Ratio Math. n. 2
(1991), 29 - 56

IL PROBLEMA DELLA PROTEZIONE DELL'INFORMAZIONE II. LA GESTIONE E LO SVILUPPO DEI SISTEMI INFORMATIVI: SICUREZZA CONTRO COSTO ? (*)

Arrigo Bonisoli e Franco Eugeni

Sunto. Nell'attuale scenario dei sistemi informativi pubblici e privati le problematiche economico-gestionali che si presentano sono enormi e svariate. Ad alcune di queste può, se vuole, accostarsi con profitto il matematico. La trattazione dell'informazione può riguardarlo da vicino allora che si pensi ai seguenti tre fondamentali problemi: protezione crittografica del messaggio-informazione, certificazione/firma elettronica dello stesso, rivelazione/correzione di errori in un canale disturbato.

Nella prima parte dell'articolo illustriamo alcuni prodotti crittografici oggi disponibili, i loro limiti e le problematiche che nascono nel loro contesto. Nella seconda parte ci occupiamo da vicino di una classe di codici correttori interessanti soprattutto per le proprietà che ne facilitano l'implementazione.

1. L'ATTUALE SCENARIO DEI SISTEMI INFORMATIVI

In questi ultimi anni abbiamo vissuto e stiamo vivendo una profonda e continua evoluzione dei sistemi informativi. Ci si rivolge ad un gruppo di ingegneri del software, si definisce chiaramente tutto ciò che il sistema deve fare, quindi si reperisce nell'ampio e ricco mercato quanto occorre. Le attuali tendenze nelle medie e piccole aziende sono tutte proiettate verso il miglioramento dell'integrazione uomo-computer e della gestione attraverso sistemi/strumenti computerizzati con utilizzo di chip sempre più potenti. La favola del

(*) lavoro svolto nell'ambito del GNSAGA del CNR con il contributo del MURST

computer delle vecchie generazioni miniaturizzato in un piccolo personal da tavolo è di fatto una realtà. Tra i vari strumenti sono da ricordare i numerosi pacchetti applicativi per specialisti (ingegneri, architetti, medici) o per la gestione (programmi per la contabilità, l'editing, i fogli elettronici di gestione finanziaria). Così per la documentazione, l'archiviazione e la consultazione si è in grado di fare ricerche velocissime e operazioni "taglia e cuci" a livelli veramente alti. E' inoltre possibile lo scambio in tempi reali di grosse moli di documenti tra sedi diversamente ubicate.

Un altro aspetto che ci interessa evidenziare è quello dell'avvento dei "sistemi esperti". Un tale sistema può essere definito come un programma in cui sono condensate le conoscenze di un enorme numero di esperti in un particolare argomento o settore e che in più ha capacità di fare uso di procedure di inferenza. In altre parole un sistema esperto oltre ad immagazzinare notizie che gli arrivano da tutto il mondo le elabora utilizzando "l'intelligenza" della macchina ad integrazione dell'esperienza umana. Un tale sistema ha capacità di scelte e decisioni nel senso che segue. Un sistema esperto—strutturato in un magazzino di conoscenze più una capacità elaborativa di tipo inferenziale—acquisendo informazioni su un problema e utilizzando la propria banca dati è in grado di pervenire a risposte finali, cioè decide e sceglie; inoltre è in grado di fornire all'utente tutti i dettagli sia sui dati utilizzati che sulle procedure applicate. Sistemi esperti sono stati progettati per dedurre situazioni da dati provenienti da sensori, prevedere conseguenze da una serie di azioni, formulare diagnosi a partire da dati rilevati, configurare oggetti dal punto di vista del design tenuto conto di vincoli e richieste opportune e addirittura per gestire risorse finanziarie.

Nella gestione delle aziende moderne, quasi come in un organismo militare, sono ancora di importanza crescente i modelli strategici per la produzione, i piani di sviluppo e della globale gestione delle risorse. Certo è ancora costume imperante per molti quadri d'impresa operare scelte e strategie "a caldo" basate su intuizioni, su sensibilità di mercato ed esperienza. Tuttavia è sempre più importante, a causa della sempre crescente internazionalizzazione del mercato, per la competizione derivante dalla tendenza in atto ad una maggiore omogeneizzazione dei consumi nei più diversi paesi, operare in modo più consapevole. Nasce così in questo mondo economico sempre in fermento il marketing strategico, basato su analisi rigorose e complete. Alla base di queste analisi vi è la banca dati della nostra azienda, le accurate statistiche per le analisi di mercato, l'utilizzo dei sistemi esperti per la gestione delle informazioni che via via vengono immagazzinate per l'ottimizzazione della pianificazione e delle funzioni di produzione. Così la strategia che la grande azienda sviluppa ha alla base un patrimonio divenuto sempre più importante: l'analisi, la gestione, la manipolazione, la trasmissione delle informazioni che possiede.

Nell'attuale scenario vediamo dunque l'informazione al centro dei sistemi economico-gestionali viaggiare da un angolo all'altro del mondo a velocità

incredibili ed ai più diversi livelli. L'informazione si muove all'interno di una direzione aziendale o entro un gruppo di progettazione, da un edificio ad un altro, tra città, regioni, stati; inoltre essa viaggia ed interconnette vari campi dello scibile: per l'informazione non vi è il senso del confine, della distanza, del tempo. I sistemi di posta elettronica sono sempre più efficienti ed affidabili: ciascun utente immette la sua "password" di identificazione e invia i propri messaggi, i quali, codificati in modo più o meno robusto, viaggiano lungo i cavi del telefono e raggiungono rapidamente un qualunque nodo della rete.

Anche la moneta è oggi considerata o almeno trattata alla stregua di un'informazione. Si stanno effettuando studi sempre più numerosi sul problema delle transazioni di tipo economico per via elettronica. Naturalmente prima che la moneta elettronica possa avere l'intero diritto di cittadinanza occorre garantire l'utente con alcuni standard di sicurezza e con una legislazione adeguata. Si pensi ad un messaggio adeguatamente protetto dalle frodi e/o manipolazioni, autenticato, anzi, firmato elettronicamente (chi lo riceve è in grado di provare ad un terzo l'identità del mittente). Un tale messaggio potrebbe essere una transazione finanziaria fatta come davanti ad un notaio, ma in tempi di secondi.

La nuova moneta elettronica è una carta dello stesso aspetto del classico Bancomat ma contenente al suo interno un microprocessore ed una memoria da 8 kilobyte (8.000 caratteri per intenderci). La carta ha capacità elaborative e non può essere contraffatta. Ben nota in Francia (carte a memoire) e nei paesi anglosassoni (smart card), recentemente è stata introdotta anche in Italia dalla Eni Data, come libretto Universitario alle Università di Roma e Bologna. Per maggiori dettagli si veda [**9**]. Si pensi all'uso di punti di vendita elettronici, i cosiddetti POS (Point Of Sale), con i quali ogni transazione finanziaria viene registrata. Controllo su tutto, dalle tangenti ai ricatti. Potrebbe essere la fine del denaro sporco! Ma certo inventerebbero qualche cosa di nuovo.

Un obiettivo fondamentale nei moderni canali di comunicazione è il trasferimento di informazioni con elevato livello di sicurezza sia ai fini della loro corretta ricezione sia ai fini delle intercettazioni non autorizzate. Si pensi ai canali utilizzati dai satelliti artificiali, dalla posta elettronica, dai sistemi di telecontrollo e di trasferimento elettronico di fondi.

L'impiego di tecniche crittografiche è uno dei metodi più pratici ed economici per la protezione dei dati archiviati e trasmessi. In generale nella progettazione dei sistemi informatici è riconosciuta importanza notevole al problema della sicurezza dei dati e della loro protezione. Deve essere naturalmente essere tenuto in conto che la trasmissione, memorizzazione e l'utilizzo di un sistema crittografico non deve presentare grandi difficoltà ed occorre una attenta analisi costi-benefici per la scelta di algoritmi di protezione, di controllo, di rivelazione di errori o di correzione degli stessi.

Spesso il fallimento di un sistema dipende da cattiva gestione da parte degli utenti, ad esempio uso di una stessa chiave per tempi più lunghi di quanto

raccomandato dai progettisti, riutilizzo di vecchie chiavi, trasmissione di testi chiari dei quali vi sia in giro una versione crittografata.

Naturalmente ogni azienda dovrà classificare il proprio materiale da proteggere, dovrà valutare i rischi connessi con la distruzione dei dati, la loro modificazione e l'eventuale divulgazione. Inoltre dovrà anche ottimizzare certe situazioni, ad esempio minimizzando il numero di persone che negli anni possano acquisire la conoscenza completa di materiale riservato.

Non possiamo concludere questo paragrafo, illustrante appunto lo scenario nel quale si elabora l'informazione, senza parlare di un antico problema di garanzia per un qualsiasi prodotto commerciale, cioè della creazione di standard. Le esigenze che sorgono con la necessità di proteggere dati trattati elettronicamente conducono anche alla loro standardizzazione, almeno per quei sistemi per i quali un hardware oppure un software è di dominio pubblico. Per standard intendiamo quei prodotti che rispecchiano alcune norme tecniche emanate da enti nazionali od internazionali, o anche da privati che abbiano il potere politico o commerciale di farlo. Uno standard offre naturalmente sicurezza perché è sottoposto a verifiche e certificazioni ed anche economicità dividendosi tra molti i costi generali della sua produzione. L'ente internazionale che si occupa di questo problema è l'INTERNATIONAL ORGANIZATION OF STANDARDIZATION (ISO) che raggruppa più di 75 Nazioni. In Italia abbiamo l'ENTE NAZIONALE ITALIANO DI UNIFICAZIONE (UNI) che rappresenta l'Italia presso l'ISO. In ambito UNI l'ente che si occupa del settore informatico è l'UNIPREA presso il quale sono stati proposti vari standard per la protezione dell'informazione.

I due Paragrafi che seguono saranno dedicati allo studio di alcuni codici crittografici in uso oggi, in particolare i ben noti sistemi DES, che è uno standard della IBM, e il codice RSA basato sulla cattiva conoscenza che abbiamo dei numeri primi; quest'ultimo non è uno standard, ma è al centro di vaste ed interessanti polemiche interagenti perfino con la teoria dei numeri.

Lo schema generale della trasmissione di un'informazione sarà per noi quello dato dal seguente schema, del quale tratteremo i vari punti separatamente:

— Il trasmettitore prepara il **messaggio in chiaro,** lo crittografa trasformandolo in **testo cifrato** (eventualmente coincidente con il chiaro) ed appronta un secondo messaggio, che accompagnerà il primo: **l'autenticatore** (eventualmente vuoto).
— La coppia testo cifrato-autenticatore viene inviata nel canale.
— All'arrivo i due messaggi sono sottoposti ad un procedimento di correzione automatica degli errori.
— Utilizzando l'autenticatore ci si rende conto dell'identità del mittente, se l'autenticatore è una firma numerica si appronta una prova d'identità per un terzo.
— Infine si decifra il messaggio utilizzando il sistema crittografico pubblico o privato che sia.

2. CODICI CRITTOGRAFICI SEGRETI ANTICHI E MODERNI

Il significato della dicitura "protezione dell'informazione" è cambiato col passare del tempo. Infatti un tempo proteggere l'informazione significava soltanto impedire ad un non autorizzato la lettura di un messaggio, oggi vuol dire anche impedirne la manipolazione.

Per renderci conto di questa necessità è sufficiente pensare ad una transazione finanziaria che avviene a distanza, oppure ad un telecomando di un impianto strategico. La protezione dell'informazione oggi si attua con metodi crittografici.

Iniziamo questo paragrafo con uno sguardo al passato. Bisogna premettere che nella storia la parola codice è stata veramente usata con molti e diversi significati. Uno di questi è appunto il significato di codice segreto per inviare messaggi segreti.

Sul finire del Medio Evo, con l'inizio delle relazioni diplomatiche tra i vari stati, i codici segreti diventano una necessità. Sembra che l'uso sistematico dei codici segreti abbia inizio nella corte papale, nelle repubbliche e nelle signorie italiane a partire dal 1300. Si diffonde l'uso dei nomenclatori cioè di vocabolari con parole cammuffate e simboli nuovi rispetto alle lettere: è un codice efficiente ma abbastanza primitivo. La difficoltà principale è il trasporto del nomenclatore e nasce l'idea di crittografare con regole e chiavi.

In realtà un codice segreto di questo tipo era di fatto passato alla storia: il codice di Cesare. Scriviamo due alfabeti su due righe successive in modo che le lettere del secondo alfabeto siano slittate di un certo numero di posti rispetto a quelle del primo. Chiamiamo il primo alfabeto **alfabeto in chiaro** ed il secondo **alfabeto cifrante.** Il metodo di cifratura è ovvio, così come la sua rottura. La chiave di un codice di Cesare è la permutazione circolare usata. Dunque le chiavi possibili sono 26. E sufficiente che una persona ipotizzi che un messaggio in italiano sia stato cifrato (o **crittato**) con il codice di Cesare e, pur non conoscendo la chiave segreta usata, riesce a **forzare** il codice. La crittoanalisi di un codice di Cesare può essere anche basata sullo studio della frequenza delle lettere. Infatti in ogni lingua ogni lettera compare con una sua frequenza propria. Segue che, se il messaggio è abbastanza lungo, la lettera che in esso compare con una maggiore (minore) frequenza probabilmente corrisponde alla lettera che nella lingua usata ha una maggiore (minore) frequenza.

Le cose si complicano notevolmente usando in luogo di una permutazione circolare, una permutazione qualsiasi. In tal caso le chiavi possibili sono $26!$. Un codice segreto in cui si usi una qualunque permutazione dell'alfabeto si chiama codice a sostituzione.

Esiste nella pratica un altro tipo di permutazione usata nei codici, la trasposizione. Si tratta di una permutazione a blocchi. Per costruire un tale codice si fissa un intero n ed una permutazione P dei primi n interi. Si suddivide

il testo in blocchi di lunghezza n ed in ognuno di essi si opera con la permutazione P. Il numero delle sue chiavi è maggiore di quello di un codice a sostituzione, quindi è più difficile trovare la chiave.

Il codice di Cesare è, come suole dirsi, monoalfabetico. E' sulla fine del medio evo che la **crittografia** ha una grossa evoluzione. Troviamo un cambiamento radicale, infatti nascono i primi codici segreti che non usano un solo alfabeto cifrante: i **codici polialfabetici.** Non abbiamo lo spazio per illustrare in dettaglio i primi due magnifici codici del tempo: quello costruito da Leon Battista Alberti su commissione di un segretario pontificio intorno al 1466 e quello di Giambattista Della Porta (1563), l'inventore della camera oscura, al quale si deve l'introduzione della **parola chiave.** Troviamo nel tempo vari altri codici polialfabetici come quelli di G. Cardano e di Bellaso. Per la descrizione dei codici suddetti rimandiamo ad un articolo di E. Ambrisi e F. Eugeni [**1**] e ci occupiamo ora di un codice importante ancora oggi.

Si tratta del codice costruito dal francese Blaise di Vigénère. Egli nel 1586, riciclando e ripulendo da marchiani errori un'idea dell'abate Tritemius, fa uso di una tavola quadrata riportata sotto, nota appunto come **tavola di Vigénère.** La tavola in sostanza è la tabella additiva del gruppo $\mathbf{Z}_{26}$ e la codifica e decodifica sono legate all'addizione modulo 26.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Esempio di codifica:

parola chiave : **wernhei swern heiswer nh eiswernhei ...**
testo in chiaro : **accordi italo sloveni in trattative ...**
testo cifrato : **wgtbyhp ...**

Il segreto di questo codice è tutto nella parola chiave. Il destinatario del messaggio conosce la parola chiave. Allora è in grado di decifrare il messaggio in arrivo usando il procedimento al contrario (ovvero la sottrazione modulo 26).

Il codice di Vigénère è stato usato da eserciti e stati in lungo e in largo e ha resistito a tutti gli "attacchi" dei crittoanalisti per circa tre secoli. Fu decrittato da un ufficiale prussiano di nome Kasiski (1863) con un procedimento noto come **test di Kasiski.** Esso si può applicare se il testo cifrato è abbastanza lungo e la parola chiave abbastanza corta. L'idea di base è quella di scoprire la lunghezza ℓ della parola chiave. Si considerano poi la prima lettera, la $(\ell + 1)$-esima lettera, la $(2\ell + 1)$-esima lettera, e così via, ottenendo un messaggio parziale che, essendo codificato con la stesso alfabeto, cioè con la stessa riga della tavola di Vigénère, ovvero con un codice di Cesare, si può decrittare facilmente.

E' possibile "complicare" il codice di Vigénère per avere un codice sicuro? Un codice completamente sicuro è il **codice di Vernam** (1926). Questo è un codice del tipo di Vigénère nel quale si usa una parola chiave avente una lunghezza pari alla lunghezza del messaggio. Oggi in tempi di calcolatori si usa di fatto un metodo noto come "one-time-pad". Si prendono un messaggio binario ed una sequenza binaria completamente casuale: codifica e decodifica avvengono sommando modulo 2 le singole componenti. Siamo nel pieno della Teoria di Shannon [**46**].

Esempio di codifica:

parola chiave:	11001101011100101011101010111000010
testo in chiaro:	00111001010111001010110101010101010
testo cifrato:	11110100001011100001011111101101000

Una sequenza casuale molto lunga può essere memorizzata ma non si trasmette molto bene. Spesso si usano sequenze pseudo-casuali cioè costruibili con una legge prefissata, cioè programmabili. Inoltre si possono scambiare tra gli utenti tabelle di chiavi o "vermi" lunghissimi e segnalare di volta in volta con qualche criterio la chiave usata o il punto del verme. .

Nel 1949 Shaunon [**46**] ha provato che se la parola chiave o verme è lungo come il testo ed è completamente casuale allora il sistema è indecifrabile. Questo spiega la sempre crescente ricerca di sequenze pseudocasuali, intendendo con questa locuzione sequenze generate al computer con qualche

sistema, quindi non casuali, ma soddisfacenti ad una serie di test statistici che le avvicinino abbastanza alla teorica casualità, il cui unico torto è appunto la non riproducibilità.

Sequenze pseudocasuali $(a_1, a_2, ..., a_n)$ si costruiscono usando ricorrenze del tipo:

$$a_{n+1} = f(a_1, a_2, ..., a_n)$$

presupponendo di solito che la legge f sia lineare. Una tecnica di questo tipo sta diventando tuttavia sempre meno sicura, così che vi è una certa tendenza a fare uso di ricorrenze non lineari, come ad esempio nella modernissima tecnica DES.

Ad un testo in chiaro si possono applicare successivamente svariate sostituzioni e trasposizioni. Naturalmente non a caso perché, come ben noto, non è detto che componendo due codici si abbia come risultato un codice più robusto. Un codice siffatto si dice appunto codice composto. I codici composti sono i tipici codici dei nostri giorni. Il DES (Data Encryption Standard) è un codice che è stato pubblicato dalla IBM nel 1977. Il DES è usato principalmente in campo bancario. Esso opera su blocchi di 64 bit. Ognuno di essi è suddiviso in 8 blocchi di 8 bit ciascuno (byte) in cui l'ottavo simbolo è un carattere di controllo. Dunque in ogni blocco ci sono 56 simboli significativi. Tale codice opera alternando trasposizioni e sostituzioni per 16 volte, ed ha 2^{56} chiavi possibili. L'algoritmo DES è stato pubblicato: ne segue che tutta la sicurezza/segretezza del DES risiede nella chiave, scelta su un totale di 2^{56} chiavi. Veramente tante!

Il DES è inoltre uno standard di mercato, esso quindi possiede i vantaggi di economicità e di compatibilità di un codice universale. Lo svantaggio di un metodo standard sta nel fatto che tutti i crittoanalisti stanno cercando di forzare il DES e in caso di riuscita essi potrebbero venire a conoscenza di segreti importanti. In tal caso ogni utente del DES dovrebbe cambiare l'intero sistema protettivo con enormi spese.

Ciò nonostante il DES è oggi uno dei codici più usati. Ad esempio nelle carte a banda magnetica (tipo Bancomat) si usa il DES. Una carta Bancomat ha una striscia (banda magnetica) in cui si possono memorizzare alcuni dati (pochi). Quando una banca consegna ad un suo cliente una carta Bancomat, comunica al cliente anche un numero segreto: il PIN (Personal Identification Number), composto da 5 cifre in numerazione decimale. La banca associa al cliente una stringa N formata dal numero di conto corrente del cliente assieme ad un proprio numero di codice. Il numero N è scritto sulla carta. La banca critta il numero N con il DES, cioè scrive N in forma binaria e poi applica a questa sequenza il DES, usando una chiave segreta K; l'algoritmo produce un numero a cinque cifre, il PIN$= f(N, K)$. Quando il cliente introduce la carta nel terminale, nel breve tempo di "attendere prego, operazioni in corso ..." il terminale

legge N sulla carta identificando sia il cliente che il codice della banca che ha emesso la carta. Dal codice della banca è facile per il terminale desumere da un archivio in suo possesso la chiave K usata dalla banca, quindi applicando il DES ad N e a K è in grado di ricalcolare il PIN della carta e controllare se coincide con quello che deve digitare l'utente. Se una persona volesse scoprire il PIN di una carta, dovrebbe fare mediamente 50.000 tentativi, cioè 200 ore di sportello. I punti deboli sono il codice segreto, che non si dovrebbe mostrare a nessuno, e l'archivio on-line al quale qualcuno potrebbe riuscire ad accedere. Protezioni del Bancomat: cambiare spesso i numeri e limitare la somma da ritirare giornalmente.

3. CODICI PER L'AUTENTICA DEL MITTENTE E CODICI A CHIAVE PUBBLICA

In genere sono molti i significati che acquista la parola "codice". In questo paragrafo prescinderemo dalla realtà del soggetto che costituisce l'informazione, essa può essere musicale, scritta, parlata, gestuale, mentre il codice è qualcosa di più concreto: può ad esempio essere costituito da simboli su carta, da impulsi elettrici traducibili in voce al telefono o scritti da macchine stampanti, oppure da segnali di vario tipo quali i Morse, le bandierine, e così via. Assumiamo che Mr. **B**onisoli trasmetta dati a Mr. Eugeni. Mr. **E** vuole essere certo che i dati ricevuti siano proprio quelli spediti da Mr. **B** e non quelli che proditoriamente o per fare uno scherzo ha voluto spedire un certo Mr. **H** (probabilmente il Prof. **H**eise). Si tratta allora di usare un codice che garantisca il messaggio. Un codice di questo tipo si può ottenere accludendo al messaggio un certo numero di simboli ridondanti, costruiti come funzione dei simboli del messaggio. Tale funzione deve essere naturalmente un segreto, anzi il segreto del sistema. E' dunque importante che il malintenzionato o scherzoso Mr. **H** non possa alterare il messaggio in suo favore. Ad impedire che ciò accada i nostri Mr. **B** e Mr. **E** si accordano su una chiave segreta C ed un algoritmo F. Ora Mr. **B** spedisce un messaggio m assieme ad un **autenticatore** che altro non è che il messaggio ridondante e crittografato $a = F(m, C)$. Cosa succede a Mr. **E**? Egli riceve un messaggio costituito dalla coppia (m', a'), che può essere la modifica di (m, a) escogitata da Mr. **H**. Il nostro Mr. **E** forma l'espressione $a^* = F(m', C)$; solo se a^* coincide con a' egli considera autentico il messaggio. Cosa può fare Mr. **H**? Non conoscendo F e C egli può solo fare dei tentativi per rompere il codice. Le sue possibilità di indovinare la coppia giusta sono espresse dal seguente teorema di Gilbert-MacWilliams-Sloane **[28]**:

TEOREMA. *Supponiamo equiprobabili i messaggi e le chiavi possibili. Allora in ogni codice di autenticazione la probabilità di vittoria di un estraneo è almeno pari al reciproco della radice quadrata del numero totale delle chiavi.*

Uno schema di autenticazione nel quale la probabilità di successo dell'estraneo sia uguale al valore minimo è detto **perfetto.** Un codice in cui la probabilità suddetta è dello stesso ordine di grandezza del valore minimo si dice **praticamente perfetto.**

Vediamo un esempio di codice perfetto di autenticazione dovuto a Gilbert, MacWilliams e Sloane [**28**]. Fissiamo il piano proiettivo finito $PG(2,q)$ di ordine q sul campo $GF(q)$. Fissiamo una retta L. Definiamo come **messaggi** i $q+1$ punti della retta L, le **chiavi** sono i q^2 punti esterni alla retta L, inoltre **l'autenticatore** di un dato messaggio m in una fissata chiave è la retta a che li congiunge. Segue che un decrittatore che dall'esterno veda passare la coppia (m, a) e la voglia alterare in (m^*, a^*) può solo tentare di scegliere un punto di $A \setminus L$ sperando che sia proprio la chiave. Ciò può essere fatto allora in esattamente q modi. Essendo q la radice del numero delle chiavi segue che lo schema è perfetto.

Vi sono alcuni esempi per lo più dovuti a Beutelspacher [**15**], [**16**] e anche a Tallini e Zanella [**20**]. Recentemente un semplicissimo schema perfetto basato sui disegni divisibili è stato costruito da Berardi ed Eugeni [**9**].

Gli schemi di autenticazione più interessanti sono quelli derivanti dai **crittosistemi a chiave pubblica,** dei quali vogliamo ora occuparci.

L'idea nuova del codice a chiave pubblica è che la chiave non è segretamente inviata ma è nota, come su un elenco del telefono. Alla base di questi codici vi sono le **funzioni unidirezionali.** Si tratta di funzioni calcolabili in tempi rapidi, teoricamente invertibili ma con funzione inversa calcolabile in tempi tanto lunghi da risultare proibitivi. Sia C una funzione unidirezionale facilmente calcolabile e sia P la sua inversa, calcolabile in tempi lunghissimi. Mr. **B** e Mr. **E** fanno ora parte di un sistema di scambi gestionali e in un elenco sono riportate le rispettive funzioni C_B e C_E, note a tutti. Ognuno di essi ha un segreto: l'inversa P_B è il segreto di Mr. **B**. Quando Mr. **B** vuole mandare il messaggio m a Mr. **E** gli manda non m ma $C_E(m)$. Può farlo perché legge C_E nell'elenco dei codici che è pubblico. Ma solo Mr. **E** conosce la funzione P_E, e solo lui può decifrare.

Una funzione unidirezionale può essere usata anche **come firma numerica non riutilizzabile** nel modo seguente. Il nostro Mr. **B** concorda una frase breve con un certo numero di persone: dichiara ad esempio che firmerà con la parola $A = ARRIGO$. Invia dunque il messaggio m crittografato o no ad un secondo personaggio assieme alla frase $P_B(A)$. Egli solo conosce la funzione P_B, difficile a chiunque da trovare, ed egli solo è in grado di costruire $P_B(A)$. Ma chiunque abbia accesso all'elenco conosce C_B e se questi conosce anche A è in grado di calcolare $C_B[P_B(A)]$ e vedere se coincide con la parola A prescelta, la firma numerica non riutilizzabile di Mr. **B**, che per un nuovo messaggio cambierà parola. Esempi di funzioni unidirezionali sono il **logaritmo discreto e la funzione prodotto di numeri primi,** naturalmente per valori elevati dei numeri che consideriamo. La funzione prodotto di primi viene utilizzata nel codice RSA, costruito da Rivest, Shamir ed Adelman nel 1978, utilizzato per scopi

bancari, essenzialmente come firma o per messaggi corti. E' un sistema molto costoso.

Per descrivere tale codice si consideri un sistema di utenti (organizzazioni civili o private) che partecipino al codice. Ogni utente sceglie due numeri N ed H che pubblica su un elenco ed un terzo numero D che rimane segreto. I numeri sono costruiti con i seguenti criteri. Il numero N è il prodotto di due primi p e q. Allora l'utente sa calcolare il numero di Eulero $E = (p - 1)(q - 1)$ e può scegliere un numero D primo con E. Ora può prendere un intero H tale che $(DH - 1)/E$ sia un intero e renderlo pubblico. Un messaggio ora è una sequenza di numeri. Sia m il generico numero della sequenza e vediamo come si critta.

Se vogliamo mandare un messaggio m ad un utente, ci procuriamo i suoi numeri N ed H e costruiamo il crittogramma c di m, ponendo:

$$c = \min \{X \mid (X - m^H) / N \text{ è un intero}\};$$

l'utente calcola

$$m' = \min \{Y \mid (Y - c^D)/ N \text{ è un intero}\}$$

ed è $m' = m$.

Le cifre che si usano attualmente per i primi che formano N sono dell'ordine delle duecento cifre. Recentemente è stata pubblicata a riguardo una notizia che è sembrata un grido d'allarme per il codice RSA. Precisamente si è saputo che un gruppo di matematici guidati da Mark Manasse del Centro di ricerca della Digital Equipment Corporation di Palo Alto e Arjen Lenstra della Chicago University facendo lavorare 400 calcolatori localizzati negli Stati Uniti, in Australia e in Olanda per circa un mese ed in contemporanea hanno scomposto un numero di cento cifre. Esso è risultato il prodotto di due primi rispettivamente di 41 e 60 cifre. Questo risultato sembrerebbe intaccare la sicurezza del codice RSA. Invero, sempre a proposito di questa notizia, non molto preoccupato si è dichiarato il Prof. Silvio Micali del MIT, uno dei big della crittografia internazionale. Avrebbe rilasciato la seguente significativa dichiarazione: "Un pericolo reale potrà venire solo da una nuova idea matematica e non dalla forza bruta di molti computer". D'altro canto il precedente primato mondiale di decomposizione di un numero era stato la decomposizione di un numero di 93 cifre decimali. Tale primato stabilito nel 1988 da Riele, Lioen e Winter richiese 95 ore di calcolo su di un supercomputer. Pertanto l'aumento di complessità tra 93 a cento cifre osservato su un caso da Manasse e Lenstra ci sembra veramente notevole! Tuttavia una nuova generazione di computer superveloci potrebbe spostare il problema molto avanti: quindi in guardia con la tecnologia!

4. STRATEGIE NEL MONDO REALE

Consideriamo un qualsiasi ambiente EDP (Electronic Data Processing). I problemi concreti che si pongono sono molti: impedire il contatto da parte di estranei con il materiale da proteggere, impedire di copiare ed elaborare in qualsiasi modo, ma anche il non permettere ad estranei di vedere o poter fare congetture plausibili. Naturalmente tutto questo era più facile da fare nel mondo militare e lo è un po' meno nel mondo civile. Ovviamente alla base di ogni progetto crittografico vi deve essere una protezione fisica del sistema che si presuppone sempre progettata ed organizzata a monte. E' impensabile ragionare sulla complessità di un sistema crittografico se poi le chiavi si mettono a disposizione di chiunque, oppure proteggere da falsificazioni senza avere da qualche parte una copia dell'originale, e questi sono soltanto esempi banali.

Un ulteriore aspetto da considerare è quello tecnologico. I nostri messaggi possono e devono essere pensati come oggetti di natura elettromagnetica, quindi altamente miniaturizzati e capaci di viaggiare a velocità altissime, con trasmissioni in tempo reale. Per comprendere la natura e la mole fisica di questi messaggi pensiamo che 10 cartelle dattiloscritte, per un totale di 5.000 caratteri circa, sono registrate su un tratto di nastro magnetico di circa 6 centimetri e che bastano 30-40 secondi di registrazione da un normale telefono per ottenerne una copia. L'attacco ad un centro di archiviazione ed elaborazione dati è oggi dunque più facile di ieri. Si pensi all'estensione che può avere una rete di terminali e ai molteplici punti in cui l'ascolto è possibile. Si pensi alla possibilità per un ascoltatore di captare brani di messaggi di cui conosce il corrispondente testo in chiaro, alla possibilità di ottenere statistiche, correlazioni, informazioni di ogni genere. Un'attenta raccolta e le giuste elaborazioni, come l'esperienza prova, possono dare risultati insperati.

Un ostacolo fino ad oggi è stata di sicuro l'incompatibilità di alcuni sistemi, la continua modifica degli hardware, la mancanza di standard, non ultima la buona retribuzione del poco personale altamente specializzato. Ma ragioni legali di certificazione dei sistemi di sicurezza impongono l'uso degli standard. La tecnologia attuale sembra tenda a stabilizzarsi (a parte i futuribili laser-computer che potrebbero dar luogo a nuove rivoluzioni per la loro enorme velocità), il numero dei competenti tende a crescere, insomma tutto fa presagire una crescente necessità di dure protezioni crittografiche.

Da quanto ci risulta, per la protezione di messaggi trasmessi in linea in entrambi i versi sono in uso macchine-hardware autonome, diciamo collegate con i modem e capaci di crittare e decrittare, usualmente con chiave segreta. Altri sistemi proteggono archivi e funzionano con chiavi in possesso di personaggi autorizzati, quali ad esempio programmi applicativi ovvero sistemi operativi.

5. LA CODIFICAZIONE A CORREZIONE NELLA TRASMISSIONE DEI MESSAGGI

La teoria dei codici correttori studia il problema della spedizione di messaggi lungo una linea disturbata, facendo in modo che il messaggio ricevuto coincida per quanto più possibile con quello originario, [2], [32], [33], [35], [36], [45]. Lo scopo è in fondo proprio l'opposto di quello della crittografia. Si pensi al telegrafo, al telefono, alla comunicazione con i satelliti e così via.

In una descrizione unitaria di questi fenomeni si è soliti rappresentare i messaggi da trasmettere come sequenze di simboli provenienti da un **alfabeto** F di cardinalità finita, tipicamente l'alfabeto binario $\{0, 1\}$. La linea di trasmissione prende il nome di **canale** ed è rappresentata da una matrice stocastica detta appunto **matrice di canale**; un elemento di questa matrice indica la probabilità condizionata $p(\beta/\alpha)$ che un certo simbolo β dell'alfabeto esca dal canale dopo che in ingresso era stato spedito un fissato simbolo α. L'ingresso nel canale di un simbolo α darà luogo a un **errore di trasmissione** (cioè all'emissione di un simbolo diverso dal simbolo originario) con probabilità $\sum_{\beta \in F\setminus\{\alpha\}} p(\beta|\alpha)$ che in generale è diversa da zero.

Per diminuire la probabilità di errori di trasmissione si ricorre alla **codificazione** che nella maggior parte dei casi può essere schematizzata come segue. I simboli in ingresso nel canale vengono suddivisi in blocchi, ciascuno costituito da k simboli. A ciascun blocco $x_1x_2 \dots x_k$ di **simboli d'informazione** vengono aggiunti r **simboli di controllo** $y_1y_2 \dots y_r$, mediante una opportuna regola; la sequenza o **parola** $\mathbf{c} = x_1x_2 \dots x_ky_1y_2 \dots y_r$ deve appartenere a un sottoinsieme $\mathbf{C}$ di F^{k+r} che prende il nome di **codice.** Il meccanismo che realizza questo procedimento prende il nome di **codificatore.** Il codificatore spedisce i simboli d'informazione e i simboli di controllo lungo il canale il quale emette una parola $\mathbf{c'} = x'_1x'_2 \dots x'_ky'_1y'_2 \dots y'_r$ che in generale non coinciderà con la parola $\mathbf{c}$ originariamente spedita. Scopo del **decodificatore** è di riuscire a ricostruire la parola originaria $\mathbf{c}$ a partire dalla parola ricevuta $\mathbf{c'}$. In generale il decodificatore mediante una opportuna **regola di decodificazione** determinerà a partire da $\mathbf{c'}$ una parola $\mathbf{c''} = x''_1x''_2 \dots x''_ky''_1y''_2 \dots y''_r$ appartenente al codice $\mathbf{C}$; emetterà poi la sequenza $x''_1x''_2 \dots x''_k$ come stima del blocco d'informazione originario. Di solito il decodificatore nella sua ricerca si basa su un criterio di **massima verosimiglianza,** cioè vuole che la parola $\mathbf{c''}$ abbia la più alta probabilità "a posteriori" di essere stata immessa nel canale dopo che in uscita è stata osservata la parola $\mathbf{c'}$. Per i canali più comunemente studiati — i cosiddetti canali "simmetrici" — ciò si attua facendo in modo che la parola di codice $\mathbf{c''}$ differisca dalla parola ricevuta $\mathbf{c'}$ nel minor numero possibile di componenti; se chiamiamo **distanza di Hamming** di due parole $\mathbf{a}$ e $\mathbf{b}$ il numero $\rho(\mathbf{a}, \mathbf{b})$ delle posizioni in cui esse hanno componenti diverse, allora possiamo dire che il decodificatore cerca una parola $\mathbf{c''} \in \mathbf{C}$ per la quale risulti minima la distanza di Hamming $\rho(\mathbf{c'}, \mathbf{c''})$.

E' abbastanza chiaro a questo punto quali siano i parametri che misurano le qualità del nostro schema di codificazione- decodificazione. Se indichiamo con $n := k+r$ la **lunghezza del codice C**, il numero $R(\mathbf{C}) := k/n$ *si* chiama **tasso d'informazione** del codice e misura il contenuto d'informazione utile che si trova mediamente in una parola di codice. Al contrario il rapporto r/n rappresenta la **ridondanza** del codice **C**, ossia l'informazione aggiuntiva che in un certo senso siamo stati costretti a sprecare con i simboli di controllo per migliorare la sicurezza della trasmissione. All'aumentare della ridondanza aumenta la necessità di memorizzare dati, diminuisce pertanto la velocità della trasmissione e quindi aumentano i costi. Se indichiamo con $d := d(\mathbf{C})$ la **distanza minima** del codice **C**, cioè il numero $d(\mathbf{C}) := \min \{ \rho(\mathbf{c}_1, \mathbf{c}_2); \mathbf{c}_1, \mathbf{c}_2 \in \mathbf{C}, \mathbf{c}_1 \neq \mathbf{c}_2 \}$, allora si definisce il **tasso di correzione** $\lambda(\mathbf{C})$ come il rapporto d/n. La regola di decodificazione precedentemente illustrata sarà in grado di correggere tanti più errori avvenuti nel canale quanto maggiore è la distanza minima del codice e quindi quanto maggiore è il suo tasso di correzione; quest'ultimo ci fornisce dunque una buona misura dell'affidabilità del sistema.

La contemporanea presenza di un alto tasso d'informazione e di un alto tasso di correzione in uno stesso codice è un evento che si verifica raramente: nelle famiglie di codici che si conoscono a tutt'oggi almeno uno dei due parametri tende a zero al divergere della lunghezza. Un famoso risultato di C. Shannon [**45**] asserisce che in corrispondenza di un prefissato $\varepsilon > 0$ è possibile trovare un intero $n(\varepsilon)$ tale che per ogni $n > n(\varepsilon)$ esiste un codice **C** di lunghezza n con il quale si ottiene, in presenza di un decodificatore a massima verosimiglianza, una probabilità di errori di trasmissione inferiore a ε. La natura di questo risultato è per l'appunto puramente esistenziale: a parte le informazioni sui suoi parametri, nulla conosciamo della struttura del codice **C** e quindi non sappiamo se sia effettivamente possibile trovare per esso delle regole di codificazione e decodificazione efficienti e realizzabili in pratica. Ci troviamo nella situazione conflittuale di dover garantire la sicurezza della trasmissione con apparecchiature dal costo accettabile.

Da questo punto di vista diventa allora importante avere a disposizione dei codici dotati di molte proprietà descrivibili in termini matematici. Un buon codice è generalmente ricco di "simmetrie" e ha una certa "omogeneità". Nei Paragrafi che seguono descriveremo una classe di codici lineari binari, i *codici di Reed-Muller*, che raggiunsero una certa fama all'inizio degli anni '70 quando furono utilizzati dalla NASA per equipaggiare le sue sonde spaziali. La *decodificazione a soglia* o *a maggioranza* introdotta per questi codici fu estesa ad altre classi di codici e ha il vantaggio di una semplice realizzazione strumentale, [**37**]. Ciò non significa necessariamente che le proprietà di questi codici utilizzate per la decodificazione siano semplici o comunque evidenti. In quest'ottica è molto istruttiva la lettura della prefazione del libro di Elwyn R. Berlekamp [**12**] che spiega come lo stesso fenomeno accada per un'altra importantissima classe di codici, i cosiddetti codici BCH.

6. LA SOMMA MISTA DI DUE CODICI LINEARI BINARI

Se n è un intero positivo indicheremo con $V_n(q)$ lo spazio vettoriale di tutte le n-uple a coefficienti nel campo finito $GF(q)$; ci interesserà particolarmente il caso binario $q = 2$: gli elementi di $GF(2)$ si chiamano **bit**. Se $\mathbf{x} \in V_n(q)$ è una parola arbitraria chiameremo **peso (di Hamming)** di $\mathbf{x}$ il numero $\gamma(\mathbf{x})$ delle componenti non nulle di $\mathbf{x}$. La distanza di Hamming su $V_n(q)$ è **invariante per traslazioni,** cioè per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V_n(q)$ si ha sempre $\rho(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x}+\mathbf{z}, \mathbf{y}+\mathbf{z})$. Ne segue che la distanza di Hamming $\rho(\mathbf{x}, \mathbf{y})$ di due qualunque parole $\mathbf{x}, \mathbf{y} \in V_n(q)$ è uguale al peso di Hamming $\gamma(\mathbf{x} - \mathbf{y})$ della loro differenza.

Un qualunque sottospazio vettoriale k-dimensionale $\mathbf{C}$ di $V_n(q)$ verrà chiamato (n, k)-**codice lineare su** $GF(q)$ (nel caso $q = 2$ parleremo di codici lineari **binari).** Una matrice k x n a coefficienti in $GF(q)$ le cui righe formino una base del sottospazio vettoriale $\mathbf{C}$ si dirà una **matrice generatrice** di $\mathbf{C}$. Si verifica subito che la distanza minima di un codice lineare è uguale al peso minimo delle sue parole non nulle.

Siano $\mathbf{C}_1$ e $\mathbf{C}_2$ due codici lineari binari di lunghezza n e di rispettive dimensioni k_1 e k_2. La **somma mista** di $\mathbf{C}_1$ e $\mathbf{C}_2$ è definita da $\mathbf{C}_1 \,\&\, \mathbf{C}_2 := \{(\mathbf{x}_1, \mathbf{x}_1 + \mathbf{x}_2); \mathbf{x}_1 \in \mathbf{C}_1, \mathbf{x}_2 \in \mathbf{C}_2\}$.

Se G_1 e G_2 sono rispettivamente matrici generatrici per $\mathbf{C}_1$ e $\mathbf{C}_2$ allora la matrice

$$G_1 \,\&\, G_2 = \begin{pmatrix} G_1 & | & G_1 \\ \text{—} & | & \text{—} \\ 0 & | & G_2 \end{pmatrix}$$

è una matrice generatrice della somma mista $\mathbf{C}_1 \,\&\, \mathbf{C}_2$. Non è difficile verificare che se $\mathbf{C}_1$ ha distanza minima d_1 e $\mathbf{C}_2$ ha distanza minima d_2 allora la somma mista $\mathbf{C}_1 \,\&\, \mathbf{C}_2$ è un $(2n, k_1 + k_2)$-codice lineare binario con distanza minima $d := \min\{2d_1, d_2\}$.

Il tasso d'informazione $R(\mathbf{C}_1 \& \mathbf{C}_2) = (k_1 + k_2)/(2n)$ della somma mista $\mathbf{C}_1 \,\&\, \mathbf{C}_2$ si calcola come media aritmetica dei tassi d'informazione $R(\mathbf{C}_1) = k_1/n$ di $\mathbf{C}_1$ e $R(\mathbf{C}_2) = k_2/n$ di $\mathbf{C}_2$. La somma mista è particolarmente promettente quando risulta $d_2 = 2d_1$. In tal caso infatti $\mathbf{C}_1 \,\&\, \mathbf{C}_2$ ha lo stesso tasso di correzione d_1/n di $\mathbf{C}_1$.

La somma mista si può naturalmente applicare anche a codici $\mathbf{C}_1$ e $\mathbf{C}_2$ di lunghezza diversa; in tal caso basta estendere il codice di lunghezza minore fino a raggiungere la lunghezza dell'altro, eventualmente aggiungendo brutalmente degli zeri nelle componenti mancanti.

7. DEFINIZIONE COMBINATORIA DEI CODICI DI REED-MULLER

Utilizziamo formalmente la stessa formula ricorsiva dei coefficienti binomiali abbinata alla somma mista per definire, in corrispondenza agli interi m,s con $0 \leq s \leq m$ il **codice** (binario) di **Reed-Muller** $RM(m, s)$ **dell's-esimo ordine e di lunghezza** 2^m:

$$RM\ (m+1, s+1) := RM\ (m, s+1)\ \&\ RM\ (m, s) \quad \text{per } 0 \leq s \leq m.$$

Come valori iniziali per la ricorsione definiamo i codici di Reed-Muller

$$RM(m, 0) := \{00\ldots0, 11\ldots1\} \subseteq V_{2^m}(2), \quad RM(m, m) := V_{2^m}(2) \quad \text{per } m \geq 0.$$

I codici di Reed-Muller dello 0-esimo ordine sono codici ripetitivi, mentre il codice di Reed-Muller dell'm-esimo ordine e di lunghezza 2^m coincide con lo spazio vettoriale di tutte le 2^m-uple a componenti in $GF(2)$. Le distanze minime $d := d(RM(m,s))$ e le dimensioni $k := \dim(RM(m,s))$ si calcolano ricorsivamente a partire dai valori iniziali $d(RM(m, 0)) = \dim(RM(m, m)) = 2^m$, $d(RM(m, m)) = \dim(RM(m, 0)) = 1$, $m \geq 0$, mediante le formule ricorsive

$$d\ (RM(m+1, s+1)) = \min\{2d(RM\ (m, s+1)), d(RM(m, s))\}$$

e

$$\dim(RM(m+1, s+1)) = \dim(RM(m, s+1)) + \dim(RM(m, s))$$

per $0 \leq s \leq m$. Per induzione su m si riconoscono le relazioni

$$d = d(RM(m, s)) = 2^{m-s}, \qquad k = \dim(RM(m, s)) = \sum_{i=0}^{s} \binom{m}{i}.$$

(Per la somma $\sum_{i=0}^{s} \binom{m}{i}$, $s < m$, non esiste un'espressione più semplice).

8. MATRICI GENERATRICI

Definiamo innanzitutto ricorsivamente per $m \geq 0$ una matrice generatrice da indicarsi con $G(m)$ del codice di Reed-Muller banale $RM\ (m, m) = V_n(2)$ dell'm-esimo ordine e lunghezza $n := 2^m$. L'unica matrice generatrice di $RM(0, 0)$ è la matrice $G(0) := (1)$ di tipo 1 x 1. Per tutti gli $m \geq 0$ utilizziamo il metodo illustrato precedentemente per costruire ricorsivamente la matrice generatrice $G(m+1)$ del codice $RM(m+1, m+1) = RM(m,m)\&RM(m,m)$ a partire dalla matrice generatrice $G(m)$ del codice $RM(m, m)$:

$$G(m+1) := \left(\begin{array}{c|c} G(m) & G(m) \\ \hline 0 & G(m) \end{array}\right)$$

La matrice $G(m)$ è una matrice triangolare superiore di tipo n x n. Associamo ora ordinatamente a ciascuna delle $n = 2^m$ righe di $G(m)$ uno degli n sottoinsiemi ø, {0}, {1}, {1,0}, {2}, {2,0}, {2,1}, {2,1,0}, {3}, {3,0}, ... (ordinamento lessicografico) dell'insieme $M := \{0,1, ..., m - 1\}$; tale insieme verrà detto **l'insieme caratteristico** della riga. (Se non vi è pericolo di equivoci tralasciamo nella scrittura degli insiemi caratteristici le parentesi graffe "{"e"}" nonché le virgole, p.es. {3,2,0} = 320.) Se numeriamo le righe di $G(m)$ dall'alto verso il basso con gli interi 0, 1, ..., n - 1, allora il numero assegnato alla riga avente l'insieme caratteristico $A \subseteq M = \{0,1, . . ., m - 1\}$ risulta essere $\sum_{j \in A} 2^j$. Gli elementi di un insieme caratteristico indicano pertanto quelle posizioni nelle quali la rappresentazione binaria del numero di riga contiene un 1; p.es. numero di riga $13 = 2^3 + 2^2 + 2^0 = 1101$ (rappr. binaria), insieme caratteristico $\{3, 2, 0\}$. Per brevità indichiamo con $\mathbf{v}(A)$ la riga cui è associato l'insieme caratteristico $A \subseteq M = \{0,1, ..., m - 1\}$, p.es. $\mathbf{v}(320) = 0000000000000101$.

Descriviamo le righe della matrice $G(m+1)$ ricorsivamente con l'aiuto delle righe della matrice $G(m)$, m = 1,2, ...; una riga di $G(m+1)$ il cui insieme caratterisitco non contenga l'elemento m si ottiene scrivendo due volte di seguito la riga di $G(m)$ che pure possiede A come insieme caratteristico; una riga di $G(m + 1)$ il cui insieme caratteristico contenga l'elemento m si ottiene scrivendo il vettore nullo 2^m-dimensionale seguito dalla riga di $G(m)$ avente l'insieme caratteristico $A \setminus \{m\}$.

Da questa descrizione ricorsiva si ricava subito la seguente descrizione non ricorsiva delle righe di $G(m)$. Determiniamo le componenti $x(0), x(1), ..., x(n - 1)$ della riga $\mathbf{v}(A)$ di $G(m)$ di cui sia noto l'insieme caratteristico $A \subseteq M = \{0,1, ..., m - 1\}$. Per $j = m - 1, m - 2, . . . , 1, 0$ si ha che se $j \in A$ allora risulta

$$\begin{array}{l} x(0) \quad = x(1) \quad = \ldots = x(2^j - 1) \quad = 0 \\ x(2 \cdot 2^j) = x(2 \cdot 2^j + 1) = \ldots = x(3 \cdot 2^j - 1) = 0 \\ x(4 \cdot 2^j) = x(4 \cdot 2^j + 1) = \ldots = x(5 \cdot 2^j - 1) = 0 \\ x(6 \cdot 2^j) = x(6 \cdot 2^j + 1) = \ldots = x(7 \cdot 2^j - 1) = 0 \\ \quad \vdots \qquad \vdots \qquad \vdots \end{array}$$

Le componenti che alla fine non risultano ancora determinate hanno il valore 1. A parole: Se per un dato insieme caratteristico $A \subseteq M$ vogliamo determinare la corrispondente riga $\mathbf{v}(A)$ di $G(m)$ procediamo come segue. Per ogni $j \in A$ poniamo una dopo l'altra uguali a zero le prime 2^j componenti, lasciamo indeterminate le successive 2^j componenti, poniamo di nuovo uguali a zero le

successive 2^j componenti, lasciamo indeterminate le successive 2^j componenti e così via. Alla fine poniamo uguali a 1 tutte le componenti che sono rimaste indeterminate.

Da questa descrizione non ricorsiva si riconosce che se $\mathbf{v}(A)$ e $\mathbf{v}(B)$ sono righe di $G(m)$, allora la riga $\mathbf{v}(A \cup B)$ ha componente 1 esattamente nelle posizioni in cui sia $\mathbf{v}(A)$ che $\mathbf{v}(B)$ hanno componente 1. Grazie a questa semplice regola possiamo costruire una qualunque riga $\mathbf{v}(A)$ di $G(m)$ a partire dalle righe $\mathbf{v}(0)$, $\mathbf{v}(1),\ldots, \mathbf{v}(m-1)$ mediante la relazione $\mathbf{v}(A) = \mathbf{v}(\cup_{j \in A} \{j\})$; pertanto la riga $\mathbf{v}(320)$ ha componente 1 esattamente nelle posizioni in cui le righe $\mathbf{v}(3)$, $\mathbf{v}(2)$ e $\mathbf{v}(0)$ hanno in comune la componente 1. Sfrutteremo questa proprietà per la caratterizzazione algebrica dei codici di Reed-Muller del Paragrafo 5.

Per $s = 0, 1, 2, \ldots, m$ indichiamo ora con $G(m, s)$ la matrice che consiste di tutte le righe di $G(m)$ il cui insieme caratteristico ha al più s elementi. Dalla precedente descrizione ricorsiva delle righe di $G(m+1)$ possiamo dedurre una formula ricorsiva per le matrici $G(m, s)$: per tutti gli interi m, s con $0 \leq s < m$ risulta

$$G(m+1, s+1) := \left(\begin{array}{c|c} G(m, s+1) & G(m, s+1) \\ \hline 0 & G(m, s) \end{array} \right)$$

Per ogni intero $m \geq 0$ la matrice $G(m, m)$ coincide con la matrice generatrice $G(m)$ del codice $RM(m, m)$, mentre la matrice $G(m, 0)$ consiste del solo vettore 2^m-dimensionale $\mathbf{v}(0) = \mathbf{1} = 11 \ldots 1$ ed è pertanto una matrice generatrice del codice $RM(m, 0)$. Ne ricaviamo che $G(m+1, s+1)$ è una matrice generatrice della somma mista del codice generato da $G(m, s+1)$ e del codice generato da $G(m, s)$ (somma eseguita nell'ordine indicato) e quindi che la matrice $G(m, s)$ è una matrice generatrice del codice di Reed-Muller $RM(m, s)$ che chiameremo **matrice generatrice canonica** di $RM(m, s)$.

Richiamiamo l'attenzione su due proprietà dei codici di Reed-Muller $RM(m, s)$ che si ottengono direttamente dalla rappresentazione della loro matrice generatrice canonica come sottomatrice di $G(m)$ formata scegliendone opportunamente certe righe:

1. $RM(m, s) \subseteq RM(m, s+1)$ per $0 \leq s < m$,
2. $k = \dim (RM(m,s)) = \sum_{i=0}^{s} \binom{m}{i}$ per $0 \leq s \leq m$.

9. CARATTERIZZAZIONE ALGEBRICA

Siano m, s interi con $0 \leq s \leq m$ e sia $n := 2^m$ Definiamo in $V_n(2)$ il **prodotto**

per componenti

$$\times: \begin{cases} V_n(2) \times V_n(2) & \to & V_n(2); \\ (\mathbf{x} = x_1 x_2 \ldots x_n, \mathbf{y} = y_1 y_2 \ldots y_n) & \mapsto & \mathbf{x} \times \mathbf{y} = (x_1 y_1, x_2 y_2, \ldots, x_n y_n) \end{cases}$$

Il prodotto per componenti $\mathbf{x} \times \mathbf{y}$ di due vettori $\mathbf{x}$ e $\mathbf{y}$ ha componente 1 esattamente nelle posizioni in cui sia $\mathbf{x}$ che $\mathbf{y}$ hanno componente 1. Per tutti i vettori $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V_n(2)$ e tutti gli scalari $\lambda \in GF(2)$ valgono banalmente le proprietà algebriche

$$(\mathbf{x} \times \mathbf{y}) \times \mathbf{z} = \mathbf{x} \times (\mathbf{y} \times \mathbf{z}), \quad \mathbf{x} \times \mathbf{y} = \mathbf{y} \times \mathbf{x}, \quad \mathbf{x} \times \mathbf{1} = \mathbf{x},$$
$$(\mathbf{x} + \mathbf{y}) \times \mathbf{z} = \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z}, \quad \lambda (\mathbf{x} \times \mathbf{y}) = (\lambda \mathbf{x}) \times \mathbf{y} \quad \text{e} \quad \mathbf{x} \times \mathbf{x} = \mathbf{x}.$$

Dalla descrizione non ricorsiva delle righe di $G(m)$ del precedente Paragrafo ricaviamo che per ogni sottoinsieme $A = \{j_1, j_2, \ldots, j_s\} \subseteq M$ con $|A| = s$ la riga di $G(m)$ avente A come insieme caratteristico può essere scritta come

$$\mathbf{v}(A) = \mathbf{v}\left(\cup_{i=1}^{s} \{j_i\}\right) = \mathbf{v}(j_1) \times \mathbf{v}(j_2) \times \ldots \times \mathbf{v}(j_s),$$

p.es. $\mathbf{v}(320) = \mathbf{v}(3) \times \mathbf{v}(2) \times \mathbf{v}(0)$. Le righe di $G(m)$ aventi insiemi caratteristici con al più s elementi formano una base del codice di Reed-Muller $RM(m, s)$ dell's-esimo ordine e lunghezza $n = 2^m$ Osserviamo che quando il prodotto per componenti è privo di fattori si ottiene il vettore $\mathbf{v}(0) = \mathbf{1} = 11\ldots1$. Abbiamo così la

CARATTERIZZAZIONE ALGEBRICA DEI CODICI DI REED-MULLER. *Il codice di Reed-Muller $RM(m, s)$ dell's-esimo ordine consiste delle combinazioni lineari dei prodotti per componenti di al più s delle righe $\mathbf{v}(0), \mathbf{v}(1), \ldots, \mathbf{v}(m-1)$ della matrice $G(m)$.*

Le righe $\mathbf{v}(0), \mathbf{v}(1), \ldots, \mathbf{v}(m-1)$ formano una base di $RM(m, 1)$. Un semplice calcolo mostra che il prodotto per componenti di al più s vettori di $RM(m, 1)$ è una parola di $RM(m, s)$. Pertanto il codice $RM(m, s)$ si può caratterizzare anche come lo spazio delle combinazioni lineari dei prodotti per componenti di al più s parole di $RM(m, 1)$.

10. CARATTERIZZAZIONE GEOMETRICA

Per lo studio di certe proprietà dei codici di Reed-Muller è più appropriata una descrizione geometrica di quella combinatoria o quella algebrica. Ricapitoliamo innanzitutto alcuni concetti fondamentali della geometria.

Sia F un campo, sia m un intero positivo e sia $V_m(F)$ lo spazio vettoriale di tutte le m-uple a componenti in F. Per ogni vettore $\mathbf{a} \in V_m(F)$ la **traslazione**

$$\tau_{\mathbf{a}} : V_m(F) \to V_m(F); \qquad \mathbf{x} \mapsto \tau_{\mathbf{a}}(\mathbf{x}) := \mathbf{x}+\mathbf{a}$$

è una biiezione. Le traslazioni di $V_m(F)$ rispetto alla composizione di applicazioni formano un gruppo isomorfo al gruppo additivo di $V_m(F)$. Sia ora $U \subseteq V_m(F)$ un sottospazio vettoriale di dimensione $s := \dim U$; un **traslato** di U è una classe laterale $\tau_{\mathbf{a}}(U) = \mathbf{a} + U$ dove $\mathbf{a}$ è un vettore di $V_m(F)$; i traslati di tutti i sottospazi vettoriali s-dimensionali di $V_m(F)$ si chiamano **sottospazi affini di dimensione** s di $V_m(F)$. I sottospazi affini di dimensione rispettivamente $0, 1, 2, m-2, m-1$ si chiamano rispettivamente **punti, rette, piani, iperrette, iperpiani.** Due sottospazi affini U, W della stessa dimensione si dicono **paralleli** e si scrive $U \| W$ se esiste una traslazione $\tau_{\mathbf{a}}$ di $V_m(F)$ tale che risulti $\tau_{\mathbf{a}}(U) = W$. Per ogni dato sottospazio vettoriale U e ogni punto $\mathbf{a} \in V_m(F)$ esiste esattamente un sottospazio affine W (della stessa dimensione di U) parallelo a U e tale che $\mathbf{a} \in W$, si tratta precisamente del sottospazio $W = \tau_{\mathbf{a}}(U) = \mathbf{a} + U$ (*postulato euclideo della parallela*). Due sottospazi affini paralleli o coincidono oppure sono disgiunti. I sottospazi affini paralleli a un dato sottospazio affine U formano una partizione di $V_m(F)$, cioè ricoprono tutto $V_m(F)$ senza sovrapporsi.

In un contesto geometrico, invece che dello spazio vettoriale $V_m(F)$ si parla dello **spazio affine** $AG_m(F)$ (se $F = GF(q)$ si scrive anche $AG_m(q)$) e si utilizzano terminologie geometriche come "un punto $\mathbf{a}$ è *incidente* un sottospazio U" invece di "$\mathbf{a} \in U$", "due rette G e H si *intersecano* in un punto $\mathbf{a}$" invece di "$\{\mathbf{a}\} = G \cap H$" e così via.

Comunque presi due punti distinti $\mathbf{a}, \mathbf{b} \in AG_m(F)$ esiste un'unica retta incidente entrambi, vale a dire $\{\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}); \lambda \in F\}$. Più in generale: comunque dati $s+1$ punti, tali che s qualunque di essi non giacciano mai in uno stesso sottospazio affine di dimensione $s-1$, esiste un unico sottospazio affine s-dimensionale incidente ciascuno di essi.

L'intersezione di una qualunque famiglia di sottospazi affini è ancora un sottospazio affine. L'intersezione di s iperpiani è vuota oppure è un sottospazio affine di dimensione almeno $(m-s)$. Ogni sottospazio affine $(m-s)$-dimensionale si può rappresentare come intersezione di s opportuni iperpiani.

Un ben noto conteggio mostra che $V_m(q)$ contiene $\begin{bmatrix} m \\ s \end{bmatrix}_q := \prod_{i=0}^{s-1}(q^{m-i}-1)/(q^{s-i}-1)$ sottospazi vettoriali s-dimensionali; pertanto lo spazio affine $AG_m(q)$ contiene esattamente $q^{m-s}\begin{bmatrix} m \\ s \end{bmatrix}_q$ sottospazi affini s-dimensionali, in quanto ciascuno dei q^{m-s} traslati di un sottospazio vettoriale s-dimensionale di $V_m(q)$ è un sottospazio affine s-dimensionale di $AG_m(q)$

Per la descrizione dei codici di Reed-Muller utilizziamo lo spazio affine m-dimensionale $AG_m(2)$ su $GF(2)$. Ogni sottospazio vettoriale $(m-1)$-dimensionale di $V_m(2)$ consiste di 2^{m-1} vettori; l'insieme complementare di un iperpiano

in $AG_m(2)$ è dunque un iperpiano ad esso parallelo. La *differenza simmetrica* $(U \cup V) \setminus (U \cap V)$ di due iperpiani U e V di $AG_m(2)$ o è un iperpiano oppure è l'insieme vuoto (nel caso U = W) oppure è tutto lo spazio affine (nel caso $U \| W$ con $U \neq W$). In generale tuttavia la differenza simmetrica di due sottospazi affini non è un sottospazio affine.

Scriviamo ora in un ordine particolare gli $n = 2^m$ punti dello spazio affine $AG_m(2)$ come colonne $\mathbf{s}(0), \mathbf{s}(1), \ldots \mathbf{s}(n-1)$ di una matrice $m \times n$ che denoteremo con $S(m)$: le righe di $S(m)$ devono risultare uguali ai complementi dei vettori $\mathbf{v}(0), \mathbf{v}(1), \ldots, \mathbf{v}(m-1)$ cioè i vettori $\mathbf{v}^*(0) := \mathbf{v}(0) + \mathbf{1}$, $\mathbf{v}^*(1) := \mathbf{v}(1) + \mathbf{1}, \ldots,$ $\mathbf{v}^*(m-1) := \mathbf{v}(m-1) + \mathbf{1}$. Fra l'altro per $j = 0,1, \ldots, n-1$ la colonna nr. j di $S(m)$ dà proprio la rappresentazione binaria del numero $n - j - 1$.

Possiamo ora rappresentare un qualunque sottoinsieme X di $AG_m(2)$ mediante il suo **vettore d'incidenza** $x_0 x_1 \ldots x_{n-1} \in V_n(2)$: poniamo $x_j = 1$ se e soltanto se $s(j) \in X$. Se $\mathbf{x}$, $\mathbf{y}$ sono rispettivamente i vettori d'incidenza dei sottoinsiemi X, Y di $AG_m(2)$ allora la somma $\mathbf{x} + \mathbf{y}$ e il prodotto per componenti $\mathbf{x} \times \mathbf{y}$ sono rispettivamente i vettori d'incidenza della differenza simmetrica $(X \cup Y) \setminus (X \cap Y)$ e dell'intersezione $X \cap Y$.

Per $i = 0,1, \ldots, m-1$ la riga $\mathbf{v}(i)$ di $G(m)$ è proprio il vettore d'incidenza dell'insieme di tutti quei punti $\mathbf{s}(j) = (s_{j,0}, s_{j,1}, \ldots s_{j,m-1})^T$ di $AG_m(2)$ per i quali risulti $s_{j,i} = 0$; in altre parole $\mathbf{v}(i)$ è il vettore d'incidenza di un sottospazio vettoriale $(m-1)$ dimensionale $U(i)$ di $V_m(2)$, cioè di un iperpiano di $AG_m(2)$ contenente il vettore nullo $\mathbf{s}(n-1) = \mathbf{0}^T$. Più in generale: per ogni insieme caratteristico $A \subseteq M := \{0,1, \ldots, m-1\}$ con $|A| = s$ la riga $\mathbf{v}(A)$ di $G(m)$ è il vettore d'incidenza di quel sottospazio vettoriale $(m-s)$-dimensionale $U(A)$ di $V_m(2)$ che contiene tutti i punti $\mathbf{s}(j)$ tali che risulti $s_{j,i} = 0$ per ogni $i \in A$.

Comunque presi due insiemi caratteristici $A, B \subseteq M = \{0,1, \ldots, m-1\}$ risulta

$$1)\ A \subseteq B \Leftrightarrow U(A) \supseteq (B),$$
$$2)\ U(A \cup B) = U(A) \cap U(B),$$
$$3)\ U(A \cap B) = U(A) + U(B).$$

In particolare si ha $U(M) = \{\mathbf{0}\}$ e $U(\emptyset) = V_m(2)$. L'applicazione che ad ogni insieme caratteristico associa il sottospazio vettoriale $U(A)$ di $V_m(2)$ è dunque un anti-isomorfismo dal reticolo $\mathcal{P}(M)$ dei sottoinsiemi di M al sottoreticolo $\{U(A); A \subseteq M\}$ del reticolo dei sottospazi vettoriali di $V_m(2)$.

Il codice di Reed-Muller del prim'ordine $RM(m, 1)$ contiene assieme ai vettori d'incidenza $\mathbf{v}(0), \mathbf{v}(1), \ldots, \mathbf{v}(m-1)$ dei sottospazi vettoriali $(m-1)$-dimensionali $U(0), U(1), \ldots, U(m-1)$ di $V_m(2)$ anche i loro complementi cioè i vettori d'incidenza $\mathbf{v}^*(0), \mathbf{v}^*(1), \ldots, \mathbf{v}^*(m-1)$ degli iperpiani ad essi paralleli. A parte la parola $\mathbf{v}(\emptyset) = \mathbf{1}$ e il vettore nullo $\mathbf{0}$, tutte le rimanenti $2(m-1)$ parole di $RM(m, 1)$, in quanto vettori d'incidenza di differenze simmetriche d'iperpiani, sono esse stesse vettori d'incidenza di iperpiani.

Lo spazio affine $AG_m(2)$ possiede esattamente $2 \cdot [\begin{smallmatrix} m \\ m-1 \end{smallmatrix}]_2 = 2(n-1)$ iperpiani; ne segue che il codice di Reed-Muller $RM(m, 1)$ consiste dei vettori d'incidenza di tutti gli iperpiani di $AG_m(2)$, del vettore d'incidenza $\mathbf{v}(0) = \mathbf{1}$ di tutto lo spazio $AG_m(2)$ e del vettore d'incidenza $\mathbf{0}$ dell'insieme vuoto $\emptyset$. Da questa descrizione delle parole del codice di Reed-Muller $RM(m, 1)$ e tenendo in mente la caratterizzazione algebrica si deduce che il codice $RM(m, s)$ contiene i vettori d'incidenza di tutte le intersezioni di al più s iperpiani di $AG_m(2)$.

CARATTERIZZAZIONE GEOMETRICA DEI CODICI DI REED-MULLER. *Numerando opportunamente i punti di $AG_m(2)$, il codice di Reed-Muller $RM(m,s)$ consiste dei vettori d'incidenza di tutti i sottospazi affini almeno $(m-s)$-dimensionali di $AG_m(2)$ e delle loro differenze simmetriche.*

11 . DECODIFICAZIONE A MAGGIORANZA

Presentiamo qui la decodificazione a maggioranza del codice di Reed-Muller $RM(m, s)$ dell's-esimo ordine e di lunghezza $n := 2^m$, introdotta da I.S. Reed nel 1954.

Il codificatore prende il blocco d'informazione

$$\mathbf{a}(s) = a_\emptyset a_0 a_1 a_{1,0} a_2 a_{2,0} a_{2,1} \ldots a_{m-1,\, m-2,\, \ldots,\, m-s}$$

costituito da $k := \sum_{i=0}^{s} \binom{m}{i}$ bit a_A (che indiciamo con gli insiemi caratteristici $A \subseteq M := \{0,1, \ldots, m-1\}$, $|A| \leq s$, ordinati lessicograficamente come nel Paragrafo 5) e lo codifica mediante la matrice generatrice canonica $G(m, s)$ nella parola

$$\mathbf{x}(s) := x_{s,0} x_{s,1} \ldots x_{s,n-1} := \sum_{A \subseteq M;\, |A| \leq s} a_A \, \mathbf{v}(A)$$

di $RM(m, s)$ e la manda al canale. Qui la parola $\mathbf{x}(s)$ viene (eventualmente) disturbata in alcune componenti e viene trasformata in una parola $\mathbf{y}(s) = y_{s,0} y_{s,1} \ldots y_{s,n-1}$ a partire dalla quale il decodificatore dovrebbe riuscire a ricostruire il blocco d'informazione originario $\mathbf{a}(s)$. Il decodificatore è strutturato in $s + 1$ *stadi* concatenati. Il primo stadio si chiama stadio nr. s, il secondo si chiama stadio nr. $s - 1, \ldots$, lo stadio $(s + 1)$-esimo si chiama stadio nr. 0. Nello stadio nr. $t = s, s - 1, \ldots, 0$ il decodificatore calcola per ogni insieme caratteristico $A \subseteq M$ con $|A| = t$ un bit b_A che rappresenta la sua stima del bit d'informazione originale a_A . Quando la procedura di decodificazione ha superato l'ultimo stadio nr. 0, il decodificatore consegna al ricevitore la sequenza di bit

$$\mathbf{b} = b_\varnothing b_\varnothing b_1 b_{1,0} b_2 b_{2,0} b_{2,1} \dots b_{m-1,\, m-2,\, \dots,\, m-s}$$

che costituisce la sua stima del blocco d'informazione $\mathbf{a}(s)$.

Ciascuno stadio nr. $t = s, s\text{-}1, \dots, 0$ eredita dallo stadio precedente nr. $t+1$ (il primo passo nr. $t = s$ eredita dal canale) una parola $\mathbf{y}(t) = y_{t,\varnothing} y_{t,1} \dots y_{t,n-1}$ calcola a partire dalle sue componenti i bit b_A con $A \subseteq M$, $|A| = t$ (come questo calcolo si svolga effettivamente verrà spiegato più avanti) e consegna la parola

$$\mathbf{y}(t\text{-}1) = y_{t-1,0} y_{t-1,1} \dots y_{t-1,n-1} := \mathbf{y}(t) - \sum_{\substack{A \subseteq M; \\ |A| = t}} b_A \, \mathbf{v}(A)$$

al successivo stadio nr. t-1 (con l'eccezione dell'ultimo stadio nr. $t = 0$).

La parola $\mathbf{y}(s)$ è una versione disturbata della parola $\mathbf{x}(s)$ del codice $RM(m, s)$. Più in generale possiamo interpretare la parola $\mathbf{y}(t)$ per $t = s, s\text{-}1, \dots, 0$ come una versione disturbata della parola

$$\mathbf{x}(t) := \mathbf{x}(s) - \sum_{\substack{A \subseteq M; \\ t < |A| \le s}} a_A \, \mathbf{v}(A) = \sum_{\substack{A \subseteq M; \\ |A| \le t}} a_A \, \mathbf{v}(A)$$

del codice di Reed-Muller $RM(m, t)$ del t-esimo ordine; i bit b_A sono in effetti l'approssimazione dei bit d'informazione a_A. Possiamo scrivere la parola $\mathbf{x}(t) \in RM(m, t)$ nella forma $\mathbf{a}(t) \cdot G(m, t)$ dove il blocco $\mathbf{a}(t) = a_\varnothing a_0 a_1 \dots a_{m-1,\, m-2,\, \dots,\, m-t}$ è una parte di lunghezza $\Sigma_{i=0}^{t} \binom{m}{i}$ del blocco d'informazione $\mathbf{a}(s)$. Il lavoro dello stadio nr. t del nostro $RM(m, s)$-decodificatore non si differenzia per nulla dal lavoro del primo stadio (che è pure lo stadio nr. t) di un $RM(m, t)$-decodificatore in un sistema di comunicazione che utilizzi il codice di Reed-Muller $RM(m, t)$ del t-esimo ordine e nel quale il codificatore codifichi il blocco d'informazione $\mathbf{a}(t)$ nella parola $\mathbf{x}(t) \in RM(m, t)$ ottenendo la parola $\mathbf{y}(t)$ in uscita dal canale.

Esaminiamo ora come vengono ricostruiti i bit b_A in un passo nr. t del decodificatore a partire dalle componenti $y_{t,0}, y_{t,1}, \dots, y_{t,n-1}$ della parola $\mathbf{y}(t)$.

Sia $A \subseteq M$ uno degli $\binom{m}{t}$ insiemi caratteristici con $|A| = t$, sia $A' = M \setminus A$ il suo insieme complementare e sia $U(A)$ risp. $U(A')$ il sottospazio vettoriale $(m\text{-}t)$-dimensionale risp. t-dimensionale di $V_m(2)$ dato dal vettore d'incidenza $\mathbf{v}(A)$ risp. $\mathbf{v}(A')$. Indichiamo con $U_1(A'), U_2(A'), \dots, U_{2^{m-t}}(A')$ i 2^{m-t} sottospazi affini di $AG_m(2)$ paralleli a $U(A')$. Per $i = 1,2, \dots, 2^{m-t}$ il decodificatore calcola la somma $b_A(i)$ di tutte le componenti $y_{t,j}$ di $\mathbf{y}(t)$ per le quali il punto $\mathbf{s}(j)$ appartiene al sottospazio affine $U_i(A')$:

$$b_A(i) := \sum_{\substack{0 \le j \le n-1; \\ \mathbf{s}(j) \in U_i(A')}} y_{t,j}; \qquad i = 1, 2, \dots, 2^{m-t}.$$

Dopo aver calcolato tutte queste 2^{m-t} *somme di controllo* il decodificatore prende una *decisione di maggioranza*: quando il numero delle somme di controllo $b_A(i)$ che hanno il valore 0 raggiunge o supera il *valore di soglia* $2^{m-t}/2 = 2^{m-t-1}$ allora il decodificatore pone $b_A := 0$, altrimenti pone $b_A := 1$. (La decisione per il bit 0 nel caso di un pareggio è del tutto arbitraria; talora in questi casi di patta un decodificatore un po' più complicato è in grado di prendere decisioni migliori).

Il codice $RM(m,s)$ ha distanza minima $d = 2^{m-s}$. L'algoritmo di decodificazione di Reed qui descritto garantisce una decodificazione corretta, cioè $\mathbf{b} = \mathbf{a}(s)$, se la parola di codice $\mathbf{x}(s)$ è stata disturbata in meno di $d/2$ componenti, cioè se la distanza di Hamming $\rho(\mathbf{x}(s), \mathbf{y}(s))$ risulta minore di 2^{m-s-1}. A tale scopo dimostriamo innanzitutto che per ogni insieme caratteristico $A \subseteq M$ di cardinalità t e ogni indice $i = 1, 2, \ldots, 2^{m-t}$ si ha sempre

$$a_A = \sum_{\substack{0 \le j \le n-1; \\ s(j) \in U_i(A')}} x_{t,j}.$$

In primo luogo abbiamo

$$\sum_{\substack{0 \le j \le n-1; \\ s(j) \in U_i(A')}} x_{t,j} = \sum_{\substack{0 \le j \le n-1; \\ s(j) \in U_i(A')}} \sum_{\substack{B \subseteq M;\ |B| \le t; \\ s(j) \in U(B)}} a_B = \sum_{\substack{B \subseteq M; \\ |B| \le t}} \sum_{\substack{0 \le j \le n-1 \\ s(j) \in U_i(A') \cap U(B)}} a_B.$$

Dal momento che stiamo calcolando in $GF(2)$, per tutti i sottoinsiemi caratteristici $B \subseteq M$ con $|B| \le t$ per i quali il sottospazio affine $U_i(A') \cap U(B)$ contenga un numero pari di punti, cioè per i quali risulti $\dim(U_i(A') \cap U(B)) \neq 0$, abbiamo la relazione

$$\sum_{\substack{0 \le j \le n-1 \\ s(j) \in U_i(A') \cap U(B)}} a_B = 0$$

L'intersezione del sottospazio affine t-dimensionale $U_i(A')$ col sottospazio vettoriale almeno $(m-t)$-dimensionale $U(B)$ consiste di un unico punto se e soltanto se i due insiemi caratteristici A' e B sono l'uno il complementare dell'altro, cioè se risulta $A = B$. Dalla relazione

$$\sum_{\substack{0 \le j \le n-1 \\ s(j) \in U_i(A') \cap U(A)}} a_A = a_A$$

segue l'asserto.

Se nel canale la parola di codice $\mathbf{x}(s)$ è stata disturbata in un numero e di posizioni minore di 2^{m-s-1}, allora il decodificatore nel primo stadio nr. s del calcolo dei bit b_A con $|A| = s$ prende sempre la decisione esatta, cioè risulta $b_A = a_A$. Pertanto anche la parola $\mathbf{y}(s-1)$ differisce dalla parola di codice $\mathbf{x}(s-1)$ in esat-

tamente e posizioni. Poiché e a maggior ragione risulta minore del valore di soglia 2^{m-s} del secondo stadio nr. s-1, allo stesso modo il decodificatore calcolando i bit b_A con $|A| = s$-1 determina sempre correttamente i bit d'informazione $b_A = a_A$. E così di seguito.

12 . UN ESEMPIO

Illustriamo l'algoritmo di decodificazione di Reed nel caso del (32, 6)-codice di Reed-Muller RM (5, 1).

Il codificatore prende un blocco d'informazione $\boldsymbol{a}$ (1) = $a_\emptyset a_0 a_1 a_2 a_3 a_4$ costituito da sei bit e lo codifica nella parola di codice

$$\mathbf{x}(1) = x_{1,0} x_{1,1} \ldots x_{1,31} = a_\emptyset \mathbf{v}(\emptyset) + \sum_{j=0}^{4} a_j \mathbf{v}(j).$$

Nel canale questa parola viene trasformata nella parola $\mathbf{y}(1) = y_{1,0} y_{1,1} \ldots y_{1,31}$. Il decodificatore inizia il suo lavoro:

1.0 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,1}, \quad y_{1,2} + y_{1,3}, \quad \ldots, y_{1,30} + y_{1,31}$$

hanno il valore 0 si ponga $b_0 := 0$, altrimenti si ponga $b_0 := 1$;

1.1 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,2}, \quad y_{1,1} + y_{1,3}, \quad \ldots, y_{1,29} + y_{1,31}$$

hanno il valore 0 si ponga $b_1 := 0$, altrimenti si ponga $b_1 := 1$

1.2 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,4}, \quad y_{1,1} + y_{1,5}, \quad \ldots, y_{1,27} + y_{1,31}$$

hanno il valore 0 si ponga $b_2 := 0$, altrimenti si ponga $b_2 := 1$;

1.3 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,8}, \quad y_{1,1} + y_{1,9}, \quad \ldots, y_{1,23} + y_{1,31}$$

hanno il valore 0, si ponga $b_3 := 0$, altrimenti si ponga $b_3 := 1$;

1.4 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,16}, \quad y_{1,1} + y_{1,17}, \quad \ldots, y_{1,15} + y_{1,31}$$

hanno il valore 0 si ponga $b_4 := 0$, altrimenti si ponga $b_4 := 1$.

Si costruisca la parola $\mathbf{y}(0) := y_{0,0}\, y_{0,1} \ldots y_{0,31} := \mathbf{y}(1) - \Sigma_{j=0}^{4} b_j\, \mathbf{v}(j)$.

2. Se almeno 16 dei 32 bit

$$y_{0,0}, y_{0,1}, \cdots, y_{0,31}$$

hanno il valore 0 si ponga $b_\varnothing := 0$, altrimenti si ponga $b_\varnothing := 1$.

Emettere il blocco $b_\varnothing b_0 b_1 b_2 b_3 b_4$.

BIBLIOGRAFIA

1. E. Ambrisi, F. Eugeni, *Il problema della protezione dell'informazione I: cenni storici e metodi statistici; per la decrittazione di sistemi di cifratura classici*, Ratio Math. **1** (1990), 1-29.
2. E. Angeleri, "Trasmissione dati", Editoriale Delfino, Milano, 1972.
3. L. Berardi, *A new secret code based on Steiner system*, Preprint (1989).
4. L. Berardi, *Some remarks about an Electronic Signature Derived from a Generalized RSA-code*, J. of Info. & Opti. Sci. **11** (1990), 189-194.
5. L. Berardi, A. Beutelspacher, *I buoni angeli custodi, ovvero i protettori di un messaggio*, Archimede **2**-3 (1989), 129-140.
6. L. Berardi, M . Di Fonso, *Protezione dell'informazione su personal computer*, Atti dei I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia (SPRCI) (1987), Fondazione Bordoni.
7. L. Berardi, M . Di Fonso, F. Eugeni, *Threshold Schemes based on Criss-Cross Block Designs*, J. of Info. & Opti. Sci. **11** (1990), 153-160.
8. L. Berardi, F. Eugeni, *Strutture geometriche, crittografia e sistemi di sicurezza richiedenti un quorum*, Atti del I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia, SPRCI (1987), 127-133, Fondazione Bordoni.
9. L. Berardi, F. Eugeni, *Smart Cards and Authentication schemes*, 5-th Meeting Euro Working Group on Financial Modelling, Acireale (CT) 1989. Rend. di Mat., Roma. In corso di stampa.
10. L. Berardi, F. Eugeni, *Blocking sets e teoria dei giochi*, Atti Sem. Mat. Fis. Univ. Modena **34** (1988), 165-196 .
11. L. Berardi, B. Rizzi, *Generalizziamo il codice RSA e la Funzione di Eulero*, Atti del I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia (SPRCI) (1987), Fondazione Bordoni.
12. E.R. Berlekamp, "Algebraic Coding Theory", McGraw-Hill, New York, 1968.
13. A. Beutelspacher, *La scuola elementare della teoria dei codici*, Quad. n.1 (suppl. didatt.) Sem. Geom. Comb. L'Aquila (1985).

14. A. Beutelspacher, "Kryptologie", Vieweg-Verlag, Braunschweig & Wiesbaden, 1988.
15. A. Bautelspacher, *Enciphered Geometry. Some applications of Geometry to Cryptography*, Ann. Disc. Math. **37** (1988), 59-68.
16. A. Beutelspacher, *Perfect and essentially perfect authentication systems*, in "Advances in Cryptology - EUROCRYPT '87", Lecture Notes Comp. Science, 304, D. Chaum, W.L. Price (Eds.), Springer, Berlin, 1988, pp. 167-170.
17. A. Beutelspacher, *How to say "no"*, Preprint (1989).
18. A. Beutelspacher, F. Eugeni, *Geometrie finite e Crittosistemi: stato dell'arte e problematiche*, Atti del II Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia (SPRCI) (1989), Fondazione Bordoni.
19. A. Beutelspacher, U. Rosenbaum, *Geometric authentication systems*, Ratio Math. **1** (1990), 30-41.
20. A. Bautelspacher, G. Tallini, C. Zanella, *Examples of essentially s-fold secure geometric authentication system with large s*, Rend. di Mat., Roma. In corso di stampa.
21. A. Beutelspacher, K. Vedder, *Geometric structures as threshold schemes*, IMA, Proceedings Cirencester, Cambridge Univ. Press. In corso di stampa.
22. G.R. Blakley, *Safeguerding Criptographic keys*, Proc. NCC **48** (1979), 313-317, AFIPS Press, Montvale, NJ .
23. O. Brugia, *Sistemi crittografici a chiave pubblica*, (Quad. n.4673 (1985), Scuola Superiore "G. Reiss Romoli", L'Aquila.
24. O. Brugia, S. Improta, W. Wolfowicz, *Segretezza ed autenticazione nelle moderne reti di comunicazione*, Rel. Int. 2B 63385 (1985), Fondazione Bordoni, Roma.
25. M. Cerasoli, F. Eugeni, M. Protasi, "Matematica Discreta", Zanichelli, Bologna, 1988.
26. M. De Soete, K. Vedder, *Some new classes of geometric threshold schemes*, in "Advances in Cryptology-EUROCRYPT '88", Lecture Notes Comp. Science, Springer, Berlin.
27. W. Diffie, M.E. Hellman, *New directions in Cryptography*, IEEE Trans. Inform. Theory **22** (1976), 644-654.
28. E.N. Gilbert, F.J. MacWilliams, N.J.A. Sloane, *Codes which detect deception*, Bell Syst. Techn. J. **53** (1974), 405-424.
29. M. Gionfriddo, *Sui sistemi di Steiner*, Quad. n.9 Sem. Geom. Comb. Univ. L'Aquila (1986), 3-28.
30. M. Gionfriddo, F. Eugeni, *On the minimum number of blocks of a maximal partial spread in STS(v) and SQS(v)*, J. of Geometry **36** (1989), 37-48.
31. W. Heise, P. Quattrocchi, "Informations - und Codierungstheorie," Springer, Berlin, 1988.
32. W . Heise, P. Quattrocchi, *Teoria algebrica dei codici*, Boll. Un. Mat. Ital. (7) **1**-A (1987), 313-331.

33. G. Longo, "Teoria dell'informazione," Boringhieri, Torino, 1980.
34. G. Longo, *Introduzione alla teoria dei codici algebrici*, Archimede **34**-3 (1982), 99-107.
35. L. Lunelli, "Teoria dell'informazione e dei codici," Clup, Milano, 1981.
36. F.J. MacWilliams, N.J.A. Sloane, "The theory of error-correcting codes", North-Holland, Amsterdam, 1977.
37. J.L. Massey, "Threshold decoding", MIT Press, Cambridge, MA, 1963.
38. A. Maturo, *Messaggi cifrati per mezzo di numeri pseudo-causali ottenuti a partire da successioni in algebre di supporto*, Atti del I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia, SPRCI (1987), Fondazione Bordoni.
39. D.E. Muller, *Application of Boolean algebra to switching circuit design and to error detection*, IEEE Trans. Computers **3** (1954), 6-12.
40. I.S. Reed, *A class of multiple-error-corretting codes and the decoding scheme*, IEEE Trans. Inform. Theory **4** (1954), 38-49.
41. A. Rizzi, *Alcune considerazioni sulla crittografia*, Atti del I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia, SPRCI (1987), Fondazione Bordoni.
42. L. Sacco, "Manuale di Crittografia," 1947. Ristampa Anastatica.
43. A. Sgarro, "Crittografia", Muzzio, Padova, 1986.
44. A. Shamir, *How to share a secret*, Communications ACM **11** (1979), 612-613.
45. C.E. Shannon, *A Mathematical Theory of communication*, Bell Syst. Techn. J. **27** (1948), 379-423.
46. C.E. Shannon, *Communication Theory of Secrecy Systems*, Bell Syst. Techn. J. **28** (1949), 656-715.

A. Bonisoli, Istituto di Matematica, Università della Basilicata, via N. Sauro 85, I-85100 Potenza. F. Eugeni, Dipartimento di Scienze e Storia dell'Architettura, Università "G. D'Annunzio", viale Pindaro 42, I-65127 Pescara.