Public Key Encryption without Random Oracle Made Truly Practical

Puwen Wei[1,⋆], Xiaoyun Wang[1,2,⋆], and Yuliang Zheng[3]

[1] Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education, Shandong University, Jinan 250100, China
weipuwen@mail.sdu.edu.cn
[2] Institute for Advanced Study, Tsinghua University, Beijing 100084, China
xiaoyunwang@mail.tsinghua.edu.cn
[3] Department of Software and Information Systems, University of North Carolina at Charlotte, Charlotte, NC 28223, USA
yzheng@uncc.edu

Abstract. An important research area in the past decade is to search for efficient cryptographic schemes that do not rely for their security on the controversial random oracle assumption. In this paper, we continue this line of endeavors and report our success in identifying a very efficient public key encryption scheme whose formal security proof does not require a random oracle. Specifically, we show how to modify a universal hash based public key encryption scheme proposed by Zheng and Seberry at Crypto'92, in such a way that the resultant scheme not only preserves efficiency but also admits provable security against adaptive chosen ciphertext attack without a random oracle. We also compare the modified Zheng-Seberry scheme with related encryption schemes in terms of efficiency and underlying assumptions, supporting our conclusion that the modified Zheng-Seberry scheme is preferable to its competitors.

Keywords: random oracle, universal hash, public key encryption.

1 Introduction

The notion of chosen ciphertext security was introduced by Naor and Yung [24]. Then, Rackoff and Simon [26] provided a stronger notion called indistinguishability under adaptive chosen ciphertext attack (IND-CCA2). Adaptive chosen ciphertext security has since become a standard notion for the security of public key encryption. A considerable amount of work on the construction of adaptive chosen ciphertext secure encryption have been presented. Some of these research results are based on non-interactive zero knowledge proofs [13], which are not quite practical in real world applications. To construct an efficient encryption scheme, many encryption techniques have been proposed in the so-called random oracle model [4][14][3]. The random oracle model, however, is one of the most

⋆ Supported by the National Natural Science Foundation of China (NSFC Grant No. 60525201) and 973 Project (No.2007CB807902).

S. Qing, C.J. Mitchell, and G. Wang (Eds.): ICICS 2009, LNCS 5927, pp. 107–120, 2009.

controversial issues in cryptography. A notable argument against the random oracle model was made by Canetti, Goldreich and Halevi [8] who demonstrated that there exist cryptographic schemes that are secure in the random oracle model but insecure for any instantiation of the random oracle. Recently, Leurent and Nguyen [23] showed that instantiations of full domain hash functions (random oracles) proposed in the literature are insecure. They also advocated to assess carefully the impact of potential flaws in random oracle instantiations on a system that relies on the instantiations.

To address the concern over random oracles, an obvious approach is to design an adaptive chosen ciphertext secure public key encryption scheme that does not rely on a random oracle. The often cited adaptive chosen ciphertext secure encryption scheme proposed by Cramer and Shoup [10] represents the first concrete result in this line of research. A multiple number of techniques have since been proposed and studied by many researchers. Most of these techniques, however, share a common drawback that impedes their possible adoption in practice, that is, they generally require at least a few times more computation than their random oracle based counterparts.

Given the computational superiority of random oracle based encryption, it is a shared view amongst most researchers that alternative encryption techniques without random oracles will not be able to win over practitioners unless these alternatives afford a computational speed comparable to that enjoyed by random oracle based techniques.

A major advantage of random oracle based schemes [4][14][3] lies in its simplicity. To preserve the simplicity while not relying on a random oracle for security proofs, new computational assumptions have been examined. One such effort is made by Pandey, Pass and Vaikuntanathan [25] who introduce a few complexity theoretical hardness assumptions that abstract out concrete properties of a random oracle. Based on these assumptions, they are able to solve a number of open problems, including the construction of a non-interactive concurrently non-malleable string commitment. Their results point to an interesting approach towards designing efficient and provably secure cryptographic schemes without random oracles. We note that although these assumptions are stronger than traditional cryptographic hardness assumptions, they seem quite reasonable and it is conceivable that, like many other assumptions in the field such as the decisional Diffie-Hellman assumption, this type of new assumptions may gain wider acceptance after further screening by peers in the field.

1.1 Our Contribution

The goal of this paper is to search for a public key encryption scheme that (1) does not rely on a random oracle, (2) is adaptive chosen ciphertext secure, and (3) is truly practical in that it requires no more exponentiations of large integers than does a comparable random oracle based scheme. To achieve our goal, we examine a variant of Pandey et al.'s assumption [25], called the adaptive DDH assumption. Based on the adaptive DDH assumption, a modified version of

Zheng and Seberry's encryption scheme proposed in [30] is proved to be adaptive chosen ciphertext secure without a random oracle.

Zheng and Seberry [30] proposed three simple methods for immunizing public key cryptosystems against chosen ciphertext attacks. The nature of the three methods is the same. They immunized a public key cryptosystem by appending to each ciphertext a tag that is correlated to the message to be encrypted. Soldera, Seberry and Qu [29] showed the insecurity of the first scheme, denoted by Zheng-Seberry$_{1wh}$, on some special circumstances and attempted to modify Zheng-Seberry$_{1wh}$ resulted on an El Gamal variant. Based on the Gap Diffie-Hellman assumption (GDH), Baek and Zheng [2] provided a security proof for the slightly modified version of Zheng-Seberry$_{1wh}$, in the random oracle model, leaving as an open problem proofs for the other two schemes. The focus of this paper is to modify the second scheme in [30], denoted by Zheng-Seberry$_{uh}$, so that the resultant scheme is adaptive chosen ciphertext secure (see Section 4). The scheme Zheng-Seberry$_{uh}$ is worth studying for the following reasons: First, the scheme immunizes public key encryption against adaptive chosen ciphertext attacks with the help of a universal hash function. This allows the scheme to steer clear of a one-way hash function with non standard output size, whereby successfully averting potential risks recently discovered in [23]. Second, the input length of a plaintext can be arbitrary, while the overhead of the corresponding ciphertext is a constant. As a result, the ratio between the length of the ciphertext and that of the plaintext can be close to 1 as the length of the plaintext increases.

1.2 Related Work

Hybrid encryption, which is also known as the KEM-DEM approach, applies a public key cryptosystem to encapsulate the key of a symmetric cryptosystem and the symmetric cryptosystem is subsequently used to conceal data. Cramer and Shoup first generalized the notion in their work [27][11]. Kurosawa and Desmedt [22] later presented a more efficient hybrid encryption scheme by using a KEM which is not necessarily adaptive chosen ciphertext secure. More recently, Kiltz et al. [20] improved on the Kurosawa-Desmedt technique and proposed a new approach to design adaptive chosen ciphertext secure hybrid encryption schemes without a random oracle. Compared with Kiltz et al.'s concrete scheme [20] which relies on the DDH assumption and AE-OT[1] secure symmetric encryption, our modified Zheng-Seberry$_{uh}$ scheme is conceptually much simpler and relies only on the adaptive DDH assumption. More important, this newly modified scheme requires less computation time than Kiltz et al.'s.

Another important progress was made by Hofheinz and Kiltz [18] recently. They proposed a new public key encryption scheme based on factoring. Their scheme requires only roughly two exponentiations in encryption and roughly one exponentiation in decryption. (Here, "roughly" two or one exponentiation means

[1] According to [20], a symmetric cipher is AE-OT secure if it satisfies (one-time) ciphertext indistinguishability (IND-OT) and (one-time) ciphertext integrity (INT-OT).

two or one full exponentiation and additional exponentiations with small exponents.) While for the encryption schemes based on discrete logarithm, DHIES [1] is one of the most efficient schemes without random oracle. Compared with DHIES which relies on the Oracle Diffie-Hellman (ODH) assumption together with the security of symmetric encryption and a message authentication code, our modified scheme relies only on the adaptive DDH assumption and preserves the computational efficiency of Zheng-Seberry$_{uh}$. However, it is fair to say that our modified Zheng-Seberry scheme and DHIES are comparable, each having its own pros and cons in practice. With DHIES, all three assumptions on symmetric encryption, MAC and ODH are responsible for the security of DHIES and it is relatively easy to select proper candidates to realize each function of the assumption. With our modified Zheng-Seberry scheme, the adaptive DDH assumption which is solely responsible for the security of the scheme is slightly stronger than the ODH assumption required by DHIES.

2 Preliminaries

Notation and Definition. $|X|$ denotes the length of a binary string X or the size of (or number of elements in) a set X. $x \stackrel{R}{\leftarrow} X$ denotes picking an element x from X uniformly at random. $x \leftarrow A(x_1)$ denotes the experiment of running an algorithm A on input x_1 and outputting x. $x||y$ denotes the concatenation of strings x and y. A function $\mu : \mathbb{N} \rightarrow \mathbb{R}$ is called negligible in n if for every positive polynomial $p(\cdot)$ and all sufficiently large n's, we have $\mu(n) < 1/p(n)$.

Universal hashing [9]. A family of functions $H : \{0,1\}^P \rightarrow \{0,1\}^l$ is a universal family of hash functions if, for every $x_1 \neq x_2 \in \{0,1\}^P$ and every $y_1, y_2 \in \{0,1\}^l$, the number of functions in H mapping x_1 to y_1 and x_2 to y_2 is precisely $|H|/2^{2l}$, where $|H|$ denotes the number of functions in H.

For the security proof in this paper, we need the following lemma whose proof can be found in [28].

Lemma 1. *[28] Let S_1, S_2, and S' be events defined on a probability space such that $\Pr[S_1 \wedge \neg S'] = \Pr[S_2 \wedge \neg S']$. Then we have $|\Pr[S_1] - \Pr[S_2]| \leq \Pr[S']$*

3 New Assumptions

In this section, we give the definitions of the adaptive DDH assumption and other related assumptions. First, we recall the definition of an adaptive one-to-one one-way function introduced in [25]. In the definition, an adversary picks an index tag^* and is given $y^* = f_{tag^*}(x^*)$ for a random x^* in the domain of $f_{tag^*}(x)$. The aim of the adversary is to compute x^*. The difference between the traditional definition for an one-way function and the one in [25] is that the adversary in [25] has access to a "magic oracle" $\mathcal{O}_{tag}(\cdot,\cdot)$ that on input (tag, y) with $tag \neq tag^*$, returns $f^{-1}_{tag}(y)$. The security requirement is that the adversary

can compute x^* only with a negligible probability, even if the adversary can get help from the "magic oracle". Similarly to the definition of adaptive one-way function, the definition of adaptive pseudorandom generator G_{tag} in [25] requires that the adversary can not tell the output of G_{tag^*} from the random string, even if the adversary can get help from a magic oracle that, on input (tag, y) with $tag \neq tag^*$, returns 0 or 1 depending on whether y is in the range of G_{tag} or not. Formal definitions of the adaptive one-way function and the adaptive pseudorandom generator (PRG) in [25] are given in Appendix.

Combining definitions of the adaptive one-way function and the adaptive PRG, we have the definition for a variant of the adaptive PRG. The variant is similar to the definition of the adaptive PRG except that, the adversary A has some auxiliary information $f_{tag}(x)$ on a seed x and interacts with the oracle $\mathcal{O}_{tag}(\cdot,\cdot,\cdot)$.

Definition 1. *(Auxiliary adaptive PRG) Let*

$$\mathcal{G} = \{G_{tag} : \{0,1\}^n \to \{0,1\}^{s(n)}\}_{tag \in \{0,1\}^n}$$

be a pseudorandom generator (PRG). And let $\mathcal{O}_{tag}(\cdot,\cdot,\cdot)$ denote an oracle that, on input $(tag', f_{tag'}(x), y)$ such that $tag' \neq tag$, $|tag'| = |tag|$, outputs the seed x if

- *$y = G_{tag'}(x)$, and*
- *x is consistent with its auxiliary information $f_{tag'}(x)$.*

$\mathcal{O}_{tag}(\cdot,\cdot,\cdot)$ outputs $\perp$ otherwise.

We say that the PRG $\mathcal{G}$ is adaptively secure if, for any probability polynomial-time adversary A which has the auxiliary information $f_{tag}(x)$ on the seed x and has access to the oracle $\mathcal{O}_{tag}(\cdot,\cdot,\cdot)$, there exists a negligible function μ such that for all n and for all tags $tag \in \{0,1\}^n$,

$$|Adv_A^{real} - Adv_A^{rand}| \leq \mu(n)$$

where Adv_A^{real} denotes $\Pr[x \leftarrow U_n : A^{\mathcal{O}_{tag}(\cdot,\cdot,\cdot)}(f_{tag}(x), G_{tag}(x)) = 1]$, Adv_A^{rand} denotes $\Pr[y \leftarrow U_{s(n)} : A^{\mathcal{O}_{tag}(\cdot,\cdot,\cdot)}(f_{tag}(x), y) = 1]$ and the probability is over the random choice of y and x, and the coin-tosses of A.

Definition 1 is a combination of definitions of the adaptive one-way function and the adaptive PRG in that the auxiliary information on x in Definition 1 can be replaced by an one-way function $f(x)$ and the inversion oracle $\mathcal{O}_{tag}(\cdot,\cdot,\cdot)$ plays the role of $\mathcal{O}_{tag}(\cdot,\cdot)$ in the definition of adaptive one-way function. In addition, Definition 1 implies that the adversary can not invert the one-way function $f(x)$ even with the help from $\mathcal{O}_{tag}(\cdot,\cdot,\cdot)$. A candidate construction for an auxiliary adaptive PRG, based on AES , is defined by $G_{tag}(x) = AES_x(tag||0)||AES_x(tag||1)$.

From Definition 1 and the specific number theory assumption DDH, we derive the definition of the adaptive DDH assumption.

Let $\mathcal{G}$ be a group with prime order q. $g \in \mathcal{G}$ is the generator. $G_{tag}(\cdot) : \mathcal{G} \to \{0,1\}^*$ is a pseudorandom generator. $G_{tag}(\cdot)_{[i,\ldots,j]}$ denotes the substring from the i-th bit to the j-th bit of the output of $G_{tag}(\cdot)$.

Definition 2. *(Adaptive DDH assumption) Given*

$$\{g, g^a, g^b, G_{g^a,g^b}(g^c)_{[1,\ldots,P+W]}\}$$

where $a, b, c \in Z_q$, it is computationally infeasible for any PPT distinguisher D to tell whether $c = ab$, even if D has access to an oracle $\mathcal{O}_{g^a,g^b}(\cdot,\cdot,\cdot)$, where P and W are polynomials in a security parameter.

The oracle $\mathcal{O}_{g^a,g^b}(\cdot,\cdot,\cdot)$, on input $(g^{a'}, g^{b'}, G_{g^{a'},g^{b'}}(g^{c'})_{[P+1,\ldots,P+W]})$, outputs $g^{a'b'}$ if the input $(g^{a'}, g^{b'}, G_{g^{a'},g^{b'}}(g^{c'})_{[P+1,\ldots,P+W]})$ satisfies:

- $(g^{a'}, g^{b'}, G_{g^{a'},g^{b'}}(g^{c'})_{[P+1,\ldots,P+W]}) \neq (g^a, g^b, G_{g^a,g^b}(g^c)_{[P+1,\ldots,P+W]})$
- $G_{g^{a'},g^{b'}}(g^{a'b'})_{[P+1,\ldots,P+W]} = G_{g^{a'},g^{b'}}(g^{c'})_{[P+1,\ldots,P+W]}$

Otherwise, the oracle outputs $\perp$.

That is, for all PPT D, there is a negligible function μ such that

$$|\Pr_{a,b,c\xleftarrow{R}Z_q}[D^{\mathcal{O}_{g^a,g^b}(\cdot,\cdot,\cdot)}(S) = 1] - \Pr_{a,b\xleftarrow{R}Z_q}[D^{\mathcal{O}_{g^a,g^b}(\cdot,\cdot,\cdot)}(S') = 1]| \leq \mu(n)$$

where $S = (g, g^a, g^b, G_{g^a,g^b}(g^c)_{[1,\ldots,P+W]})$, $S' = (g, g^a, g^b, G_{g^a,g^b}(g^{ab})_{[1,\ldots,P+W]})$, and n is the security parameter. A quadruple $\{g, g^a, g^b, G_{g^a,g^b}(g^c)_{[1,\ldots,P+W]}\}$ satisfying $c = ab$ is called an adaptive DDH quadruple.

Remark. Comparing with Definition 1, (g^a, g^b) is not only a tag but also represents some auxiliary information on g^{ab}. Note that it is not required that the length of the substring of $G_{g^{a'},g^{b'}}(g^{c'})$ in the adversary's query be equal to that of $G_{g^a,g^b}(g^c)_{[1,\ldots,P+W]}$. However, the length of $G_{g^{a'},g^{b'}}(g^{c'})_{[P+1,\ldots,P+W]}$, that is W, should be large enough to guarantee that the adversary can guess a "right" query only with a negligible probability. Intuitively, that means, in almost all cases, the oracle does not provide any "useful" help for the adversary. However that does not mean the adversary can not provide the right query with a non-negligible probability. In fact, the adversary can randomly pick a', b' and generate the "right" query $(g^{a'}, g^{b'}, G_{g^{a'},g^{b'}}(g^{a'b'})_{[P+1,\ldots,P+W]})$ by himself. Although the oracle's answer to such a query does not provide any useful information for the adversary, it is important for the simulation in the security proof, which will be explained later.

3.1 Relationships with Other Assumptions

HDH, ODH and SDH assumptions. Abdalla et al. introduce three related notions, which are the hash Diffie-Hellman assumption (HDH), the oracle Diffie-Hellman (ODH) assumption and the strong Diffie-Hellman assumption (SDH) [5] [1]. It seems that the ODH assumption and the adaptive DDH assumption are similar in flavor. But the adversary's power in the adaptive DDH assumption is much more restricted, as the adversary can get the help of the oracle only if it can produce a useful and "right" query, which happens with only a negligible probability.

Non-malleable pseudorandom generator. In order to prove the security ($NM-CPA) of OAEP without random oracle, Boldyreva and Fischlin [7] fully instantiated OAEP by assuming special properties of the two pseudorandom generators G and H in OAEP. To be more precise, G is a near-collision resistant trapdoor pseudorandom generator, which can recover the pre-image s of $G(s)$ according to the k least significant bits of $G(s)$; H is a non-malleable pseudorandom generator. Our adaptive DDH assumption is closely related to their assumption. To some extent, the adaptive DDH combines the above properties of G and H, and takes advantage of concrete algebra structures to replace the random oracle.

4 Modified Zheng-Seberry$_{uh}$ Scheme

First, we give the description of the modified Zheng-Seberry$_{uh}$ in Table 1. Assume that $H : \{0,1\}^* \to \{0,1\}^l$ is a family of universal hash functions. Each function in H is specified by a string of exactly Q bits. Denote by h_s the function in H that is specified by a string $s \in \{0,1\}^Q$. L denotes an encryption label, which consists of public data. In addition, m denotes a plaintext to be encrypted. Our major modification to the original Zheng-Seberry$_{uh}$ scheme [30] is to increase the output length of the pseudorandom generator by W bits. These additional W bits play the role of a tag for an ephemeral key y_A^x and will be

Table 1. The modified Zheng-Seberry$_{uh}$ Scheme

Modified Zheng-Seberry$_{uh}$ Scheme

Public parameters: A label L, a universal class of hash functions $H : \{0,1\}^* \to \{0,1\}^l$, a group $\mathcal{G}$, a generator g of $\mathcal{G}$ with order q, and an adaptively secure pseudorandom generator $G_{tag} : \mathcal{G} \to \{0,1\}^*$.
Key Generation: Choose x_A randomly in $\mathbb{Z}_q^*$ and compute $y_A = g^{x_A}$. The public key is y_A and the private key is x_A.

Encryption $E_{uhf}(y_A, m, L)$

1. $x \overset{R}{\leftarrow} \mathbb{Z}_q^*$, $r = y_A^x$
2. $c_1 = g^x$. Let $tag = (y_A, c_1)$.
3. $s = G_{tag}(r)_{[1,\ldots,Q]}$, $t = h_s(m||L)$
4. $z = G_{tag}(r)_{[Q+1,\ldots,Q+P+W]}$, $c_2 = z \oplus (m||t||0^W)$

Output the ciphertext (c_1, c_2).

Decryption $D_{uhf}(x_A, y_A, c_1, c_2, L)$

1. $r' = {c_1}^{x_A}$, $s' = G_{tag}(r')_{[1,\ldots,Q]}$, $z' = G_{tag}(r')_{[Q+1,\ldots,Q+P+W]}$,
2. $m'||t' = (c_2 \oplus z')_{[1,\ldots,P]}$, where $m' = (c_2 \oplus z')_{[1,\ldots,P-l]}$, $t' = (c_2 \oplus z')_{[P-l+1,\ldots,P]}$
3. if $h_{s'}(m'||L) = t'$ and $z'_{[P+1,\ldots,P+W]} = c_{2[P+1,\ldots,P+W]}$, then output m' as a plaintext; otherwise output $\perp$.

sent to a recipient as part of a ciphertext. In practice, in order to minimize the impact of these additional bits on the efficiency of the scheme, W should be chosen to be as short as practical. For a security level of 2^{80}, we suggest $W \geq 160$. Additionally, the pseudorandom generator $G(\cdot)$ is required to be a adaptively secure pseudorandom generator $G_{tag}(\cdot)$, where $tag = (y_A, c_1)$.

Other modifications. A public label L is employed in Table 1. Using such a label is a widely adopted practice and does not affect the security proof. Besides, the universal hash value is encrypted together with a message, which allows the use of a broader range of universal hash functions that may not necessarily hide all the information on a message.

4.1 Security Proof of the Modified Zheng-Seberry$_{uh}$ Scheme

Theorem 1. Assuming the adaptive DDH assumption holds, the modified Zheng-Seberry$_{uh}$ scheme is secure against adaptive chosen ciphertext attacks.

Proof. The main idea of the security proof is to construct three adaptive chosen ciphertext attack games, which are denoted by Game 1, Game 2 and Game 3, and prove that the adversary's views in these games are indistinguishable.

Game 1: Game 1 is a real run of a standard adaptive chosen ciphertext attack game. After the adversary submits a pair of plaintexts (m_0, m_1) in the challenge phase, the challenger creates a target ciphertext as follows: $c^* = (c_1^*, c_2^*) = (g^{x^*}, z^* \oplus (m_\beta || t^* || 0^W))$, where $t^* = h_{s^*}(m_\beta || L)$, and $\beta \overset{R}{\leftarrow} \{0,1\}$.

Game 2: Game 2 is similar to Game 1 except that the target ciphertext is modified to $c_+^{**} = (g^{x^*}, z^{**} \oplus (m_\beta || t^{**} || 0^W))$, where $s^{**} = G_{y_A, g^{x^*}}(r^{**})_{[1,\ldots,Q]}$, $z^{**} = G_{y_A, g^{x^*}}(r^{**})_{[Q+1,\ldots,Q+P+W]}$, $r^{**} \overset{R}{\leftarrow} \mathcal{G}$, $t^{**} = h_{s^{**}}(m_\beta || L)$.

Game 3: Game 3 is similar to Game 2 except that the target ciphertext is modified to $c_+^* = (g^{x^*}, u_3 \oplus (m_\beta || t_+^* || 0^W))$, where $u_2 \overset{R}{\leftarrow} \{0,1\}^Q$, $u_3 \overset{R}{\leftarrow} \{0,1\}^{P+W}$, $t_+^* = h_{u_2}(m_\beta || L)$. Since the distribution of c_+^* is independent of the choice of β, the probability that the adversary can guess β correctly in Game 3 is 1/2. That is $\Pr[Game\ 3] = 1/2$, where $\Pr[Game\ i]$ denotes the probability that the adversary wins Game i, for $1 \leq i \leq 3$.

Next, we will prove that $|\Pr[Game\ 1] - Pr[Game\ 2]| \leq \mu(k)$, where $\mu(k)$ is a negligible function. Assume for contradiction that there exists a polynomial $p(k)$ such that, for infinitely many k's, $|\Pr[Game\ 1] - \Pr[Game\ 2]| \geq 1/p(k)$, which means there exists an adversary B for Game 1 and Game 2 such that $|\Pr[Game\ 1] - \Pr[Game\ 2]|$ is non-negligible. We show how to construct a PPT algorithm A to break the adaptive DDH assumption using B, by explicitly constructing an experiment of statistical test for the adaptive DDH problem.

Given $\{g, g^a, g^b, G_{g^a, g^b}(g^c)_{[1,\ldots,Q+P+W]}\}$, A sets $y_A = g^a$ and simulates the adaptive chosen ciphertext attack game for the adversary B in the following experiment.

Experiment: A sets the target ciphertext (c_1^*, c_2^*) to

$$(g^b, G_{g^a,g^b}(g^c)_{[Q+1,\ldots,Q+P+W]} \oplus (m_\beta || h_{G_{g^a,g^b}(g^c)_{[1,\ldots,Q]}}(m_\beta||L)||0^W))$$

and uses the oracle $\mathcal{O}_{g^a,g^b}(\cdot,\cdot,\cdot)$ to answer the decryption query. Notice that the oracle $\mathcal{O}_{g^a,g^b}(\cdot,\cdot,\cdot)$ would output $\perp$ if the challenger does not propose the "right" query. More precisely, when the challenger receives the decryption query (c_1, c_2), he computes $T = c_{2[P+1,\ldots,P+W]}$ and decrypts as follows

1. If $c_1 \neq c_1^*$, the challenger makes the query (g^a, c_1, T) to the oracle $\mathcal{O}_{g^a,g^b}(\cdot,\cdot,\cdot)$.
   - If the oracle returns the answer r, the challenger can compute

$$m||t = c_{2[1,\ldots,P]} \oplus G_{g^a,g^b}(r)_{[Q+1,\ldots,Q+P]}$$
$$s = G_{g^a,g^b}(r)_{[1,\ldots,Q]}$$

     and check whether $t = h_s(m||L)$. If $t = h_s(m||L)$, the challenger returns the plaintext m. Otherwise, the challenger outputs $\perp$.
   - If the oracle outputs $\perp$ which means the (g,g^a,c_1,T) is not an adaptive DDH quadruple and the corresponding ciphertext is not valid, then the challenger outputs $\perp$.
2. If $c_1 = c_1^*$, $c_2 \neq c_2^*$, $T = T^*$, where $T^* = c^*_{2[P+1,\ldots,P+W]}$, the challenger can not get help from the oracle $\mathcal{O}_{g^a,g^b}(\cdot,\cdot,\cdot)$ and outputs $\perp$. Let $\Pr[Bad]$ denote the probability that (c_1, c_2) is a valid ciphertext such that $c_1 = c_1^*$, $c_{2[1,\ldots,P-l]} \neq c^*_{2[1,\ldots,P-l]}$, $T = T^*$. $\Pr[Bad]$ is negligible, because the adversary needs to find a c_2 satisfying

$$(c_2 \oplus z^*)_{[P-l+1,\ldots,P]} = h_{s^*}((c_2 \oplus z^*)_{[1,\ldots,P-l]}||L)$$
$$(c_2^* \oplus z^*)_{[P-l+1,\ldots,P]} = h_{s^*}((c_2^* \oplus z^*)_{[1,\ldots,P-l]}||L)$$

   According to the definition of the universal hash functions, if h is chosen uniformly from the universal class H, for every $c_{2[1,\ldots,P]}$, $c^*_{2[1,\ldots,P]} \in \{0,1\}^P$ with $c_{2[1,\ldots,P]} \neq c^*_{2[1,\ldots,P]}$, $c_{2[P-l+1,\ldots,P]}$ and $c^*_{2[P-l+1,\ldots,P]}$ are uniformly and independently distributed over $\{0,1\}^l \times \{0,1\}^l$. Therefore, the adversary can find such a c_2 only with negligible probability $1/2^l$. Otherwise, it would imply that h is not chosen uniformly from H. That means, the pseudorandom string s^* could be distinguished from a random string by an efficient algorithm with a non-negligible advantage. This is a contradiction.
3. Otherwise, the challenger outputs $\perp$.

Let $\Pr[Exp]$ denote the probability that the adversary B wins the above game in the experiment. The following claims, Claim 1 and Claim 2, show that, if $|\Pr[Game\ 1] - \Pr[Game\ 2]|$ is non-negligible, then whether $\{g,\ g^a,\ g^b,\ G_{g^a,g^b}(g^c)_{[1,\ldots,P+Q+W]}\}$ is an adaptive DDH quadruple or not can be decided with a non-negligible advantage. Due to Claim 1 and Claim 2, we have Claim 3. More details of proofs of Claim 1, Claim 2 and Claim 3 are given in Appendix.

Claim 1. If $\{g, g^a, g^b, G_{g^a,g^b}(g^c)_{[1,\ldots,P+Q+W]}\}$ is an adaptive DDH quadruple, then $|\Pr[Game\ 1] - \Pr[Exp]|$ is negligible and $|\Pr[Game\ 2] - \Pr[Exp]|$ is non-negligible.

Claim 2. If $\{g, g^a, g^b, G_{g^a,g^b}(g^c)_{[1,\ldots,P+Q+W]}\}$ is not an adaptive DDH quadruple, then $|\Pr[Game\ 2] - \Pr[Exp]|$ is negligible and $|\Pr[Game\ 1] - \Pr[Exp]|$ is non-negligible.

Claim 3. $|\Pr[Game\ 1] - \Pr[Game\ 2]| \leq \mu(k)$ if the adaptive DDH assumption holds, where $\mu(k)$ is a negligible function.

Finally, $|\Pr[Game\ 3] - \Pr[Game\ 2]|$ is also negligible if G is a secure pseudorandom generator. (Otherwise, Game 2 would serve as an efficient algorithm to distinguish the output distribution of G from the uniform distribution.) Hence, we obtain the following claim:

Claim 4. $|\Pr[Game\ 3] - \Pr[Game\ 2]| \leq \mu'(k)$ if G_{tag} is a secure pseudorandom generator, where $\mu'(k)$ is a negligible function.

From Claim 3 and Claim 4, we have

$$\begin{aligned} |\Pr[Game\ 1] - 1/2| &= |\Pr[Game\ 1] - \Pr[Game\ 3]| \\ &\leq |\Pr[Game\ 1] - \Pr[Game\ 2]| + \\ &\quad |\Pr[Game\ 2] - \Pr[Game\ 3]| \\ &\leq \mu(k) + \mu'(k) \end{aligned}$$

where $\mu(k) + \mu'(k)$ is a negligible function.

That is, the adversary can win the standard adaptive chosen ciphertext attack game with only a negligible advantage. This completes the proof of Theorem 1.

5 Instantiation

First we note that an ϵ-AXU hash function [12] can be used in place of a universal hash function. One may also use an efficient universal hash function family proposed by Bernstein [6]. Such a substitution almost does not affect the security proof. In fact, only minor revisions need to be made in the security proofs. Specifically, in Case 2 of the experiment for the security proof of the modified Zheng-Seberry$_{uh}$ scheme, the probability that the adversary can find c_2 satisfying $(c_2 \oplus z^*)_{[P-l+1,\ldots,P]} = h_{s^*}((c_2 \oplus z^*)_{[1,\ldots,P-l]}||L)$ and $(c_2^* \oplus z^*)_{[P-l+1,\ldots,P]} = h_{s^*}((c_2^* \oplus z^*)_{[1,\ldots,P-l]}||L)$ needs to be changed. To instantiate the adaptively secure pseudorandom generator $G_{tag}(\cdot)$, we can use the HMAC-based key derivation function (KDF) [21], which follows the extract-then-expand paradigm.

6 Comparison

For the modified Zheng-Seberry$_{uh}$ scheme, the length of a ciphertext is $|m| + |p| + 320$, where $|p|$ denotes the binary length of a element in $\mathcal{G}$. Thanks to the use of the pseudorandom generator and the universal hash function, the input length of the plaintext can be flexibly adjusted. With the increase in the length of a plaintext m, the ratio between the length of a ciphertext and a plaintext, $\alpha = \frac{|m|+|p|+320}{|m|}$, becomes even closer to 1. Table 2 shows a comparison of the modified Zheng-Seberry$_{uh}$ schemes with a few of the relevant encryption schemes.

Table 2. Efficiency comparison of the modified Zheng-Seberry$_{uh}$ schemes with some relevant encryption schemes. "trapdoor permutation$^+$" denotes trapdoor permutations that are uninvertible with access to a H-inverting oracle. "one-way hash$^+$" denotes adaptively secure perfectly one-way hash. "SPD-OW" denotes set partial domain one-wayness. "SKE" denotes secure symmetric encryption. "MAC" denotes secure message authentication code. "Enc Exp" ("Dec Exp") denotes the number of exponentiations or double exponentiations in encryption (decryption).

	Enc Exp	Dec Exp	Assumption	RO
Modified Zheng-Seberry$_{uh}$	2	1	adaptive DDH	No
Cramer-Shoup [10]	4	3	DDH	No
Kurosawa-Desmedt [22]	3	1	DDH, SKE	No
Hofheinz-Kiltz [17]	3	1	DDH	No
Hofheinz-Kiltz [18]	roughly 2	roughly 1	Rabin's trapdoor OWP	No
DHIES [1]	2	1	ODH, SKE, MAC	No
Pandey-Pass-Vaikuntanathan [25]	-	-	trapdoor permutation$^+$, one-way hash$^+$	No
Zheng-Seberry$_{1wh}$[2]	2	1	GDH	Yes
OAEP [15]	1	1	SPD-OW	Yes
Bellare-Rogaway [3]	-	-	trapdoor OWP	Yes

7 Concluding Remarks

We have proved the adaptive chosen ciphertext security of a modified version of Zheng and Seberry's encryption scheme that employs universal hashing. The scheme investigated in this work is based on discrete logarithms in a subgroup. A possible interesting area for further research is to investigate whether similar results can be obtained with schemes built on other computationally hard problems, such as the integer factorization problem.

References

1. Abdalla, M., Bellare, M., Rogaway, P.: The oracle Diffie-Hellman assumptions and an analysis of DHIES. In: Naccache, D. (ed.) CT-RSA 2001. LNCS, vol. 2020, pp. 143–158. Springer, Heidelberg (2001)

2. Baek, J., Zheng, Y.: Zheng and Seberry's public key encryption scheme revisited. International Journal of Information Security (IJIS) 2(1), 37–44 (2003)
3. Bellare, M., Rogaway, P.: Random oracles are practical: A paradigm for designing efficient protocols. In: First ACM Conference on Computer and Communication Security, pp. 62–73. Association for Computing Machinery (1993)
4. Bellare, M., Rogaway, P.: Optimal asymmetric encryption—how to encrypt with RSA. In: De Santis, A. (ed.) EUROCRYPT 1994. LNCS, vol. 950, pp. 92–111. Springer, Heidelberg (1995)
5. Bellare, M., Rogaway, P.: Minimizing the use of random oracles in authenticated encryption schemes. In: Han, Y., Quing, S. (eds.) ICICS 1997. LNCS, vol. 1334, pp. 1–16. Springer, Heidelberg (1997)
6. Bernstein, D.J.: Polynomial evaluation and message authentication (2007), http://cr.yp.to/papers.html#pema
7. Boldyreva, A., Fischlin, M.: On the security of OAEP. In: Lai, X., Chen, K. (eds.) ASIACRYPT 2006. LNCS, vol. 4284, pp. 210–225. Springer, Heidelberg (2006)
8. Canetti, R., Goldreich, O., Halevi, S.: The random oracle methodology, revisited. In: STOC 1998, pp. 209–218. ACM Press, New York (1998)
9. Carter, J.L., Wegman, M.N.: Universal classes of hash functions. Journal of Computer and System Sciences 18(2), 143–154 (1979)
10. Cramer, R., Shoup, V.: A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 13–25. Springer, Heidelberg (1998)
11. Cramer, R., Shoup, V.: Universal hash proofs and a paradigm for adaptive chosen ciphertext secure public-key encryption. In: Knudsen, L.R. (ed.) EUROCRYPT 2002. LNCS, vol. 2332, pp. 45–64. Springer, Heidelberg (2002)
12. den Boer, B.: A simple and key-economical unconditional authentication scheme. Journal of Computer Security 2(1), 65–71 (1993)
13. Dolev, D., Dwork, C., Naor, M.: Non-malleable cryptography. SIAM Journal on Computing 30(2), 391–437 (2000)
14. Fujisaki, E., Okamoto, T.: Secure integration of asymmetric and symmetric encryption schemes. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 537–554. Springer, Heidelberg (1999)
15. Fujisaki, E., Okamoto, T., Pointcheval, D., Stern, J.: RSA-OAEP is secure under the RSA assumption. In: Kilian, J. (ed.) CRYPTO 2001. LNCS, vol. 2139, pp. 260–274. Springer, Heidelberg (2001)
16. Gennaro, R., Shoup, V.: A note on an encryption scheme of Kurosawa and Desmedt (2005), http://www.shoup.net/papers/kdnote.pdf
17. Hofheinz, D., Kiltz, E.: Secure hybrid encryption from weakened key encapsulation. In: Menezes, A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 553–571. Springer, Heidelberg (2007)
18. Hofheinz, D., Kiltz, E.: Practical chosen ciphertext secure encryption from factoring. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 313–332. Springer, Heidelberg (2009)
19. Impagliazzo, R., Levin, L.A., Luby, M.: Pseudo-random generation from one-way functions. In: STOC 1989, pp. 12–24. ACM Press, New York (1989)
20. Kiltz, E., Pietrzak, K., Stam, M., Yung, M.: A new randomness extraction paradigm for hybrid encryption. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 589–608. Springer, Heidelberg (2009)
21. Krawczyk, H.: On extract-then-expand key derivation functions and an HMAC-based KDF (2008), http://www.ee.technion.ac.il/~hugo/kdf/

22. Kurosawa, K., Desmedt, Y.: A new paradigm of hybrid encryption scheme. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 426–442. Springer, Heidelberg (2004)
23. Leurent, G., Nguyen, P.Q.: How risky is the random oracle model. In: Halevi, S. (ed.) CRYPTO 2009. LNCS, vol. 5677, pp. 445–464. Springer, Heidelberg (2009)
24. Naor, M., Yung, M.: Public-key cryptosystems provably secure against chosen cipher-text attacks. In: STOC 1990, pp. 14–16. ACM Press, New York (1990)
25. Pandey, O., Pass, R., Vaikuntanathan, V.: Adaptive one-way functions and applications. In: Wagner, D. (ed.) CRYPTO 2008. LNCS, vol. 5157, pp. 57–74. Springer, Heidelberg (2008)
26. Rackoff, C., Simon, D.R.: Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack. In: Feigenbaum, J. (ed.) CRYPTO 1991. LNCS, vol. 576, pp. 433–444. Springer, Heidelberg (1992)
27. Shoup, V.: Using hash functions as a hedge against chosen ciphertext attack. In: Preneel, B. (ed.) EUROCRYPT 2000. LNCS, vol. 1807, pp. 275–288. Springer, Heidelberg (2000)
28. Shoup, V.: OAEP reconsidered. In: Kilian, J. (ed.) CRYPTO 2001. LNCS, vol. 2139, pp. 239–259. Springer, Heidelberg (2001)
29. Soldera, D., Seberry, J., Qu, C.: The analysis of Zheng-Seberry scheme. In: Batten, L.M., Seberry, J. (eds.) ACISP 2002. LNCS, vol. 2384, pp. 159–168. Springer, Heidelberg (2002)
30. Zheng, Y., Seberry, J.: Immunizing public key cryptosystems against chosen ciphertext attacks. IEEE journal on selected areas in communications 11(5), 715–724 (1993)

Appendix

Definition 3. *(Family of adaptive one-to-one one-way functions) [25]. A family of injective one-way functions* $\mathcal{F} = \{f_{tag} : D_{tag} \rightarrow \{0,1\}^*\}_{tag \in \{0,1\}^n}$ *is called adaptively secure if*

- *There is an efficient randomized domain sampler* D*, which on input* $tag \in \{0,1\}^n$*, outputs a random element in* D_{tag}*. There is a deterministic polynomial algorithm* M *such that for all* $tag \in \{0,1\}^n$ *and for all* $x \in D_{tag}$*,* $M(tag, x) = f_{tag}(x)$*.*
- *Let* $\mathcal{O}_{tag}(\cdot,\cdot)$ *denote an oracle that, on input* tag' *and* y*, outputs* $f^{-1}_{tag'}(y)$ *if* $tag' \neq tag$*,* $|tag'| = |tag|$ *and* $\perp$ *otherwise. The family* $\mathcal{F}$ *is adaptively secure if, for any probabilistic polynomial time adversary* A *which has access to the oracle* $\mathcal{O}_{tag}(\cdot,\cdot)$*, there exists a negligible function* μ *such that for all* n*, and for all tags* $tag \in \{0,1\}^n$*,*

$$\Pr[x \leftarrow D_{tag} : A^{\mathcal{O}_{tag}(\cdot,\cdot)}(tag, f_{tag}(x)) = x] \leq \mu(n)$$

where the probability is over the random choice of x *and the coin tosses of* A*.*

Definition 4. *(Adaptive PRG)[25]. Let a family of functions* $\mathcal{G} = \{G_{tag} : \{0,1\}^n \rightarrow \{0,1\}^{s(n)}\}_{tag \in \{0,1\}^n}$ *be a pseudorandom generator (PRG). And let*

$\mathcal{O}_{tag}(\cdot,\cdot)$ denote an oracle that, on input (tag', y) such that $tag' \neq tag$, $|tag'| = |tag|$, outputs 1 if y is in the range of $G_{tag'}$ and 0 otherwise.

We say that $\mathcal{G}$ is an adaptively secure PRG if, for any probability polynomial-time adversary A which has access to the oracle $\mathcal{O}_{tag}(\cdot,\cdot)$, there exists a negligible function μ such that for all n and for all tags $tag \in \{0,1\}^n$,

$$|\Pr[y \leftarrow G_{tag}(U_n) : A^{\mathcal{O}_{tag}(\cdot,\cdot)}(y) = 1] - \Pr[y \leftarrow U_m : A^{\mathcal{O}_{tag}(\cdot,\cdot)}(y) = 1]| \leq \mu(n)$$

where the probability is over the random choice of y and the coin-tosses of the adversary A.

Proof of Claim 1, Claim 2 and Claim 3.

To show that Claim 1 holds, we first note that if $\{g, g^a, g^b, G_{g^a,g^b}(g^c)_{[1,\ldots,P+Q+W]}\}$ is an adaptive DDH quadruple and the event Bad does not happen, then the experiment perfectly simulates Game 1 and the adversary's views in the experiment and Game 1 are identical. Hence, we have

$$\Pr[Game\ 1 \wedge \neg Bad] = \Pr[Exp \wedge \neg Bad]$$

Applying Lemma 1, we have $|\Pr[Game\ 1] - \Pr[Exp]| \leq \Pr[Bad]$, where $\Pr[Bad]$ is negligible. On the other hand, since $|\Pr[Game\ 1] - \Pr[Game\ 2]| \geq 1/p(k)$, we have

$$\begin{aligned} & |\Pr[Game\ 1] - \Pr[Exp]| + |\Pr[Game\ 2] - \Pr[Exp]| \\ \geq\ & |\Pr[Game\ 1] - \Pr[Game\ 2]| \\ \geq\ & 1/p(k). \end{aligned}$$

Therefore, $|\Pr[Game\ 2] - \Pr[Exp]|$ is non-negligible, from which Claim 1 follows.

Using a similar argument to the correctness of Claim 1, we have Claim 2. Summing up Claim 1 and Claim 2, the adaptive DDH assumption can be compromised by observing the behavior of the adversary. Specifically, if $|\Pr[Game\ 1] - \Pr[Exp]|$ is negligible, then $|\Pr[Game\ 2] - \Pr[Exp]|$ must be non-negligible. In this case, $\{g, g^a, g^b, G_{g^a,g^b}(g^c)_{[1,\ldots,P+Q+W]}\}$ must be an adaptive DDH quadruple. Likewise, if $|\Pr[Game\ 1] - \Pr[Exp]|$ is non-negligible, then $|\Pr[Game\ 2] - \Pr[Exp]|$ must be negligible. In this case, $\{g, g^a, g^b, G_{g^a,g^b}(g^c)_{[1,\ldots,P+Q+W]}\}$ must not be an adaptive DDH quadruple. These lead to Claim 3.