Criptografía poscuántica

May 8, 2013

Luis J. Dominguez Perez
basado en trabajo de P. Barreto y N. Sendrier
ldominguez@tamps.cinvestav.mx

Cinvestav Tamaulipas,

Agenda

Introduction

Mecánica cuántica

Computadora cuántica

Criptografía basada en códigos

Funcionamiento

Ejemplos

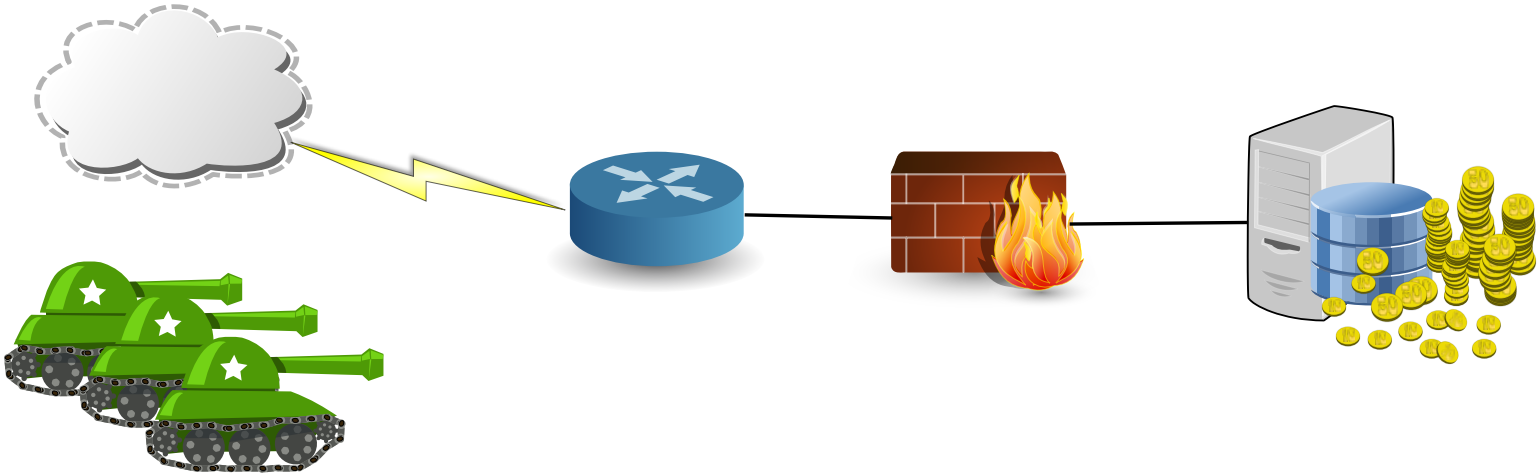
Sistemas desplegados

- Supuestos de intratabilidad convencionales:
  - Factorización entera (IFP): RSA
  - Logaritmo Discreto (DLP): Diffie-Hellman (DHP), y sus variantes bilineares: ECC y PBC

- Estos supuestos se reducen a lo que se conoce como El Problema del Subgrupo Escondido (HSP).

Balance entre la protección y lo protegido

- La principal preocupación de un sistema de seguridad es cuánto tiempo le tomará a un atacante romperlo, y cuántos recursos le son necesarios.
- El costo de romper un sistema, medido tanto en tiempo y en dinero invertido en recursos computacioneales, debe de ser mayor a el valor de la información protegida



Beneficios y tiempos

- Cuando un atacante necesitase 100 millones de USD para obtener un beneficio de 100 mil USD, un atacante típicamente desistirá

- Cuando se dice que a un atacante le tomará alrededor de 100 años en romper un sistema, es calculado al inicio del ataque.
  - Con el tiempo, nuevos métodos y mejores recursos están disponibles
  - Un atacante podría ir actualizando sus recursos y reducir drásticamente los 100 años inicialmente estimados

Calculando el costo

- En un sistema simétrico, el tiempo de vida de la protección de los datos puede ser calculado con el tiempo que nos tomaría romper el esquema mediante fuerza bruta, utilizando un diccionario de llaves, u otro método

- En “Using the cloud to determine key strengths”, Kleinjung et al. popusieron un esquema de medición para comparar el costo de romper un esquema criptográfico utilizando los servicios de Amazon.
  - Romper un esquema de 128 bits de seguridad utilizando los servidores cloud de Amazon costaría 10^{27} USD.

Algoritmo de Shor

- Factoriza un entero no primo en tiempo polinomial
- Son dos partes principales:
  - Reducción del problema de factorización a un problema de búsqueda ordenada (cómputo clásico)
  - Resolver el problema de búsqueda ordenada (cómputo cuántico)

Consecuencias: RSA podría ser roto ya que su fuerza recae en la dificultad de encontrar los dos factores primos de un número entero grande.

Ataques

- El algoritmo cuántico de Shor permite resolver en tiempo polinomial aleatorios ciertos esquemas del problema del subgrupo escondido para grupos finitos *abelianos*.

- En particular, este algoritmo puede romper los esquemas:
  - RSA
  - DSA
  - ECDSA

  en un tiempo $\mathcal{O}(\log N)^3$

Algoritmo de Shor

Algorithm 1 Algoritmo de Shor

Input: $N \in \mathbb{N}$ no primo
Output: un factor de N

1: Escoja un x aleatorio entre el rango $[2..N]$
2: Si el MCD$(x, N) \neq 1$ terminar
3: Encontrar el orden r de x mod $N(x^r \equiv 1 \bmod N)$
4: **if** r es par y $x^{r/2} \neq \pm 1 \bmod N$ **then**
5: MCD$(x^{r/2} + 1, N)$ es un factor no trivial de N
6: **else**
7: intentar con otra x
8: **end if**

Alcance

- Los algoritmos criptográficos basados en la dureza de problemas en los que el algoritmo de Shor no se puede aplicar, se conocen como **sistemas criptográficos poscuánticos**
- Estos esquemas están basados en problemas computacionalmente **NP**-completos o **NP**-duros:
  - Funciones picadillo
  - Códigos
  - Retículas
  - Ecuaciones cuadráticas multivariables
  - Llave secreta
  - Isogenias sobre curvas elípticas supersingulares
  - Grupos no abelianos
  - etc.

Mecánica cuántica

- Física a nivel atómico y subatómico
- Requiere teoría precisa
- El estado del sistema no es dado por observaciones físicas
- Es imposible conocer el estado actual del sistema
- Se pueden hacer predicciones probabilísticas

Definición

Notación $|\cdot\rangle$, se conoce como *ket*, y es un vector en el espacio de todos los estados posibles.

Conceptos

- **Matrix Hermitiana**. Es una matrix cuadrada con elementos complejos, y que es igual a su transpuesta conjugada)
- **Operador lineal**. Es una función matemática entre dos espacios vectoriales, que preserva las operaciones de adición de vectores y multiplicación por un escalar.
- **Eigen-vector**. Los vectores propios, autovectores o eigenvectores de un operador lineal son los vectores no nulos que, cuando son transformados por el operador, dan lugar a un múltiplo escalar de sí mismos, con lo que no cambian su dirección.
- **Eigen-espacio**. Un espacio propio, autoespacio, eigenespacio o subespacio fundamental asociado al valor propio λ es el conjunto de vectores propios con un valor propio común.

Principios cuánticos

- Una cantidad observable se representa como una matrix Hermitiana $\mathscr{A}$. Los posibles valores de una medición son el eigen-vector de $\mathscr{A}$. Esta matriz puede ser unitariamente diagonalizada.
  - Los eigenvectores de $\mathscr{A}$ generan una base ortogonal, tal que, cada vector $|\psi\rangle$ puede ser representado por una combinación lineal de estos eigen-vectores $|\phi_i\rangle$.

$$|\psi\rangle = c_1|\phi_1\rangle + \ldots + c_n|\phi_n\rangle$$

- Dada una matiz $\mathscr{A}$ observable, su medición nos arroja un $|\phi_i\rangle$ con una probabilidad $|c_i|^2$

Definición

Entrelazado cuántico. Es el fenómeno en el cual los estados cuánticos de dos objetos se describen como referencia entre sí, por ejemplo:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|B_1 C_2\rangle + |B_2 C_1\rangle)$$

Definiciones

- **Qubit**. Un qubit es la unidad de información cuántica. Sus dos estados computacionales clásicos son:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ y } |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

  Un qubit puede ser representado por la combinación lineal de estos dos estados.

- **Registro de qubit**. Un número de qubits entrelazados a analizar, o el producto tensorial de los qubits.

- **Compuerta cuántica**. Es un circuito cuántico que opera sobre un determinado número de qubits

Códigos lineares

Definiciones:

- El peso Hamming $w(u)$ de $u \in (\mathbb{F}_q)^n$ es el número de componentes de u que no están en cero.
- La distancia Hamming entre $u, v \in (\mathbb{F}_q)^n$ es $dist(u, v) = w(u - v)$.
- Un código linear $\mathscr{C}[n, k]$ sobre $\mathbb{F}_q$ es un vector de k dimensiones en el subespacio de $(\mathbb{F}_q)^n$.

- Un código puede ser definido por una matriz generadora $G \in (\mathbb{F}_q)^{k \times n}$ o por una matriz de verificación de paridad $H \in (\mathbb{F}_q)^{r \times n}$ with $r = n - k$:
  - $C = \{uG \in (\mathbb{F}_q)^n | u \in (\mathbb{F}_q)^n\}$
  - $C = \{v \in (\mathbb{F}_q)^n | Hv^T \in O^r\}$

  donde $HG^T = O$.

- Un vector s es llamado síndrome of v, si $Hv^T = s^T$.

Matriz generadora

- Las matrices generadores y de verificación de paridad no son únicas:
  - dada una matriz no singular $S \in (\mathbb{F}_q)^{k \times k}$ (resp. $S \in (\mathbb{F}_q)^{r \times r}$)
  - la matriz $G' = SG$ (resp. $H' = SH$) define el mismo código que G (resp. H) en otra base.

- Como consecuencia, la forma sistemática (escalonada) $G = [I_k | M]$, $H = [-M^T | I_r]$ donde $M \in (\mathbb{F}_q)^{k \times r}$ no es siempre posible.

Códigos equivalentes

- Dos códigos son equivalentes (aún permutados) si difieren por una permutación en las coordenadas de sus elementos.

- Formalmente, un código C' generado por G' es equivalente a un código C generado por G sí y solo sí $G' = SGP$ para alguna matriz de permutación $P \in (\mathbb{F}_q)^{n \times n}$, y una matriz no singular $S \in (\mathbb{F}_q)^{k \times k}$.

En esencia, se dice que $C' = CP$.

Funcionamiento

La decodificación funciona de la siguiente manera:

Decodificación General

- Dados unos enteros positivos (n, k, t), un campo finito $\mathbb{F}_q$, un código linear $\mathscr{C}[n, k] \in (\mathbb{F}_q)^n$ definido por una matrix generadora $G \in (\mathbb{F}_q)^{k \times n}$, y un vector $c \in (\mathbb{F}_q)^n$
- £existe un vector $m \in (\mathbb{F}_q)^k$ tal que $e = c - mG$ tiene peso $w(e) \leq t$?
- Encontrar tal vector e es un problema NP-completo.

Decodificación Síndrome

- Dados unos enteros positivos (n, k, t), un campo finito $\mathbb{F}_q$, un código linear $\mathscr{C}[n, k] \in (\mathbb{F}_q)^n$ definido por una matrix de paridad $H \in (\mathbb{F}_q)^{r \times n}$ con $r = n - k$, y un vector $s \in (\mathbb{F}_q)^r$
- £existe un vector $e \in (\mathbb{F}_q)^n$ de peso $w(e) \leq t$ tal que $He^T = s^T$?
- Encontrar tal vector e también es un problema NP-completo.

Decodificación Permutada

- Resolver el problema de decodificación general o el de decodificación síndrome para un código C que sea equivalente tras permutación a un código C' eficientemente decodificable, consiste en encontrar la permutación y cambio de bases entre los códigos, y utilizar el código C' como pasadillo secreto (trapdoor) para decodificar en C.
- Esto se cree que es lo *suficientemente difícil* para “ciertos códigos”.

Decodificación Recortada

- Resolver el problema de decodificación general o el de decodificación síndrome para un código C que sea equivalente tras permutación a algún subcódigo recortado (una proyección) de algún código C' eficientemente decodificable, consite en encontrar la permitación y cambio de bases entre los códigos, y usar el código C' como pasadillo secreto (trapdoor) para decodificar C.
- Decidir si el código es equivalente al código recortado es un problema NP-completo.

Ejemplos de sistemas

Existen varios esquemas criptográficos basados en códigos, por ejemplo: McEliece, Niederreiter, firmas CFS, entre otros.

Sistema McEliece

- Este criptosistema utiliza la matriz generadora para la clave pública, por lo que para el cifrado, el secreto se multiplica por ella y se le agrega la información de corrección de errores (ruido).
- Para el decifrado, se recupera la información del vector de errores y se elimina del mensaje recibido.

ejemplos de sistemas

Sistema Niederreiter

- A diferencia del anterior, en este caso se utiliza la matriz de permutación como clave pública, por lo que el transpuesto del vector de corrección de errores (que contiene el mensaje) se multiplica por esta matriz.
- Para el decifrado, el proceso es similar.

ejemplos de sistemas

Firmas CFS

- Para las firmas, la clave pública también utiliza la matriz de permutación.
- Dado un oráculo aleatorio, y hasta encontrar un síndrome decodificable con el mensaje.
- De igual manera, se extrae el vector de corrección (conteniendo el mensaje) y se utiliza como firma (junto con un valor del oráculo aleatorio).
- Para la verificación, se multiplica la firma con la matriz de permutación y se corrobora con una aplicación de la firma sobre el mensaje (y el valor del oráculo aleatorio).

Cifrado McEliece

- Generación de llaves:
  - Escoja aleatorios $[n,k]$, t-corrección de errores, un código eficientemente decodificable Γ, y una matriz aleatorio de permutación $P \in (\mathbb{F}_q)^{k\times k}$, y calcule una matriz generadora $G \in (\mathbb{F}_q)^{k\times k}$ para el código equivalente ΓP.
  - Kpriv $= (\Gamma, P)$, Kpub = (G, t).
- Cifrado de un texto plano $m \in (\mathbb{F}_q)^k$:
  - Escoja un elemento aleatorio e con error t $e \in (\mathbb{F}_q)^n$ y calcule $c = mG + e \in (\mathbb{F}_q)^n$.
- Descifrado de un texto cifrado $c \in (\mathbb{F}_q)^n$:
  - Corrija los errores en $c' = cP^{-1}$, encuentre el vector e' con t-errores $e' = eP^1$ tal que $c' - e' \in \Gamma$, y recupere m directamente de $c - e \in \Gamma P$.

Ejemplo de Cifrado McEliece

- Sea $n = 8$, $t = 1$, $k = 4$, y un código con la siguiente matriz de paridad H, y matriz generadora G:

$$H = \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$G = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{array}\right]$$

- Cifre el mensaje $m = (1100)$ con vector de error de corrección $e = (00100000)$: $c = mG + e = (11100101)$
- El cálculo del síndrome $Hc^T = (1111)^T$, la corrección del error revela e, y nos da $mG = c - e = (11000101)$.

Retos

- La criptografía basada en códigos, presenta el problema de seleccionar un tipo de código que permita que la decodificación permutada sea fuerte.
- Adicionalmente, sufre del elevado espacio de almacenamiento y transmición requeridos para la llave pública.
- La tendencia actual en investigación para este tipo de criptografía se centra en reducir la llave pública, y en la elaboración de protocolos criptográficos que utilicen códigos como su primitiva.