Optimizing Cryptographic Strength of Substitution Layers in Symmetric-Key Cryptosystems

Christopher A. Wood

May 17, 2012

Agenda

1. Motivation and background
2. Security measurements
3. Optimization problem formulation
4. Optimization solution
5. Final remarks

Motivation

- Cryptography has wide variety of practical applications
  - Online banking, e-commerce, radio and telecommunication transmissions, etc
- One of the main purposes is to encrypt sensitive data
- How can we ensure the security of such data?

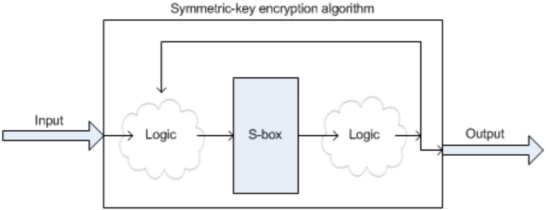
Background

Cryptographic algorithm security is measured by:

- Levels of confusion and diffusion
  - *Confusion* - Complexity of the relationship between the secret-key and ciphertext
  - *Diffusion* - Influence of single bit changes in the plaintext on the ciphertext
- Resilience to common cryptanalytic attacks
  - Linear cryptanalysis
  - Differential cryptanalysis

The substitution layer

A S-box is a common source for nonlinearity in symmetric-key cryptographic algorithms. It can be defined as function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$, where n is the number of bits needed to represent each element in the field.



S-box design

Which S-box configurations yield the highest measures of diffusion and confusion?

Security measurements

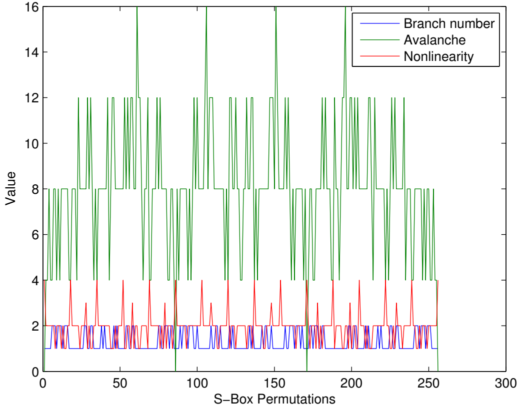
- Branch number
  - Measures lower bound on susceptibility to differential cryptanalysis
- Avalanche number
  - Measures the total number of bit changes for a single bit change in the input to the S-box
- Nonlinearity degree
  - Measures how much nonlinear "behavior" the S-box exhibits by counting the number of output elements that are directly proportional to its input.

Known optimal values

Security Measurement	Theoretical Optimal Value
Branch number (B_N)	n
Avalanche number (A_N)	$n^2 2^{n-1}$
Nonlinearity degree (P_S)	1^*

*$P_S \leq 2$ indicates that the function is "almost perfectly nonlinear", which is also a good value

Exhaustive search for 2-bit S-box



Branch number problem formulation

Branch Number - Minimize

$$B'_N(X) = -B_N(X) = -\min_{i,j \neq i}(\mathrm{wt}(i \oplus j) + \mathrm{wt}(X(i) \oplus X(j))),$$

subject to the constraints

$$0 \leq X(i) \quad \leq \quad 2^n - 1$$

where n is the number of bits needed to represent the design variables.

Avalanche number problem formulation

Avalanche Number - Minimize

$$A'_N(X) = -A_N(X) = -\sum_{i=0}^{n-1} \sum_{x \in \mathbb{F}_2^n} \mathrm{wt}(f(x) \oplus f(x \oplus 2^i))$$

subject to the constraints

$$\begin{aligned} 0 \leq X(i) &\leq 2^n - 1 \\ \sum_{i=0}^{n-1} \sum_{x \in \mathbb{F}_2^n} \mathrm{wt}(f(x) \oplus f(x \oplus 2^i)) - n2^{n-1} &\leq 0 \end{aligned}$$

where n is the number of bits needed to represent the design variables.

Nonlinearity degree problem formulation

Degree of Nonlinearity - Minimize

$$P_S(X) = \max_{0 \neq a,b} |\{x \in \mathbb{F}_2^n : S(x+a) - S(x) = b\}|,$$

subject to the constraints

$$0 \leq X(i) \leq 2^n - 1,$$

where n is the number of bits needed to represent the design variables.

Finding a shared solution

- Create a linear combination of each objective function
- Assign variable weights that correspond to the overall influence of each objective function

$$f(X) = w_1 A'_N(X) + w_2 B'_N(X) + w_3 P_S(X)$$

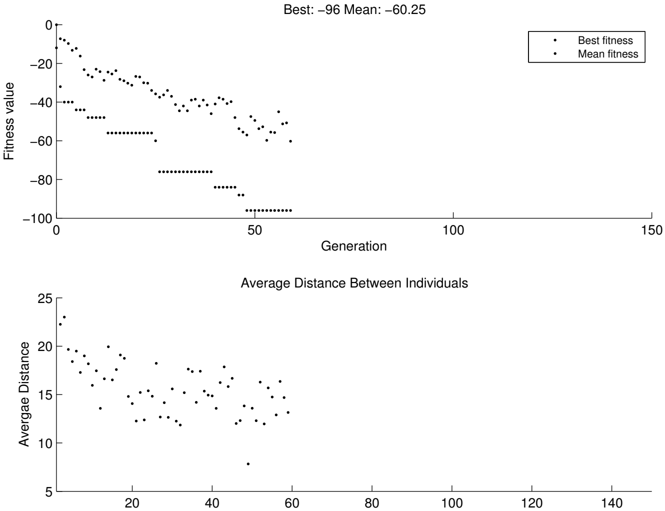
Candidate optimization methods

- Traditional MINLP methods
  - Branch and Bound (and all derivative) algorithms
- Evolutionary methods
  - *Genetic algorithm*

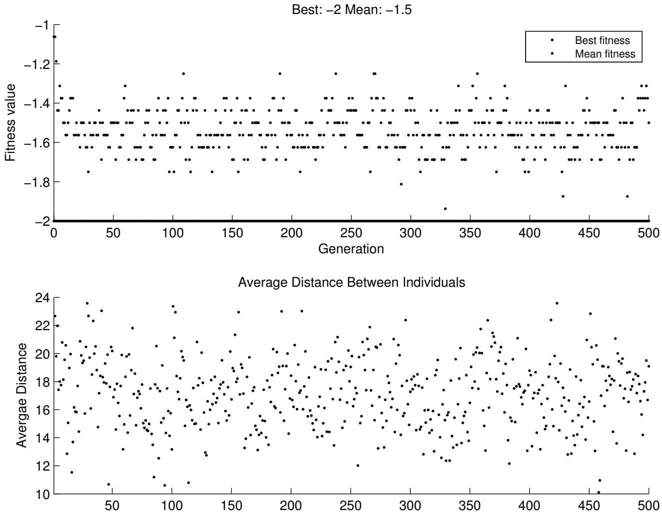
Genetic algorithm solver

- Objective functions are equivalent to fitness functions
- Randomized population generation and mutation
  - Optimal generations are chosen from a set of possible solutions
- No crossover function utilized
- Maximum of 500 generations and small stall limit

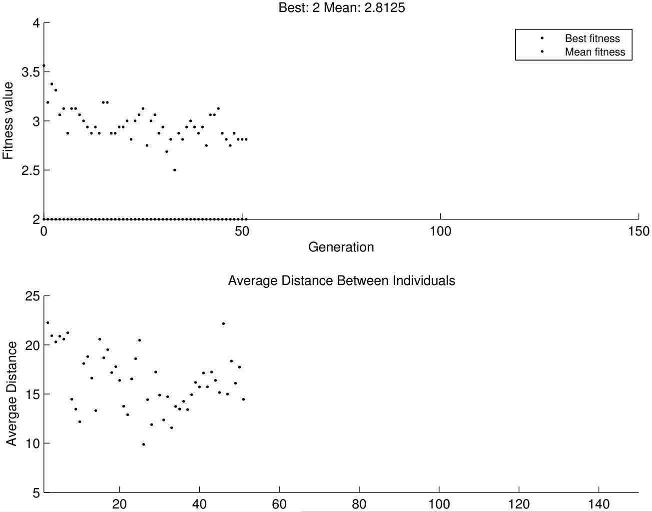
Avalanche number for 4-bit S-box



Branch number for 4-bit S-box



Nonlinearity degree for 4-bit S-box



Multi-objective solution for 3-bit S-box

w_1	w_2	w_3	S-Box Configuration	(A_N, B_N, P_S)
1	2^8	2^8	5 6 2 5 5 6 2 5	(40, 1, 8)
1	2^4	2^8	4 3 3 6 2 5 5 4	(50, 1, 6)
1	2^8	2^4	2 5 2 2 2 5 2 2	(24, 1, 8)
1	2^4	2^4	3 5 4 2 3 5 4 2	(40, 1, 8)

Solution Analysis

Measurement	**Results**
Branch number	Exponentially less effective with higher order S-boxes
Avalanche number	Logarithmically less effective with higher order S-boxes
Nonlinearity degree	Consistently effective with all order S-boxes
Multi-objective	Ineffective

Conclusions

- Evolutionary optimization algorithms are appropriate for cryptographic applications
  - Effectively finds solutions for some discontinuous and instable functions
  - Not effective at finding solution for multi-objective problems
- Difficult to find optimal S-box configurations with high security assurances
  - Does not replace hardened proofs of security