Version AGiF/2

EGT3
ENGINEERING TRIPOS PART IIB

Monday 6 May 2024 9.30 to 11.10

Module 4F5

ADVANCED INFORMATION THEORY & CODING

Answer not more than ***three*** *questions.*

All questions carry the same number of marks.

The ***approximate*** *percentage of marks allocated to each part of a question is indicated in the right margin.*

Write your candidate number ***not*** *your name on the cover sheet.*

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed
Engineering Data Book

10 minutes reading time is allowed for this paper at the start of the exam.

You may not start to read the questions printed on the subsequent pages of this question paper until instructed to do so.

You may not remove any stationery from the Examination Room.

1 (a) Determine the decimal digits $\alpha \leq 3$ and β so that

$$R_{33}(458923\alpha\beta37257834) = 25$$ [10%]

(b) Determine the inverse of 1357 in $\mathbb{Z}_{4080}$ using the extended Euclidean algorithm. [10%]

(c) Which of the following number(s) has an inverse in $\mathbb{Z}_{4080}$ and for those which have an inverse, express their inverse in Chinese Remainder notation: 531, 253, 532, 235. Note that $4080 = 15 \times 16 \times 17$, and by Chinese Remainder notation for a number x, we mean $[R_{15}(x), R_{16}(x), R_{17}(x)]$. [10%]

(d) What are the multiplicative and the additive orders of $1+X^2$ in Galois field GF(128)? [10%]

(e) For each of the following Galois fields, either give examples of one element of multiplicative order 2 and one element of additive order 2, or explain why there cannot be such an example: GF(5), GF(31), GF(128), GF(127^{25}). [10%]

(f) After many decades of bitter animosity and sending secret messages behind each other's backs, Eve and Alice have finally made up. Eve suggests to Alice that they send a secret birthday card to Bob and sign it in a manner that proves to Bob that they are now friends and writing the card together.

In a toy RSA system using small numbers, Eve's public key is $(319, 33)$, Alice's public key is $(299, 35)$ and Bob's public key is $(323, 55)$. We assume that only Eve knows that $319 = 11 \times 29$, only Alice knows that $299 = 13 \times 23$, and only Bob knows that $323 = 17 \times 19$.

(i) Bob receives the encrypted message $y = 206$, encrypted using his public key. Work out what the secret message was. [15%]

(ii) Now consider another secret message x. This is encrypted by Eve and Alice using Bob's public key to yield the encrypted message $y = 25$. Eve and Alice append their signatures $s_E = 284$ and $s_A = 64$, respectively. Show how Bob can verify Eve and Alice's signatures for the secret message he decrypted. [15%]

(iii) Do these signatures suffice to prove their friendship to Bob? Discuss the weakness of this scheme and propose a way it could be improved in a manner that proves beyond doubt that Eve and Alice both approve of the message content and confirm each other's co-authorship, without revealing their secret keys to each other. [20%]

2 A Reed-Solomon code is to be operated over the Galois field GF(41).

(a) What are the possible code lengths for a Reed-Solomon code over GF(41)? [10%]

(b) Specify the Discrete Fourier Transform matrix over GF(41) with $\alpha = 10$. [10%]

(c) An information sequence of length 3 is mapped to a codeword by prepending the information sequence with two zeros and multiplying the resulting vector by the inverse DFT matrix. What is the rate of this code and what is its parity-check matrix? [10%]

(d) What is the encoder matrix of the code? [15%]

(e) A codeword was transmitted and received with at most one error as $[12, 0, 30, 5, 0]$. Recover the encoded information sequence. [15%]

(f) Another codeword is transmitted over an erasure channel and received as $[?, 8, 16, 8, ?]$. Recover the encoded information sequence. [20%]

(g) The radio engineers who asked for this Reed-Solomon code to be designed insist that you use a systematic encoder as they want the option of recovering the first possibly noisy 3 symbols without using the decoder when the receiver operates in low power mode. Find the systematic encoder matrix for this code. [20%]

3 (a) Consider the discrete modulo-additive channel described in Fig. 1 below, with input and output alphabets $\mathcal{X} = \mathcal{Y} = \{0, 1, 2, \ldots, K\}$. Let δ denotes the probability that the noise is non-zero.

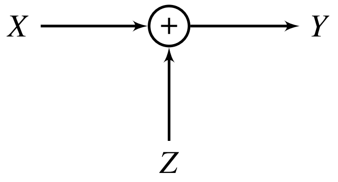

Fig. 1

(i) Write the channel transition probability matrix $P_{Y|X}$ for this channel. [10%]

(ii) We consider reliable transmission over this channel by means of a code $\mathcal{C}$ of M codewords of length n. Give an expression of the decoder that attains lowest error probability. [10%]

(iii) Show that for $\delta < \frac{1}{K}$ the decoder is equivalent to finding the nearest codeword to the received sequence $y^n \in \mathcal{Y}^n$ in terms of Hamming distance. [20%]

(iv) Explain the meaning of the function

$$E_0(\rho) = -\log \sum_y \left(\sum_x P_X(x) P_{Y|X}(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho}$$

in the context of reliable transmission and write it for the above channel assuming equiprobable inputs. [15%]

(v) Explain how to find the mutual information of the channel from $E_0(\rho)$ and provide an expression. [10%]

(vi) For what value of δ is the capacity zero? [10%]

(b) We now attempt to decode with a decoder that assumes that the probability that the noise is non-zero is γ, not necessarily equal to δ.

(i) If $\gamma < \frac{1}{K}$, show that the decoder is equivalent to finding the nearest codeword to the received sequence $y^n \in \mathcal{Y}^n$ in terms of Hamming distance. [10%]

(ii) Explain the implications of part (b)(i) in the achievable rate obtained with the new decoder. [15%]

4 (a) Consider a hypothesis testing problem between two i.i.d. distributions P_1, P_2 defined over alphabet $\mathcal{Y}$.

(i) Explain the Neyman-Pearson setting and the corresponding optimality criterion. [10%]

(ii) Provide an expression of the optimal test. [10%]

(iii) For i.i.d. observation $y^n = (y_1, \ldots, y_n)$, express the optimal test as a function of the relative entropy between the empirical distribution of the observation $\hat{P}_{y^n}$ and the testing distributions. Explain the operation of the test. [15%]

(iv) If the i.i.d. observation $y^n = (y_1, \ldots, y_n)$ is generated according to distribution Q, different from P_1, P_2, show that in the limit for $n \to \infty$ the test in part (a)(iii) decides hypothesis P_1 whenever [15%]

$$D(Q\|P_2) > D(Q\|P_1)$$

(b) Consider a discrete source V^n over alphabet $\mathcal{V}^n$.

(i) Denote by $\mathcal{A}$ the set of M most probable source sequences. Write the error probability of the fixed-length code that assigns a unique index to each element of $\mathcal{A}$ and an error index to all the rest. [5%]

(ii) Show that no other code can have a lower error probability. [5%]

(iii) Show that, for any real number $r > 1$, the probability of correct decoding, p_c, is such that

$$(p_c)^r \leq M^r \sum_{v^n \in \mathcal{A}} \frac{1}{M} P_{V^n}(v^n)^r$$

(Hint: Jensen's inequality states that for a convex function f and random variable X, $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$). [20%]

(iv) Hence show that for $-1 < \rho < 0$

$$p_c \leq M^{-\rho} \Bigg(\sum_{v^n \in \mathcal{V}^n} P_{V^n}(v^n)^{\frac{1}{1+\rho}} \Bigg)^{1+\rho}$$

[10%]

(v) Hence show that the error probability of a discrete memoryless source satisfies

$$p_e \geq 1 - e^{-n(\rho R - E_s(\rho))}$$

for $-1 < \rho < 0$, where $R = \frac{1}{n} \log M$, and $E_s(\rho) = (1+\rho) \log \sum_{v \in \mathcal{V}} P_V(v)^{\frac{1}{1+\rho}}$. [10%]

END OF PAPER