Algebraic Structures and its Applications in Cryptography

Dr. Sucheta Chakrabarti
Scientist - G
Scientific Analysis Group
DRDO
Delhi

E-mail – suchetadrdo@hotmail.com

Outline of the Presentation

- **Secure Communication & Cryptography**
- **Role of probability and entropy in secure communication from information theoretic approach**
- **Commonly used Algebraic Structures in Cryptography**
- **New Direction of (in) Cryptography based on Non-commutative / Non-associative Algebraic Structures**
- **Quaigroups**
- **Quasigroup –Based Transformations and its cryptographic applications**

Secure Communications Over Open Channels

Aim :

- To Protect Information
- To coordinate operations (*command and control*)
- To carry out online business transaction (*E- commerce*)

Service required for secure communication –

- ***Data confidentiality*** : It ensures the privacy of data i.e only the authorized person can only access the information

- ***Data Integrity*** : It ensures the protection from any unauthorized alteration i.e. no insertion, deletion or modification has been done in the information by Non-legitimate party .It provides the assurance that the data is present in its original form as it was sent by the sender.

- ***Data availability*** : This means that the data is always available for access whenever required

- ***Authentication :*** This ensures that the communication is being held among the right individuals.

- ***Non-repudiation :*** According to this, the sender or the receiver cannot deny being responsible for the data being transmitted.

Fundamental building block of security is Cryptography

1949 is the turning point for cryptography *– it turns to scientific based on mathematical grounds by the research article* ***Communication Theory of secrecy system*** *- C.E. Shannon*

- **Security needs continuous improvement / up gradation against adversary capabilities viz.**

(i) computational

- ***Computationally unbounded*** – Unconditional security (Info. theoretical or perfect secrecy)
- ***Computationally bounded*** – Computational security & Provable security (*the cryptographic primitive reduced to certain problem which is proved to be (well known)hard problem . It implies breaking of the primitive computationally infeasible*)

(ii) other capabilities -

- **Active -** can corrupt parties, inject / modify messages
- **Passive / eavesdropper** – only listens (intercepts) messages
- **Other resources** i.e. ability to decrypt some messages.

- **Security is based on Arbitrary Adversary Principle (AAP) –** *i.e it assume restrictions on adversary capabilities , but not that the adversary is using specific strategies or attacks*
- ***Secure electronic identities and information protection*** *are key for digital evolution*

In the Modern digital world

Cryptography (Crypto-primitives / algorithms) deals with information security & secure communications over insecure channels.

Mainly deals with **Confidentiality , Authenticity , Integrity & Non-repudiation**

It needs set of elements and specific operations that are applied to the elements of the set is called Algebraic Structures

Basic Components of Cryptography

- **Functions**

- **one – one**
- **one-way**
- **trapdoor one way**
- **encryption / decryption**

Encryption/Decryption function has to satisfy the following condition :

For $E \in \mathcal{E}$ and $k_e \equiv e \in \mathcal{K}$, $E_e : \mathcal{M} \rightarrow \mathcal{C}$ is a 1-1 mapping

& so there exists a corresponding $D \in \mathcal{D}$ and $k_d \equiv d \in \mathcal{K}$ such that

$D_d : \mathcal{C} \rightarrow \mathcal{M}$ and $D_d\big(E_e(m)\big) = m \ \ for\ all\ \ m\ \in \mathcal{M}$

In other words

Cryptographic Algorithms - consist of $\mathcal{M}$, $\mathcal{C}$, $\mathcal{K}$ and set

$\{E_e, e \in \mathcal{K}\}$ of encryption transformations and corresponding set

$\{D_d, d \in \mathcal{K}\}$ of decryption transformations with the property that for each $e \in \mathcal{K}$ there exists a unique , $d \in \mathcal{K}$ s.t $D_d \equiv E_e^{-1}$ i.e

$$D_d\big(E_e(m)\big) = m \ \ for\ all\ \ m\ \in \mathcal{M}$$

Domain & Codomain of Encryption / Decryption Functions

- Alphabet set - $\mathcal{A}$
- Message space - $\mathcal{M}$
- Crypt space - $\mathcal{C}$
- Key space - $\mathcal{K}$

Set of encryption and decryption functions are denoted by

$\mathcal{E}$ & $\mathcal{D}$ respectively

Cryptosystems

Three Sets : Message / Plaintext $-\ \mathcal{M}$

Ciphertext - $\mathcal{C}$

Keys - $\mathcal{K}$

Three randomized algorithms : $\langle KG, E, D \rangle$

Key generation Algo $KG\colon S^* \to \mathcal{K}$

Encryption Algo $E\colon \mathcal{K} \times \mathcal{M} \to \mathcal{C}$

Decryption Algo $D : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$

For any key $k \in \mathcal{K}$ and $m \in \mathcal{M}$ holds $D_k\big(E_k(m)\big) = m$

So a cryptosystem consists of five tuples which represent as $(\mathcal{M}, \mathcal{C}, \mathcal{K}, E, D)$

Probability & Entropy Concepts for Secure Communication

- The concept of entropy has evolved in probability theory to create information theoretical model for secure communication .
- In 1947-48 by classic work of C. Shannon gives birth of Information theory , a new branch in applied probability theory to handle practical problem of communication.
- Security generally expressed in terms of probability and amount of information (entropy)
- Here we will discuss some important concepts of discrete probabilities

Probability Space : $(\mathbb{X}, Pr)$, where

- $\mathbb{X}-$ the sample space which is a finite set of possible outcomes (events)
- Pr – a function from $\mathcal{P}(\mathbb{X}) \rightarrow [0,1]$ such that

$$(i) Pr(\mathbb{X}) = 1, (ii) Pr(\Phi) = 0, (iii) Pr(X \cup Y) = Pr(X) + Pr(Y) \text{ if } X \cap Y = \Phi$$

$$(iv)\ Pr(X \cap Y) = Pr(X)Pr(Y) \text{ if } X \cap Y = \Phi$$

Pr is called a ***probability distribution*** , a *probability measure* or just a *probability*

Pr *of* $X \in \mathcal{P}(\mathbb{X})$ *determined by* $Pr(\{x\}) \forall\ x \in X$

Joint Probabilities : Two probability spaces viz. $(\mathbb{X}, Pr_1)$ $(\mathbb{Y}, Pr_2)$

It can create joint probability space $(\mathbb{X} \times \mathbb{Y}, Pr)$ where Pr define as follows:

$$Pr(\{x, y\}) = Pr_1(\{x\})\, Pr_2(\{y\})$$

Conditional Probability

- $Pr(X|Y) = Pr(X \cap Y)/Pr(Y)$ - only defined if $Pr(Y) > 0$

- X and Y are independent if $Pr[X = x|Y = y] = \Pr[X = x]$ or $Pr[x|\, y] = Pr[x]$

  & also $Pr[X = x \cap Y = y] = \Pr[X = x] \Pr[Y = y]$ $\forall x, y$

Bayes Theorem : $Pr(X|Y) = \frac{Pr(X)Pr(Y|X)}{Pr(Y)}$

Random Variables

- A **random variable** X is a function from underlying set of probability space (*all possible outcomes* $\mathbb{X}$ **)** to some set of values (*some set of* $\mathcal{P}(\mathbb{X})$)

- Given a probability space and a random variable X, the probability that the random variable X takes value x is $Pr(\{w|X(w) = x\})$

Application to Cryptography for security analysis

Plaintext Distribution :

- X discrete random variable over the plaintext set $\mathcal{M}$
- Sender choose x from $\mathcal{M}$ based on some probability distribution
  - Let $\Pr[X = x]$ be the probability that x is chosen
  - This probability may depend on the language

Key Distribution:

Sender & Receiver agree on a key k chosen from a key set $\mathcal{K}$

- K discrete random variable over $\mathcal{K}$
- $\Pr[K = k]$,the probability that $key\ is\ \ k$

Note that here Probability space (Plaintext , Key)

Ciphertext Probability Distribution

Y is a discrete random variable over the set $\mathcal{C}$

The probability of obtaining a particular ciphertext y depends on the probability of Plaintext and key -

$$Pr[y] = \sum_{x,k|e_k(x)=y} Pr(x)\, Pr(k) = \sum_k Pr(k)\, Pr(d_k(y))$$

- *Attacker Aims to determine the plaintext* x
- *Attacker's does not know /observe ciphertext* y
  - *Probability (a* ***priori probability*** *) that the plaintext is* $x : Pr(X = x) \equiv \Pr(x)$
  - *It depends on plaintext distribution i.e language characteristics*
- *Attacker' s knows / observes ciphertext* y
  - *Probability (a* ***posteriori probability****)that the plaintext is* $x-$

$$Pr(X = x|Y = y) \equiv Pr[x|y]$$

Computation of attacker's *a posterior (conditional) probabilities*

- *Apply Bayes theorem*

$$Pr(X = x|Y = y) \equiv P\,r[x|y]$$

$$= \frac{Pr[x] \times Pr[y|x]}{Pr[y]}$$

Here $Pr[x]$- Probability of the plaintext

$Pr[y]$ - Probability of this ciphertext –It ***induced by probability of plaintext and key distributions***

$$Pr[y] = \sum_{x,k|e_k(x)=y} Pr[x]Pr\,[k]$$

$Pr[y|x]$ - probability that the y is obtained for a given x depends on the keys which provide such a mapping from plaintext domain (*Message spac*e) to ciphertext domain (*Cipher space*) -

$$Pr[y|x] = \sum_{k|e_k(x)=y\ or d_k(y)=x} Pr[k]$$

Example : A Cryptosystem is given below :

$\mathcal{M} - Message\ Space\ \{a, b, c\}$, $\mathcal{K} - Key\ Space\ \{k_1, k_2\}$ &

$\mathcal{C} - Crypt\ Space\ \{P, Q, R\}$ Plaintext Distribution

Plaintext Probability - $Pr[a] = \frac{1}{2}, Pr[b] = \frac{1}{3}, Pr[c] = \frac{1}{6}$

Key Probability - $Pr[k_1] = \frac{3}{4}, Pr[k_2] = \frac{1}{4}$

Encryption (mapping) under the keys :

$$e_{k_1}(a) = R, \qquad e_{k_1}(b) = Q, \qquad e_{k_1}(c) = P$$

$$e_{k_2}(a) = Q, \qquad e_{k_2}(b) = R, \qquad e_{k_2}(c) = P$$

Attackers knowing the system and plaintext & key probabilities can compute $Pr[y]$

$$\Rightarrow Pr[P] = \sum_{x,k|e_k(x)=y} Pr[x]Pr[k] = Pr[c]Pr[k_1] + Pr[c]Pr[k_2] = \frac{1}{6}\times\frac{3}{4} + \frac{1}{6}\times\frac{1}{4} = \frac{1}{6}$$

$$Pr[Q] = \frac{1}{3}\times\frac{3}{4} + \frac{1}{2}\times\frac{1}{4} = \frac{3}{8}$$

$$Pr[R] = \frac{1}{2}\times\frac{3}{4} + \frac{1}{3}\times\frac{1}{4} = \frac{11}{24}$$

$Pr[y|x]$, *i.e* $Pr[P|a] = 0$, $Pr[P|b] = 0$, $Pr[P|c] = Pr[k_1] + Pr[k_2] = 1$

$$Pr[Q|a] = \frac{1}{4}, \quad Pr[Q|b] = \frac{3}{4}, \quad Pr[Q|c] = 0,$$

$$Pr[R|a] = \frac{3}{4}, \quad Pr[R|b] = \frac{1}{4}, \quad Pr[R|c] = 0$$

$\Rightarrow$ **Posterior probability** $Pr[a|P] = 0, Pr[a|Q] = \frac{1}{3}, Pr[a|R] = \frac{9}{11}, Pr[b|P] = 0,$

$$Pr[b|P] = 0, Pr[b|Q] = \frac{2}{3}, \quad Pr[b|R] = \frac{2}{11},$$

$$Pr[c|P] = 1, Pr[c|Q] = 0, Pr[c|R] = 0$$

- Attacker if observes **ciphertext** *P* then he knows that the **plaintext is exactly** *c*
- Attacker if observes **ciphertext** *R* then he knows the **most probable plaintext is** *a*
- *The cryptosystem* ***not providing strong security***

To provide **perfect secrecy**, the cryptosystem has to satisfies the following condition

$$Pr(X = x) \equiv Pr[x] = Pr(X = x|Y = y) \equiv Pr[x|y] \ \forall x, y$$

i.e. the probability that the plaintext is x given that you have observed ciphertext y is the same as the probability that the plaintext is x without observing the ciphertext

In other words, **a priori probabilities = a posteriori probabilities** . It means attacker can not get any knowledge from the ciphertext about the plaintext / key

- Note that in case of perfect secrecy follows

- $Pr[y|x] = Pr[y]$
- $\forall\, x_1, x_2\ Pr[y|x_1] = Pr[y|x_2]$

Perfect secrecy has nothing to do with plaintext distribution
Crypto scheme achieve perfect secrecy without having any dependency on the PT language

A cryptosystem $(\mathcal{M}, \mathcal{C}, \mathcal{K}, E, D)$ with $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$ provides perfect secrecy iff

(i) All keys have the same probability $1/|\mathcal{K}|$ and

(ii) $\forall x \in \mathcal{M}\ \forall y \in \mathcal{C}, \exists a$ *unique key* $k \in \mathcal{K} | e_k(x) = y,$

Example –

1. One Time Pad (OTP)
2. The shift cipher where all keys have probability $1/|\mathcal{K}|$ also provides perfect secrecy if key used only once

Limitation

- Key must be at least as long as the message
- key must be changed for every time encryption
- Arise key distribution & management problems

Main question arises can we find as close as perfectly secure (practically secure) cryptosystems based on short key ?

This motivates the design of Modern cryptosystems which are computationally secure

Entropy & Secrecy of communication system

- **Entropy** - The measure of ***uncertainty*** about occurrence of any event which quantify the ***amount of information*** is given by the occurrence of that event
- Its units commonly in bits (digital communication)
- Introduced by Claude Shannon in 1948
- Built foundation of information theory
- Backbone for the digital era

Let X be a random discrete variable taking values (symbols) from the set (source) $\{x_1, x_2, \cdots, x_n\}$ associated with probabilities of occurrence of symbols $\{p_1, p_2, \cdots, p_n\}$

Information gained by observing event , say x occurred with probability p

$$= log_2 \frac{1}{p_i} = -log_2 p_i \; bits$$

Note that the amount of ***information*** we receive by observing an event occurred is ***inversely proportional to the probability of the event***

Entropy – Let X be a random discrete variable taking values (symbols) from the set (source) $\{x_1, x_2, \cdots, x_n\}$ associated with probabilities of occurrence of symbols $\{p_1, p_2, \cdots, p_n\}$

The ***entropy*** (weighted average of information) of the source , denoted by $H(X)$ *or* H which is defined as follows

$$H(X) \equiv H(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} p_i log_2 \frac{1}{p_i} = -\sum_{i=1}^{n} p_i log_2 p_i$$

We use the convention that $0 \log 0 = 0$

Note that if X *takes one value with probability* 1 *and other values with probability* 0 *then the entropy is* 0*. It clearly tells that there is* ***no uncertainty*** *since we know exactly what value* X *will take*

Note that $H(X)$ can be interpreted as follows:

- Expected amount of information from occurring of X
- Uncertainty about the outcome of X
- Expected (average) number of bits needed to represent an outcome of X

$H(X)$ has the following important property

$$0 \leq H(X) \leq log_2 n$$

When $p_1 = p_2 = \cdots p_n = 1/n$ then $H(X) = log_2 n$

➢ **One of the important application of entropy lies in source coding**

Since $H(X)$ represents the average number of bits of information per symbol from the source – It leads to ….the expectation that H bits per symbol is needed for encoding which can be uniquely decodable. Shannon in 1948 discovered this famous source coding theorem

Source Coding Theorem

(i) The average number of bits / symbol of any uniquely decodable source must be greater than or equal to the entropy H of the source

(ii) If the string of symbols is sufficiently large, there exists a uniquely decodable code for the source such that the average number of bits / symbol of the code as close to H as desired

So entropy is the bench mark for source coding. It has a great operational significance

Huffman Code (Variable length code) -

Design based on the principle : Assigned more bits to least probable events & less bits to frequent events .

It satisfies $H \leq average\ length\ of\ Huffman\ code \leq H + 1$

➢ Measuring the redundancy in a Language

Let $\mathcal{A}$ be the alphabet set of a language and $|\mathcal{A}| = N$

The maximum entropy per alphabet character considered in a language $R = log_2\ N -$

known as ***rate of the alphabet*** *(* ***absolute rate of the language****)*

Let $\mathcal{M}^n = \mathcal{A} \times \cdots \times \mathcal{A}$ *(n times) represents a set of messages of length n*

Let $\boldsymbol{M}$ *be a random variable in* $\mathcal{M}^n$

$$H(\boldsymbol{M}) = -\sum_{\boldsymbol{m}\in\mathcal{M}^n} p(\boldsymbol{m})\, log_2 p(\boldsymbol{m})$$

The entropy (average information) of the message source per alphabet symbol is

denoted by r_n and given by the rate of $\boldsymbol{M}$ as $r_n = \frac{H(\boldsymbol{M})}{n}$

Redundancy of a source (language) - Denoted it by D and defined as follows

$$D = R - r_n$$

Redundancy in English Language

Alphabet set in English languge - $|\mathcal{A}| = 26$

Absolute rate : $R = log_2\ 26 \approx 4.7\ bits\ per\ alphabet$

Entropy per alphabet – (experimentally) $r_\propto = \lim_{n \to \propto} \frac{H(\boldsymbol{M})}{n} \approx 1.5$

Redundancy of a source of the language is denoted by D and given as follows

$$D = R - r_n$$

For English when $n = 1,\ D \approx 4.7 - 1.5 \approx 3.2$

This shows that per alphabet redundancy in Eng around 70%

- *as message size increases rate reduces (infer less information) & hence redundancy increase*
- *It shows representation can be optimized*

➢ *Shannon showed in his one of the famous results , that due to redundancy in a source cryptosystem can be broken / it helps in cryptanalysis*

Joint Entropy & Conditional Entropy

Let X & Y be two discrete random variables and $p(x,y)$ *the value* of the joint probability distribution when $X = x$ & $Y = y$

Joint Entropy is given by

$$H(X,Y) = -\sum_y \sum_x p(x,y)\, log_2 p(x,y)$$

It is the average uncertainty of 2 random variables

Conditional Entropy is given by

$$H(X|Y) = -\sum_y p(y) \sum_x p(x|y)\, log_2\big(p(x|y)\big)$$

$$= -\sum_y \sum_x p(x,y)\, log_2\big(p(x|y)\big)$$

It gives the remaining uncertainty about X given Y

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$$

$H(X|Y) \leq H(X)$ with equality when X & Y are independent

There are three entropies related to a cryptosystem have to consider for analysis

Viz. $H(\mathcal{M}^n) \equiv H(\boldsymbol{M})$, $H(\mathcal{K})$, $H(\mathcal{C}^n) \equiv H(\boldsymbol{C})$, *where* n *is the length of message / cipheretext*

There are two important notions in cryptography

Message / Key Equivocation :

(a)If the attacker can observe n length ciphertexts then what uncertainty remains about the message . It is given by $H(\boldsymbol{M}|\boldsymbol{C}) = -\sum_{\boldsymbol{c}\in\mathcal{C}^n} p(\boldsymbol{c}) \sum_{\boldsymbol{m}\in\mathcal{M}^n} p(\boldsymbol{m}|\boldsymbol{c})\, log_2\big(p(\boldsymbol{m}|\boldsymbol{c})\big)$

$$= -\sum_{\boldsymbol{c}\in\mathcal{C}^n} \sum_{\boldsymbol{m}\in\mathcal{M}^n} p(\boldsymbol{m},\boldsymbol{c})\, log_2\big(p(\boldsymbol{m}|\boldsymbol{c})\big)$$

(b) If the attacker can observe n length ciphertexts then what uncertainty remains about the key .

It is given by $H(\mathcal{K}|\boldsymbol{C}) - \sum_{\boldsymbol{c}\in\mathcal{C}^n} \sum_{k\in\mathcal{K}} p(k,\boldsymbol{c})\, log_2\big(p(k|\boldsymbol{c})\big)$

It satisfies the following

$$H(\boldsymbol{M}|\boldsymbol{C}) \leq H(\boldsymbol{M}) \;\&\; H(\mathcal{K}|\boldsymbol{C}) \leq \mathrm{H}(\mathcal{K})$$

Ciphertexts does not provide more information about message and key

In terms of Entropy a system is perfectly secure iff $H(\boldsymbol{M}|\boldsymbol{C}) = \boldsymbol{H}(\boldsymbol{M})$

As n *increases* , $H(\mathcal{K}|\boldsymbol{C})$ reduces.

Formally Shannon gives the following important result

Shannon's Result : $H(\mathcal{K}|\boldsymbol{C}) \geq H(\mathcal{K}) - nD$

It leads to the other important notion in cryptography based on the redundancy of the source of the language

Unicity Distance – It is the value of length n of ciphertext for a cryptosystem which takes $H(\mathcal{K}|\boldsymbol{C}) \approx 0$

From the Shannon's result it shows that if $n \geq \frac{H(\mathcal{K})}{D}$ then $H(\mathcal{K}|\boldsymbol{C}) = 0$ i.e the uncertainty about the key might be close to zero.

It implies that the

$$\text{unicity distance} = \frac{H(\mathcal{K})}{D}$$

From practical point of view it gives a rough boarder line between the case when there are several possible solutions & the case when there is only one possible key or the message

So redundancy in source helps in cryptanalysis so compression should done before encryption to improve the security of a cryptosystem

C. Shannon identified two fundamental properties viz. confusion and diffusion of the operation of a secure crypto system in his famous paper, ***"Communication Theory of Secrecy Systems***" published in 1949 to handle the statistical properties and other relations to be used for cryptanalysis of symmetric key cryptography .

Confusion - Make the relationship between the key and plaintext bits with the ciphertext as complex as possible involving many key bits

Diffusion - Dissipate the property of redundancy in the statistics of the plaintext in the statistics of the cipher text.

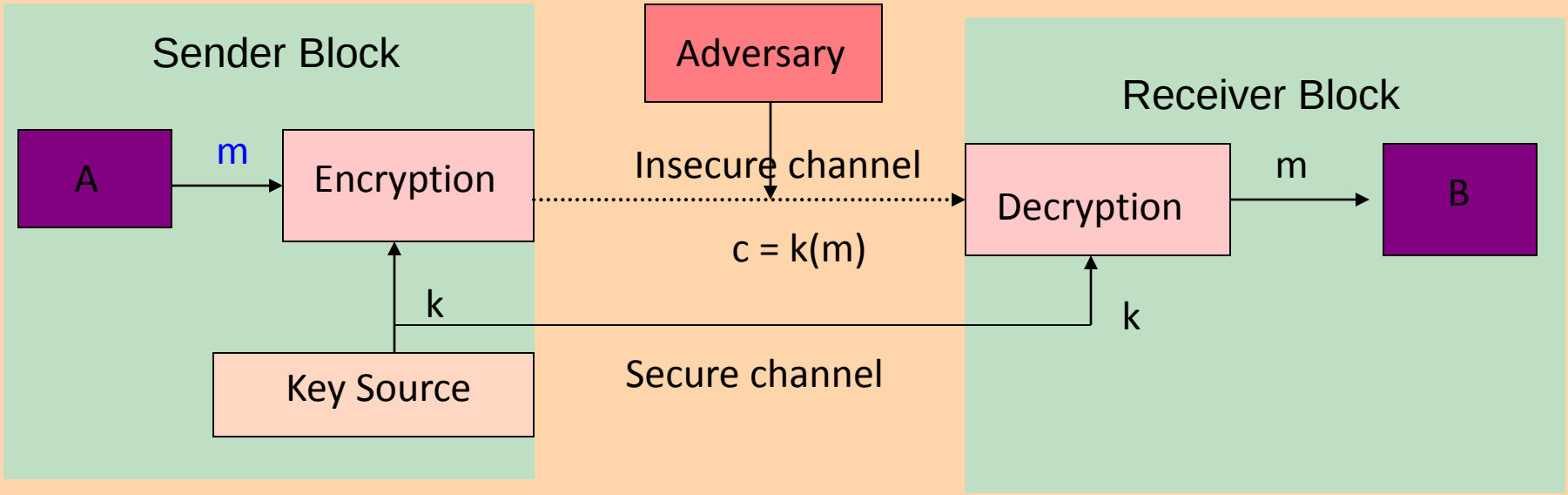
In other words, each plaintext bit or key bit affects many bits of the ciphertext

Good confusion & Diffusion functions provide computational secrecy of the cryptosystem

Symmetric(Secret / Private) key cryptography

In Symmetric key Cryptography Encryption and Decryption keys are same

i.e. e = d = k (say)

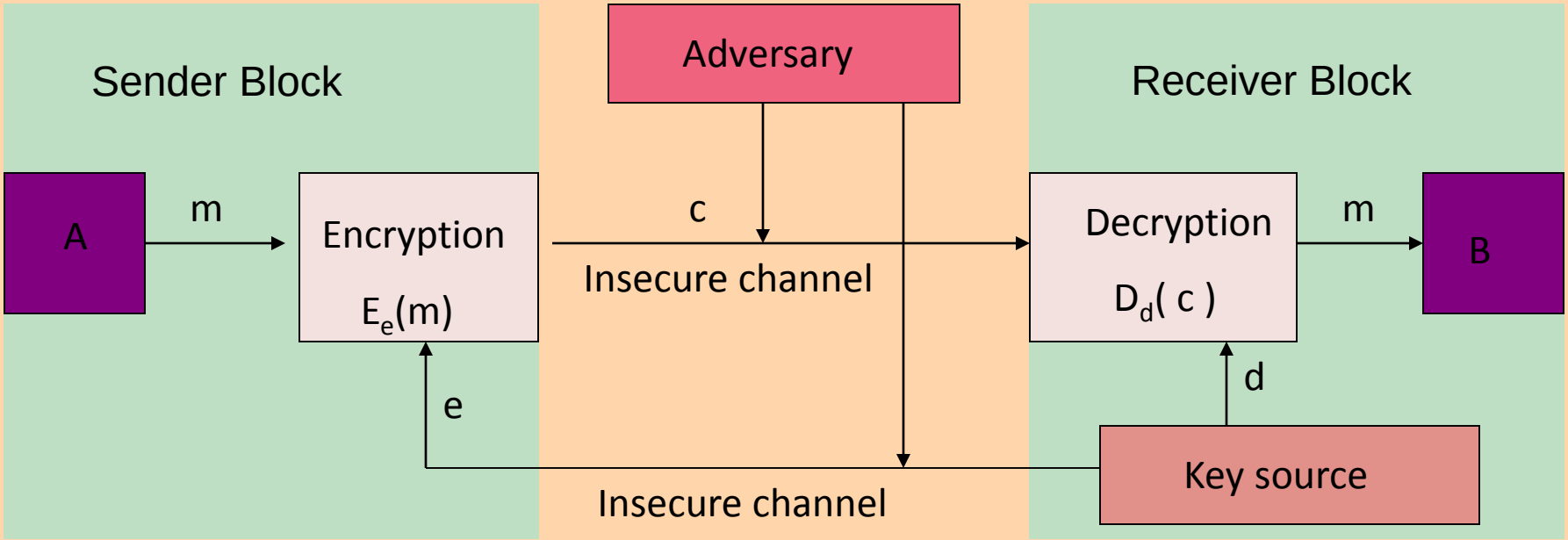


Block diag. of two party communication using symmetric key

Asymetric (Public Key) Cryptography

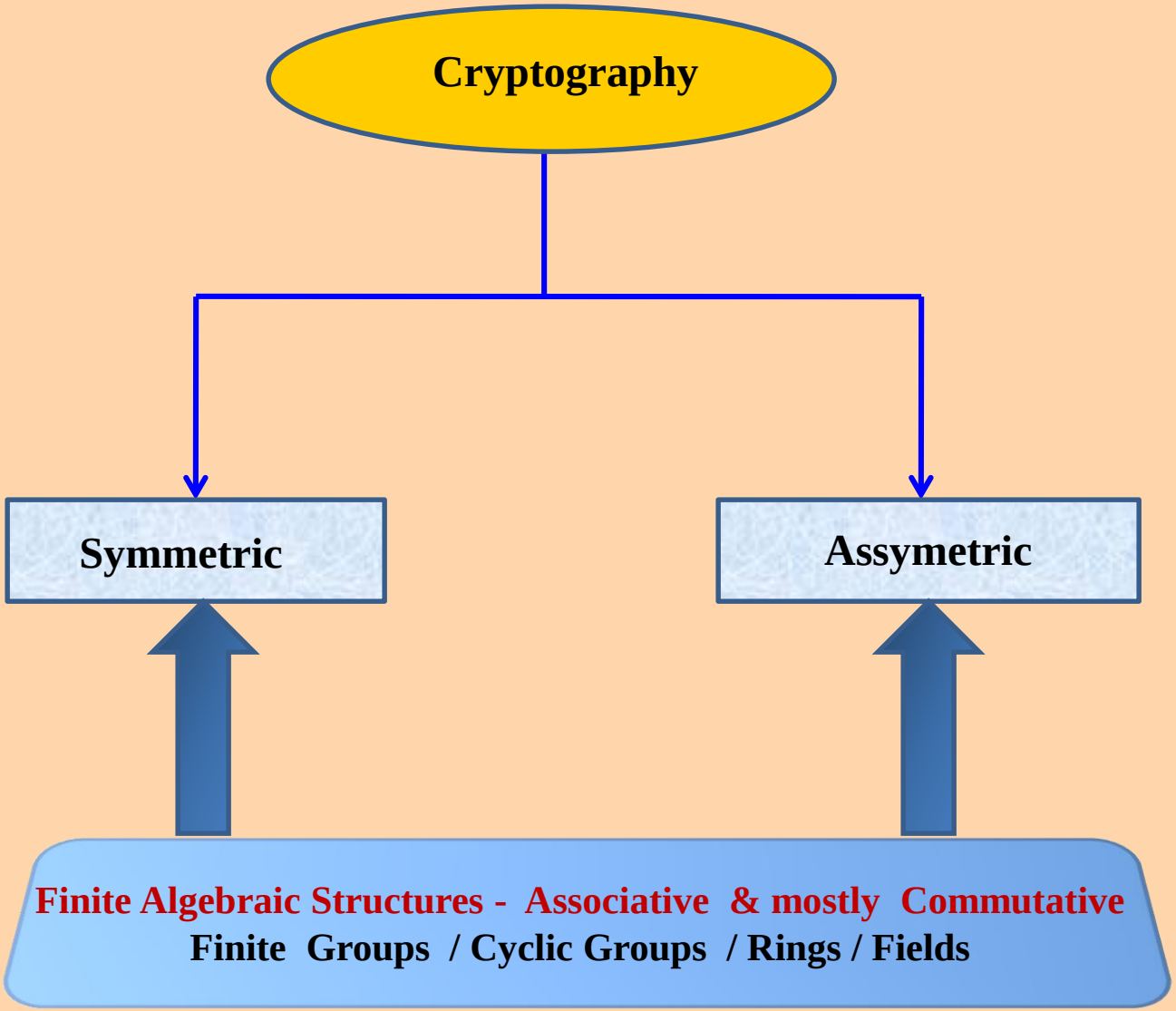
Given e (encryption key) it is infeasible to determine the corresponding decryption key d s.t $D_d(E_e(m)) = m$

E_e – being viewed here as **TOF** with d being the trapdoor information necessary to compute inverse func. and hence allow decryption –(Provable secure)



Block diagram of two party communication using public key cryptography

Commonly Used Algebraic Structures in Cryptography



- ***Finite Fields , are used mainly in Symmetric Ciphers***

- Z_n & Z_n^* *are another two important structures in cryptography*

- *In public key cryptography based on DLP mainly used* ***prime order cyclic subgroup of*** Z_p^*

  For secrecy generally use modulo large prime / $GF(2^m)$ *where* m *is quite large*

- *ECDLP also based on* ***cyclic group***

- ***Choice of the cyclic groups*** *are important for the security.*

- ***All these structures are Associative & Commutative***

New Directions in Cryptography (*Motivation / Background)*

- New directions of cryptography motivated & developed to handle the present (*ever increasing*) security requirements for secure digital communication
- One such possibilities to use the other *(not commonly used*) algebraic structures
- **In late 80's** part of crypto community (mainly European) visualizing the strong potentiality of **using non-associative / non-commutative algebraic structures** ***- a new direction in cryptography***
- *In this case order of operations matter*
- *It also broaden the used algebraic structures suitable for cryptographic purposes*
- *Obtained new crypto primitives with different properties than existing one*
- *It also able to protect against some known attacks*
- *It leads to enlarge the domain of the crypto primitives*

- In this direction, most suitable algebraic structures are **quasigroups / n-ary quasigroups**
- It has also correspondence with the combinatorial structures ***Latin Squares***
- These structures have **great potential to enhance security** based on their

  (i) Algebraic structures

  (ii) Quasigroup / n-ary quasigroup identities and

  (iii) Large number of quasigroups

  (iv) Easy to compute (QG based enc/dec functions)
- This is one of the **current research direction** to deign new crypto primitives and algorithms, PRNG, design error-correcting codes, ….

Quasigroups

- **Quasigroups** may be defined both from **combinatorial** and **algebraic** point of view
- Known as **combinatorial quasigroups & equational (algebraic) quasigroups**

respectively

Here by algebraic structures we mean generalized algebraic structures (Universal algebra)

Universal Algebra

An universal algebra $\mathbb{A}$ is a pair $\langle A, F \rangle$, with A a nonempty set , called the universe of $\mathbb{A}$, and $F = \langle f_i : i \in I \rangle$, a sequence of finitary operations on A .

The operations in F are called the basic operations of $\mathbb{A}$ and the set I is called the index set of $\mathbb{A}$

The ***type (signature)*** of $\mathbb{A}$ is the function $\tau : I \to \mathbb{N}$, where $\tau(i)$ is equal to the arity of the function f_i

The arity of an operation f on A is n , if and only if the domain of f is A^n

Two algebras are said to be similar if and only if they have the same type

Examples

1. A group $\mathbb{G} = \langle G, F = \{\cdot, ^{-1}, 1\}\rangle$ is an algebra with signature (type) $(2,1,0)$

2. A quasigroup $\mathbb{Q} = \langle Q, F = \{\cdot, /, \backslash\}\rangle$ is an algebra with signature (type) $(2,2,2)$

Definition 1 : A **(combinatorial) quasigroup** $(Q,\cdot)$ is a groupoid consisting of elements of Q with respect to a binary operation $'\cdot'$ such that for all $a, b \in Q$ there exists unique $x, y \in Q$ for which it satisfies the identities $a \cdot x = b \; and \; y \cdot a = b$

In other words, the equations, $a \cdot x = b \;\; and \;\; y \cdot a = b$***, for any given*** $a, b \in Q$ ***have unique solutions*** $x, y \in Q$

i.e. for any three elements $x, y, z \in Q$ ***specification of any two in the equation*** $x \cdot y = z$ ***determines the third element uniquely***

Latin squares : A Latin square of order m is a $m \times m$ square containing m copies of each of m symbols, arranged in such a way that no symbols is repeated in any row or column

Ex:

1	3	2	5	6	4
3	2	1	6	4	5
2	1	3	4	5	6
4	5	6	1	2	3
5	6	4	2	3	1
6	4	5	3	1	2

Fig 1 : A Latin square of order 6

Each Latin square may be bordered to yield the binary operation (multiplication) table of a quasigroup of same order.

Ex: Consider the Latin square of Fig 1. First labeling the rows and columns of Latin square by 1,. . . ,6 in order. Obtain the binary operation (multiplication) $'\cdot'$ table of a quasigroup $(Q,\cdot)$ of order 6 which is as follows :

.	1	2	3	4	5	6
1	**1**	**3**	**2**	**5**	**6**	**4**
2	**3**	**2**	**1**	**6**	**4**	**5**
3	**2**	**1**	**3**	**4**	**5**	**6**
4	**4**	**5**	**6**	**1**	**2**	**3**
5	**5**	**6**	**4**	**2**	**3**	**1**
6	**6**	**4**	**5**	**3**	**1**	**2**

Fig 2 : A Latin square yields a multiplication table

Conversely, the body of the multiplication table of a finite quasigroup Q yields a Latin square.

For any two fixed elements x, z of Q, the existence of the solution y to the equation $x.y = z$ means that the element z appears at least once in the row of multiplication table labeled by x (namely in the column labeled by y). The uniqueness of the solution z means that the element y appears at most once in the row of the multiplication table labeled by x. Similarly for columns

Definition 2 : An **(equational / algebraic) quasigroup** $(Q,\cdot,/,\backslash)$ is defined as a set Q closed under three binary operations $'\cdot'(multiplication)$, $'/'(right\ division)$ and $'\backslash'(left\ division)$, satisfying the following identities

1. $(y/x)\cdot x = y$
2. $(y\cdot x)/x = y$
3. $x\cdot(x\backslash y) = y$
4. $x\backslash(x\cdot y) = y$

From these four identities, following two more identities can also be derived

5. $x/(y\backslash x) = y$
6. $(x/y)\backslash x = y$

It is easy to prove that if $(Q,\cdot,/,\backslash)$ is an equational (algebraic) quasigroup then $(Q,\cdot)$ is a combinatorial quasigroup

Conversely, suppose that $(Q,\cdot)$ is a combinatorial quasigroup. For given elements $x, y \text{ of } Q$, define $x \setminus y$ as the unique solution of (3), and y/x as the unique solution of (1) in the Definition 2. It defines the binary operations $'/'$ and $'\setminus'$ on Q that make $(Q,\cdot,/,\setminus)$ an equational quasigroup.

Note that $(Q,/)$ *and* $(Q,\setminus)$ *are also Latin squares.*

So, usually not necessary to distinguish between the concepts of combinatorial and equational quasigroup. They are generally referred as simply quasigroups.

Advantage of Definition 2

The equational definition of quasigroups means that they form a variety and thus we can study them by the methods and concepts of Universal algebras .

- Subquasigroups
- Equivalence & Congruence relations on quasigroups
- Simple quasigroups
- Quasigroup homomorphisms / isomorphisms / Isotopies

⋮

Q_m^n - Denotes the set of *n*-ary quasigroups of order *m*

- The total number of *n*-ary quasigroups of order *m* is given by

$$|Q_m^n| = m!\,(m-1)!^{n-1}\,|R_m^n|$$

It increases asymptotically as m & n increases

we have tabulated below some cases of *m* & *n*

n	*m*	$\|R_m^n\|$	$\|Q_m^n\|$
1	4	1	24
2	4	4	576
2	5	56	161280
2	6	9408	812851200
3	4	64	55296
3	5	40246	2781803520
4	4	7132	36972288

We can generate *n*-ary quasigroups in two ways. If it is derived from any $(n-1)$ -ary quasigroups then it is called a **reducible *n*-ary quasigroup** , else it is called an **irreducible *n*-ary quasigroup.**

❖ ***key space can make as large as required by proper choice of parameters optimally***

Cryptographic Potential Quasigroup Transformations

There are different types of quasigroup transformations .

Here we discuss mainly elementary quasigroup transformations :

Let $(Q,\cdot)$ be a given QG , for a fixed element $l \in Q$, known as ***leader***

$$e_{l,\cdot}(e_l): Q^+ \to Q^+ \ where \quad Q^+ = \bigcup_{k\geq 1} Q^k \ , Q^k = \{x_1 \cdots x_k \mid x_i \in Q\}$$

given by

$$e_l(x_1,\cdots,x_k) = y_1 \cdots y_k, where \quad \begin{cases} y_1 = l \cdot x_1 \\ y_i = y_{i-1} \cdot x_i, \ i = 2,\cdots,k \end{cases}$$

Known as (left) e-transformation

Define another elementary transformation $d_{l,\backslash}(d_l)$ on $(Q,\cdot)$ with leader $l \in Q$given by

$$d_{l,\backslash}(d_l)(x_1,\cdots,x_k) = y_1 \cdots y_k, where \quad \begin{cases} y_1 = l \backslash x_1 \\ y_i = x_{i-1} \backslash x_i, \ i = 2,\cdots,k \end{cases}$$

Known as (left) d-transformation

Similarly we can define (right) e' -transformation and (right) d'-transformation and denoted by $e'_{l,.}$ and $d'_{l,/}$ respectively.

These transformations are also commonly known as ***Elmentary Quasigroup String Transformations***

Graphical presentation of these two transformations are as follows

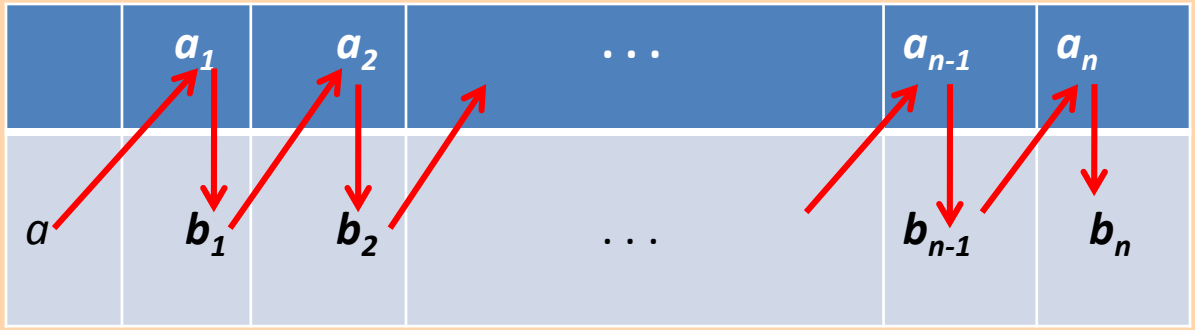


Fig1. e-transformation

Fig2. d-transformation

- Let $(Q,\cdot,/,\backslash)$ be finite quasigroup. Then for each $\alpha \in Q^{+}$and for each $l \in Q$ the e-transformation $e_{l,\cdot}$ & $d_{l,\backslash}$ are mutually inverse permutations of Q^{+}

$$i.e.,\ d_{l,\backslash}\left(e_{l,\cdot}(\alpha)\right)=\alpha=e_{l,\cdot}\left(d_{l,\backslash}(\alpha)\right)$$

It follows from the identities of of Quasigroup

$$x\cdot(x\backslash y)=y \ \&\ x\backslash(x\cdot y)=y$$

- Note that basesd on these properties we can construct the stream & Block cipher with the $\left(e_{l,\cdot}, d_{l,\backslash}\right)$ Enc / Dec function as quasigroup transformation and vice versa
- Note that we can get 6 pairs of Enc/Dec quasigroup based function

viz. $\left(e_{l,\cdot}, d_{l,\backslash}\right),\ \left(e'_{l,\cdot}, d'_{l,\backslash}\right),\ ,\left(e_{l,\cdot}, d_{l,/}\right),\ \left(e'_{l,\cdot}, d'_{l,/}\right)$

$\left(e_{l,*}, d_{l,\backslash\backslash}\right),\ \left(e'_{l,*}, d'_{l,\backslash\backslash}\right),\ \left(e_{l,*}, d_{l,//}\right),\ \left(e'_{l,*}, d'_{l,//}\right)$

- Similarly another 6 pairs of right elementary mutually inverse transformations exist

This is the advantage over Symm cipher built over GF(2) where the unique operation XOR can only be applied

❖**Not all quasigroups are suitable for Cryptographic Purposes**

Lots of research are going on to find the suitable choice of QG –

It is an important issue for security strength

❖ From algebraic structural point of view suitable choice of **Quasigroups** have to be **Polynomially (functionally) complete , no subquasigroups and high deg of non- associativity and non commutativity**

❖ **Challenging research area to test and construct good choice of quasigroups of finite order**

$(Q,\cdot)$ =

·	1	2	3	4
1	2	3	4	1
2	1	4	3	2
3	3	2	1	4
4	4	1	2	3

$$\alpha = 1212121212 \quad \& l = 2$$
$$e_2(1212121212) = 1332133213$$
$$e_2^2(1212121212) = e_2(e_2(1212121212))$$
$$= e_2(1332133213) = 1424423214$$

$(Q,\bullet)$ =

•	1	2	3	4
1	1	2	4	3
2	2	1	3	4
3	4	3	2	1
4	3	4	1	2

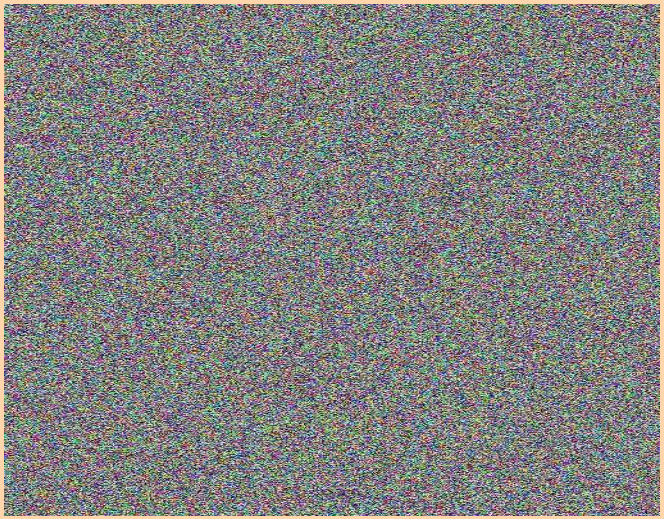
$$\alpha = 1212121212 \quad \& l = 2$$
$$e_2(1212121212) = 2112211221$$
$$e_2^2(1212121212) = e_2(e_2(1212121212))$$
$$= e_2(2112211221) = 1112111211$$

12	1	5	10	15	4	13	3	7	6	9	14	11	2	8	0
7	6	9	14	11	2	8	0	12	1	5	10	15	4	13	3
1	12	10	5	4	15	3	13	6	7	14	9	2	11	0	8
6	7	14	9	2	11	0	8	1	12	10	5	4	5	3	13
2	11	0	8	6	7	14	9	4	15	3	13	1	12	10	5
4	15	3	13	1	12	10	5	2	11	0	8	6	7	14	9
11	2	8	0	7	6	9	14	15	4	13	3	12	1	5	10
15	4	13	3	12	1	5	10	11	2	8	0	7	6	9	14
3	13	4	15	10	5	1	12	0	8	2	11	14	9	6	7
0	8	2	11	14	9	6	7	3	13	4	15	10	5	1	12
13	3	15	4	5	10	12	1	8	0	11	2	9	14	7	6
8	0	11	2	9	14	7	6	13	3	15	4	5	10	12	1
9	14	7	6	8	0	11	2	5	10	12	1	13	3	15	4
5	10	12	1	13	3	15	4	9	14	7	6	8	0	11	2
14	9	6	7	0	8	2	11	10	5	1	12	3	13	4	15
10	5	1	12	3	13	4	15	14	9	6	7	0	8	2	11

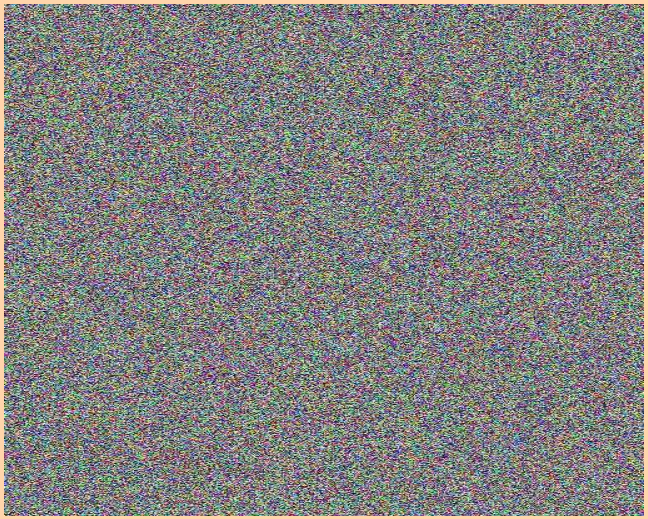
Cryptographically suitable quasigroup of order16

Original image

One round e-transformation

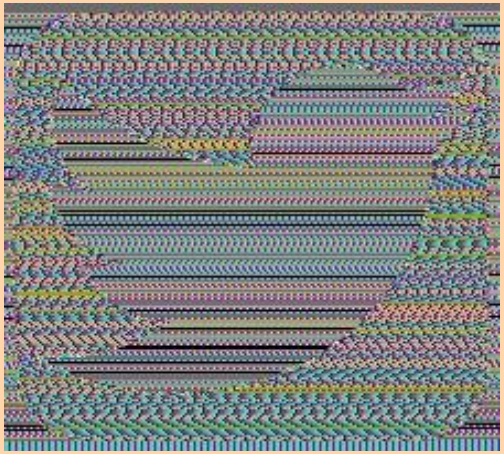


3 round e-transformation

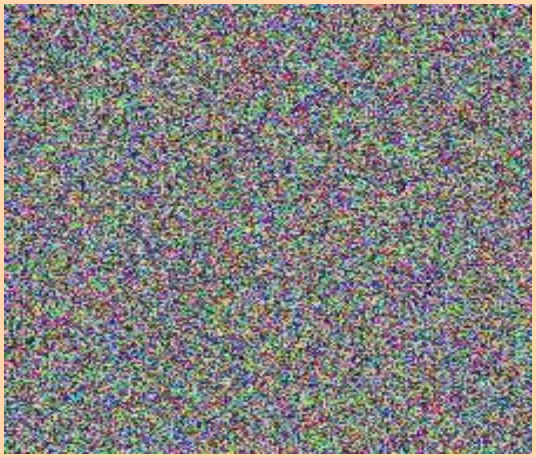


10 round e-transformation

Original image



One round e-transformation



10 round 10 e-transformation

- *Note that lots of generalized elementary transformations are developed and the research is going on to develop new transformations and use the composition of transformations in a proper way to design the new cryptographic schemes & primitives*

- *Edon80 –a stream cipher based on generalized elementary transformation*

References

1. Artamonov V. A.: Polynomially complete Algebras, Scie. Notes Orlov State Univ. (Sci. Journal) Series Natural, Technology and Med.Sci. part 2 , pp 23-29 , 2012 (Russian)

2. Artamonov V. A., Chakrabarti S., Gangopadhyay S., Pal S. K.: On Latin Squares of Polynomially Complete Quasigroups and Quasigroups generated by Shifts, Quasigroups and Related Systems, Vol 21, No. 2, 117-130, 2013.

3. Artamonov V. A., Chakrabarti S., Pal S. K.: Characterization of Polynomially Complete Quasigroups based on Latin Squares for Cryptographic Transformations, Discrete Applied Mathematics, may,2015.

4. Artamonov V. A., Chakrabarti S., Pal S. K Characcterizations of Highly Non-associative Quasigroups and Associative Triples, Quasigroups and Related Systems, 25(2017)1-19.

5. Artamonov V.A., Chakrabarti S, Markov V.T , Pal S.K : Constructions of Polynomially Complete Quasigroups of Arbitrary Order, Journal of Algebra and Its Applications, accepted Aug,2020

6. Artamonov V.A., Chakrabarti S, Tiwari S.K Markov V.T : Algebraic Properties of Subquasigroups And Construction of Cryptographically Suitable Finite Quasigroups *Submitted to Discrete Applied Mathematics, Aug 2020.*

7. Chakrabarti S., Pal S. K., Gangopadhyay S.: An Improved 3-ary quasigroup Based Encryption Scheme, ICT Innovations Conference 2012 on Secure and Intelligent Systems, Macedonia,Web proceedings ISSN 1857-7288, 173-184, 2012.

8. Chakrabarti S., Pal S.K On Increasing Key Space of Quasigroup Based Ciphers, presented in National Workshop on Cryptology, India, 2013.

9. Denes J., Keedwell A.D. :LatinSquares. New Development in the Theory and Applications, Vol 46, Annals of Discrete Mathematics, North –Holland, 1991

10. Glukhov M. M.: On Application of Quasigroups in Cryptology, Applied Discrete Mathematics, 2, 28-32, 2008 (Russian).

11. Gligorski D., Dimitrova V. and Markovski S. : Quasigroups as Boolean functions, their Equation Systems and Groebner Bases, M Sala, T Mora etc(Eds), Groebner Bases, Coding and Cryptography, Springer, 2009.

12. Gligorski D., Markovski S., and Knapskog S. J. : The Stream Cipher Edon 80, Stream Cipher Designs : The eSTREAM Finalists , LNCS, Vol. 4986, pp. 152-169, 2008.

13. Mileva A.: Cryptographic Primitives with Quasigroup Transformation, PhD Thesis, Faculty of natural science, Ss Cyril and Methodius University in Skopje, Republic of Macedonia, 2010.

14. Mileva A.: Chapter on New Developments in Quasigroup-Based Cryptography, Multidisciplinary perspectives in cryptology and information security, Edited by Sadkhan S. B.,etc, IGI,Global, 2014.

15. Menezes A.J , VanOorschot P. C– Handbook of Applied Cryptography

16. Stanley B, sankappanavar H : A Course in unversal Algebra, Springer

17. Shannon C.E. – A Mathematical Theory of Communication –BSTJ -1948

18. Shannon C.E. – Communication Theory of Secrecy Systems – BSTJ – 1949

19. Smith J.D.H., : An Introduction to quasigroups and their representations , Chapman & Hall / CRC, 2007

20. Stanley B, sankappanavar H : A Course in unversal Algebra, Springer

THANKS

?

Learn from Yesterday

Live for Today

Hope for Tomorrow

The Important thing is not to Stop Questioning

A. Einstein