A Flexible, Cloud-based Sensor Data Service Platform

Jaewoo Shin[1], Lan Zhao[2], Carol Song[2], Jian Jin[3], Zhihang Song[3], Liangju[3] Wang[3], Ziling Chen[3]

1. Department of Computer Science, Purdue University, West Lafayette, IN 47907
2. Rosen Center for Advanced Computing, Purdue University, West Lafayette, IN 47907
3. Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN 47907

Objectives

The overall objective is to develop an integrated service platform for a diverse group of agriculture related communities to collect plant health sensing data and use it in predictive modeling, data analysis and decision support:

- Handheld crop leaf health sensor development
- Improve plant physiological prediction models by including analysis of spatial distribution of stresses on the leaf
- Digital agricultural map service with "Sensors + Smartphones + Service platform" for the U.S. agricultural community
- Improve the phenotyping result by quantitatively analyze how the phenotyping result is impacted by time, location and environmental conditions

Challenges

Although the amount of plant data collected in the fields has grown rapidly in the past decades, the data transfer and management methods are lagging behind.

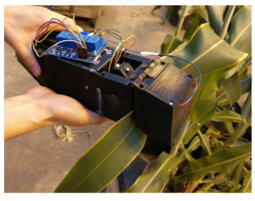
- Researchers in agricultural community still rely on the old way: saving data to local disks in simple text files or spreadsheets.
- Manipulating those data for analysis causes inaccurate or lost data and low level of productivity.

Our Solution

Develop a cloud-based data service platform for field sensor data ingestion, management, processing, and analysis.

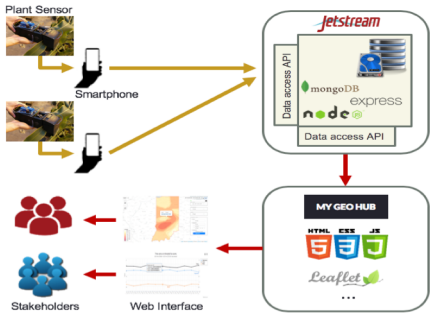
- Built on XSEDE ,HUBzero [1], and GABBs [2]
- Query, aggregate and visualize sensor data at different temporal and spatial scales in real time
- Flexible platform, can be applied to support similar use cases that utilize field data collected by mobile devices, a practice being adopted by a growing number of research projects.
- Has been used to support plant phenotyping sensor research in the field and decision making by farmers and stakeholders

Handheld Leaf Scanner



- A new low-cost precise portable hyperspectral plant sensor device easy to use by most people in the agriculture community.
- Capable of non-destructively measuring plant drought and nutritional stresses, disease symptoms caused by insects, pathogens, and various chemical spraying effects
- Report data for plant's chlorophyll level, nitrogen level, leaf moisture, chemical compounds' mode of actions, etc.

Current Implementation



The platform is built based on an open source software stack

Server backend:

- MongoDB[3] for sensor data management
- Node.js[4] for geospatial data preprocessing and data access REST API
- Deployed on the XSEDE's JetStream VM environment

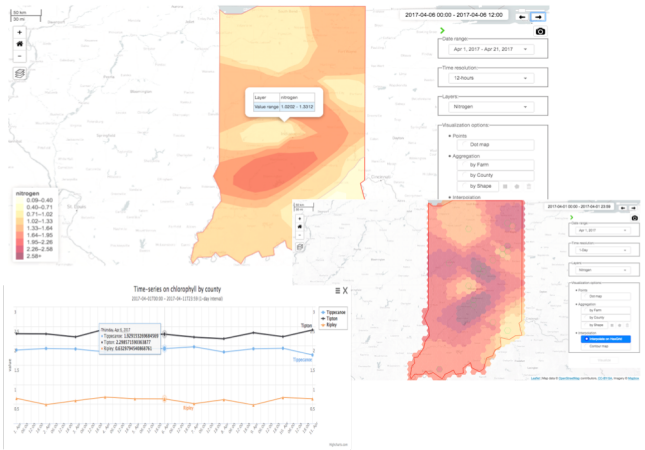
Devices in the fields:

- An Android app installed on user's smartphone which preprocess the sensor data and extract plant health measurements
- Ingest data to the MongoDB via its data service end points

Visual analytics web UI:

- A web component deployed on **mygeohub.org** [5], a HUBzero platform for online visual analytics application
- Dynamically query the MongoDB for sensor data based on time interval, location, user ID, device ID, project name, plant health variables
- Aggregate and visualize plant health data such as chlorophyll, fluorescence, nitrogen, water, normalized difference vegetation index (NDVI), nitrogen reflectance index (NRI), and soil and plant analyzer development (SPAD) at different temporal and spatial scales in real time.

Visual Analytics Web Interface



- Real time visualization and decision support information
- Stakeholders include farmers, crop re- searchers, insurance companies, local government and industries.
- Generate a Nitrogen contour map of a farmer's field and suggest to the farm owner potential hot spots
- Aggregate the data at different scales such as by county, district, or state, and alert government officials the areas that are under stress and need attention.

Future Plan

To support information collection, analysis, and decision making interactions

- Disseminate devices to farmers, researchers, and research labs in industry
- Plant health data will be collected from the leaf scanner, processed as it arrives, and used to construct a geo-referenced crop health data repository.
- Relevant environmental data for the corresponding location and time will be used to continuously calibrate prediction models and generate a comprehensive set of diagnosis, prediction and risk mitigation reports.
- The output of model predictions will be used to generate real-time visualization and decision support information for people who are in various agricultural community.
- A light-weight prediction model will be installed on farmer's mobile device to provide immediate feedback on the health status of the plant leaf scanned and suggestions on mitigation strategies.
- Apply this generic platform in use cases from other science and engineering domains.

Conclusions

Our service platform combined with the handheld sensor and environmental data will provide real-time plant growth data services and decision support for not only ordinary farmers and agricultural researchers, but also the agricultural community such as Ag insurance, policy making, climate change monitoring, market prediction, and so on.

References

1. HUBzero, https://hubzero.org
2. GABBs: Geospatial Data Analysis Building Blocks, https://mygeohub.org/groups/gabbs
3. MongoDB Inc., MongoDB, https://www.mongodb.com
4. Node.js Foundation, Node.js, https://nodejs.org
5. MyGeoHub: https://mygeohub.org

Acknowledgements

This work is part of the GABBs project, supported by the National Science Foundation (Award No. 1261727)

NSF G·A·B·B·S geospatial data analysis building blocks

XSEDE Extreme Science and Engineering Discovery Environment

HUB

MyGeoHub

PURDUE UNIVERSITY