

Graphletek funkcionális eloszlása a *C. elegans* konnektómban

Vassy Zsolt¹, Fenyves Bánk Gábor¹,

Szilágyi Gábor Sándor¹, Csermely Péter¹

¹Semmelweis Egyetem, Molekuláris Biológiai Tanszék

1094 Budapest, Tűzoltó utca 37-47.

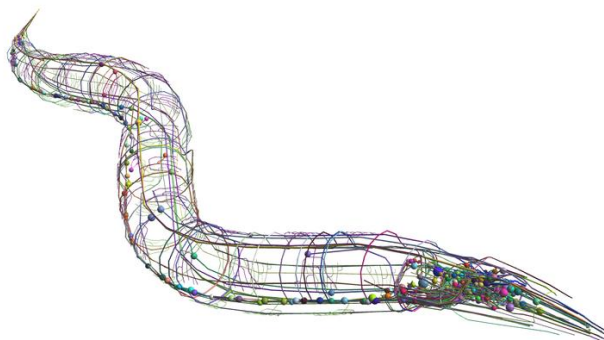
zsolt.vassy@gmail.com

Összefoglaló: A *C. elegans* talajféreg a neurobiológiai kutatások népszerű modell szervezete, neurális hálózata ismert. A neurális hálózat irányított hálózat, az ingerület terjedésének iránya van, ezért az ingerület terjedésben részt vevő graphletek szerepét akkor lehet megérteni, ha működési környezetben, funkcionálisan vizsgáljuk, hogy hogyan vesz részt az állat viselkedésében. A funkcionális elemzés számítási költségei magasak, a tanulmányban bemutatunk egy olyan módszertant, amivel ezt a költséget töredékére csökkentettük. A módszert alkalmazva, látható, hogy bizonyos típusú graphletek funkcionálisan mért eloszlása jelentősen eltér az egyszerű számbeli eloszlástól, ezeknek kiemelt szerepe lehet a jelátvitel szabályozásában.

Bevezető

A *Caenorhabditis elegans* (*C. elegans*) egy talajban élő, átlátszó fonálféreg, népszerű kutatási célpontja a biológiai kutatásoknak, mivel az egyik legegyszerűbb, soksejtű élőlény. 1963-ban Sydney Brenner (2002-ben e munkásságáért is Nobel-díjat kapott) biológus javasolta az idegfejlődés modell állatának, ezután sokan tanulmányozták a *C. elegans* egyedfejlődését, genetikáját, molekuláris biológiáját, ezáltal az állat a biológiában általános elfogadott modell szervezetté vált [1].

1. sz. Ábra



A *C. elegans* az első soksejtű élőlény volt, amelynek meghatározták a teljes genom szekvenciáját, és az első állat volt, amelynek a teljes neuron kapcsolódási hálózatát (konnectómját; ld. 1. sz. Ábra) leírták [2,3,4].

A hermafrodita *C. elegans* férgek teljes idegrendszere 302 sejtből áll [5]. Az idegrendszer részletes tanulmányozását lehetővé tette, hogy az állat átlátszó, ezért élve is könnyen tanulmányozható, laboratóriumi körülmények között olcsón szaporítható, és lefagyasztható [1].

A rendelkezésre álló részletes tudományos irodalom miatt a *C. elegans* megkerülhetetlen a neurobiológiai kutatások során. Kutatócsoportunk is ezen állat neurális hálózatának vizsgálatával foglalkozik, célunk a 3 neuronból álló graphletek ingerület átvitelben betöltött szerepét megtalálni (a graphlet egy indukált algráf: tartalmaznia kell minden él a csomópontjai között, amelyek a nagy hálózatban is jelen vannak).

A korábbi, mintakereséssel foglalkozó tanulmányok nem vették figyelembe [6, 7], hogy a neurális hálózat speciális hálózat abban az értelemben, hogy bizonyos irányokba (főleg szenzoros neurontól motorikus neuronig) gyakoribb az információterjedés mint más irányokba, illetve bizonyos terjedési irányok biológiai szempontból nézve értelmetlenek, ezért az élő szervezetben nem fordulnak elő.

Vizsgálatunk célja elemezni a három nódusból álló graphletek irányított ingerület terjedésben betöltött szerepét, a különböző típusú graphletek utak szerinti eloszlását.

Adatok

Ezen vizsgálathoz a bemeneti adatok a C. elegans konnektómot jelentik, a szakirodalom által elismert és sűrűn használt konnektóm a WormWiring Project konnektómja [7], amely 300 nódusból, 5039 irányított élből áll.

A kémiai szinapszisokat a hálózatban egyirányú élek írják le, az elektromos (elektrokémiai) kapcsolatokat leíró éleknek pedig van egy azonos súlyú párja, amelyek a másik irányba mutat.

A konnektómból épített hálózat átmérője 6. 348.276 háromszög graphlet található ebben a hálózatban.

Módszer

A neurális hálózaton két neuron között a különböző útvonalak bejárása nem azonos valószínűséggel következik be, az útvonalakhoz különböző valószínűséget rendelünk: ez az adott útvonal Markov száma.

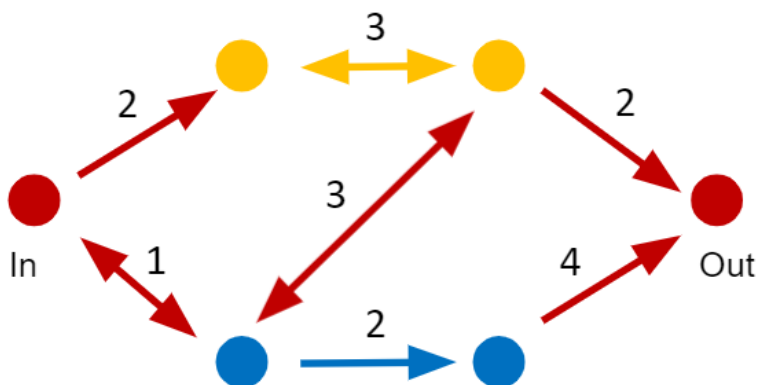
A Markov szám a hálózatban az adott Markov-lánchoz (sztochasztikus folyamathoz) rendelt valószínűség, az értékét úgy kapjuk meg, hogy összeszorozzuk az egy élhossznyi utak megtételéhez tartozó valószínűségeket. Például egy három élt tartalmazó út Markov száma:

$$P_{\text{Markov}} = P_{e_1} * P_{e_2} * P_{e_3}$$

ahol $P_{e_1}, P_{e_2}, P_{e_3}$ az e_1, e_2, e_3 éleken megtett utak valószínűsége.

Egy adott élen megtett út valószínűsége az él súlya elosztva az összes kimenő élsúly összegével.

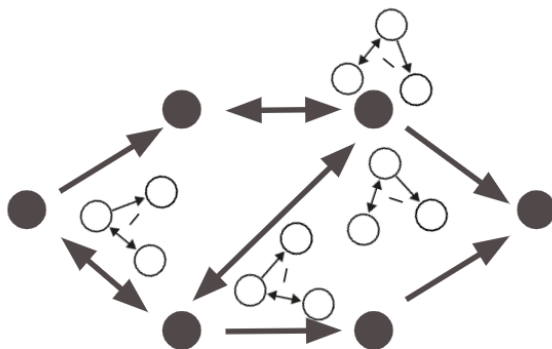
3. sz. Ábra



Egy adott út a neurális hálózaton többféle útvonalon is megvalósulhat (2. sz. ábra. Az In és Out nódusok között 7 féle út is létrejöhet) ezek Markov számát egyenként kiszámolva és összeadva megkapjuk a kumulált Markov számot.

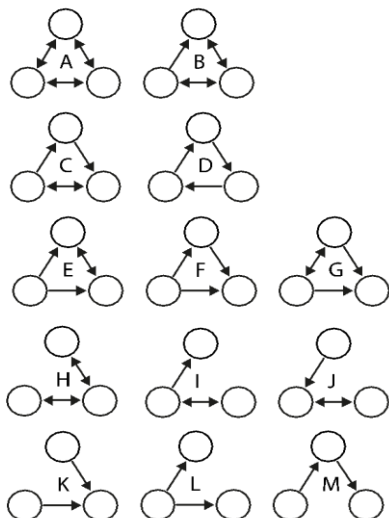
A graphletek hálózaton belüli funkcionális eloszlását az őket érintő útvonalak kumulált Markov száma alapján számítjuk, ezáltal kiemeljük a graphleteknek a valódi jelátvitelben betöltött szerepét. Ha egy graphlet egy olyan útvonalon (egy vagy több élen) csatlakozik a hálózatba, amelynek a funkcionális szerepe fontos, akkor ez a csatlakozott útvonal a graphlet funkcionális szerepét is kiemeli.

4. sz. Ábra



A 3 sz. Ábra megmutatja, hogy az előbb mutatott példa hálózatban, hol találhatóak háromszög graphletek, és hogyan illeszkednek az útvonalakra.

5. sz. Ábra



A neurobiológiai szempontból értelmes, bemenettel és kimenettel is rendelkező háromszög graphlet típusból 13 van, ezeket a 4 sz. Ábra mutatja be (a bal alsó nódus a bemeneti nódus, a jobb alsó pedig a kimeneti, fontos kritérium, hogy a két pont között kell léteznie irányított összeköttetésnek).

A konnektóm 302 neuronjából 300 nem izolált, vagyis van közöttük összeköttetés, ez azt jelenti, hogy $300 \cdot 299$ vagyis 89.700 potenciálisan összekötött neuron-pár van.

A hálózat átmérője 6, vagyis egy véletlenszerűen kiválasztott nódusból maximum 6 lépéssel elérhető mindegyik másik nódus. Egy magas átlagfokszámú hálózaton, amilyen a *C. elegans* konnektóm, minél több lépésből áll a Markov-lánc, annál kisebb lesz a Markov szám (mivel a kimenő fokszám minden esetben nagyobb mint 1 ennél a hálózatonál, ezért minden esetben egy 1-nél kisebb számmal szorozzuk minden lépésnél a meglévő értéket). Ezért a legfeljebb 6 lépésből álló útvonalakat vettük figyelembe, mivel a kumulált Markov szám kiszámításánál a rövidebb útvonalak a meghatározóak és 6 lépésnél mindig található azonos vagy rövidebb út.

A C. elegans konnektóm esetében ez így is 88.726.809.432 útvonalat jelent, ezekben kellene megnézni a 348.276 háromszög graphlet összesen 484.140 különböző útvonalának (bemenő és kimenő nódus közötti graphlet útvonalak) előfordulását. Ez összesen **4.2956198 e+16** műveletet jelentene. Ugyan a feladat jól párhuzamosítható, de a feladat extrém költség igényes.

Hamar látszott, hogy ha minden útvonalban, minden graphlet útvonalat megkeresünk, az a mi erőforrásainkkal nem kivitelezhető. Ezért megkerestük két neuron közötti útvonalak legnagyobb Markov számmal rendelkező 1 ezrelékét (az összeset megkerestük, de csak a felső 1 ezreléket tartottuk meg), és ezekre illesztettük a graphlet útvonalakat, mert a méréseink alapján kevés eltérés volt az így kapott eredmények és a teljes illesztésre kapott eredmények között. Ennek a kompromisszumos módszernek is egy szátra számolva 1080 óras futási ideje volt, ezen a viszonylag kis méretű hálózaton.

Mivel a konnektómban lévő graphleteket előre meghatároztuk és a graphletekben található útvonalak is előre ismertek voltak, rájöttünk, hogy főleg az egész hálózatot bejárni és graphlet útvonalakat keresni a bejárt útvonalakon. **Ha két nódus közötti mintaeloszlást keresünk, elegendő a kiinduló nódusból az előre megtalált graphletek bemeneti nódusáig megtalálni az útvonalakat, illetve a graphlet kimenő nódusából az útvonal záró pontjáig. A Markov szám matematikai módszertana alapján a graphleten átmenő útvonalak kumulált Markov száma meg fog egyezni a graphletbe bemenő, az azon átmenő és az abból kimenő útvonalak kumulált Markov számainak szorzatával.**

Nehézség, hogy a legfeljebb 6 hosszúságú útvonalakat keressük továbbra is, viszont a graphleten belül lehet 1-5 hosszúságú belső útvonal is. Ezért a graphletbe bemenő, abból kimenő, illetve azon átmenő útvonalakat a hosszúságuk szerint kell csoportosítani, és összegezni.

Kiválasztva a graphletek útvonalait, és csak az ehhez illeszthető útvonalakat bejárva, úgy, hogy legfeljebb 6 lépésből álljon a neurális út, a számítás erőforrás költsége a töredékére esett.

Eredmények

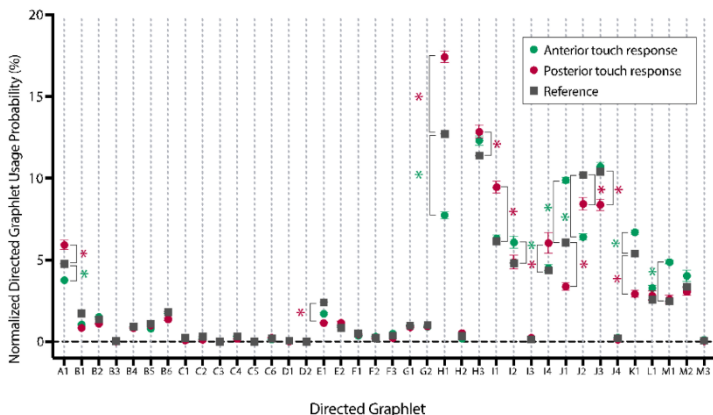
Az eredeti módszerrel végül az útvonalak legfontosabb 1 ezrelékénél kerestük meg a graphleteket (egyszálas futtatás esetében) 1080 óras futási idő alatt, ezt sikerült 12 óras futási időre leszorítani, úgy hogy az összes útvonalra megtaláljuk az illeszkedő graphleteket. **Azaz a hatékonyság javulás mértéke több mint 90.000-szeres!**

Az eredmények biológiai jelentőségét még elemezzük, de a valós utak (szenzoros neuronból motorikus neuronba) graphlet eloszlása nagyon eltérő

az átlagos graphlet eloszlástól, vagy a biológiailag invalid utak graphlet eloszlásától (motorikus neuronból szenzoros neuronba).

A valós utak graphlet eloszlása is mutat funkcionális eltéréseket. Az állat első felét ért inger hatására (anterior touch response) is hasonló választ ad, mint a hátsó felét ért behatásra (posterior touch response), vagyis ellentétes irányba menekül. A hatásra adott fizikai válasz hasonló, ugyanakkor ezek graphlet eloszlása (graphlet típusokra vetítve) eltérő (5. sz. Ábra).

6. sz. Ábra



Diszkusszió

A graphletek jelátvitelben betöltött szerepéről valódi képet akkor kapunk, ha ezek funkcionálisan betöltött szerepét is vizsgáljuk.

A hálózatok topológiája szoros összefüggésben van az általuk betöltött funkcióval, így van ez a hálózatban előforduló minták, a kisebb alhálózatok tehát a graphletek esetében is, ezek típusa meghatározza az ingerület átvitelben betöltött szerepüket.

Ez a módszer a *C. elegans* neurális hálózatán, egy gyakran elemzett neurális hálózaton tud hozni új eredményeket, de ilyen szintű hatékonyság javítással nagyobb hálózatokon is alkalmazható.

Hivatkozások

- [1] S. Brenner, „The Genetics of *Caenorhabditis elegans*”. *Genetics* 77 (1), 71–94. o., 1974, PMID 4366476.
- [2] J. G. White, E Southgate, J. N. Thomson, S. Brenner „The structure of the nervous system of the nematode *Caenorhabditis elegans*”. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 314 (1165), 1–340. o. DOI:10.1098/rstb.1986.0056. PMID 22462104.
- [3] F. Jabr: “The Connectome Debate: Is Mapping the Mind of a Worm Worth It?”. *Scientific American*, 2012. <https://www.scientificamerican.com/article/c-elegans-connectome/>
- [4] Cook et al. “Whole-animal connectomes of both *Caenorhabditis elegans* sexes” *Nature* 571, 63–71 (2019). <https://doi.org/10.1038/s41586-019-1352-7>
- [5] R. A. Kosinski, M. Zaremba „Dynamics of the Model of the *Caenorhabditis elegans* Neural Network”. *Acta Physica Polonica B* 38 (6), 2201–2210. o., 2007
- [6] O. Sporns, R. Kötter, 2004. “Motifs in Brain Networks”. *PLoS Biol* 2(11): e369. <https://doi.org/10.1371/journal.pbio.0020369>
- [7] R Milo, S Shen-Orr, S Itzkovitz, N Kashtan, D Chklovskii, U Alon “Network motifs: simple building blocks of complex networks.” *Science*. 2002 Oct 25; 298(5594):824-7. doi: <https://doi.org/10.1126/science.298.5594.824>. PMID: 12399590.
- [8] WormWiring Nematode Connectomics: <https://wormwiring.org/>