Volcanic Earthquake Timing in Wireless Sensor Networks

Guojin Liu$^1$ Rui Tan$^1$ Ruogu Zhou$^1$ Guoliang Xing$^1$ Wen-Zhan Song$^2$ Jonathan M. Lees$^3$

$^1$Dept. Computer Science & Engineering, Michigan State University $^2$Dept. Computer Science, Georgia State University $^3$Dept. Geological Sciences, University of North Carolina at Chapel Hill

Abstract—Recent years have witnessed pilot deployments of inexpensive wireless sensor networks (WSNs) for active volcano monitoring. This paper studies the problem of picking arrival times of primary waves (i.e., P-phasers) received by seismic sensors, one of the most critical tasks in volcano monitoring. Two fundamental challenges must be addressed. First, it is virtually impossible to download the real-time high-frequency seismic data to a central station for P-phase picking due to limited wireless network bandwidth. Second, accurate P-phase picking is inherently computation-intensive, and is thus prohibitive for many low-power sensor platforms. To address these challenges, we propose a new P-phase picking approach for hierarchical volcano monitoring WSNs where a large number of inexpensive sensors are used to collect fine-grained, real-time seismic signals while a small number of powerful coordinator nodes process collected data and pick accurate P-phases. We develop a suite of new in-network signal processing algorithms for accurate P-phase picking, including lightweight signal pre-processing at sensors, sensor selection at coordinators as well as signal compression and reconstruction algorithms. Tested experiments and extensive simulations based on real data collected from a volcano show that our approach achieves accurate P-phase picking while only 16% of the sensor data are transmitted.

I. INTRODUCTION

Volcanic eruptions have become a major hazard due to ever growing human population and urbanization around volcanoes. It is estimated that about 500 million people today live close to active volcanoes [1]. Existing volcano monitoring systems often employ broadband seismometers that collect high-fidelity seismic signals, but are expensive, bulky, and difficult to install. As a result, many of the most threatening volcanoes are monitored by fewer than 20 stations. Such poor spatial granularity limits scientists’ ability to study the volcano dynamics and predict eruptions.

Recent years have witnessed pilot deployments of inexpensive wireless sensor networks (WSNs) for active volcano monitoring [2]–[4]. These deployments demonstrated the potential of long-term, large-area, and fine-grained volcano coverage by deploying large numbers of low-cost sensors. Significant research has been focused on improving system robustness, time synchronization, network efficiency, and communication performance issues. In previous small-scale deployments [2]–[4], detection and analysis of volcano activity were accomplished by transmitting raw data to a base station for centralized processing. However, as the sensor signals are sampled at high frequencies (e.g., 50 to 200 Hz), it is virtually impossible to continually collect raw, real-time data from a large-scale and dense WSN. This is due primarily to severe limitations of energy and bandwidth of current WSN platforms.

The goal of this paper is to design algorithms that can accurately determine the arrival times of primary waves (i.e., P-waves) received by seismic sensors inside the network, without transmitting raw measurements to the base station for centralized processing. Earthquake signal timing is a fundamental task in seismology. P-wave arrival times (i.e., P-phases) are essential information for earthquake source localization and seismic tomography [5], which enable earth scientists to understand the physical processes inside the volcano conduit systems and implement early warning mechanisms. In volcano observatories, P-phase picking is often done by visual inspection of experienced seismologists. When the volume and rate of data capture is large, however, this process is extremely labor-intensive, time-consuming, and subject to inconsistency across different examiners. In the last two decades, automated P-phase picking algorithms have been developed in seismology community for earthquake timing [6]–[8]. However, these algorithms are designed for powerful nodes with substantial computation, storage and power resources. It remains an open question if it is possible to implement automated in-situ volcanic earthquake timing in resource-constrained WSNs without transmitting a large volume of raw sensor data.

The key contribution of this paper is the development of new in-network signal processing algorithms for P-phase picking. To balance the system lifetime and network coverage, we adopt a hierarchical network architecture that consists of low-end nodes (referred to as sensors) and high-end nodes (referred to as coordinators). A large quantity of inexpensive, mote-class sensors can provide fine-grained monitoring with long lifetime, while a small number of coordinators (e.g., Imote2 and embedded PCs like Gumstix [9]) enable advanced in-network seismological signal processing. Based on this network architecture, we develop a suite of in-network P-phase picking algorithms. (1) Lightweight algorithms are designed for sensors to coarsely pick the P-phases and estimate the signal sparsity. The coarse P-phase is an important hint of the amount of new information that the sensor can contribute. The signal sparsity determines the volume of data transmission if the sensor sends its signal to a coordinator for accurate P-phase picking. (2) A sensor selection algorithm uses signal sparsities and coarse P-phases to choose a subset of the most informative sensors to transmit compressed data subject to a given upper bound on communication overhead. The

* The first two authors are listed in alphabetic order. 1st version: July 2012. 2nd version: February 2013.
bound can be set by the network designer to meet various practical system constraints, such as bandwidth limitation, energy budget, and real-time requirement. (3) A signal compression algorithm for sensors and a reconstruction algorithm for coordinators are developed based on wavelet transform and compressive sampling theory [10]. The above algorithms work collaboratively to achieve energy-efficient and accurate in-network earthquake timing. The approach presented in this paper can be extended and then applied in various monitoring applications that need accurate signal arrival times. Moreover, it has important implications to a broader class of applications that need to accurately extract features from real-time, high-frequency signals gathered by resource-constrained WSNs.

We implement and evaluate the proposed algorithms on a testbed of TelosB motes that are loaded with real seismic data collected on Mount St. Helens. The results demonstrate the feasibility of deploying our algorithms on real volcano monitoring WSNs. We also conduct extensive simulations based on real data traces that contain 30 earthquakes. The results show that our algorithms can achieve accurate earthquake timing while only 16% of the sensor data are transmitted.

The rest of this paper is organized as follows. Section II reviews related work. Section III states the problem and approach overview. Section IV studies the sparsity of earthquake signal and presents the signal pre-processing algorithms at sensors. Section V formulates the sensor selection problem. Section VI discusses the compression/reconstruction algorithms. Section VIII presents the evaluation results. Section IX concludes this paper and discusses the future work.

II. RELATED WORK

The first attempt of using WSN for monitoring active volcano was in 2004 [2], where four MICA2 nodes [11] were deployed on Volcán Tungurahua, Ecuador. The system successfully collected the acoustic data of at least nine large explosions in three days. In 2005, the same research group deployed sixteen Tmote nodes [12] equipped with seismic and acoustic sensors on Volcán Reventador, Ecuador, for three weeks [3]. In 2007, this research group deployed eight Tmote nodes [12] on Volcán Tungurahua again and applied the Lance framework presented in [13]. Lance selects a subset of sensors such that the total value of the raw data collected from the selected sensors is maximized subject to the network lifetime constraint. In the project named Optimized Autonomous Space In-situ Sensorweb (OASIS) [4], fifteen Imote2 nodes [11] were deployed on Mount St. Helens in 2008. The project demonstrated a long-term sustainable WSN in challenging environment, and delivered a long-period (up to half a year) valuable high-fidelity sensor dataset. The design of the above volcano monitoring WSNs [2]–[4], [13] was mainly focused on the basic network services such as node sustainability, network connectivity, time synchronization and data collection. Our previous work [14] proposed a real-time and in-situ volcanic earthquake detection approach based on in-network signal processing. The testbed experiment based on TelosB motes [11] demonstrated the feasibility and advantages of in-network signal processing in volcano monitoring applications. Different from [14] that aims to detect the occurrence of earthquake, the objective of this work is to accurately pick the P-phases from the seismic signals received by sensors when an earthquake happens. To our best knowledge, this paper is the first to study volcanic earthquake timing in WSNs.

A few algorithms have been presented in seismology for picking the P-phase. These algorithms are typically based on the identification of the changes in signal characteristics such as energy, frequency and autoregressive model [7]. Most of them [6]–[8] were designed for workstation computers to centrally process the data collected from traditional seismological stations. As a result, they are ill-suited for resource-constrained wireless sensors. Among them, the STA/LTA-based approaches [6] are widely adopted. The key idea is to continuously compute the ratio of short-term average (STA) to long-term average (LTA) over a signal characteristic (e.g., the energy in the interested frequency band). A detection is raised once the ratio exceeds a threshold. Although the STA/LTA-based approaches are suitable for real-time implementation, they may not yield satisfactory performance. Moreover, these heuristic algorithms often require extensive parameters that need to be set empirically, making them difficult to adapt to different volcano environments. Another important category of picking algorithms is based on the autoregressive (AR) model [8]. Based on the assumption that the seismic signals before and after the P-phase follow different AR models, AR-based approaches pick the time instance to maximize the dissimilarity of the two AR models. AR-based approaches need few user settings and typically yield better performance than the STA/LTA-based approaches. However, as two AR models need to be trained for each time instance, they often incur high computational complexity.

III. PROBLEM DEFINITION AND APPROACH OVERVIEW

A. Problem Definition

P-phase is the arrival time of the seismic primary wave. Fig. 1 shows the seismic signals received by three sensors deployed on Mount St. Helens in the OASIS project [4] and the manually picked P-phases. We can see from the figure that sensors receive different P-phases. Such differences provide critical information for many volcano monitoring applications such as earthquake source localization and seismic tomography [5]. The task of picking the P-phases of the spatially distributed sensors is referred to as volcanic earthquake timing.
Because of the fast propagation speed of primary wave, the differences between sensors’ P-phases are often small, e.g., at most one second in Fig. 1. This imposes stringent accuracy requirements on volcanic earthquake timing. In this paper, we aim to develop a holistic and energy-efficient approach to accurate volcanic earthquake timing in WSNs. Our approach is designed to meet the following two key objectives. First, the picked P-phases achieve satisfactory precision and maximize the total information amount for the volcano monitoring applications that take P-phases as inputs, such as earthquake source localization and seismic tomography. Second, to extend the network lifetime and achieve long-term monitoring, the computation overhead of earthquake timing should be fairly distributed among sensors and the volume of seismic data transmission in each earthquake timing procedure must meet the specified energy budget.

We make the following assumptions:

**Hierarchical architecture:** The network is composed of two types of sensors distributed on the volcano, which include low-end sensors with limited resources (e.g., TelosB) and high-end sensors with more processing capability (e.g., Imote2). As discussed in Section II, most P-phase picking algorithms are not affordable for low-end sensors. However, due to low costs, a large number of low-end sensors can be deployed over the volcano to provide a high level of coverage. A small number of high-end sensors can accomplish a few critical computation and processing tasks that are not affordable to low-end sensors. According to our extensive numerical study in Appendix A, such a hierarchical architecture can reduce the per-node energy consumption by at least 50%, compared with the non-hierarchical network composed of only high-end sensors. This is mainly because the high-end sensor typically consumes much higher power than low-end sensor in standby state when there is no earthquake for most of the time. Moreover, such a hierarchical architecture is consistent with the architecture of many practical large-scale WSN systems [15].

**Clustering:** The network is organized into one or multiple clusters. Each cluster is composed of a number of low-end sensors and a high-end sensor as the cluster head. As our approach can be applied under any cluster formation scheme, we do not assume a specific cluster formation scheme. We refer to Appendix C for extensive discussion of two practical clustering schemes and the setting of cluster size for volcano monitoring application. The rest of this paper is focused on one cluster.

**Earthquake onset time:** The network can detect the occurrence of volcanic earthquake and the earthquake onset time. The earthquake onset time is a coarsely estimated time instance, typically in second precision, at which the earthquake process starts. Many earthquake detection algorithms such as the widely adopted STA/LTA approach [3], [6] and a recent Bayesian approach [14] can detect the earthquake onset time. By the earthquake onset time, sensors can largely narrow the range of searching for the P-phases.

---

**Fig. 2.** Algorithm framework of energy-efficient volcanic earthquake timing. Solid line represents data flow; dotted line represents control flow; white blocks are the components at low-end sensors; shadowed blocks are the components at the cluster head.

**B. Approach Overview**

Because most P-phase picking algorithms are not affordable for low-end sensors as discussed in Section II, under the hierarchical network architecture, an important design option of our approach is to shift the P-phase picking to the cluster head. We propose a collection of algorithms running at sensors and the cluster head, which work together to achieve the requirements discussed in Section III-A. The algorithm framework is shown in Fig. 2. When an earthquake is detected, each sensor chooses a segment of seismic signal around the detected earthquake onset time and applies wavelet transform to the signal. The wavelet transform can make the signal more sparse, which can be exploited to reduce the volume of data transmission. Based on the transformed signal, each sensor estimates the signal sparsity and executes a lightweight P-phase picking algorithm to find a preliminary P-phase. Each sensor then sends the estimated signal sparsity and the preliminary P-phase to the cluster head. If the raw seismic signal at a sensor is requested by the cluster head for further fine-grained P-phase picking, the sensor employs compressive sampling (CS) to compress the seismic signal to reduce data volume before transmission. According to CS theory, the length of the compressed signal depends on the signal sparsity. Therefore, the cluster head can predict the volume of data transmission from the signal sparsity sent by sensors. The cluster head selects a subset of sensors such that the expected error of earthquake source localization, which is computed from the preliminary P-phases, is minimized subject to a bounded volume of data transmission. The selected sensors then compress the raw seismic signal and transmit to the cluster head. Finally, the cluster head reconstructs the seismic signals and executes high-accuracy P-phase picking algorithms and possibly other advanced seismic signal analyses. In this work, the cluster head adopts the autoregressive Akaike Information Criterion (AR-AIC) picking algorithm [8].

Our approach has the following three advantages. First, by the sensor selection, the earthquake timing process has bounded volume of data transmission. The system designer can set this bound to meet the real-time performance requirement and energy budget of the network. Second, our approach exploits CS to reduce the computation and communication overhead of sensors. Different from many general compression algorithms, CS has fixed and low complexity for low-end sensors and shifts most computation overhead to the signal...
reconstruction at the cluster head. Moreover, CS can determine the
volume of compressed signal before compression, which
enables the cluster head to select sensors and schedule the
data transmissions before sensors compress signals. As a
result, the unselected sensors can avoid expensive compression
computation. Finally, in addition to earthquake timing, the
cluster head can run various centralized seismic signal analysis
algorithms on the reconstructed signals, such as Fourier and
polarization analyses. The cluster head can also transmit the
picked P-phases and the reconstructed signals to the base
station via reliable communication links (e.g., long-distance
wireless link or wires) for joint source localization across all
clusters and visual inspection to the signals.

IV. SEISMIC PRE-PROCESSING AT SENSORS

In this section, we first explore the sparsity of seismic
signals received by sensors when an earthquake happens. The
estimated sparsity enables the exact prediction of data trans-
mission volume if the cluster head requests the compressed
signal at the sensor. We then present a lightweight preliminary
P-phase picking algorithm that is executed at sensors. The
estimated sparsity and the preliminary picks are sent to the
cluster head for selecting a subset of sensors, such that the
expected earthquake source localization error is minimized.

A. Sparsity of Volcanic Seismic Signal

In this paper, we adopt the common definition of sparsity in
signal processing [10]. Let $n$ denote the length of signal. Sup-
pose $\Psi$ is an orthonormal basis $\Psi = [\psi_1 \psi_2 \cdot \cdot \cdot \psi_n] \in \mathbb{R}^{n \times n}$
where $\psi_i$ is the $i$th column of $\Psi$, and a signal $s \in \mathbb{R}^{n \times 1}$ in
time domain is expanded in the basis $\Psi$ as $s = \Psi x$, where $x \in \mathbb{R}^{n \times 1}$ is the coefficient sequence of $s$. The signal $s$
is $k$-sparse if the number of non-zeros in $x$ is less than or
equal to $k$. The sparsity of signal $s$, denoted by $\rho$, is defined as
$\rho = k/n$. In practice, $x$ typically contains small values
rather than zeros. Considering $x(\cdot)$ obtained by keeping only
the $k$ largest coefficients of $x$ and setting others to zero, the
corresponding signal $s(\cdot)$ is $s(\cdot) = \Psi x(\cdot)$. The signal $s$
is $k$-sparse if the relative error $\| x - s(\cdot) \|_{\ell_2}$ is smaller than
a threshold, where $\| \cdot \|_{\ell_2}$ represents the $\ell_2$-norm. For the
evaluation presented in this paper, the threshold is set to be
5% unless otherwise specified.

For each sensor, we choose a signal segment for 16 seconds,
where 10 seconds before and 6 seconds after the detected
earthquake onset time. Hence, $n = 16 \cdot f_s$, where $f_s$ represents
the seismic sampling rate. Such a setting of signal length is
the minimum requirement of the AR-AIC picker [8] running
at the cluster head. As the difference between the P-phases
received by sensors is typically shorter than two seconds [3].
domain is often not sparse. For instance, for the signal shown in Fig. 4(d), the sparsity is 0.57. By applying appropriate transforms, the sparsity can be significantly reduced, which implies less wireless data transmission. In this paper, we adopt discrete wavelet transform (DWT) to reduce the signal sparsity. As DWT preserves the time-domain characteristics of the original signal, it is preferable for P-phase analysis. Moreover, the downsampling scheme of DWT allows us to develop an efficient preliminary P-phase picking algorithm in Section IV-B. The second columns of Figs. 3 and 4 show the 4-level DWT coefficients of Node01 and Node10 in two earthquakes. The vertical dotted lines represent the edges between two adjacent frequency subbands in the wavelet domain. The setting of the level of DWT will be discussed in Section IV-B. Our analysis shows that the sparsity in wavelet domain is significantly lower than that in time domain. For instance, for the four data traces shown in Figs. 3 and 4, the sparsity can be reduced by up to 75% in the wavelet domain.

1) Sparsity in Wavelet Domain: The seismic signal in time domain is often not sparse. For instance, for the signal shown in Fig. 4(d), the sparsity is 0.57. By applying appropriate transforms, the sparsity can be significantly reduced, which implies less wireless data transmission. In this paper, we adopt discrete wavelet transform (DWT) to reduce the signal sparsity. As DWT preserves the time-domain characteristics of the original signal, it is preferable for P-phase analysis. Moreover, the downsampling scheme of DWT allows us to develop an efficient preliminary P-phase picking algorithm in Section IV-B. The second columns of Figs. 3 and 4 show the 4-level DWT coefficients of Node01 and Node10 in two earthquakes. The vertical dotted lines represent the edges between two adjacent frequency subbands in the wavelet domain. The setting of the level of DWT will be discussed in Section IV-B. Our analysis shows that the sparsity in wavelet domain is significantly lower than that in time domain. For instance, for the four data traces shown in Figs. 3 and 4, the sparsity can be reduced by up to 75% in the wavelet domain.

2) Diverse Sparsity: We have the following important observations from the case study shown in Figs. 3 and 4. First, in the same earthquake, sensors receive different signal-to-noise ratios (SNRs) and different significances of P-phase. For instance, in Earthquake01 shown in Fig. 3, Node01 receives higher SNR and more significant P-phase than Node10. As seismic signal attenuates with the propagation distance, the sensors far away from the earthquake source receive weak signals and hence have lower SNRs and less significant P-phases. Second, due to the significantly variable magnitude and source location, the SNR and significance of P-phase are dynamic and unpredictable. For instance, different from Earthquake01, in Earthquake02 shown in Fig. 4, Node10 receives much higher SNR than Node01. Third, the sparsity depends on SNR and the position of P-phase. For instance, as Node01 receives higher SNR than Node10 in Earthquake01, the sparsity of Node01 is lower than Node10. However, although Node10 receives very high SNR in Earthquake02, its sparsity is just comparable to that of Node01. This is because Node10 receives P-phase much earlier than Node01, which results in more non-zeros in wavelet domain. We extensively evaluated the sparsity of transformed signals based on the data traces received by 12 nodes in 30 earthquakes in the OASIS project [4]. Fig. 5 plots the sparsity versus the threshold of the relative error \( \frac{\|s - s_{\text{thumbnail}}\|_2}{\|s\|_2} \) for determining sparsity. The result validates our hypothesis of diverse diversity. For instance, if the threshold is set to 5\%, the sparsity of 90\% data traces ranges from 0.16 to 0.63.

The above observations of dynamic, unpredictable and diverse sparsity provide important guidelines for designing WSNs for volcanic earthquake timing. First, due to the diversity of signal sparsity, it is desirable to only collect the most sparse seismic signals to meet the bandwidth limitation, real-time performance and limited node energy budget. Second, as the sparsity is dynamic and unpredictable, sensors need to calculate the sparsity when an earthquake is detected. The sparsity can be used to predict the volume of data transmission when the signal needs to be transmitted to the cluster head.

B. Preliminary P-Phase Picking at Sensors

In this section, we present a lightweight and efficient preliminary P-phase picking algorithm that runs at sensors. Due to the downsampling scheme, the lowest frequency subband in the wavelet domain is a zoomed-out version of the low-pass filtered signal. Hereafter, this subband is referred to as thumbnail of the original signal in time domain. The last columns of Figs. 3 and 4 show the thumbnails of the corresponding original signal in the first column. We can see that the thumbnails can preserve the shape of P-wave. If the seismic sampling rate is \( f_s \) and the level of DWT is \( l \), the lowest frequency subband of the wavelet domain is \([0 \text{ Hz}, \frac{f_s}{2^l} \text{ Hz}]\). By setting \( l \) such that \( \frac{f_s}{2^l} \geq 5 \text{ Hz} \), the thumbnail can preserve the shape of P-wave, which typically has a frequency lower than 5 Hz [14]. In our approach, the preliminary P-phase is picked from the thumbnail to reduce the computational complexity. For the examples shown in Figs. 3 and 4, the number of data points that a sensor needs to process is reduced from 1600 to 100. As the time resolution of thumbnail reduces to \((1000 \cdot 2^l)/f_s\) milliseconds, the P-phase picking error caused by the downsampling will be \((500 \cdot 2^l)/f_s\) milliseconds. For the examples shown in Figs. 3 and 4, the error is 80 milliseconds. Such a granularity is satisfactory for the preliminary P-phase picking.

Our lightweight preliminary P-phase picking algorithm is as follows. For a candidate P-phase \( \hat{p} \), the sensor computes the signal energies (i.e., the sample variances) of the thumbnail signal with length of two seconds before and after \( \hat{p} \), respectively. The preliminary P-phase, denoted by \( p \), is given by

\[
P = 2^l \times \mathop{\text{arg max}}_{\hat{p} \in \text{thumbnail}} \frac{\text{signal energy after } \hat{p}}{\text{signal energy before } \hat{p}}.
\]

Note that the scaling factor \( 2^l \) maps the pick in thumbnail to the original time domain. The complexity of the above algorithm is \( O(n/2^l) \), where \( l \) is the level of DWT. In contrast to the high complexities of advanced picking algorithms, e.g., \( O(n^3) \) for AR-AIC picker [8]. Note that the idea of maximizing the signal energy ratio in Eq. (1) is inspired by...
the AR-AIC picker [8] that maximizes the dissimilarity of two AR models. As the thumbnail preserves most of the P-wave energy, the performance of our preliminary picker is expected to be satisfactory. In the last columns of Figs. 3 and 4, the vertical red lines represent the preliminary P-phases. We can see that our preliminary picker can accurately pick the P-phases from the thumbnails. In Section VIII, we will conduct extensive evaluation on the accuracy of the preliminary picker.

V. SENSOR SELECTION FOR EARTHQUAKE TIMING

In this section, we first analyze the impact of timing on source localization and then formally formulate the sensor selection problem. Note that our sensor selection approach can be readily extended to address other volcano monitoring applications that take sensors’ P-phases as inputs and have theoretical error analysis.

A. Impact of Timing on Source Localization

As the propagation speed of P-wave varies with the depth in earth, the earthquake source localization often solves a joint problem of error minimization and ray tracing. Suppose a set of sensors, denoted by $S$, belongs to the cluster under consideration. Let $z_i$ and $z_o$ denote the 3-dimensional Cartesian coordinates of sensor $i$ and the earthquake source, $p_i$ and $p_o$ denote the time instances of the P-phase picked by sensor $i$ and the occurrence of earthquake at its source location, $v$ denote a list of P-wave speeds at different depths. We assume that $\{z_i|i \in S\}$ are known, e.g., by inquiring the GPS module on sensors [4]. The $z_o$ and $p_o$ are the unknowns to be estimated from the P-phases $\{p_i|i \in S\}$. We have

$$p_i - p_o = \tau(z_i, z_o|v) + \epsilon_i, \quad \forall i \in S$$

where $\tau(z_i, z_o|v)$ is the P-wave travel time from the source to sensor $i$ given the velocity model $v$, and $\epsilon_i$ is the random error experienced by sensor $i$. In this paper, we employ the ray tracing algorithm in the RSEIS R package [16] to calculate $\tau(z_i, z_o|v)$. As the data traces used in this paper were collected on Mount St. Helens, we set $v$ according to the results reported in a tomographic study to St. Helens [5]. We assume that $\epsilon_i$ follows zero-mean normal distribution with variance $\varsigma^2$. The variance $\varsigma^2$ captures the error of the P-phase picked from the seismic signal with respect to the true P-phase. As will be shown later, the localization algorithm and its accuracy analysis is independent of $\varsigma^2$. Hence, the variance $\varsigma^2$ can be unknown to the network. The unknown $p_o$ can be canceled out by subtracting Eq. (2) with $i = r$ from the same equation with $i \in S \setminus \{r\}$, yielding

$$p'_i = \tau(z_i, z_o|v) - \tau(z_r, z_o|v) + \epsilon'_i, \quad \forall i \in S \setminus \{r\},$$

where $p'_i = p_i - p_r$, $\epsilon'_i = \epsilon_i - \epsilon_r$ and sensor $r$ is the reference node. Note that $\epsilon'_i$ follows zero-mean normal distribution with variance $2\varsigma^2$. In this paper, we adopt maximum-likelihood (ML) approach to estimate $z_o$ from Eq. (3). The ML estimate of $z_o$, denoted by $\tilde{z}_o$, is given by the least-square solution:

$$\tilde{z}_o = \arg \min_{z_o} \sum_{i \in S \setminus \{r\}} (p'_i - \tau(z_i, z_o|v) + \tau(z_r, z_o|v))^2.$$
designer can set $C$ to meet bandwidth limitation, energy consumption budget and real-time performance.

In our approach, the cluster head first solves the least-square problem in Eq. (4) using the Nelder-Mead algorithm [18]. For any candidate sensor subset, we consistently use $\tilde{z}_n$ to compute $\hat{E}$. As $E$ is a non-linear and non-convex function, it is difficult to solve the sensor selection problem with polynomial complexity. In our experiments, the execution time of the Nelder-Mead algorithm on Imote2 is around 4 seconds. The brutal-force search takes 0.08 and 8.2 seconds when the cluster size is 10 and 16, respectively. Note that our numerical study shows that a cluster size of about 15 can well trade-off the localization performance and communication overhead within the cluster (c.f., Appendix C). Therefore, the computation overhead of brutal-force search is acceptable. In our future work, we plan to develop approximate solutions with polynomial complexity to be scalable for large-scale clusters.

VI. COMPRESSIVE SAMPLING FOR EARTHQUAKE TIMING

This section presents our approach of compressing and collecting the seismic signals from the selected sensors based on compressive sampling (CS) [10]. We first briefly review the CS theory. Let $y \in \mathbb{R}^{n \times 1}$ denote the compressed signal and $A \in \mathbb{R}^{m \times n}$ denote the random projection matrix, where $m < n$. The compression is expressed as $y = Ax$, where $x$ is the wavelet coefficients of the original signal. If $x$ is $k$-sparse and $A$ complies the restricted isometry property (RIP) of order $k$, the original signal $s$ can be exactly reconstructed from $y$ [10]. The wavelet transform of the reconstructed signal, denoted by $\tilde{x}$, is given by $\tilde{x} = \arg \min_x ||x||_2$ subject to $y = Ax$. The above optimization can be solved by various algorithms such as the iterative hard thresholding method [19]. With $\tilde{x}$, the reconstructed seismic signal, denoted by $\tilde{s}$, is given by inverse DWT, i.e., $\tilde{s} = \Psi \tilde{x}$.

We now discuss the design of CS for earthquake timing on resource-constrained sensors. We adopt the binary random projection matrix [20] that is promising for the implementation on resource-constrained sensors. Specifically, only the positions of ‘1’s need be to stored and the multiplication $Ax$ is simply the sum of the elements of $x$ at these positions. The binary random matrix complies RIP of order $k$ if $m \geq h \cdot k \cdot \log(n/k)$, where $h$ is a unknown constant [20]. From the results shown in Fig. 5, the sparsity $\rho$ of volcanic seismic signal typically ranges from 0.1 to 0.6. Hence, $\log(n/k) = \log(1/\rho)$ ranges from $\log(1.67)$ to $\log(10)$. We define $\eta = \log(10) \cdot h$. If $m \geq \eta \cdot \rho \cdot n$, the RIP condition must be satisfied. Therefore, we let

$$m(\rho) = \eta \cdot \rho \cdot n.$$  

Many studies have reported that $\eta = 4$ is a safe setting that ensures satisfactory reconstruction [10]. However, as the sparsity $\rho$ estimated in Section IV-A does not follow the strict definition of sparsity (i.e., the ratio of non-zeros), the setting of $\eta = 4$ might be overly conservative for earthquake timing, which may result in excessive data transmission. In Section VIII, we will extensively evaluate the impact of $\eta$ on the quality of seismic signal reconstruction as well as the P-phase picking. The results show that the setting of $\eta = 1.5$ can well trade-off the volume of data transmission and the P-phase picking error caused by reconstruct error. We note that the CS-compressed signal can be further compressed by other data compression algorithms if computation resource allows.

VII. IMPLEMENTATION

We have implemented the proposed algorithms for earthquake timing in TinyOS 2.1.0 on TelosB platform and conducted testbed experiments in laboratory. Several important implementation details are presented as follows.

Data acquisition: To improve the realism of testbed experiments, we create a volume of 320KB on mote’s flash and load it with the seismic data traces collected in OASIS [4]. We implement a nesC module that provides the standard ReadStream interface to read seismic data from flash to process.

Seismic signal processing: We implement all the four seismic signal processing algorithms running at low-end sensors, i.e., DWT, sparsity estimation, preliminary P-phase picker and CS. A major challenge is that TelosB mote has limited on-board resources, including a 8 MHz CPU and only 10 KB RAM. We conducted extensive code optimization to the all the signal processing algorithms in the system. First, we adopt fixed point arithmetic, which can speed up the decimal computation by up to 10 folds on TelosB with respect to the default float point arithmetic. Second, we maintain a single input/output data buffer for the four pipelined algorithms and wire the output of each algorithm back to the buffer. This pipeline implementation significantly reduces the RAM usage.

Cluster head: We use a laptop computer to simulate the high-end cluster head and implement all the cluster head algorithms in ANSI C. To evaluate the computation overhead of these algorithms on embedded platforms such as Imote2, we cross-compile the programs and run them in the SimIt-ARM 3.0 [21], which can simulate the XScale processor on Imote2.

VIII. PERFORMANCE EVALUATION

We conduct testbed experiments and extensive simulations based on real data traces collected by 12 nodes in the OASIS project [4]. The testbed experiments evaluate the feasibility of the proposed signal processing algorithms on low-end sensor platforms. The trace-driven simulations extensively evaluate the performance of each algorithm in our earthquake timing approach. The data set used in our evaluation spans 5.5 months and comprises 30 significant earthquakes. In each earthquake, the data traces collected by 12 nodes for 10 minutes are used. The onset time of each earthquake is detected by a Bayesian approach [14].

A. Testbed Experiments

We evaluate the computation and storage overhead of the algorithms running on low-end sensors in a testbed of 12 TelosB motes. We load the 12 TelosB motes with the real data traces sampled at 100 Hz by 12 nodes nodes in the OASIS
The preliminary picker can provide a significantly refined P-phase. The mean and standard deviation of the error of earthquake onset time are 280 milliseconds and 1310 milliseconds. Therefore, the preliminary picker can be approximated as a zero-mean error picker with respect to the AR-AIC picker. The standard deviation of the picking error is 8.5 milliseconds. Therefore, the preliminary P-phase picking error based on 100 sensor data traces. The mean of picking error dramatically drops when \( \eta \) increases from 0.75 to 1.5 and becomes flat after 1.5. Therefore, \( \eta = 1.5 \) is a proper setting to achieve the satisfactory gain of P-phase picking accuracy to the data transmission volume. The mean of picking error is within \([-15\text{ ms}, 15\text{ ms}]\) when \( \eta \geq 1.5 \), and hence can be safely approximated as zero-mean error. With the setting of \( \eta = 1.5 \), Fig. 10 shows the original and reconstructed signals received by a sensor in three earthquakes. We can see from the figure that the signals can be accurately reconstructed and the P-phases can be well preserved. In Appendix D, we present the source localization results of two earthquakes based on the preliminary picks and the P-phases picked by AR-AIC picker on the original/reconstructed signals. The results show that our earthquake timing approach can lead to accurate source localization.

2) Effectiveness of Sensor Selection: In the simulations, each sensor directly communicates with the cluster head. Each packet carries total ten 4-byte data points. Therefore, by setting the unit communication cost \( c_1 = 1/10 \), the upper bound of communication cost \( C \) characterizes the number of packet transmissions that are allowed in an earthquake timing process. Fig. 9 plots the error metric \( E \) (given by Eq. (5)) and the number of sensors of the optimal selection versus the upper bound of packet number \( C \). We can see that if more packet transmissions are allowed, the cluster head will select more sensors to collect data from them. Consistent with intuition, the error metric decreases with the number of packet transmissions. When \( C \in [170, 250] \), total four sensors are selected. However, the selected four sensors can be different. When more packets are allowed, the sensors with higher sparsity but more contributory to the source localization.
A in the projection matrix the cluster head reconstructs the signal, it only uses the rows probability of PRR. The cluster head can detect lost packets each link from the sensor to the cluster head has the same quake timing approach. In the simulations, we assume that sensors are subject to unreliable communication links. We are deployed in challenging environment [4], the significant loss of source localization performance.

3) Impact of Random Packet Loss: As volcano monitoring WSNs are deployed in challenging environment [4], the sensors are subject to unreliable communication links. We evaluate the impact of random packet loss on our earthquake timing approach. In the simulations, we assume that each link from the sensor to the cluster head has the same packet reception ratio (PRR). Each packet is received with a probability of PRR. The cluster head can detect lost packets from the sequence numbers in the received packets. When the cluster head reconstructs the signal, it only uses the rows in the projection matrix \( A \) that correspond to the received data points. Therefore, the effect of packet loss is similar to choosing a smaller \( m \) in CS. We compare our CS-based approach with a baseline that captures the key idea of lossy compression. The baseline transmits the largest coefficients together with their indexes in the wavelet domain. The number of transmitted coefficients is chosen to make the baseline has the same number of packets with our approach. The curves in Fig. 11 plot the relative reconstruction error versus PRR. When no packet loss happens, the baseline outperforms our approach. When the PRR is lower than 90%, our approach outperforms the baseline. The histograms in Fig. 11 plot the average P-phase picking error of our CS-based approach. As the effect of packet loss is similar to choosing a smaller \( m \) in CS, the reconstruction is resilient to packet loss when the PRR is no lower than 80%. More comparative results including the P-phase picking on the reconstructed signal in presence of packet loss can be found in Appendix E.

4) Compression Efficiency: We finally compare our CS-based compression scheme with several baselines in terms of compression ratio and execution time. The baselines are: (1) SLZW [22], which is a lossless compression algorithm designed for WSNs; (2) ALFC [23], which is a real-time predictive lossless compression algorithm and has been employed in the OASIS project [4]; (3) Lempel-Ziv coding (LZ77), which is widely employed in the traditional data-collection-based volcano monitoring systems. Compression ratio is the ratio of compressed size to the uncompressed size. Fig. 12 plots the compression ratios and relative execution times of various approaches. We can see that our CS-based scheme can further save more than 10% data transmission volume compared with the baselines. The relative execution time is calculated with respect to LZ77. We can see that our CS-based scheme is faster than LZ77 but slower than SLZW and ALFC. We note that all these baseline compression algorithms cannot predict the exact volume of compressed signal.

IX. CONCLUSION

This paper presents a holistic and energy-efficient approach to accurate volcanic earthquake timing in WSNs. We develop a collection of seismic signal processing algorithms that collaboratively pick the arrival times of seismic primary waves received by sensors. A dynamic sensor selection problem is formulated to maximize the information quality of picked arrival times for earthquake source localization subject to a bound of data transmission overhead, which can be configured to meet bandwidth limitation, energy budget and real-time performance requirement. Testbed experiments and extensive simulations based on real data traces collected on an active volcano demonstrate the feasibility and effectiveness of our approach.

REFERENCES

APPENDIX

A. Numerical Study on Network Architecture

1) Event and Detection Models: In this section, we compare two different network architectures by modeling the energy consumption of a single node under them. The first architecture is the hierarchical network discussed in Section III-A. In this architecture, the cluster adopts the approach presented in this paper to time the earthquake. In the second architecture (referred to as non-hierarchical network), a cluster is composed of only high-end nodes and each node is capable of running the advanced P-phase picking algorithms. Therefore, in the non-hierarchical network, P-phase picking runs locally at nodes. We quantify the advantage of the hierarchical network over the non-hierarchical architecture by analyzing the average energy consumption of a cluster member under the two architectures, respectively. We first make a few assumptions on the event occurrence and detection models. To simplify the discussion, we assume that each cluster member directly communicate with the cluster head. Each cluster member detects earthquake every $T$ seconds by a detection algorithm such as STA/LTA. The $T$ is referred to as detection period. In the absence of earthquake, the cluster members do not send local detection decisions to save energy. In the presence of earthquake, the cluster members send local detection decisions to the cluster head, which then fuses the results to make the final detection decision. The cluster head will send a message containing the earthquake onset time to the cluster members. We note that the earthquake onset time is an important input to the approach for the hierarchical network architecture employed in this paper (see Section IV) as well as the P-phase picking algorithm running on each cluster member in the non-hierarchical architecture. Let $K$ denote the number of positive detection results per day and $\tau$ denote the average number of detection periods that an event lasts for. Hence, averagely, there is no event during $\frac{24 \times 3600}{T} - K \cdot \tau$ detection periods per day, and there are events during $K \cdot \tau$ detection periods per day. Other symbols are defined in Table II.

2) Energy Consumption of Hierarchical Network: We first analyze the energy consumption of a cluster member in the hierarchical network. The total energy consumption of a low-end node in the absence of earthquake is

$$E_1^L = \left( \frac{24 \times 3600}{T} - K \cdot \tau \right) \cdot (t_d^L \cdot I_b^L + (T - t_d^L) \cdot I_s^L).$$

In the presence of earthquake, the node also detects event every detection period and transmits its local decision to the cluster head, consuming

$$E_2^L = K \cdot \tau \cdot (t_d^L \cdot I_b^L + t_m^L \cdot I_s^L).$$

For each earthquake event, the node receives the earthquake onset time from cluster head. The node then performs sparsity estimation, DWT and compressive sampling. It also transmits a message containing the sparsity and receives a selection message. We assume that all nodes are always selected. The compressed data is divided into $n_m$ TinyOS messages with default payload size. As a result, the energy consumption is

$$E_3^L = K \cdot (t_m^L \cdot I_r^L + t_m^L \cdot I_b^L + t_m^L \cdot I_r^L + t_m^L \cdot I_s^L + n_m \cdot t_m^L \cdot I_s^L).$$

For each earthquake event, the MCU will be busy for $(\tau \cdot t_d^L + \tau \cdot t_m^L + t_m^L + t_m^L + t_m^L + n_m) \cdot t_m^L)$ seconds. Therefore, the energy consumed in the MCU’s standby state in the presence of earthquake events is

$$E_4^L = K \cdot (\tau \cdot T - \tau \cdot t_d^L - (t_m^L + t_m^L - t_m^L - 2 \cdot t_m^L - n_m \cdot t_m^L) \cdot I_s^L).$$

The total energy consumption of MCU and radio of a low-end node per day is

$$E_{LR}^L = E_1^L + E_2^L + E_3^L + E_4^L.$$

3) Energy Consumption of Non-hierarchical Network: We then analyze the energy consumption of a cluster member in the non-hierarchical network. The total energy consumption of a high-end node in the absence of earthquake is

$$E_1^H = \left( \frac{24 \times 3600}{T} - K \cdot \tau \right) \cdot (t_d^H \cdot I_b^H + (T - t_d^H) \cdot I_s^H).$$

In the presence of earthquake, the node also detects event every detection period and transmit its local decision to the cluster head, consuming

$$E_2^H = K \cdot \tau (t_d^H \cdot I_b^H + t_m^H \cdot I_s^H).$$
TABLE II
NOTATION OF ENERGY PROFILE ANALYSIS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
<th>Setting [11], [14]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>detection period in second</td>
<td>1 s</td>
</tr>
<tr>
<td>( t_{L_d}^H / t_{H_d}^H )</td>
<td>MCU time of detection algorithm in each ( T )</td>
<td>variables in the numerical study</td>
</tr>
<tr>
<td>( t_{L_m}^H / t_{H_m}^H )</td>
<td>time for transmitting a TinyOS message</td>
<td>10 ms</td>
</tr>
<tr>
<td>( t_{L_r}^H )</td>
<td>MCU time of P-phase picking</td>
<td>1612 ms</td>
</tr>
<tr>
<td>( t_{L_c}^H )</td>
<td>MCU time of sparsity estimation, DWT, CS</td>
<td>3 s</td>
</tr>
<tr>
<td>( n_m )</td>
<td>number of TinyOS messages sent by a low-end node</td>
<td>( \frac{1600 \times 0.16 \times 2}{29} = 20 )</td>
</tr>
<tr>
<td>( n_r )</td>
<td>number of raw data TinyOS messages sent by a low-end node</td>
<td>( \frac{1600 \times 2}{29} = 110 )</td>
</tr>
<tr>
<td>( I_{L_d}^H )</td>
<td>current draw when MCU is busy</td>
<td>1.8 mA / 31 mA</td>
</tr>
<tr>
<td>( I_{L_r}^H )</td>
<td>current draw when MCU stands by</td>
<td>5.1 µA / 390 µA</td>
</tr>
<tr>
<td>( I_{L_r}^H )</td>
<td>current draw when radio is transmitting</td>
<td>19.5 mA / 44 mA</td>
</tr>
<tr>
<td>( I_{L_r}^H )</td>
<td>current draw when radio is receiving</td>
<td>21.8 mA / 44 mA</td>
</tr>
<tr>
<td>( K )</td>
<td>number of positive detection decisions per day</td>
<td>variable in the numerical study</td>
</tr>
<tr>
<td>( \tau )</td>
<td>average number of detection periods per event</td>
<td>15 s</td>
</tr>
</tbody>
</table>

* The superscripts \(^L\) and \(^H\) represent low-end and high-end sensors.

\(^\dagger\) Each data point is 2 bytes; Each TinyOS message carries 29 bytes payload; 0.16 is the compression ratio (cf. Section 8.2.5 in [24])

For each earthquake event, the node receives the earthquake onset time from the cluster head. The node then performs P-phase picking and transmits a message containing the P-phase to the cluster head. As a result, the energy consumption is

\[ E_d^H = K \cdot (t_{L_d}^H \cdot I_d^H + t_{H_d}^H \cdot I_d^H + t_{L_m}^H \cdot I_m^H). \]

For each earthquake event, the MCU will be busy for \((\tau \cdot t_d^H + \tau \cdot t_m^H + t_p^H + t_{H_m}^H)\) seconds. Therefore, the energy consumed in the MCU’s standby state in the presence of earthquake event is

\[ E_{d}^H = K \cdot (\tau \cdot T - \tau \cdot t_d^H - \tau \cdot t_m^H - t_p^H - t_{H_m}^H) \cdot I_d^H. \]

The total energy consumption of MCU and radio of a high-end node per day is

\[ E_{CR}^H = E_1^H + E_2^H + E_3^H + E_4^H. \]

4) Numerical Results: Let \( \alpha \) denote the ratio of the MCU times of high-end and low-end nodes for processing the same amount of data. Therefore, \( t_d^H = \alpha \cdot t_d^L \). We assume that the low-end and high-end sensors are the commercial off-the-shelf TelosB and Imote2. The corresponding the settings of these two platforms are listed in Table II. Figs. 13 plots the maps and contours of \( E_{CR}^H / E_{CR}^L \) under different settings of \( \alpha \). We note that the range of \( t_d^L \) is set according to the testbed experiments in our previous work [14]. Specifically, if STA/LTA algorithm or the frequency-based Bayesian approach in [14] is used to detect earthquake, \( t_d^L \) is around 10 ms and 100 ms, respectively. From the figures, we can see that the node energy consumption in the hierarchical network is significantly lower than that in the non-hierarchical network.

We now study the lifetime of the two network architectures. The results shown in Fig. 13 only consider the energy consumed in computation and radio. The energy consumption of sensor circuit is also a major part of the total energy consumption, which is also a common part of the two network architectures. We now discuss the reasonable range of sensor circuit current draw. We assume that both the low-end mote and high-end mote are connected to the data acquisition board MDA320CA that acquires data from the seismic sensor. A few seismic sensor products such as Geophone seismometer [4] do not draw power. Therefore, we only need to account for the power consumption of MDA320CA. According to the measurement results in [25], the MDA320CA consumes about 22 mW. As a result, the current draw of MDA320CA from three D-cell batteries (4.5 V) is 4.889 mA \( \simeq 5 \) mA. Therefore, 5 to 20 mA is a reasonable range for sensor circuit. In addition to sensor circuit, the false alarm rate of the earthquake detection algorithm also plays an important role in the network lifetime. The network designers often set the parameters of the detection algorithm to achieve a detection probability close to 100% to capture most earthquake events, while allowing false alarms. In this numerical study, We assume that there are 20 earthquake events per day and all these events can be detected by the detection algorithm. Moreover, let \( P_F \) denote the false alarm rate of the detection algorithm. Therefore, there will be \((24 \times 3600 - 20) P_F \) false alarms per day. Hence, \( K = \left(\frac{24 \times 3600}{5} - 20\right) P_F + 20 \). Fig. 14 shows the projected node lifetime versus the current draw of sensor circuit under various false alarm rates. From the figure, we can see that the hierarchical network has longer lifetime under various settings.

B. Numerical Comparison between CS-based Approach and Centralized Approach

In this numerical study, we compare our CS-based approach against a centralized approach. We also assume the earthquake occurrence and onset time detection models presented in
Fig. 13. Map and contours of $E_{CR}^L/E_{CR}^H$. The range of MCU time of detection algorithm is set according to [14].

Fig. 14. Projected node lifetime over 2 Alkaline D-cell batteries versus sensor current draw under various false alarm rates (MCU time of detection on TelosB is 10 ms)
Appendix A1. In the centralized approach, upon receiving the earthquake onset time from the cluster head, each low-end sensor will retrieve a signal of 1600 data points around the onset time (cf. Section IV) and transmit this signal without compression to the cluster head. Note that this centralized approach does not select sensors. In the rest of this section, we first analyze the energy consumption of this centralized approach and then compare it with that of our CS-based approach.

Under the centralized approach, $E_{CS}^L$ and $E_{CR}^L$ are the same as those in Appendix A2. We now analyze $E_{CS}^L$ and $E_{CR}^L$. For each earthquake event, the node receives the earthquake onset time from cluster head. The node then transmits the raw signal, which is encapsulated into $n_r$ TinyOS messages with default payload size. As a result, the energy consumption is

$$E_{CS}^L = K \cdot (t_m^L + I_s^L + n_r \cdot t_m^L)$$

For each earthquake event, the MCU will be busy for $(\tau \cdot t_m^L + \tau \cdot t_m^L + n_r \cdot t_m^L)$ seconds. Therefore, the energy consumed in the MCU’s standby state in the presence of earthquake events is

$$E_{CR}^L = K \cdot (\tau \cdot T - \tau \cdot t_m^L - \tau \cdot t_m^L - n_r \cdot t_m^L) \cdot I_s^L.$$

The total energy consumption of MCU and radio of a low-end node per day is

$$E_{CS,R}^L = E_{CS}^L + E_{CR}^L.$$

When we compare the lifetime of our approach and the centralized approach, the energy consumption model for the sensor circuit is different from Appendix A. Both our approach and the centralized approach are applied to the low-end sensors such as TelosB under the hierarchical architecture. The MCU of TelosB, MSP430, provides A/D channels. Therefore, the TelosB mote does not need a full-functional data acquisition board such as MDA320CA. Instead, a signal amplifier is directly connected to the A/D pins of TelosB. In the VolcanoSRI project [26], we built a signal amplifier board for Geophone seismometer based on TI LM324 Operational Amplifier. According to our measurements, this board consumes 3 mA. Therefore, in this numerical study, we set the current draw of the sensor circuit to be 3 mA. Fig. 15 plots the projected node lifetime over 2 Alkaline D-cell batteries versus false alarm rate. Our CS-based approach can increase the lifetime by about 7% and 12%, compared with the centralized approach, when the false alarm rate is 1.5% and 3%, respectively. Note that the above analysis assumes that all sensors are always selected. If only a subset of sensors are selected, the CS-based approach can further increase lifetime.

C. Sensor Clustering

In this section, we discuss the impact of sensor clustering scheme on the earthquake source localization through a numerical study. In the numerical study, 100 sensors are randomly deployed over a $6 \times 6$ km$^2$ square region. Such an area is consistent with the geographic area of Mount St. Helens. We assume that the earthquake happens beneath the center of the square region. We adopt two clustering schemes, referred to as geographic clustering and random clustering. Suppose the cluster size is $N$. In the geographic clustering scheme, the $(N - 1)$ sensors that are closest to the cluster head form a cluster. For the cluster, we calculate the source localization error metric given by Eq. (5). Fig. 16(a) plots the average source localization error metric versus the cluster size, when the cluster head is randomly selected from the whole network. In the random clustering, $N$ sensors are randomly selected from the whole network to form a cluster. Fig. 16(b) plots the average source localization error metric versus the cluster size. We can draw three important observations from the two figures. First, as the sensors in a cluster under the geographic clustering scheme are located within a smaller region compared with under the random clustering scheme, the geographic clustering scheme has larger source localization error than the random clustering scheme. However, the geographic clustering scheme has lower communication cost. Second, the localization error increases with the source depth. Note that the above two observations are consistent with many existing studies on source localization. Third, for both clustering schemes, the localization error has a rapid drop initially and then becomes flat when the cluster size increases. This result implies that adding a sensor becomes less beneficial for a larger cluster. For the geographic clustering scheme, when the cluster size is larger than 15, the decease of localization error becomes insignificant. Therefore, 15 to 20 will be satisfactory setting for cluster size. We note that the cluster size dominates the communication cost within the cluster. The optimal setting of cluster size that jointly accounts
for source localization performance and communication cost still remains an open issue.

D. Source Localization

To evaluate the effectiveness of our CS-based earthquake timing approach, we use the preliminary picking results, the P-phases picked by the AR-AIC picker on the original signal and the reconstructed signal to compute the earthquake source location. Fig. 17 shows the location of the sensors in the OASIS project on Mount St. Helens. It also shows the source localization results for two earthquakes. From the figure, we can see that the source locations based on the preliminary picks and the P-phases picked from the reconstructed signals are close to that based on the original signals. The error is within 2 km. We note that such an error is satisfactory in seismology because the granularity of the velocity model of the volcano (i.e., v in Section V-A) has a granularity of a few kilometers.

E. More Results on Packet Loss

Fig. 18 to Fig. 23 show the P-phase picking (represented by vertical line) on the original signal, reconstructed signals by the baseline and our CS-based approach, when the PRR is from 100% to 50%.
Fig. 21. PRR=70% 

Fig. 22. PRR=60% 

Fig. 23. PRR=50%