

Élet a törpe komponensekben

JORDÁN FERENC

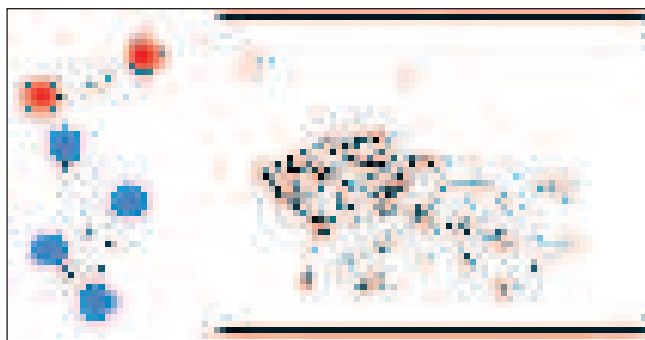
A nagy komplex hálózatok tanulmányozása során a kutatók gyakran beszélnek úgynevezett óriás komponensről (giant component). Ennek nincs különösebben pontos definíciója, egyszerűen csak annyit jelent, hogy sok hálózat pontjainak döntő többsége egyetlen nagy komponensbe tartozik, a többi pont pedig sok kicsi komponensbe tartozik. Egy komponensbe akkor tartozik két pont, ha egymásból úttal elérhetőek (az egyszerűség kedvéért beszéljünk most irányítatlan utakról és irányítatlan hálózatokról). Egy-egy komponensen belül tehát, ha valami történik az egyik ponttal (a hálózat típusától függően ez lehet például egy fehérjemolekula, egy egyed vagy egy állatfaj), akkor elvileg az hatással lesz a komponens többi pontjára, de a többi komponensbe tartozó pontokra nem. Egy hálózat állhat egyetlen komponensből is, ilyenkor azt mondjuk, összefüggő. A komplex hálózatok tehát legtöbbször több komponenset tartalmaznak, és ezek méreteloszlása egyenetlen, majdnem mindig van közöttük egy kiemelkedően nagy. Az **1. ábra** bal oldalán egy kicsi hálózatot látunk, mely két komponensből áll: a kék a nagyobb, ebbe 4 gráfpont tartozik, a kisebbik pirosba csak kettő. A piros és a kék pontok nem hatnak egymásra.

A hálózatok leírását persze sokszor a hálózat szerkezetének elemzése követi, és a hálózatelemző indexek közül több is van, amely érzékeny arra, ha több komponensből áll a hálózat. Ezek vagy torz eredményeket adnak, vagy ki sem lehet számolni őket. Ez főleg azokra az indexekre érvényes, amelyekben a pontpárok távolsága is szerepel. Különböző komponensekbe tartozó pontok távolsága definíció szerint végtelen. Emiatt a hálózatelemző vizsgálatok nagy része kimondva vagy kimondatlanul, csak az óriás komponensre vonatkozik. Ez már nyilván egy összefüggő, egyetlen komponensből álló hálózat (persze ez az eredetinek csak egy részgráfja), tehát mindent ki lehet rá számolni. Módszertanilag indokolt és persze kényelmesebb is így dolgozni. Ráadásul az óriás komponens olyan nagy (mondjuk a gráfpontok 95%-át tartalmazza), hogy a kutató csak zajnak tekint a kisebb komponensek törlésével elvesztett pontokat – és infomációt. A kisebb komponensek egyébként is sok esetben

egyetlen (tehát izolált) pontot tartalmaznak, már-már filozófiai kérdés, hogy ezek egyáltalán a hálózat részének tekinthetők-e. A pontok és a közöttük lévő relációk definiálása magának a hálózatnak a definiálása, és bizony könnyen előállhat olyan helyzet, amikor a hálózatunkban valakinek nincs szomszédja. Ha például egy osztályban valakinek nincs egyetlen barátja sem, akkor az osztály társas kapcsolathálózatában (social network) izolált pont lesz. A névsorban szerepel, de a közösségnek olyan értelemben nem tagja, hogy bármit csinál, az semmilyen hatással sincs a többiekre (ha úgy gondoljuk, hogy mégis van, akkor át kell gondolni, hogyan definiáltuk a hálózati kapcsolatokat, esetleg újra lehet

ti el, hogy a szerkezetileg (topológiailag) izolált pont valóban izolált pont-e funkcionális értelemben is (tehát valóban nincs-e semmilyen hatással a többiekre). A fő probléma az, ha megfeledekezünk a hálózatok dinamikájáról. Az esetleg jogosan törölt és nagyvonalúan elfelejtett gráfpont ugyanis később még akár kulcsszereplővé is válhatna, de akkor talán már módszertani okoknál fogva hiányzik az adatbázisunkból, és elvesztettük, nem is látjuk a lényegét. A következő példák talán érdekesek és egyúttal óvatosságra intenek. Menjünk alulról felfelé.

A fehérjék interakciós hálózatai igazán nagyok, szinte mindig látunk izolált pontokat és kicsi komponenseket. Ráadásul,



1. ábra. A bal oldalon egy két komponensből álló gráfot látunk, négy pont a kék, kettő a piros komponensbe tartozik. A piros és a kék pontok között nincs kapcsolat. A jobb oldalon egy, a Sziklás-hegységben élő mormotacsapat kapcsolathálózatát látjuk (sárgahasú mormota, *Marmotta flaviventris*). Itt két egyed izolált, a többiek benne vannak az óriás komponensben

definiálni a relációkat, de az új hálózatban talán már nem lesz izolált pont az ominózus gyerek). De ilyen egy mérgező növény is egy ökoszisztéma táplálkozási hálózatában: semmit sem esz meg és azt sem eszi meg semmi (az anyagforgalomban persze részt vesz, itt megint csak a hálózati kapcsolatok definiálása dönti el, össze kell-e kötni mérgező növényünket mondjuk a talaj szervesanyag-tartalmával, vagy nincs is olyan gráfpont).

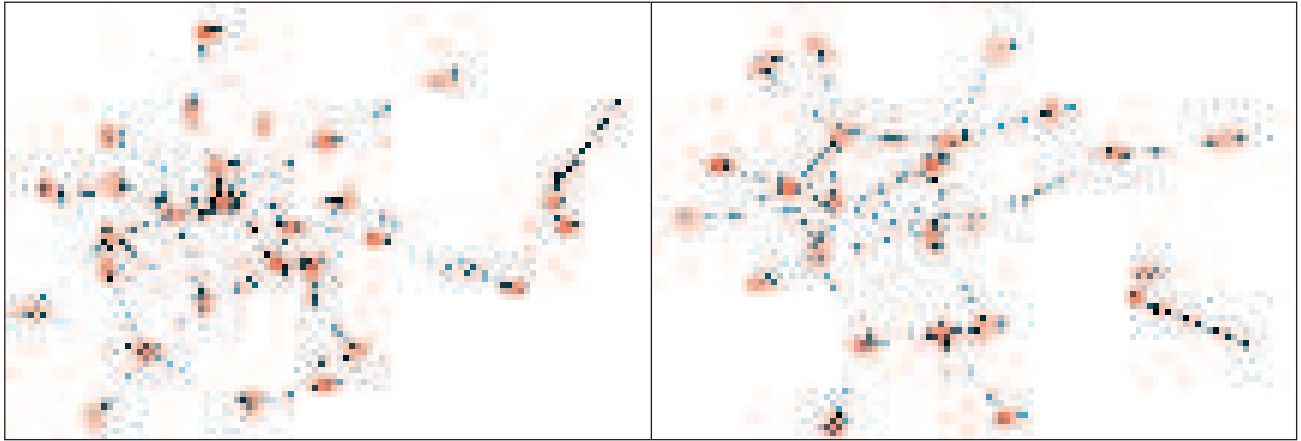
A kérdés egyszerű. Mennyire csapjuk be magunkat akkor, amikor csak az óriás komponensekkel foglalkozunk? Erre a kérdésre nyilván nincs általános válasz, minden esetben a hálózat definiálása dön-

ti el, hogy a szerkezetileg (topológiailag) izolált pont valóban izolált pont-e funkcionális értelemben is (tehát valóban nincs-e semmilyen hatással a többiekre). A fő probléma az, ha megfeledekezünk a hálózatok dinamikájáról. Az esetleg jogosan törölt és nagyvonalúan elfelejtett gráfpont ugyanis később még akár kulcsszereplővé is válhatna, de akkor talán már módszertani okoknál fogva hiányzik az adatbázisunkból, és elvesztettük, nem is látjuk a lényegét. A következő példák talán érdekesek és egyúttal óvatosságra intenek. Menjünk alulról felfelé.

A fehérjék interakciós hálózatai igazán nagyok, szinte mindig látunk izolált pontokat és kicsi komponenseket. Ráadásul, pont a méret miatt, a kutató ritkábban engedheti meg magának, hogy gondolkodjon (pardon: manuálisan gondozza a hálózati adatbázist), és nagy a csábítás arra, hogy egy gombnyomással töröljük a kisebb komponenseket. Egy 2777 fehérjemolekulát tartalmazó hálózat megalkotása során például az óriás komponensből kimaradt a P52895 kódú fehérje. Ilyenkor az

ember elgondolkozik azon, mi történik, ha megfeledekezünk róla, és a többi „kivülállóval” együtt töröljük az adatbázisból. Ez a fehérje, a dihidrodiol-dehidrogenáz-2 az AKR1C2 gén terméke, amely részt vesz a szteroid hormonok átalakításában, szerepe van az ivari fejlődésben, az emésztéssel és az elhízással is kapcsolatos. Persze minden fehérjének megvan a maga feladata, de ha az ember a fehérjehálózat elemzését éppen azért végzi, hogy ezekről a folyamatokról még többet tudjon meg, akkor hiba lehet a P52895 automatikus törlése.

Az emberek társas kapcsolathálózatában már említettük a magányos gyerek kicsit naív példáját. Itt azért érdemes vi-



2. ábra. A *Ropalidia marginata* papírdarázs két kolóniájának kapcsolathálózata. Az elsőben a királynő, a másodikban a királynői cím várományosa (post-queen) szorult ki az óriás komponensből

gyázni, mert például egy osztályközösséget ugyan egyszerű definiálni (definiálja azt az osztálynaplót), de egy-egy gyerekek nyilván lehetnek más barátai is, például akikkel együtt focizik. Az izolált gyerekek törlése az osztály csoportdinamikájának szempontjából talán jogos, de ha ezer más barátja van, akkor mégis kár róla elfeledkezni: valószínűleg nagy esély van arra, hogy egyszer az osztály közösségébe is belépjen, akár mint kulcsszereplő. Azt azért jegyezzük meg, hogy az óriás komponensek és izolált pontok problematikája általában sokkal nagyobb hálózatokban jelentkezik, az osztály példája inkább csak illusztráció.

Állatoknál talán még izgalmasabb a helyzet. A *Ropalidia marginata* trópusi papírdarászfaj egyedeinek kapcsolathálózatáról sokat tudunk. Normális, érett kolóniákban (kb. 10–50 egyedről beszélünk) a királynő általában nem tartózkodik az óriás komponensben, tehát ha töröljük az izolált pontokat, az felér egy felségstílussal. Amikor viszont új királynő kerül a kolónia élére, az az óriás komponens kellős közepén csücsül és szinte minden más egyeddel közvetlen kölcsönhatásban áll. Később, amikor már mindenki végzi a dolgát, egyre passzívabbá válik, és általában el is hagyja az óriás komponenset. A 2. ábra két érdekes darázshálózatot mutat be. Mindkettőben egy óriás és egy törpe komponens látunk. Az első hálózatban a törpe komponens egyik tagja a királynő (-L kóddal), a másodikban pedig a királynői cím várományosa, a következő királynő (-S kóddal).

Az 1. ábra jobb oldalán egy mormotacsapat kapcsolathálózatát látjuk: az amerikai kutatók adatai alapján két egyed izolált pont a hálózatban, azaz nem állnak senkivel sem kapcsolatban. A többi egyed az összefüggő, óriás komponensbe tartozik (már amennyire a mormoták hálózata óriási lehet egyáltalán: látjuk, körülbelül hú-

szan vannak). Mindketten idősebb nőstények, és ha az ábrázolt kapcsolatok szempontjából nem is fontosak, más kölcsönhatások fenntartásában még lehet szerepük. Megint a dinamika, az időbeli változatoság dönti el, mekkora hiba lehet megfedkezni róluk.

A társas kapcsolathálózatokhoz hasonlóan (gyerekek, darazsak), a táplálék-hálózatok sem tartalmaznak túlságosan sok elemet (gráfpontot). Ilyen ökológiai hálózatokból elég sokat ismerünk már, de nagyon ritkán találunk olyat, ami több komponensből állna. A mérgező növény példája is kicsit erőltetett, és ráadásul csak akkor áll elő, ha valóban eltekintünk az élettelen elemektől (pl. detritusz és anyagforgalom). Ha viszont figyelembe vesszük az időbeli változásokat, érdekes dolgokkal szembesülhetünk. Az évszakos dinamika (szezonális) eredményezheti, hogy például költöző madarak, téli álmot alvó állatok vagy éppen változatos egyedfejlődésű rovarok is megjelenjenek és eltűnjenek a hálózatból. Az éves adatok összevonásából álló hálózat unalmasnak tűnik az ilyen, kisebb idő-ablakokat bemutató hálózatok mellett. Ezekben ugyanis elvben előállhat sok érdekes izolált pont, például egy gazda nélküli parazita vagy egy préda nélküli predátor.

Az ökológiai rendszerek térbeli szerkezetét részben tájszerkezeti hálózatokkal (landscape graph) lehet bemutatni. Itt a gráfpontok élőhelyfoltokat, a gráf élei pedig közöttük kialakuló ökológiai folyosókat (tehát átjárási lehetőségeket) mutatnak be. Egy élőhelyfolt lehet izolált, tehát onnan nem juthatunk el semelyik másikba sem, legalábbis egy egyszerű, statikus megközelítés szerint. Azonban a térbeli és időbeli skálázás itt is segít abban, hogy észrevegyük, mit nem szabad elfelejtenünk. Egy izolált folt nagyon könnyen kapcsolatba kerülhet a többiekkel sokféle tájszerkeze-

ti változás hatására. Elég kivágni pár fát, elég egy árvíznek beköszöntenie, elég, ha megtanulnak valamilyen újfajta viselkedést a vizsgált állatok, és az izolált folt máris a rendszer része lesz, most már funkcionális értelemben is. A tájhálózatok nemegyszer igen nagyok, emiatt itt is csábító lehet, hogy csak az óriás komponensre koncentráljunk. Ezek a térbeli gráfok azonban ritkábban tartalmaznak egyetlen hatalmas és sok kicsi komponenset, sokkal gyakrabban több, hasonló méretű komponensből állnak. A problémát tehát itt a legnehezebb megkerülni, nem véletlen, hogy éppen a tájökölógusok foglalkoztak a legtöbbet a nem-összefüggő hálózatok elemzésének nehezségeivel. Például a távolság-mátrixot helyettesíthetjük a reciprok távolságértékek mátrixával: a különböző komponensekbe tartozó, tehát egymástól végtelen távolságra lévő m és n pontok reciprok távolságértéke 1, és a transzformált értékekkel már lehet dolgozni.

Az utolsó példa jól mutatja, hogy az egy-egy fajta hálózat elemzésére szakosodott specialisták sokszor nem tudnak egymás problémáiról – és egymás megoldásairól. Egy molekuláris rendszerbiológus, ha nem összefüggő a fehérje-fehérje kapcsolathálózat, nagyvonalúan törli a kisebb komponenseket és legfeljebb csak rosszul alszik amiatt, mit is törölt pontosan ki a hálózatból. A tájökölógus ezt nem teheti meg, ezért (még több álmatlan éjszaka után) kidolgozza a megoldást, ami viszont sajnos aztán nem feltétlenül terjed el a szűkebb szakterületen kívül.

A lényeg, hogy ha egy hálózat egyetlen komponensből áll, azonnal éljünk a gyanúval, hogy abból a hálózatból valamit már kitöröltek, és nézzünk utána, hogy az elfelejtett törpe komponens nem válhat-e óriási jelentőségűvé más körülmények között. Ebben az esetben ugyanis vissza kell tenni a hálózatba. ☛