Universal Hashing and Multiple Authentication

M. Atici[1,2] and D. R. Stinson[1,3]

[1] Computer Science and Engineering Department
University of Nebraska, Lincoln, NE 68588
[2] atici@cse.unl.edu
[3] stinson@bibd.unl.edu

Abstract. In this paper, we study unconditionally secure codes that provide authentication without secrecy. Our point of view is the universal hashing approach pioneered by Wegman and Carter in 1981. We first compare several recent universal-hashing based constructions for authentication codes. Then we generalize the theory of universal hashing in order to accommodate the situation where we would like to authenticate a sequence of messages with the same key. Unlike previous methods for doing this, we do not require that each message in the sequence have a "counter" attached to it.

Keywords: authentication code, universal hashing.

1 Introduction

In this paper, we study the application of universal hashing to the construction of unconditionally secure authentication codes without secrecy. This idea is due to Wegman and Carter [16], who gave a construction in 1981 which is extremely useful when the number of authenticators is small compared to the number of possible source states (plaintext messages). In 1991, Stinson [13] gave formal definitions of relevant classes of hash functions, and obtained some improvements to the Wegman-Carter construction. Since 1991, several authors have given improved constructions for authentication-without-secrecy that use universal hashing either implicitly or explicitly. Many of the results are in fact very similar, but do not appear so because they are presented using different notations and terminology. We give a brief comparison of the known constructions and their efficiency, as measured by the amount of secret key that has to be shared in order to authenticate a given amount of information with a given level of security.

The other main contribution of this paper is to generalize the theory of universal hashing in order to accommodate the situation where we would like to authenticate a sequence of messages with the same key. Unlike previous methods for doing this, we do not require that each message in the sequence have a "counter" attached to it. We provide necessary definitions and theory, and then give a construction which acheives our goals.

The remainder of this paper is organized as follows. Section 2 is a brief review of the necessary background of authentication codes. Section 3 gives relevant definitions from universal hashing. We also compare known authentication codes in

N. Koblitz (Ed.): Advances in Cryptology - CRYPTO '96, LNCS 1109, pp. 16-30, 1996.

this section. Section 4 reviews counter-based multiple authentication. In Section 5 multiple authentication without counters is introduced. Section 6 provides composition constructions for the relevant hash families. Finally in Section 7 we use our constructions to obtain some specific families of codes for multiple authentication.

2 Authentication Codes

Authentication codes were invented in 1974 by Gilbert, MacWilliams and Sloane [5], and the general theory of unconditional authentication was developed by Simmons (see, e.g., [11]) In this section we will give a brief review of standard terminology and basic results on authentication-without-secrecy.

In the usual model for authentication, there are three participants: a *transmitter*, a *receiver*, and an *opponent*. The transmitter wants to communicate some information to the receiver using a public communications channel. The *source state* (i.e., plaintext) is concatenated with an *authenticator* to obtain a *message* which is sent through the channel. An *authentication rule* (or *key*) e defines the authenticator $e(s)$ to be appended to the source state s. We assume the transmitter has a key source from which he obtains a key. Prior to any message being sent, this key is communicated to the receiver by means of a secure channel.

We will use the following notation. Let $\mathcal{S}$ be a set of k source states; let $\mathcal{A}$ be a set of n authenticators; define $\mathcal{M} = \mathcal{S} \times \mathcal{A}$; and let $\mathcal{E}$ be a set of authentication rules. Each authentication rule $e : \mathcal{S} \to \mathcal{A}$.

Assume that the same key is used to authenticate up to w consecutive source states, where w is some fixed positive integer. Suppose an opponent observes $i \leq w$ distinct messages which are sent using the same key. The opponent has the ability to introduce new messages into the channel and/or to modify existing messages. Assume the opponent places a message $m' = (s', a')$ into the channel by either of these methods, where m' is distinct from the i messages already sent. That is, if e is the key being used, then the opponent is hoping that $a' = e(s')$. In [9], Massey calls this a *spoofing attack* of order i.

The special cases $i = 0$ and $i = 1$ have received the most attention. The case $i = 0$ is called *impersonation*, and the case $i = 1$ is called *substitution*.

The receiver and transmitter will choose a probability distribution for $\mathcal{E}$, called an *authentication strategy*. It is assumed that the opponent knows the authentication strategy being used. Then, for each $i \geq 0$, it is possible to compute Pd_i, which is the probability that the opponent can deceive the transmitter/receiver with a spoofing attack of order i. The following lower bound on Pd_i is given in [9].

Theorem 1. *Suppose we have an authentication code (without secrecy) with n authenticators. Then $Pd_i \geq 1/n$ for all $i \geq 0$.*

3 Universal Hashing

In this paper, we are interested in authentication codes obtained from universal hash families. We recall some definitions from [12] of various types of relevant hash families.

Definition:

- An $(N; m, n)$ *hash family* is a set $\mathcal{F}$ of N functions such that $f : A \to B$ for each $f \in \mathcal{F}$, where $|A| = m, |B| = n$. There will be no loss in generality in assuming $m \geq n$.
- An $(N; m, n)$-hash family is ϵ-*universal* provided that for any two distinct elements $x_1, x_2 \in A$, there exist at most ϵN functions $f \in \mathcal{F}$ such that $f(x_1) = f(x_2)$. We will use the notation ϵ-U as an abbreviation for ϵ-*universal*.
- An $(N; m, n)$ hash family is ϵ-*almost-strongly-universal* provided that the following two conditions are satisfied:
  1. for any $x \in A$ and any $y \in B$, there exist exactly N/n functions $f \in \mathcal{F}$ such that $f(x) = y$.
  2. for any two distinct elements $x_1, x_2 \in A$ and for any two (not necessarily distinct) elements $y_1, y_2 \in B$, there exist at most $\epsilon N/n$ functions $f \in \mathcal{F}$ such that $f(x_i) = y_i$, $i = 1, 2$.

  We will use the notation ϵ-ASU as an abbreviation for ϵ-almost-strongly-universal.
- An $(N; m, n)$-hash family $\mathcal{F}$ of functions from A to B is *strongly-universal* provided that, for any two distinct elements $x_1, x_2 \in A$, and for any two (not necessarily distinct) elements $y_1, y_2 \in B$, we have

$$|\{f \in \mathcal{F} : f(x_i) = y_i, i = 1, 2\}| = \frac{N}{n^2}.$$

  We will use the notation SU as an abbreviation for strongly-universal.

It is not difficult to see that a hash family is SU if and only it is $\frac{1}{n}$-ASU. ϵ-ASU hash families can be used in an obvious way for authentication, where each function in the family corresponds to a key. If we have such a class $\mathcal{F}$ of hash functions from A to B, then we can think of the elements of A as source states and the elements of B as authenticators. Each hash function gives rise to an authentication rule, and the authentication rules are used with equal probability. The proof of the following theorem is straightforward.

Theorem 2. *[12] If there exists an ϵ-$ASU(N; m, n)$ hash family, $\mathcal{F}$, then there exists an authentication code without secrecy for m source states, having n authenticators and N authentication rules, such that $Pd_0 = 1/n$ and $Pd_1 \leq \epsilon$.*

We see from Theorem 1 that SU families achieve the minimum possible deception probability Pd_1. The observation of Wegman and Carter [16] is that it is possible to construct ϵ-ASU hash families, having ϵ a bit larger than $1/n$, that

are much smaller than SU hash families. In terms of the resulting authentication codes, this means that if we allow a slightly larger deception probability Pd_1, then we can reduce the key length very significantly.

Many papers have used this approach, either implicitly or explicitly, for example Wegman and Carter [16], Stinson [12], den Boer [4] , Taylor [15], Bierbrauer, Johansson, Kabatianskii and Smeets [3], Krawczyk [7], Stinson [13], Krawczyk [8], Rogaway [10] and Bierbrauer [1].

In fact, the construction of ASU hash families typically is accomplished by one of two means:

- composition of a U family and a (smaller) ASU family (this is the approach used by Wegman and Carter [16])
- composition of a ΔU family [14] (also known as an AXU family [10]) with a one-time pad (this approach was first used by Krawczyk [7]).

Further discussion and examples of these two techniques can be found in the expository paper by Stinson [14].

3.1 Comparison of Authentication Codes

In this section, we briefly compare authenticator length and key length of for several constructions of authentication codes. To be specific, we consider the problem of authenticating an a-bit plaintext with a b-bit authentication tag. The number of key bits is denoted by ℓ. (In other words, we have an ϵ-$ASU(2^\ell; 2^a, 2^b)$ hash family.) In every code mentioned, $Pd_0 = 1/2^b$, but various values of Pd_1 are obtained, depending on the construction used.

1. Wegman-Carter ([16, §3], 1981).
   Here $s = b + \lceil log(log\ a) \rceil$, $\ell = 4s\ log\ a$ and $Pd_1 = 1/2^{b-1}$.
2. Stinson ([12, Theorem 6.2], *CRYPTO '91*).
   Here $a = b2^i$, $\ell = (i+2)b$ and $Pd_1 = (i+1)/2^b$.
3. Taylor ([15, §2], *EUROCRYPT '94*).
   This is identical to the previous construction of Stinson.
4. den Boer ([4, §2], 1993).
   Here $a = bi$, $\ell = 2b$ and $Pd_1 = i/2^b$.
5. Bierbrauer, Johanson, Kabatianskii, and Smeets[4] ([3, p. 336], *CRYPTO '93*).
   Here $a = (b+s)(2^s+1)$, $\ell = 3b + 2s$ and $Pd_1 = 1/2^{b-1}$.
6. Stinson ([13, Theorem 6.3], 1994).
   Here $s = b + \lceil log(log\ a) \rceil$, $r = \lceil log(a/s) \rceil$, $\ell = (r+1)s + b$ and $Pd_1 = 1/2^{b-1}$.
7. Krawczyk ([7, Theorem 7] Theorem 7, *CRYPTO '94*).
   Here $\ell \approx 2b - log\ b$ and $Pd_1 = (a+b)/2^{b-1}$.
8. Krawczyk ([7, Theorem 8] Theorem 8, *CRYPTO '94*).
   Here $\ell \approx 3b - log\ b$ and $Pd_1 = a/2^{b-1}$.

[4] In [6] (*CRYPTO '96*), Helleseth and Johansson give some constructions that achieve identical and/or slightly better results. Their approach also has the advantage that the parameters are a bit more flexible than this construction.

9. Rogaway ([10, Theorem 11], *CRYPTO '95*).
   Here $a = wA$, $b = wB$ (where $A \leq B^3/6$), $\ell \approx 3A \log B + wB$ and $Pd_1 \approx 3348/(B^6 - 6B^3A)$. (Note: Since [10, Theorem 11] produces a U family (actually a ΔU family), a one-time pad is also needed to obtain the authentication code. This accounts for the "extra" $b = wB$ key bits.)

Remarks:

- Constructions 1–6 all use the Wegman-Carter approach. Constructions 7–9 use the idea of composing a ΔU family with a one-time pad.
- Constructions 1, 5 and 6 have $Pd_1 = 1/2^{b-1}$, so the security level depends only on the length of the authentication tag. In constructions 2, 3, 4, 7, 8 and 9, the security level depends on the length of the authentication tag and on the length of the plaintext. In these situations, one would start with a given plaintext length a and a given security level, say ϵ, and then determine the minimum b such that $Pd_1 \leq \epsilon$.
- Constructions 7–9 were designed with the goal of efficient software implementation. Constructions 7 and 8 achieve a short key length, but construction 9 is not competitive with the other constructions in terms of deception probabilities and key length.
- Bierbrauer [1] gives some constructions using geometric codes that achieve extremely short key lengths. However, there are some paremetric restrictions on when they can be applied, and they would probably be more difficult to implement than the other constructions mentioned above.

In Table 1, we tabulate b and ℓ, for $a = 2^8, 2^{16}, 2^{32}, 2^{64}$ and 2^{128} and $\epsilon = 2^{-20}$, obtained using the different constructions. In Table 2, we list b and ℓ for the same values of a when $\epsilon = 2^{-40}$. (We have computed b and ℓ for various combinations of a and ϵ, and the these tables are typical of the results obtained.)

From Tables 1 and 2, we see that the construction from [3] best combines a small key length with a short authenticator.

4 Counter-based Multiple Authentication

We will be generalizing the theory of universal hashing so that it can be applied to authentication of a sequence of w messages using one key. First, however, we review the approach used by Wegman and Carter in [16], which is a method to authenticate multiple messages using any ϵ-ASU class of hash functions. To apply this technique, the ith message in the sequence must be labeled with a counter having the value i, $1 \leq i \leq w$.

Let $\mathcal{F}$ be an ϵ-$ASU(N; m, n)$ hash family, where each function in $\mathcal{F}$ has domain A and range B, and suppose we want to authenticate a sequence of at most w source states. We will also assume that B is an abelian group. A key e is specified by a function $f \in \mathcal{F}$, together with a $(w-1)$-tuple $(b_1, \ldots, b_{w-1}) \in B^{w-1}$. (This $(w-1)$-tuple will act like a sequence of $w-1$ one-time pads.)

Construction	a	2^8	2^{16}	2^{32}	2^{64}	2^{128}
	ϵ	2^{-20}	2^{-20}	2^{-20}	2^{-20}	2^{-20}
1	b	21	21	21	21	21
	ℓ	768	1600	3328	6912	14336
2	b	23	24	25	26	27
	ℓ	138	336	750	1612	3402
4	b	24	32	47	78	141
	ℓ	48	64	94	156	282
5	b	21	21	21	21	21
	ℓ	71	85	117	179	305
6	b	21	21	21	21	21
	ℓ	141	346	775	1668	3521
7	b	30	38	54	86	150
	ℓ	56	71	103	166	293
8	b	29	37	53	85	149
	ℓ	83	106	154	249	440
9	b	1248	1312	29792	48393888	1.28×10^{14}
	ℓ	1375	34229	4×10^9	3.5×10^{19}	1.34×10^{39}

Table 1. Parameters for authentication codes when $\epsilon = 2^{-20}$

Let s_i denote the ith source state in the sequence. The authenticator for (i, s_i) is defined to be

$$e(i, s_i) = \begin{cases} f(s_i) & \text{if } i = 1 \\ f(s_i) + b_{i-1} & \text{if } 2 \leq i \leq w. \end{cases}$$

Note that the authentication function depends in an essential way upon the position of each source state within the sequence of w source states. We also remark that this is essentially the method suggested by Wegman and Carter in [16], except that we have omitted a one-time pad for the first source state since it is not necessary. (This approach has also been used by other researchers, e.g., [10].)

The following theorem can be proved in a manner similar to [16]. The proof is omitted from this Extended Abstract.

Theorem 3. *Suppose there exists an ϵ-$ASU(N; m, n)$ hash family, and let $w \geq 1$. Then there exists an authentication code without secrecy for m source states, which can be used to authenticate a sequence of up to w source states, having n authenticators and Nn^{w-1} authentication rules, such that $Pd_0 = 1/n$ and $Pd_i \leq \epsilon$, $1 \leq i \leq w$.*

Observe that this counter-based scheme is much more efficient than simply

Construction	a	2^8	2^{16}	2^{32}	2^{64}	2^{128}
	ϵ	2^{-40}	2^{-40}	2^{-40}	2^{-40}	2^{-40}
1	b	41	41	41	41	41
	ℓ	1408	2880	5888	12032	24576
2	b	42	44	45	46	47
	ℓ	210	572	1305	2806	5875
4	b	43	51	66	98	161
	ℓ	86	102	132	196	322
5	b	41	41	41	41	41
	ℓ	129	145	175	239	365
6	b	41	41	41	41	41
	ℓ	217	581	1329	2861	5993
7	b	50	58	74	106	170
	ℓ	95	111	142	206	333
8	b	49	57	73	105	169
	ℓ	142	166	213	309	500
9	b	12576	12576	29856	48393888	1.28×10^{14}
	ℓ	12783	51075	4×10^9	3.5×10^{19}	1.34×10^{39}

Table 2. Parameters for authentication codes when $\epsilon = 2^{-40}$

using w independent keys, since we need only add $\log n$ new key bits for each extra message to be authenticated

Although this counter-based scheme provides a nice method for multiple authentication, it has some drawbacks. For example, if a message is lost in transmission, then subsequent (valid) messages will not authenticate properly. (This would also be the case if w independent keys were used.) Hence, we believe there is some interest in achieving multiple authentication without counters. We pursue this theme in the remainder of the paper.

5 Multiple Authentication without Counters

In this section, we give some new definitions of hash families that we will use for multiple authentication.

Definition:

- An $(N; m, n)$-hash family $\mathcal{F}$ of functions from A to B is *ϵ-universal-w* (or ϵ-$U(N; m, n, w)$) provided that, for all distinct elements $x_1, x_2, \ldots, x_w \in A$, we have

$$|\{f \in \mathcal{F} : f(x_i) \neq f(x_j), 1 \leq i < j \leq w\}| \geq (1 - \epsilon)N.$$

- An $(N; m, n)$-hash family $\mathcal{F}$ of functions from A to B is *ϵ-almost-strongly-universal-w* (or ϵ-$ASU(N; m, n, w)$) provided that, for all distinct elements $x_1, x_2, \ldots, x_w \in A$, and for all (not necessarily distinct) $y_1, y_2, \ldots, y_w \in B$, we have

$$|\{f \in \mathcal{F} : f(x_i) = y_i, 1 \leq i \leq w\}| \leq \epsilon \times |\{f \in \mathcal{F} : f(x_i) = y_i, 1 \leq i \leq w-1\}|.$$

- (see [17]) An $(N; m, n)$-hash family $\mathcal{F}$ of functions from A to B is *strongly-universal-w* (or $SU(N; m, n, w)$) provided that, for all distinct $x_1, x_2, \ldots, x_w \in A$, and for all (not necessarily distinct) elements $y_1, y_2, \ldots, y_w \in B$, we have

$$|\{f \in \mathcal{F} : f(x_i) = y_i, 1 \leq i \leq w\}| = \frac{N}{n^w}.$$

We observe that the definition of ϵ-$U(N; m, n, 2)$ given above is the same as the definition of ϵ-$U(N; m, n)$ that we gave in Section 3. Similarly, the definition of ϵ-$SU(N; m, n, 2)$ given above is the same as the definition of ϵ-$SU(N; m, n)$ from Section 3. As well, a hash family that is both ϵ-$ASU(N; m, n, 2)$ and $(1/n)$-$ASU(N; m, n, 1)$ (as defined above) is ϵ-$ASU(N; m, n)$ (as defined in Section 3).

The following lemma describes the relation between ASU and SU families.

Lemma 4. *Let w be a positive integer. An $(N; m, n)$-hash family is $SU(N; m, n, w)$ if and only if it is $\frac{1}{n}$-$ASU(N; m, n, j)$ for $1 \leq j \leq w$.*

Proof. Suppose $\mathcal{F}$ is an $SU(N; m, n, w)$. Pick any j, where $1 \leq j \leq w$. Let $x_1, x_2, \ldots, x_j$ be distinct elements of A and let $y_1, y_2, \ldots, y_j$ be not necessarily distinct elements of B. Then we have

$$\frac{|\{f : f(x_i) = y_i, 1 \leq i \leq j\}|}{|\{f : f(x_i) = y_i, 1 \leq i \leq j-1\}|} = \frac{N/n^j}{N/n^{j-1}} = \frac{1}{n}.$$

Hence $\mathcal{F}$ is a $\frac{1}{n}$-$ASU(N; m, n, j)$ hash family, for $j = 1, 2, \ldots, w$.

Conversely, suppose $\mathcal{F}$ is an $\frac{1}{n}$-$ASU(N; m, n, j)$ for $j = 1, 2, \ldots, w$. Let $x_1, x_2, \ldots, x_w$ be distinct elements of A and let $y_1, y_2, \ldots, y_w$ be not necessarily distinct elements of B. Then we have

$$\begin{aligned}
|\{f : f(x_i) = y_i, 1 \leq i \leq w\}| &\leq \frac{1}{n}|\{f : f(x_i) = y_i, 1 \leq i \leq w-1\}| \\
&\leq \frac{1}{n^2}|\{f : f(x_i) = y_i, 1 \leq i \leq w-2\}| \\
&\vdots \\
&\leq \frac{1}{n^{w-1}}|\{f : f(x_i) = y_i\}| \\
&\leq \frac{N}{n^w}.
\end{aligned}$$

Since this is true for all $y_1, y_2, \ldots, y_w \in B$, we have

$$\sum_{\{y_1, y_2, \ldots, y_w \in B\}} |\{f : f(x_i) - y_i\}| \leq n^w \frac{N}{n^w} = N,$$

and, since each hash function is used at least once, we have

$$\sum_{\{y_1,y_2,\ldots,y_w \in B\}} |\{f : f(x_i) = y_i\}| \geq N.$$

Hence

$$|\{f : f(x_i) = y_i, 1 \leq i \leq w\}| = \frac{N}{n^w}.$$

□

We also have the following lemma which shows that ϵ-U hash families are also ϵ'-U-w families for some $\epsilon' > \epsilon$.

Lemma 5. *Suppose $\mathcal{F}$ is an ϵ-$U(N;m,n)$ hash family. Then $\mathcal{F}$ is an $\epsilon\binom{w}{2}$-$U(N;m,n,w)$ hash family for any integer w such that $\epsilon\binom{w}{2} \leq 1$.*

Proof. Since $\mathcal{F}$ is an ϵ-$U(N;m,n)$ family, for any two distinct elements of A, say x_1, x_2, we have

$$|\{f \in \mathcal{F} : f \text{ is not 1-1 on } x_1, x_2\}| \leq \epsilon N.$$

Therefore for any w distinct element of A, say $x_1, x_2, \ldots, x_w$, we have

$$\begin{aligned} |\{f \in \mathcal{F} : f \text{ is not 1-1 on } x_1, x_2, \ldots, x_w\}| &\leq \sum_{1 \leq i < j \leq w} |\{f \in \mathcal{F} : f(x_i) = f(x_j)\}| \\ &\leq \binom{w}{2} \epsilon N. \end{aligned}$$

Hence, we have

$$|\{f : f \text{ is 1-1 on } x_1, x_2, \ldots, x_w\}| \geq (1 - \binom{w}{2}\epsilon)N.$$

□

ϵ-$ASU(N;m,n,w)$ hash families can be used for authentication of a sequence of $w-1$ distinct source states, without the need for counters. The following result is immediate.

Theorem 6. *If there exists an ϵ_w-$ASU(N;m,n,w)$ hash family, then there exists an authentication code without secrecy for m source states, having n authenticators and N authentication rules, such that $Pd_{w-1} \leq \epsilon_w$.*

6 Composition Constructions

In this section, we present the composition constructions that we will use to achieve multiple authentication without counters. First, we present a method which generalizes a construction from Stinson [13] of combining hash families.

Theorem 7. *Suppose $\mathcal{F}_1$ is an $\epsilon_1(j)$-$U(N_1; m_1, n_1, j)$ hash family from A_1 to B_1, and suppose $\mathcal{F}_2$ is an $\epsilon_2(j)$-$ASU(N_2; n_1, n_2, j)$ hash family from B_1 to B_2, for all j, $1 \leq j \leq w$. Then there exists an $\epsilon(j)$-$ASU(N; m_1, n_2, j)$ hash family $\mathcal{F}$ of hash functions from A_1 to B_2, where*

$$\epsilon(j) \leq \frac{\epsilon_1(j)[1 - \epsilon_2(2)\epsilon_2(3)\ldots\epsilon_2(j)] + \epsilon_2(2)\ldots\epsilon_2(j)}{(1 - \epsilon_1(j-1))\epsilon_2(2)\ldots\epsilon_2(j-1)} \qquad j = 2, 3, \ldots, w,$$

$$\epsilon(1) \leq \epsilon_2(1), \qquad \textit{and}$$

$$N = N_1 N_2.$$

Proof. Let $1 \leq j \leq w$. We need an upper bound on

$$|\{f : f(x_i) = y_i, 1 \leq i \leq j\}|$$

and a lower bound on

$$|\{f : f(x_i) = y_i, 1 \leq i \leq j-1\}|.$$

We proceed as follows:

Upper bound

Let $x_1, x_2, \ldots, x_j \in A_1$ (all distinct) and $y_1, y_2, \ldots, y_j \in B_2$. Let p denote the probability that for some i, k, $(1 \leq i < k \leq j)$, x_i, x_k collide under a hash function from $\mathcal{F}_1$. If $f_1 \in \mathcal{F}_1$ and f_1 is one-to-one on $x_1, x_2, \ldots, x_j$, the number of hash functions $f \in \mathcal{F}$ such that $f(x_i) = y_i$ for $i = 1, 2, \ldots, j$ is

$$(1-p)N_1 N_2 \epsilon_2(1)\epsilon_2(2)\ldots\epsilon_2(j).$$

If $f_1 \in \mathcal{F}_1$ and f_1 is not one-to-one on $x_1, x_2, \ldots, x_j$, then the number of hash functions $f \in \mathcal{F}$ such that $f(x_i) = y_i$ for $i = 1, 2, \ldots, j$ is at most

$$pN_1 N_2 \epsilon_2(1).$$

Therefore, the number of hash functions $f \in \mathcal{F}$ such that $f(x_i) = y_i$ for $i = 1, 2, \ldots, j$ is at most

$$N_1 N_2[p\epsilon_2(1) + (1-p)\epsilon_2(1)\epsilon_2(2)\ldots\epsilon_2(j)].$$

Hence, we have

$$|\{f : f(x_i) = y_i, 1 \leq i \leq j\}| \leq N_1 N_2 \epsilon_2(1)[p + (1-p)\epsilon_2(2)\epsilon_2(3)\ldots\epsilon_2(j)].$$

Lower bound

Let $x_1, x_2, \ldots, x_{j-1} \in A_1$ (all distinct) and let $y_1, y_2, \ldots, y_{j-1} \in B_2$. Let p' denote the probability that for some i, k, $(1 \leq i < k \leq j-1)$, x_i, x_k collide under a hash function from $\mathcal{F}_1$. Since we only need a lower bound, we will look the case where $f_1 \in \mathcal{F}_1$ is one-to-one on $x_1, x_2, \ldots, x_{j-1}$. Hence we have

$$|\{f : f(x_i) = y_i, 1 \leq i \leq j-1\}| \geq N_1 N_2 (1-p')\epsilon_2(1)\epsilon_2(2)\ldots\epsilon_2(j-1).$$

We now combine the upper and lower bounds. We obtain the following:

$$\begin{aligned}\frac{|\{f : f(x_i) = y_i, 1 \leq i \leq j\}|}{|\{f : f(x_i) = y_i, 1 \leq i \leq j-1\}|} &\leq \frac{\epsilon_2(1)[p + (1-p)\epsilon_2(2)\epsilon_2(3)\ldots\epsilon_2(j)]}{(1-p')\epsilon_2(1)\epsilon_2(2)\ldots\epsilon_2(j-1)} \\ &\leq \frac{\epsilon_1(j)[1-\epsilon_2(2)\epsilon_2(3)\ldots\epsilon_2(j)] + \epsilon_2(2)\ldots\epsilon_2(j)}{[1-\epsilon_1(j-1)]\epsilon_2(2)\ldots\epsilon_2(j-1)},\end{aligned}$$

since $p \leq \epsilon_1(j)$ and $p' \leq \epsilon_1(j-1)$. □

Corollary 8. *Suppose $\mathcal{F}_1$ is an $\epsilon_1(j)$-$U(N_1; m_1, n_1, j)$ hash family from A_1 to B_1, and suppose $\mathcal{F}_2$ is an $SU(N_2, n_1, n_2, w)$ hash family from B_1 to B_2, $1 \leq j \leq w$. Then there exists an $\epsilon(j)$-$ASU(N; m_1, n_2, j)$ hash family $\mathcal{F}$ from A_1 to B_2, where $N = N_1 N_2$ and*

$$\epsilon(j) = \frac{\epsilon_1(j) n_2^{j-1} - \epsilon_1(j) + 1}{(1-\epsilon_1(j-1))n_2},$$

for $j = 1, 2, \ldots, w$.

Proof. Apply Lemma 4 and Theorem 7. Note that $\epsilon(1) = \frac{1}{n_2}$ by this formula. □

7 Multiple Authentication without Counters

We now use the tools of the previous section to obtain our multiple authentication codes. We could generalize many of the constructions that were mentioned in Section 3.1. The method we have chosen to use is inspired by the construction from [3] (i.e., construction 5 in Section 3.1). We need two ingredients to accomplish this. First, Bierbrauer gave a construction for orthogonal arrays that gives us SU-w hash families.

Lemma 9. *[2] Let q be a prime power and let S, T be integers such that $S \geq T$. Then there exists an $SU(q^{(w-1)S+T}; q^S, q^T, w)$ hash family, where $w \leq q^S$.*

The second ingredient is the U hash families that are obtained from Reed-Solomon codes [3].

Lemma 10. *[3] Let Q be a prime power, and let $k \leq Q$. Then there is a $\frac{k-1}{Q}$-$U(Q; Q^k, Q)$ hash family.*

Applying Lemma 5, the following is obtained.

Lemma 11. *Let Q be a prime power, let $k \leq Q$, and suppose $\binom{j}{2}\frac{k-1}{Q} \leq 1$. Then there is a $\binom{j}{2}\frac{k-1}{Q}$-$U(Q; Q^k, Q, j)$ hash family.*

Now, let a, b and w be given, as usual. Let s be an integer such that

$$a \leq ((w-1)b + s)(2^s + 1).$$

Then take

$$Q = 2^{(w-1)b+s}$$

and

$$k = 2^s + 1$$

in Lemma 11, and restrict the resulting hash functions to a domain of size 2^a. In this way, we obtain a

$$\binom{j}{2}2^{-(w-1)b}\text{-}U(2^{(w-1)b+s}; 2^a, 2^{(w-1)b+s}, j)$$

hash family, for all j such that $1 \leq j \leq w$.

Next, use Corollary 8 to compose this family with an

$$SU(2^{(w-1)((w-1)b+s)+b}; 2^{(w-1)b+s}, 2^b, w)$$

hash family obtained from Lemma 9 with $S = (w-1)b+s$ and $T = b$. The result is an

$$\epsilon(j)\text{-}ASU(2^{(w^2-w+1)b+ws}; 2^a, 2^b, j)$$

hash family ($1 \leq j \leq w$), where

$$\epsilon(j) \leq \frac{j(j-1)2^{-b(w-1)}(2^{b(j-1)} - 1) + 2}{[2 - (j-1)(j-2)2^{-b(w-1)}]2^b},$$

$1 \leq j \leq w$.

Phrasing our construction in terms of authentication codes, we obtain the following result.

Theorem 12. *Let a, b, and w be integers, and let s be an integer such that $((w-1)b+s)(1+2^s) \geq a$. Then there exists an authentication code for an a-bit source, having a b-bit authenticator and requiring $\ell = (w^2 - w + 1)b + ws$ bits of key, in which*

$$Pd_{j-1} \leq \frac{j(j-1)2^{-b(w-1)}(2^{b(j-1)} - 1) + 2}{[2 - (j-1)(j-2)2^{-b(w-1)}]2^b},$$

for $j = 1, 2, \ldots, w$.

We remark that in the case $w = 2$, Theorem 12 is identical to construction 5 in Section 3.1, due to [3].

In Theorem 12 the security level depends on the length of the authentication tag, on the length of the plaintext and number of messages that are being sent. Hence, one would start with a given plaintext length a and a given security level, say ϵ, and then determine the minimum b such that $Pd_{w-1} \leq \epsilon$. Once b is determined, we can proceed to compute s, and then apply Theorem 12.

In Tables 3, 4 and 5, we tabulate the length of authentication tag and the length of the key for given a, w, and ϵ values of the authentication codes that are constructed in this way from Theorem 12.

a	2^8	2^{16}	2^{32}	2^{64}	2^{128}	2^8	2^{16}	2^{32}	2^{64}	2^{128}
ϵ	2^{-20}	2^{-20}	2^{-20}	2^{-20}	2^{-20}	2^{-40}	2^{-40}	2^{-40}	2^{-40}	2^{-40}
b	22	22	22	22	22	42	42	42	42	42
ℓ	163	187	232	328	517	300	324	372	46	657

Table 3. Authentication codes for $w = 3$

a	2^8	2^{16}	2^{32}	2^{64}	2^{128}	2^8	2^{16}	2^{32}	2^{64}	2^{128}
ϵ	2^{-20}	2^{-20}	2^{-20}	2^{-20}	2^{-20}	2^{-40}	2^{-40}	2^{-40}	2^{-40}	2^{-40}
b	23	23	23	23	23	43	43	43	43	43
ℓ	307	339	403	531	783	563	595	659	787	1043

Table 4. Authentication codes for $w = 4$

a	2^8	2^{16}	2^{32}	2^{64}	2^{128}	2^8	2^{16}	2^{32}	2^{64}	2^{128}
ϵ	2^{-20}	2^{-20}	2^{-20}	2^{-20}	2^{-20}	2^{-40}	2^{-40}	2^{-40}	2^{-40}	2^{-40}
b	26	26	26	26	26	46	46	46	46	46
ℓ	2376	2456	2606	2926	3566	4186	4266	44426	4746	5376

Table 5. Authentication codes for $w = 10$

8 Summary

We have generalized the theory of universal hashing to construct authentication codes that allow the authentication of a sequence of (distinct) source states without the use of counters. It can be seen that the construction we have given (Theorem 12) requires considerably more key bits than the counter-based method described in Section 4. More efficient constructions (without counters) would therefore be of considerable interest.

Acknowledgements

The authors' research is supported by NSF Grant CCR-9402141 and by the Center for Communication and Information Science at the University of Nebraska.

References

1. J. Bierbrauer, Universal hashing and geometric codes, to appear in *Designs, Codes and Cryptography.*
2. J. Bierbrauer, Construction of orthogonal arrays, to appear in *Journal of Statistical Planning and Inference.*
3. J. Bierbrauer, T. Johansson, G. Kabatianskii and B. Smeets, On families of hash functions via geometric codes and concatenation, in "Advances in Cryptology – CRYPTO '93", D. R. Stinson, ed., *Lecture Notes in Computer Science* **773** (1994), 331-342.
4. B. den Boer, A simple and key-economical unconditional authentication scheme, *Journal of Computer Security* **2** (1993), 65-71.
5. E. N. Gilbert, F. J. MacWilliams and N. J. A. Sloane, Codes which detect deception, *Bell System Technical Journal* **53** (1974), 405-424.
6. T. Helleseth and T. Johansson, Universal hash functions from exponential sums over finite fields and Galois rings, in "Advances in Cryptology – CRYPTO '96", N. Koblitz, ed, *Lecture Notes in Computer Science* (1996).
7. H. Krawczyk, LFSR-based hashing and authentication, in "Advances in Cryptology – CRYPTO '94", Y. G. Desmedt, ed., *Lecture Notes in Computer Science* **839** (1994), 129-139.
8. H. Krawczyk, New hash functions for message authentication, in "Advances in Cryptology – EUROCRYPT '95", L. C. Guillou and J.-J. Quisquater, eds., *Lecture Notes in Computer Science* **921** (1995), 301-310.
9. J. L. Massey, Cryptography – a selective survey, in "Digital Communications", E. Biglieri and G. Prati, eds., North-Holland, 1986, 3-21. [Also published in *Alta Frequenza* **55** (1986), 4-11.]
10. P. Rogaway, Bucket hashing and its application to fast message authentication, in "Advances in Cryptology CRYPTO '95", D. Coppersmith, ed., *Lecture Notes in Computer Science* **963** (1995), 29-42.
11. G. J. Simmons, A survey of information authentication, in "Contemporary Cryptology, The Science of Information Integrity", G. J. Simmons, ed., IEEE Press, 1992, 379-419. [Preliminary version appeared in *Proceedings of the IEEE* **76** (1988), 603-620.]

12. D. R. Stinson, Universal hashing and authentication codes, in "Advances in Cryptology CRYPTO '91", J. Feigenbaum, ed., *Lecture Notes in Computer Science* **576** (1992), 74-85.
13. D. R. Stinson, Universal hashing and authentication codes, *Designs, Codes and Cryptography* **4** (1994), 369-380.
14. D. R. Stinson, On the connections between universal hashing, combinatorial designs and error-correcting codes, to appear in *Congressus Numerantium* **115** (1996). [Also appears in *Electronic Colloquium on Computational Complexity*, Report TR95-052.]
15. R. Taylor, Nearly optimal unconditionally secure authentication, in "Advances in Cryptology – EUROCRYPT '94", A. De Santis, ed., *Lecture Notes in Computer Science* **950** (1995), 244-253.
16. M. N. Wegman and J. L. Carter, New hash functions and their use in authentication and set equality, *Journal of Computer and System Sciences* **22** (1981), 265-279.
17. Y. Zheng, T. Hardjono and J. Pieprzyk, Sibling intractable function families and their applications, in "Advances in Cryptology – ASIACRYPT '91", H. Imai, R. L. Rivest and T. Matsumoto, eds., *Lecture Notes in Computer Science* **739** (1993), 124-138.