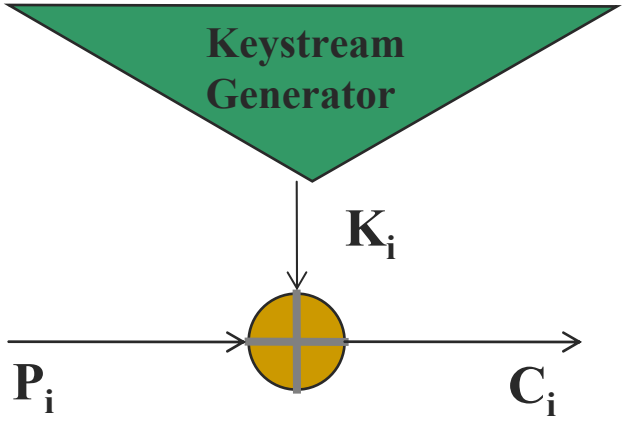
Κρυπτογραφία

Κεφάλαιο 2

Αλγόριθμοι ροής - Stream ciphers

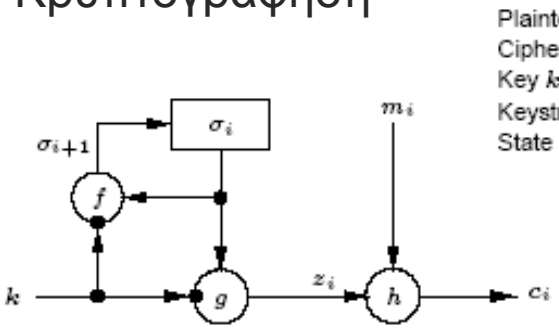
Γενικά χαρακτηριστικά



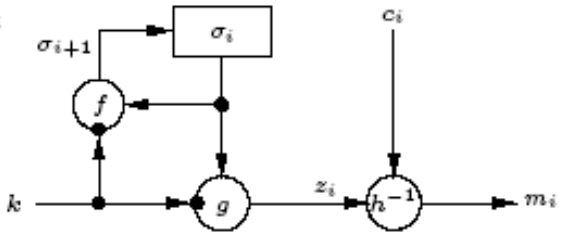
- Δουλεύουν πάνω σε ένα ρεύμα από bits (ή bytes)
- Απαιτούν μία γεννήτρια ψευδοτυχαίας ακολουθίας bits (keystream generator) – αυτή η ακολουθία που παράγεται λέγεται κλειδοροή (keystream)
- Τα bits του κλειδιού γίνονται XOR με τα bits του μηνύματος για να προκύψει έτσι το κρυπτόγραμμα, και αντίστροφα
- Η περίοδος της ακολουθίας του κλειδιού πρέπει αν είναι όσο γίνεται πιο μεγάλη
- Άλλες επιλογές
  - **Τα bits του κλειδιού μπορούν να εξαρτώνται από προηγούμενα bits του κρυπτογράμματος (ασύγχρονο σύστημα)**

Σύγχρονοι stream ciphers

Κρυπτογράφηση



Αποκρυπτογράφηση

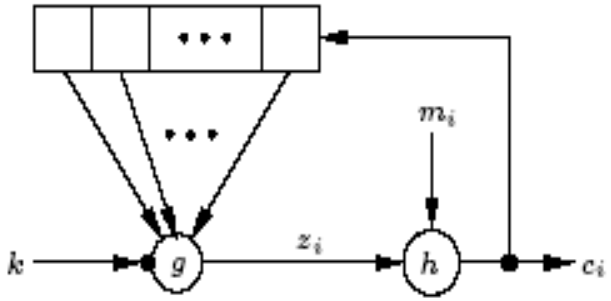


$\sigma_{i+1} = f(\sigma_i,k)$, $z_i = g(\sigma_i , k)$, $c_i = h(m_i, z_i)$

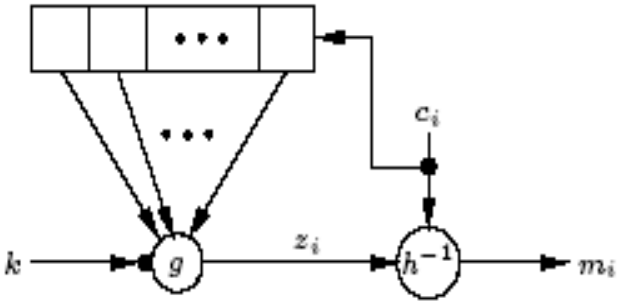
- Αποστολέας και παραλήπτης πρέπει να είναι συγχρονισμένοι. Αν κάποιο τμήμα του κρυπτογράμματος χαθεί κατά τη μετάδοσή του, ο συγχρονισμός χάνεται και απαιτούνται πρόσθετες τεχνικές επανασυγχρονισμού.
- Η παραποίηση (αλλοίωση) ενός ψηφίου του κρυπτογράμματος κατά τη μετάδοση δεν έχει ως αποτέλεσμα περαιτέρω λανθασμένη αποκρυπτογράφηση επόμενων ψηφίων (no error propagation)
- “Ενεργές επιθέσεις” προκαλούν σοβαρά προβλήματα (π.χ. έλλειψη συγχρονισμού), για αυτό απαιτούνται τεχνικές για πιστοποίηση της γνησιότητας και της ακεραιότητας του μηνύματος

Ασύγχρονοι stream ciphers

Κρυπτογράφηση



Αποκρυπτογράφηση



$\sigma_i = (c_{i-t}, c_{i-t+1}, \ldots, c_{i-1})$ $z_i = g(\sigma_i, k)$, $c_i = h(m_i, z_i)$

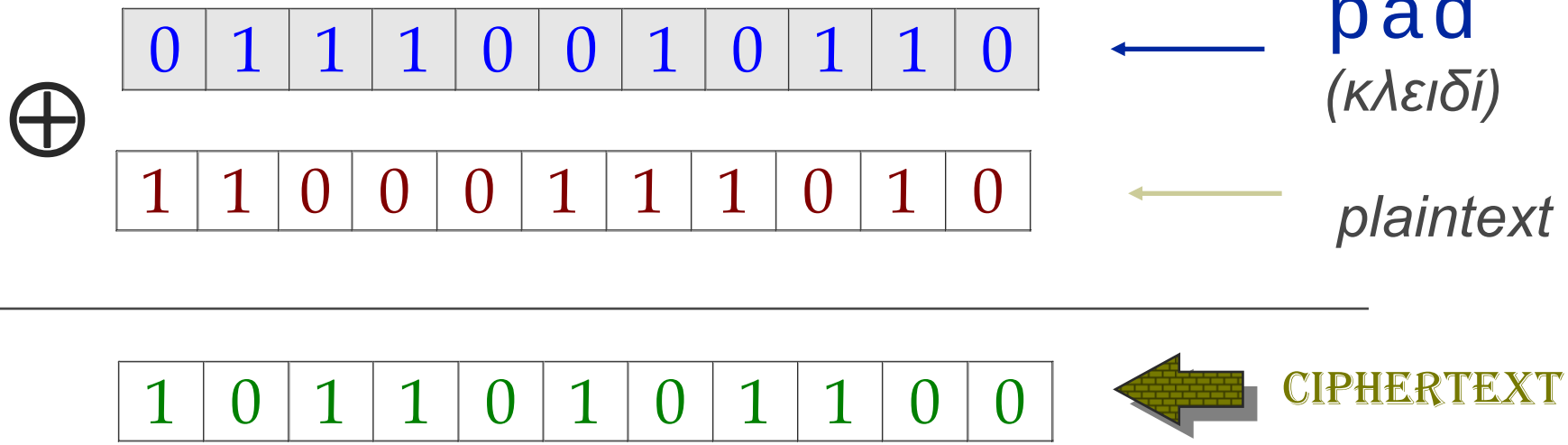
- Τα ασύγχρονα συστήματα ανακτούν από μόνα τους το συγχρονισμό όταν αυτό χαθεί, μετα από την παρέλευση κάποιων λανθασμένων αποκρυπτογραφήσεων (self – synchronization)
- Αν συμβεί μία λάθος αποκρυπτογράφηση, αυτή διαδίδεται για ορισμένα επόμενα ψηφία και μετά στο σύστημα επανέρχεται σε σωστή λειτουργία (limited error propagation)
- «Ενεργές επιθέσεις» ανιχνεύονται πιο δύσκολα από ό,τι στα σύγχρονα συστήματα – απαιτούνται επίσης τεχνικές για πιστοποίηση γνησιότητας και ακεραιότητας του μηνύματος
- Κάθε bit του μηνύματος επιδρά σε πολλά bits του κρυπτογράμματος, άρα οι στατιστικές ιδιότητες του μηνύματος «χάνονται» μέσα στο κρυπτόγραμμα

Σημειωματάριο μιας χρήσης (One-time pad)

- Ως σημειωματάριο μιας χρήσης (one-time pad) αποκαλείται το ιδανικό εκείνο κρυπτοσύστημα, όπου το κλειδί είναι τυχαία σειρά bits με απεριόριστα μεγάλη περίοδο (στην ουσία το μήκος του κλειδιού είναι ίσο με το μήκος του μηνύματος). Επιπρόσθετα, το ίδιο κλειδί δεν επαναχρησιμοποιείται ποτέ (κάθε νέο μήνυμα κρυπτογραφείται με διαφορετικό κλειδί) και τα στοιχεία του κλειδιού δεν σχετίζονται μεταξύ τους.
- Ο Shannon (1948) απέδειξε ότι το one-time pad είναι απεριόριστα ασφαλές.

  ("Communication Theory of Secrecy Systems", Claude Shannon, Bell Syst. Tech. J. 28, 656-715, 1949)
- Στην πράξη δεν μπορούμε να έχουμε τυχαίες ακολουθίες, παρά μόνο ψευδοτυχαίες (που προέρχονται από ντετερμινιστικές μηχανές).
- Ειδική περίπτωση: αν ο κρυπταλγόριθμος είναι αυτός του Vigenere και το μήκος του κλειδιού είναι όσο το μήκος του μηνύματος, τότε το κρυπτοσύστημα ονομάζεται **κρυπτοσύστημα Vernam**.

One-time pad (σχηματική αναπαράσταση)



Σημείωση: Λειτουργία του τελεστή XOR ⊕ :

- *a ⊕ b = 0 αν τα a,b είναι ίδια,*
- *a ⊕ b = 1 αν τα a,b είναι διαφορετικά.*

Αντίστοιχα, για πολλές μεταβλητές (π.χ. a ⊕ b ⊕ c ⊕ …..), αν άρτιος αριθμός από αυτές είναι 1 τότε το αποτέλεσμα είναι 1, αλλιώς το αποτέλεσμα είναι 0.

Βασικά στοιχεία Θεωρίας Πληροφορίας

- Αν X τυχαία μεταβλητή, με τιμές στο σύνολο $\{x_1, x_2, \ldots, x_n\}$ και αντίστοιχες πιθανότητες $p_1, p_2, \ldots, p_n$, η εντροπία της ορίζεται ως:

$$H(X) = \sum_{i=1}^{n} p_i \log_2 (\frac{1}{p_i})$$

Η μέγιστη τιμή της εντροπίας είναι $\log_2 n$ και εμφανίζεται αν και μόνο αν $p_i=1/n$ για κάθε i.

- Η συνδυασμένη εντροπία δύο τυχαίων μεταβλητών X, Y δίνεται από:

$$H(X,Y) = -\sum_{x,y} P(X = x, Y = y) \log_2 (P(X = x, Y = y))$$

Θεωρία Shannon

- Η υπό συνθήκη πιθανότητα να εστάλη κρυπτόγραμμα c, δοθέντος του αρχικού μηνύματος m, ισούται με:

$$p(c \mid m) = \sum_{k:E_k(m)=c} p_k$$

Όπου το κλειδί θεωρείται τυχαία μεταβλητή K, που μπορεί να πάρει διάφορες τιμές k με αντίστοιχες πιθανότητες p_k.

- Η πιθανότητα να εστάλη κρυπτόγραμμα c ισούται με:

$$p(c) = \sum_{\{k,m:E_k(m)=c\}} p_k p(m)$$

Θεωρία Shannon (II)

- Η υπό συνθήκη πιθανότητα να εστάλη μήνυμα m, δοθέντος του λαμβανομένου κρυπτογράμματος c, ισούται με:

$$p(m/c) = \frac{p(m)\,p(c\,|\,m)}{p(c)}$$

- Αν M συμβολίζει την τυχαία μεταβλητή που αντιστοιχεί στα πιθανά μηνύματα, τότε η υπό συνθήκη εντροπία του M δοθέντος συγκεκριμένου κρυπτογράμματος c (p(c) > 0) ισούται με:

$$\mathrm{H}(\mathrm{M}\,|\,c) = -\sum_{m} p(m\,|\,c)\log_2\, p(m\,|\,c)$$

(είναι γενίκευση του κλασικού ορισμού της εντροπίας, για υπό συνθήκη πιθανότητες)

Θεωρία Shannon (III)

- Η υπό συνθήκη εντροπία του M δοθέντος C (όπου C συμβολίζει την τυχαία μεταβλητή που αντιστοιχεί στα κρυπτογράμματα) ισούται με:

$$H(M \mid C) = \sum_{c} p(c) H(M \mid c)$$

- Η διαφορά H(M) – H(M|C) εκφράζει την ελάττωση της αβεβαιότητας του μηνύματος, που προέρχεται από γνώση του κρυπτογράμματος. Ονομάζεται και **αμοιβαία πληροφορία (mutual information)** και συμβολίζεται με I(M|C)

Ένα σύστημα χαρακτηρίζεται ως απολύτως ασφαλές αν γνώση του κρυπτογράμματος δεν προσφέρει καμία πληροφορία για το αρχικό μήνυμα. Αυτό σημαίνει p(m|c) = p(m) για όλα τα m και c. Με άλλα λόγια, οι μεταβλητές M, C είναι ανεξάρτητες, που σημαίνει I(M|C)=0.

Απόδειξη του Shannon για την απεριόριστη ασφάλεια του σημειωματαρίου μιας χρήσης

Για το one-time pad: Για κάθε μήνυμα $m=m_1m_2...m_n$ και κρυπτόγραμμα $c=c_1c_2\ldots c_n$ υπάρχει ένα μοναδικό κλειδί $k=k_1k_2\ldots k_n$ τέτοιο ώστε $E_k(m)=c$ (δηλαδή, αν κάνουμε xor το μήνυμα m με το κλειδί k θα πάρουμε το κρυπτόγραμμα c). Συνεπώς:

$$p(m \mid c) = \frac{p(m)\,p(c \mid m)}{p(c)} = \frac{p(m)\,p_k}{\sum_{m,k} p(m)\,p_k} = p(m)$$

Νόμος του Bayes (από θεωρία πιθανοτήτων)

αφού $p_k=1/2^n$ για όλα τα k και $\sum_m p(m) = 1$

Τι εννοούμε ως τυχαίο κλειδί?

- Πότε μία ακολουθία μπορεί να χρησιμοποιηθεί σαν κλειδί σε stream ciphers? Πρέπει να είναι «τυχαία». Τι εννοούμε όμως με τη λέξη «τυχαία»? (Για παράδειγμα, είναι προφανές διαισθητικά ότι η ακολουθία 11111111110 δεν μπορεί να θεωρηθεί τυχαία). Από την άλλη μεριά, μπορούμε να ορίσουμε πότε ακριβώς είναι τυχαία η ακολουθία?
- Έχουν προταθεί κάποια κριτήρια από τον Golomb, τα οποία αν τα πληρεί μία ακολουθία μπορεί να θεωρηθεί τυχαία.

Ιδιότητες ψευδοτυχαίων ακολουθιών (Αξιώματα Golomb)

Για μία ακολουθία $a_0a_1a_2a_3$…. Με περίοδο $N=2^n-1$, πρέπει να ισχύουν τα εξής:

- Ισοκατανεμημένο πλήθος 0 και 1 (Balance property)
- Αν ως διαδρομή (run) ορίζουμε ένα τμήμα οσοδήποτε μήκους μίας ακολουθίας που αποτελείται μόνο από μηδενικά ή μόνο από άσους, τότε σε μία περίοδο οι μισές διαδρομές έχουν μήκος 1, το ¼ των διαδρομών έχουν μήκος 2, το 1/8 μήκος 3 κ.ο.κ. Η ισχύς της συνθήκης εξετάζεται όσο ο αριθμός των διαδρομών είναι μεγαλύτερος ή ίσος από 2^l, όπου l το μήκος της διαδρομής. (run property)
- Η συνάρτηση αυτοσυσχέτισης

$$C(\tau) = \sum_{i=0}^{N-1} (-1)^{a_i \oplus a_{i+\tau}}$$

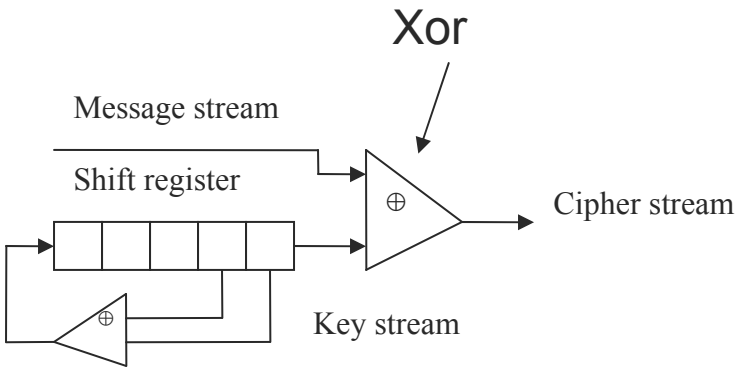
για την ακολουθία a_0a_1… περιόδου N μπορεί να πάρει μόνο δύο τιμές: να είναι σταθερή (ίση με K) για $\tau \neq 0$, και τιμή N για $\tau=0$. (two-level autocorrelation)

Παράδειγμα ψευδοτυχαίας ακολουθίας

- Έστω η ακολουθία 110100111101000 που επαναλαμβάνεται περιοδικά.
- Έχει περίοδο $15=2^4-1$.
- Σε μία περίοδο έχει 8 άσους και 7 μηδενικά, άρα είναι ισομοιρασμένα.
- Έχει 8 διαδρομές: Οι μισές (τέσσερις) έχουν μήκος 1 (το τρίτο, τέταρτο, ενδέκατο και δωδέκατο bit). Το ένα τέταρτο (δύο) των διαδρομών έχουν μήκος 2 (τα ζευγάρια bits 1-2 και 5-6). Το ένα όγδοο των διαδρομών (μία) έχει μήκος 3 (τα τρία τελευταία bits). Για διαδρομή μήκους 4 δεν εξετάζουμε, μια που $8 < 2^4$.
- Για τ=0, η συνάρτηση αυτοσυσχέτισης ισούται με 15, ενώ για οποιοδήποτε άλλο τ ισούται με -1.

Συστήματα στην πράξη

- Ως γεννήτρια ψευδοτυχαίων bits χρησιμοποιήθηκε αρχικά ένας **γραμμικός καταχωρητής ολίσθησης με ανάδραση (LFSR)**
- Έχουν καλή μαθηματική περιγραφή και οι ιδιότητές τους αναλύονται εύκολα

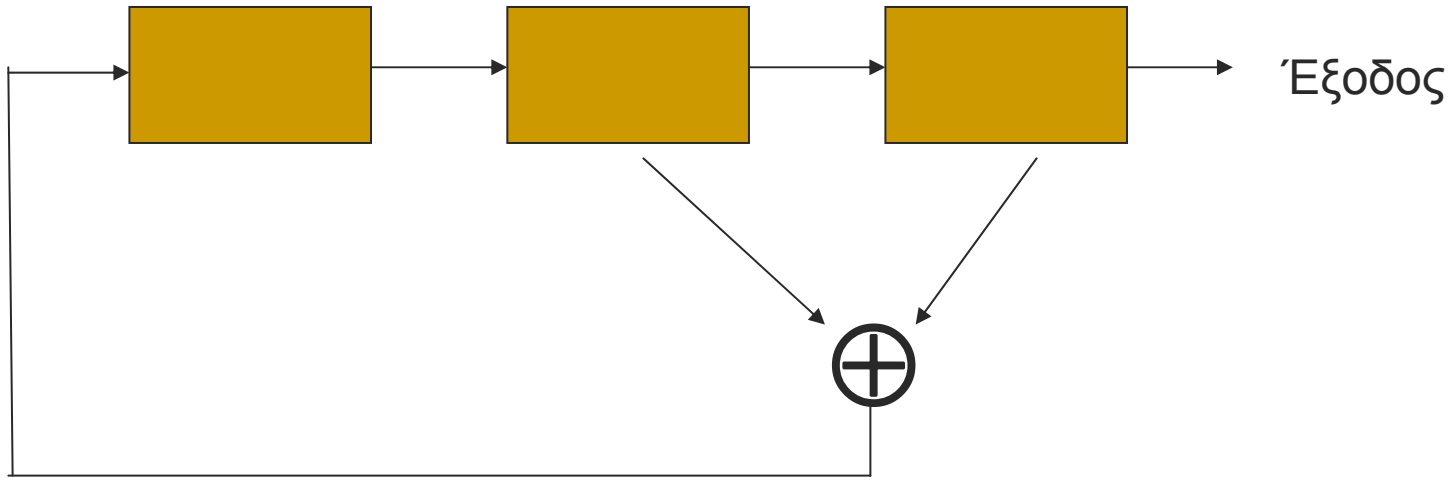


Λειτουργία ενός LFSR

- Αποτελείται από Ν βαθμίδες (θέσεις μνήμης): το περιεχόμενο κάθε μιας είναι είτε 0 είτε 1. Κάποιες από τις βαθμίδες αυτές γίνονται xor και το αποτέλεσμα πηγαίνει πίσω στην πρώτη βαθμίδα. Αν ο LFSR βρίσκεται σε μία κατάσταση (δηλαδή οι βαθμίδες του έχουν μία συγκεκριμένη τιμή), τότε η επόμενη κατάστασή του προσδιορίζεται εύκολα από τον ακόλουθο κανόνα:
  - Όλες οι βαθμίδες (η τιμή τους δηλαδή) ολισθαίνουν κατά μία θέση δεξιά
  - Η νέα τιμή για την πρώτη βαθμίδα είναι το αποτέλεσμα της παραπάνω XOR πράξης

Σχηματική αναπαράσταση της λειτουργίας ενός LFSR

Παράδειγμα: LFSR τριών βαθμίδων



Αν η αρχική κατάσταση είναι 001, τότε η έξοδος είναι 1 (η δεξιότερη βαθμίδα).

Την επόμενη χρονική στιγμή, η κατάσταση θα είναι 100 και η έξοδος 0. Το 100 προκύπτει ως εξής: το «1» είναι το XOR που είχαν αρχικά η δεύτερη και η τρίτη βαθμίδα (που ήταν 0 και 1 αντίστοιχα), ενώ το «00» είναι απλά ολισθημένες οι τιμές που είχαν αρχικά η πρώτη με τη δεύτερη βαθμίδα.

Σχηματική αναπαράσταση της λειτουργίας ενός LFSR (II)

- Στον προηγούμενο LFSR, αν θεωρήσουμε ότι η αρχική κατάσταση είναι η 001, οι διαδοχικές καταστάσεις από τις οποίες περνάει (και η αντίστοιχη έξοδος που παράγεται) είναι:

Κατάσταση	**Έξοδος**
00**1**	**1**
10**0**	**0**
01**0**	**0**
10**1**	**1**
11**0**	**0**
11**1**	**1**
01**1**	**1**
00**1**	**1**

Η 001 έχει ξαναεμφανιστεί, οπότε οι καταστάσεις επαναλαμβάνονται. Άρα, ο συγκεκριμένος LFSR παράγει την ακολουθία 1001011, η οποία επαναλαμβάνεται περιοδικά

Ιδιότητες LFSRs

- Ένα LFSR μήκους L μπορεί να περάσει από 2^L -1 διαφορετικές καταστάσεις, άρα μπορεί να γεννήσει ακολουθίες με μέγιστη περίοδο 2^L -1.
- Γενικά, η ακολουθία εξόδου ενός LFSR εξαρτάται τόσο από την ανάδρασή του όσο και από την αρχική του κατάσταση.
- **LFSRs που επιτυγχάνουν μέγιστη περίοδο ονομάζονται πρωταρχικοί (primitive). Οι ακολουθίες που παράγονται από τέτοιους καταχωρητές ονομάζονται ακολουθίες μέγιστου μήκους (maximal-length sequences ή m-sequences).** Ο LFSR στην προηγούμενη διαφάνεια παράγει προφανώς ακολουθία μέγιστου μήκους (έχει μήκος 3 και παράγει ακολουθία με μήκος 7). Η έξοδος έχει πάντα μέγιστη περίοδο σε πρωταρχικούς LFSRs, ανεξάρτητα της αρχικής τους κατάστασης.
- **Ορισμός:** Το μήκος του ελάχιστου LFSR που μπορεί να γεννήσει μία συγκεκριμένη ακολουθία ονομάζεται **γραμμική πολυπλοκότητα** (linear complexity) της ακολουθίας.

Παράρτημα: πότε ένας LFSR είναι πρωταρχικός (δηλαδή, πότε παράγει ακολουθίες μέγιστου μήκους? *

- Πρωταρχικό πολυώνυμο σε ένα πεπερασμένο σώμα GF(2^L) λέγεται κάθε ανάγωγο πολυώνυμο που διαρεί το $x^{2^L-1}-1$, αλλά δεν διαιρεί κανένα της μορφής x^n-1 για n < 2^L-1. Κάθε πεπερασμένο σώμα έχει τουλάχιστον ένα πρωταρχικό πολυώνυμο.
- Αν το πολυώνυμο ανάδρασης ενός LFSR με L βαθμίδες είναι πρωταρχικό ως προς το πεπερασμένο σώμα GF(2^L), τότε η παραγόμενη ακολουθία έχει περίοδο τη μέγιστη δυνατή, δηλαδή 2^L-1. Το αντίστροφο επίσης ισχύει.

* *Εκτός ύλης η συγκεκριμένη διαφάνεια*

Είναι κατάλληλος ένας LFSR?

- Παρόλο τις καλές ιδιότητες των LFSR (εύκολη υλοποίηση, παράγουν ακολουθίες με καλά χαρακτηριστικά ψευδοτυχαιότητας, «ελέγξιμοι» όσον αφορά το να επιλέγουμε LFSR που να παράγουν ακολουθίες με μέγιστη περίοδο), εν τούτοις δεν είναι καλή ιδέα να χρησιμοποιοιούνται έτσι απλά ως γεννήτριες κλειδοροής:
  - Αν ξέρουμε ένα τμήμα του μηνύματος, τότε στην ουσία ξέρουμε το αντίστοιχο τμήμα της κλειδοροής, άρα ξέρουμε το αντίστοιχο τμήμα της τρέχουσας κατάστασης του LFSR!
  - Πρέπει λοιπόν, ως σύστημα παραγωγής της κλειδοροής να έχουμε μία πιο σύνθετη δομή, έτσι ώστε ακόμα κι αν γνωρίζουμε τμήμα της κλειδοροής να μην μπορούμε να υπολογίσουμε τμήμα της κατάστασης της γεννήτριας.
  - Όπως θα δούμε αμέσως τώρα, δεν αρκεί μία σύνθετη δομή: πρέπει η ακολουθία που παράγεται να μην μπορεί να παραχθεί από LFSR μικρού μήκους!!

Berlekamp – Massey αλγόριθμος

- Για μία ακολουθία, η οποία μπορεί να προέκυψε από οποιονδήποτε τρόπο (π.χ. από μη γραμμικό FSRs), υπάρχουν πάντα κάποιοι LFSR που μπορούν επίσης να την παράγουν.
- Το μήκος του μικρότερου LFSR που την παράγει (δηλαδή η γραμμική της πολυπλοκότητα), παρουσιάζει εξαιρετικό ενδιαφέρον κρυπτογραφικά λόγω του ακόλουθου:
  - **Ο ελάχιστος LFSR που μπορεί να παράγει μία ακολουθία υπολογίζεται σε πολυωνυμικό χρόνο με τον αλγόριθμο Berlekamp-Massey (1969).**
  - **Αν η περίοδος μίας ακολουθίας είναι N και η γραμμική της πολυπλοκότητα είναι L<N/2, τότε ο μικρότερος LFSR μήκους L που την παράγει είναι μοναδικός. Για να υπολογιστεί από τον αλγόριθμο Berlekamp-Massey αυτός ο LFSR, του αρκεί να γνωρίζει 2L διαδοχικά bits της ακολουθίας.**

Πού καταλήγουμε?

- **Αν η γραμμική πολυπλοκότητα μίας ακολουθίας είναι L, τότε ο αλγόριθμος Berlekamp-Massey χρειάζεται μόνο 2L διαδοχικά bits για να υπολογίσει τον ελάχιστο LFSR που την παράγει!!**
- Συμπέρασμα: **Μία ακολουθία για να χρησιμοποιηθεί ως κλειδί στην κρυπτογραφία, πρέπει να έχει όσο γίνεται υψηλότερη γραμμική πολυπλοκότητα.**
  - Πράγματι, αν έχει μικρή γραμμική πολυπλοκότητα L, τότε αν ξέρουμε μόνο 2L διαδοχικά bits της ακολουθίας μπορούμε με τον αλγόριθμο Berlekamp-Massey να βρούμε τον ελάχιστου μήκους LFSR που την παράγει. Άρα, ο κρυπταναλυτής μπορεί να κατασκευάσει μία ισοδύναμη γεννήτρια της κλειδοροής (τον LFSR) οπότε να επιτύχει πλήρη ανάκτηση του μηνύματος

Ακολουθίες μεγίστου-μήκους: μπορούν να χρησιμοποιηθούν σαν κλειδιά σε κρυπτογραφικά συστήματα?

- Οι ακολουθίες μέγιστου μήκους (που παράγονται από πρωταρχικούς καταχωρητές) **πληρούν όλα τα κριτήρια ψευδοτυχαιότητας του Golomb**!! Επιπλέον, για μεγάλη τιμή του L, έχουν μεγάλη περίοδο (2^L-1). Όλες αυτές οι ιδιότητες τις κατέστησαν κατ' αρχήν εφαρμόσιμες στην κρυπτογραφία .
- Ωστόσο, αφού έχουν την ελάχιστη δυνατή γραμμική πολυπλοκότητα, ο αλγόριθμος Berlekamp-Massey τις καθιστά τελείως ακατάλληλες για την κρυπτογραφία!!!!
  - Για παράδειγμα, ας σκεφτούμε μία ακολουθία μέγιστου μήκους με περίοδο 2^{128}-1 (πολύ μεγάλη περίοδος – είναι αυτή που χρησιμοποιείται σήμερα). Αν και ψευδοτυχαία με βάση τα κριτήρια Golomb, η γραμμική της πολυπλοκότητα είναι μόλις 128 (υπάρχει δηλαδή LFSR μήκους 128 που την παράγει – συγκεκριμένα, ο αντίστοιχος πρωταρχικός LFSR με 128 βαθμίδες). Συνεπώς, αν κανείς γνωρίζει μόλις 256 διαδοχικά bits της κλειδοροής, μπορεί να υπολογίσει ολόκληρη τη κλειδοροή, δηλαδή όλα τα 2^{128}-1 bits!!!

Βιβλιογραφία LFSRs

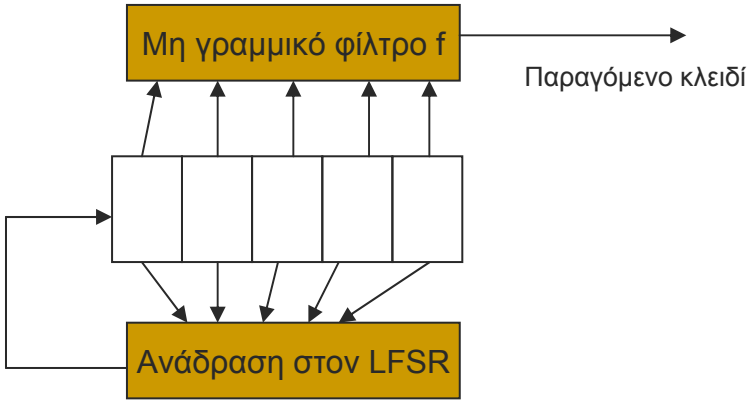
- “Shift Register Synthesis”, S.W. Golomb, Holden-Day, San Francisco, 1969
- “Finite Fields”, R. Lidl and H. Niederreiter, vol. 20 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1996 (chapter 8).

Αναφορές για τον αλγόριθμο Berlekamp-Massey:

1. “Algebraic Coding Theory”, E. R. Berlekamp, New York: McGraw-Hill, 1968
2. “Shift Register Sequences and BCH Decoding”, J. L. Massey, IEEE Trans. On Information Theory, vol. IT-15, pp. 122-127, Jan. 1969
3. Handbook of Applied Cryptography (chapter 6 – διαθέσιμο στο Internet)
4. Τεράστιο πλήθος ιστοσελίδων στο διαδίκτυο, σχετικά με τον αλγόριθμο (περιγραφή, υλοποιήσεις κτλ.)

Γεννήτριες ψευδοτυχαίων ακολουθιών – μη γραμμικά φίλτρα

- Για να πάρουμε ακολουθίες με μεγάλη γραμμική πολυπλοκότητα, καταφεύγουμε στην πράξη σε κάποιο από τα ακόλουθα σχήματα:
  1. Εφαρμογή μη γραμμικής συνάρτησης f (φίλτρου) στις βαθμίδες ενός LFSR (filter function generators – Key (1973))

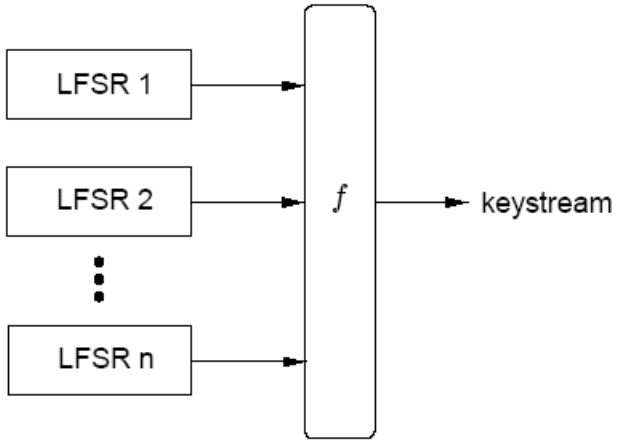


Η μέγιστη τιμή που μπορεί να έχει η γραμμική πολυπλοκότητα του παραγόμενου κλειδιού (keystream) είναι: $L_m = \sum_{i=1}^{m} \binom{L}{i}$

όπου m το μέγιστο πλήθος βαθμίδων του LFSR που συνδυάζονται σε ένα γινόμενο από την f

Γεννήτριες ψευδοτυχαίων ακολουθιών – συνδυαστικές συναρτήσεις

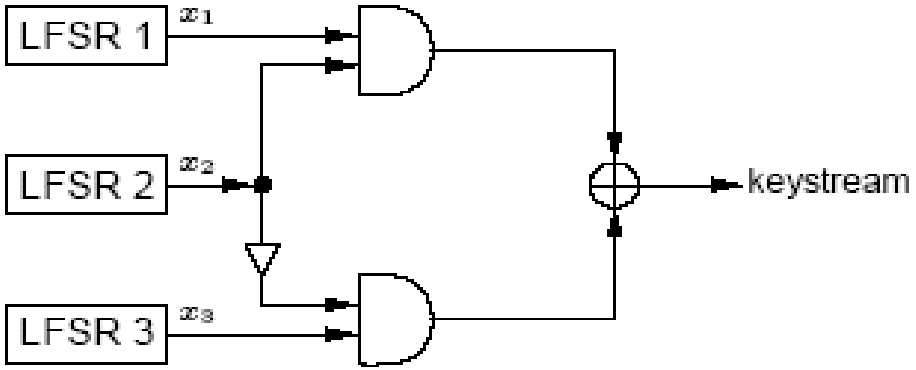
2. Οι έξοδοι πολλών LFSRs περνούν από μία μη γραμμική συνάρτηση (combinatorial function generators)



Αν $L_1, L_2,\ldots,L_n$ είναι οι γραμμικές πολυπλοκότητες (ανά δύο διαφορετικές μεταξύ τους) των ακολουθιών που παράγονται από τους καταχωρητές, τότε η γραμμική πολυπλοκότητα του παραγόμενου κλειδιού (keystream) ισούται με: $f(L_1,L_2,\ldots,L_n)$

Για παράδειγμα, αν n=4 και $f=x_1x_3 + x_2 + x_2x_3x_4$, τότε: $L_{keystream}=L_1L_3 + L_2 + L_2L_3L_4$

Παράδειγμα – Geffe generator



- Ο κάθε LFSR έχει μήκος L_i, i=1,2,3, πρώτα μεταξύ τους, και μέγιστη περίοδο 2^{Li}-1
- Έξοδος $z(t)=x_1x_2 \oplus x_2x_3 \oplus x_3$
- Η ακολουθία εξόδου z έχει μεγάλη περίοδο: $(2^{L_1}-1)(2^{L_2}-1)(2^{L_3}-1)$
- Η γραμμική πολυπλοκότητα της εξόδου ισούται με: $L=L_1L_2 + L_2L_3 + L_3$

Ασφάλεια Geffe generator

- Όχι ασφαλής: Ας παρατηρήσουμε ότι:

$$P(z(t)=x_1(t))=P(x_2(t)=1)+P(x_2(t)=0)P(x_3(t)=x_1(t))=$$

$$=\frac{1}{2}+\frac{1}{2}\frac{1}{2}=\frac{3}{4}$$

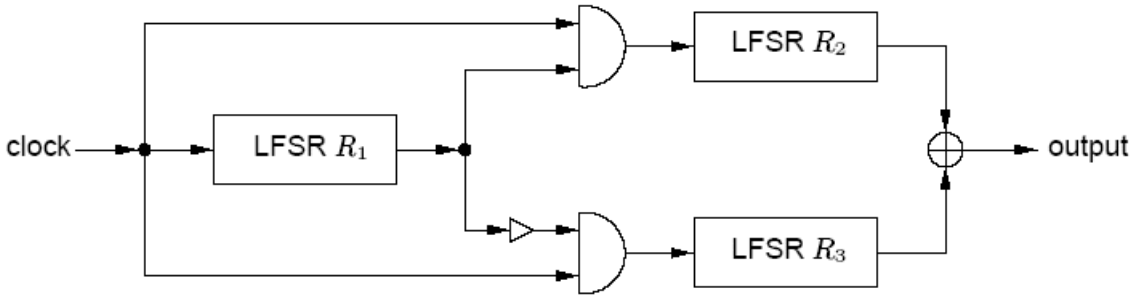
Ομοίως μπορεί να δειχτεί ότι $P(z(t)=x_3(t))=3/4$

- Υπόκειται σε correlation κρυπτανάλυση λόγω της παραπάνω παρατήρησης (Μέθοδος που στηρίζεται στην εξάρτηση της εξόδου κάποιας συνάρτησης από κάποιες συγκεκριμένες εισόδους). Βασική της ιδέα είναι η εξής: αν η έξοδος ταυτίζεται με την έξοδο κάποιου από τους καταχωρητές με πιθανότητα $p > ½$, τότε αν ένα ικανοποιητικά μεγάλο τμήμα του κλειδιού είναι γνωστό, μπορεί να ευρεθεί η αρχική κατάσταση του εν λόγω καταχωρητή.
- Συμπέρασμα: η συνάρτηση f σε τέτοιου τύπου γεννήτριες ψευδοτυχαίων ακολουθιών δεν πρέπει να εμφανίζει στατιστική εξάρτηση της εξόδου με κάποιο υποσύνολο των εισόδων της.

  (αυτές οι συναρτήσεις χαρακτηρίζονται ως ανεπηρέαστες στη συσχέτιση – correlation immune)

Γεννήτριες ψευδοτυχαίων ακολουθιών – ελεγχόμενες από ρολόι

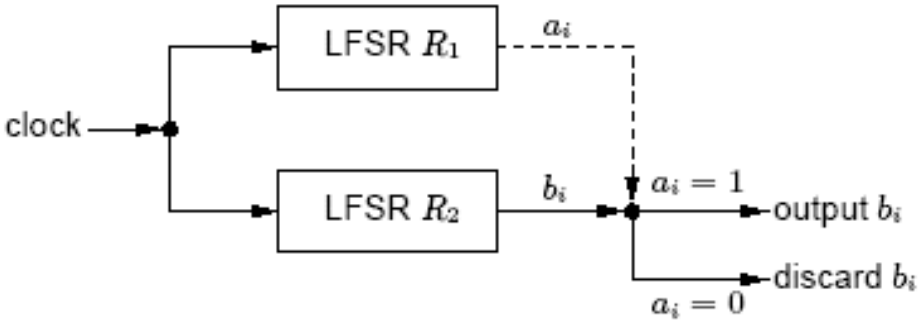
3. Το ρολόι ενός LFSR είναι η έξοδος κάποιου άλλου LFSR (clocked-controlled generators)



Για καλή έξοδο, οι καταχωρητές πρέπει να είναι όλοι πρωταρχικοί, και παραπλήσιου μεγέθους μεταξύ τους.

Γεννήτριες ψευδοτυχαίων ακολουθιών – γεννήτριες σμίκρυνσης (shrinking generators)

4. Ένας LFSR καθορίζει το αν θα λαμβάνεται ή όχι υπ'όψιν η έξοδος ενός άλλου LFSR (shrinking generator – Coppersmith – Krawczys – Mansour (1993))



Αν L_1, L_2 οι γραμμικές πολυπλοκότητες των R_1, R_2 αντιστοίχως, τότε για τη γραμμική πολυπλοκότητα L του παραγομένου κλειδιού ισχύει η σχέση:

$$L_2 \cdot 2^{L_1-2} \leq L \leq L_2 \cdot 2^{L_1-1}$$

Αναφορές (για τις προηγούμενες μεθόδους)

- “Generation of binary sequences with controllable complexity”, E. J. Groth, IEEE Trans. Information Theory, vol. 17, pp. 288-296, 1971
- “An analysis of the structure and complexity of nonlinear binary key generators”, E. L. Key, IEEE Trans. Information Theory, vol. 22, pp. 732-736, 1976.
- “The stop-and-go generator”, T. Beth and F. C. Piper, Proceedings of Eurocrypt 84, pp. 88-92, 1985.
- “The shrinking generator”, Coppersmith – Krawczys – Mansour, Advances in Cryptology – Crypto ’93, pp. 22-94, 1994

Αρχή του Kerchoff στους κρυπταλγόριθμους ροής

- Σε όλα τα συστήματα στην πράξη, ο τρόπος γέννησης της ψευδοτυχαίας ακολουθίας της κλειδοροής είναι γνωστός (δηλαδή το ποιοι ακριβώς καταχωρητές χρησιμοποιούνται, με τι συναρτήσεις κ.ο.κ.).
- Αυτό που παραμένει κρυφό (και στην ουσία παρέχει την ασφάλεια) είναι η αρχική κατάσταση των καταχωρητών. Αυτό είναι και το κλειδί στους αλγορίθμους ροής. Βέβαια, οι καταχωρητές που χρησιμοποιούνται στην πράξη είναι πολλών βαθμίδων (το λιγότερο 128), έτσι ώστε να μη μπορεί να βρει κανείς την αρχική κατάστασή τους με εξαντλητικές δοκιμές.

Κάποιοι σημαντικοί stream ciphers που χρησιμοποιούνται σήμερα

- **SEAL** (το paper υπάρχει στο: http://www.cs.ucdavis.edu/~rogaway/papers/seal-abstract.html)
- **RC4** (χρησιμοποιείται σε πολλά προϊόντα, όπως Oracle Secure SQL, Apple's Computer's AOCE). Επίσης στο πρωτόκολλο SSL στο Internet.

  http://www.wisdom.weizmann.ac.il/~itsik/RC4/rc4.html για περιγραφή και περαιτέρω αναφορές
- **A5/1** (χρησιμοποιείται στις GSM επικοινωνίες) (http://cryptome.org/a51-bsw.htm)
- **E0** (χρησιμοποιείται στην bluetooth επικοινωνία)
- Λίστα stream ciphers, με πολύ βιβλιογραφικό υλικό: http://homes.esat.kuleuven.be/~jlano/stream/designs.htm
- Πολύ πρόσφατοι αλγόριθμοι (από το eStream Project) που «άντεξαν» σε όλες τις γνωστές κρυπταναλυτικές επιθέσεις: Trivium, Grain, MICKEY, HC-128, Rabbit, Salsa 20/12, Sosemanuk (με βάση την αναφορά από το eStream Project, Σεπτέμβριος 2008)