© de Gruyter 2010
J. Math. Crypt. **4** (2010), 1–24 DOI 10.1515 / JMC.2010.xxx

The power of primes: security of authentication based on a universal hash-function family

Basel Alomair, Andrew Clark and Radha Poovendran

Communicated by xxx

Abstract. Message authentication codes (MACs) based on universal hash-function families are becoming increasingly popular due to their fast implementation. In this paper, we investigate a family of universal hash functions that has been appeared repeatedly in the literature and provide a detailed algebraic analysis for the security of authentication codes based on this universal hash family. In particular, the universal hash family under analysis, as appeared in the literature, uses operation in the finite field $\mathbb{Z}_p$. No previous work has studied the extension of such universal hash family when computations are performed modulo a non-prime integer n. In this work, we provide the first such analysis. We investigate the security of authentication when computations are performed over arbitrary finite integer rings $\mathbb{Z}_n$ and derive an explicit relation between the prime factorization of n and the bound on the probability of successful forgery. More specifically, we show that the probability of successful forgery against authentication codes based on such a universal hash-function family is bounded by the reciprocal of the smallest prime factor of the modulus n.

Keywords. Cryptography, authentication, finite integer rings, universal hash-function families.

AMS classification. 11T71, 94A60, 94A62.

1 Introduction and related work

Message authentication code (MAC) algorithms can be categorized, based on their security, into unconditionally and conditionally secure MACs. While the security of the former category of MACs is unconditional, the latter is only secure against computationally bounded adversaries. The first unconditionally secure MAC was introduced by Gilbert *et al.* in [18]. The first deployment of universal hash-function families for the design of authentication codes was introduced by Wegman and Carter for the purpose of designing unconditionally secure authentication [12, 49, 13, 50]. Since then, the study of unconditionally secure message authentication based on universal hash-function families has been attracting research attention, both from the design and analysis viewpoints (see, e.g., [8, 3, 21, 37, 9, 2]).

The use of universal hash-function families is not confined to the design of unconditionally secure MACs. Cryptographists have realized that universal hash functions can be used to construct efficient, computationally secure, MACs. Traditional computationally secure MACs are usually block cipher based (see, e.g., [46, 4, 22, 16, 34, 28]). Compared to block cipher based MACs, however, universal hash-function families based MACs usually offer better performances (to date, the fastest MACs are based on universal hash-function families [47]). The basic idea behind universal hash-function families based MACs is to compress the message to be authenticated (using a univer-

sal hash function) and then encrypt the compressed image (e.g., using one-time pad ciphers, stream ciphers, or pseudorandom functions). Universal hash-function families based MACs include, but are not limited to, [10, 6, 7, 31, 19, 17, 24].

An important branch of the area of message authentication codes is the study of their security. In particular, substantial efforts have been devoted to bounding the probabilities of deception (forgeability) of authentication codes. The significance of such analysis is that it provides a metric for measuring and comparing the reliability of different MAC algorithms. Valuable contributions that investigate the security of different authentication codes include, but are not limited to, [14, 39, 23, 27, 29, 30, 41, 33, 32, 5, 35].

The security of many universal hash-function families rely on the fact that computations are performed over finite fields (see, e.g., [18, 50, 38, 15, 21, 26, 19, 17, 2]). In this work, we investigate a universal hash-function family that belongs to this class of universal hash families. Unlike previous analysis, however, we will consider the effect of performing operations over finite integer rings, as opposed to fields, on the security of authentication codes based on this universal hash family.

To give an example of the universal hash family under study, let a message m be divided into equal-length blocks $m_i \in \mathbb{Z}_p$, where p is a pre-specified prime integer. Given the secret hashing keys $k_i \in \mathbb{Z}_p$, compute the hashed image of the message m as $h(m) = \sum_i k_i m_i \pmod p$. Then, the authentication tag of m is simply an encryption of its hashed image. There have been multiple proposals in the literature of message authentication that were based on variants of this approach (see, e.g., [38, 19, 17, 2]). When the multiplication is performed modulo a prime integer, it has been proven that such proposals provide message integrity. However, the effect of using non-prime moduli on the security of such proposals has not been previously investigated.

CONTRIBUTIONS. In this paper, we investigate the use of a class of universal hash-function families that has been used for message authentication. In particular, we will analyze the security of authentication based on this class of universal hash-function families when the operations are performed over arbitrary finite integer rings instead of fields, where they have been shown to be secure. We derive tight bounds on the probabilities of deception for all choices of finite integer rings $\mathbb{Z}_n$. We show the direct relation between the prime factorization of the modulus n and the security of authentication. More precisely, we prove that the probability of deception is bounded by the reciprocal of the smallest prime factor of the modulus n.

Since the derivation of the main result is quite lengthy, we attempt to clarify it by breaking the proof into a series of lemmas (Lemma 5.5 - Lemma 5.10) leading to the final theorem. One particular result that is generally interesting (not only for this paper) is the result of Lemma 3.1. In Lemma 3.1 we prove what can be viewed as an extension to Bézout's lemma for finite integer rings. It is a well-known fact in algebra and number theory that, if $\gcd(a, n) = d$ then, there exists an integer x such that $x \cdot a \equiv d \pmod n$. What we show in Lemma 3.1 is that for an $a \in \mathbb{Z}_n \backslash \{0\}$ such that $\gcd(a, n) = d$, not only there exists an element $x \in \mathbb{Z}_n$ such that $x \cdot a \equiv d \pmod n$ but, further, there exists an *invertible element* $x \in \mathbb{Z}_n^*$ such that $x \cdot a \equiv d \pmod n$. This result is essential to generalize our bounds to any finite integer ring and, to the best of

our knowledge, has not appeared in the literature of mathematics.

ORGANIZATION. The rest of the paper is organized as follows. Section 2 provides a list of used notations and relevant definitions. In Section 3, we formally state and prove our extension to Bézout's lemma along with some basic properties of the finite integer ring $\mathbb{Z}_n$. Section 4 gives two examples of the use of the studied universal hash-function family for the construction of computationally secure MACs and the construction of codes with secrecy. Section 5 is devoted to the security analysis. Section 6 provides a summary of the choices of moduli and their security ramifications. In Section 7 we conclude our paper.

2 Notations and definitions

In this section we list the notations and definitions that are relevant to the presentation of the paper.

2.1 Notations

The following notations will be used throughout the rest of the paper.

- For the ring $\mathbb{Z}_n := \{0, 1, ..., n-1\}$ with the usual addition and multiplication modulo n, the subset $\mathbb{Z}_n^*$ is defined to be the set of integers in $\mathbb{Z}_n$ that are relatively prime to n.
- If S is a set, then $|S|$ is defined to be the cardinality of the set. If r is an integer, then $|r|$ is defined to be the length of r in bits.
- The function $\varphi(n)$ (the *Euler totient function*) is defined to be the number of positive integers less than n that are relatively prime to n. Equivalently, $\varphi(n) = |\mathbb{Z}_n^*|$.
- For any two strings a and b, $(a \mid\mid b)$ denotes the concatenation operation.
- For two integers a and b, we say $a \mid b$, read as a divides b, if there exists an integer c such that $b = c \times a$.
- For two integers a and b, we say $a \nmid b$, read as a does not divide b, if there is no integer c such that $b = c \times a$.
- For the rest of the paper, $(+)$ and $(\times)$ represent addition and multiplication over $\mathbb{Z}_n$, even if the $\pmod n$ part is dropped for simplicity.
- For any two integers a and b, $\gcd(a, b)$ is the greatest common divisor of a and b.
- For an element a in a ring R, the element a^{-1} denotes the multiplicative inverse of a in R, if it exists.
- Throughout the rest of the paper, random variables will be represented by bold font symbols, whereas the corresponding non-bold font symbols represent specific values that can be taken by these random variables.

2.2 Definitions

One definition that will be used in the paper is the notion of perfect secrecy in Shannon's information-theoretic sense. An encryption algorithm is said to be information-theoretically secure if the ciphertext gives no information about the plaintext, i.e., the ciphertext and the plaintext are statistically independent. Formally, perfect secrecy can be defined as:

Definition 2.1. [44][Perfect secrecy] For a plaintext m and its corresponding ciphertext ψ, the cipher is said to achieve perfect secrecy if

$$\Pr(\boldsymbol{m} = m|\boldsymbol{\psi} = \psi) = \Pr(\boldsymbol{m} = m)$$

for all plaintext m and all ciphertext ψ. That is, the a posteriori probability that the plaintext is m, given that the ciphertext ψ is observed, is identical to the a priori probability that the plaintext is m.

Another definition that is relevant to this work is the definition of universal hash-function families. A family of hash functions $\mathcal{H}$ is specified by a finite set of keys $\mathcal{K}$. Each key $k \in \mathcal{K}$ defines a member of the family $\mathcal{H}_k \in \mathcal{H}$. As opposed to thinking of $\mathcal{H}$ as a set of functions from A to B, it can be viewed as a single function $\mathcal{H} : \mathcal{K} \times A \to B$, whose first argument is usually written as a subscript. A random element $h \in \mathcal{H}$ is determined by selecting a $k \in \mathcal{K}$ uniformly at random and setting $h = \mathcal{H}_k$. Different notions of universal hash families have appeared in the literature (see, e.g., [49, 12, 43, 25, 26, 19]), we give below one such definition.

Definition 2.2. [43, 10] [Universal hash families] Let $\mathcal{H} = \{h : A \to B\}$ be a family of hash functions and let $\epsilon \geq 0$ be a real number. We say that $\mathcal{H}$ is ϵ-almost universal, denoted ϵ-AU, if for all distinct $M, M' \in A$, we have that $\Pr_{h \leftarrow \mathcal{H}}[h(M) = h(M')] \leq \epsilon$.

3 Preliminaries

For any nonzero integers a and n with $\gcd(a, n) = d$, by Bézout's lemma [45], there exist two integers x and y so that $ax + ny = d$. Otherwise stated, for any nonzero integers a and n with $\gcd(a, n) = d$, by Bézout's lemma, there exists an integer x so that

$$ax \equiv d \pmod{n}. \tag{3.1}$$

It is further known that the x satisfying equation (3.1) is not necessarily unique. In particular, for a nonzero $a \in \mathbb{Z}_n$, there are $d = \gcd(a, n)$ distinct elements in $\mathbb{Z}_n$ satisfying equation (3.1), given by

$$\{x_0,\ x_0 + \frac{n}{d},\ x_0 + 2\frac{n}{d},\ \cdots,\ x_0 + (d-1)\frac{n}{d}\}, \tag{3.2}$$

where x_0 is the smallest integer in $\mathbb{Z}_n$ satisfying equation (3.1) [45]. The significance of the following lemma is the statement that at least one of the d elements of the set in equation (3.2) *must be* invertible in $\mathbb{Z}_n$.

Lemma 3.1. *In any finite integer ring $\mathbb{Z}_n$, for any $\delta \in \mathbb{Z}_n \backslash \{0\}$, if* $\gcd(\delta, n) = d$*, then there exists an invertible element $\alpha \in \mathbb{Z}_n^*$ such that $\alpha \times \delta \equiv d \pmod{n}$.*

Proof. Let $\gcd(\delta, n) = d$, then by Bézout's lemma [45], there exists an integer α_0 such that

$$\alpha_0 \times \delta \equiv d \pmod{n}. \tag{3.3}$$

Further, all integers in the infinite set

$$A = \{\alpha_k | \alpha_k = \alpha_0 + k\frac{n}{d}, \quad \forall\, k \in \mathbb{Z}\} \tag{3.4}$$

are valid solutions to equation (3.3) [45]. The lemma states that, not only there exists an integer that satisfies equation (3.3), but there exists an *invertible* element in $\mathbb{Z}_n$ that satisfies equation (3.3). We will prove the lemma by finding an integer k such that $\alpha_k \in A$ is relatively prime to n.

If $\gcd(\delta, n) = 1$ then $\alpha_0 = \delta^{-1} \in \mathbb{Z}_n^*$ does exist and is the invertible solution to equation (3.3). Assume, however, that $\gcd(\delta, n) = d > 1$ and write n in its prime factorization as

$$n = \prod_{i=1}^{\ell_1} p_i^{e_i} \prod_{i=1}^{\ell_2} \gamma_i^{e_{\gamma_i}} \prod_{i=1}^{\ell_3} \zeta_i^{e_{\zeta_i}}. \tag{3.5}$$

Assume further that δ can be written in its prime factorization form as

$$\delta = \prod_{i=1}^{\ell_1} p_i^{e'_i} \prod_{i=1}^{\ell_2} \gamma_i^{e'_{\gamma_i}} \prod_{i=1}^{\ell_4} r_i^{e_{r_i}}, \tag{3.6}$$

where $e'_i \geq e_i,\ \forall\, i = 1, \cdots, \ell_1$, and $e'_{\gamma_i} < e_{\gamma_i},\ \forall\, i = 1, \cdots, \ell_2$, with the ζ_i's and r_i's being distinct primes. Then, $d = \prod_{i=1}^{\ell_1} p_i^{e_i} \prod_{i=1}^{\ell_2} \gamma_i^{e'_{\gamma_i}}$ and, by Bézout's lemma, there exists an α_0 such that

$$\alpha_0 \times \delta \equiv d \pmod{n}. \tag{3.7}$$

Which is equivalent to

$$\alpha_0 \times \prod_{i=1}^{\ell_1} p_i^{e'_i - e_i} \prod_{i=1}^{\ell_4} r_i^{e_{r_i}} \equiv 1 \pmod{\prod_{i=1}^{\ell_2} \gamma_i^{e_{\gamma_i} - e'_{\gamma_i}} \prod_{i=1}^{\ell_3} \zeta_i^{e_{\zeta_i}}}. \tag{3.8}$$

Equation (3.8) implies that α_0 is relatively prime to $\prod_{i=1}^{\ell_2} \gamma_i^{e_{\gamma_i} - e'_{\gamma_i}} \prod_{i=1}^{\ell_3} \zeta_i^{e_{\zeta_i}}$, which implies that none of the γ_i's nor the ζ_i's divides α_0. Furthermore, by equation (3.4), none of the γ_i's nor the ζ_i's will divide α_k for any $k \in \mathbb{Z}$. Therefore, to prove that an $\alpha_k \in A$ is relatively prime to n, since the prime factorization of n consists only of p_i's, γ_i's, and ζ_i's, it suffices to show that none of the p_i's divides α_k.

Define $\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} := \prod_{i=1}^{\ell_2} \gamma_i^{e_{\gamma_i} - e'_{\gamma_i}} \prod_{i=1}^{\ell_3} \zeta_i^{e_{\zeta_i}}$, where $q_i = \gamma_i, e_{q_i} = e_{\gamma_i} - e'_{\gamma_i}$, for $i = 1, \cdots, \ell_2$ and $q_{\ell_2+i} = \zeta_i, e_{q_{\ell_2+i}} = e_{\zeta_i}$, for $i = 1, \cdots, \ell_3$. Then, equation (3.8) can be rewritten as

$$\alpha_0 \times \prod_{i=1}^{\ell_1} p_i^{e'_i - e_i} \prod_{i=1}^{\ell_4} r_i^{e_{r_i}} \equiv 1 \pmod{\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}}}, \tag{3.9}$$

where none of the q_i's divides any α_k for any $k \in \mathbb{Z}$.

Now, if none of the p_i's divides α_0 then $\gcd(\alpha_0, n) = 1$ and we are done. Assume, however, that some of the p_i's, for $i = 1, \cdots, \ell_1$ divide α_0, and let p_1 be one such prime dividing α_0. Then α_0 can be written as $\alpha_0 = m_1 p_1$, where m_1 is relatively prime to all q_i's (since, otherwise, some of the q_i's will divide α_0). Then, from equation (3.4), we know that

$$\alpha_1 = \alpha_0 + \frac{n}{d} = m_1 p_1 + \prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} \tag{3.10}$$

also satisfies equation (3.3). Therefore, $p_1 \nmid \alpha_1$ since it does not divide $\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}}$ (also none of the q_i's divides α_1 since none of them divides $m_1 p_1$).

Assume, however, that some of the other p_i's divide α_1, and let p_2 be such a prime. Then α_1 can be written as $\alpha_1 = m_2 p_2$ for some m_2 relatively prime to p_1 and all the q_i's. Then, by equation (3.4),

$$\alpha_2 \overset{\text{(b)}}{=} m_2 p_2 + \prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} \overset{\text{(a)}}{=} m_1 p_1 + 2 \prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} \tag{3.11}$$

also satisfies equation (3.3). Therefore, by equality (b), $p_2 \nmid \alpha_2$ since it does not divide $\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}}$ and, by equality (a), $p_1 \mid \alpha_2$ iff $p_1 = 2$. Assume that $p_1 = 2$ and write $\alpha_2 = m_3 p_1$ for an m_3 that is relatively prime to p_2 and the q_i's, then

$$\alpha_3 \overset{\text{(b)}}{=} m_3 p_1 + \prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} \overset{\text{(a)}}{=} m_2 p_2 + 2 \prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} = m_1 p_1 + 3 \prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}}. \tag{3.12}$$

Thus, since $p_2 \neq 2$, $p_1 \nmid \alpha_3$ and $p_2 \nmid \alpha_3$ by equalities (b) and (a) respectively, and $q_i \nmid \alpha_3 \ \forall i$ by construction.

Assume now that there exists an α_k such that $p_i \nmid \alpha_k \forall\, i = 1, \cdots, \ell_1 - 1$ and $q_i \nmid \alpha_k \forall i$, but $p_{\ell_1} \mid \alpha_k$. Then write $\alpha_k = m_k p_{\ell_1}$ for some m_k relatively prime to all q_i's and all p_i's except possibly p_{ℓ_1}. Then α_{k+1} can be expressed as

$$\alpha_{k+1} = m_k p_{\ell_1} + \prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} = \cdots \overset{\text{(b)}}{=} m_2 p_2 + c_2 \prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} \overset{\text{(a)}}{=} m_1 p_1 + c_1 \prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}}, \tag{3.13}$$

for some constants $c_i \in \mathbb{N}$. Recall that all the m_i's are relatively prime to the q_i's by construction. Therefore, to complete the proof, it suffices to show that there exists an integer $h \geq 1$ such that α_{k+h} is not divisible by any p_i. As a function of the p_i's and the c_i's, we conclude the proof by showing how to iteratively find such an h.

ITERATION 1. Assume that p_1 divides the α_{k+1} in equation (3.13). This implies, by equality (a), that $p_1 \mid c_1$. However, if $p_1 \mid c_1$ then $p_1 \nmid (c_1 + 1)$, and α_{k+2} can be written

as

$$\alpha_{k+2} = m_k p_\ell + 2\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} = \cdots \overset{\text{(c)}}{=} m_3 p_3 + (c_3+1)\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}}$$
$$\overset{\text{(b)}}{=} m_2 p_2 + (c_2+1)\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} \overset{\text{(a)}}{=} m_1 p_1 + (c_1+1)\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}}. \tag{3.14}$$

Therefore, by equality (a) in equation (3.14), we get $p_1 \nmid \alpha_{k+2}$.

ITERATION 2. Now, assume that $p_2 \mid \alpha_{k+2}$. By equality (b) in equation (3.14), this implies that $p_2 \mid (c_2+1)$. However, if $p_2 \mid (c_2+1)$ then $p_2 \nmid (c_2+1+p_1)$, and α_{k+2+p_1} can be written as

$$\alpha_{k+2+p_1} = m_k p_\ell + (2+p_1)\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} = \cdots \overset{\text{(c)}}{=} m_3 p_3 + (c_3+1+p_1)\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}}$$
$$\overset{\text{(b)}}{=} m_2 p_2 + (c_2+1+p_1)\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}} \overset{\text{(a)}}{=} m_1 p_1 + (c_1+1+p_1)\prod_{i=1}^{\ell_2+\ell_3} q_i^{e_{q_i}}. \tag{3.15}$$

Then, by equality (b) in equation (3.15), $p_2 \nmid \alpha_{k+2+p_1}$ and, by equality (a) in equation (3.15), $p_1 \nmid \alpha_{k+2+p_1}$.

ITERATION 3. Similarly, if p_3 divides α_{k+2+p_1} in equation (3.15), by equality (c), $p_3 \mid (c_3+1+p_1)$. However, if $p_3 \mid (c_3+1+p_1)$ then $p_3 \nmid (c_3+1+p_1+p_1p_2)$ and, by writing $\alpha_{k+2+p_1+p_1p_2}$ as

$$\begin{aligned}
\alpha_{k+2+p_1+p_1p_2} &= m_k p_\ell + (2+p_1+p_1p_2)\prod_i q_i^{e_{q_i}} \\
&= \cdots \\
&\overset{\text{(c)}}{=} m_3 p_3 + (c_3+1+p_1+p_1p_2)\prod_i q_i^{e_{q_i}} \\
&\overset{\text{(b)}}{=} m_2 p_2 + (c_2+1+p_1+p_1p_2)\prod_i q_i^{e_{q_i}} \\
&\overset{\text{(a)}}{=} m_1 p_1 + (c_1+1+p_1+p_1p_2)\prod_i q_i^{e_{q_i}},
\end{aligned} \tag{3.16}$$

one can see that neither p_3 nor p_2 nor p_1 divides $\alpha_{k+2+p_1+p_1p_2}$ by equalities (c), (b), and (a) in equation (3.16), respectively.

ITERATION ℓ_1. After the ℓ_1^{th} iteration, for an h given by:

$$h = 1 + \beta_1 + \beta_2 p_1 + \beta_3 p_1 p_2 + \cdots + \beta_{\ell_1} \prod_{i=1}^{\ell-1} p_i, \tag{3.17}$$

where $\beta_i = 1$ if in the i^{th} iteration $p_i \mid (m_i p_i + c_i \prod_i q_i^{e_{q_i}})$ and zero otherwise, α_{k+h} will not be divisible by any p_i. Hence, we have found, by construction, an α_{k+h} with $\gcd(\alpha_{k+h}, n) = 1$ that satisfies equation (3.3). The residue of α_{k+h} modulo n is an invertible element of $\mathbb{Z}_n$ that satisfies equations (3.3), and the lemma follows. □

For any finite integer ring $\mathbb{Z}_n$, $\mathbb{Z}_n \backslash \mathbb{Z}_n^*$, the complement of $\mathbb{Z}_n^*$, will be the set of elements that are not relatively prime to n. The following result holds for the set of integers that are not relatively prime to n.

Lemma 3.2. *In any finite integer ring $\mathbb{Z}_n$, for any $\alpha \in \mathbb{Z}_n \backslash \mathbb{Z}_n^*$ and any $\beta \in \mathbb{Z}_n$, $\alpha \times \beta \in \mathbb{Z}_n \backslash \mathbb{Z}_n^*$.*

The proof of this lemma can be found in [36].

Lemma 3.3. *Given an integer $k \in \mathbb{Z}_n^*$, for an r uniformly distributed over $\mathbb{Z}_n^*$, the value δ given by:*

$$\delta \equiv r \times k \pmod{n} \tag{3.18}$$

is uniformly distributed over $\mathbb{Z}_n^*$.

A more general result of Lemma 3.3 can be stated as follows.

Lemma 3.4. *Let G be a finite group and $\boldsymbol{X}$ a uniformly distributed random variable defined on G, and let $k \in G$. Let $\boldsymbol{Y} = k * \boldsymbol{X}$, where $*$ denotes the group operation. Then $\boldsymbol{Y}$ is uniformly distributed on G.*

Therefore, Lemma 3.3 follows directly from this general result in probability theory. The following is also a general result from number theory.

Lemma 3.5. *For any positive integer n with a prime factor p, $\varphi(n) \geq p - 1$, with equality iff $n = p$.*

This Lemma is a standard result for integers and its proof can be found in most books in number theory (see, e.g., [20]).

4 Examples of constructions

In this section, we give two examples of authentication codes based on the universal hash-function family under analysis. The first example is a construction of a computationally secure message authentication code (MAC) algorithm, while the second construction is an example of authentication codes with secrecy.

4.1 Constructing computationally secure MACs

In computationally secure MACs, the message to be authenticated is first compressed using a universal hash function and then the compressed image is processed with a cryptographic function (such as one-time pad ciphers, stream ciphers, or pseudorandom function).

Assume the message to be authenticated can be divided into b blocks, i.e., $m = (m_1, \cdots, m_b)$, where $m_i \in \mathbb{Z}_p^*$ for $i = 1, \cdots, b$. Let the key of the universal hash function be $k = (k_1, \cdots, k_b)$, where the k_i's are drawn uniformly at random from the multiplicative group $\mathbb{Z}_p^*$. Then, the compressed image of m is computed as

$$h(m) = \sum_{i=1}^{b} k_i \times m_i \pmod{p}. \tag{4.1}$$

Note that the key need not to be as long as the message, otherwise, such constructions will be impractical. That is, there are standard techniques so that the same key can be used to hash messages of arbitrary lengths (see, e.g., [50, 19, 10] for the description of such techniques).

The security of universal hash-function families based MACs depends on the probability of message collision. That is, if two distinct messages m and m' hash to the same image (i.e., $h(m) = h(m')$), then they will have the same authentication tag. Consequently, for a message-tag pair, if an adversary can come up with a different message that hashes to the same value, successful forgery can be accomplished with high probabilities. Therefore, the most important security property of universal hash functions is their probabilities of message collisions.

Carter and Wegman suggested the hash function of equation (4.1) with the primes $p = 2^{16} + 1$ or $p = 2^{32} - 1$ [19]. Halevi and Krawczyk later suggested the same equation with any prime $2^{32} < p < 2^{32} + 2^{16}$. They designed their MMH family, one of the fastest universal hash-function families, with $p = 2^{32} + 15$, the smallest prime between 2^{32} and $2^{32} + 2^{16}$ [19]. Etzel *et al.* proposed a variant of the MMH family of [19] that can be faster in some applications [17].

When the hash function is computed modulo a prime integer, equation (4.1) is known to be $(p-1)^{-1}$-AU. In fact, it is shown to be $(p-1)^{-1}$-AΔU in [19] (the notion of ϵ-AΔU is a stronger notion than ϵ-AU; interested readers may refer to [19] for the precise definition of ϵ-AΔU hash families).

The security proofs of all such constructions rely on the fact that computations are performed over integer fields, i.e., the moduli must be prime integers. To the best of our knowledge, no previous work has studied the security of such constructions when the computations are performed over finite integer rings, i.e., not restricting the moduli to prime integers. We aim to provide the first such analysis.

4.2 Constructing codes with secrecy

In this section, we describe a construction of codes with secrecy based on the same principle of Section 4.1; that is, the security of the construction restricts the computa-

tions to be performed over an integer field. What we will describe here is a generalization of the construction appeared in [2], in which we allow operations to be performed over a finite integer ring instead of a field. (Similar constructions have also appeared in [38, 1]). As in the computationally secure constructions discussed in Section 4.1, the codes in [38, 2, 1] demand that operations must be performed over the integer field $\mathbb{Z}_p$; no previous work has studied the probability of deception of such codes when computations are performed over arbitrary finite integer rings. Other codes with secrecy include, but are not limited to, [40, 42, 48, 15].

Let the legitimate users agree on an ℓ-bit long positive integer n, where ℓ is a security parameter. The users share a secret key $k = k_1||k_2$, where k_1 and k_2 are drawn *uniformly* and *independently* from $\mathbb{Z}_n$ and $\mathbb{Z}_n^*$, respectively.

For any message $m \in \mathbb{Z}_n^*$, define $\psi_{k_1}(m) : \mathbb{Z}_n^* \to \mathbb{Z}_n$ and $\psi_{k_2}(m) : \mathbb{Z}_n^* \to \mathbb{Z}_n^*$ as follows:

$$\psi_{k_1}(m) \equiv k_1 + m \pmod{n}, \tag{4.2}$$

$$\psi_{k_2}(m) \equiv k_2 \times m \pmod{n}. \tag{4.3}$$

Equivalently, the exclusive-or operation can be used instead of the addition operation in equation (4.2) without affecting the cipher's security properties [2]. We will refer to $\psi_{k_1}(m)$ and $\psi_{k_2}(m)$ as the ciphertext and authentication tag, respectively. Then, as a function of the key k, the output of the system, $\psi_k(m)$, is the concatenation of the ciphertext and the authentication tag. That is,

$$\psi_k(m) = \psi_{k_1}(m) \,||\, \psi_{k_2}(m). \tag{4.4}$$

A block diagram to implement the described authenticated encryption scheme is depicted in Figure 1 (a).

Upon receiving a ciphertext $\psi'_k(m)$, the legitimate receiver extracts the plaintext m' as follows:

$$m' = \psi'_{k_1}(m) - k_1 \pmod{n}. \tag{4.5}$$

The integrity of the extracted m' is verified by the following check:

$$m' \times k_2 \stackrel{?}{\equiv} \psi'_{k_2}(m) \pmod{n}. \tag{4.6}$$

The notations $\psi'_k(m)$ and m' are to reflect the possibility that the received ciphertext and the extracted plaintext are different than the transmitted ones. The ciphertext is considered valid if and only if the integrity check of equation (4.6) is passed.

A block diagram describing the decryption and integrity check of the scheme is shown in Figure 1 (b).

5 Security analysis

This section will be dedicated to analyzing the security of the authentication with secrecy detailed in Section 4.2, although the bounds on deception probabilities applies to both constructions of Section 4.1 and Section 4.2.

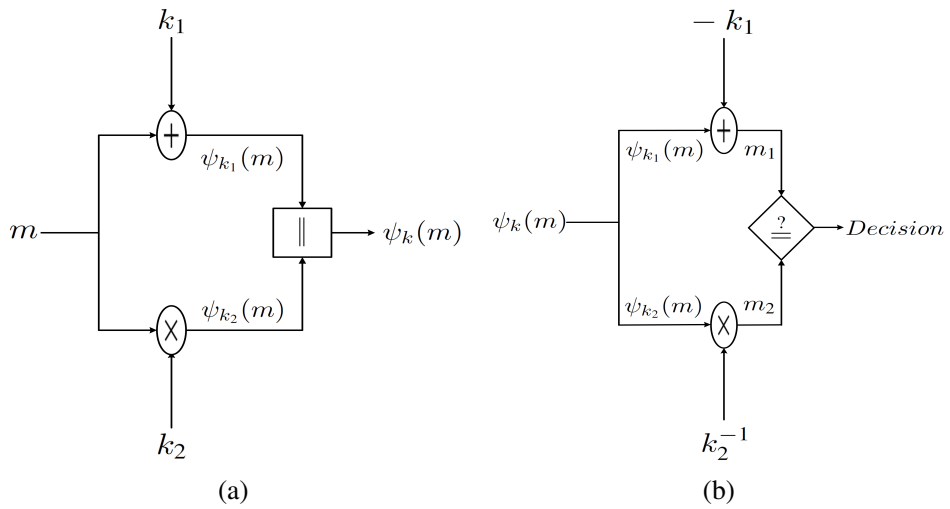

Figure 1. (a) A block diagram to implement the authenticated encryption scheme, and (b) A block diagram implementing the decryption and the validity check of the studied scheme. The addition and multiplication operations are performed over the ring $\mathbb{Z}_n$.

The scheme described in Section 4.2 is designed to achieve two security objectives, confidentiality and integrity. More specifically, by restricting computations to be performed over integer fields, the scheme in Section 4.2 achieves Shannon's perfect secrecy in addition to message integrity [2]. Even though the main emphasis of this work is to analyze the effect of working with arbitrary finite integer rings on the integrity of the scheme, we will show in Section 5.1, for completeness of presentation, the effect on the confidentiality of the scheme when computations are allowed to be performed over arbitrary integer rings. In Section 5.2 we address the main focus of the paper, namely, the bounds on the probabilities of successful message forgery.

5.1 Perfect secrecy

Corollary 5.1. *If encrypted messages are restricted to belong to $\mathbb{Z}_n^*$, the scheme of Section 4.2 achieves perfect secrecy (in Shannon's sense).*

Corollary 5.1 is a direct consequence of Lemma 3.4. To see this, observe that the results of equations (4.2) and (4.3) are defined on a group $G = \mathbb{Z}_n \times \mathbb{Z}_n^*$.

Remark 5.2. Restricting the message m to be relatively prime to n does not impose a significant limitation on the system since, for example, any non-trivial message will satisfy the condition when n is a prime integer. For an arbitrary positive integer n, the message can be padded to be relatively prime to n. Moreover, the system will still work without this restriction; however, perfect secrecy is not achieved.

To illustrate how perfect secrecy is violated when messages are not restricted to the multiplicative group, consider an arbitrary message $m \in \mathbb{Z}_n$ to be encrypted. If

$m \in \mathbb{Z}_n \backslash \mathbb{Z}_n^*$, by Lemma 3.2, the resulting ψ_{k_2} will be in $\mathbb{Z}_n \backslash \mathbb{Z}_n^*$. On the other hand, since $k_2 \in \mathbb{Z}_n^*$, if $m \in \mathbb{Z}_n^*$, by Lemma 3.3, the resulting ψ_{k_2} will be in $\mathbb{Z}_n^*$. Therefore, an adversary observing the authentication tag ψ_{k_2} can determine a subset of the message space that the encrypted message belongs to (if $\psi_{k_2} \in \mathbb{Z}_n \backslash \mathbb{Z}_n^*$ then $m \in \mathbb{Z}_n \backslash \mathbb{Z}_n^*$ and if $\psi_{k_2} \in \mathbb{Z}_n^*$ then $m \in \mathbb{Z}_n^*$); thus, revealing partial information about the encrypted message. Otherwise put,

$$\Pr(\boldsymbol{m} = m | \psi_{k_2} \in \mathbb{Z}_n^*) = \begin{cases} \frac{1}{|\mathbb{Z}_n^*|} & \text{if } m \in \mathbb{Z}_n^*, \\ 0 & \text{if } m \in \mathbb{Z}_n \backslash \mathbb{Z}_n^*, \end{cases} \tag{5.1}$$

and similarly for the case where $\psi_{k_2} \in \mathbb{Z}_n \backslash \mathbb{Z}_n^*$. Therefore,

$$\Pr(\boldsymbol{m} = m | \boldsymbol{\psi_{k_2}} = \psi_{k_2}) \neq \Pr(\boldsymbol{m} = m) \tag{5.2}$$

for all plaintext m and all ciphertext ψ_{k_2}; a clear violation of Definition 2.1 of perfect secrecy.

5.2 Message integrity

In what follows, we address message integrity of authentication codes based on the universal hash family under analysis. Even though the analysis applies to both schemes described in Section 4, we will use the notations of Section 4.2.

As discussed in Section 4.2, the main purpose of ψ_{k_2} is to serve as an authentication tag (MAC) for the encrypted message m. Thus, there are two cases to be considered, modifying ψ_{k_1} alone, and modifying both ψ_{k_1} and ψ_{k_2}. Modifying ψ_{k_2} alone, since it serves as a MAC, does not lead to extracting a false plaintext.

- CASE I. MODIFYING THE CIPHERTEXT ONLY

  Assume that ψ_{k_1} has been modified, by a man in the middle, to ψ'_{k_1}. Since k_1 is known to the receiver, this modification will lead to the extraction of an m' different than the encrypted m; that is, $m' = \psi'_{k_1} - k_1 \pmod n$. Let $m' = m + \delta \pmod n$, for some $\delta \in \mathbb{Z}_n \backslash \{0\}$. To be accepted by the receiver, m' must satisfy the following integrity check:

$$\begin{aligned} m' \times k_2 &\equiv (m + \delta) \times k_2 \pmod n && (5.3) \\ &\equiv (m \times k_2) + (\delta \times k_2) \pmod n && (5.4) \\ &\stackrel{?}{\equiv} \psi_{k_2} \pmod n && (5.5) \\ &\equiv m \times k_2 \pmod n. && (5.6) \end{aligned}$$

  Equivalently, the integrity check in equation (5.5) is satisfied if and only if the following condition holds:

$$\delta \times k_2 \equiv 0 \pmod n. \tag{5.7}$$

That is, modification of ψ_{k_1} alone will go undetected if and only if it is modified by a δ that satisfies equation (5.7). Section 5.2.1 provides detailed probabilistic analysis of equation (5.7).

- CASE II. MODIFYING BOTH THE CIPHERTEXT AND THE MAC

  In a different scenario, the adversary may attempt to modify both ψ_{k_1} and ψ_{k_2} so that a false message will be validated. Assume that ψ_{k_1} has been modified so that the extracted message becomes $m' = m + \delta \pmod n$, for some $\delta \in \mathbb{Z}_n \backslash \{0\}$. Also, assume that ψ_{k_2} has been modified to $\psi'_{k_2} = \psi_{k_2} + \epsilon \pmod n$, for some $\epsilon \in \mathbb{Z}_n \backslash \{0\}$. The integrity of m' is verified using the received ψ'_{k_2} as follows:

$$\begin{aligned} \psi_{k_2} + \epsilon &\equiv \psi'_{k_2} \pmod n && (5.8)\\ &\stackrel{?}{\equiv} m' \times k_2 \pmod n && (5.9)\\ &\equiv (m+\delta) \times k_2 \pmod n && (5.10)\\ &\equiv (m \times k_2) + (\delta \times k_2) \pmod n && (5.11)\\ &\equiv \psi_{k_2} + (\delta \times k_2) \pmod n. && (5.12) \end{aligned}$$

  Equivalently, the false m' will be accepted if and only if the following condition is satisfied:

$$\epsilon \equiv \delta \times k_2 \pmod n. \tag{5.13}$$

  That is, modification of ψ_{k_1} by a value δ and ψ_{k_2} by a value ϵ will go undetected if and only if δ and ϵ satisfy equation (5.13). Section 5.2.2 provides detailed probabilistic analysis of equation (5.13).

5.2.1 Analysis of modifying ciphertext only

As derived above, an adversary modifying the ciphertext ψ_{k_1} in order to make the legitimate receiver authenticate a false message is successful if and only if she can solve the congruence

$$\delta \times k_2 \equiv 0 \pmod n \tag{5.14}$$

for an unknown k_2 uniformly distributed over $\mathbb{Z}_n^*$. To analyze the adversary's ability to solve this congruence for an arbitrary finite integer n, we start with the following lemma.

Lemma 5.3. *Let n be any fixed finite integer. For any nonzero elements α and β in $\mathbb{Z}_n$, if n divides $\alpha \times \beta$, then both α and β must belong to $\mathbb{Z}_n \backslash \mathbb{Z}_n^*$. Formally, the following one-way implication must hold:*

$$\alpha \times \beta \equiv 0 \pmod n \;\Rightarrow\; \{\alpha, \beta \in \mathbb{Z}_n \backslash \mathbb{Z}_n^*\}. \tag{5.15}$$

Lemma 5.3 is a corollary of more general results shown by Schwarz in [36]. Given Lemma 5.3, the adversary's chances of tampering with the ciphertext ψ_{k_1} in a way undetected by the legitimate receiver is stated in the following theorem.

Theorem 5.4. *Any modification of the ciphertext ψ_{k_1} alone will be detected by the legitimate receiver with probability one.*

Proof. Recall that the modification of ψ_{k_1} will be verified only if

$$\delta \times k_2 \equiv 0 \pmod{n}. \tag{5.16}$$

Lemma 5.3, however, states that equation (5.16) can be satisfied only if both δ and k_2 belong to $\mathbb{Z}_n \backslash \mathbb{Z}_n^*$. Since, by design, k_2 is chosen from $\mathbb{Z}_n^*$, equation (5.16) can never be satisfied. Therefore, any modification of the ciphertext ψ_{k_1} will be detected by its MAC with probability one. □

Next, we analyze the possibility of modifying both the ciphertext and MAC, ψ_{k_1} and ψ_{k_2}, in order to make the legitimate receiver authenticate a false message.

5.2.2 Analysis of modifying both the ciphertext and the MAC

This section constitutes the main contribution of this paper. All previous results stated in this paper were either already known or follow directly from known results. The result of this section, on the other hand, has not appeared in the literature; it will show the direct relation between the prime factorization of the modulus n and the security of any authentication code based on the use of the universal hash family discussed in Section 4.

Recall that the adversary has to find a solution to the congruence

$$\epsilon \equiv \delta \times k_2 \pmod{n}, \tag{5.17}$$

where n is an arbitrary fixed modulus and k_2 is chosen uniformly at random from $\mathbb{Z}_n^*$, in order to make the legitimate receiver authenticate a modified message. To be able to analyze the adversary's ability to solve the congruence in equation (5.17), we start by stating a sequence of lemmas.

The first lemma specifies a necessary and sufficient condition for the existence of a k_2 that satisfies equation (5.17).

Lemma 5.5. *Let n be any finite positive integer. Then, for any nonzero $\epsilon, \delta \in \mathbb{Z}_n$, there exists $k \in \mathbb{Z}_n^*$ satisfying*

$$\epsilon \equiv k \times \delta \pmod{n} \tag{5.18}$$

if and only if

$$\gcd(\epsilon, n) = \gcd(\delta, n). \tag{5.19}$$

Proof. Let $\gcd(\epsilon, n) = \gcd(\delta, n) = r$. By lemma 3.1, there exist two invertible elements $\alpha, \beta \in \mathbb{Z}_n$ so that, $\epsilon \equiv r \times \alpha^{-1} \pmod{n}$ and $\delta \equiv r \times \beta^{-1} \pmod{n}$. Then,

$$\begin{aligned} \epsilon &\equiv r \times \alpha^{-1} \pmod{n} && (5.20)\\ &\equiv r \times \alpha^{-1} \times \beta^{-1} \times \beta \pmod{n} && (5.21)\\ &\equiv \alpha^{-1} \times \beta \times \delta \pmod{n}. && (5.22) \end{aligned}$$

Hence, $k \equiv \alpha^{-1} \times \beta \pmod{n}$ satisfies equation (5.18). Further, $k \in \mathbb{Z}_n^*$ by Lemma 3.3. Therefore, equation (5.19) implies equation (5.18).

Now, suppose that $\epsilon \equiv k \times \delta \pmod{n}$ for some $k \in \mathbb{Z}_n^*$. Let $r = \gcd(\epsilon, n)$ and $s = \gcd(\delta, n)$ and suppose, without loss of generality, that $r > s$. Again, by Lemma 3.1, there exist $\alpha, \beta \in \mathbb{Z}_n^*$ satisfying $\epsilon \equiv r \times \alpha^{-1} \pmod{n}$ and $\delta \equiv s \times \beta^{-1} \pmod{n}$. Then,

$$\begin{aligned} r \times \alpha^{-1} &\equiv \epsilon \pmod{n} && (5.23)\\ &\equiv k \times \delta \pmod{n} && (5.24)\\ &\equiv k \times s \times \beta^{-1} \pmod{n}, && (5.25) \end{aligned}$$

and multiplying both sides by α yields,

$$r \equiv s \times (\alpha \times \beta^{-1} \times k) \pmod{n}. \quad (5.26)$$

Also, since $r \mid n$, there exists an $\ell \in \mathbb{Z}_n$ such that $\ell \cdot r = n$. Multiplying both sides of equation (5.26) by ℓ yields,

$$0 \equiv (\ell \times s) \times (\alpha \times \beta^{-1} \times k) \pmod{n}. \quad (5.27)$$

Since $s < r$ by hypothesis, the first factor on the right hand side is strictly less than $n = \ell \cdot r$; hence, $(\ell \times s)$ is a nonzero element in $\mathbb{Z}_n$. By Lemma 3.3, the second factor belongs to $\mathbb{Z}_n^*$; a contradiction to Lemma 5.3, which states that for the product of two nonzero integers to be congruent to zero modulo n, both integers must be in $\mathbb{Z}_n \backslash \mathbb{Z}_n^*$. Therefore, $r = s$, and the lemma follows. □

Lemma 5.5 specifies a necessary condition for the successful forgery by modifying the ciphertext by a $\delta \in \mathbb{Z}_n \backslash \{0\}$ and the MAC by an $\epsilon \in \mathbb{Z}_n \backslash \{0\}$. Namely, $\gcd(\delta, n)$ must be equal to $\gcd(\epsilon, n)$; otherwise, there does not exist a shared key $k_2 \in \mathbb{Z}_n^*$ that could possibly satisfy equation (5.17) for the chosen δ and ϵ.

Assume now that an adversary has chosen nonzero δ and ϵ that satisfy the necessary condition of Lemma 5.5. Given the value of δ, what is the probability that the chosen ϵ will satisfy equation (5.17). To be able to answer this question, we introduce the following set.

Definition 5.6 (The set of common gcd's).
For any fixed integer δ, define $T(\delta)$ to be the set of ϵ's that satisfy equation (5.17) for at least one $k \in \mathbb{Z}_n^*$. That is,

$$T(\delta) := \{\epsilon \in \mathbb{Z}_n : \exists\, k \in \mathbb{Z}_n^* \text{ such that } \epsilon \equiv \delta \times k \pmod{n}\}. \quad (5.28)$$

By Lemma 5.5, this set is equal to the set of ϵ's in $\mathbb{Z}_n$ such that $\gcd(\epsilon, n) = \gcd(\delta, n)$. Therefore, it can be written as,

$$T(\delta) = \{\epsilon \in \mathbb{Z}_n : \gcd(\epsilon, n) = \gcd(\delta, n)\}. \tag{5.29}$$

For the rest of the paper, the representations in equations (5.28) and (5.29) of the set of common gcd's will be used interchangeably to define the set $T(\delta)$.

To be able to quantify the adversary's probability of successful forgery, we need to answer the following question: For an $\epsilon \in T(\delta)$, how many possible secret keys k_2's can satisfy equation (5.17) for the given (δ, ϵ) pair? More importantly, for two distinct ϵ's in $T(\delta)$, say ϵ and ϵ', what is the relation between the number of k_2's in $\mathbb{Z}_n^*$ that satisfy equation (5.17) for each of them? This question is important since, for a given δ, an intelligent adversary will choose the ϵ that maximizes her probability of successful forgery. The following lemma addresses this question.

Lemma 5.7. *Fix any $\delta \in \mathbb{Z}_n$ and let $\epsilon, \epsilon' \in T(\delta)$. Define the set K_ϵ to be the set of all k's in $\mathbb{Z}_n^*$ that satisfy equation (5.17) for the given δ and ϵ. Similarly, define the set $K_{\epsilon'}$ to be the set of all k's in $\mathbb{Z}_n^*$ that satisfy equation (5.17) for δ and ϵ'. That is, $K_\epsilon := \{k \in \mathbb{Z}_n^* : \delta \times k \equiv \epsilon \pmod n\}$ and $K_{\epsilon'} := \{k \in \mathbb{Z}_n^* : \delta \times k \equiv \epsilon' \pmod n\}$. Then $|K_\epsilon| = |K_{\epsilon'}|$, i.e., the sets K_ϵ and $K_{\epsilon'}$ have the same cardinality.*

Proof. Without loss of generality, assume $|K_\epsilon| < |K_{\epsilon'}| = \ell$, and let $K_{\epsilon'} = \{k_1, \ldots, k_\ell\}$, for distinct k_i's. Since $\epsilon \in T(\delta)$, there exists an r satisfying $r \times \delta \equiv \epsilon \pmod n$. Also, since $k_1 \in K_{\epsilon'}$, $\delta \equiv k_1^{-1} \times \epsilon' \pmod n$. Now, for $i = 1, \ldots, \ell$, define r_i as,

$$r_i = r \cdot k_1^{-1} \cdot k_i. \tag{5.30}$$

Then, every r_i satisfies,

$$\begin{aligned} r_i \times \delta &\equiv r \times k_1^{-1} \times k_i \times \delta \pmod n && (5.31)\\ &\equiv r \times k_1^{-1} \times \epsilon' \pmod n && (5.32)\\ &\equiv r \times \delta \pmod n && (5.33)\\ &\equiv \epsilon \pmod n. && (5.34) \end{aligned}$$

Furthermore, the r_i's are distinct: if $r_i = r_j$, then

$$r \times k_1^{-1} \times k_i \equiv r \times k_1^{-1} \times k_j \pmod n. \tag{5.35}$$

Since k_1^{-1} and r are invertible, by cancellation we have $k_i = k_j$, implying that $i = j$. Therefore, the set K_ϵ contains at least ℓ distinct elements, a contradiction to the hypothesis that $|K_\epsilon| < |K_{\epsilon'}|$. Therefore, $|K_\epsilon| = |K_{\epsilon'}|$. □

Lemma 5.7 implies that any ϵ which has the same greatest common divisor with n as δ will have the same number of keys as possible candidates for successful forgery. That is, from the adversary's standpoint, there is no advantage of picking one particular $\epsilon \in T(\delta)$ over the others. The following lemma formalizes this argument.

Lemma 5.8. *Suppose that k is an unknown integer, randomly drawn from $\mathbb{Z}_n^*$. Then for any fixed $\delta \in \mathbb{Z}_n \backslash \{0\}$, the probability of selecting ϵ satisfying $\epsilon \equiv k \times \delta \pmod{n}$ is at most $1/|T(\delta)|$.*

Proof. By the definition of $T(\delta)$ and Lemma 5.5, all valid ϵ's are in $T(\delta)$, and any ϵ in $T(\delta)$ is a valid choice. Also, by Lemma 5.7, the number of possible values of k that map δ to any ϵ is the same, so there is no advantage in picking one ϵ over another, i.e., the ϵ's are uniformly distributed in $T(\delta)$. Hence, for a given $\delta \in \mathbb{Z}_n \backslash \{0\}$, the probability of selecting an $\epsilon \in T(\delta)$ that satisfies equation (5.17) is $1/|T(\delta)|$. □

Lemma 5.8 implies that the adversary's best strategy for successful forgery is to choose the δ that minimizes $|T(\delta)|$. (Observe that the cardinality of $T(\delta)$ is at least one since $\delta \in T(\delta)$ for any $\delta \in \mathbb{Z}_n$.) The next two lemmas address the problem of minimizing $|T(\delta)|$.

We start with a lemma that relates the cardinality of the set T with the Euler totient function φ.

Lemma 5.9. *For any integer α that divides n, $|T(n/\alpha)| = \varphi(\alpha)$. More explicitly, the set $T(n/\alpha)$ can be expressed as,*

$$T(n/\alpha) = \frac{n}{\alpha} \{\beta \in \mathbb{Z}_\alpha : \gcd(\beta, \alpha) = 1\}. \tag{5.36}$$

Proof. The fact that $\alpha | n$ implies that $\gcd(n/\alpha, n) = n/\alpha$. Therefore, by the definition of T in equation (5.29),

$$T(n/\alpha) = \{\epsilon \in \mathbb{Z}_n : \gcd(\epsilon, n) = \gcd(n/\alpha, n) = n/\alpha\}. \tag{5.37}$$

Now, for any $\beta \in \mathbb{Z}_\alpha$ such that $\gcd(\beta, \alpha) = 1$, using the fact that $\gcd(ka, kb) = k \gcd(a, b)$ [11], we get:

$$\gcd(\beta, \alpha) = 1 \iff \gcd(\frac{n}{\alpha}\beta, \frac{n}{\alpha}\alpha) = \frac{n}{\alpha} \tag{5.38}$$

$$\iff \gcd(\frac{n}{\alpha}\beta, n) = \frac{n}{\alpha} \tag{5.39}$$

$$\overset{(5.37)}{\iff} \frac{n}{\alpha}\beta \in T(n/\alpha). \tag{5.40}$$

Furthermore, for distinct $\beta_1, \beta_2 \in \mathbb{Z}_\alpha$, $\frac{n}{\alpha}\beta_1$ and $\frac{n}{\alpha}\beta_2$ are distinct elements of $\mathbb{Z}_n$. This is because $n > \max\{\frac{n}{\alpha}\beta_1, \frac{n}{\alpha}\beta_2\}$ (since $\alpha > \max\{\beta_1, \beta_2\}$). Therefore, there is a one-to-one correspondence between the set $\{\beta \in \mathbb{Z}_\alpha : \gcd(\beta, \alpha) = 1\}$ and the set $\{\gamma \in \mathbb{Z}_n : \gamma \in T(n/\alpha)\}$. □

We can now state the relation between the cardinality of $T(\delta)$, for any δ, and the choice of the underlying integer ring. More specifically, the following lemma emphasizes the effect of the prime factorization of n on the cardinality of the smallest $T(\delta)$.

Lemma 5.10. *If p is the smallest prime factor of n, then $|T(\delta)| \geq |T(n/p)|$ for any $\delta \in \mathbb{Z}_n$.*

Proof. Let $\delta \in \mathbb{Z}_n$ and let p be the smallest prime factor of n. By Lemma 5.9, $|T(n/p)| = \varphi(p) = p - 1$. Now, recall that:

$$T(\delta) = \{\epsilon \in \mathbb{Z}_n : \gcd(\epsilon, n) = \gcd(\delta, n)\}. \tag{5.41}$$

Then, if $\gcd(\delta, n) = 1$, by equation (5.41),

$$|T(\delta)| = |\mathbb{Z}_n^*| = \varphi(n); \tag{5.42}$$

and we know, by Lemma 3.5, that

$$\varphi(n) \geq p - 1; \tag{5.43}$$

and, by Lemma 5.9, that

$$p - 1 = |T(n/p)|. \tag{5.44}$$

Thus, $|T(\delta)| \geq |T(n/p)|$ for all δ's that are relatively prime to n.

It remains to show that the same is true for δ's that are not relatively prime to n. Let $\gcd(\delta, n) = d > 1$, then, by equation (5.41),

$$T(\delta) = T(\gcd(\delta, n)) = T(d). \tag{5.45}$$

Therefore, we can assume, without loss of generality, that $\delta | n$ (since for any δ such that $\gcd(\delta, n) = d$, $T(\delta) = T(d)$ and $d|n$). Now, write $\delta = n/\alpha$ and let α be written in its prime factorization form as $\alpha = \prod_i p_i^{e_i}$, where the p_i's are distinct primes. Then, $\varphi(\alpha) = \prod_i (p_i - 1) p_i^{e_i - 1}$. Since $\alpha | n$, and p is the smallest prime factor of n, $p \leq p_i$ for any i. Hence, by Lemma 5.9,

$$|T(\delta)| = |T(n/\alpha)| = \varphi(\alpha) \geq p - 1 = |T(n/p)|. \tag{5.46}$$

Therefore, for any $\delta \in \mathbb{Z}_n$, $|T(\delta)| \geq |T(n/p)|$. □

We can now state the main theorem analyzing the adversary's probability of successful forgery by modifying both ciphertext ψ_{k_1} and the MAC ψ_{k_2}.

Theorem 5.11. *Let p be the smallest prime factor of n. Then, an adversary modifying both the ciphertext ψ_{k_1} and the MAC ψ_{k_2} will be successful with probability at most $1/(p-1)$.*

Proof. Recall that an adversary modifying ψ_{k_1} and ψ_{k_2} will be successful only if she can choose δ, ϵ such that:

$$\epsilon \equiv \delta \times k_2 \pmod{n}. \tag{5.47}$$

By Lemma 5.8, the probability of choosing δ, ϵ that satisfy equation (5.47) is given by $1/|T(\delta)|$. To maximize the probability of successful forgery, the adversary can choose δ that minimizes the size of $T(\delta)$. By Lemma 5.10, the best choice of δ that minimizes $T(\delta)$ is $\delta = n/p$, where p is the smallest prime factor of n. Finally, by Lemma 5.9, $|T(n/p)| = p - 1$, and the theorem follows. □

6 Choice of integer rings

The described authenticated encryption schemes is designed to achieve two main objectives, message confidentiality and integrity. In this section we summarize the effect of the underlying integer ring on the security properties of the scheme.

It has been shown, in Section 5.1, that reducing message space to the multiplicative group of integers modulo n is a necessary condition for the scheme to achieve perfect secrecy. Consequently, the choice of the underlying integer ring will be a factor for the number of possible messages that can be encrypted with perfect secrecy.

In Section 5.2.1, it was shown that an adversary modifying ψ_{k_1} only will be successful only if ψ_{k_1} is perturbed by an integer δ that satisfies

$$\delta \times k_2 \equiv 0 \pmod{n}. \tag{6.1}$$

Moreover, it was shown that choosing k_2 from the multiplicative group $\mathbb{Z}_n^*$ is a sufficient condition to guarantee that no nonzero $\delta \in \mathbb{Z}_n$ will satisfy equation (6.1). Therefore, the choice of the underlying integer ring does not play an important role in the protection against modifying ψ_{k_1} only, other than restricting k_2 to be chosen from the multiplicative group $\mathbb{Z}_n^*$.

The choice of the underlying integer ring has its most impact when an adversary modifies both ψ_{k_1} and ψ_{k_2}. As discussed in Section 5.2.2, the adversary is successful in tampering with the message, in a way undetected by the legitimate receiver, only if she can select ϵ, δ satisfying:

$$\epsilon \equiv \delta \times k_2 \pmod{n}. \tag{6.2}$$

The proof of Theorem 5.11 describes the following attack on the scheme. Suppose that the scheme designer chooses a modulus with prime factorization given by $n = p_1^{e_1} \cdots p_k^{e_k}$, where the p_i's are ordered increasingly. Assuming the adversary is able to factor n, then she can choose $\delta = n/p_1$ to maximize her probability of successful forgery. The resulting $\delta \times k_2 \pmod{n}$ will be, from the adversary's perspective, a random element in the set of multiples of n/p_1 (excluding 0 because k_2 is known to be relatively prime to n). Consequently, by randomly choosing an integer ϵ from the set $\{m\frac{n}{p_1} \pmod{n}, \text{ for } m = 1, \ldots, p_1 - 1\}$, the adversary can tamper with the message without detection with probability $1/(p_1 - 1)$. The following numerical example illustrates the attack.

Example 6.1. Let $n = 45 = 3^2 \times 5$. According to Theorem 5.11, the adversary can maximize her probability of successful forgery by choosing $\delta = n/3 = 15$. Moreover, the secret key k_2 is restricted to belong to the multiplicative group $\mathbb{Z}_{45}^*$, that is, $k_2 \in \{1, 2, 4, 7, 8, 11, 13, 14, 16, 17, 19, 22, 23, 26, 28, 29, 31, 32, 34, 37, 38, 41, 43, 44\}$. Th-

erefore, the resulting $\delta \times k_2 \pmod{45} := \epsilon$ is equal to

$$15 \times 1 \equiv 15 \pmod{45}, \tag{6.3}$$
$$15 \times 2 \equiv 30 \pmod{45}, \tag{6.4}$$
$$15 \times 4 \equiv 15 \pmod{45}, \tag{6.5}$$
$$15 \times 7 \equiv 15 \pmod{45}, \tag{6.6}$$
$$15 \times 8 \equiv 30 \pmod{45}, \tag{6.7}$$
$$15 \times 11 \equiv 30 \pmod{45}, \tag{6.8}$$
$$15 \times 13 \equiv 15 \pmod{45}, \tag{6.9}$$
$$15 \times 14 \equiv 30 \pmod{45}, \tag{6.10}$$
$$15 \times 16 \equiv 15 \pmod{45}, \tag{6.11}$$
$$15 \times 17 \equiv 30 \pmod{45}, \tag{6.12}$$
$$15 \times 19 \equiv 15 \pmod{45}, \tag{6.13}$$
$$15 \times 22 \equiv 15 \pmod{45}, \tag{6.14}$$
$$15 \times 23 \equiv 30 \pmod{45}, \tag{6.15}$$
$$15 \times 26 \equiv 30 \pmod{45}, \tag{6.16}$$
$$15 \times 28 \equiv 15 \pmod{45}, \tag{6.17}$$
$$15 \times 29 \equiv 30 \pmod{45}, \tag{6.18}$$
$$15 \times 31 \equiv 15 \pmod{45}, \tag{6.19}$$
$$15 \times 32 \equiv 30 \pmod{45}, \tag{6.20}$$
$$15 \times 34 \equiv 15 \pmod{45}, \tag{6.21}$$
$$15 \times 37 \equiv 15 \pmod{45}, \tag{6.22}$$
$$15 \times 38 \equiv 30 \pmod{45}, \tag{6.23}$$
$$15 \times 41 \equiv 30 \pmod{45}, \tag{6.24}$$
$$15 \times 43 \equiv 15 \pmod{45}, \tag{6.25}$$
$$15 \times 44 \equiv 30 \pmod{45}. \tag{6.26}$$

That is, the resulting ϵ will be 15 or 30 with equal probability. (Similarly, one can show that, by choosing $\delta = n/5 = 9$, the resulting ϵ is uniformly distributed over $\{9, 18, 27, 36\}$). In any case, the resulting ϵ will be uniformly distributed over the multiples of n/p, where p is a prime factor of n, and the p that minimizes the cardinality of the set of possible ϵ's is the smallest prime factor of n.

As an illustration of the importance of the underlying integer ring, in what follows, we show the best and the worst choices of integer rings in terms of security against man in the middle attacks.

6.1 Prime moduli

When the used modulus is a prime integer p, the underlying integer ring $\mathbb{Z}_p$ becomes a field. Not surprisingly, the use of a prime modulus gives the best security performances against message corruption attacks. Since the smallest prime factor of p is p itself, by Theorem 5.11, the adversary's probability of successful forgery is $1/(p-1)$. That is, there is no advantage of choosing a δ over another. In other words, no matter what the value of δ an adversary chooses, the resulting ϵ will be uniformly distributed over the entire set of nonzero element $\{1, 2, \cdots, p-1\}$.

6.2 Even moduli

Even moduli give the worst security against message modification. An active adversary can take advantage of the even modulus to make the intended receiver authenticate a false message with probability one. This is due to the fact that the smallest prime factor of n is 2. Therefore, by Theorem 5.11, the adversary's probability of successful forgery is $1/(2-1)$. To illustrate the attack, let the adversary choose $\delta = n/2$. Since $k_2 \in \mathbb{Z}_n^*$, and n is an even integer, k_2 must be an odd integer, which can be written in the form $2r+1$ for some positive integer r. Then,

$$\epsilon \equiv \delta \times k_2 \pmod{n} \tag{6.27}$$

$$\equiv (\frac{n}{2}) \times (2r+1) \pmod{n} \tag{6.28}$$

$$\equiv \frac{n}{2} \pmod{n}. \tag{6.29}$$

Therefore, choosing $\delta = \epsilon = n/2$ guarantees that the modification will go undetected with probability one. Consequently, even moduli cannot be used to implement the described scheme since an active adversary can always perturb both ψ_{k_1} and ψ_{k_2} in a way undetected by the legitimate receiver.

7 Conclusion

In this paper, we investigated authentication based on a class of universal hash-function families that have been appeared in the literature. Although the studied universal hash-function family has appeared in many places, computations have always been performed modulo prime integers. In this work, we analyzed the security of message authentication when computations are performed over arbitrary finite integer rings. We derived a direct relation between the security of authentication and the underlying integer ring $\mathbb{Z}_n$. Specifically, we showed that the bound on successful forgery is proportional to the reciprocal of the smallest prime factor of the used modulus n.

Acknowledgments. We would like to thank anonymous reviewers for their valuable comments regarding the presentation and the technical contents of the paper. They

brought to our attention some missed references discussing similar results. Their comments have helped us improving the paper significantly.

References

[1] B. Alomair, L. Lazos, and R. Poovendran, *Securing Low-cost RFID Systems: an Unconditionally Secure Approach*, Journal of Computer Security - Special Issue on RFID System Security (2010).

[2] B. Alomair and R. Poovendran, *Information Theoretically Secure Encryption with Almost Free Authentication*, Journal of Universal Computer Science 15 (2009), pp. 2937–2956.

[3] M. Atici and D. Stinson, *Universal hashing and multiple authentication*. Advances in Cryptology–Crypto'96, 1109, pp. 16–30, Lecture Notes in Computer Science. Springer, 1996.

[4] M. Bellare, R. Guerin, and P. Rogaway, *XOR MACs: New methods for message authentication using finite pseudorandom functions*. Advances in Cryptology–Crypto'95, 963, pp. 15–28, Lecture Notes in Computer Science. Springer, 1995.

[5] M. Bellare, J. Kilian, and P. Rogaway, *The Security of the Cipher Block Chaining Message Authentication Code*, Journal of Computer and System Sciences 61 (2000), pp. 362–399.

[6] D. Bernstein, *Floating-point arithmetic and message authentication*, Unpublished manuscript. Available at `http://cr.yp.to/papers.html#hash127`.

[7] ———, *The Poly1305-AES message-authentication code*. Fast Software Encryption–FSE'05, 3557, pp. 32–49, Lecture Notes in Computer Science. Springer, 2005.

[8] J. Bierbrauer, *A2-codes from universal hash classes*. Advances in Cryptology–Eurocrypt'95, 921, pp. 311–318, Lecture Notes in Computer Science. Springer, 1995.

[9] ———, *Universal hashing and geometric codes*, Designs, Codes and Cryptography 11 (1997), pp. 207–221.

[10] J. Black, S. Halevi, H. Krawczyk, T. Krovetz, and P. Rogaway, *UMAC: Fast and Secure Message Authentication*, Advances in Cryptology–Crypto'99 1666 (1999), pp. 216–233.

[11] D. Burton, *Elementary Number Theory*. McGraw Hill, 2002.

[12] J. Carter and M. Wegman, *Universal classes of hash functions*. Symposium on Theory of Computing–STOC'77, pp. 106–112, ACM, 1977.

[13] ———, *Universal hash functions*, Journal of Computer and System Sciences 18 (1979), pp. 143–154.

[14] L. Casse, K. Martin, and P. Wild, *Bounds and characterizations of authentication/secrecy schemes*, Designs, Codes and Cryptography 13 (1998), pp. 107–129.

[15] M. De Soete, *Some constructions for authentication-secrecy codes*. Advances in cryptology–Eurocrypt'88, 330, pp. 57–75, Lecture Notes in Computer Science. Springer, 1988.

[16] M. Dworkin, National Institute of Standards, and Technology (US), *Recommendation for block cipher modes of operation: The CMAC mode for authentication*. US Dept. of Commerce, Technology Administration, National Institute of Standards and Technology, 2005.

[17] M. Etzel, S. Patel, and Z. Ramzan, *Square hash: Fast message authentication via optimized universal hash functions*, Advances in Cryptology–Crypto'99 1666 (1999), pp. 786–786.

[18] E. Gilbert, F. MacWilliams, and N. Sloane, *Codes which detect deception*, Bell System Technical Journal 53 (1974), pp. 405–424.

[19] S. Halevi and H. Krawczyk, *MMH: Software message authentication in the Gbit/second rates*. Fast Software Encryption–FSE'97, 1267, pp. 172–189, Lecture Notes in Computer Science. Springer, 1997.

[20] G. Hardy and E. Wright, *An Introduction to the Theory of Numbers*. Clarendon Press, 1979.

[21] T. Helleseth and T. Johansson, *Universal hash functions from exponential sums over finite fields and Galois rings*. Advances in cryptology–Crypto'96, 1109, pp. 31–44, Lecture Notes in Computer Science. Springer, 1996.

[22] T. Iwata and K. Kurosawa, *OMAC: One-Key CBC-MAC*. Fast Software Encryption–FSE'03, 2887, pp. 129–153, Lecture Notes in Computer Science. Springer, 2003.

[23] T. Johansson, *Lower bounds on the probability of deception in authentication with arbitration*, IEEE Transactions on Information Theory 40 (1994), pp. 1573–1585.

[24] J. Kaps, K. Yuksel, and B. Sunar, *Energy scalable universal hashing*, IEEE Transactions on Computers 54 (2005), pp. 1484–1495.

[25] H. Krawczyk, *LFSR-based hashing and authentication*. Advances in Cryptology–Crypto'94, 839, pp. 129–139, Lecture Notes in Computer Science. Springer, 1994.

[26] ———, *New hash functions for message authentication*. Advances in cryptology–Eurocrypt'95, 921, pp. 301–310, Lecture Notes in Computer Science. Springer, 1995.

[27] K. Kurosawa, *New bound on authentication code with arbitration*, Advances in Cryptology–Crypto'94 839 (1994), pp. 140–149.

[28] K. Kurosawa and T. Iwata, *TMAC: Two-Key CBC MAC*, Topics in Cryptology–CT-RSA'03 2612 (2003), pp. 33–49.

[29] K. Kurosawa and S. Obana, *Combinatorial Bounds on Authentication Codes with Arbitration*, Designs, Codes and Cryptography 22 (2001), pp. 265–281.

[30] K. Kurosawa, K. Okada, H. Saido, and D. Stinson, *New combinatorial bounds for authentication codes and key predistribution schemes*, Designs, Codes and Cryptography 15 (1998), pp. 87–100.

[31] D. McGrew and J. Viega, *The security and performance of the Galois/Counter Mode (GCM) of operation*, Progress in Cryptology–Indocrypt'04 3348 (2004), pp. 343–355.

[32] M. Naor, G. Segev, and A. Smith, *Tight bounds for unconditional authentication protocols in the manual channel and shared key models*, IEEE Transactions on Information Theory 54 (2008), pp. 2408–2425.

[33] D. Pei, *Information-theoretic bounds for authentication codes and block designs*, Journal of Cryptology 8 (1995), pp. 177–188.

[34] E. Petrank and C. Rackoff, *CBC MAC for real-time data sources*, Journal of Cryptology 13 (2000), pp. 315–338.

[35] B. Preneel and P. Van Oorschot, *On the security of iterated message authentication codes*, IEEE Transactions on Information Theory 45 (1999), pp. 188–199.

[36] S. Schwarz, *The role of semigroups in the elementary theory of numbers*, Mathematica Slovaca 31 (1981), pp. 369–395.

[37] V. Shoup, *On fast and provably secure message authentication based on universal hashing*. Advances in Cryptology–Crypto'96, 1109, pp. 313–328, Lecture Notes in Computer Science. Springer, 1996.

[38] B. Smeets, P. Vanroose, and Z. Wan, *On the construction of authentication codes with secrecy and codes withstanding spoofing attacks of order* $L \geq 2$. Advances in cryptology–Eurocrypt'90, 921, pp. 307–312, Lecture Notes in Computer Science. Springer, 1990.

[39] M. Soete, *New bounds and constructions for authentication/secrecy codes with splitting*, Journal of Cryptology 3 (1991), pp. 173–186.

[40] D. Stinson, *A construction for authentication/secrecy codes from certain combinatorial designs*, Journal of Cryptology 1 (1988), pp. 119–127.

[41] ———, *Some constructions and bounds for authentication codes*, Journal of Cryptology 1 (1988), pp. 37–51.

[42] ———, *The combinatorics of authentication and secrecy codes*, Journal of Cryptology 2 (1990), pp. 23–49.

[43] ———, *Universal hashing and authentication codes*, Designs, Codes and Cryptography 4 (1994), pp. 369–380.

[44] ———, *Cryptography: Theory and Practice*. CRC Press, 2006.

[45] J. Tignol, *Galois'Theory of Algebraic Equations*. World Scientific, 2001.

[46] US National Bureau of Standards, *DES Modes of Operation*, Federal Information Processing Standard (FIPS) Publication 81 (1980).

[47] H. Van Tilborg, *Encyclopedia of cryptography and security*. Springer, 2005.

[48] T. Van Tran, *On the construction of authentication and secrecy codes*, Designs, Codes and Cryptography 5 (1995), pp. 269–280.

[49] M. Wegman and J. Carter, *New classes and applications of hash functions*. Foundations of Computer Science–FOCS'79, pp. 175–182, IEEE, 1979.

[50] ———, *New hash functions and their use in authentication and set equality*, Journal of Computer and System Sciences 22 (1981), pp. 265–279.

Received xxx; revised xxx

Author information

Basel Alomair,
Center of Excellence in Information Assurance (CoEIA), King Saud University, Riyadh, Saudi Arabia
&
Network Security Lab (NSL), University of Washington, Seattle, USA.
Email: alomair@uw.edu

Andrew Clark,
Network Security Lab (NSL), University of Washington, Seattle, USA.
Email: awclark@uw.edu

Radha Poovendran,
Network Security Lab (NSL), University of Washington, Seattle, USA.
Email: rp3@uw.edu