

Design of A Faulty Data Recovery System based on Sensor Network

金星鎬[†] · 李榮森* · 陸宜洙**

(Sung-Ho Kim · Young-Sam Lee · Yui-Su Youk)

Abstract - Sensor networks are usually composed of tens or thousands of tiny devices with limited resources. Because of their limited resources, many researchers have studied on the energy management in the WSNs(Wireless Sensor Networks), especially taking into account communications efficiency. For effective data transmission and sensor fault detection in sensor network environment, a new remote monitoring system based on PCA(Principle Component Analysis) and AANN(Auto Associative Neural Network) is proposed. PCA and AANN have emerged as a useful tool for data compression and identification of abnormal data. Proposed system can be effectively applied to sensor network working in LEA2C(Low Energy Adaptive Connectionist Clustering) routing algorithms. To verify its applicability, some simulation studies on the data obtained from real WSNs are executed.

Key Words : Sensor network, PCA, AANN, , Data Compression, Sensor fault detection

1. 서 론

유비쿼터스 사회를 구현하기 위한 한 방향으로 최근 무선 센서네트워크가 대두되고 있다. 무선 통신 기술과 반도체 설계 기술의 발전으로 계측, 연산 및 무선 통신기능을 갖춘 저가격, 저전력의 다양한 센서 노드들이 개발되고 있다. 이러한 센서 노드들은 사람이 접근하기 어려운 작업공간에 살포되며 센서 노드들은 자체적으로 네트워크를 형성하여 센서 노드에서 측정된 내용을 자체적으로 형성된 망을 통해 기간망으로 전달함으로써 사용자들이 원격지에서 자연환경의 변화내용 혹은 감시 내용을 파악할 수 있는 센서 네트워크에 사용된다. 이러한 기능을 갖는 센서 네트워크는 환경, 생태 모니터링, 지진 감시 및 군사용 등에 폭넓게 도입되고 있다[1-2].

센서 네트워크는 무선 네트워크의 한 종류이나 QoS의 최적화를 위해 라우팅 및 mobility를 고려하는 종래의 셀룰러 네트워크 및 MANET 등과는 많은 차이점을 보인다. 기존의 ad-hoc 네트워크는 노드의 mobility가 존재하는 경우에도 효율적으로 자체 네트워크 구성 능력을 갖도록 하는 프로토콜의 개발이 중요한 반면 센서 네트워크에서는 제한된 전력으로도 네트워크의 생존 시간을 최대화시킬 수 있는 기법의 개발이 보다 중요하다. 따라서 최근에 센서 네트워크 분야의 많은 연구자들은 전력소모를 최소화할 수 있는 다양한

형태의 라우팅 프로토콜 및 MAC 프로토콜 개발에 관한 연구를 진행하고 있다[3-5].

측정 현장의 센서 노드들은 여러 가지 요소(노드의 배터리 고갈, 사고로 인한 보드의 파손, 센서의 노화)로 문제가 발생하게 되며 이는 정상적으로 동작하고 있는 무선 센서네트워크의 운용에 악영향을 미치게 된다. 따라서 무선 센서 네트워크에 큰 영향을 미치는 센서의 신뢰도는 모니터링, 진단 및 제어시스템 설계의 중요한 요소로 간주되고 있다. 일반적으로 신뢰성 높은 데이터를 얻기 위해서는 고장 난 센서의 교체 및 주기적인 보정과 같은 센서 유효화 과정이 필연적으로 요구된다. 그러나 이러한 과정은 무선 센서네트워크의 적용분야가 사람이 쉽게 접근하기 어려운 지역에 설치되기 때문에 시스템의 효율적인 운영을 위해서 고장 센서노드의 정확한 검출과 성능 기반 자동 센서 유효화 기법의 개발이 절실히 요구된다[6-7].

현재 많은 연구자들에 의해 고차 다변수 시스템 변수들간의 복잡한 관계를 새로이 생성된 저차의 변수들로 표현함으로써 변수들 간의 중요한 변량을 효율적으로 표현하고자 하는 주성분분석(Principle Component Analysis)이 고장검출 및 진단에 효과적으로 도입되고 있다[8]. PCA는 측정 데이터들간에 존재하는 공분산 정보를 이용하여 PCA 모델을 간단히 만들 수 있으며 PCA 모델과 계통으로부터의 출력을 비교함으로써 계통의 고장을 효율적으로 검출할 수 있다는 특성을 갖는다. 온라인 센서 유효화기법 또한 활발하게 연구되고 있으며 특히 Hines 등은 Jang이 제안한 ANFIS(Adaptive Neural Fuzzy Inference System) 및 Kramer가 제안한 AANN(Auto Associative Neural Network)을 기반으로 온라인 센서 유효화 기법을 제안하고 이의 유용성을 입증한 바 있다[9-10].

[†] 교신저자, 正會員 : 群山大學校 電子情報工學部 教授

* 正會員 : 群山大學校 電子情報工學部 博士課程

** 正會員 : 群山大學校 電子情報工學部 博士課程

接受日字 : 2006년 11월 7일

最終完了 : 2006년 12월 4일

본 연구에서는 다수의 센서노드들로부터의 측정 데이터를 받아들이며 원격의 서버로 전송함과 동시에 측정 데이터에 발생할 수 있는 데이터 오류를 효율적으로 검출 및 복원할 수 있는 PCA 및 AANN 기반 데이터 압축 및 복원 시스템의 유용성을 살펴보고자 한다.

2. 데이터 압축 기법 및 이상 데이터 검출 기법

2.1 주성분 분석(Principle Component Analysis)[8]

주성분 분석은 변수들의 선형 결합을 통해 변수들 간에 존재하는 중요한 변량의 방향을 결정하는 통계학적 모델링 기법으로 서로 상관되어 있는 변수들 사이의 복잡한 관계를 단순화하기 위하여 사용된다. 즉, 상관관계가 있는 변수들을 결합하여 서로 독립적인 새로운 저차의 새로운 변수들을 생성하는 것으로 생각할 수 있으며 이때 새로이 생성된 변수를 주성분이라 부른다. PCA의 기본 개념에 대해 설명하면 다음과 같다.

n 개의 변수들에 대한 m 번의 측정으로 얻어진 데이터 행렬을 $X(m \times n)$ 이라고 정의하고 X 는 0 인 평균값과 크기가 1인 분산을 갖는다고 가정한다. 이 경우 X 의 공분산 행렬은 다음과 같이 정의된다.

$$cov(X) = \frac{X^T X}{m-1} \quad (1)$$

PCA는 데이터 행렬 X 를 다음과 같이 분해할 수 있으며

$$X = \hat{X} + E \quad (2)$$

여기서

$$\hat{X} = \sum_{i=1}^r t_i p_i \quad r \leq n \quad (3)$$

식(2)에서 E 는 모델 예측시의 편차를 나타낸다. 또한 가중치 벡터 t_i 는 프로세스에 존재하는 잡음에 대한 정보를 포함하며 벡터 p_i 는 변수들 사이의 상관관계에 대한 정보를 포함한다. 벡터 p_i 는 다음과 같이 공분산 행렬의 고유벡터이기도 하다.

$$cov(X)P_i = \lambda_i P_i \quad (4)$$

식(3)에서 사용되는 고유벡터는 가장 큰 고유값에 대한 고유벡터부터 r 개까지 내림차순으로 사용된다. 또한 각 고유벡터에 대한 가중치 벡터 t_i 는 다음과 같이 구할 수 있다.

$$t_i = X p_i \quad (5)$$

만일 또 하나의 새로운 측정이 이루어질 경우, 새로운 측정 벡터 x 에 대한 모델 에러 e 는 다음과 같이 구할 수 있으며

$$e = (x - \hat{x}) \quad (6)$$

새로운 측정 벡터 x 에 대한 잔차 R 은 다음과 같이 구해진다.

$$R = ee^T \quad (7)$$

따라서 식(7)를 이용한다면 정상으로 동작하는 시스템에 대한 PCA 모델을 구하는 것이 가능하게 된다. 이는 새로운 측정 벡터가 정상시와는 다른 양상을 나타낼 경우 식(7)로 표현되는 잔차 R 이 신뢰성 한계(confidence limit)를 초과하게 되고 이로부터 효율적인 고장의 검출이 가능하게 된다.

2.2 AANN(Auto Associative Neural Network)[9]

PCA(Principle Component Analysis)는 선형 매핑에 기반한 기법으로 비선형 데이터에의 직접 적용은 어렵게 된다. 따라서 비선형 데이터의 분석에 적용될 수 있는 기법들이 많은 연구자들에 의해 개발되었다. 이중 Kramer등에 의해 제안된 AANN은 자기연관 신경망을 이용하여 비선형 PCA를 구현하고자 한 것으로 5개층(입력층, 매핑층, 바틀넥층, 디매핑층, 출력층)으로 구성되며 구조를 나타내면 그림 1과 같다.

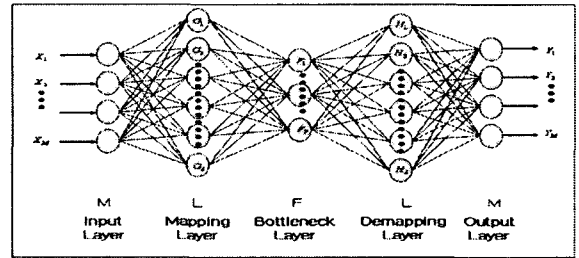


그림 1 AANN의 구조

Fig. 1 The Structure of AANN

그림 1의 입력층은 M -차원의 데이터(X_1, \dots, X_M)를 입력으로 받아들인다. 입력된 데이터들은 한개의 계층 스테이션으로부터 측정된 다양한 형태의 센서값이 될 수 있으며 이들은 상호 연관성을 갖고 있어야 한다. 매핑층은 M -차원의 입력 데이터를 이보다 적은 F -차원의 특징 데이터로 축약하는 기능을 수행하는 L -차원의 뉴런으로 구성되며 이의 입출력 특성을 나타내면 식(8)과 같다.

$$F = G(X) \quad (8)$$

식 (8)에서 G 는 매핑층 L -차원의 뉴런 활성화함수(G_1, G_2, \dots, G_L)들을 요소로 갖는 벡터이며 X 는 $1 \times M$ 차원의 측정변수를 의미하며 F 는 $1 \times F$ 차원의 축약된 데이터를 의미한다. 이러한 특성을 통해 M -차원의 입력데이터로부터 F -차원의 데이터로의 압축이 가능하며 바틀넥 층의 출력을 통해 압축된 데이터를 얻을 수 있다. 디매핑층은 축약된 $1 \times F$ 차원의 데이터를 원래의 차원($1 \times M$)으로 복구하는 기능을 수행하는 L -개의 뉴런으로 구성되며 이의 특성을 나타내면 식(9)과 같다.

$$Y = H(F) \quad (9)$$

식(9)에서 H 는 다매핑층의 L -차원의 뉴런 활성화함수 (H_1, H_2, \dots, H_L)를 요소로 갖는 벡터이며 Y 는 자기상관 신경망의 입력으로 사용된 것과 동일한 $1 \times M$ 차원 출력 데이터이다. 식(8), (9)의 뉴런 활성화함수 G, H 에는 학습시 조정되어야 할 결합강도가 포함되며 이들은 일반적인 교차학습방법에 의해 다음의 잔차가 최소가 되도록 학습된다.

$$R = X - Y \quad (10)$$

2.2.1 AANN 기반 센서노드 고장 검출 기법

본 연구에서는 Hines 등이 제안한 바 있는 AANN을 기반으로 하는 센서 고장 검출 기법을 사용하고자 하며 이를 나타내면 그림 2와 같다[3-5]. AANN은 다른 일반적인 인공신경망의 경우와 마찬가지로 다차원 비선형 데이터의 상관정보가 학습을 통해 뉴런 간 결합강도에 분산, 저장되며 따라서 학습된 데이터와는 다른 데이터가 AANN에 인가된다 할지라도 원래의 데이터를 네트워크에 저장된 정보로부터 추정하는 것이 가능하게 된다. 이러한 AANN의 데이터 추정 기능을 이용한다면 이상 데이터의 효율적인 검출할 수 있다.

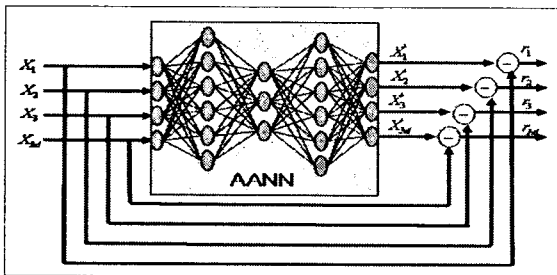


그림 2 AANN의 데이터 추정능력을 이용한 이상 데이터 검출 특성
Fig. 2 Characteristics of abnormal data detection using data estimation characteristic of AANN

여기서 X_1, X_2, X_3, X_M 는 실제의 계측 데이터를 의미하며 X'_1, X'_2, X'_3, X'_M 는 입력 계측데이터에 대한 AANN의 출력을 의미한다. 또한 r_1, r_2, r_3, r_M 는 실제 계측 데이터와 AANN 출력간의 잔차를 의미한다. 일반적으로 학습된 AANN은 입력 데이터의 차수 보다 적은 뉴런으로 구성되는 바틀벡층의 도입으로 학습시 사용된 데이터와 다른 데이터가 입력된 경우에도 데이터의 효율적 추정이 가능하게 된다. 따라서 그림 2와 같이 AANN을 사용한다면 학습시와 다른 특성을 갖는 데이터의 발생 여부를 간단히 검출할 수 있다.

2.3 제안된 기법의 센서 네트워크에의 적용

본 논문에서는 계층적 라우팅 알고리즘 중 LEA2C 라우

팅 프로토콜로 운영되고 있는 센서 네트워크에서 클러스터 헤드로부터 싱크노드로의 데이터 전송시의 에너지 절약을 위한 데이터 압축 기법 및 압축 전송된 데이터를 기반으로 한 네트워크내의 고장 센서 노드 검출을 가능케 하는 새로운 형태의 시스템을 제안하고자 한다. 제안된 시스템은 클러스터 헤드에서 싱크노드로의 에너지 데이터 전송 및 모니터링 서버에서 압축 전송된 데이터를 기반으로 하는 센서노드의 고장검출을 위해 PCA 와 AANN을 도입하고자 하며 제안된 시스템의 전체 구성은 그림 3과 같다.

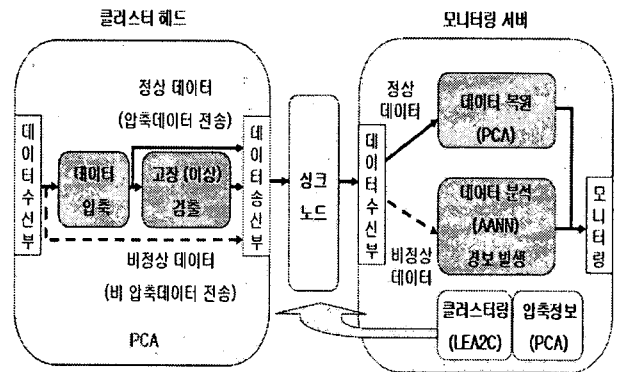


그림 3 제안된 센서 네트워크 기반 원격 모니터링 시스템의 구조

Fig. 3 The Structure of remote monitoring system based on WSN

2.3.1 센서노드의 하드웨어 플랫폼

센서노드는 크게 데이터 처리부, RF 송수신부, 센서부, 외장장치 인터페이스부로 구성된다. 데이터 처리부는 RF 장치, 센서장치, 외부 인터페이스 장치를 제어하는 역할을 하며 이를 위해 ATMEL사의 ATMEGA128L 저전력 8비트 RISC 프로세서를 사용하였다. 이 칩은 내부에 128Kbyte의 FLASH와 4Kbyte의 SRAM 과EEPROM을 내장하고 있어 TinyOS를 구동하기에 충분하며 전력소비를 줄이기 위한 6가지 슬립모드를 지원한다. RF 송수신부는 네트워크를 형성하는 센서노드들 간의 통신을 담당한다. RF transceiver 칩으로는 Cipcon사의 CC1000을 사용하였으며 지원하는 300MHz에서 1000MHz 사이의 대역 중 915MHz의 고정 주파수를 사용하였다. 이 칩은 SPI와 I2C 방식의 인터페이스를 이용하여 CPU와 연결되며 최소한의 전력소비를 지원하기 위한 power down mode를 지원한다.

센서부는 센서노드에 부착된 센서의 데이터를 취득하는 역할을 담당하며 온도, 습도, 조도, 초음파 센서등의 기타 여러가지 센서들을 부착하여 사용 가능하다. 부착된 센서들은 ATMEGA128L에 내장된 10비트 해상도의 AD컨버터와 I2C 방식이나 SPI 방식을 이용하여 취득가능하다.

마지막으로 인터페이스부는 외장장치와의 연결 및 보드 확장을 위해 사용된다. 이를 위해 2개의 UART 포트와 각종 I/O 포트를 지원한다. UART 포트는 PC와 임베디드 장치와의 연결 하기 위해 사용되어지며 I/O 포트는 역추제이터 노드를 구성하거나 다른 소자를 이용하여 보드를 확장 사용할 수 있도록 한다.

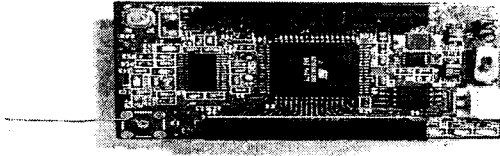


그림 4 ATMEGA128L 기반 센서 노드
Fig. 4 ATMEGA128L Sensor Node

3.2.2 LEA2C(Low Energy Adaptive Connectionist Clustering)

센서네트워크는 응용분야에 대한 의존성이 강한 특성을 갖는다. 응용분야에 따라 요구하는 정보 전송 유형이 달라 지므로 그것에 적합한 라우팅 방식이 선택되어질 수 있다. 본 연구에서는 제한한 기법을 보다 효율적으로 운용하기 위해서 SOMs(self-organizing maps)과 같은 자율 연결주의자 (unsupervised connectionist) 학습 기술과 LEACH-C의 다른 속성에 바탕을 둔 LEA2C 클러스터링 기법을 이용한 라우팅 알고리즘을 사용하였다[11].

LEA2C 라우팅 알고리즘은 다음과 같은 사항을 전제 조건으로 필요로 한다.

- ① 네트워크의 주어진 영역에 같은 종류의 센서노드 N 개가 임의로 설치되며 이 주어진 영역과 싱크 노드간의 거리는 멀리 떨어져 있다.
- ② 센서노드는 싱크 노드에 그들의 위치정보를 제공할 수 있는 GPS 시스템을 갖추고 있다.
- ③ 센서노드에 적용되는 에너지 소비 모델은 LEACH와 LEACH-C 프로토콜의 경우와 같다.
- ④ 싱크 노드는 생존해 있는 센서 노드들의 위치정보에 따라 클러스터링을 반복하여 실시한다.
- ⑤ 센서 노드의 클러스터링 기법은 SOMs을 사용한다.
- ⑥ 각 클러스터의 헤드 선정 기준은 다음과 같다.
 - 클러스터 내에서 가장 많은 에너지를 갖는 노드
 - 싱크 노드에 가장 가까이 설치된 노드
 - 클러스터 중심에 가장 가까운 센서노드
- ⑦ 센서 노드는 단지 클러스터 헤드에만 센서 데이터를 전송 한다.
- ⑧ 클러스터 헤드는 데이터 압축 및 병합을 하고 싱크 노드에 데이터를 전송한다.
- ⑨ 센서노드는 시간이 지남에 따라 에너지 고갈로 인해 기능을 상실한다.
- ⑩ 네트워크 안에 살아 있는 센서노드들이 존재하는 한 이러한 작업을 반복한다.

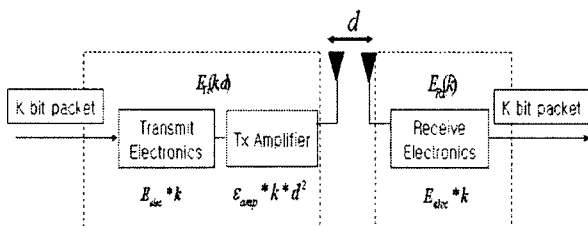


그림 5 무선 센서 네트워크에서 통신 에너지 소비 모델
Fig. 5 Communication power consumption model in WSNs

$$E_{Tx}(k,d) = E_{Tx-elec}(k) + E_{Tx-amp}(k,d) \quad (11)$$

$$E_{Tx}(k,d) = E_{elec} * k + \epsilon_{amp} * k * d^2 \quad (12)$$

또한 k 비트의 메시지를 수신하는데 소비되는 에너지는 식(12)와 같다.

$$E_{Rx}(k) = E_{Rx-elec}(k) = E_{elec} * K \quad (13)$$

여기서 $E_{Tx-elec}$, $E_{Rx-elec}$ 은 송수신 회로에서 소비되는 에너지이고 $E_{Tx-elec} = E_{Rx-elec} = E_{elec} = 50nJ/bit$ 이다.

ϵ_{amp} 는 메세지 전송을 위해 증폭에 소비되는 에너지로 $100pJ/bit/m^2$ 이다. 또한 센서노드 내에 연산을 위해 소비되는 에너지는 식(3.4)과 같다.

$$E_{DA}(k) = 5nJ/bit/msg \quad (14)$$

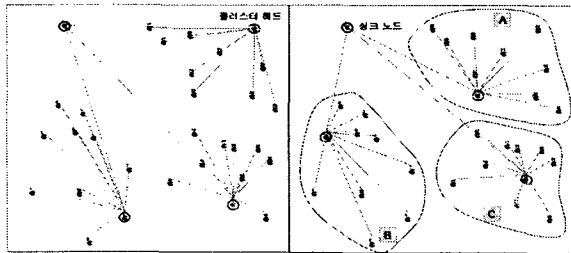
상기와 같은 전제하에서 구동되는 LEA2C 라우팅 알고리즘 다음과 같다.

- 1) 초기화
 - 센서노드들은 같은 양의 에너지를 갖으며 주어진 공간에 랜덤하게 설치된다.
- 2) 클러스터링
 - 2.1) 각각의 센서노드들의 위치정보를 바탕으로 SOM을 이용하여 첫 번째 클러스터링을 수행하며 그 결과를 K-means 클러스터링에 적용한다.
 - 2.2) 주어진 클러스터 헤드 선정 기준에 따라 모든 클러스터의 헤드를 선정한다.
 - 2.3) 각각의 센서노드에 역할을 할당한다.
- 3) 데이터 전송
 - 3.1) 센서노드들은 그들의 클러스터 헤드에 데이터를 전송한다. 소비된 에너지는 다음과 같은 식(11), 식(12)에 의해 계산된다.
 - 3.2) 클러스터 헤드에서 사용된 에너지를 계산한다.
 - 데이터를 수신시 - 식(13)
 - 데이터 처리시 - 식(14)
 - 싱크 노드로 전송시 - 식(12)
 - 3.3) 데이터 전송 후 클러스터 헤드는 최대 잔존 에너지(Maximum energy)에 의해 다시 선정된다.
 - 3.4) 에너지 고갈 또는 고장으로 인해 센서노드가 더 이상 동작하지 않을 때까지 3.1)에서 3.3) 단계를 반복한다.

4) 네트워크 안에 모든 센서노드들이 동작하지 않을 때까지 2), 3) 단계를 반복한다.

3. 적용 실험

본 연구에서 제안한 알고리즘의 유용성을 확인하기 위해 30개의 센서노드를 임의로 설치하여 센서 네트워크를 구성하였다. LEA2C 라우팅 프로토콜의 모의실험을 위해 HUT(Helsinki University of Technology)에 의해 제안된 SOM Toolbox 2.0[12]를 이용하여 클러스터링 하였고 TOSSIM 시뮬레이터를 이용하여 LEA2C 라우팅 알고리즘에 의해 운영되는 센서 네트워크는 그림 6과 같이 나타내었다. 그림 6의 (a)는 임의의 시간 t에 클러스터 헤드를 나타내고 (b)는 시간 t+1에 변화된 클러스터 헤드를 나타내고 있다.

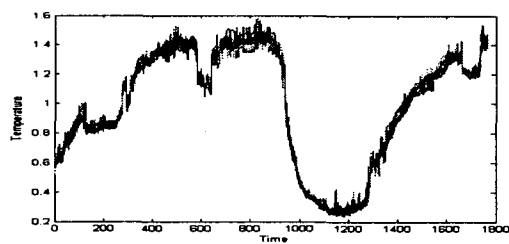


(a) Cluster head nodes at time t (b) Cluster head nodes at time t+1

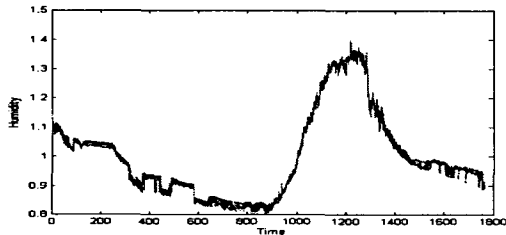
그림 6 LEA2C 알고리즘에 의한 클러스터 헤드 변화 과정
Fig. 6 Dynamic clusters of LEA2C

3.1 PCA 기반 압축 및 고장검출기법 적용

각각의 센서노드들의 온도와 습도를 약 1분 간격으로 24 시간 동안 측정하였으며 클러스터 A에서 측정된 데이터를 나타내면 그림 7과 같다. 각각의 온도와 습도 데이터는 같은 시간에 측정된 데이터로 압축 및 고장검출 표본 데이터로 사용된다.



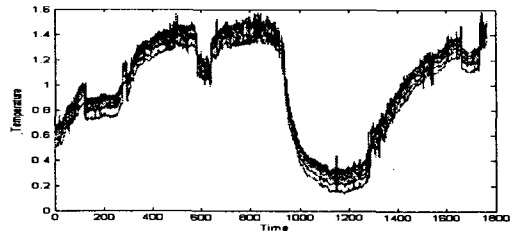
(a) Temperature



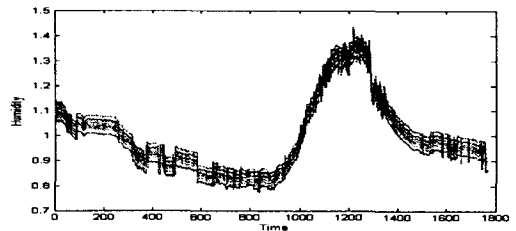
(b) Humidity

그림 7 각 센서의 측정 데이터
Fig. 7 Measurement data from each sensor

본 절에서는 앞에서 언급한 PCA기법을 이용하여 제안된 알고리즘의 유용성을 확인하고자 한다. 센서노드들로부터 데이터를 수신한 클러스터 헤더는 싱크노드로 데이터를 전송하기 전에 PCA를 이용하여 데이터 압축 및 고장검출 기법에 적용한다.



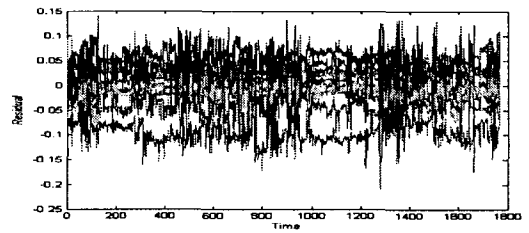
(a) Temperature



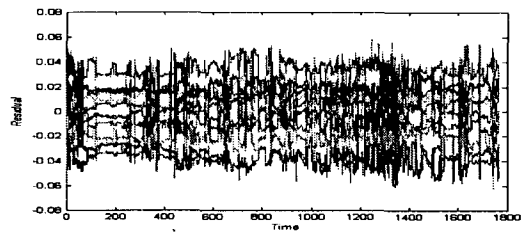
(b) Humidity

그림 8 4개의 고유벡터로 복원된 온도 데이터
Fig. 8 Reconstructed data by PCA with 4 eigen vectors

그림 8은 표본 데이터 그림 7에 PCA 알고리즘을 적용하여 복원된 데이터이다. 온도와 습도 총 20개의 센서로부터 측정된 데이터를 4개의 주성분으로 축소하여 복원하였다. 실제 데이터와 복원 데이터와의 잔차는 그림 9와 같다.



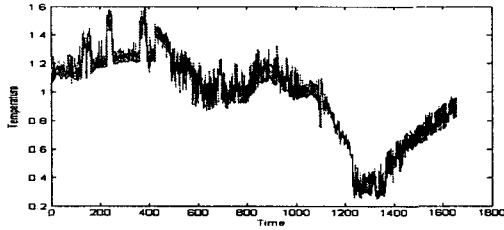
(a) Temperature



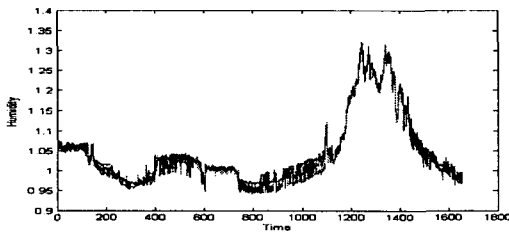
(b) Humidity

그림 9 실측 온도 데이터와 복원된 데이터 간의 편차
Fig. 9 Residual between measurement data and output of PCA

PCA 기법의 압축 및 복원 성능을 확인하기 위하여 구해진 주성분을 그림 10과 같은 새로운 데이터에 적용하여 보았다. 측정된 데이터는 마찬가지로 각각의 센서노드들의 온도와 습도를 약 1분 간격으로 24시간 동안 측정하였다.

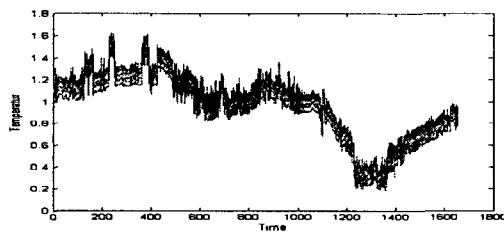


(a) Temperature

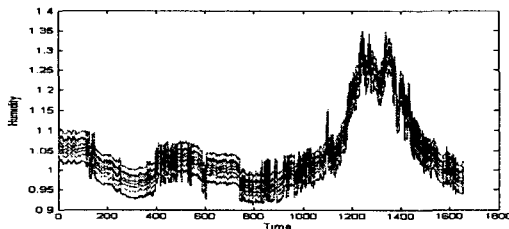


(b) Humidity

그림 10 공분산 행렬 계산에 사용되지 않은 새로운 데이터
Fig. 10 Measurement data not used in calculating covariance matrix

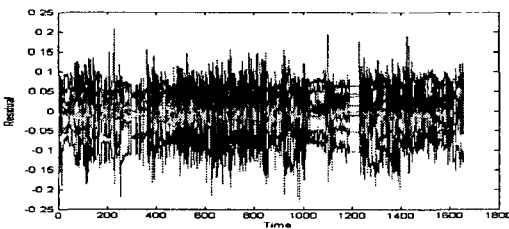


(a) Temperature

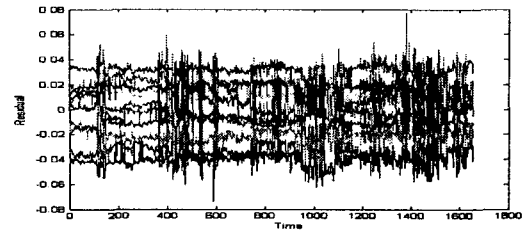


(b) Humidity

그림 11 4개의 고유벡터로 복원된 온도 데이터
Fig. 11 Reconstructed data by PCA with 4 eigen vectors



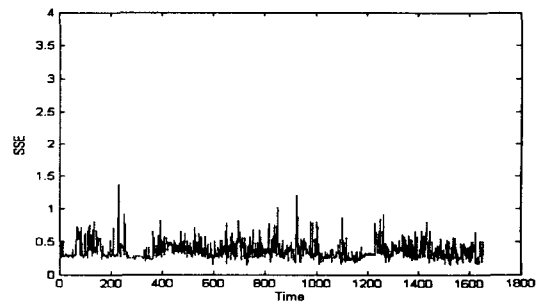
(a) Temperature



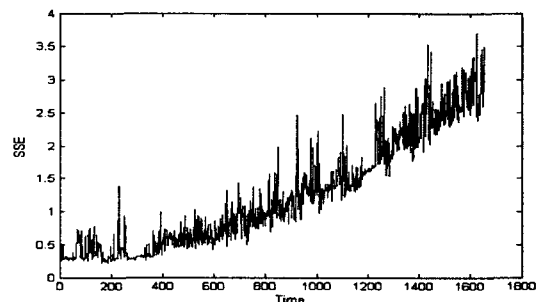
(b) Humidity

그림 12 실측 온도 데이터와 복원된 데이터 간의 편차
Fig. 12 Residual between measurement data and output of PCA

또한 PCA를 이용한 고장검출 유용성을 확인하기 위해 100번째 스텝에서 램프 형태의 바이어스를 인가하여 고장 상태를 가정하였다. 이때 복원알고리즘에 사용된 주성분의 개수 m 을 4로 하였다. 그림 13 (a)는 정상상태에서 습도센서와 PCA 압축알고리즘에 의해 복원된 데이터와의 SSE(Sum of Square Errors)를 나타내고 있다. 정상상태 일 때는 약 0.5 값의 잔차가 존재한다. 그림 13 가정된 고장이 발생했을 경우 고장검출을 위한 SSE를 나타내고 있다. 여기서 신뢰성 한계(confidence limit)값을 1로 설정하였고 이로 인해 공정에서 발생하는 잔차가 1을 초과할 때 고장이 검출된다.



(a)Normal data



(b)Abnormal data

그림 13 바이어스를 2번 센서노드 습도센서에 인가된 데이터와 PCA 출력과의 SSE
Fig. 13 SSE of between measurement data and output of PCA (Humidity sensor of 2th sensor node)

이와 같은 방법으로 2번 센서노드의 온도센서의 고장검출 성능을 확인하였다.

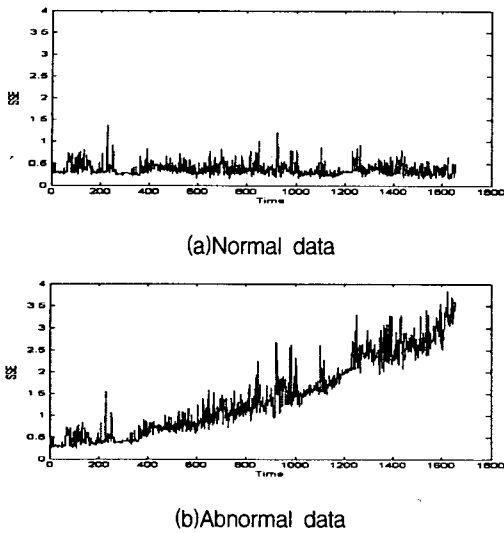
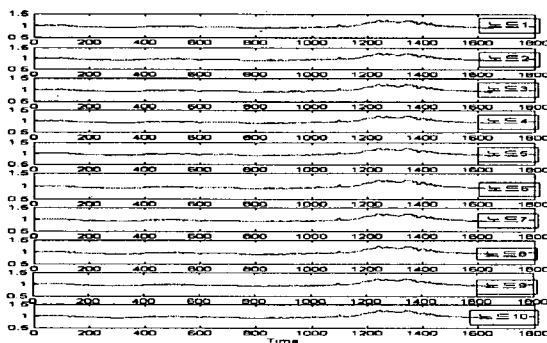


그림 14 바이어스를 2번 센서노드 온도센서에 인가시 인가된 데이터와 PCA 출력과의 SSE
 Fig. 14 SSE of between measurement data and output of PCA (Temperature sensor of 2th sensor node)

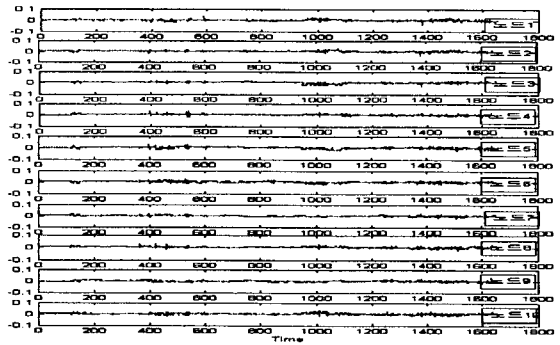
3.2 AANN의 데이터 추정특성을 이용한 고장 진단

AANN의 학습에 사용되는 입력력 패턴은 동일하며 따라서 AANN은 1:1 매핑기능을 갖는다. 그러나 앞서 언급한 바와 같이 입력력 뉴런의 수보다 적은 바틀넥층의 도입으로 인해 데이터를 압축 할 수 있으며 또한 비선형 PCA의 특성을 갖게 되어 학습시와는 다른 데이터(이상 데이터)가 입력되어도 정상 동작시의 값을 추정하는 것이 가능하게 된다. 본 연구에서는 그림 7의 습도 데이터 10개를 학습에 사용하였다. 학습에 사용되는 AANN의 입력층은 10개의 뉴런, 매핑층은 14개의 뉴런, 바틀넥층은 3개, 디매핑층은 14개의 뉴런으로 구성하였다. AANN의 학습은 Matlab 신경망들의 NEWFF (Feedforward Backpropagation Network)함수를 사용하였다. AANN을 이용한 데이터 압축 및 복원 효과를 위해 AANN을 분리한 형태의 신경망을 구현 하였으며 성능 확인을 위해 Matlab 시뮬링크를 사용하였다.

학습이 완료된 AANN의 성능의 확인을 위해 그림 10의 습도 데이터를 AANN에 인가하였으며 이 경우 AANN으로부터 출력되는 데이터와 그에 따른 잔차는 그림 15와 같다.

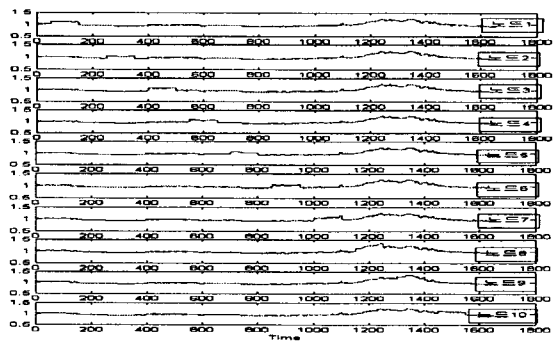


(a) Output of AANN

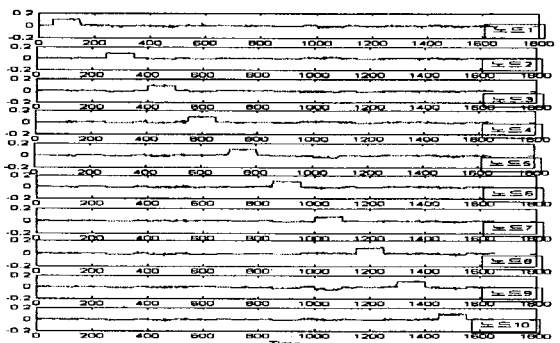


(b) Residual between measurement data and output of AANN
 그림 15 학습된 AANN의 응답특성(정상상태)
 Fig. 15 Response characteristic of learned AANN (normal state)

또한 제안된 시스템의 이상 데이터에 대해서 효율적인 이상의 진단 및 추정 특성을 확인하기 위해 10개의 센서 각각에 크기 0.1의 바이어스를 인가하였다. 50에서 150번 스텝사에 첫 번째 온도 센서에 바이어스를 인가하였으며 250에서 350까지는 2번 센서, 400에서 500까지는 3번 센서, 550에서 650까지는 4번 센서, 700에서 800까지는 5번 센서에 바이어스를 인가하였다. 그리고 850에서 950까지는 6번 센서, 1000에서 1100까지는 7번 센서, 1150에서 1250까지는 8번 센서, 1300에서 1400까지는 9번 센서, 1450에서 1550까지는 10번 센서에 바이어스를 인가하였다. 이 경우 각각의 이상 데이터에 대한 AANN출력을 나타내면 그림 16과 같다.



(a) Output of AANN



(b) Detection of abnormal data

그림 16 학습된 AANN의 응답특성(비정상상태)
 Fig. 16 Response characteristic of learned AANN (abnormal state)

그림 16의 (a)는 학습이 완료된 AANN에 바이어스가 포함된 센서값을 입력하였을 경우 AANN의 출력을 나타낸 것으로 AANN의 출력은 대부분 바이어스가 없을 경우의 정상 센서 값을 출력함을 알 수 있다.

또한 AANN과 실제 계측 센서 값과의 잔차를 나타내면 Fig 16의 (b)와 같으며 이와 같은 AANN의 데이터 추정능력을 이용하여 통해 센서 고장 검출 및 진단을 수행 할 수 있다.

4. 결 론

본 연구에서는 저전력 요구하는 무선센서네트워크에서 측정 데이터를 효율적으로 전송하고 고장검출 및 진단할 수 있는 알고리즘을 제안하였다. 제안된 PCA와 AANN 알고리즘은 변수들 간에 존재하는 상관관계를 제거함으로써 고차의 다변수 데이터를 정보의 손실을 최소화하면서 최소 차원의 데이터로 변환하는 알고리즘으로 압축 및 고장검출에 응용될 수 있다. 제안된 알고리즘은 데이터 융합(data aggregation) 특성을 고려할 수 있는 계층적 라우팅 알고리즘 가운데 LEA2C 라우팅 프로토콜을 적용하였다.

제안된 알고리즘의 유용성을 확인하기 위해 센서노드를 임의로 30개 설치하여 무선센서네트워크를 구성하고 제안된 알고리즘을 적용하여 다음과 같은 결과를 얻을 수 있었다.

1. 다변수 데이터의 해석에 널리 사용되고 있는 PCA을 기반 데이터 전송 및 복원 알고리즘의 유용성을 확인.
2. PCA 알고리즘을 이용하여 클러스터헤더에서의 효율적인 이상 데이터 검출 성능 확인
3. AANN 알고리즘을 이용하여 모니터링 서버에서의 고장검출 및 진단 성능 확인

제안된 알고리즘은 온습도 데이터와 같이 데이터 상호간의 선형적 특성을 갖는 데이터는 선형결합을 통해 데이터 분석할 수 있는 PCA 알고리즘으로 적용 가능하지만 비선형 특성을 나타내는 시스템은 적합하지 않아 이에 관한 연구가 요망된다.

감사의 글

본 연구는 산업자원부의 지역혁신 인력양성사업의 연구결과로 수행되었음

참 고 문 헌

[1] B.Krishnamachari, D. Estrin and S. Wicker, "Modelling Data-Centric Routing in Wireless Sensor Networks," IEEE INFOCOM'02, June 2002
 [2] W.Ye, J. Heidemann and D.Estrin "An Energy-Efficient MAC Protocol for Wireless Sensor Networks" IEEE INFOCOM'02, June 2002
 [3] S. Ranganathan, A.D. George, R.W. Todd, Matthew

C.Chidester, "Gossip-Style Failure Detection and Distributed Consensus for Scalable Heterogeneous Clusters", HCS Research Laboratory, 2000

[4] M. Barborak, A. Dahbura, M. Malek, "The Consensus problem in Fault-Tolerant Computing, ACM Computing Surveys", Vol 25, No 2, 171-220, 1993.
 [5] Ian F. Akyildiz, Weilian Su, Y. Sankar-subramaniam, Erdal Cayirci, "A Survey on Sensor Networks", IEEE Communications Magazine, August 2002, pages 102-114
 [6] MacGregor, J.F., K. Costas, "Process monitoring and diagnosis by multiblock PLS method", AIChE J. , 40(5), 826-838, 1994
 [7] R. Isermann, "Process Fault Detection Based on Modelling and Estimation Methods-A survey", Automatica, Vol. 20. 4, pp. 387-404, 1984
 [8] Dunia, R. "Identification of faulty sensors using principle component analysis", AIChE J., 42(10), 2797-2812, 1996
 [9] Kramer, M.A., "Autoassociative neural networks," Computers in Chemical Eng., Vol.16, No.4, pp.313-328, 1992
 [10] J. W. Hines, D. J. Wrest, and R. E. Uhring, "Plant Wide Sensor Calibration Monitoring:", published in the proceedings of The 1996 IEEE International Symposium on Intelligent Control, Sept. 15-18, pp.378-383,1996
 [11] F. K. L. Dehni and Y. Bennani. Power control and clustering in wireless sensor networks. In Proceedings of Med-Hoc-Net, 2005.
 [12] J.Vesanto, J. Himberg, E. Alhoniemi, J. Parhankangas" SOM Toolbox for Matlab 5", Copyright (C) 2005.(http://www.cis.hut.fi/projects/ somtoolbox/)

저 자 소 개



김 성 호(金星鎬)

1984년 고려대학교 공과대학 학사. 1986년 고려대학교 대학원 석사. 1991년 고려대학교 대학원 박사. 1988-1990년 고려대 생산기술연구소 연구원. 1995-1996년 JAPAN HIROSHIMA UNIVERSITY POST-DOC. 1997-현재 군산대학교 전자정보공학부 교수

E-mail : shkim@kunsan.ac.kr



이 영 삼 (李 榮 森)

2001년 군산대학교 공과대학 학사. 2003년
군산대학교 대학원 석사. 2003년-현재 군
산대학교 대학원 박사과정

Tel : 019-631-2032

E-mail : 03uncle@hanmail.net



육 의 수 (陸 宜 洙)

2004년 군산대학교 공과대학 학사. 2006년
군산대학교 대학원 석사. 2006년-현재 군
산대학교 대학원 박사과정

Tel : 011-653-2849

E-mail : sixofnum@hanmail.net