MS Thesis Preproposal

Large Substitution Boxes with Efficient Implementations

Christopher A. Wood
Committee Chair: Professor Stanisław Radziszowski
Department of Computer Science, Rochester Institute of Technology

1 Problem Statement

S(ubstitution)-boxes are the most common nonlinear building block of symmetric-key block ciphers. An S-box can be defined as a function $F : GF(2^n) \rightarrow GF(2^m)$. To measure the cryptographic strength of these functions it is common to represent them as vectorial Boolean functions, where $F(x) : (f_1(x), \ldots, f_m(x))$ and $f_i : GF(2^n) \rightarrow GF(2)$. Such a representation enables one to measure its nonlinearity (i.e. measure of distance from affine functions), algebraic immunity (i.e. difficulty measure of annihilator-based algebraic attacks), resiliency (i.e. balancedness and correlation immunity between input and output of the function), and differential uniformity (i.e. difficulty measure of differential cryptanalysis).

Naturally, S-box security comes at the price of computational efficiency. In resource constrained systems, such as VLSI circuits and embedded platforms, traditional LUT-based implementations of S-boxes are not feasible. In such systems, combinational logic circuits or software functions are commonly used to implement the S-box. For this reason, cryptographers have worked for many years to find a balance between the strength and efficiency of S-boxes. In addition, with the selection of the Advanced Encryption Standard, all of this research has been focused on 8-bit S-boxes (i.e. $F : GF(2^8) \rightarrow GF(2^8)$).

To our knowledge, there has not been any work that studies the strength and implementation efficiency (in terms of hardware area or total gates required) of higher order S-boxes. Therefore, in this work, we break away from tradition and focus our attention on 16-bit S-boxes defined as $F : GF(2^{16}) \rightarrow GF(2^{16})$. The problem is to study these S-boxes in the context of Boolean functions to determine their strength, and for each ideal candidate S-box, attempt to find optimal hardware and software implementations either through the use of Boolean function minimization or composite field arithmetic.

2 Proposed Approach

The first part of our research will focus on the cryptographic strength of various S-box definitions using analysis methods for Boolean functions. Clearly, exhaustively searching all $GF(2^{16^{16}})$ S-box definitions is infeasible, so our analysis will be constrained to S-boxes defined by operations in Galois Fields (i.e. $F(x) = F(x)^{-1}$, $x \in GF(2^{16})$) and built using the Maiorana-McFarland Boolean function construction technique. Time permitting, we may also explore Boolean function construction techniques targeted towards specific cryptographic properties, such as resiliency and algebraic immunity. The metrics collected for this part of the research include the nonlinearity, algebraic immunity, resiliency, and differential uniformity of each S-box candidate.

The second thread of our research will be to implement candidate S-boxes in FPGA hardware. Software versions of these functions will also be implemented and optimized to measure the throughput performance on low-end platforms (i.e. 32-bit PowerPC hardcore processors and 8-bit Atmel AVR microcontrollers). The metrics collected for this part of the research include hardware area (i.e. FPGA slices and total synthesized logic gates) and throughput (cycles per byte), as well as the software memory footprint for S-box functions and their throughput (cycles per byte).

3 Evaluation

To evaluate our work we will first formalize our experimental methodology and testbed for 8-bit S-boxes and compare the collected metrics against those published in the literature. Once verified for correctness, we will then scale up the S-boxes to 16-bits. Since this is a new line of research, we have no benchmark against which to compare our results. The primary outcome of this thesis will be the establishment of such a benchmark.