

# Verifying a shuffle in zero-knowledge

Chris Patton

February 11, 2016

## 1 Introduction

A mixnet (sometimes called a “mix-network”) is a service designed to facilitate anonymous communications over the internet [1]. Each of a set of users  $\mathcal{Q} = \{P_1, \dots, P_n\}$  wish to send a message to a user in  $\mathcal{Q}$ . Each encrypts their message  $M$  and the identity of the intended recipient  $P_i$  under the public key of the mixnet server (hereafter referred to as the “mix”) and transmits the ciphertext. The mix server waits until it receives  $m \leq n$  ciphertexts from  $m$  distinct users. It then decrypts each message and destination pair under its secret key and transmits each message to its intended destination. Crucially, the mix transmits the messages in a random order so as to hide the correspondence between senders and recipients. As long as the mix is trusted and the user’s ciphertexts are faithfully transmitted to the mix, this simple approach suffices to preserve anonymity of its participants.

Normally we think of the mixnet being composed of a cascade of mix servers, each decrypting, shuffling, then transmitting to the next mix in the cascade until finally the last server outputs the users’ messages. This is meant to distribute trust so that as long as the messages pass through at least one honest mix, anonymity is achieved. Still, the trust model here is rather strong. For example, the mix could send a fake message on behalf of one of the senders, or redirect it to a different recipient. We consider the problem of publically verifying that the outputs are indeed a permutation of the inputs. Moreover, we require that the correspondence not be revealed. Andrew Neff describes in [6] a method for proving a shuffle in zero-knowledge.

Neff’s solution is based on the ElGamal public key encryption scheme, which is known to be semantically secure assuming the discrete log problem (decisional Diffie-Hellman, or DDH) is hard [2]. Based on this assumption, Neff proves his protocol achieves a property known as honest verifier zero-knowledge (HVZK) [5]. Informally speaking, this means that as long as the verifier faithfully executes the protocol, the mix can prove the outputs are a permutation of the inputs without revealing the permutation.

I formulate and define security for the problem and present Neff’s protocol. For simplicity, I limit the discussion to a mixnet with a single mix server. In this scenario, users encrypt their messages under the public key of the mix. The mix waits until it receives a full batch of ciphertexts, decrypts them under its private key, and shuffles the output. In section 2, I specify ElGamal encryption, describe the DDH assumption, and

sketch the HVZK notion. In section 3, I present the protocol and sketch its security. I implement the general  $k$ -shuffle as described in this section in the Go programming language.<sup>1</sup>

## 2 Preliminaries

If  $X$  is a finite set, let  $x \leftarrow X$  denote uniformly sampling an element  $x$  from the set. Let  $y \leftarrow A(x)$  denote running  $A$  on input  $x$  and assigning its output to  $y$ . Let  $y \leftarrow A(x; r)$  denote the execution of a probabilistic algorithm with the sequence of coins  $r \in \{0, 1\}^\infty$ . Let  $y \leftarrow A(x)$  denote choosing  $r \leftarrow \{0, 1\}^\infty$  and executing  $y \leftarrow A(x; r)$ . Let  $[i..j]$  where  $i \leq j$  denote the set of integers from  $i$  to  $j$  inclusive. Let  $y \leftarrow \langle P, V \rangle(x)$  denote the probabilistic execution of interactive Turing machines  $P$  and  $V$  on common input  $x$ . When the protocol finishes, the output of  $V$  is assigned to  $y$ .  $x$  is called the common input of  $P$  and  $V$ . Finally a function  $\Delta(n)$  is negligible if for all positive polynomials  $p(\cdot)$  and sufficiently large  $n$ ,  $\Delta(n) < \frac{1}{p(n)}$ .

### 2.1 ElGamal encryption

Fix a prime number  $p$  and let  $\mathbb{Z}_p$  denote the modular ring of order  $p$ . In what follows, we assume arithmetic is performed in the modular ring  $\mathbb{Z}_p$ . Let  $g \in \mathbb{Z}_p$  such that  $g^q = 1$  where  $q$  is prime. Then  $\langle g \rangle = \mathbb{Z}_q^*$  is a subgroup of  $\mathbb{Z}_p^*$  and  $q|(p-1)$ . We call  $(p, q, g)$  the *public parameters*. The *secret-key* is chosen uniformly from the set of powers of  $g$  mod  $p$ : let  $k \leftarrow [1..q-1]$ . Let  $K = g^k$  be the *public-key*. Key generation is denoted  $(K, k) \leftarrow \mathcal{G}(p, q, g)$ . Encryption and decryption operate on group elements.<sup>2</sup> Encryption of  $M \in \mathbb{Z}_p^*$  proceeds as follows: Let  $r \leftarrow [1..q-1]$ ,  $R = g^r$ ,  $S = K^r$ , and  $C = M \cdot S$ . Output the tuple  $(C, R)$ . This is denoted  $(C, R) \leftarrow \mathcal{E}_K(M)$ . Decryption of a ciphertext  $(C, R) \in \mathbb{Z}_p^* \times \mathbb{Z}_p^*$  proceeds as follows: let  $S = R^k$  and  $M = C \cdot S^{-1}$ . Output  $M$ . This is denoted  $M \leftarrow \mathcal{D}_k(C, R)$ .

ElGamal is known to be *semantically secure* under the decisional Diffie-Hellman assumption [2, 4]. Informally speaking, this means that if given  $K$ , it is infeasible for any adversary to determine  $k$ , then given  $(C, R)$ , it is infeasible for any adversary to learn anything meaningful about  $M$  (nor  $m$ ). As we're not interested in the security of ElGamal from a privacy standpoint, we won't formalize semantic security. However, we define the DDH assumption in the next subsection.

The mixnet operates as follows: let  $M_1, \dots, M_n \in \mathbb{Z}_p^*$  be a sequence of plaintexts. Let  $(C_1, R_1), \dots, (C_n, R_n) \in \mathbb{Z}_p^* \times \mathbb{Z}_p^*$  be a sequence of corresponding ElGamal ciphertexts encrypted under the public-key  $K$ . The mix server, who possesses the secret-key  $k$ , takes as input the ciphertexts and outputs a sequence  $M_{\pi(1)}, \dots, M_{\pi(n)}$  where  $\pi$  is a permutation on  $[1..n]$ . We remark that there is a natural extension of this scheme for mixnets with a cascade of servers, each possessing its own decryption key [7]. Indeed,

<sup>1</sup> Freely available at [github.com/cjpatton/shuffle](https://github.com/cjpatton/shuffle).

<sup>2</sup> Public-key encryption schemes are typically defined on bit strings. It is easy to messages as elements of  $\mathbb{Z}_p^*$ , but the number of messages that can be encoded depends on the modulus  $p$ .

Camenisch and Mityagin use the Neff proof in order to verify the output of each mix in the cascade. If a dishonest mix is detected, the protocol is run without without using it.

## 2.2 Decisional Diffie-Hellman

Let  $(p, q, g)$  be the public parameters of the ElGamal cryptosystem. The game the adversary  $D$  plays is defined as follows: let  $a, b, r \leftarrow [1..q-1]$  and let  $A = g^a$ ,  $B = g^b$ ,  $Y_0 = g^r$ , and  $Y_1 = g^{ab}$ . Choose a random bit  $b \leftarrow \{0, 1\}$ . On input  $(A, B, Y_b)$ , the adversary outputs a bit  $b' \in \{0, 1\}$  and wins if  $b = b'$ . The advantage of  $D$  is defined as

$$\Delta(q) = |\Pr[D(A, B, Y_1) = 1] - \Pr[D(A, B, Y_0) = 1]|$$

The DDH assumption is that, for every polynomial-time adversary  $D$ , the function  $\Delta(q)$  is negligible.

## 2.3 Honest verifier zero-knowledge

Zero-knowledge is defined in [5] with respect to the class of formal languages with interactive proof systems. A language  $L$  has an interactive proof system if there exists a pair of probabilistic algorithms  $(P, V)$  such that for all  $x \in L$ ,  $\Pr[\langle P, V \rangle(x) = 1] = 1$ , and for all  $x \notin L$ ,  $\Pr[\langle P, V \rangle(x) = 1] \leq 1/2$ . The class of languages with interactive proof systems is denoted **IP**.

Let  $L \in \mathbf{IP}$  exhibited by  $(P, V)$ . We define security with respect to a game associated to adversary  $D$ , verifier  $V^*$ , and simulator  $M^*$ . Suppose that  $L$  is a finite set<sup>3</sup> and let  $x \leftarrow L$ . Let  $y_1 \leftarrow \langle P, V^* \rangle(x)$  and  $y_0 \leftarrow M^*(x)$  choose a bit  $b \leftarrow \{0, 1\}$ . On input  $(x, y_b)$ , adversary  $D$  outputs a bit  $b'$  and wins if  $b = b'$ . Let  $k = |L|$  and define the advantage of  $D$  as

$$\Delta(k) = |\Pr[D(x, y_1) = 1] - \Pr[D(x, y_0) = 1]|$$

We say that  $(P, V)$  is *computational zero-knowledge* (or just *zero-knowledge*) if for every probabilistic polynomial-time adversary  $D$  and for every probabilistic polynomial-time verifier  $V^*$ , there exists probabilistic polynomial-time simulator  $M^*$  such that  $\Delta(k)$  is negligible. The security notion achieved by Neff's protocol is somewhat weaker: a proof system  $(P, V)$  is *honest verifier zero-knowledge* (HVZK) if for every probabilistic polynomial-time adversary  $D$ , there exists a probabilistic polynomial-time simulator  $M^*$  such that

$$|\Pr[D(x, \langle P, V \rangle(x)) = 1] - \Pr[D(x, M^*(x)) = 1]|$$

is a negligible function of  $k$ .

## 3 The Neff shuffle

We now give the syntax of an interactive proof system for mixnets. The prover  $P$  (nominally the mix server) possesses the secret-key  $k$  and the verifier  $V$  possesses the public-key

<sup>3</sup>This is without loss of generality, since we consider polynomial-time adversaries.

$K$ . The common input of  $P$  and  $V$  is all of the messages passed through the mix server. The mix takes as input a sequence of ciphertexts  $(C_1, R_1), \dots, (C_n, R_n) \in \mathbb{Z}_p^* \times \mathbb{Z}_p^*$  and outputs a sequence of plaintexts  $M_1, \dots, M_n \in \mathbb{Z}_p^*$ . Let  $x \in \{0, 1\}^*$  denote the encoding of these messages. We say  $x \in L$  if and only if there exists a bijection  $\pi : [1..n] \rightarrow [1..n]$  such that for every  $i \in [1..n]$ ,  $M_{\pi(i)} = \mathcal{D}_k(C_i, R_i)$ .

A trivial proof system for  $L$  is one in which the prover transmits to the verifier its secret-key and the verifier decrypts each of the ciphertexts. Of course, this completely exposes  $\pi$ . Our goal is to specify a proof system  $(P, V)$  in which  $P$  convinces  $V$  of the existence of  $\pi$  without revealing the permutation or the secret key. A proof system that is HVZK is sufficient to achieve this goal, assuming the DDH problem is hard. Such a proof system is given by the general  $k$ -shuffle protocol defined in section 3.3. It is built upon two simpler protocols, which we define in sections 3.1 and 3.2.

### 3.1 The basic protocol and its security

Not useful yet ...

### 3.2 Simple $k$ -shuffle

Not useful yet ...

### 3.3 General $k$ -shuffle

This is what we actually implement.

## References

- [1] David Chaum, 1981. "Untraceable electronic mail, return addresses, and digital pseudonyms." <http://freehaven.net/anonbib/cache/chaum-mix.pdf>
- [2] Taher Elgamal, "A public key cryptosystem and a signature scheme based on discrete logarithms." Appeared in *CRYPTO'84*.
- [3] D. Chaum and T.P. Pedersen. "Wallet databases with observers." Appeared in *CRYPTO'92*.
- [4] Yiannis Tsiounis and Moti Yung 1998. "On the security of ElGamal based encryption." Appeared in *PKC'98*.
- [5] Oded Goldreich, 2001. *Foundations of Cryptography*.
- [6] C. Andrew Neff, 2001. "A verifiable secret shuffle and its application to e-voting." <http://freehaven.net/anonbib/cache/shuffle:ccs01.pdf>
- [7] Jan Camenisch and Anton Mityagin, 2006. "Mix-network with stronger security."