SDP 03 CENS Deployment Center (CENSDC)

SDP 03.1 People

- Principle Investigator: Christine Borgman
- Faculty: Christine Borgman
- Staff: Mike Taggart
- Graduate Students: Matthew Mayernik
- Undergraduates: Erick Romero

SDP 03.1 Overview

CENS researchers are developing flexible wireless sensing technologies that can be used in a variety of scientific applications. These technologies are used to produce valuable scientific data. CENS data are largely collected on real-world deployments where sensing systems are deployed in particular locations where phenomena of scientific interest exist. CENS deployments are highly variable. Researchers cannot fully predict what they will encounter when they reach a field location. There may be unexpected temperature fluctuations, excess moisture or dust, or unpredictable flora and fauna, all of which affect the functionality of both equipment and people. Researchers have deployment goals and objectives prior to going out in the field, but decisions have to be made on site about where, when, and how to deploy equipment and sensors based on local conditions and the state of equipment. As a deployment proceeds, researchers adjust their activities in response to the situation, to time constraints, and to the available sensors and power. Trade-offs between these factors are made in real time, affecting the data that are captured during a deployment and the ways that they are interpreted post-deployment.

As CENS researchers participate in deployments, they build up knowledge about potential problems they may encounter and how to solve them. Community knowledge of deployment best practices is a valuable asset for CENS. Our central research questions are how to facilitate deployment knowledge transfer in the collaborative CENS research setting, and how information about deployment activities can be leveraged to describe CENS data.

We have expanded CENSDC development efforts in two main ways, first through the development of a multi-site deployment module, and second to enable researchers to collect information about deployment activities while deployments are taking place. Early user feedback suggested that the inability access the internet while in field locations created a significant gap in our system. In our work this year, we have approached this gap in two ways: 1. developing handheld methods for collecting deployment information as deployments are taking place, and 2. interfacing with streaming data. This led to a successful proposal to Microsoft Research to develop an application for a handheld device for in-field use, which is discussed more in later sections of this report, and in greater detail in a separate report.

SDP 03.2 Approach

The CENS Deployment Center (CENSDC) was designed to leverage CENS deployment knowledge by providing a central location for researchers to document deployment activities through the creation of pre-deployment plans and post-deployment feedback/notes. By allowing users to describe their deployment experiences, including lessons learned, troubleshooting techniques, and provide guidance for future deployments, we are attempting to capture the tacit knowledge about equipment setups, deployment locations, and field preparations that play a critical role in data collection techniques. As CENS technologies mature and current researchers gain deployment experience, new students face a steeper learning curve when joining a project.

A parallel goal of the CENSDC is to add value to CENS data by providing a source of descriptive and contextual information surrounding the data collection. CENS provides a great cyberinfrastructure test case on data reuse. The data from CENS deployments are spatially and temporally located, having scientific value both to immediate research questions and long-term longitudinal studies, and are therefore irreproducible. With any research
endeavor, understanding the context of data collection is critical to the ultimate evaluation and interpretation of results. The challenge of capturing data collection context is particularly difficult when data are collected on highly variable and unpredictable real-world deployments. The process of collecting and articulating information about deployments in the CENSDC involves valuable reflection on research methods, sample selection, and troubleshooting techniques. This information can assist researchers in writing papers, proposals, and reviews, as well as in maintaining their data and leveraging them for reuse by others.

**SDP 03.3 System(s) Description and/or Experiments**

System development in the past year has primarily focused on implementing a module that records information about multi-site deployments. The CENSDC system was designed to address the needs of singular campaign-style deployments. We have expanded the system to incorporate a multi-site deployment module uses a parent-child model where a number of individual deployment sites are constituent parts of a single deployment. The development of the multi-site module was undertaken to specifically address the complex deployments performed by the CENS seismic group. A multi-site module has been implemented and is being used to collect regular updates from the seismic deployment.

The second main development effort has focused on interfacing with regular status updates from ongoing deployments. As the Peru seismic deployment is progressing, more and more data are being streamed to UCLA, including various kinds of technical data, along with the seismic data itself. The technical data provide various measures of the health of the seismic stations, as well as of the wireless networks that connect them to each other and to the internet. These measures are meant to enable members of the seismic team to characterize the status of each station, and to identify problems as they arise. We have created a Google Maps interface that pulls data from the seismic team database to map the quality of the wireless links between stations along the deployment transect. The map interface is shown in Figure 1. The link quality data on which the map is based are updated every day, enabling members of the seismic team to view the current status of the station-to-station wireless links. The mapping feature is being evaluated with members of the seismic team, and our intension is to expand the feature to show other kinds of data. Next steps will include mapping data flows between stations.

**SDP 03.4 Accomplishments**

Members of the CENSDC team have taken part in five deployments or research trips in the past year. CENS research takes on many forms. Each research team has their own ways of performing field research, and collaborations between research teams add to the complexity. Sensor network deployments occur in different types, reflecting this diversity of research methods, and fit within the larger scope of research projects. Deployment planning and organization also reflects the research diversity within CENS, as each team coordinates and organizes within the dynamic of their own group.

In expanding use of the CENSDC, we
have found that the system is most suited for projects that conduct repeated or long-term data collection deployments. For projects with repeated deployments, equipment and data collection processes are often similar from deployment to deployment, but the specific differences between deployments are important. The CENSDC “make like” function enables researchers to create new deployment records from past deployment information, illustrating the continuity between deployments. For long-term sensor deployments, monitoring and maintenance of deployment equipment and sites are critical throughout the lifetime of the project. Logs of these activities in the CENSDC serve as data annotation, giving an overall picture of the problems encountered and solutions during the deployment. Additionally, the most active users of the CENSDC have been involved with deployments that take place in remote locations, such as in Peru and Bangladesh, as the organizational requirements for performing remote deployments are much greater than for local deployments. Members of local deployments can adapt to changing plans or deployment needs with more flexibility.

To connect deployment information in the CENSDC with the data that results from CENS deployments, we are linking from deployment records to online data sources. Most deployments in the CENSDC do not have associated projects in Sensorbase. None of the recent NAMOS, Bangladesh, Merced River, or Seismic data are held in Sensorbase. However, some of these projects have data available online separate from Sensorbase, such as the Merced River and NAMOS projects. We are linking to these additional CENS data sources where possible, and have plans to create more direct Sensorbase-CENSDC crossovers, as detailed in the next section.

SDP 03.5 Future directions

Future work on the CENSDC will focus on increasing the visibility and use of the system by creating Sensorbase-CENSDC crossover functionality, performing user studies and evaluations of the CENSDC website and functionalities, and continuing to work on collecting and displaying information about deployments as they are taking place.

One of the goals of CENS is to make high quality, well documented data sets from CENS deployments available to other researchers. Sensorbase provides researchers with a platform for data sharing, but as mentioned above, Sensorbase projects typically do not have much description or documentation associated with them. Detailed documentation is not necessary for all Sensorbase projects, as for various reasons many projects are not intended to be used in contexts outside of their immediate use. But for projects where the data owners would like to share their data, or at least make it available to others, additional description may be beneficial. We have been in discussion with the Sensorbase team regarding ways that some CENSDC features might be integrated into Sensorbase, specifically to enable users to give more information about the dates and locations that data were collected, the people involved in the project, and equipment used to collect data. Additionally, we plan on enabling CENSDC users to create Sensorbase projects when creating new deployment plans, as a means to reduce redundancies across the two systems.

A further benefit to building bridges between the two systems is that Sensorbase projects are not indexed by Google or other search engines. CENSDC project pages that describe Sensorbase data would provide more visibility for CENS data, as the CENSDC pages are indexed by the major search engines. CENSDC descriptions of Sensorbase data will not replace the need for researchers to personally discuss the specifics of sharing data, but they can be a way for outside users to quickly assess the potential usefulness of CENS data, and find out who to contact for more information. In this sense the CENSDC pages would serve as the public face to privately held data.

In parallel with the integration efforts, we will conduct user studies of the existing CENSDC system. User studies will focus on two main aspects of the system. The first aspect is collaborative knowledge transfer. We will evaluate the utility of the CENSDC in helping new members of a research team to get up to speed on the kinds of deployment problems and issues to expect, as well as the utility of the CENSDC in keeping current team members updated on the status of a deployment. The second aspect we will evaluate is the utility of the information captured in the CENSDC for data description and discovery. Data is often not self-describing, in particular Sensorbase projects typically do not have much description associated with them. Many different kinds of metadata schemas are available to describe different kinds of data, but are often difficult to understand and use. Our goal is to enable
researchers to collect deployment information in the CENSDC that can be used to supplement existing metadata, or when metadata is absent to serve as a data description that can be used to share and discover data resources. We will evaluate the effectiveness of CENSDC deployment information in describing CENS data through surveys and informal interviews with CENS researchers.

Further deployment and use of CENSDC will focus on the types of deployments described that we have found to be most suited to the system architecture, those with similar and repeated deployments and those with long term sensor installations. We will also look at ways that we can expand the integration of real time deployment displays like we have created for the seismic deployment. The need to know the status of existing equipment installations is not unique to the seismic deployment, other deployments have similar needs. Developing similar deployment displays will be based on the needs of individual deployments, with the goal of producing generalizable best practices for displaying deployment status information.

Additional future work centers on the development and deployment of the handheld application for deployment data collection. This work, described in more detail in a separate annual report, feeds into the CENSDC development by providing a mechanism for collecting information about field activities as they are occurring. The lack of in-field access was identified as a significant gap in our existing system, and we hope to address it through a handheld-based data collection and note-taking tool.