COMPSCI 690T Coding Theory and Applications April 3, 2017

Lecture 17

Instructor: Arya Mazumdar *Scribe: Names Redacted*

1 Secret Sharing Scheme

In the last class, we saw how codes are used in group testing and compressed sensing. In this class, we will look at secret sharing schemes and how codes can be used to construct various secret sharing schemes.

In a secret sharing scheme, a secret "s" is distributed among many users. The information that a user gets is called a share. We define a secret sharing scheme such that for some sets of the users, when they get together, the secret is revealed and for some sets of users, the secret can't be revealed from their shares. For the rest of the lecture, we will assume the following notation : Let a secret s be distributed among n users.

Shamir and Blakely independently defined a notion of a secret sharing scheme called Threshold secret sharing scheme.

1.1 Threshold Secret Sharing Scheme

If the secret s is distributed such that the following conditions hold, then it is called Threshold secret sharing scheme.

1 If any set of k or more users get together, then the secret will be revealed.

2 For all sets of $k-1$ or less users, the secret will not be revealed.

It is denoted by (n, k)-Threshold scheme where n is the number of users and k is the threshold.

1.2 Access and Block Structures

Let $[n] = \{1, 2, \ldots, n\}$ and $2^{[n]}$ be power set of $[n]$.

Access Structure $\mathcal{A} \subseteq 2^{[n]}$ is such that any set $a \in \mathcal{A}$ can reveal secret s.
For eg. $\mathcal{A} = \{\{1, 2\}, \{2, 3, 4\}\}$, then when users 1, 2 or users 2, 3, 4 get together, they can reveal s.

Block Structure $\mathcal{B} \subseteq 2^{[n]}$ is such that any set $b \in \mathcal{B}$ cannot reveal the secret s.

The above definitions are a generalization of any secret sharing scheme.

Monotonicity of a secret sharing scheme :

We say that a secret sharing scheme is Monotonic if its Access Structure $\mathcal{A}$ and Block Structure $\mathcal{B}$ satisfy the following properties :

1 $a \in \mathcal{A} \Rightarrow$ for any $a' \supseteq a : a' \in \mathcal{A}$

2 $b \in \mathcal{B} \Rightarrow$ for any $b' \subseteq b : b' \in \mathcal{B}$

Perfect Scheme :

If $\mathcal{B} = 2^{[n]} \setminus \mathcal{A}$, then the secret sharing scheme is called Perfect Scheme.

An example of a (n, k)-Threshold secret sharing scheme which is monotonic and perfect is given by:

$$\mathcal{A} = \{a \subseteq [n] \mid |a| \geq k\}$$
$$\mathcal{B} = \{b \subseteq [n] \mid |b| \leq k-1\}$$

Observe that $\mathcal{B} = 2^{[n]} \setminus \mathcal{A}$

2 Shamir's construction

We will see a construction of an (n, k)-Threshold scheme due to Shamir.

Let the secret $s \in \mathbb{F}_q$ and a polynomial $f(x) \subseteq \mathbb{F}_q[x]$.
$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{k-1} x^{k-1}$

Set $a_0 = s$ and $a_1, a_2, \ldots a_{k-1}$ are picked uniformly at random from $\mathbb{F}_q$.

We will pick n distinct points $\alpha_1, \alpha_2, \ldots \alpha_n$ from $\mathbb{F}_q \backslash \{0\}$ such that $n < q$ and distribute $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)$ among n users. The points of evaluation α_i's are public knowledge i.e, every user knows them.

Claim : The above construction is a (n, k)-Threshold scheme.
Proof : When k or more evaluations of the polynomial f are known, we should be able to obtain a_0 which is secret s.

Consider a set of k points from $\alpha_1, \alpha_2, \ldots, \alpha_n$ given by $\alpha_{i1}, \alpha_{i2}, \ldots \alpha_{ik}$.

$$\begin{aligned}
f(\alpha_{i1}) &= a_0 + a_1 \alpha_{i1} + \cdots + a_{k-1} \alpha_{i1}^{k-1} \\
f(\alpha_{i2}) &= a_0 + a_1 \alpha_{i2} + \cdots + a_{k-1} \alpha_{i2}^{k-1} \\
&\vdots \\
f(\alpha_{ik}) &= a_0 + a_1 \alpha_{ik} + \cdots + a_{k-1} \alpha_{ik}^{k-1}
\end{aligned}$$

We have k unknowns $a_0, a_1, \ldots, a_{k-1}$ and k equations. If the equations are independent, then we have a unique solution and can recover a_0.

$$M = \begin{bmatrix} 1 & \alpha_{i1} & \alpha_{i1}^2 & \ldots & \alpha_{i1}^{k-1} \\ 1 & \alpha_{i2} & \alpha_{i2}^2 & \ldots & \alpha_{i2}^{k-1} \\ & \vdots & & & \\ 1 & \alpha_{ik} & \alpha_{ik}^2 & \ldots & \alpha_{ik}^{k-1} \end{bmatrix}$$

We know that M is a Vandermonde matrix and determinant of M is non-zero if α_{ij}'s are distinct which is the case with our construction. So, the equations are independent and we have a unique solution. If we have $\geq k+1$ equations(i.e a set with more than k users) and k unknowns, we have a unique solution and can obtain a_0.

Suppose, we have a set of $k-1$ evaluations. We must show that the secret is not revealed. So, the users

should be able to believe that any element in $\{0, 1, \ldots, q\}$ is equally likely for a_0.

$$\begin{aligned} f(\alpha_{i1}) &= a_0 + a_1\alpha_{i1} + \cdots + a_{k-1}\alpha_{i1}^{k-1} \\ f(\alpha_{i2}) &= a_0 + a_1\alpha_{i2} + \cdots + a_{k-1}\alpha_{i2}^{k-1} \\ &\vdots \\ f(\alpha_{i(k-1)}) &= a_0 + a_1\alpha_{i(k-1)} + \cdots + a_{k-1}\alpha_{i(k-1)}^{k-1} \end{aligned}$$

We have k unknowns $a_0, a_1, \ldots, a_{k-1}$ and $k-1$ equations. As the number of unknowns are more than the number of equations we don't have a unique solution. For every value in $\{0, 1, \ldots, q\}$ for a_0, we have a unique solution and each such solution is equally likely for the group of $k-1$ users. Therefore, they will not be able to reduce the solution space of a_0. Actually the solution space forms a coset and any of the solutions is equally likely when only $k-1$ equations are known. Hence this scheme is perfect.

3 Linear Codes and Access Structure

In the last section we have seen how Reed Solomon codes help us to realize a perfect secret sharing scheme. In this section we will describe how we can also use linear codes for secret sharing schemes and the access structure can be much more arbitrary than regular k-sets. We will show an example and then we will describe the thought process behind it. Here we are going to describe a (S, τ) access structure where secret is transformed into S shares such that

1. knowledge of the shares in any set in τ determines the secret
2. knowledge of the shares in a set not in and having no subset in τ reveals no information whatsoever about the secret.

Let us say that there are 4 users and we want an access structure $\{\{A, B\}, \{B, C, D\}\}$. The codewords of the linear code will describe the shares of the users. Each codeword will be of length 5 since the first element will be the secret. Now let us provide the definition of minimal codewords. We will say that a codeword x is minimal if it satisfies the following two properties

1. x is a non-zero codeword whose leftmost element is 1
2. x covers no other codeword whose leftmost non-zero component is a 1.

Now for the given parity check matrix

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

we have the Generator matrix as

$$G = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

We can easily see that the columns v_1, v_2, v_3 are linearly dependent and $v_2, v_4, v5$ are linearly dependent as well. Hence we have $v_1 + v_2 + v_3 = 0$ and $v_2 + v_4 + v_5 = 0$ which means that for any codeword C of the dual code (codeword generated with H as the generator matrix), $C[2]$ and $C[3]$ can always be used to decipher $C[1]$ and also $C[2], C[4], C[5]$ can be used to decipher $C[1]$ as well. These sets of columns of the generator matrix is called information set. There are no other combinations of the elements of C except a union of the 2 that we have already mentioned that would lead us to decipher $C[1]$ and hence this also agrees with the (S, τ) secret sharing scheme. But this exactly corresponds to our required access structure with $C[1]$ as the secret and hence this linear code works for our secret sharing

scheme. So backtracking, we basically want the access structure to be defined by the minimal codewords of the dual code. From the example access structure we can see that we want our minimal codewords to be $[11100, 10111]$ and H is chosen such that it forms a generator matrix for those codewords. Hence the codewords generated by G will work as the shares of the secret. Let us show this as a formal theorem

Theorem *The access structure of the secret-sharing scheme for $n-1$ users corresponding in the manner just specified to the linear (n, k) code V consists precisely of those share sets corresponding to those minimal codewords in the dual code $V^{\perp}$ having 1 as their first component in the manner that the share set specified by such a minimal codeword contains the share x_i $(2 \leq i \leq n)$if and only if the i^{th} component of this codeword is non-zero and any codeword of V can be the secret and the shares of the secret*

Proof Please see [1] (Page 4, Proof of Proposition 2)

References

[1] Massey, James Lee. Some applications of coding theory in cryptography. Codes and Ciphers: Cryptography and Coding IV (1995): 33-47.