Real-World Sensor Network for Long-Term Volcano Monitoring: Design and Findings

Renjie Huang, Wen-Zhan Song, Mingsen Xu, Nina Peterson, Behrooz Shirazi, Richard LaHusen

Abstract—This paper presents the design, deployment and evaluation of a real-world sensor network system in an active volcano - Mount St. Helens. In volcano monitoring, the maintenance is extremely hard and system robustness is one of the biggest concerns. However, most system research to date have focused more on performance improvement and less on system robustness. In our system design, to address this challenge, automatic fault detection and recovery mechanisms were designed to autonomously roll the system back to the initial state if exceptions occur. To enable remote management, we designed a configurable sensing and flexible remote command and control mechanism with the support of a reliable dissemination protocol. To maximize data quality, we designed event detection algorithms to identify volcanic events and prioritize the data, and then deliver higher priority data with higher delivery ratio with an adaptive data transmission protocol. Also, a light-weight adaptive linear predictive compression algorithm and localized TDMA MAC protocol were designed to improve network throughput. With these techniques and other improvements on intelligence and robustness based on a previous trial deployment, we air-dropped 13 stations into the crater and around the flanks of Mount St. Helens in July 2009. During the deployment, the nodes autonomously discovered each other even in-the-sky and formed a smart mesh network for data delivery immediately. We conducted rigorous system evaluations and discovered many interesting findings on data quality, radio connectivity, network performance, as well as the influence of environmental factors.


1 INTRODUCTION

Wireless sensor networks have been attracting increased interest from the research community for a broad range of applications [17]. Wireless sensor networks have the potential to greatly enhance the understanding of volcanic hazards by permitting large distributed deployments of sensor nodes in difficult-to-reach or hazardous areas [16]. Wireless networking allows sensor nodes to communicate with each other and to a central base station via a self-healing mesh network, allowing intelligent real-time data reduction, data archival, as well as the re-tasking of the array after deployment. In remote volcano monitoring, there is typically no infrastructure available and maintenance is extremely hard. The sensor network must be able to run continuously with zero-maintenance for a long period in a hostile volcanic environment.

In this paper, we present the design, deployment and evaluation of a real-world sensor network system for long-term volcano hazard monitoring. Our sensor network has been deployed and tested on Mount St. Helens since July 2009, as part of the Optimized Autonomous Space In-situ Sensorweb (OASIS) [1] system. The OASIS is a prototype system that provides scientists and decision-makers with a tool composed of a smart ground sensor network integrated with smart space-borne remote sensing assets to enable prompt assessments of rapidly evolving geophysical events in a volcanic environment. This paper describes the design and deployment of ground sensor networks. In 2008, we conducted a trial deployment [16] with 5 stations as a proof-of-concept with basic functions including UTC-time synchronized data acquisition, agile data collection routing, and reliable command dissemination. Learning lessons from that deployment, we have significantly improved the system functions and intelligence, and successfully conducted a larger scale deployment into the crater and around the flanks of Mount St. Helens in July 2009. In this paper, we comprehensively review our system design and deployment experience and lessons, especially after trial deployment [16].

The rest of this paper is organized as follows. Section 2 presents the system architecture of our field deployment on Mount St. Helens. Section 3 describes the hardware design of sensor nodes. Section 4 and Section 5 present the software design. Section 6 presents the rigorous system evaluation and various interesting findings. Finally, we discuss related works in Section 7 and conclude the paper in Section 8 by pointing out our future work.

2 SYSTEM OVERVIEW

Figure 1 illustrates the end-to-end configuration of the full OASIS system. The ground sensor network delivered real-time volcanic signals to the sink nodes at JRO (Johnston Ridge Observatory) through multi-hop relays. The
sink nodes are connected to the gateway through serial connection. The gateway (MOXA device server DE-304) relayed the data stream to a WSUV server through a microwave link of 50 miles. In the lab, a customized TinyOS tool SerialForwarder in WSUV server forwards the data between the sensor network and the Internet. Multiple control clients may connect to it, access the sensor data stream, and control the network in real time. V-alarm is a volcano activity alarm system, which can automatically identify earthquake events from the raw data stream. Once an event is triggered, V-alarm can send event alerts via email or text messages to the corresponding scientists in charge. The Command & Control center is for situation awareness and integration of in-situ sensor network and space observations from EO-1 satellite. It incorporates existing real-time volcano monitoring and data-processing tools used by the USGS (U.S. Geological Survey) and makes real-time autonomous operational decisions to control the sensor network according to local and remotely sensed environmental changes.

The real-time data stream from seismic, infrasonic, lightning, and GPS sensors, as well as RSAM, battery voltage and RSSI/LQI data, are imported into a MYSQL database with UTC timestamps of millisecond resolution. In connection with the database, a web application VALVE [2] was developed to display not only current, but also historical data to a large number of distributed users. Our database is well integrated with USGS’s science analysis software (e.g. VALVE) to manage and visualize the volcano monitoring data. VALVE (Volcano Analysis and Visualization Environment) is a on-demand client/server system for visualization of serving, graphing, and mapping nearly every type of historical data collected by the sensor network. It greatly facilitates our analysis of the data quality and network status. Our data is shared with the community through the VALVE web client. It allows users to visualize or download the volcanic data of a specific period from any location on the Internet. The system collected about 60 GB volcanic data in the first 6 months during the deployment.

3 HARDWARE DESIGN

The hardware design of the OASIS station has considered various environmental challenges in volcano hazard monitoring with the direct involvement of experienced USGS engineers. Figure 3 shows the OASIS station. A wireless mote iMote2 is the core of OASIS station. iMote2’s PXA271 processor is configured to operate in a low voltage (0.85 V) and low frequency (13 MHz) mode. Each station contains a GPS receiver (U-Blox LEA-4T L1) to pinpoint the exact location and measure subtle ground deformation, a seismometer to detect earthquakes, an infrasonic sensor to detect volcanic explosions, and a lightning sensor to detect eruption clouds. The accuracy of the GPS receiver’s time pulse is up to 50 ns providing accurate timing for TDMA and data timestamp. The infrasonic sensor (All Sensor Model 1 INCH-D-MV) is a low range differential pressure sensor to record low frequency acoustic waves generated during explosive events. Lightning typically accompanies heavy ash emissions, so lightning is a useful parameter for monitoring volcanic activity. In this deployment, we used two types of seismometers: low-cost MEMS accelerometer (Silicon Designs Model 1221J-002) and geophone sensor (Geospace HS-1). The geophone sensor has been used as a seismometer by USGS and has ultra low noise level. However, they are difficult to be deployed since they must be hand-placed in a perfect vertical position.
The accelerometer sensor is suitable for fast air-drop deployment, but it has higher noise level compared to the geophone sensor. The measured average energy consumption of the OASIS station is about 375 mW. Powered by AIR-ALKALINE batteries of 1200 Ah Capacity at a voltage of 3V, the estimated lifetime of the OASIS station is about 400 days. The station communicates with each other via a 802.15.4 radio.

Learning lessons from the trial deployment [16] in 2008, we upgraded the radio amplifier and added a lightning protector. We originally used the 2.4 GHz bi-directional Amplifier SmartAmp to increase the transmission range. However, it does not work well because the required minimum TX input power is 0 dBm, while the output power from the iMote2 mote is typically about −3 dBm after attenuation of connector. Moreover, the current consumption power of SmartAmp is as high as 540 mA in TX mode and 60 mA in RX mode at 7.5 V. To meet the energy requirement of one-year unattended operation, we built the amplifier customized for OASIS using a cost-effective and high performance RF front end CC2591. The current consumption of CC2591 in TX and RX mode are typically 100 mA and 1.7 mA at 3 V respectively, while obtaining similar gains compared to SmartAmp. We added a coaxial lighting arrestor (NextTek Surgeguard) to reflect the lightning energy after we found that lightning strikes can destroy the amplifier. In addition, previous deployment a 900 MHz Freewave radio modem connected the sink station to the gateway over a 6-mile radio link. Later the field test showed that the 802.15.4 radio with our customized amplifier can achieve similar transmission range, so the 900 MHz Freewave radio was removed from the hardware configuration to conserve energy.

**Fig. 3.** The whole OASIS station.

4 ROBUSTNESS AND REMOTE NETWORK MANAGEMENT

To survive unforeseen software faults, our sensor node automatically detects and self-recover from software failures. Also, considering the longevity and remoteness of environmental monitoring, online reconfiguration of the network and motes is highly desirable for system management. Thus, we developed a comprehensive remote network management mechanism that provides interactions between users and the network in the field.

4.1 Automatic Fault Detection and Recovery

All of our nodes are in rugged terrain and only reachable by helicopter. The field maintenance is difficult, if not impossible. Thus, software dependability and reliability is a major concern. Nasty bugs may occur after deployments [7], [20]. It is, therefore, crucial to have an exception-handling mechanism to recover nodes automatically from software and hardware failures. We exploited the benefits of watchdog. The iMote2’s hardware watchdog can restart the node under exceptions such as dead loop, memory errors, and stack overflow. In addition, software failures can also be caused by unexpected logic errors. For example, corrupted packets may result in time desynchronization, and corruption of communication protocols. We further developed a software watchdog to enable self-recovery from erroneous states. Each node monitors the most important internal logic states and calls a command to reboot by stopping the watchdog timer when it detects erroneous states. However, the reboot operation will cause a node to discard its system state information in memory. To reduce the reboot cost to the minimum, important parameters and states, such as sampling rate and RF channel/power, are written to Flash when configured by remote users, and are restored once a node reboots. With those fault tolerant mechanisms, our system is able to continuously operate normally after the deployment.

4.2 Remote Command and Control

The remote command and control is based on a flexible Remote Procedure Call (RPC) mechanism [22]. It allows a PC to access the exported functions and any global variables of a statically-compiled program on sensor nodes at run time. To ensure the reliable dissemination of RPC messages over multi-hop paths, we have designed a reliable data dissemination protocol Cascades [12]. The RPC mechanism gives users great flexibility to read/write system variables and run any exported functions. Operations such as set/get sampling rate, beacon interval, power level, radio channel, and event report level threshold are provided to remote clients. The RPC mechanism also provides visibility into network failures and helps to correct bugs.

4.3 Configurable Sensing

As an instrument to enhance scientific explorations, the OASIS stations are designed to be smart and configurable with the sensing parameters adjustable based on environmental conditions and mission needs. Our sensor driver performs synchronized sampling operation and maintains sensing parameters, such as sampling rate,
ADC channel, and data priority. All these parameters could be tuned according to environmental and resource situations to conserve energy or increase fidelity. When energy conservation becomes a priority, users can remotely close a non-critical data channel by simply setting the sampling rate to 0. Currently, we collect all raw data for scientific analysis, but users also can change the base data priority to 0, then the OASIS station will send out event data only. If a sensor is broken or the hardware interface is disconnected, its channel can be closed to save energy and bandwidth.

Besides the configurable parameters, it is also useful to configure the data processing tasks for different data types. To support this, each processing algorithm is indexed in a task queue through function pointers. Each data block is associated with a taskCode, indicating the processing tasks to be performed on the raw data samples. By configuring the taskCode, users can flexibly alter the in-network processing for a specific data type on specific nodes. With this mechanism, we can remotely choose to run different event detection and compression algorithms on different types of data without reprogramming the nodes.

4.4 Over-the-air Network Reprogramming

After a field deployment, the network functionality may need improvement or fix new software failures. Thus, it is important to support remote software upgrades. Deluge is the de facto network reprogramming protocol that provides an efficient method for disseminating code update over the wireless network and having each node program itself with the new image. Deluge originally does not support the iMote2 platform, and it is not trivial to port it to support the iMote2 platform [11] (see supplement). We also improve Deluge to ensure that it could handle some adverse situations. (1) Image integrity verification. If a node reboots during the download phase, we have to ensure it correctly resumes the download. To address this issue, we implemented a mechanism where we verify the image integrity during startup. If the image has been completely downloaded, then we continue with the normal operations; otherwise we erase the entire downloaded image and reset the meta data to enable a fresh re-download. (2) Image version consistency. The original Deluge is based on sequence number. However, if the gateway lost track of the sequence number and did not use a higher sequence number, then Deluge will not respond to new request of code update. We fixed this problem by using the compilation timestamp to differentiate new image from old image.

5 Quality-aware Data Collection

For such a high data rate application, a key challenge is how to collect the high-fidelity data subject to the limited bandwidth available to sensor nodes. Adaptive data transmission protocols were designed to ensure higher priority data with higher reliability. Additionally, light-weighted compression and localized TDMA MAC protocol was designed to improve network throughput.

5.1 Priority-aware Data Delivery

In the volcano network, the sensor data is not equally important; thus, we need to treat the data accordingly and control the Quality of Service (QoS). Firstly, we used the STA/LTA (short-term average over long-term average) algorithm [25], [26] to detect seismic events and assign those event data with higher priority. More details about this algorithm can be found in the supplement. Then, we designed a Tiny-Dynamic Weighted Fair Queuing algorithm (Tiny-DWFQ) [13] to assign proper QoS for each packet based on the data priorities and network situations. Once the QoS is assigned, Tiny-DWFQ ensures that the packets are sent throughout the network in a way that the desired QoS requirements are upheld. The dynamic nature of Tiny-DWFQ is exemplified in the assignment of the priority of each packet, called the Dynamic Weighted Priority (DWP). The DWP is a weighted combination of the node and data priorities assigned based on the current context of the environment. Once a packet has been assigned a DWP, it must be scheduled in accordance with its DWP and the current network congestion level. The Tiny-DWFQ algorithm ensures that the high priority packets are transmitted, even in the midst of congestion.

5.2 Reliable Event Data Collection

Seismic event and RSAM data are important for volcano studies, and volcanologists expect them to be reliably collected. Thus, we developed a Reliable Data Transfer (RDT) protocol, on top of the priority-aware data delivery approach in section 5.1. Notice that, the resource and system constraints of sensor network demand a light-weighted design, where TCP [24] can not be directly migrated. Firstly, the bandwidth is severely constrained and may be insufficient to deliver all seismic and RSAM event data during some active periods. Secondly, the feedback control traffic from the data server to sensor nodes shall be as few as possible. The feedback control traffic directly competes the bandwidth of sensor data traffic due to the radio broadcast nature. When the bandwidth is severely limited, the data delivery with RDT could be worse than that without RDT. Realizing those new system challenges, we made several innovative design choices as described in the supplement.

5.3 Network Throughput Improvement

Existing low power MAC protocols, such as B-MAC [14], are designed for low duty cycle applications. To provide high channel capacity utilization and low congestion ratio, we developed a new TDMA MAC protocol called TreeMAC [15] to regulate the channel access. The design
of TreeMAC is based on a key observation of multi-hop data collection networks: the bandwidth allocation of any node shall be no less than that of its subtree, so that the nodes closer to the sink have enough bandwidth to forward data packets for the nodes in its subtree. TreeMAC divides a time cycle into multiple frames and each frame into 3 slots. Parent nodes assign non-overlapping frames to their children nodes based on their proportional bandwidth demands. Each node calculates its own slot assignment based on its hop-count to the sink. This innovative 2-dimensional frame-slot assignment algorithm has the following nice theory properties. Firstly, given any node, at any time slot, there is at most one active sender in its neighborhood (including itself). Secondly, the packet scheduling with TreeMAC is bufferless, which therefore minimizes the probability of network congestion. Thirdly, the data throughput to the gateway is at least 1/3 of the optimum assuming reliable links.

To reduce the bandwidth demands and maximize the data return over the unreliable and low rate radio links, we designed an Adaptive Linear Filtering Compression (ALFC) [9] algorithm to compress the seismic raw data. It is a lightweight compression algorithm tailored for sensor networks with code size only 768 bytes. Considering the relatively modest computational power of existing sensor platforms, ALFC does not use floating-point operations and has very low computation and energy cost. More details about ALFC are presented in the supplement.

6 System Evaluation and Findings

After the deployment, we have conducted rigorous system evaluations and discovered many interesting findings on data quality, radio connectivity, network performance, as well as the influence of environmental factors. Due to space limit, additional system evaluation and findings can be found in the supplement.

6.1 Data Quality Evaluation

One important aspect of evaluating a real-world sensor network system is the data quality. To assess the quality of our data we compared it with that from broadband station VALT, which is the state-of-the-art instrument in seismology. VALT sits inside the crater of Mount St. Helens volcano, and is the closest to OASIS node 1. The analog-to-digital converter for VALT station is 24 bits while that of our OASIS stations is 16 bits. To compare the SNR between VALT station with OASIS node 1, we scaled the seismic reading from VALT by \( \frac{1}{16} \) (e.g., from 24-bit to 16-bit). Figure 4 shows 50-minute seismic raw reading from OASIS node 6, node 1, and VALT station during the time period from UTC 07/20/2009 18:20 to UTC 07/20/2009 19:10. After scaling, the noise level of VALT and OASIS station is almost the same. We can see that the OASIS station with geophone seismic sensor can achieve similar data quality. It is worth mentioning that the seismic sensor of the broadband station costs about $10,000 ($25,000 for the whole station), while the OASIS station only costs about $3000 (including radios and other sensors). In addition, the deviation of the noises of OASIS node 6 is much higher than OASIS node 1 as expected. The reason is that OASIS node 1 is equipped with a high-fidelity geophone seismic sensor that has a low noise level, while OASIS node 6 uses a low-cost MEMS accelerometer as the seismic sensor. The poor SNR (signal-noise-ratio) of seismic MEMS sensors does not meet the requirements from seismologists very well.

The infrasonic sensor is to record infrasound, low-frequency acoustic waves generated during explosive events. During the first 2 months after the deployment, no explosions happened in Mount St. Helens volcano, thus the infrasonic sensor did not detect any explosive events. However, heavy storms were correctly detected by the infrasonic sensors. The lightning sensors also recorded several lightning strikes on 08/26/2009, 08/28/2009, 09/01/2009, and 09/04/2009, as shown in Figure 8. When analyzing the data, we found that the lightning strikes were accompanied by storms. For example, Figure 7 shows that OASIS node 3 triggered lightning events during a heavy storm indicated by the infrasonic RSAM on 08/25/2009. The lightning strikes also destroyed node 3’s radio amplifier, and caused its data stream to stop around 07:30. We added a lightning arrester to protect the amplifier after we realized this problem. Correlating different sensors can help to correctly identify events. For example, with the infrasonic sensor and seismic sensor together, we can tell a volcano explosion (with infrasonic event) from an ordinary earthquake (without infrasonic event).

6.2 Event Detection Accuracy

Broadband station VALT co-locates with OASIS node 1, so it is chosen as the ground truth for comparison. From
Figure 4, we can see that OASIS node 1 and VALT detect seismic events at the same time. Figure 5 compares the number of events detected during an active period from 07/21/2009 to 07/26/2009. OASIS node 1 triggers 140 seismic events while VALT detects 160. The VALT station is more sensitive, because it has a very low frequency response up to 0.01 Hz, while the corner frequency of the OASIS station is 2 Hz. Therefore, VALT can detect more subtle events than the OASIS station, but it is also more expensive as said earlier.

When our STA/LTA algorithm identifies seismic events, it assigns the highest priority 7 to the seismic signals, as shown in Figure 6 (Top). The default STA/LTA ratio 2 works well in the lab test with real seismic data as input. However, post-deployment analysis shows that 2 is not always the optimal STA/LTA ratio. Due to the high sensitivity and low noise level of the geophone sensor, many small activities are also recognized as earthquake events. If we choose 3 as the STA/LTA threshold, those small volcanic activities will not be considered as events, as illustrated in Figure 6 (Bottom). We reconfigured the parameter remotely via RPC commands avoid false event triggering and result in better bandwidth allocation. The right monitoring parameter value highly depends on the status of the volcano. It may need to be tuned in long-term monitoring as the volcanoes status changes. This exemplifies the importance of being remotely configurable for systems like ours.

### 6.3 System Failures and Diagnosis

In this section, we present our evaluation and diagnosis on several problems during deployment. The uptime evaluation is fully end-to-end, that is to say, a node is considered to be up only if its data is successfully logged in the database, no matter where the failure is in-between. Figure 9 shows the status of each node during the period from 07/15/2009 to 09/01/2009. The uptime of the nodes varies from 34% to 93.6% with different types of failures. From 08/16, the UPS for the Internet router at the control center was down for two days, and caused failure of data importing. On 08/23/2009, due to an exception in the data importer tool, the network branch 2 was offline for 1 day. Node 1 has the shortest uptime due to the Deluge failure. We used Deluge to reprogram the network with a new version of node software after fixing a bug. Unfortunately, node 1 failed to reboot from the designated image slot.

Radio related problems is another challenge to the system robustness. We describe how the problem was exposed and solved as follows. From 08/25/2009 to 09/04/2009, 9 nodes surprisingly disappeared one after another. However, the RSSI/LQI report in our VALVE database shows that the sink nodes can normally receive the beacon packets from those nodes. We suspected that the radio amplifier of those nodes may always stay in TX mode, and thus could not receive beacon packets to form a valid path to the sink. To verify our conjecture, we remotely reprogrammed the sink node 0 with the
TOSBaseLQI (based on TOSBase in TinyOS) program. The snooped beacon packets show that those nodes restarted every 30 seconds (our software watchdog resets the node once it fails to send out any data packets for 30 seconds), and further ensure us that the amplifier is the reason for the node failures. On 09/11/2009, we got a chance to visit several failed nodes by sharing a helicopter with another USGS mission. The field test shows that the power supply to the amplifier is only 0.33 V (normally it should be 3 V), and the current consumption of the amplifier is as high as 164 mA. That indicates that the amplifier is latched in TX mode. A few days later, when we scrutinized the data logged in the database, we observed a pattern that most nodes detected lightning strikes before they disappeared, as shown in Figure 8. Node 11 recorded at least 20 strikes in the last minute before it died. We confirmed that our losses were due to lightning strikes from the weather patterns on 8/25/2009, 8/27/2009, 9/1/2009, and 9/3/2009. The RF receiver on the TI chip CC2591 is sensitive to lightning, and got damaged by lightning strikes. Afterward we added a coaxial lightning arrestor to CC2591 to reflect the lightning energy without decreasing the RF SNR, and replaced the broken RF amplifiers.

![Node Status (%)](image)

**Fig. 9. The node status during the first 47 days.**

### 6.4 Link Quality and Network Connectivity

We conducted evaluation on link quality (RSSI/LQI) and packet loss. The end-to-end packet loss is measured as the ratio of the number of lost data bits over the total number of bits expected in each time unit. Each node in the network reports the RSSI (Received Signal Strength Indicator) and LQI (Link Quality Indicator) of beacon packets from its neighbors every 5 minutes.

Figure 10 plots the change of the RSSI and LQI of links $2 \rightarrow 0$ (node 2 is the sender and node 0 is the receiver) and $5 \rightarrow 0$ over a time period of 36 hours. We can see that the RSSI of node 5 is higher than that of node 2 by about 13 dBm. The LQI of node 5 is also constantly higher than that of node 2 due to the directional antenna on node 5 that significantly increases the signal strength and link quality. We also observed a diurnal drop in the signal strength and link quality. In Figure 10, the RSSI of both nodes are high during UTC time 06:00 to 15:00 (23:00 to 8:00 in Pacific time), and decreases by more than 10 dBm during UTC time 18:00 to 00:00. Figure 10 shows the diurnal change in LQI. The LQI of node 2 also experiences a diurnal drop, similar to the RSSI. The LQI of node 5 does not experience significant fluctuation over time, because the link quality (e.g., LQI) stays high when the RSSI is way above the threshold. The above analysis reveals that there exists certain diurnal environmental parameters that affect radio propagation. From the data logged in our database, this diurnal phenomenon repeats almost everyday. Similar phenomenon is also reported in [23]. From the above observations we conclude that it is necessary to provide a certain amount of redundancy in the link quality to survive the diurnal drop when deploying a network. The network connectivity can become intermittent in extreme conditions. For example, our node 12 occasionally lost its communication connection to the network when the link quality dropped (See Figure 14 in the supplement).

### 7 Related Works

Deployment lessons [18] have stacked up with the increasing quantity of reported deployments. One of the first deployments was conducted in 2002 at Great Duck Island to monitor the habitat of the Leach’s Storm Petrel [17]. Other habitat monitoring deployments include tracking the micro-climate of a redwood tree [19] and Cane Toad populations [3]. These applications typically have low-duty-cycle and low-data-rate characteristics. Like our application requirements, some recent deployments involved high-data-rate signals, such as monitor-
uring industrial processes [5], long-span bridge [10], and railway [4]. Those deployments shed light on a number of design issues of sensor networks.

However, long-term viability in harsh and remote environments remains a challenging issue. A recent deployment in the Swiss Alps [6] reported that the temperature response differences for the processor and radio oscillators caused a loss of clock synchronization under temperature variations, and resulted in communication failures. FireWxNet [8], a multi-tiered portable wireless system for monitoring wildland fire, is distinguished by the rugged mountainous terrain over which it was spread, similar to the volcanic environment of OASIS. The deployment life of FireWxNet was expected to be roughly 3 weeks, while OASIS is designed to operate unattended for at least one year. All of our nodes were only reachable by helicopter, and the field maintenance is extremely hard. Thus software dependability and reliability are our critical concerns. Harvard has done the pioneering work [21] in volcano monitoring and deployed a sensor network on an active Ecuador volcano in 2005 to monitor seismic activities. They used an event-detection algorithm to trigger on interesting volcanic activity and initiate reliable data transfer to the base station. During the 19-day sensor network deployment, the network recorded 229 earthquakes, eruptions, and other seismoacoustic events. However, they found their event detection accuracy was only 1%, which justifies the requirement of real-time continuous raw data delivery by USGS. Their network reliability and uptime was relatively low: the mean node uptime was only 69% after factoring out base station failures. This work reveals many hard lessons in volcano monitoring, and has greatly benefited our design of OASIS system. In our deployment, we have achieved better performance in those aspects and learned many new lessons.

8 CONCLUSION AND FUTURE WORK

Our successful design and deployment in Mount St. Helens demonstrated that a low-cost sensor network system can provide real-time continuous monitoring in harsh environments and greatly promoted the confident use of sensor networks. It is an achievement of the whole sensor network community. USGS has been planning to utilize this design for other volcano monitoring and geological survey missions.

Sustainability and reliability of sensor networks in extreme environments remain a major research challenge. In extreme situations, a predictable and stable path may never exist, the network connectivity is intermittent (as observed in this study), a node could suddenly appear or disappear, and the rare upload opportunity and unpredictable node disruptions often result in data loss. One of our future works is to design a collaborative communication and storage middleware that cooperatively configures resources to increase disruption resilience, data persistence and network lifetime, and capture the intermittent connectivity for data delivery.

REFERENCES


Renjie Huang is a PhD student in computer science from Washington State University. His research interest mainly focuses on communication scheduling and optimization in wireless sensor networks. He received his B.S. degree from South China University of Science & Technology and received his Master degree from Huazhong University of Science & Technology.

Wen-Zhan Song is an associate professor in computer science and director of Sensorweb Research Laboratory at Georgia State University. Dr. Song is an active researcher on sensor networks, smart grid and pervasive computing and has received more than $2 million funding support from NSF, NASA, USGS and Boeing during 2005 - 2010. He is a recipient of NSF CAREER award in 2010 and Chancellor Research Excellence Award in WSU Vancouver in 2010. He received PhD from Illinois Institute of Technology and MS/BS from Nanjing University of Science and Technology.

Mingsen Xu is a PhD student in computer science from Georgia State University. He currently is a research assistant of sensorweb research laboratory. His research interest focuses on data correlation and reduction on sensor networks. He received his MS and BS degree from Beijing University of Posts and Telecommunications.

Nina Peterson is currently an Assistant Professor of Computer Science in the Natural Sciences and Mathematics Division at Lewis-Clark State College. She received her Doctorate from Washington State University and worked as a Research Assistant in the Pervasive Computing Laboratory.

Richard LaHusen received the BS degree from the University of California, Davis, and completed 4 years of graduate work at Humboldt State University. Thereafter, he has worked for the last 20 years as part of the USGS Volcano Hazards Team and is currently a senior instrumentation engineer at the USGS Cascades Volcano Observatory. Throughout the 20-year period, he has developed instrumentation, telemetry systems, and software for the research and monitoring of volcanic processes and hazards (Murray et al., 1996). He developed the Acoustic Flow Monitor (AFM), an innovative system that incorporates in situ analysis of seismic signals at remote nodes for the real-time detection and warning of volcanic debris flows that is in use around the world (LaHusen, 1996). He also developed a low-powered, high-resolution earth deformation monitoring system that optimizes allocation of limited resources for GPS data acquisition and communication needs (LaHusen and Reid, 2000). Most recently, he introduced and applied aerial deployment of self-contained volcano monitoring instrumentation stations during the 2004-2005 eruption of Mount St. Helens (LaHusen et al., 2006).

Behrooz A. Shirazi is currently the Huie-Rogers chair professor and director of the School of Electrical Engineering and Computer Science. His research interests include the areas of pervasive computing, software tools, distributed real-time systems, scheduling and load balancing, and parallel and distributed systems. He has received grant support totaling more than $8 million from federal agencies, including the US National Science Foundation (NSF), the US Defense Advanced Research Projects Agency (DARPA), and the AFOSR, and private sources, including Texas Instruments and Mercury Computer Systems. He has received numerous teaching and research awards. He is currently the Editor-in-Chief for the Special Issues for Pervasive and Mobile Computing Journal and has served on the editorial boards of the IEEE Transactions on Computers and the Journal of Parallel and Distributed Computing.
9 Supplemental System Design Details

In this section, we present more system design details as a supplement of section 4 and section 5 in the main paper.

9.1 Over-the-air Network Reprogramming (continued)

This is a supplement of section 4.4 in the main paper. Deluge works on Tmote Sky and MicaZ, but it does not support the iMote2 platform, so we ported Deluge to the iMote2 platform. The challenges of porting Deluge to iMote2 result from the need of a sophisticated mechanism to access the Flash due to the increased size of both the RAM (32 MB) and the flash (32 MB), which use a Linux file system. In addition, the bootloader of iMote2 is much more sophisticated than that of Tmote Sky. Here we summarize the major modifications made in the implementation of Deluge for iMote2.

iMote2 does not have sectors that can be used for allocating space for the three different Deluge images. So we created three files in the user area for storing the code updates. The bootloader is modified so that it can read the new code from the location specified in the user area and load the code to the bootable location. This was achieved by adding additional attributes to the shared attribute table including the location and size of the image. Whenever a node receives a reboot command, it updates these attributes in the shared attribute table and reboots the device. Prior to each reboot, the bootloader first checks the shared attributes to see if these attributes are enabled. If enabled, it then loads the image from the designated location; if no location is recorded, it loads the image from the pre-defined primary and secondary locations.

9.2 Event Detection and Data Prioritization

The seismic event data is critical for volcano studies, and domain scientists require reliable delivery of the data with the highest priority. We used the STA/LTA (short-term average over long-term average) algorithm [4], [5] to detect seismic events. The prototype of this algorithm is presented in [1] and we describe it here briefly. The STA/LTA algorithm is based on RSAM (Realtime Seismic Amplitude Measurement), which is calculated to detect seismic events. The prototype of this algorithm is presented in [1] and we describe it here briefly.

The seismic event data is critical for volcano studies, and domain scientists require reliable delivery of the data with the highest priority. We used the STA/LTA (short-term average over long-term average) algorithm [4], [5] to detect seismic events. The prototype of this algorithm is presented in [1] and we describe it here briefly. The STA/LTA algorithm is based on RSAM (Realtime Seismic Amplitude Measurement), which is calculated

\[
X_i = \frac{1}{\sum_{j=0}^{n-1} x_i - x_j}, \quad \text{where } x_i \text{ is } i\text{-th-second RSAM; } n \text{ is the STA or LTA time window size. LTA gives the long term background signal level while the STA responds to short term signal variation. In our implementation, the STA window is 8 seconds; the LTA window is 30 seconds. The ratio of STA over LTA is constantly monitored. Once the ratio exceeds the trigger threshold (by default 2), the start of an seismic event is declared, and the LTA value is frozen so that the reference level is not affected by the incoming signals. The end of the event is declared when the STA/LTA ratio is below the de-trigger threshold. The event data is then assigned the highest priority 7 for reliable delivery. Seismologists expect the onset of the event to be reliably collected for analysis purpose; the data at the margin of the event-triggered window should be included in the event period. We used a pre-event buffer to retain the data for a period of \( T_{pre} = 4 \) seconds by default. Once a seismic event is detected, the pre-event buffer data is also assigned the highest priority. The event data will be reliably delivered to the gateway by the reliable data transfer protocols in section 9.3. The default parameter values of the STA/LTA algorithm are suggested by USGS scientists and are configurable via RPC mechanism.

9.3 Reliable Event Data Collection (continued)

A key challenge of reliable data collection in sensor network is that the feedback control traffic directly competes the bandwidth of sensor data traffic due to the radio broadcast nature. When the bandwidth is severely limited, the data delivery with RDT could be worse than that without RDT. This has been overlooked in the literature and we made the following innovative design choices.

Firstly, we used bitmap to represent multiple ACK/NAKs, namely ANK, to reduce feedback traffic. An ANK packet has 3 fields (startSeqNo, validBits, ankBitmap), where startSeqNo is the starting sequence number to ANK, and validBits indicates number of valid bits in ankBitmap (e.g., number of packets to ANK). In ankBitmap, bit-1 denotes a received packet while bit-0 denotes a lost packet. For example, an ANK packet (startSeqNo = 10, validBits = 4, ankBitmap = 1001···) means that packet 10 and 13 are received, while packet 11 and 12 are lost. In this way, a single ANK packet can ACK/NAK many packets. We also applied a mechanism to regulate ANK traffic with a controlled interval. This mechanism was not considered in our original design. After the deployment, we observed that the data stream in the database was intermittent with gaps of several minutes. Later we found that the reason was because sometimes the receiver sent ANK packets too frequently and caused buffer overflow in the MOXA device server. After we prolonged the ANK interval, the system worked normally. Secondly, we removed

\[
\sum_{i=1}^{n-1} s_i \text{ is the average seismic sample value in the } (i-1)\text{th and } i\text{th second respectively, then } c_{i-1} = \frac{\sum_{i=1}^{n-1} s_i}{m} \text{ is the average seismic sample value in the } (i-1)\text{th second. The } i\text{th-second RSAM } x_i \text{ is calculated with the equation: } x_i = \frac{\sum_{k=i-m}^{i-1} x_i - c_{i-1}}{m}. \]

The STA or
sender’s retransmission timer. In the previous example that ANK can not reach the sender, a sender with retransmission timer would retransmit useless old packets wasting bandwidth. In addition, the timeout interval would be very hard to estimate in lossy wireless networks, comparing to wired Internet. If a sender receives an ANK, the buffer space of those ACKed packets will be freed for reuse. If a sender’s buffer is full, a new packet will simply replace the oldest one, even if it has not been acknowledged. Because if this situation happens, it means either no packet loss as no ANK comes back, or bandwidth is severely limited - we would rather drop old packets than new packets. Those designs are important to ensure that using RDT will be at least as good as the case without RDT. When the receiver stops sending ANK, the system will be same as without RDT.

9.4 Adaptive Linear Filtering Compression

To reduce the bandwidth demands and maximize the data return over the unreliable and low rate radio links, we designed an Adaptive Linear Filtering Compression (ALFC) [3] algorithm to compress the seismic raw data. It is a lightweight compression algorithm tailored for sensor networks with code size only 768 bytes. Considering the relatively modest computational power of existing sensor platforms, ALFC does not use floating-point operations and has very low computation and energy cost.

Our method relies on adaptive prediction, which eliminates the need to determine a priori of prediction coefficients and, more importantly, allows the compressor to dynamically adjust to a changing source. This is particularly important for seismic data because the source behavior can vary so dramatically depending on seismic activities. Predicted sample values are used to losslessly encode source samples using a variable length coding scheme. We map each sample value to a non-negative integer and then encode the resulting sequence using Golomb codes. This general strategy is used in the Rice entropy coding algorithm and the LOCO-I image compressor, among a myriad of other applications. Given the low power characteristics of wireless sensor networks, wireless links are typically lossy, demanding a compression scheme that allows packets to be decompressed even when preceding packets have been lost. We alter the prediction approach for the first few samples in the packet so that it does not rely on sample values in preceding packets. Predicted sample values are used to losslessly encode source samples using a variable-length coding scheme.

10 Field Deployment Experience

The deployment of the sensor network of 13 OASIS stations was conducted in Mount St. Helens volcano in July 2009. The OASIS stations were lowered by cable from a helicopter hovering about 100 feet up and gently put in hot spots inside the crater and around the flank.

Fig. 11. The spatial distribution of the triggered events from 07/15/2009 to 07/31/2009.

We monitored the network connectivity in real time at JRO by connecting a laptop to the sink node and gave feedback via satellite phones to the crew in the helicopter to ensure network connectivity. The real-time feedback was very useful in a field deployment. The 13-node deployment took us about 6 hours. Installing stations on the flank turned out to be more difficult than inside the crater due to the long distance between stations (up to 4 miles) and the rugged terrain. The diameter of the covered area is about 6 miles. The customized amplifier worked remarkable well. Node 6 has a reliable link to the sink node 0 at JRO with a distance of approximately 4.6 miles.

The deployed in-situ sensor network has two branches. Each branch operates with a separate data collection sink and radio channel. The first branch network (nodes 1−6) is mostly placed inside the crater. The second branch network (nodes 8−14) is deployed around the flank forming a semicircle. Some OASIS stations co-locate with existing USGS stations including VALT, SEP, and NED, which serve as ground truth to evaluate the data quality of OASIS stations. The two sink nodes 0 and 7 are installed at JRO so they are easy to access. This makes the network failures caused by sink easier to fix. Also, the sink now has a reliable source of power. During the trial deployment in 2008, the sink node was placed in the crater with battery and equipped with a 900 MHz radio for telemetry to the gateway, which was power hungry and depleted the sink node in 4 months.

11 Supplemental System Evaluation and Findings

In this section, we present more system evaluations and findings as a supplement of section 6 in the main paper.

11.1 Data Quality Evaluation (continued)

This is a supplement of Section 6.1 in the main paper. During our trial deployment in 2008, we found that some seismic data samples lost their 6 LSBs (Least Significant Bit) and have distortions, as shown in Figure 12 (Top). Eventually, we figured out that the 6.5 MHz clock rate of iMote2 driving the ADC driver was the cause of the data distortion. The ADC chip ADS8344 [2] can only work normally at 2.4 MHz clock mode. In other words, the external clock cycle should be no less than
400 ns (2.4 MHz) to correctly accomplish the conversion. Originally the SPI clock rate was configured to be 6.5 MHz (typical iMote2 clock rate), which is too high for ADS8344. Thus we changed the SPI Serial clock rate to 2.6 MHz by configuring a higher clock divisor. This SPI clock rate was still slightly higher than the specification, but ADS8344 worked normally and the data distortion was eliminated (See Figure 12 (Bottom)).

11.3 Data Prioritization and Compression Effects

Next we evaluated the performance of our data prioritization scheme Tiny-DWFQ and data compression algorithm ALFC.

Figure 15 shows the end-to-end packet reception ratios with different data priorities based on node 4’s 24-hour data stream on 08/02. We can see that data with a higher priority accordingly has a higher chance to reach the gateway.

11.2 Event Detection Accuracy (continued)

This is a supplement of Section 6.2 in the main paper. We also investigated the spatial distribution of the triggered events. Figure 11 shows the number of seismic events triggered by each OASIS station from 07/15/2009 to 07/31/2009. We can see that the network branch (nodes 1 – 4, and 6) inside the crater detects more events than the branch (nodes 8 – 14) along the volcano flank, with a lead of 138%. That sheds light on the volcanic hot spot distribution and benefits the refinement of our future deployment strategy. Additionally, OASIS node 1 detected the most events due to its advantages in the seismic sensor. Some small events can only be detected by highly sensitive sensors.

Not all seismic events indicate earthquakes. For example, a small rock fall can generate small seismic events but not earthquake. Thus, we used USGS analysis software to pick out real earthquake events. Figure 13 (Top) shows the location and magnitude of the 187 earthquakes that happened in the first 6-month deployment period from 07/18/2009 - 01/13/2010. The size of the circles denotes the earthquake magnitude. From Figure 13 (Top) we can see that while the crater is most active area, some strong earthquakes also took place on the flanks. Figure 13 (Bottom) shows the earthquake time. We observed that earthquakes are more frequent during the first 3 month.

A unique innovation of OASIS is feeding back information from EO-1 into the in-situ element. High spatial resolution data generated by EO-1’s Hyperion spectrometer is fed through a thermal analysis element to detect a region of thermal activity on the target area, analyzes the data, and pushes results to the ground segment to re-prioritize bandwidth allocation through Command & Control. Figure 16 shows the space-to-ground triggering and the prioritized data delivery mechanism in the ground network. For example, as snow accumulates in the Mount St. Helens, OASIS node 4 was buried under snow gradually. During that process, the data stream of node 4 experienced more and more packet loss. However, once the data priority was raised to highest...
Fig. 13. (Top) The spatial distribution of earthquake events. (Bottom) The temporal distribution of earthquake events.

Fig. 14. The intermittent data delivery of OASIS node 12 between 11/09/2009 and 11/10/2009.

level due to space triggering, the data during that period were reliably delivered.

The ALFC algorithm in each OASIS station losslessly compresses the real-time seismic data. The decompression is performed at the gateway before the data is stored into the database.

Figure 17 illustrates the average compression ratio of the seismic data on each OASIS station based on a 4-hour continuous data stream. Node 0 and node 7 serve as two sink nodes without connection to seismic sensors, and node 5 is not connected to the network due to the problem of asymmetric links. So these 3 nodes are not included in Figure 17. An observable margin in the compression ratio can be found in node 1. It has an average compression ratio of about 2.2, which is about 30% better than other nodes. Node 1’s Geophone-based seismometer has lower noise level than the MEMS-based seismometer, thus its data is more compressible. Additionally, the standard error on node 1 indicates that its compression ratio experiences a larger fluctuation than other nodes. That is because node 1’s data contains more seismic event signals. According to the data logged in our database, node 1 experienced about 20 seismic
Fig. 16. Space-to-ground triggering raised the data priority to the highest level, resulting in reliable delivery of the data stream from node 4.

events in this 4-hour period, while other nodes detected fewer, even none.

Fig. 17. Averaged compression ratio of each node based on a 4-hour seismic data stream

Fig. 18. The number of bits per data sample after compression. Each data sample has 16 bits before being compressed. The lower number of bits per sample after compression denotes better compression performance. Figure 18 plots node 1 and node 6, the two nodes with the most seismic events. The bit number per sample is calculated every 280 samples to reflect the variation over time series. There are two facts observed from Figure 18. Firstly, the average number of bits per sample of node 1 is noticeably lower than node 6. With ALFC, it costs node 1 about 7 bits to encode each sample on average, while for node 6 more than 9 bits are required to encode each sample. That is because the data from node 1 have lower noise level than that from node 6 due to the equipped geophone seismic sensor as explained above. Secondly, the spikes in the curves for each node have a strong correlation with the seismic events. Figure 4 plots the seismic waveform data for node 6 and node 1. We can see that each seismic event has a corresponding spike in the curves in Figure 18. The seismic event data have high dynamics and thus result in low compression ratio. Moreover, Figure 18 illustrates that the compression ratio of node 1 has larger fluctuation than that of node 6, which is indicated by the number and the amplitude of spikes on each curve. That exactly explains why the compression ratio of node 1 has a larger variation as shown in Figure 17. The results above show that ALFC performs effectively in a field network deployed on an active volcano where the seismic data generally have high dynamics. They also reveal that the noise floor of sensors has a significant impact on the compression performance of ALFC.

REFERENCES


