Regular paper http://dx.doi.org/10.13067/JKIECS.2015.10.7.769

사물인터넷을 위한 센서 네트워크에서 효율적인 데이터 분산 기법

김정원*

An Effective Data Distribution Scheme in Sensor Network for Internet of Things

Jeong-Won Kim*

요 약

사물인터넷의 하부 구조로 센서네트워크는 네트워크의 신뢰성이 중요한 이슈이다. 센서네트워크의 각 노드는 제한된 메모리와 파워로 인해 데이터의 신뢰성을 향상시켜야 한다. 본 연구에서는 센서 노드가 생성한 데이터를 이웃 노드에 복제하는 기법을 제안한다. 네트워크의 신뢰성을 향상시키기 위해서 주변 노드의 메모리 및 파워 잔량이 높은 노드에 사본을 저장하는데 1 홉 이상의 노드들에도 저장하여 공간적인 재해에도 대비한다. 실험을 통해 적절한 사본을 유지하는 것은 네트워크의 수명에 부정적인 영향을 미치지 않으면서도 사물인터넷의 신뢰성을 향상시킴을 확인하였다.

ABSTRACT

Sensor network as an infrastructure of IoT(Internet of Things) has reliability issue because sensor nodes have limited memory as well as bounded battery. To improve the reliability of network, this paper proposes a data distribution scheme. The proposed algorithm distributes the data which each sensor node periodically produces into neighbor nodes that have enough memory as well as battery. This distribution process goes on more than 1 hop for overcoming unexpected spatial crash. Through simulation, we have confirmed that the proposed scheme can improve the resilience of IoT without affecting the life time of sensor network.

키워드

IoT, Sensor Network, Replication, Distribution
사물인터넷, 센서네트워크,데이터 복제, 데이터 분산

Ⅰ. 서 론

사물인터넷의 개념은 인간이 접하는 모든 사물에 IP를 부여하고 사물을 상호 연결하여 완전한 제어 및 부가적인 정보 창출을 가능하게 하는 것이다. 이들 사물은 RFID(Radio-Frequency Identification), M2M (Machine to Machine), 그리고 WSN(Wireless Sensor Network) 으로 구성되는데 스마트 그리드, 스마트 워크 등 다양한 애플리케이션으로 응용된다. 이와 관련된 다수의 연구들이 제안되고 있으나 표준 아키텍처나 IPv6 기반 프로토콜 등 해결해야할 많은 문제들이 남아 있는 실정이다[1-3].

사물인터넷을 위한 WSN는 주위 환경 데이터를 수집하는 센서 노드와 수집된 데이터를 서버로 전송하는 싱크 노드로 구성될 수 있는데 센서 노드와 싱크 노드사이의 통신은 즉각적으로 발생될 수 없으며 특

* 교신저자 (corresponding author) : 신라대학교 컴퓨터정보공학부(jwkim@silla.ac.kr)
접수일자 : 2015. 06. 10 심사(수정)일자 : 2015. 07. 13 게재확정일자 : 2015. 07. 23

히 독립된 센서네트워크에서는 싱크 노드의 신뢰성 또한 보장할 수 없다. 만약 애플리케이션이 데이터를 실시간으로 요구하지 않는 경우에는 수집된 데이터의 전송에 필요한 에너지를 감소시킬 수 있고 이는 네트워크의 수명을 연장시킬 수 있다. 이 경우 모바일 단말기를 통해 데이터를 주기적으로 수집할 수 도 있으나 부정기적인 데이터 수집으로 인해 센서노드의 메모리는 오버플로우가 발생할 수 있고 이것은 정보의 손실을 야기할 수도 있다. 따라서 데이터의 분산이 요구된다.

사물인터넷의 하부 기술로 WSN는 장시간 오프라인상태로 있을 가능성이 있으므로 각 노드의 수명 및 저장된 데이터의 신뢰성 보장이 중요하다. 이러한 점을 해결하기 위해서는 데이터의 분산 및 복제가 요구되는데 각 센서가 수집한 데이터를 주변의 노드에 사본을 만드는 것이다. 한편 이러한 과정은 저장 대상 노드를 찾는데 통신 오버헤드 및 저장 공간의 증가가 필연적으로 요구된다. 즉 시스템의 신뢰성 향상 측면에서의 데이터 분산 및 복제와 네트워크의 수명은 상호 반작용의 관계이다. 신뢰성이 향상되면 노드의 배터리 소모량이 증가하고 분산하지 않으면 신뢰성은 낮아지는 것이다.

본 연구에서는 사물인터넷을 위한 WSN에서 이 두 가지 요소를 모두 만족시킬 수 있는 기법을 제안하고자 한다. 신뢰성 향상을 위해서 각 노드는 생산한 데이터를 이웃 노드에 복제하는데 이웃 노드의 거리, 통신 오버헤드, 노드의 메모리 잔량, 노드의 에너지 잔량을 고려한다. 또한 네트워크 전체의 에너지 잔량과 저장 공간의 잔량을 고려하여 특정 범위를 초과하면 가장 오래된 데이터를 삭제하여 데이터의 신뢰성을 제고한다.

본 논문의 구성은 다음과 같다. 2장에서는 WSN에서 데이터 분산 및 복제에 대한 기존 연구를 분석하고 본 논문의 기여를 설명한다. 3장에서는 제안하는 분산 및 복제 기법에 대한 모델을 설명한다. 4장에서는 본 연구에서 제안한 모델을 검증하기 위해 구축한 시뮬레이션 환경과 실험 결과를 설명한다. 마지막으로 5장에서는 본 연구의 결론을 제시한다.

II. 관련 연구

지난 수년간 WSN에서 데이터의 분산 및 복제에 관한 많은 연구들이 제안되었는데 노드들은 데이터를 효율적으로 분산하기 위해 상호 협력한다. 이 방법은 크게 데이터 중심 기법과 완전 분산 기법으로 분류될 수 있다. 먼저 데이터 중심 기법은 해시 함수 등에 의해 결정된 특수 노드가 데이터를 전적으로 수집한다. 특수 노드에 집중되는 부하를 분산하기 위해서 밀도가 높은 지역의 데이터 손실을 최소화하기 위한 기법과 질의에 대한 지연을 최소화하기 위하여 시·공간적으로 유사도가 높은 데이터는 동일 공간에 저장하는 기법들이 제안되었다[4-5]. 이 기법들은 노드간의 협력에 기반을 두므로 특정 노드에 다른 노드들이 생산한 데이터가 집중되므로 데이터 분산이 미흡한 것으로 판단된다[6].

완전분산기법에서는 모든 노드가 데이터 수집 및 저장을 수행하며 각 노드들은 우선 자신의 저장공간에 저장을 시도하고 지역 메모리가 부족하면 이웃 노드에게 위임하는 기법이다. Data Farm[7]에서는 주기적인 데이터 수집 및 분산을 수행하고 저장 노드는 에너지 소비 모델에 의해 결정되는 기법을 제안하였다.

데이터 복제 관련 연구들도 다수 제안되었는데 주로 노드 실패에 대비하기 위한 것이다. ProFlex는 에너지가 부족한 노드대신 에너지가 많이 남아 있는 노드가 데이터 수집 및 복제를 전담하여 노드의 실패를 완화시키고자 한다. TinyDSM에서는 복제의 수와 밀도에 기반하는 모델에 의해 복사본을 노드에 임의적으로 배치하며 연결성, 에너지 잔량, 메모리 잔량의 요소에 기반하여 복제 노드를 선택하기도 한다[8].

상기의 연구들은 데이터 분산과 복제를 별개로 취급하는 경향이 많은 반면 본 연구는 두 가지 요소를 동시에 고려한 기법으로서 분명히 차이가 존재한다. 제안하는 모델은 네트워크의 전체 에너지 및 저장공간이 고갈되는 시점을 고려하여 데이터를 분산하고 복제하며 각 노드에서 복사본을 저장할 노드의 거리까지 고려하여 네트워크의 신뢰성을 향상시키고자 한다.

III. WSN의 신뢰성 향상을 위한 데이터 분산 기법

WSN의 신뢰성 향상을 위한 데이터 분산 기법을 제안하기 전에 네트워크의 동작에 대하여 먼저 설명한다. 각 센서 노드가 수집한 데이터는 주기적으로 싱크노드로 전달되고 전송된 데이터는 최신의 데이터만 유지하고 나머지는 삭제한다. 이 주기적인 데이터 전달은 각 노드의 메모리와 배터리 사용의 효율성을 위해 필연적이다. 본 연구는 WSN의 신뢰성 향상이 목적이므로 데이터 수집보다는 데이터 복제 및 분산에 초점을 두는데 데이터의 분산은 다음과 같은 패턴을 따른다. 센서 노드가 생산한 데이터는 이웃 노드에 사본을 저장하는데 후보노드의 선택과 사본의 수를 결정해야 한다. 사본의 수를 증가시키면 신뢰성을 향상시킬 수 있지만 메모리 및 통신 오버헤드로 인해 네트워크의 수명을 단축시킬 수 있으므로 적절한 수를 결정해야 한다. 후보노드를 선택하는 방법은 메모리 잔량과 통신비용, 그리고 배터리 잔량을 고려한다.

각 노드는 주기적으로 자신의 메모리 및 배터리 잔량을 이웃노드들에게 방송하는데 각 노드는 이 정보를 테이블에 유지하여 후보노드를 선택할 때 판단 기준으로 이용한다. 그리고 각 노드의 방송 주기는 데이터 생산 주기와 맞추는데 이것은 데이터가 생산되면 메모리 및 배터리 잔량이 변화하기 때문이다. 일단 후보노드가 선택되면 후보노드로 사본을 보내고 후보노드는 자신의 후보노드를 다시 선택하여 사본을 전송하는데 이 횟수는 사본의 수에 의해 결정된다. 만약 이 전송의 매 단계에서 후보노드가 존재하지 않으면 전송을 중단한다. 생산 노드에서 후보노드가 없다면 생산 노드의 메모리에만 저장하고 사본은 저장되지 못한다.

표 1은 본 연구의 WSN에서 사용하는 주요 기호들이다. N 개의 노드가 임의의 지역에 배치되어 있고 각 센서는 d 의 통신 반경을 가진다. 이 반경 내에서 각 센서 노드는 1 홉 거리 내에 있는 이웃 노드들이 SC_i 개 존재한다. 각 노드는 T_i 주기 마다 데이터를 수집하여 싱크노드에게 전송하는데 이때 메모리 잔량 m_i, 배터리 잔량 p_i 를 이웃 노드에게 전송한다. 데이터가 생산되면 자신의 메모리에 저장하고 선택된 후보 노드 c_i 에게 사본을 전송하고 노드 c_i 는 다시 자신의 후보노드에게 사본을 전송하는데 이 과정은 R 번 반복된다.

표 1. WSN의 주요 기호
Table 1. Symbols for WSN

Symbol	Description
N	Total number of sensor node
d	Transmission range
SC_i	Number of 1-hop neighbors
m_i	Remaining memory
p_i	Remaining battery
c_i	Candidate node
T_i	Period of data transmission
T_{br}	Period of advertisement
R	Maximum number of replication
t_i	time of each sensor node

각 노드는 이웃 노드로부터 받은 최신의 메모리, 배터리 상태를 자신의 메모리에 유지하는데 이 테이블에는 이웃 노드당 한 개의 행을 가지며 주기 T_{br} 마다 정보를 갱신하게 된다. 또한 순서를 유지하여 중복되거나 이전 방송은 무시한다. 센서노드의 메모리는 고정되어 있으므로 네트워크가 밀집한 지역에서는 이 테이블이 순식간에 고갈되므로 최적의 이웃들만 유지한다.

각 센서노드는 데이터를 생산하면 R 개의 사본을 유지하는데 자신의 지역 메모리에 유지하고 R-1 개의 사본을 후보 노드에게 전송해야 한다. 노드의 특정 시간 t_i 에 유효한 데이터를 생산하였다고 가정하면 1 홉 거리에 있는 이웃노드 SC_i 중 최적의 노드, 즉 후보노드 c_i 에게 사본을 전송한다. 이때 후보 노드를 선택하는 기준은 이웃 노드의 메모리 잔량과 배터리 잔량이 모두 우수한 노드이다. 다음 수식 (1) 은 최적의 후보노드를 선택하는 규칙이다.

$$c_i(t_i) = \max \frac{m_j(t_j) \times p_j(t_j)}{t_j - t_i}, (j \in SC_i) \quad (1)$$

센서 노드 i 의 특정 시간 t_i 에서의 후보노드는 자신의 메모리 테이블에 유지하고 있는 이웃 노드의 집

합 SC_i에서 메모리 잔량과 배터리 잔량이 가장 높은 노드가 선택된다. 이것은 메모리 잔량이 높은 노드가 사본을 저장해야 하고, 메모리 잔량은 높은데 배터리가 부족하면 데이터 손실이 발생할 가능성이 있으므로 배터리 잔량 또한 높은 노드를 선택한다. 수식 (1)에서 이웃 노드가 보낸 시간이 데이터를 생산한 노드의 시간보다 최근인 경우($t_j > t_i$)의 노드만을 대상으로 후보노드를 선택한다. 만약 이웃 노드의 시간이 생산 노드의 시간보다 이전 이면($t_j < t_i$) 음수의 값이 발생하여 후보 대상에서 제외된다.

다음 그림은 R 이 3인 경우의 데이터 분산 과정의 예이다. 각 노드의 메모리 테이블에는 {노드 id, 메모리잔량, 배터리잔량, 생산시간} 정보를 유지하고 있다.

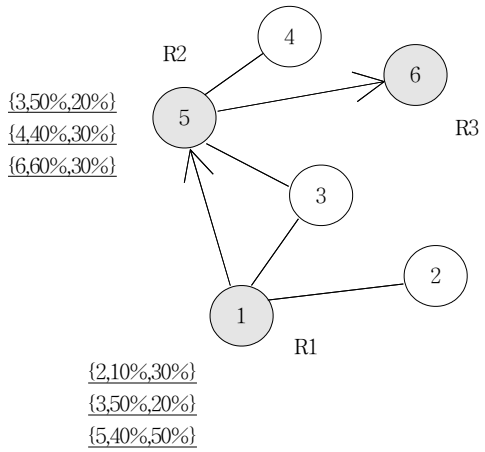

그림 1. 2-홉 데이터 분산
Fig. 1 2-hop data distribution

그림 1의 경우 노드 1 이 데이터를 생산하여 사본 1(R1)을 자신의 메모리에 저장하고 이웃 노드의 메모리와 배터리 용량 테이블을 검사한 결가 노드 3 이 50%로 메모리는 가장 여유가 있지만 수식 (1) 에 의하면 배터리 잔량을 고려한 결과 c_i 값이 가장 높은 노드는 노드 5 이다. 따라서 R2는 노드 5에 저장된다. 노드 5에서 다시 수식 (1)을 적용하면 노드 6의 c_i 값이 가장 높으므로 R3은 노드 6에 저장된다. 이 경우 마지막 사본이 노드 1에서 2 홉 거리에 저장된다.

그림 2는 노드 2에서 R1을 저장하고 노드 6에서 R2를 저장한다. 노드 6에서 이웃 노드는 노드 2, 3, 4 인데 c_i 값이 가장 높은 노드는 노드 3이다. 노드 3은 노드 2에서 1 홉이므로 R2, R3가 모드 노드 2에서 1 홉에 분산됨을 알 수 있다.

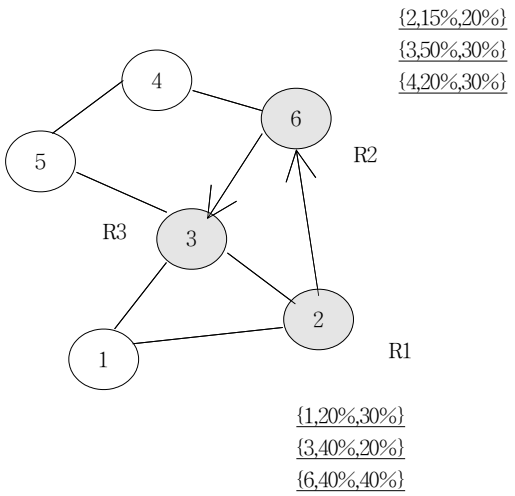

그림 2. 1-홉 데이터 분산
Fig. 2 1-hop data distribution

Ⅳ. 실험

본 논문에서 제안하는 기법의 효율성을 검증하기 위하여 시뮬레이션을 진행하였다. 실험의 환경은 가로 및 세로가 1 킬로미터인 지역에 1,000 개의 센서를 임의의 좌표에 배치하여 환경 값을 측정하는 것이다. 센서 모델은 MICA2로서 저전력 무선 센서네트워크의 3세대 모델이다[9]. MCU는 Atmel ATmega 128L기반의 MPR400CB로서 아날로그 및 디지털 입출력, I2C, SPI, UART 인터페이스를 지원한다.

또한 868/916MHz 의 다중 채널 라디오, 38.4Kbps 데이터 전송속도, 전송시 최대 27mA, 수신시 10mA, 그리고 대기모드일 경우 1uA 이하의 배터리를 소모한다. 데이터 저장 플래시 메모리는 512KB이다. 각 센서는 두 개의 AA 배터리가 장착되는데 25mA 소모시 각 배터리는 120시간의 수명을 제공할 수 있다. 본 실험에서 사용한 MICA2의 속성은 메모리, 송수신시 파

워소모량, 그리고 데이터 전송률이다.

상기의 실험환경에서 매트랩과 시뮬링크에서 실험을 진행하였다. 실험은 2000 가상시간에 각 센서 노드들은 데이터를 생산하여 주위의 노드에게 수식 (1)에 의해 사본을 전송하게 되는데 'replication 1' 은 자신의 노드에만 저장하고 'replication 2' 는 1개의 사본을, 'replication 5'는 네 개의 노드에 사본을 저장한다.

그림 3은 시간이 지남에 따른 사본의 수를 나타낸 그래프이다. 본 실험에서는 0에서 4개 까지의 사본을 이웃 노드에 저장하는 실험을 진행하였다. 'replication 1'의 경우는 이웃 노드에 저장하지 않으므로 배터리 소모량은 낮으며 시간이 지날수록 가장 완만한 메모리 소모량을 보인다. 반면 사본을 가장 많이 저장하는 'replication 5'의 경우 전체 시뮬레이션의 중간 시점에서 사본의 수가 최고점에 다다르고 있다. 이것은 다수의 사본을 저장하므로 선택된 이웃 노드가 더 이상 사본을 수용하지 않기 때문이다.

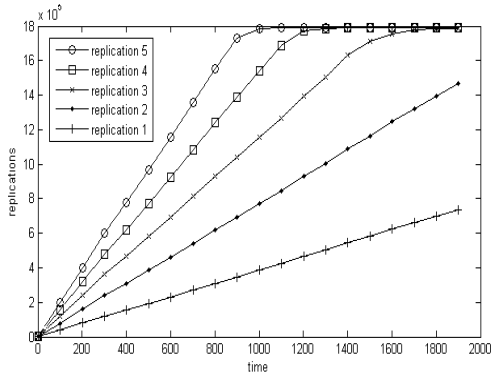

그림 3. 시간의 변화에 따른 사본의 수
Fig. 3 Replication number as functions of time

그림 4는 본 연구의 효율성을 검증하기 위해 지진, 홍수, 화재 등의 재해로 인해 특정 지역의 센서들이 동작하지 않을 경우 해당 노드에 대한 질의시 성공 가능한 회수를 측정하였다. 실험을 편의를 위해 재해 지역을 'dead zone'으로 명명하고 각각 100, 200, 300 미터의 영역을 임의로 설정하여 해당 지역에 포함된 노드의 특정 시점에서의 데이터에 대한 질의를 실험 영역의 노드들에게 할 경우 성공한 회수를 측정하였다. 당연히 'dead zone 300'의 경우 모든 기법에서 질의 성공회수가 낮은 것을 알 수 있으며 사본을 저장하지 않는 'replication 1'의 경우는 성공률이 0%이다. 따라서 전체 네트워크의 메모리와 파워를 크게 소모하지 않는 범위에서 사본의 저장은 네트워크의 신뢰성 향상에 긍정적인 요인을 작용함을 알 수 있다.

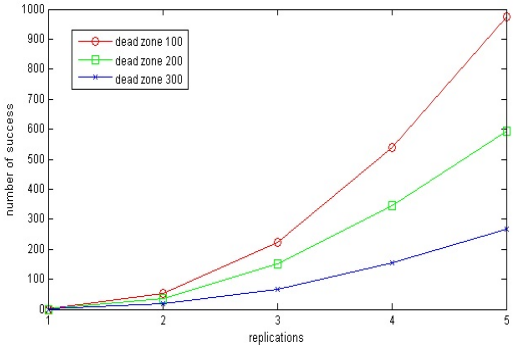

그림 4. 데드 존 발생시 질의 성공 회수
Fig. 4 Query successes at dead zone

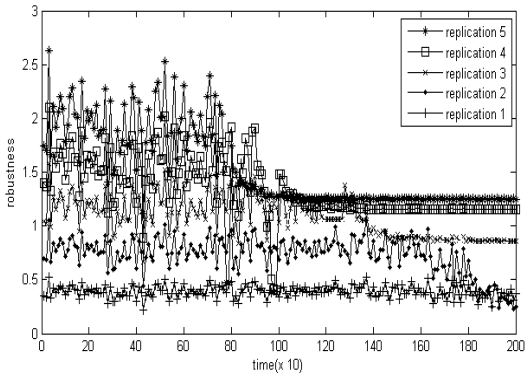

그림 5. 사본의 수에 따른 네트워크 안정성
Fig. 5 Network stability as functions of replicas

그림 5는 전체 시뮬레이션 시간 동안의 네트워크 신뢰도를 표현한 것으로서 특정 시각 t 에서 네트워크가 신뢰도는 시각 t 까지 각 센서가 생성한 데이터에 대하여 시작부터 시각 t 까지의 적어도 하나의 사본이 저장된 회수를 합한 비율이다. 이 값이 높다는 것은 'dead zone'의 외부에서 데이터를 발견할 확률이 높다는 것을 의미하므로 네트워크의 신뢰도는 높아 질 것이다. 그림 5에서 사본을 저장하지 않는 'replication 1'의 경우는 1에 미치지 못하며 사본을 가장 많이 저장

하는 replication 5' 는 신뢰도 값이 2를 보이고 있다. 신뢰도의 값은 초기에는 높은 값을 보이다가 중후반에는 다소 떨어지고 있다. 이것은 각 노드의 버퍼가 여유가 없거나 배터리 소모로 인해 노드가 사본을 받지 못하는 경우로서 네트워크가 포화 상태로 진입하고 있음을 의미한다.

상기의 실험의 결과 적절한 사본의 유지는 재해 발생시 네트워크의 신뢰성을 향상시킬 수 있음을 확인하였다. 또한 전체 네트워크의 생존 시간은 사본을 유지하지 않는 경우가 가장 높지만 질의 성공확률을 높이기 위해서는 2홉 이상의 노드에 사본을 유지해야 하며 사본의 개수는 전체 네트워크의 자원에 의해 결정되어야 한다.

V. 결 론

본 연구에서는 사물인터넷에서 각종 센서가 생성한 데이터의 신뢰성을 향상시키기 위한 데이터 복제 기법을 제안하였다. 센서들은 자신의 메모리뿐만 아니라 이웃 노드에 사본을 저장하는데 이웃 노도의 효율적인 선택 방법과 수를 결정하는 것이 사물인터넷의 신뢰성을 향상시킬 수 있다. 본 논문에서는 이웃 노드는 메모리 및 배터리 잔량을 기준으로 선택하고 2 홉 이상의 노드를 선택해야 전체 네트워크의 수명을 단축시키지 않고 신뢰성을 향상시킬 수 있음을 실험을 통해 확인하였다. 향후 연구로는 실제 센서 노드를 배치하여 실험과 실측을 비교하는 것이다.

References

[1] K. Nam, "A Study on the Office Management Service Platform based on M2M/IoT," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 9, no. 12. 2014, pp. 1405-1413.

[2] D. Ryu, "Development of BLE Sensor Module based on Open Source," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 10, no. 3. 2015, pp. 419-424.

[3] M. Kang, "Platform Design of Unity Launcher for the IoT Beacon based 3D Position," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 10, no. 4. 2015, pp. 477-482.

[4] M. Albano, S. Chessa, F. Nidito, and S. Pelagatti, "Dealing with nonuniformity in data centric storage for wireless sensor networks," *IEEE Trans. on Parallel and Distributed Systems*, vol. 22, no. 8, 2011, pp. 1398 - 1406.

[5] H. Shen, L. Zhao, and Z. Li, "A distributed spatial - temporal similarity data storage scheme in wireless sensor networks," *IEEE Trans. on Mobile Computing*, vol. 10, no. 7, 2011, pp 982 - 996.

[6] K. Ahmed, and M. A. Gregory, "Techniques and challenges of data centric storage scheme in wireless sensor network," *J. of Sensor and Actuator Networks*, vol. 1, no. 1, 2012, pp. 59 - 85.

[7] A. Omotayo, M. Hammad, and K. Barker, "A cost model for storing and retrieving data in wireless sensor networks," In *Proc. IEEE 23rd Int. Conf. on Data Engineering Workshop*, Istanbul, Turkey, 2007, pp. 29 - 38.

[8] K. Piotrowski, P. Langendoerfer, and S. Peter, "tinyDSM: A highly reliable cooperative data storage for wireless sensor networks," In *Int. Symp. on Collaborative Technologies and Systems*, Maryland, USA, 2009, pp. 225 - 232.

[9] http://www.eol.ucar.edu/isf/facilities/isa/internal/CrossBow/DataSheets/mica2.pdf

저자 소개

김정원(Jeong-won Kim)

1995년 부산대학교 전자계산학과 졸업(공학사)
1997년 부산대학교 대학원 전자계산학과 졸업(공학석사)
2000년 부산대학교 대학원 전자계산학과 졸업(공학박사)
2000년 기술보증기금 기술평가역
2002년 신라대학교 컴퓨터정보공학부 교수
※ 관심분야 : 사물인터넷, 임베디드시스템