Criptografia i seguretat de dades

per Llorenç Huguet

La paraula telemàtica, sortida de l'informe de Nora i Minc publicat el 1978, que avui tothom coneix malgrat que la seva definició formal no és gaire assumida, es refereix a l'aliança entre la informàtica i la telecomunicació. Si l'espina dorsal de la telemàtica la constitueixen les xarxes de comunicació de dades (especialment digitals), és clar que la protecció del flux i el contingut de la informació transmesa ha de ser assegurada. Aquest article presenta algunes solucions per a la protecció del secret i l'autenticitat de la informació tractada en una xarxa de sistemes distribuïts.

Llorenç Huguet i Roiger va néixer a Ferreries (Menorca) l'any 1953. Ha cursat estudis en les Universitats Autònoma de Barcelona i Catòlica de Louvain (Bèlgica). Llicenciat en Matemàtiques i ciències aplicades en informàtica i gestió industrial. El 1981 es va doctorà en ciències per la UAB on actualment és professor adjunt del departament d'informàtica.

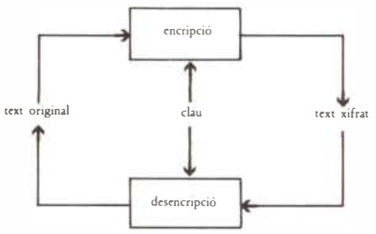
Mots bàsics en criptografia:

CRIPTOGRAFIA: ciència i art d'escriure, tot guardant inintel·ligible, el contingut del text escrit.

ENCRIPCIÓ: procés de transformació del text original (*cleartext*) en el text xifrat (*ciphertext*).

DESENCRIPCIÓ: procés de transformació del text xifrat en el text original.

CLAU: paràmetres que controlen els processos d'encripció i de desencripció.



CRIPTOANÀLISI: ciència i estudi dels mètodes de descobrir els textos xifrats.

CRIPTOLOGIA: coneixement que engloba la criptografia i la criptoanàlisi.

L'evolució dels sistemes informàtics, des de fa ja algun temps, apunta cap a la utilització de xarxes de sistemes substituint el clàssic sistema informàtic enclaustrat en un recinte tancat, i aïllat o individual. Avui ja es parla de sistemes distribuïts.

Per sistema entenem tant ordinadors, com terminals i/o concentradors, etc.; però la veritable importància està en el fet de ser distribuïts, és a dir, amb tendència a compartir tota classe de recursos: *hardware,* discos, cintes, processadors, memòria, informació emmagatzemada en fitxers, etc. Es pot pensar ja en sistemes interconnectats i cadascun d'ells realitzant una tasca específica que podrà ser aprofitada per tots els usuaris interconnectats a la xarxa. Però aquesta possibilitat de compartir recursos planteja el problema de la seguretat i autenticitat de la informació. Seguretat pel que fa la confidencialitat de certa informació, i autenticitat o integritat referent al valor de la informació tractada; protegint-la de malintencionades manipulacions.

A tall d'il·lustració, estudiem un cas típic de xarxa de sistemes i vegem els problemes que se'ns plantegen respecte a la seguretat i autenticitat (vegeu fig. 1).

En una xarxa de sistemes, un element imprescindible és el mitjà de la interconnexió, que en el llenguatge informàtic es coneix pel nom de línia de comunicació, i dels quals n'existeixen de diferents tipus: telefònica, fibra òptica, via satèl·lit, etc. Un dels primers fronts d'atac per part d'un hipotètic "espia" seria la línia de comunicació, que pot interceptar qualsevol missatge transmès entre els diferents components de la xarxa.

Si un usuari vol treballar amb la xarxa, aquesta li exigeix la seva identificació (LOGIN), actualment ho sol·licita amb un porocés d'identificació mitjançant "mots" de pas (*passwords*). D'aquesta manera es mantenen certs privilegis per a certs usuaris per poder accedir a un determinat tipus d'informació, o poder llançar certs programes, connectar-se amb d'altres xarxes, etc. I aquest seria el segon front a protegir, perquè el coneixement, per part de l'"espia", del mot de pas d'algun usuari el possibilitaria per poder realitzar certes funcions associades només a aquell usuari.

Per altra banda, dintre d'un sistema s'ha de protegir tant programes com informació en les bases de dades per tal d'evitar que algú estrany a la xarxa, o al mateix sistema, pugui executar aquells programes o accedir a la informació que eren reservats, com podria ser les taules del sistema, per exemple. Aquest seria el tercer front a protegir.

En resum, el nostre objectiu davant el secret i autenticitat associada a una xarxa de sistemes és el d'evitar que un "espia" pugui violar o eliminar la protecció del sistema pel que fa a:

–línies de comunicació
–connexió a la xarxa (mots de pas)
–utilització dels recursos d'un determinat sistema

És la criptografia, aquesta nova ciència que en els darrers anys ha tingut un reguany d'interès, tant teòric com pràctic, qui s'ocupa de la protecció dels missatges transmesos sobre una línia de comunicació i de la informació emmagatzemada en sistemes digitals; contemplant els aspectes de seguretat i autenticitat esmentats abans.

Pel que respecta a la informació transmesa sobre una línia de comunicació, considerem dos tipus de vulneracions: la produïda per un "espia" passiu (*passive wiretapping* o *eavesdropping*), el qual només escolta els missatges transmesos, i la produïda per un "espia" actiu (*active wiretapping* o *tampering*), que introdueix modificacions sobre la informació que s'està transmetent; tal com queda il·lustrat en la fig. 2.

Obviament, el mètode d'encripció utilitzat (vegeu el vocabulari) haurà de protegir tant l'escolta i possibles modificacions de la informació transmesa com possibles insercions o esborrades de missatges.

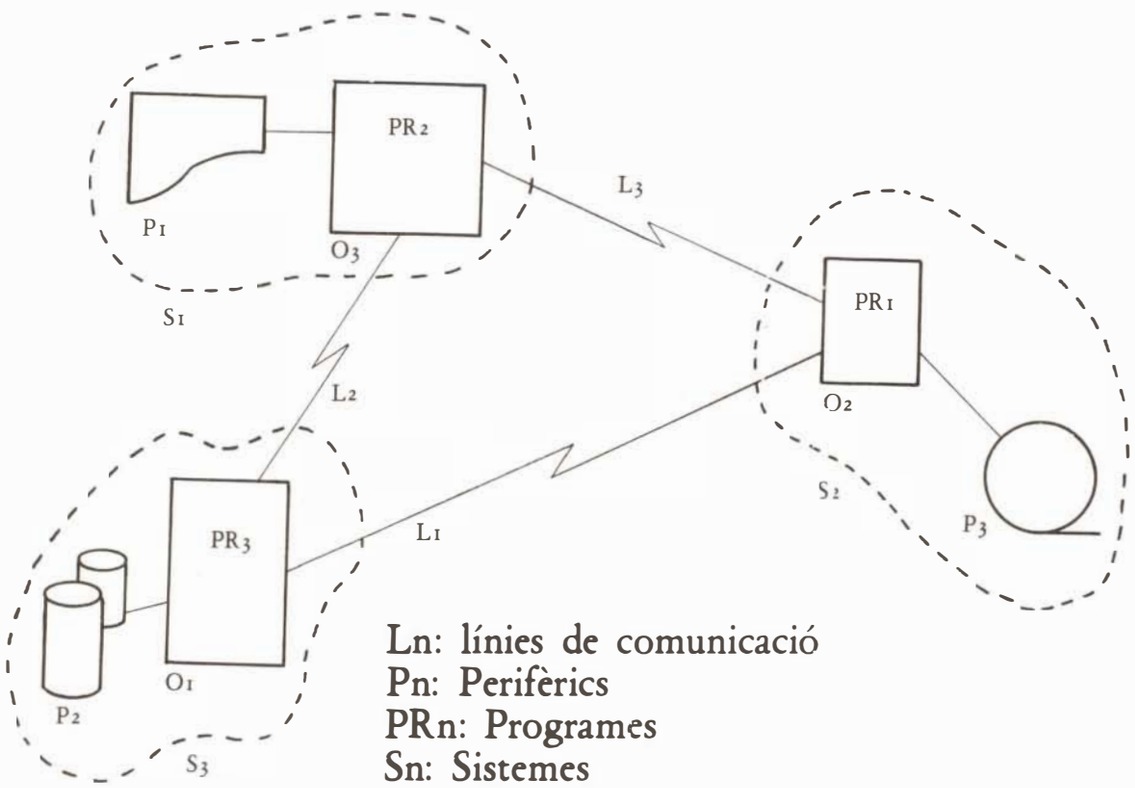
Per altra banda, la vulnerabilitat dels sistemes d'ordinadors, com a processadors d'informació, pot ser atacada, pel que fa al secret, mitjançant:

1- **recerca d'informació** en memòria principal o auxiliars. En aquest cas són

i Clemente Rodríguez

Clemente Rodríguez Lafuente va néixer a Miranda del Ebro (Burgos) l'any 1957. És llicenciat amb grau per la Universitat Autònoma de Barcelona i actualment és professor contractat del departament d'informàtica de la mateixa universitat.

Fig. 1: Cas típic de xarxa de sistemes



dues les diferències amb les operacions realitzades sobre una línia de comunicació:

–la informació emmagatzemada té una vida més llarga que no pas la informació que es transmet

–l'accés a la línia és gairebé immediata, malgrat la utilització d'accessos de control per tal de limitar els camps de treball de cada usuari

nota: cal observar que la criptografia ens podrà evitar que aquesta informació emmagatzemada sigui legible, però no evitarà possibles manipulacions de distorsió.

2- **transmissió de dades** a usuaris no autoritzats, mitjançant processos amb accés legítim a aquestes dades. Per evitar-ho caldrà suplementar la criptografia i el control d'accessos mitjançant informació de control de flux.

3- **utilització d'estadístiques** per tal de deduir, a partir de dades globals, dades confidencials concernent un individu. Per evitar aquest atac, necessitarem controls especials de la informació.

4- **modificació d'informació guardada.** La criptografia ens podrà ajudar a detectar aquests possibles canvis, però no serà una mesura preventiva. Seran necessaris accessos de control i mètodes de *backup* per fer el *recovery*.

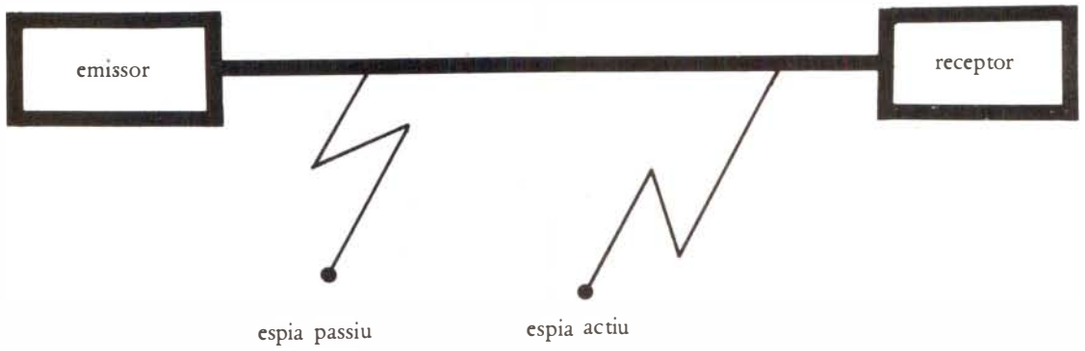


Fig. 2: Vulneracions d'un canal de transmissió

5- **destrucció accidental.**
Aquí la criptografia no ens ajuda ni molt ni poc. També seran necessaris mètodes de *backup.*

6- **accés al sistema** per compte d'un altre usuari, per poder accedir als seus fitxers i altra informació particular.
Un sistema d'encripció de mots de pas i de signatures digitals evitaria aquest tipus d'atac.
Vistes les possibilitats de manipulacions no desitjades que com a usuaris d'un sistema o xarxa podem sofrir, presentem ara les nocions i característiques d'un sistema criptogràfic.

Sistema criptogràfic

Tot sistema criptogràfic, també dit criptosistema, constarà de cinc components: M,C,K,E i D.
M, que serà el conjunt de missatges originals que volem transmetre. C, que serà el conjunt de missatges xifrats. K, que serà el conjunt de claus utilitzables. E, que serà el conjunt de mètodes d'encripció E_k: M — C per tot $k \varepsilon K$. D, que serà el conjunt de mètodes de desencripció D_k: C — M per tot $k \varepsilon K$.
Cada transformació o mètode d'encripció de E és definida per un algorisme, el qual és comú per a totes les transformacions de E, i per a una clau $K \varepsilon K$, la qual distingirà l'algorisme corresponent a cada transformació E_k.
La mateixa consideració és certa per a les transformacions D_k de D. Per a una clau donada K, la transformació D_k és la inversa de E_k; és a dir: $D_k (E_k (M)) = M$ per a tot missatge $M \varepsilon M$.
Tot criptosistema, tal com mostra la figura precedent, ha de complir, almenys, aquests tres requisits:

1.- Totes les transformacions E_k i D_k han de ser "fàcilment" calculables.
2.- Els algorismes de les transformacions E_k i D_k han de ser "fàcilment" implementables.
3.- La seguretat del sistema solament pot dependre del secret de les claus $k \varepsilon K$, i no dels algorismes de les transformacions corresponents de E i D;
i sempre tenir en compte els objectius de "seguretat" i "autenticitat", cristal·litzats amb els aspectes que ara considerem:
"seguretat": incapacitat, per a un criptoanalista, de determinar el text original del text xifrat que s'hagi pogut interceptar.

Aquest objectiu de seguretat exigeix dos requeriments:

1.- Que, des d'un punt computacional, sigui impossible, per a un criptoanalista, determinar la transformació de desencripció D_k a partir del missatge xifrat C, tot i conèixer el missatge original M.
2.- Que, des d'un punt de vista computacional, sigui impossible, per a un criptoanalista, determinar el missatge original M, sistemàticament, de la sola intercepció del missatge xifrat C.

EL SECRET REQUEREIX ÚNICAMENT LA NO REVELACIÓ DE LA TRANSFORMACIÓ DE DESENCRIPCIÓ D_K.
mentre que E_k pot ser revelada públicament

"autenticitat": incapacitat, per a un criptoanalista, de substituir un text xifrat fals, C', en el lloc d'un text xifrat real C, sense que sigui detectat.
I formalment són també dos els requeriments de l'autenticitat:

1.- Que, des d'un punt de vista computacional, sigui impossible determinar la transformació E_k corresponent a un missatge xifrat C, encara que sigui conegut el missatg original M. $(C = E_k (M))$
2.- Que, des d'un punt de vista computacional, sigui impossible determinar, per a un criptoanalista, un missatge xifrat C' tal que $D_k(C') \varepsilon M$; és a dir, que el receptor pugui trobar, en desencriptar C', un missatge del seu conjunt M.

L'AUTENTICITAT REQUEREIX ÚNICAMENT LA NO REVELACIÓ DE LA TRANSFORMACIÓ D'ENCRIPCIÓ E_K.
mentre que D_k pot ser revelada públicament.

Evidentment, la conjunció de la "seguretat" i l'"autenticitat" ens exigirà la no revelació de totes dues transformacions d'encripció i desencripció. Malgrat aquestes precaucions sobre la indesxifrabilitat i/o la detecció de modificacions dels missatges xifrats, queda encara la protecció davant de possibles esborrats o eliminació de part de la informació. Aquest darrer cas haurà de ser tractat mitjançant un protocol de comunicacions entre l'emissor i el receptor. Com a exemple podem citar els protocols de seqüencialització de missatges en la transmissió de dades, com són els protocols HDLC, SDLC...
Actualment es consideren dos tipus de criptosistemes, segons la utilització i gerència de les transformacions d'encripció i desencripció. El criptosistema clàssic o convencional, en el qual la clau corresponent a l'encriptador i desencriptador és la mateixa, o en qualsevol cas l'una fàcilment deduïble de l'altra, dóna excel·lents resultats quan es tracta de la utilització de fitxers privats. En aquest cas, cada usuari disposa d'una parella E_A, D_A que li són particulars i cap altre usuari no té accés a la informació guardada si no disposa de E_A i D_A (vegeu fig. 4). En aquest cas podem parlar de "seguretat" i "autenticitat" de la informació de cada usuari, malgrat que sobre un

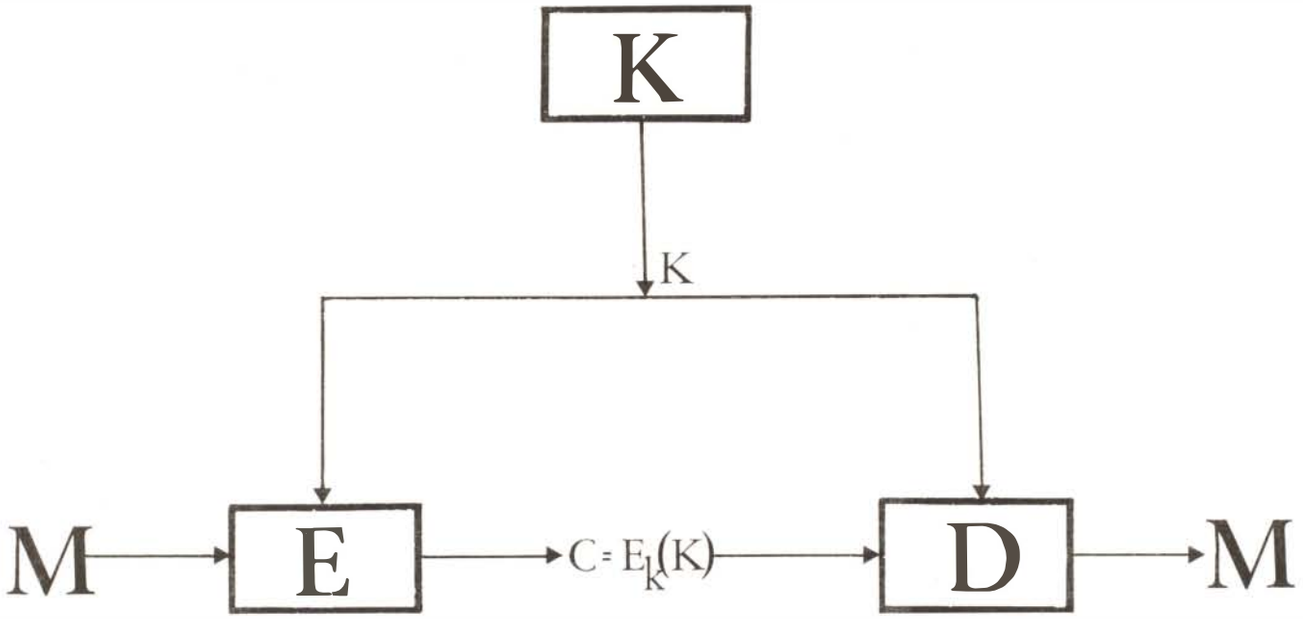


Fig. 3: Criptosistema

mateix fitxer hi tenen accés altres.

És el 1976 quan entra en joc un nou concepte de criptosistema, proposat per Diffie i Hellman, anomenat a doble clau (o criptosistema asimètric, ja que el cas clàssic també se'n diu criptosistema simètric), caracteritzat pel fet que conèixer la transformació E_k no dóna informació sobre la transformació D_k, o viceversa. És a dir, una de les claus E_k o D_k pot ser revelada públicament sense perill que l'altra sigui deduïda. Aquest segon enunciat ens permetrà la protecció de la informació fluctuant sobre un fitxer públic, de manera que escriure informació sobre ell no té per què implicar poder-la llegir. La clau d'escriptura i lectura, malgrat estar relacionades, roman independent des del punt de vista de trobarne una a partir de l'altra (vegeu fig. 5).

El propòsit de Diffie i Hellman és considerar un criptosistema a doble clau de manera que cada usuari disposa d'una parella de transformacions E_A i D_A de les quals una és feta pública i l'altra és privada. Aquest model de criptosistema rep el nom de criptosistema a clau pública. Estudiem-ne ara les propietats i requeriments, per acabar donant un suggestiu i fascinant exemple.

En un criptosistema a clau pública, cada usuari A disposa d'una transformació d'encripció E_A, la qual pot ser enregistrada en un director públic, i una transformació D_A, de la qual solament ell n'és coneixedor. La transformació D_A és descrita en termes d'una clau secreta, mentre que E_A per una clau pública mitjançant un algorisme o funció que no reveli, que sigui computacionalment impossible determinar, la transformació D_K (aquest tipus de funcions són conegudes pel nom de *one-way function*).

En aquest cas, el "secret" i l'"autenticitat" són proporcionats per algorismes diferents. Per exemple, si un usuari A vol enviar un cert missatge M a un usuari B, només ha de cercar, en el directori públic, la transformació d'encripció E_B i enviar el missatge xifrat $C = E_B(M)$. Si B vol desxifrar C, només ha de prendre la transformació de desencripció D_K, que tan sols ell coneix, i calcular $D_K(C) = D_K(E_K(M)) = M$, retrobant M. En aquest cas, el secret queda assegurat; això no obstant, l'autenticitat no queda protegida, ja que en conèixer E_B, qualsevol usuari podrà substituir el missatge xifrat C per un altre $C' = E_B(M')$, sense que B se n'adoni.

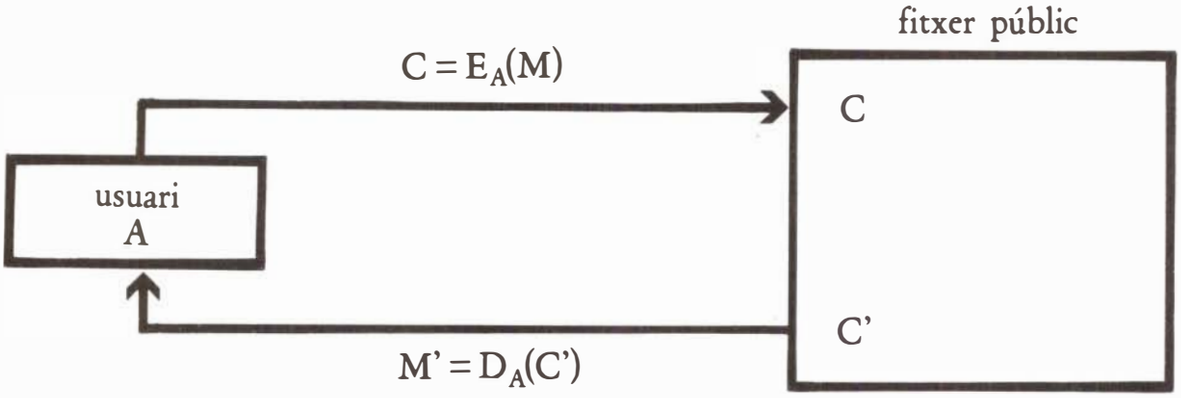
Per protegir l'autenticitat haurem d'exigir a la parella de transformacions E_K i D_K que verifiquin, a més, $E_K(D_k(M)) = M$ per a tot $M \varepsilon M$.

En aquest cas, l'usuari A pot signar els seus missatges mitjançant la transformació secreta D_A. En efecte, si A vol enviar un missatge M a B, certificant la seva procedència, podrà signar aplicant $D_A(M)$. Diguem que S és la signatura. Aleshores, el missatge xifrat que enviarà A és $C = E_B(S)$; el receptor B desencriptarà el missatge xifrat rebut mitjançant la seva transformació secreta D_B: $D_B(C) = D_B(E_B(S)) = S$; resultat que roman inintel·ligible per a B, però que pot disposar de la transformació E_A consignada al fitxer públic. Assabentat que el missatge li prové de A, només ha de calcular $E_A(S) = E_A(D_A(S)) = M$, per retrobar el missatge original que li volia transmetre A (vegeu fig. 5).

Clarament, existeixen moltes i importants aplicacions dels criptosistemes a clau pública, com el correu electrònic entre entitats bancàries, on el flux de divises necessita un secret i, sobretot, una autenticitat rigorosa.

L'exemple de criptosistema a clau pública que volem presentar és degut a Rivest, Shamir i Adleman, pertanyents al MIT de Standford. La validesa d'aquest criptosistema està, encara avui, en la dificultat de poder factoritzar, en producte de nombres primers, un cert nombre prudentment llarg, tot i saber que aquest és producte de dos nombres primers.

Pel que respecta al criptosistema R.S.A., les transformacions E_k i D_k són explicitades a la fig. 6, i òbviament compleixen els requisits de secret i autenticitat ($E_K(D_k(M)) = D_K(E_K(M)) = M$ per a tot $M \varepsilon M$). Observeu que els números n, e i



M i M' són missatges originals; C i C' són els corresponents missatges xifrats només intel·ligibles si es posseeix la parella E_A i D_A.

Fig. 4: Criptosistema tradicional

d són obtinguts de la següent manera: n és el producte de dos nombres primers p i q, i és aquest n el que farem públic, mentre p i q seran guardats en secret per tal de no revelar $\psi(n) = (p-1).(q-1)$; e i d seran dos números, un invers de l'altre a l'anell mòdul $\psi(n)$ (vegeu fig. 6).

L'article original presenta els corresponents algorismes per a cada una de les transformacions i eleccions de paràmetres, calculables de manera ràpida. Rivest, Shamir i Adleman, donen també l'exemple que transcrivim.

Consideren el cas $p = 47$, $q = 59$; aleshores

$n = 2773$ i $\psi(n) = 2668$.

Agafant $d = 157$, retroben, mitjançant els algorismes que presenten,

$e = 17$ ($17 \times 157 = 1$ mòd $\psi(n)$).

Mitjançant $n = 2773$ poden codificar dos lletres diferents per bloc, suposant cada lletra associada al seu ordre decimal dintre de l'alfabet

(A = 00, b = 01, ..., Z = 26).

El missatge original a encriptar és: *it's all greek to me;* és a dir:

0920 1900 0112 1200 0718 0505 1100 2015 0013 0500

Car e = 10001, en binari, el primer bloc 920 és encriptat:

$M = 920$; $E_k(M) = M^{17}$ mòd. 2773,

és a dir:

$x = (((((1)^2 M)^2)^2)^2)^2 M = 948$,

i repetint aquestes operacions trobaríem que el missatge global ja encriptat és:

0948 2342 1084 1444 2663 2390 0778 0774 0219 1655

El receptor desencriptaria els missatges xifrats mitjançant:

$x = 948$, $D_k(x) = x^{157}$ mòd 2773,

és a dir:

$M = 948^{157} = 920$, etc.

i fent la reconversió decimal a alfabètica retrobaria que el primer bloc és *it .. .. ..*

Llorenç Huguet i Clemente Rodríguez

Referències

W. Diffie, M.E. Hellman: *Privacy and Authentication: An introduction to Cryptograpfy.* Proceedings of the I.E.E.E. Vol 67 nº 3, March 1979.

M.E. Hellman: *The Mathematics of Public-Key Cryptography.* "Scientific American", August 1979.

R.L. Rivest, A. Shamir, L. Adleman: *A method for obtaining Digital Signatures and Public-Key Cryptosystems.* Sec. Com. A.C.M. 21 nº 2 Feb 1978.

L. Huguet: *Códigos Regulares: propiedades combinatóricas y aplicación al "Wire-Tap Channel".* Tesi doctoral U.A.B. (juny 1981)

D.E. Robling: *Cryptography and Data Security.* Addison-Wesley 1982.

C. Macchi, J.-F. Guilbert: *Téléinformatique.* Ed. Dunod (1979)

D.W. Dvies i altres: *Computer Networks and their Protocols.* John Wiley and Sons (1979).

P.E. Green: *Computer Network Architectures and Protocols.* Plenum Press. N.Y. (1982)

MÈTODE D'ENCRIPCIÓ I DECRIPCIÓ R.S.A.

CLAU PÚBLICA (n,e);

ENCRIPCIÓ DEL MISSATGE M:

$X = M^e$ mòdul n

FUNCIÓ DE DECRIPCIÓ (guardada secreta)

x^d mòdul n

Prenent e i d tals que $e.d = 1$ mòdul $(p-1).(q-1)$, resulta

$M^{e.d} = M$ mòdul n

El SECRET del sistema roman en la DIFICULTAT de factoritzar n.

La taula següent dóna el nombre d'operacions necessàries, i el temps invertit, per factoritzar un número n segons l'algorisme de Schroeppel.

dígits (decimals)	*n.º operacions*	*temps*
50	1.4×10^{10}	3.9 hores
75	9×10^{12}	104 dies
100	2.3×10^{15}	74 anys
200	1.2×10^{23}	3.8×10^{9} anys
300	1.5×10^{29}	4.9×10^{15} anys
500	1.3×10^{39}	4.2×10^{25} anys

bagatge matemàtic

$\varphi(n)$ és l'indicador d'Euler, nombre de nombres sencers i positius més petits que n i que són relativament primers amb ell.

$\varphi(p) = p-1$ si p és un nombre primer

$\varphi(n) = (p-1).(q-1)$

Si d és relativament primer amb $\varphi(n)$, aleshores d posseeix una inversa e en l'anell dels sencers mòdul $\varphi(n)$: $e.d = 1$ mòdul $\varphi(n)$

Per a tot M es compleix

Fig. 6

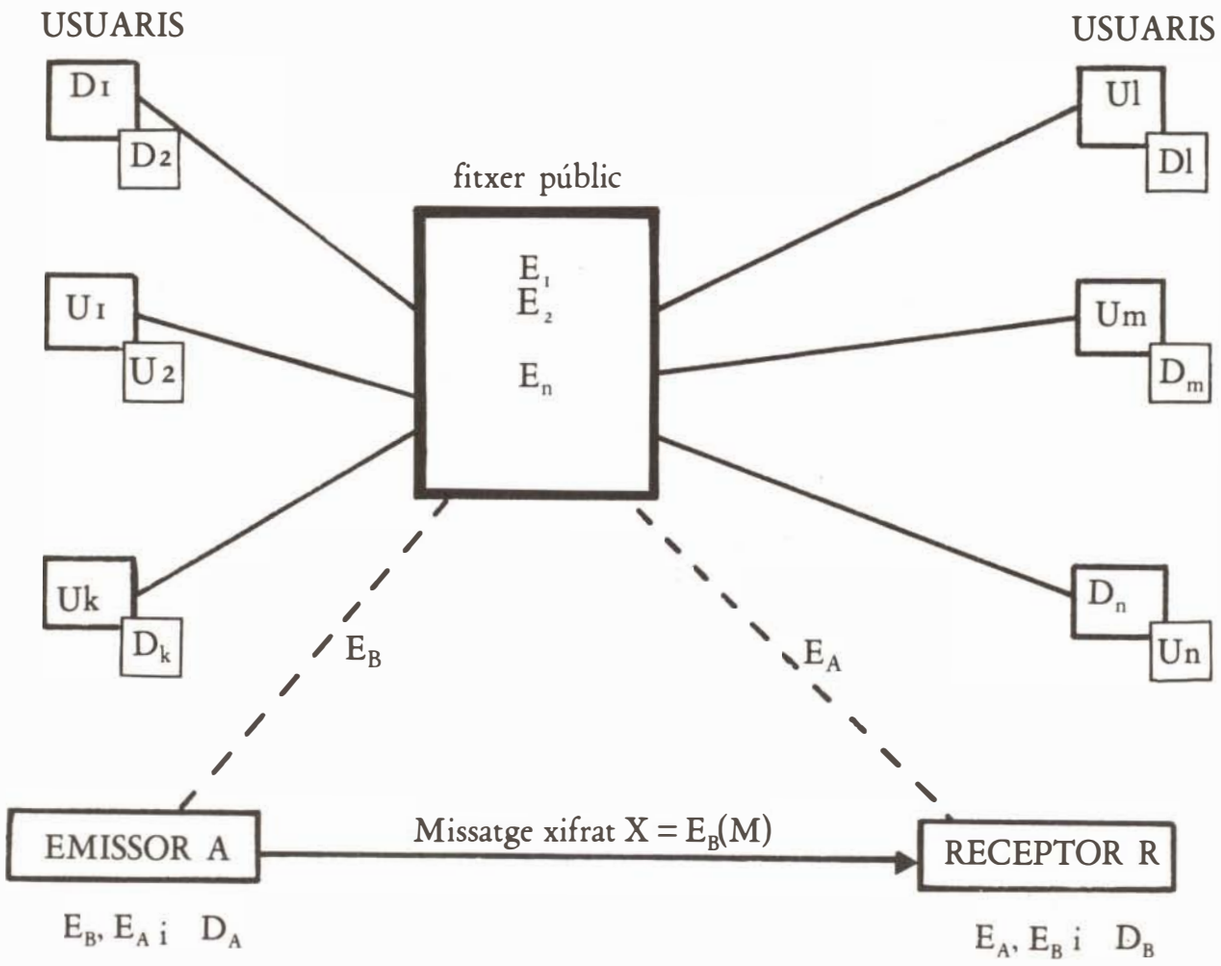
Missatge a transmetre	M
Missatge xifrat	$X = E_B(M)$
Descripció de X	$D_B(X) =$
	$D_B(E_B(M)) = M$

Signatura
Signatura

Signatura del missatge M	$S = D_A(M)$
Missatge xifrat	$X = E_B(D_A(M))$
Descripció de X	$D_B(X) = S$

S és indesxifrable per B; no obstant això, pot disposar de E_A i reconstruir el missatge $M = E_A(S)$

Fig. 5



CRIPTOSISTEMA A CLAU PÚBLICA