Universitatea Politehnica Bucuresti

Facultatea de Electronica Telecomunicatii si Tehnologia Informatiei

 **Codurile LDPC**

 (Low-Density Parity-Check Codes)

 -Retele de Calculatoare-

Broasca Valentina-Giorgiana

441A - CTI

**Cuprins**

1. Introducere

 1.1 Generalitati

 1.2 Tipuri de coduri LDPC

2. Codarea codurilor LDPC

 2.1 Coduri Structurate

 2.2 Structuri de Cod Ciclic

3. Decodarea codurilor LDPC

 3.1 Decodarea hard decision

 3.2 Decodarea soft decision

4. Concluzii si aplicatii ale codurilor LDPC

5. Bibliografie

**1.Introducere**

* **Generalitati**

 In acest proiect voi face o scurta prezentare a codurilor LDPC, coduri corectoare de erori in sistemele moderne de transmisiuni numerice.

In cea mai mare parte a sistemelor comerciale de transmisiuni sunt utilizate codurile LDPC, codurile Turbo, codurile convolutionale si codurile Reed-Solomon.

 Toate codurile corectoare de erori se bazeaza pe acelasi principiu: informatiei ii este adaugata redundanta, pentru a incerca corectarea eventualelor erori ce pot aparea in procesul de stocare sau transmisie. In functie de metoda prin care se adauga redundanta mesajului, codurile corectoare de erori se impart in doua clase: coduri bloc si coduri convolutionale, ambele avand aplicatii practice. Cele mai bune coduri corectoare de erori cunoscute astazi sunt blocurile cod. Acestea sunt coduri neregulate cu densitate mica si control al paritatii.

 Codurile LDPC sunt apropiate de limita codurilor corectoare de erori a lui Shannon. Recent codurile LDPC au fost adoptate in standarde precum IEEE 802.16e wireless MAN1, IEEE 802.11n wireless LAN si transmisi video digitale DVB-S2. Modelarea codurilor LDPC este mult mai flexibila, astfel pot fi modelate pentru a obtine codari si decodari eficiente.Interesul in cazul codurilor LDPC a aparut datorita potentialului de a obtine un randament (debit) foarte bun (datorita paralelismului intrinsec al algoritmului de decodare) totodata mentinand si o buna performanta de corectare a erorilor si o complexitate redusa a decodarii.

 Codurile LDPC au fost introduse in 1963 de Robert G. Gallager si reprezinta o familie de coduri bloc liniare care se obtin pornind de la o matrice rara de control al paritatii H, ce contine un numar mic de elemente nenule si un numar mare de elemente de 0.

 Codurile LDPC pot fi privite ca si coduri bloc liniare, descrise de matricea de control a paritatii H, ale carei dimensiuni sunt M × N. Fiecare linie a matricii H defineste o ecuatie de control a paritatii care trebuie sa fie satisfacuta de fiecare cuvant de cod v. Rezolvand sistemul format de aceste ecuatii se determina bitii de control specifici fiecarui cuvant de cod.

 Matricile de control a paritatii care determina codurile LDPC au urmatoarele proprietatii specifice:

• dimensiunile M si N ale matricii H sunt mult mai mari fata de dimensiunile unei matrici de control specifica unui cod Hamming. S-au impus aceste dimensiuni astfel incat, prin codare, sa ne apropiem cat mai mult de limita teoretica de codare impusa de teorema lui Shannon, in care se presupune un cod corector cu lungimea infinita a cuvantului de cod

• matricile H sunt definite, prin constructie, intr-o forma nesistematica si contin un numar de elemente de 1 (pentru codurile binare) mult mai mic decat numarul de elemente de 0;

• matricea de control a paritatii ce defineste un cod LDPC, H(j, k) are exact j elemente de 1 pe fiecare coloana si exact k elemente de 1 pe fiecare linie. Acestea inseamna ca fiecare bit al cuvantului de cod intra in j ecuatii de control, iar fiecare ecuatie de control contine k biti.

 Putem descrie matricea H prin intermediul unui graf bipartit cu doua tipuri de noduri:

• N noduri de simbol care corespund fiecarui bit din cuvantul de cod si

• M noduri de control care corespund ecuatiilor de control de paritate, noduri reprezentate de liniile matricii.

 Codurile definite prin astfel de grafuri pot fi decodate cu algoritmul Suma-Produs (Sum Product Algorithm – SPA sau Message-Passing –MP). Decodarea cu acest algoritm asigura performantele unui decodor bazat pe principiul ML (Maximum-Likelihood) numai daca grafurile sunt fara cicli cu lungime mica, sau altfel spus, sunt de tipul „4-cycle free”. Aceasta cerinta este echivalenta cu conditia ca matricea H sa nu aiba doua linii (oricare doua linii) in care sa existe elemente de 1 suprapuse pe mai mult de o singura pozitie. Pornind de la aceasta conditie, limitele ce se impun dimensiunilor M si N ale matricii H sunt interconditionate de relatia:

 N <= M(M-1)/ [j(j-1)]

Exemplu de matrice de paritate:

 x1 x2 x3 x4 x5 x6 x7

n(1-R) 1 1 1 0 1 0 0

 1 1 0 1 0 1 0

 1 0 1 1 0 0 1

x5=x1+x2+x3

x6=x1+x2+x4

x7=x1+x3+x4

* **Tipuri de coduri LDPC**

 Putem defini mai multe clase de coduri LDPC in functie de modul de generare a matricii de control H. Exemple:

-construite aleator

-pe baza unor constructii combinatoriale

-pe baza unor constructii geometrice

-construite pe baza unor structuri regulate (‘’array.based’’)

-construite pe baza grafurilor

-cuasi-ciclice

**2. Codarea codurilor LDPC**

 Codarea LDPC este realizata astfel: sa alegem de exemplu cateva noduri variabile unde sa plasam biti mesajului.In cel de-al doilea pas sa calculam valorile lipsa ale celorlalte noduri. O solutie evidenta pentru asta il reprezinta rezolvarea ecuatiei de paritate. Aceasta contine operatii ce implica intreaga matrice de control al paritati si astfel complexitatea ar creste exponential in concordanta cu lungimea blocului.In practica cu toate acestea sunt folosite metode eficiente care sa asigure un timp mult mai mic.

 Codurile LDPC fac parte din categoria codurilor bloc liniare. Pentru a intelege mai bine algoritmiice vor urma, vom definii cateva proprietati ale codurilor.

-Sursa informatiei: impartim o secventa cu sursa primara intr-un sir de vectori de lungime k inainte de codificare. Un astefl de vector il notam cu ‘u’.

-Codarea: se face pe baza lungimii unei diagrame aplicand u pe un sir de vectori cifrat cu x. Un cod (n,k) are lungimea de cod n.

-Cod sistematic: un cod in care informatia vector apare ca o parte a cuvantului de cod sistematic.

-Proprietatea de liniaritate: codul C este un cod liniar binar daca si numai daca C formeaza un subspatiu vectorial peste+. Deci, suma oricaror doua cuvinte de cod trebuie sa fie in sine un cuvant de cod.

-Dimensiunea codului: este dimensiunea spatiului vectorial care ii corespunde.

-Cuvant de cod diferit de zero: o consecinta directa a proprietatii de liniaritate este aceea ca numele de cod total=0, x=0, este un membru al oricarui cod liniar.

-Coeficientul codului: r=k/n, indica proportia de informatie transferata pe canal utilizat.

-Matricea generatoare: proprietatea de liniaritate implica existenta unei baze pentru cod. Sa construim o matrice generatoare pentru cod, denumita G, folosind baza independenta de cuvantul de cod k pe post de sir de vectori G. Matricea G are dimensiunea k\*n si poate fi folosita pentru a codifica vectorul de informatie. Matricea generatoare descrie structura codului, totusi nu este definita numai pentru cod.

-Matricea controlului paritatii: reprezinta un alt mod de a descrie sructura codului.

-Distributia greutatii: definim vectorul de greutate x, ca numarul elementelor diferite de zero pe care le contine. Distributia greutatii unui cod este o lista cu numarul de cuvinte la fiecare greutate, pentru toate greutatile cuvantului cod.

-Distanta Hamming dintre doua cuvinte cod este numarul pozitiilor in care ele difera.

-Distanta minima: se noteaza cu dmin si este cea mai mica distanta Hamming dinre doua cuvinte de cod din multimea tuturor cuvintelor.

-Teorema Massey: Distanta minima a unui cod cu diagrama liniara binara este egala cu numarul minim de coloane diferite de zero in matricea sa de control al paritatii care insumeaza zero.

 Un cod de control al paritatii cu densitate mica este un cod cu o matrice liniara care are o matrice de control al paritatii slab populata. Codurile LDPC reprezinta o clasa a codurilor bloc liniare. Astfel ele pot fi codificate folosind matricea generatoare insa in absenta unei structuri in interiorul codului, me masura ce dimensiunea blocului de date creste aceasta metoda impune decodorului stocare mare si cerinte legate de procesare.

**Coduri structurate**

 Metoda cu matricea de control al paritatii cu structura triunghiulara, permite ca cea mai mare parte din bitii de paritate sa fie calculati in timp liniar folosind operatii putine, de exemplu functiile care nu implica decat un numar mic de variabile. Experimentele arata ca astfel de coduri au o performanta care e comparabila cu cea a codurilor regulate LDPC.

Exemplu de matrice de control al paritatii triunghiulara:



Un caz special de triungiularitate este intalnit atunci cand Hp are structura duala diagonala expusa. Aici substitutia regresiva urmeaza o structura simpla si poate fi efectuata folosind un acumulator. Exemplu:

O structura in forma de s pentru H:

 1 1 0 0 0 0

 0 1 1 0 0 0

Hp= 0 0 1 1 0 0

 0 0 0 1 1 0

 0 0 0 0 1 1

 0 0 0 0 0 1

Se observa ca daca matricea de control al paritatii are vreo coloana de greutatea j>1, si este strict triunghiulara, atunci structura codului este neaparat neregulata.

**Structuri de cod ciclic**

Pentru a asigura codificarea liniara de timp, putem folosi structurile ciclice si cvasi-ciclice. Un cod ciclic are proprietatea ca orice deplasare ciclica a unui cuvant de cod este in sine un cuvant de cod. Un cod cvasi-ciclic are proprietatea ca pentru marimea fixa de deplasare, o deplasare ciclica a oricarui cuvant de cod de acea marime este in sine un cuvant de cod. Putem construi matricea de control al paritatii pentru un cod LDPC cvasi-ciclic prin unirea pe orizontala a matricelor rare circulare, fiecare avand dimensiunea mxm.

 O matrice circulara este o matrice patrata in care fiecare sir este o deplasare ciclica a sirului anterior. Putem observa ca permutarea liniilor si coloanelor H nu schimba structura codului ci doar redenumeste controalele sau respectiv variabilele.



Metoda imparte matricea H in trei componente circulare mxn:

-Hp corespunde bitilor de paritate

-Hu1 si Hu2 corespund bitilor de informare.

Apoi multiplicam H cu Hp-1 pentru a obtine forma sistematica Hsys, dupa care permutam coloanele apartinand Hu(sys) in ordinea 1, m+1, 2, m+2 ,…. pentru a redefini bitii de informare corespunzator.

 Se poate extinde matricea de control al paritatii pentru a construi coduri de un randament mai mare prin adunarea orizontala a altor matrice circulare. Alegerea unei rate de control este restransa de relatia r=(n0-1)/n0.

**3.Decodarea codurilor LDPC**

 Deoarece in 1963 dupa ce Gallager a publicat teoria sa, specialistii nu aveau la dispozitie calculatoare performante, metoda aceasta nu a fost folosita. Un numar mic de cercetatori au continuat insa sa foloseasca codurile Gallager timp de cateva decade dupa publicarea tezei lui. In anul 1980 Tanner a furnizat o reprezentare grafica a LDPC si a altor scheme de codificare. El a propus ca problema decodificarii sa fie rezolvata prin descompunerea codurilor in coduri mai simple din structura lor si a introdus notiunea care este astazi cunoscuta sub denumirea de algoritm de decodificare a sumelor minime.

**Decodarea cu doua nivele de decizie (hard-decision decoding)**

 Acest algoritm nu are acelasi nivel de performanta ca algoritmii de decizie cu doua nivele, totusi, simplitatea lor permite executarea lor rapida si a dus de asemenea la cateva rezultate analitice interesante.

 Pentru a explica acest algoritm vom realiza un exemplu. Un cuvant de cod receptionat fara erori ar fi de exemplu c=[1 0 0 1 0 1 0 1].Sa consideram cuvantul receptionat cu o eroare la bitul c1 (c1 devine din 0 - 1) .

1. La primul pas toate nodurile variabile notate ci trimit un mesaj tuturor nodurilor de control notate cu fi (intotdeauna ci–fi este doi in exemplu nostru) ce contine un bit pe care il considera corect. In aceasta etapa doar informatiile continute de un nod variabil,este corespunzator bitului i receptionat din mesajul c, yi. Asta inseamna de exemplu, c0 trimite un mesaj ce contine 1 la f1 si f3, nodul c1 trimite mesajul y1(1)la f0 si f1, si asa mai departe.



1. In etapa doi fiecare nod de control fi calculeaza un raspuns pentru nodurile variabile la care este conectat. Mesajul raspuns contine bitul pe care fi il considera corect pentru nodul variabil ci considerand ca restul nodurilor variabile conectate la fi sunt corecte. Cu alte cuvinte daca ne uitam la exemplu fiecare nod fi este conectat la 4 noduri variabile.Astfel un nod de control fi analizeaza mesajele primite de la 3 noduri variabile si calculeaza bitul pe care ar trebui sa il receptioneze de la cel de-al 4 nod pentru a indeplini conditia de paritate.In continuare vom prezenta grafic acesti pasi.



Ceea ce este important este faptul ca in acest punct se poate incheia decodarea.Aceasta se realizeaza daca toate ecuatiile de control sunt indeplinite.O alta modalitate de a termina decodarea este depasirea numarului de iteratii(bucle) permise.

1. Urmatoarea etapa consta in: nodurile variabile receptioneaza mesaje de la nodurile de control si folosesc aceste informatii pentru a decide daca biti receptionati initial au fost receptionati corect.O metoda simpla pentru a face acest lucru este votul majoritar.Revenind la exemplul nostru asta inseamna ca fiecare nod variabil are trei surse de informatie referitoare la bitul sau.Bitul original este receptionat impreuna cu cele doua sugestii de la nodurile de control.Tabelul 2 ilustreaza acesti pasi.Acum nodurile variabile pot trimite un alt mesaj prin intermediul hard decision pentru nodurile variabile cu valoarea corecta.
2. Se revine la pasul 2.

In exemplul nostru la a doua trecere prin pasul doi algoritmul de decodare ar lua sfarsit deoarece c1 a votat 0 in ultima etapa.Aceasta corecteaza eroarea de transmisie si toate ecuatiile de control sunt acum satisfacute.



Tabel 2 explica pasul 3 al algoritmului de decodare.Nodurile v folosesc mesajele receptionate de la nodurile c pentru a realiza un vot majoritar asupra valorii bitului,

 **Decodarea soft-decision**

 Decodarea soft-decision al codurilor LDPC este bazata pe conceptul de belief propagation, are un randament de decodare mare si este astfel metoda preferata.Ideea fundamentala este asemanatoare cu cea din decodarea hard decision. Inainte de a prezenta algoritmul sa introducem cateva notatii:

* pi = p (ci = 1/yi)
* qij este un mesaj trimis de nodul variabil ci nodului de control fj. Fiecare mesaj contine intotdeauna perechea qij(0) si qij(1) care reprezinta probabilitatea ca yi sa fie “ 0 ” sau “1”
* Rij este mesajul trimis de nodul de control fj catre nodul variabil ci.Din nou avem rij(0) si rij(1) care indica probabilitatea ca yi este “0” sau “1”
1. Toate nodurile variabile trimit mesajele qij.Din moment ce nu exista nici o informatie disponibila in aceasta etapa, qij(1) = Pi si qij(0)= 1 - Pi



Figura 2 : a) ilustreaza calculul rji(b) iar b) calcul qij(b)

1. Nodurile de control calculeaza mesajele de raspuns rji2

 

 

Astfel calculeaza probabilitatea ca sa existe un numar par de “1” in nodurile variabile exceptand ci(asta reprezinta exact Vi / i).Aceasta probabilitate este egala cu probabilitatea rij(0) ci sa fie 0.

1. Nodurile variabile isi actualizeaza mesajul de raspuns pentru nodurile de control. Actualizarea este realizata conform ecuatiilor:





Unde constantele kij sunt alese astfel incat sa asigure relatia qij(0) + qij(1) = 1. Ci\jacum semnifica toate nodurile de control cu exceptia nodului fj.

 In acest moment nodurile variabile de asemenea actualizeaza estimarea curenta ĉi al variabilei ci. Aceasta se realizeaza prin calcularea probabilitatii pentru un 0 sau 1 si votand pentru cel mai mare. Ecuatia folosita este similara cu cea folosita pentru calculul q ij(b) dar acum sunt folosite toate informatiile de la nodurile de control.







Daca cuvantul de cod estimat indeplineste acum conditia de paritate algoritmul se incheie.In caz contrar finalizarea algoritmului este asigurata printr-un numar maxim de iteratii.

1. Se reia pasul 2.

Aceasta varianta pentru decodare este o varianta simpla, optima pentru canalele BSC si poate fi modificata pentru a aduce imbunatatiri.Pe langa problema performantei mai este si problema statistic- numerica datorita multiplicarilor repetate al probabilitatilor. Rezultatul va tinde spre zero pentru blocuri de lungime mare. Pentru a preveni asta se poate modifica domeniul si anume sa se adopte domeniul logaritmic, de aici rezulta ca multiplicarile vor fi inlocuite cu adunari. Rezultatul este un algoritm mult mai stabil care are si un castig in performanta deoarece adunarile sunt mai putin „avantajoase” decat multiplicarile.

**4.Concluzii si Aplicatii ale codurilor LDPC**

 Codurile LDPC sunt mult mai atractive din punctul de vedere al puterii de calcul deoarece ofera un nivel mai înalt al paralelismului intrinsec.Aceasta se realizeaza deoarece ecuatiile de verificare a paritati pot fi actualizate independent.În timp ce complexitatea interconexiunilor este mare, toate nodurile ce pot fi actualizate simultan si toate informatiile extrinseci pot fi transmise simultan nodurilor variabile.Similar toate nodurile variabile pot fi actualizate simultan.Acest lucru este în contrast cu codarea turbo, unde într-un cod constituent, toti biti de informatie sunt continuti de un trellis si intrinsec la fiecare moment trellis are nevoie sa obtina informatia de stare din stagiile adiacente trellis-ului.

 In ultima decada codurile LDPC au primit multa atentie datorita performantelor superioare de corectie de erori, si au fost adoptate de mai multe standarde recente de comunicatie, cum ar fi 10 Gigabit Ethernet (10GBASE-T), broadcast video digital (DVB-S2 –satelit- , DVB-T2 – terestru-), DVB-C2 – cablu- , Wimax (802.16e), Wi-Fi (802.11n) si 60 GHz WPAN (802.15.3c).

 Cei mai importanti parametrii care afecteaza castigul codarii asigurat de un cod LDPC sunt:

• tipul codului, dictat de modalitatea de constructie a matricii H

• rata codului

• lungimea cuvantului de cod

• valorile parametrilor k si j

• diametrul buclei minime (girth); acesta este corelat cu valorile parametrilor j si k

• numarul maxim de iteratii efectuate de algoritmul de decodare

 Dintre clasele de coduri mentionate, codurile random asigura castigurile cel mai mari. Exista coduri LDPC random care, pentru rate scazute se apropie de la cateva zecimi de decibeli de limta teoretica a lui Shannon si au performante comparabile (sau chiar mai bune decat) turbocodurile. Codurile “array-based” asigura castiguri mai mici decat cele random, diferentele depinzand de lugimea si rata studiate. Castigul codarii asigurat creste odata cu scaderea ratei codului. La o rata data, castigul codarii creste (de obicei) cu cresterea lungimii cuvantului de cod. Castigul codarii e maxim pentru valori mici ale parametrului j, dar trebuie ca j > 2. Pentru un cod dat, castigul codarii creste cu cresterea numarului de iteratii efectuate de algoritmul de decodare MP; trebuie însa mentionat ca la marirea numarului de iteratii efectuate peste 25, castigul codarii suplimentar este extrem de mic, fata de cel obtinut cu 25 de iteratii, nejustificand cresterea timpului de decodare indusa.

 Acum ca am prezentat pe scurt performantele codurilor LDPC, vom prezenta pe scurt domeniile in care sunt aplicate acestea.

**Domenii de aplicatii:**

-*in comunicatiile optice*

-*in comunicatiile in spatiul cosmic*. Comparativ cu comunicatiile standard terestre si satelitare, comunicatiile in spatiul cosmic prezinta o serie de probleme pentru transmisia de informatii, cum ar fi distantele foarte lungi care trebuie parcurse de semnal, raport semnal-zgomot foarte mic, intarzieri mari in propagarea semnalelor, rate de coruptie a datelor ridicate sau largimi de banda asimetrice. In studierea spatiului, pe langa tehnologia extraordinar de avansata folosita la emitator, receptor sau antene, un rol foarte important il au si codarea la sursa imaginii, codarea de canal, modulatia sau demodulatia.

-*in sistemele de comunicatii RF/microunde*. In acest caz, semnalul care se propaga prin atmosfera va experimenta fluctuatii in amplitudine si in faza datorita turbulentelor atmosferice (ceata, ploaie). Pentru a se combate macar o parte din aceste fluctuatii si a se imbunatatii performantele sistemelor de transmisiune, au fost propuse mai multe tehnici de codare.

-*in sisteme de comunicatii satelitare*. In acest caz atenuarea datorata ploii si umbrirea sunt principalii factori care deterioreaza performantele sistemului, pierderea semnalului radio, scaderea fiabilitatii si a performantelor link-urilor satelitari. Pentru a combate aceste probleme este necesara o metoda de codare eficienta pentru transmisiuni cu rata mare de date.

**5. Bibliografie**

* Robert G. Gallager: ‘*Low-Density Parity-Check Codes’*
* Jose M.F. Moura, Jin Lu, Haotian Zhang: *‘Structured Low-Density Parity-Check Codes’*
* Herman Myburg: *‘ LDPC Codes:An Investigation’*
* Ichimoaiei Stefan Cosmin: *’Modul de simulare de coduri corectoare de erori LDPC in canale magnetice’* (lucrare de dizertatie)