Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Privacy protection in on-line shopping for electronic documents ☆

Yu-Chi Chen, Gwoboa Horng *, Chang-Chin Huang

Department of Computer Science and Engineering, National Chung Hsing University, Taiwan

ARTICLE INFO

Article history:
Received 31 December 2009
Received in revised form 17 September 2013
Accepted 10 February 2014
Available online xxxx

Keywords:
Cryptography
Privacy protection
Blind decoding
RSA
Oracle attack

ABSTRACT

Blind decoding schemes are proposed as tools for protecting customers' privacy in on-line shopping for electronic documents in such a way that the documents' owner has no way of knowing which documents the customers have purchased. However, most of the blind decoding schemes suffer from the oracle problem where an adversary may obtain useful information, such as the signature of a message, by interacting with the documents' owner (the oracle) using tricky requests. In this paper, a secure blind decoding scheme based on RSA scheme is proposed. Its security against one more decoding key and oracle attacks is formally proved.

1. Introduction

Data encryption and digital signature are two important cryptographic primitives. The former ensures secrecy of data and the later provides authenticity of messages. In many application scenarios, such as on-line shopping and electronic voting, the property of untraceability of messages is also essential. In 1982, blind signature schemes were introduced to provide authenticity and untraceability of messages in realizing digital cash systems [6]. A similar concept called blind decoding was proposed for protecting customers' privacy in on-line shopping for electronic documents [22].

Let M be the message or ciphertext to be blindly signed or decoded, Alice be the sender of the message or ciphertext, and Bob be the signer or decoder. It is possible that Alice might prepare a tricky message M and try to deduce Bob's secret key from the value returned by Bob, or transform it into desired information related to Bob's secret key which she cannot compute alone. Therefore, Bob must never let himself be used as an oracle [1]. He must assure himself that the message contains accurate information to prevent abusing or the oracle attack [14]. The well-known cut-and-choose methodology [7] can be used to prevent abusing in blind signature schemes. However, it is not very useful in defeating the oracle attack in blind decoding schemes. Solutions utilizing the transformability of digital signatures of the ElGamal scheme [13] are proposed in [14,11]. Another solution based on RSA scheme [18] is proposed in [11]. In this paper, a new blind decoding scheme based on RSA scheme is proposed.

Another direction of research motivated by the growing concern about the user's privacy when querying a large commercial database is the notion of private information retrieval (PIR for short). PIR was introduced by Chor, Goldreich, Kushilevitz and Sudan [9]. Ideally, the PIR problem consists of devising a communication protocol involving just two parties, the

☆ A preliminary version appeared in ISA 2009 [8].

* Corresponding author. Tel.: +886 4 22840497x912; fax: +886 4 22853869.
E-mail addresses: wycchen@ieee.org (Y.-C. Chen), gbhorng@cs.nchu.edu.tw (G. Horng), phd9308@cs.nchu.edu.tw (C.-C. Huang).

http://dx.doi.org/10.1016/j.ins.2014.02.065
0020-0255/© 2014 Elsevier Inc. All rights reserved.

database and the user, each having a secret input. The database's secret input is called the data string, an n-bit string $B = b_1b_2\cdots b_n$. The user's secret input is an integer i between 1 and n. The protocol should enable the user to learn b_i in a communication-efficient way and at the same time hide i from the database. The main research goal in private information retrieval protocols considered in the last several years is to reduce the communication complexity [5,10,19]. We note that the bit string B is not encrypted in PIR.

The rest of this paper is organized as follows. In Section 2, we review a blind decoding scheme proposed in [11]. It utilizes two pairs of public-secret exponents for the same modulus. An attack against the scheme and a potential weakness of the scheme are also discussed. In Section 3, we propose a new scheme based on RSA systems and analyze its security in Section 4. Finally, conclusions are given in Section 5.

2. Related work

A blind decoding scheme based on RSA scheme is very similar to blind signature scheme developed by Chaum [6]. Let n be the modulus, e be the public exponent of Bob's RSA signature scheme. Chaum's RSA blind signature scheme with parameter (e, n) is as follows, where Alice is the requester of a blind signature for message M. Following [15], we denote by $x \bmod n$ to mean the unique remainder $b, 0 \leqslant b \leqslant n-1$ when integer x is divided by n and denote by $x \equiv y \bmod n$ to mean $x \bmod n = y \bmod n$.

- Step 1. Alice picks an integer r at random and sends X to Bob where $0 < r < n$ and $X = r^e h(M) \bmod n$.
- Step 2. Upon receiving X, Bob computes $Y = X^d \bmod n$ and returns Y to Alice.
- Step 3. Finally, Alice obtains the signature $S = h(M)^d \bmod n$ by computing $r^{-1}Y \bmod n$ where $rr^{-1} = 1 \bmod n$.

Assume Bob runs a pay magazine system on-line. It contains many documents encrypted with Bob's public key. However, the abstracts of these documents are not encrypted and are free to all potential customers. Suppose Alice decides to purchase a document C after reading its abstract. Let n be the public modulus, e be the public exponent, and d be the secret exponent of Bob's RSA cryptosystem. Then Alice can pay Bob and use the following blind decoding scheme to ask him to decode (decrypt) it. Note that C is encrypted with Bob's public key. That is, $C = M^e \bmod n$ where M is the plaintext of the document.

- Step 1. Alice randomly chooses an integer r and sends X to Bob where $0 < r < n$ and $X = r^eC \bmod n$.
- Step 2. Bob computes $Y = X^d \bmod n$ and sends Y to Alice.
- Step 3. Alice recovers M by computing $M = r^{-1}Y \bmod n$.

The above scheme achieves perfect untraceability. That is, Bob has no way to know which document Alice has acquired. However, Bob is vulnerable to the oracle attack. If Bob uses the same set of parameters for creating digital signatures then Alice can trick him to sign any message chosen by her since decrypting a message is equivalent to signing it in RSA system. Therefore, the request from Alice must be restricted. The notion of transformable digital signature was introduced to serve this purpose [14,11]. Transformability is a property of a digital signature such that one valid signature can be transformed into another valid signature of the same scheme. The basic idea is to ensure the correctness of Alice's request based on the transformable signature derived from an original signature issued by Bob. Hence a malicious request from Alice will be discovered by Bob before the blind decryption step.

A scheme utilizes two pairs of RSA public-secret exponents for the same modulus, one for encrypting and the other for signing messages is proposed in [11]. The scheme goes as follows.

First Bob prepares two pairs of RSA exponents (e_1, d_1) and (e_2, d_2) for the same modulus n such that $e_1d_1 \equiv 1\ mod\ \phi(n),\ e_2d_2 \equiv 1\ mod\ \phi(n)$, and $\gcd(e_1, e_2) = 1$. The public exponent e_1 is for encryption and e_2 is for signature verification.

Then Bob encrypts document m_i using public exponent e_1 and secret exponent d_2 to produce a pair of integers (c_i, s_i), where $c_i = m_i^{e_1} \bmod n$ and $s_i = c_i^{d_2} \bmod n$. That is, c_i is the ciphertext of m_i and s_i is the signature of c_i. The pairs of all (c_i, s_i) are made known to all potential customers.

To blindly decode a document c_i, Alice and Bob perform the following steps interactively.

- Step 1. Alice randomly selects an integer B as the blind factor. Then she computes the transformation (u, v) from (c_i, s_i) such that $u = c_iB^{e_1e_2} \bmod n$ and $v = s_iB^{e_1} \bmod n$.
  The pair (u, v) is sent to Bob.
- Step 2. Upon receiving (u, v), Bob first checks the signature for u to see if $u = v^{e_2} \bmod n$. If the check is not valid, Bob stops. Otherwise, Bob computes $w = u^{d_1} \bmod n$ and returns w to Alice.
- Step 3. Alice recovers the document m_i from w and B since $m_i = wB^{-e_2} \bmod n$.

The above scheme achieves perfect untraceability. That is, Bob has no way to know which document Alice has acquired as long as the blind factor B is kept confidential since every m_j is possible if $B = (s/s_j)^{d_1} \bmod n$. However, it is shown that Alice can generate a valid signature S for any message M corresponding to the public key (e_1, n). The attack goes as follows.

Please cite this article in press as: Y.-C. Chen et al., Privacy protection in on-line shopping for electronic documents, Inform. Sci. (2014), http://dx.doi.org/10.1016/j.ins.2014.02.065