Opponensi vélemény

***Galambos Gábor*: Competitive Algorithms in Discrete Optimization**

című MTA doktori értekezéséről

Bevezetés

A dolgozat témája versenyképes algoritmusok diszkrét optimalizálási problémák megoldására a ládapakolástól kezdve az adattömörítésig. Az online algoritmusoknak kiemelten fontos szerepünk van a műben. Online algoritmust olyan számítási feladatra adunk, ahol a bemenet folyamatosan válik ismertté az algoritmus futása során, és az input beolvasásával egyidejűleg alakul ki az algoritmus eredménye. Egy klasszikus példa, ami a dolgozatban is központi szerepet játszik, a ládapakolási feladat, ahol különböző méretű elemeket kell ládákba helyezni, és amikor egy új elem érkezik, azonnal ládához kell rendelni (vagy egy korábban már megkezdett ládához, amiben még jut hely az új elemnek, vagy egy új, üres ládához). A cél az elemek minél kevesebb ládához rendelése. Az online algoritmusok működését többféle szempont szerint lehet elemezni, a legtöbbször a feladat optimumához hasonlítják az online algoritmus eredményét a legrosszabb eseteket nézve, amit versenyképességi elemzésnek is neveznek. Ugyanakkor a gyakorlati működést jobban mutatja az átlagos esetek vizsgálata, ami rendszerint egy technikailag nehezebb valószínűségi elemzést igényel. Mindkettőre van példa a dolgozatban.

A dolgozat a szerző munkásságának egy több mint harminc éves időszakába ad betekintést, a 80-as évektől napjainkig. A bemutatott eredmények az ezen időszakba eső cikkekből származnak. A régebbi eredmények jelentőségét, és hasznosságát jól mutatja a cikkekre kapott nagyszámú hivatkozás.

A dolgozat felépítése

A dolgozat összesen 9 fejezetből áll, az első fejezet bemutatja a területet, a vizsgált problémákat, az alkalmazott elemzési technikákat, és áttekintést nyújt a dolgozat hátralevő részéről. A többi 8 fejezet foglalja össze a szerző tudományos eredményeit. Mindegyik fejezet alapja egy vagy több, referált nemzetközi folyóiratban megjelent tudományos cikk. A dolgozat angol nyelven íródott, jó felépítésű.

Az eredmények értékelése

A 2. fejezet témája az 1-dimenziós ládapakolási feladat, és fő eredményei új alsó korlát számítási technikák online algoritmusok versenyképességére. Az új módszer alapja a 2.2. Tétel, ami van Vliet egy korábbi eredményének általánosítása. Van Vliet lineáris programozási technikák segítségével 1.54014-es alsó korlátot adott online ládapakolási feladatok aszimptotikus hibájára (André van Vliet (1992), An improved lower bound for on-line bin packing algorithms, Information Processing Letters, vol. 43, pp. 277-284). A 2.2. tételt használva egy új bizonyítás kerül bemutatásra, alkalmasan választott pakolási mintákat használva (2.3. Tétel). Továbbá, szintén a 2.2. tételt használva, 1.54014-nél kicsit erősebb alsó korlátot ad a szerző online ládapakolási algoritmusok versenyképességére. Az új eredmény alapja a pakolási minta technika, ahol összesen k különböző méretű elemekből álló lista alkalmas megválasztására volt szükség.

A szerző megemlíti, hogy a 2.4. tétel bizonyítása kivitelezhető volna van Vliet LP-alapú módszerével is, amivel az 1.54014-es alsó korlátot kapta (és a 2.3. tétel ezzel ekvivalens korlátot ad). Érdekes lett volna megvizsgálni, hogy mi a kapcsolat a dolgozatban kapott képlet (2.3. tételben a (13) formulába formálisan behelyettesítve a β_j, α_j, U_j paramétereket), és van Vliet 1992-es cikkében, a Theorem 4-ben kapott formula között.

A 3. fejezet témája az 1-dimenziós félig-online ládapakolási problémák. A 3.1. alfejezet fő eredménye egy, a korábbiaknál jobb alsó korlát arra az esetre, ha az inputban az elemek hossza csökkenő sorozatot alkot. A bizonyítás alapja ismét a 2.2. Tétel. A 3.2. alfejezet olyan algoritmus osztályt vizsgál, ahol legfeljebb k aktív láda lehet, és új elem elhelyezésekor megengedett az aktív ládák között az elemek mozgatása. A fő eredmény egy olyan, legfeljebb 3 aktív ládát használó eljárás, aminek versenyképessége megegyezik az alsó korláttal, tehát optimális! A 3.3. alfejezetben egy olyan modellt vizsgál a szerző, ahol egy új elem ládához rendelésekor legfeljebb k, már korábban elhelyezett elem mozgatható, másik ládába rakható. A szerző definiálja és elemzi a HR-k algoritmust, valamint alsó és felső korlátot ad annak versenyképességére. Az alsó és felső korlátok sajátossága, hogy k növelésével a távolságuk csökken. A 3.4. alfejezetben pedig egy technikailag nagyon szép alsó korlátot ad a k elem átpakolását megengedő modellben az algoritmusok versenyképességére.

A 4. fejezetben alsó korlátokat ad a szerző online algoritmusok versenyképességére 2-dimenziós ládapakolási problémák esetében. Egy egyszerűbb alsó korlát bemutatása után, Van Vliet alsó korlát eredményét használva egy egy lényegesen erősebb alsó korlát kerül ismertetésre. Érdemes megjegyezni, hogy a Theorem 4.3 bizonyításában a szerző a Theorem 2.3-ra hivatkozik, ami szerintem inkább a Theorem 2.2, de mivel $\alpha_j = 1$, ezért lényegében a Theorem 2.1 speciális eset is elegendő.

Az 5. fejezet ládapakolási, illetve ládafedési algoritmusok véletlenségi elemzését tartalmazza, ahol az elemek nagyságát véletlen változók határozzák meg egy adott eloszlás szerint. Az 5.1 fejezet az 1 dimenziós-, az 5.2 a két-dimenziós ládapakolási feladattal, míg az 5.3 fejezet az 1-dimenziós ládafedési feladattal foglalkozik, tehát ahol minden ládába a kapacitást meghaladó összméretű elemeket kell rakni, és a "túltöltött" ládák számát maximalizálni kell. Az 5.1 fejezetben a "Next Fit Decreasing" algoritmus által használt ládák várható értéke, ill. az algoritmus várható versenyképessége kerül meghatározásra abban az esetben, ha az elemek mérete egyenletes eloszlású a 0-1 intervallumon, majd egy becslést kapunk az algoritmus által használt ládák eltérésére a várható értéktől, végül egy határeloszlás tétel kerül bizonyításra. Az elemzések természetesen erősen kihasználják, hogy az elemek egyenletes eloszlásúak. Az 5.2 fejezetben a "Hybrid Next Fit" algoritmus kerül elemzésre. Az elemzési technika hasonló az 5.1 fejezetben bemutatotthoz. A Theorem 5.3, ami a fő eredmény a [43]-as cikkben jelent meg, jelen dolgozatban mégis a [48]-hoz köti a szerző. Végül az 5.3 fejezetben 3 algoritmus kerül elemezésre, de mindenek előtt az optimum várható értékére ad felső becslést a szerző.

A 6. fejezetben a párhuzamos gépes ütemezési problémához online algoritmusok versenyképességére a gépek számától függő alsó korlátok kerülnek ismertetésre, majd pedig a jól ismert $2-1/m$ felső korlátot javítása következik $(2 - 1/m - \varepsilon_m)$ -re, ahol m a gépek száma, és $\varepsilon_m \to 0$, ha $m \to \infty$. A szerző meg is jegyzi, hogy a cikk megjelenése után egy egész sor új eredmény született, 2-nél konstanssal kisebb versenyképességgel tetszőleges számú gépre, illetve kevés gép esetén is javítva a korábbiakat.

A 7. fejezet témája egy kétgépes open shop ütemezési probléma, ahol minden feladat megmunkálási ideje 1, a feladatoknak van egy legkorábbi kezdési ideje, egy határideje, és a cél a késő feladatok számának minimalizálása. Egy új polinomiális idejű algoritmus kerül bemutatásra. Az algoritmus érdekessége, hogy a gráfelméletből ismert 2-faktor probléma egy általánosítására vezeti vissza az ütemezési problémát.

A 8. fejezet témája ismét egy ütemezési probléma, amit magyarra talán "feladat páros" ütemezési problémának lehetne fordítani. Ebben a modellben feladat párokat kell beütemezni egyetlen gépen, és a művelet pár két tagja között előre adott időnek kell eltelni, se többnek, se kevesebbnek. Egy olyan speciális esetre mutat pszeudo-polinomiális idejű algoritmust a szerző, ahol minden feladat párban az első feladatok hossza ugyanaz az a szám, a második feladatok hossza ugyanaz a b szám, és a két feladat távolsága ugyanaz az L szám. A cél a maximális befejezési idő minimalizálása. Az ismertetett algoritmus dinamikus programozáson alapul, és futási ideje exponenciális L-ben. Ugyanakkor a mellékelt futási eredmények szerint gyorsan lefutott már a 90-es évek személyi számítógépein is, akár $n = 1000$ feladat pár esetén. Megjegyezzük, hogy a fenti eredményt tovább fejlesztette Philippe Baptiste a "A note on scheduling identical coupled tasks in logarithmic time, Discrete Applied Mathematics 158 (2010) 583–58." cikkében a futási időt $O(\log n)$-re csökkentette fix a, b, L esetén. A feladat bonyolultságát ugyanakkor ő sem tudta megválaszolni.

A 9. fejezetben online adattömörítési eljárások versenyképességét elemzi a szerző. A fejezetben 4 különböző algoritmus versenyképességét vizsgálja meg 4 különböző kódkönyvtár tekintetében. Ez a fejezet önmagában is igen gazdag eredményekben.

Összegzés

A bemutatott eredmények a diszkrét optimalizálási problémák széles köréből kerülnek ki, az algoritmikus megközelítések között találunk online, gráfelméleti, illetve jó értelemben vett ad-hoc megoldásokat. Az elemzési technikák tekintetében is igen gazdag eszköztárat vonultat fel a szerző (pakolási minták, LP-alapú elemezés, valószínűségszámítási meggondolások, gráfelméleti bizonyítások, kódoláselméleti technikák).

Az eredmények magas színvonalúak, és elegendőnek tartom azokat az MTA doktori cím megszerzéséhez, a védés kitűzését javaslom.

Kis Tamás, PhD.

Budapest, 2017. május 10.