2 TECNICHE CRITTOGRAFICHE PER LA PROTEZIONE DEI DATI

UNITÀ DI APPRENDIMENTO

L 1 Principi di crittografia

L 2 Dalla cifratura monoalfabetica ai nomenclatori

L 3 Crittografia bellica

L 4 Crittografia simmetrica (o a chiave privata)

L 5 Crittografia asimmetrica (o a chiave pubblica)

L 6 Certificati e firma digitale

hoepliscuola.it

hoepliscuola.it

OBIETTIVI

- Conoscere il significato di cifratura
- Avere il concetto di chiave pubblica e privata
- Conoscere gli elementi essenziali di "matematica per la criptografia"
- Sapere le tecniche monoalfabetiche per trasposizione e sostituzione
- Sapere le tecniche polialfabetiche di Alberti e Vigenere
- Apprendere i metodi poligrafici e i nomenclatori
- Conoscere il ruolo avuto dalla crittografia nelle due Guerre Mondiali
- Conoscere le macchine crittografiche e l'avvento della crittografia elettronica
- Conoscere la crittografia a chiave simmetrica e pubblica
- La firma digitale, l'algoritmo MD5 e i certificati digitali

ATTIVITÀ

- Saper utilizzare il:
  - il Playfair cipher
  - il cifrario bifido di Delastelle
  - la Cifra campale germanica
  - il Cifrario di Vernam
- Distinguere il cifrario DES, 3-DES e IDEA
- Conoscere l'algoritmo RSA
- Utilizzare le funzioni crittografiche in PHP
- Crittare file e volumi con TrueCrypt
- Firmare i documenti con la CNS
- Conoscere i possibili utilizzi della firma digitale

LEZIONE 1

PRINCIPI DI CRITTOGRAFIA

IN QUESTA UNITÀ IMPAREREMO...

- il significato di cifratura
- il concetto di chiave pubblica e privata
- gli elementi essenziali di "matematica per la criptografia"

La sicurezza nelle reti

Il problema della sicurezza nelle reti riveste una grande importanza dato che le reti per loro natura non sono sicure: basta un ◀ analizzatore di rete ▶ come un semplice *packet sniffer* tipo wireshark per intercettare le informazioni che viaggiano su di essa.

◀ Analizzatore di rete
Un analizzatore di rete è un programma che permette di esaminare il traffico tra due stazioni qualsiasi della rete. ▶

L'utilizzo della rete come strumento di transazioni commerciali, e quindi come mezzo di "*trasferimento di denaro*", ha incontrato come ostacolo alla piena diffusione la non completa fiducia da parte degli utenti di Internet verso gli strumenti telematici per comunicare i propri dati segreti per l'accesso ai conti correnti o per l'utilizzo di carte di credito online.
Sulle reti circolano anche documenti riservati, come accordi commerciali, relazioni tecniche su nuovi studi e ricerche, sia in ambito commerciale che scientifico, previsioni e analisi di mercato, ecc. e i *malintenzionati* sono sempre in agguato per cercare di intercettare tutte queste informazioni per farne usi illeciti.

Esistono diversi tipi di *malintenzionati* e diverse *motivazioni* per le quali questi cercano di intromettersi sulla rete. Le riportiamo nella seguente tabella:

SOGGETTO MALINTENZIONATO	SCOPO
◀ hacker ▶	violare e danneggiare
studente	curiosare nella posta altrui e non solo…
uomo d'affari	strategie di mercato
progettista	appropriarsi di progetti altrui
ex dipendente	danneggiare
bancario	furto

truffatore	rubare numeri di carte di credito
terrorista	rubare segreti strategici
spia	rubare segreti militari e civili
spionaggio/controspionaggio	intercettare messaggi e inviare messaggi falsi

◀ **Hacker** Hacker is a term used in computing for someone who accesses a computer system by circumventing its security system. ▶

È quindi di estrema importanza poter garantire la sicurezza della rete.

Possiamo individuare diversi aspetti connessi al *problema della sicurezza*:
- la segretezza;
- l'autenticazione;
- l'affidabilità dei documenti.

Con segretezza si intende l'aspetto più classico cioè che le informazioni siano leggibili e comprensibili solo a chi ne ha i diritti, cioè solo alle persone autorizzate: è necessario che gli altri non le possano intercettare o, comunque, non siano in grado di comprenderle.

Con autenticazione si intende il processo di riconoscimento delle credenziali dell'utente in modo di assicurarsi dell'identità di chi invia messaggi o esegue operazioni evitando che qualche malintenzionato si spacci per qualcun altro.

Con l'affidabilità dei documenti si intende di avere la garanzia e la certezza che un documento sia originale, cioè che il suo mittente sia certo (ad esempio mediante l'apposizione su di esso di una firma digitale) e che non sia stato letto e/o alterato e modificato da altre persone non autorizzate.

Le misure da intraprendere per ottenere la segretezza possono essere anche affrontate in diversi livelli della pila protocollare: a livello fisico si può cercare di impedire che avvengano intercettazioni di dati, a livello di data link si possono introdurre codifiche dei dati trasmessi per renderli incomprensibili agli hacker.
È comunque il livello di applicazione che può gestire gli altri due problemi ed è su quello che noi concentreremo la nostra attenzione.

Possiamo riassumere in due aspetti le richieste degli utenti di Internet:
- **A** la possibilità di codificare i dati scambiati per renderli incomprensibili;
- **B** la garanzia di integrità e autenticazione del mittente.

CRITTOGRAFIA
La scrittura segreta in codice o cifrata (Gabrielli, dix. Lingua italiana).

Entrambi hanno come base la crittografia (o ◀ criptografia ▶) cioè:

◀ Criptografia La parola criptografia deriva dal greco: κρυπτος ("kriptós" = nascosto) εγραφειν ("gráphein" = scrivere). Il termine che ne deriva significa dunque "scrittura nascosta". Essa, in altre parole, si occupa, dei metodi per rendere un messaggio non leggibile o non comprensibile a persone che non siano autorizzate a leggerlo. ▶

Lo studio della crittografia e della criptanalisi si chiama comunemente «criptologia».

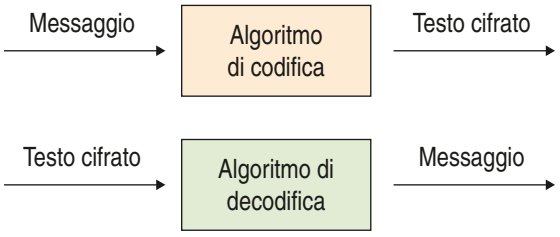
Crittografia

Il desiderio di mantenere nascosti messaggi tra due interlocutori agli occhi di terzi si perde nella notte dei tempi e non è una necessità nata con Internet; possiamo trovare tracce di tentativi di occultamento delle informazioni già su geroglifici di 4500 anni fa e la prova di un primo messaggio in codice è una tavoletta babilonese del 500 a.C. dove sono state tolte le prime consonanti di alcune parole e in altre sono state sostituite con simboli poco utilizzati rendendo a prima vista incomprensibile il messaggio.

CIFRATURA

Con cifratura intendiamo il processo mediante il quale un messaggio viene trasformato mediante un insieme di regole di codifica (algoritmo di cifratura – ◀ **Secrecy system** ▶) in formato tale da essere incomprensibile per occhi indiscreti.
Un algoritmo di cifratura è un metodo per stabilire una corrispondenza tra simboli in chiaro e simboli cifrati: i simboli in chiaro vengono utilizzati per creare il messaggio in chiaro (plain text message) che, una volta cifrato, diviene un testo crittografato o criptato (cipher text).
Il messaggio cifrato prende anche il nome di "crittogramma".

Naturalmente le regole di cifratura devono essere note sia al mittente del messaggio che al destinatario, in modo che quest'ultimo possa, alla sua ricezione, effettuarne il decriptaggio e comprenderne il significato.



Gli esperti di crittografia utilizzano anche il termine codifica, attribuendogli un significato diverso da quello degli informatici: con codifica essi intendono un *"metodo di scrittura in chiave che consiste nel sostituire alcune parole con altre"* distinguendolo dalla cifratura che più precisamente *"sostituisce lettere o caratteri"*.

ESEMPIO

Se volessimo trasferire una parola come "AIUTO" utilizzando i due metodi potremmo:

- **A** trasmettere "HELP" oppure "SOCCORSO", cioè sostituire completamene la parola con un'altra, magari in una lingua diversa, non conosciuta a chi potrebbe intercettarla (codifica) (la lingua degli Indiani Navajo fu usata nella II guerra mondiale per le operazioni nel Pacifico);
- **B** trasmettere "BLVUP", cioè sostiture a ciascuna lettera quella che la segue nell'alfabeto (cifratura).

Quest'ultimo esempio ci permette di effettuare una osservazione: la regola di cifratura è generalmente composta da due elementi:

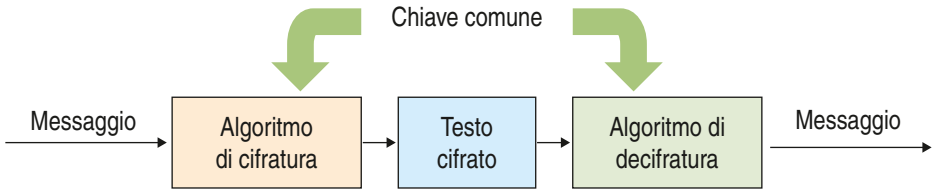
- **A** la regola vera e propria (l'algoritmo utilizzato), che in questo caso consiste nella "sostituzione di un carattere con un altro";

B uno (o più) parametri, in questo caso la posizione del carattere da prendere (nell'esempio il successivo a quello in chiaro: se invece si fosse trasmesso CMZVQ le posizioni sarebbero state due in avanti).

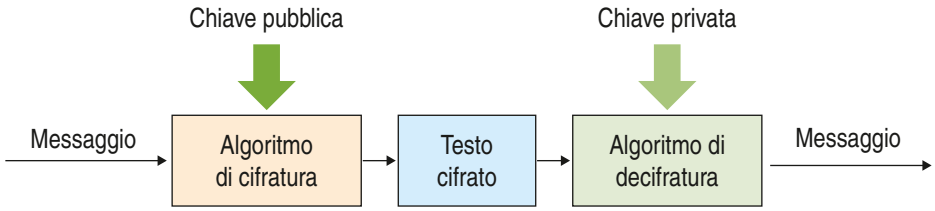
La distinzione tra *regola* e *parametro* è fondamentale nella crittografia: l'hacker deve conoscerli sempre entrambi e, quindi, basta modificarne uno periodicamente o in base ad accordi prestabiliti per aumentare la complessità di intercettazione.

La *regola* prende il nome di algoritmo di criptazione mentre il *parametro* di chiave.

Quando la chiave di cifratura coincide con quella di decifratura lo schema crittografico si dice simmetrico e la chiave prende il nome di chiave comune.



Quando la chiave di cifratura è invece diversa da quella usata per la decifratura lo schema crittografico si dice asimmetrico e le due chiavi si chiamano chiave pubblica quella usata per la cifratura, che è comune a tutti i mittenti e di pubblico dominio, e chiave privata quella utilizzata per la decifratura, che è segreta e di conoscenza solo del destinatario del messaggio.



Questo procedimento è alla base della moderna sicurezza delle reti: il mittente non deve comunicare col destinatario o accordarsi preventivamente, ma utilizza la chiave pubblica del destinatario che, proprio perché pubblica, è a disposizione di tutti, e con essa prepara il messaggio da trasmettere criptandolo in modo tale che solo chi è in possesso della chiave privata lo può decriptare.

È necessario che chiave pubblica e chiave privata siano *diverse*: per migliaia di anni la possibilità di cifrare un messaggio con una chiave e decifrarlo con una seconda chiave diversa dalla prima sembrava un assurdo, ma oggi, come vedremo in seguito, questo è possibile ed è utilizzato regolarmente nella pratica giornaliera.

◀ **Secrecy system** A secrecy system is defined abstractly as a set of transformations of one space (the set of possible messages) into a second space (the set of possible cryptograms). Each particular transformation of the set corresponds to enciphering with a particular key. The transformations are suppose dreversible (non-singular) so that unique deciphering is possible when the key is known." *Shannon, C.E., "Communication Theory of Secrecy System"* ▶

Crittoanalisi

Generalmente l'algoritmo di cifratura è noto e standardizzato, quindi conosciuto e soggetto a ◀ crittoanalisi ▶ da parte dei malintenzionati per individuare la chiave utilizzata e quindi decriptare il messaggio.

◀ Crittoanalisi La parola crittoanalisi proviene dal greco κρυπτοσ ("kryptós" = nascosto) αναλυειν ("analùein" = scomporre) e comporta lo studio dei metodi per ottenere il significato di informazioni cifrate: tipicamente si tratta delle operazioni effettuate alla ricerca della chiave segreta. ▶

ATTACCO

L'azione di un crittoanalista mirata a violare (rompere, sfondare) il crittosistema prende il nome di crittoanalisi.

Per poter effettuare un attacco gli intrusi devono essere in possesso dei messaggi cifrati e la natura stessa di Internet rende questa operazione molto semplice: lo scopo della crittografia è quello di rendere difficile la decifrazione del messaggio.

Il principio di Kerckhoffs (1835-1903) stabilisce che è la chiave l'elemento fondamentale per la sicurezza di un sistema informatico: la sua prima formulazione, nel trattato "La Cryptographie Militaire (1883)", diceva:
"È necessario che il sistema non richieda segretezza, e che possa senza problemi cadere in mano nemica."
Un altro modo di definire il principio di Kerckhoffs è riportato a fianco.
Come corollario al principio, Shannon aggiunse la frase: "il nemico conosce il sistema".

PRINCIPIO DI KERCKHOFFS

La sicurezza di un crittosistema deve dipendere solo dalla segretezza della chiave e non dalla segretezza dell'algoritmo usato.

A partire dal principio di Kerckhoffs si sono nel tempo aggiunte alcune riflessioni che hanno avuto un ruolo fondamentale nella moderna crittografia: le chiavi sono generalmente semplici e possono essere cambiate più frequentemente dell'algoritmo quindi, una volta determinato un buon algoritmo, è sufficiente "concentrarsi" sulla loro gestione.
Si è anche arrivati alla definizione di un cifrario assolutamente sicuro, il cosiddetto ◀ **one-time pad** ▶: in esso le due parti in comunicazione condividono un blocco (pad) di chiavi di sostituzione alfabetica generate con un procedimento casuale, e *cambiano la chiave a ogni lettera*.

◀ **One time pad** The one-time pad is the only encryption technique that has been mathematically proven to be uncrackable. While hard to use, it has often been the choice for highly sensitive traffic. Soviet spies used one-time pads in the 1940s and -50s. The Washington-Moscow "hot line" also uses one-time pads. However, the technique is hard to use correctly. ▶

Se il messaggio viene intercettato l'intruso vede solamente una sequenza di caratteri casuali e non è in grado di decifrare il messaggio.
Questo meccanismo trova implementazione nel cifrario di Vernam che descriveremo in seguito che, come dimostrato da Shannon, richiede per la sua realizzazione una chiave lunga quanto il messaggio stesso e necessita che il mittente e il destinatario siano sincronizzati per essere sicuri di partire dalla stessa posizione del blocco: risulta inoltre complesso realizzare il blocco chiave in modo che sia anch'esso sicuro.

Oggi esistono dei sistemi che generano e utilizzano una password "usa e getta", soprattutto per transazioni bancarie, che possiamo dire ispirati al "one time pad": possono definirsi sistemi "one

time key" in quanto la chiave può essere utilizzata una sola volta dato che generalmente ha una validità temporale modesta (10-20 secondi) dopo che è generata da un dispositivo elettronico (key generator) che è sincronizzato con il sistema di controllo di accesso al servizio.

Conclusioni

Alla base della criptografia c'è la matematica, in particolare:
Ⓐ l'artimetica modulare, con lo studio dei resti delle divisioni aritmetiche;
Ⓑ la teoria dei numeri, in particolare quella dei numeri primi.

Artimetica modulare

Nella aritmetica modulare il quoziente nell'operazione di divisione è irrilevante mentre unica importanza lo assume il resto, e viene così indicato:
Q il quoziente della divisione fra il dividendo **X** e il divisore **m**, mentre è **R** il resto e viene indicato con la seguente notazione:

$$X(\text{mod } m) = R$$

che si legge: "X modulo m è uguale a R" e si dice anche "R è congruo a X modulo m".

ESEMPIO

Vediamo alcuni esempi:
- 14(mod 4) = 2
- 79(mod 7) = 2
- 21(mod 33) = 21 dalla quale deduciamo che X(mod m) = X se X < m
- 37(mod 37) = 0 quindi m(mod m) = 0
- 27(mod 1) = 0 quindi X(mod 1) = 0
- 77(mod 76) = 1 quindi (m + 1)(mod m) = 1

Possiamo fare tre osservazioni sul resto R:
- vale sempre la relazione R < m;
- tutti i possibili resti sono in numero pari a m e con valori compresi fra 0 e m – 1, e l'insieme dei resti viene indicato con $Z_n = \{0, 1, 2, \ldots, n-1\}$;
- se X < m allora X(mod m)=X.

CLASSE DI RESTI

Dato un numero intero positivo *X*, i numeri interi si distribuiscono in *X* classi di resto modulo m, a seconda del resto che danno quando vengono divisi per *m*.

Valgono inoltre le seguenti due equivalenze:

$(X + Y)(\text{mod } m) = X(\text{mod } m) + Y(\text{mod } m)$, e cioè: ***il resto di una somma è pari alla somma dei resti***
$(X\ Y)(\text{mod } m) = X(\text{mod } m)\ Y(\text{mod } m)$, e cioè: ***il resto di un prodotto è pari al prodotto dei resti***.

L'equivalenza sul prodotto conduce alla importante equivalenza sul quadrato:

il resto di un quadrato è pari al quadrato del resto

$$X^2(\text{mod } m) = (X \cdot X)(\text{mod } m) = x(\text{mod } m) \cdot x(\text{mod } m) = R \cdot R = R^2$$

Grazie a questa equivalenza sarà possibile determinare resti di divisioni fra numeri con un incalcolabile numero di cifre, base della crittografia a chiave pubblica che utilizza i numeri primi.

ESEMPIO

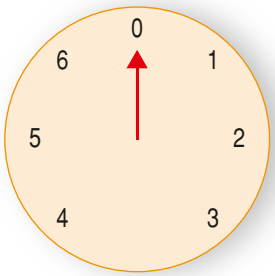
Ⓐ $13^2(\text{mod } 11) = 169(\text{mod } 11) = 4 = 13(\text{mod } 11) \cdot 13(\text{mod } 11) = 2 \cdot 2 = 4$
Ⓑ $25^2(\text{mod } 7) = 625(\text{mod } 7) = 2 = 25(\text{mod } 7) \cdot 25(\text{mod } 7) = 4 \cdot 4 = 16$

16, essendo maggiore di m, deve essere ulteriormente elaborato ottenendo:

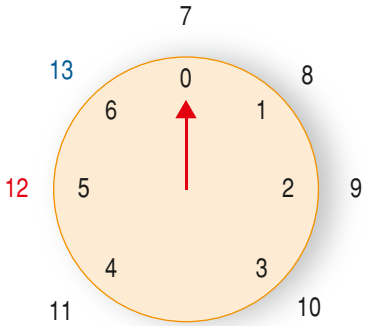
$$16(\text{mod } 7) = 2$$

L'aritmetica in modulo viene anche chiamata aritmetica *dell'orologio* in quanto è possibile ottenere il risultato considerando un orologio con m ore e muovendo la lancetta su di esso fino a che non si raggiunge il numero X di ore.

Vediamo ad esempio come risolvere $12(\text{mod } 7) =$



Dopo il primo giro della lancetta dell'orologio sono trascorse 7 ore, quindi procediamo fino a raggiungere la 12_{ima} ora, che corrisponde al numero 5, che è anche il nostro risultato.



Numeri primi

I numeri primi sono stati oggetto di studio dai matematici di ogni periodo storico: tutti sanno che un numero primo non è rappresentabile come prodotto di interi che lo precedono e si dice primo se è divisibile esattamente solo per 1 e per se stesso.
Ancora oggi il metodi più semplici per trovare tutti i numeri primi risale a qualche millennio fa, cioè al ben noto crivello di Eratostene.

I numeri primi sono stati utilizzati da Euclide che enunciò due teoremi su di essi:

TEOREMI DI EUCLIDE SUI NUMERI PRIMI

Primo Teorema di Euclide: ogni numero intero N si scrive in modo unico (a parte l'ordine) come prodotto di numeri primi.
Secondo Teorema di Euclide: i numeri primi formano una successione infinita.

Anche Eulero li studio ed enunciò un famoso teorema dal quale Fermat arrivo a promulgare il suo famoso piccolo teorema (dimostrato in seguito proprio da Eulero).

Noi non entriamo in particolare nella trattazione dei numeri primi ma ci limitiamo a sottolineare che questi sono alla base della criptologia e si lascia l'approfondimento a chi è interessato allo sviluppo degli algoritmi di cifratura.

Simbologia utilizzata

Prima di proseguire riportiamo la simbologia che viene normalmente utilizzata nei testi di crittografia.

Generalmente plaintext e ciphertext si indicano rispettivamente con le lettere m (come "**m**essaggio") e c (come "**c**odice"); la chiave con il simbolo k ("**k**ey").
La funzione di cifratura viene indicata con il simbolo f, con f^{-1} quella di decifratura (alcuni testi riportano la lettera E (Encrypt)).

Possiamo scrivere quindi la procedura di cifratura con la seguente espressione:

$$c = F_k(m)$$

oppure

$$c = E_k(m)$$

e per la decifratura

$$m = f_k^{-1}(c)$$

o

$$m = E_k^{-1}(c)$$

Nel caso di chiave simmetrica, dato che si utilizza la stessa chiave per la cifratura e la decifratura si può scrivere:

$$m = f_k^{-1}(f_k(m))$$

e quindi:

$$m = E_k^{-1}(E_k(m))$$

Nel resto della nostra trattazione utilizzeremo la seconda notazione, cioè:
- per la cifratura $c = E_k(m)$
- per la decifratura $m = E_k^{-1}(c)$

Verifichiamo le conoscenze

Esercizi a scelta multipla

1 **Quale tra i seguenti non è un aspetto connesso al *problema della sicurezza*:**

a) la segretezza
b) gli errori di trasmissione
c) l'autenticazione
d) l'affidabilità dei documenti

2 **La cifratura si differenzia dalla codifica in quanto:**

a) la cifratura sostituisce alcune parole con altre
b) la cifratura sostituisce lettere o caratteri
c) la codifica sostituisce alcune parole con altre
d) la codifica sostituisce lettere o caratteri

3 **La chiave pubblica è:**

a) usata per la cifratura nello schema crittografico simmetrico
b) usata per la decifratura nello schema crittografico simmetrico
c) usata per la cifratura nello schema crittografico asimmetrico
d) usata per la decifratura nello schema crittografico asimmetrico

4 **La chiave privata è:**

a) usata per la cifratura nello schema crittografico simmetrico
b) usata per la decifratura nello schema crittografico simmetrico
c) usata per la cifratura nello schema crittografico asimmetrico
d) usata per la decifratura nello schema crittografico asimmetrico

5 **La chiave comune è:**

a) usata per la cifratura nello schema crittografico simmetrico
b) usata per la decifratura nello schema crittografico simmetrico
c) usata per la cifratura nello schema crittografico asimmetrico
d) usata per la decifratura nello schema crittografico asimmetrico

6 **Qual è delle seguenti espressioni è errata:**

a) $25 \pmod 4 = 1$
b) $36 \pmod 7 = 1$
c) $41 \pmod{21} = 20$
d) $27 \pmod{27} = 0$
e) $17 \pmod 1 = 0$
f) $625 \pmod 7 = 1$
g) $1000 \pmod 7 = 6$
h) $1296 \pmod 7 = 1$

Test vero/falso

1 Un documento è affidabile se ne conosciamo il mittente. V F
2 A livello datalink è possibile gestire la segretezza. V F
3 A livello datalink è possibile gestire l'affidabilità. V F
4 La crittografia consiste nella scrittura segreta in codice o cifrata. V F
5 Un algoritmo di cifratura prende anche il nome di crittogramma. V F
6 La regola con la quale si effettua la cifratura prende il nome di algoritmo di criptazione. V F
7 Il parametro che viene modificato nell'algoritmo di criptazione prende il nome di chiave. V F
8 Nello schema crittografico simmetrico la chiave di cifratura coincide con quella di decifratura. V F
9 Nello schema crittografico asimmetrico la chiave di cifratura è la chiave privata. V F
10 La stessa chiave pubblica può essere utilizzata da più persone contemporaneamente. V F
11 La chiave pubblica e la chiave privata per lo stesso utente sono tra loro reciproche. V F
12 Una buona segretezza richiede che l'algoritmo di cifratura sia segreto. V F
13 L'azione di un crittoanalista mirata a violare il crittosistema prende il nome di attacco. V F