

KOMPLEX RENDSZEREK MODELLEZÉSE FEJLŐDŐ HÁLÓZATOKKAL

A doktori értekezés tézisei

CSÁRDI GÁBOR

MTA KFKI Részecske és Magfizikai Kutatóintézet,
Biofizika osztály

Témavezető: Dr. Érdi Péter

Eötvös Loránd Tudományegyetem, Informatikai Kar
Informatika Doktori Iskola
Információs rendszerek doktori program

A Doktori Iskola vezetője: Dr. Demetrovics János
Programvezető: Dr. Benczúr András

A disszertációról

Dolgozatom kettő részből áll. Az első részben egy fejlődő gráfmodellt vizsgáltam meg, amelyben a csomópontok egyedi, időben állandó tulajdonságokkal bírnak. Célom a tulajdonságok hálózat szerkezetére gyakorolt hatásának feltérképezése volt, a tulajdonságok nélküli modellt használva viszonyítási pontként. A második részben fejlődő hálózatok inverz problémájára adok megoldást, vagyis egy modellkeret valós adatokra való illesztésére módszereket.

Mindkét rész a gráfelmélet módszereit használja. A gráfelméleti modellezés azt jelenti hogy a vizsgált rendszereket bináris relációk segítségével írjuk le. Ilyen relációkat gyakran találunk a természet és a társadalom vizsgálatakor, innen ered a gráfmodell univerzalitása.

Tulajdonság alapú hálózatok

Bevezetés

Számos gráfelméleti modellt definiáltak már, általában a valódi hálózatokban talált meglepő struktúrák megmagyarázására, mint például a kis-világ vagy a skálamentes tulajdonság. Ezek a modellek általában nem különböztetik meg az egyes csomópontokat, minden egyes, a hálózat kialakításában résztvevő ágens ugyanazzal a belső programmal rendelkezik. Tanulmányom azt vizsgálja hogy milyen hatással van a hálózat struktúrájára ha ezek az ágensok különbözőek.

Célkitűzések

Konkrétabban, a tulajdonság alapú hálózat foksám eloszlása és komponens méret eloszlása érdekelt. A foksám eloszlás talán a legalapvetőbb strukturális jellemzője egy hálózatnak, több tanulmány is megmutatta hogy jelentős hatással van a hálózaton lejátszódó dinamikai folyamatokra. Sok hálózatra jellemző hogy egy nagy összefüggő, ún. óriáskomponens birtokolja a csomópontok jelentős részét. Célom az volt hogy megállapítsam milyen hatással vannak a csomópontok belső tulajdonságai az óriáskomponens kialakulására.

Módszerek

Definiáltam egy növekvő hálózatmodellt, amelyben a csomópontok a lehetséges kapcsolataikat ugyan véletlenszerűen választják ki, de a kapcsolatok realizációja (azaz az élek behúzása) a résztvevő csomópontok tulajdonságainak függvénye. Minden ilyen tulajdonság alapú modellhez készíthető egy ún. redukált modell, amelyben az átlagos kapcsolódási valószínűség ugyanakkora mint a teljes modellben, de nem különböztetjük meg az egyes csomópontokat.

Numerikus szimulációkkal megmutattam hogy a foksámeloszlás időben stacionárius, mindkét modellben. Analitikusan kiszámítottam a foksám eloszlást, mindkét modellben. Iteratív egyenletek megoldásával és generátorfüggvények segítségével, numerikus közelítéseket is felhasználva kiszámítottuk a komponens méret eloszlást a redukált modellben. Numerikus szimulációk segítségével ugyanezt megtettem a teljes modellben.

Eredmények

T 1. *Megmutattam hogy a hálózat foksám eloszlása exponenciális, a csomópontok tulajdonságaitól függetlenül. Tehát ez az eloszlás modellkeret következménye, mivel az véletlen választásra épül.*

T 2. *Fázisátalakulást találtunk mind a teljes, mind a redukált modellben. A tulajdonság alapú modellben a fázisátalakulás küszöbe mindig alacsonyabb mint a redukált modellben.*

Konklúzió

Megmutattam hogy egy általam definiált, sztochasztikus fejlődő hálózatmodellben a hálózat kvalitatív tulajdonságait a csomópontok különböző tulajdonságainak figyelembe vétele nem változtatja meg, de hatással van a hálózat finom szerkezetére: csökkenti az óriáskomponens kialakulásának küszöbét. Emiatt a tulajdonságoknak fontos szerepük lehet a hálózaton zajló dinamikai folyamatokra, mint például egy betegség vagy trend terjedése, vagy hogy melyek a hatékony keresési technikák a hálózatban.

Fejlődő hálózatok visszafejtése

Bevezetés

Egy hálózat fejlődése azt jelenti hogy csomópontokat és éleket adunk hozzá és/vagy törölünk ki a hálózatból. A fejlődő hálózatok inverz problémája azt jelenti hogy egy adott hálózat lehető legjobb dinamikai leírását keressük, a (matematikai) leírás paramétereit illesztve az evolúciós adatokra. Másként fogalmazva, a feladat a hálózat strukturális változásainak megjósolása.

Célkitűzések

Céлом egy olyan keretrendszer létrehozása volt, amelyben az inverz probléma a matematika nyelvén megfogalmazható, és a megfelelő módszerekkel megoldható. A keretrendszernek lehetővé kell tennie egy rendszer alternatív leírásainak kvantitatív összevetését. Valamint, a modellkeretet egy adott hálózatra illesztő algoritmusoknak elég hatékonyaknak kell lenniük nagyon nagy hálózatokon is.

Módszerek

Definiáltam egy diszkrét idejű, sztochasztikus keretrendszert, fejlődő hálózatok leírására. Ebben a keretben a strukturális változásokat magfüggvények (ill. a legegyszerűbb esetben csak egyetlen magfüggvény) írják le. Ezek valós értékű függvények, amelyek a csomópontok strukturális vagy nem strukturális tulajdonságaitól függenek. Definiáltam egy magfüggvény „jóóságát”,

ez a mennyiség lehetővé teszi egy rendszer viselkedését leíró modellek összevetését.

Az inverz probléma megoldására kettő módszert fejlesztettem ki, az ún. frekventista és a maximum likelihood módszereket. Mindkét módszer alkalmazhatóságára elégséges feltételeket adtam. A módszereket egy nyílt forráskódú függvénykönyvtárban implementáltam. A kifejlesztett módszertan segítségével több hálózatot vizsgáltam. Az Amerikai Egyesült Államok szabadalmi hálózatának vizsgálatakor megmutattam hogy a jogrendszerben a kilencvenes évek elején bevezetett változások mérhető hatással voltak a hálózat dinamikájára.

Eredmények

T 3. *Definiáltam az ún. magfüggvény-alapú modellkeretet, ez lehetővé teszi egy fejlődő hálózat alternatív leírásainak összevetését.*

T 4. *Algoritmusokat fejlesztettem ki a keretrendszer paramétereinek, vagyis a magfüggvényeknek egy adott hálózatra való illesztéséhez. Elégséges feltételeket adtam a kifejlesztett algoritmusok használhatóságára, megmutattam megoldásaik unicitását.*

T 5. *Megmutattam hogy a lineáris elsőbbségi kapcsolódás (preferential attachment) a legjobb be-fok alapú modell több citációs és kollaborációs hálózat leírására.*

T 6. *Megmutattam hogy az Amerikai Egyesült Államok szabadalmi hálózatának dinamikája – a magfüggvény alapú modellkeretben – jelentősen megváltozott a kilencvenes évek elején. Az elsőbbségi kapcsolódás kitevőjének csökkenő trendje megállt és azóta egy növekvő trend van jelen.*

T 7. *Megmutattam hogy skálamentes hálózat esetén a legjobb (a magfüggvény jósága szerint legjobb) be-fok alapú modell mindig a lineáris elsőbbségi kapcsolódás. Hasonlóan, a legjobb be-fok alapú kernel függvény exponenciális fokszám eloszlású hálózat esetén mindig az uniform véletlen kapcsolódás.*

Konklúzió

A kifejlesztett módszertan alkalmas különféle fejlődő rendszerek dinamikájának kvantitatív leírására, a csomópontok strukturális és/vagy nem strukturális tulajdonságai alapján. Azt gondolom hogy a bemutatott módszerek hiánypótlóak és fontos szerepet töltenek be a komplex hálózatok elemzésének területén.

Publikációs lista

A tézisek alapjául szolgáló közlemények

- 2007 **Csárdi G**, Strandburg KJ, Zalányi L, Tobochnik J, Érdi P. Modeling innovation by a kinetic description of the patent system. *Physica A*, 374, 783–793.
- 2006 **Csárdi G**. Dynamics of citation networks. *International Conference on Artificial Neural Networks*, Lecture Notes in Computer Science, 698–709.
- 2006 Strandburg KJ, **Csárdi G**, Tobochnik J, Érdi P, Zalányi L. Law and the science of networks: an overview and an application to the „patent explosion.”, *Berkeley Technology Law Journal*, 21:4, 1293–1362.
- 2006 **Csárdi G**, Strandburg KJ, Zalányi L, Tobochnik J, Érdi P. Estimating the dynamics of kernel-based evolving networks. *International Conference on Complex Systems*.
- 2003 Zalányi L, **Csárdi G**, Kiss T, Lengyel M, Warner R, Tobochnik J, Érdi P. Properties of a random attachment growing network. *Physical Review E* 68, 066104.

A dolgozat témájában megjelent egyéb közlemények

- 2006 **Csárdi G**, Nepusz T. The igraph software package for complex network research. *International Conference on Complex Systems*.
- 2004 **Csárdi G**, Young M, Sager J, Hága P. Self-repairing peer-to-peer networks. Kutatási jelentés, Santa Fe Komplex Rendszerek Nyári Iskola.
- 2004 Young M, Sager J, **Csárdi G**, Hága P. An agent-based algorithm for detecting community structure in networks. Kutatási jelentés, Santa Fe Komplex Rendszerek Nyári Iskola.