ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

ΚΡΥΠΤΟΓΡΑΦΙΑ

2η Διάλεξη

Κωνσταντίνου Ελισάβετ

Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων

Άδειες Χρήσης

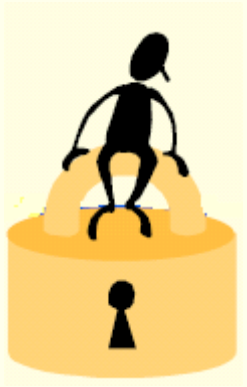
- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αιγαίου**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

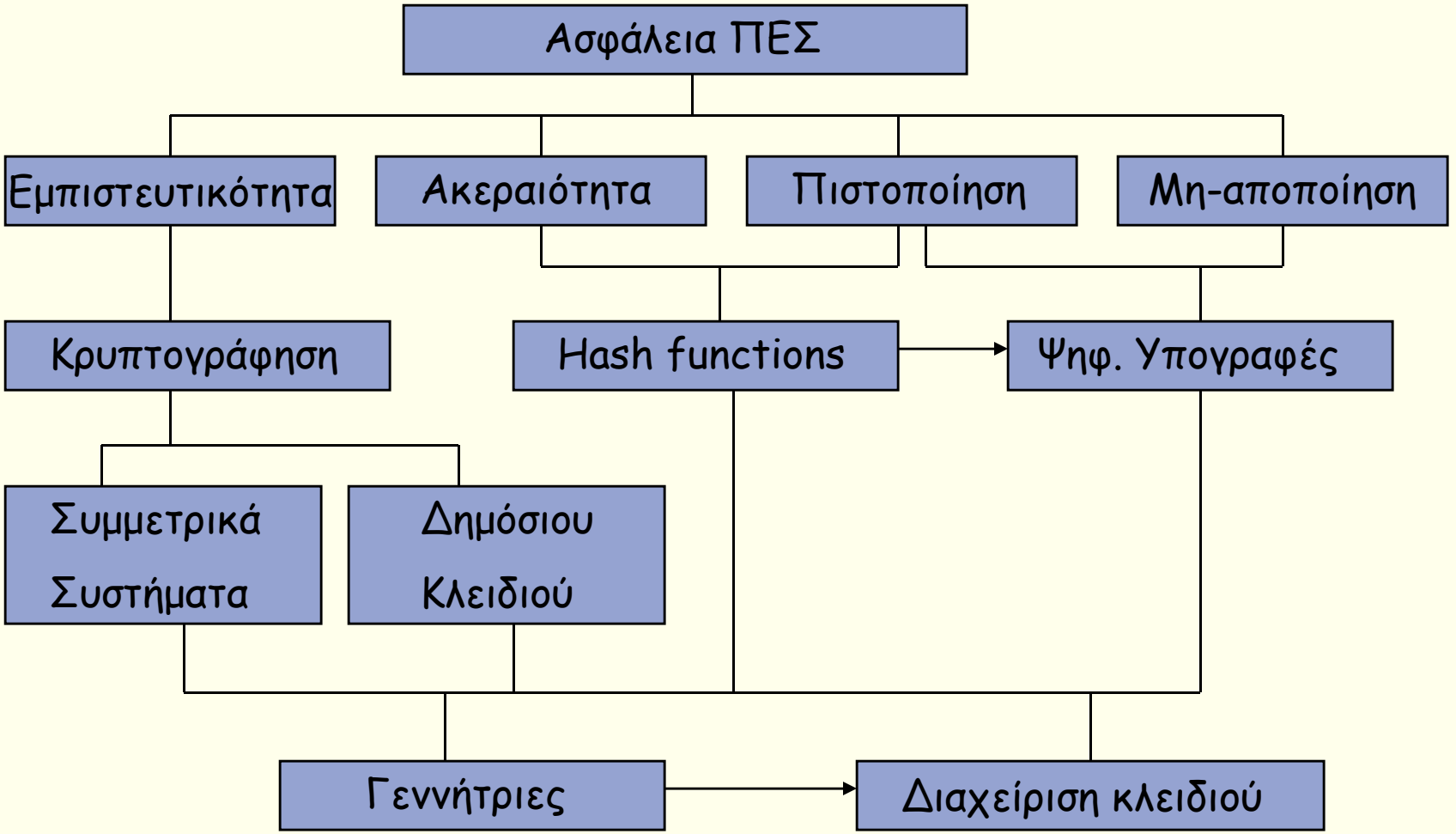
Κρυπτογραφία



Κωνσταντίνου Ελισάβετ
ekonstantinou@aegean.gr

http://www.icsd.aegean.gr/ekonstantinou

Συνολικό Πλαίσιο



Βασικοί Συμβολισμοί

A : Αλφάβητο, αποτελείται από τα σύμβολα με τα οποία δομούνται τα plaintexts.

M : message space, αποτελείται από συμβολοσειρές στοιχείων του A , δηλαδή από όλα τα πιθανά plaintexts.

C: ciphertext space, είναι το σύνολο όλων των πιθανών ciphertexts.

K : key space, δηλαδή κάθε στοιχείο του K αποτελεί και ένα κλειδί.

E_e : κρυπτογράφηση με το κλειδί e

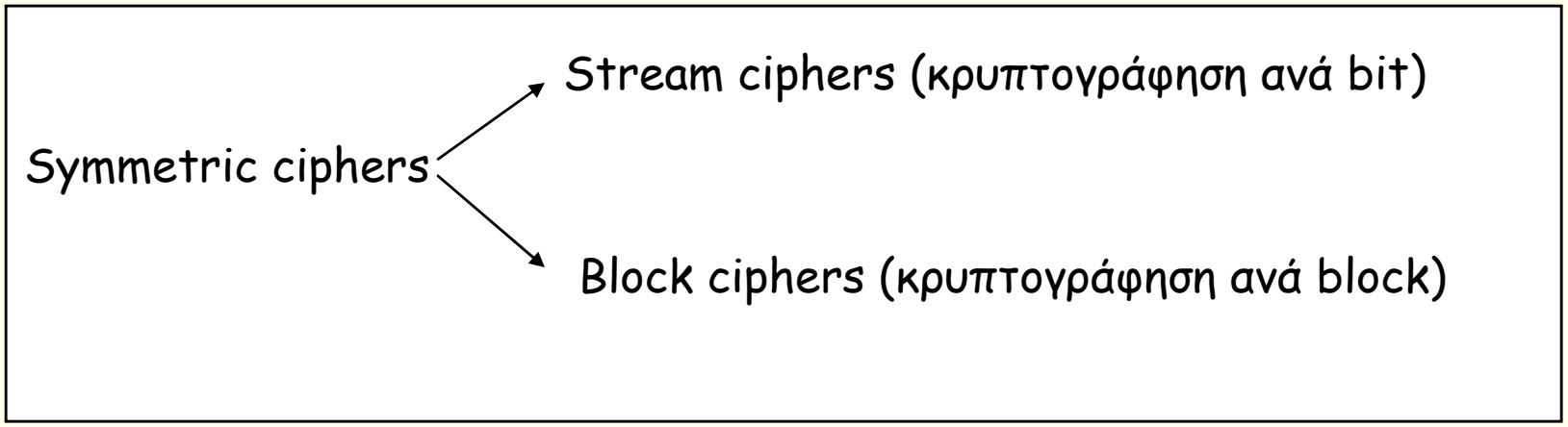
D_d: αποκρυπτογράφηση με το κλειδί d

Συμμετρική Κρυπτογραφία

Από τους αρχαιότερους συμμετρικούς αλγορίθμους, είναι ο αλγόριθμος του Καίσαρα:

$$e = \begin{pmatrix} A & B & C & D & E & \dots & Z \\ D & E & F & G & H & & C \end{pmatrix}$$

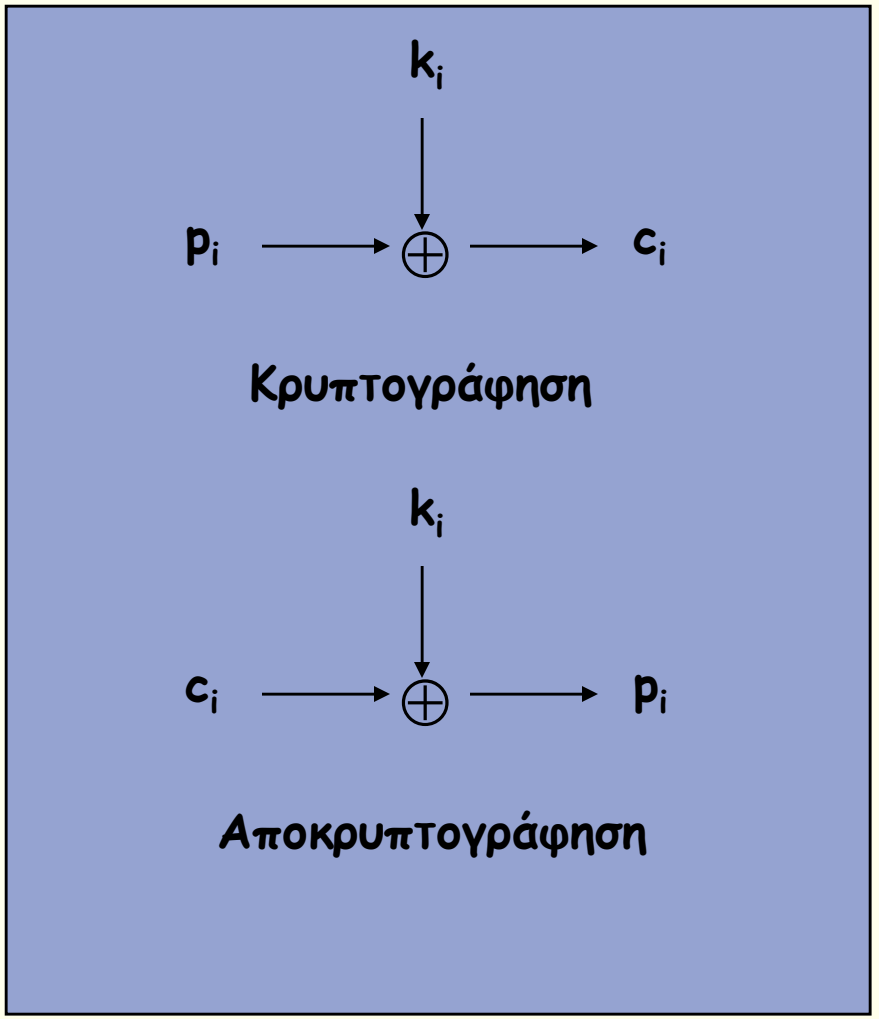
$d = e^{-1}$



Stream ciphers

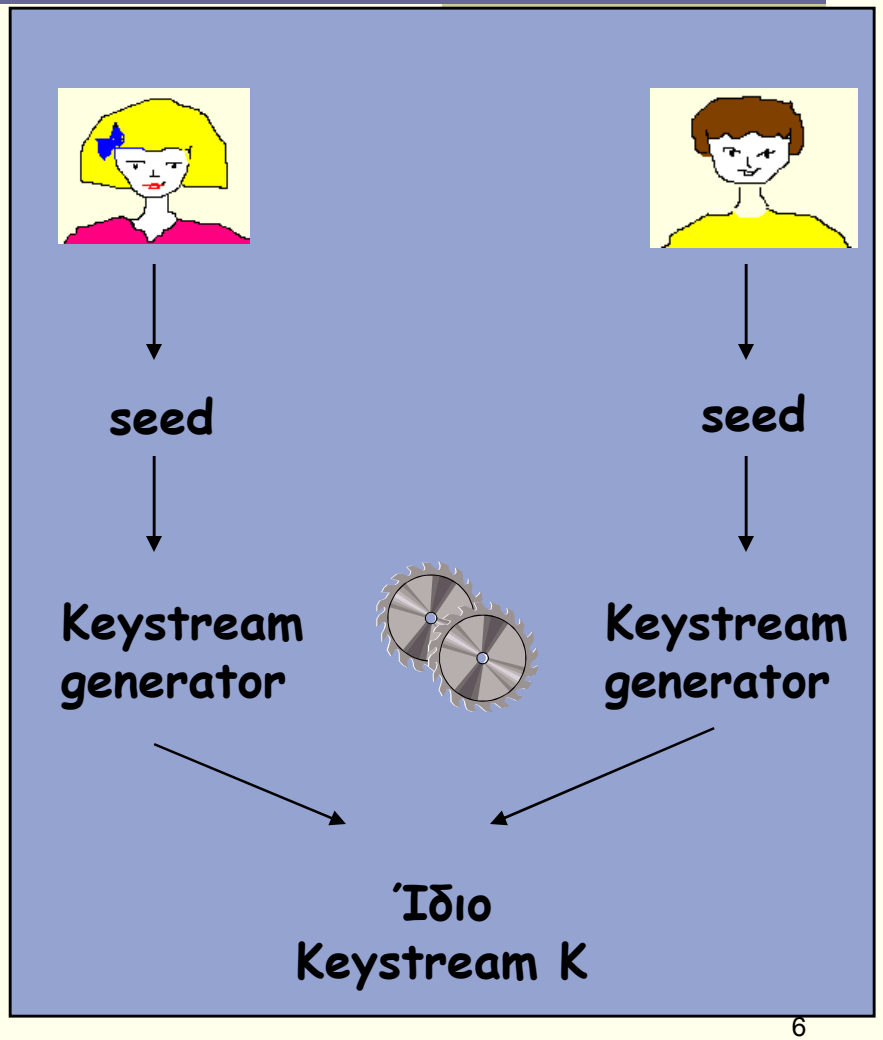
Η διαδικασία κωδικοποίησης για έναν stream cipher συνοψίζεται παρακάτω:

1. Το αρχικό μήνυμα (plaintext) μετατρέπεται σε δυαδική ακολουθία.
2. Επιλέγεται ένα κλειδί κρυπτογράφησης το οποίο μετατρέπεται σε δυαδική ακολουθία επίσης.
3. Το αρχικό μήνυμα και το κλειδί προστίθενται σύμφωνα με την πράξη XOR για να προκύψει το ciphertext.

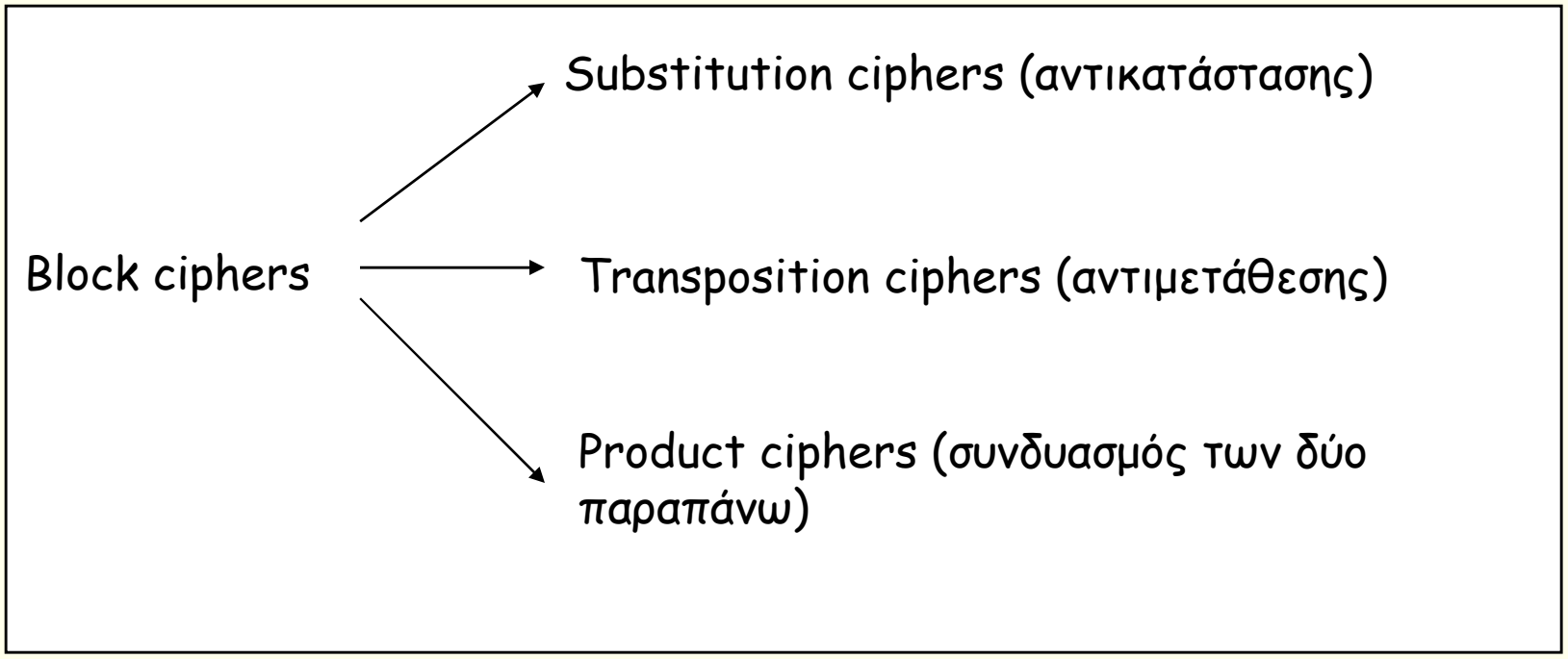


Stream ciphers

- Αν η συμβολοσειρά κλειδιού που χρησιμοποιείται για ένα plaintext δεν ξαναχρησιμοποιείται τότε έχουμε το σύστημα one-time pad.
- Η σειρά των κλειδιών που χρησιμοποιούνται καλείται keystream. Αυτή πρέπει να είναι όσο το δυνατόν περισσότερο τυχαία για να μην μπορεί να προβλεφθεί. Προκύπτει από μια γεννήτρια που καλείται keystream generator.



Block ciphers



Substitution ciphers

Α) Απλοί αλγόριθμοι αντικατάστασης (simple substitution ciphers) ή μονοαλφαβητικοί αλγόριθμοι.

Ορισμός: Έστω A ένα αλφάβητο από q σύμβολα και M (message space) το σύνολο όλων των συμβολοσειρών μήκους t πάνω στο A. Το K (key space) είναι το σύνολο όλων των μεταθέσεων του συνόλου A. Τότε για κάθε κλειδί $e \in K$ ορίζεται κρυπτογράφηση $E_e(m) = (e(m_1) e(m_2) \dots e(m_t)) = (c_1 c_2 \dots c_t) = c$. Η αποκρυπτογράφηση γίνεται με την αντίστροφη μετάθεση $d = e^{-1}$.

Παράδειγμα:

abcdefghijklmnopqrstuvwxyz
sdfghjklopiuytrewqazxcvbnm ⟶ **Τυχαία μετάθεση = κλειδί**

hello ⟶ lhuue

Substitution ciphers

Απλοί μονοαλφαβητικοί αλγόριθμοι -> αλγόριθμος του Καίσαρα

Πόσα κλειδιά υπάρχουν? Πώς σπάει ο αλγόριθμος?

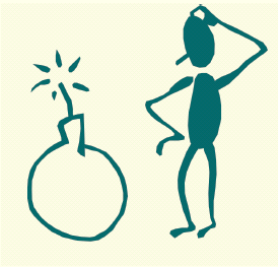
Γενικοί μονοαλφαβητικοί αλγόριθμοι -> οποιαδήποτε μετάθεση των στοιχείων του αλφαβήτου

Αν το αλφάβητο έχει q στοιχεία, πόσα είναι τα πιθανά κλειδιά που υπάρχουν? Πώς σπάει ο αλγόριθμος?

Substitution ciphers

Β) Ομόφωνοι αλγόριθμοι αντικατάστασης (homophonic substitution ciphers).

Ορισμός: Με κάθε σύμβολο α του αλφαβήτου Α , συσχετίζεται ένα σύνολο Η(α) από συμβολοσειρές από t σύμβολα (t είναι το μέγεθος των μηνυμάτων του message space). Όλα τα σύνολα Η(α) είναι ανά δύο ανεξάρτητα. Ο ομόφωνος αλγόριθμος αντικατάστασης αντικαθιστά κάθε σύμβολο α του αρχικού κειμένου με μια τυχαία συμβολοσειρά του Η(α). Το κλειδί αποτελείται από όλα τα σύνολα Η(α).



Substitution ciphers

Παράδειγμα: Έστω A = {a, b}, H(a) = {00, 01} και H(b) = {10, 11}.
Τότε η κρυπτογραφημένη μορφή των αρχικών κειμένων aa, ab, ba και bb μπορεί να είναι μία από τις ακόλουθες τετράδες:

aa ⟶ {0000, 0001, 0100, 0101}
ab ⟶ {0010, 0011, 0110, 0111}
ba ⟶ {1000, 1001, 1100, 1101}
bb ⟶ {1010, 1011, 1110, 1111}

Κάθε συμβολοσειρά των 4-bits αντιστοιχεί σε ένα μόνο plaintext.

Καλύτεροι αλγόριθμοι από τους απλούς αλγορίθμους αντικατάστασης γιατί κάνουν πιο ομοιόμορφη την κατανομή των συμβόλων, με κόστος βέβαια την επέκταση των δεδομένων (data expansion). Επιπλέον η αποκρυπτογράφηση δεν είναι τόσο απλή.

Substitution ciphers

Γ) Πολυαλφαβητικοί αλγόριθμοι αντικατάστασης (polyalphabetic substitution ciphers).

Ορισμός: Ένας πολυαλφαβητικός αλγόριθμος αντικατάστασης είναι ένας block cipher που κρυπτογραφεί block μήκους t χαρακτήρων και έχει τις εξής ιδιότητες:

i) Το key space K αποτελείται από όλα τα σύνολα των t μεταθέσεων $(p_1, p_2, \ldots, p_t)$ όπου κάθε μετάθεση p_i ορίζεται στο σύνολο A.

ii) Η κρυπτογράφηση ενός μηνύματος $m = (m_1 m_2 \ldots m_t)$ με κλειδί το $e = (p_1, p_2, \ldots, p_t)$ είναι η $E_e(m) = (p_1(m_1)\ p_2(m_2) \ldots p_t(m_t))$.

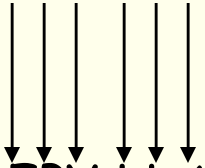
iii) Η αποκρυπτογράφηση γίνεται με το κλειδί $d = (p_1^{-1}, p_2^{-1}, \ldots, p_t^{-1})$.

Substitution ciphers

Παράδειγμα (Vigenere cipher):

Έστω A = {A, B, …, Z}, t = 3 και e = (p_1, p_2, p_3), όπου το p_1 μετατοπίζει κάθε γράμμα 3 θέσεις δεξιά, το p_2 7 θέσεις και το p_3 10 θέσεις.

Αν m = CIP HER τότε



c = FPY LLA

Οι αλγόριθμοι αυτοί είναι πιο ασφαλείς από τους μονοαλφαβητικούς, αφού τώρα το ίδιο γράμμα κωδικοποιείται σε διαφορετικό ανάλογα με την θέση του στο block. Παρόλα αυτά, και οι αλγόριθμοι αυτοί σπάνε εύκολα (κάνοντας ανάλυση συχνότητας για κάθε ομάδα p_i).

Transposition ciphers

Απλά αλλάζουν τη σειρά των χαρακτήρων στο block.

Ορισμός: Έστω ότι το μήκος του block είναι t και K είναι το σύνολο των μεταθέσεων στο {1, 2, …, t}. Για κάθε κλειδί e η κρυπτογράφηση γίνεται ως εξής:

$E_e(m) = (m_{e(1)}, m_{e(2)}, \ldots, m_{e(t)})$.

Παράδειγμα: Έστω t = 3 και e = (3, 1, 2). Τότε το κείμενο abc κρυπτογραφείται στο cab.

Σπάει πολύ εύκολα, αν όμως χρησιμοποιηθούν αρκετοί μαζί στη σειρά η δυσκολία αυξάνει δραματικά.

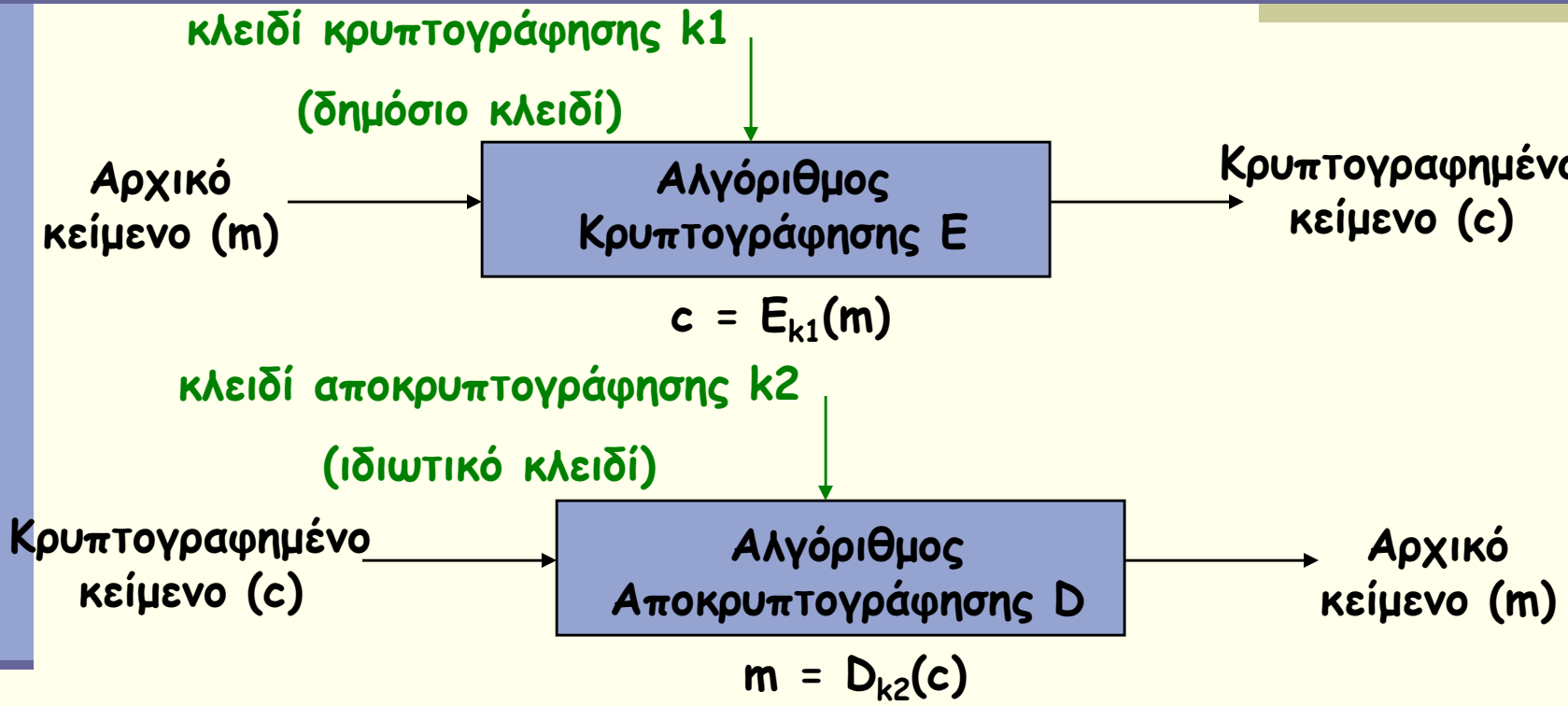
Product ciphers

Οι απλοί αλγόριθμοι αντικατάστασης και αντιμετάθεσης δεν είναι επαρκώς ασφαλείς. Ωστόσο, συνδυάζοντας τους δύο τύπους αλγορίθμων μπορούμε να φτιάξουμε πολύ ισχυρούς block ciphers.

Ουσιαστικά ένας Product cipher είναι μια σύνθεση $t > 1$ μετασχηματισμών $E_1E_2...E_t$, όπου κάθε E_i είναι ένας αλγόριθμος αντικατάστασης ή αντιμετάθεσης.

Round (γύρος) καλείται η σύνθεση ενός αλγορίθμου αντικατάστασης με έναν αλγόριθμο αντιμετάθεσης. Οι περισσότεροι σύγχρονοι block ciphers είναι product ciphers και χρησιμοποιούν ένα πλήθος από rounds για να κρυπτογραφήσουν το αρχικό κείμενο.

Ασύμμετρα Κρυπτοσυστήματα



Το κλειδί αποκρυπτογράφησης βρίσκεται δύσκολα από το αντίστοιχο κλειδί κρυπτογράφησης. Η δυσκολία βασίζεται σε κάποιο μαθηματικό πρόβλημα.

Μονόδρομες Συναρτήσεις

Ορισμός: Μια συνάρτηση f: X -> Y καλείται μονόδρομη (one-way) αν η τιμή f(x) είναι «εύκολο» να υπολογιστεί για όλα τα $x \in X$, αλλά για τυχαία στοιχεία $y \in Y$ είναι υπολογιστικά αδύνατο (computationally infeasible) να βρεις ένα x τέτοιο ώστε

f(x) = y.

Παραδείγματα:

1) $f(x) = 3^x \bmod 17$
2) $f(x) = x^3 \bmod n$, όπου n = pq με p, q πρώτους αριθμούς. Αν οι δύο πρώτοι αριθμοί δεν είναι γνωστοί, τότε η συνάρτηση είναι δύσκολο να αντιστραφεί, διαφορετικά υπάρχει αποδοτικός αλγόριθμος που βρίσκει κυβικές ρίζες.

Μονοδρομές Συναρτήσεις με Καταπακτή

Ορισμός: Μια συνάρτηση f: X -> Y καλείται μονόδρομη συνάρτηση με καταπακτή (trapdoor one-way function) αν είναι μονόδρομη και έχει επιπλέον την ιδιότητα ότι δοθείσης κάποιας πληροφορίας (trapdoor information) είναι «εύκολο» να βρεις για κάθε τιμή y, το x για το οποίο f(x) = y.

Για παράδειγμα, η $f(x) = x^3 \bmod n$ είναι μια trapdoor one-way function, όπου η trapdoor πληροφορία είναι οι παράγοντες p και q του n. Αν p και q είναι μεγάλοι πρώτοι αριθμοί, τότε είναι υπολογιστικά δύσκολο να τους βρεις γνωρίζοντας μόνο το n.

⟶ Integer Factorization Problem

Τι είναι εύκολο ή δύσκολο υπολογιστικά?

Για δύο συναρτήσεις f και g λέμε ότι $f(n) = O(g(n))$ αν υπάρχει σταθερά c και ένας θετικός ακέραιος n_0 τέτοιος ώστε

$0 \le f(n) \le cg(n)$ **για όλα τα** $n \ge n_0$.

Ένας αλγόριθμος καλείται πολυωνυμικός αν ο χρόνος χειρότερης περίπτωσης του είναι της μορφής $O(n^k)$, όπου n είναι το μέγεθος της εισόδου και k είναι μια σταθερά. Κάθε αλγόριθμος του οποίου ο χρόνος εκτέλεσης δεν μπορεί να φραγεί από ένα τέτοιο όριο, καλείται εκθετικός.

Εύκολο υπολογιστικά → πολυωνυμική πολυπλοκότητα

Δύσκολο υπολογιστικά → εκθετική πολυπλοκότητα

Κρυπτογραφία Δημόσιου Κλειδιού

Δοθέντος ενός δημόσιου κλειδιού e είναι υπολογιστικά αδύνατο να βρεθεί το ιδιωτικό κλειδί d.

Αλγόριθμος κρυπτογράφησης E_e = trapdoor one-way function με trapdoor information = d.

Π.χ. η $f(x) = x^3 \bmod n$ είναι μια trapdoor one-way function, όπου η trapdoor πληροφορία είναι οι παράγοντες p και q του n. Άρα, αν το μυστικό κλειδί του Bob είναι τα p, q και δημόσιο το n, τότε η Alice μπορεί να κρυπτογραφήσει το μήνυμα x ως $c = x^3 \bmod n$ και να το στείλει πίσω στον Bob. Στη συνέχεια, μόνο ο Bob που έχει την trapdoor πληροφορία μπορεί να βρει το x (δηλαδή να αντιστρέψει τη συνάρτηση).

Κρυπτογραφία Δημόσιου Κλειδιού

$c = x^3 \bmod n$

(s_A, P_A)

x = μήνυμα

(s_B, P_B)

$s_B = p, q$
$P_B = n$

Μόνο ο Bob μπορεί να αποκρυπτογραφήσει που κατέχει τα p, q. Αν γνωρίζεις το n (δημόσιο κλειδί), δεν μπορείς να βρεις τα p, q (ιδιωτικό κλειδί).

Integer factorization problem!!

Ψηφιακές Υπογραφές + Κρυπτογράφηση Δημόσιου Κλειδιού

Υποθέστε ότι M = C (message space = ciphertext space).

Επειδή $D_d(E_e(m)) = E_e(D_d(m)) = m$ μπορεί να προκύψει το εξής σχήμα ψηφιακών υπογραφών:

1. Signer ⟶ Υπολογίζει $s = D_d(m)$ και στέλνει (s, m) στον Verifier (d είναι το ιδιωτικό κλειδί του Signer).
2. Verifier ⟶ Αν $E_e(s) = m$ τότε επιστρέφει 1 (true), διαφορετικά 0 (false).

Είναι δυνατόν να μην σταλεί το μήνυμα m?

Ψηφιακές Υπογραφές με Ανάκτηση Μηνύματος

Ο Signer πρέπει να υπογράφει μηνύματα που έχουν μια συγκεκριμένη μορφή.

Π.χ. αν το M (message space) αποτελείται από όλες τις συμβολοσειρές των 8-bits, τότε ο υπογράφων μπορεί να υπογράφει 8άδες στις οποίες τα 4 πρώτα bits είναι ίδια με τα 4 τελευταία.

Άρα αν ο Verifier διαπιστώσει ότι το αποτέλεσμα $E_e(s)$ έχει την επιθυμητή μορφή, δέχεται την υπογραφή.

Μια απλή επίθεση...

Η προηγούμενη τροποποίηση μπορεί να προστατεύσει από την εξής επίθεση (γιατί??):

(d, e)

Επιλέγει τυχαία ένα s ως υπογραφή, υπολογίζει u = $E_e(s)$

(s, u)

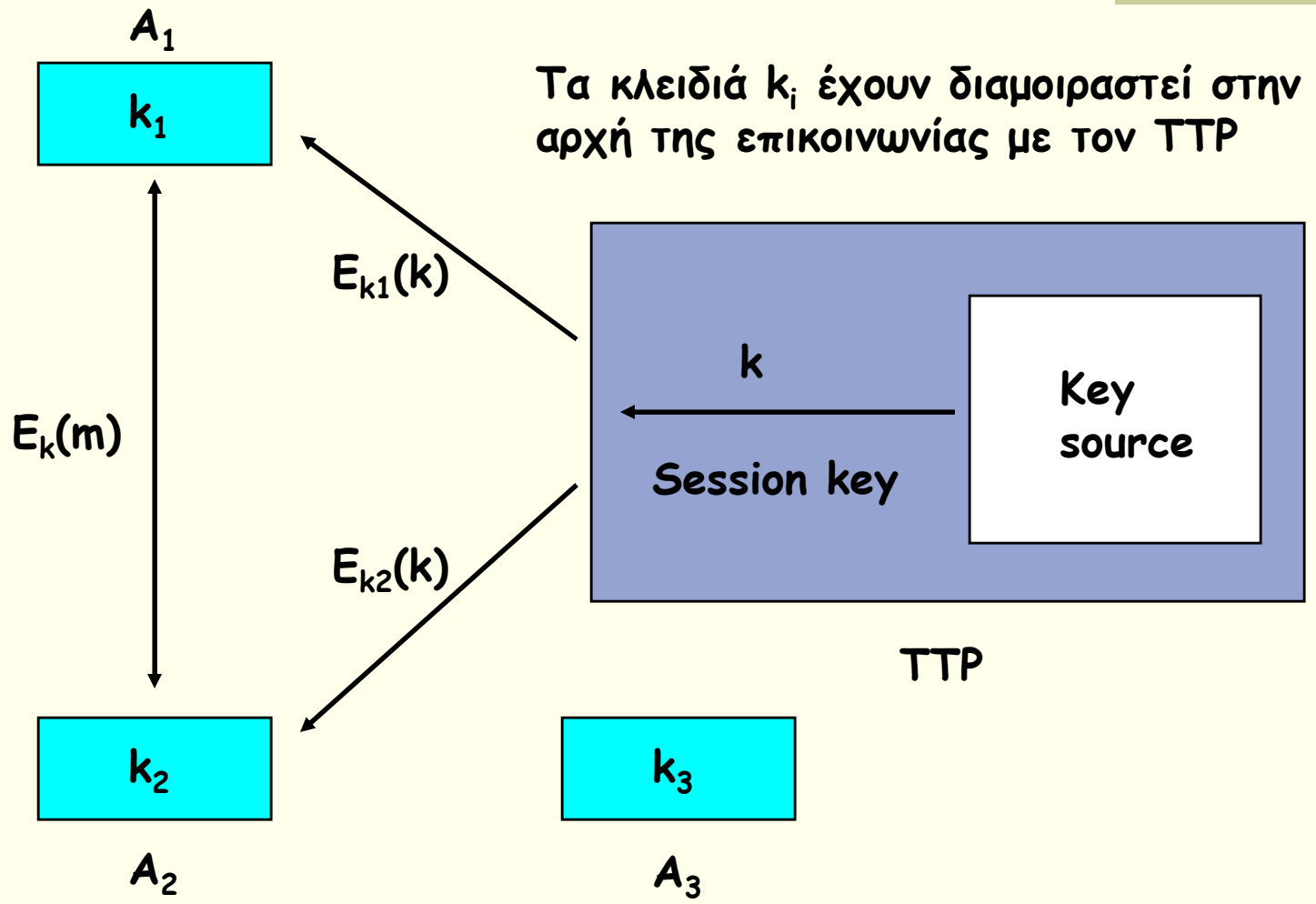
δηλαδή s = υπογραφή

u = μήνυμα

Επαληθεύει ότι

u = $E_e(s)$!!!

Διαχείριση κλειδιών μέσω συμμετρικών τεχνικών



Διαχείριση κλειδιών μέσω συμμετρικών τεχνικών

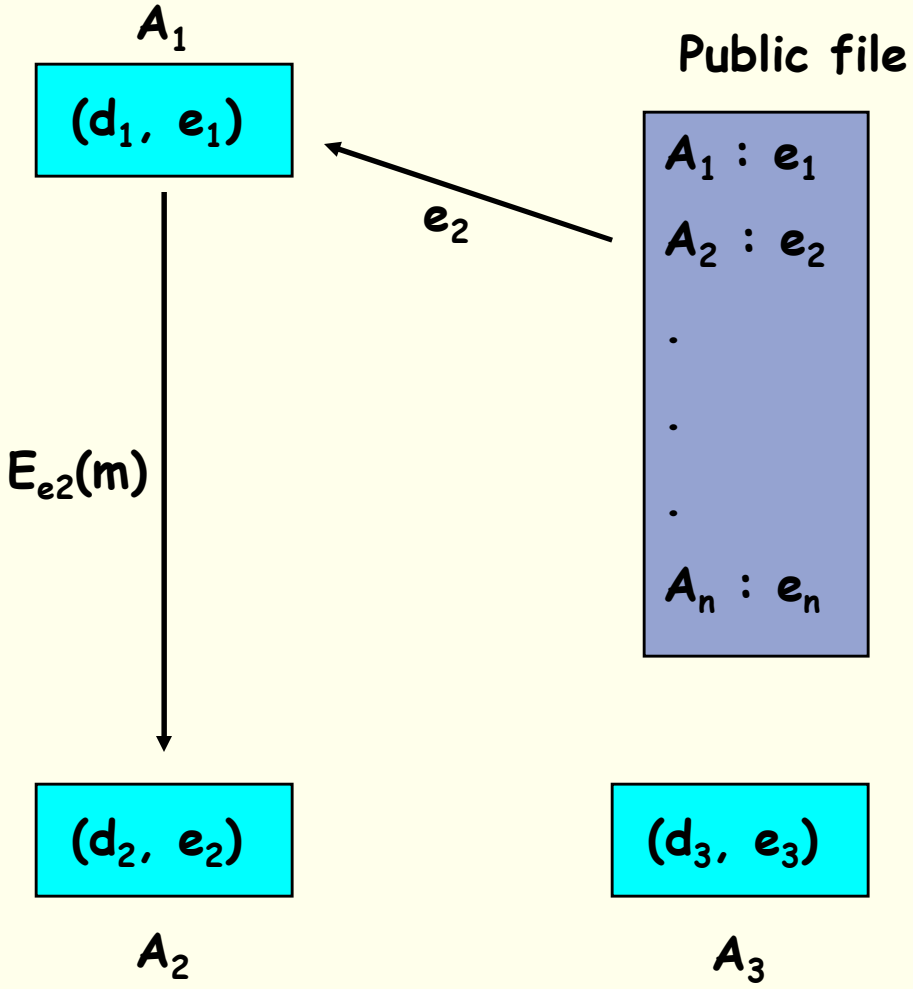
Πλεονεκτήματα:

1) Πολύ εύκολο να διαγραφεί ή να προστεθεί κάποιος στο δίκτυο.
2) Κάθε χρήστης αποθηκεύει μόνο ένα κλειδί.

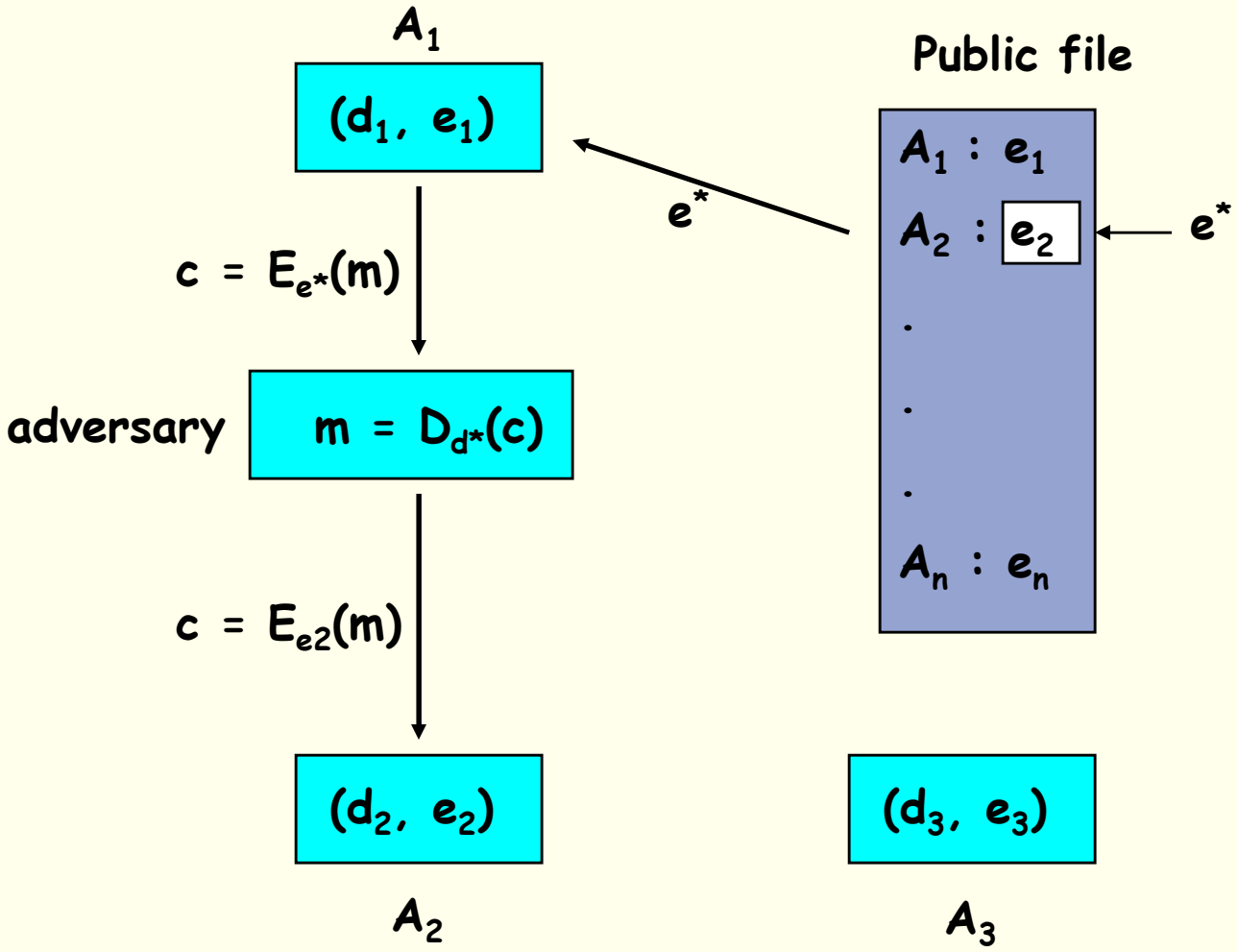
Μειονεκτήματα:

1) Ο TTP πρέπει να συμμετέχει σε όλες τις επικοινωνίες.
2) Ο TTP αποθηκεύει n κλειδιά.
3) Ο TTP μπορεί να διαβάσει όλα τα μηνύματα.
4) Αν ο TTP βρεθεί υπό τον έλεγχο ενός κακόβουλου χρήστη, η ασφάλεια όλου του δικτύου καταρρέει.

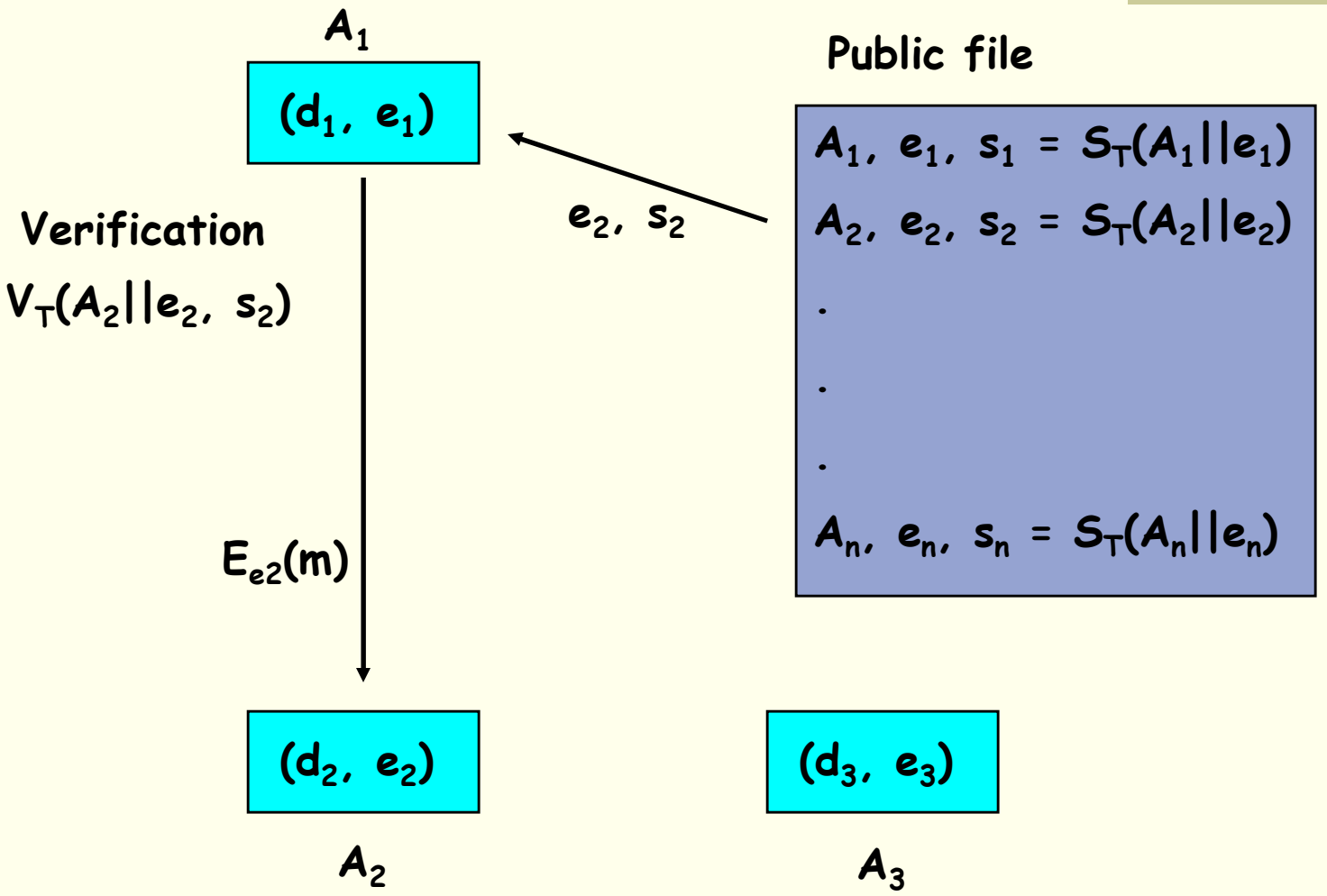
Διαχείριση κλειδιών μέσω ασύμμετρων τεχνικών



Διαχείριση κλειδιών μέσω ασύμμετρων τεχνικών (επίθεση)



Διαχείριση κλειδιών μέσω ασύμμετρων τεχνικών (αντιμετώπιση)



Διαχείριση κλειδιών μέσω ασύμμετρων τεχνικών με TTP

Πλεονεκτήματα:

1) Αποφεύγονται οι ενεργές επιθέσεις.
2) Ο TTP δεν μπορεί να παρακολουθήσει τις επικοινωνίες μεταξύ των χρηστών του δικτύου.

Μειονεκτήματα:

1) Αν βρεθεί το κλειδί υπογραφής του TTP το σύστημα καταρρέει

Διάβασμα….

Κεφάλαιο 1 του Handbook of Applied Cryptography