Gestión de la Seguridad del Software

Dr. Félix Oscar Fernández Peña

ffernandez1@utmachala.edu.ec

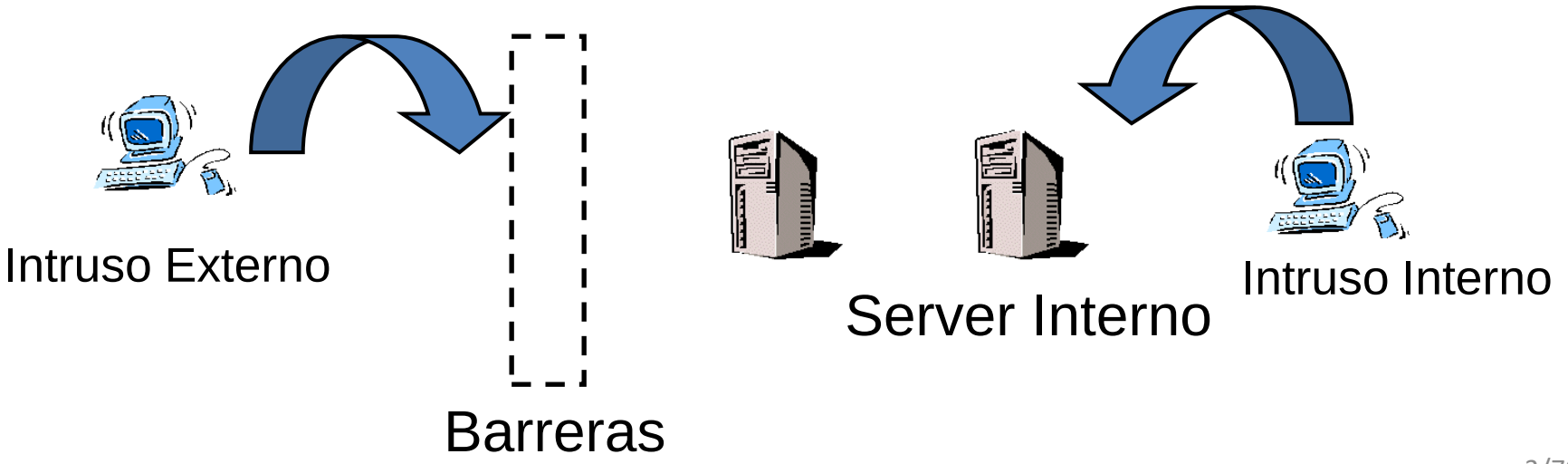
Tipos de Ataque

- Ataque pasivo: quien escucha en el canal de comunicación.

- Ataque activo: quien modifica los mensajes legítimos por un interés particular.

Tipos de Ataque

- Ataque interno.
- Ataque externo.



Aspectos Legales

- En Francia, la criptografía no gubernamental es prohibida, a no ser que el gobierno tenga las claves empleadas.

- Phil Zimmermann, autor de PGP fue acusado de exportación de material de guerra.

Tanenbaum. “Redes de Computadoras”. 3ª edición, pp. 621.

Áreas del Problema

- Confidencialidad: mantener la información accesible solo a usuarios autorizados.

Cifrado

- Autenticidad: ¿con quién se está hablando?

Autenticación

Certificado Digital

- No repudio: certificación de origen.

Firma Digital

- Control de integridad: validación de contenido.

Funciones Hash.

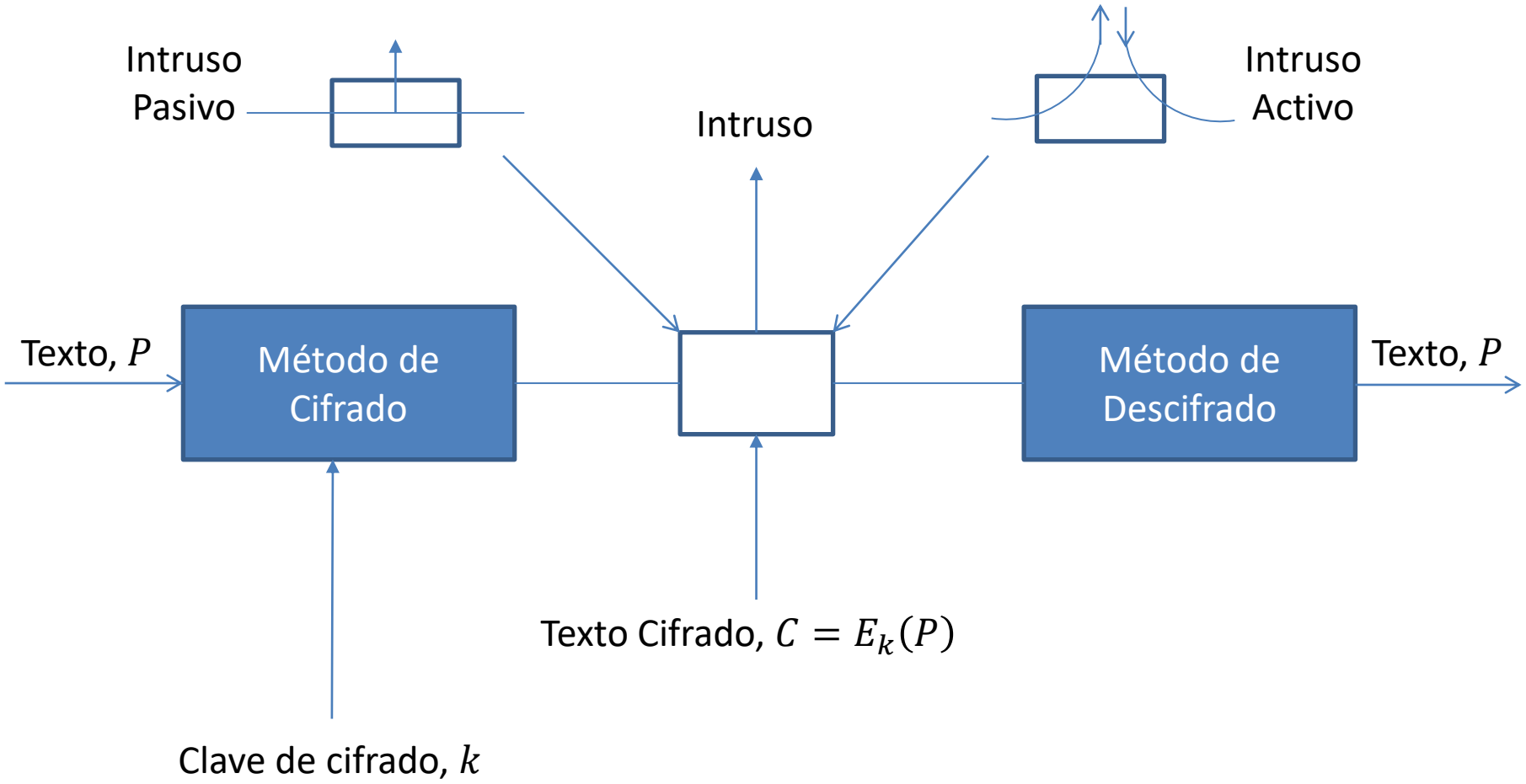
No todo es lo que parece…

¿Cuántos estudiantes se necesitan en un grupo antes de que la probabilidad de tener dos personas con el mismo cumpleaños exceda ½?

Cant. parejas = $\frac{n(n-1)}{2}$

Para n=20 → Cant. parejas = 190, P= $190*\frac{1}{365}>0.5$

Modelo de Cifrado



Pilares de la Seguridad Informática

- La teoría de la información.
  - Estudio de la cantidad de información contenida en los mensajes y claves, así como su entropía.
- La teoría de los números
  - Estudio de las matemáticas discretas y cuerpos finitos que permiten las operaciones de cifrado y descifrado.
- La teoría de la complejidad de los algoritmos
  - Estudio de la clasificación de los problemas como computacionalmente tratables o intratables.

Científico del Día

Claude Elwood Shannon (April 30, 1916 – February 24, 2001) was an American mathematician, electronic engineer, and cryptographer known as "the father of information theory".

A Mathematical Theory of Communication, 1948.

http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html

Información

Conjunto n de datos o mensajes inteligibles creados con un lenguaje de representación y que debemos proteger ante las amenazas del entorno, durante su transmisión o almacenamiento, usando técnicas criptográficas entre otras herramientas.

000/lunes.
001/martes.
010/miércoles.
011/jueves.
100/viernes.
101/sábado.
110/domingo.

Días de la Semana :: 3 bits de información.
Sexo :: 1 bit de información.

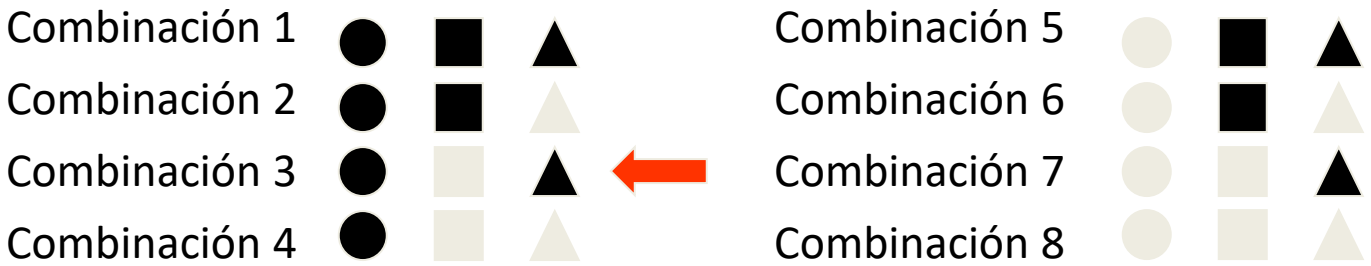
$log_2 n$

Cantidad mínima
de bits necesarios
para codificar un alfabeto
con *n* elementos.

Teoría de la Información

- Evalúa la incertidumbre. Ante varios mensajes posibles, en principio todos equiprobables, aquel que tenga una menor probabilidad de aparición será el que contenga una mayor cantidad de información.
- Mide la cantidad de información que contiene un mensaje a través del número medio de bits necesario para codificar todos los posibles mensajes con un codificador óptimo.

Incertidumbre



$$H = log_2 8 = 3$$

Con sólo tres preguntas "*más o menos inteligentes*" podemos pasar de la incertidumbre total a la certeza:

- Pregunta 1: ¿Está entre la opción 1 y la 4? $\Rightarrow$ Sí
- Pregunta 2: ¿Está entre la opción 1 y la 2? $\Rightarrow$ No
- Pregunta 3: ¿Es la opción 4? $\Rightarrow$ No

¿Dónde le da alegría a su cuerpo Macarena?

- Respuesta 1: En un país de Europa.
- Respuesta 2: En una ciudad de España.
- Respuesta 3: En los números 1 y 3 de la calle Sierpes en Sevilla, España.

¿Dónde hay una mayor cantidad de información?

Está en función de la probabilidad del mensaje

Variable Aleatoria

Sea X una variable aleatoria con n estados posibles con $X = x_i$ una ocurrencia iésima:

$$X = \{x_1, x_2, x_3, \ldots, x_{n-1}, x_n\}$$

$$p_1 = p(x_1),\ p_2 = p(x_2),\ \ldots,\ p_n = p(x_n)$$

Como:

$$0 \leq p_i \leq 1 \quad \text{para } i = 1, 2, \ldots, n$$

Entonces:

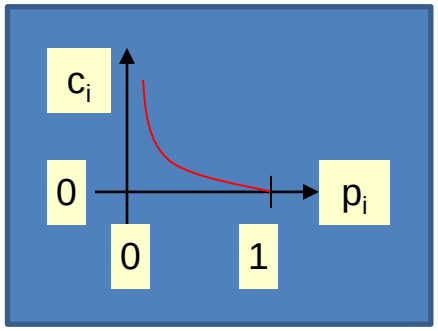
$$\sum_{i=1}^{n} p_i = 1$$

La probabilidad de que ocurra p_1 o p_2 o p_3, etc. será siempre la unidad porque seguro será uno de ellos.

Cantidad de Información

Ci: Cantidad de información del estado i de la variable X.

$$C_i = -log_2(p_i)$$



$p(x_i)$ = 1 $\Rightarrow$ no hay incertidumbre: c_i = 0

$p(x_i) = 0 \Rightarrow$ máxima incertidumbre: $c_i \to \infty$

- Signo: $p(xi) < 1 \Rightarrow \log p(xi)$ será negativo
- Base 2: Un fenómeno binario $\Rightarrow$ dos estados (bit)

Entropía

La entropía de un mensaje X, que se representa por H(X), es el valor medio ponderado de la cantidad de información de los diversos estados del mensaje.

$$H(X) = -\sum_{i=1}^{k} p(x_i) \log_2 p(x_i)$$

- Es una medida de la incertidumbre media acerca de una variable aleatoria y el número de bits de información.

Propiedades de la Entropía

- Toma valores no negativos.
- $H = 0 \; ssi \;\; Px_i = 1$
- La entropía será máxima (hay mayor incertidumbre del mensaje) cuando exista una equiprobabilidad en todos los valores de la variable X.

$H(X) = - \Sigma \; p_i \log_2 p_i = - n(1/n) \log_2 (1/n) = - (\log_2 1 - \log_2 n)$

$H(X)_{máx} = \log_2 n$ ← Valor máximo.

Codificador Óptimo

- Codificador óptimo es aquel que para codificar un mensaje X usa el menor número posible de bits.

$$H(X) = \sum_i p(x) \log_2 [1/p(x)]$$

- La expresión $\log_2 [1/p(x)]$ representará el número necesario de bits para codificar el mensaje x en un codificador óptimo.

Codificador Óptimo y Entropía

Calcular el número de bits óptimo de codificación para el mensaje

M = LELA ELLA de 8 caracteres.

Solución:

p(L) = 0,5; p(E) = 0,25; p(A) = 0,25; y obviamente Σ p(L, E, A) = 1,0.

Para codificar L necesitaremos 1 bit: $\log_2 [1/ P(L)] = \log_2 2 = 1$

Para codificar E necesitaremos 2 bits: $\log_2 [1/ P(E)] = \log_2 4 = 2$

Para codificar A necesitaremos 2 bits: $\log_2 [1/ P(A)] = \log_2 4 = 2$

Luego, si L se codifica como 0, E como 10 y A como 11, el mensaje M se codificará como: 0 10 0 11 10 0 0 11, es decir se transmiten 12 bits.

H(M) = 1,5 y al mismo valor se llega con el concepto de número medio de bits: para codificar un mensaje M de 8 elementos, hemos usado 12 bits. Luego 12/8 = 1,5 bits por elemento.

Conclusiones

- La gestión de la seguridad del software está estrechamente relacionada con la cantidad de información disponible.
- La incertidumbre es un componente fundamental en la seguridad de un software.
- La teoría de la información es aplicable a:
  - La definición de claves.
  - La creación de estrategias de protección de los datos.
  - La definición de políticas de seguridad.

Control de Aprendizaje

1. Al despertar ponemos la radio y escuchamos noticias que no nos llaman la atención. ¿Por qué decimos que no había información?
2. ¿Por qué usamos la base 2 en el logaritmo que define c_i?
3. ¿Cuál es el número mínimo -e inteligente- de preguntas que hay que hacer para pasar de la incertidumbre a la certeza en un sistema de n estados equiprobables?
4. ¿Por qué la entropía es no nula y se anula si y sólo si uno de los estados de la variable es igual a la unidad?