Xuejia Lai Moti Yung Dongdai Lin (Eds.)

Information Security and Cryptology

Sixth International Conference, Inscrypt 2010
Shanghai, China, October 2010
Short Paper Proceedings

State Key Laboratory of Information Security
Chinese Association for Cryptologic Research
Shanghai Jiaotong University

2010
Inscrypt

SCIENCE PRESS
Beijing

Xuejia Lai   Moti Yung   Dongdai Lin (Eds.)

Information Security and Cryptology

Sixth International Conference, Inscrypt 2010
Shanghai, China, October 2010
Short Paper Proceedings

SCIENCE PRESS
Beijing

内 容 简 介

本书是 2010 年 10 月在上海召开的第六届中国密码学与信息安全国际会议(The 6th China International Conference on Information Security and Cryptology-Inscrypt 2010)的短论文文集。Inscrypt 系列国际会议是由信息安全国家重点实验室发起，与中国密码学会联合举办的高水平国际会议，每年在中国举办一次，该会议论文集由 Springer 出版社出版。本书收录了这次会议的短文 13 篇。主要内容包括公钥和椭圆曲线密码学、密码系统构造、系统安全等。

本书可供从事密码学、信息安全、通信与信息系统、计算机应用技术等专业的科技人员和高等院校师生参考。

图书在版编目(CIP)数据

信息安全与密码学 = Information Security and Cryptology: 英文/来学嘉等主编. —北京: 科学出版社, 2011

ISBN 978-7-03-031194-8

Ⅰ.①信… Ⅱ.①来… Ⅲ.①信息安全-国际学术会议-文集-英文 ②密码-理论-国际学术会议-文集-英文 Ⅳ.①TP309-53 ②TN918.1-53

中国版本图书馆 CIP 数据核字 (2011) 第 098611 号

责任编辑: 鞠丽娜 / 责任校对: 耿 耘
责任印制: 吕春珉 / 封面设计: 三函设计

科学出版社出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

双青印刷厂印刷

科学出版社发行 各地新华书店经销

*

2011 年 6 月第 一 版 开本: 小16 (720 × 1000)
2011 年 6 月第一次印刷 印张: 13
印数: 1—1 500 字数: 260 000

定价: 50.00 元

(如有印装质量问题，我社负责调换〈双青〉)

Preface

The 6th China International Conference on Information Security and Cryptology (Inscrypt 2010) was held in Shanghai, China during 20–23, October 2010. The conference is a leading annual international event in the area of cryptography and information security taking place in China. Inscrypt continues to get the support of the entire international community, reflecting the fact that the research areas covered by the conference are important to modern computing, where increased security, trust, safety and reliability are required.

Inscrypt 2010 was co-organized by the State Key Laboratory of Information Security and by the Chinese Association for Cryptologic Research, in cooperation with Shanghai Jiaotong Univeristy and the International Association for Cryptologic Research (IACR). The conference was further sponsored by the Institute of Software, the Graduate University of the Chinese Academy of Science and the National Natural Science Foundations of China.

The scientific program of the conference covered all areas of current research in cryptography and security, with sessions on central subjects of cryptographic research and on some important subjects of information security. The international Program Committee of Inscrypt 2010 received a total of 125 submissions from more than 29 countries and regions, from which only 35 submissions were selected for presentation in the regular papers track and 13 submissions in the short papers track. Short track papers appear in these proceedings. All anonymous submissions were reviewed by experts in the relevant areas and based on their ranking, technical remarks and strict selection criteria the papers were chosen to the various tracks. The selection to both tracks was a highly competitive process. We further note that due to the conference format, many good papers were regrettably not accepted. Besides the contributed papers, the program also included two invited presentations by Bart Preneel and Moti Yung.

Inscrypt 2010 was made possible by a joint effort of numerous people and organizations worldwide. We take this opportunity to thank the Program Committee members and the external experts they employed for their invaluable help in producing the conference program. We further thank the Conference Organizing Committee, the various sponsors, and the conference attendees. Last but not least, we express our great gratitude to all the authors who submitted papers to the conference, the invited speakers, and the session Chairs.

Xuejia Lai
Moti Yung
December 2010

Inscrypt 2010

6th China International Conference on Information Security and Cryptology

Shanghai, China
October 20–23, 2010

Sponsored and organized by

State Key Laboratory of Information Security
(Chinese Academy of Sciences)
Chinese Association for Cryptologic Research

in cooperation with

Shanghai Jiaotong University
International Association for Cryptologic Research

Steering Committee

Dengguo Feng	SKLOIS, Chinese Academy of Sciences, China
Dongdai Lin	SKLOIS, Chinese Academy of Sciences, China
Moti Yung	Google Inc and Columbia University, USA
Chuankun Wu	SKLOIS, Chinese Academy of Sciences, China

General Chairs

Dengguo Feng	SKLOIS, Chinese Academy of Sciences, China

Program Committee

Co-chairs

Xuejia Lai	Shanghai Jiaotong University, China
Moti Yung	Google Inc and Columbia University, USA

Members

Vladimir Anashin	Moscow State University
Frederik Armknecht	Institute for Computer Science at the University of Mannheim, Germany
Rana Barua	Indian Statistical Institute, Kolkata, INDIA
Zhenfu Cao	Shanghai Jiao Tong University
Claude Carlet	Université Paris 8, France
Luigi Catuogno	Università degli Studi di Salerno - ITALY
Lily Chen	National Institute of Standards and Technology (NIST), USA
Liqun Chen	HP Laboratories, UK

Organization

Ed Dawson	QUT, Australia
Robert Deng	Singapore Management University
Jintai Ding	Cincinnati University, USA
Jean-Charles Faugere	INRIA, France
Keith Frikken	Miami University
Alejandro Hevia	University of Chile, Chile
James Hughes	Huawei N. America, USA
Jiwu Jing	GUCAS, Chinese Academy of Sciences, China
Brian King	Indiana University Purdue University Indianapolis
Albert Levi	Sabanci University, Turkey
Chao Li	National University of Defence Tech., China
Peng Liu	Penn State University, USA
Javier Lopez	University of Malaga, Spain
Masahiro Mambo	University of Tsukuba, Japan
Atsuko Miyaji	Japan Advanced Institute of Science and Technology, Japan
Yi Mu	University of Wollongong, Australia
Peng Ning	North Carolina State University, USA
Olivier Pereira	Universit́e catholique de Louvain
Ludovic Perret	LIP6, France
Svetla Petkova-Nikova	K.U. Leuven, Belgium and UTwente, The Netherlands
Raphael Phan	Loughborough University, UK
Josef Pieprzyk	Macquarie University, Australia
Kui Ren	Illinois Institute of Technology, USA
Yannis Stamatiou	University of Ioannina, Greece
Tsuyoshi Takagi	Kyushu University, Japan
Jacques Traore	Orange Labs, France
Wen-Guey Tzeng	National Chiao Tung University, Taiwan
Daoshun Wang	Tsinghua University, China
Huaxiong Wang	Nanyang Technological University, Singapore
Wenling Wu	Institute of Software, CAS, China
Shouhai Xu	University of Texas at San Antonio, USA
Nong Ye	Arizona State University, USA
HeungYoul Youm	Soonchunhyang University, Korea
Meng Yu	Virginia Commonwealth University
Erik Zenner	Technical University of Denmark
Rui Zhang	AIST, Japan
Yuliang Zheng	University of North Carolina at Charlotte, USA
Jianying Zhou	Institute for Infocomm Research, Singapore

Proceedings Co-editors

Xuejia Lai	Shanghai Jiaotong University, China
Moti Yung	Google Inc and Columbia University, USA
Dongdai Lin	SKLOIS, Institute of Software, Chinese Academy of Sciences, China

Organizing Committee

Co-chairs

Kefei Chen	Shanghai Jiaotong University, China
Chuankun Wu	SKLOIS, Institute of Software, Chinese Academy of Sciences, China

Members

Feng Liu	Institute of Software, CAS, China
Yanfei Zheng	Shanghai Jiaotong University, China
Xianping Mao	Shanghai Jiaotong University, China

Publication Chair

Dongdai Lin	SKLOIS, Institute of Software, Chinese Academy of Sciences, China

WEB Master

Jinyuan Tang	Institute of Software, CAS, China

Conference Secretary

Yi Qin	Institute of Software, CAS, China

Author Index

Table of Contents

I Public Key and Elliptic Curve Cryptography

II Cryptographic Construction

III System Security

Part I

Public Key and Elliptic Curve Cryptography

Attacking Code/Lattice-based Cryptosystems Using Partial Knowledge

Robert Niebuhr[1], Pierre-Louis Cayrel[2], Stanislav Bulygin[2], Johannes Buchmann[1,2]

1 Technische Universität Darmstadt Fachbereich, Informatik Kryptographie und Computeralgebra, Hochschulstraße 10, 64289 Darmstadt Germany
rniebuhr@cdc.informatik.tu-darmstadt.de
2 CASED – Center for Advanced Security Research Darmstadt,
Mornewegstrasse 32, 64293 Darmstadt Germany
{pierre-louis.cayrel,Stanislav.Bulygin}@cased.de

Abstract. Code-based cryptographic schemes are promising candidates for post-quantum cryptography since they are fast, require only basic arithmetic, and because their security is well understood. While most analyses of security assume that an attacker does not have any information about the secret key, we show that in certain scenarios an attacker can gain partial knowledge of the secret key. We present how this knowledge can be used to improve the efficiency of an attack, and give new bounds for the complexity of such an attack. In this paper, we analyze two types of partial knowledge including concrete scenarios, and give an idea how to prevent the leak of such knowledge to an attacker.

Keywords: Information set decoding, Partial knowledge, Codes, Post quantum, Cryptography

1 Introduction

In 1994, P. Shor [18] has shown that quantum computers can break most or all "classical" cryptosystems, e.g. those based on RSA or elliptic curves. Therefore, it is crucial to develop cryptosystems that are resistant to quantum computer attacks. Code-based cryptography is a very promising candidate for post-quantum cryptography since the cryptographic schemes are usually fast and do not require special hardware, specifically no cryptographic co-processor. The first application was the McEliece encryption scheme [13] which was published in 1978. It is as old as RSA and has resisted cryptanalysis to date (except for a parameter adjustment).

To analyze the security of code-based schemes, cryptanalysts develop and improve (generic as well as specific) attacks, or propose lower bounds for such attacks. All of these developments, however, assume that an attacker does not have any knowledge about the private key.

Partial Knowledge

In some scenarios though an attacker can obtain partial knowledge of the private information and exploit this to improve the efficiency of an attack. Examples are:

1) **Scheme that uses a restricted error vector domain**: Some cryptographic schemes, e.g. NTRU [11], restrict the domain of the error vector. In this example, while the scheme itself is defined over $\mathbb{F}_q$ for $q = 128$ or $q = 256$, the error vector e is ternary, i.e. $e \in \{0, 1, -1\}^n$. Information-set decoding (ISD) algorithms can be used to attack such systems, and we will show how to exploit this knowledge. We will analyze this example in more detail in section 5.

2) **Schemes that leak information about the error vector entries**: In Stern's identification (ID) scheme [20], the prover sends a random permutation of the private vector to the verifier. This reveals the non-zero values of the vector, while their positions remains secret. While this information is useless when binary codes are used (as for the original scheme), it *does* give the attacker an advantage when codes over $\mathbb{F}_q$ are used. We will analyze this example in more detail in section 4.

Another type of partial knowledge would be the use of error vectors with a certain structure, e.g. regular words which are used for the FSB hash function [1]. However, in this specific context this knowledge does not necessarily help an attacker, since this restricts both the set of error vectors that need to be searched as well as the number of vectors that lead to a successful attack. Therefore, we will not consider this type of partial knowledge in this paper.

Our Contributions

In this paper we will analyze two types of partial knowledge an attacker can obtain in certain scenarios. We will show that they can be used to improve the efficiency of an attack by restricting the space that need to be searched and prove new lower bounds for these cases. As a first examples, we analyze a q-ary version of Stern's ID scheme to apply our results and discuss a method to prevent the information leak. A second example shows how to attack the NTRU scheme using our modified ISD algorithm.

Organization of The Paper

In section 2 we review some concepts and notation from code-based cryptography, and describe the Information Set Decoding algorithm on which our analysis is based. section 3 covers the analysis of the above types of partial knowledge. In the following section 4 we apply our results to using the example of Stern's ID scheme and to NTRU in section 5. We conclude in section 6.

2 Review

2.1 Coding Theory over $\mathbb{F}_q$

In this paper, we consider linear error-correcting codes over a finite field $\mathbb{F}_q$. A linear code C is a k-dimensional subspace of an n-dimensional vector space over $\mathbb{F}_q$ and is called an $[n, k]$ code. The elements of a code are called codewords. The (Hamming) weight of a vector is the number of its non-zero entries, and the (Hamming) distance of two vectors is the weight of their difference. The minimum distance d of a code is the minimum distance between any two distinct codewords; a code with these properties is denoted as an $[n, k, d]$ code. Codes that are able to correct up to t errors are denoted (n, k, t)-codes.

Another common notation is the *co-dimension* r of a code where $r = n - k$.

Definition 1 (Generator and Parity Check Matrix). *Let C be a linear code over $\mathbb{F}_q$. A generator matrix G of C is a matrix whose rows form a basis of C:*

$$C = \{xG : x \in \mathbb{F}_q^k\}$$

Two generator matrices generate equivalent codes *if one is obtained from the other by a linear transformation. Therefore, we can write any generator matrix G in* systematic form $G = [I_k|R]$ *with $R \in \mathbb{F}_q^{k \times r}$, which allows a more compact representation.*

A parity check matrix H of C is defined by

$$C = \{x \in \mathbb{F}_q^n : Hx^{\mathrm{T}} = 0\}$$

and generates the dual space of C. If C is generated by $G = [I_k|R]$, then a parity check matrix for C is $H = [-R^{\mathrm{T}}|I_r]$ (sometimes H is transformed so that the identity submatrix is on the left hand side).

The problems which cryptographic applications rely upon can have different numbers of solutions. For example, public key encryption schemes usually have exactly one solution, while digital signatures often have more than one possible solution. The uniqueness of solutions can be expressed by the Gilbert-Varshamov (GV) bound:

Definition 2 (q-ary Gilbert-Varshamov bound). *If*

$$\sum_{i=0}^{d-2}(q-1)^i\binom{n}{i} \leq q^{n-k}$$

there exist a $[n, k, d]$ code over $\mathbb{F}_q$.

For NTRU and Stern's scheme we have $k = \frac{n}{2}$, and

$$\sum_{i=0}^{d-2}(q-1)^i\binom{n}{i} \leq q^{n/2}$$

Random codes, which are used in Stern's scheme, on average satisfy this bound ($\sum_{i=0}^{d-2}(q-1)^i\binom{n}{i} = q^{n-k}$).

2.2 Information Set Decoding (ISD)

Information Set Decoding algorithms are often the most efficient generic attack against code-based cryptosystems like McEliece, the CFS signature scheme [9], the FSB hash function [1], and others. Over the years, there have been many improvements and generalizations of this attack, e.g. Lee-Brickell [12], Stern [19], Canteaut-Chabaud [7], Bernstein et al. [6], Finiasz-Sendrier [10], Peters [16], Niebuhr et al. [15] and Peters [17].

ISD algorithms solve the problem of decoding codewords with errors. More specifically, if m is a cleartext and $c = mG + e$ a ciphertext, where e is a random vector of weight t, then ISD algorithms take c as input and recover m (or, equivalently, e). Since $Hc^{\mathrm{T}} = H(mG + e)^{\mathrm{T}} = He^{\mathrm{T}}$, the problem is often formulated using a parity check matrix:

Problem 1 (The q-ary syndrome decoding problem). Given a matrix H and a vector s, both over $\mathbb{F}_q$, and a non-negative integer t; find a vector $x \in \mathbb{F}_q$ of weight t such that $Hx^{\mathrm{T}} = s$.

This problem was proved to be NP-complete, in 1978 for binary codes [5] and in 1994 for codes over all finite fields ([3, in russian] and [2]).

If the number of errors that have to be corrected is smaller than the GV bound, then there is on average only one solution. Otherwise, there can be several solutions.

A basic version of an ISD algorithm works as follows: A random permutation P is applied to H in the hope that all columns corresponding to error positions in e are moved to the left hand side of the matrix (the first $n-k$ columns). Then Gaussian elimination is used to transform H into the form $H' = [I_{n-k}|R]$, where I_{n-k} is the identity submatrix, and the same steps are performed on s to get s'. If s' has a weight not exceeding t, the algorithm has succeeded; we can read of the error positions from s' and get $e = P^{-1}[s'|0]$. Otherwise, the algorithms restarts.

Most advanced ISD versions make use of the birthday paradox: They allow a certain (usually small) number p of errors in the last k columns of H. Then lists of column sums of H are used to find these error positions. If we split the right hand part of H into $[H_1|H_2]^{\mathrm{T}}$, and write $e = [e_1|e_2]$, then we search for a vector e_2 of weight p such that $s - H_2 e_2^{\mathrm{T}}$, has weight $t - p$, and the non-zero positions of $s - H_2 e_2^{\mathrm{T}}$ show the remaining $t - p$ error positions.

Since our paper modifies the ISD algorithm described in [15], we will review some of the concepts and notations.

In each step, we randomly re-arrange the columns of the parity check matrix H and transform it into the form

$$H = \left(\begin{array}{c|c} I_{n-k-l} & H_1 \\ \hline 0 & H_2 \end{array}\right) \tag{1}$$

where I_{n-k-l} is the identity matrix of size $(n-k-l)$. Usually, the columns are chosen adaptively to guarantee the success of this step. The variables l and p (see next step) are algorithm parameters optimized for each attack.

The error vector we are looking for has p errors in the column set corresponding to H_1 and H_2, and the remaining $(t-p)$ errors in the first $(n-k-l)$ columns. We first check all possible error patterns of p errors in the last $k+l$ columns such that the sum S of those p columns equals the syndrome s in the last l rows. We do this by searching for collisions between the two sets L_1 and L_2, where

$$L_1 = \{H_2 e^{\mathrm{T}} : e \in W_1\} \tag{2}$$

$$L_2 = \{s - H_2 e^{\mathrm{T}} : e \in W_2\} \tag{3}$$

where $W_1 \subseteq \mathcal{W}_{k+l;\lfloor p/2 \rfloor;q}$ and $W_2 \subseteq \mathcal{W}_{k+l;\lceil p/2 \rceil;q}$ are given to the algorithm, and $\mathcal{W}_{k+l;p;q}$ is the set of all q-ary words of length $k+l$ and weight p. Writing $e = [e'|e_1+e_2]$ and $s = [s_1|s_2]$ with s_2 of length l, this means we search for vectors e_1 and e_2 of weight $\lfloor p/2 \rfloor$ and $\lceil p/2 \rceil$, respectively, such that

$$H_2 \cdot [e_1 + e_2]^{\mathrm{T}} = s_2^{\mathrm{T}}$$

If this succeeds, we compute the difference $S - s$; if this does not have weight $t-p$, the algorithm restarts. Otherwise, the non-zero entries correspond to the remaining $t-p$ errors:

$$\begin{aligned} He^{\mathrm{T}} &= \left(\begin{array}{c|c} I_{n-k-l} & H_1 \\ \hline 0 & H_2 \end{array}\right) \begin{pmatrix} e' \\ e_1 + e_2 \end{pmatrix} \\ &= \begin{pmatrix} I_{n-k-l} \cdot e'^{\mathrm{T}} + H_1 \cdot (e_1+e_2)^{\mathrm{T}} \\ H_2 \cdot (e_1+e_2)^{\mathrm{T}} \end{pmatrix} \\ &= \begin{pmatrix} I_{n-k-l} \cdot e'^{\mathrm{T}} \\ 0 \end{pmatrix} + S \\ &\overset{!}{=} \begin{pmatrix} s_1^{\mathrm{T}} \\ s_2^{\mathrm{T}} \end{pmatrix} \end{aligned}$$

Therefore, we have

$$I_{n-k-l} \cdot e'^{\mathrm{T}} = s_1^{\mathrm{T}} - H_1 \cdot (e_1 + e_2)^{\mathrm{T}}$$

revealing the remaining columns of e.

3 Impact of Partial Knowledge on ISD

We will analyze two types of partial knowledge:

1) Error values come from a set $E \subsetneq \mathbb{F}_q$.
2) The entries of e are known but their positions are not.

There are various situations where an attacker has some partial knowledge of these types. For example, the NTRU cryptosystem can be attacked by an ISD-like algorithm; although the cryptosystem is defined over $\mathbb{F}_q$ (common values for q are 128 or 256), the random vector required to break the scheme is only ternary.