Eötvös Loránd Tudományegyetem
Természettudományi Kar

Kombinatorikus batch kódok

Szakdolgozat

Készítette:
Tardos Jakab

Témavezető:
Frank András

Budapest, 2016.

Köszönetnyilvánítás: *Szeretném megköszönni témavezetőmnek, Frank Andrásnak, segítségét és útmutatását a témakörben és a matematikában általában.*

Tartalomjegyzék

1. Bevezető

A batch kódok optimalizálása a kódelmélet egy újabban megjelent kérdésköre. A feladatot egy informatikai probléma motiválja: egy nagy méretű adatbázist kell szerverek segítségével lekódolni úgy, hogy az felhasználók számára elérhető legyen a szerverek túlterhelése nélkül. A szakdolgozat a batch kódok egy speciális változatáról, a kombinatorikus batch kódokról szól. Ez, a nevének megfelelően, már egy teljesen kombinatorikai feladat. A batch kódokat a 2. fejezetben precízen definiáljuk, mint adott paraméterekhez tartozó speciális adottságokkal rendelkező, adott tulajdonságú halmazrendszerek. Ekkor a feladat a halmzarendszer össz-méretének minimalizálása.

Az optimális kombinatorikus batch kódok méretére még nem ismert megoldás általános paraméterek mellett. Ezért a dolgozat nagy részében a paraméterek speciális beállításai mellett vezetjük le az optimális kód méretét. Eközben különböző módszereket és megfigyeléseket mutatunk be a batch kódok tanulmányozására.

A 3. fejezetben levezetünk egy általános alsóbecslést a batch kódok méretére, amit néhány konstrukcióval el is érünk rendkívül nagy adatbázisméret esetén. Valamint bemutatunk egy kapcsolatot a kombinatorikus batch kódok és a konstans súlyú bináris Hamming kódok között. A 4. fejezetben optimalitás-tartó transzformációkkal egyszerűsítjük az ismeretlen optimális kód struktúráját, amíg elég könnyen kezelhető lesz, hogy éles alsó becslést bizonyítsunk a méretére. A 6. fejezetben, többek között, bevezetünk egy irányított segédgráfot, amelyen átláthatóvá válik a batch kódok struktúrája egy egyszerű esetben. Ezen kívül bemutatjuk a kis adatbázis esetén ismert optimális megoldásokat. A 7. fejezetben a batch kódok transzverzális matroidjait vizsgáljuk, és ennek a módszernek a segítségével kapunk is egy alternatív bizonyítást az előző fejezet legnehezebb tételére. A 8. fejezetben röviden bemutatjuk az uniform batch kódok feladatkörét. A 9. fejezetben a véges testek feletti affin síkokat alapul véve konstruálunk érdekes batch kódokat. Ezek a paraméterek egy eddig feltérképezetlen állása mellett adnak optimális értékeket és mutatják meg a 3. fejezetben kimondott alsóbecslés élességét.

A dolgozat célja, hogy összefoglalja a témakörben eddig elért eredményeket és szemléltesse a batch kódok optimalizásának különböző trükkjeit és módszereit.

2. Alapok

A batch kódok fogalmát először Y. Ishai, E. Kushiletitz, R. Ostrovsky és A. Sahai vezették be 2004-es cikkükben [13], ahol a következő problémát vetették fel: Tegyük fel, hogy egy adatbázis n tárgyat (vagy bitet) tárol, melyek egy Σ ábécé elemei. A tárolt adatokat elérhetővé szeretnénk tenni felhasználók számára. Ennek érdekében a tárgyakban tárolt információt lekódolva szétosztjuk m darab szerver között (szintén Σ-beli betűkkel kódolunk). Egy felhasználó egyszerre legfeljebb k darab tárgyat szeretne megtekinteni, viszont nem szeretnénk, hogy a szerverek terhelése túl nagy le-

gyen. Ez okból kikötjük, hogy minden felhasználó minden szerverről legfeljebb t tárgyat olvashat le.

Ha például $t \geq k$, akkor egyszerű dolgunk van, hiszen a tárgyakat mind ráírhatjuk egy szerverre és így sem ütközünk problémába, hiszen egy felhasználó nem tud annyi tárgyra rákérdezni, hogy túlterhelje az egyetlen szervert. A másik véglet, ha $m \geq n$, azaz elég szerver áll a rendelkezésünkre, hogy minden egyes tárgy külön szervert kapjon. Ekkor a szerverek terhelése legfeljebb egy lehet, mert csak egy-egy tárgyat tárolnak.

A batch kód méretének mondjuk a szervereken tárolt összes információt és N-nel jelöljük. Egy batch kód rátája n/N és általában egy batch kód annál jobb, minél nagyobb a rátája. A feladat általában adott (n, k, m, t, Σ) mellett a lehető legkisebb N megtalálása. A fent említett esetekben elértük az 1 rátát, amit természetesen nem lehet meghaladni. Azonban az általános esetben a feladat ennél sokkal bonyolultabb.

Defináljuk precízen a batch kódok fogalmát. A triviális eset kizárásának érdekében kikötünk néhány egyszerű összefüggést a változók között.

2.1. Definíció ([13]). *Legyenek n, m, k és t egészszámok, amikre $1 \leq t \leq k$ és $k \leq tm \leq n$. Továbbá legyen Σ egy véges ábécé. Legyen $x_1x_2...x_n \in \Sigma^n$ betűsorozat lekódolva az $y_1, y_2, ..., y_m \in \Sigma^*$ tetszőlegesen hosszú betűsorozatokkal. Ezt batch kódnak hívjuk, ha tetszőleges $i_1, i_2, ..., i_k \in \{1, ..., n\}$ indexhalmaz esetén $x_{i_1}, x_{i_2}, ..., x_{i_k}$ dekódolható $y_1, y_2, ..., y_m$ néhány betűjéből úgy, hogy minden y_i betűsorozatból legfeljebb t betűt használunk.*

Ezután két egyszerűsítést végzünk a defíniált feladaton. Ellőször is, a lekódolásnak csak egy nagyon egyszerű fajtájával fogunk foglalkozni: a másolás-alapú kódolással. Azaz az adatbázis tárgyait egyszerűen átmásoljuk a szerverek egy részére. Az információ a felhasználó által való dekódolása ekkor egyszerű olvasás, aminek feltétele annyi, hogy a kiolvasott betűk között ott legyen a keresett tárgy másolata. Ezzel egy teljesen kombinatorikai feladatot kaptunk és az ilyen batch kódokat kominatorikus batch kódoknak nevezzük. A definíciójuk az eredetinél lényegesen egyszerűbb, például megfigyelhető, hogy független lesz a Σ ábécétől, ezért ezt el is hagyjuk.

2.2. Definíció ([14]). *Legyenek n, m, k és t egészszámok, amikre $1 \leq t \leq k$ és $k \leq tm \leq n$. Legyen $X = \{x_1, x_2, ..., x_n\}$ a tárgyak halmaza. $Y_1, Y_2, ..., Y_m \subseteq X$ szerverek kombinatorikus batch kódot alkotnak, ha tetszőleges $\{i_1, i_2, ..., i_k\}$ indexhalmazra léteznek $Z_1 \subseteq Y_1, Z_2 \subseteq Y_2, ..., Z_m \subseteq Y_m$ kiolvasandó részhalmazok, hogy*

$$\forall i : |Z_i| \leq t$$

$$\{x_{i_1}, x_{i_2}, ..., x_{i_k}\} \subseteq \cup_{i=1}^{m} Z_i.$$

Ekkor a kód mérete $N = \sum_{i=1}^{m} |Y_i|$ és $CBC(n, N, k, m, t)$-vel jelöljük. Adott (n, m, k, t) mellett a lehető legkisebb N jelölése $N(n, k, m, t)$.

Itt a kombinatorikus batch kódot mint egy speciális tulajdonsággal rendelkező halmazrendszert defíniáltuk. Egy batch kód megadható a duális halmazrendszerével is [10], a következőképpen. A változókat a fenti definícióval azonos módon kötjük meg. Legyen $Y = \{y_1, y_2, ..., y_m\}$ a szerverek halmaza. Egy kombinatorikai batch kód duális rendszere $F_1, F_2, ..., F_n \subseteq Y$, ahol F_i azon szerverek halmaza, amelyek tárolják x_i tárgy egy példányát a kódban. A kombinatorikus batch kódok ilyen megadása sok esetben hasznosabbnak bizonyul mint az eredeti, ezért számos cikk ezt használja.

Mi egy gráfokon alapuló reprezentációt fogunk használni a dolgozat során. Adott kombinatorikus batch kód esetén definiálunk egy G páros gráfot az (S, T) csúcsosztályokon. $|S| = m$ a szerverek halmaza és $|T| = n$ az adatbázis tárgyainak halmaza. Egy adott pár akkor van benne az E élhalmazban, ha az adott szerver tárolja az adott tárgy egy másolatát. A batch kódok helyessége könnyen átfogalmazható a reprezentáló párosgráfra.

2.3. Állítás. *Egy $G = (S, T; E)$ párosgráf pontosan akkor reprezentál egy $CBC(n, N, k, m, t)$-t, ha $|S| = m$, $|T| = n$, $|E| = N$ és tetszőleges k méretű $X \subseteq T$ halmazhoz létezik egy részgráf G-ben, aminek minden S-beli fokszáma legfeljebb t és minden X-beli fokszáma legalább egy (feltehető, hogy pontosan egy). Ekkor az egyszerűség kedvéért azt mondjuk, hogy G egy batch kód, ahelyett hogy G egy batch kódot reprezentál.*

Ebből a grafikus reprezentációból kiolvasható, a batch kód másik két reprezentációja. A definíció szerinti Y_i halmazok megjelennek, mint az S-beli csúcsok szomszédságai, a duális halmazrendszer (az F_i halmazok) pedig megjelennek, mint a T-beli csúcsok szomszédságai.

(G-ben egy csúcsot egy másik csúcs szomszédjának nevezünk, ha a kettő között fut él. Ennek megfelelően egy $u \in S \cup T$ csúcs szomszédsága azon csúcsok halmaza, amik össze vannak vele kötve. Ezt $\Gamma(u)$-val jelöljük. Egy $X \subseteq S \cup T$ halmaz szomszédsága $\Gamma(X) = \cup_{u \in X} \Gamma(u)$. Ha $\Gamma(X) = \{v\}$, azt mondjuk, hogy $\Gamma(X) = v$. Ha $X \subseteq T$, $\Gamma(X) \subseteq S$ és fordítva. $|\Gamma(u)| = d(u)$ u fokszáma vagy foka.)

A második egyszerűsítés az, hogy a cikk nagy részében a $t = 1$ esettel fogunk foglalkozni. Ezt majdnem minden batch kódokkal foglalkozó cikk így teszi, mivel ez a speciális eset is nehéz és mély problémakört takar. Ezzel leegyszerűsítjük a jelölést is: Legyen $CBC(n, N, k, m, 1) = CBC(n, N, k, m)$ és $N(n, k, m, 1) = N(n, k, m)$. Ennek ellenére visszatérünk az általános esetre az utolsó fejezetben. A $t = 1$ esetben az 1.1 állításban jellemzett részgráf pontosan egy X-et fedő párosítás lesz. A Hall tétel segítségével a kód szabályosságának egy nagyon hasznos ekvivalens feltételéhez jutunk, ami megalapozza szinte az összes optimalitási bizonyítást.

2.4. Állítás ((Hall feltétel[14]). *G párosgráf pontosan akkor batch kód, ha minden $i \leq k$ méretű $X \subseteq T$ halmazra $|\Gamma(X)| \geq |X|$.*

Bizonyítás. A feltétel láthatóan szükséges, mert ha egy X halmazra nem teljesül, akkor semmilyen X-et tartalmazó halmaz nem fedhető párosítással. Ha viszont teljesül a Hall feltétel, akkor tetszőleges legfeljebb k méretű X lefedhető párosítással az $X \cup S$-re leszűkített gráfra alkalmazott Hall tétel alapján.

□

Ha egy halmaz nem teljesíti a rá vonatkozó Hall feltételt, akkor azt mondjuk a halmaz Hall-hiányos vagy hiányos. Még érdemes kimondani a Hall feltétel S-re vonatkozó ekvivalens változatát. Egyes bizonyításoknál hasznos ezt felírni az eredeti T-re vonatkozó Hall feltétel helyett.

2.5. Állítás. *G párosgráf pontosan akkor batch kód, ha minden $i > m - k$ méretű $Y \subseteq S$ halmazra $|\Gamma(Y)| \geq |Y| + n - m$.*

Bizonyítás. Ha $X \subseteq T$ hiányos halmaz, akkor $|\Gamma(X)| < k$, ezért $|Y| = |S \backslash \Gamma(X)| > m-k$, viszont $|\Gamma(Y)| = |T \backslash X| < n-|\Gamma(X)| = |n-m+|Y|$. Tehát $Y = S \backslash \Gamma(X)$ megszegi az S-re vonatkozó Hall feltételt. Hasonlóan belátható, hogy ha $Y \subseteq S$ hiányos, akkor $T \backslash \Gamma(Y)$ is az, így a kétféle Hall feltétel ugyanakkor teljesül G-n.

□

3. Optimális kódok nagy n esetén

Az első speciális eset, amit megvizsgálunk, az m-hez képest relative nagy n esete. Az $n \geq \binom{m}{k-2}$ határ felett már ismertek optimális konstrukciók minden m és k esetén. E határ alatt is tárgyalunk egy közel-optimális konstrukciót. A fejezetet egy hasznos általános becsléssel kezdjük. Ez teszi lehetővé a konstrukciók optimalitásának bizonyítását nagy n-ek esetén. Kisebb n értékekre nem feltételnül éles, bár a 9. fejezetben belátjuk, hogy végtelen sok esetben az.

Könnyen látható, hogy a T-ben lévő csúcsok fokai fölösleges hogy k-nál nagyobbak legyenek. Ha egy $CBC(n, m, k)$ egy $t \in T$ csúcsából több mint k él indul ki, akkor ezen élek közül kitörölhetünk néhányat úgy, hogy még k darab megmaradjon. Ekkor T k-nál kisebb részhalmazai közül csak a t-t tartalmazóknak csökkenhet a szomszédságának mérete, de ezek nem sérthetik meg a Hall-feltételt, mert $|\Gamma(t)| = k$. Ebből következik, hogy optimális batch kódoknál mindnen T-beli csúcs foka legfeljebb k.

[14] Tekintsünk egy tetszőleges optimális batch kódot. Osszuk a T osztályt komponensekre a csúcsok fokszáma alapján: T_i azon T-beli csúcsok halmaza legyen, amelyeknek fokszáma i $(1 \leq i \leq k)$. Legyen $A_i = |T_i|$. Ekkor a batch kód mérete

$$N = \sum_{i=1}^{k} iA_i.$$

Valamint

$$n = \sum_{i=1}^{k} A_i.$$

Végezzünk kettős leszámlálást az (S', t) párokon, ahol S' S-nek $k-1$ elemű részhalmaza, $t \in T$ és $\Gamma(t) \subseteq S'$. Tetszőleges $t \in T$ esetén $\Gamma(t)$-t S-nek éppen $\binom{m-d(t)}{(k-1)-d(t)}$ darab $k-1$-részhalmaza tartalmazza. Viszont semelyik S' nem tartalmazhatja T k különböző csúcsának a szomszédságát, mert ekkor ez a k csúcs olyan halmazt alkotna, ami megszegi a Hall feltételt. Ebből a következő egyenlőtlenség olvasható ki:

3.1. Állítás ([14]).

$$\sum_{i=1}^{k-1} \binom{m-i}{k-1-i} A_i \leq (k-1)\binom{m}{k-1}$$

[14] Ebből levezethető egy N-re vonatkozó alsó becslés. (Megjegyzés: Definíció szerint $\binom{n}{k} = 0$, ahol $k < 0$ vagy $k > n$, így a fenti egyenlőtlenségben a szumma tartománya kiterjeszthető k-ig.)

$$N = \sum_{i=1}^{k} iA_i = \sum_{i=1}^{k} (k-(k-i))A_i = kn - \sum_{i=1}^{k} (k-1)A_i$$

$$= kn - \sum_{i=1}^{k} (k-i)A_i + \left(\sum_{i=1}^{k-1} \binom{m-i}{k-1-i} A_i - (k-i)\binom{m}{k-1} \right)$$

$$= kn - (k-1)\binom{m}{k-1} + \sum_{i=1}^{k} \left(\binom{m-i}{k-1-i} - (k-i) \right) A_i$$

A_i együtthatója minden $1 \leq i \leq k$ esetén nemnegatív, ezért a szumma alulról becsülhető nullával. Valóban, $i = k$ és $i = k-1$ esetén láthatóan $\binom{m-i}{k-1-i} = k-i$. Ezután, ha i-t egyel csökkentjük $k-i$ egyel nő, de $\binom{m-i}{k-1-i}$ $\binom{m-i}{k-i}$-vel nő meg. Ekkor a következőt kapjuk:

3.2. Becslés ([14]).

$$N \geq kn - (k-1)\binom{m}{k-1}$$

A szummában minden A_i együtthatója nemnegatív és éppen A_k és A_{k-1} együtthatója 0. Ez azt jelenti, hogy egy $CBC(n, m, k)$ csak akkor teljesítheti egyenlőséggel a 3.2 becslést, ha minden T-beli csúcs foka k vagy $k-1$. Kiderül, hogy ez az $n \geq (k-1)\binom{m}{k-1}$ tartományban megvalósítható.

3.3. Konstrukció ([14]). *S minden $k-1$ méretű részhalmazhoz válasszunk $k-1$ darab csúcsot T-ből, amik éppen ennek a részhalmaznak a csúcsaival lesznek összekötve. (Ezt megtehetjük, mivel $|T| \geq (k-1)\binom{m}{k-1}$.) T maradék csúcsait kössük össze S tetszőleges k csúcsával.*

A 3.3 konstrukcióban T minden csúcsának fokszáma legalább $k-1$, ezért a Hall-feltételt csak k méretű halmaz szegheti meg. Továbbá, hiányos halmaz nem tartalmazhat k fokú csúcsot, így csak k darab $k-1$-fokú csúcsból állhat. Azonban semelyik k $k-1$-fokú csúcsnak nem egyezik meg pontosan a szomszédsága, ezért minden ilyen halmaz szomszédsága legalább k méretű. A Hall feltétel nem sérül meg, így a konstrukció egy batch kódot alkot. A 3.3 konstrukció mérete

$$N = \left(n - (k-1)\binom{m}{k-1}\right)\cdot(k) + \left((k-1)\binom{m}{k-1}\right)\cdot(k-1)$$
$$= nk - (k-1)\binom{m}{k-1}$$

Tehát a kód optimális.

3.4. Tétel ([14]). *Ha $n \geq (k-1)\binom{m}{k-1}$, akkor $N(n,k,m) = kn - (k-1)\binom{m}{k-1}$.*

Ez az érték azonban már nem érhető el $(k-1)\binom{m}{k-1}$-nél kisebb n-ekre. Ha n kisebb, mint $(k-1)\binom{m}{k-1}$, akkor a fenti módszer általánosításával kaphatunk értékes alsóbecslést.

3.5. Becslés ([2]). *Tetszőleges $1 \leq c \leq k-1$ egészszám esetén*

$$N \geq nc - \frac{(k-c)(U_{m,k,c} - n)}{m-k+1},$$

ahol $U_{m,k,c} = \frac{(k-c)\binom{m}{c}}{\binom{k-1}{c}}$.

Bizonyítás. **[2]** Ismét a 2.1 egyenlőtlenséget fogjuk használni. Azonban most súlyozzuk $\frac{1}{\binom{m-c}{k-c}}$-vel.

$$N = \sum_{i=1}^{k} iA_i = \sum_{i=1}^{k}(c-(c-i))A_i = nc - \sum_{i=1}^{k}(c-i)A_i$$

$$\geq nc - \sum_{i=1}^{k}(c-i)A_i + \frac{1}{\binom{m-c}{k-c}}\left(\sum_{i=1}^{k-1}\binom{m-i}{k-1-i}A_i - (k-1)\binom{m}{k-1}\right)$$

$$= nc - \sum_{i=1}^{k}(c-i)A_i + \sum_{i=1}^{k-1}\left(\frac{\binom{m-i}{k-i-1}}{\binom{m-c}{k-c}} - \frac{k-c}{m-k+1}\right)A_i - \frac{(k-c)(U_{m,k,c}-n)}{m-k+1}$$

Itt az utolsó lépésben kihasználtuk, hogy $\sum_{i=1}^{k} A_i = n$, valamint, hogy

$$\frac{(k-1)\binom{m}{k-1}}{\binom{m-c}{k-c}} = \frac{(k-1)m!(k-c)!(m-k)!}{(k-1)!(m-k+1)!(m-c)!} = \frac{(k-c)U_{m,k,c}}{m-k+1}$$

Majd a kifejezést rendezzük.

$$nc - \frac{(k-c)(U_{m,k,c}-n)}{m-k+1} + \sum_{i=1}^{k}\left(\frac{\binom{m-i}{k-i-1}}{\binom{m-c}{k-c}} - \frac{k-c}{m-k+1} - (c-i)\right)A_i$$

Az előző esethez hasonlóan itt is igaz lesz, hogy minden A_i együtthatója nemnegatív, azonban ez kevésbé nyilvánvaló. A szummát itt is alulról becsülhetjük nullával.

3.6. Lemma. *Minden $0 < i$ esetén, $f(i) = \frac{\binom{m-i}{k-i-1}}{\binom{m-c}{k-c}} - \frac{k-c}{m-k+1} - (c-i) \geq 0$.*

Bizonyítás.

$$f(i) - f(i-1) = \frac{\binom{m-i}{k-i-1} - \binom{m-i+1}{k-i}}{\binom{m-c}{k-c}} - 1 = \frac{-\binom{m-i}{k-i}}{\binom{m-c}{k-c}} - 1$$

$$= \frac{\binom{m-c}{m-k} - \binom{m-i}{m-k}}{\binom{m-c}{k-c}}$$

Látható, hogy ez $i < c$-re pozitív, $i = c$-re 0 és $i > c$-re negatív, azaz az $f(i)$ kifejezés $i = c$-ben és $i = c-1$-ben veszi fel a minimumát. Ezeken a helyeken éppen 0.

□ □

Ebből a levezetésből az is kijön, hogy egy $CBC(n, m, k)$ csak akkor teljesítheti a c-hez tartozó becslést egyenlőséggel, ha minden T-beli csúcs fokszáma c vagy $c-1$. Még eldöntendő kérdés, hogy n-től függően melyik c értékre a legerősebb a 3.5 becslés. Vizsgáljuk a $c+1$-hez és a c-hez tartozó becslések különbségét!

$$n(c+1) - \frac{(k-c-1)(U_{m,k,c+1}-n)}{m-k+1} - nc + \frac{(k-c)(U_{m,k,c}-n)}{m-k+1}$$

$$= n + \frac{-(k-c-1)\frac{m-c}{k-c-1}U_{m,k,c} + (k-c)(U_{m,k,c}-n)}{m-k+1} = \frac{(k-m)(U_{m,k,c}-n)}{m-k+1}$$

Itt kihasználtuk, hogy $U_{m,k,c+1} = \frac{(k-1)\binom{m}{c+1}}{\binom{k-1}{c+1}} = \frac{(k-1)\binom{m}{c}\frac{m-c}{c+1}}{\binom{k-1}{c}\frac{k-c-1}{c+1}} = \frac{m-c}{k-c-1}U_{m,k,c}$. Ebből az is kiderül, hogy $U_{m,k,c}$ monoton nő c-ben. A kapott kifejezésből kiolvasható, hogy a 3.5 becslés nő amíg $n > U_{m,k,c}$, ezután pedig csökken. Így az optimális becslést a legkisebb olyan c adja, amire $n \leq U_{m,k,c}$. (Kivéve az $m = k$ esetben. Ezt az esetet külön tárgyaljuk.)

3.7. Becslés ([2]). *Ha $k \neq m$ és $U_{m,k,c-1} < n \leq U_{m,k,c}$, akkor*

$$N \geq nc - \left\lfloor \frac{(k-c)(U_{m,k,c} - n)}{m-k+1} \right\rfloor .$$

A hányadosnak vehetjük az alsó egészrészét, mert N és nc is egész. Speciálisan a $c = k-1$ esetre létezik a 3.7 becslést egyenlőséggel teljesítő konstrukció.

3.8. Konstrukció ([2][9]). *A feltétel alapján $\binom{m}{k-2} \leq n \leq (k-1)\binom{m}{k-1}$. Először vegyük T $\left\lfloor \frac{(k-1)\binom{m}{k-1}-n}{m-k+1} \right\rfloor$ darab csúcsát és mindegyiket kössük össze S egy különböző $k-2$ méretű részhalmazával. Ezt megtehetjük, mert*

$$\left\lfloor \frac{(k-1)\binom{m}{k-1} - n}{m-k+1} \right\rfloor \leq \frac{(k-1)\binom{m}{k-1} - \binom{m}{k-2}}{m-k+1} = \frac{(k-1)\binom{m}{k-2}\frac{m-k+2}{k-1} - \binom{m}{k-2}}{m-k+1} = \binom{m}{k-2}.$$

Ezek a csúcsok fogják alkotni T_{k-2}-t. T maradék csúcsainak mindegyikét kössük össze $k-1$ darab csúcssal miközben vigyázunk arra, hogy S semelyik $k-1$ elemű részhalmaza se tartalmazza k darab különböző csúcs szomszédságát is. (Ez egy hiányos halmaz létezését jelentené T-ben.) Tehát S minden S' $k-1$ elemű részhalmazához válasszunk legfeljebb $k-1-|\{t \subseteq T_{k-2} : \Gamma(t) \in S'\}|$ darab csúcsot T-ből és ezeket kössük össze S' minden csúcsával. Az így kiválasztott T-beli csúcsok alkotják T_{k-1}-et.

$$\sum_{\substack{S' \subseteq S \\ |S'| = k-1}} (k-1-|\{t \in T_{k-2} : \Gamma(t) \subseteq S'\}|) + A_{k-2}$$

$$= (k-1)\binom{m}{k-1} - (m-k+2)A_{k-2} + A_{k-2}$$

$$= (k-1)\binom{m}{k-1} - (m-k+1)\left\lfloor \frac{(k-1)\binom{m}{k-1} - n}{m-k+1} \right\rfloor \geq n$$

Így elérhető, hogy T minden csúcsának fokszáma $k-1$ vagy $k-2$ legyen.

Megmutatjuk, hogy a fenti konstrukció valóban batch kódot alkot. Tegyük fel, hogy egy $X \in T$ csúcshalmaz megszegi a Hall feltételt. Mivel T-ben minden csúcs fokszáma legalább $k-2$, ezért $\Gamma(X)$ csak $k-2$ vagy $k-1$ lehet. $|\Gamma(X)| = k-2$ esetén X nem lehet hiányos, mert minden $k-2$-fokú csúcsnak különböző a szomszédsága. $|\Gamma(X)| = k-1$ esetén X nem tartalmazhat $k-2$-nél több csúcsot T_{k-2}-ből, arra pedig vigyáztunk T_{k-1} konstruálása során, hogy $\Gamma(X)$ ne tartalmazza k különböző csúcs szomszédságát. A kód mérete

$$(n - A_{k-2}) \cdot (k-1) + A_k \cdot (k-2) = n(k-1) - A_{k-2} = n(k-1) - \left\lfloor \frac{(k-1)\binom{m}{k-1} - n}{m-k+1} \right\rfloor ,$$

tehát a kód optimális.

□

3.9. Tétel ([2][9]). *Ha* $\binom{m}{k-2} \leq n \leq (k-1)\binom{m}{k-1}$, *akkor*

$$N(n,k,m) = n(k-1) - \left\lfloor \frac{(k-1)\binom{m}{k-1} - n}{m-k+1} \right\rfloor$$

A $c = k-2$ esetben a tartománynak már csak egy részére ismert a 2.7 becslést egyenlőséggel teljesítő konstrukció. A konstrukció Hamming kódokat használ.

Hamming kódok [3]. Egy Hamming kód azonos (m) hosszúságú $(0,1)$-sorozatokból (kódszavakból) álló halmaz. Két kódszó Hamming-távolsága azon 1 és m közötti i számok száma, amelyekre igaz, hogy a két kódszó eltér az i. számjegyben. Általában egy kód esetén megkövetelünk valamilyen minimum távolságot bármely két kódszó között. Egy kódszó Hamming-súlya a 0-sorozattól való távolsága. A konstrukció minél nagyobb méretű konstans súlylú Hamming kódokat használ. Ez azt jelenti, hogy minden kódszó súlya azonosan egy előre megszabott érték. Jelölje a legnagyobb m-hosszú, w súlyú szavakból álló Hamming kód méretét, ahol a kódszavak távolsága legalább $2d$, $A(m,2d,w)$. A konstrukcióban speciálisan az $A(m,4,k-3)$. Ennek pontos értéke nem ismert, de egy létező becslés $A(m,4,w) \geq \frac{1}{m}\binom{m}{w}$.

3.10. Konstrukció. *A konstrukció az* $\binom{m}{k-2} - (m-k+1)A(m,4,k-3) \leq n \leq \binom{m}{k-2}$ *intervallumban valósítható meg, valamint feltesszük, hogy* $k \geq 5$. *(A* $k < 5$ *esetet külön tárgyaljuk a következő fejezetben.)*

Legyen H *egy* $\left\lfloor \frac{\binom{m}{k-2}-n}{m+k-1} \right\rfloor$ *méretű, 4 távolságú, konstans* $k-3$ *súlyú Hamming kód. Ilyen létezik, mivel*

$$\binom{m}{k-2} - (m-k+1)A(m,4,k-3) \leq n$$

$$\binom{m}{k-2} - n \leq (m-k+1)A(m,4,k-3)$$

$$\left\lfloor \frac{\binom{m}{k-2} - n}{m+k-1} \right\rfloor \leq A(m,4,k-3).$$

Számozzuk S *csúcsait* 1*-től* m*-ig. Ekkor minden* H*-beli kódszó tekinthető úgy, mint* S *egy részhalmazának a karakterisztikus vektora. Legyen* S_H S *azon részhalmazainak a halmaza, amelyeket* H *egy kódszava ír le, mint karakterisztikus vektor. Minden* $S' \in S_H$*-hoz halmazhoz vegyünk két csúcsot, amelyeket éppen ennek a halmaznak a csúcsaival kötünk össze. Ezek a csúcsok alkotják* T_{k-3}*-at.*

T maradék csúcsainak mindegyikét kössük össze $k-2$ darab csúccsal úgy, hogy semelyik ilyen csúcs szomszédsága se tartalmazzon egy S_H-beli halmazt és semelyik két ilyen csúcs szomszédsága ne egyezzen meg pontosan. Ezt megtehetjük, mert minden S_H-beli halmaz S $m-k+3$ $k-2$ méretű részhalmazát zárja ki, és

$$2 \cdot |S| + \left(\binom{m}{k-2} - |S|(m-k+3) \right) = \binom{m}{k-2} - |S|(m-k+1) \leq n,$$

tehát elérhető, hogy T minden csúcsának fokszáma $k-3$ vagy $k-2$ legyen.

Megmutatjuk, hogy a konstrukció valóban batch kódot alkot. Tegyük fel, hogy egy $X \in T$ csúcshalmaz megszegi a Hall feltételt. Ha $\Gamma(X) < k-1$, X legfeljebb egy T_{k-2}-beli és legfeljebb kettő T_{k-3}-beli csúcsot tartalmaz, így könnyen látható, hogy nem lehet hiányos. Feltehető, hogy $\Gamma(X) = k-1$. Legyen $X_{k-3} = T \cap T_{k-3}$ és $X_{k-2} = T \cap T_{k-2}$. Ha $t_1 \in X_{k-3}$, akkor $\Gamma(t_1)$ $\Gamma(X)$ egy $k-3$-részhalmaza, így létezik két darab $k-2$ méretű halmaz $\Gamma(X)$-ben, ami tartalmazza. Legyenek ezek egyike $\Gamma'(t_1)$. Mivel T_{k-3}-at egy 4 távolságú Hamming kód alapján definiáltuk, és minden kódszóhoz csak két csúcs tartozik, Γ' definiálható úgy, hogy $\Gamma'(t_1) \neq \Gamma'(t_2)$ különböző $t_1, t_2 \in X_{k-3}$ esetén. Továbbá, T_{k-2} konstrukciójából látható, hogy $\Gamma'(t_1) \neq \Gamma(t)$, ha $t \in X_{k-2}$.

$$k-1 = \binom{k-1}{k-2} \geq |\{\Gamma'(t) : t \in X_{k-3}\}| + |\{\Gamma(t) : t \in X_{k-2}\}|$$

$$= |X_{k-3}| + |X_{k-2}| = X$$

Tehát X nem hiányos. A kód mérete

$$A_{k-2} \cdot (k-2) + (n - A_{k-2})(k-3) = n(k-2) - 2 \left\lfloor \frac{\binom{m}{k-2} - n}{m+k-1} \right\rfloor.$$

Ez megegyezik a 3.7 alsóbecsléssel, ha $0 \leq \binom{m}{k-2} - n < \frac{m-k+1}{2}$ $mod(m-k+1)$ és egyel nagyobb nála, ha $\frac{m-k+1}{2} \leq \binom{m}{k-2} < m-k+1$ $mod(m-k+1)$. Tehát a konstrukció optimális az esetek felében.

□

c kisebb értékeire nem ismert a 2.7 becslést egyenlőséggel teljesítő konstrukció. n $\binom{m}{k-2} - (m-k+1)A(m,4,k-3)$ alatti értékeire nem ismert optimális konstrukció általános m és k mellett, kivéve ha n nagyon kicsi, azaz $n \leq m+2$. Ezzel az esettel a 5. fejezet foglalkozik.

4. Optimális kódok kis k esetén

A kis konstans k esete természetesen adódik speciális esetként a 3. fejezet tételeiből. A $k = 1, 2, 3$ esetekben az $n \geq \binom{m}{k-2}$, ami felett tökéletesen jellemezhető $N(n, k, m)$, annyira alacsony, hogy feltétel nélkül teljesül. Sokkal tanulságosabb a $k = 4$ eset bizonyítása. Az itt kapott képlet az optimális batch kódra nagyon bonyolult: 8 esetet különböztet meg n és m értékei szerint, és nem olvasható ki belőle általánosítható szabály. A $k \geq 5$ esetén nem ismert általános optimális batch kód.

$k = 1$ esetén könnyen látható, hogy szükséges és elégséges feltétel az, hogy T-ben ne legyen izolált csúcs. Ebből azonnal következik erre az esetre az optimális kód mérete.

4.1. Tétel. $N(n, 1, m) = n$.

A $k = 2$ és $k = 3$ esetek már nem ilyen nyilvánvalóak. Azonban ezek az esetek még n minden lehetséges értékére kiolvashatóak a 3.4 és 3.9 tételekből. Ha $k = 2$, $U_{m,2,1} = m$, így a 3.4 tétel minden esetben alkalmazható. Behelyettesítve $k = 2$-t megkapjuk N optimális értékét.

4.2. Tétel ([9]). $N(n, 2, m) = 2n - m$.

Ha $k = 3$, $U_{m,3,2} = m^2 - m$ és $U_{m,3,1} = m$, így a két intervallumot külön kell vizsgálni a 2.4 illetve a 2.7 tétel segítségével. Behelyettesítve $k = 3$-at megkapjuk N optimális értékét.

4.3. Tétel ([9]).

$$N(n, 3, m) = \begin{cases} 2n - m + \left\lfloor \frac{n-3}{m-2} \right\rfloor & ha \quad n \leq m^2 - m \\ 3n - m^2 + m & ha \quad n \geq m^2 - m \end{cases}$$

A $k = 4$ eset ezekhez képest meglepően bonyolult. Ebben az esetben $U_{m,4,3} = 3\binom{m}{3}$ és $U_{m,4,2} = \binom{m}{2}$. Tehát ha $n \geq \binom{m}{2}$, akkor a 3.4 és 3.9 tételek segítségével kiszámítható az optimális N érték. Ha $n \geq 3\binom{m}{3}$, akkor $N = 4n - 3\binom{m}{3}$ és ha $n \geq \binom{m}{2}$, akkor $N = 3n - \left\lfloor \frac{3\binom{m}{3} - n}{m-3} \right\rfloor = 3n - \left\lfloor \frac{m^3 - 3m^2 + 2m - 2n}{2(m-3)} \right\rfloor = 3n - \left\lfloor \frac{m^2}{2} - \frac{n-m}{m-3} \right\rfloor$.

[9] Ha $n < \binom{m}{2}$, akkor a $CBC(n, 4, m)$-eket struktúrálisan kell vizsgálni. Először különböző optamilitás-tartó transzformációkkal egyszerűbbé, kezelhetőbbé tesszük a vizsgált kódot. Definiáljuk a $G_2 = (V_2, E_2)$ segédgráfot. $V_2 = S$ és $uv \in E_2$, ha létezik $t \in T_2$, hogy $\Gamma(t) = \{u, v\}$. Továbbá legyen $\Gamma(T_1) = V_1 \subseteq V_2$. Látható, hogy minden V_1-beli csúcsnak pontosan egy szomszédja van T_1-ben. Fogalmazzuk meg a Hall feltételt a G_2 gráfra.

(a) Nincs háromszoros él.

(b) Párhuzamos élek végpontja nem lehet V_1-ben.
(c) Párhuzamos élek végpontja nem lehet 1 távolságra V_1-beli csúcstól.
(d) Semelyik csúcsból nem indulhat ki két különböző párhuzuamos élpár.
(e) Háromszög egyik oldala sem lehet párhuzamos él.
(f) Két V_1-beli csúcs nem lehet szomszédos és szomszédságuk (a G_2 gráfban) diszjunkt.
(g) Háromszög egyik csúcsa sem lehet V_1-ben.

Ha ezek a feltételek mind teljesülnek, akkor a G_2 gráf által meghatározott T_1 és T_2 halmazuk önmagukban nem szegik meg a Hall feltételt. Ekkor a maradék $n - A_1 - A_2$ T-beli csúcshoz tartozó élek behúzásával batch kódot kaphatunk. (Például ha az összes többi T-beli csúcsból legalább 4 él indul ki.)

4.4. Állítás ([9]). *Adott G_2 segédgráfhoz tartozó optimális batch kód mérete csak $|V_1|$-től és $|E_2|$-től függ.*

Bizonyítás. **[9]** Minden $T \backslash T_1 \backslash T_2$-beli csúcs fokszáma 3 vagy 4 lesz. A cél az, hogy minél több csúcs foka 3 legyen. Az alábbi feltétel szükséges ahhoz hogy szabályos batch kódot kapjunk:
(h) Minden $S' \subseteq S$ 3-elemű halmazhoz legfeljebb $3 - |\{t \in T_1 : \Gamma(t) \in S'\}| - |\{t \in T_2 : \Gamma(t) \subseteq S'\}|$ darab T-beli csúcs tartozik, aminek a szomszédsága éppen S'.
Ez elégséges is: tegyük fel hogy (h) mindig teljesül, de $X \in T$ hiányos halmaz. Ha $X \in T_1 \cup T_2$, akkor már a G_2 nem teljesítette az (a)-(g) feltételek valamelyikét. Ha létezik $t \in T_3 \cap X$, akkor X nem lehet hiányos a $\Gamma(t)$-re vonatkozó (h) feltétel alapján.
Tehát a legjobb esetben

$$min\Big(n - A_1 - A_2, \sum_{\substack{S' \subseteq S \\ |S'| = 3}} (3 - |\{t \in T_1 : \Gamma(t) \in S'\}| - |\{t \in T_2 : \Gamma(t) \subseteq S'\}|)\Big)$$

darab T-beli csúcs fokszáma lesz 3 és a maradéké 4 lesz.

$$\begin{aligned} &3 - |\{t \in T_1 : \Gamma(t) \in S'\}| - |\{t \in T_2 : \Gamma(t) \subseteq S'\}| \\ &= \binom{m}{3} - \binom{m-2}{1} A_2 - \binom{m-1}{2} A_1, \end{aligned}$$

tehát csak $A_1 = |V_1|$-től és $A_2 = |E_2|$-től függ.

□

Tehát, ha adott egy optimális batch kódhoz tartozó segédrgáf, akkor ezen a gráfon végezhetünk olyan transzformációkat, amik nem rontják el az (a)-(g) feltételeket és nem változtatják $|E_2|$-t.

A Transzformáció: Legyen u_1v_1 és u_2v_2 két diszjunkt kétszeres él. Ha az $\{u_1, v_1\}$ és $\{u_2, v_2\}$ közötti élek közül legalább három be lenne húzva, akkor itt sérülne az (e) feltétel, tehát legfeljeebb két él lehet behúzva. Ez azt jelenti, hogy $(V_2, \overline{E_2})$-ben, a segédgráf komplementerében van $\{u_1, v_1\} - \{u_2, v_2\}$ párosítás. Ennek a párosításnak a két élét vegyük be E_2-be, u_1v_1 és u_2v_2 multiplicitását pedig csökkentsük egyre. Könnyen ellenőrizhető, hogy az (a)-(g) feltételek nem romolhattak el. A párhuzamos élek száma csökkent

B Transzformáció: Legyen uv az egyetlen párhuzamos él G_2-ben. Ha létezik uv-hez kapcsolódó csúcs (x), akkor feltehető, hogy $ux \in E_2$, de $vx \notin E_2$. Ekkor E_2-höz hozzávesszük vx-et és uv multiplicitását csökkentjük egyre. Ha nincs ilyen x, akkor tetszőleges ux élt veszünk hozzá x-hez és uv ismét egyre csökkentjük. Itt is ellenőrizhető, hogy az (a)-(g) feltételek megmaradnak. G_2-ben nem maradt párhuzamos él.

C Transzformáció: Tegyük fel, hogy G_2-ben nincs párhuzamos él. Legyen $x \in V_1$ foka legalább kettő, tehát $ux, vx \in V_1$. Ekkor (g) alapján $uv \notin V_1$. uv-t vegyük be E_2-be és ux-et hagyjuk ki. Az (f) feltétel miatt $v \notin V_1$, így a V_1-beli csúcsok összfokszáma csökkent.

4.5. Lemma ([9]). *Minden $n < \binom{m}{2}$-re létezik optomális $CBC(n, 4, m)$, amire G_2-ben nincs párhuzamos él, és minden V_1-beli csúcs foka legfeljebb egy.*

Bizonyítás. [9] Vegyünk tetszőleges optimális $CBC(n, 4, m)$-et és tekintsük a segédgráfját. Amíg több mint egy párhuzamos éltl tartalmaz ismételten végezzük el rajta az **A** transzformációt. Mivel minden lépésben csökken a párhuzamos élek száma, egy idő után nulla vagy egy párhuzamos él marad. Ha pontosan egy párhuzamos él van G_2-ben akkor elvégezhetjük a **B** transzformációt. Ekkor mindenképpen olyan G_2-t kapunk, amiben nincs párhuzamos él.

Amíg van olyan V_1-beli csúcs, aminek legalább kettő a fokszáma, végezzük el egy ilyen csúcson a **C** transzformációt. Mivel minden lépésben csökken a V_1-beli csúcsok összfokszáma, egy idő után minden ilyen csúcsnak a foka legfeljebb egy lesz.

Így a lemma feltételének megfelelő segédgráfot kaptunk. A 4.4 állítás alapján ehhez tartozik optimális $CBC(n, 4, m)$.

D Transzformáció: Legyen $t_1, t_2 \in T$ két csúcs, amelyekre $\Gamma(t_1) \subseteq \Gamma(t_2)$. Továbbá legyen $A \subseteq \Gamma(t_2) \backslash \Gamma(t_1)$. Ekkor az A és t_2 közötti élek lecserélhetők A és t_1 közötti élekre. Az (f) feltétel miatt $v \notin V_1$, így a V_1-beli csúcsok összfokszáma csökkent.

A transzformáció nyilvánvalóan megtartja a kód méretét. Bizonyítandó, hogy az így kapott gráf még mindig egy batch kódot ír le. Legyen X egy hiányos halmaz T-ben a transzformáció után. X-nek tartalmaznia kell t_1-et vagy t_2-t, különben a transzformáció előtt is hiányos lett volna. Ha X csak t_1-et tartalmazza, akkor $\Gamma(X)$ mérete nem csökkent a transzformáció során. Ha X tartalmazza t_1-et és t_2-t is,

akkor $\Gamma(X)$ nem változott. Mindkét eset ellentmond annak, hogy a transzformáció előtt szabályos volt a kód. Ha X csak t_2-t tartalmazza, akkor a transzformáció előtt $X - t_2 + t_1$ hiányos lett volna, ami szintén ellentmondás.

E Transzformáció: Tegyük fel hogy a batch kódhoz tartozó segédgráfra teljesülnek a 3.5 lemmában leírt tulajdonságok. Legyen $t \in T_3$, $\Gamma(t) = \{u, v, w\} \subseteq V_2 \backslash V_1$. Továbbá legyen $x \in V_1 \backslash \Gamma(T_3$. A (g) feltétel miatt az uvw háromszögnek nem lehet minden oldala E_2. Feltehető, hogy $vw \notin E_2$. A batch kódban töröljük ki a tu élt és helyettesítsük a tx éllel. Ez a transzformáció csak úgy ronthatná el a batch kódot, ha egy t-t tartalmazó X halmaz hiányossá válna, ahol $\Gamma(X) = \{v, w, x\}$. Tekintsük a (h) feltételt a $\{v, w, x\}$ halmazra. $|\{t \in T_1 : \Gamma(t) \in S'\}| = 1$ az (f) feltétel miatt és $|\{t \in T_2 : \Gamma(t) \subseteq S'\}| \leq 1$ a 3.5 lemma feltételei miatt, így X nem lehet hiányos.

4.6. Lemma ([9]). *Minden $n < \binom{m}{2}$-re létezik optimális $CBC(n, 4, m)$, amire G_2-re teljesülnek a 4.5 lemma feltételei és T-ben minden csúcs foka legfeljebb kettő.*

Bizonyítás. [9] Mivel $n < \binom{m}{2}$, létezik egyszerű gráf, m csúcson n éllel. Egy ilyen G_2 $V_1 = \emptyset$-zal egy szabályos batch kódot ad, aminek mérete $2n$. Tehát $N(n, 4, m) \leq 2n$ és egy optimális kódban T-ben az átlag fokszám legfeljebb kettő. Vegyünk egy optimális batch kódot, ami megfelel a 4.5 lemma feltételeinek.

Tegyük fel, hogy létezik 4-fokú csúcs T-ben (t_2). Ekkor léteznie kell egy egyfokú t_1 csúcsnak is. t_2 nem lehet benne hiányos halmazban, ezért feltehető, hogy $\Gamma(t_1) \subseteq \Gamma(t_2)$. Ekkor elvégezhetjük a **D** transzformációt erre a két csúcsra $|A| = 1$ mellett. Ezzel csökkent a 4-fokú csúcsok száma. Ezt ismételjük amíg nem marad ilyen T-ben.

Tegyük fel, hogy létezik 3-fokú csúcs T-ben. Ekkor léteznie kell egy egyfokú csúcsnak is. Ha nincs olyan 3-fokú csúcs, aminek a szomszédsága belemetsz V_1-be, akkor vegyünk tetszőleges $t \in T$ 3-fokú csúcsot és $x \in V_1$-et. Ezekre teljesülnek az **E** transzformáció feltételei, így elvégezhetjük azt. Ha létezik $t \in T_3$ és $y \in T_1$, amelyekre $\Gamma(y) \in \Gamma(t)$, akkor ezekre elvégezhetjük a **D** transzformációt $|A| = 1$ mellett. Ezzel csökkent a 3-fokú csúcsok száma. Ezt ismételjük, amíg T-ben minden csúcs foka legfeljebb kettő.

□

[9] Mivel feltehető, hogy T-ben minden csúcs fokszáma egy vagy kettő, innentől kezdve elég G_2-t tekinteni, amiben feltehtő, hogy a 4.5 lemma feltételei teljesülnek:

(i) Nincsenek párhuzamos élek.

(j) Minden V_1-beli él foka legfeljebb egy.

A kód mérete $2|E_2| + |V_1| = 2A_2 + A_1 = 2n - A_1$. A kód akkor optimális, ha $|V_1|$ minimális. Ha $|V_1|$ adott, akkor megbcsülhetjük $|E_2|$-t. $V_2 \backslash V_1$-en belül legfeljebb $\binom{m-A_1}{2}$ él lehet az (i) feltétel miatt. A V_1 és V_2 közötti élek párosítást alkotnak a (j) és (f) feltételek alapján. Emiatt itt legfeljebb $min(A_1, m - A_1)$ él lehet. V_1-en belül

pedig nem lehet él szintén az (f) feltétel miatt. Tehát

$$|E_2| \leq \binom{m - A_1}{2} + min(A_1, m - A_1).$$

A becslés éles, ami a bizonyításából könnyen látszik. Ha $|E_2|$-re teljesül a becslés, akkor szabályos G_2 gráf konstruálható: Válasszuk ki V_2-ből tetszőlegesen a V_1 halmazt, majd húzzunk be megfelelő mennyiségű élt úgy, hogy csak $V_2 \backslash V_1$-en belüli éleket és egy $V_2 \backslash V_1 - V_1$ párosítás éleit használjuk. Könnyen ellenőrizhető, hogy ez szabályos segédgráf lesz.

$$n = A_1 + A_2$$
$$n \leq \binom{m - A_1}{2} + min(A_1, m - A_1) + A_1$$

4.7. Állítás ([9]). *Ha $n < \binom{m}{2}$, akkor $N(n, 4, m) = 2n - x$, ahol x a legnagyobb olyan egészszám 0 és m között, amire*

$$0 \leq \binom{m - x}{2} + min(x, m - x) + x - n.$$

[9] Ezután $N(n, 4, m)$ megtalálása már csak számolás kérdése.

Először keressünk ilyen x-et a $[\frac{m}{2}, m]$ intervallumban. Ekkor a feltétel:

$$0 \leq \binom{m - x}{2} + m - n$$
$$0 \leq m^2 - 2xm + x^2 + x + m - 2n$$
$$0 \leq x^2 + (1 - 2m) + (m^2 + m - 2n)$$
$$x_{12} = \frac{2m - 1 \pm \sqrt{(2m-1)^2 - 4(m^2 + m - 2n)}}{2}$$
$$x_{12} = \frac{2m - 1 \pm \sqrt{8n - 8m + 1}}{2}$$

Ha $n = m$, akkor $x_2 = m$ így ez az x érték megfelelő és optimális. Nagyobb n-ekre azonban $x_2 > m$, ezért $x \leq x_1$-nek kell teljesülnie. Az optimális x tehát

$$\left\lfloor \frac{2m - 1 - \sqrt{8n - 8m + 1}}{2} \right\rfloor,$$

de csak akkor, ha ez benne van az $[\frac{m}{2}, m]$ intervallumban. Esetszétválasztás m paritása szerint. Ha m páros, a következő kell.

$$\frac{m}{2} \leq \frac{2m - 1 - \sqrt{8n - 8m + 1}}{2}$$

$$\sqrt{8n-8m+1} \le m-1$$
$$8n-8m+1 \le m^2-2m+1$$
$$n \le \frac{m^2+6m}{8}$$

Ha m páratlan, akkor a következő kell.

$$\frac{m+1}{2} \le \frac{2m-1-\sqrt{8n-8m+1}}{2}$$
$$\sqrt{8n-8m+1} \le m-2$$
$$8n-8m+1 \le m^2-4m+4$$
$$n \le \frac{m^2+4m+3}{8}$$

Ha ezek a feltételek nem teljesülnek, akkor nincs megfelelő x az $[\frac{m}{2}, m]$ intervallumban, ezért a $[0, \frac{m}{2}]$ intervallumban kell keresni. Innentől kezdve feltehetjük, hogy $n \ge \frac{m^2+6m+8}{8}$, ha m páros, és $n \ge \frac{m^2+4m+11}{8}$, ha m páratlan. Ekkor a feltétel:

$$0 \le \binom{m-x}{2} + 2x - n$$
$$0 \le m^2 - 2mx + x^2 + 5x - m - 2n$$
$$0 \le x^2 + (5-2m)x + (m^2 - m - 2n)$$
$$x_{12} = \frac{2m-5 \pm \sqrt{(2m-5)^2 - 4(m^2-m-2n)}}{2}$$
$$x_{12} = \frac{2m-5 \pm \sqrt{8n-16m+25}}{2}$$

Itt, ha m páros

$$x_2 \ge \frac{2m-5+\sqrt{m^2+6m+8-16m+25}}{2} = \frac{2m-5+\sqrt{(m-5)^2+8}}{2}$$
$$\ge \frac{2m-5+3}{2} > \frac{m}{2}.$$

Ha pedig páratlan

$$x_2 \ge \frac{2m-5+\sqrt{m^2+4m+11-16m+25}}{2} = \frac{2m-5+|m-6|}{2}$$
$$\ge \frac{2m-5+1}{2} > \frac{m-1}{2}.$$

Tehát az optimális x

$$\left\lfloor \frac{2m-5-\sqrt{8n-16m+25}}{2} \right\rfloor,$$

ha ez benne van a $[0, \frac{m}{2}]$ intervallumban.

$$\frac{2m-5-\sqrt{8n-16m+25}}{2} \geq \frac{2m-5-\sqrt{8\binom{m}{2}-16m+25}}{2}$$
$$= \frac{2m-5-\sqrt{4m^2-4m-16m+25}}{2} = \frac{2m-5-|2m-5|}{2} = 0$$

Ha m páros a következő kell.

$$\frac{m}{2} \geq \frac{2m-5-\sqrt{8n-16m+25}}{2}$$
$$\sqrt{8n-16m+25} \geq m-5$$
$$8n-16m+25 \geq m^2-10m+25$$
$$n \geq \frac{m^2+6m}{8},$$

ami mindig teljesül. Ha m páratlan, a következő kell.

$$\frac{m+1}{2} > \frac{2m-5-\sqrt{8n-16m+25}}{2}$$
$$\sqrt{8n-16m+25} > m-6$$
$$8n-16m+25 > m^2-12m+36$$
$$n > \frac{m^2+4m+11}{8},$$

ami teljsesül egy eset kivételével: ha $n = \frac{m^2+4m+11}{2}$. Ebben a kivételes esetben minden $x \in [0, \frac{m-1}{2}]$ jó, ezért az optimális x $\frac{m-1}{2}$.

Minden esetet megvizsgáltunk, így kimondhatjuk az $N(n,4,m)$-re vonatkozó tételt.

4.8. Tétel ([9]).

$$N(n,4,m) = \begin{cases} n & ha \quad n = m \\ 2n-m+\left\lceil \frac{1+\sqrt{8n-8m+1}}{2} \right\rceil & ha \quad \begin{cases} m < n \leq \frac{m^2+6m}{8}, & 2|m \\ m < n \leq \frac{m^2+4m+3}{8}, & 2 \not| m \end{cases} \\ 2n-\frac{m-1}{2} & ha \quad n = \frac{m^2+4m+11}{8} \\ 2n-m+\left\lceil \frac{5+\sqrt{8n-16m+25}}{2} \right\rceil & ha \quad \begin{cases} \frac{m^2+6m}{8} < n < \binom{m}{2}, & 2|m \\ \frac{m^2+4m+11}{8} < n < \binom{m}{2}, & 2 \not| m \end{cases} \\ 3n-\left\lfloor \frac{m^2}{2}-\frac{n-m}{m-3} \right\rfloor & ha \quad \binom{m}{2} \leq n \leq 3\binom{m}{3} \\ 4n-3\binom{m}{3} & ha \quad 3\binom{m}{3} \leq n \end{cases}$$

Látható, hogy az előző esetekkel ellentétben a $k = 4$ már nagyon monyolult és n sok intervallumát kell külön-külön vizsgálni. A $k = 5$ és nagyobb esetekre nem ismert N értéke.

5. A $k = m$ eset

Az alábbi rövid fejezetben a $k = m$ esetet tárgyaljuk. Ez egyedülálló azon esetek között, ahol k megközelíti m-et. Itt N optimális értéke könnyen megadható, de $k \leq m - 1$ esetén nem ismert általánosan az optimális batch kód mérete egészen $k = 4$-ig, amit a 4 fejezetben tárgyalunk.

5.1. Tétel. $N(n, k, k) = kn - k^2 + k$

Bizonyítás. Az S-re vonatkozó Hall feltételek szerint minden $Y \subseteq S$-re $|\Gamma(Y)| \geq |Y| + n - m$. Emiatt minden $s \in S$-re $d(s) \geq n - m + 1$ és $N \geq m(n - m + 1)$. Ez az érték el is érhető a következő konstrukcióval.

Legyen $T_1 \subseteq T$, $|T_1| = m$ és $T' = T \backslash T_1$. S és T_1 között vegyünk teljes párosítást, majd S és T' között vegyünk teljes párosgráfot. Az így kapott batch kód szabályosságát az S-re vonatkozó Hall feltételek ellenőrzésével bizonyítjuk. Tetszőleges $Y \subseteq S$ esetén Y T_1-beli szomszédságának mérete $|Y|$ és $T' \subseteq \Gamma(Y)$ ezért Y teljes szomszédságának mérete pontosan $|Y| + n - m$. A kód mérete $m + m(n - m) = mn - m^2 + m$.

□

A $k = m - 1$ egyes speciális eseteiről még esik szó a 9 fejezetben.

6. Optimális kódok kis n esetén

Ebben a fejezetben azt a speciális esetet tárgyaljuk, ahol m relative kicsi m-hez képest. Itt az $n = m, m+1, m+2$ esetek megoldásai ismertek. Mellékesen bevezetjük a batch kódhoz tartozó $D = (V, A)$ irányított gráfot, ami egy általánosságban is hasznos segédeszköz.

Az $n = m$ eset könnyen kezelhető. Ebben az esetben az (T, S) teljes párosítás megfelelő minden k-ra és ez optimális is, mivel minden k esetén minden T-beli csúcs fokának legalább egynek kell lennie.

6.1. Tétel ([14]). $N(m, k, m) = m$

Az $n = m + 1$ esetben kezdjünk egy kézenfekvő konstrukcióval.

6.2. Konstrukció ([14]). *Válasszunk ki egy tetszőleges t csúcsot T-ből. Vegyünk egy $(T - t, S)$ teljes párosítást, majd t-t kössük össze S tetszőleges k csúcsával. Látható, hogy T semmilyen X halmaza nem szegheti meg a Hall feltételt, mert, ha $t \in X$, akkor $\Gamma(X) \geq k$, de ha $t \notin X$, akkor X benne van a párosításban.*

Kiderül, hogy a fenti konstrukció optimális minden k esetén.

6.3. Tétel ([14]). $N(m+1,k,m) = m+k$

Bizonyítás. [14] Elég belátni, hogy nem létezik $m+k$-nál kisebb kód. Tegyük fel, hogy létezik egy $CBC(m+1, m+j, k, m)$ batch kód, ahol $j < k$. Rendezzük fokszám szerinti sorrendbe S csúcsait: $(s_1, s_2, ..., s_m)$, ahol $d(s_1) \geq d(s_2) \geq ... \geq d(s_m)$ Ekkor

$$\sum_{s \in S} d(s) = N = m + j,$$

tehát az $A = \{s_1, ..., s_j\}$ halmazból kiinduló élek száma legalább $2j$, míg a $B = \{s_j+1, ..., s_m\}$ halmazból kiinduló élek száma legfeljebb $m-j$. Legyen $\Gamma(B) = C \subseteq T$ és $T \backslash C = D$.

$$|C| \leq \sum_{s \in B} d(s) \leq m - j$$

$$|D| = n - |C| \geq j + 1$$

Azonban $\Gamma(D) = A$, aminek a mérete csak j, tehát D-nek tetszőleges $j+1 \leq k$ méretű részhalmaza megszegi a Hall feltételt, ami ellentmondás.

□

Bár ez a bizonyítás rövid és elég egyszerű, érdemes megemlíteni egy másik gondolatmenetet, amiből jobban meg lehet érteni az ilyen batch kódok struktúráját. Először belátunk egy egyszerű, de általános esetben is hasznos állítást.

6.4. Állítás ([8]). *Ha $k > 1$, optimális batch kódban mindig létezik S-et fedő párosítás.*

Bizonyítás. [8] Tegyük fel, hogy S nem fedhető párosítással. Ekkor létezik $A \subseteq S$ Hall-hiányos halmaz, azaz $|B| = |\Gamma(A)| < |A|$. Legyen $C = T \backslash B$ és $D = S \backslash A$. A és C között nem fut él ezért ha a gráfot leszűkítjük a $C \cup D$ csúcshalmazra szabályos batch kód marad. Ahhoz tehát, hogy az eredeti batch kód szabályos legyen elég, hogy legyen benne egy B-t fedő, B és A közötti párosítás. Mivel feltettük, hogy a batch kód optimális, A és B között csak egy ilyen párosítás élei lehetnek behúzva.

$|A| > |B|$, ezért A egy s csúcsa izolált. Ekkor T bármely t csúcsára az összes t-ből induló élt helyethesíthetnénk egyetlen ts éllel. Ezzel a kód mérete csökken, ellentmondva az optimalitásnak, kivéve, ha $d(t) \leq 1$, tehát T minden csúcsának fokszáma legfeljebb egy. Ez azonban $k > 1$ mellett csak akkor lehetséges, ha a gráf egy (S, T) teljes párosítás, ami ellentmond a kezdeti feltevésnek.

□

[8] Visszatérve az $n = m + 1$ esetre, rendezzük sorba az S és T csúcshalmazokat: $(s_1, s_2, ..., s_m)$ és $(t_1, t_2, ..., t_{m+1})$. Minden, az s_1t_1, s_2t_2, s_mt_m párosítást tartalmazó batch kódhoz hozzárendelhető egy $D = (T, A)$ irányított gráf a következőképpen. A t_it_j $(i \neq j)$ irányított él pontosan akkor van benne A-ban, ha a batch kódban benne van a t_is_j él. Az előző állítás alapján már tudjuk, hogy van ilyen párosítást tartalmazó optimális batch kód is. Jellemezzük a T csúcshalmazon azokat az irányított gráfokat, amik egy szabályos batch kódhoz tartoznak.

(Megjegyzendő, hogy egy batch kódhoz sok különböző irányított gráfot is rendelhetünk, attól függően, hogy melyik párosítást, vagy S és T elemeinek melyik sorrendjeit választottuk ki. Rögzített párosítás mellett azonban már minden batch kódhoz egyértelműen rendelünk egy-egy különböző irányított gráfot. Az így megkapható irányított gráfok halmazát szeretnénk jellemezni.)

6.5. Állítás ([8]). *$k > 1$. Egy $D = (T, A)$ irányított gráf pontosan akkor tartozik batch kódhoz, ha t_{m+1} befoka nulla és t_{m+1}-ből legalább k darab másik csúcs elérhető irányított úton.*

Bizonyítás. Nyilvánvaló, hogy D-ben $t_m + 1$-be nem futhat be él, hiszen S-ben nincs s_{m+1} csúcs. Bizonyítandó, hogy t_{m+1}-ből elérhető k másik csúcs. Tegyük fel, hogy t_{m+1}-ből csak k-nál kevesebb csúcs érhető el irányított úton D-ben. Legyen a t_{m+1}-ből elérhető csúcsok halmaza $X \subseteq T$, $|X| = l + 1 \leq k$. Ha $X = \{t_{i_1}, t_{i_2}, ..., t_{i_l}, t_{m+1}\}$, akkor $\Gamma(X)$ tartalmazza az $s_{i_1}, s_{i_2}, ..., s_{i_l}$ pontokat, de ezeken kívül mást nem, mert ez D-ben egy X-ből kimutató élt jelentene. $k \geq |X| > |\Gamma(X)|$, tehát X hiányos halmaz lenne; D-hez nem tartozik batch kód.

Tegyük fel, hogy t_{m+1}-ből legalább k másik csúcs elérhető, de a D-hez tartozó batch kód nem szabályos. Ekkor létezik egy $X \subseteq T$ hiányos halmaz. Ha $t_{m+1} \notin X$ azonnal ellentmondásba ütközünk, mert X-et fedi a kiválasztott párosítás. Ha $t_{m+1} \in X$, akkor X nem tartalmazhat minden t_{m+1}-ből elérhető csúcsot (mert legfeljebb k méretű), ezért létezik X-ből kilépő él D-ben. Legyen egy ilyen élnek a végpontja t_{i_0}. Legyen továbbá $X = \{t_{i_1}, t_{i_2}, ..., t_{i_l}, t_{m+1}\}$, (ahol $|X| = l + 1 \leq k$). Ekkor $s_{i_1}, s_{i_2}, ..., s_{i_l}, s_{i_0} \in \Gamma(X)$, tehát $|\Gamma(X)| \geq |X|$ és X nem lehet hiányos halmaz.

□

Ebből azonnal látszik, hogy pontosan a t_{m+1}-gyökerű, k-élű fenyőkhöz tartozó batch kódok optimálisak. (Fenyőnek nevezzük az olyan irányított fákat, amikben egy adott gyökércsúcsból minden más csúcs elérhető irányított úton.) Ebből adódik az 6.3 tétel. (A kihagyott $k = 1$ esetben a 4.1 tételből adódik az 6.3 tétel.) A $D = (T, A)$ irányított gráf bevezetése hasznos ötlet általában is, a $k = m$ esetben például kihozható a 5.1 tétel a D gráf vizsgálatával. A módszer alkalmazható $n = m + 2$ és nagyobb n-ek esetén is. Ezekben az esetekben is jellemezhető a batch kódokhoz tartozó irányított gráfok struktúrája, azonban az optimális batch kód méretét már nem

lehet ilyen könnyen kiolvasni.

Az $n = m+2$ esethez a 4. fejezethez hasonlóan optimalitás tartó transzformációkat kell alkalmazni, hogy a vizsgált kód struktúrája egyszerűbbé váljon. S csúcsait fokszám alapján osztályozzuk. Legyen S_1, S_2 és S_3 S-ben rendre az egy, kettő és háromnál nagyobb fokszámú csúcsok halmaza. A $k = 1$ esettől eltekintünk, mivel ezt már a 4 fejezetben vizsgáltuk. Ekkor az 6.4 állítás alapján minden S-beli csúcsból indul ki él, így $S = S_1 \cup S_2 \cup S_3$.

6.6. Lemma ([10]). *$k > 1$. Ha optimális $CBC(m+2, k, m)$-ben $s \in S_1$, akkor $\Gamma(s) \in T_1$.*

Bizonyítás. [10] Ha ez nem teljesül, $t = \Gamma(s)$ összes ts-től különböző élét kitörölhetnénk, ezzel csökkentve a kód méretét.

□

6.7. Becslés ([10]). *$k > 1$, $m > k$. Ha egy optimális $CBC(m+2, k, m)$-ben $S_1 \neq \emptyset$, akkor $N(m+2, k, m) \geq N(m+1, k, m-1) + 1$.*

Bizonyítás. [10] Vegyünk egy $s \in S_1$ csúcsot. Az előző állítás alapján ennek egyetlen szomszédja $t \in T_1$. Tehát $\{s, t\}$ a gráf egy összefüggőségi komponenese, amit elhagyva egy $CBC(m+1, k, m-1)$-et kapunk. Ennek élszáma nagyobb, mint $N(m+1, k, m-1)$, tehát a kód teljes mérete legalább $N(m+1, k, m-1) + 1$.

□

A Transzformáció: Legyen $s_1 \in S_2$, $\Gamma(s_1) = \{t_1, t_2\}$ és $s_2 \in \Gamma(t_2)$, ahol $s_1 \neq s_2$. Ekkor a batch kódban helyettesíthetjük t_2s_2-t t_1s_2-vel. Ezzel a transzformációval csak t_2-t tartalmazó X hiányos halmazt hozhatunk létre. Ha $t_1 \in X$, akkor $\Gamma(X)$ nem változott a transzformáció során. Ha viszont $t_1 \notin X$, akkor $s_1 \notin \Gamma(X - t_1)$, ezért $X - t_1$ a transzformáció előtt hiányos lett volna, ami ellentmondás.

□

(Megjegyzendő, hogy nem kell feltételezni, hogy t_1s_2 nem volt eredetileg a kódban. Ha benne volt, akkor a transzformáció során csak elhagyjuk a t_2s_2 élt, ezzel csökkentve a kód méretét. Szintén fontos észrevétel, hogy a transzformáció semmilyen S-beli csúcs fokszámát nem változtatja.)

6.8. Becslés ([10]). *$k > 1$. Ha egy optimális $CBC(m+2, k, m)$-ben $S_1 = \emptyset$, akkor*

$$N(m+2, k, m) \geq 2m + \left\lfloor \frac{k}{m-k+1} \right\rfloor.$$

Bizonyítás. [10] Először az A transzformáció segítségével elérjük, hogy minden S_2-beli csúcsnak legyen T_1-beli szomszédja. Ha létezik $s \in S_2$, amire $\Gamma(s) \cap T_1 = \emptyset$. Ekkor az textbf A transzformációval csökkenthető $\Gamma(t_1)$ mérete. Ezt addig ismételhetjük, amíg $\Gamma(t_1)$ már csak s-et tartalmazza. Ezzel egyel csökkent azon S_2-beli csúcsok száma, amelyek szomszédsága nem metszi T_1-et. Ezt megismételve minden ilyen csúcsra olyan optimális CBC-t kapunk, amiben minden S_2-beli csúcsnak van T_1-beli szomszédja és $S_1 = \emptyset$ még mindig teljesül.

Ebből azonnal látszik, hogy $|T_1| \geq |S_2|$. Tegyük fel, hogy egy $t \in T \backslash T_1$-re $|\Gamma(t) \cap S_2| > m - k$. Ekkor $Y = \Gamma(t) \cap S_2$ szomszádságának mérete csak egyel nagyobb, mint Y mérete, ami ellentmond az Y-ra vonatkozó Hall feltételnek. Minden S_2-beli csúcsnak pontosan egy $T \backslash T_1$-beli szomszédja van és minden $T \backslash T_1$-beli csúcsnak legfeljebb $m-k$ S_2-beli szomszédja van, ezért $(m-k)|T \backslash T_1| \geq |S_2|$. Ebből levezethető az N-re vonatkozó becslés.

$$n = |T_1| + |T \backslash T_1| \geq |S_2| + \frac{|S_2|}{m-k} = \frac{m-k+1}{m-k} \cdot |S_2|$$

$$|S_3| = m - |S_2| = m - \frac{m-k}{m-k+1} \cdot n = \frac{m(m-k+1) - (m-k)(m+2)}{m-k+1} = \frac{2k-m}{m-k+1}$$

$$N \geq 2|S_2| + 3|S_3| \geq 2m + |S_3| \geq 2m + \left\lceil \frac{2k-m}{m-k+1} \right\rceil = 2m + \left\lfloor \frac{k}{m-k+1} \right\rfloor$$

□

Két konstrukcióval megmutatjuk, hogy az 6.7 és 6.8 becslés is teljesíthető egyenlőséggel. Mivel a két becslés közül az egyik mindenképpen teljesül $N(m+2, k, m)$-re attól függően, hogy az optimális kódban S_1 üres-e, ezért $N(m+2, k, m)$ értéke mindig éppen a két becslés minimuma lesz. Az 6.7 becslés könnyen teljesíthető egyenlőséggel.

6.9. Konstrukció ([10]). $k > 1$. *Vegyünk egy optimális $CBC(m+1, k, m-1)$-t és adjunk hozzá S-hez és T-hez is egy új csúcsot, majd az új csúcsokat kössük össze. Könnyen látható hogy ez szabályos batch kódot alkot, melynek mérete $N(m+1, k, m-1)+1$.*

6.10. Konstrukció ([10]). $k > 1$. *Az 6.8 becslés bizonyításából látható, a becslés egyenlőséggel való teljesüléséhez kell $|T_1| = |S_1|$, $|T \backslash T_1| \approx \frac{|S_2|}{m-k}$, valamint hogy S_3-ban minden csúcs foka 3 legyen. Osszuk S-et és T-t két-két részre. $T = T_1 \cup T'$ és $S = S_2 \cup S_3$, ahol*

$$|T_1| = n \cdot \left\lfloor \frac{m-k}{m-k+1} \right\rfloor$$

$$|T'| = \left\lceil \frac{n}{m-k+1} \right\rceil$$

$$|S_2| = |T_1|$$

$$|S_3| = |T'| - 2.$$

Az élek a következőek: T_1 és S_2 között vegyünk be egy teljes párosítást. T' és S_3 között vegyünk be egy optimális $CBC\,(|T'|, |T'|-2, |T'|-2)$-t. A 4.1 tételből tudjuk, hogy ennek mérete $N(|T'|, |T'|-2, |T'|-2) = 3(|T'|-2)$ és lehet minden S_3-beli csúcs foka 3. Ezután S_2 minden csúcsát kössük össze egy T'-beli csúccsal úgy, hogy minden T'-beli csúcsnak legfeljebb $\frac{1}{m-k}$ S_2-beli szomszédja legyen. Ez megtehető, mert

$$(m-k)\cdot|T'| \geq (m-k)\frac{n}{m-k+1} \geq |S_2|.$$

A batch kód szabályosságának bizonyításához az S-re vonatkozó Hall feltételt fogjuk ellenőrizni. Tegyük fel, hogy $Y \subseteq S$ hiányos halmaz. Legyen $Y_2 = Y \cap S_2$ és $Y_3 = Y \cap S_3$. Ha $Y_3 \neq \emptyset$, akkor Y_3 T'-beli szomszédságának mérete legalább $|Y_3|+2$ a T' és S_3 közötti batch kódra vonatkozó Hall feltétel miatt. Tobábbá Y_2 T_1-beli szomszédságának mérete legalább $|Y_2|$, mert a kód tartalmaz teljes (Y_2, T_1) párosítást. Tehát $|\Gamma(Y)| \geq |Y_3|+2+|Y_2| = |Y|+2$, így Y nem hiányos. Ha $Y_3 = \emptyset$, akkor $|Y_2| > m-k$, ezért van legalább kettő T'-beli szomszédja. Az előző esethez hasonlóan Y_2 T_1-beli szomszédsága legalább $|Y_2| = |Y|$. $|\Gamma(Y)| \geq |Y|+2$, így Y ismét nem hiányos.

A kód mérete

$$2\cdot|S_2| + 3\cdot(|T'|-2) = 2\cdot|T_1| + 3\cdot(m-|T_1|) = 2m + (|T'|-2)$$

$$= 2m + \left\lceil \frac{m+2-2(m-k+1)}{m-k+1} \right\rceil = 2m + \left\lfloor \frac{2k-m}{m-k+1} \right\rfloor = 2m + \left\lfloor \frac{k}{m-k+1} \right\rfloor.$$

□

Ezzel beláttuk, hogy ha $k > 1$, akkor $N(m+2,k,m) = min(N(m+1,k,m-1)+1, 2m = \lfloor \frac{k}{m-k+1} \rfloor)$. Innen egyszerű számolással megkapható $N(m+2,k,m)$ explicit képlete.

6.11. Tétel ([9]).

$$N(m+2,k,m) = \begin{cases} 2m + \lfloor \frac{k}{m-k+1} \rfloor & ha \quad k \leq m \leq k+\sqrt{k} \\ m+k-2+\lceil 2\sqrt{k+1} \rceil & ha \quad m > k+\sqrt{k} \end{cases}$$

Bizonyítás. **[9]** A $k=1$ esetben a 4.1 tétel alapján ellenőrizhető, hogy valóban helyes a képlet. Ha $m=1$, $2m + \lfloor \frac{k}{m-k+1} \rfloor = 3 = N(3,1,1)$. Ha $m=2$, $2m + \lfloor \frac{k}{m-k+1} \rfloor = 4 = N(4,1,1)$. Végül, ha $m > 2 = k+\sqrt{k}$, $m+k-2+\lceil 2\sqrt{k+1} \rceil = m+2 = N(m+2,1,m)$.

Innentől feltehető, hogy $k > 1$. m-re vonatkozó teljes indukcióval bizonyítunk. Ha $m = k$, akkor a 5.1 tétel alapján ellenőrizhető a képlet: $2m + \lfloor \frac{k}{m-k+1} \rfloor = 3m = N(m+2,m,m)$.

Az indukciós lépéshez ellőször vizsgáljuk meg a $\left\lfloor \frac{k}{l} \right\rfloor - \left\lfloor \frac{k}{l+1} \right\rfloor$ különbséget, ahol $l > 0$ egészszám.

$$\frac{k}{l} - \frac{k}{l+1} - 1 = \frac{k - l(l+1)}{l(l+1)}$$

Tehát, ha $k \geq l(l+1)$, akkor $\frac{k}{l} - \frac{k}{l+1} \geq 1$ és $\left\lfloor \frac{k}{l} \right\rfloor - \left\lfloor \frac{k}{l+1} \right\rfloor \geq 1$. Ha viszont $k \leq l(l+1)$, akkor $\frac{k}{l} - \frac{k}{l+1} \leq 1$ és $\left\lfloor \frac{k}{l} \right\rfloor - \left\lfloor \frac{k}{l+1} \right\rfloor \leq 1$. Ennek segítségével becsülhetjük a $\left\lfloor \frac{k}{m-k} \right\rfloor - \left\lfloor \frac{k}{m-k+1} \right\rfloor$ különbséget. Ha $m \leq k + \sqrt{k} - 1$, akkor

$$(m-k)(m-k+1) \leq (\sqrt{k}-1)(\sqrt{k}) \leq k$$

$$\left\lfloor \frac{k}{m-k} \right\rfloor - \left\lfloor \frac{k}{m-k+1} \right\rfloor \geq 1.$$

Ha viszont $m \geq k + \sqrt{k}$, akkor

$$(m-k)(m-k+1) \geq (\sqrt{k})(\sqrt{k}+1) \geq k$$

$$\left\lfloor \frac{k}{m-k} \right\rfloor - \left\lfloor \frac{k}{m-k+1} \right\rfloor \leq 1.$$

Visszatérve az indukciós lépéshez, az első eset, ha $m \leq k + \sqrt{k} - 1$.

$$2m + \left\lfloor \frac{k}{m-k+1} \right\rfloor \leq 2(m-1) + \left\lfloor \frac{k}{m-k} \right\rfloor + 1 = N(m+1, k, m-1) + 1$$

Tehát $N(m+2, k, m) = 2m + \left\lfloor \frac{k}{m-k+1} \right\rfloor$.

A második eset, ha $k + \sqrt{k} - 1 < m \leq k + \sqrt{k}$.

$$(m-k+1)(m-k-1) \leq (\sqrt{k}+1)(\sqrt{k}-1) = k - 1 < k$$

$$\frac{k}{m-k+1} \geq m-k-1$$

$$\left\lfloor \frac{k}{m-k+1} \right\rfloor \geq m-k-1$$

$$(m-k)^2 > (\sqrt{k}-1)^2 > k$$

$$\frac{k}{m-k} < m-k$$

$$\left\lfloor \frac{k}{m-k} \right\rfloor \leq m-k$$

$$\Longrightarrow \left\lfloor \frac{k}{m-k} \right\rfloor - \left\lfloor \frac{k}{m-k+1} \right\rfloor \geq 1$$

Így az első esethez hasonlóan belátható, hogy $N(m+2,k,m) = 2m + \left\lfloor \frac{k}{m-k+1} \right\rfloor$.

A harmadik eset, ha $k + \sqrt{k} < m \leq k + \sqrt{k} + 1$, azaz $m = k + \lfloor \sqrt{k} \rfloor$.

$$2m + \left\lfloor \frac{k}{m-k+1} \right\rfloor \geq 2(m-1) + \left\lfloor \frac{k}{m-k} \right\rfloor + 1 \geq N(m+1,k,m) + 1$$

Ezért:

$$N(m+2,k,m) = N(m+1,k,m-1) + 1 = 2(m-1) + \left\lfloor \frac{k}{m-k} \right\rfloor + 1$$

$$= 2k + 2\lfloor \sqrt{k} \rfloor + 1 + \left\lfloor \frac{k}{\lfloor \sqrt{k} \rfloor + 1} \right\rfloor$$

Bizonyítandó, hogy ez egyenlő $m+k-2+\lceil 2\sqrt{k+1} \rceil$-el. Legyen a pozitív egészszám, amire $a^2 \leq k < (a+1)^2$.

$$2k + 2\lfloor \sqrt{k} \rfloor + 1 + \left\lfloor \frac{k}{\lfloor \sqrt{k} \rfloor + 1} \right\rfloor = 2k + 2a + 1 + \left\lfloor \frac{k}{a+1} \right\rfloor$$

Ha $a^2 \leq k < a^2 + a$, akkor $\left\lfloor \frac{k}{a+1} \right\rfloor = a - 1$ és

$$a^2 + 1 \leq k + 1 \leq a^2 + a$$

$$a < \sqrt{k+1} < a + \frac{1}{2}$$

$$\lceil 2\sqrt{k+1} \rceil = 2a + 1.$$

Ha $a^2 + 1 \leq k < (a+1)^2$, akkor $\left\lfloor \frac{k}{a+1} \right\rfloor = a$ és

$$a^2 + a < k + 1 \leq (a+1)^2$$

$$a + \frac{1}{2} < \sqrt{k+1} \leq a + 1$$

$$\lceil 2\sqrt{k+1} \rceil = 2a + 2.$$

Mindkét esetben

$$2k + 2\lfloor \sqrt{k} \rfloor + 1 + \left\lfloor \frac{k}{\lfloor \sqrt{k} \rfloor + 1} \right\rfloor = 2k + \lfloor \sqrt{k} \rfloor - 1 + \lceil 2\sqrt{k+1} \rceil = m + k - 2 + \lceil 2\sqrt{k+1} \rceil,$$

tehát $N(m+2,k,m) = m + k - 2 + \lceil 2\sqrt{k+1} \rceil$.

A negyedik eset, ha $m > k + \sqrt{k} + 1$. A harmadik esethez hasonlóan belátható, hogy $N(m+2,k,m) = N(m+1,k,m-1)+1$. Tehát $N(m+2,k,m) = m+k-2+\lceil 2\sqrt{k+1} \rceil$.

□

Az $n > m+2$ nem ismert az optimális batch kód mérete általános esetben, kivéve n a 3. fejezetben tárgyalt nagy értékeire.

7. Transzverzális matroidok

[8] A kombinatorikus batch kódok optimalizálásának egy érdekes módszere a batch kódhoz tartozó transzverzális matroid vizsgálata. Az (S,T) párosgráfhoz tartozó transzverzális matroid T alaphalmazon értelmezett martoid, amiben a párosítással lefedhető részhalmazok a függetlenek. A Hall feltétel alapján a transzverzális matroidból megállapítható, hogy a batch kód szabályos-e: pontosan akkor szabályos, ha minden legfeljebb k méretű részhalmaz független.

Bár minden párosgráf egyértelműen meghatározza transzverzális matroidját, ez fordítva nem igaz. Az adott M transzverzális matroidhoz tartozó párosgráfokat M prezentációinak hívjuk, Az 6.4 állítás alapján feltehető, hogy M alaphalmazának mérete n és rangja m, így a prezentációkat értelmezhetjük az $S \cup T$ csúcshalmazon. (A $k = 1$ esettől eltekintünk, ezt a 4 fejezetben kimerítettük.)

M maximális prezentációja S permutációja erejéig egyértelmű, legyen ez a $G_M = (S, T; E_M)$ gráf. Tehát M bármely $G = (S, T; E)$ prezentációjánál, S csúcsainak megfelelő permutációja esetén, $E \subseteq E_M$. M minimális prezentációja nem egyértelmű S permutációja erejéig sem, de az élszáma egyértelmű, és meghatározható a G_M maximális prezentációból. Pontosan, ha $G_m = (S, T; E_m)$ minimális prezentáció, akkor

$$|E_m| = \sum_{i=1}^{m} (d_M)(s_i) + r(T \backslash \Gamma_M(s_i))) - m^2 + m,$$

ahol $S = \{s_1, ..., s_m\}$, d_M a G_M-beli fokszámfüggvény, hasonlóan Γ_M a G_M-beli szomszédság és r M rangfüggvénye.

A transzverzális matroidok prezentáciról szóló fenti állításokat itt nem bizonyítjuk. A transzverzális matroidok és prezentációik különböző tulajdonságairól a [3], [4], [5], [6], [7] cikkekben lehet olvasni.

Ha optimális batch kódról van szó, akkor következik, hogy a saját matroidjának minimális prezentációja. Ennek méretét meg tudjuk határozni, ha ismerjük a matroid maximális prezentációját.

7.1. Tétel ([8]). *$k > 1$. Ha egy optimális $CBC(n, k, m)$ matroidja M és M maximális prezentációja G_M, akkor*

$$N(n, k, m) = \sum_{i=1}^{m} (d_M(s_i) + r(T \backslash \Gamma_M(s_i))) - m^2 + m.$$

Ennek a módszernek a felhasználásával többek között levezethetjük az $n = m + 2$ esetben az 6.11 tételt. Itt az M^* duális matroidot vizsgáljuk, amelyről tudjuk, hogy kettő rangú, ezért struktúrája egyszerűen jellemezhető. Legyen $X_0 \subseteq$ a hurkok halmaza. Két pontot párhuzamosnak mondunk, ha kételemű kört alkotnak. Ez egy ekvivalencia-reláció. Legyenek a párhuzamosság legalábbb kételemű ekvivalencia-osztályai $X_1, ..., X_p \subseteq T$ és $X_{p+1} = T \backslash \cup_{i=0}^{p}$. Ekkor M^* körei X_0 elemei, minden $1 \leq$

$i \leq p$-re X_i elempárjai és minden háromelemű halmaz, ami ezeket nem tartalmazza, mert M^* rangja kettő. Ez egyértelműen meghatározza M^*-ot és így M-et is. M^* bázisai azok a kételemű halmazok, amik diszjunktak X_0-tól és minden $1 \leq i \leq p$-re legfeljebb egy elemet tartalmaznak X_i-ből. Feltétel, hogy létezzen legalább egy ilyen, tehát $p + |X_{p+1}| \geq 2$. Ennek megfelelően M bázisai azon $B \subseteq T$ m méretű halmazok, amikre $X_0 \subseteq B$ és minden $1 \leq i \leq p$-re $|X_i \backslash B| \leq 1$. Legyen minden i-re $|X_i| = n_i$, feltehető, hogy $n_1 \geq n_2 \geq ... \geq n_p$. Ekkor $\sum_{i=0}^{p+1} n_i = n$.

A fenti definíciók mellet belátjuk, hogy minden $m + 2$ méretű, m rangú matroid transzverzális és kiszámoljuk a maximális, majd a minimális prezentáció méretét.

7.2. Lemma ([8]). *M transzverzális matroid és egy maximális prezentációja a következő G_M gráf: S csúcsait osszuk az $Y_0, Y_1, ..., Y_{p+1}$ halmazokra, ahol $|Y_0| = n_0$, minden $1 \leq i \leq p$-re $|Y_i| = n_i - 1$ és $|Y_{p+1}| = n_{p+1} + p - 2$. Az élekbe vegyük be minden $0 \leq i \leq p$ esetén az (X_i, Y_i) teljes párosgráf éleit, valamint a (T, Y_{p+1}) teljes párosgráf éleit.*

Bizonyítás. [8] Legyen $T' \subseteq T$ és $|T'| = 2$. Kérdés, hogy $T \backslash T'$ mikor fedhető párosítással, azaz mikor létezik $(S, T \backslash T')$ teljes párosítás. Ha T' tartalmaz X_0-beli csúcsot, akkor Y_0-nak nincs elég szomszédja és nem létezik teljes párosítás. Ha $1 \leq i \leq p$-re $T' \subseteq X_i$, akkor Y_i-nek nincs elég szomszédja és nem létezik teljes prosítás. Ha ezek nem teljesülnek, akkor minden $0 \leq i \leq p$-re Y_i lefedhető (X_i, Y_i) párosítással, Y_{p+1} pedig tetszőleges élekkel lefedhető, ezért létezik teljes párosítás.

Ezzel beláttuk, hogy M transzverzális matroid és G_M prezentációja. Bizonyítandó, hogy ez maximális. Vegyünk hozzá G_M-hez egy új e élt, legyen e S-beli csúcsa Y_i-ben $(0 \leq i \leq p)$. Ha $i = 0$, létezik olyan párosítással fedhető halmaz az új gráfban, ami nem tartalmazza X_0-t, ezért már nem M prezentációja. Ha $i > 0$, létezik olyan párosítással fedhető halmaz az új gráfban, amiből két X_i-beli csúcs is hiányzik, ezért már nem M prezentációja.

□

[?] Ha M optimális $CBC(m + 2, k, m)$ matroidja, akkor a tétel alapján:

$$N(m+2,k,m) = \sum_{i=1}^{m} (d_M(s_i) + r(T \backslash \Gamma_M(s_i)) - m + 1)$$

$$= n_0 \cdot (n_0 + r(\cup_{j=1}^{p+1} X_j) - m + 1) + \sum_{i=1}^{p} (n_i - 1) \cdot (n_i + r(\cup_{j \neq i} X_j) - m + 1)$$

$$+ (n_{p+1} + p - 2) \cdot (m + 2 - r(\emptyset) - m + 1)$$

$$= n_0 \cdot (n_0 + (m - n_0) - m + 1) + \sum_{i=1}^{p} (n_i - 1) \cdot (n_i + (m + 1 - n_i) - m + 1)$$

$$+ (n_{p+1} + p - 2) \cdot (m + 2 - m + 1)$$

$$= n_0 + 2\sum_{i=1}^{p} n_i + 3(n_{p+1} + p - 2) = 2m - 2 - n_0 + n_{p+1} + p$$

Így ha $\mathcal{M}$ a szabályos $CBC(m+2, k, m)$-ek transzverzális matroidjainak halmaza,

$$N(m+2, k, m) = min_{M \in \mathcal{M}}(2m - 2 - n_0 + n_{p+1} + p).$$

A fentiek alapján egy $m+2$ elemű, m rangú matroid pontosan akkor van benne $\mathcal{M}$-ben, ha minden függő halmaza legalább $k+1$ elemű. T egy részhalmaza éppen akkor függő, ha nem részhalmaza M egyetlen bázisának sem, azaz M^* minden bázisába belemetsz. Eszerint a minimális függő halmazok kétféleképpen nézhetnek ki: $T \backslash X_0 \backslash \{t\}$, ahol $t \in X_{p+1}$ vagy $T \backslash X_0 \backslash X_i$, ahol $1 \leq i \leq p$.

Ha $p = 0$, akkor a legkisebb méretű függő halmaz $T \backslash X_0 \backslash \{t\}$ alakú és mérete $n - n_0 - 1$. A szabályosság feltétele

$$n - n_0 - 1 \geq k + 1$$

$$n_0 \leq n - k + 2 = m - k.$$

Tehát

$$2m - 2 - n_0 + n_{p+1} + p = 2m - 2 - n_0 + n_1 = 3m - 2n_0 \geq 3m - 2(m+k) = m - 2k.$$

A bonyolultabb eset, ha $p \geq 1$. Ekkor a legkisebb méretű függő halmaz $T \backslash X_0 \backslash X_1$, mert feltettük, hogy $n_1 \geq n_2 \geq ... \geq n_p$. Ennek mérete $n - n_0 - n_1$, tehát a szabályosság feltétele

$$n - n_0 - n_1 \geq k + 1$$

$$n_0 + n_1 \leq n - k - 1 = m - k + 1.$$

Ebben az esetben a $2m - 2 - n_0 + n_{p+1} + p$ kifejezést minimalizáló matroidot szabályosság tartó transzfotmációkkal hozzuk egyszerűbb alakra.

A Transzformáció: Ha $n_{p+1} \geq 2$, akkor X_{p+1}-et kettéosztjuk a X'_{p+1} és X_{p+2} halmazokra, ahol az új X'_{p+1} kételemű. Ezzel p értékét megnöveltük egyel, a matroid továbbra is szabályos kódhoz tartozik és a minimalizálandó kifejezés nem nőtt.

B Transzformáció: Ha $p \geq 2$, $n_0 > 0$ és $n_{p+1} \geq 1$, akkor vegyünk egy $x_0 \in X_0$ és egy $x \in x_{p+1}$ elemet. x_0-t áthelyezve X_1-be és x-et X_2-be a matroid szabályos marad, mert $n_0 + n_1$ nem nő, továbbá a minimalizálandó kifejezés sem nő.

7.3. Lemma ([8]). *Létezik $2m - 2 - n_0 + n_{p+1} + p$-t minimalizáló matroid, amiben X_{p+1} üres és $n_0 + n_1 = m + 1 - k$.*

Bizonyítás. [8] Vegyünk egy optimális M matroidot. Ha $n_0 + n_1 < m + 1 - k$, akkor $T \backslash X_0 \backslash X_1$ tetszőleges elemét tegyük át X_0-ba. Ezzel a minimalizálandó kifejezést csökkentettük, mert n_0 egyel nőtt és ha n_{p+1} egyel nőtt, akkor p egyel

csökkent. A matroid továbbra is szabályos marad, ami ellentmond a feltételnek, hogy M optimális. Feltehető, hogy $n_0 + n_1 = m - k + 1$.

Ha $n_{p+1} \geq 2$ végezzük el az **A** transzforációt, és ezt ismételjül, amíg $n_{p+1} < 2$ nem teljesül. A transzformáció tartja az optimalitást és az $n_0 + n_1 = m - k + 1$ egyenlőséget. Ha $n_{p+1} = 1$, $n_0 > 0$ és $p \geq 2$, akkor végezzük el a **B** transzformációt, ami szintén meghagyja az $n_0 + n_1 = m - k + 1$ feltételt. Ha $n_{p+1} = 1$, $n_0 = 0$ és $p \geq 2$, akkor $n_1 = m-k+1 > n_2$. Ekkor X_{p+1} egyetlen eleme hozzáadható X_2-höz az X_i halmazok rendezésének elrontása nnélkül. Mindkét esetben a lemmának megfelelő optimális matroidot kaptunk, amiben $n_{p+1} = 0$. $n_{p+1} = p = 1$ nem lehetséges, mert ekkor $m - k + 1 = n_0 + n_1 = n - 1$.

□

[8] Ha $\mathcal{M}'$ azon $m+2$ méretű, m rangú matroidok halmaza, ahol $n_0+n_1 = m-k+1$ és $X_{p+1} = \emptyset$, akkor

$$N(m+2, k, m) = min_{M \in \mathcal{M}'}(2m - 2 - n_0 + p)$$

n_0 képletét behelyettesítve a következő kifjezést kell minimalizálni:

$$2m - 2 - n_0 + p$$

$$2m - 2 - (m - k + 1) + n_1 + p$$

$$n_1 + p$$

$X_1, X_2, ..., X_p$ egy $m + 2 - n_0$ méretű halmaz partíciója, ahol minden rész mérete legalább kettő és legfeljebb n_1. Emiatt, rögzített n_1 mellett az optimális p $\left\lceil \frac{m+2-n_0}{n_1} \right\rceil$, tehát a következő kifejezész kell minimalizálni:

$$n_1 + \left\lceil \frac{m + 2 - n_0}{n_1} \right\rceil = \left\lceil \frac{n_1 + k + 1}{n_1} \right\rceil = n_1 + 1 + \left\lceil \frac{k+1}{n_1} \right\rceil$$

$$n_1 + \left\lceil \frac{k+1}{n_1} \right\rceil,$$

arra vigyázva, hogy $0 \leq n_0 = m - k + 1 - n_1$.

Tehát keressük az $f(x) = x + \left\lceil \frac{C}{x} \right\rceil$ függvény minimumát egész x-re. Könnnyen látható, hogy az $x + \frac{C}{x}$ kifejezés $\sqrt{C}$-ben minimális és ezen pont előtt és után rendre monoton csökken és nő. $f(x) = \left\lceil x + \frac{C}{x} \right\rceil$ és a felsőegészrész-függvény monoton növő, ezért f az $\left(0, \lfloor \sqrt{C} \rfloor\right]$ és $\left[\lceil \sqrt{C} \rceil, \infty\right)$ intervallumokon rendre monoton csökken és nő, tehát a minimuma $\lfloor \sqrt{C} \rfloor$ vagy $\lceil \sqrt{C} \rceil$.

A 6.11 tétel bizonyításában látottakhoz hasonló esetszétválasztást alkalmazunk. Legyen a egész szám, ahol $a^2 < C \leq (a+1)^2$. Ha $a^2 < C \leq a(a+1)$, akkor

$a < \sqrt{C} < a + \frac{1}{2}$, tehát f minimuma a vagy $a + 1$.

$$a < \frac{C}{a} \leq a + 1$$

$$f(a) = a + a + 1 = 2a + 1$$

$$a - 1 < \frac{C}{a+1} \leq a$$

$$f(a+1) = a + 1 + a = 2a + 1$$

Ekkor f minimuma $2a + 1 = \lceil 2\sqrt{C} \rceil$.

Ha $a^2 + a + 1 \leq C < (a+1)^2$, akkor $a + \frac{1}{2} < \sqrt{C} \leq a + 1$, tehát f minimuma a vagy $a + 1$.

$$a + 1 < \frac{C}{a} \leq a + 2$$

$$f(a) = a + a + 2 = 2a + 2$$

$$a < \frac{C}{a+1} < a + 1$$

$$f(a) = a + 1 + a + 1 = 2a + 2$$

Ekkor f minimuma $2a + 2 = \lceil \sqrt{C} \rceil$.

Végül, ha $C = (a+1)^2$, akkor $f(a+1) = a + 1 = a + 1 = 2a + 2 = \lceil 2\sqrt{C} \rceil$. Tehát minden esetben igaz, hogy f a minimumát felveszi $\lceil \sqrt{C} \rceil$-ben, ami $\lceil 2\sqrt{C} \rceil$. Továbbá f ezen pont előtt és után rendre monton csökken és nő.

Ezek szerint az oprimális n_1 $\lceil \sqrt{k+1} \rceil$ lenne, ami akkor valósítható meg, ha ebben az esetben $n_0 \geq 0$:

$$0 \leq n_0 = m - k + 1 - n_1 = m - k + 1 - \lceil \sqrt{k+1} \rceil$$

$$m \geq k - 1 + \lceil \sqrt{k+1} \rceil = k + \lceil \sqrt{k} \rceil$$

Ha az optimális n_1 érték nem megengedett, akkor f monotonitási tulajdonságai miatt a legnagyobb megengedett n_1 lesz optimális, azaz $n_0 = 0$ és $n_1 = m - k + 1$. Ezzel megkaptuk $N(m+2, k, m)$ értékét minden esetben. Ahogy vártuk, éppen a 6.11 tételt keptuk vissza.

$$m \leq k + \sqrt{k} \Longrightarrow 2m - 2 - n_0 + p = 2m - p = 2m - 2 + 1 + \left\lceil \frac{k+1}{m-k+1} \right\rceil = 2m + \left\lfloor \frac{k}{m-k+1} \right\rfloor$$

$$m \geq k + \sqrt{k} \Longrightarrow 2m - 2 - n_0 + p = 2m - 2 - (m - k + 1) + n_1 + p$$

$$= m + k - 2 + (n_1 + p - 1) = m + k - 2 + \lceil 2\sqrt{k+1} \rceil$$

$$N(m+2, k, m) = \begin{cases} 2m + \left\lfloor \frac{k}{m-k+1} \right\rfloor & ha \quad k \leq m \leq k + \sqrt{k} \\ m + k - 2 + \lceil 2\sqrt{k+1} \rceil & ha \quad m > k + \sqrt{k} \end{cases}$$

8. Uniform batch kódok

Egy batch kódot c-uniformnak nevezünk, ha a T-beli csúcsok fokszáma azonosan c. Uniform batch kódok esetén a kézenfekvő feladat adott m, k és c mellett maximális olyan n keresése, amire létezik c-uniform $CBC(n, cn, k, m)$. Ezt a számot nevezzük $n(m, c, k)$-nak. A fejezetet egy, az 3. fejezetben látotthoz hasonló, általános becsléssel kezdjük, ami a $c = k-1, k-2$ esetkre optimális konstrukciókat ad. Ennél kisebb c-re azonban nem ismert $n(m, c, k)$. Érdekes speciális eset $c = 2$, ami egy gráfelméleti feladatra vezethető vissza.

[14] A 3.1-es állítás bizonyításához hasonlóan végezzünk kettős leszámlálást az (S', t) párokon, ahol S' S-nek $k-1$ elemű részhalmaza, $t \in T$ és $\Gamma(t) \subseteq S'$. A nem-uniform esettel azonos módon semelyik S' nem tartalmazhatja T k különböző csúcsának szomszédságát, mert ekkor ezek a csúcsok hiányos halmazt alkotnának. Viszont speciálisan az uniform esetben tudjuk, hogy minden $t \in T$-hez pontosan $\binom{m-c}{k-1-c}$ darab S' halmaz tartozik. Ebből következik:

$$n\binom{m-c}{k-1-c} \leq (k-1)\binom{m}{k-1}$$

$$\frac{n}{k-1}\binom{k-1}{c} \leq \frac{\binom{m}{k-1}\binom{k-1}{c}}{\binom{m-c}{k-1-c}} = \frac{m!(k-1)!(k-1-c)!(m-k+1)!}{(k-1)!(m-k+1)!c!(k-1-c)(m-c)!} = \binom{m}{c}$$

8.1. Becslés ([14]).

$$n(m, c, k) \leq \frac{(k-1)\binom{m}{c}}{\binom{k-1}{c}}$$

Könnyen látható, hogy ez a becslés egyenkőséggel teljesíthető $c = k-1$ és $c = k-2$ esetén.

8.2. Tétel ([14]). $n(m, c, c+1) = c\binom{m}{c}$

Bizonyítás. [14] A 3.3 konstrukció éppen c-uniform lesz $n = c\binom{m}{c}$, $k = c+1$ esetén: S minden c méretű részhalmazához választunk c darab csúcsot T-ből és ezeket összekötjük egymással. A kód a 8.1 becslést egyenlőséggel teljesíti:

$$\frac{c\binom{m}{c}}{\binom{c}{c}} = c\binom{m}{c} = n,$$

ezért optimális.

□

8.3. Tétel ([14]). $n(m, c, c+2) = \binom{m}{c}$

Bizonyítás. **[14]** Itt az optimális uniform konstrukció egy $CBC(\binom{m}{c}, c\binom{m}{c}, c+2, m)$. S minden c méretű részhalmazához választunk egy T-beli csúcsot, amit ennek a részhalmaznak a csúcsaival kötünk össze.

Tegyük fel, hogy $X \subseteq T$ hiányos halmaz. Ekkor $c \leq |\Gamma(X)| \leq k = c+2$. Ha $|\Gamma(X)| = c$, akkor $|X| = 1$, ezért nem lehet hiányos. Ha $|\Gamma(X)| = c+1$, akkor $|X| \leq \binom{c+1}{c} = c+1$ és szintén nem lehet hoánypos. A kód a 8.1 becslést egyenlőséggel teljesíti:

$$\frac{(c+1)\binom{m}{c}}{\binom{c+1}{c}} = \binom{m}{c} = n,$$

ezért optimális. Ez a konstrukció valójában a 3.8 konstrukció $n = \binom{m}{c}$, $k = c+2$ speciális esete.

□

A $c \leq k-3$-ra nem ismert $n(m,c,k)$ optimális esetben. Térjünk át a $c=2$ esetre. [14] Itt T minden csúcsának fokszáma kettő ezért a kód teljesen jellemezhető a 4. fejezetben definiált G_2 segédgráffal. Azaz legyen $G_2 = (V_2, E_2)$ multigráf (párhuzamos élek lehetnek, de hurokélek nem), ahol $V_2 = S$ és uv éppen annyiszor van benne E_2-ben, amennyi $t \in T$-nek a szomszédsága éppen $\{u,v\}$. A Hall feltételből könnyen látható, hogy egy G_2 gráf éppen akkor tartozik szabályos $CBC(n, 2n, k, m)$-hez, ha nincs benne $i < k$ csúcs, ami több mint i élt feszít ki.

A 8.2 és 8.3 tételek alapján adódik, hogy $n(m,2,3) = 2\binom{m}{2}$ és $n(m,2,4) = \binom{m}{2}$, ezért az első érdekes kérdés $n(m,2,5)$, azaz legfeljebb hány éle lehet egy m-csúcsú gráfnak, ha $i < 5$ csúcs nem feszíthet ki i-nél több élt.

8.4. Tétel ([1]). $n(m,2,5) = \left\lfloor \frac{m^2}{4} \right\rfloor$

Bizonyítás. **[1]** Könnyen ellenőrizhető, hogy az $(\lfloor \frac{m}{2} \rfloor, \lceil \frac{m}{2} \rceil)$ csúcsú teljes egyszerű páros gráfban két csúcs legfeljebb egy, három csúcs legfeljebb kettő és négy csúcs legfeljebb négy élt feszíthet ki. Emiatt $n(m,2,5) \geq \left\lfloor \frac{m^2}{4} \right\rfloor$ és elég belátni, hogy egy ennél több élű gráfban már van rossz részgráf.

Ezt m-re vonatkozó teljes indukcióval bizonyítjuk. Ha $m = 4$, akkor $n > 4$ miatt az egész gráf rossz. Ha $m > 4$ feltehető, hogy $n = \left\lfloor \frac{m^2}{4} \right\rfloor + 1$. Ekkor $n < \frac{1}{2} \cdot \left(\lfloor \frac{m}{2} \rfloor + 1 \right) \cdot m$, ezért létezik olyan csúcs, aminek fokszáma legfeljebb $\lfloor \frac{m}{2} \rfloor$. Ezt a csúcsot elhagyva $m-1$ csúcsú, legalább $\left\lfloor \frac{m^2}{4} \right\rfloor + 1 - \lfloor \frac{m}{2} \rfloor = \left\lfloor \frac{(m-1)^2}{4} \right\rfloor + 1$ élű multigráfot kapunk. Ebben az indukciós feltevés alapján van rossz részgráf.

□

$k > 5$ esetén nem ismert $n(m,2,k)$ pontos értéke.

9. Batch kódok affin síkokból

A véges testek feletti affin síkok struktúráját kihasználva konstruálható néhány batch kód. Így optimális és közel-optimális konstrukciókat kapunk az uniform és a nem-uniform feladatra is, a $k = m$ és $k = m - 1$ nagyon speciális eseteire.

Még a konstrukciók előtt felírunk egy egyszerű képletet, ami a batch kódok szabályosságának bizonyításakor számtalanszor felhasználásra kerül majd.

Megfigyelés. [15] A q-rendű affin sík irányainak nevezzük a párhuzamosság ekvivalenci-osztályait. A q-rendű affin síkon tehát $q+1$ irány van és minden irányban q egyenes van. Ha két különböző irányból veszünk rendre x és y darab egyenest, akkor ezek összesen $qx + qy - xy$ pontot fognak le. Valóban, mivel minden egyenes q pontot fog le, párhuzamos egyeneseknek nincs közös pontja és minden két nempárhuzamos egyenesnek pontosan egy közös pontja van.

9.1. Konstrukció ([15]). *Az első konstrukció a legegyszerűbb: Legyen $q \geq 3$ prímhatvány és S jelölje egy q-rendű affin sík pontjait, T pedig az egyeneseit. Két csúcsot akkor kötünk össze, ha illeszkednek egymásra. Ekkor $n = q^2 + q$, $m = q^2$ és T-ben minden csúcs foka q, ezért $N = q^3 + q^2$. Az állítás, hogy a konstrukció egy q-uniform $CBC(q^2 + q, q^3 + q^2, q^2, q^2)$.*

[15] Az affin síkra átfogalmazva ez azt jelenti, az egyenesek közül bármely $i \leq q^2$-re legalább i pont illeszkedik. Ha $i = q^2$, akkor skatulya elv alapján létezik olyan irány, amiből az összes q egyenes ki van választva. Ekkor ezek az egyenesek az összes pontot lefogják. Ellenkező esetben i felírható $a(q + 1) + b$ alakban, ahol $0 \leq a \leq q - 2$ és $1 \leq b \leq q + 1$. Ha létezik olyan irány, amiből legalább $a + 2$ egyenes ki van választva, akkor az ezen egyenesek által lefogott pontok száma $q(a+2) = aq+2q > aq+a+b = i$. Ha minden irányból legfeljebb $a+1$ egyenes van kiválasztva, akkor legalább b irányból pontosan $a + 1$ ($i = a(q + 1) + b$ miatt).

Ha pontosan egy irányból van csak $a + 1$ kiválasztott egyenes, akkor következik, hogy $j = 1$ és a maradék q irányból mind pontosan a egyenes van kiválasztva. Vegyünk egy $a + 1$ és egy a egyenest tartalmazó irányt és tekintsük ezeket az egyeneseket. Ezek összesen $(a + 1)q + aq - (a + 1)a \geq aq + q + a > i$ pontot lefognak.

Ha legalább két irányból van $a + 1$ egyenes kiválasztva, akkor tekintsük két ilyen irány kiválasztott egyeneseit. Ezek az előző esethez hasonlóan legalább $(a + 1)q + (a + 1)q - (a + 1)^2 \geq (a + 1)q + (a + 1) = a(q + 1) + (q + 1) \geq i$ pontot lefognak.

□

Ez a konstrukció egyenlőséggel teljesíti a 8.1 becslést, így optimális q-unifrom kód és $n(q^2, q, q^2) = q^2 + q$.

$$\frac{(k-1)\binom{m}{c}}{\binom{k-1}{c}} \frac{(q^2-1)\binom{q^2}{q}}{\binom{q^2-1}{q}} = q(q+1) = n$$

9.2. Konstrukció ([15]). *A 9.1 konstrukcióban T egy tetszőleges t csúcsát hagyjuk el, valamint $\Gamma(t)$-t hagyjuk el S-ből. Ekkor $n = q^2 + q - 1$, $m = q^2 - q$ és S-ben minden csúcs foka $q + 1$, ezért $N = q^3 - q$. Az állítás, hogy a konstrukció egy $CBC(q^2 + q - 1, q^3 - q, q^2 - q - 1, q^2 - q)$. (Ez már nem uniform.)*

[15] Ennél a konstrukciónál tehát az affin síkból hiányzik egy t egyenes az összes pontjával együtt. Nevezzük t irányát vízszintesnek. A hiányzó pontok miatt itt a fejezet elején szerepelt megjegyzés kicsit módosított változata lesz érvényes: Ha a két irány közül egyik sem vízszintes, akkor az egyenesek legalább (nem feltétlenül pontosan) $(q-1)x + (q-1)y - xy$ pontot fognak le. Ha az x-hez tartozó irány vízszintes, akkor pontosan $qx+(q-1)y-xy$-t. Az állítás az, hogy a többi egyenesből bárhogy választunk ki $i \leq q^2 - q - 1$-et, ezekre legalább i pont fog illeszkedni.

Legyen az i kiválasztott egyenesből i_1 nem vízszintes és i_2 vízszintes, továbbá legyen $i_1 = aq + b$, ahol $0 \leq a \leq q - 2$ és $0 \leq b \leq q - 1$. Legyen az egyik nem vízszintes irány, amiből maximálisan sok egyenes van kiválasztva függőleges, és legyen a kiválasztott függőleges egyenesek száma M. Tudjuk, hogy $M \geq a$. Innen esetszétválasztást alkalmazunk M szerint.

Az első eset, ha $M \geq a+2$. Ekkor ha $a = q-2$, a függőleges egyenesek önmagukban legalább $(a + 2)(q - 1) = q^2 + q > i$ pontot fognak le. Ha viszont $a < q - 2$, akkor a vízsintes egyenesek és $a + 2$ darab függőlegegs egyenes összesen $(a+2)(q-1)+i_2q-i_2(a+2) = aq + q + i_2(q - a - 2) \geq aq + b + i_2 = i$ pontot fog le.

A második eset, ha $M = a + 1$. Ekkor kell hogy létezzen legalább egy nem vízszintes, nem függőleges irány, amiből legalább a egyenes van kiválasztva (ellenkező esetben nem teljesülhetne $i_1 = aq + b$). Az ilyen irányú és függőleges egyenesek összesen $(a+1)(q-1)+a(q-1)-a(a+1) = aq+q-1+a(q-a-3)$ pontot fognak le. Ez láthatóan nagyobb vagy egyenlő mint i, ha $a \leq q-4$ és $i_2 \leq a$. Másrészt viszont, a függőleges és vízszintes egyenesek összesen legalább $(a+1)(q-1)+i_2q-(a+1)i_2 = aq+q-a-1+i_2(q-a-1) = aq+q-1+(i_2-a)+i_2(q-a-2)$ pontot fognak le. Ez nagyobb vagy egyenlő mint i, ha $i_2 \geq a$ és $a \leq q - 3$. Tudjuk, hogy $a \leq q - 2$ így már csak az $a = q - 2$ és az $a = q - 3$, $i_2 \leq q - 4$ esetkre nincs belátva az állítás.

Tegyük fel, hogy $a = q - 2$ és $M = q - 1$. Ha létezik nem vízszintes, nem függőleges irány, amiből legalább $q - 1$ egyenes van kiválasztva, akkor mind ebben az irányban, mind a függőleges irányban csak egy egyenes marad ki. Emiatt a kiválasztott egyenesek minden pontot lefognak, kivéve a két hiányzó egyenes legfeljebb egy közös pontját. Ez legalább $q^2 - q - 1 \geq i$ pont. Ha azonban minden nem vízszintes, nem függőleges irányban legfeljebb $q - 2$ pont van kiválasztva, akkor $i_1 \leq (q - 1) + (q - 2)(q - 1) = aq + 1$ és $b \leq 1$. Ebben az esetben a vízszintes és függőleges egyenes együtt legalább $(q - 1)(q - 1) + i_2q - i_2(q - 1) = aq + 1 + i_2 \geq i$ pontot fognak le.

Tegyük fel, hogy $a = q - 3$, $M = q - 2$ és $i_2 \leq q - 4$. Ha létezik nem vízszintes, nem függőleges irány, amiből legalább $q - 2$ egyenes van kiválasztva, akkor mind ebben az irányban, mind a függőleges irányban csak két egyenes marad ki. Emiatt

a kiválasztott egyenesek minden pontot lefognak, kivéve a két-két hiányzó egyenes legfeljebb négy metszéspontját. Ez legalább $q^2 - q - 4 \geq (q-3)q + q + (q-4) \geq i$ pont. Ha azonban minden nem vízszintes, nem függőleges irányban legfeljebb $q-3$ pont van kiválasztva, akkor $i_1 \leq (q-2) + (q-3)(q-1) = aq + 1$ és $b \leq 1$. Ebben az esetben feltétlenül létezik nem vízszintes, nem függőleges irány, amiből $q-3$ egyenes van kiválasztva, különben i_1 túl kicsi kenne. Egy ilyen irány egyenesei és a függőleges egyenesek összesen legalább $(q-2)(q-1) + (q-3)(q-1) - (q-2)(q-3) = (q-2)(q-1) + q - 3 \geq aq + 1 + (q-4) \geq i$ pontot fognak le.

A harmadik eset, ha $M = a$. Ekkor minden nem vízszintes irányból pontosan a egyenesnek kell kiválasztva lennie, $i_1 = aq$ és $b = 0$. Bármely két nem vízszintes irány egyenesei összesen $a(q-1) + a(q-1) - a^2 = aq + a(q-a-2)$ pontot fognak le. Ez nagyobb vagy egyenlő mint i, ha $i_2 = 0$, vagy $a \leq q-3$ és $i_2 \leq a$. Másrészt viszont a függőleges és vízszintes egyenesek összesen legalább $a(q-1) + i_2 q - i_2 a = aq + i_2 + i_2(q-a-1) - a$ pontot fognak le. Ez legfeljebb i, ha $i_2 \geq a$.

Ekkor az egyetlen megvizsgálatlan eset, ha $M = a = q-2$ és $1 \leq i_2 \leq q-3$. $q \geq 4$, mert ellenkező esetbne i_2-nek nem lehetne megfelelő értéke. Vegyünk ellőször két tetszőleges nem vízszintes irányt. Mindkét irányban csak két-két egyenes marad ki a kiválasztottak közül. Emiatt a kiválasztott egyenesek minden pontot lefognak, kivéve a két-két hiányzó egyenes legfeljebb négy metszéspontját. Ez legalább $q^2 - q - 4 \geq (q-2)q + q - 4$, ami nagyobb vagy egyenlő mint i, kivéve ha i_2 éppen $q-3$. Feltehető, hogy $i_2 = q-3$, azaz midnden irányból (még a vízszintesből is) pontosan kettő egyenes nincs kiválasztva és legalább 5 irány van. Ez összesen $q^2 - q - 3$ egyenes, tehát azt kell belátni, hogy legfeljebb 3 pont nincs az egyenesek által lefedve. Tegyük fel, hogy van 4 lefedetlen pont. Ha ezek közül három egy egyenesen van, akkor minden ettől különböző irányú egyenesekből legalább háromra illeszkedik egy ilyen pont, ami ellentmondás. Ha nincs a 4 pont közül semelyik három egy egyenesen, akkor minden irányhoz tartozik a 4 pontnak egy párosítása úgy, hogy a párok által meghatározott egyenes az adott irányba mutat. Azonban 4 pontot csak háromféleképpen lehet összepárosítani, de legalább 5 irány van az q rendű affin síkon: ez ellentmondás. Így az utolsó esetben is beláttuk, hogy a kiválasztott egyenesek legalább i pontot lefognak.

□

[15] Levezethető, hogy ez a konstrukció egyenlőséggel teljesíti a 3.5 becslést. A konstrukcióban T minden csúcsának fokszáma q vagy $q-1$. Ahogy azt a 3.5 becslés bizonyításában is megjegyeztük, ez a becslés csak akkor teljesülhet egyenlőséggel, ha T minden csúcsának fokszáma c vagy $c-1$, ezért a $c = q$-hoz tartozó becsléssel érdemes próbálkozni.

$$U_{m,k,q} = \frac{(k-1)\binom{m}{q}}{\binom{k-1}{q}} = \frac{(q^2-q-2)\binom{q^2-q}{q}}{\binom{q^2-q-2}{q}} = \frac{(q^2-q-2)(q^2-q)!q!(q^2-2q-2)!}{(q^2-q-2)!q!(q^2-2)!}$$

$$= \frac{(q^2-q)(q^2-2q)(q^2-q-2)}{(q^2-2q)(q^2-2q-1)} = \frac{(q^2-1)(q^2-q-1)}{q^2-2q-1}$$

$$nq - \left\lfloor \frac{(k-q)(U_{m,k,q}-n)}{m-k+1} \right\rfloor = (q^2+q-1)q - \left\lfloor \frac{(q^2-q-1-q)\left(\frac{(q^2-1)(q^2-q-1)}{q^2-2q-1} - (q^2+q-1)\right)}{q^2-q-(q^2-q-2)} \right\rfloor$$

$$= (q^2+q-1)q + \left\lfloor \frac{(q^2-1)(q^2-q-1) - (q^2+q-1)(q^2-2q-1)}{2} \right\rfloor$$

$$= (q^2+q-1)q + \left\lfloor \frac{2q^2}{2} \right\rfloor = q^3+q^2-q-q^2 = q^3-q = N$$

Tehát a 9.2 konstrukció optimális és $N(q^2+q-1, q^2-q-1, q^2-q) = q^3-q$. Ez azért különösen érdekes, mert idáig a 3.5 becslésekről csak a $c = k-1, k-2$ esetben tudtuk, hogy éles. A 9.2 konstrukcióból következik, hogy c tetszőlegesen nagy értékeire is éles, bár az egyenlőség teljesülését csak egy nagyon speciális esetben láttuk be. A következő konstrukció előtt felítrunk egy újabb megjegyzést, ami az következő szabályosság-bizonyításban kerül többször alkalmazásra.

9.3. Konstrukció ([15]). *A 9.2 konstrukcióban módosítsuk a vízszintes egyenesekhez tartozó csúcsokat. Minden ilyen csúcsnál hagyjunk el a gráfból egy tetszleges őt tartalmazó élt. Ezzel elértük, hogy a konstrukció uniform legyen. Ezután hagyjunk el a gráfból két tetszőleges vízszintes egyeneshez tartozó csúcsot. (Ekkor a 9.1 konstrukcióhoz képest három vízszintes egyenes van elhagyva.) Ekkor $n = q^2+q-3$, m még mindig q^2-q és a kó $q-1$-uniform, ezért $N = (q-1)(q^2+q-3)$. Az állítás, hogy a konstrukció egy $CBC(q^2+q-3, (q-1)(q^2+q-3), q^2-q-1, q^2-q)$.*

[15] A szabályosság bizonyítása nagyon hasonló lesz a 9.2 konstrukció szabályosságának bizonyításához. Ugyanazt a jelölést használjuk: Válasszunk ki testszőleges $i \leq q^2-q-1$ darab egyenest a hiányos affin síkon. Legyen a nem vízszintes egyenesek száma i_1, a vízszinteseké i_2. Ennél a konstrukciónál $i_2 \leq q-3$ teljesül, mivel három vízszintes egyenes hiányzik. Legyen $i_1 = aq+b$, ahol $0 \leq a \leq q-2$ és $0 \leq b \leq q-1$. Legyen az egyik nem vízszintes irány, amiből maximálisan sok egyenes van kiválasztva függőleges, és legyen M darab kiválasztott függőleges egyenes. A 9.2 konstrukcióhoz tartozó bizonyításhoz hasonlóan M szerint végzünk esetszétválasztást.

Ahol nem vízszintes irányú egyenesekkel fedtünk le megfelelő számú pontot, ott a bizonyítás továbbra is érvényes. Azonban mivel a vízszintes egyenesekre ebben a konstrukcióban csak $q-1$ pont illeszkedik x darab vízszintes és y darab egyirányú nem vízszintes egyenesre már csak az igaz, hogy összesen legalább $(q-1)x+(q-1)y-xy$ pontot fognak le.

Az első eset, ha $M \geq a+2$. Ekkor ha $a = q-2$, a függőleges egyenesek önmagukban lefognak minden pontot. Ha viszontn $a \leq q-3$, akkor van olyan nem vízszintes, nem függőleges irány, amiből legalább a egyenes van kiválasztva (ellenkező esetben $i_1 =$

$aq+b$ nem teljesülne). Tekintsük egy ilyen I irányból a és $a+2$ függőleges egyenes által lefogott pontokat. Mivel nincs az összes függőleges egyenes kiválasztva: legyen egy nem kiválasztott függőleges egyenes e. A hiányos affin síkon egy kivételével minden I irányú egyenesnek van e-vel közös pontja és ezek a pontok mind különbözőek. Ezért a függőleges egyenesek által lefogott $(a+2)(q-1)$ pont mellett még le van fogva legalább $a-1$ e-beli pont. Az összesen lefogott pontok száma legalább $(a+2)(q-1)+a-1 \geq aq+(q-1)+(q-3) \geq i$.

A második eset, ha $M = a-1$. Ez annyiban tér el a 9.2 konstrukció megfelelő esetétől, hogy a függőleges és vízszintes egyenesek összesen csak $(a+1)(q-1)+i_2(q-1)-(a+1)i_2 = aq+(q-1)+i_2(q-a-2)-a$ pontot fognak le, ami legalább i, ha $a \leq q-4$ és $a \leq x_2$. Itt is csak az $a = q-2$ és $a = q-3$ esetek maradnak hátra.

Az összes aleset ugyanúgy bizonyítható, mint a 9.2 konstrukciónál, kivéve ha $a = q-2$, $M = q-1$, minden nem vízszintes, nem függőleges irányból legfeljebb $q-2$ egyenes van kiválasztva és emiatt $i_1 \leq aq+1$, $b \leq 1$. Ekkor a függőleges egyenesek által lefogott pontok száma $(q-1)(q-1)$ és a hiányzó függőleges egyenes pontjai közül legalább $q-3$ le van fogva a többi egyenes által, mert van olyan nem vízszintes, nem függőleges irány, amiből pontosan $q-2$ egyenes van kiválasztva. Ez összesen $(q-1)^2+q-3 = aq+1+(q-3) \geq i$ pont.

A harmadik eset, ha $M = a$. Ez annyiban tér el a 9.2 konstrukció megfelelő esetétől, hogy a függőleges és vízszintes egyenesek összesen csak $a(q-1)+i_2(q-1)-ai_2 = aq+(q-1)+i_2(q-a-1)-a$ pontot fognak le, ami legalább i, ha $a \leq q-3$ és $a \leq i_2$. Itt az egyetlen kimaradó eset az $M = a = q-2$, $q \leq i_2 \leq q-3$, ami a 9.2-vel azonosan kezelhető.

□

Ez a konstrukció már nagyon megközelíti a 8.1 becslés alsó egészrészét, de még nem éri el.

$$\left\lfloor \frac{(k-1)\binom{m}{c}}{\binom{k-1}{c}} \right\rfloor - n = \left\lfloor \frac{(q^2-q-2)\binom{q^2-q}{q-1}}{\binom{q^2-q-2}{q-1}} \right\rfloor - (q^2+q-3)$$

$$= \left\lfloor \frac{(q^2-q)(q^2-q-1)(q^2-q-2)}{(q^-2q+1)(q^2-2q)} \right\rfloor - (q^2+q-3)$$

$$= \left\lfloor \frac{q^3-2q-1}{q-1} \right\rfloor - (q^2+q-3) = q^2+q-\left\lceil \frac{q+1}{q-1} \right\rceil - (q^2+q-3) = 1$$

Ez kis módosítással megjavítható.

9.4. Konstrukció ([15]). *A 9.3 konstrukcióban S minden csúcsának fokszáma q vagy $q+1$ attól függően, hogy a megfelelő ponton átmenő vízszintes egyenes benne van-e a konstrukcióban. Hagyjuk el S egy tetszőleges q-fokú csúcsát az összes szomszédjával*

együtt. Ekkor $n = q^2 - 3$, $m = q^2 - q - 1$ *és a kód még mindig* $q-1$*-uniform, ezért* $N = (q-1)(q^2-3)$. *A konstrukció még mindig szabályos* $CBC(q^2-3, (q-1)(q^2-3), q^2-q-1, q^2-q-1)$, *mert semelyik* T*-beli csúcsnak nem változott a szomszédsága a 9.3-hoz képest.*

Ez a konstrukció már eléri a 8.1 becslés alsó egészrészét.

$$\left\lfloor \frac{(k-1)\binom{m}{c}}{\binom{k-1}{c}} \right\rfloor = \left\lfloor \frac{(q^2-q-2)\binom{q^2-q-1}{q-1}}{\binom{q^2-q-2}{q-1}} \right\rfloor = \left\lfloor \frac{(q^2-q-2)(q^2-q-1)}{q^2-2q} \right\rfloor$$

$$= \left\lfloor \frac{q^3-2q-1}{q} \right\rfloor = q^2 - 3 = n$$

A batch kód optimális $q-1$-uniform, így $n(q^2-q-1, q-1, q^2-q-1) = q^2-3$. A 9.1 és 9.3 konstrukció is annak bizonyítéka, hogy a 8.1 becslés éles egyes $c = 2, k-2, k-1$-től különböző esetkben is, sőt tetszőlegesen nagy c-re lehet éles.

10. Optimális kódok $t \geq 1$ esetén

A dolgozat során idáig az egyszerűség kedvéért csak a $t = 1$ speciális esettel foglalkoztunk. Itt, az utolsó fejezetben az eddig bemutatott eredmények egy részét általánosítjuk a $t \geq 1$-re. Ehhez először ki kell mondanunk néhány egyszerű állítás, ami az általános kombinatorikus batch kódokra vonatkozik.

Az 2. fejezetben kimondott Hall feltételes jellemzése egy batch kód szabályosságának alapozta meg a további fejezetekben következő legtöbb tételt. Ennek megfelelően az általános esetben is kimondunk egy hasonló, bár valamivel bonyolultabb állítást.

10.1. Állítás (Hall feltétel). *Egy párosgráf* (S, T)*-n pontosan akkor szabályos batch kód* k *és* t *paraméterekkel, ha* T *bármely legfeljebb* k *méretű* X *részhalmazára* $|\Gamma(X)| \geq |X|/t$.

Bizonyítás. A feltétel szükségessége nyilvánvaló. Ha létezik egy hiányos $X \subseteq T$ halmaz, akkor ezt nem lehet lefedni olyan részgráffal, aminek S-ben mindenhol legfeljebb t a foka. Ugyanis egy ilyen gráfnak legfeljebb $t \cdot |\Gamma(X)| < |X|$ éle lehetne.

Az elégségesség bizonyításához tegyük fel, hogy az állítás nem teljesül és létezik egy gráf, amiben nincs hiányos halmaz, de az $X \subseteq T$ halmazt mégsem lehet megfelelő típusú részgráffal lefedni. A gráfhoz adjunk hozzá egy a forrás és egy z nyelő csúcsot és az $\{a, z\} \cup X \cup S$ csúcshalmazon definiáljunk egy hálózatot a következőképpen: Az (X, S)-re leszűkített eredeti gráf éleit irányítsuk X-től S felé végtelen kapacitással. A forrásból irányítsunk egy kapacitású éleket X minden csúcsába és S minden csúcsából irányítsunk t kapacitású éleket a nyelőbe. Látszik, hogy az így kapott hálózatban egy pontosan az $|X|$ méretű folyamok lesznek a megfelelő típusú X-et fedő részgráfok. Az

indirekt feltevés alapján ilyen nem létezik ezért a Ford-Fulkerson tétel szerint létezik k-nál kisebb vágás a hálózatban.

Legyen C egy k-nál kisebb vágás. Legyen $C \cup X = Y$ és $C \cup S = Z$. Mivel az X és S közötti élek kapacitása végtelen, $\Gamma(Y) \subseteq Z$-nek mindenképpen teljesülnie kell. A vágás teljes mérete:

$$k > |X \backslash Y| + t \cdot |Z| \geq k - |Y| + t \cdot |\Gamma(Y)|$$

$$|Y| > |\Gamma(Y)|$$

Ekkor Y Hall-hiányos, ami ellentmond az indirekt feltevésnek.

□

[11] A 10.1 Hall-feltétel ekvivalens azzal, hogy S semelyik $l < k/t$ méretű részhalmaza nem tartalmazhatja T-nek több, mint lt csúcsának teljes szomszédságát. Valóban, ha ez megtörténne, akkor $lt + 1 \leq k$ ilyen csúcs hiányos halmazt alkotna T-ben. Fordítva pedig, minden $X \subseteq T$ Hall-hiányos halmaz szomszédsága megszegi a fenti fetételt.

Ez az egyszerű áltfogalmazás azért fontos, mert l egészszám, így $l < k/t$ helyett elég megkötni, hogy $l < \lceil k/t \rceil$. Ezzel a kód szabályosságának egy olyan ekvivalens jellemzését kapjuk, ami csak t-től és $\lceil k/t \rceil$-től függ. Tehát ha $\lceil k_1/t \rceil = \lceil k_2/t \rceil$, akkor $N(n, k_1, m, t) = N(n, k_2, m, t)$ minden n-re és m-re. Innentől legyen $\lceil k/t \rceil = K$. Feltehető, hogy $k = tK$ [11].

Egyszerű összefüggés az általános és a $t = 1$ eset között, hogy ha adott egy $CBC(n, N, K, m, 1)$, ebből könnyen konstruálható $CBC(tn, tN, tK, m, t)$ a T-beli csúcsok meg-t-szerezésével. Látható, hogy ez valóban batch kód lesz a megfelelő paraméterekkel, tehát $N(tn, tK, m, t) \leq t \cdot N(n, K, m, 1)$. Sajnos ehhez a becsléshez fel kell tenni, hogy azáltalános feladathoz tartozó n osztható t-vel. Általános n esetén $CBC(\lceil n/t \rceil, N, K, m, 1)$-ből konstruálható $CBC(t \cdot \lceil n/t \rceil, tN, tK, m, t)$, ami lecsökkenthető $CBC(n, N', tK, m, t)$-vé néhány csúcs elhagyásával. Legfeljebb $t - 1$ darab csúcsot kell elhagynunk, ezért feltehető, hogy minden elhagyott csúcs a $CBC(\lceil n/t \rceil, K, m, 1)$-ből a legnagyobb fokszámú T-beli csúcs másolata.

10.2. Becslés. $N(n, tK, m, t) \leq t \cdot N(\lceil n/t \rceil, K, m, 1) - (t \cdot \lceil n/t \rceil) \cdot \delta$,
ahol $\delta = max_{u \in T} d(u)$ *egy optimális* $CBC(\lceil n/t \rceil, N, K, m, 1)$*-ben.*

A 3. fejezetben leírtak egy nagy része természetesen általánosítható t tetszőleges értékére, ezzel pontos értékeket adva $N(n, k, m, t)$-re elég nagy n esetén. Az első lépés a 3.1 becslés általánosítása, ami itt is fontos szerepet fog játszani az optimalitások bizonyításában.

[11] Ha optimális kódokról beszélünk, feltehetjük hogy nincs T-beli csúcs, aminek fokszáma nagyobb, mint K. Ez fölösleges lenne, mert már K-fokú csúcs sem lehet benne hiányos halmazban. Az első fejezethez hasonlóan végezzünk kettős leszámlálást

az (S', u) párokon, ahol $S' \subseteq S$ $K-1$ elemű halmaz, $u \in T$ és $\Gamma(u) \subseteq S'$. Itt S' nem tartalmazhatja T $t(K-1)+1$ különböző csúcsának szomszédságát. Viszont tetszőleges $u \in T$ esetén $\Gamma(u)$-t S-nek éppen $\binom{m-d(t)}{(K-1-d(t)}$ darab $K-1$-részhalmaza tartalmazza. Ebből következik a 3.1 állítás általánosítása.

10.3. Állítás ([11]).

$$\sum_{i=1}^{K-1} \binom{m-i}{K-1-i} A_i \leq t(K-1)\binom{m}{K-1}$$

[11] Ebből itt is levezethetünk egy nagy N-re vonatkozó becslést, ami elég nagy n-ek esetén optimálisnak fog bizonyulni.

$$N = \sum_{i=1}^{K} iA_i = \sum_{i=1}^{K}(K-(K-i))A_i = Kn - \sum_{i=1} A_i$$

$$= Kn - \sum_{i=1} K(K-i)A_i + \left(\sum_{i=1}^{K} \binom{m-i}{K-1-i} A_i - t(K-1)\binom{m}{K-1}\right)$$

$$= Kn - t(K-1)\binom{m}{K-1} + \sum_{i=1}^{K} \left(\binom{m-i}{K-1-i} - (K-i)\right) A_i$$

Ugyanúgy, mint a 3.2 becslés levezetésében, itt is nemnegativ a szummás tagban szereplő összes együttható, ezért az egész tag alulról becsülhető nullával. A 3.2 becslés általánosítását kapjuk:

10.4. Becslés ([11]).

$$N \geq Kn - t(K-1)\binom{m}{K-1}$$

Tekintsük az $n \geq t(K-1)\binom{m}{K-1}$ esetet. Ekkor $\lceil n/t \rceil \geq (K-1)\binom{m}{K-1}$, ezért a 10.2 becslés alkalmazásakor tekinthetjük a 3.3 konstrkciót, mint optimális $CBC(\lceil n/t \rceil, N, m, k, 1)$-t. Ebben a konstrukcióban a maximális T-beli fokszám K (kivéve az $\lceil n/t \rceil = (K-1)\binom{m}{K-1}$ esetben, de itt $t \cdot \lceil n/t \rceil - n = 0$). A 10.2 felsőbecslésből azt kapjuk, hogy

$$N \leq Kn - t(K-1)\binom{m}{K-1},$$

mert itt t-szer annyi T-beli csúcs foka $K-1$, mint a 3.3 konstrukcióban és a többi csúcs foka K. Ez éppen annyi, mint a 10.4 alsóbecslésben.

10.5. Tétel ([11]). *Ha* $n \geq t(K-1)\binom{m}{K-1}$, *akkor* $N(n, tK, m, t) = Kn - t(K-1)\binom{m}{K-1}$.

10.6. Becslés ([11]).

$$N \geq n(K-1) - \left\lfloor \frac{t(K-1)\binom{m}{K-1} - n}{m+K-1} \right\rfloor$$

Bizonyítás. [11] Itt is a 10.3 egyenlőtlenséget használjuk, de most $\frac{1}{m-K+1}$-el súlyozva. A bizonyítás a 3.5 becslés bizonyításával azonosan működik.

$$N = \sum_{i=1}^{K} iA_i = \sum_{i=1}^{K}(k-1-(k-1-i))A_i = n(K-1) - \sum_{i=1}^{K}(k-1-i)A_i$$

$$\geq n(k-1) - \sum_{i=1}^{K}(k-1-i)A_i + \frac{1}{m-K+1}\left(\sum_{i=1}^{K}\binom{m-i}{k-1-i}A_i - t(K-1)\binom{m}{K-1}\right)$$

$$= n(K-1) - \frac{t(K-1)\binom{m}{K-1} - n}{m-K+1} + \sum_{i=1}^{K}\left(\frac{\binom{m-i}{K-1-i}}{m-K+1} - \frac{1}{m-K+1} - (K-1-i)\right)A_i$$

Itt kihasználtuk, hogy $n = \sum_{i=1}^{K} A_i$. A 3.6 Lemma $c = K-1$ speciális esete alapján minden A_i együtthatója nemnegatív, ezért az egész szummát alulról becsülhetjük nullával. A baloldalon a törtnek vehetjük alsó egészrészét mert a kifejezés másik két tagja egész.

□

Megjegyzendő, hogy a 3.5 becslés általánosítását tetszőleges c esetén hasonlóan le lehet vezetni, de ezt itt nem tesszük meg, mert nem vezet ismert optimális batch kódhoz. Ha $t\binom{m}{K-2} \leq n \leq t(K-1)\binom{m}{K-1}$, akkor a 10.2 becslés alkalmazása már nem ad optimális értéket. Ezért a 3.8 konstrukciót általánosítjuk, hogy elérjük a 10.6 becslést.

10.7. Konstrukció ([11]). *Feltesszük, hogy* $t\binom{m}{K-2} \leq n \leq t(K-1)\binom{m}{K-1}$. *Először vegyük* $T\left\lfloor \frac{t(K-1)\binom{m}{K-1}-n}{m-K+1} \right\rfloor$ *csúcsát és mindegyiket kössük össze* S *egy* $k-2$*-részhalmazával úgy, hogy minden* $k-2$ *részhalmazot legfeljebb* t*-szer használjunk. Ezt megtehetjük, mert*

$$\left\lfloor \frac{t(K-1)\binom{m}{K-1}}{m-K+1} \right\rfloor \leq \frac{t(K-1)\binom{m}{K-1} - t\binom{m}{K-2}}{m-K+1}$$

$$= \frac{t(K-1)\binom{m}{K-2}\frac{m-K+2}{K-1} - t\binom{m}{K-2}}{m-K+1} = t\binom{m}{K-2}.$$

A 3.8 konstrukcióval azonosan ezek a csúcsok alkotják T_{K-2}*-t a többi csúcs pedig* T_{k-1}*-et. A* T_{K-1}*-beli csúcsok szomszédságait arra vigyázva választjuk ki, hogy semelyik*

$K-1$ méretű $S' \subseteq S$ ne legyen több mint $t(K-1) - |\{u \in T_{K-2} : \Gamma(u) \subseteq S'\}|$. Bizonyítandó, hogy ezzel T minden csúcsát be tudjuk rakni T_{K-2}-be vagy T_{K-1}-be.

$$\sum_{\substack{S' \subseteq S \\ |S'| = K-1}} (t(K-1) - |\{u \in T_{K-2} : \Gamma(u) \subseteq S'\}|) + A_{K-2}$$

$$= t(K-1)\binom{m}{K-1} - (m-K+2)A_{K-2} + A_{K-2}$$

$$= t(K-1)\binom{m}{K-1} - (m-K+1)\left\lfloor \frac{(K-1)\binom{m}{K-1} - n}{m-K+1} \right\rfloor \geq n$$

[11] Tegyük fel, hogy a fenti konstrukció nem szabályos batch kód és $X \subseteq T$ Hall-hiányos halmaz. Mivel T minden csúcsának szomszédsága legalább $K-2$, $|\Gamma(X)| \geq k-2$. Ha $|\Gamma(X)| = K-2$ X nem lehet hiányos, mert definíció szerint legfeljebb t különböző T_{K-2}-beli csúcsnak lehet a pontos szomszédsága. Ha $|\Gamma(X)| = K-1$, X szintén nem lehet hiányos, mert legfeljebb $t\binom{K-1}{K-2} = tK-1$ T_{K-2}-beli csúcs szomszédságát tartalmazhatja, arra pedig T_{k-1} konstrukciója során vigyáztunk, hogy $\Gamma(X)$ ne tartalmazza túl sok csúcs szomszédságát. A kód mérete

$$n(K-2) - A_{K-2} = n(K-1) - \left\lfloor \frac{t(K-1)\binom{m}{K-1} - n}{m-K+1} \right\rfloor.$$

Ez éppen a 10.6 becslés értéke.

□

10.8. Tétel ([11]). *Ha $t\binom{m}{K-2} \leq n \leq t(K-1)\binom{m}{K-1}$, akkor*

$$N(n, tK, m, t) = n(K-1) - \left\lfloor \frac{t(K-1)\binom{m}{K-1} - n}{m-K+1} \right\rfloor.$$

A fejezet ezután következő részében még levezetjük az optimális batch kód méretét a legegyszerűbb speciális esetekre. Ezek az esetek analógak már korábban tárgyalt esetekkel, csak itt nem tesszük fel, hogy $t = 1$. Ahogy a 2. fejezetben is, fel van téve, hogy $tm \leq n$. A $tm > n$ eset megegyezik a $tm = n$ esettel.

10.9. Tétel. $N(n, t, m, t) = n$

Bizonyítás. Nyilvánvalóan szükséges, hogy T-nek ne legyen izolált csúcsa, de ebben az esetben ez elégséges is.

10.10. Tétel ([11]). $N(n, 2t, m, t) = 2n - tm$

Bizonyítás. [11] Itt $tm \leq n$ miatt mindig teljesül a 8.5 tétel feltétele. Ennek alkalmazásával kapjuk az eredményt.

□

10.11. Tétel ([11]).

$$N(n, 3t, m, t) = \begin{cases} 2n - mt + \left\lceil \frac{n-mt}{m-2} \right\rceil & ha \quad tm \leq n \leq 2t\binom{m}{2} \\ 3n - 2t\binom{m}{2} & ha \quad 2t\binom{m}{2} \leq n \end{cases}$$

Bizonyítás. [11] $tm \leq n$ miatt $t\binom{m}{K-2} \leq n$ mindig teljesül, ezért a 8.5 vagy 8.8 tétel egyike mindig alkalmazható. A 8.8-ból következik az első eset, a 8.5-ből a második.

□

10.12. Tétel ([11]). $N(tm, tK, m, t) = tm$

Bizonyítás. [11] Nyilvánvalóan szükséges, hogy T-nek ne legyen izolált csúcsa. Itt létezik is olyan kód, amiben T minden csúcsának foka 1. Csoportosítsuk T csúcsait m darab t méretű csoportba, majd mindegyik csoportot kössük össze S egy különböző csúcsával.

□

10.13. Tétel ([11]). $N(n, tm, m, t) = nm - tm(m-1)$

Bizonyítás. [11]

Ha $s \in S$ fokszáma kevesebb lenne, mint $n - t(m-1)$, akkor $T \backslash \Gamma(s)$ hiányos halmaz lenne, mert a mérete nagyobb, mint $t(m-1)$, de a szomszédsága legfeljebb $|S - s| = m - 1$ méretű. Tehát minden S-beli csúcs fokszáma legalább $n - t(m-1)$ és $N \geq nm - tm(m-1)$.

Ez el is érhető, amit a következő konstrukció mutat: Legyen $T_1 \subseteq T$, $|T_1| = mt$ és $T' = T \backslash T'$. S és T_1 között vegyünk egy optimális $CBC(tm, tm, tK, m, t)$-t (a 8.12 tétel alapján), S és T' között pedig vegyünk teljes párosgráfot. Ekkor T'-beli csúcs nem lehet hiányos halmazban, mert minden ilyen csúcs szomszédsága S. T_1 viszont nem tartalmazhat hiányos halmazt, mert (S, T_1) batch kód. A kód mérete $tm + (n - tm)m = m - tm(m-1)$.

□

11. Összefoglalás

A szakdolgozatban a kombinatorikus batch kódok optimalizálásának témakörét jártuk körül. A dolgozat nagy része a paraméterek speciális állássai mellett mutatott be optimális batch kódokat. Pontosan, ismertettünk optimális vagy közel-optimális kódokat az $n \geq \binom{m}{k-2} - (m-k+1)A(m,4,k-3)$, valamint az $n = m, m+1, m+2$ esetekben. Továbbá, kimerítettük a $k \leq 4$ és $k = m$ eseteket is.

Az itt fel nem sorolt esetek nyitottak. Különösen érdekes például az $m = k-1$ és $n = m+3$ esetek, amelyekre a dolgozatban részletezett módszerek használatával el lehet indulni.

Szintén bemutatásra került a batch kódok tanulmányozásának számos különböző módszere, mint az optimalitástartó transzformációk, a G_2 segédgráf vizsgálata és a batch kód transzverzális matroidjának vizsgálata. Az is kiderült, hogy a batch kódok konstruálásában segíthet a bináris Hamming-kódok és a véges affin síkok struktúrája.

Végül kitértünk két alternatív feladatra: Az uniform batch kódok és az általános ($t \geq 1$) kombinatorikus batch kódok optimalizására. Az utóbbiról mindössze egy cikk ([11]) szól, melynek tartalmát teljes egészében közöltük a 10. fejezetben. Tehát sok $t = 1$-re vonatkozó tételek általánosítása nyitott feladat marad.

A kombinatorikus batch kódok vizsgálatának említendő ága $N(n,k,m)$ és $n(m,k,c)$ aszimptotikus viselkedésének leírása különböző paraméterekben. Az uniform kódok aszimptotikus viselkedésének meghatározásáról kiderült, hogy a Turán szám probléma egy speciális esete, így erre léteznek is eredmények. Ezzel a feladatkörrel nem foglalkoztunk a dolgozat során, de erről szólnak például a [1] és [12] cikkek.

Hivatkozások

[1] Balachandran N, Bhattacharya S. On an extremal hypergraph problem related to combinatorial batch codes, Discrete Applied Mathematics, 162 (2014), 373-380.

[2] Srimanta Bhattacharya, Sushmita Ruj and Bimal Roy, Combinatorial batchcodes: A Lower Bound and Optimal Constructions, arXiv:1102.4951v1 (2011).

[3] J.A. Bondy, Presentations of tranversal matroids, J. London Math. Soc., (2), 5 (1972), 289–292.

[4] J.A. Bondy, Transversal matroids, base orderable matroids, and graphs, Quart. J. Math. (Oxford), 23 (1972), 81–89.

[5] J.A. Bondy and D.J.A. Welsh, Some results on tranversal matroids and constructions for identically self dual matroids, Quart. J. Math. (Oxford), 22 (1971), 435-451.

[6] R.A. Brualdi, Transversal Matroids, Chapter 5 in Combinatorial Geometries, N. White ed., Cambridge Univ. Press, Cambridge, 1987, 72–97.

[7] R.A. Brualdi and G.W. Dinolt, Characterizations of transversal matroids and their presentations, J. Comb. Theory, 12 (1972), 268–286.

[8] R. A. Brualdi, K. P. Kiernan, S. A. Meyer and M. W. Schroeder, Combinatorial Batch Codes and Transversal Matroids, Adv. Math. Commun., 4 (2010), 419-431, Erratum ibid. p. 597.

[9] Cs. Bujtás and Zs. Tuza, Optimal batch codes: Many items or low retrievalrequirement, Adv. Math. Commun., 5 (2011), 529-541.

[10] Cs. Bujtás and Zs. Tuza, Optimal combinatorial batch codes derived from dualsystems, Miskolc Math. Notes 12 (1) (2011), 11-23.

[11] Cs. Bujtás and Zs. Tuza, Relaxations of Hall's Condition: Optimal batch codes with multiple queries, Applicable Analysis and Discrete Mathematics, 6 (1) (2012), 72-81.

[12] cs. Bujtás and Zs. Tuza, Turán numbers and batch codes, Discrete Applied Mathematics, 186 (2015), 45-55

[13] Y. Ishai, E. Kushiletitz, R. Ostrovsky and A. Sahai, Batch codes and theirapplications, In: Proceedings of the 36th Annual ACM Symposium on Theory of Computing, ACM Press, New York, 2004, 262–271.

[14] M. B. Paterson, D. R. Stinson and R. Wei, Combinatorial batch codes, Adv.Math. Commun., 3 (2009), 13–27.

[15] Natalia Silberstein and Anna Gál, Optimal Combinatorial Batch Codes based on Block Designs, Designs, Codes and Cryptography 78 (2) (2014) 409-424.