Smart Seismic Network for Shallow Subsurface Imaging and Infrastructure Security

Maria Valero*, Fangyu Li and WenZhan Song

Center for Cyber-Physical Systems,
University of Georgia,
United States
E-mail: maria.valero@uga.edu
E-mail: fangyu.li@uga.edu
E-mail: wsong@uga.edu
*Corresponding author

Abstract: The use of seismic arrays as a tool for imaging subsurface infrastructures and monitoring the corresponding underground activities enables real-time subsurface security and surveillance applications. However, the existing approach relies on manual data collection and centralized computing, which may not work in communication-denied environments. These approaches may take a long time to get useful results. In this paper, we present a real-time smart seismic imaging system based on Ambient Noise Imaging on Networks (ANION) for a variety of subsurface infrastructure imaging applications. The proposed approach integrates in-situ signal processing techniques as well as inter-nodes communication and cooperation to obtain reliable velocity maps for subsurface characterization and monitoring. It generates real-time subsurface images by taking advantage of collective computation power in sensor networks while avoiding transferring all raw data to a central place or server. ANION system is autonomous, self-healing, scalable and independent of external interventions. Field tests demonstrate that such a system can detect underground pipelines and potentially its leakage that has important implications on infrastructure security. The uses can be extended to other applications like border security, building monitoring, underground water detection for agriculture, etc. An exhaustive evaluation regarding bandwidth utilization and communication cost were conducted to highlight the benefits of the proposed approach.

Keywords: Smart Sensor Network, Subsurface security, infrastructure detection, activity monitoring, ambient noise.

Reference to this paper should be made as follows: Valero, M., Li, F., and Song. W. (2019) 'Smart Seismic Network for Shallow Subsurface Imaging and Infrastructure Security', *International Journal of Sensor Networks*, Vol. x, No. x, pp.xxx–xxx.

Biographical notes: Maria Valero received her MS in Computer Science from University of Los Andes (Venezuela) in 2009. She is PhD candidate at the College of Engineering, University of Georgia. Her research interests include distributed computing, signal processing, wireless sensor networks, cyber-physical systems.

Fangyu Li is a postdoctoral associate with the College of Engineering, University of Georgia. He received his PhD in Geophysics from University of Oklahoma, and MS and BS degrees in Electrical Engineering from Tsinghua University and Beihang University, respectively. His research interests include signal processing, seismic imaging, geophysical interpretation, machine learning, deep learning, distributed computing, and cyber-physical systems.

WenZhan Song received Ph.D. in Computer Science from Illinois Institute of Technology (2005). He is a Chair Professor of Electrical and Computer Engineering in the University of Georgia. Dr. Songs research focuses on cyber-physical systems and security and their applications in energy, environment, food and health sectors. He received NSF CAREER award in 2010.

1 Introduction

Seismic array is an effective tool for monitoring subsurface structures and activities. The ability to "see" through the ground would enable many infrastructure security applications. For instance, water companies are under constant pressure to ensure the water leakage keeps to a minimum. Many techniques to detect water leakage require an invasive approach where sensors need to be installed directly on the pipelines (Almeida

et al., 2015). A non-invasive and efficient way to detect pipelines and capture the velocity variation produced by external elements like water is crucial for civil engineering security applications. Furthermore, if it is possible to see under the ground, many subsurface issues can be addressed; for example, tunnel detection for homeland/border security, building foundation status monitoring for infrastructure security, etc.

Seismic tomography is one way to imaging subsurface based on the analysis of the subsurface waveform propagation. The type of tomography based on ambient noise cross-correlation is called Ambient Noise Tomography (ANT) (Ritzwoller et al., 2011). In ANT, the data recovered from ambient seismic noise, which implies no need for active energy sources like earthquakes. By correlating the ambient noise signal from two or more different locations, seismic sensors can estimate a subsurface image of the area.

The problem is that existing ANT solutions rely on a centralized approach for processing the raw data captured by seismic sensors. The data are gathered manually or relayed by expensive broadband stations, and then processed at a base station. These approaches are time-consuming, and the final tomography is obtained in weeks, even months. Furthermore, collecting data in one central place may not work in communication-denied environments where the external computing infrastructure cannot be relied on. Real-time seismic imaging within sensor networks (Kamath et al., 2016; Zhao et al., 2015; Kamath et al., 2015; Shi et al., 2013; Kamath et al., 2015; Valero et al., 2017, 2018; Song et al., 2019) enables unprecedented Intelligence, Surveillance, Target Acquisition, and Reconnaissance (ISTAR) of underground/underwater structures, targets and activities.

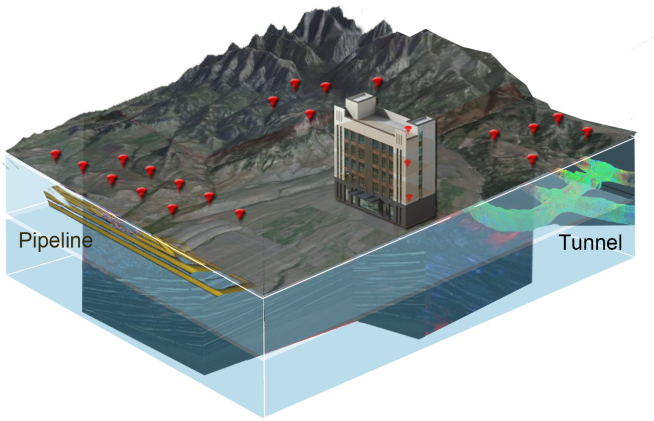


Figure 1 Some potential subsurface security applications using seismic sensor networks system for subsurface imaging: pipeline detection, tunnel detection, infrastructure security monitoring.

In this paper, we present a smart wireless seismic network system that uses seismic sensors to perform a tomography of the subsurface based on ambient noise cross-correlation measurements. The system is able to produce a subsurface image that represents the underground velocity variations when smart seismic sensors are deployed over the area to study (Fig. 1), and they use wireless network communication and in-situ computing. There is no need of gathering data to a central place since the sensors use collaborative and distributed computing to share only pre-processed information to generate the final tomography. The system has been called ANION, ambient noise seismic imaging on networks, and it will be referred with this name in the rest of the paper.

ANION is an end-to-end system, fully configurable, that can be used to imaging shallow subsurface infrastructures and provide information for security purposes. For example, ANION can provide real-time images of underground pipelines; this information can be useful for determining if there exists a water leakage in that pipe. ANION can provide images of shallow tunnels, and this information can be important for border security decision-making. Many other applications on subsurface infrastructure security can also be benefited.

Our system integrates in-situ signal processing techniques and inter-nodes communication and cooperation to obtain reliable velocity maps for geophysics and engineering investigations. We generate real-time images by taking advantage of the new generation of sensors and without the need of transferring all raw data to a central place or server.

The main contributions of this paper can be summarized as follows: (i) a system that performs subsurface imaging and achieves real-time results using in-situ computing and distributed sensor collaboration; this implies eliminating the need to send raw data to a server for post-processing, (ii) a system that can be easily adapted to a broad range of subsurface security/monitoring applications with minimal user intervention or system modifications, (iii) a system that takes advantage of ambient noise theory instead of using active sources (earthquakes or explosions) for illuminating the subsurface, (iv) a system that can meet network bandwidth constraints while reduces communication cost between sensors.

The rest of the paper is organized as follows: section section 2 provides needed information of ANT and surveys different kind of applications that use ambient noise data for security purposes, section 3 provides details of the system design, fundamental principles of the proposed solution and the details of our algorithm, section 4 introduces the hardware and software we use to develop ANION system. Section 5 shows the in-field experiments of our system. We discuss the results and highlight the benefits of the system regarding bandwidth utilization, communication cost, and link reliability. We conclude the paper in section 6.

2 Background

In this section, we provide a needed background on the basis of ANT and the potential and current applications for infrastructure security purposes.

2.1 Ambient Noise Tomography (ANT)

Ambient noise tomography has become one of the fastest growing research areas in seismology and exploration geophysics. Compared to earthquake-based seismic tomography methods, ambient noise tomography is particularly useful in imaging shallow earth structures (Lin et al., 2008; Moschetti et al., 2010). It can be applied to regions with non-existent seismicity and produces reliable measurements at frequencies that are particularly difficult using earthquakes or explosions due to scattering and attenuation. This advantage represents an attractive cheap scenario since producing active energy sources (explosions) in non-seismic areas is very costly. Moreover, because of the persistent nature of the seismic background noise, temporal variation of the earth structure can be analyzed and monitored by studying the variation in the noise cross-correlation function (Brenguier et al., 2008; Duputel et al., 2009).

To perform tomography with ambient noise, many methods can be used, for example eikonal tomography (Lin et al., 2009b), straight-ray tomography (Barmin et al., 2001), seismic interferometry (Nicolson et al., 2012). All of them have their own properties and mathematical formulation. Even though these approaches have been successfully applied, they lack real-time results. The cross-correlation process needs at least several days for collecting data, and then manual extraction is needed to gather the information to a central server. Thus velocity maps may take days or months in being generated. In this paper, we took the principles of Eikonal Tomography, and we modify the procedure to make them to perform in-situ and distributed fashion. We developed ANT in networks leading to ANION creation. With ANION approach, many potential applications can be done in real-time and with minimum user intervention.

2.2 Potential Applications

Traditional use of ambient noise imaging includes the study of the subsurface at many different scales for geophysical exploration or environmental remediation (Mordret et al., 2013; de Ridder and Dellinger, 2011; Tomar et al., 2018; Brenguier et al., 2007). And, there is an increasing interest in structures monitoring and anomaly detection.

Passive seismic monitoring based on ambient noise assessment has been used for risk management and reduction in many engineering applications. Stork et al. (2018) investigated the potential to monitor seismic velocity changes following a hypothetical leak of CO_2 from Aquistore storage site in Saskatchewan, Canada using passive monitoring methods. They showed the map of near-surface velocities obtained with ambient noise techniques could be useful for near-surface static corrections when using active-source seismic reflection surveys to image and monitor the reservoir. However, further similar studies are required to assess leak detection at other CO_2 storage sites.

Liu et al. (2014) conducted an experiment to test the idea of using ambient noise to characterize building vulnerability for strong ground motions. By placing one geophone on each floor in a building (from the basement to the seventh floor), they were able to determine shear wave velocities, which implies that the ambient noise approach appears sensitive enough to reveal some wave propagation difference and variations in a civil infrastructure. Mordret et al. (2017) also measured the shear velocity and the apparent attenuation factor of a building using ambient noise techniques. They linked the velocity variations with weather parameters. They determined that the variation in velocities in the building is intrinsically related to air humidity.

Olivier et al. (2015) also took advantage of ambient noise techniques to investigate an active underground mine (Garpenberg, Sweden) by cross-correlating seismic noise generated by mining activities. Their study revealed the existence of a high-velocity zone and a low-velocity zone that corresponded with known ore bodies inside the mine.

All these geophysics and engineering solutions using ambient noise imaging make us argue that the ambient noise can be treated as a new source that is economical, practical, and particularly valuable for seismic hazard mitigation and anomaly and activities detection in urban and non-urban areas. It can also be viewed as a unified source to characterize the near-surface sediment and the infrastructure simultaneously. The main drawback of these applications relies on the off-line data processing. All of these solutions require gathering the ambient noise data to a central place for further processing. This may imply days or even months to get a result.

3 Algorithm and System Design

ANION is based on ambient noise tomography (Lin et al., 2008; Moschetti et al., 2010), which studies the temporal variation of the earth structure by analyzing the variation in the noise cross-correlation function of the signals in different sensors (Brenguier et al., 2008; Duputel et al., 2009). As a widely used technique in geophysical exploration for investigating structures under the earth's surface, ANION is based on the recorded background raw ambient noise data. Lin et al. (2008) discussed how to use surface wave (Rayleigh wave) to image the shallow structures.

ANION design is built on top of a distributed system architecture as shown in Fig. 2. In this section, we explain the fundamental principles behind the architecture work-flow of ANION. Every sensor performs

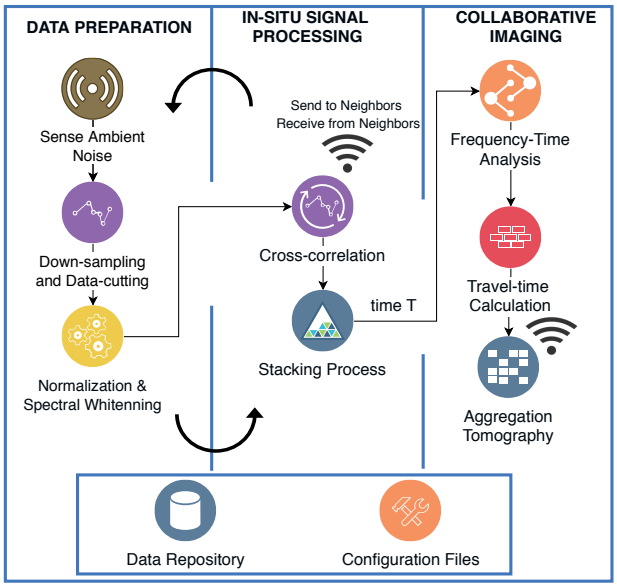


Figure 2 Proposed distributed architecture for ANION. One sensor perspective.

three stages in ANION: (i) data preparation, (ii) in-situ signal processing, and (iii) collaborative imaging. The majority of steps inside each stage are executed in-situ, and some of them require communication with neighbors via UDP (User Datagram Protocol). In Fig. 2, process with wireless symbol represents inter-node communication.

The data preparation and in-situ signal processing are performed continuously during the time the sensors are over the field and working. The collaborative imaging is executed every T time (for example, every 2 hours). T is a configured value in the configuration file.

3.1 Data Preparation

Every node senses the ambient noise from the medium. We use smart seismic sensors (Section 4.1) that sense ambient noise and record the raw data internally in a database (for future analysis if needed). During the data preparation, sensors perform a down-sampling and data-cutting. That is, sensors take t seconds or minutes or data to analysis each time. For example, if $t = 5$ minutes, then every 5 minutes of data, the sensors perform a down-sampling, a normalization and a spectral whitening. t is another configurable parameter in the configuration file. The purpose of normalization and spectral whitening is to accentuate ambient noise by attempting to remove earthquake signals and instrumental irregularities that tend to hide the ambient noise. We use a running-absolute-mean normalization (Bensen et al., 2007b). This method computes the running average of the absolute of the waveform in a normalization time window of fixed length and weight the waveform at the center of the window by the inverse of this average. Given a discrete time-series f, the normalization weight is $w_n = \frac{1}{2N+1}\sum_{i=n-N}^{n+N}|f_i|$, and the normalized datum is $\tilde{f}_n = f_n/w_n$. The width of the normalization window is $2N+1$. An example of data preparation can be seen in Fig. 3.

3.2 In-situ signal processing

After data preparation, sensors perform data cross-correlation with its neighbor nodes. For doing so, every node broadcasts its prepared data to its neighbor nodes. Prepared data includes a compression technique for improving communication cost. The noise cross-correlation C_{AB} between two stations is performed as:

$$\begin{aligned} C_{AB}(t) &= \int_{-\infty}^{\infty} u_A(\tau)u_B(t+\tau)d\tau \\ &= \int_{-\infty}^{\infty}[-G_{AB}(\tau)+G_{AB}(-\tau)]d\tau. \end{aligned} \quad (1)$$

where u_A and u_B are the prepared signals at locations A and B (Bensen et al., 2007a). Theoretical studies have shown that if the noise wavefield is sufficiently diffusive, the cross-correlation between two stations can be used to approximate the Green's function G_{AB} between the two sensors or locations (Lobkis and Weaver, 2001; Snieder, 2004). Based on the noise cross-correlations, the period-dependent surface wave phase and group travel time can be determined between each pair of sensors.

Focusing on using the Rayleigh wave information, we only calculate the cross-correlations between the vertical components of a 3-component geophone. We stack the cross-correlation result every t time until complete T. Usually, the stacking/averaging enhances the signal-to-noise ratio (SNR) and also effectively suppresses the source distribution inhomogeneity. An example of stacked cross-correlation can be seen in Fig. 3.

3.3 Collaborative Imaging

After stacking cross-correlation and the time T is completed, sensors begin the collaborative imaging stage. First, they perform a Frequency-Time Analysis (FTAN) of the stacked cross-correlations. FTAN generates the dispersion curve of the Rayleigh wave phase velocity (Bensen et al., 2007b). A whole FTAN process includes a series of Gaussian band-pass filters to get Green's function with different central frequencies and transformation processes to get the envelope function and phase function of time series data. For each frequency, we can obtain the estimated travel time between sensors using FTAN.

The eikonal and Helmholtz tomography methods is adopted to determine 2D phase velocity maps based on empirical wavefield tracking (Lin et al., 2009a; Lin and Ritzwoller, 2011). For each event i, it measures surface wave phase velocities at each location directly by the spatial derivatives of the observed wavefield:

$$\frac{1}{c_i^2(\mathbf{r})} = |\nabla\tau(\mathbf{r}_i,\mathbf{r})|^2 - \frac{\nabla^2 A_i(\mathbf{r})}{A_i(\mathbf{r})\omega^2}, \quad (2)$$

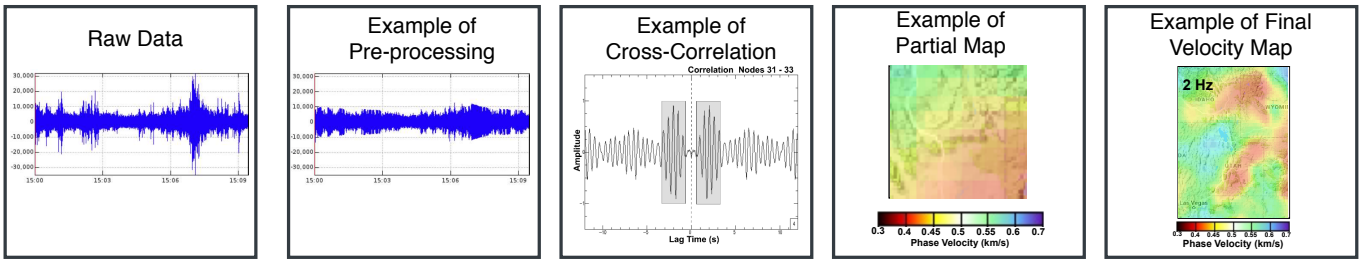


Figure 3 Examples of output of each stage in ANION. From left to right: (i) raw data sensed by smart sensor; (ii) example of data pre-processing with normalization. No down-sampling nor compression was applied in this figure for readers' understanding of ambient noise preservation; (iii) example of cross-correlation. Grey area represents group velocity arrival of the wave signal; (iv) example of partial velocity constructed on one sensor; (v) example of final velocity map of the interested area.

where τ and A represent phase travel time and amplitude measurements, and $\widehat{\mathbf{k}}_i \cong \nabla\tau(\mathbf{r}_i, \mathbf{r})c_i(\mathbf{r})$, c and ω are direction of wave propagation, phase velocity and angular frequency, respectively. $\widehat{\mathbf{k}}_i$ can be derived directly by solving 2D Helmholtz wave equation also called eikonal equation, can be derived from equation 2 under infinite frequency approximation. While the above equations are defined for 'events,' it's important to note that the cross-correlation method from equation 1 effectively turns each sensor into an 'event' recorded at every other sensor, and so the wavefield from virtual sources at each sensor, as well as the spatial derivatives in equations 2 can be approximated from the set of cross-correlations with that sensor. This generates a 2D partial velocity map at each sensor.

To aggregate the all partial velocity maps into the final subsurface image, sensors form a tree structure starting from a root node. The root node is configured in the configuration file. The tree is constructed as a Breadth-First Search (BFS) and depending on the tree formed, the nodes leaf, intermediate and root aggregate the maps. The final map is generated in the root node after aggregation. The amount of data transmitted for aggregating the final velocity maps is significantly less than transmit raw data in the network. An example of a final velocity map can be seen in Fig. 3.

3.4 Algorithm

We adapt the ANION methodology to compute the velocity map in-situ, that is generating the velocity map without sending the raw data to a central place (Valero et al., 2017). We integrated all steps of the ANION methodology in a system that allows automatic subsurface imaging based on surface wave. The system is an array of sensors that collect raw data from the medium, performs in-situ signal processing, and generates the subsurface imaging result (velocity map) by using collaborative inter-nodes communication and aggregation. In the end, we are able to distinguish velocities differences that may indicate different elements and/or activities under the subsurface by analyzing the subsurface velocity map.

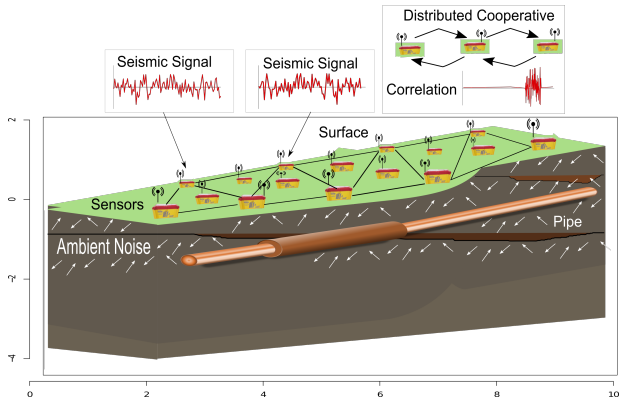


Figure 4 Sensor network performing ANION techniques to generate a velocity map for pipeline detection.

The ANION system sketch is shown in Fig. 4. A mesh network is formed and it (i) uses seismic sensors to measure the vibration of the ambient noise; (ii) calculates the cross-correlation of the signal waves with neighbors and performing a frequency-time analysis to obtain phase traveltime measurements of the ambient noise signal; and (iii) employs a method known as eikonal tomography (Lin et al., 2009b) to build velocity maps. Algorithm 1 describes the distributed subsurface imaging process based on the surface wave, which is executed by every sensor in the network.

3.4.1 Input description

The input of Algorithm 1 is a pre-processed sensor data after data preparation. Every node, after gathering raw data, applies an in-situ pre-processing to prepare the data to a suitable form for cross-correlation (Bensen et al., 2007b). This preparation was explained in detail in Section 3.1. Then, a compression process is required to make the signal suitable to be transmitted to other nodes. The configuration file is used to know the parameters of the system; for instance, frequency of interest and size of the grid for the final map.

Algorithm 1: Distributed Subsurface Imaging based on Surface Wave

Input: Pre-processed data $\hat{v}_0$ and configuration file
Output: Collaborative velocity map (M).

```
1  repeat
2      Broadcast v̂0 to neighbors |η| ;
3      Received from i ∈ |η| the pre-processed data v̂i;
4      Apply cross-correlation (cc0i) between v̂0 and
         v̂i ∀i ∈ |η| ;
5      if t%T == 0 then
6          Apply Frequency-time Analysis to all cc0i
             where i ∈ |η| ;
7          Obtain traveltime measurements τ̂0i ;
8          Interpolate τ̂0i onto a grid R of size x × y ;
9          Calculate slowness vector ŝ for each point in R
             ;
10         Form a tree structure with other nodes in the
             network ;
11         if the node is leaf then
12             Send via TCP ŝ to its parent
13         end
14         if the node is intermediate node then
15             Receive via TCP ŝj from children nodes j;
16             Aggregate ŝj from all children nodes j with
                 ŝ and get new ŝ;
17             Send via TCP ŝ to its parent
18         end
19         if the node is root then
20             Receive via TCP ŝj from children nodes j;
21             Aggregate ŝj from all children nodes j with
                 ŝ and get new ŝ;
22             Calculate M = 1/ŝ;
23             Output M
24         end
25     end
26 until total time of experiment is completed;
```

3.4.2 In-situ signal processing

The compressed signal is transmitted to neighbor nodes via UDP communication using broadcasting (line 2). Once every node receives their neighbors pre-processed data, it performs the cross-correlation between its own signal and its neighbors signal (line 4). With the cross-correlation results, the node performs frequency-time analysis (line 6), which allows the traveltime measurements calculation between itself and its neighbors (line 7). With the traveltime measurements, the node is able to construct a partial view of the subsurface by calculating and interpolating the slowness (1/velocity) vector respecting to its neighbors (line 9).

3.4.3 Collaborative Imaging

Lines from 10 to 23 of algorithm 1 describe the collaborative process of the subsurface imaging. We borrow the idea of data aggregation process in ambient noise from Valero et al. (2018). A tree structure if forming between all nodes inside the network with a pre-defined root. The root can be selected by the user in the configuration file of the system. The tree structure enables aggregation of the partial maps (slowness vectors) in a bottom-top fashion. Depending on the node role in the tree, the algorithm shows the node procedure. If the node is a leaf (line 11), it just sends its partial map to its parent. If the node is an intermediate node (line 14), it aggregates the partial maps of its children with itself, and then sends the aggregated map to its parent. If the node is the root (line 19), it waits for all its children partial maps, performs the aggregation with itself, and then output the final map. The system manages a MAX_TIME to define the maximum time of waiting for a node child response, after which the waiting is discarded assuming the child node is down.

3.4.4 Output description

The final map, calculated in line 23, is the collaborative velocity map which represents the subsurface image. The system then plots the image for experts interpretations.

4 System Setup

In this section, we describe the main hardware and software components of ANION to resolve subsurface imaging. We implemented and deployed ANION system on real devices and perform experimental tests on real field scenarios.

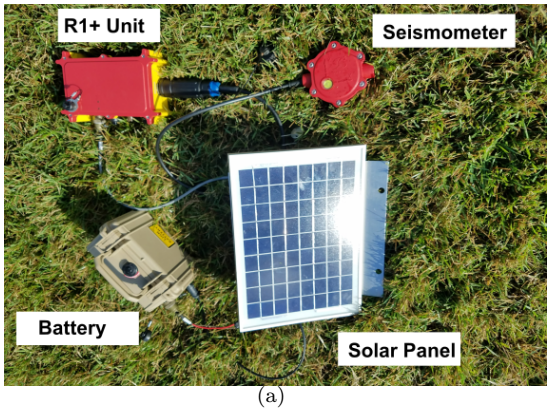
4.1 Hardware

Every node or sensor in the network has a global positioning system (GPS), three channel/component seismometer (geophone), a Raspberry Pi 3 board, a battery and a solar panel as shown in Fig. 5. Some hardware components are housed into a waterproof box called R1+ for protecting them from the harsh environment. The low-power GPS interface provides the geo-location of the sensor node and a time-stamp is used for the system to collect, synchronize and process the seismic data. The three channels geophone is incorporated into the system to detect the velocity of ground movements. Each channel records its own data with respect to its axis N, E, and Z or directions North, East and Depth (vertical). The single board computer (Raspberry Pi) is the core of the system because is in charge of collecting and storing data, processing data analytics, communicating with other units and providing raw and processed information to a visualization tool. We also integrate a waterproof battery 11V and 99.9 Wh. The battery is connected to a 10 Watt solar panel for giving to the system renewable energy.

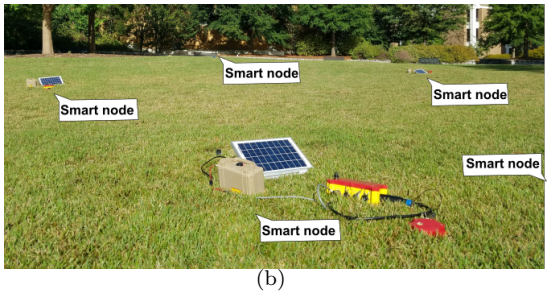
The detailed specifications of the main single-board computer inside R1+ are presented in Table 1

4.2 Software

Our system relies on five main software modules shown in Fig. 6, and we define them as follows: (i) the *acquisition*



(a)



(b)

Figure 5 (a) Node hardware details. (b) Deployment.

Table 1 Single-board computer specifications

Raspberry Pi 3 Model B	
CPU	1.2GHz 64-bit quad-core ARMv8
Memory	1 GB SDRAM
USB 2.0 ports	4 (via the on-board 5-port USB hub)
On-board storage	32 Gb Micro SDHC
On-board network	10/100 Mbit/s Ethernet, 802.11n wireless, Bluetooth 4.1

module is responsible for gathering the raw data from the medium. We use C++ language for developing the functions for reading geophone measurements and processing data. We also have a MySQL database inside each node to store the data for future use if needed. The *acquisition module* also helps with the synchronization of each node; (ii) the *signal processing module* performs the pre-treatment of the signal, the cross-correlation between signals, and the frequency-time analysis to obtain phase traveltime measurements. This module was developed using C++; (iii) the *communication module* was developed mainly in C++ and is responsible to communicate and cooperate with neighbor nodes. We use User Datagram Protocol (UDP) to broadcast the information and Transmission Control Protocol (TCP) to send data to neighbors in order to aggregate the final map; (iv) the *map generation module* is responsible for generating the final velocity map. C++ was used for developing this module; (v) finally, for the *visualization module*, we developed a Web front-end in PHP and JavaScript to enable users to monitor in real-time what is happening inside each node and the final map.

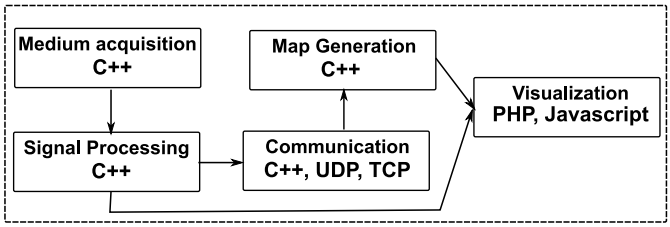


Figure 6 Programming languages used on software modules.

5 Experiments and Evaluation

For evaluation purposes, we deployed our system using 13 sensor nodes and performed a field test. Our test was made on April 10th, 2018. The deployment location was an open space in University XXXcampus between geography and chemistry buildings. We chose this location because this area is known for having underground pipelines. Our goal was to measure the correct functionality of the system and to investigate whether we could generate a velocity map that illustrates the differences in velocity between the ground and the pipelines. Fig. 7 shows the location of the sensors in this experiment.

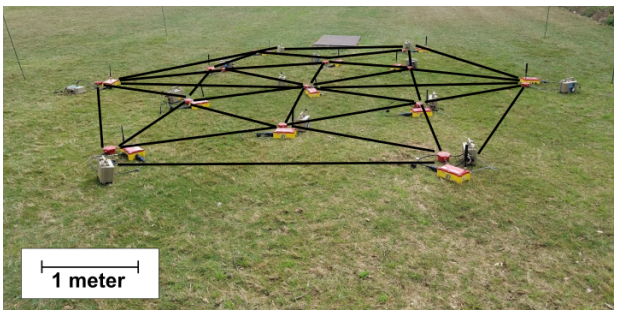


Figure 7 Deployment and sensor location in the field test.

The system ran for five hours, and it was configured to (i) perform cross-correlation between neighbor nodes every five minutes and stack the results, (ii) run collaborative imaging for generating the velocity map every five (5) hours and (iii) utilize an appropriate frequency range for detecting small object, in this case, the pipe. All these parameters are setup before the deployment in a configuration file.

First, we examined the ambient noise data recorded by each node. Fig. 8 shows an example of the data gathering by the geophone in sensor number 1. Notice that even though we are using the vertical component of the geophone (Z direction), we can also monitor the other two components (East direction and North direction).

Once we ensure the acquisition module functionality, we checked the signal processing module. This was done by looking into the cross-correlation and Green functions (equation 1) of the signal between each pair of nodes. Each node cross-correlate its signal with its neighbor signal every t time and stacks them together. As said,

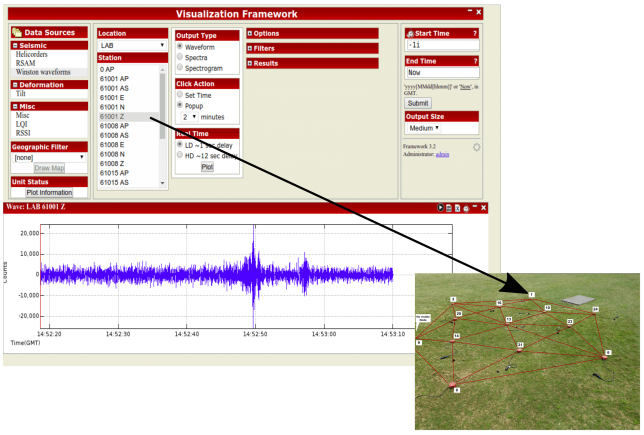


Figure 8 Raw data monitoring example from sensor node 1 in our deployment.

this procedure enhances the SNR. In this case, we configure t to be 5 minutes. Then, the system correlates every 5 minutes of data and stack it for a long period of 5 hours. Fig. 9 shows an example of symmetry cross-correlation we obtain between sensors 3 and 6.

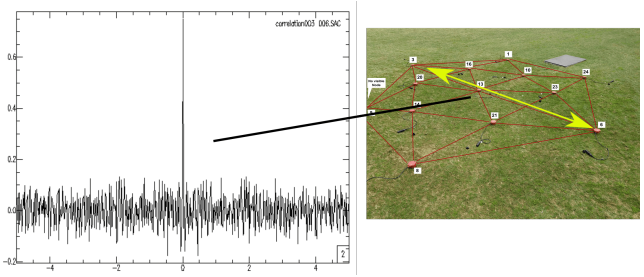


Figure 9 Symmetric cross-correlation result from nodes 3 and 6 in our deployment.

Every $T = 5$ hours, every node performs the FTAN procedure. In this procedure, each cross-correlation measure is analyzed to determine the phase traveltime (also known as delay time). When the traveltime is calculated, we are able to generate the slowness vector $\hat{s}$. Each node generates a "node partial map". Examples of node partial maps are shown in Fig. 10. Green and black diamonds represent the location of the sensors in the field; the green diamond is the node that is plotting the partial map. Axis x and y represent the size of the interpolated grid in the ambient noise seismic process described in algorithm 1, line 8.

These maps have significant errors because they are just a partial view of the subsurface. For example, in Fig. 10(a), node 1 is located at the upper-right part, and it seems to do not have a complete vision of the subsurface, which is predictable since the node is located in a corner. The same happens with node 6 and node 8 (Fig. 10(b) and (c)). Fig. 10(d) shows the partial map of node 16 that is located near to the central node. However, there are still irregularities in the partial map. These irregularities are due to in partial map

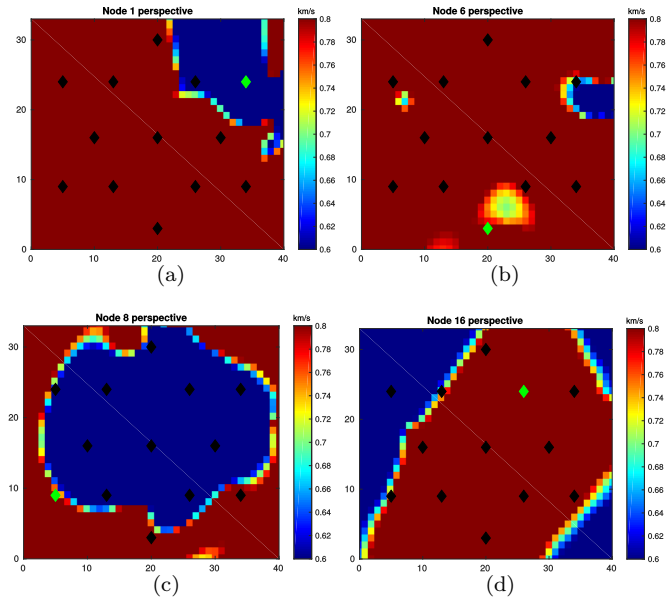


Figure 10 Partial maps from (a) Node 1, (b) Node 6, (c) Node 8, and (d) Node 16.

the correct frequency range is still not set up, and the different spectral components may interfere with each other. The system performs a collaborative aggregation process to generate the final map and mitigate errors through validations and frequency selection.

The final velocity maps after collaborative aggregation for our field experiment is shown in Fig. 11(a) and Fig. 11(b). The central frequency used was 125 Hz. By using a high frequency, we ensure to evaluate the shallow surface(Artman, 2006) since the wavelength of the seismic wave decreases with increasing frequency. We can observe from Fig. 11(a) that when we set up a velocity range between 500 m/s and 800 m/s the map looks diffuse but we can still perceive a difference in the center of the map.

To obtain a clearer image of the shallow subsurface, we modify the display range to 740 m/s and 756 m/s to further increase the contrast in the velocity map. We can observe that the surrounding velocities are all below 740 m/s, but the central structure is stable along with a horizontal line and between 752 m/s and 756 m/s. The area with high velocity in the map indicates that it should be an isolated structure/facility, corresponding to the wanted pipeline we are aiming to detect under our deployment. This result shows we are able to see structures under the subsurface and detecting some security issues (for example, a broken pipeline).

5.1 Results Discussion

For instrument limitations, we chose 125 Hz as the dominant frequency in our experiments. The sampling rate of our sensors is 500 Hz. Based on the Nyquist-Shannon sampling theorem, only the first 250 Hz are usable. Furthermore, to avoid aliasing effect(**?**), we adopted up to 125 Hz frequencies. From Figs. 11(a)

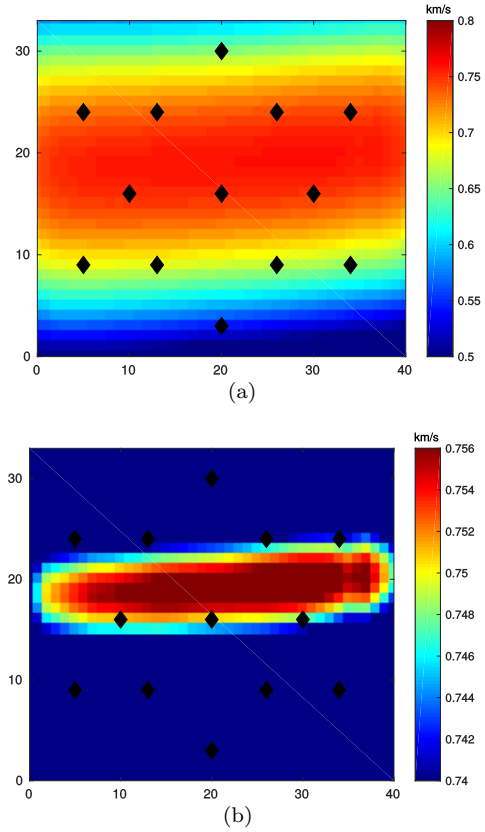


Figure 11 (a) Velocity map with a velocity range of 500 m/s and 800 m/s. (b) Final velocity map with velocity range between 740 m/s and 756 m/s.

and 11(b), in our application, the shallow subsurface velocity is around 750 m/s. Considering a central frequency of 125 Hz, the wavelength Λ ($\Lambda = c/\omega$, where c is velocity and ω is frequency) will be about 6 m/s. Then, the seismic resolution is calculated by $\Lambda/4$, resulting in our resolution being about 1.5 m, which is not optimal for a pipeline detection, whose diameter is about 20 cm. This is the reason why the pipeline image looks thick in Fig. 11(b). In addition, according to depth sensitivity kernel theory (Chen and Lee, 2015), the maximum depth for a frequency of 125 Hz is approximately 5-7 m, which differs at different locations with various geological conditions. Thus, the velocity map we generate is an average map between depth 0 and 5 m. Since the pipe we want to detect is located 1 meter under the ground, it should be detected in the imaging result. However, if the pipeline is not the only underground facility in this depth range, our result may be degraded. The solution to improve the resolution and image shallower subsurface is to increase the sampling frequency, which is the reason why ground-penetrating radar (GPR) can do its job.

To sum up, the application results show that our ANION system can detect thick pipelines with a relatively acceptable resolution when the underground network is not complicated. And, it may present better results for underground tunnel detection, as the tunnel usually has a larger diameter and within our depth sensitivity range. We envision to extend this system to other potential applications explained in section 2. We also expect to continue improving our algorithms to obtain higher resolution images. Potentially, we can include more algorithms for detecting pipe leakages and other security problems. In addition, we plan to incorporate shear-wave inversion algorithms to obtain a 3D velocity map for further interpretations in detail.

5.2 Bandwidth and Communication Cost Evaluation

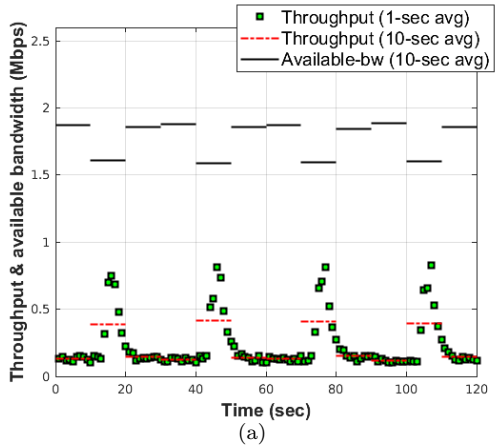
We conducted an evaluation in terms of bandwidth/throughput utilization and communication cost to highlight the benefits of our distributed approach with a centralized approach in which all nodes send the raw data to a central place (central coordinator). Manual data collection is not considered for fair comparison.

First, we calculate the throughput of the network and the available bandwidth of both approaches. The local available bandwidth of a sensor node is measured locally at the MAC (media-access-control) level: the percentage of time channel is sensed idle and the node is not back-off state, multiplied by the channel capacity. In a mesh network, every "*hop*" (link) between sensors will decrease the bandwidth by half (Jain and Dovrolis, 2003). This happens because wireless links can only do one thing at a time - transmit or receive. Our instruments are based on a Raspberry Pi 3 as computer board. The wireless communication bandwidth of Raspberry Pi 3 is estimated at ∼10Mbps (Megabytes per second) (Upton, 2016). Due to the number of links in our topology (some nodes may have 5 or 6 links, which reduced the available bandwidth), we based our observations on a maximum available bandwidth of ∼2Mbps.

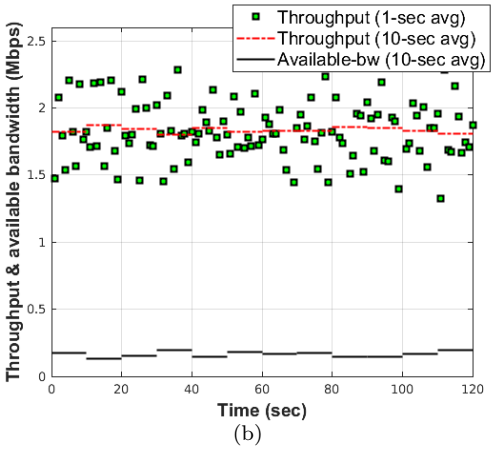
Fig. 12 shows the throughput and available bandwidth for (a) distributed ANION presented in this paper, and (b) a centralized approach. This throughput was recorded for 120 seconds in which nodes in the distributed approach exchange information with the neighbors every $t = 20s$ to perform cross-correlation later. From Fig. 12(b), it is possible to note that the average available bandwidth is very low since all nodes send information to a central place. In contrast, ANION improves the bandwidth utilization due to in-situ processing and minimal communication every 20 seconds. Our approach meets the bandwidth limitations, and the sent packages are small due to data preparation and compression.

Second, we conducted a study of the communication cost of ANION compared to a centralized approach. To analyze the communication cost, we measured the cost in terms of the number of received packages on each sensor in the network. We calculated the received package after 1 hour of cross-correlation in the network.

Fig. 13 shows a 5×5 grid of sensor nodes and the color-bar denotes the number of messages received by each node. We can see that the communication is balanced in ANION and each node receives roughly

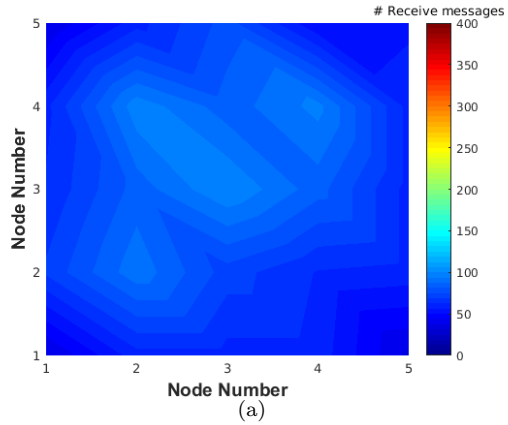


(a)

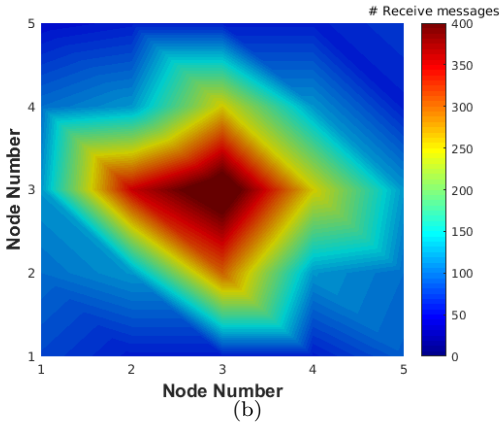


(b)

Figure 12 Throughput and available bandwidth analysis. (a) Distributed ANION. (b) Centralized approach where all nodes send continuously data to a central coordinator.



(a)



(b)

Figure 13 Communication cost in terms of number of messages received by each node. (a) Distributed ANION. (b) Centralized approach where all nodes send continuously data to a central coordinator.

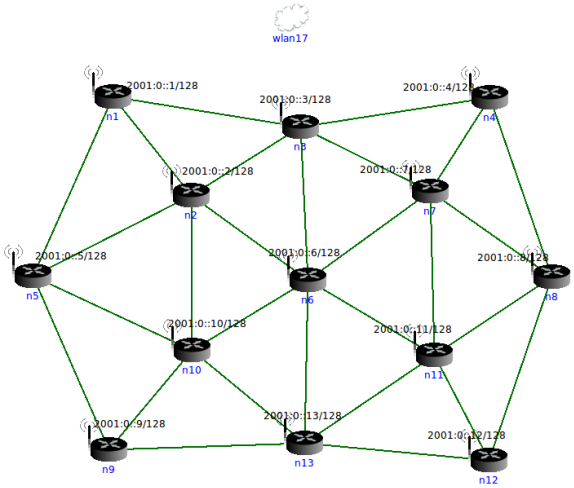
around 100 messages. Whereas in the centralized approach since we use center node as a collector of information, we observe an imbalance communication pattern. There is an increased message overhead near the center node roughly accumulating to 400 messages. Building the infrastructure with a special coordinator, routing protocols and synchronization adds extra overhead and is not suitable for decentralized systems such as mesh sensor network. This also has an impact in the energy consumption of each node. According to Pottie and Kaiser (2000), the energy of transmitting 1KB a distance of 100m is approximately the same that executing 3 million of instruction by processor. Hence, local data processing is crucial for also saving sensors energy. This implies that our approach besides reducing communication cost, it also helps to avoid extra energy utilization.

5.3 Link reliability and node disruption analysis

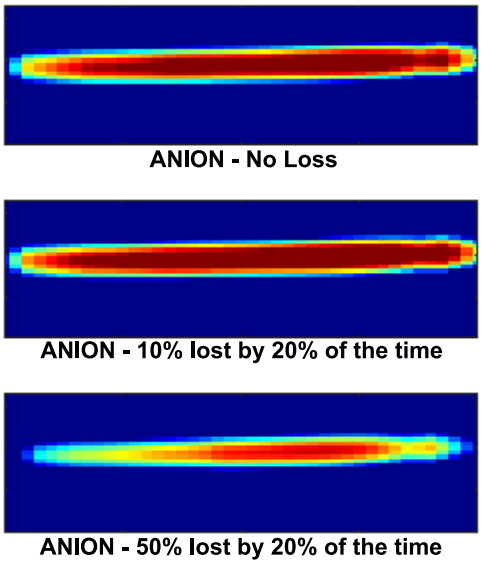
The system behavior under link reliability and node disruption is analyzed in this section. With the data collected for the nodes in the real deployment, we run an experiment on CORE network emulator (Ahrenholz et al., 2011). Every node emulates the real device we deployed on the field by reading from an internal database we previously filled with the information we collected in the field. For testing link reliability, we configured the node links in the emulator to loss 20% and 50% of the packages 20% of the time. Because we are stacking the cross-correlations that are calculated on each node for a long period T, even if some packages are lost, the final result will no be significantly affected. However, if the stacking time T is too short, the final velocity maps could present important inconsistent. In the collaborative imaging step, nodes aggregate information. If some nodes fail or packages are lost too frequently, the final result can be highly impacted. However, we designed a resilient mechanism that helps to diminish the impact of package loss in aggregation.

The system has been designed to restart automatically the operations after failure. In this scheme, after a sensor is automatically started with a system service, and it has been synchronized with the other sensors via its GPS, the sensor checks if there a velocity calculation and imaging has been done during the time it was down. This is done by checking is the current time is greater than the time the process supposes to be performed. If this happens,

the node sends a request to others for recalculation of the velocities and interpolation. The cross-correlation process also starts in any case. With this scheme, we introduce resilience to the system, and we aim to guarantee that the results will be computed with the maximum numbers of available sensors.



(a)



(b)

Figure 14 Link reliability analysis (a) CORE emulator scenario with data from real deployment. (b) Pipeline detection result from: (top) ANION with no loss; (middle) ANION with 10% of package loss by 20% of the time; (bottom) ANION with 50% of package loss by 20% of the time.

Fig. 14(a) shows CORE emulator with the same node topology we used in field. Fig. 14(b) illustrates the final pipeline detection through velocity map for the three scenarios of package loss. Note that, when the package loss is severe, the final velocity map is impacted even with our recovery scheme. However, the detection of the infrastructure is still possible.

In ANION method, although we are able to distribute the computation part, the communication still relies on a root node to complete the aggregation of the velocity map. We are currently investigating methods to perform aggregation using gossip methods (Boyd et al., 2006) that also involves information exchange only with neighbors. Since the variables (slowness/velocity) need to be aggregated and averaged, we intend to use gossip methods that converge to average consensus (Boyd et al., 2006). In this way, we can achieve a decentralized way to compute aggregation that besides to meet bandwidth constraints and reduce communication cost, it will be capable to be resilient to link and node disruptions.

6 Conclusion

In this paper, we presented a seismic sensor system for imaging subsurface infrastructures that can be used in wide variety of applications. Our approach is based on ambient noise seismic imaging, a geophysics technique that uses the cross-correlation of the ambient noise measurements of the earth to calculate the velocity under the surface. This technique is non-invasive, economical and no need of active energy sources. We leverage the capabilities of current sensor networks to provide a system that can perform in-situ signal processing and inter-nodes communication and cooperation to generate real-time velocity maps for security analysis. Our field tests demonstrate the system is able to calculate velocity maps that illustrate the differences in speed between underground structures. We envision that ambient noise techniques in real time can be treated as a new source that is economical, practical, and particularly valuable for seismic hazard mitigation, and anomalous activity detection in urban and non-urban areas. Our bandwidth and communication cost analysis indicate that ANION is suitable for distributed environments since it reduces communication while meets bandwidth constraints.

References

Ahrenholz, J., T. Goff, and B. Adamson (2011). Integration of the CORE and EMANE Network Emulators. pp. 1870–1875.

Almeida, F. C. L., M. J. Brennan, P. F. Joseph, S. Dray, S. Whitfield, and A. T. Paschoalini (2015). Towards an in-situ measurement of wave velocity in buried plastic water distribution pipes for the purposes of leak location. *Journal of Sound and Vibration 359*, 40–55.

Artman, B. (2006). Imaging passive seismic data. *Geophysics 71*(4), SI177–SI187.

Barmin, M. P., M. H. Ritzwoller, and A. L. Levshin (2001). A Fast and Reliable Method for Surface Wave

Tomography. *Pure and Applied Geophysics 158*, 1351–1375.

Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007a, June). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International 169*(3), 1239–1260.

Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007b). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International 169*(3), 1239–1260.

Boyd, S., A. Ghosh, B. Prabhakar, and D. Shah (2006). Randomized gossip algorithms. *IEEE transactions on information theory 52*(6), 2508–2530.

Brenguier, F., N. M. Shapiro, M. Campillo, V. Ferrazzini, Z. Duputel, O. Coutant, and A. Nercessian (2008, January). Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience 1*(2), 126–130.

Brenguier, F., N. M. Shapiro, M. Campillo, A. Nercessian, and V. Ferrazzini (2007). 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlations. *Geophysical research letters 34*(2).

Chen, P. and E.-J. Lee (2015). *Full-3D Seismic Waveform Inversion: Theory, Software and Practice*. Springer.

de Ridder, S. and J. Dellinger (2011). Ambient seismic noise eikonal tomography for near-surface imaging at Valhall. *The Leading Edge 30*(5), 506–512.

Duputel, Z., V. Ferrazzini, F. Brenguier, N. M. Shapiro, M. Campillo, and A. Nercessian (2009, July). Real time monitoring of relative velocity changes using ambient seismic noise at the Piton de la Fournaise volcano (La Réunion) from January 2006 to June 2007. *Journal of Volcanology and Geothermal Research 184*(1-2), 164–173.

Jain, M. and C. Dovrolis (2003). End-to-end available bandwidth: measurement methodology, dynamics, and relation with TCP throughput. *IEEE/ACM Transactions on Networking (TON) 11*(4), 537–549.

Kamath, G., P. Ramanan, and W.-Z. Song (2015). Distributed Randomized Kaczmarz and Applications to Seismic Imaging in Sensor Network. In *The 11th IEEE International Conference on Distributed Computing in Sensor Systems (IEEE DCOSS), Fortaleza, Brazil*.

Kamath, G., L. Shi, E. Chow, and W.-Z. Song (2015). Distributed Tomography with Adaptive Mesh Refinement in Sensor Networks. *International Journal of Sensor Network*.

Kamath, G., L. Shi, W.-Z. Song, and J. M. Lees (2016). Distributed Travel-time Seismic Tomography in Large-Scale Sensor Networks. *Journal of Parallel and Distributed Computing 89*.

Lin, F. and M. H. Ritzwoller (2011). Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure. *Geophysical Journal International 186*, 1104–1120.

Lin, F., M. H. Ritzwoller, and R. Snieder (2009a, June). Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array. *Geophysical Journal International 177*(3), 1091–1110.

Lin, F.-C., M. P. Moschetti, and M. H. Ritzwoller (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. *Geophysical Journal International 173*(1), 281–298.

Lin, F.-C., M. H. Ritzwoller, and R. Snieder (2009b). Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array. *Geophysical Journal International 177*(3), 1091–1110.

Liu, L., Q.-f. Chen, W. Wang, and E. Rohrbach (2014). Ambient noise as the new source for urban engineering seismology and earthquake engineering: a case study from Beijing metropolitan area. *Earthquake Science 27*(1), 89–100.

Lobkis, O. I. and R. L. Weaver (2001). On the emergence of the Green's function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America 110*(6), 3011.

Mordret, A., M. Landès, N. M. Shapiro, S. C. Singh, P. Roux, and O. I. Barkved (2013). Near-surface study at the Valhall oil field from ambient noise surface wave tomography. *Geophysical Journal International 193*(3), 1627–1643.

Mordret, A., H. Sun, G. A. Prieto, M. N. Toksöz, and O. Büyüköztürk (2017). Continuous Monitoring of High-Rise Buildings Using Seismic Interferometry. *Bulletin of the Seismological Society of America 107*(6), 2759–2773.

Moschetti, M. P., M. H. Ritzwoller, F. Lin, and Y. Yang (2010, October). Crustal shear wave velocity structure of the western United States inferred from ambient seismic noise and earthquake data. *Journal of Geophysical Research 115*(B10).

Nicolson, H., A. Curtis, B. Baptie, and E. Galetti (2012). Seismic interferometry and ambient noise tomography in the British Isles. *Proceedings of the Geologists' Association 123*(1), 74–86.

Olivier, G., F. Brenguier, M. Campillo, R. Lynch, and P. Roux (2015). Body-wave reconstruction from ambient seismic noise correlations in an underground mine. *Geophysics 80*(3), KS11–KS25.

Pottie, G. J. and W. J. Kaiser (2000). Wireless integrated network sensors. *Communications of the ACM 43*(5), 51–58.

Ritzwoller, M. H., F.-C. Lin, and W. Shen (2011). Ambient noise tomography with a large seismic array. *Comptes Rendus Geoscience 343*(8), 558–570.

Shi, L., W.-Z. Song, M. Xu, Q. Xiao, J. M. Lee, and G. Xing (2013). Imaging Seismic Tomography in Sensor Network. In *IEEE SECON*.

Snieder, R. (2004, April). Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E 69*(4), 046610.

Song, W., F. Li, M. Valero, and L. Zhao (2019). Toward Creating Subsurface Camera. *Sensors 19*(2), 301.

Stork, A. L., C. Allmark, A. Curtis, J.-M. Kendall, and D. J. White (2018). Assessing the potential to use repeated ambient noise seismic tomography to detect CO 2 leaks: Application to the Aquistore storage site. *International Journal of Greenhouse Gas Control 71*, 20–35.

Tomar, G., E. Stutzmann, A. Mordret, J.-P. Montagner, S. C. Singh, and N. M. Shapiro (2018). Joint inversion of the first overtone and fundamental mode for deep imaging at the Valhall oil field using ambient noise. *Geophysical Journal International 214*(1), 122–132.

Upton, E. (2016). Raspberry Pi 3. *URL https://www.raspberrypi.org/products/raspberry-pi-3-model-b*.

Valero, M., G. Kamath, J. Clemente, F.-C. Lin, Y. Xie, and W. Song (2017). Real-time Ambient Noise Subsurface Imaging in Distributed Sensor Networks. In *The 3rd IEEE International Conference on Smart Computing (SMARTCOMP 2017)*.

Valero, M., F. Li, S. Wang, F.-C. Lin, and W. Song (2018). Real-time Cooperative Analytics for Ambient Noise Tomography in Sensor Networks. *IEEE Transactions on Signal and Information Processing over Networks*.

Zhao, L., W.-Z. Song, L. Shi, and X. Ye (2015). Decentralized Seismic Tomography Computing In Cyber-Physical Sensor Systems. *Cyber-Physical Systems, Taylor and Francis*.