EE728
Προχωρημένα Θέματα Θεωρίας Πληροφορίας
6η διάλεξη

Δημήτρης-Αλέξανδρος Τουμπακάρης

Τμήμα ΗΜ&ΤΥ, Πανεπιστήμιο Πατρών

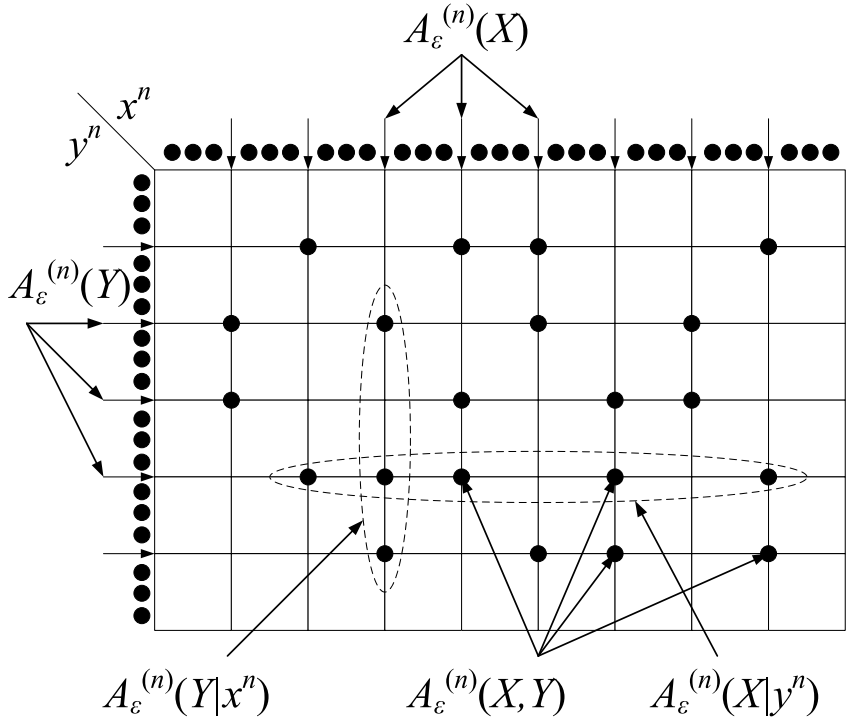
24 Μαρτίου 2010

Περιεχόμενα σημερινού μαθήματος

Αντιστοιχία με συγγράμματα

- Cover & Thomas: 7.5 – 7.7
- Gallager: Διαφορετική απόδειξη του ευθέος (όχι με τυπικότητα) στο Κεφ. 5. Περισσότερα σε επόμενο μάθημα.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης



Στο σχήμα δίνεται ένα παράδειγμα από κοινού τυπικού συνόλου. Υπάρχουν περίπου $2^{nH(X)}$ τυπικές ακολουθίες τ.μ. X και περίπου $2^{nH(Y)}$ τυπικές ακολουθίες τ.μ. Y. Ωστόσο, οι από κοινού τυπικές ακολουθίες είναι περίπου $2^{nH(X,Y)}$, δηλαδή, υπάρχουν ζεύγη τυπικών X^n με τυπικά Y^n τα οποία *δεν* είναι από κοινού τυπικά.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (2)

- Από την 3η ιδιότητα, η πιθανότητα ένα ζεύγος (X^n, Y^n) το οποίο επιλέγεται τυχαία και του οποίου οι συνιστώσες είναι (μεμονωμένα τυπικές) να είναι και από κοινού τυπικό, ισούται περίπου με $2^{-nI(X;Y)}$.
- Επομένως, στο σχήμα της προηγούμενης διαφάνειας, κατά μέσο όρο πρέπει να θεωρήσουμε περίπου $2^{nI(X;Y)}$ ζεύγη μεμονωμένα τυπικών X^n και Y^n έως ότου εμφανιστεί ένα τυπικό ζεύγος.
- Ισοδύναμα, εάν θεωρήσουμε μια ακολουθία Y^n η οποία αποτελεί την έξοδο καναλιού με είσοδο X^n, υπάρχουν περίπου $2^{nH(X|Y)}$ υπό συνθήκη τυπικές ακολουθίες X^n. Η πιθανότητα να διαλέξουμε μια ακολουθία X'^n η οποία είναι τυπική με την Y^n αλλά δεν είναι η ακολουθία X^n η οποία μεταδόθηκε ισούται, περίπου, με $2^{nH(X|Y)}/2^{nH(X)} = 2^{-nI(X;Y)}$. Επομένως, και πάλι, κατά μέσο όρο πρέπει να θεωρήσουμε περίπου $2^{nI(X;Y)}$ ακολουθίες X^n έως ότου εμφανιστεί ακολουθία που αποτελεί τυπικό ζεύγος με την Y^n.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (3)

- Συνεπώς, διαισθητικά, μπορούμε να μεταδώσουμε περίπου $2^{nI(X;Y)}$ διακριτές ακολουθίες στο κανάλι χωρίς να υπάρξει σύγχυση.
- Θα αποδείξουμε ότι είναι εφικτή η μετάδοση έως και $2^{nI(X;Y)}$ διακριτών ακολουθιών με αυθαίρετα μικρή πιθανότητα σφάλματος για $n \to \infty$. Θα αποδείξουμε, επίσης, ότι εάν προσπαθήσουμε να μεταδώσουμε περισσότερες από $2^{nI(X;Y)}$ διακριτές ακολουθίες, η πιθανότητα σφάλματος τείνει στο 1.

Το Θεώρημα Κωδικοποίησης Καναλιού

Θεώρημα Κωδικοποίησης Καναλιού – εισαγωγή

- Το Θεώρημα Κωδικοποίησης Καναλιού (Channel Coding Theorem) αποτελεί το πιο βασικό και το πιο διάσημο αποτέλεσμα της Θεωρίας Πληροφορίας.
- Σύμφωνα με το Θεώρημα Κωδικοποίησης Καναλιού, είναι εφικτή η μετάδοση σε κανάλια χωρίς μνήμη με ρυθμό αυθαίρετα κοντά στη χωρητικότητα και με αυθαίρετα μικρή πιθανότητα σφάλματος. Αντίστροφα, δεν είναι εφικτή μετάδοση με αυθαίρετα μικρή πιθανότητα σφάλματος εάν ο ρυθμός μετάδοσης υπερβαίνει τη χωρητικότητα του καναλιού.

Θεώρημα Κωδικοποίησης Καναλιού – εισαγωγή (2)

- Στη συνέχεια, θα διατυπώσουμε με την απαραίτητη λεπτομέρεια και θα αποδείξουμε το Θεώρημα Κωδικοποίησης Καναλιού.
- Το Θεώρημα Κωδικοποίησης Καναλιού (ευθύ - achievability) μπορεί να αποδειχτεί είτε με χρήση αποκωδικοποίησης Μέγιστης Πιθανοφάνειας (Maximum Likelihood decoding -- Gallager) είτε με χρήση Από Κοινού Τυπικών ακολουθιών (Cover).
- Στο μάθημα θα εξετάσουμε την απόδειξη με χρήση Από Κοινού Τυπικότητας η οποία είναι μάλλον πιο απλή.
- Το αντίστροφο του Θεωρήματος Κωδικοποίησης Καναλιού θα αποδειχτεί με χρήση της ανισότητας Fano.

Θεώρημα Κωδικοποίησης Καναλιού – εισαγωγή (3)

- Το βασικό ερώτημα (και, εκ πρώτης όψεως, παράδοξο) είναι το εξής: Πώς είναι δυνατόν να μεταδώσουμε με αυθαίρετα μικρή πιθανότητα σφάλματος σε ένα κανάλι που εισάγει σφάλματα με μη μηδενική πιθανότητα και με τυχαίο τρόπο;
- Για να απαντήσει στο ερώτημα, ο Shannon χρησιμοποίησε ένα διαφορετικό τρόπο σκέψης:
  - Δεν προσπάθησε να εκμηδενίσει την πιθανότητα σφάλματος, απλώς να την περιορίσει σε αυθαίρετα μικρές τιμές.
  - Βασίστηκε σε πολλές διαδοχικές χρήσεις του καναλιού ώστε να εκμεταλλευτεί το Νόμο των Μεγάλων Αριθμών.
  - Χρησιμοποίησε κώδικες οι οποίοι δημιουργούνται τυχαία και υπολόγισε τη μέση πιθανότητα σφάλματος.
- Αυτός ο τρόπος σκέψης διέπει τόσο την απόδειξη με χρήση τυπικότητας όσο και την απόδειξη με αποκωδικοποίηση Μέγιστης Πιθανοφάνειας.

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί

Ένας κώδικας (M, n) για το Διακριτό Κανάλι Χωρίς Μνήμη $(\mathcal{X}, p(y|x), \mathcal{Y})$ αποτελείται από

1. Ένα σύνολο δεικτών $\{1, 2, \ldots, M\}$.
2. Μια συνάρτηση κωδικοποίησης $X^n : \{1, 2, \ldots, M\} \to X^n$ η οποία παράγει κωδικές λέξεις (codewords) $x^n(1), x^n(2), \ldots, x^n(M)$. Το σύνολο των κωδικών λέξεων ονομάζεται βιβλίο κωδίκων (codebook).
3. Μια συνάρτηση αποκωδικοποίησης $g : \mathcal{Y}^n \to \{1, 2, \ldots, M\}$, η οποία αποτελεί ένα νομοτελειακό κανόνα ο οποίος αντιστοιχίζει ένα εκτιμώμενο δείκτη μεταδοθέντος μηνύματος, $\hat{m}$, σε κάθε ληφθείσα ακολουθία.

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί (2)

- Υπό συνθήκη πιθανότητα σφάλματος δεδομένου ότι εστάλη το μήνυμα με δείκτη i:

$$\lambda_i = \Pr\{g(Y^n) \neq i | X^n = x^n(i)\} = \sum_{y^n} p(y^n|x^n(i)) I(g(y^n) \neq i),$$

όπου $I(\cdot)$ η συνάρτηση δείκτης (ισούται με 1 όταν το όρισμά της αληθεύει, αλλιώς με 0).

- Η Μέγιστη Πιθανότητα Σφάλματος $\lambda^{(n)}$ κώδικα (M, n) ορίζεται ως

$$\lambda^{(n)} = \max_{i \in \{1,2,\ldots,M\}} \lambda_i.$$

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί (3)

- Η μέση (αριθμητικά) πιθανότητα σφάλματος $P_e^{(n)}$ κώδικα (M, n) ισούται με

$$P_e^{(n)} = \frac{1}{M} \sum_{i=1}^{M} \lambda_i.$$

- Όταν ο δείκτης μηνύματος W ακολουθεί ομοιόμορφη κατανομή, $P_e^{(n)} = \Pr\{W \neq g(Y^n)\}$, όπου Y^n η ακολουθία που λαμβάνεται στην έξοδο καναλιού όπου έχει μεταδοθεί η ακολουθία $X^n = x^n(W)$.
- Επίσης, $P_e^{(n)} \leq \lambda^{(n)}$.

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί (4)

Ορισμός Ο ρυθμός (rate) R κώδικα (M, n) ισούται με

$$R = \frac{\log M}{n} \text{ bits ανά μετάδοση.}$$

- Ένας ρυθμός R είναι εφικτός (achievable) όταν υπάρχει ακολουθία κωδίκων $\left(\lceil 2^{nR} \rceil, n\right)$ για την οποία η μέγιστη πιθανότητα σφάλματος $\lambda^{(n)}$ τείνει στο 0 καθώς το n τείνει στο άπειρο.

Ορισμός Η Χωρητικότητα λειτουργίας (operational capacity) ενός καναλιού ισούται με το μέγιστο ρυθμό ο οποίος είναι εφικτός.

- Το Θεώρημα Κωδικοποίησης Πηγής αποδεικνύει ότι η χωρητικότητα λειτουργίας $\max_R$ εφικτός R ισούται με την πληροφοριακή χωρητικότητα $\max_{p(x)} I(X;Y)$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού – Εισαγωγή

- Θα αναφερθούμε στην απόδειξη η οποία χρησιμοποιεί την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint AEP).
- Η ιδέα:
  - Στέλνουμε στο κανάλι ακολουθία $X^n = x^n(W)$ μήκους n.
  - Στην έξοδο του καναλιού λαμβάνουμε ακολουθία Y^n η οποία εξαρτάται από τη X^n, καθώς και από τον πίνακα μετάβασης, $p(y|x)$, του καναλιού.
  - Στο δέκτη αναζητούμε ακολουθία $\hat{X}^n$ η οποία να είναι από κοινού τυπική με την Y^n. Εάν υπάρχει, ο δέκτης θεωρεί ότι η $\hat{X}^n$ είναι η ακολουθία που μετέδωσε ο πομπός.
  - Από την Ιδιότητα από κοινού Ασυμπτωτικής Ισοδιαμέρισης, με μεγάλη πιθανότητα η ληφθείσα ακολουθία θα είναι από κοινού τυπική με τη μεταδοθείσα.
  - Ωστόσο, υπάρχει η πιθανότητα η Y^n να μην είναι από κοινού τυπική με καμία από τις πιθανές κωδικές λέξεις X^n ή να είναι από κοινού τυπική με άλλη ακολουθία από αυτή που μεταδόθηκε. Στην περίπτωση αυτή εμφανίζεται σφάλμα μετάδοσης.
  - Θα αποδείξουμε ότι, εάν $R < C$, καθώς το n τείνει στο άπειρο, η πιθανότητα σφάλματος τείνει στο 0.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού

Θεώρημα Θεώρημα Κωδικοποίησης Καναλιού:

- Σε ένα Διακριτό Κανάλι Χωρίς Μνήμη, όλοι οι ρυθμοί οι οποίοι είναι μικρότεροι από την πληροφοριακή χωρητικότητα είναι εφικτοί. Δηλαδή, για κάθε ρυθμό $R < C$, υπάρχει ακολουθία κωδίκων $(\lceil 2^{nR} \rceil, n)$ με μέγιστη πιθανότητα σφάλματος $\lambda^{(n)} \to 0$.
- Αντίστροφα, για οποιαδήποτε ακολουθία από κώδικες $(\lceil 2^{nR} \rceil, n)$ με $\lambda^{(n)} \to 0$ πρέπει να ισχύει $R \leq C$.

- Απόδειξη (ευθύ).

  Για απλοποίηση και χωρίς απώλεια γενικότητας υποθέτουμε ότι ο αριθμός κωδικών λέξεων $\lceil 2^{nR} \rceil$ είναι ακέραιος.

  Θεωρούμε δεδομένη πιθανότητα συμβόλων εισόδου $p(x)$ και δημιουργούμε 2^{nR} τυχαίες κωδικές λέξεις x^n μήκους n θεωρώντας ανεξάρτητες ομοίως κατανεμημένες (i.i.d.) τ.μ. x_i. Η πιθανότητα να δημιουργήσουμε μια συγκεκριμένη κωδική λέξη (ακολουθία) x^n ισούται με $p(x^n) = \prod_{i=1}^{n} p(x_i)$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (2)

- Οι 2^{nR} κωδικές λέξεις αποτελούν τις γραμμές του πίνακα

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ x_1(2) & x_2(2) & \dots & x_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(2^{nR}) & x_2(2^{nR}) & \dots & x_n(2^{nR}) \end{bmatrix}$$

- Η πιθανότητα να δημιουργηθεί ένας συγκεκριμένος τυχαίος κώδικας (πίνακας) $\mathcal{C}$ ισούται με $\Pr(\mathcal{C}) = \prod_{w=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(w))$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (3)

- Θεωρούμε την παρακάτω ακολουθία βημάτων
  1. Δημιουργείται ένας *τυχαίος* κώδικας $\mathcal{C}$ σύμφωνα με την κατανομή $p(x)$ όπως περιγράφτηκε παραπάνω.
  2. Ο κώδικας ανακοινώνεται στον πομπό και στο δέκτη. Επίσης, τόσο ο πομπός όσο και ο δέκτης γνωρίζουν τον πίνακα μετάβασης του καναλιού, $p(y|x)$.
  3. Επιλέγεται ένα μήνυμα W σύμφωνα με ομοιόμορφη κατανομή $\Pr\{W = w\} = 2^{-nR}$, $w = 1, 2, \ldots, 2^{nR}$.
  4. Στέλνεται στο κανάλι η $w-$οστή κωδική λέξη $X^n(w)$ η οποία αντιστοιχεί στη w-οστή γραμμή του πίνακα $\mathcal{C}$.
  5. Ο δέκτης λαμβάνει ακολουθία Y^n με δεσμευμένη κατανομή $p(y^n|x^n(w)) = \prod_{i=1}^{n} p(y_i|x_i(w))$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (4)

6. Ο δέκτης εκτιμά ποιο μήνυμα έχει σταλεί. Ο βέλτιστος δέκτης χρησιμοποιεί ανίχνευση Μέγιστης Πιθανοφάνειας (δεδομένου ότι θεωρούμε ομοιόμορφη κατανομή μηνυμάτων). Ωστόσο, όπως αναφέρθηκε, για την απόδειξη θα θεωρήσουμε ανίχνευση με βάση την από κοινού τυπικότητα. Παρόλο που ο δέκτης αυτός δεν είναι βέλτιστος, θα αποδείξουμε ότι, και σε αυτήν την περίπτωση, $\lambda^{(n)} \to 0$ για $n \to \infty$ (ο δέκτης είναι *ασυμπτωτικά* βέλτιστος). Ο δέκτης αποφασίζει (εκτιμά) ότι εστάλη το μήνυμα $\hat{W}$ εάν ικανοποιούνται ταυτόχρονα οι εξής δύο συνθήκες:
   a. Το ζεύγος ακολουθιών $\left(X^n(\hat{W}), Y^n\right)$ είναι από κοινού τυπικό.
   β. Δεν υπάρχει άλλος δείκτης μηνύματος $W' \neq \hat{W}$ για τον οποίο να ισχύει $\left(X^n(W'), Y^n\right) \in A_\epsilon^{(n)}$. Δηλαδή, δεν υπάρχει ακολουθία $X^n(W')$ που αντιστοιχεί σε μήνυμα $X^n(W') \neq \hat{W}$ (δηλαδή ανήκει στο βιβλίο κωδίκων) η οποία να είναι από κοινού τυπική με την Y^n.
7. Εάν $\hat{W} \neq W$, εμφανίζεται σφάλμα ανίχνευσης. Έστω $\mathcal{E}$ το ενδεχόμενο $\{\hat{W} \neq W\}$.

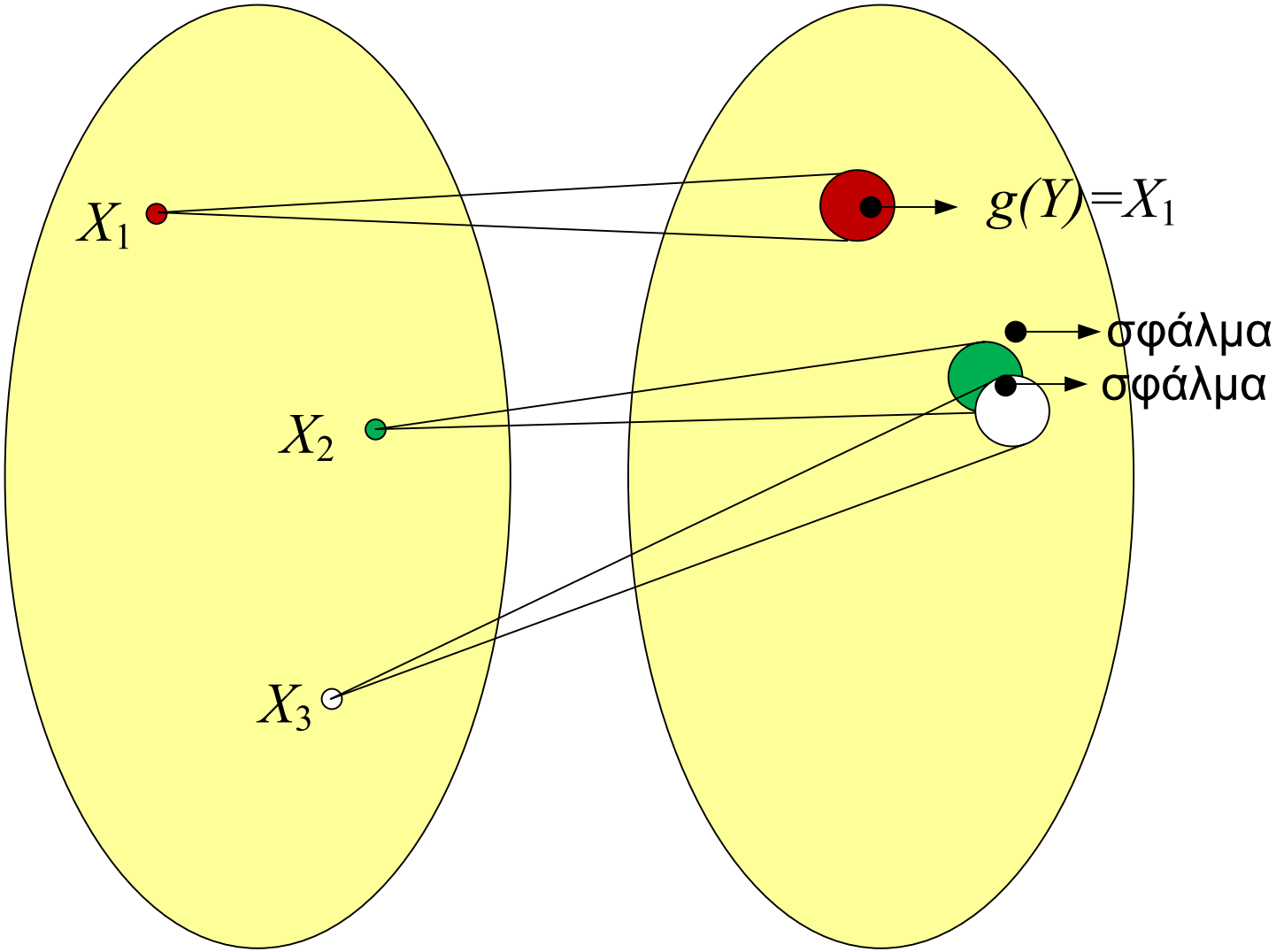
Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (5)
Ανάλυση της πιθανότητας σφάλματος – Εισαγωγή

- Η ιδέα: Αντί να υπολογίσουμε την πιθανότητα σφάλματος για ένα συγκεκριμένο κώδικα, θα υπολογίσουμε τη μέση πιθανότητα σφάλματος για τυχαία δημιουργία κωδίκων.
- Όταν χρησιμοποιείται αποκωδικοποίηση με χρήση από κοινού τυπικότητας, υπάρχουν δύο πηγές σφάλματος: Είτε η έξοδος Y^n δεν είναι από κοινού τυπική με την ακολουθία που εκπέμπει ο πομπός ή υπάρχει τουλάχιστον μια ακόμα κωδική λέξη η οποία είναι από κοινού τυπική με την Y^n.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (6)
Ανάλυση της πιθανότητας σφάλματος – Εισαγωγή

- Από την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης, η πιθανότητα η ληφθείσα ακολουθία να είναι από κοινού τυπική με την εκπεμφθείσα τείνει στο 1 για $n \to \infty$. Επίσης, η πιθανότητα η ληφθείσα ακολουθία να είναι από κοινού τυπική με ακολουθία διαφορετική από την εκπεμφθείσα ισούται περίπου με $2^{-nI(X;Y)}$. Επομένως, μπορούμε να χρησιμοποιήσουμε περίπου 2^{nI} κωδικές λέξεις και, ταυτόχρονα, να διασφαλίσουμε μικρή πιθανότητα σφάλματος.
- Στη συνέχεια θα αποδείξουμε τα παραπάνω και με την απαραίτητη μαθηματική αυστηρότητα.

Αποκωδικοποίηση με χρήση από κοινού τυπικότητας



Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (7)
Υπολογισμός Πιθανότητας Σφάλματος (I)

- Έστω ότι το μήνυμα W που εκπέμπεται επιλέγεται με ομοιόμορφη κατανομή από τα 2^{nR} πιθανά μηνύματα. $\mathcal{E} \triangleq \{\hat{W}(Y^n) \neq W\}$ είναι το ενδεχόμενο σφάλματος.
- Θα υπολογίσουμε τη μέση πιθανότητα σφάλματος για όλα τα πιθανά βιβλία κωδίκων.

$$\Pr\{\mathcal{E}\} = \sum_{\mathcal{C}} \Pr(\mathcal{C}) P_e^{(n)}(\mathcal{C}) =$$

$$= \sum_{\mathcal{C}} \Pr(\mathcal{C}) \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \lambda_w(\mathcal{C}) = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{\mathcal{C}} \Pr(\mathcal{C}) \lambda_w(\mathcal{C}).$$

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (8)
Υπολογισμός Πιθανότητας Σφάλματος (II)

- Δεδομένου ότι η αντιστοίχιση μηνυμάτων σε κωδικές λέξεις γίνεται τυχαία και επειδή για όλους τους πιθανούς κώδικες το μήνυμα W θα αντιστοιχίζεται κάθε φορά σε διαφορετική κωδική λέξη, η ποσότητα $\sum_{\mathcal{C}} \Pr(\mathcal{C})\lambda_w(\mathcal{C})$ είναι ανεξάρτητη του μηνύματος w. Επομένως, μπορούμε να υποθέσουμε, χωρίς απώλεια της γενικότητας, ότι επιστάλη η κωδική λέξη με δείκτη $w = 1$.
- Επομένως, η $\Pr(\mathcal{E})$ ισούται με

$$\begin{aligned}\Pr\{\mathcal{E}\} &= \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{\mathcal{C}} \Pr(\mathcal{C})\lambda_w(\mathcal{C}) \\ &= \sum_{\mathcal{C}} \Pr(\mathcal{C})\lambda_1(\mathcal{C}) = \Pr(\mathcal{E}|W = 1).\end{aligned}$$

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (9)
Υπολογισμός Πιθανότητας Σφάλματος (III)

- Ορίζουμε τα ενδεχόμενα $E_i = \{(X^n(i), Y^n) \in A_\epsilon^{(n)}\}$, $i \in \{1, 2, \ldots, 2^{nR}\}$, δηλαδή τα ενδεχόμενα η κωδική λέξη $X^n(i)$ (που αντιστοιχεί στο μήνυμα i) να είναι από κοινού τυπική με τη ληφθείσα ακολουθία Y^n η οποία προήλθε από μετάδοση της κωδικής λέξης $X^n(1)$.
- Συνεπώς,

$$\Pr(\mathcal{E}) = \Pr(\mathcal{E}|W=1) = P(E_1^c \cup E_2 \cup E_3 \cup \ldots \cup E_{2^{nR}}|W=1)$$
$$\leq P(E_1^c|W=1) + \sum_{i=2}^{2^{nR}} P(E_i|W=1).$$

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (10)
Υπολογισμός Πιθανότητας Σφάλματος (IV)

$$\Pr(\mathcal{E}) \leq P(E_1^c|W=1) + \sum_{i=2}^{2^{nR}} P(E_i|W=1).$$

- Από την ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης, η πιθανότητα η Y^n να μην είναι από κοινού τυπική με τη $X^n(1)$ τείνει στο 0 για $n \to \infty$: Επομένως, για κάθε $\epsilon > 0$ υπάρχει n_0 τέτοιο ώστε $P(E_1^c|W=1) \leq \epsilon$, για $n > n_0$.
- Επίσης, από τον τυχαίο τρόπο δημιουργίας του κώδικα, οι κωδικές λέξεις $X^n(1)$ και $X^n(i)$ είναι ανεξάρτητες μεταξύ τους για $i \neq 1$, με αποτέλεσμα η Y^n να είναι ανεξάρτητη από τις $X^n(i)$ για $i \neq 1$. Από την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης, η πιθανότητα οι $X^n(i)$ και Y^n να είναι από κοινού τυπικές ενώ επιλέχθηκαν ανεξάρτητα είναι $\leq 2^{-n(I(X;Y)-3\epsilon)}$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (11)
Υπολογισμός Πιθανότητας Σφάλματος (V)

- Συνδυάζοντας όλα τα παραπάνω,

$$\Pr(\mathcal{E}) \leq P(E_1^c|W=1) + \sum_{i=2}^{2^{nR}} P(E_i|W=1) \leq \epsilon + \sum_{i=2}^{2^{nR}} 2^{-n(I(X;Y)-3\epsilon)}$$
$$= \epsilon + \left(2^{nR}-1\right) 2^{-n(I(X;Y)-3\epsilon)} \leq \epsilon + 2^{-n(I(X;Y)-3\epsilon-R)} \leq 2\epsilon.$$

  Η τελευταία ανισότητα ισχύει εφόσον $n > n_1$ και $R < I(X;Y) - 3\epsilon$.

- Επομένως, εάν $R < I(X;Y)$, μπορούμε να επιλέξουμε n τέτοιο ώστε η μέση πιθανότητα σφάλματος υπολογισμένη επάνω σε όλους τους πιθανούς κώδικες και σε όλες τις πιθανές κωδικές λέξεις να μην υπερβαίνει το 2ϵ, για οποιοδήποτε $\epsilon > 0$.

- Δεν τελειώσαμε ακόμα... Πρέπει να δείξουμε ότι η μέγιστη πιθανότητα σφάλματος $\lambda^{(n)} \to 0$ και, επίσης, ότι υπάρχει τουλάχιστον ένας κώδικας με $\lambda^{(n)} \to 0$.