

센서 네트워크에서의 데이터 수집을 위한 라우팅 기법

(Energy-Efficient Routing for Data Collection in Sensor Networks)

송인철[†] 노요한[†] 현동준[†] 김명호^{**}
 (In Chul Song) (Yohan Roh) (Dongjoon Hyun) (Myoung Ho Kim)

요약 연속 질의는 센서 네트워크에서 일반적으로 사용되는 질의로서, 한 번 요청되면 일정한 주기로 여러 번 실행되어 그 질의의 결과를 베이스 스테이션으로 모은다. 이것은 지속적으로 많은 메시지 전송을 유발시키므로 베이스 스테이션으로 데이터를 모으는 데 드는 통신 비용을 줄이는 것이 중요하다. 센서 네트워크에서 네트워크내 프로세싱 기법은 중간 노드에서 집계 질의의 결과 데이터들에 대해 부분 집계를 수행하거나 그 데이터들을 하나의 메시지에 담아 보냄으로써 메시지 전송 횟수를 줄이며, 이것은 결국 통신 비용을 줄인다.

본 논문에서는 질의별로 구성되는 라우팅 트리인 질의 기반 라우팅 트리(Query Specific Routing Tree: QSRT)를 제안한다. QSRT의 핵심 아이디어는 네트워크내 프로세싱의 기회를 최대화하는 것이다. QSRT는 질의가 전달되는 과정을 통해 생성되며, 질의의 결과들이 베이스 스테이션으로 전달되는 도중 가능한 한 빨리 서로 만나게 만들어 부분 집계 및 패킷 합병의 기회를 최대화시킨다. 수행된 실험의 결과는 제안된 QSRT가 기존의 라우팅 트리보다 메시지 전송 횟수를 18% 이상 줄일 수 있음을 보여 준다.

키워드 : 센서 네트워크, 질의 처리, 라우팅

Abstract Once a continuous query, which is commonly used in sensor networks, is issued, the query is executed many times with a certain interval and the results of those query executions are collected to the base station. Since this comes many communication messages continuously, it is important to reduce communication cost for collecting data to the base station. In sensor networks, in-network processing reduces the number of message transmissions by partially aggregating results of an aggregate query in intermediate nodes, or merging the results in one message, resulting in reduction of communication cost.

In this paper, we propose a routing tree for sensor nodes that qualify the given query predicate, called the query specific routing tree(QSRT). The idea of the QSRT is to maximize in-network processing opportunity. A QSRT is created separately for each query during dissemination of the query. It is constructed in such a way that during the collection of query results partial aggregation and packet merging of intermediate results can be fully utilized. Our experimental results show that our proposed method can reduce message transmissions more than 18% compared to the existing one.

Key words : Sensor Networks, Query Processing, Routing

· 본 연구는 한국과학재단 특장기초연구(R01-2006-000-10809-0)의 지원으로 수행되었음

[†] 학생회원 : 한국과학기술원 전산학과
 icsong@dbserver.kaist.ac.kr
 yhroh@dbserver.kaist.ac.kr
 djhyun@dbserver.kaist.ac.kr
^{**} 종신회원 : 한국과학기술원 전산학과 교수
 mhkim@dbserver.kaist.ac.kr
 논문접수 : 2005년 8월 17일
 심사완료 : 2006년 1월 7일

1. 서론

컴퓨팅 기기의 눈부신 발전으로 인하여 작은 기기 내에 컴퓨팅 기능과 통신 및 센싱 기능을 끌고루 갖춘 센서 노드들이 출현하였다[1]. 이러한 센서 노드들은 다양한 환경에 배치된 후 환경 변화를 감지하여 사람들에게 유용한 정보를 제공한다. 예를 들어, 동물들의 서식지를 관찰하거나[2,3], 빌딩의 안전성을 검사하는 등의 응용에

사용된다[4]. 베이스 스테이션(base station)은 사용자로부터 SQL과 비슷한 형태를 가진 질의를 요청 받아 센서 노드들에게 전달한다. 센서 노드들은 센싱한 데이터들 중 전달 받은 질의의 조건을 만족하는 데이터들을 멀티홉(multi-hop) 무선 통신 방식을 통해 베이스 스테이션 쪽으로 전송한다.

센서 노드들은 보통 소형의 배터리를 이용하여 동작하므로 사용할 수 있는 에너지원이 제한적이다. 예를 들어, 미카 모트[1]는 AA 배터리 2 개를 에너지원으로 사용하며, 길게는 6개월 동안, 그러나 최대 전력 소모 상태로는 고작 6일 동안만 동작이 가능하다[5]. [2]나 [3]에서와 같이 서식지를 관찰하는 용도로 사용될 때, 센서 노드는 6개월 이상 사람의 간섭 없이 서식지 동물들의 행태를 관찰하고 그 결과를 보고할 수 있어야 한다. 따라서 제한된 양의 에너지를 효율적으로 사용하는 것이 중요하다.

센서 노드들 간에 메시지를 전송하는 데 소모되는 에너지량이 애플리케이션 단에서 계산을 위한 CPU 프로세싱에 소모되는 에너지량 보다 수십 배 이상 많으므로 통신에 소모되는 에너지량을 줄이는 것이 중요하다[6]. 환경 모니터링의 용도로 사용되는 센서 네트워크에서 일반적인 질의 형태는 반복적이고 지속적으로 수행되는 연속 질의(continuous query)의 형태이다. 연속 질의는 한 번 요청되면 센서 노드들로부터 질의의 조건을 만족하는 질의의 결과 데이터들이 지속적으로 베이스 스테이션으로 전송되어야 한다. 따라서 결과 데이터들이 베이스 스테이션으로 한 번 전송되는 데 드는 통신 비용을 줄이는 것이 매우 중요하다.

[7,8]에서는 MIN, MAX, SUM, AVG, COUNT 등, 집계 질의(aggregate query)의 결과 데이터들이 베이스 스테이션으로 전송될 때 부분 집계(partial aggregation) 및 패킷 합병(packet merging)을 수행하여 메시지 전송 횟수를 줄임으로써 통신 비용을 줄이는 네트워크내 프로세싱(in-network processing) 기법을 제안하였다. 부분 집계는 각 센서 노드에서 여러 결과 데이터들을 모아 집계를 수행하여 하나의 결과 데이터를 만들어 내는 것을 말한다. 예를 들어, MIN 질의의 결과 데이터들에 대한 부분 집계는 결과 데이터들 중에서 가장 작은 값만을 남기고, SUM 질의에 결과 데이터들에 대한 부분 집계는 결과 데이터 값들을 합한 값 하나만을 남긴다. 각 센서 노드는 자신이 받은 여러 결과 데이터들을 개별적으로 메시지에 담아 베이스 스테이션 쪽으로 전송하는 것이 아니라, 그 결과 데이터들을 모아 부분 집계를 수행하여 하나의 결과 데이터를 만들고 그것을 하나의 메시지로 전송한다. 패킷 합병은 서로 다른 그룹에 속하여(GROUP BY 절에 의해) 하나의 결과 데

이타로 부분 집계될 수 없는 결과 데이터들을 하나의 메시지 내에 담아 보냄으로써 메시지 내에서 큰 비중을 차지하는 메시지 헤더 부분이 여러 번 전송되는 오버헤드를 줄인다. 부분 집계 및 패킷 합병에 의해 각 센서 노드가 작은 크기의 하나의 메시지를 전송하는 것은 여러 개의 결과 데이터들을 개별적으로 전송하는 경우보다 에너지 소모량을 크게 줄일 수 있다.

질의의 결과 데이터들을 베이스 스테이션으로 모으기 위해 베이스 스테이션을 루트로 하는 라우팅 트리(routing tree)가 사용된다. 라우팅 트리는 라우팅 트리 구성 메시지가 베이스 스테이션으로부터 센서 네트워크 전체로 전달되는 과정을 통해 구성된다. 이 때 각 센서 노드는 자신에게 가장 먼저 라우팅 트리 구성 메시지를 보내 온 센서 노드를 부모 노드로 선택한다[7-9]. 기존의 라우팅 트리 구성 기법에서 각 노드가 자신에게 가장 먼저 라우팅 트리 구성 메시지를 보내 온 노드를 부모 노드로 선택하는 이유는 가장 가까이 위치한 노드와의 통신 품질이 가장 좋을 것이라고 가정하고 있기 때문이다. 그러나 [10]에서 지적된 바와 같이 패킷 손실이 반드시 거리에 비례하는 것은 아니다. 따라서 더 가까이 위치한 노드와의 통신 품질이 더 좋지 않아 잦은 메시지 재전송으로 인해 에너지를 더 많이 소모할 수 있다. 또한 기존의 라우팅 트리는 질의와 관련 없이 구성된다. 일단 한 번 구성된 라우팅 트리는 여러 질의에 대한 결과 데이터들을 베이스 스테이션으로 모으는 데 사용된다. 따라서 기존의 라우팅 트리는 질의별로 해당 질의의 결과 데이터를 서로 다른 부모 노드 쪽으로 전송함으로써 부분 집계 및 패킷 합병의 기회를 증대시킬 수 있는 기회를 얻지 못한다.

본 논문에서는 질의별로 구성되는 라우팅 트리인 질의 기반 라우팅 트리(Query Specific Routing Tree: QSRT)를 제안한다. QSRT는 결과 데이터들이 베이스 스테이션으로 전송되는 도중 가능한 한 서로 빨리 만나게 만들어 부분 집계 및 패킷 합병의 수행 시점을 앞당기고 그 기회를 최대화시킴으로써 전송되는 메시지 횟수를 더욱 줄인다.

그림 1은 기존의 라우팅 트리와 QSRT를 비교한 예이다. 그림 1에서 회색으로 표시된 3 번과 4 번 센서 노드는 질의의 결과 데이터를 베이스 스테이션 쪽으로 전송하고자 하는 노드들이며 화살표는 라우팅 트리를 따라 메시지가 전송되는 경로를 나타낸다. 그림 1의 (a)는 기존의 라우팅 트리 구성 기법에 의해 4 번 노드가 2 번을 부모 노드로 선택한 라우팅 트리를, (b)는 4 번 노드가 3 번 노드에서 결과 데이터가 전송될 것이라는 정보를 기반으로 3 번 노드를 선택한 QSRT를 보여준다. 그림 1의 (a)에서는 3 번의 메시지 전송이, (b)에서는 3

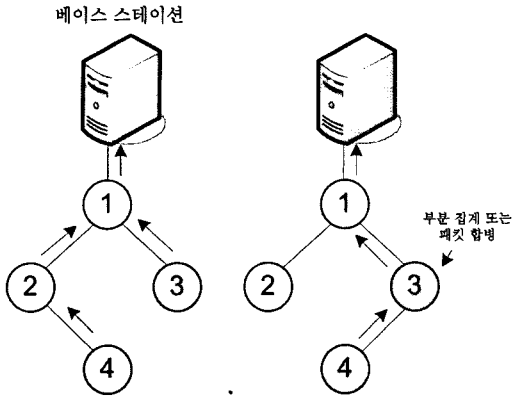


그림 1 기존의 라우팅 트리와 QSRT의 비교

번 노드에서 3 번과 4 번의 두 결과 데이터들이 만나 부분 집계 또는 패킷 합병이 수행되어 하나의 메시지에 담겨져 전송되었으므로 두 번의 메시지 전송이 발생하였다. 그림 1을 통해 알 수 있듯이 QSRT는 부분 집계 및 패킷 합병의 기회를 촉진시켜 메시지 전송 횟수를 더욱 줄인다.

본 논문의 공헌은 다음과 같이 요약된다.

- 질의별로 구성되는 라우팅 트리인 QSRT를 제안한다. QSRT는 결과 데이터들이 베이스 스테이션으로 전송되는 도중 가능한 한 서로 빨리 만나게 만들어 부분 집계 및 패킷 합병의 수행 시점을 앞당기고 그 기회를 최대화시킨다.
- 에너지 효율적인 QSRT의 구성 기법을 제안한다. QSRT는 질의 메시지가 전달되는 과정을 통해 구성된 질의 메시지에 추가적으로 최단 거리에 위치한 결과 데이터를 생성할 가능성이 있는 노드까지의 거리 정보가 기록되어 전달된다. 질의 전달 과정은 한번만 수행되므로 이 정보는 한 번만 전달되면 되고 이 또한 QSRT를 결과 데이터들을 베이스 스테이션으로 모으는 데 사용할 경우 얻을 수 있는 지속적인 메시지 전송 감소 효과에 비하면 무시할 수 있는 비용이다.
- 다양한 실험을 통해 기존의 라우팅 트리와 QSRT의 성능을 비교하였다. 실험의 결과는 QSRT를 사용할 경우 기존 라우팅 트리보다 메시지 전송 횟수를 18% 가까이 줄일 수 있음을 보여 주며, 센서 노드들이 조밀하게 배치된 센서 네트워크 환경에서는 메시지 전송 횟수를 28% 이상 줄일 수 있음을 보여 준다.

본 논문의 구성은 다음과 같다. 먼저, 2절에서 센서 네트워크에 대한 배경 지식과 관련 연구에 대해 설명한다. 3절에서는 QSRT를 정의하고 QSRT 구성 절차를 설명한다. 4절에서 QSRT와 기존의 라우팅 트리의 성능

을 비교하며, 마지막으로 5절에서 본 논문의 연구 결과를 요약한다.

2. 배경 지식 및 관련 연구

2.1 센서 네트워크 개요

일반적으로 센서 네트워크에서 발생하는 데이터는 *sensors*라 불리는 가상적인 릴레이션으로 모델링된다 [7,8]. 이 가상 릴레이션의 튜플들은 센서 네트워크 상의 각 센서 노드들이 생성한다. 각 센서 노드에서 생성되는 튜플은 노드의 아이디, 위치, 빛, 온도, 습도 등의 애트리뷰트들을 가진다. *베이스 스테이션(base station)*은 사용자에게 *sensors* 릴레이션에 대한 질의 인터페이스를 제공한다. 사용자는 센서 네트워크로부터 정보를 얻기 위해 베이스 스테이션에서 SQL과 유사한 형태를 가지는 질의를 요청한다. *연속 질의(continuous query)*는 반복적이고 지속적으로 실행되는 질의로서 동물의 서식지 관찰과 같은 지속적인 환경 모니터링이 필요한 센서 네트워크의 응용에서 일반적으로 사용되는 질의 형태이다. 연속 질의는 질의가 반복적으로 실행되는 주기와 질의의 수행 기간을 포함한다. 한 번 요청된 연속 질의는 각 센서 노드에 저장되었다가 일정한 주기마다 한 번씩 실행되고 질의의 결과 데이터들은 라우팅 트리를 따라 베이스 스테이션으로 수집된다. 센서 네트워크에서는 연속 질의 중에서 결과 값들의 분포를 알아 보고자 하는 COUNT, MIN, MAX, SUM, AVG 등의 *집계 질의(aggregate query)*가 많이 사용된다[8,11]. 예를 들어, 빌딩 내의 화재 발생 여부를 모니터링하기 위해 최고 온도 값을 알아 보거나 비어 있는 방을 알아내기 위해 방별 평균 음량 값을 알아 보기 위해 집계 질의를 사용할 수 있다. 각 집계 질의의 의미는 표준 SQL에서의 의미와 동일하다.

본 논문에서는 집계 질의의 결과 데이터들을 에너지 효율적으로 베이스 스테이션으로 모을 수 있는 라우팅 트리 구성 방안에 대해 논의한다. 추가적으로 본 논문에서는 집계 질의들 중에서도 고정 값을 가지는 애트리뷰트가 조건절에 하나 이상 포함되는 질의 처리를 논의한다. 고정 값을 가지는 애트리뷰트들의 예로는 센서 노드의 아이디, 센서 노드의 위치(x, y 축 정보), 고도(층수) 등이 있다. 이러한 질의는 센서 네트워크의 특정 지역이 어떤 상황인지를 알아 보기 위해 빈번히 사용되는 질의 형태이다. 따라서 이러한 질의들에 대한 보다 효과적인 질의 처리 방안이 요구된다.

센서 네트워크에서의 질의 처리는 *질의 분배(dissemination)* 단계와 *데이터 수집(collection)* 단계로 이루어진다[7]. 질의 분배 단계에서는 사용자의 질의가 센서 네트워크 상의 모든 센서 노드들에게 전달된다. 데이

타 수집 단계에서는 질의의 결과 데이터들이 라우팅 트리를 따라 베이스 스테이션으로 전송된다. 데이터 수집 단계에서 연속 질의 처리는 다음과 같이 진행된다. 먼저, 각 센서 노드는 질의 분배 단계에서 전달 받은 질의를 로컬 스트리지에 저장한다. 그 다음으로 각 센서 노드는 질의에 지정된 주기마다 한 번씩 깨어나 환경 정보를 센싱하고 센싱된 데이터가 질의의 조건을 만족할 경우 그 데이터를 라우팅 트리를 따라 베이스 스테이션으로 전송한다.

데이터 수집 단계에서는 결과 데이터들을 베이스 스테이션 쪽으로 모으는 데 드는 메시지 전송 횟수를 줄이기 위해 네트워크내 프로세싱이 수행된다. 네트워크내 프로세싱은 라우팅 트리의 하단부에서 시작되어 루트 노드까지 레벨별로 수행된다. 각 센서 노드는 레벨에 따라 메시지를 송신하고 수신하는 시점을 할당 받는다[5]. 이에 따라, 레벨 l 의 센서 노드들이 메시지를 송신할 때 레벨 $l-1$ 의 센서 노드들은 메시지를 수신한다. 각 센서 노드는 자신이 메시지를 송신할 시점이 되면, 자식 노드들로부터 받은 데이터들과 자신의 데이터를 모아 부분 집계 및 패킷 합병을 수행하고 그 결과를 하나의 메시지에 담아 부모 노드에게 전송한다.

2.2 기존의 라우팅 트리 구성 절차

기존의 라우팅 트리는 다음의 절차를 통해 구성된다 [7,8]. 먼저 베이스 스테이션은 라우팅 트리 구성을 위해 라우팅 트리 구성 메시지를 준비하여 루트 노드에게 전달한다. 라우팅 트리 구성 메시지의 포맷은 다음과 같다 (라우팅 트리 구성 절차의 논의와 상관 없는 필드들은 표시하지 않았음).

$\langle dest_id, src_id, level \rangle$

라우팅 트리 구성 메시지에서 $dest_id$ 와 src_id 필드는 각각 라우팅 트리 구성 메시지를 받을 센서 노드의 아이디와 보내는 센서 노드의 아이디를 저장하고, $level$ 필드는 라우팅 트리 구성 메시지를 보내는 센서 노드의 레벨을 저장한다. 루트 노드는 라우팅 트리 구성 메시지의 src_id 와 $level$ 필드를 각각 자신의 아이디와 레벨 (루트 노드의 레벨은 0)로 설정하고, $dest_id$ 필드를 방송 주소(*broadcast address*)로 설정한 후 이 메시지를 이웃 노드들에게 방송한다. 루트 노드를 제외한 각 센서 노드는 레벨이 정해지지 않은 상태에서 이웃 노드로부터 라우팅 트리 구성 메시지를 받으면, 라우팅 트리 구성 메시지의 $level$ 필드의 값에 1을 더한 값을 자신의 레벨로 설정하고 라우팅 트리 구성 메시지를 보낸 센서 노드를 자신의 부모 노드로 선택한다. 그리고 라우팅 트리 구성 메시지의 src_id 와 $level$ 필드를 각각 자신의 아이디와 레벨로 설정한 후 이 메시지를 이웃 노드들에게 방송한다. 이러한 과정은 모든 센서 노드들의 레벨과

부모 노드가 정해질 때까지 계속된다. 본 논문에서는 기존의 라우팅 트리를 FRT(Flooding-based Routing Tree)라 부르겠다.

2.3 관련 연구

[5]에서 제안된 Semantic Routing Tree(SRT)는 질의 분배 단계에서 질의 전달에 드는 통신 비용을 줄이기 위해 사용되는 라우팅 트리이다. SRT의 핵심 아이디어는 질의가 모든 센서 노드들에게 전달될 필요가 없다는 것이다. 따라서 질의를 전달 받을 필요가 없는 센서 노드들에게 질의를 전달하지 않으므로써 질의 전달에 드는 통신 비용을 절약한다. SRT는 질의 분배 단계에서 사용되는 메시지 전송 횟수를 줄이기 위한 라우팅 트리인 반면, QSRT는 데이터 수집 단계에서 네트워크내 프로세싱의 효과를 최대화하기 위한 라우팅 트리이다.

[12]에서 제안된 Group-Aware Network Configuration (GaNC)은 같은 그룹에 속하는 센서 노드들이 라우팅 트리의 같은 경로 상에 모이도록 라우팅 트리를 구성한다. 이렇게 하는 이유는 부분 집계 시, 서로 다른 그룹에 속하는 결과 데이터들에 대한 부분 집계 결과들은 하나의 데이터로 합칠 수 없고 따로 유지해야 하기 때문이다. 따라서 가능한 한 동일한 그룹에 속하는 결과 데이터들끼리 부분 집계에 참여하게 만드는 것이 보다 적은 개수의 부분 집계 결과들이 메시지 내에 담겨 전송될 것이므로, 전송되는 메시지의 크기를 더 줄일 수 있다. GaNC와 QSRT 구성 기법의 차이점은 GaNC는 결과 데이터가 속하는 그룹 정보를 활용하는 반면, QSRT 구성 기법은 결과 데이터가 위치한 곳까지의 거리 정보를 활용한다는 점이다. GaNC에 의해 구성되는 라우팅 트리는 가능한 한 동일한 그룹에 속하는 결과 데이터들끼리 부분 집계 및 패킷 합병에 참여하게 만들고, QSRT는 부분 집계 및 패킷 합병이 수행되는 시점을 앞당기고 그 기회를 촉진한다.

3. 질의 기반 라우팅 트리

그림 1의 예를 통해 살펴 보았듯이 QSRT는 데이터 수집 단계에서 결과 데이터들이 베이스 스테이션 쪽으로 전달되는 도중 서로 만나게 만들어 부분 집계 및 패킷 합병의 기회를 촉진시킨다. 이에 덧붙여 QSRT는 결과 데이터들이 서로 만나는 시점을 앞당긴다. 이렇게 하는 이유를 다음의 예를 통해 설명한다.

예를 들어, 센서 노드 i 가 어떤 질의에 대한 결과 데이터를 라우팅 트리를 따라 베이스 스테이션과 연결된 루트 노드까지 전달해야 하는 경우를 생각해 보자. 센서 노드 i 의 결과 데이터가 루트 노드까지 전송될 때는 그것이 다른 결과 데이터와 만나서 부분 집계 및 패킷 합

병에 참여하는 경우와 그렇지 않은 경우가 있다. 그 예로는 각각 그림 1의 (a)에서 4 번 노드의 결과 데이터가 루트까지 전송되는 과정과 (b)에서 루트 노드까지 전송되는 과정을 들 수 있다. 예를 들어, 센서 노드 i 의 결과 데이터가 센서 노드 j 에 이르러 부분 집계 및 패킷 합병에 참여하였다고 하면, 그곳에서부터는 센서 노드 j 의 결과 데이터와 합쳐져 센서 노드 j 부터는 하나의 메시지로 전송된다. 즉, 센서 노드 i 의 결과 데이터는 센서 노드 j 에서 전송될 메시지에 무임승차하는 효과를 얻을 수 있다. 여기서 센서 노드 i 는 무임승차 효과로 인해 자신의 결과 데이터를 루트 노드까지 전송하는 데 드는 메시지 전송 횟수 중에서 센서 노드 j 의 메시지가 루트 노드까지 전송되는 데 드는 메시지 전송 횟수만큼, 즉 센서 노드 j 의 레벨 수만큼을 절약할 수 있다. 따라서 센서 노드 i 는 자신의 결과 데이터를 또 다른 결과 데이터와 만날 수 있는 노드 쪽으로 전송하되, 가능한 한 최단 거리에 위치한 상위 레벨의 노드 쪽으로 보내는 것이 무임승차의 효과를 최대화할 수 있는 것이다.

이렇게 결과 데이터들이 가능한 한 빨리 서로 만날 수 있게 하기 위해서 QSRT를 구성할 때 센서 노드들 간에 교환되는 핵심 정보는 최단 거리에 위치한 결과 데이터 생성 노드까지의 거리 정보이다. 각 노드는 이 거리 정보를 바탕으로 자신의 결과 데이터를 가장 빨리 다른 결과 데이터 생성 노드까지 전송해 줄 수 있는 부모 노드를 선택하여 QSRT를 구성한다.

질의마다 질의의 조건에 따라 결과 데이터 생성 노드들의 위치가 달라질 수 있으므로, 질의별로 서로 다른 QSRT가 구성되고 각 센서 노드는 질의별로 서로 다른 부모 노드 쪽으로 해당 질의의 결과 데이터를 전송한다. 기존의 라우팅 트리는 여러 질의 처리에 동일하게 사용되는 것으로서 질의별로 서로 다른 부모 노드 쪽으로 결과 데이터를 전송하여 부분 집계 및 패킷 합병의 기회를 증대시킬 수 있는 기회를 얻지 못한다.

QSRT는 질의 분배 단계를 통해 구성된다. 각 센서 노드는 질의 분배 단계에서 질의 메시지를 이웃 노드에게 전달할 때 자신을 통해 결과 데이터를 루트 노드 쪽으로 전달할 경우(즉 자신을 부모 노드로 선택할 경우) 그 결과 데이터가 몇 홉 이후에 다른 결과 데이터 생성 노드와 만나는지에 대한 정보를 질의 메시지에 기록하여 전달한다. 각 센서 노드는 질의 메시지에 기록된 정보를 바탕으로 이웃 노드들 중에서 자신의 결과 데이터를 가장 빨리 다른 결과 데이터와 만나도록 전송해 주는 센서 노드를 자신의 부모 노드로 선택하여 QSRT를 구성한다.

3.1 QSRT 정의

센서 노드 i 의 이웃 노드는 센서 노드 i 의 통신 반경

내에 존재하는 센서 노드를 의미한다. 센서 노드 i 의 레벨 l_i 는 루트 노드로부터 떨어진 거리(홉 수)를 의미한다 (루트 노드의 레벨은 0.) 각 센서 노드는 이웃 노드들 중 몇 개를 후보 부모 집합으로 유지하고 그 중에서 하나를 특정 질의에 대한 결과 데이터를 전송할 부모 노드로 선택한다. 센서 노드 i 의 후보 부모 집합 CP_i 는 다음과 같이 정의된다.

$$CP_i = \{ j : \text{센서 노드 } j \text{는 센서 노드 } i \text{의 이웃 노드이고 } l_j = l_i - 1 \}$$

즉, 후보 부모 집합은 센서 노드 i 의 레벨 보다 하나 작은 레벨을 가지는 이웃 노드들이다.

예제 1. 그림 2는 각 센서 노드의 후보 부모 집합을 보여 준다. 그림에서 각 노드는 후보 부모 노드들과 선으로 연결되어 있으며 각 노드의 오른쪽에는 후보 부모 집합이 표시되어 있다. 예를 들어, 레벨 2의 5 번 노드는 이웃 노드들인 4, 2, 3, 6, 7 번 노드들 중에 레벨이 하나 작은 2, 3 번 노드를 후보 부모 집합으로 가진다. □

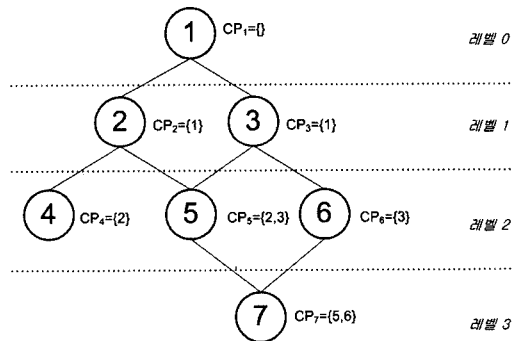


그림 2 센서 노드들의 후보 부모 집합

온도 애틀리뷰트와 같이 그 값이 시간에 따라 변하는 애틀리뷰트를 동적 애틀리뷰트(dynamic attribute)라고 하고 위치 애틀리뷰트와 같이 고정된 값을 가지는 애틀리뷰트를 정적 애틀리뷰트(constant attribute)라고 한다. 어떤 질의 Q의 WHERE 절은 정적 애틀리뷰트를 하나 이상 포함한다고 가정한다. 질의 Q 후보 노드는 질의 Q WHERE 절의 정적 애틀리뷰트에 대한 조건을 만족하는 노드다.

예제 2. 다음 질의 Q₁은 빌딩 내에 빈 방의 여부를 알아 보는 질의다[5].

```
Q1: SELECT AVG(volume), room FROM sensors
WHERE floor = 6
GROUP BY room
HAVING AVG(volume) > threshold
SAMPLE PERIOD 30s
```

질의 Q_i 는 6 층에 있는 센서 노드들을 방별로 그룹 짓고 방별 평균 음량 값이 threshold를 넘는 방을 보고 한다. 여기서 정적 애트리뷰트인 floor 애트리뷰트에 대한 조건을 만족하는 6 층에 위치한 센서 노드들이 질의 Q_i 후보 노드가 된다. □

지금부터 QSRT를 구성하기 위해 센서 노드들 간에 교환되는 핵심 정보인 **최단 거리(Minimum Distance: MD)**를 정의한다. 센서 노드 i 의 질의 Q 에 대한 최단 거리 $MD_{i,Q}$ 는 센서 노드 i 를 포함한, 상위 레벨에 있는 최단 거리에 위치한 질의 Q 후보 노드까지의 거리를 의미한다. $MD_{i,Q}$ 는 다음과 같이 정의된다.

$$MD_{i,Q} = \begin{cases} 0, & \text{센서 노드 } i \text{가 루트 노드이거나, 질의 } Q \text{ 후보 노드일 경우} \\ \min \{MD_{i,Q'} : j \in CP_j\} + 1, & \text{이외의 경우} \end{cases}$$

센서 노드 i 가 루트 노드일 경우 상위 레벨에 존재하는 노드가 없으므로 $MD_{i,Q}$ 를 0으로 정의한다. 센서 노드 i 가 질의 Q 후보 노드이면, $MD_{i,Q}$ 는 자신까지의 거리, 즉 0이다. 센서 노드 i 가 질의 Q 후보 노드가 아니라면 $MD_{i,Q}$ 를 후보 부모 노드들의 MD 값 중 가장 작은 값에 1을 더한 값으로 정의한다.

예제 3. 그림 3은 그림 2와 동일한 설정에서 센서 노드들의 질의 Q 에 대한 MD 값을 보여 준다. 그림에서 회색으로 표시된 2, 6, 7 번 노드들은 질의 Q 후보 노드이며 각 노드의 오른쪽에는 MD 값이 표시되어 있다. 1 번 노드는 루트 노드이므로 MD 값이 0이다. 2 번 노드는 자신이 질의 Q 후보 노드라서 MD 값이 0이고 3 번 노드는 질의 Q 후보 노드가 아니므로 MD 값은 후보 부모 노드인 1 번 노드의 MD 값에 1을 더한 1이다. 5 번 노드의 경우 질의 Q 후보 노드가 아니므로 후보 부모 노드 2, 3 번 노드 중에 가장 작은 MD 값, 즉 2 번 노드의 MD에 1을 더한 1을 MD 값으로 가진다. □

센서 노드 i 의 후보 부모 MD 집합 $CP-MD_{i,Q}$ 는 질의 Q 에 대한 후보 부모 노드들의 MD 값을 저장한다. $CP-MD_{i,Q}$ 는 (후보 부모 노드 아이디, MD 값) 쌍들을

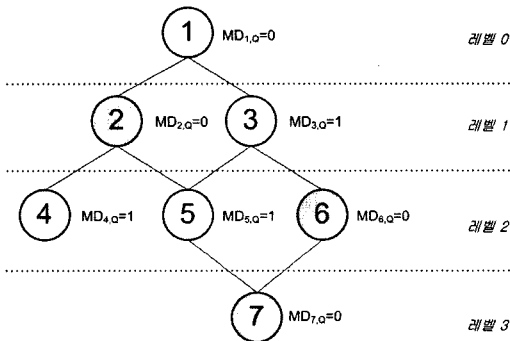


그림 3 센서 노드들의 질의 Q 에 대한 MD 값

저장한다. 예를 들어, 그림 3의 5 번 노드의 후보 부모 MD 집합 $CP-MD_{5,Q}$ 는 $\{(2,0), (3,1)\}$ 이다.

질의 Q 에 대한 센서 노드 i 의 부모 노드 $P_{i,Q}$ 는 후보 부모 노드들 중 가장 작은 MD 값을 가진 노드다. 즉, $P_{i,Q} = \text{MinDistId}(CP-MD_{i,Q})$ 이다. 여기서 $\text{MinDistId}()$ 함수는 $CP-MD_{i,Q}$ 집합에서 가장 작은 MD 값을 가진 노드의 아이디를 반환한다. 이 때, 가장 작은 MD 값을 가진 노드가 두 개 이상일 경우에는 무작위로 그 중 하나를 선택하여 반환한다.

센서 노드가 1, ..., n 까지 있다고 할 때 질의 Q 에 대한 QSRT T_Q 는 다음과 같이 정의된다.

$$T_Q = (V, E), \quad V = \{1, \dots, n\}, \quad E = \{(i, P_{i,Q}) : i, P_{i,Q} \in V\}.$$

예제 4. 그림 4는 그림 3에 이어 각 센서 노드가 부모 노드를 선택하여 구성된 질의 Q 에 대한 QSRT T_Q 를 보여 준다. 그림에서 5 번 센서 노드는 2 번과 3 번 후보 부모 노드 중에서 MD 값이 작은 2 번 노드를 부모 노드로 선택하였고, 7 번 센서 노드는 5 번과 6 번 센서 노드 중에서 MD 값이 작은 6 번 센서 노드를 부모 노드로 선택하였다. □

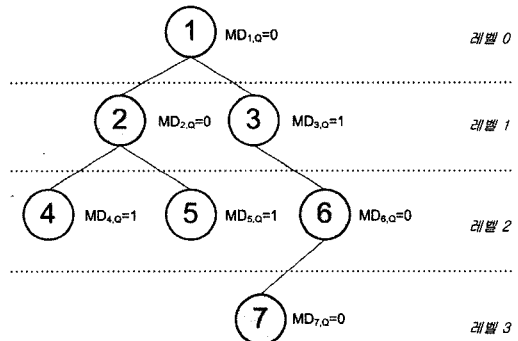


그림 4 질의 Q 에 대한 QSRT T_Q

MD 값에 기반하여 구성되는 T_Q 는 다음과 같은 특성을 가진다.

특성 1. T_Q 를 질의 Q 의 결과 데이터를 모으는 데 사용할 경우 루트 노드를 제외한 센서 노드 i 에서 부모 노드 p 로 전송된 질의 Q 의 결과 데이터는 $MD_{p,Q}$ 이후에 질의 Q 후보 노드를 만난다.

증명. 수학적 귀납법을 사용하여 쉽게 증명할 수 있으므로 생략한다. □

질의별 QSRT는 각 센서 노드에 라우팅 테이블의 형태로 저장된다. 각 센서 노드는 여러 질의에 대해 각 질의에 대한 결과 데이터를 전송할 부모 노드 정보를 저장하는 라우팅 테이블을 유지한다. 센서 노드 i 의 라우팅 테이블 RT_i 는 (질의 아이디, 부모 노드 아이디) 쌍들을 저장한다.

3.2 QSRT 구성 절차

본 절에서는 질의 메시지가 전달되는 과정을 통해 QSRT를 구성하는 절차를 설명한다. 각 질의에 대한 QSRT는 다음 두 단계를 통해 구성된다.

- 후보 부모 집합 결정 단계: 본 단계에서는 각 센서 노드의 후보 부모 집합이 결정된다.
- 질의 전달 및 QSRT 구성 단계: 본 단계에서는 질의 메시지가 전달되는 과정을 통해 QSRT가 구성된다. 이를 위해, 각 센서 노드는 자신의 MD 값을 계산하고 그것을 질의 메시지에 기록하여 이웃 노드들에게 전달한다. 또한 각 센서 노드는 후보 부모 노드들 중에서 가장 작은 MD 값을 가진 노드를 자신의 부모 노드로 선택한다.

3.2.1 후보 부모 집합 결정 단계

후보 부모 집합 결정 단계에서는 각 센서 노드의 후보 부모 집합이 결정된다. 본 단계는 후보 부모 집합 결정 메시지가 센서 네트워크 전체로 전달되는 과정을 통해 수행된다. 후보 부모 집합 결정 메시지의 포맷은 다음과 같다(후보 부모 집합 결정 단계의 수행 절차를 기술하는 데 관련 없는 필드들은 표시하지 않았음.)

<dest_id, src_id, level>

후보 부모 집합 결정 메시지의 dest_id와 src_id 필드는 각각 후보 부모 집합 결정 메시지를 받을 센서 노드의 아이디와 보내는 센서 노드의 아이디를 저장하고 level 필드는 후보 부모 집합 결정 메시지를 보내는 센서 노드의 레벨을 저장한다. 그림 5는 후보 부모 집합 결정 단계의 수행 절차를 기술한다.

그림 5의 후보 부모 집합 결정 단계의 수행 절차는 3번과 4번 단계에서 후보 부모 집합을 갱신하는 부분을 제외하고는 2.2 절에서 설명된 기존의 라우팅 트리 구성 절차와 거의 비슷하다. 그림 6은 후보 부모 집합 결정 단계가 수행되는 예를 보여 준다. 그림 6에서 화살표는 후보 부모 집합 결정 메시지가 전송되는 경로를 나타내고, 각 센서 노드의 오른쪽에는 후보 부모 집합이 표시되어 있다.

후보 부모 집합 결정 단계는 센서 노드의 고장이나 통신 품질의 변화에 따른 센서 네트워크의 구조 상의 변화에 대응하기 위해 주기적으로 수행되거나, [8]에서와 같이 센서 네트워크의 구조 상의 변화로 인한 데이

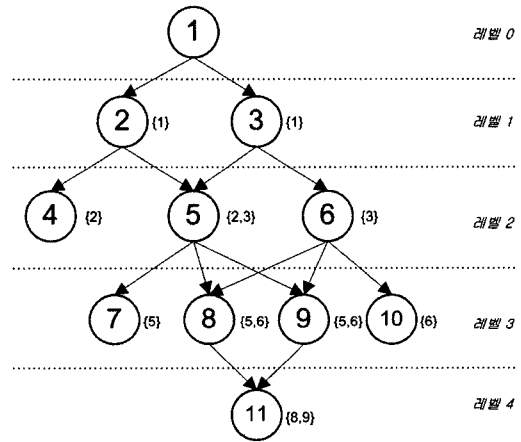


그림 6 후보 부모 집합 결정 단계의 수행 예

1. 베이스 스테이션은 후보 부모 집합 결정 단계를 수행하기 위해 후보 부모 집합 결정 메시지를 준비하여 루트 노드에게 전달한다.
2. 루트 노드는 후보 부모 집합 결정 메시지를 받으면 src_id와 level 필드를 각각 자신의 아이디와 자신의 레벨(루트 노드의 레벨은 0)로 설정하고 dest_id 필드를 방송 주소로 설정한 후 이 메시지를 이웃 노드들에게 방송한다.
3. 센서 노드 i가 후보 부모 집합 결정 메시지를 받았고 아직 레벨이 정해지지 않았다면, 후보 부모 집합 결정 메시지의 level 필드의 값에 1을 더한 값을 자신의 레벨 l_i 로 설정하고 (즉, $l_i \leftarrow level + 1$), src_id 필드의 값을 후보 부모 집합 CP_i 에 추가한다. 또한 센서 노드 i는 후보 부모 집합 결정 메시지의 src_id와 level 필드를 각각 자신의 아이디와 레벨 l_i 로 설정하고 이 메시지를 이웃 노드들에게 방송한다.
4. 센서 노드 i가 후보 부모 집합 결정 메시지를 받았고 자신의 레벨 값이 정해진 상태라면, 후보 부모 집합 결정 메시지의 level 필드가 자신의 레벨 값보다 하나 작은 값일 경우(즉, $level = l_i - 1$), src_id 필드의 값을 후보 부모 집합 CP_i 에 추가한다.
5. 3번과 4번 과정은 모든 센서 노드들의 레벨 값과 후보 부모 집합이 결정될 때까지 반복된다.

그림 5 후보 부모 집합 결정 단계의 수행 절차

타의 손실로 인해 수집된 질의의 결과 데이터의 개수가 사용자가 허용한 개수 이하로 떨어질 경우(즉, 에러율이 높을 경우) 수행될 수도 있다.

3.2.2 질의 전달 및 QSRT 구성 단계

후보 부모 집합 결정 단계를 통해 각 센서 노드의 후보 부모 집합이 결정된 상태에서, 질의 요청이 들어왔을 때 질의 전달 및 QSRT 구성 단계가 수행되어 해당 질의에 대한 QSRT가 구성된다. 질의 전달 및 QSRT 구성 단계에서는 질의 메시지가 센서 네트워크 전체로 전달되는 과정을 통해 각 센서 노드의 MD 값이 계산되어 질의 메시지에 기록되어 전달되고, 각 센서 노드는 후보 부모 노드들 중에서 가장 작은 MD 값을 가지는 노드를 부모 노드로 선택하여 라우팅 트리를 구성한다. 질의 메시지의 포맷은 다음과 같다(질의 전달 및 QSRT 구성 단계의 수행 절차를 기술하는 데 관련 없는 필드들은 표시하지 않았음.)

$\langle dest_id, src_id, md, query_info \rangle$

질의 메시지의 $dest_id$ 와 src_id 필드는 각각 질의 메시지를 받을 센서 노드의 아이디와 보내는 센서 노드의 아이디를 저장하고, md 필드는 질의 메시지를 보내는 센서 노드의 MD 값을, $query_info$ 필드는 질의의 내용을 저장한다. 그림 7은 질의 전달 및 QSRT 구성 단계의 수행 절차를 기술한다.

그림 7의 질의 전달 및 QSRT 구성 단계의 수행 절

차에서, 각 센서 노드는 3 번 단계를 통해 후보 부모 노드들의 MD 값을 갱신한다. 각 센서 노드는 모든 후보 부모 노드들의 MD 값을 갱신하고 나면 4 번 단계에서 부모 노드를 선택한다. 각 센서 노드는 후보 부모 노드들 중에서 가장 작은 MD 값을 가진 노드를 부모 노드로 선택한다. 4 번 단계에서 사용된 $MinDistId()$ 함수는 3.2 절에서 설명되었다. 각 센서 노드는 부모 노드를 선택한 후 5 번 단계에서 자신의 MD 값을 계산하여 질의 메시지를 통해 이웃 노드들에게 전달한다. 이렇게 질의 전달 및 QSRT 구성 단계가 수행되고 나면 질의 Q에 대한 QSRT가 구성된다.

그림 8은 그림 6의 예를 이어 어떤 질의 Q의 요청이 들어왔을 때 질의 전달 및 QSRT 구성 단계가 수행되는 예를 보여 준다. 그림 8에서 회색으로 표시된 노드들은 질의 Q 후보 노드이다. 그림 8의 (a)는 질의 메시지를 통해 MD 값이 전달되는 과정을 보여 준다. 여기서 화살표는 질의 메시지가 전달되는 경로를, 화살표 위에 표시된 숫자는 질의 메시지의 md 필드의 값을 나타낸다. 그림 8의 (b)는 각 센서 노드가 MD 값에 기반하여 부모 노드를 선택하여 구성된 QSRT의 모습을 보여 준다.

3.3 데이터 수집 단계에서 QSRT의 활용

질의 전달 및 QSRT 구성 단계에서 질의 메시지가 전달되는 과정을 통해 구성된 QSRT는 질의의 결과 데이터들을 베이스 스테이션으로 모으는 데이터 수집 단

1. 베이스 스테이션은 질의 Q를 요청 받으면 질의 Q에 대한 질의 메시지를 준비하여 루트 노드에게 전달한다.
2. 루트 노드는 베이스 스테이션으로부터 질의 Q에 대한 질의 메시지를 받으면 질의 메시지의 src_id 와 md 필드를 각각 자신의 아이디와 MD 값(루트 노드의 MD 값은 0)으로 설정하고 $dest_id$ 필드를 방송 주소로 설정한 후 이 메시지를 이웃 노드들에게 방송한다.
3. 후보 부모 노드 MD 값 갱신: 센서 노드 i 가 후보 부모 노드로부터 질의 Q에 대한 질의 메시지를 받았다면, 즉 $src_id \in CP_i$ 라면, (src_id, md) 를 $CP-MD_{i,Q}$ 에 추가한다.
4. 부모 노드 선택: 센서 노드 i 가 모든 후보 부모 노드들에 대한 MD 값을 갱신하였다면, 즉 $|CP-MD_{i,Q}| = |CP_i|$ 라면, 후보 부모 노드들 중에서 가장 작은 MD 값을 가진 노드를 질의 Q에 대한 부모 노드 $P_{i,Q}$ 로 선택한다. 즉, $P_{i,Q} \leftarrow MinDistId(CP-MD_{i,Q})$. 또한 $(Q, P_{i,Q})$ 를 라우팅 테이블 RT_i 에 추가한다.
5. MD 값 계산: 센서 노드 i 의 부모 노드 $P_{i,Q}$ 가 결정되면, 센서 노드 i 는 질의 Q에 대한 자신의 MD 값 $MD_{i,Q}$ 를 계산한다. 센서 노드 i 가 질의 Q 후보 노드라면, $MD_{i,Q}$ 는 0이다. 센서 노드 i 가 질의 Q 후보 노드가 아니라면, $MD_{i,Q}$ 는 선택된 부모 노드의 MD 값에 1을 더한 값이 된다. $MD_{i,Q}$ 가 결정되고 나면, 센서 노드 i 는 질의 메시지의 src_id 와 md 필드를 각각 자신의 아이디와 $MD_{i,Q}$ 로 설정하고 이 메시지를 이웃 노드들에게 방송한다.
6. 3 번에서 5 번까지의 과정은 모든 센서 노드들의 부모 노드가 결정될 때까지 반복된다.

그림 7 질의 전달 및 QSRT 구성 단계의 수행 절차

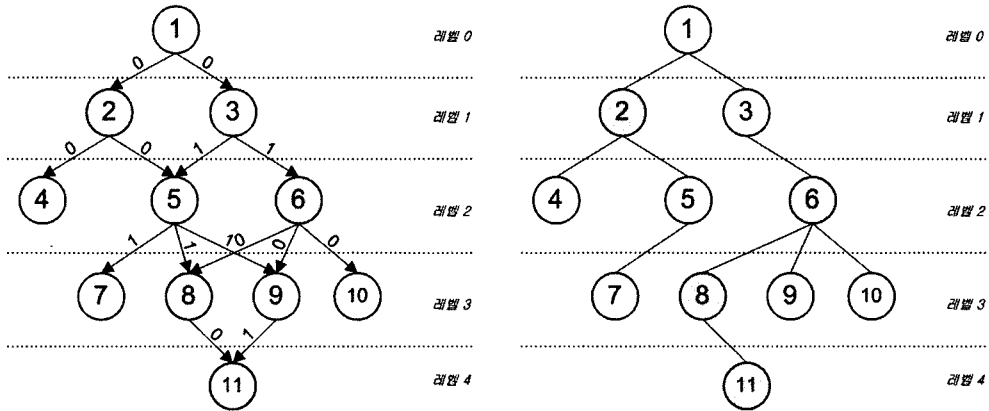


그림 8 질의 전달 및 QSRT 구성 단계의 수행 예

계에서 사용된다. 2.1 절에서 설명된 바와 같이 데이터 수집 단계에서는 메시지 전송 횟수를 줄이기 위해 네트워크내 프로세싱이 수행된다. 네트워크내 프로세싱은 라우팅 트리의 하단부에서 시작되어 루트 노드까지 레벨별로 수행된다. 센서 노드 i 는 자신이 데이터를 송신해야 할 시점이 되면, 자식 노드들로부터 받은 결과 데이터와 자신의 결과 데이터를 모아 부분 집계 또는 패킷 합병을 수행하고 그 결과를 하나의 메시지 내에 담는다. 센서 노드 i 는 이 메시지를 보낼 부모 노드를 라우팅 테이블 RT_i 에서 찾은 후 그 부모 노드를 타겟으로 하여 메시지를 전송한다.

예를 들어 센서 노드 i 의 라우팅 테이블 RT_i 가 표 1과 같다고 하자. 센서 노드 i 가 4 번 질의의 결과 데이터들에 대해 부분 집계 또는 패킷 합병을 수행한 결과를 메시지에 담아 전송하고자 할 때 라우팅 테이블 RT_i 로부터 4 번 질의에 대한 메시지는 5 번 노드로 보내야 한다는 것을 알 수 있다. 따라서 센서 노드 i 는 5 번 노드를 타겟으로 하여 메시지를 전송한다.

표 1 센서 노드 i 에서 라우팅 테이블 RT_i

질의 아이디	부모 노드 아이디
1	4
4	5
10	8

4. 실험

본 논문에서는 시뮬레이션 실험을 수행하여 데이터 수집 단계에서 QSRT와 FRT를 사용하였을 때의 성능 차이를 비교하였다. 성능 비교에 사용된 수치는 데이터 수집 단계에서 질의의 결과 데이터들을 한 번 모으는데 필요한 메시지 전송 횟수이다.

실험에서는 다음의 두 라우팅 트리의 성능을 비교 평가한다.

- FRT: 2.2 절에서 설명된 FRT 구성 절차에 따라 구성되는 라우팅 트리이다. FRT는 각 노드가 자신에게 가장 먼저 질의 메시지를 전달한 노드를 부모 노드로 선택하여 구성된다. 실험에서는 네트워크의 상황에 따라 이웃 노드들 중 가장 먼저 질의를 전달하는 노드가 달라질 수 있다는 사실을 반영하기 위해 센서 노드가 자신의 레벨 값 보다 하나 작은 레벨 값을 가지는 이웃 노드들 중에서 무작위로 한 노드를 부모 노드로 선택하게 하였다.
- QSRT: 본 논문에서 제안하는 QSRT이다. QSRT는 각 노드가 후보 부모 노드들 중에서 가장 작은 MD 값을 가진 노드를 자신의 부모 노드로 선택하여 구성된다.

4.1 실험 환경

본 실험에서는 두 라우팅 트리의 성능 비교를 위해 센서 네트워크 환경을 시뮬레이션하였다. 센서 네트워크는 가로 길이 w 미터, 세로 길이 h 미터인 사각형의 형태를 가진다. 센서 노드들은 이 센서 네트워크 상에 무작위로 배치된다. 센서 네트워크의 가로 길이를 w 미터, 세로 길이를 h 미터, 센서 노드의 통신 반경을 r 미터, 밀도를 d (통신 반경 내에 존재하는 센서 노드의 개수)이라고 할 때 센서 네트워크 상에 존재하는 노드의 개

수는 $N = \frac{d}{\pi \times r^2} \times w \times h$ 개이다. 센서 네트워크는 주어진 파라미터들로부터 센서 노드 개수인 N 값을 구하고 센서 네트워크 상에 N 개의 노드를 무작위로 배치하여 구성된다. 베이스 스테이션은 센서 네트워크의 한 가운데 위치한다.

두 라우팅 트리의 성능 비교 실험을 동일한 센서 네트워크 상에서 수행하기 위해 센서 네트워크 구성 정보

를 일단 파일에 기록한 후 실험 시에 파일로부터 동일한 센서 네트워크 구성 정보를 읽어 들여 실험을 수행하였다. 서로 다른 10 개의 센서 네트워크를 생성하고 각 실험을 이 10 개의 센서 네트워크에 대해 반복적으로 수행하였으며 실험의 결과 값으로 10 번의 반복 실험 수행에서 나온 결과 값들의 평균 값을 사용하였다.

두 라우팅 트리의 성능 비교의 수치로는 메시지 전송 횟수를 사용하였다. 각 실험에서는 주어진 센서 네트워크 환경에서 각 라우팅 트리를 사용할 때 데이터 수집 단계에서 질의 결과 데이터들을 한 번 모으는 데 필요한 메시지 전송 횟수를 세었다. 데이터 수집 단계에서는 네트워크내 프로세싱이 수행된다. 이 때 각 센서 노드는 결과 데이터들에 대해 부분 집계 또는 패킷 합병하고 그 결과를 하나의 메시지로 전송한다고 가정한다.

표 2는 실험에 사용된 파라미터들을 나타낸다. 통신 반경 파라미터는 센서 노드들의 통신 반경을 지정한다. 센서 노드들은 통신 반경 내의 이웃 노드들과 서로 통신 할 수 있으며 통신은 무손실(lossless) 통신을 가정한다. 밀도 파라미터는 센서 노드의 통신 반경 내에 몇 개의 센서 노드가 존재하는지 지정한다. 질의 선택도 파라미터는 전체 노드들 중에서 질의 결과 데이터를 생성하는 노드들이 몇 퍼센트가 있는지를 지정한다. 본 실험에서는 질의 Q 후보 노드들이 모두 결과 데이터를 생성한다고 가정한다. 센서 네트워크 크기 파라미터는 센서 네트워크의 가로·세로 크기를 나타낸다. 표 2에서 주어진 파라미터들의 기본 값을 사용하여 센서 네트워크를 구성하면 생성되는 라우팅 트리는 레벨이 10 정도가 된다.

우리는 앞서 설명된 실험 환경에서 다음의 세 가지 실험을 수행하였다.

- 질의 선택도 실험: 질의 선택도의 변화에 따른 두 라우팅 트리의 성능 변화를 알아 보는 실험이다. 본 실험에서는 두 라우팅 트리의 성능에 질의 선택도가 미치는 영향을 알아 본다.
- 네트워크 크기 실험: 센서 네트워크의 크기 변화에 따른 두 라우팅 트리의 성능 변화를 알아 본다. 좁은 지역에 센서 노드들이 배치된 경우(빌딩 내)에서부터, 넓은 지역에 센서 노드들이 배치된 경우(동물들의 서

식지)까지, 센서 네트워크의 크기를 변화시키며 두 라우팅 트리의 성능 변화를 알아 본다.

- 밀도 실험: 밀도 변화에 따른 두 라우팅 트리의 성능 변화를 알아 보는 실험이다. 센서 노드들이 드문드문 배치된 센서 네트워크에서부터 조밀하게 배치된 센서 네트워크까지, 밀도 값을 변화시켜 가면서 두 라우팅 트리의 성능 변화를 알아 본다.

4.2 실험 결과 및 분석

4.2.1 질의 선택도 실험

질의 선택도 실험에서는 질의 선택도를 1%부터 100%까지 변화시켜가며 두 라우팅 트리의 성능을 비교하였다. 그림 9는 질의 선택도 실험의 결과를 보여 준다. 질의 선택도가 낮을 때(그래프의 왼쪽 편)는 두 라우팅 트리의 성능 차이가 거의 나지 않는다. 그 이유는 질의 선택도가 낮을 때는 질의 결과 데이터를 전송하는 노드 수가 적어 전송되는 메시지 수가 적기 때문에 부분 집계 또는 패킷 합병 자체가 많이 일어나지 않기 때문이다. 그러나 질의 선택도가 커질수록 QSRT는 부분 집계 및 패킷 합병을 최소화하므로 FRT에 비해 성능이 더 좋아진다. 질의 선택도가 100% 쪽으로 가까워지면 두 라우팅 트리의 성능 차이가 다시 줄어들기 시작한다. 그 이유는 큰 질의 선택도에서는 대부분의 노드가 질의 결과 데이터를 생성하므로 QSRT와 FRT 간에 부분 집계 또는 패킷 합병의 수행 횟수의 차이가 크지 않기 때문이다. 질의 선택도 실험의 결과 그래프는 QSRT가 FRT에 비해 최고 18% 정도까지 메시지 전송 횟수를 줄일 수 있음을 보여 준다.

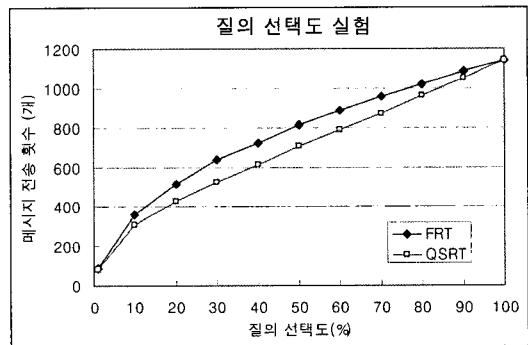


그림 9 질의 선택도 실험

표 2 실험에 사용된 파라미터들

파라미터	값	기본 값
통신 반경(미터)	30 (대략 100 feet[13])	
밀도	6~20	9
질의 선택도 (%)	1~100	25
센서 네트워크 크기(미터×미터)	120×120~1080×1080	600×600

질의 선택도 실험의 결과는 질의 선택도가 1%에 가깝거나 100%에 가까운 경우를 제외하고는 QSRT의 성능이 FRT에 비해 우수함을 보여 준다. 일반적으로 사용자는 질의를 통해 관심 있는 센서 노드들의 데이터들만을 선택적으로 알아 보고자 할 것이므로 질의 선택도는 양 극단의 값을 가지지는 않을 것이다. 따라서 일반적인 질의 처리에 QSRT가 기존의 라우팅 트리보다 효과적으로 사용될 수 있다.

4.2.2 네트워크 크기 실험

네트워크 크기 실험에서는 센서 네트워크의 크기를 120×120부터 1080×1080까지 변화시켜가며 두 라우팅 트리의 성능을 비교하였다. 네트워크 크기 실험은 센서 노드들이 배치된 지역의 크기가 변함에 따라 두 라우팅 트리의 성능이 어떻게 변하는지를 알아 본다. 그림 10은 네트워크 크기 실험의 결과를 보여 준다. 그림 10의 그래프에서 가로 축의 수치는 센서 네트워크의 가로 길이를 나타내며, 센서 네트워크의 세로 길이는 가로 길이와 동일하다. 네트워크 크기 실험은 다른 두 실험 파라미터(질의 선택도, 밀도)를 고정된 상태에서 실행하였다. 따라서 센서 네트워크의 크기가 커질수록 센서 노드의 개수가 많아지고 라우팅 트리의 레벨이 커진다.

네트워크 실험 결과 그래프에서 네트워크 크기가 증가함에 따라 두 라우팅 트리의 성능 차이는 점점 증가하다가 네트워크 크기가 480 이상이 될 때부터는 그 차이가 더 증가하지 않고 16%에서 18%에 머무른다(네트워크 크기 1080 이상의 실험에서도 마찬가지임). 네트워크 크기 480 이하에서 네트워크 크기가 증가함에 따라 두 라우팅 트리의 성능 차이가 커지는 이유는 QSRT에서는 결과 데이터들이 곧 서로 만나 무임승차의 효과를 얻는 반면, FRT에서는 결과 데이터들이 다른 결과 데이터를 만나지 못하고 전달되어 그만큼의 전송 비용이 더 들기 때문이다. FRT에서 결과 데이터들이 다른 결과 데이터를 만나지 못하고 전달될 수 있는 최대 홉 수는 라우팅 트리의 레벨만큼이다. 따라서 네트워크 크기가 증가하여

라우팅 트리의 레벨이 커지면 FRT의 전송 비용이 더욱 증가하여 두 라우팅 트리의 성능 차이가 커지는 것이다. 그러나 네트워크 크기가 480 이상이 되면 두 라우팅 트리의 성능 차이는 더 이상 증가하지 않는다. 그 이유는 라우팅 트리의 레벨이 어느 이상이 되면 FRT에서도 결과 데이터들이 라우팅 트리의 레벨만큼 전송되기 전에 다른 결과 데이터를 만나기 때문인 것으로 보인다.

이러한 네트워크 실험의 결과는 확률적인 방법으로 설명할 수 있다. 네트워크 실험에서 질의 선택도와 밀도가 고정되어 있고 FRT를 생성할 때 각 노드는 후보 부모 노드들 중 무작위로 한 노드를 부모 노드로 선택하므로 선택된 부모 노드가 결과 데이터를 생성하는 노드일 확률은 모든 노드에서 동일하다. 이 확률을 p라고 하자. 그러면 한 결과 데이터가 FRT를 따라 전송될 때 다음 노드에서 결과 데이터를 생성하는 노드를 만날 확률 역시 p이다. 한 결과 데이터가 결과 데이터를 생성하는 노드까지 도달하는 데 걸리는 홉 수를 확률 변수 X로 두면, X는 기하분포(geometric distribution)를 따르고 평균 E(X)는 1/p이다. 즉, FRT에서 한 결과 데이터가 결과 데이터를 생성하는 노드까지 도달하는 데 걸리는 평균 홉 수는 라우팅 트리의 레벨과 무관하게 1/p라는 고정값을 가진다. 네트워크 크기가 480 이하일 때는 FRT의 레벨이 1/p보다 작아 한 결과 데이터가 다른 결과 데이터를 만나지 못하고 진행할 수 있는 홉 수는 최대 FRT의 레벨만큼이다. 그러므로 라우팅 트리의 레벨이 커질수록 FRT에서의 전송 비용이 증가한다. 그러나 네트워크 크기 480 이상일 때는 라우팅 트리의 레벨이 1/p보다 커져 FRT에서도 결과 데이터들이 라우팅 트리의 레벨만큼 진행하기 전인 1/p 홉 수 이후에 다른 결과 데이터를 만나기 때문에 네트워크 크기 증가에 따른 더 이상의 전송 비용의 증가는 없으므로 두 라우팅 트리의 성능 차이는 더 이상 증가하지 않는다.

네트워크 실험 결과 그래프는 QSRT가 FRT에 비해 최고 18% 정도까지 메시지 전송 횟수를 줄일 수 있음을 보여 준다. 네트워크 실험 결과를 통해 센서 노드들이 한 건물 내에 설치되어 빈방 여부를 알아 보는 센서 네트워크 응용에서부터 넓은 지역에 배치되어 동물들의 서식지를 관찰하는 응용까지 다양한 환경에서의 질의 처리에 QSRT가 기존의 라우팅 트리보다 효과적으로 사용될 수 있다는 것을 알 수 있다.

4.2.3 밀도 실험

밀도 실험에서는 밀도 값을 6에서부터 20까지 증가시켜가며 두 라우팅 트리의 성능을 비교하였다. 그림 11은 밀도 실험의 결과를 보여 준다. 밀도가 작을 때는 센서 네트워크 상에 존재하는 센서 노드들의 수가 적고 질의에 참여하는 노드의 수가 많지 않아 부분 집계 및 패킷

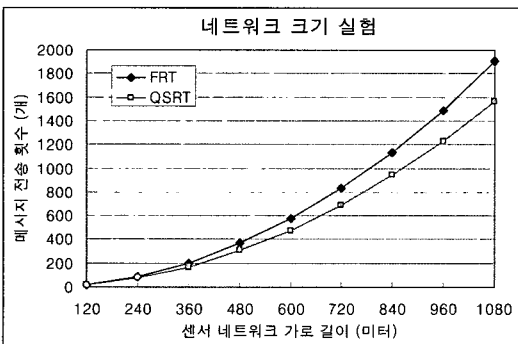


그림 10 네트워크 크기 실험

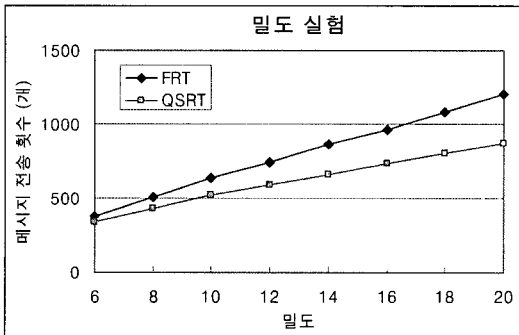


그림 11 밀도 실험

합병이 많이 일어나지 않으므로 두 라우팅 트리의 성능이 비슷하다. 그러나, 밀도가 커져 센서 노드들의 수가 증가함에 따라 QSRT가 FRT보다 더 좋은 성능을 보인다.

그림 11에서 두 라우팅 트리에 대한 메시지 전송 횟수의 증가세는 기울기가 일정한 선형적인 모습을 띈다. 두 그래프 간의 기울기의 차이를 통해 알 수 있듯이, 밀도가 커짐에 따라 QSRT와 FRT의 성능 차이는 더욱 커진다. 그림 11의 밀도 실험 결과 그래프는 밀도 값이 20일 때 QSRT가 FRT에 메시지 전송 횟수를 28% 가까이 줄일 수 있음을 보여 주며, 추가적인 실험을 통해 밀도가 20 이상인 경우 둘 간의 성능 차이가 더욱 커짐을 확인하였다.

밀도 실험의 결과는 QSRT가 센서 노드들이 드문드문 배치된 센서 네트워크에서만 뿐만 아니라 센서 노드들이 조밀하게 배치된 센서 네트워크까지 다양한 환경에서의 질의 처리에 좋은 성능을 낸다는 것을 보여 준다.

5. 결론

본 논문에서는 질의별로 구성되는 라우팅 트리인 QSRT를 제안하였다. QSRT는 질의의 결과 데이터들이 베이스 스테이션으로 전송되는 도중 가능한 한 서로 빨리 만나게 만들어, 부분 집계 및 패킷 합병의 수행 시점을 앞당기고 그 기회를 최대화시킴으로써 메시지 전송 비용을 줄인다. 본 논문에서는 시뮬레이션 실험을 통해 QSRT의 우수성을 보였다. 질의 선택도 실험 결과는 센서 노드들 중 일부 노드들의 정보를 알아 보고자 하는 일반적인 형태의 질의 처리에 QSRT가 기존의 라우팅 트리보다 특히 효과적으로 사용될 수 있음을 보여 준다. 또한 네트워크 실험 결과는 작은 규모의 센서 네트워크(빌딩 내)에서뿐만 아니라 큰 규모의 센서 네트워크(동물의 서식지)에서의 질의 처리에도 QSRT가 기존의 라우팅 트리보다 효과적으로 사용될 수 있음을 보여 준다. 마지막으로 밀도 실험은 센서 노드들이 드문드문 배치

된 센서 네트워크에서의 질의 처리에 QSRT가 기존의 라우팅 트리에 비해 효과적으로 사용될 수 있을뿐만 아니라 센서 노드들이 조밀하게 배치된 센서 네트워크에서의 질의 처리에는 더욱 효과적으로 사용될 수 있음을 보여 준다. QSRT를 데이터 수집에 사용할 경우 기존의 라우팅 트리에 비해 메시지 전송 횟수를 18%까지 줄일 수 있고 센서 노드들이 조밀하게 배치된 센서 네트워크에서는 28% 이상도 줄일 수 있다.

참고 문헌

- [1] J. Hill and D. Culler, "Mica: a wireless platform for deeply embedded networks," *IEEE Micro* 22(6):12-24, 2002.
- [2] A. Cerpa, J. Elson, D. Estrin, L. Girod, M. Hamilton, and J. Zhao, "Habitat monitoring: Application driver for wireless communications technology," In *ACM SIGCOMM Workshop on Data Communications in Latin America and the Caribbean*, 2001.
- [3] P. Juang, H. Oki, Y. Wang, M. Martonosi, L. S. Peh, and D. Rubenstein, "Energe-efficient computing for wildlife tracking: design tradeoffs and early experiences with zebnet," In *Proc. of ACM ASPLOS Conf.*, 2002.
- [4] D. Pescovitz. *Smart Buildings Admit Their Faults. Lab Notes*, <http://www.coe.berkeley.edu/labnotes/1101smartbuildings.html>.
- [5] S. R. Madden, M. J. Franklin, and J. M. Hellerstein, "TinyDB: an acquisitional query processing system for sensor networks," *ACM Trans. on Database Systems*, Vol. 30, No. 1, 122-173, 2005.
- [6] V. Shnayder, M. Hempstead, B. Chen, G. W. Allen, and M. Welsh, "Simulating the Power Consumption of Large-Scale Sensor Network Applications," In *Sensys*, 2004.
- [7] S. Madden, M. Franklin, J. Hellerstein, and W. Hong, "TAG: a tiny aggregation service for ad-hoc sensor networks," In *Proc. of OSDI*, 2002.
- [8] Y. Yao and J. Gehrke, "Query processing for sensor networks," In *Proc. of CIDR Conf.*, 2003.
- [9] A. Woo, T. Tong, and D. Culler, "Taming the underlying challenges or reliable multihop routing in sensor networks," In *SenSys*, 2003.
- [10] A. Woo and D. Culler, "A transmission control scheme for media access in sensor networks," In *Proc. of ACM MobiCom Conf.*, 2001.
- [11] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," In *Proc. of ACM MobiCom Conf.*, 2000.
- [12] M. A. Sharaf, J. Beaver, A. Labrinidis, and P.K. Chrysanthis, "Balancing energy efficiency and

quality of aggregate data in sensor networks," The VLDB Journal, Vol. 13, No. 4, 384-403, 2004.

- [13] D. J. Abadi, S. Madden, and W. Lindner, "REED: Robust, Efficient Filtering and Event Detection in Sensor Networks," In Proc. of VLDB Conf., 2005.



송 인 철

2004년 아주대학교 정보 및 컴퓨터공학부 학사. 2004년~현재 한국과학기술원 전자전산학과 석사과정. 관심분야는 데이터베이스, 센서 네트워크, 스트림 데이터베이스



노 요 한

2002년 경북대학교 컴퓨터공학과 학사 2002년~2004년 한국과학기술원 전자전산학과 석사. 2004년~현재 한국과학기술원 전자전산학과 박사과정. 관심분야는 데이터베이스, 센서 네트워크, 시멘틱 웹, Pervasive Computing



현 동 준

1999년 한국과학기술원 전산학과 학사 2000년 한국과학기술원 전자전산학과 석사. 2000년~현재 한국과학기술원 전자전산학과 박사과정. 관심분야는 데이터베이스, 데이터마이닝, 센서 네트워크

김 명 호

정보과학회논문지 : 데이터베이스
제 33 권 제 1 호 참조