New Hash Functions for Message Authentication

Hugo Krawczyk

IBM T.J. Watson Research Center
Yorktown Heights, NY 10598
(hugo@watson.ibm.com)

Abstract. We show that Toeplitz matrices generated by sequences drawn from small biased distributions provide hashing schemes applicable to secure message authentication. This work extends our previous results from Crypto'94 [4] where an authentication scheme based on Toeplitz matrices generated by linear feedback shift registers was presented.
Our new results have as special case the LFSR-based construction but extend to a much wider and general family of sequences, including several simple and efficient constructions with close to optimal security. Examples of the new constructions include Toeplitz matrices generated by the Legendre symbols of consecutive integers modulo a prime (of size significantly shorter than required by public-key modular arithmetic) as well as other algebraic constructions. The interest of these schemes extends beyond the proposed cryptographic applications to other uses of universal hashing (including other cryptographic applications).

1 Introduction

In *Crypto'94*, we introduced [4] a new scheme for hash functions suitable for message authentication in the symmetric key model. The scheme uses a linear feedback shift register sequence for the generation of a Boolean Toeplitz matrix that in turn is used to hash the message using matrix-vector multiplication. When combined with a one-time pad encryption of the hash value this schemes gives (provable) unconditional security.

The efficiency and simplicity of that construction makes it attractive conceptually as well as for practical use. In this paper we generalize the above scheme so that it can be used with a variety of different "weakly random" sequences as an alternative to LFSRs. We prove that any sequence taken from an *ε-biased distribution* of sequences can be used to generate a Toeplitz matrix that has all the security properties for authentication that the original LFSR-based construction had (and essentially the same strength of a completely random matrix!).

ε-biased distributions (introduced by Naor and Naor [7]) on bit sequences of length ℓ are characterized by the property that for any Boolean vector $\alpha \neq 0$ of length ℓ and for a sequence s chosen from that distribution, the probability of (α, s) (i.e., the scalar product modulo 2 of α and s) being 1 deviates from 1/2 by at most ε. (See Section 2.3 for a formal definition). These sequences prove to be useful for replacement of true randomness in different applications (see [7]). The advantage of these sequences over purely random ones is that they allow for easy generation,

L.C. Guillou and J.-J. Quisquater (Eds.): Advances in Cryptology - EUROCRYPT '95, LNCS 921, pp. 301-310, 1995.

simplicity and short description. All these properties translate in our case to a variety of attractive schemes for secure message authentication: simple, efficient, requiring short keys and short authentication tags.

In addition to LFSRs, which are a particular case of ε-biased sequences, this class of sequences includes several constructions [7, 1] from error correcting codes as well as other algebraic sequences; e.g., the sequence of Legendre symbols (or quadratic character) of consecutive integers modulo a prime (of size significantly shorter than required by public-key modular arithmetic).

Before stating our main result, we recall that in the setting of message authentication, we use Toeplitz matrices to first hash the message and then encrypt the hash value using a one-time pad (or stream cipher); the resultant encrypted hash value is the authentication tag for the message. Therefore, the communicating parties need to share a description of a particular hashing matrix as their shared secret key, and need to append to each transmitted message the corresponding authentication tag. This implies the need for a family of matrices that can be generated efficiently out of a short seed (the secret key), result in short hash values (the authentication tag) and still have the required security properties. The schemes we construct in this paper out of ε-biased sequences satisfy all these requirements with essentially the same security as provided by a completely random matrix (but while the later may require millions of random bits to describe the function, the ε-biased approach can do with keys in the order of 100 bits).

Next, we state our main theorem in terms of *otp-secure* hash functions, a notion defined in [4] (see Section 2.2 below). Informally, a family of hash functions is δ-otp-secure if the probability of an adversary to defeat the authentication is no more than δ where a message is authenticated by encrypting its hash value under a one-time pad (otp).

Main Theorem: *Let $\mathcal{T}_S$ be a family of Toeplitz matrices corresponding to sequences selected from an ε-biased distribution S. Then the hashing scheme that uses multiplication of the message by the Toeplitz matrix as the message hash is $(\frac{1}{2^n} + \varepsilon)$-otp-secure, where n is the length of the hash output.*

By using constructions of ε-biased sequences introduced by Alon et al. [1], we get explicit realizations of the above theorem. In particular, we show δ-otp-secure constructions where δ can be as close as desired to the optimal value 2^{-n} (this is optimal since the adversary can always guess the n-bit authentication tag with probability 2^{-n}). As an example, for any values of n and m, we present schemes that authenticate messages of length m with authentication tags of length n, and have security $\delta = 2^{-n+1}$. The shared key for describing such an authentication function is of length $2 \cdot (n + \log(n+m))$ (i.e., increases only logarithmically with the message which is typically much longer than n). These results are presented in Section 4.

Our proofs use as basic tools a characterization theorem from [4] that determines sufficient (and necessary) conditions for a family of hash functions to be secure for message authentication; and Discrete Fourier analysis for proving that ε-biased sequences induce Toeplitz hashing with the stated properties.

The interest of our work extends beyond the proposed cryptographic applications to other uses of universal hashing (including other cryptographic applications).

Our work extends the work of Krawczyk [4] that in turn follows the approach introduced by Carter and Wegman [10] of basing message authentication on hash functions. The same approach was followed by several authors, e.g., [8, 2]. We refer to [4] for a more complete survey of the relevant works.

ORGANIZATION: Section 2 presents the technical background and basic notions used throughout the paper. Section 3 presents the proof of the Main Theorem. Section 4 describes explicit constructions and their properties for message authentication.

2 Technical Background

In this section we introduce the main notions used throughout the paper as well as the technical background and tools.

2.1 Toeplitz matrices

Toeplitz matrices are characterized by having *fixed diagonals*. More precisely, each left-to-right diagonal is fixed, i.e., if $k - i = l - j$ for any indices $1 \leq i, k \leq n$, $1 \leq j, l \leq m$, then $A_{i,j} = A_{k,l}$. See Figure 1 for an example. Notice that an $n \times m$ Toeplitz matrix is fully described by its first column and first row (i.e., by $n+m-1$ elements).

Notation: *To any given sequence s of $n + m - 1$ bits we associate an $n \times m$ Toeplitz matrix T_s, where the elements of s determine the first column and first row of T_s, and therefore the whole matrix. We map the first n elements of s into the first column of T_s starting from the bottom (i.e., $T_s(n,1) = s_1, \ldots, T_s(1,1) = s_n$) and then the last m bits of s into the first row of T_s (i.e., $T_s(1,1) = s_n, T_s(1,2) = s_{n+1}, \ldots, T_s(1,m) = s_{n+m-1}$). See Figure 1. We say that s* generates *T_s*.

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Fig. 1. The 4×8 Toeplitz matrix $T_{110011010001}$

Toeplitz matrices of dimension $n \times m$ can be used to hash messages of length m by multiplying the message (seen as a column vector) by the matrix. The resultant hash value has length n. It is well-known that the family of Toeplitz matrices T_s with s chosen at random constitutes a strongly universal$_2$ family of hash functions (see [6]). One of the main results in [4] is to show that when the sequence s is generated out of only n random bits using a random irreducible LFSR, the resultant family is still almost universal (or ε-balanced in the terminology of [4] – see Section 2.2 below). This is especially important when, as it is the case in practice, $n << m$. Here we show that this result extends to *any* family of sequences s which are ε-biased distributed. Explicit constructions are presented in Section 4.

2.2 ε-balanced hashing and authentication

Following the approach initiated by Carter and Wegman [10] (and developed by several authors – e.g., [8, 2, 4]) we study the message authentication model in which the communicating parties share a specific hash function h_k, chosen out of a family of hash functions H, and use this particular function h_k to authenticate multiple messages. The function h_k is described by a (secret) key k shared by the parties. Each message to be authenticated is first hashed using h_k and the resultant hash value encrypted by xor-ing it with a secret one-time pad known only to the legitimate parties. We use m to denote the length of the message M, and n for the length of the resultant hash value $h_k(M)$. Notice that all the functions described here can work, in principle, with arbitrary and variable length messages (typically, the security of the authentication degrades logarithmically with the length of messages). The output length n is the same for all messages and can be thought of as a security parameter.

The task of an adversary A that tries to break the authentication is to intercept a message M, sent between the legitimate parties, together with its legitimate authentication tag $h_k(M) \oplus r$ and replace it with another message M' (which may depend on M and other messages exchanged between the parties in the past [1]) for which A can produce the legal authentication tag $h_k(M') \oplus r$. Notice that we assume that A does not know k or r. (The above is called a *substitution* attack; an *impersonation* attack where the adversary initiates the sending of a fake message without the legitimate parties having a communication between them trivially has the minimal probability 2^{-n} to succeed because of the use of one-time pads).

The following definitions and theorem are from [4].

Definition 1. A family H of hash functions is called *ε-otp-secure* if for any message M, and for h_k chosen according to the distribution on H, no adversary succeeds in the above scenario with probability larger than ε.

Definition 2. A family of hash functions H with a probability distribution attached to it is called *ε-balanced* if

$$\forall M \neq 0,\ \forall b,\quad Prob_h(h(M) = b) \leq \varepsilon$$

where the probability is taken for h chosen according to the distribution on H.

The relation between the above two definitions is given by the following theorem.

Theorem 3 [4]. *If H is a family of linear functions relative to the bitwise XOR operation then H is ε-otp-secure if and only if H is ε-balanced.*

Our goal is to prove that the families of functions that we build in this paper (i.e., Toeplitz hashing generated by small biased sequences) are secure for use in message authentication, namely, they are ε-otp-secure for a very small ε. Since these functions use (Boolean) matrix multiplication for hashing then they are linear relative to XOR, and by virtue of the above theorem our task reduces to prove that they are ε-balanced for small ε. This proof is presented in Section 3.

[1] we assume a *chosen message attack* in which messages exchanged between the legitimate parties can be chosen by the adversary

2.3 ε-biased distributions

ε-biased distributions were introduced by Naor and Naor [7] (following the work of Vazirani [9]) as a tool for constructing small sample spaces, or more generally as a tool for replacement of truly random sequences with more "compact" and easier to generate sequences. This approach works in a variety of applications as shown in [7] and subsequent works. In this paper we show how ε-biased distributions are useful for generating efficient and short-key authentication functions. More precisely, we prove that Toeplitz matrices generated out of ε-biased sequences preserve essentially the same properties for hashing and authentication as completely random matrices.

Definition 4. Let S be a distribution on sequences of length ℓ. Let (α, s) denote the scalar product modulo 2 of $\alpha \in \{0,1\}^\ell$ and $s \in \{0,1\}^\ell$. Then,

1. S is said *to pass the linear test α with bias ε* if $|Prob((\alpha, s) = 1) - \frac{1}{2}| \leq \varepsilon$ (the probability taken over the choice of s from the distribution S).
2. S is said to be an *ε-biased distribution* if it passes *all* linear tests $\alpha \neq 0$ with bias ε.

The case $\varepsilon = 0$ corresponds to the uniform distribution; therefore, ε-biased distributions can be viewed as approximations to the uniform distribution. However, even for very small ε, there may be significant distinctions between ε-biased and uniform distributions. In the negative side, ε-biased distributions can be weakly random (such is the case, e.g., of LFSR sequences that are exponentially small biased but can be predicted from a very short prefix, or the set of all binary strings with Hamming weight divisible by three, proven to be exponentially small biased in [3]). Fortunately, there is a positive side. First there exist some very simple and efficient constructions for ε-biased sequences. Examples of such constructions, due to Alon et al. [1], are presented in Section 4. Second, in some applications these weakly random sequences can replace truly random bits with the advantage of simplicity and efficiency of generation. The results in this paper are an example of such an application.

Multiple tests If S is an ε-biased distribution then by definition S passes all (non-zero) linear tests with bias at most ε. Now, let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be k elements from $\{0,1\}^\ell$ and $b_1, b_2, \ldots, b_k$ be k bits. What can we say about the *simultaneous* probability that $(\alpha_1, s) = b_1, (\alpha_2, s) = b_2, \ldots, (\alpha_k, s) = b_k$? Notice that if the set of α_i's is linearly independent and s is chosen uniformly from $\{0,1\}^\ell$ then the above probability would be 2^{-k}. The following theorem states that for s drawn from the ε-biased distribution S this probability is at most $2^{-k} + \varepsilon$.

Theorem 5. *Let S be an ε-biased distribution on sequences of length ℓ. Let A be a $k \times \ell$ matrix of (full) rank k, and let b be a column vector of length k then*

$$Prob_s(A \cdot s = b) \leq \frac{1}{2^k} + \varepsilon$$

for s taken from the distribution S.

Notice that this is an extension of Definition 4, which corresponds to the particular case of $k = 1$. In other words, the power of passing single linear tests extends to the power of passing multiple simultaneous tests.

As we will show in Section 3, the proof of our Main Theorem reduces to the above theorem.

Naor and Naor [7] prove this property using a result by Vazirani [9, 1] that connects between ε-biased distributions and k-wise independence; here we present a direct proof using Discrete Fourier analysis (and get Vazirani's Theorem as a corollary). We present the proof of Theorem 5 in Section 3.2.

2.4 Discrete Fourier Transform

In this section we bring some minimal technical background and known facts on discrete Fourier transform, and its relationship to ε-biased distributions,that we use in our proof of Theorem 5

The set of real valued functions defined on the Boolean domain $\{0,1\}^\ell$, i.e., functions $f : \{0,1\}^\ell \to \mathbf{R}$, has an orthonormal basis composed of the following functions χ_σ, $\sigma \subseteq \{1, \cdots, \ell\}$:

$$\chi_\sigma(x_1, \cdots, x_\ell) = \begin{cases} +1 & \text{if } \sum_{i \in \sigma} x_i \text{ is even} \\ -1 & \text{if } \sum_{i \in \sigma} x_i \text{ is odd} \end{cases}$$

In other words, each real valued function over $\{0,1\}^\ell$ can be written as a (unique) linear combination of the functions χ_σ. The respective coefficients are denoted by $\hat{f}(\sigma)$, i.e., $f = \sum_\sigma \hat{f}(\sigma)\chi_\sigma$. This representation is called the *(discrete) Fourier transform* of the function f.

Property 1 *For Boolean f (i.e., $f : \{0,1\}^\ell \to \{0,1\}$), the coefficients $\hat{f}(\sigma)$ have the following special form:*

$$\hat{f}(\sigma) = Pr[f(x) = \oplus_{i \in \sigma} x_i] - Pr[f(x) \neq \oplus_{i \in \sigma} x_i] = 2 \cdot Pr[f(x) = \oplus_{i \in \sigma} x_i] - 1$$

where $x = (x_1, x_2, \ldots, x_\ell)$ is chosen uniformly at random.

We need the following definition and property of the L_1-norm of the function f.

Definition 6. Let f be a function from $\{0,1\}^\ell$ to the real numbers. Define $L_1(f) = \sum_\sigma |\hat{f}(\sigma)|$.

Property 2 *Let f and g be functions from $\{0,1\}^\ell$ to the real numbers. Then, $L_1(fg) \leq L_1(f)L_1(g)$.*

The connection between ε-biased distributions and Fourier transform is given by the following Lemma. Notice that probability distributions over the set of strings $\{0,1\}^\ell$ are real valued functions and therefore their Fourier transform is well-defined.

Lemma 7. *Let μ be an ε-biased probability distribution over $\{0,1\}^\ell$. Then, for every subset $\sigma \subset \{1 \ldots \ell\}$, $|\hat{\mu}(\sigma)| \leq \varepsilon 2^{-\ell}$.*

The following lemma due to Kushilevitz and Mansour [5] relates ε-biased distributions and the norm $L_1(f)$, and plays a central role in the proof of our Main Theorem.

Lemma 8 [5].

$$|E_U[f] - E_\mu[f]| \leq \varepsilon L_1(f)$$

where E denotes expectation, U is the uniform distribution and μ is an ε-biased distribution.

3 Main Result

This section is devoted to prove the Main Theorem of this paper (see the Introduction), namely, that Toeplitz matrices generated by ε-biased sequences constitute a family of $(2^{-n} + \varepsilon)$-otp-secure hash functions. By Theorem 3, this task reduces to proving the following result.

Theorem 9. *Let S be distribution on sequences of length $n+m-1$ bits. Let $\{T_s\}_s$ be the set of $n \times m$ Toeplitz matrices generated by the elements $s \in S$. If the distribution S is ε-biased then the family $\{T_s\}_s$ with the distribution induced by S is a $(2^{-n}+\varepsilon)$-balanced family of hash functions.*

3.1 Proof of Theorem 9

Following the definition of ε-balanced hash functions, we need to show that, for any given vector b of length n and message $M \neq 0$ of length m, the probability that $T_s \cdot M = b$ is at most $(2^{-n} + \varepsilon)$. This probability is taken over the distribution on T_s induced by choosing s from the ε-biased distribution S.

We start by transforming the representation of the problem. Notice that in a Toeplitz matrix each row is shifted (to the right) relative to the previous row, with a new element set to the first position of the row. This allows as to "swap" the roles of the sequence s and the message M in the following way.

We generate a new $n \times (n+m-1)$ matrix A_M which is cyclic and is defined by its first row containing $n-1$ zeros and then the entries of the vector M (viewed now as a row vector). Each new row in A_M is defined as a cyclic shift from right to left relative to the previous row (e.g., the second row of A_M contains the vector M but this time prepended by $n-2$ zeros and appended with one zero). It is easy to see that the following relationship, where the sequence s is represented by a $n+m-1$ column vector, holds. (See Figure 2 for an illustration of $A_M \cdot s$).

Lemma 10.

$$T_s \cdot M = b \quad \textit{if and only if} \quad A_M \cdot s = b\ .$$

Therefore, the proof of Theorem 9 (and then of the Main Theorem) reduces to prove that for any $M \neq 0$, $Prob(A_M \cdot s = b) \leq 2^{-n} + \varepsilon$, for s chosen according to the distribution S. But due to the special form of A_M this is a special case of Theorem 5 with $k = n$ and $\ell = n+m-1$. We conclude the proof of our main result by proving Theorem 5.

3.2 Proof of Theorem 5

We start by defining the following function.

$$f_{A,b}(s) = \begin{cases} 1 \text{ if } A \cdot s = b \\ 0 \text{ otherwise} \end{cases}$$

This definition can be specialized to $f_{\alpha,\beta}(s)$ where α is a vector of length ℓ and β a bit.

Lemma 11. $L_1(f_{\alpha,\beta}) = 1$

Proof. The proof uses Property 1 and the definition of the norm L_1. Details are omitted.

Lemma 12. $L_1(f_{A,b}) \leq 1$

Proof. Using Property 2 and the fact that $f_{A,b} = \prod_{i=1}^{k} f_{\alpha_i,b_i}$, where $\alpha_1, \ldots, \alpha_k$ are the rows of the matrix A and $b = (b_1, \ldots, b_k)^T$, we get

$$L_1(f_{A,b}) \leq \prod_{i=1}^{k} L_1(f_{\alpha_i,b_i}) = 1$$

where the last equality is derived from Lemma 11.

Now, we can conclude the proof of the theorem. Notice that, by the definition of $f_{A,b}$, for any distribution on the sequences s we have the expectation $E(f_{A,b}) = Prob(A \cdot s = b)$. For the uniform distribution this gives $E_U(f_{A,b}) = Prob_U(A \cdot s = b) = 2^{-k}$, where the last equality uses the fact that A has full rank k. Then, using Lemma 8 and Lemma 12, we get:

$$\left|Prob_S(A \cdot s = b) - 2^{-k}\right| = \left|E_S(f_{A,b}) - E_U(f_{A,b})\right| \leq \varepsilon \cdot L_1(f_{A,b}) \leq \varepsilon.$$

That is, $Prob_S(A \cdot s = b) \leq 2^{-k} + \varepsilon$, which proves the theorem. ·♡

4 Explicit Constructions

Our Main Theorem tells us that Toeplitz hashing generated out of ε-biased sequences, and combined with a one-time pad, provides a secure message authentication scheme in which no adversary can break the system with chance better than $2^{-n} + \varepsilon$. (Naturally, if the one-time pad is generated using a pseudorandom generator or stream cipher the security of the whole authentication scheme reduces to that of the pseudorandom generator.)

The advantage of using ε-biased sequences (to define the Toeplitz matrices) as opposed to purely random bits is that the former can be generated efficiently out of a short random seed (typically, in the order of 100 bits while a pure random Toeplitz matrix would require millions of random bits in order to hash a few megabits of

information). In particular, this implies a short shared authentication key for specifying a particular matrix. Moreover, the process of generation of the whole matrix is efficient and the resultant authentication tags short. Also, let us stress that in practical applications there is no need to simultaneously generate or work with the whole matrix, but only small portions of it (e.g., one column at a time).

Here we present several examples of explicit constructions of ε-biased sequences due to Alon et al. [1], with each of these constructions translating, using our results, into efficient authentication schemes. The first example is intended to show how the result of [4] about the quality of LFSR-based Toeplitz hashing for authentication can be derived as a special case of our results. We use ℓ to denote the length of the ε-biased sequences in these constructions and r to denote the number of random bits required to generate a sequence. In our setting of $n \times m$ Toeplitz matrices we have $\ell = n + m - 1$, and r is the length of the key that determines the specific shared matrix.

LFSR CONSTRUCTION. Sequences are determined by a random seed of length $r/2$ and an irreducible polynomial over GF(2) of degree $r/2$. The sequence of ℓ bits is generated using an LFSR loaded initially with the above seed and the connections corresponding to the irreducible polynomial.

LEGENDRE SYMBOL CONSTRUCTION. Let p be a fixed prime number of length r. The sequences s are generated out of a random number $x \in \{0, \ldots, p-1\}$. For $1 \leq i \leq \ell$, s_i is defined as 1 if $x + i$ is a quadratic residue modulo p or 0 otherwise. (We note that the prime p is not part of the secret key but a public value).

SCALAR PRODUCT CONSTRUCTION. Sequences s are determined by two random elements $x, y \in GF(2^{r/2})$. For $1 \leq i \leq \ell$, s_i is defined as the scalar product (x^i, y) where x^i is the i-th power of x as an element of $GF(2^{r/2})$.

Theorem 13 [1]. *The above three constructions produce ε-biased distributions on sequences of length ℓ with $\varepsilon = \frac{\ell}{2^{r/2}}$. Each sequence is generated out of r initial bits.*

Combining this with our Main Theorem we get the following result.

Theorem 14. *The families of $n \times m$ Toeplitz matrices resulting from each of the above three constructions by putting $l = n + m - 1$ constitute δ-otp-secure hashing schemes with $\delta = \frac{1}{2^n} + \frac{n+m-1}{2^{r/2}}$ and with each matrix described by a key of length r.*

Notice that for any fixed value of n, one can choose r such that δ gets as close as desired to the optimal value of 2^{-n}. In particular, for $r = 2n + 2\log(n + m - 1)$ one gets $\delta = 2^{-n+1}$ (notice that r increases only logarithmically with the size of messages). By choosing $r = 2n$ one gets $\delta = \frac{n+m}{2^n}$. In particular, for the LFSR construction the later parameters coincide exactly with the construction in [4]. Interestingly enough, the general bound proved here results in a just slightly larger bound than the $m/2^n$ bound derived in [4] specifically for the LFSR construction.

Finally, we remark that additional constructions of ε-biased distributions exist, some of them based on techniques from error correcting codes. In particular, using dual BCH codes. See [7, 1].

$$\begin{pmatrix} 0&0&0&1&1&1&1&1&1&1&1 \\ 0&0&1&1&1&1&1&1&1&1&0 \\ 0&1&1&1&1&1&1&1&1&0&0 \\ 1&1&1&1&1&1&1&1&0&0&0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Fig. 2. The product $A_M \cdot s$ corresponding to $M = 1^8$ and $s = 11001101001$

Acknowledgments: I thank Oded Goldreich for helpful conversations regarding small biased distributions.

References

1. Noga Alon, Oded Goldreich, Johan Hastad, and Rene Peralta. Simple constructions of almost k-wise independent random variables. In 31^{th} *Annual Symposium on Foundations of Computer Science, St. Louis, Missouri*, pages 544–553, October 1990.
2. Bierbrauer J., Johansson T., Kabatianskii G., and Smeets, B., "On Families of Hash Functions via Geometric Codes and Concatenation", *Crypto'93*
3. G. Even. *Construction of small probability spaces for simulation.* M.Sc. thesis, Dept. of Computer Science, Technion, August 1991.
4. Krawczyk, H., "LFSR-based Hashing and Authentication", *Advances in Cryptology - CRYPTO 94 Proceedings, Lecture Notes in Computer Science Vol. 839, Springer-Verlag, Y. G. Desmedt, ed* 1994, pp. 129-139.
5. E. Kushilevitz and Y. Mansour. "Learning decision trees using the Fourier spectrum", *SIAM Journal on Computing* 22(6) 1331-1348, December 1993.
6. Mansour, Y., Nisan, N., and Tiwari, P., "The Computational Complexity of Universal Hash Functions", *Theoretical Computer Science*, 107(1):121–133, 1993.
7. Joseph Naor and Moni Naor. Small bias probability spaces: efficient construction and applications. *SIAM Jour. on Computing*, Vol. 22, No. 4, 1993, pp. 838-856.
8. Stinson, D.R., "Universal hashing and authentication codes", *Proc. of Crypto'91*, pp. 74-85.
9. Vazirani, U.V., "Randomness, Adversaries and Computation", Ph.D. Thesis, EECS, UC Berkeley, 1986.
10. Wegman, M.N., and Carter, J.L., "New Hash Functions and Their Use in Authentication and Set Equality", *JCSS*, 22, 1981, pp. 265-279.