

Developing World II: Sensing with Farmers

Mutembesa Daniel and Dr. Ernest Mwebaze
Makerere University
School of Computing and IT

3/27/18

These notes are based off a presentation by Mutembesa Daniel and Dr. Ernest Mwebaze for the section on the developing world in the Mechanism Design for Social Good Reading Group. The notes are taken by members of the reading group with some figures and texts taken from the presentations and project website. Questions and comments from reading group members during the presentation are labeled as such. Please contact the reading group organizers with any questions or comments.

1 Introduction

Real-time surveillance is key for effective crop health monitoring and disease detection in developing nations. In Uganda, viral disease attacks on crops are viewed as one of the leading causes of food insecurity and poverty, and they disproportionately affect small-holder farmers. For instance, the Cassava Brown Streak disease and Cassava Mosaic diseases can cause up to a zero percent cassava yield if the disease spread is not curtailed in time. These disease result in up to 2 Billion USD loss for farmers in the East African region alone.

Enabling experts to conduct surveillance tasks more effectively can help ensure that interventions are deployed in a timely manner. Crowdsourcing surveillance data can play a key role here. Traditional surveys can be prohibitively expensive to conduct frequently and at scale. As the disease incidences, severities and pest infestation varies in both space and time, single stations of expert surveyors or annual pest and disease survey are not sufficient to produce such a detailed map.

The Cassava Adhoc Surveillance Project research work aims to understand the tools and techniques needed for real-time surveillance of serious viral diseases in cassava. The data collected through this surveillance technology can then be used for spatial analysis and modeling of the different cassava diseases and pests. This will enable appropriate interventions to be developed by concerned Agricultural government bodies.

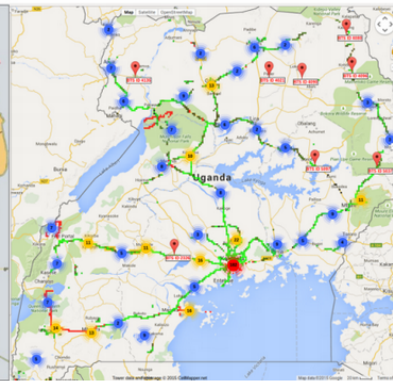
There is far more coverage by cell phone towers of areas of Uganda than agricultural research centers. This, coupled with the fact that many farmers own feature or smart phones, allows for the augmentation of expert surveys with farmer contributions.

LOCATION OF THE PUBLIC AGRICULTURAL RESEARCH INSTITUTES

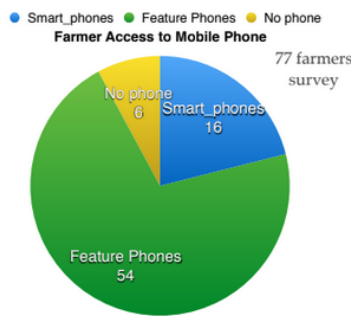


16

Cell Phone towers spread for major telecom

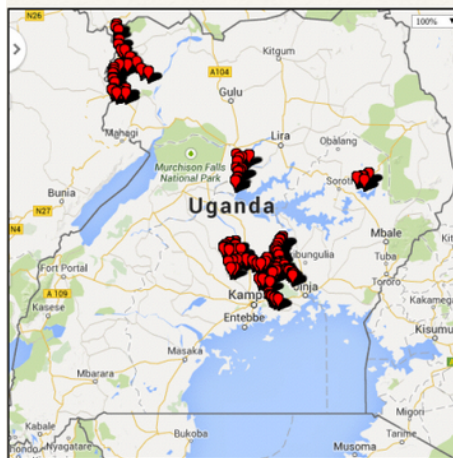


42

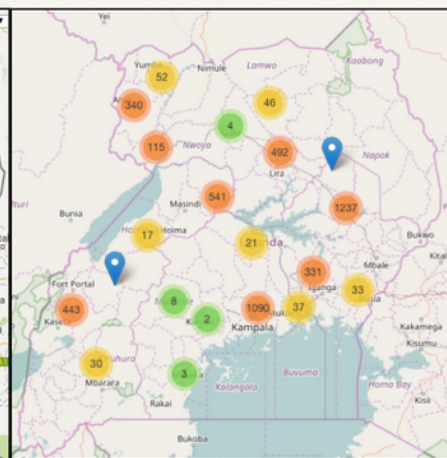


The group did a small pilot in 2016 using phones to collect data on occurrence of crop diseases. Data was collected through expert surveys along main roads and in places that are accessible by highway. They found that smartphone-based crowd-sources surveillance has substantial coverage and can provide a rich source of data that can be used to mitigate the burden of crop diseases.

National Expert Cassava Pest & Disease Survey 2014



Mobile Ad Hoc Surveillance Live Map



1.1 AdSurv Tool: Opportunities and Challenges

The AdSurv tool is a data collection tool used by farmers. Farmers can use this tool to report information (e.g. see figure below) and this information is directly added to a live map.



This tool presents several opportunities:

- Monitor pest and disease in real time. Knowing incidence and severity can be inform early warning systems.
- Automate pest and symptom measurements.
- Monitor harvests and seed systems.
- Monitor crop varieties in agro-ecologies.
- Understand farmer communities. e.g., stress of socioeconomic factors such as women being the main farmers yet men having control over revenues in many instances.
- Supplement humanitarian relief.
- Explore the use of innovative low-cost spectrometry.
- Assess the effectiveness of coupling mobile survey technology with the use of spatio-temporal modeling techniques to monitor and predict the spread of disease.

This platform is set up as a game where the players are farmers, experts using the ad-hoc surveillance tool, and collaborators. The objective is to collect truthful, high-value data. We are especially interested in the incidence rate or severity of certain pests and diseases. Truthfulness is key here. High-value locations are those that are hard-to-reach, and are therefore of high-interest.

For this task, players take pictures, label the pictures with what it is, its severity, and GPS location, and transmit this information. We influence participation through different incentives such as:

- Providing phone equipment and internet data,
- Follow-up calls,
- Monetary rewards,
- Subject surveillance. (i.e. asked for weekly contributions to a specific subject of interest to the national agricultural organization.)

The pilot contained about 30 agents who participated for a period of 2 years. The plan is to scale to a network of 200 farmers. Note, people are interested in various aspects including: balancing technology (images, quality and quantity of monitoring, quickly changing incentives, etc). This is an ongoing project and the pipeline is still being built. There have been various mechanisms taking into consideration participation efforts.

There are several additional projects including:

- African Cassava Whitefly: a project dedicated to automating the whitefly pest count, which carry tons of devastating cassava diseases
- Cassava diagnostic project: rapid diagnosis of crop disease in fields based on crop images, using computational techniques in machine learning implemented on a mobile phone.
- and many others!

Question: You are trying to estimate coverage. Is the challenge in trying to infer the true distributions of these effects in a mechanism where you can only incentivize certain individuals? Are there issues that come up with incentives to report measurement, and you report coverage against benchmark of being able to measure everything yourself?

Answer: Our first objective was to understand whether it is possible to extend a reporting tool to farming communities that can report in real time. So the first objective is participation. Can we draw interest and sustain participation? In our first cohort, most were trustworthy and were able to train other farmers. We weren't worried about truthfulness of the data per se. The mechanisms incentivized participation. We were worried about quality (e.g., completeness of the task), which is a participation issue as well. There are also a variety literacy across locales, and we wanted to know if they can collect quality data. The second objective was to collect data widely, including in areas where data collection had previously been sparse. Mechanisms increase the weight on a region of interest. The third objective was quantity of data. Each week, ask the farmers to send 20 reports.

Question: Have you looked at connections with citizen science research? There is work at the Ornithology Lab at Cornell where they're trying to figure out bird migration patterns through crowdsourcing. They also face similar issues including differential coverage, mis-measurement, gaps in the data, and so on.

Answer: With the 200 farmer network, one thing we're thinking about is to leverage the network for citizen sensing. Some people are interested in nutrition, crowd-sourcing market prices, health/immunization. Once we have the backbone set up, there are many opportunities to leverage the network.

Question: How do you decide the 200 people in the network? Is it easy to get people to join the network?

Answer: In the last 3-4 months, we've done participation enrollment and verification exercises. We look at enrollee education, phone use experience, etc. It's complicated because there are many interest groups who want to join, but we want to select the right group. Joining is not a problem – people want free phones and also value the possible benefits of the program.

The group website contains details about methods about these projects (as well as many others!) and corresponding datasets.

2 Mechanism Design and AI in the Developing World

The above presentations was followed by a discussion with Dr. Ernest Mwebaze.

Question: What other problems, outside of agriculture, does the lab focus on in the developing world?

Answer: The goal is to develop technological/computational tools to solve problems in developing countries. We have looked at problems related to expert tasks, which is a big problem in developing nations. We've looked into automating processes, speeding up work pipelines, and other related tasks.

For instance, for health, diagnosis often requires expensive equipment and human resources. There are phone-based systems that can look through a microscope at a blood sample and make an inference about sickness. For traffic, we can mount low-cost cameras to check road images and estimate traffic and speed. Kudu is a project that came out of this lab.

We have been looking at other datasets in health, mapping with satellite imagery (e.g. where people are growing crops, roofs on different types of houses), and many other domains. We study high-impact problems, where lack of human experts is a problem and where computational tools may fill in the gap.

Question: How long has the lab been running?

Answer: The lab has been running for 5-7 years. It began as a small group with myself, my advisor, and a couple of students. It has now grown to 5-8 researchers and 15-20 students.

Question: Are there opportunities for people who are not based at Makerere University to come and visit and potentially explore problems that your lab is interested in?

Answer: Yes, we've had a collaboration with MIT undergrads. A visitor is coming this June from another university. We usually discuss online about the types of problems that are of mutual interest and do the preliminary work and make progress during visits.

Question: Starting from an online discussion is an interesting contrast between how projects start in universities in the US when people may be co-located. Have you had

experience with how that works out when there isn't geographic co-location in early stages of projects?

Answer: Previously, what happened was that visitors would come, and it would take time to get set up (e.g., download software, get datasets, etc). So, we have found that working online beforehand is more efficient since it can cut out 2-3 weeks of setup time.

Question: Are there some research questions which are purely observational, or do they all involve an intervention at the end?

Answer: We've found that interventions work best. Scoping out a very small project that can be done in a month can be very useful.

Question: Do you have a sense of the types of questions you want to see answered and are you looking for collaborators? Or are the projects more exploratory where you go back and forth with potential collaborators?

Answer: Some of both!

Question: Are you exploring any directions related to education?

Answer: Not yet! It's something we've been wanting to do. We have done some brainstorming, but no concrete ideas.

Question: What research directions do you see related to primary/secondary education in the developing world?

Answer: We have been thinking about how you evaluate literacy at this level. Most interventions are about changing curriculum, teacher learning, and so on. We're interested in how you evaluate learning. Can we do this through crowd-sourcing or voice recording? E.g. students could read sentences and we can have software that analyzes the speed, time, gaps, spacing, and intonation.

Question: It is interesting that you mention literacy evaluation. In this group, we covered a section on education about how to properly evaluate students, teachers, schools, etc. Another thing we discussed is resource allocation within an education system. I imagine the issues in the developing world may be different from those in the US. What resource allocation problems show up within education in a developing world context?

Answer: Resource allocation is very challenging in this context. We have allocation of human resources, e.g., teachers to schools. The government rolled out universal education and allocates funds/resources to different schools based on student counts. These numbers tend to be exaggerated. Also, payment is based on number of students, number of lectures by teachers, and other such metrics, but people just fill out the maximum value. Even with things like books, furnitures, and other small equipments, there are political dynamics. People don't report the right amount. There is probably a good project to explore there.

Speaker Bios

Mutembesa is a Project Head and Lead Researcher for an Adhoc Surveillance Project focusing on the smart-phone application tool kits. These kits have been developed by the mcrops team to facilitate and provide automated diagnosis and improved vector and symptom measurement for Cassava viral disease and pests. The team also looks at crowdsourced Crop health data through an Ad-hoc Surveillance system. The goal is to use crowdsourcing techniques to supplement experts who used to do this in a manual and infrequent manner.