

■ 2009년도 학생논문 경진대회 수상작

PBFiltering: 무선 센서 네트워크에서 우선순위 기반 상향식 필터링을 이용한 에너지 효율적인 스카이라인 질의 처리 기법

(PBFiltering: An Energy Efficient Skyline Query Processing Method using Priority-based Bottom-up Filtering in Wireless Sensor Networks)

성 동 옥 [†] 박 준 호 [†] 김 학 신 [†] 박 형 순 [†]
 (Dong Ook Seong) (Jun Ho Park) (Hak Sin Kim) (Hyoung Soon Park)

노 규 종 [†] 여 명 호 [†] 유 재 수 ^{**}
 (Kyu Jong Roh) (Myung Ho Yeo) (Jae Soo Yoo)

요 약 센서 네트워크에서 병합 질의를 효율적으로 처리하기 위한 다양한 인-네트워크 질의 처리 기법들이 제안되었다. 스카이라인 질의는 일반적인 병합 질의와 달리 다차원 데이터에 대한 총괄적인 비교를 요구한다. 따라서, 부분적인 데이터만을 이용할 수 있는 인-네트워크 방식으로 처리하기 어렵다. 스카이라인 질의를 에너지 효율적으로 처리하기 위해서 불필요한 데이터의 전송을 제거하는 것이 중요하다. 기존에 제안된 스카이라인 처리 기법은 전체 네트워크에 필터를 배포함으로써 불필요한 데이터 전송을 차단한다. 하지만 많은 긍정 오류(False Positive) 발생에 따른 불필요한 데이터 전송과 필터 배포시 발생하는 에너지 소모로 인해 네트워크의 수명이 단축된다. 본 논문에서는 필터 배포에 따른 에너지 소모를 줄이기 위한 방법으로 상향식 필터 설정을 통한 스카이라인 질의 처리 기법과 필터링 성능을 향상시키는 PBFiltering 기법을 제안한다. 제안하는 기법은 필터를 미리 배포하지 않고 하위 노드로부터 기지국으로 데이터를 수집하는 과정에서 스카이라인 필터 테이블(SFT)을 만들고 필터링을 수행한다. 그리고 여기서 제안하는 우선순위 맵을 이용한 선 필터링(Pre-filtering) 기법을 통해 필터링 효율을 증가시킨다. 제안하는 알고리즘의 우수성을 보이기 위해 기존에 제안된 MFTAC 기법과의 시뮬레이션을 통해 비교 평가하였다. 그 결과 기존 기법에 비해 다수의 긍정 오류의 발생을 감소시키고, 네트워크 수명이 연장됨을 보였다.

키워드 : 스카이라인 질의, 센서 네트워크, 필터링, 상향식 필터, 선-필터링

Abstract In sensor networks, many methods have been proposed to process in-network aggregation effectively. Unlike general aggregation queries, skyline query processing compares multi-dimensional

* 이 논문은 2009년도 교육과학기술부로부터 지원(지역거점연구단육성사업/충북 BIT연구중심대학육성사업단)과 한국과학재단(교육과학기술부)의 지원을 받아 수행된 연구임 (No. 20090080279)

논문접수 : 2009년 5월 13일
 심사완료 : 2009년 9월 21일

[†] 학생회원 : 충북대학교 정보통신공학과
 seong.do@gmail.com
 jhpark@netdb.cbnu.ac.kr
 hskim@netdb.cbnu.ac.kr
 soon@netdb.cbnu.ac.kr
 kjrho@netdb.cbnu.ac.kr
 mhyeo@netdb.cbnu.ac.kr

^{**} 종신회원 : 충북대학교 정보통신공학과 교수
 yjs@chungbuk.ac.kr
 (Corresponding author임)

Copyright©2009 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 제36권 제6호(2009.12)

data for the result. Therefore, it is very difficult to process the skyline queries in sensor networks. It is important to filter unnecessary data for energy-efficient skyline query processing. Existing approach like MFTAC restricts unnecessary data transitions by deploying filters to whole sensors. However, network lifetime is reduced by energy consumption for many false positive data and filters transmission. In this paper, we propose a bottom up filtering-based skyline query processing algorithm of in-network for reducing energy consumption by filters transmission and a PBFiltering technique for improving performance of filtering. The proposed algorithm creates the skyline filter table (SFT) in the data gathering process which sends from sensor nodes to the base station and filters out unnecessary transmissions using it. The experimental results show that our algorithm reduces false positives and improves the network lifetime over the existing method.

Key words : skyline query, sensor network, filtering, bottom-up filter, pre-filtering

1. 서론

최근에 신호 처리 기술, 소형 전자 장치 개발 기술, 무선 통신 기술이 발전함에 따라 센서 네트워크의 많은 응용들이 가능하게 되었다. 연구 목적의 환경 감시 혹은 군수 물자의 흐름 파악, 전장 상황 파악과 같은 다양한 분야에서 활용된다[1,2]. 일반적으로 센서 노드는 온도, 습도, 조도, 영상 등의 다양한 환경 정보를 수집할 수 있는 센서 소자들과 단거리 RF통신이 가능한 통신 모듈을 내장하고 있다. 이를 통해 센서들 간의 네트워크를 구축하여 관심 지역에 대한 다양한 데이터를 수집할 수 있다. 센서는 일반적으로 내장 배터리를 이용해 동작한다. 만약 배터리를 모두 소모하게 되면 해당 센서 노드는 더 이상 동작할 수 없으며, 구성된 네트워크는 활용 불가능해진다. 이로 인해 배터리 전력을 얼마나 오래 유지하느냐에 따라 센서 네트워크의 수명이 결정된다.

현재 센서 노드의 배터리 전력을 절약하기 위한 다양한 기법들이 연구되고 있다. 그 중 센서에서 수집되는 데이터에 대한 질의를 효율적으로 처리하기 위하여 인-네트워크 질의 처리 기법들이 제안되었다. 기존 방식의 경우 기지국의 로컬 스토리지에서 질의를 처리하기 위해 수집되는 모든 데이터를 기지국으로 전송한다. 이 방식은 큰 데이터 전송 비용이 소모되어 네트워크의 수명을 줄인다. 하지만 인-네트워크 질의 처리 기법은 데이터 라우팅 도중에 네트워크 내에서 데이터를 병합(aggregation)하여 전송량을 줄임으로써 전송 에너지를 절약한다. 보통 데이터 병합 기법들은 일차원의 속성값에 대한 처리만을 지원한다. 하지만 환경 데이터의 특성상 여러 속성값을 조합해야 가치 있는 데이터를 얻을 수 있는 경우가 많다. 예를 들어, “대기 오염이 심한 지역을 알려달라.”와 같은 질의는 대기오염 정도를 판단하기 위해 CO₂, CO, SO₂ 등의 다양한 값들을 고려해야 한다. 이러한 경우 스카이라인 질의를 이용해 처리할 수 있다. 일반적으로 센서 네트워크에서 환경을 대상으로 하는 응용들은 다중 속성에 대한 질의 처리가 많이

발생하므로 센서 네트워크에서의 효율적인 스카이라인 질의 처리 기법이 요구된다.

스카이라인 질의 처리에 대한 기존 연구들은 모든 데이터가 동일한 저장공간에 위치한 상황만을 고려하였다. 하지만 센서 네트워크 환경에서 인-네트워크 방식으로 질의 처리하는 상황에서는 질의 대상이 되는 데이터들이 각 센서들에 분산되어있다. 따라서 기존에 제안된 스카이라인 질의 처리 기법을 센서 네트워크 환경에 적용하기 힘들다. 본 논문에서는 효율적인 인-네트워크 스카이라인 질의 처리를 위해 스카이라인 질의 결과에 만족하지 않는 데이터를 필터링하는 기법에 대하여 연구를 진행하였다.

본 논문은 다음과 같은 순서로 기술하였다. 2장에서는 기존에 제안된 스카이라인 질의 처리 기법과 문제점을 살펴보고, 3장에서는 제안하는 스카이라인 질의 처리 기법을 설명한다. 4장에서는 제안하는 기법의 우수성을 보이기 위해 기존에 제안된 질의 처리 기법과 시뮬레이션을 통해 비교 평가한다. 마지막으로 5장에서는 결론에 대하여 기술한다.

2. 관련연구

TAG[3]와 같은 인-네트워크 병합 질의 처리 기법은 센서 네트워크 내부에서 병합된 결과만을 기지국으로 전송하기 때문에 에너지 소모를 감소시킨다. 특히, SUM, MIN, MAX와 같은 병합 질의의 경우, 모든 데이터를 수집하지 않고 병합된 결과만을 수집하기 때문에 에너지 소모를 줄이는데 더 효과적이다[4]. 하지만 다중 속성을 이용하는 스카이라인 질의의 경우, 모든 센서 데이터를 대상으로 모든 차원의 속성에 의해 배제되지 않는 데이터만 추출해야 하므로 부분적인 결과를 통해 전체 결과를 도출하는 것은 어려운 일이다[5]. 따라서, 결과에서 불필요한 데이터를 미리 제거하여 일부의 주요 데이터들만을 통해 질의처리가 이루어지도록 하는 것이 중요하다.

Block Nested Loop(BNL)[6]는 순차적으로 전체 데이터를 스캔하고, 메모리에 저장된 후보 데이터와 비교한다. Sorted Filter Skyline(SFS)[7]은 연산을 통해 모든 속성을 하나의 단일 값으로 만들고 정렬을 통해 비교 대상의 수를 줄인다. Divide and Conquer(D&C)[6]는 데이터의 집합을 많은 영역으로 나누고, 각 영역에서 스카이라인을 계산한다. 그 다음, 각 영역의 스카이라인으로부터 최종 스카이라인을 결정한다. 하지만, 이 방법들은 중앙처리를 통해 수행되므로 센서 네트워크와 같은 분산 환경에 적합하지 않다.

[8]과 같은 분산 환경을 고려한 스카이라인 질의 처리 기법은 웹 정보 시스템을 대상으로 하고 있기 때문에 쉽게 센서 네트워크에 적용할 수 없다. [8]은 모바일 애드혹 네트워크의 경량 디바이스를 위한 분산 스카이라인 질의 처리 기법이다. 질의 결과에서 불필요한 데이터를 제외하기 위한 필터링 기법을 제안한다. 하지만, 필터링 정보는 상위 노드에 의해 결정되기 때문에 다른 센서 노드의 데이터는 필터 설정에 영향을 끼치지 않는다. [5]는 [8]보다 향상된 필터링 기법을 제안한다. 그림 1은 [5]의 MFTAC 예를 나타낸다. 먼저 센싱한 데이터를 이용하여 가장 많은 데이터를 필터링할 수 있는 MFT(Min-Score Filtering Tuple)를 계산하고, MFT를 센서 노드에 배포함으로써 불필요한 데이터의 전송을 줄인다. 또한, MFT 배포를 위한 비용을 분산하기 위해서 HEED[9] 클러스터링 방법을 이용하여 클러스터 헤드가 MFT를 계산하고 배포하도록 한다. 하지만 클러스터 헤드에 의해 MFT 배포 비용이 분산된다 하더라도 연속적인 스카이라인 질의를 처리하기 위해서는 매 주기마다 MFT를 클러스터 멤버로 배포해야 한다. 결국, 전체 센서 노드가 MFT의 송/수신에 참여하게 된다. 또한, 클러스터나 라우팅 트리에 의한 MFT의 계산

은 지역적인 데이터를 반영하고 있기 때문에 많은 불필요한 데이터들이 기지국으로 전송된다. 예를 들면, $\{S_2, S_4, S_5\}$ 로 구성된 클러스터는 스카이라인 질의의 결과로 $\{S_2, S_4\}$ 를 보고한다. 하지만, 기지국에 계산된 최종 스카이라인 결과에서 $\{S_2, S_4\}$ 는 S_3 에 의해 제거된다. 결국 $\{S_2, S_4\}$ 는 불필요한 데이터이다. 효과적인 스카이라인 질의 처리를 위해서 필터의 배포 비용을 줄이고, 지역적인 필터링을 통해 긍정 오류를 효과적으로 줄이기 위한 연구가 필요하다.

3. 제안하는 스카이라인 질의 처리 기법

본 장에서는 필터 배포 비용과 불필요한 데이터 전송을 효과적으로 줄이는 우선순위 필터링을 스카이라인 질의 처리 기법을 제안한다. 기존에 제안된 질의 처리 기법은 각 센서 노드에 필터를 설정하기 위해서 라우팅 트리의 최상위 노드로부터 말단 노드로 필터 값을 배포한다. 전체 센서 노드에 필터를 설정하기 위해서는 네트워크의 모든 노드가 적어도 한번 필터 값을 송/수신해야 한다. 또한 기존기법에서 사용하는 필터인 MFT는 전송 경로상의 부분 데이터만을 반영하고 있기 때문에 필터의 성능이 우수하지 않다. 이로 인해 많은 긍정 오류 데이터가 발생되고, 이것은 네트워크 수명을 단축시키는 주된 원인이 된다. 본 논문에서는 필터를 구성하는 데이터와 기지국으로 전송되는 데이터가 동일하다는 점을 착안하여 데이터를 수집하는 과정에서 필터를 설정하는 상향식 필터 설정 기법을 제안한다. 그리고, 상향식 필터의 필터링 성능을 높이기 위한 다양한 기법을 제안한다. 먼저 상향식 필터 설정 기법을 기술하고, 상향식 필터 설정기법 적용에 따른 문제점과 이를 해결하기 위한 우선순위 기반 데이터 수집 기법을 기술한다.

3.1 상향식 필터 설정 기법(BUF: Bottom-Up Filtering)

제안하는 기법은 필터를 배포하기 위해 필요한 추가적인 통신비용을 줄이기 위해 센싱 데이터를 기지국으로 전송하는 과정에서 필터를 설정한다. 최초 각 센서 노드들은 자신의 수집 데이터를 필터로 설정한다. 그리고 라우팅 경로 상에서 가장 단말노드를 시작으로 상위 노드로 센싱 데이터를 전송한다. 센싱 데이터를 수신한 상위노드는 자신의 필터값과 비교하여 수신된 데이터가 배제되는지 판단한다. 만약 필터에 의해 배제되면 해당 데이터를 제거하고, 배제되지 않는다면 자신의 필터에 업데이트하고, 해당 데이터를 다시 상위 노드로 전송한다. 이러한 과정을 모든 데이터가 기지국에 도달할 때까지 반복한다. 최종적으로 기지국은 긍정 오류가 어느 정도 제거된 데이터만을 수신한다. 그림 2를 보면 단말 노드인 A, B노드의 센싱 데이터 $\{30, 40\}$, $\{10, 20\}$ 와 C노드의 필터값 $\{20, 40\}$ 은 C노드에서 상호 배제유무를

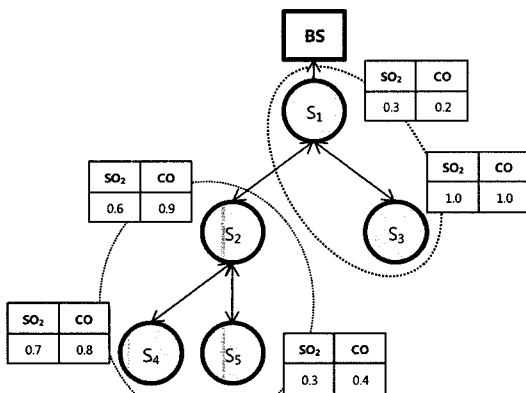


그림 1 MFTAC의 예

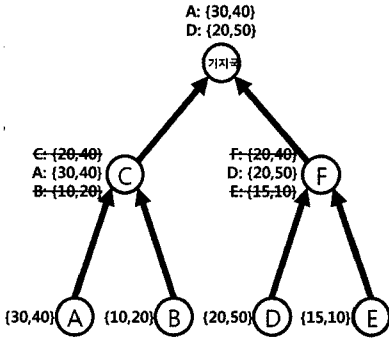


그림 2 상향식 필터 설정 과정

판단할 수 있다. 그 결과 B, C 노드의 데이터는 배제되고, A 노드의 센싱 데이터만 최종적으로 기지국으로 전송된다.

3.1.1 수집 우선순위를 이용한 성능향상

앞서 제안한 상향식 필터 설정 기법은 한가지 문제점을 내포하고 있다. 각 센서 노드에서 동일한 데이터가 수집되더라도 전송되는 순서에 따라 데이터 전송횟수가 틀려진다. 그림 3의 예를 보면, 센서 노드 A, B, C는 각각 {10, 10}, {5, 5}, {1, 1}의 데이터를 수집하였다. 여기서 데이터를 전송 순서가 C→B→A일 경우, A노드는 1번, B노드는 2번 그리고 C노드는 1번의 전송이 발생한다. 하지만 전송 순서가 A→B→C일 경우, A노드는 1번, B노드는 1번 그리고 C노드는 1번의 전송이 발생한다. 그 이유는 전송 순서에 따라서 B노드의 최종 필터 값이 설정 되는 시점이 달라지기 때문이다. A노드의 (10, 10)의 데이터가 가장 먼저 전송될 경우 B노드의 필터값으로 (10, 10)이 설정된다. 이로 인해 B, C노드의 (5, 5), {1, 1}은 기지국으로 전송되기 전에 B노드에서 배제 시킬 수 있다. 즉, 스카이라인의 결과에 만족할 가능성이 높은 순서대로 수집순서를 결정하면 필터의 성능을 최적화시킬 수 있다.

본 논문에서는 단조 점수(Monotone Score)를 이용하여 각 센서 노드의 수집 데이터를 바탕으로 전송 우선순위를 결정하는 기법을 제안한다. 단조 점수는 식 (1)을 이용하여 다차원 값을 단일 스코어로 변환한 값이다. m 은 차원의 수를 말하고, p_j 는 각 차원 별 속성값을 의미한다. 변환 결과, 그림 4와 같이 스카이라인 결과에

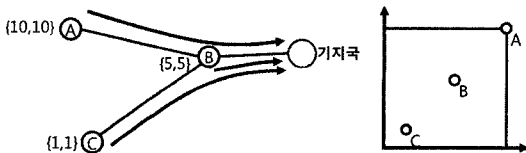


그림 3 수집 순서에 따른 성능 변화

만족할 가능성이 높을수록 낮은 값을 보인다. 따라서 단조 점수에 따라 전송 시각을 결정하게 되면 스카이라인 결과에 포함될 가능성이 높을수록 먼저 전송하여 필터의 성능을 높일 수 있다. 식 (2)를 이용해 전송 시간을 결정할 수 있다. R_c 는 현재 라운드 수를 의미하고, T 는 1라운드의 시간 범위를 나타낸다. S_{max} 는 단조 점수의 최대값을 나타내고, S_{own} 은 현재 값의 단조 점수 값을 나타낸다.

$$Monotone\ Score = \sum_{j=1}^m \frac{1}{p_j} \quad (1)$$

$$Transmit\ Time = ((R_c - 1) \cdot T) + \left(\frac{T \cdot S_{own}}{S_{max}} \right) \quad (2)$$

Sensor	Real Data (co., so.)	Monotone Score
S ₁	(0.7, 0.5)	3.43
S ₂	(0.9, 0.2)	6.11
S ₃	(0.1, 0.6)	11.66
S ₄	(0.5, 0.4)	4.5
S ₅	(1.0, 0.3)	4.33
S ₆	(0.2, 0.8)	6.25
S ₇	(0.3, 1.0)	4.33

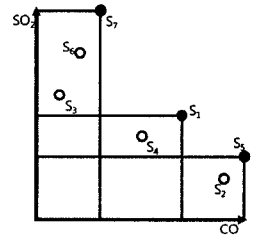
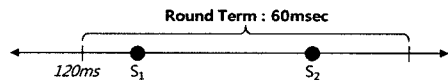


그림 4 단조 점수에 따른 스카이라인 결과

그림 5는 식 (2)를 이용해 수집된 데이터의 전송 시간을 계산한 것이다. 현재 라운드는 3이며, 1라운드 별 시간 간격은 60ms이다. 계산 결과 단조 점수가 낮은 S₁의 전송시간이 S₂보다 약 22ms 빠른 값을 산출하는 것을 확인할 수 있다.

ID	Real Data	Monotone Score
S ₁	{1.0, 0.3}	4.33
S ₂	{0.1, 0.6}	11.66



$$S_1\ \text{의 전송시간} = 120\text{msec} + (60\text{msec} \times 4.33) / 20 = 132.99\text{ms}$$

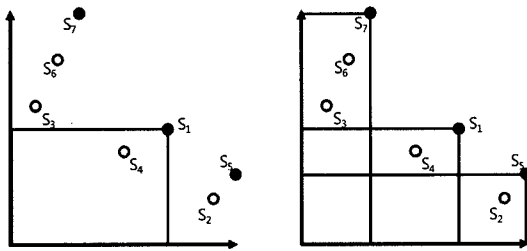
$$S_2\ \text{의 전송시간} = 120\text{msec} + (60\text{msec} \times 11.66) / 20 = 154.98\text{ms}$$

그림 5 단조 점수에 따른 전송 시간 결정

3.1.2 SFT(Skyline Filter Table)

일반적인 필터를 배포하는 기법들이나 기존에 제안된 기법에서는 단일 값을 필터로서 활용한다. min, max, top-k 등의 일차원 질의 경우 단일 값으로도 적절히 활용될 수 있지만, 스카이라인 질의와 같이 다차원을 대상으로 하는 경우 단일 값은 적절하지 않다. 그림 6과 같이 스카이라인 질의는 여러 차원에 걸쳐 동시에 서로 배

제되지 않는 값을 따져야 하므로 단일 값을 필터로 사용하는 것은 한계가 있다. 따라서 본 기법에서는 스카이라인 결과를 효과적으로 필터링 할 수 있는 스카이라인 필터 테이블(SFT)을 사용한다. SFT에는 서로 배제되지 않은 다차원 속성값을 테이블 구조로 저장하고 있다.



(a) 단일 값에 의한 필터링 (b) SFT에 의한 필터링
그림 6 다중 속성값 기반 SFT

3.2 오버히어링(Overhearing) 기반 필터링 향상 기법

앞에서 제안한 상향식 필터링 기법의 경우 부가적인 필터의 배포비용을 제거하고, 필터의 성능을 높여 많은 성능향상을 보인다. 하지만 상향식 필터링 기법의 경우 각 노드의 필터를 설정하기 위해서는 자신의 하위 노드로부터 데이터들을 수신해야 한다. 즉, 그림 7과 같이 라우팅 경로의 단말 노드는 매 주기마다 반드시 한번의 데이터를 전송해야 한다. 이로 인해 단말 노드의 에너지 소모가 증가하여 네트워크의 수명이 단축된다.

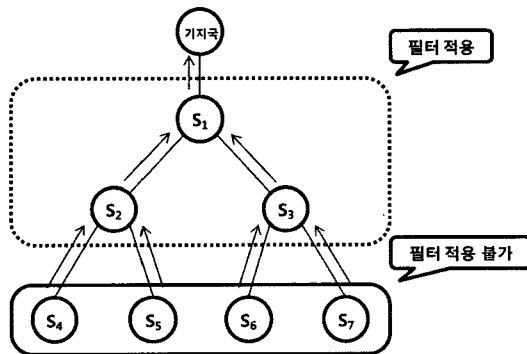


그림 7 기존 상향식 필터링 기법의 문제점

센싱 데이터는 일반적으로 지역적 연관성(Locality)을 가진다. 즉, 지리적으로 인접한 센서들의 경우 유사한 데이터를 수집할 가능성이 높다. 본 장에서는 기존 상향식 필터링 기법의 문제점을 해결하고, 데이터의 지역적 연관성을 고려하여 오버히어링 기반 필터링 향상 기법을 제안한다.

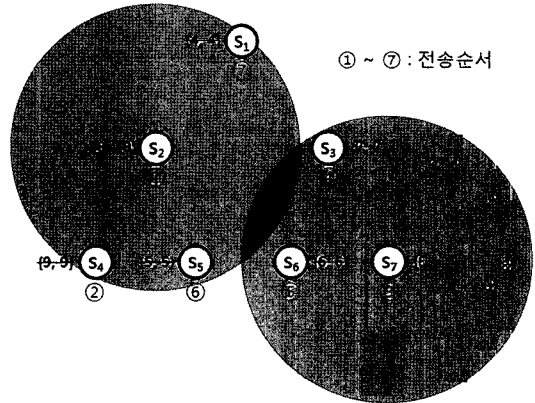


그림 8 오버히어링을 통한 이웃노드 필터링

그림 8은 상향식 필터링 기법에 오버히어링을 적용한 예이다. ①~⑦는 각 센서의 수집 데이터의 단조 점수를 기반으로 결정된 전송순서이다. 스카이라인 질의 결과에 만족할 가능성이 가장 높은 데이터를 수집한 S2노드가 제일먼저 데이터 전송을 시도한다. 이 과정에서 S2노드의 통신 범위에 S1, S4, S5노드가 포함되는 것을 그림에서 확인할 수 있다. S1, S4, S5는 S2의 데이터를 오버히어링 하고, 자신의 수집 데이터의 필터링 여부를 판단한다. 전송한 바와 같이 센서의 수집 데이터는 지역적 연관성을 가지고 있다. 따라서 인근 센서 노드는 S2와 비슷한 데이터를 수집할 가능성이 높다. 그리고 단조 점수에 의해 전송순서가 결정되므로 인근 데이터를 배제시킬 가능성이 높은 데이터가 먼저 전송된다. 따라서 S1, S4, S5의 데이터는 S2의 데이터를 오버히어링한 결과 배제됨을 판단하고, 자신의 전송시간이 되어도 데이터를 전송하지 않는다. S3과 S6 또한 S7에 의해서 배제되어 전송하지 않는다. 최종적으로 S2와 S7의 데이터만이 전송된다.

오버히어링 기법에서도 상향식 필터링 기법과 동일하게 각 노드별로 SFT를 유지한다. 따라서 가장 먼저 자신의 수집 값을 자신의 SFT에 등록하고, 이후 수신되는 데이터를 바탕으로 자신의 SFT를 업데이트한다. 그림 9의 상황에서 상향식 필터링 기법은 왼쪽 라우팅 경로와 오른쪽 라우팅 경로를 완전히 구분되어 각 경로상에 존재하는 데이터를 필터로 활용할 수 없다. 즉, 오른쪽 경로 상에서 S3의 데이터를 이용하면 S6의 데이터를 필터링 할 수 없지만 S2의 데이터를 이용하면 필터링 할 수 있다. 따라서 오버히어링 기법에서는 다른 경로의 데이터라고 할지라도 자신의 필터 값으로 활용할 수 있게 하였다. 그림의 S3의 초기 SFT에는 {7, 7}의 값이 등록되어 있지만 S2의 {10, 10}을 오버히어링하여 자신의 필터를 업데이트 한다. 이로써 S6의 {8, 8}데이터를 필터링할 수 있다.

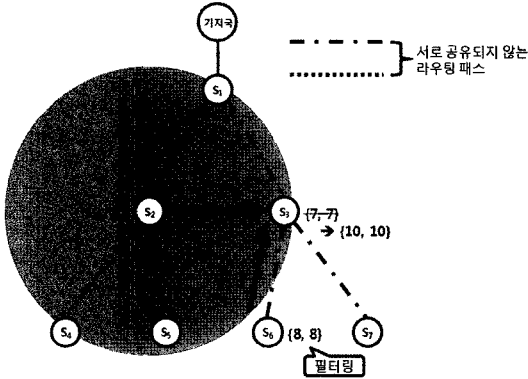
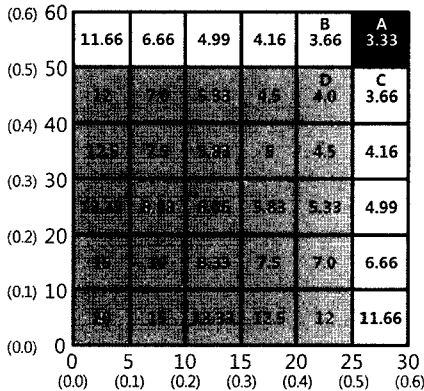


그림 9 오버헤어링을 통한 필터 업데이트

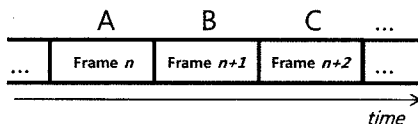
3.3 우선순위 맵을 이용한 선-필터링 기법

센서 노드에서 발생하는 에너지 소모는 다양한 동작에서 나타나지만 데이터 전송에 가장 큰 에너지 소모를 보인다. 하지만 데이터의 수신 비용 또한 적지 않은 에너지 소모를 보인다. 앞서 제안한 오버헤어링을 이용한 기법은 많은 긍정 오류 전송을 줄일 수 있다. 하지만 데이터를 오버헤어링하기 위해서 주변 노드들은 많은 수신 비용을 소모하게 된다. 따라서 본 장에서는 오버헤어링을 하지 않고, 많은 긍정 오류를 제거할 수 있는 기법을 제안한다.

우선순위 맵을 이용한 선-필터링 기법은 크게 두 단계로 진행된다. 첫 번째 단계는 우선순위 맵을 이용한



(a) 우선순위 맵



(b) 필터링 비컨 전송 프레임

그림 10 우선순위 맵을 이용한 필터링

선-필터링 단계이고, 두 번째 단계는 필터링 단계를 거친 데이터의 수집 단계이다. 우선순위 맵은 수집할 데이터 범위의 각 차원을 다차원 공간에 매핑하고, 해당 공간을 그리드 구조로 분할한다. 그리고 그리드 각 셀에 할당된 값을 이용해 단조 점수를 계산하여 그림 10의 (a)와 같이 만든다. 계산된 점수를 기반으로 각 셀 별 전송시간을 시간 프레임 단위로 할당한다. 즉, 그림 10의 (b)처럼 점수가 가장 낮은 A셀을 n번째 프레임에 할당하고, 그 다음 낮은 B, C 셀을 n+1번째와 n+2번째 프레임에 할당한다.

데이터를 수집한 각 센서 노드들은 자신의 데이터가 어떠한 셀에 포함되는지 판단한다. 모든 센서들을 추가적인 통신 없이 우선순위 맵과 전송 프레임의 시간 단위만을 이용해 자신의 데이터가 포함되는 전송 프레임을 산출할 수 있다. 데이터 수집 첫 번째 단계에서 이와 같은 방식으로 모든 센서들은 자신의 전송 프레임을 결정하고, 해당 프레임이 전송할 시점이 되면 필터링 비컨 신호를 전체 네트워크에 방송(Broadcast) 한다. 비컨 신호를 수신한 센서들은 필터링 작업을 수행한다. 그림 11의 상황을 예로 들면, 센서 노드 S₁이 데이터 {56, 29}를 수집하였고, S₂가 데이터 {45, 21}를 수집하였다. S₁의 전송 프레임은 A셀의 전송 프레임이고, S₂의 전송 프레임은 D셀의 전송 프레임이다. S₁은 전송 프레임의 전송 순서에 따라 Frame n의 시간에 비컨 신호를 방송하고, S₂는 자신의 전송 시간까지 대기한다. S₂는 대기 도중에 S₁이 방송한 비컨 신호를 수신한다. S₂는 비컨 신호를 수신한 시간 정보를 통해 현재 A셀의 범위에 포함되는 데이터를 수집한 센서의 유무를 판단할 수 있다. 즉, Frame n의 전송시간 범위에 비컨 신호를 수신하게 되면 정확히 어떤 노드가 어떤 데이터를 수집하였는지는 판단하지 못하더라도 하나 이상의 임의의 노드가 {60~50, 30~25}의 범위에 포함되는 데이터를 수집하였

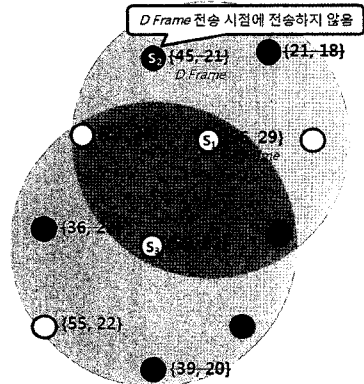


그림 11 우선순위 맵을 이용한 필터링 단계

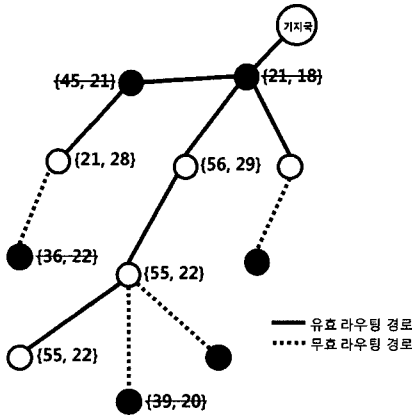


그림 12 데이터 라우팅 단계

```

sensors = {s1, s2, s3, ... sn} // 네트워크를 구성하는 센서 노드
// 기지국에서의 10x10 크기의 우선순위 맵 생성
mapResolution = 10x10
priorityMap[] = createPriorityMap(mapResolution.x, mapResolution.y)
// 우선순위 맵의 각 셀 별 단위점수 생성 및 저장
WHILE map_x = 0 TO mapResolution.x DO
  WHILE map_y = 0 TO mapResolution.y DO
    // 단위 셀 별 대표점수 추출
    cv_x = (max_v_x / mapResolution.x) * map_x
    cv_y = (max_v_y / mapResolution.y) * map_y
    // 단위 셀 별 단위점수 계산 및 저장
    priorityMap[map_x][map_y].V_priority = (1/cv_x) + (1/cv_y)
  END WHILE
END WHILE
// 기지국에서 전체 센서들로 우선순위 맵 배포
BSBroadcastToNetwork(priorityMap)
Round = 0
WHILE network == live DO
  Round++
  FOR EACH s = sensors[i] DO
    v_x = x-value of s
    v_y = y-value of s
    // 각 센서마다 자신이 수집한 데이터에 따른 필터링 비전용 송신할 프레임 계산
    FrameNo = 0
    FOR EACH cell = priorityMap[] DO
      IF (1/v_x + 1/v_y) >= cell.V_priority THEN
        s.FBeaconTrFrame = FrameNo
        BREAK
      END IF
      FrameNo++
    END FOR
    // 각 센서마다 자신이 수집한 데이터에 따른 실제 데이터 송신할 시간 계산
    time_v of s = (Round - 1) * T_round + ((T_round * (1/v_x + 1/v_y)) / (1/max_v_x + 1/max_v_y))
  END FOR
  // 해당 프레임에서 각 센서의 필터링 비전 발생
  FOR EACH frame = transmitFrame(i) DO
    FOR EACH s = sensors[i] DO
      IF s.FBeaconTrFrame == frame THEN
        s.Broadcast(beacon)
      END IF
    END FOR
  END FOR
  .....
  // 실제 데이터 수집 단계 생략 /
  .....
END WHILE

```

그림 13 우선순위 맵 기반 데이터 필터링

음을 판단할 수 있다. A셀의 가장 왼쪽 하단의 꼭지점은 전체 데이터 범위상에 {0~50, 0~25}의 데이터 범위를 배제할 수 있다. 따라서 {0~50, 0~25}의 범위에 포함되는 센서들은 비컨 신호를 통해서 자신의 수집 데이터를 필터링 하게 된다.

데이터 수집 두 번째 단계에서는 그림 12와 같이 기본 라우팅 경로상에서 필터링 된 센서 노드들이 연결된 경로는 무효 라우팅 경로로서 사용되지 않고 필터링 되지 않은 센서 노드들의 데이터만을 라우팅 한다. 수집의 기본 전략은 앞서 제안한 상황식 필터링 기법과 동일한 방식을 이용한다.

그림 13은 제안하는 기법에서 우선순위의 맵을 이용한 데이터 필터링 기법을 보여준다.

4. 성능평가 및 분석

4.1 실험 환경

제안하는 기법의 우수성을 보이기 위해 기존 기법인 MFTAC와 시뮬레이션을 통해 성능을 비교 및 평가 하였다. 비교한 기법은 제안하는 기법(PBFiltering), 상황식 필터 설정 기법(BUF), 그리고 MFTAC 이다. 시뮬레이션은 표 1과 같은 환경 변수를 바탕으로 그림 14와 같은 센서 네트워크 토폴로지로 구성하여 진행하였고, 시뮬레이션 툴은 J-SIM을 이용하였다[10].

표 1 성능 평가 환경

파라미터	값
센서 개수	40~400
네트워크 크기	50x50m~150x150m
센서 식별자	4byte
센싱 데이터의 크기	4byte
센싱 속성의 차원수	2
센서 통신 반경	10m~30m
우선순위 맵 해상도	{1x1}~{10x10}
초기 센서 에너지	0.1J

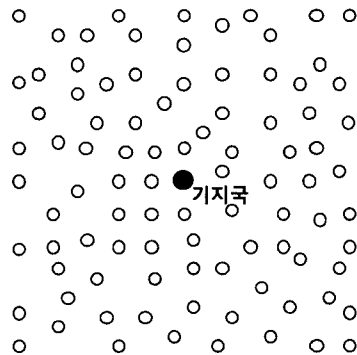


그림 14 다중-홉 랜덤 네트워크 토폴로지

센서 노드의 메시지 송수신에 소모되는 에너지 식 (3) 과 식 (4)의 모델을 사용하였다[11,12].

$$T_{cost} = 50nJ/b, \quad R_{cost} = 50nJ/b, \quad T_{amp} = 100pJ/b/m^2$$

$$SEND_{cost} = MSG_{size} \cdot (T_{cost} + T_{amp} \cdot T_{dist}^2) \quad (3)$$

$$RECV_{cost} = MSG_{size} \cdot R_{cost}$$

표 2 에너지 소모 모델에 사용된 파라미터

파라미터	설명
$SEND_{cost}$	총 전송 에너지
MSG_{size}	메시지 크기
T_{cost}	전송 비용
T_{amp}	증폭 비용
T_{dist}	전송 거리
$RECV_{cost}$	총 수신 에너지
R_{cost}	수신 비용

시뮬레이션에서 사용된 데이터는 미국 워싱턴주에서 측정된 온도와 습도 데이터이며, 여러 개의 센서 노드로 구성된 센서 네트워크를 시뮬레이션하기 위해서 데이터 모델의 위상차를 두어 각 센서 노드가 서로 다른 데이터를 수집하도록 설정하였다. 센서 데이터는 그림 15와 같은 분포를 보인다[13].

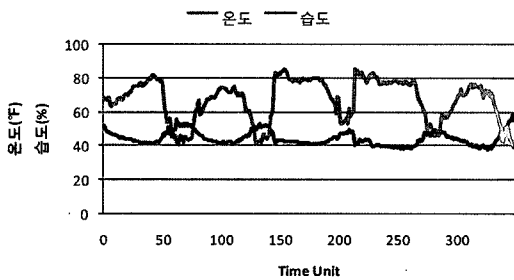


그림 15 시뮬레이션에 사용한 데이터 모델

4.2 성능평가 결과

4.2.1 센서 배포밀도에 따른 성능평가

그림 16과 그림 17은 센서 네트워크의 노드 배포밀도에 따른 네트워크 수명과 평균 긍정 오류에 대하여 성능 평가한 결과이다. MFTAC의 경우, 앞서 설명한 바와 같이 단일 값 기반의 필터를 활용한다. 따라서 다차원, 다중 속성 값을 고려하여 처리하는 스카이라인 질의에 적합하지 않다. 이로 인해 불필요한 데이터 전송을 적절히 제거하지 못해 다수의 긍정 오류 데이터가 지구적으로 전송된다. BUF의 경우, 데이터를 전송하는 과정에서 필터를 설정하는 상향식 필터링 기법을 통해 많은 수의 긍정 오류를 줄였다. 하지만, 단일 노드의 데이터를 필터링하지 못하는 단점을 가지고 있다. PBFiltering

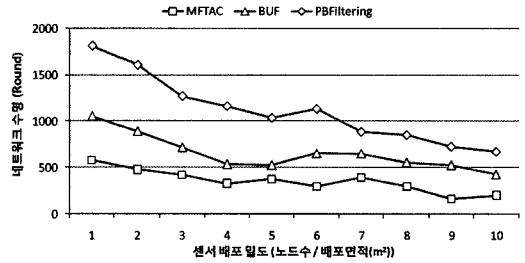


그림 16 센서 배포 밀도에 따른 네트워크 수명 변화

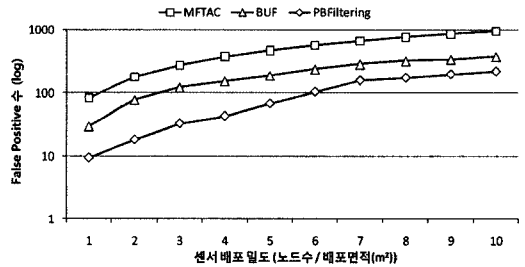


그림 17 센서 배포 밀도에 따른 평균 긍정 오류 수

의 경우 상향식 필터를 설정하기 이전에 우선순위 맵을 이용한 선-필터링 과정을 수행함으로써 BUF의 단점을 보완하였다. 성능평가 결과, MFTAC에 대비하여 BUF는 평균 긍정 오류의 수가 약 59% 감소하였고, 네트워크의 수명이 약 195% 증가하였다. 선-필터링 과정을 적용한 PBFiltering은 평균 긍정 오류의 수가 약 83% 감소하였고, 네트워크 수명이 약 326% 증가하였다.

4.2.2 네트워크 크기에 따른 성능평가

그림 18과 그림 19는 센서 네트워크의 배포 면적에 따른 네트워크 수명과 평균 긍정 오류에 대하여 성능 평가한 결과이다. 배포 면적을 증가 비율에 따라 배포된 센서의 수를 증가시켜 동일한 배포밀도를 유지시켰다. 성능평가 결과, MFTAC에 대비하여 BUF는 평균 긍정 오류의 수가 약 36% 감소하였고, 네트워크의 수명이 약 153% 증가하였다. 선-필터링 과정을 적용한 PBFiltering은 평균 긍정 오류의 수가 약 86% 감소하였고, 네트워크 수명이 약 277% 증가하였다.

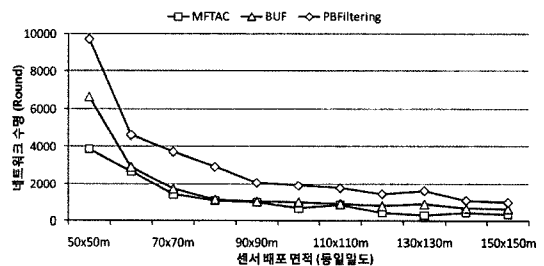


그림 18 센서 네트워크 크기에 따른 네트워크 수명 변화

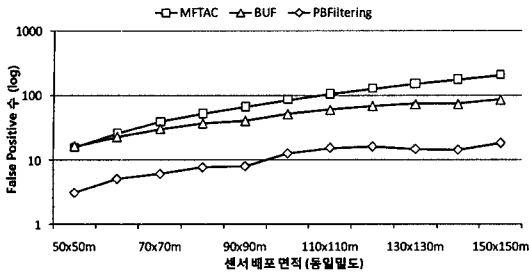


그림 19 센서 네트워크 크기에 따른 평균 긍정 오류 수

4.2.3 우선순위 맵의 해상도에 따른 필터링 성능평가

그림 20은 PBFitering에 사용하는 우선순위 맵의 해상도에 따른 긍정 오류 발생량을 평가한 결과이다. 1x1은 셀 하나의 크기가 데이터 전체 범위를 포함하는 크기 이므로 선-필터링 과정에서 걸리지는 긍정 오류는 존재하지 않는다. 10x10은 데이터 전체 범위를 100개의 셀로 분할하여 선-필터링을 수행한다. 성능평가는 이러한 해상도를 1x1에서 10x10까지 순차적으로 증가시켜 가며 필터링 되는 긍정 오류 수의 변화를 나타낸 것이다. 성능평가 결과, 10x10의 해상도를 가지는 우선순위 맵을 이용할 경우 약 96%까지 긍정 오류를 제거할 수 있다.

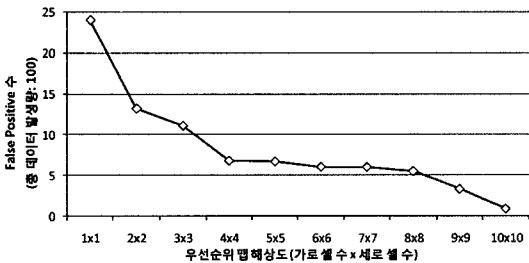


그림 20 우선순위 맵의 해상도에 따른 긍정 오류 수

5. 결론

본 논문에서는 기존에 제안된 스카이라인 질의 처리기법의 문제점을 분석하고, 에너지 효율적인 연속 스카이라인 질의 처리기법을 제안한다. 기존의 질의 처리 기법의 경우, 결과에 포함되지 않는 데이터의 전송을 막기 위해 MFT를 네트워크 전체에 배포한다. 이 과정에서 많은 양의 에너지 소모가 발생하고, 배포된 필터의 필터링 기능이 뛰어나지 않아 다수의 긍정 오류 데이터 전송이 발생한다. 이러한 문제점을 해결하기 위하여 필터를 배포하는 단계를 없애고, 데이터 수집과정에서 필터를 설정하는 상황식 필터 설정 기법을 제안하였다. 단조 점수를 기반으로 스케줄링 된 데이터 전송 순서에 따라 다

중 MFT기반 SFT를 설정하여 필터링 성능을 최적화한다. 또한 필터링 성능을 높이기 위하여 오버헤더링을 이용한 필터링 기법과 우선순위 맵을 이용한 선 필터링 기법을 제안하여 불필요한 데이터의 전송을 최소화한다. 실험 결과, 기존 스카이라인 질의 처리기법에 비해 네트워크의 수명이 향상되었고, 기지국에 전송되는 스카이라인 질의에 포함되지 않은 데이터의 수가 감소하였다.

참고 문헌

- [1] R. Szweczyk, A. Mainwaring, J. Polastre, and D. Culler, "An Analysis of a Large Scale Habitat Monitoring Application," *Proc. ACM Conf. Embedded Networked Sensor Systems (SenSys '04)*, Nov. 2004.
- [2] R. Szweczyk, E. Osterweil, J. Polastre, M. Hamilton, A. Mainwaring, and D. Estrin, "Habitat Monitoring with Sensor Networks," *Comm. ACM*, vol.47, no.6, pp.34-40, Jun. 2004.
- [3] S. Madden, M. Franklin, J. Hellerstein, and W. Hong, "TAG: A Tiny Aggregation Service for Ad Hoc Sensor Networks," *Proc. Usenix Fifth Symp. Operating Systems Design and Implementation (OSDI '02)*, pp.131-146, Dec. 2002.
- [4] A. Sharaf, J.Beaver, A. Labrinidis, and K. Chrysanthis, "Balancing Energy Efficiency and Quality of Aggregate Data in Sensor Networks," *VLDB*, vol.13, no.4, pp.384-403, 2004.
- [5] K. Yoon, J. Choi, Y. Chung, and S. Lee, "In-Network Processing for Skyline Queries in Sensor Networks," *IEICE Trans. COMMUN.*, vol.E90-B, no.12, Dec. 2007.
- [6] S. Borzsonyi, D. Kossmann, and K. Stocker, "The skyline operator," *Proc. International Conference on Data Engineering*, pp.421-430, 2001.
- [7] J. Chomicki, P. Godfrey, J. Gryz, and D. Liang, "Skyline with presorting," *Proc. International Conference on Data Engineering*, pp.717-719, 2002.
- [8] Z. Huang, C.S. Jensen, H. Lu, and B.C. Ooi, "Skyline queries against mobile lightweight devices in MANETs," *Proc. International Conference on Data Engineering*, p.66, 2006.
- [9] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. Mobile Computing*, vol.3, no.4, pp.366-379, 2004.
- [10] The J-SIM Simulator, <http://www.j-sim.org>, Accessed in Jan. 2007.
- [11] W. Heinzelman, "Application-Specific Protocol Architectures for Wireless Networks," PhD dissertation, Massachusetts Inst. Of Technology, 2000.
- [12] X. Tang, and J. Xu, "Extending Network Lifetime for Precision-Constrained Data Aggregation in Wireless Sensor Networks," *Proc. IEEE INFOCOM*, Apr. 2006.

[13] Live from Earth and Mars (LEM) Project, <http://www-k12.atmos.washington.edu/k12/grayskies/>, 2006.

여 명 호

한국정보과학회논문지 : 데이터베이스
제 36 권 제 4 호 참조

성 동 욱

한국정보과학회논문지 : 데이터베이스
제 36 권 제 3 호 참조

유 재 수

한국정보과학회논문지 : 데이터베이스
제 36 권 제 4 호 참조



박 준 호

2008년 2월 충북대학교 정보통신공학과
공학사. 2008년 3월~현재 충북대학교
정보통신공학과 석사과정. 관심분야는 무선
센서 네트워크, RFID 시스템, LMS/
LCMS, 데이터베이스 시스템 등



김 학 신

2008년 2월 충주대학교 컴퓨터공학과 공
학사. 2008년 3월~현재 충북대학교 정
보통신공학과 석사과정. 관심분야는 무선
센서 네트워크, 데이터베이스 시스템, RFID,
FLASH 메모리 파일 시스템



박 형 순

2008년 2월 충북대학교 정보통신공학과
공학사. 2008년~현재 충북대학교 정
보통신공학과 석사과정. 관심분야는 무선
센서 네트워크, 데이터베이스 시스템, LMS/
LCMS, FLASH 메모리 저장 시스템 등



노 규 중

2008년 2월 충북대학교 정보통신공학과
공학사. 2008년 3월~현재 충북대학교
정보통신공학과 석사과정. 관심분야는 무
선 센서 네트워크, 데이터베이스 시스템,
RFID, FLASH 메모리 파일 시스템