Introducing Biasedness in NSGA-II to Construct Boolean Function Having Best Trade-Off Among Its Properties

Rajni Goyal(✉) and Anupama Panigrahi

Department of Mathematics, University of Delhi, Delhi 110016, India
goyalrajni2584@gmail.com, anupama.panigrahi@gmail.com

Abstract. To construct Boolean function, many heuristic techniques have been used like NSGA-II, PSO, Ant Colony Method etc., but results are good only for few variables and complexity of these methods are very high. So, to reduced the complexity and to get desired results instead of all solutions, we have introduced a new concept of biasedness in our proposed method. We have used NSGA-II as our heuristic technique with concept of biasedness and got desired Boolean functions for 6 and 7 variables.

Keywords: MOOP · Genetic algorithms · Boolean functions · Biasedness · Nonlinearity

1 Introduction

In literature, there are many methods (heuristic as well as concatenation) to construct optimized Boolean functions. But heuristic techniques are mostly used as complexity of these methods are comparatively less and we can generate Boolean functions on higher variables also. In [1] Aguirre et al. have given a very good approach for multiobjectives. They took two and three objectives, and compared the results with two stage optimization. In [2], Camion et al. mentioned a new approach based on orthogonal arrays and constructed Boolean functions having good correlation immunity. In [3], Clark et al. gave a new two stage method based on simulated annealing, and result listed in this paper was better than previous results. But that method was able to get the optimum Boolean functions only for some limited variables. In [10,13,14], Maitra et al. constructed correlation immune functions keeping their nonlinearity optimal. First time they have constructed 1-resilient Boolean function with maximum nonlinearity for 8 variables and that method was based on concatenation. In [6,8,9], Clark et al. have found some functions with best tradeoff among Boolean function's properties. In [11,12,15], there are some construction methods but these methods are not applicable for multiobjectives and complexity of these methods is not considerable. In [7], we have given a method based on multiobjective optimization (based on genetic algorithms) but were able to get the functions only for 4, 5, 6 and 7 variables and complexity of method was high.

K. Deep et al. (eds.), *Proceedings of Sixth International Conference on Soft Computing for Problem Solving*, Advances in Intelligent Systems and Computing 546, DOI 10.1007/978-981-10-3322-3_11

There are many others heuristic and other types of techniques available in literature. But to find good trade-off among the properties, only a heuristic technique is not sufficient. We want a technique that is having less complexity and should work for large number of variables. If we want to optimize many properties simultaneously, technique should be multiobjective also. So, by introducing biasedness concept [4] in heuristic technique, we tried to find good Boolean functions with less complexity. So, in present paper, we have given a new concept (biasedness) and got some optimum results.

2 Some Definition [1, 7]

2.1 Boolean Function

Any function $g : \mathbb{K}_2^n \to \mathbb{K}_2$ is called a Boolean function of n-variables. $\mathbb{K}_2^n$ is vector space (n-dimensional) over $\mathbb{K}_2$ where $\mathbb{K}_2$ represents a field of two elements. $\mathcal{Z}_n$ is called the set of all Boolean functions (n-variables).

2.2 Balancedness

If number of 0's in truth table representation is same as the number of 1's, than function is called balanced and the property is known as the balancedness.

2.3 Walsh Hadamard Transform

Boolean function can be represent in term of Walsh Hadamard Transform (WHT) also. If L_λ is linear function, specified by $\lambda \in \mathbb{K}_2^n$, the we denote WHT by $H_g(\lambda)$ and can be defined as

$$H_g(\lambda) = \sum_{x \in \mathbb{K}_2^n} (-1)^{g(x) \oplus \lambda . x}. \tag{1}$$

2.4 Non-linearity

Nonlinearity of a Boolean function is minimum hamming distance of that function from the set of all affine functions. It can be given by

$$nl(g) = (2^n - \max_{\lambda \in \mathbb{K}_2^n} |H_g(\lambda)|)/2. \tag{2}$$

2.5 Autocorrelation

The derivative of Boolean function g(x), with respect to a vector s, is defined as $g(x) \bigoplus g(x+s)$, where x and $s \in \mathbb{K}_2^n$. So, in polar form, derivative can be defined as $\widehat{g}(x)\widehat{g}(x+s)$. The autocorrelation of a function g is denoted by $A_g(s)$ and is defined by

$$A_g(s) = \sum_{x \in \mathbb{K}_2^n} \hat{g}(x)\hat{g}(x+s),$$

where $\hat{g}(x) = (-1)^{g(x)}$.

For a good Boolean function g, value of A_g should small.

2.6 Correlation Immunity

A Boolean function $g \in \mathcal{B}_n$ is said to be correlation immune (order m) if $H_g(\alpha) = 0$ for all $\alpha \in \mathbb{K}_2^n$ such that $1 \leq w_H(\alpha) \leq m$. Moreover, if g is balanced than it is called the m-resilient.

3 Non-dominated Sorting Genetic Algorithm II (NSGA-II) with Biasedness

Deb et al. [5] doveloped NSGA-II, that is a generational Multiobjective Optimization Evolutionary Algorithm (MOEA). It is based on three modules and we have explained the method in [7]. We have applied the algorithms on our developed method and got some good Boolean function [7]. But only this technique was not sufficient to get desired Boolean functions as complexity of method was comparatively high. Deb [4] discussed a sharing approach which uses a biased distance metric. By introducing biasedness means we give extra weightage to some specific objective function by introducing a constraint (same as objective function) into MOOP. In present paper, we have introduced a new concept of biasedness in NSGA-II to reduce the complexity. In our MOOP, we have formed first objective to optimize nonlinearity and nonlinearity is most important property here to optimize. So, in our MOOP, first objective and first constraint are same.

4 Formulation of MOOP

It consists of (i) Introduction of biasedness into MOOP and (ii) Application of NSGA-II.

(i) **Formulation of MOOP with biasedness:** Our main task is to form objective functions. To get optimum value of Nonlinearity, balancedness, autocorrelation and resiliency is our motive. We have formed first objective to optimize nonlinearity, second to optimize resiliency, and we have optimized autocorrelation by third objective. To get balanced functions, we have introduced two constraints. Nonlinearity is very important property. Hence to give extra weightage to first objective we have introduced concept of biasedness and added another constraints that is same as first objective.

First objective function: Based on the definition of nonlinearity [1,7]

$$nl = 2^{(n-1)} - 1/2(\max_{\lambda} H_g(\lambda)),$$

We know maximum value of nonlinearity for 6 variables is 48 and for 7 variables is 56. So, to form first objective function we have introduced a new constant say T. Now, we want nl to take the value equal to T. So, first objective can be formed as follows:

$$g^1 = |nl - T|, \tag{3}$$

g^1 is our the first objective function, where T is constant value for a fixed number of variables. (Here we take its value as 48 for 6 variables and as 56 for seven variables.)

Second objective function: Second objective is to optimize autocorrelation. So, we have directly assigned the value of autocorrelation equal to second objective.

To formulate second objective, we have used definition of autocorrelation (Definition 2.5). According to the above definition of autocorrelation, we have formulated

$$A_g(\lambda) = \sum_{x \in \mathbb{K}_2^n} (-1)^{g(x) \oplus g(x+\lambda)},$$

and $A_g(0)$ is maximum,

So,

$$g^2 = \max_{\lambda} |A_g(\lambda)| \tag{4}$$

is our second objective function, where $\lambda \in \mathbb{K}_2^n$ and $\lambda \neq zero$

Now,

$$g^1 = |nl - T|,$$

$$g^1 = |2^{(n-1)} - 1/2 \sum_{x \in \mathbb{K}_2^n} (-1)^{g(x) \oplus \lambda . x} - T|, \tag{5}$$

Similarly, for all $\lambda \in \mathbb{K}_2^n$,

$$g^2 = \max_{\lambda} \sum_{x \in \mathbb{K}_2^n} (-1)^{g(x) \oplus g(x+\lambda)}. \tag{6}$$

Now

Third objective function: According to the definition 2.6, for a Boolean function to be m resilient, value of Walsh Hadamard Transform should be zero corresponding to all $x \in \mathbb{K}_2^n$ having weight $\leq$ m. So, to form out third objective, we take all WHT corresponding to all such $x \in \mathbb{K}_2^n$. We added all WHT and assigned them to the third objective. Now our purpose is to minimize this third objective (equal to zero). This is because with zero value of third objective, we will get m-resilient functions. So, our third objective is,

$$g^3 = \sum_{\lambda} |H_f(\lambda)| \tag{7}$$

where $w_H(\lambda) \leq m$ for $\lambda \in \mathbb{F}_2^n$

So, we design MOOP as:

$$\left.\begin{array}{l} \min F = (g^1, g^2, g^3) \\ subject \quad to \\ \sum_{x \in \mathbb{K}_2^n} g(x) = 2^{n-1}, \\ nl = T. \end{array}\right\} \tag{8}$$

$\sum_{x \in \mathbb{K}_2^n} g(x)$ should be equal to 2^{n-1} for balanced function. To use biasedness sharing technique, the second constraint $nl = T$ is taken to give more weightage to the first objective.

(ii) **Application of above method:** After applying above method (with biasedness concepts) to the MOOP, we get the desired results. Results are given in Sect. 5. The list of parameters are listed in Table 2 (for 6 variables) and 3 (7 variables).

5 Result and Discussion

We got desired results by applying our method (In Sect. 4) on MOOP and got some good Boolean functions from cryptography point of view. These balanced functions have the best trade-off among non-linearity, autocorrelation and resiliency. In Table 1, we have listed those functions for 6 and 7 variables and parameters are given in Table 2 respectively. We have compared our results with literature [1,3] and can conclude that our results are at least as better.

Table 1. Obtained results

No. of variables	Previous results	Our results
6	nl = 48, $A_g = 8$, resiliency = 1	nl = 48, $A_g = 4$, resiliency = 1.
7	nl = 56, $A_g = 8$, resiliency = 1	nl = 56, $A_g = 8$, resiliency = 1

Table 2. Parameters

Parameters	For 6 variables	For 7 variables
Size of generation	2000	4000
Size of population	500	2000
Probability of crossover	0.8	0.8
Probability of mutation	0.1	0.11
Random seed number	0.9876	0.9976
Number of bits (for binary variables)	1	1
How many objective functions	3	3
How many constraints	2	2

6 Conclusion

In present paper, we have developed a new method to design good Boolean functions from cryptography point of view. We got Boolean functions for 6 and 7 variables that are better or at least comparable with [1,3]. So, we can conclude, our method is at least as better as the methods available in the literature.

References

1. Aguirre, H., Okazaki, H., Fuwa, Y.: An evolutionary multiobjective approach to design highly non-linear boolean functions. In: GECCO 2007, pp. 749–756 (2007)
2. Camion, P., Carlet, C., Charpin, P., Sendrier, N.: On correlation-immune functions. In: Feigenbaum, J. (ed.) CRYPTO 1991. LNCS, vol. 576, pp. 86–100. Springer, Heidelberg (1992). doi:10.1007/3-540-46766-1_6
3. Clark, J.A., Jacob, J.L., Stepney, S., Maitra, S., Millan, W.: Evolving Boolean functions satisfying multiple criteria. In: Menezes, A., Sarkar, P. (eds.) INDOCRYPT 2002. LNCS, vol. 2551, pp. 246–259. Springer, Heidelberg (2002). doi:10.1007/3-540-36231-2_20
4. Deb, K.: Multi-objective evolutionary algorithms: introducing bias among pareto-optimal solutions. In: Ghosh, A., Tsutsui, S. (eds.) Advances in Evolutionary Computing, pp. 263–292. Springer, New York (2003)
5. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm. IEEE Trans. Evol. Comput. **6**(2), 182–197 (2002)
6. Filiol, E., Fontaine, C.: Highly nonlinear balanced Boolean functions with a good correlation-immunity. In: Nyberg, K. (ed.) EUROCRYPT 1998. LNCS, vol. 1403, pp. 475–488. Springer, Heidelberg (1998). doi:10.1007/BFb0054147
7. Goyal, R., Yadav, S.P.: Design of Boolean functions satisfying multiple criteria by NSGA-II. In: Deep, K., Nagar, A., Pant, M., Bansal, J.C. (eds.) SocProS 2011. AISC, vol. 130, pp. 461–468. Springer, Heidelberg (2011). doi:10.1007/978-81-322-0487-9_45
8. Dobbertin, H.: Construction of bent functions and balanced Boolean functions with high nonlinearity. In: Preneel, B. (ed.) FSE 1994. LNCS, vol. 1008, pp. 61–74. Springer, Heidelberg (1995). doi:10.1007/3-540-60590-8_5
9. Kurosawa, K., Satoh, T.: Design of SAC/PC (l) of order k Boolean functions and three other cryptographic criteria. In: Fumy, W. (ed.) EUROCRYPT 1997. LNCS, vol. 1233, pp. 434–449. Springer, Heidelberg (1997). doi:10.1007/3-540-69053-0_30
10. Maitra, S., Pasalic, E.: Further constructions of resilient boolean functions with very high nonlinearity. IEEE Trans. Inf. Theory **48**(7), 1825–1834 (2002)
11. Millan, W., Clark, A., Dawson, E.: Heuristic design of cryptographically strong balanced Boolean functions. In: Nyberg, K. (ed.) EUROCRYPT 1998. LNCS, vol. 1403, pp. 489–499. Springer, Heidelberg (1998). doi:10.1007/BFb0054148
12. Chee, S., Lee, S., Lee, D., Sung, S.H.: On the correlation immune functions and their nonlinearity. In: Kim, K., Matsumoto, T. (eds.) ASIACRYPT 1996. LNCS, vol. 1163, pp. 232–243. Springer, Heidelberg (1996). doi:10.1007/BFb0034850
13. Sarkar, P., Maitra, S.: Nonlinearity bounds and constructions of resilient boolean functions. In: Bellare, M. (ed.) CRYPTO 2000. LNCS, vol. 1880, pp. 515–532. Springer, Heidelberg (2000). doi:10.1007/3-540-44598-6_32
14. Su, S., Tang, X., Zeng, X.: A systematic method of constructing Boolean functions with optimal algebraic immunity based on the generator matrix of the ReedMuller code. Des. Codes Crypt. **72**(3), 653–673 (2014)
15. Wang, Q., Tan, C.H.: A new method to construct Boolean functions with good cryptographic properties. Inf. Process. Lett. **113**(14–16), 567–571 (2013). Elsevier