CS486C – Senior Capstone Design in Computer Science

Project Description

Project Title: Project SilvaFlux: Big data computing and interface for tropical forest regeneration	
Sponsor Information: 	Alexander (Allie) Shenkin, Assistant Research Professor SICCS alexander.shenkin@nau.edu https://al.shenkin.org

Project Overview:

Reforestation is a critical activity for limiting global warming. If we reforest to the maximum realistic extent possible, we can limit peak global warming by 10%. Perhaps the best tool to encourage reforestation is by making it profitable on the carbon market.

The carbon market works in the following way: project developers invest in reforesting land, and they then sell the carbon that the new trees capture and store in their wood (aka carbon credits) to companies who want to reduce their carbon footprints. Unfortunately, while this has enabled some reforestation to occur, it just isn't profitable enough at current carbon prices to encourage widespread reforestation, especially in the tropics. We aim to address this shortfall, and this project will play an important role in that effort.

Figure 1. A 3D scan (combined TLS and SfM) of the world's tallest tropical tree.

By combining novel 3-dimensional models of trees (fig. 1) and forests with new ecological data, we have discovered a new climate cooling service provided by forests. This service increases our estimates of the carbon taken up by tropical reforestation projects by approximately 30%. Thus, we aim for these projects to be able to sell 30% more carbon credits, and hence become more profitable. In turn, we hope that reforestation will significantly expand as a result.

The process we've discovered is still confidential (you'll find out what it is if you join this project!), but we have all the ecological models together to estimate its magnitude across the globe. What is really needed now is a cutting-edge computing architecture, UI, and API for clients to use to get these estimates for their project areas (or potential project areas) in a fluid, dynamic, and secure way.

The estimates require crunching very large datasets of global scale raster data. As such, you will learn how to manage and analyze 'big data' datasets efficiently. Getting those critical estimates back and forth from a

server (Monsoon) to a front-end will require architectural decision making. Designing the UI to be slick, user friendly, and attractive will require coding of dynamic elements. Designing the API will give you experience in real world API design to make them logical, functional, and secure.

We currently have a prototype hosted on Google Earth Engine, but it is too slow (that gives you an idea of the amount of data we're talking about), and not flexible enough for our use. Your application and architecture will go beyond what we have been able to do so far.

A bit about me: I study the relationship between forest structure and forest function, mostly in the tropics but also in temperate forests such as those here in Flagstaff. Questions we ask include, for example, how the structure of a forest relates to how much carbon it takes up from the atmosphere, how structure can help forests be resilient to climate change, and how the architecture of individual trees relates to their life history strategies. I use UAV's to study forests from above with hyperspectral optical sensors, and from below with terrestrial LiDAR scanning (TLS), Structure from Motion (SfM), and microenvironmental sensing techniques (see my brief talk about TLS here, and some 3D tree models here). This work is yielding new insights into how forests function around the globe. I also have a background in electrical and software engineering (see my CV here), and I always look for opportunities to combine these fields, such as this project demonstrates.

Key features for a **minimum viable product** would include:

- Block diagram of the planned computing architecture for the system
- A working global prediction system on Monsoon
    - Raster layers stored in appropriate format
    - Python code successfully running with client-supplied model specification to predict climate cooling process across entire globe. There will be significant thought required to keep processing times to acceptable levels, given the size of the data to be processed.
- User interface running on a public-facing system that will pull data from Monsoon.
    - The UI will present a zoomable global map to the user.
    - It will prompt the user to draw or upload a polygon.
    - Once uploaded and confirmed, it will then run the prediction for the specified area and present raster and summary results to the user.
- A simple system for user access control

A **useful system** would include:

- A Partner REST API (likely python-based) that would allow submission of polygons and retrieval of results (raster blocks, summarized data)
    - This would include a test webpage where queries could be tested.
- A more sophisticated user account system that allows users to manage their own accounts, and would keep a history of users' queries.

Stretch goals could include:

- An even more sophisticated user account system that would allow for the tracking of how users use the system, to enable usage-based billing.
- Build a module to conduct the environmental modeling on Monsoon itself in Python
- Further performance upgrades such as tiling rasters.
- Ability to upload shape files (.shp) instead of drawing polygons.
- Ability to select countries from a dropdown list instead of a polygon.
- Developing an ArcGIS plugin that would connect this service to ESRI's ArcGIS tool.

A successful project would have lasting impacts on the trajectory of climate change by making an attractive and functional tool that would encourage developers to take advantage of our discovery to increase profits from reforestation.

If you choose to pursue this project, you will gain experience in: handling big environmental data on large state-of-the-art clusters; building interfaces for global environmental data; API development; user management; ecological modeling; and applications for the carbon market.

Knowledge, skills, and expertise required for this project:

- Interest in working with big environmental data to help tackle climate change.
- Python programming
- Linux
- Interested in learning how to develop high-speed and high-capacity computing architecture

Equipment Requirements:

- There should be no equipment or software required other than a development platform and software/tools freely available online.

Software and other Deliverables:

- A strong as-built report detailing the design and implementation of the product in a complete, clear and professional manner. This document should provide a strong basis for future development of the product.
- Complete professionally-documented codebase, delivered both as a repository in GitHub, BitBucket, or some other version control repository; and as a physical archive on a USB drive.