Paper Reference No.: PN-205

Title: An Infrastructure for Mobile Weather Data Management
Authors: Wenqiang Wang, Hock Beng Lim, Mudasser Iqbal, Yuxia Yao

Name of the Presenter: Wang Wenqiang

Mailing Address:
Intelligent Systems Centre
Nanyang Technological University
Research Techno Plaza, BorderX Block, Level 7
50 Nanyang Drive, Singapore 637553

Tel: +65-67906678
Fax: +65-63164480

Biography:

Wenqiang Wang obtained his BSc (Hons) degree in Computer and Information Science from the National University of Singapore in 2000 and the PhD degree in Computer Science from the National University of Singapore in 2006. Currently, he is a research fellow in the Intelligent Systems Center, Nanyang Technological University. His research interests include sensor grid infrastructure design and implementation, and data management for large-scale wireless sensor networks.

Hock Beng Lim is program director of the Intelligent Systems Center at Nanyang Technological University, Singapore. He received his BS in Computer Engineering, MS in Electrical Engineering, and PhD in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign, and his MS in Management Science and Engineering from Stanford

University. His research interests include sensor networks and sensor grids, cyber-physical systems, ambient intelligence, parallel and distributed computing, wireless and mobile networks, computer architecture, embedded systems, performance evaluation, e-Science and high-performance computing.

Mudasser Iqbal received his Ph.D. in Computer Science in 2008 from Monash University, Australia and his B.Sc. (Hons) in Computer Science in 2003 from the University of Engineering and Technology, Lahore, Pakistan. He is currently a research fellow in the Intelligent Systems Center at Nanyang Technological University. His research interests include design and implementation of real life sensor grids for environmental monitoring, earth sciences and health care; middleware and embedded solutions for data acquisition, processing and communication; service oriented and semantic web architectures in sensor networks; cyber-physical systems and large-scale ambient smart environments.

Yuxia Yao received the BSc degree in Computer Science from Huazhong University of Science and Technology, Wuhan, China in 2003, and the PhD degree from the School of Computer Engineering at Nanyang Technological University, Singapore, in 2008. She is currently a research fellow in the Intelligent Systems Center, Nanyang Technological University. Her research interests include mobile computing, data management in wireless sensor networks and spatial databases.

An Infrastructure for Mobile Weather Data Management

Wenqiang Wang, Hock Beng Lim, Mudasser Iqbal, Yuxia Yao
Intelligent Systems Centre, Nanyang Technological University
[wqwang ,limhb, mmiqbal, yxyao]@ntu.edu.sg

Abstract

With the rapid development of wireless communication technologies and mobile devices, accessing the web and Internet using smart phones or PDAs is becoming increasingly popular. However, due to the limitations of current technologies, mobile internet connection still cannot be compared with traditional fixed line connection in terms of upload/download bandwidth and stability. It is still a major challenge to deliver real-time information to a large number of mobile devices effectively and efficiently. We believe that a well designed mobile data management infrastructure can play a crucial role in solving this issue. In this paper, we will present such a data management infrastructure specifically for the management of weather data and the delivery of these data to mobile devices.

Under the National Weather Study Project (NWSP), hundreds of mini weather stations have been deployed in schools across Singapore. These weather stations generate weather data in different formats and at different rates. Our infrastructure collects data from these weather stations, processes it and delivers weather information to mobile devices in real time. We address several major challenges in mobile data delivery such as network instability and low bandwidth using our infrastructure. The infrastructure is based on our work on sensor grid technologies to facilitate sensor data processing, data management, and data query processing.

We have also developed a mobile web portal and applications for users to easily access weather data using mobile devices. We have also developed several services for users to make use of the weather data. A set of APIs has been developed and published for developers who are interested in accessing the weather data. Besides managing weather data, our data management infrastructure is capable of handling other types of data from heterogeneous sensors due to its novel design and implementation.

1. Introduction

By the end of 2008, the number of mobile phones in use had reached four billion world-wide [1]. Among all the mobile phone users, over 600 million are 3G subscribers and this number is increasing rapidly as more and more countries issue 3G licenses to the telecom market. Today's mobile devices, such as Google phone, iPhone, etc., are much more powerful than their predecessors in terms of memory, processor speed and various user features. Thus, they are not only simply used as communication tools, but also as personnel portals to receive information through the 3G network. People can browse web pages, watch on-line videos, and use on-line applications on their mobile devices the same way as they do on normal computers. There is no doubt that accessing the web and Internet using mobile devices and PDAs is becoming increasingly popular. Mobile devices are also ideal platforms for real-time information dissemination and update, which are gaining increasing importance in practical applications such as early warning of natural disasters, traffic management, business advertisements, etc. Currently, the primary means of providing such information is through SMS, which is costly and lacks in interactivity. With the development of wireless broadband networks, we can expect better solutions.

However, mobile internet connection still cannot be compared with traditional fixed line connection in terms of upload/download bandwidth and stability. Current 3G networks support a maximum of 14.4Mbit/s data rate [2] for stationary users. In comparison, for fixed line networks, 100MB/s is considered as the basic requirement. Moreover, environmental factors such as blockage by tall buildings and natural phenomena such as heavy thunderstorms can also greatly weaken the wireless signal strength causing network instability or even unexpected disconnection from the 3G network.

Therefore, it is still a major challenge to deliver real-time information to a large number of mobile devices effectively and efficiently. We believe that a well designed mobile data management infrastructure can play a crucial role in solving this issue. In this paper, we discuss the practical scenario of delivering real time weather information to mobile devices. Hundreds of mini weather stations have been deployed in schools across Singapore under the National Weather Study Project (NWSP) [3]. These weather stations generate weather data in different formats and at different rates. We have designed and implemented a data management infrastructure specifically for managing weather data and delivering them to mobile devices. The infrastructure has the following features:

- The infrastructure collects weather data from mini weather stations, processes it and delivers weather information to mobile devices in real time.

- It is based on our previous works on sensor grid technology to ensure scalability of weather data processing.

- It provides APIs for external applications to access weather data in an easy yet efficient manner. We have developed a mobile web portal and several user applications using these APIs.

- The infrastructure is capable of handling other types of data from heterogeneous sensors with its novel design and implementation.

The rest of this paper is organized as follows. Section 2 presents the background information of this work. We discuss the key design challenges for the proposed infrastructure in Section 3. In Section 4, we present the infrastructure with its architecture, major components and also applications and services built on top of it. Finally, we conclude this paper in Section 5.

2. Background

2.1 The National Weather Study Project (NWSP)

News and articles regarding global climate change appear in press and media more and more frequent in recent years. Not only scientists and leaders, but common people also have started to worry about global warming, El Niño phenomenon, melting of icecap and many other phenomena caused by global climate change. As an island country, Singapore is vulnerable to the hazardous consequences of global climate change and is therefore very much concerned about this issue. In 2006, Senoko Power, one of the largest local power companies, initiated the National Weather Study Project [3]. This project is a community-based environmental initiative that aims to promote the awareness of Singapore's younger population, especially school students, about climate change, global warming and their relationship with the environment. Under this project, each participating school has been provided with a mini weather station for the school students to carry out various projects related to the study of the climate. At this point of time, over two hundred mini-weather stations have already been installed in schools across Singapore with the Vantage Pro 2 weather stations deployed in primary and secondary schools and WeatherHawk weather stations deployed in junior colleges.

2.2 Sensor Grid

Sensor grid is the integration of wireless sensor networks and grid computing infrastructures – the two cutting edge state-of-the-art technologies. The sensor grid enables the collection, processing, sharing, visualization, archival and searching of large amount of sensor data. We have developed a sensor grid architecture framework, called the *Scalable Proxy-based aRchItecture for seNsor Grid* (*SPRING*) [4]. It is a flexible framework that is not constrained by the characteristics and requirements of specific target applications. By using proxies as interfaces between the sensor networks and the grid fabric, the SPRING architecture can support a wide range of sensor devices, including the less computationally powerful ones. Based on the SPRING framework, we have designed and implemented the *National Weather Sensor Grid* (NWSG) [5] for real time collection, archival, processing and visualization of weather data from the network of weather stations under the NWSP.

3. Design Challenges

This section discusses the important issues and challenges in the design of the proposed data management infrastructure.

3.1 Mobile Network Connectivity and Bandwidth

In conventional fixed line networks, connections to the Internet are reasonably reliable and unexpected disconnections do not occur very frequently. However, this is not the case in the case of accessing the internet via mobile devices. The mobile Internet connection is dynamic in nature, and at times, it may get very intermittent and susceptible to faults due to noise and signal degradation caused by environmental factors and natural phenomena.

International Telecommunication Union (ITU) states that 3G networks are expected to provide the data rate of minimum 2 Mbit/s and maximum of 14.4 Mbit/s for stationary users, and 348 Kbit/s in moving vehicles. However, these theoretical bandwidths are rather difficult to achieve in practice because multiple users simultaneously share the same bandwidth. Besides, upload bandwidth is much less than the download bandwidth in most situations.

The data management framework needs to take these issues into consideration when supporting mobile applications which require frequent data exchange. It should provide support to applications so that they can avoid failure or stop responding due to network disconnection or high latency.

3.2 Scalability

Scalability is the ability to add additional data resources to the infrastructure in order to increase the capacity of data collection, without substantially changing the underlying framework. Since weather data collection from different types of weather stations may vary substantially in terms of format and type, the infrastructure should be able to provide them to applications in a unified form so that the underlying data formats and types are kept transparent to applications.

3.3 Security

Similar to the fixed line networks, mobile networks are also vulnerable to internet attacks such as denial of service, etc. Mobile network security needs to be taken into consideration in the framework design. In addition, if the data is confidential, the user applications should obtain authenticated before the data is delivered to the mobile device. Thus, it is necessary to ensure that proper authentication mechanism is implemented in the infrastructure.

3.4 Quality of Service

Quality of Service (QoS) is a key issue that determines whether the proposed infrastructure can efficiently provide the weather data on demand. Weather stations under the NWSP project are all installed in schools; therefore, their availability is subject to the school's IT/power management policy. Moreover, there are a number of areas in Singapore where there is no school where a weather station can be deployed, such as the central containment area, etc. In the situation that a user application requires weather data from a specific area at a specific resolution (i.e. in terms of sampling rate), the infrastructure should be able to determine which weather station (or group of weather stations) to query in order to meet application's data needs.

3.5 The kNN problem

The classical kNN problem has its application in designing such an infrastructure. Finding the nearest weather stations around the mobile devices is likely to be the first step in the work flow of many weather related mobile applications. Moreover, as mentioned in Section 3.4, the weather stations under the NWSP cannot guarantee 100% availability. The status of weather station and its availability patterns must be taken into consideration when employing kNN algorithms.

4. The Infrastructure

The proposed mobile data management infrastructure is designed in a layered, proxy-based approach based on our previous work on the sensor grid. In the following sub sections, we will introduction its architecture, major components and services build on top of it.

4.1 Architecture

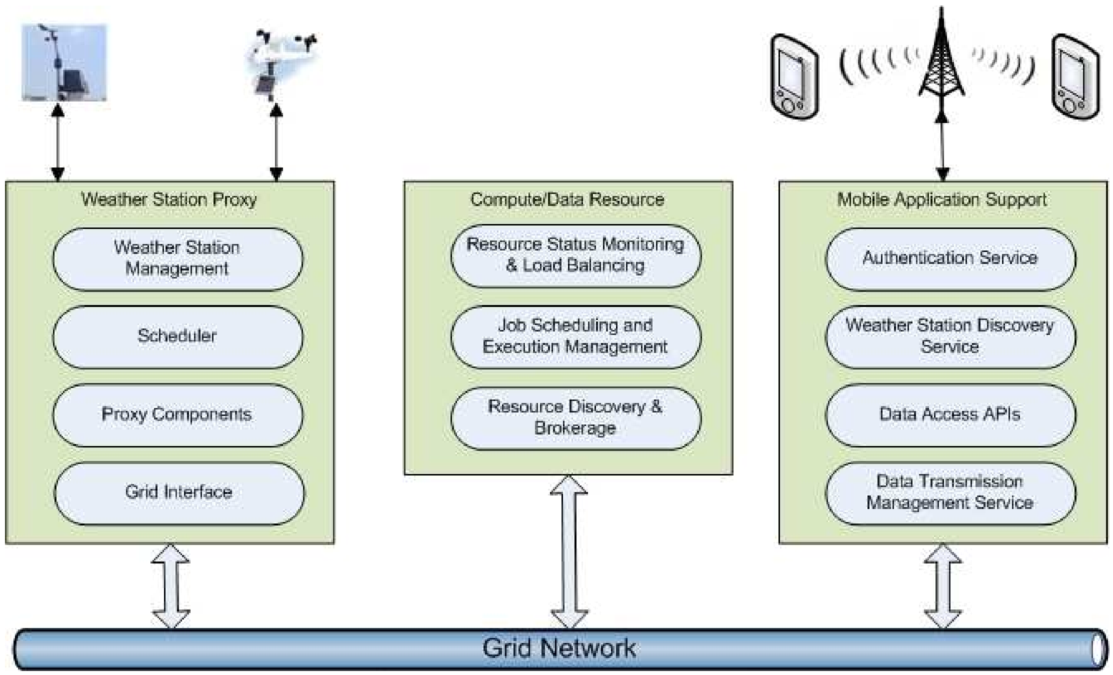

Figure 1. Architecture of the proposed infrastructure

Figure 1 shows the architecture of the proposed mobile weather data management infrastructure. The infrastructure makes use of proxies to interface weather data resource with the underlying sensor grid framework. The proxy serves several important functions. First, it exposes the data resources as grid services which can be discovered and accessed by higher level applications. It also translates weather data from its native format to a suitable data exchange format such as XML. Second, the proxy coordinates the connection between weather stations and the grid. It uses techniques like caching, buffering and link management to minimize the possibility of connection failure due to unexpected network disconnection or high network latency. Third, it greatly improves the scalability of the infrastructure. New types of weather stations can be integrated into the existing system by adding new proxies. Finally, it also provides an interface for various underlying services such as station status monitoring, active data collection, etc.

Compute/data resources are responsible for storing, analyzing and processing weather data transmitted through the proxies. The infrastructure takes care of issues such as resource discovery and brokerage, resource monitoring and load balancing, job scheduling, job executing management, etc.

Mobile application support component plays the major role in delivering real time weather information to mobile devices. It has the following modules.

- *Authentication*

This module is in charge of handling authentication requests from mobile user application. It interprets and sends these requests to the authentication engine of different compute/data resources in the infrastructure.

- *Weather Station Discovery*

  This module is responsible for locating most relevant weather station resources to answer incoming requests. The requests may come from the mobile applications directly or from other modules/services within the infrastructure. There are a number of factors to consider when determining the most relevant stations, such as geographical locations, current status, data availability, sampling frequency, etc.

- *Data Access*

  The data access module provides pull/push based access to real time as well as archived weather data. The data is provided in various formats depending on the application requests. Plain ASCII test, XML, light semi-structured format, or even binary streams are different data formats supported by the proposed infrastructure.

- *Data Transmission Management*

  This module manages the procedure of transmitting large amounts of weather data to mobile devices. It uses techniques such as caching and buffering to minimize the data loss due to unexpected network disconnections during the data transmissions.

4.2 Implementation

Currently, the *Weather Station Proxies* have been implemented using two approaches. In schools that have suitable conditions, a PC is used as the proxy. Programs running on the PC are responsible for collecting weather data from weather stations on a pre-determined interval and convert them into suitable formats before sending them to the sensor grid. This approach is cost-effective for schools because the proxy software components require minimum system resources and are easy to install and maintain. However, due to the school's IT/power management policy, it's very difficult to keep the PC powered on at all times. During the period when the PC is switched off, no real time weather information can be collected. Data collection only resumes when the PC is switched on again. To address this issue, we have also developed an embedded device to replace the PC. This embedded device is very power efficient so it can be turned on all the time. In addition to running the similar proxy components, the key advantage that the embedded device provides is the remote access to the weather station through secure SSH tunnels. This remote access allows the sensor grid to reconfigure various weather station parameters (such as the sampling rate) and also query the data as and when needed by different mobile applications.

The *Compute/Data Resource* component consists of various servers and databases. Several important services have been implemented to ensure that these resources are effectively and efficiently used to handle application requests and data management jobs. The *Resource*

Discovery and Brokerage service provides seamless access to weather data and computing resources, irrespective of their specifications and locations. It determines the most relevant weather station to resolve applications' data requests and also discovers computational and storage resources when needed. *The Monitoring and Load Balancing* service watches system health and detects possible failure based on tested schemes like heartbeats. It also assigns computational/storage jobs based on system loads such as average CPU utilization, to different servers/databases.

The *Mobile Application Support* component provides support for mobile applications to access weather data from the infrastructure. In the current phase of implementation, web service is used as the main method for data access. Due to the characteristics of mobile network connection mentioned in previous sections, changes to the default web service mechanism are necessary. For example, depending upon the available bandwidth, weather data should be automatically aggregated to smaller sizes and a lighter data format used so that long transmission delays can be avoided.

4.3 API

An Application Programmers' Interface (API) is designed to provide weather data to mobile applications such that the underlying data collection and processing are kept transparent to the applications. It provides the following key functions:

- *GetStationList*

  This function simply returns the unique ids and names of all the weather stations currently connected to the infrastructure.

- *GetStationMetaData*

  This function returns parameters that are defined to describe a weather station.

- *GetStationInfo*

  This function returns the meta information of a specific weather station. The returned results are a set of (*name*, *value*) pairs.

- *GetNearestStation*

  This function returns the nearest *k* weather stations with respect to a fixed location. It takes a *(location, k)* pair and a set of *(name, operator, value)* pairs as input, where the (*name, operator, value*) are used to specify meta information of a weather station. Unique identifiers of the weather stations that satisfy both the kNN condition and the meta data conditions are returned as results.

- *GetWeatherData*

This function serves the main method of providing weather data to mobile applications. It takes a station id and a set of *(name, operator, value)* pairs as input, where the (*name, operator, value*) pairs specify the parameters that are sought as part of the of the weather data.

- *UploadWeatherData*

  We foresee that it will not be long before the mobile devices are equipped with some basic weather or other environmental monitoring sensors. This function is designed to handle weather data uploaded directly from such devices. Currently, it takes a chunk of fixed format weather data as input, process it and makes it available to the entire infrastructure.

4.4 Web Portal and Application

Limited processor power, low volume of on board RAM and small screen sizes of mobile devices pose challenges in the development of a mobile weather data portal. The development of a mobile web portal is different from normal web portal design in several aspects. First, the interfaces need to be friendlier than normal web sites as QWERTY keyboards and mice are not available on most mobile devices. Navigation mechanism and input methods should be redesigned to accommodate normal mobile phone keyboard. Second, the layout of the web pages requires redesign due to small screen sizes of mobile devices. Third, the use of web designing techniques such as java scripts, flash objects, etc needs to be reconsidered as the browsers used on mobile devices may not be powerful enough to supports these features.

Figure 2. Screenshots from the web portal

Figure 3. Android Map Viewer Application

Figure 2 shows a few screen shots from the web portal that we are developing for the mobile devices. From left to right, the screenshots show weather summary interface, station information interface and data download interface. We use geo-centric web interfaces such as Google map for visualization purposes. Users can obtain real-time weather information and relevant weather parameter values from the portal. For users that need archived weather data for research purposes, the portal provides a flexible data download services. The web portal can be accessed at "http://nwsp.ntu.edu.sg" using mobile phones or PDAs.

For users using Google Phone, we have also developed an easy-to-use android application for users to access weather data on a map-base interface. Figure 3 show a screenshot of this application.

5. Conclusion

In this paper, we have presented a data management infrastructure that aims to provide weather information to mobile devices. This infrastructure is based on the SPRING sensor grid framework and is capable of real-time collection, processing and delivery of weather data to mobile devices. We have examined key design issues and challenges for such an infrastructure under mobile networking environments. A prototype of the proposed infrastructure was implemented and a web portal and various applications were developed on top of it.

Based on the current prototype, we plan to develop more mobile applications or games to exploit the capabilities of the infrastructure to its full potential. We will also develop more features and services for the mobile web portal. In addition to the weather data, we plan to integrate real-time data feeds, such as air quality and environmental haze, into our infrastructure so that they can be used for a wide range of complex mobile applications.

Acknowledgements

This work was supported by the Singapore National Research Foundation (NRF) under grant number NRF2008IDM-IDM001-005.

References

[1] Wikipedia: List of Countries by Number of Mobile Phones in Use, http://en.wikipedia.org/wiki/List_of_countries_by_number_of_mobile_phones_in_use

[2] "Cellular Standards for the Third Generation". ITU. 2005-12-01.

[3] "National Weather Study Project", http://nwsp.ntu.edu.sg

[4] H. B. Lim, Y. M. Teo, P. Mukherjee, V. T. Lam, W. F. Wong, and S. See, "Sensor Grid: Integration of wireless sensor networks and the grid," *Proc of the 30th IEEE Conference on Local Computer Networks (LCN 2005)*, pp. 91-98, Nov 2005.

[5] H. B. Lim, M. Iqbal, W. Wang, and Y. Yao, "The National Weather Sensor Grid: A large-scale cyber-sensor infrastructure for environmental monitoring," To appear in the *International Journal of Sensor Networks (IJSNet), Special Issue on Environmental Sensor Networks*, Inderscience, 2009.