



### Bírálat Karsai Márton MTA doktori értekezéséről

A doktori mű célkitűzése az emberi viselkedés, mint dinamikai rendszer megfigyelése, modellezése és megértése a hálózattudomány, a komplex rendszerek fizikája, a számítás- és az adattudomány eszköztárának segítségével.

Komplex rendszerek területén a vizsgálatok fő célja annak megértése, hogy hogyan alakulnak ki makroszkopikusan megfigyelhető kollektív jelenségek nagyszámú ágens vagy egyén kölcsönhatásának eredményeként, úgy hogy azok egyéni tulajdonságai nem elégségesek a jelenségek megmagyarázására. Ilyen jelenségeket a statisztikus fizika tradicionálisan fázisátmenetnek nevez, vagy kritikus jelenségként azonosít. Az egyéni és kollektív viselkedési minták pontos megfigyelése a társadalmi rendszerekben a közelmúltban vált lehetővé, amikor nagy mennyiségű, pontos adat összegyűjtése társadalmi méretekben elvégezhető lett nagy teljesítményű számítási erőforrásokat igénybe vevő “socialcomputing” rendszerek vagy a statisztikai alapokra épülő modern gépi tanulási módszerek segítségével.

A disszertáció három fejezetben mutatja be a szerző eredményeit az emberi viselkedés dinamikai modellezése (computational human dynamics) területén, nyolc tézispontba csoportosítva.

A bevezetést követő második fejezet az úgynevezett „bursty” jelenség vizsgálatával foglalkozik, és két tézispontot tartalmaz. Ez a kifejezés egyéni és csoportos viselkedési minták azon jellegzetességét takarja, amikor az események sorozatában az egymást követő események nem hasonló időközönként jelennek meg, hanem rövid idő alatt sok esemény történik, és ezt követően a rendszer hosszú ideig inaktív marad. Ez a dinamikai viselkedés jellemezi többek között az egyéni kommunikációs folyamatokat, a neuronok aktivitását vagy a földrengések sorozatát. A matematikai leírásban az egymást követő események között eltelt időintervallumok hosszának eloszlása játszik meghatározó szerepet. Ez az eloszlás lassú lecsengésű hatványfüggvénnyel jellemezhető, ahol az események között feltételezett korrelációkat az ún. autokorrelációs függvény jellemez. Valós adatok elemzésével a jelölt megmutatta, hogy ez a leírás nem alkalmas a jelenségek megmagyarázására, és új megközelítést vezetett be. Ennek során a közeli eseményeket úgynevezett vonatokba csoportosította, és ezek hosszának eloszlásáról megmutatta, hogy hatványfüggvénnyel jellemezhetők. A megfigyelt jelenségek magyarázatára bevezetett egy memória vezérelt kétállapotú esemény-generáló modellt, amely lényegében azt adja meg, hogy adott hosszúságú vonatot még egy esemény követ, vagy a vonat véget ér. Ennek segítségével sikerült nem-markovi korrelációkhoz jutnia, amelyek így jól reprodukálták az empirikus eredményeket.

A társadalmi hálózatok szerkezetében meghatározó az időbeli változás, ezért ezeket nem statikus, hanem temporális hálózatként szokás modellezni. A disszertáció harmadik fejezete ezek vizsgálatának van szentelve, és három tézispontot tartalmaz. Az itt ismertetett első tézispont [T4] egy mintegy 6 millió felhasználó 600 millió mobiltelefonhívásának elemzésével készült temporális hálózati struktúra elemzésén alapul. Az adatok alapján a jelölt létrehozott egy aktivitás vezérelt temporális hálózati modellt az adatokból szintetizált memóriafolyamat értelmezésével, ami azt határozza meg, hogy egy adott pillanatban aktív csomópont mekkora valószínűséggel fog kapcsolódni egy véletlenül kiválasztott másik csomóponthoz. Az így



Eötvös Loránd Tudományegyetem  
Természettudományi Kar  
Matematikai Intézet  
Alkalmazott Analízis és Számításmatematikai Tanszék

---

definiált temporális hálózati modell sok szempontból reprodukálni tudta a valós megfigyeléseket. A következő tézispont [T5] a temporális hálózatok gyökeresen új matematikai leírását adja. Hagyományosan a temporális hálózat éleihez időpontok is tartoznak, ilyen módon egy él nemcsak egy csúcspár, hanem a két csúcsból és egy időpontból álló hármas. Az új, magasabb szintű reprezentáció egy olyan úgynevezett eseménygráf, amelyben az ilyen hármasok alkotják a csúcsokat, és közöttük akkor van él, ha a hármasban van közös csúc. Ráadásul ehhez az élhez súlyt is rendelünk aszerint, hogy a két hármasban szereplő idő komponensek mennyire térnek el egymástól. Meghatározó új eredmény, hogy a temporális gráfon zajló folyamatot jól lehet jellemezni az eseménygráfon perkolációs módszerekkel.

A disszertáció [T6], [T7] és [T8] tézise a kollektív hálózati jelenségeket tárgyaló negyedik fejezetben kerül bemutatásra. A Skype több mint 300 millió felhasználójának kommunikációs hálózatán a termékadaptációs dinamikája képezte a vizsgálatok alapját. A felhasználók hálózatán terjedő termékvásárlás, mint egy társadalmi terjedési folyamat írható le, és a modellezés során a nem várt jelenségek is magyarázhatók. Az adatokat elemezve igazolható, hogy az egyén szempontjából az adaptálást kényszerítő külső nyomás lineárisan növekszik a befolyásoló ismerősök számával. Ezeket a megfigyeléseket egy olyan matematikai modellbe lehet építeni, amelyben hálózatba kötött entitások egy terjedő állapot adaptálása vagy elhagyása érdekében befolyásolták egymást. A modell heterogén átlagtér egyenletekkel való közelítése analitikusan vizsgálható. A hagyományos terjedési modellek alapján a társadalmi befolyás által gerjesztett terjedési folyamatok lavinaszerű fejlődése várható, azonban a megfigyelések alapján számos társadalmi terjedési folyamat sokkal lassabb dinamikát követ. Ennek magyarázatához ismét a Skype online társadalmi hálózatát használta a szerző. Ebben a hálózatban a befolyásolási küszöbérték eloszlását elsőként mérte meg sikeresen, és a korábbi elméleti feltevésekkel ellentétben egy heterogén, log-normális eloszlást talált. Ennek segítségével sikerült igazolni, hogy nem azok a felhasználók indukálták a globális terjedési folyamatot, akiknek csak egy befolyásoló szomszédra volt szükségük, hanem azok, akik több szomszéd együttes hatásának következményeként vásárolták meg a terméket. Egy új terjedési modellt bevezetve a termék elterjedésének valóságghűbb szimulációját sikerült megmutatni. Ebben a csomópontok, azon túl, hogy üresek vagy betöltöttek lehettek, a folyamat kezdetétől blokkolt állapotot is felvehettek, vagyis sosem “fertőződhetek” meg a terjedő állapottal. Ez a blokkolt populáció ritkító hatásként lépett fel a folyamat során, és képes volt jelentősen lelassítani a terjedési folyamatot. A jelenség analitikusan és numerikusai is vizsgálható közelítő mester- egyenletek stabilitásanalízisével.

Az értekezés záró fejezetében a szerző nemcsak összefoglalja a tárgyalt eredményeket, hanem átfogó képet ad a bemutatott kutatási terület mai helyzetéről, és a körvonalazódó új kutatási irányokról.

Az értekezéssel kapcsolatos kérdéseim a következők.

1. Az értekezés jelentős értéke, hogy nagy mennyiségű adat elemzése nyomán hozta létre a szerző elméleti megközelítéssel a modelleket. Ilyen mennyiségű adat elemzése óriási munkát jelent. Mennyire támasztják alá a modelleket a mások által elemzett adatok, az irodalomban található-e erre vonatkozó eredmény? Az adatok előkészítése befolyásolhatja-e a modell létrehozását? Például az 1.3.2 szakasz DS2



Eötvös Loránd Tudományegyetem  
Természettudományi Kar  
Matematikai Intézet  
Alkalmazott Analízis és Számításmatematikai Tanszék

---

adatbázisában a 110 millió felhasználó 8 milliárd hívásából számos törlésre került, hogy csak az emberi interakciók kerüljenek be a vizsgálatba, és így egy jelentősen kisebb adatbázis jött létre 80 millió felhasználóval és „csak” egymilliárd összeköttetéssel.

2. Milyen lehetőségeket lát a jelölt a most kialakulóban levő, szinte teljesen adatalapú, gépi tanulási módszerekre támaszkodó kutatásokban, amelyekben a sokunk által szeretett modellalkotásnak várhatóan csak periférikus szerep juthat?

A fent ismertetett eredmények azt mutatják, hogy a jelölt kutatásai során komoly technikai eszközkészlet segítségével erős és mély elméleti eredményeket ért el, és a gyakorlati alkalmazásokban is hasznosítható sikeres vizsgálatokat végzett. A doktori mű angol nyelven készült, alaposan és világosan mutatja be az elért eredményeket. A magyarul írt téziszfüzet az előírt szerkezetben tartalmazza a 8 tézisben megfogalmazott eredményeket, amelyeket új tudományos eredménynek ismernek el. A doktori munka mintegy 400 hivatkozást tartalmaz, jelezve a szerző alapos tájékozottságát kutatási területén.

A jelölt közleményei túlnyomórészt folyóirat cikkek formájában jelentek meg, és a Fizikai Tudományok Osztálya csoportosítása szerinti lefeljebb ötszerős kategóriába tartoznak. Így a mintegy 90 publikációra kapott 2200 hivatkozás sokszorosan felülmúlja a doktori minimumkövetelményeket. Publikációi igen rangos nemzetközi folyóiratokban jelentek meg. Több dolgozata neves külföldi kutatókkal közös munka, amely szintén a jelölt nemzetközi elismertségét támasztja alá, valamint nyilvánvalóan mutatja, hogy aktívan részt vesz a nemzetközi tudományos életben

Összefoglalva, az értekezés egy fontos, aktuális témában mutat be értékes új eredményeket, amelyek eléréséhez új módszerek kidolgozására, jelentős technikai nehézségek legyőzésére volt szükség. Karsai Márton a hálózatkutatás valamint a kollektív társadalmi viselkedés modellezésének és vizsgálatának nemzetközileg elismert képviselője. Számos új eredményt publikált a téma legkiválóbb folyóirataiban. Munkásságával lényegesen hozzájárult e terület további fejlődéséhez.

Az értekezésben bemutatott eredmények messzemenően megfelelnek az MTA doktori disszertációkkal szemben támasztott követelményeknek, ezért határozottan javaslom az értekezés nyilvános vitára bocsátását, és sikeres védelem esetén az MTA doktora fokozat odaítélését.

Budapest, 2023. szeptember 23.

Simon Péter  
tanszékvezető egyetemi tanár  
ELTE, TTK, Matematikai Intézet  
Alkalmazott Analízis és  
Számításmatematikai Tanszék