Secure Group Authentication Using a Non-perfect Secret Sharing Scheme based on Controlled Mixing

Kannan Karthik
Electronics and Communication Engineering Dept.
Indian Institute of Technology Guwahati
Guwahati, Assam 781039
Email: k.karthik@iitg.ernet.in

Dimitrios Hatzinakos
Electrical and Computer Engineering Dept.
University of Toronto
Toronto, Ontario, M5S3G4
Email: dimitris@comm.utoronto.ca

***Abstract*—In perfect secret sharing schemes, invalid coalitions of shares do not reveal any information regarding the secret. However, in secret sharing constructions which are non-perfect in nature, every illegitimate coalition leaks out information not just pertaining to the parent secret but also related to other shares in circulation. In this paper the information flow and leakage through shares generated by a non-perfect scheme called MIX-SPLIT is examined with the help of a geometric model, developed with a 3-out-of-3 sharing example in view. Finally the principle of controlled leakage has been applied towards a secure biometric-based group authentication system. No fingerprint template storage is required and access to the respective user biometric PINs are made completely secure.**

I. Introduction

Traditionally shared access control schemes [1], [2] have been used to restrict access to top secret documents such as government records and sensitive areas such as document centres, to only selective groups of individuals. These shared access schemes are usually implemented with the threshold perfect secret sharing (n, n) or (k, n) constructions [1], [2]. In an (n, n) construction, a particular secret K, is split into n different shares, $S_1, S_2, ..., S_n$ using for example the polynomial sampling/reconstruction approach of Shamir's [1], where every share represents one of n distinct sample points $S_i :: [x_i, f(x_i)]$ derived from a polynomial, $f(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + ... + a_{n-1}x^{n-1}$. The coefficient a_0 is the secret K. To retrieve the secret all n points must be put together and a polynomial must be interpolated through them, in the process extracting the buried constant $a_0 = K$. If less than n shares are provided, absolutely no information pertaining to K is disclosed. Hence the scheme is perfect. The set, $A = \{S_1, S_2, ..., S_n\}$ is called an access set, while, any subset of A, i.e. $B_t = \{S_{i_1}, S_{i_2}, .., S_{i_t}\}$ with $t < n$ is termed as an invalid coalition. In the information theoretic sense for perfect (n, n) schemes, the following conditional entropies hold: $H(K/A) = 0$, $H(K/B_t) = H(K)$ and for an arbitrary share $S_i \in A$, $H(S_i/B_t) = H(S_i)$. Invalid coalitions of the type B_t do not leak out any information pertaining to either the original secret K or other subsets of shares.

A non-perfect scheme in contrast consists of three different types of sets of shares: (i) Access sets which constitute subsets of users who have full information about the secret, (ii) Non-access sets in which users do not have any information about the secret and (iii) Partial access sets which include sets of users who possess partial information about the secret. If B is a partial access set, B will have partial information on K but cannot recover K and the conditional entropy $H(K/B)$ takes on values between 0 and $H(K)$. Non-perfect secret sharing is still very much an open area of research with several unanswered questions. We raise a few.

- How does one utilize the information leaking out of the partial access sets?
- If every partial access set is forced to reveal a different portion of the secret, implying a unique group inheritance, can one use this feature for group authentication?

A non-perfect secret sharing algorithm called MIX-SPLIT [3] is presented in Section II. A geometric view of MIX-SPLIT focussing on information leakages from invalid coalitions of shares is introduced in Section III. Further, it is also shown that collusion of any two shares S_i, S_j in a $(3, 3)$ scheme have a tendency to reveal a unique component of the parent secret K. This principle has been applied towards secure group authentication in Section IV.

II. MIX-SPLIT

Let $\vec{X} = [x_1, x_2, .., x_L]$ and $\vec{Y} = [y_1, y_2, .., y_L]$ represent two statistically similar but independent discrete binary sequences, with $x_i, y_j \in \{0, 1\}$. Furthermore we assume that the respective sequences comprise of independent and identically distributed (I.I.D.) random variables with $\Pr(x_i = 1) = \Pr(y_j = 1) = p$. Let $S = [s_1, s_2, .., s_L]$ denote a mixture sequence constructed such that,

$$\begin{aligned} s_i &= x_i \text{ with Prob. } \alpha \\ &= y_i \text{ with Prob. } 1 - \alpha \end{aligned} \tag{1}$$

Note the following about the new sequence S,

- The sequence S is statistically similar to both parent sequences $\vec{X}$ and $\vec{Y}$. Given a specific realization of S, it is impossible to separate the $\vec{X}$ and $\vec{Y}$ components convincingly. Thus S can be treated as a secure derived share of $\vec{X}, \vec{Y}$.
- Such mixtures can be constructed by random partitioning, fragmentation and then fusion of the parent sequences. Let $P = \{1, 2, 3, .., L\}$ be a set of all possible bit

positions in the original parent sequences $\vec{X}$ and $\vec{Y}$. Let $P_A \subset P$ and $P_B \subset P$ be two disjoint partitions of P, i.e. $P_A \cup P_B = P$ and $P_A \cap P_B = \phi$. The partition lengths are chosen such that $|P_A|/L \approx \alpha$. The child string S is a concatenation and mixture of the pieces $\vec{X}(P_A)$ and $\vec{Y}(P_B)$, the child bits retaining their original parent positions.

- Several controlled and correlated mixtures can be constructed using codebooks such as $C_{3,3}$ and $C_{5,5}$ [4].

$$C_{3,3} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \tag{2}$$

In the above equation which represents the $C_{3,3}$ codebook, '1' represents an affiliation with the $\vec{X}$ sequence and '0' an affiliation with the $\vec{Y}$ sequence. Every row represents an abstraction of one of three shares of the parent secrets $\vec{X}, \vec{Y}$, for example,

$$S_1 = \text{MIX}\left[\vec{X}(P_1), \vec{X}(P_2), \vec{Y}(P_3)\right] \tag{3}$$

abstracted as $S_1 \equiv [1, 1, 0]$, where P_1, P_2, P_3 are random partitions of P. Closure is effected when all three shares are either fused by Majority (MAJ_VOTE) or Minority (MIN_VOTE) bit votes. The share generation and secret retrieval (closure) portions of MIX-SPLIT are shown in Algorithm. 1 and Algorithm. 2 respectively.

```
input  : Codebook $C_{3,3}$, Sequences $\vec{X}, \vec{Y}$)
         Permutation keys $K_{p1}, K_{p2}, K_{p3}$
output: Shares $S_1, S_2, S_3$
Pos ← {1,2,..,L};
for $j$ ← 1 to 3 do
    Pos ← Permute(Pos, $K_{pi}$);
    $P_j$ ← Pos(1:L/3);
    Pos ← SetDifference(Pos, $P_j$);
end
for $i$ ← 1 to 3 do
    for $j$ ← 1 to 3 do
        if $C_{3,3}(i,j) = 1$ then
            $S_i(P_j) \leftarrow \vec{X}(P_j)$;
            else if $C_{3,3}(i,j) = 0$ then
                $S_i(P_j) \leftarrow \vec{Y}(P_j)$ ;
            endif
        endif
    end
end
```

Algorithm 1: MIX-SPLIT (share generation)

```
input  : Shares $S_1, S_2, S_3$
output: Secrets $\vec{X}, \vec{Y}$
$\vec{X}$ ← MAJ_VOTE($S_1$,$S_2$,$S_3$);
$\vec{Y}$ ← MIN_VOTE($S_1$,$S_2$,$S_3$);
```

Algorithm 2: MIX-SPLIT (retrieval)

III. Geometric Interpretation

A geometric view of the MIX-SPLIT scheme is obtained through the following simplified projections:

(i) Constraining the choice of $\vec{Y}$-sequence by imposing,

$$\vec{Y} = BIT\_CMP[\vec{X}] \tag{4}$$

where, $BIT\_CMP[.]$ is the bit complement of an arbitrary binary string. The loss of independence of $\vec{Y}$ converts the 2-secret sharing scheme to single-secret sharing.

(ii) Restricting the codebook size to 3-out-of-3,

$$C_{3,3} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \tag{5}$$

Three disjoint equal-sized random partitions, P_1, P_2 and P_3 of $\{1, 2, 3, 4, ...L\}$ are created with sizes $L_p = \text{length}(\vec{X})/3 = L/3$.

(iii) Every bit string contained in the partitions is mapped to a point in the interval $[0, 1]$. Let $\vec{X}_j = \vec{X}(P_j) = [x_{j_1}, x_{j_2}, x_{j_3}, ..., x_{j_{L_p}}]$, where $x_{j_k} \in \{0, 1\}$. The point corresponding to the partition P_j is,

$$\begin{aligned} a_j &= \text{String2Point}(\vec{X}_j) \\ &= \sum_{r=1}^{L_p} x_{j_r} \cdot \left[\frac{1}{2}\right]^r \end{aligned} \tag{6}$$

The corollary of (i) and (iii) is,

$$\begin{aligned} \text{IF } \vec{X}(P_j) &\equiv a_j \\ \text{THEN } \vec{Y}(P_j) &\equiv 1 - a_j \end{aligned} \tag{7}$$

Secrecy is examined for the cases: (a) $B_t = B_1 = \{S_i\}$ (only one share available to the traitors), (b) $B_t = B_2 = \{S_i, S_j\}$ (some two out of three shares available) and (c) $B_t = B_3 = A = \{S_1, S_2, S_3\}$ (all three shares available). In all the three cases it is assumed that the random partitions P_1, P_2, P_3 are destroyed as soon the shares are created.

A. Single share

The invisibility of all the partitions are preserved when only one share S_i is available. The absence of no prior information forces the attacker to fabricate arbitrary disjoint sets $\hat{P}_A, \hat{P}_B, \hat{P}_C$ to split S_i into three components $S_A = S_i(P_A)$, $S_B = S_i(P_B)$ and $S_C = S_i(P_C)$ corresponding to the three co-ordinate estimates,

$$\begin{aligned} \hat{c}_A &= \text{String2Point}(S_A) \\ \hat{c}_B &= \text{String2Point}(S_B) \\ \hat{c}_C &= \text{String2Point}(S_C) \end{aligned} \tag{8}$$

where, $\hat{c}_A, \hat{c}_B, \hat{c}_C \in [0, 1]$. The number of ways in which partitions P_A, P_B, P_B can be chosen is,

$$N_{parts} = \binom{L}{L_p} \times \binom{2L_p}{L_p} \tag{9}$$

where, $L = 3L_p$. Given two different partition selections for the first co-ordinate, P_A^1 and P_A^2, the probability that the two extracted strings $S_i(P_A^1) \equiv \hat{c}_A^1$ and $S_i(P_A^2) \equiv \hat{c}_A^2$ will point to the same estimate $\hat{c}_A$ is,

$$\Pr(\hat{c}_A^1 = \hat{c}_A^2) = [1 - 2 \cdot p \cdot (1 - p)]^{L_p} \tag{10}$$

which becomes $(\frac{1}{2})^{L_p}$ when $p = 0.5$. When applied to data such as compressed binary images typical partition lengths L_p exceed 1000 bits, hence, the chances of overlap are virtually NIL. This consequently implies that almost all N_{parts} are distinct and since the realization is derived from an I.I.D. sequence, the estimates (of say co-ordinate 1) $\hat{c}_A^1, \hat{c}_A^2, \hat{c}_A^3, ..., \hat{c}_A^{Nparts}$ are uniformly spread over the interval [0,1]. The same argument holds for the other estimates $\hat{c}_B$ and $\hat{c}_C$. Hence, the secret can be any point within a *unit cube*. This is reflected in Fig. 1(a).

B. Two shares

When two shares S_i, S_j are fused, exactly one of the partitions P_{vis} becomes visible corresponding to the positions r_{eq}, where,

$$P_{vis} \quad = \quad \{r_{eq}\}, \text{ s.t. } S_i(r_{eq}) = S_j(r_{eq}) \qquad (11)$$

i.e. P_{vis} is X-dominant as the position r_{eq} can be clearly linked to sequence $\vec{X}$ (refer to any two rows in the $C_{3,3}$ codebook and examine the column with common ones - Eqn. 5), $P_{vis} = P_{XX}$. The remaining two partitions will be hidden owing to the chequered pattern in the codebook. For instance if the two available shares were S_1 and S_2, the corresponding code-representation would be,

$$\begin{bmatrix} S_1 \\ S_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \qquad (12)$$

Hence, the set of positions, P_{XY},

$$P_{XY} \quad = \quad \{r_{neq}\}, \text{ s.t. } S_i(r_{neq}) \neq S_j(r_{neq}) \qquad (13)$$

will leave an ambiguity pertaining to the affiliation of a particular position r_{neq} to either $\vec{X}$ or $\vec{Y}$. Let $S_{col(i,j)}$ be a collusion of S_i, S_j constructed in the following fashion:

- Extraction of the visible co-ordinate: $S_{col(i,j)}(P_{XX}) = S_i(P_{XX}) \equiv a_{vis}$. The point a_{vis} can be any one amongst a_1, a_2, a_3.
- Estimating the diffused co-ordinates: Split the set P_{XY} into two disjoint equal-sized sets P_{XYX} and P_{XYY}. The number of ways in which this split can be executed is,

  $$N_{XY} \approx \binom{2L_p}{L_p} \qquad (14)$$

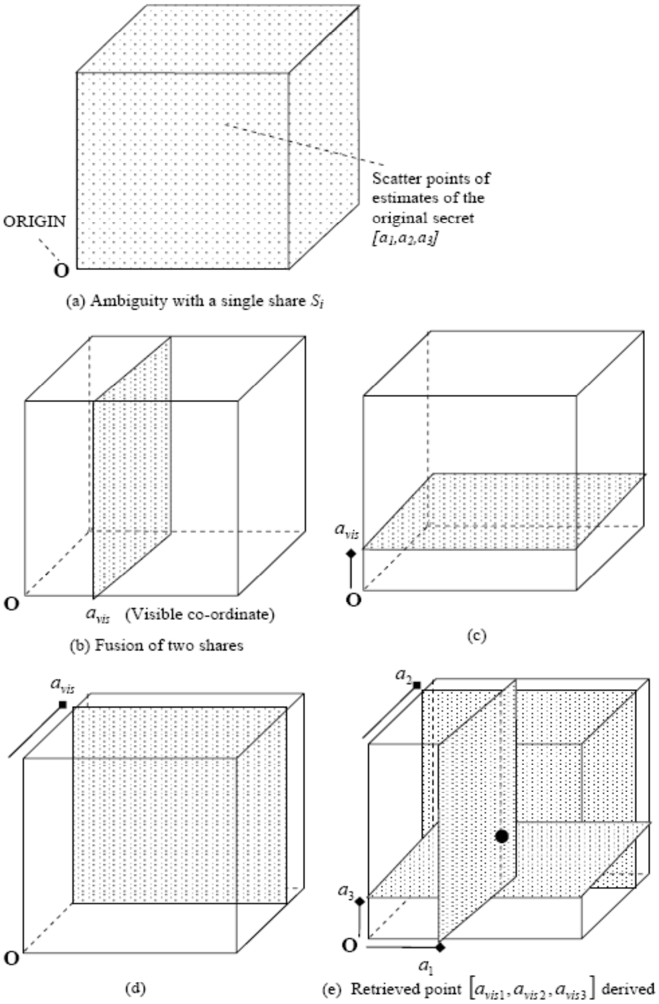
  The second component of the share $S_{col(i,j)}(P_{XY})$ is obtained by simply lifting the first segment from share S_i, i.e. $S_i(P_{XYX})$ and fusing it with the version from S_j, i.e. $S_j(P_{XYY})$.

  $$S_{col(i,j)}(P_{XY}) = S_i(P_{XYX}) || S_j(P_{XYY}) \qquad (15)$$

- Estimate of secret $\vec{X}$ is,

  $$\hat{X} = S_i(P_{XX}) || S_i(P_{XYX}) || S_j(P_{XYY}) \qquad (16)$$

Note in the second step there are N_{XY} ways in which the 'split + fuse' can be performed. This leaves N_{XY} choices for the estimates $[\hat{a}_i, \hat{a}_j]$ uniformly scattered over a UNIT-plane intersecting the co-ordinate axes at one point. Based on which two shares $[S_i, S_j] \subset \{S_1, S_2, S_3\}$ are combined, exactly one coordinate is revealed $a_{vis} \in \{a_1, a_2, a_3\}$. This is illustrated in Fig. 1(b,c,d).


Scatter points of estimates of the original secret $[a_1,a_2,a_3]$
ORIGIN
O
a_{vis} (Visible co-ordinate)
O
a_{vis}
O
a_{vis}
O
a_2
a_3
O
a_1

(a) Ambiguity with a single share S_i

(b) Fusion of two shares

(c)

(d)

(e) Retrieved point $[a_{vis1}, a_{vis2}, a_{vis3}]$ derived by fusing three shares $[S_i, S_j, S_k]$.

Fig. 1. Geometric interpretation of MIX-SPLIT. (a) With a single share the region of uncertainty is the entire unit cube, (b,c,d) Fusion of any two shares reveals exactly one co-ordinate (region of uncertainty is a plane), (c) Fusion of all three shares by majority vote reveals the secret $[a_1, a_2, a_3]$.

C. All three shares

The three partitions become fully visible with the following comparisons across the shares $\{S_i, S_j, S_k\}$:

$$\begin{aligned} P_{visA} &= \{r_A\} \text{ s.t. } [S_i(r_A) = S_j(r_A)] \neq S_k(r_A) \\ P_{visB} &= \{r_B\} \text{ s.t. } S_i(r_B) \neq [S_j(r_B) = S_k(r_B)] \\ P_{visC} &= \{r_C\} \text{ s.t. } [S_i(r_C) = S_k(r_C)] \neq S_j(r_C) \end{aligned} \qquad (17)$$

The corresponding points are:

$$\begin{aligned} a_{vis1} &= \text{String2Point}[S_i(P_{visA})] \\ a_{vis2} &= \text{String2Point}[S_j(P_{visB})] \\ a_{vis3} &= \text{String2Point}[S_k(P_{visC})] \end{aligned} \qquad (18)$$

The triplet $\hat{T} = [a_{vis1}, a_{vis2}, a_{vis3}]$ is none other than the ordered set $[a_1, a_2, a_3]$ or a permutation of it and thus the original secret $\vec{X} \equiv [a_1, a_2, a_3]$ can be fully recovered as shown in (Fig. 1(e)). It is the point of intersection of the three planes.

IV. Secure group authentication

Consider a group access control and authentication application in which only individuals in groups of two or more are allowed to enter a restricted area within an organization. The entire user space is partitioned into groups of three users (the number of members in a particular mini-project). Atleast two users are required to be present during the project work in the organization at any given time. Hence, single user entry is denied. Access is granted based on a biometric PIN verification mechanism. A biometric PIN is essentially a compressed biometric template of a particular user whose loss does not result in the compromise of the template itself, mainly because the hash compression is lossy and of one way type [5], [6].

A. Registration

Three users u_1, u_2, u_3 submit their biometric templates (e.g. fingerprints) $Temp(u_1), Temp(u_2)$ and $Temp(u_3)$ respectively to the center. The center computes the corresponding PINs as $\vec{X}_1, \vec{X}_2$ and $\vec{X}_3$ respectively,

$$\vec{X}_i \quad = \quad \text{BioHash}(Temp(u_i)) \tag{19}$$

All the three strings $\vec{X}_i, i = 1, 2, 3$ are designed to be of equal length and can be modeled as I.I.D. binary sequences. The center first constructs a group verification sequence,

$$\vec{V}_{123} \quad = \quad \vec{X}_1 \oplus \vec{X}_2 \oplus \vec{X}_3 \tag{20}$$

Note that since all the PINs are I.I.D. sequences, this verification string $\vec{V}_{123}$ is as secure as a one time key pad [7]. Let L_p be the length of each PIN $\vec{X}_i$ and $P = \{1, 2, 3, ..., L\}$, where $L = 3L_p$. A group parent sequence $\vec{X}$ of length $L = 3L_p$ is constructed by *mixing* the three PINS. If P_1, P_2, P_3 are random partitions of P, a one to one correspondence (permutation map) between $\vec{X}$ and $\vec{X}_1, \vec{X}_2, \vec{X}_3$ is established as follows:

$$\begin{aligned} \vec{X}(P_1) &= \vec{X}_1 \\ \vec{X}(P_2) &= \vec{X}_2 \\ \vec{X}(P_3) &= \vec{X}_3 \end{aligned} \tag{21}$$

Geometrically, the three PINs $[\vec{X}_1, \vec{X}_2, \vec{X}_3]$ correspond to the point $[a_1, a_2, a_3]$ inside the *unit cube*, where a_i = String2Point($\vec{X}_i$). As discussed in the geometric model, a constrained $\vec{Y}$ sequence as the bit complement of the $\vec{X}$ sequence is generated and the pair is submitted to the MIX-SPLIT algorithm with a slight modification to the $C_{3,3}$ codebook (Eqn. 22) to produce shares $S_i, i = 1, 2, 3$.

$$C_{3,3(p)} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \tag{22}$$

$$[S_1, S_2, S_3] \quad = \quad \text{MIX-SPLIT}(\vec{X}, \left[\vec{X}\right]^c, C_{3,3(p)}, P_1, P_2, P_3)$$

The generated shares S_i are given to each user u_i through a smart card, while the group verification code $\vec{V}_{123}$ is the only secure data stored at the center.

B. Verification

If a single user arrives (say u_1), the only useful information that can be computed is the user's PIN $\vec{X}_1$. The other two user's PINs are unavailable and so the verification data cannot be created from a single share as a stand alone share remains anonymous. When two users arrive (say u_1, u_3) and submit their biometric templates $Temp(u_1), Temp(u_3)$ and shares S_1, S_3, three strings are released: (1) PIN of user u_1: $\hat{X}_1$ directly from the template $Temp(u_1)$; (2) PIN of user u_3: $\hat{X}_3$ directly from the template $Temp(u_1)$; (3) The fusion of S_1 and S_3 makes the co-ordinate a_2 visible and thus a portion of user u_2's share is revealed, thereby releasing $\hat{X}_2 \equiv a_2$. Now with the virtual presence of the third user, a verification code can be constructed as

$$\hat{V} \quad = \quad \hat{X}_1 \oplus \hat{X}_2 \oplus \hat{X}_3 \tag{23}$$

Now $\hat{V}$ is compared with the stored code $\vec{V}_{123}$ for authentication. Note that if the PIN generation is stable, then any two out of the three users can pass the authentication process. The biometric PIN storage is completely secure as the PINs are stored in the encrypted form. The 2-out-of-3 share fusion reveals the exact subgroup trying to gain entry (subgroup inheritance), which reinforces the PIN authentication process.

V. Conclusions

In non-perfect secret sharing schemes like MIX-SPLIT, information leakages from invalid coalitions can be controlled and used to form unique associations with the parent secret. As per the geometric model for a $(3, 3)$ sharing example, the secret is some point within an unit cube $\vec{X} \equiv [a_1, a_2, a_3]$. With a single share, there is virtually no information regarding the position of the secret. With two shares, exactly one co-ordinate is made visible confining the secret's position to a plane. Thus, every subset of two shares has a unique group inheritance. This has been applied towards secure group authentication, which does not require biometric template storage and the verification process is implemented with the help of encrypted PINs.

References

[1] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, pp. 612–613, 1979.

[2] G. Blakley, "Safeguarding cryptographic keys," in *Proceedings of AFIPS 1979 National Computer Conference*, vol. 48, pp. 313–317, 1979.

[3] K. Karthik and D. Hatzinakos, "Multimedia Encoding for Access Control with Traitor Tracing: Balancing Secrecy, Privacy and Traceability," *VDM Verlag Dr. Muller*, 2008, ISBN: 978-3-8364-3638-0.

[4] K. Karthik and D. Hatzinakos, "A unified approach to construct non-perfect secret sharing and traitor tracing schemes," in *Int. Conference on Security and Management*, pp. 83–89, Las Vegas, Nevada, 2007.

[5] A. S. R. C. Soutar, D. Roberge and B. V. Kumar, "Biometric encryption using image processing," vol. 3314, pp. 178–188, SPIE, 1998.

[6] S. P. U. Uludag, S. Pankanti and A. K. Jain, "Biometric cryptosystems: issues and challenges," *Proceedings of the IEEE*, vol. 92, no. 6, pp. 948–960, 2004.

[7] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, Oct 1949.