Deniable Encryption from the McEliece Assumptions

Bernardo M. David*, Anderson C. A. Nascimento†

March 22, 2011

Abstract

The first construction of sender-deniable encryption with negligible detection probability and a single encryption scheme was introduced by Dürmuth and Freeman. Their construction is based on *samplable public key encryption*, which mainly differs from general semantically secure cryptosystems in that it allows the owner of a secret key to *extract* the randomness used in generating a given ciphertext. However, it was left as an open problem to construct samplable public key encryption based on assumptions other then the hardness of factoring.

We show that a semantically secure variant of the McEliece cryptosystem is samplable. Being based on the McEliece assumptions, our scheme is the first not to rely on the hardness of factoring and the first to achieve post quantum security. Our construction takes advantage of specific properties of a semantically secure variant of the McEliece cryptosystem introduced by Nojima *et al.*, which allows us to construct a randomness extraction algorithm.

1 Introduction

Deniable encryption was introduced by Canetti *et. al.* in [2]. This primitive allows the sender or the receiver of a ciphertext to "fake" the encrypted message in order to reveal an arbitrary plaintext message without being detected by a coercing adversary. Such a primitive is interesting for adaptively secure multi-party computation and applications such as electronic voting, where the committing property of most encryption schemes may be problematic. However, the constructions of [2] did not offer negligible detection, allowing an adversary to detect whether a party is producing "fake" messages with probability $1/O(n)$, where n is the security parameter.

This problem was recently solved in [3], which presents a general construction of sender-deniable encryption with negligible detection. The authors show that this construction is implied by semantically secure samplable public key cryptosystems, which allow the holder a secret key to extract the randomness used in encrypting a given ciphertext. Nevertheless, the only examples of such an encryption scheme presented in [3] are based on assumptions related to the hardness of factoring. It was left as an open problem to construct deniable encryption schemes from other assumptions, such as the hardness of Learning With Errors (LWE).

Our Contribution: We show that a semantically secure variant of the McEliece cryptosystem [6] is samplable, thus obtaining a sender-deniable encryption scheme with negligible probability based on the McEliece assumptions. Our scheme is the first based on assumptions other than the hardness of factoring and is also a good candidate for the post quantum scenario.

2 Preliminaries

In this section we define the notation used throughout this work and provide the formal description Samplable Public Key Encryption, the McEliece cryptosystem and the security assumptions on which it relies.

Hereupon, we will denote by $x \in_R D$ an uniformly random choice of element x over its domain D; by $\oplus$ a bit-wise exclusive OR of strings; and by $a \mid b$ the concatenation of string a with string b. For a set X an a distribution Δ on X, we denote by $x \leftarrow \Delta_X$ an element of X sampled according to the distribution Δ. All logarithms are to the base 2.

If X and Y are families of distributions indexed by a security parameter λ, we use $X \overset{s}{\approx} Y$ to mean the distributions X and Y are *statistically close*, *i.e.*, for all polynomials p and sufficiently large λ, we

*Email: bernardo.david@redes.unb.br, Department of Electrical Engineering, University of Brasilia. Campus Universitario Darcy Ribeiro,Brasilia, CEP: 70910-900, Brazil

†Email: andclay@ene.unb.br, Department of Electrical Engineering, University of Brasilia. Campus Universitario Darcy Ribeiro,Brasilia, CEP: 70910-900, Brazil

have $\sum_x |Pr[X = x] - Pr[Y = x]| < 1$. Two sequences $X_n, n \in \mathbb{N}$ and $Y_n, n \in \mathbb{N}$ of random variables are said to be *computationally indistinguishable*, denoted by $X \stackrel{c}{=} Y$, if for every non-uniform probabilistic polynomial-time distinguisher D there exists a negligible function $\epsilon(\cdot)$ such that for every $n \in \mathbb{N}$,

$$| Pr[D(X_n) = 1] - Pr[D(Y_n) = 1] | < \epsilon(n)$$

2.1 McEliece Public-Key Cryptosystem

The following definition has been taken from [6]. The McEliece cryptosystem [5] consists of a triplet of probabilistic algorithms $ME = (Gen_{ME}, Enc_{ME}, Dec_{ME})$ and $M = \{0,1\}^k$.

- Key generation algorithm: The PPT key generation algorithm Gen_{ME} works as follows:
  1. Generate a $k \times n$ generator matrix $\mathbf{G}'$ of a binary Goppa code, where we assume that there is an efficient error-correction algorithm `Correct` which can always correct up to t errors.
  2. Generate a $k \times k$ random non-singular matrix $\mathbf{S}$.
  3. Generate a $n \times n$ random permutation matrix $\mathbf{P}$.
  4. Set $\mathbf{G} = \mathbf{SG'P}$, and outputs $pk = (\mathbf{G}, t)$ and $sk = (\mathbf{S}, \mathbf{G}', \mathbf{P})$.
- The encryption algorithm: The PPT encryption algorithm Enc_{ME} takes a plaintext $m \in \{0,1\}^k$ and the public-key pk as input and outputs ciphertext $c = m\mathbf{G} \oplus e$, where $e \in \{0,1\}^n$ is a random vector of hamming weight t.
- The decryption algorithm: The polynomial-time algorithm Dec_{ME} works as follows:
  1. Compute $c\mathbf{P}^{-1} = ((m\mathbf{S})\mathbf{G}' \oplus e\mathbf{P}^{-1})$, where $\mathbf{P}^{-1}$ denotes the inverse matrix of $\mathbf{P}$.
  2. Compute $m\mathbf{S} = \texttt{Correct}(c\mathbf{P}^{-1})$.
  3. Output $m = (m\mathbf{S})\mathbf{S}^{-1}$.

2.2 Semantically Secure McEliece Public-Key Cryptosystem

In our application we use a variant of the original McEliece PKC that was proved semantically secure [6]. We are interested in encrypting single bit messages, instead of strings.

In the semantically secure version of McEliece PKC, we pad the message $m \in \{0,1\}$ with a random string $r \in_R \{0,1\}^{k-1}$, obtaining $(r \mid m)\mathbf{G} \oplus e$ as the desired ciphertext.

2.3 The McEliece Assumptions

The security of the McEliece cryptosystem is based on the assumptions introduced and briefly discussed in this subsection. The first assumption is that there is no algorithm which can distinguish the scrambled generating matrix of a binary Goppa code (matrix G generated by the algorithm Gen_{ME} described before) from a random matrix of the same size.

Definition 1. *There is no PPT algorithm which can distinguish the public-key matrix P of the McEliece cryptosystem from a random matrix of the same size with non-negligible probability.*

The hardness of decoding random linear codes is the second assumption on which the McEliece PKC and our protocol are based. This problem was proved to be NP-complete [1] and currently the best algorithm to solve this problem still runs in exponential time. This problem was proved to be equivalent to the Learning Parity with Noise (LPN) problem [7], a fact that is central to our scheme's security.

Definition 2. *The Syndrome Decoding Problem is hard for every PPT algorithm.*

Notice that a recent attack introduced in [4] breaks the McEliece assumptions for case where a high rate Goppa code is used. Therefore, a low rate Goppa code should be used in the McEliece cryptosystem instantiation in order to guarantee that the assumptions above hold.

3 Achieving Deniable Encryption

In this section, we formally define samplable public key encryption and show that the semantically secure variant of the McEliece cryptosystem [6] fits this definition.

3.1 Sampleable Public Key Encryption

Samplable public key encryption was defined in [3]. A samplable public key cryptosystem is basically a bit encryption scheme which allows the holder of the secret key to recover the randomness used in the encryption. Although the recovered randomness is not required to be the exact randomness used to encrypt the message, it must be indistinguishable from the real randomness. We now present a formal definition based on [3].

Definition 3. *Let $\mathcal{PK}$, $\mathcal{SK}$, $\mathcal{R}$ and $\mathcal{C}$ denote the spaces of public keys, secret keys, sender randomness, and ciphertexts, respectively. A samplable public key cryptosystem consists of four polynomial time algorithms $SPK = (Gen_S, Enc_S, Dec_S, SampleRand)$ with the following characteristics:*

- *$Gen_S(1^n)$: A PPT key generation algorithm that outputs a a public key pk and a secret key sk upon input of a security parameter n.*
- *$Enc_S(pk, b, r)$: A PPT encryption algorithm that receives as input a public key pk and a bit b, outputting a ciphertext c generated with randomness r.*
- *$Dec_S(pk, m)$: A deterministic polynomial time decryption algorithm that receives as input a secret key sk and a ciphertext c, outputting a bit b.*
- *$SampleRand(sk, c)$: An efficient algorithm that outputs a value $r \in \mathcal{R}$ upon input of a secret key sk and a ciphertext c.*

This ensemble of algorithms is considered to be a samplable cryptosystem if it has the following properties:

1. ***Correctness:** For all $pk \in \mathcal{PK}$, all $sk \in \mathcal{SK}$ and all $b \in \{0,1\}$, it holds that $Dec_S(sk, Enc_S(pk, b, r)) = b$ with overwhelming probability over $r \leftarrow \Delta_{\mathcal{R}}$.*
2. ***Uniform distribution of ciphertext on $\mathcal{C}$:** The set $\mathcal{C}$ is finite and the distribution on $\mathcal{C}$ given by*

$$(Enc_S(pk, b, r) : b \in_R \{0,1\}, r \leftarrow \Delta_{\mathcal{R}}) \tag{1}$$

   is statistically indistinguishable from the uniform distribution on $\mathcal{C}$.

3. ***Samplability**: The values $\tilde{r} \in \mathcal{R}$ outputted by the algorithm $SampleRand(sk, c)$ is such that, for every key pair (sk, pk), every $b \in \{0,1\}$, every $r \leftarrow \Delta_{\mathcal{R}}$ and every $Enc_S(pk, b, r)$, it holds that:*

$$(sk, c, r) \approx (sk, c, \tilde{r}) \tag{2}$$

   Note that this property implies that for all but a negligible fraction of $c \in \mathcal{C}$, we have

$$Enc_s(sk, Dec_s(b), SampleRand(sk, c)) = c.$$

3.2 Obtaining samplable public key encryption from the McEliece Assumptions

Now we show that the semantically secure McEliece cryptosystem is samplable, thus implying in sender deniable encryption with negligible detection. Our technique consists in simply extracting the random pad and the random error vector used in the encryption algorithm and outputting these values as the result of *SampleRand*.

Construction 4. *1. $Gen(1^n)$: Run the McEliece key generation algorithm Gen_{ME} with input 1^n, obtaining a secret key $sk = (\mathbf{S}, \mathbf{G}', \mathbf{P})$ and a public key $\mathbf{G} = (\mathbf{SG}'\mathbf{P}, t)$.*

2. *$Enc(pk, b, r)$: Run $c = Enc_{ME}(pk, (r \mid b))$ and output c.*
3. *$Dec(sk, c)$: Run $m = Dec_{ME}(sk, c)$ and output the last bit of m.*
4. *$SampleRand(sk, c)$: Run $m = Dec_{ME}(sk, c)$. Compute $e = m(\mathbf{SGP}) \oplus c$. Denote by m' the $k-1$ first bits of m, output (m', e).*

The triplet of algorithms (Gen, Dec, Enc) was proved to be a semantically secure cryptosystem in [6].

Proposition 5. *The public key encryption scheme of Construction 4 is samplable.*

Proof. The first property of Definition 3 is clearly satisfied, since the basic correctness condition holds for the semantically secure McEliece cryptosystem.

In the notation of Definition 3, the set $\mathcal{R}$ is the set $\mathcal{PAD} \times \mathcal{E}$, where $\mathcal{PAD} = \{0,1\}^{k-1}$ and $\mathcal{E} = \{e \in \{0,1\}^t \mid w_h(e) = t\}$. The set $\mathcal{C}$ is the set of all codewords C_w with t errors. The distribution $\Delta_{\mathcal{R}}$ is the uniform distribution on $\mathcal{PAD} \times \mathcal{E}$. Since the randomness used in the encryption algorithm is chosen uniformly in $\mathcal{R}$ and any bir in the message space $\{0,1\}$ can be encrypted with equal probability, any codewords C_w with t errors is outputted by the encryption algorithm with equal probability. It follows that the distribution (1) is the uniform distribution on $\mathcal{C}$, satisfying the second property of Definition 3.

It is clear that that the third property of Definition 3 is satisfied, since *SampleRand* is able to extract the exact randomness used for generating each ciphertext. □

4 Conclusion

We introduced the first construction of samplable public key encryption based on assumptions other than the hardness of factoring. Our result implies in sender-deniable encryption with negligible detection, being the first construction of such a primitive under the McEliece assumptions. Hence, it is also the first construction of this primitive to achieve post quantum security without resorting to weaker definitions where there are different honest and dishonest encryption algorithms. However, it remains to determine whether it is possible to construct interactive bi-deniable encryption based on the same assumptions, which we leave as a future work.

References

[1] Elwyn R Berlekamp, Robert McEliece, and Henk C A van Tilborg. On the inherent intractability of certain coding problems (corresp. *IEEE Transactions on Information Theory*, (24), 1978.

[2] Ran Canetti, Cynthia Dwork, Moni Naor, and Rafail Ostrovsky. Deniable encryption. In Burton S. Kaliski Jr., editor, *CRYPTO*, volume 1294 of *Lecture Notes in Computer Science*, pages 90–104. Springer, 1997.

[3] Markus Duermuth and David Mandell Freeman. Deniable encryption with negligible detection probability: An interactive construction. Cryptology ePrint Archive, Report 2011/066, 2011. `http://eprint.iacr.org/`.

[4] Jean-Charles Faugère, Valérie Gauthier, Ayoub Otmani, Ludovic Perret, and Jean-Pierre Tillich. A distinguisher for high rate mceliece cryptosystems. Cryptology ePrint Archive, Report 2010/331, 2010. http://eprint.iacr.org/.

[5] R J McEliece. A public-key cryptosystem based on algebraic coding theory. dsn progress report. In *Jet Propulsion Laboratories, CALTECH*, pages 42–44, 1978.

[6] Ryo Nojima, Hideki Imai, Kazukuni Kobara, and Kirill Morozov. Semantic security for the mceliece cryptosystem without random oracles. *Des. Codes Cryptography*, 49(1-3):289–305, 2008.

[7] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 84–93, New York, NY, USA, 2005. ACM.