

# Simulated annealing algorithm for optimal capital growth

Yong Luo<sup>a,c,\*</sup>, Bo Zhu<sup>b</sup>, Yong Tang<sup>a,d</sup>

<sup>a</sup> School of Management and Economics, University of Electronic Science and Technology, 610054 Chengdu, PR China

<sup>b</sup> School of Finance, Southwestern University of Finance and Economics, 610074 Chengdu, PR China

<sup>c</sup> College of Science, Ningbo University of Technology, 315211 Ningbo, PR China

<sup>d</sup> Department of Physics, University of Fribourg, Chemin du Musée 3, CH-1700 Fribourg, Switzerland

---

## H I G H L I G H T S

- Extension of the capital growth under multi games case.
  - Dynamic optimal capital growth of a portfolio was investigated.
  - A general framework that one strives to maximize growth was developed.
  - Simulated annealing algorithm was investigated to solve the framework.
  - Performance and risk parameter based on real financial data was calculated.
- 

We investigate the problem of dynamic optimal capital growth of a portfolio. A general framework that one strives to maximize the expected logarithm utility of long term growth rate was developed. Exact optimization algorithms run into difficulties in this framework and this motivates the investigation of applying simulated annealing optimized algorithm to optimize the capital growth of a given portfolio. Empirical results with real financial data indicate that the approach is inspiring for capital growth portfolio.

---

## 1. Introduction

For an investor, there are two aspects to improve a trading strategy. The first and the most important goal is to achieve a positive expected return. Once this has been achieved, the investor needs to know what percentage of his capital to risk on each trade, this task is often known as asset allocation or position sizing. The principle of asset allocation is maximizing the expected value of the logarithm of wealth after each period, which was originally developed for gambling [1–11]. Breiman's 1961 paper proved that optimal strategy based on log utility will beat any different strategies almost surely in the long run [12].

The gamblers often bet on several games at once, it is interested in blackjack when a player bets on multiple hands or more players share a common bankroll. Simultaneous bets at different tables are independent but at the same table they have a correlation, this should reduce the fraction per hand. Practical applications to long sequences of wagers are especially

---

\* Corresponding author at: School of Management and Economics, University of Electronic Science and Technology, 610054 Chengdu, PR China. Tel.: +86 13678089339.

E-mail address: [cbiluoy@gmail.com](mailto:cbiluoy@gmail.com) (Y. Luo).

appropriate for the strategy. Hedge fund trading that enters and exits in a few seconds is an application of this. However, when applied to investment with many assets, optimization algorithms such as quadratic programming run into difficulties.

We investigate the problem of dynamic optimal capital growth of a portfolio. A general framework that one strives to maximize the expected log utility of long term growth rate was developed. The original approach ignores many of the constraints faced by real world investors: trading limitations, budget constraint, no short sales, etc. It is very difficult to solve with these constraints. In this study, we investigate applying the simulated annealing algorithm to solve the problem of maximizing long term growth.

Simulated annealing (SA) is a local search technique for combinatorial optimization, such as optimizing functions with multiple variables. Its convergence properties and its ability to escape local optimal have made it a popular technique over the past years [13,14]. 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.

This paper is organized as follows. In Section 2, the capital growth strategy was reviewed. A general framework that one strives to maximize the expected log utility of long term growth rate was developed in Section 3. Finally in Section 4, we present the empirical results with real financial data. We show that in consequence, the approach is inspiring for capital growth portfolio.

## 2. Brief review of the capital growth strategy

We assume that a gambler has found a positive expectation return and he is able to play this game repeatedly, we let the initial bankroll be  $W_0$ , and after  $n$  iterations the bankroll is  $W_n$ . The winning probability is  $p$ , the probability of losing is  $1 - p$ . We define the game return as  $r_i = (W_i - W_{i-1})/W_{i-1}$ , where  $W_i$  is the wealth after  $i$  turns.

The amount of money he could make depends only on how much he chooses to bet. How much would he bet? Further, suppose the gambler bets fraction  $f_i$  of the actual wealth in  $i$  turn. After  $n$  turns the gambler's wealth equals

$$W_n = W_0 \prod_{i=1}^n (1 + f_i r_i). \quad (1)$$

The returns are independent, so the average wealth after  $n$  turns can be written as

$$\langle W_n \rangle = W_0 \prod_{i=1}^n (1 + E[f_i r_i]). \quad (2)$$

Since the game has a positive expectation,  $E[f_i r_i] > 0$  in this situation, in order to maximize  $\langle W_n \rangle$ , we would maximize  $E[f_i r_i]$  at each turn. The optimal strategy is to stake all capital in each trial, however, the probability of ruin is given by  $1 - p^n$  and with  $p < 1$ ,  $\lim_{n \rightarrow \infty} (1 - p^n) = 1$  so ruin is almost sure. Thus maximization of  $\langle W_n \rangle$  is not a good criterion for a long run investment.

In 1956, an asymptotically optimal strategy was proposed by Kelly [1], who made use of a quantity  $G$  called the exponential rate of growth of the gambler's wealth. Without affecting the results, in our analysis we use natural logarithms

$$G = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \frac{W_n}{W_0} \quad (3)$$

as a criterion for investment optimization. Due to the multiplicative character of  $W_n$ ,  $G$  can be rearranged as

$$G = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \prod_{i=1}^n (1 + f_i r_i) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \ln(1 + f_i r_i) = \langle \ln(1 + f_i r_i) \rangle. \quad (4)$$

Imagine that a gambler is faced with an infinitely wealthy opponent who will wager bets made on repeated independent trials of a biased coin. For the risk game introduced above is  $p \ln(1 + f) + (1 - p) \ln(1 - f)$  which is maximized by the investment fraction

$$f^* = 2p - 1. \quad (5)$$

Moreover,  $G(f_c) = 0$  so we get unique number  $f_c > 0$ , where  $0 < f^* < f_c < 1$ . The gambler's wealth will exceed the initial value when  $f$  is chosen in the interval  $(0, f_c)$ . But, if  $f > f_c$ , the ruin is almost sure. In order to maximize wealth, we should maximize  $E[\ln(1 + f_i r_i)]$  by choosing the optimal fraction  $f^*$  at each trial although the probabilities change from one trial to the next.

The gambler introduced here follows a different criterion from the classical gambler. Because of the logarithm which is additive in repeated bets and to which the law of large numbers applies. At every bet he maximizes the expected value of the logarithm of his capital. The criterion asymptotically maximizes the expected growth rate of wealth, which is often called the capital growth strategy.

A criticism applied to the strategy is that capital is not infinitely divisible but multiples of a minimum unit. If the minimum bet allowed is small relative to the gambler's initial capital, the probability of ruin in the standard sense is negligible. In the security markets, the minimum unit can be as small as desired.

### 3. Optimization of simultaneous risky investments

Consider that there are  $n + 1$  investment opportunities  $r_0, r_1, \dots, r_n$  at time  $t$ , which can be done simultaneously in each time step,  $r_0$  represent risk free asset. In opportunities  $i$  ( $i = 0, 1, \dots, n$ ), the investor allocates the fraction  $f_i$  of the total wealth. Let  $F^* = (f_0^*, f_1^*, \dots, f_n^*)$  be optimal investment strategy. We assume that  $f_0^* + f_1^* + \dots + f_n^* \leq 1$  which means there is no borrowing. Let the probability of the outcome  $r_0, r_1, \dots, r_n$  be  $p_0, p_1, \dots, p_n$ . The capital growth function

$$G(f_0, f_1, \dots, f_n) = \sum [p(\cdot) \ln(1 + f_0 r_0 + f_1 r_1 + \dots + f_n r_n)], \quad (6)$$

where  $p(\cdot)$  is joint distribution of  $n$  assets with correlated outcomes,  $p(\cdot) = p_0 p_1 \dots p_n$  if the returns  $r_i$  are independent. Note that concave function  $G(f_0, f_1, \dots, f_n)$  is defined if and only if  $1 + f_0 r_0 + f_1 r_1 + \dots + f_n r_n > 0$ . Computational procedures for finding an optimal fixed fraction strategy are based on the theory of concave function. To find the optimal  $f_i^*$ , let  $\partial G / \partial f_i = 0$ . For portfolios with many securities are extremely difficult to be optimized due to the computation complexity.

For our simultaneous risky investment, the optimal fraction would be the solution of the following optimization problem

$$\max G(f_0, f_1, \dots, f_n) \quad (7)$$

subject to the following constraints

$$\begin{aligned} \sum_{i=0}^n f_i &= 1 \\ f_i &\geq 0 \\ 1 &\leq i \leq n. \end{aligned} \quad (8)$$

The first constraint, called budget constraint, requires that all of the budget to be invested in the risky portfolio and risk free asset. The non-negative constraints express requires that no short sales are allowed. Note that the optimal growth function

$$\begin{aligned} G(f_0, f_1, \dots, f_n) &= \sum [p(\cdot) \ln(1 + f_0 r_0 + f_1 r_1 + \dots + f_n r_n)] \\ &= E \left\{ \ln \left[ \frac{W(f_0, f_1, \dots, f_n)}{W_0} \right]^{1/t} \right\} \\ &= (1/t) E[\ln W(f_0, f_1, \dots, f_n)] - (1/t) \log W_0. \end{aligned} \quad (9)$$

The wealth of portfolio at time  $t$

$$W(f_0, f_1, \dots, f_n) = W_0 \prod_{t=1}^T \left[ 1 + \sum_{i=1}^n f_i r_i \right]. \quad (10)$$

So for a fixed period  $t$ , maximizing  $G(f_0, f_1, \dots, f_n)$  is the same as maximizing  $(1/t) E[\ln W(f_0, f_1, \dots, f_n)]$ . In such situations, classical quadratic programming does not work efficiently and heuristic optimization techniques may be considered as a choice and worth investigation.

It may seem surprising that the portfolio should depend on the past, because the future has no relationship to the past. Indeed the stock sequence is arbitrary, and the malicious nature can make future to take advantage of past beliefs. Cover shows that there exists a universal portfolio strategy, which is only based on the past historical data, that will perform asymptotically as well as the optimal portfolio based on foreknowledge of the price.

$$\lim_{n \rightarrow \infty} \ln \frac{\hat{W}_n}{W_n^*} = 0 \quad (11)$$

where  $\hat{W}_n$  represents the wealth of universal portfolio strategy based on historical data after  $n$  periods,  $W_n^*$  denotes the maximum wealth of the optimal portfolio with foreknowledge [15].

For financial return series, the sample mean is a poor estimator for the expected return. The sample mean is the *best linear unbiased estimator* of the population mean for distributions which are not heavy tailed. Furthermore, the resulting estimator has a large estimation error, which significantly influences the portfolio selection. There are different ways to address this issue. On the estimation side, one can try to produce robust estimates of the input parameters for the optimization problems and this can be achieved by using backtesting.

Backtesting is used in financial industries referring to testing a trading strategy using existing historic data. Instead of applying a strategy for the forward testing, an investor can do a simulation of his or her trading strategy on past data in order to gauge the effectiveness of the strategy. The most common backtesting platforms include Matlab, Tradestation and WealthLab. Our empirical results are based on QuantDeveloper and C#.

Simulated annealing (SA) is a local search technique for combinatorial optimization problems, especially for optimizing functions with multiple variables. Its convergence properties and its use of hill climbing moves to escape local optima have

made it a popular technique over the past years. It is based on an analogy of thermodynamics with the way metals cool and anneal. A particle configuration is a solution to the problem, the energy of a configuration is the value of the objective function where a minimum energy configuration corresponds to an optimal solution.

The major advantage of SA over classical other local search methods is its ability to avoid getting trapped in local optimization while searching for a global optimization. Other local searches such as hill climbing strategy might converge to a local optimal and fail to get the global optimal. In many cases, SA takes less CPU time than genetic algorithm (GA) in solving optimization problems, because it finds the optimal solution using point by point iteration rather than a search over a population of individuals.

#### Objective function

Let  $W(f_0, f_1, \dots, f_n)$  be the wealth of a portfolio. We note that

$$\begin{aligned} G(f_0, f_1, \dots, f_n) &= E \left\{ \ln \left[ \frac{W(f_0, f_1, \dots, f_n)}{W_0} \right]^{1/t} \right\} \\ &= (1/t)E[\ln W(f_0, f_1, \dots, f_n)] - (1/t) \ln W_0. \end{aligned} \quad (12)$$

The wealth of portfolio at time  $t$

$$W(f_0, f_1, \dots, f_n) = W_0 \prod_{t=1}^T \left[ 1 + \sum_{i=1}^n f_i r_i \right]. \quad (13)$$

So for a fixed period  $t$ , maximizing  $G(f_0, f_1, \dots, f_n)$  is the same as maximizing  $(1/t)E[\ln W(f_0, f_1, \dots, f_n)]$ .

#### Statement of SA algorithm

Suppose that the solution space,  $\mathbb{F}$ , is the set of all possible solutions. Let  $G : \mathbb{F} \rightarrow \mathbb{R}$  be an objective function defined on the solution space. Our goal is to find a global optimal  $f^*$ ,  $f^* \in \mathbb{F}$  such that  $G(f^*) \geq G(f)$  for all  $f \in \mathbb{F}$ , where solution vector be  $f = (f_0, f_1, \dots, f_n)$ , best available solution vector is  $f^* = (f_0^*, f_1^*, \dots, f_n^*)$ . The objective function must be bounded to ensure that  $f^*$  exists. A neighboring solution  $f'$  is generated either randomly or using some pre specified rule. The candidate solution is accepted as the current solution based on the acceptance probability

$$P(\cdot) = \begin{cases} e^{-\frac{G(f')-G(f)}{T_k}}, & \text{if } G(f') - G(f) < 0 \\ 1, & \text{if } G(f') - G(f) \geq 0 \end{cases} \quad (14)$$

where  $T_k$  is the temperature parameter at iteration  $k$ , such that  $T_k > 0$  for all  $k$  and  $\lim_{k \rightarrow +\infty} T_k = 0$ . The generic simulated annealing algorithm performs the following steps.

- Choose an initial solution  $f_0$ , set the temperature change counter  $k = 0$ , set a temperature cooling schedule  $T_k$ , set an initial temperature  $T_0$ , set  $M_k$  as the number of iterations executed at each temperature  $T_k$ .
- Repeat until a stopping criterion is fulfilled.
  - Set repetition counter  $m = 0$ .
  - Repeat until  $m = M_k$ .
    - \* Draw a solution  $f'$  in the neighborhood.
    - \* Calculate  $\Delta_{f,f'} = G(f') - G(f)$ .
    - \* If  $\Delta_{f,f'} \geq 0$ , then  $f \leftarrow f'$ .
    - \* If  $\Delta_{f,f'} < 0$ , then  $f \leftarrow f'$  with probability  $e^{-\frac{G(f')-G(f)}{T_k}}$ .
    - \*  $m \leftarrow m + 1$ .
  - $k \leftarrow k + 1$ .

Due to the generality of the concepts that it involves, simulated annealing can be applied to a wide range of optimization problems. In particular, no specific requirements need to be imposed on the objective function nor on the solution space. The simulated annealing pseudo code is outlined in the following.

```

1  m = number of moves to attempt;
2  t = current temperature;
3  s = stop temperature;
4  a = constant factor;
5  f_0 = initial solution;
6  g_old = old solution of objective function;
7  g_new = new solution of objective function;
8  if(t > s){
9    for(i = 1; i < m; i++){
10     Generate a random solution;
11     Evaluate the change in objective function, g_new - g_old;
12     if(g_new - g_old > 0){
13       accept this move and update configuration;

```

```

14     } else{
15         accept with probability P;
16         update configuration if accepted;
17     }
18 }
19 t = a * t;
20 }

```

In order to apply this technique, a serial of annealing parameters must be defined including the initial value of temperature, the cooling schedule, the number of iterations to be performed at each temperature and the stopping criterion.

The theory of SA indicates that SA converges to a globally optimal solution with probability 1. However, the asymptotic convergence, such as cooling schedule and stop criterion, cannot be met in practice. Thus, it is often critical to adjust values of those parameters based on problem characteristics. We determine which parameters to be used considering the performance and running time of the algorithm. In order to obtain good performance, we need to specify these values carefully.

#### *Initial value of temperature*

Theoretically speaking, the initial temperature should be high enough so that all movements are acceptable. However, it is necessary to control initial temperature because very high initial temperature could consume too much time. A proper initial value of temperature is chosen so that it can cover the entire solution space. A very high initial temperature increases the solution space and requires a large number of iterations.

Many researches have proposed different ways to select the initial temperature, but no conclusion has been made about the approach of selection of the initial temperature. Following the recommendations of many authors, we set the initial temperature in such a way that during the first cooling stage, the probability of acceptance of a move is relatively  $\chi_0$ . We set the initial probability equal to 0.9. In preliminary phase, the algorithm runs  $M$  steps without rejecting any moves. The average increase of the objective function over this phase as  $\Delta$ .  $T_0$  equal to

$$T_0 = \frac{-\Delta}{\ln \chi_0}. \quad (15)$$

#### *Cooling schedules*

Simulated annealing convergence theory has been used to design effective cooling schedules. The earliest annealing schedules based on the analogy with physical annealing. Cooling schedules are grouped into static schedules and adaptive schedules. The first must be completely specified before the algorithm begins, and the second will adjust the decrease rate of temperature based on the information obtained during the algorithm's execution. Strenski and Kirkpatrick's paper suggests that optimal cooling schedules are not monotone decreasing in temperature [16]. They also show that in their testing problem, geometric and linear cooling schedules perform better than logarithmic cooling schedules.

In proportional temperature cooling schedules, the temperature is kept fixed at each stage. After each stage, the temperature is multiplied by a constant factor  $\alpha \in (0, 1)$ . Finally, temperature becomes very small and it does not search any smaller energy level, which is called frozen state. The cooling schedule is given as

$$T_{k+1} = \alpha T_k. \quad (16)$$

The geometric ratio, used commonly in practice, use  $\alpha$  between 0.1 and 0.99.

#### *Number of iterations*

The chain's size at each temperature is chosen so that the system is close to the stationary distribution at that temperature. If the size is small compared to the total solution space, then the Markov chain cannot move around the solution space fast enough to find the optimal solution in reasonable time. On the other hand, a very large size has the algorithm randomly in a large solution space, and thus, is unable to focus on specific areas of the solution space.

The number of iterations  $M_k$  are difficult to quantify precisely. The value of the number of iterations depends on the nature of the problem. There is no general agreement about it [17].

#### *Stopping criterion*

Stopping criterion is the condition that the algorithm stops, and it stops theoretically at the point that converges to temperature of 0. Because the algorithm has to spend a long time under low temperatures in this case, it is typical to stop the algorithm if the performance does not improve after some iterations.

If the number of iterations exceeds the allowed number of iterations, we terminate the algorithms,  $10^8$  in our experiment.

Our algorithm terminates if temperature reaches the minimum value of temperature. The idea is generated with the fact that the chance of improvement in a solution is rare once the temperature is close to zero. At a very low temperature, moves will be trapped in the neighborhood of the current solution.

#### *Choice of neighborhood*

Neighborhood search is a key problem in SA, which is an iterative procedure to move from an admissible schedule to a new one until some admissible schedule satisfying a certain stopping criterion. The efficiency of simulated annealing is highly influenced by the neighborhood search.

**Table 1**  
Optimal allocation of the pool portfolio.

Symbol	CSCO	AMZN	AAPL	YHOO	MSFT	
	C	ADS	BRK.A	BAC	AFL	
	KO	SBUX	JNJ	KFT	PFE	
	CBS	TWX	DIS	TRI	AOL	
	DUK	RIO	XOM	AEP	COP	
	CL	GE	PG	AA	DD	
	DIA	XLE	XLV	USO	FXI	
	CPHI	SNP	BIDU	PTR	CHL	
	AES	ORCL	QCOM	EBAY	GOOG	
	GS	MA	WMT	FDX	RTN	
	Fraction	0.0032	0.0164	0.0296	0.0179	0.0012
		0.0023	0.0497	0.0021	0.0031	0.0007
		0.0484	0.0221	0.0041	8.65E-05	0.0062
0.0137		0.0007	0.0034	0.0023	0.0486	
0.0006		0.0461	0.0052	0.0463	0.0498	
0.0003		0.0005	0.0004	0.0073	0.0001	
0.0037		0.0011	0.0033	0.0139	0.0028	
0.0407		0.0153	0.0489	0.0345	0.0002	
0.0338		0.0004	0.035	0.046	0.0046	
0.0001		0.0022	0.0036	0.0001	0.0111	

Instead of predefining the random disturbance as a constant, we generate random disturbance from Cauchy distributed random numbers. Cauchy mutation is more likely to generate an offspring that is further away from its parent than Gaussian mutation due to its long flat tails. It is expected to have a higher probability of escaping from a local optimum or moving away from a plateau. On the other hand, the smaller hill around the center indicates that Cauchy mutation spends less time in exploiting the local neighborhood.

The one dimensional Cauchy density function centered at the origin is defined by

$$f(x) = \frac{1}{\pi} \frac{t}{t^2 + x^2}, \quad -\infty < x < \infty \quad (17)$$

where  $t$  is a scale parameter.

#### How to handle constraints

For better risk controlling, most portfolio managers limit the fraction allocate to a single asset. If all of our money put in a basket, market crash will make us go bankruptcy. So, in our optimal model, we consider the constraint of asset fraction, which is also similar to fraction Kelly strategy. But how do we ensure that the final solution satisfies all the constraints? First, we should make sure the neighborhood of current solution consists of feasible solutions.

In our implementations, we selected all of the feasible approaches, in which the neighborhood of the current solution may contain solutions satisfying the constraints. The algorithm draws a direction and takes a small step in this direction away from the current solution. The direction is computed to meet the constraints, so all of the feasible approaches will not lost time to deal with infeasible solutions.

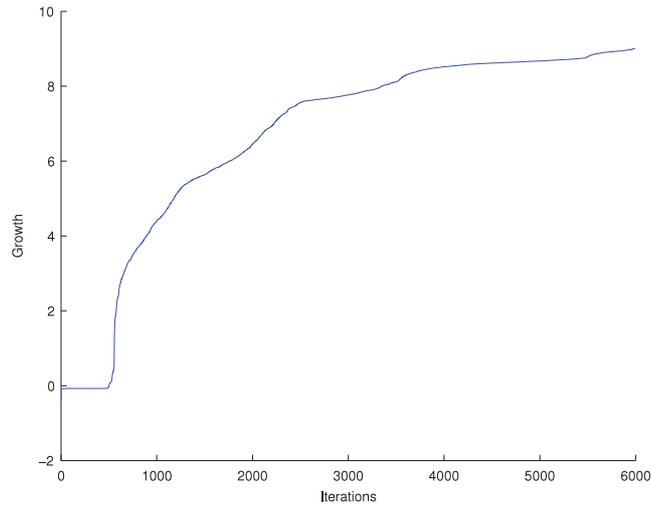
## 4. Empirical results with real financial data

The algorithms described above have been implemented in C# and run on a personal computer. All the figures in this paper are drawn with Matlab. In our implementations of simulated annealing, we adopted the geometric cooling schedule defined in the last section. In order to describe more completely this cooling schedule, we need to specify the value of the parameters. Unless otherwise stated, the parameter settings for the simulated annealing are defined as follows: number of tries per on circle of temperature  $M = 250$ ; start temperature in the annealing schedule  $T_0 = 100$ ; stop temperature in the annealing schedule  $S = 0.0001$ ; temperature reduction factor  $\alpha = 0.1$ ; max number of objective function calls is 10 million.

For the sake of constructing realistic problem instances, we use the daily data extracted from our database, which is from USA stock market covering the period 1st January 2000–31st December 2011. The stock is obtained from these sectors: technology, finance, consumer goods, health care, services, basic materials, industrial goods. Using a different sector division can efficiently minimize the ratio of correlations and smooth the capital growth curve. For simplicity, we assume that there have no transaction cost, tax and short sell, which will be investigated in a separate work. To illustrate the performance of the optimal asset allocation, a trading model named Turtle [18] was applied.

It is assumed that the stocks were drawn at random from asset pool.<sup>1</sup> The optimized fraction for the algorithms is shown in Table 1. If we allocate on AAPL, we always allocate 2.96% of the total bankroll at time  $t$ . It can be seen that the algorithm

<sup>1</sup> The pool portfolio include 50 assets from USA market: CSCO, AMZN, AAPL, YHOO, MSFT, C, ADS, BRK.A, BAC, AFL, KO, SBUX, JNJ, KFT, PFE, CBS, TWX, DIS, TRI, AOL, DUK, RIO, XOM, AEP, COP, CL, GE, PG, AA, DD, DIA, XLE, XLV, USO, FXI, CPHI, SNP, BIDU, PTR, CHL, AES, ORCL, QCOM, EBAY, GOOG, GS, MA, WMT, FDX, RTN.



**Fig. 1.** Iterations of the pool portfolio.

**Table 2**  
Optimal allocation of the capital growth portfolio.

Symbol	AAPL COP	ADS CPHI	KO BIDU	SBUX PTR	AOL AES	RIO QCOM	AEP EBAY
Fraction	0.0299 0.0465	0.0401 0.0224	0.0407 0.0455	0.0196 0.0459	0.0459 0.0225	0.0412 0.0315	0.0471 0.047

**Table 3**  
Execution time.

Computer	Portfolio size	Execution time (h)
PC (Intel i5-3450 3.1G processor, 4G memory)	50	8.9
PC (Intel i5-3450 3.1G processor, 4G memory)	14	2.3

produces most unequal sizes of fraction. The difference between the fraction is that some have a higher value of winning or R multiple and, for some, the value is smaller. We note that the total committed is considerably less than 100% of the bankroll.

For the value of optimal fraction, Fig. 1 illustrates the evolution of the portfolio growth rate in the course of iterations. The simulation was repeated 5995 times and there are 1360 investment decision points. Convergence to solution is fast for the simulated annealing algorithm, although error term does not reach zero for each circle. No further optimization occurs after 4000 iterations, which suggests that the growth function has reached a good optimum.

#### *Performance and risk analysis of the capital growth portfolio*

The stock was selected if the fraction is larger than 2% in the pool portfolio.<sup>2</sup> The optimized fraction of the selected portfolio is shown in Table 2, which means there are 14 stocks in the portfolio. If we allocate on AAPL, we always allocate 2.99% of the total bankroll at time  $t$ . Its can be seen that the sum of fraction is 52.58%, which means 47.42% of total capital will be allocated to money market.

Considering the data described in Table 4, the investor's initial wealth is 1 million and there are 333 investment decision points, average annual return is 460.6%, the final wealth is more than 134 821 times the initial wealth. Indeed, Kelly criterion provides an enormous growth rate but a very high volatility of wealth levels. The final wealth is very likely to be higher than other strategies, but the volatility will be very bumpy. Drawdown controlling will be investigated in a separate work. From the Skewness and Kurtosis data shown in the table, the return of the portfolio does not satisfy the assumption of normality in classical mean variance models.

Table 3 shows the performance of different portfolio in the same computer. Fig. 2(a) is the curve of equity. For the value of optimal fraction, Fig. 3(a) illustrates the simulation was repeated 1682 times, which is evolution of the portfolio growth rate in the course of iterations.

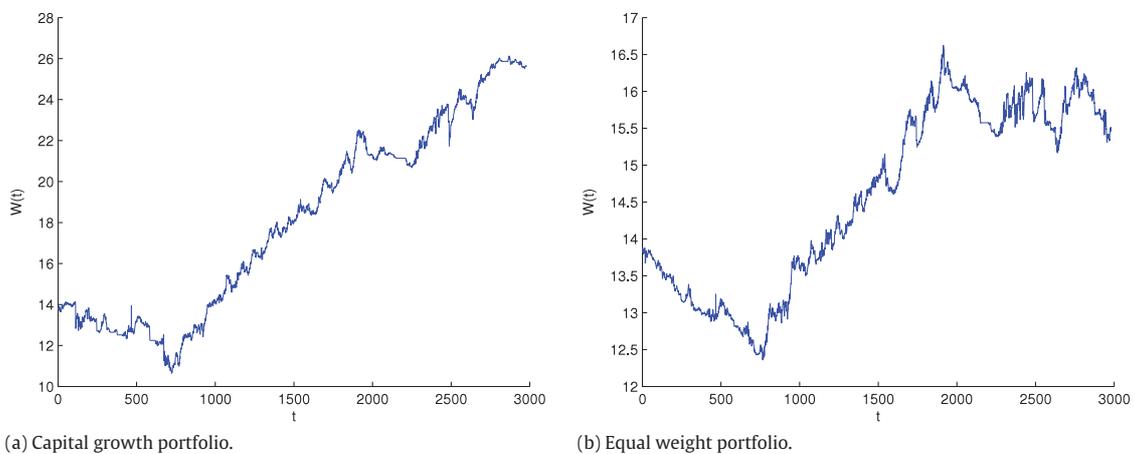
#### *Comparison with the equal weight portfolio*

Assume that each stock has an equal weight, consequently, the optimal fractions are identical and the investment optimization is simplified to a one variable problem. We optimize the problem with a simple downhill optimization technique.

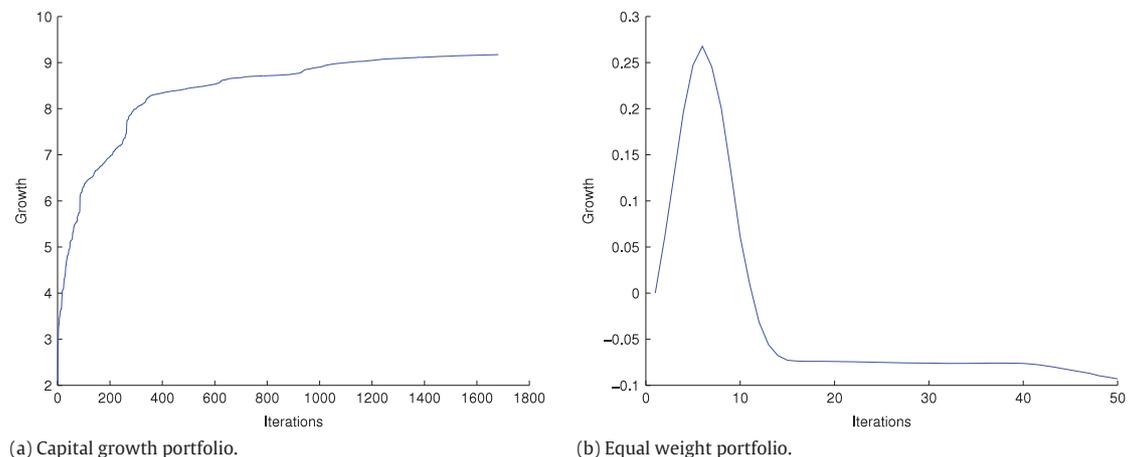
<sup>2</sup> The selected portfolio include 14 assets from USA stock market: AAPL, ADS, KO, SBUX, AOL, RIO, AEP, COP, CPHI, BIDU, PTR, AES, QCOM, EBAY.

**Table 4**  
Performance and risk analysis.

Statistics	Capital growth portfolio	Equal weight portfolio
Initial Wealth	1 million	1 million
Final Wealth	134 821 times	5 times
Average Annual Return (%)	460.6	40.45
Median Annual Return (%)	414.88	29.8
Maximum Annual Return (%)	1281.68	210.77
Minimum Annual Return (%)	-79.53	-51.23
Average Monthly Return (%)	19.16	2.66
Median Monthly Return (%)	5.01	-1.1
Maximum Monthly Return(%)	260.64	64.86
Minimum Monthly Return (%)	-84.08	-41.81
Compound Average Return	306.47	61.43
Minimum DrawDown	-0.96	-0.78
Number Of RoundTrips	333	1398
Skewness	0.22	-0.64
Kurtosis	13.43	8.04



**Fig. 2.** Equity of the portfolio (log scale).



**Fig. 3.** Iterations of the portfolio.

The method is feasible only for small problems, since the number of possible states of the system increases exponentially with the number of dimensions. The case of unequal weight is impossible achieved by straightforward optimization technique when the portfolio include 50 stocks. In fact, many hedge funds have the fastest real time large machine or distributed computing power in high frequency quantitative trading. Further systematic comparison with other heuristic techniques such as genetic algorithm is a complex topic, which will be investigated in a separate work.

The statistics of performance and risk are described in Table 4, the final wealth is more than 5 times to the initial wealth. Fig. 2(b) provides a visual representation of equity. Fig. 3(b) illustrates the evolution of the capital growth in the course of iterations.

## 5. Conclusions

We investigated the problem of dynamic optimal capital growth of a portfolio. A general framework that one strives to maximize the expected log utility of long term growth rate was developed. Classical optimization algorithms run into difficulties in this framework, it motivates the investigation of the algorithm of simulated annealing. Simulated annealing is a powerful tool for the solution of many optimization problems. Its main advantages are its ability to approach global optimality.

The result shows that the investor's wealth will exceed the initial value when the fraction is chosen less than critical value. But, if larger than the value, ruin is almost sure. Empirical results with real financial data show the performance and risk, final wealth of the capital growth strategy far more than the equal weight strategy. In order to maximize wealth, we should choose the optimal fraction at each trade.

## Acknowledgments

This work was supported by The National Natural Science Foundation of China (project 71103146). We thank the anonymous reviewers for their comments and suggestions, which helped us to improve this work. We appreciate Liang Zhang for discussions and many stimulating comments from professor Zhang.

## References

- [1] J.L. Kelly, A new interpretation of information rate, *Bell Syst. Tech. J.* 35 (1956) 917–926.
- [2] H.A. Latane, Criteria for choice among risky ventures, *J. Political Econ.* 67 (1959) 144–155.
- [3] E.O. Thorp, Optimal gambling systems for favorable games, *Rev. Int. Statist. Inst.* 37 (1969) 273–293.
- [4] E.J. Elton, M.J. Gruber, S.J. Brown, *Modern Portfolio Theory and Investment Analysis*, seventh ed., Wiley, New York, 2006.
- [5] S. Browne, W. Whitt, Portfolio choice and the bayesian kelly criterion, *Adv. Appl. Probab.* 28 (1996) 1145–1176.
- [6] W. Poundstone, *Fortune's Formula*, Hill and Wang, New York, 2005.
- [7] E.O. Thorp, The Kelly criterion in blackjack sports betting and the stock market, in: *Handbook of Asset and Liability Management*, Vol. 1, 2006, pp. 387–428.
- [8] E.W. Piotrowski, M. Schroeder, Kelly criterion revisited: optimal bets, *Eur. Phys. J. B* 57 (2) (2007) 201–203.
- [9] W.T. Ziemba, D.B. Hausch, The Dr. z betting system in England, in: *Efficiency Racetrack Betting Markets*, Vol. 1, 2008, pp. 567–574.
- [10] MacLean, Thorp, Zhao, Medium term simulations of the full Kelly and fractional Kelly investment strategies, in: *Kelly Capital Growth Investment Criterion: Theory and Practice*, Vol. 1, 2010, pp. 543–562.
- [11] M.A.H. Dempster, I.V. Evstigneev, K.R. Schenk-Hppe, Growing wealth with fixed-mix strategies.
- [12] L. Breiman, Optimal gambling systems for favorable games, in: *Fourth Berkeley Symposium on Probability and Statistics*, vol. 1, 1961, pp. 65–78.
- [13] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, *Science* 220 (1983) 671–680.
- [14] Y. Crama, M. Schyns, Simulated annealing for complex portfolio selection problems, *European J. Oper. Res.* 150 (2003) 546–571.
- [15] T.M. Cover, Universal portfolios, *Math. Finance* 1 (1991) 1–29.
- [16] K.A. Strenski, S. Kirkpatrick, Analysis of finite length annealing schedules, *Algorithmica* 6 (1991) 346–366.
- [17] D.S. Johnson, C.R. Aragon, L.A. McGeoch, Optimization by simulated annealing: an experimental evaluation, *Oper. Res.* 37 (1989) 865–892.
- [18] C.M. Faith, *Way of the Turtle*, McGraw Hill, New York, 2007.