

# 무선 센서 네트워크에서 우선순위 필터링을 이용한 스카이라인 질의 처리 기법

성동욱\*, 여명호\*, 박준호\*, 유재수\*

\*충북대학교 정보통신공학과

e-mail : {sergei, mhyeo, jhpark}@netdb.cbnu.ac.kr, yjs@chungbuk.ac.kr

## Priority Filtering-based Skyline Query Processing in Wireless Sensor Networks

Dong-Ook Seong\*, Myung-Ho Yeo\*, Jun-Ho Park\*, Jae-Soo Yoo\*

\*Dept of Information and Communication Engineering, Chung-buk National University

### 요 약

센서 네트워크에서 병합 질의를 효율적으로 처리하기 위한 다양한 인-네트워크 질의 처리 기법이 제안되었다. 스카이라인 질의는 일반적인 병합 질의와 달리 다차원 데이터에 대한 비교를 요구하므로 인-네트워크 처리가 쉽지 않다. 스카이라인 질의를 에너지 효율적으로 처리하기 위해서 불필요한 데이터의 전송을 제거하는 것이 중요하다. 기존에 제안된 스카이라인 질의 처리 기법은 전체 네트워크에 필터를 배포함으로써 불필요한 데이터 전송을 차단한다. 하지만 많은 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 기법을 제안한다. 제안하는 기법은 데이터를 수집하는 과정에서 스카이라인 필터 테이블(SFT)설정하는 상향식 필터링을 수행한다. 그리고 선-필터링(Pre-filtering) 기법을 통해 필터 효과를 증가시킨다. 제안하는 알고리즘의 우수성을 보이기 위해 시뮬레이션을 통해 기존에 제안된 MFTAC 기법과 비교하였으며, 그 결과 평균 False Positive가 평균 84.44% 감소하였고, 네트워크 수명이 약 75.99% 증가하였다.

### 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 정보를 수집할 수 있는 센서 소자와 단거리 RF 통신이 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 더 이상 센서로서 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. 기

존 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 CO<sub>2</sub>, CO, SO<sub>2</sub> 등의 다양한 값들을 고려해야 한다. 이러한 경우 스카이라인 질의를 이용해 처리 할 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 되

이 논문은 2008년도 교육과학기술부로부터 지원(지역거점연구단육성사업/충북 BIT 연구중심대학육성사업단)과 산업자원의 지역혁신 인력양성사업의 연구결과로 수행되었음.

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

2. 관련연구

Block Nested Loop(BNL)[4]는 순차적으로 전체 데이터를 스캔하고, 메모리에 저장된 후보 데이터와 비교한다. Sorted Filter Skyline(SFS)[5]은 연산을 통해 모든 속성을 하나의 단일 값으로 만들고 정렬을 통해 비교 대상의 수를 줄인다. Divide and Conquer(D&C)[4]는 데이터의 집합을 많은 영역으로 나누고, 각 영역에서 스카이라인을 계산한다. 그 다음, 각 영역의 스카이라인으로부터 최종 스카이라인을 결정한다. 하지만, 이 방법들은 중앙처리를 통해 수행됨으로 센서 네트워크와 같은 분산 환경에 적합하지 않다.

[6]과 같은 분산 환경을 고려한 스카이라인 질의 처리 기법은 웹 정보 시스템을 대상으로 하고 있기 때문에 쉽게 센서 네트워크에 적용할 수 없다. [6]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. [3]는 [6]보다 향상된 필터링 기법을 제안한다. [3]의 MFTAC 는 먼저 센싱한 데이터를 이용하여 가장 많은 데이터를 필터링할 수 있는 MFT(Min-Score Filtering Tuple)를 계산하고, MFT 를 센서 노드에 배포함으로써 불필요한 데이터의 전송을 줄인다. 또한, MFT 배포를 위한 비용을 분산하기 위해서 HEED 클러스터링 방법을 이용하여 클러스터 헤드가 MFT 를 계산하고 배포하도록 한다. 하지만 클러스터 헤드에 의해 MFT 배포 비용이 분산된다 하더라도 연속적인 스카이라인 질의를 처리하기 위해서는 매 주기마다 MFT 를 클러스터 멤버로 배포해야 한다. 결국, 전체 센서 노드가 MFT 의 송/수신에 참여하게 된다. 또한, 클러스터나 라우팅 트리에 의한 MFT 의 계산은 지역적인 데이터를 반영하고 있기 때문에 많은 불필요한 데이터들이 기지국으로 전송된다. 예를 들면, {S<sub>2</sub>, S<sub>4</sub>, S<sub>5</sub>}로 구성된 클러스터는 스카이라인 질의의 결과로 {S<sub>2</sub>, S<sub>4</sub>}를 보고한다. 하지만, 기지국에 계산된 최종 스카이라인 결과에서 {S<sub>2</sub>, S<sub>4</sub>}는 S<sub>3</sub> 에 의해 제거된다. 결국 {S<sub>2</sub>, S<sub>4</sub>}는 불필요한 데이터이다. 효과적인 스카이라인 질의 처리를 위해서 필

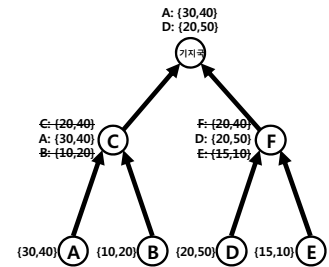
터 배포비용을 줄이고, 효과적으로 False Positive 를 줄이기 위한 연구가 필요하다.

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 MFT 는 전송 경로상의 부분 데이터만을 반영하고 있기 때문에 필터의 성능이 우수하지 않다. 이로 인해 많은 False Positive 데이터가 발생하여 네트워크 수명을 단축시킨다. 본 논문에서는 필터를 데이터 수집 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.

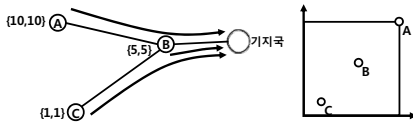
3.1 상향식 필터 설정 기법

제안하는 기법은 필터를 배포하기 위해 필요한 추가적인 통신비용을 줄이기 위해 센싱 데이터를 기지국으로 전송하는 과정에서 필터를 설정한다. 최종 각 센서 노드들은 자신의 수집 데이터를 필터로 설정한다. 그리고 라우팅 경로상에서 단말노드를 시작으로 상위노드로 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 도달 할 때까지 반복한다. 최종적으로 기지국은 False Positive 를 어느 정도 제거된 데이터 만을 수신된다. 그림 1 을 보면 단말 노드인 A, B 노드의 센싱 데이터 {30,40}, {10,20}와 C 노드의 필터값 {20,40}은 C 노드에서 상호 배제유무를 판단할 수 있다. 그 결과 B, C 노드의 데이터는 배제되고, A 노드의 센싱 데이터만 최종적으로 기지국으로 전송된다.



(그림 1) 상향식 필터 설정 과정

앞서 제안한 상향식 필터 설정 기법은 한가지 문제점을 내포하고 있다. 각 센서 노드에서 동일한 데이터가 수집된다면 전송되는 순서에 따라 데이터 전송 횟수가 틀려진다. 그림 2의 예를 보면, 센서 노드 A, B, C는 각각 {10,10}, {5,5}, {1,1}의 데이터를 수집하였다. 여기서 데이터를 전송 순서가 C→B→A 일 경우, A 노드는 1 번, B 노드는 2 번 그리고 C 노드는 1 번의 전송이 발생한다. 하지만 전송 순서가 A→B→C 일 경우, A 노드는 1 번, B 노드는 1 번 그리고 C 노드는 1 번의 전송이 발생한다. 그 이유는 전송 순서에 따라서 B 노드의 최종 필터값이 설정 되는 시점이 달라지기 때문이다. A 노드의 {10,10}의 데이터가 가장 먼저 전송될 경우 B 노드의 필터값으로 {10,10}이 설정된다. 이로 인해 B, C 노드의 {5,5}, {1,1}은 기지국으로 전송되기 전에 B 노드에서 배제시킬 수 있다. 즉, 스카이라인의 결과에 만족할 가능성이 높은 순서대로 수집순서를 결정하면 필터의 성능을 최적화시킬 수 있다.



(그림 2) 수집 순서에 따른 성능 변화

본 논문에서는 Monotone Score 를 이용하여 각 센서 노드의 수집 데이터를 바탕으로 전송 우선순위를 결정하는 기법을 제안한다. Monotone Score 는 식 1 을 이용하여 다차원 값을 단일 스코어로 변환한 값이다. m 은 차원의 수를 말하고, p<sub>j</sub> 는 각 차원 별 속성값을 의미한다. 변환 결과, 스카이라인 결과에 만족 할 가능성이 높을수록 낮은 값을 보인다. 따라서 Monotone Score 에 따라 전송 시각을 결정하게 되면 스카이라인 결과에 포함될 가능성이 높을수록 먼저 전송하여 필터의 성능을 높일 수 있다. 식 2 를 이용해 전송 시간을 결정 할 수 있다. R<sub>c</sub> 는 현재 라운드 수를 의미하고, T 는 1 라운드의 시간 범위를 나타낸다. S<sub>max</sub> 는 Monotone Score 의 최대값을 나타내고, S<sub>own</sub> 은 현재 값의 Monotone Score 값을 나타낸다.

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

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

일반적인 필터를 배포하는 기법들이나 기존에 제안된 기법에서는 단일 값을 필터로서 활용한다. min, max, top-k 등의 일차원 질의 경우 단일 값으로도 적절히 활용 될 수 있지만, 스카이라인 질의와 같이 다차원을 대상으로 하는 경우 단일 값은 적절하지 않다. 스카이라인 질의는 여러 차원에 걸쳐 동시에 서로 배제되지 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. SFT 에는 서로 배제되지 않은 다차원 속성값을 테이블 구조로 저장하고 있다.

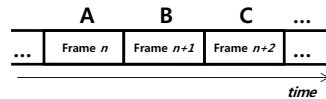
3.2 Priority Map 을 이용한 선-필터링 기법

앞에서 제안한 상향식 필터링 기법은 라우팅 경로의 단말 노드는 필터가 적용되지 않아 매 주기마다 데이터를 전송해야 한다. 이로 인해 단말 노드의 에너지 소모가 증가하여 네트워크의 수명이 단축된다. 본 장에서는 기존 상향식 필터링 기법의 문제점을 해결하는 Priority Map 을 이용한 선-필터링 기법을 제안한다.

Priority Map 을 이용한 선-필터링 기법은 크게 두 단계로 진행된다. 첫 번째 단계는 Priority Map 을 이용한 선-필터링 단계이고, 두 번째 단계는 필터링 단계를 거친 데이터의 수집 단계이다. Priority Map 은 수집할 데이터 범위의 각 차원을 다차원 공간에 매핑하고, 해당 공간을 그리드 구조로 분할한다. 그리고 그리드 각 셀에 할당된 값을 이용해 Monotone Score 를 계산하여 그림 3 의 (a)와 같이 만든다. 계산된 Score 를 기반으로 각 셀별 전송시간을 시간 프레임 단위로 할당한다. 즉, 그림 10 의 (b) 처럼 Score 가 가장 낮은 A 셀을 n 번째 프레임에 할당하고, 그 다음 낮은 B, C 셀을 n+1 번째와 n+2 번째 프레임에 할당한다.

(0.6)	60	11.66	6.66	4.99	4.16	B 3.66	A 3.33	
(0.5)	50	12	7.0	5.33	4.5	D 4.0	C 3.66	
(0.4)	40	12.5	7.5	5.83	5	4.5	4.16	
(0.3)	30	13.33	8.33	6.66	5.83	5.33	4.99	
(0.2)	20	15	10	8.33	7.5	7.0	6.66	
(0.1)	10	20	15	13.33	12.5	12	11.66	
(0.0)	0	0	0	0	0	0	0	
		(0.0)	(0.1)	(0.2)	(0.3)	(0.4)	(0.5)	(0.6)

(a) Priority Map



(b) 전송 프레임

(그림 3) Priority Map 을 이용한 필터링

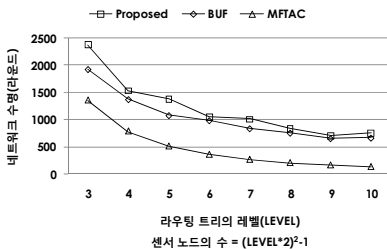
데이터를 수집한 각 센서 노드들은 자신의 데이터가 어떠한 셀에 포함되는지 판단한다. 모든 센서들을 추가적인 통신 없이 Priority Map 과 전송 프레임의 시간 단위만을 이용해 자신의 데이터가 포함되는 전송 프레임을 산출할 수 있다. 데이터 수집 첫 번째 단계에서 이와 같은 방식으로 모든 센서들은 자신의 전송 프레임을 결정하고, 해당 프레임이 전송할 시점이 되면 Beacon 신호를 전체 네트워크에 브로드캐스트 한다. Beacon 신호를 수신한 센서들은 필터링 작업을 수행한다. 예를 들어, 센서 노드 S<sub>1</sub> 이 데이터 {56,29}를 수집하였고, S<sub>2</sub> 가 데이터 {45,21}를 수집하였다. S<sub>1</sub> 의 전송 프레임은 A 셀의 전송 프레임이고, S<sub>2</sub> 의 전송 프레임은 D 셀의 전송 프레임이다. S<sub>1</sub> 은 전송 프레임의 전송 순서에 따라 Frame n 의 시간에 Beacon 신호를 브로드캐스트 하고, S<sub>2</sub> 는 자신의 전송 시간까지 대기한다. S<sub>2</sub> 는 대기 도중에 S<sub>1</sub> 이 브로드캐스트한 Beacon 신호를

수신한다.  $S_2$  는 Beacon 신호를 수신한 시간 정보를 통해 현재 A 셀의 범위에 포함되는 데이터를 수집한 센서의 유무를 판단할 수 있다. 즉, Frame n 의 전송시간 범위에 Beacon 신호를 수신하게 되면 정확히 어떤 노드가 어떤 데이터를 수집하였는지는 판단하지 못하더라도 하나 이상의 임의의 노드가 {60-50,30-25}의 범위에 포함되는 데이터를 수집하였음을 판단 할 수 있다. A 셀은 전체 데이터 범위상에 {0-50,0-25}의 데이터 범위를 배제할 수 있다. 따라서 {0-50,0-25}의 범위에 포함되는 센서들은 Beacon 신호를 통해서 자신의 수집 데이터를 필터링 하게 된다. 데이터 수집 두 번째 단계에서는 기본 라우팅 경로상에서 필터링 된 센서 노드들이 연결된 경로는 무효 라우팅 경로로서 사용되지 않고 필터링 되지 않은 센서 노드들만 데이터를 라우팅 한다. 수집 전략은 앞서 제안한 상향식 필터링 기법과 동일한 방식을 이용한다.

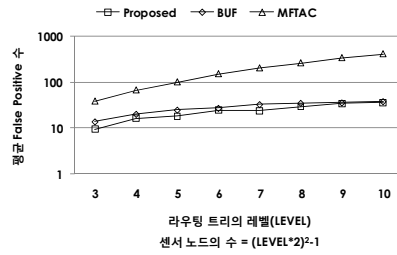
4. 성능평가 및 분석

제안하는 알고리즘의 우수성을 보이기 위해 기존에 제안된 연속적인 스카이라인 질의 처리 알고리즘과 성능을 비교평가 하였다. 센서는 40~400 개를 배포하였고, 데이터 차원수는 2, 데이터 크기는 4byte 로 설정하였다. 사용된 데이터는 미국 워싱턴주에서 측정된 온도와 습도 데이터이다.

그림 4 와 그림 5 은 센서 노드 수가 증가됨에 따라 제안하는 기법과 BUF, MFTAC 의 네트워크 수명과 평균 False Positive 를 비교한 결과이다. MFTAC 의 경우, 단일값 기반의 필터를 활용한다. 따라서 다차원, 다중값의 특성을 가지는 스카이라인 질의 데이터를 처리하는데 적합하지 않다. 이로 인해 불필요한 데이터 전송을 적절히 제거되지 못해 다수의 False Positive 데이터가 기지국으로 전송된다. BUF 의 경우, 데이터를 전송하는 과정에서 필터를 설정하는 상향식 필터링 기법을 통해 많은 수의 False Positive 를 줄였다. 하지만, 단일 노드의 데이터를 필터링하지 못하는 단점을 가지고 있다. 본 논문에서 제안하는 기법은 상향식 필터를 설정하기 이전에 비컨을 통해 필터를 배포함으로써 상향식 필터링 기법의 단점을 보완하였다. 그 결과, 기존 기법에 비해 평균 False Positive 의 수가 약 84.44% 감소하였고, 네트워크의 수명이 약 75.99% 증가하였다.



(그림 4) 네트워크 수명



(그림 5) 평균 False Positive 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 MFT 를 네트워크 전체에 배포한다. 이 과정에서 많은 양에 에너지 소모가 발생하고, 배포된 필터의 필터링 기능이 뛰어나지 않아 다수의 false positive 데이터 전송이 발생한다. 이러한 문제점을 해결하기 데이터 수집과정에서 필터를 설정하는 상향식 필터 설정기법을 제안하였다. Monotone Score 를 기반으로 스케줄링 된 데이터 전송 순서에 따라 다중 MFT 기반 SFT 를 설정하여 필터링 성능을 최적화 한다. 또한 필터링 성능을 높이기 위하여 Priority Map 을 이용한 선 필터링 기법을 제안하여 불필요한 데이터의 전송을 최소화 한다. 실험 결과, False Positive 가 평균 84.44% 감소하였고, 네트워크 수명이 약 75.99% 증가하였다

참고문헌

- [1] R. Szweczyk, A. Mainwaring, J. Polastre, and D. Culler, "An Analysis of a Large Scale Habitat Monitoring Application", 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] 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.
- [4] S. Borzsonyi, D. Kossmann, and K. Stocker, "The skyline operator", Proc. International Conference on Data Engineering, pp.421-430, 2001.
- [5] J. Chomicki, P. Godfrey, J. Gryz, and D. Liang, "Skyline with presorting," Proc. International Conference on Data Engineering, pp.717-719, 2002.
- [6] 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.