Kódolástechnika

forrás $\xrightarrow{u}$ kódoló $\xrightarrow{c}$ csatorna $\xrightarrow{v}$ dekódoló $\xrightarrow{u'}$ vevő u, c, v, u' mind vektorok

Jelölje $d(v,c)$ azon i pozíciók számát melyre $c_i \neq v_i$, ezt c és v vektor Hamming távolságának hívjuk. A feladat tehát azt a c' kód megtalálása minden $d(u,v)$ közül a lehetséges vektoret keressük, akkor megkeressük azt a c' értéket melyre $d(c',v) = \min_{c \in C} d(c,v)$. A feladat ezt a legegyszerűbb de legrosszabb megoldás kiszámolása kérdés, ezt előre is megtehetjük: minden v-hez minden c-t kiszámolunk 1 egy táblázatban eltároljuk, ez a ...

Ismétléses kódolás

$u \quad c_1 c_2 c_3$
$0 \to 000$
$1 \to 111$

A kód egy hibát képes javítani, hisz ilyenkor (...) a $d(c,v)$ megfelelő. Legyen az ismétlések száma $n \geq 3$ plus szám, ekkor a kód $\frac{(n-1)}{2}$ hibát képes javítani.
A kód egy hibát képes javítani
2 vagy hibát jelezni.

Egyszerű paritáskód

$u_1 u_2 \quad c_1 c_2 c_3$
$0\,0 \to 0\,0\,0$
$0\,1 \to 0\,1\,1$
$1\,0 \to 1\,0\,1$
$1\,1 \to 1\,1\,0$

A kód egy hibát képes jelezni, de nem képes hibát javítani.
A kódoló legfontosabb tulajdonsága a C kód egy paramétere melyet kódtávolságnak nevezünk s d_{min}-nel jelölünk: $d_{min} = \min_{c \neq c',\ c,c' \in C} d(c,c')$, $d_{min} > d(v,c)$ tehát a hibák
A kódoló 1 független (...) tulajdonsága a C kód egy paramétere ... csak detektálni akarjuk a hibát, ezt csak akkor tudjuk megtenni ha v nem kódszó.
A hibajelzés azon feladat, ... kódtávolságú kód max $d_{min}-1$ hibát tud jelezni.
száma $d_{min} \geq t \Rightarrow$ egy d_{min} kódtávolságú ... visszaállítható legyen tehát: $\forall c' \in C$ $c' \neq c$-re $d(v,c') > d(v,c)$
A hibajavítás esetén azt kérdezzük hogy t hiba esetén a v-ből c mindig visszaállítható legyen tehát ... $d(v,c') \geq d(c,c') - d(v,c) > d(v,c)$ / $+ d(v,c)$
A Hamming táv teljesíti a Δ egyenlőtlenséget, így $d(v,c') \geq d(c,c') - d(v,c)$ tehát $d(v,c') > d(v,c)$ úgy biztosítható ha $d(c,c') \geq d(v,c) > d(v,c)$
$d(c,c') > 2d(v,c)$
$\frac{d_{min}}{2} > d(v,c)$

geometriai szemléltetés: $\bullet c \quad d_{min} \quad \bullet c'$

Tehát egy ... esetén $\left\lfloor \frac{d_{min}-1}{2} \right\rfloor$ hiba javítható.
q a C halmaz dimenziója $C \in Q^n$

Tétel: Singleton korlát: Egy M kódszóból álló n hosszú s d_{min} kódtávolságú kódra $M \leq q^{n-d_{min}+1}$

Biz: Legyen k egy természetes szám melyre $q^{k-1} < M \leq q^{k}$. Mivel a k-1 hosszú különböző sorozatok száma q^{k-1} ezért $q^{k-1} < M$ miatt létezik két c, c' kódszó melyeknek első k-1 koordinátájuk megegyezik $\Rightarrow d(c,c') \leq n-k+1$, tehát $d_{min} \leq n-k+1$ azaz $M \leq q^{k} \leq q^{n-d_{min}+1}$ ... a kódoló k hosszú ... n hosszú ... kódunk (n,k) paraméterű ... a Singleton korlát alapján $d_{min} \leq n-k+1$
Jellegzetes esetben $M = q^{k}$ vagyis ... egyenlőség áll MDS (Maximum Distance Separable) kódnak nevezzük.

Def: Azon kódot melyre a Singleton korlátban egyenlőség áll MDS (...)
Def: Egy lineáris kód lineáris ha a C halmaza lineáris tér azaz $\forall c, c' \in C: c + c' \in C$
Def: $g_1, g_2, \dots g_k \in C$ vektorok lineárisan függetlenek ha $\lambda_i \in \{0,1\}$ mellett $\sum_{i=1}^{k} \lambda_i g_i = 0$ csak úgy állhat elő ha $\forall i$-re $\lambda_i = 0$
Def: $g_1, g_2 \dots g_k \in C$ vektorok a C lineáris tér egy bázisát alkotják ha lineárisan függetlenek, továbbá igaz hogy minden $c \in C$ kódszó előállítható $c = \sum_{i=1}^{k} u_i g_i$ alakban ahol $u_i \in \{0,1\}$ $\forall i$-re. Ez felírható mátrixalakban $c = uG$ ahol $u = (u_1, u_2, \dots u_k)$ G pedig a bázis vektorai által alkotott mátrix. A C kódot G generálja.
A $c = uG$ egyenlőség egy k dimenziós s n dimenziós vektort rendel össze. ... megfelelő k hosszú üzenetet
Def: Egy (n,k) paraméterű kód szisztematikus, ha minden kódszóra igaz, hogy annak első n-k szimbólumait elhagyva éppen a neki megfelelő k hosszú üzenetet kapjuk. Ez a üzenethez tartozó C kódszó szerkezete tehát:
Szisztematikus kód esetén $G = (I_k, B)$ ahol I_k a $k \times k$-s egységmátrix B pedig $k \times (n-k)$ méretű mátrix
$c = (u_1, u_2, \dots u_k, c_{k+1}, c_{k+2} \dots c_n)$
Def: Ha egy ... sorból s n oszlopból álló H mátrixra $Hc^T = 0 \Rightarrow$ H-t a C kód paritásellenőrző mátrixának nevezzük.
Tétel: Ha G^T s H ugyanazon lineáris kódhoz tartozik akkor $HG^T = 0$
$c \in C$ melyre $c = uG$. Ugyanakkor $Hc^T = 0$ tehát
Biz: Jelölje Q^k a k hosszú lineáris sorozatok halmazát $\Rightarrow \forall u \in Q^k$ -hoz $\exists c \in C$ melyre $c = uG$.
$Hc^T = H(uG)^T = HG^T u^T = 0$
A fenti tétel alapján könnyen előállíthatunk H-t. Legyen $H = (A, I_{n-k})$ alakú. Ekkor $HG^T = (A, I_{n-k})(I_k, B)^T = A + B^T = 0 \Rightarrow A = -B^T$
Def: Egy vektor súlya a koordinátái közt lévő nemnulla elemek száma. Jele $w(c)$
Def: Egy kód minimális súlyán a $w_{min} = \min_{c \neq 0,\ c \in C} w(c)$ számot értjük

Tétel: Ha C lineáris kód $\Rightarrow d_{min} = w_{min}$ $c'' \in C$ mert C lineáris
Biz: $d_{min} = \min_{c \neq c'} (c,c') = \min_{c \neq c'} (c - c') = \min_{c'' \neq 0} (c'') = w_{min}$
Def: Az $s = eH^T$ mennyiséget szindrómának nevezzük ... Ekkor: $Hv^T = H(c+e)^T = Hc^T + He^T = He^T$
Legyen az adott kódszó c s a vett szó v. Az $e = v - c$ vektort hibavektornak nevezzük. ... a $Hv^T = He^T$ szindrómát, ezalapján megbecsüljük e-t majd innen $c = v - e$ ... Valamilyen e hiba által generált halmaz
A dekódolás legegyszerűbb alakja a szindróma dekódolás: a v vett szóból kiszámítjuk a ... standard elrendezési táblázat.
A szindrómák ... történő kiszámítása ... ha a v vett szó szindrómája $s^{(i)}$ akkor $\hat{c} = v - e^{(i)}$
(mellékosztály) az $e + c$, $c \in C(n,k)$ vektorok halmaza $A\ w(e^{(i+1)}) \geq e^{(i)}$ $e^{(0)} = 0$ a szokásos sorrend.
Az $e^{(i)}$ mellékosztály vezetőket javítható hibamintáknak nevezzük, ugyanis ... ezt a módját táblázatos dekódolásnak nevezzük.
kódszó-a döntünk. A szindróma ...

$s^{(0)}$	$e^{(0)} = 0$	$c^{(1)}$	$\dots$	$c^{(2^k-1)}$
$s^{(1)}$	$e^{(1)}$	$c^{(1)} + e^{(1)}$	$\dots$	$c^{(2^k-1)} + e^{(1)}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
$s^{(2^{n-k}-1)}$	$e^{(2^{n-k}-1)}$	$c^{(1)} + e^{(2^{n-k}-1)}$	$\dots$	$c^{(2^k-1)} + e^{(2^{n-k}-1)}$

mellékosztály elemek

szindróma mellékosztály vezető

Tétel: C lineáris kód távolsága legalább d^* akkor s csak akkor ha a paritásellenőrző mátrixa tetszőlegesen választott d^*-1 oszlopa lineárisan független
A fenti tétel alapján 1 hibát javító lineáris kódot kapunk ha H tetszőleges két oszlopa különböző. Mivel a különböző nemzérus n-k hosszú bináris vektorok száma $2^{n-k}-1$ ezért: $n = 2^{n-k}-1$. Ez azt is jelenti, hogy a kapott kód perfekt tulajdonságú. Ennek alapján a bináris Hamming-kód paraméterei az alábbi számpárok
$d_{min} = 3$: $n = 3$, $k = 1$
$7 \quad 4$
$15 \quad 11$
$31 \quad 26$

... két műveletet melyeket összeadásnak s szorzásnak nevezünk s Q rendelkezik a következő tulajdonságokkal

Def: Egy Q halmazt testnek nevezünk ha értelmezve van tetszőleges két elem között ... két művelet
1) Q az összeadásra kommutatív csoport: a) $\forall \alpha, \beta \in Q$ esetén $\alpha + \beta \in Q$ összeadásra zárt
b) $\forall \alpha, \beta, \gamma \in Q$ esetén $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ asszociatív
c) $\exists\, 0$ elem melyre $\forall \alpha \in Q$-ra $0 + \alpha = \alpha$. 0-t nullelemnek nevezzük
d) $\forall \alpha \in Q$-ra $\exists\, \beta \in Q: \alpha + \beta = 0$. β-t α additív inverzének nevezzük s $-\alpha$-val jelöljük
e) $\forall \alpha, \beta \in Q$-ra $\alpha + \beta = \beta + \alpha$ kommutatív

2) $Q \setminus \{0\}$ a szorzásra kommutatív csoport: a) $\forall \alpha, \beta \in Q \setminus \{0\}$ esetén $\alpha\beta \in Q$ zártság
b) $\forall \alpha, \beta, \gamma \in Q \setminus \{0\}$ esetén $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ asszociatív
c) $\exists\, 1$ elem melyre $\forall \alpha \in Q \setminus \{0\}$-ra $1\alpha = \alpha$. 1-t egységelemnek nevezzük
d) $\forall \alpha \in Q \setminus \{0\}$-ra $\exists\, \beta \in Q \setminus \{0\}: \alpha\beta = 1$. β-t α multiplikatív inverzének nevezzük s α^{-1}-el jelöljük
e) $\forall \alpha, \beta \in Q \setminus \{0\}$-ra $\alpha\beta = \beta\alpha$ kommutatív

3) $\forall \alpha, \beta, \gamma \in Q$-ra $\alpha \cdot 0 = 0 \cdot \alpha = 0$ s $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ disztributivitás
Ezen értelmezhető az összeadás: $\alpha - \beta = \alpha + (-\beta)$ s az osztás: $\frac{\alpha}{\beta} = \alpha\beta^{-1}$ ha β nem 0. Egy q elemszámú Q testet véges testnek nevezzük s $GF(q)$-val jelöljük
Tétel: Egy $GF(q)$ esetén $q = p^m$ alakú ahol p prím
Lemma: $\forall\, 0 \neq a \in GF(q)$-ra $a^{q-1} = 1$
Lemma: $\forall\, 0 \neq a \in GF(q)$-ra létezik egy legkisebb m természetes szám amit az elem rendjének nevezünk, s $a^m = 1$ s az $a^1, a^2 \dots a^m$ elemek mind különbözők, illetve m osztója $q-1$-nek
Def: Egy $\alpha \in GF(q)$-t a $GF(q)$ primitív elemének nevezzük ha α rendje $q-1$.
Tétel: $\forall\, GF(q)$-ban $\exists$ primitív elem
Tétel: A $G = \{0,1 \dots p\}$ halmaz modulo p aritmetikai ... egy p prímszám esetén véges test azaz: $a + b = b + a \mod p$, $ab = ba \mod p$
... egy a $GF(q)$ elem α alapú logaritmusát az $\alpha^{\log_\alpha a} = a$ egyenlet egyértelmű megoldásával ahol $a \neq 0$. Ha a, b a $GF(q)$ ...
Ha ... a $GF(q)$ egy primitív eleme akkor bevezethetjük
0 elemei $\Rightarrow a \cdot b = \alpha^{\log_\alpha a}\, \alpha^{\log_\alpha b} = \alpha^{\log_\alpha a + \log_\alpha b}$
... vektorok reprezentációira s szorzásuk ... a polinom reprezentáció. ... polinom ha $a_i \in GF(q)$ $i = 0,1 \dots m$ $a_m \neq 0$, $x \in GF(q)$. A polinom fokszámát $\deg(a(x))$ jelöli
$GF(q)$ feletti ... $GF(q)$ feletti m-ed fokú polinom
Def: $a(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m$
Def: $a(x) = b(x)$ ha $a_i = b_i$ $\forall i$-re
Polinomok összeadása $c(x) = a(x) + b(x)$ tagonként történik: $c_i = a_i + b_i$ s $\deg(c(x)) = \max\{\deg(a(x)), \deg(b(x))\}$
Polinomok szorzása $c(x) = a(x)b(x)$ minden tagot minden taggal szorzunk majd az azonos fokú tagokat csoportosítjuk: $c_i = \sum_{j=0}^{\min\{i, \deg(a(x))\}} a_j b_{i-j}$. A polinom fokszáma $\deg(c(x)) = \deg(a(x)) + \deg(b(x))$
Tétel: Euklideszi osztás polinomokra. Adott $a(x)$ s $d(x) \neq 0$ esetén egyértelműen létezik olyan $q(x), r(x)$ hogy $a(x) = q(x)d(x) + r(x)$ s $\deg(r(x)) < \deg(d(x))$
Def: $r(x)$-et az $a(x)$ $d(x)$-re vonatkozó maradékának nevezzük. Jele $a(x) \equiv r(x) \mod d(x)$
Def: $a(x)$ osztja $d(x)$-et ha $a(x) \equiv 0 \mod d(x)$. Jele $d(x) \mid a(x)$
Def: $c \in GF(q)$ gyöke az $a(x)$ polinomnak ha $a(c) = 0$
Tétel: Ha c $a(x)$ polinom gyöke akkor az előáll $a(x) = b(x)(x - c)$ alakban
Tétel: Egy k fokú polinomnak max k gyöke lehet
A $GF(p^m)$ beli elemek legyenek a $0, 1, \dots p^m - 1$ számok melyeknek m hosszú vektorokat feleltetünk meg, ahol a koordináták $GF(p)$ beliek. Ezt akár a $GF(p^m)$ beli elemek m hosszú vektor által reprezentált ... koordinátánként $GF(p)$ beli összegként értjük. Szorzás során a két m hosszú vektor által ... definiáljuk. Ezt viszont ... kiválasztunk egy speciális irreducibilis polinom által ... maradékot, képezzük, ez a ... m hosszú ... az m-1-ed fokú ... $GF(p)$ feletti polinom szorzatára
polinomot összeszorozzuk. Az eredmény ... prímszám $GF(p)$-ben
polinom ugyanolyan szerepet játszik mint a ... irreducibilis polinomnak nevezzük, ha nem bontható fel két nála alacsonyabb fokú $GF(p)$ feletti polinom szorzatára
Def: A $GF(p)$ feletti nem nulladfokú $P(x)$ polinomot irreducibilis polinom. pl $GF(2)$-re 1. fokú irreducibilis polinomok: x, $x+1$, ekkor minden 2. fokú polinom -egy
Minden nagyobb ... tetszőleges fokszámú mindegyiket ... irreducibilis polinom
különbözik az x^2, $x(x+1)$ s $(x+1)^2$ polinomok ... $P(x)$ egy $GF(p)$ feletti m-edfokú irreducibilis polinom s $Q = \{0,1 \dots p^m - 1\}$. Egy $a \in Q$ s $b \in Q$-nak kölcsönösen egyértelműen
Tétel: Legyen p egy prím, m egy természetes szám, $P(x)$ egy $GF(p)$ feletti legfeljebb (m-1) fokú $a(x)$ s $b(x)$ polinomot. a+b definíció szerint az a $c \in Q$ melynek megfelelő $c(x)$ polinomra:
$c(x) = a(x) + b(x)$
$a \cdot b$ az a $d \in Q$ melynek megfelelő $d(x)$ polinomra
$d(x) = a(x)b(x) \mod P(x)$
Ezzel az aritmetikával Q egy $GF(p^m)$
Def: Egy C kód lineáris ha a C halmaz lineáris tér $GF(q)$ felett azaz $\forall c, c' \in C$-re $c + c' \in C$ illetve $\beta \in C$-re $\beta c \in C$
C lineáris kódhoz létezik egy k lineárisan független sorból s n oszlopból álló G mátrix melyre $c = uG$ ahol a k hosszú u üzenethez a c kódszó tartozik s G mátrixot a C kód generátormátrixának nevezzük
Tetszőleges ... s n oszlopból álló H mátrixot paritásmátrixnak nevezünk ha $Hc^T = 0$
C lineáris kódnak egyik sorból s n oszlopból álló paritásellenőrző mátrixa
Tétel: Minden C lineáris kódnak van paritásellenőrző mátrixa
Bináris esetben elég csak a hiba helyét ismerni, nembináris esetben viszont a hely mellett a hiba értékét is ismernünk kell. Ezért a H mátrix oszlopait úgy választjuk, hogy azok nem 0-k s mind különbözőek legyenek s az első nem 0 elem minden oszlopban 1 legyen. Ekkor ha egy hiba esetén ... adódó s értéke e_i akkor a szindróma $s = e_i a_i$ ahol a_i^T a H i. oszlopa. Tehát a hiba értéke ... éppen a szindróma első nem 0 értéke, míg $a_i = \frac{s}{e_i}$ amiből i visszafejthető
Ha H tartalmazza az összes fenti módon megengedett oszlopvektort akkor $n = \frac{q^{n-k}-1}{q-1}$ azaz $1 + n(q-1) = q^{n-k}$ másrészt Hamming korlát miatt $1 + n(q-1) \leq q^{n-k}$ tehát a maximális hosszúságú nembináris Hamming kód perfekt kód.
Tétel: Az $(n, n-2)$ paraméterű nembináris Hamming kód MDS-kód.
Def: Legyenek $\alpha_0, \alpha_1 \dots \alpha_{n-1}$ $GF(q)$ különböző elemei $(n \leq q)$ s $u = (u_0, u_1, \dots u_{k-1})$ $(u_i \in GF(q))$ a k hosszúságú üzenetszegmens ... az $u(x) = u_0 + u_1x + \dots + u_{k-1}x^{k-1}$ üzenetpolinomot rendeljük. Ekkor a Reed-Solomon kódnak az a üzenethez tartozó n hosszú c kódszavát a következő módon állítjuk elő: $c_0 = u(\alpha_0)$, $c_1 = u(\alpha_1)$, $\dots$ $c_{n-1} = u(\alpha_{n-1})$
Belátható hogy a Reed-Solomon kód lineáris s generátormátrixa:

$$G = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha_0 & \alpha_1 & \dots & \alpha_{n-1} \\ \vdots & & & \vdots \\ \alpha_0^{k-1} & \alpha_1^{k-1} & \dots & \alpha_{n-1}^{k-1} \end{pmatrix}$$

Tétel: Az (n,k) paraméterű Reed-Solomon kód távolsága $d_{min} = n-k+1$ vagyis a Reed-Solomon kód maximális távolságú.

Az (n,k) paraméterű Reed-Solomon kód esetén $n-k$ hibát tud jelezni $\lfloor \frac{n-k}{2} \rfloor$ egyszerű hibát tud javítani. Az n ismeretlenre vonatkozó $c=uG$ n darab egyenletből bármelyik $n-k$ egyenletet elhagyjuk ... egy egyértelműen megoldható egyenletrendszer marad

Konstr: Legyen α a $GF(q)$ egy nemnulla eleme, melynek rendje m, $m \geq n$ és legyen $\alpha_0=1, \alpha_1=\alpha, \dots \alpha_n=\alpha^{n-1}$. Ekkor a generátormátrix:

$$G=\begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & \alpha & \dots & \alpha^{n-1} \\ 1 & \alpha^2 & \dots & \alpha^{2(n-1)} \\ \vdots & & & \\ 1 & \alpha^{k-1} & \dots & \alpha^{(k-1)(n-1)} \end{pmatrix}$$

Def: Egy $c=(c_0, c_1 \dots c_{n-1})$ vektor ciklikus eltoltja az $Sc=(c_{n-1}, c_0, c_1 \dots c_{n-2})$ S-et a ciklikus eltolás operátorának nevezzük

Def: A kódot ciklikusnak nevezzük ha bármely kódszó ciklikus eltoltja is kódszó

Def: Rendeljünk polinomot az egyes kódszavakhoz a következő módon $C=(c_0, c_1 \dots c_{n-1}) \to c(x)=c_0+c_1x+\dots+c_{n-1}x^{n-1}$ ekkor $c(x)$-et kódszópolinomnak nevezzük. Ezek halmazát a $C(x)$-el jelöljük.

Lemma: Legyen $c(x)$ a c kódszó. Sc eltoltjának megfelelő kódszópolinom ekkor $c(x) \equiv x c(x) \mod (x^n-1)$

Tétel: Minden (n,k) paraméterű ciklikus C kódban a nem azonosan nulla kódszópolinomok között egyértelműen létezik egy minimális fokszámú $g(x)$ polinom melynek legmagasabb fokú tagjának együtthatója 1. $g(x)$ fokszáma $n-k$ és egy $c \in C(x)$ $g(x)|c(x)$ azaz $\exists u(x)$: $g(x)=c(x)u(x)$

Def: $g(x)$-et a kód generátor polinomjának nevezzük

Tétel: Minden ciklikus lineáris kód $g(x)$ generátorpolinomja $g(x)|x^n-1$, illetve ha egy $g(x)$ főpolinom $g(x)|x^n-1 \Rightarrow$ létezik hozzá ciklikus kód.

Def: Egy $g(x)$ generátorpolinomú lineáris ciklikus kód esetén a $h(x)=\frac{x^n-1}{g(x)}$ polinomot paritásellenőrző polinomnak nevezzük

Tétel: Egy lineáris ciklikus kódra $c(x)$ kódszó $\Leftrightarrow c(x)h(x)=0 \mod (x^n-1)$ és $\deg(c(x)) \leq n-1$

Egy lineáris ciklikus kódot lehet például a generátorpolinom s üzenetpolinom szorzásával generálni. Ez a módszer megfeleltethető olyan G generátormátrix segítségével is melynek legyen sorai ... ha G sorai $g(x)$ eltoltjai.

Tétel: Minden lineáris ciklikus kód generálható szisztematikusan

Biz: Vegyük észre, hogy $g_0 \neq 0$ mert ha 0 lenne akkor g-t leosztva ... x-el 1-el kisebb fokszámú kódszópolinomot kapnánk ami lehetetlen. Gauss elimináció ... szisztematikus generátormátrixot kapunk. Legyen $u(x)$ egy legfeljebb $(k-1)$-edfokú üzenetpolinom s $c(x)=u(x)x^{n-k}-[u(x)x^{n-k}] \mod g(x)$

A szisztematikus generálás egy praktikus módszere a következő: ...

akkor $c(x)$ kódszópolinom mivel $c(x)=0 \mod g(x)$ ... a korábbi tétel alapján

A Reed-Solomon kódok legfontosabb gyakorlati alkalmazási módja az előbbi kódszóhossz egyenlő az ott szereplő α elem rendjével. Ekkor a kód ciklikus s generátorpolinomja: $g(x)\prod_{i=1}^{n-k}(x-\alpha^i)$

Tétel: A Reed-Solomon kód esetén legyen ha kódszóhossz egyenlő ... $h(x)=\prod_{i=n-k+1}^{n}(x-\alpha^i)$ tehát a nem rövidített Reed-Solomon kód ciklikus.

speciális paritásellenőrző polinomja ... Egy hibát javító Reed-Solomon kód generátorpolinomja $g(x)=(x-\alpha)(x-\alpha^2)$. Tetszőleges $c(x)$ kódszó esetén $c(\alpha)=c(\alpha^2)=0$. A vett szó polinoma alakban $v(x)=c(x)+e(x)$

Az egy hiba javítás esetét célozzuk $e(x)=ex^i$ ahol e a hiba értéke s i a hiba helye. Adottak, hogy $S=(S_1,S_2)$ szindrómavektor komponenseit ekvivalensen kiszámíthatjuk a következő módon: $S_1=v(\alpha)$, $S_2=v(\alpha^2)$. Innen $S_1=v(\alpha)=c(\alpha)+e(\alpha)=0+e(\alpha)=e(\alpha)=e \cdot \alpha^i$

$S_2=v(\alpha^2)=c(\alpha^2)+e(\alpha^2)=0+e(\alpha^2)=e(\alpha^2)=e \cdot \alpha^{2i}$ Zh anyaga

Így $\frac{S_2}{S_1}=\alpha^i$ amiből az i hibahelyet kapjuk majd ebből meghatározzuk az $e=S_1\alpha^{-i}$ hibaértéket.

Jelöljön X egy diszkrét val. változót, mely az $x_1, x_2 \dots x_n$ véges halmazból veszi értékeit. A halmaz elemeit forrásábécének hívjuk, ugyanígy egy Y halmaz elemeit kódábécének. Y^* jelölje Y elemeiből álló véges sorozatok halmazát. Y^* elemeit kódszavaknak nevezzük. Az $f: X \to Y^*$ függvényt kódnak nevezzük, ha ennek értékkészlete különböző hosszú kódszavakból áll $\Rightarrow$ változó hosszúságú kódolással foglalkozunk.

Def: Az $f: X \to Y^*$ kód egyértelműen dekódolható, ha minden véges kódbetűsorozat legfeljebb egy üzenetnek kódolásaként állhat elő, azaz ha $u \in X^*$, $v \in X^*$, $u=u_1u_2 \dots u_k$, $v=v_1v_2 \dots v_l$, $u \neq v$, akkor $f(u_1)f(u_2) \dots f(u_k) \neq f(v_1)f(v_2) \dots f(v_l)$

Def: Az f kód prefix, ha a lehetséges kódszavak közül egyik sem folytatása a másiknak, vagyis bármely kódszó végéből önmagában szegmenst levágva nem kapunk egy másik kódszót, az ilyen kódszavak ... A kódfa gyökeréből induló egyértelmű utat jelölnek ki egy levélhez

A prefix kódok egyértelműen dekódolhatóak is mint egy adott kódszó ... dekódolható kódhoz létezik vele ekvivalens (azonos kódszóhosszú) prefix kód, tehát vehetjük s ez lesz a dekódolhatóság ... tulajdonságot az egyértelmű dekódolhatóság helyett.

A az egyszerűbb prefix ... értékű val. változó amit kódolni akarunk, vezessük be a $p: X \to [0,1]$ függvényt: $p(x)=P\{X=x\}$ $x \in X$. A $p(x)$ tehát x forrásbetűhöz annak

Legyen X egy valószínűségét rendeli.

Def: Az X val. változó entrópiáját $H(X)$-el $H(X)=E(-\log p(X))=-\sum_i^n p(x_i)\log p(x_i)$ összeggel definiáljuk

Def: Egy kód átlagos kódszóhosszán az $E|f(X)|=\sum_i^n p(x_i)|f(x_i)|$ várható értéket értjük

Tétel: Egy egyértelműen dekódolható $f: X \to Y^*$ kódra $E|f(X)| \geq \frac{H(X)}{\log s}$

Tétel: Tetszőleges ... val. változóra melynek ... X s Y halmazból veszi értékeit

Tétel: Legyenek X s Y val. változók a különböző értékeket ... pozitív valószínűséggel, akkor $0 \leq H(X) \leq \log n$ s a bal oldali egyenlőség csak akkor áll fenn ha az X val. változó 1 valószínűséggel konstans, a jobb oldalon pedig csak akkor, ha X egyenletes eloszlású.

1) ... X s Y diszkrét val. változókra $H(X,Y) \leq H(X)+H(Y)$ s az egyenlőség szükséges s elégséges feltétele az, hogy X s Y független legyen

2) X s Y diszkrét val. változó $g(x)$ függvényére $H(g(X)) \leq H(X)$ s itt az egyenlőség szükséges s elégséges feltétele az, hogy g invertálható legyen

3) Az X tetszőleges ... forrásra mint ... generátornak tekintünk akkor a ... nem függ a ... időtől

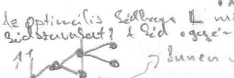
Adattömörítéskor ... Digital Communication ...

$P(X(k)=x_i)=P_i(k)=p_i$

A forrásnak vannak:

- szimbólumai: $x_1, x_2, \dots x_n \to X$
- valószínűség eloszlás: $p_1, p_2 \dots p_n \to p(X)$
- kódszavak: $z_1, z_2 \dots z_n \to z(X)$
- kódszóhossz: $l_1, l_2 \dots l_n \to l(X)$
- átlagos kódszóhossz, jele: $L=E(l(X))=\sum_x p(x)l(x)$

Az optimális kódhossz L minimális $\Rightarrow$ minimális az adattároló sebesség. Változó kódhosszal egyértelműen dekódolható?, hogyan rendeljük a szimbólumokhoz a kódszavakat? A kód egyértelműen dekódolható, ha prefixmentes $\to$ innen egy prefix fa következik.



$\to$ innen nem megérkező levelek száma: $2^{L-l(x)} \Rightarrow \sum_x 2^{L-l(x)} \leq 2^L$

$\sum_x 2^{-l(x)} \leq 1 \to$ Kraft egyenlőtlenség, prefix mentes kódot csak ilyen kódszavakkal lehet ...

2 eloszlás távolsága: $D(p\|q)=\sum_x p(x)\,\mathrm{ld}\frac{p(x)}{q(x)} \geq 0$ $\quad -D(p\|q) \leq 0$ $\quad -D(p\|q)=\sum_x p(x)\,\mathrm{ld}\frac{q(x)}{p(x)} \leq \sum_x p(x)\left(\frac{q(x)}{p(x)}-1\right)=\sum_x q(x)-\sum_x p(x)=1-1=\emptyset$

Információ: $I(x)=\mathrm{ld}\frac{1}{p(x)} \to$ attól függ, hogy mennyire "meglepő" az információ

Entrópia (átlagos információ): $H(X)=E(I(X))=\sum_x p(x)\,\mathrm{ld}\frac{1}{p(x)}$ $\quad 0 \leq H(X) \leq \mathrm{ld}N \to$ csak akkor éri el ha egyenletes a forrás eloszlása

$p(x), q(x)=\frac{1}{N}$ $\quad D(p\|q)=\sum_x p(x)\,\mathrm{ld}\frac{p(x)}{\frac{1}{N}}=\sum_x p(x)\,\mathrm{ld}(p(x)N)=\sum_x p(x)\{\mathrm{ld}\,p(x)+\mathrm{ld}N\}=\underbrace{\sum_x p(x)\,\mathrm{ld}\,p(x)}_{-H(X)}+\mathrm{ld}N\underbrace{\sum_x p(x)}_{1}=-H(X)+\mathrm{ld}N \geq 0$

$\mathrm{ld}N \geq H(X)$

Forráskódolási tétel: A tömörítés alsó határa az entrópia, $H(X) \leq L$

Biz: $p(x)$, $q(x)=\frac{2^{-l(x)}}{\sum_y 2^{-l(y)}}$, $0 \leq q(x) \leq 1$; $\sum_x q(x)=1$

$$0 \leq D(p\|q)=\sum_x p(x)\,\mathrm{ld}\frac{p(x)}{q(x)}=\sum_x p(x)\,\mathrm{ld}\left(\frac{p(x)\overbrace{\sum_y 2^{-l(y)}}^{\leq 1}}{2^{-l(x)}}\right) \leq \sum_x p(x)\,\mathrm{ld}(p(x)2^{l(x)})=\sum_x p(x)\{\mathrm{ld}\,p(x)+\mathrm{ld}\,2^{l(x)}\}=\sum_x p(x)\,\mathrm{ld}\,p(x)+\sum_x p(x)l(x)=-H(X)+L \Rightarrow H(X) \leq L$$

Shannon-Fano kód

$H(X)=\sum_x p(x)\,\mathrm{ld}\frac{1}{p(x)}$ } jó közelítés

$l(x)=\lceil \mathrm{ld}\frac{1}{p(x)} \rceil$ $\quad L=\sum p(x)\lceil \mathrm{ld}\frac{1}{p(x)} \rceil$

$\sum_x 2^{-l(x)}=\sum_x 2^{-\lceil \mathrm{ld}\frac{1}{p(x)} \rceil} \leq \sum_x 2^{-\mathrm{ld}\frac{1}{p(x)}}=\sum_x 2^{\mathrm{ld}\,p(x)}=\sum_x p(x)=1 \Rightarrow$ tehát kielégíti a Kraft egyenlőtlenség

$H(X) \leq L=\sum_x p(x)\lceil \mathrm{ld}\frac{1}{p(x)} \rceil \leq \sum_x p(x)(\mathrm{ld}\frac{1}{p(x)}+1)=\sum_x p(x)\,\mathrm{ld}\frac{1}{p(x)}+\sum_x p(x)=H(X)+1$

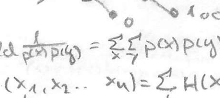
Algoritmus: Adott a forráseloszlás $\to$ innen kódszóhosszak $\to$ ezekre bináris fa

Optimális kódok, Huffman kód

Az optimálisság szükségességi kritériumai: 1) $p(x)>p(y) \Rightarrow l(x)<l(y)$

2) $p_1>p_2>\dots p_N \Rightarrow l_{N-1}=l_N$

Huffman kód példa: 0,4 0,2 0,15 0,15 0,1 $\to$ 0,4 0,2 0,15 0,25 $\to$ 0,4 0,35 0,25 $\to$ 0,4 0,6 $\to$ 1 $\Rightarrow$



Az egybeolvasztott elemek 1 lépéssel... számítunk

Blokk kód

$H(X,Y)=\sum_x\sum_y p(x,y)\,\mathrm{ld}\frac{1}{p(x,y)} \overset{x,y \text{ függetlenek}}{=} \sum_x\sum_y p(x)p(y)\cdot \mathrm{ld}\frac{1}{p(x)p(y)}=\sum_x\sum_y p(x)p(y)\{\mathrm{ld}\frac{1}{p(x)}+\mathrm{ld}\frac{1}{p(y)}\}=\left(\sum_y p(y)\right)\sum_x p(x)\,\mathrm{ld}\frac{1}{p(x)}+\left(\sum_x p(x)\right)\sum_y p(y)\,\mathrm{ld}\frac{1}{p(y)}=H(X)+H(Y)$

Kiterjesztés $\Rightarrow$ független: $X_1, X_2, \dots X_n \to H(X_1, X_2 \dots X_n)=\sum_i H(X_i)$ $\quad$ azonos eloszlás-ra $n \cdot H(X)$

Összevonunk M db szimbólumot 1 db szimbólummá

Y	$x(k)$	$x(k-1)$	...	$x(k-M+1)$	
y_1	x_1	x_1	...	x_1	p_1^M
y_2	x_1	x_1	...	x_2	$p_1^{M-1}p_2$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
y_n	x_n	x_n	...	x_n	p_n^M

$H(Y) \leq L^{(M)} \leq H(Y)+1$

$MH(X) \leq L^{(M)} \leq M H(X)+1$

$H(X) \leq \frac{L^{(M)}}{M} \leq H(X)+\frac{1}{M}$

$\frac{L^{(M)}}{M}=\lambda$

$H(X) \leq \lambda \leq H(X)+\frac{1}{M}$ $\quad$ ha $\varepsilon:=\frac{1}{M} \Rightarrow$ tetszőlegesen közel tudunk jutni az entrópiához

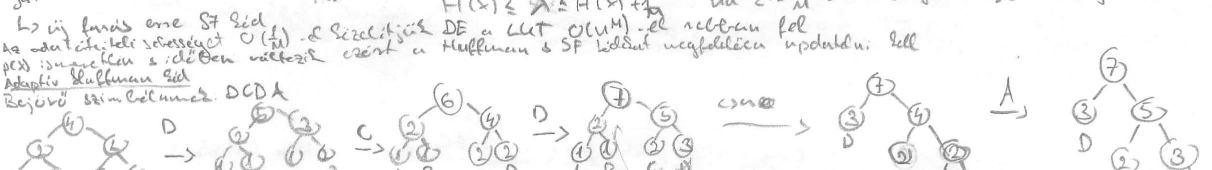
$\to$ új forrás erre SF kód

Az adatátviteli sebességet $O(\frac{1}{M})$-el közelítjük DE a LUT $O(N^M)$-el nebban fel

$p(x)$ ismeretlen s időben változik ezért a Huffman s SF kódot megfelelően updatelni kell

Adaptív Huffman kód

Bejövő szimbólumok: DCDA



LZ77 algoritmus: közelíti a Huffman kódot. A szövegre egy search s egy look up buffert illesztünk melyeknek hossza rendre n_s s n_c

Aszimptotikusan

$n_s=7$ $n_c=6$ cabracadabrarrarrad... $\quad$ Output: $0,0,f(d)$ (0 hosszú egyezés, 0. helyen, az új szimbólum d)

cabracadabrarrarrad... $\quad$ Output: $7,4,f(r)$

7, 4, 3, 5

cabracadabrarrarrad $\quad$ Output: $3,5,f(d)$

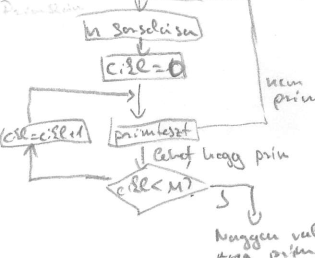
LZ78 algoritmus

A szöveg egymás után jövő részeiből szótárt alkotunk

0101011100110101101111 ... $\to$ (0,0,0); (0,0,1); (0,0,1); (0,1,1); (1,0,0); (1,0,1); (1,1,1);

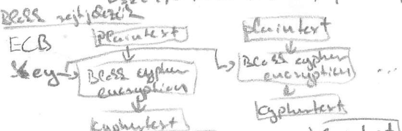
m: nyílt szöveg $m \in M$
c: rejtett szöveg $c \in C$
k: kulcs, $k \in K$
rejtjelező kódolás: $E_{k_1}(m) = c$ $\}$ $D_{k_2}(E_{k_1}(m)) = m$
dekódolás: $D_{k_2}(c) = m$
A kulcs lehet szimmetrikus és aszimmetrikus is
Cézár kód exists. Formalizálva a kódolás: $c = a * m + b$ mod 26
A támadás esetei: • Csak nyílt-rejtett párok alapján történő támadás
• Chosen plain-text: bármely nyílt szöveghez ismeri a rejtettet
• Chosen text: bármely nyílt szöveghez ismeri a rejtettet & fordítva
Kerckhoffs feltevés, hogy a támadó ismeri az algoritmust & chosen text támadást végez, ha egy kulcsot néhányszor a kulcsból ellen felfedte
Lineáris block rejtjelező
Ha $y = Ax + b$ -vel rendelt jelezzük ahol A egy n-es lineáris mátrix & x, y, b egy n hosszú oszlopvektor
Statisztikai analízist is alkalmazhatunk: a leggyakoribb betű az angol ábécében az e
One Time Pad
x bináris sorozat, k egy szimmetrikus bináris kulcs melyet pénzfeldobással állítunk elő, innen y a rejtett szöveg $y = x + k$ mod 2
Tökéletes rejtjelező → nyílt & zárt szöveget függetlennek $P(y = y_i | x = x_i) = P(y = y_j | x = x_j)$...
Shamir háromlépéses protokollja → A kódot kulcsként használja
Ilyenek kommutatív tulajdonságú rejtjelezés & Bellgate típusú támadó kell
Integritás védelem: az m db üzenetblokkot mod 2 összegezzük & így egy extra ellenőrző blokkot képezünk, majd ezt az egészet rejtjelezzük. Támadható-e ez? A támadó időközben módosítani próbálja az üzenetet
Jelszó azonosításhoz egyirányú leképezést használunk
Kulcsmenedzsment feladatai: • generálás
• tárolás
• szétosztás
• frissítés
• visszavonás

RSA kódolás
1) választunk 2 nagy prímszámot: p1, p2
2) $m = p1 \cdot p2$ $\Phi(m) = (p1-1)(p2-1)$
e: $1 \le e \le \Phi(m)$ $(\Phi(m), e) = 1$
3) $d = e^{-1}$ (mod $\Phi(m)$)
4) $k^P = (m, e)$ $k^S = (d, p1, p2)$
Kódolás: $y = x^e$ (mod m) $1 \le x \le m$
Dekódolás: $x = y^d$ (mod m) $1 \le y \le m$
Miért előnyös az $e = 2^k + 1$ alakú választás? → 1 db (moduláris) szorzás & k db négyzetre emelés
Példa: $e = 2^k + 1 \Rightarrow x^e = x((\dots(x)^2 x)^2 \dots)^2$
Az RSA-hoz nagy prímszámokat is kell keresnünk. M: a valószínűsége annak hogy egy m bit hosszú n szám prím? $(2^{m-1} < n < 2^m)$. Csebisev számelméleti tétele: $\pi(n)$ az n-nél kisebb prímek számának nagyságrendje $\pi(n) = \frac{n}{\ln n}$. Innen $P = \frac{\pi(2^m) - \pi(2^{m-1})}{2^m - 2^{m-1}} \approx \frac{2^{m-1} \frac{1}{(m-1)\ln 2}}{2^{m-1}} = \frac{1}{(m-1)\ln 2}$

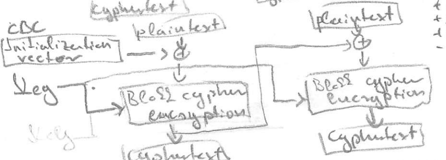


Kis Fermat tételen alapszik
Fermat tétel: Ha a c egész nem osztható a p prímmel $\Rightarrow c^p = 1$ mod p
Egy n összetett szám álprím b bázisra nézve ha $b^{n-1} = 1$ mod n → ez jön ki prímre is
Miller-Rabin teszt: Véletlenszerűen választunk
$\frac{1}{4}$ esélyen rossz $\{$ n prímgyanús ha $b \in R_n$,
n összetett ha $b \in R_n$ ahol R_n azon $b \in Z_n$ számok halmaza amelyekre vagy $b^u = 1$ mod n,
$n - 1 = 2^v u$, u páratlan vagy $\exists j$ $0 \le j < v$ melyre $b^{2^j u} = -1$ mod n

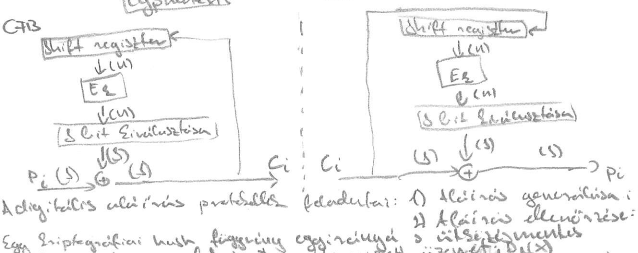
Ha p1 & p2 nincs túl messze egymástól akkor a faktorizáció hatékonyan végrehajtható
Fermat faktorizáció:
$n = ab$ $a, b > 0$
$t = \frac{a+b}{2}$ $s = \frac{a-b}{2}$ $\Rightarrow a = t + s$, $b = t - s$ $\Rightarrow n = t^2 - s^2 = (t+s)(t-s)$
Ha a-b kicsi ⇒ s is kicsi, (a & b közel vannak egymáshoz) $\Rightarrow t \approx n^{\frac{1}{2}}$
Próbálgatás t-re: $t_1 = int(n^{\frac{1}{2}}) + 1$, $t_2 = int(n^{\frac{1}{2}}) + 2$...
Ellenőrzés: $t_i^2 - n$ négyzetszám?
Ha igen $\Rightarrow t_i^2 - n = s^2 \to t$, $s \to a, b$
El-Gamal rejtjelező
G multiplikatív csoport g generátor elemmel
Kulcspár: $pk = X$, $sk = x$ ahol $X = g^x$, x véletlen elem $S = \{1, 2, \dots |G|\}$ halmazból
Rejtjelezés: $m \in S$
$E_{pk}(m) = (Y, b)$ ahol $Y = g^y$, y véletlen elem S halmazból
$b = Km$, $K = X^y$
Dekódolás: (Y, b) rejtett szöveg
$D_{sk}(Y, b) = m$, $K = Y^x$, $m = \frac{b}{K}$ $(X^y = Y^x)$
Blokk rejtjelező



A nyílt szöveg mintáit nem rejti megfelelően
A blokk rejtjelező kimenete nem randomizált
Szótár alapú támadás lehetséges
A rejtjeles blokkok felcserélhetők, törölhetők, helyettesíthetők



+ A nyílt szöveg mintáit rejti az előző blokkal XORolás
+ Azonos szövegeket különböző IV-el rejtjelezve különböző szövegeket kapunk
+ Korlátozott mértékben detektálni lehet a blokkok felcserélését, törlését helyettesítését
- Kiinduló & szótár támadások lehetségesek
Dekódolni ennek tükrével lehet



+ A nyílt szöveg mintáit rejti
+ A blokkrejtjelező kimenete véletlen
+ A blokkrejtjelező különböző rejtett szöveg lehetséges
+ Azonos nyílt szövegre korlátozott mértékben detektálható a rejtjelezésbeli módosítás
+ Korlátozott mértékben detektálható
- Utolsó karakter bitjei manipulálhatók

A digitális aláírás protokoll feladatai: 1) Aláírás generálása: $A \to B$ $X, D_A(X)$
2) Aláírás ellenőrzése: B: $X = E_A(D_A(X))$
Egy kriptográfiai hash függvény egyirányú & ütközésmentes
1 blokk méretű üzenetek: • Formázott üzenet: $D_A(X)$
• Nem formázott üzenet: $X, D_A(X)$
• Hashelve: $[X, D(H(X))]$
Damgard Merkle padding: Ha van f kompressziós fgv ütközés ellenálló tulajdonsággal. Ha az $X = [X_1, X_2 \dots X_{n-1}]$ üzenetet kiegészítjük egy X_n blokkal amely tartalmazza az X üzenet bithosszát, az f fgv-re épülő iteratív hash fgv CRHF lesz (ütközés ellenálló)