

센서의 불균일한 배포밀도를 고려한 효율적인 데이터 중심 저장기법

An Efficient Data Centric Storage Scheme with Non-uniformed Density of Wireless Sensor Networks

성동욱, 이석재, 송석일*, 유재수
충북대학교 정보통신공학과, 충주대학교 컴퓨터공학과*

Seong dong-ook, Lee seek-jae, Song seek-il*,
Yoo jae-soo
Dept of Information & Communication Engineering,
Chung-Buk University, Dept. of Computer Engineering,
Chungju National University*

요약

최근 자연 환경 조사, 군사 응용 시스템, 환경 변화 감시 등의 여러 응용 분야에서 무선 센서 네트워크 자체를 데이터 저장 장치로 활용하는 데이터 중심 저장기법(DCS)에 대한 연구들이 진행되고 있다. 기존에 제안된 데이터 중심 저장기법들은 데이터들의 지역성이 고려되지 않아 범위질의 처리에 비효율적인 문제점(Q-NiGHT)과 센서 배포밀도가 불균일한 경우 각 센서의 데이터에 대한 저장 부하가 불균등해지는 문제점(KDDCS)을 갖는다. 본 논문에서는 이러한 기존의 문제점을 해결하고자, 데이터 저장 위치의 지역성(Locality)를 만족시키며 센서 배포 밀도를 고려한 새로운 데이터 중심 저장 기법을 제안한다. 제안하는 기법은 전체 센서 영역을 그리드로 나누고, 각 셀의 센서 밀도정보를 비트로 표현하는 배포 밀도 비트맵을 만들어 배포한다. 센서들은 데이터를 저장하거나 탐색할 때 배포 밀도 비트맵을 이용하여 데이터 저장 위치를 계산함으로써 별도의 통신 없이 최적의 저장 위치를 선정할 수 있다. 본 논문에서는 시뮬레이션을 통해 기존 연구와의 성능 평가 하여 센서들의 배포밀도가 균일한 환경뿐만 아니라 불균일한 환경에서도 각 센서들의 데이터 저장부하를 고르게 분산시키고, 범위질의 처리 또한 효율적으로 처리됨을 보였다.

Abstract

Recently Data Centric Storage (DCS) schemes are variously studied for several applications (e.g. natural environment investigation, military application systems and environmental changes monitoring). In DCS scheme, data is stored at nodes within the network by name. There are several drawbacks in the existing schemes. The first is the inefficiency of the range query processing on not considered the locality of store point. the second is the non-homogeneity of store load of each sensors in case of the sensor distribution density is non-uniformed. In this paper, we propose a novel data centric storage scheme with the sensor distribution density which satisfied with the locality of data store location. This scheme divides whole sensor network area using grid and distributes the density bit map witch consist of the sensor density information of each cell. sensors use the density bit map for storing and searching the data. We evaluate our scheme with existing schemes. As a result, we show improved load balancing and more efficient range query processing than existing schemes in environment which sensors are distributed non-uniform.

I. 서론

오늘날 군사시스템, 위치 탐색, 환경 감시 등의 다양한 분야에서 무선 센서 네트워크의 활용에 대한 연구가 진행되고 있다. 일반적으로 센서 네트워크를 활용한 서비스는 센서에서 수집한 데이터를 가공 처리하여 제공하는 서비스이다. 따라서 센서에서 수집된 데이터의 효율적인 저장, 관리는 데이터 기반 서비스의 효율적인 운용과 직결된다. 데이터 중심 저장기법은 센서의 하드웨어적 한계인 제한적인 전력량을 극복하기 위한

에너지 효율적인 센싱 데이터 저장 및 질의 처리를 제공하기 위한 기법이다.

데이터 저장 기법은 데이터 저장 위치에 따라 크게 외부 저장 방식(External Storage: ES), 내부 저장 방식(Local Storage: LS), 데이터 중심 저장 방식(Data Centric Storage: DCS)[6]으로 구분할 수 있다. ES 방식은 발생한 센싱 데이터를 싱크노드에 저장하는 방식이다. 이 방식은 센싱 데이터를 매번 싱크노드로 전송해야 하므로 라우팅하는 비용이 크게 증가한다. LS 방식은 센싱 데이터를 수집한 노드 자체의 저장 공간에 저장하는 방식이다. 이 방식은 원하는 데이터의 저장

이 논문은 2006년도 정부(과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임(No. R01-2006-000-1080900)

위치를 특정화 할 수 없으므로 질의 처리를 위해서는 질의를 모든 노드들에게 전송해야 질의 결과를 얻을 수 있다. 따라서 질의를 노드들에게 라우팅하는 비용이 크게 증가한다. DCS 방식은 데이터 값의 범위마다 저장을 담당할 센서를 할당하고, 센싱 데이터가 발생하면 할당된 센서에 데이터를 저장하는 방식이다. 이 방식은 질의 처리에 필요한 데이터가 저장되어 있는 위치를 알 수 있어 질의를 해당 위치로 정확히 라우팅할 수 있다. 이로 인해 질의의 불필요한 확산을 막을 수 있다.

기존에 제안된 대표적인 데이터 중심 저장방식으로는 GHT, Q-NiGHT, DIM, KDDCS 등이 있다. Q-NiGHT 기법[5]은 GHT 기법[4]의 공간 해싱 기법을 이용하여 데이터가 저장될 지리적인 위치를 선정하고, 데이터의 저장, 탐색, 질의를 위해 해당 위치로 패킷을 GPSR 기법[3]으로 최단 경로 라우팅하는 데이터 중심 저장 기법이다. Q-NiGHT 기법은 모든 공간을 균일하게 해싱하는 GHT 기법과 달리 센서의 배포 밀도를 고려한 불균일 공간 해싱(Non-uniform Hashing)을 적용하여 해싱되는 위치가 센서의 배포밀도가 높은 지역에는 많이 해싱되고, 낮은 지역은 적게 해싱되도록 하였다.

GHT와 Q-NiGHT에서 사용하는 공간 해싱은 해싱의 특성에 따라 랜덤한 위치를 생성한다. 따라서 인접한 데이터 범위의 값이라 할지라도 해싱된 저장 위치는 지리적으로 연관성이 없는 위치에 저장한다. 이 때문에 범위질의 처리를 위해서는 센서 네트워크 내의 흩어진 여러 노드에게 부분질의를 전송해야 하므로 많은 라우팅 비용을 소모한다.

KDDCS 기법은 센서가 배포된 지리적 영역에 각 차원별로 데이터 범위를 할당하고, 센서의 배포 위치를 기반으로 K-D 트리 공간 분할을 통해 하나의 센서 노드가 하나의 분할영역에 남게 될 때까지 분할한다. 이렇게 분할된 공간은 전체 데이터 범위에서 차지하는 위치와 크기만큼을 각 공간의 센서에 저장할 데이터 범위로 할당하는 기법이다. 이 기법은 인접한 데이터 범위를 지리적으로 인접한 위치에 할당함으로써 범위질의를 부분질의로 분할하지 않고 처리 가능하다. KDDCS 기법은 DIM 기법의 핫스팟 발생에 따른 노드별 부하 불균등 문제를 데이터 저장 범위를 결정하는 분할 축을 동적으로 이동시킴으로써 어느 정도 해결하였다. 그러나 KDDCS 기법은 기본적으로 센서의 배포밀도를 고려하지 않고 단순히 배포된 지리적 위치를 기반으로 공간을 분할하기 때문에 센서가 불균일하게 배포된 경우 분할 축 이동만으로 해결할 수 없을 정도로(분할 축의 이동범위는 인접한 센서 노드들의 범위 이내로 제한됨) 분할된 데이터 범위간의 불균등이 심화된다.

KDDCS에서의 가장 치명적인 약점은 라우팅 기법에 있다. KDDCS에서의 라우팅은 분할 축이 동적으로 이동함에 따라 원하는 데이터의 저장을 담당하고 있는 위치를 정확히 알아낼 수 없으므로 K-D 트리를 기반으로 이루어져야 한다. 상위

노드로 갈수록 라우팅 집중도가 높아지게 되어, 상위 노드의 통신 에너지 소모가 커지게 되고, 이는 네트워크의 수명을 단축시키는 결과를 야기한다.

본 논문은 앞에서 지적한 Q-NiGHT의 비효율적인 범위질의 처리 문제와 KDDCS의 배포밀도를 고려하지 않은 저장위치 선정 문제를 해결하기 위해 인접한 데이터 범위들 간의 지역성을 만족시키고, 균일한 배포밀도를 가지는 환경뿐만 아니라 불균일한 배포밀도를 가지는 센서 네트워크 환경에서도 저장위치를 모든 센서에 고르게 지정하는 방법을 제안한다.

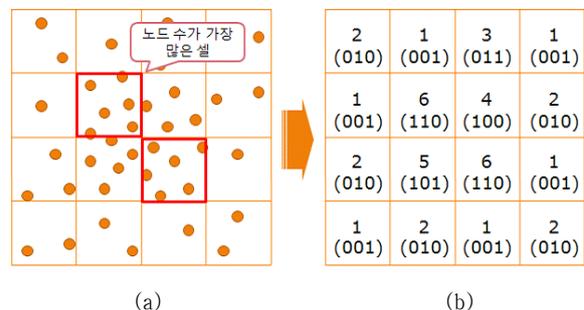
본 논문의 구성은 2장에서 본 논문에서 제안하는 데이터 중심 저장 기법에 대해서 기술하고, 3장에서는 시뮬레이션을 통해 제안된 기법의 성능을 평가한다. 마지막으로 4장에서 결과와 향후 연구의 방향을 제시한다.

II. 제안하는 데이터 중심 저장기법

본 논문은 제안하는 기법의 밀도 비트맵을 이용하여 센서 네트워크의 배포밀도와 인접한 데이터 범위의 저장 위치들 간의 지역성을 고려한 데이터 중심 저장 기법을 제안한다.

1. 밀도 비트맵

기존에 제안된 KDDCS 기법은 센서의 지리적 위치에 따라 각 센서가 저장할 데이터 범위가 결정된다. 따라서 센서의 배포밀도에 따라 데이터 범위의 분할과 할당이 불균등해질 수 있게 되고, 이는 일부 노드에 과도한 저장 부하를 야기하여 전체 네트워크 수명에 직접적으로 악영향을 끼치게 된다. 본 논문에서는 밀도 비트맵을 이용하여 저장 위치를 선정하는 방법을 제안한다.



▶▶ 그림 1. 센서 수 40, 분산 3.0 인 센서 네트워크의 4x4 센서 밀도 비트맵

밀도 비트맵은 전체 센서 네트워크의 배포밀도를 적은 용량의 비트를 이용해 나타낸다. 그림 1(a)와 같이 센서가 배포된 공간을 그리드로 분할하고, 분할된 각 셀에 포함된 센서의 개수를 비트로 표현하여 그림 1(b)와 같은 비트맵을 만든다. 각

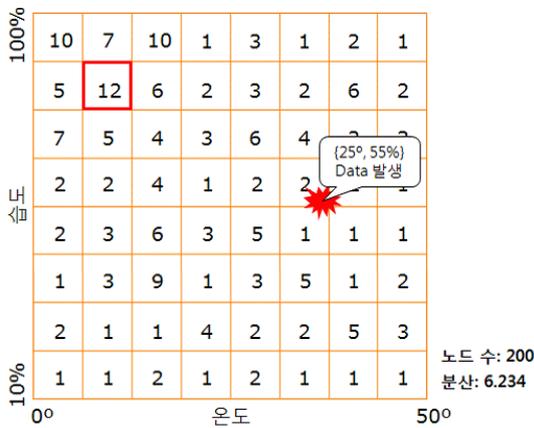
셀에 할당되는 비트 수는 모든 셀 중에 가장 많은 센서를 포함하는 셀의 센서 수를 표현할 수 있는 최소 비트 수이다. 그림 1(a)에서 가장 많은 센서를 포함하는 셀의 센서 개수는 6개이며, 6을 나타낼 수 있는 최소의 비트 수는 3비트 ($6 < 2^3$) 이다.

밀도 비트맵의 전송 부하는 무시 할 수 있을 정도로 작으며, 센서들을 배포할 때 단 한번 모든 센서들에게 전송되며 네트워크의 토폴로지가 바뀌지 않는 한 재전송하지 않는다. 그림 1(a)에서 예를 든 센서 네트워크(분산: 3, 센서 수: 40, 그리드 해상도: 4x4)의 분포 밀도를 나타내는 비트맵의 총 용량은 6바이트 이다.

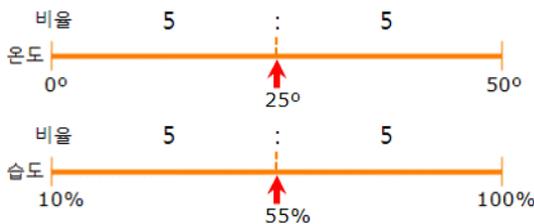
2. 밀도를 고려한 저장위치 선정

본 논문에서 제안하는 데이터 중심 저장 기법은 센서 네트워크의 밀도 비트맵을 이용하여 데이터 저장 위치를 선정한다. 그림 2는 앞서 설명한 방식을 이용해 전체 수가 200개이고 분산이 6.23인 센서 네트워크의 밀도 비트맵을 나타낸 것이다.

그림 2의 공간에 {온도, 습도} 2차원 데이터를 저장한다고 가정하였다(가로 차원: 온도의 범위 매핑, 세로 차원: 습도의 범위 매핑). 센서가 배포된 영역에서 센싱 데이터 {25°, 55%}가 발생했을 때 저장할 좌표는 다음의 순서로 계산한다.

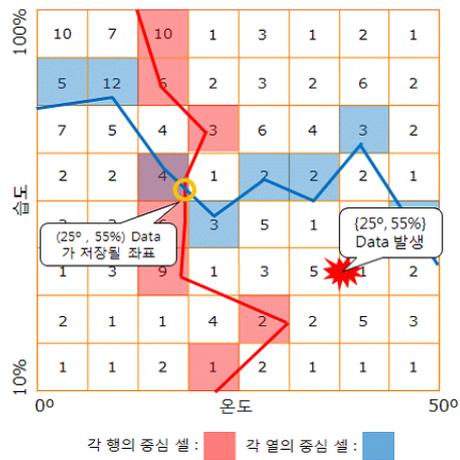


▶▶ 그림 2. 저장할 데이터{20°, 55%}의 발생

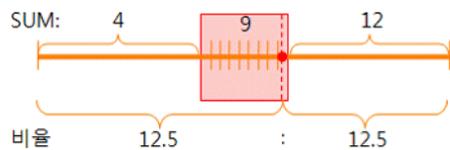


▶▶ 그림 3. 저장할 데이터{20°, 55%}의 각 차원 별 비율

1. {온도 25°, 습도 55%} 데이터 발생
2. 각 차원의 전체 범위에서 저장할 데이터의 비율 계산한다.
<ex. 그림 3, 온도 5:5, 습도 5:5>
3. 그림 4와 같이 밀도 비트맵의 각 행과 열마다 5:5로 센서의 개수를 양분 하는 셀을 선정한다.
4. 그림 5와 같이 각 열과 행의 양분하는 셀의 세부 분할 지점을 선정한다.
5. 각 행과 열마다 세부 분할 지점들의 연결하여 차원별 분할 선을 만든다.
6. 두 차원별 분할 선의 교차 지점 좌표를 데이터를 저장할 위치의 좌표로 선정한다.



▶▶ 그림 4. 센서 배포 밀도를 고려한 좌표 선정



▶▶ 그림 5. 분할 셀의 세부 분할 지점 선정

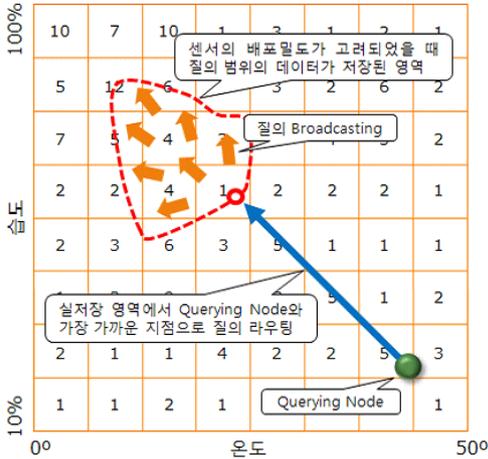
3. 질의 처리

본 논문에서 제안하는 데이터 중심 저장 기법의 동작은 센싱 데이터를 저장하기 위한 put 메소드, 저장되어있는 데이터를 탐색하기 위한 get 메소드 그리고 범위 질의를 처리하기 위한 range 메소드를 통해 이루어진다.

put 메소드가 수행되면 로컬 메모리상의 센서 배포 밀도 비트맵과 센싱된 값을 이용하여 저장할 위치 좌표<x,y>를 알아내고, 수집한 데이터 d와 해당 센서의 id를 합쳐 패킷 <id, <x,y>,d>를 생성하고, GPSR 기법을 이용해 <x,y>위치로 라우팅한다.

get 메소드가 수행되면 로컬 메모리상의 센서 배포 밀도 비트맵과 찾고자 하는 값을 이용하여 탐색할 위치 좌표<x,y>를

알아내고, 질의 노드의 위치 좌표 $\langle r, z \rangle$ 를 합쳐 패킷 $\langle \langle x, y \rangle, \langle r, z \rangle \rangle$ 를 생성한다. 이후 $\langle x, y \rangle$ 위치로 질의를 라우팅 하고, 데이터를 저장하고 있는 노드가 질의를 수신하게 되면 질의 결과를 $\langle r, z \rangle$ 위치로 라우팅한다.



▶▶ 그림 6. range 메소드를 이용한 범위질의 처리

range 메소드가 수행되면 그림 6과 같이 질의 범위에 해당하는 데이터가 저장된 영역을 계산한다. 그리고 질의를 내리는 센서에서 저장 영역내의 가장 가까운 지점의 좌표인 $\langle x, y \rangle$ 를 계산한다. 해당 좌표로 질의를 GPSR 기법을 이용하여 라우팅한다. 저장영역내의 특정 센서 노드가 질의를 수신하면 해당 질의를 저장 영역내의 센서들에게 브로드캐스트한다. 이때, 질의를 수신한 센서들은 자신의 데이터 저장 범위가 질의 범위내의 포함여부를 판단하여 만족하면 수신된 질의를 이웃 센서들로 재배포하고, 만족하지 않으면 수신된 질의를 무시한다. 이를 통해 범위를 만족하는 모든 센서 노드들에게 질의를 전달할 수 있으며, 범위 외의 센서 노드들의 불필요한 통신 에너지 소모를 막을 수 있다. 범위질을 영역내의 모든 센서들이 수신하게 되면 질의 결과에 대해 데이터 병합(agggregation)하여 질의를 내린 센서로 라우팅한다.

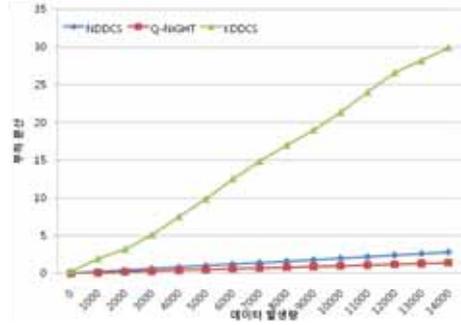
III. 성능 평가

성능평가는 가상의 환경을 구축하여 시뮬레이션 하였으며, 시뮬레이션 환경은 다음의 표 1과 같다.

[표 1] 시뮬레이션 환경

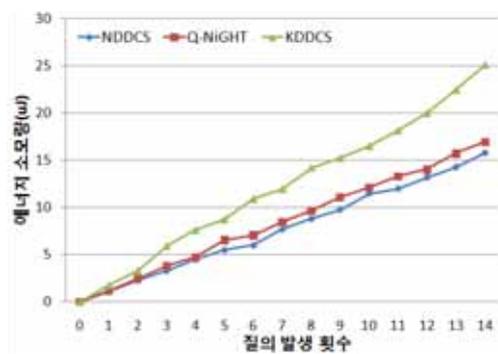
센서 노드의 수	2000개
센서노드의 통신반경	10m
필드 크기	500m×500m
센서 배포 환경	랜덤 가우시안 분포
비트맵 해상도	10×10

제안하는 데이터 중심 저장 기법의 우수성을 보이기 위해서 대표적인 데이터 중심 저장 기법인 Q-NiGHT기법, KDDCS 기법과 비교하여 각 노드별 저장 부하의 편차와 질의 처리 시 발생하는 전체 노드의 에너지소모량을 비교하였다.



▶▶ 그림 7. 데이터의 저장 부하 분산

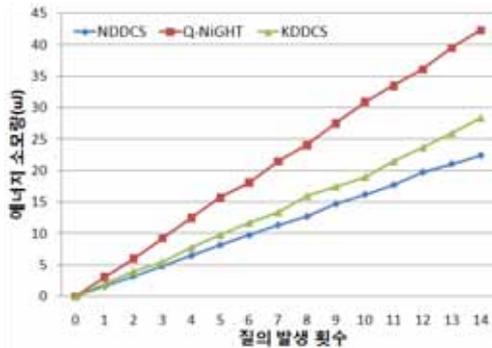
그림 7은 전체 데이터 범위에 걸쳐 저장할 센싱 데이터가 크게 발생할 때 센서 노드들의 저장 부하의 분산을 나타낸다. 제안하는 데이터 중심 저장 기법의 경우, 데이터가 계속적으로 발생하더라도 부하분산이 거의 일정하게 나타난다. 이에 비해 KDDCS 기법의 경우 저장된 데이터가 늘어남에 따라 부하 분산이 점점 커진다. 그 이유는 데이터 저장 범위를 분할할 때 센서의 절대적인 공간적 위치를 기반으로 나뉘지므로 밀도가 불균일할 경우 배포 밀도가 낮은 곳에 위치한 센서 노드의 데이터 저장 범위가 배포밀도가 높은 곳에 위치한 센서 노드의 데이터 저장 범위에 비해 넓다. 제안하는 데이터 중심 저장 기법과 Q-NiGHT의 경우, 센서 노드들의 배포밀도를 고려하여 저장을 담당할 센서 노드를 선정하므로 배포밀도의 영향을 받지 않고 전 영역의 센서 노드들에게 고르게 저장부하를 할당한다.



▶▶ 그림 8. 데이터 저장에 따른 에너지 소모

그림 8은 센싱 데이터를 저장하고, 탐색하는데 소모되는 에너지양을 나타낸다. 제안하는 데이터 중심 저장 기법과 Q-NiGHT의 경우, GPSR 기법을 이용해 최단 경로로 라우팅하므로 에너지 효율적이다. KDDCS기법의 경우 분할축의 이

중에 따라 저장될 위치를 계산해 낼 수 없기 때문에 K-D 트리 기반의 라우팅을 수행하여야 하므로 배포 밀도가 불균일할 경우 트리 또한 불균일하게 구축되어 최단 라우팅 경로를 벗어나게 된다. 따라서 상대적으로 많은 노드를 거쳐 라우팅되므로 통신비용이 커진다.



▶▶ 그림 9. 범위질의에 따른 에너지 소모

그림 9는 저장된 데이터에 대한 범위질의에 따라 소모되는 에너지양을 나타낸다. 제안하는 데이터 중심 저장 기법의 경우, 지역성이 고려되어 데이터들이 저장되어 있으므로 범위질의를 분할하지 않고 라우팅 가능하다. 따라서 하나의 범위질의 패킷만이 라우팅되므로 범위질의에 따른 에너지 소모를 줄일 수 있다. Q-NIGHT의 경우 지역성이 고려되지 않아 범위질의를 처리하기 위해서는 범위를 여러 부분 범위로 분할하여 라우팅한다. 이로 인해 하나의 범위질을 처리하기 위해 여러 다른 경로로 라우팅이 일어나므로 통신비용이 증가한다.

IV. 결 론

본 논문에서는 센서 네트워크의 배포밀도와 인접한 데이터 범위들 간의 지역성을 만족시키는 데이터 중심 저장 기법을 제안했다. 기존의 데이터 중심 저장 기법들의 경우 센서네트워크의 배포밀도와 인접한 데이터 범위들 간의 지역성 두 가지를 모두 만족시키지 않아 비효율적인 범위질의 처리나 일부 센서에 핫스팟을 발생시켜 네트워크의 수명을 단축시키는 문제점을 가지고 있었다.

본 논문은 밀도 비트맵을 바탕으로 센서 네트워크의 배포밀도가 고려하여 데이터를 저장하는 센서 노드를 선정한다. 따라서 센서가 균일하게 배포된 상황은 물론 불균일하게 배포된 상황에서도 센서 네트워크 모든 노드의 저장 부하를 균등하게 할당하여 일부 이용되지 않는 센서 노드의 저장 공간을 전체적으로 고루 사용할 수 있게 한다. 그리고 지역성을 고려한 데이터 저장을 통해 에너지 효율적인 범위질의 처리를 가능케 한다.

■ 참고 문헌 ■

- [1] X. Li, Y. J. Kim, R. Fovidan, and W. Hong, "Multi-dimensional range queries in sensor networks", *SenSys'03*, 2003.
- [2] M. Aly, K. Pruhs, P. K. Chrysanthis, "KDDCS: a load-balanced in-network data-centric storage scheme for sensor networks", *CIKM'06*, pp. 317-326 2006.
- [3] B. Karp, H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless sensor networks", *ACM Mobicom*, 2000.
- [4] S. Ratnasamy, B. Karp, S. Shenker, D. Estrin, R. Govindan, L. Yin, F. Yu, "Data-centric storage in sensornets with GHT, a geographic hash table", *MONET* vol. 8, no. 4, 2003.
- [5] M. Albano, S. Chessa, F. Nidito, S. Pelagatti, "Q-NIGHT: Adding QoS to Data Centric Storage in Non-Uniform Sensor Networks", *Mobihoc 2006*, vol. 1, pp. 1-3, 2006.
- [6] M. Aly, P. K. Chrysanthis, and K. Pruthis, "Decomposing data-centric storage query hot-spots in sensor networks", *MobiQuitous*, 2006.