Az emberi agygráf felépítésének jellemzése

Hallgató: Hegedűs Dániel (hegedani1226@gmail.com)

Témavezető: Dr. Grolmusz Vince (grolmusz@pitgroup.org)

Tartalomjegyzék

1. Bevezetés **2**

2. Az agygráfok kiszámolása **2**

3. Az átlagolt agygráf **3**

4. Agyterületek fontossága **3**
4.1. Célkitűzés és eszközök 4
4.2. A különböző fontossági szempontok korrelálása 5
4.2.1. Egy egyszerű kontroll 6

5. Permutációk összehasonlítása inverziószámokkal **6**
5.1. Inverziószámokkal kapcsolatos alaptulajdonságok 6
5.2. A csúcsok fontossági szempontjainak összehasonlítása inverziószámokkal . . 7

6. Az élek fontosságainak korrelálása **8**
6.1. Az élek vastagság és hossz szerinti korrelálása 8
6.2. Az élek vastagság és gyakoriság szerinti korrelálása 9

7. Agyterületek elhelyezkedése **10**
7.1. Csúcsok súlypontjai és súlyozott súlypontjai 10
7.2. Csúcsok súlypontjainak elhelyezkedése 12
7.2.1. A csúcsok súlypontjainak egyenestől vett távolságának korrelálása a fontossággal 13

8. Maximális élekből álló gráfok **14**
8.1. Maximális élekből álló agygráf 14
8.2. A komponensek legfontosabb csúcsainak gráfelméleti helyzete 15
8.3. Bővített maximális élekből álló gráf 15
8.4. Véletlen gráfok 17

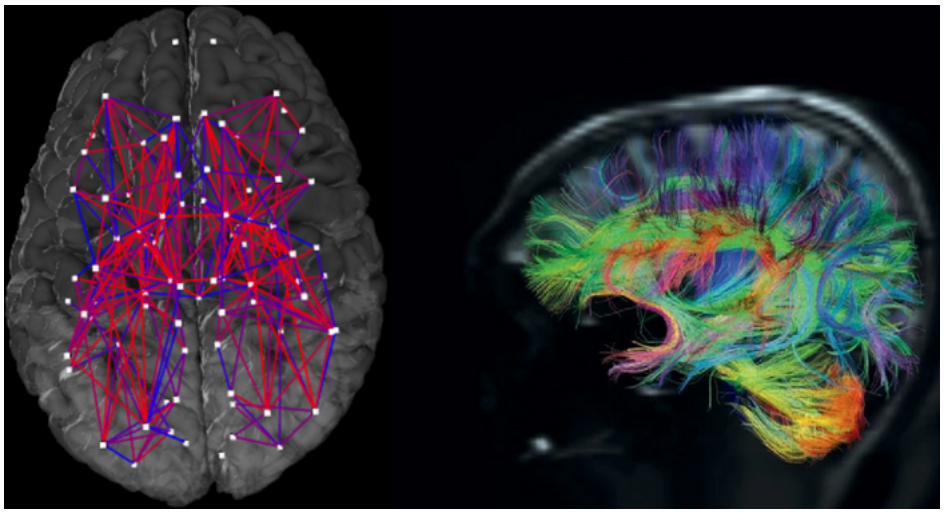
9. Összegzés **18**

1. Bevezetés

Az emberi agy az általunk ismert legkomplexebb szerv, aminek működése és felépítése évezredek óta foglalkoztatja az emberiséget. Ezen munkámban az agy szerkezetét elemzem agygráfok segítségével. Az agygráfok tényleges matematikai gráfok: csúcsai a különböző agyterületek, míg élei a köztük fennálló kapcsolatokat jellemzik. A legfőbb eredményeim az agyterületek különböző fontossági szempontok szerinti rendezéseinek összehasonlítása és az agyi kapcsolatok összefüggőségének jellemzése véletlen gráfok használatával.

2. Az agygráfok kiszámolása

A modern technológia lehetővé teszi precíz MRI-k használatát. Ennek köszönhetően a Human Connectome Project [1] nyilvánosságra hozott 1064 emberről készült diffúziós MRI-t. (Ez a mágnesesrezonancia-képalkotásnak egy olyan fajtája, ahol a vízmolekulákat vizsgálják azt használva, hogy az axonok közelében máshogy mozognak.) Ezután a Lausanne2008 atlas [2] segítségével az agy(felszín) 1015 agyterületre lett osztva, a részek mérete kb. 1cm^2. Az agygráf csúcsai az előbb említett agyterületek, melyeknek a különböző emberekre is egységes címkéi vannak 1-től 1015-ig. A csúcsok közé súlyozott élek lettek behúzva aszerint, hogy a két agyterület között milyen sok idegsejt fut (vagyis, hogy milyen vastag az idegrost). Ezen kívül használni fogjuk még az idegsejtek hosszát. Fontos megjegyezni, hogy az él hossza nem egyezik meg a két csúcs mértani helyének távolságával, hiszen a mértani helyek mindösszesen 3-3 koordinátát jelentenek, viszont a vizsgált agyterületek nem pontszerűek. (Sőt, maguk az idegsejtek sem is hasonlíthatók geometriailag egy szakaszhoz.) Az agyterületek koordinátáit meghatározó koordináta-rendszer tengelyei: x-tengely a balra-jobba, y tengely előre-hátra és z-tengely a fel-le irány. Minél nagyobb az adott koordináta, annál előrébb, fentebb vagy balrább helyezkedik el a csúcs. (Az idegrost frakcionális anizotrópiája is el van mentve a gráfokba, de ezzel a TDK dolgozatom során nem fogok foglalkozni. Ezzel a víz diffúziós profilja jellemezhető, nekünk erre most nincs szükségünk.) Az egyedi gráfok elkészítését nem én végeztem: már a munkám elkezdése előtt Grolmusz Vince és csoportja készítette el őket az MRI-k alapján. [3] Egy agygráf rávetítve az agy egy keresztmetszetére és az idegrostok az agyban: [4] [5] [6]:



1., 2. ábra. Az első ábrán egy emberi agyra vetített agygráf. A fehér pöttyök jelzik a különböző agyterületek pozícióit. Akkor kötünk össze két agyterületet éllel, ha idegrostot

fedeztünk fel köztük. A szín az adott él vastagságával kapcsolatban hordoz információt. A piros-kék színskálán minél pirosabb egy él, annál vastagabb. A második ábrán az idegrostok fizikai helyzete is megfigyelhető, a különböző színek a különböző irányokat jelzik.

3. Az átlagolt agygráf

Az 1064 agygráfot külön-külön elemezve az emberre jellemző általános következtetéseket levonni nehezebb, mintha a gárfokból egy kiátlagolt konszenzusgráfot elemeznénk. Ezért a gráfokat kiátlagoltam az alábbi módon. Minden csúcsnak a helyét egyszerű átlagolással számoltam, így egy átlagos méretű agyban helyeztem el a különböző agyterületeket. (A helyzeteket természetesen most is három koordinátával adjuk meg: az origó az agy hátuljának jobb alsó sarkában van.) Ezután minden él tulajdonságait átlagoltam, ennek első lépése az, hogy összeadtam az összes emberre a kapott értékeket. Az első nem triviális kérdés az, hogy az élek hosszát és vastagságát különböző módon kell-e átlagolni. Az adott él hosszainak összegét nincs értelme 1064-gyel osztani, mert az élek többsége nem szerepel minden emberben. Ebben az esetben viszont elvesztené az információtartalmát, hiszen egy aránytalanul kicsi számot kapnánk hossznak. Így tehát a hossz esetében mindenképpen annyival kell osztanunk, ahány emberben szerepel az adott él. Viszont vastagság esetén az is információ, hogy az emberek mekkora részében 0 a vastagsága az adott élnek. Tehát vastagság esetén mindenképp 1064-gyel kell osztanunk. Így tehát nem kell minden esetben elvégezni egy számolást 1064 gráfon, hanem elég csak az átlagolt gráfon dolgoznunk. (Fontos meggondolni, hogy az átlagolt gráfban minden él szerepel, ami legalább egy emberben előfordult, hiszen az átlagolásnál mindkét esetben csak egy pozitív számmal osztunk.) Tehát ha $\#(i,j)$-vel az (i,j) él előfordulásának számát jelöljük és rendre $s_{i,j,k}$ és $h_{i,j,k}$ a k-adik emberben az (i,j) él súlya és hossza, akkor az átlagolt gráf éleire:

$$s_{i,j} = \frac{\sum_{k=1}^{1064} s_{i,j,k}}{1064} \tag{1}$$

$$h_{i,j} = \frac{\sum_{k=1}^{1064} h_{i,j,k}}{\#(i,j)} \tag{2}$$

4. Agyterületek fontossága

A különböző csúcsok fontosságát többféle szempont szerint is mérhetjük. Vizsgálhatjuk a csúcsok fokszámát, vagy nézhetjük az éleinek vastagságát és hosszát is. Súly és hossz szerint ráadásul háromféle opciónk is van: legnagyobb, átlagos vagy összegzett. Mindegyik értelmes, hiszen pl. az átlagolt súlyok azt mutatják, hogy az idegrostok átlagosan milyen vastagok, míg a súlyainak összege azt mutatja, hogy összesen hány idegsejten kommunikál a többi agyterülettel. (Természetesen ez a két mérce nem ugyanaz, hiszen a sok kis súlyú élnek az összege lehet ugyanannyi, mint kevés vastagabb élnek, az átlaguk mégsem ugyanannyi.) Az élek hosszának vizsgálatakor hasonlóan arról tudunk meg információt, hogy mennyire egyenletes az, hogy az adott csúcs mennyire messzire kommunikál.

4.1. Célkitűzés és eszközök

Fontos kiemelni, hogy a célunk nem az, hogy megállapítsuk a fontos csúcsokat a különböző szempontok szerint. Helyette a cél az, hogy a különböző szempontok szerint rendezzem a csúcsokat és ennek segítségével hasonlítsam össze a különböző szempontokat. Más szóval arról szeretnék nyilatkozni, hogy ugyanazok-e a csúcsok a terjeszkedők, (tehát messzebbre kommunikálók,) mint amik idegrostjai vastagok.

Vegyük tehát a hétféle fontossági szempontot és mindegyik szerint rendezzük csökkenő sorrendbe a csúcsokat. (Pontosabban a hozzájuk tartozó címkéket/indexeket, ahogy minden agyterülethez egy-egy index lett rendelve.) Természetesen azokat a csúcsokat nevezem fontosabbnak, amik a különböző tulajdonságok szerinti indexsorrendekben előrébb szerepelnek. Minden fontossági szempont szerint így kaptunk egy-egy indexsorrendet, amik az 1015-nél nemnagyobb pozitív egészekből, mint elemekből képzett permutációk. (Fontos meggondolni, hogy kihasználjuk azt, hogy semelyik szempont szerint sincs egyenlőség semelyik két csúcs közt sem, emiatt lesz egyértelmű a rendezés.)

Az ilyen indexsorrendek összehasonlítására ideális a Spearman-féle korrelációs ϱ együttható [7]. Az együttható két tulajdonság korrelálásról ad információt, használva a két tulajdonság szerinti sorrendeket. Minden elemnek lesz tehát egy-egy indexe a két tulajdonság szerint: hogy hányadik az adott rendezésben. Az együttható kiszámítására van egyszerű képlet, ha feltehetjük, hogy az indexek egyértelműek, vagyis nincs két elem, amik valamelyik tulajdonság szerint éppen egyenlők. (Ahogy az előbb láttuk, az agygráf esetében ez teljesül.) Ha n elem kétféle tulajdonságát számoljuk ki és d_i-vel jelöljük az i-edik elem két indexének különbségét, akkor

$$\varrho = 1 - \frac{6 \cdot \sum_{i=1}^{n} d_i^2}{n^3 - n} \tag{3}$$

Az együtthatóra teljesül, hogy $-1 \leq \varrho \leq 1$, ahol az 1 jelenti a tökéletes korrelációt és a -1 jelenti a tökéletes ellentétességet. (Utóbbi könnyen bizonyítható annak segítségével, hogy az első n négyzetszám összegére ismert n függvényében köbös zárt képlet: $\frac{n(n+1)(2n+1)}{6}$. A teljes ellentétettségnél az első $\frac{n}{2}$ páratlan négyzetszám összegét kell kiszámolnunk, ami könnyen megkapható a nem használt páros négyzetszámok összegéből, hiszen ez 4-szerese, az első ennyi négyzetszám összegének.) Minél közelebb van az együttható a 0-hoz, annál kevésbé beszélhetünk bármiféle korrelálásról. Viszont ahogy n nő, úgy a kisebb abszolút értékű ϱ-k is már szignifikánsak lehetnek. Így tehát a p-értéket nem csupán a ϱ határozza meg, hanem az n és a ϱ együtt. A kapott p-érték azt jelzi, hogy mennyi az esély a nullhipotézis melletti ilyen mértékű korrelálásra, tehát azt feltételezve, hogy a két változó random viselkedik egymáshoz képest.

A Spearman-együttható precíz kiszámítását és a hozzátartozó p-érték számolását SciPy csomag `spearmanr` függvényével végeztem [8]. (Az egész munkám eredményeit Python nyelven, Pycharm környezetben írt programjaimból kaptam.) A táblázatokban előre megadjuk az egyik sorozatot: ez van a felső sorban, míg a másik sorozat a korreláció vizsgálatban rendre a második sorban található. A harmadik sorban a hozzájuk tartozó Spearman-együttható, az alsó sorban a p-érték. (A 0.0-s p-érték azt jelzi, hogy a Python már nem tudja kezelni az annyira kis értéket. Természetesen a többi értékben is követhet el elhanyagolható kerekítési hibákat. Ez persze nem számít, hiszen 10^{-2}-es p-értékek alatt már szignifikáns korrelálásról beszélhetünk.)

4.2. A különböző fontossági szempontok korrelálása

Korrelációk, ha egyik sorozat a fokszám

	Fokszám					
	Össz.-súly	Max.-súly	Átl.-súly	Össz.-hossz	Max.-hossz	Átl.-hossz
ϱ	0.88	0.84	0.79	0.98	0.51	0.76
p	0.0	10^{-278}	10^{-220}	0.0	10^{-68}	10^{-194}

1. táblázat. A Pearson-féle korrelációs r együtthatóval ellentétben a 0.5-ös ϱ szignifikánsnak számít, mivel aránylag nagy adathalmazról beszélünk, hiszen $n = 1015$. Így leolvashatjuk, hogy a magas fokszámú csúcsok erősen korrelálnak azzal, hogy ha bármilyen szempont szerint veszünk vastag vagy hosszú éllel rendelkező csúcsokat. Ez abban az esetben nem meglepő, amikor a tulajdonság maximumát vagy összegét vettük, de az átlagoltnál azt jelenti, hogy a vastag élekkel rendelkező csúcsok élei nem csak kiugrásszerűen vastagok. A középső sorban a Spearman-együttható, az alsó sorban a p-érték található.

Korrelációk, ha egyik sorozat az súlyösszeg; súlymax

	Súlyösszeg			Súlymaximum		
	Össz.-hossz	Max.-hossz	Átl.-hossz	Össz.-hossz	Max.-hossz	Átl.-hossz
ϱ	0.86	0.42	0.63	0.83	0.41	0.62
p	10^{-302}	10^{-45}	10^{-112}	10^{-262}	10^{-43}	10^{-110}

Korrelációk, ha egyik sorozat az súlyátlag

	Súlyátlag		
	Össz.-hossz	Max.-hossz	Átl.-hossz
ϱ	0.77	0.37	0.54
p	10^{-199}	10^{-33}	10^{-79}

2., 3. táblázat. Megfigyelhetjük, hogy tetszőleges hossz és vastagság szerinti rendezés korrelál. Ez persze szintén nem intuitív, hiszen semmi sem garantálja, hogy ugyanazok a csúcsok lesznek a terjeszkedők, mit a legvastagabb idegrostokkal rendelkezők. A középső sorban a Spearman-együttható, az alsó sorban a p-érték található.

Korrelációk a súlyos tulajdonságokkal; hosszos tulajdondágokkal

	Súlyos tulajdonságok			Hosszos tulajdonságok		
	Össz.-Max.	Össz.-Átl.	Max.-Átl.	Össz.-Max.	Össz.-Átl.	Max.-Átl.
ϱ	0.97	0.98	0.96	0.58	0.86	0.70
p	0.0	0.0	0.0	10^{-91}	10^{-296}	10^{-148}

4. táblázat. Leolvashatjuk, hogy a különböző súly szerinti rendezések és a különböző hossz szerinti rendezések is korrelálnak egymással. Ez persze következik az előző táblázatokból, hiszen az ilyen mértékű korrelálás "tranzitív". A középső sorban a Spearman-együttható, az alsó sorban a p-érték található.

A táblázatokról általánosságban tehát leolvashatjuk, hogy 1015 csúcsra nézve borzasztóan magas együtthatókat és elhanyagolhatóan alacsony p-értékeket kaptunk. Ez pedig természetesen azt jelenti, hogy rendkívül erős korreláció van az agyterületek különböző szempontok szerint vett rendezései közt.

Ekkor tehát megkaptuk, hogy a hétféle szempontja a fontosságnak korrelál. Más szóval ugyanazok az agyterületek fontosak terjeszkedés szempontjából, mint amik vastag idegrostokkal, tehát erős/gyors kapcsolatokkal rendelkeznek. Ebből természetesen nem következik, hogy az élekre is elmondható hasonló: nem feltétlenül igaz az, hogy a hosszabb élek vastagabbak is. Az élek vizsgálatával később foglalkozunk.

4.2.1. Egy egyszerű kontroll

A teljesség és biztosság kedvéért érdemes megemlíteni, hogy a randomnak várt tulajdonságokról az előző vizsgálattal tényleg azt kapjuk, hogy random viselkednek. Ilyen pl. a csúcsok alapvető címkéje és a csúcsok éleinek átlagsúlya 0.01-es ϱ-val és 0.65-ös p-értékkel. Ez a nagyságú p-érték is csak annak köszönhető, hogy az indexelés sem teljesen random. A két agyfélteke csúcsai egymás utána vannak megsorszámozva: az első kb. 500 csúcs az egyik agyféltekében helyezkedik el, a második kb. 500 csúcs a másik agyféltekében (ugyanolyan sorrendben). Ekkor viszont az, hogy a legfontosabb néhány csúcs az egyikben késői sorszámot kapják, a sorszámozás miatt magával vonja azt is, hogy ugyanezen fontos csúcsok másik agyféltekében található megfelelői is nagy sorszámokat kapnak. Természetesen így sem beszélünk semmiféle korrelációról, ezzel csak megbizonyosodtunk az algoritmusunk helyességéről.

5. Permutációk összehasonlítása inverziószámokkal

Megkaptuk tehát, hogy a korrelálás borzasztóan erős a fontossági szempontok közt, de semmi kézzel fogható, vagy olyan eredményt nem kaptunk, ami miatt jobban értenénk, hogy miért is van így. Érdemes tehát csúcsonként megnézni, hogy mennyire különbözik a permutációkban elfoglalt helyük. Pontosabban a Spearman-féle módszerrel ellentétben itt nem azt fogjuk megnézni, hogy a helyzete a két permutációban mennyire tér el, hanem az inverziószámokat vizsgáljuk meg. Természetesen a két módszer ugyanazt az eredményt adja, de ez inkább beszél arról, hogy a csúcsok egymáshoz képesti helyzete mennyire eltérő a különböző tulajdonságok szerint. Lentebb az inverziószámok főbb tulajdonságait szedjük össze, a teljesség kedvéért be is bizonyítjuk őket. Az alábbiakban amikor permutációkról beszélünk, akkor mindig azon n hosszúságúakról lesz szó, amiknek elemei pontosan az egész számok 1-től n-ig. Jelöljük ezen permutációk halmazát S_n-nel.

5.1. Inverziószámokkal kapcsolatos alaptulajdonságok

5.1. Definíció. Két adott permutáció esetén két elem *inverzióban áll egymással*, ha sorrendjük ellentétes a két permutációban.

5.2. Definíció. Két permutáció *inverziószáma* azon elempárok száma, amik inverzióban állnak egymással. Egy elem *inverziója* azon tőle különböző elemek száma, amikkel inverzióban áll.

5.1. Lemma. Két S_n-beli permutáció közti inverziószám várható értéke $\frac{n(n-1)}{4}$.

Bizonyítás. Vizsgáljunk egy rögzített (i, j) nem rendezett számpárt. Ekkor szimmetriai okokból a köztük levő inverzió várható értéke $\frac{1}{2}$, hiszen minden permutációnak bijektíven létezik egy párja, amik közt csak az (i, j) transzpozíció a különbség. (Transzpozíciónak hívjuk azt a függvényt, ami egy permutációban két adott elemet megcserél.) Ismert, hogy $\binom{n}{2} = \frac{n(n-1)}{2}$ rendezetlen számpár választható az n elem közül. Használva a várható érték linearitását megkapjuk, hogy az inverziószám várható értéke tényleg $\frac{1}{2} \cdot \frac{n(n-1)}{2} = \frac{n(n-1)}{4}$. □

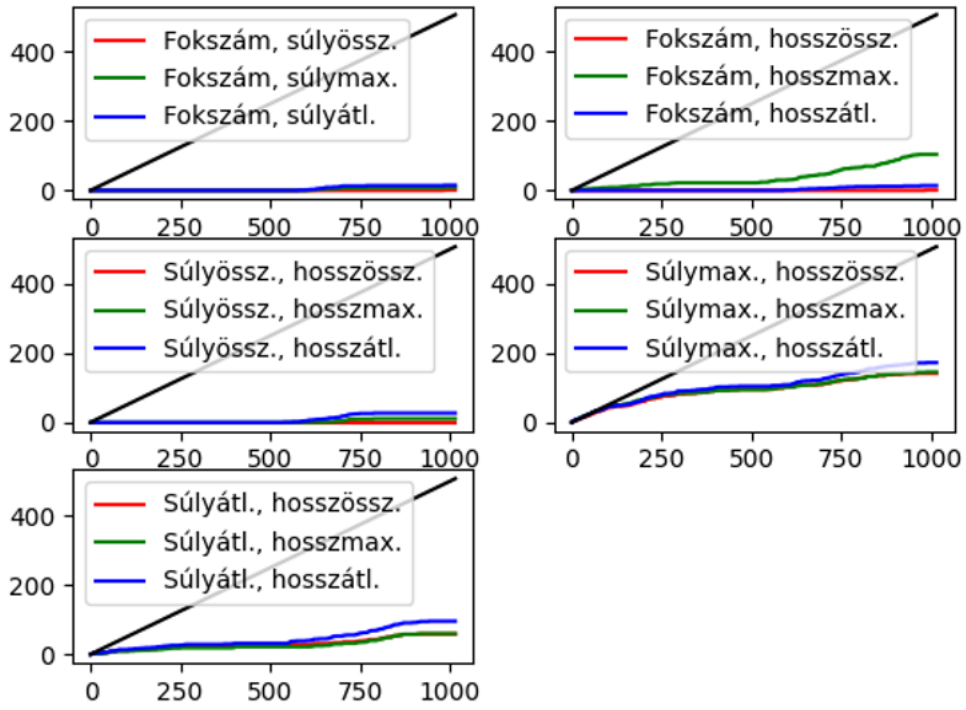
Következmény. Minden elem várható inverziója $\frac{n-1}{2}$.

Bizonyítás. Ismételten használhatjuk az 1. Lemmára a várható érték linearitását. Mivel n elem van és minden inverziót mindkét elemére számolunk, ezért (a kettős leszámlálás, más néven tehén-módszer miatt) elemenként a várható érték $\frac{n(n-1)}{4} \cdot \frac{2}{n} = \frac{n-1}{2}$. □

Alternatív bizonyítás. Állítsuk párba bijektíven a permutációkat a fordított permutációjukkal. Ekkor minden párban bármely (i, j) nem rendezett elempár pontosan az egyik permutációban van felcserélve. Ekkor tehát tetszőleges i elemre $n-1$-féle tőle különböző j létezik és mindegyikkel a várható értéke $\frac{1}{2}$, vagyis az adott elem várható inverziója $\frac{n-1}{2}$. □

5.2. A csúcsok fontossági szempontjainak összehasonlítása inverziószámokkal

Lentebb diagramokat láthatunk, amiken azt ábrázoltam, hogy hány csúcs esetén lesz nagyobb az inverziószám a várható értéknél a két megadott fontossági szempont szerint. (A várható inverzió az előbb kiszámolt $\frac{n-1}{2}$. Most azt vizsgáljuk, hogy hány csúcs inverziója nagyobb ennél.) A fekete egyenes jelzi a várható értékét a nagy inverziójú csúcsok számának, a többi szín pedig a különböző tulajdonságpárokat mutatja. Az x-tengely n pontjában az van, ha az egyik fontossági szempont szerint az első n legfontosabb csúcs halmazát nézzük. Az y-tengelyen az, hogy ezen csúcshalmazban hány olyan van, amik inverziója nagyobb a várható értéknél (az 1015 hosszú permutációkban). Ezzel azt is könnyen leolvashatjuk, hogy nem figyelhető meg lényegi különbség a fontos és a kevésbé fontos csúcsok inverziói közt.



3. ábra. Magas inverzióval rendelkező csúcsok száma a különböző tulajdonságpárok szerint. Az x-tengely n pontjában az van, ha az egyik fontossági szempont szerint az első n legfontosabb csúcs halmazát nézzük. Az y-tengelyen az, hogy ezen csúcshalmazban hány olyan van, amik inverziója nagyobb a várható értéknél (az 1015 hosszú permutációkban). Minden diagramban a fekete egyenes jelzi az n csúcs esetén fenn álló $\frac{n}{2}$-es várható értéket. Az ábrákról leolvasható, hogy a vizsgált 15-féle tulajdonságpár mindegyikére igaz az, hogy a várható értéknél lényegesen kevesebb csúcs esetén van nagy eltérés az indexsorrendekben elfoglalt pozíciói közt.

Ezzel tehát tényleg beláttuk, hogy a különböző fontossági szempontok erősen korrelálnak. A csúcsoknak hasonló csoportja lesz nagy fokszámú, mint amik vastag, vagy

hosszú élekkel rendelkeznek. Ezek közül természetesen a fokszám és a hosszok összege közti kapcsolat nem váratlan, ahogy több éllel várhatóan az hosszuk összege is nagyobb lenne. Az igazán meglepő eredmény az, hogy a súlyok átlaga is korrelál a fokszámmal. Az, hogy ezt vizsgáltuk egyedi gondolkodás, hiszen az orvostudományban legtöbbször ennek az ellenkezőjére helyeznek hangsúlyt. A céljuk az, hogy megállapítsák, hogy melyik csúcs, miért felelős. Jelen dolgozatban nem ez volt az érdeklődésünk középpontjában. Természetesen az elemzések során a csúcsok címkéiből könnyen visszakaphatjuk a hozzájuk tartozó agyterületeket. Ezáltal meg tudjuk nevezni az agy legfontosabb területeit a fenti szempontok szerint, a teljesség kedvéért, ezt meg is teszem. A néhány legfontosabb agyterület a fokszámhoz és élvastagsághoz kapcsolódó sorszámaik összege alapján (a jobb és bal agyféltekében egyaránt): Caudate, Putamen, Thalamus-Proper, Hippocampus, Pallidum, azaz az agykéreg alatti magok.* Megjegyzendő, hogy ezek az agyterületek közel vannak a két agyféltekét elválasztó síkhoz, így azon ritka csúcsok közé is tartoznak, amikből egyaránt fut él mindkét agyféltekébe.

Fontos megjegyezni, hogy a fentiek alapján nem következik funkcionális karakterizáció. Sőt, a puszta agygráfok ismeretében ez lehetetlen, hiszen a gráfban csak az agyterületek helyei és a köztük lévő kapcsolatok vannak elmentve. Ahhoz, hogy a konkrét csúcsok funkcióiról kapjunk eredményt, funkcionális MRI-ket (fMRI-ket) kellene összehasonlítani. Ott az MRI-k készítése közben az alany valamilyen egyszerű feladatot végez (pl. számol vagy olvas). A különböző feladatok elvégzésekor az agy más-más területe fog aktívabban működni, innen megtudva, hogy melyik területeknek van nagyobb szerepe az adott feladat elvégzésében.

Kiemelendő, hogy a dolgozat eddigi részéből a kari TDK idején Grolmusz Vince témavezetőmmel közös cikket írtunk [9].

6. Az élek fontosságainak korrelálása

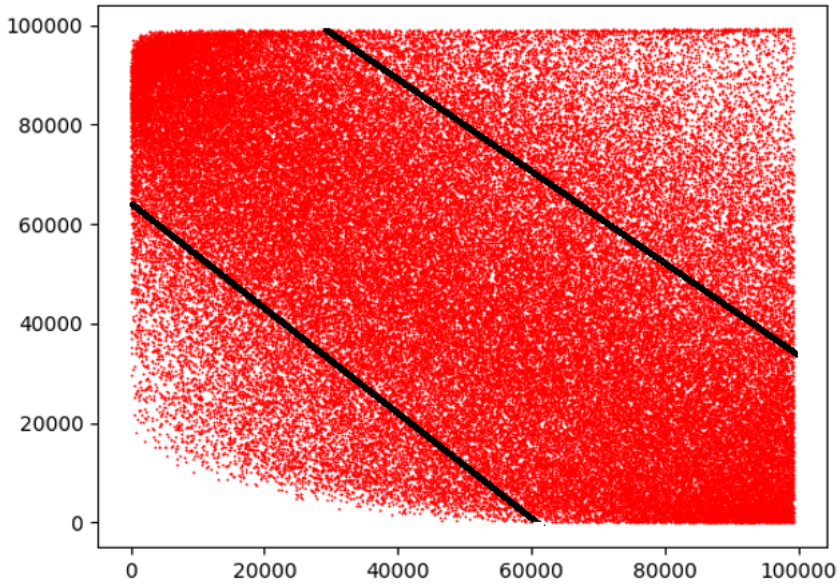
Az élekre természetesen az átlagolt agygráfban csak két tulajdonságot tudunk vizsgálni: súly és hossz. Ez amiatt van, hogy az agygráf értelme az agyi kapcsolatok rendszerének jellemzése, jelen esetben pedig magukat a kapcsolatokat szeretnénk összehasonlítani. Tehát így csak az a két tulajdonság áll rendelkezésre, ami alapvetően is el van mentve a gráfba. Rendezhetjük tehát az éleket a hosszuk és a súlyuk szerint.

6.1. Az élek vastagság és hossz szerinti korrelálása

Vizsgáljuk meg az összes él súlyából és hosszából kapott indexsorrendek közti Spearman-korrelációt. Ezt arra a 99171 élre végezzük el, amik az átlagolt agygráfban szerepelnek (nemnulla súllyal). A két sorozat a vastagság és a hossz szerinti csökkenő sorrend. Azt kaptuk, hogy a Spearman-együttható -0.51 és a hozzá tartozó p-érték 0.0. (Fontos megjegyezni itt is, hogy van olyan nagy a mostani adathalmazunk, hogy ehhez az aránylag kis abszolút értékű együtthatóhoz is egy elhanyagolható nagyságú p-érték tartozik.) Ez persze látszólag ellentmond az eddig leírtaknak, miszerint az élek hosszúsága és vastagsága pozitívan korrelálna, hiszen most azt kaptuk, hogy a rövidebb élek a vastagabbak. Ennek megértése érdekében készítettem egy diagramot, aminek a pontjai az élek. Minden pont

*Az egyértelműség kedvéért a nemzetközileg egységes, angol nyelv által átvett latin kifejezéseket használjuk, melyek Magyarországon is elfogadott szakszavak.

x-koordinátája az, hogy hányadik legvastagabb az adott él, az y-tengelyen pedig az, hogy hányadik leghosszabb az adott él.

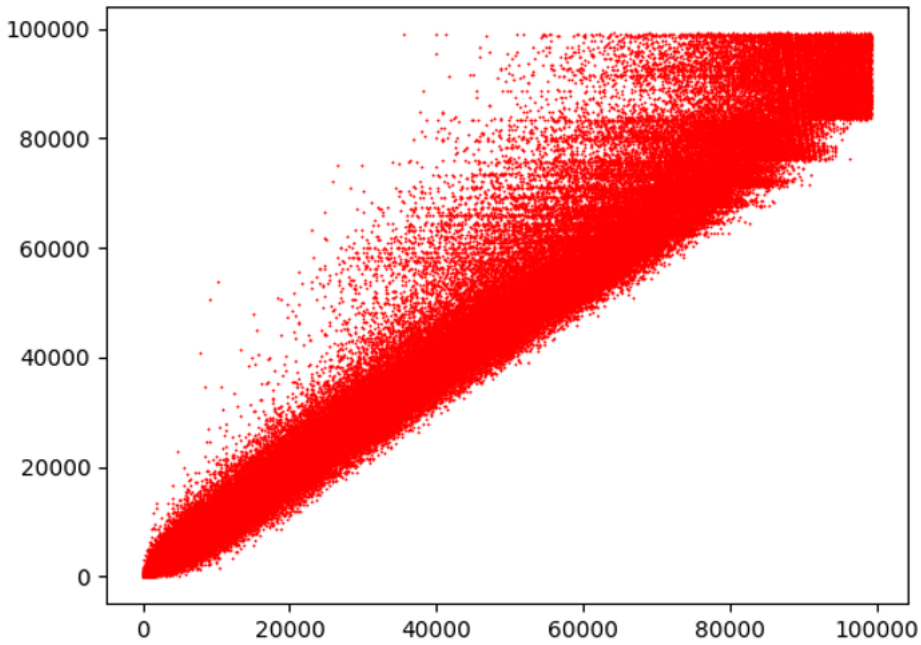


4. ábra. Az élek vastagság és hossz szerinti indexeinek összehasonlítása. A diagramon minden él egy pontnak felel meg, amik x-koordinátája a vastagság szerinti, y-koordinátája a hossz szerinti rendezésbeli sorszáma. Az erős korrelálást az adja, hogy az élek túlnyomó többsége a két fekete vonal között helyezkedik el. A Spearman-együttható -0.51, a hozzátartozó p-érték 0.0.

Láthatjuk, hogy az erős korrelálást az adja, hogy a legvastagabb élek között nincs olyan, ami az egyik leghosszabb is lenne. Másrészt az is ritka, hogy valami egyszerre rövid és vékony. Ebből következhet tehát az ellentétes korrelálás: minél vastagabb egy él, annál jellemzőbb az, hogy rövid is és hasonlóan minél hosszabb egy él, annál jellemzőbb, hogy vékonyabb is. Ez persze csak még erősebben kelt látszólagos ellentmondást: a csúcsok fontosságai közt korrelálás volt. Hogy lehet az, hogy mégis a vastagabb élek a rövidebbek? Úgy, hogy az élek két rendezése közt olyan értelemben gyenge a korrelálás, hogy pl. az első 1000 élt vizsgálva lokálisan semmilyen korrelálást nem kapunk. Tehát a látszólagos ellentmondás abból az anomáliából ered, hogy úgy is lehet két tulajdonság közt korrelálás, hogy azok közt lokálisan lényegében alig van.

6.2. Az élek vastagság és gyakoriság szerinti korrelálása

Az éleknek van egy tulajdonságuk, amit az átlagolt agygráfból nem olvashatunk le: az, hogy hány emberben szerepeltek. Ahhoz, hogy ezt kiszámoljuk, még az átlagolt gráf elkészítésekor azt is elmentjük, hogy az adott él hány emberben fordult elő. Természetesen adja magát az ötlet, hogy az előfordulási szám korreláljon azzal, hogy mekkora a súlya az adott élnek. Ebből következne az is, hogy ellentétesen korrelál az él hosszúsága és az előfordulási száma. Ez pedig fenn is áll: az előfordulás és vastagság közti korrelálás Spearman-együtthatója 0.97, a hozzá tartozó p-érték 0.0, míg a hosszúsággal való összehasonlítás esetén az együttható -0.47, a hozzátartozó p-érték 0.0. Ez tehát borzasztóan erős korrelálást jelent mindkét, de főleg az előbbi esetén (a kiemelkedően magas együttható miatt). Ennek szemléltetésére érdemes most is megnézni az előzőhöz hasonló diagramot:



5. ábra. Az élek vastagság és előfordulás szerinti indexeinek összehasonlítása. A diagramon minden él egy pontnak felel meg, amik x-koordinátája a vastagság szerinti, y-koordinátája a gyakoriság szerinti rendezésbeli sorszáma. A diagram felső részén megfigyelhető vízszintes sávozottságot az adja, hogy a gyakoriság szerinti rendezésben sok él indexe azonos a skatulya-elv miatt, ahogy lényegesen több él van (99171), mint lehetséges előfordulási szám (1064). A Spearman-együttható 0.97, a hozzátartozó p-érték 0.0.

7. Agyterületek elhelyezkedése

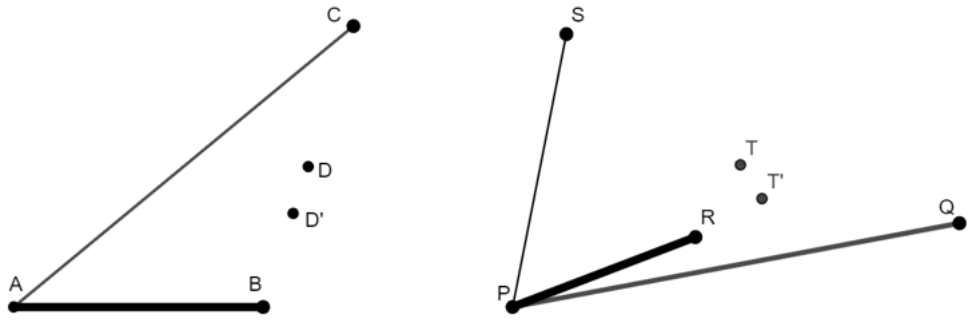
Észrevehetjük, hogy a dolgozat ezen részéig semmit sem foglalkoztunk azzal, hogy az adott agyterületek hordoznak geometriai információt is. Pedig egy olyan szerv esetében, mint az agy, koránt sem szabad megfeledkeznünk arról, hogy az agyterületek hogyan helyezkednek el egymáshoz képest. Azt ugyan vizsgáltuk, hogy az idegsejtek hossza mekkora, de ezt csak csúcspárokra tettük meg, nem pedig az egész agyhoz viszonyítva.

7.1. Csúcsok súlypontjai és súlyozott súlypontjai

Adott i címkéjű csúcshoz rendeljük hozzá a belőle kiinduló élek másik végpontjainak H_i halmazát.

7.1. Definíció. Az i-edik csúcs *egyszerű súlypontja* legyen a H_i-beli csúcsok helyeinek számtani közepe.

7.2. Definíció. Az i-edik csúcs *súlyozott súlypontja* legyen a H_i-beli csúcsok helyeinek súlyozott átlaga, ahol a végpontok helyeinek súlya legyen az oda mutató él vastagsága.



6. ábra. Az ábrán megfigyelhetjük, hogy az A csúcsból kiinduló AB és AC élek végpontjainak egyszerű súlypontja a D pont, de AB súlyát kétszeresre növelve már a $D' \neq D$ a súlyozott súlypont. Hasonlóan a P pontból a PQ, PR, PS élek végpontjainak egyszerű súlypontja T, de PQ súlyát kétszeresre, míg PR súlyát háromszorosra változtatva már a $T' \neq T$ pont a súlyozott súlypont. Az agygráf esetében ennél nagyságrendekkel nagyobb eltérés is gyakori az élsúlyok közt, így mindenképp érdemes mindkét féle súlypontot vizsgálni.

Természetesen minden csúcs egyszerű súlypontja arról ad információt, hogy átlagosan hova kommunikál az adott agyterület. A súlyozott súlypont esetén pedig számításba vesszük a kommunikáció erősségét is. Utóbbi nyilván jobban jellemzi azt, hogy a csúcsból közölt információk átlagosan hol érnek célba vagy honnan érkeznek bele információk.

Felmerülhet a kérdés, hogy a csúcsok milyen messze vannak a súlypontjuktól. Mivel azt kaptuk, hogy a rövidebb élek a vastagabbak, ezért azt várjuk, hogy a súlyozott súlypontok közelebb lesznek a csúcshoz, mint az egyszerű súlypontok. Ez pedig tényleg így is van, mind koordinátánként, mind euklideszi távolságban számolva. A táblázatból leolvashatjuk az átlagos távolságokat mm-ben (egyszerű és módosított súlypont az x, y, z-koordináták szerint és euklideszi távolság):

Súlyponttól vett távolságok

Egysz_x	Egysz_y	Egysz_z	$\text{Egysz}_{\text{Eukl}}$	Mod_x	Mod_y	Mod_z	Mod_{Eukl}
5.38	7.92	6.12	12.80	4.04	4.77	4.08	8.46

5. táblázat. Leolvashatjuk mm-ben mérve a csúcsok átlagos távolságát az egyszerű és súlyozott súlypontoktól koordinátánkénti és euklideszi távolságot használva.

Az cseppet sem meglepő, hogy az y-koordináta szerint a legnagyobb a távolság. Hiszen az agy fizikai adottságait tekintve előre-hátra irányban nyújtott, így az élek y-koordináta szerinti komponense a legnagyobb. Viszont a koordináták közti különbség számottevően kisebb a módosított súlypontok esetén: 2.4 mm-es különbség helyett csak 0.7 mm-es különbség van. Ez arra enged következtetni, hogy a kis súlyú, de azért előforduló élek leginkább az y-tengellyel párhuzamos irányban terjeszkednek.

Érdemes meggondolni, hogy eddig persze a megszokott távolságot használtuk, de ehelyett használhatnánk előjeles távolságot is. Első ránézésre gondolhatnánk azt, hogy a különbségek összege az a nullvektor, de ez nincs így. Hiszen egy adott él két végpontjának különböző lehet a foka, ezzel együtt persze más súllyal fog számolódni a kétféle irányításban. Az előjeles távolsággal nem fogok foglalkozni a dolgozatom során.

7.2. Csúcsok súlypontjainak elhelyezkedése

A vizsgált agyterületek az agykéreg szürkeállományában vannak, míg az idegrostok a fehérállományban futnak. Teljességgel jogos kérdés lehet az, hogy mennyire szimmetrikusan futnak az élek a csúcsokból.

7.3. Definíció. Az összes csúcs koordinátáinak súlyozatlan átlagát az alábbiakban nevezzük az *agy közepének*.

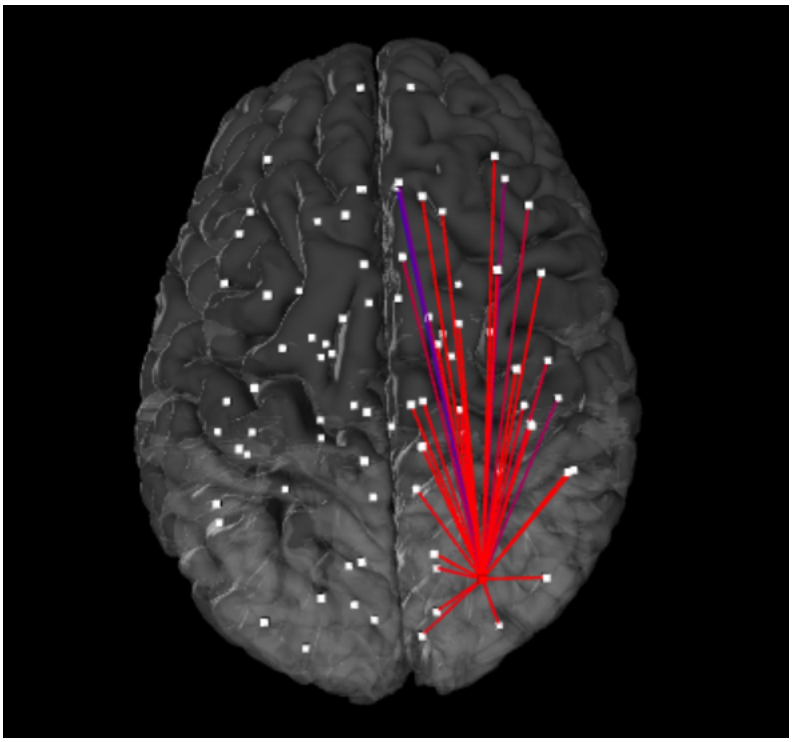
Ez egy megfelelő definíció, mert azon kívül, hogy az agyterületek "egyenletesen" vannak elosztva, pont azt az információt fogja hordozni, hogy a csúcsok átlagosan hol vannak. Az érthetőség kedvéért például, ha egy csúcs minden másik csúccsal össze van kötve, akkor a súlyozatlan súlypontja minimális eltéréssel az agy közepe.

7.4. Definíció. Minden csúcshoz rendeljük hozzá azt az egyenest, amin rajta van a csúcs és az agy középpontja.

Megjegyezzük, hogy ez az egyenes sosem lesz elfajult, ahogy az agy középpontja nem esik egybe egyik csúccsal sem, mivel a csúcsok többsége az agykérgen helyezkedik el.

A kérdés az, hogy a súlypont milyen messze van az előbb definiált egyenestől. Az 1015 csúcsra átlagolva az egyszerű súlypont átlagosan 7.24 mm-re, míg a súlyozott csak 4.98 mm-re van az egyenestől. A különbség szembetűnő, a módosított súlypont bő két milliméterrel közelebb van az egyeneshez, mint az egyszerű.

Az előző egy érdekes következménye annak, hogy a rövidebb élek a vastagabbak. Először is azt kell meggondolni, hogy a két agyfélteke (és az enyhén megnyújtottság) miatt az agy nem forgásszimmetrikus. A két agyféltekét átívelő él viszonylag kevés van, így a csúcsok súlypontjának komponensei nagyrészt a saját agyféltekéjükből származnak. Ekkor viszont a nagyobb eltérést eredményező hosszabb élek kisebb súllyal vannak számolva a súlyozott súlypont esetén, tehát tényleg közelebb lesz a súlyozott súlypont az egyeneshez. (Megjegyzendő, hogy azt használjuk ki, hogy a csúcsok többségének a hosszabb élei messzebb esnek az agy középpontjától.) A következő ábrán egy csúcs élei láthatók egy agygráfon, amin jól megfigyelhetjük a jelenséget.



7. ábra. Egy agyterület és a hozzá tartozó élek. Megfigyelhető, hogy a két agyfélteke közt aránylag kis helyen (a Corpus Callosumban, azaz a kérgestestben) futó élek miatt a csúcs leghosszabb élei nem az agy közepe irányában futnak.

Felmerül a lehetőség, hogy az egyenestől vett távolság és a csúcsok fontosságai közt van-e korrelálás. Mit is jelentene ez? Azt, hogy a fontos csúcsok nem úgy lesznek fontosak, hogy egy-egy adott irányba kommunikálnak erősen, hanem a kommunikációjuk valamilyen értelemben szimmetrikus is.

7.2.1. A csúcsok súlypontjainak egyenestől vett távolságának korrelálása a fontossággal

Nézzük meg a csúcsokra egyesével, hogy milyen távol van a súlypontjuk az egyenestől és ezt hasonlítsuk össze azzal, hogy a csúcsok milyen fontosak. (Természetesen ahogy a különböző fontossági szempontok erősen korrelálnak egymással, így bármelyiket választhatnánk. Jelen esetben használjuk a csúcsok éleinek vastagságainak összege szerinti rendezést.)

Az egyszerű súlypont esetén a Spearman-együttható és a hozzátartozó p-érték -0.20 és $3.66 \cdot 10^{-10}$, míg a módosított súlypont esetén -0.06 és 0.075. Azt látjuk tehát, hogy az egyszerű súlypont esetén a korrelálás erős, míg a módosított súlypont esetén gyengébb (kb. $92,5\%$-os megbízhatóság). Ez azt jelenti, hogy azok a csúcsok a fontosabbak, amik élei súlytól függetlenül szimmetrikusan helyezkednek el. Míg a fontosságot kevésbé határozza meg az, hogy ha az éleket a vastagságaikkal számoljuk. Ez tehát azt jelenti, hogy a fontosabb csúcsokra az igaz, hogy az élei aránylag szimmetrikusan helyezkednek el az agyban. (Fontos megjegyezni, hogy a csúcsok többé-kevésbé az agy kéregállományából vannak véve. Tehát az, hogy az egyenestől vett távolság kicsi, tényleg azt jelenti, hogy az agy középpontja felé helyezkedik el a súlypont és nem az ellenkező irányba.)

8. Maximális élekből álló gráfok

Szeretnénk az agyat nem csak összességében vizsgálni, hanem megnézni a legvastagabb élek szerkezetét is. Így az eredményeinket esetlegesen torzító vékony éleket kiszűrjük és a legfontosabb élekkel fogunk foglalkozni.

8.1. Definíció. Legyen adott egy tetszőleges súlyozott egyszerű gráf páronként különböző súlyokkal. Hívjuk a hozzá tartozó *maximális gráfnak* azt a gráfot, aminek csúcsai az eredeti gráfunk csúcsai és az élei azok pontosan azok, amik valamelyik végpontjának a legnagyobb súlyú éle. Hívjuk *bővített maximális gráfnak* azt, aminek élei pontosan minden csúcs két-két legnagyobb súlyú éle.

Az agy esetében a maximális gráfnál multiplicitással számolva 1015 élt választottunk ki, de persze lehetnek élek, amiket kétszer is kiválasztottunk. Először lássunk be egy általános gráfelméleti lemmát.

8.1. Lemma. Vegyünk egy tetszőleges n csúcsú egyszerű gráfot súlyozott élekkel, ahol minden súly különböző. Állítás: a hozzá tartozó maximális gráf körmentes, tehát nem tartalmaz legalább három különböző élből álló kört.

Bizonyítás. Indirekten tegyük fel, hogy a gráfban van kör, legyen a hossza $k \leq n$. A kör élein haladva számozzuk meg a kör csúcsait rendre az $1, 2, ..., k$ számokkal.

Tudjuk, hogy bármelyik élnek két végpontja van, így bármelyik kiválasztott él valamelyik végpontjának a legnagyobb éle. Mivel egy körben minden élnek mindkét végpontja körbeli, így mindenképp egy körbeli csúcs miatt lett kiválasztva minden él. Tudjuk, hogy minden csúcshoz pontosan egy él lett kiválasztva, ezért bijekció áll a körbeli csúcsok és a kiválasztott élek közt aszerint, hogy melyik csúcshoz, melyik élt választottuk ki.

Mivel az $(1, 2)$ él ki lett választva, az általánosság megsértése nélkül feltehetjük, hogy az 1-es csúcsnál választottuk ki. Viszont ahogy a 2-es csúcsnak két éle van a körben és az egyik élének már megtaláltuk a bijekcióbeli párját, ezért a $(2, 3)$ él mindenképp a 2-es csúcsnál lett kiválasztva. Ez viszont azt jelenti, hogy a $(2, 3)$ él súlya nagyobb, mint az $(1, 2)$ él súlya, hiszen különben a 2-es csúcsnál nem a $(2, 3)$ él lett volna kiválasztva. A körön haladva az előbbi indoklást elmondva kapjuk, hogy minden új él súlya nagyobb, mint az előző. De ekkor az (i, j) él súlyát $s(i, j)$-vel jelölve azt kaptuk, hogy $s(1, 2) < s(2, 3) < ... < s(k - 1, k) < s(k, 1) < s(1, 2)$. Ez pedig ellentmondás, hiszen $s(1, 2) < s(1, 2)$ nem teljesülhet. (Az előző bizonyítás elmondható úgy is, hogy a kör legnagyobb súlyú éléből indulunk ki, amiről az általánosság megsértése nélkül feltesszük, hogy az 1-es csúcshoz kiválasztott $(1, 2)$ él. De ekkor a 2-es csúcsnál való kiválasztás miatt a $(2, 3)$ élnek nagyobb súlyúnak kellene lennie, ami szintén ellentmondás.) □

Megjegyzendő, hogy a bizonyítás során csak azt használtuk ki, hogy semelyik csúcsnak sincs két azonos súlyú éle. Tehát az a feltétel, hogy minden súly különböző, lecserélhető arra az általánosabbra, hogy ez csúcsonként külön-külön igaz.

8.1. Maximális élekből álló agygráf

Mivel az átlagolt agygráfunkban szerencsére minden él súlya különböző, így alkalmazható rá az előbbi lemma. Megtudtuk tehát, hogy az agyban (is) ez a gráf egy erdő. Természetesen minden erdő független fákból áll, az átlagolt agygráf 27-ből. A komponensek nem azonos nagyságúak: három 200-as csúcsszám feletti komponens van. Az érdekesség

az, hogy egyik lefedi a jobb agyfélteke nagyrészét, a másik a bal agyfélteke nagyrészét, a harmadik pedig mindkettőben szimmetrikusan jelen van.

8.2. A komponensek legfontosabb csúcsainak gráfelméleti helyzete

Mind a 27 komponensnek kiválaszthatjuk a legfontosabb csúcsát. A fontossági szempont a csúcsokba futó élek súlyainak összege (egyenlőség itt sem áll fenn).

8.2. Definíció. Egy (V, E) gráf *centruma* azon v csúcsok halmaza, melyekre $\max_{u \in V} d(u, v)$ minimális, ahol a d függvény a súlyozatlan gráfra vonatkozó gráfelméleti távolságot jelöli. (A definíció értelmes, ahogy a távolság egyértelmű bármely két csúcsra, valamint véges sok csúcspár van.)

8.2. Lemma. Fagráfok esetén a centrum legfeljebb két csúcsból áll.

Bizonyítás. Indirekten tegyük fel, hogy a gráfnak három csúcsa is a centrumban van. Ekkor ahogy a gráf körmentes, van köztük kettő, amik nem szomszédosak, legyenek ők u és v. Ekkor az egyértelmű uv útban kell lennie egy $s \neq u, v$ csúcsnak. Ahogy u és v is centrumbeli, így mindkettőnek a legnagyobb távolságát tanúsító csúcshoz vezető útnak át kell mennie a másik csúcson. (Hiszen különben, ha ez v-re nem teljesül, akkor u-t eggyel v felé mozgatva csökkenne a maximális távolsága, tehát u mégsem lenne centrumbeli.) Ez viszont azt jelenti az uv út belsejének nemüressége miatt, hogy az út belsejében levő s-nek a maximális távolsága kisebb lenne, mint u, v-nek. Ez viszont ellentmondás, hiszen ekkor tetszőleges u, v centrumbeli csúcsnak szomszédosnak kellene lennie. □

Biológiailag az látszik előnyösebbnek, ha minden komponensben a centrumbeli csúcsok a legfontosabbak, hiszen így a fizikai adottságokhoz képest a lehető legtöbb csúcs a lehető legközelebb van a legfontosabbhoz. Azt kaptuk, hogy a 27 komponensből 19 esetén a legfontosabb csúcs benne van a centrumban. Ez borzasztóan erős korrelációt jelent, hiszen egy n csúcsú gráf esetén erre legfeljebb $\frac{2}{n}$ az esély. (Megjegyzendő, hogy a maradék 8 komponensből is 6 esetén a legfontosabb csúcsnak van centrumbeli szomszédja, tehát szintén a komponens aránylag közepén helyezkedik el.)

Ez az eredmény alátámasztja a hierarchiás felépítettségét az agynak. A legfontosabb csúcsokhoz alá vannak rendelve kevésbé fontosak, amik természetesen rétegesen helyezkednek el aszerint, hogy mekkora a gráfelméleti távolságuk a legfontosabb csúcstól. Tehát ha egy feladatot a legfontosabb csúcs lebont több, kisebb kiértékelendő döntésre, akkor a kevésbé fontos csúcsok, csúcscsoportok kiszámolják lokálisan a saját feladatukat, de az eredménynek vissza kell jutnia a legfontosabb csúcshoz. Az elvégzési idő természetesen akkor lesz a lehető legkevesebb általánosságban, ha a csúcsok távolságai minél egyenletesebben vannak elosztva a legfontosabb csúcstól. Ez pedig természetesen egybeesik azzal, hogy a legfontosabb csúcs a centrumban van, vagy attól nem messze. (Természetesen fizikai adottságok is korlátozzák az agy lehetőségeit, pl. adott csúcsból nem futhat ki tetszőlegesen sok él. Másrészt a fontos csúcsok "döntései" is időt vesznek igénybe, így az sem előnyös, ha túl sok él fut a legfontosabb csúcsokba.) Ez tehát egy újabb bizonyíték az agyterületek hierarchikus felépítettségére és az agy ebből a szempontból rendkívül előnyös kialakulására.

8.3. Bővített maximális élekből álló gráf

A bővített maximális gráfok nem erdők, hiszen ismert, hogy minden fának van levele, jelen esetben viszont minden csúcshoz legalább két különböző él ki van választva. Azt

kaptuk, hogy az agygráfra ez a gráf már összefüggő.

Felmerül az ötlet, hogy ez a gráf lehet akár kétszeresen (csúcs)összefüggő is. Azt kaptuk, hogy az 1015 csúcsból csak két olyan csúcs van, amit (az éleikkel együtt) elvéve már nem lenne összefüggő a gráf, tehát két csúcs híján kétszeresen összefüggő a gráf. Megjegyzendő, hogy ez a két csúcs a legfontosabb csúcsok közé tartozik és elvételükkel az agygráf egy kb. 1000 csúcsból álló komponensre és néhány kevésbé fontos csúcs kisebb komponensére esik szét. Ez is mutatja azt, hogy az agy nagyrésze még ezek hibásodásával is erős éleken tud kommunikálni. Ezen kívül megerősíti azt az elméletünket is, hogy az agy hasonló szisztémájú, mint az erős párhuzamos számolásokat végző számítógépek. Hasonló hierarchikus felépítettség van a csúcsok közt: valamilyen értelemben kiemelt (fontos) csúcsok alá tartoznak csúcscsoportok, amik egymással erősebben kommunikálnak, majd a "főnökük" felé szintén erős élen át vissza tudják juttatni a kapott eredményt, kiértékelést.

Fontos megjegyezni ezen kívül, hogy ez egy újabb bizonyíték a Kerepesi, Szalkai, Varga, Grolmusz cikkében leírtakra [10]. Leegyszerűsítve az áll benne, hogy ha a teljes agygráfnak csak a legvastagabb/leggyakoribb éleit vesszük a gráfunkba, akkor az terjeszkedő szerkezetű lesz. Ez alatt az értendő, hogy amikor a vastagsági alsó korlátot csökkentjük, akkor az új élek ritkán lesznek külön komponensben. Tehát ahogy folytonosan csökkentjük az alsó határt ahonnan az élek bekerülnek a gráfba, úgy a gráf mindig majdnem összefüggő lesz. Ez pedig az előző erős összefüggőség miatt teljesül, hiszen minden csúcsra a két legvastagabb éle kerül be először a gráfba, amik által meghatározott gráf már összefüggő. (Persze előfordulhatna, hogy a második legvastagabb élek súlyai minimálisak, így mégsem egyenletesen kerülnek be a gráfba a legvastagabb élekkel, de ez jelen esetben nem áll fenn. Ezen kívül a legvastagabb élek is egy kevés komponensű gráfot határoznak meg, így a szélsőséges esetek sem okoznak problémát.)

Felmerül a kérdés, hogy az összefüggőségi tulajdonságoknak milyen biológiai szerepe van az agyban. Gondoljuk meg, hogy miért fontos az agynak, hogy az erős élei összefüggően épülnek fel. Az összefüggőség azt jelenti, hogy bármely két csúcs tud kommunikálni egymással a megadott idegrostokon keresztül. Természetesen az agy akkor lesz a legeffektívebb, hogy ha a kommunikáció bárhonnan bárhova a lehető legjobb/leggyorsabb éleken keresztül történik. Az előbb említett tulajdonság persze pont akkor teljesül, ha a legfontosabb élek behálózzák az egész agyat úgy, hogy bármelyik két csúcs tud kommunikálni egymással.

A kétszeresen összefüggőség azt jelenti, hogy bármelyik csúcs hibásodása, vagy gyengülése esetén is el tudunk jutni bármely csúcsból bármely másikba. Ez pedig az agynak egy nagyon fontos tulajdonsága. Hiszen mind betegségek miatt, mind az öregség folyamán fontos képesség az, hogy adott csúcsok szerepét az agy mellőzni tudja, néhány csúcs át tudja venni legalább részben más csúcsoknak a szerepét (agyi plaszticitás vagy neuroplaszticitás).

Eszünkbe juthat, hogy az átlagolt agygráf nem feltétlenül ugyanolyan tulajdonságokkal rendelkezik, mint az egyénenkénti agygráf. Egyrészt lehetne probléma abból is, hogy az átlagolt agygráfban vannak olyan pici élek is, amik csak egy-egy emberben vannak jelen. Ez most nem okoz gondot, hiszen az ilyenek egyébként sem lesznek kiválasztva, mint maximális élek. Ettől függetlenül érdemes megnézni az 1064 ember agygráfjaira különkülön, hogy milyen összefüggőségi adatokat kapunk.

Emberenként kiszámoltuk, hogy a maximális élekből álló gráfja hány komponensű,

ezen kívül azt, hogy a bővített maximális gráfja hány csúcs híján kétszeresen összefüggő. Azt kaptuk, hogy átlagosan 64.04 komponensű az emberek maximális gráfja (a legkisebb száma a komponenseknek 36, míg a legnagyobb 100). Átlagosan a bővített maximális gráf 63.9 csúcs híján összefüggő. Azt kaptuk tehát, hogy az egyének eredményei számottevően rosszabbak, mint az átlagolt agy eredménye. Természetesen szignifikáns eredményről beszélünk, hiszen mind az 1064 alany agya rosszabb eredményt adott, mint az átlagolt gráf. Ennek a jelenségnek a magyarázatára a dolgozat későbbi részében térek vissza.

8.4. Véletlen gráfok

A célunk az, hogy valamilyen módon az agygráfot hasonlítsuk össze véletlen gráfokkal. Természetesen nem érdemes abszolút véletlen módon összekötött gráfokkal összehasonlítani, hiszen a fizikai adottságok miatt valamennyire rögzített az agygráf valamilyen értelemben vett szerkezete (pl. két agyfélteke). Emiatt rögzítsük le az agy 1015 indexekkel ellátott csúcsát és azt is, hogy melyik csúcsok között fut él. Tehát pontosan azon csúcspárok közt legyen él a véletlen gráfokban, mint az átlagolt agyban. A random faktor az legyen, hogy melyik élnek mekkora a súlya. Természetesen számolhatnánk abszolút random számokkal, de a realitáshoz ragaszkodva a súlyok legalább akkorák legyenek, mint az agyban előforduló legkisebb él súlya (0.0009) és legfeljebb akkorák, mint a legnagyobb él súlya (3115.94). (Megjegyzendő, hogy tényleg teljesen random számokat is sorsolhatnánk súlyoknak. Ugyanis a maximális gráfok összefüggőségi vizsgálatánál csak az számít, hogy az élek csökkenő sorrendje micsoda, szóval valójában az éleknek egy random permutációját vesszük. Tehát lényegében a random gráf esetében az élek vastagságának egy sorrendjét sorsoljuk ki.) Fontos megemlíteni, hogy azzal, hogy az agy szerint rögzítjük a random gráfunk éleit, azzal már a random gráfoknak egy jól terjeszkedő csoportját választom ki. Így, ha az agy ezeknél is effektívebben épül fel, az tényleg az agy evolúciójának ebből a szempontból rendkívül előnyös kialakulására utal.

Az előbb leírt módon generáltunk 5000 véletlen gráfot és megnéztük, hogy azokra behúzva a csúcsok legvastagabb éleit a gráf hány komponensű lesz. Azt kaptuk, hogy átlagosan 224.76 komponensből áll egy ilyen gráf, ami lényegesen több, mint az agy esetében kapott 27. Ez persze borzasztóan nagy különbség, főleg azt is figyelembe véve, hogy a random gráfokra kapott komponens számoknak kicsi a szórása. A legkisebb komponens szám 196, míg a legnagyobb komponens szám 255. Másrészt amikor a két legvastagabb élt húzzuk be, akkor annak átlaga, hogy hány csúcs híján kétszeresen összefüggő a gráf az 2.97. Ebben az esetben viszont pusztán az átlagot látni csalóka, mert az átlag úgy alakul ki, hogy a gráfok többsége legfeljebb 3-4 csúcs híján már kétszeresen összefüggő. Viszont vannak olyan gráfok is, amik már alapvetően sem összefüggőek, így pedig bármely csúcsát elhagyva sem lesz összefüggő. Így ezek 1015 csúcs híján összefüggőek, ami torzítja az átlagot.

Természetesen meg kell győződnünk arról, hogy az először kapott eredményünk statisztikailag szignifikáns eredményt ad. Jelen esetben a tipikus statisztikai próbák nem megfelelőek, hiszen nem ismerjük sem a várható értéket, sem az eloszlásfüggvényt. Ezért azt fogjuk vizsgálni, hogy mekkora annak az esélye, hogy mind az 5000 random gráf több komponensből áll, mint az agy. A cél az, hogy szignifikáns eredményt tudjunk levonni arról, hogy az agy a legeffektívebb gráfok hány százalékába eshet, tehát megtudjuk, hogy mennyire a lehető legjobban alakult ki az agy. Az előre rögzített szignifikanciaszintünk 99%. Ekkor tehát feltesszük, hogy az agyra kapott értéknél a gráfok x-ede több komponensből áll. Akkor kapunk tehát szignifikáns eredményt, ha $x^{5000} < 0.01$, vagyis

ha $x < 0.999$. Ekkor tehát szignifikánsan megkaptuk, hogy a gráfok 99.9%-ánál az agy effektívebben van felépítve ebből a szempontból.

Mit kaptunk az eddigiekből? Azt, hogy a random gráfoknak az agyhoz hasonló összeköttöttségére kicsi az esély. Persze ez annyira nem meglepő, de ez még egy erős bizonyíték az agy optimális kialakulására, hiszen a valószínűség tényleg elhanyagolható és a random gráfok ráadásul ismerten jól terjeszkedőek. Itt fontos kitérni arra a jelenségre, hogy az emberek agygráfjai külön-külön rosszabb eredményeket adnak, mint az átlagolt agygráf. Ennek oka szintén az evolúciós tökéletesség iránya: az emberiség átlagos agya jobban teljesítene, mint a külön-külön egyedeké. Természetesen a génszelekció miatt a születő új egyedek valamilyen értelemben a keverékei lesznek a felmenőinek. Ennek okán az emberiség átlagos agya hasonló szerkezetű, mint ami felé az evolúció során fejlődik az agy. Természetesen a folyamatot erősen befolyásolja a génszelekció és az, hogy az idő múlásával más-más gének lesznek előnyösek. Ebből kifolyólag nem tisztán csak átlagolódni fognak a jövőbeli egyedek génjei. Emiatt az evolúciós okoskodás miatt tehát érthető, hogy az átlagos agy jobban teljesít ebből a szempontból, mint külön-külön az egyének agya, hiszen az evolúció során is a fajok fejlődnek, előnyösebb tulajdonságokat vesznek fel.

Eszünkbe juthat, hogy az, hogy a súlypontok helyzete összefügg az adott csúcs fontosságával az köszönhető a korábban említett optimális hierarchikus szerkezetnek. A párhuzamos számolások optimális processzor-felépítésében is a processzorok közt megfigyelhető egy hierarchia. Minél inkább fontosabb egy csúcs, annál inkább fentebb helyezkedik el a hierarchiában. Ez pedig azt jelentheti az agynál, hogy azok a csúcsok, amik jobban szétosztják a feladatokat, információkat, azok a fontosabbak. Megjegyzendő, hogy az egymáshoz közel álló csúcsok sokszor ugyanazon agyterülethez tartoznak, így sokszor az adott területet irányító csúcs össze van kötve az összes, ezen a helyen levő csúccsal. Másrészt, ha a fontos csúcsok egyesével lennének felelősek adott agyterületekért, nem pedig valamennyire megosztva, akkor az agyterület hibásodása, gyengülése esetén nagy kár keletkezhetne az agyban. Ekkor tehát az agynak az az előnyösebb, ha a kitüntetett felelős csúcs mellett az adott agyterülettel számottevően kommunikálnak más csúcsok is. Tehát a szimmetrikus kommunikáció oka összefügg az agy neuroplaszticitásával is.

9. Összegzés

- Beláttuk, hogy a különböző agyterületek intuitív és kevésbé intuitív fontossági szempontjai is korrelálnak. Láttuk, hogy a sok éllel rendelkező csúcsok ugyanazok, mint amiknek több értelemben hosszú vagy vastag élei vannak.

- Összehasonlítottuk az élek vastagság, hossz és előfordulás szerinti rendezéseit.

- Foglalkoztunk azzal, hogy az adott csúcsok szomszédainak súlypontja hol helyezkedik el. Összevetettük az itt kapott eredményeket a fontossági szempontokkal és összefüggést kaptunk köztük.

- Megvizsgáltuk azt, hogy a gráf csúcsonkénti legfontosabb élei hogyan helyezkednek el az agyban. Egy általános lemma bizonyítása után láttuk, hogy a vizsgált véletlen gráfok döntő többségénél az agy ebből a szempontból lényegesen előnyösebben alakult ki. Láttuk, hogy az előbbi témák eredményeit összeköti a biológiailag rendkívül fontos agyi plaszticitás. Megmutattuk azt is, hogy a gráf csúcsonkénti legfontosabb

élei által alkotott részgráfban a komponensekben a legfontosabb csúcsok szinte mindig gráfelméletileg középen helyezkednek el.

Köszönetnyilvánítás

Köszönöm Csikvári Péternek a kutatási lehetőség megszervezését és Grolmusz Vincének a témavezetését. Köszönet illeti Varga Bálintot, hogy az általa kiszámított gráfokkal dolgozhattam.

Hivatkozások

[1] Jennifer A. McNab, Brian L. Edlow, Thomas Witzel, Susie Y. Huang, Himanshu Bhat, Keith Heberlein, Thorsten Feiweier, Kecheng Liu, Boris Keil, Julien Cohen-Adad, M Dylan Tisdall, Rebecca D. Folkerth, Hannah C. Kinney, and Lawrence L. Wald. The Human Connectome Project and beyond: initial applications of 300 mT/m gradients. Neuroimage, 80:234–245, Oct 2013. doi: 10.1016/j.neuroimage.2013.05.074. URL http://dx.doi.org/10.1016/j.neuroimage.2013.05.074.

[2] Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, et al. Mapping the structural core of human cerebral cortex. PLoS Biol. 2008 Jul;6(7):e159. Available from: https://doi.org/10.1371/journal.pbio.0060159

[3] Csaba Kerepesi, Balázs Szalkai, Bálint Varga, Vince Grolmusz: The braingraph.org Database of High Resolution Structural Connectomes and the Brain Graph Tools, Cognitive Neurodynamics Vol. 11 No. 5, pp. 483-486 (2017) http://dx.doi.org/10.1007/s11571-017-9445-1

[4] Balázs Szalkai, Csaba Kerepesi, Bálint Varga, Vince Grolmusz, The Budapest Reference Connectome Server v2.0, Neuroscience Letters, Vol. 595 (2015), Pages 60-62, http://dx.doi.org/10.1016/j.neulet.2015.03.071.

[5] Balázs Szalkai, Csaba Kerepesi, Bálint Varga, Vince Grolmusz, Parameterizable Consensus Connectomes from the Human Connectome Project: The Budapest Reference Connectome Server v3.0, , Cognitive Neurodynamics, (2016), http://dx.doi.org/10.1007/s11571-016-9407-z. https://pitgroup.org/connectome/

[6] Bardin, J. Neuroscience: Making connections. Nature 483, 394–396 (2012). https://doi.org/10.1038/483394a

[7] Spearman, C. "The Proof and Measurement of Association between Two Things." The American Journal of Psychology, vol. 15, no. 1, 1904, pp. 72–101. JSTOR, https://doi.org/10.2307/1412159.

[8] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.spearmanr.html

[9] https://arxiv.org/abs/2212.00168

[10] Kerepesi C, Szalkai B, Varga B, Grolmusz V (2016) How to Direct the Edges of the Connectomes: Dynamics of the Consensus Connectomes and the Development of the Connections in the Human Brain. PLoS ONE 11(6): e0158680. https://doi.org/10.1371/journal.pone.0158680