

STABILITY AND SCALABILITY ASSESSMENT OF COMPLEXITY ESTIMATION BASED WORK LOAD BALANCING APPROACH FOR A PARALLEL LIDAR WAVEFORM DECOMPOSITION ALGORITHM

¹⁾ **Jinha Jung, Melba M. Crawford, ²⁾ Sanghoon Lee**

¹⁾ Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN 47906
(jinha@purdue.edu and mcrawford@purdue.edu)

²⁾ Department of Industrial Engineering, Kyungwon University, Seongnam-shi, Kyunggi-do, Korea
(shl@mail.kyungwon.ac.kr)

ABSTRACT

LIDAR is an active remote sensing technology which performs range measurements from the sensor and converts them into 3D coordinates of the Earth's surface. Recent advances in LIDAR hardware make it possible to digitize full waveforms of the returned energy. LIDAR waveform decomposition involves separating the return waveform into a mixture of Gaussians which is then used to characterize the original data. It plays an important role in LIDAR data processing because the resulting components are expected to represent reflection surfaces within waveform footprints and ultimately affect the interpretation of the data. Computational requirements in the waveform decomposition process result from two factors; (1) estimation of the number of components in a mixture and the resulting parameter estimates are inter-related and cannot be solved separately, (2) parameter optimization does not have a closed form solution, and thus needs to be solved iteratively. A current state-of-the-art airborne LIDAR system acquires more than 50,000 waveforms per second, and the number of waveforms easily exceeds tens of millions even for small area. Therefore, decomposing the enormous number of waveforms is challenging using traditional single processor architecture. Four work load balancing approaches – (1) a no weighting (NW), (2) a linear weighting based on the decomposition results (DRLW), (3) a squared weighting based on the decomposition results (DRSW), and (4) a linear weighting based on the decomposition time (DTLW) of sampled waveforms - for a parallel LIDAR waveform decomposition were assessed in terms of the scalability and stability. The DTLW approach yielded the best efficiency when the number of processors is small, and the NW approach showed the most scalable and stable results as the number of processors gets larger.

INTRODUCTION

LIDAR (LIght Detection And Ranging) is an active remote sensing technique which provides 3D coordinates of the Earth's surface by performing range measurements from the sensor. Hardware limitation prohibited early LIDAR systems from recording the continuous back-scattered energy, and resulted in recording only multiple discrete returns by filtering the return signal. These discrete returns are combined with the location and the attitude of the sensor to generate 3D coordinates of the Earth's surface by simple vector calculations. However, recent advances in hardware design now make it possible to record high volume of data in short period of time, and enable a full waveform LIDAR system which digitizes the continuous return signal, which is called a waveform. The full waveform LIDAR system has recently attracted attention of researchers because it contains more information than traditional discrete returns LIDAR system. Most discrete return LIDAR systems not only use proprietary algorithm to detect peaks so the end-user has no way to assess the quality of the results, but also limits number of returns (usually from 2 to 5) from a waveform. However, full waveform LIDAR data provide the end-user raw data to extract more accurate and meaningful information. Researchers (Reitberger et al., 2008) reported that a much higher point density was achieved by decomposing waveforms than conventional discrete return LIDAR system and higher classification accuracy was achieved. Other researchers (Duong et al., 2008) reported better extraction of canopy and ground elevations using the ICESat waveforms even in heavily forested areas.

A LIDAR waveform can be modeled as the convolution of the outgoing signal and the vertical structure within the waveform footprints. LIDAR waveform decomposition refers to the process of decomposing a return waveform into a mixture of components which are then used to characterize the original waveform data. It plays an important role in LIDAR waveform processing because the resulting decomposed components are assumed to represent reflection surfaces within waveform footprints. Hence the decomposition results ultimately affect the interpretation of LIDAR waveform data. The most common statistical mixture model used for the process is the Gaussian mixture, whose parameters include mixing coefficients and the mean and standard deviation of each component. Various researchers utilized a Gaussian mixture model to decompose LIDAR waveform into components by utilizing various optimization techniques such as a Gauss-Newton, a Levenberg-Marquardt, and the EM (Expectation-Maximization) algorithms (Chauve et al., 2007; Hofton et al., 2000; Jung and Crawford, 2008; Persson et al., 2005). Waveform decomposition is an unsupervised machine learning problem, and a

computationally intensive process. Computational requirements in the waveform decomposition process result from two factors; (1) estimation of the number of components in a mixture and the resulting parameter estimates are inter-related and cannot be solved separately, and (2) the parameter optimization problem does not have a closed form solution, and thus needs to be solved iteratively. The current state-of-the-art airborne LIDAR system acquires more than 50,000 waveforms per second (Mallet et al, 2009), so decomposing the enormous number of waveforms is challenging using traditional single processor architecture. Furthermore, there may be a situation in which LIDAR waveform data need to be processed near real-time. A parallel LIDAR waveform decomposition algorithm with four different work load balancing approaches - (1) no weighting (NW), (2) a decomposition results based linear weighting (DRLW), (3) a decomposition results based squared weighting (DRSW), and (4) a decomposition time based linear weighting (DTLW) – were developed and tested using various number of processors (from 8 to 256) in the previous study (Jung et al., 2009). However, the scalability and stability of each work load balancing approach have not been studied yet. The goal of this study is to assess the scalability and stability of those approaches using larger number of processors (up to 1,024) with different processor selection scheme from each node such as 1 processor per node, 4 processors per node, and 8 processors per node.

DATA AND COMPUTATIONAL PLATFORM

Full Waveform Data

The Freeman Ranch is a research site located near San Marcos, TX (USA) and managed by Texas State University. It contains a mixture of rangeland and woodlands. Topography is primarily low hills divided by small creeks, except with steep slopes along drainage channels. An Optech ALTM (Airborne Laser Terrain Mapper) 1225 small footprint LIDAR system with a full waveform digitizer, which is owned and managed by the University of Texas at Austin (UT), was flown over Freeman Ranch on 12 August 2005. The UT LIDAR laser system operates at 1064 nm with a pulse rate of 25 kHz. Its waveform sampling rate is 1 ns, which corresponds roughly to 15 cm in the vertical dimension. Five flight lines were acquired at an altitude of 650 - 720 m above ground level with a resulting footprint diameter of approximately 13 - 14 cm (Neuenschwander et al., 2008). About 21 million waveforms were acquired in five strips, but only data from 4th strip, which contains 2,867,200 waveforms, were used in this study.

Computational Platform

The Steele community cluster, which is managed by RCAC (Rosen Center for Advanced Computing) at Purdue University, was used for the study. It consists of 893 nodes and 7,144 processors. Each node has two quad-core processors and either 16 GB or 32 GB of memory. They are inter-connected by either Gigabit Ethernet or Infiniband. All nodes run Red Hat Enterprise Linux 4 and use PBSPro 9 for job management. The parallel LIDAR waveform decomposition algorithm was implemented in the C programming language using GSL (GNU Scientific Library) for the implementation of the nonlinear least squares algorithm and MPI (Message Passing Interface) library for communication among processors.

EXPERIMENTS

The proposed work load balancing approaches for a parallel waveform decomposition algorithm in the previous study are composed of two main steps; (1) complexity estimation, and (2) mapping waveforms onto multiple processors (Jung et al., 2009). The NW approach does not perform complexity estimation and groups waveforms into subsets so that each subset contains the same number of waveforms. The other three approaches (DRLW, DRSW, DTLW) perform complexity estimation and assign waveforms to subsets based on the estimated complexity so that each subset contains the same amount of complexity. Here, complexity is estimated from the sampled waveforms. The DRLW and the DRSW approaches perform complexity estimation using the estimated number of components of the sampled waveforms, while the DTLW approach estimates complexity from the decomposition run time of the sampled waveforms. The DRLW approach estimates the complexity of the subset as the estimated number of components of the sampled waveform, and the DRSW approach estimates the complexity of the subset as the square of the estimated number of components of the sampled waveform. The DTLW approach estimates the complexity of the subset as the decomposition run time of the sampled waveform. Once waveforms are divided into subsets based on their own criteria, the subsets are mapped onto processors for parallel execution.

Resources of the Steele cluster community cluster are shared by lots of researchers, and the performance of the parallel algorithms is affected seriously especially when multiple computationally extensive processes are running in the same node at the same time. The main goal of this study is to assess the stability and scalability of four work load balancing approaches. For the assessment of the scalability and stability, three processor selection schemes - (1) selecting 1 processor per node, (2) selecting 4 processors per node, and (3) selecting 8 processors per node - were designed and tested with a various number of processors (from 8 to 1,024 processors). Since each node in the Steele cluster has 8 processors, selecting

8 processors per node would be the case in which all nodes participating in the computation are not interrupted by other users during the decomposition process except the hardware failure, and selecting 1 processor per node would be the case in which nodes participating the computation are highly likely to be affected by other computationally intensive processes.

These processor selection schemes were applied to each work load balancing approach with different number of processors from 8 to 1,024. In order to better assess the stability of each experiment, each experiment was repeated 40 times, and participating nodes and processors are randomly selected for every experiments.

RESULTS AND DISCUSSION

The efficiency of parallel algorithm is defined as the ratio between run time of the serial execution and the effective run time of the parallel execution (Eq. 1).

$$E = \frac{T_s}{pT_p} \quad (1)$$

Among 40 experiments for each work load balancing approach with different processor selection scheme, the best and the worst efficiency were calculated and plotted (Fig. 1, Fig. 2). The best efficiency is computed using the minimum parallel run time, and the worst efficiency is computed using the maximum parallel run time of the 40 experiments.

In general, the efficiency of parallel algorithm goes down as the number of processors increases especially when a parallel algorithm depends heavily on communication among processors. However, a parallel LIDAR waveform decomposition algorithm developed in the previous study has very little communication overhead. The only communication occurring during the parallel execution is the gathering operation after the sampled waveforms are decomposed. Therefore, the only factor which affects the efficiency of the parallel algorithm is the work load balancing among processors.

In general, the work load balancing approaches based on the complexity estimation showed significant efficiency drop as the number of processor gets large (Fig. 1, Fig. 2). The reason for this significant efficiency drop is because the complexity is estimated from the sampled waveforms and accuracy of the estimated complexity is severely affected when the number of the sampled waveforms decreases to estimate complexity of the subsets. 2,867,200 waveforms were utilized in this study, and the average number of waveforms assigned to each processor is approximately 2,800 when 1,024 processors are utilized for the parallel execution. 1% sampling rate is used in this study, and approximately only 28 waveforms were used to estimate the complexity in this case. Using smaller number of sampled waveforms can be problematic especially when the complexity varies significantly along scan lines.

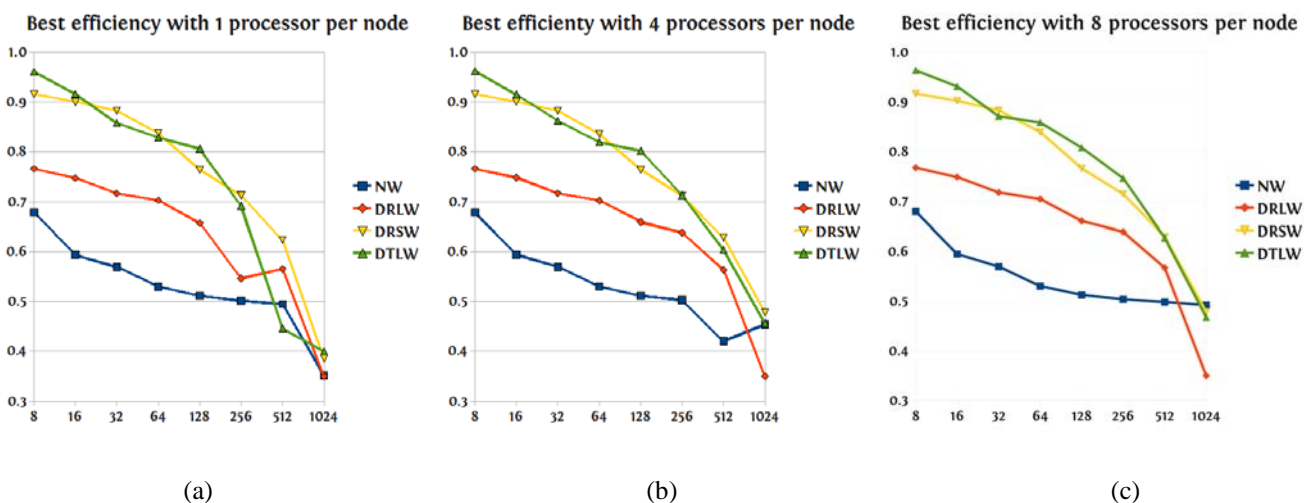


Figure 1. Best efficiency of four (NW, DRLW, DRSW, DTLW) work load balancing approaches among 40 experiments when (a) 1 processor per node, (b) 4 processors per node, and (c) 8 processors per node selection scheme is used.

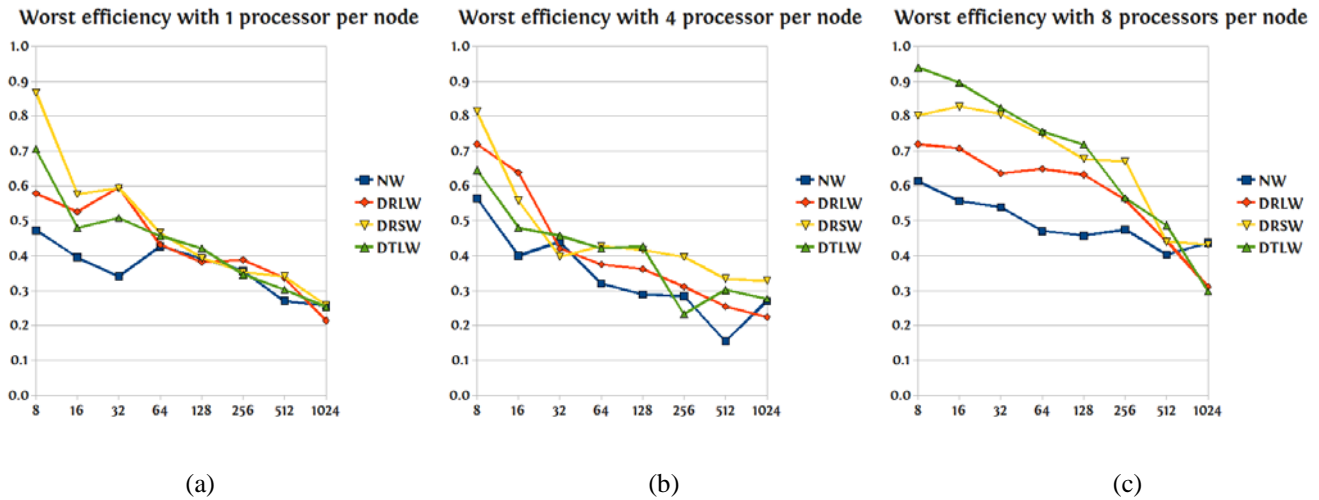


Figure 2. Worst efficiency of four work load balancing approaches (NW, DRLW, DRSW, DTLW) among 40 experiments when (a) 1 processor per node, (b) 4 processors per node, and (c) 8 processors per node selection scheme is used.

Fig. 1 showed that the DTLW and the DRSW approach generally performed the best among the four work load balancing approaches, and the NW approach yielded the worst performance. Fig. 1 also indicated that the efficiencies of the work load balancing approaches based on the complexity estimation (DRLW, DRSW, DTLW) were severely affected as the number of processors increases, while the NW approach yielded the most scalable and stable results as it maintains its efficiency stable as the number of processors increases. It is also noted that the NW approach even performed the best when 1,204 processors are utilized in the parallel execution and 8 processors were selected from each node.

Fig. 2 indicated that the efficiency of work load balancing approaches based on the complexity estimation (DRLW, DRSW, DTLW) showed bigger variation of the efficiencies, while the NW approach showed smaller variation. When 1 processor is selected from each node, the efficiencies of all approaches yielded similar performance if the number of processors is larger than 64. When 4 processors are selected from each node, the DRLW, DRSW, DTLW approaches yielded better performance than the NW approach in general, even though there are some cases where the efficiencies were lower than that of the NW such as DRSW with 32 processors and DTLW with 256 processors. When 8 processors are selected from each node, the efficiency showed similar pattern as Fig. 1, that is the DTLW and DRSW approaches yielded the best efficiency in general and the NW yielded the worst efficiency while the most stable and scalable efficiency.

CONCLUSION

LIDAR waveform decomposition plays an important role in LIDAR data processing because the resulting decomposed components are assumed to represent reflection surfaces within waveform footprints. Hence the decomposition results ultimately affect the interpretation of LIDAR waveform data. LIDAR waveform decomposition is also computationally intensive process, so decomposing the enormous number of waveforms is challenging using traditional single processor architecture. Furthermore, there may be a situation in which LIDAR waveform data need to be processed near real-time. Four work load balancing approaches (NW, DRLW, DRSW, and DTLW) were tested with different number of processors and with different processor selection schemes from each node. The NW approach yielded the worst efficiency when the number of processors is small, but it maintains the efficiency stable well as the number of processors increases. The other three work load balancing approaches which are based on the complexity estimation yielded better efficiency than the NW approach, but their efficiencies were significantly affected as the number of processors increases. In sum, the NW approach is the most scalable and stable work load balancing approach, and the DRSW and DTLW approaches are the most efficient work load balancing approach when the number of processors is small and computation is not affected by other users' processes.

ACKNOWLEDGMENT

This research was supported by a grant (07KLSGC03) from Cutting-edge Urban Development - Korean Land Spatialization Research Project funded by Ministry of Land, transport and Maritime Affairs of Korean government.

REFERENCES

- Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Deseilligny, M.P., and Peuch, W., 2007, Processing full-waveform LIDAR data: Modelling raw signals. *ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007*, pp. 102-107.
- Duong, V. H., Lindenbergh, R., Pfeifer, N., and Vosselman, G., 2008. Single and two epoch analysis of ICESat full waveform data over forested areas. *International Journal of Remote Sensing*, 29 (5):1453-1473.
- Hofton, M.A., Blair, J.B., and Minster, J., 2000, Decomposition of Laser Altimeter Waveforms. *IEEE Transactions on Geoscience and Remote Sensing*, 38(4), pp. 1989-1996.
- Jung, J., Crawford, M.M., Lee, S., 2009. Complexity Estimation Based Work Load Balancing for a Parallel Lidar Waveform Decomposition Algorithm, *Korean Journal of Remote Sensing*, 25(6):1-10.
- Mallet, C. and Bretar, F., 2009. Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1):1-16.
- Neuenschwaner, A.L., Urban, T.J., Gutierrez, R., and Schutz, B.E., 2008. Characterization of ICESat/GLAS waveforms over terrestrial ecosystems: Implications for vegetation mapping. 2008, *American Geophysical Union*, 113.
- Persson, A., Soderman, U., Topel, J., and Ahlberg, S., Visualization and analysis of full-waveform airborne laser scanner data. *ISPRS Workshop on Laser scanning 2005*, pp. 103-108.
- Reitberger, J., Krzystek, P., and Stilla, U., 2008. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees, *International Journal of Remote Sensing*, 29 (5):1407-1431.