Criptografia: anàlisis d'una simulació de l'algorismeRSA

Alumne:Víctor Guerrero Corbi

Classe: 2n BATX A

Seguiment: Dep. Matemàtiques

Tutora: Carmen Cervantes

Centre: IES Jaume Balmes

Curs: 2006-2007

ÍNDEX

I. PREÀMBUL

1.1. Introducció 1

1.2. Motivacions i objectius 3

1.3. Estructura 4

APARTAT TEÒRIC

II. INTRODUCCIÓ A LA CRIPTOGRAFIA

2.1. Conceptes bàsics 5

2.2. Regles de Kerckhoffs 8

2.3. Tipus d'atacs 9

2.4. Secret perfecte 10

III. LA CRIPTOGRAFIA FINS ELS NOSTRES DIES

3.1. Primers mètodes criptogràfics 12

3.2. Quadre Vigenère 14

3.3. Mecanització de la criptografia: Enigma 16

3.4. Criptografiant el llenguatge 19

IV. LA CRIPTOGRAFIA A L'ACTUALITAT

4.1. Necessitat de la criptografia 22

4.2. Criptografia de clau secreta 24

4.2.1. Transposició i substitució 25

- Substitució monoalfabètica 26
- Substitució polialfabètica 26

4.2.2. Xifrat en bloc 28

- Algorisme DES 29
- Rijndael 31

4.2.3. Xifrat en flux 31

4.3. Criptografia de clau pública 32

4.3.1. Inicis 33

4.3.2. Criptosistema RSA 33

- Protocol RSA 35

Generació de les claus 35

Xifrat de missatges 36

Desxifrat de missatges 37

Exemple 38

- Criptoanàlisis elementals al RSA 39

Tipus d'atacs 39

Elecció dels primers p i q 40

Elecció de l'exponent de xifrat *e* *42*

Elecció de l'exponent de xifrat *d* *44*

Propietats del missatge 44

- Criptoanàlisis especialitzat al RSA 44

4.4. Aplicacions criptogràfiques 45

4.4.1. Autentificació 45

4.4.2. Firma digital 45

4.4.3. Certificació de clau pública 47

APARTAT PRÀCTIC

V DISSENY D'UNA APLICACIÓ INFORMÀTICA

5.1. Base matemàtica: l'algorisme 49

5.2. Llenguatges de programació 50

5.3. Fases de programació 51

5.4. Ordinograma 52

VI ANÁLISIS D'UNA APLICACIÓ PER SIMULAR EL CRIPTOSISTEMA RSA

6.1. Presentació de l'aplicació 54

6.2. Primer apropament a l'aspecte visual 55

6.3. Apropament al codi font 57

6.4. Dibuix de l'algorisme 63

6.5. Proves 64

APARTAT FINAL

VII MATERIALS, METODOLOGIA, CONCLUSIONS

7.1. Materials i metodologia utilitzada 66

7.2. Resultats i conclusions 66

VIII. RESSENYA BIBLIOGRÀFICA

Llibres 69

Pàgines webs 69

Altres 70

ANNEXES

XIX. CONCEPTES

9.1. Notacions i apunts matemàtiques 71
9.2. Glossari específic 72
9.3. Línia del temps 75

I PREÀMBUL

1.1. Introducció

> *Durant quant de temps vols que siguin secrets aquells missatges? – li va preguntar Randy en l'últim missatge abans d'abandonar San Francisco-. Cinc anys? Deu anys? Vint-i-cinc anys?*
> *Després d'arribar a l'hotel aquella tarda, Randy va desxifrar i va llegir la resposta d'Avi. Encara la té penjada davant els seus ulls, com la imatge remanent d'un flash. "Vull que segueixin sent secrets mentre els homes siguin capaços del mal."*
> ***Neal Stephenson, "Kriptonomicón"***

Des de temps antics fins avui en dia l'evolució humana ha estat constant i grans canvis han aparegut per canviar el transcurs de la història; però hi ha una cosa que no ha canviat: les ànsies de l'ésser humà per amagar les seves coses més íntimes i secretes. La criptografia ha estat una mena de resposta a aquesta necessitat. Una necessitat que va començar ja amb Juli Cèsar i la seva xifra on intercanviava les lletres del seu text original per la lletra que estava unes determinades posicions més endavant. També altres personatges històrics com els espartans que aprengueren també la importància de no permetre que les seves informacions quedessin als ulls de l'enemic van crear sistemes com la Scitala cap l'any 400 a.C.

Amb el desenvolupament de les ciències, però sobretot amb el de les matemàtiques, la criptografia va anar creixent poc a poc. L'any 1466, León Battista Alberti va crear el xifrat polialfabètic basat en la rotació d'unes rodes. Va haver de passar un segle fins que Giovani Battista Belaso inventés la clau criptogràfica basada en una paraula o text que es transformava lletra a lletra sobre el missatge original.

Però si un punt va marcat un abans i un després en la història de la criptografia va ser la Segona Guerra Mundial, la guerra d'Enigma i de la criptografia. Ja a principis del segle XX s'havien ingeniat màquines que eren capaces de transformar directament un text original en un de xifrat. Aquestes màquines estaven basades en les rodes concèntriques que al segle XV havien rondat la ment d'Albert. Aquestes màquines, però anaven més enllà i utilitzaven un sistema molt complexe. Un sistema amb el qual ja no hi havia suficient amb el seu criptoanàlisis, un sistema aparentment infranquejable: Enigma. Durant els anys 30 i 40 va ser aquest nom el causant de mals de caps a tots dos bàndols: en un principi pels que l'havien de criptoanalitzar, que no sabien com fer front a aquest nou i complexe sistema d'encriptació i posteriorment als mateixos alemanys que van perdre la guerra per la gran confiança que tenien en la seva creació i va ser aquesta mateixa confiança el que va fer que no s'adonessin

que l'enemic havia aconseguit trencar enigma: els secrets quedaven al descobert i per tant, podien espiar les seves comunicacions. Aquesta proesa es va dur a terme gràcies a un equip de criptoanalistes, ments privilegiades i matemàtics que l'any 1942 van poder trencar enigma gràcies a les seves bombes, aparells de càlcul mecànic que s'encarregaren de trencar el xifrat alemany. Aquest fet canviaria el transcurs de la guerra i el de la història mateixa.

Posteriorment amb naixement de la informàtica i dels criptosistemes informàtics es va produir un altre canvi radical en la forma d'enfocar el món de la criptografia: de cop hi volta l'augment de criptosistemes va augmentar tant en nombre com en complexitat. Gràcies a aquestes aparicions s'ha aconseguit arribar als punts on ens trobem ara per ara. També ha permès, en gran part, gràcies a Zimmerman que la criptografia estigui en mans de les persones del carrer, que no es quedi només com a eina dels governs o dels poderosos. Així doncs, ara per ara, qualsevol de nosaltres té (almenys) accés al dret a l'intimitat informativa pels mitjans electrònics. No cal dir que cal destacar el paper de grans informàtics i matemàtics com Shamir, Adleman o Rivest amb el seu RSA, Diffie i Hellman amb els primers plantejaments de clau pública i molts altres, alguns coneguts i altres coneguts en un futur ja que el món de la criptografia està envoltat d'aquest secretisme governamental.

1.2. Motivacions i objectius:

La realització d'aquest treball es deu al caràcter dels estudis universitaris que vull fer: enginyeria de telecomunicacions. I és precisament en aquest àmbit on el desenvolupament de la criptografia ha permès arribar a on s'ha arribat. Per això volia aprendre un tema nou per mi, un tema que servís de vincle amb l'imminent pas a la vida universitària. També destacar que l'elecció d'aquest tema ha estat moguda per l'interès d'aprofundir en un llenguatge matemàtic que potser fins ara no havia integrat del tot en la meva vida quotidiana perquè no n'havia fet ús i també perquè em semblava una eina necessària i molt útil en el nou camí que se'm presenta. Així doncs, el treball de recerca representa una oportunitat per poder fer tot allò que tenia en ment fins ara, però que potser no havia fet per manca de temps o d'imposició personal; no pas per desinterès.

Els objectius que busco amb aquest treball són: l'**analitzar l'evolució** de la criptografia i com aquesta ha vingut donada sobretot en temps de guerra. Com aquesta ciència i altres veuen el seu desenvolupament màxim durant les guerres on els diners invertits en la matèria d'investigació d'elements que puguin ajudar augmenta i per tant, com la criptografia ha jugat un paper importantíssim en el transcurs de les guerres. Un altre objectiu és el de la **realització d'una aplicació** basada en el criptosistema RSA. Aquest mateix objectiu s'explica perfectament si s'enfoca des del punt de vista mencionat abans de voler fer telecomunicacions i ser precissament en aquesta carrera on durant el primer any de carrera s'hi duu a terme l'assignatura d'introducció als ordinadors on s'aprèn un llenguatge de programació. Per això un objectiu implícit en la creació d'aquest programa és el d'**introduir-me en el món de la programació** i l'adquisició de nous coneixements matemàtics. Tot aquest objectius juntament amb l'objectiu d'adquirir un hàbit de creació de treballs més elevat que la resta de treballs realitzats fins ara, són els principals objectius a seguir durant el meu treball de recerca.

1.3. Estructura

L'estructura que segueix aquest treball és la que segueix la majoria d'estudis: una part teòrica i una part pràctica.

En el meu cas la part teòrica la he subdividida en tres parts. Aquestes parts venen predeterminades pels següents temes: primer es fa una introducció al llenguatge i termes en què es parla la criptografia i després hi ha les altres dues que representen la part més evolutiva. La primera és purament històrica i la segona en molts aspectes més matemàtica tot mostrant com amb l'aparició de la criptografia de clau pública i dels ordinadors el món de la criptografia queda reduït a un camp més matemàtic i informàtic.

A l'apartat pràctic s'hi recull com s'ha dut a terme la creació del programa i una introducció a teoria sobre programació. Encara ser teoria, aquesta s'ha introduït en l'apartat pràctic perquè respon al que es fa en aquest apartat. Aquest segon bloc representa el resultat de l'aplicació de la teoria recollida en el primer apartat. En el cas d'aquest treball com el que es vol aconseguir és fer un programa basat en el criptosistema RSA es pot veure com s'ha fet un anàlisi exhaustiu d'aquest sistema en l'apartat teòric.

Finalment s'hi ha inclòs un apartat amb les conclusions obtingudes, metodologia, etc. i uns annexes que inclouen un recull de notació matemàtica, glossari i una cronologia sobre la història de la criptografia.

APARTAT TEÒRIC

II INTRODUCCIÓ A LA CRIPTOGRAFIA

2.1. Conceptes bàsics

Es diu que el **llenguatge humà** neix per satisfer una sèrie de necessitats en l'àmbit comunicatiu que van anar apareixent a causa de l'evolució d'un sistema de comunicació primitiu en un principi, als complexes sistemes lingüístics que podem trobar avui en dia. Aquesta evolució del sistema de comunicació va marcar certament la història de la humanitat. Al parlar de la prehistòria i el pas a història estem evidenciant aquest gran pas: l'escriptura. Un mode de comunicació que deixava constància i que perdurava amb el temps. Un sistema de transmissió de coneixement més enllà del boca-orella dut a terme fins llavors.

Analitzant aquest tipus de comunicació podem trobar una sèrie de parts implicades, també anomenats **elements de la comunicació**, que cal precisar i definir abans d'entrar a aprofundir en el tema de la **criptografia**:

Per poder entendre la **criptografia** hem d'entendre en primer lloc que és l'acte comunicatiu i quins elements hi estan implicats. Així podem definir tot **procés comunicatiu** com el procés on un emissor transmet a un receptor un missatge conceptualitzat a partir d'un codi que coneixen tan ell com el receptor i que arriba al receptor a través d'un canal. Podem distingir-hi en aquest els següents elements comunicatius:

- **Emissor**: És la persona que transmet el missatge. Escull el que vol comunicar i ho transmet.
- **Receptor**: És la persona que rep i entén el missatge.
- **Missatge**: És la conceptualització que ha fet l'emissor del que volia transmetre.
- **Canal**: És el lloc a través del qual viatja la informació que es transmet.
- **Codi**: És l'idioma o conjunt de regles que coneixen tant l'emissor com el receptor i que serveix per poder entendre aquell missatge que volia transmetre l'emissor.

Per poder entendre millor aquests diferents elements podem entendre el cas següent que no és més que un dels molts casos de comunicació que podem trobar durant el dia a dia:

> *"Un avi es troba pel carrer amb el seu net i li explica el que ha fet avui a ioga"*

En aquesta situació podem observar que l'**emissor** seria l'avi, el qual explica el que ha fet a ioga, que seria el **missatge**. El seu net que està escoltant faria de **receptor** i el llenguatge que els dos dominen, en aquest cas, segurament el català, el **codi**. Finalment, el **canal** seria l'aire a través del qual viatgen les ones de so emeses pel seu avi.

Un cop hem entès el concepte de la situació on es duu a terme la criptografia, passarem a entendre que és en sí mateixa.

Si analitzem etimològicament la paraula **criptografia**, veurem que prové del grec **κρυπτός (kriptos)**, que significa ocult i **γραφή (graphos)**, que podem traduir com escriure. Així doncs, podem tenir una idea clara de la seva definició clàssica: *art d'escriure missatges en clau secreta.*

Aquesta visió de criptografia com a art va ser acceptada fins que **Shannon** va publicar l'any 1949 la "*Teoria de les comunicacions secretes*". És precisament en aquest instant que comença a ser considerada com una **ciència aplicada** ja que està relacionada amb altres ciències com poden ser la teoria de nombres, l'estadística o altres teories com les que es poden observar en la figura següent:

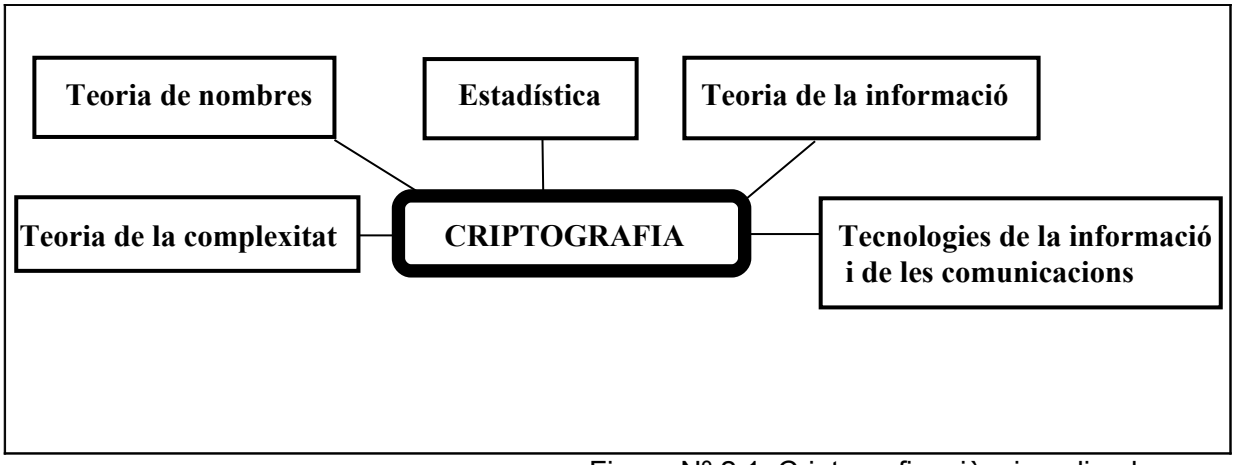

Figura Nº 2.1: Criptografia, ciència aplicada

Però cal especificar que la **criptografia** apareix ja que es precisa d'un secret en les comunicacions: existeix una desconfiança o perill que el missatge que es vol transmetre sigui interceptat per un enemic. Aquest enemic, si és que existeix, utilitzarà tots els mitjans que puga per desxifrar aquests missatges secrets. Amb aquest propòsit emprarà un conjunt de tècniques i mètodes que constitueixen una ciència coneguda com **criptoanàlisi.** A aquestes dues ciències, criptografia i criptoanàlisis, se les coneix per **criptologia**.

Des que publiqués, **Shannon**, la seva teoria de la informació, que s'utilitza la seva teoria i nomenclatura per l'estudi teòric de la criptografia. Així trobem que

ell ens parla d'aquestes dues ciències com a compartidores de molts conceptes i resultats, però essencialment, matèries diferents. En ambdós casos l'objectiu d'estudi o protecció és la informació. La **teoria de la informació** a diferència de la **criptografia**, estudia el procés de la transmissió sorollosa d'un missatge; per un altre banda, la criptografia estudia el procés de desxifrat d'un text. Per entendre-ho millor, en la teoria de la informació, la distorsió no és intencionada; sinó que és causada pel canal. A més a més, el receptor intenta recuperar el missatge transmès utilitzant el missatge rebut i una descripció probabilística d'aquest canal sorollós. En la criptografia, la distorsió sí és intencionada, i el receptor intentarà obtenir el text original utilitzant el text xifrat i un coneixement relatiu del procés de xifrat.

Si observem el punt de vista dels objectius de cadascuna, veurem com en la **teoria de la informació** el que es pretén és transmetre el missatge amb la més claredat (absència de distorsió per part del canal) possible, en canvi, en la **criptografia** el que es pretén és fer que el possible receptor enemic no pugui entendre que s'està comunicant.

La protecció de la informació que es vol aconseguir en la criptografia s'aconsegueix variant la seva forma. S'anomena **xifrat** (*o transformació criptogràfica*) a una transformació del **text original** (anomenat també *text inicial*) que el converteix en el text xifrat o **criptograma**. Anàlogament, a la transformació que permet recuperar el text original a partir del text xifrat s'anomena **desxifrat**.

Cada una d'aquestes transformacions ve determinada per un paràmetre anomenat **clau**. El conjunt dels seus possibles valors s'anomena **espai de claus K**. La família de transformacions criptogràfiques s'anomena sistema criptogràfic $\mathbf{T=\{T_k / k \in K\}}$, on anomenem $\mathbf{T_k}$ a l'operació de xifrat i com a $\mathbf{D_k}$ al desxifrat amb clau k. L'espai de missatges originals que constitueix el conjunt de partida dels xifrats s'anomena **M**, mentre que el conjunt imatge format pels textos xifrats s'anomena **C**.

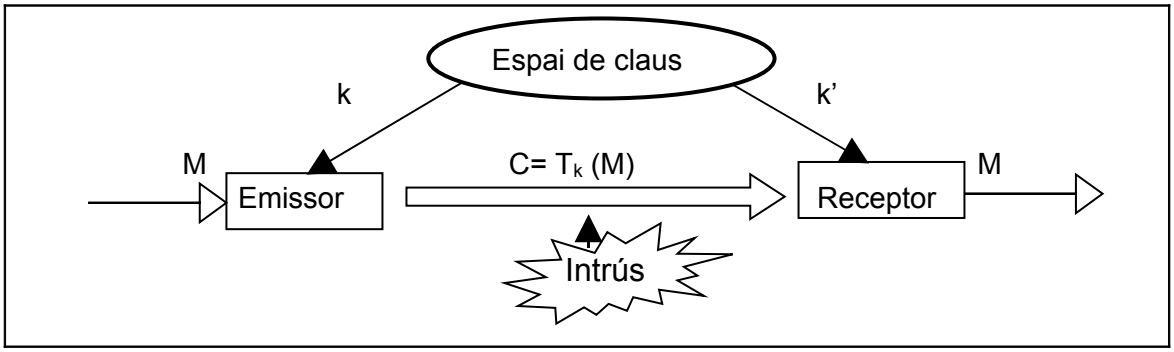

Figura Nº 2.2: Representació sistema criptogràfic

Destacar, finalment, que podem fer una primera classificació del xifrat segons la font que proporciona els textos: **xifrat digital** o **analògic**. *Digital* quan els missatges estan fets per grups d'elements que formen part d'una col·lecció finita anomenada alfabet. *Analògica* si genera un continuat de valors. Comentar que el xifrat de fonts analògiques no resulta segur, pel que normalment es tendeix a digitalitzar primer aquest tipus de missatges.

En els xifrats digitals, tant el text original com el xifrat resulten de les lletres d'un alfabet. Es consideren a més de l'alfabet $\mathbf{A=\{a_1,a_2,...,a_m\}}$, els alfabets formats per les diferents unions de les seves lletres $\mathbf{A^n = \{(a_1,...^{(n},a_1),...,(a_m,...^{(n},a_m)\}}$. Els seus $\mathbf{m^n}$ elements són les anomenades n-paraules. Per motius pràctics es tendeix a substituir les lletres per nombres. Aquesta substitució, principalment, es fa en dos alfabets numèrics: el alfabet binari $\mathbf{Z_2=\{ 0,1\}}$ i el decimal $\mathbf{Z_{10}=\{ 0,1,...,9\}}$ que queden generalitzats en l'expressió d'alfabets de nombres enters $\mathbf{Z_m=\{ 0,1,...m-1 \}}$.

Teòricament podem descriure aquest *sistema criptogràfic* com a $\mathbf{T=\{T_k^{(n)}:1\leq n<\infty\}}$, on $\mathbf{T_k^{(n)}(x)=y}$, on **y** és la n-paraula xifrada que correspon a la n-paraula original **x**.

Per evitar alguns problemes o ambigüitats es fan una sèrie de suposicions durant tot el desenvolupament teòric:

- $\mathbf{T_k^{(n)}}$ és bijectiva.
- S'utilitzen el mateix alfabet per ambdós textos, original i xifrat.
- Es defineix el xifrat de totes les possibles paraules, independentment de si existeixen o no.
- Cada n-paraula es defineix en una n-paraula, fent així que el xifrat no canviï la longitud del text original.

2.2. Regles de Kerckhoffs

Aquestes **regles de Kerkhoffs** (*s. XIX*) van aparèixer com a recomanacions per aconseguir una bona encriptació en el seu llibre titulat "*La criptografia militar"* i posteriorment van ser adoptades per gran part de les persones que tenien a veure amb la criptografia.

1) No ha d'existir ninguna forma de recuperar mitjançant el criptograma el text inicial o la clau.
2) Tot sistema criptogràfic ha d'estar compost per dos tipus d'informació: pública, com pot ser la família d'algorismes que el defineixen i privada, com és la clau que s'utilitza en cada xifrat particular
3) La forma d'escollir la clau ha de ser fàcil de recordar i de modificar.
4) Ha de ser factible la comunicació del criptograma amb els mitjans de transmissió habituals.

5) La complexitat del procés de recuperació del text original ha de correspondre amb el benefici obtingut.

Encara que puguin semblar regles un tant obvies, cal matisar que són de gran ús per evitar caure en errors bàsics. Així la primera regla que pot semblar bastant obvia serveix per explicar l'error dels sistemes criptogràfics que poden ser descoberts per l'anàlisi de freqüència o altres mètodes. La segona regla s'ha d'entendre en quant que tal com diu el principi de Kerckhoffs: *"la seguretat d'un criptosistema s'ha de medir imaginant que l'enemic coneix els processos de xifrat i desxifrat que s'han utilitzat."* Els altres punts tenen a veure amb qüestions de rendiment o relació cost-benefici de l'acte de l'encriptament.

2.3. Tipus d'atacs

Podem trobar que hi ha diferents problemes de seguretat tant en el moment de l'emmagatzemament de la informació com en el moment de la seva transmissió. En el primer cas el problema recau sobre el robatori del suport del missatge o directament l'accés no autoritzat a la informació, mentre que en el cas de la seva transmissió, podem trobar-nos amb el perill d'una possible intervenció, modificació d'aquests o interrupció de la comunicació i inclús la creació de missatges falsos.

Generalitzant les amenaces podem parlar que hi ha dos tipus bàsics:

- **Actives**, com pot ser la seva modificació, interrupció o falsificació.
- **Passives**, com és el seu robatori, l'accés no autoritzat o la seva intercepció.

Podem veure com segons el perill que trobem el que perilla en cada cas varia: si mirem el cas d'intercepcions i accessos no autoritzats el que perilla és la confidencialitat, en canvi, en els atacs per modificació i falsificació es posen en perill la integritat i la autenticitat dels missatges respectivament.

Si apliquem la primera regla de Kerckhoffs, podem distingir dos tipus bàsics de secrets: el **secret teòric** o incondicional i el **secret pràctic** o computacional. El primer on la informació disponible per l'enemic no és suficient per trencar el sistema. Per exemple quan el criptoanalista no disposa d'una quantitat de criptograma suficient per poder aplicar l'anàlisi de freqüència. En canvi, el secret pràctic es mesura d'acord amb la complexitat computacional del criptoanàlisi. Aquest fet ve donat per les necessitats actuals: s'ha de transmetre una gran quantitat d'informació i per això es busca el protegir aquesta informació de forma que el seu desxiframent a través de càlculs sigui complexa i llarga.

També trobem un altre nivell de seguretat anomenat secret probable per aquells criptosistemes que han resistit els atacs, però que no s'ha aconseguit demostrar formalment la base del seu secret.

Per l'estudi dels diferents sistemes criptogràfics és adequat posicionar a l'enemic. Així es poden descriure les següents situacions:

1. **Atac només amb text xifrat**: Aquesta és la pitjor situació possible pel criptoanalista ja que només coneix el criptograma
2. **Atac amb text original conegut**: En aquesta situació el criptoanalista té informació d'una part o de tot el text inicial i el xifrat. Podem trobar aquesta situació quan es coneix el tema del que parla el missatge ja que això crea una correspondència entre les paraules més habituals i els texts xifrats més repetits.
3. **Atac amb text original escollit**: Ara l'enemic pot obtenir a més del criptograma que intenta desxifrar, el xifrat de qualsevol altre text que ell escolleixi (entenem que ell no és que sàpiga xifrar-lo sinó que l'obté ja xifrat).
4. **Atac amb text xifrat escollit**: aquí l'enemic pot obtenir el text original corresponent a determinats textos xifrats de la seva elecció.

2.4. Secret perfecte

En aquest apartat exposaré els inconvenients i conclusions que se n'extreuen de la creació d'un suposat secret perfecte.

Si realment es pogués crear un secret perfecte veuríem com es produirien els següents fets:

- Un criptosistema té *secret perfecte* quan el text xifrat no proporciona ninguna informació sobre el text original.

  Si existeixen al menys tantes claus en K com n-paraules del text inicial de probabilitat positiva, llavors el sistema criptogràfic té *secret perfecte*.

- L'únic sistema existent amb les condicions de *secret perfecte* resulta poc pràctic ja que la clau és tan llarga com el missatge inicial i, a més, només es pot utilitzar un cop.

  Donat un criptosistema on el nombre de n-paraules inicials, el nombre de n-paraules xifrades i el nombre de claus coincideixen, existeix *secret perfecte* si, i només si: existeix només una clau que transformi cada n-paraula inicial en una n-paraula xifrada i totes les claus són equiprobables.

També podem concloure que a major quantitat de text xifrat, major possibilitat de recuperar el text original, per tant, un *secret perfecte* hauria de ser relativament curt, fent que es condicionés la quantitat d'informació que es volgués transmetre.

III LA CRIPTOGRAFIA FINS ELS NOSTRES DIES

3.1. Primers mètodes criptogràfics

En molts casos mai podem saber exactament el dia en què va aparèixer una determinada cosa. Així mateix passa amb el món de la criptografia del qual no es sap el seu punt de partida, però si que es coincideix en establir un inici basat en les descobertes fetes. Aquestes troballes ens parlen de dos mètodes que van néixer per donar resposta a les necessitats que representava la *situació bèl·lica* al qual feien front: un primer, durant el segle VI a. C., emprat per **Damarato**, personatge era un rei espartà que va idear un sistema per comunicar als lacedemonis del projecte que tenia Jerjes, despòtic líder dels perses, per envair Grècia. Aquest mètode consistia en fer un gravat del missatge que es volia transmetre en **tabletes** i després es cobrien amb cera per amagar el missatge. Altre sistemes semblants consistien en tallar el cabell a persones, escriure-hi el missatge i esperar a que creixés el cabell per amagar el missatge. El problema d'aquests mètodes era que no eren del tot eficaços i que no es tractava d'una encriptació pròpiament dita, sinó que el que feien era amagar el missatge. Aquests primers sistemes van ser recollits per **Herodoto** a "*Les Històries*", que representa una crònica dels conflictes entre Grècia i Pèrsia durant el segle V a.C. A aquests sistemes se'ls coneixerà per **esteganografia,** derivat de les paraules gregues **esteganos**, encobert i **graphein**, escriure.

Figura nº 3.1: tableta amb gravats

Si el que volem és trobar un primer mètode pròpiament criptogràfic es podria situar durant el segle V a. C., quan Espartans i Atenencs estaven en guerra. El mètode que van emprar consistia en l'escriptura dels símbols verticalment sobre una cinta enrotllada en un rotllo, de tal forma que, al desenrotllar-la els símbols del missatge quedaven desordenats, fent que el missatge fos incomprensible per aquell que no tingués informació sobre el grossor del rotllo

que s'havia fet servir per escriure-ho. Aquest sistema representa el *primer mètode de criptografia basat en la transposició* i s'anomenava **Scitala**.

Figura nº 3.2: Scitala

Un altre mètode que es sol destacar com un dels *primers sistemes* que es van emprar dins d'aquest camp va ser el de la **xifra del Cèsar**, que consistia en la substitució de determinades lletres per altres lletres segons una regla fixa. Aquest mètode aparegué durant el segle I a.C. i seguint en la línea del temps podem trobar altres exemples semblants en la Bíblia, on es poden trobar texts xifrats amb **atbash hebreu**, encara que aquests texts es creuen d'èpoques fins i tot anteriors a la *xifra del Cèsar*. Per ser exactes del 600-500 a.C.

Si busquem a la història documents que parlin expressament sobre *teoria de la criptologia* podem trobar com a un dels llibres més antics on se'ns parla d'aquest tema precisament és al **Kamasutra**, on entre les 64 arts que han de ser estudiades hi podem trobar el de l'art de l'escriptura secreta per raons conjugals, ja que suposava l'eina per les dones per mantenir relacions d'amagades de l'espòs amb uns amants i poder comunicar-s'hi sense haver de patir per si l'home se n'adonava. Però en aquest cas, com he dit, es tracta d'una menció dins d'un bloc. No serà fins el segle XIV que trobem l'obra més antiga que existeix sobre criptografia, "*Liber Zifrorum*" de **Cicco Simoneta**. En aquesta obra trobem un estudi de diferents sistemes basats en simples substitucions de lletres.

Al segle següent trobem la figura important d'**Alberti** (1404-1472) que va destacar en l'àmbit del criptoanàlisi i per la creació d'un aparell que estava format de dos cercles concèntrics que, tenint radis diferents, estaven fixats en un eix comú. Cadascun dels anells tenia 24 divisions: l'anell més gran contenia

20 lletres de l'alfabet (*no hi apareixien les lletres h, j, k, u, w, i*) i els quatre primers dígits i en l'altre anell hi podíem trobar 23 lletres desordenades (*no es feia ús de j, ñ, u, w*) i la conjunció et *("i" en francès*). El procés a seguir era el d'escollir una clau determinada. Amb aquesta clau es substituïen les lletres del text normal (*les que es trobaven en el cercle menor*) per les lletres de l'anell exterior que representava al text xifrat. Per aquest i els seus estudis com a criptoanalista se'l considera el **pare de la criptologia** (*recordar que s'entén per criptologia a la suma dels dos sistemes: encriptació i criptoanàlisi. Fins ara no s'havien dut a terme estudis seriosos sobre com poder esbrinar el que s'encriptava*).

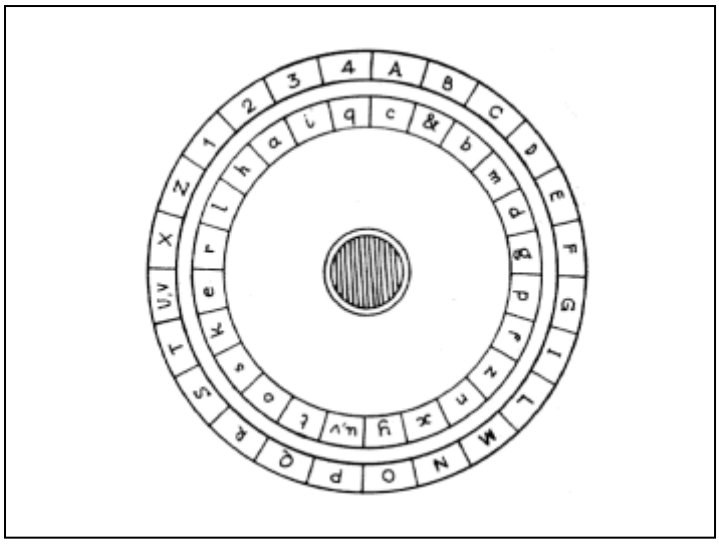

Figura nº 3.3: Roda d'Alberti

Aquest personatge donaria pas a un segle marcat per l'aparició d'un sistema d'encriptació que no s'havia vist fins ara:

3.2. Quadre Vigenère

Va ser durant el segle XVI que es va *generalitzar l'ús de la criptografia* en l'àmbit diplomàtic i cap al final d'aquest període, l'any 1586, **Blaise de Vigenère** va presentar la seva obra "*Traicté des Chiffres*" on ens presenta un tipus de criptografia que intentava donar resposta al l'avanç per part dels criptoanalistes àrabs (*primer text on se'ns descriu un criptoanàlisi més antic que es coneix basat en l'anàlisi de freqüència al "Al-Kindi", 801-873*) i posteriorment pels europeus que havien trobat la manera de guanyar la situació a la simple substitució o transposició mitjançant l'anàlisi de freqüència dels

texts xifrats. Així doncs aparegué el que es va anomenar "*Le chiffre indéchiffrable*" (*la xifra indesxifrable*) que intentava guanyar la partida als criptoanalistes. Aquest sistema estava descrit en el llibre de *Vigenère*, però no era pas idea seva. Tot va començar amb Alberti que va ser el creador de la substitució polialfabètica (*utilitzant diversos alfabets per xifrar*). El que va millorar *Vigenère* va ser el crear una substitució polialfabètica a partir d'una paraula clau. Així trobem el **quadre Vigenère** que mostra com funciona aquest sistema:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
1	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
2	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
3	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
4	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
5	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
6	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
7	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
8	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
9	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
10	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
11	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
12	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
13	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
14	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
15	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
16	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
17	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
18	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
19	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
20	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
21	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
22	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
23	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
24	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
25	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
26	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Figura nº 3.5: Quadre de Vigènere

Per exemple si agaféssim com a paraula clau "casa" i volguéssim encriptar "Les oques van descalces" primer hauríem de repetir tants cops com lletres té el missatge, la nostra clau obtenint:

Clau: C A S A C A S A C A S A C A S A C A S A
Missatge: L E S O Q U E S V A N D E S C A L C E S

Ara el que cal fer és mirar la lletra que hi ha sobre de la lletra del nostre missatge: en el primer cas la lletra "C" que servirà per indicar-nos quina fila haurem d'utilitzar: la 3, ja que la lletra "C" ocupa la tercera posició. Ara per saber la columna mirarem la lletra del nostre missatge, la "L" que servirà per obtenir la columna. Amb la fila i la columna busquem on coincideixen al quadre de dalt i observem que ens surt la lletra "O". Així anem fent amb cadascuna de les lletres que conformen el nostre missatge. En l'exemple donat obtindríem el següent missatge encriptat:

Text encriptat: O F L P T V X T Y B G E H T V B O D X T

Com podem veure es tracta d'un text sense sentit. Un cop el missatge ha estat transmès per poder-ho desencriptar com el receptor coneix la clau el que ha de fer és el pas contrari, és a dir, situa la clau sobre el text que ha rebut tot repetint-la fins a donar a cada lletra una lletra. Un cop ha fet això observa quina lletra hi ha sobre la primera lletra del seu text i la fa correspondre amb la fila que està marcada pel número que representa aquella lletra. Ara resseguim en horitzontal aquesta fila tot buscant la lletra del nostre text encriptat. Un cop l'hem trobada mirem quina lletra correspon a aquella columna i ja tenim la nostra lletra del text desencriptat.

Aquest sistema que es creia com a indesxifrable va ser desxifrada per **Charles Babbage** que al segle XIX va poder trobar el punt feble d'aquest sistema. Així fent un anàlisis de freqüència, aplicant altre informació basada en combinacions de lletres i sentit comú va poder resoldre l'enigma.

3.3. Mecanització de la criptografia: Enigma

El que havia representat la **roda d'Alberti** en el segle XV, una mena de *mecanització del sistema criptogràfic* va anar un pas més enllà a partir del segle XX que va veure com el món de la criptografia havia d'avançar a marxes forçades a causa de les **dues guerres mundials**. Aquest fet va venir donat ja que s'havia de poder establir *comunicacions militars i diplomàtiques* secretes utilitzant les noves tecnologies de comunicació: la *telegrafia* i la *radiotècnia*. A la **primera guerra mundial** podem destacar el fet del desencriptament del conegut com **telegrama Zimmermann**, amb el qual el ministre alemany amb aquest nom pretenia convèncer a Mèxic i Japó per envair els Estats Units.

Aquest telegrama va ser entès gràcies als aliats britànics i va ser la clau per convèncer als Estats Units d'entrar a la guerra.

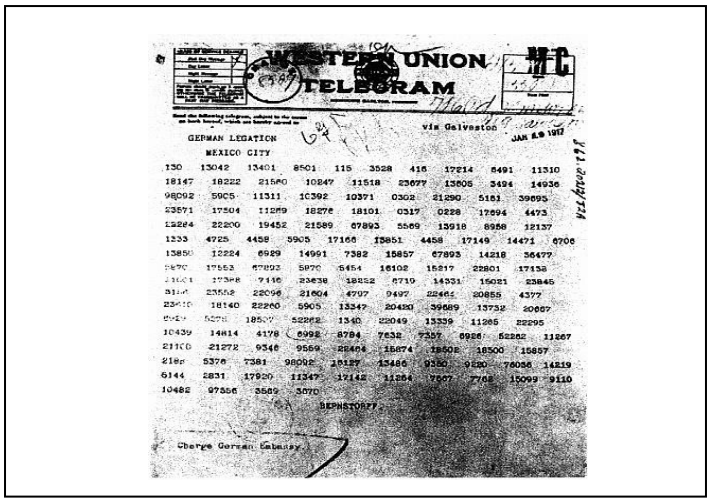

Figura nº 3.5: Telegrama Zimmerman

En la **segona guerra mundial** destaca el paper important de l'**Enigma**, la qual té els seus inicis en el paper que van fer durant tot el segle XX personatges com *Jefferson* que van intentar agafar la idea d'Alberti, però desenvolupant-la, això vol dir que van crear unes màquines que tenien uns rotors que es movien a mesura que es premien les tecles, tot seguint una seqüència. Aquestes màquines aconseguien que enlloc d'haver de fer el dur treball d'anar transcrivint el text encriptat manualment, dedicant en molts dels casos molta estona, tot aquest procés es fes automàticament per la màquina. Aquestes màquines van rebre el nom de **traductors mecànics**.

Figura nº 3.6: Màquina Enigma

Tornant a **Enigma**, aquesta va ser patentada per **Koch** i **Scherbius** l'any 1919 i va ser utilitzada pels *nazis alemanys* com a màquina per enviar les seves comunicacions encriptades. Aquesta màquina basava el seu funcionament en tres cilindres a través dels quals rotava cada lletra. Quan una lletra passava pel primer cilindre, aquest girava una posició. Al haver girat 26 cops, ho feia llavors el segon cilindre i després de 26 altres pulsacions o feia el tercer. Depenent de la posició inicial dels tres cilindres sortia un tipus de combinació o una altre i la seguretat d'aquest sistema residia en la col·locació dels mateixos i del complex funcionament intern (el que s'ha explicat ara no és més que una simplificació del seu funcionament). *Enigma* va representar una *arma de doble fil*, ja que la confiança dels alemanys en aquesta màquina va fer que no pensessin en la seva possible desencriptació, però això no va ser així i cap al 1941 ja representava una ajuda per derrotar-los. Aquesta ajuda venia per part de l'equip de l'organització secreta que treballava pel desxiframent d'Enigma. En aquesta organització hi podem trobar a ***Alan Turing***, un important personatge que ajudaria al seu desencriptament que va venir en forma de màquines que provaven les diverses combinacions. Aquestes màquines van rebre el nom de **bombes** pel seu tamany i pel soroll que feien. Podríem dir que aquest tipus de màquines eren les precursores dels ordinadors, però de forma mecànica (*Alan Turing tenia un treball on havia ideat una mena d'ordinador, aquesta idea era massa avançada pel moment i no es va dur a terme, però existia la teoria o la idea que entreveia el que serien els primers ordinadors*).

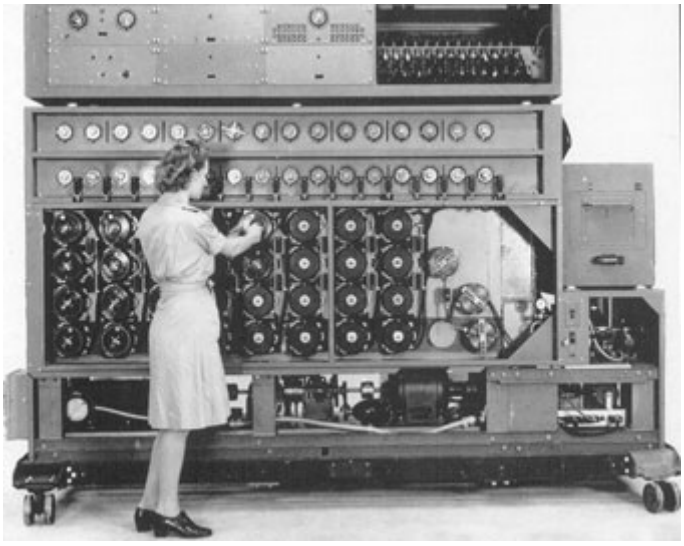

Figura nº 3.7: màquines Bomba

Encara que aquí sembli molt senzill, tot el procés va ser extremadament difícil i va comportar molts mals de cap a molts dels que hi treballaven. Un cop descobert el funcionament d'Enigma i la forma de vèncer-la es van poder llegir i saber molta informació sobre els plans dels alemanys i aquests van ser utilitzats per guanyar la segona guerra mundial.

Així podem veure dues coses: com les guerres formen un cop més part de les causes dels grans avenços de la criptografia i l'important paper que la criptografia va tenir en les dues guerres mundials. Un paper que en molts dels casos roman desconegut per moltes persones.

3.4. Criptografiant el llenguatge

En aquest apartat s'explica no pas un gran avenç criptogràfic, sinó més bé com el **llenguatge** en sí és un *sistema criptogràfic* si s'observa des del punt de vista que tota llengua és un seguit de símbols arbitraris que no tenen sentit més enllà dels que la dominen. Així posem per exemple un català que no sàpiga anglès es troba amb un text en anglès on hi posa: "*My house is big*" Com no sap anglès, no entén que hi diu i necessita d'un *diccionari* que seria com la *clau* que es necessita per desencriptar o en aquest cas traduir al propi llenguatge el que hi diu. En aquest cas com la gramàtica no difereix gaire no caldria gran cosa més, però en cas de llengües on diferís molt l'estructuració d'oracions i altres aspectes ens caldria més coses per poder "*desencriptar*" el que hi posa al missatge.

Aquest fet va ser aprofitat pels ***americans*** en la guerra i van buscar dins del seu territori aquells membres que complissin les condicions de tenir un *llenguatge desconegut i aïllat*, ja que si buscaven entre una llengua provinent d'una altre podria coincidir en molts aspectes amb aquesta. Aquest cas el van buscar entre les poblacions d'indis americans. Els escollits van ser els **indis navajos**. El seu llenguatge va conformar un codi que va ser essencial per guanyar varies batalles. La complexitat del codi va deixar perplexos als japonesos i va demostrar que era impossible de ser desxifrat per varies raons: hi ha múltiples sons per la representació de la mateixa lletra, així una paraula diferia molt depèn de com es pronunciés, també per la complexitat intrínseca del propi idioma i la forma com es va utilitzar: es va crear vocabulari específic per poder parlar de termes militars i per comunicar-se es feien servir dos sistemes: algunes expressions eren utilitzades descriptivament (*ASI-UN-LA-IH en el seu idioma significava "una estrella", però feia referència a la insignia d'un General de Brigada o CHA-LE-GAI que es traduïa com "les gorres blanques" que volia referir-se als mariners*), altres vegades el que es feia era enviar paraules que fessin referència a objectes, la primera lletra del qual conformés la paraula que es volia transmetre. Un altre aspecte important de la seva seguretat era que aquesta tribu índia havia estat aïllada a les seves reserves

de Nou Mèxic i Arizona i tampoc hi havia cap diccionari, ni tant se vol del idioma navajo a l'anglès. Finalment el fet que els japonesos no fossin conscients que s'enfrontaven a una llengua viva i no pas a un codi xifrat va dificultar la feina per desxifrar-ho. Cosa que no es va arribar a realitzar mai.

Gràcies al poder d'aquesta llengua es va poder guanyar les **batalles del Pacífic** com la del 7 d'agost de 1942 a *Guadalcanal*, o a *Iwo Jima* entre altres llocs. Així els militars americans van reconèixer la seva importància: el *Comandant Howard Connor* de la 5a Divisió dels marins va dir que "*sense els navajos no hagués estat possible aconseguir Iwo Jima*" Per fer-nos una idea de la seva participació mencionar que quan es va aterrar a la platja d'Iwo Jima el 19 de febrer de 1945, els joves navajos van transmetre en 48 hores més de 800 missatges sense cap error.

A més a més aquest tipus de sistema tenia un altre *avantatge*: era molt *més ràpid* que el sistema tradicional ja que no s'havia de codificar el missatge, enviar-lo en morse, transcriure'l i tornar a descodificar-lo que en moltes ocasions duia a tarda una hora o més. El codi basat en la seva llengua índia es transmetia en uns minuts.

Aquest apartat pretén mostrar com els *llenguatges* poden ser també una *eina útil per a l'"encriptació"*. Altres exemples de llengües que es podrien haver fet servir serien l'*euskera* que també satisfaria (en el seu moment, no pas ara) en certa manera les condicions que es buscaven. L'altre cosa que vol aquest apartat és donar a conèixer el paper d'aquest grup de persones que en el seu moment no van rebre la consideració necessària ja que mentre el seus companys marines se'ls va donar una benvinguda triomfal, la tornada dels indis americans va ser molt dura ja que ningú sabia el que havien fet perquè va romandre en secret durant temps després que s'acabessin les guerres. Com ells, molts altres personatges que van romandre, romanen o romandran en l'*anonimat* i *secretisme* que comporta el camp de la criptografia.

Figura nº 3.8: utilització del codi navajo en les selves de Bougainville en 1943

IV LA CRIPTOGRAFIA A L'ACTUALITAT

4.1. Necessitat de la criptografia

Durant la *segona meitat del segle XX* el món va començar a canviar, una nova realitat prenia forma: la cultura informàtica i Internet. Amb aquest canvi el que es va produir va ser un increment exponencial de la informació que estava en mans dels usuaris: la **societat de la informació**. Aquest canvi ha estat tant important que en la portada del mes de gener d'aquest any del **Time** on s'escull a la persona o tipus de persona més important de l'any, s'ha decidit escollir a tots nosaltres, els formants d'aquesta era de la informació:

Figura nº 4.1: Portada de Gener de 2007 de Time

El *desenvolupament de la informàtica* va causar també un important canvi en la forma com es veia la **seguretat dels sistemes criptogràfics**. Això es produïa ja que aquells sistemes que costava molt de temps de treure manualment es podien desxifrar mitjançant l'ordinador que era capaç de provar i calcular moltes combinacions per segon. Aquest fet ha fet que la seguretat dels sistemes clàssics es veiés substituïda per una *seguretat matemàtica i computacionalment demostrable*. El que ha causat això és que el nivell de la criptografia actual s'entengui en termes diferents i que aquesta no estigui en mans de qualsevol. Així veiem com fins aquest punt del treball, la criptografia ha anat per vies més o menys comprensibles per tothom i que els mètodes que es seguien també es podien dur a terme per qualsevol persona que volgués, és més fins aquest punt molts dels *amants de la criptografia* havien estat entesos escriptors interessats pel tema com **Edgar Allan Poe** (*1809-1849*) que va llençar el repte als seus lectors que era capaç de desxifrar qualsevol missatge basat en xifres de substitucions monoalfabètiques. En aquest punt, això desapareix i la criptografia passa a estar en mans d'uns pocs matemàtics i informàtics que s'encarregaran de buscar nous sistemes criptogràfics. Estem doncs davant d'un punt que marca un abans i un després de la criptografia.

En aquesta nova etapa hi destaquen dos fets importantíssims: la publicació dels **treballs de Shannon** que van significar la formulació matemàtica definitiva de la teoria de les comunicacions i la seva aplicació criptogràfica i després l'aparició de la criptografia de **clau pública** de la mà de **Diffie** i **Hellman** l'any 1976.

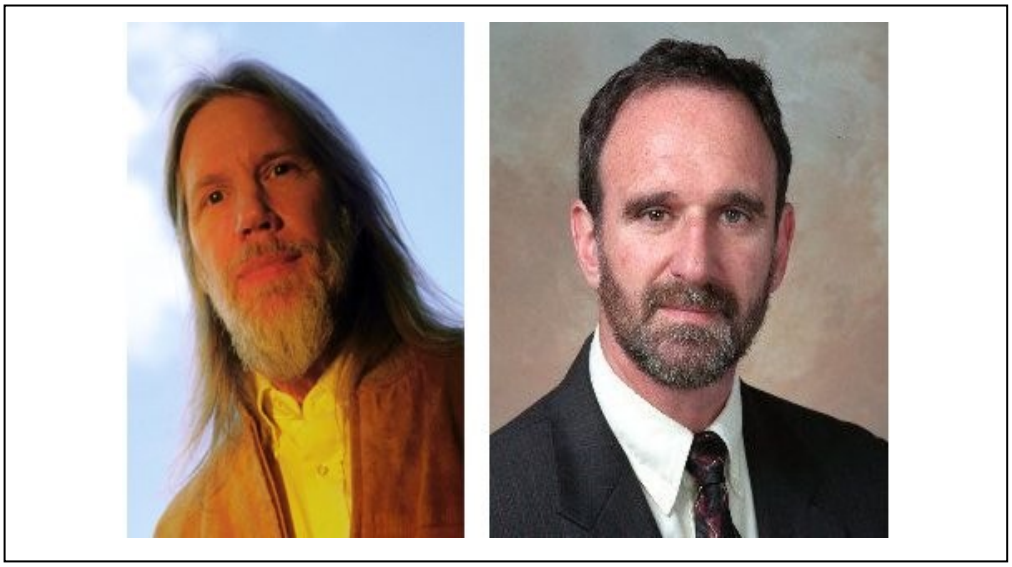

Figura nº 4.2: Per ordre d'aparició White Diffie i Hellman

En aquest punt es descriuen els sistemes que van sorgir a partir d'aquesta segona meitat del segle XX i les seves bases teòriques. És per això que

resulten uns punts més matemàtics i teòrics en comparació al que ha estat fins ara l'anàlisis evolutiu de la criptografia.

El que es presenta a continuació és una *classificació* que va néixer amb l'aparició de la criptografia de clau pública. Així ara es parlarà de **criptografia de clau secreta** que conté els *mètodes clàssics* i nous sistemes que apareixen en aquest segle com el *DES*, *Rijndael* entre altres i la **criptografia de clau pública** que és totalment nova i que no s'havia vist fins ara. Precisament aquesta idea neix contradient una de les coses que es creia indispensable en el món de la criptografia: *la clau ha de ser secreta* i l'han de conèixer els dos participants de la comunicació, amb la criptografia de clau pública s'aconsegueix un *sistema per poder compartir la clau públicament* i el que és més, es va poder aconseguir d'idear un sistema per poder identificar que l'emissor era realment qui deia ser i que el text no havia estat retocat.

4.2. Criptografia de clau secreta

Parlem de *criptografia de clau secreta* quan la **clau de xiframent** és un secret compartit exclusivament per l'emissor i pel receptor.

El problema principal d'aquest sistema és l'alta redundància de la font, és a dir, com més llarg sigui el missatge més pistes estarem donant sobre com s'ha encriptat el missatge ja que el mètode a seguir per encriptar cada lletra segueix els mateixos procediments. Precisament, **Shannon** va suggerir que per evitar el **criptoanàlisis estadístic** que aprofiti aquesta debilitat es podia utilitzar la *difusió* i la *confusió*.

La **difusió** intenta anular la influència de la redundància de la font sobre el text xifrat, dissipant-lo: això vol dir separant les estructures que vinguin donades per l'encriptament per evitar combinacions de lletres usuals que ajudin al criptoanàlisis de freqüència. La forma de dur a terme aquesta difusió, la **transposició**, que evita el criptoanàlisi basat en la freqüència de les paraules. Un altre opció és el de fer que a cada lletra del text xifrat li corresponguin un gran nombre de lletres del text original.

Per altre banda, la **confusió** consisteix en fer que la relació clau-text xifrat sigui la més complexa possible, fent així que les estadístiques del text xifrat no estiguin molt influïdes pel text original. Això s'aconsegueix mitjançant la tècnica de **substitució**.

Comentar que en solitari cap dels dos mètodes garantitzen la seguretat de cap xifrat.

4.2.1. Transposició i substitució

Dins de la criptografia de clau secreta podem trobar una primera classificació en:

- **Transposició**: consisteix a crear un text xifrat simplement desordenant les unitats que formen el text original a partir d'una clau
- **Substitució**: consisteix a reemplaçar les unitats del text original per unes altres a partir d'una clau.

Cal destacar que tant la primera com la segona opció tenen una cosa en comú: la clau és secreta i compartida per emissor i receptor. També podem parlar d'una acció que permet millorar una mica el que volem aconseguir amb el xifrat d'aquest tipus: combinar substitucions i transposicions alhora. És més, anomenem **xifrat producte** a l'aplicació iterativa d'operacions de xifrat sobre texts ja xifrats. A aquest procés que matemàticament parlant anomenaríem composició de xifrats, vindria definit com a $\mathbf{T(M)=T_1\,(T_2\,\ldots\,(T_r\,(M))\,\ldots\,)}$

Entrant en l'explicació de la **transposició** parlarem que el procés que es segueix és el següent:

1. Dividim el missatge original en blocs de la longitud escollida.
2. S'aplica a cada bloc la transposició determinada per la clau escollida.

En canvi, direm que la **substitució** definida mitjançant la clau $\mathbf{k=(\gamma_o, \gamma_1,\ldots, \gamma_n,\ldots)}$ és la transformació criptogràfica $\mathbf{T_k}$, que xifra la n-paraula original $\mathbf{(x_0,\ldots,x_{n-1})}$ com la n-paraula $\mathbf{(y_0,\ldots,y_{n-1})}$ amb $\mathbf{y_i=\gamma_i(x_i)}$, $\mathbf{0\leq i<n}$, $\forall$ **n=1,2,...**

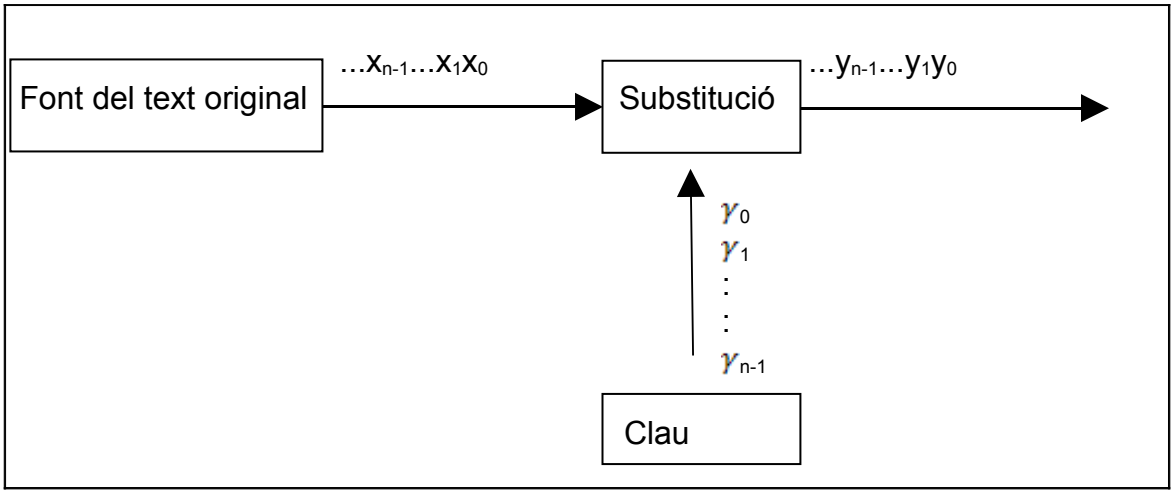

Figura nº 4.3: Esquematització funcionament mètode substitució

- **Substitució monoalfabètica**

Dins d'aquest grup podem distingir entre substitucions de lletres i substitucions d'n-paraules.

- ***Substitució de lletres***: un exemple clàssic d'aquest tipus d'acte seria el de la **Substitució del Cèsar** (*50 a.C.)* on canviava una lletra per un altre que estava 3 posicions per davant. Així tenim que el sistema generalitzat per a l'expressió d'aquest tipus de substitució de lletres vindria donat com: donat un alfabet de **m** lletres i un enter fixe **b**, s'anomena transformació desplaçament a la funció $\mathbf{T_b}$ definida mitjançant $\mathbf{T_b(M) \equiv M + b \pmod m}$. Per desxifrar una unitat del missatge xifrat **C**, es calcula $\mathbf{M = T_b^{-1}(C) \equiv C - b \pmod m}$. Així doncs, el xifrat de Cèsar no és més que un desplaçament on **b = 3** (*substituïa la lletra per la corresponent 3 llocs per davant*) i **m = 27** (*utilitzava les 27 lletres del seu abecedari*).

  El *problema* d'aquest tipus de substitució és que qualsevol enemic pot criptoanalitzar-lo amb un **anàlisis de freqüència**. Amb aquest anàlisis de freqüència el que es pretén és fer coincidir la lletra més repetida en el text xifrat per la més habitual en l'idioma que s'ha escrit el missatge. Així doncs amb una mica de paciència i sentit comú es pot treure més o menys fàcilment aquest tipus de xifrat.

- ***Substitució d'n-paraules****:* en aquest cas el que tenim és que el canvi en comptes de ser d'una lletra per una altre, el que canviem ara són **n** lletres per combinacions d'altres lletres cosa que fa que no sigui tan evident l'anàlisi de freqüència per al segon bloc de lletres. Exemples d'aquest tipus de substitució els trobem en la xifra **Playfair** *(1854)*, en els ***xifrats fraccionals*** o en els ***matricials***.

- **Substitució polialfabètica**

La substitució polialfabètica es defineix mitjançant una clau $\mathbf{k=(\gamma_0,\gamma_1,...)}$, que conté com a mínim dos substitucions de γ_i diferents. Dins d'aquest grup podem trobar dos xifrats com el de **Vernam** o el de **Vigenère**.

- ***Xifrat de Vernam*** *(1917):* es pot veure un anàlisis d'aquest xifrat més endavant en el punt 4.2.2. Com per tenir una idea del que és la substitució polialfabètica amb un exemple ho podem tenir molt clar, explicaré el més conegut també com a quadre Vigènere que va suposar la introducció d'un nou alfabet de xifrat.

- *El* ***xifrat Vigènere*** *(1586):* també anomenat quadre Vigènere, les claus d'aquest tipus de xifrat es defineixen mitjançant la repetició d'una

seqüència de lletres anomenada llavor $\mathbf{k_0,....k_{r-1}}$, fent que $\mathbf{k= k_0,...,k_j,...}$ amb $\mathbf{k_j=k_{j(mod\ r)}}$, $\mathbf{j=0,1,...}$

Així el xifrat Vigènere es defineix mitjançant aquesta clau de la forma $\mathbf{T_k(x_0,...,x_{n-1})= ((x_0+k_0)(mod\ m),...,(x_{n-1}+k_{n-1})(mod\ m))}$. Aquest es pot considerar una generalització de la xifra de Cèsar.

El *problema principal* de la substitució polialfabètica la qual es creia més segura a causa de la aparició de nous alfabets que feien augmentar la dispersió de la coincidència de lletres de la simple substitució monoalfabètica va ser que no resultava ser perfecte ja que s'aconseguia trencar mitjançant un criptoanàlisis recollit en el **mètode Kasiski** *(1863)* i en l'**índex de coincidència.**

Per poder aplicar aquests mètodes primer calia saber dues coses:

1) Descobrir el període **r** de la clau
2) Recuperar les **r** substitucions $\mathbf{(\gamma_0,...,\gamma_{r-1})}$

Un cop s'han obtingut aquestes dues dades només cal aplicar un anàlisis de freqüència de les lletres i les paraules. Per descobrir el període de la clau abans mencionat es poden aplicar els mètodes abans mencionats:

- ***Mètode Kasiski***: aquest sistema utilitza les repeticions en el text xifrat per obtenir pistes sobre el període, que és justament on es centra la seguretat de la substitució polialfabètica.

Si tenim davant el següent exemple:

Text original: POLVO ERES Y POLVO SERÁS
Clau: NOTAS NOTA S NOTAS NOTAS

Text xifrat: CCEVG RFXS Q CCEVG FSKAK

De l'observació d'aquest text xifrat podem intuir que el període sigui 10 o un divisor de 10. Això ho podem saber ja que les repeticions que observem en el text xifrat passen ja que el text original té repeticions a una distància igual a un múltiple de la longitud de la clau. Això vol dir que si a l'observar el text xifrat veiem l'estructura "*CCEVG*" dues vegades és molta casualitat que hagin coincidit totes les lletres, això ens indica que segurament la clau i les paraules del text original eren les mateixes cosa que ens pot servir per entreveure la possible llargària de la clau. Aquesta podrà ser o una clau de 10 lletres tot pensant que les lletres que no es repeteixen en el text xifrat són degudes a que el text original es repeteix en intervals de cinc lletres o bé que la clau és de cinc lletres i es va repetint i en aquest cas la variació del text xifrat es deu que el text

original té dues paraules de 5 lletres iguals i entremig d'aquestes unes altres de 5 lletres que són diferents. En aquest cas podríem veure que té més coherència la segona opció, cosa que ens donaria informació sobre la llargària de clau que com veiem, hem encertada (la llargària). A continuació s'aplicarien els sistemes exposats a continuació per poder fer un estudi que permetés esbrinar la clau exacta.

- ***Índex de coincidència o IC***: aquest sistema el que fa és donar una mesura de variació entre les freqüències de les lletres del text xifrat. Si el període de la clau fos 1 (substitució de lletres), tindríem un alt ranc de variació en la freqüència de les lletres, pel que el valor de IC seria alt. En canvi, a mesura que el període augmenta, la variació en la freqüència de lletres disminueix i per tant el IC també baixa. Aquest coeficient es defineix exactament com a:

IC = $\sum_{i=1}^{m} \frac{Fi(Fi-1)}{N(N-1)}$, sent $\mathbf{F_i}$ la freqüència absoluta de la i-èssima lletra del alfabet en un text xifrat de longitud **N**. El IC estima la probabilitat de que dos lletres escollides a l'atzar dins d'un text coincideixin.

L'aplicació conjunta d'aquestes dues tàctiques permeten una ajuda a l'hora del criptoanàlisi d'aquest tipus de xifrat.

Per acabar aquest apartat, mencionar que la substitució i la transposició com ja he dit abans no resulten molt efectius utilitzats individualment, però que els dos junts constitueixen la base de sistemes molt més difícils de criptoanalitzar. Alguns d'aquests esquemes van ser emprats en els anys vint pel disseny de màquines de rotor (*implementacions de xifrats de Vigènere amb claus llargues*). Entre aquests usos hi trobem precisament la **Enigma**, que es va utilitzar en la **Segona Guerra Mundial**.

4.2.2. Xifrat en bloc

També podem trobar un altre tipus de classificació segons la forma en què ve presentada la font xifrada. Així doncs, parlarem de dos grans blocs: *xifrat en bloc* i *xifrat en flux*.

El **xifrat en bloc** consisteix en les fonts que generen paraules o blocs de lletres. Aquest treballa sobre texts formats per paraules, convertint cada paraula en una de nova. Una de les característiques d'aquest tipus de xifrat és que cada bloc xifrat és independent tant de la seva posició com dels blocs adjacents. Així, si un bloc apareix repetit durant el missatge original, sempre es xifra de la mateixa forma, i es pot desxifrar part del missatge sempre i quan correspongui a blocs complets.

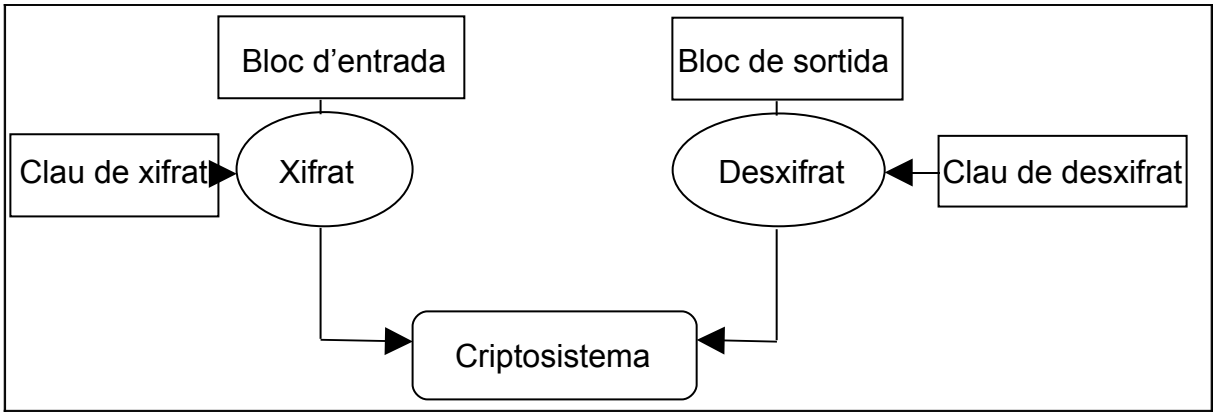

Figura nº 4.4: Generalització funcionament xifrat en bloc

Destacar dos xifrats en blocs importants: el **DES**, el més conegut durant el segle XX i **Rijndael** que va aparèixer durant l'any 2000. La importància d'aquests dos xifrats radica en el fet que van ser escollits pels EEUU com a estàndards del seu moment.

- **Algorisme DES**

A finals dels quaranta **Shannon** va plantejar noves idees per futurs sistemes criptogràfics. Les seves idees es movien en l'ús d'operacions múltiples que barregessin les mencionades anteriorment: *transposició* i *substitució*. Aquestes idees van ser aprofitades per la **IBM** en els setanta, al crear **LUCIFER** que va ser un sistema criptogràfic informatitzat que intentava aplicar aquests coneixements. Va ser l'any 1976, però, quan el govern dels EEUU va adoptar com estàndard un *sistema de xifrat* basat en el **LUCIFER** que es va anomenar **DES** (*Data Encryption Standard*). A partir d'aquesta decisió, casi tots els governs del món van decidir utilitzar el mateix tipus de xifrat en les seves comunicacions de xarxes bancàries i comercials.

L'**algorisme DES** matemàticament parlant es tracta d'un *algorisme de 64 bits* de clau dels quals 56 bits formen part de la part de la clau de xifrat pròpiament dita, mentre els 8 que sobren s'utilitzen per correcció d'errors.
El seu funcionament el següent: primer, es tradueix el missatge a una llarga sèrie de dígits binaris. Segon, la sèrie es divideix en blocs de 64 dígits, i es realitza la codificació separadament per cadascun dels blocs. Tercer, centrant-nos en només un dels blocs, els 64 dígits es remouen i després es divideixen en semiblocs de 32, denominats $Esquerra^0$ i $Dreta^0$. Els dígits de la $Dreta^0$ es sotmeten a una funció de deformació, que canvia els dígits segons una complexa substitució. La $Dreta^0$ deformada s'afegeix a l'$Esquerra^0$ per crear un nou semibloc de 32 dígits anomenat $Dreta^1$. La $Dreta^0$ original s'anomena ara $Esquerra^1$. Aquesta sèrie d'operacions s'anomena una ronda. Es repeteix el procés sencer en una segona ronda, però començant amb els nous semiblocs,

Esquerra[1] i Dreta[1], i acabant amb Esquerra[2] i Dreta[2]. Aquest procés es repetirà fins a 16 rondes en total.

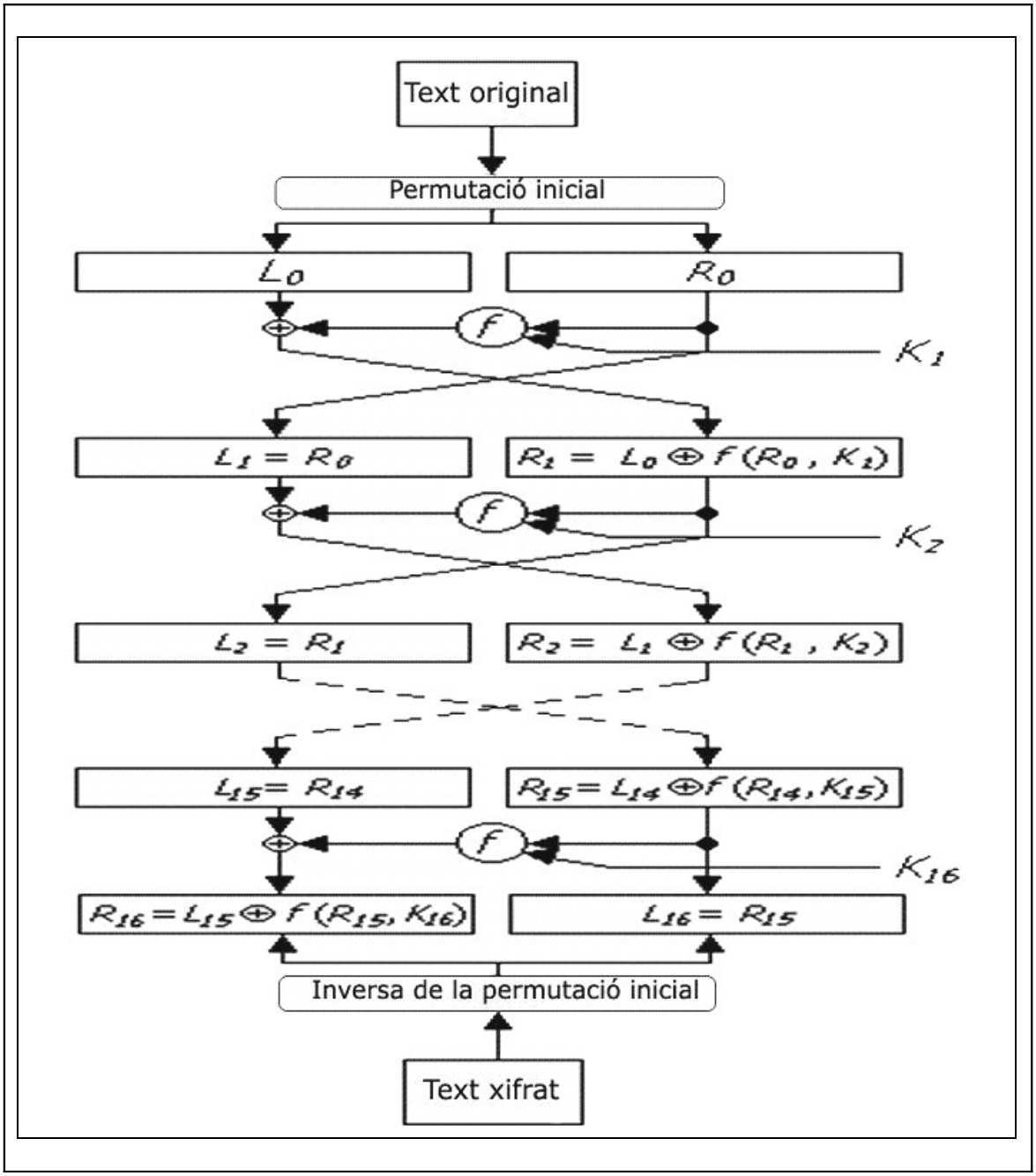

Figura nº 4.5: Funcionament algorisme DES

Mencionar que **DES** actualment no és un estàndard criptogràfic ja que va ser trencat el Gener de 1999 amb un sistema de computació que analitzava 250 · 10^9 claus per segon. El seu principal avantatge és la rapidesa de càlcul i la facilitat de la seva aplicació. Els seus principals defectes, però: la poca longitud

de clau amb que funciona, juntament amb la incapacitat de treballar amb claus de longitud variable.

- **Rijndael**

L'any 2000 el **NIST** (*National Institute for Standards and Technology*) va declarar com nou estàndard avançat de xifrat el conegut com **Rijndael**, per les inicials dels seus autors: els belgues *Rijmen* i *Daemen*. L'adopció d'aquest nou estàndard va ser realitzada després d'una convocatòria pública en la que es van proposar i analitzar obertament diversos algorismes.

Aquest tipus de xifrat no és com el de **DES**, treballa amb longituds de clau i bloc variables, en tots dos casos entre 128 i 256, d'aquí que el *criptoanàlisis exhaustiu* quedi descartat.

Podríem definir l'algorisme aplicat a Rijndael de la següent manera: si anomenem **B** al bloc de text a xifrar i **S** la matriu de l'estat, l'algorisme Rijndael amb **n** iteracions com:

Pas 1: Calcular les **n** subclaus $\mathbf{K_1,K_2,...K_n}$ a partir de la clau original **K**.
Pas 2: Obtenim $\mathbf{S:=B+K_0}$
Pas 3: Des de **i:=1** fins **n-1** fer
S:=Byte-sub(S)
S:=Desplaçar-fila(S)
S:=Barrejar-columnes(S)
$S:=K_i+S$
Pas 4: Fer
S:=Byte-sub(S)
S:=Desplaçar-fila(S)
$S:=K_n+S$

Un cop hem fet aquest últim pas, el bloc xifrat es troba comprès en la matriu d'estat **S**.

Si mirem la seguretat d'aquest algorisme jove, podem veure com fins ara ha passat totes les proves que se li han fet fins ara i els seus autors afirmen que no posseeix claus febles.

4.2.3. Xifrat en flux

El **xifrat en flux** és un xifrat per substitució que s'aplica sobre els caràcters o lletres dels texts combinant-los amb un **flux de bits secrets** (*seqüència xifrant*) mitjançant un operador * determinat. El desxifrat es duu a terme d'una forma

anàloga, combinant mitjançant l'operador $*^{-1}$ a les lletres del text xifrat amb la seqüència xifrant, produint així les lletres de l'original.

El primer que va proposar un xifrat on el text xifrat resultés de combinar mitjançant una operació eficient el text original amb un flux de bits secrets va ser **Vernam** *(1917).* El xifrat de Vernam es basa en l'operador o-exclusiu (suma de mòdul 2). Aquesta operació resulta especialment indicada per aquest tipus de xifrat ja que $*^{-1}=*$

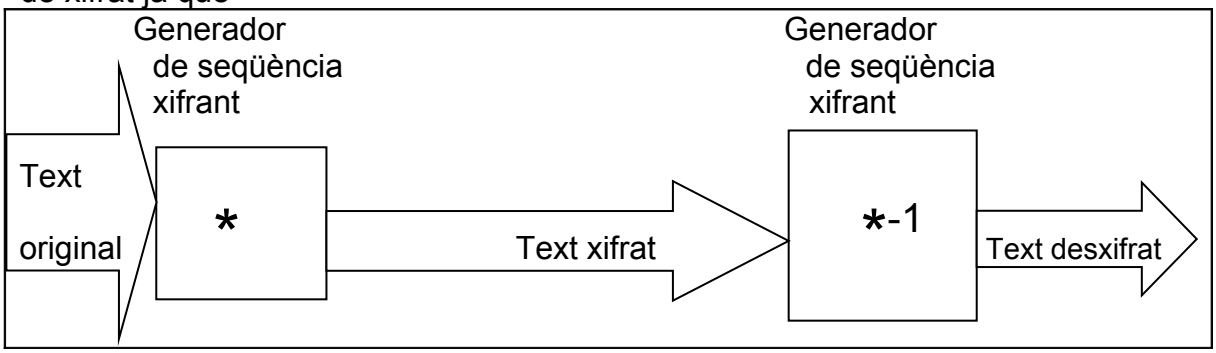

Figura Nº 4.6: Funcionament xifrat en flux

En el *xifrat de Vernam*, primer es codifica cada lletra de l'abecedari segons el codi de **Baudot** (1874), és a dir, A=11000,...,Z=10001, i després es suma en mòdul 2 la clau expressada en binari:

Missatge original:	01101001
Clau aleatòria:	11010011
Criptograma:	10111010

Shannon va demostrar que el *xifrat de Vernam* té *secret perfecte*, és a dir, és irrompible. Encara que cal matisar, la dificultat principal d'aquest procés ve donava per la dificultat de generar la mateixa seqüència per al desxifrat, ja que sent aleatòria, l'única forma d'obtenir-la seria mitjançant l'enviament des del propi emissor a través d'un canal segur.

4.3. Criptografia de clau pública

Entenem per **criptografia de clau pública** aquells *criptosistemes asimètrics* que a diferència dels simètric, per poder desencriptar un text encriptat hem d'utilitzar a més d'una *clau pública* una altre *clau que és privada*. Aquest fet permet enviar informació secreta a una altre persona sense necessitat de compartir cap clau secreta, solventant el problema logístic que representava la transmissió de claus en els criptosistemes clàssics o simètrics.

Els *desaventatges* que tenen aquest tipus de criptosistemes és que es tarden més a dur a terme que els simètrics ja que precisen de més operacions. Un altre problema d'aquest tipus de criptosistemes és l'autenticitat de la clau

pública, és a dir, com saber si la clau pública que fa referència a una persona és realment d'aquella persona i no estem enviant el nostre missatge directament a l'enemic.

Els aventatges que va suposar són a més dels que he mencionat abans el de l'aparició de l'autentificació de l'usuari que va permetre que es poguessin solucionar els problemes per assegurar que el missatge no havia estat alterat i que era de qui deia ser l'emissor.

La *seguretat* d'aquests tipus de sistemes es basen principalment en la dificultat computacional que comporta la factorització de nombres grans. Així la seva seguretat ve donada per la gran quantitat de temps que cal emprar per provar totes les combinacions possibles.

4.3.1. Inicis

Els inicis d'aquest tipus de criptografia el trobem en la ment de **Diffie** i **Hellman**. Aquests personatges van ser dos persones peculiars que veien la importància de la criptografia en la recent societat de la informació. Es van adonar que el sistema criptogràfic que es volia presentar com a sistema de seguretat, el **DES**, no era més que una forma de controlar els missatges de la gent, de poder saber més fàcilment per part de la **NSA** (*National Security Agency*) el que s'estava comunicant tot violant la intimitat de les persones. Aquests actes eren coneguts per *Diffie* i *Hellman* que sabien com aquesta organització es dedicava a interceptar i desencriptar suposades comunicacions perilloses. Això els va dur a pensar en una alternativa millor al DES i sent conscients dels problemes de la *criptografia simètrica* aplicada fins ara i de les necessitats que podrien sorgir en aquesta nova era de la informació, va sorgir la idea d'un tipus d'operació que permetés encriptar fàcilment, però que fos molt difícil de desencriptar pel qui no conegués la clau. Així va ser com se'ls va passar per la ment el buscar algun tipus d'operació que permetés fer això, una mena de representació de la facilitat de trencar un plat (*encriptar*) i la dificultat o directament, impossibilitat de poder a ajuntar-lo perfectament (*desencriptar*). Després d'anys d'investigació i temps de dedicació, l'any 1976 van topar amb una operació que semblava servir pels seus propòsits. Es tractava de les funcions unidireccionals, que eren tan difícils d'invertir com calcular un logaritme discret, que precisa al voltant d'un milió de milions de quadrilions més d'operacions que la exponenciació per poder transformar nombres.

Així va ser com va néixer la idea i com es va definir el camí a seguir. Aquests personatges van seguir investigant en el tema i difonent les seves troballes. Això va fer que més matemàtics i persones s'interessin en el tema cosa que al final va fer que no fossin ells els que trobessin el algorisme que buscaven.

Figura nº 4.7: Ralph Merkle, Martin Hellman, Whitfield Diffie

Ells van arribar fins a cert punt a solucionar el problema, però van ser *Ron Rivest, Adi Shamir i Leonard Adleman* qui l'any 1978 amb donar amb el que seria el criptosistema de clau pública més estès: el **criptosistema RSA**. Nom que ve donat per les inicials dels cognoms dels que van crear aquest algorisme.

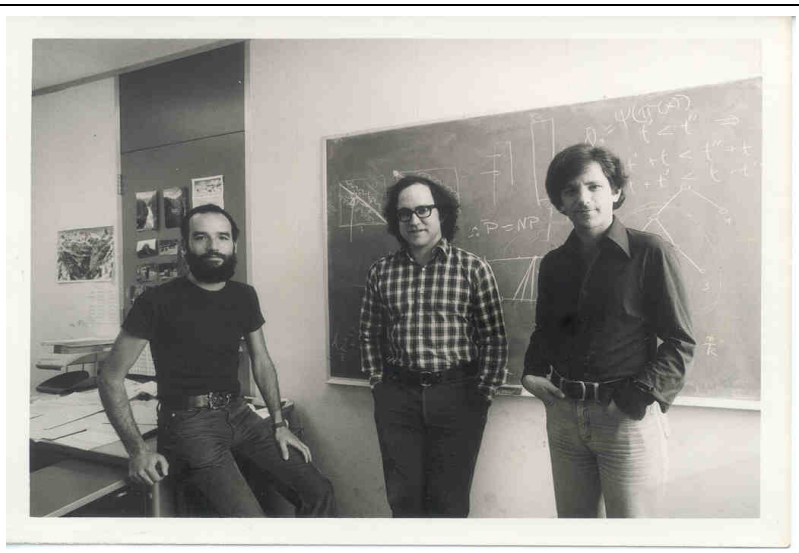

Figura nº 4.8: Adi Shamir, Ron Rivest i Leonard Adleman

4.3.2. Criptosistema RSA

Aquest és un dels criptosistemes de **clau pública** més estesos i utilitzats. Va representar una revolució en el camp de la criptografia trencant tot el que s'havia establert fins ara. Aquest sistema va ser introduït per *Rivest*, *Shamir* i *Adleman* l'any 1978 i és a ells que deu aquest nom tan peculiar.

La **seguretat** d'aquest sistema criptogràfic es basa en la dificultat que representa el repte de *factoritzar nombres grans*. Aquesta tasca és la que ha permès mantenir fins ara l'ús d'aquest criptosistema. Un criptosistema que deixarà de ser viable quan el món de la informàtica hagi avançat suficientment com per fer un gran nombre de càlculs per segon: el *món de la **informàtica quàntica***. Tot i així, la seva desaparició no serà, si és que és per la via de la informàtica quàntica, per haver pogut solucionar el problema de la factorització, sinó per poder aplicar atacs a **força bruta** que no és més que intentar totes les combinacions possibles per intentar trobar la combinació de la clau. Aquest tipus d'atac ara per ara no té gaire sentit ja que ens hi podríem estar molt de temps, però amb la realitat dels ordinadors quàntics, aquest temps seria reduït a períodes inimaginablement molt més petits que amb els ordinadors actuals.

- **Protocol RSA**

El Protocol RSA consta de tres parts:

  - **Generació de les claus**

En aquest primer pas cada membre **M** ha d'escollir les seves claus, tant la pública com la privada. Per fer-ho haurà de seguir el procés descrit a continuació:

- cada membre **M** escull dos nombres primers grans **p** i **q**. Amb aquests dos el que caldrà fer serà calcular el valor del seu producte **n= p · q**. El conjunt de missatges que emprarà cada membre **M** serà el grup multiplicatiu $\mathbf{Z_n^*}$. L'ordre d'aquest grup serà el nombre $\varphi(n)$

- a continuació el membre **M** escull un enter positiu **e, 2 < e <** $\varphi(n)$, de tal forma que sigui primer amb l'ordre del grup $\mathbf{Z_n^*}$. Això voldrà dir que **mcd (e,** $\varphi(n)$**) = 1**, on φ és *l'indicador d'Euler* que aplicat en el cas del producte de dos nombres primers (*p i q*) ve determinat per l'expressió:

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

- fent ús de *l'algorisme d'Euclides*, cada membre **M** calcula l'invers d'e a $Z^*_{\varphi(n)}$, **d**; on **d** serà l'únic enter que verifiqui que **1 < d <** $\varphi(n)$ de tal forma que se'ns compleixi que:

$$e \cdot d \equiv 1 \ (mod \ \varphi \ (n))$$

- ara el que tenim és que cada membre **M** té la seva *clau pública* formada per la parella (**n**,**e**) i la seva *clau privada* **d**.

Mencionar que també haurà de romandre en secret **p, q** i $\varphi \ (n)$ ja que si es descobreixen alguns d'aquests termes és pot fer tot el procés per calcular igualment la *clau privada*.

- **Xifrat de missatges**

Un cop descrit el procés per generar les claus el que cal fer és encriptar el missatge. Per fer-ho es seguiran els passos següents:

- primer de tot hem d'obtenir la *clau pública* **(n_D,e_D)** del nostre destinatari **D**. Per fer-ho haurem de seguir alguns dels mètodes exposats més endavant per aconseguir aquesta *clau pública* de llocs segurs.

- el pas següent és representar el nostre missatge m com un element $Z^*_{n_D}$, és a dir, com un enter en el ranc **{0,1,...,n_D-1}** i primer amb n_D

- finalment, calculem el valor del criptograma: $c = m^e_D \ (mod \ n_D)$ que enviarem al destinatari **D**.

Cal puntualitzar en el segon punt d'aquest procés que els missatges s'envien com a elements de Z^*_n ja que si no es fes així i volguéssim enviar un missatge més llarg, l'hauríem de dividir en trossos, fent que cadascuna d'aquestes parts fos d'un element d'aquest Z^*_n. Això es deu a que si el missatge **m** no fos primer amb **n**, com abans hem dit que $n = p \cdot q$, si apliquem que **m** no és primer amb **n**, per tant, que $m = p \cdot k$, per algun membre enter de **k**; tindríem doncs, que el criptosistema seria:

$c = m^e \ (mod \ n)$, d'aquí podem extreure que: $c = m^e + h \cdot n = p^e \cdot k^e + h \cdot n = p \cdot l$, d'on podem obtenir que $mcd \ (c, n) = p$ fent que el criptosistema quedés automàticament trencat, ja que coneixent **p**, podríem treure que $q = \frac{n}{p}$ i a partir d'aquests valors seria conegut el valor de $\varphi \ (n) = \ (p-1)(q-1)$ i per tant, el de la *clau privada* **d**, també.

- **Desxifrat de missatges**

Per acabar el destinatari **D**, quan ha rebut el missatge **m** xifrat, haurà de fer el càlcul següent:

Emprant la seva clau privada $\mathbf{d_D}$, calcula:

$$c^{d_D} = (m^{e_D})^{d_D} \text{ (mod } n_D) = m^{e_D d_D} \text{ (mod } n_D) \equiv m \text{ (mod } n_D)$$

Per poder entendre el procés de càlcul anterior i com s'arriba a la simplificació final cal entendre la proposició següent:

Donada la clau **(n,e)**, sent $\mathbf{n = p \cdot q}$, amb clau privada **d**, llavors per qualsevol enter $\mathbf{c = m^e \in Z_n}$, es verifica que:

$$C^d = (m^e)^d = m^{ed} \equiv (mod\ n)$$

Demostració: com que $\mathbf{e \cdot d \equiv 1 \cdot (}$ **mod** $\varphi\ (n)$**),** llavors existeix un enter **k** tal que:

$$e{\cdot}d = 1 + k\ \varphi\ (n) = 1 + k(p-1)(q-1)$$

com que **mcd (m,n) = 1**, obtenim que **mcd (m,p) = 1**, i en la *congruència de Fermat*:

$$m^p \equiv m(\text{mod } p)$$

simplifiquem dividint per **m**:

$$\frac{m^p}{m} \equiv \frac{m}{m} \text{ (mod p) i d'aquí}$$

$$m^{p-1} \equiv 1 \text{ (mod p)}$$

elevant ambdós membres d'aquesta congruència a **k(q-1)** i multiplicant-los per **m**, obtenim:

$$(m^{p-1})^{k(q-1)}\ m \equiv 1^{k(q-1)}\ m \text{ (mod p)} \text{ d'on:}$$

$$m^{k(p-1)(q-1)+1}$$

com que abans teníem que $\mathbf{e \cdot d = 1 + k(p-1)(q-1)}$, tenim que:

$$m^{k(p-1)(q-1)+1} = m^{ed} \equiv m \text{ (mod p)}$$

Ídem per demostrar que $\mathbf{m^{ed} \equiv m\ (\ mod\ q)}$. Com sabem que **p** i **q** són primers diferents i que $\mathbf{n = p \cdot q}$, aplicant el *Teorema xinès del residu* que diu que siguin $\mathbf{n_1,n_2,...,n_k}$ enters primers entre ells dos a dos, és a dir, on $\mathbf{mcd\ (n_i,n_j) = 1}$, $\mathbf{i = j}$; llavors el sistema de congruències simultànies definides per:

$$\begin{aligned} x &\equiv a_1 \pmod{n_1} \\ x &\equiv a_2 \pmod{n_2} \\ &\vdots \\ x &\equiv a_k \pmod{n_k} \end{aligned}$$

tindrà una solució única mòdul $\mathbf{n = n_1 \cdot n_2 \cdot ... \cdot n_k}$; podrem obtenir que:

$\mathbf{m^{ed} \equiv m\ (\ mod\ n)}$, per tant:

$$\mathbf{c^d \equiv (m^e)^d\ (mod\ n) \equiv m\ (\ mod\ n)}$$

- **Exemple**

Per poder entendre millor tot el **procés teòric** explicat fins ara exposo el següent exemple on es pot apreciar amb més o menys dificultat el procés a seguir pas a pas en una situació hipotètica i amb valors no comparables amb els emprats en una situació real quan fem servir el **criptosistema RSA**. Aquesta elecció no té altre raó que la d'evitar un excés numèric innecessari que pot acabar duent al desinterès. Alhora, l'ús de xifres grans pot derivar en la situació on ens haguem de creure al peu de la lletra tots els càlculs fets. És per tot això que com he mencionat abans, aquest exemple ha de ser entès com a tal, un exemple amb xifres petites que no és vàlid per demostrar l'alta eficàcia del nivell de seguretat d'aquest criptosistema. Així doncs, passem a l'exemplificació:

1. Busquem els dos nombres primers:

   p=43 i **q=53**

2. Amb els dos primers obtinguts anteriorment trobem el mòdul **n**:

   $\mathbf{n = p \cdot q = 2279}$

3. D'aquestes dades, calculem el valor de $\varphi(n)$**:**

   $\varphi(n)$ **= (p - 1) · (q - 1) = 2184**

4. Ara trobem un valor **e** que sigui menor que el mòdul **n = 2279** i també que **e** i $\varphi(n)$ siguin primers entre si:

e = 19

5. Per últim calculem **d** sabent que **(e · d – 1)** és divisible entre $\varphi(n)$.

(19 · d –1) / 2184 = 1 → d = (2184 + 1) / 19 = 115

6. La clau pública **(n,e)** serà **(2279,19)**, i la privada, **(n,d) (2279,115)**.

- **Criptoanàlisis elementals al RSA**

Tot seguit es mostraran els aspectes generals que cal tenir en compte a l'hora d'utilitzar aquest criptosistema:

- **Tipus d'atacs**

Per poder entendre millor la seguretat i els punts on s'ha de prestar més atenció alhora de dur a terme aquest criptosistema, val la pena, primer de tot, ficar-se en la pell de l'atacant o criptoanalista que és qui intentarà esbrinar el contingut del nostre missatge. Passem doncs a la classificació de l'atac:

Primer de tot remarcarem els aspectes que decidiran o confeccionaran les dues classificacions:

- Segons la disposició o actitud de l'atacant, és a dir, la relació que guarda l'atacant amb el missatge. Així dins d'aquest grup podem fer l'agrupació següent:

  - **Passius**: en aquesta situació l'atacant atempta contra la confidencialitat de les dades transmeses, però limitant-se només a observar el canal de comunicació per fer-ho.

  - **Actius**: aquest segon cas representa quan l'atacant té la intenció d'afegir, canviar o esborrar de la forma que sigui la transmissió que s'està duent a terme. Per tant, aquest atacant està atemptant contra l'autenticitat, la integritat i la confidencialitat de les dades.

- Segons el tipus d'atac que es dugui a terme podem diferenciar-hi dos grups:

  - **Atacs passius**: aquí es pretén obtenir un text clar a partir d'un text encriptat, és a dir, obtenir la clau de desxifrat. Podem distingir-hi els diferents casos:

- *Atac al text encriptat*: l'atacant intenta deduir la clau a partir del text xifrat.
- *Atac al text desencriptat*: l'adversari té a les seves mans una determinada quantitat de texts desencriptats i els seus texts encriptats corresponents.
- *Atac al text desencriptat escollit*: el criptoanalista escull un text desencriptat i obté també el text encriptat corresponent, de tal manera que la informació que pugui deduir d'aquest coneixement pugui aplicar-se per tractar de desencriptar un nou criptograma o obtenir la clau
- *Atac al text encriptat escollit*: l'enemic selecciona un text encriptat i obté la seva correspondència en text desxifrat.
- *Atac adaptable al text desencriptat escollit*: com en el cas de l'atac al text desencriptat escollit, però podent seleccionar de nou un altre text en funció dels resultats de l'encriptament previ.
- *Atac adaptable al text encriptat escollit*: com en el cas anterior, però seleccionant el text encriptat.

- **Atacs actius**: en aquest segon cas el que es vol és alterar el contingut del missatge. Hi podem diferenciar:

  - *Atac a meitat de camí*: el criptoanalista aconsegueix interceptar la comunicació. Intercepta el missatge de l'emissor i el canvia pel seu propi missatge enviant-lo al receptor tot fent-se passar per l'emissor.
  - *Atac de la temporalització*: en aquest cas el criptoanalista busca aconseguir la informació referent al temps que tarda en executar-se l'algorisme i tot el procés de desencriptació, com també del nombre d'operacions que aquest realitza. Amb la possessió d'aquests coneixements pot arribar a conèixer informació sobre la llargària de la clau privada o altres aspectes que el poden encaminar al descobriment d'aquesta clau.

o Elecció dels primers p i q

Alhora de realitzar la implementació del **criptosistema RSA**, el primer pas a seguir era el d'escollir dos nombres **p** i **q** primers. Aquests dos nombres però, han de complir una sèrie de característiques si volem que sigui òptima la seguretat que ens proporciona aquest sistema ja que com amb el cas d'**Enigma** (1923), la seva desencriptació va ser facilitada per errors humans externs a la pròpia seguretat del sistema. Per tant, en aquest primer pas hauríem de tenir en consideració els aspectes següents per aconseguir dificultar el màxim possible el procés de factorització d'**n**:

1. **p** i **q** han de tenir la mateixa longitud

   Els dos nombres escollits han de ser grans i tenir, si fa o no fa, la mateixa longitud. Això és debut a què si un dels dos nombres és molt més gran que l'altre, el petit serà més fàcil de trobar i per tant, per trobar l'altre nombre només ens caldria dividir **n** per aquest nombre.

2. El **mcd (p-1,q-1)** ha de ser petit

   Aquest aspecte té a veure amb que si tenim que **mcd (p-1,q-1)** és un nombre gran, tindríem:

   $$\mathbf{u = mcm\ (p-1,q-1) = \frac{\varphi(n)}{mcd\ (p-1,q-1)}}$$

   obtindrem una **u** petita en comparació amb $\varphi(n)$. Així si apliquem que qualsevol invers d'**e** amb mòdul **u**, **d'**; fent que $\mathbf{e \cdot d' \equiv 1\ (mod\ u)}$ ens servirà d'exponent de desxifrat ja que:

   $$\mathbf{m^{ed'} \equiv 1\ (mod\ n)}$$

   per tot missatge **m**. Doncs, com podem veure, si tenim una **u** petita, s'empra aquesta propietat per trencar el criptosistema seguint els passos següents:

   - Escollim un valor **u** i calculem l'invers d'**e** mòdul **u**
   - Encriptem diversos missatges amb el criptosistema i provem a desxifrar-los utilitzant l'exponent **d'**.
   - Si en algun dels intents del segon pas aconseguim desencriptar el missatge, el criptosistema ha quedat trencat. Si no ho aconseguim, tornem a repetir des del primer pas.

3. **p** i **q** han de ser robustos

   Això vol dir que compleix aquestes tres condicions:

   - **p - 1** té un factor primer gran **u**
   - **p + 1** té també un factor primer gran **v**
   - **u - 1** té un factor primer gran **w**

   Aquestes condicions s'entenen gràcies al fet que amb aquest compliment s'aconsegueix evitar (*o s'intenta*) els atacs al criptosistema RSA mitjançant els algorisme **p – 1** de *Pollard* (1974) i **p + 1** de *Williams* (1982)

4. **p-1** i **q-1** han de tenir factors primers grans

   Això és perquè si no es compleix aquesta condició s'obtindria que els factors de $\varphi(n)$ **= (p-1)(q-1)** serien petits també. Així si tenim per cas que tots els factors d'$\varphi(n)$ són menors a una cota **J**, seria factible trobar tots els possibles candidats **w** a ser $\varphi(n)$ i alhora intentar veure si el missatge xifrat elevat a la potència d' **(w + 1) / e**, fent que doni un nombre enter, proporcioni un missatge comprensible.

5. **p** i **q** no han d'estar massa pròxims

   Cal complir aquesta condició ja que sinó es podria aplicar el següent plantejament:

   Si **p<q**, tenim doncs que $\frac{q-p}{2}$ seria petit perquè com hem dit abans, ens trobem en el cas d'haver agafat dos nombres primers propers. Així tindríem que:

   $$\frac{p+q}{2} \cong \sqrt{n}$$

   de tal manera que:

   $$\left(\frac{p+q}{2}\right)^2 - n = \left(\frac{p+q}{2}\right)^2$$

   doncs aquesta diferència seria petita. Com el segon membre és un quadrat perfecte, podem esbrinar amb més o menys intents, un enter **x >** $\sqrt{n}$, tal que $\mathbf{x^2 - n = y^2}$, d'on obtenim que:

   $$\mathbf{p = x - y} \quad \text{i} \quad \mathbf{q = x + y}$$

   fent que quedés factoritzat.

- **Elecció de l'exponent de xifrat e**

Normalment un cop s'ha obtingut un valor gran pel *mòdul RSA*, el que es sol fer és escollir un *exponent de xifrat* **e** petit, ja que a l'hora de realitzar el procés d'encriptació del missatge haurà de dur a terme l'operació $\mathbf{m^e \pmod n}$, i el que volem és facilitar la feina de computació perquè sigui eficient.

Així com amb l'elecció dels nombres **p** i **q** hi havia unes certes recomanacions, trobem per aquest unes característiques especials:

1. L'exponent de xifrat ha de ser petit

   Com deia al principi d'aquest apartat l'exponent **e** es recomana agafar-lo petit. És més, directament es recomana que sigui **e = 3** o **e = 65537**. La raó d'aquests dos nombres la trobem en els següents fets:

   - Tant un com l'altre són nombres primers i l'expressió binaria d'aquests dos nombres és molt senzilla i facilita l'estructuració de forma que el càlcul de valors de $\mathbf{m^e \pmod n}$ sigui més fàcil. Això el que farà serà augmentar la velocitat de càlcul d'aquest algorisme ja que també hem de trobar viable i senzilla la forma d'encriptació. Encara que un ordinador pugui semblar que mai s'equivoca, si té algun error de programació pot generar errors. És per això que s'intenta fer que aquest exponent faciliti el màxim possible la feina del processador.

2. Exponent de xifrat comú

   Aquest títol fa referència al fet que diversos usuaris poden tenir el mateix valor d'**e**, però això sí, tots han de tenir diferents valors d' **n**. Si no fos així, com un usuari sap les seves dues claus (**e,d**) podria saber l'exponent de desxifrat de tots els altres usuaris, ja que tots tindrien la mateixa clau.

3. Exponent petit amb missatges iguals

   Quan ens trobem davant d'un cas com és el d'enviar un mateix missatge a diferents persones es recomana no utilitzar un valor petit d'**e**. Per veure-ho més clar: si intentéssim enviar un missatge a 4 persones diferents amb valor d'**e = 3**, tenint 4 mòduls diferents $\mathbf{n_1}$, $\mathbf{n_2}$, $\mathbf{n_3}$ i $\mathbf{n_4}$, tindríem els següents criptogrames: $\mathbf{c_i = m^3 \pmod{n_i}}$. Amb això un criptoanalista que interceptés els missatges es podria plantejar un sistema de congruències simultànies com aquest:

$$\mathbf{x \equiv c_1 \pmod{n_1}}$$
$$\mathbf{x \equiv c_2 \pmod{n_2}}$$
$$\mathbf{x \equiv c_3 \pmod{n_3}}$$
$$\mathbf{x \equiv c_4 \pmod{n_4}}$$

   aplicant el *Teorema xinès del residu* tal i com s'ha mencionat abans, si els mòduls RSA fossin primers entre sí dos a dos, obtindríem que $\mathbf{x = m^4 < n_1 n_2 n_3 n_4}$. Traient l'arrel quarta d'**x**, recuperaríem el missatge original.

4. Exponent petit amb missatges curts

   Cal evitar finalment, xifrar missatges curts quan fem ús d'exponents de xifrat petits ja que si se'ns verifica que $\mathbf{m < n^{1/e}}$, llavors **m** es pot recuperar utilitzant l'expressió $\mathbf{c = m^e \pmod n}$ calculant només l'arrel e-èssima d'**c** pels nombres naturals, que existiria ja que **m** és precisament aquesta arrel.

- **Elecció de l'exponent de xifrat d**

L'exponent de desencriptació **d** ha de tenir una longitud semblant a la d'**n**. Per fer-ho més exactament s'estableix la relació següent:

$$\text{Longitud en bits } (\mathbf{d}) \leq \frac{1}{4} \text{ longitud en bits } (\mathbf{n})$$

- **Propietats del missatge**

En alguns casos alguns detalls del missatge faciliten la feina als criptoanalistes. Així com amb **Enigma** la descoberta de paraules clau que van ser utilitzades per desencriptar més ràpidament els missatges i per trobar característiques sobre aquests, el mateix passa amb l'RSA. El que hem de tenir en compte és evitar missatges petits o continguts previsibles o els mateixos processos descrits en l'apartat que tracta sobre el missatge perfecte.

- **Criptoanàlisis especialitzat al RSA**

Per acabar parlarem de l'existència d'un **criptoanàlisis especialitzat**. Així com fins ara s'han presentat coses a tenir en compte a l'hora d'escollir els diferents membres del **criptosistema RSA** ja que eren susceptibles a patir anàlisis fent ús de *Teoremes generals*, també existeixen uns sistemes que analitzen la comunicació des del punt de vista específic de l'RSA. Trobem, doncs, diferents tipus de criptoanàlisis que busquen explotar algunes debilitats de l'RSA que en molts dels casos no són tant conegudes com les nombrades amb anterioritat.

Mencionar l'existència d'aquest tipus de criptoanàlisis i sobretot, la de l'anomentat *criptoanàlisi de Wiener* que va presentar l'any 1990 un algorisme eficient, és a dir, en quant a temps polinòmic; que ens permet estudiar la seguretat del criptosistema RSA aplicant-lo per calcular l'exponent de desxifrat o clau privada **d** del criptosistema.

4.4. Aplicacions criptogràfiques

El sentit de la **criptografia** com s'ha mencionat des d'un principi és el de poder dur a terme comunicacions segures sobre canals insegurs. Però aquest sentit inicial que va prendre la criptografia es va veure millorat amb l'aparició de la criptografia de *clau pública*. La **criptografia de clau pública** va permetre un fet que fins ara no s'havia pogut aconseguir: l'autentificació de l'emissor. Aquest fet havia causat molts problemes a nivell històric amb suplantacions que van significar la mort de reis, de soldats i en definitiva de gent que va morir per no poder corroborar si el missatge que rebia era realment de l'emissor. Així podríem dir que aquesta nova criptografia, la pública, va venir també per donar pas al futur que ja s'entreveia: *comerç electrònic* i *món de les comunicacions*. Una nova societat que precisava de seguretat i d'unes garanties per a poder desenvolupar-se. És, precisament gràcies a aquesta criptografia de clau pública que tenim ara el desenvolupament que tenim a nivell comerç cibernètic i tota la cultura d'Internet.

Podem aplicar en aquest sentit l'ús nou que va aportar la criptografia de clau pública i l'RSA:

4.4.1. Autentificació

Aquesta característica de l'RSA i de la *criptografia pública* en general no és més que la que ens permet detectar possibles intrusos. Al parlar d'intrusos es fa referència al tipus d'**atacs actius**, on el criptoanalista canviarà l'estat original del missatge. Així a partir de la firma digital podrem assegurar-nos de dues coses:

1. El missatge prové efectivament de l'emissor que creiem
2. Podrem saber si algú n'ha variat alguna part

Així direm que l'**autentificació** en general prevé estafes fetes per tercers i la firma digital, permet a més a més d'això, l'estafa feta per comunicants il·legítims.

4.4.2. Firma digital

La **firma digital** com s'ha mencionat abans és la resposta a l'*avenç en el camp de les telecomunicacions*. Aquesta resposta ha vingut donant suport al *comerç electrònic*, enviament d'informacions confidencials entre moltes altres coses.

El nom de firma digital el rep ja que, tal i com el seu nom indica estem davant de la *versió digitalitzada de la nostra firma*, actualment amb la mateixa validesa que aquesta. Per donar un exemple: amb els nous **DNIs digitals** el que es fa és la introducció d'una firma digital personal e intransferible que permetrà

encara més el desenvolupament de la societat informàtica, una societat basada en el món de la informació i les telecomunicacions. Per poder arribar a aquesta societat desenvolupada cal, però una seguretat pels usuaris que ve amb la forma de la firma digital. Aquesta té unes característiques principals descrites a continuació:

- *No és falsificable*: o en el seu cas, l'intent de falsificació ha de dur a la resolució d'un problema numèric intractable.
- *Fàcil d'autenticar*: és a dir, podent qualsevol receptor establir la seva autenticitat encara després de molt de temps després d'haver rebut el missatge.
- *Irrevocable*: l'autor de la firma no pot negar la seva autoria.
- *Barata i fàcil de generar*.

Mencionar que clarament ha de ser única.

Així en el cas del criptosistema RSA tenim que la seva firma digital ve definida pel procés que segueix:

Siguin ($\mathbf{n_D}$,$\mathbf{e_D}$) i $\mathbf{d_D}$ les *claus pública* i *privada* d'un destinatari **D**, i siguin les pròpies del remitent **R**, ($\mathbf{n_R}$, $\mathbf{e_R}$) i $\mathbf{d_R}$. Si el destinatari **D** vol enviar, junt amb el criptograma, **c**, corresponent al missatge **m**, la seva firma digital l'haurà de calcular fent:

1. **D** calcula el valor de la seva firma pel missatge **m**. Per fer-ho busca el valor de:

$$r \equiv m^{d_D} \ (mod\ n_D)$$

2. Després xifra el valor anterior amb la clau pública d'**R**,

$$s \equiv r^{e_R} (mod\ n_R)$$

El missatge firmat que **D** envia a **R** és la parella formada per (**c, s**), sent c el criptograma corresponent al missatge **m**. No cal dir que només **D** pot firmar el missatge anterior, ja que ell és l'únic que coneix la seva clau privada, $\mathbf{d_D}$.

Ara, per fer que **R** pugui verificar que la firma correspon a **D** només ha de fer el següent:

1. **R** recupera la firma de **D** pel missatge **m** calculant:

$$s^{d_R} \ (mod\ n_R) \equiv r$$

Això ho podem aplicar ja que:

$$s^{d_R}\ (mod\ n_R) \equiv (r^{e_R}(mod\ n_R))^{d_R}\ (mod\ n_R) \equiv\ r^{e_R d_R}(mod\ n_R) \equiv r$$

2. Després, **R** comprova que la firma xifrada coincideixi amb el missatge:

$$r^{e_D}(mod\ n_D) \equiv m^{d_D e_D}(mod\ n_D) \equiv m$$

Si a l'hora d'aplicar el punt 2 no coincidís el valor d'**m** obtingut amb el procés de desencriptació, això voldria dir que el missatge rebut no serà vàlid ja que ha estat manipulat en algun punt de la transmissió o perquè el remitent no és qui diu ser.

4.4.3. Certificació de clau pública

Per aconseguir les característiques anteriors es tendeix a recórrer a terceres companyies, en el cas d'Espanya, a la **Fàbrica Nacional de Moneda i Timbre**-*Real Casa de la Moneda* és l'encarregat d'assegurar aquestes característiques. És a dir, el que fa aquest organisme és donar seguretat a l'emissor perquè vegi que al receptor al qual vol enviar el missatge és realment el que creu i no pas un personatge que el suplanti. Per aconseguir una **firma digital** s'haurà de demanar a aquest organisme el qual demanarà *documentació oficial* i el *pagament d'una quota* pel manteniment i oferiment del servei. A canvi, tu dones més seguretat a l'emissor i a tu mateix.

Tot aquest procés també vol acabar amb un dels problemes de la criptografia pública: la **suplantació**. Aquest problema es detecta quan una persona es fa passar per un altre donant una clau pública que no és de la persona que diu ser. Així, l'emissor tot confiat envia el missatge a aquest personatge pensant que l'està enviant a la persona que deia ser, quan la realitat era ben diferent. Així neixen el que s'anomenen **Autoritats de Certificació**. Aquestes Autoritats de Certificació no són més que terceres parts de confiança que garanteixen l'associació entre la identitat de l'usuari i la seva clau pública. El format de certificat digital més estès és l'estàndard **X.509** que podríem definir com un procés pel qual es defineixen diversos camps com: firma digital de l'autoritat certificadora, temps de validesa de la firma digital, nom i altres dades que poden resultar necessàries per a la confirmació de l'usuari.

Finalment destacarem l'aplicació que aquests processos han tingut en la nostra societat ja que en 1999 es va començar a fer la declaració de la Renta a través d'Internet. Aquest acte va ser possible gràcies a l'existència d'aquestes característiques que garantien la seguretat de l'usuari i també per part del govern que ha mostrat en aquest camp un interès especial fent menció a l'aparició del *Real Decreto Ley 14/1999 sobre Fima Electrònica* on s'expressen

les ganes que aquestes noves tecnologies criptogràfiques s'apliquin ràpidament en el territori espanyol. Podríem dir que en aquest aspecte es va posar al país a la vanguardia de la legislació en matèria criptogràfica amb la finalitat, com deia abans, de promoure el desenvolupament de la societat de la informació. També el mencionat **DNI digital** que forma part del nou pas que pretén fer el govern en matèria criptogràfica ja que aquest no seria possible sense una bona seguretat criptogràfica.

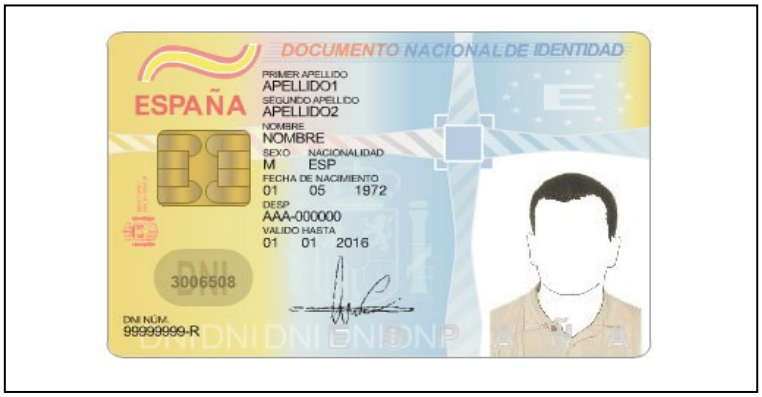

Figura nº 4.9: nou DNI electrònic

APARTAT PRÀCTIC

V DISSENY D'UNA APLICACIÓ INFORMÀTICA

5.1. Base matemàtica: l'algorisme

En aquest punt m'aturaré per explicar la base matemàtica que gira al voltant del procés de creació d'un programa o programació: l'algorisme.

Per **algorisme** entenem tal i com podem trobar al *Diccionari de la Llengua de la Gran Enciclopèdia Catalana*:

"Un algorisme és un procediment de càlcul que consisteix a acomplir un seguit ordenat i finit d'instruccions que condueix, un cop especificades les dades, a la solució que el problema genèric en qüestió té per a les dades considerades."

Anàlogament, doncs, podem comparar o entendre un algorisme com un conjunt de passos que hem de seguir per aconseguir resoldre un problema, com una espècie de recepta de cuina. Per poder entendre millor el concepte d'algorisme el que haurem d'entendre també són els conceptes d'*entorn*, *acció*, *procés* i *processador*.

- **Entorn**: conjunt d'objectes necessaris per a portar a terme una tasca determinada. Entenem per estat de l'entorn com a la descripció en un moment precís de l'estat dels objectes de l'entorn en aquell precís moment ja que un algorisme fa que vagin canviant progressivament l'estat del seu entorn.
- **Acció**: esdeveniment finit en el temps, que té un efecte definit i previst. L'acció, el que fa és actuar sobre un entorn modificant-lo,
- **Procés**: execució d'una o diverses accions. L'algorisme marca unes pautes a seguir per a resoldre el problema. L'encarregat de realitzar aquests processos és el processador.
- **Processador**: entitat capaç de comprendre i executar eficaçment un algorisme.

Un cop entesos aquests processos només ens falta entendre com queda registrat aquest **algorisme**, és a dir, quin llenguatge s'utilitza per a la seva creació. Analitzant el que caldria que fos aquest llenguatge hauríem de trobar un llenguatge que ens permetés fer una descripció no ambigua i precisa de les accions que caldria que realitzés el algorisme desitjat. Precisament per poder satisfer aquestes necessitats es crea el ***llenguatge algorísmic*** que es diferencia amb el nostre llenguatge o *llenguatge natural* ja que aquest és molt més reduït i precís. Si intentéssim plantejar els processos que ha de seguir un algorisme amb el nostre llenguatge natural veuríem com la seva intrínseca

complexitat i ambigüitat farien que no es poguessin dur a terme amb precisió i claredat les accions que ens haguéssim proposat.

Finalment, comentar que les qualitats que a de tenir un bon algorisme són:

- **Correctesa**: l'algorisme ha de fer allò que es demana.
- **Intel·ligibilitat**: l'algorisme ha de ser clar i fàcil de comprendre ja que ha de quedar obert a futures revisions i millores.
- **Eficiència**: l'algorisme ha de fer el que se li ha programat en un temps raonable.
- **Generalitat**: l'algorisme s'ha d'adaptar a altres problemes semblants amb pocs canvis.

5.2. Llenguatges de programació

Podem entendre un **llenguatge de programació** com un conjunt de símbols i caràcters o notació que es combinen entre si seguint unes regles marcades per una sintaxi predefinida. La raó d'aquesta acció; el poder permetre la transmissió d'instruccions a un ordinador. Aquests símbols i caràcters són traduïts per l'ordinador a un conjunt de senyals elèctriques representades en sistema binari (només dos valors: 0 i 1). Aquest procés s'entén ja que el processador només entén aquest tipus de llenguatge, el qual és conegut amb el nom de **llenguatge màquina**.

Dins dels llenguatges de programació hi podrem distingir dos tipus:

- **Llenguatges de baix nivell**: en aquest grup hi trobarem els llenguatges que per les seves característiques es troben més propers a la naturalesa interna de l'ordinador com poden ser el *llenguatge màquina* o el *llenguatge de l'assemblador*.

  El *llenguatge màquina* és aquell llenguatge molt bàsic el qual necessita un ordinador per poder entendre les nostres ordres formulades en un llenguatge natural. Precisament la característica d'aquest llenguatge és que és l'únic comprensible per l'ordinador. I el *llenguatge assemblador* és una mena d'evolució del llenguatge màquina. Basat en l'ús d'abreviatures de paraules que indiquen noms d'instruccions, el que pretén és ocupar molt menys espai de memòria i executar-se més ràpidament que amb un llenguatge de nivell alt; però amb el problema principal que són tan específics que no són transferibles a altres ordinadors amb processadors diferents ja que han estat creats específicament per una arquitectura determinada.

- **Llenguatges d'alt nivell**: a diferència de l'altre grup, en aquest trobarem tots els llenguatges de programació que per les seves característiques s'assemblin més al nostre llenguatge natural.

  El que fa importants aquests llenguatges és que són independents de l'arquitectura de l'ordinador. Això fa que un programa es pugui fer funcionar en ordinadors amb processadors diferents. Aquest fet permet també que el programador no hagi de conèixer a fons totes les característiques de l'ordinador pel qual ha de fer el programa, sinó que es pot abstreure dels detalls de baix nivell. Ara, aquesta mateixa característica no evita que s'hagi de transformar a llenguatge màquina perquè pugui ser entès per l'ordinador. Així doncs, s'haurà de dur la traducció de llenguatge d'alt nivell a llenguatge màquina amb l'ús de compiladors.

  Al llarg de la història han aparegut diferents llenguatges d'alt nivell, en aquest cas, es poden observar en la figura que hi ha a continuació:

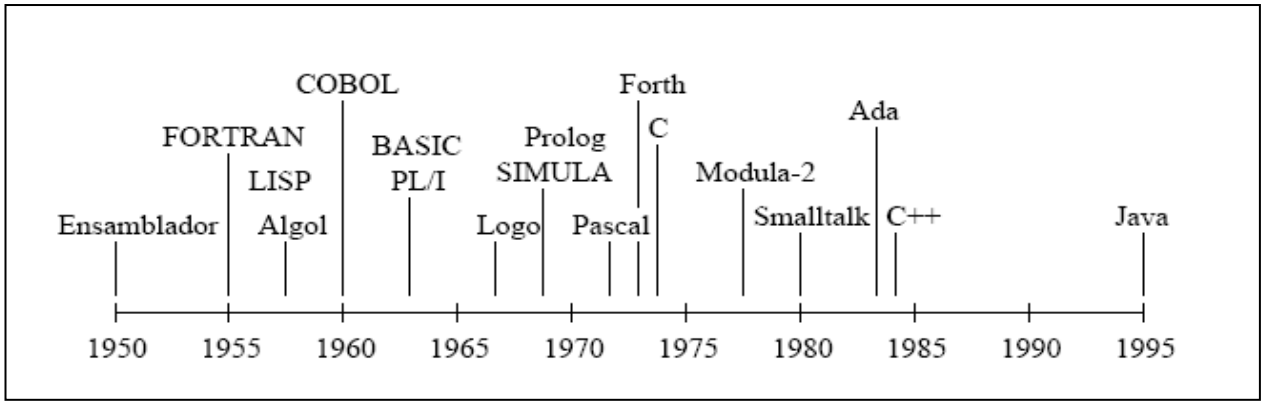

Figura nº 5.1: Cronologia d'alguns llenguatges de programació

5.3. Fases de programació

En el procés de creació d'un programa hi podem distingir les etapes següents:

1. **Definició del problema. Anàlisi de requeriments**: en aquesta etapa el que farem serà plantejar el problema que vulguem solucionar. Primerament, haurem de dur a terme un anàlisi del problema. Un cop hem comprés el problema el que cal fer és un anàlisi minuciós dels requeriments que ajudaran a definir el problema que volem resoldre.
2. **Disseny de l'algorisme**: en aquest punt del procés el que cal fer és la representació del disseny mitjançant el llenguatge algorísmic. Aquesta

etapa és una de les més importants i costoses dins el desenvolupament d'un programa.

3. **Implementació del programa**: ara el que es farà serà portar a terme el programa, és a dir, convertirem el resultat del disseny de l'algorisme d'un llenguatge algorísmic a un de programació que l'ordinador pugui entendre. Per fer-ho es procedirà de forma mecànica a la traducció del disseny ja realitzat per poder obtenir un executable per a l'ordinador.
4. **Proves**: arribant al final del projecte el que haurem de fer serà buscar possibles errors en el nostre programa. En aquest pas no buscarem pas d'assegurar-nos que l'algorisme és correcte ja que per fer-ho hauríem de provar totes les possibles entrades, cosa en la majoria dels casos impossible. Així doncs el que es mira és de veure que les proves no donin cap error.
5. **Manteniment i millores**: finalment, en aquest punt mirarem de mantenir el programa al dia, amb això vull dir: adaptar-lo a les noves necessitats que puguin sorgir, solucionar problemes que sorgeixin en el futur,etc. Per poder dur a terme fàcilment aquesta part del treball haurem d'haver fet prèviament un disseny del programa de forma intel·ligible i clara ja que un programa s'escriu varies vegades, però és llegit molts cops més. Sinó es fa tal i com he mencionat es podrien tenir problemes d'eficiència i eficàcia en aquesta etapa.
   Afegir que el programa haurà d'estar ben documentat en un arxiu extern al programa que permeti tenir tota la informació necessària sobre el programa

5.4. Ordinograma

Els **ordinogrames** són la representació gràfica del procés marcat per un algorisme. La funció de l'ordinograma és doncs la de *mostrar com evolucionen les variables* durant el conjunt de processos que marca el nostre programa.

Aquest ordinograma es durà a terme abans de fer el pas de **codificació** en l'apartat d'implementació de l'algorisme. Podríem dir que és el pas previ per poder estructurar visual i mentalment allò que haurem d'expressar en el llenguatge de programació.

Els *elements de representació gràfica* que podem trobar en un ordinograma són diversos, entre els quals podem trobar els que hi ha a continuació:

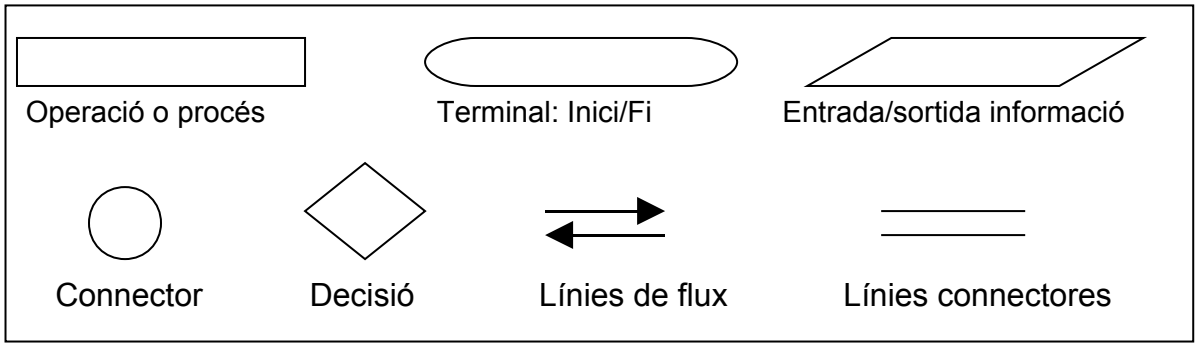

Figura nº 5.2: Elements bàsics d'un ordinograma de flux

- **Operació o procés**: serveix per senyalar l'operació que ha de fer l'ordinador
- **Terminal: Inici/fi**: indicador del començament i final del programa
- **Entrada/Sortida informació**: en aquest camp s'especificarà l'introducció d'una variable mitjançant el teclat i la sortida d'informació a través de la impressió a la pantalla.
- **Connector**: en aquest punt es troben els dos camins que s'han obert a partir d'una decisió.
- **Decisió**: bifurcació del camí que distribueix el camí segons si es compleix una condició de cert o fals preestablerta anteriorment.
- **Línies de flux**: indicacions del sentit d'execució de les operacions.
- **Línies connectores**: uneixen dos símbols que són diferents

Considerar també les següents regles a l'hora de crear un ordinograma:

- Tot **flux** ha de tenir un principi i un final com a mínim.
- A tot **símbol** (*excepte terminal i caràcters*) ha d'arribar com a mínim una connexió i d'aquest ha de partir una direcció.
- Quan s'empra un **símbol**, s'ha d'escriure dins d'ell
- La **direcció de flux** s'executa cap avall, cap a la dreta, cap a l'esquerra i en algunes ocasions cap amunt.
- Per cada **acció** s'utilitza el seu símbol corresponent.
- Les **variables** que s'utilitzin en un flux, han de ser prèviament omplertes amb informació.
- Quan un problema és molt complicat, s'escriu primer l'ordinograma general.
- Els **ordinogrames** s'escriuen de dalt cap avall.
- El **connector** s'ha d'evitar fins on sigui possible.

VI ANÀLISIS D'UNA APLICACIÓ PER SIMULAR EL CRIPTOSISTEMA RSA

En aquest apartat farem un anàlisis d'una aplicació per **simular el criptosistema RSA**. Aquesta elecció es deu a voler respondre a la realitat afrontada: aprendre un llenguatge de programació no és gens fàcil i partir de zero tot creant una aplicació basada en un criptosistema RSA no és feina del tot senzilla.

És per això que hem volgut dur a la pràctica una altre de les fases del procés de programació: la de proves i la de manteniment i millores. El que es farà doncs serà cercar una creació existent i analitzar-la.

6.1. Presentació de l'aplicació

L'aplicació que hem escollit per fer la part pràctica del treball de recerca és una creació feta per uns estudiants de la carrera d'Informàtica de la Universitat de la Palma de Gran Canària. Les raons per les quals s'ha escollit aquest treball són les exposades a continuació:

- La creació està **feta en JavaScript**, és a dir, un tipus de llenguatge de programació especial. Especial en el sentit que és dels anomenats llenguatges interpretats o el que és el mateix: no necessita ser compilat. Això facilita molt la feina ja que l'aplicació o simulació sol ser utilitzada en pàgines html i es pot fer servir en qualsevol ordinador que tingui un explorador que entengui el llenguatge: Internet Explorer 6, Netscape Navigator, Opera 7 o Mozilla en qualsevol de les seves versions.
- No s'està trencant ni fent **cap violació de drets de llicència** ja que la seva creació està lligada a un tipus de llicència creative commons que permet la seva alteració i publicació tot complint unes condicions.
- En la seva pàgina web expressen el **valor didàctic** de la seva creació. Encara que no la especifiquin en aquest sentit expressament ja que no posen el codi font directament, si que han creat la simulació per mostrar com funciona a nivell senzill l'algorisme RSA
- Enllaçant amb el punt anterior: com deia es tracta d'una creació que utilitza una **simplificació** de l'algorisme que s'utilitza a l'RSA. A més està bastant limitat ja que no deixa l'elecció dels nombres primers, sinó que estan prefixats. Tot això farà que el codi font no sigui d'un nivell molt

elevat i el que es vol, després de tot, és veure com s'aplica l'algorisme RSA en el camp de la programació; cosa que si ens poséssim a analitzar una creació molt complexa no compliríem gaires dels objectius que s'intenten aconseguir.

6.2. Primer apropament a l'aspecte visual

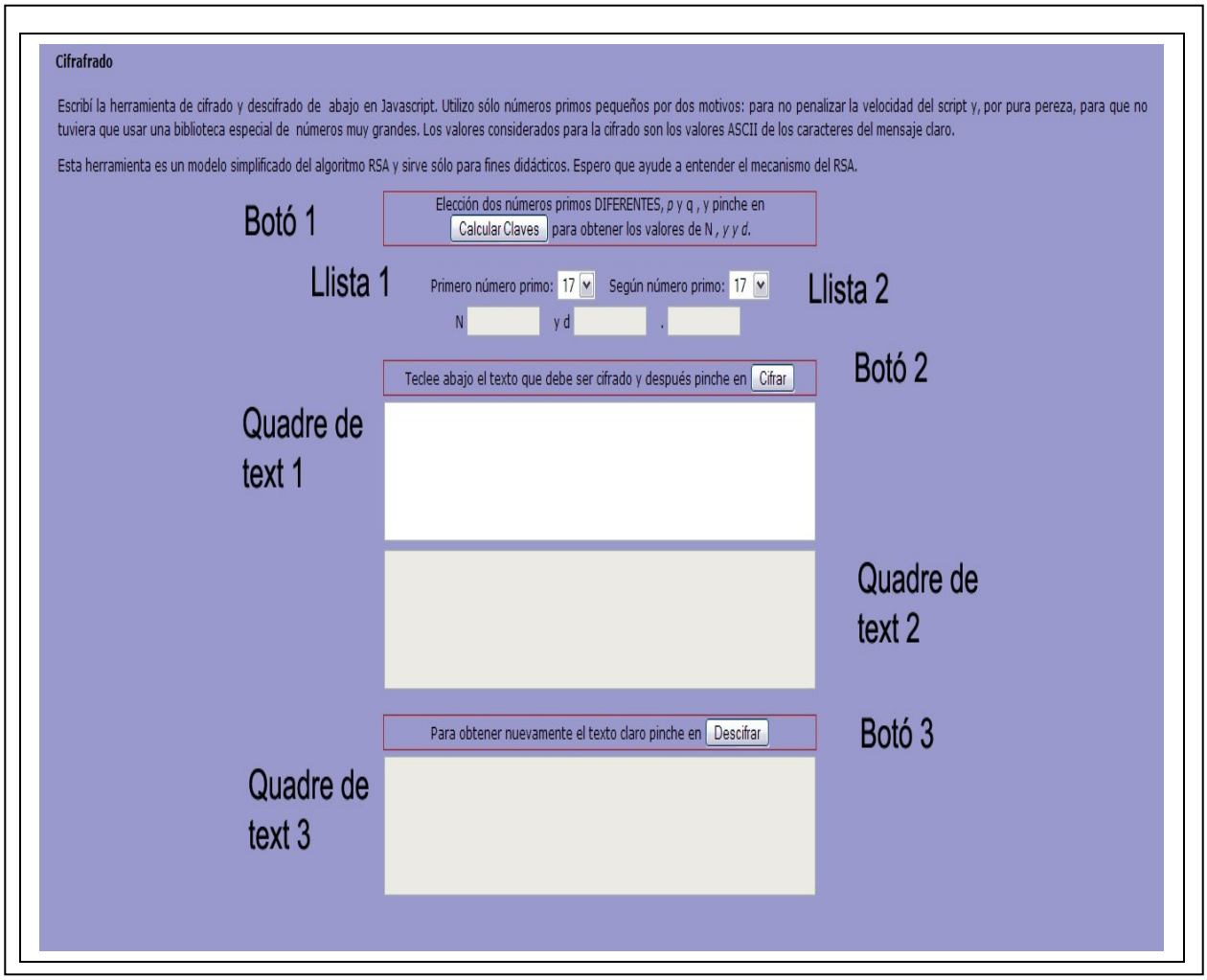

Figura nº 6.1: Aspecte de la creació original i els seus elements

Com podem veure la simulació original disposa dels elements següents:

- **Botó 1**: aquest botó serveix per calcular les diferents claus un cop hem escollit els dos nombres primers de les llistes.

- **Botó 2**: la funció que té aquest botó és al de xifrar el missatge un cop s'ha escrit en el quadre de text 1 algun missatge i s'han escollit els dos nombres primers de les llistes corresponents.
- **Botó 3**: aquest botó serveix per comprovar que el que hem xifrat es correspongui amb la realitat
- **Llista 1**: aquesta llista conté 5 nombres primers prefixats per agilitar la velocitat de funcionament del codi.
- **Llista 2**: aquesta llista conté els 5 mateixos nombres primers prefixats i representa l'altre nombre primer que cal escollir per poder fer servir el criptosistema RSA.
- **Quadre de text 1**: en aquest espai hem de col·locar el missatge que volem codificar.
- **Quadre de text 2**: en aquest requadre ens sortirà la codificació en nombres del nostre missatge.
- **Quadre de text 3**: aquest quadre ens mostra el text desxifrat per comprovar que no hi hagi cap error en l'aplicació de l'algorisme.

Un cop hem vist l'aparença inicial del simulador el que farem és la seva alteració i adaptació a les nostres necessitats. Per això s'ha utilitzat un programa de creacions de pàgines web per poder accedir al codi font d'aquesta pàgina web (*recordar que el JavaScript funciona com a llenguatge interpretat i no necessita ser compilat*). Un cop es disposa del **codi font** s'ha alterat la seva presentació i s'han canviat aspectes com la interfície, llenguatge de la presentació, s'hi ha col·locat el símbol de la llicència de *creative commons* tal i com correspon, entre altres alteracions totes elles de caire visual.

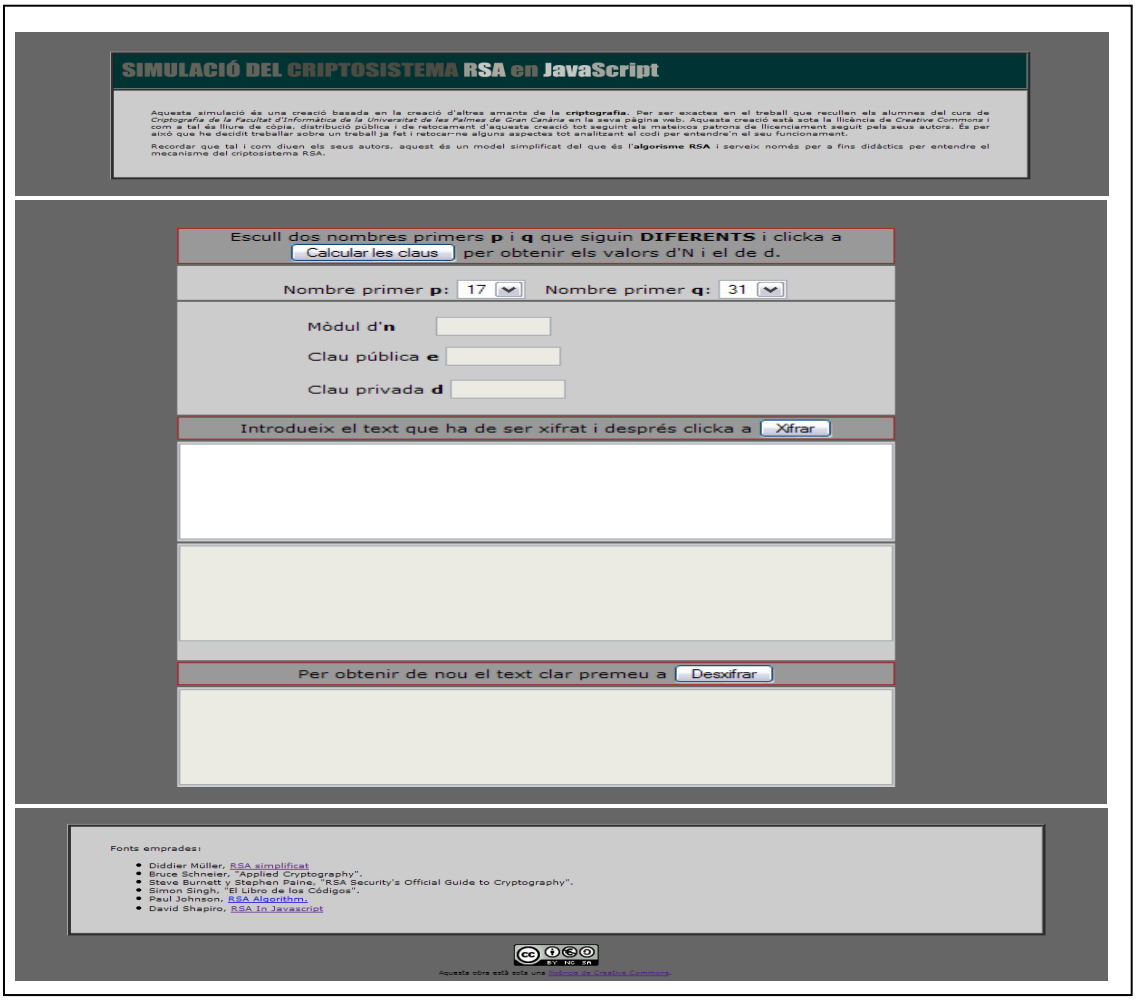

Figura nº 6.2: Aspecte de la creació retocada

6.3. Apropament al codi font

Ara que hem vist la part externa de la creació passarem a interpretar la part interna d'aquesta, és a dir, el **codi font** o llenguatge que s'ha fet servir. Aquest anàlisi no serà de tota la pàgina en si, ja que també hi ha llenguatge específic de creació de pàgines web: **html**, sinó que s'anirà directament a analitzar el codi que fa referència a la simulació en **JavaScript**. La visualització d'aquests codis en JavaScript s'han fet amb el programa *Macromedia Dreamweaver 8* ja que disposa d'un bon sistema de visualitzador de codi que marca amb diferents colors les diferents instruccions que composen el codi tot seguint com a criteri

el tipus de funció que té cada cosa. Així passem a analitzar pas per pas el codi que composa la simulació:

```
<body><script language="javascript"><!--

function GCD(e,PHI) {
  if (e > PHI) {
    while (e%PHI != 0) {
      a = e%PHI;
      e = PHI;
      PHI = a;
    }
    great = PHI;
  } else {
    while (PHI%e != 0) {
      a = PHI%e;
      PHI = e;
      e = a;
    }
    great = e;
  }
  return great;
}
```

Figura nº 6.3: Scripts del simulador

```

function tofindE(PHI,P,Q) {
  great = 0;
  e = 2;
  while (great != 1) {
    e = e + 1;
    great = GCD(e,PHI);
    PHI = (P - 1) * (Q - 1);
  }
  return e;
}

```

Figura nº 6.4: Scripts del simulador

```
40  function extend(E,PHI) {
41    u1 = 1;
42    u2 = 0;
43    u3 = PHI;
44    v1 = 0;
45    v2 = 1;
46    v3 = E;
47
48    while (v3 != 0) {
49      q = Math.floor(u3/v3);
50      t1 = u1 - q * v1;
51      t2 = u2 - q * v2;
52      t3 = u3 - q * v3;
53
54      u1 = v1;
55      u2 = v2;
56      u3 = v3;
57
58      v1 = t1;
59      v2 = t2;
60      v3 = t3;
61      z = 1;
62    }
63    uu = u1;
64    vv = u2;
65
66    if (vv < 0) {
67     inverse = vv + PHI;
68    } else {
69      inverse = vv;
70    }
71    return inverse;
72  }
```

Figura nº 6.5: Scripts del simulador

```
74  function chaves() {
75    var P = (document.form1.primo1.options[document.form1.primo1.selectedIndex].value);
76    var Q = (document.form1.primo2.options[document.form1.primo2.selectedIndex].value);
77    PHI = (P - 1) * (Q - 1);
78    E = tofindE(PHI,P,Q);
79    document.form1.modulo.value = P * Q;
80    document.form1.publ.value = E;
81    document.form1.priv.value = extend(E,PHI);
82  }
```

Figura nº 6.6: Scripts del simulador

-

```
84  function char2ascii (c)
85  {
86    // restrict input to a single character
87    c = c . charAt (0);
88
89    // loop through all possible ASCII values
90    for (var i = 0; i < 256; ++ i)
91    {
92      // convert i into a 2-digit hex string
93      var h = i . toString (16);
94      if (h . length == 1)
95        h = "0" + h;
96
97      // insert a % character into the string
98      h = "%" + h;
99
100     // determine the character represented by the escape code
101     h = unescape (h);
102
103     // if the characters match, we've found the ASCII value
104     if (h == c)
105       break;
106   }
107   return i;
```

Figura nº 6.7: Scripts del simulador

```
110 function ascii2char (v)
111 {
112   var h = v . toString (16);
113   if (h . length == 1)
114     h = "0" + h;
115
116   // insert a % character into the string
117   h = "%" + h;
118
119   // determine the character represented by the escape code
120   return unescape (h);
121 }
```

Figura nº 6.8: Scripts del simulador

```
123  function cifra() {
124    var msg = document.form1.texto.value;
125    var E = document.form1.publ.value;
126    var N = document.form1.modulo.value;
127    var cifra = "";
128    var quebra = 0;
129
130    for(var i = 0; i < msg.length; i++) {
131      var M = char2ascii(msg.charAt(i));
132      if (M > - 1) {
133        if ( E % 2 == 0) {
134          G = 1;
135          for ( var ii = 1; ii <= E/2; ii++) {
136            F = (M*M) % N;
137            G = (F*G) % N;
138          }
139        } else {
140          G = M;
141          for ( ii = 1; ii <= E/2; ii++) {
142            F = (M*M) % N;
143            G = (F*G) % N;
144          }
145        }
146        xc = "" + G;
147        while(xc.length < 3) xc = '0' + xc;
148        cifra = cifra + xc;
149        quebra = quebra + 1;
150        if(quebra % 20 == 0) cifra = cifra + " ";
151      }
152    }
153    document.form1.cifrado.value = cifra;
154  }
```

Figura nº 6.9: Scripts del simulador

```
156  function decifra() {
157    var cfr = document.form1.cifrado.value;
158    var N = document.form1.modulo.value;
159    var D = document.form1.priv.value;
160    var txt = "";
161
162    var i = cfr.indexOf(" ");
163    while(i > -1){
164      cfr = cfr.replace(" ", "");
165      i = cfr.indexOf(" ");
166    }
167
168    for(var i = 0; i < cfr.length; i = i+3) {
169      var M = cfr.charAt(i) + cfr.charAt(i+1) + cfr.charAt(i+2);
170      if ( D % 2 == 0) {
171        G = 1;
172        for ( var ii = 1; ii <= D/2; ii++) {
173          F = (M*M) % N;
174          G = (F*G) % N;
175        }
176      } else {
177        G = M;
178        for ( ii = 1; ii <= D/2; ii++) {
179          F = (M*M) % N;
180          G = (F*G) % N;
181        }
182      }
183      txt = txt + ascii2char(G);
184    }
185    document.form1.decifrado.value = txt;
186  }
187
188  --></script>
```

Figura nº 6.10: Scripts del simulador

6.4. Dibuix de l'algorisme

6.5. Proves

A continuació el seguit de proves per verificar el bon funcionament del programa. Aquestes proves es fan per veure que no hi hagi cap error en el disseny de l'algorisme. El nostre paper a l'hora d'analitzar aquesta creació seria el després d'haver analitzat i entès el codi mirar que no se'ns hagi passat res per alt tot sometent la simulació a diferents exercicis per veure si hi ha alguna irregularitat que ens pugui guiar cap a algun error o en el cas contrari, no se'ns presenti cap error i signifiqui que, de moment, sembla ser correcte. Sembla ser, ja que un programa s'ha de mantenir sempre actualitzat al dia ja que ens podem trobar en casos molt puntuals que donin errors. En el cas que se'ns presenta l'error observat està a l'hora d'escollir els dos nombres primers iguals. Això fa que se'ns alteri el missatge que volem codificar. Els propis creadors ja eren conscients d'aquests problema i per això remarquen el que han de ser diferents ja que sinó no funcionarà bé. És per això que el terme "DIFERENTS" està ressaltat. Mantenint la solució que ells van utilitzar, les altres proves evidencien un bon funcionament de l'algorisme que fan servir. Per tant, no serà retocat.

- **Prova de minúscules/nombres:**

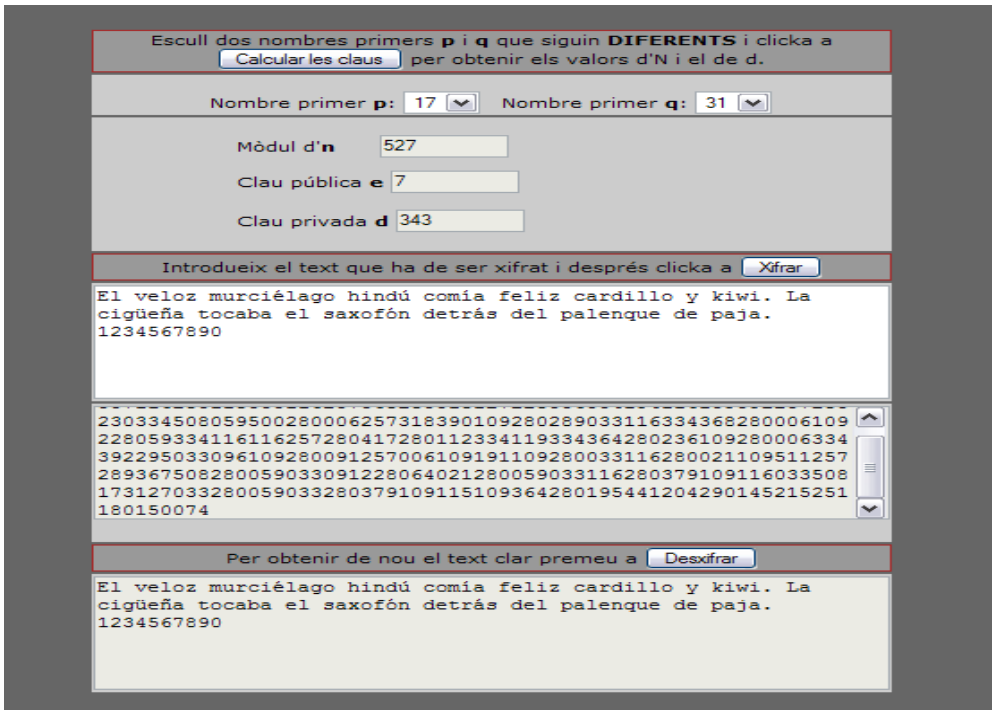

Figura nº 6.11: Prova amb minúscules i nombres

- **Prova de majúscules:**

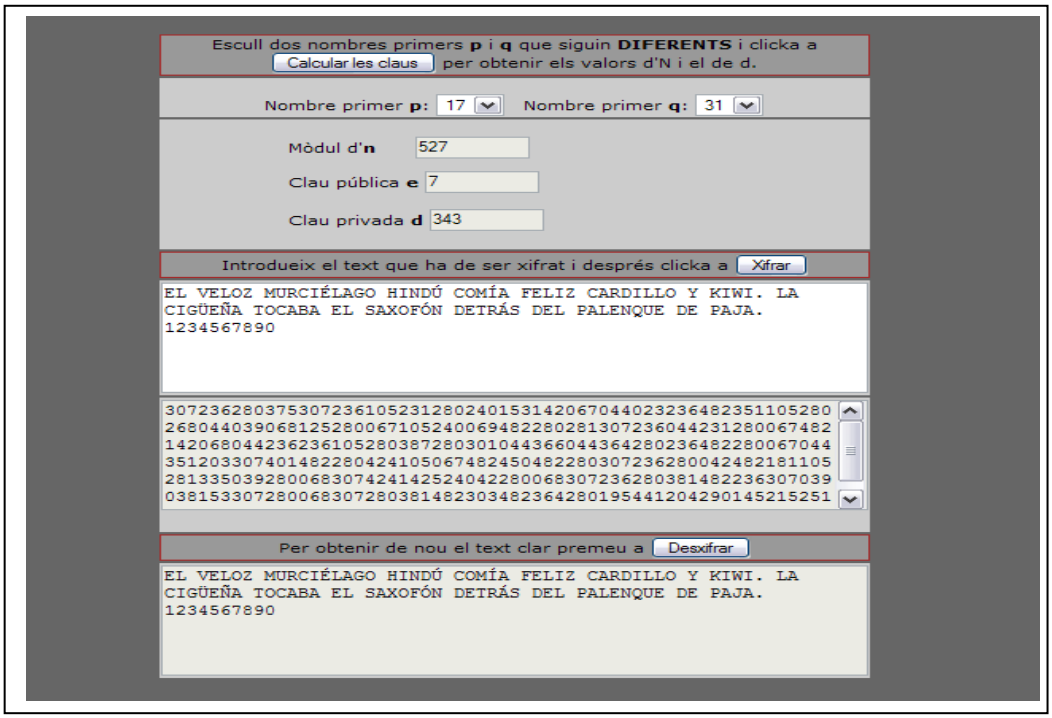

Figura nº 6.12: Prova amb majúscules i nombres

- **Prova d'accents:**

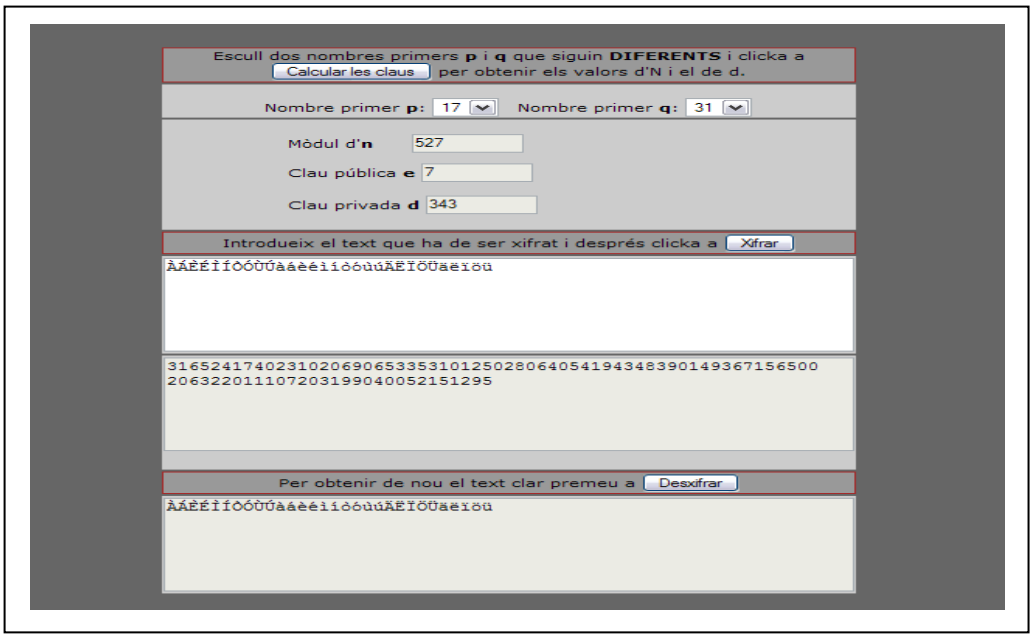

Figura nº 6.13: Prova d'accents

APARTAT FINAL

VII MATERIALS, METODOLOGIA, CONCLUSIONS

7.1. Materials i metodologia utilitzada

Els materials que s'han utilitzat per la realització del Treball de Recerca han estat:

- Ordinadors
- Material d'escriptura: llapis, bolígraf, etc.
- Programes informàtics com:
    - Microsoft Word
    - Adobe Photoshop CS
    - Macromedia Dreamweaver 8
    - Internet Explorer 7.0
    - Mozilla Firefox 1.5
    - Open Office
- Suports de gravació:
    - Cds
    - Pen Drive

7.2. Resultats i conclusions

Els resultats obtinguts amb aquest treball són principalment l'obtenció de la memòria que representa el treball i l'adquisició d'uns coneixements abans no posseïts sobre el tema de la criptografia. És més el que s'ha aconseguit és realment sorprenent ja que la visió que puc tenir ara un cop acabat el treball respecte a la visió que tenia en un principi sobre el tema de la criptografia s'ha vist augmentat enormement. Per posar un exemple no gaire erroni: el que creia en un principi sobre la criptografia no devia arribar ni a un 1 % del que era en realitat. Per això dic que el resultat en el sentit d'aprenentatge ha estat molt satisfactori. Alhora ha estat la primera vegada que prenia contacte directe amb llenguatges de programació i sobre creació de programes. Aquest contacte ha fet que intentant resoldre el problema d'aprendre un llenguatge de programació

a la perfecció fes que es replantegés la part pràctica del treball de recerca en un altre sentit: en comptes de crear una aplicació, s'analitzaria una ja d'existent tot intentant entendre el funcionament del llenguatge de programació utilitzat sense necessitat d'haver d'aprendre a la perfecció el llenguatge fent que es tingui una vertadera sensació d'aproximació al llenguatge de programació. A més aquesta interpretació no precisa de gaires coneixements sobre aquest perquè amb uns coneixements bàsics de programació i coneixement el funcionament del criptosistema RSA s'ha pogut intuir amb més o menys dificultat el funcionament de cadascuna de les instruccions del programa. Així podem dir doncs que els coneixements adquirits en la part teòrica han donat com a resultat el poder interpretar un llenguatge de programació tot sabent com funciona el que s'està analitzant.

Pel que respecta a les conclusions podem dir que:

- La **criptologia** ha estat una ciència que ha estat sempre en l'anonimat i pel que sembla encara ho seguirà estant encara que jugui un paper molt important en la vida quotidiana.
- L'**eficàcia del Criptosistema RSA** durarà el que es tardi en dessarrollar una forma viable d'informàtica quàntica i ordinadors quàntics o que es descobreixi la forma de factoritzar nombres primers grans de forma ràpida.
- L'**evolució de la criptografia** ve donada efectivament per les guerres i per la necessitat de protecció de la informació dels bàndols respectius. En l'actualitat representada per la constant guerra que representen les empreses: una guerra econòmica per veure qui surt endavant i qui no.
- L'**inmersió de la ciència i la tecnologia** en el camp de la criptografia ha fet que aquest experimenti una evolució molt gran fent que en l'actualitat apareguin nous models criptogràfics amb més freqüència que en l'antiguitat.
- Sense la criptografia no podríem gaudir d'una **vida cibernètica** amb la seguretat que nosaltres voldríem cosa que faria impossible coses com el negoci per Internet o altres afers que tinguin a veure amb estaments oficials: declaració de la Renta per Internet, etc.
- Podríem dividir la **història de la criptografia** en dues etapes: una més manual i simple i una altre més mecanitzada i complexa. Parlem doncs, d'una matematització del món de la criptografia que ha fet que es converteixi en un camp molt exclusiu.

- Al tractar d'aquesta segona criptografia més matematitzada ens trobem amb unes **limitacions de coneixements** ja que la teoria o coneixements adquirits fins ara no són del tot suficients per poder entendre la forma de treballar dels complexes algorismes de la criptografia actual.
- Finalment, que encara que no es complís el propòsit inicial de creació d'una aplicació partint de zero, és igual d'educatiu **analitzar un sistema** ja creat ja que permet posar en pràctica altres àmbits com el de l'associació. Potser no s'aplica tanta creativitat matemàtica, però si reconeixement de la teoria reunida en la primera part aplicada en la part pràctica d'aquest treball.

VIII RESSENYA BIBLIOGRÀFICA

- **Llibres:**

BERLINSKI, D. (2006). *Ascenso infinito.* Barcelona: Debate
CABALLERO, P. (2002). *Introducción a la criptografia..* Madrid: Ra-ma.
DURÁN, R.; HERNÁNDEZ, L.; MUÑOZ, J. (2005) *El criptosistema RSA.* Madrid: RA-MA.
GALENDE, J. (1995). *Criptografia, historia de la escritura cifrada*. Madrid: Editorial Complutense.
JUHER, D. (2001). *Introducció a la criptografia*. Girona: Publicacions docents.
JOYANES, L. (1996). *Fundamentos de la programación: Algoritmos y estructuras*. Madrid: Mc
Graw-Hill.
LEVY, S. (2002). *Cripto.* Madrid: Alianza Editorial.
SINGH, S. (2000). *Los códigos secretos.* Barcelona: Debate.

- **Pàgines webs:**

DAVE, O. (2007) *RSA en JavaScript* Visitat el http://www.ohdave.com/rsa/
ELLISON, C.(2004) *Cronologia* Vist el 23 de Desembre de 2006 a http://world.std.com/%7Ecme/html/timeline.html
MANUEL, G. (2007) *Blog sobre criptografía*. Vist el 14 d'Octubre de 2006 a http://www.kriptopolis.com/
MONJAS, M. (1998). *Servicios de seguridad*. Vist el 24 de Septembre de 2006 a http://www.dat.etsit.upm.es/~mmonjas/cripto/
MÜLLER, D. (2006) *Apunts criptografía* Vist el 23 de Septembre de 2006 http://serdis.dis.ulpgc.es/index.php
RINCONQUEVEDO (2002). *Introducció a la criptografia*. Vist el 5 d'Agost de 2006 http://rinconquevedo.iespana.es/rinconquevedo/criptografia/
RODRÍGUEZ, A. (2006). *Mètodes classics Vist el 6 d'Agost de 2006 a* http://www.epsilones.com/paginas/t-historias1.html#historias-criptografia
SCHNEIER, B. (2003). *Criptografia - Historia*. Vist el 5 d'Agost de 2006 a http://www.redsegura.com/Temas/CRhistoria.html
WIKIPEDIA (2005) *Criptografia.* Vist el 20 d'Agost de 2006 a http://es.wikipedia.org/wiki/Criptograf%C3%ADa

- **Altres:**

(2001) *Introducción a la programación en C*. Barcelona: Edicions UPC
(2004) *Curso de criptografia infromática*. Internet
(2004) *Fonaments de programació.* Barcelona: UOC

ANNEXES

XIX CONCEPTES

9.1. Notacions i apunts matemàtiques

- Definicions matemàtiques:

  - **N** representa el conjunt dels nombres naturals, és a dir,
    N = { 0,1,2,...}

  - **Z** representa el conjunt dels nombres enters, això vol dir,
    Z = {..., -2, -1, 0, 1, 2,...}

  - **Q** representa el conjunt dels nombres racionals, per exemple,

    $$\mathbf{Q} = \{ a/b : a, b \in \mathbf{Z}, b \neq 0 \}$$

  - R representa el conjunt dels nombres reals

  - P representa el conjunt dels nombres primers

  - Un número enter **p** s'anomena *primer* si no té divisors propis, és a dir, si només es divisible per ell mateix i per 1. Un nombre enter s'anomena *compost* si no és primer.

  - Dos nombres enters a i b són considerats primers entre si o coprimers si verifiquen que **mcd** (a,b) = 1

  - El càlcul del **mcd** (*màxim comú divisor*) de dos nombres enters es pot aconseguir multiplicant els seus factors comuns elevats al menor exponent.

  - El càlcul del **mcm** (*mínim comú múltiple*) de dos nombres enters es pot aconseguir multiplicant els seus factors comuns i els no comuns amb l'exponent més gran.

  - Per qualsevol parell de nombres enters, a i b, es verifica que

    $$a \cdot b = \text{mcd}\,(a,b) \cdot \text{mcm}\,(a,b)$$

  - Donats a, b $\in$ **Z**, es diu que a i b són congruents mòdul n si la seva diferència és un múltiple d'n, és a dir, n | (a-b). Dit d'una altre manera, si a i b tenen el mateix residu al ser dividits per n. L'enter n és el mòdul de la congruència.

- Simbologia:

  ≤ menor o igual que

  ≥ més gran o igual que

  ≠ no igual a

  { } delimitador dels elements

  ∑ suma d'un conjunt de nombres o elements

  ∀ per tot

  = igual

  < més petit que

  > més gran que

  ≡ idèntic a

  ∃ existeix

  ∄ no existeix

  ∋ conté com a element

  ∈ és membre de

  ⋮ fins a

  ⋯ fins a

9.2. Glossari específic

Algorisme: conjunt d'operacions elementals que s'han d'efectuar per obtenir un resultat esperat.

Algorisme RSA: el més popular sistema criptogràfic de clau pública, inventat per Rivest, Shamir i Adleman l'any 1977.

ASCII: (American Standard Code Information Interchange) Codi Patró Americà per l'intercanvi d'informació que tradueix els nombres dels caràcters d'un alfabet a nombres. Per exemple, la lletra "A" es traduida per 65.

Atac: intent de criptoanàlisis.

Clau: el component d'un criptosistema que determina com serà mesclat el missatge. Aplicant una clau al missatge original s'obté el text xifrat; la mateixa clau el durà a la seva forma original. En el cas específic dels sistemes de clau pública, la seva clau secreta associada).

Clau privada: en un sistema de clau pública la clau privada és la component de la parella de claus que s'ha de mantenir en secret: és l'única clau que pot desxifrar el missatge xifrat amb la clau pública corresponent i firmar els missatges per contrastar que va ser el propietari de la clau pública qui el va enviar.

Clau pública: el component de la parella de claus que permet a altres persones enviar missatges al seu propietari. També s'utilitza per verificar les firmes digitals. Pot ser distribuïda lliurement sense comprometre la seguretat del sistema.

Clau simètrica: utilitzada en la criptografia convencional, s'empra la mateixa clau per fer que el remitent xifri el missatge i per que el receptor el desxifri.

Clau asimètrica: utilitzada en la criptografia de clau pública, s'empra una clau pública per encriptar i una clau privada (diferent) per desencriptar el missatge que ens han enviat.

Codi: Sistema de símbols (paraules, noms, símbols, etc.) que substitueixen paraules senceres.

Criptoanàlisis: desxifrat de codis, l'art de tornar el text xifrat a la seva forma original sense disposar de la clau de xifrat.

Criptografia de clau pública: revolucionari sistema descobert per Diffie i Hellmann l'any 1975 que substituïa les claus simètriques per una parella de claus.

Criptologia: és l'estudi i les matemàtiques dels codis secrets i les xifres.

Criptosistema: els mitjans per xifrar les dades i realitzar altres funcions criptogràfiques; freqüentment s'utilitza com a sinònim de l'algorisme que realitza el xifrat.

Diccionari: llista de paraules i expressions més utilitzades que serveixen per buscar estructures xifrades prefixades.

Dipòsit de claus: una porta trasera introduïda intencionadament en un criptosistema que permet a les autoritats desxifrar ràpidament el missatge, teòricament sense comprometre la seva seguretat de ninguna altre manera.

Encriptació: l'acció de xifrar la informació per impedir que el missatge interceptat pugui ser llegit.

Estàndar de Xifrat de Dades (Data Encrypted Standard – DES): un criptosistema creat per IBM basat en el Lucifer. Encara que inicialment va ser molt qüestionat pels crítics, aquest sistema convencional va demostrar ser fiable; el seu únic problema: la insuficiència de la longitud de la seva clau.

Factorització: l'acte matemàtic d'agafar un nombre generat per multiplicació de dos nombres més petits i trobar aquests nombres originals. Aquesta funció unidireccional és la base de l'algorisme RSA.

Firma digital: dades criptogràfiques generades matemàticament que relacionen el missatge amb el seu autor.

Força bruta: és un atac que consisteix en provar totes les claus possibles fins arribar a trobar la correcta. No és un bon mètode d'accès pel temps que pot implicar: dies, mesos, anys o fins i tot que mai s'arribi a aconseguir.

Freqüència: percentatge d'aparició d'una lletra o paraula en una determinada llengua. Calcular la freqüència d'aparició és una de les primeres etapes de la desencriptació.

Funció unidireccional: una operació matemàtica fàcil de calcular però molt difícil d'invertir.

Javascript : llenguatge d'instruccions multiplataforma que es interpretat per aplicacions client, normalment un navegador. Creat per Netscape, el codi JavaScript s'addereix directament en una pàgina HTML. JavaScript no és un llenguatge de programació pròpiament dit, sinó un llenguatge escrit orientat a millorar aspectes relacionats amb pàgines web. La seva sintaxi es semblant a la dels llenguatges Java i C.

Longitud de clau: quan més llarga és una clau, més difícil serà de desxifrar pel procediment de "força bruta". El conjunt de totes les claus s'anomena espai de claus.

Lucifer: criptosistema convencional dissenyat per Horst Feistel d'IBM a principis dels setanta

Text xifrat: l'estat del missatge un cop ha estat xifrat.

Text original: el missatge abans de ser xifrat.

Xifra: també coneguda per algorisme criptogràfic, és la funció matemàtica utilitzada per xifrar i desxifrar els missatges.

9.3. Línia del temps

- **3500 a.C. – 1 a.C.:**

  - 3500 - Els sumeris desenvolupen l'escriptura cuneiforme I els egipcis l'escriptura jeroglífica
  - 1500 – Els fenicis desenvolupen un alfabet
  - 600-500 – Els erudits hebreus fan ús de substitucions simples monoalfabètiques com la xifra Atbash
  - 400 – Els espartans utilitzen la scytala.
  - 400 – Herodotus descriu l'ús de l'esteganografia en les guerres entre grecs i perses: un tatuatge en un cap rapat.
  - 100-1 CE – Notables xifres romanes com la xifra del Cèsar.

- **1 - 1799:**

  - 1000 – L'anàlisi de freqüència que duu a tècniques per trencar xidres de substitució monoalfabètiques semblen haver estat inspirades en l'anàlisi textual del Corà.
  - 1450 – Els xinesos creen blocs de fusta gravats amb lletres que representa un inici de la impressió de caràcters.
  - 1450-1520 – El manuscript Voynich, un exemple d'una possible encriptació pictórica, és escrit.
  - 1466 - Leone Battista Alberti crea la xifra polialfabètica, també coneguda primerament com la xifra mecànica.
  - 1518 – Llibre de Johannes Trithemiu sobre criptologia.
  - 1553 - Belaso inventa la xifra Vigenère
  - 1585 – Llibre de Vigenère sobre diferents xifres
  - 1586 – Criptoanàlisis utilitzat per l'espia Sir Francis Walsingham per implicar a la Reina Maria d'Escòcia en el complot de Babington d'assassinar la Reina Elizabeth I d'Anglaterra. La Reina Maria va ser executada.
  - c. 1645 – El Mercuri de Wilkins, llibre anglès sobre criptologia.
  - 1793 – Claude Chappe estableix la primera línea llarga de telègraf.
  - 1795 - Thomas Jefferson inventa la xifra de disc de Jefferson, reinventada de nou més de 100 anys més tard per Etienne Bazeries.

- **1800 – 1900:**

  - 1809-14 Treball de George Scovell sobre la xifra napoleònica Durant la Guerra Peninsular.
  - 1831 - Joseph Henry proposa i construeix un telègraf electric.
  - 1835 - Samuel Morse desenvolupa el codi morse.
  - 1854 - Wheatstone inventa la xifra Playfair.
  - 1854 – Mètode de Babbage per trencar les xifres polialfabèriques (publicat per Kasiski l'any 1863)
  - 1855 – Pel bàndol angles a la Guerra Crimea, Charles Babbage trenca la xifra autoclau de Vigenère (la "xifra indesxifrable" de l'època) així com la més feble xifra pel que avui s'entén a la xifra Vigenère. A causa del secretisme també va ser descobert i atribuït al prussià Friedrich Kasiski.
  - 1883 – El llibre "La Criptografia militar" d'Auguste Kerchoffs va ser publicat, contenent les seves conegudes lleis sobre la criptografia.
  - 1885 – Xifra Beale publicada.
  - 1894 – L'assumpte de Dreyfus a França implica l'ús de la criptografia.

- **1900 – 1949:**

  - 1915 - William Friedman aplica l'estadística en el criptoanàlisi (index de coincidència, etc.)
  - 1917 - Gilbert Vernam crea la seva primera implementació pràctica d'una xifra teletip. Posteriorment, amb Joseph Mauborgne va ser coneguda per la llibreta d'un sol ús.
  - 1917 – Telegrama Zimmermann interceptat i desxifrat, avançant l'entrada dels Estats Units a la Primera Guerra Mundial.
  - 1919- L'Oficina Estrangera Alemana de Weimar adopta l'ús manual de les llibretes d'un sol ús en certs tipus de missatges.
  - 1919 - Edward Hebern inventa i patenta el primer disseny de màquina de rotors – Damm, Scherbius i Koch també patenten la seva Enigma el mateix any.
  - 1921 – Conferència Naval a Washington – l'equip de negociació dels Estats Units són ajudats per l'equip de desxiframent dels telegrames japonesos de caire diplomàtic.
  - 1924 - MI8 (Herbert Yardley) aconsegueixen trencar diversos tipus d'intercepcions per ajudar als Estats Units durant les negociacions a la Conferència Naval de Washington.
  - 1932 – primer cop que es va trencar l'Enigma alemanya per Marian Rejewski a Polònia.

- 1929 – El secretari d'Estat nord-americà Henry L. Stimson tanca el departament estatal de criptoanàlisi també anomenada la cambra negra, dient "Senyors, no llegissin el correu dels altres"
- 1931 – "La Cambra Negra Americana" per Herbert O. Yadley és publicat, revelant moltes coses sobre la criptografia Americana.
- 1940 – s'aconsegueix trencar la màquina japonesa d'encriptació, PURPLE, per l'equip SIS.
- Desembre 7, 1941 – La base naval dels Estats Units a Pearl Harbour es sorpresa per l'atac japonès, encara que els Estats Units havien desxifrat els codis japonesos, entren a la Segona Guerra Mundial.
- June 1942 – Batalla de Midway on els Estats Units semblen obtenir la victòria sobre els japonesos.
- 1943 – L'almirant Yamamoto, creador de l'atac de Pearl Harbour, és assassinat per les forces americanes que coneixen les seves implicacions a partir dels missatges descodificats.
- 1943 - Max Newman, Wynn-Williams, i el seu equip (incloent Alan Turing) a l'Escola Guvernamental de Codis i Xifres (Estació X), Bletchley Park, a Anglaterra completen el "Cap Robinson". Aquesta és una màquina especialitzada en trencament de xifres.
- Desembre de 1943 – El Colossus va ser construït, per Thomas Flowers a l'Oficina Postal d'Investigacions Laboratòries a Londres, per crackejar la xifra alemanya Lorenz. Colossus va ser utilitzat a Bletchley Park Durant la Segona Guerra Mundial – com a successor de l'antiga màquina Robinson. També va ser construïts deu més, desafortunadament van ser destruïdes immediatament després d'acabar la seva feina – eren tan avançades que no hi havia d'haver cap possibilitat de caure en les mans equivocades.
- 1944 – Es va fer la patent pel codi de la màquina SIGABA utilitzada pels Estats Units a la Segona Guerra Mundial. Mantinguda en secret fins el 2001.
- 1946 – Primer trencament de les comunicacions dels espies russos per part de la màquina VENONA.
- 1948 - Claude Shannon escriu les bases matemàtiques de la teoria de la informació.
- 1949 – "Teoria dels Sistemes de Comunicació Secrets" de Shannon va ser publicat al Bell Labs Technical Journal.

- **1950 – 1999:**

  - 1951 – Va ser fundada l'Agència Nacional de Seguretat dels Estats Units (NSA).
  - 1957 – Primera comanda de producció de sistemes d'encriptació electròniques.
  - 1964 - David Kahn publica el llibre "The Codebreakers"
  - Agost de 1964 – L'incident del Golf de Tonkin porta als Estats Units a la guerra del Vietnam, possiblement degut a una mala interpretació dels senyals per part de la NSA.
  - 1968 - John Anthony Walker ven informació a l'embaixada de la Unió Soviètica a Washington sobre la xifra KL-7. Els seus actes van durar fins el 1985.
  - 1969 – Els primers amfitrions d'ARPANET, l'antecessor d'Onternet, estan conectats.
  - 1974 - Horst Feistel crea la red Feistel d'encriptació per blocs.
  - 1976 – el DES va ser publicat com un Estandard Federal de Processament d'Informació (FIPS) als Estats Units.
  - 1976 - Diffie and Hellman publiquen "Noves Direccions en la Criptografia"
  - 1977- Criptosistema RSA de clau pública creat.
  - 1981 - Richard Feynman proposa els ordinadors quàntics. L'aplicació principal que tenia en ment era la simulació de sistemes quàntics, però també va mencionar la possibilitat de resoldre altres problemes.
  - 1985 - Walker és descobert, quedant el codi KL-7 for a de servei.
  - 1986 – Després d'augmentar el nombre d'intents per entrar als ordinadors del govern i de corporacions, el Congrés dels Estats Units crea una llei que defineix aquest tipus d'actes com a delictes. La llei, però, no afecta als menors.
  - 1988 – Primer xip òptic desenvolupat, utilitza llum en comptes d'electricitat per incrementar la velocitat de processament.
  - 1989 - Tim Berners-Lee i Robert Cailliau construeixen el prototip de sistema que es va convertir en la World Wide Web a CERN.
  - 1991 - Phil Zimmermann publica el programa PGP d'encriptació de clau pública juntament amb el seu codi font, que ràpidament s'estenen per Internet.
  - 1994 – El protocol Secure Sockets Layer (SSL) es utilitzat per Netscape.
  - 1994 - Peter Shor inventa un algorisme que permet als ordinadors quàntics determinar la llargària de factorització dels nombres ràpidament. Aquest és el primer problema interessant pel qual els

ordinadors quàntics prometen un significant avenç, i per això es va generar un gran interès en ordinadors quàntics.

- 1994 – Crackers russos traspassen 10 milions de dollars del Citibank I el transfereixen a diferents contes bancàries d'arreu del món.
- 1995 - NSA publica el algorisme hash SHA1 com una part del seu estàndard de firma digital.
- Juliol de 1997 - OpenPGP surt a la llum.
- Octubre de 1999 - DeCSS, un programa d'ordinador capaç de desxifrar el contingut d'un DVD, és publicat a Internet.

- **2000 i endavant:**

  - 14 de Gener del 2000 – El govern dels Estats Units anuncia que les restriccions en l'àmbit de l'exportació de criptografia han disminuït. Això fa que moltes empreses americanes parin de fer la dura feina de crear versions per Amèrica i pel mercat Internacional totalment diferents del seu software.
  - Març de 2000 – El president Clinton diu que no utilitza el correu e-mail per comunicar-se amb la seva filla, Chelsea Clinton, a la universitat perquè no creia que el mitjà fos segur.
  - 6, de Setembre de 2000 - RSA Security Inc. mostren el seu algorisme RSA al domini públic, uns dies després que la seva Patent 44058829 s'acabés. Juntament amb el canvi de política en les restriccions d'exportació criptogràfica farà que s'acabin una de les últimes barreres per la distribució mundial de gran part dels nous programes basats en sistemes criptogràfics.
  - 2001 – L'algorisme belga Rijndael va ser escollit pels Estats Units com a Estandard d'Encriptació Avançada (AES) després de cinc anys de busca pública duta a terme per Institut Nacional d'Estandards i Tecnologia (NIST) dels Estats Units.
  - 2001 - Scott Fluhrer, Itsik Mantin i Adi Shamir publiquen un atac a la capa de seguretat que hi ha a les xarxes WiFi.
  - 11 de Setembre de 2001 – Els Estats Units culpen dels atacs terroristes a la manca de comunicacions segures.
  - 2005 – atacs potencials sobre el SHA1 demostrats.
  - 2005 – agents de l'FBI demostren la seva facilitat per evadir la seguretat WEP de xarxes WiFi amb programes a l'abast de tothom.