Pergamon

Computers Math. Applic. Vol. 27, No. 5, pp. 71–82, 1994

0898-1221/94 $6.00 + 0.00

0898-1221(94)E0007-7

Bin-Packing by Simulated Annealing

R. L. Rao and S. S. Iyengar*
Robotics Research Laboratory, Department of Computer Science
Louisiana State University, Baton Rouge, LA 70803, U.S.A.

(Received and accepted May 1993)

Abstract—The process of gradually settling a combinatorial system into configurations of globally minimum energy has variously been called simulated annealing, statistical cooling, and so on. In the past, very large combinatorial optimization problems have been solved using this technique. It has also been shown that this method is effective in obtaining close-to-optimal solutions for problems known to be NP complete. Further, this technique is applicable to a whole class of problems that satisfy a few requirements.

The purpose of this paper is to illustrate an efficient version of the simulated annealing method as applied to a variant of the bin-packing problem. The computational complexity of the method is linear in input size similar to various well-known heuristic methods for the problem. The solutions obtained, however, are much better than any of the heuristic methods. The particular variant of the bin-packing problem we consider has several practical applications such as static task allocation in process scheduling and batch processing. At the time of this writing, we have not yet seen a stochastic solution to the bin-packing problem in the literature.

One of the distinguishing features of our research is the high quality of solutions obtained by our method. Extensive simulation experiments we have carried out show that the solutions obtained by the stochastic method show a significant improvement over those obtained by any of the well-known heuristic methods.

Keywords—Bin-Packing, Combinatorial optimization, Global minimum, Monte Carlo methods, NP completeness, Simulated annealing, Statistical cooling.

1. INTRODUCTION

1.1. The Bin-Packing Problem

The classical definition of the bin-packing problem involves packing a list of items of (possibly) different sizes into the smallest number of bins, each of which has a given maximum capacity. Coffman *et al.* [1] is a good survey of several approximation algorithms for bin-packing that yield quick sub-optimal solutions.

The variant of the classical problem we solve, deals with a fixed number of bins each with an unlimited capacity and the objective is to pack the items into these bins so that each bin has about the same total allocation. In other words, our attempt is to find the most *equable* distribution of items to bins. Our problem differs from the classical problem in the following two ways.

1. The bin sizes are not constrained. The rationale behind this particular variation of the problem is the fact that this model is appropriate in certain real-world situations. For example, in a batch processing environment, it is sometimes necessary to complete a fixed

This work was supported in part by ONR Grant No. N000014-91-J1306.
*E-mail addresses: `ramana@bit.csc.lsu.edu` and `iyengar@bit.csc.lsu.edu` for R. L. Rao and S. S. Iyengar, respectively.

A typical instance of the annealing experiment involves two important choices for the control parameter corresponding to the high temperature and the low temperature regimes of a physical system and a method of varying the control parameter so that the system is driven to optimality. A set of choices for these critical parameters is known as an *annealing schedule*. There is evidence in the literature that the global optimum of a combinatorial optimization problem can be found with probability one provided that the annealing schedule satisfies certain conditions [8–10].

In practical applications such as circuit placement in VLSI, the computing resources needed to obtain 'good' solutions are excessive. To remedy this, several approaches have been proposed in the literature. Greene *et al.* [6] have suggested a schedule which involves annealing without rejected moves. Their method offers a significant speed-up at the expense of increased memory usage. Lam *et al.* [5], White [11], and Huang *et al.* [12] have proposed techniques to make annealing schedules efficient.

The various approaches that attempt to remedy the massive computational time required by the annealing method, may be grouped into three broad categories—parallel annealing techniques [13], efficient annealing schedules [5,6,11,12] and controlled move generation methods. This last category of methods is usually problem-instance dependent and is not widely applicable to combinatorial optimization. The more popular techniques in the literature employ efficient annealing schedules and this paper also describes one such schedule.

2. PROBLEM FORMULATION

We define an instance of the bin-packing problem as consisting of

1. M bins, each of which has an unlimited capacity;
2. N items (sizes) $t_1, t_2, \ldots, t_N$; $0 \le t_i \le t_{\max}$, $1 \le i \le N$;
3. an objective function (also refered to as the cost or the energy function) defined as

$$\mathcal{C}(\{a_i\}) = \sum_{j=1}^{M} (B_j - \overline{T})^2, \tag{1}$$

where B_j is the sum of the sizes of the items allocated to the j^{th} bin and $\{a_i\}$ is an allocation sequence $a_1, a_2, \ldots, a_N$; $1 \le a_i \le M$. An allocation sequence determines which bin each item is allocated to. Thus, every allocation sequence represents a feasible solution to the problem. The resulting distribution of items to bins is also called a configuration (state) of the problem (system). $\overline{T}$ is the total allocation at each bin that globally minimizes the cost function. Thus,

$$\overline{T} = \frac{1}{M} \sum_{i=1}^{N} t_i; \tag{2}$$

in general, though a particular instance of the problem may not render itself to such a perfect allocation scheme. Hence, there is the minimization problem.

The following lemma follows directly from the problem formulation.

LEMMA 1. *The magnitude of the objective function does not exceed* $M(M-1)\overline{T}^2$.

The maximum value for the objective function is obtained by allocating all the items in the item list to one bin and leaving the other $M-1$ bins empty. Such an allocation yields an objective function of $(M\overline{T} - \overline{T})^2 + (M-1)\overline{T}^2$ which is the same as $M(M-1)\overline{T}^2$. It is easy to see that no other allocation can yield a larger value for the objective function since moving any item from the bin to which all items are allocated now, would decrease the objective function at both the bin to which it is allocated now and the bin to which it is being moved.

4. ANNEALING SCHEDULE

The annealing schedule is described by quantitative choices for the three parameters—the starting value of the temperature T_∞, the stopping value of the temperature T_0, and the decrement function $\mathcal{F}(t)$ which determines the profile of the temperature from the beginning till the end of the annealing process.

The annealing curve obtained from a good schedule typically displays three broad areas of interest—a high energy area characterized by high temperature, an intermediate energy area, and a low energy area characterized by low temperature. The intermediate energy area is well defined and spans a relatively small portion of the temperature axis of the curve. We will refer to the behavior of the system described by the high energy region of the curve as the *High Temperature Regime* and that described by the low energy region of the curve as the *Low Temperature Regime.* These are further described in the following sections.

During the actual simulation, we first carry out an exploratory search of the configuration space where we assume that the temperature is infinite and accept each generated configuration. From this data we obtain fundamental statistical quantities about the problem. In particular, we are interested in the average value of the cost $\langle C(T) \rangle$ and the standard deviation σ of the the density of states distribution. The density of states information may later be used to select an appropriate starting value for the temperature parameter.

4.1. High Temperature Regime

This region of the annealing curve (and the corresponding behavior of the system) is marked by the acceptance of most generated states. The value of the temperature parameter is so high that the Metropolis criterion is always satisfied. Thus, the average energy in this regime is very high. Just how high the starting temperature must be, for a good annealing schedule, is usually determined by monitoring the acceptance ratio at each temperature. The acceptance ratio (the fraction of generated states that are accepted) is arbitrarily fixed at some high value such as 0.9 and the temperature is increased to a value where the acceptance ratio is high enough.

While this serves as a problem-independent method of fixing the starting value of the temperature, often it yields a temperature value that is too high thus yielding an annealing schedule that is wasteful of computational resources. For the bin-packing problem, Lemma 1 gives the theoretical maximum for the objective function. If the control parameter is just high enough to accept the configuration with this maximum energy, then it follows that the temperature is high enough to accept any configuration. This is the technique we use to arrive at the high temperature limit for the schedule.

If t_k is the item with the largest size in the item list, then the configuration that allocates t_k alone to a bin and all the other items to another bin has the property that it is within one move[3] of the maximum energy configuration. The energy of this configuration is given by

$$C = (M\overline{T} - t_k - \overline{T})^2 + (M-2)\overline{T}^2 + (t_k - \overline{T})^2 = M(M-1)\overline{T}^2 - 2t_k(M\overline{T} - t_k), \qquad (5)$$

and the difference in energy, ΔC, between this configuration and the maximum energy configuration is given (from Lemma 1) by

$$\Delta C = 2t_k(M\overline{T} - t_k).$$

Assuming without loss of generality that there exists only one configuration with the maximum energy, an uphill move from a configuration with energy given by equation (5) will be accepted only if $e^{-\Delta C/T} \leq 1/N$ (only one out of N possible moves results in the maximum energy configuration). This gives the high temperature condition as

$$T_\infty \simeq \frac{2t_k(M\overline{T} - t_k)}{\ln(N)}. \qquad (6)$$

[3]Moves are described in Section 4.3.

1. relocation of a single randomly selected item from the bin to which it is currently allocated to a randomly selected bin, the associated change in energy being given by

$$\begin{aligned}\Delta C &= (B_{a_i} - t_i - \overline{T})^2 + (B_j + t_i - \overline{T})^2 - (B_{a_i} - \overline{T})^2 - (B_j - \overline{T})^2 \\ &= 2\,t_i\,(B_j - B_{a_i} + t_i);\end{aligned} \tag{8}$$

2. randomly selecting two items currently allocated to two different bins and exchanging their positions, the associated change in energy in this case being

$$\begin{aligned}\Delta C &= (B_{a_i} - t_i + t_j - \overline{T})^2 + (B_j + t_i - t_j - \overline{T})^2 - (B_{a_i} - \overline{T})^2 - (B_j - \overline{T})^2 \\ &= 2\,(t_j - t_i)\,(B_{a_i} - B_{a_j} - t_i + t_j).\end{aligned} \tag{9}$$

At high temperatures, $e^{-\Delta C/T} \simeq 1$ and state changes involving relocation of items with large sizes are likely to be accepted. At low temperatures, $e^{-\Delta C/T} \simeq 0$ and generated configurations are likely to be accepted only when they have a smaller energy.

Typically, the choice of an acceptance criterion such as the Metropolis criterion means that the acceptance ratio[5] tends to become very small at low temperatures. To counteract this effect, our annealing schedule uses both kinds of move generation strategies above. At high temperature, we use moves of Type 1 above which coarsely optimize allocation, while at low temperatures we switch to moves of Type 2 which make finer adjustments to the objective function. We have also tested schedules which employ both kinds of moves all the time, increasing the proportion of Type 2 moves applied at lower temperatures. The simulation results indicate that this method of hybrid move generation yields only slightly better solutions when compared to schedules that employ only one kind of move generation.

4.4. Temperature Decrement

The rate at which the control parameter is varied has a profound impact on the quality of the final solution obtained by annealing. Too slow a rate wastes computational time while too fast a cooling rate quenches the system and yields local minima. The optimal cooling rate is hard to determine although there have been schedules in the literature [5] using dynamically determined temperature decrements. Typically, the temperature is decremented according to a logarithmic scheme [12]. The idea is that in the absence of a good guideline, our best bet is to ensure that the average cost decreases smoothly, thus, increasing the chance of obtaining a good annealing curve.

We will use a temperature decrement function of the form

$$\mathcal{F}(T) = \gamma T, \tag{10}$$

where γ lies in the interval $[0.9, 1.0)$. The closer it is to 0.9, the faster is the rate of cooling and the closer it is to 1.0, the slower is the cooling rate. In our simulation experiments, we have obtained very good solutions with a γ value of about 0.95.

5. SIMULATION RESULTS

We have applied the algorithm described in Section 3 to the bin-packing problem and have run extensive simulation experiments. In this section, we present a representative cross section of our results.

[5]This is the fraction of generated configurations that are accepted according to the Metropolis criterion.

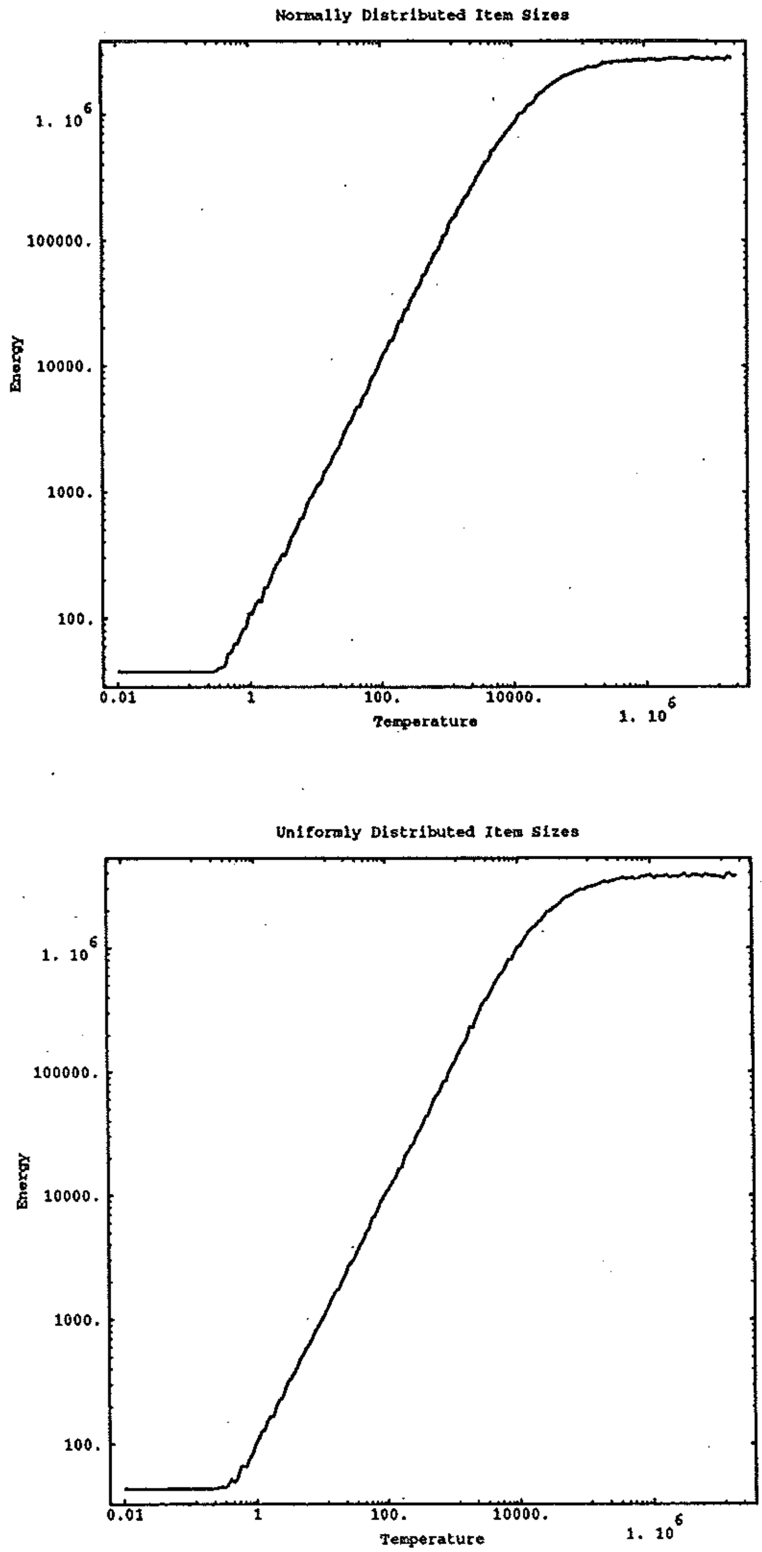

Figure 3. Annealing curves.

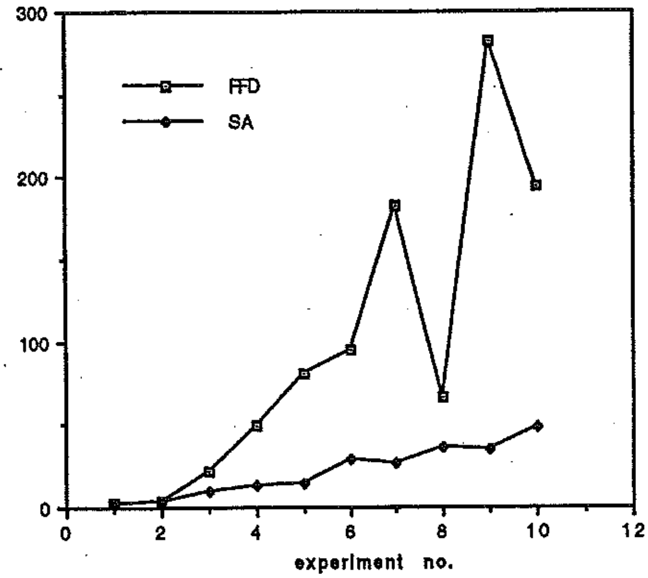

Figure 6. Performance comparison (uniform distribution).

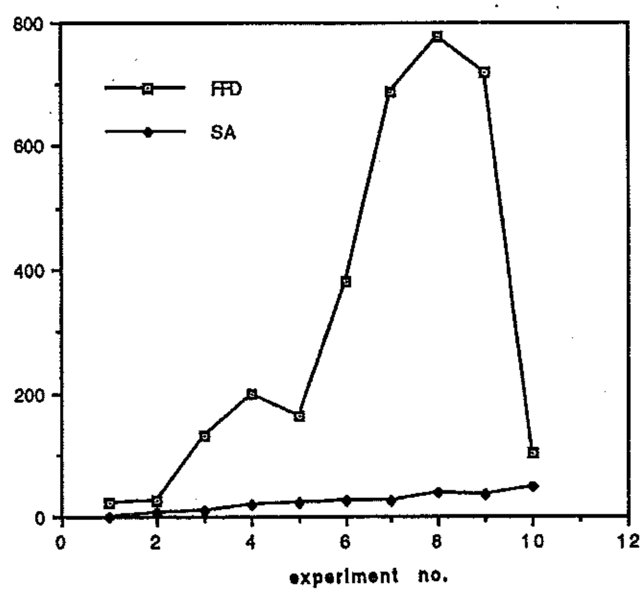

Figure 7. Performance comparison (normal distribution).