 또한 동일한 무선 통신 범위를 가진다고 가정한다.

라우팅 경로 설정은 베이스 노드가 이웃 노드들에게 경로설정 메시지(PathMsg)를 전송함으로써 시작된다. PathMsg에는 다음과 같은 정보를 포함한다.

- Hops : 베이스 노드로부터의 홉 수 (초기값 = 0)
- (ID, Loc) : PathMsg를 전파하는 센서의 식별자 및 위치 정보

이 메시지를 수신하는 노드는 위치 정보를 고려하여 자신의 부모 노드를 결정하며, 다시 자신의 ID 및 위치 정보를 담은 PathMsg를 전파하게 된다. 일정 시간이 경과한 후, 단말 노드로부터 통지 메시지(NotiMsg)를 자신이 결정한 부모노드에게 전파하게 된다. NotiMsg

에는 자식노드의 ID 및 m개의 속성을 위한 m개의 AVB 정보가 포함되어 있다. 이 정보를 받은 노드는 자신의 AVB 정보와 함께 OR 연산을 수행한 후, 다시 자신의 부모노드에게 m 개의 AVB 정보를 전파하도록 한다.

앞서 PathMsg에 포함된 Hops 정보는 일정시간이 경과한 후 단말노드로부터 NotiMsg를 전송할 수 있도록 시간을 설정하기 위해 이용된다. 즉, 각 노드는 이 값을 이용하여 언제 통지 메시지를 전파해야하는지 시작시간까지의 대기시간(w_t)의 추정값을 계산할 수도 있다. 즉, 네트워크 내의 Hops의 최대값(max_{hop})을 상수로 정의해두었다면, 식 (1)에 의해 각 노드의 최대 대기 시간을 결정한다.

$$w_t \leftarrow (max_{hop} - Hops_a) \times \kappa \times 2 \quad (1)$$

수식에서 $Hops_a$ 는 센서 노드 a가 PathMsg 메시지로 부터 얻은 Hops 값이며, 상수 κ 는 각 센서에서 PathMsg 메시지를 처리하는데 걸리는 시간이다. 따라서 베이스 노드의 경우, 최대 대기시간인 $max_{hop} \times \kappa \times 2$ 만큼 기다려야 자신의 자식노드들로부터 부모선택 메시지를 받게 될 것이다. 만약 어느 노드가 w_t 시간만큼 기다렸는데도, 어떤 노드로부터도 부모 선택 메시지를 받지 못한다면, 자신을 단말노드(leaf node)로 결정한 후 NotiMsg 메시지를 전송한다.

각 센서는 자신이 수집한 값만 자신의 부모노드에게 전송하는 것이 아니라, 자신을 부모노드로 선택한 모든 자손노드로부터 전송된 속성값도 함께 전송해야 한다. 하지만 자손 노드의 개수와 상관없이 각 속성별로 하나의 AVB만 전송하도록 함으로써, 각 노드가 전송하거나 관리해야 할 데이터의 크기는 크게 줄일 수 있다.

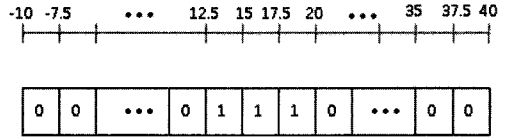
3.2 질의 처리

사용자의 질의는 베이스노드를 거쳐 전체 센서 네트워크로 전달된다. 이 질의를 받은 노드들은 먼저 질의 조건을 자신 및 자신의 자손들이 만족할 수 있는지를 결정한다. 이때 자신이 관리하고 있는 자식 노드들의 AVB 값을 이용한다. 이를 위해 앞서 그림 1에서 사용된 예제를 이용하기로 한다. 사용자의 질의는 다음과 같다.

```
select sensor.location, sensor.온도
from All sensors
where sensor.온도 > 14 and sensor.온도 < 18 ;
```

이 질의처리를 위해 먼저 조건절에 있는 14와 18 값에 해당하는 임의의 AVB_{temp}를 만든다. 이때 만들어진 AVB_{temp}는 14와 18사이에 해당하는 부분만 1로 설정되고 나머지는 0으로 설정된다(그림 2).

그런데 현재의 예제에서는 온도는 2.5°C 간격으로 구



AVB_{temp}
그림 2 그림 1을 이용하여 질의 처리과정 중 생성된 임시 AVB

분되므로, 실제로는 12.5부터 20까지가 모두 조건으로 설정된다. 이렇게 얻어진 AVB_{temp}와 자신 및 자손들의 AVB를 AND 연산을 수행한 후 그 값이 0이 아니라면 이 질의를 자손 노드들에게 전파한다.

3.3 갱신 처리

앞서 초기에 전체 네트워크내의 모든 센서들은 자신의 자손들에 대한 속성값에 대한 영역 범위(AVB)를 유지하게 된다. 하지만, 실제 시간이 지남에 따라 각 센서가 담당하는 지역의 대상 속성 값은 변하게 되며, 이를 센서들은 주기적으로 수집하게 된다. 만약 수집된 값이 이전에 수집된 값과 다르다면 갱신이 발생한 것이므로 이를 부모노드에게 알려주어야 하며, 이는 라우팅 경로상의 모든 노드들에게 갱신 메시지를 보내야 함을 의미한다. 이 갱신메시지는 자신이 받은 자식노드의 AVB에 자신의 새로운 속성 값을 OR 연산한 결과인 AVB를 보내면 되는 것이다.

하지만 제안한 AVB 기법은 부모 노드가 자식 노드들의 속성에 대한 영역 정보만 유지하도록 한다. 따라서 앞선 예제를 이용하여 다음과 같은 두 가지 경우에 따라 갱신 메시지가 상당히 줄어들 수 있음을 설명한다.

먼저, 노드 s₁에서 측정된 온도가 10°C에서 11°C로 증가하였다고 가정하자. 현재 예제에서 온도의 간격은 2.5°C이며, 10~12.5°C가 하나의 등분을 구성한다. 따라서 10°C에서 2.5°C가 증가하더라도 이는 부모노드에게 메시지를 보내지 않아도 된다. 즉, 갱신된 값이 부모에게 알려준 AVB_{temp}를 변화시키지 않는다면 부모에게 통지하지 않아도 된다(그림 3(a)).

다음으로 노드 s₁에서의 갱신을 부모노드 s₂에게 전달해야 하는 경우를 고려하자. 예를 들면, 온도가 10°C에서 9°C로 떨어진 경우이다. 부모 노드 s₂는 여러 자식 및 자손 노드들로부터 받은 AVB_{temp}를 유지하고 있다. 만약 s₁을 포함한 어느 자손 노드에서 이미 7.5~10°C 사이의 영역 값을 가지고 있다면, s₂ 노드는 자신의 부모인 s₃에게 이 사실을 알려줄 필요가 없다(그림 3(b)). 이처럼 속성값의 갱신을 처리하는 방식은 간단하지만, 실제 영역 정보를 기반으로 하므로 발생하는 메시지는 그렇게 많지 않을 수 있다.

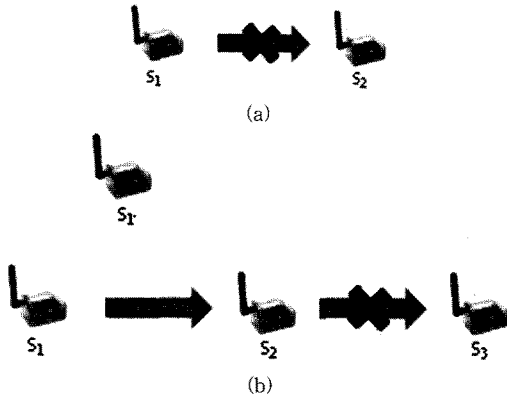


그림 3 갱신처리 과정 중, 불필요한 메시지를 줄이는 경우:

- (a) 갱신값을 부모노드에게 전달할 필요가 없는 경우,
- (b) 부모노드가 그 값을 자신의 부모에게 전달할 필요가 없는 경우

3.4 논의사항

제안하는 질의처리 방식에 대해 다음과 같은 사항들을 고려할 수 있다.

첫째, 주어진 예제에서는 각 등분의 간격이 커서 많은 메시지가 전파될 것으로 보여진다. 하지만 각 등분의 간격이 좁아진다면, 즉 각 속성에 대하여 더 많은 비트들이 할당된다면 센서들이 지역적으로 관리하는 정보의 정확도가 높아짐으로써 불필요한 메시지의 전파를 상당히 억제할 수 있다. 특히 대부분의 센서들이 비슷한 값을 가지고 있다면, 더 많은 비트가 할당되더라도 정보의 효율성이 떨어질 수 있다. 하지만 이런 경우는 굳이 센서를 이용하여 정보를 수집하는 의미가 희석될 수 있으므로, 본 기법에서는 응용의 대상으로 고려하지 않는다.

둘째, 조건이 두 가지 이상의 속성에 대한 AND 조건이라면, 불필요하게 생성되어 전파될 메시지의 수가 상대적으로 크게 줄어든다. 이는 지역적으로 유지하는 정보의 정확도가 낮더라도, 조건의 의미에 따라 그 정보의 유용성은 크게 높아질 수 있음을 의미한다.

셋째, 지역적으로 유지하는 정보의 정확성을 높이면, 값의 갱신으로 인한 부하가 커질 수 있다. 따라서 정보의 정확성과 갱신으로 인한 부하에 맞게, 속성별로 적절한 AVB 크기를 결정해야 한다.

이러한 논점에 대해서 실험을 통하여 자세하게 분석하기로 한다.

4. 성능 평가

4.1 실험 모델

본 연구에서는 무선 센서 네트워크에서 센서들의 에너지 효율성을 제고하기 위한 방법으로서 질의처리에

표 1 실험의 초기 설정값

변수	초기값	의미
<i>NumOfSensors</i>	1000	센서의 총개수
<i>RangeOfComm</i>	5	센서의 통신 반경
<i>MaxAttr_i</i>	50	속성 <i>i</i> 에 대한 최대값
<i>MinAttr_i</i>	-20	속성 <i>i</i> 에 대한 최소값
<i>Avg_i</i>	20	속성 <i>i</i> 에 대한 전체 평균
<i>StdDev_i</i>	20	속성 <i>i</i> 에 대한 표준편차
<i>QueryCondition</i>	5 %	질의 조건에서 속성값의 범위

필요한 메시지 수를 줄이는 알고리즘을 제안하였다. 즉 부모-자식 관계기반으로 하여 각 센서들간의 라우팅 경로를 설정하였으며, 이후 질의처리 과정에서 다양한 환경 요소를 변화시키면서 제안한 기법을 분석하도록 하였다. 이를 위해 Java 언어를 이용하여 시뮬레이터를 구현하였으며, 실험은 Intel 듀얼코어 CPU 3.2GHz의 Windows XP 기반 시스템에서 수행하였다. 실험과 관련된 내부 설정은 다음 표 1과 같다.

실험에서는 먼저 500×500 영역내에 1000개의 센서들이 배치되어 있으며, 센서들의 배치는 균등분포(uniform)되어 있다고 가정하였다. 표에서는 하나의 속성에 대한 최대값과 최소값, 그리고 표준편차를 보여주고 있다. 본 논문에서 하나를 제외한 모든 실험 결과는 하나의 속성에 대한 실험으로부터 얻었다. 이는, 질의 처리과정에서 하나의 속성만 이용하더라도 제안기법의 성능이나 경향을 이해할 수 있기 때문이다. 실험에서, 질의는 일반적인 SQL 형식(select-from-where)으로 구성된다. where 절에서는 속성값의 범위를 넘겨주며, 이 범위를 *QueryCondition*이라고 하며, 초기값은 ((최대값 - 최소값)의 5%라고 설정하였다. 예를 들어, “where attr₁ > 20 and attr₁ < 27”라고 하면, 질의 조건 범위는 7(= 27 - 20)가 되며, 이는 속성 *i*의 영역 범위(최대값 - 최소값 = 50 - (-20))의 10%에 해당한다. 마지막으로, 각 센서들은 무선 통신을 이용하여 상호간에 정보를 주고받을 수 있으며, 통신 반경은 5라고 설정하였다.

사용자들은 외부에서 베이스노드에게 일반적인 속성 질의를 전달하며, 이 질의는 초기에 설정된 라우팅 경로를 따라 네트워크로 전달된다. 이때 지역적으로 저장된 AVB 값을 이용하여 불필요한 메시지를 발생시키지 않도록 한다. 따라서 실험 결과는 대부분 질의 처리과정에서 질의를 전달하는데 사용된 메시지의 개수만을 보여주고 있다.

마지막으로 이 실험에서는 센서 노드의 실패나 통신 과정에서의 실패 및 잡음 등의 효과는 고려하지 않았으며, 모든 센서의 성능은 동일하다고 가정한다.

4.2 실험결과

(1) 센서들간 속성값의 유사도

첫 번째 실험은 센서들이 측정한 값들이 이웃한 노드

들간에 유사성이 있는지에 따라 메시지의 발생량을 평가하였다. 즉, Random 이란 노드들이 수집한 값이 완전히 서로 다른 경우를 의미하며, Normal은 그 반대 경우를 의미한다. 이때 Normal은 전체 영역을 4등분한 후, 각 등분면의 한 점을 중심으로 유사한 값들을 갖도록 설정하였다. 그림 4에서 x-축은 속성에 대한 AVB 구조에 할당된 비트의 개수를 의미하며, y-축은 메시지의 개수를 나타낸다.

Random한 경우에는 부모-자식 간에 특별한 값의 유사도가 없으며, 또한 전체적으로 보아도 값들의 분포가 제각각이다. 이에 반해, Normal인 경우에는 일정 노드들간에 값의 차이가 적으며¹⁾, 따라서 한 부모노드의 모든 자손노드들이 가진 값들은 어느 구간 내에 몰려 있게 된다. 제안한 AVB 기법은 기본적으로 자손노드들의 정확한 속성값들을 유지하는 대신, 값들의 범위를 유지하도록 한다. 따라서 이 실험은 속성값에 대한 각 값의 영역이 작을수록, 즉 속성에 대하여 더 많은 비트를 할당할수록 성능에 미치는 영향을 평가하기 위한 것이다. 당연히 그림 4의 결과에서 알 수 있듯이, 속성에 대한 최대값-최소값 구간을 좁게 등분할수록 질의처리 과정에서 발생하는 메시지의 수는 크게 감소함을 알 수 있다. 또한 부모-자식 노드들간에 값의 유사성이 많을수록 (Normal), 그 효과는 커짐을 알 수 있다.

실험에서는 속성에 대하여 최소 10비트에서 최대 40 비트를 할당한다고 설정하였는데, 이는 실제 무선 센서 환경에서도 그렇게 큰 값은 아니다. 예를 들어, 가장 대표적인 센서 시리즈인 Mote는 16~128Kbytes 플래시 메모리와 외부 저장공간을 가지고 있으며, 6LoWPAN 802.15.4 방식으로 무선 통신을 수행한다[12]. 이런 환경에서 제안기법은 40비트 정보를 이용하여 상당수의 메시지를 줄일 수 있다.

(2) 질의 조건에 포함된 속성값의 범위

이번 실험에서는 표 1에서 언급하였던 QueryCondition 값의 변화에 따른 질의 성능을 평가하였다. 본 연구에서 질의는 단순 속성 질의에 초점을 맞추고 있으며, 따라서 질의 조건에 포함된 속성값의 범위가 커질수록 보다 많은 수의 센서들이 이 질의처리 과정에 참여해야 한다. 그림 5와 그림 6은 실험 결과를 보여주고 있다. 여기서 x-축은 속성값 범위에 대하여 질의 조건이 몇 %를 차지하는지를 보여준다. 예를 들면, 최대값 50, 최소값 -20인 속성(전체 범위 : 70)에서 QueryCondition이 10%라면 조건 범위가 7이 됨을 의미한다. 이 실험부터는 모든 센서들의 속성값은 Normal(그림 4 참조)

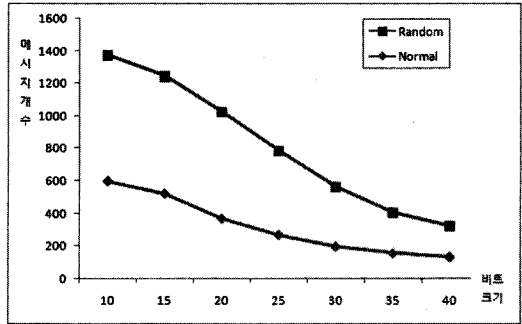


그림 4 속성값의 유사도에 따른 실험

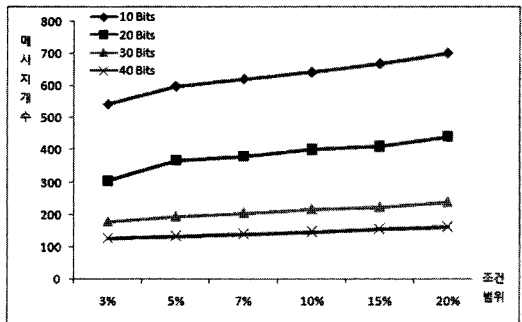


그림 5 질의에 포함된 속성값의 범위

한 경우만 실험하였다. 그림 5는 질의에 포함된 속성이 하나인 경우를 보여주고 있으며, 그림 6은 2개의 속성이 질의 조건에서 AND로 결합되어 있는 결과이다.

먼저 그림 5의 경우, 질의조건 범위가 커질수록 그리고 적은 비트값이 할당될수록 질의처리를 위해 많은 메시지가 전파되었음을 알 수 있다. 특히 지역 저장장소에 저장된 정보가 부정확한 경우(속성당 10비트 할당), 약 70%의 센서들이 질의를 전파하고 있음을 알 수 있다. 이에 반해 40비트가 할당된 경우에는 겨우 16%의 센서들이 질의를 전파하며, 두 차이에 해당하는 54%의 센서들은 자신의 자손들에게 전파할 필요가 없다고 판단하여 메시지를 보내지 않을 수 있었다. 결국 그만큼의 에너지를 낭비하지 않게 되었다.

이 실험에서 한 가지 더 언급할 점은, 자신이 유지한 정보가 정확해질수록 속성값의 범위가 커지더라도 메시지 수의 증가폭이 상대적으로 적다는 점이다. 40비트의 경우, 3%에서 20%로 증가하더라도 메시지의 수는 겨우 4%가 증가되었다. 이는 표 1에서 설명한 실험 환경 인자 중 StdDev의 영향 때문이다. 즉 본 실험은 센서들의 측정값이 주변 노드 값과 어느 정도 영향이 있으며 (Normal), 또한 그 값의 차이는 대략적으로 표준편차를 따르기 때문이다. 만약 이 실험을 Random 상황에서 실험하였다면, 그림 5보다는 상대적으로 더 많은 메시지가

1) 모든 센서들이 수집한 값들은 평균값을 기준으로 표준편차를 따르도록 되어 있다.

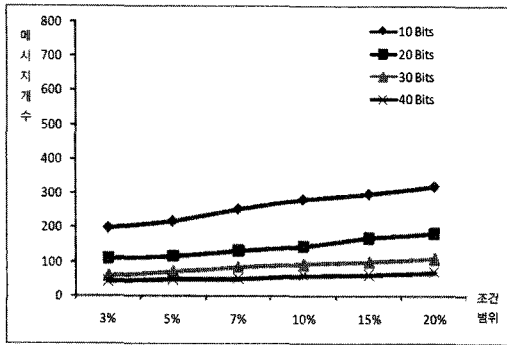


그림 6 두 개의 속성에 대한 AND조건 질의에서 속성값의 범위

필요하게 된다. 실제 실험을 수행한 결과, 20%의 경우 약 32%의 센서들이 메시지를 전파하여야 했다.

그림 6은 두 개의 속성을 대상으로 실험을 수행하였다. 실험의 편의를 위해 두 속성은 모두 동일한 최대값, 최소값, 평균, 표준편차를 가진다고 하였으며, 두 속성에 대한 질의 범위도 동일하다고 하였다. 그림은 이전 하나의 속성에 대한 실험(그림 5)과 비교를 위해 y-축의 범위를 동일하게 하였다. 두 속성에 대한 조건을 AND 연산을 수행하였으므로, 당연히 이에 해당하는 센서의 수는 크게 줄어들게 된다. 하지만 이러한 정보를 센서들이 지역 저장장소에 적은 비용으로 저장하고 있어야만 실제 혜택을 얻을 수 있게 된다. 본 실험에서는 속성당 최대 40비트를 할당하더라도, 두 속성이면 80비트(10 바이트)정도면 충분하다.

그림 5와 비교하면 모든 경우에서 메시지의 수가 상당히 줄어들었음을 알 수 있다. 특히 그 정보가 비록 부정확한 속성당 10비트를 할당했다라도 20% 범위에서 전체의 32% 센서들만 질의처리에 참여해도 되었다. 40비트인 경우에는 6.8%만 참여하였다. 결국 여러 속성을 다루는 경우에는 정확한 값이 아닌 값의 추정범위만 가지고 있더라도 상당한 성능 향상을 가져올 수 있음을 알 수 있다.

(3) 각 센서에서 속성값의 갱신처리 비용

이전 두 실험은 초기에 라우팅 경로를 설정하는 과정에서, 자식이 부모에게 자신 및 자손의 속성값을 의미하는 AVB 정보를 전달하도록 하였다. 모든 노드들은 자신의 자손 노드에 대한 추정값을 지역적으로 유지하게 되는 것이다. 하지만, 실제 응용에서는 시간에 지남에 따라 센서들이 주기적으로 혹은 다른 조건에 의해 속성값을 새롭게 측정해야 하며, 이때 이전의 값과 차이가 나게 된다. 즉, 속성값에 대한 갱신이 발생하였으며, 이를 모든 부모 노드들에게 전달해야 한다. 따라서 이번 실험은 각 센서들에서 갱신이 발생하였을 때 발생하는

비용을 평가한다.

그림 7에서 x-축은 센서가 수집한 새로운 값이 이전의 값과 비교했을 때 얼마만큼 갱신이 되었는지를 의미한다. 예를 들면, 1이라는 값의 의미는 이전에 비해 속성값이 1만큼 증가했음을 의미한다²⁾. 본 실험도 센서들의 값들은 Normal한 경우를 가정하며, 이를 위해 전체 영역이 크게 4등분되어 있다고 가정한다. 각 등분의 중심부분을 중심으로 인접한 노드들은 유사한 값을 가질 확률이 비교적 높다고 가정한다.

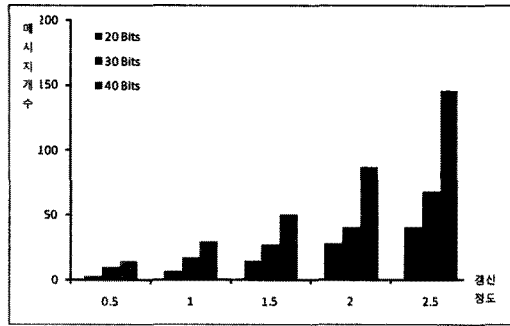
센서 S_0 의 속성값이 변한 경우에는 이웃한 노드의 값도 변경되었을 확률이 높다. 이를 처리하기 위해 4등분된 전체 실험 영역에서 S_0 가 위치한 등분의 일부 노드들이 영향이 받았다고 처리한다. 그림 7의 (a)는 가령 1/4분면에서 한 센서가 측정값이 변경되었다면, 그 센서를 중심으로 대략적으로 10%(전체로 보면 2.5%)의 센서들의 값들도 일정 비율로 변경되었다. 그림 7(b)는 20%(전체로 보면 5%)가 변경됨을 의미한다. 또한 갱신의 발생은 Poisson 분포를 따른다.

그림에서 얻어진 결과는 일반적인 내용을 담고 있다. 즉, 갱신되는 비율이 커질수록, 그리고 더 많은 인접한 노드들이 갱신할수록 갱신을 위한 메시지의 수는 증가한다. 그리고 지역적으로 유지하는 AVB의 값이 부정확할수록 갱신비용은 적게 발생한다. 이는 AVB 정보 자체가 값의 영역을 저장하는 방식이기 때문이다. 즉, AVB의 한 비트가 표현하는 값의 범위가 커질수록 센서들이 갱신값을 부모노드에게 전달할 필요성이 적어지기 때문이다.

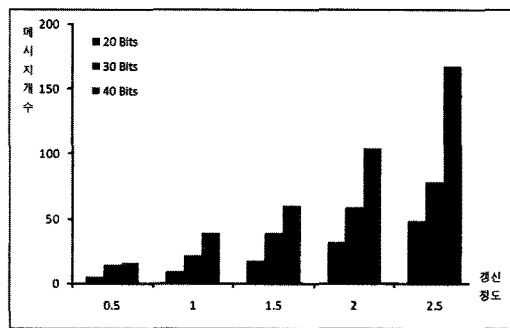
40비트를 할당하는 경우, 하나의 비트는 대략 1.75 범위를 표현한다. 따라서 1이하의 범위까지는 갱신이 발생하더라도 크게 증가하지 않는다. 하지만 그 이상이 되기 시작하면 갱신 메시지를 보내는 회수가 증가하게 된다. 그렇더라도 하나의 갱신마다 전체 노드의 2.5-5%까지 갱신이 발생하는 것이 비하면 갱신 메시지 처리 비용은 상대적으로 적을 수 있다.

여기서 한가지 언급할 점은, 그림 4의 결과에서는 비트의 크기가 클수록 메시지의 수가 줄어든다고 실험결과가 보여주고 있다. 왜냐하면 그 경우에는, 질의처리 과정에서 각 노드가 자신의 자손노드의 속성정보를 기반으로 하여 현재의 질의는 더 이상 전송할 필요가 없다고 결정하면, 보내지 않는 방식이다. 메시지의 수가 줄기 위해서는 정보가 정확해야 하므로, 비트의 크기가 큰 경우가 좋은 성능을 보인다. 이에 반해 그림 7의 경우는 질의처리과정이 아니라 네트워크에서 지속적으로

²⁾ 본 실험에서는 값의 감소는 처리하지 않고 단지 값이 x-축의 값만큼 증가된 경우만 고려하였다.



(a)



(b)

그림 7 각 센서에서 값의 갱신으로 인한 메시지 발생량

속성값들이 변경되는 경우 이 값들을 네트워크의 노드들로 전송하는 메시지의 수를 측정하였다. 비트의 크기가 크다는 것은 각 비트가 표현하는 값의 범위가 작아지며, 따라서 센서의 수집값이 조금만 변경되어서 부모에게 알려야할 필요성이 커지기 때문에 메시지의 수가 증가한 것이다.

5. 결론 및 향후 과제

기술의 발전과 함께 응용범위가 넓어지고 있는 센서 네트워크는 이전의 다른 시스템 환경과 다른 여러 제약 조건을 가지고 있다. 그 중에서 에너지의 제약조건은 많은 연구에서 고려되고 있다. 본 연구에서는 절의처리과정에서 에너지 소모에서 큰 영향을 차지하는 전파 메시지의 수를 줄이기 위하여 AVB 구조를 제안하였다. 이 구조는 부모와 자식 관계를 이용하여 각 노드가 자신의 자손 노드들의 속성값에 대한 영역정보를 저장하도록 하는 방식이다. 이 기법은 특별한 라우팅 기법이나 복잡한 정보의 교환을 필요로 하지 않으며, 지역적으로 저장해야 하는 정보의 양도 상대적으로 적다. 특히 속성값의 갱신이 발생하더라도 메시지의 폭주를 막을 수 있어서 자원의 제약이 중요한 환경에서는 그 의미를 가질 수 있다.

본 연구는 알고리즘을 구현하는 과정에서 설정한 가정 중에서 노드 및 통신과정에서의 실패를 고려하지 않았다. 본래 무선 센서 환경은 여러 가지 이유에 의해 실패가 발생할 수 있다고 가정하며, 수집하는 정보 자체도 오류가 발생하기 쉽다. 따라서 향후에는 이러한 실패에 대하여 효율적으로 처리할 수 있는 기법도 개발하려고 한다.

또한 제안한 AVB 기법은 속성당 무조건 많은 공간을 할당하는 것이 좋은 것은 아니다. 특히 갱신이 빈번하게 발생하며, 갱신되는 값의 범위가 커지는 응용범위에서는 차라리 비트 하나가 표현하는 영역의 범위가 큰 것이 유리할 수도 있다. 따라서 이러한 여러 경우에 최적의 값을 찾을 수 있는 수학적 모델을 개발하여 응용을 개발할 때 참고할 수 있도록 하려고 한다.

참고 문헌

- [1] D. Abadi, S. Madden, and W. Lindner, "Reed: Robust, efficient filtering and event detection in sensor networks," In *VLDB*, pp.769-780, 2005.
- [2] A. Silberstein, K. Munagala, and J. Yang, "Energy-efficient monitoring of extreme values in sensor networks," In *SIGMOD* conference, pp.169-180, 2006.
- [3] D. Chu, A. Deshpande, J.M. Hellerstein, and W. Hong, "Approximate data collection in sensor networks using probabilistic models," In *ICDE*, p.48, 2006.
- [4] A.Deshpande, C. Guestrin, W. Hong and S. Madden, "Exploiting correlated attributes in acquisitional query processing," In *ICDE*, pp.143-154, 2005.
- [5] X. Yang, H.-B. Lim, M.T. Özsu, and K.-L. Tan, "In-network execution of monitoring queries in sensor networks," In *SIGMOD* conference, pp.521-532, 2007.
- [6] L. Pan, J. Luo and J. Li, "Probing queries in wireless sensor networks," In *ICDCS*, pp.546-553, 2008.
- [7] A. Silberstein, R. Braynard, C.S. Ellis, K. Munagala, and J. Yang, "A sampling-based approach to optimizing top-k queries in sensor networks," In *ICDE*, p.68, 2006.
- [8] J. Considine, F. Li, G. Kollios, and J. Byers, "Approximate aggregation techniques for sensor databases," In *ICDE*, pp.449-460, 2004.
- [9] Y. Yao and J. Gehrke, "The cougar approach to in-network query processing in sensor networks," In *SIGMOD* record, vol.31, no.3, pp.9-18, 2002.
- [10] N. Sadagopan and B. Krishnamachari, "Maximizing data extraction in energy-limited sensor networks," In *IEEE INFOCOM*, pp.1717-1727, 2004.
- [11] I.F. Khan and M.Y. Javed, "A Survey on Routing

Protocols and Challenge of Holes in Wireless Sensor Networks," In *ICACTE*, pp.161-165, 2008.

[12] <http://robotics.eecs.berkeley.edu/~pister/SmartDust/>



김 성 석

1997년 고려대학교 학사. 1999년 고려대학교 석사. 2003년 고려대학교 박사(이학박사). 2003년~현재 서경대학교 컴퓨터과학과 조교수. 관심분야는 분산정보처리 시스템, 센서 네트워크 데이터베이스, 유비쿼터스 정보시스템



김 형 순

1995년 고려대학교 학사. 1997년 고려대학교 석사. 2009년 고려대학교 박사(이학박사). 1997년~현재 한국정보화진흥원 수석연구원. 관심분야는 Seamless mobility, IPv6, BcN(NgN), 센서 네트워크, 동시성 제어 및 회복 기법



양 순 우

1995년 고려대학교 학사. 2002년 고려대학교 컴퓨터학과 석사. 2006년 고려대학교 컴퓨터학과 박사(이학박사). 2006년 2007년 고려대학교 공과대학 정보통신연구소 연구교수. 2008년 연세대학교 전기전자공학부 BK21 TMS정보기술사업단 연구교수. 2009년~현재 세종대학교 전자정보공학대학 컴퓨터공학과 초빙교수. 관심분야는 차세대 네트워크 이동성 관리, 센서 네트워크 데이터처리, 유비쿼터스 정보시스템