Unconditionally secure chaffing and winnowing with short authentication tags

D.R. Stinson*
David R. Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario N2L 3G1, Canada
dstinson@uwaterloo.ca

March 23, 2007

There are just three tasks.
Firstly, I would like to move this pile from here to there,
but I'm afraid that all I have is this tiny tweezers.
Secondly, I would like to empty this well and fill the other;
but I have no bucket, so you'll have to use this eye dropper.
And, lastly, I must have a hole through this cliff,
and here is a needle to dig it.
Norton Juster, *The Phantom Tollbooth.*

Abstract

Rivest proposed the idea of a chaffing-and-winnowing scheme, in which confidentiality is achieved through the use of an authentication code. Thus it would still be possible to have confidential communications even if conventional encryption schemes were outlawed. Hanaoka *et al.* constructed unconditionally secure chaffing-and-winnowing schemes which achieve perfect secrecy in the sense of Shannon. Their schemes are constructed from unconditionally secure authentication codes.

In this paper, we construct unconditionally secure chaffing-and-winnowing schemes from unconditionally secure authentication codes in which the authentication tags are very short. This could be a desirable feature, because certain types of unconditionally secure authentication codes can provide perfect secrecy if the length of an authentication tag is at least as long as the length of the plaintext. The use of such a code might be prohibited if encryption schemes are made illegal, so it is of interest to construct chaffing-and-winnowing schemes based on "short" authentication tags.

*research supported by NSERC discovery grant 203114-06

1 Introduction

The idea of chaffing-and-winnowing was suggested by Rivest [5]. The hypothetical motivating scenario is that encryption schemes might be outlawed at some future time, while the use of message authentication codes (i.e., MACs) could still remain legal. The basic idea of a chaffing-and-winnowing scheme is to use a MAC to provide confidentiality, thus circumventing the hypothetical ban against encryption. Typically, a sender (Alice) and a receiver (Bob) share a secret key K. Alice prepares a large number of "authenticated messages", each having the form $m = (x, a)$ where each x is an unencrypted plaintext and a is an authentication tag. Then Alice sends all the authenticated messages to Bob. Bob only accepts the message(s) having authentication tags that are valid under the key K. An observer O has no way to distinguish between valid and invalid authentication tags, so O cannot determine the plaintext(s) that Alice is communicating to Bob.

This intriguing idea has not received much study to date. There are two main papers investigating theoretical aspects of chaffing-and-winnowing subsequent to [5], namely Bellare and Boldyreva [1] and Hanaoka *et al.* [4]. The paper [1] gives formal definitions and security treatments of chaffing-and-winnowing schemes, based on the notion of "find-then-guess" security of encryption schemes. The paper [4] studies chaffing-and-winnowing in the setting of unconditional security. The desire is to provide perfect secrecy, based on a suitable unconditionally secure authentication code. That paper also considers "non-malleability" properties of chaffing-and-winnowing schemes. For a paper that discusses practical issues regarding the implementation of chaffing-and-winnowing schemes, see Clayton and Danezis [2].

In this paper, we continue the study of unconditionally secure chaffing-and-winnowing schemes. One possible difficulty with the schemes constructed in [4] is that the underlying authentication codes have the property that the authentication tag has the same entropy as the plaintext. Thus the authentication code might already provide perfect secrecy. Such an authentication code might not be considered "legal" in our motivating scenario.

We are interested in building unconditionally secure chaffing-and-winnowing schemes that are constructed from underlying authentication codes that cannot provide perfect secrecy. In fact, we base our construction on authentication codes that employ only one-bit authenticators. Such authenticators clearly cannot provide perfect secrecy for any plaintext space of cardinality greater than two, so it seems to be an interesting result that we can manufacture unconditionally secure chaffing-and-winnowing schemes from them.

1.1 Our Contributions

The rest of the paper is organized as follows. In Section 1.2, we briefly present some background theory about authentication and secrecy codes. In Section 2, we decribe our model for unconditionally secure chaffing-and-winnowing schemes. In Section 3, we review and discuss the main construction from [4]. In Section 4, we give our new construction for chaffing-and-winnowing schemes based on one-bit authenticators. We prove our scheme is correct, present a modification and discuss its efficiency. We also show that our scheme is optimal among chaffing-and-winnowing schemes with one-bit authentication tags. As well, we present a "hybrid" scheme based on our one-bit chaffing-and-winnowing schemes. A few final comments are made in Section 5.

1.2 Unconditionally Secure Authentication and Secrecy

Unconditionally secure authentication codes were first studied by Gilbert, MacWilliams and Sloane [3]. Simmons developed the theoretical foundations (for a survey, see [7]) and many constructions have been provided over the years. We briefly summarize some definitions, notation and basic results.

An *authentication code* is a four-tuple $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{F})$ such that the following conditions are satisfied:

1. $\mathcal{X}$ is a set of *plaintexts*,
2. $\mathcal{A}$ is a set of *authentication tags*,
3. $\mathcal{K}$ is a set of *keys*, and
4. $\mathcal{F} = \{f_K : K \in \mathcal{K}\}$ is a set of *authentication functions*, where $f_K : \mathcal{X} \to \mathcal{A}$ for every $K \in \mathcal{K}$.

A key $K \in \mathcal{K}$ is chosen uniformly at random[1] by Alice and communicated to Bob over a secure channel. When Alice wants to send an authenticated message to Bob, she chooses a plaintext $x \in \mathcal{X}$ and computes the authentication tag $a = f_K(x)$. The *message* $m = (x, a)$ is sent to Bob over an insecure channel. Bob *accepts* a message $m = (x, a)$ if and only if $a = f_K(x)$; such a message is said to be *valid* under the key K. In this basic model, each key K is to be used to transmit only one message.

An opponent, denoted by O, may try to cheat Bob by causing him to accept a message that was not constructed and transmitted by Alice. If O introduces a message (x', a') into the channel before Alice sends a message to Bob, then we say that O is attempting to *impersonate* Alice. On the other hand, O might intercept a valid message (x, a) that was transmitted by Alice, and replace it with a message (x', a') where $x' \neq x$. This scenario is termed *substitution*. In each case, O is hoping that Bob will accept the bogus message (x', a').

Let P_I denote the *impersonation probability* of the authentication code, which denotes the maximum probability with which O can fool Bob in an impersonation attack. The following bound on P_I is well-known (see, for example, [8, Theorem 4.14]).

Theorem 1.1. *Suppose that $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{F})$ is an authentication code. Then $P_I \geq 1/|\mathcal{A}|$. Further, equality occurs if and only if*

$$|\{K \in \mathcal{K} : f_K(x) = a\}| = \frac{|\mathcal{K}|}{|\mathcal{A}|}$$

for every $x \in \mathcal{X}$ and for every $a \in \mathcal{A}$.

The next result is an immediate corollary of Theorem 1.1.

Corollary 1.2. *Suppose that $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{F})$ is an authentication code with $P_I = 1/|\mathcal{A}|$. Then $|\mathcal{K}| \geq |\mathcal{A}|$, and equality occurs if and only if*

$$|\{K \in \mathcal{K} : f_K(x) = a\}| = 1$$

for every $x \in \mathcal{X}$ and for every $a \in \mathcal{A}$.

[1] It is possible to consider codes where keys are not chosen equiprobably. Most of the results we state can be appropriately modified to apply to the more general setting. However, it is simpler and more convenient to restrict our attention to the simplified setting.

An authentication code in which $P_I = 1/|\mathcal{A}|$ and $|\mathcal{K}| = |\mathcal{A}|$ will be termed *optimal*.

There are also many results about *substitution probabilities* of authentication codes; however, we do not need to consider them in this paper.

Now we turn to unconditionally secure secrecy codes, the theory of which was developed by Shannon [6]. A *secrecy code* is a five-tuple $(\mathcal{X}, \mathcal{Y}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ such that the following conditions are satisfied:

1. $\mathcal{X}$ is a set of *plaintexts*,
2. $\mathcal{Y}$ is a set of *ciphertexts*,
3. $\mathcal{K}$ is a set of *keys*,
4. $\mathcal{D} = \{d_K : K \in \mathcal{K}\}$ is a set of *decryption functions* such that $d_K : \mathcal{Y} \to \mathcal{X}$ for every $K \in \mathcal{K}$,
5. $\mathcal{E} = \{e_K : K \in \mathcal{K}\}$ is a set of *encryption functions* such that $e_K : \mathcal{X} \to \mathcal{Y}$ for every $K \in \mathcal{K}$, and
6. $$d_K(e_K(x)) = x \tag{1}$$ for all $x \in \mathcal{X}$ and for all $K \in \mathcal{K}$.

Observe that (1) implies that e_K is injective for every $K \in \mathcal{K}$.

As before, a key $K \in \mathcal{K}$ is chosen uniformly at random by Alice and communicated to Bob over a secure channel. When Alice wants to send a message to Bob, she chooses a plaintext $x \in \mathcal{X}$ and computes the ciphertext $y = e_K(x)$, which is sent to Bob over an insecure channel. Bob decrypts the ciphertext y to the plaintext $x = d_K(y)$.

A secrecy code is said to provide *perfect secrecy* if $\mathsf{Pr}[x|y] = \mathsf{Pr}[x]$ for all plaintexts x and all ciphertexts y. That is, the *a priori* probability of plaintext x is the same as the *a posteriori* probability of x given that the ciphertext y is observed. We will assume that $\mathsf{Pr}[x] > 0$ for all x. In this case, we can apply Bayes' Theorem, which states that

$$\mathsf{Pr}[y|x] = \frac{\mathsf{Pr}[x|y] \times \mathsf{Pr}[y]}{\mathsf{Pr}[x]},$$

and it is easily seen that we have perfect secrecy if and only if $\mathsf{Pr}[y|x] = \mathsf{Pr}[y]$ for all plaintexts x and all ciphertexts y.

Shannon gave the following characterization of secrecy codes that provide perfect secrecy.

Theorem 1.3. *[6] A secrecy code $(\mathcal{X}, \mathcal{Y}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ provides perfect secrecy if and only if, for every $y \in \mathcal{Y}$, there exists a non-negative integer r_y such that*

$$|\{K \in \mathcal{K} : e_K(x) = y\}| = r_y$$

for every $x \in \mathcal{X}$.

Remark. It is easy to verify that $\mathsf{Pr}[y|x] = r_y/|\mathcal{K}|$ if we have perfect secrecy.

The following corollary is useful; for a detailed proof, see [8, Theorem 2.4]).

Corollary 1.4. *Suppose that a secrecy code $(\mathcal{X}, \mathcal{Y}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ provides perfect secrecy. Then $|\mathcal{K}| \geq |\mathcal{Y}| \geq |\mathcal{X}|$. Further, $|\mathcal{K}| = |\mathcal{Y}| = |\mathcal{X}|$ if and only if, for every $y \in \mathcal{Y}$ and every $x \in \mathcal{X}$, there exists a unique key $K \in \mathcal{K}$ such that $e_K(x) = y$ (i.e., if $r_y = 1$ for all $y \in \mathcal{Y}$).*

Protocol 2.1: Unconditionally Secure Chaffing-and-Winnowing Scheme Suppose $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ is a chaffing-and-winnowing scheme.

1. A secret key $K \in \mathcal{K}$ is chosen randomly by Alice and communicated to the receiver, Bob, over a secure channel. K is the decryption key.

2. Later, Alice wants to encrypt a plaintext $x \in \mathcal{X} = \{0, \ldots, n-1\}$ to send to Bob. Alice chooses an encryption function $e \in \mathcal{E}(K, x)$ uniformly at random. Then Alice computes $a_j = e(j)$ for all j, $0 \leq j \leq n-1$. The list of n ordered pairs,

$$y = ((0, a_0), \ldots, (n-1, a_{n-1})),$$

   is sent to Bob; y is the *ciphertext*.

3. Bob computes $b_j = f_K(j)$ for all j, $0 \leq j \leq n-1$. Bob decrypts y to the plaintext x if and only if $\{j : b_j = a_j\} = \{x\}$. (Because (2) holds, there will be exactly one ordered pair $m = (x, a)$ such that a is a valid authentication tag under the key K. This plaintext element x is defined to be the decryption of y.)

2 Unconditionally Secure Chaffing-and-Winowing

In this section, we develop necessary and sufficient conditions for a chaffing-and-winnowing scheme to achieve perfect secrecy. First, we give a careful description of the structure of a chaffing-and-winnowing scheme.

Let $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{F})$ be an authentication code, and for convenience, let $\mathcal{X} = \{0, \ldots, n-1\}$ be the set of possible *plaintexts*. A secret key $K \in \mathcal{K}$ will be used as a *decryption key* in the chaffing-and-winnowing scheme. Given any decryption key $K \in \mathcal{K}$ and any plaintext $x \in \mathcal{X}$, there will be a set $\mathcal{E}(K, x)$ of possible *encryption functions*. For each encryption function $e \in \mathcal{E}(K, x)$, $e : \mathcal{X} \to \mathcal{A}$.

In order for unambiguous decryption to be possible, we will require that the following property must hold for all $K \in \mathcal{K}$, for all $x \in \mathcal{X}$, and for all $e \in \mathcal{E}(K, x)$:

$$\{j : 0 \leq j \leq n-1 \text{ and } f_K(j) = e(j)\} = \{x\}. \tag{2}$$

We denote the set of all the encryption keys by $\mathcal{E}$. The set

$$\mathcal{E} = \bigcup_{K \in \mathcal{K}, x \in \mathcal{X}} \mathcal{E}(K, x).$$

We will assume, for simplicity, that all the sets $\mathcal{E}(K, x)$ have the same cardinality, which we denote by S.

A chaffing-and-winnowing scheme is denoted by a five-tuple $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$. The chaffing-and-winnowing schemes we study in this paper follow the general structure set out in Protocol 2.1. This is essentially the model for unconditionally secure chaffing-and-winnowing introduced in [4].

Before continuing, we make a few observations and remarks about Protocol 2.1.

Remarks.

1. As is common in the unconditionally secure setting, each key is to be used for only one encryption or decryption.

2. In the unconditionally secure setting, we generally consider plaintexts that are chosen from a finite set of possible plaintexts (e.g., the set $\{0, \ldots, n-1\}$ for some fixed integer n, or all bitstrings of some fixed length). In the computationally secure setting, it is more common to consider plaintexts of arbitrary length.

3. In the computationally secure setting, it is not necessary for a ciphertext to include all possible plaintexts. However, this is clearly required if we hope to attain perfect secrecy in the setting of unconditional security.

4. In a later section, we will consider schemes based on breaking a plaintext into smaller blocks and using multiple keys. This was suggested by Rivest [5] in his original paper, and it can be done in the unconditionally secure setting as well.

5. An interesting feature of chaffing-and-winnowing schemes is that the encryption key is chosen *after* the plaintext to be encrypted is chosen.

6. None of the schemes we construct in this paper are very effficient from the point of view of message expansion, key size, etc. This is because we are attempting to achieve a strong type of confidentiality with a tool that is deliberately chosen in order to *not* provide confidentiality (at least, it does not provide confidentiality when it is used in its intended manner).

Lemma 2.1. *Suppose a chaffing-and-winnowing scheme $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ satisfies (2). Let $K \in \mathcal{K}$ and suppose $x_1, x_2 \in \mathcal{X}$, $x_1 \neq x_2$. Then $\mathcal{E}(K, x_1) \cap \mathcal{E}(K, x_2) = \emptyset$.*

Proof. Suppose that $e \in \mathcal{E}(K, x_1) \cap \mathcal{E}(K, x_2)$. Because $e \in \mathcal{E}(K, x_1)$, property (2) implies that

$$\{j : 0 \leq j \leq n-1 \text{ and } f_K(j) = e(j)\} = \{x_1\}.$$

However, because $e \in \mathcal{E}(K, x_2)$, property (2) implies that

$$\{j : 0 \leq j \leq n-1 \text{ and } f_K(j) = e(j)\} = \{x_2\}.$$

It follows that $x_1 = x_2$, a contradiction. □

2.1 Security of the Scheme

Consider an observer O who sees a ciphertext y in the channel. We do not want O to be able to determine any information about the value of the plaintext x. We assume that O has complete information about the chaffing-and-winnowing scheme, but O does not know the values of K, x and e that are used. Also, O has unlimited computing power, since we are working in the model of unconditional security.

Remember that we are considering chaffing-and-winnowing schemes $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ in which $|\mathcal{E}(K, x)| = S$ for all $x \in \mathcal{X}$ and for all $K \in \mathcal{K}$. Fix an excryption function $e \in \mathcal{E}$. For $x \in \mathcal{X}$ and $e \in \mathcal{E}$, define $r_{e,x} = |\{K \in \mathcal{K} : e \in \mathcal{E}(K, x)\}|$. The following theorem gives a characterization of perfect secrecy in terms of the quantities $r_{e,x}$ (there is an obvious similarity between the following theorem and Theorem 1.3).

Theorem 2.2. *Suppose that $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ is a chaffing-and-winnowing scheme in which $|\mathcal{E}(K, x)| = S$ for all $x \in \mathcal{X}$ and for all $K \in \mathcal{K}$. Then the scheme achieves perfect secrecy if and only if $r_{e,x}$ is independent of x for every $e \in \mathcal{E}$.*

Proof. In this proof, we make use of the fact that there is an obvious correpondence between the encryption functions and ciphertexts defined by the following mapping:

$$e \mapsto y = ((0, e(0)), \ldots, (n-1, e(n-1))).$$

It is clear that the above-defined mapping is a bijection. In the rest of this proof, we will denote the encryption function corresponding to a ciphertext y by e_y.

Perfect secrecy is achieved for a chaffing-and-winnowing scheme if and only if $\mathsf{Pr}[y|x] = \mathsf{Pr}[y]$ for every $x \in \mathcal{X}$ and every $y \in \mathcal{Y}$. Equivalently, we have perfect secrecy if and only if $\mathsf{Pr}[y|x]$ is independent of x, for every $y \in \mathcal{Y}$.

For a ciphertext $y \in \mathcal{Y}$, let $e = e_y$ and consider the values $r_{e,x}$ $(x \in \mathcal{X})$ as defined above. Now, it is easy to see that

$$\mathsf{Pr}[y|x] = \frac{r_{e,x}}{S|\mathcal{K}|}.$$

Therefore we have perfect secrecy if and only if $r_{e,x}$ is independent of x. □

3 The Hanaoka *et al.* Construction

We describe the basic construction from [4], which we term the H^3WI scheme. Suppose that $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{F})$ is an authentication code. Define two keys K and K' to be *disjoint* if $f_K(x) \neq f_{K'}(x)$ for all x, $0 \leq x \leq n-1$. In the H^3WI construction, we require that there exist $|\mathcal{X}| = n$ mutually disjoint keys, which implies that $|\mathcal{A}| \geq |\mathcal{X}| = n$. The paper [4] suggests to use an optimal authentication code that is secure against impersonation. In view of Theorem 1.1 and Corollary 1.2, this is equivalent to saying that $|\mathcal{K}| = |\mathcal{A}| = |\mathcal{X}|$ and the n keys are mutually disjoint.

For convenience, suppose that $\mathcal{K} = \{K_0, \ldots, K_{n-1}\}$. We need to describe the sets of encryption functions $\mathcal{E}(K_t, x)$, $0 \leq t \leq n-1$, $0 \leq x \leq n-1$. Given t and x, let $\mathcal{P}(t, x)$ denote the set of all permutations π of $\{0, \ldots, n-1\}$ subject to the constraint that $\pi(x) = t$. For each permutation $\pi \in \mathcal{P}(t, x)$, there is an associated encryption function $e = e(\pi)$, which is defined as follows:

$$e(j) = f_{K_{\pi(j)}}(j)$$

for all j, $0 \leq j \leq n-1$. Then we define

$$\mathcal{E}(f_K, x) = \{e_\pi : \pi \in \mathcal{P}(t, x)\}.$$

The property (2) is clearly satisfied, so unambiguous decryption is achieved. We now use Theorem 2.2 to verify that the chaffing-and-winnowing scheme provides perfect secrecy. It is easy to see that $|\mathcal{E}| = n!$, since there is a one-to-one correspondence between the encryption keys and permutations of $\{0, \ldots, n-1\}$. For any $K \in \mathcal{K}$, the n sets $\mathcal{E}(K, x)$ $(x \in \mathcal{X})$ form a partition of $\mathcal{E}$ into n sets of size $(n-1)!$. Then it is easy to see, for any $x \in \mathcal{X}$, that the n sets $\mathcal{E}(K, x)$ $(K \in \mathcal{K})$ form a partition of $\mathcal{E}$ into n sets of size $(n-1)!$. Applying Theorem 2.2, it follows that the scheme $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ achieves perfect secrecy.

This chaffing-and-winnowing scheme can be modified so that the encryption keyspace is smaller. This reduces the number of random bits required by Alice to choose an encryption function. Given

t and x, we define $\mathcal{P}^*(t,x)$ to consist of only one permutation, namely the permutation π such that $\pi(j) = t - x + j \bmod n$ for all j, $0 \le j \le n-1$. Then let

$$\mathcal{E}^*(f_K, x) = \{e_\pi : \pi \in \mathcal{P}^*(t,x)\}$$

and

$$\mathcal{E}^* = \bigcup_{K \in \mathcal{K}, x \in \mathcal{X}} \mathcal{E}^*(K,x).$$

It is easy to prove the following facts about the chaffing-and-winnowing scheme $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}^*, \mathcal{F})$:

1. $|\mathcal{E}^*| = n$,
2. for any $K \in \mathcal{K}$, the n sets $\mathcal{E}^*(K,x)$ $(x \in \mathcal{X})$ form a partition of $\mathcal{E}^*$ into n sets of size one, and
3. for any $x \in \mathcal{X}$, the n sets $\mathcal{E}^*(K,x)$ $(K \in \mathcal{K})$ form a partition of $\mathcal{E}^*$ into n sets of size one.

Applying Theorem 2.2, it follows that the scheme $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}^*, \mathcal{F})$ achieves perfect secrecy.

Here are two examples to illustrate the scheme.

Example 3.1. Suppose that $f_{K_j}(i) = i - j \bmod n$ for all i and j. This authentication code is optimal (cf. Theorem 1.1). Then it is easy to see that the ciphertext is

$$y = ((0, x-t), (1, x-t), \ldots, (n-1, x-t)).$$

We illustrate with the case $n = 4$. First we present the four decryption functions (i.e., authentication functions) and then we present the encryption functions in each $\mathcal{E}^*(K_t, x)$. All encryption and decryption functions are written as 4-tuples.

K_j	f_{K_j}
K_0	$(0,1,2,3)$
K_1	$(3,0,1,2)$
K_2	$(2,3,0,1)$
K_3	$(1,2,3,0)$

t	$x=0$	$x=1$	$x=2$	$x=3$
0	$(0,0,0,0)$	$(1,1,1,1)$	$(2,2,2,2)$	$(3,3,3,3)$
1	$(3,3,3,3)$	$(0,0,0,0)$	$(1,1,1,1)$	$(2,2,2,2)$
2	$(2,2,2,2)$	$(3,3,3,3)$	$(0,0,0,0)$	$(1,1,1,1)$
3	$(1,1,1,1)$	$(2,2,2,2)$	$(3,3,3,3)$	$(0,0,0,0)$

Since $a_0 = \cdots = a_{n-1}$ in this scheme, it would be sufficient to simply transmit one copy of this common value, which we denote by a, to Bob. Then, Bob can decrypt the ciphertext by computing $x = a + t \bmod n$. Now, it is easy to see from Corollary 1.4 that the constituent authentication code already provides perfect secrecy. Therefore it could be argued plausibly that this is not a permissible chaffing-and-winnowing scheme.

Example 3.2. Suppose that $f_{K_j}(i) = j \bmod n$ for all i and j. This is also an optimal authentication code. It is easy to see that the ciphertext is

$$y = ((0, t-x), (1, t-x+1), \ldots, (n-1, t-x-1)).$$

We illustrate the chaffing-and-winnowing scheme in the case $n = 4$, as we did in the previous example:

K_j	f_{K_j}
K_0	$(0,0,0,0)$
K_1	$(1,1,1,1)$
K_2	$(2,2,2,2)$
K_3	$(3,3,3,3)$

t	$x=0$	$x=1$	$x=2$	$x=3$
0	$(0,1,2,3)$	$(3,0,1,2)$	$(2,3,0,1)$	$(1,2,3,0)$
1	$(1,2,3,0)$	$(0,1,2,3)$	$(3,0,1,2)$	$(2,3,0,1)$
2	$(2,3,0,1)$	$(1,2,3,0)$	$(0,1,2,3)$	$(3,0,1,2)$
3	$(3,0,1,2)$	$(2,3,0,1)$	$(1,2,3,0)$	$(0,1,2,3)$

In this chaffing-and-winnowing scheme, Bob can decrypt y to be the unique value x such that $a_x = t$. This authentication code does not provide perfect secrecy. However, it is still the case that $|\mathcal{A}| = |\mathcal{X}| = n$. As mentioned earlier, in a strict prohibition of encryption schemes, it might not be permitted to have $|\mathcal{A}| \geq |\mathcal{X}|$.

These two examples illustrate that the authentication codes used in the construction from [4] may or may not achieve perfect secrecy. However, in both cases, the entropy of the authentication tag is equal to the entropy of the plaintext space. As we have noted already, it is possible to achieve perfect secrecy solely from an authentication code in this situation.

We are interested in finding constructions where the underlying authentication codes cannot possibly achieve perfect secrecy. This is pursued in the next section.

4 A New Chaffing-and-winnowing Scheme

We describe a chaffing-and-winnowing scheme that we denote by $\mathsf{CW}(n)$, where $n \geq 2$ is an integer. For our construction of $\mathsf{CW}(n)$, we use an authentication code $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{F})$ in which $\mathcal{X} = \{0, \ldots, n-1\}$, $\mathcal{K} = \{0,1\}^n$ and $\mathcal{A} = \{0,1\}$. For every key $K = (\kappa_0, \ldots, \kappa_{n-1}) \in \{0,1\}^n$, we have an authentication function f_K where $f_K(i) = \kappa_i$ for $0 \leq i \leq n-1$. Note that each key basically specifies the list of authenticators of all n possible plaintexts.

It is not hard to see that this authentication code has impersonation probability $P_I = 1/2$. The value of P_I meets the bound given in Theorem 1.1. Of course, $P_I = 1/2$ does not provide much security against impersonation.

For each $K = (\kappa_0, \ldots, \kappa_{n-1}) \in \mathcal{K}$ and each $x \in \mathcal{X}$, we define $\mathcal{E}(K, x) = \{e_{K,x}\}$, where

$$e_{K,x}(j) = \begin{cases} \kappa_j & \text{if } j = x \\ 1 - \kappa_j & \text{if } j \neq x. \end{cases} \tag{3}$$

That is, each set $\mathcal{E}(K, x)$ consists of one encryption function, denoted $e_{K,x}$. The authentication function f_K and the encryption function $e_{K,x}$ are "complements" of each other, except for the input x, where they agree.

The resulting chaffing-and-winnowing scheme is denoted $\mathsf{CW}(n)$. The following easily verified properties show that $\mathsf{CW}(n)$ provides perfect secrecy:

1. $|\mathcal{E}| = 2^n$,
2. for any $K \in \mathcal{K}$, the n sets $\mathcal{E}(K, x)$ $(x \in \mathcal{X})$ are disjoint, and
3. for any $x \in \mathcal{X}$, the 2^n sets $\mathcal{E}(K, x)$ $(K \in \mathcal{K})$ form a partition of $\mathcal{E}$ into 2^n sets of size one.

Applying Theorem 2.2, it follows that the chaffing-and-winnowing scheme $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ achieves perfect secrecy.

4.1 A Modified Scheme

It is possible to prove that a modified version of $\mathsf{CW}(n)$ remains unconditionally secure. In the modified scheme, we restrict the set of decryption keys to be

$$\mathcal{K}_E = \left\{ K = (\kappa_0, \ldots, \kappa_{n-1}) \in \{0,1\}^n, \sum_{i=0}^{n-1} \kappa_i = 0 \bmod 2 \right\}.$$

We are reducing the number of decryption keys by a factor of two by only using keys with even hamming weight.

We denote the sets of encryption functions in the modified scheme by $\mathcal{E}_E$ and $\mathcal{E}_E(K,x)$. The modified scheme is denoted by $\mathsf{CW}_E(n)$.

The following properties show that $\mathsf{CW}_E(n)$ provides perfect secrecy:

1. $|\mathcal{E}_E| = 2^{n-1}$,
2. for any $K \in \mathcal{K}$, the n sets $\mathcal{E}_E(K,x)$ are disjoint, and
3. for any $x \in \mathcal{X}$, the 2^{n-1} sets $\mathcal{E}_E(K,x)$ $(K \in \mathcal{K}_E)$ form a partition of $\mathcal{E}$ into 2^{n-1} sets of size one.

Applying Theorem 2.2, it follows that the chaffing-and-winnowingscheme $(\mathcal{X}, \mathcal{A}, \mathcal{K}_E, \mathcal{E}_E, \mathcal{F})$ achieves perfect secrecy.

Example 4.1. In the case $n = 4$, we present the sets $\mathcal{E}_E(K,x)$ in the scheme $(\mathcal{X}, \mathcal{A}, \mathcal{K}_E, \mathcal{E}_E, \mathcal{F})$:

K	$x=0$	$x=1$	$x=2$	$x=3$
$(0,0,0,0)$	$(0,1,1,1)$	$(1,0,1,1)$	$(1,1,0,1)$	$(1,1,1,0)$
$(0,0,1,1)$	$(0,1,0,0)$	$(1,0,0,0)$	$(1,1,1,0)$	$(1,1,0,1)$
$(0,1,0,1)$	$(0,0,1,0)$	$(1,1,1,0)$	$(1,0,0,0)$	$(1,0,1,1)$
$(0,1,1,0)$	$(0,0,0,1)$	$(1,1,0,1)$	$(1,0,1,1)$	$(1,0,0,0)$
$(1,0,0,1)$	$(1,1,1,0)$	$(0,0,1,0)$	$(0,1,0,0)$	$(0,1,1,1)$
$(1,0,1,0)$	$(1,1,0,1)$	$(0,0,0,1)$	$(0,1,1,1)$	$(0,1,0,0)$
$(1,1,0,0)$	$(1,0,1,1)$	$(0,1,1,1)$	$(0,0,0,1)$	$(0,0,1,0)$
$(1,1,1,1)$	$(1,0,0,0)$	$(0,1,0,0)$	$(0,0,1,0)$	$(0,0,0,1)$

We will measure the efficiency of a chaffing-and-winnowing scheme in terms of the length of a ciphertext and the length of a key (in bits). Suppose that $|\mathcal{X}| = n = 2^k$. Then there are 2^{n-1} possible decryption keys in the improved scheme, so the length of a decryption key is $n-1 = 2^k - 1$ bits. For exammple, a decryption key $K = (\kappa_0, \ldots, \kappa_{n-1}) \in \mathcal{K}_E$ is determined uniquely from the $n-1$ values $\kappa_0, \ldots, \kappa_{n-2}$, because

$$\kappa_{n-1} = \sum_{i=0}^{n-2} \kappa_i \bmod 2.$$

A ciphertext has the form $y = ((0,a_0), \ldots, (n-1, a_{n-1}))$. Since the first co-ordinates are $0, \ldots, n-1$, in that order, we can eliminate them from the ciphertext if desired, and just transmit the vector of authenticators, namely, $\vec{a} = (a_0, \ldots, a_{n-1})$. Under this assumption, the length of the ciphertext is $n = 2^k$.

Summarizing, we have shown the following.

Theorem 4.1. *For any integer $k \geq 1$, the scheme $\mathsf{CW}_E(2^k)$ is an unconditionally secure chaffing-and-winnowing scheme for k-bit plaintexts, based on 1-bit authenticators, in which a decryption key consists of $2^k - 1$ bits and a ciphertext consists of 2^k bits.*

4.2 Optimality of the Schemes $\mathsf{CW}_E(n)$

We now show that the schemes $\mathsf{CW}_E(n)$ are optimal in the set of all chaffing-and-winnowing schemes that have 1-bit authenticators for a plaintext space of cardinality n. To show this, we prove that there must be at least 2^{n-1} keys in such a scheme.

The optimality proof is based on the observation that the encryption function sets $\mathcal{E}(K, x)$ and $\mathcal{E}_E(K, x)$ used in the schemes $\mathsf{CW}(n)$ and $\mathsf{CW}_E(n)$ (respectively) provide the only way to ensure unambiguous decyption. In fact, this follows immediately from equation (2). Therefore, we have the following.

Lemma 4.2. *Suppose $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ is any chaffing-and-winnowing scheme in which $|\mathcal{X}| = 2$. Then, for every $x \in \mathcal{X}$ and every $K \in \mathcal{K}$, $\mathcal{E}(K, x) = \{e_{K,x}\}$, where $e_{K,x}$ is as defined in equation (3).*

Lemma 4.3. *Suppose $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ is any chaffing-and-winnowing scheme in which $|\mathcal{X}| = 2$. Suppose that $K = (\kappa_0, \ldots, \kappa_{n-1}) \in \mathcal{K}$, $K' = (\kappa'_0, \ldots, \kappa'_{n-1})$ and $\mathsf{dist}(K, K') = 2$, where $\mathsf{dist}(\cdot, \cdot)$ denotes the hamming distance between two vectors. Then $K' \in \mathcal{K}$.*

Proof. Suppose that $\{j : \kappa_j \neq \kappa'_j\} = \{x, x'\}$. We have that $\mathcal{E}(K, x) = \{e\}$, where

$$e(j) = \begin{cases} \kappa_j & \text{if } j = x \\ 1 - \kappa_j & \text{if } j \neq x. \end{cases}$$

Now, $\{e\} = \mathcal{E}(K'', x'\}$ for some $K'' = (\kappa''_0, \ldots, \kappa''_{n-1}) \in \mathcal{K}$. However, it must be the case that

$$e(j) = \begin{cases} \kappa''_j & \text{if } j = x' \\ 1 - \kappa''_j & \text{if } j \neq x'. \end{cases}$$

From this, it is immediate that $K'' = K'$, and therefore $K' \in \mathcal{K}$. □

Theorem 4.4. *Suppose $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ is any chaffing-and-winnowing scheme in which $|\mathcal{X}| = 2$. Then $\mathcal{K}$ must consist of all the binary n-tuples of even weight, all the binary n-tuples of odd weight, or all the binary n-tuples.*

Proof. Let $K \in \mathcal{K}$ be any key. Applying Lemma 4.3, it follows that every binary n-tuple that has hamming distance two from K is also a decryption key. From this, it is easily seen that $\mathcal{K}$ contains all the binary n-tuples that have even hamming distance from K. The desired result follows. □

Corollary 4.5. *Suppose $(\mathcal{X}, \mathcal{A}, \mathcal{K}, \mathcal{E}, \mathcal{F})$ is any chaffing-and-winnowing scheme in which $|\mathcal{X}| = 2$. Then $|\mathcal{K}| \geq 2^{n-1}$.*

4.3 A Hybrid Scheme

Suppose we have an ℓ-bit plaintext, where $\ell = rk$, and we break it into r blocks, each of which contains k bits. Each k-bit block is then encrypted using a scheme $\mathsf{CW}_E(2^k)$. In total, we have r independent schemes $\mathsf{CW}_E(2^k)$, each of which has an independently chosen key. Each possible ℓ-bit plaintext receives an r-bit authenticator, which is the concatenation of the 1-bit authenticators of each of the r blocks in the plaintext. This hybrid scheme, which will be denoted by $\mathsf{HCW}(r, k)$, has the following properties.

Theorem 4.6. *For integers $k, r \geq 1$, the scheme $\mathsf{HCW}(r, k)$ is an unconditionally secure chaffing-and-winnowing scheme for rk-bit plaintexts, based on r-bit authenticators, in which a decryption key consists of $r(2^k - 1)$ bits and a ciphertext consists of $r\,2^k$ bits.*

Theorem 4.6 illustrates a trade-off between authenticator size and efficiency. The parameter k denotes the ratio between the length of a plaintext and the length of an authenticator in the scheme $\mathsf{HCW}(r, k)$. The situation when $k = 1$ is problematic because the underlying authentication schemes might already provide perfect secrecy (a plaintext and authentication tag have the same length when $k = 1$). Clearly, as k is increased, the system becomes less and less efficient, so the "best value" of k to use is the smallest one that would be permitted by law.[2]

For example, when we set $k = 1$, our scheme $\mathsf{HCW}(r, 1)$ has r-bit plaintexts and r-bit authenticators, an r-bit decryption key and a $2r$-bit ciphertext. This is a two-fold message expansion, as compared to the classical one-time pad or any other optimal secrecy code that provides perfect secrecy (see Corollary 1.4). For the next case, $k = 2$, we have $2r$-bit plaintexts and r-bit authenticators, a $3r$-bit decryption key and a $4r$-bit ciphertext. For larger values of k, the situation becomes progressively worse.

5 Conclusion

We do not claim that our new chaffing-and-winnowing scheme is "practical." The interesting contribution of our paper is that we can still obtain perfect secrecy even when we are forced to use particularly ill-suited tools, namely, authentication codes with very short authentication tags.

It would be of interest to develop bounds (i.e., necessary conditions) on the parameters of unconditionally secure chaffing-and-winnowing schemes. We have shown that our schemes $\mathsf{CW}_E(n)$ are optimal in the set of all chaffing-and-winnowing schemes that have 1-bit authenticators and k-bit plaintexts. We ask if there are more efficient "r-bit schemes" than the schemes $\mathsf{HCW}(r, k)$ that we have constructed.

References

[1] M. Bellare and A. Boldyreva. The security of chaffing and winnowing. *Lecture Notes in Computer Science* **1976** (2000), 517–530 (ASIACRYPT 2000).

[2] R. Clayton and G. Danezis. Chaffinch: confidentiality in the face of legal threats. *Lecture Notes in Computer Science* **2578** (2003), 70–86 (Information Hiding 2002).

[2] Of course, encryption is currently legal, so this discussion is purely hypothetical.

[3] E.N. Gilbert, F.J. MacWilliams and N.J.A. Sloane. Codes which detect deception. *Bell System Tech. J.* **53** (1974), 405–424.

[4] G. Hanaoka, Y. Hanaoka, M. Hagiwara, H. Watanabe and H. Imai. Unconditionally secure chaffing-and-winnowing: a relationship between encryption and authentication. *Lecture Notes in Computer Science* **3857** (2006), 154–162 (AAECC-16).

[5] R.L. Rivest. Chaffing and winnowing: confidentiality without encryption. *CryptoBytes* **4-1** (1998), 12–17.

[6] C.E. Shannon. Communication theory of secrecy systems. *Bell System Tech. J.* **28** (1949), 656–715.

[7] G.J. Simmons. A survey of information authentication. In "Contemporary cryptology", IEEE Press, New York, 1992, pp. 379–419.

[8] D.R. Stinson. *Cryptography Theory and Practice, Third Edition*, Chapman & Hall/CRC, 2006.