

Web-Based Data Analysis Service for Smart Farms

Jimin Jung[†] · Jihyun Lee^{††} · Hyemin Noh^{†††}

ABSTRACT

Smart Farm, which combines information and communication technologies with agriculture is moving from simple monitoring of the growth environment toward discovering the optimal environment for crop growth and in the form of self-regulating agriculture. To this end, it is important to collect related data, but it is more important for farmers with cultivation know-how to analyze the collected data from various perspectives and derive useful information for regulating the crop growth environment. In this study, we developed a web service that allows farmers who want to obtain necessary information with data related to crop growth to easily analyze data. Web-based data analysis service developed uses R language for data analysis and Express web application framework for Node.js. As a result of applying the developed data analysis service together with the growth environment monitoring system in operation, we could perform data analysis what we want just by uploading a CSV file or by entering raw data directly. We confirmed that a service provider could provide various data analysis services easily and could add a new data analysis service by newly adding R script.

Keywords : Precision Agriculture, Smart Farm, Growth Data Analysis, Data Analysis-as-a-Web-Service

스마트팜을 위한 웹 기반 데이터 분석 서비스

정지민[†] · 이지현^{††} · 노혜민^{†††}

요약

농업에 정보 통신 기술을 접목한 스마트팜은 단순한 생육 환경 모니터링에서 벗어나 작물 생육을 위한 최적의 환경을 발견하고 자율제어가 가능한 농업의 형태로 나아가고 있다. 이를 위해서는 관련 데이터를 수집하는 것도 중요하지만, 재배 경험과 지식을 가진 농업인 사용자들이 수집된 데이터를 다양한 관점에서 분석하여 작물 생육 환경 제어에 유용한 정보를 도출해야 할 필요가 있다. 본 연구에서는 작물 생육과 관련된 데이터를 가지고 필요한 정보를 얻고자 하는 농업인 사용자가 쉽게 데이터 분석을 할 수 있는 웹 서비스를 개발하였다. 개발한 웹 기반 데이터 분석 서비스는 데이터 분석을 위하여 R 언어를 사용하며 Node.js를 위한 익스프레스 웹 애플리케이션 프레임워크를 기반으로 개발하였다. 데이터 분석 서비스를 운영 중인 생육 환경 모니터링 시스템과 함께 적용해 본 결과 사용자는 웹 상에서 CSV 형식의 파일을 입력하거나 직접 데이터 입력함으로써 서버가 제공하는 데이터 분석을 위한 R 스크립트를 실행하여 데이터 분석을 수행할 수 있었다. 서비스 제공자는 다양한 데이터 분석 서비스를 쉽게 제공할 수 있었고, R 스크립트만 새로 추가하면 애플리케이션에 대한 수정 없이 새로운 데이터 분석 서비스 추가가 용이함을 확인하였다.

키워드 : 정밀 농업, 스마트팜, 생육 데이터 분석, 웹 서비스로서의 데이터 분석

1. 서론

켈리 비즈니스 스쿨은 정밀농업 분야의 비즈니스 현황을 분석하고 진화 방향을 제시하면서, 정밀농업은 기존 정밀농업이 수행해 온 농기계의 자동화, 센서 기술과 농작물 작황 및 기후

의 예측 그리고 변량 시비 기술을 통한 친환경적 농업 지원에서 실시간 농업 데이터 수집과 분석으로 전환될 것임을 예측하였고[1], 이후 정밀농업은 이 예측대로 발전하고 있다.

정밀농업의 한 축인 스마트팜은 온실에 설치된 각종 센서와 장비를 통해 환경정보, 생육정보, 에너지정보, 농작업 정보 등을 통합적으로 수집하고 분석하여 최적의 생육환경을 제공할 뿐만 아니라 개별 농장들을 하나의 서버로 연계하여 생산성의 극대화를 지향한다. 온실에 설치된 센서는 생육 환경 데이터를 수집하여 데이터 수집 시스템에 일정 주기마다 전송한다. 이렇게 전송된 데이터를 기반으로 다양한 기술이 융복합된 정밀농업은 자원의 효율적 이용이 가능하게 만들 뿐만 아니라 스마트팜 농업생산 시스템을 정밀하게 자동화하여 손실

※ 이 논문은 2021년 한국정보처리학회 ACK 2021의 우수논문으로 "온실의 미기후 및 근권 데이터 활용을 위한 R 웹 애플리케이션 연구"의 제목으로 발표된 논문을 확장한 것임.

† 비 회 원 : 전북대학교 소프트웨어공학과 학사과정

†† 정 회 원 : 전북대학교 소프트웨어공학과 교수

††† 비 회 원 : 전북대학교 소프트웨어공학과 강의초빙교수

Manuscript Received : December 28, 2021

First Revision : February 3, 2022

Accepted : March 8, 2022

* Corresponding Author : Jihyun Lee(jihyun30@jbn.ac.kr)

을 감소시키고 실패 가능성을 최소화할 수 있게 해준다[2].

국외의 PRIVA, Hoogendoorn, HortiMax 등은 각종 센서를 기반으로 작물 생육 환경을 모니터링하고 다양한 현장 정보를 종합하여 보다 효과적이고 전문적인 생육 환경 설정에 대한 의사결정을 지원한다. 이 시스템들은 각종 설비의 원격 제어를 지원하는 시스템 및 관련 솔루션을 제공하고 있다. 또한, IBM, 구글, GE 등도 농작업 관련 의사결정 기술 개발에 참여하고 있다[1, 3]. 그러나 이들 시스템은 고가이거나 국내 사용자가 쉽게 접근하여 사용할 수 있는 형태가 아니다.

국내의 경우, 최근 공공데이터와 연구구역에서 발굴된 데이터 공개 정책에 따라 온실 등 작물 생육 시설로부터 수집된 데이터들이 대중에게 공개되고 있다. 그러나 관련 데이터들이 지속적으로 늘어나고 있는 반면, 데이터들의 품질에 대한 검증이나 데이터를 분석하여 작물 생육 환경을 제어하기 위한 유용한 정보로 활용하려는 노력은 부족한 실정이다[4]. 특히, 사용자가 공개된 데이터를 활용하여 작물 생장에 필요한 정보를 얻을 수 있도록 지원하는 애플리케이션의 개발은 더욱 더디다.

이에 본 논문은 각종 센서들로부터 수집된 생육환경 데이터를 기반으로 사용자가 원하는 유형의 데이터 분석을 수행해 주는 웹 기반 데이터 분석 서비스를 제안한다. 이 웹 서비스를 통하여 스마트팜 농가나 해당 분야의 연구자는 (1) 통계 분석에 대한 전문 지식이 없어도 수집했거나 보유하고 있는 공개된 원시 데이터를 토대로 통계 분석을 수행할 수 있다. (2) 개발한 웹 기반 데이터 분석 서비스는 웹 애플리케이션에 대한 수정 없이도 사용자의 요구에 따라 새로운 데이터 분석 서비스를 쉽게 추가할 수 있다. (3) 실시간성은 없지만 데이터 분석의 정확도를 높일 수 있다. 사용자는 데이터 분석을 통하여 작물 생장에 대한 정보를 얻을 수 있고, 관련 지식에 대한 축적뿐만 아니라 관련 이해관계자들의 지식을 간접적으로 수집할 수 있다.

논문의 구성은 다음과 같다: 2장에서는 관련 연구로 스마트팜을 위한 데이터 분석 지원 기술과 관련 애플리케이션의 현황을 설명한다. 3장에서는 웹 기반 데이터 분석 서비스의 설계 및 구현 결과를 기술하고, 4장에서는 실제 수집한 온실의 미기후 및 근권 데이터 분석에 제안한 웹 기반 데이터 분석 서비스를 활용한 경험을 기술한다. 마지막으로 5장에서는 결론과 향후 연구 방향을 기술한다.

2. 연구 배경 및 관련 연구

2.1 스마트팜에서의 데이터 수집 및 분석

스마트팜 기술에서 가장 중요한 기술은 데이터 수집 및 분석인데, 우리나라의 경우 유용한 농업 빅데이터가 많지 않다[5]. 데이터 수집 기술은 정밀농업이라는 개념이 등장한 1980년대 이후 꾸준히 발전되어 왔다. IoT, 클라우드, 빅데이터, 인공지능 등 관련 기술이 급격한 발전을 이루면서 최근에는 관련 데이터베이스 구축 및 데이터를 기반으로 한 효율

적 의사결정을 위한 기술들이 개발되고 있다[7]. 이에 클라이 및 코퍼레이션, 존 디어, IBM 등은 토양 정보나 기후에 대한 빅데이터 분석을 지원하여 생산성 향상 및 수확량 개선에 도움을 주고 있다[5].

유럽연합이 추진한 IoF 2020의 채소 분야 세부 프로젝트들에서 시도한 센서와 데이터 분석을 결합하여 재배 과정을 자동화하려면[6] 온실부터 노지에 이르기까지 다양한 생육환경 데이터 수집과 데이터 분석 기술이 결합되어야 한다. 농업 빅데이터 처리는 데이터 전처리와 데이터 분석의 두 흐름으로 구분할 수 있다. 데이터 전처리는 데이터의 품질 및 빅데이터 수집원과의 지속적 통합에, 데이터 분석은 다양한 이해관계자들을 겨냥한 비즈니스 모델과 솔루션 개발을 지원할 플랫폼에 중점을 두는 흐름이다[8]. 데이터 분석을 위한 솔루션 개발을 지원할 플랫폼 개발 연구들이 함께 진행되어 왔다[9-12].

이 과정에서 중요한 일은 농업인, 작물과 환경과의 관계를 연구하는 연구자들, 또는 관련 기관들 등과 같은 이해관계자들이 생육환경의 근권 정보나 미기후 상태를 확인하는 것이다. 이들 이해관계자들이 손쉽게 현재의 생육환경 정보와 생육 상태를 확인하고 의사결정을 내리도록 지원해 줄 수 있는 용이한 방법 중 하나는 웹 서비스 형태이다. 이를 통해 사용자는 공개되거나 실시간으로 수집한 생육 환경 데이터로부터 농작업에 보다 유용한 정보를 추출하고 지식을 축적할 수 있다.

2.2 웹 기반 데이터 분석 서비스

X. Hu 등[13]의 분석에 의하면 스마트팜 지원 애플리케이션은 작물 파종부터 성장, 수확, 판매에 이르는 전체 공급망에 대한 운영 지원으로 확장해 나가고 있다. 반면 작물 재배를 지원하기 위한 이론적 지식의 축적이 부족하다고 지적하고 있다. 스마트팜 기술 확산의 또 다른 중요한 장애는 이 기술을 활용한 농업인들의 관련 기술에 대한 인식이다[14]. 이는 농업인들을 포함한 이해관계자들의 지식을 자산화하고 접근이 용이한 데이터 분석 서비스를 제공하는 것이 중요함을 의미한다.

[15]는 농장으로부터 수집한 원시 데이터를 사용자의 요청에 따라 시간별, 일별, 또는 월별로 통계 분석을 수행할 수 있는 웹 서비스 플랫폼을 제안하였다. 그리고 [16]은 센서로부터 온도, 습도, 토양 수분 상태 정보를 수집하여 관수를 제어할 수 있도록 하는 웹 애플리케이션을 소개하고 있다. 이들은 데이터 수집 및 분석을 지원한다는 점에서 유사하다. [15]가 제안한 플랫폼은 온도, 습도, CO₂ 등 단일 파라미터만을 대상으로 한 추이 분석 결과를 제공하며, [16]은 생육환경 제어까지 지원하고 있다. 그러나 생육환경 제어를 위한 의사결정을 내리려면 이러한 생육환경 데이터들 간의 상관관계 또한 중요한 정보이다.

[17]은 데이터 소비자가 간단한 HTTP 요청만으로 인터넷을 통해 분석하고자 하는 데이터를 분석할 수 있도록 지원하는 웹 기반 데이터 분석 서비스를 제안하였다. 이 웹서비스는 CSV 파일 형태로 센서로부터 수집한 데이터를 입력받아 분석 도구에 전달한 후 결과를 사용자에게 제공한다. 실험 데이터

가 생육환경 데이터 분석 서비스는 아니지만, 데이터 분석을 웹 서비스화하고자 하는 시도를 했다는 점에서 의의가 있다. 그러나 이 연구는 웹 서비스로서의 데이터 분석의 가능성에만 중점을 두고 있으며 데이터 분석이 사용자에게 줄 수 있는 이득은 본 연구의 웹 서비스 설계에서 고려하고 있지 않다.

2.3 R을 이용한 데이터 분석

KDnuggets[19]의 조사에 의하면 2019년 기준으로 66%의 데이터 분석 도구들이 파이썬 API를 사용하고 있으며 2017년부터 수행된 조사 결과의 추이를 볼 때 파이썬은 매년 증가하는 추세인 반면 R[18]은 줄고 있다. 이는 많은 애플리케이션들이 기계 학습을 이용한 지능적인 의사결정 지원 요구를 반영해야 하기 때문으로 분석된다. 반면, R을 이용한 데이터 분석은 CRAN(Comprehensive R Archive Network)에서 제공하는 패키지를 사용한다. CRAN은 통계학자들의 연구와 검증을 통하여 그 정확성이 확인된 패키지이기 때문에 파이썬의 API보다 높은 신뢰성을 기대할 수 있다[20, 21].

R을 이용하여 데이터 분석을 지원하는 웹 서비스를 개발하고자 하는 경우 R Studio의 개방형 패키지인 R Shiny 웹 애플리케이션 프레임워크를 이용할 수 있다[22]. 이 패키지는 R의 모든 기능과 연산 능력이 포함되어 있으며 클라우드 또는 물리적 서버에서 활용할 수도 있다. 이 패키지를 사용하면 HTML 및 CSS를 사용하여 사용자의 정의에 따라 대시보드를 작성할 수도 있다. R Shiny의 가장 큰 장점은 사용자와의 유연한 상호 작용이다. 사용자는 필드, 텍스트 상자와 같은 다양한 위젯을 사용하여 실시간으로 값을 입력하고 기본 가정을 변경할 수 있다. 하지만 Shiny로 작성된 애플리케이션의 경우 배포에 제한이 있다. 애플리케이션의 배포는 “Shinyapps.io” 사이트를 통해서만 이루어지며, 한 달에 25시간의 사용 제한과 메모리 공간 제한 등과 같은 몇 가지 제한 사항들이 존재한다. 즉, 추가 결제를 통해 이러한 제한들을 해제해야 한다는 것과 애플리케이션이 “Shinyapps.io” 플랫폼에 종속적이라는 것이다.

따라서 본 연구에서는 R shiny를 사용하지 않았고, 웹 서비스의 실제 사용자가 농업 관련 종사자라는 점과 사용자의 분석 요구가 다양하다는 점을 고려하여 Node.js와 R을 기반으로 하는 웹 기반 데이터 분석 서비스를 개발하였다. 사용자가 필요로 하는 데이터 분석 서비스는 사용자 수만큼 다양할 수 있다. R은 스크립트 형식으로 분석 코드를 작성하기 때문에 수정이 쉬울 뿐만 아니라 새로운 데이터 분석 요구가 발생하면 스크립트를 작성하기만 하면 되기 때문에 새로운 분석 서비스를 추가하는 것이 용이하다. 이 웹 서비스를 통하여 데이터 분석 서비스를 받는 사용자는 데이터 분석에 사용할 데이터를 입력하기만 하면 되는 방식이다.

3. 웹 기반 데이터 분석 서비스의 아키텍처와 상세 설계

웹 기반 데이터 분석 서비스의 전체 개념적 아키텍처는 Fig. 1과 같이 데이터 서버, 애플리케이션 서버, 클라이언트로 구성된 3-티어 아키텍처 스타일을 가진다. 본 장에서는 제안하는 데이터 분석 서비스의 아키텍처와 세부 설계를 데이터 수집을 위한 데이터 서버와 애플리케이션 서버(백엔드)로 나누어 설명한다.

3.1 데이터 서버

데이터는 온실 및 노지의 생육환경으로부터 수집된다. 생육 환경 정보를 수집하기 위해 설치될 수 있는 IoT(Internet of Things) 디바이스는 온도, 습도, EC(Electrical Conductivity), pH(potential of Hydrogen) 등 그 유형이 제한되어 있지 않으며, 디바이스의 개수 또한 제한이 없다. 각 IoT 디바이스가 수집하는 데이터 타입은 센서 별로 보정된 값으로 디바이스가 수집하는 데이터 타입에 제한은 없다. 이에 기존 시스템은 Fig. 1의 IoT 디바이스, 데이터 서버, 애플리케이션 서버, 클라이언트 부분과 같이 멀티-티어 구조로 설계되어 있

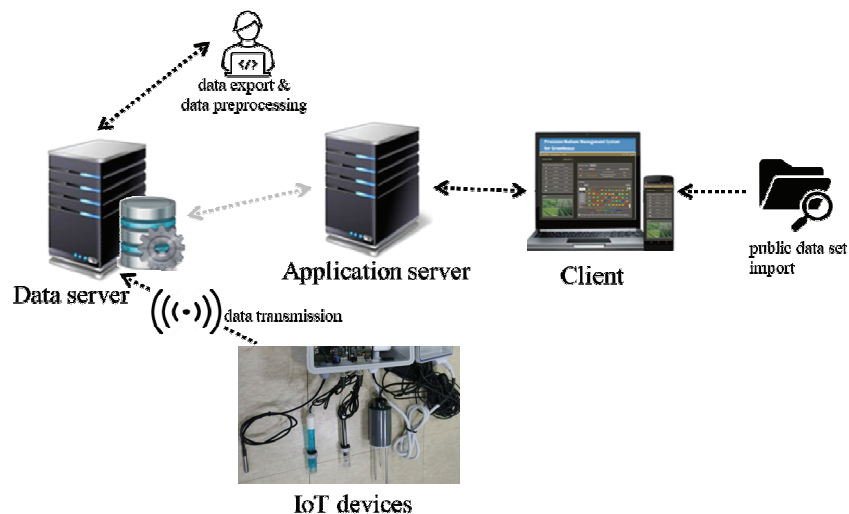


Fig. 1. Conceptual Architecture of the Web-based Data Analysis Service

다. 본 연구에서 추가하고자 하는 데이터 분석 서비스에서 분석 서비스를 제공할 데이터는 변경될 수 있다. 데이터 분석 서비스가 데이터 종류와 값의 타입, 설치된 IoT 디바이스 개수에 영향을 받지 않아야 한다. 게다가 사용자는 새로운 데이터 분석 서비스의 추가를 요구할 수 있다. 따라서 전체 시스템 설계는 다음과 같은 아키텍처 드라이버를 만족시켜야 한다:

- a) 확장성(Extendability): 사용자가 새로운 데이터 분석 서비스를 요구하면 쉽게 추가될 수 있어야 한다.
- b) 수정용이성(Modifiability): IoT 디바이스가 변경, 제거, 또는 추가되어도, 데이터 타입이 변경되더라도 나머지 시스템은 영향을 받지 않아야 한다.

신규로 추가하고자 하는 웹 기반 데이터 분석 서비스 역시 현재 운영 중인 시스템 아키텍처를 그대로 따른다. Fig. 1과 같이 분석할 데이터는 데이터 서버 tier에 두고 애플리케이션 서버 tier는 사용자 분석 요구를 디스패치하고 데이터 분석을 수행한 후 사용자가 원하는 형태로 분석 결과를 표현하여 사용자에게 전달하는 역할을 담당한다. 기존 시스템에 변화가 있거나 사용자가 직접 분석할 데이터를 입력하여 분석하고자 해도 영향을 받지 않도록 하기 위해 분석할 데이터는 데이터 서버에 별도로 저장한 후 분석하도록 설계하였다. 현재 운영 중인 시스템으로 사용자가 원하는 데이터 분석을 수행할 수 있으나 원시 데이터에는 이상치, 결측치 등이 다수 포함되어 있어 분석의 정확도가 떨어진다. 따라서, 웹 기반 데이터 분석 서비스는 데이터를 추출하여 전처리 후 데이터 저장소에 저장한 결과로 수행하도록 설계하였다. 사용자는 데이터 저장소로부터 데이터를 선택하거나 외부 공개 데이터 세트를 다운로드 하여 원하는 형태의 데이터 분석을 수행할 수 있다.

3.2 백-엔드 애플리케이션 아키텍처

R 언어는 통계 컴퓨팅, 데이터 분석, 분석 결과에 대한 그래픽 표현 및 보고를 지원하는 소프트웨어 환경이다. 그러나 R은 통계 지향 프로그래밍 언어로 여러 애플리케이션과의 연동이 용이하지 않으며, 관련 전문 지식이 부족한 농가의 개인 사용자들이 R에서 제공하는 유용한 통계 분석 기능을 능숙하게 사용하기란 쉽지 않다. 따라서 본 연구에서는 사용자가 웹을 통해서 데이터 분석을 손쉽게 수행하여 유용한 정보를 획득하고, 데이터 분석 서비스 제공자는 새로운 데이터 분석 서비스를 시스템에 쉽게 추가할 수 있도록 지원하는 웹 기반 데이터 분석 서비스 개발을 목표로 한다.

본 연구에서 구현한 웹 기반 데이터 분석 서비스는 계층 아키텍처(Layered architecture)와 MVC(Model View Control) 아키텍처 스타일에 따라 설계되었다(Fig. 2).

뷰(View)는 프론트-엔드 애플리케이션(Front-end application)에 전달될 표현 형태를 결정하며, 뷰를 통해 클라이언트가 요청한 데이터의 분석 결과를 제공한다. 컨트롤러(Controller)는 라우터 컨트롤러(Router controller)로 표현하며, 컨트롤러를 통해 클라이언트의 요청에 따라 모델(Model)로

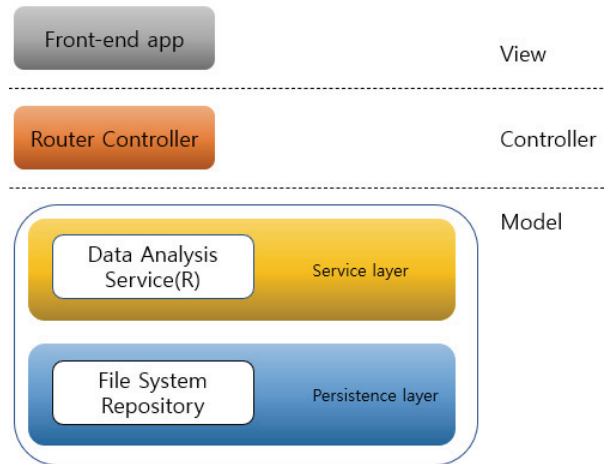


Fig. 2. Back-end Application Architecture

보낸 후, 처리된 데이터를 뷰로 반환한다. 모델은 서비스 계층(Service layer)과 영구저장 계층(Persistence layer)으로 나뉜다. 서비스 계층은 컨트롤러의 모든 요청을 처리하는 역할을 하며, 이 계층에서 사용자로부터 입력받은 데이터를 처리 및 분석하기 위한 R 스크립트를 구동하는 데이터 분석 서비스(Data analysis service)를 지원한다. 영구저장 계층은 처리할 데이터와 처리 결과를 보관하기 위한 파일 시스템 저장소의 역할을 한다.

백-엔드 애플리케이션을 개발하기 위하여 우리는 Node.js에서 서버 프로그램을 구현하는데 사용하는 웹 애플리케이션 프레임워크인 익스프레스(Express)를 사용하였다. 백-엔드 애플리케이션은 R Studio가 R 언어로 작성된 R 스크립트를 실행하는 방식인 "R CMD BATCH" 형식의 콘솔 명령어를 활용한다. 이때 사용자가 데이터를 파일 형식으로 업로드할 경우에는 클라이언트 측에서 첨부된 파일을 서버 측에서 "R CMD BATCH" 명령에 직접적으로 전달할 수 없다. 이러한 데이터 접근 문제와 기밀성 문제를 해결하기 위해서 제공 받은 데이터는 사용자의 고유한 IP주소(Internet Protocol address)를 키 값(Key value)으로 하여 서버의 파일 시스템에 저장된다. 개발한 웹 애플리케이션 서버는 IP주소를 효율적으로 처리하기 위해 정수형(number type)으로 변환한 후, 이 IP주소와 파일명을 매개변수로 하여 선택한 분석 서비스에 맵핑되는 R 스크립트를 실행한다.

3.3 웹 애플리케이션의 파일 시스템 개요

Fig. 3은 웹 기반 데이터 분석 서비스의 파일 시스템 구조를 보여준다.

파일 시스템의 구조는 Node.js 익스프레스 프로젝트의 기본적인 구조와 동일하다. "bin"은 서버를 실행하는 스크립트로 포트 번호의 변경과 같은 행위를 다룬다. "routes"는 페이지 라우팅과 관련된 파일을 저장하며 URL 별로 실행되는 실제 서버 로직을 구현하고 있다. 그리고 프론트엔드를 위한 "views" 폴더에는 ".html"과 같은 템플릿 파일들을 저장하며

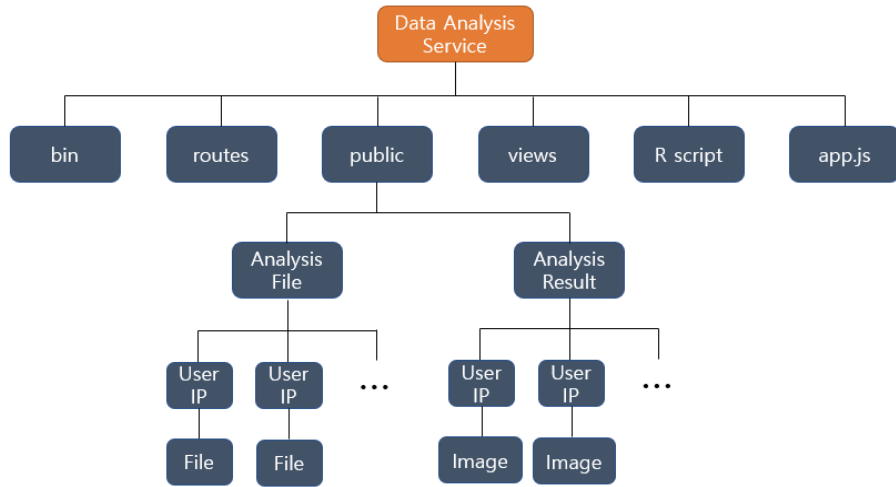


Fig. 3. File System Structure for the Web Service

웹 서버 사용 시 이 폴더의 파일들을 사용한다. 그리고 데이터를 분석하고 처리하기 위한 R 스크립트가 존재하며, 또한 “app.js”라는 프로젝트의 미들웨어 관리를 통한 핵심적인 서버 역할을 하는 라우팅의 시작점 파일이 존재한다. 그리고 외부에서 접근 가능한 파일들을 저장하는 “public” 폴더가 존재한다. 이 “public” 폴더에는 다양한 자원들이 적재된다. 본 연구에서는 사용자의 입력 데이터를 저장하기 위한 “Analysis File”과 분석 결과를 저장하기 위한 “Analysis Result”로 각 데이터의 사용처를 기준으로 나누어 구성하였다. “Analysis File”과 “Analysis Result”는 사용자의 IP주소를 구분자로 하여 저장되고, 데이터에 접근할 때도 사용자의 IP주소를 통해 이루어진다.

4. 웹 기반 데이터 분석 서비스 활용

이 장에서는 개발한 데이터 분석 웹 서비스를 실제 IoT 디바이스들이 설치된 온실로부터 수집한 데이터 분석에 적용한 사례를 설명한다.

4.1 데이터 수집

본 활용에서는 온실로부터 수집한 미기후 및 근권 데이터를 이용하여 수행되었다. 현재 설치된 IoT 디바이스는 온도, 습도, 조도, CO₂, 배지 무게 변화, EC, pH 관련 데이터를 수집하는 센서들이며 Fig. 1의 IoT 디바이스 부분에서 보여지는 것과 같이 구성되어 있다. IoT 디바이스들은 매 1분마다 데이터를 수집하여 전송하고, 그 결과는 데이터베이스에 저장된다. 현재 수집된 미기후 및 근권 데이터 추이는 사용자가 스마트팜 모니터링 플랫폼¹⁾을 통하여 실시간으로 확인할 수 있다.

1) Smart Software Lab. Inc.의 스마트팜 모니터링 플랫폼은 <http://www.ezsmartfarm.com/>에서 확인할 수 있다.

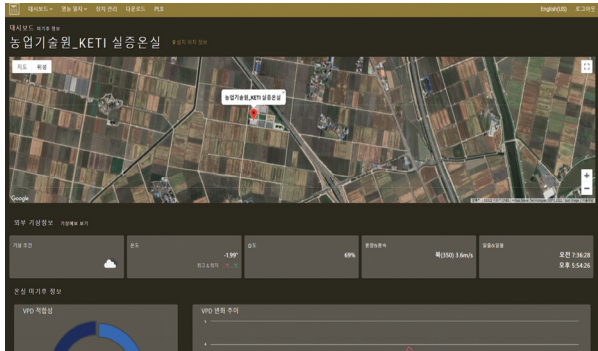
이 서비스 플랫폼은 온실 정보를 실시간으로 확인할 수 있으나 사용자의 필요에 따라 다양한 데이터 분석 서비스는 제공하고 있지 않다. 본 논문이 기존 서비스의 확장을 위해 제안한 웹 기반 데이터 분석 서비스의 활용성을 검증하기 위해 기존 플랫폼과 통합하지 않고, Fig. 4와 같이 데이터베이스에 누적된 데이터를 원시 데이터 상태의 “.csv” 파일로 저장하였다. 사용자는 저장된 파일을 액세스하여 원하는 데이터 분석을 수행할 수 있다. 그러나 Fig. 1과 같이 만일 사용자가 외부 데이터 제공자로부터 전처리된 데이터(예, 공공 데이터 세트)를 가지고 있다면 해당 데이터를 분석의 입력으로 활용할 수 있다.

데이터는 온실에서 재배 중인 파프리카, 토마토, 감초의 세 가지 작물에 대하여 수집한 미기후 및 근권 데이터이다. 세부 온실 환경과 수집 데이터, 수집 방법은 다음과 같다:

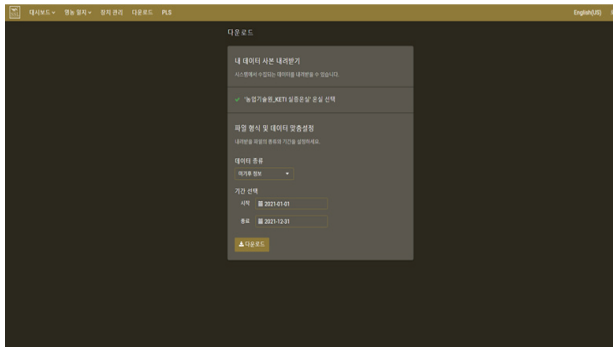
- a) 데이터 수집 온실 유형: 배지에 양액을 점적 관수하여 재배하는 온실과 스프링클러를 이용하는 토경 재배 온실 두 가지 유형을 기본으로 진행하며, 토경 재배 온실의 경우에는 온실의 특성상 미기후 데이터만 수집하였다.
- b) 온실 내 미기후 데이터 수집 항목: 분당 온도, 습도, 조도, CO₂ 값을 센서로부터 수집하였다.
- c) 배지 내 근권 데이터 수집 항목: 분당 배지 무게 변화, 지온, EC, pH 값을 센서로부터 수집하였다.
- d) 데이터 수집 방법: 파프리카, 토마토, 감초를 재배하는 세 개의 온실로부터 데이터를 수집(약 100만 레코드, 1,300 만건)하였으며, 분석을 위하여 격주로 CSV 파일로 내려받아 전처리를 수행하였다.

4.2 웹서비스를 이용한 데이터 분석

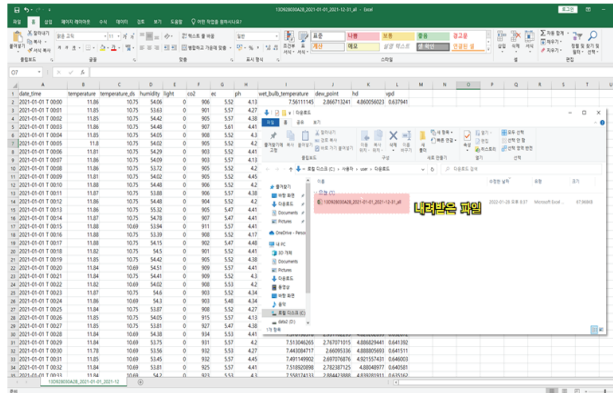
사용자는 본 연구에서 개발한 웹 기반 데이터 분석 서비스를 통해 원하는 분석 결과를 얻을 수 있다. Fig. 5는 수집된 데이터에서 온도 항목에 대한 데이터의 분포를 히스토그램으로 시각화한 것이다. 온실 온도는 대부분의 경우 15℃-20℃



(a) Screenshot of the user interface



(b) Downloading raw data for a particular time period



(c) Extracted raw data

Fig. 4. Data Extraction from the Smart Farm Monitoring Platform

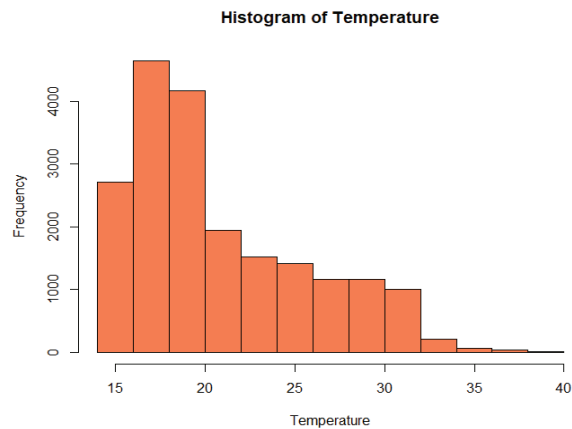


Fig. 5. Analysis of Temperature Values for Two Weeks

로 유지되었으며 30℃ 이상으로 높아지는 경우도 자주 있었음을 확인할 수 있다. 사용자는 온도뿐만 아니라 습도, CO₂ 농도, 광량 등 수집 데이터의 분포를 확인할 수 있다.

Fig. 6은 수집된 데이터에서 온도 항목에 대하여 날짜별, 시간 흐름에 따른 온도의 변화를 꺾은선 그래프로 나타낸 것이다. 대체로 15℃에서 30℃ 내에서 일교차로 인해 유사한 모양의 파형이 생기는 것을 확인할 수 있다.

Fig. 7은 수집 데이터 항목 간에 서로 어떠한 관계를 가지고 있으며, 어느 정도의 관련성이 있는지를 상관분석 차트를 통해 시각화한 것이다. 온도와 습도의 경우 -0.88로 높은 음의 관련성을 가지고 있다. 이것은 온도가 올라갈수록 습도는 내려가며, 반대의 경우도 동일하다는 뜻이다.

Fig. 8은 사용자가 새로운 데이터 분석으로 요구한 온실 내부와 온실 외부의 온도와 습도 간의 상관관계 분석 결과이다. 새로운 데이터 분석 요구는 R 스크립트를 새로 추가하기만 하면 된다. 생성된 히스토그램 및 차트를 바탕으로 웹 상에서 정렬하여 사용자에게 제공한다.

Fig. 9는 웹 기반 데이터 분석 서비스의 사용자 인터페이스이다. 사용자가 Fig. 9의 왼쪽 부분과 같은 인터페이스를 통하여 데이터 파일과 분석할 데이터 항목을 지정하면 관련된 데이터의 처리와 분석이 웹 애플리케이션 서버에서 수행되고 분석의 결과는 Fig. 9와 같이 오른쪽에 출력된다.

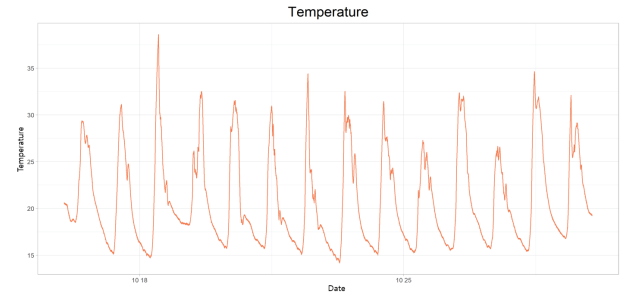


Fig. 6. Temperature Trend Analysis

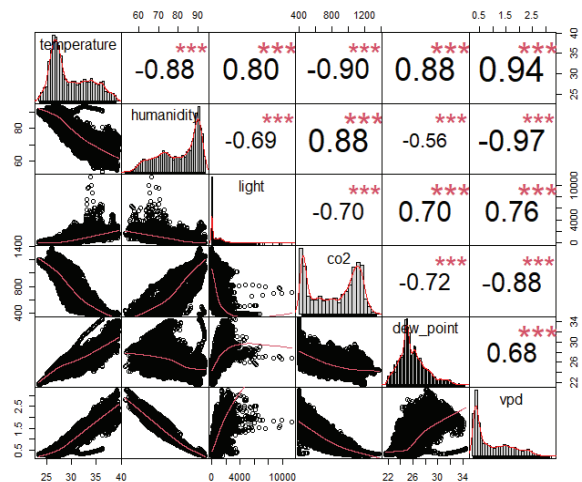


Fig. 7. Correlation Analysis Between Data Items

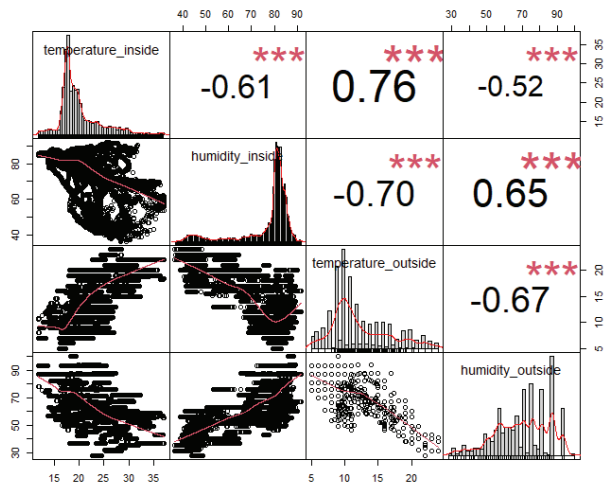


Fig. 8. Correlation Analysis Between Temperature and Humidity of Inside and Outside of Greenhouse

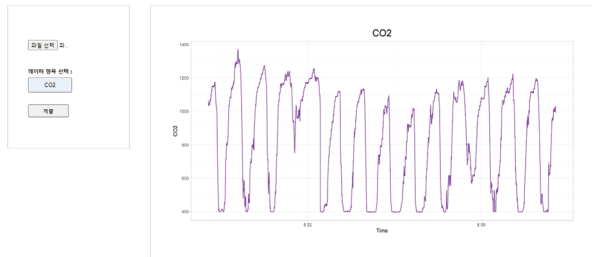


Fig. 9. User Interface of the Web-based Data Analysis Service

5. 결론 및 향후 연구

지금까지 작물 성장 관련 데이터 처리 및 분석을 위한 웹 기반 데이터 분석 서비스를 설계하고 구현한 결과를 기술하였다. 웹 기반 데이터 분석 서비스는 확장성과 수정용이성을 고려하여 멀티-티어 아키텍처 스타일을 채택하였으며 사용자의 다양한 시각적 표현 요구를 용이하게 반영할 수 있게 하기 위하여 MVC 아키텍처 스타일을 채택하였다. 또한, 개발한 웹 기반 데이터 분석 서비스는 R의 장점인 다양한 통계 분석 및 시각화 능력을 기반으로 사용자가 웹을 통하여 분석하고자 하는 데이터를 전송하기만 하면 데이터의 분석 결과를 쉽게 확인할 수 있도록 하였다. 이 방법은 데이터 분석 유형에 대한 요구는 사용자 수만큼 다양할 수 있기 때문에 새로운 요구를 기존 구현에 영향을 미치지 않으면서 빠르게 추가할 수 있다는 장점이 있었다. 향후 연구로 개발한 웹 기반 데이터 분석 서비스를 기존 생육 환경 모니터링 시스템과 통합할 계획이다.

References

[1] X. Pham and M. Stack, "How data analytics is transforming agriculture," *Business Horizons*, Vol.61, No.1, pp.125-133, 2018.

[2] Rural Development Administration, "A study on productivity improvement model and gathering big data of smart farm in vegetable grown in facilities," Final Report (TRKO201900 016054), 2019.

[3] Y. K. Yim, "Precision agriculture: Agricultural management system for minimizing environmental pollution while maximizing crop productivity," Innovative Growth Item Report, 2021.

[4] Y. C. Choi, "Smart farm and big data," *TTA Journal*, Vol.11/12, 2018.

[5] H. Yeo, "Application status of agricultural big data in foreign countries," *World Agriculture*, No.226, pp.37-52, 2019.

[6] IoF 2020, Digital ecosystem utilization [Internet], <https://www.iof2020.eu/use-case-catalogue/vegetables/digital-ecosystem-utilisation>, last accessed in Dec. 2021.

[7] J. A. Delgado, N. M. Short Jr., D. P. Robers, and B. Vandenberg, "Big data analysis for sustainable agriculture on a geospatial cloud framework," *Frontiers in Sustainable Food Systems*, Vol.3, pp.1-13, 2019.

[8] S. Wolfert, L. Ge, C. Verdouw, and M.-J. Bogaardt, "Big data in smart farming - a review," *Agricultural Systems*, Vol.153, pp.69-80, 2017.

[9] K. Park, M. C. Nguyen, and H. Won, "Web-based collaborative big data analytics on big data as a service platform," in *Proceedings of the 17th International Conference on Advanced Communication Technology (ICACT)*, Pyeong-Chang, Korea, pp.564-567, 2015.

[10] S. Barmounakis et al., "Management and control applications in Agriculture domain via a Future Internet Business-to-Business platform," *Information Processing In Agriculture*, Vol.2, pp.51-63, 2015.

[11] S. Li and Y. Zhang, "Construction of big data processing platform for intelligent agriculture," in *Proceedings of the International Conference on Big Data Analytics for Cyber-Physical-Systems*, pp.1206-1212, 2020.

[12] S. Stoyanov, J. Todorov, I. Stoyanov, V. T.-Komsalova, L. Doukovska, and A. Dukovski, "ZEMELA - an intelligent agriculture platform," in *Proceedings of the Big Data, Knowledge and Control Systems Engineering (BdKCSE)*, Sofia, Bulgaria, pp.1-6, 2021.

[13] X. Hu, L. Sun, Y. Zhou, and J. Ruan, "Review of operational management in intelligent agriculture based on the Internet of Things," *Frontiers of Engineering Management*, Vol.7, No.3, pp.309-322, 2020.

[14] Market and Markets, "Agriculture analytics market by application area (Farm analytics, Livestock analytics, and Aquaculture analytics), component (solution and services), farm size (small, medium-sized, and large), deployment type, and region-global forecast to 2025," 2019.

- [15] M. Colezea, G. Nusat, F. Pop, C. Negru, and A. Dumitrascu, "CLUeFARM: Integrated web-service platform for smart farms," *Computers and Electronics in Agriculture*, Vol.154, pp.134-154, 2018.
- [16] J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbang-pong, A. Wanichsombat, and P. Nillaor, "IoT and agriculture data analysis for smart farm," *Computers and Electronics in Agriculture*, Vol.156, pp.467-474, 2019.
- [17] A. Ahrabian, S. Kolozali, S. Enshaeifar, C. Cheong-Took, and P. Barnaghi, "Data analysis as a web service: A case study using IoT sensor data," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, New Orleans, Louisiana, USA, pp.6000-6004, 2017.
- [18] John Verzani "Using R for Introductory Statistics," 2nd Ed. CRC Press.
- [19] KDnuggets, "Data Science Tools Popularity, animated," last accessed on 29 Jan. 2022, [Internet], <https://www.kdnuggets.com/2020/06/data-sci-ence-tools-popularity-animated.html>.
- [20] CRAN, "The Comprehensive R Archive Network," last accessed on 29 Jan. 2022, [Internet], <https://cran.r-project.org/>
- [21] Python 3.10.2, "Distributing Python Modules," last accessed on 29 Jan. 2022, [Internet], <https://docs.python.org/ko/3/distributing/index.html>
- [22] Carson Sievert, "Interactive Web-Based Data Visualization with R, plotly, and shiny," CRC Press, 2019.



정 지 민

<https://orcid.org/0000-0002-7335-2977>
e-mail : jimin606@naver.com
2017년 ~ 현 재 전북대학교
소프트웨어공학과 학사과정
관심분야 : Web Application & Data
Analysis



이 지 현

<https://orcid.org/0000-0003-4512-806X>
e-mail : jihyun30@jbnu.ac.kr
1993년 전북대학교 정보통신공학과(학사)
2000년 전북대학교 전자계산교육(석사)
2005년 전북대학교 컴퓨터과학과(박사)
2016년 ~ 현 재 전북대학교
소프트웨어공학과 교수
관심분야 : Software Product Line & Architecture
Reconstruction



노 혜 민

<https://orcid.org/0000-0003-1150-3570>
e-mail : hmino@jbnu.ac.kr
2000년 전북대학교 컴퓨터과학과(학사)
2002년 전북대학교 전산통계학과(석사)
2006년 전북대학교 컴퓨터통계정보학과
(박사)
2011년 ~ 현 재 전북대학교 소프트웨어공학과 강의초빙교수
관심분야 : Precision Agriculture & Big Data Analysis