

# 시설 원예작물 생장예측모델을 위한 데이터 수집 시스템 설계사례

안성철, 김희성, 권혜은  
가림정보기술(주)  
e-mail:sungchul@galimit.com

## Design Case on Data Collection System for the GreenHouse Horticultural Crops Growth Forecasting Model

Sung-Chul Ahn, Hee-Sung Kim, Hye-Eun Kwon  
GaLim Information Technology

### 요 약

생장예측모델이란 작물의 생장 시스템 내에서 일어나는 기작이나 생산과정을 수식으로 묘사하는 것이다. 신뢰성 있는 생장예측모델을 만들기 위해서는 생장과 관련된 대량의 데이터가 필요하다. 본 논문에서는 IT와 농업을 융합한 시설 원예작물 생장예측모델을 위한 생장 및 생장환경 데이터 수집 시스템 설계사례를 소개하고자 한다.

### 1. 서론

그동안 우리나라의 시설재배는 크게 발전하였다. 재배면적이 증가하고 시설의 구조형태나 자재와 함께 재배기술이 개선되어 품질이 향상되고 수량도 상당히 증가하였다. 그러나 아직도 선진국에 비하면 개선의 여지가 많이 남아 있다. 시설재배는 다양하면서 수준 높은 재배기술이 요구된다. 일부 재배자는 스스로 연구하고 분석하면서 시설을 경영하기도 한다. 경우에 따라서는 독자적으로 개발한 독특한 기술을 가지고 있을 정도이다. 그러나 전반적으로 기술수준이 대단히 낮아 각 작물의 생산능력을 충분히 발휘시키지 못하고 있다.[1] 철저하고 정확한 환경관리기술과 시설의 특이환경에 적합한 재배관리기술을 터득하여 생리장해 등을 억제하고 생산성을 향상시키기 위한 신뢰성 있는 생장예측모델을 만들어야 한다.

지난 20여 년 동안 컴퓨터의 급속한 발달로 아주 복잡하고 큰 생장예측모델도 손쉽게 구동할 수 있게 되었고 단지 일부 과학자들의 전유물이었던 컴퓨터 사용이 일반화됨에 따라 작물모델의 이용성도 확대 되었다. 생장예측모델은 생육 시스템 내에서 일어나는 기작이나 생산과정을 수식으로 묘사하는 것이다. 이런 설명과 묘사는 반드시 과학적인 이론과 가설에 근거하여 이루어진다. 모델을 개발하기 위해서는 작물이 자라는 시스템을 잘 분석하고 그 기작이나 과정이 수치적으로 계산되어서 양적인 변화를 나타낼 수 있어야 한다.[2]

이와 같은 생장예측모델을 만들기 위해서는 실제 데이터를 바탕으로 통계적으로 유의성 있는 데이터를 도출하여야 한다. 그러나 데이터를 실험실에서 직접 작물을 재배하

여 데이터를 수집하기 위해서는 많은 시간과 노력이 필요하며 이는 생장예측모델을 연구 개발하는 실험실에서는 한계가 따르기 때문에 최근 증가 추세에 있는 시설작물 재배 농가의 도움이 필요하다.

본 논문에서는 다양한 시설작물 재배 농가에서 작물의 생장과 관련된 데이터를 효율적으로 수집을 위한 시스템 설계 사례를 소개한다.

### 2. 관련 연구

#### 2.1. 데이터 마이닝

데이터 마이닝은 대규모 데이터 저장소에서 유용한 정보를 자동적으로 탐색하는 과정으로 데이터베이스를 검색하여 새롭고 유용한 패턴을 탐색하기 위해 적용된다. 데이터 마이닝의 작업은 다른 속성들의 값을 기반으로 특정 속성의 값을 예측하는 예측작업과 데이터에 숨어있는 관련성을 요약하여 패턴을 찾는 서술작업으로 구분된다.[3] 예측작업은 연속형 값에 대한 함수를 모형화 하는 것으로 수치예측을 위해 가장 많이 사용되는 통계적 기법은 회귀분석이다. 선형회귀분석은 두 개의 속성에 하나의 속성이 다른 하나의 속성을 예측하는데 이용될 수 있도록 최적의 직선식을 찾는 것을 포함한다.[4]

#### 2.2 작물의 생장

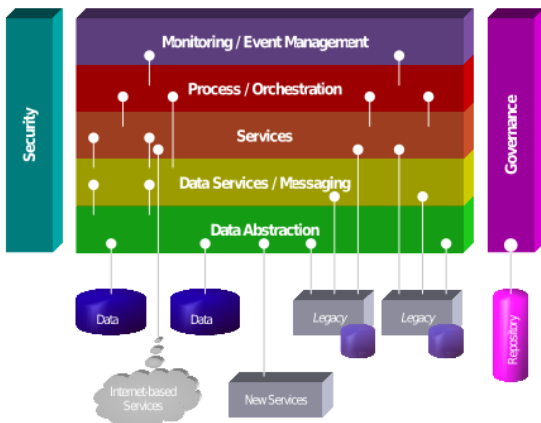
작물의 생산성에 영향을 미치는 요인은 크게 3가지 범주로 나눌 수 있다. 첫 번째 범주는 최대 생산성을 결정하는 요인들로 빛, 온도, 그리고 작물의 생리적 특성 등이다. 두 번째 요인으로는 수량 감소를 유발하는 요인들로 물과 양

분의 이용성 등이며, 세 번째 범주는 잡초, 병, 충, 해일 같은 재해들이다.[2]

농업 분야에서 작물의 생장이나 과수별 만개기, 생산량 등에 대한 예측은 각 연구자가 수많은 실험을 통해 데이터를 얻어내고, 얻어낸 데이터의 분석을 통해 모델식을 도출한 다음 현재의 상태를 입력으로 그 결과를 계산해 내는 방식으로 진행된다.

### 2.3 Service Oriented Architecture

서비스 지향 아키텍처(Service Oriented Architecture, SOA)란 대규모 컴퓨터 시스템을 구축할 때의 개념으로 업무상에 일 처리에 해당하는 소프트웨어 기능을 서비스로 판단하여 그 서비스를 네트워크상에 연동하여 시스템 전체를 구축해 나가는 방법론이다. 업무 처리 변화를 시스템에 빠르게 반영하고자 하는 수요를 대응하기 위해 2004년부터 IT업계에서 주목을 하고 있다. SOA의 대표적인 예인 SOAP기반의 웹 서비스에서는 서로 다른 이용자들이 서로 다른 방식으로 서비스와 의사소통을 하면서도 통합 관리되는 서비스들을 사용할 수 있다. [5]



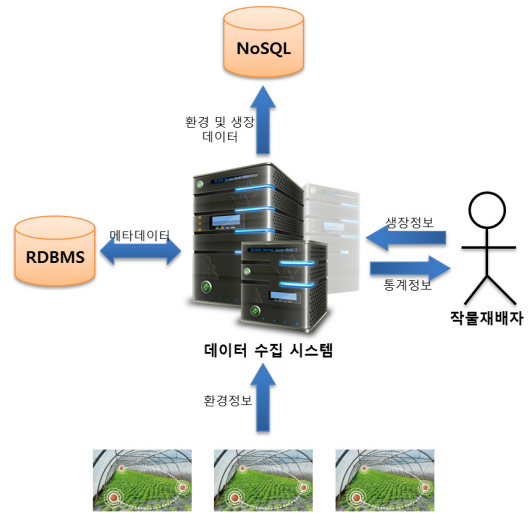
<그림 1> SOA meta-model, The Linthicum Group, 2007

본 논문에서는 추후 서비스의 확장성을 고려하여 서비스 지향 아키텍처를 채용하여 설계 하였다.

## 3. 설계사례

### 3.1. 시스템 구조

다음은 데이터 수집 시스템 전체 개념도이다. 시스템은 시설 하우스에서 환경정보와 작물재배자로부터 생장정보를 받고 데이터베이스에 수집된 값을 저장한다.



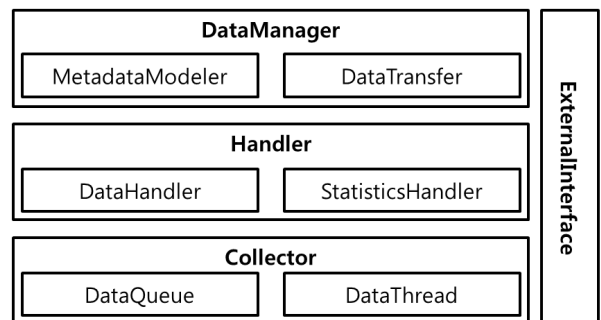
<그림 2> 데이터수집 시스템 개념도

시설환경에서는 작물의 생장과 밀접한 관련이 있는 자동으로 수집이 가능한 환경정보 데이터를 받으며 작물재배자로부터 작물의 생장상태 또는 생산량 등 자동으로 수집할 수 없는 데이터를 받고 실시간 통계정보를 제공한다.

실제 데이터를 처리하는 데이터 수집 시스템은 다수의 시설환경과 해당 시설환경을 관리하는 사용자로부터 받는 대량의 데이터를 효율적으로 수집하여 데이터베이스에 저장하며 이렇게 수집된 대량의 데이터는 추후 연구기관에서 생장예측모델을 도출하는데 사용될 수 있다.

생장예측모델을 도출하는데 사용될 생장데이터를 저장할 데이터베이스는 무수히 많은 정형 또는 비정형 데이터 세트를 유연하고 더욱 빠르게 처리하기 위해 NoSQL 데이터베이스와 관계형 데이터베이스를 병행하여 사용한다.

다음은 데이터 수집 시스템의 전체 구조이다. 시스템은 다음 그림과 같이 구성된다.



<그림 3> 데이터수집 시스템 구조

ExternalInterface는 시설환경 데이터와 사용자의 입력 데이터를 받기위한 외부 시스템과의 인터페이스를 담당하는 모듈로 실제 연동은 SOAP(Simple Object Access Protocol) 메시지를 주고받는 WebService 인터페이스를 제공한다.

DataManager는 수집되는 데이터를 관리하는 모듈로 비

정형 데이터를 수집할 수 있도록 메타정보를 정의하는 MetadataModeler와 수집된 데이터를 저장소로 전송하는 DataTransfer로 구성된다.

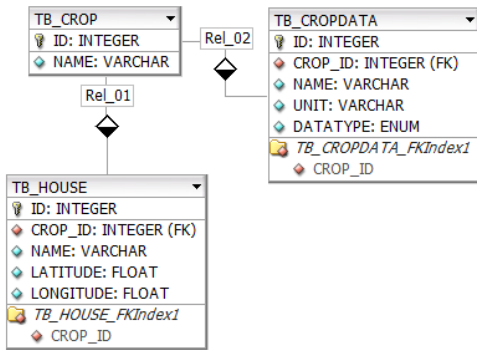
Handler는 시설에서 수집된 데이터의 처리를 담당하는 DataHandler와 사용자의 입력으로 받아온 성장정보를 처리하는 StatisticsHandler로 구성되며 DataHandler는 데이터 처리하는 기능을 담당하며 StatisticsHandler는 통계정보를 제공한다.

마지막으로 Collector는 DataQueue와 DataThread를 사용하여 시스템에서 대량 데이터의 수집과 빠른 응답속도를 위해 존재한다.

### 3.2. 데이터 모델

다수 작물의 성장예측모델을 도출하기 위해서는 작물별로 성장에 영향을 주는 환경 및 성장상태를 구분하여 데이터를 수집해야 하며 수시로 변경될 수 있어야 한다.

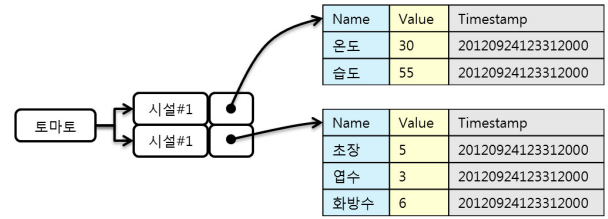
다음은 수집될 데이터의 메타정보를 나타낸다.



<그림 4> 데이터 메타모델

메타모델은 각기 다른 작물의 비정형 데이터세트를 저장하고 관리하기 위하여 정의된다. TB\_CROP은 데이터를 수집하는 작물의 종류를 나타내며 TB\_HOUSE는 작물이 어느 위치에 있는 시설에서 재배하는지를 나타내는 정보를 가지며. 마지막으로 TB\_CROPDATA는 작물의 환경 및 성장정보 항목을 나타낸다. 구축된 메타모델은 예를 들어 “토마토 작물은 전라남도 담양군 담양읍 향교리 산 37-6번지 3번 하우스에서 재배되며 온도, 습도 등의 환경 정보와 초장, 엽수, 화방수 등의 성장정보를 수집 한다”라는 문장으로 설명될 수 있다.

메타모델을 기반으로 실제 수집될 데이터는 다음과 같은 형태로 수집되며 NoSQL 데이터베이스를 사용하여 저장한다.



<그림 5> 수집 데이터

### 3.3. 데이터 수집 시스템 서비스

실제 작물의 생육 정보를 수집하기 위해서는 시설을 관리하고 작물을 재배하는 작물재배자로부터 직접 수집되어야 한다. 작물재배자의 원활한 협조를 위해서는 작물 재배에 도움이 되는 통계정보를 제공해야 한다. 데이터 수집 시스템은 대량의 데이터를 기반으로 생육정보를 입력하면 시스템에서 필요한 정보는 수집하고 현 시점의 통계를 산출하여 필요한 시점에 농가에 제공한다.

### 3.4. 향후 활용방안

대량의 데이터가 수집이 완료되면 우선 데이터 중 정규 분포를 따르는 불필요한 데이터를 정제하고 빅데이터 분석 기술 중에 하나인 군집 분석을 활용하여 작물생장에 가장 최적화된 재배패턴을 발굴하여 재배작물의 결과 농가 소득에 도움을 줄 수 있을 것이다.

## 3. 결론

본 논문에서는 시설 원예작물 성장예측모델을 위한 데이터 수집 시스템의 설계사례를 소개하였다. 현재는 데이터를 효율적으로 수집하기 위한 시스템 설계단계이며 실제 구현 단계에서 기능 및 성능 검증이 필요한 단계이다. 또한 작물재배자의 많은 협조를 위하여 실제 필요로 하는 통계정보를 도출할 계획이다.

### 참고문헌

- [1] 이병일, 문원, 손정익 공저 “시설재배학”
- [2] 임준택, “작물생육 모델링의 이론과 실제”
- [3] 용환승, 나연목 외 역, “데이터 마이닝”
- [4] Jiawei Han, Micheline Kamber, “Data Mining”
- [5] 한국정보통신기술협회(TTA) “용어사전”