DYNAMIC REAL TIME DISTRIBUTED SENSOR NETWORK BASED DATABASE MANAGEMENT SYSTEM USING XML, JAVA AND PHP TECHNOLOGIES

D. Sudharsan¹, J. Adinarayana¹, S. Ninomiya², M. Hirafuji³ and T. Kiura⁴

¹CSRE, Indian Institute of Technology, Mumbai, India
sudharsan@iitb.ac.in adi@iitb.ac.in
²ISAS, GSALS, The University of Tokyo, Tokyo, Japan
snino@isas.a.u-tokyo.ac.jp
³NARO, National Agricultural Research Center for Hokkaido Region, Kasai-gun, Japan
hirafuji@affrc.go.jp
⁴NARO, National Agricultural Research Center, Tsukuba, Ibaraki, Japan
kiura@affrc.go.jp

ABSTRACT

Wireless Sensor Network (WSN) is well known for distributed real time systems for various applications. In order to handle the increasing functionality and complexity of high resolution spatio-temporal sensory database, there is a strong need for a system/tool to analyse real time data associated with distributed sensor network systems. There are a few package/systems available to maintain the near real time database system/management, which are expensive and requires expertise. Hence, there is a need for a cost effective and easy to use dynamic real-time data repository system to provide real time data (raw as well as usable units) in a structured format. In the present study, a distributed sensor network system, with Agrisens (AS) and FieldServer (FS) as well as FS-based Flux Tower and FieldTwitter, is used, which consists of network of sensors and field images to observe/collect the real time weather, crop and environmental parameters for precision agriculture. The real time FieldServer-based spatio-temporal high resolution dynamic sensory data was converted into Dynamic Real-Time Database Management System (DRTDBMS) in a structured format for both raw and converted (with usable units) data. A web interface has been developed to access the DRTDBMS and exclusive domain has been created with the help of open/free Information and Communication Technology (ICT) tools in Extensible Markup Language (XML) using (Hypertext preprocessor) PHP algorithms and with eXtensible Hyper Text Markup Language (XHTML) self-scripting. The proposed DRTDBMS prototype, called GeoSense DRTDBMS, which is a part of the ongoing Indo-Japan initiative ‘ICT and Sensor Network based Decision Support Systems in Agriculture and Environment Assessment’, will be integrated with GeoSense cloud server to provide database (dynamic real-time weather/soil/crop and environmental parameters) and modeling services (crop water requirement and simulated rice yield modeling). The GeoSense cloud server has been developed with Opera-Unite for file sharing, web server (with cost effective web server tools Windows Apache MySQL PHP-WAMP and X-any operating systems Apache MySQL PHP and Perl –XAMPP), file upload, and web proxy functionalities. Currently, the GeoSense DRTDBMS is useful to the rural farming community for ubiquitous decision making in precision agriculture aspects. In future, this DRTDBMS system could be used in climate/environmental systems to understand the micro-climatic variations in real-time mode.

Keywords: Real-time distributed database management system, xml, JAVA, phpMyAdmin, open source system, precision agriculture, wireless sensor network and OSC standards.
1. INTRODUCTION

In today’s expeditious world, the development of information and communication Technology (ICT) and wireless sensor network (WSN) applications are becoming increasingly important and real-time in current scenario, particularly in precision agriculture aspects. Precision agriculture is concerned with integrated, productive and sustainable use of biological, physical and financial capital at varying geographic and temporal scales. It has been observed since two decades that there is a rapid growth in utility of modern technologies in precision agriculture, where one needs “where, what, when” types of solutions. Today, most farmers in developed countries own and use computers. [8] Cite a figure of 75% for Australian grain growers. Policies on ICT utility in agriculture are improving in many developing countries; Ministry of Agriculture and ICT of Government of India have set new policies with an intention to supply adequate and quality inputs to farmers in a timely manner [16] and to strengthen the agriculture sector with database to ensure greater reliability of estimates and forecasting which will help in the process of planning and policy making [5].

Many agricultural decision support packages are readily available and affordable. But most of the systems are working on offline data/database system (data processing and database development) either stand-alone or web based system (e-Sagu, 2012., GramyaVikas, 2008., aAQUA, 2012). In recent years, distributed sensor network system has emerged as a popular way to obtain location specific real-time weather and crop information’s corresponding Database Management System (DBMS) (U-Agri, 2012 and COMMONsense Net, 2012). These dynamic real-time parameters are particularly needed in micro-management precision agriculture system. Keeping in view the importance of the location specific dynamic real-time crop, soil, weather and environmental parameters in precision agriculture, an attempt has been made with multi-mode (short/long range and twitter environment) distributed sensor network system in the fragile semi-arid tropics in Indian peninsular region. This integrated system, called GeoSense [7], is developed with advanced and simple embedded systems, which will collect, upload and update in the DRTDBMS in periodic manner. This much needed DRTDBMS will be used in real-time crop yield, water requirement and to understand the impact of climate/seasonal change effects.

Currently, dynamic sensory data are collected at every one minute interval and after every five minute interval the data is uploaded in the remotely placed GeoSense centralized server. At the same, the DRTDBMS is updated at time every minute. The dynamic functionalities like real-time distribution/transmission/storage of data using open-source systems are challenging tasks. Many commercial systems are available to fulfill the above said functionality but are very expensive and require expertise to maintain the system. Using a few open source ICT tools, the proposed objectives (DRTDBMS) has been successfully completed.

The sensory data and database system has been developed in XML (eXtensible Markup Language), a standard information exchange tool on the World Wide Web [22]. One can extract/use/convert the XML based information, which can be a platform independent and could be compatible with user-based/defined model/database system. At present, many agricultural RDBMSs rely an off-the shelf technology (off-line) (DSSAT, APSIM and CRAM, etc.,). In this paper, the potentiality of wireless sensor network (WSN) as information gathering and dissemination technology and developing DRTDBMS using modern languages such as XML, JAVA and PHP, which will help the user community (extension workers/researchers/farmers) for ubiquitous decision making were discussed.

2. DISTRIBUTED WIRELESS SENSOR NETWORK SYSTEMS

WSN is an emerging technology, which has revolutionized the data collection in agricultural research in obtaining real time data from the test bed, which will improve the decision–making
process to a large extent and help user community to draw contingent measures [23]. The research presented here is a part of an ongoing project of the Indo-Japan initiative to develop a real-time push-based decision support system (DSS), called GeoSense, for precision agriculture. GeoSense consists of three different low cost distributed WSN systems such as Agrisens [17], Fieldserver and FieldTwitter/Open FieldServer [12] to obtain micro-climatic parameters in real time mode. The dynamic real-time database management system, developed for the above spatial-temporal proximal systems, will improve the agricultural decision making and apply coping strategies to combat the threats from climate changes and extensively used in real-time agricultural monitoring for various aspects.

2.1. Agrisens (AS)

Agrisens (AS) is a zigbee based self-organizing short-range wireless sensor network communication system. It comprises several sensors to obtain environmental/weather/agricultural parameters. It is a pre-defined self-program system (close-loop) that communicates with neighboring AS and ultimately to a central AS, called Stargate [4]. Stargate stores the sensory data and sends it to the remotely placed centralized GeoSense server in FTP (GPRS service) mode. In the centralized server, the received AS data was formatted in to the defined structure (Txt/PostgreSQL/Excel) using C++ and JAVA languages. In addition, PHP algorithm was used for generating graphical data from the dynamic database. The sensory data is collected automatically at regular intervals, which can be customized depending on the sampling user requirement. Figure 1 illustrates the AS with its data processing flow.

![Figure 1. Agrisens and its data flow](image)

2.2. FieldServer (FS)

The 3rd generation FieldServer (FS) [9] has been used in the present study, which is a WiFi (long-range communication) based self-organizing distributed sensing device with 24 channels to sense various weather, agricultural and environmental parameters (Figure 2). Sensors used in the present study include air-temperature, humidity, relative humidity, solar radiation, leaf wetness, soil moisture and CO2 concentration. The sensory data collected at customized regular intervals gets transmitted and stored in the main/parent FS in the form of XML and with front end in HTML Java interfaces (Futkatsu and Hirafuji, 2005). The parent FS, which is equipped with Fit2PC [3] (agent box), will store and transmits/shares all files through virtual private network (VPN) (Packetix, 2012) and Opera-Unite based cloud services [14] to the GeoSense centralized server. The received FS raw data (XML) is appended to the sensory database at every one minute interval. To execute this function, the PHP based algorithm was developed to read new files. This new file is identified on the basis of file name i.e. while storing the file in the database; modify the file name with system (Fit2PC) date and time in such a way that file can be identified and appended in the FS database. Also, algorithms were developed with PHP and Java languages to
convert raw (analog to digital) to real (usable units) sensory data; and these data (raw as well as real) are stored in phpMyAdmin (SQL) database. A provision has been made to export the converted real data into open source consortium (OSC) data formats such as CodeGen, CSV, Excel, Word, Latex, Open Document Spreadsheet, Open Document Text, PDF, SQL, Texy Text, XML, YAML. Figure 2 also depicts the data flow in FS.

2.2.1. Fs-based FluxTower (FLTs)

Two FLT's were deployed in maize field (Figure 3) to study the weather profiles and partitioning of energy into different fluxes (Latent Heat Flux, Sensible heat Flux, Ground Heat Flux). Each Flux Tower consists of three sensor modules with temperature, relative humidity and CO$_2$ concentration sensors at 03 different heights (1m, 2m and 3 m). Real time knowledge of weather profiles and energy fluxes allow farming community to calculate water requirement (ET), irrigation scheduling, pest and disease management, etc. [10]. Flux tower sensors were embedded with FieldServer Engine (FSE) board (one of the components of FS) and are in parallel connection with FS with registered jack (RJ) 45 (RJ45) connectors. The associated FS collects and transmits sensory data to the designated server in the same manner as FS.
2.2.2. FieldTwitter (FT)

FieldTwitter (FT) (Hirafuji et al., 2011) comprises (i) Arduino [1] (ii) signal (transmitting to the Internet clouds) through Fon (iii) Algorithm process for FieldTwitter data to the twitter environment.

(i) Arduino: Arduino is an Open-Hardware electronics prototyping platform based on flexible, easy-to-use hardware and software (Arduino, 2011). Arduino is attached with an external handmade soil moisture sensor (probe) at a depth of 15 cm. This is the first attempt in the world in developing an open-hardware based cost effective sensing system and is particularly useful in developing countries where WSN is still a novice and costly technology (Figure 4).

(ii) Signal transmission: In FT, the communication mode consumes more power than any other parts, as it has been customized into WiFi based communication system by using Fon (router) that helps in receiving the internet pockets (3G) from the FieldServer. Subsequently, it Tweets/Transmits the attached sensory data either through gateway or in twitter (Hydbot01) environment. Anyone can follow the FieldTwitter sensory data in Twitter social network [20] in the name of “Hydbot01”.

(iii) Software development: FT sensory data was stored in twitter database [19] in the form of webpage, with XML syntax, which could be useful to maintain FT database. In FT web interface, the sensory data is available in raw format (analog to digital conversion – ADC). PHP-based algorithms have been developed for converting raw data into usable format (units) and store the data in GeoSense database.

2.3. Integrated (AS, FS, FLT and FT) Communication System

The integrated distributed wireless sensor network system, consisting of ASs, FSs, FLTs and FT, was deployed in the test bed, christened as GeoSense. This small to medium scale GeoSense network, for monitoring weather, agriculture and environment parameters, includes 11 distributed sensing devices (6 AS, 2 FS, 2 FLT and 1 FT), 1 Stargate and 1 Fit-PC II. Drivers GPRS/Broadband/3G network were used for accessing/sharing the data/system at a field/farm level. The sensory data from the test bed is transferred to the GeoSense server through a dedicated asymmetric digital subscriber line (ADSL). This is also the first approach to combine wireless broadband/3G and WSN technologies in India. Figure 5 illustrates the overall architecture of GeoSense. The deployed GeoSense system provides a continuous and dynamic communication between sensing devices (AS/FS) without pocket lost (communication signal) and centralized GeoSense server. These servers process raw sensory data into the real/usable format in dynamic manner. This real-time dynamic distributed sensory data (raw as well as real/usable) can be
accessed through GeoSense web portal (GeoSense, 2012) (for FS, FLT, AS and FT) by any authorized/registered member (rural extension/farming community/decision makers).

Figure 5. GeoSense Communication Flow

3. STUDY AREA & EXPERIMENTS

Distributed sensing systems were deployed in a research form in semi-arid tropics of southern India. In order to study the precision agriculture (crop yield modeling and irrigation aspects), experiments were carried out on maize (Dekalb Super 900m cultivar), groundnut (TMV-2) and rice (MTU 1010) crops. Standard agriculture experimental design was laid out in the test bed, with different irrigation systems (rainfed, ridge & furrow, drip irrigation as per crop water requirement and life saving irrigation) in maize (Kari/t/monsoon agriculture season) and groundnut (Rabi/post-monsoon agriculture season) crops for precision irrigation and crop-yield modeling aspects. In addition, long-term rice experiments were also carried out with different dates of sowing and nitrogen levels to arrive at proper decision making to overcome climate change risks. The distributed wireless sensing systems (FieldServer, Agrisens, Flux Towers and FieldTwitter) were deployed under different crop experiments: FieldServers in rice, groundnut and maize fields for yield modeling. Agrisens were deployed in maize for nitrogen management and groundnut field for pest management, Flux Tower in maize crop field for crop energy balance studies and FieldTwitter in the groundnut field for different date of sowing observations. Weather parameters from the weather station, which is close proximity to the experimental site, will augment the GeoSense researches and also help in validating the sensory data.

4. PROTOTYPE DATABASE MANAGEMENT

Database management application is common in nearly all walks of life. However, maintaining cost effective dynamic real-time farm level database management system is complex and challenging tasks. In this paper, main emphasis is given on the open source cost effective
database management system [15] used to maintain the distributed sensory database from the robust FS and its related systems (FLTs and FTs) in a real-time manner.

4.1. XML in Distributed Sensor Network

Recently, XML has become an emerging standard for information exchange on the World Wide Web and has gained great attention as a database management among database communities. As XML is a self-describing language, one can issue many kinds of queries against its documents from heterogeneous sources and get the necessary information [2]. In addition, it supports high resolution spatio-temporal systems such as GeoSense distributed sensor network (DSN).

4.1.1. XML Protocol

The XML protocol was designed as a generic communication protocol between devices utilizing XML-live tagging. The proximal benefits of an XML protocol are its open and flexible design as well as the ease with which it can be read (Figure 6).

![Figure 6. FieldServer Data Encryption & Decryption of FS Data](image-url)
All data has been processed as ASCII text organized within a user defined tagging structure. An XML parser can interpret messages of different structures provided it can recognize the tagging design. This means that devices can send different message contents, in a different order and understand each other if they use the same XML message tagging design. ASCII messages are an open and portable means of transferring data which is both human and machine readable. The data observed by a FieldServer output is displayed in HTML web page and stored in single HTML file per day [11]. This HTML file contains day wise XML file which consists date and time, location and individual sensory information. The individual XML files are extracted by using string operation.

4.2. XML Based Database Development

Several processing tasks are involved in XML. Unlike in relational databases, where there are some principles in general-purpose language SQL [21], the XML-world is much more diverse. Different types of programming languages for different kinds of tasks exist and for most tasks several competing language proposals are available [18]. The data conversion (any extension file to XML) involves unloading tuples (collection of information about attributes) of relations into sequential files, and formatting them into object-oriented databases (provide persistent storage for objects; a query language, indexing, transaction support with rollback and commit – distributing object into many servers) with constraints preservation [6]. In the present study, XML data are processed using PHP language to store in the distributed sensory networks relational database. The working model (flow chart) is illustrated in Figure 7.

4.2.1. Database Application Programming Interface

A user friendly Application Programming Interface (API) is developed using PHP and Extensible Hypertext Markup Language (XHTML) scripts. This API is designed to the novice registered users (rural extension/farming community/ decision makers) to access the GeoSense DRTDBMS. Figure 8 illustrates the GeoSense (GUI) DRTDBMS (front end) with PHP script (back end).
This user-friendly GeoSense GUI provides an option to select any one of the distributed sensory system database i.e. FS, AS, FLT, FT. The users need to write/mention the SQL statement to obtain selected data. The SQL syntax based queries are often somewhat complex to the novice users. Keeping in view of these users, a query interface (XHTML with PHP) has been developed to formulate the on-demand SQL syntax in simple template format; a functionality with which the user can execute without the SQL knowledge. In addition, the GeoSense GUI additionally provides annual/month/daily data to download on demand with simple selection procedure. The user may also access the query while executing another query. For example, Figure 9 illustrates a screen snapshot of the GeoSense GUI and SQL query (console/command). When the user clicks SEARCH button, the PHP program automatically generates an equivalent SQL query to be processed. The users also submit the split SQL scripts into individual statements that could possibly generate multiple result sets. The result is in hierarchical format for easy reading and navigation.
4.2.2. Compatibility and Interoperability

The GeoSense DRTDBMS real-time sensory data has been developed to the OSC standards [13] (Figure 10), which facilitates the user community to translate/export entire/partial/specific database to the different formats such as CodeGen, CSV, Excel, Word, Latex, Open Document Spreadsheet, Open Document Text, PDF, SQL, Texy Text, XML, YAML.

- CodeGen, it is intermediate representation of source code.
- CSV, The “comma-separated values” file format is a set of file formats used to store tabular data in which numbers and text are stored in plain textual form that can be read in a text editor. Lines in the text file represent rows of a table, and commas in a line separate what are fields in the table’s row.
- Excel, it helps to develop features calculation, graphing tools, pivot tables and a macro programming language called Visual Basic for Applications. It has been a very widely applied spreadsheet for most of the platforms.
- Word, (word processor)
- LaTeX, is a document preparation system for high-quality typesetting. It is the most often used for medium-to-large technical/scientific documents and can also be used in almost any form to publish documents.
- Open Spreadsheet, open office spreadsheet
- Open Document Text, open office text document
- PDF, portable document format for easy handling
- SQL, structure query language (SQL) is a standard language for accessing and manipulating databases.
- YAML, it is a recursive acronym for "YAML Ain't Markup Language". Early in its development, YAML was said to mean "Yet Another Markup Language". It purpose as data-oriented, rather than document markup and XML.

Figure 10. GeoSense DRTDBMS with OSC Standard Output Option Page
5. SUMMARY & CONCLUSIONS
Keeping in view of the importance of DRTDBMS in dynamic agricultural systems and its novice stakeholders, a web-based user-friendly DRTDBMS prototype has been developed with potential WSN. The system, which is christened as GeoSense DRTDBMS, will augment the decision making processes ubiquitously and effectively with the high resolution spatio-temporal dynamic crop/weather/environmental sensory data. Moreover, the GeoSense DRTDBMS has been developed with OSC standards, which will help in modeling services i.e. crop water requirement, simulation model for rice and weather relation, energy flux studies, pest & disease forecasting, etc. for interoperability. Although some of the users are familiar with SQL syntax to obtain the required sensory parameters, quite a few of them are also unable to write a proper SQL syntax to execute the required parameters in a customized manner. To help such users, a queries based GUI (template type) has been developed, a functionality with which the user can generate SQL syntax (without SQL knowledge) with ease and can access/view/download/modify dynamic databases. The cost-effective DRTDBMS is fast and easy to analyze the data in object oriented format. The GeoSense DRTDBMS displays real-time sensory data output in a separate interface with user defined option in hour/day/week/monthly/season/annual manner. This provision could be useful in precision agriculture system, where one needs data dynamically, and identify the suitable agricultural model under the present climate change scenarios. It also helps the user community in day to day decision making processes such as irrigation scheduling, crop yield modeling and fertigation, etc.

ACKNOWLEDGEMENTS

REFERENCES
OpenFS (Open Field Server) for High-throughput Phenotyping”, Proceedings of the SICE Annual Conference, IEEE Catalog No. CFP11765-DVD, pp 2090-2092.


Author:

D. Sudharsan

Sudharsan. is a research scholar, Centre of Studies in Resource Engineering, Indian Institute of Technology Bombay, with extensive experience and management skills in Geographical Information and Communication Technology (Geo-ICT) and Wireless Sensor Network System (WSN) technologies. In his doctoral research work, he carried out development of cost effective dynamic real time sensory network/communication/database management based decision support system. Sudharsan has M.Sc. in Geology from the Presidency College, Chennai Tamilnadu, India (1994) and M.Tech in Bharathidasan University, Trichy, Tamilnadu, India (2006). He also has research interaction with International teams.