PC クラスタ環境における並列シミュレーテッド・アニーリング計算法

正員 趙悦
正員 銭飛
正員 呉継軍
非会員 青井秀樹
非会員 何愛国

(広島国際学院大学)
(広島国際学院大学)
(広島国際学院大学)
(広島国際学院大学)
(株式会社アイヴィス)

A Parallel Simulated Annealing Algorithm in PC-Cluster Environments

Yue ZHAO, Member, Fei QIAN, Member, Jijun WU, Member, Hideki AOI, Non-member (Hiroshima Kokusai Gakuin University), Aiguo HE, Non-member (IVIS. INC.)

Parallel Simulated Annealing(PSA) is an effective method for some commbinatorial optimization problems. In this letter, We apply PSA to the graph partitioning problem and the maximum clique problem in a PC-cluster environment, and show the results of our simulations.

キーワード : Simulated Annealing, Combinatorial Optimization, Distributed Computation, Parallel Computation

1. まえがき

シミュレーテッド・アニーリング法 (Simulated Annealing) (SA 法と略記) は大規模組合せ最適化問題に使われている一般的な解法として広く利用されているが、計算時間が非常に長いという致命的な欠点がある。この欠点を克服する方法の一つは、並列計算を通じて計算時間の縮減が考えられる。本稿ではまず、各温度ステップで、部分同期過程⁽¹⁾の概念に基づいた並列 SA 法のアルゴリズムを記述した後、PC クラスタの環境で、グラフ分割問題、および最大クリーク抽出問題に応用し、従来の SA 法⁽²⁾ 及び平均場法 (MFA)⁽³⁾ との比較によって計算時間の短縮効果を示す。

2. 目的関数の分解と構造グラフ

$\mathbf{x}=(x_1,x_2,...,x_N)$ を変数ベクトルとし、 $f(\mathbf{x})$ を組合せ最適化問題の評価関数とする。 $f(\mathbf{x}) = g_1(\mathbf{x}^1) + g_2(\mathbf{x}^2) + ... + g_M(\mathbf{x}^M)$ を $f(\mathbf{x})$ の分解と呼ぶ。ここで、 $\mathbf{x}^1,...,\mathbf{x}^M$ は $\mathbf{x}$ のサブベクトルであり、 $g_1(\mathbf{x}^1),...,g_M(\mathbf{x}^M)$ を $f(\mathbf{x})$ の成分と呼ぶ。それに対して、 $f(\mathbf{x})$ の構造グラフは次のように定義する。

(i)各変数 $x_1,...,x_N$ にそれぞれノードを対応づける。
(ii)変数 x_i と x_j とが同一成分に属しているとき、ノード x_i と x_j とを無向枝で結ぶ。

明かに $f(\mathbf{x})$ の一つの成分は構造グラフの一つのクリーク (完全サブグラフ) に対応する。したがって

$$f(\mathbf{x}) = \sum_{c \in C} f_c(\mathbf{x}_c) \tag{1}$$

と書ける。ここで C は構造グラフにおける $f(\mathbf{x})$ の成分に対応するクリークの集合であり、 $f_c(\mathbf{x}_c)$ はクリーク

c に対応する $f(\mathbf{x})$ の成分である。

3. 並列的シミュレーテッド・アニーリング法 (PSA)

式 (1) の $f(\mathbf{x})$ の関数構造には、 C が 1-クリーク (単一のノード) と 2-クリーク (リンク) しか含んでいなければ、各温度での部分同期空間過程の定常分布は Gibbs 分布に帰着するされることが証明できる。⁽¹⁾ この部分同期空間過程に基づいた並列シミュレーテッド・アニーリング法のアルゴリズムを次のように記述する。

アルゴリズム：

① ノードの集合 $\{1, 2, ..., N\}$ を適当に M 個の部分集合 $A^1, A^2, ..., A^M$ に分割する。

② 各ノード $j(j = 1, 2, ..., N)$ の状態 x_j にランダム的に初期値を与え、温度パラメタ T の初期値を $T_0 = T_{initial}$ に設定する。 $k = 0, t = 0$ 。

③ 部分集合 $A^1, A^2, ..., A^M$ からランダム的に A^α を選んで、 A^α に属する全てのノード $j(j \in A^\alpha)$ に対して同時に (同期的に) を実行する。

④ 現時点でのシステムの様相は $\mathbf{x} = (x_1, x_2, ..., x_N)$ とする。次の確率でノード j の状態を $y_j(y_j \in X_j)$ に遷移させる。

$$P(y_j|\mathbf{x}) = \frac{1}{Z} \exp\{\frac{1}{2T} \sum_{c \in C(j)} [f_c(x_c) - f_c(T_j^{y_j} x_c)]\}$$

$$Z = \sum_{m \in X_j} \exp\{\frac{1}{2T} \sum_{c \in C(j)} [f_c(x_c) - f_c(T_j^m x_c)]\}$$

ここで $C(j)$ は j を含むクリークの集合で、 X_j は x_j の取りうる状態の集合であり、また、 $T_j^m x$ は j 番目のノードの状態 x_j のみ m にする演算子である。

- ⑤ $t < t_{final}$ (各温度での反復回数) であれば, $t = t + 1$, にもどる。そうでなければに進む。
- ⑥ $T_k < T_{final}$ であれば, アルゴリズムが終る。そうでなければ $T_{k+1} = \mu T_k$; $k = k + 1$; $t = 0$; にもどる。ここで, μ は温度の下げ率で, $(0 < \mu < 1)$ である。

このアルゴリズムでは, ③で選ばれた部分集合 A^α に属するすべてのノードに対して同時に④の計算を実行し, ④では式 (2) で j ノードの遷移確率を計算するときに現時点でのシステムの様相 x と j ノード自身の遷移先の状態 y_j のみに依存するので, A^α 中でノード別の並列計算が可能であり, 複数のプロセッサを使用することによって, 計算時間を短縮することが期待できる。

4. 組合せ最適化問題への応用

PSA をいくつかの組合せ最適化問題に応用して, 以下の PC クラスタ並列計算環境で実験を行った。クラスタ数: 16 (CPU 性能: Celeron300 $\times$ 8, Celeron400 $\times$ 8), クラスタ間のネットワーク構成: 100Mbps DUAL, プログラミング言語: MPI/C, OS: Linux 2.2

〈4.1〉グラフ分割問題 グラフ分割問題 (GPP: Graph Partitioning Problem) は無向グラフのノード集合をいくつかの部分集合に分割する問題である。評価基準としては

- (i) 部分集合間にまたがる枝数の総数を最小にする。
- (ii) 各部分集合のノード数なるべく均一にする。
- ノード i の状態を x_i で表す。即ち $x_i = m$ は, ノード i が部分集合 m に割り当てられることを意味する。GPP の目的関数は次の 2 次形式で表すことができる。

$$f(x) = \sum_{i,j} \phi(x_i, x_j) + 2\lambda \sum_{i < j} \xi(x_i, x_j) \quad (4)$$

$$\phi(x_i, x_j) = \begin{cases} 0 & \text{if } x_i = x_j \\ w_{ij} & \text{otherwise} \end{cases}$$

$$\xi(x_i, x_j) = \begin{cases} 0 & \text{if } x_i \neq x_j \\ 1 & \text{otherwise} \end{cases}$$

ここで, λ は制約条件の強さを表す定数である。ノード数が 900 でランダムに枝を生成したグラフについて, $\lambda = 1.5$ での実験結果を表 1 に示す。

表 1. グラフ分割問題の実験結果 (10 回試行の平均値)
Table 1. Simulation results of the graph partitioning problem

par	SA		MFA		PSA	
	cost	time	cost	time	cost	time
10	2360.4	81593	2361.1	6730	2358.0	782
20	6321.4	100145	6327.1	8656	6322.5	911
50	14132.1	114954	14141.5	10811	14125.7	985
100	37219.0	139597	37231.4	12312	37201.3	1115

par: 分割数。cost: 目的関数値。time: 計算時間 (秒)

SA の実験パラメータ: $T_0 = 100.0$, $T_{final} = 0.0001$, $\mu = 0.975$, 各温度での反復回数: 2500, 終了条件: $T < T_{final}$ または同じ値を 100 回以上連続出力した場合。

MFA の実験パラメータ: $T_0 = 100.0$, $T_{final} = 0.0001$, $\mu = 0.975$, 各温度での反復回数: 2500, 終了条件: $T < T_{final}$ 。

PSA の実験パラメータ: グループ数: $M = 20$, プロセッサ数: 9, $T_0 = 100.0$, $T_{final} = 0.0001$, $\mu = 0.975$, 各温度での反復回数: 2500, 終了条件: $T < T_{final}$ または同じ値を 100 回以上連続出力した場合。

〈4.2〉最大クリーク抽出問題への応用 最大クリーク抽出問題 (MCP: Maximum Clique Problem) は無向グラフに含まれている最大クリーク (最大完全部分グラフ) を求める問題である。ノード i の状態を x_i で表す。 i が選択されたクリークに属するとき $x_i = 1$, 属さないときは $x_i = 0$ とする。MCP の目的関数は次の 2 次形式で表すことができる。

$$f(x) = \sum_{i,j} (1 - e_{i,j}) x_i x_j + \sum_i x_i \quad (5)$$

ここでノード i と j の間に枝があるとき $e_{i,j} = 1$ で, 枝がないときは $e_{i,j} = 0$ である。実験結果を表 2 に示す

表 2. 最大クリーク抽出問題の実験結果 (10 回試行の平均値)
Table 2. Simulation results of the Maximum Clique Problems

N/MC	SA		MFA		PSA	
	clique	time	clique	time	clique	time
200/10	10.0	371.4	10.0	93.3	10.0	81.4
400/25	25.0	1123.4	25.0	173.4	25.0	151.2
600/30	28.0	3261.7	26.4	461.8	30.0	379.9
800/35	34.0	7294.1	28.2	910.5	35.0	984.1
1600/50	49.2	7294.1	48.2	4211.1	50.0	2934.6

N: ノード数 MC: 最大クリーク clique: 結果 time: 計算時間 (秒)

SA の実験パラメータ: $T_0 = 100.0$, $T_{final} = 0.0001$, $\mu = 0.975$, 各温度での反復回数: 5000, 終了条件: $T < T_{final}$ または同じ値を 100 回以上連続出力した場合。

MFA の実験パラメータ: $T_0 = 100.0$, $T_{final} = 0.0001$, $\mu = 0.975$, 各温度での反復回数: 3500, 終了条件: $T < T_{final}$ 。

PSA の実験パラメータ: グループ数: $M = 20$, プロセッサ数: 10, $T_0 = 100.0$, $T_{final} = 0.0001$, $\mu = 0.975$, 各温度での反復回数: 2500, 終了条件: $T < T_{final}$ または同じ値を 100 回以上連続出力した場合。

5. むすび

シミュレーテッド・アニーリング法の解の品質を保ちながら計算時間を短縮することが重要な課題の一つである。本稿は一つの試みとして PSA をグラフ分割問題と最大クリーク抽出問題に応用し, PC クラスタの並列計算環境で解の品質と計算時間との両面で若干の成功を収めた。今後の課題としては, 同期させるノードの数と解の品質との関係, 特に完全同期時の解の劣化について解明することがあげられる。

(平成 12 年 01 月 31 日受付, 同 12 年 02 月 23 日再受付)

文 献

- (1) Y. ZHAO, T. FUKAO, "Distributed Computing of a Stochastic Algorithm for Combinatorial Optimizations", SYSTEM MODELLING AND OPTIMIZATION Springer-Verlag Berlin Heidelberg New York. pp.318-327, 1989
- (2) S. KIRKPATRICK, C.D. GELATT, JR., M.P. VERCCHI, "Optimization by Simulated Annealing", Science, Vol. 220, pp.671-680, 1983.
- (3) 呉、原田、深尾, "組合せ最適化問題に対する臨界温度を考慮した適応平均場近似アルゴリズム", 電子情報通信学会論文誌, Vol. J79-A, NO.9, pp.1582-1589, 1996.