ΔΗΜΗΤΡΙΟΣ Μ. ΠΟΥΛΑΚΗΣ

ΚΡΥΠΤΟΓΡΑΦΙΑ

Η ΕΠΙΣΤΗΜΗ ΤΗΣ ΑΣΦΑΛΟΥΣ ΕΠΙΚΟΙΝΩΝΙΑΣ

ΕΚΔΟΣΕΙΣ ΖΗΤΗ
ΘΕΣΣΑΛΟΝΙΚΗ

Περιεχόμενα

Πρόλογος

Η κρυπτογραφία, πριν μόλις μία τριακονταετία, είχε ως μοναδικά πεδία εφαρμογής της τον στρατό και την διπλωματία. Οι τεχνικές οι οποίες χρησιμοποιούσε ήταν αρκετά απλές. Σήμερα, το πεδίο εφαρμογής της έχει επεκταθεί σημαντικά και περιλαμβάνει όλους τους τομείς στους οποίους η ασφαλής μετάδοση μηνυμάτων παίζει πρωτεύοντα ρόλο. Η ανάπτυξη της κοινωνίας της πληροφορίας και των δικτύων επικοινωνίας δημιούργησαν καινούργιες ανάγκες τις οποίες η κρυπτογραφία καλείται να καλύψει. Η σύγχρονη κρυπτογραφία δεν ασχολείται μόνο με την ασφαλή μετάδοση μηνυμάτων αλλά και με άλλα θέματα εξ ίσου σημαντικά, όπως η ακεραιότητα, αυθεντικότητα και αδυναμία αποκήρυξης των μηνυμάτων. Η ανάπτυξη που γνωρίζει σήμερα η κρυπτογραφία οφείλεται κατά ένα μεγάλο μέρος στη χρήση της Θεωρίας Αριθμών, της Άλγεβρας και της Θεωρίας Αλγορίθμων. Ειδικότερα, οι πρώτοι αριθμοί, τα πεπερασμένα σώματα και οι ελλειπτικές καμπύλες βρήκαν ένα πολύ ενδιαφέρον πεδίο εφαρμογής τους. Επίσης, η μεγάλη αύξηση των δυνατοτήτων των ηλεκτρονικών υπολογιστών έπαιξε καθοριστικό ρόλο. Χωρίς αυτή, η αποτελεσματικότητα πολλών αλγορίθμων ή πρωτοκόλλων δεν θα ήταν δυνατόν να δοκιμαστεί.

Το παρόν σύγγραμμα απευθύνεται κυρίως σε φοιτητές και αποφοίτους τμημάτων Μαθηματικών και Πληροφορικής, αλλά και σ΄ οποιονδήποτε ενδιαφέρεται για την ασφαλή επικοινωνία. Βασίζεται κατά μεγάλο μέρος σ΄ ένα μάθημα που δίδαξα το χειμερινό εξάμηνο του πανεπιστημιακού έτος 2003-4 στη κατεύθυνση "Θεωρητική Πληροφορική και Θεωρία Συστημάτων" του Μεταπτυχιακού Προγράμματος του Τμήματος Μαθηματικών του ΑΠΘ. Απαραίτητες γνώσεις για την κατανόησή του είναι η βασική Θεωρία Αριθμών και η βασική Άλγεβρα. Περιέχει 9 κεφάλαια. Στο πρώτο κεφάλαιο περιγράφονται μερικά απλά κλασσικά κρυπτοσυστήματα έτσι, ώστε ο αναγνώστης να εξοικειωθεί με τις πρώτες έννοιες της κρυπτογραφίας. Το δεύτερο κεφαλαιο είναι αφιερωμένο στην ασφά-

λεια ενός κρυπτοσυστήματος κατά τον *C. Shannon* και στους καταγραφείς μετατόπισης με ανάδραση. Στο τρίτο κεφάλαιο παρουσιάζονται βασικά θέματα από την Υπολογιστική Θεωρία Αριθμών. Το αντικείμενο του τέταρτου κεφαλαίου είναι το διάσημο πλέον κρυπτοσύστημα *RSA* και το κρυπτοσύστημα του *Rabin*. Στο πέμπτο κεφάλαιο περιγράφονται μερικές κλασσικές μέθοδοι πιστοποίησης πρώτου και δίνεται μία εκδοχή του αλγόριθμου *AKS*. Το έκτο κεφάλαιο ασχολείται με την παραγοντοποίηση των ακεραίων. Το έβδομο κεφάλαιο είναι αφιερωμένο σε κρυπτοσυστήματα που βασίζονται στο πρόβλημα του διακριτού λογάριθμου και παρουσιάζονται μερικοί αλγόριθμοι για την επίλυσή του. Οι ψηφιακές υπογραφές είναι το θέμα του όγδοου κεφαλαίου. Τέλος, στο ένατο κεφάλαιο περιγράφονται μερικά κρυπτογραφικά πρωτόκολλα. Επίσης, για την ευκολία του αναγνώστη, μερικές στοιχειώδεις έννοιες από την Θεωρία των Πιθανοτήτων περιλαμβάνονται σ΄ ένα παράρτημα.

Θεσσαλονίκη 2004 Δημήτριος Πουλάκης

Συμβολισμοί - Ορολογία

Υποθέτουμε ότι ο αναγνώστης μας είναι εξοικειωμένος με την βασική Θεωρία Αριθμών και την βασική Άλγεβρα. Θα χρησιμοποιούμε τα συνήθη σύμβολα της Θεωρίας Συνόλων: $\in$, $\subseteq$, $\subset$, $\cap$ και $\cup$. Αν X και Y είναι υποσύνολα του ίδιου συνόλου, τότε συμβολίζουμε με $X - Y$ το σύνολο των στοιχείων του X που δεν ανήκουν στο Y. Ας είναι $f : A \to B$ μία απεικόνιση. Η f καλείται ένεση, αν για κάθε $x, y \in A$ με $x \neq y$ έχουμε $f(x) \neq f(y)$ και έφεση αν για κάθε $z \in B$ υπάρχει $x \in A$ με $f(x) = z$. Επίσης, η f καλείται αμφίεση, αν είναι ένεση και έφεση. Θα συμβολίζουμε με $\mathbb{N}$ το σύνολο των φυσικών αριθμών $\{0, 1, 2, \ldots\}$, με $\mathbb{Z}$ το σύνολο των ακεραίων αριθμών και με $\mathbb{Q}$ και $\mathbb{R}$ τα σύνολα των ρητών και πραγματικών αριθμών, αντίστοιχα.

Αν a και b είναι ακέραιοι, τότε θα συμβολίζουμε με (a, b) τον μέγιστο κοινό τους διαιρέτη (μ.κ.δ.). Ας είναι n ένας θετικός ακέραιος. Συμβολίζουμε με $\mathbb{Z}_n$ τον δακτύλιο των κλασεων των ακεραίων modulo n και με $\mathbb{Z}_n^*$ την ομάδα των αντιστρεψίμων στοιχείων του $\mathbb{Z}_n$. Ας είναι a ένας ακέραιος. Συμβολίζουμε με $\bar{a}$ την κλάση του a modulo n. Αν ισχύει $(a, n) = 1$, τότε θα γράφουμε $\mathrm{ord}_n(a)$ για την τάξη του a modulo n. Με ϕ συμβολίζουμε, ως συνήθως, την συνάρτηση του Euler. Στη περίπτωση όπου $\mathrm{ord}_n(a) = \phi(n)$, ο a καλείται αρχική ρίζα modulo n. Για περισσότερες πληροφορίες σχετικά με τα αποτελέσματα της Θεωρίας Αριθμών που θα χρησιμοποιήσουμε ο αναγνώστης μπορεί να ανατρέξει στο βιβλίο μας "Θεωρία Αριθμών" ενώ για την Άλγεβρα στο βιβλίο "Εισαγωγή στη Σύγχρονη Άλγεβρα" του καθηγητή Σ. Μποζαπαλίδη (βλέπε Βιβλιογραφία).

Κεφάλαιο 1

Κλασσική Κρυπτογραφία

1.1 Κρυπτογράφηση

Ας υποθέσουμε ότι δύο άτομα, ο A και ο B, επικοινωνούν δια μέσου ενός τυχόντος διαύλου επικοινωνίας, π.χ. συνηθισμένο ταχυδρομείο, ηλεκτρονικό ταχυδρομείο. Λέμε ότι η επικοινωνία των A και B έχει *ασφαλή μετάδοση της πληροφορίας* αν ικανοποιεί τις εξής ιδιότητες:

1. *Εμπιστευτικότητα*: Κανένας τρίτος να μη μπορεί να λάβει γνώση των μηνυμάτων ή μέρους των που ανταλλάσσονται μεταξύ του A και του B.

2. *Ακεραιότητα*: Κανένας τρίτος να μην μπορεί να τροποποιήσει τα μηνύματα που ανταλλάσσονται μεταξύ του A και του B.

3. *Αυθεντικότητα*: Καθένας από τους A και B πρέπει να είναι σίγουρος για την ταυτότητα του άλλου, την προέλευση των μηνυμάτων, την ημερομηνία τους και τα λοιπά δεδομένα του μηνύματος.

4. *Αδυναμία αποκήρυξης*: Αδυναμία άρνησης των A και B της αποστολής ή υπογραφής κάποιου προηγουμένου μηνύματός των.

Εφαρμογές που είναι απαραίτητη η ασφαλής μετάδοση της πληροφορίας είναι οι βάσεις δεδομένων, εμπιστευτικές βιομηχανικές πληροφορίες, ηλεκτρονικό ταχυδρομείο, μετάδοση στρατιωτικών απορρήτων κ.λπ.

Καλούμε *κρυπτογραφία* την μελέτη των μαθηματικών μεθόδων που εξασφαλίζουν την ασφαλή μετάδοση της πληροφορίας δια μέσου μη ασφαλών διαύλων επικοινωνίας.

Ένα *κρυπτοσύστημα* είναι μία πεντάδα (P, C, K, E, D), όπου P, C, K είναι πεπερασμένα σύνολα που καλούνται, αντίστοιχα, *χώρος των καθαρών κειμένων*, *χώρος των κρυπτογραφημένων κειμένων* και *χώρος των κλειδιών*. Τα σύνολα E και D είναι οικογένειες συναρτήσεων της μορφής $E_k : P \to C$ και $D_k : C \to P$, $k \in K$, αντίστοιχα. Οι E_k και D_k καλούνται *συνάρτηση κρυπτογράφησης* και *συνάρτηση αποκρυπτογράφησης* που αντιστοιχουν στο κλειδί k. Επίσης, για κάθε $e \in K$ υπάρχει $d \in K$ έτσι, ώστε για κάθε $p \in P$ να ισχύει $D_d(E_e(p)) = p$. Οπότε, η συνάρτηση E_e είναι ένεση. Σε κάθε ζεύγος (e, d) που ικανοποιεί αυτήν την σχέση, το e καλείται *κλειδί κρυπτογράφησης* και το d *κλειδί αποκρυπτογράφησης*.

Αν ο A επιθυμεί να στείλει ένα κρυπτογραφημένο μήνυμα στον B, τότε χρησιμοποιεί ένα κλειδί κρυπτογράφησης e και ο B χρησιμοποιεί το αντίστοιχο κλειδί αποκρυπτογράφησης d για να το διαβάσει. Ένα κρυπτοσύστημα καλείται *συμμετρικό*, αν το κλειδί κρυπτογράφησης e είναι το ίδιο με το κλειδί αποκρυπτογράφησης d, ή πιο γενικά αν το d είναι δυνατόν να υπολογιστεί πολύ εύκολα από το e. Αν δύο χρήστες επιθυμούν να χρησιμοποιήσουν ένα συμμετρικό σύστημα, τότε θα πρέπει ν' ανταλλάξουν το κλειδί e πριν ν' αρχίσουν την επικοινωνία. Η ασφαλής ανταλλαγή του κλειδιού αποτελεί βασικό πρόβλημα της ασφάλειας του συστήματος και φυσικά το κλειδί θα πρέπει να κρατείται μυστικό.

Ένα κρυπτοσύστημα καλείται *ασύμμετρο*, αν ο υπολογισμός του d από το e είναι πρακτικά αδύνατος. Σε τέτοια κρυπτοσυστήματα το κλειδί κρυπτογράφησης e είναι δυνατόν να δημοσιοποιήται. Αν ένας χρήστης επιθυμεί να λαμβάνει κρυπτογραφημένα μηνύματα δια μέσου ενός τέτοιου συστήματος, τότε καταχωρεί το κλειδί e σ' ένα δημόσιο κατάλογο στον οποία τα κλειδιά είναι καλά προστατευμένα, ώστε να μένουν ανέπαφα. Το αντίστοιχο κλειδί αποκρυπτογράφησης d κρατείται κρυφό. Σ' αυτή την περίπτωση, οποιοσδήποτε, χρησιμοποιώντας το e, μπορεί να στείλει κρυπτογραφημένο μήνυμα στον εν λόγω χρήστη ο οποίος το αποκρυπτογραφεί με την χρήση του d το οποίο μόνο αυτός γνωρίζει. Τα ασύμμετρα κρυπτοσυστήματα καλούνται επίσης *κρυπτοσυστήματα δημόσιου κλειδιού*.

Ένα βασικό μειονέκτημα των κρυπτοσυστημάτων δημόσιου κλειδιού είναι ότι η κρυπτογράφηση μεγάλης ποσότητας δεδομένων απαιτεί πολύ περισσότερο χρόνο από ότι στα συμμετρικά κρυπτοσυστήματα. Η ταχύτητα κρυπτογράφησης είναι ικανοποιητική μόνο στην περίπτωση μικρής ποσότητας δεδομένων. Γι' αυτό τον λόγο, στην πράξη, πολλές φορές χρησιμοποιείται ένα συμμετρικό κρυπτοσύστημα για την κρυπτο-

γράφηση ενός μηνύματος και ένα κρυπτοσύστημα δημοσίου κλειδιού για την αποστολή του κλειδιού που θα χρησιμοποιηθεί.

Για την γραφή μηνυμάτων χρησιμοποιούμε στοιχεία από ορισμένα σύνολα που καλούνται *αλφάβητα*. Ένα *αλφάβητο* Σ είναι ένα μη κενό πεπερασμένο σύνολο. Τα στοιχεία του καλούνται *σύμβολα* ή *γράμματα*. Συνήθη παραδείγματα αλφαβήτων είναι τα αλφάβητα των φυσικών γλωσσών. Άλλα παραδείγματα αλφαβήτων αποτελούν τα σύνολα $\mathbb{Z}_n$. Συχνά ταυτίζουμε τα γράμματα ενός αλφαβήτου n γραμμάτων με τα στοιχεία του συνόλου $\mathbb{Z}_n$. Καλούμε *λέξη* με γράμματα από το Σ κάθε πεπερασμένη ακολουθία γραμμάτων από το Σ. Η κενή ακολουθία συμβολίζεται με ϵ και καλείται *κενή λέξη*.

1.2 Κρυπτανάλυση

Καλούμε *κρυπτανάλυση* την μελέτη των μαθηματικών μεθόδων οι οποίες επιχειρούν την αναίρεση της ασφαλούς μετάδοσης της πληροφορίας. Η κρυπτανάλυση βασίζεται εκτός των μαθηματικών και στο γεγονός ότι ο κρυπταναλυτής κατέχει ικανό πλήθος κρυπτογραφημένων κειμένων που κρυπτογραφήθηκαν με τον ίδιο τρόπο. Καλούμε *κρυπτολογία* την μελέτη της κρυπτογραφίας και της κρυπτανάλυσης.

Ο τρόπος σχεδίασης ενός σύγχρονου κρυπτοσυστήματος βασίζεται στην εξής αρχή:

Αρχή του Kerckhoff. *Η ασφάλεια ενός κρυπτοσυστήματος δεν πρέπει να εξαρτάται από την μυστική διαφύλαξη της μεθόδου κρυπτογράφησης, αλλά μόνο από την μυστική διαφύλαξη του κλειδιού.*

Επομένως, υποθέτουμε πάντα ότι ο κρυπταναλυτής γνωρίζει τον τύπο του κρυπτοσυστήματος που θέλει να προσβάλει. Οι πλέον συνήθεις τύποι προσβολής ενός κρυπτοσυστήματος είναι οι εξής:

1. *Προσβολή κρυπτογραφημένου κειμένου.* Ο κρυπταναλυτής έχει στη κατοχή του ένα μέρος ή όλο το κρυπτογραφημένο μήνυμα.

2. *Προσβολή γνωστού καθαρού κειμένου.* Ο κρυπταναλυτής έχει στη κατοχή του ένα μέρος του μηνύματος που στάλθηκε και το αντίστοιχο κρυπτογραφημένο.

3. *Προσβολή επιλεγμένου καθαρού κειμένου.* Ο κρυπταναλυτής έχει πρόσβαση για κάποιο χρονικό διάστημα στο μηχάνημα κρυπτο-

γράφησης. Έτσι, μπορεί να κρυπτογραφεί επιλεγμένα καθαρά κείμενα χωρίς όμως να γνωρίζει το κλειδί της κρυπτογράφησης.

4. *Προσβολή επιλεγμένου κρυπτογραφημένου κειμένου.* Ο κρυπταναλυτής έχει πρόσβαση για κάποιο χρονικό διάστημα στο μηχάνημα αποκρυπτογράφησης. Έτσι, μπορεί να αποκρυπτογραφεί επιλεγμένα κρυπτογραφημένα κείμενα χωρίς όμως να γνωρίζει το κλειδί της αποκρυπτογράφησης.

Σε κάθε περίπτωση ο σκοπός του κρυπταναλυτή είναι η εύρεση του κλειδιού. Στις επόμενες ενότητες θ΄ ασχοληθούμε κυρίως με τον ασθενέστερο τρόπο προσβολής του κρυπτοσυστήματος, την προσβολή κρυπτογραφημένου κειμένου. Στην περίπτωση όπου το καθαρό κείμενο είναι γραμμένο σε μία φυσική γλώσσα, πολλές τεχνικές της κρυπτανάλυσης χρησιμοποιούν τις στατιστικές ιδιότητες αυτής της γλώσσας. Η πιθανότητα εμφάνισης των γραμμάτων της αγγλικής γλώσσας έχει υπολογιστεί από πολλούς ερευνητές δια μέσου των περιοδικών, εφημερίδων, βιβλίων κ.λ.π. Δίνουμε παρακάτω ένα πίνακα, ο οποίος οφείλεται στους Beker και Piper, και παρουσιάζει την πιθανότητα εμφάνισης των γραμμάτων του αγγλικού αλφαβήτου σ΄ ένα κείμενο.

ΠΙΝΑΚΑΣ 1
Πιθανότητα εμφάνισης των γραμμάτων του αγγλικού αλφαβήτου

Γράμμα	Πιθανότητα	Γράμμα	Πιθανότητα
A	.082	N	.067
B	.015	O	.075
C	.028	P	.019
D	.043	Q	.001
E	.127	R	.060
F	.022	S	.063
G	.020	T	.091
H	.061	U	.028
I	.070	V	.010
J	.002	W	.023
K	.008	X	.001
L	.040	Y	.020
M	.024	Z	.001

Με βάση τον παραπάνω πίνακα οι Beker και Piper ταξινόμησαν τα γράμματα του αγγλικού αλφαβήτου σε πέντε ομάδες:

1. E, με πιθανότητα εμφάνισης 0.127.
2. T, A, O, I, N, S, H, R με πιθανότητα εμφάνισης μεταξύ 0.06 και 0.09.
3. D, L με πιθανότητα εμφάνισης 0.043 και 0.040 αντίστοιχα.
4. C, U, M, W, F, G, Y, P, B με πιθανότητα εμφάνισης μεταξύ 0.015 και 0.028.
5. V, K, J, X, Q, Z με πιθανότητα εμφάνισης μικρότερη του 0.01.

Επίσης, είναι χρήσιμο να αναφέρουμε και την πιθανότητα εμφάνισης ζευγών ή τριάδων γραμμάτων. Έτσι, τα 30 ζεύγη γραμμάτων με την μεγαλύτερη πιθανότητα εμφάνισης είναι κατά σειρά συχνότερης εμφάνισης τα εξής: TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI και OF. Τέλος, οι 12 τριάδες γραμμάτων με την μεγαλύτερη πιθανότητα εμφάνισης είναι κατά σειρά συχνότερης εμφάνισης οι εξής: THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR και DTH. Ας σημειωθεί ότι ανάλογα αποτελέσματα υπάρχουν και για άλλες γλώσσες. Στις επόμενες ενότητες θα μελετήσουμε μερικά απλά συμμετρικά κρυπτοσυστήματα.

1.3 Κρυπτοσύστημα Μετατόπισης

Το αλφάβητο που χρησιμοποιήται στο κρυπτοσύστημα μετατόπισης είναι ο δακτύλιος $\mathbb{Z}_n$. Ο χώρος των καθαρών κειμένων P, ο χώρος των κρυπτογραφημένων κειμένων C και ο χώρος των κλειδιών K είναι επίσης το σύνολο $\mathbb{Z}_n$. Για κάθε $k \in \mathbb{Z}_n$ ορίζουμε την συνάρτηση κρυπτογράφησης

$$E_k : \mathbb{Z}_n \longrightarrow \mathbb{Z}_n,\ x \longmapsto x + k$$

και την συνάρτηση αποκρυπτογράφησης

$$D_k : \mathbb{Z}_n \longrightarrow \mathbb{Z}_n,\ x \longmapsto x - k.$$

Για $k = \overline{3}$ προκύπτει το κρυπτοσύστημα του Ιούλιου Καίσαρα.

Παράδειγμα 1.1 Ας αντιστοιχήσουμε τα γράμματα του Ελληνικού αλφαβήτου με τα στοιχεία του $\mathbb{Z}_{24}$, ή καλύτερα με το πλήρες σύστημα υπολοίπων mod 24, $\{0, 1, \ldots, 23\}$, ως εξής:

$$\text{A} \longleftrightarrow 0, \ \text{B} \longleftrightarrow 1, \ \ldots, \Omega \longleftrightarrow 23.$$

Οπότε το κείμενο

"ΕΠΙΘΕΣΗ"

αντιστοιχεί στην ακολουθία ακεραίων

"4 15 8 7 4 17 6".

Ας υποθέσουμε ότι $k = \bar{8}$. Τότε

$$E_{\bar{8}}(\bar{4}) = \overline{12}, \ E_{\bar{8}}(\overline{15}) = \overline{23}, \ E_{\bar{8}}(\bar{8}) = \overline{16},$$

$$E_{\bar{8}}(\bar{7}) = \overline{15}, \ E_{\bar{8}}(\overline{17}) = \bar{1}, \ E_{\bar{8}}(\bar{6}) = \overline{14}.$$

Θεωρούμε τους αντιπροσώπους των παραπάνω κλάσεων που βρίσκονται μέσα στο σύνολο $\{0, 1, \ldots, 23\}$ και έτσι προκύπτει το κρυπτογραφημένο κείμενο:

12 23 16 15 12 1 14.

Συνεπώς, ο παραλήπτης θα λάβει το μήνυμα

"ΝΩΡΟΝΒΞ".

Για να το αποκρυπτογραφήσει θα ακολουθήσει την αντίστροφη πορεία. Θα το μετατρέψει σε ακολουθία ακεραίων και κατόπιν χρησιμοποιώντας την συνάρτηση $D_{\bar{8}}$ θα βρεί την ακολουθία ακεραίων που είναι το καθαρό κείμενο. Τέλος θα βρεί τα γράμματα που αντιστοιχούν στους ακέραιους της ακολουθίας και έτσι θα προκύψει το μήνυμα.

Παρατηρούμε ότι το κρυπτοσύστημα μετατόπισης δεν είναι ασφαλές. Καθώς υπάρχουν n δυνατότητες για το κλειδί, στην περίπτωση όπου ο n δεν είναι μεγάλος, π.χ. $n = 24$, τότε μπορούμε να το βρούμε εύκολα δοκιμάζοντας όλες τις περιπτώσεις.

1.4 Ομοπαραλληλικό Κρυπτοσύστημα

Το ομοπαραλληλικό κρυπτοσύστημα είναι μία γενίκευση του κρυπτοσυστήματος μετατόπισης. Το αλφάβητο που χρησιμοποιείται σ΄ αυτό είναι ο δακτύλιος $\mathbb{Z}_n$. Ο χώρος των καθαρών κειμένων P και ο χώρος των κρυπτογραφημένων κειμένων C είναι επίσης ο δακτύλιος $\mathbb{Z}_n$. Ο χώρος των κλειδιών K είναι το καρτεσιανό γινόμενο $\mathbb{Z}_n^* \times \mathbb{Z}_n$. Το πλήθος των στοιχείων του ισούται με $\phi(n)n$. Για κάθε $(a,b) \in K$ ορίζουμε την συνάρτηση κρυπτογράφησης

$$E_{(a,b)} : \mathbb{Z}_n \longrightarrow \mathbb{Z}_n,\ x \longmapsto ax + b$$

και την συνάρτηση αποκρυπτογράφησης

$$D_{(a,b)} : \mathbb{Z}_n \longrightarrow \mathbb{Z}_n,\ x \longmapsto a^{-1}x - a^{-1}b.$$

Για κάθε $x \in \mathbb{Z}_n$ έχουμε

$$(D_{(a,b)} \circ E_{(a,b)})(x) = D_{(a,b)}(ax + b) = a^{-1}(ax + b) - a^{-1}b = x.$$

Παράδειγμα 1.2 Ας αντιστοιχήσουμε, τα γράμματα του Ελληνικού αλφαβήτου με τους ακεραίους $0, 1, \ldots, 23$, όπως στο προηγούμενο παράδειγμα. Επίσης, ας αντιστοιχήσουμε στο κενό διάστημα τον ακέραιο 24 και στο ερωματηματικό ";" τον 25. Έτσι, έχουμε $n = 26$. Τότε το μήνυμα

"ΘΑ ΕΡΘΕΙΣ;"

αντιστοιχεί στην ακολουθία ακεραίων

"7 0 24 4 16 7 4 8 17 25".

Ας είναι $k = (\bar{7}, \bar{4})$. Τότε η συνάρτηση κρυπτογράφησης είναι

$$E_{(\bar{7},\bar{4})} : \mathbb{Z}_{26} \longrightarrow \mathbb{Z}_{26},\ x \longmapsto \bar{7}x + \bar{4}.$$

Έχουμε

$$E_{(\bar{7},\bar{4})}(\bar{7}) = \bar{1},\ E_{(\bar{7},\bar{4})}(\bar{0}) = \bar{4},\ E_{(\bar{7},\bar{4})}(\overline{24}) = \overline{16},\ E_{(\bar{7},\bar{4})}(\bar{4}) = \bar{6}$$

$$E_{(\bar{7},\bar{4})}(\overline{16}) = \overline{12},\ E_{(\bar{7},\bar{4})}(\bar{8}) = \bar{8},\ E_{(\bar{7},\bar{4})}(\overline{17}) = \overline{19},\ E_{(\bar{7},\bar{4})}(\overline{25}) = \overline{23},$$

απ΄ όπου, παίρνουμε την ακολουθία ακεραίων:

"1 4 16 6 12 1 6 8 19 23".

Επομένως το κρυπτογραφημένο μήνυμα είναι:

"ΒΕΡΘΝΕΙΥΩ".

Καθώς $\bar{7}^{-1} = \overline{15}$, η συνάρτηση αποκρυπτογράφησης είναι

$$D_{(\bar{7},\bar{4})} : \mathbb{Z}_{26} \longrightarrow \mathbb{Z}_{26},\ x \longmapsto \overline{15}(x - \bar{4}).$$

Χρησιμοποιώντας την συνάρτηση $D_{(\bar{7},\bar{4})}$, ο παραλήπτης παίρνει την αρχική ακολουθία ακεραίων και κατά συνέπεια διαβάζει το κείμενο του μηνύματος.

Στη συνέχεια θα δείξουμε μ' ένα παράδειγμα την μέθοδο με την οποία γίνεται η κρυπτανάλυση του ομοπαραλληλικού συστήματος με την χρήση στατιστικών δεδομένων.

Παράδειγμα 1.3 Το παρακάτω κρυπτογραφημένο κείμενο κρυπτογραφήθηκε με την χρήση του ομοπαραλληλικού κρυπτοσυστήματος με $n = 26$ και την αντιστοίχιση των γραμμάτων του αγγλικού αλφαβήτου με τους ακέραιους $0, 1, \ldots, 25$.

"*FMXVEDKAPHFERBNDKRXRSREFMORUDSD*

KDVSHVUFEDKAPRKDLYEVLRHHRH".

Η συχνότητα εμφάνισης κάθε γράμματος δίνεται στον παρακάτω πίνακα:

Γράμμα	Συχνότητα	Γράμμα	Συχνότητα
A	2	*N*	1
B	1	*O*	1
C	0	*P*	2
D	7	*Q*	0
E	5	*R*	8
F	4	*S*	3
G	0	*T*	0
H	5	*U*	2
I	0	*V*	4
J	0	*W*	0
K	5	*X*	2
L	2	*Y*	1
M	2	*Z*	0

Υπάρχουν μόνο 57 γράμματα στο παραπάνω κείμενο. Αυτό είναι συνήθως αρκετό για την κρυπτανάλυση ενός ομοπαραλληλικού συστήματος. Καθώς το R εμφανίζεται τις περισσότερες φορές, λαμβανομένου υπ' όψιν του πίνακα της πιθανότητας εμφάνισης των γραμμάτων του αγγλικού αλφαβήτου, υποθέτουμε ότι το E αντιστοιχεί στο R. Μετά το R το D εμφανίζεται τις περισσότερες φορές και επομένως υποθέτουμε ότι αντιστοιχεί στο T. Αν $E_{(a,b)}$ είναι η συνάρτηση κρυπτογράφησης, τότε

$$E_{(a,b)}(\bar{4}) = \bar{4}a + b = \overline{17}, \quad E_{(a,b)}(\overline{19}) = \overline{19}a + b = \bar{3},$$

απ' όπου παίρνουμε $a = \bar{6}$, $b = \overline{19}$. Καθώς όμως $(6, 26) = 2$, η κλάση $\bar{6}$ δεν είναι αντιστρέψιμη μέσα στο $\mathbb{Z}_{26}$ και συνεπώς η υπόθεση που κάναμε δεν είναι σωστή.

Στη συνέχεια υποθέτουμε ότι το E αντιστοιχεί στο R και το T στο E. Έτσι, έχουμε

$$\bar{4}a + b = \overline{17}, \quad \overline{19}a + b = \bar{4},$$

απ' όπου παίρνουμε $a = \overline{13}$ το οποίο δεν είναι αντιστρέψιμο μέσα στο $\mathbb{Z}_{26}$. Επόμενη υπόθεσή μας είναι ότι το E αντιστοιχεί στο R και το T στο H. Βρίσκουμε $a = \bar{8}$, τιμή που ομοίως δεν είναι αποδεκτή. Συνεχίζοντας, υποθέτουμε ότι το E αντιστοιχεί στο R και το T στο K. Αυτό μας δίνει $a = \bar{3}$ και $b = \bar{5}$, τιμές που είναι αποδεκτές. Επομένως, η συνάρτηση αποκρυπτογράφησης δίνεται από τον τύπο

$$D_{(\bar{3},\bar{5})}(x) = \bar{9}x + \bar{7}.$$

Εφαρμόζουμε την $D_{(\bar{3},\bar{5})}$ στο κρυπτογραφημένο κείμενο και παίρνουμε:

"*ALGORITHMSAREQUITEGENERALDEFINITIONS*

OFARITHMETICPROCESS"

Καθώς το αποτέλεσμα είναι ένα αποδεκτό κείμενο υποθέτουμε ότι βρήκαμε το σωστό κλειδί. Ας σημειωθεί ότι στη περίπτωση $n = 26$, ο χώρος κλειδιών έχει 312 στοιχεία. Η μέθοδος της δοκιμής καθενός από αυτά είναι δυνατόν να απαιτήσει πολύ περισσότερο χρόνο.

1.5 Κρυπτοσύστημα του *Hill*

Ας είναι A ένας $m \times m$-πίνακας με στοιχεία από το $\mathbb{Z}_n$. Αν ο A είναι αντιστρέψιμος, τότε υπάρχει πίνακας B με στοιχεία από το $\mathbb{Z}_n$ έτσι,

ώστε $AB = I_m$, όπου I_m είναι ο μοναδιαίος $m \times m$-πίνακας. Παίρνοντας τις ορίζουσες, έχουμε

$$(\det A)(\det B) = 1$$

και κατά συνέπεια το στοιχείο $\det A$ είναι αντιστρέψιμο μέσα στο $\mathbb{Z}_n$. Αντιστρόφως, ας υποθέσουμε ότι το στοιχείο $\det A$ είναι αντιστρέψιμο μέσα στο $\mathbb{Z}_n$. Ας είναι A_{ji} ο πίνακας που προκύπτει από τον A, αν απαλείψουμε την j-γραμμή και την i-στήλη. Συμβολίζουμε με A^* τον πίνακα που έχει στη (i,j)-θέση το στοιχείο $(-1)^{i+j}\det A_{ji}$. Τότε, όπως και στη περίπτωση των πινάκων με στοιχεία πραγματικούς αριθμούς, παίρνουμε

$$AA^* = (\det A)I_m.$$

Καθώς το στοιχείο $\det A$ είναι αντιστρέψιμο, έπεται ότι ο A είναι αντιστρέψιμος και ο αντίστροφός του είναι

$$A^{-1} = (\det A)^{-1}A^*.$$

Άρα ο A είναι αντιστρέψιμος, αν και μόνον αν το στοιχείο $\det A$ είναι αντιστρέψιμο μέσα στο $\mathbb{Z}_n$. Σ' αυτή την περίπτωση η απεικόνιση

$$\alpha_A : \mathbb{Z}_n^m \longrightarrow \mathbb{Z}_n^m,\ x \longmapsto xA$$

είναι αμφίεση.

Στη συνέχεια θα ορίσουμε ένα κρυπτοσύστημα που εφευρέθηκε στα 1929 από τον Lester S. Hill. Ας είναι m ένας ακέραιος ≥ 2. Ο χώρος P των καθαρών κειμένων και ο χώρος C των κρυπτογραφημένων κειμένων είναι το σύνολο $\mathbb{Z}_n^m$. Ο χώρος κλειδιών K είναι το σύνολο όλων των $m \times m$-πινάκων με στοιχεία από το $\mathbb{Z}_n$ οι οποίοι είναι αντιστρέψιμοι. Για κάθε $A \in K$ η συνάρτηση κρυπτογράφησης είναι

$$E_A : \mathbb{Z}_n^m \longrightarrow \mathbb{Z}_n^m,\ x \longmapsto xA$$

και η συνάρτηση αποκρυπτογράφησης

$$D_A : \mathbb{Z}_n^m \longrightarrow \mathbb{Z}_n^m,\ x \longmapsto xA^{-1}.$$

Παράδειγμα 1.4 Αντιστοιχούμε τα γράμματα του αγγλικού αλφαβήτου με τους αριθμούς $0, 1, \ldots, 25$ όπως κάναμε στα προηγούμενα παραδείγματα. Ας είναι

$$A = \begin{pmatrix} \bar{2} & \bar{3} \\ \bar{7} & \bar{8} \end{pmatrix}.$$

Η ορίζουσα του πίνακα A ισούται με $\overline{21}$ που είναι αντιστρέψιμο στοιχείο του $\mathbb{Z}_{26}$ και επομένως ο A είναι αντιστρέψιμος. Η λέξη NO αντιστοιχεί στο ζεύγος 13, 14. Έχουμε

$$(\overline{13}, \overline{14}) \begin{pmatrix} \bar{2} & \bar{3} \\ \bar{7} & \bar{8} \end{pmatrix} = (\overline{20}, \overline{21}).$$

Επομένως, το κρυπτογραφημένο μήνυμα είναι το 20, 21 που αντιστοιχεί στα γράμματα U, V. Για ν' αποκρυπτογραφήσει ο παραλήπτης αυτό το μήνυμα θα πρέπει να υπολογίσει τον αντίστροφο πίνακα του A που είναι

$$A^{-1} = \overline{21}^{-1} \begin{pmatrix} \bar{8} & -\bar{3} \\ -\bar{7} & \bar{2} \end{pmatrix}.$$

Οπότε

$$(\overline{20}, \overline{21})\bar{21}^{-1} \begin{pmatrix} \bar{8} & -\bar{3} \\ -\bar{7} & \bar{2} \end{pmatrix} = (\overline{13}, \overline{14}).$$

Το κρυπτοσύστημα του Hill είναι εύκολο να κρυπταναλυθεί με προσβολή γνωστού καθαρού κειμένου. Ας υποθέσουμε ότι ο κρυπταναλυτής γνωρίζει την τιμή του m και ότι έχει στη κατοχή του m ζεύγη καθαρών κειμένων

$$X_j = (x_{j1}, \ldots, x_{jm}) \quad (j = 1, \ldots, m)$$

και αντιστοίχων κρυπτογραφημένων κειμένων

$$Y_j = (y_{j1}, \ldots, y_{jm}) \quad (j = 1, \ldots, m).$$

Αν A είναι ο πίνακας κλειδί που αναζητούμε, τότε $E_A(X_j) = Y_j$ $(j = 1, \ldots, m)$. Θεωρούμε τους πίνακες $X = (x_{ij})$ και $Y = (y_{ij})$. Οπότε έχουμε την εξίσωση πινάκων:

$$Y = XA.$$

Αν ο πίνακας X είναι αντιστρέψιμος, τότε $A = X^{-1}Y$. Αν ο X δεν είναι αντιστρέψιμος, τότε προσπάθούμε να βρούμε άλλα ζεύγη καθαρών - κρυπτογραφημένων κειμένων. Στη περίπτωση όπου $n = 2$ και ο κρυπταναλυτής έχει στη διάθεσή του ένα αρκετά μεγάλο κρυπτογραφημένο μήνυμα ή μέρους αυτού, τότε με την χρήση στατιστικών δεδομένων είναι δυνατόν να προσδιορίσει δύο ζεύγη καθαρών - κρυπτογραφημένων κειμένων.

Παράδειγμα 1.5 Ας υποθέσουμε ότι το κρυπτοσύστημα του Hill με $m = 2$ χρησιμοποιήται για την κρυπτογράφηση της λέξης

$$\text{“}FRIDAY\text{”}$$

και δίνει

$$\text{“}PQCFKU\text{”}.$$

Αν A είναι ο πίνακας κλειδί, τότε

$$E_A(\bar{5}, \overline{17}) = (\overline{15}, \overline{16}), \quad E_A(\bar{8}, \bar{3}) = (\bar{2}, \bar{5}), \quad E_A(\bar{0}, \overline{24}) = (\overline{10}, \overline{20}).$$

Από τις πρώτες δύο σχέσεις, παίρνουμε

$$\begin{pmatrix} \overline{15} & \overline{16} \\ \bar{2} & \bar{5} \end{pmatrix} = \begin{pmatrix} \bar{5} & \overline{17} \\ \bar{8} & \bar{3} \end{pmatrix} A.$$

Έχουμε

$$\begin{pmatrix} \bar{5} & \overline{17} \\ \bar{8} & \bar{3} \end{pmatrix}^{-1} = \begin{pmatrix} \bar{9} & \bar{1} \\ \bar{2} & \overline{15} \end{pmatrix}$$

και επομένως

$$A = \begin{pmatrix} \bar{9} & \bar{1} \\ \bar{2} & \overline{15} \end{pmatrix} \begin{pmatrix} \overline{15} & \overline{16} \\ \bar{2} & \bar{5} \end{pmatrix} = \begin{pmatrix} \bar{7} & \overline{19} \\ \bar{8} & \bar{3} \end{pmatrix}.$$

1.6 Κρυπτοσύστημα Μετάθεσης

Ένα κρυπτοσύστημα που χρησιμοποιήθηκε για αρκετούς αιώνες είναι το κρυπτοσύστημα μετάθεσης που θα ορίσουμε στη συνέχεια. Ας είναι m ένας θετικός ακέραιος. Ο χώρος καθαρών κειμένων P και ο χώρος κρυπτογραφημένων κειμένων C είναι το σύνολο $\mathbb{Z}_n^m$. Ο χώρος κλειδιών K είναι το σύνολο όλων των μεταθέσεων του συνόλου $\{1, 2, \ldots, m\}$. Για κάθε $\sigma \in K$ η συνάρτηση κρυπτογράφησης είναι

$$E_\sigma : \mathbb{Z}_n^m \longrightarrow \mathbb{Z}_n^m, \ (x_1, \ldots, x_m) \longmapsto (x_{\sigma(1)}, \ldots, x_{\sigma(m)})$$

και η συνάρτηση αποκρυπτογράφησης

$$D_\sigma : \mathbb{Z}_n^m \longrightarrow \mathbb{Z}_n^m, \ (x_1, \ldots, x_m) \longmapsto (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(m)}).$$

Παράδειγμα 1.6 Ας είναι $m = 6$ και το κλειδί η παρακάτω μετάθεση:

x	1	2	3	4	5	6
$\sigma(x)$	3	5	1	6	4	2

Αντιστοιχούμε τα γράμματα του Ελληνικού αλφαβήτου, με τον συνήθη τρόπο, στους ακέραιους $0, 1, \ldots, 23$ και στο κενό διάστημα τον 24. Για να κρυπτογραφήσουμε το μήνυμα

"ΕΠΙΘΕΣΗ ΣΤΙΣ ΔΕΚΑ"

το χωρίζουμε σε ομάδες των έξι συμβόλων και το μετατρέπουμε σε μία ακολούθία ακεραίων. Καθώς έχουμε μόνο 17 σύμβολα, στο τέλος προσθέτουμε το σύμβολο για το κενό διάστημα. Έτσι παίρνουμε τις τρείς εξάδες ακεραίων

(4,15,8,7,4,17), (6,24,17,18,8,17), (24,3,4,9,0,24).

Εφαρμόζοντας σε κάθε μία εξάδα την μετάθεση σ προκύπτουν οι εξάδες:

(8,4,4,17,7,15), (17,8,6,17,18,24), (4,0,24,24,9,3).

Μετατρέποντας τις εξάδες ακεραίων σε σύμβολα, παίρνουμε

"ΙΕΕΣΘΞΣΙΗΣΤ ΕΑ ΚΔ" .

Η αποκρυπτογράφηση γίνεται με τον ίδιο τρόπο χρησιμοποιώντας την μετάθεση σ^{-1}.

Στη συνέχεια θα δείξουμε ότι το κρυπτοσύστημα μετάθεσης είναι ειδική περίπτωση του κρυπτοσυστήματος του Hill. Σε κάθε μετάθεση σ του $\{1, \ldots, m\}$ αντιστοιχούμε έναν $m \times m$-πίνακα $K_\sigma = (k_{ij})$, όπου $k_{ij} = 1$ αν $i = \sigma(j)$ και $k_{ij} = 0$ αν $i \neq \sigma(j)$. Έχουμε $K_\sigma^{-1} = K_{\sigma^{-1}}$ και

$$(x_1, \ldots, x_m)K_\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(m)}).$$

Για παράδειγμα η μετάθεση σ του παραπάνω παραδείγματος αντιστοιχεί στον πίνακα

$$K_\sigma = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

Επίσης

$$K_{\sigma^{-1}} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

1.7 Κρυπτοσύστημα Αντικατάστασης

Ένα από τα πλέον γνωστά και ευρέως χρησιμοποιημένα κρυπτοσυστήματα επί σειρά αιώνων είναι το κρυπτοσύστημα αντικατάστασης. Ο χώρος καθαρών κειμένων P και ο χώρος κρυπτογραφημένων κειμένων C είναι το σύνολο $\{0, 1, \ldots, n-1\}$. Ο χώρος κλειδιών K είναι το σύνολο των μεταθέσεων του $\{0, 1, \ldots, n-1\}$. Για κάθε $\sigma \in K$ ορίζουμε την συνάρτηση κρυπτογράφησης

$$E_\sigma : \{0, 1, \ldots, n-1\} \longrightarrow \{0, 1, \ldots, n-1\},\ x \longmapsto \sigma(x)$$

και την συνάρτηση αποκρυπτογράφησης

$$D_\sigma : \{0, 1, \ldots, n-1\} \longrightarrow \{0, 1, \ldots, n-1\},\ x \longmapsto \sigma^{-1}(x).$$

Παράδειγμα 1.7 Ας υποθέσουμε ότι $n = 26$ και ότι έχουμε ταυτίσει τα γράμματα του αγγλικού αλφαβήτου με τους ακεραίους $0, 1, \ldots, 25$ ως συνήθως. Θεωρούμε την παρακάτω μετάθεση γραμμάτων:

A	B	C	D	E	F	G	H	I	J	K	L	M
Y	M	I	H	B	A	W	C	X	V	D	N	O

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
J	K	U	Q	P	R	T	F	E	L	G	Z	S

Χρησιμοποιώντας το κρυπτοσύστημα αντικατάστασης που αντιστοιχεί σ' αυτή την μετάθεση, κρυπτογραφούμε το μήνυμα

$$\text{“}COMEATONCE\text{”}$$

και παίρνουμε

$$\text{“}IKOBYTKJIB\text{”}.$$

Παρατηρούμε ότι ο χώρος κλειδιών έχει $n!$ στοιχεία. Έτσι, στη περίπτωση του λατινικού αλφαβήτου ($n = 26$) έχουμε περισσότερα από $4{\times}10^{26}$ κλειδιά. Επομένως, ακόμη και για έναν υπολογιστή η μέθοδος της δοκιμής όλων των δυνατών κλειδιών για την αποκρυπτογράφηση ένος κειμένου δεν είναι πρακτικώς δυνατή. Στη συνέχεια θα δείξουμε με ένα παράδειγμα πως επιτυγχάνεται η κρυπτανάλυση ενός κρυπτογραφημένου κειμένου μ΄ αυτό το σύστημα.

Παράδειγμα 1.8 Ας υποθέσουμε ότι το παρακάτω κρυπτογραφημένο κείμενο κρυπτογραφήθηκε με την χρήση του συστήματος αντικατάστασης και των γραμμάτων του αγγλικού αλφαβήτου:

"*YIFQFMZRWQFYVECFMDZPCVMRZWNMDZ*

VEJBTXCDDUMJNDIFEFMDZCDMQZKCEY

FCJMYRNCWJCSZREXCHZUNMXZNZUCDR

JXYYSMRTMEYIFZWDYVZVYFZUMRZCRW

NZDZJJXZWGCHSMRNMDHNCMFQCHZJMX

JZWIEJYUCFWDJNZDIR".

Η συχνότητα εμφάνισης κάθε γράμματος δίνεται στον παρακάτω πίνακα:

Γράμμα	Συχνότητα	Γράμμα	Συχνότητα
A	0	*N*	9
B	1	*O*	0
C	15	*P*	1
D	13	*Q*	4
E	7	*R*	10
F	11	*S*	3
G	1	*T*	2
H	4	*U*	5
I	5	*V*	5
J	11	*W*	8
K	1	*X*	6
L	0	*Y*	10
M	16	*Z*	20

Καθώς το πλήθος εμφανίσεων του Z είναι πολύ μεγαλύτερο από αυτό των άλλων γραμμάτων, υποθέτουμε ότι $D_k(Z) = E$. Από τα υπόλοιπα γράμματα του κειμένου αυτά που εμφανίζονται τουλάχιστον 10 φορές το καθένα είναι τα C, D, F, J, M, R, Y. Έτσι, υποθέτουμε ότι αυτά είναι οι κρυπτογραφήσεις κάποιων από τα T, A, O, I, N, S, H, R. Όμως οι συχνότητες εμφάνισής τους δεν διαφέρουν αρκετά, ώστε να έχουμε σαφή ένδειξη για την αντιστοιχία.

Σ΄ αυτό το σημείο θα θεωρήσουμε ζεύγη γραμμάτων της μορφής $-Z$ ή $Z-$. Τα πλέον εμφανιζόμενα ζεύγη γραμμάτων στο κείμενο είναι τα DZ και ZW (4 φορές το καθένα), NZ, ZU (3 φορές το καθένα) και $RZ, HZ, XZ, FZ, ZR, ZV, ZC, ZD, ZJ, RW$ (2 φορές το καθένα). Παρατηρούμε ότι ενώ το ZW εμφανίζεται τέσσερεις φορές, το WZ δεν εμφανίζεται καμμία. Από την άλλη πλευρά, τα ζεύγη γραμμάτων που εμφανίζονται συχνότερα σ΄ ένα κείμενο και αρχίζουν από το E είναι τα ER, ED, ES, EN, EA, και ET. Καθώς τα RE, SE, TE ανήκουν και αυτά στην ίδια κατηγορία, τότε υποθέτουμε ότι το ZW θα πρέπει ν΄ αποκρυπτογραφείται σε κάποιο από τα ED, EN, EA. Το γεγονός ότι η συχνότητα εμφάνισης του W στο κείμενό μας δεν είναι μεγάλη (8 φορές) και η πιθανότητα εμφάνισης των N και A σ΄ ένα αγγλικό κείμενο είναι μεγαλύτερη απ΄ αυτή του D μας οδηγούν να υποθέσουμε ότι $D_k(W) = D$. Στη συνέχεια βλέπουμε ότι το DZ εμφανίζεται στο κείμενο τέσσερεις φορές ενώ το ZD δύο. Τότε θα πρέπει το $D_k(D)$ να είναι κάποιο από τα γράμματα R, S, T.

Παρατηρούμε ότι το ζεύγος RW εμφανίζεται δύο φορές στο κείμενο. Από την άλλη πλευρά, από τα ζεύγη γραμμάτων με μεγάλη πιθανότητας εμφάνισης σ΄ ένα αγγλικό κείμενο είναι τα ED και ND. Καθώς το Z αποκρυπτογραφείται στο E, υποθέτουμε ότι $D_k(R) = N$. Το ζεύγος NZ εμφανίζεται τρείς φορές στο κείμενο ενώ το ZN μία. Επίσης, το HE είναι ένα από τα πλέον εμφανιζόμενα ζεύγη σε αγγλικό κείμενο ενώ το EH όχι. Συνεπώς, υποθέτουμε ότι $D_k(N) = H$. Σ΄ αυτό το σημείο έχουμε την εξής εικόνα:

$-$	$-$	$-$	$-$	$-$	$-$	E	N	D	$-$	$-$	$-$	$-$	$-$	$-$	$-$
Y	I	F	Q	F	M	Z	R	W	Q	F	Y	V	E	C	F
$-$	$-$	E	$-$	$-$	$-$	$-$	N	E	D	$-$	$-$	$-$	E	$-$	$-$
M	D	Z	P	C	V	M	R	Z	W	N	M	D	Z	V	E
$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$
J	B	T	X	C	D	D	U	M	J	N	D	I	F	E	F

–	–	*E*	–	–	–	–	*E*	–	–	–	–	–	–	–	–
M	*D*	*Z*	*C*	*D*	*M*	*Q*	*Z*	*K*	*C*	*E*	*Y*	*F*	*C*	*J*	*M*
–	*N*	–	–	*D*	–	–	–	*E*	*N*	–	–	–	–	*E*	–
Y	*R*	*N*	*C*	*W*	*J*	*C*	*S*	*Z*	*R*	*E*	*X*	*C*	*H*	*Z*	*U*
–	–	–	*E*	–	*E*	–	–	–	*N*	–	–	–	–	–	–
N	*M*	*X*	*Z*	*N*	*Z*	*U*	*C*	*D*	*R*	*J*	*X*	*Y*	*Y*	*S*	*M*
N	–	–	–	–	–	–	*E*	*D*	–	–	–	*E*	–	–	–
R	*T*	*M*	*E*	*Y*	*I*	*F*	*Z*	*W*	*D*	*Y*	*V*	*Z*	*V*	*Y*	*F*
E	–	–	*N*	*E*	–	*N*	*D*	–	*E*	–	*E*	–	–	–	*E*
Z	*U*	*M*	*R*	*Z*	*C*	*R*	*W*	*N*	*Z*	*D*	*Z*	*J*	*J*	*X*	*Z*
D	–	–	–	–	–	*N*	–	–	–	–	–	–	–	–	–
W	*G*	*C*	*H*	*S*	*M*	*R*	*N*	*M*	*D*	*H*	*N*	*C*	*M*	*F*	*Q*
–	–	*E*	–	–	–	–	*E*	*D*	–	–	–	–	–	–	–
C	*H*	*Z*	*J*	*M*	*X*	*J*	*Z*	*W*	*I*	*E*	*J*	*Y*	*U*	*C*	*F*

D	–	–	–	*E*	–	–	*N*
W	*D*	*J*	*N*	*Z*	*D*	*I*	*R*

Το κομμάτι $NE-NDHE$ σε συνδιασμό με την μεγάλη πιθανότητα εμφάνισης της τριάδας AND και των ζευγών EA και AN μας προτρέπουν να υποθέσουμε ότι $D_k(C) = A$. Μετά το Z το γράμμα που εμφανίζεται τις περισσότερες φορές στο κείμενο είναι το M. Η τριάδα RNM που αποκρυπτογραφήσαμε ως $NH-$ μας υποβάλλει ότι από το $H-$ αρχίζει μία λέξη και επομένως είναι πιθανό το M ν΄ αντιστοιχεί σ΄ ένα φωνήεν. Συνεπώς $D_k(M) = O$ ή I. Το ζεύγος AI εμφανίζεται συχνότερα σ΄ ένα αγγλικό κείμενο από το AO και καθώς το CM αποκρυπτογραφείται σε $A-$, υποθέτουμε ότι $D_k(M) = I$.

Στη συνέχεια θα προσπαθήσουμε να προσδιορίσουμε ποιανού γράμματος του κειμένου η αποκρυπτογράφηση είναι το O. Επειδή το O εμφανίζεται συχνά στα αγγλικά κείμενα, υποθέτουμε ότι η αποκρυπτογράφηση κάποιου από τα D, F, I, Y θα δίνει το O. Η πλέον κατάλληλη επιλογή είναι το Y γιατί διαφορετικά θα προέκυπτε η τριάδα AOI από την αποκρυπτογράφηση του CFM, CJM ή CDM. Έτσι, υποθέτουμε ότι $D_k(Y) = O$.

Τα εναπομείναντα πλέον συχνά εμφανιζόμενα γράμματα του κειμένου είναι τα D, F, J. Υποθέτουμε ότι αντιστοιχούν στα υπόλοιπα πλέον συχνά εμφανιζόμενα γράμματα στα αγγλικά κείμενα R, S, T. Καθώς

η τριάδα NMD εμφανίζεται δύο φορές, υποθέτουμε ότι $D_k(D) = S$ και επομένως η αποκρυπτογράφηση της τριάδας NMD δίνει την λέξη HIS. Επίσης, η πεντάδα $HNCMF$ θα ήταν δυνατόν να ήταν η κρυπτογράφηση της λέξης $CHAIR$. Οπότε $D_k(F) = R$, $D_k(H) = C$ και επομένως $D_k(J) = T$. Έτσι, έχουμε:

O	$-$	R	$-$	R	I	E	N	D	$-$	R	O	$-$	$-$	A	R
Y	I	F	Q	F	M	Z	R	W	Q	F	Y	V	E	C	F

I	S	E	$-$	A	$-$	I	N	E	D	H	I	S	E	$-$	$-$
M	D	Z	P	C	V	M	R	Z	W	N	M	D	Z	V	E

T	$-$	$-$	$-$	A	S	S	$-$	I	T	H	S	$-$	R	$-$	R
J	B	T	X	C	D	D	U	M	J	N	D	I	F	E	F

I	S	E	A	S	I	$-$	E	$-$	A	$-$	O	R	A	T	I
M	D	Z	C	D	M	Q	Z	K	C	E	Y	F	C	J	M

O	N	H	A	D	T	A	$-$	E	N	$-$	$-$	A	C	E	$-$
Y	R	N	C	W	J	C	S	Z	R	E	X	C	H	Z	U

H	I	$-$	E	H	E	$-$	A	S	N	T	$-$	O	O	$-$	I
N	M	X	Z	N	Z	U	C	D	R	J	X	Y	Y	S	M

N	$-$	I	$-$	O	$-$	R	E	D	S	O	$-$	E	$-$	O	R
R	T	M	E	Y	I	F	Z	W	D	Y	V	Z	V	Y	F

E	$-$	I	N	E	A	N	D	H	E	S	E	T	T	$-$	E
Z	U	M	R	Z	C	R	W	N	Z	D	Z	J	J	X	Z

D	$-$	A	C	$-$	I	N	H	I	S	C	H	A	I	R	$-$
W	G	C	H	S	M	R	N	M	D	H	N	C	M	F	Q

A	C	E	T	I	$-$	T	E	D	$-$	$-$	T	O	$-$	A	R
C	H	Z	J	M	X	J	Z	W	I	E	J	Y	U	C	F

D	S	T	H	E	S	$-$	N
W	D	J	N	Z	D	I	R

Στη συνέχεια δεν είναι δύσκολο ν' αποκρυπτογραφήσουμε ολόκληρο το κείμενο που είναι το εξής (αφού προσθέσαμε τα διαστήματα μεταξύ των λέξεων και τα σημεία στίξης):

> Our friend from Paris examined his empty glass with surprise, as if evaporation had taken place while he wasn't looking. I poured some more wine and he settled back in his chair, face titled up towards the sun.

1.8 Κρυπτοσύστημα του *Vigenère*

Σ' αυτή την ενότητα θα μελετήσουμε ένα κρυπτοσύστημα που οφείλεται στον Γάλλο διπλωμάτη Blaise de Vigenère. Ας είναι m ένας θετικός ακέραιος. Ο χώρος καθαρών κειμένων P, ο χώρος κρυπτογραφημένων κειμένων C και ο χώρος κλειδιών K είναι το σύνολο $\mathbb{Z}_n^m$. Για κάθε $k = (k_1, \ldots, k_m) \in K$ η συνάρτηση κρυπτογράφησης είναι

$$E_k : \mathbb{Z}_n^m \longrightarrow \mathbb{Z}_n^m,\ (x_1, \ldots, x_m) \longmapsto (x_1 + k_1, \ldots, x_m + k_m)$$

και η συνάρτηση αποκρυπτογράφησης

$$D_k : \mathbb{Z}_n^m \longrightarrow \mathbb{Z}_n^m,\ (x_1, \ldots, x_m) \longmapsto (x_1 - k_1, \ldots, x_m - k_m).$$

Ας σημειωθεί ότι το κρυπτοσύστημα αυτό διατυπώθηκε για πρώτη φορά στα 1586 για την περίπτωση $m = 26$.

Παράδειγμα 1.9 Ας υποθέσουμε ότι θέλουμε να στείλουμε το μήνυμα

$$\text{“}STRIKEATEVENING\text{”}.$$

Αντιστοιχούμε με τον συνήθη τρόπο τα γραμματα του αγγλικού αλφαβήτου και το κενό διάστημα με τους ακεραίους $0, \ldots, 26$ και μετατρέπουμε το μήνυμα στην ακολουθία ακεραίων

$$\text{“18 19 17 8 10 4 0 19 4 21 4 13 8 13 6”}.$$

Θα χρησιμοποιήσουμε το κρυπτοσύστημα του Vigenère με κλειδί $k = (\bar{4}, \bar{3}, \overline{10}, \bar{5})$

Τα σύμβολα του μηνύματος είναι 15. Επισυνάπτουμε στο τέλος της ακολουθίας το σύμβολο του κενού διαστήματος, τον 26. Κατόπιν χωρίζουμε την ακολουθία σε τετράδες και θεωρούμε τις αντίστοιχες κλάσεις του $\mathbb{Z}_{27}$. Οπότε, παίρνουμε τα εξής στοιχεία του $\mathbb{Z}_{27}^4$:

$$e_1 = (\overline{18}, \overline{19}, \overline{17}, \bar{8}),\quad e_2 = (\overline{10}, \bar{4}, \bar{0}, \overline{19}),$$

$$e_3 = (\bar{4}, \overline{21}, \bar{4}, \overline{13}),\quad e_4 = (\bar{8}, \overline{13}, \bar{6}, \overline{26}).$$

Έχουμε

$$e_1 + k = (\overline{22}, \overline{22}, \bar{0}, \bar{13}),\quad e_2 + k = (\overline{14}, \bar{7}, \overline{10}, \overline{24}),$$

$$e_3 + k = (\bar{8}, \overline{24}, \overline{14}, \overline{18}),\quad e_4 + k = (\overline{12}, \overline{16}, \overline{16}, \bar{4}).$$

Έτσι, παίρνουμε την ακολουθία ακεραίων

"22 22 0 13 14 7 10 24 8 24 14 18 12 16 16 4 ",

απ' όπου προκύπτει το κρυπτογραφημένο μήνυμα:

"*WWANOHKYIYOSMQQE*".

Για την αποκρυπτογράφησή του ο παραλήπτης ακολουθεί την αντίστροφη πορεία. Πρώτα το μετατρέπει σε ακολουθία ακεραίων και το χωρίζει σε τετράδες. Κατόπιν, θεωρεί τις αντίστοιχες τετράδες του $\mathbb{Z}_{27}^4$ και προσθέτει σ' αυτές την τετράδα $-k = (\overline{23}, \overline{24}, \overline{17}, \overline{22})$. Έτσι, προκύπτει μία ακολουθία ακεραίων της οποίας η μετατροπή σε γράμματα επιτρέπει στον παραλήπτη την ανάγνωση του μηνύματος.

Στη συνέχεια θ' ασχοληθούμε με δύο τεχνικές κρυπτανάλυσης του συστήματος του Vigenère, το κριτήριο του Kasiski και το κριτήριο του Friedman.

Το πρώτο κριτήριο περιγράφηκε από τον Friedrich Kasiski στα 1863, αν και είχε ανακαλυφθεί στα 1854 από τον Charles Babbage. Αυτό βασίζεται στη παρατήρηση ότι δύο ίδια κομμάτια μηνύματος τα οποία απέχουν d θέσεις με $d \equiv 0 \pmod{m}$ δίνουν ίδια κομμάτια κρυπτογραφημένου μηνύματος. Επομένως, το κριτήριο του Kasiski εφαρμόζεται ως εξής: Αναζητούμε στο κρυπτογραφημένο μήνυμα ίδια κομμάτια μήκους ≥ 3 και καταγράφουμε την απόσταση των θέσεων εμφάνισης του πρώτου γράμματός των. Αν $d_1, \ldots, d_r$ είναι τέτοιες αποστάσεις, μπορούμε να εικάσουμε ότι ο m διαιρεί τον μέγιστο κοινό διαιρέτη των $d_1, \ldots, d_r$.

Περισσότερες πληροφορίες για τον m μπορούμε να πάρουμε από το δεύτερο κριτήριο το οποίο διατυπώθηκε στα 1925 από τον William Frederick Friedman. Θα υποθέσουμε ότι $n = 26$ και ότι τα αποστελλόμενα μηνύματα είναι γραμμένα με γράμματα του αγγλικού αλφαβήτου.

Ας είναι $p_0, \ldots, p_{25}$ οι πιθανότητες εμφάνισης των γραμμάτων A, B,...,Z σ' ένα αγγλικό κείμενο όπως δίνονται στον Πίνακα 1. Αν επιλέξουμε τυχαία δύο γράμματα από ένα αγγλικό κείμενο, τότε η πιθανότητα και στις δύο θέσεις να βρίσκεται το A είναι κατά προσέγγιση ίση με p_0^2. Όμοια, η πιθανότητα και στις δύο θέσεις να βρίσκεται το B είναι κατά προσέγγιση ίση με p_1^2, κ.ο.κ. Συνεπώς, η πιθανότητα του ενδεχομένου δύο τυχόντα γράμματα ενός αγγλικού κειμένου να συμπίπτουν είναι ίση με

$$\sum_{i=0}^{25} p_i^2 \approx 0.065.$$