ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ (ΠΜΣ)

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΔΙΔΑΚΤΙΚΗ ΤΗΣ ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΨΗΦΙΑΚΑ ΣΥΣΤΗΜΑΤΑ

ΚΑΤΕΥΘΥΝΣΗ : ΨΗΦΙΑΚΕΣ ΕΠΙΚΟΙΝΩΝΙΕΣ ΚΑΙ ΔΙΚΤΥΑ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

«Εφαρμογές της Θεωρίας Πληροφορίας στην ασφάλεια δικτύων»

ΦΟΙΤΗΤΡΙΑ : Δεμερτζή Χριστίνα ΜΕ/08045

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ : Δρ. Χρήστος Ξενάκης

Αθήνα, Μάιος 2011

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

0.0. ΠΕΡΙΛΗΨΗ 4

0.1. ABSTRACT 5

1. ΓΙΑΤΙ ΕΙΝΑΙ ΑΝΑΓΚΑΙΑ Η ΚΩΔΙΚΟΠΟΙΗΣΗ ΚΑΝΑΛΙΟΥ; 7

2. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ 9

3. ΑΠΟΣΤΑΣΗ ΚΑΙ ΒΑΡΟΣ HAMMING 12

4. ΤΕΧΝΙΚΕΣ ΚΑΙ ΚΩΔΙΚΕΣ ΑΝΙΧΝΕΥΣΗΣ ΚΑΙ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΩΝ 14

4.1. Automatic Repeat Request (ARQ) και Forward Error Correction (FEC) 15

4.2. Κωδικας ελεγχου απλης ισοτιμιας, κωδικας απλης επαναληψης και δισδιαστατος κωδικας ελεγχου ισοτιμιας 16

4.3. Επαναληπτικοι κωδικες 17

4.4. Αθροισματα ελεγχου (check sums) 17

4.5. Ορθογωνιοι και τριγωνικοι κωδικες 18

5. ΚΑΤΗΓΟΡΙΕΣ ΚΩΔΙΚΩΝ ΚΑΝΑΛΙΟΥ 21

6. ΓΡΑΜΜΙΚΟΙ ΚΩΔΙΚΕΣ (LINEAR CODES) 23

7. ΚΩΔΙΚΕΣ ΔΟΜΗΣ (BLOCK CODES) 27

8. ΚΩΔΙΚΕΣ HAMMING 30

9. ΚΩΔΙΚΕΣ ΧΑΜΗΛΗΣ ΠΥΚΝΟΤΗΤΑΣ ΕΛΕΓΧΟΥ ΙΣΟΤΙΜΙΑΣ (LOW DENSITY PARITY CHECK CODES, LDPC) 35

10. ΚΥΚΛΙΚΟΙ ΚΩΔΙΚΕΣ 38

11. ΚΩΔΙΚΕΣ BCH (BOSE-CHAUDHURI-HOCQUENGHEM) ΚΑΙ RS (REED-SOLOMON) 40

12. ΣΥΝΕΛΙΚΤΙΚΟΙ Η ΣΥΓΚΕΡΑΣΤΙΚΟΙ ΚΩΔΙΚΕΣ (CONVOLUTIONAL CODES) 41

13. ΚΩΔΙΚΕΣ TURBO 44

14. ΕΠΙΛΟΓΟΣ 47

15. ΑΣΚΗΣΕΙΣ 48

16. ΕΡΩΤΗΣΕΙΣ ΠΟΛΛΑΠΛΗΣ ΕΠΙΛΟΓΗΣ 54

ΒΙΒΛΙΟΓΡΑΦΙΑ 58

ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ

ΕΙΚΟΝΑ 1. ΔΙΣΔΙΑΣΤΑΤΟΣ ΚΩΔΙΚΑΣ ΕΛΕΓΧΟΥ ΙΣΟΤΙΜΙΑΣ 16
ΕΙΚΟΝΑ 2. ΠΑΡΑΔΕΙΓΜΑ ΟΡΘΟΓΩΝΙΟΥ ΚΩΔΙΚΑ. 19
ΕΙΚΟΝΑ 3. ΠΑΡΑΔΕΙΓΜΑ ΤΡΙΓΩΝΙΚΟΥ ΚΩΔΙΚΑ. 19
ΕΙΚΟΝΑ 4. ΚΑΤΗΓΟΡΙΕΣ ΚΩΔΙΚΩΝ ΚΑΝΑΛΙΟΥ 22
ΕΙΚΟΝΑ 5. HAMMING ΚΩΔΙΚΑΣ ΓΙΑ R=3 32
ΕΙΚΟΝΑ 6. ΤΟ ΜΕΤΑΛΛΙΟ ΜΕ ΤΟ ΟΠΟΙΟ ΤΙΜΗΘΗΚΕ Ο HAMMING ΑΠΟ ΤΟ IEEE. 33
ΕΙΚΟΝΑ 7. ΣΥΝΕΛΙΚΤΙΚΟΣ ΚΩΔΙΚΟΠΟΙΗΤΗΣ 41
ΕΙΚΟΝΑ 8. ΣΥΝΕΛΙΚΤΙΚΟΣ ΚΩΔΙΚΟΠΟΙΗΤΗΣ ΜΕ (N, K, K) = (2, 1, 3) 42
ΕΙΚΟΝΑ 9. ΚΩΔΙΚΟΠΟΙΗΣΗ TURBO 45
ΕΙΚΟΝΑ 10. ΑΠΟΚΩΔΙΚΟΠΟΙΗΣΗ TURBO. 46

0.0. Περίληψη

Στο σύγχρονο κόσμο η αποτελεσματική και σωστή μεταφορά της πληροφορίας από την πηγή στον προορισμό αποτελεί κλειδί για τη σωστή λειτουργία και ανάπτυξή του. Κάποια απλά και καθημερινά παραδείγματα είναι η μεταφορά οικονομικών πληροφοριών μέσω τηλεφωνικών γραμμών, η μεταφορά δεδομένων από έναν ηλεκτρονικό υπολογιστή σε άλλο ή από τη μνήμη στην κεντρική μονάδα επεξεργασίας και η μεταφορά δεδομένων από δορυφόρους στη Γη. Το φυσικό μέσο με το οποίο μεταδίδεται η πληροφορία ονομάζεται κανάλι επικοινωνίας και παραδείγματα καναλιών είναι οι τηλεφωνικές γραμμές και η ατμόσφαιρα. Δυστυχώς, στα κανάλια επικοινωνίας επενεργεί θόρυβος, ο οποίος προκαλεί αλλοίωση της μεταδιδόμενης πληροφορίας.

Η κωδικοποίηση καναλιού είναι μια καλά υπολογισμένη χρήση πλεονασμού με επιδίωξη την ανίχνευση και τη διόρθωση σφαλμάτων που προκαλούνται από το θόρυβο κατά τη μετάδοση της πληροφορίας στο κανάλι επικοινωνίας. Στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας δικαιολογείται η αναγκαιότητα της χρήσης κωδικοποίησης καναλιού, αναφέρονται βασικές έννοιες ώστε τα γραφόμενα να γίνονται εύκολα αντιληπτά από τον αναγνώστη, εξετάζονται διεξοδικά τεχνικές και κώδικες ανίχνευσης και διόρθωσης λαθών και τέλος αναλύονται οι βασικές κατηγορίες των κωδίκων καναλιού συμπεριλαμβάνοντας τα πλεονεκτήματα και μειονεκτήματά τους με απώτερο σκοπό τη βελτιστοποίηση των ήδη υπαρχόντων κωδίκων, την ανεύρεση καλύτερων συνδυασμών τους και γιατί όχι την ανακάλυψη νέων που θα μηδενίζουν την πιθανότητα αλλοίωσης της πληροφορίας από το θόρυβο.

0.1. Abstract

Nowadays, the effective and correct transport of information from the source to the destination constitutes key for its correct operation and development. Certain simple and daily examples are the transport of economic information via telephone lines, the transport of data from a computer to another or from the memory to the central processor unit and the transport of data from satellites to the Earth. Communication channel is the natural means which transmits the information and some examples of channels are the telephone lines and the atmosphere. Unfortunately, noise acts upon the communication channels and causes alteration of transmitted information.

The channel coding is a well calculated use of pleonasm with objective the error detection and correction which is caused by the noise during the transmission of information in the communication channel. In this master's thesis the necessity of use of channel coding is justified, basic significances are reported in order to make the text easy perceptible from the reader, techniques and codes for error detection and correction are examined and finally the main categories of channel codes are analyzed including their advantages and disadvantages with final aim the optimisation of already existing codes, the recovery of better code combinations and why not the discovery of new codes that will annihilate the probability of alteration of information because of the noise.

Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα καθηγητή μου Δρ. Χρήστο Ξενάκη για την ανάθεση της συγκεκριμένης διπλωματικής εργασίας, τη βοήθεια που μου παρείχε και το χρόνο που αφιέρωσε κατά τη διάρκεια της εκπόνησής της.

Τέλος, θα ήθελα να ευχαριστήσω το σύζυγό μου Αλέξανδρο για τη διαρκή συμπαράσταση, ενθάρρυνση και υποστήριξή του καθ' όλη τη διάρκεια των σπουδών μου.

1. Γιατί είναι αναγκαία η κωδικοποίηση καναλιού;

Ο πιο απλός τρόπος για να παρουσιάσει κανείς δυαδική πληροφορία είναι να χρησιμοποιήσει μία συγκεκριμένη τάση σε Volts για να απεικονίσει το 1 και τη μηδενική τάση για να απεικονίσει το 0. Η μορφή αυτή ονομάζεται απλή δυαδική (pure binary) κωδικοποίηση και είναι απόλυτα ικανοποιητική, όταν χρησιμοποιείται σε τερματικές συσκευές ή άλλα συστήματα. Παρουσιάζει προβλήματα όμως, όταν χρησιμοποιείται για τη μετάδοση της πληροφορίας σε απόσταση που ξεπερνά τις μερικές δεκάδες μέτρα. Το κενό αυτό καλύπτεται από ένα μεγάλο αριθμό κωδίκων καναλιού.

Οι δύο βασικοί λόγοι που χρησιμοποιούμε κώδικες καναλιού είναι η ανίχνευση σφαλμάτων και η διόρθωση σφαλμάτων. Η ανίχνευση σφάλματος σημαίνει ότι ο δέκτης έχει τη δυνατότητα να «καταλαβαίνει» αν η ληφθείσα κωδική λέξη είναι σφάλμα, δηλαδή διαφέρει από τη κωδική λέξη που μεταδώσαμε. Στην περίπτωση που ο δέκτης ανιχνεύει σφάλμα, ζητάει από τον πομπό την επανάληψη της μετάδοσης της κωδικής λέξης, μέχρι αυτή να φτάσει στο δέκτη χωρίς σφάλματα. Η διόρθωση σφάλματος, από την άλλη πλευρά, αποσκοπεί στην ελάττωση των σφαλμάτων χωρίς να χρειάζεται επαναμετάδοση της κωδικής λέξης.

Σχετικά με το κανάλι είναι απαραίτητες δύο παραδοχές που είναι καθοριστικές για την ανάπτυξη της θεωρίας κωδικοποίησης. Σύμφωνα με την πρώτη παραδοχή, μια κωδική λέξη μήκους *n* δυαδικών ψηφίων, που εισέρχεται στο κανάλι, λαμβάνεται στην έξοδό του ως λέξη μήκους *n* δυαδικών ψηφίων, αν και η ακολουθία εισόδου του καναλιού μπορεί να διαφέρει από αυτή της εξόδου του καναλιού. Επίσης, διαπιστώνεται από το δέκτη η αρχή της πρώτης λέξης μιας ακολουθίας κωδικών λέξεων που μεταδίδεται μέσω του καναλιού. Για παράδειγμα, αν στο κανάλι μεταδίδεται η δυαδική ακολουθία 010011, τότε στην έξοδό του λαμβάνεται η ακολουθία 010011 ή κάποια άλλη του ιδίου μήκους, όχι όμως η ακολουθία 10011 ή κάποια άλλη μικρότερου μήκους, επειδή χάθηκε το 1ο ψηφίο (το «0») της πρώτης λέξης της ακολουθίας. Άρα, η πρώτη παραδοχή αναφέρεται στη δυνατότητα του δέκτη να λάβει όλες τις λέξεις που μεταδόθηκαν, με ή χωρίς σφάλματα.

Η δεύτερη παραδοχή αναφέρεται στο ότι τα σφάλματα, δηλαδή ο θόρυβος, εμφανίζονται διασκορπισμένα κατά τυχαίο τρόπο και όχι σε συστάδες (ή καταιγισμούς, bursts). Με άλλα λόγια, η πιθανότητα να αλλοιωθεί ένα bit κατά τη μετάδοση εξαιτίας του θορύβου είναι η ίδια με αυτή οποιουδήποτε άλλου bit και δεν επηρεάζεται από σφάλματα σε γειτονικά δυαδικά ψηφία. Αυτή η παραδοχή δεν είναι ιδιαίτερα ρεαλιστική, αν λάβουμε υπόψη φυσικά φαινόμενα όπως αστραπές ή ακόμα και «γρατζουνιές» στα CD, που οδηγούν σε καταιγισμούς σφαλμάτων.

Κώδικας καναλιού είναι κάθε μέθοδος κωδικοποίησης ψηφιακής πληροφορίας που διευκολύνει τη μετάδοσή της μέσα από αναλογικά και ψηφιακά μέσα μετάδοσης. Περαιτέρω λόγοι που κάνουν απαραίτητη την ύπαρξη των κωδίκων καναλιού είναι οι παρακάτω:

• Είναι απαραίτητο να αφαιρείται από το αποστελλόμενο σήμα η συνεχής συνιστώσα τάσης που μπορεί αυτό να έχει, λόγω του ότι το μέσο μετάδοσης δεν μπορεί να τη μεταδώσει.

• Η ανάγκη να είναι ενήμερος ο δέκτης για τη χρονική στιγμή που ξεκινάει η μετάδοση και τη διάρκειά της.

• Η ανάγκη βέλτιστης χρήσης του εύρους ζώνης του συγκεκριμένου καναλιού επικοινωνίας.

• Η ανάγκη ύπαρξης τρόπου εντοπισμού και διόρθωσης λαθών (error detection and error correction) που παρουσιάζονται κατά τη μετάδοση της πληροφορίας.

• Η ανάγκη μείωσης της παραμόρφωσης.

• Η μείωση της πιθανότητας παρουσίασης διαφωνίας (crosstalk).

2. Βασικές έννοιες

Στο δυαδικό συμμετρικό κανάλι (Binary Symmetric Channel, BSC) η πιθανότητα το «0» να μεταφέρεται από το κανάλι ως «0» είναι ίση με την πιθανότητα το «1» να μεταφέρεται ως «1».

Η **αξιοπιστία** του καναλιού είναι ο πραγματικός αριθμός *p* με $0 \le p \le 1$, όπου *p* είναι η πιθανότητα της ορθής μεταφοράς ενός δυαδικού ψηφίου μέσω του καναλιού. Ένα κανάλι χαρακτηρίζεται πιο αξιόπιστο από ένα άλλο αν η πιθανότητα *p*, δηλαδή η αξιοπιστία του, είναι πιο υψηλή. Ωστόσο, αν $p = 1$ ή $p = 0$, τότε δεν υπάρχει περίπτωση σφάλματος ή υπάρχει πάντα σφάλμα κι επομένως δε θα ασχοληθούμε με αυτό το κανάλι. Επειδή κάθε κανάλι αξιοπιστίας *p* με $0 < p \le \frac{1}{2}$ μπορεί να μετατραπεί σε ένα κανάλι με $\frac{1}{2} \le p < 1$ θα ασχοληθούμε με δυαδικά συμμετρικά κανάλια με $\frac{1}{2} < p < 1$, διότι η περίπτωση $p = \frac{1}{2}$ δεν επιτρέπει την εξαγωγή οποιουδήποτε αξιοποιήσιμου αποτελέσματος.

Ο **ρυθμός πληροφορίας** ενός κώδικα είναι το ποσοστό της κωδικής λέξης που μεταφέρει το μήνυμα.
Ο ρυθμός πληροφορίας ενός δυαδικού κώδικα *C* μήκους *n* είναι :

$$R = \frac{\log_2 |C|}{n} \qquad (1)$$

Αφού $1 \le |C| \le 2^n$, ο ρυθμός πληροφορίας παίρνει τιμές μεταξύ 0 και 1, την τιμή 1 αν $|C| = 2^n$ δηλαδή κάθε λέξη n δυαδικών ψηφίων είναι κωδική λέξη και την τιμή 0 αν $|C| = 1$.

Το **βάρος Hamming** ή **βάρος** $wt(x)$ μιας λέξης x μήκους και n ψηφίων είναι το πλήθος των ψηφίων της λέξης που είναι ίσα με «1». Το βάρος παίρνει τιμές από 0 έως n .

Η **απόσταση Hamming** ή **απόσταση** $d(x, y)$ μεταξύ δύο λέξεων x και y του ίδιου μήκους n είναι το πλήθος των θέσεων στις οποίες οι δύο λέξεις εμφανίζουν ασυμφωνία του δυαδικού ψηφίου. Η απόσταση παίρνει τιμές από 0 έως n . Για παράδειγμα, οι κωδικές λέξεις 11000010 και 10010010 έχουν απόσταση Hamming 2, αφού διαφέρουν στο δεύτερο και στο τέταρτο δυαδικό ψηφίο. Δύο κωδικές λέξεις είναι ίδιες αν και μόνο αν η απόσταση Hamming μεταξύ τους είναι μηδέν.

Το βάρος και η απόσταση Hamming αναλύονται περισσότερο σε επόμενη παράγραφο.

Κωδικές λέξεις από διαφορετικά σύμβολα πηγής πρέπει να «απέχουν» με απόσταση τουλάχιστον 1, αλλιώς ο κώδικας θα είναι ίδιος. Οι κώδικες που θεωρούνται ανθεκτικοί στα λάθη αποτελούνται από λέξεις που «απέχουν» μεγαλύτερες αποστάσεις. Οι περισσότερες διαδικασίες αποκωδικοποίησης χρησιμοποιούν το κριτήριο της απόστασης, ώστε να επιλέξουν την πιθανότερα αναμενόμενη κωδική λέξη. Συγκεκριμένα, όταν λαμβάνεται μια λέξη, ο αποκωδικοποιητής ψάχνει για την έγκυρη κωδική λέξη που είναι πιο κοντά σε αυτήν, δηλαδή αυτή με τη μικρότερη απόσταση Hamming. Αν υπάρχει μία μόνο κοντινή λέξη, υποθέτει ότι είναι η αναμενόμενη κωδική λέξη. Αν υπάρχουν περισσότερες λέξεις με τη μικρότερη απόσταση, το λάθος απλά ανιχνεύεται, αλλά δε διορθώνεται.

Παράδειγμα 1

Δίνονται οι κωδικές λέξεις μιας πηγής 4 συμβόλων $x_1 = 0000$, $x_2 = 1101$, $x_3 = 0110$ *και* $x_4 = 1011$.

Ας υποθέσουμε ότι λαμβάνεται η ακολουθία 0001. Για να την αποκωδικοποιήσουμε, ψάχνουμε την κωδική λέξη που είναι πλησιέστερη σε αυτή. Η πλησιέστερη κωδική λέξη είναι η $x_1 = 0000$, *που έχει απόσταση 1 από τη λαμβανόμενη ακολουθία. Όλες οι υπόλοιπες κωδικές λέξεις έχουν απόσταση μεγαλύτερη από 1 από τη λαμβανόμενη ακολουθία.*

Ας υποθέσουμε τώρα ότι λαμβάνεται η ακολουθία 0100. Υπάρχουν 2 κοντινές λέξεις, η $x_1 = 0000$ *και η* $x_3 = 0110$, *που έχουν απόσταση 1. Για το λόγο αυτό, αν και ένα λάθος μπορεί να ανιχνευτεί, δεν μπορεί να διορθωθεί χρησιμοποιώντας το κριτήριο της ελάχιστης απόστασης. Γενικά, η ελάχιστη απόσταση καθορίζει το βαθμό στον οποίο ένας κώδικας είναι αδιάβλητος από λάθη.*

Για να προσδιορίσουμε τη σχέση μεταξύ της ελάχιστης απόστασης ενός κώδικα και της ανθεκτικότητάς του σε λάθη ξεχωρίζουμε την ανίχνευση πολλαπλών λαθών και τη διόρθωση πολλαπλών λαθών. Για έναν ακέραιο e, ο όρος e-error detection σημαίνει ότι αν υπάρχουν το πολύ e λάθη στη λαμβανόμενη λέξη, ο παραλήπτης μπορεί να ανιχνεύσει ότι η λέξη έχει λάθη, αλλά δεν μπορεί απαραίτητα να διορθώσει αυτά τα λάθη. Για έναν ακέραιο f, ο όρος f-error correction σημαίνει ότι αν υπάρχουν το πολύ f λάθη στη λαμβανόμενη λέξη, ο παραλήπτης μπορεί και ανιχνεύσει και να διορθώσει αυτά τα λάθη χρησιμοποιώντας το κριτήριο της ελάχιστης απόστασης.

Είναι εύκολο να καταλάβουμε τη διαφορά μεταξύ των δύο εννοιών στην περίπτωση ενός λάθους. Όταν υπάρχει μόνο ένα λάθος, μπορεί να ανιχνευθεί

αν η ελάχιστη απόσταση μεταξύ των κωδικών λέξεων είναι $d = 2$. Αυτό γίνεται γιατί όταν το λάθος είναι ένα, η λαμβανόμενη λέξη έχει απόσταση 1 από την αναμενόμενη λέξη, αλλά δεν υπάρχει κωδική λέξη τόσο κοντά, οπότε το λάθος ανιχνεύεται. Παρ' όλα αυτά, αυτό το λάθος δεν μπορεί πάντα να διορθωθεί όταν $d = 2$, διότι μπορεί να υπάρχουν περισσότερες από μια έγκυρες κωδικές λέξεις με απόσταση 1 από τη λαμβανόμενη λέξη. Η ελάχιστη απόσταση $d = 3$ εξασφαλίζει ότι ο παραλήπτης μπορεί να ανιχνεύσει και να διορθώσει ένα λάθος. Αν η λαμβανόμενη λέξη έχει απόσταση 1 από την αναμενόμενη λέξη (λόγω μόνο ενός λάθους), θα έχει απόσταση τουλάχιστον 2 από οποιαδήποτε άλλη κωδική λέξη. Ως εκ τούτου, υπάρχει μοναδική κωδική λέξη πλησιέστερη στη λαμβανόμενη λέξη και αυτή πρέπει να είναι η αναμενόμενη λέξη.

3. Απόσταση και Βάρος Hamming

Σε οποιαδήποτε κωδική λέξη ορίζεται ένας συγκεκριμένος αριθμός ο οποίος καλείται βάρος Hamming (Hamming weight). Η ποσότητα αυτή αντιστοιχεί στον αριθμό των μη μηδενικών στοιχείων της κωδικής λέξης. Ως απόσταση Hamming (Hamming distance) d, μεταξύ δύο κωδικών λέξεων ορίζεται ο αριθμός των θέσεων στις οποίες διαφέρουν οι δύο λέξεις.

Για παράδειγμα, έστω οι κωδικές λέξεις $\nu = 11010$ και $u = 10111$. Παρατηρούμε ότι οι λέξεις αυτές διαφέρουν στα δυαδικά ψηφία a_0, a_2 και a_3, όπου το bit a_0 αντιστοιχεί στο ελάχιστης σημασίας bit (Least Significant Bit, LSB), δηλαδή σε τρεις θέσεις, συνεπώς η απόσταση Hamming είναι $d = 3$.

Ως ελάχιστη απόσταση (minimum distance) $d_{\min}$, ενός κώδικα δομής ορίζεται η ελάχιστη απόσταση Hamming μεταξύ δύο ζευγών κωδικών λέξεων του κώδικα. Η ελάχιστη απόσταση αποτελεί σημαντική παράμετρο καθώς καθορίζει την ικανότητα ανίχνευσης και διόρθωσης σφαλμάτων του κώδικα. Συγκεκριμένα, σε ένα γραμμικό κώδικα δομής με ελάχιστη απόσταση $d_{\min}$ αποδεικνύεται ότι σε κάθε κωδική λέξη υπάρχει η δυνατότητα να ανιχνευθούν $s \le (d_{\min} - 1)$ σφάλματα και να διορθωθούν αντίστοιχα.

Σημαντική είναι και η σχέση που υπάρχει μεταξύ του πίνακα ελέγχου ισοτιμίας και της ελάχιστης απόστασης, η οποία μπορεί οριστεί ως ο ελάχιστος αριθμός στηλών του πίνακα H, που έχουν άθροισμα ίσο με το μηδέν. Συγκεκριμένα, η σχέση που αναφέρεται προκύπτει από τα ακόλουθα:

Ένας γραμμικός κώδικας δομής ορίζεται από το σύνολο των κωδικών λέξεων που ικανοποιούν τη σχέση $\nu \cdot \mathrm{H}^{\mathrm{T}} = 0$.

Μπορούμε να χρησιμοποιήσουμε για τον πίνακα ελέγχου ισοτιμίας την έκφραση $\mathrm{H} = [h_0, h_1, ..., h_i, ..., h_{n-1}]$, όπου κάθε τιμή h_i αντιστοιχεί στην i-οστή στήλη του πίνακα. Έτσι λοιπόν ο κώδικας περιγράφεται από τη σχέση που ακολουθεί:

$$\mathrm{u}_0 \cdot \mathrm{h}_0 + \mathrm{u}_1 \cdot \mathrm{h}_1 + \ldots + \mathrm{u}_i \cdot \mathrm{h}_i + \ldots + \mathrm{u}_{n-1} \cdot \mathrm{h}_{n-1} = 0 \quad (2)$$

όπου η τιμή u_i αναπαριστά το i-οστό στοιχείο της κωδικής λέξης ν.

Γίνεται εύκολα κατανοητό ότι για να ισχύει η προηγούμενη εξίσωση, θα πρέπει η κωδική λέξη ν να έχει μη μηδενικά στοιχεία σε κατάλληλες θέσεις τέτοιες ώστε οι αντίστοιχες στήλες του πίνακα ελέγχου ισοτιμίας να έχουν σαν άθροισμα το μηδενικό διάνυσμα. Από την άλλη πλευρά, η ελάχιστη απόσταση ενός κώδικα είναι ίση με τον ελάχιστο αριθμό των μη μηδενικών στοιχείων

μιας κωδικής λέξης. Για το λόγο αυτό λοιπόν, η ελάχιστη απόσταση ενός γραμμικού κώδικα δομής ισούται με τον ελάχιστο αριθμό στηλών του πίνακα Η των οποίων το άθροισμα είναι ίσο με το μηδέν.

Παράδειγμα 2

Έστω ότι έχουμε τον εξής πίνακα ελέγχου ισοτιμίας:

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

Παρατηρούμε ότι καμία στήλη δεν είναι μηδενική. Μπορούμε να διακρίνουμε επίσης, ότι δεν υπάρχει κανένα ζευγάρι στηλών των οποίων το άθροισμα να δίνει αποτέλεσμα ίσο με το μηδέν. Το ελάχιστο σύνολο στηλών δηλαδή, που δίνει μηδενικό άθροισμα είναι τρεις. Σύμφωνα με τις παρατηρήσεις αυτές λοιπόν, αντιλαμβανόμαστε ότι η ελάχιστη απόσταση του κώδικα ισούται με 3.

4. Τεχνικές και κώδικες ανίχνευσης και διόρθωσης σφαλμάτων

Σύμφωνα με τον θεμελιωτή της «Σύγχρονης Θεωρίας Κωδίκων», Claude Shannon, κατά τη διάδοση μίας λέξης μέσω ενός δυαδικού Συμμετρικού Καναλιού (Binary Symmetric Channel, BSC), δεν χάνονται ούτε προστίθενται δυαδικά ψηφία. Έτσι λοιπόν, μία κωδική λέξη η οποία έχει μήκος n bits και η οποία μεταδίδεται διαμέσου ενός BSC, θα φτάσει στον προορισμό της με το ίδιο μήκος ψηφίων. Αυτό όμως δεν σημαίνει ότι η λέξη που λαμβάνει ο δέκτης είναι και η επιθυμητή, διότι μπορεί να έχει υποστεί αλλοιώσεις λόγω θορύβου.

Για παράδειγμα, αν ο παραλήπτης δεχτεί την ψηφιοσειρά (bitstream) 001111101010 και ο κώδικας έχει μήκος 4 (δηλαδή κάθε κωδική λέξη αποτελείται από 4 ψηφία), τότε γνωρίζει ότι έλαβε τις εξής τρεις κωδικές λέξεις: 0011, 1110 και 1010. Όπως έχει αποδείξει ο Shannon, δεν είναι δυνατό να ληφθεί πληροφορία με μήκος διαφορετικό από κάποιο πολλαπλάσιο του 4. Αυτό θα σήμαινε ότι θα έχουν «χαθεί» ή «προστεθεί» ψηφία, γεγονός που δεν ευσταθεί.

Η λογική της κωδικοποίησης έγκειται στη σύγκριση της κάθε λέξης που λαμβάνει ο δέκτης με ένα γνωστό σύνολο κωδικών λέξεων. Προσπαθεί δηλαδή να αντιπαραβάλει τις λαμβανόμενες κωδικές λέξεις με αυτές κάποιου αλφαβήτου. Αν δεν κατορθώνει να βρει λέξη όμοια με αυτές που περιλαμβάνονται στο αλφάβητο, συμπεραίνουμε ότι έχουν υπεισέλθει σφάλματα. Σε αντίθετη περίπτωση, εξάγεται το συμπέρασμα ότι έχει γίνει ορθή λήψη της πληροφορίας.

Έστω για παράδειγμα ότι έχουμε τον κώδικα $C_1 = \{00,\ 01,\ 10,11\}$. Όπως φαίνεται, κάθε κωδική λέξη έχει μήκος 2 και ο κώδικας περιλαμβάνει όλους τους δυνατούς συνδυασμούς των ψηφίων «0» και «1». Συνεπώς, κάθε λέξη των δύο ψηφίων που καταλήγει στον δέκτη αποτελεί κωδική λέξη και για το λόγο αυτό δεν υπάρχει η δυνατότητα ανίχνευσης κάποιου λάθους.

Στην περίπτωση όμως του κώδικα $C_2 = \{000000,\ 010101,\ 101010,\ 111111\}$, ο οποίος είναι επαναληπτικός, κάθε κωδική λέξη του C_1 επαναλαμβάνεται τρεις φορές. Η μέθοδος αυτή κωδικοποίησης, παρουσιάζει το πλεονέκτημα της ευκολότερης ανίχνευσης πιθανών λαθών. Αν ληφθεί για παράδειγμα η λέξη «011101», με απλή αντιπαραβολή της με το σύνολο των γνωστών λέξεων του κώδικα, παρατηρείται ότι δεν ανήκει στο συγκεκριμένο λεξικό, δεν αποτελεί δηλαδή κωδική λέξη, άρα περιέχει ένα τουλάχιστον σφάλμα.

Υποθέτοντας ότι σε κάθε κωδική λέξη είναι δυνατό να υπάρξει το πολύ ένα λάθος, συμπεραίνουμε ότι τουλάχιστον δύο από τις τρεις επαναλήψεις των

λέξεων που ελήφθησαν είναι σωστές. Η διαδικασία της επανάληψης όμως έχει και το μειονέκτημα ότι ο βαθμός πληροφορίας του κώδικα μειώθηκε στο 1/3, από 1 που ήταν στην πρώτη περίπτωση. Για να αποφευχθεί η ατέλεια αυτή, εισάγεται στον αρχικό κώδικα C_1 ένα επιπλέον ψηφίο. Συγκεκριμένα, στο τέλος κάθε κωδικής λέξης προστίθεται ένα «1» ή ένα «0» , έτσι ώστε κάθε κωδική λέξη να περιλαμβάνει άρτιο αριθμό «1».

Κατά συνέπεια προκύπτει ένας νέος, πιο αποτελεσματικός κώδικας $C_3 = \{000,\ 011,\ 101,\ 110\}$. Το επιπλέον δυαδικό ψηφίο που προστέθηκε στο τέλος της κωδικής λέξης ονομάζεται ψηφίο ελέγχου ισοτιμίας (parity bit). Έτσι, αν ληφθούν τα στοιχεία «010», τα οποία δεν αποτελούν κωδική λέξη, γίνεται αντιληπτό ότι έχει προκύψει κάποιο σφάλμα κατά τη μετάδοση. Η προσπάθεια διόρθωσης έγκειται στην αλλαγή των λιγότερων κατά το δυνατό δυαδικών ψηφίων, ώστε να προκύψει κωδική λέξη. Γενικά, οι κώδικες πρέπει να σχεδιάζονται με τέτοιο τρόπο ώστε το κόστος κωδικοποίησης να είναι σχετικά χαμηλό. Η διαδικασία αυτή περιλαμβάνει κάποιες τεχνικές ανίχνευσης, αλλά και διόρθωσης σφαλμάτων, προκειμένου να μην απαιτείται επανάληψη της μετάδοσης της πληροφορίας και συνεπώς ύπαρξη διαύλου ανάδρασης.

4.1. Automatic Repeat Request (ARQ) και Forward Error Correction (FEC)

Σε περίπτωση εμφάνισης σφάλματος στο δέκτη μπορεί να εφαρμοστεί για την αντιμετώπισή του η Αυτόματη Αίτηση Επανεκπομπής (Automatic Repeat Request, ARQ). Σε ένα τέτοιο σύστημα ο δέκτης εκτελεί ανίχνευση των σφαλμάτων και απλά ζητά από τον πομπό επανεκπομπή των δεδομένων. Η συγκεκριμένη τεχνική συμβάλλει στην αξιοπιστία της λαμβανόμενης πληροφορίας, αν και έχει αυξημένη πολυπλοκότητα, αφού απαιτεί την ύπαρξη ενός καναλιού ανάδρασης, το οποίο όμως δεν είναι πάντα διαθέσιμο, καθιστώντας την έτσι μη πρακτική για αρκετές εφαρμογές.

Μία δεύτερη τεχνική που μπορεί να εφαρμοστεί είναι η Forward Error Correction (FEC), σύμφωνα με την οποία δέκτης σε περίπτωση ανίχνευσης σφάλματος προβαίνει και στη διόρθωσή του ακολουθώντας τους κανόνες κωδικοποίησης. Η τεχνική αυτή, αν και δυσκολότερη στην εφαρμογή από την ARQ, δεν απαιτεί δίαυλο ανάδρασης.

Βασική διαφορά των δύο τεχνικών αποτελεί η διόρθωση σφαλμάτων, διαδικασία που συμβαίνει μόνο στη FEC. Η τεχνική FEC περιλαμβάνει δύο μεγάλες κατηγορίες κωδίκων, τους κώδικες δομής (block codes) και τους

συνελικτικούς κώδικες (convolutional codes), οι οποίοι αναλύονται σε επόμενες παραγράφους.

4.2. Κώδικας ελέγχου απλής ισοτιμίας, κώδικας απλής επανάληψης και δισδιάστατος κώδικας ελέγχου ισοτιμίας

Ο απλούστερος κώδικας ανίχνευσης σφάλματος είναι ο κώδικας ελέγχου απλής ισοτιμίας (simple parity check code), $C(n, n-1, 2)$, $q = 2$. Ένα bit ισοτιμίας προστίθεται στο μήνυμα. Αυτός ο κώδικας είναι ικανός να ανιχνεύει ένα μοναδικό σφάλμα. Για παράδειγμα ο κώδικας :

$$C = \{0000,\ 0011,\ 0101,\ 0111,\ 1001,\ 1010,\ 1100,\ 1110\}$$

έχει παραμέτρους $C(4,3,2)$.

Ο απλούστερος κώδικας διόρθωσης σφάλματος είναι ο κώδικας απλής επανάληψης (simple repetition code), $C(n,1,n)$, $q = 2$. Αυτός ο κώδικας αποτελείται από δύο μόνο κωδικές λέξεις, η μια με όλα τα ψηφία της 0 και η άλλη με όλα 1. Το μήνυμα είναι ίσο με ένα bit, αν είναι 0 τότε μεταδίδεται η κωδική λέξη που αποτελείται μόνο από μηδενικά. Για παράδειγμα ο κώδικας: $C = \{00000,\ 11111\}$ έχει παραμέτρους $C(5,1,5)$ και μπορεί να διορθώσει 2 σφάλματα ή να ανιχνεύσει 4. Ο ρυθμός του είναι χαμηλός: $R = \frac{1}{5}$.

Ο δισδιάστατος κώδικας ελέγχου ισοτιμίας (2-dimensional parity check code), είναι εξίσου απλός, αλλά εμφανίζει βελτιωμένη αποδοτικότητα. Το μήνυμα τοποθετείται μέσα σε έναν πίνακα. Οι στήλες και οι γραμμές του πίνακα επιμηκύνονται κατά το αντίστοιχο bit ισοτιμίας. Θεωρήστε, για παράδειγμα, έναν μεγέθους 3×3 πίνακα μηνύματος και προσθέστε μια 4η γραμμή και μια 4η στήλη από bits ισοτιμίας όπως φαίνεται στην Εικόνα 1.

1	1	0	*0*
0	1	1	*0*
1	1	1	*1*
0	*1*	*0*	***1***

Εικόνα 1. Δισδιάστατος κώδικας ελέγχου ισοτιμίας

Το bit στη κάτω δεξιά γωνία καλείται ισοτιμία των ισοτιμιών (parity of parities), το οποίο μπορεί ισοδύναμα να υπολογιστεί ως την ισοτιμία της γραμμής με

bits ισοτιμίας με την ισοτιμία της στήλης με bits ισοτιμίας. Αυτός ο κώδικας έχει απόσταση 4, κατά συνέπεια είναι ικανός να διορθώσει ένα σφάλμα. Ο αλγόριθμος διόρθωσης είναι επίσης πολύ απλός: οι ενδείξεις (indices) της γραμμής με τις αντίστοιχες της στήλης με σφάλμα ισοτιμίας σηματοδοτούν την είσοδο με σφάλμα.

4.3. Επαναληπτικοί κώδικες

Αν υπάρχουν δύο σύμβολα Α και Β σε μια πηγή αλφαβήτου όπου χρησιμοποιείται δυαδική κωδικοποίηση, ο πιο άμεσος τρόπος να τα στείλουμε είναι να αντιστοιχίσουμε για παράδειγμα το 0 στο Α και το 1 στο Β. Αυτός ο τρόπος βέβαια είναι άμεσος, αλλά δεν προστατεύει από λάθη. Η προστασία από λάθη μπορεί εύκολα να επιτευχθεί με κάποιον επαναληπτικό κώδικα. Στην πιο απλή του μορφή κάθε χαρακτήρας του κώδικα επαναλαμβάνεται, ώστε το 0 να αντικατασταθεί από το 00 και το 1 από το 11. Αν συμβεί κάποιο λάθος, ο παραλήπτης μπορεί για παράδειγμα να λάβει 01 κι έτσι μπορεί να εντοπίσει την ύπαρξη λάθους. Παρόλα αυτά, ο παραλήπτης δε θα μπορέσει να διορθώσει το λάθος, αφού το 01 θα έπρεπε να είναι 00 ή 11. Μεγαλύτερη επανάληψη εγγυάται μεγαλύτερη αξιοπιστία.

4.4. Αθροίσματα ελέγχου (check sums)

Ένας πιο αποτελεσματικός τρόπος για ανίχνευση και διόρθωση λαθών είναι η χρήση αθροισμάτων ελέγχου.

Ένα bit ελέγχου ισοτιμίας (parity check bit) είναι ένα bit που «προσκολλάται» στη δυαδική κωδική λέξη κι εξασφαλίζει ότι το άθροισμα των bits θα είναι μια καθορισμένη τιμή. Για παράδειγμα, ένα bit μπορεί να προστεθεί στο τέλος της κωδικής λέξης 0110 ώστε να κάνει άθροισμα 0. Η νέα λέξη θα είναι 01100 όπου το υπογραμμισμένο bit είναι το bit ισοτιμίας. Το άθροισμα των ανεξάρτητων bits θα είναι μηδέν (αριθμητικά το υπόλοιπο της διαίρεσης με το 2) ή ισοδύναμα υπάρχει ζυγός αριθμός άσων. Αν η κωδική λέξη ήταν 1110, με την ίδια διαδικασία θα παραγόταν η νέα λέξη 11101 για να κρατήσει ζυγό τον αριθμό των άσων. Αυτό ονομάζεται άρτια ισοτιμία (even parity), ενώ υπάρχει και η σπανιότερα χρησιμοποιούμενη περιττή ισοτιμία (odd parity) που κρατά το άθροισμα των άσων μονό αριθμό.

Παράδειγμα 3

Στα περισσότερα βιβλία είναι τυπωμένος ο αριθμός ISBN (International Standard Book Number). Συνήθως είναι ένας 10-ψήφιος αριθμός που ορίζεται από τον εκδότη. Για παράδειγμα, ένας αριθμός ISBN μπορεί να είναι ο ***0-691-12418-3****, αν και οι παύλες μπορεί να διαφέρουν από βιβλίο σε βιβλίο. Το πρώτο ψηφίο υποδεικνύει τη γλώσσα στην οποία είναι γραμμένη το βιβλίο (το 0 αντιστοιχεί στα Αγγλικά). Τα επόμενα 2 ή 3 ψηφία υποδεικνύουν τον εκδότη (το 691 αντιστοιχεί στο Princeton University Press). Τα επόμενα 5 ή 6 ψηφία υποδεικνύουν τον αριθμό του βιβλίου που έχει οριστεί από τον εκδότη. Το τελευταίο ψηφίο είναι ένα ψηφίο ελέγχου που μπορεί να πάρει τις τιμές 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, X με το X να αντιστοιχεί στο 10. Το ψηφίο ελέγχου έχει σχεδιαστεί ώστε το 0 modulo 11 να είναι ένας γραμμικός συνδυασμός όλων των ψηφίων, δηλαδή ο συνδυασμός να είναι πολλαπλάσιο του 11. Πιο συγκεκριμένα:*

$$x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 + 6x_6 + 7x_7 + 8x_8 + 9x_9 + 10x_{10} = 0 \bmod 11$$

Το άθροισμα στο αριστερό μέρος της ισότητας είναι ένα ακέραιο πολλαπλάσιο του 11. Για τον αριθμό ISBN που δίνεται παραπάνω έχουμε:

$$1\cdot 0 + 2\cdot 6 + 3\cdot 9 + 4\cdot 1 + 5\cdot 1 + 6\cdot 2 + 7\cdot 4 + 8\cdot 1 + 9\cdot 8 + 10\cdot 3 = 198 = 11\cdot 18 = 0 \pmod{11}$$

Να σημειωθεί εδώ ότι το -10 είναι αριθμητικά το ίδιο με το 1 mod 11, αφού

$$-10 = -1\cdot 11 + 1 \ \kappa\alpha\iota \ 1 = 0\cdot 11 + 1$$

Άρα, μπορούμε να λύσουμε ως προς x_{10}

$$x_{10} = -10x_{10} = x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 + 6x_6 + 7x_7 + 8x_8 + 9x_9 \bmod 11$$

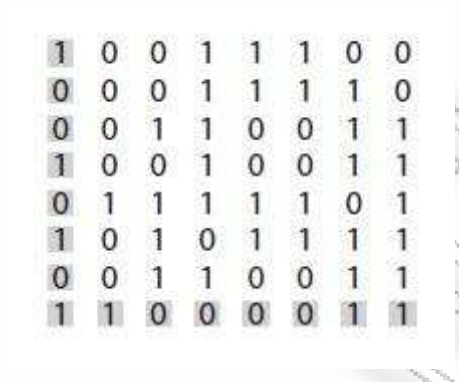
Ως εκ τούτου για το συγκεκριμένο παράδειγμα:

$$x_{10} = 1\cdot 0 + 2\cdot 6 + 3\cdot 9 + 4\cdot 1 + 5\cdot 1 + 6\cdot 2 + 7\cdot 4 + 8\cdot 1 + 9\cdot 8 = 168 = 11\cdot 15 + 3 = 3 \ (\bmod 11)$$

Η ιδιότητα του ISBN κώδικα είναι ότι μπορεί να ανιχνεύσει ένα λάθος σε οποιοδήποτε ψηφίο ή μια ανταλλαγή δύο ψηφίων. Στο άθροισμα ελέγχου οι συντελεστές είναι αυτοί που επιτρέπουν την ανίχνευση μιας ανταλλαγής. Αυτός είναι πρακτικός κώδικας που προέκυψε από την εμπειρική παρατήρηση ότι τα πιο κοινά λάθη που γίνονται στους αριθμούς ISBN είναι τα λάθη ενός ψηφίου και οι μεταθέσεις.

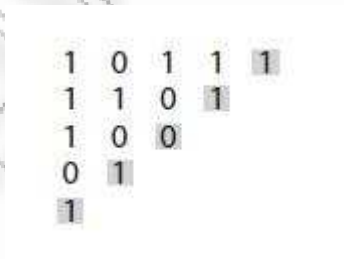
4.5. Ορθογώνιοι και τριγωνικοί κώδικες

Πάνω από ένα αθροίσματα ελέγχου μπορούν να ενσωματωθούν σε έναν κώδικα με το κάθε άθροισμα να ελέγχει διαφορετικό συνδυασμό συμβόλων κωδίκων ή να προσδίδει διαφορετικά βάρη στα σύμβολα. Γενικά, κάθε άθροισμα παρέχει επιπρόσθετη διόρθωση λαθών και ικανότητα διόρθωσης.



Εικόνα 2. Παράδειγμα ορθογώνιου κώδικα.

Μια ενδιαφέρουσα κατηγορία αποτελείται από τους ορθογώνιους κώδικες, ένα παράδειγμα των οποίων φαίνεται στην Εικόνα 2. Τα σύμβολα του μηνύματος είναι τα μη χρωματισμένα και σχηματίζουν ορθογώνιο. Τα χρωματισμένα με γκρι bits στην αριστερή στήλη είναι bits ισοτιμίας ορισμένα με τέτοιο τρόπο ώστε το άθροισμα της αντίστοιχης γραμμής να έχει άρτια ισοτιμία. Ομοίως, τα χρωματισμένα με γκρι bits στην τελευταία γραμμή είναι ορισμένα με τέτοιο τρόπο ώστε το άθροισμα κάθε στήλης να έχει άρτια ισοτιμία. Ο ορθογώνιος κώδικας μπορεί να διορθώσει ένα λάθος οπουδήποτε κι αν εμφανιστεί στον πίνακα. Για παράδειγμα, αν το λάθος είναι στην τρίτη γραμμή και τέταρτη στήλη, ο έλεγχος ισοτιμίας της τρίτης γραμμής και της τέταρτης στήλης δε θα είναι άρτιος και ως εκ τούτου αυτοί οι έλεγχοι θα δείξουν τη θέση του λάθους. Επίσης, μπορεί να διορθωθεί κι ένα λάθος σε κάποιο από τα bits ισοτιμίας. Πάνω στην ίδια ιδέα βασίζεται και ο τριγωνικός πίνακας που φαίνεται στην Εικόνα 3.



Εικόνα 3. Παράδειγμα τριγωνικού κώδικα.

Εδώ τα bits ισοτιμίας ορίζονται έτσι ώστε το άθροισμα των στοιχείων σε γραμμή και στήλη του bit να έχει άρτια ισοτιμία. Για παράδειγμα, το bit στο τέλος της δεύτερης γραμμής είναι 1, ώστε το άθροισμα της δεύτερης γραμμής και το τέταρτο στοιχείο της πρώτης γραμμής να είναι άρτιο.

5. Κατηγορίες κωδίκων καναλιού

Οι κώδικες καναλιού χωρίζονται σε γραμμικούς και μη γραμμικούς. Η δυσκολία όμως στην εξέταση και ανάλυση των μη γραμμικών κωδίκων, καθώς και η διεξαγωγή μη ασφαλών συμπερασμάτων οδήγησε στο να δοθεί μεγαλύτερη βαρύτητα στους γραμμικούς κώδικες, Οι γραμμικοί κώδικες χωρίζονται σε block κώδικες, συνελικτικούς και turbo.

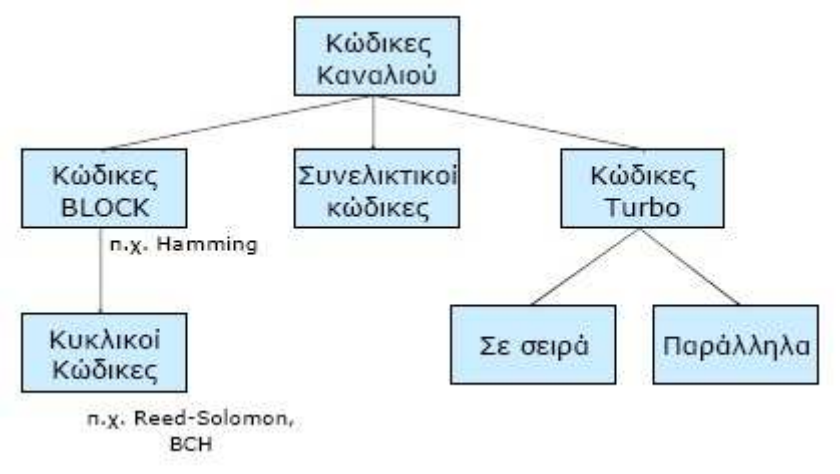
Σε έναν κώδικα block η ακολουθία των bits πληροφορίας έχει κατατμηθεί σε διαδοχικά block μήκους k και κάθε block έχει απεικονιστεί στην είσοδο του καναλιού με ένα block μήκους n κωδικών ψηφίων. Η απεικόνιση αυτή είναι ανεξάρτητη από τα προηγούμενα block, δηλαδή δεν υπάρχει μνήμη από ένα block προς ένα άλλο επόμενο. Ένα παράδειγμα των block κωδίκων είναι οι κώδικες Hamming και μια βασική υποκατηγορία τους είναι οι κυκλικοί κώδικες. Ένας κυκλικός κώδικας είναι ένας γραμμικός κώδικας block με την πρόσθετη ιδιότητα ότι αν c είναι μια κωδική λέξη, τότε μια κυκλική ολίσθηση των ψηφίων της είναι επίσης κωδική λέξη. Η δομή των κυκλικών κωδίκων βασίζεται στην αναπαράσταση της κωδικής λέξης, αλλά και του γεννήτορα πίνακα με πολυώνυμα. Οι κυκλικοί κώδικες μπορούν να κατασκευαστούν με ολισθητές καταχωρητές. Ένας εύκολα πραγματοποιήσιμος κυκλικός κώδικας βασίζεται στην παρατήρηση ότι το πολυώνυμο οποιασδήποτε κωδικής λέξης επιτυγχάνεται με τον πολλαπλασιασμό του πολυωνύμου γεννήτριας $g(p)$ επί το πολυώνυμο $X(p)$ που αντιστοιχεί στην ακολουθία της πληροφορίας εισόδου. Σε αυτήν την κατηγορία ανήκουν οι κώδικες Reed-Solomon. Τέλος, θα πρέπει να αναφερθεί ότι οι κωδικοποιητές και οι αποκωδικοποιητές των κυκλικών κωδίκων υλοποιούνται εύκολα.

Σους συνελικτικούς κώδικες υπάρχει ένας ολισθητής καταχωρητής μήκους $k_0 L$. Τα bits πληροφορίας μπαίνουν στον ολισθητή καταχωρητή k_0 bits κάθε φορά και στην έξοδο του κωδικοποιητή βγαίνουν n_0 bits, τα οποία είναι ένας γραμμικός συνδυασμός διάφορων bits του ολισθητή και στη συνέχεια μεταδίδονται μέσω του καναλιού. Τα n_0 bits της εξόδου δεν εξαρτώνται μόνο από τα περισσότερα k_0 bits που μπήκαν στον ολισθητή, αλλά επίσης και από τα $(L-1)k_0$ προηγούμενά του που περιέχονται στον ολισθητή και αποτελούν την κατάστασή του. Η ποσότητα L καλείται μήκος περιορισμού (constraint length) του συνελικτικού κώδικα και ο αριθμός δυνατών καταστάσεών του είναι $2^{(L-1)k_0}$. Ο κωδικός ρυθμός του συνελικτικού κώδικα ορίζεται ως $R_c = \frac{k_0}{n_0}$.

Συμπερασματικά, μπορούμε να πούμε ότι η κύρια διαφορά μεταξύ των

κωδίκων block και των συνελικτικών κωδίκων είναι ότι οι δεύτεροι έχουν μνήμη.

Οι κώδικες turbo είναι ένας εξελιγμένος και σύγχρονος συνδυασμός με πολλές δυνατότητες των δύο προηγούμενων κατηγοριών. Στην παρακάτω εικόνα απεικονίζεται η κατηγοριοποίηση των κωδίκων καναλιού.



Εικόνα 4. Κατηγορίες κωδίκων καναλιού

6. Γραμμικοί κώδικες (Linear codes)

Όλοι οι κώδικες στους οποίους αναφερόμαστε είναι δυαδικοί. Αυτό σημαίνει ότι έχουν στοιχεία από το πεδίο

$$F = GF(2) = \{0,1\} \qquad (3)$$

Στο πεδίο αυτό ισχύει ότι $1+1=0$ και $+1=-1$.

Τα στοιχεία του F καλούνται bits. Το F^n δηλώνει το διάνυσμα των n-διαστάσεων που αποτελείται από όλα τα δυαδικά διανύσματα μήκους n. Οι κώδικες με κωδικές λέξεις m που έχουν το ίδιο μήκος κωδικής λέξης n λέγονται ομοιόμορφοι δυαδικοί κώδικες (uniform binary codes). Αυτοί είναι κατά κανόνα γραμμικοί κώδικες εκτός αν ορίζονται διαφορετικά.

Ένας γραμμικός κώδικας C είναι ένα γραμμικό υποσύνολο του F^n, έχει ίσα ψηφία ίσα με n και είναι κακοί κώδικες διόρθωσης σφαλμάτων. Ισοδύναμα, ένας (δυαδικός) κώδικας είναι γραμμικός αν το άθροισμα των οποιωνδήποτε κωδικών λέξεων είναι επίσης μια κωδική λέξη.

Ένας (n, M, d) κώδικας είναι μια ομάδα από M δυαδικά διανύσματα μήκους n, που τα καλούμε κωδικές λέξεις, έτσι που κάθε δύο κωδικές λέξεις διαφέρουν μεταξύ τους τουλάχιστον κατά d θέσεις. Το n καλείται μήκος block του κώδικα και το d είναι η ελάχιστη απόσταση του κώδικα. Αναλυτικότερα, στους κώδικες block ένα block από k bits πληροφορίας ακολουθείται από μια ομάδα από c bits ψηφία ισοτιμίας που προκύπτουν κάθε φορά από το block των ψηφίων πληροφορίας.

Ακολουθούν κάποια παραδείγματα για το διαχωρισμό των κωδίκων.

Παράδειγμα 4

0 0 0 0 0

1 1 1 1 1

Ο παραπάνω είναι ένας (5, 2, 5) γραμμικός κώδικας όπου

$n=5$ *μήκος λέξης*

$M=2$ *αριθμός λέξεων στον κώδικα*

$d=5$ *απόσταση*

Επεξήγηση

$3+2=5$

$k+c=n=5$ *σύμβολα* $=5$ *ψηφία*

$k \mid c$ *Block κώδικας (χωριστά τα k με τα c σύμβολα)*

0 0 0 | 0 0 Διαφορές $5=d$

1 1 1 | 1 1 Επαναλαμβάνονται τα 1 διαδοχικά 5 φορές

$M=2$ *λέξεις*

Απόδοση $R=\frac{\log_2}{n}=\frac{1}{5}=0.2$

Παράδειγμα 5

0 0 0

0 1 1

1 0 1

1 1 0

Είναι ένας (3, 4, 2) γραμμικός κώδικας όπου

$n=3$ *μήκος λέξης*

$M=4$ *αριθμός λέξεων στον κώδικα*

$d=2$ *απόσταση*

Παράδειγμα 6

0 0 0 0 0 0 0

1 1 1 0 1 0 0

0 1 1 1 0 1 0

0 0 1 1 1 0 1

1 0 0 1 1 1 0

0 1 0 0 1 1 1

1 0 1 0 0 1 1

1 1 0 1 0 0 1

Είναι ένας (7, 8, 4) γραμμικός κώδικας (Κώδικας Hamming – Simplex)

Παράδειγμα 7

0 0 0 0 0 0 0

1 1 1 0 1 0 0

0 1 1 1 0 1 0

0 0 1 1 1 0 1

1 0 0 1 1 1 0

0 1 0 0 1 1 1

1 0 1 0 0 1 1

1 1 0 1 0 0 1

0 0 0 1 0 1 1

1 0 0 0 1 0 1

1 1 0 0 0 1 0

0 1 1 0 0 0 1

1 0 1 1 0 0 0

0 1 0 1 1 0 0

0 0 1 0 1 1 0

1 1 1 1 1 1 1 *Είναι ένας (7, 16, 3) γραμμικός κώδικας Hamming*

Παράδειγμα 8

0 0 0 0 0 0 0 0 *1 1 1 1 1 1 1 1*

1 1 0 0 0 0 0 0 *0 0 1 1 1 1 1 1*

1 0 1 0 0 0 0 0 *0 1 0 1 1 1 1 1*

1 0 0 1 0 0 0 0 *0 1 1 0 1 1 1 1*

1 0 0 0 1 0 0 0 *0 1 1 1 0 1 1 1*

1 0 0 0 0 1 0 0 *0 1 1 1 1 0 1 1*

1 0 0 0 0 0 1 0 *0 1 1 1 1 1 0 1*

1 0 0 0 0 0 0 1 *0 1 1 1 1 1 1 0*

Είναι ένας (8, 16, 2) όχι και τόσο καλός κώδικας (αν και αρκετά σπουδαίος). Ίδια n αλλά η μία λέξη προστιθέμενη στην άλλη δε δίνει μία τρίτη λέξη, δηλαδή δεν είναι κυκλικός κώδικας. Έτσι, αυτός ο κώδικας δεν είναι γραμμικός.

Αν πάρουμε ένα μέγιστο σύνολο από γραμμικά ανεξάρτητες κωδικές λέξεις από τον C, για παράδειγμα $x^{(1)}...x^{(k)}$, τότε ο C εμπεριέχει όλους τους γραμμικούς συνδυασμούς:

$$a_1 X^{(1)} + a_k X^{(k)}, a_1 \in F \qquad (4)$$

Η $k \times n$ δυαδική μήτρα

$$G = \begin{bmatrix} x^{(1)} \\ \vdots \\ x^{(k)} \end{bmatrix} \qquad (5)$$

καλείται γεννητριακή (generator) μήτρα του κώδικα. Ο κώδικας είναι ο γραμμικός χώρος του G.

7. Κώδικες δομής (Block Codes)

Σε ένα κώδικα δομής, ο κωδικοποιητής δέχεται στην είσοδό του μία σταθερού μήκους ακολουθία πληροφοριακών ψηφίων, το μήνυμα (message), που αποτελείται από k bits. Στη συνέχεια, μετατρέπει την πληροφορία αυτή σε μία ακολουθία n bits και σχηματίζει την κωδική λέξη (codeword). Το σημαντικό χαρακτηριστικό των κωδικοποιητών δομής είναι ότι αποτελούν διατάξεις χωρίς μνήμη, διότι δε χρησιμοποιούν ψηφία από προηγούμενα μπλοκ. Το block που παράγεται στην έξοδο εξαρτάται μόνο από το μπλοκ εισόδου.

Σε έναν (n,k) κώδικα δομής υπάρχουν 2^k ξεχωριστά μηνύματα. Αφού λοιπόν σε κάθε μήνυμα αντιστοιχεί μία και μόνο κωδική λέξη, θα υπάρχουν και 2^k ξεχωριστές κωδικές λέξεις, καθεμιά από τις οποίες έχει μήκος n. Ο ρυθμός R ενός κώδικα block ορίζεται ως εξής:

$$R = \frac{k}{n} \qquad 0 \le \mathrm{R} \le 1 \qquad (9)$$

Όσο υψηλότερος είναι ένας ρυθμός, τόσο λιγότερο πλεονασμό έχουμε. Άρα, ο ρυθμός του κώδικα καθορίζει και την ποσότητα του πλεονασμού.

Οι κώδικες δομής καθορίζονται από τρεις βασικές παραμέτρους : το μήκος της ομάδας n, το μήκος του μηνύματος k και την ελάχιστη απόσταση κώδικα $d_{\min}$. Ένας τέτοιος κώδικας συμβολίζεται ως $C(n,k,d_{\min})$. Όπως έχει αναφερθεί σε προηγούμενη παράγραφο, η απόσταση Hamming $d(x,y)$ μεταξύ δύο κωδικών λέξεων είναι ο αριθμός των θέσεων στις οποίες διαφέρουν και η ελάχιστη απόσταση κώδικα ορίζεται ως η μικρότερη απόσταση Hamming ανάμεσα στα δυνατά ζευγάρια κωδικών λέξεων. Για παράδειγμα ας θεωρήσουμε ότι ο κώδικας (10) έχει $\min = 3$.

$$C = \{00000,\ 10101,\ 01111,\ 11010\} \qquad (10)$$

Ας υποθέσουμε ότι ο κώδικας έχει μεταδοθεί και ο δέκτης έχει λάβει τη λέξη r:

$$(r_0, r_1, ..., r_{n-1}) = (c_0, c_1, ..., c_{n-1}) + (e_0, e_1, ..., e_{n-1}) \qquad (11)$$

,όπου η διαφορά μεταξύ r και c ονομάζεται διάνυσμα σφάλματος (error vector). Για παράδειγμα, έστω ότι έχει μεταδοθεί η κωδική λέξη 01111 και ο δέκτης έχει λάβει τη λέξη 01011. Στην περίπτωση αυτή έχουμε ένα σφάλμα στο τρίτο δυαδικό ψηφίο της κωδικής λέξης και το διάνυσμα σφάλματος είναι $\mathrm{e} = (00100)$.

Ένας κώδικας δομής είναι γραμμικός (linear) αν το άθροισμα δύο κωδικών λέξεων αποτελεί μία άλλη κωδική λέξη και αν ο κώδικας περιέχει και τη μηδενική κωδική λέξη. Ένας κώδικας δομής παράγεται από ένα σύνολο k γραμμικώς ανεξάρτητων n-διάστατων διανυσμάτων $g_0, g_1, ..., g_{k-1}$. Οι κωδικές λέξεις αποτελούν γραμμικό συνδυασμό αυτών των k n-διάστατων διανυσμάτων. Συνεπώς, η κωδική λέξη ενός μηνύματος $c = (c_0, c_1, ..., c_{k-1})$ μπορεί να αναπαρασταθεί με τη μορφή

$$v = c_0 \cdot g_0 + c_1 \cdot g_1 + \ldots + c_{k-1} \cdot g_{k-1} \qquad (12)$$

Τα k n-διάστατα διανύσματα που δημιουργούν τον κώδικα $g_0, g_1, ..., g_{k-1}$ μπορούν να αποτελέσουν τις γραμμές ενός πίνακα G διάστασης $k \times n$ όπως φαίνεται παρακάτω:

$$G = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \\ g_{k-1} \end{bmatrix} = \begin{bmatrix} g_{00} & g_{01} & \cdots & g_{0,k-1} \\ g_{10} & g_{11} & \cdots & g_{1,k-1} \\ g_{20} & g_{21} & \cdots & g_{2,k-1} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k-1,0} & g_{k-1,1} & \cdots & g_{k-1,n-1} \end{bmatrix}$$

Ο πίνακας G λέγεται γεννήτορας πίνακας ή γεννητριακή μήτρα (generator matrix) του κώδικα. Έτσι λοιπόν, η κωδική λέξη v για το μήνυμα c γράφεται ως εξής:

$$v = c \cdot G = c_0 \cdot g_0 + c_1 \cdot g_1 + \ldots + c_{k-1} \cdot g_{k-1} \qquad (13)$$

Δεδομένου του γεννήτορα πίνακα δηλαδή, είμαστε σε θέση να υπολογίσουμε από την προηγούμενη σχέση τις διακριτές κωδικές λέξεις που αντιστοιχούν σε όλες τις δυνατές ακολουθίες «0» και «1» ενός μπλοκ.

Για παράδειγμα, ένας (8,4) κώδικας δομής μπορεί να παραχθεί από τον γεννήτορα πίνακα:

$$G = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Τότε, το μήνυμα $c = (0\ 1\ 1\ 0)$ κωδικοποιείται ως εξής:

$$
\begin{aligned}
v &= c \cdot G = \\
&= 0 \cdot (01101010) + 1 \cdot (00011100) + 1 \cdot (10110011) + 0 \cdot (11000101) = \\
&= (00000000) + (00011100) + (10110011) + (00000000) = \\
&= (11001111)
\end{aligned}
$$

8. Κώδικες Hamming

Σύμφωνα με τη «Θεωρία της πληροφορίας» έχει αποδειχτεί ότι είναι εφικτή η αξιόπιστη μετάδοση πληροφορίας με τον ελάχιστο απαιτούμενο αριθμό πλεοναζόντων δυαδικών ψηφίων (redundancy). Αρχικά, είχε γίνει αποδεκτό ότι για την έγκυρη μετάδοση της πληροφορίας, έπρεπε να ενσωματωθούν στο προς αποστολή μήνυμα τα ψηφία ισοτιμίας (parity bits) με καθορισμένο τρόπο, δηλαδή 1 ψηφίο ισοτιμίας για κάθε ψηφίο του μηνύματος. Στη συνέχεια όμως, παρατηρήθηκε ότι η συγκεκριμένη μέθοδος, η οποία είχε προταθεί από τους θεμελιωτές της Θεωρίας των κωδίκων, απλά ανίχνευε τα σφάλματα και δεν προέβαινε στη διόρθωση αυτών. Η μόνη οδός για σωστό ανασχηματισμό του μηνύματος σε αυτή την περίπτωση, ήταν η εκ νέου μετάδοση του τμήματος εκείνου της πληροφορίας που περιείχε σφάλματα, από την πηγή πληροφορίας. Η επανάληψη της αποστολής όμως κοστίζει τόσο σε χρόνο όσο και σε πόρους, καθώς απαιτείται δίαυλος ανάδρασης, γεγονός που καθιστά τη συγκεκριμένη μέθοδο μη πρακτική.

Η έρευνα πέρασε στα χέρια του διακεκριμένου στη Θεωρία Κωδικοποίησης επιστήμονα της εποχής, Richard Hamming. Ο Hamming είναι πιο γνωστός από την εργασία του πάνω στους κώδικες Hamming για διόρθωση λαθών και για την έννοια της απόστασης Hamming, η οποία είναι κεντρική στη θεωρία κωδικοποίησης. Τα δεδομένα στα ψηφιακά συστήματα συνήθως αποθηκεύονται, εκπέμπονται και επεξεργάζονται σε δυαδική μορφή ως ομάδες από bits. Αν ένα εκ των bit είναι λάθος, το μήνυμα διαστρεβλώνεται ή ο υπολογισμός αποτυγχάνει. Σε μεγάλης κλίμακας υπολογιστές ή τηλεφωνικά συστήματα, είναι απαραίτητο ένα τεράστιο πλήθος υπολογισμών χωρίς ούτε ένα λάθος στο τελικό αποτέλεσμα. Ο Hamming ανέλαβε να κάνει τον ίδιο τον υπολογιστή να εντοπίζει και να διορθώνει τα μεμονωμένα λάθη, ώστε ο υπολογισμός να μπορεί να συνεχιστεί απρόσκοπτα με πιο αποδοτικό τρόπο από τον τότε ισχύοντα, ο οποίος ήταν να γίνεται το ίδιο πράγμα τρεις φορές και να επιλέγεται ένα προσεγγιστικό τελικό αποτέλεσμα.

Η προσέγγισή του εντασσόταν στη γενίκευση του ελέγχου ισοτιμίας (parity checking). Επαναλαμβάνουμε εν συντομία πώς δουλεύει ένας έλεγχος ισοτιμίας. Υποθέτουμε ότι έχουμε ένα σύνολο (block) από n bits και προσθέτουμε ένα επιπλέον bit, ώστε να γίνουν $n+1$ τα bits με σκοπό ολόκληρο το μήνυμα να έχει άρτιο πλήθος άσων μέσα του. Αυτό ονομάζεται έλεγχος άρτιας ισοτιμίας. Εξετάζοντας το μήνυμα ο αποδέκτης του, αν δεν υπάρχει άρτιο πλήθος άσων στο μήνυμα, βγάζει το συμπέρασμα ότι πρέπει να υπάρχει περιττό πλήθος λαθών στο μήνυμα. Αν τα λάθη στα bits συμβαίνουν ανεξάρτητα και αν το μήνυμα είναι μικρό και ο ρυθμός λαθών περιορισμένος, τότε πιθανότατα περιλαμβάνει μόνο ένα λάθος, αλλά δε γνωρίζουμε ποιο bit είναι λανθασμένο.

Οι κώδικες Hamming χρησιμοποιούν πολλαπλούς ελέγχους ισοτιμίας προκειμένου να προσδιορίσουν και να διορθώσουν λάθη του ενός bit. Κάθε έλεγχος τώρα είναι ένα άθροισμα ορισμένων bits σε επιλεγμένες θέσεις. Στην απλούστερη περίπτωση λέξεις μηνύματος μήκους $2^r - r - 1$, όπου r ακέραιος, πρόκειται να σταλούν μαζί με r bits ελέγχου, ώστε κάθε κωδική λέξη (τα bits του μηνύματος μαζί με τα bits ελέγχου) να περιλαμβάνει $2^r - 1$ bits. Οι θέσεις στην κωδική λέξη αριθμούνται από αριστερά προς τα δεξιά. Το πρώτο bit ελέγχου είναι στη θέση 1 και είναι ένας έλεγχος ισοτιμίας για τις θέσεις που έχουν το 1 ως λιγότερο σημαντικό (least significant) bit της δυαδικής τους αναπαράστασης (αυτές είναι οι θέσεις 1, 3, 5, 7, …). Το δεύτερο bit ελέγχου είναι στη θέση 2 και είναι ένας έλεγχος ισοτιμίας για τις θέσεις που έχουν το 1 στο δεύτερο λιγότερο σημαντικό (least significant) bit της δυαδικής τους αναπαράστασης (αυτές είναι οι θέσεις 2, 3, 6, 7, …). Το τρίτο bit ελέγχου είναι στη θέση 3 και είναι ένας έλεγχος ισοτιμίας για τις θέσεις που έχουν το 1 στο τρίτο λιγότερο σημαντικό (least significant) bit της δυαδικής τους αναπαράστασης (αυτές είναι οι θέσεις 4, 5, 6, 7, 12, …), κ.ο.κ.. Αν κανένας έλεγχος ισοτιμίας δεν αποτύχει, τότε ο κώδικας θεωρείται σωστός. Αν ένα bit στην κωδική λέξη είναι λάθος, το λάθος βρίσκεται στη θέση που η δυαδική αναπαράσταση ισούται με το πρότυπο του αποτυχημένου ελέγχου ισοτιμίας.

Η Εικόνα 5 δείχνει τον κώδικα για $r = 3$. Η πρώτη, η δεύτερη και η τέταρτη στήλη είναι οι θέσεις 1, 2 και 4 κάθε κωδικής λέξης. Οι τιμές τους μπορούν να υπολογιστούν από τα υπόλοιπα bits του μηνύματος, που αναπαριστούν τους αριθμούς από το 1 έως και το 15 στο δυαδικό σύστημα αρίθμησης.

Θέση **1 2 3 4 5 6 7**	**Δεκαδική** **Τιμή**
0 0 0 0 0 0 0	0
1 1 0 1 0 0 1	1
0 1 0 1 0 1 0	2
1 0 0 0 0 1 1	3
1 0 0 1 1 0 0	4
0 1 0 0 1 0 1	5
1 1 0 0 1 1 0	6
0 0 0 1 1 1 1	7

1 1 1 0 0 0 0	8
0 0 1 1 0 0 1	9
1 0 1 1 0 1 0	10
0 1 1 0 0 1 1	11
0 1 1 1 1 0 0	12
1 0 1 0 1 0 1	13
0 0 1 0 1 1 0	14
1 1 1 1 1 1 1	15

Εικόνα 5. Hamming κώδικας για r=3

Όπως φαίνεται στην Εικόνα 6, στο μετάλλιο που δόθηκε στον Hamming υπάρχει ο πίνακας με τους ελέγχους ισοτιμίας, ειδικότερα για την περίπτωση του πίνακα της Εικόνας 5 έχουμε :

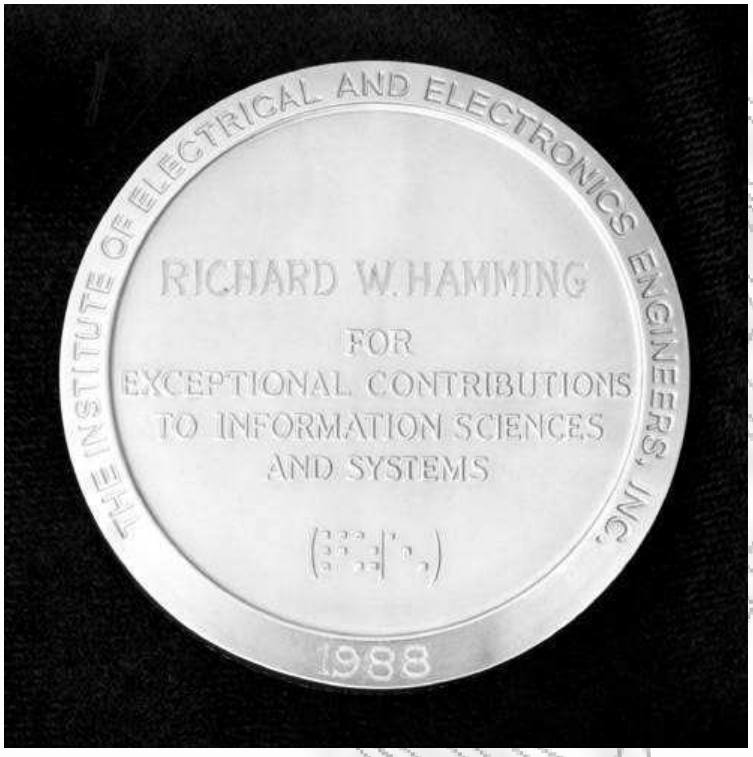
$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Ο πίνακας H χρησιμοποιείται ως εξής.

Έστω r ένα δυαδικό διάνυσμα μήκους 7 που αναπαριστά κάθε λαμβανόμενη λέξη, όχι απαραίτητα μια κωδική λέξη. Χρησιμοποιώντας αριθμητική Boolean υπολογίζουμε το

$$s = H \bullet r^T \qquad (14)$$

όπου s είναι ένα δυαδικό διάνυσμα μήκους 3. Αν $s = 0$, τότε το r είναι μια κωδική λέξη. Διαφορετικά, το s θα συμφωνεί με την i-οστή στήλη του H, τότε το i-οστό bit του r είναι λάθος και πρέπει να αντιστραφεί έτσι ώστε να λάβουμε τη σωστή κωδική λέξη.



Εικόνα 6. Το μετάλλιο με το οποίο τιμήθηκε ο Hamming από το IEEE.

Οι κώδικες Hamming μπορούν να ερμηνευθούν γεωμετρικά. Ορίζουμε την απόσταση Hamming ανάμεσα σε δύο κωδικές λέξεις ως τον αριθμό των θέσεων που διαφέρουν οι δύο κωδικές λέξεις. Η μικρότερη απόσταση Hamming μεταξύ των κωδικών λέξεων στον πίνακα της Εικόνας 5 είναι 3. Εφόσον ένα μεμονωμένο λάθος μετακινεί μια ληφθείσα λέξη σε απόσταση 1 από τη σωστή λέξη, μεμονωμένα λάθη μπορούν χωρίς αμφιβολία να διορθωθούν αλλάζοντας τη ληφθείσα λέξη στην κοντινότερη κωδική λέξη. Επιπλέον, οι κώδικες Hamming είναι τέλειοι από τη σκοπιά ότι κάθε ληφθείσα λέξη απέχει το πολύ απόσταση 1 από μια κωδική λέξη. Είναι εύκολο να πιστοποιήσουμε ότι ο αριθμός των κωδικών λέξεων επί τον αριθμό των λέξεων που δεν απέχουν μεγαλύτερη απόσταση από 1 από μια κωδική λέξη, ισούται με το συνολικό αριθμό δυνατών λέξεων. Αυτό σημαίνει ότι όταν κάθε τύπος ελέγχου αποτυγχάνει, στην πραγματικότητα έχουμε παραπάνω από ένα λάθος στη λέξη που έχει εκπεμφθεί. Πιο σύγχρονοι κώδικες, που προσπαθούν να διορθώσουν παραπάνω από ένα λάθος, είναι σπανίως τέλειοι, δηλαδή μερικοί τύποι λαθών που μπορεί να συμβούν στα bits δεν οδηγούν σε ξεκάθαρη αποκωδικοποίηση χωρίς αμφιβολίες.

Ο Hamming θεωρούσε ότι ένας κώδικας με ελάχιστη απόσταση Hamming μήκους $2t+1$, όπου t οποιοσδήποτε ακέραιος μπορούσε να διορθώσει t λάθη και ένας κώδικας με ελάχιστη απόσταση $2t+2$ μπορούσε να διορθώσει t λάθη και να εντοπίσει αλλά να μη διορθώσει $t+1$ λάθη.

Οι κώδικες που περιγράφηκαν παραπάνω είναι κώδικες που διορθώνουν ένα μεμονωμένο λάθος. Προσθέτοντας ένα επιπλέον bit σε κάθε λέξη μπορούν να εντοπίζουν δύο λάθη. Οι κώδικες αυτοί έλυσαν ένα μεγάλο μέρος του προβλήματος συντήρησης του εξοπλισμού των τηλεφωνικών εταιριών, ενώ τα «Hamming bits» εισήχθησαν στις μνήμες των υπολογιστών στα τέλη της δεκαετίας του ’50, όπως για παράδειγμα στον IBM 7030 Stretch supercomputer.

Μπορούμε να πούμε εν συντομία ότι οι κώδικες Hamming σχετίζονται με μικρές οικογένειες κωδίκων διόρθωσης πολλαπλών λαθών, που χρησιμοποιούνται σήμερα. Γενικά, ένας γραμμικός κώδικας διόρθωσης λαθών μπορεί να χαρακτηριστεί από μια διατεταγμένη τριάδα αριθμών (n,k,d), όπου n ο αριθμός των συμβόλων στις κωδικές λέξεις, k ο αριθμός των συμβόλων στις λέξεις του μηνύματος και d η μικρότερη απόσταση. Συνεπώς, ο κώδικας του πίνακα της Εικόνας 5 είναι ένας κώδικας $(7,4,3)$.

9. Κώδικες Χαμηλής Πυκνότητας Ελέγχου Ισοτιμίας (Low Density Parity Check Codes, LDPC)

Τις τελευταίες δεκαετίες το ενδιαφέρον των επιστημόνων της Θεωρίας Κωδίκων έχει επικεντρωθεί στην κατασκευή κωδίκων διόρθωσης σφαλμάτων πολύ καλά δομημένων, οι οποίοι διαθέτουν μεγάλη ελάχιστη απόσταση, $d_{\min}$. Η υψηλή ποιότητα των κωδίκων ως προς τη δομή τους, καθιστά αντιμετωπίσιμη την πολυπλοκότητα της αποκωδικοποίησης, ενώ παράλληλα η μεγάλη ελάχιστη απόσταση φέρεται να εξασφαλίζει την υψηλή απόδοση του κώδικα. Ωστόσο, η προσέγγιση αυτή δεν θα μπορούσε να μην έχει και ορισμένα μειονεκτήματα. Αρχικά, για να είναι ένα σχήμα κωδικοποίησης αξιόπιστο, η επιλογή των κωδίκων θα πρέπει να γίνει τυχαία. Το γεγονός αυτό όμως, έρχεται σε αντίθεση με το στόχο της θεωρίας κωδίκων, την κατασκευή δηλαδή πολύ καλά δομημένων κωδίκων, οι οποίοι παράλληλα χαρακτηρίζονται από ένα απλό σχήμα αποκωδικοποίησης. Επίσης, συγκρινόμενη με τη χωρητικότητα του διαύλου (channel capacity), η ελάχιστη απόσταση, σε πρακτικό επίπεδο, αποτελεί μία μικρότερου ενδιαφέροντος παράμετρο σχετικά με την απόδοση του κώδικα.

Από το 1993 και μετά, οι νεότερες τεχνικές κωδικοποίησης επέτρεψαν την κατασκευή κωδίκων των οποίων η απόδοση σε δίαυλο Προσθετικού Λευκού Γκαουσσιανού Θορύβου (AWGN) προσέγγιζε το όριο του Shannon με απόκλιση 1 dB. Οι μέθοδοι αυτές, όπως για παράδειγμα οι Turbo και οι LDPC κώδικες, χρησιμοποιούν μία εντελώς διαφορετική φιλοσοφία βασισμένη στα επαναληπτικά σχήματα κωδικοποίησης.

Οι κώδικες Χαμηλής Πυκνότητας Ελέγχου Ισοτιμίας (LDPC) αποτελούν μία κατηγορία γραμμικών κωδίκων δομής. Οι κώδικες αυτοί, καθώς και ο σχετικός επαναληπτικός αλγόριθμος κωδικοποίησης, προτάθηκαν από τον R. G. Gallager το 1960 στη διδακτορική του διατριβή, όμως δεν αξιοποιήθηκαν παρά μόνο στις αρχές της δεκαετίας του 1990. Ο λόγος για τον οποίο οι LDPC κώδικες είχαν κατά κάποιο τρόπο τεθεί στο περιθώριο ήταν το υπερβολικά μεγάλο, για τα δεδομένα της εποχής, υπολογιστικό κόστος που απαιτούσαν, καθώς και το γεγονός ότι οι υπολογιστικές μηχανές της εποχής δεν ήταν σε θέση να ανταπεξέλθουν στην πολυπλοκότητα του αλγορίθμου στον οποίο βασίζονταν οι LDPC κώδικες. Εξαίρεση αποτελεί το έργο του Tanner το 1981, ο οποίος εισήγαγε την αναπαράσταση των LDPC κωδίκων με γράφους, οι οποίοι ονομάζονται Γράφοι Tanner (Tanner Graphs) ή αλλιώς Διμερείς Γράφοι (Bipartite Graphs).

Οι κώδικες Χαμηλής Πυκνότητας Ελέγχου Ισοτιμίας (LDPC) μπορούν να αναπαρασταθούν με δύο τρόπους. Όπως το σύνολο των κωδίκων δομής, δύνανται να περιγραφούν μέσω πινάκων. Υπάρχει όμως και μία εναλλακτική μέθοδος απεικόνισης, η οποία χρησιμοποιεί γράφους Tanner. Εμείς θα ασχοληθούμε με τον πρώτο τρόπο αναπαράστασης.

Εφόσον οι κώδικες που μελετάμε ανήκουν στην κατηγορία των γραμμικών κωδίκων δομής, προκύπτουν από έναν γεννήτορα πίνακα G διάστασης $k \times n$, ενώ οι αριθμοί k και n αντιστοιχούν στον αριθμό ψηφίων του προς μετάδοση μηνύματος και της κωδικής λέξης αντίστοιχα. Όπως αναφέρθηκε και σε προηγούμενο κεφάλαιο, ο πίνακας αυτός δημιουργείται από ένα σύνολο k γραμμικώς ανεξάρτητων n-διάστατων διανυσμάτων, $g_0, g_1, \ldots, g_{k-1}$. Ο γεννήτορας πίνακας συνδέει το προς μετάδοση μήνυμα c με την κωδική λέξη ν, αφού κάθε κωδική λέξη γράφεται $\mathrm{v} = \mathrm{c} \cdot G$. Ο γεννήτορας πίνακας έχει τη μορφή $\mathrm{G} = [\mathrm{P}\ \mathrm{I}_k]$, δηλαδή αποτελείται από τον $\mathrm{k} \times (\mathrm{n} - \mathrm{k})$ πίνακα P και από τον $\mathrm{k} \times \mathrm{k}$ μοναδιαίο πίνακα I_k.

Γενικά, οι γραμμικοί κώδικες δομής, περιγράφονται κυρίως από τον Πίνακα Ελέγχου Ισοτιμίας (Parity Check Matrix), H. Ο πίνακας αυτός έχει τη μορφή $\mathrm{H} = [\mathrm{I}_{n-k}\ \mathrm{P}_T]$, δηλαδή προκύπτει από το συνδυασμό του μοναδιαίου $(\mathrm{n} - \mathrm{k}) \times (\mathrm{n} - \mathrm{k})$ I_{n-k} πίνακα με τον ανάστροφο του πίνακα P διάστασης $(\mathrm{n} - \mathrm{k}) \times \mathrm{k}$. Όπως γίνεται εύκολα αντιληπτό, ο πίνακας H έχει διάσταση $(\mathrm{n} - \mathrm{k}) \times \mathrm{n}$. Συγκεκριμένα, το πλήθος των γραμμών του αντιστοιχεί στο πλήθος των πλεοναζόντων ψηφίων ελέγχου που εισάγονται με την κωδικοποίηση, ενώ ο αριθμός των στηλών του ισούται με τον αριθμό των ψηφίων από τον οποίο αποτελείται μία κωδική λέξη. Ο πίνακας H πραγματοποιεί $m = n - k$ ελέγχους ισοτιμίας σε κάθε κωδική λέξη που φτάνει στον αποκωδικοποιητή. Οι LDPC κώδικες αποτελούν μία συγκεκριμένη κατηγορία γραμμικών κωδίκων δομής, της οποίας το βασικό χαρακτηριστικό συνίσταται στην χαμηλή πυκνότητα του πίνακα ελέγχου ισοτιμίας σε μη μηδενικά στοιχεία («1»). Αυτό σημαίνει ότι ο πίνακας H της συγκεκριμένης κατηγορίας κωδίκων αποτελείται κυρίως από μηδενικά στοιχεία και μόνο από έναν πολύ μικρό αριθμό άσων. Από το χαρακτηριστικό αυτό προκύπτει και η ονομασία των συγκεκριμένων κωδίκων (χαμηλής πυκνότητας).

Ακολούθως, δίνεται ένας πίνακας χαμηλής πυκνότητας ελέγχου ισοτιμίας για έναν (8,4) κώδικα, δηλαδή για κώδικα με μεταδιδόμενη πληροφορία αποτελούμενη από 4 ψηφία και με 4 ψηφία ελέγχου.

$$H = \begin{bmatrix} 1\,0\,0\,1\,1\,0\,1\,0 \\ 0\,0\,1\,0\,0\,1\,1\,1 \\ 1\,1\,1\,0\,0\,1\,0\,0 \\ 0\,1\,0\,1\,1\,0\,0\,1 \end{bmatrix}$$

Δύο σημαντικά μεγέθη που πρέπει να λαμβάνονται υπόψη σε έναν πίνακα ελέγχου ισοτιμίας είναι το πλήθος των μη μηδενικών στοιχείων σε κάθε γραμμή του πίνακα, w_r και το πλήθος των μη μηδενικών στοιχείων σε κάθε στήλη του πίνακα w_c. Τα μεγέθη αυτά ονομάζονται βαθμός γραμμής και βαθμός στήλης αντίστοιχα. Για να μπορεί να χαρακτηριστεί ένας πίνακας ως χαμηλής πυκνότητας πίνακας (low-density), θα πρέπει να ικανοποιούνται οι συνθήκες:

$$w_c \ll n \quad \text{και} \quad w_r \ll m \qquad (15)$$

Πιο αναλυτικά, θα πρέπει ο αριθμός των στοιχείων «1» σε μία στήλη του πίνακα να είναι κατά πολύ μικρότερος από το πλήθος των στηλών, δηλαδή από το μήκος της κωδικής λέξης και αντίστοιχα ο αριθμός των στοιχείων «1» σε μία γραμμή του πίνακα να είναι κατά πολύ μικρότερος από το πλήθος των γραμμών, δηλαδή από το μήκος του προς μετάδοση μηνύματος. Για την ικανοποίηση των ανωτέρω συνθηκών, ο πίνακας ελέγχου ισοτιμίας πρέπει να είναι πολύ μεγάλος. Συνεπώς, ο πίνακας του παραδείγματος δεν μπορεί να χαρακτηριστεί ως πίνακας χαμηλής πυκνότητας, απλά χρησιμοποιείται για την κατανόηση των χαρακτηριστικών ενός τέτοιου πίνακα.

10. Κυκλικοί Κώδικες

Η κυκλική μετατόπιση $\kappa(x)$ μιας λέξης x είναι η λέξη y που έχει ως πρώτο ψηφίο της το τελευταίο ψηφίο της x και τα υπόλοιπα ψηφία της προκύπτουν με απλή μετατόπιση κατά μία θέση προς τα δεξιά όλων των ψηφίων της x. Για παράδειγμα, $\kappa(010011) = 101001$ και $\kappa(101001) = 110100$. Με τη βοήθεια της συνάρτησης της κυκλικής μετατόπισης μπορούμε να ορίσουμε τους κυκλικούς κώδικες.

Ένας γραμμικός κώδικας C καλείται κυκλικός αν η κυκλική μετατόπιση κάθε κωδικής λέξης είναι και αυτή κωδική λέξη.

Παράδειγμα 6

Θεωρούμε τον κώδικα $C = \{000,\ 110,\ 101,\ 011\}$ *και ζητείται να εξεταστεί αν είναι κυκλικός κώδικας.*

Πρώτα ελέγχουμε αν είναι γραμμικός κώδικας, δηλαδή αν το άθροισμα οποιωνδήποτε δύο ή περισσότερων κωδικών λέξεων του C *είναι επίσης κωδική λέξη του* C*, που στην περίπτωσή μας ισχύει και κατόπιν εξετάζουμε τις κυκλικές μετατοπίσεις όλων των κωδικών λέξεων. Οι κυκλικές μετατοπίσεις είναι οι εξής:*

$\kappa(000) = 000$

$\kappa(110) = 011$

$\kappa(101) = 110$

$\kappa(011) = 101$

Αφού όλες είναι κωδικές λέξεις, ο κώδικας είναι κυκλικός.

Αναφορικά με τη συνάρτηση κυκλικής μετατόπισης $\kappa(.)$ ισχύει:

$$\kappa(x+y) = \kappa(x) + \kappa(y) \text{ και } \kappa(\alpha x) = \alpha\kappa(x) \qquad (16)$$

όπου x, y λέξεις και $\alpha \in \mathrm{K} = \{0,1\}$

Επομένως, για να δείξουμε ότι ένας γραμμικός κώδικας C είναι κυκλικός, αρκεί να δείξουμε ότι $\kappa(x) \in C \ \ \forall x$ που περιέχεται σε μία βάση του C. Έτσι, αν θέλουμε να κατασκευάσουμε έναν κυκλικό κώδικα C μήκους n, αρκεί να σχηματίσουμε το υποσύνολο S αποτελούμενο από μία λέξη x μήκους n και τις $(n-1)$ κυκλικές της μετατοπίσεις, $S = \{x, \kappa(x), \kappa(\kappa(x)), \ldots\}$. Αν ο κώδικας C είναι το γραμμικό ανάπτυγμα του S, δηλαδή $C = <S>$, τότε αφού το S περιέχει μια βάση του C, ο C είναι σύμφωνα με τα προηγούμενα κυκλικός κώδικας.

Στο Παράδειγμα 6, θα μπορούσαμε να ξεκινήσουμε από τη λέξη $x = 110$ και να σχηματίσουμε το $S = \{110, \kappa(110) = 011, \kappa(011) = 101\}$, του οποίου το γραμμικώς ανεξάρτητο υποσύνολο $\{110,\ 011\}$ είναι μια βάση του $C = \{000,\ 110,\ 101,\ 011\}$.

Η λέξη x που απαρτίζει μαζί με τις $(n-1)$ κυκλικές της μετατοπίσεις το S, γραμμικό ανάπτυγμα του οποίου είναι ο κώδικας C, ονομάζεται γεννήτορας του κυκλικού κώδικα C. Κάθε κώδικας μπορεί να έχει πολλούς γεννήτορες.

Οι κωδικές λέξεις μπορούν να παρασταθούν με πολυώνυμα. Ιδιαίτερα στην περίπτωση των κυκλικών κωδίκων, παρατηρούμε ότι αν μια λέξη u παριστάνεται από το πολυώνυμο $u(x)$, τότε η κυκλική μετατόπιση $\kappa(u)$ αναπαριστάται από το πολυώνυμο $x \cdot u(x) \bmod(1 + x^n)$.

<u>Θεώρημα</u>

Αν C είναι ένας κυκλικός κώδικας μήκους n, $\gamma(x)$ το πολυώνυμο γεννήτορας και $n-k$ ο βαθμός του, τότε ισχύουν τα ακόλουθα:

- Ο κώδικας C είναι διάστασης k
- Οι λέξεις που αντιστοιχούν στα πολυώνυμα $\gamma(x), x\gamma(x), x^2\gamma(x), \ldots, x^{k-1}\gamma(x)$ αποτελούν μια βάση του C
- Μία λέξη *c* ανήκει στον C αν και μόνο αν το αντίστοιχο πολυώνυμο $c(x)$ είναι το γινόμενο του γεννήτορα $\gamma(x)$ με κάποιο πολυώνυμο $a(x)$, δηλαδή αν $c(x) = a(x)\gamma(x) \bmod(1 + x^n)$.

11. Κώδικες BCH (Bose-Chaudhuri-Hocquenghem) και RS (Reed-Solomon)

Στην αλγεβρική θεωρία κωδικοποίησης οι λέξεις του μηνύματος, όπως και οι κωδικές λέξεις, αναπαριστώνται με πολυώνυμα με συντελεστές που λαμβάνουν τιμές από ένα Galois field $\mathrm{GF[q]}$ τάξεως q, όπου q είναι μια πρώτη δύναμη. Το πολυώνυμο $c[X]$ που αναπαριστά μια κωδική λέξη παράγεται από τον πολλαπλασιασμό του πολυωνύμου του μηνύματος $m[X]$ με ένα σταθερό πολυωνυμικό γεννήτορα $\mathrm{g[X]}$. Δύο σχετικές οικογένειες κωδίκων, οι οποίες επινοήθηκαν το 1960, επιτρέπουν τη διόρθωση τυχαίου αριθμού λαθών με τη χρήση κατάλληλου πλεονασμού (redundancy).

Οι κώδικες Bose-Chaudhuri-Hocquenghem (BCH) μπορούν να δημιουργηθούν πάνω στο $\mathrm{GF[2]}$. Ένας δυαδικός BCH με μήκος κωδικής λέξης $2^r - 1$ κι ελάχιστη απόσταση το λιγότερο 2t+1, για να μπορεί να διορθώσει t λάθη πάντα μπορεί να δημιουργηθεί με το πολύ $r \bullet t$ ψηφία ελέγχου. Τότε, ο κώδικας θα έχει απόδοση το λιγότερο

$$\left[2^r - 1, 2^r - 1 - r \bullet t, 2t + 1\right] \qquad (17)$$

Όταν $\mathrm{t = 1}$, αυτός ο BCH κώδικας είναι ισοδύναμος με τον κώδικα Hamming για διόρθωση ενός λάθους ανά συνδυασμό.

Οι κώδικες Reed-Solomon (RS) δημιουργούνται πάνω στο field $\mathrm{GF[q]}$ τάξεως q, όπου q είναι μια πρώτη δύναμη και $\mathrm{q > 2}$. Οι κώδικες Reed-Solomon είναι μέγιστης απόστασης διαχωρίσιμοι κώδικες, δηλαδή δημιουργούνται πάνω στο $\mathrm{GF[q]}$ με ελάχιστη απόσταση d, όπως περιγράφεται από την παρακάτω διατεταγμένη τριάδα.

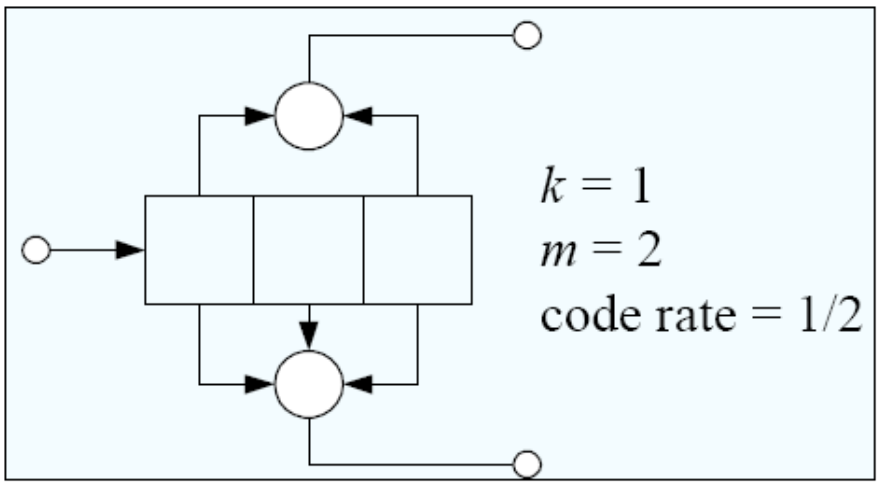
$$\left[q - 1, q - d, d\right]$$

Για παράδειγμα η NASA χρησιμοποιεί έναν κώδικα $[255, 233, 33]$ RS πάνω σε ένα $\mathrm{GF}[2^8]$ για διαπλανητική επικοινωνία. Οι κώδικες $[32, 28, 5]$ και $[28, 24, 5]$ χρησιμοποιούνται ευρέως για τη διόρθωση σε λάθη μήκους έως και 4000 bits στους οπτικούς δίσκους (compact disks).

12. Συνελικτικοί ή Συγκεραστικοί Κώδικες (Convolutional Codes)

Οι συνελικτικοί κώδικες είναι περισσότερο πολύπλοκοι από τους κώδικες δομής. Σε αντίθεση με τους κώδικες δομής που συγκεντρώνουν τα εισερχόμενα bits σε ομάδες (blocks) και παράγουν μεγαλύτερες ομάδες στην έξοδό τους, οι συνελικτικοί κώδικες μεταχειρίζονται τα δεδομένα εισόδου σαν μια συνεχή ακολουθία δεδομένων. Για κάθε k bits στην είσοδο, ο κωδικοποιητής παράγει m bits στην έξοδο.

Στο παρακάτω σχήμα φαίνεται ένας συνελικτικός κωδικοποιητής που χρησιμοποιεί το τρέχον bit και τα δύο προηγούμενα.



Εικόνα 7. Συνελικτικός κωδικοποιητής

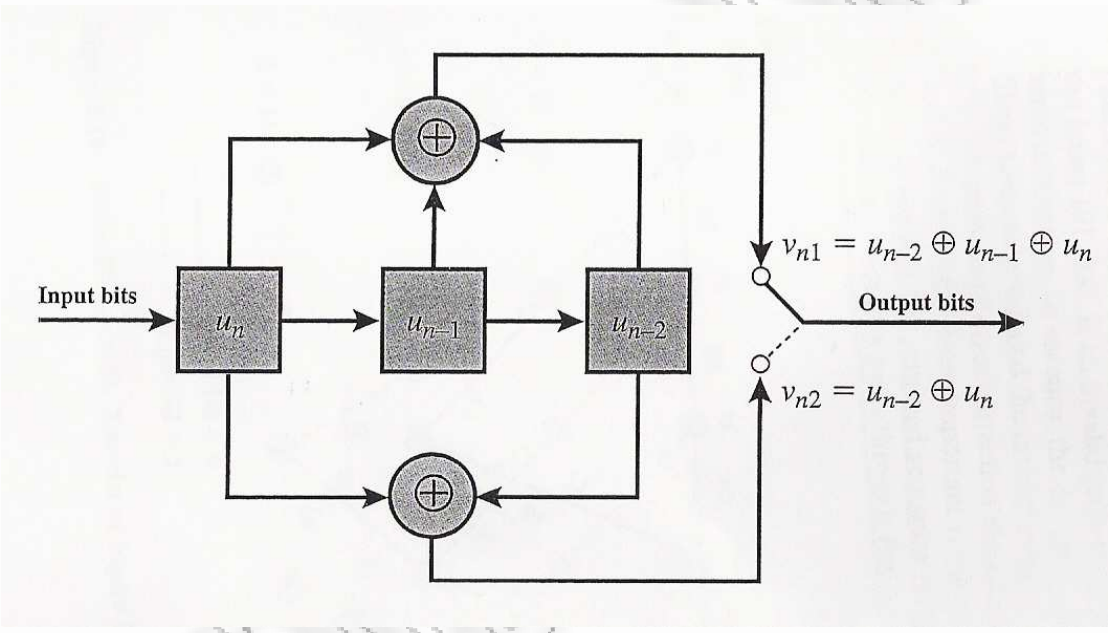
Ο ρυθμός R ενός συνελικτικού κώδικα είναι:

$$R = \frac{k}{m} \qquad 0 \le \mathrm{R} \le 1 \qquad (18)$$

Ένας συνελικτικός κώδικας ορίζεται από τρεις παραμέτρους : n, k, K. Ένας (n,k,K) κώδικας επεξεργάζεται κάθε φορά k ψηφία πληροφορίας εισόδου και παράγει μία έξοδο n ψηφίων για κάθε k εισερχόμενα ψηφία. Στη περίπτωση ενός συνελικτικού κώδικα, οι παράμετροι n και παίρνουν γενικά μικρές τιμές. Η διαφορά με τους κώδικες δομής είναι ότι οι συνελικτικοί κώδικες έχουν μνήμη, η οποία χαρακτηρίζεται από τον περιοριστικό

παράγοντα K (constraint factor). Στην ουσία, η τρέχουσα έξοδος n ψηφίων ενός (n,k,K) κώδικα εξαρτάται όχι μόνο από την τιμή του τρέχοντος συνόλου k ψηφίων εισόδου, αλλά και από τα προηγούμενα $K-1$ σύνολα των k ψηφίων εισόδου. Έτσι, η τρέχουσα έξοδος των n ψηφίων είναι μία συνάρτηση των τελευταίων $K*k$ ψηφίων εισόδου.

Στην Εικόνα 8 απεικονίζεται ένα συγκεκριμένο παράδειγμα ενός $(2,1,3)$ συνελικτικού κωδικοποιητή.



Εικόνα 8. Συνελικτικός κωδικοποιητής με (n, k, K) = (2, 1, 3)

Για έναν (n,k,K) κώδικα, ο καταχωρητής ολίσθησης περιέχει τα πιο πρόσφατα $K*k$ ψηφία εισόδου. Η αρχική κατάσταση των καταχωρητών είναι η μηδενική. Ο κωδικοποιητής παράγει n ψηφία εξόδου, μετά από τα οποία τα k παλαιότερα ψηφία του καταχωρητή αποβάλλονται και k νέα ψηφία εισάγονται. Επομένως, αν και η έξοδος των n ψηφίων εξαρτάται από τα $K*k$ ψηφία εισόδου, ο ρυθμός κωδικοποίησης είναι $\frac{k}{n}$. Οι πιο ευρέως χρησιμοποιούμενοι κωδικοποιητές έχουν $k=1$ και συνεπώς μήκος

καταχωρητή ολίσθησης K. Στο παραπάνω παράδειγμα ο κωδικοποιητής μετατρέπει ένα ψηφίο εισόδου u_n σε δύο ψηφία εξόδου v_{n_1} και v_{n_2}, χρησιμοποιώντας τα τρία πιο πρόσφατα ψηφία.

Για μία δοσμένη ακολουθία εισόδου k ψηφίων, υπάρχουν $2^{k(K-1)}$ διαφορετικές συναρτήσεις που μετατρέπουν τα k ψηφία εισόδου σε n ψηφία εξόδου. Τα τελευταία $(K-1)$ ψηφία εισόδου καθορίζουν το ποια συνάρτηση θα χρησιμοποιηθεί. Μπορεί επομένως να παρουσιαστεί ένας συνελικτικός κωδικοποιητής χρησιμοποιώντας μία μηχανή πεπερασμένης κατάστασης. Η μηχανή έχει $2^{k(K-1)}$ καταστάσεις και η μετάβαση από την μία κατάσταση στην άλλη καθορίζεται από τα k πιο πρόσφατα ψηφία εισόδου.

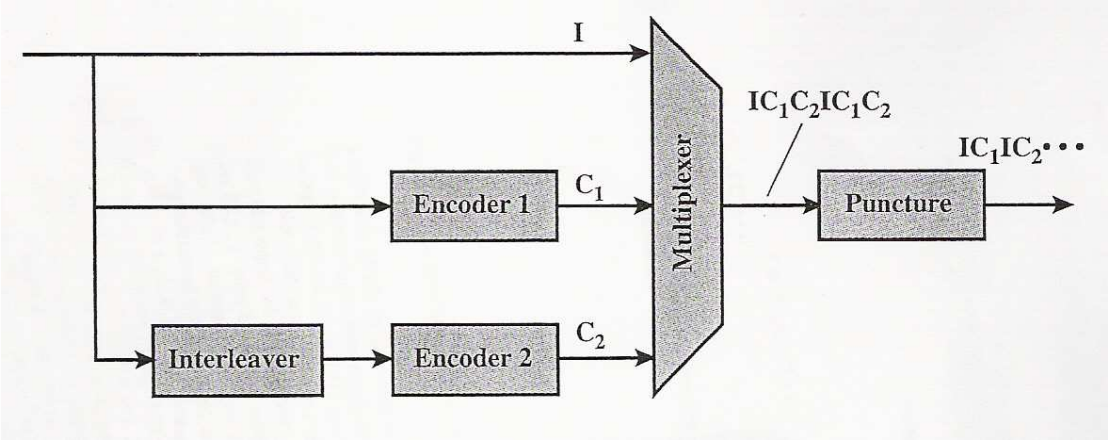
13. Κώδικες Turbo

Πριν από δεκαπέντε περίπου χρόνια, ο Berrou, ο Glavieux και ο Thitimajshima εισήγαγαν μια νέα προσέγγιση στην κωδικοποίηση διόρθωσης λαθών που έφερε επανάσταση στη θεωρία και στις τεχνικές κωδικοποίησης. Ανακάλυψαν ένα ψηφιακό σχήμα κωδικοποίησης που μπορούσε να παρέχει εικονικά αλάνθαστες επικοινωνίες σε μεγαλύτερους ρυθμούς δεδομένων και αποδόσεις μεταδιδόμενης ισχύος από ότι θεωρούσαν δυνατόν οι ειδικοί. Το σχήμα που ανακάλυψαν, το οποίο ονομάστηκε κώδικες turbo (turbo codes) σε αναλογία με τη μηχανή turbo και την αποτελεσματική χρήση της ανατροφοδότησης, οδήγησε σε τεχνικές κωδικοποίησης οι οποίες προσεγγίζουν τα απόλυτα όρια απόδοσης.

Η καινούρια τάξη κωδικών αποτελεί σημαντική επιλογή για τα ασύρματα συστήματα τρίτης γενιάς και συγκεκριμένα χρησιμοποιείται από τα συστήματα UMTS (Universal Mobile Telecommunications Systems) και τα κυψελωτά συστήματα τρίτης γενιάς cdma2000. Οι διαφορετικοί τύποι των κωδικοποιητών και αποκωδικοποιητών turbo στηρίζονται στη συνελικτική κωδικοποίηση.

Στην Εικόνα 9 φαίνεται ένας κωδικοποιητής turbo, ο οποίος αποτελείται από δύο όμοιους κωδικοποιητές. Ο πρώτος κωδικοποιητής (encoder 1) δέχεται μια ακολουθία ψηφίων εισόδου και παράγει για κάθε εισερχόμενο ψηφίο ένα ψηφίο ελέγχου C_1 (check bit). Η είσοδος στο δεύτερο κωδικοποιητή (encoder 2) είναι μια αναδιαταγμένη εκδοχή της ακολουθίας ψηφίων εισόδου παράγοντας έτσι μία ακολουθία ψηφίων ελέγχου C_2. Τα αρχικά ψηφία εισόδου σε συνδυασμό με τα δύο ψηφία ελέγχου πολυπλέκονται στη συνέχεια και παράγουν την ακολουθία $I_1C_{11}C_{21}I_2C_{12}C_{22},\ldots$, η οποία αποτελείται από το πρώτο ψηφίο εισόδου, που ακολουθείται από το πρώτο ψηφίο του πρώτου κωδικοποιητή, που ακολουθείται από το πρώτο ψηφίο του δεύτερου

κωδικοποιητή, κ.ο.κ. Η τελική ακολουθία έχει ρυθμό κωδικοποίησης $\frac{1}{3}$. Λαμβάνοντας μόνο τα μισά από τα ψηφία ελέγχου και εναλλάσσοντας τις εξόδους των δύο κωδικοποιητών, μπορεί να επιτευχθεί ρυθμός κωδικοποίησης $\frac{1}{2}$. Αυτή η διαδικασία ονομάζεται διάτρηση (puncturing). Οι ρυθμοί κωδικοποίησης $\frac{1}{3}$ και $\frac{1}{2}$ συναντώνται στα συστήματα τρίτης γενιάς.

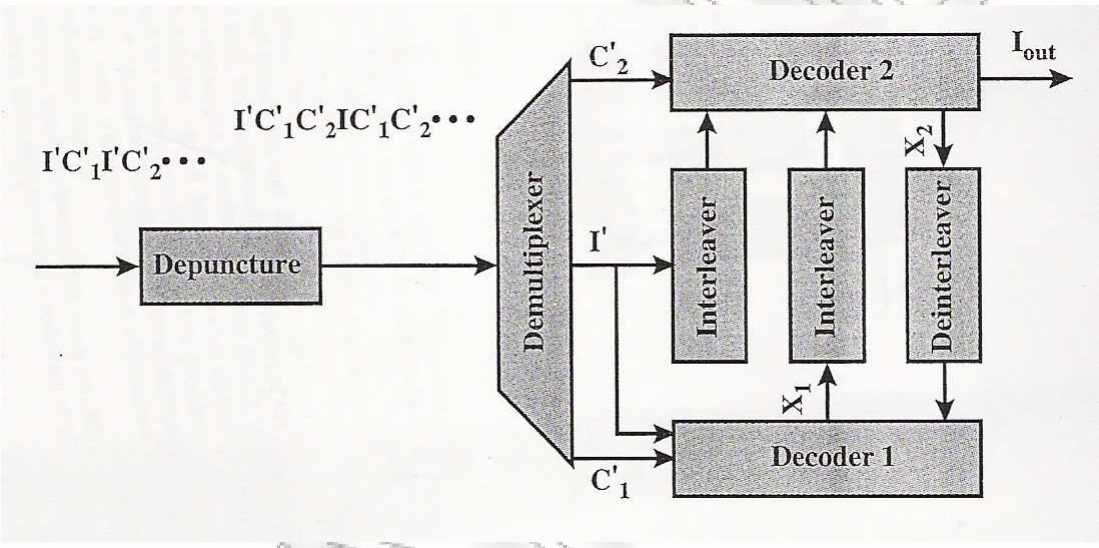


Εικόνα 9. Κωδικοποίηση turbo

Πρέπει να σημειωθεί ότι κάθε κωδικοποιητής παράγει ένα μόνο ψηφίο ελέγχου για κάθε ψηφίο εισόδου, καθώς επίσης και ότι το ψηφίο εισόδου διατηρείται. Για την κωδικοποίηση turbo χρησιμοποιείται μια παραλλαγή της συνελικτικής κωδικοποίησης, γνωστή ως επαναληπτική συστηματική συνελικτική κωδικοποίηση RSC (Recursive Systematic Convolutional).

Στην Εικόνα 10 φαίνεται ένα γενικό διάγραμμα του αποκωδικοποιητή turbo. Στα λαμβανόμενα δεδομένα εφαρμόζεται η αντίστροφη διαδικασία της διάτρησης (depuncturing), αν αυτό είναι απαραίτητο, εκτιμώντας τα ψηφία ελέγχου που λείπουν ή θέτοντάς τα ίσα με το μηδέν. Ο πρώτος αποκωδικοποιητής (decoder 1) λειτουργεί πρώτος χρησιμοποιώντας τις I' και PC_1' τιμές και παράγει τα ψηφία διόρθωσης X_1. Τα ψηφία I' και X_1 τροφοδοτούνται στο δεύτερο αποκωδικοποιητή, σε συνδυασμό με τις τιμές

C_2'. Για την ευθυγράμμιση των ψηφίων πρέπει να πραγματοποιηθεί αναδιάταξη (interleaving). Ο δεύτερος αποκωδικοποιητής (decoder 2) χρησιμοποιεί όλες του εισόδους και παράγει τις τιμές διόρθωσης X_2. Αυτές τροφοδοτούνται στον πρώτο αποκωδικοποιητή (decoder 1) για μια δεύτερη επανάληψη του αλγόριθμου αποκωδικοποίησης, αφού πρώτα πραγματοποιηθεί η αντίστροφη διαδικασία αναδιάταξης (deinterleaving). Μετά από αρκετές επαναλήψεις, παράγεται από τα I' και X_2 ένα ψηφίο εξόδου.



Εικόνα 10. Αποκωδικοποίηση turbo.

14. Επίλογος

Στην παρούσα διπλωματική εργασία παρουσιάστηκε διεξοδικά η κωδικοποίηση καναλιού και δόθηκε ιδιαίτερη έμφαση στους κώδικες καναλιού που χρησιμοποιούνται. Αναλύθηκαν οι κώδικες block, οι συνελικτικοί κώδικες και οι κώδικες turbo που θεωρούνται οι πιο σύγχρονοι κι εξελιγμένοι.

Μακάρι κι αυτή η διπλωματική εργασία να συμβάλει όσο το δυνατό περισσότερο στη μελέτη και περαιτέρω εξέλιξη του τομέα της πληροφορίας και των μεθόδων που χρησιμοποιούνται, ώστε να εξαλειφθούν τα προβλήματα που προκύπτουν από τη μετάδοσή της με λάθη ή από την ανελάστικότητά της σε τυχόν υποκλοπές.

15. Ασκήσεις

Άσκηση 1

Ζητούνται οι ρυθμοί πληροφορίας των κωδίκων $C_1 = \{00,10,01,11\}$, $C_2 = \{000,010,101,111\}$ και $C_3 = \{000000,000010,110001,111111\}$.

Απάντηση

Τα μήκη των κωδικών λέξεων των κωδίκων C_1, C_2 και C_3 είναι 2, 3 και 6 αντίστοιχα.
Το πλήθος των κωδικών λέξεων όλων των κωδίκων ισούται με 4.
Άρα, οι αντίστοιχοι ρυθμοί πληροφορίας είναι:

$$R_1 = \frac{\log_2 |C_1|}{n_1} = \frac{\log_2 4}{2} = 1$$
$$R_2 = \frac{\log_2 |C_2|}{n_2} = \frac{\log_2 4}{3} = \frac{2}{3}$$
$$R_3 = \frac{\log_2 |C_3|}{n_3} = \frac{\log_2 4}{6} = \frac{2}{6} = \frac{1}{3}$$

Άσκηση 2

Δίνονται οι λέξεις $x_1 = 000000$, $x_2 = 000010$, $x_3 = 110001$ και $x_4 = 111111$.
Ζητούνται τα βάρη όλων των λέξεων και οι αποστάσεις $d(x_1, x_2)$, $d(x_1, x_3)$, $d(x_2, x_3)$ και $d(x_3, x_4)$.

Απάντηση

Η λέξη $x_1 = 000000$ δεν περιέχει «1» κι επομένως $wt(000000) = 0$.
Η λέξη $x_2 = 000010$ περιέχει ένα «1» κι επομένως $wt(000010) = 1$.
Με τον ίδιο τρόπο $wt(110001) = 3$ και $wt(111111) = 6$.

Αναφορικά με την απόσταση $d(x_1, x_2)$, παρατηρούμε ότι οι λέξεις $x_1 = 000000$ και $x_2 = 000010$ διαφέρουν μόνο σε μία θέση (την πέμπτη), άρα $d(000000, 000010) = 1$.
Με τον ίδιο τρόπο, $d(000000,110001) = 3$, $d(000010,110001) = 4$ και $d(110001,111111) = 3$.

Άσκηση 3

Ποιοι από τους κώδικες $C_1 = \{101,\ 111,\ 011\}$, $C_2 = \{0000,\ 1001,\ 0110,\ 1111\}$, $C_3 = \{00000,\ 11100,\ 00111,\ 11011\}$ και $C_4 = \{000000,\ 101010,\ 010101,\ 111111\}$ είναι γραμμικοί και ποιες είναι οι αποστάσεις των γραμμικών κωδίκων;

Απάντηση

Ο κώδικας $C_1 = \{101,\ 111,\ 011\}$ δεν είναι γραμμικός, αφού η λέξη $010 = 101\ +\ 111$ δεν είναι κωδική.

Ο κώδικας $C_2 = \{0000,\ 1001,\ 0110,\ 1111\}$ είναι γραμμικός, αφού το άθροισμα κάθε δυνατού ζεύγους κωδικών λέξεων οδηγεί σε κωδική λέξη.

$1001 = 0000 + 1001$
$0110 = 0000 + 0110$
$1111 = 0000 + 1111$
$1111 = 1001 + 0110$
$0110 = 1001 + 1111$
$1001 = 0110 + 1111$

Ο κώδικας $C_3 = \{00000,\ 11100,\ 00111,\ 11011\}$ είναι γραμμικός, αφού το άθροισμα κάθε δυνατού ζεύγους κωδικών λέξεων οδηγεί σε κωδική λέξη.

$11100 = 00000 + 11100$
$00111 = 00000 + 00111$
$11011 = 00000 + 11011$
$11011 = 11100 + 00111$
$00111 = 11100 + 11011$
$11100 = 00111 + 11011$

Ο κώδικας $C_4 = \{000000,\ 101010,\ 010101,\ 111111\}$ είναι γραμμικός, αφού το άθροισμα κάθε δυνατού ζεύγους κωδικών λέξεων οδηγεί σε κωδική λέξη.

$101010 = 000000 + 101010$
$010101 = 000000 + 010101$
$111111 = 000000 + 111111$
$111111 = 101010 + 010101$
$010101 = 101010 + 111111$
$101010 = 010101 + 111111$

Άρα, όλοι οι κώδικες είναι γραμμικοί εκτός του C_1.

Η απόσταση ενός γραμμικού κώδικα είναι ίση με το ελάχιστο από τα βάρη των μη μηδενικών κωδικών λέξεων. Πολύ εύκολα βλέπουμε ότι οι μη μηδενικές κωδικές λέξεις 1001 και 0110 έχουν το πιο μικρό βάρος, ίσο με 2, το οποίο είναι και η απόσταση του κώδικα.
Οι αποστάσεις του κώδικα C_2 είναι $d_2 = 2$.

Ομοίως, $d_3 = 3$ και $d_4 = 3$.

Άσκηση 4

Ποιες είναι οι γεννητριακές μήτρες για τους κώδικες των παραδειγμάτων 4 με 7 ;

Απάντηση

Γεννήτρια μήτρα παραδείγματος 4.

$$G_{11.3} = [11111]$$

Γεννήτρια μήτρα παραδείγματος 5.

$$G_{11.4} = \begin{bmatrix} 01|1 \\ 10|1 \end{bmatrix}$$

Γεννήτρια μήτρα παραδείγματος 6.

$$G_{11.5} = \begin{bmatrix} 1110100 \\ 0111010 \\ 0011101 \end{bmatrix}$$

Γεννήτρια μήτρα παραδείγματος 7.

$$G_{11.6} = \begin{bmatrix} 1110100 \\ 0111010 \\ 0011101 \\ 1111111 \end{bmatrix}$$

Άσκηση 5

Θεωρούμε έναν πίνακα ελέγχου ισοτιμίας ενός κώδικα Hamming, μήκους 7 $(r = 3)$.

$$H = \begin{bmatrix} 1\,0\,1 \\ 0\,1\,1 \\ 1\,1\,1 \\ 1\,1\,0 \\ 1\,0\,0 \\ 0\,1\,0 \\ 0\,0\,1 \end{bmatrix}$$

Να βρεθεί η απόστασή του και ο ρυθμός πληροφορίας του.

Απάντηση

Μπορούμε να σχηματίσουμε ένα γεννήτορα πίνακα G του κώδικα.
Έτσι, από τον H μπορούμε να σχηματίσουμε τον ακόλουθο G.

$$G = \begin{bmatrix} 1\,0\,0\,0\,1\,0\,1 \\ 0\,1\,0\,0\,0\,1\,1 \\ 0\,0\,1\,0\,1\,1\,1 \\ 0\,0\,0\,1\,1\,1\,0 \end{bmatrix}$$

Επομένως, ο κώδικας έχει διάσταση *k* = 4.

Η απόσταση είναι $d = 3$ (ελάχιστο βάρος).
Το πλήθος των κωδικών λέξεων είναι $|C| = 16 = 24$ (αφού η διάσταση του κώδικα είναι 4).
Ο ρυθμός πληροφορίας είναι $R = \frac{4}{7}$, αφού από τα 7 bits μόνο τα 4 χρησιμοποιούνται για την παράσταση της πληροφορίας.

Άσκηση 6

Έστω $n = 4$ και $\gamma(x) = 1 + x^2$ το πολυώνυμο γεννήτορας ενός κώδικα C. Ζητείται η διάσταση και μία βάση του C.

Απάντηση

Σύμφωνα με το θεώρημα που αναφέρεται στην παράγραφο των κυκλικών κωδίκων, αφού ο βαθμός του $\gamma(x)$ είναι 2, η διάσταση του κώδικα C είναι $k = n - (n-k) = 4 - 2 = 2$ και μια βάση του είναι $\{1010,\ 0101\}$, αφού $\gamma(x) \leftrightarrow 1010$ και $x\gamma(x) = x + x^3 \leftrightarrow 0101$. Το γραμμικό ανάπτυγμα της βάσης είναι ο κώδικας $C = <\{1010,\ 0101\}> = \{0000,\ 1010,\ 0101,\ 1111\}$.

Άσκηση 7

Θεωρούμε τον κώδικα $C=\{0000,\ 1010,\ 0101,\ 1111\}$. Ζητείται ένας γεννήτορας πίνακας του C.

Απάντηση

Το πολυώνυμο ελάχιστου βαθμού που αντιστοιχεί σε μη μηδενική λέξη του και επομένως το πολυώνυμο γεννήτορας του C είναι $\gamma(x)=1+x^2$.
Αφού $k=n-\beta\alpha\theta\mu\acute{o}\varsigma(\gamma(x))=4-2=2$, ο γεννήτορας πίνακας του C έχει ως γραμμές τις λέξεις που αντιστοιχούν στα πολυώνυμα $\gamma(x)$ και $x\cdot\gamma(x)$, δηλαδή

$$P=\begin{bmatrix}\gamma(x)\\ x\cdot\gamma(x)\end{bmatrix}=\begin{bmatrix}1010\\ 0101\end{bmatrix}$$

Άσκηση 8

Θεωρούμε τον κώδικα $C=\{0000,\ 1010,\ 0101,\ 1111\}$. Ζητείται η κωδικοποίηση των ψηφίων πληροφορίας 01 και 10.

Απάντηση

Οι ακολουθίες των ψηφίων πληροφορίας αντιστοιχούν στα πολυώνυμα 1 και x και το γινόμενό τους με το πολυώνυμο – γεννήτορα $\gamma(x)=1+x^2$ είναι $1+x^2$ και $x+x^3$ αντίστοιχα. Επομένως, οι αντίστοιχες κωδικές λέξεις είναι 1010 και 0101.

Άσκηση 9

Θεωρούμε ένα κώδικα C μήκους $n=7$ με πολυώνυμο – γεννήτορα $\gamma(x)=1+x^2+x^3$. Ζητείται ένας γεννήτορας πίνακας του C, καθώς και η κωδικοποίηση των μηνυμάτων 1110 και 0110.

Απάντηση

Αφού ο βαθμός του πολυωνύμου – γεννήτορα είναι 3, η διάσταση του κώδικα C είναι $k=n-3=4$.

Επομένως, ένας πίνακας – γεννήτορας του C έχει ως γραμμές το $\gamma(x)=1+x^2+x^3$ και τις τρεις πρώτες κυκλικές μετατοπίσεις του, δηλαδή τα πολυώνυμα

$x\gamma(x)= x+x^3+x^4$

$x^2\gamma(x)= x^2+x^4+x^5$ και

$x^3\gamma(x)= x^3+x^5+x^6$.

Συνεπώς, ένας πίνακας – γεννήτορας του C είναι ο:

$$G=\begin{bmatrix}1\,0\,1\,1\,0\,0\,0\\0\,1\,0\,1\,1\,0\,0\\0\,0\,1\,0\,1\,1\,0\\0\,0\,0\,1\,0\,1\,1\end{bmatrix}$$

Τα μηνύματα 1110 και 0110 παριστάνονται από τα πολυώνυμα $1+x+x^2$ και $x+x^2$ και επομένως το γινόμενό τους με το πολυώνυμο – γεννήτορα είναι

$$\left(1+x+x^2\right)\gamma\left(x\right)=\left(1+x+x^2\right)\left(1+x^2+x^3\right)=1+x^2+x^3+x+x^3+x^4+x^2+x^4+\ x^5=1+x+x^5$$

και

$$\left(x+x^2\right)\gamma\left(x\right)=\left(x+x^2\right)\left(1+x^2+x^3\right)=x+x^3+x^2+x^5=x+x^2+x^3+\ x^5$$

που αντιστοιχούν στις κωδικές λέξεις 1100010 και 0111010.

Λήφθηκε υπόψη ότι $x^i+x^i=0$, αφού $1+1=0$ στο $K=\{0,\ 1\}$.

16. Ερωτήσεις πολλαπλής επιλογής

Ερώτηση 1

Γιατί είναι αναγκαία η κωδικοποίηση καναλιού;

Α) Για την ανίχνευση σφαλμάτων.

Β) Για τη διόρθωση σφαλμάτων.

Γ) Για τη βέλτιστη χρήση του εύρους ζώνης του καναλιού επικοινωνίας.

Δ) Όλα τα παραπάνω.

Ερώτηση 2

Ποιά συνθήκη ισχύει στον τύπο ρυθμού πληροφορίας δυαδικού κώδικα C μήκους n ($R = \frac{\log_2 |C|}{n}$);

Α) $|C| \geq 2^n$

Β) $1 \leq |C| \leq 2^n$

Γ) $|C| = 2^n$

Δ) $1 < |C| < 2^n$

Ερώτηση 3

Δίνονται οι λέξεις $x_1 = 000000$, $x_2 = 000010$, $x_3 = 110001$ και $x_4 = 111111$.
Ποιά είναι τα βάρη των λέξεων;

Α) 1, 2, 2 και 4 αντίστοιχα.

Β) 0, 2, 3 και 5 αντίστοιχα.

Γ) 0, 1, 3 και 6 αντίστοιχα.

Δ) 3, 4, 5 και 6 αντίστοιχα.

Ερώτηση 4

Πώς βρίσκουμε την ελάχιστη απόσταση ενός γραμμικού κώδικα δομής;

Α) Μετράμε πόσα ψηφία απέχει ο πρώτος άσος της πρώτης γραμμής από τον τελευταίο.

Β) Μετράμε πόσα ψηφία απέχει ο πρώτος άσος της πρώτης στήλης από τον τελευταίο.

Γ) Βρίσκουμε τον ελάχιστο αριθμό γραμμών του πίνακα των οποίων το άθροισμα ισούται με μηδέν.

Δ) Βρίσκουμε τον ελάχιστο αριθμό στηλών του πίνακα των οποίων το άθροισμα ισούται με μηδέν.

Ερώτηση 5

Δίνονται οι λέξεις $x_1 = 000000$, $x_2 = 000010$, $x_3 = 110001$ και $x_4 = 111111$. Ποιές είναι οι ελάχιστες αποστάσεις των λέξεων $d(x_1, x_2)$, $d(x_1, x_3)$, $d(x_2, x_3)$ και $d(x_3, x_4)$;

Α) 1, 2, 4 και 3 αντίστοιχα.

Β) 2, 1, 3 και 4 αντίστοιχα.

Γ) 2, 1, 4 και 3 αντίστοιχα.

Δ) 1, 3, 4 και 3 αντίστοιχα.

Ερώτηση 6

Πώς μπορεί να βελτιωθεί ένας επαναληπτικός κώδικας;

Α) Με την τεχνική Automatic Repeat Request (ARQ).

Β) Με την τεχνική Forward Error Correction (FEC).

Γ) Με την προσθήκη ψηφίου ελέγχου ισοτιμίας (parity bit).

Δ) Με τη χρήση ορθογώνιου κώδικα.

Ερώτηση 7

Οι κώδικες Hamming είναι υποδιαίρεση των κωδίκων δομής (block) και οι κώδικες Reed-Solomon είναι υποδιαίρεση των κυκλικών κωδίκων.

Α) Σωστό.

Β) Λάθος.

Ερώτηση 8

Δίνεται ο κώδικας $H = \begin{bmatrix} 000 \\ 011 \\ 101 \\ 110 \end{bmatrix}$.

Α) Ο κώδικας είναι κυκλικός.

Β) Ο κώδικας δεν είναι γραμμικός.

Γ) Ο κώδικας είναι γραμμικός και έχει απόσταση $d = 1$.

Δ) Ο κώδικας είναι γραμμικός και έχει απόσταση $d = 3$.

Ερώτηση 9

Μπορούν οι κώδικες Hamming να ερμηνευθούν γεωμετρικά;

Α) Ναι.

Β) Όχι.

Ερώτηση 10

Γιατί οι κώδικες Χαμηλής Πυκνότητας Ελέγχου Ισοτιμίας (LDPC), ενώ προτάθηκαν από το 1960, αξιοποιήθηκαν στις αρχές της δεκαετίας του ’90;

Α) Λόγω μεγάλου υπολογιστικού κόστους για τα δεδομένα της εποχής.

Β) Οι υπολογιστικές μηχανές της εποχής δεν μπορούσαν να ανταπεξέλθουν στην πολυπλοκότητα του αλγορίθμου στον οποίο βασίζονταν.

Γ) Όλα τα παραπάνω.

Δ) Κανένα από τα παραπάνω.

Ερώτηση 11

Ποιές συνθήκες πρέπει να ικανοποιούνται για να χαρακτηριστεί ένας πίνακας ως χαμηλής πυκνότητας (low density) πίνακας;

Α) $w_c << n$ και $w_r >> m$

Β) $w_c << n$ και $w_r << m$

Γ) $w_c << m$ και $w_r << n$

Α) $w_c << m$ και $w_r >> n$

Ερώτηση 12

Ποιά είναι η βασική διαφορά ανάμεσα στους κώδικες δομής και τους συνελικτικούς κώδικες;

Α) Οι κώδικες δομής είναι γραμμικοί, ενώ οι συνελικτικοί κώδικες δεν είναι.

Β) Οι κώδικες δομής δεν είναι γραμμικοί, ενώ οι συνελικτικοί κώδικες είναι.

Γ) Στους κώδικες δομής η έξοδος του κωδικοποιητή κάθε στιγμή εξαρτάται όχι μόνο από την τρέχουσα πληροφορία εισόδου, αλλά και από block δυαδικών ψηφίων που προηγήθηκαν, ενώ στους συνελικτικούς κώδικες κάθε διαδικασία κωδικοποίησης εξαρτάται μόνο από την τρέχουσα πληροφορία εισόδου.

Δ) Στους κώδικες δομής κάθε διαδικασία κωδικοποίησης εξαρτάται μόνο από την τρέχουσα πληροφορία εισόδου, ενώ στους συνελικτικούς κώδικες η έξοδος του κωδικοποιητή κάθε στιγμή εξαρτάται όχι μόνο από την τρέχουσα πληροφορία εισόδου, αλλά και από block δυαδικών ψηφίων που προηγήθηκαν.

Ερώτηση 13

Για την κωδικοποίηση turbo, η παραλλαγή ποιάς κωδικοποίησης χρησιμοποιείται;

Α) Της κωδικοποίησης δομής.

Β) Της κυκλικής κωδικοποίησης.

Γ) Της κωδικοποίησης LDPC.

Δ) Καμιάς από τις παραπάνω.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Wiley.Interscience.Elements.of.Information.Theory.Jul.2006

[2] Information Theory, Inference, and Learning Algorithms, David J.C. MacKay, University Press 2003

[3] INFORMATION SCIENCE, DAVID G. LUENBERGER, PRINCETON UNIVERSITY PRESS, Princeton and Oxford

[4] Θεωρία Πληροφορίας και Κωδικοποίησης, ΒΑΣΙΛΕΙΟΣ ΖΟΡΚΑΔΗΣ, ΠΑΤΡΑ 2002

[5] Ψηφιακές Επικοινωνίες, Γιώργος Φούσκας, Πάτρα 2002

[6] Θεωρία Πληροφοριών – Κώδικες, Δημ. Χ. Βούκαλης, Εκδόσεις Ίων 2009

[7] A Mathematical Theory of Communication, By C. E. SHANNON

[8] Θεωρία πληροφορίας και κωδίκων. Δρ. Βασίλης Διακολουκάς,

[9] Information Theory, INFORMATION THEORY AND THE DIGITAL AGE AFTAB, CHEUNG, KIM, THAKKAR, YEDDANAPUDI, Institute of Technology