On the Security Properties of OAEP as an All-or-Nothing Transform[⋆]

Victor Boyko[⋆⋆]

MIT Laboratory for Computer Science
545 Technology Square, Cambridge, MA 02139
boyko@theory.lcs.mit.edu

Abstract. This paper studies All-or-Nothing Transforms (AONTs), which have been proposed by Rivest as a mode of operation for block ciphers. An AONT is an unkeyed, invertible, randomized transformation, with the property that it is hard to invert unless all of the output is known. Applications of AONTs include improving the security and speed of encryption. We give several formal definitions of security for AONTs that are stronger and more suited to practical applications than the original definitions. We then prove that Optimal Asymmetric Encryption Padding (OAEP) satisfies these definitions (in the random oracle model). This is the first construction of an AONT that has been proven secure in the strong sense. Our bound on the adversary's advantage is nearly optimal, in the sense that no adversary can do substantially better against the OAEP than by exhaustive search. We also show that no AONT can achieve substantially better security than OAEP.

Key words: all-or-nothing transforms, encryption modes, OAEP, random oracles, polynomial indistinguishability, semantic security, exact security.

1 Introduction

The concept of an *All-or-Nothing Transform (AONT)* was introduced by Rivest [18] to increase the cost of brute force attacks on block ciphers without changing the key length. As defined in that paper, an AONT is an efficiently computable transformation f, mapping sequences of blocks (i.e., fixed length strings) to sequences of blocks, which has the following properties:

- Given all of $f(x_1, \ldots, x_n) = (y_1, \ldots, y_{n'})$, it is easy to compute $x_1, \ldots, x_n$.
- Given all but one of the blocks of the output (i.e., given $y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_{n'}$ for any $1 \leq j \leq n'$), it is infeasible to find out any information about any of the original blocks x_i.

[⋆] This version of the paper has been truncated due to page limits. The full version is available from the author's homepage [5].

[⋆⋆] Supported by an NSF Graduate Research Fellowship and DARPA grant DABT63-96-C-0018.

Michael Wiener (Ed.): CRYPTO'99, LNCS 1666, pp. 503–518, 1999.

As mentioned by Rivest [18], an AONT should be randomized, so that a known message does not yield a known output.

An AONT itself does not perform any encryption, since there is no secret key information involved. However, if its output is encrypted, block by block, with a block cipher, the resulting scheme will have the following interesting property: An adversary cannot find out any information about any block of the message without decrypting all the blocks of the ciphertext. Now if the adversary attempts to do an exhaustive search for the key, she will need to perform n' decryptions before determining whether a given key is correct. Thus, the attack will be slowed down by a factor of n', without any change in the size of the secret key. This is particularly important in scenarios where the key length is constrained to be insecure or marginally secure (e.g., because of export regulations).

Another very important application of AONTs, as proposed by Johnson et al. [12] for inclusion in the IEEE P1363a standard, is to make fixed-blocksize encryption schemes more efficient. Instead of encrypting the whole message block by block, we can apply AONT to it, and encrypt just some of the blocks of the output. This will be an improvement if the AONT is more efficient than the cipher. It is especially useful if the cipher is a public key cryptosystem, such as RSA [17] or ElGamal [6]. This way we can, for instance, use RSA to securely encrypt messages longer than the key size, without need for a symmetric cipher. This gives an even greater improvement for elliptic-curve cryptosystems, which typically have a block length that is too small to efficiently use the traditional approach of encrypting a symmetric session key, together with padding and redundancy (see Johnson and Matyas [11]). A similar application of AONTs, as proposed by Rivest [19], would be to reduce communication requirements, in case the encryption function greatly expands its input.

The use of AONT with encryption can be particularly useful for remotely keyed encryption, i.e., applications where the part of the system that contains the keys is separate, and where bandwidth restrictions prevent us from sending the whole message from the insecure to the secure component [4]. An example of such a scenario would be the case where the keys are stored in a smartcard, and the user wishes to encrypt or decrypt large files. Through the use of AONT, we can completely eliminate any encryption components from the host system, and restrict such operations to the smart card (this is a generalization of the scheme of Jakobsson et al. [10], substituting general AONTs for the OAEP-like construction used in that paper). The host would transform the message with an AONT, and send one block to the smartcard. The smartcard would encrypt that block, and return it to the host. The encryption of the message will then be the output of the AONT, with one block encrypted. Assuming the block encryption is secure, the whole message will be secure. Note that since the host system does not contain any encryption algorithms, it might not be subject to export regulations.

The major problem with the definition of Rivest [18] is as follows: That definition only speaks about the amount of information that can be learned about a *particular* message block. It does not, however, address the issue of information about the message as a whole (e.g., the XOR of all the blocks). To make the AONT truly useful, we would want it to hide *all* information about the input if any part of the output is missing (we will refer to this as the *semantic security model*). For instance, if an AONT is used for the purpose of slowing down exhaustive search of the key space, a relation between several blocks of the plaintext may provide enough information to the adversary for the purpose of detecting an invalid key.

Another disadvantage of Rivest's model [18] is that it does not consider the relation between the number of bits of AONT output that the adversary has, and the information that is leaked about the input. That model only considers the cases when the adversary has the whole output (in which case she should be able to completely determine the input), and when at least one complete block of the output is missing (in which case it should be infeasible to determine any block of the input). It would be interesting to consider exactly how much information about the input can be determined by looking at all but a certain number l bits of the AONT output, and how much effort is required to obtain that information.

1.1 This Work

The goal of this paper is to provide an AONT construction that is provably secure in the strong sense described above. Our contributions are as follows:

- We give new formal definitions of AONT security in terms of semantic security and indistinguishability. These definitions address the concerns mentioned above and provide the security needed for practical applications. They are parallel to the two notions of security for public-key cryptosystems, defined by Goldwasser and Micali [9]. We consider both the non-adaptive scenario (where the positions of the bits that are removed from AONT output are fixed before the experiment), and the adaptive scenario (where the adversary can choose the positions).
- We prove that OAEP (see Sect. 1.2), a construction originally introduced by Bellare and Rogaway in a different context, satisfies these definitions (in the random oracle model).
- We give an upper bound on the adversary's advantage in getting information about OAEP input when given all but l bits of OAEP output, as opposed to having none of the output. The bound is exact, i.e., does not involve asymptotics. It does not use any computational assumptions and relies only on the properties of random oracles. The bound is directly proportional to the number of adversary's queries to the random oracle and is inversely exponential in the number of bits of OAEP output that are withheld from the adversary.
- We then show that our upper bound is nearly optimal, in the sense that no adversary can do substantially better against OAEP than by exhaustive search. In addition, it will follow that no AONT can achieve substantially better security (i.e., upper bound on the adversary's advantage) than OAEP.

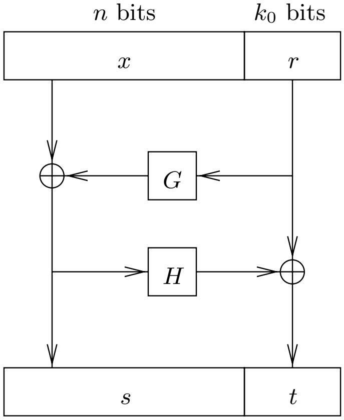

Fig. 1. A diagram of the OAEP

1.2 OAEP

Optimal Asymmetric Encryption Padding (OAEP) was originally introduced by Bellare and Rogaway [2] for the purpose of constructing semantically secure and plaintext-aware public-key encryption schemes from arbitrary trapdoor permutations. For parameters n and k_0, "generator" $G : \{0,1\}^{k_0} \to \{0,1\}^n$, and "hash function" $H : \{0,1\}^n \to \{0,1\}^{k_0}$, the transform OAEP : $\{0,1\}^n \times \{0,1\}^{k_0} \to \{0,1\}^{n'}$, for $n' = n + k_0$, is defined as

$$\mathrm{OAEP}^{G,H}(x,r) = x \oplus G(r) \,\|\, r \oplus H(x \oplus G(r)) \ ,$$

where $\|$ denotes concatenation. Here x is the message and r is a random string. In applications, n would be the length of a message, and k_0 would be the security parameter, e.g., $k_0 = 128$. We will often refer to the first half of the OAEP output (i.e., $x \oplus G(r)$) as s, and to the second half (i.e., $r \oplus H(s)$) as t. Here $|s| = n$ and $|t| = k_0$. We may also write $\mathrm{OAEP}^{G,H}(x)$, implying that r is chosen uniformly at random from $\{0,1\}^{k_0}$.

A diagram of the OAEP appears in Fig. 1.

Functions G and H are "random oracles," as introduced in by Bellare and Rogaway [1]. Bellare and Rogaway [2] show that if G and H are "ideal," i.e., they are random functions, and $f : \{0,1\}^{k_0+n} \to \{0,1\}^{k_0+n}$ is a trapdoor permutation, then the encryption scheme

$$\mathcal{E}^{G,H}(x) = f(\mathrm{OAEP}^{G,H}(x,r)) \ ,$$

with r chosen at random for each encryption, is semantically secure, in the sense of Goldwasser and Micali [9]. They also show that a small modification of that scheme provides plaintext-awareness (introduced by Bellare and Rogaway [2]).

1.3 Previous Work

Rivest [18] has proposed the following construction ("the package transform") as a candidate AONT:

- Let E be a block cipher. Let K_0 be a fixed, publicly known key for E.
- Let the input message be the sequence of blocks $m_1, m_2, \ldots, m_s$.
- Choose at random a key K' for E.
- Compute the output sequence $m'_1, m'_2, \ldots, m'_{s'}$, for $s' = s + 1$, as follows:
  - Let $m'_i = m_i \oplus E(K', i)$ for $i = 1, 2, \ldots, s$.
  - Let

    $$m'_{s'} = K' \oplus h_1 \oplus h_2 \oplus \cdots \oplus h_{\mathrm{s}},$$

    where

    $$h_i = E(K_0, m'_i \oplus i)$$

    for $i = 1, 2, \ldots, s$.

No formal proof was given that this construction is actually an AONT. The heuristic argument for security is based on the idea that if any block of the output is unknown, then K' cannot be computed, and so it is infeasible to compute any message block. Rivest [18] mentions that "the package transform" can be viewed as a special case of the OAEP, for $G(x) = E(x, 1) \| E(x, 2) \| \cdots \| E(x, s)$ and $H(x) = \bigoplus_{i=1}^{s} E(K_0, x_i \oplus i)$. However, no claims about OAEP itself are made in that paper.

Johnson et al. [12], in their contribution for the IEEE P1363a standard, give an OAEP-like transform that uses four rounds of hash applications instead of two. A heuristic analysis of the security of that construction is given by Matyas et al. [13]. Using an informal assumption about the hardness of the underlying hash functions, they argue that the number of operations required to determine the secret bits in the input message grows exponentially with the number of unknown bits. However, we are not aware of any formal proof of security of the transform of Johnson et al. [12]. In any case, the analysis of Matyas et al. [13] is not directly applicable if there are fewer than four rounds, so it does not work for OAEP.

Stinson [20] gives a treatment of AONTs from the point of view of unconditional security. Similarly to Rivest [18], Stinson's definition only considers the amount of information leaked about a particular block of the message, as opposed to the whole message. He uses a straightforward formalization of Rivest's definition above, suitably modified for information-theoretic security. Stinson then goes on to propose some constructions for AONTs using linear transforms, which can be proven secure in that model. The basic idea of these constructions is to use the function $\phi(\mathbf{x}) = \mathbf{x}M^{-1}$, where $\mathbf{x}$ is a vector of s message blocks (considered as elements of $GF(q)$, for some prime power q), and M is an invertible s by s matrix over $GF(q)$, such that no entry of M is equal to 0. It is easy to see that each component of $\mathbf{x}$ linearly depends on all the components of $\mathbf{y} = \phi(\mathbf{x})$ (since $\mathbf{x} = \mathbf{y}M$).

It is conceivable that "the package transform" of Rivest [18] would be secure in the semantic security model (with sufficiently strong assumptions about the block cipher). The construction of Johnson et al. [12] may also be secure, although no formal proof has been given. However, the linear constructions of Stinson [20] would definitely not be secure in that model, since it is easy to come up with linear relations among the elements of $\mathbf{x}$ by looking at just a few elements of $\phi(\mathbf{x})$ (in fact, since ϕ is linear and deterministic, *every* output of $\phi(\mathbf{x})$ gives a linear relation on elements of $\mathbf{x}$). Even if the message is padded with random blocks, it is still possible to extract partial information about the message if the number of known outputs is larger than the number of random blocks.

It is interesting to note that the relationship between the number of missing bits and adversary's required effort has come up in other contexts. Merkle [14], in one of the first papers on public key cryptography, defines the concept of a "puzzle," which is a cryptogram that requires $\Theta(N)$ work to break, where N is some number depending on the security parameters (the total amount of work put in by the communication parties is going to be $\Theta(N)$). Merkle's proposed construction of such "puzzles" is to take a block cipher and restrict the size of the key space, by varying only $\Theta(\log N)$ bits of the key and fixing the rest. It is assumed that breaking a cryptogram of the underlying cipher, when all but $\Theta(\log N)$ bits of the key are known, requires $\Theta(N)$ work.

Even et al. [7] assume the existence of a "uniformly secure" block cipher for their construction of a contract signing protocol. They consider a block cipher "uniformly secure" if it is infeasible to find a key for a given plaintext-ciphertext pair when no information about the key is known; but if the first i bits of the key are known, then there is an algorithm for breaking the cryptogram in time $t(k-i)$, for some function $t(\cdot)$, and no algorithm can do it faster than in time $\frac{1}{2}t(k-i)$. Here k is the key length.

Both Merkle [14] and Even et al. [7] conjecture that standard block ciphers, such as Lucifer [8] or DES [16], satisfy their assumptions. However, uniform security is probably not a common consideration in block cipher design, as almost all applications of these primitives assume the whole key to be secret. Thus it may be unsafe to make such an assumption about standard block ciphers. In fact, this is, in effect, one of the criticisms given by Ben-Or et al. [3] of the work of Even et al. [7]. It seems to us, however, that the methods of our paper can be used to give a simple construction that will turn any block cipher that is secure in the regular sense into one which is uniformly secure. See Sect. 5 for details.

1.4 Outline

The outline of the rest of the paper is as follows. Section 2 describes the notation and model. In Sect. 3, we give formal definitions of security for AONTs. Section 4 presents the results on the security of OAEP as an AONT. Section 5 discusses open problems.

2 Notation and Model

Let us speak briefly about our notation and model. All algorithms used are oracle Turing machines, possibly randomized. Oracle queries execute in unit time. If A is a randomized algorithm, we may write $A(x_1, \ldots)$ to mean the distribution of A's output on certain inputs. We may also specify the coins explicitly, as in $A(r_A, x_1, \ldots)$, in which case the notation will refer to the fully determined output.

We will write $x \overset{R}{\leftarrow} X$ to mean that a variable x is to be chosen at random according to distribution X. As a shorthand, $x_1, x_2 \overset{R}{\leftarrow} X$ denotes $x_1 \overset{R}{\leftarrow} X$, $x_2 \overset{R}{\leftarrow} X$. On the other hand, $x \leftarrow X$ will mean that x is to be set to the result of evaluating expression X (which is not random). If S is a set, then we will write $x \overset{R}{\leftarrow} S$ to mean that x is chosen uniformly at random from S. We will write $\Pr[x \overset{R}{\leftarrow} X; y \overset{R}{\leftarrow} Y; z \leftarrow Z; \ldots : p(x, y, z, \ldots)]$ to mean the probability of predicate $p(x, y, z, \ldots)$, when x is chosen at random according to distribution X, y is chosen at random according to distribution Y, z is set to the result of evaluating expression Z (possibly a function of x and y), etc. Similarly, we will write $E[x \overset{R}{\leftarrow} X; \ldots : f(x, \ldots)]$ to mean the expected value of $f(x, \ldots)$ when x is chosen at random according to distribution X, etc.

To specify the distribution of a random function ("random oracle"), such as G and H for OAEP, we will use notation like $G, H \overset{R}{\leftarrow} \Omega$, where Ω is the set of all maps from the set $\{0,1\}^*$ of finite strings to the set $\{0,1\}^\infty$ of infinite strings. The notation should be interpreted as appropriate in its context, restricting the input and truncating the output of the function as necessary.

For $x \in \{0,1\}^*$, $1 \leq i \leq |x|$, and $0 \leq l \leq |x| - i + 1$, let $\mathrm{substr}(x, i, l)$ denote the substring of x starting at bit i (with the leftmost bit being 1) and having length l.

For any integer m and $L \subseteq [1, m]$, we define $h_{m,L} : \{0,1\}^m \to \{0,1\}^{m-|L|}$ as follows: $h_{m,L}$ takes a bit string of length m and throws out ("*h*ides") the bit positions indicated by L. More precisely, if we let $\bar{L}_i$, for $1 \leq i \leq m - |L|$ denote the ith smallest element of $\bar{L} = [1, m] \setminus L$, then

$$h_{m,L}(x) = \mathrm{substr}(x, \bar{L}_1, 1) \| \mathrm{substr}(x, \bar{L}_2, 1) \| \cdots \| \mathrm{substr}(x, \bar{L}_{m-|L|}, 1) \ .$$

For $n' \geq l \geq 0$, let $\left\{ {n' \atop l} \right\} = \{L \subseteq [1, n'] : |L| = l\}$.

3 Definitions

Our definitions of security for AONTs are patterned after the notions of security for encryption defined by Goldwasser and Micali [9]: polynomial security (polynomial indistinguishability) and semantic security.[1] We also try to define security "exactly," as in Bellare and Rogaway [2]: instead of concerning ourselves

[1] To prevent confusion, we note that while Bellare and Rogaway [2] talk about semantic security (for encryption), the definition they give is actually stated in terms

with asymptotics (i.e., showing that the adversary's advantage is "negligible" in the security parameters), we are interested in giving an exact bound on the adversary's advantage, as a function of the adversary's running time, the number of bits of AONT's output given to the adversary, etc.

For simplicity, we will formulate the definitions in terms of a single random oracle Γ. No generality is lost, since a single random oracle can be used to simulate several, by constructing the query as the concatenation of the oracle index and the original query. For instance, we could use Γ to simulate random oracles G and H by translating query x to G into query $0\|x$ to Γ and query y to H into $1\|y$. In addition, it would be easy to change the definitions for the case of no random oracles.

The *non-adaptive indistinguishability scenario* is as follows: Let L be an arbitrary set of l bit positions. The adversary runs in two stages:

1. **Find stage:** The adversary is given L and access to Γ. She outputs $x_0 \in \{0,1\}^n$, $x_1 \in \{0,1\}^n$, and $c_{\mathsf{f}} \in \{0,1\}^*$.
2. **Guess stage:** The adversary is given c_{f} and, for random bit b, $\mathrm{AONT}^{\Gamma}(x_b)$ with bit positions L missing. The adversary has access to Γ. She has to guess b.

Note that x_0 and x_1 do not need to be explicitly passed to the guess stage, since they may be included in c_{f}. We may view c_{f} as the saved state of the adversary at the end of the find stage.

We want the adversary's probability of correctly guessing b to be as close as possible to $\frac{1}{2}$. The formal definition is as follows:

Definition 1 (Non-adaptive indistinguishability). *Let* AONT *be a randomized transform mapping n-bit messages to n'-bit outputs and using random oracle Γ. Let l be between* 1 *and n'. An adversary A is said to succeed in* (T, q_Γ, ϵ)-**distinguishing** AONT *with l missing bits if there exists $L \in \left\{ {n' \atop l} \right\}$ such that*

$$\Pr[\Gamma \xleftarrow{R} \Omega; (x_0, x_1, c_{\mathsf{f}}) \xleftarrow{R} A^{\Gamma}(L, \mathsf{find}); b \xleftarrow{R} \{0,1\};$$
$$y \xleftarrow{R} \mathrm{AONT}^{\Gamma}(x_b) : A^{\Gamma}(h_{n',L}(y), c_{\mathsf{f}}, \mathsf{guess}) = b] \geq \frac{1}{2} + \epsilon \ ,$$

and, moreover, in the experiment above, A runs for at most T steps, and makes at most q_Γ queries to Γ.

It follows from this definition that in order for an AONT to be secure in the sense of non-adaptive indistinguishability for certain choices of parameters, it needs to be that for every adversary and every L, the adversary's advantage has to be less than ϵ.

of indistinguishability. This is acceptable in their context, since the two notions are known to be equivalent for encryption (see Micali et al. [15]). In our context, however, we state and analyze each one separately, since no equivalence has yet been proven.

The *adaptive indistinguishability scenario* is as follows: The adversary runs in three stages. The first stage chooses a value of L, while the last two stages are same as in the non-adaptive indistinguishability scenario. The adversary runs as follows:

1. **Select stage:** The adversary is given l and access to Γ. She selects l bit positions and outputs $L \in \left\{ {n' \atop l} \right\}$ and $c_{\mathsf{s}} \in \{0,1\}^*$.
2. **Find stage:** The adversary is given c_{s} and access to Γ. She outputs $x_0 \in \{0,1\}^n$, $x_1 \in \{0,1\}^n$, and $c_{\mathsf{f}} \in \{0,1\}^*$.
3. **Guess stage:** The adversary is given c_{f} and, for random bit b, $\mathrm{AONT}^{\Gamma}(x_b)$ with bit positions L missing. The adversary has access to Γ. She has to guess b.

Similarly to the remark about x_0 and x_1 above, we note that L does not need to be explicitly passed to the find and guess stages, since it may be included in c_{s}, and then put into c_{f}.

In the formal definition, we will assume that the adversary's select stage will always output a valid value of $L \in \left\{ {n' \atop l} \right\}$ (this can be implemented by having a suitable encoding).

Definition 2 (Adaptive indistinguishability). *Let* AONT *be a randomized transform mapping n-bit messages to n'-bit outputs and using random oracle Γ. Let l be between* 1 *and n'. An adversary A is said to succeed in (T, q_Γ, ϵ)-* **adaptively-distinguishing** AONT *with l missing bits if*

$$\begin{aligned}\Pr[&\Gamma \overset{R}{\leftarrow} \Omega; (L, c_{\mathsf{s}}) \overset{R}{\leftarrow} A^{\Gamma}(l, \mathsf{select}); (x_0, x_1, c_{\mathsf{f}}) \overset{R}{\leftarrow} A^{\Gamma}(c_{\mathsf{s}}, \mathsf{find});\\ &b \overset{R}{\leftarrow} \{0,1\}; y \overset{R}{\leftarrow} \mathrm{AONT}^{\Gamma}(x_b) : A^{\Gamma}(h_{n',L}(y), c_{\mathsf{f}}, \mathsf{guess}) = b] \geq \frac{1}{2} + \epsilon \ ,\end{aligned}$$

and, moreover, in the experiment above, A runs for at most T steps, and makes at most q_Γ queries to Γ.

Note that for the application of speeding up encryption that was mentioned above in Sect. 1, it is sufficient for the AONT to be secure for a fixed choice of the missing part of the output (since the user decides which part will be encrypted). Thus, for that application, it is sufficient for the AONT to be secure in the non-adaptive scenario. However, when an AONT is used to increase the cost of exhaustive search, it needs to be secure in the adaptive scenario, since then the adversary has a choice of which blocks to decrypt.

For the adaptive and non-adaptive indistinguishability scenarios, we will assume, without loss of generality, that A never asks the same oracle query more than once (A can accomplish this by remembering the history of past queries; this history can be passed between stages through c_{s} and c_{f}).

The *non-adaptive semantic security scenario* is as follows: Let L be an arbitrary set of l bit positions and $f : \{0,1\}^n \to \{0,1\}^*$ be an arbitrary deterministic function. The adversary runs in two unconnected stages (each stage can be viewed as a separate algorithm):

- **Find stage:** The adversary is given L and access to Γ. She outputs $x \in \{0,1\}^n$.
- **Guess stage** (no data is passed from the find stage): The adversary is given L and $\mathrm{AONT}^{\Gamma}(x)$ with bit positions L missing. The adversary has access to Γ. She has to guess $f(x)$.

In the context of the traditional definition of semantic security for encryption, the adversary's find stage may be seen as the sampling algorithm for a distribution of messages, and the guess stage as the actual predicting algorithm. We want the adversary not to be able to do substantially better than always outputting the most probable value of $f(x)$. The formal definition is as follows:

Definition 3 (Non-adaptive semantic security). *Let* AONT *be a randomized transform mapping n-bit messages to n'-bit outputs and using random oracle Γ. Let l be between* 1 *and n'. Let $f : \{0,1\}^n \to \{0,1\}^*$ be any deterministic function. An adversary A is said to succeed in* (T, q_Γ, ϵ)-**predicting** *f from* AONT *with l missing bits if there exists $L \in \left\{ {n' \atop l} \right\}$ such that*

$$\Pr[\Gamma \overset{R}{\leftarrow} \Omega; x \overset{R}{\leftarrow} A^{\Gamma}(L, \mathsf{find}); y \overset{R}{\leftarrow} \mathrm{AONT}^{\Gamma}(x) : \\ A^{\Gamma}(L, h_{n',L}(y), \mathsf{guess}) = f(x)] \geq p_{A,f} + \epsilon \ , \quad (1)$$

where

$$p_{A,f} = E[\Gamma \overset{R}{\leftarrow} \Omega : \max_z \Pr[x \overset{R}{\leftarrow} A^{\Gamma}(L, \mathsf{find}) : f(x) = z]] \ ,$$

and, moreover, in the experiment (1)*, A runs for at most T steps, and makes at most q_Γ queries to Γ.*

The expectation in the definition of $p_{A,f}$ is necessary to handle the possibility that the adversary may choose x to be a function of Γ (e.g., x could be set to the result of querying Γ on some fixed input). This would result in perfect prediction (both the find and guess stages can compute the same x), even though the output of the find stage will appear random, for random Γ. Thus, the quantity

$$\max_z \Pr[\Gamma \overset{R}{\leftarrow} \Omega; x \overset{R}{\leftarrow} A^{\Gamma}(L, \mathsf{find}) : f(x) = z]$$

could be much smaller than the adversary's success probability. However, for any fixed Γ, this adversary would always output the same x, so $p_{A,f} = 1$. Thus, this adversary's advantage ϵ will have to be zero.

In the semantic security scenario (both adaptive and non-adaptive), no information is passed between the adversary's find and guess stages, except $h_{n',L}(\mathrm{AONT}^{\Gamma}(x))$ (otherwise, the find stage could simply pass the value of $f(x)$). We will therefore remove the assumption that A can't make the same query to Γ more than once. We will still assume, though, that all queries are unique within a single stage.

The *adaptive semantic security scenario* is same as the non-adaptive one, except for the addition of the select stage before the find stage, in which the adversary outputs L. The formal definition is as follows:

Definition 4 (Adaptive semantic security). *Let* AONT *be a randomized transform mapping n-bit messages to n'-bit outputs and using random oracle Γ. Let l be between* 1 *and n'. Let $f : \{0,1\}^n \to \{0,1\}^*$ be any deterministic function. An adversary A is said to succeed in (T, q_Γ, ϵ)-***adaptively-predicting*** *f from* AONT *with l missing bits if*

$$\Pr[\Gamma \overset{R}{\leftarrow} \Omega; (L, c_{\mathsf{s}}) \overset{R}{\leftarrow} A^\Gamma(l, \mathsf{select}); x \overset{R}{\leftarrow} A^\Gamma(c_{\mathsf{s}}, \mathsf{find}); y \overset{R}{\leftarrow} \mathrm{AONT}^\Gamma(x) : \\ A^\Gamma(h_{n',L}(y), c_{\mathsf{s}}, \mathsf{guess}) = f(x)] \geq p_{A,f} + \epsilon \ , \quad (2)$$

where

$$p_{A,f} = E[\Gamma \overset{R}{\leftarrow} \Omega; (L, c_{\mathsf{s}}) \overset{R}{\leftarrow} A^\Gamma(l, \mathsf{select}) : \max_z \Pr[x \overset{R}{\leftarrow} A^\Gamma(c_{\mathsf{s}}, \mathsf{find}) : f(x) = z]] \ ,$$

and, moreover, in the experiment (2), *A runs for at most T steps, and makes at most q_Γ queries to Γ.*

Note that since information may be passed from the select stage to the find and guess stages (through c_{s}), we can assume that no query from the select stage is repeated in any of the other stages. There is no danger in passing c_{s} to the guess stage, since c_{s} is generated before x is chosen (note that $p_{A,f}$ involves an expectation over c_{s}, so the adversary will not gain any advantage by choosing $(x, f(x))$ at the select stage and then passing it to the other stages).

4 Security Results

Throughout most of this section we will be using two random oracles G and H. As mentioned above, we can still use our definitions, since Γ could be used to simulate G and H. We will write $A^{G,H}$ in place of A^Γ. We will also use notation $(T, q_{\mathrm{G}}, q_{\mathrm{H}}, \epsilon)$-$\cdots$ (e.g., "an adversary $(T, q_{\mathrm{G}}, q_{\mathrm{H}}, \epsilon)$-distinguishes") as a shorthand for $(T, q_{\mathrm{G}} + q_{\mathrm{H}}, \epsilon)$-$\cdots$, with the additional condition that at most q_{G} queries are made to G and at most q_{H} queries are made to H.

4.1 Non-adaptive Indistinguishability: Upper Bound

Theorem 1. *Suppose $l \leq k_0$ and $k_0 \geq 14$. Suppose that there exists an adversary A that $(T, q_{\mathrm{G}}, q_{\mathrm{H}}, \epsilon)$-distinguishes* OAEP *with l missing bits, where $q_{\mathrm{G}} \leq 2^{k_0 - 1}$. Then*

$$\epsilon \leq 8 q_{\mathrm{G}} \frac{k_0}{\log_2 k_0} 2^{-l} \ .$$

The proof has been omitted due to page limits and can be found in the full version of this paper [5]. The intuition behind the result is as follows: Let r_0 be the value of r that was used to generate $\tilde{y} = h_{n',L}(\mathrm{OAEP}^{G,H}(x_b))$ in a particular experiment. Then, the adversary cannot find out any information about x_b unless she queries G for $G(r_0)$ (since x_b only appears in $\mathrm{OAEP}^{G,H}(x_b, r_0)$ as $x_b \oplus G(r_0)$). There are $\sim 2^l$ possible values of r_0, corresponding to the 2^l values of y that are consistent with $\tilde{y}$. Thus we would expect the probability that any of the adversary's queries to G are equal to r_0 to be bounded by approximately $q_{\mathrm{G}} 2^{-l}$. The complication is that there may be fewer than 2^l possible values of r_0 and that these values may not be equally probable, given $\tilde{y}$. These possible variations in probability cause the term $O(\frac{k_0}{\log k_0})$.

Note that this result, like all the others in this paper, does not use any computational assumptions and the bound is information theoretic, based on the properties of random oracles. In fact, the bound does not directly depend on T, the adversary's running time. It does, however, have implications for the running time, since $T \geq q_{\mathrm{G}} + q_{\mathrm{H}}$ (every oracle query takes unit time).

4.2 Non-adaptive Indistinguishability: Lower Bound

To see how good our upper bound is, let us try to give a lower bound on the adversary's advantage, by estimating the success of exhaustive search. This lower bound applies to any AONT.

Theorem 2. *Let* AONT *be a randomized transform mapping n-bit messages to n'-bit outputs and using random oracle Γ. Let l be between* 1 *and* $n-3$*. Then, for any $L \in \left\{ {n' \atop l} \right\}$ and any N between* 1 *and* 2^l*, there exists an adversary that $(NT, Nq_\Gamma, \epsilon)$-distinguishes* AONT *with l missing bits, with*

$$\epsilon \geq \frac{1}{16} N 2^{-l}.$$

Here T and q_Γ are the time and number of queries to Γ, respectively, taken by a single evaluation of AONT*.*

The proof has been omitted due to page limits and can be found in the full version of this paper [5]. The idea of the proof is as follows: The exhaustive search algorithm that achieves the advantage of at least $\frac{1}{16} N 2^{-l}$ works by choosing x_0 and x_1 independently at random in the find stage. The guess stage tries random values of the missing bits, up to N times, and, if the inverse AONT returns $x_{b'}$ for $b' \in \{0, 1\}$, produces b' as the guess. If none of the trials has succeeded, a random bit is returned. The idea of the analysis of this algorithm is that every trial in the guess stage has probability of at least 2^{-l} of succeeding with the correct value of b (since there exists a choice of the missing bits, namely the values that actually appeared in y, that leads to x_b). On the other hand, since x_{1-b} is chosen uniformly and independently, the probability of getting x_{1-b} in any particular trial is $2^{-n} \leq 2^{-l-3}$.

We see from Theorems 1 and 2, that no adversary can improve by a factor of more than $O(\frac{k_0}{\log k_0})$ over exhaustive search. Since, for large l, this factor is negligible compared to 2^{-l}, our bounds for OAEP are nearly optimal.

We also see that no AONT can be substantially more secure than OAEP, in the sense that no AONT can have an upper bound that is better than OAEP's by a factor of more than $O(\frac{k_0}{\log k_0})$.

4.3 Adaptive Indistinguishability

Theorem 3. *Suppose $l \leq k_0$, $l \leq \frac{n}{2}$, and $k_0 \geq 14$. Suppose that there exists an adversary A that $(T, q_{\mathrm{G}}, q_{\mathrm{H}}, \epsilon)$-adaptively-distinguishes* OAEP *with l missing bits, where $q_{\mathrm{G}} \leq 2^{k_0-1}$. Then*

$$\epsilon \leq 8(q_{\mathrm{G}} + q_{\mathrm{H}}) \frac{k_0}{\log_2 k_0} 2^{-l} \ .$$

The proof has been omitted due to page limits and can be found in the full version of this paper [5]. It is very similar to the proof of Theorem 1, with the only major difference being that we have to consider the possible correlation between L and H (since L may be chosen by the adversary to depend on H). This is taken care of by showing that with large probability (depending on q_{H}), the queries made in the select stage will not constrain H enough to spoil those properties of it that are used in the proof of Theorem 1.

We can easily see that for the adaptive indistinguishability scenario, as for the non-adaptive one, our bound is optimal within a factor of $O(\frac{k_0}{\log k_0})$ of the advantage given by exhaustive search for an arbitrary AONT (in this scenario, exhaustive search would choose a random L in the select stage).

4.4 Non-adaptive Semantic Security

Theorem 4. *Suppose $l \leq k_0$ and $k_0 \geq 14$. Suppose that there exists a deterministic function $f : \{0,1\}^n \rightarrow \{0,1\}^*$ and an adversary A that $(T, q_{\mathrm{G}}, q_{\mathrm{H}}, \epsilon)$-predicts f from* OAEP *with l missing bits, where $q_{\mathrm{G}} \leq 2^{k_0-1}$. Then*

$$\epsilon \leq 8q_{\mathrm{G}} \frac{k_0}{\log_2 k_0} 2^{-l} \ .$$

The proof has been omitted due to page limits and can be found in the full version of this paper [5]. It is a simple modification of the proof of Theorem 1, with the difference being the estimation of the adversary's success probability in the case where she has not queried G for $G(r_0)$ (where r_0 is the value of r used to compute $\mathrm{OAEP}^{G,H}(x)$ in the experiment). In the case of Theorem 1 that success probability was $\frac{1}{2}$, while here it can be easily seen to be less than or equal to $p_{A,f}$.

We have not yet shown a lower bound for the semantic security scenarios. We still expect our upper bound for these scenarios to be nearly optimal, as for the indistinguishability scenarios.

4.5 Adaptive Semantic Security

Theorem 5. *Suppose $l \leq k_0$, $l \leq \frac{n}{2}$, and $k_0 \geq 14$. Suppose that there exists a deterministic function $f : \{0,1\}^n \rightarrow \{0,1\}^*$ and an adversary A that $(T, q_{\mathrm{G}}, q_{\mathrm{H}}, \epsilon)$-adaptively-predicts f from* OAEP *with l missing bits, where $q_{\mathrm{G}} \leq 2^{k_0-1}$. Then*

$$\epsilon \leq 8(q_{\mathrm{G}} + q_{\mathrm{H}}) \frac{k_0}{\log_2 k_0} 2^{-l} \ .$$

The proof of Theorem 5 is a simple combination of the proof of Theorem 1 with the modifications needed for Theorems 3 and 4 (see the full version of this paper for details [5]).

5 Open Problems

The first open problem that comes to mind is to improve our bounds. The best would be to bring the upper bounds within a constant factor of exhaustive search, or to devise an algorithm that does better than exhaustive search. Also, it would be interesting to give lower bounds for the semantic security scenarios.

Another open problem is to show equivalence (or non-equivalence) of our definitions, trying to carry over the exact bounds as much as possible. There are also other possible models to consider, such as a scenario where, instead of specifying the positions of the missing bits in advance, the adversary is allowed to ask for bits after seeing the value of other bits. Also, just as there are several variations of the definition of semantic security for encryption, one might consider other definitions for AONTs, and whether they are equivalent to the ones in this paper.

One of the most interesting open problems related to AONTs is to construct a secure AONT (in the sense of our definitions) without the use of random oracles. One might start by trying to modify the OAEP by replacing either G or H by a deterministic function. In any case, until formal results about deterministic AONTs are obtained, it would be fruitful to investigate the practical security concerns that arise in the use of OAEP as an AONT, when the random oracles are instantiated with deterministic hash functions.

Another interesting question is whether there is any relation between the properties of OAEP as an AONT, and its original proposed use for constructing secure cryptosystems. One might ask, for instance, if OAEP could be replaced by an arbitrary AONT in the construction of Bellare and Rogaway [2].

One could also look into the possibility of generalizing the definitions of AONT security, so that instead of getting a certain number of bits of the output, the adversary gets the equivalent amount of information through other means, i.e., by seeing the value of some transformation of the output that reduces its entropy by l. The function $h_{n',L}$ is just one example of such a transformation.

Finally, now that we have a provably secure AONT, it would be of great interest to find new applications of this primitive. We hope that its usefulness extends far beyond its original applications.

We note that an AONT that satisfies our definitions, such as OAEP, can be used to implement the “puzzles” of Merkle [14], or the notion of uniform security of Even et al. [7]. The “puzzles” could be made by publishing a certain number of bits of AONT output on a bit string of sufficient redundancy (so that, with overwhelming probability, only one such string corresponded to the published information). Similarly, cryptograms of uniform security could be achieved by publishing a part of AONT output on the key used in the cryptogram. It seems, though, that a simpler construction would suffice: Let $E(K, M) \rightarrow C$ be a regular symmetric cryptosystem, and let H be a random oracle. Then, it seems to us that, using the methods of this paper, it can be shown that $E(H(K), M) \rightarrow C$ is a uniformly secure cryptosystem. On the other hand, the AONT construction for the “puzzles” has the advantage that it does not use encryption, which could put it outside the scope of export regulations. It would be interesting to investigate these issues further.

Acknowledgments

I would like to express my deepest gratitude to Ronald Rivest for his direction, support, and frequent discussions. This work would have been impossible without him. I am also greatly indebted to Shafi Goldwasser for her numerous comments on the manuscript, helpful advice, and discussions. In addition, I would like to thank the following for their help, comments, and advice: Yevgeniy Dodis, Markus Jakobsson, Burt Kaliski, Daniele Micciancio, David Molnar, Kazuo Ohta, Alexander Perlin, Zulfikar Ramzan, Leo Reyzin, Phillip Rogaway, Julien Stern, and Yiannis Tsiounis.

References

[1] Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In *Proceedings of the First Annual Conference on Computer and Communications Security*. ACM, 1993. Latest version can be obtained from `http://www-cse.ucsd.edu/users/mihir/papers/ro.html`.

[2] Mihir Bellare and Phillip Rogaway. Optimal asymmetric encryption. In Alfredo De Santis, editor, *Advances in Cryptology—EUROCRYPT 94*, volume 950 of *Lecture Notes in Computer Science*, pages 92–111. Springer-Verlag, 1995, 9–12 May 1994. Latest version can be obtained from `http://www-cse.ucsd.edu/users/mihir/papers/pke.html`.

[3] Michael Ben-Or, Oded Goldreich, Silvio Micali, and Ronald L. Rivest. A fair protocol for signing contracts. *IEEE Transactions on Information Theory*, 36(1):40–46, 1990.

[4] Matt Blaze. High-bandwidth encryption with low-bandwidth smartcards. In Dieter Grollman, editor, *Fast Software Encryption: Third International Workshop*, volume 1039 of *Lecture Notes in Computer Science*, pages 33–40, Cambridge, UK, 21–23 February 1996. Springer-Verlag.

[5] Victor Boyko. On the security properties of OAEP as an all-or-nothing transform. Available from `http://theory.lcs.mit.edu/~boyko/aont-oaep.html`. Full version of the current paper.

[6] T. ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Trans. Inform. Theory*, 31:469–472, 1985.

[7] S. Even, O. Goldreich, and A. Lempel. A randomized protocol for signing contracts. *Communications of the ACM*, 28:637–647, 1985.

[8] H. Feistel. Cryptography and computer privacy. *Scientific American*, 228:15–23, May 1973.

[9] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, April 1984.

[10] Markus Jakobsson, Julien Stern, and Moti Yung. Scramble all, encrypt small. In *Fast Software Encryption: 6th International Workshop*, Rome, Italy, 24–26 March 1999.

[11] Don Johnson and Stephen Matyas. Asymmetric encryption: Evolution and enhancements. *CryptoBytes*, 2(1):1–6, Spring 1996. Available from `ftp://ftp.rsa.com/pub/cryptobytes/crypto2n1.pdf`.

[12] Don Johnson, Stephen Matyas, and Mohammad Peyravian. Encryption of long blocks using a short-block encryption procedure. Available from `http://grouper.ieee.org/groups/1363/contributions/peyrav.ps`, November 1996. Submitted for inclusion in the IEEE P1363a standard.

[13] Stephen Matyas, Mohammad Peyravian, and Allen Roginsky. Security analysis of Feistel ladder formatting procedure. Available from `http://grouper.ieee.org/groups/1363/contributions/peyrav2.ps`, March 1997.

[14] R. C. Merkle. Secure communications over insecure channels. *Communications of the ACM*, 21:294–299, April 1978.

[15] S. Micali, C. Rackoff, and R. H. Sloan. The notion of security for probabilistic cryptosystems. *SIAM J. Computing*, 17(2):412–426, April 1988.

[16] National Institute of Standards and Technology (NIST). *FIPS Publication 46: Announcing the Data Encryption Standard*, January 1977. Originally issued by National Bureau of Standards.

[17] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[18] Ronald Rivest. All-or-nothing encryption and the package transform. In Eli Biham, editor, *Fast Software Encryption: 4th International Workshop*, volume 1267 of *Lecture Notes in Computer Science*, pages 210–218, Haifa, Israel, 20–22 January 1997. Springer-Verlag. Also available from `http://theory.lcs.mit.edu/~rivest/fusion.ps`.

[19] Ronald Rivest. Chaffing and winnowing: Confidentiality without encryption. Available from `http://theory.lcs.mit.edu/~rivest/chaffing-980701.txt`, July 1998.

[20] D.R. Stinson. Some observations on all-or-nothing transforms. Available from `http://cacr.math.uwaterloo.ca/~dstinson/papers/AON.ps`, 31 August 1998.