Two Chosen-Plaintext Attacks on the Li-Wang Joint Authentication and Encryption Scheme

Johan van Tilburg*

Abstract. In [LW91], Li and Wang proposed a joint authentication and encryption scheme based on algebraic coding theory. They claimed that their scheme is as secure as the Rao-Nam scheme [RN89]. However, in contrast with their claim, it will be shown that this joint authentication and encryption scheme appears to be less secure. In this paper two inherently different chosen-plaintext attacks are presented.

The first attack is based on the linearity of the bit selection function, and obtains a $k \times n$ matrix equivalent to the encryption matrix in $O(k)$ encryptions. If the set of error vectors $\mathcal{Z}$ is randomly chosen, about $|\mathcal{Z}|$ encryptions are necessary to obtain a corresponding set of error vectors. With knowledge of only $r$ error vectors one can always encrypt and decrypt $r^2 k^{-n}$ messages.

The second attack makes use of the non-linearity of the error function, and always obtains $2k - n$ rows of the encryption matrix after $O(k^2)$ encryptions. Hereafter $|\mathcal{Z}|$ encryptions are required to create a cryptosystem equivalent to the Li-Wang scheme.

Some extensions of the scheme are discussed, and a general question raised by Brickell and Odlyzko [BO88] related to the Rao-Nam scheme is settled in a negative way.

1 Introduction

In [LW91], Li and Wang introduced a joint authentication and encryption scheme based on algebraic coding theory. They claim that their secret-key cryptosystem, which we call LW-scheme, can be used to authenticate and encrypt messages simultaneously. Under the assumption that the LW-scheme is secure, this means that two users sharing the same secret key are able to exchange secret messages over an insecure communication channel and are able to detect if the received ciphertext was tampered with. We refer to [Sim88] and [Mass88] for a detailed treatment of this subject.

In this paper, the LW-scheme is described for code rates $R \geq 1/2$. However, the results obtained are also valid for the less interesting case $R < 1/2$. For the original description we refer to [LW91].

The LW-scheme is based on a secret-key cryptosystem, called RN-scheme, proposed by Rao and Nam [RN89]. The RN-scheme is a secret-key variant of the McEliece public-key cryptosystem [McE78]. All these schemes are based on

* PTT Research, Dept. of Applied Mathematics and Signal Processing, P.O. Box 421, 2260 AK Leidschendam, the Netherlands.
algebraic coding theory and allow high-speed implementations. Without details, the RN-scheme can be summarized as follows. A message $z$ is encrypted by randomly selecting an error vector $z$ from a predefined set of error vectors $Z$ and forming the ciphertext $y = zE + z$, where $E$ is the encryption matrix. As a consequence of the special way the set $Z$ was generated, unique decoding is possible. For more details, we refer the reader to [RN89].

In [Mass88], Massey remarks that most cryptosystems in use today are intended by their designers to be secure against at least a chosen-plaintext attack, even if it is hoped that the enemy cryptanalyst will never have the opportunity to mount more than a cipher-text-only attack. In [LW91], Li and Wang conclude that their secret-key cryptosystem is secure under a chosen-plaintext attack by a work factor of about $|Z|^k$, where $|Z|$ denotes the cardinality of the set of error vectors. It should be noted that the set $Z$ is part of the secret key. In the RN-scheme each error vector can be used to encrypt any given message, but in the LW-scheme each message corresponds to only one error vector. Besides, each error vector corresponds to several messages. It will be shown in Sections 4 and 5 that this is a serious weakness of the LW-scheme and enables us to mount two attacks with a work factor of about $|Z|$ chosen-plaintexts. Because the RN-scheme does not have this weakness, the proposed attacks do not apply to the RN-scheme.

In Section 6, some extensions of the LW-scheme will be discussed. Moreover, we will settle a general question raised by Brickell and Odlyzko in their survey article [BO88]. Their question was related to the security of the RN-scheme: Could the security of the RN-scheme be improved if the scheme is slightly modified by using a pseudo-random function $f$, and letting $z = f(z)$ so that there is only one encryption for each message $z$?

Li and Wang suggest to use the set $Z$, if not used for authentication, for purpose of encryption and improving the code rate of their scheme. And in their conclusion, they state that a similar modification of the McEliece scheme [McEl78] is possible to get a new public-key cryptosystem for purpose of both encryption and authentication. These two situations will be discussed in Section 6.

Finally, conclusions can be found in Section 7.

2 Notations and Definitions

In this section notations and definitions are introduced, necessary to describe the LW-scheme as presented in [LW91] more strictly.

Let $C$ be a linear $[n, k]$ code over $\mathbb{F}_2$ with code length $n$ and dimension $k$. Let $G$ be a $k \times n$ generator matrix of the code $C$ and let $H$ be an $(n - k) \times n$ parity check matrix, so $GH^T = O_{k,(n-k)}$. The code $C$ has exactly $2^k$ codewords, all are linear combinations of the rows of $G$. Let $y \in \mathbb{F}_2^n$. The set $y + C = \{y + z \mid z \in C\}$ is called a coset of $C$. Note that two vectors $y$ and $z$ are in the same coset if and only if $y - z \in C$. Each coset contains $2^k$ vectors, and there are exactly $2^{n-k}$ different cosets. The vector $s = yH^T$ is called the syndrome of $y$ with respect