Ideal Threshold Schemes from MDS Codes

Josef Pieprzyk and Xian-Mo Zhang

Centre for Advanced Computing – Algorithms and Cryptography
Department of Computing, Macquarie University
Sydney , NSW 2109, AUSTRALIA
josef,xianmo@ics.mq.edu.au

Abstract. We observe that MDS codes have interesting properties that can be used to construct ideal threshold schemes. These schemes permit the combiner to detect cheating, identify cheaters and recover the correct secret. The construction is later generalised so the resulting secret sharing is resistant against the Tompa-Woll cheating.

Key Words: Secret Sharing, Threshold Schemes, Cheating Detection and Identification

1 Introduction

In this paper we use MDS codes, i.e., maximum distance separable codes, to construct ideal threshold schemes. Based on the properties of MDS codes, in these ideal threshold schemes, cheating can be detected, cheaters can be identified and the correct secret can be recovered.

The work is structured as follows. The basic concepts of perfect and ideal secret sharing schemes/threshold schemes are introduced in Section 2, In Section 3, we briefly introduce MDS codes. In Section 4, we use MDS codes to construct ideal threshold schemes. We then propose a general construction of ideal threshold schemes in Section 5. The construction not only provides ideal threshold schemes but also protects secret sharing against the Tompa-Woll cheating. In Section 6, we prove that all these ideal threshold schemes, constructed from MDS codes, have an ability to detect incorrect shares, recover correct shares and identify the cheaters. We illustrate our considerations in Section 7. We compare this work with previous works in Section 8. Conclusions close the work.

2 Access Structures and Threshold Structures

A secret sharing scheme is a method to share a secret among a set of participants $\mathbf{P} = \{P_1, \ldots, P_n\}$. Let $\mathbf{K}$ denote the set of *secrets* and $\mathbf{S}$ denote the set of *shares*. The secret sharing has two algorithms: the distribution algorithm (dealer) and the recovery algorithm (combiner). The dealer assigns shares $s_1, \ldots, s_n \in \mathbf{S}$ to all the participants $P_1, \ldots, P_n$ respectively. Assume that ℓ

participants $P_{j_1}, \ldots, P_{j_\ell}$ are active, i.e., they currently have trading, then they submit their shares $s_{j_1}, \ldots, s_{j_\ell}$ to the combiner so as to recover a secret while other participants have no trading. Shares $s_{j_1}, \ldots, s_{j_\ell}$ together can determine a secret $K \in \mathbf{K}$ if and only if $\{P_{j_1}, \ldots, P_{j_\ell}\}$ is a qualified subset of $\mathbf{P}$. The collection of all qualified sets is called the *access structure* Γ. The access structure should be *monotone*: if $\mathcal{A} \in \Gamma$ and $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathbf{P}$ then $\mathcal{B} \in \Gamma$.

An access structure $\Gamma = \{\mathcal{A} \mid \#\mathcal{A} \geq t\}$, where $\#X$ denotes the cardinality of the set X, i.e., the number of elements in the set X and t is an integer with $0 < t \leq n$, is called a (t, n)*-threshold access structure*. A secret sharing scheme with a (t, n)-threshold access structure is called a (t, n)-threshold scheme. The parameter t is called the *threshold*.

We say that secret sharing based on an access structure Γ is *perfect* if the following two conditions are satisfied [3]:

(1) if $\mathcal{A} \in \Gamma$, then the participants in $\mathcal{A}$ can determine a secret,
(2) if $\mathcal{A} \notin \Gamma$, then the participants in $\mathcal{A}$ can determine nothing about a secret (in the information theoretic sense).

Alternatively, we say that a (t, n)-threshold scheme is *perfect* if the following two conditions are satisfied:

(1') if $\#\mathcal{A} \geq t$ then the participants in $\mathcal{A}$ can determine a secret,
(2') if $\#\mathcal{A} < t$ then the participants in $\mathcal{A}$ can determine nothing about a secret (in the information theoretic sense).

It is known [3] that for perfect secret sharing, the size of the shares has to be no smaller than the size of the secrets or $\#\mathbf{K} \leq \#\mathbf{S}$. In particular, secret sharing is said to be *ideal* if it is perfect and the size of the shares is equal to the size of the secrets or $\#\mathbf{K} = \#\mathbf{S}$. Thus ideal secret sharing is a special case of perfect secret sharing. Without loss of generality, we can assume that $\mathbf{S} = \mathbf{K}$ for ideal secret sharing.

Threshold schemes were first introduced by Blakley [1] and Shamir [9]. Ito *et al* [4] generalised secret sharing for arbitrary monotone access structure.

3 MDS Codes

In this section we introduce MDS codes, that will be used to construct ideal threshold schemes. Let $q = p^v$ where p is a prime number and v is a positive integer. We write $GF(q)$ or $GF(p^v)$ to denote the finite field of $q = p^v$ elements, and $GF(q)^n$ or $GF(p^v)^n$ to denote the vector space of n tuples of elements from $GF(q)$. Then each vector $\alpha \in GF(q)^n$ can be expressed as $\alpha = (a_1, \ldots, a_n)$ where $a_1, \ldots, a_n \in GF(q)$. We write $HW(\alpha)$ to denote the *Hamming weight* of α, i.e., the number of nonzero coordinates of α. The *Hamming distance* of two vectors α and β in $GF(q)^n$, denoted by $dist(\alpha, \beta)$, is the Hamming weight of $\alpha - \beta$.

A set $\Im$ of R vectors in $GF(q)^n$ is called an $(n, R, d)_q$ *code* if $\min\{dist(\alpha, \beta) \mid \alpha, \beta \in \Im, \alpha \neq \beta\} = d$. The parameter n is called the *length* of the code. Each vector in $\Im$ is called a *codeword* of $\Im$. In particular, if $\Im$ is a t-dimensional subspace of $GF(q)^n$, then the $(n, q^t, d)_q$ code is called *linear* and it is denoted by $[n, t, d]_q$. Since an $[n, t, d]_q$ code is a subspace of $GF(q)^n$, a linear $[n, t, d]_q$ code $\Im$ can be equivalently defined as a t-dimensional subspace of $GF(q)^n$ such that $\min\{HW(\alpha) \mid \alpha \in \Im, \alpha \neq 0\} = d$. In this work we focus our attention on linear codes. Let $\Im$ be an $[n, t, d]_q$ code. Set $\Im^\perp = \{\beta \mid \langle \beta, \alpha \rangle = 0 \text{ for all } \alpha \in \Im\}$ where $\langle \beta, \alpha \rangle$ denotes the inner product between two vectors $\beta = (b_1, \ldots, b_n)$ and $\alpha = (a_1, \ldots, a_n)$, i.e., $\langle \beta, \alpha \rangle = b_1 a_1 + \cdots + b_n a_n$. The set $\Im^\perp$ is an $(n - t)$-dimensional linear subspace of $GF(q)^n$ and it is called the *dual code* of $\Im$.

There are two methods to determine a linear code $\Im$: a generator matrix and a parity check matrix. A *generator matrix* of a linear code $\Im$ is any $t \times n$ matrix G whose rows form a basis for $\Im$. A generator matrix H of $\Im^\perp$ is called a *parity check matrix* of $\Im$. Clearly, the matrix H is of the size $(n - t) \times n$. Hence $\alpha = (a_1, \ldots, a_n) \in \Im$ if and only if $H\alpha^T = 0$.

For any $[n, t, d]_q$ code, the following inequality holds and it is known as the *Singleton bound* [7], [8], [10], $t + d \leq n + 1$. In particular, if $t + d = n + 1$ then the $[n, t, d]_q$ code is called *maximum distance separable (MDS)* [7], [10]. Clearly we can rewrite an $[n, t, d]_q$ MDS code as $[n, t, n - t + 1]_q$.

MDS codes have interesting properties, that will be used in this work. From [7], [10], we assert the validity of the lemma given below.

Lemma 1. *Let $\Im$ be an $[n, t, d]_q$ code. Then the following statements are equivalent:*

(i) $\Im$ *is an* $[n, t, n - t + 1]_q$ *MDS code,*
(ii) any t *columns of a generator matrix of* $\Im$ *are linearly independent,*
(iii) $\Im^\perp$ *is an* $[n, n - t, t + 1]_q$ *MDS code.*

The following property of MDS codes is known [7], [8], [10].

Lemma 2. *Let $\Im$ be an $[n, t, n - t + 1]_q$ MDS code. Then $n - q + 1 \leq t \leq q - 1$.*

4 Ideal Threshold Schemes from MDS Codes

Construction 1 *Let D be a generator matrix of an $[n + 1, t, n - t + 2]_q$ MDS code. Thus D is a $t \times (n + 1)$ matrix over $GF(q)$ satisfying (ii) of Lemma 1. Set*

$$(K, s_1, \ldots, s_n) = (r_1, \ldots, r_t)D \tag{1}$$

where each $r_j \in GF(q)$. For any fixed $r_1, \ldots, r_t \in GF(q)$, $K, s_1, \ldots, s_n$ can be calculated from (1). We define $s_1, \ldots, s_n$ to be the shares for participants $P_1, \ldots, P_n$ respectively, and define K to be the secret corresponding to the shares $s_1, \ldots, s_n$.

Lemma 3. *The secrets and shares, defined in Construction 1, satisfy Conditions (1') and (2') so the resulting secret sharing is a perfect (t, n)-threshold scheme.*

Proof. Index $n+1$ columns of D by $0, 1, \ldots, n$, and write $D = [\eta_0, \eta_1, \ldots, \eta_n]$, where η_j is the jth column of D. Let $P_1, \ldots, P_n$ be all the participants and $P_{j_1}, \ldots, P_{j_\ell}$ be all the currently active participants, where $1 \leq j_1 < \cdots < j_\ell \leq n$.

We first verify Condition (1'). Let $\ell \geq t$. Assume that the dealer sends shares $s_1, \ldots, s_n$ to $P_1, \ldots, P_n$ respectively, where $(s_1, \ldots, s_n)$ is **created according to (1)**. Thus $P_{j_1}, \ldots, P_{j_\ell}$ have their shares $s_{j_1}, \ldots, s_{j_\ell}$ respectively. Consider a $t \times \ell$ submatrix $D_1 = [\eta_{j_1}, \ldots, \eta_{j_\ell}]$. From (1), we get

$$(s_{j_1}, \ldots, s_{j_\ell}) = (r_1, \ldots, r_t)D_1 \tag{2}$$

Recall that D is a generator matrix of an $[n+1, t, n-t+2]_q$. Due to the statement (ii) of Lemma 1, when $\ell \geq t$, the rank of D_1 is t and then according to the properties of linear equations, $(r_1, \ldots, r_t)$ is uniquely identified by $(s_{j_1}, \ldots, s_{j_\ell})$. It follows that K is uniquely determined by $K = (r_1, \ldots, r_t)\eta_0$. This proves (1').

We next verify Condition (2'). Let $0 < \ell < t$. Consider a $t \times (1+\ell)$ submatrix $D_0 = [\eta_0, \eta_{j_1}, \ldots, \eta_{j_\ell}]$. For **any arbitrary** $K, s_{j_1}, \ldots, s_{j_\ell} \in GF(q)$, consider the system of equations on $r_1, \ldots, r_t$:

$$(K, s_{j_1}, \ldots, s_{j_\ell}) = (r_1, \ldots, r_t)D_0 \tag{3}$$

Due to (ii) of Lemma 1, when $\ell < t$, the rank of D_0 is $1 + \ell (\leq t)$. Thus, using the properties of linear equations, we conclude that (3) has solutions on $(r_1, \ldots, r_t)$ and the number of solutions is $q^{t-\ell-1}$. This number is independent to the choice of K. Thus the secret K can take any element in $GF(q)$ at an equal probability and thus there is no information on the secret. We then have proved that the scheme satisfies Condition (2'). Summarising Conditions (1') and (2'), we have proved that the secret and shares, defined in Construction 1, form a perfect (t, n)-threshold scheme. □

Corollary 1. *The secrets and shares, defined in Construction 1, form an ideal (t, n)-threshold scheme.*

Proof. According to Lemma 3, the (t, n)-threshold scheme, defined in Construction 1, is perfect. Note that each column vector η_j $(0 \leq j \leq n)$ of matrix D is nonzero. Thus $(r_1, \ldots, r_t)\eta_0$ takes all elements in $GF(q)$ when $(r_1, \ldots, r_t)$ takes all vectors in $GF(q)^t$. This implies that $\mathbf{K} = GF(q)$. On the other hand, for each j with $1 \leq j \leq n$, $(r_1, \ldots, r_t)\eta_j$, takes all elements in $GF(q)$ when $(r_1, \ldots, r_t)$ takes all vectors in $GF(q)^t$. This implies that $\mathbf{S} = GF(q)$. By definition, we know that the scheme is ideal. □

We now explain how the scheme works. The matrix D is public but $(r_1, \ldots, r_t)$ is chosen secretly by the dealer. From $(r_1, \ldots, r_t)$, the dealer (distribution algorithm) computes $(s_1, \ldots, s_n)$ based on (1). The dealer sends the shares $s_1, \ldots, s_n$ to participants $P_1, \ldots . P_n$ respectively via secure channels. Assume that $P_{j_1}, \ldots$, P_{j_ℓ} are the currently active participants, where $1 \leq j_1 < \cdots < j_\ell \leq n$. $P_{j_1}, \ldots, P_{j_\ell}$ submit their shares to the combiner (recovery algorithm). The combiner recovers the secret. There are two cases: $\ell \geq t$ and $\ell < t$. According to Lemma 3 and its proof, if $\ell \geq t$, then the combiner can uniquely determine

$(r_1, \dots, r_t)$ and then identify the secret $K = (r_1, \dots, r_t)\eta_0$, while in the case of $\ell < t$, the secret can be any element in $GF(q)$ with the same probability so the combiner knows nothing about the secret.

5 More General Constructions of Ideal Threshold Schemes

In this section, we generalise Construction 1.

Construction 2 *Let D be a generator matrix of an $[n+1, t, n-t+2]_q$ MDS code. Thus D is a $t \times (n+1)$ matrix over $GF(q)$ satisfying (ii) of Lemma 1. Let $\pi_0, \pi_1, \dots, \pi_n$ be permutations on $GF(q)$. Set*

$$(K, s_1, \dots, s_n) = (r_1, \dots, r_t)D \tag{4}$$

and

$$(K^*, s_1^*, \dots, s_n^*) = (\pi_0(K), \pi_1(s_1), \dots, \pi_n(s_n)) \tag{5}$$

where each $r_j \in GF(q)$. For any fixed $r_1, \dots, r_t \in GF(q)$, $K^*, s_1^*, \dots, s_n^*$ can be calculated from (4) and (5). We define $s_1^*, \dots, s_n^*$ to be the shares for participants $P_1, \dots, P_n$ respectively, and define K^* to be the secret corresponding to the shares $s_1^*, \dots, s_n^*$.

Theorem 1. *The secrets and shares, defined in Construction 2, form not only a perfect but also an ideal (t, n)-threshold scheme.*

Proof. Let $P_1, \dots, P_n$ be all the participants and $P_{j_1}, \dots, P_{j_\ell}$ be all the currently active participants, where $1 \le j_1 < \dots < j_\ell \le n$.

We first verify Condition (1'). Let $\ell \ge t$. Assume that the dealer sends the shares $s_1^*, \dots, s_n^*$ to $P_1, \dots, P_n$ respectively where $(s_1^*, \dots, s_n^*)$ is **created according to (5)**. Then $P_{j_1}, \dots, P_{j_\ell}$ have their shares $s_{j_1}^*, \dots, s_{j_\ell}^*$ respectively. Clearly, there uniquely exists a $(s_{j_1}, \dots, s_{j_\ell})$ such that $s_{j_1}^* = \pi_{j_1}(s_{j_1}), \dots, s_{j_\ell}^* = \pi_{j_\ell}(s_{j_\ell})$. Due to the same reasoning as in the proof of Lemma 3, $(r_1, \dots, r_t)$ is uniquely identified by $(s_{j_1}, \dots, s_{j_\ell})$. It follows that K is uniquely determined by $(r_1, \dots, r_t)$. Thus $K^* = \pi(K)$ is uniquely determined. This proves (1').

We next verify Condition (2'). Let $0 < \ell < t$. For **any arbitrary** K^*, $s_{j_1}^*$, $\dots$, $s_{j_\ell}^* \in GF(q)$, there uniquely exists a $(s_{j_1}, \dots, s_{j_\ell})$ such that $s_{j_1}^* = \pi_{j_1}(s_{j_1})$, $\dots$, $s_{j_\ell}^* = \pi_{j_\ell}(s_{j_\ell})$. Due to the same reasoning as in the proof of lemma 3, for these $s_{j_1}, \dots, s_{j_\ell}$, (3) has solutions on $(r_1, \dots, r_t)$, and the number of solutions is $q^{t-\ell-1}$. This number is independent to the choice of K, and thus K can take any element in $GF(q)$ at an equal probability. It follows that K^* can take any element in $GF(q)$ at an equal probability, and then there exists no information on the key. We have proved that the scheme satisfies Condition (2'). Summarising Conditions (1') and (2'), we have proved that the secret and shares, defined in Construction 2, form a perfect (t, n)-threshold scheme. Due to Corollary 1, we know that this scheme is ideal. □

Clearly the schemes in Construction 1 are special schemes in Construction 2 when $\pi_0, \pi_1, \ldots, \pi_n$ are all the identity permutation on $GF(q)$.

We now explain how the scheme works. The matrix D and the $n+1$ permutations $\pi_0, \pi_1, \ldots, \pi_n$ are public but $(r_1, \ldots, r_t)$ is chosen secretly by the dealer. From $(r_1, \ldots, r_t)$, the dealer (distribution algorithm) computes $(s_1, \ldots, s_n)$ based on (4), then $(s_1^*, \ldots, s_n^*)$ based on (5). After that, the dealer sends the shares $s_1^*, \ldots, s_n^*$ to participants $P_1, \ldots, P_n$ respectively, via the secure channels. Assume that $P_{j_1}, \ldots, P_{j_\ell}$ are the currently active participants, where $1 \leq j_1 < \cdots < j_\ell \leq n$, and they wish to recover the secret. They submit their shares to the combiner (recovery algorithm). There are two cases: $\ell \geq t$ and $\ell < t$. According to Theorem 1, if $\ell \geq t$, then the combiner can uniquely determine $(r_1, \ldots, r_t)$ from (4), identify K from (4), and finally identify the secret $K^* = \pi_0(K)$ from (5). In the case when $\ell < t$, the secret may take any element in $GF(q)$ with uniform probability so the secret cannot be determined.

In contrast to Construction 1, Construction 2 not only provides ideal threshold schemes but also improves the schemes in Construction 1. In fact, all the possible share vectors $(s_1, \ldots, s_n)$ in a (t, n)-threshold scheme by Construction 1 form a linear subspace of $GF(q)^n$ as MDS codes are linear codes. Usually, this is not a desirable property from a point of information security as this case gives a chance to the Tompa-Woll attack [11]. To remove this drawback, we consider schemes in Construction 2. For example, we choose $\pi_0, \pi_1, \ldots, \pi_{t-1}$ to be the identity permutation on $GF(q)$ but we require the permutations $\pi_t, \ldots, \pi_n$ on $GF(q)$ to satisfy $\pi_t(0) \neq 0, \ldots, \pi_n(0) \neq 0$. It is easy to verify that all the possible share vectors $(s_1^*, \ldots, s_n^*)$ in the (t, n)-threshold scheme by Construction 2 do not form a linear subspace of $GF(q)^n$, as $(s_1^*, \ldots, s_n^*)$ cannot take $(0, \ldots, 0) \in GF(q)^n$.

6 Cheating Detection and Cheater Identification

In this section, we show that the ideal threshold schemes constructed in Construction 2 have an ability to find whether the shares, submitted by participants to the combiner, are correct, or in other words, the modified shares can be detected. The (t, n)-threshold schemes, defined in Construction 2, have the following property.

Theorem 2. *Let $K^*, s_1^*, \ldots, s_n^*$, $K, s_1, \ldots, s_n$ and $r_1, \ldots, r_t$ satisfy (4) and (5), and $K'^*, s_1'^*, \ldots, s_n'^*$, $K', s_1', \ldots, s_n'$ and $r_1', \ldots, r_t'$ also satisfy (4) and (5). If $(r_1, \ldots, r_t) \neq (r_1', \ldots, r_t')$ then the Hamming distance between $(K^*, s_1^*, \ldots, s_n^*)$ and $(K'^*, s_1'^*, \ldots, s_n'^*)$ is at least $n - t + 2$.*

Proof. Recall that $K^* = \pi_0(K)$, $s_1^* = \pi_1(s_1)$, …, $s_n^* = \pi_n(s_n)$, and $K'^* = \pi_0(K')$, $s_1'^* = \pi_1(s_1')$, …, $s_n'^* = \pi_n(s_n')$. Thus we know that

$$K^* = K'^* \text{ if and only if } K = K', \tag{6}$$

$$s_j^* = s_j'^* \text{ if and only if } s_j = s_j' \ (j = 1, \ldots, n) \tag{7}$$

Since $(r_1, \ldots, r_t) \neq (r'_1, \ldots, r'_t)$ and the rank of the matrix D in (4) or (1) is equal to t, we know that $(K, s_1, \ldots, s_n)$ and $(K', s'_1, \ldots, s'_n)$ are two distinct codewords of an $[n+1, t, n-t+2]_q$ MDS code. Thus the Hamming distance between $(K, s_1, \ldots, s_n)$ and $(K', s'_1, \ldots, s'_n)$ is at least $n-t+2$. On the other hand, according to (6) and (7), we know that the Hamming distance between $(K^*, s^*_1, \ldots, s^*_n)$ and $(K'^*, s'^*_1, \ldots, s'^*_n)$ is equal to the Hamming distance between $(K, s_1, \ldots, s_n)$ and $(K', s'_1, \ldots, s'_n)$. This proves the theorem. □

The following property [10] of codes will be used in this work:

Lemma 4. *Let $\Im$ be an $(n, R, d)_q$ code. For any j with $1 \leq j \leq n$, the code $\Im_0$, obtained by removing the jth coordinate from all codewords of $\Im$, is a code $(n-1, R, d-1)_q$ or $(n-1, R, d)_q$.*

Given an $[n+1, t, n-t+2]_q$ MDS code $\Im$ with a generator matrix D and $n+1$ permutations $\pi_0, \pi_1, \ldots, \pi_n$. According to Theorem 1, we have an ideal threshold scheme defined in Construction 2. Let $P_1, \ldots, P_n$ be the participants. We keep using all the notations in Sections 4 and 5. The dealer selects $r_1, \ldots, r_t \in GF(q)$ then computes $s_1, \ldots, s_n \in GF(q)$ by (4), and then $s^*_1, \ldots, s^*_n \in GF(q)$ by (5). The dealer sends the shares $s^*_1, \ldots, s^*_n$ to $P_1, \ldots, P_n$ respectively. Let $P_{j_1}, \ldots, P_{j_\ell}$ be all the currently active participants, where $1 \leq j_1 < \cdots < j_\ell \leq n$.

Consider a $t \times \ell$ submatrix D_1 consisting of ℓ columns of D, indexed by $j_1, \ldots, j_\ell$. Set

$$W_0 = \{(s^*_{j_1}, \ldots, s^*_{j_\ell}) = (\pi_{j_1}(s_{j_1}), \ldots, \pi_{j_\ell}(s_{j_\ell})) \mid (s_{j_1}, \ldots, s_{j_\ell}) = (r_1, \ldots, r_t)D_1, \\ r_1, \ldots, r_t \in GF(q)\} \qquad (8)$$

According to Theorem 2 and Lemma 4, we state

Lemma 5. *Any two distinct vectors in W_0, defined in (8), have a Hamming distance at least $\ell - t + 1$.*

6.1 Cheating Detection

Assume that $P_{j_1}, \ldots, P_{j_\ell}$ submit their modified shares $s^*_{j_1} + \delta_1, \ldots, s^*_{j_\ell} + \delta_\ell$ to the combiner (recovery algorithm) where each $\delta_j \in GF(q)$. Thus P_{j_i} is honest if $\delta_i = 0$, otherwise he cheats. We write

$$\beta = (s^*_{j_1}, \ldots, s^*_{j_\ell}),\ \delta = (\delta_1, \ldots, \delta_\ell) \text{ and } \tilde{\beta} = \beta + \delta \qquad (9)$$

Assume that $HW(\delta_1, \ldots, \delta_\ell) \leq \ell - t$. Clearly

$$dist(\tilde{\beta}, \beta) = HW(\delta) \leq \ell - t \qquad (10)$$

Theorem 3. *Given an $[n+1, t, n-t+2]_q$ MDS code with a generator matrix D and $n+1$ permutations $\pi_0, \pi_1, \ldots, \pi_n$. According to Theorem 1, we have an ideal (t,n)-threshold scheme defined in Construction 2. Let $P_1, \ldots, P_n$ be all the participants and $P_{j_1}, \ldots, P_{j_\ell}$ $(t < \ell \leq n)$ be all the participants who are currently active. Assume that no more than $\ell - t$ cheaters who submit incorrect shares. Then $\tilde{\beta}$, where $\tilde{\beta}$ has been defined in (9), is correct if and only if $\tilde{\beta} \in W_0$, where W_0 has been defined in (8), or in other words, the combiner can find that $\tilde{\beta}$ is correct or incorrect according to $\tilde{\beta} \in W_0$ or $\tilde{\beta} \notin W_0$.*

Proof. Assume that $\tilde{\beta}$ is correct, or in other words, $\delta = (\delta_1, \ldots, \delta_k) = (0, \ldots, 0)$ where δ has been defined in (9). Thus $\tilde{\beta}$ is identical with the β. In this case $\tilde{\beta} = \beta \in W_0$. Conversely, assume that $\tilde{\beta} \in W_0$. We now prove by contradiction that $\tilde{\beta} = \beta$. Assume that $\tilde{\beta} \neq \beta$. According to Lemma 5, $\tilde{\beta}$ and β have a Hamming distance at least $\ell - t + 1$. This contradicts (10). The contradiction proves that $\tilde{\beta}$ must be identical with β and thus $\tilde{\beta} = \beta$ is correct. Thus we have proved that $\tilde{\beta}$ is correct if and only if $\tilde{\beta} \in W_0$. □

6.2 Cheater Identification

In Section 6.1 the combiner can detect incorrect shares sent by participants, however there is no guarantee that it can identify the cheaters or reconstruct the correct shares (and the secret). In this section we consider how to identify the cheaters and how to recover the correct shares. We keep using all the assumptions and the notations in Section 6.1. We additionally suppose that $\delta = (\delta_1, \ldots, \delta_\ell)$ satisfies

$$0 < HW(\delta) \leq \lfloor \frac{1}{2}(\ell - t) \rfloor \tag{11}$$

where $\lfloor r \rfloor$ denotes the maximum integer no larger than r.

Due to (11) and Theorem 3, the combiner knows that $\tilde{\beta}$ is incorrect by the fact $\tilde{\beta} \notin W_0$. The combiner further determines a vector $\gamma_0 \in W_0$ such that

$$dist(\tilde{\beta}, \gamma_0) = \min\{dist(\tilde{\beta}, \gamma) \mid \gamma \in W_0\} \tag{12}$$

We now prove by contradiction that γ_0 is identical with β. Assume that $\gamma_0 \neq \beta$. Since $\gamma_0, \beta \in W_0$, due to Lemma 5, we know that

$$dist(\gamma_0, \beta) \geq \ell - t + 1 \tag{13}$$

Recall that $dist(\tilde{\beta}, \beta) = HW(\delta) \leq \lfloor \frac{1}{2}(\ell - t) \rfloor$, we have

$$dist(\tilde{\beta}, \gamma_0) = \min\{dist(\tilde{\beta}, \gamma) \mid \gamma \in W_0\} \leq dist(\tilde{\beta}, \beta) \leq \lfloor \frac{1}{2}(\ell - t) \rfloor \tag{14}$$

Clearly $dist(\gamma_0, \beta) \leq dist(\gamma_0, \tilde{\beta})$ $+dist(\tilde{\beta}, \beta)$. Thus $dist(\gamma_0, \beta) \leq dist(\gamma_0, \tilde{\beta}) + HW(\delta)$. Due to (14), we have

$$dist(\gamma_0, \beta) \leq \lfloor \frac{1}{2}(\ell - t) \rfloor + \lfloor \frac{1}{2}(\ell - t) \rfloor \leq \ell - t < \ell - t + 1 \tag{15}$$

Obviously, (15) contradicts (13). The contradiction disproves the assumption that $\gamma_0 \neq \beta$. Therefore γ_0 and β must be identical.

After knowing γ_0, i.e., β, the combiner can identify the δ as he has received the vector of $\tilde{\beta} = \beta + \delta$. So we can formulate the following theorem.

Theorem 4. *Given an $[n+1, t, n-t+2]_q$ MDS code with a generator matrix D and $n+1$ permutations $\pi_0, \pi_1, \ldots, \pi_n$. According to Theorem 1, we have an ideal (t, n)-threshold scheme defined in Construction 2. Let $P_1, \ldots, P_n$ be all the participants and $P_{j_1}, \ldots, P_{j_\ell}$ $(t < \ell \leq n)$ be all the participants who are currently active. If the number of cheaters is less than or equal to $\lfloor \frac{1}{2}(\ell - t) \rfloor$ then the combiner can identify the cheaters who submitted incorrect shares also recover the correct shares by determining the vector $\gamma_0 \in W_0$ where W_0 has been defined in (8) and γ_0 satisfies (12).*

Summarising Theorems 3 and 4, the combiner first checks whether the share vector $\tilde{\beta}$, that he received from the active participants, is correct. If $\tilde{\beta}$ is incorrect, the combiner further determines who are cheaters and reconstructs the correct shares. We notice that both Theorems 3 and 4 require the parameter ℓ to be greater than t.

7 Examples

Example 1. There exists an MDS code $[18, 9, 10]_{25}$, that is also a quadratic residue code (Chapter 4 of [8]). Let D denote a general matrix of this code. For any permutations $\pi_0, \pi_1, \ldots, \pi_{17}$ on $GF(25)$, according to Theorem 1, we can construct an ideal $(9, 17)$-threshold scheme over $GF(25)$ in Construction 2. Let ℓ $(9 < \ell \leq 17)$ denote the number of currently active participants. Due to Theorems 3 and 4, this scheme has the ability to detect cheating and identify cheaters. More precisely, if there are no more than $\ell - 9$ participants who submit incorrect shares then the incorrect shares can be detected. Furthermore, if there are no more than $\lfloor \frac{1}{2}(\ell - 9) \rfloor$ participants submitting incorrect shares then all the cheaters can be identified and the correct shares can be recovered.

Example 2. Let $GF(q) = \{0, \lambda_1, \ldots, \lambda_{q-1}\}$ and t be an integer with $2 \leq t \leq q-1$. Set

$$E = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 0 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_{q-1} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_{q-1}^2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{t-1} & \lambda_2^{t-1} & \cdots & \lambda_{q-1}^{t-1} & 0 & 1 \end{bmatrix} \tag{16}$$

From [7], [10], E is a generator matrix of a $[q+1, t, q-t+2]_q$ MDS code. For any permutations $\pi_0, \pi_1, \ldots, \pi_q$ on $GF(q)$, according to Theorem 1, we can construct an ideal (t, q)-threshold scheme over $GF(q)$ in Construction 2. Let ℓ

$(t < \ell \leq n)$ denote the number of currently active participants. Due to Theorems 3 and 4, this scheme has the ability to detect cheating and identify cheaters. More precisely, if there are no more than $\ell - t$ participants who submit incorrect shares then the incorrect shares can be detected. Furthermore, if there are no more than $\lfloor \frac{1}{2}(\ell - t) \rfloor$ participants submitting incorrect shares then all the cheaters can be identified and the correct shares can be recovered.

8 Comparing This Work with Previous Results

Comparing Shamir scheme [9] with the ideal threshold scheme in Example 2, we can find: (a) k in Shamir scheme is corresponding to t in Example 2, (b) the coefficients $a_0, a_1, \ldots, a_{k-1}$ of the polynomial $q(x)$ in Shamir scheme are corresponding to $r_1, \ldots, r_t$ in Example 2 respectively, (c) the shares $D_1 = q(1), \ldots, D_n = q(n)$ in Shamir scheme are corresponding to $s_1, \ldots, s_n$ in Example 2 respectively, (d) if we remove the last two columns of E in Example 2 and change the entries of E, then we obtain

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 2^{t-1} & \cdots & n^{t-1} \end{bmatrix} \tag{17}$$

where the entries are elements in the residue modulo class of prime p $(t \leq n \leq p-1)$, then we regain Shamir scheme. This shows that the Lagrange interpolation suggested in [9] can be re-obtained from Example 2.

McEliece and Sarwate [6] generalised Shamir's construction as they allowed the elements in the Lagrange interpolation to be from a finite field, instead of only elements in a prime filed. They also indicated that the share vectors form Reed-Solomon codes and then their schemes can correct modified shares. As known, Reed-Solomon codes are special MDS codes and MDS codes are not necessarily Reed-Solomon codes. Thus Constructions 1 and 2 are more general.

Karnin, Greene and Hellman obtained a similar result (Theorem 2 of [5]) to Construction 1. There is, however, a basic difference between this work and their work. The difference is in the definitions of (t, n) threshold schemes. In our definition, we allow t or more participants to collaborate in recovery of the secret. In fact, the cheating detection relies on the existence of redundant shares so they can be used to identify incorrect ones (then identify cheaters) and to recover the correct secret. Karnin *et al* considered threshold schemes in which the number of active participants is precisely equal to t. However, as mentioned in Theorem 6 of [5], cheating detection is impossible in this case.

Summarising the above discussions, the above previous schemes are all special cases in Construction 1. However Construction 1 is a special case of Construction 2. In addition, according to Theorem 1, we are sure that all the threshold schemes in Constructions 1 and 2 are ideal. However this property was not mentioned in the above papers.

9 Conclusions

Using interesting properties of MDS codes, we have constructed ideal threshold schemes and indicated that incorrect shares can be detected and the cheaters can be identified, furthermore the correct secret can be recovered. We have further suggested a general construction that not only provides more ideal threshold schemes but also prevents Tompa-Woll attack.

Acknowledgement

The work was partially supported by Australian Research Council grant A00103078.

References

1. G. R. Blakley. Safeguarding cryptographic keys. In *Proc. AFIPS 1979 National Computer Conference*, pages 313–317. AFIPS, 1979. E.F. Brickell and D.R. Stinson.
2. E. F. Brickell and D. M. Davenport. On the Classification of Ideal Secret Sharing Schemes. J. Cryptology, 4: 123 - 134, 1991.
3. E. F. Brickell and D.R. Stinson. Some Improved Bounds on Information Rate of Perfect Sharing Schemes J. Cryptology, 5: 153 - 166, 1992.
4. M. Ito, A. Saito, and T. Nishizeki. Secret sharing scheme realizing general access structure. In *Proceedings IEEE Globecom '87*, pages 99–102. IEEE, 1987.
5. E.D. Karnin, J.W. Greene, and M.E. Hellman. On secret sharing systems. *IEEE Transactions on Information Theory*, IT-29:35–41, 1983.
6. R.J. McEliece and D. V. Sarwate. *On Sharing Secrets and Reed-Solomon Codes.* Communications of the ACM, Vol. 24, 1981, pp 583-584.
7. F.J. MacWilliams and N.J.A. Sloane. *The theory of error-correcting codes.* North-Holland, Amsterdam, Seventh Impression 1992.
8. V. C. Pless and W. C. Huffman, Editors. Handbook of Coding Theory, Elsevier Science B. V., 1998.
9. A. Shamir. How to share a secret. *Communications of the ACM*, 22:612–613, November 1979.
10. S. Roman. Coding and Information Theory. Springer-Verlag, Berlin, Heidelberg, New York, 1992.
11. M. Tompa and H. Woll. How to share a secret with cheaters. *Journal of Cryptology*, 1(2):133–138, 1988.