

Simulated Annealing Articles summaries

William Bezuidenhout

Tuesday, 25 May 2010

An Unsupervised Intrusion Detection Method Combined Clustering with Chaos Simulated Annealing

```
@INPROCEEDINGS{4370702,
author={Lin Ni and Hong-Ying Zheng},
booktitle={Machine Learning and Cybernetics, 2007 International Conference on},
title={An Unsupervised Intrusion Detection Method Combined Clustering with Chaos Simulated Annealing},
year={2007},
month={19-22},
volume={6},
number={},
pages={3217 -3222},
keywords={chaos simulated annealing algorithm;computer networks security; hardware failures;malicious attacks; near-optimal partitioning clustering; optimization technique;software flaws;tentative probing; unsupervised clustering intrusion detection method;chaos;computer networks;pattern clustering; security of data;simulated annealing;telecommunication security;},
doi={10.1109/ICMLC.2007.4370702},
ISSN={},}
```

Simulated Annealing is a powerful optimization technique that attempts to find a global minimum of a function using concepts borrowed from Statistical Mechanics. The algorithm was originally intended for simulating the evolution of a solid in a heat bath to thermal equilibrium. As it was first described, the algorithm starts with a “substance” composed of many interacting individual molecules arranged in a random fashion. Then, small random perturbations to the structure of the molecules are attempted, and each perturbation is accepted with a probability based on the associated

“energy” increase, ΔE . If ΔE is at least 0, then the perturbation is accepted with probability $\exp(\frac{\Delta E}{T})$. If ΔE is less than 0, then the perturbation is accepted with probability 1. Eventually, after a large number of trial perturbations, the energy settles to equilibrium for the temperature. SA uses both the high- and the low-temperature properties of the Metropolis algorithm to find low energy, regardless of the initial structure. The cooling is made slow to overcome the high dependence of low temperature equilibrium energies on the initial state. Simulated annealing exploits the obvious analogy between process annealing and combinatorial optimization problems, where the “molecules” are the variables in the data structure and the “energy” function is the objective function.

A Simulated Annealing Algorithm with Constant Temperature for Discrete Stochastic Optimization

A Simulated Annealing Algorithm with Constant Temperature for Discrete Stochastic Optimization

Author(s): Mahmoud H. Alrefaei and Sigrn Andradttir

Source: Management Science, Vol. 45, No. 5 (May, 1999), pp. 748-764

Published by: INFORMS

Stable URL: <http://www.jstor.org/stable/2634729>

Accessed: 09/03/2010 05:10

Kirkpatrick et. al (1983) and Cerny (1985) introduced the idea of simulated annealing to the field of combinatorial optimization. Their procedure resembles the model of Metropolis et al. (1953) for how in the (physical) annealing process, particles of a solid arrange themselves into thermal equilibrium at a given temperature. More specifically, consider the optimization problem (1). Suppose that exact objective function values are available and that the objective function may have multiple local optimal solutions. Simulated Annealing allows “hill-climbing” moves in order to avoid local optimal solutions. This method starts with an arbitrary state in φ is the estimated optimal solution. If $x \in \varphi$ is the estimated optimal solution in iteration k , then a new state x' is randomly selected from a neighborhood of x , $N(x)$. if $f(x') \leq f(x)$, so that x' is better state than x , then the method accepts the move from x to x' , and x' becomes the new estimated optimal solution. On the other hand, if $f(x') > f(x)$, then there is a chance that x' will be accepted even though it is a worse state than x ; this is because good points might be “hidden behind” x' . More specifically, a uniform random number

$U_k \sim U[0, 1]$ is generated (i.e, U_k is uniformly distributed on the interval $[0, 1]$). If $U_k \leq \exp[-[f(x') - f(x)]^+/T_k]$, where $T_k > 0$ is a controlling parameter called “temperature” and $y^+ = \max\{0, y\}$ for all $y \in \mathbb{R}$, the state x' is accepted as the new estimate of the optimal solution; otherwise x remains the estimate of the optimal solution. Note that the higher the temperature T_k , the more likely it is that a “hill-climbing” move from x to x' with $f(x') > f(x)$ is accepted. The simulated annealing algorithm assumes $T_k > 0$ for all k and that $T_k \rightarrow 0$ as $k \rightarrow \infty$. The initial temperature T_0 and the rate at which the sequence $\{T_k\}$ (the annealing schedule) is decreased play important roles in the convergence of the simulated annealing algorithm and the quality of the final estimate of the optimal solution.

Most authors studying simulated annealing have focused on determining an appropriate annealing schedule in order to obtain convergence in probability to the set of global optimal solutions φ^* . Hajek (1988), Theorem 1, shows that for annealing schedules $\{T_k\}$ of the form

$$T_k = \frac{C}{\ln(1+k)}, \forall k \in \mathbb{N} \quad (1)$$

a necessary and sufficient condition on C for the algorithm to converge in probability to the set φ^* is that $C \geq d^*$, where d^* is the maximum depth of the local optimal solutions that are not globally optimal. Earlier results on the convergence of simulated annealing are given by Geman and Geman (1984) and by Mitra et al. (1986).

The original simulated annealing algorithm assumes that the objective function values can be evaluated exactly. However, in my practical, it is impossible to evaluate the objective function exactly, and instead the valuation of the objective function values may include some noise. This is, for example, the case when values of the objective function are the expected performance of a complex stochastic system under different system configurations and are estimated using simulation. Despite the wide use of simulated annealing procedure can be viewed as an ordinary hill-descent method with artificial noise added.

A simulated annealing approach for curve fitting in automated manufacturing systems

Title: A simulated annealing approach for curve fitting in automated manufacturing systems

Author(s): Hsien-Yu Tseng, Chang-Ching Lin

Journal: Journal of Manufacturing Technology Management

SA is a stochastic search technique. It is designed to lead jumping out local optimum during the search process. SA is a very effective combinatorial optimization method, which is successfully applied to VLSI design, schedule, plant layout and etc. and the combinatorial optimization problem related to production while extending its application to the continuous variable optimization problem.

Similar to statics mechanism, the search process of SA is also operated in accordance with transition probability. This transition probability depends on control temperature and the changing amount of objective function. Since, SA possesses stochastic search policy, it can be “uphill” move by using the solution of a larger objective function as the present solution. Moving the present solution to an inferior solution under a controlled probability may allow SA jump out local optimum and may be a better “downhill” path will be found, which further obtains a better solution. The “uphill” move is controlled carefully by the temperature. When temperature is too high, “uphill” move probability will increase; when temperature decreases gradually, “uphill” move probability will decrease accordingly. Kirkpatrick demonstrated that SA is capable of obtaining the real optimum under a very long search time. In actual execution, this real optimum is not obtainable due to the restriction of calculation time, but an approximation of the real optimum can be obtained.

When SA is applied to one problem, four basic components should be defined. These are:

- *Configuration* — represents the possible solution of problem.
- *Move set* — Is an allowable move, which can let us achieve all feasible configuration, this set have to be easy for calculation.
- *Cost Function* — Is used to measure the quality of configuration
- *Cooling schedule* — Sets the initial temperature and the cooling regulation to determine when and how many degrees to decrease the present temperature, and when to end the annealing.

The SA method that solves the continuous function problem proposed by Corana et al. (1987) faces a tough problem that needs long copious calculation. The curve fitting optimization model mentioned in the previous

section is immense and complicated. The method mentioned by Corana et al. is not suitable to apply to the optimization model solving. So to decrease the SA calculation demand, the PSSA optimization method is proposed in this current paper uses PS as the move generating function to accelerate the search process.

0.0.1 Patter search

The PSSA optimization method recommended in this paper uses Pattern Search(PS) as move generation mechanism. The operation of PS includes two kinds of move: exploratory move and pattern move. Exploratory move examines the local behaviour of function, and finds out the direction of “down-hill” path. Pattern move utilizes the information generated by exploratory move to reach the valley rapidly.

Exploratory move starts search from initial point and moves along, the axial direction of each variable according to step size, and decides whether neighboring move is set as present solution in accordance with objective function. If setting neighboring solution as present solution is unacceptable (no improvement in objective function), it will search the opposite direction. During exploratory move, each move only refers to one variable, and explores each variable in proper order to check whether there is improvement in the objective function. If there is an improvement, it means that a direction “downhill” path (or pattern direction) exists. After an exploratory move ends, it will check whether a pattern direction exists. If it does exist, make the pattern move according to this pattern direction and this move can rapidly reach the minimum of the function (the valley) and continue executing the pattern move until no objective function improves anymore.

Changing step size or not depends on pattern move existence after each exploratory move ends. If a pattern direction does not exist, then change the step size. PS uses exploratory move and pattern move repeatedly until the result condition is satisfied.

There are many cooling schedules discussed in the literature, and geometric cooling schedule is very fast and very effective

A Survey of Simulated Annealing as a Tool for Single and Multiobjective Optimization

@article{2006,
jstor_articletype = {primary_article},

```

title = {A Survey of Simulated Annealing as a Tool for Single and Multiobject
author = {Suman, B. and Kumar, P.},
journal = {The Journal of the Operational Research Society},
jstor_issuetitle = {},
volume = {57},
number = {10},
jstor_formatteddate = {Oct., 2006},
pages = {1143--1160},
url = {http://www.jstor.org/stable/4102365},
ISSN = {01605682},
language = {},
year = {2006},
publisher = {Palgrave Macmillan Journals on behalf of the Operational Research
copyright = {Copyright 2006 Operational Research Society},
}

```

Simulated Annealing (SA) is a compact and robust technique, which provides excellent solutions to single and multiple objective optimization problems with a substantial reduction in computation time. It is a method to obtain an optimal solution of a single objective optimization problem and to obtain a Pareto set of solutions for a multi-objective optimization problem. It is based on an analogy of thermodynamics with the way metals cool and anneal. If a liquid metal is cooled slowly, its atoms form a pure crystal corresponding to the state of minimum energy for the metal. The metal reaches a state with higher energy if it is cooled quickly. SA has received significant attention in the last two decades to solve optimization problems, where a desired global minimum/maximum is hidden among many poorer local minima/maxima. Kirkpatrick et al (1983) and Cerny (1985) showed that a model for simulating the annealing of solids, proposed by Metropolis et al (1953), could be used for optimization problems, where the objective function to be minimized corresponds to the energy of states of the metal. These days SA has become one of the many heuristic approaches designed to give a good, not necessarily optimal solution. It is simple to formulate and it can handle mixed discrete and continuous problem with ease. It is also efficient and has low memory requirement. SA takes less CPU time than genetic algorithm (GA) when used to solve optimization problems, because it finds the optimal solution using point-by-point iteration rather than a search over a population of individuals.

Initially, SA has been used with combinatorial optimization problems. Maffioli (1987) showed that SA can be considered as one type of randomized

heuristic approaches for combinatorial optimization problems.

In SA, instead of this strategy, the algorithm attempts to avoid being trapped in a local minimum by sometimes accepting even the worse move. The acceptance and rejection of the worse move is controlled by a probability function. The probability of accepting a move, which causes an increase δ in f , is called the acceptance function. It is normally set to $\exp(\frac{-\delta}{T})$, where T is a control parameter, which corresponds to the temperature in analogy with the physical annealing. The acceptance function implies that the small increase in f is more likely to be accepted than a large increase in f . When T is high most uphill moves will be rejected. Therefore, SA starts with a high temperature to avoid being trapped in local minimum. The algorithm proceeds by attempting a certain number of moves at each temperature and decreasing the temperature.

Like other heuristic optimization techniques, there is a chance of revisiting a solution multiple times in SA as well. It leads to extra computational time without any improvement in the optimal solution. Mingjun and Huanwen (2004) have proposed chaos simulated annealing (CSA) by introducing chaotic systems to SA. The CSA is different from SA as chaotic initialization and chaotic sequences replace the Gaussian distribution. CSA is more likely to converge to the global optimum solution because it is random, stochastic and sensitive on the initial conditions. It has been shown that CSA improves the convergence and is efficient, applicable and easy to implement.

0.0.2 Annealing Schedule

The setting of the parameters for the SA-based algorithm determines the generation of the new solution. The precise rate of cooling is an essential part of SA as it determines the performance of the SA-based algorithm. A high cooling rate leads to poor results because of lack of the representative states, while a low cooling rate requires high computation time to get the result. The following choices must be made for any implementation of SA and they constitute the annealing schedule: initial value of temperature (T), cooling schedule, number of iterations to be performed at each temperature and stopping criterion to terminate the algorithm.

Initial value of temperature (T) Initial temperature is chosen such that it can capture the entire solution space. One choice is a very high initial temperature as it increases the solution space. However, at a high initial temperature, SA performs a large number of iterations, which may be without giving better results. Therefore, the initial temperature is chosen by

experimentation depending upon the nature of the problem. van Laarhoven et al (1988) have proposed a method to select the initial temperature based on the initial acceptance ratio χ_0 , and the average increase in the objective function, $\Delta f_0 : T = -\frac{\Delta f_0}{\ln(\chi_0)}$ where χ_0 is defined as the number of accepted bad moves divided by the number of attempted bad moves. A similar formula has been proposed by Sait and Youssef (1999) with the only difference being the definition of χ_0 . They have defined χ_0 as the number of accepted moves divided by the number of attempted moves. A simple way of selecting initial temperature has been proposed by Kouvelis and Chiang (1992). They have proposed to select the initial temperature by the formula $P = \exp(-\frac{\Delta s}{T})$