SRC: Accelerating the Calculation of Minimum Set of Viewpoints for Maximum Coverage over Digital Elevation Model Data by Hybrid Computer Architecture and Systems

Chenggang Lai, Miaoqing Huang
CSCE Department
University of Arkansas
Fayetteville, Arkansas 72701, USA
Email: {cl004,mqhuang}@uark.edu

Xuan Shi
Department of Geosciences
University of Arkansas
Fayetteville, AR 72701, USA
Email: xuanshi@uark.edu

ABSTRACT

This paper introduces how to accelerate the calculation of the minimum set of viewpoints for the maximum coverage over digital elevation model data using Intel's Xeon Phi and a computer cluster equipped with Intel's Many-Integrated-Core (MIC) coprocessors. This data and computation intensive process consists of a series of geocomputation tasks, including 1) the automatic generation of control viewpoints through map algebra calculation and hydrological modeling approaches; 2) the creation of the joint viewshed derived from the viewshed of all viewpoints to establish the maximum viewshed coverage of the given digital elevation model (DEM) data; and 3) the identification of a minimum set of viewpoints that cover the maximum terrain area of the joint viewshed. The parallel implementation on the hybrid computer cluster was able to achieve more than 100x performance speedup in comparison to the sequential implementation. The outcome of the computation has broad societal impacts since the research questions and solutions can be applied to real-world applications and decision-making practice.

1. INTRODUCTION

Identifying a minimum set of observational viewpoints that can cover the maximum area of a given terrain has high values in many applications including civil engineering, infrastructure optimization and management, and military operations. Theoretically this minimum set problem can be elaborated as given a set U of n elements, and a collection I = [S_1, S_2, ..., S_m] of m subsets of U such that the union of S equals U. The set cover problem is to identify the smallest subset of S whose union covers U. Such an optimization problem is NP-hard [1]. More formally, no polynominal solution has been identified for such a set coverage problem.

Approximate solutions can be explored using heuristic strategies, which typically take a very long process. In order to reduce the computation time, we use computer clusters with coprocessors/accelerators to parallelize the application. To achieve the goal of this research, three computation tasks have to be implemented. *Firstly*, for any given DEM data, all potential control viewpoints will be extracted automatically through map algebra calculation and hydrological modeling approaches. *Secondly*, the viewshed calculation has to be implemented on each viewpoint to generate the joint viewshed of all viewpoints to establish the maximum viewshed coverage of the given DEM. The R3 [2] and the sweep line [3] algorithms are implemented in this study. *Thirdly*, the minimum set coverage computation is to derive the minimum set of viewpoints that have their joint viewshed equals to the maximum coverage, as shown in Algorithm 1.

Algorithm 1. Finding a Minimum Set of Viewpoints for the Maximum Coverage of Digital Terrain.

```
Initialize the solution set S to empty;
while (joint viewshed criterion is not satisfied) do
    for (each viewshed Pi in the potential points P) do
        Compute its overlap fractions and Euclidean distances
    between viewpoints and S;
        if (overlap fractions > overlap criterion or Euclidean
    distances < distance criterion)
            Pi cannot be added to S;
        Calculate the joint coverage;
        if (joint coverage > maximum joint coverage)
            Maximum joint coverage := joint coverage;
    end for;
    Add Pi to S and remove it from P;
end while;
```

2. METHODS AND EXPERIMENTS

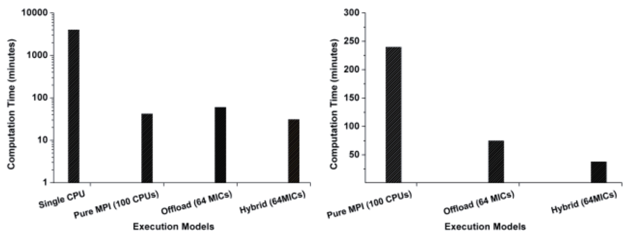
We conducted our experiments on two platforms, the NSF sponsored Arkansas High Performance Computing Center (AHPCC) computer cluster, which is a CPU cluster (i.e., Xeon E5-2670 8-core 2.6 GHz processors), and Beacon supercomputer, which is a hybrid cluster containing both CPUs (i.e., Intel Xeon E5-2670 8-core 2.6 GHz processors) and Intel MIC coprocessors (i.e., Intel Xeon Phi 5110P). Task 1 (i.e., identifying control points) runs on the CPU sequentially and only takes 1 minute. Both Task 2 (i.e., viewshed computation) and Task 3 (i.e., finding the minimum set of viewpoints) are time-consuming and thus have to be parallelized. The parallel solutions are implemented using the following three models.

- MPI: AHPCC cluster is employed for the viewshed calculation and the parallel minimum set calculation through MPI commands. In this implementation, a single-thread MPI process is directly executed on a CPU core. We used sweep line algorithm for viewshed calculation on CPU.
- MIC+Offload: In this model, the MPI processes are hosted on the CPU cores, which offload the computation including data to the MIC processors on Beacon. The host MPI process on CPU issues multiple threads to the MIC card using OpenMP so that each thread works on one or more coefficient vectors depending on the number of participating MIC cards. The computation tasks are done on MIC processors, while CPU cores just wait for the results. We used R3 algorithm for viewshed calculation on

MIC cards as sweep line algorithm has data dependency issue. Algorithm 1 is applied for the minimum set calculation.

- MIC+Hybrid: In this model, both CPUs and MICs are utilized for data processing on Beacon. First the workload is distributed to CPUs through MPI. Then a host CPU will offload part of the workload to a MIC card using OpenMP. On the host CPU, we also use OpenMP to spawn multiple threads for parallel processing. The R3 algorithm is used for viewshed calculation, while Algorithm 1 is applied for the minimum set calculation.

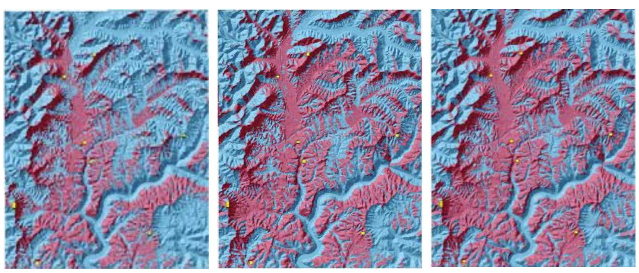
3. RESULTS AND DISCUSSION



(a) Task 2: viewshed computation.

(b) Task 3: minimum set calculation.

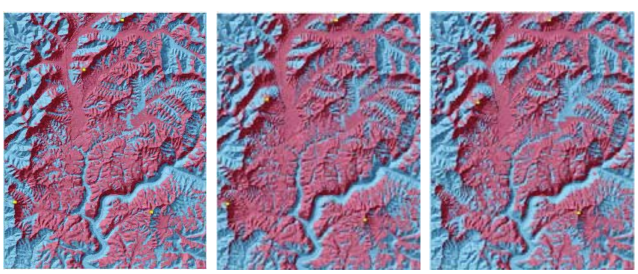
Figure 1. Performance comparisons.



(a) 1 foot, 10 FCC antennas, 33% coverage

(b) 30 feet, 10 FCC antennas, 46% coverage

(c) 60 feet, 10 FCC antennas, 49% coverage



(d) 1 foot, 4 viewpoints, 50% coverage

(e) 30 feet, 4 viewpoints, 50% coverage

(f) 60 feet, 3 viewpoints, 50% coverage

Figure 2. Visibility coverage at different offset heights.

The 3-meter resolution DEM data for West Virginia is used in this study. Federal Communications Commission’s antenna data are used for validation and comparison. Although multiple DEM tiles were tested in this experiment, only the result on *Summersville* is reported. This DEM tile has 3,646 × 4,626 pixels, in which 10 FCC antennas are installed in this DEM tile. 4,106 viewpoints are automatically derived from Task 1 calculation. The joint viewshed of these 4,106 viewpoints can cover about 99.5% of this tile of DEM. The remaining task is to identify the minimum set of 4,106 viewpoints that can cover the same joint viewshed area.

The performance of Task 2 (i.e., viewshed computation) under different execution models are shown in Figure 1(a). For the offload model, each host CPU will host one MPI process, which offloads the computation including data to the MIC coprocessors. We schedule 240 threads to a MIC card. For the hybrid model, both CPUs and MICs are allocated for data processing. We run 4 threads on the host CPU and evenly divide the workload between a host CPU and its corresponding MIC coprocessor. We also schedule 240 threads to a MIC card. From the result, the hybrid model has the best performance. However, the performance of pure MPI model (100 CPUs on AHPCC) is better than that of offload model, since the sweep line algorithm is more efficient than R3 and there is about 10 times performance difference between them. From Figure 1(b) we can see that the execute time of minimum set implementations under offload model and hybrid model are much shorter than the time under the pure MPI model.

The proposed workflow successfully derives the minimum set of 1,217 viewpoints to achieve the goal of the maximum coverage of 100%, which means the selected minimum set of viewpoints can achieve the same coverage of 4,106 viewpoints on the tile of Summersville DEM. Obviously such a minimum set still contains a large number of viewpoints because many single cells in the DEM grid can only be seen by one viewpoint. When the criteria of maximum coverage are changed, the number of minimum set can be reduced significantly.

Figures 2(a)-2(c) display the visibility coverage of current locations of FCC antennas at different offset heights. Even when the height of the antennas is set to 60 feet, these 10 FCC antennas can only cover 49.74% of this area. Figures 2(d)-2(f) display the result derived from the minimum set calculation. Only 4 antennas are required to cover 50% of the area even when the offset height is set to 1 foot, or only 3 antennas are required when the offset height is set to 60 feet.

4. CONCLUSIONS

While a few relevant works [4] were conducted in the past decades, we resume this challenging research on generating a minimum set of viewpoints for the maximum coverage over large-scale digital terrain data. The comprehensive workflow has been implemented and validated with satisfactory results in comparison to the current locations of FCC antennas. The computational bottleneck of the proposed workflow mainly lies in viewshed/joint viewshed calculation, counting visible pixels, computing the ratio of overlaid viewshed, and minimum set calculation. Although deploying CPU clusters can help reduce the computational time, modern accelerator technologies can achieve better efficiency and scalability when large volumes of high-resolution DEM data are to be processed.

5. REFERENCES

[1] Feige, U. 1996. A threshold of ln n for approximating set cover. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp. 314-318. ACM

[2] Sorensen, P.A. and Lanter, D.P. 1993. Two algorithms for determining partial visibility and reducing data structure induced error in viewshed analysis. *Photogrammetric Engineering and Remote Sensing*. 59(7):1149-1160.

[3] Warn, S. 2011. High performance geospatial analysis on emerging parallel architectures. PhD Thesis, U. of Arkansas.

[4] Franklin, W. R. 2002. Siting Observers on Terrain. In Advances in Spatial Data Handling: 10th International Symposium on Spatial Data Handling, edited by D. Richardson and P. van Oosterom, 109–120.