# 환경 모니터링 센서 네트워크에서의 이상 데이터 검출을 위한 시스템 설계

박수용*, 전민재**, 신용태***
*숭실대학교 컴퓨터학과
***숭실대학교 컴퓨터학부
*e-mail:waterdragon@soongsil.ac.kr

# System Design for Detecting Abnormal Data in the Monitoring Sensor Network

Soo-Yong Park*, Jeon-Min Jae**, Yong-Tae Shin*
*Dept of Computer Science, Soongsil University.

요 약

센서 네트워크는 IoT, 스마트 팜 등의 다양한 분야에 활용되고 있다. 그러나 센서 자체의 결함과 환경적인 요인으로 인해 데이터의 변화가 발생하는 단점을 가지고 있다. 제안하는 시스템은 기존의 센서 네트워크에 중앙 서버를 배치하고 중앙 서버에 STL Module를 추가한다. STL Module은 센서로부터 측정된 데이터를 분석하고 이상 데이터를 검출하여 제거한다. 중앙 서버에 추가적으로 Web Server를 구축하여 관리자가 실시간으로 모니터링할 수 있다.

## Ⅰ. 서론

환경 모니터링 시스템은 센서 네트워크를 이용하며 IoT, 스마트 팜 등 다양한 분야에 활용되고 있다. 그러나 센서 노후화, 센서 성능 미흡 등의 요인으로 인해 센서 데이터의 손실 및 변화의 가능성이 존재한다. 스마트 팜의 경우 센서 데이터에 이상이 발생하면 농작물을 재배할 수 있는 최적의 환경에 대한 모델을 도출할 수 없다. 따라서 센싱 데이터의 신뢰성을 검증할 수 있는 시스템이 필요하다. 이에 기존의 센서 네트워크 시스템에 이상 데이터 검출을 할 수 있는 모듈을 추가하여 센싱 데이터의 신뢰도를 높이고 이상 데이터를 생성하는 센서를 찾을 수 있는 시스템을 제안한다.

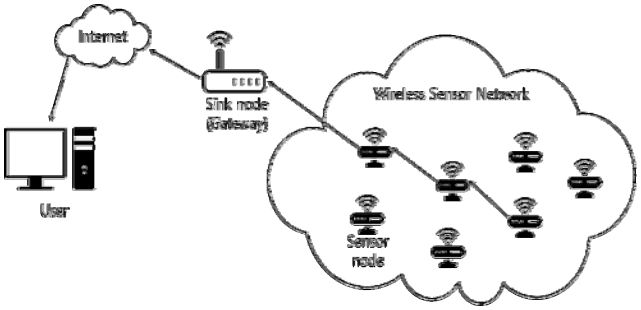
본 논문의 구성은 다음과 같다. 2장에서는 기반 기술인 무선 센서 네트워크와 STL(Seasonal decomposition of Time-series by Loess)에 대해 연구한다. 3장에서는 제안하는 시스템의 구체적인 설계를 제시한다. 4장에서는 설계한 시스템에 대한 결론을 도출한다.

## Ⅱ. 관련 연구

본 장에서는 센서 네트워크에서의 이상 데이터 검출을 위한 시스템을 설계하기 위한 관련 기술인 무선 센서 네트워크와 이상 데이터 검출 기법인 STL(Seasonal decomposition of Time-series by loess)에 대하여 연구한다.

### 2.1. 무선 센서 네트워크

무선 센서 네트워크는 대규모 환경 모니터링을 위한 저비용의 차세대 무선 네트워크 기술이다. (그림 1)은 무선 센서 네트워크의 토폴로지를 나타낸다.



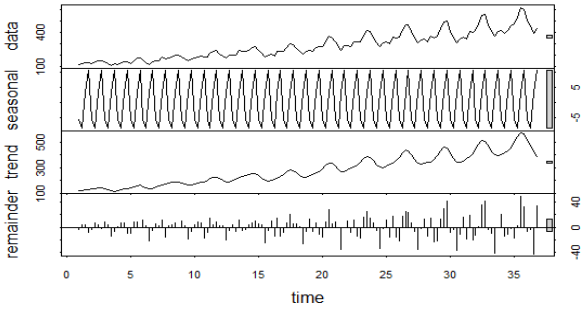

(그림 1) 무선 센서 네트워크의 토폴로지

무선 센서 네트워크는 기본적으로 센서 노드와 싱크 노드로 구성된다. 센서 노드는 데이터를 측정하는 역할을 한다. 싱크노드는 센싱 한 데이터를 수집하고 외부 네트워크로의 게이트웨이 역할을 수행한다[1]. 무선 센서 네트워크에는 수집된 데이터를 이용하여 분석 및 시각화를 위한 중앙 서버가 위치하는 경우도 있다.

센서 간의 무선통신에는 주로 ZigBee가 사용된다. ZigBee는 IEEE 802.15를 기반으로 만들어진 소형, 저전력의 개인 통신망을 구축하기 위한 기술이다[2]. ZigBee 장치는 메시 네트워크 방식을 사용하여 여러 중간 노드를 거쳐 목적지로 데이터를 전송하기 때문에 넓은 범위의 통신이 가능하다. 환경 모니터링 시스템은 스마트팜과 같은 광범위한 곳에 구축되기 때문에 ZigBee를 사용하기 적합하다[3].

### 2.2. STL

STL은 시계열 데이터가 갖는 시간 주기를 알고 있는 경우 데이터를 Seasonal, Trend, Remainder 요소로 분해하여 분석하는 기법이다[3]. (그림 2)는 시계열 데이터를 Seasonal, Trend, Remainder 요소로 분해한 예를 나타낸다.



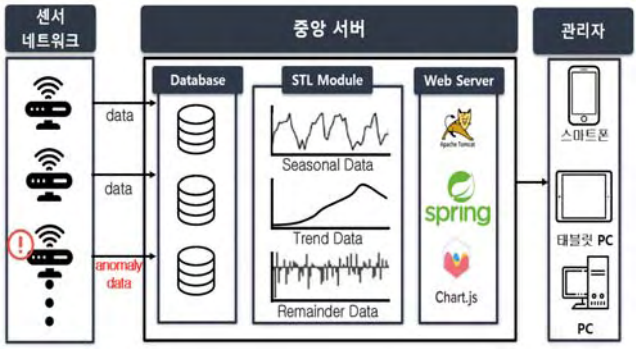

(그림 2) 시계열 데이터를 STL 요소로 분해한 예

Seasonal 요소는 계절적인 변화 또는 연간 1회의 주기를 갖는 데이터를 의미한다. Trend 요소는 데이터의 장기적인 수준의 변화를 의미한다. Remainder 요소는 전체 데이터에서 Trend 및 Seasonal 요소를 제거한 나머지를 의미한다. Remainder 요소는 정규 분포 특성을 가지고 있기 때문에 평균 표준편차 대비 크게 벗어나는 데이터를 이상 데이터로 검출할 수 있다[4].

환경 모니터링 센서 네트워크에서 수집하는 데이터는 온도, 습도 등과 같은 환경 데이터이다. 환경 데이터는 시간에 따라 변화하는 시계열의 특징을 가지고 있기 때문에 STL 기법을 적용하기 적합하다.

## III. 시스템 설계

본 논문에서 제안하는 시스템은 센서 네트워크, 중앙 서버로 구성된다. (그림 3)은 제안하는 시스템의 구조를 나타낸다.




(그림 3) 시스템 구조

센서 네트워크는 ZigBee가 장착된 센서 노드를 측정하고자 하는 환경에 망형으로 배치하고 중앙 서버로 데이터를 전달하기 위한 싱크 노드를 추가한다.

센서 노드는 싱크 노드로 데이터를 전달할 때 센서 노드의 ID를 포함하여 전송한다. 싱크 노드는 중앙 서버로 데이터를 전달할 때 센서 노드의 ID도 같이 전달하여 어느 센서에서 이상 데이터가 검출되었는지 알 수 있게 한다.

중앙 서버는 Database, STL Module, Web Server로 구성된다. Database에는 수집 데이터 테이블을 생성한다. <표 1>은 수집 데이터 테이블을 나타낸다.

<표 1> 수집 데이터 테이블

| 센서 ID | 측정 시간 | 온도 | 습도 |
|---|---|---|---|
| ID01 | 2019.03.12 12:00 | 19 | 60 |
| ID02 | 2019.03.12 12:00 | 20 | 61 |
| ID03 | 2019.03.12 12:00 | 19 | 62 |
| ID04 | 2019.03.12 12:00 | 30 | 60 |
| ID05 | 2019.03.12 12:00 | 18 | 80 |
| ID06 | 2019.03.12 12:00 | 17 | 75 |

수집 데이터 테이블은 센서ID, 측정 시간, 온도, 습도 등의 데이터를 저장한다. <표 1>의 “ID04” 센서와 “ID05” 센서는 각각 온도와 습도에 대해 이상 데이터를 생성했다.

STL Module은 수집 데이터 테이블에 저장된 시계열 데이터에서 Trend, Seasonal 데이터를 추출한다. STL Module은 원본 데이터에서 Trend, Seasonal 데이터를 제거하여 Remainder 데이터를 생성한다. Remainder 데이터의 평균 표준편차 대비 벗어나는 데이터를 이상 데이터로 판단한다. STL Module는 이상 데이터로 판단된 데이터를 수집 데이터 테이블에서 삭제한다. <표 2>는 이상 데이터가 삭제된 수집 데이터 테이블을 나타낸다.

<표 2> 이상 데이터가 제거된 수집 테이블

| 센서 ID | 측정 시간 | 온도 | 습도 |
|---|---|---|---|
| ID01 | 2019.03.12 12:00 | 19 | 60 |
| ID02 | 2019.03.12 12:00 | 20 | 61 |
| ID03 | 2019.03.12 12:00 | 19 | 62 |
| ID04 | 2019.03.12 12:00 | NULL | 60 |
| ID05 | 2019.03.12 12:00 | 18 | NULL |
| ID06 | 2019.03.12 12:00 | 17 | 75 |

Web Server는 관리자가 실시간으로 모니터링할 수 있는 웹 어플리케이션을 제공한다. Tomcat과 Spring을 사용하여 수집 테이블로부터 데이터를 가져올 수 있는 RestAPI 서버를 구축한다. Chart.js는 자바스크립트에서 제공하는 그래프 제작 라이브러리이다. 관리자에게 보이는 웹 페이지는 Chart.js를 이용하여 RestAPI로부터 전송받은 데이터를 월별, 주별, 일별 그래프로 시각화하여 출력

한다. (그림 4)는 그래프로 시각화된 관리자 페이지를 나타낸다.

(그림 4) 관리자 페이지

중앙 서버는 센서 노드들의 모든 ID를 가지고 있고 일정한 시간 간격으로 센서에게 응답을 요청하여 센서의 작동 여부를 판단한다. 만약 센서로부터 응답이 없을 경우 응답이 없는 센서의 정보를 웹 어플리케이션에 출력한다.

## Ⅳ. 결론

기존의 시스템은 센서로부터 수집한 데이터의 이상 데이터를 검출하는 기능이 없어서 잘못된 데이터를 맹신하는 오류가 발생하였다. 본 논문에서 제안하는 시스템은 기존의 센서 네트워크에 STL Module을 추가하여 이상 데이터를 검출하여 신뢰성 있는 데이터를 제공할 수 있다. 관리자는 검증된 데이터를 활용하여 최적의 환경 모델을 도출할 수 있다. 또한 실시간으로 환경을 모니터링할 수 있기 때문에 신속한 대처가 가능하고 센서와의 지속적인 통신을 통해 결함 센서를 검출할 수 있다. 그러나 STL 기법을 적용하여 데이터 분석을 하기 위해선 데이터가 1년 이상 쌓여야 한다는 단점이 있다. 향후 STL Module의 이상 데이터 검출 성능에 대한 추가적인 연구가 필요하다.

### Acknowledgement

본 논문은 중소벤처기업부에서 지원하는 2018년도 산학연협력 기술개발사업(No. S2632389)의 연구수행으로 인한 결과물임을 밝힙니다.

### 참고문헌

[1] 채동현, 한규호, 임경수, 안순신, “센서 네트워크의 개요 및 기술동향”, 정보과학학회지, 제22권, 제12호, pp. 5-12, 2004

[2] 제현우, 양오, “지그비 통신을 이용한 태양광 인버터의 원격 모니터링 시스템”, 한국정보기술학회 논문지, 제10권, 제2호, pp.94-101, 2012

[3] 최혀현, “지그비 네트워크에서의 전송 시간 할당 방법”, 한국컴퓨터정보학회 논문지, 제16권, 제4호, pp.119-126, 2011

[3] Robert B. Cleveland, William S. Cleveland, Jean E. McRae and Irma Terpenning, “STL: A Seasonal-Trend Decomposition Procedure Based on Loess” Journal of Official Statistics, Vol. 6, No. 1, pp.3-73, 1990

[4] 연한별, 김석연, 장윤, “Seasonal-Trend Decomposition과 시계열 상관관계 분석을 통한 비정상 이벤트 탐지 시각적 분석 시스템”, 한국컴퓨터종합학술대회 논문집, pp.1285-1287, 2014