

OTKA Zárójelentés

Nyilvántartási szám: 81493

Témavezető: Tuza Zsolt

Elvégzett kutatásainkból összesen 47 publikációt készítettünk. Ezek közül 28 SCI folyóiratban jelent meg vagy lett elfogadva, és további 6 került benyújtásra hasonló fórumokhoz; ez utóbbiaknak az impakt faktorát egyelőre sajnos nem tudtuk megadni az EPR felületen, de SCI folyóiratokban fognak megjelenni ezek is, az OTKA azonosító szám feltüntetésével.

Eredményeink között több olyan is van, amellyel évtizedek óta nyitott problémákat oldottunk meg.

Az alábbiakban témakörök szerint csoportosítva foglalkozunk össze elvégzett kutatásainkat.

Folyamattervezés és optimalizálás

Az informatika üzleti alkalmazásinak jelentős része az üzleti folyamatok támogatása. Ha ezen szoftverek optimalizáló képességekkel bírnak, akkor alkalmazásuk hasznossága jelentősen nőhet az optimalizáló algoritmusok minőségétől függően. Ezirányú kutatásaink során olyan algoritmusokat és szoftvereket dolgoztunk ki, melyek folyamattervezés és optimalizálás során a matematikai programozással szinergiában gráf algoritmusokat is futtatnak. A gráf algoritmusok és matematikai programozási feladat megoldási lépései kölcsönösen hasznosítják egymás eredményeit, így javítva eredményességüket. Ezen modellek kiterjesztésre kerültek többcélú optimalizálásra, a költség mellett célként tekintve a szintetizált folyamatok megbízhatóságát, időigényét, és környezeti terhelését. Mindemellett újabb eredményeink születtek szabványos üzleti folyamat leírások formális modellekké alakításában. Így számos hozzájárulást tettünk ahhoz, hogy a számítógéppel támogatott üzleti folyamattervezés és lebonyolítás kiegészíthető legyen garantált optimumra törekvő optimalizáló modulokkal. A javasolt modulok alternatív megoldásokat is biztosítanak, és figyelembe vesznek időbeli, megbízhatósági és fenntarthatósági szempontokat is.

A megbízhatóságra törekvés támogatására eljárásokat alkottunk, melyek egyrészt – a szintézis közbülső fázisában – becslik a felépítés alatt álló folyamat megbízhatóságát, másrészt redundanciák beépítésével automatikusan növelik annak értékét. A folyamat időigényének számításához szintén kiterjesztettük a folyamatok gráf modelljét, hogy a párhuzamos folyamatok időbeli szinkronizáltságát vagy annak hiányát meg tudjuk különböztetni.

A környezeti terhelés mérése a tervezett folyamatok fenntarthatósági elemzése szempontjából elengedhetetlen. Ennek elsősorban a gyakorlati jelentősége nagy, például szállítmányozásban, amely az üvegházhatású gázok legnagyobb és ma is növvő hányadát adja világszerte. A javasolt és piaci alkalmazásban is tesztelt jármű hozzárendelés optimalizáló eljárás a korábbi módszernél akár 25%-kal kedvezőbb megoldást adott, költség és széndioxid kibocsátás szempontjából is.

Az ipari gyártórendszerek ütemezésben korábban kettéváltak a matematikai programozásra és gráf reprezentáció alapú megoldó algoritmusok. Ezeknek egyfajta szinergiáját alkottuk meg egy szuperstruktúra alapján generálható vegyes egész matematikai programozási modell kidolgozásával, mely hagyományos MILP megoldókkal és gráf leírást használó megoldókkal is megoldható.

Publikáltuk jármű hozzárendelési feladatok átfogalmazását folyamatszintézis feladatra. Az eredményül kapott leírás előnye, hogy egyrészt vegyes egész matematikai programozási feladatra is konvertálható, másrészt gráfalgoritmusokkal alternatív (n-legjobb) jármű hozzárendelés is generálható.

Üzleti folyamatok és ellátási láncok tervezésekor gyakran időkorlátokat is figyelembe kell venni. A láncok egyes lépései adott megvalósításban időben függenek egymástól, míg más megvalósításban nem. Ha lánc tervezésekor minden potenciális összefüggést egyértelműen egy gráfon szeretnénk reprezentálni, akkor a hálózat szerkezetét leíró gráfot további csúcsokkal és élekkel kell bővíteni. A folyamatok szerkezetét leíró gráfok automatikus bővítésére javasolunk algoritmust. Eredményképpen, egyértelműen leírható, hogy egyazon közbülső eredményeket előállító és felhasználó lépések közül melyeknek kell megvárnia egymást és melyeknek nem.

A metabolikus hálózatok elemzése, bár matematikai szempontból nagyon hasonlít a reakcióút-azonosításra, más megoldó módszereket használ, melyek főként konvex elemzésre épülnek annak érdekében, hogy a reakció lépéseinek irányát (nem megfordíthatóságát) is figyelembe lehessen venni. Formálisan bizonyítottuk, hogy a metabolikus hálózatokhoz generált úgynevezett „extrém utak” illetve a lineáris algebrai fogalmakkal definiált „direkt utak” ekvivalensek a gráf leírás alapján szintetizált „strukturálisan minimális utakkal”. Ennek az algoritmusok továbbfejlesztési lehetőségei miatt van jelentősége. Az eredmény fontossága, hogy lineáris algebrai és gráf leírás alapú megoldó módszerek integrációja ezután elméletileg is indokolható.

Közöltünk továbbá egy módszert, ami biztosítja, hogy folyamathálózatok alternatív struktúráinak generálása közben olyan redundáns struktúrák is szülessenek, melyek robusztussága bizonyítottan nagyobb, mint a minimális költségű megoldásoké.

Hálózatok partíciói

Különböző gyakorlati problémákhoz kapcsolódóan vizsgáltuk a hálózatok csúcs (csomópont) illetve él (link) halmazainak partícióit.

Első modellünkben („ k -improper C-színezés”) olyan csúcspartíciókat vizsgáltuk, ahol egy-egy csúcs szomszédságában csak korlátozottan fordulhatnak elő más partíciós osztályba sorolt csúcsok. Közvetlenül látható, hogy ez a tulajdonság alkalmas olyan problémák modellezésére, amikor egy hálózatban bizonyos processzorok között direkt kommunikációs kapcsolatot kell biztosítani úgy, hogy minden processzor legfeljebb egy rögzített k számú tőle eltérő típusú protokollt használó processzorral állhat kapcsolatban.

A hálózat különböző paramétereinek függvényeként alsó és felső korlátokat bizonyítottunk a partíciós osztályok elérhető maximális számára vonatkozóan. Bizonyítottuk, hogy a partíciós osztályok maximális számára Nordhaus-Gaddum-típusú korlát érvényes, továbbá hogy elegendően nagy r esetén az r -dimenziós hiperkocka csúcshalmazán vett maximum a dimenzió nagyságától függetlenül 2^k . Szintén vizsgáltuk a maximum meghatározásának

algoritmikus nehézségét. Ez számos hálózatosztályra nézve NP-nehéznek bizonyult, más esetekre pedig polinomiális idejű algoritmusokat tudtunk megadni.

Másrészt bizonyítottuk, hogy a vizsgált partíciós feltételek szoros kapcsolatban állnak a hálózatok bizonyos tulajdonságú elvágó élhalmazainak létezésével. Ezért az általunk bevezetett k -improper C -színezés modellezi azt, hogy egy összefüggő hálózat mekkora és hány részre eshet szét, ha csomópontonként legfeljebb k kapcsolódás hibásodik meg.

Analóg összefüggés alapján vezettük be és vizsgáltuk a hálózatok olyan élpartícióit, ahol bármely élhez legfeljebb egyik végpontjában csatlakozhatnak eltérő partíciós osztályhoz tartozó linkek. Ez az úgynevezett „3-c-élszínezés” többek között azon probléma modelljéül szolgál, amikor egy hálózatban a meghibásodások egymással közvetlen kapcsolatban nem álló csomópontokban lépnek fel. A partíciós osztályok maximális számára bizonyítottunk alsó és felső korlátokat a hálózatot jellemző paraméterek függvényében, továbbá a maximum meghatározásának algoritmikus nehézségét vizsgáltuk különböző gráfosztályokon.

Vizsgálataink során igazoltuk, hogy egy hálózat 3-c-csúcspartíciói és a linkek metszés-gráfjának 3-c-élpártíciói bijektív módon megfeleltethetők egymásnak (amennyiben minden csúcs fokszáma legalább 2). Ennek következményeként bizonyos típusú elvágó csúcs- és élhalmazok között kapcsolatokat mutattunk ki.

Az előző modellek releváns kérdése a partíciós osztályok maximális számára vonatkozott. A további tanulmányozott problémaköröknél olyan csúcs- és élpártíciókat tekintettünk, amelyekben a cél a felbontás osztályainak számát minimalizálni, illetve listaszínezések esetén minél rövidebb listák mellett is garantálni a megoldás létezését. A „proper connection” az utóbbi években sokat vizsgált „rainbow connection” probléma rokona, ily módon a biztonságos adattovábbítási protokollok motivációjából indult ki. Számos becslést adtunk és bizonyos feltételek mellett a pontos optimumokat is meghatároztuk.

Adatkezelés kombinatorikus optimalizálása

A „kombinatorikus batch kód” fogalmát 2004-ben vezette be Ishai, Kushilevitz, Ostrovsky és Sahai. Ez a privát információ visszakeresésnek (PIR) olyan kombinatorikus modellje, ahol az információt tároló szerverek nem kommunikálnak egymással, továbbá adott n számú információ-darabot úgy kell tárolnunk m számú szerveren (egy darabot akár több helyen is), hogy bármely legfeljebb k információ-darab visszanyerhető legyen oly módon, amikoris bármely szerverről legfeljebb t számú darabot olvasunk ki összességében. Cél, hogy a szerverek $N(n,k,m,t)$ optimális (vagyis minimális) összterheltségét meghatározzuk a paraméterek függvényében.

A szakirodalomban megjelent első cikkek a $t = 1$ esetet tárgyalták. Első két cikkünkben mi is erre az esetre bizonyítottunk új eredményeket. Pontos képleteket adtunk meg az n és m paraméterek függvényében a szerverek $N(n, 3, m, 1)$ és $N(n, 4, m, 1)$ minimális összterheltségére, vagyis azokra az esetekre, amikor a visszanyerendő adatok száma nem lépi túl a $k = 3$ illetve a $k = 4$ értéket. Továbbá pontos képletet bizonyítottunk arra, amikor az adatok száma elegendően nagy a szerverek számához képest. Másrészt új optimális konstrukciókat terveztünk és olyan transzformációkat írtunk le, amelyek egy optimális adatelrendezésből újabb optimális elrendezéseket állítanak elő. Ezek segítségével azon esetek is jól kezelhetővé váltak kombinatorikus módszerekkel, ahol az adatok száma csak kicsivel haladja meg a szerverek számát.

A kutatás következő fázisában az általános esetet vizsgáltuk, ahol a t paraméter értéke tetszőleges pozitív egész lehet. Képletet bizonyítottunk az elérhető minimális összterheltség $N(n, k, m, t)$ értékére a $k < 3t + 1$ feltétel mellett, továbbá azokra az esetekre is, ahol az információ-darabok n száma elegendően nagy a többi paraméter értékéhez képest.

A kombinatorikus batch kódok uniform változata a teljesítendő kikötéseket azzal egészíti ki, hogy minden információ-darabot pontosan r különböző szerveren kell tárolnunk. Adott n , k , r és $t = 1$ pozitív egészek esetén cél a feltételek figyelembe vételével tárolható információ-darabok maximális $m(n, k, r)$ számának meghatározása illetve becslése. Az optimalizálási feladat felfogható egy hipergráfokra vonatkozó Turán-típusú problémaként is. Ebből kiindulva egy általánosabb kérdésfelvetést vizsgáltunk, ahol az eredményeink speciális eseteként az $m(n, k, r)$ értékére is a korábban ismerteknél szignifikánsan jobb felső korlátot kaptunk.

Hálózatok lefogása csomópontokkal és domináló részek

A hálózatok minimális méretű domináló halmazainak elmélete kiterjedt szakirodalommal és fontos alkalmazásokkal bír. Új területnek számít, de hasonlóan fontos az informatikai problémák modellezésében ennek a halmazrendszerekre (hipergráfokra) vonatkozó változata. Vagyis egy alaphalmazon bizonyos részhalmazok vannak kijelölve és a cél az, hogy minimális számú elemet válasszunk ki (domináló halmaz) az alaphalmazból úgy, hogy bármely ki nem választott elem közvetlen kapcsolatban álljon (vagyis a halmazrendszerben közös kijelölt részhalmazban legyen) valamely domináló halmazbeli elemmel. A problémát olyan halmazrendszerekre vizsgáltuk, ahol a halmazrendszerben kijelölt halmazok mindegyike legfeljebb k elemet tartalmaz. Többfajta általános felső korlátot adtunk a minimális domináló halmaz méretére (dominálási szám), többek között szoros összefüggést mutattunk ki a dominálási számra és egy másik fontos paraméterre, a csúcslefogási számra vonatkozó felső korlátok között. Továbbá, teljes körűen jellemeztük a $k = 2$ és $k = 3$ esetekre vonatkozóan a csúcs- és élszám függvényében felírható lineáris felső korlátokat.

A csúcslefogási és a dominálási szám kapcsolatára vonatkozóan vizsgáltuk még azokat az eseteket is, amikor a két paraméter megegyezik a halmazrendszerekben, vagyis amikor valamely csúcslefogás egyúttal minimális méretű domináló halmaz is lesz. A probléma „öröklődő” változatát is tanulmányoztuk; bizonyítottuk, hogy ha a kiválasztott részalmazok méretére korlátot adunk, akkor csak véges sokfajta részstruktúrát kell kitiltani a halmazrendszerből az egyenlőség „öröklődő” teljesüléséhez. Többfajta komplexitási eredményt igazoltunk a lefogási és a dominálási szám egyenlőségének eldöntési problémájára vonatkozóan.

Emellett tanulmányoztuk a lokálisan optimális lefogásokat, a halmazrendszerek többszörös lefogásának egy általános modelljét, a halmazrendszerek úgynevezett totális dominálását, a hálózatok dominálásának játékelméleti megközelítését, az adott távolságú (általánosított) dominálást, valamint azt, hogy adott számú csomópont meghibásodása mekkora mértékben módosíthatja az optimális lefogások ill. komplementáris szempontból nézve a legnagyobb független halmazok méretét. Utóbbiakra erős nem approximálhatósági eredményeket is bizonyítottunk.

Általános modellek

2007-ben vezettünk be egy igen általános színezési modellt „stably bounded” (más jelöléssel STAB) hipergráf néven. Ebben olyan csúcspartíciókat vizsgálunk, ahol az alaphalmaz kijelölt részein (a hipergráf élein) színkorlátozó függvényekkel előírt feltételeket kell teljesíteni a metszett partíciós osztályok számára és a legnagyobb bemetszés méretére vonatkozóan. Ezzel a kikötésrendszerrel modellezhető például a klasszikus frekvenciakiosztási probléma, a hibatűrő rendszerek tervezésénél felmerülő többirányú feltételrendszer, és kiemelt speciális esetként a sokszálú optikai hálózatok hullámhossz-kiosztási problémája (wavelength assignment problem [WAP] in multifiber optical networks).

A fenti általános partíciós probléma megoldhatóságának algoritmikus (idő-) komplexitását vizsgáltuk a színkorlátozó függvények lehetséges kombinációira intervallum-hipergráfok és hiperfák osztályain. Bizonyos strukturális és a használt függvényekre vonatkozó megszorítások esetére polinomiális idejű algoritmusokat terveztünk, bővebb osztályokra pedig kimutattuk, hogy a feladat egy magasabb komplexitási osztályba tartozik.

A STAB hipergráfoknak a történetileg talán legfontosabb részosztályára, a „vegyes hipergráfok” osztályára vonatkozóan megoldottunk egy régi (az 1990-es évek közepéről származó) problémát arra vonatkozóan, hogy milyen összefüggés mutatható ki az osztályt leíró kétféle korlátozás (C-él, D-él) felcserélése során adódó két struktúra között. A megoldásban kulcsszerepet játszik egy új, polinom idejű algoritmus a hipergráfoknak egy részosztályán.

A csak C-típusú éleket tartalmazó hiperfák osztályán a partíciós osztályok maximális számának meghatározását illetően approximálhatósági és nem-approximálhatósági eredményeket bizonyítottunk, továbbá terveztünk egy – korlátozott foksám esetén polinomiális futásidejű – algoritmust a paraméter pontos meghatározására és egy optimális partíció előállítására. Ezen kívül elkezdtük tanulmányozni halmazrendszerek felbonthatóságát paritási feltételek mellett is, a vegyes hipergráfokkal rokon szemléletben.

Egy másik általános modellünk a „graph-bin packing”, ami egy sor nevezetes optimalizálási problémát kezel egységes keretek között. Ebben adott egy feltételeket kifejező gráf, csúcsain súlyokkal, élein pedig alsó-felső korlátokkal. A gráf csúcsait egy hálózat csomópontjaihoz kell rendelni úgy, hogy az egyes csomópontokra előírt kapacitás-korlát teljesüljön, továbbá minden esetben az első gráf élein adott korlátok között legyen a végpontokhoz rendelt csomópontok távolsága a hálózatban. Ezzel kapcsolatban az egyik kérdés (pakolhatóság) a feltételeknek megfelelő pakolás létezésére, a másik kérdés pedig (pakolhatóság esetén) az optimális, minél kisebb részt elfoglaló pakolás nagyságára vonatkozik. Ez a modell közvetlenül alkalmazható például az ütemezési problémák és a frekvenciakiosztási problémák vizsgálatánál.

Szükséges illetve elégséges feltételeket adtunk meg pakolhatóságra, korlátokat bizonyítottunk az optimális pakolás méretére vonatkozóan, különböző feltételek mellett vizsgáltuk az optimális megoldás előállításának illetve adott arányú approximálásának (idő-) komplexitását. A gráf köreibben szereplő élkorlátokra és a csomópontok kapacitásaira teljesülő meghatározott feltételek mellett polinomiális futásidejű algoritmust adtunk a pakolásra. A csomópontok kapacitásait elegendően nagynak feltételezve, a körök négyzetrács-hálózatba pakolhatóságát teljesen jellemeztük, továbbá lineáris futásidejű algoritmust terveztünk az optimális (bizonyos esetekben majdnem optimális) pakolás előállítására.

Hálózatok teljes részei közti összefüggések

Egy hálózatban egy k -pontú teljes részgráf k darab szoros kapcsolatban levő, vagyis páronként összekötött csúcsot tartalmaz. Három páronként összekötött csúcsot háromszögnek nevezünk; míg ha a hálózat háromszögeit csúcsokkal reprezentáljuk, amelyek közül kettő akkor van összekötve, ha a reprezentált háromszögek egyik „oldala” közös, a hálózat háromszöggráfját kapjuk. Ha ugyanezt a reprezentációt analóg módon k -pontú teljes részgráfokra alkalmazzuk, akkor az úgynevezett k -élgráfokat kapjuk.

Egyik cikkünkben azokat a háromszöggráfokat és k -élgráfokat jellemeztük, amelyek két nemtriviális összefüggő gráf Descartes-szorzataként állnak elő. Ezzel háromszöggráfokra megoldottunk egy Bagga által 2004-ben felvetett problémát, továbbá a karakterizációt is

felhasználva komplexitási eredményeket bizonyítottunk a háromszöggráfok, a k -élgráfok és az úgynevezett k -anti-Gallai gráfok felismerhetőségére vonatkozóan.

Másik ide tartozó kutatásunkban a projekt vezetőjének 30 év óta nyitott (és intenzíven kutatott) sejtését vizsgáltuk és bizonyítottuk különböző típusú hálózatokra. A sejtés szerint, ha egy hálózatban nem jelölhető ki t -nél több páronként élfüggetlen háromszög, akkor legfeljebb $2t$ darab éllel a hálózat összes háromszöge lefogható.

Kémiai rendszerek kezelése kombinatorikus módszerekkel

Gráfelméleti és lineáris programozási módszerek alkalmazásával vizsgáltuk kémiai reakció-hálózatok differenciálegyenleteinek struktúráját és hatékony algoritmust adtunk „gyengén reverzibilis” megoldások megtalálására (valamint annak igazolására, ha nincs ilyen megoldás). Ily módon egy három évtizede nyitott kérdésre is választ adtunk. Nem optimalizálási feladatot old meg, de kémiai egyenletek szerkezetének vizsgálatára vonatkozó alkalmazása miatt itt említhető az a cikkünk is, amelyben lineáris egyenletrendszerek minimális megoldásait tanulmányozzuk.

Adatok partíciói

Az e témában tanulmányozott kombinatorikus optimalizálási problémák a ládapakolásként ismert (és a fentebb említett „graph-bin packing” formában általánosított) feladat tárgyköréhez kapcsolódnak, annak különféle feltételek melletti szigorításai. Alapvetően három megközelítésben vizsgáltuk az algoritmusok elérhető hatékonyságát: ha a teljes input ismert (offline); ha az input mindenkor következő darabjának érkezésekor visszavonhatatlan döntést kell hozni, de semmi információ nem áll rendelkezésre a későbbi darabokról (online); illetve e két modell között, ha a döntéseket a teljes információ birtokában, de adott sorrendben kell meghozni (restricted offline).

Az eredmények közül itt azt a terjedelmes munkát emeljük ki, amely a „First Fit Decreasing” eljárás egészen pontos legrosszabb-eset vizsgálatát adja és az ezirányú kutatásoknak négy évtizede kezdődött sorát zárja le.