Újszerű párosítás feladatok

Önálló projekt II.
Beszámoló

Tóth Sára Hanna
Alkalmazott matematikus MSc

Témavezető:
Madarasi Péter
ELTE Matematikai Intézet,
Operációkutatási Tanszék

Eötvös Loránd Tudományegyetem
Természettudományi Kar
Budapest, 2022

1. Bevezetés

Az Önálló projekt tárgy keretei között a félévben folytattam a megszorított párosítás feladatokkal kapcsolatos kérdések vizsgálatát. Már a BSc szakdolgozatom középpontjában is ez a témakör állt, ennek köszönhetően pedig az előző félévben egy TDK dolgozat is született az eddigi eredményekből. Ebben a beszámolóban a vizsgált feladatok bemutatása után először a korábbi félévek eredményeit vázolom, majd bemutatom az idén vizsgált új irányokat, az ezekkel kapcsolatos észrevételeinket, valamint a további terveket.

2. A vizsgált feladatok

Feleltessük meg egy $G = (S, T; E)$ páros gráfban az S csúcsait egymást követő egész napos eseményeknek, T csúcsait biztonsági őröknek. Egy s nap és egy t őr között él megy, ha t beosztható az s napra. A feladat minél több naphoz hozzárendelni egy-egy őrt, ügyelve arra, hogy valamennyi őr minden egymást követő d napon legfeljebb egyszer dolgozhat. Ekkor egy megengedett beosztás a következő élhalmaznak felel meg [1]:

2.1. Definíció (d-távolságú párosítás)**.** *Adott egy $G = (S, T; E)$ páros gráf, melyre $S = \{s_1, s_2, \dots, s_n\}$, és egy $d \in \mathbb{Z}_+$ szám. Egy $M \subseteq E$ élhalmazt d-távolságú párosításnak (d-distance matching) nevezünk, ha az S-beli csúcsok foka M-ben legfeljebb* 1*, és minden $t \in T$-re, ha $s_i t, s_j t \in M$, akkor $|i - j| \geq d$ teljesül.*

A *maximális súlyú d-távolságú párosítás* feladatban adott egy $w : E \to \mathbb{R}_+$ súlyfüggvény, és olyan M d-távolságú párosítást keresünk, amelynek az összsúlya, azaz $w(M) = \sum_{e \in M} w_e$ maximális. A probléma a [2]-ben tárgyalt szimultán hozzárendelési feladat (simultaneous assignment problem) egy speciális esete, de megfogalmazható frekvencia kiosztási feladatként (frequency assignment problem) is, amely egy széles körben kutatott problémakör [3].

A feladat egy másik változata a *ciklikus d-távolságú párosítás* probléma, amely annyiban különbözik a fent definiált változattól, hogy az S csúcsait ciklikus sorrendben vesszük, tehát ekkor a fentiek mellett még az $|i - j| \leq |S| - d$ feltételnek is teljesülnie kell minden $s_i t, s_j t \in M$-re.

Mindkét problémát általánosíthatjuk úgy, hogy az inputban adott még a $b_S : S \to \mathbb{Z}_+$ fokszámkorlát függvény is, és olyan élhalmazt keresünk, amelyben minden $s \in S$ csúcs foka legfeljebb $b_S(s)$ lehet, valamint teljesíti a fenti távolságfeltételeket is. Ez a *(ciklikus) d-távolságú b_S-párosítás* feladat.

3. Eredmények

3.1. Közelíthetetlenség

A BSc szakdolgozatom középpontjában a d-távolságú párosítás feladat egy általánosítása, az úgynevezett k-szoros párosítás feladat állt. A dolgozat fő eredményeként bebizonyítottuk, hogy a 2-szeres párosítás feladatot NP-nehéz α-közelíteni, ha $\alpha < \frac{950}{949}$. A 2-szeres párosítás feladat visszavezethető a d-távolságú párosítás feladatra, így ennek a közelíthetetlensége is belátható.

3.1. Tétel. *A maximális d-távolságú párosítás feladatot NP-nehéz α-közelíteni, ha $\alpha < \frac{950}{949}$.*

A tétel következményeként kapjuk a ciklikus változat közelíthetetlenségét is.

3.2. Tétel. *A maximális ciklikus d-távolságú párosítás feladatot NP-nehéz α-közelíteni, ha $\alpha < \frac{950}{949}$.*

3.2. Egészértékűségi hézag

Az önálló projekt keretei között első félévben főként a d-távolságú párosítás feladattal foglalkoztunk. Az egyik fontos kutatási irány az egészértékűségi hézaggal kapcsolatos kérdések voltak.

Bebizonyítottuk az alábbi felső becslést a d-távolságú párosítás feladatra vonatkozó egészértékűségi hézagra:

3.3. Tétel. *Ha $d \geq 2$, akkor a maximális súlyú d-távolságú párosítás probléma bármely példányára felírt természetes IP feladatra és annak LP relaxáltjára*

$$\frac{OPT_{LP}}{OPT_{IP}} \leq 2 - \frac{2}{d}.$$

Ebből az is következik, hogy $d = 2$ esetben a természetes IP modell relaxáltjához tartozó poliéder egész, így polinomiális időben megoldható a feladat. A tétel bizonyításából egy $\left(2 - \frac{2}{d}\right)$-közelítő algoritmus is adódik.

A ciklikus változatra vonatkozó egészértékűségi hézagra egy oszthatósági feltétel mellett adtunk felső becslést, amelyről az is belátható, hogy éles.

3.4. Tétel. *A maximális súlyú ciklikus d-távolságú párosítás problémára, ha $(2d-1)$ osztója az S elemszámának, akkor a feladat természetes IP modelljére és annak LP relaxáltjára*

$$\frac{OPT_{LP}}{OPT_{IP}} \leq 2 - \frac{1}{d},$$

és ez a becslés éles.

A bizonyítás ebben az esetben is szolgáltat egy $\left(2 - \frac{1}{d}\right)$-közelítő algoritmust a megfelelő példányokon.

3.3. Optimális permutációk

A d-távolságú (b_S-)párosítás feladatok kapcsán természetesen adódó kérdés – akár a ciklikus, akár a nem ciklikus esetben –, hogy ha csak az adott, hogy melyik őr melyik eseményre osztható be (tehát egy páros gráf), akkor milyen sorrendben tartsuk meg az eseményeket ahhoz, hogy a lehető legtöbbre közülük találjunk őrt a fenti feltételeknek megfelelően. Ekkor tehát az S csúcshalmaz csúcsainak egy olyan permutációját keressük, amelyre a maximális (ciklikus) d-távolságú (b_S-)párosítás a lehető legnagyobb. A kérdés persze a súlyozott esetben is értelmes, ekkor olyan permutációt keresünk, amelyre a maximális súlyú (ciklikus) d-távolságú (b_S-)párosítás a legnagyobb.

Ezzel a kérdéskörrel főként az előző félévben foglalkoztunk. Beláttuk, hogy a $b_S \equiv 1$ esetben a feladatot polinom időben meg tudjuk oldani a ciklikus és a nem ciklikus esetben is, ami az alapfeladat nehézségének ismeretében meglepő eredmény.

3.5. Tétel. *Egy $G = (S, T; E)$ páros gráfra, $w : E \to \mathbb{R}$ költségfüggvényre és adott d számra polinom időben megtalálható az S csúcshalmaz csúcsainak egy olyan permutációja, amelyre a maximális súlyú (ciklikus) d-távolságú párosítás a legnagyobb.*

Azt is bebizonyítottuk, hogy $b_S \geq 2$-re mindkét feladat nehézzé válik. Pontosabban:

3.6. Tétel. *Ha minden $s \in S$-re $b_S(s) \geq 2$, akkor NP-nehéz annak eldöntése, hogy egy $G = (S, T; E)$ páros gráfban van-e az S halmaz csúcsainak olyan permutációja, amelyhez van teljes (ciklikus) d-távolságú b_S-párosítás. Teljesség alatt most az S-ben pontosan b_S-fokú élhalmazt értjük.*

3.4. Kicsit sértő megoldások

Egy új megközelítés a d-távolságú párosítás feladattal kapcsolatban, hogy – az approximációs algoritmusokkal ellentétben – nem olyan élhalmazt keresünk, amely az optimálisnál nem sokkal kisebb súlyú, de megengedett, hanem olyat, amely legalább olyan súlyos mint az optimum, de egy kicsit sértheti a feltételeket. A második félévben főként ezzel a kérdéskörrel foglalkoztunk.

Az intervallumfeltételek megsértésére több különböző mód is felmerül. Egyrészt próbálkozhatunk olyan élhalmaz keresésével, amely valamely d-nél kisebb d'-re egy megengedett d'-távolságú párosítás, a súlya pedig legalább akkora, mint az optimális d-távolságú párosítás súlya. Észrevettük, hogy ha kicsit módosítjuk a d-távolságú párosítás feladat nehézségének bizonyításához használt visszavezetést, akkor belátható, hogy a feladat az olyan páros gráfokra megszorítva is nehéz marad, amelyekben nincsen pontosan d-széles cseresznye (tehát olyan 2-hosszú út, amely valamely i-re s_i és s_{i+d} között megy). Egy ilyen gráfban pedig világos, hogy a megengedett $(d-1)$-távolságú párosítások, valamint a megengedett d-távolságú párosítások pontosan ugyanazok az élhalmazok. Tehát ha tudnánk olyan $(d-1)$-távolságú párosítást találni, amelynek a súlya legalább akkora, mint az optimális d-távolságú párosítás súlya, akkor így megkapnánk az optimális d-távolságú párosítást, ami nem lehetséges, hiszen ez egy NP-nehéz feladat. Ezt a gondolatmenetet általánosítva bebizonyíthatjuk a következőt:

3.7. Tétel. *Egy $G = (S, T; E)$ páros gráfra, melyre $S = \{s_1, \ldots, s_n\}$, $w : E \to \mathbb{R}$ költségfüggvényre és adott d számra, ha $1 \leq c < 2$, akkor NP-nehéz olyan $\frac{d}{c}$-távolságú párosítást találni, amely legalább olyan súlyos, mint a maximális súlyú d-távolságú párosítás.*

Ennél a megközelítésnél tehát csak olyan sértő megoldás polinom időben való megtalálásában reménykedhetünk, amely valamely $d' \leq \frac{d}{2}$-re d'-távolságú, és legalább olyan súlyos, mint az optimális d-távolságú párosítás. Továbbra is nyitott kérdés, hogy tudunk-e polinom időben ilyen élhalmazt keresni. A $d' = \frac{d}{2}$ esetben (amely a legnehezebb a $d' \leq \frac{d}{2}$ esetek közül) olyan $\frac{d}{2}$-távolságú megoldást tudunk találni – egy már korábban is használt módszerünk módosításával –, amelynek súlya legalább a $\frac{2}{3}$-része az optimális d-távolságú párosítás súlyának.

Továbbra is a kicsit sértő megoldások keresésénél maradva kiindulhatunk a d-távolságú párosítás feladat azon megfogalmazásából, miszerint olyan élhalmazt keresünk, amelyben minden $t \in T$ csúcsnak legfeljebb 1 szomszédja lehet az $S = \{s_1, \ldots, s_n\}$ minden d-hosszú intervallumából. Egy kicsit sértő megoldástól most azt követeljük meg, hogy minden d-hosszú intervallumból legfeljebb 2 szomszédja lehessen minden T-beli csúcsnak, azonban a súlya legyen legalább akkora, mint az optimális megoldás súlya. Ezzel a megközelítéssel kapcsolatban pozitív eredményről is beszámolhatunk, ugyanis ilyen élhalmazt polinom időben tudunk keresni, sőt, a kapott megoldásra az a kicsit erősebb állítás is teljesül, hogy minden T-beli csúcsnak minden $(d+1)$-hosszú intervallumból legfeljebb 2 szomszédja lehet.

3.8. Tétel. *Egy $G = (S, T; E)$ páros gráfra, melyre $S = \{s_1, \ldots, s_n\}$, $w : E \to \mathbb{R}$ költségfüggvényre és adott d számra polinom időben megtalálható egy olyan M élhalmaz, amelyben minden $t \in T$-re, és minden $i = 1, 2, \ldots, (n-d)$-re t-nek legfeljebb 2 szomszédja van az $\{s_i, s_{i+1}, \ldots, s_{i+d}\}$ halmazból, és M legalább olyan súlyos, mint a maximális súlyú d-távolságú párosítás.*

További természetes ötlet, hogy a távolságfeltételeket nem szeretnénk sérteni, azonban megengedjük, hogy az S-beli csúcsok foka 1-nél valamivel nagyobb legyen, és így keresünk legalább olyan súlyos megoldást, mint az azonosan 1 fokszámkorláttal kapott optimum. Ez a kérdés – sajnos – továbbra is a nyitott kérdések fejezetben kap helyet.

3.5. Egy általánosítás

A d-távolságú párosítás feladat azon általánosításával is foglalkoztunk, amelyben adott egy $g : T \times \mathcal{I} \to \mathbb{Z}_+$ függvény, ahol $\mathcal{I}$-vel az S-beli d-hosszú intervallumok halmazát jelöljük. Ebben az általánosabb feladatban olyan maximális súlyú élhalmazt keresünk, amelyben minden $t \in T$-nek az $I \in \mathcal{I}$ intervallumból legfeljebb $g(t, I)$ darab szomszédja lehet. Az eredeti d-távolságú párosítás feladat tehát a $g \equiv 1$ eset.

A feladatra adhatunk egy $(2 - \frac{1}{d})$-közelítő algoritmust abban az esetben, ha g monoton, ami alatt azt értjük, hogy minden $t \in T$-re ha az I intervallum kezdőpontja az I' intervallum kezdőpontjától balra van, akkor $g(t, I) \geq g(t, I')$ (tehát az idő előrehaladtával egyre jobban korlátozzuk, hogy legfeljebb mennyit dolgozhat egy őr).

4. További tervek

A témakörben még számos érdekes nyitott kérdés van, amelyekkel tovább szeretnénk haladni a következő félévben.

A kicsit sértő megoldásokkal kapcsolatban már említett kérdés, hogy tudunk-e valamely $c \geq 2$-re olyan $\frac{d}{c}$-távolságú párosítást keresni, amelynek a súlya legalább akkora, mint a maximális súlyú d-távolságú párosítás súlya. Érdekes lenne továbbá olyan kicsit sértő megoldást találni, amely a távolságfeltételeknek eleget tesz, de S-ben 1-nél valamivel nagyobb fokszámokat is megengedünk. Elképzelhető, hogy bizonyos fokszámkorlátokra ez a probléma NP-nehéz, ekkor persze ezt lenne jó bizonyítani. Felmerül az is, hogy kombináljuk az eddigi megközelítéseket, és megpróbáljunk olyan elég súlyos megoldást keresni, amely a fokszámkorlátokat és a távolságfeltételeket egyaránt sértheti egy kicsit (tehát például $\frac{d}{2}$-távolságú, és S-ben minden fokszám legfeljebb 2).

Vizsgálhatjuk tovább az egészértékűségi hézag becslését a d-távolságú párosítás feladatok kapcsán. A ciklikus változatban érdemes lenne (lehetőleg éles) becsléseket bizonyítani a $(2d-1) \nmid n$ esetekre is, a nem ciklikus változatra pedig javíthatnánk tovább a felső korlátot. Legjobban persze itt is egy éles becslést szeretnénk.

Egy lazábban kapcsolódó ismert probléma az úgynevezett nem-interferáló hálózati folyam feladat (non-interfering network flow problem) [4]. Itt olyan maximális folyamot keresünk egy hálózatban, amelyben ahelyett, hogy azt korlátoznánk, hogy egy adott élen keresztül mekkora folyamérték mehet, azt írjuk elő minden élre, hogy mekkora lehet a folyamérték az él "közelében". A probléma önmagában is érdekes, valamint a távolságfeltételekből adódó hasonlóság miatt érdemes lehet vizsgálni a d-távolságú párosítással való kapcsolatát is.

Ebben a félévben elkezdtük az eddigi eredményeket angolra fordítani, a cél, hogy egy cikk szülessen belőle. Ennek elkészüléséhez még sok munkára van szükség.

Hivatkozások

[1] P. Madarasi. Matchings under distance constraints I. *Ann Oper Res*, 305, 137–161, 2021.

[2] P. Madarasi. The Simultaneous Assignment Problem. arXiv:2105.09439, 2021.

[3] K. I. Aardal, S. P. M. Van Hoesel, A. M. Koster, C. Mannino, A. Sassano. Models and solution techniques for frequency assignment problems. *Ann Oper Res*, 153, 79–129, 2007.

[4] C. McDiarmid, B. Reed, A. Schrijver, B. Shepherd. Non-interfering network flows. *Scandinavian Workshop on Algorithm Theory. Springer, Berlin, Heidelberg. pp. 245-257.* 1992.