A Novel Encryption Method with Its Application in the Copyright Protection of Digital Data*

MA Jian-feng[1,2], Chiam Teechye[2], Kot Chichung Alex[2]

[1](*School of Computer, Xidian University, Xi'an* 710071, *China*);

[2](*School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore*)

E-mail: jianfma@pub.xoaonline.com; ejfma@hotmail.com; {etcchiam,eackot}@ntu.edu.sg

http://www.xidan.edu.cn

Received September 15, 2000; accepted October 17, 2001

Abstract: A novel public key encryption scheme with multiple private keys is presented. The security of this encryption method depends on the difficulty of the decoding problem for block codes over finite fields. Based on the encryption method, a pirate tracing scheme is proposed for the copyright protection of digital data. In this tracing scheme, for each legal copy of digital data there is a codeword buried in the corresponding decryption software or box. The scheme can extract the codeword efficiently, and uncover all authorized users involved in making the illegal copy by using a proper decoding algorithm if the number of such users is not more than the error-correcting capacity of a given code. Compared with known tracing schemes, this scheme is efficient in performance, and easy in implementation. Potential applicable areas of the encryption method include the protection of copyrights of various forms of digital data such as computer software and audio/video products.

Key words: linear code; copyright protection; pirate tracing scheme; public key encryption; decoding

To ensure the security and privacy of sensitive messages over non-secure channels, cryptographic techniques have been widely used. In fact, cryptographic techniques can also be used to protect the copyrights of digital data[1,2]. Public key cryptography is a well-established method for the private transmission of sensitive data, which uses a public key for message encryption and a private key for message decryption. In this case privacy means that if two parties use public key cryptography to exchange a message, an intermediary, although capable of intercepting the message, will not be able to read, or modify the message or fabricate a realistic-looking substitute message. Compared with non-public-key cryptographic techniques the encryption method has several advantages; in particular, it eliminates the problem of key distributions over the network, and provides the digital signature to verify the authorized sender and check whether any message has been altered. It is known that the decoding problem of linear block codes over finite fields is NP-Complete. This property can be used to build cryptosystems or identification systems[3~5]. In this paper, we propose a new public key cryptosystem based on linear codes[5,6]. The cryptosystem has multiple public keys and multiple private keys. The security of this encryption method depends on the difficulty of the decoding problem for block codes over finite fields. Because of its salient properties, this new cryptosystem can be used in the copyright protection of digital data.

This paper is organized as follows. In section 1 we present the public key cryptosystem. In section 2 we discuss applications of the novel encryption system in the copyright protection of digital data. Conclusions are given in

∗ **MA Jian-feng** was born in 1963. He is a professor and doctoral supervisor of Xidian University. He received his Ph.D. degree in communications engineering from Xidian University in 1995. His research areas include information security, channel coding, fault tolerant computing, and network management.

section 3.

1 Encryption Scheme Based on Linear Code

Let V be an $[n,k]$ linear block code over $GF(q)$, and for the dual of the code V there exists fast decoding algorithm, where q is a large prime number.

Encryption system setup

Perform the following steps.

Let H be a parity check matrix of the linear block code V used, choose a random matrix $R=(r_{ij})_{(n-k)\times n}$ over $GF(q)$. For the matrix $R_{(n-k)\times n}$, there must be some codewords of the code V satisfying that every one of these codewords is not orthogonal to any row of the matrix R. These codewords will be used to generate private keys of the encryption scheme.

For each codeword $v=(v_1,...,v_n)\in V$ suitable for generating private key, compute $\theta_i = \alpha_i \Big/ \sum_{j=1}^{n} r_{ij}v_j$, $1\le i\le n-k$, where α_i is at random chosen from $GF(q), 1\le i\le n-k$.

Select an $(n-k)\times(n-k)$ invertible matrix S over $GF(q)$ and an $n\times n$ permutation matrix P. Let $R'=diag(\alpha_1^{-1},...,\alpha_{n-k}^{-1})RP$, and P' be another $n\times n$ permutation matrix over $GF(q)$. Suppose the public key is (H',R',P',t), where $H'=SHP$, and t is the error-correcting capability of the code with the generator matrix H, that is, which is the dual of the code V with the matrix H as its parity check matrix.

Let the private key be $[S,H,P,(\theta_1,...,\theta_{n-k}),v]$.

Encryption operations: To encrypt a message $M=(M_1,...,M_{n-k})$ in $[GF(q)\backslash\{0\}]^{n-k}$, select randomly $n-k$ elements $\beta_1,...,\beta_{n-k}$ from $GF(q)\backslash\{0\}$ and a random matrix $T_{(n-k)\times(n-k)}$ over $GF(q)$. Then set the ciphertext to be

$$TH'+E \quad \text{and} \quad \begin{pmatrix} M_1\beta_1 & \\ ... & diag(\beta_1,...,\beta_{n-k})(H'+R')+EP' \\ M_{n-k}\beta_{n-k} & \end{pmatrix},$$

where E is a random matrix over $GF(q)$, of which the weight of each row is not greater than t. Therefore the ciphertext comprises two parts, which can be denoted by A and B.

Decryption operations: To decrypt the ciphertext $\{A,B\}$ such as the above using a user's secret key, compute $AP^{-1}=(TH'+E)P^{-1}=TSH+EP^{-1}$, where $A=TH'+E$ and

$$B=\begin{pmatrix} C_1 & \\ ... & W_{(n-k)\times n} \\ C_{n-k} & \end{pmatrix}=\begin{pmatrix} M_1\beta_1 & \\ ... & diag(\beta_1,...,\beta_{n-k})(H'+R')+EP' \\ M_{n-k}\beta_{n-k} & \end{pmatrix}.$$ Decode every row of the matrix AP^{-1} using the proper decoding algorithm of the code $V^{\perp}$. The code $V^{\perp}$ is the dual of the code V. Then E is obtained. We calculate

$$B(v_1,...,v_n)^t \quad =\{[diag(\beta_1,...\beta_{n-k})(H'+R')+EP']-EP'\}(v_1,...,v_n)^t$$

$$=diag(\theta_1^{-1}\beta_1,...,\theta_{n-k}^{-1}\beta_{n-k})=diag(u_1,...,u_{n-k}).$$

Thus $$M_i=(C_i \big/ \theta_i\beta_i),\quad 1\le i\le n-k.$$

The security of the above encryption scheme: To decrypt a ciphertext, one must find out E from the equation

$A=TH'+E$ first, then compute the plaintext from B. However, solving E from that equation is as difficult as decoding a linear code. It is well known that the decoding problem of linear codes is NP-Complete[3~5]. Therefore, it is impossible to recover the plaintext M without a valid private key, that is, our encryption scheme is secure.

Because for a given linear code there is one more parity check matrix, and for a given chosen matrix R and a given parity check matrix there are some codewords that can be used in generating private keys, the encryption scheme has multiple public keys and multiple private keys.

2 A Collusion Secure Pirate Tracing Scheme for Digital Data Based on the New Encryption Method

Consider the distribution of digital content to users over a broadcast channel. Typically the distributor gives each authorized user a hardware or software decoder (“box”) containing a secret decryption key. The distributor then broadcasts an encrypted version of the digital content. Authorized users are able to decrypt and use the content. This scenario comes up in the content of pay-per-view television, and more commonly in web based electronic commerce (e.g. broadcast of online stock quotes or broadcast of proprietary market analysis). However, nothing prevents a legitimate user from giving copies of their decryption software to someone else. Worse is that authorized users might try to expose the secret key buried in their decryption boxes and make copies of the keys freely available. The pirates would thus make all of the distributor's broadcasts freely available to non-authorized users. Pirate tracing schemes of digital data discourage authorized users from giving away their keys. But a coalition of pirates might try to mix keys from many boxes to create a new piratic box that can still decrypt but cannot be traced back to them. Therefore pirate tracing schemes should be collusion-secure. In the following, we present an efficient tracing scheme based on the second encryption method mentioned above.

2.1 Proposed traitor tracing scheme

In the scheme, two linear codes are needed. Let the first code V' be an $[n,k,d]$ code with the generator matrix G' and the parity check matrix H', and the second V an $[2(n-k),n-k]$ code with the generator matrix G and the parity check matrix H, both are over $GF(q)$, where q is a large prime number. For the first code V', let the parity check matrix $H'=(h_1',...,h_n')$. Corresponding to n columns of the matrix H', there are n codewords in the code V. Among these codewords there are n-k linear independent codewords which form the rows of the generator matrix G of the code V. In accordance with the code V, an encryption system is established, but the codewords used to generate private keys are chosen only from those n codewords of the code V corresponding to the columns of the parity check matrix H' of the code V'. In the encryption scheme every row of the chosen matrix R is required not to be orthogonal to each of these n codewords of the second code V. Thus, the related encryption scheme has n private keys, where n is the codelength of the first code V', so is also the number of authorized users of some given digital data. Let M be the message to be distributed, which is denoted by $M=(M_1,...,M_{n-k})$, where $M_i\in Z_q^*, 1\le i\le n-k$. The message M will be encrypted using the second encryption scheme with multiple public keys and multiple private keys in Section 1. For the sake of simplicity, we assume that the second code is systematic, the first $n-k$ components of every codeword constitute its information part. In order to recover the coalition of users involved in making an illegal decryption box, it is desirable to enable the tracer to extract the codeword used by the decryption box by observing its behavior on a chosen ciphertext. That is, the piratic decryption box or software is used as an oracle. Given this oracle the pirate tracing scheme to be designed must output the keys of the pirates. In the following, we give a detailed description of the related pirate tracing scheme to illegal decryption boxes of some given digital data.

Step 1. Using given piratic decryption box, we observe its behavior for a specified invalid ciphertext, where a chosen invalid ciphertext C is obtained by encrypting a plaintext message with a specified encryption scheme which is similar to the authorized one used to encrypt given digital data. In this encryption scheme, the matrix related R is chosen as $(A_{(n-k)\times(n-k)}, O_{(n-k)\times(n-k)})$, where $\det(A)\neq 0$, and $O_{(n-k)\times(n-k)}$ is a zero matrix. From the output of the decryption box, the following set of linear equations can be derived,

$$A(v_1,\dots,v_{n-k})^t=(\sigma_1,\dots,\sigma_{n-k})^t\ ,$$

where σ_i, $1\le i\le n-k$ can be obtained from some known quantities.

Therefore,

$$(v_1,\dots,v_{n-k})^t=A^{-1}(\sigma_1,\dots,\sigma_{n-k})^t\ .$$

Step 2. As the information part of the codeword $(v_1,\dots,v_{2(n-k)})$, the vector $(v_1,\dots,v_{n-k})$ must be the linear combination of some columns of the parity check matrix H' of the first linear code V'. In order to determine these columns of the matrix H', we must find a vector e of Hamming weight at most $t=\left\lfloor {(d-1)}/{2} \right\rfloor$ satisfying $eH'=(v_1,\dots,v_m)$, where t is the error-capacity of the code V', is also the largest size of the coalition of users making the illegal decryption box, and d is the minimum Hamming distance of the first code V'. By the vector e and the relationship between the columns of the matrix H' and the private keys of the used encryption scheme for the digital content, the coalition of users can be determined.

Find a vector u over Z_q such that $uH'=(v_1,\dots,v_{n-k})$. Many such vectors exist. Choose one arbitrarily. Since $(u-e)H'=0$, $u-e$ is a codeword of the first code V'.

By employing the proper decoding algorithm of the first code V', we can find the codeword $u-e$ from u, which results in the required vector e.

In order to recover the vector e, the first code V' should have a fast and efficient decoding algorithm. The time complexity of the above tracing algorithm depends on that of the decoding algorithm of the first code V'.

The security of the above scheme depends on that of the encryption scheme used, its efficiency is determined by the encryption scheme for the copies of digital data, the decoding algorithms of the related linear codes in the tracing scheme and the extraction process of the codeword buried in the piratic decryption box or software. Obviously, the above pirate tracing scheme is deterministic, and is, in principle, able to deal with any case occurring in practice if the size of the coalition of users making an illegal decryption box is not more than some specified limit. Therefore, it is possible to apply the tracing scheme in practical copyright of digital information such as pay-per-vision television.

3 Conclusion

In this paper, we present a novel public key cryptosystem based on linear codes, and have demonstrated that the security of the encryption scheme is determined by the intractability of the decoding problem of block codes over finite fields. In particular the proposed encryption scheme has multiple private keys, and can be implemented efficiently. Finally, we discuss the application of the encryption scheme in the copyright protection of digital data, and present a pirate tracing scheme for illegal copy of digital data. If the number of users creating some illegal copy is not more than the error-capacity of some given linear code, the above scheme can effectively uncover these users. By checking the output of the related piratic decryption box or software for a specified ciphertext, the proposed pirate tracing scheme can extract the required codewords buried in the piratic decryption boxes or software efficiently.

References:

[1] Boneh, D., Franklin, M. An efficient public key traitor tracing scheme. In: Wiener, M., ed. Advances in Cryptology: Proceedings of the Crypto'99. Lecture Notes in Computer Science, Vol.1666. Berlin: Springer-Verlag, 1999. 338~353.

[2] Pfitzmann, B., Sadeghi, A.R. Coin-Based anonymous fingerprinting. In: Stern, J., ed. Advances in Cryptology: Proceedings of the EUROCRYPT'99. Lecture Notes in Computer Science, Vol.434. Berlin: Springer-Verlag, 1999. 150~164.

[3] Chaband, F. On the security of some cryptosystems based on error-correcting codes. In: Santis, A.D., ed. Advances in Cryptology: Proceedings of the EUROCRYPT'94. Lecture Notes in Computer Science, Vol.950. Berlin: Springer-Verlag, 1995. 131~139.

[4] Tilborg, H.C.A. Coding theory at work in cryptology and vice versa. In: Pless, V.S., Huffman, W.C., eds. Handbook of Coding Theory (Part 2, Connections). Amsterdam: Elsevier Science B.V., 1998. 1195~1227.

[5] Pless, V. Introduction to the Theory of Error-Correcting Codes. 3rd ed, New York: John Wiley & Sons, Inc., 1998. 17~36.

[6] Peterson, W.W., Weldon, Jr. E.J. Error-Correcting Codes. 2nd ed, Cambridge: The MIT Press, 1988. 40~115.

一种新的加密方法及在数字数据版权保护中应用

马建峰[1,2], Chiam Teechye[2], Kot Chichung Alex[2]

[1](西安电子科技大学 计算机学院,陕西 西安 710071);

[2](南洋理工大学 电气电子工程学院,新加坡 639798,新加坡)

摘要: 提出了一种新的具有多个私钥的加密方法.这种加密方法的安全性有赖于有限域上分组码译码问题的困难性.根据所提的加密算法,给出了一种适用于数字数据版权盗版跟踪方案.对于每一个数字数据的合法拷贝,给出的跟踪方案都分配一个分组码的码字与该拷贝的解密软件或者相关装置相对应.所提的方案能够有效地提取与解密软件或者装置相对应的码字,并进而利用所用分组码的译码算法找到所有参与非法拷贝制作的合法授权用户,其中,假设参与数字数据非法拷贝的合法用户数不超过所用码的纠错能力.与其他已有方案相比,所提方案具有更好的性能,并且容易实现.所提加密方案的潜在应用领域包括计算机软件、视频/音频等产品的数字数据版权保护方面.

关键词: 线性码;版权保护;盗版跟踪方案;公钥加密;译码

中图法分类号: TP309 文献标识码: A