William Stallings
Data and Computer Communications

Ανίχνευση και διόρθωση σφαλμάτων

Εισαγωγή

- Κανένα επικοινωνιακό σύστημα δεν είναι τέλειο.
  - Η ύπαρξη θορύβου στο κανάλι επικοινωνίας, αλλά και στα άλλα τμήματα ενός επικοινωνιακού συστήματος, οδηγεί με στατιστική βεβαιότητα στη δημιουργία σφαλμάτων.
- Οι συνέπειες της δημιουργίας σφαλμάτων σε ένα μεταδιδόμενο σήμα εξαρτώνται από τη φύση του σήματος αυτού.
  - Υπάρχουν περιπτώσεις - όπως για παράδειγμα στη μετάδοση ενός ραδιοφωνικού σήματος - όπου η ύπαρξη σφαλμάτων και η παραμόρφωση του μεταδιδόμενου σήματος είναι μέχρι ενός σημείου αποδεκτή.
- Στην περίπτωση όμως της επικοινωνίας υπολογιστικών συστημάτων συνήθως απαιτείται διασφάλιση της ορθότητας των μεταδιδόμενων δεδομένων σε απόλυτο βαθμό.

Εισαγωγή

- Ένα σφάλμα συμβαίνει όταν η τιμή ενός bit αλλάζει μεταξύ μετάδοσης και λήψης
- Λάθη μονού bit (Single bit errors)
  - Αλλαγή σε ένα bit
  - Παρακείμενα bit δεν επηρεάζονται
- Λάθη ριπής (Burst errors)
  - Μήκος *B*
  - Ακολουθία από *B* bits που περιέχουν εσφαλμένα bits
    - Όχι κατ'ανάγκη όλα τα bits εσφαλμένα
  - Προκαλούνται από
    - Κρουστικό θόρυβο
    - Σκίαση σε ασύρματα περιβάλλοντα
    - Μεγαλύτερη επίδραση σε υψηλότερους ρυθμούς δεδομένων
  - Π.χ κρουστικός θόρυβος ή σκίαση διάρκειας 1 μs προκαλεί αντιστροφή 10 bits σε ρυθμό μετάδοσης 10 Mbps και 100 bits σε ρυθμό μετάδοσης 100 Mbps

Εισαγωγή

- Αν υποθέσουμε ότι τα δεδομένα μεταδίδονται σε **ομάδες (πλαίσια) των F bits** και ότι P_b είναι η πιθανότητα να φθάσει ένα bit στον αποδέκτη με σφάλμα (Bit Error Rate, BER),
  - τότε η πιθανότητα να φθάσει ολόκληρο το πλαίσιο χωρίς κανένα σφάλμα είναι:

$$P_c=(1- P_b)^F$$

- Η πιθανότητα να φθάσει το πλαίσιο με ένα ή περισσότερα σφάλματα είναι:

$$P_e=1-P_c = 1- (1- P_b)^F$$

Εισαγωγή

- Υπάρχουν δύο δυνατοί τρόποι για την εξασφάλιση της ορθότητας των μεταδιδόμενων δεδομένων σε ένα επικοινωνιακό σύστημα:
- **α) Η ανίχνευση του σφάλματος και η επαναμετάδοση** της ομάδας δεδομένων στην οποία ανιχνεύτηκε το σφάλμα.
  - Η ανίχνευση του σφάλματος βασίζεται στην επισύναψη στο μεταδιδόμενο μήνυμα πρόσθετης πληροφορίας.
  - Η επεξεργασία της πληροφορίας αυτής σε συνδιασμό με το ληφθέν μήνυμα επιτρέπει στον αποδέκτη να διαπιστώσει την ύπαρξη ή μη σφάλματος.
  - Ο υπολογισμός της πρόσθετης πληροφορίας από τον μεταδότη καθώς και η επεξεργασία της πληροφορίας αυτής από τον αποδέκτη πραγματοποιούνται με βάση αλγόριθμους κωδικοποίησης που ονομάζονται κώδικες ανίχνευσης σφάλματος (error detection codes).

Εισαγωγή

- **β) Η ανίχνευση του σφάλματος και η διόρθωσή του**, χωρίς να απαιτηθεί η επαναμετάδοση του μηνύματος.
  - Στην περίπτωση αυτή, η πρόσθετη πληροφορία που επισυνάπτεται στο μήνυμα επιτρέπει όχι μόνο την ανίχνευση, αλλά και τη διόρθωση του σφάλματος.
  - Οι αλγόριθμοι κωδικοποίησης που χρησιμοποιούνται για τον υπολογισμό της πρόσθετης πληροφορίας από το μεταδότη αλλά και για την επεξεργασία της πληροφορίας αυτής από τον αποδέκτη προκειμένου να εντοπιστούν και να διορθωθούν τυχόν σφάλματα ονομάζονται κώδικες διόρθωσης σφάλματος (error correction codes).

Εισαγωγή

- Ο **ρυθμός πληροφορίας** ενός κώδικα είναι το ποσοστό των ψηφίων της κωδικής λέξης που μεταφέρουν αρχική πληροφορία (δηλ. ψηφία πληροφορίας πριν την κωδικοποίηση).
  - Ο ρυθμός πληροφορίας ενός δυαδικού κώδικα C μήκους n είναι ίσος με $(1/n)\log_2|C|$, κείται δε μεταξύ 0 (όταν $|C|=1$) και 1 (όταν $|C| = 2^n$)
    - π.χ. εάν C1 = {00, 10, 01, 11}, τότε ο ρυθμός πληροφορίας είναι: $(1/n)\log_2|C1| = (1/2)\log_2 4 = 1$).
    - Εάν C2 = {000, 101, 011, 110}, τότε ο ρυθμός πληροφορίας είναι: $(1/n)\log_2|C2| = (1/3)\log_2 4 = 2/3$).

  Έστω ότι μεταδίδεται η κωδική λέξη 10 και παραλαμβάνεται 11. Δεν υπάρχει δυνατότητα ανίχνευσης στην περίπτωση του C1.
  - Αν όμως χρησιμοποιήσουμε τον C2 όπου το τελευταίο ψηφίο κάθε κωδικής λέξης είναι το ψηφίο ελέγχου ισοτιμίας τότε όταν στείλουμε την κωδική λέξη 000 και παραλάβουμε τη λέξη 001, αυτή δεν αποτελεί κωδική λέξη και έτσι καταλαβαίνουμε ότι παραλάβαμε λάθος.
  - Δεν είναι πάντοτε δυνατή η ανίχνευση και ακόμα δυσκολότερη η διόρθωση

Εισαγωγή

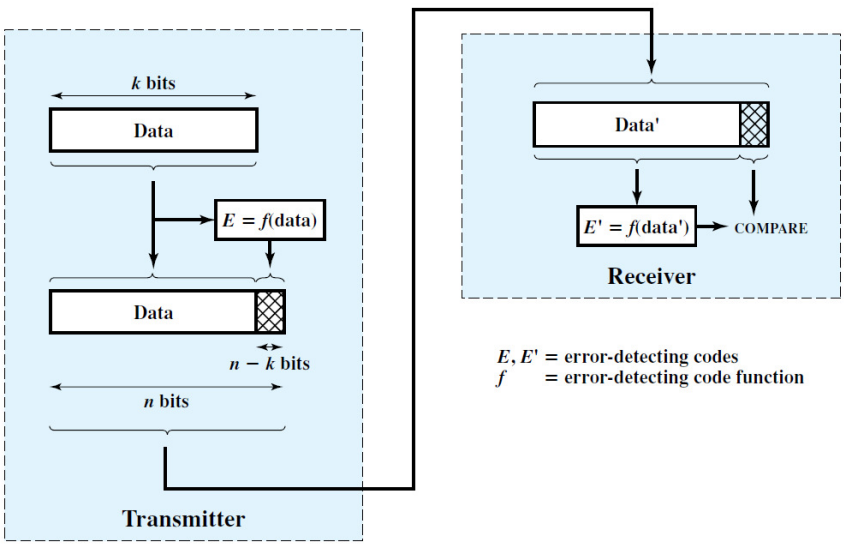
- **Ορισμός**: ***Βάρος (Weight) Hamming*** ή απλά βάρος, wt(x), μιας λέξης x μήκους n ψηφίων ονομάζεται το πλήθος των ψηφίων της λέξης, τα οποία είναι ίσα με το «1». Το βάρος παίρνει τιμές από 0 έως n
  - **Παράδειγμα:** Για x_1 = 000000, x_2 = 000010, x_3 = 110001 έχουμε: $wt(x_1)=0$, $wt(x_2)=1$, $wt(x_3)=3$
- **Ορισμός**: ***Απόσταση (Distance) Hamming*** ή απλά απόσταση, d(x, y), μεταξύ δύο λέξεων x και y του ιδίου μήκους n ονομάζεται το πλήθος των θέσεων, στις οποίες οι δύο λέξεις εμφανίζουν ασυμφωνία του δυαδικού ψηφίου.
  - Η απόσταση παίρνει τιμές από 0 έως n (όπου n το μήκος του κώδικα).
- Ως ***Απόσταση ενός κώδικα*** **C** ορίζεται η μικρότερη από τις (Hamming) αποστάσεις όλων των δυνατών ζευγών κωδικών λέξεων του κώδικα.

Εισαγωγή

Γενικά για κωδικοποίηση

- Πρώτα επιλέγεται ένας θετικός ακέραιος k (το μήκος κάθε δυαδικής λέξης που αντιστοιχεί σε ένα μήνυμα).
- Στη συνέχεια, επιλέγεται το πλήθος των n-k δυαδικών ψηφίων που θα προστεθούν σε κάθε λέξη (πλεονασμός) έτσι ώστε να μπορεί να ανιχνεύεται ή και να διορθώνεται το επιθυμητό πλήθος σφαλμάτων. **Έτσι προκύπτουν οι κωδικές λέξεις που αντιστοιχούν στα δυνατά μηνύματα, μήκους n ψηφίων**,
  - Τα n-k bits πλεονασμού, ονομάζονται bits ελέγχου σφάλματος ή bits ισοτιμίας (parity)

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ



ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

- **Ανίχνευση Σφαλμάτων:**

Πρότυπο σφάλματος είναι ένας δυαδικός αριθμός **ε** με αριθμό bits ίσο με αυτόν του μηνύματος, με «1» στις θέσεις που υπάρχει σφάλμα και «0» στις υπόλοιπες θέσεις.

- ***Θεώρημα Ανίχνευσης Σφαλμάτων***
  - Ένας κώδικας **C απόστασης d ανιχνεύει όλα τα μη μηδενικά πρότυπα σφάλματος ε βάρους μικρότερου ή ίσου του d – 1**, ενώ υπάρχει τουλάχιστον ένα πρότυπο ε βάρους d που δεν ανιχνεύει ο κώδικας C.

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

- **Ανίχνευση Σφαλμάτων:**

- **Παράδειγμα:** Ο κώδικας C = {000, 111} έχει απόσταση d= wt(000+111) = wt(111) =3
- Άρα ανιχνεύεται από τον κώδικα κάθε πρότυπο σφάλματος ε βάρους 1 ή 2 (όπως τα 011 , 100 κλπ.).
- Δεν ανιχνεύεται όμως το πρότυπο σφάλματος «111» βάρους 3 (π.χ. y= 000+ε= 000+111=111 $\in$ C)

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Ελεγχος Ισοτιμίας

- Ένα απλός τρόπος ανίχνευσης σφάλματος σε ένα επικοινωνιακό σύστημα είναι με την προσθήκη ενός επιπλέον bit σε κάθε μεταδιδόμενο μήνυμα.
  - —Η τιμή του bit αυτού είναι τέτοια ώστε: ο συνολικός αριθμός των bits του μηνύματος (συμπεριλαμβανομένου και του πρόσθετου bit) που έχουν την τιμή "1"
    - να είναι άρτιος (άρτια ισοτιμία)
    - ή περιττός (περιττή ισοτιμία).

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Ελεγχος Ισοτιμίας

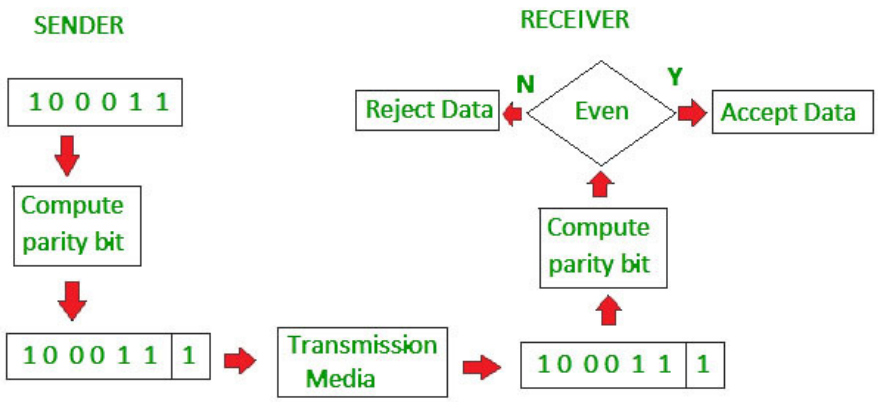
- Ο αποδέκτης εξετάζει τον συνολικό αριθμό των "1" του λαμβανόμενου μηνύματος.
  - — Αν πληρείται ο εφαρμοζόμενος κανόνας ισοτιμίας το μήνυμα γίνεται αποδεκτό.
  - — Αν το μήνυμα παραβιάζει τον εφαρμοζόμενο κανόνα ισοτιμίας τότε το μήνυμα απορρίπτεται ως εσφαλμένο και ζητείται από τον μεταδότη η επαναμετάδοσή του.
- Η μέθοδος της ισοτιμίας είναι απλή στην υλοποίησή της και δεν επιβαρύνει το επικοινωνιακό σύστημα με μεγάλο αριθμό πρόσθετων bits.
- Δεν είναι όμως ιδιαίτερα ισχυρή καθώς αδυνατεί να ανιχνεύσει άρτιο αριθμό σφαλμάτων.

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Ελεγχος Ισοτιμίας-Παράδειγμα

- Ο πομπός θέλει να μεταδώσει το μήνυμα 100011.
  - —Με χρήση **άρτιας ισοτιμίας** προσθέτει στο τέλος ένα 1 και στέλνει το 1000111.
- Ο δέκτης εξετάζει το εισερχόμενο μήνυμα (έστω 1100111) το οποίο έχει περιττό αριθμό 1
  - —και διαπιστώνει οτι υπήρξε σφάλμα στη λήψη
- Για λήψη του 1100101 (2 εσφαλμένα bits) ο δέκτης δεν διαπιστώνει σφάλμα λήψης

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ



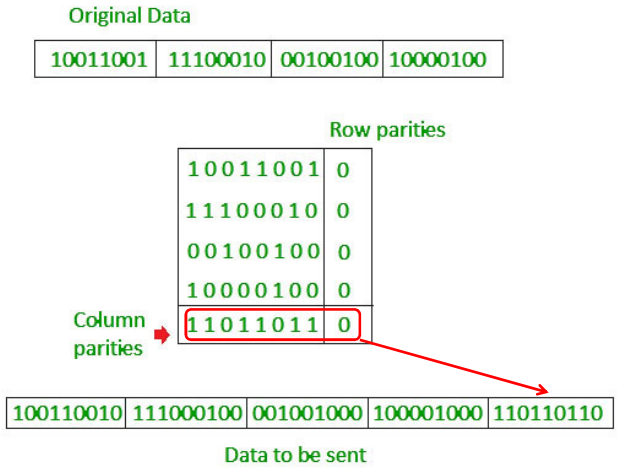
ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Διδιάστατη Ισοτιμία

- Τα προς μετάδοση δεδομένα ομαδοποιούνται σε **δισδιάστατους πίνακες**.
- **Σε κάθε γραμμή και σε κάθε στήλη του πίνακα προστίθεται ένα bit ισοτιμίας**
  - το οποίο ελέγχει με περιττή ή άρτια ισοτιμία τη συγκεκριμένη γραμμή ή στήλη.
- Όταν όλα τα δεδομένα του πίνακα (μαζί με τα bits ισοτιμίας) φθάσουν στον αποδέκτη,
  - τότε ανασυγκροτείται ο πίνακας
  - και εξετάζονται τα bits ισοτιμίας που ελέγχουν την κάθε γραμμή και την κάθε στήλη του πίνακα.

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Διδιάστατη Ισοτιμία



ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Κυκλικοί κώδικες (Cyclic Redundancy Check, CRC)

- Για κάθε μπλοκ δεδομένων D των k bits, ο μεταδότης δημιουργεί μια ακολουθία ελέγχου σφάλματος F των n-k bits (Frame Check Sequence, FCS),
  - τέτοια ώστε **η συνολική ακολουθία T των n bits που προκύπτει να διαιρείται ακριβώς με κάποιον προκαθορισμένο αριθμό P των n-k+1 bits.**
- Όταν η ακολουθία T των n bits φθάσει στον αποδέκτη,
  - τότε η ορθότητά της ελέγχεται διαιρώντας την με τον προκαθορισμένο αριθμό.
  - **Αν από τη διαίρεση αυτή δεν προκύψει υπόλοιπο**, τότε το πλαίσιο γίνεται αποδεκτό.
  - **Αν προκύψει υπόλοιπο**, τότε συνάγεται ότι το πλαίσιο έχει αλλοιωθεί και ζητείται η επαναμετάδοσή του.

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

- Αξίζει να σημειωθεί ότι για τον υπολογισμό του FCS καθώς και για την εξακρίβωση της ορθότητας του ληφθέντος πλαισίου **χρησιμοποιείται αριθμητική modulo-2.**
  - Δηλαδή δυαδική αριθμητική στην οποία όμως δεν υπάρχουν κρατούμενα ή δανεικά.
- **Λόγοι χρησιμοποίησης αριθμητικής modulo-2**
  - η απλότητα που χαρακτηρίζει την αριθμητική αυτή και η συνεπαγόμενη ευκολία στην υλοποίησή της.
  - Η διαίρεση modulo-2 αφήνει υπόλοιπο κατά 1 bit μικρότερο σε σχέση με την κανονική διαίρεση
    - γεγονός που οδηγεί σε ελαφρά μείωση των μεταδιδόμενων bits ελέγχου σφαλμάτων που επιβαρύνουν το επικοινωνιακό σύστημα.

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

modulo-2 αριθμητική

- Δυαδική πρόσθεση/αφαίρεση χωρίς κρατούμενο
  - —**Ουσιαστικά πρόκειται για την πράξη XOR**

```
  1111          1111          11001
 +1010         -1010            x11
 -------       -------       ---------
  0101          0101          11001
                              11001
                              -----------
                              101011
```

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Υπολογισμός CRC

- Έστω:
  - —**D** η προς μετάδοση ακολουθία δεδομένων των **k bits**,
  - —**F** η ακολουθία FCS των **n-k bits**,
  - —**T** η ακολουθία των **n bits** που πρόκειται να μεταδοθεί
  - —**P** ο προκαθορισμένος αριθμός των **n-k+1 bits** με τον οποίο θα πρέπει να είναι διαιρέσιμη η ακολουθία T.
    - Θα πρέπει προφανώς πρώτο και τελευταίο bit του να είναι 1
- Η ακολουθία **T** μπορεί να γραφτεί ως: $T = 2^{n-k} D + F$
  - —Πριν μεταδοθεί η ακολουθία T πρέπει να υπολογιστεί ακολουθία ελέγχου F.

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Υπολογισμός CRC

- Για τον υπολογισμό της ακολουθίας ελέγχου σφάλματος **F** χρησιμοποιείται ο παρακάτω αλγόριθμος:
  - —Τοποθετούμε n-k μηδενικά στα δεξιά του D έτσι ώστε να προκύψει το 2^{n-k} D
  - —Διαιρούμε το 2^{n-k} D με τον P
  - —Χρησιμοποιούμε ως **F** το **υπόλοιπο R** της παραπάνω διαίρεσης

$$\frac{2^{n-k}D}{P} = Q + \frac{R}{P}$$

$$T = 2^{n-k}D + R$$

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Υπολογισμός CRC

- Μπορεί να αποδειχθεί ότι η διαίρεση (modulo-2) T/P δεν αφήνει υπόλοιπο. Έχουμε:

$$\frac{T}{P} = \frac{2^{n-k}D + R}{P} = \frac{2^{n-k}D}{P} + \frac{R}{P}$$

Επειδή όμως όπως είδαμε στην προηγούμενη διαφάνεια:

$$\frac{2^{n-k}D}{P} = Q + \frac{R}{P}$$

Προκύπτει ότι:

$$\frac{T}{P} = Q + \frac{R}{P} + \frac{R}{P} = Q$$

($\frac{R}{P} + \frac{R}{P} \to 0$)

Ο.Ε.Δ.

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Υπολογισμός CRC

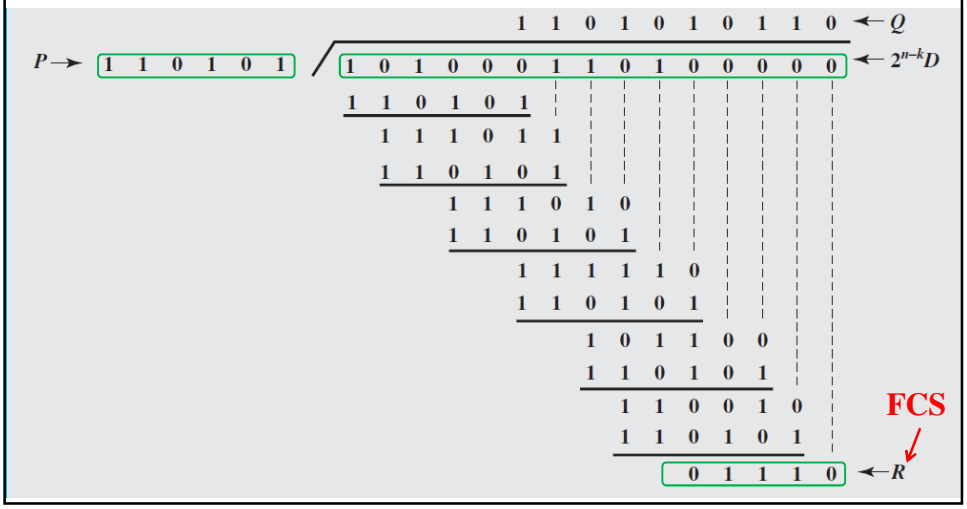
ΠΑΡΑΔΕΙΓΜΑ ΥΠΟΛΟΓΙΣΜΟΥ ΤΟΥ FCS ΣΤΟ ΜΕΤΑΔΟΤΗ ΚΑΙ ΕΛΕΓΧΟΥ ΟΡΘΟΤΗΤΑΣ ΣΤΟΝ ΑΠΟΔΕΚΤΗ

Έστω: $D = 1010001101$ (10 bits)
$P = 110101$ (6 bits)
$n = 15, k = 10$ και άρα $(n - k) = 5$.

ΜΕΤΑΔΟΤΗΣ: ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ R (5 bits) ΩΣΤΕ ΝΑ ΧΡΗΣΙΜΟΠΟΙΗΘΕΙ ΩΣ ΑΚΟΛΟΥΘΙΑ ΕΛΕΓΧΟΥ ΣΦΑΛΜΑΤΟΣ **F**

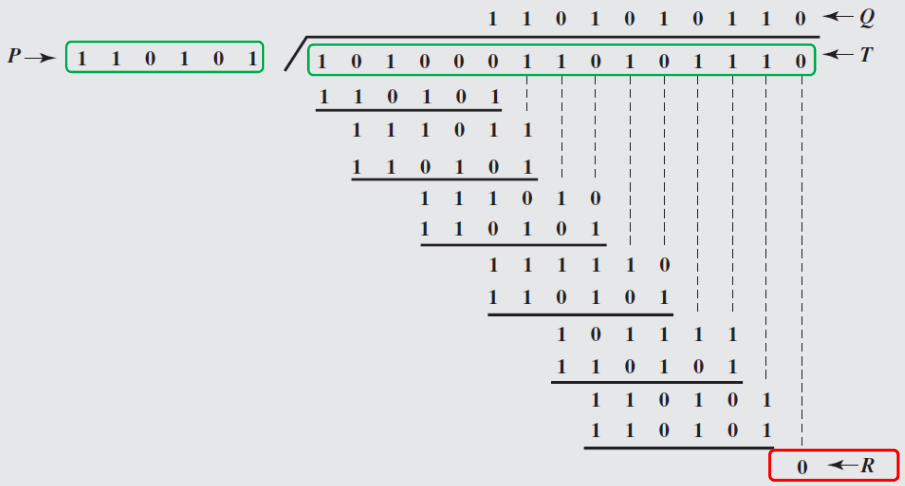
ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

ΥΠΟΛΟΓΙΣΜΌΣ ΤΟΥ FCS ΣΤΟ ΜΕΤΑΔΟΤΗ



ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

ΕΛΕΓΧΟΣ ΣΤΟΝ ΑΠΟΔΕΚΤΗ



ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Απόδοση CRC-σχόλια

- Έστω οτι ο πομπός μετέδωσε το μήνυμα T και ο δέκτης έλαβε το T_r
  - Για να ΜΗΝ ανιχνεύσει ο δέκτης το σφάλμα, θα πρέπει το T_r να διαιρείται από το P
    - Όμως T_r=T XOR E, όπου E το πρότυπο σφάλματος (διάνυσμα με 1 στη θέση όπου έχει συμβεί bit error και 0 στις υπόλοιπες)
  - Ισοδύναμα, θα πρέπει το T XOR E να διαιρείται από το P
    - Δηλ. το E να διαιρείται από το P
      - Κάτι που και διαισθητικά καταλαβαίνουμε οτι είναι σπάνιο

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Οι κυκλικοί κώδικες ως πράξεις πολυωνύμων

- Ένας εναλλακτικός τρόπος να περιγραφεί η διαδικασία του CRC είναι η έκφραση όλων των δυαδικών αριθμών που παίρνουν μέρος σε αυτήν
  - ως πολυωνύμων μιας ψευδομεταβλητής X των οποίων συντελεστές είναι τα ψηφία των δυαδικών αριθμών.
- Για παράδειγμα,
  - ο αριθμός D=110011 αναπαριστάται ως: $D(X)=X^5+X^4+X+1$
  - ενώ ο αριθμός P=11001 μπορεί να αναπαρασταθεί ως: $P(X)=X^4+X^3+1$.
- Οι πράξεις μεταξύ των δυαδικών αριθμών αναπαριστώνται ως πράξεις πολυωνύμων.
- Οι πράξεις είναι πάντα modulo-2.

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Οι κυκλικοί κώδικες ως πράξεις πολυωνύμων

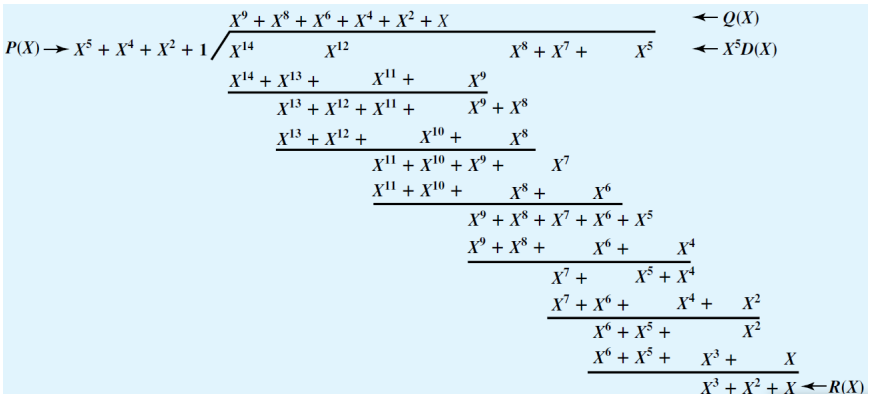
- Αν υιοθετήσουμε αυτόν τον εναλλακτικό τρόπο αναπαράστασης η διαδικασία του CRC μπορεί να περιγραφεί ως:

$$\frac{X^{n-k}D(X)}{P(X)} = Q(X) + \frac{R(X)}{P(X)}$$

$$T(X) = X^{n-k}D(X) + R(X)$$

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Παράδειγμα διαίρεσης πολυωνύμων



ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Ανιχνεύσιμα είδη σφαλμάτων

- Με κατάλληλη επιλογή του P(X) όλα τα παρακάτω είδη σφαλμάτων καθιστούν μια λανθασμένη λήψη του T(X) μη διαιρέσιμη διά P(X) και άρα είναι ανιχνεύσιμα:
  - Όλα τα σφάλματα σε ένα μόνο bit.
  - Όλα τα λάθη σε δύο bits, εφόσον το P(X) είναι ένα ειδικό είδος πολυωνύμου που καλείται πρωταρχικό (primitive) πολυώνυμο με μέγιστο εκθέτη L και το μήκος πλαισίου είναι μικρότερο από 2^L-1
  - Κάθε περιττό αριθμό σφαλμάτων, εφόσον το P(X) περιέχει έναν συντελεστή (X+1).
  - Κάθε μαζικό (bursty) σφάλμα σε συνεχόμενα bits, εφόσον ο αριθμός των bits του μαζικού σφάλματος είναι μικρότερος ή ίσος από το μήκος του FCS (δηλ. μικρότερος ή ίσος από το n-k).
  - Ποσοστό των μαζικών λαθών μήκους μεγαλύτερου ή ίσου με το μήκος του P(X) (δηλ. n-k+1).

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Τα πιο συχνά χρησιμοποιούμενα CRC είναι τα ακόλουθα:

- CRC-12: $P(X) = X^{12}+X^3+X^2+X+1$
- CRC-16: $P(X) = X^{16}+X^{15}+X^2+1$
- CRC-CCITT: $P(X) = X^{16}+X^{12}+X^5+1$
- CRC-32: $P(X) = X^{32}+X^{26}+X^{23}+X^{22}+X^{16}+X^{12}+X^{11}+X^{10}+X^8+X^7+X^5+X^4+X^2+X+1$

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Υλοποίηση CRC σε υλικό

- Η υλοποίηση του CRC πραγματοποιείται συνήθως σε hardware. Η λύση αυτή έχει τα πλεονεκτήματα
  - υψηλή ταχύτητα
  - δεν οδηγεί σε σημαντική αύξηση του κόστους
    - γιατί η χρησιμοποιούμενη ψηφιακή λογική είναι εξαιρετικά απλή.
- Το hardware που χρησιμοποιείται για την υλοποίηση του CRC αποτελείται από αποθηκευτικά στοιχεία του ενός bit που απαρτίζουν έναν καταχωρητή ολίσθησης (shift register) και από έναν αριθμό πυλών XOR.

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Υλοποίηση CRC σε υλικό

- **1.** Ο καταχωρητής περιέχει n-k bits (όσα και το μήκος του FCS).
- **2.** Υπάρχουν μέχρι n-k πύλες XOR.
- **3.** Η παρουσία ή η απουσία πύλης XOR αντιστοιχεί στην παρουσία ή στην απουσία του αντίστοιχου όρου στο πολυώνυμο διαιρέτη P(X), με εξαίρεση τους όρους 1 και X^{n-k}

ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Υλοποίηση CRC σε υλικό

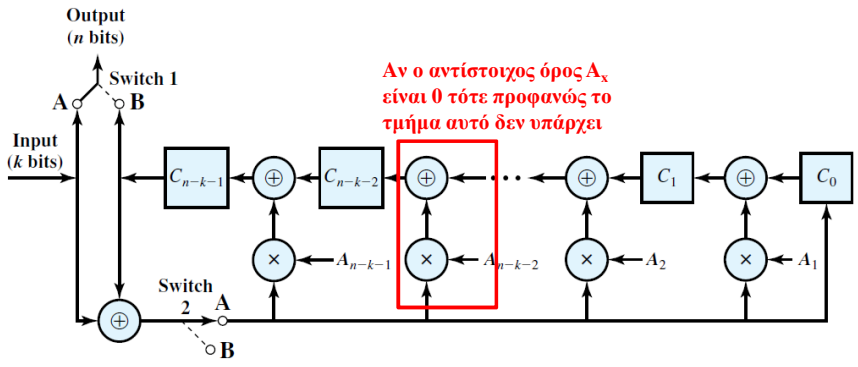


Figure 6.6 General CRC Architecture to Implement Divisor $(1 + A_1X + A_2X^2 + \cdots + A_{n-1}X^{n-k-1} + X^{n-k})$

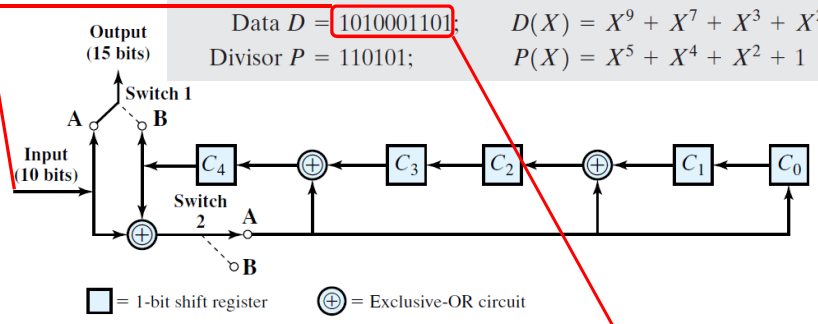
ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΣΦΑΛΜΑΤΟΣ

Υλοποίηση CRC σε υλικό

- Δίνουμε αρχικά μηδενική τιμή στα C_i (i=0,...,n-k-1)
- Εισάγουμε διαδοχικά στην είσοδο input τα bits του μηνύματος D
- **Μετά από την εισαγωγή και του τελευταίου μηδενικού ο καταχωρητής ολίσθησης θα περιέχει το FCS.**

Data D = 1010001101; $D(X) = X^9 + X^7 + X^3 + X^2 + 1$

Divisor P = 110101; $P(X) = X^5 + X^4 + X^2 + 1$



(a) Shift-register implementation

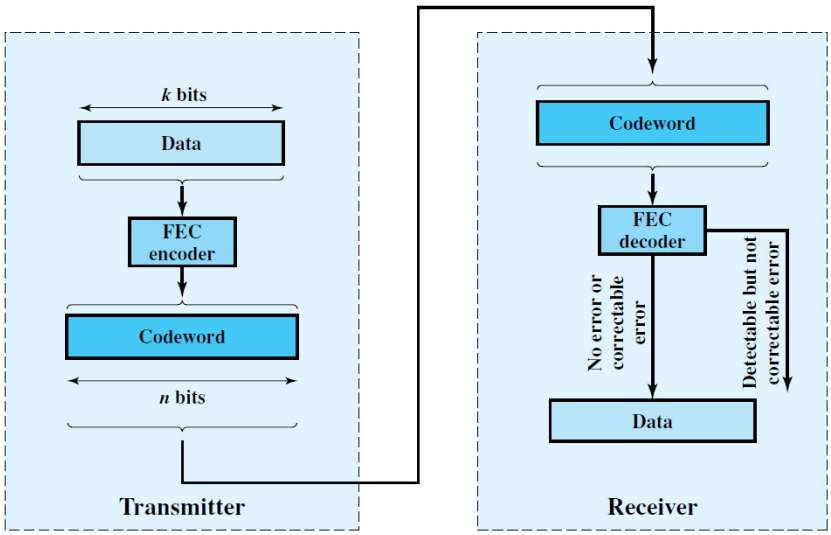
	C_4	C_3	C_2	C_1	C_0	$C_4 \oplus C_3 \oplus I$	$C_4 \oplus C_1 \oplus I$	$C_4 \oplus I$	I = input
Initial	0	0	0	0	0	1	1	1	1
Step 1	1	0	1	0	1	1	1	1	0
Step 2	1	1	1	1	1	1	1	0	1
Step 3	1	1	1	1	0	0	0	1	0
Step 4	0	1	0	0	1	1	0	0	0
Step 5	1	0	0	1	0	1	0	1	0
Step 6	1	0	0	0	1	0	0	0	1
Step 7	0	0	0	1	0	1	0	1	1
Step 8	1	0	0	0	1	1	1	1	0
Step 9	1	0	1	1	1	0	1	0	1
Step 10	0	1	1	1	0				

Message to be sent

FCS

(b) Example with input of 1010001101

ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ



ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

- Γενικά, οι κώδικες διόρθωσης σφαλμάτων επιβάλλουν τη μετάδοση πολλών πρόσθετων bits, κάτι που ελαττώνει την απόδοση του επικοινωνιακού συστήματος.
- Συνήθως, είναι προτιμότερη ανίνευση του σφάλματος και η διατύπωση αιτήματος προς το μεταδότη για επαναμετάδοση των εσφαλμένων μηνυμάτων.
- Μόνο σε περιπτώσεις που η διατύπωση τέτοιου αιτήματος είναι αδύνατη χρησιμοποιούνται κώδικες διόρθωσης σφαλμάτων.
- Χαρακτηριστικό παράδειγμα χρησιμοποίησης κώδικα διόρθωσης σφαλμάτων είναι η περίπτωση που το κανάλι επικοινωνίας είναι μονής κατεύθυνσης (simplex), γεγονός που δεν επιτρέπει την αποστολή μηνύματος από τον αποδέκτη προς το μεταδότη.

ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

- **Γενικά για διόρθωση:**
- Έστω x,y η αποσταλείσα/ληφθείσα λέξη. Εάν y ≠x και το y δεν είναι κωδική λέξη, y∉ C (ανίχνευση λάθους), τότε για λήψη απόφασης χρησιμοποιείται η **Αποκωδικοποίηση Μεγίστης Πιθανότητας (ΑΜΠ)**:
- Αν υπάρχει μόνο μία κωδική λέξη x, η οποία εμφανίζει την μικρότερη απόσταση από τη λέξη y, σε σύγκριση με όλες τις άλλες κωδικές λέξεις, τότε η y αποκωδικοποιείται ως x *(μεγιστοποίηση της πιθανότητας p(x,y))*
- Εάν υπάρχουν περισσότερες κωδικές λέξεις με ίδια απόσταση από τη y:
  - Είτε ο αποδέκτης αποκωδικοποιεί αυθαίρετα τη ληφθείσα λέξη ως μία από αυτές τις κωδικές λέξεις
  - Είτε επαναλαμβάνεται η μετάδοση
    - Επανάληψη μετάδοσης μπορεί να ζητηθεί και όταν,υπάρχει μια μόνο κωδική λέξη εγγύτερη στη λέξη y, με πολύ μεγάλη απόσταση

Εισαγωγή στην Κωδικοποίηση

- **Παράδειγμα**:Εάν |M|=2, k = 1, n = 3 και C={000, 111} :
  - Ο αποδέκτης αποφασίζει για «111», εφόσον είτε y= 111 ή y= «011», «101 ή «110», καθώς έχουν μικρότερη απόσταση προς την «111» από ό,τι προς «000».
  - Συμπεραίνει την κωδική λέξη «000» σε οποιαδήποτε από τις υπόλοιπες τέσσερις δυαδικές ακολουθίες. Η περίπτωση της ίσης απόστασης της ληφθείσας λέξης δεν είναι δυνατή με τον κώδικα αυτό

Εισαγωγή στην Κωδικοποίηση

Θεώρημα Διόρθωσης Σφαλμάτων

- Ένας κώδικας C απόστασης d **διορθώνει** όλα τα πρότυπα σφάλματος βάρους **μικρότερου ή ίσου του t=⌊(d – 1)/2⌋** και υπάρχει τουλάχιστον ένα πρότυπο σφάλματος βάρους t+1 που δεν διορθώνει ο κώδικας C.
  - ***Ισοδύναμη διατύπωση***: Ένας κώδικας C για να μπορεί να διορθώσει t σφάλματα θα πρέπει να έχει απόσταση **d≥2t+1**

Εισαγωγή στην Κωδικοποίηση

- **Παράδειγμα:** Ποια πρότυπα σφάλματος διορθώνονται από C = {000, 111};
- Σύμφωνα με το θεώρημα, αφού η απόσταση του κώδικα είναι d= 3, **διορθώνει κάθε πρότυπο σφάλματος βάρους** t ≤ ⌊(3 – 1)/2⌋= **1**
- Όπως εδείχθη στο προηγούμενο παράδειγμα, ο αποδέκτης συμπεραίνει την κωδική λέξη «111», εφόσον (λάβει αυτήν ή λάβει μία από τις «011», «101 ή «110», δηλαδή διορθώνει:
  - τα πρότυπα σφάλματος ε1= 111 + 011 = 100, ε2=111 + 101 = 010 και ε3=111 + 110 = 001 στην περίπτωση μετάδοσης της λέξης «111».
- Ομοίως ο αποδέκτης συμπεραίνει την «000», εφόσον λάβει αυτήν ή λάβει μία από τις «001», «010» ή «100», δηλαδή διορθώνει:
  - Τα πρότυπα σφάλματος 000 + 100 = 100, 000 + 010 = 010 και 000 + 001 = 001.
- **Συνεπώς ο κώδικας C διορθώνει τα πρότυπα σφάλματος 100, 010 και 001,**
- **ο C δεν διορθώνει πρότυπα σφάλματος βάρους μεγαλύτερου του 1.**
  - π.χ., κατά τη μετάδοση της λέξης «111», αν εμφανιστεί το πρότυπο σφάλματος «110», δηλαδή ληφθεί η λέξη «001», ο κώδικας τη διορθώνει εσφαλμένα σε «000».

ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

- Ο συχνότερα χρησιμοποιούμενος κώδικας διόρθωσης σφαλμάτων είναι ο κώδικας Hamming.
- Σύμφωνα με τον κώδικα αυτόν τα bits του μηνύματος αριθμούνται από τα αριστερά προς τα δεξιά, αρχίζοντας από το 1.
  - Τα bits που οι θέσεις τους είναι δυνάμεις του 2 (1,2,4,8,...) είναι bits ελέγχου.
  - Στις υπόλοιπες θέσεις τοποθετούνται τα bits δεδομένων.
- Κάθε bit ελέγχου, έχει τέτοια τιμή ώστε να υπάρχει άρτια (ή περιττή) ισοτιμία σε κάποιο σύνολο bits στα οποία συμπεριλαμβάνεται και ο εαυτός του.
- Κάθε bit δεδομένων ελέγχεται από περισσότερα του ενός bits ελέγχου.
- Για να υπολογίσουμε ποιά bits ελέγχου, ελέγχουν ένα συγκεκριμένο bit δεδομένων αναλύουμε τον αριθμό που εκφράζει τη θέση του συγκεκριμένου bit δεδομένων ως άθροισμα δυνάμεων του 2.

ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

- Για παράδειγμα, το bit δεδομένων που βρίσκεται στη θέση 13 (13=8+4+1) ελέγχεται από τα bits ελέγχου που βρίσκονται στις θέσεις 1, 4 και 8.
- Όταν φθάνει στον αποδέκτη ένα μήνυμα κωδικοποιημένο κατά Hamming, τότε ο αποδέκτης εξετάζει καθένα από τα bits ελέγχου (1,2,4,...) προκειμένου να διαπιστώσει αν έχουν τις σωστές τιμές.

ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

- Για κάθε bit ελέγχου που δεν πληρεί τον κανόνα της άρτιας (ή περιττής) ισοτιμίας με τα bits δεδομένων που ελέγχει,
  - **προστίθεται ο αριθμός της θέσης του λανθασμένου bit ελέγχου σε έναν μετρητή.**
- Αν μετά την ολοκλήρωση της εξέτασης όλων των bits ελέγχου
  - δεν βρεθεί σφάλμα, τότε το μήνυμα έχει ληφθεί σωστά.
  - **Αλλιώς, ο μετρητής περιέχει τη θέση του bit οπου υπάρχει σφάλμα**.
- Αν για παράδειγμα έχουν βρεθεί λανθασμένα τα bits ελέγχου που βρίσκονται στις θέσεις 1,4 και 8,
  - τότε συνάγεται ότι το bit που βρίσκεται στη θέση 13 είναι εσφαλμένο,
    - διότι είναι το μόνο bit που ελέγχεται από τα bits ελέγχου 1,4 και 8.
  - Αλλάζοντας την τιμή του bit αυτού διορθώνουμε τα σφάλμα και αποκαθιστούμε το μήνυμα στην ορθή του μορφή.

ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

- Αν m είναι ο αριθμός των bits δεδομένων που πρόκειται να μεταδοθούν
  - τότε ο αριθμός r των bits του κώδικα Hamming που πρέπει να χρησιμοποιηθούν
    - είναι ο ελάχιστος αριθμός που ικανοποιεί τη συνθήκη: $m+r+1 \leq 2^r$.
    - Αν, για παράδειγμα, m=13 τότε πρέπει r=5, έτσι ώστε $m+r+1=19 \leq 2^r=32$.

Το μεταδιδόμενο μήνυμα έχει τη μορφή:

H1 H2 D3 **H4** D5 D6 D7 **H8** D9 D10 D11 D12 D13 D14 D15 **H16** D17 D18

ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

Κάθε bit δεδομένων Di ελέγχεται από τα εξής στου κώδικα Hamming:

- **D3**: H1,H2
- **D5**: H1,H4
- **D6**: H2,H4
- **D7**: H1,H2,H4
- **D9**:H1,H8
- **D10**: H2,H8
- **D11**: H1,H2,H8
- **D12**: H4,H8
- **D13**: H1,H4,H8
- **D14**: H2,H4,H8
- **D15**: H1,H2,H4,H8
- **D17**: H1,H16
- **D18**: H2,H16

ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

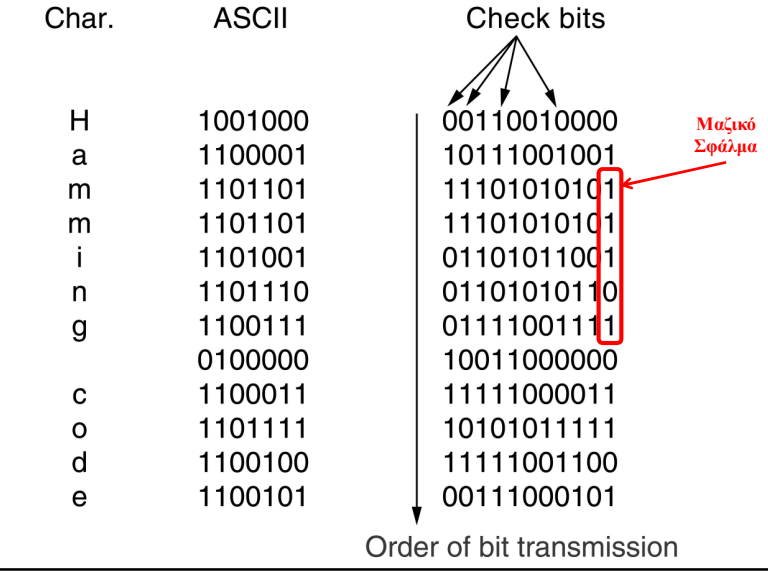
Κάθε bit του κώδικα Hamming H_i (i=1,2,4,…) ελέγχει τα παρακάτω bits δεδομένων:

- **H1**: D3, D5, D7, D9, D11, D13, D15, D17
- **H2**: D3, D6, D7, D10, D11, D14, D15, D18
- **H4**: D5, D6, D7, D12, D13, D14, D15
- **H8**: D9, D10, D11, D12, D13, D14, D15
- **H16**: D17, D18

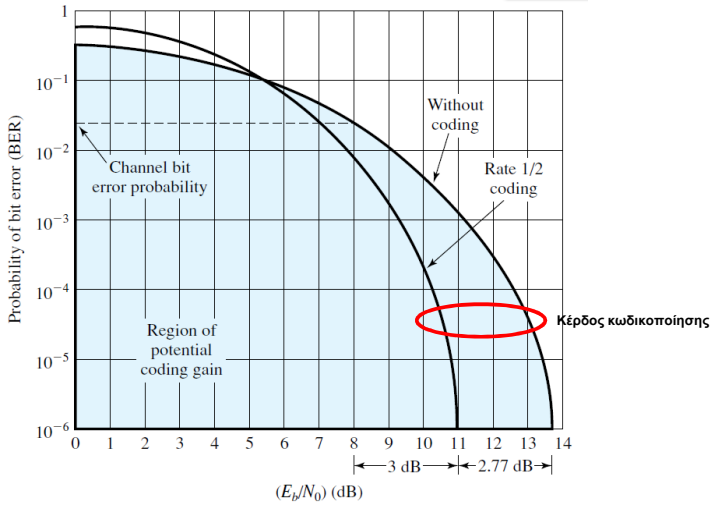
ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

- Ένα μειονέκτημα του κώδικα Hamming είναι ότι αδυνατεί να διορθώσει μαζικά (bursty) σφάλματα.
  - Επειδή είναι σύνηθες το φαινόμενο μια "ριπή" θορύβου να αλλοιώσει πολλά συνεχόμενα bits του μηνύματος, εύκολα γίνεται αντιληπτό το πόσο σημαντικό είναι το μειονέκτημα αυτό.
- Μια τακτική που ακολουθείται προκειμένου να αντιμετωπιστεί αυτό το πρόβλημα είναι η μετάδοση ενός συνόλου κωδικοποιημένων κατά Hamming μηνυμάτων κατά στήλες
  - Έτσι μια "ριπή" θορύβου που αλλοιώνει πολλά συνεχόμενα στη σειρά μετάδοσης bits, δεν προκαλεί αλλοίωση παρά σε ένα μόνο bit του κάθε μηνύματος.

ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ



ΚΩΔΙΚΕΣ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ



ΠΑΡΑΔΕΙΓΜΑΤΑ

Παράδειγμα 1

Η ένταξη ενός bit ισοτιμίας σε κάθε χαρακτήρα ενός μηνύματος θα μπορούσε να αλλάξει τη πιθανότητα να ληφθεί σωστό ένα μήνυμα;

Παράδειγμα 1 (Λύση)

- Η ένταξη ενός bit ισοτιμίας **αυξάνει το μήκος του μηνύματος.**
- Επομένως, υπάρχουν **περισσότερα bits που ενδέχεται να είναι λάθος** (με δεδομένο ότι περιλαμβάνεται σε αυτά και το bit ισοτιμίας).
- **Το bit ισοτιμίας μπορεί να είναι λάθος**, ακόμα και όταν δεν υπάρχουν λάθη στα αντίστοιχα bits δεδομένων.
- **Άρα, η ένταξη ενός bit ισοτιμίας σε κάθε χαρακτήρα θα άλλαζε την πιθανότητα λήψης ενός σωστού μηνύματος.**

Παράδειγμα 2

Θεωρείστε ένα frame που αποτελείται από δύο χαρακτήρες των 4 bits ο καθένας. Υποθέστε ότι η πιθανότητα σφάλματος σε ένα bit (BER) είναι 10^{-3} και είναι ανεξάρτητη για κάθε bit.

α) Ποιά είναι η πιθανότητα το λαμβανόμενο frame να περιέχει ένα τουλάχιστον λάθος;

β) Προσθέστε ένα bit ισοτιμίας σε κάθε χαρακτήρα. Ποια είναι τώρα η πιθανότητα το λαμβανόμενο frame να περιέχει ένα τουλάχιστον λάθος;

Παράδειγμα 2 (Λύση)

α) Έχουμε:

Pr[ένα bit να ληφθεί λάθος] = 10^{-3}

Pr[ένα bit να ληφθεί σωστά] = $1-10^{-3}=0.999$

Pr[να μη ληφθεί λάθος κανένα από τα 8 bits του frame] = $(1-10^{-3})^8 = (0.999)^8 = 0.992$

Pr[να υπάρχει τουλάχιστον ένα λάθος στο frame]= $1-(1-10^{-3})^8 = 0.008$

β) Pr[να υπάρχει τουλάχιστον ένα λάθος στο frame]= $1-(1-10^{-3})^{10} = 1-(0.999)^{10} = 0.01$

Παράδειγμα 3

Αν για τον έλεγχο σφαλμάτων χρησιμοποείται CRC με διαιρέτη P=10111 και τα προς μετάδοση δεδομένα είναι M= 1010011110:

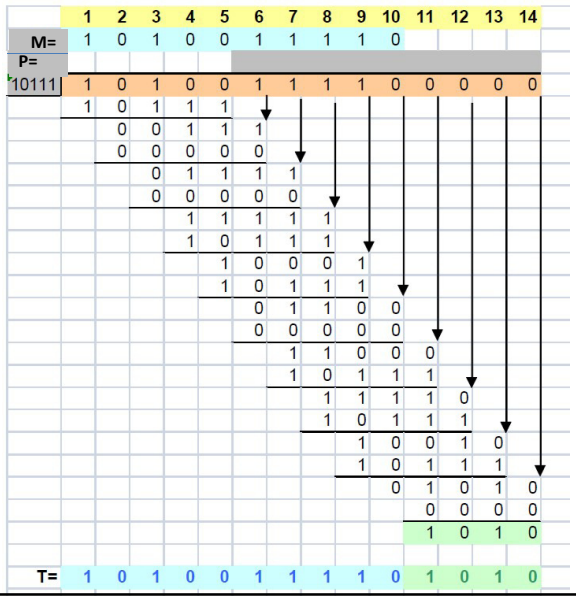
α) Υπολογίστε την ακολουθία T που θα μεταδοθεί.

β) Δείξτε τη διαδικασία ελέγχου σφάλματος στον αποδέκτη, υποθέτοντας ότι δεν υπάρχουν σφάλματα.

γ) Αν υπάρχει λάθος στο 6ο bit (μετρώντας από αριστερά προς τα δεξιά), δείξτε τη διαδικασία ελέγχου σφάλματος στον αποδέκτη, και διαπιστώστε αν το λάθος ανιχνεύεται.

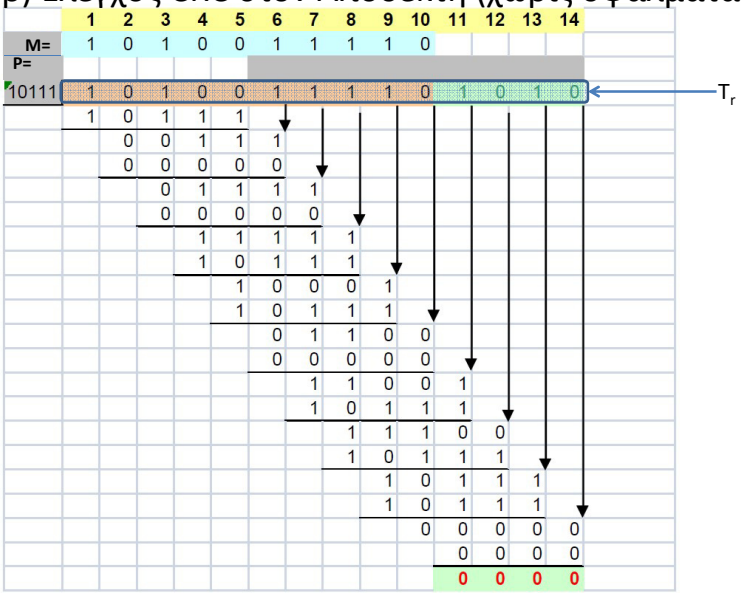
Παράδειγμα 3 (Λύση 1/3)

α) Υπολογισμός T στο Μεταδότη



Παράδειγμα 3 (Λύση 2/3)

β) Ελέγχος CRC στον Αποδέκτη (χωρίς σφάλματα)



Παράδειγμα 3 (Λύση 3/3)

γ) Ελέγχος CRC στον Αποδέκτη (σφάλμα στο 6ο bit)