Real-time streaming of environmental field data

Enrique R. Vivoni*, Richard Camilli

Ralph M. Parsons Laboratory, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 20 March 2002; received in revised form 21 October 2002; accepted 20 November 2002

Abstract

Field measurements in the environmental sciences still depend upon the pencil and paper notebook for data collection. Although robust, this method is labor-intensive and susceptible to recording and georeferencing errors during transcription. Recent advances in mobile computing and wireless communications allow the geoscientist to process and transmit data while still in the field, thereby minimizing human errors and time delays. We describe an integrated system developed for environmental and geolocation data acquisition that is intended to streamline the collection process. The system consists of software applications and hardware components that enable wireless, mobile and Internet computing during field campaigns. In particular, two-way transfer and display of collected data is achieved between the field site and a remote location, a concept referred to as field data streaming. A prototype system has been tested in field trials in Cambridge, Massachusetts, USA and Newcastle, New South Wales, Australia. Field studies demonstrate the noticeable gains in efficiency and precision achieved with the use of the field streaming technology. Potential applications include biogeochemical and hydrologic studies, water quality monitoring, emergency response to water-borne disasters and intensive field sampling campaigns.

r 2003 Elsevier Science Ltd. All rights reserved.

Keywords: Mobile computing; Wireless communications; Environmental monitoring; Geographic information systems; Field data collection

1. Introduction

The recent advances in mobile information technology have created opportunities for developing tools that bring computing power to field workers. As a result, applications for mobile devices have been developed in various disciplines, for example, in geographical information systems (Pundt and Brinkkötter-Runde, 2000), resource management (Stern and Kendall, 2001), geology (Briner et al., 1999), archaeology (Ancona et al., 1999), geomorphology (Cornelius et al., 1994) and ecology (Pascoe et al., 1999). These disciplines share a common need to associate field measurements with unique locations in space and time. However, the analysis required to place sampled data within a spatial–temporal context is usually performed after field study completion. For the geoscientist, the use of mobile computers can facilitate data collection (Loudon, 2000) and also provide a means for discovering data relationships while still in the field.

In situ field mapping, data analysis and sharing is currently achievable by integrating environmental and geolocation sensors with mobile, wireless computers. A mobile data collection system for these purposes should:

1. Efficiently collect and store various parameters;
2. Acquire and merge data from multiple teams;
3. Accurately map data within a geographic context.

*Corresponding author. Tel.: +1-617-253-1969; fax: +1-617-258-8850.
E-mail addresses: vivoni@mit.edu (E.R. Vivoni), romeo@mit.edu (R. Camilli).
and methods for sharing, analyzing and displaying data over a network. Application software resident on a remote server can provide data access, mapping and analysis functionality to the field worker through the same network. Interactive data collection and mapping has been shown to significantly enhance the scientific discovery process (Carver et al., 1995).

This paper presents the development of a field data collection system that streamlines the collection process and provides data sharing between multiple field teams and remote locations. The collection and transmission of environmental field data is executed without the loss of precision that typifies manual data transfer. By using a geographic information system (GIS), the system adds a new dimension of spatial localization and provides the user with cartographic displays of relevant field data. The exchange of georeferenced environmental data between a field site and a remote location provided by the system is referred to here as field data streaming (Fig. 1).

This study focuses on the collection of geospatial and water quality data during a watershed field study. To this end, the system is tailored to sampling biological, chemical and hydrologic parameters along river cross-sections. We describe the field testing of the ENVIT Integrated Data Collection System during trials in Cambridge, USA and Newcastle, Australia. First, we present the technologies required to store, acquire, transmit and display environmental field data. We then assess the utility of the integrated system for a water quality study based on field tests in the Williams River. Finally, we discuss the key merits of the data collection system and potential applications areas.

2. ENVIT integrated data collection system

The field data collection system consists of a series of ruggedized mobile computers and a roving station equipped with a field server, wireless router and mobile telephone. The system components are similar to the “tele-geomatics” framework outlined by Zingler et al. (1999) for wireless field data collection. Fig. 2 presents a generalized schematic of the integrated system. Table 1 presents the components of the prototype system while specific details are discussed in the following sections.

2.1. Field data acquisition and storage

The process of gathering field data in many geosciences is typically a time-consuming task due to the need for measurement and transcription of data often in remote regions. To be worthwhile, mobile computing should streamline the data collection process and allow users to rapidly acquire, record and submit data to a broader system infrastructure. With this in mind, a software application, known as ENVIT-Note, was developed for the hand-held device. The application consists of a graphical user interface (GUI) with features that lead the user sequentially through the following

<table>
<thead>
<tr>
<th>Field User</th>
<th>Remote User</th>
</tr>
</thead>
<tbody>
<tr>
<td>sensor</td>
<td>Data Services</td>
</tr>
<tr>
<td>Mobile Device</td>
<td>GIS Map Services</td>
</tr>
<tr>
<td>sensor</td>
<td>Application Services</td>
</tr>
<tr>
<td>Mobile Device</td>
<td>Process/Map/Display/Provide Data</td>
</tr>
<tr>
<td>Collect/Store/Transmit/Display Data</td>
<td></td>
</tr>
<tr>
<td>Field Server</td>
<td>Remote Web Server</td>
</tr>
<tr>
<td>Transmit Raw Data</td>
<td>Integrate Data</td>
</tr>
<tr>
<td>Field Server</td>
<td>Transmit Processed Data</td>
</tr>
<tr>
<td>Remote Web Server</td>
<td>Data Services</td>
</tr>
</tbody>
</table>

Fig. 1. Conceptual diagram of field data streaming technology.
The mobile software encapsulates a suite of applications that provide several functions to the field worker, including: manual input of velocity measurements using a hand-held flow meter; manual input of biological samples using a field biology kit; manual input of chemical samples using a field-portable spectrophotometer; manual input of water quality measurements using a multi-parameter probe; manual and automated input of geographic position using a mobile GPS; manual input of field comments and geometric measurements; data submittal to an independent mobile GIS application; and review and submittal of local database contents.

In a watershed or estuarine study, the river cross-section is the basic geometric unit across which discharges and biogeochemical fluxes are typically computed. For this reason, chemistry, biology or velocity measurements are entered into the system through a graphical representation of a river cross-section (Fig. 4A). Manual input occurs, for example, when entering velocity data along a cross-section. After the user specifies the geometry, the software application computes the required sampling locations through standard stream gauging procedures (Wahl et al., 1995). Upon completion of the measurements, the computed channel discharge is displayed. Similar

---

functionality is provided for collecting biology or chemistry samples and computing chemical fluxes.

Using this cross-section sampling strategy, sensor data can be mapped to a specific location within each section. Measurement is further automated by using signal-averaged GPS data at each stream bank to calculate the stream cross-sections (Fig. 4B). For increased accuracy, a differential GPS measurement could potentially be
used with the existing system. Combining the sampling location with the position of the banks is a means for efficiently determining the geographic position of each sample without the need for including position fixes at each point (Fig. 4C).

Georeferencing the environmental parameters is automated within the system database application, which consists of a master copy operating on a field laptop and local copies on the mobile devices (Fig. 5). A key feature is the capability of the master database to simultaneously gather and process newly acquired data from multiple mobile devices. The design is made robust and expandable by abstracting the components of the field data collection process. In general, the database consists of data tables describing the project, location, field operator, equipment and data. The database can be tailored to the sampling study and the availability of specific system components. Preconfiguring the master database allows entering descriptive properties of the field study into the database prior to the actual expedition. In addition, the preconfigured database supports user interaction with the system.

2.2. Field data transmission and integration

In order to simultaneously share data amongst geoscientists in both the field and at a remote location, the ENVIT System includes a portable, wireless, radio-frequency communication network. This network transmits data from each mobile device to a field laptop functioning as the central data server. To connect the portable network, an outdoor wireless field router, operating at 2.4 GHz,\(^7\) is mounted on a roving van along with a 1 W amplifier and omnidirectional antenna.\(^8\) The wireless router communicates with a series of wireless PC cards\(^9\) attached through the PCMCIA port to the hand-held devices.\(^10\) A signal radius of tens of kilometers is potentially available, depending on the availability of line-of-sight and the signal amplification provided. This range permits various field teams to update the central database simultaneously (Fig. 6).

Merging of data collected from field teams requires transmission from each mobile device to the master relational database on the field laptop. To ensure optimal networking efficiency, only updates of changes between the local and server databases are transmitted.

---


thus eliminating redundant transfers. Local database copies provide system back-ups in case of potential problems during transmission. Upon receiving data from each team, the server database organizes these records, associates parameters with their sampling location, and provides a field data summary table in a format accessible to other software. In order to create a summary table, the database utilizes key identifiers that relate the measurement records with a sampling location (Table 2).

Data transmission and integration can be complicated by the lack of a wireless transceiver in each mobile device. In the prototype system, for example, the GPS sensor and wireless card both utilize the PCMCIA expansion pack on the hand-held device,11 precluding simultaneous use of all sensors on a single device. With alternative hardware components or wireless capability, however, this situation can be easily circumvented. Given the constraints of the prototype, a method for transmitting data across mobile devices without direct access to the radio-frequency network is required. Therefore, the mobile device infrared sensors are utilized to transmit the local database contents between hand-held computers. Infrared transmission was shown to be an effective means to transfer data onto a mobile device that has long range, radio-frequency, capability.

Field data streaming requires that the deployed field workers have the capacity to transmit and receive data from a remotely located server. Therefore, data transfer to an off-site location requires an alternative communications network. In principle, each mobile device could function as a client to a remote host. Given the associated costs of hardware and services, however, a more efficient solution is to have a single access point and utilize the wireless field network to distribute and gather data. In the integrated system, the field laptop is connected via a mobile phone to the Internet. Based on available coverage, a dual-frequency mobile phone (e.g., GSM/GPRS service at 900 MHz and 1.9 GHz) is used to enable connection across different mobile standards.12 Wireless transmission of the database contents is periodically made to a remote web server where it is captured by a customized web services application. The field data is subsequently displayed in tabular or cartographic format on an Internet site available to the field laptop for content distribution to the multiple devices.

2.3. Field data display and mapping

Visual representation of the environmental field data obtained during a field campaign is typically produced after completion of the study and analysis of the recorded data. In the ENVIT System, a major innovation is the capability to immediately stream processed data to an off-site user or back to the field worker. In order to display data to the mobile user, two mechanisms are implemented: (1) local map generation utilizing a mobile GIS (see footnote 6); and (2) uploading of a customized web mapping service through the wireless network. For a remote user, the web services application served as a means for querying, displaying and mapping the environmental data collected in the field.

Map generation on the mobile device is achieved through a customization of the mobile GIS application to graphically display the database summary table (Fig. 7). Overlaid on thematic map layers (e.g., topography, streams), the sampling points reveal the spatial

---


and temporal variation of the measured parameters in the field site. This data may be a product of a single team or the integrated efforts of many teams, depending on whether the summary table is generated by the mobile device or the field laptop. In either case, the capability for extracting the spatial–temporal context of the field data and its relationship to regional characteristics is enhanced.

Mobile GIS also offers limited spatial analysis capabilities for extracting information from the collected data. Advanced processing, however, requires more powerful GIS software.\textsuperscript{13,14} To meet this need, a web services application captures the transmitted data from the mobile phone and displays it on a web browser. The cartographic display consists of a background image depicting the topography and streams overlaid by color-coded points representing the georeferenced environmental parameters (Fig. 7). This service provides various other functionalities including: (1) selection of the parameter to be displayed; (2) mapping of a user-provided geographical location; (3) zoom and pan

\begin{table}[h]
\centering
\begin{tabular}{ll}
\hline
Parameter type & Parameter name \\
\hline
Geolocation & Longitude \\
& Latitude \\
Hydrology & Velocity \\
Water quality & pH \\
& Dissolved oxygen \\
& Conductivity \\
& Oxygen reduction potential \\
& Turbidity \\
& Temperature \\
Geometry & Depth \\
& Width \\
Chemistry & Nitrate \\
& Nitrite \\
& Phosphate \\
Biology & Total coliform \\
& \textit{E. coli} \\
\hline
\end{tabular}
\caption{Measurement parameters in Williams River field study}
\end{table}

\textsuperscript{13} ESRI, Inc. Arcview 3.2 GIS \url{http://www.esri.com/arcview/}.
\textsuperscript{14} ESRI, Inc. ArcIMS 3 \url{http://www.esri.com/arcims/}. 

Fig. 6. Diagram of multiple-team field deployment in a watershed study.
capabilities; and (4) tabular view of the summary output. Both tabular and graphic displays provide capabilities for quickly analyzing the spatial and temporal trends in the field data from any location during an on-going study. Access to this data by an off-site manager or decision-maker allows for rapid evaluation of progress and preliminary results. Providing this data to the mobile device browser closes the field streaming loop by allowing field workers immediate access to processed data (Fig. 1).

3. Application and field study deployment

Prior to full-scale field deployment, the integrated data collection system was pilot-tested in a study of the Charles River Basin in Cambridge, Massachusetts. Two sampling teams tested various system components including the mobile device software, geopositional sensor, water quality probe and wireless data transmission. A roving van equipped with the field laptop, wireless outdoor router, amplifier and antenna served as the central data server. This trial study verified that environmental field data could be successfully collected, stored and transmitted utilizing the integrated system. Further testing determined a maximum range of one kilometer for the radio-frequency network given clear line-of-sight. Coverage beyond 1 km range or under obstructed conditions necessitates a higher broadcasting power normally requiring a license.

The site chosen for carrying out a full-scale field test was the Williams River watershed in New South Wales, Australia (Fig. 8). This catchment is an area of active hydrological and water quality monitoring\(^\text{15}\) and research and is confronted with issues of non-point source pollution. A detailed GIS database for the Williams River was available prior to the field study (A. Krause, pers. comm., 2001). In addition, a network of water quality and flow gauging stations is maintained within the watershed.\(^\text{16}\) The availability of the preexisting infrastructure and historical data was an attraction for testing the integrated system within the Williams River watershed.

The catchment, 1260 km\(^2\) in size, is located within the lower Hunter Valley, about 30 km north of Newcastle. The watershed is characterized by a mountainous, forested region in the north and a series of low, rolling hills, used for cattle grazing, in the central and southern portions. The warm, temperate climate is modulated by an orographic effect and a maritime influence due to its coastal proximity (Wooldridge et al., 2001a). Previous

---

\(^\text{15}\) NSW Department of Land and Water Conservation \texttt{http://www.dlwc.nsw.gov.au/}.

\(^\text{16}\) Hunter Integrated Telemetry System \texttt{http://hits.nsw.gov.au/}.\n
hydrologic studies have documented the correlation of rainfall and runoff to large-scale climate phenomenon (Wooldridge et al., 2001b), modeled the effect of land-use on hydrologic response (e.g., Wooldridge and Kalma, 2001) and studied the spatial distribution of soil moisture (Walker et al., 2001).

The Williams River deployment consisted of three field teams of six to seven members. Team members were responsible for tasks such as water quality and hydrologic sampling, operation of sensors and computers, and data transmission and display. Each team was equipped with two ruggedized hand-held computers and associated environmental sensors (Fig. 6). The field teams traveled independently visiting various predetermined sampling locations. Upon sampling at each location, the collected data (Table 2) was transmitted to the central data server in the roving van. From this field server, data was periodically transmitted back to the field teams, as well as to a remote web server in Cambridge, MA. Data sharing between the multiple teams included the transmission of summary tables and remotely processed maps of the sampled parameters.

For field deployment, the system also provided equipment ruggedization and extended battery life. Modifications to a waterproof case17 protected the mobile computers from dust, submergence and impact, while allowing user interaction with minimal interference including data transmission via infrared and radio frequencies (Fig. 9). To extend battery life, a lightweight, rechargeable battery was used to power the mobile devices and water quality probe from a single source. Given the different power requirements, power management circuitry was designed, built and integrated into the battery pack to provide approximately 40h of continuous field use. This efficient power system was essential since the mobile computers have high consumption rates.

4. Results and discussion

The pilot-scale and full-scale field tests of the ENVIT System successfully demonstrated the field data streaming technology. For the Williams River study, a thorough test of the system was conducted within an intensive field campaign in a remote setting. During the four-day campaign, the Williams River watershed was sampled at 35 locations spanning several tributaries throughout the watershed (Fig. 10). Various deployment configurations were tested, including a dispersal of the three-field teams at distances greater than two kilometers. In addition, an impromptu estuarine study was conducted to assess the longitudinal variation of water quality parameters at 500-m intervals along a 10 km

---

The field data collection system proved versatile enough to accommodate this new sampling strategy. As the field study progressed, environmental data collected from the various sensors provided an increasingly resolved view of the spatial variation of
hydrological, biological and chemical features within the watershed. This allowed the teams to plan future sampling as well as navigate through the watershed in the context of the sampled data. In addition, visualization of the data from multiple field teams in Australia documented progress for remote viewers worldwide through the web-based map and photographic displays. The field streaming concept was realized by transferring georeferenced environmental data from individual teams into a processed map product accessible to the field workers and any interested party while the study was in progress.

Although the ENVIT System was successfully used during the field campaign, various potential improvements to system performance were identified, including use of: (1) a higher powered radio-frequency network to overcome line-of-sight and range limitations across complex terrain; (2) satellite transmission in remote areas not covered by the mobile network; (3) improved integration of system components including GIS and mobile application user interfaces; and (4) continuous transmission, updating and sharing of raw and processed environmental data products. Nevertheless, the integrated sensors and mobile computers expedited gathering environmental and GPS data at each river cross-section. Data transfer rates to the field roving station were adequate for the content exchanged between the mobile and master databases. Overall system reliability across the hilly pastureland and streams during dry, summer conditions was satisfactory.

The field study experience indicates that the ENVIT System improved the efficiency, timeliness and transfer of field data into scientific results. The accuracy of the geolocation data have been verified with available GIS thematic layers (e.g., soils, landuse, topography, hydrography) for the watershed region. In addition, results suggest that the field portable environmental sensors provided measurements within expected ranges relative to the historical data for the Williams River. A detailed analysis of the water quality, hydrologic and geolocation data and their spatial variation is currently in progress. The quality-controlled, georeferenced environmental data will serve various purposes including augmentation of historical databases and water quality and hydrologic modeling of the watershed using a semi-distributed basin model.18

Linking environmental field data streaming with modeling is a key extension of the current framework. The field data streaming technology can accelerate not only the process of data collection and scientific discovery, but also the incorporation of observations into predictive models. Real-time transmission of environmental data from field sites allows for updating models of complex environmental systems with in situ measurements. For example, field discharge measurements along a river cross-section may serve to update model forecasts of stream flow. Conversely, real-time model predictions could serve to target sampling efforts. For example, in emergency response to chemical spills, model-directed field data collection can expedite plume characterization.

In addition, field-streaming technology can enable the continuous monitoring of environmental systems and function as a distributed sensor network.19,20 The prototype system provides the key elements for distributed collection, transmission and display of environmental field data. Further enhancements may include features such as: (1) collection and documentation of photographic data; (2) data transmission in regions without mobile phone coverage using satellite technology; (3) incorporation of additional environmental sensors using automated or manual input; and (4) web services dedicated to providing additional support such as communications coverage, model or data retrieval and advanced geospatial processing.

5. Potential applications

Although specifically tailored for water quality assessments, the ENVIT System design is sufficiently flexible to address other environmental data collection in a variety of contexts, including research, monitoring, resource management and emergency response. One potential application is the response to water-borne chemical or biological disasters. Field scale testing has demonstrated the system’s capability to serve in situations that require a coordinated response by multiple field teams. The advantages of this system are evident in emergency response situations with: (1) a critical time constraint, (2) an expansive study area, (3) remotely located decision support and management, and/or (4) multiple individuals requiring processed field data. A second potential application is for in situ sampling during intensive field campaigns designed to verify data from aerial or satellite platforms. Validating remote sensing data requires a coordinated effort for deploying multiple teams to measure ground-based environmental parameters over large sensor footprints (e.g., Famiglietti et al., 1999).

6. Conclusions

This paper provides an overview of the development and application of an integrated field data collection

---

19 High Performance Wireless Research and Education http://hpwren.ucsd.edu/.
20 Cal(IS2) http://www.calit2.net/.
system designed to acquire, store, display and transmit georeferenced environmental data during field campaigns. The system encompasses advanced mobile, wireless and Internet computing technologies that together facilitate the sharing of field data between the study site and remote locations in real-time. The collection of field data was tested within the context of a water quality study in the Williams River catchment in New South Wales, Australia. Results from the field study suggest that the prototype system is most useful for time-critical, intensive field campaigns carried out by multiple research teams. Furthermore, the general use of distributed wireless sensors was shown to be advantageous in environmental field research. Potential research avenues include the integration of new sensing capacities, enhanced server-client processing and improved integration with environmental models.

Acknowledgements

The development of the ENVIT Integrated Data Collection System was carried out as part of a MIT/ Microsoft Alliance project. We acknowledge the support of Sheila Frankel, Eric Adams, Dara Entekhabi, Daniel Sheehan and Mario Rodriguez. In addition, various students were involved in product development and testing: Neeraj Agarwal, James Brady, Nancy Choi, Chrissy Dobson, Arthur Fitzmaurice, Eric Lau, Anna Leos, Linda Liang, Brian Loux, Aurora Kagawa, Kris Kolodziej, Patricia McAndrew, Kevin Richards, Laura Rubiano, Kim Schwing, Russ Spieler, Chin-Huei Tsou, Lisa Walters and Amy Watson. We also acknowledge Garry Willgoose, Jetse Kalma and Andrew Krause (University of Newcastle, NSW). We also thank an anonymous reviewer as well as Steve Carver for their insightful comments on the manuscript.

References