Noname manuscript No.
(will be inserted by the editor)

Linear Secret Sharing for Hierarchical Access Structures

Ali Aydın Selçuk · Ramazan Yılmaz

the date of receipt and acceptance should be inserted later

Abstract In this paper, we focus on the problem of constructing secret sharing schemes realizing disjunctive hierarchical access structures. We propose two schemes for this problem. The first scheme gives a perfect solution with an overwhelming probability, while the solutions provided by the second scheme, which is an extension of the first one, is always perfect. Moreover, both schemes are ideal. The proposed schemes are based on simple linear algebra and are easy to understand and implement.

Keywords secret sharing · hierarchical access structures · ideal secret sharing

1 Introduction

A secret sharing scheme defines a method to assign shares for a secret among a set of participants such that only qualified subsets of participants can recover the secret. A secret sharing scheme is said to be *perfect* if a non-qualified subset gains no information about the secret. The set of authorized subsets, conventionally denoted by Γ, is called the *access structure*. An access structure is called *monotone* if for all subsets $W \in \Gamma$, $W \subset W'$ requires $W' \in \Gamma$. Ito et al. [5] showed that there exists a perfect secret sharing scheme realizing any monotone access structure. A secret sharing scheme is called *ideal* if the size of the set that the share of any participant belongs to is equal to the size of the set that the secret is taken from.

A specific access structure form is threshold access structures, in which the secret is recoverable by a coalition if and only if the size of that coalition reaches the threshold value. The threshold secret sharing schemes introduced

This work is supported in part by the Turkish Scientific and Technological Research Agency (TÜBİTAK), under grant number 108E150.

Department of Computer Engineering, Bilkent University, Ankara, 06800, Turkey
E-mail: {selcuk, ryilmaz}@cs.bilkent.edu.tr

by Shamir [6] and Blakley [2] are very useful schemes, and these can also be used for constructing schemes that realize more than just basic threshold secret sharing. One of these more advanced structures is the disjunctive hierarchical access structure introduced by Simmons [7]. In a hierarchical access structure, each participant is assigned a level and each level is assigned a threshold. Contrary to the basic threshold access structures, participants in such a scheme are not all equivalent. A participant from a higher level can take place in a coalition of a lower level, but not vice versa. The secret is recoverable if and only if the number of participants from a particular level meets the corresponding threshold value.

In this paper, we propose two ideal secret sharing schemes for disjunctive hierarchical access structures. The first scheme is the basic one and it is perfect with an overwhelming probability. The second scheme is an extension of the first one, and it is always guaranteed to be perfect.

In the rest of the paper, all values and computations are in $\mathbb{Z}_p$ for some large prime p, and vectors are denoted as row matrices, unless otherwise is stated.

2 Background

In this section, we give an overview of threshold secret sharing and hierarchical access structures.

2.1 Threshold Secret Sharing

Let U denote the set of all participants. The access structure of a threshold secret sharing scheme is defined as

$$\Gamma = \{W \subset U : |W| \geq t\}$$

where t is the threshold value.

Blakley secret sharing [2] is one of the best-known threshold secret sharing schemes, which is based on linear algebra. The dealer selects a random vector X in $\mathbb{Z}_p^t$, for some prime p, whose first coordinate is equal to the secret. The dealer also selects a random vector $A_u = (a_{u,1}, a_{u,2}, ..., a_{u,t}) \in \mathbb{Z}_p^t$ for each participant u, and gives $y_u = A_u X^T$ as a share to u. In other words, the dealer assigns a hyperplane equation,

$$a_{u,1}x_1 + a_{u,2}x_2 + \ldots + a_{u,t}x_t = y_u$$

to each participant u. The A_u vector is made public, and y_u is the secret share of user u. In this paper, we will represent the hyperplane of a participant u by (A_u, y_u).

For A denoting the $t \times t$ coefficient matrix

$$\begin{bmatrix} A_{u_1} \\ A_{u_2} \\ \vdots \\ A_{u_t} \end{bmatrix}$$

formed by a t-member coalition $\{u_1, u_2, \ldots, u_t\}$, and y denoting the column vector $(y_{u_1}, y_{u_1}, \ldots y_{u_t})^T$, this coalition can recover the secret by solving the linear system

$$A\,x = y,$$

provided that A is nonsingular.

Shamir secret sharing is another popular threshold secret sharing scheme. In this scheme, the dealer selects a random degree $t-1$ polynomial $f(x)$, such that the secret is $f(0)$, i.e. the constant term of the polynomial. For each participant u, the dealer selects a random non-zero x_u and gives $f(x_u)$ as the private share to u. x_u is made public. When any coalition of size t is present, they can calculate the polynomial by Lagrange interpolation and find the secret $f(0)$.

Shamir secret sharing can be seen as a special case of Blakley secret sharing where the dealer generates the A_u vector as $a_{u,i} = r_u^{i-1}$ for some value r_u. Note that the coefficient matrix formed by a qualified subset is always a Vandermonde matrix, hence is guaranteed to be nonsingular. Shamir secret sharing is always perfect.

2.2 Hierarchical Access Structures

Let U be the set of all participants, and let m nested subsets L_i, $1 \leq i \leq m$ be the levels of a hierarchy satisfying $L_i \subset L_j$ if $i < j$ and $L_m = U$. The access structure is defined as

$$\Gamma = \{W \subset U : |W \cap L_i| \geq t_i \text{ for some } i, 1 \leq i \leq m\}$$

where $0 < t_1 < t_2 < ... < t_{m-1} < t_m$ are the threshold values of the levels.

Brickell [3] proposed several schemes for hierarchical access structures. The main scheme is based on Shamir secret sharing scheme: The dealer determines a Shamir polynomial of degree $t_m - 1$. Let the coefficients be a_i, $0 \leq i \leq t_m - 1$, and the secret is a_0. For each level i, the dealer defines Shamir polynomials $f_i(x) = \sum_{j=0}^{t_i - 1} a_j x^j$ where t_i is the threshold value for the ith level. Note that the secret is the same for all polynomials. The drawback of this scheme is that the nonsingularity of the coefficient matrix is not guaranteed, so the dealer needs to check exponentially many matrices.

Ghodosi et al. [4] studied compartmented and hierarchical access structures, and they proposed a Shamir based secret sharing scheme for hierarchical access structures: For each level i, the dealer selects a polynomial $f_i(x)$. These

polynomials are selected such that for a participant $u \in L_i$, $f_j(x_u) = y_u$ for all $i \leq j \leq m$. In this way, u can participate in qualified coalitions of level j for $i \leq j \leq m$. The degrees of the polynomials are defined recursively: the degree of $f_{i+1}(x)$ depends on not only thresholds t_i, but also on the degree of $f_i(x)$ and $|L_{i+1} - L_i|$. Because of this, the scheme is not dynamic. A new participant cannot be added to any level, except the last level, without changing the existing participants' shares.

Belenkiy [1] proposed another scheme for hierarchical access structures. In this scheme, the dealer selects a degree $t_m - 1$ polynomial $f(x)$ with the secret s as the coefficient of x^{t_m-1} term, and gives values on this polynomial to the participants in the last level of the hierarchy. For the other levels, the dealer takes multiple derivatives of $f(x)$ and uses resulting polynomials for assigning values to the participants. For a user u with identity x_u in the ith level, the dealer computes $f_i(x) = f^{(t_m-t_i)}(x)$ and gives $f_i(x_u)$ as its share to u. Note that all polynomials $f_i(x)$ contains the secret as a coefficient. When any t_i participants from the ith level are present, they have t_i equations with t_i unknowns (coefficients), and find the secret.

More recently, *conjunctive* hierarchical access structures and schemes realizing such access structures have been introduced by Tassa [8] and Tassa and Dyn [9].

3 Proposed Schemes

In this section, we propose two secret sharing schemes for disjunctive hierarchical access structures. The first scheme, which is almost surely perfect, is based on Blakley secret sharing. The second scheme is an extension of the first one such that it is always perfect.

3.1 Basic Scheme

3.1.1 Share Generation

The dealer selects m random points $X_1, X_2, ..., X_m$ over $\mathbb{Z}_p^{t_m}$. The selection of these points are subject to two conditions:

- The first coordinate of all points are equal to the secret.
- The points are affinely independent.

Let C_i denote the set difference $L_i - L_{i-1}$, with $C_1 = L_1$. For a participant $u \in C_i$, the dealer finds a hyperplane (A_u, y_u) passing through X_j for all $i \leq j \leq m$. A_u is made public and y_u is the private share of u.

For each point X_i, the last $t_m - t_i$ coordinates are made public. Only the first t_i coordinates, including the secret, are private. Hence, to solve the private coordinates of X_i, a coalition needs to have t_i hyperplanes passing through X_i.

The reason of the first condition of the selection of points is clear; qualified coalitions of all levels should compute the same secret. For the second condition, assume $X_{i_1}, X_{i_2}, X_{i_3}$ are affinely dependent for some $i_1 < i_2 < i_3$. Then a hyperplane (A_u, y_u) assigned to $u \in C_{i_2}$ will pass through X_{i_1} too, which is not desired.

3.1.2 Reconstruction

When any t_i participants from L_i come together, they will have t_i hyperplanes passing through X_i. Since only the first t_i coordinates of X_i are private, they will compute X_i by solving the $t_i \times t_i$ linear system they have and find the secret $s = x_{i,1}$.

3.1.3 Perfectness

A secret sharing scheme is said to be perfect if

- an unqualified subset gains no information about the secret, and
- a qualified subset can compute the secret.

We show that the proposed scheme is perfect with an overwhelming probability in the following lemmas and theorems.

Lemma 1 *For $1 \leq i < j \leq m$, we have $t_j - t_i \geq j - i$.*

Proof We have $t_i < t_{i+1} < ... < t_{j-1} < t_j$. So

$$\begin{aligned} t_j - t_{j-1} &\geq 1 \\ t_{j-1} - t_{j-2} &\geq 1 \\ &\vdots \\ t_{i+2} - t_{i+1} &\geq 1 \\ t_{i+1} - t_i &\geq 1 \end{aligned}$$

Adding up the inequalities proves the desired result.

Lemma 2 *In the share generation phase, the degree of freedom of the linear system $X_j A_u^T = y_u$, for $i \leq j \leq m$, which the dealer needs to solve for A_u and y_u for user $u \in C_i$, is at least t_i.*

Proof In the linear system,

$$\begin{aligned} X_i A_u^T &= y_u \\ X_{i+1} A_u^T &= y_u \\ &\vdots \\ X_m A_u^T &= y_u \end{aligned}$$

we have $t_m + 1$ unknowns to solve in A_u and y_u.

The number of linear equations is $m-i+1$. Therefore, the degree of freedom is at least $(t_m+1)-(m-i+1)$. By Lemma 1, we have $t_m - t_i \geq m-i$; hence the degree of freedom is at least t_i.

In the following theorems, for a given subset W, l_i denotes $|W \cap L_i|$ and c_i denotes $|W \cap C_i|$.

Theorem 1 *Let W be an unqualified user set of size l, and let P_W denote the probability of W not being able to construct the secret. We have,*

$$P_W \geq (1 - \frac{1}{p})^l.$$

Proof We will first develop the linear system W has on each level i, $1 \leq i \leq m$, and then develop the system over all levels.

W has l_i equations regarding X_i, for $1 \leq i \leq m$. For $u \in L_i$, if the hyperplane assigned to u is (A_u, y_u), we have

$$A_u X_i^T = y_u \tag{1}$$

Since the last $t_m - t_i$ coordinates of X_i are public, this can be written as

$$A_u^{'} X_i^{'T} = y_u^{(i)} \tag{2}$$

where $X_i^{'}$ denotes the $1 \times t_i$ private section of X_i, $A_u^{'}$ is the corresponding, first t_i coefficients in A_u, and

$$y_u^{(i)} = y_u - \sum_{j=t_i+1}^{t_m} a_j x_{i,j} \tag{3}$$

for $A_u = (a_1, a_2, \ldots, a_{t_m})$. W has l_i such equations for each $1 \leq i \leq m$. When these equations are written in matrix form, W has

$$A^{(i)} X_i^{'T} = Y_i, \tag{4}$$

for $1 \leq i \leq m$, where the $l_i \times t_i$ matrix $A^{(i)}$ is formed by the $A_u^{'}$ row vectors in (2), and the $l_i \times 1$ column vector Y_i is formed by the $y_u^{(i)}$ values in (3).

Let D_i denote the first column of $A^{(i)}$, and E_i denote the remaining $l_i \times (t_i - 1)$ part of $A^{(i)}$. Hence $A^{(i)} = [D_i \; E_i]$. Similarly, $X_i^{'} = [s \; V_i]$, for s denoting the secret and V_i denoting the last $t_i - 1$ coordinates of $X_i^{'}$. Then, (4) can be written as

$$[D_i \; E_i][s \; V_i]^T = Y_i.$$

When all equations are combined into a single system, we get:

$$\left[\begin{array}{c|ccccc} D_1 & E_1 & 0 & 0 & \ldots & 0 \\ D_2 & 0 & E_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ D_m & 0 & \ldots & \ldots & 0 & E_m \end{array}\right] \left[\begin{array}{c} s \\ V_1 \\ V_2 \\ \vdots \\ V_m \end{array}\right] = \left[\begin{array}{c} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{array}\right]$$

The coalition W can compute the secret s if and only if the rows of the coefficient matrix above span the unit vector $(1, 0, \ldots, 0)$. That requires the E matrix

$$E = \begin{bmatrix} E_1 & 0 & 0 & \ldots & 0 \\ 0 & E_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & 0 & E_m \end{bmatrix}$$

to have linearly dependent rows (i.e. is not full-rank). If E is not full-rank, then E_i is not full-rank for some i.

Therefore, W can find the secret only if E_i is not full-rank for some i. If E_i matrices are all full-rank, then W cannot find the secret. The probability of all E_i matrices being full-rank is bounded from below by $(1 - \frac{1}{p})^l$, as we show in Appendix A.1. Hence, $P_W \geq (1 - \frac{1}{p})^l$.

Theorem 2 *Given that an unqualified set W cannot find the secret, W gains no information about the secret.*

Proof Assume an unqualified set W satisfies $|W \cap L_i| = t_i - 1$ for some i. Let the share of a participant $v \notin W$, $v \in L_i$, be y_v. W has t_i equations regarding X_i, and one of them is $A_v X_i^T = y_v$. When they solve the system of equations, they will have $s = k_1 y_v + k_2$ for some $k_1, k_2 \in \mathbb{Z}_p, k_1 \neq 0$. Hence, all values are possible for the secret for an unknown y_v. The situation is more clear when $|W \cap L_i| < t_i - 1$.

Theorem 3 *For a qualified subset W, let i be the smallest integer satisfying $l_i \geq t_i$, and let $\bar{P}_W$ denote the probability of W being able to construct the secret. We have*

$$\bar{P}_W \geq \left(1 - \frac{1}{p^2}\right)^{l_{i-1}} \left(1 - \frac{1}{p}\right)^{c_i}. \tag{5}$$

Proof We have $l_j < t_j$, for $j < i$, and $l_i \geq t_i$. We will consider only the first l_i participants of W that are in L_i and take $l_i = t_i$, for the sake of simplicity. As in (4), W has the linear system

$$A^{(i)} X_i^{'T} = Y_i$$

with $A^{(i)}$ being of size $t_i \times t_i$ this time. W can compute the secret if $A^{(i)}$ is nonsingular. The probability that $A^{(i)}$ is nonsingular is related to the probability we computed in Appendix A.1 for Theorem 1. Following a similar methodology, we compute the desired bound (5) for $\bar{P}_W$ in Appendix A.2.

As a final remark for the basic scheme, we would like to note that for $m = 1$ (i.e., when there is only one level of users), the scheme we have proposed here becomes identical to the Blakley threshold secret sharing scheme.

3.2 Extended Scheme

The second scheme extends the basic scheme by adding new dimensions to the space worked in: The dealer chooses m points over $\mathbb{Z}_p^t$, where $t = t_m + m - 1$, instead of over $\mathbb{Z}_p^{t_m}$. In this way, the coordinates used to solve the final linear system to recover the secret will be separate from the coordinates solved to arrange that the hyperplane of a user in level i passes through the points $X_i, \ldots, X_m$. Moreover, the hyperplane coefficients for the coordinates used to solve the final linear system are generated in a Vandermonde-like fashion so that the final system will always be nonsingular.

3.2.1 Share Generation

The dealer selects m random points over $\mathbb{Z}_p^t$, where the ith point is represented as $X_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,t})$, according to the following conditions:

- The first coordinate of every point X_i, $1 \leq i \leq m$, is equal to the secret; i.e. $x_{i,1} = s$, for all $1 \leq i \leq m$.
- The points are affinely independent.
- For X denoting the $m \times m$ matrix containing the last $m - 1$ coordinates of the selected points and -1 as its rows,

$$X = \begin{bmatrix} x_{1,t_m+1} & x_{1,t_m+2} & \ldots & x_{1,T} & -1 \\ x_{2,t_m+1} & x_{2,t_m+2} & \ldots & x_{2,T} & -1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ x_{m,t_m+1} & x_{m,t_m+2} & \ldots & x_{m,T} & -1 \end{bmatrix} \tag{6}$$

  the matrix X is nonsingular.

As in the basic scheme, the dealer publishes the last $t - t_i$ coordinates of each X_i, $1 \leq i \leq m$; and the first t_i coordinates, including the secret, are kept private.

Also just as in the basic scheme, for a participant $u \in C_i$, the dealer finds a hyperplane (A_u, y_u) passing through X_j for all $i \leq j \leq m$. The difference is that, the dealer will select a random value $r_u \in \mathbb{Z}_p^*$ and set $a_{u,j} = (r_u)^{j-1}$ for $1 \leq j \leq t_m$, for $A_u = (a_{u,1}, a_{u,2}, \ldots, a_{u,t})$. Then y_u and the remaining $m - 1$ coordinates of A_u will be selected such that

$$A_u X_j = y_u \tag{7}$$

for $i \leq j \leq m$. Note that the number of equations in this linear system is at most m, and the number of unknowns is m.

The motivation for the first two conditions of selecting the X_i points is the same as that of the basic scheme, and the third condition is needed to guarantee the existence of a solution in (7) for the last $m - 1$ coordinates of

A_u and y_u: Assume $u \in C_i$; then the dealer needs to solve the system,

$$\begin{bmatrix} X_i \\ X_{i+1} \\ \vdots \\ X_m \end{bmatrix} A_u^T = \begin{bmatrix} y_u \\ y_u \\ \vdots \\ y_u \end{bmatrix}$$

to generate the hyperplane (A_u, y_u) for user u. The dealer selects the random $r_u \in \mathbb{Z}_p$ and sets the first t_m coordinates of A_u as $a_{u,j} = (r_u)^{j-1}, 1 \leq j \leq t_m$. Then the system becomes

$$\begin{bmatrix} X_i^{'} \\ X_{i+1}^{'} \\ \vdots \\ X_m^{'} \end{bmatrix} A_u^{'T} - \begin{bmatrix} y_u \\ y_u \\ \vdots \\ y_u \end{bmatrix} = \begin{bmatrix} b_{u,i} \\ b_{u,i+1} \\ \vdots \\ b_{u,m} \end{bmatrix}$$

where $X_j^{'}$ and $A_u^{'}$ denote the last $m-1$ coordinates of X_j and A_u respectively, and $b_{u,k} = -\sum_{j=1}^{t_m} x_{k,j} r_u^{j-1}$ for $i \leq k \leq m$. By including y_u in the vector of unknowns, the dealer has the linear system,

$$\underbrace{\begin{bmatrix} X_i^{'} & -1 \\ X_{i+1}^{'} & -1 \\ \vdots & \vdots \\ X_m^{'} & -1 \end{bmatrix}}_{X'} \begin{bmatrix} A_u^{'T} \\ y_u \end{bmatrix} = \begin{bmatrix} b_{u,i} \\ b_{u,i+1} \\ \vdots \\ b_{u,m} \end{bmatrix} \tag{8}$$

Note that X' is a submatrix of X in (6), and it is just equal to X for $i = 1$. Hence, we have the third condition in the selection of the X_i points during the share generation phase in order to guarantee that the system (8) always has a solution for A'_u and y_u.

3.2.2 Reconstruction

The reconstruction of the secret is the same as that of the basic scheme. Additionally, if desired, Lagrange interpolation can also be used as in Shamir secret sharing: Assume a qualified subset W satisfying $|W \cap L_i| \geq t_i$ for some i is present. Let $f(z)$ denote the degree $t_i - 1$ polynomial, $\sum_{j=1}^{t_i} x_{i,j} z^{j-1}$. Since the last $t - t_i$ coordinates of X_i are public, each participant $u \in W$ can compute $f(r_u)$ as $y_u - \sum_{j=t_i+1}^{t} x_{i,j} a_{u,j}$. Since the coalition W has t_i points on polynomial f, they can compute $f(0) = x_{i,1} = s$.

3.2.3 Perfectness

As explained in Section 3.2.2, a qualified set will have t_i points over a degree $t_i - 1$ polynomial. Just like in Shamir secret sharing, the coefficient matrix will be a Vandermonde matrix, which is always nonsingular. A qualified subset will always be able to compute the secret uniquely.

When a non-qualified subset W is present, the E_i matrices defined in Section 3.1.3 will be truncated Vandermonde matrices, with fewer rows than columns, so they are always full-rank. Hence, a non-qualified subset will not be able to find the secret. As in the basic scheme, all values in $\mathbb{Z}_p$ will be equally likely for the secret.

We would also like to note that the extended scheme reduces to the Shamir threshold secret sharing scheme when there is only one level, i.e. $m = 1$.

4 Conclusion

In both schemes, a single hyperplane is assigned to a user $u \in C_i$ which passes through $m - i + 1$ given points. Since there is a single hyperplane equation and a single secret share y_u per user, the scheme is ideal.

In the extended scheme, instead of choosing the points from a t_m dimensional space, we added new dimensions to be used in solving the hyperplane coefficients and increased the number of dimensions to $t_m + m - 1$. By adding these new dimensions, for each user $u \in U$, the dealer can set the first t_m entries of A_u such that the coefficient matrix formed by a qualified subset of participants is always a Vandermonde matrix. This guarantees that the extended scheme is always perfect.

Note that the affine independence condition for the selection of the points in both schemes can be dropped, since the participants will not know that the points are affinely dependent even if they are. So, even in that case, all values of the secret are equally likely for an unqualified subset.

References

1. M. Belenkiy. Disjunctive multi-level secret sharing. Cryptology ePrint Archive, Report 2008/018, 2008.
2. G. Blakley. Safeguarding cryptographic keys. In *AFIPS National Computer Conference*, 1979.
3. E.F. Brickell. Some ideal secret sharing schemes. In *EUROCRYPT'89*, volume 434 of *LNCS*, pages 468–475. Springer-Verlag, 1990.
4. H. Ghodosi, J. Pieprzyk, and R. Safavi-Naini. Secret sharing in multilevel and compartmented groups. In *ACISP'98*, volume 1438 of *LNCS*, pages 367–378, London, UK, 1998. Springer-Verlag.
5. M. Ito, A. Saito, and T. Nishizeki. Secret sharing scheme realizing general access structure. In *GLOBECOM'87*, pages 99–102. IEEE Press, 1987.
6. A. Shamir. How to share a secret? *Communications of the ACM*, 22(11):612–613, 1979.
7. G. J. Simmons. How to (really) share a secret. In *CRYPTO'88*, volume 403 of *LNCS*, pages 390–448, London, UK, 1988. Springer-Verlag.

8. T. Tassa. Hierarchical threshold secret sharing. *Journal of Cryptology*, 20(2):237–264, 2007.
9. T. Tassa and N. Dyn. Multipartite secret sharing by bivariate interpolation. *Journal of Cryptology*, 22(2):227–258, 2008.

A Lower Bounds for Theorems 1 and 3

Let $P^{(p)}_{(m,n)}$, for $m \leq n$, denote the probability of a randomly generated $m \times n$ matrix over $\mathbb{Z}_p$ to be full-rank. We have the following lower bound regarding $P^{(p)}_{(m,n)}$:

Lemma 3

$$P^{(p)}_{(m,n)} \geq \left(1 - \frac{1}{p}\right)^m .$$

Proof The first row of a full-rank matrix can be anything except for all zeros; so we have $p^n - 1$ possible choices for the first row. The second row cannot be a scalar multiple of the first row; so we have $p^n - p$ possible choices for the second row. In general, the ith row cannot be a linear combination of the first $i-1$ rows; so we have $p^n - p^{i-1}$ possible choices for the ith row. Therefore, the proportion of full-rank matrices among all $m \times n$ matrices is,

$$\begin{aligned} P^{(p)}_{(m,n)} &= \frac{(p^n-1)(p^n-p)\ldots(p^n-p^{m-1})}{(p^n)^m} \\ &= \frac{p^n-1}{p^n}\,\frac{p^n-p}{p^n}\cdots\frac{p^n-p^{m-1}}{p^n} \\ &\geq \left(\frac{p^n-p^{m-1}}{p^n}\right)^m \\ &\geq \left(\frac{p^n-p^{n-1}}{p^n}\right)^m \\ &= \left(1-\frac{1}{p}\right)^m . \end{aligned}$$

Let M be an $m \times n$ matrix over $\mathbb{Z}_p$, for $m \leq n$, such that the first m_1 rows of M are given to be linearly independent and the remaining $m_2 = m - m_1$ rows are generated randomly. Let $P^{(p)}_{(m_1,m_2,n)}$ denote the probability that all the rows of M are linearly independent. We have the following lower bound for $P^{(p)}_{(m_1,m_2,n)}$:

Lemma 4

$$P^{(p)}_{(m_1,m_2,n)} \geq \left(1 - \frac{1}{p^{n-m+1}}\right)^{m_2} .$$

Proof For the selection of the (m_1+j)th row, $1 \leq j \leq m_2$, there are $p^n - p^{m_1+j-1}$ possible choices given that the previous (m_1+j-1) rows are linearly independent. Therefore the proportion of the full-rank M matrices, given the first m_1 rows are linearly independent, is

$$\begin{aligned} P^{(p)}_{(m_1,m_2,n)} &= \prod_{j=1}^{m_2} \frac{p^n - p^{(m_1+j-1)}}{p^n} \\ &\geq \left(\frac{p^n - p^{(m-1)}}{p^n}\right)^{m_2} \\ &= \left(1 - \frac{1}{p^{n-m+1}}\right)^{m_2} . \end{aligned}$$

Note that Lemma 3 is a special case of Lemma 4 for $m_1 = 0$ and $m_2 = m$.

In the following discussion, $C_i = L_i - L_{i-1}$, with $C_1 = L_1$. For a given subset W, we take $l_i = |W \cap L_i|$ and $c_i = |W \cap C_i|$.

A.1 A Lower Bound for Theorem 1

Let W be an unqualified set of size l, hence $l_m = l$. Let Q_i denote the probability of all E_j matrices obtained by W, for $1 \leq j \leq i$, being full-rank. We have the following lower bound regarding Q_i:

Lemma 5

$$Q_i \geq \left(1 - \frac{1}{p}\right)^{l_i}.$$

Proof For the first level, note that the degree of freedom in generation of the hyperplane for a user $u \in C_1$ is at least t_1 by Lemma 2; and the rows of $A^{(1)}$ are of size t_1; therefore, $A^{(1)}$ is completely random. Since E_1 is a submatrix of $A^{(1)}$, it is completely random too. Then by Lemma 3, we have,

$$Q_1 = P^{(p)}_{(l_1, t_1 - 1)} \geq \left(1 - \frac{1}{p}\right)^{l_1} = \left(1 - \frac{1}{p}\right)^{c_1}. \tag{9}$$

For $i \geq 2$, first note that $u \in W \cap L_{i-1}$ implies $u \in W \cap L_i$. We can assume that the first l_{i-1} rows of E_i come from $W \cap L_{i-1}$, and E_i contains E_{i-1} as its upper-left corner submatrix. For R_i denoting the probability that E_i is full-rank given that E_{i-1} is full-rank, we have,

$$Q_i = Q_{i-1} R_i. \tag{10}$$

To calculate R_i, note that the degree of freedom in generation of the hyperplane for a user $u \in C_i$ is at least t_i, by Lemma 2, and the rows of $A^{(i)}$ are of size t_i too. Therefore, the rows of $A^{(i)}$, hence the rows of E_i, that come from C_i (i.e. those after E_{i-1}) are completely random. So we have,

$$\begin{aligned} R_i &= P^{(p)}_{(l_{i-1}, c_i, t_i - 1)} \\ &\geq \left(1 - \frac{1}{p^{(t_i - l_i)}}\right)^{c_i}. \end{aligned}$$

Since we always have $l_i < t_i$ for an unqualified set W, we have,

$$R_i \geq \left(1 - \frac{1}{p}\right)^{c_i} \tag{11}$$

By substituting (11) in (10) recursively with the base case (9) for Q_1, and by the fact that $\sum_{j=1}^{i} c_j = l_i$, we get,

$$Q_i \geq \left(1 - \frac{1}{p}\right)^{l_i}.$$

For the particular case $i = m$, we have the result needed in Theorem 1:

$$Q_m \geq \left(1 - \frac{1}{p}\right)^{l_m} = \left(1 - \frac{1}{p}\right)^{l}.$$

A.2 A Lower Bound for Theorem 3

Let W be a qualified set of users W. As given in (4), Section 3.1.3, W has a linear system of equations $A^{(j)} {X'_j}^T = Y_j$ for each level j. Let Q'_j denote the probability of all $A^{(k)}$, $1 \leq k \leq j$, to be full-rank for a given j. For i being the smallest integer satisfying $l_i \geq t_i$, we have the following lower bound regarding Q'_i:

Lemma 6 *For a qualified set of users W, and i denoting the smallest integer satisfying $l_i \geq t_i$ in W, we have*

$$Q'_i \geq \left(1 - \frac{1}{p^2}\right)^{l_{i-1}} \left(1 - \frac{1}{p}\right)^{c_i}.$$

Proof As stated in the proof of Lemma 5, the matrix $A^{(1)}$ is completely random. Then,

$$Q'_1 = P^{(p)}_{(l_1,t_1)} \geq \left(1 - \frac{1}{p}\right)^{l_1} = \left(1 - \frac{1}{p}\right)^{c_1}. \tag{12}$$

As in the proof of Lemma 5, again, $A^{(j-1)}$ can be seen as the upper-left corner submatrix of $A^{(j)}$. For R_j denoting the probability that $A^{(j)}$ is full-rank given that $A^{(j-1)}$ is full-rank, we have,

$$Q'_j = Q'_{j-1} R_j. \tag{13}$$

By Lemma 2, the degree of freedom in generation of the hyperplane for a user $u \in C_j$ is at least t_j, which is equal to the size of the rows of $A^{(j)}$. Therefore, the rows of $A^{(j)}$ that come from C_j (i.e. those after $A^{(j-1)}$) are completely random. Hence,

$$\begin{aligned} R_j &= P^{(p)}_{(l_{j-1},c_j,t_j)} \\ &\geq \left(1 - \frac{1}{p^{(t_j - l_j + 1)}}\right)^{c_j}. \end{aligned}$$

For levels $j < i$, we have $l_j < t_j$. Therefore,

$$R_j \geq \left(1 - \frac{1}{p^2}\right)^{c_j}. \tag{14}$$

For level i, we have $l_i = t_i$, and therefore,

$$R_i \geq \left(1 - \frac{1}{p}\right)^{c_i}. \tag{15}$$

By substituting (15) and (14) in (13) with the base case (12), and by the fact that $\sum_{j=1}^{i-1} c_j = l_{i-1}$, we get,

$$Q'_i \geq \left(1 - \frac{1}{p^2}\right)^{l_{i-1}} \left(1 - \frac{1}{p}\right)^{c_i}.$$