Check for updates

Generalized Secret Sharing Schemes Using N$^{\mu}$MDS Codes

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

[1] Goldman Sachs Services Pvt Ltd, Bangalore, India
sanyam.mehta12@gmail.com
[2] Robert Bosch Engineering & Business Solutions Pvt Ltd, Bangalore, India
vishal.saraswat@gmail.com

Abstract. Mehta et al. [11] recently proposed an NMDS code-based secret sharing scheme having a richer access structure than the traditional (t, n) threshold secret sharing schemes, and is based on two mutually nonmonotonic sets of user groups of sizes t and $t-1$ respectively, where $n \geq t > 1$ corresponds to the total number of users. We give a full generalization of their scheme with complete security proofs. We propose an efficient generalized secret sharing scheme constructed using N$^{\mu}$MDS codes with time complexity of $O(n^3)$. The scheme accepts an access structure constructed using $\mu + 1$ mutually nonmonotonic sets of user groups with sizes, $t, t-1, \ldots, t-\mu$, respectively, where $1 \leq \mu < t$, and the parameter t defines the threshold such that all user groups of size greater than t can recover the secret. The proposed secret sharing scheme is perfect and ideal and has robust cheating detection and cheater identification features.

Keywords: Secret sharing schemes · Generalized access structure · Near MDS codes · Almost MDS codes

1 Introduction

Secret sharing schemes allow a dealer, D, to split a secret s into n shares $s_1, \ldots, s_n$ and distribute these shares to a set $\mathcal{P}$ of n users, $P_1, \ldots, P_n$, according to an *access structure* $\Gamma \subset 2^{\mathcal{P}}$ such that a subset $\mathcal{A} \subseteq \mathcal{P}$ of users can form the secret using their shares if and only if $\mathcal{A} \in \Gamma$. Moreover the secret sharing scheme is called a (t, n) *threshold secret sharing scheme* if the access structure Γ is defined by

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

for some $t \in \{1, 2, \ldots, n\}$. Otherwise it is called a *generalized secret sharing scheme*.

Blakley [2] and Shamir [13] independently proposed secret sharing schemes in 1979. Shamir's scheme utilises the standard Lagrange interpolation and linear algebra whereas Blakley's scheme uses the concept of intersection of hyperplanes in finite geometries. Both of these schemes were threshold secret sharing schemes,

© Springer Nature Switzerland AG 2020
D. Slamanig et al. (Eds.): MACIS 2019, LNCS 11989, pp. 234–248, 2020.
https://doi.org/10.1007/978-3-030-43120-4_18

that is, they restricted users in such a way that if and only if the number of users exceeds the threshold, they could recover the secret. Ito et al. [8] introduced the notion of a secret sharing scheme with a generalized access structure. A generalized access structure consists of arbitrary subsets of users (irrespective of each subset's size), who could find the secret. They proposed a scheme in which the dealer assigned several copies of a (t, n)-threshold secret sharing scheme to every user. The dealer chooses two positive integers m and t and a prime power q satisfying $t \leq m < q$ and

- chooses $\alpha_{t-1} \in \mathrm{GF}(q) - \{0\}$ and $\alpha_1, \ldots, \alpha_{t-2}$ from $\mathrm{GF}(q)$ and computes $f(x) = s + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_{t-1} x^{t-1}$, where $\mathrm{GF}(q)$ is the Galois Field of order q and $f(0) = s \in \mathrm{GF}(q)$ is the secret;
- chooses $x_1, \ldots, x_m \in \mathrm{GF}(q) - \{0\}$ and computes $s_j = f(x_j)$ $(1 \leq j \leq m)$;
- and finally, assigns a subset $S_i \subset \{(x_1, s_1), \ldots, (x_m, s_m)\}$ to the user P_i, $1 \leq i \leq n$.

The access structure of this scheme contains all those sets for which the size of the union of the users' shares $\geq t$. In the worst case, the share size is exponential in the size of the set of users. Benaloh and Leichter [1] proposed a secret sharing scheme with a generalized access structure which was simpler than that of the Ito et al.'s scheme [8]. Their construction utilizes the monotonicity property inherent in secret sharing schemes. They create a composition of multiple schemes with simple access structures and realize all access structures which can be defined using a small monotone formula. Although this scheme is simpler and more efficient than Ito et al.'s scheme [8], the share length is still exponential in the number of users.

Considering the secret sharing scheme proposed by Shamir once again, note that although a cheating user can not recover the secret by providing an incorrect share, but by getting a wrong key, he can misguide the honest users. Various ways of detecting and correcting the secret have been suggested by scholars. Some consider that there are only t shareholders for secret recovery and to check that the shares are not fake, the dealer gives an additional information such as using some check vectors to which will act like some kind of certificate for each user. Others have suggested to use error correcting codes where fake shares can be assumed to be errors and corrected like error correction of codes. Most of the initial schemes had concerns over cheater detection and identification and use of trusted third parties (combiners and dealers). Lein et al. [6] proposed a modification of Shamir's scheme [13] which allowed for cheater detection and identification. If $m > t$ users come together, where t is the threshold, then there are $\binom{m}{t}$ ways for the users to pool their shares and for each such way, a $t-1$ degree recovery polynomial can be constructed through interpolation. The original polynomial can be then compared with the interpolated polynomial. Users who could not recover the original polynomial and are in the majority of groups are marked as possible cheaters and then the shares are corrected recursively until no cheater is left. This cheater detection and identification algorithm trades off space and time-complexities for secret recovery.

Researchers also observed that instead of using arbitrary matrices, using linear codes provided the following advantages

- A single generator matrix is sufficient to represent them.
- They enable easy transmission and easier error detection.
- Even though features for cheater detection, identification, and verification were added, schemes were still efficient.

McEliece and Sarwate [10] constructed a secret sharing scheme from Reed-Solomon codes and showed it to be essentially the same as the Shamir threshold scheme [13]. Later, Massey [9] gave a general construction of linear secret sharing schemes from linear codes (or linear matroids). Blakley and Kabatiansky [3] and Dijk [4] gave a generalization of Massey's scheme to multidimensional subspaces instead of vectors. Pieprzyk and Zhang [12] used Maximum Distance Separable (MDS) codes to construct a secret sharing scheme in which, an Maximum Distance Separable matrix G of dimension $(t \times n)$ along with a message vector $\mathbf{v}$ of dimension $1 \times t$ is chosen by the dealer. The dealer then finds the desired codeword by computing $\mathbf{v} \times G$. The secret is the first element of the codeword.

It was shown in [9] that the access structure of the resulting secret sharing schemes is determined by the minimal codewords in the dual code. However, determining the minimal codewords in a linear code and hence, the access structure, is hard. Dodunekov [5] proposed using NMDS codes instead of MDS codes to construct a secret sharing scheme while observing the following advantages:

- They are less space consuming and easier to implement.
- Their access structure is richer than MDS secret sharing.
- The generator matrix of the code is hard to identify by an adversary.
- Shares the same properties of cheating detection and cheater identification with MDS codes based schemes.

Mehta et al. [11] proposed an NMDS code-based secret sharing scheme having a richer access structure than the traditional (t, n) threshold secret sharing schemes and an access structure constructed using two mutually nonmonotonic sets of user groups having sizes, t and $t-1$ respectively, where n corresponds to the total number of users.

1.1 Our Contribution

We have proposed an efficient generalized secret sharing scheme based on N$^{\mu}$MDS codes. The use of the N$^{\mu}$MDS matrices allows us to have *authorized* sets of varying sizes thus allowing the scheme to have a generalized and richer access structure. The proposed secret sharing scheme is perfect and ideal and has robust cheating detection and cheater identification features. The time complexity for the share distribution and share recovery phases is just $O(n^3)$, where n is the order of users. The proposed scheme has a finer access structure and provides a direction towards a fully generalized secret sharing scheme. The scheme constructs the access structure using $\mu + 1$ mutually nonmonotonic sets of user

groups of sizes, $t, t-1, \ldots, t-\mu$, respectively, where $1 \leq \mu < t$, and the parameter t defines the threshold such that all user groups of size greater than t can recover the secret.

2 Preliminaries

We denote the Galois Field, GF(q), of order q where $q = p^m$ is a prime power by $\mathbb{F}_q$. For $a_i \in \mathbb{F}_q$, $1 \leq i \leq n$, $(a_1, \ldots, a_n)$ denotes a vector in $\mathbb{F}_q^n$. We will also use the same notation, $(a_1, \ldots, a_n)$, to denote to denote a $n \times 1$ matrix (column) over $\mathbb{F}_q$. On the other hand, $[\ a_1 \ \ a_2 \ \ \ldots \ \ a_{n-1} \ \ a_n\ ]$ denotes a $1 \times n$ matrix (row) over $\mathbb{F}_q$. For vectors $\mathbf{v}_i = (v_{i1}, \ldots, v_{it}) \in \mathbb{F}_q^t$, $1 \leq i \leq n$, $[\ \mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \ldots \ \ \mathbf{v}_{n-1} \ \ \mathbf{v}_n\ ]$ denotes the $t \times n$ matrix over $\mathbb{F}_q$ formed by considering $\mathbf{v}_i$ as columns. For a $t \times n$ matrix G over $\mathbb{F}_q$, the ith column of G is denoted $G[i] \in \mathbb{F}_q^t$, $0 \leq i \leq n$.

2.1 Coding Theory

Definition 1. *A non-empty subset* $\mathbf{C}$ *of* $\mathcal{A}^n$*, where* $\mathcal{A} = \{a_0, \ldots a_{q-1}\}$*, is called a* q-ary block code of length n *over* $\mathcal{A}$*, and a string in* $\mathbf{C}$ *is called a* codeword.

Definition 2. *The number of positions in which* x *and* y *differ is known as* Hamming distance $d(x, y)$ *between* x *and* y*. The* minimum distance *of a code* $\mathbf{C}$ *is defined as*

$$d(\mathbf{C}) = \min_{x \neq y \in \mathbf{C}} d(x, y)\,.$$

Definition 3. *A* linear code, $\mathbf{L}$, of length n *is a linear subspace of* $\mathbb{F}_q^n$*. If dimension of* $\mathbf{L}$ *is* t *then we call it an* $[n, t]$*-code (over* $\mathbb{F}_q$*). Further, if the minimum distance of* $\mathbf{L}$ *is* d *then we call it an* $[n, t, d]$*-code (over* $\mathbb{F}_q$*).*

Definition 4. *The set of non-zero coordinate positions of a codeword* $c \in \mathbf{C}$ *is called its* support, Supp(c)*. The* support *of a code* $\mathbf{C}$, Supp($\mathbf{C}$)*, is defined as*

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

Definition 5. *The* r^{th} generalized Hamming distance, $d_r(\mathbf{C})$*, is the cardinality of the minimum support of an* $[n, r]$*-subcode of* $[n, t]$*-code* $\mathbf{C}$*, where,* $1 \leq r \leq t$*.*

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

Remark 1. The Hamming Distance of $\mathbf{C}$ $d(\mathbf{C}) = d_1(\mathbf{C})$.

Definition 6. *For an* $[n, t, d]$*-code* $\mathbf{C}$*, the* Singleton bound *states that the parameters of* $\mathbf{C}$ *must satisfy*

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

In other words, $d \leq n - t + 1$.

Definition 7. *The* r^{th} generalized Singleton bound $d_r(\mathbf{C})$ *states that*

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

Definition 8. *A* maximum distance separable (MDS) *code is an* $[n, t]$*-linear code which achieves the Singleton bound, that is, it is an* $[n, t, n - t + 1]$*-code.*

Proposition 1. *For an* $[n, t, d]$ MDS *code* $\mathbf{L}$ *over* $\mathbb{F}_q$*, let* H *be any of its parity check matrix of* $\mathbf{L}$ *and let* $G = (I_t \mid A)$ *be any of its generator matrix in standard form (ref. Remark 2). Then*

1. *Any* $n - t$ *columns of* H *are linearly independent.*
2. *Any* t *columns of* G *are linearly independent.*
3. *Any square submatrix of* A *is non singular.*

Definition 9. *The class of* $[n, t]$*-codes with*

$$d_1(\mathbf{C}) = n - t$$

are called almost-MDS (AMDS) codes.

Definition 10. *The class of* $[n, t]$*-codes with*

$$\begin{aligned} d_1(\mathbf{C}) &= n - t, \\ \text{and } d_i(\mathbf{C}) &= n - t + i, \qquad \text{for } i = 2, 3, \ldots, t, \end{aligned}$$

are called near-MDS (NMDS) codes.

Definition 11. *The class of* $[n, t]$*-codes with*

$$\begin{aligned} d_i(\mathbf{C}) &= n - t + 2i - \mu - 1, \qquad \text{for } i = 1, 2, \ldots, \mu \\ \text{and } d_i(\mathbf{C}) &= n - t + i, \qquad \text{for } i = \mu + 1, \ldots, t, \end{aligned}$$

are called N$^\mu$MDS codes.

Remark 2. For the purposes of this work, we will assume that the generator matrices G are in their standard form, that is, $G = (I_t \mid A)$, where I_t is the identity matrix of size $t \times t$. Moreover, the MDS (or the N$^\mu$MDS) matrices correspond to the matrix A.

A detailed characterization of N$^\mu$MDS codes was provided in [14]. The relevant properties of N$^\mu$MDS matrices required for this paper are as follows.

Proposition 2 (Properties of N$^\mu$MDS **Codes).** *The matrix characterization of an* N$^\mu$MDS *code with a generator matrix* G *is as follows:*

1. *For all* $i = 1, 2, \ldots, \mu$,
   (i) for $i < l \leq min\{d_i - 1, t\}$*, every* $(l - 2i + 2 + \mu, l)$ *submatrix has rank* $\geq (l - i + 1)$.

(ii) there exists an l, $i < l \leq min\{d_i, t\}$, *and an* $(l - 2i + 1 + \mu, l)$ *submatrix with rank equal to* $(l - i)$.

2. *For all* $i = \mu + 1, \ldots, t$,
   (i) for $1 < l \leq min\{(n - t), (t - \mu)\}$, *every* $(l, l + \mu)$ *submatrix has rank* l.

Corollary 1 (Properties of N$^{\mu}$MDS **Matrices.)** *The standard generator matrix for an* $[n, t]$ N$^{\mu}$MDS *code has the following properties:*

1. *Any* $t - \mu + 2i$ *columns of the generator matrix have rank* $\geq t - \mu + i$, *where* $i = 0, 1, \ldots, \mu - 1$.
2. *There exists a set of* $t - \mu + 2i + 1$ *columns with rank* $t - \mu + i$, *for* $i = 0, 1, \ldots, \mu - 1$.
3. *Any* $t + \mu$ *columns of the generator matrix have rank* t *and are linearly independent.*

2.2 Secret Sharing

Let $\mathcal{P} = P_1, \ldots, P_n$ be a set of n *users*. We call a subset $\mathcal{A}$ of $\mathcal{P}$ a *group* of users.

Definition 12. *A collection* $\Gamma \subseteq 2^{\mathcal{P}}$ *is called* monotone *if* $\mathcal{A} \in \Gamma$ *and* $\mathcal{A} \subseteq \mathcal{B}$ *then* $\mathcal{B} \in \Gamma$.

Definition 13. *We call two collections (sets)* $\mathcal{G}^i, \mathcal{G}^j \subseteq 2^{\mathcal{P}}$ mutually nonmonotonic sets *if for all* $\mathcal{A} \in \mathcal{G}^i$, *there is no* $\mathcal{B} \in \mathcal{G}^j$, *such that* $\mathcal{B} \subset \mathcal{A}$ *and vice versa.*

Definition 14. $\Gamma \subseteq 2^{\mathcal{P}}$ *is called an* access structure *if it is a monotone collection such that only the subsets of users in* Γ *are* authorized *to recover the secret. Subsets not in* Γ *are termed to be* unauthorized sets.

Definition 15. *A* distribution scheme *is denoted by* Π *with* $\mathcal{S}$, *the domain of secrets, and* $\mathcal{R}$, *a set of strings. For a secret* $t \in \mathcal{S}$ *and a string* $r \in \mathcal{R}$ *sampled randomly observing* Δ, *where* Δ *is the probability distribution on* $\mathcal{R}$, *a share vector* $\Pi(t, r) = (s_1, s_2, \ldots, s_j)$ *is computed and each share* s_j *is communicated to* P_j *via a secure channel.*

Definition 16. *A distribution scheme along with domain of secrets* $\mathcal{S}$ *realizing access structure* Γ *is called a* secret sharing scheme $\Sigma = \langle \Pi, \Delta \rangle$.

Definition 17. *A secret sharing scheme is* correct *if an authorized subset of users can always recover the secret. In other words, for any set* $\mathcal{A} \in \Gamma$, *there exists a recovery function or algorithm* SRA *such that for a key* $k \in \mathcal{S}$,

$$\Pr[\mathsf{SRA}(\mathcal{A}) \text{ is } k] = 1\,.$$

Definition 18. *If* $\mathcal{T}$ *is the set of all possible shares and* $\mathcal{S}$ *is the set of all possible secrets, then the* information rate ρ *of the secret sharing scheme is defined to be*

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

Definition 19. *A secret sharing scheme is* ideal *if the set of all secrets, $\mathcal{S}$, and the set of all shares, $\mathcal{T}$, are of same cardinality. That is, a secret sharing scheme is* ideal *if its information rate is one.*

Definition 20. *A secret sharing scheme is* perfect *if an unauthorized group of users, $\mathcal{C}$, cannot obtain any information about the secret from their pool of shares. That is, the probability of $\mathcal{C}$ recovering the secret using their pool of shares is equivalent to the probability of recovering the secret without using their pool of shares. In other words, for any subset $\mathcal{B} \notin \Gamma$, two secrets b and $c \in \mathcal{S}$ and every possible share vector $\langle s_j \rangle_{P_j \in \mathcal{B}}$,*

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

Definition 21. *A secret sharing scheme Σ is said to be* linear *over $\mathbb{F}_q$ if there exists a vector $\mathbf{v} = (v_0, v_1, \ldots, v_{t-1}) \in \mathbb{F}_q^t$ and a matrix $A \in \mathbb{F}_q^{t \times n}$, such that $\mathbf{v} \times A = (s_0, s_1, \ldots, s_{n-1})$ where s_0 is the secret and $(s_1, \ldots, s_{n-1})$ is the share vector.*

Definition 22. *During the secret recovery phase of a secret sharing scheme by an authorized subset of users $\mathcal{A}_c$, if a user P_i provides a wrong share, $\hat{s}_i$, instead of the correct one, s_i, it was assigned by the dealer during the share distribution phase, then the subset may fail to recover the secret, or worse, recover a wrong secret. Such a user is called a* cheater *and detection of occurrence of such an attack is called* cheating detection.

Definition 23. *Identification, with negligible error probability ϵ, of the user(s) providing wrong inputs while recovering the secret is called* cheater identification.

3 Proposed Secret Sharing Scheme

Though the scheme proposed in [11] has a richer access structure than the traditional (t, n) threshold secret sharing schemes, it only allows an access structure consisting of two mutually nonmonotonic sets of user groups of sizes, t and $t-1$, respectively. We propose a secret sharing scheme which admits a finer access structure based on $\mu + 1$, $1 \leq \mu \leq n - t$, mutually nonmonotonic sets of user groups of sizes, $t - \mu + 1 + i$, $1 \leq i \leq \mu + 1$, respectively. The proposed scheme is based on the properties of N$^\mu$MDS matrices which allow us to have an access structure which is richer and independent of the field size.

3.1 Access Structure

The access structure of the proposed secret sharing scheme is definded using the properties of N$^\mu$MDS matrices [14] and is a generalization of the one proposed in [11]. Let

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

be a standard generator matrix of an $[n+1, t, n-t-\mu+2]$ N$^{\mu}$MDS code over $\mathbb{F}_q$ where $G[i] \in \mathbb{F}_q^t$, $0 \leq i \leq n$.

Given a set $\mathcal{P}$ of n users, $P_1, \ldots, P_n$, we say that the column $G[i]$ corresponds to the user P_i and we define an *access structure* $\Gamma_\mu \subset 2^{\mathcal{P}}$ consisting of $\mu+1$ mutually nonmonotonic sets, namely, $\mathcal{G}^0$, $\mathcal{G}^1$, $\ldots$, $\mathcal{G}^\mu$ defined as as follows:

1. $\mathcal{G}^i$, $i < \mu$, consists of all $(t-\mu+i)$ users whose corresponding columns in G, along with the first column, form $t-\mu+i+1$ linearly dependent columns, and for all $\mathcal{A} \in \mathcal{G}^i$, there is no $\mathcal{B} \in \mathcal{G}^j, j < i$, such that $\mathcal{B} \subset \mathcal{A}$.
2. $\mathcal{G}^\mu$ consists of all (t) users whose corresponding columns in G are linearly independent, and for all $\mathcal{A} \in \mathcal{G}^\mu$, there is no $\mathcal{B} \in \mathcal{G}^j, j < \mu$, such that $\mathcal{B} \subset \mathcal{A}$.

Note that the access structure Γ_μ as defined above is a generalized access structure and satisfies the monotonicity property. Thus, the secret sharing scheme based on Γ_μ is a generalized secret sharing scheme.

3.2 Share Construction

To compute the n shares of a given secret $s_0 \in \mathbb{F}_q$, the dealer chooses $t-1$ random elements $\alpha_1, \ldots, \alpha_{t-1}$ from $\mathbb{F}_q$ and computes the codeword $(s_0, s_1, \ldots, s_n)$ by multiplying the generator matrix G by the t-length vector $(s_0, \alpha_1, \ldots, \alpha_{t-1})$. That is,

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

The elements $s_i \in \mathbb{F}_q$, $1 \leq i \leq n$, are the shares of the users $P_1, \ldots, P_n$ respectively. We say that the first column of G, $G[0]$, corresponds to the secret s_0 and the remaining columns $G[i]$, $1 \leq i \leq n$, correspond to the shares s_i of the users P_i.

3.3 Secret Recovery

The secret recovery algorithm SRA_μ is similar to the method proposed in [11] with modifications in the algorithm to allow for recovery of secret by user subsets of various sizes. Given a set of m users $\mathcal{B} = \{P_{j_1}, \ldots, P_{j_m}\} \in \Gamma_\mu$ and their respective shares $\{s_{j_1}, \ldots, s_{j_m}\}$, SRA_μ computes the secret as follows:

1. Construct the matrix

$$G' = \left[\; G[j_1] \;\; \ldots \;\; G[j_m] \;\; G[0] \;\right]$$

   formed by the columns which correspond to the shares of the users and the column which corresponds to the secret.
2. Row-reduce the matrix G' to make its first m (or t, whichever is minimum) rows and columns an identity matrix and denote the last column of this row-reduced matrix G' by $G[0]'$.
3. If $m < t$, add $t-m$ zeros to construct the pooled codeword

$$\mathsf{pool} = (s_{t_0}, s_{t_1}, \ldots, s_{t_{m-1}}, 0, \ldots, 0)$$

   and multiply pool to $G[0]'$ to obtain the secret.

4. Else multiply its sub-codeword $(s_{t_0}, s_{t_1}, \ldots, s_{t_{t-1}})$ to $G[0]'$ to obtain the secret.

Here, t_i's correspond to the t (or m) columns forming an identity matrix.

4 Analysis of the Proposed Scheme

Lemma 1. *For any $(t-\mu+2i+1)$ linearly dependent columns of an $[n, t, n-t-\mu+1]$ N$^\mu$MDS matrix, G, with rank $(t-\mu+i)$ where $0 \leq i \leq \mu-1$, each of the remaining $n-(t-\mu+2i+1)$ columns is linearly independent of them.*

Proof. Without loss of generality, suppose the given $(t-\mu+2i+1)$ linearly dependent columns with rank $(t-\mu+1)$ are $G[0], G[1], \ldots, G[t-\mu+2i]$ and let $0 \leq j \leq (t-\mu+2i)$ be such that

$$G[j] = \sum_{i=0, i\neq j}^{t-\mu+2i} a_i G[i], \text{ not all } a_i = 0\,.$$

Now, let $G[\ell]$ be a column from the remaining $n-(t-\mu+2i+1)$ columns of the matrix which is linearly dependent on the given $(t-\mu+2i+1)$ columns. That is,

$$G[\ell] = \sum_{i=0}^{t-\mu+2i} b_i G[i], \text{ not all } b_i = 0\,.$$

Substituting the value of $G[j]$, we get

$$G[\ell] = \sum_{i=0, i\neq j}^{t-\mu+2i} (a_i b_j + b_i) G[i],$$

where $0 \leq j \leq t-\mu+2i$ and not all $a_i = 0$ and not all $b_i = 0$. Hence $G[\ell]$ is a linear combination of the remaining $(t-\mu+2i)$ columns $G[i]$ $(0 \leq i \leq t-\mu+2i, i \neq j)$.

Since both the columns $G[j]$ and $G[\ell]$ are a linear combination of remaining the $(t-\mu+2i)$ columns, it makes the rank of these $(t-\mu+2i+2)$ columns less than or equal to $(t-\mu+i)$. But, from Property 1 of N$^\mu$MDS codes, any $(t-\mu+2i+2)$ columns have rank $\geq (t-\mu+i+1)$. Thus, our hypothesis is wrong and $G[\ell]$ must be linearly independent of the given $(t-\mu+2i+1)$ columns.

Proposition 3. *There exists a group of $(t-\mu+2i+1)$ users, $0 \leq i \leq \mu-1$ which is unauthorized.*

Proof. By Lemma 1, for any $(t-\mu+2i+1)$ linearly dependent columns

$$\{G[j_1], G[j_2], \ldots, G[j_{t-\mu+2i+1}]\}$$

with rank $(t-\mu+i)$, the column $G[0]$ is linearly independent of them. Thus the secret s_0 cannot be recovered using just the shares

$$\{s_{j_1}, s_{j_2}, \ldots, s_{j_{t-\mu+2i+1}}\}.$$

Hence the users

$$\{P_{j_1}, \ldots, P_{j_{t-\mu+2i+1}}\}$$

form an unauthorized set.

Proposition 4. *There exists a group of* $(t-\mu+2i)$ *users,* $0 \leq i \leq \mu-1$ *which is unauthorized.*

Proof. If we take all columns except $G[j_\ell]$, $(0 \leq \ell \leq (t-\mu+2i+1))$, from the previous construction, we will get $(t-\mu+2i)$ linearly dependent columns

$$\{G[j_1], \ldots, G[j_{\ell-1}], G[j_{\ell+1}], \ldots, G[j_{(t-\mu+2i+1)}]\}$$

with rank $(t-\mu+i)$, with the secret's column $G[0]$ being linearly independent from these $(t-\mu+2i)$ columns. Thus, the $(t-\mu+2i)$ users

$$\{P_{j_1}, \ldots, P_{j_{\ell-1}}, P_{j_{\ell+1}}, \ldots, P_{j_{(t-\mu+2i+1)}}\}$$

form an unauthorized set.

Theorem 1. *The proposed secret sharing scheme* Σ_μ *is correct.*

Proof. Let $\mathcal{B} \in \Gamma_\mu$. Then $\mathcal{B}$ is an authorized set and we show that $\mathcal{B}$ can correctly recover the secret. Let $s_{j_1}, \ldots, s_{j_m}$ be the shares of the users in $\mathcal{B}$, and s_0 be the secret.

Case 1: $\mathcal{B}$ **is from** $\mathcal{G}^i, i < \mu$**:** Note that, the column $G[0]$ which corresponds to the secret s_0 is linearly dependent on the columns which correspond to the users in $\mathcal{B}$. Therefore, the algorithm SRA_μ can find the coefficients a_i's (by row-reducing the matrix formed by these columns and the column $G[0]$) such that

$$s_0 = a_1 s_{j_1} + a_2 s_{j_2} + \ldots a_{t-\mu+i} s_{j_{t-\mu+i}}$$

and find the secret s_0.

Case 2: $\mathcal{B}$ **is from** $\mathcal{G}^\mu$**:** Since columns which correspond to the users in $\mathcal{B}$ are t linearly independent columns of G, any other column of G, including the column $G[0]$, must be linearly dependent on them. Thus, the algorithm SRA_μ can find the coefficients a_i's (by row-reducing the matrix formed by these columns and the column $G[0]$) such that

$$s_0 = a_1 s_{j_1} + a_2 s_{j_2} + \ldots a_t s_{j_t}$$

and find the secret s_0.

***Case 3:* $\mathcal{B}$ is a superset of a group in $\mathcal{G}^i$ or $\mathcal{G}^\mu$:** If $\mathcal{B}$ is a superset of a group in $\mathcal{G}^i$, the users in $\mathcal{B}$ have at least $t-\mu+i$ linearly independent columns in G with the column $G[0]$ being linearly dependent on them by definition of $\mathcal{G}^i$. Therefore the algorithm SRA_μ, as in Case 1, can find the secret s_0. Otherwise, if $\mathcal{B}$ is a superset of a group in $\mathcal{G}^\mu$, then we already have t linearly independent columns in G which correspond to the group in $\mathcal{G}^\mu$ and the algorithm SRA_μ, as in Case 2, can find the secret s_0.

Hence, if $\mathcal{B}$ is an authorized set, then $\Pr[\mathsf{SRA}_\mu(\mathcal{B}) = s_0] = 1$ and hence the secret sharing scheme Σ_μ is correct.

Theorem 2. *The proposed secret sharing scheme Σ_μ has perfect privacy.*

Proof. Let $\mathcal{B}$ be an unauthorized set of m users which try to recover the secret. Since the secret $s_0 \xleftarrow{\$} \mathbb{F}_q$, the probability of randomly guessing the secret is $1/q$. Also, since N$^\mu$MDS matrices have a high diffusion property, whenever a vector $\mathbf{v} \in \mathbb{F}_q^t$ is multiplied to its submatrix formed by its m columns, the output generated is uniformly distributed in $\mathbb{F}_q^m$. Hence, for any share s_i, $1 \leq i \leq n$, the probability of randomly guessing s_i is $1/q$.

Case 1: $m \leq t-\mu-1$: Note that, by Property 1 of N$^\mu$MDS matrices, the $m+1 \leq t-\mu$ columns in G which correspond to these m users along with the column $G[0]$ are linearly independent. Therefore the column $G[0]$ cannot be obtained as a linear combination of m columns which correspond to these users, that is, $\mathsf{SRA}_\mu(\mathcal{B}) \neq s_0$. Thus $\mathcal{B}$ will require at least one more correct share to compute the secret. But the probability of $\mathcal{B}$ guessing the correct secret (or another correct share) is $1/q$. Thus the probability of $\mathcal{B}$ obtaining the secret is less than or equal to $1/q$.

Case 2: $m = t-\mu+i, 0 \leq i < \mu$: Since $\mathcal{B}$ is unauthorized, it neither belongs in $\mathcal{G}^i$ nor is a superset of a group in $\mathcal{G}^j, j < i$. This implies that the column $G[0]$ is linearly independent of the columns which correspond to the users in $\mathcal{B}$. Therefore the column $G[0]$ cannot be obtained as a linear combination of m columns which correspond to these users, that is, $\mathsf{SRA}_\mu(\mathcal{B}) \neq s_0$. Thus $\mathcal{B}$ will require at least one more correct share, or replace one of the pooled shares with a forged share, to compute the secret. But the probability of $\mathcal{B}$ guessing the correct secret (or another correct share) is $1/q$. Thus the probability of $\mathcal{B}$ obtaining the secret is less than or equal to $1/q$.

Case 3: $m = t+i, 0 \leq i < \mu$: Since $\mathcal{B}$ is unauthorized, it neither belongs in $\mathcal{G}^\mu$ nor is a superset of a group in $\mathcal{G}^j, j \leq \mu$. This implies that the columns which correspond to $\mathcal{B}$ are linearly dependent and the column $G[0]$ is independent of them (rendering any subset of $\mathcal{B}$ not a part of $\mathcal{G}^j$). Therefore the column $G[0]$ cannot be obtained as a linear combination of m columns which correspond to these users, that is, $\mathsf{SRA}_\mu(\mathcal{B}) \neq s_0$. Thus $\mathcal{B}$ will require at least one more correct share, or replace one of the pooled shares with a forged share, to compute the secret. But the probability of $\mathcal{B}$ guessing the correct secret (or another correct share) is $1/q$. Thus the probability of $\mathcal{B}$ obtaining the secret is less than or equal to $1/q$.

Note that, on an input of a random set of shares to SRA_{μ}, the probability of SRA_{μ} generating the correct secret s_0 is $1/q$. Therefore,

$$\Pr[\mathsf{SRA}_{\mu}(\mathcal{B}) = s_0] = \Pr[\mathsf{SRA}_{\mu}(\mathcal{B}) = \overline{s_0}]$$

and hence Σ_{μ} has perfect privacy.

Theorem 3. *The proposed secret sharing scheme Σ_{μ} is ideal.*

Proof. Since both the secret and the shares are elements of $\mathbb{F}_q$, the information rate ρ is

$$\rho = \frac{\log \mid \mathbb{F}_q \mid}{\log \mid \mathbb{F}_q \mid} = 1$$

and hence Σ_{μ} is ideal.

Theorem 4. *The proposed secret sharing scheme Σ_{μ} is a linear secret sharing scheme.*

Proof. By Definition 21 of a linear secret sharing scheme, and by the construction of the shares as in Subsect. 3.2, it is clear that the proposed secret sharing scheme is linear.

Proposition 5. *The time-complexity for the share construction and the secret recovery phase of the proposed scheme is $\mathcal{O}(n^3)$.*

Proof. That the complexity of the setup phase is $\mathcal{O}(n^3)$ is straight forward. We show that the complexity of the secret reconstruction phase is $\mathcal{O}(n^3)$.

The Step 2 of Algorithm SRA_{μ} computes the reduced row echelon form of the matrix G' constructed in Step 1. Since $m \leq n$, G' is at most a $(t \times n)$ matrix. Since row reduction of a $(t \times n)$ matrix can be done in $\mathcal{O}(t^2 n)$ operations and since $t \leq n$, the complexity of this step is $\mathcal{O}(n^3)$. That is the most complex step of the code because the remaining steps are linear in the size of the matrix. Hence, the complexity of the reconstruction phase is $\mathcal{O}(n^3)$.

4.1 Cheating Detection and Cheating Identification

The proofs in this section Σ_{μ} have been adapted from [11]. The following two lemmas, Lemmas 2 and 3, state standard properties of linear codes which we will use in this section. We refer the reader to [7] for the proof of Lemma 3.

Lemma 2. *Given an $[n, t, n-t-\mu+1]$ N$^{\mu}$MDS code and its generator matrix G, if*

$$(s_0, s_1, \ldots, s_{n-1}) = (\alpha_0, \alpha_1, \ldots, \alpha_{t-1}) \cdot G$$

and

$$(\hat{s}_0, \hat{s}_1, \ldots, \hat{s}_{n-1}) = (\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_{t-1}) \cdot G$$

such that

$$(\alpha_0, \alpha_1, \ldots, \alpha_{t-1}) \neq (\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_{t-1}),$$

then

$$d((s_0, s_1, \ldots, s_{n-1}), (\hat{s}_0, \hat{s}_1, \ldots, \hat{s}_{n-1})) \geq n - t - \mu + 1 .$$

Proof. Since $(\alpha_0, \alpha_1, \ldots, \alpha_{t-1})$ and $(\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_{t-1})$ are distinct, they generate different codewords of the N$^\mu$MDS code. Hence, they generate different codewords $(s_0, s_1, \ldots, s_{n-1})$ and $(\hat{s}_0, \hat{s}_1, \ldots, \hat{s}_{n-1})$ are distinct. Thus, the Hamming distance between them must be greater than or equal to $n - t - \mu + 1$, the minimum distance of the code.

Lemma 3. *Let* $\mathbf{C}$ *be an* $[n, t, d]$ *linear code over* $\mathrm{GF}(q)$. *Let* $\mathbf{C}^i$ *be the* punctured *code defined by dropping the* i^{th} *coordinate,* $1 \leq i \leq n$, *from the codewords of* $\mathbf{C}$. *Then,* $\mathbf{C}^i$ *is an* $[n-1, \tilde{t}, \tilde{d}]$ *code where*

- $\tilde{t} = t$ *and* $\tilde{d} = d$ *if* $\mathbf{C}$ *does not have any codeword of weight* d *with a nonzero* i^{th} *coordinate;*
- $\tilde{t} = t$ *and* $\tilde{d} = d - 1$ *if* $d > 1$ *and* $\mathbf{C}$ *has a codeword of weight* d *with a nonzero* i^{th} *coordinate;*
- $\tilde{t} = t - 1$ *and* $\tilde{d} \geq d$ *if* $d = 1$, $t > 1$ *and* $\mathbf{C}$ *has a codeword of weight* d *with a nonzero* i^{th} *coordinate.*

Theorem 5. *The proposed scheme allows cheating detection if the number of cheaters in a group* m *users is less than* $m - t - 1$.

Proof. Suppose $P_{j_1}, \ldots, P_{j_m}$ submit the shares $\hat{s}_{j_1} = s_{j_1} + \delta_1, \ldots, \hat{s}_{j_m} = s_{j_m} + \delta_m$, $\delta_j \in GF(q)$, to the reconstruction algorithm. Then if $\delta_i = 0$, P_{j_i} is honest, and if $\delta_i \neq 0$, P_{j_i} is a cheater. Let G' be the $t \times m$ submatrix formed by the m columns of G indexed by $j_1, j_2, \ldots j_m$. Let

$$H_0 = \{(s_1, \ldots, s_m) \mid (s_1, \ldots, s_m) = (\alpha_0, \alpha_1, \ldots, \alpha_{t-1}) \cdot G',\ \alpha_i \in GF(q)\}\,.$$

Let $\mathbf{s} = (s_{j_1}, \ldots, s_{j_m})$, $\delta = (\delta_1, \ldots, \delta_m)$ and $\hat{\mathbf{s}} = \mathbf{s} + \delta = (\hat{s}_{j_1}, \ldots, \hat{s}_{j_m})$.

By Lemma 3, any two distinct codewords in H_0 have a Hamming distance of at least $m - t - 1$. Now, if the Hamming weight of δ is less than $m - t - 1$, then the Hamming distance between $\hat{\mathbf{s}}$ and $\mathbf{s}$ is less than $m - t - 1$. Thus by Lemma 2, $\hat{\mathbf{s}} \in H_0$ if and only if $\hat{\mathbf{s}} = \mathbf{s}$, that is, when $\delta = 0$. Hence, if the number of cheating users is less than $m - t - 1$, cheating by them can be detected.

Theorem 6. *The proposed scheme allows cheater identification if the number of cheaters in a group* m *users is less than* $\lfloor \frac{m-t-1}{2} \rfloor$.

Proof. Let P_{j_i}, $1 \leq i \leq m$, G', H_0, $\mathbf{s}$, δ and $\hat{\mathbf{s}}$ be as in Theorem 5. Let the Hamming weight of δ is less than $\lfloor \frac{m-t-1}{2} \rfloor$. Then the Hamming distance $d(\hat{\mathbf{s}}, \mathbf{s})$ is less than $\lfloor \frac{m-t-1}{2} \rfloor$. For any $\tilde{\mathbf{s}} \neq \mathbf{s} \in H_0$, by Lemma 3, $d(\mathbf{s}, \tilde{\mathbf{s}}) \geq m - t - 1$. Hence using the triangle inequality, we get

$$\begin{aligned} d(\hat{\mathbf{s}}, \tilde{\mathbf{s}}) &\geq d(\mathbf{s}, \tilde{\mathbf{s}}) - d(\hat{\mathbf{s}}, \mathbf{s}) \\ &\geq (m - t - 1) - \left\lfloor \frac{m-t-1}{2} \right\rfloor = \left\lceil \frac{m-t-1}{2} \right\rceil \geq \left\lfloor \frac{m-t-1}{2} \right\rfloor = d(\hat{\mathbf{s}}, \mathbf{s})\,. \end{aligned}$$

Hence, $d(\hat{\mathbf{s}}, \mathbf{s}) = \min\{d(\hat{\mathbf{s}}, \tilde{\mathbf{s}}) \mid \tilde{\mathbf{s}} \in H_0\}$. Thus standard error decoding techniques for linear codes can be used to decode $\hat{\mathbf{s}}$ to recover the secret $\mathbf{s}$. Then by computing $\delta = \hat{\mathbf{s}} - \mathbf{s}$, the user P_{j_i} is determined to be a cheater if $\delta_i \neq 0$.

Hence, if the number of cheating users is less than $\lfloor \frac{m-t-1}{2} \rfloor$, the secret can be reconstructed correctly and all the cheating users can be identified.

5 Conclusion and Future Work

We have proposed an efficient ideal and perfect generalized secret sharing scheme based on N$^{\mu}$MDS codes with desirable security features of cheating detection and cheater identification. The use of the N$^{\mu}$MDS matrices allows us to have authorized sets of varying sizes thus allowing the scheme to have a generalized and richer access structure. The proposed scheme allows an access structure consisting of $\mu + 1$ mutually nonmonotonic sets of user groups of sizes, $t, t - 1, \ldots, t - \mu$, respectively, where $1 \leq \mu < t$, where n is the number of users and the parameter t for the access structure is independent of the field size. The proposed scheme admits a finer access structure and provides a direction towards a fully generalized secret sharing scheme. We believe a fully generalized secret sharing scheme realizing arbitrary access structures should be possible with *almost* MDS codes. We are studying the properties of these codes and working on generating an almost MDS code for any given access structure.

Acknowledgments. The authors acknowledge the support of the Department of Mathematics, BITS Goa, Indian Institute of Technology, Jammu, and R. C. Bose Centre for Cryptology and Security, ISI Kolkata.

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. Blakley, G.: Safeguarding cryptographic keys. In: AFIPS, pp. 313–317. AFIPS Press (1979)
3. Blakley, G., Kabatiansky, G.: Generalized ideal secret-sharing schemes and matroids. Probl. Peredachi Informatsii **33**(3), 102–110 (1997)
4. Dijk, M.: A linear construction of perfect secret sharing schemes. In: De Santis, A. (ed.) EUROCRYPT 1994. LNCS, vol. 950, pp. 23–34. Springer, Heidelberg (1995). https://doi.org/10.1007/BFb0053421
5. Dodunekov, S.: Applications of near MDS codes in cryptography. In: Enhancing Cryptographic Primitives with Techniques from Error Correcting Codes, NATO Science for Peace and Security Series - D: Information and Communication Security, vol. 23, pp. 81–86. IOS Press (2009)
6. Harn, L., Lin, C.: Detection and identification of cheaters in (t, n) secret sharing scheme. Des. Codes Cryptograph. **52**(1), 15–24 (2009)
7. Huffman, W.C., Pless, V.: Fundamentals of Error-Correcting Codes. Cambridge University Press, Cambridge (2010)
8. 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)
9. Massey, J.: Minimal codewords and secret sharing. In: Sixth Joint Swedish-Russian Workshop on Information Theory, Molle, Sweden, pp. 276–279 (1993)
10. McEliece, R., Sarwate, D.: On sharing secrets and Reed-Solomon codes. Commun. ACM **24**(9), 583–584 (1981)

11. Mehta, S., Saraswat, V., Sen, S.: Secret sharing using near-MDS codes. In: Carlet, C., Guilley, S., Nitaj, A., Souidi, E.M. (eds.) C2SI 2019. LNCS, vol. 11445, pp. 195–214. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-16458-4_12
12. 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
13. Shamir, A.: How to share a secret. Commun. ACM **22**(11), 612–613 (1979)
14. Viswanath, G., Rajan, B.S.: Matrix characterization of generalized Hamming weights. In: IEEE International Symposium on Information Theory, p. 61. IEEE (2001)