Comparative Study of Several Intelligent Algorithms for Knapsack Problem

Jing Zhang
Basic Courses Department Beijing Union University Beijing, China
e-mail: zhang1jing4@sina.com

Abstract

Knapsack problem in the field of combinatorial optimization is a typical, easy-to-describe but difficult to deal with NP-complete problems. Knapsack problem for large-scale, are not yet very effective way, how quickly and effectively solve the knapsack problem has important theoretical and practical significance, is widely used in engineering practice. We introduce the knapsack problem, discusses the current knapsack problem several intelligent algorithm is more effective (simulated annealing algorithm, ant colony algorithm, taboo search method), a simple illustration of their solution process, respectively, and pointed out their advantages and disadvantages the prospects for solving knapsack problem raised outlook.

© 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of the Intelligent Information Technology Application Research Association.

Keywords: Knapsack problem; intelligent algorithms; simulated annealing; ant colony algorithm; taboo search

1. Introduction

The problem of the backpack, KP also noted is a combinatorial optimization problem. Objects placed in the backpack must maximize the total value without exceeding the maximum weight. It models a situation analogous to filling a backpack, unable to bear more than a certain weight, with all or part of a given set of objects each having a weight and value.

This problem is at the base of the first asymmetric encryption algorithm (or "public key") presented by Martin Hellman, Ralph Merkle and Whitfield Diffie at Stanford University in 1976 [1]. However, even if the idea is due to the problem of the backpack, RSA is considered the first true asymmetric encryption algorithm. It is intensively studied since the mid-twentieth century and there are references as early as 1897, in an article by George Ballard Mathews [2]. The formulation of the problem is very simple, but its resolution is more complex. The problem of the backpack is one of 21 NP-complete problems by Richard Karp, outlined in his 1972 paper. Existing algorithms can solve instances of large size. However, the unique structure of the problem, and therefore its presence as a sub problem of other more general problems, make it a matter of choice for research.
NP-hard version of this problem has been used in primitive and cryptographic protocols, such as Merkle-Hellman cryptosystem or Chor-Rivest cryptosystem. Their advantage over asymmetric cryptosystems based on the difficulty of factoring is their speed encryption and decryption. However, the algorithm for Hellman, Merkle and Diffie is subject to the "backdoors" algorithms, which implies that it is "broken", that is to say cryptanalysis. The problem of the backpack is a classic example of misunderstanding regarding the relationship between the NP-completeness and cryptography. A revised version of the algorithm with an iteration of the problem of the backpack was then presented, soon to be broken. Its main idea is to assume that someone has a large number of items, the weight of the different. By choosing this person as part of the secret items and place them in a backpack to encrypt the message. Additional restrictions on certain conditions, given the weight, and to list the items may be in the calculation is not implemented. Total weight of the items in the backpack is open, is open all possible items, but items in the backpack is confidential. Asymmetric encryption algorithms based on the backpack were all broken so far, the latest being that of Chor-Rivest. However, most systems were cracked a backpack, so now rarely use it. Knapsack problem is not well-known computational problems, backpack system with its encryption, decryption, fast and eye-catching.

2. Simulated Annealing

Simulated annealing algorithm is a simulation of the object from the molten state to gradually cool down to the final state of the physical process of crystallization of the approximate global optimization algorithm. Simulated annealing principle is also similar to the principles and metal annealing. It uses the solution of the problem solving process and melting objects, the similarity of the annealing process, annealing process using stochastic simulation objects to complete the problem solving, that is, the control parameter (temperature) parameters under the effect of adjusting the value until it finds a can reach the global minimum of energy function parameters. We will apply thermodynamics to statistical theory, every point within the search space thought of as the air molecules. In the annealing process in order not to lose the minimum internal energy state, when the initial temperature to cool down sufficiently high rate of decrease is smaller. Molecular energy is the kinetic energy of its own; and search every point of space, but also the same as the air molecules with the "energy" to the point of the proposition that the appropriate level. Algorithm is first to search space for an arbitrary starting point: each first select a "neighbor", and then calculated from the existing location to reach, "neighbors" of probability.

Simulated annealing algorithm for knapsack problem the basic steps are as follows:

1. Select a starting point \( y \), and set a higher initial temperature \( t \), so that the number of iterations \( n = 1 \).
2. Find the objective function (energy function) \( C = f(x) \) the function value.
3. In accordance with the generating function \( h(\Delta x, t) \) to determine the probability of selection \( \Delta x \), so that the new point \( x_k = x + \Delta x \).
4. Calculate the new objective function value \( C_k = f(x_k) \). In accordance with the receiver function \( g(\Delta C, t) \) to determine the probability of \( x \) is set to \( x_k \), where, \( \Delta C = C_k - C \).
5. In accordance with the timetable for lower annealing temperature \( t \). Increase the number of iterations \( k \), if \( k \) reaches the maximum number of iterations, stop iteration; otherwise return to Step (3).

Solutions calculated with the new objective function corresponding to worse. By a generating function from the current solution to generate a new solution in the solution space; order to facilitate the subsequent calculations and accepted method to reduce time-consuming, often choose a new solution by the current transformation through a simple method to generate new solutions, such as the composition Solutions for all or part of the new elements replacement, exchange, etc., noted that the transformation method to generate new solutions of the decision of the current structure of the new neighborhood solution, and thus the selection of the cooling schedule to a certain extent. The fact that, for most applications, this is the objective function calculated the fastest way worse. Because the objective
function difference is only part of the transformation generated, so the calculation of the objective function difference is best calculated by increments.

When the new solution was identified accepted, instead of the current solution with the new solution, which simply corresponds to the current solution to generate new solutions in the transform part of the time to be achieved, while the objective function value can be modified. Determine whether the new solution is accepted, the judge is based on an acceptance criteria, the most commonly used criteria for the acceptance criteria Metropolis. When the new solution is judged to be discarded, then the original basis of the current solution to continue to the next round of tests. At this point, the current solution to achieve iteration. On this basis, can start the next round of tests.

3. Ant Colony Algorithm

Ant colony algorithm is a novel simulated evolutionary algorithm, by Italian scholars Dorigo M. et al 90 first proposed in [7], which simulated ants find food in the searching embodied ability to solve practical difficulties in the optimization problem. Ants are most common on Earth, one of the largest numbers of insect species, often in groups in the daily life of the human environment. These groups of insects, biological intelligence features, attracted the attention of some scholars. Italian scholar M. Dorigo, V. Maniezzo and others in observing the foraging behavior of ants found the ant nest and always find the shortest path between food sources. Ant chemical communication is the basic information exchange to take one of the ways the living habits of the ants play an important role. The study found that ants of the collaboration features of such groups are left in their travel through a path called pheromone of the volatile chemicals to the communication and coordination. Ant path more traveled, the higher the concentration of pheromone left. And when the time of finding a path of ants was will choose the greater chance of pheromone doctrinal path. Foraging behavior of ants through the study, they found that the ant colony pheromone is through this collaboration with each other to form a positive feedback, so that the number of ants on the path are gradually piece together the shortest path to the. Ants in the feeding process through which the road will be left in different concentrations of pheromone, while the longer path to lower the concentration of pheromone while pheromone which will be volatile over time. The process to form a positive feedback, this way you can eventually find an optimal path. The basic idea of ant colony algorithm:

(1) In the initial state, a group of ants to go out, no information at this time element, then each will randomly select a path.

(2) The next state, each ant to reach the different points, these points from the initial point to the left between the pheromone, the ants continue to walk, the ant has reached the target return, while the next batch of ants out, they will follow the path of each selected line the amount of pheromone, pheromone prefer more direct path.

(3) Then went to the next state, just not on the route through which the ant pheromone evaporates different degrees, while the line just after the ant pheromone enhanced. Then sent a group of ants, repeat step 2.

Ant colony algorithm is that it also has adaptive, self-organization, positive feedback, essentially parallel features, but the search time is longer, and easy because the process by searching a better solution of the early detection of the impact of falling into local optimal solution, so the search stagnation. Each state to the next change of state called the first iteration after the iteration times; there will be a pheromone on a path much more than the other path, which usually is an optimal path.

4. Taboo Search

Tabu Search (or taboo Search) is a sub-heuristic search technique [8], by Glover first proposed in 1986, thus the formation of a complete set of algorithms [9, 10]. TS algorithm by introducing a flexible storage
structure and the corresponding tabu search criteria to avoid the roundabout, and pardon some of the criteria by contempt fine was taboo status, thus ensuring effective diversification to explore the final global optimization. It is a local neighborhood search for an extension, is a global optimization algorithm step by step, the process is a simulation of human intelligence.

The basic idea is simple TS algorithm is: given a current solution (initial solution) and a neighborhood, and then in the neighborhood of the current solution to determine a number of candidate solutions; if the best candidate solution is better than the target value corresponding to "best so far "condition, the neglect of its taboo characteristics, with its alternative solutions and the current" best so far "condition, and the corresponding object into the tabu list, tabu list also changes the term of office of each object. If there is the candidate solution, candidate solution is chosen to select the best non-taboo for the new current solution, while ignoring the merits of it with the current solution, while the corresponding object into the tabu list, tabu list and modify the term of office of each object. Neighborhood function, tabu objects, tables and contempt tabu criteria tabu search algorithm constitute the key. The iterative search process until the stopping criteria to meet.

Tabu search algorithm based on the general design principle, a typical combinatorial optimization problem knapsack problem, the algorithm can be achieved by the following programs:

1. Randomly generated problems can also be based on information generated through a number of heuristic methods in order to guarantee the initial performance.

2. If a state with better performance than the "best so far" condition, the neglect of its taboo property, select it directly to the current state.

3. The length of proposed adaptive method attempts, for example, the situation according to the updated target information to the appropriate frequency or taboo to increase or reduce the taboo list length.

4. Common method is to swap, insert, reverse and other operations. Take the current solution is usually the neighborhood of a subset of the solution set as the candidate solution set, and take them to meet the criteria for contempt or taboo status as the best candidate for the optimal solution.

5. Steps were used in a certain number of iterations re-initialized based on the best state and its neighborhood chemotaxis multi-step re-search and random initialization of the algorithm or frequency information is based on some known object punishment. Given the optimal state of maximum sustainable consecutive iterations remain the same steps.

Numerous studies show that the tabu search algorithm simulated annealing performance quite intelligent optimization algorithms, or even superior.

5. Comparison of Advantages and Disadvantages of Various Methods

Since the typical knapsack problem in the past few decades a great many people to study the solution method. In addition to solving the above described methods and their various modified methods, there are some other ways to solve the knapsack problem can be, for example: greedy algorithms, n-opt method, hill climbing, backtracking, branch and bound, EP, chaotic search, fuzzy optimization.

Ant colony algorithm is proposed by the early 90s was only a new algorithm, the hybrid optimization strategy due to lack of rigorous and extensive analysis of theoretical and efficiency is only in recent years the development and application. Generally speaking, simulated annealing algorithm experimental performance with high quality, robust initial, common and easy realization of the advantages, the biggest drawback is that the optimization process is often long. Taboo search algorithm is a strong local search ability of the global iterative optimization algorithm; the deficiency is a strong dependence on the initial solution and serial iterative search process.

Nature of ant colony algorithm is a parallel algorithm, but its search time comparison long, also often trapped in local optimal solution, making the search stagnation. Since few papers knapsack problem to provide some specific examples of computing time, but only reported the number of function evaluation,
which makes comparison between different methods slightly difficult because of the different operators have different time complexity.

6. Conclusions

In this paper, a classical NP-complete knapsack problem and its variants, made a summary of the research. In recent years, researchers have tried to use various methods to solve knapsack problem, however, because of the knapsack problem to deepen the understanding of properties, replaced by a variety of approximate methods; trying to use the exact algorithm of knapsack problem the basic disappeared, combined with the use of a variety of methods of study gradually occupy the mainstream, attempt knapsack problem using a single method for solving the problem of the decline. Throughout recent years, research results, the researchers mainly used the following methods were studied on the knapsack problem: use a variety of pure mathematics to construct polynomial time complexity of the approximate algorithm. Neural networks and self-organization chart as a self-learning, associative memory function and high-speed optimal solution the ability to use them, and the combination of other methods by the researcher’s attention. Because the ant colony algorithm has strong search ability groups, but at the same time there is likely to fall into local optimum, so researchers usually other search algorithms and the combination of these algorithms to construct a more efficient hybrid algorithm. The use of ant colony algorithms simulated annealing, taboo search algorithms.

Knapsack problem as the practical application of a wide range of backgrounds, coming quite a long time algorithm for knapsack problem will remain a hot issue in the field. But unless there is a new algorithm for solving combinatorial optimization problems framework for the emergence of all kinds of imitation of nature combined with local optimization algorithm ideas will continue to research. Imitation of nature in the current algorithm used for the study of knapsack problem, solutions that approach seems to have become a bottleneck limit break, if we can design a new solution easier to produce solutions of better representation, and accordingly design solving method is a breakthrough algorithm for knapsack problem. Practical problems with the design of appropriate and local operators of optimization strategies to construct the hybrid algorithm will continue to be an important way to solve the knapsack problem.

Acknowledgment

The work is supported by Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality (PHR (IHLB)). Thanks for the help.

References

[1] R.C. Merkle et M.E. Hellman, "Hiding Information and Receipts in Trap Door Knapsacks", Internal Symposium on Information Theory, Cornell University, Ithaca, New York, October 1977.
[2] G.B. Mathews, On the partition of numbers, Proceedings of the London Mathematical Society, 28:486-490, 1897.
[3] A. Shamir, A Polynomial-Time Algorithm for Breaking the Basic Merkle-Hellman Cryptosystem, IEEE Transactions on Information Theory, Vol. IT-30, pp. 699-704, 1984.
[4] Skiena, S. S. (September 1999). "Who is Interested in Algorithms and Why? Lessons from the Stony Brook Algorithm Repository". AGM SIGACT News 30 (3): 65–74.
[5] Gallo, G.; Hammer, P. L.; Simeone, B. (1980). "Quadratic knapsack problems". Mathematical Programming Studies 12: 132–149.
[6] G. Plateau, M. Elkihel, A hybrid algorithm for the 0-1 knapsack problem, Methods of Oper. Res., 49:277–293, 1985.
[7] M. Dorigo, V. Maniezzo, A. Colomi. The ant system: Optimization by a colony of cooperation agents. IEEE Transactions on Systems, Man, and Cybernetics Part B, 1996, 26(1):29–41.

[8] F. Glover. Future paths for integer programming and links to artificial intelligence. Computers and Operations Research. 1986, 13: 533–549.

[9] F. Glover. Tabu Search: part I. ORSA Journal on Computing, 1989, 1: 190–206.

[10] F. Glover. Tabu Search: part II. ORSA Journal on Computing, 1990, 2: 4–32.

[11] Ferreira AG. Efficient parallel algorithms for the knapsack problem. In: Cosnard M, Barton MH, Vanneschi M, ed. Proceedings of the IFIP WG 10.3 Working Conference on Parallel Processing. North-Holland: Elsevier Science Publishers, 1988. 169–179.

[12] S. Martello, D. Pisinger, P. Toth, Dynamic programming and strong bounds for the 0-1 knapsack problem. Management Sci., 45:414–424, 1999.

[13] Dantchev S. Improved sorting-based procedure for integer programming. Mathematical Programming, Serial A, 2002,92:297–300.

[14] Pisinger, D. 2003. Where are the hard knapsack problems? Technical Report 2003/08, Department of Computer Science, University of Copenhagen, Copenhagen, Denmark.

[15] Chang HK-C, Chen JJ-R, Shyu S-J. A parallel algorithm for the knapsack problem using a generation and searching technique. Parallel Computing, 1994,20(2):233–243.

[16] S. Martello, P. Toth, A mixture of dynamic programming and branch-and-bound for the subset-sum problem, Management Sci., 30:765–771.

[17] Feuerman, Martin; Weiss, Harvey (April 1973). "A Mathematical Programming Model for Test Construction and Scoring". Management Science 19 (8): 961–966.

[18] Ferreira AG. A parallel time/hardware tradeoff T(H=O(2n/2) for the knapsack problem. IEEE Transactions on Computing, 1991, 40(2):221–225.

[19] Chor B, Rivest RL. A knapsack-type public key cryptosystem based on arithmetic in finite fields. IEEE Transactions on Information Theory, 1988,34(5):901–909.

[20] S. Martello, P. Toth, Knapsack Problems: Algorithms and Computer Implementation, John Wiley and Sons, 1990.

[21] Schroeppel R, Shamir A. A T=O(2n/2), S=O(2n/4) algorithm for certain NP-complete problems. SIAM Journal of Computing, 1981, 10(3):456–464.

[22] Mathews, G. B. (25 June 1897). "On the partition of numbers". Proceedings of the London Mathematical Society 28: 486–490.