

TEMA 2. CRIPTOGRAFÍA 2.1 INTRODUCCIÓN A LOS CRIPTOSISTEMAS

Criptografía y seguridad informática Seguridad en las tecnologías de la información @ COSEC LAB

Curso 2016-2017

ÍNDICE

2.1 Introducción a los criptosistemas

- Criptografía
 - Definición
 - Modelo de criptosistema
 - Características de los sistemas criptográficos
 - Codificadores vs cifradores
- Criptoanálisis
- Teoría de la información
 - Entropía
 - ► Entropía condicionada
- Aleatoriedad
- Complejidad algorítmica



Definición de criptografía

Definición clásica (2000 a.c – 1949)

Disciplina que estudia los principios, métodos y medios de transformar los datos para ocultar su significado

Definición moderna (desde 1976) (m les quevas.

Disciplina que estudia los principios, métodos y medios de transformar los datos para ocultar su significado, garantizar su integridad, establecer su autenticidad y prevenir su repudio

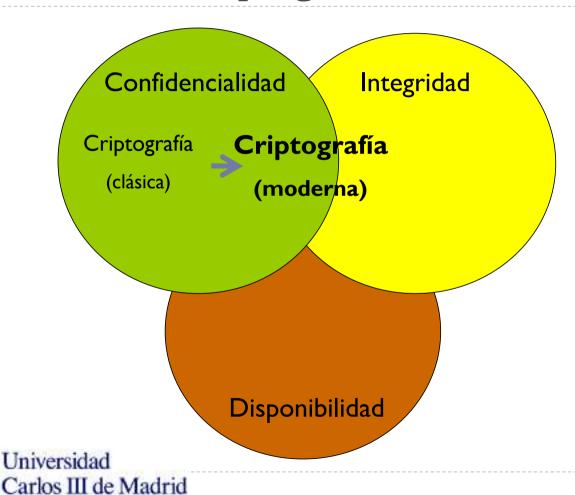
Geno han Si do maniplado L. Asegurar quien asido el autor

al tener clave privad



Definición de criptografía

COSEC LAB · Dpto. Informática



Modelo de criptosistema

Espacio de mensajes " Heusejes originales, sin tours former.

$$M = \{m_1, m_2, ...\}$$

Espacio de cifrados: Meniajes transformados

$$C = \{c_1, c_2, ...\}$$

- Espacio de claves: Sirve para configurar el algoritmo de cifrado /descifrado $K = \{k_1, k_2, ...\}$
- Familia de transformaciones de cifrado

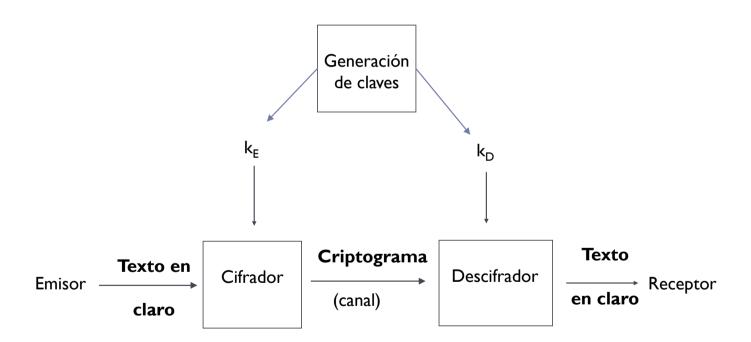
$$E_k: M \to C$$

Familia de transformaciones de descifrado

$$D_k:C\to M$$



Modelo de criptosistema



 $k_E y k_D$ pueden o no ser iguales



Características de los sistemas criptográficos

- ▶ Se caracterizan con tres dimensiones independientes:
 - Tipo de operaciones realizadas
 - En general, substituciones y transposiciones. No puede perderse información. Los más comunes usan el producto de varias ops.
 - Número de claves usadas
 - Simétricos o con una clave (también conocido como algoritmos de clave secreta)
 - Asimétricos o con dos claves (también conocido como algoritmos de clave pública) Se cifra con la publica se descifra cen la privada.
 - Tipo de procesamiento del texto en claro
 - Por bloques (algoritmos de cifrado en bloque) grandes tross.
 - Como un flujo continuo de bytes o de bits (algoritmos de cifrado en flujo)



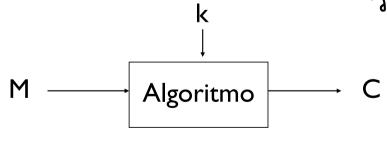
Codificadores vs Cifradores

Codificador: Se sique una función, que siempre en iguel. Ejen: Marse, ASCII



$$C=f(M)$$

Cifrador: Se signe un proceso que no conocido por todos, se necesito una clave.



$$C=E(k, M)=E_k(M)$$



Codificadores vs Cifradores





ÍNDICE

2.1 Introducción a los criptosistemas

- Criptografía
 - Definición
 - Modelo de criptosistema
 - Codificadores vs cifradores
- Criptoanálisis
- Teoría de la información
 - Entropía
 - ► Entropía condicionada
- Aleatoriedad
- Complejidad algorítmica



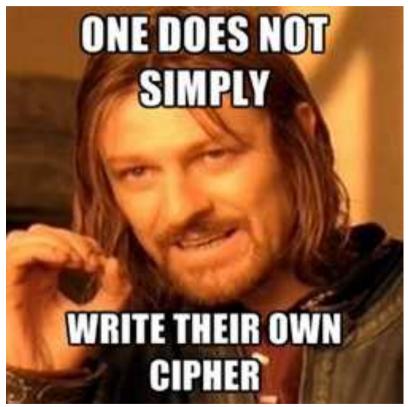
- Ciencia que trata de frustrar las técnicas criptográficas
 - Principio de Kerckoff

La seguridad del cifrado debe de residir, exclusivamente, en el secreto de la clave

La cryptographie militaire, 1883. Auguste Kerckhoffs von Nieuwenhof (1835-1903)

- No a la seguridad por falta de claridad
- Los ataques se basan en el conocimiento del algoritmo y, quizá, en información adicional sobre el texto en claro





- Dbjetivo del criptoanalista:
 - Principal: Recuperar la clave de descifrado 2 1 Junes
 - Secundario: Descifrar un texto cifrado concreto
- Aproximaciones del criptoanalista/atacante:



Ataque de fuerza bruta



Ataques al algoritmo

Carlos III de Madrid COSEC LAB · Dpto. Informática

Ataque	Conocido por el atacante (además de algoritmo)	Dificultad		
Texto cifrado	Criptograma	+4		
Texto en claro conocido	Criptograma + uno o más pares (texto en claro, texto cifrado) con la misma clave			
Texto en claro escogido	Criptograma + uno o más pares (texto en claro escogido, texto cifrado) con la misma clave			
Texto cifrado escogido	Criptograma + uno o más criptogramas escogidos por el atacante junto con sus correspondientes textos en claro, con la misma clave			
Texto escogido	Criptograma + uno o más pares (texto en claro escogido, texto cifrado) con la misma clave + uno o más criptogramas escogidos por el atacante junto con sus correspondientes textos en claro, con la misma clave	-		
(III) Universidad				

- Algoritmo de cifrado incondicionalmente seguro
 - No se filtra información adicional a la conocida por el atacante independientemente de la longitud del texto cifrado C
 - Solo el cifrador de Vernam es incondicionalmente seguro

 Lo Es el unico que lo cumple, pero en ciertas andiciones funciona.

 Selo quedoria la fuerza brota
- Algoritmo de cifrado matemáticamente vulnerable
 - Si al aumentar la longitud de C se filtra información
 - El resto de algoritmos de cifrado excepto Vernam son matemáticamente vulnerables

Cifrado de Vernam. One-time-pad

Cifrado:
$$E(M) = M \oplus K = m_1 \oplus k_1, m_2 \oplus k_2, ..., m_n \oplus k_n$$

$$1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 1$$

$$0 \ 0 \ 1 \ 0 \ 0 \ 1$$

$$1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0$$

$$E(M) = M \oplus K = m_1 \oplus k_1, m_2 \oplus k_2, ..., m_n \oplus k_n$$

$$M_{\text{tailed}} \text{ or } i$$

$$M_{\text{tailed}} \text{ or } i$$

$$G \ 0 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0$$

$$C \ i_{\text{tailed}} \text{ or } i$$

- Descifrado: $M = E(M) \oplus K$
- Shannon demostró que el cifrado de Vernam es incondicionalmente seguro si la clave K:
 - Es realmente aleatoria
 - Se usa una sola vez



Es de longitud igual o mayor que M ~ Tiere que hale dreus Universidad

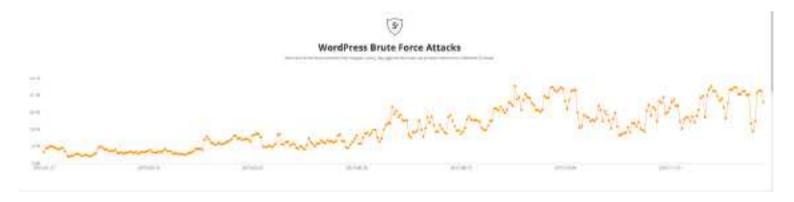
- Los cifradores incondicionalmente seguros, como Vernam, NO SON PRÁCTICOS
- > Seguridad computacional (o "No es vulnerable en la práctica"): Para ere origina andisis se requierent operaciones y dificulto en realizable

 en mayor que el tiempe visl. Parlo que y a da segual.

 El criptoanálisis del sistema requiere al menos toperaciones
 - - El tiempo de criptoanalizar el algoritmo excede el tiempo de vida útil de la información
 - El coste de criptoanalizar el algoritmo excede el valor de la información
- Para cifradores simétricos
 - No existe un algoritmo capaz de criptoanalizar el cifrador con una complejidad menor que la de un ataque de fuerza bruta



- Ataque de fuerza bruta
 - Probar todas las claves posibles
 - En media, se deben probar la mitad de las posibilidades para tener éxito

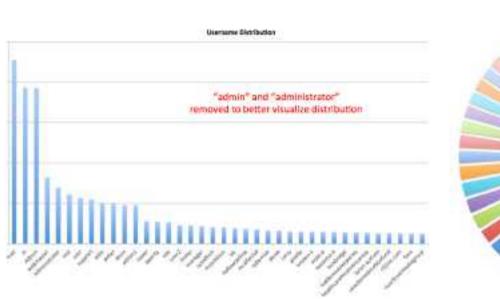


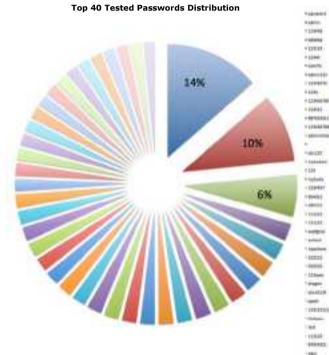
Número de intentos fallidos (Sitios WordPress protegidos por Sucuri)

https://sucuri.net/security-reports/brute-force/



Ataque de fuerza bruta





https://blog.sucuri.net/2014/03/understanding-denial-of-service-and-brute-force-attacks-wordpress-joomla-drupal-vbulletin.html



```
Trying username: 'ashishl' with password: '1212
- failed to login as 'ashishl' with password '1212

    Trying username: 'ashishl' with password: '123321'

  failed to login as 'ashishl' with password '123321'
  Trying username: 'pelowo' with password: '12121
  failed to login as 'gelowo' with password '12121'
  Trying username: 'gelowo' with password: 'asdad
  failed to login as 'gelowo' with password 'asdad'
- Trying username: 'gelowo' with password: 'asdasd'
  failed to login as 'gelowo' with password 'asdasd'
- Trying username: 'gelowo' with password: 'asdas
- failed to login as 'gelowo' with password 'asdas'
- Trying username: 'gelowo' with password: '1212'
  failed to login as 'gelowo' with password '1212'
  Trying username: 'gelowo' with password: '123321'
  failed to login as 'gelowo' with password '123321'
  Trying username: 'gelowo' with password: 'hello
  failed to login as 'gelowo' with password 'hello'
  Trying username: 'root' with password: '12121'
  Trying username: 'root' with password: 'asdasd
  failed to login as 'root' with password 'asdasd'
  failed to login as 'root' with password 'asdas'
- Trying username: 'root' with password: '1212
                          with password: "123321"
                            with password '123321'
                            th password: 'hello'
```



Tiempo medio requerido para realizar una búsqueda exhaustiva de la clave (ataque de fuerza bruta)

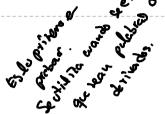
Suposición razonable

Supuesto procesamiento masivo paralelo

Tamaño de la clave (bits)	Número de claves posibles	Tiempo requerido supuesto 1 descifrado/µs	Tiempo requerido supuesto 10 ⁶ descifrados/µs
32	$2^{32} = 4.3 \cdot 10^9$	$2^{31}\mu s = 35,8 \text{ minutos}$	2,15 milisegundos
56	$2^{56} = 7.2 \cdot 10^{16}$	$2^{55}\mu s = 1142 \text{ años}$	10,01 horas
128	$2^{128} = 3.4 \cdot 10^{38}$	$2^{127}\mu s = 5.4 \cdot 10^{24} \text{ años}$	5,4 · 10 ¹⁸ años
168	$2^{168} = 3.7 \cdot 10^{50}$	$2^{167}\mu s = 5.9 \cdot 10^{36} \text{ años}$	5,9 · 10 ³⁰ años
26 caracteres (permutación)	$26! = 4 \cdot 10^{26}$	$2 \cdot 2^{26} \mu s = 6,4 \cdot 10^{12}$ años	6,4 · 10 ⁶ años



Ataque de diccionario



DICTIONARY ATTACK!

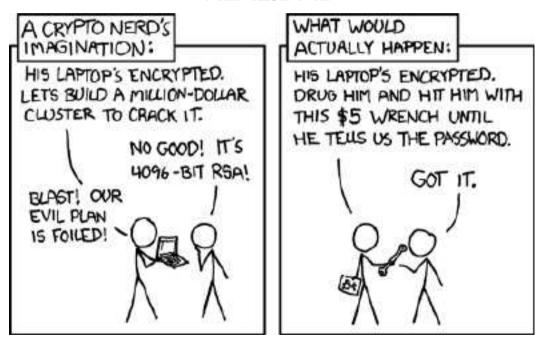






Seguridad incondicional vs seguridad computacional vs...

REALIDAD





https://xkcd.com/538/

▶ O lo que es peor...



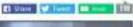


RISK ASSESSMENT / SECURITY & HACKTIVISM

Hacked French network exposed its own passwords during TV interview

Post-if note on wall remailed network's passwords for YouTube, instagram

by sain Machineoth - Apr 18-2016 to 57 am CEST







They recgo. David Deep.

O III

While French authorities continued investigating how the TVIM and exelviors had TL of its statutors. signals interrupted the right before, one of its staffers proved just how likely a basic positivior their. might have led to the incident.

http://arstechnica.com/security/2015/04/hacked-french-network-exposed-its-own-passwords-during-tv-interview/



ÍNDICE

2.1 Introducción a los criptosistemas

- Criptografía
 - Definición
 - Modelo de criptosistema
 - Características de los sistemas criptográficos
 - Codificadores vs cifradores
- Criptoanálisis
- Teoría de la información
 - Entropía
 - ► Entropía condicionada
- Aleatoriedad
- Complejidad algorítmica



TEORÍA DE LA INFORMACIÓN

- Bases matemáticas. (Claude E. Shannon)
 - A mathematical theory of communication, Bell Syst. Tech. J., vol.23. 1948
- Fundamentos teóricos de la criptografía: Criptología científica





TEORÍA DE LA INFORMACIÓN

- Establece una métrica para evaluar el secreto de un cifrador
- > Se basa en la incertidumbre que sobre el texto en claro tiene un criptoanalista que intercepta un texto cifrado

Cifrador incondicionalmente seguro

No se filtra nada, independientemente de la longitud de C (Vernam)

Cifrador matemáticamente vulnerable

Cuanta mayor sea la longitud de C, mayor cantidad de información se filtra (y por tanto está disponible para el criptoanalista)



CANTIDAD DE INFORMACIÓN

Sea M={m₁, m₂, ..., m_n} una fuente de mensajes estadísticamente independientes cuyas probabilidades de ocurrencia respectivas son:

Preb.

October p(m₁),..., p(m_n) con
$$\Sigma$$
 p(m_i)=1

La cantidad de información (c_i) de un mensaje m_i es:

A mayor p(mi), menor ci



Entropía de una fuente M es la cantidad promedio de información transportada por un mensaje perteneciente a dicha fuente

Entropía de la fuente M:

$$H(M) = -\sum_{suma jai perijilidad perijo} | bits$$

Bit: entropía de una fuente con 2 mensajes equiprobables

$$-(1/2 \log_2 \frac{1}{2} + 1/2 \log_2 \frac{1}{2}) = 1/2 \log_2 2 + 1/2 \log_2 2) = 1$$
 bit



- Previsible ganar tras la aparición de un mi
- La entropía de M mide la incertidumbre que, a priori, tiene un observador acerca de la aparición de un m_i
- A mayor entropía, mayor incertidumbre sobre M

Entropía cero = incertidumbre cero = $p(m_i)$ = I para algún i



- Sea M={m₁, m₂, ..., m_n} con Σ p(m_i)=1
- Propiedades
- 0 \le H(M) \le log_n minime siso prob. de apari es /(segro)
 - 2. H(M)= 0 si y sólo si p(m_i)= I para algún i
 - $H(M) = \log_2 n$ si y sólo si $p(m_i) = 1/n$ para $1 \le i \le n$ maximo solosi son equipobables



Fig. Considere una fuente con 2 elementos $M = \{m_1, m_2\}$ con $p(m_1) = 1/3$ y $p(m_2) = 2/3$. Calcule la entropía de M $H(M) = - Σ p(m_i) log_2 p(m_i) = 1/3 log_2 3 - 2/3 log_2 2/3 = 0.52 + 0.38 = 0.9$

Fig. Considere una fuente con 2 elementos $M = \{m_1, m_2\}$ con $p(m_1) = 0.4$ y $p(m_2) = 0.6$. Calcule la entropía de M $H(M) = - Σ p(m_i) log_2 p(m_i) = -0.4 log_2 0.4 - 0.6 log_2 0.6 = 0.52 + 0.44 = 0.96$



ENTROPÍA CONDICIONADA

- Cuando existe alguna relación entre las apariciones de dos mensajes consecutivos n; (de una fuente N) y m; (de una fuente M), la presencia del primero disminuye la incertidumbre del segundo

 Conoar derinaro ne falliba encuentrar d'aquado, de reduce mi incertido del incertido del segundo.
- La entropía de M condicionada por N, **H(M|N)**, se define como el valor medio de la cantidad de información de M conocido N

$$H(M|N) = -\sum_{j} p(n_{j}) \sum_{i} p(m_{i}|n_{j}) \log_{2} p(m_{i}|n_{j})$$



ENTROPÍA CONDICIONADA

- ▶ Ej. M= $\{m_1, m_2, m_3, m_4\}$, p (m_1) =p (m_2) =p (m_3) = p (m_4) =1/4 y N= $\{n_1, n_2\}$, p (n_1) =p (n_2) =1/2. N= n_1 ⇒ M= m_1 ó m_2 (equiprobablemente) N= n_2 ⇒ M= m_3 ó m_4 (equiprobablemente)
- ► H(M)=2 y
- $H(M|N) = \frac{1}{2}(\frac{1}{2} \lg_2 2 + \frac{1}{2} \lg_2 2) + \frac{1}{2}(\frac{1}{2} \lg_2 2 + \frac{1}{2} \lg_2 2) = 1$
- El conocimiento de N hace disminuir la entropía resultante de M



ENTROPÍA CONDICIONADA

- Los métodos criptográficos tratan de maximizar H(M|N) siendo M el conjunto de textos en claro y N el de los cifrados
- ▶ Todos los cifradores (menos Vernan) filtran alguna información sobre el texto en claro al texto cifrado, y según la longitud del texto cifrado crece, mayor es la información filtrada

ÍNDICE

2.1 Introducción a los criptosistemas

- Criptografía
 - Definición
 - Modelo de criptosistema
 - Características de los sistemas criptográficos
 - Codificadores vs cifradores
- Criptoanálisis
- Teoría de la información
 - Entropía
 - ► Entropía condicionada
- 18/61 Aleatoriedad
 - Complejidad algorítmica

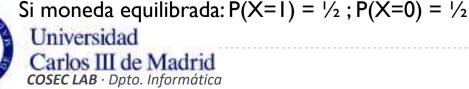


VARIABLE ALEATORIA

- Sea S un espacio muestral con distribución de probabilidad P (cada posible valor que X puede tomar en S tiene asociada una determinada probabilidad)
- ▶ Una variable aleatoria X es una función de S al conjunto de los números reales $X : S \longrightarrow E = |R|$
- Ejemplo de variable aleatoria discreta

$$S = \{cara, cruz\}$$

$$X(s) = \begin{cases} I & \text{si } s = cara \\ 0 & \text{si } s = cruz \end{cases}$$



SECUENCIA ALEATORIA

Múltiples usos

- Distribución de claves
- Protocolos de autenticación mutua
- Generación de claves de sesión
- Generación de claves para RSA
- Generación de flujos de bits para algoritmos de cifrado simétrico de flujo
- Criterios de aleatoriedad:
 - Distribución uniforme: La frecuencia de aparición de l's y

 0's debe ser aproximadamente la misma

 Se delse comporter de estamente la misma
 pavana pode de delicital resto de la secuencia.
 - Independencia: Ninguna subsecuencia puede ser inferida de otras



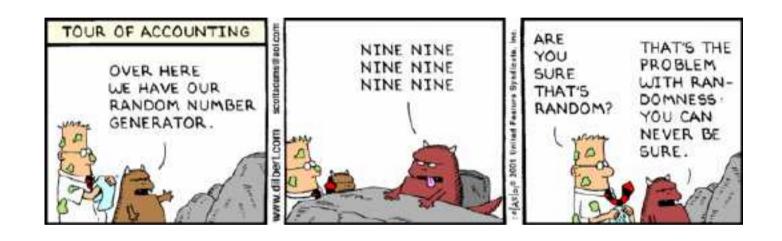
SECUENCIA ALEATORIA andita Paent repin

- Baterías de tests
 - Existen test para probar distribución uniforme
 - No existen test para probar independencia Existen test para demostrar la no independencia esindependencia

 - Si no pasa tests, aleatoriedad descartada
 - Si pasa todos los tests, no se puede garantizar aleatoriedad
 - ▶ El Maurer Universal Test no es definitivo



SECUENCIA ALEATORIA



http://dilbert.com/strip/2001-10-25



ALEATORIEDAD SECUENCIA ALEATORIA

Generadores pseudoaleatorios: pascu las pruebas pero fiesencu vecanísmo. Olganismo

- Las aplicaciones criptográficas generalmente utilizan algoritmos para generar números "aleatorios"
- Aunque una secuencia verdaderamente aleatoria no puede estar generada por un algoritmo dado que éste por definición es determinista
- Diferencia:
 - Pseudoaleatoriedad (PRNG)
 - Algoritmo
 - Aleatoriedad (TRNG) [Uso de fuentes no deterministas]
 - Fuente de entropía tomadad de ciertos procesos naturales
 - Eliminación del sesgo con funciones resumen



ÍNDICE

2.1 Introducción a los criptosistemas

- Criptografía
 - Definición
 - Modelo de criptosistema
 - Codificadores vs cifradores
- Criptoanálisis
- Teoría de la información
 - Entropía
 - ► Entropía condicionada
- Aleatoriedad
- Complejidad algorítmica

COMPLEJIDAD ALGORÍTMICA

 Campo de la matemática que estudia los algoritmos bajo la dificultad de su resolución

Clasifica los algoritmos según su complejidad



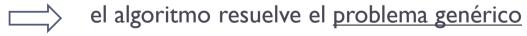
PROBLEMAS Y ALGORITMOS

Problema

Planteamiento de una tarea en un determinado contexto

Algoritmo

- Conjunto finito de <u>operaciones</u>, que realizadas en un determinado <u>orden</u>, resuelven un <u>problema</u>
- Los algoritmos pueden trabajar sobre un ejemplo particular de problema (problemas particulares)
- Si un algoritmo resuelve todos los problemas particulares





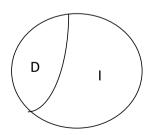
PROBLEMAS Y ALGORITMOS

- Turing demuestra que no todos los problemas tienen un algoritmo que los resuelva
- > ¡No todos los problemas tienen solución!



PROBLEMAS Y ALGORITMOS

- Una primera clasificación de los problemas
 - Indecidibles (I)
 - No resolubles mediante un algoritmo
 - Decidibles (D)
 - Cuentan con al menos un algoritmo para su resolución

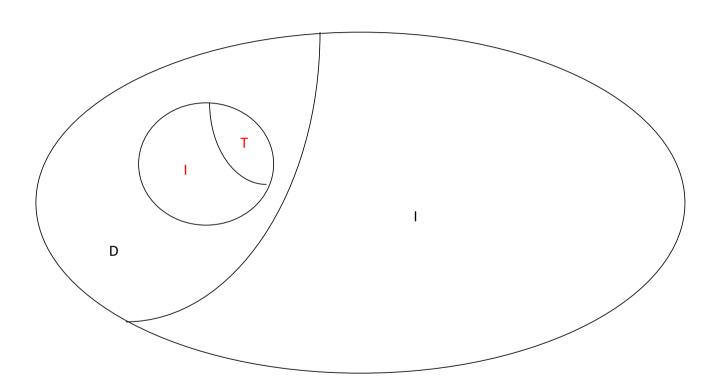


Problemas Tratables e Intratables

- Existen problemas cuya solución es inabordable por el elevado número de operaciones a realizar
- Una segunda clasificación de los problemas:
 - Intratables (I)
 - No es factible obtener su solución en un tiempo razonable con potencia de cálculo actual
 - Tratables (T)
 - Existen al menos un algoritmo que resuelve cualquier problema particular en tiempo razonable



Problemas Tratables e Intratables





TIEMPO DE EJECUCIÓN

- La dificultad para resolver una instancia de un problema se mide según su tiempo de ejecución (t)
- Es función del tamaño de la entrada (n)
- Se analiza el comportamiento del algoritmo cuando n crece (comportamiento asintótico)
 - Se dice que un algoritmo presenta una complejidad polinómica si el tiempo t es de orden polinómico o menor
 - Logarítmico $O(\log n)$: Ej. t = 5 log n $O(\log n)$
 - Potencia de n (polinómico) $O(n^c)$: Ej. $t = 2n^3 + 6n$ $O(n^3)$
 - vs. complejidad exponencial si el tiempo t es de orden mayor que polinómico
 - Exponencial $O(c^n)$: Ej. $t = 3^n + 4n$ $O(3^n)$
 - Factorial O(n!): Ej. t = $5n! + 6^n$ O(n!)



COSEC LAB · Dpto. Informática

TIEMPO DE EJECUCIÓN

▶ En un ordenador con l millón de operaciones por segundo

Tamaño n	log ₂ n (t)	n (t)	n ² (t)	2 ⁿ (t)
10	3·10 ⁻⁶ s	10 ⁻⁵ s	10 ⁻⁴ s	10 ⁻³ s
10 ²	7·10 ⁻⁶ s	10 ⁻⁴ s	10 ⁻² s	10 ¹⁴ siglos
10 ³	10·10 ⁻⁶ s	10 ⁻³ s	1 s	Muy grande
10 ⁴	13·10 ⁻⁶ s	10 ⁻² s	1,7 min	Muy grande
10 ⁵	17·10 ⁻⁶ s	10 ⁻¹ s	2,8 h	Muy grande



CLASES DE COMPLEJIDAD **ALGORÍTMICA**

- Un problema puede resolverse por distintos algoritmos
- Los problemas se clasifican en clases de complejidad según el tiempo en el que pueden ser resueltos:
 - ► Clase P (Polynomial time) polinomial con algoritmo deforminista
 - Clase NP (Non deterministic Polynomial time) polinemial
 - Otras clases...



CLASES DE COMPLEJIDAD ALGORÍTMICA

- Problemas de Clase P (Polynomial time)
 - Son problemas <u>Tratables</u>
 - Se resuelven mediante <u>algoritmos polinómicos</u> (buenos algoritmos)
 - Los algoritmos utilizados son deterministas
 - En cada paso de computación se determina de forma única el siguiente paso
 - La concatenación de dos algoritmos P es otro algoritmo P



CLASES DE COMPLEJIDAD ALGORÍTMICA

- Problemas de clase NP (Non deterministic Polynomial time)
 - Contiene problemas <u>Intratables</u> (y tratables)
 - → ;P ⊂ NP?
 - Los problemas intratables se resuelven mediante algoritmos no polinomiales (malos algoritmos), como los exponenciales
 - Los algoritmos utilizados son no deterministas
 - En cada paso de computación necesitan una selección entre diferentes opciones
 - Ejemplos:
 - □ Problema del logaritmo discreto → DH, EG
- - Universidad

COSEC LAB · Dpto. Informática