Chapter 8

Secret Sharing Schemes

8.1 Introduction

According to *Time Magazine*, May 4, 1992, control of nuclear weapons in Russia involves a two-out-of-three mechanism. In order to launch a nuclear missile, the cooperation of at least two parties out of three are needed. The three parties involved are the president, the Defence Minister, and the Defence Ministry.

A similar situation can occur in a bank, where there is a vault which must be opened every day. The bank employs a number of senior tellers, who are trusted enough to participate in the opening of the vault, but not trusted to the tent that they themselves own the combination to the vault. We would like to design a system where for example any two of the senior tellers together can open the vault, but no individual alone can do so.

We have given examples of shared secrets and shared control. The above problems can be solved by means of a *secret sharing scheme*, the topic of this chapter. Secret sharing schemes were independently introduced by Blakley [**?**] and Shamir [**?**] in 1979. We will first consider a special type of secret sharing scheme, called a threshold scheme, before we proceed to the general case.

8.2 Threshold schemes

We start by an informal definition of a threshold scheme.

Definition 8.1. *Let k and n be positive integers, $k \leq n$. A (k, n)-threshold scheme is a method of sharing a secret K among a set of n participants in such a way that any k participants can compute the value of the secret, but no group of $k-1$ or fewer can do so.*

The examples in the previous section are threshold schemes. We will give a more precise definition later.

Let the set of participants be denoted by $\mathcal{P}$. The value of the secret K is chosen by the *dealer*, denoted D, who is a special participant not in $\mathcal{P}$. When D wants to share the secret K among the participants in $\mathcal{P}$, D gives each participant some partial information, called a *share*. The shares are distributed secretly, so no participant knows any other participant's share.

At a later time, some qualified subset of participants $\mathcal{B} \subseteq \mathcal{P}$ would like to compute the secret K. They will then pool their shares together in an attempt to compute K.

We will now present the most famous construction of a (k, n)-threshold scheme, called the *Shamir Threshold Scheme*, invented in 1979. First some notations. Let $\mathcal{P} = \{P_1, P_2, \ldots, P_n\}$ be the set of participants, $\mathcal{K}$ the set of possible secrets, and $\mathcal{S}$ the set of possible shares (if different participants have different sizes of shares, $\mathcal{S}$ denotes the largest of them). In the Shamir Threshold Scheme $\mathcal{K} = \mathbb{Z}_p$, where $p \geq n+1$ is a prime. Also, $\mathcal{S} = \mathbb{Z}_p$. Hence both the secret and the shares are elements of $\mathbb{Z}_p$. The construction is now as follows.

Initilization:

D chooses n distinct non-zero elements from $\mathbb{Z}_p$, denoted x_i, $1 \leq i \leq n$. All these values are public.

Share Distribution:

1. D wants to share the secret $K \in \mathbb{Z}_p$. D randomly chooses $k-1$ elements of $\mathbb{Z}_p$, denoted $a_1, a_2, \ldots, a_{k-1}$. Furthermore, $a_0 = K$.

2. D computes $y_i = a(x_i)$, for $1 \leq i \leq n$, where

$$a(x) = \sum_{j=0}^{k-1} a_j x^j \mod p.$$

3. D gives participant P_i the share y_i.

In short, the dealer constructs a random polynomial of degree at most $k-1$ in which the constant term is the secret K, i.e., $a_0 = K$. Every participant obtains a point (x_i, y_i) on this polynomial.

We now have to verify two different properties, firstly that any k participants can reconstruct the polynomial $a(x)$ and, hence, calculate the secret, and secondly that any group of $k-1$ participants cannot do so.

Let us start by looking at how k participants can reconstruct the polynomial $a(x)$. This is basically done by means of polynomial interpolation. Without loss of generality we can assume that $\mathcal{B} = \{P_1, P_2, \ldots, P_k\}$ is the set that is to reconstruct the secret. The participants in $\mathcal{B}$ know

$$y_i = a(x_i), \quad 1 \leq i \leq k,$$

where $a(x) \in \mathbb{Z}_p[x]$ is the secret polynomial chosen by D. The polynomial $a(x)$ has degree at most $k-1$ and, hence, can be written

$$a(x) = a_0 + a_1 x + \cdots + a_{k-1} x^{k-1},$$

where the coefficients $a_0, a_1, \ldots, a_k$ are unknown elements of $\mathbb{Z}_p$ and $a_0 = K$ is the secret. Each participant knowing $y_i = a(x_i)$ can obtain a linear equation in the k unknowns

$a_0, a_1, \ldots, a_{k-1}$. So the group $\mathcal{B}$ has k linear equations at its disposal. If the equations are all linearly independent, there will be a unique solution, and a_0 will be revealed as the key.

The system of linear equations is the following:

$$\begin{array}{rcl} a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_{k-1} x_1^{k-1} & = & y_1, \\ a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_{k-1} x_2^{k-1} & = & y_2, \\ \vdots & \vdots & \vdots \\ a_0 + a_1 x_k + a_2 x_k^2 + \cdots + a_{k-1} x_k^{k-1} & = & y_k, \end{array}$$

which can be written in matrix form as follows:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{k-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_k & x_k^2 & \ldots & x_k^{k-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix}.$$

The coefficient matrix, we call it A, is a so-called Vandermonde matrix. There is a well-known formula for the determinant of a $k \times k$ Vandermonde matrix, namely

$$\det A = \prod_{1 \leq i < j \leq k} (x_i - x_j) \mod p.$$

Since the x_i's are all distinct, no term in the product equals 0. Because of the fact that $\mathbb{Z}_p$ is a field, the product of non-zero elements is again a non-zero element, and we have that $\det A \neq 0$. Since the coefficient matrix has a non-zero determinant, the system of linear equations has a unique solution over $\mathbb{Z}_p$. This establishes that any k participants can recover the polynomial $a(x)$ and, hence, the secret K. Here is an example.

Example 8.1. *Suppose that $p = 17$, $k = 3$, $n = 5$, and the public x-coordinates are $x_i = i$, for $1 \leq i \leq 5$. Suppose that $\mathcal{B} = \{P_1, P_3, P_5\}$ pool their shares, which are 8, 10, and 11, respectively. Writing the polynomial $a(x)$ as*

$$a(x) = a_0 + a_1 x + a_2 x^2,$$

and computing $a(1)$, $a(3)$, $a(5)$ in the unknowns a_0, a_1, a_2 results in the following three linear equations in $\mathbb{Z}_{17}$,

$$\begin{array}{rcl} a_0 + a_1 + a_2 & = & 8, \\ a_0 + 3a_1 + 9a_2 & = & 10, \\ a_0 + 5a_1 + 8a_2 & = & 11. \end{array}$$

Solving this system of linear equations over $\mathbb{Z}_{17}$ gives us the polynomial $a(x) = 13 + 10x + 2x^2$, and therefore, the secret is $K = a_0 = 13$.

We now consider the second issue, namely what will happen if a group of $k-1$ participants tries to compute the secret. Proceeding as above, they will end up with a system of $k-1$

linear equations in k unknowns. We will now show that with this information only, the secret can have any value in $\mathbb{Z}_p$. We do as follows. Suppose that the secret K has the value y_0. Since the secret $K = a_0 = a(0)$, we have

$$y_0 = a(0)$$

and this will yield a kth linear equation. This linear equation together with the previous $k-1$ linear equations will result in a linear system of k equations in k unknowns, where the coefficient matrix is again a Vandermonde matrix. As before, there is a unique solution. So for every possible value y_0 of the secret K, there is a unique polynomial $a_{y_0}(x)$ such that

$$y_i = a_{y_0}(x_i),$$

for $1 \le i \le k-1$, and such that

$$y_0 = a_{y_0}(0).$$

This means that no value of the secret can be ruled out by the group of $k-1$ participants, and they have obtained no information about the secret.

We continue by giving an alternative method of calculating the polynomial $a(x)$, which is based on the Lagrange interpolation formula for polynomials. The Lagrange interpolation formula is an explicit formula for the unique polynomial $a(x)$ of degree at most $k-1$, when the values in k different points $y_1 = a(x_1), y_2 = a(x_2), \ldots, y_k = a(x_k)$ are given. It is as follows:

$$a(x) = \sum_{i=1}^{k} y_i \prod_{1\le j\le k, j\ne i} \frac{x - x_j}{x_i - x_j}.$$

We can verify the correctness of the formula by substituting $x = x_i$, and all terms in the sum vanish except for the ith term, which is y_i.

Hence, any group of k participants can calculate $a(x)$ using the interpolation formula. But the group is only interested in calculating the secret $K = a_0$, and we can then make a simplification. Since $K = a_0 = a(0)$, we substitute $x = 0$ into the Lagrange interpolation formula and get

$$K = \sum_{i=1}^{k} y_i \prod_{1\le j\le k, j\ne i} \frac{x_j}{x_j - x_i}.$$

So we end up with an explicit formula for the secret.

Example 8.2. *Consider the previous example. The set of participants $\{P_1, P_3, P_5\}$ can calculate the key according to the formula above and gets*

$$\begin{aligned} K &= \sum_{i=1}^{k} y_i \prod_{1\le j\le k, j\ne i} \frac{x_j}{x_j - x_i} \\ &= 8 \cdot \frac{3\cdot 5}{(-2)\cdot(-4)} + 10 \cdot \frac{1\cdot 5}{2\cdot(-2)} + 11 \cdot \frac{1\cdot 3}{4\cdot 2} \mod 17 \\ &= 8\cdot 4 + 10\cdot 3 + 11\cdot 11 \mod 17 \\ &= 13 \mod 17. \end{aligned}$$

We end this section by yet an example, giving a simplified construction for (k,k)-threshold schemes.

Example 8.3. *Consider the following (k, k)-threshold scheme.*

1. *The dealer D chooses k random shares $y_1, y_2, \ldots, y_k$ from $\mathbb{Z}_m$, and gives y_i to participant i as the share.*
2. *The secret K is chosen to be*

$$K = \sum_{i=1}^{k} y_i \mod m.$$

The k participants can recover the key by adding all their shares, but $k - 1$ participants have no information about the key.

8.3 Access structures and general secret sharing

In the previous section, we had a k-out-of-n structure on the subsets qualified to recover the secret. In a more general situation, we would like to specify exactly which subsets that are qualified to recover the secret. Let Γ be a set of subset of $\mathcal{P}$, $\Gamma \subseteq 2^{\mathcal{P}}$. The subsets in Γ specify those subset of participants that are qualified to compute the secret. Then Γ is called an *access structure* and the subsets in Γ are called *qualified subsets* or *authorized subsets*. If $\mathcal{B} \subseteq \mathcal{P}$, then we denote by B the set of shares for the participants in subset $\mathcal{B}$. Using the notation from the previous section, we formally define a general secret sharing scheme as follows.

Definition 8.2. *A secret sharing scheme for the access structure Γ is a method of sharing a secret K among a set of n participants in such a way that the following two properties hold:*

1. *If $\mathcal{B} \in \Gamma$, then $H(K|B) = 0$.*
2. *If $\mathcal{B} \notin \Gamma$, then $H(K|B) \geq log\ \alpha$,*

where α is a fixed value and $\alpha > 1$.

It is clear from the definition that any qualified subset can recover the secret, whereas any non-qualified subset $\mathcal{B}$ has some uncertainty about the secret but can guess the correct value of the secret with probability at least $2^{-H(K|B)}$.

We see that a (k, n)-threshold scheme has access structure

$$\Gamma = \{\mathcal{B} \subseteq \mathcal{P}; |\mathcal{B}| \geq k\}.$$

Definition 8.3. *A secret sharing scheme for which $\alpha = |\mathcal{K}|$ is called a* perfect *secret sharing scheme.*

In a perfect secret sharing scheme any non-qualified subset of participants gets no information about the secret, and has thus no advantage over an outsider in guessing the correct secret. Here an even stronger definition.

Definition 8.4. *A secret sharing scheme is called* ideal *if it is perfect and if* $|\mathcal{S}| = |\mathcal{K}|$.

An ideal secret sharing scheme is perfect and has also the smallest possible size of each participants share. This is an important aspect since the share is secret and thus needs to be as small as possible. We see that the Shamir Threshold Scheme is an ideal secret sharing scheme.

Next we consider some properties of the access structure Γ. Suppose that $\mathcal{B}$ is a qualified subset ($\mathcal{B} \in \Gamma$) and we add yet another participant to the set $\mathcal{B}$. Then the resulting subset must also be qualified, since it contains a qualified subset. Thus, the access structure must have the following property,

$$\text{if } \mathcal{B} \in \Gamma \text{ and } \mathcal{B} \subseteq \mathcal{C} \subseteq \mathcal{P}, \text{ then } \mathcal{C} \in \Gamma.$$

Such an access structure is called a *monotone access structure*. We define the *closure* of an access structure Γ, denoted $\mathrm{cl}(\Gamma)$, as

$$\mathrm{cl}(\Gamma) = \{\mathcal{C} \subseteq \mathcal{P}; \mathcal{B} \subseteq \mathcal{C}, \mathcal{B} \in \Gamma\}.$$

It is easy to see that an access structure is monotone if and only if

$$\Gamma = \mathrm{cl}(\Gamma).$$

It is common to give an access structure Γ for a problem, when it is really the monotone access structure $\mathrm{cl}(\Gamma)$ that is meant. So from now on we assume that if Γ is not monotone, we mean $\mathrm{cl}(\Gamma)$.

The next topic is to show that it is possible to construct a perfect secret sharing scheme for any access structure. We will do this by giving a simple construction due to Benaloh and Leichter [?].

Let Γ_0 be an access structure and suppose that we want to construct a perfect secret sharing scheme for the monotone access structure $\mathrm{cl}(\Gamma_0)$. The construction is as follows.

Initilization:

D chooses any access stucture Γ such that

$$\mathrm{cl}(\Gamma) = \mathrm{cl}(\Gamma_0).$$

Write Γ in the form

$$\Gamma = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_l\},$$

where $\mathcal{B}_i \subseteq \mathcal{P}$ for $1 \leq i \leq l$. The ordered sequence of sets $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_l$ is public information.

Share Distribution:

Let $\mathcal{K} = \mathbb{Z}_m$. For each $\mathcal{B}_j = \{P_{i_1}, P_{i_2}, \ldots, P_{i_k}\}$, the dealer does the following:

1. D constructs a (k,k)-threshold scheme for $\mathcal{K}$ with participants $\mathcal{B}_j$, meaning that D chooses k random values $a_{j1}, a_{j2}, \ldots, a_{jk}$ such that

$$K = a_{j1} + a_{j2} + \cdots + a_{jk} \mod \mathbb{Z}_m.$$

2. D gives the value a_{jm} to participant P_{i_m}, $1 \leq m \leq k$, as a piece of that participant's share.

The share for participant i is then the set of pieces that P_i has received. The number of pieces is the same as the number of times that P_i can be found in the subsets $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_l$.

We illustrate this construction with an example.

Example 8.4. *Let*

$$\Gamma_0 = \{\{P_1, P_2, P_4\}, \{P_1, P_3, P_4\}, \{P_2, P_3\}\} = \{\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3\},$$

and $\mathcal{K} = \mathbb{Z}_m$. Choose $\Gamma_0 = \Gamma$. For each of the subsets in Γ, we design a (k,k)-threshold scheme. Starting with $\mathcal{B}_1$, we choose $a_1, a_2, a_3 \in \mathbb{Z}_m$ such that

$$K = a_1 + a_2 + a_3 \mod \mathbb{Z}_m.$$

We give the piece a_1 to P_1, a_2 to P_2, and a_3 to P_4. For $\mathcal{B}_2$, we choose $b_1, b_2, b_3 \in \mathbb{Z}_m$ such that

$$K = b_1 + b_2 + b_3 \mod \mathbb{Z}_m.$$

We give the piece b_1 to P_1, b_2 to P_3, and b_3 to P_4. Then, for $\mathcal{B}_3$ we choose $c_1, c_2 \in \mathbb{Z}_m$ such that

$$K = c_1 + c_2 \mod \mathbb{Z}_m.$$

We give the piece c_1 to P_2 and c_2 to P_3. The distribution of pieces has resulted in the following shares: P_1 has (a_1, b_1); P_2 has (a_2, c_1); P_3 has (b_2, c_2); and P_4 has (a_3, b_3).

We now prove that this construction gives perfect secret sharing schemes. We omit a formal proof and give only a sketch. For any qualified subset $\mathcal{B}$ there is a $\mathcal{B}_i \in \Gamma$ such that $\mathcal{B}_i \subseteq \mathcal{B}$. So the participants in $\mathcal{B}_i$ can pool their shares together and recover the secret K.

For any non-qualified subset $\mathcal{B}$, there can be *no* $\mathcal{B}_i \in \Gamma$ such that $\mathcal{B}_i \subseteq \mathcal{B}$. Because the threshold schemes are independent, it is enough to consider one of these. Since the non-qualified subset $\mathcal{B}$ is not qualified in any of these threshold schemes, $\mathcal{B}$ will obtain no information about the key in any threshold scheme.

The obvious disadvantage of this construction is that the sizes of the shares will be quite large. One of the main problems in secret sharing is to find constructions for which the shares given to the participants are as small as possible. The best we can hope for is to find ideal secret sharing schemes, but it can for many access structures be proved that they cannot be realized by an ideal secret sharing scheme.

8.4 Exercises

11.1 The Shamir Threshold Scheme is described for $\mathcal{K} = \mathbb{Z}_p$, but it works equally well over any field $\mathbb{F}_q$. Design a $(2,4)$-threshold scheme for $\mathcal{K} = \mathbb{F}_{2^3}$, when $K = 1$.

11.2 Suppose that $p = 19$, $k = 3$, $n = 6$, and the public x-coordinates are $x_i = i$, for $1 \leq i \leq 6$ in a Shamir Threshold Scheme. Suppose that $\mathcal{B} = \{P_2, P_3, P_6\}$ pool their shares, which are 8, 18, and 11, respectively. Calculate the secret K.

11.3 Construct a perfect secret sharing scheme for the access structure

$$\Gamma_0 = \{\{P_1, P_2, P_4, P_5\}, \{P_1, P_2, P_3, P_4\}, \{P_1, P_3\}, \{P_3, P_5\}\}.$$

11.4 Construct an ideal secret sharing scheme for the access structure

$$\Gamma_0 = \{\{P_1, P_2\}, \{P_1, P_3\}, \{P_2, P_3\}, \{P_4, P_5\}\}.$$