XVI. FIATAL MŰSZAKIAK TUDOMÁNYOS ÜLÉSSZAKA

Kolozsvár, 2011. március 24–25.

GRÁF-ELMÉLETI TOPOLÓGIAI DESZKRIPTOROK DISZKRIMINÁCIÓS KÉPESSÉGÉNEK ELEMZÉSE

dr. RÉTI Tamás, dr. BITAY Enikő

Abstract

In the fields of chemical graph theory and structural chemistry the applications of graphs to the description of quantitative structure-property relationships implies the representation of molecules by so-called molecular descriptors often referred to as topological indices. The topological indices which often have a direct structural interpretation are numerical parameters of a connected undirected graph which characterize its topology. In present study, we have performed comparative tests on a set of simple biregular graphs to analyse and assess the discrimination performance of various topological descriptors.

Key words:

topological descriptors, biregular graphs, Zagreb indices

Összefoglalás

A kémiai gráfelméletben, illetve a szerkezeti kémiában elterjedten alkalmaznak gráfokat a molekul szerkezet és a tulajdonságok közötti összefüggések leírására, és e célra ún. molekuláris deszkriptorokat használnak, amelyeket szokás szerint topológiai indexeknek neveznek. A topológiai indexek - ezeknek gyakran szemléletes jelentés is tulajdonítható -, a gráf topológiai szerkezetét jellemző számszerű paraméterek. Jelen dolgozatban összehasonlító vizsgálatokat végeztünk bireguláris gráfok egy véges halmazán azzal a céllal, hogy elemezzük és értékeljük a különféle topológiai deszkriptorok diszkriminációs képességét.

Kulcsszavak:

topológiai deszkriptorok, bireguláris gráfok, Zágráb indexek

1. Bevezetés

A szerkezeti kémiában (structural chemistry) általános az törekvés, hogy a szerves vegyületek (molekulák) fizikai és kémiai tulajdonságai valamint e vegyületek kémiai képletének topológiai szerkezete között (pl. szénhidrogén vegyületek szerkezeti képlete és forráspontja között) „viszonylag erős” korrelációval jellemezhető kvantitatív összefüggést, egyfajta függvény kapcsolatot találjanak. E célra ún. topológiai deszkriptorokat (topological descriptor) szokás definiálni illetve használni, ezeknek számtalan változata ismert a szakirodalomból [1–8]. A legrégebbi gráfelméleti deszkriptor az 1947-ben bevezetett Wiener-index, amely nevét felfedezőjéről kapta [2]. Jelen dolgozatban néhány ismert topológiai deszkriptor diszkriminációs képességének (discriminating performance) elemzésével foglalkozunk, és arra kívánjuk felhívni a figyelmet, hogy ezek diszkriminációs képessége bizonyos esetekben

meglehetősen korlátozott. Különösképp igaz ez a megállapítás az elterjedten használatos ún. Zágráb-indexekre, amelynek két legismertebb változatára az első M_1 és a második M_2 Zágráb indexként szokás hivatkozni a szakirodalomban. Tekintettel arra, hogy a topológiai deskriptorok értelmezése alapvetően gráfelméleti megfontolásokon alapul, a következőkben – a tárgyalás megkönnyítése végett – bevezetünk néhány, a témakörbe vágó matematikai definíciót.

2. Gráfelméleti definíciók

A szerkezeti képletükkel jellemzett karbon-vegyületeket (szerves molekulákat) véges összefüggő gráfokkal reprezentáljuk. Jelölje V és E a G gráf csúcsainak és éleinek halmazát. Feltételezzük, hogy az összefüggő gráfnak n számú csúcsa és m számú éle van, $n \geq 2$. Egy összefüggő gráf D átmérője két leg-távolabbi csúcsának távolsága. Definíció szerint a gráf egy u csúcsának $d(u)$ valenciája (fokszáma) az u csúcsban található élek száma, a gráf átlagos valencia-száma $[d] = 2m/n$. Egy gráf csúcsainak, éleinek számát, átmérőjét, átlagos valencia-számát a gráf globális topológiai invariánsainak szokás tekinteni.

A továbbiakban r -csúcsnak nevezzük a gráf azon csúcsát, amelynek valenciája r , valamint (r,s) élnek pedig a gráf azon élét, amelynek végpontjaiban a csúcsok fokszáma r és s , ahol $r \geq s \geq 1$. A gráfban az r -csúcsok számát n_r -el, az (r,s) típusú élek számát pedig $g(r,s)$ -el jelöljük.

Egy gráfot regulárisnak (R -regulárisnak) nevezzük, ha valamennyi csúcsának fokszáma azonos, nevezetesen $R \geq 1$. Egy G gráfot páros gráfnak nevezünk, ha csúcsainak V halmaza felosztható V_a és V_b diszjunkt részhalmazokra úgy, hogy a gráf minden egyes élének egyik végpontja V_a -ban, a másik pedig V_b -ben van. Teljes páros gráfnak (complete bipartite graph) nevezünk valamely $G(V_a, V_b)$ páros gráfot, ha bármely $u_a \in V_a$ és $u_b \in V_b$ csúcspárra létezik (u_a, u_b) él.

A gráfot biregulárisnak nevezünk, ha csúcsainak valenciája pontosan két különböző értékű, nevezetesen Δ és δ , ahol $\Delta > \delta$ és $\delta \geq 1$. Egy összefüggő gráfot erősen páros bireguláris gráfnak (strongly bipartite biregular graph) nevezünk, ha minden éle (Δ, δ) típusú, azaz minden egyes él egyik végpontjának valenciája Δ , a másik végpontjának valenciája pedig δ . E definícióból következik, egyrészt az, hogy az erősen páros bireguláris gráfokra $m = g(\Delta, \delta)$ teljesül, másrészt az, hogy a bireguláris teljes páros gráfok az erősen páros bireguláris gráfok részhalmazát képezik. Könnyű belátni, hogy van olyan összefüggő, erősen páros bireguláris gráf, amely nem tartozik a teljes páros gráfok halmazába. Példa erre az **1. ábrán** látható síkbeli bireguláris gráf, amelynek 8 darab $\delta=3$ fokszámú valamint 6 darab $\Delta=4$ fokszámú csúcsa van. E gráfnak 24 éle van és valamennyi él $(4,3)$ típusú. (Itt jegyezzük meg, hogy az **1. ábrán** látható sík-gráf a rombikus dodekaéder (rhombic dodecahedron) elnevezésű poliéder élgráfja, e poliédernek 12 rombusz lapja van.)

3. Topológiai deskriptorok

A szakirodalomban a szerves molekulákat reprezentáló gráfok struktúrájának számszerű jellemzésre

sokféle topológiai invariánst (deszkriptort) javasoltak az elmúlt két évtizedben. Ezek között a legismertebbek a M_1 és M_2 ún. első és második Zágráb indexek, amelyek egy véges összefüggő G gráfra nézve az alábbi formulákkal definiáltak:

$$M_1 = M_1(G) = \sum_{u \in V(G)} d^2(u) = \sum_r r^2 n_r \quad (1)$$

és

$$M_2 = M_2(G) = \sum_{(u,v) \in E(G)} d(u)d(v) = \sum_r \sum_{s \leq r} g(r,s)rs \quad (2)$$

1. ábra A rombikus dodekaéder élgráfja

Kiegészítésképpen definiáljuk az M_3 topológiai deszkriptort

$$M_3 = M_3(G) = \sum_{u \in V(G)} d^3(u) = \sum_r r^3 n_r \quad (3)$$

amelyet az egyszerűség kedvéért a harmadik Zágráb indexnek nevezünk. Könnyen igazolható, hogy M_2 és M_3 között fenn áll

$$M_3 = \sum_r r^3 n_r = \sum_r \sum_{s \leq r} g(r,s)(r^2 + s^2) \geq 2 \sum_r \sum_{s \leq r} g(r,s)rs = 2M_2 \quad (4)$$

egyenlőtlenség.

A következőkben egy G gráfra vonatkozóan definiáljuk a topológiai deszkriptorok egy általános osztályát (halmazát), ennek elemeit „valencia alapú éldeszskriptoroknak” nevezzük. A $\Phi_D(G)$ valencia alapú éldeszskriptort (továbbiakban VE-deszkriptort) az alábbi képlettel értelmezzük:

$$\Phi_D = \Phi_D(G) = \sum_r \sum_{s \leq r} g(r,s) W_{r,s}(r,s,n,m) \quad (5)$$

Az (5) formulában $W_{r,s}(r,s,n,m)$ az (r,s) típusú élekhez rendelt nem-negatív súlyfüggvények. Ezek 4-változós függvények, és alapvető jellegzetességük, hogy az r és s változókra nézve szimmetrikus függvények, következésképpen a $W_{r,s}(r,s,n,m) = W_{r,s}(s,r,n,m)$ összefüggés teljesül. Megjegyezzük, hogy a $W_{r,s}$ súlyfüggvények általánosabb formában is definiálhatók, azaz az n csúcsszámon és az m élszámon kívül több más globális topológiai paramétert (pl. D gráfátmérő, a gráf átlagos valencia-száma, Wiener indexe, stb.) is tartalmazhatnak. Könnyen belátható, hogy M_1 , M_2 és M_3 Zágráb indexek mindegyike VE-deszkriptor. A szakirodalomból számtalan különféleképpen értelmezett VE-deszkriptor ismert [4–8]. Példaként említjük az $S(G)$ ún. Gordon-Scantlebury indexet [1], amely az alábbi alakban írható fel:

$$S(G) = \frac{1}{2} \sum_{u \in V(G)} [d(u)(d(u)-1)] = \frac{1}{2} \sum_{u \in V(G)} d^2(u) - m = \sum_r \sum_{s \leq r} g(r,s) \frac{(r+s-2)}{2} \quad (6)$$

4. Bireguláris gráfok strukturális jellemzése topológiai deszkriptorokkal

Gyakorta felmerülő kérdés, hogy az egyes topológiai deszkriptorok milyen hatékonysággal alkalmazhatók a „hasznos jellegű”, de nem izomorf gráfok közötti strukturális különbségek kimutatására, számszerű minősítésére. Ilyen jellegű vizsgálatok céljára előnyösen jönnek számításba a bireguláris gráfok, közülük is főként azok, amelyek azonos csúcsszámmal (n), élszámmal (m), valamint azonos maximális (Δ) és minimális (δ) valencia-számmal rendelkeznek. Ezen megfontolásból kiindulva, a topológiai deszkriptorok diszkriminációs képességének összehasonlító elemzésére kis csúcsszámú bireguláris gráfokat választottunk. Itt jegyezzük meg, hogy bireguláris gráfok a kémiai gráfelmélet egy fontos területét reprezentálják, ezen belül is központi jelentőségű a benzoid szénhidrogének tulajdonságainak predikciója [9]. Mint ismeretes, a bireguláris gráf négy topológiai alapparamétere (Δ , δ , n , m) az M_1 Zágráb index értékét egyértelműen meghatározza [10].

$$M_1 = 2m(\Delta + \delta) - n\Delta\delta \quad (7)$$

A fenti formulából közvetlenül adódik, hogy az M_1 Zágráb index diszkriminációs képessége erősen behatárolt, alkalmatlan az azonos (Δ , δ , n , m) paraméterekkel rendelkező bireguláris gráfok megkülönböztetésére. A **2. ábrán** látható bireguláris gráfok közös jellemzője, hogy azonos ($\Delta=4$, $\delta=2$, $n=6$, $m=8$) paraméterekkel rendelkeznek, és mind a négy gráfra vonatkozóan azonos az első Zágráb index

értéke, $M_1=48$. Amint az **1. táblázat** adataiból kitűnik, a H_1 és H_2 gráfoknak $D=3$, a H_3 és H_4 gráfoknak $D=2$ az átmérője. Megállapítható továbbá, hogy M_1 -hez képest az M_2 deszkriptor illetve a D gráf-átmérő hatékonyabb a gráfok közötti strukturális különbségek detektálására, továbbá M_2 értéke – a négy gráf közül – a H_4 teljes páros gráfra a legkisebb, $M_2(H_4)=64$. A vizsgálatokat a bireguláris gráfok egy másik, bővebb csoportjára kiterjesztve, arra következtettünk, hogy az M_2 deszkriptor „szelektivitása” sem tökéletes, a diszkrimináció hatékonysága olykor ugyancsak korlátozott. A **3. ábrán** látható gráfok nem izomorfak, jóllehet globális topológiai paramétereik közül hat ($\Delta=4$, $\delta=3$, $n=7$, $m=12$, $D=2$, $M_1=84$) azonos.

2. ábra Négy nem izomorf bireguláris gráf ($n=6$ és $m=8$).

A **2. táblázat** adataiból kitűnik, hogy az M_2 deszkriptor alapján a G_a , G_b és G_c gráfok egymástól nem különböztethetők meg, hasonló megállapítás érvényes a G_d és G_e gráf-párra is. Ez esetben is a G_f teljes páros gráfra nézve (amely nem síkgráf) a legkisebb a második Zágráb index értéke ($M_2(G_f)=144$).

1. táblázat Négy bireguláris gráf topológiai paramétere (Δ=4, δ=2, n=6, m=8, M₁=48)

Gráf	Topológiai paraméterek				
	g(2,2)	g(4,2)	g(4,4)	D	M ₂
H ₁	3	2	3	3	76
H ₂	2	4	2	3	72
H ₃	1	6	1	2	68
H ₄	0	8	0	2	64

Ez esetben is a G_f teljes páros gráfra nézve (amely kivételesen nem síkgráf) a legkisebb a második Zágráb index értéke (M₂(G_f)=144)

3. ábra Hat nem izomorf bireguláris gráf (n=7 és m=12).

Könnyen konstruálható bireguláris síkgráfoknak (pontosabban fogalmazva gráf-pároknak) olyan végtelen sorozata, melyben a gráf-párok az M₁ és M₂ Zágráb indexek alapján nem különböztethetők meg egymástól annak ellenére, hogy e gráf-párok struktúrája meglehetősen különböző. A 4. ábrán feltüntetett G_A(k) és G_B(k) gráf-párok több topológiai paramétere – fix k≥3 esetében, ahol k a központi sokszög oldalszáma – megegyezik.

2. táblázat Hat bireguláris gráf topológiai paramétereit ($\Delta=4$, $\delta=3$, $n=7$, $m=12$, $M_1=84$)

Gráf	Topológiai paraméterek				
	$g(3,3)$	$g(4,3)$	$g(4,4)$	D	M_2
G_a	3	6	3	2	147
G_b	3	6	3	2	147
G_c	3	6	3	2	147
G_d	2	8	2	2	146
G_e	2	8	2	2	146
G_f	0	12	0	2	144

4. ábra $G_A(k)$ és $G_B(k)$ bireguláris gráf-párok végtelen sorozata ($k \geq 3$)

Nevezetesen a $G_A(k)$ és $G_B(k)$ gráf-párookra $\Delta=4$, $\delta=2$, $n=3k$, $g(2,2)=k$, $g(4,2)=2k$, $g(4,4)=k$, $m=4k$. Következésképpen tetszőleges $k \geq 3$ esetében a Zágráb indexek is azonosak lesznek: $M_1(G_A(k))=M_1(G_B(k))=24k$, $M_2(G_A(k))=M_2(G_B(k))=36k$ és $M_3(G_A(k))=M_3(G_B(k))=80k$. Ebből adódik az a felismerés, hogy $G_A(k)$ és $G_B(k)$ gráf-párok a hagyományos valencia alapú $\Phi_D(G)$ éldeszcriptorokkal - például a Gordon-Scantlebury indexszel - nem különböztethetők meg egymástól.

5. Topológiai deszkriptorok diszkriminációs képességének elvi korlátai

Bireguláris gráfokra vonatkozóan a hagyományosan topológiai deszkriptorok jelentős hányadára érvényes az a megállapítás, hogy ezek diszkriminációs képességének elvi korlátai vannak. Ennek igazolására definiáljuk a G_B bireguláris gráfra vonatkozóan a

$$\Psi(G_B) = g(\Delta, \Delta)W_1(\Delta, \delta, n, m) + g(\Delta, \delta)W_2(\Delta, \delta, n, m) + g(\delta, \delta)W_3(\Delta, \delta, n, m) \quad (8)$$

topológiai deszkriptort, ahol $W_1(\Delta, \delta, n, m)$, $W_2(\Delta, \delta, n, m)$ és $W_3(\Delta, \delta, n, m)$ nem-negatív súlyfüggvények. Fontos megemlíteni, hogy az (5) képletben szereplő $W_{r,s}$ súlyfüggvények definíciója nem azonos a $W_1(\Delta, \delta, n, m)$, $W_2(\Delta, \delta, n, m)$ és $W_3(\Delta, \delta, n, m)$ súlyfüggvényekével. Az utóbbiak definíciója némileg általánosabb, ugyanis W_1 , W_2 és W_3 nem szükségképpen szimmetrikus függvényei a Δ és δ változóknak.

Legyenek G_1 and G_2 nem-izomorf összefüggő bireguláris gráfok, amelyekre nézve az $(\Delta, \delta, n, m, M_2)$ topológiai mennyiségek azonosak. (Nevezetesen tételezzük fel, hogy a G_1 és G_2 gráfokra $\Delta(G_1)=\Delta(G_2)$, $\delta(G_1)=\delta(G_2)$, $n(G_1)=n(G_2)$, $m(G_1)=m(G_2)$ és $M_2(G_1)=M_2(G_2)$ összefüggések teljesülnek.)

Állítás: Ez esetben a G_1 és G_2 bireguláris gráfokra fenn áll a $\Psi(G_1)=\Psi(G_2)$ azonosság, függetlenül a $W_1(\Delta, \delta, n, m)$, $W_2(\Delta, \delta, n, m)$ és $W_3(\Delta, \delta, n, m)$ súlyfüggvények megválasztásától.

Bizonyítás: Elég megmutatni, hogy tetszőleges G_B bireguláris gráfra értelmezett $\Psi(G_B)$ deszkriptor felírható

$$\Psi(G_B) = X * W_1(\Delta, \delta, n, m) + Y * W_2(\Delta, \delta, n, m) + Z * W_3(\Delta, \delta, n, m) \quad (9)$$

alakban, ahol X , Y és Z egyértelműen definiált 5-változós függvényei a $(\Delta, \delta, n, m, M_2)$ topológiai mennyiségeknek. Kiindulásul tekintsük a bireguláris gráfokra érvényes alábbi azonosságokat:

$$m = m(G_B) = g(\Delta, \Delta) + g(\Delta, \delta) + g(\delta, \delta) \quad (10)$$

$$M_2 = M_2(G_B) = \Delta^2 g(\Delta, \Delta) + \Delta \delta g(\Delta, \delta) + \delta^2 g(\delta, \delta) \quad (11)$$

$$n = n(G_B) = \frac{2}{\Delta} g(\Delta, \Delta) + g(\Delta, \delta) \left(\frac{1}{\Delta} + \frac{1}{\delta} \right) + \frac{2}{\delta} g(\delta, \delta) \quad (12)$$

A (10) és (11) egyenletekből

$$m\Delta^2 - M_2 - \Delta(\Delta - \delta)g(\Delta, \delta) = (\Delta^2 - \delta^2)g(\delta, \delta) \quad (13)$$

valamint a (10) és (12) formulákból kapjuk

$$\delta(\Delta + \delta)(\Delta n - 2m) = (\Delta^2 - \delta^2)g(\Delta, \delta) + 2(\Delta^2 - \delta^2)g(\delta, \delta) \quad (14)$$

Másrészt a (13) és (14) egyenletekből következik

$$2m\Delta^2 + \delta(\Delta + \delta)(2m - \Delta n) - 2M_2 = (\Delta - \delta)^2 g(\Delta, \delta). \quad (15)$$

Könnyen belátható, hogy a $g(\Delta, \Delta)$, $g(\Delta, \delta)$ és $g(\delta, \delta)$ mennyiségek kifejezhetők $(\Delta, \delta, n, m, M_2)$ topológiai paraméterek öt-változós függvényeként az alábbi megfontolások alapján:

A (15) formulából adódóan

$$g(\Delta, \delta) = \frac{2m\Delta^2 + \delta(\Delta + \delta)(2m - \Delta n) - 2M_2}{(\Delta - \delta)^2} = Y(\Delta, \delta, n, m, M_2) \quad (16)$$

Továbbá a (14) egyenlet szerint

$$\delta(\Delta n - 2m) = (\Delta - \delta)g(\Delta, \delta) + 2(\Delta - \delta)g(\delta, \delta) \quad (17)$$

Mivel $g(\Delta, \delta) = Y(\Delta, \delta, n, m, M_2)$ ezért

$$g(\delta, \delta) = \frac{\delta(\Delta n - 2m)}{2(\Delta - \delta)} - \frac{g(\Delta, \delta)}{2} = Z(\Delta, \delta, n, m, M_2) \quad (18)$$

A fentiekből következően a bizonyítás utolsó lépéseként az alábbi

$$g(\Delta, \Delta) = m - Y(\Delta, \delta, n, m, M_2) - Z(\Delta, \delta, n, m, M_2) = X(\Delta, \delta, n, m, M_2) . \quad (19)$$

végeredményhez jutunk.

Fontos következményként adódik: az azonos $(\Delta, \delta, n, m, M_2)$ topológiai paraméterekkel jellemzett összefüggő bireguláris gráfok a hagyományos VE topológiai deskriptorokkal egymástól meg nem különböztethetők. Egy további következmény, hogy (16) egyenletből az

$$M_2 = \frac{1}{2} \{ 2m\Delta^2 + \delta(\Delta + \delta)(2m - \Delta n) - (\Delta - \delta)^2 g(\Delta, \delta) \} \quad (20)$$

összefüggés származtatható. Mint megállapítható, azonos (Δ, δ, n, m) paraméterekkel jellemzett bireguláris gráfokra a második Zágráb index értéke kizárólag a $g(\Delta, \delta)$ mennyiség függvénye. Mivel összefüggő bireguláris gráfokra $g(\Delta, \delta) \leq m$, ezért a (20) formulából

$$M_2 \geq \frac{1}{2} \{ 2m\Delta^2 + \delta(\Delta + \delta)(2m - \Delta n) - m(\Delta - \delta)^2 \} \quad (21)$$

becslés adódik eredményül. Egyenlőség itt csak abban a speciális esetben áll fenn, ha a bireguláris gráf erősen páros, ugyanis $g(\Delta, \delta) = m$ csak ez utóbbi esetben teljesül. Erősen páros bireguláris gráfra $M_2 = m\Delta\delta$. (Szemléletes példa erre az 1. ábrán látható bireguláris poliéder-gráf, melyre $g(4,3) = m = 24$ és $M_2 = m\Delta\delta = 24 \cdot 4 \cdot 3 = 288$.)

Irodalom

- [1] N. Trinajstić: *Chemical Graph Theory*, 2nd revised ed. CRC Press, Boca Raton, USA, 1992.
- [2] B. Lučić, S. Nikolić and N. Trinajstić: *Távolságfüggő molekuláris deszkriptorok*, Magyar Kémiai Folyóirat, Vol. 114 (2008) p. 171–176.
- [3] I. Gutman and K. Ch. Das: *The first Zagreb indices 30 years after*, MATCH Commun. Math. Comput. Chem. Vol. 50 (2004) p. 83–92.
- [4] B. Zhou and N. Trinajstić: *On a novel connectivity index*, J Math Chem Vol 46, (2009) p. 1252–1270.
- [5] J. Zheng: *The general connectivity indices of catacondensed hexagonal systems*, J Math Chem. Vol. 47 (2010) p. 1112–1120.
- [6] B. Zhou and N. Trinajstić: *On a novel connectivity index*, J Math Chem Vol 46, (2009) p. 1252–1270.
- [7] B. Fortula, A. Graovac and D. Vukičević: *Atom-bond connectivity index of trees*, Discrete Applied Mathematics, Vol. 157 (2009) p. 2828–2835.
- [8] D. Vukičević and B. Fortula: *Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges*, J Math Chem (2009) Vol 46 (2009) p. 1369–1376.
- [9] R. Ponc, S. Fias, P. Bultinck, I. Gutman and S. Stankovic: *Benzoid Szénhidrogének aromás és lokális aromás tulajdonságairól*, Magyar Kémiai Folyóirat, Vol. 114 (2008) p. 177–182.
- [10] K.C. Das: *Maximizing the sum of the squares of the degrees*, Discrete Mathematics, Vol. 285, (2004) p. 57–66.

dr. Bitay Enikő, egyetemi docens

Sapientia – Erdélyi Magyar Tudományegyetem
Műszaki és Humántudományok Kar, Marosvásárhely
E-mail: ebitay@gmail.com

dr. Réti Tamás, egyetemi tanár

Óbudai Egyetem, Budapest
H-1081 Budapest, Népszínház u. 8, Hungary
E-mail: reti.tamas@bgk.uni-obuda.hu