AN EXACT RANDOM NUMBER GENERATOR FOR VISITING DISTRIBUTION IN GSA

JYHJENG DENG, CHUNLAN CHANG AND ZHONGXIAH YANG

Industrial Engineering and Technology Management Department, DaYeh University,
112 Shan-Jiau Rd., Da-Tsuen, Changhua, Taiwan 51505.
E-mail: jdeng@mail.dyu.edu.tw

Abstract: Global optimization is widely used in various aspects both in industry and academia. Although different methods in operations research have been proposed to solve the global optimization, simulated annealing and its variations have been favored for their simplicity in the implementation. In order to reach the global optimum from the current solution, a walking mechanism must be planned so that global optimum can be reached stochastically with high probability. The walking mechanism is the stochastic rule to find the next solution from the current solution. One of the most promising walking mechanisms is represented as Tsallis distribution. In this study, an exact simulation method is proposed for the Tsallis distribution that is used in performing generalized simulated annealing (GSA), a global optimization algorithm suggested by Tsallis (1996). First, the Tsallis variate (random variable) is transformed linearly into a standard Tsallis variate with one parameter q_v. Then the structure of a t distribution variate is imitated to obtain the standard Tsallis variate. A ratio of a standard normal variate over a square root function of the Gamma variate with proper chosen parameters' values renders the standard Tsallis distribution variate. This method is suitable for all q_v in the range 1 to 3. Theoretical justification is provided in supporting the proposed method. A numerical example, given originally by Tsallis and Stariolo, on the GSA is studied on the effect to which the q_v has on the convergence to the optimal solution. A conclusion can be drawn from the results that the development of the exact simulation method is crucial in obtaining the correct results by using GSA.

***Keywords*:** Methodology; Optimization; Simulation; Statistics

1. INTRODUCTION

Simulated annealing, derived from the metallurgical annealing, is a simple tool to search the global optimum. It is a stochastic method with the guarantee, in the probability sense, that the global optimum can be found. However, the search time is infinite. That is the convergence to the global optimum could be very slow for complex problems. One of the key factors which influence the convergence is the walking mechanism. A walking mechanism is a procedure to search the next possible solution from the current solution. If the walking mechanism can allow the current solution to jump very far in a wise way, then it is possible to reach the global optimum in a very short time. The walking mechanism can be described in a probability density function (pdf) with variable x representing the jump distance. One of the most promising probability density functions which control the move of the walker in the randomly search is the Tsallis distribution. Although Tsallis distribution in combination with generalized simulated annealing (GSA) has been proposed by C. Tsallis and D. Stariolo in 1996 and has been proved to have a better convergence to the global optimum in a theoretical way, a working methodology to implement the GSA is lacking. This is due to the fact only ad-hoc methods are available to simulate the Tsallis distribution. In order to use the GSA to its fullest capacity, a theoretical sound algorithm must be provided to simulate the Tsallis distribution.

This paper is a follow-up study to an article published in 2004 (Deng *et al*, 2004) with the aim to provide a theoretical sound algorithm. The Tsallis distribution was first introduced by C. Tsallis and D. Stariolo (1996). They unified the classical ("Boltzmann machine") simulated annealing (CSA) (Kirkpatrick *et al*, 1983) and fast ("Cauchy machine") simulated annealing (FSA) (Szu *et al*, 1987) using the generalized thermostatistics (Tsallis, 1988) with an extra benefit of providing an algorithm which is even faster than the FSA. The GSA mechanics consists of three parts: (1) the visiting distribution, (2) the acceptance probability, and (3) the cooling schedule. The visiting distribution is the walking mechanism to allow the walker walk in a wise way in the solution space. The characteristic of a good walking mechanism is that when the annealing temperature is high, the walker should be able to jump very far from the current solution; whereas when the temperature approaches to zero the solution will converge to the global optimum. In this way, no matter where the initial solution is the walking mechanism is able to locate the optimal solution region in a very short time and eventually converges to the global optimum. The acceptance probability is the rule to accept a new move after the visiting distribution is performed. It is always good to accept a "better", in the sense of lower objective function, solution. However when a poor solution is encountered, care must be taken to decide whether it is accepted or rejected. The acceptance of a poor move is not always allowed, this is where the acceptance probability comes in. The reason to accept a poor solution is to allow the walker to jump out of the local optimum. Finally the cooling schedule determines the timing to cool down the annealing temperature. The higher the temperature, the easier it is to jump farther; whereas the lower the temperature, the walker becomes less active and it just jumps locally around the neighborhood of the current solution. When the temperature reaches zero, the walker stops jumping. A good cooling schedule will save the unnecessary moves around the local optimum region and can shorten the simulation time required to reach the global optimum. The visiting distribution gives [1,2]

$$g_{q_v}(\Delta x_t) = \left(\frac{q_v - 1}{\pi}\right)^{\frac{D}{2}} \times \frac{\Gamma\left(\frac{1}{q_v - 1} + \frac{D-1}{2}\right)}{\Gamma\left(\frac{1}{q_v - 1} - \frac{1}{2}\right)} \times \frac{\left[T_{q_v}^{V}(t)\right]^{\frac{-D}{(3-q_v)}}}{\left\{1 + (q_v - 1)\frac{(\Delta x_t)^2}{\left[T_{q_v}^{V}(t)\right]^{\frac{2}{(3-q_v)}}}\right\}^{\frac{1}{(q_v-1)} + \frac{(D-1)}{2}}} \tag{1}$$

where Δx_t is the trial jump distance from location x_t to the next location x_{t+1}, i.e., $\Delta x_t = x_{t+1} - x_t$, D is the degree of dimension of Δx_t, q_v is a parameter of visiting distribution and is in the range of (1.0, 3.0), $T_{q_v}^{V}(t)$ is the positive controlling temperature at time t, and $\Gamma(x)$ is the Gamma function. When D=1, it shows that in the limit of $q_v \to 1$, the visiting distribution becomes a normal distribution with mean equal to zero and variance given by $T_{q_v}^{V}(t)/2$ (Deng et al, 2004). However, when $q_v = 2$, the visiting distribution reduces to the Cauchy distribution of

$$g_2(\Delta x_t) = \frac{1}{\pi} \times \frac{T_{q_v}^{V}(t)}{(T_{q_v}^{V}(t))^2 + (\Delta x_t)^2} \tag{2}$$

The second part of the GSA is the probability of accepting a new move (Tsallis et al, 1996 and Franz et al, 2003)

$$P_{q_A}(x_t \to x_{t+1}) = \begin{cases} 1, & \text{if } \Delta E_t \le 0 \\ \left(1 - (1 - q_A)\frac{\Delta E_t}{T_{q_A}^{A}(t)}\right)^{1/(1-q_A)}, & \text{if } \Delta E_t > 0 \text{ and } (1 - q_A)\frac{\Delta E_t}{T_{q_A}^{A}(t)} \le 1 \\ 0, & \text{if } \Delta E_t > 0 \text{ and } (1 - q_A)\frac{\Delta E_t}{T_{q_A}^{A}(t)} > 1 \end{cases} \tag{3}$$

depending on two additional positive parameters $q_A \in R \setminus \{1\}$ and positive $T_{q_A}^{A}(t)$. The temperature parameter $T_{q_A}^{A}(t)$ in (3) can differ from its counterpart $T_{q_v}^{V}(t)$ in (1); similarly for the parameters q_A and q_v. For the sake of simplicity, these parameters are treated respectively as identical. I.e., $T_{q_A}^{A}(t) = T_{q_v}^{V}(t) = T_q(t) \quad \forall t$ where the subscript q in $T_q(t)$ could be q_v or q_A depending on the context. When the context is clear, $T_q(t)$ can be further simplified to T. As for q_A=1, equation (3) is not defined. But one can show that in the limit of $q_A \to 1$, the acceptance probability (3) converges to the Metropolis probability (Geman *et al*, 1984). I.e.,

$$P_{Me}(x_t \to x_{t+1}) = \begin{cases} 1, & \text{if } \Delta E_t \le 0 \\ e^{-\Delta E_t / T_{qA}(t)} & \text{if } \Delta E_t > 0 \end{cases} \tag{4}$$

Last of all is the cooling schedule. For an arbitrary D, the equation is written as

$$T_{q_v}(t) = T_{q_v}(1) \times \frac{2^{q_v - 1} - 1}{(1+t)^{q_v - 1} - 1} \tag{5}$$

In the limit of $q_v \to 1$ the logarithmic cooling schedule is recovered as $T_1(t) = T_1(1)\ln 2 / \ln(1+t)$; whereas in $q_v = 2$ the reciprocal cooling schedule is recovered as $T_2(t) = T_2(1)/t$ (Szu et al, 1987).

The energy function in the GSA is assumed to be greater than or equal to zero. The goal of the GSA is to find the global minimum of the energy function. For the sake of completeness, the whole algorithm to find the global minimum of a given cost function $E(x)$ is summarized as follows (Tsallis *et al*, 1996):

(i) Designate a pair of predetermined values of (q_A, q_v) and proceed, at t = 1, with an arbitrary value x_1 and a value high enough for $T_{q_v}(1)$ and calculate $E(x_1)$.

(ii) Then randomly generate x_{t+1} from x_t according to equation (1) to determine the size of the jump Δx_t. And isotropically determine its direction. Isotropically mean the direction is randomly chosen in each parameter in the polar coordinate system when the variate is represented in the polar form.

(iii) Next, calculate $E(x_{t+1})$ and determine whether or not to accept the new move x_{t+1} by equation (3). We randomly pick a value, say v, from a (0, 1) uniform random variate. If v is less than the accepting probability in (3), we accept the new move, otherwise reject it.

(iv) Calculate the new temperature T_{q_v} using equation (5) and then go back to (ii) until the minimum value of $E(x)$ is reached within the desired precision.

To make the GSA work, an accurate and efficient random number generator is required to simulate the visiting distribution in (1). These visiting distribution functions are not analytically inversible; hence, a numerical procedure is necessary to generate random numbers (See Deng *et al*, 2004 and references cited therein for discussion of different classical simulation methods and their inappropriateness.) Since the cumulative probability distribution of the Tsallis distribution cannot be derived analytically and the numerical integration to calculate the cumulative probability distribution is very time consuming, Monte Carlo technique is not suitable for this type of application. Inspired by the simulation concept of Levy's stable process mentioned in Mantegna (1994) [8], Deng et al have made an attempt to solve this simulation problem with some progress in their previous paper. They proposed a two-step algorithm to simulate the Tsallis variate which outperformed the Tsallis and Stariolo's algorithm in terms of matching condition between the corresponding simulated probability density function (pdf) and the theoretical counterpart (it is measured by the Kolmogorov's statistic (Birnbaum, 1952 and Massey, 1951).) However, one major problem concerning Deng and et al's algorithm is that it generates a two-peak effect when $q_v \geq 2.6$, as shown in Figure 1. It shows that the empirical probability density (□) and the theoretical probability density (-) have parameters T =1 and q_v =2.6. A two-hump phenomenon of the simulated distribution stands out and jeopardizes the matching quality between the simulated probability density function (pdf) and theoretical pdf. Note that the ordinate in Figure 1 is in a logarithmic scale on the base of 10. Heading of Figure 1 indicate the proper chosen parameters values, $k(\alpha)$ and $C(\alpha)$, determined in the two-step algorithm. (Check Deng *et al*, 2004 for the meaning and determination of those parameters.) Such an inconsistency between empirical distribution and the theoretical distribution gives rise to the investigation of the current research. And the main objective of this study is to overcome the problem by providing an exact random number generator to simulate Tsallis variate.

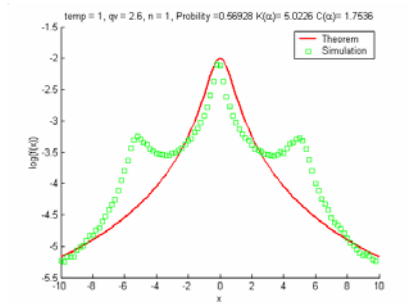

Figure 1. Empirical probability density, (□) together with theoretical probability (-), with parameters =1 and =2.6 in two-step algorithm.

2. Development of an Exact Random Number Generator

The need to develop an accurate and efficient algorithm for Tsallis distribution is evident in the number of researchers wanting to use it. See, for example, Penna (1995) [11], Okamoto and Hansmann (2001) [12] and Xiang and Gong (2002) [13]. Without such an algorithm the GSA will be useless. In order to simulate the Tsallis distribution variate, the form in (1) is first simplified. Regarding D as 1, Δx_t as x, $T_{q_v}^{V}(t)$ as T, the visiting distribution becomes

$$g(x) = \left(\frac{q_v - 1}{\pi}\right)^{\frac{1}{2}} \times \frac{\Gamma\left(\frac{1}{q_v - 1}\right)}{\Gamma\left(\frac{1}{q_v - 1} - \frac{1}{2}\right)} \times \frac{T^{\frac{-1}{(3-q_v)}}}{\left\{1 + (q_v - 1)\frac{x^2}{T^{\frac{2}{(3-q_v)}}}\right\}^{\frac{1}{(q_v - 1)}}} \qquad (6)$$

Note that the distribution has two parameters q_v and T. Let variate $y = \frac{x}{T^{\frac{1}{3-q_v}}} = g_1(x)$ be a linear transformation of variate x, then it renders $g_1'(x) = \frac{1}{T^{\frac{1}{3-q_v}}}$. Therefore, the probability density function (pdf) of y is

$$f_y(y) = \frac{f_x(x)}{|g_1'(x)|} = \frac{\left(\frac{q_v-1}{\pi}\right)^{\frac{1}{2}}}{\frac{1}{T^{\frac{1}{3-q_v}}}} \times \frac{\Gamma\left(\frac{1}{q_v-1}\right)}{\Gamma\left(\frac{1}{q_v-1}-\frac{1}{2}\right)} \times \frac{T^{\frac{-1}{3-q_v}}}{\{1+(q_v-1)y^2\}^{\frac{1}{(q_v-1)}}}$$

$$= \left(\frac{q_v-1}{\pi}\right)^{\frac{1}{2}} \times \frac{\Gamma\left(\frac{1}{q_v-1}\right)}{\Gamma\left(\frac{1}{q_v-1}-\frac{1}{2}\right)} \times \frac{1}{\{1+(q_v-1)y^2\}^{\frac{1}{(q_v-1)}}} \quad (7)$$

which is the standard Tsallis distribution with $T=1$. The term "standard" is coined because any Tsallis distribution can be transformed into a simple Tsallis distribution by a linear function. Just like an ordinary normal distribution with mean μ and standard deviation σ can be converted into the standard normal distribution with mean zero and standard deviation one. The work is then focused on the y variate, which has only one parameter q_v, and with the inverse linear transform a general Tsallis distribution with two parameters T and q_v can be obtained. Looking into both the numerator and denominator in (7), it is self evident that each one can be represented using a Gamma function. Recall that the probability density function (pdf) of a *t* variate has a similar form and its pdf is given by

$$f_t(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \times \frac{1}{\sqrt{v\pi}} \times \left(1+\frac{x^2}{v}\right)^{\frac{-(v+1)}{2}} \quad (8)$$

where *v* is the degree of freedom greater than or equal to 2 (Bain et al, 1987). Note that variate *t* is a ratio formed by standard normal variate over a square root function of an independent Gamma variate formulated as

$$t_v = \frac{Z}{\sqrt{\frac{V}{v}}} \quad (9)$$

where $Z \sim N(0,1)$ and $V \sim Gamma(2, \frac{v}{2})$ are independent of each other. Where the pdf of a $Gamma(\theta,\kappa)$ variate gives

$$f(x) = \frac{x^{\kappa-1}e^{-\frac{x}{\theta}}}{\theta^{\kappa}\Gamma(\kappa)} \quad (10)$$

with parameters $\kappa > 0$ and $\theta > 0$. The analysis of equations (7-9) indicates that a mimicking t variate, which is a ratio of a standard normal variate over a square root function of Gamma variate, can be used to represent a standard Tsallis variate. We thus obtain the following Lemma.

Lemma 1. Let $Z \sim N(0,1)$ and $V_1 \sim Gamma(\theta,\kappa)$ be two independent variates with parameters $\theta = 2$, $\kappa = v - \frac{1}{2}$ where $v = \frac{1}{q_v - 1}$. Then $T_s = \frac{Z}{\sqrt{\frac{V_1}{v}}} = \frac{Z}{\sqrt{V_1(q_v-1)}}$ is the standard Tsallis variate.

Proof:

The pdf of $Z \sim N(0,1)$ is

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}z^2}, \quad (11)$$

whereas the pdf of $V_1 \sim Gamma(2, v-\frac{1}{2})$ is

$$f_{V_1}(v_1;\theta,k) = f_{V_1}\left(v_1; 2, v-\frac{1}{2}\right) = \frac{1}{2^{v-\frac{1}{2}}\Gamma\left(v-\frac{1}{2}\right)} v_1^{v-\frac{3}{2}} e^{-v_1/2} \quad (12)$$

The random variates (RVs) *Z* and V_1 are independent of each other and with a joint density

$$f_{Z,V_1}(z,v_1) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} \times \frac{1}{2^{v-\frac{1}{2}}\Gamma\left(v-\frac{1}{2}\right)} v_1^{v-\frac{3}{2}} e^{-v_1/2} \quad (13)$$

We form two functions $T_s = \frac{Z}{\sqrt{\frac{V_1}{v}}}$, $W = V_1$ and can show that it is a one-to-one transformation from (Z, V_1) to (T_s, W). The joint density $f_{T_s,W}(t,w)$

of RVs T_s and W can be expressed as the joint density $f_{Z,V_1}(z,v_1)$, involving the Jacobian $J(z,v_1)$ of the transformation (Papoulis, 1990), as

$$f_{T_s,W}(t,w)=\frac{f_{Z,V_1}(z,v_1)}{|J(z,v_1)|} \tag{14}$$

where $J(z,v_1)$ is by definition the determinant

$$J(z,v_1)=\begin{vmatrix}\frac{\partial t}{\partial z} & \frac{\partial t}{\partial v_1}\\ \frac{\partial w}{\partial z} & \frac{\partial w}{\partial v_1}\end{vmatrix}=\begin{vmatrix}\frac{1}{\sqrt{\frac{v_1}{v}}} & -\frac{1}{2}\frac{z}{\left(\frac{v_1}{v}\right)^{\frac{3}{2}}v}\\ 0 & 1\end{vmatrix}=\frac{1}{\sqrt{\frac{w}{v}}} \tag{15}$$

Substituting (15) into (14), have

$$f_{T_s,W}(t,w)=\frac{2^{-v}e^{\left(\frac{-1}{2}\times\frac{w(t^2+v)}{v}\right)}w^{v-1}}{\sqrt{v\pi}\Gamma\left(v-\frac{1}{2}\right)},\qquad -\infty<t<\infty,0<w<\infty \tag{16}$$

Hence, using Maple (Monagan *et al*, 2001) the marginal density of variate T_s is maintained as

$$f_{T_s}(t;v)=\int_0^\infty f_{T_s,W}(t,w)dw =\frac{v^{v-\frac{1}{2}}\Gamma(v)(t^2+v)^{-v}}{\sqrt{\pi}\Gamma\left(v-\frac{1}{2}\right)}. \tag{17}$$

Rearranging (17), have

$$f_{T_s}(t;v)=\frac{v^v\cdot v^{-\frac{1}{2}}\Gamma(v)(t^2+v)^{-v}}{\sqrt{\pi}\Gamma\left(v-\frac{1}{2}\right)}=\frac{(v^{-1})^{-v}\cdot v^{-\frac{1}{2}}\Gamma(v)(t^2+v)^{-v}}{\sqrt{\pi}\Gamma\left(v-\frac{1}{2}\right)}$$

$$=\frac{v^{-\frac{1}{2}}\Gamma(v)\left(1+\frac{t^2}{v}\right)^{-v}}{\sqrt{\pi}\Gamma\left(v-\frac{1}{2}\right)}=\frac{\Gamma(v)\left(1+\frac{t^2}{v}\right)^{-v}}{\sqrt{\pi v}\cdot\Gamma\left(v-\frac{1}{2}\right)} \tag{18}$$

Substituting $v=\frac{1}{q_v-1}$ into (18), obtain

$$f_{T_s}(t;v)=\frac{\Gamma\left(\frac{1}{q_v-1}\right)}{\Gamma\left(\frac{1}{q_v-1}-\frac{1}{2}\right)}\times\sqrt{\frac{q_v-1}{\pi}}\times\left(1+(q_v-1)t^2\right)^{-\frac{1}{q_v-1}} \tag{19}$$

which is identical to the pdf of the standard Tsallis variate in (7) if the notation is changed from *t* to *y*. This completes the proof.

The method mentioned above can be called the mimicking t variate method since the structure and pdf of the t variate are mimicked to obtain the standard Tsallis variate. An example of the standard Tsallis distribution with parameter q_v =2.6 is simulated by applying Lemma 1 and its empirical probability density plotting together with the theoretical counterpart. Figure 2 shows the empirical probability density (□) and the theoretical probability density (-) match perfectly. A close-up picture of Figure 2 in the region roughly between –2 to 2 also reveals a perfect match between the empirical probability density and the theoretical probability density. For the sake of simplicity, this figure is omitted. Note that a previous two-hump effect is completely eliminated from the figure.

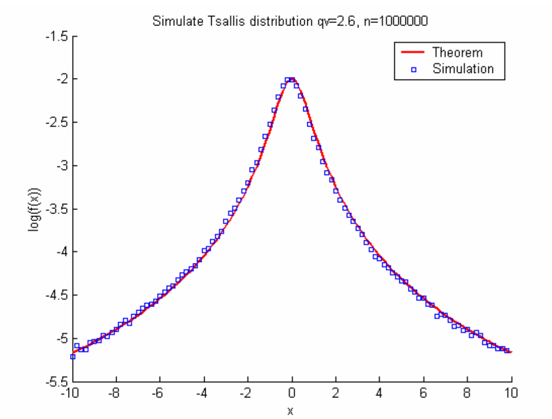

Figure 2. Empirical probability density (□) and theoretical probability (-) with parameters T =1 and q_v =2.6 by Lemma 1.

3 A Numerical Example

In this section how the values of q_v affect the speed of convergence to the optimum is studied. A simple example considering four variables to be optimized, is proposed to illustrate this effect. This example was originally proposed by Tsallis and Stariolo (1996) [2] to show how the procedures work in GSA. Therein they made comments on the best choice of q_v . This example is used again for purpose of comparing their results and ours. The energy (cost) function to be minimized is

$$E(x) = \sum_{i=1}^{4} \left(x_i^2 - 8\right)^2 + 5\sum_{i=1}^{4} x_i + E_0 \qquad (20)$$

where $E_0 \cong 57.3276$ such that $E(x) \geq 0 \;\; \forall x$. It is clear that the minimum of the function is zero. The parameters in Tsallis and Stariolo's paper are repeated here. Thus having set $T(1) = 100$, q_A =1.0 the idea is to study the effect of the q_v parameter on the speed of convergence to the optimal value. Three thresholds of 0.1, 0.01, and 0.001 are used as measures of the desired precision. That is if the value of the cost function is less than the threshold, the optimum is regarded to be reached. The goal of the study is to find out what is the best parameter value for q_v in terms of the number of iterations needed to reach the optimum. The smaller is the number of iterations, the better. Twenty samples are run for each parameter with 5000 iterations as the cap for each run. The range of q_v is [1.0, 1.1, 1.2, …, 2.7, 2.8, 2.9]. The box-plot (Vardeman, 1993) of the twenty simulated numbers of iterations for each parameter q_v is drawn and displayed in Figures 3-5. Based on the values of the mean, standard deviation of the number of iterations, therein it is seen clearly that the best choice for q_v is unanimously 2.3. This value is totally different from the best value q_v =2.7 obtained by Tsallis and Stariolo using their heuristic algorithm. The example indicates the importance of using an accurate algorithm to generate the Tsallis distribution. An inaccurate algorithm could lead to wrong conclusions.

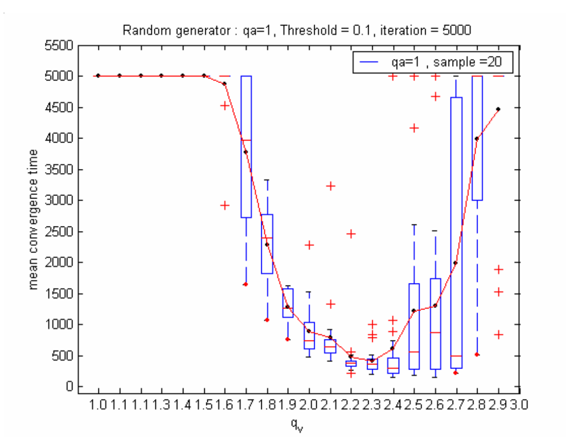

Figure 3. Box plot of the number of iterations to reach the optimum with threshold=0.1.

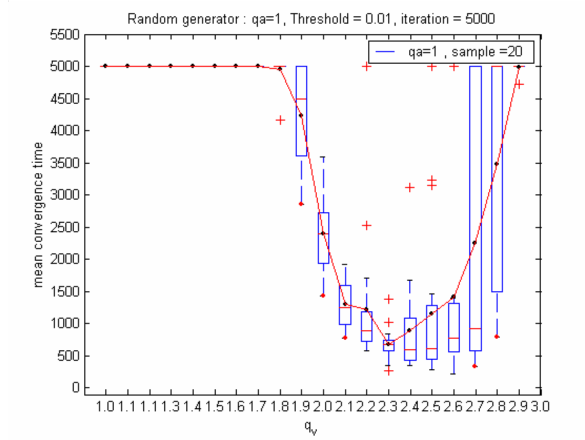

Figure 4. Box plot of the number of iterations to reach the optimum with threshold=0.01.

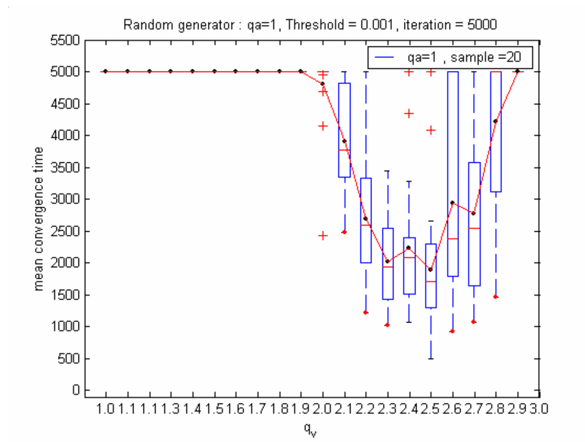

Figure 5. Box plot of the number of iterations to reach the optimum with threshold=0.001.

These three figures show that the standard deviation of the numbers of iterations is large for q_v =2.7, which indicates it is not stable to reach the optimum. For example, in the case of threshold 0.01, the smallest number of iterations is less than 500 while the largest number of iterations is 5,000. This means some of the twenty simulations the GSA could not reach the optimum within the cap limit of 5,000 iterations. In practice, it always prefers a reliable algorithm to reach the optimum, thus the value of q_v =2.7 surely is not a good choice. Another consideration worth pointing out therein is that both the classical simulated annealing (CSA), with q_v =1.0, and the fast simulated annealing (FSA), with q_v =2.0, are not suitable to reach the optimum because of the very large number of the iterations required. For the case of CSA, the average number of iterations is above 5,000 even for the roughest precision (threshold = 0.1). This means none of the energy functions in twenty simulations can reach a precision beyond 0.1 within 5,000 moves. Whereas for the case of FSA, in only four out of twenty simulations runs does the energy function reach the optimum with 0.001 accuracy. However, for the case of q_v =2.3, the energy function reaches the optimum with threshold 0.001 for all the twenty simulations. The worst result (which takes the most iteration) is one that takes less than 3,500 iterations to reach the optimum. With the proper chosen value for q_v, the out-performance of GSA over CSA and FSA is obvious.

4 Conclusions

Global optimization is widely used in a variety of applications both in the industry and in academia. Although different methods in operations research have been proposed to perform global optimization, simulated annealing and its variations have been favored for their simplicity in the implementation. In order to reach the global optimum from the current solution, a walking mechanism must be planned so that global optimum can be reached stochastically with high probability. One of the most promising walking mechanisms is represented as the Tsallis distribution. When Tsallis distribution combines with suitable acceptance probability and cooling schedule, they become a powerful tool, known as generalized simulated annealing (GSA), to solve the global optimization problem.

Although GSA is a powerful tool, its capacity has been hindered by the lacking of proper algorithm to simulate the Tsallis distribution. The purpose of this paper is to bridge this gap. A method to generate Tsallis variate has been proposed. Given the probability density function of ordinary Tsallis variate x with two parameters q_v and T as shown in equation (6), first linearly transform the Tsallis variate by $y = \frac{x}{T^{\frac{1}{3-q_v}}}$ and render a standard Tsallis variate with one parameter q_v as shown in equation (7). Where q_v is in the range of (1.0, 3.0), and T is greater than zero. Then using two independent variates Z ~ N(0,1) and $V_1 \sim Gamma(\theta,\kappa)$ with parameters $\theta = 2$, $\kappa = v - \frac{1}{2}$ where $v = \frac{1}{q_v - 1}$, it is shown that $T_s = \frac{Z}{\sqrt{\frac{V_1}{v}}} = \frac{Z}{\sqrt{V_1(q_v - 1)}}$ is the standard Tsallis variate. After generating the standard Tsallis variate, an inverse linear transform of the standard Tsallis variate can be used to restore the ordinary Tsallis variate. This method is inspired by the structure and the probability density function of a t variate. Thus this method is christened mimicking t variate method.

The significance of the proposed method is demonstrated by a numerical method which was originally proposed by Tsallis and Stariolo to show how the procedures of GSA work and the effect of q_v has on the convergence to the optimum. The energy function is shown in equation (20) and its optimum is zero. To show the effect of q_v on the convergence to the optimum, three different threshold values, 0.1, 0.01, and 0.001, are used to signal the arrival of optimum in the GSA operations. That is when the value of energy function is lower than the threshold the optimum is considered being reached. The number of iterations required to reach the optimum is recorded for twenty simulations and for each simulation the cap of 5,000 iterations is being set. The range of q_v in this study is [1.0, 1.1, 1.2, …, 2.7, 2.8 2.9]. A box-plot for every value of q_v is drawn based on twenty simulations and the results are shown in Figures 3-5. In terms of the mean and standard deviation of the number of iterations they show unanimously that the best choice for q_v is 2.3, that which is different from the reported best value, 2.7, by Tsallis and Stariolo. In the three Figures 3-5, it is shown clearly there is at least one run in twenty simulations, in each of the three cases, can not reach the optimum when q_v =2.7; whereas when q_v =2.3 all of the twenty simulations reach the optimum. Indeed, even for the strictest threshold, 0.001, the worst solution is to take less than 3,500 iterations to reach the optimum when q_v =2.3. This shows the out-performance of q_v =2.3 over q_v =2.7. This example demonstrates the importance of accurate algorithm of Tsallis distribution. An erroneous algorithm, such as the one proposed by Tsallis and Stariolo, could lead us to an incorrect conclusion. In additions, q_v =2.3 is also superior to both the classical simulated annealing (CSA) (q_v =1.0) and the fast simulated annealing (FSA) (q_v =2.0) in this case study. In the case of CSA, none of the twenty simulations reach the optimum even for threshold to be 0.1. Whereas for FSA, the achievement is better, however, it is not as good as that of q_v =2.3. For the case of FSA with threshold 0.001, sixteen runs cannot reach the optimum.

Acknowledgements

This research was partially supported by the National Science Council of Taiwan (Project No. NSC 92-2213-E-212-018). Simulation runs in part were performed at a Compaq alpha station, model DS20E, 500 MHz with 2GB memory at the National Center of High-Performance Computing (NCHC), Taiwan. The center's generous support for CPU time is greatly appreciated. The authors also wish to express appreciation to Dr. Cheryl Rutledge for her editorial assistance.

REFERENCES:

Deng J, Chen H, Chang C and Yang Z (2004), A Superior Random Number Generator for Visiting Distribution in GSA, *International journal of computer mathematics*, 81(1): 103-120.

Tsallis C and Stariolo D A (1996), Generalizes simulated annealing, Physical A, 233: 395-406.

Kirkpatrick S C, Gelatt D and Vecchi M P (1983), Optimization by simulated annealing, *Science*, 220: 671 – 680.

Szu H and Hartley R (1987), Fast simulated annealing, *Physics Letters A*, **122**: 157 – 162.

Tsallis C (1988), Possible generalization of Boltzmann-Gibbs statistics, *Journal of Statistical Physics*, **52**: 479-487.

Franz A and Hoffmann K H (2003), Threshold accepting as limit case for a modified Tsallis statistics, *Applied Mathematics Letters*, **16**: 27 - 31.

Geman S and Geman D (1984), Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAMI-6**(6): 721 - 741.

Mantegna R N (1994), Fast accurate algorithm for numerical simulation of Levy stable stochastic process, *Physical Review E*, **49**: 4677-4683.

Birnbaum Z W (1952), Numerical tabulation of the distribution of Kolmogorov's statistic for finite sample size, *Journal of the American Statistical Association*, **47**: 425 - 441.

Massey J F (1951), The Kolmogorov-Smirnov test for goodness of fit, *Journal of the American Statistical Association*, **46**: 68 - 78.

Penna T J P (1995), Traveling salesman problem and Tsallis statistics, *Physical Review E*, **51**: R1-R3.

Okamoto Y and Hansmann U H E (2001), Protein folding simulations by a generalized-ensemble algorithm based on Tsallis statistics, in: *Lecture Notes in Physics: Nonextensive Statistical Mechanics and Its Applications*, Springer Verlag, Berlin, pp 259 - 274.

Xiang Y and Gong X G (2002), Efficiency of generalized simulated annealing, *Physical Review E*, **62**: 4473 - 4476.

Bain L J and Engelhardt M (1987), *Introduction to Probability and Mathematical Statistics*, Duxbury Press, Boston.

Papoulis A (1990), *Probability & Statistics*, Prentice Hall. Englewood Cliffs, NJ.

Monagan M B, Geddes K O, Heal K M, Labahn G, Vorkoetter S M, McCarron J and DeMarco P (2001), *Maple 7 Programming Guide*, Waterloo Maple Inc., Canada.

Vardeman S B (1993), *Statistics for engineering problem solving*, PWS Publishing, Boston.

BIOGRAPHIES:

Jyh-jeng Deng is a professor in industrial engineering and technology management department at DaYeh University, Taiwan. He graduated from National Tsing-Hua University, Taiwan in 1982 with a BS in Industrial Engineering and earned both his MS and PhD in Industrial and Manufacturing Systems Engineering from Iowa State University, USA in 1988 and 1993 respectively. His current research is in the zero weight problems on Non-Uniform Rational B-spline Surface (NURBS) and Russian "theory of inventive problem solving" TRIZ.

Both ChunLan Chang and ZhongXiah Yang are MS graduates from industrial engineering and technology management department at DaYeh University, Taiwan.