Secret Sharing Using Near-MDS Codes

Sanyam Mehta[2], Vishal Saraswat[1(✉)], and Smith Sen[2]

[1] Robert Bosch Engineering and Business Solutions Pvt. Ltd. (RBEI/ESY), Bangalore, India
vishal.saraswat@gmail.com

[2] Birla Institute of Technology and Science, Pilani, K. K. Birla Goa Campus, Goa, India
{f20140526,f20140896}@goa.bits-pilani.ac.in

Abstract. We propose a generalized secret sharing scheme based on NMDS codes. The proposed scheme is efficient and the computational complexity for setup and reconstruction phase is only $O(n^3)$, where n is the number of participants. The scheme admits an access structure based on two mutually exclusive sets of participant combinations of sizes t and $t-1$ respectively. The parameter t for the access structure is independent of the field size. The proposed scheme is ideal and perfect and has desirable security features of cheating detection and cheater identification. We also provide a cryptanalysis of the $(t+1, n)$ threshold secret sharing scheme based on NMDS codes proposed in [12]. We show that their scheme is insecure and that there always exists a set of m participants, where $m < t+1$, which can reconstruct the secret.

Keywords: Secret sharing · Generalized access structure · MDS codes · Near-MDS codes

1 Introduction

Secret sharing schemes were independently proposed by Blakley [2] and Shamir [11] in 1979. The scheme by Shamir is based on linear algebra and the standard Lagrange's interpolation while the scheme given by Blakley is built upon the idea of finite geometries, particularly on the concept of intersection of hyperplanes. These were the first threshold secret sharing schemes which allowed a *secret* s to be split into n *shares* which could be distributed among n *participants* (or *users*), $\mathcal{P} = \{P_1, \ldots, P_n\}$, in such a way that for some threshold t, $1 \leq t \leq n$, any group of t or more participants could pool in their shares to *reconstruct* the secret but if the number of participants in a group is less than the threshold t, then that group does not get any extra information about the secret.

Note that a threshold secret sharing scheme makes the authorized sets rigid. Consider a hypothetical situation of a firm, where we have three levels of the workforce, namely, directors, managers and employees. To access the key for a

C. Carlet et al. (Eds.): C2SI 2019, LNCS 11445, pp. 195–214, 2019.
https://doi.org/10.1007/978-3-030-16458-4_12

certain new product launched, the firm does not want every employee to open up the locker. So naturally more power has to be given to the directors and managers than the employees. This in turn means, maybe less directors can pool their shares to find the secret. Comparatively more managers and even more employees should be required to find the secret. This type of flexibility is not directly possible with a threshold access structure.

To overcome the problems arising in the scenario above, Ito et al. [7] introduced the concept of a generalized access structure which contains all the possible sets of participants who can reconstruct the secret. They proposed a secret sharing scheme which admitted a generalized access structure irrespective of the size of each set. Every participant is assigned several shadows of a (t, n)-threshold secret sharing scheme by the dealer. For integers t, m and q satisfying $t \leq m < q$, q being a prime power, the dealer

- samples $\alpha_1, \ldots, \alpha_{t-2}$ from GF(q) and α_{t-1} from GF(q) $-$ {0} and constructs $f(x) = s + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_{t-1} x^{t-1}$, where the secret is $f(0) = s$.
- samples $x_1, \ldots, x_m$ from GF(q) $-$ {0} and computes $s_j = f(x_j)$ $(1 \leq j \leq m)$.
- chooses $S_i \subset \{(x_1, s_1), \ldots, (x_m, s_m)\}$ and assigns S_i to each participant P_i $(1 \leq i \leq |\mathcal{P}|)$.

In this scheme, the access structure Λ can be defined as the sets in which the union of all the shares of the participants has cardinality greater than t. In case the minimal subsets are big, this scheme turns out to be ineffective. At the same time, for this access structure, Shamir's Scheme gives each party a share equal to the size of the secret.

Later, Benaloh and Leichter [1] proposed a simpler and relatively efficient secret sharing scheme exploiting the *monotonicity* property of access structures in secret sharing schemes. The proposed scheme begins with multiple schemes for simple access structures and creates a scheme for composition of those access structures. Thus, the scheme by Benaloh and Leichter efficiently realizes every access structure that can be described by small monotone formula. Even though this construction is more efficient and generalizes the scheme proposed by Ito et al. [7], the length of shares become exponential in the number of parties.

All the above schemes were linear secret sharing schemes which can be modelled using *monotone span programs* [8]. These are essentially matrices describing a linear mapping of a linear scheme. These variations are very efficient to implement. Around the same time, the following advantages of using linear codes, instead of arbitrary matrices, for designing secret sharing schemes were observed.

- Easier to detect errors and easy transmission.
- Can be defined using a single generator matrix.
- Schemes were still efficient although features for verification and cheating detection and identification were included.

Some of such constructions are based on Maximum Distance Separable (MDS) matrices [10] where the dealer chooses an MDS matrix A of dimension $k \times n$ and a vector $\mathbf{v}$ of dimension $1 \times k$, and computes the codeword $\mathbf{v} \cdot A$ whose first element is the desired secret.

These schemes have been extensively used as they are easier to implement. It has been observed that NMDS matrices have better computational performance than MDS matrices. NMDS matrices require less storage than MDS matrices, do not require additional clock cycles and have sub-optimal branch numbers [9]. Some lightweight block ciphers have already been using NMDS matrices for their diffusion layer which includes ciphers like PRINCE, FIDES, PRIDE and MANTIS due to its benefits of low power, low energy and low latency in implementations. Due to all these, it will be useful to implement a secret sharing scheme using NMDS codes as well. The advantages of using NMDS codes for making a secret sharing scheme are as follows,

- Easy to implement and less space consuming
- Does not require additional clock cycles
- Difficult to identify the generator matrix of the code for an adversary
- Richer access structure than MDS secret sharing
- Has the property of cheating detection and identification like MDS secret sharing

Considering Shamir's secret sharing scheme, one can observe that a dishonest participant cannot find the secret by giving a wrong share but can misguide the honest participants by getting a wrong key which ultimately results in failure for the authorized set of participants to obtain the correct secret. Numerous solutions have been suggested in literature to solve this issue and retrieve the accurate secret. Some propose error correcting codes where a tampered share is treated as an error and is corrected using the error correcting property of code. While, some other propose to use a protocol where dealer validates individual shares in an authorized set to detect and rectify any tampering of shares. One plausible way is to use check vectors that dealer uses as certification for each participant. Some such schemes have been reviewed in Sect. 3.

Most of the initial secret sharing schemes had issues with trusted third parties (dealers and combiners) as well as cheating detection and identification. One of the modifications of Shamir's scheme for cheating detection and correction is proposed by Lein et al. in [5]. It is assumed that $m > t$ number of participants have to come up with their shares, where t is the threshold, giving the participants $\binom{m}{t}$ ways to pool their shares. For each way we get a degree $t-1$ polynomial by interpolation which can be checked with the original polynomial. Participants who are present in the majority of groups and couldn't recover the same polynomial are grouped as possible cheaters and shares are corrected recursively unless there is no cheater left in the group of participants. They provide the algorithm for cheating detection and cheating correction by trading off the time and space-complexities for computing the secret to prevent cheating.

1.1 Organization

In this paper, we have introduced some essential definitions, problems and assumptions related to the scheme in Sect. 2. Further we have discussed in Sect. 3

some of the previous work done on secret sharing schemes formed using MDS (Maximum Distance Separable) and NMDS (Near-Maximum Distance Separable) codes. In Sect. 4, we have analyzed and discussed the shortcomings of the scheme proposed by Zhou et al. [12]. The proposed secret sharing scheme is presented in Sect. 5 and its security analyzed in Sect. 6. Finally, we provide a concrete instantiation of the proposed scheme in the Appendix A.

2 Preliminaries

The assignment operator is denoted by "$\leftarrow$". In particular, the operation of running a randomized or deterministic algorithm A with input x and storing the result to the variable y is denoted by $y \leftarrow A(x)$. The operation of choosing an element x of a set X randomly at uniform from X is denoted by $x \xleftarrow{\$} X$ and the operation of choosing an element x of X according to a distribution μ on a set X is denoted by $x \xleftarrow{\mu} X$. For a given function $f : N \to [0, 1]$ is said to be *negligible in* n if $f(n) < 1/p(n)$ for a polynomial p and having sufficiently large n.

2.1 Coding Theory

Definition 1 (Block Codes). *Let* $\mathbb{A}^n$ *be the set of all strings of length* n *over* $\mathbb{A}$*, where* $\mathbb{A} = \{a_0, \ldots a_{q-1}\}$*. Any nonempty subset* $\mathbf{C}$ *of* $\mathbb{A}^n$ *is called a* q-ary block code *of* length n *and each string in* $\mathbf{C}$ *is called a* codeword*.*

Definition 2 (Hamming Distance). *If* $\mathbf{c}$ *and* $\mathbf{d}$ *are two codewords of length* n*, the* hamming distance $d(\mathbf{c}, \mathbf{d})$ *is the number of positions at which* $\mathbf{c}$ *and* $\mathbf{d}$ *differ. The* hamming weight $d(\mathbf{c})$ *of a codeword* $\mathbf{c} \in \mathbf{C}$ *is defined to the number of its non-zero coordinate positions.*

Definition 3 (Minimum Distance). *The* minimum distance *of a code* $\mathbf{C}$ *is defined as*

$$d(\mathbf{C}) = \min_{\mathbf{c},\mathbf{d} \in \mathbf{C}} d(\mathbf{c}, \mathbf{d}) .$$

Definition 4 (Linear Code). *A* linear code*,* $\mathbf{L}$*, of length* n *is a linear subspace of* $\mathbb{F}_q^n$ *where* $\mathbb{F}_q = \mathrm{GF}(q)$ *is the Galois Field of order* $q = p^m$ *for some prime* p *and power* $m \geq 1$*. If* $\mathbf{L}$ *has a dimension* k*, we say it is an* $[n, k]$ code*. Further, if its minimum distance is* d*, we say* $\mathbf{L}$ *is an* $[n, k, d]$-code*.*

Definition 5 (Generator Matrix). *A* generator matrix*, of a linear* $[n, k, d]$ *code,* $\mathbf{L}$*, is a* $m \times n$ *matrix,* $m \geq k$*, whose rows span* $\mathbf{L}$*. That is, every codeword* $\mathbf{c} \in \mathbf{L}$ *is a linear combination of the rows of the generator matrix,* G*, and for every* $\mathbf{v} \in \mathbb{F}_q^m$*,* $\mathbf{v} \cdot G$ *is a codeword in* $\mathbf{L}$*.*

Remark 1. For the purposes of this paper, we will assume that the number of rows in generator matrices is equal to the dimension of the code. That is, if G is a generator matrix of a linear $[n, k, d]$ code, $\mathbf{L}$, then G is a $k \times n$ matrix and its rows form a basis of $\mathbf{L}$.

Definition 6 (Standard Form of a Generator Matrix). *A generator matrix, G, of a linear $[n, k, d]$ code, $\mathbf{L}$, is said to be in the* standard form *if $G = [I_k \mid A]$, where I_k is the identity matrix of size $k \times k$ and A is a $k \times (n-k)$ matrix. For such G, the code $\mathbf{L}$ is said to* systematic *in its first k coordinate positions.*

Remark 2. Any $k \times (n-k)$ matrix, A, defines a linear $[n, k, d]$ code, $\mathbf{L}$, completely by defining the standard form of a generator matrix, $G = [I_k \mid A]$, for $\mathbf{L}$.

Definition 7 (Support of a Code). *Let $\mathbf{C}$ be a code over $\mathbb{F}_q$. The* support*, $\mathrm{Supp}(\mathbf{c})$, of a codeword $\mathbf{c} \in \mathbf{C}$ is defined to be the set of its non-zero coordinate positions. The* support*, $\mathrm{Supp}(\mathbf{C})$, of the code $\mathbf{C}$ is defined to be*

$$\mathrm{Supp}(\mathbf{C}) = \cup_{\mathbf{c} \in \mathbf{C}} \mathrm{Supp}(\mathbf{c})\,.$$

Definition 8 (Generalized Hamming Distance). *The minimum cardinality of the supports of the $[n, r]$-subcodes of $\mathbf{C}$, for $1 \leq r \leq k$, is called the r*th generalized hamming distance *$d_r\mathbf{C}$. That is,*

$$d_r\mathbf{C} = \min\{|\mathrm{Supp}\,\mathbf{D}| \colon \mathbf{D} \text{ is } [n, r]_q \text{ subcode of } \mathbf{C}\}\,.$$

Remark 3. Note that $d_1(\mathbf{C}) = d(\mathbf{C})$ is the Hamming Distance of $\mathbf{C}$.

Proposition 1 (Hierarchy of Hamming Weights). *For every linear $[n, k]_q$ code $\mathbf{C}$,*

$$0 < d_1(\mathbf{C}) < d_2(\mathbf{C}) < \cdots < d_k(\mathbf{C}) \leq n\,.$$

Definition 9 (The Singleton Bound). *The* singleton bound *states that any $[n, k, d]$-code must satisfy*

$$q^k \leq q^{n-d+1}\,.$$

In particular, $d \leq n + 1 - k$.

Definition 10 (The Generalized Singleton Bound). *The r*th generalized singleton bound *is given by $d_r(\mathbf{C})$*

$$d_r(\mathbf{C}) \leq n - k + r \text{ where } r = 1, 2, \ldots, k\,.$$

Definition 11 (Maximum Distance Separable Codes). *A linear $[n, k, n + 1 - k]$-code, that is, an $[n, k]$-code with largest possible minimum distance, is called a* maximum distance separable (MDS) *code.*

Proposition 2 (Properties of MDS Matrices). *Given an $[n, k, d]$ MDS code, $\mathbf{L}$, over $\mathbb{F}_q$, the accompanying explanations are proportionate.*

1. *If $G = [I_k \mid A]$ is a generator matrix of $\mathbf{L}$ in standard form, then every square submatrix of A is non-singular.*
2. *Any k columns of a generator matrix for $\mathbf{L}$ are linearly independent.*
3. *Any $n - k$ columns of a parity check matrix for $\mathbf{L}$ are linearly independent.*

Definition 12 Almost-MDS Codes). *The class of codes with* $d_1(\mathbf{C}) = n - k$ *are called* almost-MDS (AMDS) *codes.*

Definition 13 (Near-MDS Codes). *The class of codes with* $d_1(\mathbf{C}) = n - k$ *and* $d_i(\mathbf{C}) = n - k + i$*, for* $i = 2, 3, \ldots, k$*, are called* Near-MDS (NMDS) codes. *Equivalently, a code is NMDS if and only if* $d_1(\mathbf{C}) = n - k$ *and* $d_1(\mathbf{C}^{\perp}) = k$.

Proposition 3 (Properties of Near-MDS Matrices). *A linear* $[n, k]$ *code is Near-MDS if and only if its generator matrix satisfies the following conditions*

1. *Any* $k - 1$ *columns of the generator matrix are linearly independent.*
2. *Any* $k + 1$ *columns of the generator matrix are of rank* k.
3. *There exists a set of* k *linearly dependent columns in the generator matrix.*

2.2 Secret Sharing

A secret sharing scheme is a shared control scheme in which a *dealer* D splits a *secret* s into n *shares* $s_1, \ldots, s_n$ and distributes these to a set $\mathcal{P} = \{P_1, \ldots, P_n\}$ of n *participants* (or *users*) such that while certain *groups* of participants (subsets of $\mathcal{P}$) can reconstruct the secret from their shares (possibly by submitting their shares to a *combiner* C), others cannot.

Definition 14 (Access Structure). *An* access structure Λ *on a set of participants* $\mathcal{P}$ *is a subset of* $2^{\mathcal{P}}$ *such that sets in* Λ *can reconstruct the secret but sets not in* Λ *cannot. Elements of* Λ *are termed to be* authorized *sets while the sets not in* Λ *are called* unauthorized *sets.*

Definition 15 (Monotone Property). *A collection* $\Lambda \subseteq 2^{\mathcal{P}}$ *is called* monotone *if for all* $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{P}$*, if* $\mathcal{A} \in \Lambda$ *then* $\mathcal{B} \in \Lambda$*. In other words, if* $\mathcal{A}$ *is in the access structure* Λ *and* $\mathcal{B}$ *is a superset of* $\mathcal{A}$ *then* $\mathcal{B}$ *is also present in* Λ.

Definition 16 (Distribution Scheme). *A* distribution scheme*,* $\Pi_{\mathcal{P},\mathcal{S},\mathcal{R},\mu}$*, with a domain of secrets* $\mathcal{S}$*, a set of strings* $\mathcal{R}$ *and a probability distribution* μ *on* $\mathcal{R}$*, is a system which on input a secret* $s \in \mathcal{S}$*, chooses a random string* $r \xleftarrow{\mu} \mathcal{R}$ *sampled in agreement with* μ*, and computes a vector of shares* $\Pi_{\mathcal{P},\mathcal{S},\mathcal{R},\mu}(s) = (s_1, \ldots, s_n)$*, and communicates each share* s_j *to* P_j *via a secure channel.*

Definition 17 (Secret Sharing Scheme). *A* secret sharing scheme*,* $\Gamma_{\mathcal{P},\mathcal{S},\mathcal{R},\mu,\Lambda}$*, is a distribution scheme* $\Pi_{\mathcal{P},\mathcal{S},\mathcal{R},\mu}$ *along with a* reconstruction function*,* RECON*, realizing the access structure* Λ.

Remark 4. When any of $\mathcal{P}$, $\mathcal{S}$, $\mathcal{R}$, μ and Λ are clear from the context, we may not specify the respective subscripts in $\Pi_{\mathcal{P},\mathcal{S},\mathcal{R},\mu}$ and $\Gamma_{\mathcal{P},\mathcal{S},\mathcal{R},\mu,\Lambda}$.

Definition 18 (Threshold Secret Sharing Scheme). *If the access structure* $\Lambda \subseteq 2^{\mathcal{P}}$ *is defined by*

$$\mathcal{A} \in \Lambda \iff |\mathcal{A}| \geq t\,,$$

for some $t \in \{1, 2, \ldots, n\}$*, then we call the secret sharing scheme a* (t, n) threshold *secret sharing scheme. That is, in a* (t, n)*-threshold secret sharing scheme, any set of at least* t *participants should be able to retrieve the secret but any set of* $t - 1$ *or less participants must not be able to find the secret.*

Definition 19 (Generalized Secret Sharing Scheme). *If the access structure $\Lambda \subseteq 2^{\mathcal{P}}$ does not have restrictions such as in the case of threshold secret sharing schemes, then we call the secret sharing scheme a* generalized *secret sharing scheme.*

Definition 20 (Correctness). *A secret sharing scheme $\Gamma_{\mathcal{P},\mathcal{S},\mathcal{R},\mu,\Lambda}$ is said to be* correct *if the secret can be reconstructed by any authorized set of parties by pooling in their shares. That is, for the access structure Λ, for any set $\mathcal{A} \in \Lambda$, and for any key $s \in \mathcal{S}$, $\Pr[\mathsf{RECON}(\mathcal{A}) = s] = 1$.*

Definition 21 (Perfect Secret Sharing). *A secret sharing scheme is said to be* perfect *if, in information theoretic sense, an unauthorized set is unable to learn anything about the secret from their pool of shares. Formally, for every possible pair of secrets $a, b \in \mathcal{S}$, for any set $\mathcal{A} \notin \Lambda$ and a vector of shares $\langle s_j \rangle_{P_j \in \mathcal{A}}$,*

$$\Pr[\Pi(a, r)_{\mathcal{A}} = \langle s_j \rangle_{P_j \in \mathcal{A}}] = \Pr[\Pi(b, r)_{\mathcal{A}} = \langle s_j \rangle_{P_j \in \mathcal{A}}]$$

That is, the probability of finding a secret by an unauthorized set is equivalent to the probability of finding the secret randomly from the set of secrets $\mathcal{S}$.

Definition 22 (Information Rate). *The* information rate, *ρ, of a secret sharing scheme is the ratio between the length of the secret to that of the shares which will be allocated to the participants. That is, if $\mathcal{S}$ is the set of all possible secrets and $\mathcal{T}$ is the set of all possible shares, then*

$$\rho = \frac{\log |\mathcal{S}|}{\log |\mathcal{T}|}.$$

Definition 23 (Ideal Secret Sharing Scheme). *A secret sharing scheme is said to be* ideal *if the key space and the share space are same. More specifically, a scheme is considered to be ideal if information rate is equal to one, that is, the maximum possible value of ρ.*

Definition 24 (Linear Secret Sharing Scheme). *A secret sharing scheme Γ is called* linear *over $\mathrm{GF}(q)$, if there exists a matrix $G \in \mathrm{GF}(q)^{k \times n}$ and a vector $\mathbf{v} = (v_0, v_1, \ldots, v_{k-1}) \in \mathrm{GF}(q)^k$, such that $\mathbf{v} \cdot G = (s_0, s_1, \ldots, s_{n-1})$ gives the vector of shares.*

Definition 25 (Cheating Detection). *In the secret reconstruction phase of the scheme, when a participant or group of participants, $\mathcal{A}_c$, give wrong share(s), an authorized set might fail to retrieve the secret, or an unauthorized set might be able to find the secret. The security against such kind of attacks is known as* cheating detection.

Definition 26 (Cheater Identification). *If any participant P_i produces an incorrect share $s_i' \neq s_i$ during the reconstruction phase of the secret, then with the error probability of ϵ, P_i will be identified as a cheater and will be put in the set of cheaters $\mathcal{A}_c$. This is known as* cheater identification. *Cheater identification claims that the error probability ϵ is negligible.*

3 Related Work

In this section, we review the secret sharing scheme proposed by Zhou et al. [12] which is based on NMDS (Near-Maximum Distance Separable) codes. To the best of our knowledge, this is the first scheme based on NMDS codes. The scheme claims to be perfect and have an access structure such that at least k participants must come together to construct the secret. Unfortunately, the claims are not correct as we show in Sect. 4.

3.1 Share Construction

In this scheme, the dealer constructs the shares for each participant by selecting a generator matrix G of an $[n+1, k, n+1-k]$ NMDS code and a random vector $\mathbf{v} = (v_0, v_1, \ldots, v_{k-1})$ of length k. Using the vector, the dealer calculates the codeword $(s_0, s_1, \ldots, s_n)$ by multiplying $\mathbf{v}$ and G. That is,

$$(s_0, s_1, \ldots, s_n) = (v_0, v_1, \ldots, v_{k-1}) \cdot G\,.$$

Here s_0 is considered as the secret and s_i's are the shares of participants P_i's, where $1 \leq i \leq n$.

3.2 Secret Reconstruction

The reconstruction of the secret is based on the property of NMDS matrices that in a generator matrix of a NMDS code any $k+1$ columns have rank k.

Let $P_{j_0}, P_{j_1}, \ldots, P_{j_{m-1}}$, $m > k$, be the participants who come together to find the secret. They pool their shares forming the codeword $(s_{j_0}, s_{j_1}, \ldots, s_{j_{m-1}})$ and construct the corresponding submatrix

$$G' = (\mathbf{g}_{j_0}, \mathbf{g}_{j_1}, \ldots, \mathbf{g}_{j_{m-1}})$$

where $\mathbf{g}_{j_i}$ is the $(j_i + 1)$th column of the matrix G.

Since G is an NMDS matrix and $m \geq k+1$, the submatrix G' is of full rank, the system of linear equations

$$(s_{j_0}, s_{j_2}, \ldots, s_{j_{m-1}}) = (v_0, v_1, \ldots, v_{k-1}) \cdot G'$$

can be uniquely solved for $(v_0, v_1, \ldots, v_{k-1})$.

Then the secret s_0 can be calculated as $\mathbf{s}[0]$ where

$$\mathbf{s} = (v_0, v_1, \ldots, v_{k-1}) \cdot \mathbf{g}_{0G}$$

where $\mathbf{g}_{0G}$ is the first column of G.

Algorithm 1. Secret Reconstruction for the Secret Sharing Scheme in [12]

Input: Generator matrix G, number of participants m, collected set of shares

1: **if** $(m < k+1)$ **then return** "The secret cannot be recovered!" **end if**
2: $G' \leftarrow$ [columns of G corresponding to first $k+1$ shares available]
3: $G' \leftarrow$ [first k linearly independent column vectors of G']
4: $(s_{j_0}, s_{j_1}, \ldots, s_{j_{k-1}}) \leftarrow$ [shares corresponding to the columns of G']
5: $(v_0, v_1, \ldots, v_{k-1}) \leftarrow (s_{j_0}, s_{j_1}, \ldots, s_{j_{k-1}}) \cdot (G')^{-1}$
6: $\mathbf{s} \leftarrow (v_0, v_1, \ldots, v_{k-1}) \cdot G'$
7: **secret** $\leftarrow \mathbf{s}[0]$

Output: Secret **secret**

4 Attack on the Scheme in [12]

It has been stated in [12] that the scheme is ideal and perfect. This scheme is also claimed to have security characterization of cheating detection and cheating verification. However, we show that the scheme is not perfect and will show that there exist unauthorized sets of participants, that is, sets with k or less number of participants who are able to reconstruct the secret. Since the rank of the submatrix formed using any $k+1$ columns in an $[n+1, k, n+1-k]$ NMDS matrix is k, there will also exist k participants among the given $k+1$ participants, who can generate the secret independently. In fact, we show that there may exist sets of only $k-1$ participants who can generate the secret.

As an illustration of the flaw, we give an instantiation of the scheme formed using the $[12, 6, 6]$ NMDS matrix G in $\mathbb{F}_5$, given in [4].

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 4 & 2 & 0 & 3 & 1 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 3 & 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 2 & 4 & 4 & 3 & 3 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 4 & 1 & 2 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 4 & 2 & 4 & 2 \end{bmatrix}.$$

According to the given scheme, the minimum number of participants in an authorized set is 7 since $k+1 = 7$. The secret corresponds to the first column. If we denote the $(i+1)$th row by $\mathbf{r}_i$, then a codeword formed from a $[n+1, k, n+1-k]$ NMDS matrix is of the form $\alpha_0\mathbf{r}_0 + \alpha_1\mathbf{r}_1 + \cdots + \alpha_{k-1}\mathbf{r}_{k-1}$ where the secret is α_0. Therefore, the codeword $\mathbf{s}$ formed from the matrix G is

$$\begin{aligned} \mathbf{s} = (&\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \\ &\alpha_0 + 4\alpha_1 + \alpha_2 + 2\alpha_3 + 4\alpha_4, \\ &\alpha_0 + 2\alpha_1 + 3\alpha_2 + 4\alpha_3 + \alpha_4 + \alpha_5, \\ &\alpha_0 + \alpha_2 + 4\alpha_3 + 2\alpha_4 + 4\alpha_5, \\ &\alpha_0 + 3\alpha_1 + 3\alpha_3 + \alpha_4 + 2\alpha_5, \\ &\alpha_0 + \alpha_1 + 2\alpha_2 + 3\alpha_3 + 3\alpha_4 + 4\alpha_5, \\ &\alpha_0 + 2\alpha_1 + 2\alpha_2 + 2\alpha_3 + 2\alpha_4 + 2\alpha_5)\,. \end{aligned}$$

Share Reconstruction with only 6 *Participants:* $\{P_1, P_2, P_3, P_4, P_5, P_7\}$.

The row reduced form of the columns corresponding to these participants is

$$G' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

where the last column corresponds to the secret, that is the first column of G.

Note that, if $\mathbf{g}_i$ denotes the $(i+1)$th column of G', then $\mathbf{g}_0 = 3\mathbf{g}_1 + 2\mathbf{g}_2 + \mathbf{g}_3 + 4\mathbf{g}_4 + 4\mathbf{g}_5 + \mathbf{g}_7$. Therefore the shares corresponding to these columns also satisfy the same equation, that is, $\alpha_0 = 3\alpha_1 + 2\alpha_2 + \alpha_3 + 4\alpha_4 + 4\alpha_5 + \alpha_7$ where $\alpha_7 = \alpha_0 + 2\alpha_1 + 3\alpha_2 + 4\alpha_3 + \alpha_4 + \alpha_5$. Therefore, the secret α_0 can be recovered by the given set of 6 participants.

Share Reconstruction with only 5 Participants: $\{P_1, P_2, P_3, P_4, P_6\}$.

The row reduced form of the columns corresponding to these participants is

$$G' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

where the last column corresponds to the secret, that is the first column of G.

Note that, here $\mathbf{g}_0 = \mathbf{g}_1 + 4\mathbf{g}_2 + 3\mathbf{g}_3 + \mathbf{g}_4 + \mathbf{g}_6$. Therefore the shares corresponding to these columns also satisfy the same equation, that is, $\alpha_0 = \alpha_1 + 4\alpha_2 + 3\alpha_3 + \alpha_4 + \alpha_6$ where $\alpha_6 = \alpha_0 + 4\alpha_1 + \alpha_2 + 2\alpha_3 + 4\alpha_4$. Therefore, the secret α_0 can be recovered by the given set of 5 participants.

Finally, we observe that, for any secret sharing scheme constructed as in [12], the secret can also be constructed with k or, sometimes even, $k-1$ participants. Thus the mentioned scheme cannot be a perfect secret sharing scheme.

5 Proposed Secret Sharing Scheme

5.1 Motivation for the Scheme

It is observed that according to [12], the minimum number of participants required to find the secret is $k+1$. But in most of the cases, the secret can also be found with either k participants or with $k-1$ participants. This motivates us to make a more generalized access structure for secret sharing scheme based on the properties of Near-MDS matrices which is vaster and has a rich access structure.

5.2 Access Structure

The formation of the access structure is based on the design and properties of the near-MDS matrix [3,4] and is on similar lines as proposed in [12]. Let

$$G = \left[\ G[0]\ \ G[1]\ \ \ldots\ \ G[k-1]\ \ G[k]\ \ \ldots\ \ G[n]\ \right]$$

be a generator matrix of an $[n+1, k, n+1-k]$ NMDS code over $\mathbb{F}_q$ such that $\mathbf{g}_{iG} := \mathbf{g}_i := G[i] \in \mathbb{F}_q^k$, $0 \leq i \leq n$, is the ith column of G. We assume that G is in the standard form, that is, $G = \left[I_k \mid A_{k\times(n+1-k)}\right]$.

As noted in [3], since $G[0], G[1], \ldots, G[k-1]$ are linearly independent, for all i, $0 \leq i \leq n-k$, there exist $a_j \in \mathbb{F}_q$, such that

$$G[k+i] = \sum_{j=0}^{k-1} a_j G[j]\,.$$

The access structure for the scheme has the monotonicity property and can be defined with the help of two mutually exclusive sets, namely *Group I* and *Group II*. If G is the generator matrix for the scheme, with the first column of G corresponding to the secret and rest of the columns corresponding to the shares of the participants, then the groups can be defined as follows:

Group I consists of all $k-1$-tuples of participants whose corresponding columns in G, along with the first column, form k-linearly dependent columns.

Group II consists of all k-tuples of participants which aren't a superset of a $k-1$-tuple in *Group I*, and their corresponding columns in G are linearly independent.

Let us denote this access structure with Λ_0. Note that due to its monotonicity property, any k participants whose corresponding columns in G are linearly independent as well as any $k+1$ or more participants are authorized to recover the secret. Moreover, Λ_0 has two special groups, namely, *Group I* which needs just $k-1$ participants to generate the secret and *Group II* which need just k participants to generate the secret and no more. This scheme is not a threshold secret sharing scheme but has a more generalized access structure.

5.3 Share Construction

The codeword $(s_0, s_1, \ldots, s_n)$ is formed by multiplying G by a chosen vector $(\alpha_0, \alpha_1, \ldots, \alpha_{k-1})$ of length k.

$$(s_0, s_1, \ldots, s_n) = (\alpha_0, \alpha_1, \ldots, \alpha_{k-1}) \cdot G\,.$$

Here, s_0 forms the secret and rest of the s_i's corresponds to the shares of the participants. Let us denote this distribution algorithm by Π.

Algorithm 2. Pseudocode for Π

Input: Standard Near-MDS Generator matrix G, Secret s_0, Random field element generator $\mathcal{R}$

$\alpha[0] \leftarrow s_0$
for $(i \leftarrow 1; i < k; i \leftarrow i + 1)$ **do**
    $\alpha[i] \leftarrow \mathcal{R}$
end for
$\mathbf{s} \leftarrow \alpha \cdot G$

Output: Vector $\mathbf{s}$ containing the secret s_0 and the shares $s_1, \ldots, s_n$.

5.4 Secret Reconstruction

The reconstruction of the secret is similar to the way proposed in [12]. Given a set of m participants $\mathcal{A} = \{P_{j_0}, P_{j_1}, \ldots, P_{j_{m-1}}\} \in \Lambda_0$ and their pooled shares forming the *pooled codeword* $\mathsf{pcw} = (s_{j_0}, s_{j_1}, \ldots, s_{j_{m-1}})$, the secret can be reconstructed as follows:

1. Find the submatrix G' corresponding to the shares of the participants, such that $G' = (\mathbf{g}_{j_0}, \mathbf{g}_{j_1}, \ldots, \mathbf{g}_{j_{m-1}}, \mathbf{g}_{0G})$, where $\mathbf{g}_{0G}$ is the first and $\mathbf{g}_{j_i}$ is the $(j_i + 1)$th column of the matrix G.
2. Reduce G' using elementary row operations to make its k (or m, whichever is minimum) rows and columns, an identity matrix and get the modified column $\mathbf{g}'_{0G}$ corresponding to the secret.
3. If $m = k-1$, multiply the pooled codeword $\mathsf{pcw} = (s_{j_0}, s_{j_1}, \ldots, s_{j_{m-1}}, 0)$. Else, multiply its sub-codeword $(s_{j_0}, s_{j_1}, \ldots, s_{j_{k-1}})$ to $\mathbf{g}'_{0G}$ to obtain the secret. Here, $\mathbf{g}_i$'s corresponds to the k columns forming an identity matrix.

Algorithm 3. Pseudocode for RECON

Input: The pooled set of m shares from $\mathcal{A}$

if $(m < k - 1)$ **then return** "Unauthorized set!" **end if**
$G_{\mathcal{A}} \leftarrow$ [columns of G corresponding to the available shares]
$G' \leftarrow [\, G_{\mathcal{A}} \mid G[0] \,]$ (where $G[0]$ denotes the first column of G)
$G' \leftarrow$ reduced row echelon form of G'
if $(m = k - 1)$ **then**
    if $(\text{rank}(G') = k)$ **then return** "Unauthorized set!"
    else $\mathsf{pcw} \leftarrow$ [shares related to $k - 1$ columns of $G' \mid 0$] **end if**
else if $(m = k$ and $\text{rank}(G) = k - 1)$ **then return** "Unauthorized set!"
else $\mathsf{pcw} \leftarrow$ [shares corresponding to k columns of G' forming identity matrix]
end if
$\mathbf{s} \leftarrow \mathsf{pcw} \cdot G'[m-1]$
secret $\leftarrow \mathbf{s}[0]$

Output: Secret **secret**

6 Analysis of the Proposed Scheme

Lemma 1. *Given k linearly dependent columns in an $[n+1, k, n+1-k]$ NMDS matrix, each of the remaining $n+1-k$ columns will be linearly independent of them.*

Proof. From Property 3 of NMDS matrices, we know that any $k+1$ columns have rank k. Since the given k columns are linearly dependent, we can write

$$\mathbf{g}_j = \sum_{i=0, i\neq j}^{k-1} a_i \mathbf{g}_i, \text{ not all } a_i = 0 \text{ and } 0 \leq j \leq k-1\,.$$

Now, for a column $\mathbf{g}_l$ of the matrix, let's consider $\mathbf{g}_l$ along with given the k columns, $\mathbf{g}_0, \mathbf{g}_1, \ldots, \mathbf{g}_{k-1}$. Assuming that $\mathbf{g}_l$ is linearly dependent on given k columns, that is,

$$\mathbf{g}_l = \sum_{i=0}^{k-1} b_i \mathbf{g}_i, \text{ not all } b_i = 0\,.$$

Substituting the value of $\mathbf{g}_j$, we get

$$\mathbf{g}_l = \sum_{i=0, i\neq j}^{k-1} (a_i + b_i)\mathbf{g}_i, \text{ not all } a_i = 0, \text{ not all } b_i = 0 \text{ and } 0 \leq j \leq k-1\,.$$

Note that, the above equation makes $\mathbf{g}_l$ a linear combination of $k-1$ columns. Since both $\mathbf{g}_j$ and $\mathbf{g}_l$ can be expressed as a linear combination of rest of the $k-1$ columns, it makes the rank of these $k+1$ columns equal to $k-1$. But, since the rank of these $k+1$ columns formed needs to be k, our assumption is wrong and $\mathbf{g}_j$ is linearly independent from the given k columns.

Lemma 2. *In the secret sharing scheme Γ_0 on the access structure Λ_0, if a k-participant tuple $\mathcal{A}$ is a superset of a* Group I *tuple, then columns corresponding to $\mathcal{A}$ will have rank k.*

Proof. Note that, the $k-1$ columns corresponding to a *Group I* tuple along with the secret's column form k linearly dependent columns. Now, because of Lemma 1, any other column of the matrix will be linearly independent of these k columns and thus, linearly independent of the $k-1$ columns corresponding to rest $k-1$ participants in $\mathcal{A}$. Therefore, columns corresponding to the participants in $\mathcal{A}$ will have rank k.

Contrapositive: If a set of columns corresponding to k participant tuple $\mathcal{A}$ does not have rank k, then it cannot be a superset of any *Group I* tuple.

Proposition 4. *There exists an unauthorized tuple of $k-1$ participants.*

Proof. Let us take a tuple $\mathcal{A} = \{P_{j_1}, \ldots, P_{j_{k-1}}\} \in$ *Group I*. By definition, the column corresponding to the secret $\mathbf{g}_0$ is linearly dependent on the $k-1$ columns $\{\mathbf{g}_{j_1}, \mathbf{g}_{j_2}, \ldots, \mathbf{g}_{j_{k-1}}\}$ corresponding to the participants.

Now we will replace one of the participants P_{j_x} in $\mathcal{A}$ with the participant P_{j_y} from rest of the participants. As a result, from Lemma 1, the column $\mathbf{g}_{j_y}$ is linearly independent from the columns $\{\mathbf{g}_0, \mathbf{g}_{j_1}, \ldots, \mathbf{g}_{j_{x-1}}, \mathbf{g}_{j_{x+1}}, \ldots \mathbf{g}_{j_{k-1}}\}$ as well, therefore the secret's column $\mathbf{g}_0$ will also be linearly independent of the new $k-1$ columns, that is, $\{\mathbf{g}_{j_1}, \ldots, \mathbf{g}_{j_{x-1}}, \mathbf{g}_{j_{x+1}}, \mathbf{g}_{j_{k-1}}, \mathbf{g}_{j_y}\}$. Thus, we have constructed an unauthorized set $\mathcal{A}' = \{P_{j_1}, \ldots, P_{j_{x-1}}, P_{j_{x+1}}, \ldots P_{j_{k-1}}, P_{j_y}\}$ consisting of $k-1$ participants.

Proposition 5. *There exists an unauthorized tuple of k participants.*

Proof. From Lemma 1, we know that taking any k linearly dependent columns $\{\mathbf{g}_{j_1}, \mathbf{g}_{j_2}, \ldots, \mathbf{g}_{j_k}\}$, the secret's column $\mathbf{g}_0$ will be linearly independent from them. Thus, the k participants $\{P_{j_1}, \ldots, P_{j_k}\}$ form an unauthorized set.

Theorem 1. *The secret sharing scheme Γ_0 on the access structure Λ_0 is a linear secret sharing scheme.*

Proof. We know that the multiplication by a matrix is a linear operation. From the definition of linearity and construction of the scheme we can conclude that the proposed scheme is linear secret sharing scheme.

Theorem 2. *The secret sharing scheme Γ_0 on the access structure Λ_0 is correct, that is, every authorized set $\mathcal{A}$ in Λ_0 can correctly generate the secret.*

Proof. Let $s_1, \ldots, s_m$ be the shares of the participants in $\mathcal{A}$, and s_0 be the secret.

Case 1: $\mathcal{A}$ is from *Group I*: Note that, the column in G corresponding to the secret s_0 is linearly dependent to the columns corresponding to the participants in $\mathcal{A}$. Therefore, the algorithm of RECON can row reduce the columns to find the coefficients a_i's such that $s_0 = a_1 s_1 + a_2 s_2 + \ldots a_{k-1} s_{k-1}$ and find the secret s_0.

Case 2: $\mathcal{A}$ is from *Group II*: Since participants in $\mathcal{A}$ have linearly independent k columns in G, every other column including the secret's column will be linearly dependent on these k columns. Thus, RECON can row reduce the columns to recover the secret s_0.

Case 3: $\mathcal{A}$ forms a superset of a tuple in *Group I* or *Group I*: Note that if $\mathcal{A}$ is a superset of a tuple from *Group I*, then from Lemma 2, the participants in $\mathcal{A}$ have k linearly independent columns in G. Otherwise, if $\mathcal{A}$ is a superset of a tuple from *Group II*, then since the number of participants is greater than or equal to $k+1$, from Property 3 there exist k linearly independent columns in G corresponding to the participants. Therefore, in both the instances the algorithm RECON will row reduce the columns in a similar way as Case 2 and find the secret s_0. Hence, if $\mathcal{A}$ is an authorized set, then $\Pr[\mathsf{RECON}(\mathcal{A}) = s_0] = 1$ and from Definition 20, the secret sharing scheme $Gamma_0$ is correct.

Proposition 6. *The complexity of the scheme for setup and secret reconstruction phase is of* $\mathcal{O}(n^3)$.

Proof. Note that the algorithm of RECON requires the matrix to be reduced in a reduced row echelon form. This operation is the most complex section of the RECON. Since the number of participants is less than n, the algorithm requires row reduction of an $(k \times n)$ matrix. We know that reduced row echelon form of an $(k \times n)$ matrix requires $\mathcal{O}(k^2 n)$ operations. Since $k \leq n$, the complexity of the reconstruction algorithm RECON is $\mathcal{O}(n^3)$.

Theorem 3. *The secret sharing scheme* Γ_0 *on the access structure* Λ_0 *is ideal.*

Proof. Note that both the secret and the shares belong to $\mathrm{GF}(q)$. Therefore, the information rate ρ is

$$\rho = \frac{\log |\mathrm{GF}(q)|}{\log |\mathrm{GF}(q)|} = 1$$

Hence, from Definition 22 of the ideal secret sharing scheme, Γ_0 is ideal.

Theorem 4. *The secret sharing scheme* Γ_0 *on the access structure* Λ_0 *is a perfect secret sharing scheme.*

Proof. Let an unauthorized set $\mathcal{A}$ of m participants come together to construct the secret. Note that, since the secret $s_0 \in \mathrm{GF}(q)$, the probability of anyone randomly finding the secret is $1/q$.

Case 1: $m \leq k-2$: Note that columns in G corresponding to these m participants along with the secret's column form less than $k+1$ columns which are linearly independent because of Property 1. Therefore they cannot form the secret s_0 on their own, that is, $\mathsf{RECON}(\mathcal{A}) \neq s_0$. Thus they will need at least one more share to form an authorized set. If they forge one share, the probability of them finding the secret is equal to the probability of them forging a correct secret which is again $1/q$ since the shares also belong to $\mathrm{GF}(q)$. This makes the probability of $\mathcal{A}$ finding the secret greater than or equal to $1/q$.

Case 2: $m = k-1$: Note that since $\mathcal{A}$ is unauthorized, from Lemma 2 we can say that $\mathcal{A}$ does not belong to *Group I*. This implies that the secret's column is linearly independent from the corresponding columns in $\mathcal{A}$ and therefore the participants cannot form the secret s_0 with no additional information. They will need to either forge at least one more share or replace one pooled share with a forged share to form an authorized set. Therefore, the probability of $\mathcal{A}$ finding the secret follows from Case 1 and is at least $1/q$.

Case 3: $m = k$: Since $\mathcal{A}$ is an unauthorized tuple and thus not in *Group II* or a superset of a tuple in *Group I*, from the contrapositive of Lemma 2 we know that columns corresponding to the participants in $\mathcal{A}$ are linearly dependent and the secret's column is linearly independent of these columns. Therefore, in a similar way as Case 2, they also need to either forge one more share or replace one share of their own participant with a forged share to form an authorized set, and the probability of $\mathcal{A}$ finding the secret follows.

Note that, when giving a set of shares as an input to RECON, the probability of RECON generating some other secret $\overline{s_0}$ from the set of secrets $\mathcal{S}$ is $1/q$. That is so because as mentioned in [9], NMDS matrices have a high diffusion property. Therefore, whenever a vector $\mathbf{v} \in \mathrm{GF}(q)^k$ is multiplied to its submatrix, the output generated is uniformly distributed. Therefore,

$$\Pr[\mathsf{RECON}(\mathcal{A}) = s_0] = \Pr[\mathsf{RECON}(\mathcal{A}) = \overline{s_0}]\,.$$

Hence, from Definition 21, no unauthorized set can learn anything about the secret, and the secret sharing scheme Γ_0 is a perfect secret sharing scheme.

6.1 Cheating Detection and Cheating Identification

The proofs for the safety features of cheating detection and cheater identification for the secret sharing scheme Γ_0 are adopted from [12].

We require the following property of linear codes [6] in this work.

Lemma 3. *Let* $\mathbf{C}$ *be an* $[n, k, d]$ *linear code over* $\mathbb{F}_q$*. For any* i*,* $1 \leq i \leq n$*, let* $\mathbf{C}^\star$ *be the code formed by removing the* i*th coordinate from all codewords of* $\mathbf{C}$*.*

- *If* $d > 1$*,* $\mathbf{C}^\star$ *is an* $[n-1, k, d^\star]$ *code where* $d^\star = d - 1$ *if* $\mathbf{C}$ *has a minimum weight codeword with a nonzero* i*th coordinate and* $d^\star = d$ *otherwise.*
- *If* $d = 1$*,* $\mathbf{C}^\star$ *is an* $[n-1, k, 1]$ *code if* $\mathbf{C}$ *has no codeword of weight* 1 *whose nonzero entry is in coordinate* i*.*
- *Otherwise, if* $k > 1$*,* $\mathbf{C}^\star$ *is an* $[n-1, k-1, d^\star]$ *code with* $d^\star \geq 1$*.*

Remark 5. Note that, the minimum distance between two codewords in $\mathbf{C}^\star$ is at least $d - 1$.

Lemma 4. *Given an* $[n+1, k, n+1-k]$ *NMDS code with generator matrix* G*, if*

$$\mathbf{s} = (s_0, s_1, \ldots, s_n) = (\alpha_0, \alpha_1, \ldots, \alpha_{k-1}) \cdot G$$
$$\text{and } \mathbf{s}' = (s'_0, s'_1, \ldots, s'_n) = (\alpha'_0, \alpha'_1, \ldots, \alpha'_{k-1}) \cdot G$$

such that $(\alpha_0, \alpha_1, \ldots, \alpha_{k-1}) \neq (\alpha'_0, \alpha'_1, \ldots, \alpha'_{k-1})$*, then*

$$d((s_0, s_1, \ldots, s_n), (s'_0, s'_1, \ldots, s'_n)) \geq n + 1 - k\,.$$

Proof. Since $(\alpha_0, \alpha_1, \ldots, \alpha_{k-1})$ and $(\alpha'_0, \alpha'_1, \ldots, \alpha'_{k-1})$ form different codewords of the NMDS code, the hamming distance between $\mathbf{s}$ and $\mathbf{s}'$ would be greater than or equal to the minimum distance of the code, that is, $n + 1 - k$.

We prove the following results for the secret sharing scheme Γ_0 by applying the same method as in [10].

Theorem 5. *The secret sharing scheme* Γ_0 *on the access structure* Λ_0 *has the security characterization of cheating detection when the cheaters are less than* $m - k$ *where* m *is the number of active participants.*

Proof. Assume that $P_{j_1}, \ldots, P_{j_m}$ submit their modified shares $s^\star_{j_i} = s_{j_i} + \delta_i$, $\delta_i \in \mathrm{GF}(q)$, $1 \le i \le m$ to the recovery algorithm. Note that P_{j_i} is honest if and only if $\delta_i = 0$, otherwise he is a cheater. Now consider the $k \times m$ submatrix G' consisting m columns of G, indexed by $j_1, j_2, \ldots j_m$ and let

$$\mathbf{D} = \{(s_1, \ldots, s_m) \mid (s_1, \ldots, s_m) = (\alpha_0, \alpha_1, \ldots, \alpha_{k-1}) \cdot G',\ \alpha_i \in \mathrm{GF}(q)\}\,.$$

Let $\mathbf{s} = (s_{j_1}, \ldots, s_{j_m})$, $\mathbf{d} = (\delta_1, \ldots, \delta_m)$ and $\mathbf{s}^\star = \mathbf{s} + \mathbf{d}$. From Lemma 3, any two distinct shares in $\mathbf{D}$ have a Hamming distance of at least $m - k$. Hence, if $d(\mathbf{d}) = d(\mathbf{s}, \mathbf{s}^\star) < m - k$, then $\mathbf{s}^\star \in \mathbf{D}$ if and only if $\mathbf{s}^\star = \mathbf{s}$, that is, when $\mathbf{d} = 0$. Therefore cheating can be detected if the cheaters are less than $m - k$.

Theorem 6. *In the secret sharing scheme Γ_0 on the access structure Λ_0 cheaters can be identified when their number is less than $\lfloor \frac{m-k}{2} \rfloor$.*

Proof. Using the same notations as in the previous proof let $\mathbf{s} = (s_{j_1}, \ldots, s_{j_m})$, $\mathbf{d} = (\delta_1, \ldots, \delta_m)$ and $\mathbf{s}^\star = \mathbf{s} + \mathbf{d}$. Now, if $d(\mathbf{d}) < \lfloor \frac{m-k}{2} \rfloor$, then for $\mathbf{s}' \in \mathbf{D}$, we have the following relation when $\mathbf{s}' \neq \mathbf{s}$

$$d(\mathbf{s}^\star, \mathbf{s}') \ge d(\mathbf{s}^\star, \mathbf{s}) + d(\mathbf{s}, \mathbf{s}') \ge (m - k) - \left\lfloor \frac{m-k}{2} \right\rfloor = \left\lceil \frac{m-k}{2} \right\rceil .$$

Since $d(\mathbf{s}^\star, \mathbf{s}) = \lfloor \frac{m-k}{2} \rfloor$, we get $d(\mathbf{s}^\star, \mathbf{s}) = \min\{d(\mathbf{s}^\star, \mathbf{s}') \mid \mathbf{s}' \in \mathbf{D}\}$. By decoding $\mathbf{s}^\star$ to $\mathbf{s}$ using error decoding algorithms of linear codes and deriving $\mathbf{d} = \mathbf{s}^\star - \mathbf{s}$, we can say that the participant P_{j_i} is a cheater if $\delta_i \neq 0$. Therefore when the number of cheaters is less than $\lfloor \frac{m-k}{2} \rfloor$, the secret can be recovered successfully, and the cheaters can be identified.

Acknowledgements. This work has been partially supported by DST-FIST Level-1 Program, Grant No. SR/FST/MSI-092/2013. The authors would like to thank Department of Mathematics, BITS Goa, R. C. Bose Centre for Cryptology and Security, ISI Kolkata, and Indian Institute of Technology, Jammu, for their support.

A An Instantiation of the Proposed Scheme

Consider the following NMDS matrix G having elements over $\mathbb{F}_5$, as mentioned in [4].

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 4 & 2 & 0 & 3 & 1 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 3 & 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 2 & 4 & 4 & 3 & 3 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 4 & 1 & 2 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 4 & 2 & 4 & 2 \end{bmatrix} .$$

If we denote the i-th row by $\mathbf{r}_{i-1}$, and the chosen vector by $(\alpha_0, \alpha_1, \ldots, \alpha_{k-1})$, then the codeword formed is of the form $\alpha_0\mathbf{r}_0 + \alpha_1\mathbf{r}_1 + \cdots + \alpha_{k-1}\mathbf{r}_{k-1}$. Therefore, the codeword c formed from the matrix G is

$$\begin{aligned} c = (&\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \\ &(\alpha_0 + 4\alpha_1 + \alpha_2 + 2\alpha_3 + 4\alpha_4), \\ &(\alpha_0 + 2\alpha_1 + 3\alpha_2 + 4\alpha_3 + \alpha_4 + \alpha_5), \\ &(\alpha_0 + \alpha_2 + 4\alpha_3 + 2\alpha_4 + 4\alpha_5), \\ &(\alpha_0 + 3\alpha_1 + 3\alpha_3 + \alpha_4 + 2\alpha_5), \\ &(\alpha_0 + \alpha_1 + 2\alpha_2 + 3\alpha_3 + 3\alpha_4 + 4\alpha_5), \\ &(\alpha_0 + 2\alpha_1 + 2\alpha_2 + 2\alpha_3 + 2\alpha_4 + 2\alpha_5))\,. \end{aligned}$$

Hence, the first element of the codeword, that is, α_0 forms the secret while the rest of the elements become the shares for the participants.

A.1 Secret Reconstruction

Now any 5 participants from *Group I* or any 6 participants from *Group II* or more can find the secret.

1. 5 participants: P_1, P_2, P_3, P_4 and P_6.
The pooled codeword pcw is $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_0 + 4\alpha_1 + \alpha_2 + 2\alpha_3 + 4\alpha_4)$ and the corresponding submatrix G' is:

$$G' = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

After elementary row operations,

$$(G')' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \implies \mathbf{g}'_{0G} = \begin{bmatrix} 1 \\ 4 \\ 3 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Then

$$\begin{aligned} \mathsf{pcw} \cdot \mathbf{g}_{0G} &= \alpha_1 + 4\alpha_2 + 3\alpha_3 + \alpha_4 + (\alpha_0 + 4\alpha_1 + \alpha_2 + 2\alpha_3 + 4\alpha_4) \\ &= \alpha_0 + 5\alpha_1 + 5\alpha_2 + 5\alpha_3 + 5\alpha_4 \\ &= \alpha_0 \mod 5\,. \end{aligned}$$

Hence the secret $s_0 = \alpha_0$ is recovered correctly.

2. 6 participants: P_1, P_2, P_3, P_4, P_5 and P_7.
The pooled codeword pcw is $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_0+2\alpha_1+3\alpha_2+4\alpha_3+\alpha_4+\alpha_5)$ and the corresponding submatrix G' is:

$$G' = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

After elementary row operations:

$$(G')' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \implies \mathbf{g}'_{0G} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 4 \\ 4 \\ 1 \end{bmatrix}.$$

Then

$$\begin{aligned} \mathsf{pcw} \cdot \mathbf{g}_{0G} &= 3\alpha_1 + 2\alpha_2 + \alpha_3 + 4\alpha_4 + 4\alpha_5 + \alpha_0 + (2\alpha_1 + 3\alpha_2 + 4\alpha_3 + \alpha_4 + \alpha_5) \\ &= \alpha_0 + 5\alpha_1 + 5\alpha_2 + 5\alpha_3 + 5\alpha_4 + 5\alpha_5 \\ &= \alpha_0 \mod 5. \end{aligned}$$

Hence the secret $s_0 = \alpha_0$ is recovered correctly.

3. 7 or more participants: $P_4, P_5, P_6, P_7, P_8, P_9, P_{10}$ and P_{11}.
The pooled codeword is

$$\begin{aligned} \mathsf{pcw} = (&\alpha_4, \alpha_5, \\ &\alpha_0 + 4\alpha_1 + \alpha_2 + 2\alpha_3 + 4\alpha_4, \\ &\alpha_0 + 2\alpha_1 + 3\alpha_2 + 4\alpha_3 + \alpha_4 + \alpha_5, \\ &\alpha_0 + \alpha_2 + 4\alpha_3 + 2\alpha_4 + 4\alpha_5, \\ &\alpha_0 + 3\alpha_1 + 3\alpha_3 + \alpha_4 + 2\alpha_5, \\ &\alpha_0 + \alpha_1 + 2\alpha_2 + 3\alpha_3 + 3\alpha_4 + 4\alpha_5, \\ &\alpha_0 + 2\alpha_1 + 2\alpha_2 + 2\alpha_3 + 2\alpha_4 + 2\alpha_5) \end{aligned}$$

and the corresponding submatrix G' is:

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 4 & 2 & 0 & 3 & 1 & 2 & 0 \\ 0 & 0 & 1 & 3 & 1 & 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 4 & 4 & 3 & 3 & 2 & 0 \\ 1 & 0 & 4 & 1 & 2 & 1 & 3 & 2 & 0 \\ 0 & 1 & 0 & 1 & 4 & 2 & 4 & 2 & 0 \end{bmatrix}.$$

After elementary row operations:

$$(G')' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 3 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 4 & 1 & 4 \end{bmatrix} \implies \mathbf{g}'_{0G} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 4 \\ 3 \\ 4 \end{bmatrix}.$$

Then

$$\begin{aligned} \mathsf{pcw} \cdot \mathbf{g}_{0G} &= (\alpha_4) + (\alpha_5) + (4\alpha_0 + 3\alpha_1 + 2\alpha_2 + \alpha_3 + 4\alpha_4 + 4\alpha_5) \\ &\quad + (3\alpha_0 + 3\alpha_2 + 2\alpha_3 + \alpha_4 + 2\alpha_5) + (4\alpha_0 + 2\alpha_1 + 2\alpha_3 + 4\alpha_4 + 3\alpha_5) \\ &= 11\alpha_0 + 5\alpha_1 + 5\alpha_2 + 5\alpha_3 + 10\alpha_4 + 10\alpha_5 \\ &= \alpha_0 \mod 5\,. \end{aligned}$$

Hence the secret $s_0 = \alpha_0$ is recovered correctly.

Hence in every case, the secret s_0 is recovered correctly.

References

1. Benaloh, J., Leichter, J.: Generalized secret sharing and monotone functions. In: Goldwasser, S. (ed.) CRYPTO 1988. LNCS, vol. 403, pp. 27–35. Springer, New York (1990). https://doi.org/10.1007/0-387-34799-2_3
2. George Robert Blakley: Safeguarding cryptographic keys. In: AFIPS, pp. 313–317 (1979)
3. Dodunekov, S., Landgev, I.: On Near-MDS codes. J. Geom. **54**(1), 30–43 (1995)
4. Dodunekov, S.M., Landjev, I.N.: Near-MDS codes over some small fields. Discrete Math. **213**(1–3), 55–65 (2000)
5. Harn, L., Lin, C.: Detection and identification of cheaters in (t, n) secret sharing scheme. Des. Codes Crypt. **52**(1), 15–24 (2009)
6. Huffman, W.C., Pless, V.: Fundamentals of Error-Correcting Codes. Cambridge University Press, New York (2010)
7. Ito, M., Saito, A., Nishizeki, T.: Secret sharing scheme realizing general access structure. Electron. Commun. Jpn. (Part III: Fundam. Electron. Sci.) **72**(9), 56–64 (1989)
8. Karchmer, M., Wigderson, A.: On span programs. In: Structure in Complexity Theory Conference, pp. 102–111. IEEE Computer Society (1993)
9. Li, C., Wang, Q.: Design of lightweight linear diffusion layers from Near-MDS matrices. IACR Trans. Symmetric Cryptol. **2017**(1), 129–155 (2017)
10. Pieprzyk, J., Zhang, X.-M.: Ideal threshold schemes from MDS codes. In: Lee, P.J., Lim, C.H. (eds.) ICISC 2002. LNCS, vol. 2587, pp. 253–263. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-36552-4_18
11. Shamir, A.: How to share a secret. Commun. ACM **22**(11), 612–613 (1979)
12. Zhou, Y., Wang, F., Xin, Y., Luo, S., Qing, S., Yang, Y.: A secret sharing scheme based on Near-MDS codes. In: NIDC, pp. 833–836. IEEE (2009)