Some Studies On Information Set Decoding Algorithms and Universal Hash Functions

Author: Sreyosi Bhattacharyya
Supervisor: Prof. Palash Sarkar

Abstract. This thesis presents some studies on Information Set Decoding algorithms and Universal Hash Functions. In the context of Information Set Decoding (ISD) the thesis studies time/memory trade-off of ISD algorithms and in the context of universal hash functions, the thesis studies design and efficient implementations of polynomial hash functions defined over prime order fields.

A cornerstone of ISD algorithms is the algorithm proposed by Stern and it introduced the meet-in-the-middle collision search approach to ISD algorithms. Though this algorithm is more efficient in terms of asymptotic time complexity than the preceding algorithms proposed by Prange, Lee and Brickell and Leon, the minimum time complexity of Stern's algorithm is achieved by using a large amount of memory. The algorithms which have better time complexities than Stern's algorithm like MMT and BJMM have higher space complexities than that of Stern's algorithm. The memory complexities of Stern's algorithm, MMT and BJMM are of form `O(2^{cn})` where n is the code length and the value of the constant c depends on the corresponding algorithms. Hence studying time/memory trade-off of Stern's algorithm is relevant. Hence studying time/memory trade-off of Stern's algorithm is relevant. Our first contribution is a generalisation of Stern's information set decoding (ISD) algorithm. Stern's algorithm, a variant of Stern's algorithm due to Dumer, as well as a recent generalisation of Stern's algorithm due to Bernstein and Chou are obtained as special cases of our generalisation. Our second contribution is to introduce the notion of a set of effective time/memory tradeoff (TMTO) points for any ISD algorithm for given ranges of values of parameters of the algorithm. Such a set succinctly and uniquely captures the entire landscape of TMTO points with only a minor loss in precision. We further describe a method to compute a set of effective TMTO points. As an application, we compute sets of effective TMTO points for the five variants of the Classic McEliece cryptosystem corresponding to the new algorithm as well as for Stern's, Dumer's and Bernstein and Chou's algorithms. The results show that while Dumer's and Bernstein and Chou's algorithms do not provide any interesting TMTO points beyond what is achieved by Stern's algorithm, the new generalisation that we propose provide about twice the number of effective TMTO points that is obtained from Stern's algorithm. We have also discussed the consequences of the obtained TMTO

points to the classification of the variants of Classic McEliece in appropriate NIST categories.

Polynomial hashing technique constructs a univariate polynomial using the input message such that the coefficients of the polynomial are elements of a finite field and evaluates the polynomial at a secret point, called the key, of the same finite field to obtain the digest. The security requirement on such hash functions is that all differential probabilities are provably small. Hash families with low differential probabilities are called Almost XOR Universal and are a generalisation of universal hash functions. To construct the polynomial the input message is divided into ℓ number of blocks where each block is of equal size (the last block can be smaller than the rest of the blocks) and each of the blocks is an element of the finite field. These ℓ blocks form coefficients of the polynomial. Another approach of construction and evaluation has been proposed by Bernstein based on a previous work by Rabin and Winograd and such polynomials are called BRW polynomials. BRW polynomials require half the number of field multiplications than Horner based evaluations for the same input message. Polynomial hash functions have important applications in Cryptography including the construction of authentication schemes and methods of authenticated encryption schemes. Poly1305 is a polynomial hash function which has been widely used in various real-life scenarios. The AEAD scheme ChaCha20-Poly1305 is used in TLS, SSH etc. We provide an improved SIMD implementation of Poly1305. We propose a simple algorithmic technique (which we call the Balancing Technique) which leads to noticeable and significant improvements in terms of cycles/byte for various ranges of lengths of message. The message lengths considered by us lie in the range of 49 bytes to 4KB. Our next contribution is comprehensive study of usual polynomial based hashing and hashing based on BRW polynomials, and the various ways to combine them. Several hash functions are proposed and upper bounds on their differential probabilities are derived. Concrete instantiations are provided for the primes `2^{130}-5` and `2^{127}-1`. We have done an extensive 64-bit implementation of all the proposed hash functions in assembly targeted at modern Intel processors. The timing results suggest that using the prime `2^{127}-1` is significantly faster than using the prime `2^{130}-5`. Further, a judicious mix of the usual polynomial based hashing and BRW-polynomial based hashing can provide a significantly faster alternative to only usual polynomial based hashing. In particular, the timing results of our implementations show that our final hash function proposal for the prime is much faster than the well known Poly1305 hash function defined over the prime, achieving speed improvements of up to 40%.