

Biológiai Motivációjú Komplex Rendszermodellek Struktúrális Analízise és Identifikálhatóságának Vizsgálata

Tézisfüzet



Szlobodnyik Gergely

Témavezető:
Prof. Szederkényi Gábor

Pázmány Péter Katolikus Egyetem
Információs Technológiai és Bionikai Kar
Roska Tamás Műszaki és Természettudományi Doktori
Iskola

2020

1 Bevezetés

Számos dinamikus rendszermodellhez rendelhető grafikus reprezentáció, amelyben a csomópontok jelölik a rendszer elemi komponenseit az élek pedig a funkcionális relációkat az egyes komponensek között [1, 2, 3]. Például egy génregulációs hálózatban a csomópontok jelölik az egyes géneket, az élek pedig a gének között kapcsolatokat, úgy mint aktiváció és inhibíció [6]. Egy dinamikus rendszer matematikai modellje tartalmazhatja a rendszer grafikus reprezentációját a differenciál vagy differencia egyenlet paramétereiben. Lineáris rendszermodellek esetén a rendszermátrix (állapotátmenet mátrix) az egyes komponensek interakciós mintázatát a nem-diagonális elemeiben kódolja [7]. Nem-lineáris rendszerek esetén a dinamikus egyenletek szintén hordozhatnak információt a rendszer hálózatos struktúrájáról, a kinetikus rendszerek esetén például bizonyított, hogy a differenciálegyenlet rendszer paraméterei közvetlen kapcsolatban állnak a rendszer hálózati struktúrájával [4, 5]. Az 1. ábrán illusztráljuk, hogy hálózatos struktúra és dinamikus modell (differenciál vagy differenciaegyenlet rendszer) szintén reprezentálhatja ugyanazon biokémiai folyamatot. Azt is jelöltük az ábrán, hogy a differenciálegyenlet rendszerből a hálózatos struktúra modellje is megkonstruálható.

A disszertáció különböző rendszermodellek struktúrális analízisével és identifikálhatóságával foglalkozik. A munkám fő motivációja, hogy algoritmikus módszereket adjunk biológiai háttérű rendszermodellek struktúrális analízisére, identifikálhatóságára és realizációelméleti vizsgálatára úgy, hogy a konstruált algoritmusok helyességét matematikai eszközökkel bizonyítsuk. Munkám során a hálózatos struktúra és a dinamikus viselkedés kapcsolatát is vizsgálom különböző rendszermodellek, mint például diszkrét idejű lineáris időinvariáns rendszerek és diszkrét állapotterű kémiai reakcióhálózatok (Petri hálók) esetén. A kutatómunkám fontos aspektusa annak a vizsgálata, hogy adott dinamikus rendszernek létezik-e több, struktúrálisan különböző, de dinamikusan ekvivalens realizációja. A kérdést lineáris rendszereken vizsgálva tervezek olyan algoritmust, amely képes megtalálni struktúrálisan különböző dinamikusan ekvivalens realizációit egy adott rendszernek. A struktúrális identifikálhatóság kérdését szintén vizsgálom. Egy dinamikus rendszermodellt struktúrálisan indentifikálhatónak nevezünk, ha ugyanazon dinamikus viselkedés különböző parametrizáció mellett nem lehetséges. Ha egy rendszerosztály esetén a paraméterek információt hordoznak a rendszer hálózatos struktúrájára vonatkozóan, akkor a struktúrális identifikálhatóság segítséget nyújthat a struktúra (topológikus)

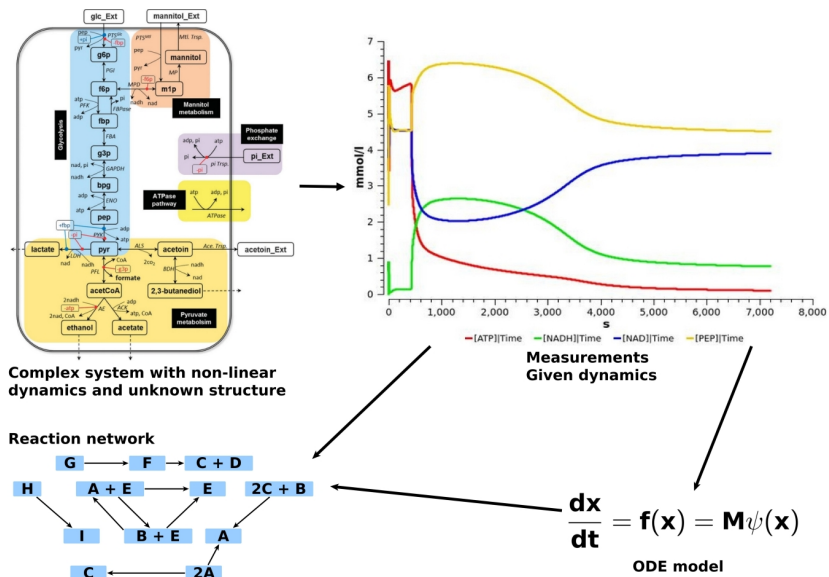


Figure 1: Komplex biológiai rendszerek különböző reprezentációinak illusztrációja. Mérési adatok alapján a rendszer dinamikus viselkedését leíró differenciálegyenlet rendszer és a funkcionális komponensek közötti interakciókat reprezentáló hálózatos struktúra is rekonstruálható.

egyértelműségének vizsgálatában. Mivel biológiai rendszermodellekben gyakran megjelenik időkéseletetés, külön feladatként kitűzöm az időkéseletetést tartalmazó dinamikus rendszerek struktúrális identifikálhatóságának vizsgálatát. Bevezetem az együttes struktúrális identifikálhatóságot, amely a konstans időkéseletetéseket és a közöséges rendszerparamétereket együttesen kibővített paramétervektoroknak tekintve azt vizsgálja, hogy a rendszer struktúrálisan identifikálható-e.

2 Elméleti háttér és Alkalmazott módszerek

2.1 Struktúrális identifikálhatóság

Tekintsünk egy folytonos idejű egy bemenetű egy kimenetű rendszert.

$$M(t, \theta, u, y, p) = 0, \quad (1)$$

ahol u a bemeneti gerjesztést, y pedig a rendszerválaszt jelöli. u és y tetszőleges számú konstans időkéseletetést is tartalmazhat, azaz

$$\begin{aligned} u &= \{u(t), u(t - T_1^u), \dots, u(t - T_{k_u}^u)\}, \\ y &= \{y(t), y(t - T_1^y), \dots, y(t - T_{k_y}^y)\}, \end{aligned} \quad (2)$$

ahol k_u és k_y a késleltetést tartalmazó komponensek számát a bemenetre és a kimenetre vonatkozóan. θ jelöli a rendszer paramétereinek halmazát, amely az időkéseletetéseket is tartalmazza:

$$T_1^u, \dots, T_{k_u}^u, T_1^y, \dots, T_{k_y}^y \in \mathbb{R}. \quad (3)$$

Feltételezzük, hogy θ a kezdeti feltételektől és az bemeneti u jeltől független. p jelöli a differenciáloperátort. A rendszeroperátor $M(\cdot)$ analitikus. Megjegyezzük, hogy az analitikus feltétel a rendszeroperátoron nem szigorú korlátozás, ugyanis számos biológiai és kémiai szempontból fontos rendszerosztály kizárólag analitikus nemlinearitásokat tartalmaz. $M(\cdot)$ tartalmazhat a θ paraméter elemeire vonatkozóan is nemlineáris elemeket.

Az (1) egyenlettel definiált kimenet-bemenet modellstruktúrához jutunk, ha egy nemlineáris állapotter modellen differenciálgberai átalakítások segítségével elimináljuk az állapotváltozókat [8, 9].

Egy identifikációs procedúra alkalmazása előtt hasznos megvizsgálni, hogy elméleti szempontból lehetséges-e egyértelműen meghatározni a rendszer paramétereit.

Definíció Az (1) egyenlettel jellemzett modellt strukturálisan globálisan identifikálhatónak nevezzük, ha

$$y(\theta) = y(\hat{\theta}) \Rightarrow \theta = \hat{\theta} \quad (4)$$

teljesül minden θ -ra, ahol $y(\theta)$ jelöli a rendszer kimenetét, amelyet θ -val parametrizáltunk.

Ha a fenti definíció egy bizonyos θ érték lokális $\mathcal{V}(\theta)$ környezetében teljesül, akkor azt mondjuk, hogy a rendszer lokálisan strukturálisan identifikálható θ körül. Ha a rendszer nem identifikálható, akkor a rendszer nem-identifikálhatónak nevezzük. Az identifikálhatóságra bevezetett definíciókat a paraméterhalmaz θ tetszőleges nem-üres $\bar{\theta} \subset \theta$ részalmazára is értelmezzük.

Megjegyezzük, hogy mivel a konstans időképletetéseket is beleértetjük a paraméterek θ halmazába, így a fenti identifikálhatósági definíciókat is együttesen értjük a közönséges rendszerparaméterekre és a konstans időképletetésekre.

A strukturális identifikálhatóság a rendszermodellre jellemző elméleti tulajdonság, amelyet a dinamikus modell (differenciálegyenlet) alakja határoz meg, független a kezdeti feltételtől és a mérési adatok minőségétől.

2.2 Lineáris időinvariáns rendszerek realizációelméleti vizsgálata

Egy diszkrét idejű lineáris időinvariáns rendszert a $\Theta = (A, B, C, D)$ állapotter realizációs mátrixokkal, mint paraméterekkel adunk meg és a kapcsolódó differenciaegyenlet rendszer a következő alakban adható meg:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k), & x(0) &= x_0, \\ y(k) &= Cx(k) + Du(k), \end{aligned} \quad (5)$$

ahol $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$ és $D \in \mathbb{R}^{p \times m}$. $x(k) \in \mathbb{R}^n$ jelöli az állapotváltozók vektorát, $u(k) \in \mathbb{R}^m$ és $y(k) \in \mathbb{R}^p$ a rendszer bemeneti és kimeneti jelei.

Noha a dinamikus (kimenet-bemenet) viselkedés, amelyet a rendszer mutat adott Θ parametrizáció mellett egyértelmű, a paraméterek, amelyekkel a rendszeroperátort megadjuk nem feltétlenül egyértelműek. Lehetséges, hogy léteznek különböző Θ és Θ' parametrizációk, amelyek mellett a fenti rendszer ugyanazon kimenet-bemenet viselkedést mutatja, amely egyértelműen azt jelenti, hogy a rendszermodell strukturálisan nem-identifikálható. Nem-identifikálható rendszermodell es-

etén hasznos lehet a matematikailag lehetséges parametrizációk halmazának leírása. A lehetséges parametrizációk halmazának kvantifikálása segítséget jelenthet mérnöki vagy biológiai szempontból releváns realizációk megtalálásában.

Definíció Egy $\Theta' = (A', B', C', D')$ parametrizációval adott rendszer dinamikus ekvivalens realizációja a (4) egyenlettel adott Θ halmazzal parametrizált rendszernek, ha Θ' -val parametrizálva a rendszer ugyanazon dinamikus (kimenet-bemenet) viselkedést produkálja, azaz $y(k|\Theta') = y(k|\Theta)$ minden $u(k)$, $k \geq 0$ bemeneti jel esetén.

Rekurzívan visszafejtve a 4-es egyenletet az állapotváltozókat kiküszöbölhetjük és a következő kimenet-bemenet egyenletet kapjuk

$$y(k) = CA^k x(0) + \sum_{i=0}^{k-1} Y_{k-i-1} u(i) + Du(k), \quad (6)$$

ahol $Y_{k-i-1} = CA^{k-i-1}B$ és D a rendszer Markov paraméterei, amelyek a kimenet-bemenet viselkedést egyértelműen karakterizálják és invariánsak tetszőleges invertálható állapottér transzformációra. Miután a Markov paraméterek – szemben az állapottér realizációs mátrixokkal – egyértelműek, szükséges és elégséges feltételeket tudunk konstruálni segítségükkel a dinamikus ekvivalenciára vonatkozóan: $\Theta' = (A', B', C', D')$ egy dinamikus ekvivalens realizációja a $\mathcal{Y} = \{Y_k = CA^k B\}_{k \geq 0}$ Markov paraméterekkel adott rendszernek, ha $Y_k = C' A'^k B'$ teljesül minden $k \geq 0$ és $D' = D$ esetén.

A strukturális nem-identifikálhatósághoz szorosan kapcsolódó, praktikus szempontból releváns probléma, annak vizsgálata, hogy léteznek-e különböző $A, A' \in \mathbb{R}^{n \times n}$ állapottár realizációs mátrixok, amelyek a nem-nulla nem-diagonális elemeik mintázatában különböznek egymástól. Feltételezve, hogy a 4-es egyenlet egy olyan hálózatos struktúrával rendelkező rendszert ír le, amelyben az A állapotátmeneti mátrix nem-diagonális elemei reprezentálják az egyes funkcionális elemek (állapotváltozók) közötti interakciókat, a különböző dinamikus ekvivalens realizációk létezése különböző nem-nulla nem-diagonálisokat tartalmazó A, A' mátrixok esetén a rendszer komponenseinek különböző kapcsolódási mintázatát jelenti, vagyis ugyanazon dinamikus viselkedés különböző interakciós mintázatok segítségével is megvalósítható.

2.3 Egészértékű programozási és döntési probléma

Egy egészértékű lineáris programot formálisan a következőképpen adhatunk meg:

$$ILP \begin{cases} \min_x \{a^\top x\} \\ \text{subject to} \\ Ax \leq b \\ x \in \mathbb{Z}^n \end{cases} \quad (7)$$

ahol x jelöli a döntési változók n -dimenziós vektorát, $a \in \mathbb{Z}^n$, $A \in \mathbb{Z}^{m \times n}$ és $b \in \mathbb{Z}^m$ pedig rögzített együtthatók. Általános esetben a fenti számítási probléma NP-nehez, amely különösen nagy dimenziószám mellett gyakorlati szempontból korlátozhatja egészértékű optimalizációs problémák hatékony megoldását.

Ha a célfüggvényt minimalizálandó/maximalizálandó döntési változók értéke nem releváns, csak a számítási probléma megoldhatósága, azaz egy $x \in \mathbb{Z}^n$ döntési változó létezése, amely az összes kényszer kielégíti, akkor a következő egészértékű döntési problémává redukálódik a kiindulási optimalizációs feladat:

$$FP \begin{cases} P = \{x \mid Ax \leq b, A \in \mathbb{Z}^{m \times n}, b \in \mathbb{Z}^m, x \in \mathbb{R}^n\} \\ P \cap \mathbb{Z}^n \stackrel{?}{=} \emptyset \end{cases} \quad (8)$$

Az egészértékű döntési probléma azt vizsgálja, hogy egy egész értékű (rács)pont a P politópban helyezkedik-e el, formálisan: $P \cap \mathbb{Z}^n \stackrel{?}{=} \emptyset$. Noha az egészértékű döntési probléma NP-nehez, a döntési változók számában, a konstansok számában és az együtthatómátrixok elemeinek abszolút értékének maximumában kifejezve létezik hatékony időkomplexitású algoritmikus eljárás. Feltételezve, hogy a döntési probléma n dimenzionalitása rögzített, egész értékű megoldás létezése polinomiális időkomplexitással eldönthető a kényszerfeltételek m számában és az együtthatómátrixok elemeinek abszolút értékének maximumában a Lenstra algoritmus segítségével [10, 11]. Továbbá a kényszerfeltételeket kielégítő egész (rács)pontok száma szintén polinom időben meghatározható m -ben és az együtthatók abszolútértékének maximumában a Barvinok algoritmus segítségével [12, 13, 14, 15]. Megjegyezzük, hogy a Barvinok algoritmusnak létezik hatékony szoftveres implementációja [16].

Szintén megjegyezzük, hogy az egészértékű lineáris program speciális esetben szintén megoldható polinomiális időkomplexitással: ha A és b egész együtthatókat tartalmaznak és A teljesen unimoduláris mátrix, akkor az egészértékű kényszerek elhagyhatóak, azaz garantált, hogy a lineáris program optimuma valós értékű döntési változó értékészlet

mellet is egész értékű lesz.

2.4 Diszkrét állapotú reakcióhálózatok

Egy diszkrét állapotú kémiai reakcióhálózatot (discrete Chemical Reaction Network, d-CRN) n különböző anyaggal, m komplexszel és l reakcióval egy $\mathcal{N} = (\mathcal{S}, \mathcal{C}, \mathcal{R})$ hármassal adunk meg a következő módon:

$$\mathcal{S} = \{s_i \mid i = 1, \dots, n\}$$

$$\mathcal{C} = \{y_j = \sum_{i=1}^n \alpha_{ji} s_i \mid s_i \in \mathcal{S}, \alpha_{ji} \in \mathbb{Z}_{\geq 0}, i = 1, \dots, n, j = 1, \dots, m\}$$

$$\mathcal{R} = \{r_v = y_{source(r_v)} \rightarrow y_{product(r_v)} \mid y_{source(r_v)}, y_{product(r_v)} \in \mathcal{C}, v = 1, \dots, l\}$$

ahol s_i az i -edik anyag, y_j a j -edik komplex és r_v a v -edik reakció a reakcióhálózatban. α_{ji} jelöli a sztöchiometriai együtthatóját az i -edik anyagnak a j -edik komplexben. Egy $r_v = y_{source(r_v)} \rightarrow y_{product(r_v)} \in \mathcal{R}$ reakciónak $y_{source(r_v)}$ és $y_{product(r_v)}$ a kiindulási- és termék-komplexét jelöli.

Minden $y_j \in \mathcal{C}$, $j \in \{1, \dots, m\}$ esetén, a sztöchiometriai együtthatók megadhatóak vektoros formában

$$\bar{y}_j = [\alpha_{j1} \ \alpha_{j2} \ \dots \ \alpha_{jn}]^\top. \quad (9)$$

Minden $r \in \mathcal{R}$ reakció esetén egy $r_{ij} \in \mathbb{Z}^n$ reakcióvektort tudunk r -hez rendelni, amely kódolja a reakció végbemenetele esetén az egyes anyagok megváltozását:

$$r_{ij} = \bar{y}_j - \bar{y}_i, \quad (10)$$

úgy hogy y_j és y_i jelölik r -nek a kiindulási-és termék-komplexét. Az r_i jelölést használjuk i -edik reakcióra és a reakcióhoz tartozó reakcióvektorra is. Feltételezzük, hogy a vizsgált reakcióhálózatokban a reakcióknak egy rögzített r_1, r_2, \dots, r_l , $l = |\mathcal{R}|$ sorrendje adott.

Egy d-CRN szintén reprezentálható egy $G = G(V, E)$ irányított gráf segítségével, amelyben a csúcspontok a komplexeknek, az élek pedig a reakcióknak felelnek meg olyan módon, hogy az él a kiindulási komplexet reprezentáló csomópontból mutat a termék-komplexet reprezentáló csomópontba.

$$V = \mathcal{C} \quad (11)$$

$$E = \mathcal{R} \quad (12)$$

Minden élhez egy súlyt is rendelhetünk amely a megfelelő reakciósebességi állandójával egyenlő.

Egy reakcióhálózathoz egy sztöchiometriai mátrixot is rendelhetünk, amely a hálózat algebrai reprezentációjában nyújt kvantitatív segítséget.

Definíció Tekintsünk egy $\mathcal{N} = (\mathcal{S}, \mathcal{C}, \mathcal{R})$ kémiai reakcióhálózatot. A rendszerhez tartozó sztöchiometriai mátrix $\Gamma \in \mathbb{Z}^{n \times l}$ a következőképpen definiálható:

$$\Gamma = [r_1 \ \dots \ r_l] \quad (13)$$

$[\Gamma]_{ij}$ kódolja s_i anyagban történő változást az r_j reakció hatására.

Az $X(t) \in \mathbb{Z}_{\geq 0}^n$ állapotvektor jelöli az egyes anyagok mennyiségét a reakcióhálózatban a t -edik időpillanatban. A diszkrét értékkészlet miatt $X(t)$ elemeiben pontosabb az anyagmennyiségek helyett molekulaszámot említeni. A reakcióhálózat időbeli viselkedése a következő diszkrét állapotegyenlet segítségével adható meg:

$$X(t) = X(0) + \Gamma N(t) \quad (14)$$

ahol $X(0)$ a rendszer kezdeti állapota. $N(t) = [N_1(t), N_2(t), \dots, N_l(t)]^\top \in \mathbb{Z}_{\geq 0}^l$, $N_k(t) \in \mathbb{Z}_{\geq 0}$ kódolja, hogy a t -edik időpillanatig az egyes reakciók hányszor játszódtak le. Megjegyezzük, hogy $N(t)$ -t általában (inhomogén) Poisson folyamatok segítségével modellezzük [18, 19].

Szintén megjegyezzük, hogy a fent vázolt diszkrét állapotterén értelmezett modellje a kémiai reakcióhálózatoknak ekvivalens az elméleti számítás-tudományban fontos modellosztállyal, a Petri hálókkkal.

3 Új tudományos eredmények

I. Tézis Egy új módszert terveztem időkésleltetést tartalmazó nemlineáris dinamikus rendszerek struktúrális identifikálhatóságának vizsgálatára. [J1].

Az egy bemenetű egy kimenetű nemlineáris konstans időkésleltetésű dinamikus rendszerek Volterra-sor reprezentációját alkalmazva adtam algebrai feltételt a konstans időkésleltetések és a közönséges rendszerparaméterek együttes struktúrális identifikálhatóságára. A Volterra magfüggvények frekvenciatartománybeli reprezentációit, az általánosított frekvenciaválasz függvényeket használva megmutattam, hogy nemlineáris algebrai egyenletek egyértelmű megoldhatóságára vezethető vissza a struktúrális identifikálhatóság vizsgálata.

II. Tézis Megmutattam, hogy a diszkrét idejű lineáris időinvariáns rendszerek lehetséges állapotátmeneti mátrixainak halmaza konvex, ha a B , C és D állapotter realizációs mátrixok rögzítettek. Kihasználva a lehetséges állapotátmenet mátrixok halmazának konvexitását, konvex optimalizáción alapuló algoritmikus eljárást adtam struktúrálisan különböző n -edrendű állapotter realizációk meghatározására. [J2].

Induktív módon bebizonyítottam, hogy a lehetséges állapotátmenet mátrixok \mathcal{A} halmaza konvex, feltéve, hogy a B , C és D mátrixok rögzítettek és a kimeneti egyenlethez tartozó C mátrix invertálható. Megmutattam, hogy a lehetséges állapotátmenet mátrixok \mathcal{A} halmazának konvexitását kihasználva konstruálható konvex optimalizációs algoritmus, amelynek segítségével különböző dinamikusan ekvivalens állapotter realizációk határozhatóak meg, melyeknek rendje a kiindulási $\Theta = (A, B, C, D)$ realizációjával egyenlő. A kinetikus rendszerek elméleti eredményeit felhasználva adtam algoritmust, amely struktúrálisan különböző n -edrendű állapotter realizációk meghatározására alkalmas.

III. Tézis Diszkrét állapotterű kémiai reakcióhálózatok elérhetőségi problémájának eldöntésére alkalmaztam egy számítási módszert, amelyben új felső korlátot adtam a körmentes állapotátmenet szekvenciák maximális hosszára [J3].

Egy egészértékű döntési probléma formájában írtam le diszkrét állapotterű kémiai reakcióhálózatok elérhetőségi problémáját, úgy, hogy a döntési változók maximális számára új felső korlátot adtam. A módszer a Lenstra algoritmuson alapszik, amelynek segítségével egészértékű döntési problémákat polinomiális időkomplexitással lehet eldönteni, feltéve, hogy a döntési változók száma rögzített. Új felső korlátot adtam a körmentes állapotátmenet szekvenciák maximális hosszára szubkonzeratív reakcióhálózati struktúrát feltételezve.

Megmutattam, hogy az elérhetőségi reláció teljesülése ekvivalens a reakcióhálózat diszkrét állapotegyenletének nemnegatív egész értékű megoldásának létezésével szubkonzeratív reakcióhálózatok esetén, feltéve hogy az állapotter kétdimenziós.

Thesis IV. Meghatároztam egy feltételrendszert, amely mellett a diszkrét állapotterű kémiai reakcióhálózatok elérhetőségi problémája ekvivalens az állapotegyenlet nemnegatív egész értékű megoldásának létezésével. Ilyen módon egy egészértékű döntési feladatot határoztam meg, amelynek feltétel-mátrixa

a kémiai reakcióhálózat sztöchiometriai mátrixa. Bebizonyítottam, hogy az egészértékű döntési probléma egy polinomiális időkomplexitású lineáris programozási feladat formájában megoldható [J4, J7].

1. Ismeretes, hogy általános esetben szubkonzervatív kémiai reakcióhálózatok állapottere korlátos, szuperkonzervatív hálózatok állapottere viszont nem korlátos, megszámlálhatóan végtelen számosságú. Bizonyítást adtam arra, hogy szuperkonzervatív hálózatok elérhetőségi problémája visszavezethető szubkonzervatív reakcióhálózatok elérhetőségi problémájára (elérhetőségi ekvivalencia), amely garantáltan véges állapotteren értelmezett.

Szub-és szuperkonzervatív kémiai reakcióhálózatok esetére a hálózati struktúrára vonatkozó feltételeket adtam, amelyek mellett az elérhetőségi reláció teljesülése ekvivalens a reakcióhálózat diszkrét állapotegyenletének nemnegatív egész értékű megoldásának létezésével. A bizonyításban a reakcióhálózatokra vonatkozó kényszerfeltételek nem tartalmazznak az állapotter dimenzionalitására vonatkozó kényszert. Az ekvivalencia maga után vonja, hogy az elérhetőségi reláció egy egészértékű döntési problémaként fogalmazható meg és a Lenstra algoritmus alkalmazható annak megoldására. [J4].

2. Megmutattam, hogy a kémiai reakcióhálózatok sztöchiometriai mátrixa teljesen unimoduláris tulajdonságú a IV/1 tézis feltételei mellett. Ismeretes, hogy egészértékű problémák teljesen unimoduláris feltételi mátrix esetén lineáris programozási feladatként megoldhatóak. Ilyen módon megmutattam, hogy a d-CRN elérhetőségi problémája visszavezethető egy lineáris programozási feladat megoldására, amely polinomiális időben megoldható [J7].

4 Alkalmazási lehetőségek és további kutatómunka

A disszertációban bemutatott elméleti és számítási eredményeknek fő célja, hogy biológiai háttérű rendszermodellek struktúrális (hálózatos) tulajdonságainak és identifikálhatóságának vizsgálatát hatékonyan tudjuk elvégezni. A kutatómunka szignifikáns aspektusa volt a dinamikus (differenciál-és differenciaegyenlet alapú) és hálózatos (topológiai, gráf alapú) tulajdonságok közötti kapcsolat vizsgálata. A bemutatott eredmények fő alkalmazási területe biológiai rendszermodellek kvantitatív

elemzése, mivel számos biológiai rendszer jellemezhető dinamikus modellekkel és hálózatos struktúrával is. A struktúrális identifikálhatósági eredményeket, amelyeket az első tézispont tartalmaz, alkalmazni lehet tetszőleges paraméterbecslési módszer előtti lépésként egy dinamikus rendszerek identifikációs vizsgálatában, mivel segítségével azonosíthatjuk azon paramétereket (beleértve a konstans időkésltetéseket is), amelyek nem identifikálhatóak, tehát elméletileg sem lehetséges egyértelműen meghatározni az értéküket. A realizációelméleti eredmények, amelyek a második tézispontban szerepelnek szoros kapcsolatban állnak az identifikálhatósággal, ugyanis a tervezett algoritmusok ugyanazon dinamikus rendszermodell paramétereiben különböző, dinamikusan ekvivalens realizációinak meghatározására alkalmasak, segítségükkel a lokális struktúrális nem-identifikálhatóság bizonyítható. A realizációelméleti eredmények szintetikus biológiai szempontból is jelentőséggel bírnak, segítségükkel lehetséges meghatározni struktúrálisan különböző dinamikusan ekvivalens realizációkat, amelyek közül ki lehet választani a biokémiai szempontból legalkalmasabbat, illetve szintetikus szempontból implementálhatóakat. A diszkrét állapotterű kémiai reakcióhálózatokra vonatkozó új eredmények alkalmasak arra, hogy számításos eszközökkel vizsgáljuk alacsony molekulaszámú molekuláris hálózatok tulajdonságait. Az eredményeknek a szintetikus biológia egy potenciális alkalmazási területe, mivel ismeretes, hogy funkcionális szintetikus biokémiai hálózatok tervezési feladata (gate implementability problem) visszavezethető a kémiai reakcióhálózatok elérhetőségi relációjának vizsgálatára [20]. Fontos megjegyezni továbbá, hogy a diszkrét állapotterű kémiai reakcióhálózatok formális modellje ekvivalens az elméleti számítástudományból ismeretes, széles körben alkalmazott Petri hálókkkal, ezért a bemutatott eredmények számítástudományi szempontból is alkalmazhatóak például elosztott rendszerek verifikációjára, konkurens és párhuzamos rendszerek elemzésére.

A következő új kutatási irányokat jelölöm meg, amelyek a fentebb vázolt elméleti és számításos eredményekhez szorosan kapcsolódnak:

1. Időkésltetést tartalmazó dinamikus rendszerek struktúrális identifikálhatósága: az általánosított frekvenciaválasz függvények alkalmazásával algebrai feltételt tudunk adtunk a rendszermodellek és konstans időkésltetések együttes struktúrális identifikálhatóságára. A bemutatott módszer egy lehetséges kiterjesztése annak vizsgálata, hogy létezik-e, illetve megállapítható-e olyan felső korlát a minimálisan megvizsgálendő általánosított frekvenciaválasz függvények számára vonatkozóan, amely mellett szük-

séges és elégséges feltétele garantált az együttes struktúrális identifikálhatóságnak. Megjegyezzük, hogy létezik rekurzív algoritmus az egymást követő frekvenciaválasz függvények kiszámítására [21]. Egy rekurzív algoritmus vizsgálata esetleg segítséget nyújthat a szükséges frekvenciaválasz függvények maximális számának meghatározásához, feltéve, hogy egy felső korlát létezése elméletileg bizonyított.

2. Diszkrét idejű lineáris időinvariáns rendszerek összes struktúrálisan különböző realizációjának számításos módszerekkel történő meghatározása: a disszertációban bemutatott sajátérték beágyazási módszer segítségével egy olyan rendszerhez jutunk, amelynek garantáltan véges sok nem-nulla Markov paramétere létezik, ezért optimalizációs algoritmust adhatunk különböző realizációk meghatározására, melyekből az eredeti – nem feltétlenül véges sok Markov paramétert tartalmazó – rendszernek különböző realizációi is származtathatóak. A bemutatott módszer praktikus és elméleti szempontból is releváns kiterjesztése lehet a sajátérték beágyazási módszer helyett egy olyan módszer alkalmazása, amelynek segítségével elméletileg garantálható, hogy az algoritmusok megtalálják az összes struktúrálisan különböző n -edrendű dinamikus ekvivalens realizációját egy adott diszkrét idejű lineáris időinvariáns rendszernek.
3. A diszkrét állapotterű kémiai reakcióhálózatok elérhetőségi analízisének kiterjesztése általánosabb kémiai reakcióhálózat osztályok esetére: a disszertáció a reakcióhálózati struktúrához kapcsolódó feltételeket ad, amelyek mellett az elérhetőségi reláció ekvivalens a rendszer diszkrét állapotegyenletének nemnegatív egész értékű megoldásának létezésével. A bemutatott eredményeknek egy lehetséges kiterjesztése annak a vizsgálata, hogy létezik-e, illetve algoritmikusan meghatározható-e olyan gráftranszformáció, amely egy adott reakcióhálózati struktúrát olyan módosított struktúrájává transzformál, amely már garantáltan kielégíti a bemutatott tételek feltételeit és bizonyítható, hogy a transzformált hálózatra adódó elérhetőségi reláció pontosan akkor teljesül, amikor az eredeti reakcióhálózat elérhetőségi relációja (reakcióhálózatok elérhetőségi ekvivalenciája).

A Szerző impakt faktoros folyóiratcikkei

- [J1] G. Szlobodnyik, G. Szederkényi, "Structural Identifiability Analysis of Non-linear Time Delayed Systems with Generalized Frequency Response Functions", *Kybernetika*, Vol. 57(6), pp. 939-957., 2021., **Impact Factor: 0.892**
- [J2] G. Szlobodnyik, G. Szederkényi, "Computing Different Realizations of Linear Dynamical Systems with Embedding Eigenvalue Assignment", *Acta Cybernetica*, Vol. 25(3), pp. 585-611., 2022., **Impact Factor: 0.38**
- [J3] G. Szlobodnyik, G. Szederkényi, M. Johnston, "Reachability analysis of subconservative discrete chemical reaction networks", *MATCH Commun. Math. Comput. Chem.*, Vol. 81(3), pp. 705-736., 2019., **Impact Factor: 2.126**
- [J4] G. Szlobodnyik, G. Szederkényi, "Reachability Analysis of Low-Order Discrete State Reaction Networks Obeying Conservation Laws", *Complexity*, Vol. 2019, Article ID 1035974, 13 pages, 2019., **Impact Factor: 2.591**
- [J5] B. Ács, G. Szlobodnyik, and G. Szederkényi, "A computational approach to the structural analysis of uncertain kinetic systems," *Computer Physics Communications*, Vol. 228, pp. 83–95, 2018., **Impact Factor: 3.309**
- [J6] G. Szlobodnyik, L. Farkas, "Data Augmentation by Guided Deep Interpolation", *Applied Soft Computing*, Vol. 111., 2021., **Impact Factor: 6.725**
- [J7] G. Szlobodnyik, G. Szederkényi, "Polynomial Time Reachability Analysis in Discrete State Chemical Reaction Networks Obeying Conservation Laws", benyújtva *MATCH Commun. Math. Comput. Chem.*

A Szerző további folyóiratcikkei

- [J8] G. Szederkényi, B. Acs, G. Szlobodnyik, "Structural analysis of kinetic systems with uncertain parameters", *IFAC-PapersOnLine*, Vol. 49., Issue 24., 2016, pp. 24-27.

- [J9] M. Vághy, G. Szlobodnyik, G. Szederkenyi, "Kinetic realization of delayed polynomial dynamical models", IFAC-PapersOnLine, Vol. 52(7), pp. 45-50., 2019.

A Szerző konferenciáinak

- [C1] G. Szlobodnyik, "Extracting Activity Patterns from Altering Biological Networks: A Sparse Autoencoder Approach", IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), 2019.
- [C2] G. Szlobodnyik, G. Szederkényi, "Reachability analysis of discrete state epidemiological models", In: H., Cherifi et al. - Complex Networks 2018. The 7th International Conference on Complex Networks and Their Applications. Book of Abstracts Cambridge: University of Cambridge, pp. 119-121., 3 p., 2018.
- [C3] G. Szlobodnyik, G. Szederkényi, "Reachability analysis of discrete state chemical reaction networks obeying conservation laws", Bioinspired Analysis of Dynamical Systems and Protein Network Analysis, Esztergom, 2019.

Hivatkozások

- [1] E. D. Kolaczyk, "Statistical Analysis of Network Data. Methods and Models", Springer, New York, 2009.
- [2] L. Chuang, et. al. "Computational network biology: Data, models, and applications", Physics Reports, Vol. 846., pp. 1-66., 2020.
- [3] E. Sontag, A. Kiyatkin, B.N. Kholodenko, "Inferring dynamic architecture of cellular networks using time series of gene expression, protein and metabolite data", Bioinformatics, Vol. 20(12), pp. 1877-1886., 2004.
- [4] G. Szederkényi, J. R. Banga, A. A. Alonso. "Inference of complex biological networks: distinguishability issues and optimization-based solutions", BMC systems biology, Vol. 5(1), 177., 2011.
- [5] G. Szederkényi, A. Magyar, K. Hangos, "Analysis and control of polynomial dynamic models with biological applications", New York, NY: Academic Press, 2018.

- [6] U. Alon, "An Introduction to Systems Biology: Design Principles of Biological Circuits", Chapman & Hall, CRC, 2007.
- [7] H. G. Frøysa, H. J. Skaug, G. Alendal, "Experimental design for parameter estimation in steady-state linear models of metabolic networks", *Mathematical Biosciences*, Vol. 319., 108291., 2020.
- [8] L. Ljung, T. Glad, "On global identifiability for arbitrary model parametrizations", *Automatica*, Vol. 30(2), pp. 265-276.,
- [9] T. Glad, "Nonlinear state space and input-output descriptions using differential polynomials", In Descusse, J., M. Fliess, A. Isidori and D. Leborgne (Eds), *Lecture Notes in Control and Information Science*, Vol. 122., Springer Berlin.
- [10] H. W. Lenstra, "Integer Programming with a Fixed Number of Variables", *Math. Oper. Res.*, 8, pp. 538–548., 1983.
- [11] M. Grötschel, L. Lovász, A. Schrijver, "Geometric Algorithms and Combinatorial Optimization", Springer-Verlag, Algorithms and Combinatorics 2, 1988.
- [12] A. J. Barvinok, "A Polynomial Time Algorithm for Counting Integral Points in Polyhedra When the Dimension is Fixed", *Math. Oper. Res.*, 19., pp. 769–779., 1994.
- [13] M. Köppe, "A primal Barvinok algorithm based on irrational decompositions", *SIAM Journal on Discrete Mathematics*, Vol. 21(1), pp. 220–236., 2007.
- [14] S. Verdoolaege, et. al., "Counting Integer Points in Parametric Polytopes Using Barvinok's Rational Functions", *Algorithmica*, Vol. 48(1), pp. 37–66., 2007.
- [15] J. B. Lasserre, "Integer programming, Barvinok's counting algorithm and Gomory relaxations", *Operations Research Letters*, Vol. 32(2), pp. 133-137., 2004.
- [16] J. A. De Loera et. al., "Effective lattice point counting in rational convex polytopes", *J. Symb. Comput.*, 38., pp. 273–1302., 2004.
- [17] T. G. Kurtz, "The Relationship between Stochastic and Deterministic Models for Chemical Reactions", *The Journal of Chemical Physics*, 57(7), 1972., 2976-2978.

- [18] D. F. Anderson et. al., "Stochastic analysis of chemical reaction networks with absolute concentration robustness", *J. R. Soc. Interface*, 11(90), 2014.
- [19] D. F. Anderson, T. G. Kurtz, "Continuous time Markov chain models for chemical reaction networks", In: H. Koepl et. al. (eds) "Design and analysis of biomolecular circuits: engineering approaches to systems and synthetic biology", Springer, Berlin, 2011., 3–42.
- [20] A. Condon, et. al., eds. "Algorithmic Bioprocesses", Springer Science and Business Media, 2009.
- [21] H. Zhang, S. A. Billings, Q. M. Zhu, "Frequency response functions for nonlinear rational models", *International Journal of Control*, Vol. 61., pp. 1073–1097., 1995.