一种获得高非线性平衡布尔函数的启发式算法*

陈 果
(重庆师范大学 影视传媒学院,重庆 400047)

摘 要:利用混沌动力系统的良好特性,把它引入传统的模拟退火中,提出一种称为混沌搜索模拟退火的新启发式算法,用于设计高度非线性平衡布尔函数。笔者分别对此方法和传统的模拟退火算法,进行多项仿真实验。实验分析表明,此算法能够更有效地避免陷入局部极小值,对于获得密码性质好的布尔函数,比单一的模拟退火具有更大优势。

关键词:混沌动力系统;模拟退火;布尔函数;非线性度
中图分类号:TP391 **文献标识码**:A **文章编号**:1672-6693(2007)01-0037-03

A Heuristic Algorithm For Obtaining Balanced Boolean Function Satisfying High Nonlinearity

CHEN Guo
(College of Movies and Media, Chongqing Normal University, Chongqing 400047, China)

Abstract: A heuristic hybrid algorithm called chaotic searching simulated annealing is here proposed, which is mixed by chaotic searching and simulated annealing to evolve balanced boolean functions satisfying high nonlinearity. Furthermore, several experiments have been made based on the novel and the traditional simulated annealing methods. The results of numerical analysis show that the novel algorithm could avoid plunging into local optimal value, more effectively and find better balanced boolean functions satisfying high nonlinearity than that of the traditional one.

Key words: chaotic dynamic system; simulated annealing; boolean function; nonlinearity

布尔函数是密码学研究领域里的核心内容,通常它的优劣直接关系到整个密码系统的安全。一直以来,人们把目光集中在用代数方法上,产生了不少布尔函数的设计方法和准则[1-3]。随着人工智能型算法的出现,一些密码学家发现,可以把这些方法用于密码学的设计。数学上,布尔函数可以表示成 $f: F_2^n \to F_2^1$,其中 F_2^n,F_2^1 分别表示伽罗华域 GF(2^n),GF(2)。根据组合数学的知识,对于 n 个比特输入的布尔函数构成了一个 $\{2^{n^n}\}$ 的集合。布尔函数的设计过程可以看成是从这个庞大的集合中寻找满足良好密码学特性的布尔函数,这就可归结为一个组合优化问题。Clark[4] 提出了一种用模拟退火算法来搜寻性质好的布尔函数。但由于模拟退火中采用的是随机搜索,为了寻找最优解,算法通常要求较高的初温,较慢的降温速率,较低的终止温度以及各温度下足够多次的抽样,因而模拟退火算法往往优化过程较长。

另一方面,近 20 年来,人们对混沌动力系统进行了大量研究,逐步认识到它的一些重要特性,如高度不可预测性、遍历性、对初始条件敏感性等。混沌运动能在一定范围内按其自身“规律”不重复地遍历所有状态。因此,如果利用混沌变量进行优化搜索[5],无疑会比随机搜索更具有优越性。笔者在总结了一些混沌优化搜索的基础上,把它嵌入到模拟退火中,替代原有的随机搜索,提出一种混沌搜索模拟退火算法,来寻找高度非线性的平衡布尔函数。

* 收稿日期:2006-07-18 修回日期:2006-10-30
资助项目:重庆市教委科技研究项目(No. KJ060804;No. KJ050802);重庆市科委自然科学基金项目(No. 8503;No. 8509)
作者简介:陈果(1980-) 男,重庆人,助教,硕士,研究方向为智能算法、信息安全。

1 布尔函数的密码学特性

1.1 平衡性

平衡性是密码学里的一个重要属性,它使攻击者难以从平衡布尔函数的输出中猜测其输入。

定义1 设 n 比特输入的布尔函数 $f(x)$: $F_2^n \to F_2$ 称为是平衡的,当且仅当满足

$$\sum_{x=0}^{2^n-1} f^*(x) = 0 \tag{1}$$

其中 $f^*(x)$ 为其极坐标形式

$$f^*(x) = (-1)^{f(x)} \tag{2}$$

1.2 非线性度

非线性度是衡量布尔函数的线性逼近程度,非线性度高的布尔函数可以有效地抵抗线性攻击。非线性度的精确定义如下。

定义2 设 $f(x)$: $F_2^n \to F_2$, $L_n = \{u \cdot x + v \mid u = (u_1, u_2, \cdots, u_n) \in F_2^n, v \in F_2\}$ 表示 F_2^n 上的所有线性布尔函数所组成的集合,其中 $u \cdot x = u_1x_1 \oplus u_2x_2 \oplus \cdots \oplus u_nx_n$。称非负整数

$$N_f = \min_{l(x) \in L_n} d_H(f(x), l(x)) \tag{3}$$

为布尔函数 f 的非线性度,其中 $d_H(f(x), l(x))$,表示 $f(x)$ 与 $l(x)$ 之间的汉明距离。为了便于分析和计算,人们引入了 Walsh 谱,把布尔函数从时域变换到频域。

定义3 设 $f(x)$: $F_2^n \to F_2$,一阶线性 Walsh 谱定义为

$$S_{(f)}(\omega) = 2^{-n} \sum_{x \in F_2^n} (-1)^{f(x)} (-1)^{\omega \cdot x} \tag{4}$$

引入了 Walsh 谱,公式(3)就有了一个等价式

$$N_f = 2^{n-1}(1 - \max_{w \in F_2^n} | S_{(f)}(\omega) |) \tag{5}$$

从密码学的角度,希望所选用的布尔函数 f 的非线性度越大越好,即 $\max | S_{(f)}(\omega) |$ 就必须尽可能小。又由 Pareseval 定理可推出

$$\sum_{\omega \in F_2^n} S_{(f)}^2(\omega) = 1 \tag{6}$$

因而可得非线性度的上界 $N_f \leqslant 2^{n-1}(1 - 2^{-n/2})$。人们把非线性度取到极值的这类函数称为 Bent 函数。但由于 Bent 函数不是平衡函数[1],一般不能直接用在密码学系统中。如何有效地构造平衡的高非线性布尔函数是目前密码学界的热点问题。

2 模拟退火

模拟退火算法是基于蒙特卡罗迭代求解策略的一种随机寻优算法,其出发点是基于物理退火过程与组合优化之间的相似性。它由某一较高初温开始,利用具有概率突跳性的 Metropolis 抽样策略在解空间中进行随机搜索,伴随温度的不断下降重复抽样过程,最终得到问题的全局最优解。标准模拟退火算法(寻找最小值)的一般步骤可描述如下。

1) 给定初温 $t = t_0$,内循环次数 N,最低度 $T_{\min}$,温度下降因子 δ。随机产生初始状态 $s = s_0$,并令 $n = 0$。

2) While($t > T_{\min}$) do

{

While($n \leqslant N$) do

{

随机产生下一个新状态 s_j = Genete(s);

If ($C(s_j) - C(s)$) $\leqslant 0$ Then $s = s_j$(C 为目标函数)

Else 以概率 $\exp(C(s) - C(s_j)/t)$ 更新 $s = s_j$;

$n = n + 1$;

}

$n = 0, t = t \times \delta$;

}

3) 输出搜索结果 s。

3 混沌搜索

混沌搜索的基本思想就是用类似载波的方法将混沌状态引入到优化变量中,并把混沌运动的遍历范围‘放大”到优化变量的取值范围,然后利用混沌变量进行搜索。由于混沌运动具有遍历性、拓扑传递性等特点,使搜索更加有效。首先选择用于载波的混沌变量,笔者选用(7)式所示的 $I = [-1, 1]$ 上的混沌 Chebyshev 映射

$$\tau(t) = \cos(k\cos^{-1}t) \tag{7}$$

它具有的不变测度为

$$f^*(\omega)\mathrm{d}\omega = \frac{1}{\pi\sqrt{1-t^2}}\mathrm{d}\omega \tag{8}$$

可以证明[6],此迭代映射满足遍历性,对称性和对初值敏感性等混沌特性。布尔函数自变量的取值范围为

$$\{x \mid x \in [0, 2^n] \wedge x \in z\} \tag{9}$$

为了使混沌搜索变得可能,每次迭代得到的结果需增加如下尺度变换

$$x = \square\ 2^n t \mid \rfloor \tag{10}$$

4　混沌搜索模拟退火算法

4.1　目标函数的确定

目标是寻找高非线性的平衡布尔函数，显然这是一个多目标规划问题。通常多目标规划的绝对最优解一般并不存在，只是存在一些相对最优的有效解。另外，多目标规划的可行解集合不是完全有序集，一般只能定义偏序关系，于是在多个有效解之间无法直接做进一步比较。但经过仔细观察发现，对于需要解决的这个问题，可以通过(11) 式，把搜索区间仅定位在平衡布尔函数上移动，从而达到把多目标转化为单目标规划。令 f 为目前状态(为一个平衡布尔函数)，x，y 为 n 比特向量，且 $f(x) \neq f(y)$。g 为新状态，若 g 满足

$$\begin{aligned} &g(x) = f(y), g(y) = f(x) \\ &\forall z \in 2^n \setminus \{x, y\} : g(z) = f(z) \end{aligned} \tag{11}$$

g 亦为平衡布尔函数。解决了平衡问题，很容易得到如下目标函数

$$\triangle E = N_f - N_g \tag{12}$$

其中 N_f，N_g 分别为布尔函数 g 和 f 的非线性度。

4.2　算法描述

在确保一定要求的优化质量基础上，为了能够提高算法的效率，笔者在此算法中增加一条退出准则。当目前状态在 K 次循环中都未能得到更新，可以认为此状态已达到最优解。具体算法如下。

1) 混沌初始化，给定一个初始值 m_0，并迭代产生 m_1。初始化当前最优解 f^*，初始化初温 t_0，最低温度 $T_{\min}$ 和内循环迭代次数 N，算法终止次数 K。并设 $n = 0, k = 0, t = t_0$。

2) While$(t > T_{\min})$ do

{$g^* = f^*$ (保存当前状态)；

While$(n \leqslant N)$ do

{迭代 Chebyshev 映射，并根据公式(10)，(11) 提供的方法，产生新的平衡布尔函数，记为 g；

根据(5)，(12) 式计算 $\triangle E$；

if $\triangle E \leqslant 0$ Then 更新 $f^* = g$；

if $\triangle E > 0$ Then 以概率 $\exp\left(-\frac{\triangle E}{t}\right)$ 更新 $f^* = g$，if$(g^* == f^*)$ Then $k = k + 1$；

$n = n + 1$，if$(k \geqslant K)$ Then goto 3)；

}

$n = 0, k = 0$，　$t = t \times \delta$；

3) 输出 f^*。

5　实验分析

在实验中，笔者用模拟退火和混沌模拟退火两种算法分别对 F_2^4，F_2^6，F_2^8 上的布尔函数进行搜索。对于这两种算法，初始温度和温度下降因子的确定，是做实验前必须考虑的问题。实验表明，初温越大，获得高质量解的机率越大，但花费的计算时间将增加。因此初温确定应折衷考虑优化质量和优化效率。文献[7] 中把遗传算法与模拟退火相结合，提出了一种获得初温的方法。在实验中，笔者采用类似的方法来确定初始温度[8]。假定取开始时接受较差点的概率为 P_r，产生一组状态函数(在模拟退火中为随机产生，在混沌模拟退火中用混沌迭代映射产生)。计算这组可行解的目标函数，获得最大和最小值，记为 $N_{\max}$ 和 $N_{\min}$，并按公式(13) 得到初温

$$T_0 = \frac{N_{\min} - N_{\max}}{\ln P_r} \tag{13}$$

实验中 P_r 取 0.5，$T_{\min}$ 为 0.1，并产生 1 000 组状态函数来获得初温，取内循环次数 N 为 3 000 次，δ 为 0.98，混沌迭代初值为 0.6，其运行结果如表 1 所示。

表 1　两种算法的实验对比

N 比特函数	4	6	8
非线性度(模拟退火)	6	20	104
非线性度 (混沌模拟退火)	6	28	112

从表 1 可以看出，在所有初始状态相同的情况下，混沌模拟退火比单一的模拟退火算法找到的平衡布尔函数非线性度要高。

又由于两种算法在退火过程都存在以概率接受较差的状态，因而重复相同的实验其结果会有差异，仅仅通过一次实验尚无法说明混沌模拟退火更优越。基于这种考虑，下面从统计分析的角度来考察。对 F_2^8 上的布尔函数，分别对于上述两种算法各测试了 100 次，各样本观察值及出现的次数如表 2 所示。

从表 2 能够看出，用模拟退火法，非线性度主要集中在 100 到 108 之间，其样本中位数和样本均值分别为 104，104.64。而用混沌模拟退火算法，非线性度主要集中在 106 到 114 之间，其样本中位数和样本均值分别为110，109.72。可见，混沌模拟退火

(下转 58 页)