

Bírálat Farkas Illés: „Átfedő modulok molekuláris biológiai kölcsönhatási hálózatokban” című doktori munkájához

A doktori dolgozat a biológia és a statisztikus fizika határterületén található kutatási témákat és eredményeket ismerteti. Napjainkban a molekuláris biológia mérési lehetőségei, valamint az így keletkező nagy adathalmazok feldolgozásának számítástechnikai képességei oda vezettek, hogy a fehérjék kölcsönhatásait soha nem látott részletességgel tudjuk leírni és megérteni. A kölcsönhatások leírására a a szociofizikából kifejlődött hálózattudomány módszereit alkalmazzák. Fehérjék kölcsönhatásán a direkt kontaktuson kívül olyan relációt is érthetünk, melynél a proteinek egy adott biológiai feladatban közösen vesznek részt. Az ilyen funkcionális csoportot is szokás modulnak nevezni. A fehérjék és a modulok kölcsönhatási hálóinak feltárása az alapvető működések megértésén kívül nagyon fontos például a gyógyszeripar vagy az orvostudomány szempontjából is. Farkas Illés MTA doktori címre benyújtott dolgozata ilyen szempontból kiemelkedően aktuális és tudományos szempontból fontos.

A doktori mű témáját képező disszertáció és összefoglaló igen gondos kiállítású, esztétikus, színes ábrákkal ellátott olvasmány. Az eredmények tárgyalása logikusan felépített, precíz munkára enged következtetni. A dolgozat négy részből áll, a bevezető 6 oldal után az első 59 oldal a fehérje-fehérje kölcsönhatási hálózatok moduljaival kapcsolatos kutatásokat mutatja be. Az első fejezet eleje ismerteti a fehérjék kölcsönhatásait mérő és kiértékelő módszereket, míg a második rész a szerző közreműködésével elért eredményeket mutatja meg. A második 25 oldal transzkripció és transláció szabályozási hálózatok moduljairól szól. Itt is először egy bevezető részt -- mely a fehérje-RNS-DNS kapcsolódásokról szól -- követnek a saját eredmények. A sör élesztőgomba transzkripció szabályozása mellett az emberi mikroRNS-ek transláció hálózati moduljainak feltárásával ismerkedhet meg az olvasó. A szerző elsősorban adatbázis és statisztikus analízisekről számol be. Végül az utolsó 10 oldal biokémiai jelek sejt felszíntől a sejtmagig történő átviteli hálózati moduljaival foglalkozik. Itt túlnyomórészt adatbázis kezelési és fejlesztési eredményekről olvashatunk, melyek elsősorban a gyógyszerkutatók számára relevánsak. A dolgozat első és második fejezetének célkitűzése döntően alapvető kutatási eredmények ismertetése, míg a harmadik fejezetben az alkalmazásokon van a hangsúly.

Az új eredmények 9 tézispontban lettek összefoglalva, a tézispontokhoz kapcsolódó tudományos közlemények száma 12, többnyire magas impakt paraméterű folyóiratokban megjelentetve. Ezek között 1 Nature és 2 Plos One publikáció is található. Ezekre a publikációkra összesen kb. 5000 független hivatkozást találtam a google scholar alapján, cikkenként 0 és 3350 közötti hivatkozással. A kiváló scientometriai adatok magukért beszélnek.

Társszerzőivel együtt a legtöbb hivatkozást eddig a „[Uncovering the overlapping community structure of complex networks in nature and society](#)” című Nature cikkre kaptam, melyben az átfedő modulok hálózatelméleti felkutatását alapozza meg. Ebben Farkas Illés részt vett a klikk perkoláción alapuló hálózati modulkereső CFinder algoritmus kidolgozásában és alkalmazásában az általa gyűjtött protein, szó asszociációs és tudományos társszerzőségi hálózatok esetén. Részletes modul méret és kapcsolati szám eloszlásokat határozott meg statisztikus analízissel a fenti rendszerekben.

A disszertáció és a tézisek gyűjteménye alapján bátran kijelenthető, hogy a mű hiteles adatokat tartalmaz és a jelölt magas szinten alkalmazza a statisztikus fizika legmodernebb eszközeit. A bemutatott kutatások ideálisan sokoldalúak: magas színvonalú, elméleti numerikus analízisek és kísérleti mérések megfelelő kombinációja. A tézisekben megfogalmazott tudományos eredményeket elfogadom és a művet nyilvános vitára alkalmasnak tartom.

A dolgozathoz az alábbi kérdéseket nyújtom be:

1. A 31. oldal lábjegyzete szerint az általánosított kisvilág modell definíciója az alacsony átmérő és a magas klaszterezettség. Ettől eltérő definíciót ismerek és használtunk saját publikációinkban, mely reguláris rácshoz hozzáadott hosszú éllel van meghatározva. Mi az eltérő definíciók oka ?
2. A 32. oldal alján Erdős Rényi véletlen gráfra az az állítás van megfogalmazva, hogy ha az átlagos fokszám nagyobb mint 1, akkor a végtelen limeszben a gráf összes csúcspontja egyetlen komponensbe fog tartozni. Nem lehetséges, hogy a véletlenszerűség miatt kimarad néhány csúcs az óriás komponensből ?
3. Az 54. oldal alján az olvasható, hogy a „CPM módszer ROC görbéje jobb (magasabban halad), mint a CPM módszer”. Kérem fejtse ki, hogy ez alatt mit kell érteni pontosan.
4. A 72. oldalon szerepel, hogy a fehérje modulok fokszám eloszlása exponenciális típusú, majd az eloszlás vége hatványfüggvény-szerű. Melyik ábra mutatja az utóbbi komplex eloszlási görbét ? Ha az exponenciális lecsengésnek az oka az alkalmazott klikk paraméter, akkor ez a modulkeresés jellemzőjének tűnik és nem a hálózatra jellemző ?
5. A fehérje hálózatok kapcsolatán túl léteznek olyan tulajdonságok, melyek ezen hálózatokon definiált elágazó, bolyongó és annihiláló terjedési modellekkel jellemezhetőek ? Ismertek-e eredmények ilyen modellekben a dinamikus viselkedésről ?

Budapest 2016. február 3

Dr. Ódor Géza
az MTA doktora