symmetry

MDPI

Article

An Efficient Identification Scheme Based on Bivariate Function Hard Problem

Boon Chian Tea [1,†], Muhammad Rezal Kamel Ariffin [1,*,†], Amir Hamzah Abd Ghafar [1,2,†], Siti Hasana Sapar [1,2,†] and Mohamat Aidil Mohamat Johari [1,2,†]

[1] Institute for Mathematical Research (INSPEM), Universiti Putra Malaysia, Serdang 43400, Selangor, Malaysia
[2] Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, Serdang 43400, Selangor, Malaysia
* Correspondence: rezal@upm.edu.my; Tel.: +60-3-9769-6838
† These authors contributed equally to this work.

Abstract: Symmetric cryptography allows faster and more secure communication between two entities using the identical pre-established secret key. However, identifying the honest entity with the same secret key before initiating symmetric encryption is vital since the communication may be impersonated. Tea and Ariffin, in 2014, proposed a new identification (ID) scheme based on the Bivariate Function Hard Problem (BFHP) that proved secure against impersonation under passive, active and concurrent attacks via the BFHP-hardness assumption. In this paper, we upgrade the ID scheme and improve some of its settings. Next, we provide the security proof against impersonation under active and concurrent attacks in the random oracle model via the hardness assumption of the One-More BFHP. Finally, we include an additional discussion about the computational efficiency of the upgraded ID scheme based on BFHP and present its comparison with other selected ID schemes.

Keywords: active and concurrent attacks; Bivariate Function Hard Problem; identification scheme; provable security; symmetric encryption

Citation: Tea, B.C.; Ariffin, M.R.K.; Ghafar, A.H.A.; Sapar, S.H.; Johari, M.A.M. An Efficient Identification Scheme Based on Bivariate Function Hard Problem. *Symmetry* **2022**, *14*, 1784. https://doi.org/10.3390/sym14091784

Academic Editors: Miodrag J. Mihaljevic, Chin-Ling Chen and Dumitru Baleanu

Received: 29 June 2022
Accepted: 19 August 2022
Published: 27 August 2022

1. Introduction

Symmetric cryptography allows users to perform encryption and decryption using the identical pre-established secret key. It enables simpler and quicker encryption since users do not have to authenticate the agreed secret key from any authority like those in public-key cryptography. Current practical symmetric-key algorithms include the Data Encryption Standard (DES) such as 3DES and the widely used Advanced Encryption System (AES) such as AES-128, AES-192, and AES-256, which target block ciphers. In contrast, RC4 deals with stream cipher in its symmetric encryption. Besides having a faster encryption speed, AES-256, in particular, is considered to be quantum resistant. Although faster and having a more secure encryption than public-key cryptography, identifying the honest users over the virtual network is crucial prior to initiating the encryption. As impersonation by an unauthorized entity may occur during the communication, both users should identify and prove to each other that they possess the same secret key before the encryption begins.

The main objective of an identification (ID) scheme is to enable an entity to prove and verify that it knows a secret to another entity without having to reveal it during the interaction. The protocol commonly begins with the commitment initiated by the prover that binds the interaction. Next, a challenge issuance by the verifier to the prover. Finally, a prover's responses to the challenge, followed by the verification and decision output by the verifier.

The first ID scheme [1] utilized the quadratic residuosity (QR) or square root modulo as their fundamental primitive. Its security rests on the difficulty of solving the integer factorization problem (IFP). Later, ref. [2] presented an ID scheme using the RSA encryption technique in the protocol with underlying security associated with the hardness of solving the eth-root of RSA. In the subsequent year, another novel ID scheme [3] featured the

utilization of the discrete logarithm problem (DLP) in the design. However, the concrete proof of security in cryptography was formally described only after 2000s. Many ID schemes were then proposed based on various primitives and their variants surrounding IFP and DLP. For example, the elliptic curve discrete logarithm (ECDLP) in [4] and the bilinear Diffie-Hellman (BDH) problem in [5] with their security proofs in the standard model.

Besides number-theoretic problems, there are non-number-theoretic problems from the NP-class were utilized in designing ID schemes, such as in [6–8], to name a few. Furthermore, some of the design proposals from these problems are categorized as post-quantum candidates. For instance, the multivariate quadratic (MQ) polynomial as in [9] relies on the intractability of the MQ polynomial. On the other hand, the lattice-based identification schemes proposed, respectively, [10,11] consider the hardness of solving the shortest vector problem (SVP) in lattices. The recently proposed code-based ID scheme by [12] features the hardness assumption of solving a variant of a syndrome decoding (SD) problem.

Ref. [13] in 2013 proposed an ID scheme based on the Diophantine Equation Hard Problem (DEHP), which claimed to be secure against impersonation under passive attack via the hardness assumption of DEHP. However, by carefully choosing the correct interval of the solutions of DEHP, successful impersonation is possible even without knowing the secret parameters, which results in acceptance with non-negligible probability. Later, ref. [14] proposed a new ID scheme based on the Bivariate Function Hard Problem (BFHP), a specific problem of two variables derived from DEHP [15]. With the hardness assumption of solving the BFHP to obtain the preferred private solution, the proposed ID scheme was proven secure against impersonation under passive, active and concurrent attacks.

In this paper, we upgrade the ID scheme based on BFHP from [14]. First, we refine the definition of the Bivariate Function Hard Problem (BFHP) and its underlying hardness assumption of One-More BFHP. Next, we present the upgraded ID scheme and the corresponding security proof against impersonation under active and concurrent attacks in the random oracle model via the difficulty of solving the One-More BFHP. Finally, we add on some efficiency analyses about our upgraded ID scheme.

The layout of the paper is as follows. Section 2 reviews preliminaries related to the ID scheme and its security model. Next, we outline the upgraded ID scheme based on BFHP, and formally describes the security proof of our ID scheme against impersonation under active and concurrent attacks in the random oracle model in Section 3. Section 4 includes a discussion about the efficiency of ID schemes. Lastly, we draw our conclusion in Section 5.

2. Preliminaries

2.1. Mathematical Background

This section reviews the Bivariate Function Hard Problem (BFHP) and all the related essential works to our proposal. We begin with the famous theorem due to Minkowski and its implication for the solution of modular linear equations [16].

Theorem 1. *(Minkowski). In an ω-dimensional lattice L, there exists a non-zero vector v with*

$$\| v \| \leq \sqrt{\omega}\, \det(L)^{\frac{1}{\omega}}.$$

Ref. [16] addressed the problem of solving modular linear equation

$$f(x_1, x_2, \ldots, x_m) = a_1x_1 + a_2x_2 + \cdots + a_mx_m \equiv 0 \pmod{N}$$

for some N with unknown factorization. Although such an equation has infinitely many solutions of $(y_1, y_2, \ldots, y_m) \in \mathbb{Z}_N^m$, one can expect a unique solution if $\prod_i^m X_i \leq N$, where X_i is the upper bound of each solution $|y_i| < X_i$ for $i = 1, \ldots, m$. This unique solution y_i can be recovered heuristically by computing the shortest vector in an m-dimensional lattice. In addition, if in turn $\prod_i^m X_i > N^{1+\epsilon}$, then the linear equation has N^ϵ many solutions, which

are exponential in the bit-size of N, and one has no hope of improving the bound in general. That is, there exists no efficient algorithm to output all the roots in polynomial time.

Next, we lay out the definition of BFHP from the previous work in [14], and further refine the hardness of the bivariate function and its corresponding assumption. We firstly review the proper analytic description related to BFHP.

Proposition 1. *Let $F(x_1, x_2, \ldots, x_n)$ be a multiplicative one-way function that maps $F : \mathbb{Z} \to (2^{m-1}, 2^m - 1)$ for some $m \in \mathbb{Z}$. Let F_1 and F_2 be such functions (either identical or non-identical) such that $A_1 = F_1(x_1, x_2, \ldots, x_n)$, $A_2 = F_2(x_1, x_2, \ldots, x_n)$ and $\gcd(A_1, A_2) = 1$. Let $\{u, v\} \in (2^{n-1}, 2^n - 1)$. Let $F(x, y) = A_1x + A_2y$. Let $k = n - m - 1$ where 2^k is exponentially large for any probabilistic polynomial time (PPT) adversary to sieve through all possible answers, then it is infeasible to determine $\{u, v\}$ over $\mathbb{Z}$ from $F(u, v)$.*

Proof. To prove that solving the Diophantine equation given by $F(u, v)$ is infeasible, consider the general solution for $F(x, y) = A_1x + A_2y$ given by $x = x_0 + A_2j$ and $y = y_0 - A_1j$ for any $j \in \mathbb{Z}$, where x_0 and y_0 are initial solutions for the linear Diophantine equation $F(x, y)$. Then $F(u, v)$ is said to be solved when the private preferred solution pair $\{u, v\}$ is found within the stipulated interval $\{u, v\} \in (2^{n-1}, 2^n - 1)$. In other words, one must search for specific integer j such that $2^{n-1} < u < 2^n - 1$ holds, that is $u = x_0 + A_2j$. This gives, $\frac{2^{n-1}-x_0}{A_2} < j < \frac{2^n-1-x_0}{A_2}$.

Then the difference between the upper and the lower bounds of j is approximately 2^{n-m-2}. Since $k = n - m - 1$ where 2^k is exponentially large for any probabilistic polynomial time (PPT) adversary to sieve through all possible answers. We conclude that the difference is very large, and finding the correct j is infeasible. This applies to the case of v too. □

We put forward an example to demonstrate the hardness of finding the preferred private solution from the exponentially many choices of j.

Example 1. *Let $A_1 = 191$ and $A_2 = 229$. Let $u = 41{,}234$ and $v = 52{,}167$, then $F(u, v) = 19{,}821{,}937$. Here we take $m = 8$ and $n = 16$. We now construct the parametric solutions for this BFHP. The initial points are $u_0 = 118{,}931{,}622$ and $v_0 = -99{,}109{,}685$. Then general parametric solutions are $u = u_0 + A_2j$ and $v = v_0 - A_1j$. There are approximately $286 \approx 2^9$, about $\left(\frac{2^{16}}{229}\right)$ values of j to try (i.e., the difference between upper and lower bounds), while at minimum, the value of $j \approx 2^{16}$. In fact, the correct value is $j = 519{,}172 \approx 2^{19}$.*

Definition 1. *(Bivariate Function Hard Problem (BFHP)). Let $F(x, y) = A_1x + A_2y$ be a bivariate function where x, y are unknown integers of n-bits, A_1, A_2 are public m-bits integers with $\gcd(A_1, A_2) = 1$ and 2^k is exponentially large for $k = n - m - 1$. Given $F(u, v) = A_1u + A_2v$, the BFHP is the problem in identifying the unique solution pair $F(u, v)$ which is known as the preferred private solution.*

Definition 2. *(BFHP Assumption). Suppose a BFHP experiment $\mathsf{BFHP}_{\mathcal{A},\mathcal{G}}(n)$ with parameter generator $\mathcal{G}$, algorithm $\mathcal{A}$ and security parameter 1^n (i.e., security parameter of length n written in unary [17]) is defined as follows:*

1. *Runs $\mathcal{G}(1^n)$ to obtain $F(x, y) = A_1x + A_2y$.*
2. *Choose F^* such that $F^* = A_1u + A_2v$.*
3. *$\mathcal{A}$ is given $(F(x, y), F^*)$ and outputs $\{u, v\} \in (2^{n-1}, 2^n - 1)$.*
4. *The output of the experiment is defined to be 1, i.e., $[\mathsf{BFHP}_{\mathcal{A},\mathcal{G}}(n) = 1]$ if $F(u, v) = F^*$ with correct preferred private solution pair $\{u, v\}$. Otherwise defined to be 0.*

Then the BFHP is hard relative to $\mathcal{G}$ if for all probabilistic polynomial-time algorithm $\mathcal{A}$, there exists a negligible function $\mathsf{negl}(n)$ such that,

$$\Pr[\mathsf{BFHP}_{\mathcal{A},\mathcal{G}}(n) = 1] \leq \mathsf{negl}(n).$$

Definition 3. *(One-More BFHP). Let* $F(x, y) = A_1x + A_2y$ *be a bivariate function with* $\{A_1, A_2\} \in (2^{m-1}, 2^m - 1)$ *such that* $\gcd(A_1, A_2) = 1$ *and* $\{x, y\} \in (2^{n-1}, 2^n - 1)$*. An adversary is given a challenge oracle* $\mathcal{O}_{cha}$ *that produces a random integer* $F_i \in (2^{m+n-1}, 2^{m+n} - 1)$ *when queried and BFHP oracle* $\mathcal{O}_{BFHP}$ *that provides the preferred private primitives* $\{x_i, y_i\}$ *correspond to the query* $F_i = A_1x + A_2$*. The adversary is said to win if after i queries to* $\mathcal{O}_{cha}$*, it can solve all the i challenges with strictly at most* $i - 1$ *queries to the* $\mathcal{O}_{BFHP}$*. Take note that* $F_i = F(x_i, y_i)$*.*

2.2. Identification Scheme and Security Model

An ID scheme comprises triple probabilistic polynomial-time algorithms, described as a series of challenge-and-response interactions between two entities: the Prover and the Verifier.

1. Setup: On input of security parameter 1^n, the algorithm generates public system and secret parameters. On termination, it publicizes the public system parameters and keeps the secrets securely.
2. Prove: A probabilistic algorithm that firstly outputs a random commitment (Cmt) to bind the interaction and subsequently replies to the challenge via a response (Rsp).
3. Verify: A probabilistic algorithm that firstly outputs a random challenge (Cha) to the bound commitment and next output decision based on the received response.

The security of an ID scheme lies in the advantage of an adversary (impersonator) being accepted after a certain number of interactions, either passively or actively. This is often described as a two-phase impersonation game between adversary and simulator. We outline the model to define an adversary's advantage in the security game.

1. Setup: A simulator takes in the security parameter 1^n and generates random public system and secret parameters. The public system parameters are given to the adversary while secrets are kept securely.
2. Phase 1 (Training): The adversary undergoes training based on different attackers' environments: (i) Passive: Adversary queries to the simulator, returned with valid conversation transcripts containing information about commitment, challenge, and response. (ii) Active and concurrent: Adversary takes the role of the cheating verifier and requests the simulator to prove itself. This environment works exactly like in the ID protocol.
3. Phase 2 (Impersonation): The adversary now acts as a cheating prover. It outputs a commitment that it wishes to impersonate. The impersonation is said to be successful if the adversary can convince the simulator and results in acceptance of the decision.

Definition 4. *An ID scheme is defined to be* (t, q_t, ϵ)*-secure against impersonation under passive, active or concurrent attacks if an adversary* $\mathcal{A}$ *who runs the ID protocol in time t and makes at most* q_t *queries has the negligible advantage such that*

$$\mathrm{Adv}_{\mathcal{A}}^{imp\text{-}pa/aa/ca}(n) \leq \epsilon(n),$$

where $\mathrm{Adv}_{\mathcal{A}}^{imp\text{-}pa/aa/c}(n) = \Pr[\mathcal{A} \text{ success impersonates}]$.

Lastly, we address the Reset Lemma by [18], which provides a better bound and simpler proof in relating two probabilities in the security game of an ID scheme. Under resetting the challenge by the verifier, the cheating prover yields two valid conversation transcripts that convince the verifier to accept him.

Lemma 1. *(Reset Lemma). Let P be a prover in a canonical ID protocol with V a verifier represented by* $(\mathsf{ChaSet}, \mathsf{Dec})$*, and* $(\mathcal{P}, \mathcal{V})$ *be inputs for the prover P and verifier V, respectively. Let* $\mathsf{acc}(\mathcal{P}, \mathcal{V})$ *be the probability that V accepts in its interaction with P, i.e.,*

$$acc(\mathcal{P}, \mathcal{V}) = \Pr[Dec_V(Cmt, Cha, Rsp) = 1]$$

and $res(\mathcal{P}, \mathcal{V})$ *be the probability that V accepts its interaction with P in the reset game, i.e.,*

$$res(\mathcal{P}, \mathcal{V}) = \Pr\left[\begin{array}{c} Dec_V(Cmt, Cha_1, Rsp_1) = 1 \wedge \\ Dec_V(Cmt, Cha_2, Rsp_2) = 1 \wedge \\ (Cha_1 \neq Cha_2) \end{array}\right].$$

Then,

$$acc(\mathcal{P}, \mathcal{V}) \leq \sqrt{res(\mathcal{P}, \mathcal{V})} + \frac{1}{|ChaSet_V|},$$

where $Cha \in ChaSet_V$.

Proof. See [18]. □

3. Results: The Design and Security Analysis

3.1. Efficient Identification Scheme Based on BFHP

This section presents the ID scheme based on BFHP by [14] with some upgrades to improve the setting. Two new hash functions are introduced in the Algorithm 1, with one used to mask the secret, and the other used for integrity check purposes.

Algorithm 1 Modified ID Scheme Based on BFHP.

Input: Security parameter 1^n.
Output: System parameters $\{H_1, H_2, v_1, v_2, v_3, x, e, f\}$.

Setup:

1: On input of security parameter 1^n, generates secret parameters of $\{v_1, v_2, x\} \in \left(2^{n-1}, 2^n - 1\right)$
2: Generates the following hash functions such that:
   (a) $H_1 : \{0,1\}^n \rightarrow \{0,1\}^n$,
   (b) $H_2 : \{0,1\}^{2n} \rightarrow \{0,1\}^{2n}$,
3: Computes the following parameters:
   (a) $e = v_1 + v_2$
   (b) $X = H_1(x)$,
   (c) $v_3 \equiv (1 - X)^{-1} \pmod{e}$,
   (d) $f = v_3 - v_1$.
4: Publicize $\{H_1, H_2, e, f\}$ as public system parameters and keep $\{v_1, v_2, v_3, x\}$ as secrets.

Identification Protocol:

1: Prover P picks a random integer $y \in \left(2^{n-1}, 2^n - 1\right)$ and computes $Y = y + v_2$. Sends Y as a commitment to Verifier V.
2: V picks a random challenge $c \in \{0,1\}$ and sends it to P.
3: Upon receiving challenge c, P responds V with:
   (a) $z = v_3 X^c - y - v_3 x$,
   (b) $\sigma = H_2(v_3 x)$.
4: V computes and verifies if one of the following holds:
   (a) For $c = 0$, if $\mathcal{V} \equiv f - z - Y \pmod{e}$ and $H_2(\mathcal{V}) = \sigma$,
   (b) For $c = 1$, if $\mathcal{V} \equiv f - z - Y \pmod{e}$ and $H_2(\mathcal{V} - 1) = \sigma$,

then accept P. Otherwise, reject P.

Proof of correctness.

$$\begin{aligned}
\mathcal{V} &= f - z - Y \\
&= (v_3 - v_1) - (v_3 X^c - y - v_3 x) - (y + v_2) \\
&= v_3 - v_1 - v_3 X^c + y + v_3 x - y - v_2 \\
&= v_3 - v_3 X^c - v_1 - v_2 + v_3 x \\
&= v_3(1 - X^c) - (v_1 + v_2) + v_3 x \\
&\equiv v_3(1 - X^c) + v_3 x \pmod{e} \\
&\equiv \begin{cases} v_3(1 - X^0) + v_3 x \pmod{e}; & \text{for } c = 0, \\ v_3(1 - X^1) + v_3 x \pmod{e}; & \text{for } c = 1, \end{cases} \\
&\equiv \begin{cases} v_3(1 - 1) + v_3 x \pmod{e}; & \text{for } c = 0, \\ v_3(1 - X) + v_3 x \pmod{e}; & \text{for } c = 1, \end{cases} \\
&\equiv \begin{cases} v_3 x \pmod{e}; & \text{for } c = 0, \\ 1 + v_3 x \pmod{e}; & \text{for } c = 1. \end{cases}
\end{aligned}$$

Since $v_1 + v_2 = e$ and $v_1 + v_2 \equiv 0 \pmod{e}$, it is easy to verify for $c = 0$, the resulting $\mathcal{V} \equiv v_3 x \pmod{e}$ produces a valid $H_2(\mathcal{V}) = \sigma$. For the case of $c = 1$, since $v_3 \equiv (1 - X)^{-1} \pmod{e}$ from Setup: 3(c), we have

$$\begin{aligned}
\mathcal{V} &= v_3(1 - X) + v_3 x \\
&= (1 - X)^{-1}(1 - X) + v_3 x \\
&\equiv 1 + v_3 x \pmod{e}
\end{aligned}$$

that produces a valid $H_2(\mathcal{V} - 1) = \sigma$. □

Remark 1. *The public system parameters $\{e, f\}$ are protected by Theorem 1 since the sizes of v_1, v_2, v_3, e, f are of 2^n and the product $v_1 \cdot v_2 > e$ and $v_1 \cdot v_3 > f$. This results infinitely many solutions to the linear equation in which finding the correct preferred private solutions of $\{v_1, v_2\}$ and $\{v_1, v_3\}$ are infeasible. This is indeed the BFHP defined in Definition 1*

Remark 2. *It is the prover's main objective to prove that he knows the secret x (i.e., step 3 in the identification protocol) without relaying it to the verifier. Since the published public system parameters $\{e, f\}$ contain secret parameters of $\{v_1, v_2, v_3\}$, it should be infeasible for an adversary to extract $\{v_1, v_2, v_3\}$ from $\{e, f\}$. Observe that from Setup:*

$$\begin{aligned}
e &= v_1 + v_2 \\
f &= v_3 - v_1,
\end{aligned}$$

each individual equation has the secret parameters protected by the BFHP with three unknown variables $\{v_1, v_2, v_3\}$. If these secret variables are solved from the public system parameters, then the preferred private solution is found.

Remark 3. *In this case, one can treat all (or part of) the secret parameters $\{v_1, v_2, v_3, x\}$ as the agreed symmetric keys (cipher keys) between two users. Once both users succeed in identifying each other that they acquired the identical secret keys, they can next perform encryption using some practical symmetric encryption available.*

3.2. Security Proof of The Identification Scheme Based on BFHP

In this section, we construct the security proof of the upgraded ID scheme based on BFHP with the refined One-More BFHP assumption (Definition 3). We considers the security against impersonation under active and concurrent attacks as this is a more relevant and stronger security notion. In addition, we refer to the Reset Lemma [18] for the probability analysis.

Theorem 2. *Let* $\varepsilon = \mathrm{Adv}_{\mathcal{I}}^{\textit{imp-aa/ca}}(n)$ *be the advantage of impersonation under active and concurrent attacks by impersonator* $\mathcal{I}$*, and* $\varepsilon' = \mathrm{Adv}_{\mathcal{A}'}^{\textit{One-More BFHP}}(n)$ *be the advantage of algorithm* $\mathcal{A}'$ *solving the One-More BFHP. Then the identification scheme based on BFHP is* (t, q_I, ε)*-secure against impersonation under active and concurrent attacks in the random oracle model if solving the One-More BFHP is* (t, q_I, ε')*-hard where:*

$$\varepsilon \leq \frac{1}{2} + \sqrt{\left(\varepsilon' - \frac{1}{2^{n-2}}\right) \cdot \left(\frac{2^{n-2}}{2^{n-2} - q_I}\right)},$$

with q_I *denotes the number of identification queries.*

Proof. We assume that if there exists an Impersonator $\mathcal{I}$ who can (t, q_I, ε)-break the scheme, then there exists an efficient probabilistic polynomial-time algorithm $\mathcal{A}'$ that can (t, q_I, ε')-solve the One-More BFHP. $\mathcal{A}'$ attempts to simulate a challenger for $\mathcal{I}$.

1. Setup: $\mathcal{A}'$ firstly access to $\mathcal{O}_{\mathsf{Cha}}$ which output the following initial challenge set:

$$\begin{aligned} X &= H_1(x) \\ e_0 &= v_{0,1} + v_2 \\ v_3 &\equiv (1 - X)^{-1} \pmod{e_0}. \\ f &= v_3 - v_{0,1} \end{aligned} \tag{1}$$

   together with two hash functions $H_1 : \{0,1\}^n \to \{0,1\}^n$ and $H_2 : \{0,1\}^{2n} \to \{0,1\}^{2n}$. The public system parameters $\{H_1, H_2, e_0, f\}$ is passed to $\mathcal{A}'$. These $\{H_1, H_2, e_0, f\}$ is then sent to $\mathcal{I}$, with H_1 and H_2 modeled as random oracles controlled by $\mathcal{A}'$. Note that $\mathcal{A}'$ does not know the secret parameters of $\{v_{0,1}, v_2, v_3, x, X\}$.
2. Query phase: The interaction between $\mathcal{A}'$ and $\mathcal{I}$ involves the simulation of random oracles and identification queries.
   (a) Random oracle query: $\mathcal{A}'$ initially prepare two empty lists of H_1-list and H_2-list, which function to receive and reply queries from $\mathcal{I}$.
      i. H_1-list: This list has the form of $\langle x_i, H_1(x_i)\rangle$. When $\mathcal{I}$ queries x_i, $\mathcal{A}'$ checks whether such $H_1(x_i)$ is in the list. If it does, it replies it to $\mathcal{I}$. Otherwise, $\mathcal{A}'$ randomly samples $H_1(x_i) \leftarrow \{0,1\}^n$ and replies it to $\mathcal{I}$. Lastly, $\mathcal{A}'$ stores the new pair of $\langle x_i, H_1(x_i)\rangle$ into H_1-list.
      ii. H_2-list: This list is in the form of $\langle x_i, H_2(x_i)\rangle$. When $\mathcal{I}$ queries x_i, $\mathcal{A}'$ checks if such $H_2(x_i)$ exists. If it does, it replies it to $\mathcal{I}$. Otherwise, $\mathcal{A}'$ randomly chooses $H_2(x_i) \leftarrow \{0,1\}^{2n}$ and replies it to $\mathcal{I}$. $\mathcal{A}'$ lastly stores the new pair of $\langle x_i, H_2(x_i)\rangle$ into H_2-list.

      Notice that $\mathcal{A}'$ simulates the above two random oracle queries perfectly, as both the replied answers are uniformly distributed.
   (b) Identification query: $\mathcal{I}$ acts as cheating verifier and requests $\mathcal{A}'$ to prove itself:
      i. Commitment: Upon query by $\mathcal{I}$, $\mathcal{A}'$ query $\mathcal{O}_{\mathsf{Cha}}$ for random challenge of $e_k = v_{k,1} + v_2$ and replies to $\mathcal{I}$.
      ii. Challenge: $\mathcal{I}$ outputs a random challenge $c_k \in \{0,1\}$ upon receiving e_k and sends it to $\mathcal{A}'$.
      iii. Response: After receiving the challenge c from $\mathcal{I}$, $\mathcal{A}'$ queries $(f - e_k - c_k - x_k)$ such that $x_k \leftarrow_R \{0,1\}^{2n}$ to the $\mathcal{O}_{\mathsf{BFHP}}$ which replied with a response $z_k = v_3(H_1(x))^{c_k} - v_{k,1} - x_k$ and $\sigma_k = H_2(x_k)$. This (z_k, σ_k) is then sent to $\mathcal{I}$. $\mathcal{A}'$ next increases k by 1.

   Referring to the Response query above, the $(f - e_k - c_k - x_k)$ queried by $\mathcal{A}'$ can be re-expressed as

$$\begin{aligned}(f - e_k - c_k - x_k) &= (v_3 - v_{0,1}) - (v_{k,1} + v_2) - c_k - x_k \\ &= v_3 - v_{k,1} - (v_{0,1} + v_2) - c_k - x_k \\ &= v_3 - v_{k,1} - e_0 - c_k - x_k + \left(v_3(H_1(x))^{c_k} - v_3(H_1(x))^{c_k}\right) \\ &= \left(v_3 - v_3(H_1(x))^{c_k}\right) + \left(v_3(H_1(x))^{c_k} - v_{k,1}\right) - e_0 - c_k - x_k \\ &\equiv \left(v_3 - v_3(H_1(x))^{c_k}\right) + \left(v_3(H_1(x))^{c_k} - v_{k,1}\right) - c_k - x_k \pmod{e_0}\end{aligned}$$

Since $c_k \in \{0, 1\}$, there are two possible cases, i.e.,

$$\begin{aligned}&\left(v_3 - v_3(H_1(x))^{c_k}\right) + \left(v_3(H_1(x))^{c_k} - v_{k,1}\right) - c_k - x_k \\ = &\begin{cases} (v_3 - v_3) + (v_3 - v_{k,1}) - 0 - x_k & ; \text{ for } c_k = 0, \\ (v_3 - v_3(H_1(x))) + (v_3 X - v_{k,1}) - 1 - x_k & ; \text{ for } c_k = 1. \end{cases} \quad (2) \\ = &\begin{cases} v_3 - v_{k,1} - x_k & ; \text{ for } c_k = 0, \\ (v_3 - v_3(H_1(x))) + (v_3 X - v_{k,1}) - 1 - x_k & ; \text{ for } c_k = 1. \end{cases} \quad (3) \\ \equiv &\begin{cases} v_3 - v_{k,1} - x_k \pmod{e_0}; & \text{for } c_k = 0, \\ v_3 X - v_{k,1} - x_k \pmod{e_0}; & \text{for } c_k = 1. \end{cases} \quad (4)\end{aligned}$$

From (1), $v_3 \equiv (1 - (H_1(x)))^{-1} \pmod{e_0}$ can be rewritten as $v_3 - v_3(H_1(x)) \equiv 1 \pmod{e_0}$, which is then used to simplify further from (3) to (4) for $c_k = 1$. Therefore, it implies that $z_k = v_3(H_1(x))^{c_k} - v_{k,1} - x_k$ and $\sigma_k = H_2(x_k)$ output by $\mathcal{O}_{\mathsf{BFHP}}$ is a valid response. This query phase is carried out for some time t until $\mathcal{I}$ readies to terminate and enter the impersonation phase.

3. Impersonation phase: Now, $\mathcal{I}$ behaves as cheating prover and tries to convince $\mathcal{A}'$ to accept it.
   (a) Commitment: $\mathcal{I}$ sends commitment Y^* in which he wishes to impersonate to $\mathcal{A}'$.
   (b) Challenge: $\mathcal{A}'$ checks if $Y^* \notin \{e_1, \ldots, e_t\}$, then it outputs a random challenge $c^* \in \{0, 1\}$ to $\mathcal{I}$. Otherwise, it aborts.
   (c) Response: $\mathcal{I}$ replies the challenge with response z^* and its corresponding signature σ^* to $\mathcal{A}'$. $\mathcal{A}'$ checks $\mathcal{V}^* = f - z^* - Y^* \pmod{e_0}$ and verifies the correctness of signature σ^*.

   Via resetting $\mathcal{I}$ to two different challenges c_1 and c_2, $\mathcal{A}'$ then obtains two valid transcripts of $\{Y^*, c_1^*, z_1^*, \sigma_1^*\}$ and $\{Y^*, c_2^*, z_2^*, \sigma_2^*\}$. The validity of the received transcripts can be verified through the protocol, $\mathcal{A}'$ next search through the H_2-list for the x_i that produced the valid signature σ_j^* for $j = 1, 2$. If such x_i is found, then $\mathcal{A}'$ proceed to search the H_1-list for the corresponding $H_1(x_i)$. This enables $\mathcal{A}'$ to extract the secret $X = H_1(x_i)$ which has the probability at least $\left(\varepsilon - \frac{1}{2}\right)^2$ following the Reset Lemma [18].
   This further enables $\mathcal{A}'$ to compute $v_{0,1}, v_2, v_3$ which finally used to solve all $v_{k,1}$ for $1 \leq k \leq t$, i.e.,

$$\begin{aligned} v_3 &\equiv (1 - X)^{-1} \pmod{e_0} \\ v_{0,1} &= v_3 - f \\ v_2 &= e_0 - v_{0,1} \\ v_{k,1} &= e_k - v_2. \end{aligned}$$

   From this point, $\mathcal{A}'$ has successfully output all the solutions to $(t + 1)$ challenges of $\{v_{0,1}, v_{1,1}, \ldots, v_{t,1}\}$ by querying only t queries of e_k to $\mathcal{O}_{\mathsf{BFHP}}$. This completes the simulation.
4. Probability study: The probability for the ID scheme against impersonation under active and concurrent attacks rests on $\mathcal{A}'$ winning the game, i.e., the successful impersonation by $\mathcal{I}$ that yield in acceptance which enables $\mathcal{A}'$ to extract the secret. Consider the following possible scenarios:

(a) Regardless of whether the game aborts, $\mathcal{A}'$ guesses the correct $X = H_1(x)$ from the interval of $\left(2^{n-1}, 2^n - 1\right)$ that output the solutions to the One-More BFHP. The corresponding probability to such an event is therefore

$$\begin{aligned} \Pr\left[\mathcal{A}' \text{ solves One}-\text{More BFHP}\right] &= \Pr\left[\mathcal{A}' \text{ solves } e_0 \text{ via guessing } H_1(x)\right] \\ &= \frac{1}{2^n - 1 - 2^{n-1}} \\ &\approx \frac{1}{2^{n-2}}. \end{aligned}$$

(b) The game does not abort, and $\mathcal{A}'$ found the corresponding secret $H_1(x)$ from the queried H_1-list that solves the One-More BFHP. This implies that $\mathcal{A}'$ successfully found the solution to the One-More BFHP with strictly less queries to $\mathcal{O}_{\mathsf{BFHP}}$ than to $\mathcal{O}_{\mathsf{Cha}}$. Via Reset Lemma,

$$\begin{aligned} & \Pr\left[\mathcal{A}' \text{ solves One}-\text{More BFHP}\right] \\ = \; & \Pr\left[\mathcal{A}' \text{ solves } e_0 \text{ via extracting } H_1(x) \;\wedge\; \neg\text{Abort}\right] \\ = \; & \Pr\left[\mathcal{A}' \text{ solves } e_0 \text{ via extracting } H_1(x) \mid \neg\text{Abort}\right] \cdot \Pr[\neg\text{Abort}] \\ \geq \; & \left(\varepsilon - \frac{1}{2}\right)^2 \cdot \left(\frac{2^{n-2} - q_I}{2^{n-2}}\right). \end{aligned}$$

Since $\Pr[\mathcal{A}' \text{ wins}] = \mathrm{Adv}_{\mathcal{A}'}^{\text{One-More BFHP}}(n) = \varepsilon'$, putting everything mathematically, we have

$$\begin{aligned} \Pr\left[\mathcal{A}' \text{ wins}\right] &= \Pr\left[\mathcal{A}' \text{ solves One}-\text{More BFHP}\right] \\ &= \Pr\left[\mathcal{A}' \text{ solves } e_0 \text{ via guessing } H_1(x)\right] \\ &\quad + \Pr\left[\mathcal{A}' \text{ solves } e_0 \text{ via extracting } H_1(x) \mid \neg\text{Abort}\right] \cdot \Pr[\neg\text{Abort}] \\ &\geq \frac{1}{2^{n-2}} + \left(\varepsilon - \frac{1}{2}\right)^2 \cdot \left(\frac{2^{n-2} - q_I}{2^{n-2}}\right) \end{aligned}$$

In other words, we have

$$\frac{1}{2^{n-2}} + \left(\varepsilon - \frac{1}{2}\right)^2 \cdot \left(\frac{2^{n-2} - q_I}{2^{n-2}}\right) \leq \varepsilon',$$

and that

$$\begin{aligned} \left(\varepsilon - \frac{1}{2}\right)^2 &\leq \left(\varepsilon' - \frac{1}{2^{n-2}}\right) \cdot \left(\frac{2^{n-2}}{2^{n-2} - q_I}\right) \\ \varepsilon - \frac{1}{2} &\leq \sqrt{\left(\varepsilon' - \frac{1}{2^{n-2}}\right) \cdot \left(\frac{2^{n-2}}{2^{n-2} - q_I}\right)} \\ \varepsilon &\leq \frac{1}{2} + \sqrt{\left(\varepsilon' - \frac{1}{2^{n-2}}\right) \cdot \left(\frac{2^{n-2}}{2^{n-2} - q_I}\right)}. \end{aligned}$$

This completes the security proof of the upgraded ID scheme. □

4. Discussion

4.1. Computational Cost of The Efficient Identification Scheme Based on BFHP

We analyze the efficiency of the ID scheme based on BFHP, considering the computational cost and its asymptotic complexity order. As the ID scheme mainly consists of simple arithmetic and modular addition and multiplication operations, Table 1 shows the corresponding computational costs for each algorithm. The scheme does not involve

modular exponentiation operation, and hence its asymptotic complexity order is bounded only by $\mathcal{O}(n^2)$, due to modular multiplication and inversion operations.

Table 1. Computational cost of the upgraded ID scheme based on BFHP.

Operation	Addition/ Subtraction	Multiplication	Modular Addition/ Subtraction	Modular Multiplication/ Inversion	Hashing
Setup	2	0	1	1	1
Prove	3	2	0	0	1
Verify	1	0	2	0	1

4.2. Comparative Analysis

We next compare our ID scheme with the five selected well-known ID schemes by [1–3,9,10]. Our choice is that the former three ID schemes were among the pioneered ID schemes that are number-theoretic based, i.e., quadratic residuosity, e^{th}-the root of RSA, and discrete logarithm problem, respectively. The latter two featured post-quantum primitives of non-number theoretic type, i.e., shortest vector problem (SVP) in lattice and multivariate quadratic (MQ) polynomial. To simplify the comparison, we tabulated all the abovementioned ID schemes by considering total computational costs and their asymptotic complexity orders in Table 2.

Table 2. Comparative analysis of selected ID schemes for selected aspects.

ID Schemes	Computational Cost	Particular Parameter Characteristic	Asymptotic Complexity Order	Underlying Security Assumption	Data Size Transmitted
[1]	$1M$, $1M_{\mathsf{Mod}}$, $3E_{\mathsf{Mod}}$	$N = pq$	$\mathcal{O}(n^3)$ where $n = \log N$	QR	$2n$
[2]	$1M$, $1M_{\mathsf{Mod}}$, $3E_{\mathsf{Mod}}$	$ed \equiv 1 \pmod{\phi(N)}$ where $N = pq$	$\mathcal{O}(n^3)$ where $n = \log N$	eth-root RSA	$2n$
[3]	$1A_{\mathsf{Mod}}$, $1M_{\mathsf{Mod}}$, $3E_{\mathsf{Mod}}$	$p \geq 2^{140}, q \geq 2^{512}$ such that $q \vert p-1$	$\mathcal{O}(n^3)$ where $n = \log q$	DLP	n
3-pass [9]	$2A_{\mathsf{Poly}}$, $6S_{\mathsf{Poly}}$	$\mathcal{MQ}(84, 80, \mathbb{F}_2)$	N/A	MQ Polynomial	29,640
5-pass [9]	$1A_{\mathsf{Poly}}$, $6S_{\mathsf{Poly}}$, $4M_{\mathsf{Scl}}$	$\mathcal{MQ}(45, 30, \mathbb{F}_{2^4})$	N/A	MQ Polynomial	26,565
[10]	$1A_{\mathsf{Mtx}}$, $3M_{\mathsf{Mtx}}$	$m = \lceil 4n \log n \rceil$	$\mathcal{O}(\lceil 4n \log n \rceil \log n)$	SVP	$\lceil 4n \log n \rceil$
Our ID Scheme	$2A$, $4S$, $2M$, $3S_{\mathsf{Mod}}$, $1I_{\mathsf{Mod}}$, $3H$	Group $\mathbb{Z}_{(2^{n-1}, 2^n-1)}$	$\mathcal{O}(n^2)$	BFHP	$2n$

Legends: (i) A = Addition, (ii) S = Subtraction, (iii) M = Multiplication, (iv) A_{Mod} = Modular Addition, (v) S_{Mod} = Modular Subtraction, (vi) M_{Mod} = Modular Multiplication, (vii) I_{Mod} = Modular Inversion, (viii) E_{Mod} = Modular Exponentiation, (ix) H = Hashing, (x) A_{Mtx} = Matrix Addition, (xi) M_{Mtx} = Matrix Multiplication, (xii) M_{Scl} = Scalar Multiplication, (xiii) A_{Poly} = Polynomial Addition, (xiv) S_{Poly} = Polynomial Subtraction, (xv) m = Matrix Dimension.

As the reader may notice from Table 2, we do not include asymptotic complexity orders for the ID scheme by [9] as its efficiencies rely on the system parameters, such as the size of public-private keys and the number of rounds performed. Readers may refer to Section 5.2 [9] for a detailed discussion.

Each of the selected ID schemes in our comparative analysis has its strengths and practical value. The first three ID schemes [1–3] based on number-theoretic hard problems are well-established. In comparison, the subsequent two ID schemes [9,10] are constructed using well-agreed post-quantum primitives, which have recently gained much attention in the field. While the primitive used in our ID scheme based on BFHP is immature in that it

is not widely studied, it can nevertheless be practical if more research is conducted on it in the future. One benefit of considering BFHP is that it allows two secrets in a single mathematical equation, i.e., $e = v_1 + v_2$ and $f = v_3 - v_1$ with many possible solutions, different from the conventional mathematical hard problems that permit only single and unique solution. Furthermore, the simpler mathematical structure comprising only modular additions and multiplications significantly reduces the computation time compared to those involving modular exponentiation operations. Hence, besides utilizing it in public-key cryptography (the proposed ID scheme in this paper), such secrets can be used in symmetric encryption schemes.

5. Conclusions

In this paper, we upgraded the ID scheme based on BFHP by [14]. We proved the security of the upgraded ID scheme against impersonation under active and concurrent attacks via the assumption that solving the One-More BFHP is difficult. In addition, the upgraded ID scheme is efficient as it does not involve mathematical operations with higher computational complexity. This is evident in the significant speed-up of the algorithm vis-à-vis the data size transmitted and the simplicity of the mathematical structure for practical deployment, which would give better efficient throughput over the bandwidth. This is clearly explained in Table 2. Once the fundamental security is proven to be secure with desirable characteristics for practical deployment, the next natural way forward would be to analyze physical attacks such as timing attacks, power analysis, power consumption attacks, side-channel analysis, and other directions.

Author Contributions: Conceptualization, B.C.T. and M.R.K.A.; methodology, B.C.T. and M.R.K.A.; validation, M.R.K.A.; formal analysis, B.C.T.; investigation, B.C.T., M.R.K.A., A.H.A.G., S.H.S. and M.A.M.J.; resources, M.R.K.A.; writing—original draft preparation, B.C.T.; writing—review and editing, B.C.T., M.R.K.A., A.H.A.G., S.H.S. and M.A.M.J.; visualization, B.C.T., M.R.K.A., A.H.A.G., S.H.S. and M.A.M.J.; supervision, M.R.K.A.; project administration, M.R.K.A.; funding acquisition, M.R.K.A. All authors have read and agreed to the published version of the manuscript.

Funding: The present research was partially supported by the Universiti Putra Malaysia Grant with Project Number GP-IPS/2018/9657300. It is also partially supported by Mediterranea Universiti of Reggio Calabria (UNIRC) Research Grant (UPM/INSPEM/700-3/1/GERANANTARABANGSA/6380071-10065).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The first author would like to further express appreciation to the Institute for Mathematical Research (INSPEM), Universiti Putra Malaysia (UPM), and Ministry of Higher Education (MOHE) for giving the opportunity to conduct this research.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BDH	Bilinear Diffie-Hellman
BFHP	Bivariate Function Hard Problem
Cha	Challenge
Cmt	Commitment
DEHP	Diophantine Equation Hard Problem
(EC)DLP	(Elliptic Curve) Discrete Logarithm Problem
ID	Identification
IFP	Integer Factorization Problem
imp-aa/ca	Impersonation under active attack/concurrent attack
MQ	Multivariate Quadratic
NP	Nondeterministic Polynomial
QR	Quadratic Residuosity

RSA	Rivest-Shamir-Adleman
SD	Syndrome Decoding
SVP	Shortest Vector Problem
Rsp	Response

References

1. Fiat, S.; Shamir, A. How to Prove Yourself: Practical Solutions to Identification and Signature Problem. In Proceedings of the Advances in Cryptology, CRYPTO'86, Santa Barbara, CA, USA, 11–15 August 1986; Volume 263, pp. 186–194. [CrossRef]
2. Guillou, L.; Quisquater, J.J. A "Paradoxical "Identity-Based Signature Scheme Resulting from Zero Knowledge. In Proceedings of the Advances in Cryptology, CRYPTO'88, Santa Barbara, CA, USA, 21–25 August 1988; Volume 403, pp. 216–231.
3. Schnorr, C.P. Efficient Identification and Signature for Smart Card. In Proceedings of the Advances in Cryptology, CRYPTO'89, Santa Barbara, CA, USA, 20–24 August 1989; Volume 435, pp. 239–252.
4. Popescu, C. An Identification Scheme Based on The Elliptic Curve Discrete Logarithm Problem. In Proceedings of the Fourth International Conference/Exhibition on High Performance Computing in the Asia-Pacific Region, Beijing, China, 14–17 May 2000; Volume 2, pp. 624–625.
5. Kim, M.; Kim, K. A New Identification Scheme Based on the Bilinear Diffie-Hellman Problem. In Proceedings of the Information Security and Privacy, Melbourne, Australia, 3–5 July 2002; Volume 2384, pp. 362–378.
6. Shamir, A. An Efficient Identification Scheme Based on Permuted Kernels (extended abstract). In Proceedings of the Advances in Cryptology, CRYPTO'89, Santa Barbara, CA, USA, 20–24 August 1989; Volume 435, pp. 606–609.
7. Stern, J. A New Identification Scheme Based on Syndrome Decoding. In Proceedings of the Advances in Cryptology, CRYPTO'93, Santa Barbara, CA, USA, 22–26 August 1993; Volume 773, pp. 13–21.
8. Pointcheval, D. A New Identification Scheme Based on the Perceptrons Problem. In Proceedings of the Advances in Cryptology, EUROCRYPT'95, St. Malo, France, 21–25 May 1995; Volume 921, pp. 319–328.
9. Sakumoto, K.; Shirai, T.; Hiwatari, H. Public-Key Identification Schemes Based on Multivariate Quadratic Polynomials. In Proceedings of the Advances in Cryptology, CRYPTO 2011, Santa Barbara, CA, USA, 14–18 August 2011; Volume 6841, pp. 706–723.
10. Lyubashevsky, V. Lattice-Based Identification Schemes Secure Under Active Attacks. In Proceedings of the Public Key Cryptography, PKC 2008, Barcelona, Spain, 9–12 March 2008; Volume 4939, pp. 162–179.
11. Akleylek, S.; Soysaldı, M. A New 3-pass Zero-Knowledge Lattice-Based Identification Scheme. In Proceeding of 2019 4th International Conference on Computer Science and Engineering (UBMK), Samsun, Turkey, 11–15 September 2019; pp. 409–413.
12. Bettaieb, S.; Bidoux, L.; Blazy, O.; Gaborit, P. Zero-Knowledge Reparation of the Véron and AGS Code-Based Identification Schemes. In Proceedings of the 2021 IEEE International Symposium on Information Theory (ISIT), Virtual, 12–20 July 2021; pp. 55–60.
13. Tea, B.C.; Ariffin, M.R.K.; Chin, J.J. An Efficient Identification Scheme in Standard Model Based on the Diophantine Equation Hard Problem. *Malays. J. Math. Sci.* **2013**, 7, 87–100.
14. Tea, B.C.; Ariffin, M.R.K. An Identification Scheme Based on Bivariate Function Hard Problem. In Proceedings of the 4th International Cryptology and Information Security Conference (CRYPTOLOGY2014), Putrajaya, Malaysia, 24–26 June 2014; pp. 48–56.
15. Ariffin, M.R.K.; Asbullah, M.A.; Abu, N.A.; Mahad, Z. A New Efficient Asymmetric Cryptosystem Based on the Integer Factorization Problem of $N = p^2q$. *Malays. J. Math. Sci.* **2013**, 7, 19–37.
16. Herrmann, M.; May, A. Solving Linear Equations Modulo Divisors: On Factoring Given Any Bits. In Proceedings of the Advances in Cryptology, ASIACRYPT 2008, Melbourne, Australia, 7–11 December 2008; Volume 5350, pp. 406–424.
17. Katz, J.; Lindell, Y. *Introduction to Modern Cryptography (Cryptography and Network Security Series)*, 2nd ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2015; pp. 312–320.
18. Bellare, M.; Palacio, A. GQ and Schnorr Identification Schemes: Proofs of Security against Impersonation under Active and Concurrent Attacks. In Proceedings of the Advances in Cryptology, CRYPTO 2002, Santa Barbara, CA, USA, 18–22 August 2002; Volume 2442, pp. 162–177.