DNA SECRET SHARING

Avishek Adhikari

Department of Pure Mathematics,

Calcutta University,

35, Ballygunge Circular Road,

Kolkata-700019,

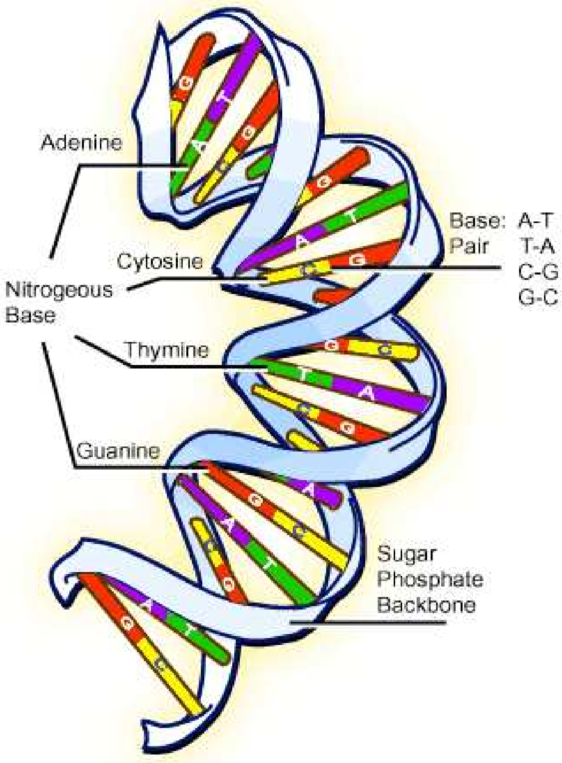
West Bengal, India

E-mail : avishek.adh@gmail.com

Aim of Our Work

Our aim is to distribute a secret binary string using DNA computing to a set $\{P_1, P_2, \ldots, P_n\}$ of n participants in such a way that certain designated set of participants can reveal the secret by pulling their shares, but no forbidden set of participants has any information about the secret binary string. In short, our aim is to sharing a secret using DNA.

Why DNA for Secret Sharing?

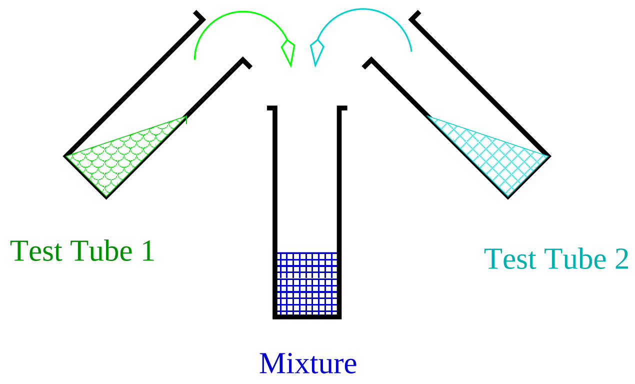


- The very small size,
- The huge storage capacity,
- Easy to carry or hide,
- Made up of A, T, G, C,
- Huge parallel computing,
- Stable as a DNA double strand,
- High longevity,
- Easy to get synthesized DNA

DNA encoding of binary strings

- a binary string can be represented as a set of integers that corresponds the positions where the bits are 1 from left to right.
- 1011 can be represented as a set $\{1, 3, 4\}$,
- each integer i can be represented $ds_i = \updownarrow S_0(GAATTGC^5)^i GAATTCS_1$, where $\updownarrow GATTC$ is the restriction site for EcoRI and S_0 and S_1 be suitable 20 to 30 base pair long DNA strand not containing $\updownarrow GAATTC$.
- if $\alpha = 1011$, then the DNA double strand representation is $T[\alpha] = \{ds_1,\ ds_3,\ ds_4\}$.

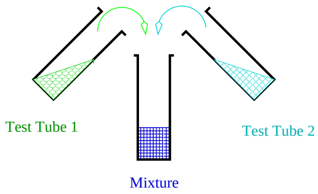
Mixing operation



- Take the content of two test tubes.
- Mixing can be done by dehydrating the tube contents (if not already in solution) and then combining the fluids together into a new tube, by pouring and pumping.

Bio-mathematical operations

- Boolean "or" operation between two binary strings
- if $\alpha = 1011$ and $\beta = 1001$,
- the binary "or" of two strings will be 1011.
- $T[\alpha]$ ($T[\beta]$) the test tube corresponding to the binary string α (β).
- pore the contents of the two test tubes to get binary "or".



Example of DNA Secret Sharing :

- Consider the secret sharing scheme on $\mathcal{P} = \{1, 2, 3, 4\}$ of 4 participants, where $\Gamma_0 = \{\{1, 2\}, \{2, 3\}, \{2, 4\}, \{1, 4\}\}$.
- $\Gamma_{Qual} = \{Y \subseteq \mathcal{P} : X \subseteq Y$ for some $X \in \Gamma_0\}$ and $\Gamma_{Forb} = 2^{\mathcal{P}} \setminus \Gamma_{Qual}$.
- let the secret binary string be $x = x_1x_2x_3 = 011$.
- Assume that the DNA encoding, the mixing procss are public.

Share Distribution :

- The dealer chooses two Boolean matrices
$$G_0 = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } G_1 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$
- Since $x_1 = 0$, the dealer considers the matrix G_0 and apply a random permutation to the columns of G_0 and produces a matrix M_1.
- Similarly, for x_2 and x_3 on G_1 to produce matrices M_2 and M_3.
- Let $M = M_1 || M_2 || M_3$.

Share Distribution :

- first row of M is $\alpha_1 = 010110101010$. Similarly $\alpha_2 = 010101100110$, $\alpha_3 = 010010001000$ and $\alpha_4 = 000100010001$.
- dealer converts them to DNA representations to get $T[\alpha_1] = \{ds_2, ds_4, ds_5, ds_7, ds_9, ds_{11}\}$, $T[\alpha_2] = \{ds_2, ds_4, ds_6, ds_7, ds_{10}, ds_{11}\}$, $T[\alpha_3] = \{ds_2, ds_4, ds_5, ds_9\}, T[\alpha_4] = \{ds_4, ds_8, ds_{12}\}$,
- $T[\alpha_i]$ is given to the participants P_i. Also the values $m = 4$ and $k = 3$ are given to each participants even through an insecure channel.

Dccryption by the Qualified participants

- Let P_1, P_2 come together.
- They use mixing procedure with test tubes $T[\alpha_1]$ and $T[\alpha_2]$ to get $T[\alpha_1] \cup T[\alpha_2] = \{ds_2, ds_4, ds_5, ds_6, ds_7, ds_9, ds_{10}, ds_{11}\}$.
- Execute automated DNA sequencing method to read the DNA double strands.
- With the knowledge of decoding the DNA representation to the binary string, the values of $k = 3$ and $m = 4$, the participants P_1 and P_2 can convert the DNA representation to the binary string $y = 010111101110$.

Dccryption by the Qualified participants

- Since, the value of m is known to the participants, P_1 and P_2 can break y as $y = (0101)(1110)(1110)$.
- Next they will find the value of w as 3 and then they will compute $z = 011$, as $BW(0101) < 3$, $BW(1110) = 3$. Thus P_1 and P_2 can recover the secret 011.

Forbidden set of participants

- Let $Y = \{P_3, P_4\}$ come together.
- They use mixing procedure with test tubes $T[\alpha_3]$ and $T[\alpha_4]$ to get $T[\alpha_1] \cup T[\alpha_2] = \{ds_2, ds_4, ds_5, ds_8, ds_9, ds_{12}\}$.
- they will convert the DNA representation to the binary string $y = (0101)(1001)(1001)$.
- Thus looking at those it is not possible to predict whether they correspond to 0 or 1.

Construction of Generating Matrices

Let us consider the following two associated system of linear equations over the binary field $\mathbf{Z_2}$,

(1) $$A\mathbf{x} = \mathbf{0}$$

(2) $$A\mathbf{x} = \mathbf{1}$$

where, A is a $2 \times n$ known Boolean matrix of rank 2; $\mathbf{x}$ is an $n \times 1$ vector of unknowns; $\mathbf{0}$ and $\mathbf{1}$ are $r \times 1$ vectors of 0's and 1's respectively. Let G_0 (G_1) be an $n \times 2^{n-2}$ Boolean matrix whose columns are all possible solutions of the system (1) ((2)).

Generating Matrices

Lemma 1 *Let* $(\Gamma'_{Qual}, \Gamma'_{Forb})$ *be a strong access structure on a set* $\mathcal{D} = \{1, 2, \ldots, p\}$ *of* p *participants with* $\Gamma'_0 = \{B_i, B_j\}$ *where* $p \leq n$*,* $B_i, B_j \subseteq \mathcal{D}$*,* $|B_i \cup B_j| = p$*,* $|B_i| \geq 2$ *and* $|B_j| \geq 2$*. Then there exists generating matrices* G_0 *and* G_1 *for* $(\Gamma'_{Qual}, \Gamma'_{Forb})$*.*

Generating Matrices

Lemma 2 *Let* G_0^1 *and* G_1^1 $(G_0^2$ *and* $G_1^2)$ *denote the generating matrices of a given access structure* $(\Gamma^1_{Qual}, \Gamma^1_{Forb})$ $((\Gamma^2_{Qual}, \Gamma^2_{Forb}))$ *on the set of participants* $X_1 = \{i_{1_1}, i_{1_2}, \cdots, i_{1_k}\}$ $(X_2 = \{i_{2_1}, i_{2_2}, \cdots, i_{2_s}\})$*. Then there exist generating matrices* G_0 *and* G_1 *for the access structure* $(\Gamma^1_{Qual} \cup \Gamma^2_{Qual}, \Gamma^1_{Forb} \cap \Gamma^2_{Forb})$ *on the set of participants* $X = X_1 \cup X_2$*.*

Main Theorem

Theorem 1 *Let* $(\Gamma_{Qual}, \Gamma_{Forb})$ *be a strong access structure on a set* $\mathcal{P} = \{1, 2, \ldots, n\}$ *of* n *participants with* $\Gamma_0 = \{B_1, B_2, \ldots, B_k\}$ *where* $B_i \subseteq \mathcal{P}, \forall\, i = 1, 2, \ldots, k$*. Let* σ *be a permutation on* $\{1, 2, \ldots, n\}$*. Then there exists generating matrices* G_0 *and* G_1 *for the access structure* $(\Gamma_{Qual}, \Gamma_{Forb})$ *on* $\mathcal{P}$*.*

Conclusion

- It is a perfectly secure scheme.
- It has low error rate and it is easy to implement.
- Huge amount of data can be shared secretly.
- the “or" operation can be carried out very efficiently.
- Finally, it is suitable for secret agents and spies.
- The results in DNA-cryptography are very few and demand attention from the researchers.