

機械学習ソフトウェアへのソフトウェア自動チューニング技術の適用

工学院大学情報学部 田中 輝雄

名古屋大学情報基盤センター 大島 聡

工学院大学情報学部 藤井 昭宏

東北大学サイバーサイエンスセンター 滝沢 寛之

工学院大学情報学専攻 多部田敏樹

工学院大学情報学専攻 藤家空太郎

プログラムを実行しながら自動的にチューニングを行うソフトウェア自動チューニングでは、対象とするプログラムの性能を決定する複数の性能パラメタの最適な組合せを効率よく探索することが重要な課題のひとつである。この課題を複数の性能パラメタの取りうる値からなる離散空間上の探索問題とし、この探索を限られた空間から徐々に範囲を広げ、離散空間上を探索する反復 1 次元探索手法を提案している。さらに、検証フェーズを設け、実行時の測定の揺らぎなどを考慮し、得られた性能パラメタ組合せの有効性を高めている。本研究においては、高性能コンピュータ環境が必須であり、これまで、東大/Reedbush, JCAHPC/Oakforest-PACS, 産総研/ABCI などを活用してきた。

本研究では、機械学習を用いた歩行者経路予測プログラムのハイパーパラメタの選択に適用する。ここでハイパーパラメタとは、学習率やドロップアウト率などの機械学習の挙動を制御する性能パラメタである。このプログラムは、学習フェーズと予測フェーズからなり、ターゲットとなる歩行者の過去の移動軌跡データから未来の経路と到着地点を予測する。評価指標値には、実際の歩行者の到達地点と予測器が予測した到達地点の誤差である FDE(Final Displacement Error)を用いる。この FDE が小さくなるハイパーパラメタの組み合わせを推定する。特にクラウド環境で行う学習フェーズでは、毎回の機械学習に時間を要すること、および膨大なハイパーパラメタの組合せが課題である。今回対象とするプログラムでは 1 回の学習で 30 分程度を要する。また、推定するハイパーパラメタは 6 種類あり、各々 5 つの値を取り得るため、ハイパーパラメタの組み合わせは全部で 15625 通りとなる。そのため、逐次実行による自動チューニングでは、ハイパーパラメタ推定に数日を要する。そこで多数の GPU 環境をノードとした環境での並列処理実行の実現が重要となる。

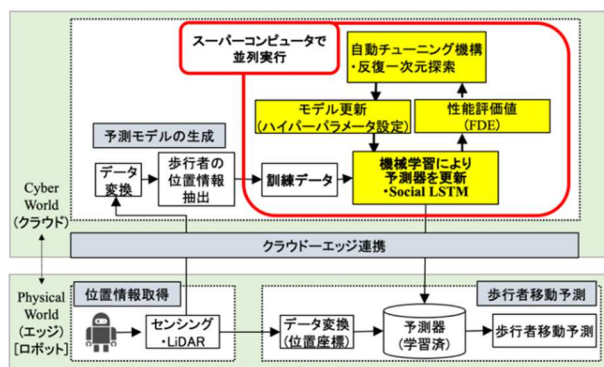


図 1 対象とするプログラムと自動チューニング

我々は好適な実行環境として、多くの GPU コアを持つ名古屋大学のスーパーコンピュータ「不老」 Type II サブシステムを用いる。ここでは1ジョブにつき1GPU を使うことにより使用するノードはユーザ間で共有する(占有せず柔軟に実行)。使用したリソースグループではジョブ同時実行可能数が 50 のため、最大で 50 のハイパーパラメタの組合せで対象プログラムを同時に実測できる。

逐次実行と並列実行による推定結果を表 1 に示す。初期パラメタは、経験則で最良として使用中のハイパーパラメタの組み合わせとした。自動チューニングにより、ハイパーパラメタの組合せを自動的に順次選択し推定を進めた。

推定フェーズでは、推定完了までは、逐次実行が約 279 時間 (11.6 日) 要したのに対し、並列実行は約 12.3 時間となった。一方、探索回数は逐次実行の 553 回に対し、並列実行では 761 回となった。これは並列実行による同時探索により探索回数が増加したからである。評価指標値 (FDE) について、推定する前(現在の経験則)は 1.21m であったが、今回の評価では逐次実行が 0.89m、並列実行が 0.81m と削減できた。

さらに、検証フェーズとして、推定した候補を含めた有望なハイパーパラメタの組合せの候補について改めて複数回の実行により、選択したハイパーパラメタの組合せの有効性を確認した。検証フェーズを含めて全体の並列化の効果として約 23 倍高速化の結果を得た。

今後、「不老」 Type II サブシステムを用いて適用アプリケーションを増やし、並列環境向けに改良した自動チューニングにおける性能パラメタ推定の有効性をさらに示していく。

【謝辞】

本研究では、東京女子大学 加藤由花教授から機械学習プログラムを提供いただきました。

【成果発表】

- [1] 藤家, 多部田, 藤井, 田中, 加藤, 大島, 片桐, GPU クラスタを用いて並列化した自動チューニングの機械学習プログラムへの適用と安定性の検証, 情報処理学会, 研究報告ハイパフォーマンクスコンピューティング(HPC), 2021-HPC-178, No. 16, pp1-8, 2021. 3
- [2] 多部田, 藤家, 藤井, 田中, 加藤, 大島, 片桐, マルチ GPU 環境における機械学習ハイパーパラメタの自動チューニング(1), 情報処理学会第 83 回全国大会, 2021. 3
- [3] 藤家, 多部田, 藤井, 田中, 加藤, 大島, 片桐, マルチ GPU 環境における機械学習ハイパーパラメタの自動チューニング(2), 情報処理学会第 83 回全国大会, 2021. 3

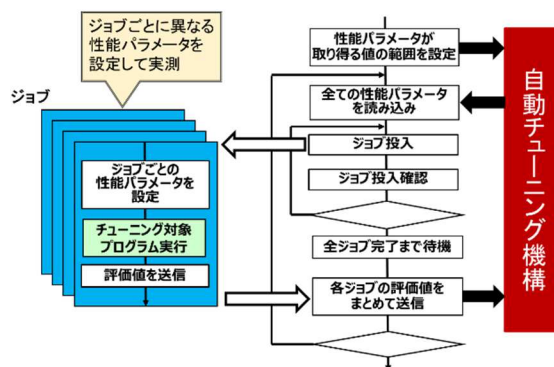


図 2 自動チューニング手順と並列実行

表 1 実行結果

		推定フェーズ	検証フェーズ	合計
逐次実行	実行時間	279時間	81時間	360時間
	実測回数	553回	160回	713回
並列実行	実行時間	12.3時間	3.6時間	15.9時間
	実測回数	761回	300回	1061回