Il Dominio della Perfetta Segretezza

Claudio Cappelli
Luca Amodeo
Lorenza De Lellls

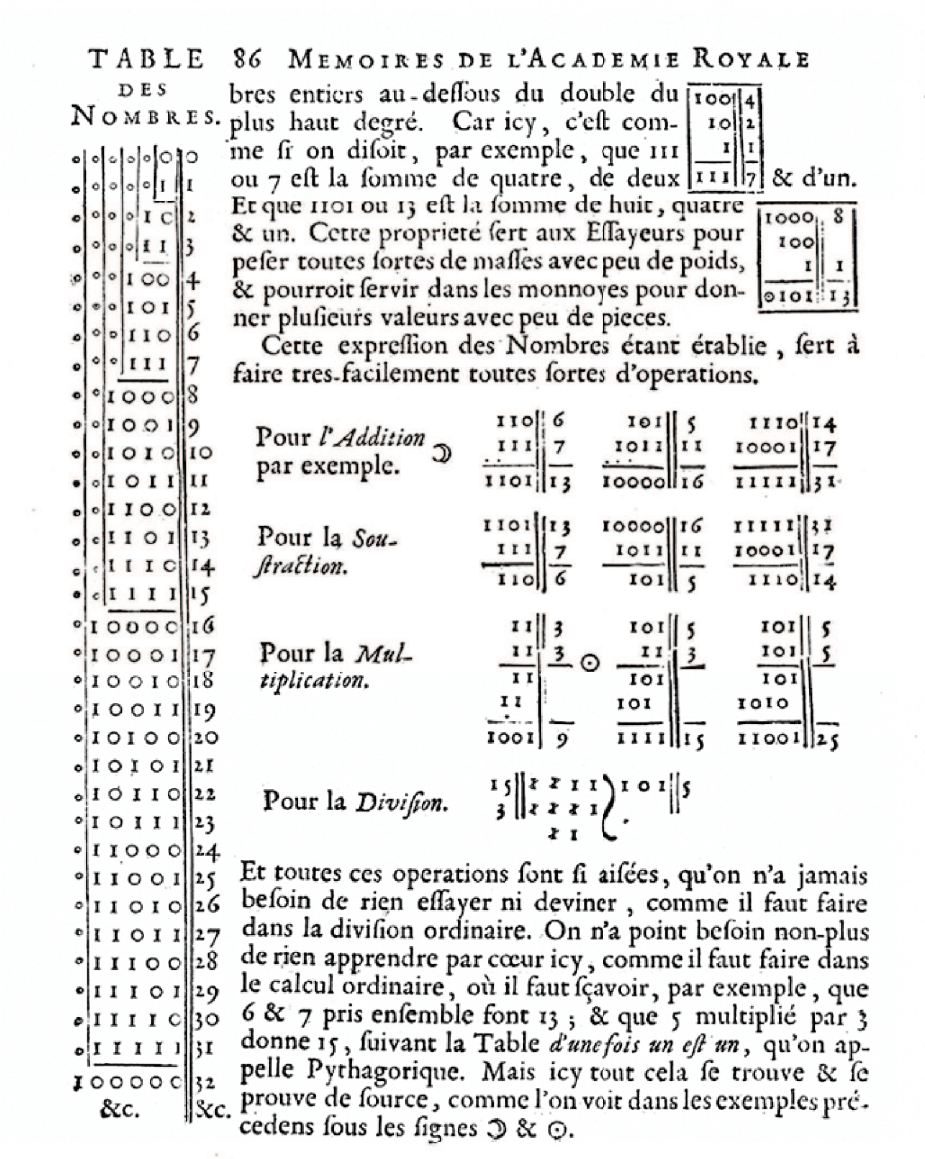


TABLE DES NOMBRES.

86 MEMOIRES DE L'ACADEMIE ROYALE

0	0
1	1
10	2
11	3
100	4
101	5
110	6
111	7
1000	8
1001	9
1010	10
1011	11
1100	12
1101	13
1110	14
1111	15
10000	16
10001	17
10010	18
10011	19
10100	20
10101	21
10110	22
10111	23
11000	24
11001	25
11010	26
11011	27
11100	28
11101	29
11110	30
11111	31
100000	32
&c.	&c.

bres entiers au-deſſous du double du plus haut degré. Car icy, c'eſt comme ſi on diſoit, par exemple, que 111 ou 7 eſt la ſomme de quatre, de deux & d'un.

100	4
10	2
1	1
111	7

Et que 1101 ou 13 eſt la ſomme de huit, quatre & un. Cette proprieté ſert aux Eſſayeurs pour peſer toutes ſortes de maſſes avec peu de poids, & pourroit ſervir dans les monnoyes pour donner pluſieurs valeurs avec peu de pieces.

1000	8
100	
1	1
1101	13

Cette expreſſion des Nombres étant établie, ſert à faire tres-facilement toutes ſortes d'operations.

Pour *l'Addition* par exemple. ☽

110	6	101	5	1110	14
111	7	1011	11	10001	17
1101	13	10000	16	11111	31

Pour la *Souſtraction.*

1101	13	10000	16	11111	31
111	7	1011	11	10001	17
110	6	101	5	1110	14

Pour la *Multiplication.* ⊙

11	3	101	5	101	5
11	3	11	3	101	5
11		101		101	
11		101		1010	
1001	9	1111	15	11001	25

Pour la *Diviſion.*

15	1 1 1 1	101	5
3	1 1 1 1		
	1 1		

Et toutes ces operations ſont ſi aiſées, qu'on n'a jamais beſoin de rien eſſayer ni deviner, comme il faut faire dans la diviſion ordinaire. On n'a point beſoin non-plus de rien apprendre par cœur icy, comme il faut faire dans le calcul ordinaire, où il faut ſçavoir, par exemple, que 6 & 7 pris enſemble font 13; & que 5 multiplié par 3 donne 15, ſuivant la Table *d'une fois un eſt un*, qu'on appelle Pythagorique. Mais icy tout cela ſe trouve & ſe prouve de ſource, comme l'on voit dans les exemples précedens ſous les ſignes ☽ & ⊙.

Consiglio Nazionale delle Ricerche
Istituto di Studi sul Mediterraneo

Il Dominio della Perfetta Segretezza

Un saggio sulla incompletezza del teorema di Shannon sui sistemi a segretezza perfetta, con alcuni cenni sui fondamenti teorici su cui basare un nuovo e innovativo metodo di crittazione

di

Claudio Cappelli[1], Luca Amodeo[2], Lorenza De Lellis[3]

[1] **Claudio Cappelli** collabora nel campo della massima sicurezza e della crittografia con imprese private.

Ha partecipato ai lavori condotti per la International Conference on DMS Distributed Multimedia Systems (San Francisco, USA, 2009, Florence, Italy, 2011).

[2] **Luca Amodeo** opera nel settore privato, in particolare presso società leader in Europa nella consulenza informatica.

E' interessato alla risoluzione di problemi logico-matematici e alla modellizzazione rivolta ad attività nel campo della cibernetica.

[3] **Lorenza De Lellis** opera presso il Dipartimento di Farmacia dell'Università Federico II di Napoli.

E' impegnata nel privato, in attività di ricerca sull'impiego delle strutture matematiche nei sistemi di sicurezza applicati al settore sanitario.

Consiglio Nazionale delle Ricerche
Istituto di Studi sul Mediterraneo
© Cnr Edizioni, 2022
Piazzale Aldo Moro, 7 - 00185 Roma

ISBN 978-88-90-95006-3 (versione elettronica)

Elaborazione ed impaginazione:

Aniello Barone, Paolo Pironti, Giovanni Ruggiero

INDICE

SECONDA PARTE

Abstract

Specialmente nel corso degli ultimi anni, sono state da noi accumulate diverse esperienze e redatti più documenti resi talora segreti, dove erano analizzati molteplici aspetti – sia di carattere pratico che teorico – sulla creazione di un sistema di segretezza dal carattere perfettamente sicuro.

Nei sistemi perfetti, per quanto scrive Claude Shannon in *Communication Theory of Secrecy Systems* (1949) abbiamo uno stato dove la conoscenza di quanto vale il crittogramma nulla aggiunge alla probabilità di risalire al messaggio; l'aver intercettato il crittogramma, non consente a chi attacca di aggiungere nuova informazione a quella che già possiede.

In proposito qui intendiamo tuttavia riprendere il lavoro svolto, stimando criticamente il saggio di Shannon così da mettere in luce criticità da affrontare e soluzioni da offrire soprattutto in merito alla difesa delle chiavi crittografiche.

E' data dunque dimostrazione:

- **(i)** dell'incompletezza del teorema di Shannon sulla crittazione a sicurezza perfetta;
- **(ii)** di un teorema originale secondo cui non dovrebbe essere la chiave a risultare un random e di maggior lunghezza[4] ma *a prescindere se funga da chiave o messaggio* almeno uno degli operandi da impiegare in fase di codifica[5].

[4] Come richiesto allo stato nei sistemi perfetti, dove sarà la chiave a essere random e di lunghezza ugual-maggiore rispetto a quella del messaggio.

[5] Suddetto teorema chiude nella Seconda Parte il corrente lavoro, ma è anticipato dai lemmi (**B**)(**C**).

Metodo

La trattazione che segue si sviluppa lungo due versanti paralleli, **uno volto a seguire alcuni aspetti meramente teorici e un altro dove si buttano le basi di un effettivo sistema di crittazione**.

Volendo mimare perciò gli sforzi di chi intenda forzare un *secrecy system*, nei prossimi capitoli si procederà in risalita da valle a monte e cioè dall'esame del crittogramma alla generazione e distribuzione di messaggi e chiavi crittografiche.
Allo scopo son dunque forniti i rudimenti di un innovativo sistema di cifratura che affronta il gravoso problema della difesa delle chiavi.
Nella prima e seconda parte del corrente saggio sul "Dominio della Perfetta Segretezza" si vorrà dunque sviscerare la questione, andando a trattare quanto segue:

1) un teorema di caratterizzazione e due lemmi originali e altri enunciati e dimostrazioni (Prima Parte)
2) lo schema di un nuovo e innovativo metodo di crittazione detto OTP++ (Prima Parte)
3) alcuni passi del saggio di Shannon sui Secrecy Systems *(*1949*)* dove è enunciato il relativo teorema sulla perfetta segretezza, del quale sarà provata l'incompletezza (Seconda Parte)
4) un nuovo teorema ed una definizione originale di perfetta sicurezza dei sistemi crittografici (Seconda Parte)

A titolo di cronaca, facciamo tuttavia presente che son stati redatti ulteriori testi di cui diamo brevemente conto sebbene siano coperti da segreto industriale e perciò esclusi dal novero della corrente pubblicazione. Essi infatti riguardano:

5) una trattazione degli oggetti e delle istruzioni che concorrono a definire il sistema di crittazione da noi offerto, dove la prassi della perfetta sicurezza è estesa alla cifratura di chiavi di lunghezza arbitraria
6) talune problematiche più specifiche come quelle che discendono dalla adozione d'una architettura paritaria o dal contrasto ai pericoli derivanti dalla malleabilità dei messaggi cifrati

Memento

Nel lavoro a seguire, testo, formule e simboli matematici, figure e tabelle, sono tutte riprodotte in tonalità di grigio.

C'è tuttavia una eccezione che riguarda le lettere con cui indichiamo (**A**) *un teorema di caratterizzazione* (**B**)(**C**) *due lemmi originali* (**D**) *un corollario,*

dove il teorema di caratterizzazione elenca i requisiti richiesti nella codifica a sicurezza perfetta così come formalizzata da Shannon;

mentre lemmi e corollario sono proposizioni illustrate per la prima volta nel 2018 nel corso d'una *due diligence* promossa da una nota azienda a partecipazione statale, e affinate in un report del 2020.

Essi lemmi offrono una più compiuta generalizzazione di quanto formulato in letteratura, spostando l'attenzione dalla destinazione d'uso attribuita al segnale (messaggio, chiave, crittogramma) alla struttura del medesimo (*aleatoria* parlando di stringhe selezionate con uniforme distribuzione di probabilità dal loro insieme di riferimento).

Talvolta parleremo pure di *files* o sequenze o flussi o bit stream, indicandoli sinteticamente come *random* o *no-random*.

Essendo che intendiamo segnalare in modo intuitivo se nel testo si sta ragionando degli enunciati del 1949 (**A**) o in rapporto a quanto riassunto nel 2020 nella lettera dei lemmi (**B**)(**C**) e del corollario (**D**),

nel caso si è optato per il grassetto rosso con simboli che banalmente ci dicono "dove siamo".

Citazioni e incisi saranno in corsivo e staccate dal treatment attraverso l'impiego d'una cospicua interlinea che li separi dal resto.

Qualora tuttavia ci siano citazioni senza fonte, saranno da implicitamente riferire al saggio di Shannon

Infine esattamente come vediamo qui sopra, ci sono righe o parti di riga poste in neretto e messe in risalto da un evidenziatore.

Pur non essendo sempre di particolare rilievo nell'economia del corrente lavoro, nondimeno riportano dettagli da tenere a mente per non perdere il filo della trattazione.

Parole Chiave

- Sicurezza Computazionale
- Sicurezza Perfetta
- One-time Pad
- Probabilità
- Casualità
- Impredicibilità
- Entropia
- Sorgenti
- Yao's XOR lemma

Glossario

Abbiamo redatto un glossario per non essere costretti a interrompere troppo spesso il filo della narrazione. Esso pertanto serve da pro memoria fissando il significato di talune espressioni da noi impiegate nel corrente lavoro.
Sono altresì fornite delle nozioni di base e delle definizioni formulate da Shannon (del cui saggio forniamo una lettura critica) e in genere in letteratura; tali nozioni e definizioni sono talora riprese e variamente sviscerate tanto nella Prima che nella Seconda Parte del saggio.

Sistema e suoi Componenti

Crittografia

E' l'arte di offuscare messaggi attraverso l'impiego di una maschera detta chiave che consente di passare *attraverso una trasformazione* dallo spazio del messaggio a quello del crittogramma.

Sistema Crittografico o Critto-Sistema

Un sistema crittografico è dato dalla quintupla ($\mathbb{M}$, $\mathbb{K}$, $\mathbb{C}$, $\mathbb{F}$, $\mathbb{G}$) dove
$\mathbb{M}$ è l'insieme di tutti i valori detti *m* che formano messaggi di lunghezza *L*
$\mathbb{K}$ l'insieme di tutti i valori detti *k* che formano chiavi di lunghezza *l*
$\mathbb{C}$ l'insieme di tutti i valori detti *Critto* che formano i relativi crittogrammi
$\mathbb{F}$ l'insieme di tutte le funzioni di codifica per le quali *f (m)* rilascia in uscita crittogramma *Critto*
$\mathbb{G}$ l'insieme di tutte funzioni di decodifica per le quali *g (Critto)* rilascia in risalita messaggio *m*

Sistema Perfetto od a Sicurezza perfetta o Incondizionata

Per Shannon è detto perfetto un critto-sistema il quale,
per ogni *m* appartenente ad $\mathbb{M}$ e per ogni *Critto* appartenente a $\mathbb{C}$,
soddisfi l'eguaglianza secondo cui P ($\mathbb{M}$ = *m* | $\mathbb{C}$ = *Critto*) = P ($\mathbb{M}$ = *m*),
dove $\mathbb{M}$ è l'insieme dei valori dei possibili messaggi,
dove $\mathbb{C}$ è l'insieme dei valori dei possibili crittogrammi.

Motivo per cui un sistema sarebbe perfetto se la conoscenza di quanto vale il crittogramma nulla aggiunge alla probabilità di un attaccante dotato di risorse infinite, di risalire da esso a chiave e messaggio.
Sul piano pratico non si ha modo di risalire al messaggio a prescindere dal tempo impiegato e dalla potenza di calcolo adoperata da chi intenda forzare il sistema.

Sistema a Sicurezza Computazionale

Diciamo a sicurezza computazionale quel sistema di crittazione per cui – tratti da insieme $\mathbb{M}$ due messaggi m_1 ed m_2 che gli appartengono – e preso da insieme $\mathbb{C}$ un crittogramma *Critto* relativo ad uno solo di codesti, un attaccante che impieghi un algoritmo probabilistico efficiente, non sarebbe in grado di distinguere l'uno dall'altro; non sarebbe cioè capace di indicare quale tra m_1 ed m_2 corrisponda al crittogramma di cui conosce il valore.
Diciamo allora in modo più formale, che per ciascun algoritmo probabilistico $\mathbb{A}$ che spenda i suoi passi in tempo polinomiale, esiste una trascurabile funzione $\varepsilon(\bullet)$ tale che per ogni $n \in \mathbb{N}$ e per ogni coppia di messaggi $m_1, m_2 \in \{0,1\}^n$ si possano distinguere le relative distribuzioni con probabilità al più pari ad $\varepsilon(n)$.

Dove essendo che $\mathbb{A}$ avrà in ogni caso almeno una probabilità P di un mezzo su uno di giungere al giusto messaggio,
comunque diremo che per ogni $\mathbb{A}$ e per ogni lunghezza finita del messaggio, si avrà una trascurabile funzione ε tale che $P \leq \frac{1}{2} + \varepsilon(n)$

In concreto un sistema a sicurezza computazionale può pur essere scalato da chi attacca, ma impiegando risorse temporali e potenze di calcolo ritenute irrealistiche giacché superiori a quelle utili allo scopo di decifrare il messaggio[6]

Si sappia poi che diremo talora "convenzionali" tali sistemi per il fatto che riguardano la stragrande maggioranza di quelli attualmente in uso.

OTP (One-time pad)

Costituisce un metodo a sicurezza perfetta di cui è stata dimostrata l'efficacia ma che ha avuto un impiego alquanto occasionale.
Esso comporta l'uso di chiavi crittografiche:

1) di lunghezza ugual-maggiore rispetto a quella del messaggio;
2) generate da sorgente discreta effettivamente aleatoria;
3) da non impiegare più di una volta.

Scendendo sul pratico si osserva che il suo impiego è frenato dal problema della distribuzione sicura di chiavi molto lunghe e perciò problematiche.

Messaggio o Messaggio in Chiaro

Muovendo da quanto illustrato da Shannon, per messaggio s'intende il segnale ingenerato da un *message-source* capace di rilasciare flussi che contengono l'informazione che mittente (Alice) intende trasmettere a ricevente (Bob).
Sono specificatamente indicati quali "messaggi in chiaro" in contrapposizione ai cosiddetti "messaggi cifrati o crittogrammi" sia quelli non ancora cifrati che quelli ormai decodificati.

Chiave Crittografica

La chiave è un segnale che funge da paradigma di un algoritmo di trasformazione $\boldsymbol{T}$ per cui si passa dallo spazio del messaggio a quello del crittogramma.
Secondo Shannon è ingenerata da una sorgente *key-source* e si frappone come una maschera tra il messaggio in chiaro in accesso e quello cifrato in uscita.

Crittogramma o Messaggio Cifrato

Il crittogramma è dato dal segnale rilasciato dall'algoritmo di trasformazione $\boldsymbol{T}$ che prende in accesso chiave e messaggio e li compone rilasciando in uscita il flusso cifrato.

Le voci a seguire sino a "tradizionale" non sono contemplate in letteratura essendo impiegate nella sola economia del corrente lavoro.

Messaggio Speciale

Per "messaggio speciale" intendiamo un segnale privo di valore semantico e sintassi, ma nondimeno trasmesso in forma di messaggio da mittente (Alice) a ricevente (Bob).
Esso è trasmesso per fungere da chiave crittografica o *nel nostro caso* per fornire in OTP++ la "materia greggia" con cui costruire chiavi crittografiche.

Messaggio Consueto o Finale

Sarebbe un messaggio in chiaro dotato di semantica e sintassi, del quale s'intenda rimarcare la distanza dal cosiddetto messaggio speciale.
Non è pertanto mai trasmesso per fungere da chiave o per fornire la "materia greggia" con cui costruire chiavi crittografiche.

Chiave della Chiave

[6] **Fatto salvo il caso di un salto tecnologico (si veda in proposito l'ultimo capitolo della Seconda Parte) che modifichi in profondità lo stato dell'arte.**

Come la chiave funge da paradigma dell'algoritmo di trasformazione che consente di passare dallo spazio del messaggio a quello del crittogramma, la cosiddetta "chiave della chiave" si frappone quale una maschera tra un messaggio speciale e il relativo crittogramma.

La particolarità sta nel fatto che tale segnale non concorre mai alla codifica di consueti messaggi ma solo di quelli da noi indicati come messaggi speciali.

OTP++ (One-time pad plus plus)

Si intende per *One-time pad plus plus*, un nuovo, innovativo sistema crittografico a sicurezza perfetta le cui premesse teoriche sono illustrate nel corrente lavoro.

Tradizionale

Con l'aggettivo "tradizionale" facciamo talora riferimento a definizioni invalse da tempo in letteratura ma per le quali offriamo un diverso approccio.
Altre volte sarà impiegato l'aggettivo "standard" (chiavi di lunghezza standard, definizione standard) che tuttavia *nel corrente lavoro* non implica il confronto con un diverso orientamentio.

Sorgenti e loro Segnali

PRNG

Con tale acronimo che vuol dire *Pseudo Random Number Generator* è indicata una sorgente SW costituita da un automa a stati finiti il quale, a partire da un seme, implementa un segnale di maggior lunghezza attraverso passi algoritmici che simulano una apparente casualità.
Tale sorgente numerica è definita dalla funzione $f : (\mathbb{Z}_2)^l \rightarrow (\mathbb{Z}_2)^L$
dove l ed L sono interi positivi,
dove $\mathbb{Z}_2$ è un campo popolato da valori binari 0, 1
dove $e \in (\mathbb{Z}_2)^l$ è il seme preso in input e sovente formato da una breve sequenza ingenerata da un processo stocastico,
dove il valore rilasciato in forma binaria da $\boldsymbol{f}(e) \in (\mathbb{Z}_2)^L$ è il flusso in uscita di lunghezza molto maggiore di quello in entrata.
Possiamo perciò dire che il processo di generazione numerica di un PRNG, è dato dall'insieme delle trasformazioni di spazio $\mathbb{Z}^l$ (cui appartengono i semi) in spazio $\mathbb{Z}^L$ (cui appartengono i flussi in uscita).
Ragion per cui a ogni elemento in entrata di spazio $\mathbb{Z}^l$ corrisponderà un distinto elemento in uscita di spazio $\mathbb{Z}^L$ per $L \gg l$
essendo che i passi di trasformazione andranno a accrescere il numero di bit presi in accesso, con la conseguenza di rappresentare più informazione di quella che effettivamente possiedono.

Segnale pseudo random

E' il segnale generato da un PRNG ed ha carattere periodico essendo che dopo un qualche lasso temporale *spesso assai ampio* si ripete uguale a se stesso.
Approssima proprietà statistiche vicine a quelle date da un flusso creato da sorgenti fisiche così da soddisfare le seguenti condizioni:

- sarà uniformemente distribuito su un intervallo specificato $[x_{min}, x_{max}]$ cosicché $f(x) = 1/(x_{max} - x_{min})$ in tale intervallo, $f(x) = 0$ fuori da tale intervallo.
  In base a tanto, possiamo perciò dare un insieme $\mathbf{X_{in}}$ di tutti i valori annoverati nell'intervallo così da scrivere la seguente funzione,

$$\boldsymbol{f}(x) = \begin{cases} 1/(x_{max} - x_{min}) & se\ x \in \mathbf{X_{in}} \\ 0 & se\ x \notin \mathbf{X_{in}} \end{cases}$$

- altresì sarà costituito di simboli tra loro indipendenti, per cui se la funzione di distribuzione per singolo simbolo è $f(x)$ quella per coppie di elementi successivi sarà $f(x, y) = f(x) f(y)$

Nondimeno diremo che in letteratura tale segnale non è ritenuto idoneo per un impiego nell'ambito dei sistemi a sicurezza perfetta.

TRNG o RNG

Con tali acronimi si allude al *True Random Number Generator* e cioè a una sorgente HW che assume in ingresso un segnale analogico che registra l'andamento di fenomeni fisici intrinsecamente aleatori impiegati come sottostante.

In genere, l'architettura di tale generatore comporta quanto segue:

(i) Un segnale analogico generato dalla sorgente che registra l'andamento dei fenomeni presi in ingresso;
(ii) Un digitalizzatore che campiona il segnale e lo converte in un flusso discreto che diremo *near random*;
(iii) Un segnale digitalizzato dato in input a un programma che lavora allo sbiancamento del flusso, così da rilasciare sequenze più qualitative di quelle assunte in ingresso.

In letteratura sono rimarcate le differenze che sembrano correre tra la struttura di un sistema di generazione numerica RNG che impone l'uso di una porta analogica, e quella del diverso sistema di generazione numerica PRNG.
Si avrà tuttavia modo di rimarcare il fatto di come siffatte differenze pur ricche di corpose conseguenze, siano più sottili di quanto si creda.

Segnale effettivamente random

E' il segnale generato da una sorgente fisica RNG e non avrà carattere periodico palesando una lunghezza arbitraria che dipende dalla durata delle osservazioni registrate dal sistema.
Sarà bene aggiungere che tale segnale intrinsecamente casuale è ritenuto idoneo per un impiego nei sistemi a sicurezza perfetta.

Tutte le voci a seguire non sono note in letteratura essendo offerte ai soli fini di quanto diremo nel corrente lavoro.

Segnale del Messaggio, Segnale della Chiave

Per segnale si intende l'informazione contenuta in una trasmissione a prescindere dal canale che la conduce.
Tanto il messaggio che la chiave sono ritenuti dei segnali per cui non si parla di "segnale del messaggio" o "segnale della chiave".
Se faremo tuttavia uso di tale espressione, è per il fatto che intendiamo riferirci al segnale come mero segnale numerico.
Nel caso non daremo risalto all'impiego cui è destinato (messaggio, chiave, crittogramma) ma alla sua informazione digitale che ci consente di mettere a nudo la struttura algebrica delle operazioni di cui vorremo trattare.

Segnale no-random o non random

Con tale sintetica dicitura indichiamo un segnale numerico poco sparso, che risulta perciò debolmente impredicibile entro la correlata soglia di complessità (per una classe che *a seconda della tolleranza adottata* risulterà ad esempio solvibile per una macchina di Turing deterministica o invece probabilistica);
talora per segnale *no-random* intenderemo un flusso che non importa (ai nostri fini) se sia stato o meno ingenerato da sorgente aleatoria.

Macrocode o Macrocode Greggio

Un ampio segnale numerico ingenerato da sorgente aleatoria, che fungerà da riserva da cui attingere file o sequenze dette gregge.

Segnale Greggio, file e sequenze Gregge

Per segnale greggio o per file e sequenze gregge intendiamo distinte parti estratte dal Macrocode, da impiegare nella costruzione di file vergini**.**
A prescindere dallo specifico talora i cosiddetti *files* saranno indicati con dei quasi-sinonimi.
Sebbene il termine infatti letteralmente significhi "archivio" è anche inteso in maniera informale come flusso di dati digitalizzati.
In effetti la voce file è talora riferita al contenitore e tal altra al contenuto (come quando si dice che un file "è cifrato" dove a esser cifrata non è certo la confezione).
Noi tuttavia impiegheremo tale voce nel senso di *data stream* da lavorare o da immagazzinare su un supporto di memoria.
D'altra parte anche il temine stream o flusso è talora indicato nel senso di canale e tal altra in quello di "ruscello" ovvero di successione numerica che può essere trattata bit a bit.

Materia Greggia o Bruta

Con tale voce intendiamo in modo generico sia il segnale numerico da cui estrarre dei file, sia quello da comprimere o espandere;
in breve, tale "materia" sarà costituita da flussi da lavorare in successive trasformazioni o da cui estrarre stringhe di minor lunghezza.

Segnale Vergine, file o sequenze Vergini

Sequenze ottenute attraverso passi di trasformazione eseguiti sui file greggi attraverso un processo detto di *digestion*;
tali sequenze saranno impiegate quali chiavi di crittazione e potranno pur essere indicate con dei quasi-sinonimi o col generico nome di segnale vergine.

Materia Vergine

Altro nome dei file vergini lavorati in fase di trasformazione.

Classi di Segnali

Insieme $\mathbb{E}$

Insieme Universo cui appartengono tutti i possibili messaggi che *a priori* dalla visuale di chi attacca saranno di arbitraria lunghezza.

Classe $\Omega\Omega$

Classe di possibili messaggi la cui numerosità è limitata dalla lunghezza *L* del crittogramma, a sua volta correlata a quella dell'operando di maggior lunghezza chiave o messaggio che sia.

Classe Ω

Classe residua di possibili messaggi la cui numerosità è correlata alla lunghezza *l* dell'operando di minor lunghezza, chiave o messaggio che sia.

Classe Ω'

Classe di possibili chiavi di crittazione.

Prima Parte

In principio si pone l'accento sul confronto tra la consueta visione del problema della segretezza crittografica
e particolarmente di quella perfetta, così come illustrata da Shannon e dai suoi epigoni,
e quella offerta da due lemmi originali che muovono da una diversa prospettiva.
In particolare ci focalizzeremo sull'ipotesi corroborata da un esperimento mentale,
che non siano le sequenze binarie o alfanumeriche delle chiavi a dover essere di maggior lunghezza,
ma in una logica più generalizzata quelle generate da sorgenti aleatorie a prescindere se fungano da chiave o messaggio.

Capitolo I

Tema

(00)

Qui procediamo col compendio di precedenti contributi da noi redatti in passato e volti a inquadrare i fondamenti teorici di un nuovo metodo crittografico OTP++

E' quindi riportato un teorema (**A**) riassuntivo dei criteri fissati in letteratura per definire i sistemi detti a sicurezza "perfetta" o "incondizionata" e due lemmi originali e un corollario che ne consentono una più compiuta generalizzazione (**B**)(**C**)(**D**)

In breve, da una parte si dà un fermo immagine del concetto di perfetta segretezza per come elaborato a partire dal secolo scorso, e dall'altra si comincia a delineare il profilo d'un nuovo e diverso approccio.

Compendio

(01)

Sulla scia del fondamentale teorema di *Shannon* offerto nel breve saggio *Communication Theory of Secrecy Systems***,** Bell System Technical Journal (1949),

è detto che abbiamo "sicurezza incondizionata" o "segretezza perfetta" quando la conoscenza[1] di quanto vale il crittogramma nulla aggiunge alla probabilità di risalire al messaggio.

Se pure qualcuno dovesse riuscire nell'intento d'intercettare il crittogramma, non vedrebbe crescere la facoltà di giungere al messaggio medesimo.

Ciò ha tuttavia luogo a patto che siano onorati taluni requisiti dedotti dal suddetto teorema[2] che qui riportiamo in una formulazione diversa da quella originale riprodotta più fedelmente nel glossario, per cui,

dato un sistema crittografico dove:

m sia il valore del messaggio in chiaro,

k quello della chiave crittografica,

Critto quello del crittogramma,

da questi non si potrà risalire ad *m* se per ogni messaggio $m \in \mathbb{M}$ avremo una ed una sola chiave *k* tale che $f(k, m) = Critto$,

a condizione che essa *k* appartenga a un insieme $\mathbb{K}$ di chiavi tra loro equiprobabili e perciò prese con uniforme distribuzione di probabilità.

Caratterizzazioni

Invero noi abbiamo che siffatta forma può essere tuttavia tradotta, come in effetti è sovente tradotta, in termini più pratici.

A ben vedere quando si dice che ciascuna chiave deve essere presa con uniforme distribuzione di probabilità dal suo insieme di riferimento, ciò necessariamente comporta che la sequenza dalla quale è fatta esponga un andamento aleatorio,

col risultato che ogni simbolo o bit che le appartiene {0,1} sarà equi-probabile così che la comparsa di questa o quell'altra sequenza di lunghezza data non prevarrà statisticamente sulle altre.

Al contempo, dire che per ogni messaggio avremo una ed una sola chiave crittografica, significa che – avendo *n* messaggi ovvero *n* messaggi appartenenti ad $\mathbb{M}$ – ci saranno sufficienti chiavi da consentire che a ciascun possibile dispaccio corrisponda una chiave capace di generare *Critto* come crittogramma.

Sulla scorta delle congetture e del teorema di Shannon se le N chiavi appartenenti a $\mathbb{K}$ tuttavia fossero più degli *n* messaggi appartenenti a $\mathbb{M}$,

[1] Talora detto *man in the middle*.

[2] Indicato come *teorema sei* nella pubblicazione del 1949.

questi non sarebbe un problema (giacché nel sistema molte chiavi sarebbero da scartare ma – per ogni missiva $m \in \mathbb{M}$ – lo stesso ci sarebbe una ed una sola chiave che consenta in uscita il rilascio del crittogramma intercettato) ma è diverso se fossero meno.

In tale circostanza si conta infatti una certa quantità di dispacci cui non corrisponde una chiave capace di dare *Critto* in uscita e quindi una quota da escludere in partenza dalla corrispondenza biunivoca con gli elementi di $\mathbb{K}$ col risultato di restringere il campo delle probabilità da prendere in considerazione avendo

$\#\mathbb{K} < \#\mathbb{M}$

In altre parole alcuni messaggi non potrebbero mai dare *Critto* come crittogramma mancando la chiave che può condurre a tale risultato.

Invero tra le altre cose ..., ciò mette in luce la correlazione che passa tra la lunghezza d'un flusso come sarebbe quello del messaggio e la quantità di istanze che ne discendono;

ciò altresì evidenzia l'analogo nesso tra la lunghezza d'una stringa come quella della chiave e il numero di configurazioni che consente.

Se gli insiemi $\mathbb{M}$, $\mathbb{K}$ devono poter rispettivamente riguardare tutti i possibili messaggi e tutte le possibili chiavi, ciò necessariamente significa che *dato un alfabeto finito di due soli simboli* 0, 1 come sarebbe un alfabeto binario,

e assegnata una misura l ai messaggi di insieme $\mathbb{M}$ ed una diversa o uguale misura L alle chiavi di insieme $\mathbb{K}$,

la cardinalità dei rispettivi spazi dipende dalla lunghezza delle stringhe (di chiave o messaggio) che gli appartengono, giacché stringhe di maggior lunghezza consentono un maggior numero di configurazioni e perciò di possibili valori in uscita[3].

Dove da ciò pertanto si deducono tre proposizioni che suonano nel seguente modo:

- **(1)** la chiave crittografica ai fini della perfetta sicurezza del sistema, dovrà essere di lunghezza uguale o anche maggiore rispetto a quella del messaggio (e quindi tendenzialmente infinita in funzione della lunghezza arbitraria del messaggio medesimo);
- **(2)** la chiave sarà effettivamente random nel senso che dovrà essere generata da una sorgente fisica aleatoria;
- **(3)** la chiave non sarà impiegata più d'una volta[4].

[3] Si tenga conto che i messaggi di insieme $\mathbb{M}$ in concreto sarebbero tutti i messaggi di lunghezza uguale a quella del dispaccio realmente redatto dal mittente; le chiavi di insieme $\mathbb{K}$ tutte le chiavi di lunghezza uguale a quella effettivamente generata dalla sorgente.

[4] Vedremo che tale ultimo punto è corollario dei precedenti.

Dette frasi si possono dunque riassumere in un unico teorema di caratterizzazione (**A**)
per cui diciamo che ai fini della perfetta sicurezza del sistema, è necessario che la sequenza binaria o alfanumerica che funge da chiave, sia formata da simboli ingenerati da sorgente discreta effettivamente aleatoria, sia di lunghezza ugual-maggiore rispetto a quella del messaggio, non sia impiegata più d'una volta.

Criticità 1949

(02)

Al dunque tuttavia abbiamo che le conclusioni suggerite da Shannon, da una parte annunciano ottime notizie ma dall'altra le negano.

Se è vero com'è vero ch'è stato conseguito l'obiettivo storico d'assicurare i messaggi in modo da renderli indecifrabili (anche per quanti siano in possesso di un futuribile computer quantico, pure per chi abbia tutto il tempo del mondo, persino per coloro che possano contare su risorse infinite), è altrettanto vero che nuove criticità s'addensano all'orizzonte.

Sin da principio alcuni aspetti della crittazione a sicurezza perfetta, furono guardati con sospetto e molti dubbi sorsero sul suo impiego; sebbene sia bene considerare che tali giudizi affondano le radici in realtà ormai lontane quando nel 1949 si buttarono le basi di quella che possiamo considerare la data di nascita della crittografia contemporanea.

Rileggendo le fonti effettivamente misuriamo con qualche stupore la distanza che corre da adesso ad allora, potendo toccare con mano come molte criticità si siano affievolite …, come più ambienti domestici e lavorativi siano divenuti desueti … e come non poche prospettive si siano capovolte alla luce del progresso tecnologico e dello sviluppo socio-economico (ai tempi chi avrebbe potuto ad esempio immaginare un uso della crittografia fuori dall'ambito militare e della sfera della massima sicurezza?).

Alcune preoccupazioni avvertite sul finire degli anni quaranta del secolo scorso, suonano perciò come improbabili e le righe che seguono ce ne danno una plastica dimostrazione.

The key must be transmitted by non-interceptible means from transmitting to receiving points. Sometimes it must be memorized.

It is therefore desirable to have the key as small as possible.

Enciphering and deciphering should, of course, be as simple as possible. If they are done manually, complexity leads to loss of time, errors, etc. If done mechanically, complexity leads to large expensive machines.

In some types of secrecy systems the size of the message is increased by the enciphering process. This undesirable effect may be seen in systems where one attempts to swamp out message statistics by the addition of many nulls, or where multiple substitutes are used.

La chiave deve essere trasmessa con mezzi non intercettabili dalla emittente ai punti di ricezione. A volte deve essere tenuta a mente.

Sarà dunque auspicabile avere la chiave più piccola possibile.

Cifrare e decifrare dovrebbe, ovviamente, essere il più semplice possibile. Se vengono eseguite operazioni a mano, la complessità porta a perdite di tempo, errori, ecc. Se eseguite meccanicamente, la complessità richiede macchine molto costose.

In alcuni tipi di sistemi di segretezza la dimensione del messaggio è aumentata dal processo di cifratura. Questo effetto indesiderato può essere presente nei sistemi in cui si tenta di inondare le statistiche dei messaggi con l'aggiunta di molti valori nulli o dove vengono utilizzati più sostituti.

A rileggere quanto scriveva imperterrito Shannon, siamo in effetti pervasi da una sensazione di sgomento essendo chiaro che ci parla da un altro mondo;
mentre noi temiamo di impiegare chiavi troppo corte che rischiano di essere forzate attraverso attacchi a forza bruta ..., il nostro si pone il problema di come registrare lunghi codici (lunghi secondo il suo metro di giudizio, molto lontano dal nostro) in tempi dove per "memoria" ancora si intendeva quella degli esseri viventi.
Anche l'idea che i calcoli si possano fare a mano, a meno di non impiegare improbabili macchinari o "*large expensive machines*", ci fa intendere in un lampo con quale medioevo informatico ci dobbiamo confrontare.
Sebbene occorra aggiungere che per quanto non si possa negare che talune questioni siano superate per le opportunità offerte dalle nuove tecnologie …, non mancano criticità che *a torto o ragione* continuano a esser regolarmente estratte dal cilindro del prestigiatore.
Per dire …, si è a lungo ragionato del carattere aleatorio dei flussi da impiegare in crittografia, nonostante il fatto che la questione della "vera" casualità fosse percepita diversamente da adesso giacché non se ne parlava o se ne parlava pochissimo;
persino Shannon disse poco o nulla in proposito (nel senso che nulla disse sul tema di quali sorgenti siano da ritenere casuali a prescindere dall'analisi a posteriori della qualità del segnale) e nessuno potrebbe giurare su cosa intendeva nel dire che "*all keys are equally likely"*.
Insomma è vero che le chiavi devono pur essere tra loro *equally likely* se vogliamo garantire efficacia a un sistema perfettamente sigillato, ma non è chiaro quali processi di selezione stocastica fossero ritenuti sufficienti allo scopo.
In altre parole, la tematica del confine che passa tra mondo delle sorgenti random e mondo di quelle pseudo random rimase sottotraccia siccome – benché da un certo momento in avanti fosse stabilito l'impiego di un segnale casuale[5] – nondimeno non era fissata la forma in cui si credeva lecito poterlo avere,
tanto che i primissimi tentativi di realizzare un cifrario dato per "incondizionatamente sicuro" introdussero taluni "trucchi" per accrescere l'ampiezza del flusso numerico.
Prendiamo comunque atto che dopo qualche incertezza, la comunità scientifica operò una scelta tra sorgenti "pseudo-random" e sorgenti "effettivamente random" chiarendo che solo le ultime sarebbero state usate nel campo della perfetta sicurezza;

[5] Fu questi il contributo dato negli anni venti del novecento, dal generale Joseph Mauborgne, allo sviluppo della cifratura a sicurezza perfetta.

perciò tutti i programmi di generazione pseudo random furono banditi dai sistemi che meritano l'ampolloso appellativo di "perfetti".
Da molteplici punti di vista il tema della aleatorietà resta tuttavia scivoloso e, sebbene venga (in parte) trattato nel terzo capitolo e altrove, ciò non toglie che assumeremo quanto segue quale stella per poterci orientare:

(1) **in via prudenziale non vorremo mai considerare casuale un segnale pseudo-random e cioè un flusso che implementi un *seed***;
il flusso così ottenuto discende da passi algoritmici che simulano, sebbene in modo realistico, l'andamento di un segnale aleatorio ma tanto non basta ai fini della perfetta segretezza del sistema;

(2) saranno pertanto da noi ritenuti casuali solo i flussi random generati per un tempo tendenzialmente infinito e quindi non periodico, così come accade quando si registrano flussi discendenti da sottostanti fenomeni fisici di tipo quantico o caotico.

Per la verità fino a non moltissimo tempo fa, i dubbi sul modo di produrre flussi "effettivamente random" parevano fondati;
ancora nel secondo dopoguerra la possibilità di creare sequenze aleatorie si limitava al lancio di dadi o alle estrazioni del lotto.
Intendiamo cioè dire che non si vedevano particolari progressi tanto che le stesse sequenze pseudo-random apparivano come rare eccezioni; talmente rare che nel 1955 diverrà un caso la prima pubblicazione cartacea di *One Million Number Digits* e cioè di un milione di numeri casuali meccanicamente ottenuti[6].
Non volendoci però dilungare nel racconto di precedenti che pure ci appassionano, tagliamo corto dicendo che le cose mutarono di botto da quando furono introdotti dispositivi capaci d'intercettare fenomeni intrinsecamente aleatori *(shot noise, beam splitter, reserve biased semiconductor juncition, reserve biased Zener diodes, photon bunching, etc.)* che consentirono l'offerta d'un ventaglio di prodotti a uso commerciale.
La lista di modelli e case di produzione di generatori di numeri effettivamente casuali, pertanto si moltiplicò nel breve volgere di un decennio;
l'ultimo salto ebbe tuttavia luogo quando furono introdotti sistemi che non erano progettati al solo scopo di creare numeri casuali …, ma a quello di poterlo fare con grande intensità nell'unità di tempo.

[6] *One million number digits with 100,000 normal deviates*, Rand Corporation, 1955.

Esempi di generatori che rilasciano flussi numerici casuali da pochi Mega a più Gigabit/sec

Produttore	Paese	Anno	Modello	Portata
ID Quantique	Switzerland	2016	Quantis	*4 Mbit/sec.*
Ubid.it	USA	2016	True RNG v3	*N/A*
Comscire	USA	2016	PQ4000KS	*N/A*
Quintessence	Australia	2012	qStream	*1 Gbit/sec.*
Letec	Giappone	2013	Grang Server	*1,2 Gbit/sec.*
Intel	USA	2013	Ivy Bridge-EP	*3 Gbit/sec.*

Da quando ciò accadde e quindi da quando furono introdotti generatori che frullano oltre un miliardo di bit al minuto secondo ..., stupisce come l'antica eccezione continui a circolare come un mantra del quale sembra essersi smarrito il significato ..., quasi a dimostrare quanto sia difficile tener vivo il contatto con la realtà dello sviluppo industriale e commerciale che non conosce soste e non attende chi si attarda non tenendo il passo[7].

[7] A nostro parere, nel campo della generazione numerica, svolgono un ruolo di rilievo gli enti di standardizzazione del segnale che creano un terreno comune tra sviluppo competitivo e ricerca scientifica.

Criticità 2020

(03)

Pur non potendo negare che esistano ulteriori rilievi sull'impiego dei sistemi perfetti, occorre dire che poco convincono assumendo talora toni talmente spicci da apparire come un freno piuttosto che uno stimolo al fiorire di nuove congetture[8];

ciò detto, noi qui non intendiamo inseguire dibattiti troppo accesi per la semplice ragione che un problema esiste e non è un problema dappoco.

Il problema dei problemi è infatti riposto nella trasmissione sicura di chiavi crittografiche,

sebbene *a differenza di quanto si creda* non sta nel "dettaglio" della lunghezza (dove sappiamo che nei sistemi perfetti ciascuna chiave sarà di misura ugual-maggiore rispetto a quella del messaggio) che si annida il diavoletto del famigerato proverbio.

Intendiamo dire che la critica assai battuta secondo cui sarebbe la lunghezza a rendere difficile la gestione dell'annoso problema della distribuzione delle chiavi di crittazione ..., risulta stravagante se consideriamo a quali sfide è chiamata la moderna tecnologia, ed a quali ponderose problematiche deve porre rimedio ogni giorno.

Se abbiamo un messaggio di lunghezza l, ed un flusso di chiave di pari lunghezza l, qualora ci sia modo di gestire tecnicamente il primo, non si comprende perché dovrebbe mancare la facoltà di gestire tecnicamente il secondo.

La verità vera è infatti diversa essendo che, volendo muovere dal teorema di caratterizzazione in (**A**) che riassume quanto accolto in letteratura,

dovremmo operare a partire dall'assunto per cui la distribuzione di chiavi della stessa misura del messaggio (perfetta segretezza) non sia possibile, dato che a sua volta richiede l'uso di ulteriori chiavi di lunghezza ugual-maggiore rispetto a quella della chiave impiegata.

Secondo quanto detto sinora, il bisogno di tutelare lunghe chiavi non discende perciò da un problema tecnico, quanto dal caso che mette a nudo una delle più battute criticità dei sistemi perfetti e cioè quella che vuole tali sistemi capaci di fornire messaggi dalla sicurezza formidabile ma dalle chiavi quasi indifendibili.

E ciò per il fatto che – a voler cifrare chiavi random di uguale o maggior lunghezza del messaggio, con altre chiavi random di uguale o maggior lunghezza della chiave – si ottiene un *loop* dove ogni codice richiede un secondo lunghissimo codice cui se ne aggiunge un altro e un altro ancora ..., procedendo attraverso passi cui si aggiungono altri passi senza che esista modo di chiudere il cerchio una volta e per sempre.

[8] Ad esempio si dice che la crittografia a sicurezza perfetta sarebbe poco pratica.

E' strano essendo che siamo ormai in grado di gestire e trasmettere ampi flussi di dati che non cambiano certo di natura a seconda se fungano da chiave o messaggio .

A ogni codifica non si farebbe null'altri che spostare di pochi passi il problema senza avere la possibilità di giungere al dunque.

Paradosso del Crittografo

(04)

E' un fatto che in precedenti lavori abbiamo messo tuttavia in dubbio la completezza di tale asserto.
In particolare si è prodotto un esperimento mentale e alcune dimostrazioni dove abbiamo provato che non tutti i vincoli usualmente dettati in letteratura sono algebricamente plausibili.
C'è infatti un caso dove i fatti hanno luogo diversamente da come li abbiamo supposti e da come vengono di solito raccontati;
un caso che occorre quando il messaggio – di solito caratterizzato da una semantica e una sintassi che comportano pattern e ricorrenze statistiche – abbia esso stesso contenuto stocastico.
In effetti si tratta di un evento poco considerato[9] ma niente affatto trascurabile, avendo luogo tutte le volte in cui non sia distribuito un normale dispaccio ma un segnale aleatorio da impiegare in forma di chiave crittografica;
è un poco ciò che accade se dentro un plico ci fosse un altro plico, dove il secondo pur morfologicamente identico a quello d'una busta postale, sia oggetto del trasferimento così da costituire il bene da dover recapitare.

Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem.
Spesso i messaggi hanno un significato; cioè si riferiscono o sono correlati, secondo un sistema, con certe entità fisiche o concettuali. Questi aspetti semantici della comunicazione sono irrilevanti per il problema ingegneristico.
(Dove presto vedremo che tale affermazione appare vera solo nei limiti in cui teniamo distinti gli aspetti di cui alla semantica in senso stretto, da quelli che attengono a una sintassi il cui andamento più o meno casuale incide su un aspetto di grande rilievo come sarebbe quello della maggiore o minore predicibilità del segnale).

Senza volerci pertanto calare in particolari su cui torneremo a momento debito ..., ci limitiamo a menzionare il fatto da cui siamo partiti in altri e diversi documenti quando dicemmo d'una *questio* da noi battezzata col nome di paradosso del crittografo.
Infatti delle tipologie di "paradosso" esistono sottili classificazioni tra cui si contempla quella d'una proposizione formulata in contrasto con quanto diffusamente accettato, ma che appare consistente se sottoposta ad analisi[10].

[9] Si tratta d'una circostanza poco trattata non perché poco interessante ma giacché ritenuta irrilevante ai fini del funzionamento di un sistema crittografico che sarebbe indifferente alla natura del messaggio.
[10] Intanto precisiamo che tale proposizione sarà in effetti data dai conseguenti lemmi originali (**B**)(**C**)

(i)

Poniamo allora che un crittografo intenda ad esempio cifrare l'incipit della Divina Commedia ma che, volendo agire in condizioni di perfetta sicurezza, impieghi una chiave fatta di sequenze di numeri casuali di lunghezza almeno uguale (se non maggiore) a quella del messaggio.

Nel caso, diremo pertanto ***m*** *il messaggio costituito da una copia digitale dell'incipit di nove endecasillabi della Commedia dantesca (nel mezzo del cammin di nostra vita mi ritrovai in una selva oscura …).*

Diremo ***k*** *la chiave random di misura ugual-maggiore rispetto a quella del messaggio.*

Diremo Critto il crittogramma che ne discende, motivo per cui, come usualmente accade nei sistemi perfetti, procederemo in XOR *avendo Critto =* ***m*** *+* ***k*** *(mod due)*

Secondo quanto ormai noto, a patto di impiegare la chiave una volta soltanto, in **(i)** *si sarà pertanto operato in condizioni di perfetta segretezza e quindi nel rispetto dei requisiti di cui al teorema di caratterizzazione* (**A**).

(ii)

Se però in una diversa occasione **(ii)** *esso crittografo intenda cifrare un messaggio dal contenuto incidentalmente identico a quello di chiave* ***k*** *di cui all'esempio in* **(i)***, avremmo lo strano caso d'un messaggio fatto di sequenze esse stesse random.*

Se poi tale messaggio fosse cifrato impiegando come chiave l'incipit della Divina Commedia (incipit immutato nella forma digitale sebbene diversamente inteso dalla mente del crittografo) avremmo un messaggio di numeri casuali cifrato con una chiave non casuale di lunghezza uguale o anche minore a quella del messaggio.

Diremo pertanto ***m'*** *tale messaggio random, per* ***m'*** *=* ***k***

Diremo ***k'*** *la chiave no-random*[11] *di misura uguale o anche minore a quella del messaggio, con* ***k'*** *=* ***m*** *dove* ***m*** *altri non sarebbe che il detto incipit della Commedia dantesca.*

Diremo Critto' il crittogramma che ne discende,

dove in somma XOR *sarà pertanto Critto' =* ***m'*** *+* ***k'*** *(mod 2)*

Avendo tuttavia ***m'*** *=* ***k*** *e* ***k'*** *=* ***m***

potremmo anche scrivere Critto' = ***m'*** *+* ***k'*** *(mod 2) =* ***k*** *+* ***m*** *(mod 2) = Critto*

Essendo pertanto esempio **(i)** *ed* **(ii)** *algebricamente identici, essi non potranno che esporre uguali proprietà,*

per cui se l'uno comporta il rilascio d'un crittogramma generato da un sistema perfettamente sicuro, dobbiamo supporre che anche l'altro implichi la creazione d'un crittogramma rilasciato in condizioni di perfetta segretezza.

[11] Per quanto il significato da noi attribuito al termine "*no-random*" o "*non random*" sia intuitivo, per un maggiore approfondimento ci riportiamo a quanto illustrato nel glossario.

Lemmi (B)(C)

(05)

Messo dunque a nudo il cosiddetto "paradosso del crittografo" di cui terremo lungamente conto per rispondere alle domande suscitate dal confronto tra **(i)** ed **(ii)**,
in proposito sono stati riferiti in più lavori tenuti sinora riservati, i seguenti lemmi **(B)(C)** che costituiscono una più generale estensione del teorema di caratterizzazione **(A)**,
che sebbene formalmente corretto finisce con l'escludere dalla sua pertinenza alcuni dei casi di maggior interesse.

(B)

Ai fini della perfetta sicurezza del sistema, sarà sufficiente che in codifica almeno uno degli operandi sia formato da simboli ingenerati da sorgente effettivamente aleatoria[12]*, a prescindere se faccia da chiave o messaggio;*

(C)

Ai fini della perfetta sicurezza del sistema, essendo che almeno un operando è formato da simboli ingenerati da sorgente aleatoria, dovrà pur essere di lunghezza ugual-maggiore rispetto a quella dell'altro operando, a prescindere se sia chiave o messaggio.

Ciò tuttavia conduce a conseguenze anti-intuitive giacché significa che qualsivoglia dispaccio composto di flussi casuali (come un codice creato da un generatore true random) non solo sarebbe "perfettamente" codificato da una stringa da noi detta *no-random* ma persino da una stringa *no-random* di misura inferiore rispetto a quella del messaggio, che poi sarebbe l'esatto contrario di quanto abbiamo sempre creduto di sapere.
E così compito di questo come dei prossimi capitoli sarà quello di rendere digeribile tale assunto dando prova dei caposaldi su cui abbiamo fondata una rilettura cui cercheremo di dar forma nelle prossime pagine.
Prendendo infatti spunto dal fatto che nei sistemi perfetti si richiede che ciascuna chiave sia random e di lunghezza ugual maggiore rispetto a quella del messaggio, prima di poterci calare nei particolari, **possiamo intanto dire che ciò non è vero tutte le volte in cui gli input da cifrare siano a loro volta aleatori**.

[12] Anche qui parliamo d'una sorgente discreta.

Esempi

Per non doverci smarrire cominciamo dunque col ricalcare le eventualità **(i)(ii)** da poco illustrate, per cui nella consueta logica dei sistemi perfetti prendiamo intanto le mosse dal caso più ovvio,
dove secondo consuetudine sarà il messaggio a non essere random e la chiave a esserlo, per una misura di lunghezza ugual-maggiore rispetto a quella del dispaccio.
A tal proposito si ricorda che la codifica XOR non richiede particolari alchimie in quanto nei sistemi perfetti, la difesa non dipende dalla complessità della risalita ma dalla pari frequenza di comparsa di ciascun messaggio cifrato tanto che è sufficiente sommare algebricamente (modulo 2) sequenza a sequenza per bloccare ogni assalto.
In altre parole, potremmo metaforicamente dire che la *coperta* data dalla profondità della chiave è abbastanza lunga da togliere il minimo riferimento statistico a chi attacca.

Esempio **(i)**

Primo Operando – Operando costituito dal messaggio in chiaro il quale è ottenuto dalla conversione in digitale dell'incipit della Commedia dantesca, reiterata sino a pareggiare la maggior lunghezza della chiave[13]

00101111001011110010111100101111001011110010111100000000000000111100000010111100101111001011110010111100101111000000000000001111000000101111001011110010111100101111000000000000000111100000010111100101111001011110010111100000000000000011110000 ...

Secondo Operando – Operando dato dalla chiave random che qui sappiamo essere di maggior lunghezza rispetto a quella del messaggio, come nei casi di scuola di perfetta sicurezza del sistema (**A**).

01101001101100100101101001010001101010011101001010100101101101001010101101001101101110011010011011001001011010010100011010100111010010101001011011010010101011010011010011101101110001110 ...

Risultato in uscita – Output che fissa la sequenza del crittogramma in uscita dalla somma a flusso tra l'operando del messaggio in chiaro e quello fornito dalla chiave di crittazione, secondo il tipico approccio di cui ai cosiddetti sistemi perfetti.

01000110100110110100101001011110100101101110001010100101101101101010010001101001101101001010010101110110110101011 ...

[13] Invero questa è una prassi che pur suonando strana, è del tutto normale nei sistemi perfetti quando secondo Shannon sarebbe la chiave a dover essere di maggior lunghezza.

La particolarità sta nel fatto che quando nei sistemi convenzionali si usa invece una chiave standard, succede il contrario essendo infatti la chiave di 126, 258 o 512 bit a esser reiterata sino a compensare la maggior lunghezza del messaggio in chiaro.

Volendo tuttavia riprendere anche il diverso caso dove era il messaggio e non la chiave a essere random, se proponiamo quanto detto in **(ii)** non dovremo nemmeno scomodare la proprietà commutativa cambiando l'ordine degli addendi …, giacché per lasciare inalterati i valori in gioco,
sarà sufficiente continuare a mettere innanzi il flusso dell'operando non casuale che qui tuttavia funge da chiave di minor lunghezza …, ed indietro a seguire quello dall'andamento effettivamente random che funge da messaggio in chiaro.
Possiamo dunque confermare che mappando bit a bit ciascun flusso …, appare facile confermare che l'operazione di cui all'esempio in **(i)** e quella di cui all'esempio in **(ii)** sono matematicamente identiche a prescindere da come saranno impiegati i segnali random e non random una volta decrittati dalla stazione ricevente.

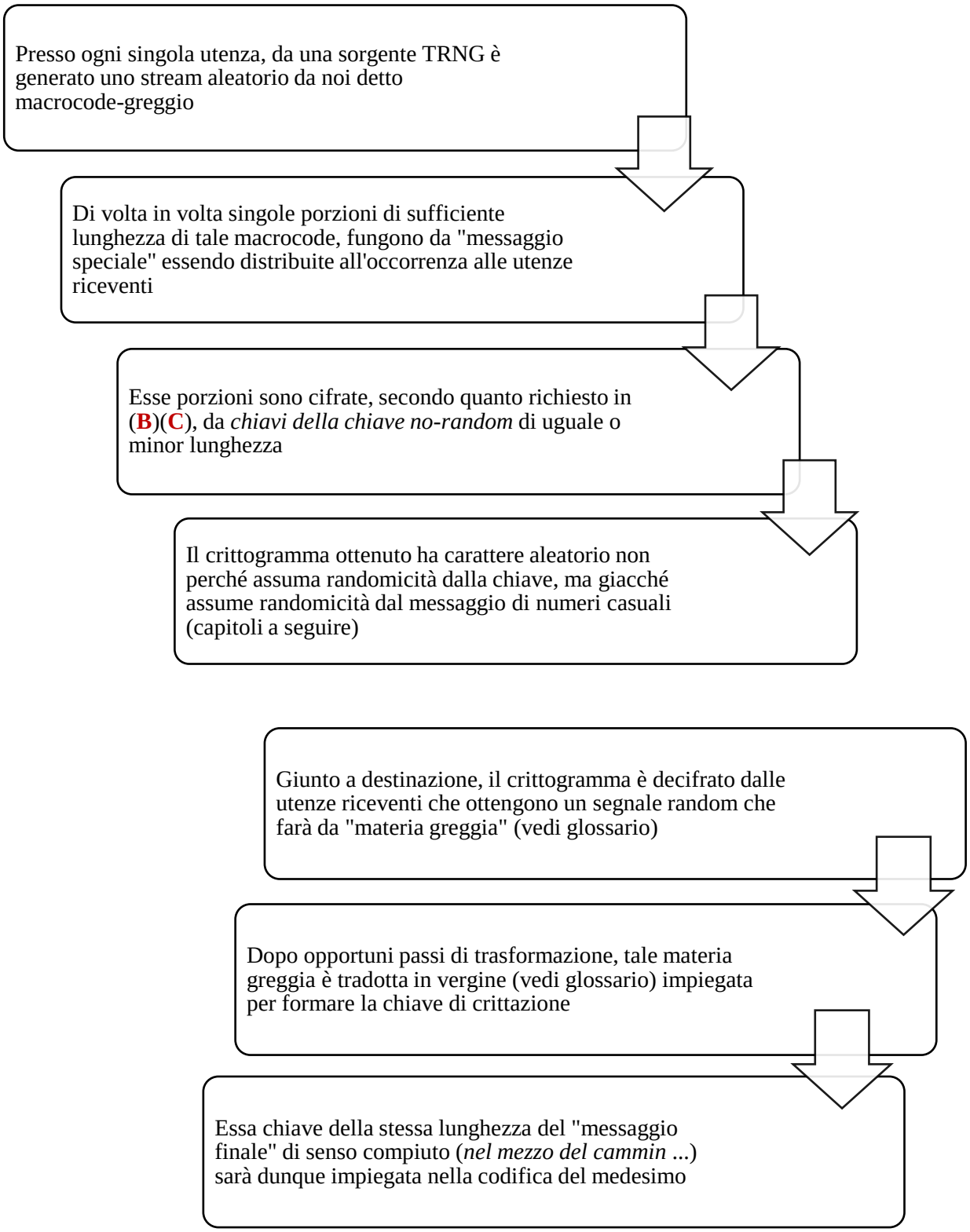
Esempio **(ii)**

Primo Operando – Operando espresso da una chiave *no-random* che incidentalmente corrisponde alla forma digitale dell'incipit della Divina Commedia, che sarà di misura inferiore rispetto a quella del messaggio essendo reiterata sino a pareggiare la lunghezza del medesimo
00101111001011110010111100101111001011110010111100000000000000011110000001011110010111
10010111100101111001011110010111100000000000000111100000010111100101111001011110010111100000000000000111100000 ...

Secondo Operando – Operando dato da un messaggio di maggior lunghezza, stavolta prodotto da un flusso generato da sorgente effettivamente aleatoria
01101001101100100101101001010001101010011101001010100101101101001010101101001101101110011010011011001001011010010100011010100111010010101001011011010010101101001101001101
01101110001011 ...

Risultato in uscita – Output che fissa il costrutto del crittogramma rilasciato dalla somma XOR tra l'operando che attiene al messaggio random e quello che attiene alla chiave *no-random*, che sarà talora da noi anche detta *chiave della chiave*
01000110100110110100101001011110100101101110001010100101101101101010010001101001101101001010011110100101101110001010100101101101101010010001101001101101001010010101110110110101011 ...

Figura 01 – Esempio di schema molto semplificato di un possibile metodo di crittazione OTP++



Corollario (D)

(06)

Se è vero come è vero che quanto abbiamo sinora evidenziato non si può trascurare consentendo di rimuovere un vincolo gravoso,

occorre tuttavia aggiungere che non andrà apprezzato oltre il dovuto così da finire col dargli più peso di quanto possieda;

Il fatto è che le opportunità offerte dal cosiddetto paradosso del crittografo, a loro volta creano una criticità non banale.

Un segnale random adoperato in modo inconsueto come messaggio e cioè quale "messaggio speciale" trasmesso da mittente a ricevente …, per non entrare in conflitto con la lettera dei lemmi (**B**)(**C**), infatti non sarà nuovamente impiegato nemmeno in forma di chiave crittografica.

Se è prescritto che nei sistemi perfetti ciascuna chiave debba essere usata una volta soltanto, qualora si dica che quanto vale per la chiave è altrettanto valido per qualsivoglia segnale aleatorio immesso nel sistema, ogni conseguente restrizione sarà necessariamente estesa anche a esso a prescindere dall'impiego cui è destinato.

Quando si scambino le proprietà di chiave e messaggio, per cui sarà questi ad essere random e quello a non esserlo, si rovesciano pure gli altri requisiti.

In ragione dell'ovvio principio secondo cui occorre guardare al carattere oggettivo dei flussi numerici e non alle mutevoli intenzioni dell'operatore umano (da cui discendono attribuzioni soggettive a seconda se stimi questa o quella sequenza come chiave o messaggio) non dovrebbe sorprenderci il fatto che qualora sia invertita la natura del segnale, si determina l'inversione di ciascuna proprietà correlata che migra da questi a quell'altro operando.

Infatti non conta come percepiamo un'attività di funzione ma come essa realmente si risolva sul piano algebrico.

Come emergerà dagli esempi di cui diremo, il reiterato uso d'una qualche successione numerica dall'andamento casuale (che faccia da chiave o invece da messaggio) non consente di soddisfare quanto richiesto in (**B**)(**C**) o nei casi che lo riguardino nel teorema di caratterizzazione in (**A**), essendo che altera il rapporto tra:

(1) lunghezza *L* della sequenza random che funge da operando della somma a flusso e che deve essere maggiore-uguale dell'altra *no-random*

(2) lunghezza *l* della sequenza che abbiam detta *no-random* e che sarebbe di misura minore o uguale.

Un fatto cruciale sta infatti nel considerare che il comportamento di due distinte successioni *no-random*, sommate a una uguale sequenza effettivamente random,

può essere assimilato a quello d'una unica concatenazione di maggior lunghezza e, per tale ragione, possiamo concludere definendo il seguente corollario (**D**) per cui,

ai fini della perfetta sicurezza del sistema, essendo che in codifica almeno un operando sarà formato di simboli generati da sorgente aleatoria e sarà di lunghezza ugual-maggiore rispetto a quella dell'altro operando, è necessario che questi sia impiegato solo una volta a prescindere se in forma di chiave o messaggio[14]

Esempio

Avendo infatti una sequenza ***y*** di variabile ***Y*** presa in modo aleatorio dal suo insieme di riferimento,
supponiamo che essa, la quale nel nostro esempio funge da chiave crittografica ***k***,
sia d'una qualche lunghezza cui attribuiamo un valore convenzionale pari a *50 bit*,
maggiore di quello della sequenza *no-random* ***x*** di variabile ***X*** che funge da messaggio ***m*** e cui
attribuiamo un valore convenzionale pari a *30 bit* dove 30 < 50
Possiamo tuttavia rilevare che – volendo impiegare una seconda volta tale ***y*** come chiave di crittazione –
il rapporto tra lunghezza della sequenza random e lunghezza di quella *non random* (dove esse stringhe fanno da distinto operando della somma a flusso) andrebbe a mutare.
Il fatto è che usare due volte lo stesso segnale aleatorio di chiave ***k*** comporta che essa

- sia prima ingaggiato da sequenza ***x*** (*no-random*) di messaggio ***m***
- e poi da altra sequenza ***x'*** (parimente *no-random*) di messaggio ***m'***

che sebbene individualmente più corte in ossequio a quanto previsto nel teorema in (**A**) e in termini più generali in (**B**)(**C**) una volta concatenate tra loro, potranno cumulativamente risultare di maggior lunghezza.

Se dovessimo infatti avere:
sequenza *no-random* ***x*** di messaggio ***m*** = *30 bit*
sequenza *no-random* ***x'*** di messaggio ***m'*** = *30 bit*
diremmo che la lunghezza cumulata di tali sequenze,
è maggiore di cinquanta che poi sarebbe la misura da noi convenzionalmente attribuita alla stringa random ***y*** di chiave ***k*** essendo 30 +30 = 60 > 50

Esempio in Numeri Binari

Esempio di sequenza di numeri random ***y*** formata da *50 bit* che funge da chiave crittografica ***k***
100100110011010011001101001100110100110010100111

Esempio di sequenza *no-random* ***x*** formata da *30 bit* che funge da messaggio ***m***
011010011001101001100110100110

[14] Più precisamente, sia che si faccia riferimento alla definizione tradizionale (**A**), sia che si faccia riferimento alla formulazione da noi offerta (**B**)(**C**),
a rigore avremmo che nei sistemi perfetti un segnale random potrebbe pur essere impiegato più d'una volta, ma a patto che la sua lunghezza resti maggiore di quella che si avrebbe concatenando le sequenze *no-random*.

Esempio di sequenza *no-random* ***x'*** parimente formata da *30 bit* che funge da messaggio ***m'***

111010011001101001100110100111

Concatenazione di maggior lunghezza binaria data dalle sequenze ***x, x'*** dei messaggi ***m***, ***m'***

011010011001101001100110100110111010011001101001100110100111

Motivo per cui si crea uno stato che non soddisfa quanto richiesto in (**A**) ma nemmeno in (**B**)(**C**) giacché il flusso aleatorio sarà di lunghezza minore rispetto a quello della concatenazione dei messaggi *no-random*, dove lo stesso vale andando a invertire l'impiego di ***x*** con quello di ***y*** qualora fosse tale ***x*** a fungere da chiave ed ***y*** da messaggio.

Capitolo II

.

Tema

(07)

Proseguendo da dove ci eravamo lasciati, si fissano le condizioni per superare le criticità emerse dalla lettera di corollario (**D**)

Allo scopo si inizia pertanto a inquadrare il senso d'alcuni passi di trasformazione che diremo di *digestion* da illustrare qui e nei prossimi capitoli alla luce dei concetti di "funzione unidirezionale" e "impredicibilità". In tale prospettiva, si dirà altresì dello XOR lemma di Yao, prima di lanciare alcune suggestioni che torneranno utili più innanzi.

Nel complesso, possiamo dire che il secondo capitolo va principalmente inteso come un capitolo introduttivo rispetto a quelli che verranno.

Trasformazioni

(08)

Per quanto detto nel precedente capitolo, abbiamo dunque che … se sfruttando le opportunità offerte dalla lettera dei lemmi (**B**)(**C**) una qualche utenza dovesse ricevere in forma di messaggio un segnale random ***y*** da noi detto greggio … assurdamente non potrebbe farne uso per trarre da esso chiavi crittografiche.
Assumendo infatti tale ***y*** per poterla adoperare prima come messaggio e poi come chiave, si verrebbe a creare un conflitto con quanto enunciato in **(D)**;
è infatti facile prevedere che dal confronto analitico tra il crittogramma che cela la sequenza random, e quello che dissimula il messaggio finale, ci sia modo di risalire al segnale che hanno in comune seppure nello svolgimento di diverse mansioni.
E appunto per questo, sequenza ***y*** giunta a utenza $\mathbf{U}_i$ non potrà essere impiegata per creare chiavi crittografiche da ingaggiare *tout court* in fase di codifica;
perché l'ingaggio abbia luogo senza abbattere la sicurezza del sistema, tale sequenza dovrà essere infatti sottoposta a operazioni che la rendano altro da sé, così da creare nuove e diverse sequenze che rispondano in forma vergine ai seguenti requisiti:

(1) **non abbiano memoria del segnale greggio dal quale discendono**;

(2) espongano un delta percentuale di differenza da questi, tale da essere indistinguibili da ogni altra e diversa sequenza presa a caso (presa con pari distribuzione di probabilità da un insieme finito di possibili stringhe);

(3) **preservino la natura random mantenendo una pari frequenza di comparsa**, continuando a rispondere per un tempo tendenzialmente infinito agli standard dei principali enti di controllo del segnale.

A tali passi di post-processing, abbiamo perciò dato il nome di *digestion* anche per figurare una procedura con qualche vaga ma significativa assonanza con quella delle funzioni di hash[15] talora impiegate in ambito crittografico;

[15] E' chiaro che il parallelo proposto è principalmente una suggestione.
Tant'è che la prima netta differenza tra qualsivoglia funzione di hash e le funzioni di trasformazione di cui abbiamo detto, sta nelle opposte finalità.
Nell'un caso si tratta di creare un'impronta del documento originale, nell'altro di dar vita ad una sequenza che, una volta conosciuta, non sia riconducibile all'originale di cui non serba memoria.
Altra differenza è che le funzioni di *hash* rilasciano un output di lunghezza data, indipendente dalla lunghezza del documento preso in accesso, mentre per le nostre funzioni di trasformazione vale esattamente il contrario.
Pur tuttavia esse hanno in comune lo spacchettamento (hash) della sequenza originale e il fatto per il quale la quantità di informazione assunta in ingresso è maggiore di quella in uscita.
Per converso, una diversità attiene alla natura dei dati in ingresso.
Da una parte si tratta di dati relativi a un documento intellegibile e quindi *no-random*, dall'altra si tratta di dati aleatori generati da una sorgente true random.
La diversità è ricca di conseguenze in quanto, nelle funzioni "hash crittografico" la pur trascurabile facoltà di risalire al documento originale, è resa teoricamente possibile dal fatto che la stragrande maggioranza delle cosiddette

a prescindere tuttavia da come si possa formulare tale processo definito da una famiglia di trasformazioni che conducono da uno spazio a altro spazio,

esso tratterà ***y*** di variabile ***Y*** come "materia bruta" da cui trarre forme compresse, che siano prive di correlazioni statistiche col segnale da cui pure discendono.

In altre parole, sequenze variamente impiegate come chiavi di crittazione (tra cui saranno anche presenti delle particolari stringhe[16] che altrove diremo come impiegare in qualità di chiavi di altre chiavi e altro ancora),

non dovranno fornire alcun indizio sul segnale assunto in partenza, giacché ogni nesso tra sequenza e sequenza sarà reciso (21); **esse rinasceranno a nuova vita essendo che non esisterà modo di poterle condurre allo stato originale**.

Figura 02 - Diagramma

"collisioni" finiscono col mettere l'attaccante davanti a un gran numero di documenti da scartare in quanto privi di senso compiuto.

Un qualche discernimento tra documento e documento è dunque possibile e, assunta la funzione $f\ \mathrm{H}(\mathrm{D}) = \boldsymbol{d}$, dove D simboleggia l'insieme dei possibili *input* (documenti in ingresso) e ***d*** l'output che ne discende,

laddove si sia venuti a conoscenza di quanto vale ***d*** (*digest*), gli oggetti appartenenti a D non sarebbero tra loro equi-probabili in quanto non tutti dotati di senso compiuto.

A scanso di equivoci, sarà bene precisare che qui non facciamo riferimento al caso di scuola dell'attaccante che cerchi una collisione tra testi dissimili, così da mettere in dubbio la traccia rilasciata dal *digest*.

Siamo invece nella diversa ipotesi di un attaccante che voglia forzare il *digest* per entrare in possesso del documento originale o perlomeno per andarci vicino.

Nelle funzioni di trasformazione di cui diremo, le cose, da questo punto di vista, sono tuttavia diverse anche per la natura random (non intellegibile) dell'input preso in ingresso.

Se assumiamo infatti la funzione di trasformazione $f\ \mathrm{T}(\mathrm{Y}) = \boldsymbol{j}$, dove Y è rappresentativo di tutti i valori che può assume una sequenza aleatoria di lunghezza *L*, conoscendo il valore in uscita $\boldsymbol{j}$

nemmeno in teoria, nemmeno tramite attacchi a forza bruta, s'avrebbe modo di risalire alla sequenza presa in ingresso giacché parimente probabile rispetto alle altre.

[16] Nel documento cui accenniamo nel paragrafo intitolato al "Metodo".

One-Way

(09)

Prima di volerci inoltrare nel racconto, è però bene darci una pausa anticipando che *d'ora innanzi* da un canto parleremo sempre e solo di forme binarie,
e dall'altro andremo a maneggiare concetti fortemente correlati come appunto sarebbero quelli di "incertezza" "memoria" e "impredicibilità" dei quali diremo.
Ci sono infatti dei casi che non si possono trattare separatamente giacché dipendono gli uni dagli altri;
se ad esempio diciamo che un flusso è fortemente **i**mpredicibile, tanto comporta che intercettati *n*-1 bit del medesimo, non saremo in grado d'indovinare l'ennesimo.
Ciò accade quando i dati raccolti rilascino poca o nessuna informazione avendo un andamento aleatorio tale da impedire ogni pronostico;
ma rimarcare che non danno informazioni è un diverso modo di dire che non hanno memoria, giacché ogni traccia è andata smarrita; d'altra parte una sequenza impredicibile nel suo andamento e quindi senza memoria è estremamente incerta e quindi difficile da invertire.
Per poterci intendere sarà tuttavia opportuno fornire delle indicazioni, che faranno da garbata cornice ai passi di trasformazione (*digestion*) da noi più innanzi fissati per soddisfare i requisiti richiesti nel precedente paragrafo.
Intendiamo cioè dire che per superare le criticità poste in (**D**) appare logico assumere come milestone quella di creare sequenze vergini da cui sia impossibile risalire a quelle gregge (in realtà ci siamo prefissi qualcosa di più radicale, quando diciamo che tali sequenze non avranno memoria delle occorrenze da cui pure provengono),
ma tale finalità non può prescindere da uno sguardo panoramico sulla natura di quelle funzioni one-way o unidirezionali che consentono di perseguire siffatto obiettivo, rendendo ardua la risalita dai valori in uscita a quelli in accesso.

Secondo la definizione standard, una funzione unidirezionale $f: \{0, 1\}^* \rightarrow \{0, 1\}^*$ è dunque una funzione facile da calcolare (come tra l'altro sarebbe una operazione di somma XOR) ma quasi impossibile da invertire.

Vale a dire che, per quanto certamente esista un algoritmo ***a*** per il calcolo in tempo polinomiale di f, avremo che per qualsiasi algoritmo ***A*** che a sua volta operi in tempo polinomiale, la facoltà di risalire ai valori in entrata è pressappoco nulla per cui si dice che la probabilità che ***A*** (f(***x***)) sia pre-immagine di f(***x***) è trascurabile.

Di tale dettato desideriamo mettere tuttavia in luce alcuni aspetti che assumono un particolare sapore entro il *framework* in cui abbiamo inquadrata la questione del segnale da cui trarre chiavi crittografiche;
occorre infatti rimarcare che spesso si dice che non tutte le funzioni unidirezionali sono interessanti, ma la verità è che molto dipende dall'uso cui sono destinate.

Prendendo ad esempio un predicato booleano facile da calcolare ma difficile da invertire, tra i più banali troveremo quelli detti con "perdita d'informazione" che pur essendo poco attraenti, sono tuttavia assai coriacei.

In *Lecture Notes on Cryptography* di *Goldwasser* e *Bellare* abbiamo ad esempio, una semplicissima funzione f(x)f(x) definita in *SHA-256* e quindi all'interno della omonima famiglia di funzioni crittografiche[17], dove avremmo in uscita una sequenza numerica yy identica a quella in accesso di xx se non fosse per la materiale rimozione del primo bit della stessa, per cui il valore che potremmo ironicamente definire come "primo bit di xx" non sarà per definizione computabile a partire da tale f(x)f(x)

In un certo senso si può notare come proprio le operazioni più rustiche dove alcuni dati sono banalmente smarriti, siano impossibili da invertire benché credute poco utili per il fatto che – perdendo informazione e non essendo iniettive – offrono output con tali elementi di indecidibilità da impedire la risalita persino a chi legittimamente in possesso della chiave.

In proposito è bene tuttavia precisare che pure le operazioni XOR impiegate in fase di *digestion* determinano (se prese come legge tra predicati indipendenti) un risultato compresso e quindi alterato[18] dal quale sarà impossibile risalire tanto che in talune condizioni (30) i valori in uscita non serberanno memoria di quelli in entrata.

(10)

Intanto un lemma assai citato in letteratura pure perché circonfuso da un alone di mistero, noto come XOR Lemma di Yao[19] (sul quale torneremo nel prossimo capitolo così da poterne tastare alcune conseguenze) ci viene in duplice soccorso sia riguardo alla crescita di aleatorietà del segnale, sia riguardo alla perdita di memoria dello stesso.

Muovendo infatti dall'ambito delle funzioni *one-way*, nel lemma si afferma che la weakly unpredictability dei predicati booleani è amplificata quando i risultati di più istanze indipendenti (x_1, x_2, ..., x_t) siano sommati XOR tra loro.

- Dove per predicato booleano è intesa una frase *qui indicata in forma di funzione f(x)* per cui preso un input come sarebbe quello dì una sequenza binaria, è rilasciato in uscita un unico bit che dia zero o uno[20];
- Dove per "debole impredicibilità del predicato*"* si intende il fatto per cui entro una correlata soglia di complessità,

[17] SHA - *Secure Hash Algoritms.*

[18] Compresso con perdita di informazione.

[19] Lo XOR lemma di Yao è stato enunciato per la prima volta in forma orale dal medesimo autore nel corso della presentazione al suo lavoro *Theory and Applications of Trapdoor Functions* .

[20] In pratica se pure la sequenza presa in input fosse di *n* bit, il risultato in uscita sarà comunque dato da un unico valore di *zero* o *uno.*

un algoritmo efficiente non sarebbe in grado di predire il valore in uscita da *f(x)* con probabilità fissata oltre un limite dichiarato (dove tale probabilità è presa con uniforme distribuzione su ogni possibile input);

- Dove per "risultati di più istanze indipendenti" sono intese le uscite (discendenti da distinti input in accesso) date in pasto a una diversa funzione di sommatoria XOR fortemente impredicibile che diremo $F(x_1, ..., x_t)$ o più semplicemente F.

In modo più formale si può anche dire che, se predicato *f(x)* è debolmente impredicibile e quindi computabile entro la correlata soglia di complessità,

per t sufficientemente grande dove t indica il numero di istanze indipendenti da sommare XOR tra loro, che abbiam dette $x = x_1, x_2, ..., x_t$, essendo $x \in \{0, 1\}^n$

tale sommatoria $F(x_1, ..., x_t)$ dove $F(x_1, ..., x_t) \stackrel{\text{def}}{=} \bigoplus_{i=1}^{t} f(\boldsymbol{x}_i)$ come supposto diventa invece quasi impredicibile entro la stessa soglia[21] nel senso che algoritmi efficienti non potrebbero dare altri che risposte aleatorie.

Il che nel recinto dei nostri intendimenti, può anche significare che dal momento in cui gli esiti rilasciati da F abbiano conseguito tale forte impredicibilità,

volendoci muovere nello spazio del nostro ambito più circoscritto, potremmo definire una grandezza $\boldsymbol{h}$ che a partire dalla *t-esima* istanza, cresca al crescere del numero dei risultati che si hanno implementando la sommatoria, essendo $\boldsymbol{h}$ negativamente correlata alla probabilità composta di indovinare la successione di valori che avremmo in uscita.

Per cui a partire da t inteso come valore di riferimento, posto dal preciso momento in cui ciascuna uscita diviene parimente probabile, potremmo pur scrivere $\boldsymbol{h}^{(z)}(\boldsymbol{ex}_t, ..., \boldsymbol{ex}_z) \stackrel{\text{def}}{=} \prod_{k=t}^{z} (\boldsymbol{ex}_k)\, a$ per ogni $z \geq t$

dove a sarebbe un fattore di probabilità che diamo per mediamente spalmato su tutti i termini della successione $\boldsymbol{ex}_t, \boldsymbol{ex}_{t+1}..., \boldsymbol{ex}_{z-1}, \boldsymbol{ex}_z$

in ragione dei quali, $\boldsymbol{h}$ crescerà al crescere del numero di uscite $\boldsymbol{ex}_t, \boldsymbol{ex}_{t+1}..., \boldsymbol{ex}_{z-1}, \boldsymbol{ex}_z$ che somma dopo somma saranno rilasciate dallo XOR, come sarebbe per il reiterato lancio d'una moneta che talora dia testa e talaltra croce[22]

In proposito, si potrà comunque specificare il fatto che $a = 2(2p)$

dove $\boldsymbol{ex}_t, ..., \boldsymbol{ex}_z$ giustappunto sarebbero le uscite dalla sommatoria (prese somma dopo somma invece che all'esito finale)

dove **2** sono gli elementi dell'alfabeto binario,

ma **2** sono anche le opzioni che esso alfabeto effettivamente concede ponendoci noi nei panni d'un indovino,

[21] A cambiare non sarebbe la soglia di complessità ma i limiti di probabilità che consentono di predire o meno la giusta soluzione.

[22] Qui non si considera il solo output finale, ma gli output che si possono avere concatenazione dopo concatenazione.

mentre *p* sarebbe la probabilità (pari ad un mezzo su uno, in caso di bit equiprobabili) che si verifichi la comparsa del simbolo annunciato dove $0 \leq p \leq 1$

Ciò detto per maggior prudenza giacché la prudenza non è mai troppa ..., se mai lo potessimo dimenticare è opportuno tenere a mente che stiamo sempre ragionando di un predicato booleano il quale, da una parte rilascia per definizione per ciascun impulso, che può pur essere un impulso dato da una successione di bit …, un unico valore 0 o 1 in uscita,
e dall'altra *qualora assuma più impulsi* andrebbe a consegnare altrettanti output a predicato *F* che li converte in input che a voler operare come usualmente concepito, a loro volta rilasciano il responso di un solo bit fortemente impredicibile.

Da ciò possiamo dunque intuire che, crescendo a monte l'impredicibilità e quindi l'aleatorietà dell'output rilasciato da operazioni condotte su istanze indipendenti**,** non può che decrescere a valle la memoria degli input presi in ingresso.
Sebbene ciò non si possa generalizzare fuori dal contesto indicato, si può generalizzare quando si abbia perdita d'informazione come quando *pescando da ripetute collezioni numeriche* somme su somme siano reiterate un numero arbitrario di volte[23];
nondimeno in una funzione difficile da predire e altrettanto difficile da invertire come sarebbe funzione *F,* per quanto la memoria statistica si possa attenuare non tutto è tecnicamente perduto motivo per cui sarebbe bene dichiarare in anticipo quali siano i dati sensibili che si intendono offuscare e quali non lo sono.

[23] Il concetto di "impredicibilità" è stato qui trattato a volo radente per cui è bene ricapitolare le differenti prospettive da cui siamo partiti.

Intanto si è definita la "debole impredicibilità" o weakly unpredictability di un predicato booleano che si ha quando un qualsiasi algoritmo efficiente non sia in grado di predire il valore in uscita con probabilità fissata oltre un limite dichiarato.

Ciò, di converso, significa che entro tale limite, esso algoritmo sarebbe capace di predire il risultato e tanto qualifica la lieve barriera opposta da una funzione "debolmente impredicibile" a una famiglia di circuiti di sufficiente complessità.

Si è poi alluso al diverso concetto di "forte impredicibilità" o hard unpredictability che si ha quando la probabilità di predire un valore in uscita non sia peggiore di quella che si avrebbe per il fortuito lancio di una moneta.

Proprio l'idea di forte impredicibilità e quella di funzione hard core (idea più volte sottintesa) ci hanno quindi condotti a fissare una grandezza $\boldsymbol{h}$ utile ai nostri fini.

Tale grandezza è negativamente correlata alla probabilità composta di indovinare consecutivamente più risultati dall'andamento aleatorio.

In concreto, essa si applica a più istanze riferibili a un medesimo predicato *f* debolmente impredicibile nel momento in cui siano assunte come input di una funzione di sommatoria *F* che sappiamo fortemente impredicibile.

Un ultimo rilievo riguarda poi il nesso tra "aleatorietà" e "mancanza di memoria" che trova un'eccezione nella cosiddetta "estrazione senza remissione" che tuttavia implica un set finito di possibili risultati (come per delle figure prese da un mazzo di carte); nel nostro caso al contrario non ci sono di tali vincoli giacché è contemplato che i valori tanto in entrata che in uscita si possano ripetere.

In effetti i valori da sommare sarebbero perciò pescati da quello che potremmo definire come un multinsieme tendenzialmente infinito.

Infine precisiamo, che quando parliamo di "somma" qui intendiamo sempre e solo somme XOR.

Sarà in questo senso che si potrà parlare di perdita di memoria in quelli che diciamo file vergini in contrapposizione con quelli greggi,
dove ciò avrà notevoli conseguenze sulla facoltà di soddisfare le condizioni richieste per superare le criticità poste in (**D**)**.**

Forza Bruta

(11)

Un'altra tessera da aggiungere al quadro da definire, attiene tuttavia a un aspetto che si collega ad una delle principali conseguenze pratiche delle speculazioni sulle funzioni hard-core e su quelle unidirezionali.

Fin quando infatti parliamo di algoritmi che operano in tempo polinomiale e di limiti di complessità, o di funzioni quasi impredicibili, stiamo in buona sostanza sottolineando come ogni impredicibilità abbia un limite, nel senso che vale fin quando non si abbia il caso di un attaccante dotato di risorse sufficienti se non infinite.

Questo può essere o non essere un problema nella stragrande maggioranza dei casi (dove ci si trova davanti a criticità più o meno gestibili) ma le cose cambiano qualora si dichiari di agire nell'ambito di sistemi a sicurezza perfetta.

Occorre infatti ricordare che in letteratura, il campo d'applicazione di tali sistemi si inscrive nel più generale panorama delle probabilità incondizionate;

assegnate infatti *w, z* si dirà che la probabilità di giungere a quest'ultima è incondizionata se p ($z \mid w$) = p (z) e cioè se la probabilità di risalire a qualche valore numerico *z* di variabile *Z* non varia sapendo quanto vale *w* di variabile *W*.

Ora se la facoltà di risalire a un'informazione come sarebbe quella di *z* non dipende da quanto vale *w*, ciò necessariamente comporta che non esiste alcuna potenza sufficientemente grande da rilevare una correlazione tra l'uno e l'altro valore numerico.

Questa è prima di tutto una sorta di provocazione, ma se stimiamo un insieme di sequenze generate da sorgenti aleatorie che sappiamo entrare in gioco (per lunghezze importanti) quando trattiamo di sicurezza perfetta, stiamo dicendo di ambiti dove non esiste algoritmo abbastanza efficiente da calcolare i flussi alla fonte.

Motivo per cui, nei sistemi perfetti, abbiamo che l'ipotesi di un attaccante dotato di risorse infinite, sia una necessità rappresentando la regola e non l'eccezione[24].

(12)

Ma vediamo come le cose possano adesso diversamente apparire, modificando di poco la nostra prospettiva, e cioè chiedendoci non tanto cosa significhi ma che comporti il fatto che talune funzioni siano "quasi" impredicibili entro una qualche soglia di complessità.

[24] E questi infatti lo stratagemma attraverso cui si vuole escludere che il messaggio – invece di essere protetto dalle difese di un sistema perfetto – sia più semplicemente tenuto indenne dalla insufficiente capacità computazionale di chi attacca.

Se ragioniamo infatti in modo informale parrebbe facile rispondere sapendo che quel "quasi" comporta che ogni impredicibilità proceda all'infinito giacché esisterà sempre un limite che si sposta in avanti e una complessità di calcolo che consente di raggiungerlo attraverso una macchina ancora più potente di quella di prima;

tanto che, se prendiamo le mosse dai risultati rilasciati da una qualche funzione difficile da invertire come quella di *F*

avremo che una volta conseguita per *t* sufficientemente grande, la condizione per cui ogni risultato XOR sia parimente probabile,

$\boldsymbol{h}$ da noi data come misura quadratica dell'incertezza di più somme cumulate di qualsivoglia numero di istanze maggiore di *t*, e quindi come reciproco della probabilità composta di indovinare gli input XOR presi a partire dal punto di loro massima incertezza,

risulterà crescente al procedere della sommatoria,

essendo $\boldsymbol{h}\{(ex_{z-t}) \oplus (\dots) \oplus (ex_{z-1})\} < \boldsymbol{h}\{(ex_{z-t}) \oplus (\dots) \oplus (ex_{z})\}$ *per ogni* $z > t$

In altre parole si può anche dire che, se dopo una qualche quantità *t* di somme XOR, ciascun esito binario diviene impredicibile come per il lancio a mezz'aria della fatidica monetina,

tale grandezza cumulata crescerà esponenzialmente al crescere di $z > t$ ma senza mai diventare infinita[25].

Ciò sembra pertanto indicare che fin quando si operi su grandezze finite,

la ricerca dei valori di input-output potrà aver sempre luogo con uno screening (tanto più impegnativo quanto più grande sia *z*) che diligentemente mappi ogni possibilità di giungere al risultato.

In altre parole si direbbe che persino quando dai valori raccolti non si abbia alcuna informazione su quelli da calcolare, una manovra a forza bruta condotta con un dispiegamento di forze sufficienti se non infinite, sarebbe coronata da successo.

Per mostrare il contrario, vediamo tuttavia di offrire un caso preso in prestito dal mondo delle favole, che contraddice nei fatti quanto detto, seguendo per intanto il proverbio secondo cui un esempio vale più di mille parole.

[25] Nel caso infatti avremmo $F(x_1, \dots, x_t, \dots, x_z)$ per *z* numero finito.

Abracadabra

(13)

Nell'abbozzare una replica alla domanda che ci siamo fatti da soli, per la quale desideriamo sapere fin quando sia possibile preservare o invece forzare ogni memoria con un attacco a forza bruta ..., mettendoci nei panni di chi offende che qui chiameremo Oracolo (e che Shannon chiamava invece nemico) diciamo che muovendo dalla sua prospettiva avremmo il vantaggio di contare su un'informazione particolarmente ostinata.

Un'informazione che permane in qualsivoglia bit-sequenza, anche quando "quasi" non abbia memoria.

La conoscenza da parte di chi attacca degli algoritmi impiegati per oscurare un segnale e il segnale medesimo, non possono infatti non fornire qualche indizio sulla misura e quindi sulla profondità degli input in accesso e tale informazione appare ineliminabile giacché può essere elusa ma mai cancellata [26].

Poniamo allora che nella favoletta che intendiamo narrare, un Oracolo in possesso di risorse sovrumane intenda indovinare una formula magica che ha la facoltà di aprire un forziere, ma che proprio per questo è stata coperta attraverso tali inganni da risultare illeggibile.

Egli è riuscito a intercettare la corsa di un araldo che teneva in custodia tale messaggio segreto, ma essendosi impossessato della missiva, non potette far altri che confessare a se stesso d'essersi impadronito d'uno scritto insensato;

uno scritto che gli avrebbe tuttavia consentito di congetturare sulla lunghezza del messaggio e quindi sul numero dei caratteri dell'abracadabra.

Tenendo infatti conto di tale ipotetica lunghezza o meglio della massima lunghezza L che poteva attribuire al messaggio oscurato, egli sarebbe stato in grado di provare tutte le $\boldsymbol{n}^{m}$ combinazioni del medesimo, declamando ciascuna possibile formula fin quando non avesse aperto il forziere che nascondeva al suo interno il tesoro.

Non c'è dubbio che avendo a disposizione forze inumane, non solo sarebbe riuscito a venire a capo del problema ma sarebbe riuscito a farlo all'istante essendo capace di pronunciare in un lampo ogni ipotetica invocazione.

La narrazione può tuttavia avere un diverso epilogo ogni qual volta la formula non dovesse andare a buon fine, come quando il chiavistello risulti bloccato o nel caso in cui pioggia e intemperie abbiano arrugginito i battenti.

Qualora per disdetta si verifichi un tale accidenti, il forziere continuerà a non dar segni di vita nemmeno davanti a tutte le formule magiche del mondo che non avrebbero effetto ... e per tale ragione l'oracolo non

[26] In effetti sarebbe bene chiarire di quali input diciamo, sia nel senso che non tutti sono interessanti per un attaccante e sia nel senso che non sempre rilasciano utili indizi.

sarebbe in grado di discernere tra la formula esatta e quelle sbagliate essendo state tutte parimente enunciate nel tentativo di trovare il tesoro.

Ciò che a noi preme non è tuttavia la poca o molta fortuna che arride all'Oracolo, quanto il fatto che non sia in grado di trovare la giusta formula ..., fin quando non venga scoperta in ragione del compiersi di qualche evento rivelatore (qualche evento che la possa tradire, come il successo della combinazione che spalanchi il forziere).

Come avremo infatti modo di vedere non mancano casi in crittografia dove ogni riscontro resta impossibile quando si operi su sequenze equi-probabili o in tutti i casi in cui non sia possibile individuare un lotto sufficientemente ristretto di possibili soluzioni.

In fin dei conti "avere memoria" altri non significa che mantenere traccia di qualche informazione che ci consenta di distinguere questo da quello.

Una Parentesi: ***Cenni sullo Stato dell'Arte*** [27]

(14)

Le difficoltà che si incontrano nel redigere (e forse anche nell'intendere) un documento come quello che stiamo trattando, molto probabilmente non stanno in questo o quell'altro passaggio logico, ma nello sforzo che richiede il costante esercizio d'una visione d'insieme.

Ora nelle righe a seguire, vorremo perciò mettere informalmente in rilievo il senso da dare al lavoro compiuto e da compiere ancora, rispetto allo stato dell'arte ed agli effetti pratici che possono discendere dalle ipotesi da noi formulate.

E per questo sarà bene ricordare il significato assunto nel mondo reale, sia dal concetto di "segretezza computazionale" che da quello di "sicurezza perfetta" giacché si tratta di due poli intorno a cui gira ogni nostra considerazione.

Sicurezza Computazionale

(15)

Tanto per cominciare, diciamo che fu sempre l'attivismo di Shannon a portare alla luce la questione della differenza tra l'inviolabilità assoluta dei cosiddetti sistemi perfetti, e quella che lui chiamava sicurezza pratica, ed a cui fu successivamente attribuito il nome di sicurezza computazionale.

Se vogliamo eludere talune sottigliezze e se guardiamo alle cose per come stanno per davvero a livello industriale e commerciale, intendiamo per sicurezza computazionale quella che attiene a tutti i sistemi attualmente in uso;

essa è violabile per definizione, sebbene entro determinati limiti di tempo e probabilità di successo, ma secondo i suoi fautori comporta un tal dispiego di risorse da rendere inutile per chi attacca ogni tentativo di tornare dal crittogramma al messaggio.

All'inizio della seconda guerra mondiale, i critto-analisti al servizio di Sua Maestà Britannica potevano contare su una messe di messaggi regolarmente intercettati sulle frequenze radio, che non avevano tuttavia il tempo materiale di decifrare pur essendo in possesso di sofisticati metodi di risalita.

Si tratta dell'epica storia della macchina Enigma che, al momento dello scoppio del secondo conflitto mondiale, risultava imbattibile in considerazione del fatto che una puntuale attività di decodifica avrebbe giustappunto richiesto risorse esorbitanti per calcolare nei secoli tutte le possibili combinazioni di simboli orto-alfabetici.

[27] Il presente scritto è tratto dalla presentazione a un report intitolato *Secrecy Systems* redatto da Claudio Cappelli nel 2020 e circolato in via riservata negli anni e nei mesi a seguire.

Ci sembra che le considerazioni qui riportate, si possano adattare alla lettera del corrente saggio.

Esse riguardano le conseguenze concrete e non gli aspetti teorici, delle scelte che comportano i diversi approcci alla sicurezza crittografica.

Invero la vicenda molto cinematografica di Enigma è tutt'altro che isolata, essendo esemplificativa di un destino che accomuna, senza eccezioni, tutti i sistemi di crittazione che abbiano adottata la prassi della cosiddetta sicurezza computazionale.
Da un certo momento in avanti, Enigma cominciò ad esser infatti bucata con tale regolarità che i messaggi trasmessi in campo nazista finirono col risultare come trasparenti alle forze alleate.
La questione può esser dunque generalizzata a partire dal seguente punto di domanda: cosa mai accomuna sistemi tanto diversi come ad esempio sarebbe Enigma medesima, l'antico cifrario di Cesare, quello di Vigenere e quello noto col nome di DES e selezionato nel 1976 come standard crittografico per il Governo degli Stati Uniti d'America?

Da una parte si può rispondere che essi sono uniti dal fatto d'essere parimente protetti dalla difficoltà del problema con cui sfidano chi attacca, ma ciò è scontato sapendo che nei fatti ciascun modello in uso è appartenuto o appartiene a tale ampia categoria crittografica.
Ciò nondimeno c'è un più sottile denominatore che mette assieme tutti i casi annoverati, giacché gli esempi prodotti sono stati tutti dei campioni invincibili in fatto di sicurezza, fin quando nuovi metodi di critto-analisi non li hanno abbattuti.
Nel gennaio 1999 alcune organizzazioni quali quelle della Distributed.net e della Electronic Frontier Foundation collaborarono per rompere in pubblico una chiave crittografica del DES, e ci riuscirono in 22 ore e 15 minuti.
Viceversa il Codice di Vigenere fu reso vulnerabile dall'approccio segretamente escogitato da Charles Babbage, pur continuando ad esser impiegato per decenni fin quando il colonnello Fredrich Kasinsky non ruppe l'incantesimo rendendo pubblico un suo metodo di critto-analisi cui era distintamente giunto nel 1883);è un fatto tuttavia che s'è ormai venuti a capo di qualsivoglia cifrario adotti il criterio della sicurezza computazionale da prima del 1990.
Ed è questa la vera criticità di tali sistemi; essa non riguarda l'offesa al singolo messaggio, ma la probabilità – pressoché certa – che prima o poi si trovi un metodo che consenta di forarli in modo sistematico, laddove la cosa può accadere tra un giorno o cinquant'anni, sempre che non sia già segretamente avvenuta senza che se ne abbia contezza[28].

[28] Anche gli algoritmi a cavallo tra vecchio e nuovo millennio lasciano intravedere delle crepe laddove un attacco ha forzato l'AES con una chiave a 256 bit e nove round (*Ferguson e altri, 2000*).
Persino il sistema a chiave asimmetrica RSA è senz'altro minacciato dall'algoritmo di Shor e dall'avvento d'una nuova generazione di computer quantistici.

Sicurezza o Segretezza Perfetta

(16)

Completamente diversi sono i principi su cui si fonda quella usualmente nota come sicurezza perfetta o "incondizionata".

Se la difesa di un sistema crittografico è spesso data dalla complessità computazionale opposta come un macigno agli sforzi di chi attacca, la sicurezza perfetta trova il suo fondamento nella numerosità dell'insieme d'ogni equiprobabile chiave (spazio della chiave) che dovrà essere ugual-maggiore rispetto alla numerosità dell'insieme dei possibili messaggi in chiaro[29]*.*

Le stesse condizioni richieste in letteratura per rientrare nell'ambito della perfetta segretezza (una chiave random di lunghezza ugual maggiore rispetto a quella del messaggio, ecc.) altri non sono che espedienti per giungere a tale risultato.

Il quadro teorico descritto per la prima volta da Shannon – ancorché meno esauriente di quanto si creda – è comunque corretto entro determinati limiti fattuali, illustrando una situazione nella quale ciascun valore del crittogramma intercettato da chi attacca, poco aggiunge alla probabilità di risalire a chiave e messaggio per il fatto che:

- **(1)** *qualsivoglia sequenza binaria appartenente a quello comunemente battezzato col nome di "spazio della chiave" risulta compatibile col crittogramma effettivamente ottenuto in uscita;*
- **(2)** *e ogni sequenza appartenente allo spazio dell'insieme dei messaggi, risulta a sua volta in accordo col medesimo.*

Sarà in tal senso che chi attacca non aggiunge nuova informazione a quella che aveva, essendo che tutti i messaggi calcolabili "a priori" e cioè prima d'averli potuti captare in forma cifrata, seppure nei limiti di cui diremo, restano parimente probabili "a posteriori" e cioè dopo che tale attività di disturbo abbia avuto luogo.

E' dunque certo che a guardare le cose alla luce della segretezza del messaggio, tra sicurezza computazionale e sicurezza perfetta non c'è storia che tenga (come si direbbe in inglese, tali distinti approcci non giocano nello stesso campionato).

E se è risaputo che l'unico cifrario che ufficialmente adotta tale pratica – noto come cifrario di Vernam o anche come One-Time Pad – non ha conosciuto una diffusione commerciale, è altrettanto vero che costituisce un punto di riferimento per la molto pubblicizzata (e altrettanto fraintesa) distribuzione quantistica e che è stato impiegato in circostanze dove la segretezza era una priorità imprescindibile.

[29] Potremmo anche dire che se, metaforicamente, la sicurezza computazionale è una montagna difficile da scalare, la sicurezza perfetta è un mare impossibile da prosciugare.

Sul concetto di "sicurezza perfetta" torneremo tuttavia più volte, anche modificando il nostro punto di vista così da seguire la questione da differenti prospettive.

In verità è proprio il caso più clamoroso di utilizzo d'un sistema perfetto, a dirci parecchio …, tanto degli scenari che possono essere aperti da tale soluzione crittografica, quanto dei limiti che ne hanno frenato lo sviluppo.

Negli anni sessanta del secolo scorso – dopo la crisi dei missili a Cuba acrobaticamente gestita da una parte all'altra del pianeta – USA e URSS decisero di attivare una linea diretta che avrebbe tenuto in collegamento il Presidente degli Stati Uniti d'America al Segretario Generale del Partito Comunista Sovietico.

Su tale linea, battezzata col nome di telefono rosso, le trasmissioni furono appunto cifrate in One Time Pad (espressione che fa riferimento al fatto che ciascuna chiave non sarà impiegata più d'una volta, essendo che one-time significa "una volta" mentre pad sarebbe il taccuino dove veniva effettivamente trascritto il valore della chiave di crittazione) e ciò dimostra che più si alza l'asticella e più cresce il bisogno di segretezza.

D'altra parte per garantire tali impenetrabili trasmissioni occorreva scambiarsi lunghe chiavi condotte oltre oceano in valigette diplomatiche, sorvegliate dai Servizi delle due Super-Potenze, e tradotte nei caveau di mezzi blindati, e ciò ben fotografa quanto fosse complicato risolvere il problema della distribuzione sicura di chiavi della stessa misura di un messaggio che non era stato ancora concepito e di cui nessuno poteva anticipatamente conoscere la lunghezza.

In effetti, nella formulazione attuale, i sistemi perfetti lasciano scoperto il fianco al fondamentale problema della messa in sicurezza delle chiavi.

Sebbene le obiezioni all'uso della cifratura a segretezza o sicurezza perfetta riguardino anche altro, il problema che ne frena l'uso attiene quasi esclusivamente alla questione delle chiavi;

risolto il busillis della sicurezza dei messaggi, appare nondimeno necessario replicare il successo con le chiavi medesime.

Distribuzione Sicura delle Chiavi [30]

(17)

Pochi lo ricordano e chi lo ricorda non è più giovanissimo, ma nel 1994 uscì nelle sale il film l'Uomo Ombra diretto da Russell Mulcahy, ispirato a un programma radiofonico dedicato agli amanti del Brivido. Secondo la trama, l'ultimo discendente dell'imperatore Gengis Khan alloggiava a New York al Monolith Hotel ma la sua lussuosa dimora appariva introvabile essendo stata resa invisibile agli occhi dei cittadini sedotti dalle sue facoltà ipnotiche.

Dobbiamo dire che quando ci troviamo a mettere in luce le incompletezze presenti nel consueto approccio al problema della sicurezza perfetta, con qualche imbarazzo ci pare di essere nel film di Mulcahy almeno tutte le volte in cui si ragioni della sicurezza delle chiavi o meglio di lunghe sequenze random da cui trarre chiavi crittografiche.

Il fatto che le operazioni algebriche che investono le sequenze numeriche di chiave e messaggio, non possano essere influenzate dalla veste che attribuiamo loro nelle nostre menti, appare talmente palese col "senno di poi" da lasciare perplessi.

Tuttavia se non si contestualizzano le cose, da un canto si rischia di apparire ingenerosi e, dall'altro, temerari.

Quando a poco più di settant'anni da adesso (1949-2022), si ponevano infatti le basi su cui stiamo lavorando tutt'ora, il panorama storico e tecnologico era quello fuggevolmente adombrato nel primo capitolo e quindi molto diverso da quello che conosciamo.

Allorché furono infatti fissati gli elementi di un critto-sistema o quando si fornirono i requisiti della chiave ideale, sembrò ovvio che talune premesse fossero date per certe (chi avrebbe pensato a delle chiavi distribuite da remoto?) così che dovette essere normale porre da subito in distinti panieri l'insieme detto "delle chiavi" e quello detto dei messaggi.

Pur tuttavia, sebbene ai tempi fosse pacifico, un certo approccio ha assunto ben altro significato al giungere della corrente era digitale.

Nessuno ci dovette badare ..., ma quando, invece di parlare di sequenze binarie o alfanumeriche che possano fungere da chiave o messaggio, si son messe le chiavi (materialmente intese) da un canto e i messaggi dall'altro, s'è passati da un livello logico-matematico (che riguarda grandezze e strutture) ad un livello che attiene allo stato dei fatti.

Per certi versi si è quasi inavvertitamente supposto che non esistano o che siano di scarso interesse, messaggi dal carattere aleatorio, tanto che di tale non banale situazione si sono trascurate le ponderose conseguenze.

[30] Il brano è stato successivamente aggiunto e non fa dunque parte del suddetto report *Secrecy Systems* di Claudio Cappelli.

A prescindere infatti dalle dimostrazioni di cui pure diremo e dal paradosso di cui abbiamo a lungo parlato, sul piano logico non avrebbe senso supporre per eguali operazioni algebriche, che la mutevole attribuzione del ruolo di chiave o messaggio a questa od a quella sequenza numerica, influisca sulla perfetta sicurezza del sistema.

Ma ciò comporta che se ***y*** *può pur essere un messaggio random da cifrare ed* ***x*** *una chiave no-random di minore o ugual lunghezza, l'epocale problema della distribuzione sicura di chiavi crittografiche potrebbe assumere forme più gestibili***.**

Capitolo III

.

Tema

(18)

Buttando le fondamenta degli sviluppi a seguire, si procede a ritroso come un gambero che muova dal crittogramma per giungere a chiave e messaggio.

Anticipando alcuni tratti del nuovo sistema OTP++, in particolare sono precisati gli obiettivi da perseguire per superare le criticità di cui abbiamo detto **(D).**

E' altresì completata la disanima dei passi di *digestion* e sono rimarcate simiglianze e differenze che corrono tra generatori true random e generatori pseudo-random.

Nel contempo è sviluppato il tema della impredicibilità-aleatorietà del segnale e di come tale impredicibilità possa essere preservata e persino implementata in fase di trasformazione (a).

Alcuni Step

(19)

Con riferimento a quanto accennato nello schema in figura uno, con l'intento di buttare un occhio a questioni più teoriche e un altro all'offerta d'un nuovo sistema di cifratura, fissiamo i seguenti passaggi molto di massima:

- **Step 0**

  Attiene alla generazione prodotta da sorgenti TRNG di un flusso cui daremo il nome di macrocode-greggio essendo che una volta memorizzato, fungerà da riserva da cui attingere file greggi[31] da impiegare nella costruzione di chiavi crittografiche.

- **Step 1**

  Riguarda il trasferimento da remoto (per la lunghezza che occorre) di tali file trattati come speciali messaggi random da cifrare con chiavi *no-random* (**B**)(**C**)

- **Step 2**

  Comporta la lavorazione in fase di *digestion* attraverso pari passi condotti presso utenza mittente e ricevente, dei file greggi da cui far discendere quelli vergini che rispondano ai requisiti di cui dicemmo[32]; i *files* così ottenute faranno da chiavi crittografiche di misura ugual-maggiore rispetto a quella del messaggio (**A**)

- **Step 3**

  Comporta l'impiego dei detti file vergini in qualità di chiavi crittografiche con cui è pertanto cifrato il messaggio detto "finale o consueto" che poi sarebbe il vero e proprio dispaccio di senso compiuto tramesso da mittente a ricevente[33].

[31] Si veda in proposito il glossario.

[32] Si veda (08)

[33] **Ricapitolando**

si può sottolineare che mentre il trasferimento di file greggi, che dopo opportuni passi di trasformazione faranno da forma vergine da cui trarre chiavi di crittazione, avrebbe luogo in base al dettato di cui ai lemmi (**B**)(**C**), e mentre i passi da eseguire in fase di digestion terranno conto delle criticità poste in (**D**),

le operazioni di codifica del messaggio finale coincidono con quelle normalmente fissate in letteratura, rispondendo oltre che alla più compiuta lettera dei lemmi (**B**)(**C**) anche a quella del teorema di caratterizzazione in (**A**)

Step (3)

(20)

S'è intanto invisibilmente tracciata una tenue linea rossa che congiunge come puntini taluni oggetti da illustrare qui e altrove; altresì diciamo che tale linea rossa è stata tuttavia resa abbastanza ampia da includere al suo interno procedure che scopriremo simili nella forma ma diverse nello scopo rispetto a quelle di cui stiamo dicendo.

Qui e nel prosieguo nell'anticipare infatti alcuni passaggi, faremo una cernita tra *passi* e *passi* del sistema, privilegiando quelli che riguardano la costruzione delle chiavi per la codifica del flusso del messaggio finale e tralasciando altri consimili.

Sebbene tali passaggi non siano gli unici ad esser praticati nella logica del nuovo cifrario e sebbene siano presi alla lontana e sebbene questo e quello …, essi forniranno una minima traccia sulla quale fissare altre e diverse procedure[34].

Tenendo dunque conto di ciò e quindi del carattere esemplare dei passi da voler illustrare, **se prendiamo le mosse dall'ultimo anziché dal primo Step incluso nel novero in** (19),

sarà per mimare un tentativo di risalita, ma anche per prendere confidenza col sistema muovendo da una fase poco problematica giacchè identica a quella ipotizzata per qualsivoglia sistema a segretezza perfetta.

Se utenza mittente U_T intende infatti inviare un dispaccio a utenza ricevente U_R, impiegherà il segnale rilasciato in *digestion* per fomare una chiave di misura almeno eguale a quella del messaggio come sarebbe volendo codificare il famigerato incipit della Divina Commedia che abbiamo preso a esempio universale di ogni messaggio finale o consueto.

E perciò nel corso di tale passaggio si avrannno in conclusione le seguenti operazioni di codifica-trasmissione-decodifica:

- **(1)** sarà ricavata una chiave, di lunghezza uguale a quella del messaggio, dal cosiddetto file vergine che possiede carattere aleatorio per quanto vagamente accennato e per quanto avremo modo di dire;
- **(2)** sarà cifrato con tale chiave random il messaggio in chiaro;
- **(3)** sulla rete di collegamento verrà dunque trasmesso in forma cifrata il messaggio che abbiamo chiamato messaggio finale;
  avremo pertanto un consueto messaggio in chiaro cifrato con una chiave random di lunghezza almeno uguale rispetto a quella del messaggio;
- **(4)** utenza U_R riceverà il crittogramma *Critto* del messaggio finale (*nel mezzo del cammin di nostra vita ci ritrovammo in una selva oscura* ...) che per quanto detto va tenuto distinto da quello speciale che sarebbe random per definizione;

[34] Lo abbiamo già detto e lo diciamo ancora per l'ultima volta, ma la pubblicazione in essere non include tutto il lavoro svolto.

Per tale ragione, si fa talora riferimento al contenuto di altri documenti che restano tuttavia riservati

(5) utenza U_R provvederà alla decodifica del crittogramma impiegando come chiave di decodifica il file vergine precedentemente ricevuto da remoto in forma grezza.

Il successo di tale Step del tutto ordinario nell'economia dei sistemi perfetti, non dipende perciò da se stesso (non presentando alcuna novità rispetto a quanto normalmente accolto in letteratura) ma dai passi che precedono,
motivo per cui un esame delle proprietà dei file vergini – per stimare se possano o meno fungere da chiavi crittografiche – sarà presto affrontato per essere nuovamente ripreso in appendice[35]

[35] E' bene dire che prove di test condotte sul campo, non solo hanno confermato la qualità e quindi l'aleatorietà dei flussi sottoposti a processi di trasformazione, ma hanno rimarcato che tale aleatorietà non è ridotta ma semmai accresciuta dai passi eseguiti in fase di *digestion*.

Step (2)

(21)

E così vediamo di indagare *per quanto detto sinora* la logica che rende necessaria oltre che efficace la conversione da file greggio a vergine,

perseguendo in un sol colpo i seguenti obiettivi di cui pure dicemmo e che coincidono con quelli del secondo step:

- **(a)** avere che nella trasformazione, i file greggi non riducano ma semmai accrescano l'aleatorietà del segnale di quelli vergini da impiegare come chiavi crittografiche;
- **(b)** avere che i file vergini diventino *altro da sé* rendendosi indistinguibili da qualsivoglia sequenza di pari lunghezza presa a caso e non serbando memoria di quelli grezzi (così da scongiurare il pericolo che nelle sequenze da impiegare una volta soltanto, sia presente qualche reminiscenza statistica dei *files* da cui pure discendono) (**D**).

Giustappunto allo scopo di perseguire tali target, essa trasformazione (*digestion*) in particolare prevede la facoltà di spacchettare le sequenze dividendole in due o più parti (sebbene da qui innanzi si ragionerà soprattutto della loro bipartizione) da sommare a flusso col risultato a sua volta concatenato XOR a un indipendente segnale *no-random.*

Data infatti sequenza detta grezza ***y*** di variabile ***Y*** di lunghezza *L*, estratta dal flusso del *macrocode* generato in origine da sorgente effettivamente aleatoria,

questa è divisa in *p* parti di pari lunghezza L/p

Tali parti son dunque sommate OR esclusivo tra loro così da rilasciare un flusso di caratteri random (output della somma tra parti numeriche generate dalla stessa sorgente).

Il risultato in uscita così ottenuto sarà dunque a sua volta sommato a un distinto segnale *no-random* e cioè poco sparso, che diremo ***x*** di lunghezza parimente uguale a L/p

Invero quella adesso fissata, non è la sola trasformazione che consenta di operare sul segnale greggio perché rilasci una sequenza intrinsecamente diversa,

ma rappresenta la forma da noi preferita, comportando sia il vantaggio dell'efficienza (per la scarsa potenza computazionale richiesta) che quello di permettere una facile dimostrazione della propria efficacia.

Adesso noi non possiamo tuttavia calcolare tutte assieme le corpose conseguenze che discendono dalle operazioni prescritte, ma volendo cominciare dall'impatto che hanno sulla aleatorietà del segnale,

andiamo intanto a rimarcare che tali passi di *digestion* comportano l'esecuzione di partizioni e somme XOR, dove tale fatto ricorda talune operazioni di "sbiancamento" le quali hanno come scopo proprio quello di aumentare la aleatorietà del segnale.

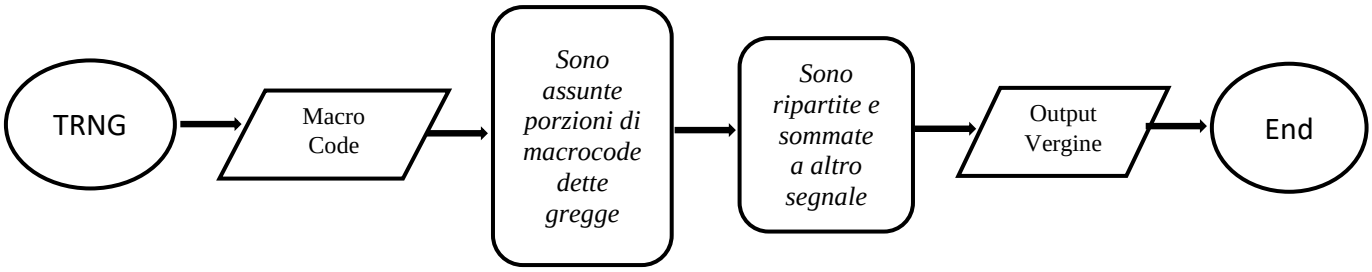
Infatti la prima domanda che rivolgiamo a noi stessi e cui cercheremo di rispondere riguarda gli effetti sui *files* greggi, delle minime operazioni che abbiamo adesso indicate,

facendo preciso riferimento alle conseguenze che possiamo prevedere sulla randomicità del segnale vergine in uscita.

In altre parole, il passaggio che richiede la partizione di una sequenza ***y*** generata da una sorgente random, la somma di parti di ***y*** originariamente e effettivamente aleatorie, l'ulteriore somma con una sequenza ***x*** che possiamo definire come debolmente impredicibile e quindi computabile entro la relativa soglia di complessità, quali risultati comporta?

Si può preservare o addirittura accrescere la impredicibilità originale?

Figura 03 - Diagramma



Sorgenti Deboli

(22)

Invero per intanto è bene sapere che i flussi ingenerati dalle cosiddette sorgenti TRNG o true random di cui vorremmo preservare o persino accrescere l'impredicibilità e quindi l'aleatorietà del segnale (**a**) dipendono sia dalla campionatura di fenomeni quantici che da quella di eventi di tipo caotico che non possono garantire da soli un flusso dall'andamento ideale;
si tratta di un fatto che richiede un fermo immagine, giacché significa che ciascun output ottenuto sarà sempre sottoposto a delle manipolazioni che consentono di affilare una qualità diversamente manchevole. Per essere espliciti, diciamo pure che senza tale intervento la maggior parte dei generatori fisici farebbe peggio in termini di randomicità, della maggior parte dei generatori pseudo-random sebbene questi siano esclusi dall'impiego nei sistemi perfetti;

se si volesse saltare perciò l'impiego di codesti correttivi detti di sbiancamento o whitening**,** che talora giustappunto contemplano delle operazioni in XOR[36], dovremmo rinunciare a quel salto che consente alle fonti true random d'essere efficaci.
In generale, tenendo distinti i diversi processi, l'architettura di un generatore effettivamente random prevede tuttavia quanto segue:

- **(i)** Un segnale analogico generato dalla sorgente che registra l'andamento di fenomeni fisici presi in ingresso da un digitalizzatore;
- **(ii)** Un digitalizzatore che campiona il segnale analogico e lo converte in un flusso discreto *near random* (segnale digitalizzato);
- **(iii)** Un segnale digitalizzato dato in input a un programma che lavora allo sbiancamento del flusso, così da rilasciare in uscita *files* più qualitativi (e quindi fortemente impredicibili) di quelli ottenuti in ingresso dal digitalizzatore.

Ora se volessimo tradurre tutto ciò algebricamente, potremmo azzardare una definizione secondo cui un generatore true random dal momento della campionatura in avanti (dove la precedente fase non ha carattere matematico investendo fenomeni naturali),
sarebbe definita dalla funzione $f : (\mathbb{Z}_2)^L \rightarrow (\mathbb{Z}_2)^l$
dove l ed L sono interi positivi,
dove $\mathbb{Z}_2$ è un campo popolato da valori binari 0, 1
dove $e \in (\mathbb{Z}_2)^L$ rappresenta il flusso discreto ottenuto in forma binaria dal digitalizzatore e preso in input dalla relazione medesima,

[36] Molti sono i metodi di sbiancamento condotti tramite algoritmi conosciuti come *estrattori di randomicità*; il reiterato impiego dello XOR è appunto uno di essi

dove il valore a sua volta dato in forma binaria da $\boldsymbol{f}(e) \in (\mathbb{Z}_2)^l$ è invece lo stream in uscita effettivamente random.

Esso sarà dunque dato dall'insieme delle trasformazioni di spazio $\mathbb{Z}^L$ (cui appartengono i flussi *near random* che supponiamo di lunghezza *L*) in spazio $\mathbb{Z}^l$ (cui appartengono i flussi *random* avuti in seguito alle operazioni di sbiancamento),

motivo per cui ad ogni flusso discreto corrisponde una sequenza di spazio $\mathbb{Z}^l$

dove $l \leq L$ in quanto i passi di trasformazione non dovranno mai accrescere il numero di bit in accesso non potendo rappresentare più informazione di quella che effettivamente possiedono**.**

Se saltiamo perciò il primo passaggio caratterizzato dal rilevamento degli effetti dati da fenomeni fisici (**i**), si determina una forma sorprendentemente simile a quella che qualifica le trasformazioni pseudo-random e ciò pare contraddire sia quanto detto in letteratura, sia quanto sembrerebbe sul piano intuitivo.

Scansando la trappola di preconcetti che si sono sedimentati negli anni e nei decenni, possiamo infatti rilevare che se sui generatori fisici (TRNG) abbiamo un accesso analogico montato su un dispositivo HW che raccoglie informazioni da sottostanti fenomeni naturali,

su quelli di stampo matematico (PRNG) avremo una situazione meno diversa di quanto si creda, cosa che tuttavia risulta evidente a patto di considerare la totalità dei processi di generazione pseudo-random e non solo quanto attiene alle attività che fanno seguito allo sviluppo su software di un seme comunque altrove generato.

Qualora il seed assunto in precedenza da un PRNG sia infatti prodotto da una sorgente autenticamente aleatoria *come spesso accade*,

è ovvio che sarà anch'esso dato dal rilevamento di fenomeni casuali assunti da una porta analogica[37], prima d'esser tradotti in forma binaria;

da tale punto di vista, l'unica palpabile differenza risiede nella dissimile architettura adottata, per la quale i generatori effettivamente random usano concentrare distinti processi all'interno di un unico dispositivo HW mentre i sistemi pseudo-random distribuiscono i loro processi in luoghi e tempi diversi,

per cui la produzione del seme è separata dalle fasi di trasformazione e ciò mimetizza l'effettiva struttura del sistema[38].

[37] Se prendiamo ad esempio il caso in cui il seme sia generato dai movimenti casualmente prodotti sul mouse dall'operatore umano, l'ingresso analogico sarebbe dato dallo stesso mouse che è poi una periferica HW capace di registrare spostamenti e clic

[38] Si badi che non vogliamo dire che sistemi *pseudo-random* e *true random* siano la stessa cosa, ma che occorre piuttosto individuare con cura il punto nel quale effettivamente divergono

Figura 04 – Schema dei tipici processi di generazione true random, i quali son tutti concentrati in un unico dispositivo HW

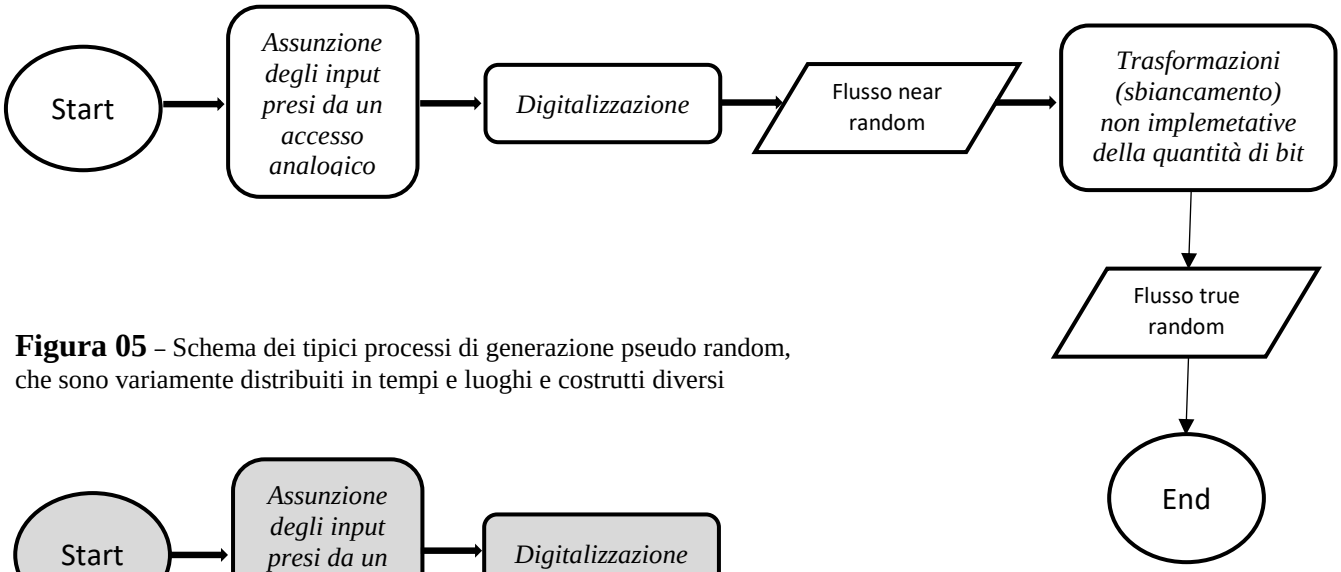
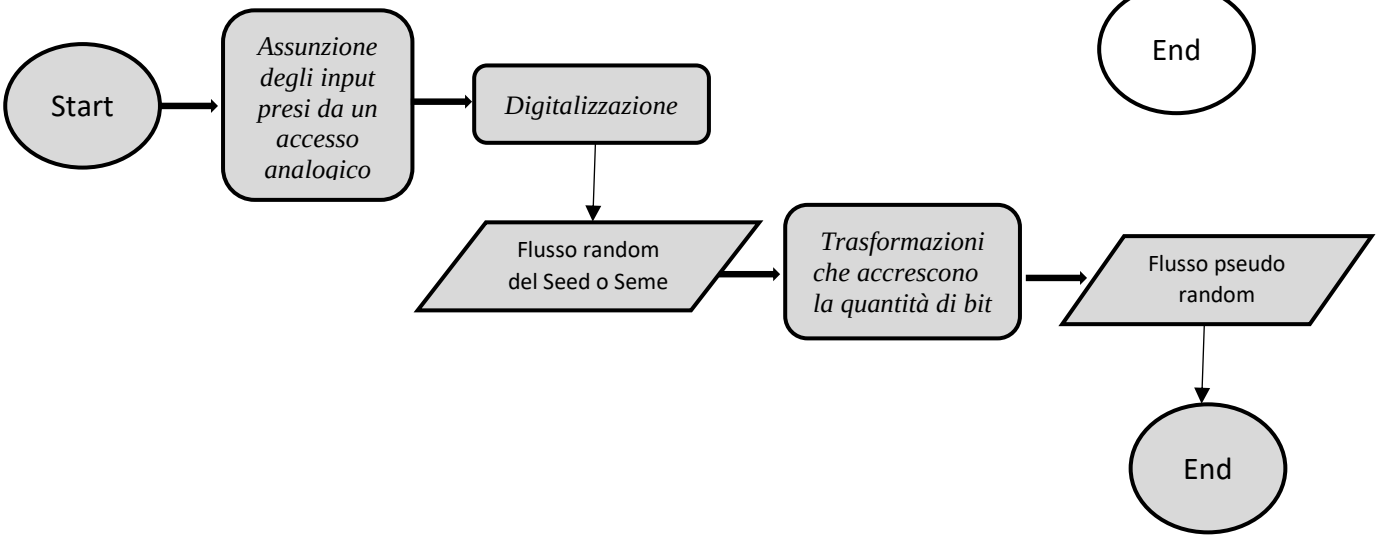


Figura 05 – Schema dei tipici processi di generazione pseudo random, che sono variamente distribuiti in tempi e luoghi e costrutti diversi



La Piccola Differenza!

(23)

Se poniamo tuttavia la questione in termini più rigorosi, avremo modo di mettere a fuoco analogie e differenze in una forma diversa di quanto fatto sinora.

Intanto è facile notare con stupore che la definizione standard di generatore pseudo-random, una volta tradotta nelle relative funzioni di trasformazione,

quasi coincide con quella da noi data per le sorgenti *true random*, almeno per come codeste sono usualmente concepite.

Un generatore PRNG, i cui flussi ai nostri fini abbiamo assimilato sin dapprincipio a quelli banalmente e puramente *non random* (03) sarà infatti definito dalla simile ma diversa funzione $f : (\mathbb{Z}_2)^{ll} \rightarrow (\mathbb{Z}_2)^{LL}$

dove pure *ll* ed *LL* sono interi positivi ma per *LL* molto maggiore di *ll*

Anche in tal caso, la formulazione impiegata non tiene conto del processo stocastico posto in origine, giacché assume in input un seme già formato che possiamo indicare come $\mathbf{s} \in (\mathbb{Z}_2)^{ll}$ dove $\mathbb{Z}_2$ sarà nuovamente un campo di valori binari.

Confrontando tuttavia le distinte funzioni, dove la precedente in (25) che si scrive $f : (\mathbb{Z}_2)^{L} \rightarrow (\mathbb{Z}_2)^{l}$ per $l \leq L$ si riferisce alle sorgenti effettivamente random,

balza agli occhi di quanto la differenza sia minima sul piano formale, sebbene nell'accezione che a tale aggettivo diede Churchill commentando il discorso alla Camera di un deputato laburista, che sudava le proverbiali sette camicie nel dire che tra sesso maschile e femminile non ci sia altri che una piccola differenza.

Un hurrà per la piccola differenza! esclamò Churchill muovendo da un banco dell'ultima fila sino al proscenio dei primissimi posti.

Nel nostro caso la distanza giustappunto risiede nel diverso rapporto di misura che corre tra lunghezze dei flussi in accesso e lunghezza dei flussi in uscita,

dove nell'un caso abbiamo che la falsa estensione di randomicità prodotta da un generatore pseudo random, comporta che ogni effettiva informazione resti concentrata nel seme da cui tutto discende[39],

mentre quando diciamo di sorgenti true random, abbiamo che ogni informazione (sebbene sottoposta a trasformazioni) risulta effettivamente distribuita sui simboli del flusso di output che non moltiplica in modo artificioso i bit di quelli in entrata.

Volendo perciò riprendere il filo della trattazione interrotta, in rapporto a quanto detto sui passi di *digestion*,

a questo punto è facile notare che a prescindere dalle dimostrazioni cui daremo conto più innanzi sino alle ultime pagine dell'appendice,

[39] Ogni effettiva informazione è concentrata nel seme da cui tutto si deduce ed è quindi invisibilmente diluita nei flussi in uscita la cui lunghzza è gonfiata attraverso artifizi

dalle prove condotte sul campo, un primo indizio del perché le operazioni XOR preservino o persino amplifichino la randomicità del segnale, ci viene dal fatto che non solo esse non simulano un accrescimento fittizio ma danno luogo a una compressione dei flussi numerici;
attività questa, che costituisce una pre-condizione necessaria ma non sufficiente alla salvaguardia di quel requisito di randomicità-imprecidibilità da noi in precedenza fissato.

Quanta Aleatorietà

(24)

Riprendendo la rotta dall'idea di memoria e da quella tempestosa di impredicibilità da stimare secondo quanto accennato qui e innanzi,
vorremo collezionare nuovi concetti e rileggere e modificare concetti precedentemente trattati girando e rigirando le cose per poterle guardare da più prospettive.
Trovandoci intanto davanti a dei costrutti logici come ad esempio sarebbero quelli di "aleatorietà" e "casualità" ne daremo un tenero assaggio sapendo di mettere in fila questioni alquanto spinose;
osserviamo infatti che se alcune "voci" sono particolarmente popolari (chi non ha mai pensato in termini di fortuna e sfortuna) ciò non significa che rispondano a teorie condivise.
Occorre infatti comprendere che la difficile questione della aleatorietà della sorgente e del segnale ingenerato, presenta non pochi risvolti pratici che ne influenzano il significato;
tanto che alquanto pratico dovette essere l'approccio di Shannon che pure non era sprovvisto di fini strumenti ermeneutici.
Il problema che cercò di risolvere tra guerra (1939-1945) e dopoguerra (1945-1950) e che vorremmo risolvere anche noi al volgere del terzo millennio ..., era legato a una delle più sentite questioni in materia di critto-analisi; si trattava di stabilire quanta aleatorietà è contenuta nella lingua inglese ed in ciascuno dei suoi ventisei caratteri alfabetici,
così da affinare gli algoritmi di risalita dal crittogramma al messaggio (almeno nei casi tenuti fuori dal recinto della sicurezza perfetta).

Come stabilire allora quanta aleatorietà è presente in una lingua scritta e parlata?

E' inconfutabile che ci sia aleatorietà nella lingua inglese e in qualsiasi altra lingua conosciuta, altrimenti in uno stato di perfetta certezza *sapremmo cosa qualcuno dice prima ancora che parli*.
Calcolare tuttavia fedelmente l'aleatorietà insita in una lingua è impossibile, e lo era ancor più quando sul finire degli anni quaranta del secolo scorso, a far di conto erano batterie di esseri umani muniti di carta e penna.
Nondimeno Shannon non si perse d'animo, e volendo stimare *how many casuality* si possa insinuare nella lingua inglese praticò un facile espediente.
Si era temporaneamente stabilito in New Jersey quando diede a leggere dei testi a un gruppo di volontari, facendo scandire una lettera alla volta e assegnando a ciascuno il compito di indovinare la successiva.
Se tra tali testi di cui, in verità, non abbiamo contezza (essendo che nutriamo forti dubbi che effettivamente sappia quali furono i titoli adottati) ci fossero ad esempio stati i primi versi del Paradiso Perduto di Milton, le cose non sarebbero andate troppo diversamente da come vanno ancora oggi quando analoghi calcoli sono affidati a elaboratori di ultima generazione che traggono le loro conseguenze dalla messe dei cosiddetti Big Data.

The Paradise Lost

1. Of man's first disobedience, and the fruit
2. Of that forbidden tree, whose mortal taste
3. Brought death into the world, and all our woe
4. With loss of Eden, till one greater man
5. Restore us, and regain the blissful seat

Avendo quindi selezionato un campione di volenterosi all'oscuro di tutto, è plausibile che la scelta della prima lettera sia stata fortuita, sebbene con una probabilità lievemente diversa da quella discendente dalla mera stima del numero di simboli alfabetici, per la semplice ragione che in inglese alcune lettere risultano notoriamente più frequenti di altre;
già la predizione del secondo carattere sarebbe poi facilitata dalla conoscenza del precedente, essendo plausibile che alla vocale "o" segua una consonante, con alcune consonanti come "n" "f" "r" più probabili di "m" "l" "z" e così andando.
A ben vedere, persino l'esame delle primissime battute dei sontuosi versi citati, ci consente di focalizzare qualcosa di notevole.
Intanto appare palese la ragione per cui l'esatto calcolo della aleatorietà d'una lingua sia impresa ardua come sarebbe quella di scalare gli scaffali della mitica *Biblioteca di Babele* di cui scrisse un grande e controverso scrittore argentino nato e vissuto nel quartiere Palermo a Buenos Aires.
Si dovrebbe calcolare la probabilità di comparsa di ciascun carattere alfabetico in ogni testo scritto o orale che sia ..., ed in qualunque frase non ancora concepita ma nondimeno concepibile da qui alla fine del mondo o perlomeno alla fine di tutti i sudditi dell'impero britannico;
ma nemmeno tale computo sarebbe definitivo, dovendo scontare la misura della frequenza di ogni testo tratto dall'insieme di tutti i testi possibili che includano da una a innumerevoli pagine che abbiano valore letterario o non ne abbiano affatto,
che siano state redatte da qualcuno o che siano semplicemente immaginate o immaginabili in qualche universo parallelo o altro ancora.

Un pensatore osservò che tutti i libri, per diversi che fossero, constavano di eguali elementi: la spaziatura, il punto, la virgola, le ventidue lettere dell'alfabeto[40]*; ebbe poi a precisare un fatto che tutti i viaggiatori confermarono: non vi sono, nella vasta Biblioteca, due soli libri identici.*
Da queste premesse che la Biblioteca è universale, e che i suoi casellari registrano tutte le possibili combinazioni dei incontrovertibili dedusse venticinque simboli ortografici (numero, anche se vastissimo, non infinito) cioè tutto ciò ch'è dato di esprimere, in tutte le lingue del mondo.
Tutto: la storia minuziosa dell'avvenire, le biografie degli arcangeli, il catalogo fedele della Biblioteca, migliaia e migliaia di cataloghi falsi, la dimostrazione della falsità di tali cataloghi, la dimostrazione

[40] In inglese le lettere sono ventisei, in italiano ventuno, nello spagnolo argentino in cui scrive Borges ventidue, sebbene sia singolare che conteggi solo il punto, la virgola e la spaziatura come ulteriori segni ortografici.

della falsità del catalogo autentico, l'evangelo gnostico di Basilide, il commento a esso evangelo, il commento del commento a questo evangelo, il resoconto veridico della tua stessa morte, la traduzione di ogni libro in tutte le lingue, le interpolazioni di ogni libro in tutti i libri (*La Biblioteca de Babel*, Jorge Borges, 1941).

Ora se tale calcolo dei calcoli è veramente assurdo, l'esperienza cui abbiamo accennato ci aiuta a mettere a nudo la correlazione tra *crescita* della quantità d'informazione[41] e *decadimento* della aleatorietà della successione ortografica.
Più innanzi nella seconda parte, avremo infatti modo d'imbatterci nel nostro cammino in una delle seguenti ipotesi da cui faremo scaturire conseguenze imprevedibili:

(1) se lo scritto prescelto dovesse apparire sensato, in quanto dotato di semantica e sintassi e quindi di un ordine riconoscibile,
man mano che si sommino informazioni su informazioni, sarà sempre possibile aggiustare la mira accrescendo la probabilità di predire dopo *n* lettere quella a seguire;

(2) se il testo fosse invece privo di qualsiasi rudimento sintattico, giacché effettivamente aleatorio come per i numeri presi da un'estrazione del lotto,
non si avrebbe alcun accumulo di informazione e la probabilità di predire l'ennesima riffa non andrebbero a salire.

Così vediamo che la predicibilità-impredicibilità di cui dicemmo, torna nuovamente in gioco illuminando sottili correlazioni prima soltanto annunciate;
se riprendiamo infatti il nostro esempio supponendo di voler leggere i primi trentatré caratteri della riga in apertura,
prendendo in tal modo atto d'una frase nondimeno incompiuta (*of man's first disobedience, and the frui…*), a meno di non essere dei completi sprovveduti non avremmo problemi a indovinare l'ultima consonante così da chiudere il primissimo verso del Paradiso Perduto.

[41] E' bene tener separati due concetti che hanno variamente fatto da implicito corollario a quanto detto in questo e in precedenti paragrafi.
Se prendiamo una stringa, quanta più informazione detiene a parità di lunghezza, tanto più risulterà aleatoria e impredicibile.
Al contrario una sequenza di *n* segni "0" che si susseguono uguali, conterrà pochissima informazione essendo facile da comprimere.
Diverso è tuttavia il concetto per cui l'impredicibilità diminuisce al crescere dell'informazione in possesso di un osservatore esterno; tale crescita si può registrare laddove chi osserva possa raccogliere informazioni utili come nel caso dei versi del Paradiso Perduto di Milton che hanno semantica e sintassi e perciò un qualche livello di maggiore o minore predicibilità statistica.
Dal punto di vista di tale osservatore, possiamo infatti diversamente dire che quanta più informazione raccoglie tanto maggiore sarà la sua capacità-probabilità di predizione.

Of man's first disobedience, and the frui (t)
Of that forbidden tree, whose mortal taste
Brought death into the world, and all our woe,
With loss of Eden, till one greater man
Restore us, and regain the blissful seat

Sulla base di quanto detto e sulla base di come i fatti si sarebbero effettivamente svolti negli anni cinquanta del secolo scorso, è dunque facilmente intuibile perché per Shannon casualità e impredicibilità in buona sostanza coincidano.
Pur volendo tralasciare la questione del calcolo dell'impredicibilità rispetto a un limite assegnato in base alla potenza di calcolo di chi attacca e alla classe di complessità di chi difende,
quando ci troviamo dinanzi al caso estremo d'una successione effettivamente random le cose si semplificano giacché la impredicibilità-aleatorietà è massima e viceversa col risultato che saremmo prossimi al caso delle estrazioni del lotto.
Occorre però dire che tale sovrapposizione di concetti, non fu condivisa da tutti ma, tra le diverse ipotesi volte a formalizzare l'idea di casualità (ipotesi anche molto intriganti come quelle distintamente avanzate da Chaitin e Kolmorogov dai lati opposti della cosiddetta cortina di ferro) quella di Shannon ha l'indubbio vantaggio di dare i numeri[42].
Ed ai nostri fini non può che costituire un imprescindibile riferimento per il semplice fatto che i test consigliati dai maggiori enti di standardizzazione del segnale, e quindi i test di aleatorietà e di sparsificazione che fanno fede nella comunità scientifica, in buona sostanza sono tutti tesi a rimarcare la maggiore o minore predicibilità di un flusso che non deve tradire ricorrenze statistiche che consentano di ottenere informazioni sulla frequenza di comparsa di questo o quel simbolo.

[42] Anche se indirettamente, l'opzione promossa da Shannon consente delle misure come dimostrano i test di randomicità promossi dai maggiori enti internazionali

Incertezza e Informazione

(25)

E' dunque l'idea di impredicibilità a legare come in pasticceria nozioni tra loro fortemente correlate fornendo le basi per comprendere i rapporti di co-varianza che muovono grandezze dissimili.

Da un canto abbiamo infatti concetti quali quelli di aleatorietà e casualità da noi spesso citati e, dall'altro concetti come quelli di "informazione" "incertezza" e "entropia" che ne conseguono come ciliegie.

Si badi bene che qui non siamo tuttavia interessati a dare definizioni rigorose quanto a arricchire il nostro vocabolario.

E' bene infatti chiarire che non intendiamo sottilizzare sul significato da dare a espressioni familiari ma dall'accezione non sempre univoca, avendo invece in animo di illustrare il modo in cui le andremo effettivamente a impiegare.

In termini generali già dicemmo qualcosa in più punti ..., ma qui torneremo sui nostri passi per toccare con mano alcuni aspetti dei metodi adottati in fase di sbiancamento e per dar conto della misura dell'incertezza d'una sorgente e dei flussi da essa ingenerati.

Avendole rigirate infatti in più salse ..., ormai sappiamo che le fonti aleatorie appartengono al novero delle sorgenti discrete essendo che a campionatura ultimata emettono simboli più o meno probabili appartenenti a un alfabeto finito *X*,

dove *X* sarà per definizione uguale ad $x_1, x_2, \ldots, x_n$

Ora ognuno di tali simboli sarà contraddistinto da una probabilità di comparsa usualmente detta P_i e da una correlata quantità di auto-informazione detta I_i

Fu sempre Shannon a fare da apripista definendo la misura dell'informazione insita in ciascun simbolo e enunciando in primo luogo la seguente formula secondo cui:

$$I_i = -\log_b P_i = \log_b \frac{1}{P_i} \text{ [simbolo]}$$

Dove se base *b* del logaritmo fosse uguale a due, come accade ogni qual volta si misuri in bit l'auto-informazione contenuta in ciascuno dei simboli alfabetici,

avremmo giustappunto quanto segue:

$$I_i = -\log_2 P_i = \log_2 \frac{1}{P_i} \text{ [bit]}$$

Quello che è bene tuttavia sottolineare è il rapporto di co-varianza inversa che corre tra la crescita della quantità di informazione misurata, che annulla a tendere ogni impredicibilità, e la dose di incertezza di ciascun simbolo binario[43].

[43] Dove a tendere, la maggior incertezza si avrà per simboli equi-probabili.

Così come un segno binario del quale abbiamo ogni informazione non potrà presentare alcuna incertezza, risultando prevedibile alla pari dell'ennesimo bit d'una qualche sequenza statisticamente piatta,
vale pure il contrario per cui un segno massimamente incerto non implica alcuna informazione aggiuntiva, dove ciò significa che stimare la quantità di informazione finisce con l'essere lo stesso che calcolare quella di incertezza in quanto dove cresce l'una diminuisce l'altra e viceversa.
Se a partire da quanto detto, volessimo però fissare anche quella "entropia della sorgente discreta" che ci consente di passare dal concetto di quantità d'informazione d'un singolo simbolo, a quello di informazione data dalla incertezza delle auto-informazioni impresse in sequenza nei simboli di un segnale (I_1, I_2) avremmo la seguente espressione:

$$H(X) \ = -\sum_{i=1}^{n} P_i \ log_{2\,(P_i)} \ = \sum_{i=1}^{n} P_i \ log_{2\ 1/P_i} [\text{bit}]$$

Per cui laddove *X* sia l'alfabeto impegnato da una sorgente binaria, i cui caratteri sono ridotti all'essenziale, dove *n* limite alto della sommatoria sarebbe uguale a due come per una frase che impieghi due elementi x_1, x_2,
s'avrebbe che per una probabilità di comparsa di questo e quell'altro simbolo binario pari a un mezzo di uno (equi-probabilità) si ottiene in uscita un'auto-informazione media di un bit,
e per una probabilità di comparsa eguale ad uno (certezza assoluta) una auto-informazione media uguale a zero essendo che nessuna informazione si aggiunge a quella che già abbiamo in uno stato di massima certezza[44].
Prendendo infatti in carico sia il limite inferiore che quello superiore dell'intervallo, abbiamo che se la probabilità che un evento si verifichi è massima e perciò pari ad 1, avremo che *H(X)* vale 0,
ma se si dovesse creare uno stato di equi-probabilità dove quella media sarà la minima possibile e quindi uguale a un ennesimo,

H(X) varrà $log_2\ n$ essendo che $\frac{1}{n}$ * *n* * $log_2\ n$ si potrà giustappunto scrivere come,

$log_2\ (1) \leq H(X) \leq log_2\ n$ che per *n* uguale a due infatti ci darebbe $0 \leq H(X) \leq 1$

E in tal modo l'entropia che si definisce come misura della quantità media di incertezza, è nei fatti pure misura inversa della quantità di informazione rilasciata dalla sorgente;
se sul piano formale incertezza e informazione sono inversamente proporzionali, anche volendo ragionare in modo intuitivo non potremmo non riconoscere le molteplici correlazioni che investono la misura della impredicibilità su cui si interrogavano i critto-analisti che hanno segretamente operato sulla lingua inglese nei laboratori dell'Università del New Jersey.

[44] Nel caso si crea la coincidenza per la quale sia la quantità di informazione insita in ciascun simbolo dell'alfabeto della sorgente, sia l'alfabeto medesimo, sono espressi in un sistema binario.

Più Aleatorietà (a)

(26)

C'è tuttavia un fatto molto preciso che ci consente d'inquadrare il tema passando da un piano meramente teorico a uno pratico;

se per certi versi non si dovrebbero infatti inquinare questioni di carattere formale con altre che attengono ai nostri comportamenti, occorre ammettere che quando ragioniamo di sorgenti numeriche, un qualche strabismo appare inevitabile nella misura in cui migriamo dal campo dell'algebra a quello dell'algebra applicata e, più significativamente, dell'algebra applicata alle trasformazioni da uno spazio a altro spazio numerico.
A tal proposito abbiamo infatti spiegato che tutte le sorgenti note come fonti naturali di randomicità, per funzionare richiedono degli artefatti e cioè dei programmi che convertono la debole impredicibilità della fonte in una impredicibilità-casualità forte.
In teoria i fenomeni registrati dal generatore dovrebbero essere caratterizzati da massimi livelli di entropia, ma per la difficoltà a tenere il segnale indenne da perturbazioni**,** ciò semplicemente non accade. In altre parole, pare quasi impossibile assumere in purezza un fenomeno caotico o quantistico e quindi intrinsecamente aleatorio, col risultato che dati alterati comportano inopportuni picchi statistici.
Concetti non indipendenti tra loro come pure sarebbero quelli di andamento ideale dei flussi numerici e entropia effettivamente misurata in uscita (output) palesano infatti dei problemi illuminati dai test di qualità del segnale.

Ironically, pseudo random numbers often appear to be more random than random numbers obtained from phisical sources. Ironicamente, i numeri pseudo random appaiono sovente più casuali dei numeri random ottenuti da fonti fisiche.

Quello citato è un breve passo della pubblicazione 2010 prodotta dal NIST che, sul punto, pare mettere le mani avanti[45];
dietro la facciata alquanto lieve si nasconde infatti una questione non da poco che lascia trapelare delle crepe che non dovrebbero esserci ma che pure ci sono, essendo messe a nudo dall'affinarsi (oltre che dal complicarsi) delle più sofisticate prove sperimentali consigliate dagli enti preposti alla standardizzazione qui e oltreoceano.

[45] *A Statistical Test Suite for Random and Pseudorandom Number Generator for Cryptographic Applications,* NIST, National Institute of Standards and Technology, U.S. Department of Commerce, April 2010.

Tale incongruenza è tuttavia più evidente nei documenti divulgati sul tema dall'omologo tedesco che, nel dare a sua volta degli standard, ha fissato tre distinte classi di qualità che vanno da un punteggio minimo a uno massimo di efficacia nella capacità di generare flussi numerici che denotano un alto livello di incertezza.
In pratica il BSI ovvero il *Bundesamt für Sicherheit in der Informationstechnik* o Istituto Federale Tedesco per la Sicurezza Aziendale assegna il miglior punteggio a quei generatori effettivamente random, che producono flussi aderenti a determinate condizioni parametriche, cui si aggiunge un'ultima sorprendente raccomandazione.
Per migliorare i risultarti ottenuti, è infatti richiesto che in fase di post-processing si proceda con una particolare manipolazione dei numeri generati dalla sorgente, tramite una susseguente codifica che una società leader del settore attua impiegando il noto algoritmo di cifratura AES che non è dunque impiegato ai fini caratteristici ma per accrescere la qualità del segnale[46]
Ora a prescindere da tale ingiunzione, è un fatto che tra i diversi metodi impiegati dai programmi di *whitening*, altresì conosciuti come estrattori di randomicità, vi sia pure quello di prendere più flussi numerici sommandoli XOR tra loro, così da ottenere una crescita della quantità media di incertezza ed un conseguente calo del *bias*[47]

A technique for improving a near random bit stream is to exclusive-or the bit stream with the output of a high-quality cryptographically secure pseudorandom number generator such as Blum Blum Shub or a strong stream cipher. (...). A related method which reduces bias in a near random bit stream is to take two or more uncorrelated near random bit streams, and exclusive OR them together.
Un metodo che riduce il bias di un flusso numerico quasi random è quello di prendere due o più decorrelati flussi quasi random e di sommarli XOR tra loro.

La base teorica celata dietro tali espedienti passati in rassegna anche dall'edizione USA d'una libera enciclopedia molto popolare[48], non ci è nuova essendo data da quello XOR-Lemma del quale dicemmo e del quale dicono molteplici pubblicazioni tra cui quelle di Noam Nisan, Avi Wigderson, Impagliazzo, Evangelos Kranakis (*Primality and Cryptography*) Oded Goldreich (*Three XOR-Lemmas, An Exposition*) e quella dello stesso Goldreich e altri (*On Yao's XOR Lemma*).
Quest'ultimo contributo è particolarmente incisivo fornendo una dimostrazioni originale e mettendo nero su bianco il principio per cui:

[46] Si veda *Functionality classes for random number generators* di Wolfgang Killmann, T-Systems GEI GmbH, Bonn, e *Werner Schindler, Bundesamt für Sicherheit in der Informationstechnik* (BSI), Bonn, 18 Settembre 2011.
[47] Si tenga conto che qui facciamo riferimento alla quantità media dove $0 \leq H \leq 1$ piuttosto che a quella cumulata delle funzioni di ripartizione.
[48] *Hardware Random Number Generator,* Wikipedia USA.

A function constructed by concatenating the values of the original function on several independent instances is much more unpredictable, with respect to specified complexity bounds, than the original function.
Una funzione data dalla concatenazione dei valori della funzione originale a diverse istanze indipendenti, diviene molto più impredicibile della funzione di partenza.

Dove quello appena ripreso è anche uno dei tanti modi di dire che a determinate condizioni *dove più output si sommano l'uno all'altro* tende a crescere l'aleatorietà del segnale che diventa sempre più impredicibile man mano che si avanza nella sommatoria.
In termini alquanto spicci che ci consentono di visualizzare da subito il senso dello XOR lemma in rapporto al nostro impegno,
possiamo tuttavia ribadire che se i valori di almeno *t* istanze (per *t* sufficientemente grande) fossero concatenati XOR tra loro, si avrebbe come effetto quello di generare in funzione $F(x_1, x_2, x_3,..., x_t)$ un unico bit estremamente difficile da invertire.
Se volessimo tuttavia reiterare tale esperienza distribuendola lungo un ventaglio di più predicati booleani invece che uno soltanto,
per $f \stackrel{\text{def}}{=} f_1 \dots f_n$ ma pure $F \stackrel{\text{def}}{=} F_1 \dots F_n$ sebbene ciascuna di tali *F* dia a sua volta per definizione in uscita un unico simbolo 0,1 ...,
ponendo i risultati nell'ultima riga d'una tabella come quella a seguire, si passerebbe da un output di un solo bit a una sequenza poco o nulla invertibile di ***n***-bit che forma una stringa dal carattere duramente aleatorio; costruiamo pertanto una tabella del seguente tipo coi valori assegnati a titolo esemplificativo per *t* istanze (relative ad *n* funzioni) da sommare XOR tra loro.

Tabella

Output f_1 Input F_1		Output f_2 Input F_2		Output f_3 Input F_3		Output f_4 Input F_4		*(…)*	Output f_n Input F_n	
$x_{1\,(1)}$	*0*	$x_{1\,(2)}$	*1*	$x_{1\,(3)}$	*0*	$x_{1\,(4)}$	*0*	*(…)*	$x_{1\,(n)}$	*0*
$x_{2\,(1)}$	*1*	$x_{2\,(2)}$	*1*	$x_{2\,(3)}$	*0*	$x_{2\,(4)}$	*0*	*(…)*	$x_{2\,(n)}$	*0*
$x_{3\,(1)}$	*1*	$x_{3\,(2)}$	*0*	$x_{3\,(3)}$	*1*	$x_{3\,(4)}$	*1*	*(…)*	$x_{3\,(n)}$	*1*
$x_{4\,(1)}$	*0*	$x_{4\,(2)}$	*0*	$x_{4\,(3)}$	*0*	$x_{4\,(4)}$	*0*	*(…)*	$x_{4\,(n)}$	*1*
(…)	*(…)*	*(…)*	*(…)*	*(…)*	*(…)*	*(…)*	*(…)*	*(…)*	*(…)*	*(…)*
$xt_{(1)}$	*1*	$xt_{(2)}$	*1*	$xt_{(3)}$	*1*	$xt_{(4)}$	*0*	*(…)*	$xt_{(n)}$	*0*
Output F_1		Output F_2		Output F_3		Output F_4		*(…)*	Output F_n	
	1		**1**		**0**		**1**	*(…)*		**0**

Dove tale esempio è creato, ponendo in colonna i distinti output riconducibili a più istanze di f per $f \stackrel{\text{def}}{=} f_1 \dots f_n$
che tuttavia si immaginano presi come input da sommare XOR in F per $F \stackrel{\text{def}}{=} F_1 \dots F_n$
Dove nella prima riga sono posti i valori $x_{1(1)}\, x_{1(2)} \dots x_{1(n)}$ *rilasciati in output da* f_1 *e presi in input da* F_1
nella seconda, i valori $x_{2(1)}\, x_{2(2)} \dots x_{2\,(n)}$ *rilasciati in output da* f_2 *e assunti in input da* F_2
nella terza, i valori $x_{3(1)}\, x_{3(2)} \dots x_{3\,(n)}$ *rilasciati in output da* f_3 *e assunti in input da* F_3
sino a giungere nella t-esima riga, ai valori $x_{t\,(1)}\, x_{t\,(2)} \dots x_{t\,(n)}$ *che sono relativi alla t-esima istanza rilasciata in output da ciascun predicato f*
per essere presa in accesso da ciascuna funzione F così come definita in (10)
Per cui i risultati che si andrebbero a ottenere, posti sull'ultima riga in basso, formano in uscita una stringa 1101 … 0 *formata di bit estremamente difficili da invertire*.

In altre parole funzione F_1 prenderà in accesso i risultati di più istanze indipendenti che sommate tra loro, non potranno che rilasciare un unico bit fortemente impredicibile trascritto nella prima cella utile dell'ultima riga in basso a sinistra**;**

parimente sarà *operando da sinistra a destra* per le susseguenti funzioni $F_2\, F_3\, F_4$ e così andando sino all'ennesima, in modo tale che concatenando i valori in uscita di ciascuna di esse, si comporrà una sequenza estremante difficile da invertire.

Diventa dunque ovvio osservare come in tale simulazione, ogni riga *in tabula* contiene una successione come sarebbe una qualche sequenza numerica che sommata XOR a quelle allocate sulle altre righe, finisce col generare una stringa fortemente impredicibile col risultato che l'aleatorietà del segnale è incrementata così come potremo anche dedurre dalla dimostrazione originale di cui daremo conto in appendice.

Capitolo IV

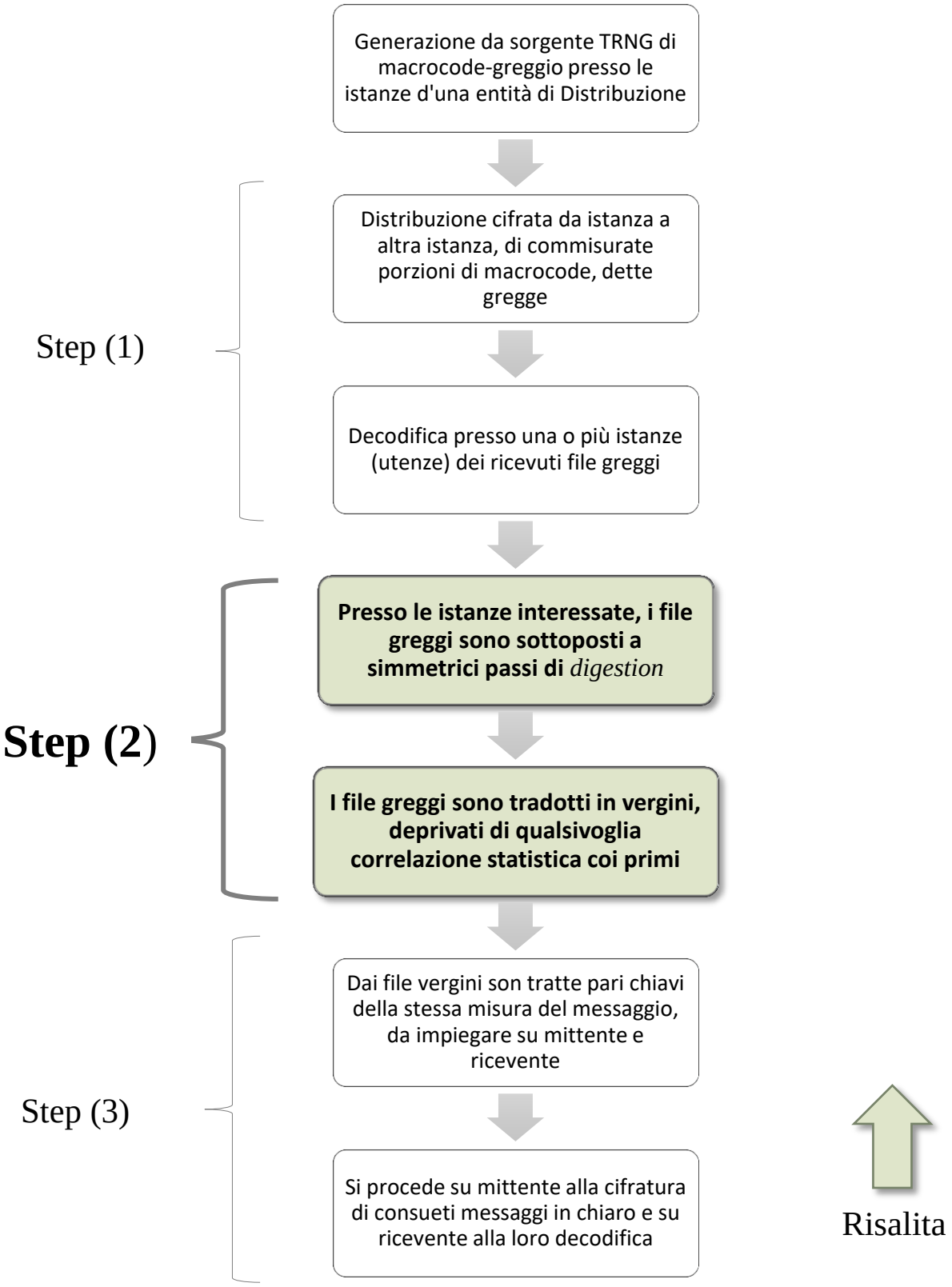
Tema

(27)

Come nelle mappe all'ingresso della Stazione Metropolitana, si mostra in figura dove siamo arrivati nel senso che stiamo tuttora trattando di problematiche che riguardano il cosiddetto secondo Step.
Per il resto, il capitolo è interamente dedicato alla questione della perdita di memoria e cioè al fatto che le operazioni di *digestion* sono in grado di recidere nelle sequenze vergini, ogni reminiscenza statistica di quelle gregge.

In proposito, è data dunque dimostrazione dell'assunto così da soddisfare quanto avevamo annunciato (b).

Figura 06 – Discesa e Risalita ma nei diversi Step



Poca Memoria

(28)

Dopo aver trattata la questione non ancora esaurita della crescente impredicibilità del segnale sottoposto a operazioni in XOR **(a)** e volendo adesso soffermarci sul diverso requisito della mancanza di memoria dei flussi ottenuti in uscita in fase di *digestion* **(b)**,
intendiamo intanto provare che il cosiddetto segnale vergine discende da operazioni che non gli consentono di serbare memoria di quello greggio.
Sappiamo infatti che nel rispetto di corollario (**D**) tale segnale greggio composto da un flusso random trasmesso da remoto**,** non sarà impiegato una seconda volta nemmeno in forma di chiave crittografica ed è perciò necessario procedere alla sua trasformazione che contempla la resezione di ogni reminiscenza statistica.
A partire intanto da quanto anticipammo in tempi non sospetti (10) dove fu detto che "*sebbene la memoria statistica si possa attenuare, non tutto è tecnicamente perduto, motivo per cui sarebbe bene dichiarare in anticipo quali siano i dati sensibili che si intendono offuscare e quali non lo sono"* per prima cosa vediamo perciò di definire quali sarebbero tali dati.
Ci siamo infatti soffermati sul perché i valori del file greggio debbano rimanere coperti (non dovendo esser collegati a quelli del file vergine, ingaggiata per trarre un segnale da usare una volta soltanto) ma ciò non significa che tale obbligo riguardi qualsivoglia informazione e, più specificamente, il flusso *no-random* $\boldsymbol{x}$ che col nome di "linea decorrelata" sarà impiegato nella somma con due o più parti del file greggio $\boldsymbol{y}$ di variabile $\boldsymbol{Y}$
Essa linea di cui celermente dicemmo (21) e di cui nuovamente diremo (30) non entra in gioco come oggetto da mascherare per la semplice ragione che non avendo carattere aleatorio, sarà suscitata *off-line* sulla base di pochi elementi condivisi.
Viepiù per tale ragione, non sarà mai distribuita su un canale di collegamento essendo che sorgerà presso ciascuna utenza *quando e dove possa servire* senza prestare il fianco a attacchi che interessano i flussi trasmessi punto punto[49].
Possiamo in effetti confessare che nella logica del nuovo cifrario saranno infatti da proteggere *e da proteggere in condizioni di perfetta sicurezza*[50] tutte e indistintamente le informazioni trasmesse da remoto sia in forma di chiave che di messaggio,

[49] Per dare un'idea sebbene alquanto fantasiosa, facciamo l'esempio del caso in cui – presso due istanze in collegamento tra loro – ci sia un uguale elenco di opere letterarie.
Essendo che si tratta di informazioni già note che possono essere ricostruite sulla base di una quantità minima di dati, sarà sufficiente trovare una intesa sul titolo per condividere una lunga sequenza come sarebbe *per dire* quella dei tre libri della Divina Commedia; nel caso non sarà il testo dei tre libri a dover essere protetto, giacché sarebbe in ogni modo sufficiente la difesa del titolo.

[50] Secondo i lemmi enunciati e secondo il teorema che andremo a illustrare a conclusione dell'ultimo capitolo.

mentre i dati da archiviare per un nano-secondo o per secoli, presso singole utenze o presso altri costrutti, potranno pur essere o meno cifrati ma tale questione esula dall'oggetto del corrente lavoro.

.

Il Piccione Viaggiatore

(29)

Continuando a battere ancora per un momento su talune questioni, abbiamo nondimeno la possibilità di scongiurare una volta per tutte quello che sarebbe un facile equivoco;
noi come mossa del cavallo abbiamo infatti riportata in apertura (01) una definizione semplice ma rigorosa di sicurezza perfetta[51], che in buona sostanza distilla quella pronunciata da Shannon in Communication Theory of Secrecy Systems:

Perfect Secrecy is defined by requiring of a system that after a cryptogram is intercepted by the enemy the a posteriori probabilities of this cryptogram representing various messages, be identically the same as the a priori probabilities of the same messages before the interception.
La segretezza perfetta è definita richiedendo a un sistema che, dopo che un crittogramma sia stato intercettato dal nemico, le probabilità a posteriori di questo crittogramma che rappresenta vari messaggi, siano identiche alle probabilità a priori degli stessi messaggi prima dell'intercettazione.

Per quanto ciò sia notevole e non occorre spaccare il capello per doverlo sottolineare, tuttavia non significa che un simile stato sia alieno da qualsivoglia attacco malevolo e tale fatto va rimarcato per mettere ordine in ciascuna casella del sistema;
bisogna infatti operare una drastica distinzione tra più problematiche giustamente e comunemente poste su piani diversi:

- **(1)** da una parte, abbiamo il problema della cosiddetta sicurezza crittografica che si esplica, non solo ma soprattutto, in fase di trasmissione;
- **(2)** dall'altra, abbiamo la distinta questione del controllo degli accessi che si esplica nella difesa di aree riservate.

Una messaggeria protetta secondo criteri di perfetta sicurezza, non potrà soccombere ai tentativi di decifrare informazioni in uscita una volta che siano state cifrate,
ma resta sensibile agli assalti alle informazioni non ancora coperte o già decifrate o perseguibili in chiaro attraverso particolari privilegi.
Affrontare il problema di una crittografia efficace ed efficiente, significa difendere allo stremo messaggi cifrati ma non significa lavorare su quelli non-cifrati che investono altri e diversi livelli di segretezza sia materiale (avverso attacchi fisici come nel caso di una rapina a mano armata) che cibernetica (avverso attacchi informatici come per una identità clonata).

[51] **Nella seconda parte tuttavia vedremo che non è impossibile rimodulare tale definizione muovendo da una diversa prospettiva.**

Pertanto a tal proposito ci piace dire d'una storia che abbiamo sentita chissà dove ma che tratta in apparenza di tutt'altro, dicendo delle avventure di un emiro e del suo piccione viaggiatore che trasmetteva messaggi oltre le linee nemiche, permettendo tuttavia a noi comuni mortali di cogliere l'occasione di spiegare come si possano riconoscere molteplici e multiformi livelli di sicurezza.

Come sappiamo dai racconti dei pellegrini che andavano dal mare alla montagna e dalla montagna al mare, la crittografia è arte antica, per cui non meraviglia che principi e condottieri ne abbiano fatto uso dal Califfato di Cordova alle corti di Navarra e Aragona a partire dalla notte dei tempi.
Si narra infatti di un emiro che possedeva un piccione invulnerabile, capace di resistere alle frecce che gli erano scagliate da ogni dove.
Una notte, egli dettò una lettera all'amanuense per inviare un messaggio al suo generale che avrebbe dovuto assalire i bivacchi cristiani al sorgere della luna nascente.
Per mantenere segreto tale dispaccio lo nascose tra le piume del suo invincibile piccione che – sebbene colpito da frombole e dardi avvelenati – passò indenne il fronte potendo raggiungere il fiume dove stava il grosso delle truppe che avrebbero dovuto condurre nottetempo l'assalto.
Il generale lesse la lettera col messaggio che nessuno aveva potuto carpire, ma quando giunse l'ora convenuta i suoi uomini caddero ugualmente in una imboscata giacché i cristiani erano stati nondimeno avvertiti.
Alcuni dissero ch'era stato il generale a tradire.
Altri l'amanuense che aveva scritto la lettera sotto dettatura, altri ancora lo stesso emiro impazzito per gli occhi assassini d'una principessa reale.
Ma come fu, come non fu, l'emiro perse prima la guerra e poi la testa in quanto fu presto condotto al patibolo sulla pubblica piazza dove all'ultimo istante scorse uno stormo di bianche colombelle levatesi in volo.

Fuor di metafora il piccione può dunque rappresentare la crittografia che cela il messaggio in chiaro e che, se perfetta, non può essere abbattuta.
Il generale è immagine delle criticità conducibili alla più ampia categoria della sicurezza cibernetica che somma alla crittografia informatica la questione del controllo degli accessi (nella disponibilità di chiunque abbia legittimi privilegi come quelli verosimilmente attribuiti al generale) mentre la follia dell'emiro e il tradimento dello scrivano rappresentano il caso di quelle minacce materiali che includono al loro interno ogni possibile punto di caduta.

Senza Memoria (b)

(30)

Chiudendo tuttavia col seducente mondo delle favole per tornare coi piedi per terra, cominciamo col dire che per fare giustizia d'ogni fraentendimento,
è bene chiarire le cose affermando che se il valore di file vergine dovesse risultare compatibile con ciascun equi-probabile valore di file greggio di lunghezza *L*,
avremmo dimostrato che ogni correlazione tra l'una e l'altra sequenza è recisa non essendovi modo d'invertire la funzione di stato del sistema.

E così dopo aver sgombrato il campo da questo e quello, intendiamo calcolare se da un determinato file sia possibile tornare al flusso di partenza e se di tale flusso permanga o meno qualche reminiscenza statistica.
Riportandoci così a quanto anticipato in (21) diciamo nuovamente ***y*** di variabile ***Y*** il file greggio di lunghezza *L* preso in ingresso dal sistema e ripartito in *p* parti che tuttavia qui supponiamo essere due per *p*=2
Diciamo ***ay, by*** tali parti di lunghezza *l* dove $l = {}^{L}/_{2}$
Diciamo $\tilde{\boldsymbol{y}}$ la risultante in uscita dalla somma XOR di parte ***ay*** con parte ***by*** [52]
per cui avremo $\boldsymbol{ay} \oplus \boldsymbol{by} = \tilde{\boldsymbol{y}}$ che essendo rilasciata dalla somma a flusso di tali operandi sarà anch'essa di lunghezza *l*
Diciamo ***x*** la linea decorrelata di pari lunghezza *l* da sommare a $\tilde{\boldsymbol{y}}$ di variabile $\tilde{\boldsymbol{Y}}$
Sarà dunque ***j*** il segnale vergine di lunghezza *l* rilasciato a sua volta in uscita.
Vediamo quindi nelle dimostrazioni a seguire, se tale sequenza vergine ***j*** mantenga o meno memoria di quella greggia ***y*** di variabile ***Y***

Dimostrazione (1)

In ultima analisi intendiamo pertanto accertare se la conoscenza del valore di ***j***, consenta o meno a un attaccante di accrescere la probabilità di tornare ad ***y***
Muovendo tuttavia in risalita come sarebbe nell'approccio tentato da un attaccante che voglia forzare il sistema, andiamo intanto a capire se conosciuta ***j***,
almeno aumenti la probabilità di compiere un primo passo cominciando col risalire da tale file vergine ***j*** a $\tilde{\boldsymbol{y}}$ di variabile $\tilde{\boldsymbol{Y}}$
Sequenza ***j*** di lunghezza *l* la cui misura sappiamo corrispondere alla metà della lunghezza *L* di file greggio ***y*** è infatti dato dall'output di somma $\tilde{\boldsymbol{y}} \oplus \boldsymbol{x}$

[52] E' ovvio che qui si suppone che file greggio ***y*** sia dato da un numero pari di simboli o bit;
se ciò non fosse sarebbe bene applicare qualche facile espediente, come quello di eliminare l'ultimo bit per pareggiare le parti della bipartizione, o altro.

dove $\tilde{y}$ di variabile $\tilde{Y}$ è una sequenza aleatoria data dalla somma tra le parti random ***ay, by*** della partizione di ***y***

e dove ***x*** di variabile ***X*** sarà invece un indipendente flusso *no-random* da noi battezzato col nome di *linea decorrelata*[53]

Sapendo allora che tali ***x*** e $\tilde{y}$ sono di pari misura binaria *l* dove $l = {}^{L}/_{2}$

costruiamo due insiemi:

insieme $\mathbb{X}$ di spazio Ω_x di variabile ***X***

insieme $\tilde{\mathbf{Y}}$ di spazio $\Omega_{\tilde{y}}$ di variabile $\tilde{Y}$

che sono di eguale numerosità in quanto in ciò dipendenti dalla pari lunghezza *l* dei rispettivi elementi che consente un pari numero di configurazioni binarie,

per cui $\#\mathbb{X} = \#\tilde{\mathbf{Y}}$

Avendo tuttavia supposto che ***j*** da un certo momento in avanti, sia conosciuta nel suo valore binario, siamo nella condizione di dire che vale *h* e che tale risultato è costante al variare di ***x*** ed $\tilde{y}$ esprimendo uno scalare fisso e non più variabile.

Se definiamo pertanto diversamente l'espressione di somma XOR dando per essa la funzione *f*: $\boldsymbol{x}, \tilde{\boldsymbol{y}} \rightarrow \boldsymbol{j}$

dove *h* valore ormai noto di ***j*** sarà a sua volta di lunghezza *l* in quanto output della somma a flusso tra due file di tale misura,

e dove $\boldsymbol{x}, \tilde{\boldsymbol{y}}$ sono giustappunto grandezze variabili con $\tilde{y}$ data da valori con probabilità uniforme di comparsa, possiamo altresì costruire insieme $\mathbb{Z}$ di tutte le possibili coppie di valori ***x*** ed $\tilde{y}$ che diano *h* come risultato.

Sapendo allora che elementi di $\mathbb{X}$ sono tutti i possibili valori di ***X*** (dove $\boldsymbol{X} = \boldsymbol{x}_1\, \boldsymbol{x}_2\, \boldsymbol{x}_3 \ldots, \boldsymbol{x}_n$) e quindi tutte le configurazione che si possono avere in sequenze di lunghezza *l*

e sapendo altresì che elementi di $\tilde{\mathbf{Y}}$ sono tutti gli equiprobabili valori (presi con pari frequenza di probabilità) assunti da variabile $\tilde{Y}$ (dove $\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{y}}_1\, \tilde{\boldsymbol{y}}_2\, \tilde{\boldsymbol{y}}_3 \ldots, \tilde{\boldsymbol{y}}_n$) e quindi tutte le configurazione che si possono avere in sequenze random della sua lunghezza medesima,

a sua volta uguale a quella di ***x*** di variabile ***X***

esisterà sempre per qualsivoglia valore di $\boldsymbol{x} \in \mathbb{X}$, uno ed un solo equiprobabile valore di $\tilde{y}$ appartenente a $\tilde{\mathbf{Y}}$ in grado di soddisfare la funzione dando tale *h* come risultato.

Per cui non solo $\#\mathbb{X} = \#\tilde{\mathbf{Y}}$ ma anche $\#\mathbb{X} = \#\tilde{\mathbf{Y}} = \#\mathbb{Z}$

Sapere che ***j*** di variabile $\boldsymbol{j}$ vale *h*, non influisce pertanto sulla facoltà di risalire ad $\tilde{y}$ la cui probabilità resta incondizionata

[53] Come trattato in altro documento (confronta "Metodo").

Esempio (1)

Riassumendo abbiam dunque specificato che, in ultima analisi, intendiamo accertare se ***j*** di variabile $\boldsymbol{j}$ abbia o meno memoria di ***y*** che sappiamo essere di lunghezza *L* per *L* che qui fissiamo a titolo esemplificativo in otto caratteri binari.

Intanto ci siamo tuttavia concentrati su un primo passaggio andando a vedere se da sequenza vergine ***j*** sia almeno possibile accrescere la probabilità di risalire ad $\tilde{\boldsymbol{y}}$ di variabile $\tilde{\boldsymbol{Y}}$ che si trova a mezza via d'un tentativo di risalita.
Diamo allora a titolo d'esempio a essa ***j*** che supponiamo ormai essere nota, un valore costante *h* dove *h* = 1111, assumendo che variabili ***X*** e $\tilde{\boldsymbol{Y}}$ di uguale lunghezza *l* dove $l = {}^{L}/_{2}$ siano perciò di lunghezza pari a quattro bit.

Assegniamo allora ogni possibile valore a variabile ***X*** che funge da primo operando
0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111,
1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111
e diamo ogni equiprobabile valore a variabile $\tilde{\boldsymbol{Y}}$ che funge da secondo operando
0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111,
1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111
Da ciò, avremo le seguenti sedici coppie di operandi XOR che, sommate modulo 2, daranno *h*= 1111 come risultato:
0000, 1111; 0001,1110; 0010, 1101; 0011, 1100; 0100, 1011; 0101, 1010; 0110, 1001; 0111, 1000;
1000, 0111; 1001, 0110; 1010, 0101; 1011, 0100; 1100, 0011; 1101, 0010; 1110, 0001; 1111, 0000
per la qual ragione, possiamo confermare che per ognuno dei possibili valori di ***X*** ci sarà sempre un equiprobabile valore di $\tilde{\boldsymbol{Y}}$ in grado di dare ***j*** = *h* = 1111 come risultato.

Da ciò quindi si evince che preso un qualche valore di ***j*** di variabile $\boldsymbol{j}$ che diciamo *h*, esso sarà sempre in accordo con qualsivoglia dei sedici possibili valori di $\tilde{\boldsymbol{Y}}$ e anche con ciascuna delle sedici coppie di incognite ***x***, ***y*** che soddisfino la funzione.

Dimostrazione (2)

Avendo dunque dimostrato che conoscendo il valore di sequenza vergine ***j*** non crescono le probabilità di tornare a $\tilde{\boldsymbol{y}}$ di variabile $\tilde{\boldsymbol{Y}}$

intendiamo procedere, chiedendoci cosa accade volendo risalire accapo dalla stessa $\tilde{\boldsymbol{y}}$ alle parti ***ay, by*** della bipartizione di ***y***.

Supponendo infatti con un mero espediente logico che $\tilde{\boldsymbol{y}}$ di variabile $\tilde{\boldsymbol{Y}}$ pur rimanendo ignota a chi attacca non essendo possibile rintracciarla, sia comunque da noi svelata nelle sue cifre binarie, potremmo pur dire che $\tilde{\boldsymbol{y}}$ vale *k* e che tale risultato è costante al variare delle incognite ***ay, by*** esprimendo uno scalare fisso e non più variabile[54].

Se dunque prendiamo ***ay*** che sarà un file dall'andamento aleatorio di lunghezza *l* in quanto parte della partizione del flusso di ***y*** ingenerato da una sorgente fisica,

e se prendiamo ***by*** che, per ugual ragione, sappiamo a sua volta aleatoria e della stessa lunghezza *l* di parte ***ay***

possiamo costruire due insiemi di pari numerosità in quanto in ciò dipendenti dall'eguale lunghezza numerica dei rispettivi elementi, tanto che avremo:

insieme $\mathbb{A}$ dello spazio $\Omega_{\mathbf{A}}$ di ***ay***

insieme $\mathbb{B}$ dello spazio $\Omega_{\mathbf{B}}$ di ***by***

dove $\#\mathbb{A} = \#\mathbb{B}$

Ora se definiamo diversamente l'espressione di somma XOR, dando per essa la funzione

f: ***ay, by*** $\rightarrow \tilde{\boldsymbol{y}}$

dove $\tilde{\boldsymbol{y}}$ di variabile $\tilde{\boldsymbol{Y}}$ vale *k* e dove ***ay***, ***by*** sono grandezze incognite di pari lunghezza *l*, soddisfatte da valori con probabilità uniforme di comparsa, e dove *k* sarà a sua volta di lunghezza *l*, in quanto output della somma a flusso in OR esclusivo delle medesime ***ay***, ***by***,

basterà ricalcare i passi della precedente dimostrazione (1) per concludere che conoscere *k* non fornisce indizi a chi attacca,

giacché il sapere che $\tilde{\boldsymbol{y}}$ vale *k* non accresce la probabilità di risalire ad ***ay*** $\in \mathbb{A}$ e ***by*** $\in \mathbb{B}$ la cui probabilità di comparsa resta incondizionata.

[54] Qui si parla di "mero espediente logico" in quanto assegnare ad $\tilde{\boldsymbol{y}}$ un valore costante, risponde a una strategia per mettere a fuoco la compatibilità che corre tra essa $\tilde{\boldsymbol{y}}$ e tutti i valori di ***ay***, ***by*** che la generano nella somma a flusso.

Esempio (2)

Se quindi nuovamente assumiamo sequenza grezza ***y*** di lunghezza *L* dove *L* viepiù corrisponda alla misura di otto bit, otterremo dalla sua bipartizione, parti ***ay***, ***by*** che saranno di pari lunghezza *l* dove *l* sarebbe della misura di quattro bit.
Volendo tuttavia riprendere il filo della narrazione dal punto in cui parte random ***ay*** e parte random ***by*** siano state sommate, rilasciando $\widetilde{\boldsymbol{y}}$ di variabile $\widetilde{\boldsymbol{Y}}$ come risultato, daremo alla medesima un valore *k* per cui diremo *k* = 0001

Assegniamo allora ogni possibile valore di ***ay*** che funge da primo operando
0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111,
1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111
e anche ogni equiprobabile valore di ***by*** che funge da secondo operando
0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111,
1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111
Si avranno le seguenti sedici equiprobabili coppie di operandi ***ay***, ***by*** che, sommati XOR tra loro, daranno $\widetilde{\boldsymbol{y}}$ = *k* = 0001 come risultato
0000 0001; 0001 0000; 0010 0011; 0011 0010; 0100 0101; 0101 0100; 0110 0111; 0111 0110;
1000 1001; 1001 1000; 1010 1011; 1011 1010; 1100 1101; 1101 1100; 1110 1111; 1111 1110
per cui per ogni valore di ***ay*** ci sarà sempre un equiprobabile valore di ***by*** in grado di dare 0001 come risultato.

In ragione di tale esempio si può tuttavia anche dire da una diversa prospettiva, che, preso come valore *k* = 0001 ma pure una qualsiasi delle altre quindici possibili permutazioni 0000, 0010, 0100, ..., 1111[55], essa si potrà accordare con qualsivoglia dei sedici equiprobabili valori di ***ay*** e con ciascuno dei sedici equiprobabili valori di ***by*** oltre che con le sedici coppie di incognite ***ay***, ***by*** che variamente soddisfino a seconda dei casi la funzione.

[55] Invero, oltre a 0001, in partenza potremmo assegnare a *k* un qualsiasi valore tra quelli possibili, per cui in questo senso possiamo pur dire $k = k_1\, k_2\, k_3 \ldots k_{16}$

In altre parole *k* è costante, ma dal solo momento in cui assume effettivamente un valore!

Dimostrazione (1)(2)

Intanto continuiamo a tenere sempre a mente che la nostra ultima ma primaria istanza ha riguardato e riguarda la questione se segnale vergine ***j*** serbi o meno memoria di ***y***
dove ciò in concreto significa chiedersi se conoscendo ***j*** o una sua frazione numerica,
un attaccante possa o meno riuscire a assumere indizi sufficienti a far crescere la probabilità di venire a capo di ***y***
Ci si vuole infatti sincerare che ***y*** prodotta in origine da una sorgente fisica true random e distribuita in forma cifrata su un canale di collegamento,
non sia erroneamente trasmessa una seconda volta sebbene in maniera residuale (in veste di chiave crittografica tratta da sequernza ***j***).
Mettendo tuttavia assieme una dopo l'altra, le due distinte dimostrazioni che abbiamo prodotte, possiamo osservare che *andando da valle a monte* dal valore in uscita di ***j*** a quello in ingresso di ***y***, ha comunque luogo quanto segue:

1) in risalita ciascun valore di ***j*** si potrà accordare con ogni equiprobabile valore di $\tilde{\boldsymbol{y}}$ giacché ciascuna $\tilde{\boldsymbol{y}}$ di lunghezza *l*, sommata al relativo valore ***x*** di pari lunghezza *l*, è in grado di soddisfare la funzione dando *h* come risultato;
2) a sua volta, ciascun valore di $\tilde{\boldsymbol{y}}$ si potrà selettivamente accordare con qualsivoglia delle distinte coppie di incognite ***ay***, ***by*** che diano in uscita un valore reso costante dal momento in cui sia effettivamente fissato.

Perciò, ciascuna $\tilde{\boldsymbol{y}}$ di variabile $\tilde{\boldsymbol{Y}}$ non si concilia con qualsivoglia coppia di incognite ***ay***, ***by*** che diremo appartenere a $\mathbb{K}$ dove $\mathbb{K}$ è un insieme la cui numerosità è data da tutti gli *n* possibili valori di parte ***ay*** moltiplicati a tutti gli altrettanto *n* possibili valori di parte ***by*** di pari lunghezza *l*

Ciascuna $\tilde{\boldsymbol{y}}$ di variabile $\tilde{\boldsymbol{Y}}$ tuttavia s'accorda con tutti i valori appartenenti ad uno ed uno soltanto dei sottonsiemi di tale $\mathbb{K}$ per $\{\mathbb{C}_1\, \mathbb{C}_2\, \mathbb{C}_3 \ldots \mathbb{C}_n\} \subset \mathbb{K}$
essendo che appartengono a tali sottoinsiemi le coppie ***ay***, ***by*** che danno *k* come risultato, a condizione di assumere $k \stackrel{\text{def}}{=} k_1\, k_2\, k_3 \ldots\, k_n$ a seconda del valore ottenuto dallo XOR e del sottinsieme $\mathbb{C}_1 \ldots \mathbb{C}_n$ cui esso appartiene.
Essendo allora, che $\tilde{\boldsymbol{y}}$ di variabile $\tilde{\boldsymbol{Y}}$ e parti ***ay***, ***by*** del segnale greggio ***y*** sono tutte di pari lunghezza *l* dove $l = {}^{L}/_{2}$
ed essendo che lo spazio delle configurazioni di $\tilde{\boldsymbol{Y}}$ e quello delle coppie di parti effettivamente random ***ay***, ***by*** sono caratterizzate da pari cardinalità,
assumendo *k* uguale a $k_1\, k_2\, k_3 \ldots\, k_n$ almeno fin quando non sia effettivamente fissato nell'esito il suo valore binario,

avremo che sequenza vergine ***j*** di lunghezza *l* presa come output del processo di *digestion*, condurrà a ciascuno degli equi-probabili 2^n valori di $\widetilde{\boldsymbol{y}}$ di lunghezza *l*,

ognuno dei quali porterà a sue corrispondenti 2^n equiprobabili coppie di incognite ***ay***, ***by*** di lunghezza *l*, parti della bipartizione di ***y*** di variabile ***Y***, che possono distintamente dare *k* come risultato della loro addizione.

Motivo per cui per tale effetto moltiplicativo, ciascun valore di ***j*** sarà parimente in accordo con qualsivoglia delle configurazioni di ***Y*** di lunghezza *L*

Segnale vergine ***j*** non avrà dunque memoria di quello greggio ***y*** di variabile ***Y***

Esempi e Conclusioni (1)(2)

Per finire di illustrare quanto detto sinora, torniamo allora per l'ultima volta agli esempi trattati dove sappiamo di avere un file greggio di lunghezza *L* pari ad otto bit (da cui son tratte due parti uguali ***ay***, ***by*** di quattro bit ciascuna).

In tal caso avremo infatti che, procedendo in risalita dal segnale vergine a quello greggio, intanto qualsivoglia valore di ***j*** dato dalla somma XOR, sarebbe in accordo con qualunque dei sedici possibili valori attribuibili ad $\widetilde{\boldsymbol{y}}$ di variabile $\widetilde{\boldsymbol{Y}}$

dove ciascuno di tali valori sarà a sua volta compatibile con ogni oggetto di uno ed uno soltanto dei sedici distinti sotto-insiemi $\{\mathbb{C}_1\ \mathbb{C}_2\ \mathbb{C}_3\ \ldots\ \mathbb{C}_{16}\}$ di $\mathbb{K}$ cui appartengono tutte le coppie di operandi ***ay***, ***by*** parti della bipartizione di ***y***.

Tenendo dunque conto di tali passaggi e assumendo per dire $\boldsymbol{j} = h = 1111$

andando a ritroso avremmo ***j*** sempre in accordo con ciascuno dei sedici possibili valori di $\widetilde{\boldsymbol{y}}$ (0000, 0001, 0010..., 1110, 1111) per come diversamente si combinano col decorrelato segnale non random da noi annotato con ***x***.

Essendo perciò che ciascuna $\widetilde{\boldsymbol{y}}$ prima di essere definitivamente fissata nel suo contenuto binario, sarà data da un qualche valore da noi detto $k = k_1\ k_2\ k_3 \ldots\ k_{16}$

dove ciascuno di tali valori discenderà dalla somma di parte ***ay*** con parte ***by*** appartenente a uno ed uno soltanto dei relativi sotto-insiemi $\{\mathbb{C}_1\ \mathbb{C}_2\ \mathbb{C}_3\ \ldots\ \mathbb{C}_{16}\}$ di insieme $\mathbb{K}$,

allora abbiamo che $\widetilde{\boldsymbol{y}} = k_2 = 0001$ potrebbe essere ad esempio compatibile con le seguenti sedici coppie ***ay***, ***by*** del relativo sotto-insieme

0000 0001; 0001 0000; 0010 0011; 0011 0010; 0100 0101; 0101 0100; 0110 0111; 0111 0110;
1000 1001; 1001 1000; 1010 1011; 1011 1010; 1100 1101; 1101 1100; 1110 1111; 1111 1110

ma se dovessimo invece dare $\widetilde{\boldsymbol{y}} = k_5 = 1000$ essa sarà diversamente in accordo con le seguenti altre e diverse sedici coppie

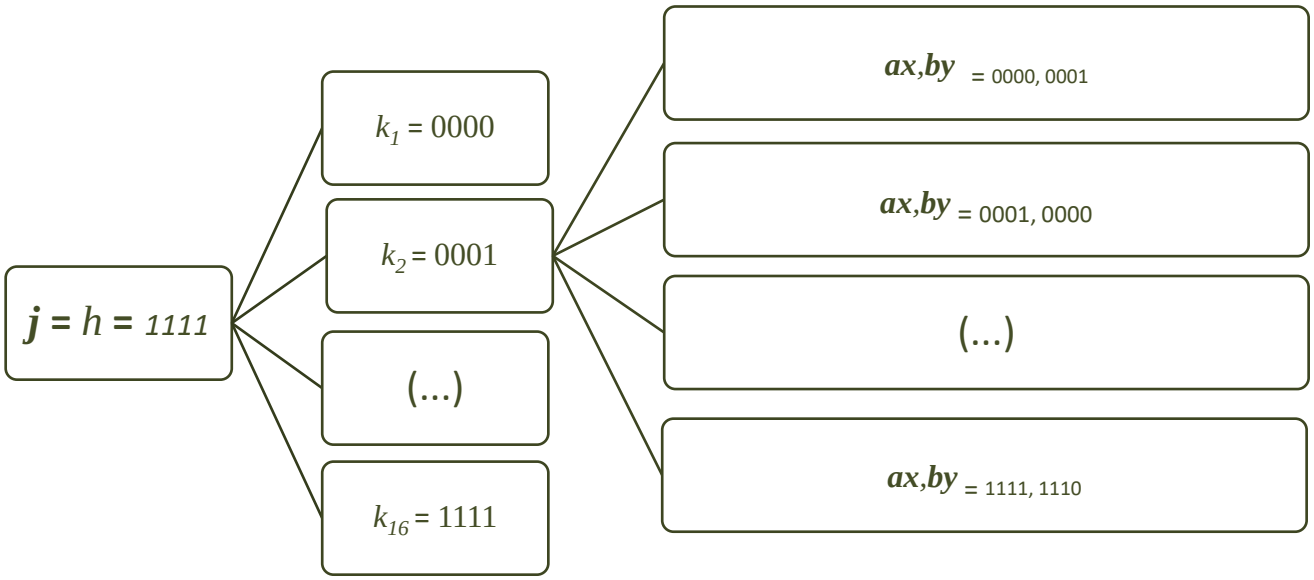
0000 1000; 0001 1001; 0010 1010; 0011 1011; 0100 1100; 0101 1101; 0110 1100; 0111 1111;
1000 0000; 1001 0001; 1010 0010; 1011 0011; 1100 0100; 1101 0101; 1110 0110; 1111 0111

per cui, assunto dalla visuale di chi attacca un valore come potrebbe essere quello per cui $\boldsymbol{j} = h = 1111$
abbiamo che tale $\boldsymbol{j}$ sarà sempre in accordo con qualsivoglia delle sedici permutazioni che possono valere per $\widetilde{\boldsymbol{y}}$ di variabile $\widetilde{\boldsymbol{Y}}$
e attraverso queste, con ciascuna delle duecento cinquantasei equiprobabili configurazioni[56] di $\boldsymbol{ay}$, $\boldsymbol{by}$
dalla cui concatenazione si torna a segnale greggio $\boldsymbol{y}$
così che qualora si venga a conoscenza che $\boldsymbol{j}$ vale $h = 1111$, **o**gni valore di $\boldsymbol{y}$ di variabile $\boldsymbol{Y}$ resterà alfine non solo possibile ma anche parimente probabile.

[56] Diversamente si può pur dire che ogni configurazione di sequenza random $\boldsymbol{y}$ corrisponde ad una diversa concatenazione di $\boldsymbol{ay}$, $\boldsymbol{by}$

Figura 07

Dove ciascun valore, di cui si venga a conoscenza, di sequenza vergine ***j*** come ad esempio sarebbe $h = 1111$, indifferentemente conduce a tutti e sedici i possibili valori $\widetilde{y}$ di variabile $\widetilde{Y}$ ognuno dei quali a sua volta conduce a sedici distinti valori dei duecento cinquantasei di ***y*** di variabile ***Y***



Avendo allora che ciascuna ***y*** di variabile ***Y*** sarà data dalla concatenazione dei relativi possibili valori di ***ax,by***
prendendo per dire ***aY***= 0000, ***bY*** = 0001, avremmo ***Y*** = 00000001
prendendo invece ***aY*** = 0001, ***bY*** = 0000, avremmo ***Y*** = 00010000 e così andando,
dove in un tentativo di risalita, ***j*** = *h* = 1111 sarebbe sempre compatibile con tutti i valori ***y*** di variabile ***Y***
per cui ***j*** non concede alcuna informazione aggiuntiva a un attaccante sull'andamento del segnale greggio da cui pure discende,
il che significa che sequenza vergine ***j*** non ha memoria statistica di quella greggia ***y*** presa originariamente in accesso dal sistema

Appendice

Supplemento n. 1
Supplemento n. 2

Tematiche

Qui si riprendono alcuni temi di cui alla fase di *digestion*
ed alle trasformazioni che comporta nel passaggio dalle frasi gregge a quelle vergini.

(1) Supplemento
In particolare si dimostra che il file vergine rilasciata in seguito ai passi di *digestion*,
rispetto a quello greggio, presenta un delta percentuale di differenza tendenzialmente uguale
allo stesso che s'avrebbe con qualsivoglia sequenza random presa a caso;

(2) Supplemento
In termini molto più generali si torna al tema della qualità del segnale che cresce
in seguito a passi di somma XOR di sequenze numeriche che non siano tra loro correlate.
Di tale assunto è data dimostrazione.

Supplemento n. 1

(Sequenze tra loro Indistinguibili)

Una delle questioni sinora trattate, ha riguardato la rescissione d'ogni correlazione statistica che passi tra la sequenza detta file vergine ***j*** e quella detta file greggio ***y*** generata da sorgente aleatoria.
Come sappiamo, la questione assume rilievo volendo noi essere certi che le condizioni dedotte da corollario (**D**) siano rispettate,
per cui ai fini della perfetta sicurezza del sistema una sequenza dal carattere random (chiave o messaggio che sia) non sarà impiegata più d'una volta.
Affinché sequenza ***y*** e sequenza ***j*** si possano considerare distinte l'una dall'altra dopo i passi di trasformazione dei quali dicemmo (*digestion*), è opportuno che non espongano correlazioni residue tali da far sospettare di condividere eccedenti pattern numerici.
E perciò a tale scopo, è necessario che sequenza ***j*** abbia due precipui attributi da noi precedentemente fissati[57]:

- **(i)** un delta percentuale di differenza da sequenza greggia ***y*** tale da renderla indistinguibile da qualsivoglia altra e diversa sequenza presa a caso;
- **(ii)** nessuna memoria di ***y***

Pertanto per quanto appena detto in **(i)** il flusso vergine di ***j*** dovrà palesare un andamento così difforme dall'originale da esibire con ***y*** e quindi con le sue parti ***ay by*** *eguaglianze* e *differenze* comparabili a quelle che mediamente si hanno con qualsivoglia stringa presa a caso.
Ricordando tuttavia le facili istruzioni che conducono da file ***y*** a file ***j***,
occorre dire che queste sono talmente concise da far dubitare che possano recidere tanto profondamente il nodo gordiano tra sequenza e sequenza.

Come infatti sappiamo, ***y*** *di lunghezza L sarebbe generata da sorgente aleatoria prima d'esser bipartita in due parti* ***ay,by*** *di pari lunghezza* $^{L}/_{2}$ *da concatenare XOR tra loro,*
così da dare in uscita $\tilde{\boldsymbol{y}}$ *a sua volta sommata a una decorrelata linea detta x di lunghezza l =* $^{L}/_{2}$ *dalla cui somma modulo due sarà alfine ottenuta sequenza vergine* ***j***

Volendo stimare pertanto l'efficacia di tali passi, ci chiediamo se si sia possibile dimostrare che rispetto a ciascuna parte di ***y***,
j effettivamente palesi in termini di differenza, un andamento indistinguibile da ciascun flusso random preso alla cieca.

[57] In (08) si fa in effetti anche riferimento al carattere random del segnale, ma si tratta d'una questione da tenere distinta e che riprenderemo nel prossimo supplemento.

Dimostrazione

(**Ipotesi Ultima**)

In riferimento a quanto detto nel corrente lavoro, diamo pertanto in ipotesi che file vergine ***j*** presenti rispetto alle parti di ***y***
un delta percentuale di differenza comparabile a quello di qualsiasi altra e diversa sequenza *di pari lunghezza binaria* assunta in modo aleatorio (diciamo pure ..., di qualsiasi sequenza presa con uniforme distribuzione di probabilità dal suo insieme di riferimento).
Dove il delta è qui calcolato confrontando a flusso le sequenze di pari lunghezza binaria *l* di cui abbiamo detto, registrando bit a bit le quantità di simboli uguali e diversi e computando le conseguenti percentuali.
Dove per parti intendiamo tanto parte ***ay*** che parte ***by*** – della bipartizione di ***y*** – che sono random in quanto sezioni d'uguale lunghezza *l* d'un flusso binario rilasciato da fonte aleatoria.

(**Ipotesi di Partenza**)

Volendoci tuttavia focalizzare in primo luogo su parte ***ay*** della bipartizione di ***y***, intanto diciamo che per la somma a flusso con ***by***,
risultante $\tilde{\boldsymbol{y}}$ (che da un canto discende dalla somma di ***ay*** con ***by*** e dall'altra partecipa alla formazione di ***j***) a sua volta esporrà un delta rispetto ad ***ay*** prossimo a *un mezzo di uno* che sarebbe il valore che mediamente si ottiene dal confronto tra distinti flussi numerici.
In altro modo possiamo pur dire che, sapendo che ciascun bit di $\tilde{\boldsymbol{y}}$ sarà di segno 0 o 1
abbinando a flusso un corrispondente bit 0, 1 di parte ***ay***
avremo che come sarebbe per il contemporaneo lancio di due monete, si otterranno
per percentuali prossime alla metà di uno coppie di valori uguali a 00, 11 da noi dette "eguaglianze"
e per percentuali prossime all'altra metà di uno coppie di valori 01, 10 da noi dette "differenze".

(**Dimostrazione**)

Andiamo pertanto a costruire la famiglia ordinata **Z** dei simboli di valore 0 di parte ***ay***, tenendo conto che tali elementi sono da selezionare – oltre che per il valore binario che li caratterizza – anche per la posizione occupata in sequenza che fornisce al segno un ulteriore attributo.
Ragion per cui il primo 0 sarà distintamente indicizzato rispetto al secondo, e questi rispetto al primo ed al terzo, e così andando su tutta la sequenza di ***ay***
così come sarebbe, supponendo che ***ay*** possa esser fatta delle seguenti cifre binarie
0001011101101010001001100101011011010110011101100101101101100101101101
dalle quali i soli simboli di valore zero, sarebbero estratti per fungere da elementi di **Z**
0001011101101010001001100101011011010110011101100101101101100101101101

(i)

Ora avendo che a ogni elemento di valore 0 di $\boldsymbol{ay}$ appartenente a $\mathbf{Z}$,

si abbina nella somma a flusso un corrispondente valore 0 o 1 di $\boldsymbol{by}$

e avendo altresì che anche la sequenza numerica di $\boldsymbol{by}$

esporrà valori che mostrano una pari frequenza di comparsa essendo parte di sequenza $\boldsymbol{y}$ generata da sorgente aleatoria,

avremo che la somma di ciascun bit di valore 0 di parte $\boldsymbol{ay}$ con ciascun bit di valore 0 od 1 di parte $\boldsymbol{by}$, non potrà che confermare o invertire ciascun valore 0 presente in $\boldsymbol{ay}$ con una frequenza prossima a quella ideale di un mezzo su uno.

Diremo pertanto PUT i singoli risultati rilasciati da tale operazione, come ad esempio si avrebbe sommando i seguenti bit di valore 0 di $\boldsymbol{ay}$

0001**0**111**0**11**0**1**0**1**0**00**0**1**00**11**00**1**0**1**0**1**0**11**0**1**0**1**0**11**00**111**0**11**00**1**0**11**0**11**0**11**00**1**0**11**0**11**0**1

coi seguenti bit corrispondenti per posizione nella somma a flusso, ma allocati in $\boldsymbol{by}$

1011**0**011**1**01**1**00**1**0**1**0**1**10**0**01**1**01**1**0**1**00**1**1**0**1**1**00**1**0**1**0**1**0**1**11**0**0**1**0**0**01**1**0**1**0**1**0**0**011

così da avere tali somme che daranno in uscita altrettanti valori di PUT

0⊕1; 0⊕0; 0⊕1; 0⊕0; 0⊕1; 0⊕1; 0⊕0; 0⊕1; 0⊕1; 0⊕0; 0⊕1; 0⊕0; 0⊕1; 0⊕1; (…)**; 0⊕1**

(ii)

Andando tuttavia a fissare due sotto-famiglie di $\mathbf{Z}$ che diremo $\mathbf{Z}_1$ e $\mathbf{Z}_2$

i cui elementi saranno dati in $\mathbf{Z}_1$ da ciascun segno "0" di parte $\boldsymbol{ay}$, al quale sia associato un segno "0" di parte $\boldsymbol{by}$ nella somma a flusso,

e in $\mathbf{Z}_2$ da ciascun segno "0" di parte $\boldsymbol{ay}$, cui sia invece associato un diverso segno "1" di parte $\boldsymbol{by}$,

considerando che nel primo caso, la somma darà un valore PUT pari a "0" e nel secondo un valore PUT pari a "1" potremo perciò scrivere la funzione:

$$\textit{Per cui } \boldsymbol{f}(\text{PUT}) = \begin{cases} 0 \textit{ se } \text{PUT} \in \mathbf{Z}_1 \\ 1 \textit{ se } \text{PUT} \in \mathbf{Z}_2 \end{cases}$$

(iii)

Se prendiamo tuttavia in esame anche la diversa famiglia ordinata $\mathbf{U}$ dei valori di segno "1" di parte $\boldsymbol{ay}$, potremo anche fissare due sotto-insiemi $\mathbf{U}_1$ e $\mathbf{U}_2$ scrivendo in tal caso l'analoga funzione:

$$\textit{Per cui } \boldsymbol{f}(\text{PUT}') = \begin{cases} 1 \textit{ se } \text{PUT}' \in \mathbf{U}_1 \\ 0 \textit{ se } \text{PUT}' \in \mathbf{U}_2 \end{cases}$$

(iv)

Dobbiamo perciò convenire che, essendo che – tanto i bit di segno "0" presi da ***ay*** e appartenenti a **Z** che quelli di segno "1" a loro volta presi da ***ay*** ma appartenenti a **U –** saranno confermati o invertiti in somma XOR con una frequenza prossima a quella di *un mezzo su uno* per la uniforme probabilità di comparsa di ciascun simbolo binario di parte random ***by*** che farà da secondo operando,
risultante $\tilde{\boldsymbol{y}}$ che ne discende non potrà che tendere a un delta percentuale di differenza d'*un mezzo su uno* da essa ***ay***

(v)

Se tuttavia proviamo altresì a confrontare parte ***ay*** con una qualsivoglia sequenza effettivamente random di egual misura binaria che chiameremo ***cy*** e che è stata a sua volta indipendentemente genearata da fonte aleatoia,
non potremo che avere un delta egualmente prossimo alla *metà di uno* giacché,
al confronto bit a bit, a ciascun "0" come a ciascun "1" corrisponderà con pari probabilità di comparsa un segno uguale o diverso da quello di frase ***cy*** per la natura aleatoria della medesima.

(vi)

In percentuali di differenza,$\tilde{\boldsymbol{y}}$ sarà dunque indistinguibile da altra sequenza random presa a caso,
in quanto mostrerà una pari frequenza di probabilità d'esporre simboli uguali o diversi da quelli di parte ***ay*** e – a voler ripercorrere i passi della presente dimostrazione – anche da quelli di parte ***by***
Non essendovi dunque modo rispetto ad ***y*** (la cui sequenza è costituita dalla semplice concatenazione di parte ***ay*** con parte ***by***), di distinguere sequenza $\tilde{\boldsymbol{y}}$ da qualsivoglia sequenza aleatoria ***cy*** presa a caso (presa con pari frequenza di comparsa dall'insieme di riferiento di tutte le configurazioni di lunghezza *l*)
viepiù dopo un'ulteriore somma con linea decorrelata ***x***
men che meno potremmo distinguere la risultante frase vergine ***j***
essendo improbabile che la somma di $\tilde{\boldsymbol{y}}$ con ***x***, possa invertire la freccia dell'entropia facendo decrescere il delta tra sequenza a sequenza[58].

[58] **Si veda in proposito il supplemento a seguire.**

Supplemento n. 2

(Incremento della Qualità del Segnale in Uscita)

Qui nella corrente appendice, intendiamo riprendere in modo originale la questione della qualità aleatoria del segnale; qualità preservata e semmai accresciuta nella fase detta di *digestion* che comporta più somme XOR

Così come abbiamo avuto modo di ribattere persino nel precedente supplemento, nel nuovo metodo di crittazione da illustrato altrove in dettaglio,

il primo passo cui è tuttavia sottoposta una sequenza grezza generata da sorgente fisica, comporta la partizione in due parti da sommare XOR tra loro.

Quindi, se prendiamo di $\boldsymbol{y}$ quella parte $\boldsymbol{ay}$ della bipartizione che farà da primo operando della somma a flusso, essa sarà impegnata in talune operazioni dove è sommata *modulo due* a altra sequenza random $\boldsymbol{by}$ e poi a una sequenza *no-random*.

Si è dunque affermato che tali passi non solo non comportano alcun deterioramento dei flussi rilasciati dalla sorgente, ma hanno come conseguenza quella d'abbattere il *bias* favorendo un aumento della impredicibilità del segnale.

Abbiamo infatti ricordato che "a fundamental lemma of Yao states that computational weak-unpredictability of Boolean predicates is amplified when the results of several independent instances are XOR together" (*On Yao's XOR Lemma, Abstract, Oded Goldreich, Noam Nisan, Avi Wigderson). "Un fondamentale lemma di Yao afferma che la debole impredicibilità dei predicati booleani* è *amplificata quando i risultati di più istanze indipendenti siano sommati XOR tra loro*"

Noi intendiamo riproporre allora adesso tale tema ma da una diversa angolazione, fornendo una dimostrazione originale dell'assunto per cui la somme XOR tra sequenze non correlate tra loro, tendenzialmente ingenera un incremento della impredicibilità-aleatorietà dei flussi in uscita.

Tutto questo perché lo XOR tra distinte sequenze tende a favorire il passaggio da uno stato relativamente certo a uno più incerto, facendo prevalere nella somma a flusso una maggior impredicibilità che equilibra verso l'alto l'entropia del sistema.

Non daremo pertanto una nuova e diversa prova dello XOR lemma di Yao, variamente e autorevolmente comprovato nel corso degli anni,

ma una più immediata dimostrazione dell'ipotesi secondo cui la somma XOR tra più stringhe indipendenti *che presentino una quantità originariamente ignota di incertezza* determina una flusso maggiormente impredicibile rispetto a quello in accesso.

Dimostrazione

(**Ipotesi**)

Diciamo in ipotesi che lo XOR tra sequenze di pari misura binaria non correlate tra loro, determina a tendere un incremento della quantità di incertezza e quindi di aleatorietà del segnale in uscita.

(Premessa)

Prendiamo pertanto due sequenze tra loro decorrelate che rispettivamente diciamo $\tilde{\boldsymbol{a}}$ e $\tilde{\boldsymbol{b}}$ di pari lunghezza S e con una quantità media di incertezza ignota,
e andiamo a mapparle attraverso un parametro che indichiamo col simbolo silllabico ***pi***, che ci consenta di poterle ripartire.
In breve, attraverso tale parametro ***pi*** si andrà infatti a fissare:

- **(a)** una lunghezza standard **s** sufficientemente lunga sebbene inferiore di *n* volte rispetto a quella S dell'intera sequenza $\tilde{\boldsymbol{a}}$ oppure $\tilde{\boldsymbol{b}}$ che sia, per cui avremo s <S
- **(b)** le *n* parti di lunghezza **s** in cui sono rispettivamente ripartite tanto sequenza $\tilde{\boldsymbol{a}}$ che sequenza $\tilde{\boldsymbol{b}}$
- **(c)** una grandezza che pesi, entro determinati limiti di tolleranza, la quantità di incertezza di ciascuna parte di sequenza $\tilde{\boldsymbol{a}}$ rispetto a ciascuna corrispondente parte di sequenza $\tilde{\boldsymbol{b}}$ e viceversa[59].

Ciò ci consente di selezionare su tali sequenze, tanto regioni (date da una o più parti tra loro contigue, per come saranno a breve definite) che diremo meno-incerte[60] che corrispondenti regioni che diremo più-incerte[61] su cui si operi a flusso con le prime.
Si potranno altresì isolare regioni che diremo quasi-neutre quando il confronto tra regione e regione dell'una e dell'altra sequenza,
manifesti quantità comparabili di incertezza.
In altre parole si suppone che ciascuna regione della sequenza di $\tilde{\boldsymbol{a}}$ od invece di $\tilde{\boldsymbol{b}}$ sia variamente stimata rispetto a quella corrispondente di $\tilde{\boldsymbol{b}}$ o piuttosto di $\tilde{\boldsymbol{a}}$
a seconda della maggiore o minore incertezza e quindi del carattere più o meno omogeneo delle frequenze di comparsa di ciascun segno binario.

[59] La quantità di incertezza e la conseguente impredicibilità, qui non è stimata rispetto a qualche standard assoluto od a una classe di complessità che ne fissi il livello,
ma in modo relativo e reciproco mettendo a confronto l'incertezza di regioni tratte da $\tilde{\boldsymbol{a}}$ con quella di corrispondenti regioni tratte da $\tilde{\boldsymbol{b}}$ o viceversa.

[60] Dove i segni binari 0,1 potranno cioè risultare più sbilanciati come nel caso di una moneta truccata che dia più frequentemente testa di croce.

[61] Dove i segni binari 0,1 risulteranno bilanciati come nel caso di una moneta non truccata che tenda perciò a far uscire testa o croce con pari frequenza di comparsa.

Figura 09

Sequenza $\widetilde{\boldsymbol{a}}$ *(sequenza che funge da primo operando di somma XOR, suddivisa in più regioni variamente evidenziate)*

000000000110000110000110000110000110000110000110000110100100110110100101101001010010101101001001101101001
0110110101000100100110110100111111010101100101001110010010000100010100100100110110100101101001010010010001
1011010010110100101001010111010110100110001010010010011011010010110100100100100101001001001101101001011010
0011101101101001011010010100101110011011010010110100101001110101011010100101101001110010010110010100100101
01101101001011010010000100100110110110101101000101001010111011101010011101001101010101010111010110010111
0
10011010101001010011011100110010101100101001001

Sequenza $\widetilde{\boldsymbol{b}}$ *(sequenza che funge da secondo operando dello XOR, suddivisa in più regioni variamente evidenziate)*

111111001010011010101010010010010110011011100100101110001010110000100001000101000111111111011111111111111
11111111011011010010000100101000000001011010010110100111011100110110100100110100101001001001101101001011
0
1000011110101111111111111110001000000000000110111100101001001001101101001011010000000000000000000000001111111
11110000000000000000010111001000000000000100000001000000001001111111111101111101111111101000100000000111111101
11111010101010101011110011111110111111101111111111000111111011101011001001111111111111111111111111111111111
1111111111111110100101101100110110100101100100 01

Voci

Sequenze: sequenze binarie $\widetilde{\boldsymbol{a}}, \widetilde{\boldsymbol{b}}$ di pari lunghezza S da sommare a flusso tra loro

Parti: moduli di misura sufficientemente lunga **s** dove tuttavia **s** <S
in cui potremmo idealmente dividere ciascuna sequenza $\widetilde{\boldsymbol{a}}, \widetilde{\boldsymbol{b}}$
e che abbiamo fissati in esempio di misura pari a sei bit[62] che per semplicità consideriamo un sottomultiplo di S

Regione: porzione di sequenza – che sia sequenza $\widetilde{\boldsymbol{a}}$ o invece $\widetilde{\boldsymbol{b}}$ **–** composta da una o più parti contigue, che saranno o tutte meno-incerte o tutte più-incerte fin quando non cambi il loro stato di maggiore o minore incertezza, fatto questo che fissa il confine da una regione all'altra.
Avendo tuttavia in esempio, ciascuna parte di lunghezza **s =** sei bit ciò significa che nel caso si potranno mappare solo regioni di sei bit o multiple di sei bit.

Le regioni meno-incerte evidenziate in nero, son dunque quelle i cui flussi numerici saranno giustappunto stimati con un inferiore livello di entropia e di omogeneità del segnale,
mentre le corrispondenti più-incerte evidenziate in giallo, denotano flussi maggiormente incerti (dove ciascun segno 0 o 1 mostra una maggiore uniformità di comparsa rispetto a quanto avremmo sulla corrispondente regione dell'altra sequenza).
In altre parole essendo che qui si parla d'una incertezza relativa, se ad esempio una qualche regione in $\widetilde{\boldsymbol{a}}$ appare più incerta la corrispondente su $\widetilde{\boldsymbol{b}}$ dovrà necessariamente risultare meno incerta.
Le regioni quasi-neutre, che in esempio non sono rimarcate in alcun modo, saranno quelle che entro dati limiti di tolleranza, appaiono d'un pari livello di incertezza su entrambe le sequenze numeriche $\widetilde{\boldsymbol{a}}$ e $\widetilde{\boldsymbol{b}}$**.**

In altre parole le regioni saranno classificate come meno-incerte, più-incerte o quasi-neutre a seconda di come risultino nel confronto con le corrispondenti,
nel senso che dove abbiamo una regione più incerta in $\widetilde{\boldsymbol{a}}$ avremo una regione meno incerta in $\widetilde{\boldsymbol{b}}$ e viceversa.

Corrispondenti: quando diremo di regioni o parti "corrispondenti" o anche di bit corrispondenti, faremo riferimento a regioni o parti o singoli bit tra loro impegnati (in conseguenza della posizione nel flusso) in reciproche operazioni di somma XOR.

[62] Invero tale misura di sei bit sarebbe ampiamente insufficiente ma noi, a fini illustrativi, abbiamo sovente preferito ragionare su valori di minor lunghezza rispetto a quelli attesi nella vita reale

(**Dimostrazione**)

Tenendo dunque conto del senso delle premesse e delle voci appena illustrate, abbiamo che volendo considerare ciascuna regione indifferentemente presa da sequenza $\tilde{\boldsymbol{a}}$ come da sequenza $\tilde{\boldsymbol{b}}$, confrontandola con quella tratta dal corrispondente flusso da sommare XOR (dove se per dire, l'una regione è tratta da $\tilde{\boldsymbol{a}}$ la corrispondente sarebbe tratta da $\tilde{\boldsymbol{b}}$) avremmo che si potrà verificare una delle seguenti condizioni:

(a) a una regione meno-incerta andrà necessariamente a corrispondere una regione più-incerta;

(b) a una regione quasi-neutra corrisponderà altra regione quasi-neutra.

Intanto diciamo che, essendo che ad una regione quasi-neutra non potrà che corrispondere altra regione quasi-neutra, intuitivamente l'entropia del segnale in uscita dalla somma XOR, non subirà variazioni significative, sebbene non sarebbe difficile dimostrare che persino nel caso si produce un tendenziale aumento di incertezza.

(i)

Noi preferiamo tuttavia soffermarci sul caso più significativo, e cioè su quello in **(a)** da cui appare lecito attendersi effetti più netti per la maggior differenza che corre sulle corrispondenti regioni tra rispettive quantità di incertezza.

Costruiamo perciò una tabella che diremo "master" con due successioni numeriche che diremo $Riga_1$ $Riga_2$ e che saranno fatte di bit assunti nel seguente modo:

(1) su $Riga_1$ si trascrivono i bit indifferentemente presi sia da $\tilde{\boldsymbol{a}}$ che da $\tilde{\boldsymbol{b}}$ di ciascuna regione indicata come meno incerta

(2) su $Riga_2$ sono invece trascritti i bit indifferentemente presi tanto da $\tilde{\boldsymbol{a}}$ che da $\tilde{\boldsymbol{b}}$ di ciascuna corrispondente regione detta più incerta**.**

Si vanno così progressivamente a formare due righe di bit disposti sopra o sotto in conseguenza della loro minore o maggiore incertezza e quindi della minore o maggiore omogeneità statistica dei flussi numerici, così come sarebbe in un esempio del seguente tipo:

Tabella *Master*

$Riga1$	0	0	1	1	0	0	1	1
$Riga2$	1	0	0	1	0	1	0	1

(ii)

Procediamo estraendo ora da tali righe ogni possibile coppia di bit costituita dai seguenti valori in colonna che saranno trascritti sempre in colonna su una diversa tabella $\mathbf{t}_A$:

(1) ciascun bit uguale a 0 preso da *Riga*1 muovendo da sinistra a destra e saltando i bit di segno 1;

(2) ogni corrispondente bit, sia esso uguale a 0 oppure a 1 tratto da *Riga*2 muovendo da sinistra a destra.

Tabella tA								
	0	0	...	...	0	0	···	···
	1	0	...	...	0	1	···	···

Diremo allora $A \in \mathbb{A}$ tali coppie prese in colonna e battezzate come $A_1\ A_2\ \dots\ A_n$ che saranno variamente formate dai simboli **01** e **00** e cioè da coppie di bit che cominciano con 0 e possono finire tanto con 0 che con 1.

Ciò fatto, estraiamo dalla precedente tabella $\mathbf{t}_0$ che sarebbe una sorta di tabella master da cui tutto discende, anche le altre coppie di bit in precedenza ignorate e diversamente costituite a da:

(3) ciascun bit uguale a 1 preso da $Riga_1$ procedendo da sinistra a destra e saltando i bit di segno 0;

(4) ciascun corrispondente bit 0 oppure 1 tratto da $Riga_2$ muovendo da sinistra a destra.

Tabella tB								
	...	...	1	1	...	...	1	1
	...	...	0	1	...	...	0	1

Diremo pertanto $B \in \mathbb{B}$ tali coppie da noi chiamate $B_1\ B_2\ \dots\ B_n$ che saranno formate dai valori incolonnati di 10 o 11 e cioè da coppie di bit che cominciano con 1 e finiscono a seconda dei casi tanto con 0 che con 1.

Per comprendere il senso delle operazioni sinora svolte, per inciso diciamo che le coppie ottenute e messe in colonna, sono tutte e solo coppie di bit che sarebbero da sommare nella somma a flusso tra sequenza $\tilde{\boldsymbol{a}}$ *e sequenza* $\tilde{\boldsymbol{b}}$*.*

Noi abbiamo solo talora cambiato l'ordine degli addendi ponendo nella prima riga in alto – sia nella tabella master che in quelle $\mathbf{t}_A$ $\mathbf{t}_B$ *che ne discendono – tutti bit provenienti da regioni meno incerte;*

abbiamo altresì diversamente allocate le coppie che cominciano con 0 e quelle che cominciano con 1 avendo cura di lasciare nondimeno vuote sulle rispettive tabelle, le celle i cui valori erano in origine diversamente incolonnati (con coppie che sul master cominciavano con 1 invece che 0, e coppie che cominciavano con 0 invece di 1).

Ciò significa che sovrapponendo una tabella all'altra non solo si ottiene nuovamente il cosiddetto master ma i risultati degli output sarebbero gli stessi, in egual modo ordinati, della somma a flusso tra sequenza $\tilde{\boldsymbol{a}}$ *e sequenza* $\tilde{\boldsymbol{b}}$*.*

Riportiamo allora per maggior chiarezza, il novero delle coppie che in esempio si sono formate come sopra.

Possibili coppie $A_1\ A_2\ \dots\ A_n$ che sono oggetti appartenenti a una famiglia ordinata detta $\mathbb{A}$

00

0 (tratto da *Riga*1);

0 (tratto da *Riga*2),

dove nella somma a flusso, avremmo 0 (nel primo operando) 0 (nel secondo operando)

01

0 (tratto da *Riga*1);

1 (tratto da *Riga*2),

dove nella somma a flusso, avremmo 0 (nel primo operando) 1 (nel secondo operando)

Possibili coppie $B_1\ B_2\ \ldots\ B_n$ che sono oggetti appartenenti ad una famiglia ordinata detta $\mathbb{B}$

10

1 (tratto da *Riga*1);

0 (tratto da *Riga*2),

dove nella somma a flusso, avremmo 1 (nel primo operando) 0 (nel secondo operando)

11

1 (tratto da *Riga*1);

1 (tratto da *Riga*2),

dove nella somma a flusso, avremmo 1 (nel primo operando) 1 (nel secondo operando)

(iii)

Per quanto detto si avrà pertanto che in somma XOR, a ogni valore 0 preso da regione meno incerta sarà abbinato un valore più equamente distribuito di segno 0 o 1 tratto dalla corrispondente regione più incerta, e ciò comporta che ciascun segno 0 tratto da regione meno incerta sarà confermato o invertito con una probabilità che non è dichiarata in assoluto,
ma che tuttavia rispetto a quella presa in input sarebbe tendenzialmente più prossima all'ideale d'un mezzo su uno.

Essendo poi che sappiamo che le coppie binarie dette B appartenenti a $\mathbb{B}$ sono a loro volta costituite da un bit 1 originariamente tratto da regioni meno incerte,
e un corrispondente bit 0 o 1 tratto da regioni più incerte,
necessariamente abbiamo che in somma XOR a ogni valore 1 preso da regione meno incerta, corrisponde un valore più equamente distribuito di segno 0 o di segno 1 proveniente da regione più incerta,
e ciò significa che anche il relativo segno 1 sarà confermato o invertito con una probabilità più prossima a quella ideale d'un mezzo su uno.

(iv)

In conclusione fissiamo allora due ulteriori famiglie ordinate che diremo $\mathbb{AB}$ e $\mathbb{BA}$ i cui elementi 0 o 1 saranno dati dai risultati di somma XOR che hanno luogo tra bit corrispondenti, diversamente appartenenti ad $\mathbb{A}$ o $\mathbb{B}$

Per cui apparterranno ad $\mathbb{AB}$ i risultati delle somme di coppie di egual valore binario 00 od 11 che dicemmo "eguaglianze" e che sono rispettivamente tratte tanto da $\mathbb{A}$ che da $\mathbb{B}$

E apparterranno a $\mathbb{BA}$ i risultati delle somme di coppie 01 o 10 che dicemmo "differenze" e che sarebbero a loro volta tratte da $\mathbb{A}$ e $\mathbb{B}$

Dicendo allora OUT i risultati di somma XOR adesso illustrati, possiamo anche scrivere la seguente funzione per cui:

$$f(\mathrm{OUT}) = \begin{cases} 0 \text{ se } \mathrm{OUT} \in \mathbb{AB} \\ 1 \text{ se } \mathrm{OUT} \in \mathbb{BA} \end{cases}$$

Essendo che tuttavia sappiamo che ogni coppia di bit da impegnare nella somma tra sequenza $\tilde{\boldsymbol{a}}$ e sequenza $\tilde{\boldsymbol{b}}$, espone una probabilità di appartenere all'uno o all'altro insieme $\mathbb{AB}$, $\mathbb{BA}$,

statisticamente più uniforme rispetto a quella offerta dalle regioni di provenienza che abbiamo dette meno incerte,

pure la frequenza di comparsa dei risultati 0 e 1 tenderà a uniformarsi alla qualità del segnale delle regioni relativamente più incerte (che pertanto palesano una più uniforme frequenza di probabilità dei simboli 0 o 1) ovunque esse siano.

Sommando perciò a flusso due stringhe $\tilde{\boldsymbol{a}}$ e $\tilde{\boldsymbol{b}}$ di pari misura binaria, abbiamo che i valori in uscita esporranno a tendere più incertezza e quindi una maggiore sparsificazione rispetto a quelli in entrata, giacché ogni output sarà costantemente modellato attorno alle frequenze meglio distribuite.

Seconda Parte

Si comincia con l'affrontare quanto attiene al cosiddetto primo Step (ultimo in risalita)
proseguendo con una disamina della definizione standard di perfetta sicurezza e del teorema di Shannon.
Se ne mette in luce l'incompletezza dimostrando che la traccia sinora offerta in letteratura
sul tema della segretezza incondizionata fa insorgere non poche contraddizioni.
Avremo dunque modo di offrire una definizione alternativa prima di giungere nell'ultimo capitolo alla
dimostrazione di un nuovo teorema.

Capitolo V

Tema

(31)

Quando ci siamo dilungati su alcune istruzioni (21)(30) abbiamo mostrato come attraverso opportuni passi di *digestion,*

da un canto i flussi numerici preservano o persino accentuano la loro qualità[1],

e dall'altro, smarriscono ogni correlazione statistica con quel segnale greggio da cui pure provengono.

Ora vorremo fotografare il momento in cui in uno stato iniziale del sistema tale segnale è trasferito, da una qualche entità di distribuzione[2] a ciascuna delle istanze di codifica, decodifica e trasmissione di messaggi cifrati.

Essendo tuttavia che la sicurezza di tale passaggio è data dalla correttezza dei lemmi (**B**)(**C**) ed essendo che questi sono un completamento del teorema di Shannon, procederemo con ulteriori analisi critiche del saggio intitolato alla *Communication Theory of Secrecy Systems*.

[1] Nel caso sono da considerare più "qualitativi" in quanto più prossimi ai valori attesi, i flussi fortemente impredicibili.

[2] Entità che potrà essere tanto centrale che distribuita, ovvero tanto inserita in un sistema centralizzato che in un sistema peer to peer.

Step (1)

(32)

Mimando allora lo sforzo di un attaccante che intenda risalire la corrente muovendo dal crittogramma al messaggio, abbiamo che tutto ha luogo a partire dalla distribuzione di sequenza greggia ***y*** generata da una sorgente aleatoria che fornirà la materia bruta da cui trarre chiavi di qualsivoglia lunghezza.

Invero il trasferimento d'un segnale random dal quale *sebbene in seguito a passi di trasformazione* saranno ricavate chiavi di misura almeno pari a quella arbitraria del messaggio ..., è un passaggio di particolare rilievo che intendiamo sviscerare non tanto sul piano d'una generica illustrazione quanto su quello delle sue implicazioni teoriche.

Intanto diciamo che il successo di tale "passaggio" dipende da due fattori che effettivamente sarebbero:

(1) la qualità del segnale generato dalla sorgente d'informazione;

(2) il fatto che sia confermata la correttezza dei lemmi (**B**)(**C**) così da provare l'efficacia delle procedure di crittazione del segnale greggio[3].

Sul primo punto c'è poco da dire nel senso che ci si dovrà opportunamente affidare a una sorgente qualitativa,

ma pure sul secondo avremmo poco da aggiungere sapendo dell'esperimento mentale condotto e della dimostrazione di cui diremo;

ma nel riprendere quanto altrove trattato[4], intendiamo tornare sui nostri passi ricalcando il tracciato condotto da Shannon sulle cui fondamenta abbiamo promesso di costruire il corpo delle nostre congetture.

.

[3] Procedure solo accennate nel corrente lavoro.

[4] *A new implementation of an extension of One Time Pad cryptography model*, Report di Claudio Cappelli, 2020.

Sicurezza 1949

(33)

Alla quarantaduesima casella *come nel gioco dell'oca* occorre tornare alla prima e cioè a quella definizione di segretezza perfetta per la quale la conoscenza di quanto vale il crittogramma nulla aggiunge alla probabilità di risalire al messaggio.

Se pure qualcuno dovesse riuscire nell'intento di intercettare il messaggio cifrato non vedrebbe aumentare le facoltà di giungere a quello in chiaro non avendo modo di sommare nuova informazione a quella che già possiede (01).

Altresì diciamo che fu sempre il vulcanico padre della teoria dell'informazione a scrivere che *a colui che egli chiamava bellicosamente nemico e noi chiamiamo Oracolo* non sarebbe stata *a priori* concessa altra informazione se non quella della lingua parlata da mittente (sebbene sia bene aggiungere che in altre parti del suo saggio sulla *Communication Theory of Secrecy Systems*, lo stesso Shannon suppone che chi attacca sappia del critto-sistema adottato).

If this is applied to the normal english, the cryptanalyst being assumed to have no knowledge of the message source other than that it is producing English text, the a priori probabilities of various messages of N *letters are merely their relative frequencies in normal English text.*

Se si applica ciò all'inglese, presumendo che il critto-analista non abbia alcuna conoscenza della fonte del messaggio a parte il fatto che generi testo inglese, le probabilità a priori di vari messaggi di N *lettere sono semplicemente le loro frequenze relative all'inglese standard.*

Qui il ricordo delle esperienze condotte col gruppo di volenterosi di cui abbiamo narrato (24), quando si voleva calcolare la misura dell'aleatorietà presente *in the normal English*, trapela in tutta evidenza nelle righe del saggio di Shannon.

E in effetti egli rimarca un fatto dirimente essendo che lucidamente ci ricorda di come esista una probabilità pregressa che sarebbe quella su cui conta il nostro Oracolo prima d'aver intercettato il crittogramma, e una probabilità condizionata accresciuta dall'eventuale conoscenza del valore del medesimo.

Nei sistemi di crittazione "convenzionali" usualmente detti a sicurezza computazionale,

si crea infatti uno stato per cui un attaccante dotato di sufficienti risorse, una volta che abbia saputo il valore del crittogramma ..., avrebbe modo di capire che la distribuzione di probabilità *a posteriori* di uno dei possibili messaggi tenderà ad 1 e cioè a quella che abbiamo indicata come la certezza che un evento succeda,

mentre quella degli altri tenderà a 0 che corrisponde al caso dove è impossibile che essi messaggi siano rilasciati dal sistema.

Il meccanismo algebrico che mette in moto tale processo sarà oggetto di molta cura, ma per adesso ci limitiamo a dire che un attaccante dotato di risorse formidabili come il mago dell'abracadabra, a furia di calcoli vedrebbe ingigantita la probabilità di comparsa di un solo messaggio e rimpicciolita quella degli altri che andrebbe man mano a scemare.

For N *fairly large, there is nearly always a unique solution to the cipher; i.e., a single good English sequence which transforms into the intercepted material (…).*
As material is intercepted, the cryptanalyst calculates the a posteriori probabilities; and as N *increases the probabilities of certain messages increase, and, of most, decrease, until finally only one is left, which has a probability nearly one, while the total probability of all others is nearly zero.*
Per N *abbastanza grande, c'è quasi sempre una soluzione unica al codice; cioè una singola buona sequenza inglese che si trasforma in materiale intercettato (…).*
Poco a poco, il crittoanalista calcola le probabilità a posteriori; e all'aumentare di N *le probabilità di certi messaggi aumentano e, nella maggioranza, descrescono fin quando alla fine ne rimane solo uno che ha una probabilità quasi 1, mentre la probabilità in totale degli altri è quasi 0.*

Se tali congetture fissano il carattere (ma anche il tallone d'Achille) di quei sistemi di crittazione da noi detti "convenzionali"
parrebbe facile ricavare una definizione alternativa di sicurezza perfetta, per la quale tale sicurezza ci sarebbe quando un attaccante non possa individuare un set sufficientemente ristretto di messaggi che incrementino la loro probabilità a discapito degli altri.
Non è tuttavia questa la strada intrapresa da Shannon e dai suoi epigoni che definiscono i sistemi perfetti nel modo da noi in precedenza citato.

Cuori Indomabili

(34)

Il fatto singolare è che quando si dice *come in effetti si dice* che la conoscenza del crittogramma non accresce nei sistemi perfetti la probabilità che un attaccante giunga al messaggio,
ciò cela un problema giacché l'informazione che il crittogramma rilascia può essere affievolita, o resa innocua, o riguardare fatti che non rilevano come quando la residua traccia sia tenuta nascosta con qualche ingegno ...**,** ma mai soppressa.
Il punto è che come abbiamo avuto modo di dire, effettivamente esiste un'informazione irriducibile insita nel crittogramma.

Poniamo allora che nella favoletta che intendiamo narrare, un Oracolo in possesso di risorse sovrumane intenda indovinare una formula magica che ha la facoltà di aprire un forziere, ma che proprio per questo è stata coperta attraverso tali inganni da risultare illeggibile.
Egli è riuscito a intercettare la corsa di un araldo che conduceva tale messaggio segreto, ma essendosi impossessato della missiva, non potette far altri che ammettere d'essersi impadronito d'uno scritto impossibile; uno scritto che gli avrebbe tuttavia consentito di congetturare sulla lunghezza del messaggio e quindi sul numero dei caratteri dell'abracadabra.

Un *core* indomabile, attiene infatti all'impronta che ciascun crittogramma rilascia sulla misura di chiave e messaggio.

Pur volendoci fermare al caso della perfetta sicurezza per come trattata nella lettera del teorema (**A**) abbiamo che nella somma a flusso dove $\boldsymbol{m} \oplus \boldsymbol{k} = Critto$, la misura di *Critto* discende da quella dell'operando di maggior lunghezza che qui sarebbe quello della chiave (tralasciando per ora quelle situazioni che pure impongono un diverso rapporto di misura tra chiave e messaggio).
Avendo infatti che nel caso $\boldsymbol{k}$ sarà necessariamente di lunghezza L ugual-maggiore di quella di $\boldsymbol{m}$ avendo
$L \geq l$
di conseguenza si converrà che dalla conoscenza di *Critto*, potrebbe un attaccante dedurre la massima lunghezza possibile che *in siffatta occasione* coincide con quella della chiave di misura almeno uguale a quella del messaggio[5].
Per cui se chiave e messaggio fossero uguali come sovente accade nei sistemi perfetti, la misura di *Critto* ci fornirebbe una prova delle loro profondità; viceversa qualora fossimo davanti al caso dove la chiave è più lunga e il messaggio meno,
pur non potendo calcolare la lunghezza di $\boldsymbol{m}$, nondimeno avremmo fissato un limite nel senso che più di tanto il messaggio non può misurare.

[5] Si confronti (58) sul caso in cui si impieghino differenti e più complessi algoritmi.

Esempio

Primo Operando – Operando random di maggior lunghezza che potrebbe essere chiave o messaggio ma che muovendoci nell'ambito del teorema di caratterizzazione in (**A**) corrisponderà a quello della chiave
01101001101100100101101001010001101010011101001010100101101101001010101101001101101110011010011011001001011010010100011010100111010010101001011011010010101011010011010011101110001011 ...

Secondo Operando – Operando non random di minor lunghezza che corrisponde a quello del messaggio, il cui segnale sarà reiterato nella somma a flusso sino a pareggiare la misura dell'operando più lungo
00101111001011110010111100101111001011110010111100000000000000011110000001011110010111100101111001011110010111100000000000000011110000001011110010111100101111001011110010111100000000000000011110000001011110010111100101111001011110010111100000000000000011110000001011110010111100101111001011110010111100000000000000011110000 ...

Risultato in *uscita* – Output che fissa la sequenza numerica del crittogramma in uscita dalla somma XOR tra operando dato dalla chiave random e operando dato dal messaggio non random il quale sarebbe d'una lunghezza che non può superare quella di *Critto*
01000110100110110100101001011110100101101110001010100101101101101010010001101001101101001010010111101001011011100010101001011011010100100011010011011010010100101011101101110101111

Invero occorre tuttavia dire che l'informazione sulla misura del messaggio non appare neutra rispetto alla probabilità di risalire ad ***m*** e tanto appare in conflitto col presupposto assegnato in letteratura alla definizione di sicurezza incondizionata o perfetta; diversamente dovremmo pretendere che chi attacca, insieme alle informazioni comunemente concesse (1.*cifrario impiegato*; 2.*lingua scritta e parlata dal mittente*) sia surrettiziamente a conoscenza anche della lunghezza del messaggio di Alice che invece non conosce.
Quello che vorremmo altresì rimarcare è che, sapendo quanto vale il crittogramma e assumendo un approccio probabilistico, si palesa il fatto che un Oracolo venuto a conoscenza del valore di *Critto*, avrà minori probabilità di indovinare un dispaccio che può sensatamente supporre di lunghezza pari a quella dei quindicimila endecasillabi della Divina Commedia rispetto a un messaggio la cui misura non superi le due lettere alfabetiche!
Tale computo è agevole perché, se prendiamo una copia digitale della commedia dantesca espressa da N caratteri binari, e la confrontiamo con un dispaccio di n caratteri dove $n \ll N$,
avremo che un attaccante ha un'infinitesima probabilità su 2^N di indovinare il giusto messaggio quando corrisponda alle Cantiche della Commedia,
e una probabilità di uno sul minor spazio di 2^n in caso contrario.

Credere infatti che nella vita reale le probabilità di giungere al messaggio non crescono se si viene a conoscenza del valore del crittogramma,
significa trascurare che potendo restringere lo spazio degli eventi da un valore arbitrario a uno più ragionevole, vedremmo crescere *a posteriori* la fortuita probabilità di successo nello stesso modo in cui puntare sul terno e più facile che tentare la cinquina.

Messaggi Impossibili

(35)

Ciò detto in tanti potrebbero tuttavia obiettare che nella definizione di perfetta sicurezza prefigurata da Shannon e affinata negli anni, si dà per acclarato il richiamo all'insieme dei messaggi possibili (possibili quando? possibili per chi?) la cui numerosità sarebbe data dalla lunghezza del messaggio segretamente redatto da Alice.

Ma ciò nulla significa, giacché è proprio tale lunghezza a essere ignota a chi attacca ed a diventare nota *seppure con qualche incertezza* nel momento e solo nel momento in cui si viene a conoscenza del valore del crittogramma.

Se perciò si crede *come appare necessario credere* che non si possa ragionare della probabilità di comparsa di questa o quella alternativa senza conoscere i confini di qualche insieme di riferimento, sembra giusto chiedersi se il pregiudiziale richiamo a un astratto spazio dei messaggi sconosciuto a chi intenda forzare il sistema,

(1) sia o meno utile ai fini dichiarati,

(2) sia in accordo con le circostanze di fatto,

(3) possa o meno risultare arbitrario.

Each particular transformation of the set corresponds to enciphering with a particular key. The transformations are supposed reversible (non-singular) so that unique deciphering is possible when the key is known.

Each key and therefore each transformation is assumed to have an a priori probability associated with it—the probability of choosing that key. Similarly each possible message is assumed to have an associated a priori probability, determined by the underlying stochastic process.

Ogni particolare trasformazione dell'insieme corrisponde alla cifratura con una particolare chiave.

Le trasformazioni sono supposte reversibili (non uni-direzionali) in modo che sia possibile una decifrazione univoca quando la chiave è nota.

Si presume che ogni chiave e quindi ogni trasformazione abbia una probabilità a priori ad essa associata: la probabilità di essere accompagnata a quella chiave.

Allo stesso modo si presume che ogni possibile messaggio abbia una probabilità a priori associata, determinata dal processo stocastico sottostante.

Il fatto è che la probabilità a priori di cui parla Shannon con qualche ottimismo..., è calata dall'alto da lui stesso essendo impossibile da calcolare per chi attacca, prima di poter conoscere con qualche approssimazione le lunghezze di chiave e messaggio,

cosa che nei sistemi perfetti non avviene senza che sia svelato il valore di *Critto* col risultato che proprio tale evento determina un salto da prima a dopo;

un salto teoricamente negato sin dalla definizione la quale postula che il crittogramma nulla aggiunge alla probabilità assegnata in partenza, cosa coerente col presupposto adottato[6] ma non corrispondente a quanto poi accade.

Abbandonando perciò ogni remora, dobbiamo intanto dire che per il nostro Oracolo gli autentici "messaggi possibili" non sarebbero mai fissati *a priori* in quantità e lunghezza ma solo *a posteriori* e entro certi limiti, una volta che sia conosciuto il crittogramma.

[6] **Probabilità fissata in base alla misura del dispaccio redatto da Alice, ma che potrebbe esser nota a priori solo alla stessa Alice.**

Alice e il Crittografo

(36)

Quanto è stato tuttavia disposto da un capo all'altro del corrente lavoro, ci pone davanti a due fatti in apparenza contraddittori che potremmo idealmente disporre ai bandoli della matassa.

Da un canto a fondamento di qualsivoglia congettura è bene archiviare il fatto che non esiste attacco che consenta di risalire nei sistemi perfetti dal valore del crittogramma a quello di chiave e messaggio, essendo che *al verificarsi di determinate condizioni* alla classe dei messaggi non possono che essere riferite molteplici equi-probabili soluzioni.

It is possible to construct secrecy systems with a finite key for certain "languages" in which the equivocation does not approach zero as $N \to \infty$. *In this case, no matter how much material is intercepted, the enemy still does not obtain a unique solution to the cipher but is left with many alternatives, all of reasonable probability.*

Si possono costruire sistemi di segretezza con una chiave finita per certi "linguaggi" in cui l'incertezza non si approssima a zero come per $N \to \infty$. *In tal caso, non importa quanto materiale sia intercettato, il nemico non ottiene comunque una soluzione univoca al cifrario ma rimane con molte alternative, tutte di ragionevole probabilità.*

D'altro canto, possiamo tuttavia constatare che tanto non implica che le questioni di cui dicemmo nello scorsa parte come nella corrente ..., si possano efficacemente trattare alla luce della definizione standard da noi più volte ripresa**.**

Se volessimo analizzare le cose buttando un occhio al processo di crittazione e ai personaggi e interpreti che lo animano ..., che poi sarebbero quelli battezzati col nome di *Bob* e *Alice* e con quello di *Oracolo* per chi attacca al sistema ..., intanto vediamo emergere una interpretazione dei fatti che non suona come ce la siamo raccontata.

Spesso ad esempio si usa dire che presso utenza mittente e cioè presso Alice siano allocate le sorgenti di chiave e messaggio,

e così infatti leggiamo nel saggio di Shannon dove si accenna a una minima architettura in cui *"there are two information sources, a message source and a key source …"*

Sebbene tale dichiarazione possa apparire innocua ponendo la fonte del messaggio in un paniere e quella della chiave in un altro, nondimeno è fuorviante;

fin quando *mondo delle chiavi* e *dei messaggi* popolano distinte dimensioni quali sarebbero quella delle onde radio in bassa frequenza e quella di un messaggero che nasconda i codici nel doppiofondo d'una valigetta diplomatica …, l'immagine scaturita dalle cronache del novecento almeno rispecchia una situazione effettiva ...

This key is transmitted by some means,…,for example by messenger, to the receiving end. The message source produces a message (the "clear") which is enciphered and the resulting cryptogram sent to the receiving end by a possibly interceptible means, for example radio.
La chiave è trasmessa in qualche modo, (…) ad esempio tramite messaggero (…). La sorgente del messaggio produce un messaggio (in chiaro) che viene cifrato e il crittogramma risultante inviato all'estremità ricevente con un mezzo presumibilmente intercettabile, come ad esempio la radio.

Ragionando tuttavia a freddo e prendendo le misure a quanto concepito nel millenovecentoquarantanove dalla mente del nostro autore preferito colto nei suoi anni migliori ..., non possiamo che ammettere che molta acqua è passata sotto i ponti e tante cose sono cambiate tra il morire del vecchio e il sorgere del nuovo millennio;
a prescidere dal tema della perfetta sicurezza ..., è ormai acclarato che anche le chiavi sono trasmesse da remoto da Alice che può tranquillamente passare dai versi danteschi (*e alfine uscimmo a riveder le stelle*) al rumore dalle sorgenti impiegate nella costruzione di chiavi crittografiche.
Il preteso "message source" non può dunque esistere come fonte coesa, essendo che sull'utenza di Alice *ora* avremo file generati da un essere umano che ci siamo figurati mentre declama versi del dolce stil novo (***X***-source),
ora altro segnale parimente trasmesso in forma di messaggio ma ottenuto dalla conversione in digitale di un processo stocastico (***Y***-source).
E' d'altronde noto che è invalso l'uso di impiegare sistemi detti "a chiave asimmetrica" per tramettere chiavi in forma di messaggio,
e sistemi "a chiave simmetrica" per i dispacci veri e propri,
così che non solo le fonti ma persino i mezzi di crittazione potrebbero essere intrinsecamente diversi da messaggio a messaggio[7].
Per quanto detto sinora, poniamo dunque di avere una stazione detta *Alice* dotata d'una sorgente ***Y***-source che rilascia flussi aleatori ***y*** di variabile ***Y*** da impiegare come chiave di crittazione[8] da trasmettere in forma di messaggio;
supponiamo d'essere altresì in possesso di altra sorgente ***X***-source che potrebbe coincidere con una fonte umana capace di generare un segnale ***x*** di variabile ***X*** da noi indicato quale messaggio *no-random*.

[7] In realtà, non sempre le chiavi sono fatte di sequenze random o pseudo-random, ma pur sempre denotano un sufficiente livello di impredicibilità.
[8] A prescindere qui dal fatto che noi abbiamo previsto di trasmettere flussi random che non saranno impiegati come chiavi, ma implicati nella costruzione delle stesse.

(i)

Se allora diciamo che dalla stazione di *Alice* si vuole inviare un consueto messaggio di senso compiuto, nell'operazione saranno necessariamente coinvolte entrambe le sorgenti così da generare un segnale $\boldsymbol{x}$ di variabile $\boldsymbol{X}$ che funge da messaggio $\boldsymbol{m}$

e un segnale $\boldsymbol{y}$ di variabile $\boldsymbol{Y}$ che giunto a destinazione sarà successivamente impiegato (semmai dopo opportune trasformazioni) per fungere da chiave $\boldsymbol{k}$.

In effetti in uno stato iniziale del sistema quando è redatto il messaggio e selezionata la chiave, tale distinzione chiave-messaggio appare ancora giustificata;

ma quando si tratterà di procedere nelle attività di codifica in senso stretto, a essere ingaggiati saranno i valori numerici di $\boldsymbol{x}$ e quelli di $\boldsymbol{y}$ così come se prendiamo da una cesta *due mele* e *tre mele* e le sommiamo tra loro, in matematica saranno le grandezze a entrare in gioco e non le consistenze del frutto.

In altre parole possiamo anche dire che presso *Alice* si darà seguito a una funzione per la quale:

$$f:\boldsymbol{X},\boldsymbol{Y} \rightarrow Critto$$

(ii)

Se tuttavia con un cambio di paradigma, diciamo che dalla stazione di *Alice* si vuole piuttosto trasmettere il flusso di $\boldsymbol{y}$ semmai perché su quella di *Bob* sono in attesa di flussi random con cui costruire chiavi di crittazione,

e se volessimo invece impiegare come chiave, una sequenza $\boldsymbol{x}$ di variabile $\boldsymbol{X}$ avendo cura di sfruttare le opportunità consentite dai lemmi (**B**)(**C**),

sarebbe impiegata in codifica una funzione del tutto identica a quella di prima per cui anche stavolta potremmo scrivere:

$$f:\boldsymbol{X},\boldsymbol{Y} \rightarrow Critto$$

In tale passaggio algebrico siamo dunque passati dal dominio delle cose a quello delle grandezze e *in tale dominio* saranno calcolate in (**i**) come in (**ii**)

le stesse identiche funzioni che danno in uscita un pari crittogramma ottenuto dalla applicazione della stessa legge che *per eguali passi* non può che implicare eguali proprietà.

Alice e non solo

(37)

Quanto dunque illustrato altri non sarebbe che un diverso modo di formalizzare quel paradosso del crittografo messo in caldo da subito, così da chiudere il cerchio illustrando le attività di funzione eseguite da coloro i quali concorrono al funzionamento del sistema.

Avendo tuttavia rassettate le idee diremo nuovamente delle diverse fasi entro cui condurremo la partita che come tutte le partite ha un primo e un secondo tempo;

con un fermo immagine abbiamo infatti immortalato due precisi frangenti per cui diventa possibile sezionare lo spazio dove sorgono le operazioni di offesa e difesa del messaggio, e cioè *prima* e *dopo* che il crittogramma sia intercettato.

In effetti stiamo parlando di cose trattate poche pagine indietro ma che desideriamo recuperare giacché ci tornano utili a voler rimarcare la linea di demarcazione che discerne tale *prima* e tale *dopo* rispetto alla pratica della distribuzione a priori.

Per distribuzione a priori s'intende infatti la ripartizione di probabilità presa sulla scorta del bagaglio di conoscenze di chi attacca; conoscenze che non discendono da informazioni carpite in corso d'opera e tantomeno dalla lettura del crittogramma; più propriamente diciamo infatti che in un sistema crittografico la *distribuzione di probabilità a priori* è data da ciò che un attaccante conosce prima d'aver intercettato il crittogramma,

così che ai nastri di partenza possa assegnare una probabilità[9] all'evento costitutivo della cosidetta alternativa favorevole che *nel caso* coincide con quella di messaggio $\boldsymbol{m}_i$

Abbiamo già accennato al fatto che Shannon nel pesare quale fosse lo stato della contesa, in alcune parti del saggio suppose che il "nemico" avrebbe potuto contare sulla conoscenza del cifrario impiegato (con chiavi di lunghezza standard o più robuste, con algoritmi dell'uno o dell'altro tipo, con livelli dissimili di efficacia e efficienza),

oltre che sulla conoscenza della lingua ingenerata dalla sorgente che nelle sue congetture sappiamo essere la lingua inglese.

Sebbene Shannon sul punto sia stato contraddittorio, in effetti la natura del lessico ha rilievo ed ha rilievo in rapporto alla frequenza di ciascun simbolo alfabetico come sarebbero vocali e consonanti del linguaggio adottato.

Al nostro eroe la lingua madre (che noi usiamo convertire in simboli binari) e quindi quel *normal English* cui dedicò i maggiori sforzi (24) gli consentiva d'operare nella dimestichezza d'un ambiente domestico dove si muoveva con indubbia padronanza.

[9] E' inutile dire, che si tratta d'una probabilità estremamente bassa ma non uguale a zero, essendo paragonabile a quella di chi tenti di indovinare qualcosa alla cieca.

Sebbene sia bene aggiungere che la concessione in via convenzionale della lingua impiegata, lanciata al nemico da Shannon come un guanto di sfida ..., rimase un'eccezione piuttosto che la regola forse giacché calata in un teatro di guerra dov'è facile immaginare la lingua impiegata così come nei film coi marines che parlano in inglese ed i giapponesi in giapponese ma coi sottotitoli.
La crittografia diffusa con cui abbiamo oggi a che fare, concede tuttavia meno appigli in quanto sarebbe facile passare da questo a quell'idioma compromettendo l'idea maturata al tramonto di quel febbrile dopoguerra.
Noi nei capitoli a venire terremo tuttavia la barra al centro, dicendo di ulteriori fattori che includono l'altra e diversa proposizione secondo cui è legge capitale della moderna crittografia quella che detta il principio zecondo cui la sicurezza di un critto-sistema "non" dipende dal tener celato il cifrario ma dal tener nascosta la chiave.

To make the problem mathematically tractable we shall assume that the enemy knows the system being used. It might be objected that this assumption is unrealistic, in that the cryptanalyst often does not know what system was used or the probabilities in question.
There are two answers to this objection:
1. The restriction is much weaker than appears at first, (...)
2. The assumption is actually the one ordinary used in cryptographic studies. It is pessimistic and hence safe, but in the long run realistic, since one must expect his system to be found out eventually.
Per rendere il problema matematicamente trattabile, assumeremo che il nemico conosca il sistema in uso. Potrebbe essere obiettato che questa ipotesi non è realistica, in quanto spesso il critto analista non sa quale sistema sia stato utilizzato o le probabilità in questione.
Ci sono due risposte a questa obiezione:
1. La restrizione è molto più debole di quanto appaia all'inizio (...)
2. L'assunto è effettivamente quello normalmente utilizzato negli studi crittografici. È pessimista e quindi (proprio per questo) sicuro, ma a lungo termine realistico, dal momento che uno deve aspettarsi che il suo sistema sia alla fine scoperto.

Più Probabilità

(38)

Invero l'ipotesi per cui il critto-sistema è tuttavia considerato di dominio pubblico, vedremo che ha impatti diversi a seconda se stiamo operando nell'ambito della sicurezza computazionale o di quella perfetta.

Siccome tuttavia parliamo di qualcosa che incide sulla probabilità di comparsa dei messaggi ed essendo che *dopo quelli di impredicibilità e entropia* altri concetti bussano alle porte e appare difficile non tenerne conto, prima di proseguire vorremmo aprire una finestra su quelle definizioni di probabilità che non trovano la pace d'una idea condivisa.

Già quando dicemmo del semplice calcolo che confronta il caso di un crittogramma compatibile con la lunghezza d'un messaggio di quindicimila endecasillabi a quello compatibile con la misura d'un dispaccio di due sole lettere alfabetiche, facemmo implicito riferimento a quella definizione classica variamente enunciata da più fonti (34)

Comunque sia, se pensiamo ai diversi approcci che si sono succeduti (e alle ribollenti correnti che li hanno attraversati) si ha come l'impressione che dovremo calarci in un pozzo che enumera più strati sovrapposti.

Definizione Classica

Intanto abbiamo che la più antica definizione di probabilità *tra più alternative possibili* è data dal rapporto che passa tra casi favorevoli che qui sarebbero quelli del solo messaggio trasmesso da mittente, e casi "possibili" che andremo a definire più tardi.

Attribuendo allora il simbolo Ω all'insieme di tutte le alternative possibili, a prescindere da quali esse siano a seconda delle circostanze che si possono creare, e dicendo #Ω la cardinalità del medesimo, indicando con $\boldsymbol{A}$ la classe dei soli eventi favorevoli come potremmo dire delle uscite vincenti del lotto e del giusto messaggio in crittografia,

annotando col simbolo $\#\boldsymbol{A}$ la cardinalità di tale classe residua, e dicendo $\mathbf{P}(\boldsymbol{A})$ la probabilità di comparsa di un evento favorevole appartenente alla stessa,

al dunque avremmo uno stato del seguente tipo,

per il quale, $\Omega = \{1, \ldots, m\}$, $\#\Omega = m$, $\#\boldsymbol{A} = n$ per $n \leq m$

per cui possiamo anche scrivere $\mathbf{P}(\boldsymbol{A}) = \frac{\#A}{\#\Omega} = \frac{n}{m}$

dove a tale proporzione sono associate alcune caratterizzazioni che possiamo così riassumere nelle righe a seguire:

(1) la probabilità che un evento appartenente ad $\boldsymbol{A}$ si verifichi, è data da un valore compreso tra 0 ed 1 dove ciò altri non sarebbe che un caso particolare di quello più generale della misura dell'incertezza che corre come vedemmo entro un intervallo per il quale $0 \leq H \leq 1$;

(2) la probabilità dell'evento favorevole certo è pari pertanto a 1 mentre la probabilità di quello impossibile è pari a 0;

(3) la probabilità che si verifichi uno di due eventi singolarmente possibili ma tra loro incompatibili, è uguale alla somma delle probabilità di comparsa di ciascuna delle due alternative che rispettivamente appartengono a sottoinsieme $\boldsymbol{A}$ e sottoinsieme $\boldsymbol{B}$, per cui abbiamo,

$$\mathbf{P}(\boldsymbol{A} \cup \boldsymbol{B}) = \frac{\#\boldsymbol{A} + \#\boldsymbol{B}}{\#\Omega} = \frac{\#\boldsymbol{A}}{\#\Omega} + \frac{\#\boldsymbol{B}}{\#\Omega}$$

Codesto approccio risulta invero assai praticato nei pronostici e nelle scommesse ma *se quando esiste una buona conoscenza degli eventi* determina calcoli efficaci ed efficienti, in altre circostanze manifesta delle criticità.

Tali criticità non emergono da procedure persino disarmanti nella loro ovvietà, ma dal sopraggiungere di errori per cui sono date per *vere* talune premesse che si dimostreranno non corrette alla prova dei fatti così da determinare l'insorgere di false aspettative.

Definizione Frequentista

In proposito esiste una storiella la quale sebbene sciupata dall'esser stata narrata troppe volte e in diverse versioni, ha nondimeno il pregio di riassumere in modo esemplare il motivo per cui certi eventi sembrano seguire traiettorie ingannevoli.

Si narra infatti del Cavaliere de Meré che era un accanito giocatore sempre sull'orlo del fallimento,
il quale aveva intuito che ottenere un doppio sei da ventiquattro doppi lanci non truccati, era poco meno probabile rispetto al caso di un unico sei da quattro lanci non truccati.
Facendo tesoro di tale minima scoperta sfidava altri contendenti meno avveduti, col poco felice risultato di perdere regolarmente !
E così scrisse a Pascal lamentando che la matematica fallisce nei casi più irriducibili e che deve arretrare davanti all'evidenza empirica sulla quale andava a sbattere la testa;
ora Pascal che era il grande matematico che sappiamo, prese molto sul serio i dolori del povero de Meré e non solo cercò di affrontare la questione facendo ampio uso del calcolo combinatorio,
ma girò il problema a Fermat il quale in una famosa epistola compilata all'alba del 29 luglio del 1654 cominciò a gettare le basi d'una diversa ipotesi[10].

[10] La storia ha troppe madri e troppi padri per essere citata, ma è un fatto che si racconta.

Si cominciò dunque a delineare il concetto di frequenza per cui, reiterando quante più volte possibile una qualche esperienza materiale ..., avendo cura di non modificare il quadro inizialmente adottato ..., si verifica come il rapporto tra eventi favorevoli e eventi possibili, si stabilizzi dopo un numero sufficietemente alto di tentativi.

Abbiamo infatti che la frequenza dell'evento favorevole **F(*A*)** progressivamente converge sino quasi a coincidere con **P(*A*)** da intendere come limite di probabilità dove l'evento atteso tendenzialmente si verifica avendo $\lim_{m \to \infty} \frac{nA}{m}$

dove *n**A*** sarebbe il numero di eventi favorevoli che effettivamente hanno storicamente luogo essendo registrati da un osservatore che ne prende nota,

e dove *m* per $m \geq n\boldsymbol{A}$ sarebbe invece il numero dei casi possibili per una quantità di esperienze che tenda a infinito.

In tale prospettiva con uno stacco rispetto a prima, la probabilità di un evento favorevole risulta quindi definita dal rapporto fra il numero *n**A*** di esperienze dall'esito benigno registrate da un osservatore (come ad esempio sarebbe il nostro Oracolo) e il complessivo numero m di esperienze per $m \to \infty$ eseguite più volte a parità di condizioni.

Il problema che il cosiddetto "frequentismo" vuol dunque rimuovere, è in realtà già tutto concentrato nelle lamentele avanzate dal *Cavalier de Meré* in quei giorni di cocenti disillusioni quando si doleva con Pascal delle sue disavventure;

egli tuttavia sbagliava a prendersela con la matematica, giacché l'inganno sta piuttosto nella difficoltà di capire o meno se determinati eventi dati per risaputi, davvero siano sufficientemente comprovati.

In proposito l'ipotesi presa per buona secondo cui i dadi delle lunghe notti trascorse a vendere l'anima al diavolo non fossero truccati,

intanto non dava garanzie di nulla (e se fossero stati effettivamente truccati?) ma, soprattutto, non concedeva alcuna alternativa a quella di reiterare all'infinito uguali esperienze finquando non si fosse o meno evidenziato un qualche errore sistemico che aveva alterato il risultato ottenuto.

Invero noi già ci siamo imbattuti in qualcosa di simile quando dicemmo che i generatori true random si presentano come sorgenti deboli di randomicità, almeno prima di sottoporre *i flussi da sé rilasciati* a operazioni di whitening.

Se tale caso non può essere assimilato in tutto e per tutto a quanto detto sinora, pure richiede l'esecuzione di un numero abbastanza grande di test che possano registrare un bias (scostamento dal risultato ideale) il quale sui grandi numeri sarà o meno confermato entro un intervallo di confidenza;

in pratica, le tecniche poste dalla logica "frequentista" consentono di giungere in modo inferenziale a informazioni che una creatura dotata di risorse inimmaginabili (come il mago delle favole ma anche il diavoletto di *Maxwell* e il demone onnisciente e onnipotente di *Laplace* e chi ne ha più ne metta) avrebbe modo d'apprendere al primo sguardo.

La ripetitività consente di avvicinarci a tendere a un valore plausibile che sarebbe il meglio a cui noi umani aspiriamo e *in ciò* rappresenta una lezione di realismo della quale Fermat cominciò a delineare i tratti salienti.

Il Prezzo della Probabilità

(39)

Morta una criticità se ne fa un'altra essendo che qualcuno dovette cominciare tuttavia a pensare che non sempre si possono fare e disfare gli stessi esperimenti, giacché esistono situazioni irripetibili che sfuggono alla definizione frequentista.

Tra esse ci piace ricordare quella del match per il campionato di scacchi che avrebbe dovuto tenersi nel 1975 a Manila tra il campione americano *Bobby Fischer* e quello sovietico *Anatolij Karpov* e che scatenò la fantasia degli allibratori.

La particolarità di tale epico incontro è che non ebbe mai luogo, giacché Fisher che era il campione in carica, aveva cercato di imporre condizioni talmente arbitrarie al suo avversario, che la Federazione mondiale dopo tortuose trattative e un ultimo disperato tentativo di mediazione, lo dichiarò decaduto dal titolo che assegnò allo sfidante sovietico.

Se in codeste condizioni qualsivoglia pronostico perde di significato, ciò nondimeno sono tuttora molteplici i tentavi di fissare un favorito tanto che son state giocate partite da programmi informatici capaci di simulare lo stile dei contendenti; tra i quali, è spesso indicato come vincente Robert James Fischer cui è assegnato un lieve vantaggio[11].

Al contempo solo l'anno prima della più leggendaria partita di scacchi mai giocata, lo statistico e matematico *De Finetti* il cui nome all'anagrafe risponde niente di meno che a quello di Bruno Johannes Leonhard Maria von Finetti, pubblicava come summa della sua opera ultradecennale, una Teoria della Probabilità o *Theory of Probability* dove tornava a battere sull'ipotesi soggettivista elaborata con *Savage* a partire dagli anni trenta del secolo scorso, e dove era avanzata la questione della irripetibilità dei fenomeni su cui sono esercitate previsioni statistiche;

fu così offerta *con un coup de théatre* una nuova definizione di probabilità per la quale un evento si ritiene più o meno probabile a seconda del prezzo che un operatore umano dotato di coerenza (per cui lo stesso evento non può esser dato per vero e falso al contempo) ritiene equo pagare, ricevendo 1 se l'evento previsto si verifica, e 0 se non si verifica.

[11] Brano tratto dal rapporto redatto nel 2020 da Claudio Cappelli e intitolato *A new implementation of an extension of One Time Pad cryptography model*.

Probabilità Soggettiva

Intanto occorre precisare che anche dalla definizione soggettiva si ricavano *sebbene alla fine di un diverso percorso* caratterizzazioni affini a quelle di cui abbiamo detto sin dalla formulazione della teoria classica, e per le quali abbiamo:

(1) $\mathbf{P}(\boldsymbol{A}) \in [0,1]$ in quanto se per assurdo $\mathbf{P}(\boldsymbol{A}) < 0$, si avrebbe un guadagno certo a prescindere dal successo del pronostico, essendo che si sarebbe disposti a pagare meno di nulla,
mentre se $\mathbf{P}(\boldsymbol{A}) > 1$, si avrebbe una perdita certa, sempre a prescindere dal risultato conseguito, essendo che si sarebbe intenzionati a pagare più di quanto si possa ricevere;

(2) $\mathbf{P}(\Omega) \stackrel{\text{def}}{=} 1$ dove se l'evento favorevole è certo giacché gli appartengono tutti i risultati possibili, sarà pur certo che paghi 1;

(3) sapendo che $\boldsymbol{A}$ e $\boldsymbol{B}$ sono distinti sottoinsiemi di Ω, qualora fosse $\boldsymbol{A} \cap \boldsymbol{B} = \varnothing$ si dovrà anche avere $\mathbf{P}(\boldsymbol{A} \cup \boldsymbol{B}) = \mathbf{P}(\boldsymbol{A}) + \mathbf{P}(\boldsymbol{B})$,
dove se gli *n* elementi di sottoinsieme $\boldsymbol{A}$ e quindi $A_1, \ldots, A_n$ sono parimente necessari e incompatibili, possiamo pur scrivere $\sum_{i=1}^{n} \mathrm{P}(A_i) = 1$
giacché si incassa 1 a prescindere da quale sia l'evento favorevole A_i che dovesse effettivamente comparire, dove tale elemento A_i è necessario tra quelli possibili in quanto appartenenti a Ω.
Considerando allora gli eventi $A_i \in \boldsymbol{A}$ e $B_i \in \boldsymbol{B}$ e l'evento complemento della loro unione, che a loro volta sarebbero tutti necessari e incompatibili per cui *uno e non più di uno* dovremmo in ogni caso ottenere,
possiamo pur scrivere,
$\mathbf{P}(\boldsymbol{A}) + \mathbf{P}(\boldsymbol{B}) + \mathbf{P}(\overline{\boldsymbol{A} \cup \boldsymbol{B}}) = 1$
per la qual cosa, essendo l'unione dei sottoinsiemi $\boldsymbol{A}$ e $\boldsymbol{B}$ e il loro complemento incompatibili, possiamo altresì diversamente dire,
$\mathbf{P}(\boldsymbol{A} \cup \boldsymbol{B}) + \mathbf{P}(\overline{\boldsymbol{A} \cup \boldsymbol{B}}) = 1$
dove tale espressione assume significato anche per eventi elementari non equiprobabili e non ripetibili giacché quello che conta è che la remunerazione in caso favorevole sia sempre uguale a 1 a fronte di un rischio sempre inferiore a 1;
un rischio da assumere a giudizio dell'operatore umano il quale nel soggettivismo ammortizza *nel momento in cui stabilisce egli stesso il prezzo* eventuali difformità che avverte nella distribuzione di probabilità degli eventi.

Il Giardino delle Probabilità

(40)

Se ci siamo tuttavia avventurati nel labirintico giardino delle probabilità, dove abbiamo lavorato di forbice sulle diverse definizioni, non è per rinfrescare la memoria ma per promuovere alcune opinioni che ci saranno utili.

Intanto è bene sottolineare che sebbene siano stati trattati più approcci collocati in letteratura lungo un ampio arco di ipotesi, sul campo essi sono sovente impiegati quali metodi di calcolo piuttosto che come costrutti *l'un contro l'altro armati*;

si pensi a quel che accade quando occorra fare un computo che muova da dati consolidati e dove si devono semplicemente discernere i casi favorevoli da quelli che non lo sono (quattro assi vincenti, per dodici figure e ventiquattro semi perdenti),

o quando ci si trovi dinanzi a un evento irripetibile come l'ammaraggio oceanico d'una navicella di cui serva calcolare il rientro;

non mancano poi fenomeni complessi da smontare in più parti alle quali applicare diverse tecniche pragmaticamente selezionate.

Non tocca dunque a noi prendere partito per scegliere tra differenti ipotesi;

così come ci siamo regolati col concetto di casualità o con la questione dei numeri pseudo-casuali generati da sorgenti matematiche che non sono impiegati nei sistemi perfetti,

parimente terremo una posizione agnostica rispetto al ponderoso concetto di probabilità che, tra l'altro, è sufficientemente battuto da molteplici angolazioni.

Prendendo infatti atto che un'idea condivisa non esiste pur essendoci un'unica teoria assiomatica che fissa il framework di riferimento delle diverse definizioni[12] intendiamo andare alla ricerca di taluni tratti salienti che ci sembrano utili ai nostri fini e trasversali a più punti di vista.

(41)

Per questo vorremmo mettere intanto in guardia prima di tutto noi stessi, da un facile equivoco nel quale è possibile incorrere.

Non si può infatti dire che *non* vi siano elementi di oggettività nel cosiddetto soggettivismo, così come non si deve pensare che la prospettiva dell'operatore umano, sia estranea alla cosiddetta teoria classica almeno per come sarebbe declinata oggigiorno.

Per quanto nel soggettivismo conti il giudizio di chi si avventura nei pronostici, è nondimeno richiesto un criterio di coerenza ed una conoscenza del problema da stimare in base alle informazioni raccolte[13];

[12] In letteratura si aderisce a un'unica teoria omnicomprensiva che coincide con quella assiomatica di Kolmogorov (1933) la quale non appare come un metodo tra altri metodi, fornendo piuttosto dei postulati indifferentemente riferibili sia a definizioni soggettive che oggettive.

[13] Si veda in proposito *Incertezza e Probabilità* di Romano Scozzafava, Zanichelli Editore, Marzo 2001.

giacché ogni impulso decisionale scatta in un secondo momento quando chi opera deve trarre le sue conclusioni non incoerenti.

In egual modo non manca un certo tipo di "soggettività" che entra in gioco nella teoria classica e non solo in quella[14]; più in generale possiamo infatti dire che la prospettiva assunta da un ipotetico Oracolo e le informazioni cadute in suo possesso, incidono sul risultato da raggiungere e ciò vale per qualsivoglia metodo adottato.

Quando tra la notte della vigilia del 2004 e le prime ore del Natale cristiano coi bambini che scartocciano regali, i tecnici di ESA e NASA scoprirono un oggetto meteorico di circa trecento metri di diametro che battezzarono col nome di Apophis 99942,

si resero subito conto che aveva una non trascurabile probabilità del due per cento di collidere con la Terra; sedici anni dopo le stesse fonti ci confermano che l'impatto è ormai scongiurato per almeno i prossimi cent'anni.

Ora è evidente che quella ch'è cambiata non è la marcia di avvicinamento nello spazio di Apophis 99942 che era segnata; è la curva delle distribuzioni di probabilità ad esser stata inizialmente fissata in base a dati parziali, da tecnici addetti al calcolo che hanno selezionato livelli di probabilità modulati sulla scia di quanto ne sapevano in quel momento.

Nella tormentata vigilia del 2004 e cioè in un tempo ***t*** *nello spazio di un luogo chiamato "pianeta terra" apparve come in sogno un insieme* ₱ *all'interno del quale fu stimata una curva di probabilità delle diverse alternative;*

ma quando si sono potuti ripetere simili calcoli in un tempo ***t'*** *dove si collocavano le nuove posizioni della Terra e della meteora in avvicinamento, la classe residua di eventi infausti si era ridotta al pizzico d'un insieme vuoto in ragione delle misure cumulate in anni di osservazioni.*

Di certo, non s'era modificata la traiettoria di Apophis nello spazio delle fasi, attraverso la parabola che lo divideva dal nostro tran tran quotidiano;

quella che cambiava e cambia era il prospetto assunto dall'osservatore umano [15] *lungo un cammino in cui aveva progressivamente aumentato il suo carico di informazioni*[16].

[14] Naturalmente trattiamo di "soggettività" differenti, nel senso che la parzialità della posizione effettivamente occupata da un osservatore è diversa dalla soggettività del suo giudizio.

Per essere tuttavia più chiari, possiamo osservare che pure quando si calcolano le probabilità dei numeri al lotto e delle relative vincite (ambo, terno, etc.) si assume il punto di vista del giocatore.

Se si assumesse quello d'una sorta di demone che sappia calcolare le forze agenti dall'interno dell'urna predicendo le uscite, diverso sarebbe pure il computo delle probabilità a prescindere dal metodo adottato che darebbe a tendere .risultati convergenti

[15] Se pure parliamo di macchine, parliamo di macchine create e applicate dall'uomo.

[16] Il brano è nuovamente tratto dal rapporto *A new implementation of an extension of One Time Pad cryptography model.*

Capitolo VI

Tema

(42)

Sono dati due apologhi per i quali focalizziamo talune problematiche che discendono dalla definizione di perfetta sicurezza così come la troviamo in letteratura.

Ci si chiede cosa effettivamente distingua i sistemi più convenzionali da quelli a sicurezza perfetta e cioè da quelli che rispondano ai requisiti richiesti dai lemmi (**B**)(**C**) che forniscono una più ampia generalizzazione di quanto enunciato nel teorema in (**A**).

Sono messi in luce alcuni aspetti che contrassegnano il concetto di segretezza a seconda che sia o meno concessa a chi attacca la facoltà di individuare classi residue che gli consentano di restringere il campo giungendo alla soluzione.

A conclusione del capitolo, formuleremo una definizione alternativa di perfetta segretezza.

Altre Favole

(43)

Diverse pagine fa nel mezzo della tempesta del secondo capitolo del trascorso documento, lasciammo un mago alle prese con innumerevoli formule declamate allo scopo di aprire un forziere dove stavano racchiuse le ricchezze d'un immenso tesoro.
Egli aveva tuttavia fallito, non perché avesse sbagliato formula ma essendo che i battenti della cassaforte lasciata per anni al freddo e al gelo si erano arrugginiti, motivo per cui non avevano risposto ai comandi.
Le cose si erano dunque messe in modo tale che due sono le conclusioni che possiamo dare banalmente per certe:

(1) in primo luogo il mago-oracolo aveva fallito;

(2) in secondo luogo, avendo pronunciato ogni possibile abracadabra si dibatteva nel dubbio domandandosi quale fosse quello effettivamente magico, non avendo trovato soddisfazione nel chiavistello rimasto bloccato.

Il nostro mago era un mago incantatore e perciò si credeva così potente da non potersi capacitare dello smacco subito, fin quando una notte in cui s'era addormentato col cuore in subbugli ... in uno sbuffo non gli apparvero tre diavoletti.
Il primo diavolo gli disse che possedeva cento forzieri e se avesse trovata la formula per poterli aprire, tutte le ricchezze sarebbero state sue;
il secondo gli disse che possedeva mille forzieri, ma se avesse trovata la formula per poterli aprire, metà delle ricchezze sarebbero state sue;
il terzo gli disse che possedeva diecimila forzieri, ma se avesse trovata la formula per poterli aprire, avrebbe avuta salva la vita …, e sennò la gola gli avrebbe tagliata.
Il primo diavoletto dal mantello bianco gli diede allora una lettera illeggibile nella quale era celato uno scritto che sebbene incomprensibile, nondimeno gli consentì di calcolare riga per riga la misura della frase da dover pronunciare.
Coi suoi poteri inumani declamò allora tutte le formule dell'esatta lunghezza riuscendo a aprire alla velocità del pensiero novantanove forzieri, giacché tale era la proporzione tra chiavistelli funzionanti e chiavistelli divenuti nel tempo inservibili.
Egli fu dunque in grado di riconoscere la giusta magia e venne così coperto di ori ed argenti che lo resero ancora più ricco di quanto era mai stato.
Anche il secondo diavoletto dal mantello rosso gli diede una lettera illeggibile ma straordinariamente lunga, per cui lunghezza e numero delle formule erano cresciuti a dismisura,
ma ciò non lo poteva impressionare per cui trascorsi novecento novantanove battiti di ciglia aprì di botto novecento novanta forzieri essendo tale la differenza che correva tra chiavistelli in uso e chiavistelli ormai irreparabili.

Poi però il terzo diavoletto dal mantello nero non aveva nulla e nulla gli diede, e si mise a ridere apostrofandolo con fare beffardo "vediamo un poco, mago dei miei stivali, se stavolta indovini o se renderai l'anima al Diavolo ..."
E perciò il mago andò a tentoni non sapendo che pesci pigliare per cui cominciò a dare numeri alla cieca ma non ebbe fortuna giacché la formula era di grandezza infinita e così perse la pazienza avendo ultimati cento miliardi di miliardi di tentativi, tanto che con un inchino fece il gesto di filarsela all'inglese lasciando tutti con un palmo di naso.
Non l'avesse mai fatto.
In un attimo il diavolo dal manto nero gli fu addosso e estratta una scimitarra gli tagliò di netto la gola; *e fu così che quel mago incantatore perse insieme alla zucca e alle sconfinate ricchezze ..., quel che gli restava per cent'anni da vivere.*

(44)

Essendo tuttavia che come con le ciliegie una storia tira l'altra ..., volendo indagare su alcune sottigliezze che metteremo a nudo nella lettera del prossimo paragrafo ..., non ci faremo pregare e diremo d'una seconda fiaba con cui intendiamo illuminare un caso limite dove le caratterizzazioni poste in (**A**) cadono nel tranello d'una particolare eccezione.

Nel contempo sotto la stessa luna ma a mille leghe dalla dimora dove il mago dell'abracadabra era andato incontro al suo triste destino ..., orsù sorgeva il Regno del Sultano Verde rinomato per i fitti palmizi e le fanciulle in fiore, ma pure giacché i suoi abitanti parlavano una lingua strana e incomprensibile.
La particolarità stava nel fatto che impiegava parole lunghissime;
articoli determinativi e indeterminativi, pronomi e preposizioni di ogni tipo, oltre a nomi comuni singolari e plurali, avverbi e aggettivi qualificativi, non avevano mai meno di cento lettere alfabetiche sebbene con una singolare eccezione.
L'unico lemma che per legge era più breve, significava "attacco" nel senso che ordinava l'assalto alle guarnigioni nemiche, e si componeva di sole tre tra vocali e consonanti.
Infatti il Sultano Verde era perennemente in guerra con quello Rosso e quando si trattava di muovere gli eserciti non c'era tempo da perdere!
Essendo tuttavia che pure tra una tenzone e l'altra molte missive finivano in mani sbagliate, il Sultano si era affidato a un sapiente che gli aveva suggerito di camuffare i dispacci;
egli avrebbe sostituita ogni lettera con un'altra, impiegando un codice di lunghezza uguale a quella del messaggio e avendo cura di far estrarre i simboli del canone segreto, dall'anima innocente di un bambino che non aveva compiuti i sette anni.

Fatto pertanto pervenire il codice così ottenuto ai riceventi, questi e solo questi avevano la possibilità di ricomporre i testi originali una volta giunti a destinazione[17].

Nelle notti languide, nei tramonti rosseggianti quando il fogliame pareva prendere fuoco ..., con tale inganno molti messaggi furono perciò inviati ai sudditi e alle amanti del Sultano, e sebbene il nemico avesse valenti matematici nessuno ne veniva a capo.

E' inutile dire che ormai gli uomini di ingegno scarseggiavano nel Reame del sultano Rosso essendo stati tutti giustiziati e buttati nella fossa dei serpenti.

Solo uno era sopravvissuto essendo un giovane di bottega, ma quando anche l'ultimo dispaccio fu carpito agli araldi sulla strada che portava dalle acque del fiume alle rocce della montagna, giunse anche per questi il suo momento.

Alì ..., tale era il nome del piccolo matematico, tremava come una foglia quando gli fu consegnato il testo che avrebbe dovuto decifrare e per il quale tutti avevano fallito, ma riprese colore appena potette dare un'occhiata a quel dispaccio segreto che gli apparve trasparente come acqua sorgiva.

Tutti i precedenti messaggi erano lunghissimi, pagine e pagine di frasi illeggibili vergate in una grafia minutissima, ma questi che era l'ultimo in ordine di tempo, appariva conciso e chiarissimo;

tre simboli tre campeggiavano infatti su una pergamena quasi vuota e d'un biancore inusitato.

Non c'erano dubbi, si trattava niente di meno che dell'ordine di attacco, l'unico che si poteva conciliare con la brevità della parola; e così Alì si presentò a palazzo colmo di fervore, spiegando che occorreva predisporre ogni difesa giacché gli eserciti nemici erano prossimi all'assalto e l'ordine già era stato impartito.

E così fu fatto.

Vennero collocate possenti difese mentre micidiali trabocchetti furono sepolti sotto due palmi di polvere ...;

frombolieri e arcieri rimasero in attesa alle bocche del deserto e duemila cavalieri si nascosero nella sabbia dorata.

Non si fecero prigionieri e la battaglia fu vinta al sorgere del sole che si levò in cielo per la gloria del Sultano Rosso ..., il più clemente e misericordioso dei sultani che aveva trafitto al cuore mille nemici mozzando la testa col saif a altri mille.

[17] In breve, ognuno avrà inteso che si usava un cifrario a sicurezza perfetta.

Non solo Favole

(45)

Con l'intento di rendere le cose più digeribili, vediamo di fornire allora una sintesi che ci consenta di mappare lo stato dei fatti, snocciolando alcuni rilievi da muovere al concetto di sicurezza per come fissato in letteratura;
si danno così le premesse per giungere a una nuova definizione di sicurezza perfetta oltre che a un teorema da illustrare in conclusione del corrente lavoro.

a) Diciamo intanto che *se per il tradizionale teorema sulla perfetta sicurezza e le caratterizzazioni che ne conseguono, Alì non avrebbe dovuto giungere all'esatto messaggio*, egli invece lo fece basandosi sulla conoscenza del crittogramma e su quella lingua parlata dal nemico ammessa giustappunto da Shannon.

Gli altri matematici che lo avevano preceduto s'erano arresi davanti a spazi popolati da moltitudini di messaggi tra loro equi-probabili;
egli poté invece selezionare una classe residua abitata da un unico messaggio, ed è appunto il successo o meno di siffatta ricerca a qualificare un attacco tanto che ogni ulteriore congettura a noi appare ridondante.

b) Assumendo che il nostro Oracolo abbia risorse infinite, nella favola dell'abracadabra si rimarca la centralità della pratica del riscontro e cioè l'importanza che riveste il fatto di poter validare o meno un evento.

La questione è anche ripresa in un sequel dove si mette a nudo che solo davanti a classi di elementi di lunghezza non solo arbitraria ma persino infinita, la prassi del riscontro perde di significato compromettendo l'approccio di chi attacca ma, in pratica, ciò non accade per cui il problema da definire coincide con quello per cui tutto si riduce alla possibilità o meno di confermare un risultato (cosa molto difficile ma fattibile quando si attenti alla sicurezza di un cifrario a sicurezza computazionale ma impossibile coi sistemi detti perfetti che rilasciano più output equamente probabili).
Se *provando e riprovando* dovessimo ad esempio trovare il valore alfanumerico d'una password, abbiamo che sarebbe giustappunto il consenso all'accesso alle aree riservate a svelare il buon esito dell'attacco;
ma qualora per assurdo, esso sistema fosse bloccato da qualche anomalia, *a prescindere dalle conseguenze pratiche* saremmo nella condizione di chi giocando al lotto, abbia perso il biglietto vincente nel senso che ogni successo rimarrebbe nella mente di Dio non concedendo a noi umani di distinguere il vero dal falso.

In effetti la necessità del "riscontro" non appare nella definizione standard di perfetta sicurezza[18] che si limita a recitare con una tautologia che la conoscenza di *Critto* nulla aggiunge alla probabilità di risalire al messaggio.

E' tuttavia possibile che ciò sia troppo e troppo poco al contempo.

- Troppo poco giacché sappiamo che la conoscenza di *Critto* aggiunge qualcosa allo stato dei fatti, consentendo di passare dal *mare magnum* di un insieme universo al più circoscritto spazio da cui pescare il messaggio.
- Troppo troppo in quanto si richiedono *tra gli altri* taluni requisiti sostanzialmente arbitrari che restringono la sfera della segretezza pefetta anche dove non dovrebbero.

Dalla lettera della definizione standard e dal teorema di Shannon sarebbe tenuto fuori uno dei casi posti nel paradosso del crittografo e più in generale tutti quelli in cui la conoscenza di *Critto* ha sì un impatto sul rapporto tra alternative possibili e alternative favorevoli,
ma un impatto che potrebbe non essere sempre superiore a quello calcolato nei casi formalmente detti a sicurezza perfetta.
Volendo ad esempio ipotizzare uno stato dove la conoscenza del crittogramma consenta di escludere da una classe $\mathbb{C}$ di messaggi cifrati di numerosità pari a 2^N
un modesto margine che corrisponde al complemento $\overline{\Omega}$ (di sottoinsieme Ω) di numerosità pari a 2^n
per *n* molto piccolo ed N molto grande,
cosi da ottenere $\Omega = \mathbb{C} - \overline{\Omega}$
dove $\mathbb{C}$ avrebbe una numerosità di 2^N
dove $\overline{\Omega}$ una molto minore numerosità di 2^n
ed Ω una numerosità pari a 2^M per $2^M = 2^{N-n}$
avremmo che seppure il numero di messaggi cifrati ottenuto dalla differenza e che diamo per equi-probabili tra loro (dove tale postulato è cruciale) fosse sufficientemente grande da impedire il computo della maggior probabilità di comparsa di uno rispetto agli altri[19],
secondo la vigente definizione che boccia ogni condizionalità, saremmo fuori dalla cerchia dei sistemi perfetti.
Se tale conclusione appare eccessiva escludendo alcuni casi di particolare interesse, essa può anche apparire arbitraria per le ragioni a seguire.

c) Dato infatti un esempio (1) dove l'esatto messaggio nuovamente appartiene a Ω, intanto si avrebbe che la fortuita eventualità per un Oracolo di indovinare *tra tanti dispacci per lui equi-probabili* quello corretto, sarebbe pari a $1/2^M$

[18] Nemmeno implicitamente.

[19] **Parliamo dei messaggi appartenenti ad Ω per i quali ribadiremo che a differenza di quanto si creda, il caso ipotizzato sarà poco probabile ma non impossibile come dimostra una delle ipotesi poste in** (04).

Avendo tuttavia in (2) un diverso set $\mathbb{C}'$ di messaggi appartenenti all'insieme di un sistema incondizionatamente sicuro ma di minor cardinalità 2^m

per assurdo avremmo una sicurezza perfetta che esprime un più ristretto spazio degli eventi da cui discende una maggior probabilità di giungere fortuitamente al messaggio[20]**.**

La numerosità data dalla lunghezza del dispaccio in (2) implica infatti una distribuzione uguale a quella d'una probabilità pari ad $1/\ 2^m$ di arrivare alla giusta soluzione,

per $\frac{1}{2^m} \gg \frac{1}{\mathbf{2^M}}$

A meno di non voler perciò porre un problema meramente nominalistico, sapendo che il nocciolo sta nel comprendere quanto sia improbabile la risalita a partire dalla conoscenza di *Critto,*

se abbiamo che una classe ridimensionata rispetto allo stato originario e che chiamiamo pertanto $\mathbb{B}$,

possa essere più ampia di quella $\mathbb{A}$ fissata quando la probabilità sia riconosciuta come incondizionata",

intuitivamente a restare perfetta è l'impossibilità di forzare matematicamente il sistema pur conoscendo il valore del crittogramma.

Intendiamo dire che se è vero che un sistema a sicurezza "incondizionata" certamente impedisce di risalire dal crittogramma al messaggio (2)

è altrettanto vero che sul piano squisitamente logico non si può escludere per definizione che un sistema la cui sicurezza sia "condizionata" (1) non possa lo stesso impedire nel suo spazio vitale ogni tentativo di risalita[21].

d) Come accennato occorre altresì rimarcare che in rapporto a quanto detto sinora, il cosiddetto paradosso del crittografo dove sarà il messaggio a essere random e la chiave a non esserlo,

non trova spiegazione né alla luce della più volte citata definizione standard, né in forza del teorema di Shannon.

[20] Qui tuttavia facciamo sempre riferimento a un colpo di fortuna essendo che, in ogni caso, parliamo di eventi tra loro equi-probabili tra i quali non è possibile discernere in modo oggettivo.

[21] Ogni tentativo che cerchi di scalfire quel "core" indomabile di cui abbiamo detto.

Sicurezza 2020

(46)

Il buon proposito di fare i bravi col giungere del nuovo anno che speriamo migliore e non peggiore di quello passato,

si scontra col desiderio di proporre dopo ulteriori progressi, una nuova e diversa definizione di perfetta sicurezza che ponga rimedio all'inconsistenza di quella sinora trattata[22].

Per tal motivo abbiamo già indicati in Bob e Alice e nel nostro Oracolo i soggetti intorno a cui costruire la forma che possa collegare mittente $\mathbf{U}_T$ a ricevente $\mathbf{U}_R$.

Adesso seguiremo perciò i passaggi prefissi mettendoci nei panni di chi attacca, avendo cura di cominciare dal caso più semplice e cioè da quello dove nuovamente si impieghi un consueto messaggio di senso compiuto.

Muovendo allora dalla traiettoria di un oggetto che dalla coda d'una stella cadente[23] attenti alla sicurezza del terzo pianeta del sistema solare, è bene rammentare quanto dicemmo sul carattere delle informazioni di cui possiamo disporre;

più in generale si può infatti dire che la prospettiva assunta da un ipotetico Oracolo e le informazioni cadute in suo possesso, incidono sul risultato che si vuole raggiungere, e ciò vale per qualsivoglia metodo adottato (41).

In altre parole il punto di vista assunto da colui che attacca rappresenta la prospettiva non immutabile ma in evoluzione, dalla quale ogni curva delle probabilità prende forma qualora si intenda risalire illegalmente al messaggio.

Di pari rilievo ai nostri fini, è tuttavia un ulteriore fatto che rende manifesto un aspetto comune a qualsivoglia teoria fiorita nel giardino delle probabilità;

gli eventi possibili sono possibili e variamente probabili (quando non equi-probabili) giacché si dividono uno spazio matematico che è lo spazio cui appartengono le diverse alternative $A_1 \dots A_m$

Il concetto di probabilità attiene dunque alla partizione in p parti[24] per p alternative possibili di tale spazio insiemistico;

si capisce allora come la ricerca del set di riferimento sia delicata e non sempre ovvia a seconda delle condizioni che si vanno a creare.

Riveste perciò particolare importanza il fatto di seguire la curva di distribuzione al crescere delle informazioni e delle quantità di calcolo prodotte dall'operatore umano o, nello specifico dell'assalto a un sistema crittografico, delle informazioni di cui gode il nostro Oracolo prima e dopo aver intercettato il crittogramma.

[22] Più precisamente la definizione classica è coerente, ma entro un determinato ambito di applicazione; fuori da esso svela delle inconsistenze.

[23] *Apophis.*

[24] Non necessariamente parti uguali.

In origine per un attaccante, quando tutto è ancora avvolto nel mistero ..., più che d'uno spazio del messaggio sarebbe opportuno parlare di un insieme universo che diremo E_1 di ampiezza arbitraria se non infinita in ragione della lunghezza altrettanto arbitraria se non infinita del messaggio che potrebbe redigere a suo piacimento Alice.

E' ovvio che in tale particolare insieme da non confondere col più ristretto spazio degli eventi che poco a poco matura agli occhi di chi guarda ..., nessuna stima è fattibile mancando nella partizione il valore di un nominatore su cui esercitare il calcolo.

Sin da subito le cose tuttavia cambiano passando dalla sfera dei sistemi perfetti a quella dei sistemi convenzionali.

Abbiamo infatti visto che nelle simulazioni offerte in letteratura, è comune concedere all'Oracolo il privilegio di conoscere il cifrario ingaggiato; sebbene occorra ricordare che non parliamo d'un privilegio vero e proprio essendo che esso discende dall'esperienza per cui rare volte chi attacca non ha contezza del critto-sistema impiegato.

(i)

Ora abbiamo che nei sistemi convenzionali l'informazione sulla lunghezza della chiave è parte integrante della più ampia conoscenza del critto-sistema;

ciò permette di determinare in partenza (a differenza di quanto accade nei sistemi dove la mutevole misura della maschera è ignota prima che sia stato intercettato il crittogramma) uno spazio Ω_1 la cui numerosità dipende dalla lunghezza standard della chiave [25], essendo che le trasformazioni concesse nel passaggio dalla forma in chiaro a quella coperta, non potranno mai superare di numero le configurazioni che abbiamo combinando i simboli del segnale della chiave già fissata in lunghezza**.**

Tutto questo tuttavia ci consente di smascherare l'ambiguità che sta dietro al concetto molto gettonato di "messaggi possibili"

In prima battuta si potrebbero infatti considerare "possibili" tutti i file-messaggio della lunghezza di N bit del dispaccio di Alice ma, in una più realistica accezione, si devono considerare "possibili" i soli messaggi concessi dalla minor grandezza della chiave che restringe la classe degli eventi che potranno effettivamente accadere;

dove nell'un caso, tale spazio sarà uguale a 2^N e nell'altro alla minor ampiezza di 2^m per $2^m \ll 2^N$.

Ciò che ci preme tuttavia sottolineare è anche il fatto che non si può impunemente fluttuare da uno scenario all'altro,

essendo che per Alice i messaggi corrispondono sempre e solo a quelli redatti da mittente, mentre per chi attacca saranno sempre e solo quelli della classe da calcolare momento dopo momento da un'angolazione fuori dal sistema.

[25] 128-256-512 bit sono le lunghezze più consuete.

A questo punto possiamo perciò immortalare il seguente stato assunto dalla prospettiva di chi osserva (il nostro Oracolo),

dove $\Omega_1 \subset \mathbb{E}_1$

dove $\mathbb{E}_1 - \Omega_1 = \overline{\Omega}_1$ cui appartengono i messaggi da escludere dal novero giacché impossibili,

per $\mathbb{E}_1 - \overline{\Omega}_1 = \Omega_1$

(ii)

Volendoci tuttavia invece fermare al caso dei sistemi perfetti, **a prescindere se trattati in base al teorema in (A) o ai lemmi (B)(C)**,

la legge più agevole da assumere e cioè la legge da assumere avendo un messaggio dotato di semantica e sintassi dove si conciliano nuovi e vecchi enunciati (nel senso che nulla cambia nel caso, adottando questa o quella ipotesi),

sappiamo che impone una chiave di lunghezza uguale se non maggiore a quella del dispaccio (*e alfine uscimmo a riveder le stelle …*) la cui misura sarebbe ignota a chi attacca almeno in una stato iniziale del sistema.

In effetti in tale strettoia, l'unico elemento conosciuto ma talmente incerto da apparire scevro da ogni conseguenza pratica, sta nel fatto che il messaggio appartiene ad un insieme $\mathbb{E}_2$ del quale nulla è però immaginabile; nessun indizio infatti si avrebbe sulla distribuzione di probabilità a priori dei messaggi appartenenti a un universo illimitato per definizione.

A posteriori occorre però dire che qualche affinamento è nondimeno fattibile sebbene sterile;

non si può infatti negare che ... nel tentativo di chi attacca di fissare un contorno interno a un insieme infinitamente ampio come sarebbe quello di $\mathbb{E}_2$ ... un qualche avanzamento può essere in qualche modo tentato.

In verità abbiamo già detto che dalla conoscenza di *Critto* si traggono informazioni utili sulla misura di chiave e messaggio ..., essendo che attraverso facili deduzioni si viene a conoscenza della lunghezza di questo o quell'altro operando di maggior lunghezza che qui nondimeno sarebbe quello della chiave.

In concreto si può tuttavia dire che difficilmente Alice impiegherà una maschera persino più lunga del messaggio stesso, essendo che ciò impaccia l'efficienza del sistema senza poterne aumentare l'efficacia che più che perfetta non potrebbe essere.

Tutte le volte in cui tale pronostico sul rapporto di misura tra operandi apparirà corretto, pertanto avremo una lunghezza del crittogramma pari a quella del messaggio che sarà a sua volta pari a quella della chiave;

se tuttavia uno degli operandi effettivamente fosse di maggior grandezza come nel caso d'un messaggio di senso compiuto e d'una chiave più profonda, comunque si andrebbe a fissare una classe di messaggi più ristretta di quella di $\mathbb{E}_2$

Diciamo pertanto $\Omega\Omega_2$ tale classe che *spesso ma non sempre* coincide con quella di quell'insieme $\mathbb{C}$ da noi pure citato[26];
essa può essere infatti descritta come una classe di messaggi la cui portata dipende dalla misura del crittogramma intercettato, che a sua volta coincide con quella dell'operando di maggior lunghezza che nel caso più tradizionale sarebbe quello della chiave.
Ciò detto, possiamo perciò fissare il seguente stato preso dalla prospettiva di chi attacca la perfezione di un sistema di sicurezza,

dove $\Omega\Omega_2 \subset \mathbb{E}_2$

dove $\mathbb{E}_2 - \Omega\Omega_2 = \overline{\Omega\Omega}_2$ cui appartengono infiniti messaggi impossibili,

per $\mathbb{E}_2 - \overline{\Omega\Omega}_2 = \Omega\Omega_2$

(iii)

Sebbene del tutto trascurata, nell'ambito dei sistemi perfetti addirittura esiste una ulteriore cernita comunque possibile sempre a partire dalla conoscenza del crittogramma.

Conosciuto infatti il valore di *Critto* e identificata una classe $\Omega\Omega_2$ di consueti messaggi *non random*[27] che non superino la lunghezza del medesimo,
una novella selezione si ha estrapolando i dispacci dotati di semantica e sintassi (dove un possibile testo scritto ha una sintassi, ma anche un brano musicale o le foto d'uno sponsale hanno una sintassi) dal calderone di quelli che ne sono sprovvisti (laddove questi sono molto più numerosi in quanto maggiormente probabili rispetto a quelli intellegibili come potemmo vedere quando dicemmo della mitica Biblioteca di Babele).
Si può infatti costruire un nuovo sotto-insieme dando fondo a ogni configurazione capace di rispettare le condizioni fissate; intendiamo dire che si potrà definire una classe di tutti i consueti messaggi di una qualche lunghezza *L'* che siano di senso compiuto.
Diciamo pertanto $\Omega\Omega'_2$ tale classe di minor grandezza che riguarda un numero contato di dispacci, dove è facile pronosticare che l'insieme di quelli esclusi assumerà una frequenza di probabilità cumulata prossima a 0, mentre la distribuzione dei messaggi appartenenti a $\Omega\Omega'_2$ comporterà che a ciascuna loro alternativa necessaria e incompatibile, non possa che corrispondere una frazione ennesima della probabilità tendente a 1,
per cui avremo che la sommatoria di ogni frequenza di probabilità in $\Omega\Omega'_2$ condurrà a un risultato secondo cui $\sum_{i=1}^{z} P(message_i) = 1$ dove $message_i$ sono gli elementi costituiti dai messaggi che gli appartengono per $i = 1, \ldots, z$

[26] Perché tale classe $\Omega\Omega$ coincida con $\mathbb{C}$, occorre che $\Omega\Omega$ e $\mathbb{C}$ misurino la stessa ampiezza, come quando sia il messaggio a fare da operando di maggior lunghezza; ciò sarà vero nei sistemi convenzionali che impiegano brevi chiavi standard, e nei sistemi perfetti quando sia il messaggio ad essere random e la chiave non random (**B**)(**C**) o anche quando chiave e messaggio siano uguali(**A**)(**B**)(**C**)

[27] Come nel glossario diciamo "consueti" i messaggi veri e propri e cioè quelli che contengono l'effettiva informazione che Alice vuol trasmettere a Bob; informazione presumibilmente dotata di semantica e sintassi.

Possiamo pertanto fissare il seguente stato preso dalla visuale di chi attacca,

per cui $\Omega\Omega_2 \subset \mathbb{E}_2$

dove $\Omega\Omega'_2 \subset \Omega\Omega_2$

dove $\Omega\Omega_2 - \Omega\Omega'_2 = \overline{\Omega\Omega}'_2$

per $\Omega\Omega_2 - \overline{\Omega\Omega}'_2 = \Omega\Omega'_2$

A differenza di quanto si dica, nel caso e soltanto nel caso in cui nei sistemi perfetti si conosca il valore del crittogramma, potremo dunque avere che sotto-insiemi di sotto-insiemi si possono tranquillamente scavare dalla cerchia degli insiemi di maggiore ampiezza;

solo che siffatto lavoro di cesello risulterà inefficace, non consentendo di forzare il sistema sino a giungere a una soluzione sufficientemente probabile se non univoca[28]**.**

[28] Per poterci meglio intendere, possiamo tornare al caso dove la lunghezza del messaggio, dedotta da quella del crittogramma, risulti essere di appena due lettere alfabetiche.

In tale frangente, quella che abbiamo indicata come classe Ω_2**'** è circoscritta e comunque molto meno ampia di quell'insieme universo che faceva da riferimento in una fase *a priori* e cioè ancor prima che il nostro Oracolo avesse intercettato il crittogramma.

Se chi attacca si dovesse ad esempio cimentare nella predizione della risposta da dare alla domanda *"hai mai letto la Divina Commedia?"* egli potrebbe supporre una replica di due sole lettere alfabetiche come suonano, in italiano, le locuzioni del *sì* e del *no*.

Qualora la risposta fosse però perfettamente segreta egli non saprebbe come andare oltre a quanto è riuscito intuitivamente a supporre dalla formulazione della domanda medesima.

Pur sapendo che la lunghezza della risposta non supera le due lettere alfabetiche, non sarebbe in grado di distinguere tra un'affermazione (*sì*) e una negazione (*no*);

anzi sul piano algebrico, non potrebbe nemmeno discernere tra tutte le altre voci di due lettere alfabetiche come pure sarebbero brevi lemmi quali *io, tu, do, re, mi, fa, la, si, lì, là, su, in, on, up, at,* e così andando in tutte le lingue scritte e parlate del mondo.

Almeno sino a un certo punto, non è la grandezza dell'insieme di riferimento a fare la differenza dove nei sistemi di crittazione convenzionali non sono esclusi spazi di maggior numerosità rispetto a quelli perfetti.

E non sarà nemmeno il fatto che tali insiemi siano sotto-insiemi di universi incommensurabili o di insiemi certamente più ampi a fare la differenza;

ciò che conta è che esista un nocciolo duro, un *core* non scalfibile, di soluzioni necessarie e incompatibili tendenzialmente equiprobabili tra loro.

Definizione 2020

(47)

Dopo ampi giri che a volte ritornano come nella canzone di Battisti, dobbiamo tuttavia dire che la definizione offerta da Shannon e affinata nei decenni a seguire, in buona sostanza punta sull'identità tra il concetto di perfetta sicurezza e quello di segretezza incondizionata.
In realtà il cammino che porta a tale conclusione, come accade sovente quando ci si avvicina all'opera del padre della teoria dell'informazione, muove dalla sua personale esperienza e cioè dal fatto che negli ambienti dello spionaggio internazionale s'era fatta strada l'idea che il cifrario inventato negli anni venti del novecento dall'ingegner Gilbert Vernam, fosse insensibile tanto alla crittoanalisi che agli attacchi a forza bruta[29].
Anche a voler tuttavia partire da tali premesse, occorre ribadire che non c'è niente che giustifichi l'identità promossa come un macigno da Shannon giacché nulla ci consente di affermare che un sistema condizionato possa essere sempre efficacemente attaccato da assalitori dalle risorse infinite a prescindere se possa o meno offrire più alternative tra loro indecidibili.
Il punto certo è che a rigore, prendendo le mosse dalla definizione che conosciamo, da un canto dovremmo concludere che sistemi imbattibili ci sono, giacché lo sappiamo per le prove accumulate sul campo e in base a facili calcoli combinatori,
e dall'altro dovremmo arrenderci al fatto per cui la sicurezza perfetta sarebbe un'utopia non essendoci sistemi di crittazione incondizionati se è vero che il crittogramma rilascia un nucleo di minime informazioni impossibili da abbattere.
La consueta formula per cui $P(\boldsymbol{m} \mid Critto) = P(\boldsymbol{m})$ sovente data per rimarcare il carattere non condizionato dei sistemi perfetti, ha senso a patto di considerare tollerabile una lieve ma stravolgente finzione**,**
per la quale ***m*** sarebbe di lunghezza conosciuta, cosa che non sarà mai vera dalla prospettiva di chi attacca quando gli manca proprio siffatta informazione.
E in effetti mentre $P(\boldsymbol{m})$ è un valore di arbitraria grandezza,
$P(\boldsymbol{m} \mid Critto)$ è un valore perfettamente fissato nella sua ampiezza numerica.
Se però il carattere incondizionato dei sistemi perfetti, non appare veritiero quando si guardi a priori allo spazio di un insieme universo,
vedremo alla fine che può rientrare dalla finestra se imputato a qualche classe residua realmente scovata dal nostro Oracolo, essendo che *nel caso* non staremmo parlando di un costrutto arbitrariamente fissato,
ma d'uno spazio rintracciato in ragione delle informazioni accumulate da chi attenti effettivamente al sistema.

[29] Se Shannon ha il merito di aver fissate le basi teoriche della crittografia contemporanea, inclusa quella a sicurezza perfetta, è Gilbert Sandford Vernam ad aver inventato quel sistema massimamente sicuro al quale fu poi attribuito il nome di one-time pad.

Pur tuttavia, quella che intendiamo fissare e che definisce il nocciolo della questione, non è la mitica incondizionalità dei sistemi a sicurezza perfetta che suona come una petizione di principio, ma la loro indifferenza agli assalti di chi conosca il valore di *Critto*.

Definizione

Per cui diremo di avere sicurezza perfetta quando un attaccante dotato di risorse infinite, pur conoscendo il valore del crittogramma, abbia una probabilità prossima a 0 di risalire da esso a messaggio ***m*** *dove* ***m*** *appartiene a una classe sufficientemente ampia di messaggi che una volta cifrati appaiono tra loro equi-probabili*[30].

[30] In effetti a monte i messaggi non sono affatto equi-probabili; lo diventano a valle dalla prospettiva del nostro Oracolo in quanto coperti dal crittogramma.

Capitolo VII

Tema

(48)

Saranno qui analizzati i momenti salienti dell'assalto a un sistema crittografico così come ha luogo dalla prospettiva di chi lo intenda forzare, mostrando come insiemi e classi residue si possano restringere fino a un certo punto nei sistemi perfetti e sino in fondo nell'attacco a un sistema convenzionale.

Dopo aver quindi marcate nel secondo capitolo più criticità di ordine logico, da qui innanzi si mostra il modo per poterle superare, anche illustrando i meccanismi algebrici per i quali un attaccante individua classi sempre più rarefatte che possano o meno portare a quella che Shannon usava indicare col nome di soluzione unica o *unique solution.*

Grandezze e Grandezze

(49)

A lungo inseguite le tracce che conducono ai sistemi perfetti, diciamo che è impossibile dire l'ultima parola senza nuovamente parlare di quelli che non lo sono;

in fin dei conti questi sono il rovescio di quelli e le ragioni che rendono invincibili gli uni, sono uguali e contrarie a quelle che fanno apparire meno resilienti gli altri[31].

Intanto in precedenza, già eravamo giunti al punto da dove vorremmo ripartire e cioè al passaggio che muove da una probabilità *a priori* ..., a quella che diviene una probabilità *a posteriori* dal momento in cui chi attacca intercetti il flusso del crittogramma.

Già infatti dicemmo che nei sistemi convenzionali è possibile determinare *a priori* l'ampiezza d'uno spazio Ω_1 del messaggio la cui numerosità dipende dalla misura standard della chiave effettivamente impiegata,

Ragion per cui abbiamo uno stato,

dove $\Omega_1 \subset \mathbb{E}_1$ essendo $\mathbb{E}_1$ un insieme universo di arbitraria numerosità,

dove $\mathbb{E}_1 - \Omega_1 = \overline{\Omega}_1$

per $\mathbb{E}_1 - \overline{\Omega}_1 = \Omega_1$

dove Ω_1 raffigura una classe di messaggi la cui numerosità è ridotta in ragione della misura della chiave medesima.

Quello che ci interessa tuttavia capire è perché quanto risulta fattibile per un attaccante alle prese coi sistemi convenzionali, diviene impossibile con quelli perfetti a prescindere dallo spazio che consentono che sarà maggiore o minore in ragione di più variabili;

si potrà infatti vedere che, fatta salva la questione dell'inversione di proprietà tra chiave e messaggio su cui torneremo appena possibile ..., esiste margine per una più compiuta formulazione di quei sistemi che *non* sono matematicamente scalabili.

Non sarà tuttavia di sicurezza computazionale e nemmeno di sicurezza perfetta *tout court* che intendiamo parlare,

giacché vorremmo intanto esplorare la terra di mezzo tra l'uno e l'altro sistema, e quindi il salto per il quale si transita da uno stato infinitamente imbattibile a uno permeabile agli attacchi (per quanto onerosi possano essere).

[31] Può a volte sembrare che ci siamo dimenticati delle difficoltà che anche i sistemi "a sicurezza computazionale" creano a chi attacca, **per cui sul punto ci riportiamo a quanto dicemmo in** (15).

(69)

Intanto sappiamo che una sostanziale differenza tra sistemi perfetti e non perfetti, sta nel rapporto di misura tra operandi dal carattere più o meno aleatorio *secondo i criteri posti nei lemmi* (**B**)(**C**) e tra chiave e messaggio in quelli forzosamente ridotti del teorema di Shannon.

Ciò detto, è tuttavia bene ricordare (attraverso ripassi utili anche a noi che non vorremmo perdere il filo del ragionamento) che sino alle ultime battute del capitolo, continueremo a riferirci a un confortevole messaggio di senso compiuto e ad una consueta chiave dal carattere aleatorio,

e cioè a uno stato nel quale i diversi criteri posti nel teorema (**A**) e nei lemmi (**B**)(**C**) sostanzialmente coincidono.

Volendo perciò fissare la nostra partenza su basi solide ..., diciamo che avendo un messaggio di N bit e una chiave random della stessa misura numerica, notoriamente saremmo nell'ambito dei sistemi perfetti a prescindere dalla definizione cui si voglia aderire;

assunta allora tale forma condivisa, essendo che desideriamo sapere come stanno effettivamente le cose cercheremo di prendere confidenza col problema cominciando col porci una domanda poco consueta ma niente affatto banale.

Se è infatti facile intendere che quando si adoperi una chiave standard di misura usualmente minore a quella del messaggio, stiamo trattando di cifrari a sicurezza computazionale, con un *ballon d'essai* possiamo lanciare la notizia che qualora un testo sia incidentalmente così stringato da essere di lunghezza inferiore a quella d'una breve chiave di crittazione ..., avremmo intanto assunta una posizione assimilabile a quella riscontrabile nei sistemi perfetti dove è la maschera (lunga o corta che sia in termini assoluti) a essere di maggior ampiezza..

Volendo tuttavia continuare a giocare coi numeri, un'altra possibile domanda potrebbe essere quella di chiederci di quanto occorra modificare un sistema per poterlo dichiarare fuori dalla segretezza perfetta o meglio ancora ..., di quanto lo dovremmo sensatamente degradare per dire d'essere usciti dalla sfera magica di cui stiamo dicendo;

noi siamo infatti interessati proprio a questo e cioè a intendere cosa accade nel guado che passa da una condizione all'altra condizione.

Volendo infatti supporre d'aver potuto calcolare la lunghezza del messaggio trasmesso da mittente, cominciamo col dire che una ipotetica chiave più piccola di un solo e unico bit, a occhio non parrebbe sufficiente a insidiare la stabilità di un sistema originariamente perfetto.

Se dovessimo avere a che fare con le solite cantiche impegnando non meno di un giga che corrisponde all'incirca a un miliardo di bit, a occhio potremmo dire che l'impiego d'una chiave dalla lunghezza infinitesimamente inferiore e perciò di 999 milioni 999mila 999 bit, provochi un degrado trascurabile.

Pure se dovessimo impiegare una minor chiave con un *byte* mancante e quindi di 999 milioni 999 mila 992 bit, forse sarebbe nondimeno ragionevole supporre che non vi siano particolari problemi e così per un impercettibile taglio di 64 bit che lo stesso consente l'uso d'una robusta chiave di 999 milioni 999 mila 936 bit al cospetto d'una messaggio di un miliardo.

Aguzzando l'ingegno, potremmo tuttavia scoprire che la lunghezza di un giga del messaggio e quella relativa della chiave che ne ridimensiona la classe di appartenenza, ci riservano qualche sorpresa. Andando infatti a fissare quanto discende dalla chiave per effetto dello spazio che rilascia ..., nel senso dell'effettivo spazio dei possibili messaggi consentiti dalla minor grandezza della maschera da sommare a flusso in codifica[32] ..., abbiamo che si determinano le seguenti condizioni per le quali:

- **(a)** laddove supponiamo che chiave e messaggio siano uguali, avremo uno spazio di numerosità pari a $2^{1000000000}$ che non è modificato dalla chiave di ugual misura;
- **(b)** qualora si assuma invece una chiave meno profonda di un solo bit, si otterrà uno spazio di ampiezza pari a $2^{999999999}$ possibili messaggi[33];
- **(c)** laddove prevediamo infine un taglio di sessantaquattro bit si andrebbe a fissare un minor spazio pari a quello di $2^{999999936}$ messaggi ancora effettivamente possibili.

Per cui, se in **(b)** lo spazio risulta della metà del precedente con uno scarto serio ma non formidabile, in **(c)** a conti fatti avremmo una ampiezza della classe residua di diciotto miliardi di miliardi di volte inferiore, con uno scarto che a quanti ne sentirono dire dal mandarino del collegio cinese di *Li-Sou-Stian*, ricorda il gioco delle torri di Hanoi il quale *per quanto si narra tra monaci buddisti* non sarebbe difficile da definire come micidiale[34].

Tabella - *Grandezze a Confronto*

Quantità bit chiave- messaggio	Spazio della Classe Residua	Rapporto Grandezze
(a) 1.000.000.000; 1.000.000.000 bit	$2^{1000000000}$ possibili messaggi	1/1
(b) 999.999.999; 1.000.000.000 bit	$2^{999999999}$ possibili messaggi	1/2
(c) 999.999.936; 1.000.000.000 bit	$2^{999999936}$ possibili messaggi	$1/2^{64}$

[32] Naturalmente potrebbe essere sommato a blocchi o in altro modo ma qui, da un canto, facciamo riferimento alla prassi usualmente impiegata nei sistemi perfetti e, dall'altro, non siamo interessati a dettagli anche importanti nella vita reale, quanto a trarre concetti che ci aiutino a fissare la mappa che divide i confini della sicurezza perfetta da quella computazionale.

[33] Qui il simbolo-bit è assunto come minima informazione al cui variare crescono gli elementi dell'insieme dei messaggi.

[34] Come risaputo, il gioco delle torri di Hanoi, noto anche come rompicapo della fine del mondo, fu effettivamente inventato da Edouard Lucas che tuttavia lo attribuisce per finta a tale mandarino del collegio cinese di *Li-Sou-Stian*.

Sistemi Imperfetti

(50)

Abbiamo cominciato dunque a vedere con degli esempi, quanto celermente si passi da uno stato che getta nello sconforto il nostro Oracolo alle prese con miliardi e miliardi di elementi irriducibili, a uno dove si cominci a ragionare.

In seguito al collasso dello spazio degli eventi, nei sistemi convenzionali si potrà infatti scartare dal cerchio magico della classe residua[35], la stragrande maggioranza dei dispacci in contrasto col crittogramma e in sovrannumero con quanto concesso dalla chiave tutte le volte in cui essa corrisponda all'operando di minor lunghezza.

Quella che abbiamo narrata resta tuttavia una storia a metà, giacché riguarda una sola faccia della medaglia;

seppure la classe modellata sulla misura della chiave fosse un infinitesimo di quella tratta in origine, dobbiamo dire che resta tuttavia un mistero tramite quali magheggi chi attacca riesca a discernere tra quelle che *non* saranno le infinite possibilità di un insieme universo ma che pure restano le innumerevoli alternative di un sottoinsieme assai numeroso.

Assumendo per dire una chiave di 128 bit ormai pressoché desueta (a meno di non voler essere facilmente bucati),

pur avendo potuto eliminare innumerabili quantità di messaggi divenuti impossibili per il meccanismo matematico appena illustrato, lo stesso avremmo uno spazio pari a due elevato alla potenza di centoventotto!

Una cifra a trentotto zeri dalle cui combinazioni emergono spropositate riserve di testi variamente possibili che tuttavia non son nulla rispetto alle infinità che si sono potute depennare dall'insieme originario una volta e per sempre.

Ora a nostro parere, la questione può essere inquadrata nel seguente modo; intanto abbiam detto $\mathbb{E}_1$ un insieme universo di arbitraria numerosità cui appartiene ogni messaggio di lunghezza ignota in quanto ignota a chi attacca.

Con più costrutto diremo $\Omega\Omega_1$ lo spazio di ampiezza pari a 2^N bit incluso in $\mathbb{E}_1$ ma compatibile con la misura del crittogramma intercettato – e di conseguenza con la misura dell'operando di maggior lunghezza che *nel caso d'una chiave di lunghezza standard* sarebbe quello del messaggio di senso compiuto[36] – al quale appartengono tutti i messaggi che abbiam fissati di lunghezza *L* dove *L* giustappunto sarebbe di N bit.

[35] Attraverso una cernita che sarebbe ad oggi tanto severa quanto laboriosa.

[36] **Insistiamo su questo concetto del "senso compiuto" per sottolineare che stiamo ancora rimandando l'appuntamento col tema dei messaggi random.**

Sapendo poi essere $N > \boldsymbol{m}$ diciamo invece Ω_1 la residua classe di ampiezza pari a $2^{\boldsymbol{m}}$ cui appartengono tanti messaggi di lunghezza *L* quanti sarebbero quelli concessi dalla numerosità dello spazio dell'operando meno profondo, che allo stato è quello d'una consueta chiave di $\boldsymbol{m}$ bit e cioè d'una misura che supponiamo inferiore a quella del messaggio medesimo,

per $\Omega_1 \subset \Omega\Omega_1$

A questo punto altresì diciamo $\boldsymbol{Z}$ il numero dei messaggi dotati di semantica e sintassi che appartengano a $\Omega\Omega_1$ e diciamo $\boldsymbol{z}$ il numero di quelli parimente dotati di semantica e sintassi, ma mediamente presenti all'interno della sola classe residua Ω_1 per $\boldsymbol{z} = \frac{\mathbf{Z}}{\mathbf{2^N/2^m}}$

Al dunque possiamo perciò logicamente affermare che nei sistemi convenzionali, si potrà tendere a una soluzione univoca tutte le volte in cui $\boldsymbol{Z}$ numero dei messaggi appartenenti a $\Omega\Omega_1$ diviso per il rapporto tra il maggior spazio di $\Omega\Omega_1$ e il minor spazio di Ω_1

possa dare $\frac{\mathbf{Z}}{\mathbf{2^N/2^m}} = \boldsymbol{z}$ per $0 < \boldsymbol{z} \leq 1$

Se consideriamo infatti $\Omega\Omega_1$ di ampiezza pari a 2^N come uno spazio ripartito in più sottoinsiemi che diciamo tutti Ω_1, di ampiezza pari a 2^m per $\Omega_1 \stackrel{\text{def}}{=} 1\Omega_1, 2\Omega_1, 3\Omega_1, \ldots, 2^{N-\boldsymbol{m}}\ \Omega_1$

si vedrà che ognuno sarà compatibile con un unico e solo possibile crittogramma a sua volta compatibile con una collezione di possibili messaggi la cui cardinalità è sempre data dallo spazio concesso della breve chiave di crittazione.

Intercettato infatti il crittogramma, di tali differenti sottoinsiemi che abbiam detti Ω_1 solo uno sopravvive così da fissare nei fatti la classe residua cui effettivamente appartiene il giusto messaggio;

e sarà per questo che nelle circostanze favorevoli al nostro oracolo,

$\boldsymbol{z}$ dovrà risultare uguale a un valore medio per cui $0 < \boldsymbol{z} \leq 1$

in quanto mettendoci nei panni di chi attacca, da un canto non vorremmo $\boldsymbol{z}$ uguale a *zero* giacché questi sarebbe il caso dove nessun consueto messaggio esiste, cosa questa che sarebbe in contrasto con le premesse che ci siamo date a inizio paragrafo,

e dall'altro saremmo tuttavia interessati a determinare uno spazio dove il giusto messaggio è così solitario nella classe residua da essere rintracciabile[37].

[37] Naturalmente il fatto se poi sia "effettivamente rintracciato" è altra storia.

Figura 08

Qui per fornire una immagine intuitiva dei sistemi a sicurezza computazionale, insieme $\Omega\Omega_1$ è irritualmente raffigurato sul piano di un quadrangolo che rappresenta uno spazio ripartito in più sottoinsiemi di messaggi, ciascuno dei quali dell'ampiezza di quello della chiave di minor lunghezza.

Insieme $\Omega\Omega_1$

Dove la misura del crittogramma fissa la "maggior lunghezza possibile" e quindi lo spazio di insieme $\Omega\Omega_1$ che qui corrisponde a quello del messaggio in chiaro.
Dove, essendo che le possibili chiavi date da tutte le configurazioni numeriche di sequenze di minor lunghezza *l*, sono meno, e essendo che i possibili messaggi dati da tutte le configurazioni di sequenze di maggior lunghezza *L*, sono più, l'ampiezza di ciascun sottoinsieme è condizionata dalla inferiore lunghezza della chiave che abbiamo supposto essere di quelle standard come ad esempio sarebbe per una misura di centoventotto bit.
Dove a ciascun sottoinsieme di spazio pari a $2^{\boldsymbol{m}}$ che abbiamo chiamato $1\Omega_1, 2\Omega_1, 3\Omega_1, \ldots, 2^{N-\boldsymbol{m}}\ \Omega_1$ corrisponde poi un distinto possibile crittogramma;
dove a ciascun crittogramma corrispondono tutte le combinazioni chiave-messaggio che danno in uscita solo e soltanto quel crittogramma medesimo.

In concreto il crittogramma intercettato consente di escludere tutte le classi che non corrispondono, lasciandone una soltanto che, dal punto di vista di chi attacca, si spera che abbia un unico messaggio riconoscibile nel senso che sia di senso compiuto.

Grandezze al Numeratore

(51)

Dopo aver pertanto fissata l'equivalenza, benigna per chi attacca, dove $\boldsymbol{z}$ sarà mediamente maggiore di 0 ma ugual-minore di uno, è bene poterla meglio definire.
Se ci siamo infatti concentrati su quanto celermente il rapporto tra la numerosità di insieme $\Omega\Omega_1$ e quella di sottoinsieme Ω_1 possa generare grandi numeri al denominatore, appare tuttavia ovvio che il valore in uscita da $\frac{\mathbf{Z}}{2^{\mathbf{N}}/2^{\boldsymbol{m}}}$ dipenda da entrambi i fattori in gioco;

volendo infatti continuare a guardare ai meccanismi che consentono a certi valori di schizzare, rimarcando il fatto che una manciata di bit crea delta miliardari negli spazi che ne discendono, procediamo inquadrando la susseguente questione del numero numeratore di siffatto rapporto tra grandezze.

If a secrecy system with a finite key is used, and N *letters of cryptogram intercepted, there will be, for the enemy, a certain set of messages with certain probabilities that this cryptogram could represent. As* N *increases the field usually narrows down until eventually there is a unique "solution" to the cryptogram; one message with probability essentially unity while all others are practically zero.*
Se viene utilizzato un sistema di segretezza con una chiave finita, e vengono intercettate N *lettere di crittogramma, ci sarà, per il nemico, un certo insieme di messaggi con determinate probabilità che questo crittogramma potrebbe rappresentare.*
All'aumentare di N*, il campo di solito si restringe fino a quando alla fine non c'è che una "soluzione" unica al crittogramma; un messaggio con probabilità essenzialmente uguale ad uno mentre tutti gli altri sono praticamente zero.*

Se il processo illustrato da Shannon *e di cui abbiamo tenuto ampiamente conto nel nostro lavoro* appare corretto, è per il fatto che la compatibilità col crittogramma non è meramente numerica.
Se fosse infatti meramente numerica, tutti i messaggi dell'insieme d'appartenenza sarebbero equi-probabili agli occhi di un osservatore giacché privi di qualsivoglia "riscontro" sintattico-semantico.
Tradotti tuttavia in lettere alfabetiche, alcuni messaggi affiorano spiccando dal rumore come gemme preziose:

?le!!9?-me-tè£ostrain-ma08/o?dati-a
?traildem—eln-vio!!memac-vio-lent?
nel-mezzo-del-cammin-di-nostra-vita
)p=?!himò!?n-op--ch-opo-;??min-l-i
?l?i!ii-ih—iopl---th-silghasqer-cc9))o

A voler tuttavia passare dalla disamina del valore del denominatore a quella del numero numeratore, che ci dice quale sarebbe la quantità dei messaggi dotati di semantica e sintassi presenti in $\Omega\Omega_1$, abbiamo che si determina una forma e diremmo una struttura tecnicamente rilevabile;
per cui diciamo che $\boldsymbol{z}$ messaggi si possono riconoscere in vario modo, ma pure attraverso procedure uguali e contrarie a quelle consigliate dai maggiori enti di standardizzazione del segnale, dei quali dicemmo parlando della qualità del bit stream ingenerato da sorgenti aleatorie.
Se nel caso si trattava di prendere per buone le sequenze random e di scartare quelle predicibili, qui saranno gli stream organizzati in linguaggio, a ottenere un riscontro negato a quelli che emergono dai flussi più disturbati[38].
Non si creda però che stiamo dicendo d'un metodo affidabile o storicamente impiegato, giacché esistono molteplici tecniche di riconoscimento meno qualitative ma più efficienti.
Ciò nondimeno, quello che interessa non è indicare questa o quella pratica tra quelle fattibili, quanto di rimarcare il carattere oggettivo di tali verifiche;
talmente oggettivo che da decenni le parabole del SETI (Search for Extra-Terrestrial Intelligence) puntano lo spazio siderale alla ricerca di linguaggi alieni captati dalle antenne di ricercatori che operano sul terzo pianeta del sistema solare, porgendo le orecchie al flusso dei segnali in arrivo notte e giorno da tramontate stelle[39].

[38] In un certo senso si tratta di discenere i flussi ingenerati da sorgenti non-random da noi dette $\boldsymbol{x}$-source da quelli che potrebbero in teoria discendere da fonti da noi dette $\boldsymbol{y}$-source.

[39] Qualunque messaggio ci possa raggiungere, proverrebbe da civiltà talmente lontane nel tempo e nello spazio, da essersi probabilmente estinte; il messaggio è reso riconoscibile proprio in quanto organizzato secondo standard logici e grammaticali.

Occorre aggiungere che se non potessimo riconoscere il carattere organizzato di talune sequenze che emergono da millanta altre, **nessun attacco a forza bruta e nessuna critto-analisi potrebbe mai avere successo.**

Sistemi Convenzionali

(52)

Se intendiamo dare tuttavia i numeri così da poterci meglio orientare, sarà necessario fornire degli indizi sul valore che può realisticamente assumere z senza cui non avremmo modo di immaginare alcun rapporto tra grandezza e grandezza; guardiamo allora ai casi esemplari che andremo a trattare nel resto del capitolo dove invece dell'inglese prenderemo in prestito la parola della lingua italiana.

(i)

Nel procedere saranno perciò dettate poche istruzioni di massima per le quali ..., adesso che ad esempio diremo dei sistemi convenzionali ..., *per facilitare le cose agendo in una sorta di mondo miniaturizzato* ogni chiave sarà immaginata per lunghezze inferiori a quelle effettive, mentre in modo analogo ciascun messaggio verrà convenzionalmente fissato d'una sola parola intellegibile così da escludere espressioni più articolate per cui sarebbe difficile azzardare un pronostico.

Diciamo dunque $\mathcal{C}_i$ un cifrario a sicurezza computazionale le cui specifiche per il principio di Kerckhoffs[40] sarebbero note a chi attacca.

Diciamo anche che in $\mathcal{C}_i$ si impieghi una immaginaria chiave di 24 bit la cui lunghezza supponiamo di dominio pubblico nel senso che sarebbe universalmente nota così come nel mondo reale lo sono quelle di 128-256-512 bit.

Altresì diremo che qualora l'attaccante e cioè il nostro Oracolo, dovesse intercettare il messaggio cifrato dalla cui misura calcola quello in chiaro, a questo punto sarebbe sia a conoscenza della lunghezza della chiave di 24 bit, che a conoscenza della lunghezza del messaggio che a titolo d'esempio supponiamo d'una misura pari a 72 bit.

Di conseguenza egli potrà calcolare l'ampiezza dello spazio del messaggio che corrisponde a quella dell'operando di maggior lunghezza per $\Omega\Omega_1 = 2^N = 2^{72}$ e potrà altresì misurare l'ampiezza dello spazio della chiave per $\Omega_1 = 2^m = 2^{24}$

Avendo dunque assodato che il critto-sistema scelto da Alice non è un mistero per nessuno e tantomeno per chi intenda venirne a capo in ogni maniera ..., nel piccolo mondo che ci siamo inventati, potremmo introdurre la regola secondo cui ogni segno o carattere alfabetico di ciascuna parola,
sarà fatto da una mini-stringa di sei bit che consente la forma di 64 simboli orto-alfa-numerici ottenuti da tutte le configurazioni concesse dalla base binaria.

Per cui si ricavano dalla scala delle possibili configurazioni:

[40] Pur senza citarne l'autore, avevamo già illustrato il principio per cui non è il critto-sistema ma la chiave a dover essere celata al nemico.

- quarantadue lettere delle quali 21 maiuscole e 21 minuscole,
- dieci cifre da zero e nove,
- un segno di *slash* e uno di spaziatura,
- dieci segni di punteggiatura che vanno dal punto alla virgola, alle virgolette e così andando.

Ciò pertanto significa che il messaggio di 72 bit sarebbe scandito da una sola parola di dodici caratteri che si deduce comparando i bit che formano il messaggio (72) con quelli impegnati nella costruzione di ciascun simbolo alfabetico (6).

Se tuttavia *è vero com'è vero* che il cosiddetto spazio del messaggio sarebbe in tal modo di 2^{72} prima di essere ridimensionato da quello della chiave di 2^{24},

è altrettanto vero che la lingua italiana contempla un numero ridotto di voci che rispondono ai detti requisiti; e cioè una quantità di parole effettivamente fatte di dodici lettere per un rapporto tra flussi casuali e parole intellegibili dove i primi hanno una molto maggiore probabilità di comparsa rispetto alle seconde.

E così a conti fatti, prendendo in carico,

- la quantità di bit di ciascun messaggio (72) così come dedotta dalla lunghezza del crittogramma;
- la quantità di bit impiegati in $\boldsymbol{6i}$ per ciascun carattere alfabetico (6) da cui si ottiene il numero di lettere che compongono il messaggio convenzionalmente fissato d'una sola parola;
- le voci che rispondono ai requisiti richiesti e che troviamo nel vocabolario Treccani della lingua italiana,

vedremo che *apriti sesamo* dal bussolotto del nostro computo[41] uscirà un numeretto per cui $\boldsymbol{Z = 9.203}$ che giustappunto sarebbero siffatte voci effettivamente presenti con tali caratteristiche in italiano,

(ii)

Dando dunque seguito alle nostre congetture scoprendo il risultato che avremmo sostituendo i valori ai simboli,

per cui $\frac{\mathbf{Z}}{\mathbf{2^N/2^m}} = \frac{\mathbf{9.203}}{\mathbf{2^{72}/2^{24}}} = \frac{\mathbf{9.203}}{\mathbf{2^{48}}} = \boldsymbol{z}$ dove $\boldsymbol{z} \ll 1$

possiamo intanto osservare a partire da tale esperienza tanto immaginaria quanto esemplare, che se *nei nostri sogni* non fossimo certi che un messaggio esiste essendo stato redatto da Alice,

e che sarà pure compatibile col crittogramma di cui rappresenta la forma palese ...,

quasi ci parrebbe impossibile pescare tale solitario dispaccio dalla classe emersa da quelle faticosamente dedotte dalla lunghezza della chiave.

In realtà il valore medio trovato intorno a numeri meno estremi di quelli solitamente impiegati nella vita reale dove i messaggi saranno in maggior numero (con crescita grosso modo lineare) ma le chiavi esponenzialmente più numerose,

[41] Che poi banalmente sarebbe la conta di tutte le parole di dodici lettere della lingua italiana.

appare talmente più piccolo di *uno*, da rendere ovvio il fatto che la stragrande maggioranza dei sotto-insiemi da scartare giacché in disaccordo col crittogramma intercettato, non potrebbe mai contemplare nemmeno un messaggio sensato.

Ben si comprende allora perché, quando in una situazione dove la chiave è per definizione meno lunga del messaggio,

col numeratore di $\frac{\mathbf{Z}}{\mathbf{2^N/2^m}}$ che cresce lentamente,

col denominatore che esplode con legge quadratica,

ci appaiono trascurabili le residue probabilità che un messaggio sintatticamente definito *non* sia quello giusto;

si crea infatti la condizione per cui il sistema è scalabile, sebbene con gran dispendio di risorse computazionali, essendo che si conferma l'ipotesi che esista una soluzione unica dissimulata tra gli elementi di Ω_1

Sistemi Perfetti

(53)

Se girovagando tuttavia tra congetture e costrutti mentali nei panni d'un novello arlecchino servo di due padroni, lasciamo la casa della sicurezza computazionale per tornare in quella dei sistemi perfetti, diremo ***cci*** un cifrario insensibile agli assalti.

Ora però prima faremo *coming out* dovendo confessare che la formuletta adottata aderisce perfettamente ai sistemi convenzionali quando un consueto messaggio (*nel mezzo del cammin di nostra vita mi ritrovai in una selva oscura...*) è cifrato da una chiave standard,

ma mostra la corda qualora si intenda migrare da tali sistemi a quelli perfetti, dove si rende opportuna una formulazione simile ma dal carattere più generale.

Sommando esempio ad esempio (fornendo casi fantasiosi ma efficaci che riproducono in sedicesimi stati realmente esistenti) intanto cominciamo a vedere cosa comporta l'opzione promossa da Shannon (**A**), dove nei sistemi perfetti da lui propugnati abbiamo un messaggio cifrato da una chiave di maggior grandezza[42].

(iii)

In primo luogo diciamo che quando la chiave è più lunga, poco cambia al numeratore **zz** che continua a ciecamente enumerare i messaggi dotati di semantica e sintassi che ora tuttavia appartengono a un insieme Ω_2 di inferiore numerosità rispetto a quello della chiave[43];

perciò sarà bene concentrarci su quanto accade al denominatore dove ci siamo imbattuti nel rapporto tra grandezza 2^N e grandezza 2^m

Sinora tale relazione *dividendo-divisore* è stata trattata pensando allo spazio d'un messaggio che fissa la numerosità dell'insieme di maggior ampiezza, e quello d'una chiave che regola la poca numerosità dell'insieme di minor grandezza,

ma quanto accade coi cifrari convenzionali non succede in quelli a sicurezza perfetta[44] su cui vorremo convergere a breve prima di dar corso a un ultimo passaggio.

Muovendo per questo come sempre dalla visuale di chi attacca vorremo proporre una diversa formula abbastanza elastica da includere sotto lo stesso tetto,

sistemi perfetti e non perfetti, convenzionali e non convenzionali ..., e cioè sistemi che *proprio in ragione del rapporto di misura tra chiave e messaggio* possano essere più o meno resiliente agli attacchi,

per cui diciamo $\frac{\mathbf{Z}}{\#\mathbf{M}/\#\mathbf{K}} = \boldsymbol{z}$ per $\boldsymbol{z} \leq 1$ tutte le volte in cui si possa giungere a una soluzione univoca.

[42] Dove se così non fosse, chiave e messaggio sarebbero uguali.
[43] Inferiore rispetto a quello del maggior spazio della chiave.
[44] Così come trattati da Shannon.

dove $\mathbb{M}$ giustappunto sarebbe l'insieme dei messaggi,

dove a tale insieme $\mathbb{M}$ appartiene ***m*** di variabile *M* che rappresenta ciascuno dei possibili dispacci siano essi random o non random,

dove $\mathbb{K}$ è l'insieme delle chiavi,

dove a essa $\mathbb{K}$ altresì appartiene ***k*** di variabile *K* che rappresenta ciascuna delle possibili chiavi di crittazione.

Per inciso noi abbiam detto e ridetto nel corso del corrente lavoro, di come non conti la distinzione tra chiave e messaggio, essendo che in fase di codifica-decodifica le sequenze adottate non possono che essere trattate quali mere grandezze numeriche.
Pur tuttavia anche dicemmo che in uno stato iniziale quando è redatto il messaggio e selezionata la chiave "la distinzione chiave-messaggio appare ancora giustificata".
Se tale tassonomia non ha senso in fase di codifica quando si procede attraverso operazioni matematiche che non lasciano spazio all'immaginazione ..., assume rilievo (e ci mancherebbe!) prima e dopo di esse, essendo che nessuno serberà gelosamente la chiave per buttare nel cestino il messaggio!
Ciò detto si rimarca che appartengono a insieme $\mathbb{M}$ tanto messaggi consueti che messaggi da noi detti speciali i quali manifestano un andamento aleatorio anche perché impiegati nella costruzione di chiavi crittografiche.
Questa davvero non sarebbe una novità, se non fosse che l'ampiezza di tale famiglia comporta conseguenze frettolosamente negate dallo stesso Shannon quando sembra rivendicare l'irrilevanza di ciò rispetto al problema ingegneristico della comunicazione cifrata. Cosa che vedemmo essere vera e non vera a seconda di come la si legga (04).

(iv)
Se voltandoci indietro abbiam fatto dunque nostro nei precedenti paragrafi, il dato per cui nei sistemi convenzionali il segnale-messaggio è *con rarissime eccezioni* più lungo di quello della chiave **(i)(ii)** per cui il relativo spazio sarebbe più ampio per 2^N maggiore di $2^{\boldsymbol{m}}$,
allargando i nostri orizzonti oltre i sistemi standard, non potremmo aderire a tale enunciato che a seconda dei casi daremo infatti per vero oppure per falso a seconda di come si presentino di volta in volta le cose,

$$\text{e infatti } f(\mathbf{M}) = \begin{cases} 2^N \textit{ se } \boldsymbol{m} \textit{ di } M \in \mathbb{L} \\ 2^{\boldsymbol{m}} \textit{ se } \boldsymbol{m} \textit{ di } M \notin \mathbb{L} \end{cases}$$

$$f(\mathbf{K}) = \begin{cases} 2^{\boldsymbol{m}} \textit{ se } \boldsymbol{m} \textit{ di } K \notin \mathbb{L}' \\ 2^N \textit{ se } \boldsymbol{m} \textit{ di } K \in \mathbb{L}' \end{cases}$$

dove **L** sarà l'insieme di tutti i messaggi di lunghezza *L* per $L \geq l$ ricordando essere *L* la misura dell'operando di maggior lunghezza,

dove **L'** invece sarebbe l'insieme delle chiavi sempre di lunghezza *L* che si hanno quando sia la chiave e non il messaggio l'operando di maggior lunghezza.

Ora se abbiamo sinora abbracciate due condizionalità (messaggio di senso compiuto, messaggio più lungo della chiave standard) che in una diversa ottica non sono tuttavia ineluttabili ..., intanto vorremmo sondare se in attesa di voltare pagina per dire dei sistemi perfetti ..., la novella formula si possa facilmente adattare alla vecchia così da non smarrire quanto avevamo acquisito.
In altre parole intendiamo vedere se la nuova formulazione abbia carattere generale come auspicato nelle scorse pagine, tornando a quei casi in precedenza risolti e che dovranno perciò apparire egualmente soluti alla luce della nuova formalizzazione.

Il che sembra ovvio essendo che $\frac{\mathbf{Z}}{\#\mathbf{M}/\#\mathbf{K}} = \boldsymbol{z}$ con poche trasformazioni si accorda facilmente a quanto detto,

avendo $\#\mathbf{M} = 2^{\mathrm{N}}$ quando ciascun messaggio sia elemento di **L**

e avendo $\#\mathbf{K} = 2^{\boldsymbol{m}}$ quando nessuna chiave sia elemento di **L'**

così da poter senz'altro scrivere $\frac{\mathbf{Z}}{\#\mathbf{M}/\#\mathbf{K}} = \frac{\mathbf{Z}}{2^{\mathrm{N}}/2^{\boldsymbol{m}}}$ che conferma ogni pregressa congettura.

(v)

Cominciando a far tuttavia cadere come in una danza autunnale la prima delle condizionalità in precedenza fissate, *giacché per il teorema di Shannon sulla perfetta segretezza non sarebbe il messaggio a essere di maggior lunghezza ma la chiave crittografica*, in effetti si mette a nudo una diversa equivalenza dove i valori cambiano sotto la linea del numeratore,

per cui $\frac{\mathbf{Z}}{\#\mathbf{M}/\#\mathbf{K}} = \frac{\mathbf{Z}}{2^{\boldsymbol{m}}/2^{\mathrm{N}}}$ in quanto diversamente da prima stavolta sarebbe chiave $\boldsymbol{k}$ uguale N bit e

messaggio $\boldsymbol{m}$ uguale $\boldsymbol{m}$ bit.

Sostituendo perciò tali valori ai simboli, dando tuttavia $\boldsymbol{m} = 72$ bit che consente parole di dodici lettere alfabetiche, e dicendo $\mathrm{N} = 73$ bit così da favorire con un esempio semplicissimo lo sviluppo intuitivo dell'espressione,

dicendo $\boldsymbol{Z}$ la quantità dei messaggi possibili che ormai sappiamo essere uguale a $\boldsymbol{9.203}$, avremo $\frac{\mathbf{Z}}{2^{\boldsymbol{m}}/2^{\mathrm{N}}} = \frac{\mathbf{9.203}}{2^{72}/2^{73}} = \frac{\mathbf{9.203}}{\mathbf{0,5}} = \boldsymbol{18.406}$ dove $\boldsymbol{18.406}$ sarebbe maggiore di 1 così che al nostro Oracolo non rimane che cercare il giusto messaggio alla cieca.

Il fatto che la numerosità sembri crescere sino a superare il numero di parole che offre la lingua italiana, svela il fatto che dalla prospettiva dell'Oracolo che conosce la sola maggior lunghezza del flusso della chiave dedotto da quello del crittogramma,
nel caso i possibili messaggi sarebbero esattamente il doppio di quelli reali; egli non ha contezza delle proporzioni essendo che tra chiave e messaggio ci potrebbe stare un diverso rapporto di misura (diverso rispetto a quello effettivo).
Per lui i messaggi sono tutti quelli consentiti dalla lunghezza del crittogramma intercettato che fissa la sola asticella che non può valicare[45].

(vi)

Facendo allora cadere come nella danza dei sette veli anche la seconda e ultima condizionalità così da tornare al cuore del nostro impegno dove c'era un messaggio random di maggior lunghezza e una chiave non random cui invertimmo i ruoli da quando dicemmo del paradosso del crittografo,

sarà di nuovo $\mathcal{CCi}$ il cifrario noto a chi attacca.

Verrà tuttavia trasmesso adesso da Alice un segnale ingenerato da sorgente aleatoria per una lunghezza che si deduce da quella del crittogramma.

Per tal motivo presumiamo che in ossequio ai lemmi (**B**)(**C**) per cifrare tale speciale messaggio sia impiegata una minima chiave non random la cui lunghezza supponiamo nota all'Oracolo non perché così sia, ma allo scopo di mettere meglio a fuoco attraverso tale concessione alcuni aspetti che ci sembrano interessanti.

Assumendo perciò che chi attacca abbia contezza della misura della chiave che sarebbe effettivamente data da una stringa di $\boldsymbol{m}$ bit,

anche ora chi attacca sarebbe in grado di calcolare tanto la maggior ampiezza di $\Omega\Omega_2$ che quella di Ω_2 che qui corrisponde allo spazio della chiave medesima.

A partire allora dall'eguaglianza per cui $\frac{\mathbf{Z}}{\#\mathbf{M}/\#\mathbf{K}} = \frac{\mathbf{Z}}{2^{\mathrm{N}}/2^{\boldsymbol{m}}} = \boldsymbol{z}$ dove $\boldsymbol{z} \leq 1$ tutte le volte in cui sia possibile giungere a una soluzione univoca,

abbiamo che $\boldsymbol{Z}$ è data da tutti i messaggi che insieme $\Omega\Omega_2$ consente, essendo che non entrano più in gioco le parole di questa o quella lingua giacché ogni configurazione sarebbe in egual modo possibile e parimente probabile per definizione.

In ciò registriamo una non banale differenza dai casi in precedenza trattati, in quanto per la prima volta il bacino di parole random entro cui cercare il messaggio non è inferiore ma uguale a quello della classe di riferimento dove $\#\Omega\Omega_2 = 2^{\mathrm{N}}$.

[45] Dal già citato *Secrecy Systems* di Claudio Cappelli.

Invero anche il risultato rilasciato in uscita da $\boldsymbol{z}$ sarebbe fornito dal numero di messaggi random (assunti con pari distribuzione di probabilità) ma non da tutti; sarebbero infatti ammessi solo quelli mediamente presenti nella classe residua Ω_2 di numerosità pari a $2^{\boldsymbol{m}}$ che sappiamo fissata dallo spazio dell'operando di minor lunghezza[46].

(vii)

A voler tuttavia tornare a fare di conto per non smarrire il senso delle cose, riprendendo esempi più volte citati ma ricicciati ai nostri fini,

assumiamo di avere un messaggio aleatorio di lunghezza uguale a 72 bit, e una chiave di minor lunghezza pari a 24 bit la cui conoscenza concedemmo al nemico senza farci troppe domande.

Si procede pertanto senza scossoni per cui abbiamo $2^{\mathrm{N}}/2^{\boldsymbol{m}}$ che fa da denominatore del rapporto di divisione dove $2^{\mathrm{N}}/2^{\boldsymbol{m}} = 2^{\boldsymbol{72}}/2^{\boldsymbol{24}} = 2^{\boldsymbol{48}}$ (denominatore)

e abbiamo $\boldsymbol{Z}$ che funge da numeratore del rapporto esprimendo tanti possibili messaggi quanti quelli concessi dall'ampiezza di $\Omega\Omega_{\boldsymbol{2}}$ per cui $\boldsymbol{Z} = 2^{\mathrm{N}} = 2^{\boldsymbol{72}}$. (numeratore)

In simboli perciò avremmo $\frac{\mathbf{Z}}{\mathbf{2^N}/\boldsymbol{2^m}} = \frac{\mathbf{2^N}}{\mathbf{2^N}/\boldsymbol{2^m}} = \frac{\mathbf{2^N}}{\mathbf{2^{N-m}}} = 2^{\boldsymbol{m}} = \boldsymbol{z}$ dove essendo $2^{\boldsymbol{m}}$ sempre maggiore di uno, il risultato esclude da ogni probabilità il fatto che ci si possa lontanamente avvicinare al caso di quella *unique solution* vagheggiata da Shannon.

[46] Invero il nostro Oracolo non avrebbe dovuto conoscere l'esatta misura della chiave giacché questa, più piccola del messaggio random, non si evince dalla maggior lunghezza del crittogramma di profondità uguale a quella del dispaccio ma non della chiave più piccola.

Se concediamo in effetti tale vantaggio, è per il fatto che intendiamo dimostrare come – pur restringendo la classe residua in rapporto alla lunghezza della chiave *no-random* che pure dovrebbe essere ignota – lo stesso si crea uno spazio abitato da moltitudini di messaggi parimente probabili e quindi indecidibili dalla visuale di chi intenda forzare il sistema.

Per chiudere il cerchio, si tenga altresì conto che dalla versione punto zero della formuletta impiegata **(i)** abbiamo lasciato cadere il limite inferiore essendo che non scriviamo più $0 < \boldsymbol{z} \leq 1$ ma $\boldsymbol{z} \leq 1$ (*continua a pag.57*);

il punto è che una volti contemplati nella stessa espressione $\frac{\mathbf{Z}}{\mathbf{\#M/\#K}}$ messaggi random e *no-random*, cambia di significato il caso dove non ci sono messaggi di senso compiuto essendo che adesso semplicemente significa che le sequenze da ricercare saranno esse stesse di carattere aleatorio.

Tabella

Tipo di Sistema	**Descrizione**	**Formula Adottata**	**Note**
A Sicurezza Computazionale	*Il caso tipico sarebbe quello dove abbiamo un messaggio di senso compiuto di maggior lunghezza rispetto alla chiave standard.* *La chiave solitamente di 128-256-512 bit, può essere random*	$\frac{\mathbf{Z}}{\#\mathbf{M}/\#\mathbf{K}} = \frac{\mathbf{Z}}{2^{\mathrm{N}}/2^{\boldsymbol{m}}} = \boldsymbol{z}$	*Qui il valore di* $\boldsymbol{Z}$ *dipende dall'effettivo numero di frasi di senso compiuto che esistono in rapporto alla lunghezza del messaggio accertata*
A Sicurezza Perfetta secondo Shannon Teorema (A)	*Il messaggio di senso compiuto è di lunghezza arbitraria.* *La chiave true random è di lunghezza ugual maggiore rispetto a quella del messaggio*	$\frac{\mathbf{Z}}{\#\mathbf{M}/\#\mathbf{K}} = \frac{\mathbf{Z}}{2^{\boldsymbol{m}}/2^{\mathrm{N}}} = \boldsymbol{z}$	*Qui il valore di* $\boldsymbol{Z}$ *dipende dall'effettivo numero di frasi di senso compiuto che esistono in rapporto alla lunghezza del messaggio che è stato possibile accertare.* *La chiave random di uguale o maggior lunghezza rende i messaggi equiprobabili a valle*
A Sicurezza Perfetta secondo i lemmi (B)(C)	*Il messaggio può pur essere esso stesso true random e, nel caso, dovrà essere di lunghezza ugual maggiore rispetto a quella della chiave.* *La chiave sarà dunque tipicamente più piccola e no-random*	$\frac{\mathbf{Z}}{\#\mathbf{M}/\#\mathbf{K}} = \frac{\mathbf{Z}}{2^{\mathrm{N}}/2^{\boldsymbol{m}}} = \frac{2^{\mathrm{N}}}{2^{\mathrm{N}-\boldsymbol{m}}} = 2^{\boldsymbol{m}} = \boldsymbol{z}$	*Qui il valore di* $\boldsymbol{Z}$ *dipende dalla sola numerosità dell'insieme di riferimento. Il messaggio esso stesso random risulta essere equiprobabile alla fonte* **Si confronti più innanzi (60)**

Capitolo VIII

Tema

(54)

Non pochi aspetti in precedenza trattati sono qui ripresi ma da un punto di vista più pratico, così da rimarcare come taluni approcci effettivamente permettano di ergere le difese di un sistema che non sarà matematicamente scalabile.

Si illustra l'importanza di quello che abbiamo chiamato "riscontro" e cioè di quella validazione che consente di distinguere tra millanta messaggi quello corretto, la cui probabilità di comparsa sarebbe da un certo momento in avanti prossima a uno.

Viene segnalato infine il fatto che tale opera di riscontro trova un ostacolo insormontabile quando il messaggio sia esso stesso aleatorio;

prendendo poi atto dell'incompletezza del teorema di Shannon sulla perfetta segretezza è formulato un teorema originale.

Sunto e Premesse (1)

(55)

Nel presente lavoro abbiamo dato più spunti e fissate molteplici milestone che *da un canto* ci hanno permesso di focalizzare le criticità presenti nella consueta definizione di perfetta segretezza, *e dall'altro* rappresentano le premesse logiche per giungere a un teorema che intendiamo esporre in dirittura d'arrivo.

Bisogna tuttavia dire che quanto da noi illustrato non sempre conduce a conclusioni perfettamente fungibili, e proprio per questo sarà nostra cura riassumere i temi trattati facendoli convergere nella lettera di un ultimo enunciato.

Contraddizioni e paradossi, a cominciare da quello del crittografo ..., ci hanno mostrato come sia dubbia l'idea di far coincidere il concetto di perfetta segretezza con quello di sicurezza incondizionata, essendo che il nocciolo della questione sta nel fatto che alcuni sistemi sono computazionalmente vulnerabili, mentre altri non lo sono persino se ingaggiati da un Oracolo dalle risorse infinite.

Questo ci ha consentito di offrire le seguenti formalizzazioni che così riassumiamo:

Lemma **(B)**

Ai fini della perfetta sicurezza di un sistema crittografico, sarà sufficiente che in codifica almeno uno degli operandi abbia un andamento aleatorio a prescindere se sia chiave o messaggio;

Lemma **(C)**

Ai fini della perfetta sicurezza di un sistema, essendo che almeno un operando ha carattere aleatorio, esso dovrà pur essere di lunghezza ugual-maggiore rispetto a quella dell'altro operando, a prescindere se sia chiave o messaggio.

Definizione

Diremo di avere sicurezza perfetta quando un attaccante dotato di risorse infinite, pur conoscendo il valore del crittogramma, abbia una probabilità prossima a 0 di risalire a messaggio m dove m appartiene a una classe sufficientemente ampia di messaggi tra loro equi-probabili.

Quest'ultima dicitura è stata poi integrata da un meccanismo algebrico che mostra come si possano individuare classi residue che talora conducono a una soluzione univoca;

cosa che resta nondimeno impossibile quando si operi nell'ambito dei sistemi perfetti come fissati nei lemmi (**B**)(**C**) e nella definizione che ne consegue oltre che nel teorema da dimostrare nel corrente capitolo.

(56)

Ciò di cui abbiamo trattato attiene alle condizioni per cui chi attacca passa da classi più ampie a più strette, man mano che accumula informazioni che gli consentono di scartare taluni messaggi per lasciar vivere altri.

E' qui che assume rilievo la questione del "riscontro" ed è qui che sorge una pesante linea di demarcazione tra sistemi perfetti e non perfetti.

Infatti tale "riscontro" può essere parziale e perciò inefficace nei sistemi non convenzionali;

tutto questo accade quando pur essendoci modo di passare da una classe più ampia a una meno, il processo si blocca giacché genera sottoinsiemi *che sarebbero comunque sufficientemente grandi* di elementi tra loro equi-probabili.

Nel caso è anche facile comprendere quanto dicemmo e cioè che la comparazione tra:

(**a**) un set $\mathbb{A}$ restato uguale a se stesso anche alla luce della conoscenza del crittogramma (sebbene nei limiti da noi prospettati),

(**b**) una classe residua tratta da un insieme $\mathbb{B}$ ridimensionato dal valore del crittogramma o da quanto si sappia sulla lunghezza della chiave e non solo,

può anche vedere tale classe esprimere più numerosità della prima, così che da essa possa discendere meno probabilità condizionata di giungere fortuitamente al giusto messaggio.

Se nell'esempio incondizionato in (**a**) abbiamo una classe di ampiezza pari a 2^{m}

e nell'esempio condizionato in (**b**) abbiamo una classe residua di ampiezza pari a 2^{N}

per $m < N$

laddove sia gli elementi dell'insieme $\mathbb{A}$ che quelli in $\mathbb{B}$ siano dati come equi-probabili e quindi indecidibili agli occhi del nostro Oracolo (e sul piano logico poco importa come e quando ciò sia possibile) quasi si direbbe superstizione non tenerne conto.

I sistemi da noi detti perfetti in base a quanto abbiamo potuto sensatamente affermare, infatti lo sono per la capitale ragione che non sono scalabili concedendo un numero sufficientemente alto di soluzioni tra loro parimente probabili.

How immune is a system to cryptanalysis when the cryptanalyst has unlimited time and manpower available for the analysis of cryptograms?

Does a cryptogram have a unique solution (even though it may require an impractical amount of work to find it) and if not how many reasonable solutions does it have?

How much text in a given system must be intercepted before the solution becomes unique?

Are there systems which never become unique in solution no matter how much enciphered text is intercepted?

Quanto è immune un sistema alla crittoanalisi quando il critto-analista ha una quantità illimitata di tempo e forza lavoro a disposizione per l'analisi dei crittogrammi?

Ha un crittogramma una one solution (anche se potrebbe richiedere un impraticabile accumulo di lavoro per trovarla) e se non, quante soluzioni ragionevoli avrebbe?
Quanto testo in un determinato sistema deve essere intercettato prima che la soluzione diventi unica?
Esistono sistemi che non danno mai una soluzione unica, non importa quanto testo cifrato sia intercettato?

(57)
Le domande formulate da Shannon mentre teneva lezioni e scriveva brevi saggi nelle sulfuree notti del 1949 ..., pur con opportuni accorgimenti sono le stesse che ci siamo fatte da soli e che riprenderemo così da accompagnare una a una le nostre conclusioni.

- *How immune is a system to cryptanalysis when the cryptanalyst has unlimited time and manpower available for the analysis of cryptograms?*
- *How much text in a given system must be intercepted before the solution becomes unique?*

Sebbene oggigiorno nessuno faccia cenno a quella forza lavoro o manpower della quale all'epoca non si diceva in senso figurato …, in cuor nostro vorremo guardare per un momento a tali argomenti anche in rapporto alla questione della disponibilità su larga scala di risorse computazionali.
In effetti qui dovremmo affrontare temi che promettono di portarci fuori strada e non intendiamo rischiare di essere dirottati dove non vogliamo ..., ma il fatto di cui intendiamo dare perlomeno il titolo sta nella stima secondo cui, sebbene la transizione dall'elaborazione elettronica in bit a quella in qbit delle macchine quantiche sia tortuosa ..., essa è inarrestabile così da mettere a repentaglio gli attuali sistemi più convenzionali.
Quindi come tante volte in passato (un passato storicamente rilevabile) sappiamo che qualcosa deve accadere sebbene non sia possibile stabilire come e quando;
invero non è nemmeno facile pronosticare la potenza di calcolo di tali mezzi futuribili, essendo che dipenderà da più variabili che attengono all'efficienza nel controllo degli errori in fase di manipolazione delle particelle quantiche,
alla loro reciproca comunicazione da punto a punto,
alla gestione delle conversioni dal mondo della meccanica classica a quello quantico e viceversa,
alla redazione di algoritmi concepiti allo scopo, che rappresentano un problema nel problema dovendo esser calati da un ambiente deterministico a quello random che permea lo stato sub-atomico.
Resta tuttavia che non sappiamo quali saranno le proprietà e le capacità di calcolo nel prossimo futuro, ma sappiamo che saranno più grandi.
E ciò vuol dire che le simulazioni dove ipotizziamo maghi dalla potenza sovrumana e oracoli dalle risorse senza alcun limite se non quello della nostra immaginazione ..., potrebbero trovare qualche mezza risposta nel giro di pochi anni.

Così che la fantasiosa ipotesi di Shannon, buttata lì per finta e secondo cui "that equivocation is a theoretical secrecy index—theoretical in that it allows the enemy unlimited time to analyse the cryptogram" "tale incertezza è un indice teorico della segretezza, teorico in quanto concede al nemico un tempo illimitato per analizzare il crittogramma" rischia d'apparire sempre più palpabile.

Sunto e Premesse (2)

(58)

- *How much text in a given system must be intercepted before the solution becomes unique?*
- *Does a cryptogram have a unique solution (even though it may require an impractical amount of work to find it) and if not how many reasonable solutions does it have?*

Non è che finora non si sia detto nulla in proposito (anzi) ma qui intendiamo armonizzare precedenti passaggi per dare una indicazione assai pratica;
si tratta cioè di mettere a fuoco dei meccanismi che sovente sfuggono anche a chi ne conosce i fondamenti logici.
Intendiamo dire che ci sono risvolti elementari che proprio per questo appaiono come invisibili, sia perché chi maneggia sistemi di crittazione non sempre possiede sufficienti basi teoriche,
e sia perché chi le possiede, sovente ha smanettato poco coi programmi di codifica e decodifica di messaggi segreti.
Non è un caso che giganti come Shannon e Turing abbiano rappresentato, sebbene in epoche affatto diverse dalla nostra, delle solenni eccezioni.
Abbiamo intanto detto che vorremo abbozzare una risposta alle domande poste settant'anni fa e restate a mezz'aria col loro punto di interrogazione;
vorremmo dare un cenno per scendere sulla terra cogliendo meglio il nesso tra "informazioni" e "possibilità" o "impossibilità" di scalare un sistema quando si operi con cifrari a sicurezza perfetta o computazionale, ricercando classi sempre più fini fin quando la trappola non scatti (quando succede) dietro alla scelta di un singolo messaggio.
Abbiamo visto come sia talora fattibile restringere il campo di chi attacca, il quale si chiederà *in primis* della misura di chiave e messaggio e dello spazio cui appartengono;
ciò mette tuttavia in risalto quanto sia arbitrario fissare uno spazio dei messaggi (della lunghezza del messaggio di Alice) nonostante sia precisamente tale spazio a essere ignoto a colui che attaccando lo dovrebbe sfruttare a proprio vantaggio.
Si è pur detto di come l'idea che esistano segnali di cui non sarebbe possibile raccogliere tracce giacché dalla incondizionata probabilità di comparsa …, si scontra col fatto secondo cui il *core* di talune informazioni non muore;
tali segnali non sono matematicamente scalabili ma nemmeno si possono definire dalla probabilità "incondizionata" essendo che a rigore lo sarebbero impiegando chiavi esse stesse infinite.

Esempio

Assegnare tuttavia di nuovo un cifrario $\mathcal{Ci}$ è un espediente per rimettere le mani in pasta attraverso l'offerta di un ultimissimo esempio;
esiste la teoria e esiste la pratica con le sue astuzie e noi vorremmo mostrare attraverso quali riscontri un attaccante effettivamente comincia quel lavoro di affinamento che lo può o meno condurre a quella soluzione con cui darà scacco al sistema.
Insomma noi dicemmo in modo temerario in (41) che le classi residue si possono ridurre "a un pizzico" e che per fare ciò occorrono imponenti risorse computazionali,
ma adesso ci chiediamo come e perché il giocatore dissimulato dietro le sembianze di qualsivoglia cripto-analista muove in un modo piuttosto che un altro i pezzi sulla sua particolare scacchiera.
A prescindere da ogni risposta sui singoli metodi di crittazione (dove qui saremmo all'altro capo di siffatta analisi) vorremmo battere su un tasto che proprio perché elementare è spesso trascurato ..., così che talune congetture ci appaiono lunari e assai lontane dalla partita che giornalmente si consuma nei luoghi dove si scontrano i buoni coi cattivi.

Esempio Molto Semplificato

Supponiamo nel quadro di riferimento di cifrario $\mathcal{Ci}$ e quindi nel quadro di un sistema convenzionale *difficile ma non impossibile da scalare* che sia stato intercettato il seguente crittogramma di otto bit che diciamo uguale a 11111111.
Altresì presumiamo con un flash che tale cifrario impieghi un algoritmo con una semplice somma XOR frapposta tra messaggio e chiave di sole quattro cifre binarie (tutte cose poco plausibili se prese alla lettera, ma che nell'economia dell'esempio ci facilitano non poco);
da ciò ricaviamo il fatto che le possibili chiavi crittografiche sarebbero date dai valori di tutte le configurazioni a quattro bit che sono quelle di 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111.
Altresì sappiamo che nella somma a flusso ciascuna chiave è reiterata sino a raggiungere la lunghezza del messaggio medesimo che sarebbe di otto bit in accordo con la profondità di quello cifrato[47] (**laddove con differenti algoritmi** *sebbene per altre vie* **nondimeno si giunge con maggiore o minore approssimazione alla misura dell'operando di maggior lunghezza**).

Avendo perciò contezza del fatto che assunta una chiave di quattro bit,
il messaggio può esser dato da uno e uno soltanto dei valori di:

[47] Si fa presente che reiterando la chiave per pareggiare gli otto bit della lunghezza del messaggio, si avrebbero tutte e solo le seguenti possibili configurazioni binarie: 00000000, 00010001, 00100010, 00110011, 01000100, 01010101, 01100110, 01110111, 10001000, 10001000, 10101010, 10111011, 11001100, 11011101, 11101110, 11111111

11111111, 11101110, 11011101, 11001100, 10111011, 10101010, 10101010, 10101010, 10001000, 01110111, 01100110, 01010101, 01000100, 00110011, 00100010, 00010001, 00000000 che esauriscono il set di quelli in grado di rilasciare 11111111 come risultato,
dobbiamo costatare che il valore del crittogramma è effettivamente compatibile con soli sedici messaggi dove gli altri duecentoquaranta che appartengono a $\overline{\Omega}_1 \subset \Omega\Omega_1$ sarebbero immediatamente da escludere, come sarebbero da escludere gli infiniti messaggi appartenenti ad $\mathbb{E}_1$ ma non a Ω_1 col risultato di avere un setaccio che potrebbe alfine portare alla soluzione finale,
cosa che sappiamo essere possibile nei sistemi a sicurezza computazionale mappando messaggio dopo messaggio,
andando cioè a escludere quelli privi dei più elementari rudimenti sintattici, così da restringere in modo sempre più sottile l'imbuto che conduce al messaggio di Alice.

(59)

- *Does a cryptogram have a unique solution (even though it may require an impractical amount of work to find it) and if not how many reasonable solutions does it have?*
- *Are there systems which never become unique in solution no matter how much enciphered text is intercepted?*

Come detto e ridetto (33) non meraviglia che sia stato lo stesso Shannon a introdurre la questione della unicità del messaggio intesa come soluzione cui tendere individuando classi sempre più sottili;
tutto ciò a patto di riconoscere che il conseguimento di tale soluzione non sarà percorribile quando l'attacco sia rivolto a un cifrario perfetto.
Stranamente nel suo saggio, il padre della teoria informazionale non percorre a sufficienza tale felice intuizione della quale non sembra pesare tutte le corpose conseguenze.

Per quanto ci riguarda, nella narrazione siamo tuttavia giunti al punto dove chi attacca ha potuta comunque rintracciare una classe residua,
avendo selezionato un certo numero di possibili dispacci in funzione della misura dell'operando meno profondo che in $\mathcal{Ci}$ sarebbe praticamente quello della chiave che sappiamo usualmente essere nei cifrari simmetrici di 128-256-512 bit.

(60)

Ora è un fatto che si sono sviscerati più aspetti del crivello condotto dal nostro Oracolo, ma qui intendiamo scorgere alcuni passaggi molto concreti attraverso cui si giunge o non si giunge alla soluzione di cui dicemmo.

Intanto quello indicato come "riscontro" può aver luogo per più motivi cui abbiamo fatto cenno nella nostra collezione di minimi apologhi,
ma esiste una ragione talmente preminente nella quale un messaggio di senso compiuto prevale sugli altri, che ci concentreremo su questa.
La classe residua figuratamente selezionata Ω (dove $\Omega = \Omega_1\ \Omega_2 \ldots$) intanto ci consente d'afferrare il nesso tra l'insieme di tutti i messaggi calcolati sulla scorta della lunghezza del crittogramma medesimo,
e il sottoinsieme di quelli ancora possibili una volta conosciuta *se conosciuta* la misura della chiave di minor lunghezza.
La portata di tali spazi cresce esponenzialmente al crescere dei valori di riferimento, dove abbiamo ΩΩ (per $\Omega\Omega = \Omega\Omega_1\ \Omega\Omega_2 \ldots$) di ampiezza pari a 2^N
e abbiamo Ω (per $\Omega = \Omega_1\ \Omega_2 \ldots$) di ampiezza pari a $2^{\boldsymbol{m}}$
Il rapporto $2^N / 2^{\boldsymbol{m}}$ da noi lungamente trattato, tra la numerosità data dall'ampiezza di ΩΩ e quella data dall'ampiezza di Ω ci dice di quante volte #Ω sarà minore di #ΩΩ, fatto salvo il diverso caso dove #Ω = #ΩΩ
Per certi versi possiamo anche dire che più grande sia il valore in uscita dal rapporto tra grandezze,
e maggiori sono le speranze d'avere una situazione gestibile, essendo che crescendo il denominatore di $\frac{\mathbf{Z}}{\mathbf{2^N}/\boldsymbol{2^m}}$ si abbatte la quantità media di possibili dispacci.

In termini pratici ciò spesso corrisponde al caso dove sia impiegata una chiave abbastanza corta per un messaggio sufficientemente lungo, per cui è più agevole giungere a rompere il sistema come accaduto pubblicamente per chiavi inferiori a 128 bit.
Ora possiamo tuttavia osservare un fatto singolare che attiene al caso del quale ci siamo appassionati, e cioè quello dove il messaggio non riguarda un segnale di senso compiuto,
ma un segnale esso stesso random per il trasferimento di flussi da impiegare nella costruzione di chiavi crittografiche.
In tale circostanza, per quanto si possa figurare un rapporto favorevole a chi attacca per N molto grande ed $\boldsymbol{m}$ molto piccola,
avremo che nessuna ulteriore cernita tra messaggio e messaggio appartenente a Ω sarà tuttavia fattibile.
Saremmo infatti giunti davanti a quel riscontro "parziale" di cui dicemmo e che si blocca sulla soglia della classe residua comunque rintracciata.
Il motivo è semplice ed è stato rappresentato in modo formale nell'espressione per cui $\frac{\mathbf{Z}}{\mathbf{2^N}/\boldsymbol{2^m}} =$

$\frac{\mathbf{2^N}}{\mathbf{2^N}/2^{\boldsymbol{m}}} = \frac{\mathbf{2^N}}{\mathbf{2^N}-\boldsymbol{m}} = 2^{\boldsymbol{m}} = \boldsymbol{z}$ per $2^{\boldsymbol{m}}$ che dando un risultato molto maggiore di uno, nemmeno si avvicina al caso d'una soluzione univoca sapendo N ed $\boldsymbol{m}$ interi positivi dove $N \geq \boldsymbol{m}$

Traducendo tutto ciò in linguaggio naturale, tuttavia si evidenziano almeno due aspetti di particolare interesse sui quali ci piace convergere.
Intanto, essendo che la ragione della incertezza a valle, nella circostanza discende dalla equi-probabilità a monte,
abbiamo che nessun riscontro è intrinsecamente possibile qualora ci siano messaggi presi da un qualche insieme di riferimento che offra alla nascita una uniforme distribuzione di probabilità;
assurdamente il trasferimento di flussi dal carattere aleatorio da sempre pittato come il problema dei problemi, degenera quasi sempre in uno stato di perfetta sicurezza dove è impossibile discernere tra caso e caso;
quando infatti flussi aleatori palesano una "hard unpredictability" alla fonte, davvero non potranno magicamente aggiungere "certezza" in uscita, e ogni riscontro sarebbe impossibile a conclusione del processo di codifica.
Nello specifico sarebbe infatti sufficiente controllare che la maschera non sia così banale da rendere trasparente il messaggio true random e, a voler parafrasare un esempio precedentemente offerto, avremmo che tornando dal digitale all'alfanumerico,
si andrebbe a creare una situazione del seguente tipo che bene ci fa intendere come ogni riscontro sia impossibile:

?le!!9?-me-tè£ostrain-ma08/o?dati-a
?traildem—eln-vio!!memac-vio-lent?
0t5789dem''lolnv0o!!meeac-dio-capp
)p=?!himò!?n-op--ch-opo-;??min-l-i
?l?i!ii-ih—iopl---th-silghasqer-cc9))o

Seppure una di tali proposizioni dovesse fungere da messaggio ..., come nella favola del mago alla ricerca della formula magica, non c'è nulla che la può tradire ai nostri occhi per farla emergere dal pulviscolo delle diverse alternative per quanti tentativi si possano tentare.
Ciò detto, è anche chiaro che le cose sono diametralmente opposte qualora un messaggio abbia senso compiuto essendo che si porterà dietro uno stigma che lo rende riconoscibile.
Nel caso il rapporto tra lunghezza della chiave e lunghezza del messaggio tornerà a essere risolutivo nella logica dei sistemi perfetti,
giacché la maschera random dovrà essere talmente ampia da occupare uno spazio perfettamente sovrapponibile a quello del messaggio non random (e abbiamo visto come minime differenze lascino molto fianco scoperto),
così da bloccare l'individuazione di classi di minor ampiezza con le sorprendenti conseguenze che abbiamo potuto tastare.

Teorema

(61)

Si suppone in ipotesi che un attaccante dotato di risorse infinite e che conosca il valore del crittogramma, alla condizione di cui al teorema a seguire,
non possa risalire a un messaggio unico o almeno a un "numero sufficientemente ristretto di messaggi" sebbene la sommatoria di ogni singola probabilità di ciascuno di tali dispacci conduca a un risultato[48] per cui $\sum_{i=1}^{z} \mathrm{P}(message_{\,i}) = 1$

dove $i = 1, \ldots, \boldsymbol{z}$

dove $\boldsymbol{z}$ è un intero positivo che fissa la quantità dei messaggi di cui abbiamo detto e che appartengono a Ω e cioè alla classe di messaggi di minor ampiezza che si sia potuta selezionare,

dove $message_{\,i}$ sono tali messaggi[49] elementi di Ω, ciascuno dei quali detiene dalla prospettiva di chi attacca una uguale parte di P,

dove P è la probabilità di corrispondere al giusto messaggio,

dove per "numero sufficientemente ristretto di messaggi" intendiamo una quantità abbastanza piccola da consentire d'abbattere in modo proficuo le alternative appartenenti a Ω;

dove ciascuna di tali alternative dovrebbe pertanto avere una non trascurabile probabilità d'apparire con frequenza maggiore di 0.

Per cui, pur riportandoci alla definizione fissata in (47) più formalmente battezziamo col nome di "segretezza perfetta" lo stato adesso descritto.

(i)

E' dunque enunciato il teorema, secondo cui si avrà "segretezza perfetta" quando si rispetti il seguente requisito per il quale, dato un sistema crittografico dove in fase di codifica[50]

$\boldsymbol{x}$ di variabile X sia il valore del flusso di lunghezza l d'uno degli operandi (chiave o messaggio che sia) che potrebbe palesare disomogenee frequenze di probabilità risultando perciò debolmente impredicibile per un efficiente algoritmo,

$\boldsymbol{y}$ di variabile Y il valore del flusso di lunghezza L dell'altro operando (chiave o messaggio che sia) generato da sorgente aleatoria e quindi fortemente impredicibile per un efficiente algoritmo,

Critto il valore del crittogramma dato dalla composizione in XOR tra $\boldsymbol{x}$ ed $\boldsymbol{y}$,

sarà $L \geq l$

[48] Si ricordi che quando abbiam parlato di una quantità media di messaggi possibili che può ben esprimere una probabilità di comparsa inferiore ad uno, ci si riferiva al numero mediamente distribuito su più sotto-insiemi di ΩΩ, dove ciascun sotto-insieme è in accordo con un unico e solo crittogramma di tutti quelli possibili.

[49] **Dove notoriamente per messaggi non intendiamo solo quelli di senso compiuto.**

[50] Dove si suppone che la codifica come ormai quasi sempre nei sistemi perfetti, sia in somma XOR.

(ii)

Intanto si presume che un attaccante abbia selezionata una classe Ω di possibili messaggi la quale, in base alle informazioni in suo possesso, sarà la meno ampia possibile.

Assumendo allora **(1)** il caso favorevole a chi attacca, per cui lo spazio di tale classe si conformi ad *l* misura dell'operando di minor lunghezza[51]

avremo che se a fungere da messaggio fosse ***y*** di variabile *Y* che sappiamo essere random e di lunghezza *L* ugual-maggiore di *l*

per il diminuire dei dispacci effettivamente possibili che si conformano alla classe selezionata,

apparterranno a Ω solo ***z*** valori di ***y*** (dove $\boldsymbol{y} = \boldsymbol{y_1}\, \boldsymbol{y_2} \ldots \boldsymbol{y_Z}$) dei $\boldsymbol{Z}$ della più ampia classe ΩΩ correlata all'operando di maggior grandezza[52].

Per cui potremo anche dire,

$\Omega\Omega - \overline{\Omega} = \Omega$

Se però invece abbiamo che **(2)** il messaggio non sia random giacché dato da ***x*** di variabile *X* che sappiamo corrispondere a un file di lunghezza *l* più breve o al più uguale rispetto a *L*

sarebbero appartenenti a Ω altrettanti ***z*** possibili valori di ***x*** (dove $\boldsymbol{x} = \boldsymbol{x_1}\, \boldsymbol{x_2} \ldots \boldsymbol{x_z}$) degli ***z*** che già appartenevano.

Per cui potremmo anche dire sebbene in modo particolarmente informale, che Ω non è ridimensionata essendo Ω = Ω

Fissando allora una ulteriore classe **Ω'** data non più dai messaggi ma dalle chiavi compatibili col valore di *Critto* e con quelli degli elementi appartenenti a Ω a prescindere se essi rientrino nel caso illustrato in **(1)** o in quello illustrato in **(2)**,

esse a loro volta forniranno ***z*** alternative, essendo che per ognuno dei ***z*** messaggi possibili nei sistemi che abbiamo detti "perfetti" sempre ci sarà una ed una sola chiave che dia *Critto* come crittogramma[53].

Perciò detto, pertanto avremo #**Ω'** uguale #Ω che esprime una pari numerosità di ***z*** elementi.

Ragion per cui possiamo anche scrivere,

#Ω'= #Ω

dove dalla visuale di chi attacca tali #**Ω'** e #Ω saranno i soli presumibili spazi di chiave e messaggio su cui occorra operare.

[51] Come in precedenti casi, qui non prendiamo in esame come l'attaccante sia giunto a tale conoscenza, essendo questi un espediente logico per mettere a nudo alcuni aspetti che ci interessano ai fini della corrente dimostrazione.

[52] **La cui misura dalla visuale di chi attenti al sistema, sappiamo dipendere da quella del crittogramma.**

[53] Dove eventuali chiavi in sovrannumero a loro volta non darebbero *Critto* in uscita.

(iii)

Formulate allora tali premesse e dando per certa la circostanza secondo cui il crittogramma sarebbe stato intercettato e reso dunque noto,

avremo che a partire da un qualche momento in avanti, tale valore detto *Critto* sarà divenuto palese a chi attacca,

così che una volta svelato, non potrà che rimanere costante al variare dei valori ***x, y*** esprimendo uno scalare fisso e non più variabile[54].

Ora se diamo la funzione di codifica f: $(X, Y) \rightarrow Critto$ dove ***y*** di variabile Y è una grandezza soddisfatta da valori con probabilità uniforme di comparsa,

e dove *Critto* valore del crittogramma è ormai costante,

possiamo costruire l'insieme Ω'' di tutte le possibili coppie di valori X, Y della funzione $f(X, Y)$ che diano *Critto* come risultato.

Essendo dunque #Ω = #Ω' ma anche #Ω = #Ω' = #Ω''

sapendo che:

se in ipotesi **(a)** gli ***z*** elementi appartenenti a Ω fossero valori di X, gli altrettanti ***z*** elementi di Ω' lo sarebbero di Y

se in ipotesi **(b)** gli ***z*** elementi appartenenti a Ω fossero invece valori di Y, gli altrettanti ***z*** elementi di Ω' lo sarebbero di X [55]

procediamo come segue.

(iv)

Per semplificare faremo dunque una scelta dando una incidentale preferenza a quanto detto in **(b)** con l'avvertenza che nulla cambia se dovessimo prendere in considerazione l'ipotesi in **(a)**.

Per far dunque chiarezza stabiliamo che gli ***z*** elementi di $\boldsymbol{Z}$ indicati con Y effettivamente appartengano a Ω e cioè alla classe residua dei messaggi, mentre gli altrettanti ***z*** elementi indicati con X effettivamente appartengano a Ω' e cioè alla classe delle chiavi,

dove se così fosse, sapendo che ***z*** frasi delle $\boldsymbol{Z}$ che diciamo ***y*** di variabile Y che fungono da messaggio dal carattere aleatorio di lunghezza L, appartengono a Ω,

e sapendo che altrettante ***z*** distinte frasi ***x*** di variabile X che fungono da chiave non random di lunghezza l, appartengono a Ω'

esisterà sempre per ogni valore di $X \in \Omega'$,

[54] Naturalmente si sa che una funzione di codifica deve essere iniettiva; pur tuttavia dal momento e solo dal momento in cui, il valore di *Critto* diviene palese a chi attacca, essa potrà esser tradotta nella forma d'una costante.

[55] Dove sappiamo che i valori di X son sempre generati da una sorgente X-source non random, ed i valori di Y da una sorgente Y-source *true random.*

uno ed un solo equiprobabile valore di *Y* tra quelli appartenenti a Ω in grado di soddisfare la funzione dando *Critto* come risultato.

Per quanto detto, noto a chi attacca che il crittogramma intercettato vale *Critto*,

questi non potrà risalire ad $\boldsymbol{x}$ e nel contempo ad $\boldsymbol{y}$, per la equa probabilità di comparsa di ciascun valore binario di quest'ultima,

motivo per cui, la probabilità pari ad 1 data dalla sommatoria $\sum_{i=1}^{\mathbf{z}} \mathrm{P}(message\ _i)$ di tutti i messaggi necessari e incompatibili appartenenti alla classe residua Ω

sarà uniformemente ripartita dalla visuale di chi attacca, così che non abbia modo di discernere tra messaggio e messaggio.

Si può tuttavia aggiungere, che adottando la condizione in precedenza fissata, assumendo quanto detto in ipotesi,

avendo mostrato che solamente $\boldsymbol{z}$ elementi di *Y* apparterranno a Ω,

e altrettanti $\boldsymbol{z}$ elementi di *X* apparterranno a Ω**',**

abbiamo che solo e soltanto in riferimento a tali classi residue, si potrà pur dire che $\mathrm{P}(\boldsymbol{y} \mid Critto) = \mathrm{P}(\boldsymbol{y}) = \frac{1}{\mathbf{z}}$ così che nei suddetti limiti anche tale requisito sarà soddisfatto e la sicurezza è perfetta[56].

[56] Invero in (47) avevamo anticipato in modo spiccio ma efficace, che il requisito della non condizionalità sarebbe uscito dalla porta per rientrare tuttavia dalla finestra.

Bibliografia

BASSHAM L.E. e altri, *A Statistical Test Suite for Random and Pseudorandom Number Generator for Cryptographic, Applications, Washington D.C.*, NIST, National Institute of Standards and Technology, U.S. Department of Commerce, 2010.00

BORGES J.L., *La Biblioteca de Babel*, si trova in forma di racconto nel volume *Ficciones*, Buenos Aires, SUR, 1944.

GOLDREICH O., NISAN N. e WIGDERSON A., *On Yao's XOR Lemma*, si trova nella raccolta *Studies in Complexity and Cryptography, Miscellanea on the Interplay between Randomness and Computation*, pp 273-301 (LNTCS, volume 6650), SpringerLink, 1995.

GOLDWASSER S. e BELLARE M., *Lecture Notes on Cryptography*, Cambridge, Massachusettes Institute of Technology, insieme di note compilate in forma di dispensa e tratte dai corsi tenuti dagli autori presso il MIT nelle estati dal 1996 al 2001.

SHANNON C. E., *Communication Theory of Secrecy Systems*, Murray Hill, Bell System Technical Journal, 1949.

SHINDLER W., *Werner Schindler's Research Works,* Bonn, Bundesamt für Sicherheit in der Informationstechnik (BSI), 18 Settembre 2011.

YAO A.C., *Theory and Applications of Trapdoor Functions*, Chicago, 23rd Annual Symposium on Foundations of Computer Science, 1982.

Un saggio sulla incompletezza del teorema di Shannon sui sistemi crittografici a sicurezza perfetta, con alcuni cenni sui fondamenti teorici su cui basare, al volgere del terzo millennio, un nuovo e innovativo metodo di crittazione.

Claudio Cappelli,
collabora nel campo della massima sicurezza e della crittografia con imprese private. Ha partecipato ai lavori condotti per la lnternational Conference on DMS Distributed Multimedia System (San Francisco, USA, 2009, Florence, ltaly, 2011).

Luca Amodeo,
opera nel settore privato, in particolare presso società leader in Europa nella consulenza informatica. È interessato alla risoluzione di problemi logico-matematici e alla modellizzazione rivolta ad attività nel campo della cibernetica.

Lorenza De Lellis,
opera presso il Dipartimento di Farmacia dell'Università Federico Il di Napoli. È impegnata nel privato, in attività di ricerca sull'impiego delle strutture matematiche nei sistemi applicati al settore sanitario.

ISBN 978-88-90-95006-3