Qüestió - V. 10, n.º 2 (juny 1986) pp. 93-97

CONSTRUCCIÓ DE CODIS CORRECTORS A PARTIR DE CONFIGURACIONS I GRAFS

J. M. BASART I MUÑOZ
UNIVERSITAT AUTÒNOMA DE BARCELONA

Es tracta de presentar algunes de les relacions de caire bàsicament combinatòric que es poden trobar entre dos tipus importants d'estructures d'incidència, com són els grafs simètrics i les t-configuracions, i una estructura en principi algebraica com es la dels codis detectors i correctors d'errors.

L'objectiu final és, basant-se en el coneixement d'algunes de les propietats fonamentals de les t-configuracions i dels grafs, determinar uns tipus de transformacions que ens permetin caracteritzar -a partir d'una classe particular d'aquestes estructures- famílies de codis - correctors. En el cas de les t-configuracions estem especialment interessats en les quasi-simètriques, i en el cas dels grafs en els fortament regulars.

Keywords: ERROR-CORRECTING-CODES, t-DESIGNS, STRONGLY-REGULAR-GRAPHS, COMBINATORICS.

1. INTRODUCCIÓ.

El problema de transmetre una certa informació, tenint l'emissor un cert marge de seguretat pel que fa a la integritat del missatge que rebrà el receptor no és, ni de bon -- tros, un problema nou. De fet des dels orígens de la Humanitat, totes les cultures i civilitzacions s'han preocupat d'obtenir mètodes més o menys sofisticats per a realitzar aquesta - tasca.

D'ençà però que la Informàtica i les Telecomunicacions s'han posat a treballar conjunta--- ment -donant origen a l'anomenada Telemàtica- per tal de donar resposta a les necessitats - actuals de tractament d'informació, el problema de la transmissió fiable de la informació ha passat a representar un paper fonamental. Per raons òbvies de seguretat i rendiment, la implantació de codis detectors i correctors - d'errors en els sistemes de comunicacions --- -públics o privats- és avui dia més que una opció una necessitat.

Els codis ens permeten a partir d'un nivell de fiabilitat prefixat, establir com han de ser els missatges associats a la informació que volem transmetre per tal d'obtenir el resultat desitjat. Naturalment, augmentar la seguretat tindrà un cost, i aquest vindrà donat pel fet que o bé els missatges hauran de portar més dígits -bits, en el cas de codis dinaris- de control, o bé el nombre de possibles missatges diferents a transmetre serà menor.

En aquest treball ens referirem sempre a codis-bloc binaris, que són aquells codis en els quals tots els missatges usen sempre - el mateix nombre de dígits del conjunt {0,1}.

2. DEFINICIONS I PROPIETATS FONAMENTALS.

2.1 Els codis lineals.

Sigui F_2 el cos de Galois d'ordre 2, i considerem l'espai vectorial F_2^n de dimensió n de tots els n-vectors sobre F_2. Aleshores un codi C de longitud n serà qualsevol subconjunt de F_2^n. Si el subconjunt C té estructura de subspai vectorial sobre F_2 direm que tenim un codi lineal C(n,k), on n és la longitud - de les paraules-codi (vectors) i k és la dimensió del subespai.

G(kxn) és una matriu generadora del codi ---

- Josep Mª. Basart i Muñoz - Universitat Autònoma de Barcelona - Facultat de Ciències - Departament d'Informàtica Bellaterra - Barcelona

- Article rebut el maig de 1986.

C(n,k) si està formada per k vectors linealment independents de longitud n del codi C. Si $C^{\perp}$ representa l'espai ortogonal a C aleshores $C^{\perp}$ és també subespai vectorial de F_2^n, de dimensió n-k, i per tant és un codi lineal que tindrà H((n-k)xn) com a matriu generadora. H s'anomena la matriu de control del codi C(n,k). Direm que un vector $\underline{a}$ pertany al codi C si i només si $\underline{a}H^t=\underline{0}$.

El pes (w) d'un vector $\underline{a}$ de F_2^n ve donat pel nombre de coordenades diferents de 0 que conté.

La distancia (d) entre dos vectors $\underline{a}$ i $\underline{b}$ de F_2^n ve donada pel nombre de coordenades en les quals no coincideixen. La capacitat d'un codi per a corregir els possibles errors ---canvis de 0 a 1, o d'1 a 0- en un vector $\underline{a}$ enviat a través d'un canal amb soroll, depèn de la distància d que hi ha entre les seves paraules-codi. De fet és conegut que si es verifica $d\geq 2t+1$ aleshores el codi C pot corregir fins a t errors produïts en un mateix vector.

2.2. Les t-configuracions.

Sigui X un conjunt de v elemnts anomenats -- punts. Una t-configuració és una col.lecció de k-subconjunts anomenats blocs, de forma que cada t-subconjunt es troba en λ blocs. Ho podem representar dient que tenim una configuració t-(v,k,λ).

Teorema 1. /2/ Si λ_i representa el nombre de blocs que contenen un i-subconjunt d'elements de X amb $0\leq i\leq t$ aleshores

$$\lambda_i = \frac{\lambda_t \binom{v-i}{t-i}}{\binom{k-i}{t-i}}.$$

Anomanarem b a λ_0, el nombre total de blocs, i r a λ_1, el nombre de blocs que contenen un element donat. D'aquesta manera, si prenem i=0 i t=1 obtenim

$$bk = vr \qquad (1)$$

El cas de les 2-configuracions ens interessa especialment, així prenent t=2 en el Teorema 1 obtenim

$$bk(k-1) = \lambda v(v-1)$$

que usant (1) podem reescriure com

$$r(k-1) = \lambda(v-1) \qquad (2)$$

La matriu d'incidencia $N=(n_{ij})$ d'una configuració amb v punts $p_1,p_2,\ldots,p_v$ i b blocs $B_1,B_2,\ldots,B_b$ esta formada per b files i v columnes on:

$$n_{ij} = \begin{cases} 1 \text{ si } P_j \in B_i \\ 0 \text{ en cas contrari.} \end{cases}$$

2.3 Els grafs fortament regulars.

Un graf està format per un conjunt finit de vèrtexs i un conjunt finit d'arcs, amb una relació d'incidència entre vèrtexs i arcs, de manera que: (i) qualsevol arc incideix - en dos vèrtexs, i (ii) dos vèrtexs qualsevol comparteixen com màxim un arc.

Un graf s'anomena complet si cada parella de vèrtexs es troba unida per un arc, i s'anomena nul si no posseeix cap arc. El complementari d'un graf G és el graf $\bar{G}$ en el qual el conjunt d'arcs és el complementari del conjunt d'arcs de G.

Per a tot vèrtex v de G, la valència de v és el nombre d'arcs que contenen v. G s'anomena regular si tots els seus vèrtexs tenen la mateixa valència, i aquesta serà la valència del graf.

Un graf fortament regular és un graf regular, -ni complet ni nul- en el qual el nombre de vèrtexs adjacents a dos donats v_1 i v_2 depèn només de si v_1 i v_2 són o no són adjacents - entre ells. Els seus parametres són:

n, el nombre de vèrtexs del graf,
a, la valència del graf,
c, el nombre de vèrtexs adjacents a tota parella de vèrtexs adjacents, i
d, el nombre de vèrtexs adjacents a tota parella de vèrtexs no adjacents.

Donat un graf G definim la seva matriu d'adjacència (1) A(G) quadrada d'ordre n com

$$a_{ij} = \begin{cases} 1 \text{ si } v_i \text{ i } v_j \text{ són adjacents} \\ 0 \text{ en cas contrari} \end{cases}$$

on v_i i v_j són vèrtexs de G.

3. CONFIGURACIONS I GRAFS.

Una 2-configuració s'anomena quasi-simètrica si el cardinal de la intersecció de dos blocs qualssevol pot prendre només dos valors. Siguin x i y aquests valors.

En les 2-configuracions quasi-simètriques amb (x,y) = (0,1) s'anomena línies als seus blocs. Així podem associar a aquestes configuracions un graf-línia que tingui les línies com a vèrtexs, on dos vertexs seran adjacents sempre que la seva intersecció no sigui el conjunt buit. Aleshores:

Teorema 2. El graf-línia d'una configuració 2-(v,k,1) amb b>v és fortament regular amb paràmetres

$$(i)\ n=b=\frac{v(v-1)}{k(k-1)}$$

$$(ii)\ a=k(r-1)=\frac{k(v-k)}{k-1}$$

(iii) $c=(r-2)+(k-1)^2$ i (iv) $d=k^2$.

Demostració:

(i) Com que sabem que n=b per construcció, només cal prendre $\lambda=1$ en l'equació (2).

(ii) Cada element e dels k elements que té un vèrtex V estara relacionat amb tants vèrtexs com blocs continguin e -és a dir r- menys un, donat que el mateix vèrtex conté e.

(iii) Veure /2/.

(iv) Dos vèrtexs no adjacents X i Y estaran relacionats amb un altre vertex Z si aquest conté un element de X i un element de Y. El total de parelles (x,y) amb $x\varepsilon X$ i $y\varepsilon Y$ és k^2 donat que el nombre d'elements de cada bloc és k. ##

Podem ara generalitzar el graf-línia de la següent forma. Si Z és una configuració quasi-simètrica amb números d'intersecció x i y (x<y) el seu graf-bloc associat està format pels blocs de Z com a vèrtexs, on dos vèrtexs són adjacents si i només si la seva intersecció té cardinal y.

Teorema 3. El graf-bloc d'una 2-configuració quasi-simètrica és fortament regular.

Anem a obtenir ara un resultat que ens permetrà fer el camí invers. És a dir, obtenir una 2-configuració a partir d'una certa classe de graf fortament regular.

Teorema 4. Sigui G un graf fortament regular amb c=0 i 1<d<a. Sigui p un vèrtex qualsevol i considerem l'estructura d'incidència D(G,p) que té el conjunt de punts G(p) i el conjunt de blocs $\bar{G}(\bar{p})$, on un punt i un bloc seran incidents si i només si són adjacents a G. Aleshores D(G,p) és una 2-configuració on

$$v=a,\quad k=d,\quad \lambda=k-1,\quad r=v-1\quad i\quad b=\frac{v(v-1)}{k}.$$

que pot contenir blocs repetits.

Demostració:

Es pot veure fàcilment que qualsevol bloc incideix en d punts donat que cap bloc no és adjacent a p. Considerem ara dos punts qualsevol de D(G,p). Seran adjacents a p, i per tant donat que c=0 no seran adjacents a cap altre dintre del bloc on es troben. A més, d-1 vèrtexs més seran també adjacents a aquests dos i es trobaran situats a $\bar{G}(\bar{p})$.

Així hem vist que k=d i que $\lambda=d-1$. Naturalment v=a, i els altres paràmetres es poden obtenir directament a partir d'aquests. ##

4. CONFIGURACIONS I CODIS.

Un sistema de Steiner és una configuració amb $\lambda=1$, i es representa usant la notació S(t,k,v).

Teorema 5. Les files de la matriu d'incidència d'un S(t,k,v) formen un codi no lineal de paràmetres n=v, i $d\geq 2(k-t+1)$ que contindrà b paraules-codi.

Demostració:

La longitud del codi i el nombre de vectors que conté s'obtenen directament. Per a justificar el valor de d podem pensar que dos blocs qualssevol no poden tenir més de t-1 punts en comú -en cas contrari hi hauria contradicció amb la definició de sistema de --

Steiner-, per tant la distància entre dos blocs és, pel cap baix, $2(k-(t-1))=2(k-t+1)$.

Un codi binari C es diu balancejat si cada paraula-codi té el mateix pes r, on $r<m$ essent m la dimensió de C. D'aquesta manera la matriu d'incidència d'una t-configuració es pot prendre com un codi C balancejat, -- tant si considerem les files com si considerem les columnes com a paraules-codi. Si de la matriu en prenem les columnes tindrem a més l'avantatge d'obtenir un codi equidis-- tant, en el qual totes les paraules-codi es troben a la mateixa distància d.

Per a calcular aquesta d notem que donats -- dos vectors $\underline{a}$ i $\underline{b}$ de F_2^n es verifica que --- $dist(\underline{a},\underline{b})=w(\underline{a})+w(\underline{b})-2(\underline{a}.\underline{b})$ on "." representa el producte escalar.

Si $\underline{a}$ i $\underline{b}$ són vectors corresponents a dos --- punts diferents de la t-configuració, aleshores $w(\underline{a})=w(\underline{b})=r$ i $\underline{a}.\underline{b}=\lambda$ de manera que $dist(\underline{a},\underline{b})=2(r-\lambda)$.

5. GRAFS I CODIS.

En tot aquest apartat considerarem que estem parlant sempre d'un graf G connex -sense -- grups de vèrtexs aïllats que té n vertexs i m arcs.

Donat G podem definir la seva matriu d'incidència M(nxm) com

$$m(i,j)=\begin{cases}1 \text{ si l'arc j conté el vèrtex i}\\ 0 \text{ en cas contrari.}\end{cases}$$

Un arbre de G és un subgraf de n-1 arcs i n vèrtexs que no conté cap camí (seqüència -- d'arcs) tancat. Fixat un arbre, una corda de G és qualsevol arc de G que no forma part de l'arbre.

Un circuit és un camí tancat que usa tres o més arcs diferents. Un tall és un conjunt d'arcs tal que si s'elimina el graf resul-- tant queda disconnex.

Considerem ara que tenim fixat un arbre --- qualssevol T de G, aleshores:

a) els m-(n-1) circuits que s'obtenen afe--- gint una corda a T s'anomenen circuits fonamentals,

(b) els n-1 talls que contenen un únic arc de T s'anomenen talls fonamentals,

(c) els circuits fonamentals es poden agrupar per files en una matriu de circuits fonamentals C(m-(n-1)xm) on

$$c(i,j)=\begin{cases}1 \text{ si l'arc j}\in\text{circuit generat per la corda i}\\ 0 \text{ en cas contrari,}\end{cases}$$

(d) els talls fonamentals es poden agrupar per files en una matriu de talls fona-- mentals T((n-1)xm) on

$$t(i,j)=\begin{cases}1 \text{ si l'arc j}\in\text{ tall generat per l'arc i}\\ 0 \text{ en cas contrari,}\end{cases}$$

(e) situant en la primera part $-M_{11}-$ de la matriu d'incidencia $M=(M_{11}M_{12})$ les columnes corresponents a les cordes de -- l'arbre, podem obtenir /4/ les matrius fonamentals de la forma següent

$$C=(I \underbrace{M_{11}^t(M_{12}^t)^{-1}}_{C_f}) \qquad i \qquad T=(C_f\ I)$$

on I representa la matriu identitat de l'ordre apropiat,

(f) tots els circuits i talls que es poden formar a G es poden obtenir com a suma mòdul 2 coordenada a coordenada dels cir cuits i talls que hem anomenat fonamentals.

En aquestes condicions ja podem establir el resultat fonamental d'aquest apartat.

Teorema 6. C i T poden ser considerades com les matrius generadora i de control respectivament d'un codi lineal C de longitud m, dimensió m-(n-1) i distancia mínima 3.

Demostració:

Com que hem col.locat, tant a C com a T, la matriu identitat, tenim assegurat que les - seves files seran linealment independents. N'hi ha prou doncs amb veure que $CT^t=0$, i això és cert donat que treballem mòdul 2 i cada circuit afectat per un tall té un nombre parell d'arestes en comú amb el tall. A més cada circuit de C conté pel cap baix 3 arcs.##

Finalment podem dir que si el graf G usat és fortament regular, podrem obtenir alguna informació més sobre el codi C i sobre el

seu $C^{\perp}$. Així per exemple, $C^{\perp}$ tindra $d_{min}=a$, donat que cada tall prendra pel cap baix -- tants arcs com tingui un dels vèrtexs.

6. CONCLUSIONS.

En aquest treball hem fet un recull d'alguns dels lligams que podem establir entre els -- grafs regulars, les configuracions i els codis correctors. Aquests lligams s'han anat - trobant de forma independent en diferents -- camps de la Combinatòria, la Informàtica Teòrica, la Teoria de Grafs i la Teoria de la - Codificació sense donar mai una visió global del tema. Aixo és degut possiblement a la seva complexitat intrínseca, que dificulta --- l'obtenció de resultats generals -la majoria dels coneguts es limiten a valors més o menys particulars dels parametres d'aquests elements.

L'estructuració d'aquests lligams ens ha -- permès obtenir una visió més global dels --- tres elements implicats, de manera que la seva anàlisi en la major part dels casos es redueix gairebé a un problema de distribució i recompte, el qual no vol pas dir, evidentment, que la seva resolució sigui immediata.

Les relacions considerades han estat les seguents:

(i) obtenció de grafs fortament regulars a partir d'un cas particular de 2-configuracions -les quasi-simètriques. Obtenció de 2-configuracions a partir de grafs fortament regulars amb c=0 i 1<d<a,

(ii) pas des del sistema de Steiner al codi a través de la matriu d'incidència de la configuració. El pas és immediat pero el codi obtingut no és lineal. En general, a partir de la matriu d'incidència d'una t-configuració obtenim -- codis balancejats i equidistants, i

(iii) obtenció de les matrius generadora i de control d'un codi lineal si organitzem adequadament les columnes de la matriu d'incidencia d'un graf connex. Les perspectives resulten particularment interessants si G és fortament regular.

7. BIBLIOGRAFIA.

/1/ BLAKE, I.F., MULLIN, R.C.: "The mathematical theory of coding". Academic Press, 1975.

/2/ CAMERON, P.J., VAN LINT, J.H.: "Graph theory, coding theory and block designs" Cambridge University Press, 1975.

/3/ MACWILLIAMS, F.J. SLOANE, N.J.A.: "The theory of error-correcting codes". North-Holland Mathematical Library, 1977.

/4/ MAYEDA, W.: "Graph theory", John Wiley & Sons, 1972.

8. APÈNDIX.

(1) Si A és la matriu d'adjacencia d'un graf G fortament regular de parametres (n,a,c,d) aleshores pot ser caracteritzada de la forma següent.

L'element (v_1,v_2) de A^2 es correspon amb el nombre de vertexs que són adjacents a v_1 i v_2. Aquest nombre és a, c o d segons v_1 i v_2 siguin respectivament iguals, adjacents o no adjacents. Així doncs podem escriure

$$(i)\quad A^2=aI+cA+d(J-I-A)$$

$$(ii)\quad AJ=JA=aJ$$

on J és la matriu "tot-uns" i I la identitat.

D'aquesta manera un g.f.r. pot ser també definit com un graf en el qual la seva matriu d'adjacencia verifica (i) i (ii).