

閾値網羅法における機密性・高速性を考慮した閾値決定法

富田 航平† 藤橋 卓也‡
†愛媛大学工学部情報工学科

遠藤 慶一‡ 小林 真也‡
‡愛媛大学大学院理工学研究科

1 はじめに

グリッドコンピューティングは、ネットワーク上に存在する多数の計算機を利用し、高い処理能力や記憶容量を安価に入手する技術である。グリッドコンピューティングの一つに、インターネット上の計算機を利用するエクスターナルグリッド（以降、単にグリッドと呼ぶ）がある。インターネット上には非常に多数の計算機が存在するため、実質的に利用できる計算資源が無限であるというメリットがある。一方、プログラムが正しく処理される保証がなく、また不正な使用を目論む悪人の手に情報が漏洩するリスクがあるというデメリットがある。

本研究では、ユーザの要求に応えられる機密性、高速性の確保を目指す。そこで、閾値網羅法の閾値がグリッドの機密性、高速性に与える影響について評価し、悪人集団による影響を排除しながら高速性を得るための最適なパラメータについて考察する。

2 セキュアプロセッシング

グリッド上で実際に処理を行う計算機を処理ノード、不正行為を行う処理ノードを悪人と呼ぶ。セキュアプロセッシングとは、グリッド上の悪人が行う不正行為への対策のために考案された技術の総称である。不正行為には、不正な解析と改竄がある。不正な解析は、悪人が渡されたデータから元のプログラムを解析する行為である。改竄は、悪人が意図的に正しくない結果を返すことである。以下に、セキュアプロセッシングの技術の一部を示す。

プログラム分割

グリッドに依頼するプログラムを複数のプログラム断片に分割する技術である。一つの処理ノードに渡る情報量が減少するため、不正な解析に対して効果がある。

多重化

多重化は多重処理と採択の2つのプロセスからなる。多重処理とは、同一のプログラム断片を複数の処理ノードに依頼することを意味する。採択とは、複数の処理ノードから返ってきた結果の中から多数決によって結果を選択することを意味する。多重化は改竄に対して効果がある。なお、同じプログラム断片を渡す処理ノードの台数を多重度と呼び、同じプログラム断片を処理する処理ノード群をまとめてブロックと呼ぶ。

3 先行処理

3.1 単純な先行処理

インターネット上には高速な計算機も低速な計算機も存在するため、同一のプログラム断片を同じように依頼したとしても、結果が返ってくるまでの時間は各ノードによって異なる。過半数による採択では、少なくとも多重度 n に対して $\lceil \frac{n}{2} \rceil$ 台のノードが結果を返すまで待機時間が生じる。

これを回避するため、最速のノードが返した結果を仮採択として後続の処理を行い、全体の高速化をはかる手法が先行処理である。しかし、最速のノードが返す結果が間違っていた場合、ロールバックが発生し高速化が阻害される。

3.2 網羅法と閾値網羅法

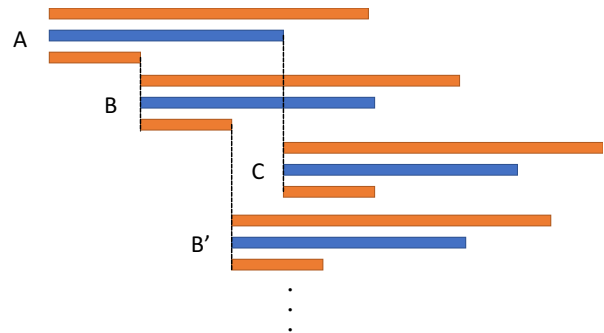


図 1: 網羅法を利用したグリッドの動作例

先行処理を改良した手法に網羅法がある。網羅法は全ての結果を仮採択して網羅的に先行処理を行うことで高速化の効果を高めている。網羅法を利用したグリッドの動作例を図 1 に示す。図 1 中の A から網羅法を開始した場合、A から分岐する枝 B, C だけではなく、B からさらに先行した枝 B' も発生する。より大きな多重度のグリッドで多数の結果が返された場合、この現象が多数の枝で発生してしまう。

このように、網羅法は安定して高速化を行える反面、参加台数が膨大になるという欠点がある [1]。これは悪人混入のリスクを高めることに直結するため、機密性の観点から好ましくない。そこで、仮採択に必要な台数を閾値として定めた上で網羅的に先行処理を行う閾値網羅法という手法が考案されている。閾値網羅法によって機密性、高速性のバランスを操作することができるが、どんな時にどのような閾値が有効かという点については議論がなされていない。

4 評価方法

閾値網羅法の閾値とグリッドの機密性、高速性の関係を明らかにするために、グリッドの動作を想定したシ

Threshold determination method considering confidentiality and processing power in comprehensive processing with threshold †K. Tomita

Department of Computer Science, Faculty of Engineering, Ehime University

‡T. Fujihashi, K. Endo, S. Kobayashi

Graduate School of Science and Engineering, Ehime University

シミュレーションを行う。シミュレーション条件は以下の通りである。

- 悪人は悪人の存在確率に応じてグリッド中に存在し、必ず正しくない結果を返す。
- グリッド中に存在するすべての悪人が共謀して同じ結果を返す。
- 各処理ノードの処理性能は、形状尺度 $k = 5$ 、尺度母数 $\theta = 2/5$ のガンマ分布に従う。
- 処理するプログラムは逐次的に実行される 100 ブロックの処理に分割され、各ブロックのサイズは 1 である。

これらの条件で多重度、閾値網羅法の閾値、悪人の存在確率を変えながらそれぞれ 1000 回ずつ試行を行い、処理時間と参加台数を計測する。なお、処理時間は高速性をはかる指標であり、参加台数は機密性をはかる指標である。

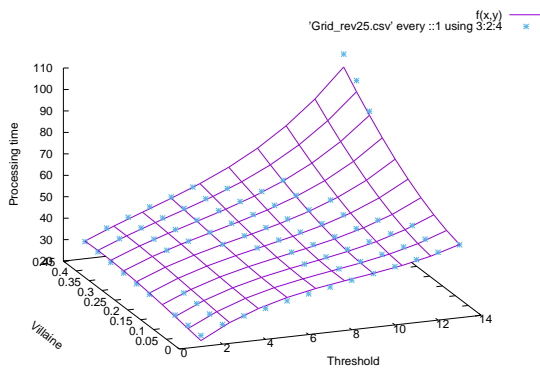


図 2: 悪人の存在確率および閾値に対する処理時間と近似曲面の関係 (多重度 25)

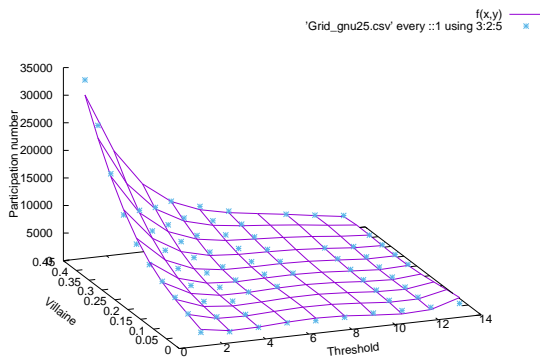


図 3: 悪人の存在確率および閾値に対する参加台数と近似曲面の関係 (多重度 25)

5 評価結果と考察

5.1 処理にかかる時間と処理に参加するノード数

多重度 25 の時について、悪人の存在確率および閾値と処理時間の関係を図 2 に、悪人の存在確率および閾

値と参加台数の関係を図 3 に示す。なお、図中の点はプロットされた実データを示し、メッシュで示した平面は実データを近似した曲面を表している。以下の (1) に処理時間に関する近似曲面、(2) に参加台数に関する近似曲面の方程式を示す。ここで、 T は処理時間、 N は参加台数、 t は閾値、 v は悪人の存在確率を示す。また、近似は最小二乗法を用いて行い、二乗平均平方根誤差はそれぞれ 2.48, 586.63 となった。

$$T = 0.028925t^3 - 50.142v^3 + 0.853233t^2v + 13.0116v^2t - 0.711318t^2 - 11.235v^2 - 11.0696tv + 7.27295t + 31.6273v + 16.2113 \quad (1)$$

$$N = 3.10258t^4 - 346527v^4 - 100.988t^3v - 15756.1tv^3 + 1434.3t^2v^2 - 81.8495t^3 + 478238v^3 + 2063.75t^2v - 19851.5tv^2 + 711.734t^2 - 32942.6v^2 - 12473.3tv - 2326.47t + 30582.3v + 4738.11 \quad (2)$$

図より、悪人が少ない環境で大きい閾値を設定した時や、悪人の多い環境で小さい閾値を設定した時、高速性や機密性が著しく低下することが分かる。また、閾値を大きくした時には高速性が低下する代わりに機密性が向上し、閾値を小さくしたときには機密性の代わりに高速性が向上する性質が定量的に示された。近似曲面については、(1) は平均 5% 程度の誤差で済むため、この方程式は十分な精度があると考えられる。(2) は平均 10% 程度の誤差が生じるため高精度とは言えない。

6 処理時間および参加台数の要求を用いた閾値の決定

(1), (2) に示した近似曲面の式を利用し、高速性・機密性に対する要求から最適な閾値を決定する。具体的には、機密性への要求は閾値の下限を、高速性への要求は閾値の上限を決定することができる。この範囲内にある閾値が、両方の要求を満たした適切な閾値の候補となる。候補となる閾値から最適なものを決定するには、閾値を処理時間と参加台数の積のような評価関数によって評価すればよい。

また、近似曲面の係数を多重度を用いて近似することで、要求に応じた多重度と閾値を合わせて決定することができる。この時、上記の評価関数では多重度が小さいほど良いと評価される。これは機密性と高速性の双方を最大化する選択としては正しいが、多重度の低下は処理結果の正しさである信頼性の低下につながるため、必ずしもグリッドとして最適な閾値とは限らない。

参考文献

- [1] 広瀬 吉隆, 稲元 勉, 樋上 喜信, 小林 真也: “セキュアプロセッシングにおける先行処理による処理時間改善に対する定量的評価”, 第 14 回情報科学技術フォーラム (FIT2015) 講演論文集, Vol.4, pp.241-242, 2015