

제품 및 장애 관리를 위한 웹기반 데이터 시각화 시스템

명제석*, 박성현, 류관희
*충북대학교 소프트웨어학과
e-mail:psh19940403@gmail.com

A Study of Web-Based Data Visualization System for Product and Fault Management

Je-Suk Myung*, Seong-Hyeon Park, Kwan-Hee Yoo
*Dept of Computer Science, Chungbuk National University

요 약

최근 4차 산업혁명이 이슈가 되면서 빅 데이터나 인공지능에 대한 연구가 활발해지고, 이를 통해 자동화 및 자율화가 제조 공정이나 차량 운행 등에서 활용되고 있다. 또한 이를 위해서 데이터를 분석하고 정제하며 시각화를 효과적으로 하는 방법에 대한 관심도 같이 늘어나고 있다. 본 논문에서는 자동화 공장의 제품을 관리함에 있어 데이터를 쉽게 이해할 수 있도록 시각화하는 방법에 대한 연구를 수행했다. 이를 위해 D3 자바스크립트 라이브러리를 통해 웹기반으로 구현한 제품과 장애를 효과적으로 관리할 수 있는 시스템을 개발했다. 제안하는 관리 시스템은 자동화 공장의 제조 공정 중 제품이나 장애 상황에 대한 이해를 빠르게 하도록 하여 의사결정 하는데 기여할 것이다.

1. 서론

독일에서 시작되어 최근까지 관심을 끌고 있는 4차 산업혁명은 정보통신기술의 발전에 따라 산업을 주도하는 시스템의 주체가 사람에서 사물로 변화했다. 이에 따라 인공지능, 사물인터넷, 빅 데이터와 같은 분야의 연구가 활발하다. 특히 이런 기술들이 발달하며 자율주행이나 스마트 팩토리나 같은 공장 자동화 기술도 급격히 발달했다.

공장의 자동화는 물론 오래전부터 진행되어왔다. 하지만 현재의 공정 과정은 파트별로 자동화가 이루어져 있기 때문에 다른 공정 라인에서는 장애가 발생하거나 품질에 문제가 생겨도 어디서 그 원인이 발생했는지 알기 어렵다. 이런 문제점이 발생하는 이유는 기계는 직접적으로 판단을 하여 시스템을 중단하거나 돌발 상황을 처리하거나 고장 및 장애를 극복할 수 없기 때문이다.

이전의 공장 자동화는 이러한 이유로 사실 기계화란 용어가 더 적합하다. 기계화는 기계가 인간의 노동력을 대체하는 것을 말하는 것이지만, 자동화는 기계가 인간의 판단력을 대체하는 방법을 개발하는 것에 중점을 둔다[1]. 따라서 현재는 아직 대부분의 자동화라는 기술이 문자의 의미 그대로의 자동화는 아니라고 할 수 있다. 공장 내의 전체 공정 과정이 연계되고 통합되어 실시간으로 모듈끼리 통신하고 협업을 하는 수준이 되어야 자동화라고 할 수 있다. 본 논문에서는 제조 공정의 운영에 대한 의사결정을 지원하기 위해 제품의 공정 관리에 대한 모니터링을 데이터 시각화 기법을 통해 하고자한다.

2. 관련 연구

2.1 데이터 시각화 시스템

의사결정을 지원하기 위해서는 제품과 관련된 공장의 위치 데이터, 제품의 제조 데이터, 장애의 발생, 인지, 처리에 대한 시간과 발생 유형 등 공장의 상황을 쉽고 빠르게 이해할 수 있도록 해야 한다. 사용자가 쉽게 데이터를 이해하려면 일반적으로 데이터를 분석하고 처리하는 과정이 필요하다. 자동화된 공장에서는 압도적으로 많은 센서 데이터를 수집하기 때문에 데이터를 있는 그대로 조회하는 것은 매우 비효율적이다.

따라서 데이터를 분석하고 정제하여 시각화로 표현하는 일련의 프로세스를 거쳐야 한다. 여기서 시각화란 정보를 시각적 요소와 매핑하는 것으로 데이터의 상관관계나 추이관계 등을 꺾은선 그래프, 산포도 등으로 나타내는 것을 예로 들 수 있다.



(그림 1) 데이터 시각화 프로세스 4단계[2]

그림 1은 데이터 분석 및 시각화 과정을 나타낸 것이다. 데이터를 시각화하기 위해서는 4단계를 거쳐야 한다. 먼저 특정 틀이나 크롤러와 같은 수집 모듈을 사용하여 데이터를 수집해야 한다. 이후 데이터를 저장하고, 알맞게 정제하는 전처리과정이 필요하다. 처리된 데이터는 다시 원하는 모델을 구축하는데 이용된다. 모델을 구축하면 마지막으로 쉽게 파악할 수 있도록 차트를 이용하여 시각화를 한다.

2.2 D3 라이브러리

데이터 시각화를 위해서는 일반적인 도표나 차트 제작 시에는 엑셀과 같이 대중화된 툴을 사용하지만, 지도 위의 지리적 데이터를 표시하는 데에는 어렵다. 지리적 데이터를 시각화하기 위해선 GIS, 구글 스프레드시트, CartoDB 등과 같은 지도 데이터 시각화 어플리케이션을 사용하거나 혹은 D3, R 등을 이용하여 직접 코드를 작성할 수 있다. R은 통계 계산과 그래픽을 위한 언어이고, 데스크톱에서 작동하기 때문에 웹에 비해서 접근성이 떨어진다[3].

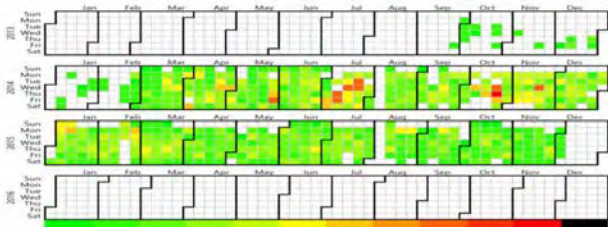
따라서 본 논문에서 제안하는 시스템은 접근성이 높은 웹 기반에서의 데이터 시각화를 위해 D3 라이브러리를 사용했다. D3 라이브러리는 자바스크립트로 만들어진 라이브러리로 Data-Driven-Document의 약자이다. D3는 다른 API처럼 미리 정의된 차트가 없기 때문에 간단한 차트도 직접 그려야 하는 번거로움이 있다. 하지만, 그런 점으로 인해서 어떤 형태든 그릴 수 있는 장점이 있다. D3는 데이터를 쿼리문같이 복잡한 인터페이스 없이도 간단하게 CSV, TSV, JSON 형식의 데이터를 로드하고 웹 요소와 바인딩할 수 있다.

3. 시스템 설계와 알고리즘

본 논문에서 개발하는 데이터 시각화 시스템은 높은 접근성을 위해서 웹을 이용해 개발한다. 시스템에서 시각화할 데이터는 제품의 관리에 필요한 위치 데이터와 제품의 상세 정보에 대한 데이터, 장애 발생 원인과 발생 시간, 인지 시간, 조치 시간 데이터이다.

먼저, 제품들의 공정 과정들을 확인 할 수 있도록 지도를 이용해 위치 데이터를 시각화한다. 지도를 먼저 그린 후에 그 위에 위치 데이터의 좌표를 통해 원으로 표시한다. 원의 색은 기본적으로 초록색을 유지시키고 장애가 발생한 경우 치명적이지 않으면 주황색, 치명적이면 빨간색으로 표시하여 의사결정에 도움을 줄 수 있도록 한다. 원에 커서를 올려놓으면 해당 지점의 상세 주소, 이름, 전화번호 및 제품들의 데이터를 표시한다.

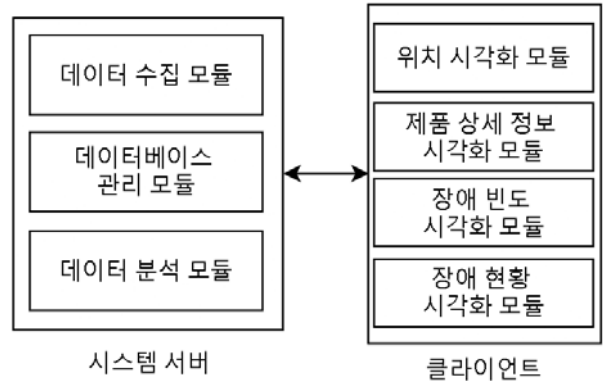
두 번째로는 제품 공정과정에서 장애의 발생 빈도를 나타내기 위해서 Calendar-Heatmap 기법을 이용한다.



(그림 2) Calendar-Heatmap의 예시[4]

이 기법은 빈도수에 따라서 색의 농도를 점점 짙게 하여 연간동안 하루에 특정 이벤트가 얼마나 발생했는지의 여부를 알 수 있다. 여기서 응용적으로 장애가 발생한 시간, 인지 시간, 조치한 시간을 기록한 데이터를 분석해서 또 다른 시각화모델을 만들 수 있다. 발생시간으로부터 인지시

간, 인지시간으로부터 조치시간까지의 데이터를 히트맵을 응용하면 장애가 발생한 이후 얼마나 빠르게 대처하는지 시각화 할 수 있다. 이 모델은 장애의 발생 현황을 쉽게 파악할 수 있도록 날짜별로 장애가 발생한 시각부터 빨간색, 인지된 시각으로부터 주황색, 조치하여 정상으로 돌아온 시각으로부터 초록색으로 표시한다.

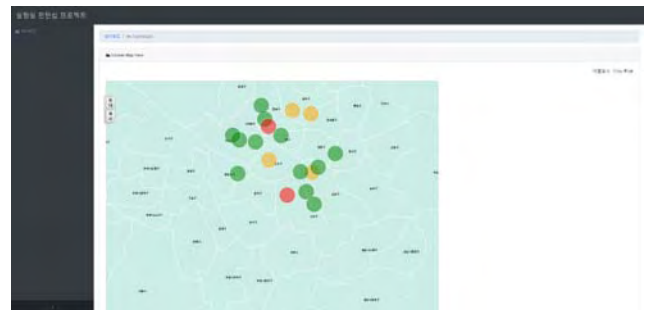


(그림 3) 데이터 시각화 시스템 구성도

그림 2는 데이터 시각화 시스템의 전반적인 구성도이다. 서버는 데이터를 각 지점으로부터 전달받는다 가정한다. 서버는 데이터를 전달받아 수집을 담당하는 모듈과 저장과 조회, 수정, 삭제 등을 담당하는 모듈, 데이터를 전처리화 하고 시각화하여 보고자 하는 모델을 구축하는 분석하는 모듈로 구성되어 있다. 클라이언트는 브라우저를 통해 시스템에 접근하는데, 이때 클라이언트의 입장에서는 자바스크립트를 통해 각각 차트를 모듈별로 구현했다.

4. 구현

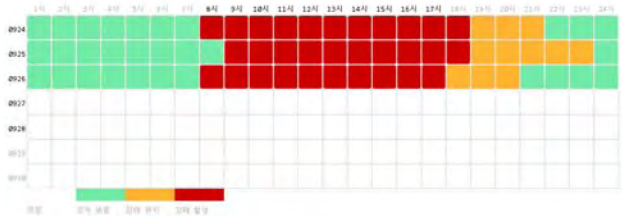
본 논문에서 제안하는 시스템은 웹 어플리케이션이기 때문에 HTML5, CSS, Javascript와 Python의 Django 웹 프레임워크를 사용했다. CSS나 대부분 HTML 코드는 부트스트랩을 사용하여 구성했다. 시각화 기법은 D3를 만든 마이크 보스타의 Github에 올려진 코드를 주로 참고했다. 시스템이 정상적으로 작동하는지 테스트하기 위해서 검증 데이터는 가상으로 생성했다.



(그림 4) 구현된 데이터 시각화 시스템 인터페이스

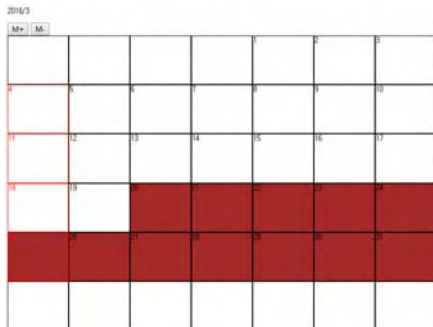
그림 4에서 볼 수 있듯이 인터페이스는 부트스트랩을 이용하여 전반적인 웹페이지의 형태를 형성하고 내용으로는

상호작용을 표현하기 위해서 대시보드를 활용했다. 대시보드 안에서는 웹의 SVG요소를 D3로 동적으로 지도를 그리고, 데이터를 그 위에 표현했다. 지도에 대한 데이터는 한국의 JSON파일을 이용해 그렸다. 실제 공장과 협업을 할 수 있는 조건이 되지 않아서, 각 공장의 위치 데이터는 데이터베이스에 저장된 가상의 좌표로 CSV파일을 만들고 가상의 좌표를 생성하여 표시했다. 각 지점들은 상태 변수값을 참조하여 원의 색을 상태별로 다른 색으로 채우게 된다.



(그림 5) 구현된 장애 발생 현황 시각화

그림 5는 Calendar Heatmap을 응용하여 장애의 발생, 인지, 조치에 대한 현황을 시각화한 것이다. 이를 통해서 장애가 언제 발생했는지 시각을 알 수 있다. 또한 인지한 시점과 조치가 된 시점까지 한 눈에 파악할 수 있어서 장애가 자주 발생하는 시간대를 파악할 수 있고, 장애가 발생하는 패턴에 대해서 파악하는 데에도 유용하다.



(그림 6) 구현된 장애 빈도수 Calendar Heatmap

그림 6에서는 일반적인 캘린더 히트맵과 같이 연간 단위가 아닌 월간 단위로 장애의 빈도수를 조회하고자 월간 캘린더 형태로 구현한 모습이다. 장애 빈도수는 다음과 같이 Calendar Heatmap을 사용하여 특정 일자로부터 어떤 장애가 발생할 경우, 해당 일자부터 빨간색으로 표시하고 빈도수에 따라 다른 농도를 부여한다.

5. 결론 및 향후 계획

본 논문에서는 D3 라이브러리를 통해 웹기반 데이터 시각화 시스템을 개발했다. 여러 시각화 기법을 통해서 제품과 장애를 쉽게 관리할 수 있도록 했다. 아직은 실제 스마트 팩토리에서 데이터를 공수하거나, 기계화된 설비들에서 데이터를 받는 것이 아니기 때문에 시스템 테스트를 위해서 가상의 데이터를 이용했던 아쉬운 점이 있다.

향후 연구 계획으로는 실제로 스마트 팩토리를 운영하는

데 이용되는 각종 센서 데이터와 위치 데이터, 장애에 관련된 데이터들을 수집해보고, 실시간으로 모니터링 할 수 있는 시간 시각화기법을 연구하여 온도같은 주변 환경도 실시간으로 모니터링할 수 있게 추가하는 것이다. 마지막으로 Weka를 사용한 연관 분석을 통해서 장애의 발생 원인들이 각각 어떤 상관관계를 가지고 있는지 파악할 수 있는 시스템으로 개선해가며 더 효과적인 시각화 기법을 연구해서 제조 공정과 관련된 의사결정을 지원하는 것이다.

ACKNOWLEDGEMENT

“본 논문은 교육부가 지원하고 충북대학교가 수행하는 지역선도대학육성사업의 지원을 받아서 수행되었습니다.”

참고문헌

- [1] 박기현 외, “우리나라의 공장자동화기술 개발현황”, 전기저널, pp 24-29, 1988.12
- [2] NIPA, “빅 데이터 - 정제와 분석 편”, <http://www.sw-eng.kr/member/customer/Webzine/BoardView.do?boardId=0000000000000040141> (2018.09.17.)
- [3] 최진 외, “웹 환경에서의 지도 기반 데이터 시각화 인터페이스 툴 개발”, 멀티미디어학회논문지, pp 1216-1223, 2017.8
- [4] 류관희, “제조 빅데이터 시스템을 위한 효과적인 시각화 기법”, 한국데이터정보과학회지, pp 1301-1311, 2017.11