

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

SZOCIÁLIS HÁLÓZATOK GEOGRÁFIAI BEÁGYAZÓDÁSA

Szakedolgozat

Készítette: Fejes Ágota
Matematika BSc, Matematikai elemző szakirány

Témavezető: Lukács András
Számítógéptudományi tanszék



2014

Budapest

Tartalomjegyzék

1. Bevezetés	4
2. Szociális hálózatok és a kisvilág-tulajdonság	5
2.1. A hálózatok fogalma	5
2.2. Szociális hálózatok	6
2.3. A hat lépés elmélet	7
2.4. Kisvilág-tulajdonsággal rendelkező gráf modellek	8
2.4.1. Az Erdős-Rényi modell	9
2.4.2. A Watts-Strogatz modell	10
2.4.3. A Barabási-Albert modell	11
3. A Kleinberg-modell	13
3.1. Az $r = 2$ esetre vonatkozó tétel bizonyítása	16
3.2. Az $r < 2$ esetre vonatkozó tétel bizonyítása	19
3.3. Az $r > 2$ esetre vonatkozó tétel bizonyítása	23
4. Összegzés	26

Köszönetnyilvánítás

Hálásan köszönöm témavezetőmnek, Lukács Andrásnak az inspirációt a szakdolgozatom elkészítéséhez, konzultációk során nyújtott tanácsait, magyarázatait. Köszönöm családomnak, akik az évek során türelemmel és megértéssel támogattak tanulmányaimban.

1. fejezet

Bevezetés

A matematika számos területével alkalmam volt megismerkedni egyetemi tanulmányaim alatt, szakdolgozatom témája mégis egy, a tananyagban nem teljesen kifejtett problémakörnek szentelem. Bár több kurzuson is elhangzott a szociális hálózatok kisvilág tulajdonsága, azonban ennek részletezése és modelljei nem hangzottak el. E jelenség népszerűsítő megfogalmazása a hat lépés távolságának elmélete, mely szerint a Földön a legtöbb ember legfeljebb hatodszintű ismerőse egymásnak. A barátaink, ismerőseink, családtagjaink által alkotott kapcsolatrendszer mindannyiunk életében fontos és funkciókban gazdag szerepet tölt be, azonban a dolgozat nem a társadalomtudományok oldaláról kívánja vizsgálni az így alkotott közösségi struktúrát, hanem annak felépítését szeretném matematikai módszerekkel elemezni.

Dolgozatom első felében szociális vagy más néven ismeretségi hálózatokhoz szorosan kapcsolódó kisvilág-tulajdonsággal rendelkező és az ehhez szorosan kapcsolódó gráfmodelleket ismertetek. Az alapvető fogalmak, modellek tisztázása után a dolgozat második felében Kleinberg Milgram-kísérletet magyarázó modelljét és vonatkozó tételeit, illetve azok bizonyítását mutatom be.

2. fejezet

Szociális hálózatok és a kisvilág-tulajdonság

2.1. A hálózatok fogalma

A *hálózatok* fogalmat a mindennapi életben sokféleképpen használjuk. Gondolhatunk itt az emberi kéz alkotta városokat összekötő út*hálózatokra* és vasút*hálózatokra*, különböző gázok és cseppfolyós anyagok szállítására alkalmas csővezeték *hálózatokra*, a természetben előforduló víz*hálózatokra*, gyökér*hálózatokra*, az emberi testben lévő ideg*hálózatra* vagy akár az ér*hálózatra*. Leggyakrabban pedig az információk közlésére alkalmas technikai hálózatok fogalmával találkozhatunk hétköznapi beszélgetések alkalmával: rádió*hálózat*, telefon*hálózat* illetve az egészét világot felölelő számítógépek*hálózata*, az internet*hálózat*. Az utóbbi évtizedekben pedig egyre elterjedtebb kifejezés a szakdolgozatoknak is alapjául szolgáló társadalmi- vagy szociális hálózat.

Általános funkción alapuló definíciót adni a *hálózat* fogalmára nem egyszerű. Ezt láthatjuk a Magyar Nyelv Értelmező Szótárában a hálózat címszó jelentéseinek sorolásában:

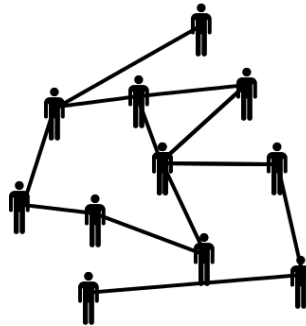
1. egymást átszelő egyenes v. görbe vonalak sűrű, szabályos szövedéke, rendszere.
2. nagy területet behálózó műszaki (közlekedési, villamossági, hírközlő, csatorna, stb.) létesítmények összefüggő rendszere.
3. (átv.) egymással összefüggésben levő, hasonló v. azonos intézmények egységesen

szervezett, kiterjedt rendszere.

Az előbb felsorolt hálózat fajták közös jellemzője azonban könnyen felismerhető, ezek mind egymással valamilyen összeköttetésben lévő különböző elemek egybefüggő rendszere. A matematika nyelvére lefordítva azonban egyértelműnek tűnik, hogy gráfokról beszélünk, amikor egy tetszőleges hálózatot próbálunk leírni. Az elemek a csúcsok, a köztük menő élek pedig az ezek közötti összeköttetés. Például legyenek a városok a csúcsok, az őket összekötő autóutak vagy vasúti vonalak pedig az élek, vagy a számítógépeket tekintjük a csúcsoknak és a köztük lévő kommunikációs összeköttetést az éleknek. Ilyen értelemben bármilyen hálózatot is tekintünk, egyértelmű, hogy a hálózatok struktúrája értékes információkat tartalmaz.

2.2. Szociális hálózatok

Amennyiben egy emberek alkotta csoportban az egyéneket tekintjük a pontoknak, a köztük lévő ismeretséget pedig az éleknek egy hálózatot kapunk. Ezt a kapcsolatrendszeri modellt nevezzük szociális vagy társadalmi hálózatnak.



2.1. ábra. Szociális hálózat ábrázolása, mint gráf

A szociális hálózat az előzőek szerint egy speciális gráf, ahol a csúcsok az egyének, a köztük menő élek pedig azt jelzik, hogy az adott emberek ismerik egymást. Egy ilyen leegyszerűsített modell azonban nem írja le hűen a valóságot, hiszen egy ismeretség tulajdonságát tekintve lehet egyirányú vagy kétirányú, időszakos vagy tartós, erős vagy gyenge.

Ezek alapján fel lehetne írni egy irányított, többszörös élekkel rendelkező súlyozott gráfot is, de jelen dolgozatban az irodalmat követve a súlyozatlan, illetve egyszerű gráfokra vezető modelleket vizsgálók.

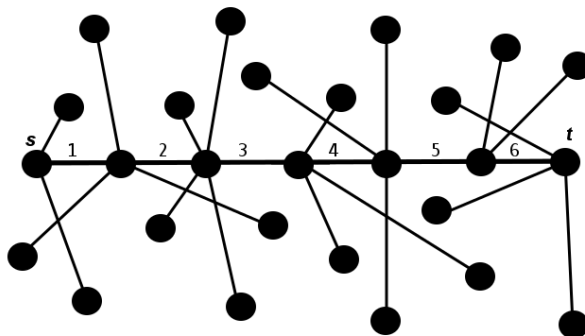
Az így felállított modellekben adódik a kérdés, hogy a hálózatban két, tetszőlegesen kiválasztott pont között hány lépésből áll a legrövidebb út. Egy konkrétan ismert hálózat esetében, például egy azonos gyárban dolgozó emberek esetében felmerülhet az az életszerű kérdés, hogy hány ismerősön keresztül juthat el az alacsony beosztású ember a vezérigazgatóhoz, amennyiben személyesen szeretne vele beszélni. Természetesen ebben az esetben egy konkrét számot fogunk kapni eredményül, azonban bonyolultabb a kérdés, ha az átlagos lépéstávolságot szeretnénk megtudni. A dolgozat szempontjából fontos, hogy a definiált szociális hálózatok rendelkeznek az úgynevezett kisvilág tulajdonsággal, azaz az átlagos távolság a csúcsok között azok számához képest kicsi. Ezt a jelenséget az angolul "six degrees of separation"-nek nevezett hat lépés elmélettel szokták bemutatni.

2.3. A hat lépés elmélet

Nem csupán a jelenkori tudósokban merült fel a kérdés, hogy egy szociális hálózat két tetszőlegesen választott csúcsa között hány, éleken át vezető lépés szükséges, hogy elérjünk egyikből a másikba. Éppen egy magyar író és költő, Karinthy Frigyes 1929-ben a *Láncszemek* című novellájában fogalmaz meg egy sejtést, miszerint a Földön két véletlenszerűen kiválasztott ember kapcsolatba kerülhet egymással az ismertségi hálózatokon keresztül maximum hat lépésben. Fontosnak tartom itt megjegyezni, hogy a Föld lakossága ma 6,6 milliárd, míg Karinthy idejében csupán 1,65 milliárd volt. Azonban arra utaló írás, miszerint ez a novella eljutott volna a későbbiekben az ezzel a témával foglalkozó szakértőkhöz, és ez alapján kezdtek volna foglalkozni a problémakörrel nincs. Így feltelezhetjük, hogy Karinthytól függetlenül jutottak hasonló megállapításra évtizedekkel később a szociális hálózatokat vizsgáló kutatók.

Évtizedekkel később, 1967-ben Stanley Milgram, amerikai szociálpszichológus elvégezte kisvilág kísérletét a Yale Egyetemen, amelyben Nebraska és Kansas államból kiválasztott embereknek kellett egy számukra ismeretlen, massachusettsi lakosnak továbbítani egy levelet úgy, hogy csak korlátozott mennyiségű információt tudtak a célszemélyről, például

a foglalkozását, de a pontos címét nem. Az üzenetet mindig egy általuk ismert személynek tudták csak tovább küldeni, aki szerintük kapcsolatban állhat a célszeméllyel. A kísérlet eredményeként Milgram azt a következtetést vonta le, hogy az Egyesült Államok szociális hálózata kisvilág tulajdonságú, hiszen a beérkezett levelekhez átlagosan csupán hat közvetítőre volt szükség, tehát az emberek közötti átlagos távolság meglepően kicsi a lakosság számához mérve. Azonban nem elhanyagolható az eredmények értelmezésekor, hogy a kísérletben résztvevők a társadalom középosztályához tartoznak, így az elszigetelten élőkre az általánosítás valószínűleg helytelen. Továbbá fontos annak figyelembevétele, hogy az elindított 296 levélből csupán 64 darab ért célba.



2.2. ábra. A hat lépés elmélet ábrázolása

Az elmélet neve ebből a kísérletből származtatható, habár Milgram maga sosem használta ezt, hanem később készített John Guarre egy színdarabot a kísérletről, melynek "A hat lépés távolság" címet adta.

2.4. Kisvilág-tulajdonsággal rendelkező gráf modellek

2.4.1. Definíció. *Egy gráf (vagy hálózat) kisvilág-tulajdonságú, ha a gráf (vagy hálózat) méretéhez képest a csúcsok közötti átlagos távolság a csúcsszám logaritmusára vagy annál kisebb nagyságrendű.*

Adódik a kérdés, hogy miként épül fel egy ilyen, kis világ tulajdonságú hálózat. Erre a

válaszra szeretnék a következőekben három különböző megközelítést adni, publikálásuk szerinti kronológiai sorrendben.

2.4.1. Az Erdős-Rényi modell

Elsőként Erdős Pál és Rényi Alfréd 1959-ben véletlen gráfok előállítására adott modellt, ekkor még nem kifejezetten a kisvilág tulajdonságot kutatva [6].

2.4.2. Definíció. *Legyen n pozitív egész, és $0 \leq p \leq 1$. Az Erdős-Rényi gráf csúcsai $1, 2, \dots, n$. Az (i, j) pontpár között p valószínűséggel van él. Ezt a gráfot $ER(n, p)$ -vel jelöljük.*

Az $1, 2, \dots, n$ csúcsú irányítatlan gráfok halmazán ez tulajdonképpen egy valószínűségi mező, ahol a gráfok az elemi események. Ez alapján a pontosan m éllel rendelkező gráf valószínűsége:

$$p^m(1-p)^{\binom{n}{2}-m}.$$

Az ő felfedezésünkkel kapcsolatban érdemesnek tartom megjegyezni a küszöbfüggvény fogalmát.

2.4.3. Definíció. *Legyen \mathcal{A} egy tetszőlegesen adott gráftulajdonság. Ekkor a $k(n) : \mathbb{Z}^+ \rightarrow \mathbb{R}^+$ függvény a \mathcal{A} küszöbfüggvénye, ha*

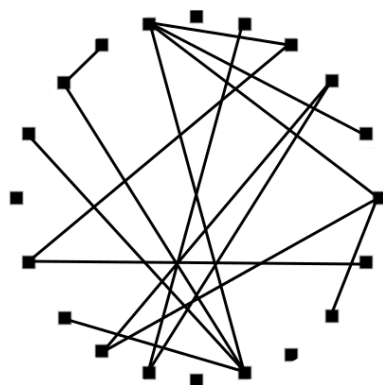
1. *bármely olyan $p(n)$ valószínűségek esetén, amelyekre $\lim_{n \rightarrow \infty} \frac{p(n)}{k(n)} = 0$ teljesül, hogy*

$$\lim_{n \rightarrow \infty} P(ER(n, p(n)) - \text{ben } \mathcal{A} \text{ igaz}) = 0$$

2. *bármely $q(n)$ valószínűségek esetén, amelyekre $\lim_{n \rightarrow \infty} \frac{q(n)}{k(n)} = \infty$ teljesül, hogy*

$$\lim_{n \rightarrow \infty} P(ER(n, q(n)) - \text{ben } \mathcal{A} \text{ igaz}) = 1.$$

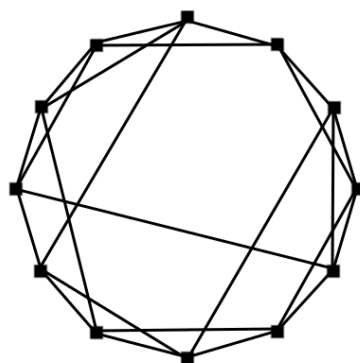
Ez azt jelenti, hogy létezik egy éles határ, amely alatti élsűrűség esetén szinte biztosan nem teljesül \mathcal{A} gráftulajdonság, míg felette szinte biztosan teljesül. Erdős és Rényi kutatásaikban bebizonyították, hogy egy gráf összefüggőségének küszöbfüggvénye $k(n) = \frac{\log n}{n}$. Ha egy Erdős-Rényi gráfra igaz, hogy $p(n) \geq \frac{\log n}{n}$, akkor bármely két csúc között van rövid út, azaz a gráf kisvilág tulajdonságú.



2.3. ábra. $p=0.1$ paraméterű Erdős-Rényi gráf

2.4.2. A Watts-Strogatz modell

A következő kisvilág-tulajdonságot mutató modellt Duncan J. Watts és Steven Strogatz 1998-ban publikált algoritmus generálja [5]. Elképzelésük szerint egy kisvilág felépítése úgy írható le, hogy kezdetben adott q érték mellett az összes, N db csúcs kapcsolatban áll az összes, tőle legfeljebb q távolságban elhelyezkedő csúccsal úgy, hogy a csúcsokat egy körvonal alapján helyezi el. Ahol $N \gg q \gg \ln(N) \gg 1$, így mindkét irányban $q/2$ szomszéddal rendelkezik minden pont. Majd az algoritmus az összes így létrehozott élen végigmegy és adott $0 < p < 1$ valószínűséggel lecserél egy q -nál távolabbi pontokat összekötőre.



2.4. ábra. $N=12$, $q=2$, $p=0.16$ paraméterű Watts-Strogatz modell alapján előállított kisvilág tulajdonságú gráf

Így az új élek biztosítják a kisvilág-tulajdonságot. A q érték növelésével a gráf egyre jobban hasonlít Erdős-Rényi gráf modelljére.

2.4.3. A Barabási-Albert modell

A harmadik elképzelés a kisvilág tulajdonságú gráfok felépítésére A Barabási-Albert modell, melyet 1999-ben publikáltak. Ez a modell arra ad rekurzív algoritmussal magyarázatot, hogy hogyan jöhet létre olyan gráf, amelyek fokszámeloszlása hatványeloszlást követ [7].

2.4.4. Definíció. *A $P(k)$ fokszámeloszlás azt a valószínűséget adja meg, amellyel a pontosan k éllel rendelkező csúcs előfordul egy gráfban.*

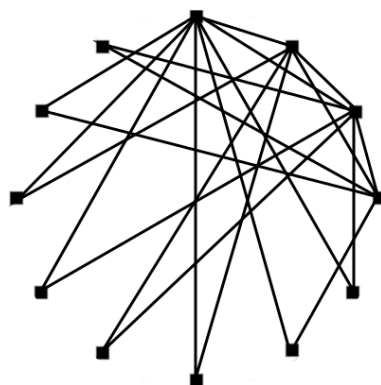
Jelöljük $BA(E, V)$ -vel a V csúcsú és EV élű Barabási-Albert gráfot. A rekurzív algoritmus $BA(E, V)$ felépítéséhez a következő:

1. A $BA(E, 1)$ gráfot egyetlen csúcs v_1 , és E hurokél alkotja.
2. A $BA(E, V)$ gráf egy új csúcs, v_n hozzáadásával keletkezik, $BA(E, V - 1)$ gráfból. Tegyük fel, hogy ez a $BA(E, V - 1)$ gráf már készen van, $V - 1$ db csúcscsal és $E(V - 1)$ éllel. Ekkor v_n csúcsot E db éllel kapcsoljuk a már kész $BA(E, V - 1)$ gráfhoz véletlenszerűen. Azaz a v_n -ből induló k -adik él a $BA(E, V - 1)$ gráf véletlenszerűen választott v_i -edik csúcsába vezet, ahol $k = 1, 2, \dots, V$. A v_i -edik csúcs választásának valószínűsége arányos a csúcs fokszámával, pontosabban ha a v_i fokszáma d_i , akkor annak a valószínűsége, hogy (v_n, v_i) él létrejön:

$$\frac{d_i}{d_1 + d_2 + \dots + d_n}.$$

Ennek a modellnek két fontos tulajdonsága a következő:

1. *Növekedés:* Egy folyamatosan bővülő hálózatot ír le.
2. *Preferenciális kapcsolódás,* azaz hogy a nagyobb fokszámú csúcsokhoz nagyobb valószínűséggel kapcsolódik új él.



2.5. ábra. $BA(2,12)$ gráf ábrázolása a rekurzív algoritmus szerint, a kezdeti hurokélek nélkül

A növekedés nem igaz az előzőekben bemutatott Erdős-Rényi és Watts-Strogatz modellre, hiszen azokban már a kiinduló állapotban adott volt az összes csúcs. A preferenciális kapcsolódás tulajdonságból adódik a hasonlóság a szociális hálózatokhoz, ahol szintén megfigyelhető, hogy van néhány, az átlagnál sokkal több ismerőssel rendelkező személy.

Ez az algoritmus skálafüggetlen hálózatot eredményez.

2.4.5. Definíció. *Egy hálózatot skálafüggetlennek nevezünk, ha a hálózat fokszámeloszlása hatványeloszlást követ.*

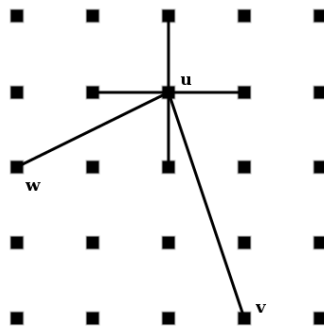
Barabási Albert László kutatásában a világháló felépítését is tanulmányozta és azt vette észre, hogy van néhány, kiugróan magas fokszámmal rendelkező csúcs. Erre az eddigiekben felsorolt modellek nem adnak magyarázatot. Az Erdős-Rényi modellben rendkívül ritka lenne az ilyen pontok előfordulása, a Watts-Strogatz modell felépítése nem engedélyezi az ilyen csomópontok létrejöttét.

Kleinberg tovább formálta a kisvilág-tulajdonságú gráfok felépítését, amikor a Milgram-kísérlet eredményét továbbgondolva nem csupán a rövid utak létezését, hanem egy hatékony és egyszerű algoritmus létét is bizonyította, amellyel az üzenet rövid úton továbbítható. Jon Kleinberg által alkotott kisvilág-tulajdonságú hálózat felépítése közelít a valóságos szociális hálózat felépítéséhez.

3. fejezet

A Kleinberg-modell

Kleinberg egy $n \times n$ -es rácshálót tekintett, melynek elemeit a következőképpen jelölte: $\{(i, j) : i \in \{1, 2, \dots, n\}, j \in \{1, 2, \dots, n\}\}$, ahol két pont közötti távolság: $d((i, j), (k, l)) = |i - k| + |j - l|$. Adott továbbá egy $p \geq 1$ konstans, u és a rácsháló összes olyan eleme között megy él, melyek között a távolság legfeljebb p . Ezeket a rácshálóban közeli kapcsolatoknak nevezi. Továbbá adott egy $q \geq 0$ konstans, mely azt határozza meg, hogy egy csúcsnak legfeljebb hány, úgynevezett távoli kapcsolata lehet, mely hosszabb mint p . Az így kapott gráf irányított, tehát u távoli ismerősének a távoli ismerőse nem egyértelműen u lesz.



3.1. ábra. $p=1$, $q=2$ paraméter mellett u közeli és távoli kapcsolatai

Annak a valószínűsége, hogy két csúcs, u és v között van él, arányos $[d(u, v)]^{-r}$ -el, ahol $r \geq 0$ adott konstans. Ezt az értéket ha normáljuk $\sum_{v \neq u} [d(u, v)]^{-r}$ -el, akkor ez a távoli kapcsolatok előfordulásának eloszlását mutatja meg.

3.0.6. Definíció. *Nevezzük a továbbiakban inverz hatvány eloszlásnak a távoli kapcsolatok előfordulásának eloszlását*

$$\frac{[d(u, v)]^{-r}}{\sum_{v \neq u} [d(u, v)]^{-r}}.$$

Kleinberg által definiált gráfban érvényes a kisvilág-tulajdonság, azaz várhatóan két tetszőleges csúcs között van rövid út, és ezt a rövid utat egy decentralizált algoritmus meg is találja. Tehát pusztán lokális információkra van szüksége az algoritmusnak ahhoz, hogy az inputból, a kiinduló és célszemély kilétéből megmondja a köztük lévő utat. Tehát az algoritmus bemenete s és t , a cél pedig eljutni s -ből t -be a lehető legkevesebb lépésben. Az algoritmus minden lépésben továbbítja az üzenetet u -tól egy távoli, vagy rövid kapcsolaton keresztül v -be, csupán lokális információk alapján. Ezek a lokális információk minden lépésben a következők:

1. A lokális kapcsolatokat,
2. A rácshálóban t elhelyezkedése,
3. A csomópontok és a távoli kapcsolatok helye a rácshálóban, amelyek kapcsolatba kerültek már az üzenettel

Tehát u -nak nincs tudomása azokról a távoli kapcsolatokról, amelyeken nem haladt át az üzenet.

Az algoritmus minden lépésben u csúcsból továbbítja az üzenetet azon v csúcsba, amelyre a $d(u, v)$ távolság a legkisebb. Ezt a lépést iterálja az algoritmus, amíg $u \neq t$.

Amennyiben minden lépésben ismernénk az egész hálózat felépítését, az összes távoli kapcsolattal együtt, akkor a feladat legegyszerűbb megoldása egy szélességi keresés lenne.

Az algoritmus várható lépésszáma a kézbesítésnél felhasznált kapcsolatok számát jelenti, ami fordított hatvány eloszlást követ a modell definiálása alapján.

Az algoritmusnak elég a felsorolt három információból csupán az első kettő, hiszen nem használja fel az üzenetnek az eddigi útját ahhoz, hogy továbbítsa a következő személynek. Tehát a csúcsoknak elegendő a saját szomszédságukat ismerni illetve bármely, t pont esetén meg kell tudniuk határozni a szomszédjaik távolságát t -tól. Ez alapján az algoritmus minden lépésben ki tudja választani azt a csúcsot, amely az üzenetet közelebb juttatja t -hez, mint a jelenlegi helyzete. Ez egy párhuzamos, vagy másnéven elosztott algoritmus. Ezen olyan algoritmust értünk, amely az adott feladatot részekre bontva több processzoron futtatva oldja meg. Ebben az esetben a processzorok az emberek, azaz a csomópontok, akik egymástól függetlenül ugyanazt a részfeladatot oldják meg, miszerint keressék meg azt a v szomszédjukat, amely legközelebb helyezkedik el t -hez viszonyítva.

Kleinberg eredményei megmutatják, hogy a hálózat felépítésétől függ az algoritmus hatékonysága. Ezzel kapcsolatban három tételt fogalmaz meg:

3.0.7. Tétel. *Adott egy $\alpha(p, q)$ konstans, és a hálózatban $r = 0$, akkor a párhuzamos algoritmus futási ideje legalább $\alpha n^{2/3}$.*

3.0.8. Tétel. *Legyen \mathcal{A} egy elosztott algoritmus, és α egy konstans, $r = 2$ és $p = q = 1$. Ekkor az elosztott algoritmus legfeljebb $\alpha(\log n)^2$ lépésben eljut s -ből t -be.*

3.0.9. Tétel. *Adott $\alpha(p, q, r)$ konstans, és*

- *ha $2 > r \geq 0$, ekkor a párhuzamos algoritmus futási ideje legalább $\alpha n^{(2-r)/3}$.*
- *ha $2 < r$, ekkor a párhuzamos algoritmus futási ideje legalább $\alpha n^{(r-2)/(r-1)}$.*

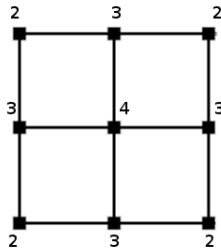
Az első állítás szerint, ha $r = 0$, azaz a hosszú távú kapcsolatok eloszlása egyenletes, akkor nem létezik jó párhuzamos algoritmus, amely megtalálná az utat s és t között. A kutatása során Kleinberg rájött, hogy az algoritmus az r paraméter növekedésével a távoli kapcsolatokat jobban ki tudja használni a hálózatban, ugyanakkor r további növelésével ezek a kapcsolatok egyre haszontalanabbá válnak az üzenet távoli mozgatásában. Ebből következik, hogy van egy érték, amikor az r értékét a legjobban ki tudja használni az algoritmus, ez az érték 2.

Az első két állítás tulajdonképpen a felállított modell tulajdonságait tükrözi, a harmadikkal együtt pedig kimondja, hogy csak $r = 2$ esetén működhet az algoritmus, tehát 2-dimenzióban. A továbbiakban ennek a három tételnek a bizonyítását részletezem.

3.1. Az $r = 2$ esetre vonatkozó tétel bizonyítása

3.1.1. Tétel. *Legyen \mathcal{A} egy párhuzamos algoritmus, és α egy konstans, $r = 2$ és $p = q = 1$. Ekkor a párhuzamos algoritmus legfeljebb $\alpha(\log n)^2$ lépésben eljut s -ből t -be.*

Az algoritmus minden lépésben úgy választja ki, hogy kinek továbbítsa az üzenetet, hogy az üzenetet birtokló u azt a v ismerőst választja, amelyre a $d(v, t)$ távolság a legkisebb. A tétel szerint $p = q = 1$ és $r = 2$, ezért minden u csúcsra igaz, hogy a rácspan 4 szomszédjával van rövid távú kapcsolata (amennyiben a rácsháló szélén van ez a szám 2, vagy 3), és egy távoli kapcsolata van, v -vel.

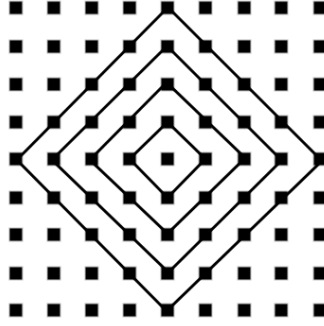


3.2. ábra. Szomszédok száma a rácshálóban

Annak a valószínűsége, hogy u a távoli kapcsolatát használva v -nek továbbítja az üzenetet a modell definiálása alapján:

$$\frac{d(u, v)^{-2}}{\sum_{(v \neq u)} d(u, v)^{-2}}$$

Adjunk felső becslést $\sum_{(v \neq u)} d(u, v)^{-2}$ értékére. Tekintsünk úgy az adott u pont körüli rácshálóra, hogy az azonos távolságra lévő pontok egy-egy héjat képeznek az alábbi ábrán látható módon. Egy adott héjon lévő pontok száma mindig $4j$, ahol j a héj sorszámát jelöli, u -tól számozva.



3.3. ábra. Az első négy héj egy csúcs körül

Ahhoz, hogy egy $n \times n$ -es rácshálót, tetszőleges u választása esetén teljesen lefedjünk ilyen héjakkal pontosan $2n - 2$ db héjra van szükségünk. Így kapjuk az alábbi felső becslést $\sum_{(v \neq u)} d(u, v)^{-2}$ értékére, ahol az összegzésben a pontokat a távolságuk szerint zárójelztük.

$$\sum_{(v \neq u)} d(u, v)^{-2} \leq \sum_{j=1}^{2n-2} (4j)(j^{-2})$$

A jobb oldalt szereplő értéket kifejtve a következőt kapjuk:

$$\sum_{j=1}^{2n-2} (4j)(j^{-2}) = 4 \sum_{j=1}^{2n-2} j^{-1} = 4 \left(\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{2n-2} \right)$$

Legyen $\sum_{k=1}^{2n-2} \frac{1}{k} = H_{2n-2}$, ahol H_{2n-2} a $2n - 2$. harmonikus számot jelöli.

3.1.2. Definíció. Ha m egy pozitív, egész szám, akkor az első m szám reciprokának az összegét harmonikus számnak nevezzük, és H_m -el jelöljük.

Kihasználva a harmonikus számnak azt a tulajdonságát, hogy $\lim_{m \rightarrow \infty} n \left(\frac{H_m}{\ln(m)} \right) = 1$ adódik a következő felső becslés:

$$4 \left(\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{2n-2} \right) \leq 4 + 4 \ln(2n-2) \leq 4 \ln(6n).$$

Tehát annak a valószínűsége, hogy u egy távoli kapcsolatot kihasználva v -nek továbbítja az üzenetet legalább

$$\frac{d(u, v)^{-2}}{d(u, v)^{-2} 4 \ln(6n)}.$$

Azt mondjuk hogy az \mathcal{A} algoritmus j -edik fázisában vagyunk, ha az aktuális csúcsra, u -ra, akinél az üzenet van igaz, hogy $2^j < d(u, t) < 2^{j+1}$. Azt mondjuk, hogy az algoritmus a 0 fázisban van, ha az üzenet távolsága t -től legfeljebb 2. Így j értékére tudunk becslést adni:

$$0 \leq j < \log(n).$$

A lépések számával a céltól való távolság értelemszerűen mindig csökken, hiszen minden lépésben u továbbítja az üzenetet azon v ismerősének amelyre, $d(v, t)$ távolság a legkisebb. Tegyük fel, hogy a j -edik fázisban vagyunk, amelyre igaz, hogy $\log(\log(n)) \leq j < \log(n)$. Ahhoz, hogy valamely j fázis az utolsó legyen, azaz az üzenet elérjen t -be szükséges, hogy az algoritmus futása során t -től adott távolságon belülre érkezzen az üzenet. Legyen \mathcal{B}_j egy 2^j sugarú gömb, t középponttal. Ennek elemszámát alulról tudjuk becsülni :

$$|\mathcal{B}_j| \geq 1 + \sum_{i=1}^{2^j} i.$$

Alkalmazva a számtani sor összegére ismert összefüggést a következőt kapjuk:

$$|\mathcal{B}_j| \geq \frac{(1 + 2^j)2^j}{2} = \frac{1}{2}2^{2j} + \frac{1}{2}2^j > 2^{2j-1}.$$

Tehát \mathcal{B}_j elemszáma legalább 2^{2j-1} .

Minden u pontra igaz, hogy

$$d(u, x)_{x \in \mathcal{B}_j} \leq d(u, t) + d(t, x) = 2^{j+1} + 2^j < 2^{j+2},$$

ahol a felső becslés a háromszögegyenlőtlenség miatt igaz. Ez alapján annak a valószínűsége, hogy u -ból a következő lépéssel távoli kapcsolaton keresztül belép \mathcal{B}_j -be az üzenet legalább

$$\frac{2^{2j-1}}{4 \ln(6n)2^{2j+4}} = \frac{2^{2j}}{2^2 \ln(6n)2^{2j}2^4} = \frac{1}{2^7 \ln(6n)}.$$

Legyen \mathcal{C}_j a j -edik fázisig megtett lépések száma, ahol \mathcal{C}_j értelemszerűen egy valószínűségi változó. Ha az algoritmus teljes lépésszáma X , akkor igaz a következő:

$$X \leq 1 + \mathcal{C}_1 + \mathcal{C}_2 + \dots + \mathcal{C}_{\log(n)}.$$

Ezért az algoritmus várható lépésszámának becsléséhez először \mathcal{C}_j várható értékét számítjuk ki.

$$E\mathcal{C}_j = \sum_{i=1}^{\infty} P(\mathcal{C}_j \geq i)$$

Mivel az algoritmus minden lépésben közelebb juttatja az üzenetet t -hez, ezért feltehetjük, hogy u -ból v -be mutató távoli kapcsolatot éppen akkor használja az algoritmus, amikor u -ból lép az algoritmus. Annak az esélye, hogy legalább i lépést tettünk meg a j -edik fázisig:

$$\left(1 - \frac{1}{2^7 \ln(6n)}\right)^{i-1}.$$

Tehát \mathcal{C}_j várható értékét a következő összeg felülről becsüli:

$$\sum_{i=1}^{\infty} \left(1 - \frac{1}{2^7 \ln(6n)}\right)^{i-1} = \frac{2^7 \ln(6n) - 1}{1 - \frac{1}{2^7 \ln(6n)}} = 2^7 \ln(6n).$$

Továbbá mivel a fázisok számának felső korlátja $\log(n)$, ezért

$$X \leq \sum_{j=0}^{\log(n)} \mathcal{C}_j,$$

ezért igaz az, hogy

$$EX \leq (1 + \log(n))(2^7 \ln(6n)).$$

Ezzel az állítást be is bizonyítottuk, hiszen

$$(1 + \log(n))(2^7 \ln(6n)) \leq \alpha(\log(n))^2$$

egy megfelelő választása α -nak.

3.2. Az $r < 2$ esetre vonatkozó tétel bizonyítása

3.2.1. Tétel. *Adott $\alpha(p, q, r)$ konstant, és ha $2 > r \geq 0$, ekkor a párhuzamos algoritmus futási ideje legalább $\alpha n^{(2-r)/3}$.*

Az alapfeladatunk ismételten az, hogy egy $n \times n$ -es rácshálóban, két tetszőlegesen választott csúcs, s és t között egy algoritmussal megtaláljuk a legrövidebb utat. A tétel azt mondja ki, hogy egy \mathcal{A} decentralizált algoritmus futási ideje függ a hálózatban lévő kapcsolatok kialakításától, és a hálózat dimenziójától. Annak a valószínűsége, hogy u a távoli kapcsolatát használva v -nek továbbítja az üzenetet a modellben definiáltak alapján:

$$\frac{d(u, v)^{-r}}{\sum_{(v \neq u)} d(u, v)^{-r}}.$$

Az előző bizonyításhoz hasonlóan most is becsüljük $\sum_{(v \neq u)} d(u, v)^{-r}$ értéket, azonban most alsó becslést adunk rá. Mivel $p \geq 1$, ezért a rövid távú kapcsolatoknak van alsó korlátja.

$$\begin{aligned} \sum_{(v \neq u)} d(u, v)^{-r} &\geq \sum_{j=1}^{n/2} (j)(j^{-r}) = \\ &= \sum_{j=1}^{n/2} j^{1-r} \geq \int_1^{n/2} x^{1-r} dx = \\ &= \left[\frac{x^{2-r}}{2-r} \right]_1^{n/2} = \frac{(\frac{n}{2})^{2-r}}{2-r} - \frac{1}{2-r} = \frac{(\frac{n}{2})^{2-r} - 1}{2-r} = \frac{n^{2-r} - 2^{2-r}}{2^{4-2r}} \end{aligned}$$

Mivel $0 \leq r < 2$, ezért igaz, hogy $2^{3-r} \leq n^{2-r}$, ezért

$$\frac{n^{2-r} - 2^{2-r}}{2^{4-2r}} \geq \frac{n^{2-r}}{(2-r)2^{3-r}}.$$

Legyen \mathcal{U} a t középpontú, pn^δ sugarú gömb, ahol $\delta = \frac{2-r}{3}$. Az előző bizonyításhoz hasonlóan, becsüljük \mathcal{U} elemszámát:

$$|\mathcal{U}| \leq 1 + \sum_{j=1}^{pn^\delta} 4j \leq 4pn^2 n^{2\delta},$$

ahol az utolsó egyenlőtlenségél kihasználtuk, hogy $2 \leq pn^\delta$, mivel $1 \leq p$ és $2 \leq n$ a rácsháló modellje alapján, illetve $2 > r \geq 0$ a tétel feltétele szerint.

Jelölje \mathcal{E}_i azt az eseményt, hogy távoli kapcsolatot használva \mathcal{U} -beli pontba érünk az i . lépésben. Legyen $\mathcal{E} = \bigcup_{i \leq \lambda n^\delta} \mathcal{E}_i$, ahol Kleinberg λ értékét $(2^{8-r} qp^2)^{-1}$ -ben határozta meg. Ez eddigiek alapján \mathcal{E}_i valószínűségét tudjuk becsülni felülről:

$$P(\mathcal{E}_i) \leq \frac{q|\mathcal{U}|}{\frac{n^{2-r}}{(2-r)2^{3-r}}} \leq \frac{(2-r)2^{5-r} p^2 q n^{2\delta}}{n^{2-r}}.$$

Mivel $\mathcal{E} = \bigcup_{i \leq \lambda n^\delta} \mathcal{E}_i$, ezért

$$P(\mathcal{E}) = \sum_{i \leq \lambda n^\delta} P(\mathcal{E}_i).$$

Továbbá \mathcal{E}_i definiálása alapján igaz, hogy

$$\sum_{i \leq \lambda n^\delta} P(\mathcal{E}_i) \leq \frac{(2-r)2^{5-r} p^2 q n^{3\delta} \lambda}{n^{2-r}} = (2-r)2^{5-r} qp^2 \lambda.$$

Ahol az egyenlőség $\delta = \frac{2-r}{3}$ adott értéke miatt igaz. Ebbe a kifejezésbe helyettesítsük be az adott λ értéket:

$$\frac{(2-r)2^{5-r}qp^2}{2^{8-r}qp^2} = \frac{(2-r)2^{5-r}}{(2^{8-r})} = (2-r)2^{-3} = 2^{-2} - r2^{-3}.$$

Tehát

$$P(\mathcal{E}) \leq 2^{-2}.$$

Jelölje \mathcal{F} azt az eseményt, hogy s és t között a távolság legalább $n/4$. Ennek a bekövetkezésekor tehát s -től az $n/4 - 1$ lépésben elérhető csúcsokat nem választhatjuk t -nek. Az $n/4 - 1$ lépésben elérhető csúcsok száma pedig

$$\sum_{j=1}^{j=\frac{n}{4}-1} (4j)(j^{-r}) = 4 \sum_{j=1}^{j=\frac{n}{4}-1} (j^{1-r}),$$

ahol a tétel feltétele szerint $0 < r < 2$. Tehát \mathcal{F} valószínűsége:

$$\frac{n^2 - 4 \sum_{j=1}^{j=\frac{n}{4}-1} (j^{1-r})}{n^2} = 1 - \frac{4 \sum_{j=1}^{j=\frac{n}{4}-1} (j^{1-r})}{n^2}.$$

Továbbá r értékétől függetlenül igaz, hogy

$$\lim_{n \rightarrow \infty} \left(\frac{4 \sum_{j=1}^{j=\frac{n}{4}-1} (j^{1-r})}{n^2} \right) = 0,$$

tehát

$$P(\mathcal{F}) = 1 - \frac{4 \sum_{j=1}^{j=\frac{n}{4}-1} (j^{1-r})}{n^2} \geq \frac{1}{2}.$$

Az eddigiekből következik, hogy $P(\bar{\mathcal{F}} \vee \mathcal{E}) \leq \frac{1}{2} + \frac{1}{4}$, ami átírva azt jelenti, hogy

$$P(\mathcal{F} \wedge \bar{\mathcal{E}}) \geq \frac{1}{4}.$$

Legyen \mathcal{G} az az esemény, hogy t éppen λn^δ lépésben éri el az üzenet. Ha \mathcal{F} bekövetkezik, de \mathcal{E} nem következik be, akkor értelemszerűen \mathcal{G} sem következhet be.

$$P[\mathcal{G} | \mathcal{F} \wedge \bar{\mathcal{E}}] = 0$$

Jelölje X valószínűségi változó az előző bizonyításhoz hasonlóan most is az algoritmus teljes lépésszámát, most is ennek a várható értékét szeretnénk becsülni. Mivel \mathcal{E} és \mathcal{F} esemény bekövetkezése azt jelenti, hogy legfeljebb λn^δ lépésben az üzenet elért a t csúcsba s -ből, amely legalább $n/4$ távolságra van, ezért X feltételes várható értékére a következő becslést tudjuk adni $\bar{\mathcal{E}}$ felhasználásával:

$$E[X|\mathcal{F} \wedge \bar{\mathcal{E}}] \geq \lambda n^\delta.$$

A teljes várható érték tétele szerint, ha Y_1, Y_2, \dots, Y_n teljes eseményrendszer alkotnak és $P(Y_i) > 0, \forall i$ -re, akkor $E(A) = \sum_{i=1}^n E(A|Y_i)P(Y_i)$, azonban $\mathcal{F} \wedge \bar{\mathcal{E}}$ nem alkotnak teljes eseményrendszer. Legyen \mathcal{Z} a következő négy esemény uniója.

$$\mathcal{F} \wedge \mathcal{E}; \mathcal{F} \wedge \bar{\mathcal{E}}; \bar{\mathcal{F}} \wedge \bar{\mathcal{E}}; \bar{\mathcal{F}} \wedge \mathcal{E},$$

így \mathcal{Z} teljes eseményrendszer alkot. Most már tudjuk alkalmazni a teljes várható érték tételét:

$$E(X) = \sum_{i=1}^4 (E(X|\mathcal{Z}_i)P(\mathcal{Z}_i)).$$

Mivel az összeg mindegyik tagja nem negatív, ezért biztosan igaz, hogy

$$E(X) \geq E[X|\mathcal{F} \wedge \bar{\mathcal{E}}]P(\mathcal{F} \wedge \bar{\mathcal{E}}).$$

A szorzat mindkét tagjára adtunk alsó becslést korábban, ezeket alkalmazva a következőt kapjuk:

$$E(X) \geq \frac{1}{4}\lambda n^\delta.$$

Az adott λ és δ értékeket behelyettesítve megkapjuk a tételben kimondott várható lépésszám alsó becslését, $\alpha = \frac{1}{4}(2^{8-r}qp^2)^{-1}$ választással.

$$E(X) \geq \frac{1}{4}(2^{8-r}qp^2)^{-1}n^{\frac{2-r}{3}} = \alpha n^{\frac{2-r}{3}}$$

Ennek a tételnek a bizonyítása tartalmazza egyben Kleiberg első tételének is a bizonyítását.

3.2.2. Tétel. *Adott egy $\alpha(p, q)$ konstant, és a hálózatban $r = 0$, akkor az elosztott algoritmus futási ideje legalább $\alpha n^{2/3}$.*

Egyértelmű, hogy $r = 0$ esetben $n^{\frac{2-r}{3}} = n^{2/3}$.

3.3. Az $r > 2$ esetre vonatkozó tétel bizonyítása

3.3.1. Tétel. *Adott egy $\alpha(p, q, r)$ konstans és ha $2 < r$, ekkor az elosztott algoritmus futási ideje legalább $\alpha n^{(r-2)/(r-1)}$.*

Az eddigiekben $r \leq 2$ esetet már bebizonyítottuk, most nézzük, hogy alakul az algoritmus futási ideje $r > 2$ esetén. Nézzük ismét azt a valószínűséget, hogy u a távoli kapcsolatát használva v -nek továbbítja az üzenetet, amelynek az értéke a modell definiálása alapján:

$$\frac{d(u, v)^{-r}}{\sum_{(v \neq u)} d(u, v)^{-r}},$$

ahol a nevezőről $r > 2$ miatt tudjuk, hogy legalább 1. Vizsgáljuk meg a $P(d(u, v) > m)$ valószínűséget, tetszőleges m érték esetén:

$$P(d(u, v) > m) \leq \sum_{j=m+1}^{2n-2} (4j)(j^{-r}),$$

ahol az egyenlőtlenség jobb oldalán álló összeget a második tétel bizonyításában leírtakhoz hasonlóan kapjuk. Az összegzést más alakban felírva egyértelmű a következő felső becslés:

$$\sum_{j=m+1}^{2n-2} (4j)(j^{-r}) = 4 \sum_{j=m+1}^{2n-2} (j^{1-r}) \leq \int_m^{\infty} j^{1-r} dx$$

Ezt a végtelen intervallumon való integrálás definíciója alapján

$$= \lim_{z \rightarrow \infty} \left(\frac{z^{2-r}}{2-r} - \frac{m^{2-r}}{2-r} \right),$$

ahol

$$\left(\frac{z^{2-r}}{2-r} \right) \rightarrow 0,$$

mivel $r > 2$. Tehát

$$\int_m^{\infty} j^{1-r} dx \leq \frac{m^{2-r}}{r-2}.$$

Legyen $\varepsilon = r - 2$, ezzel a helyettesítéssel az előbbi érték $\varepsilon^{-1} m^{-\varepsilon}$ alakba írható. Tetszőlegesen nagy n értékre igaz, hogy $n^{\frac{1}{1-r}} \geq p$, hiszen $r > 2$ a tétel feltétele alapján. Definiáljuk \mathcal{E}_i eseményt, ami azt jelenti, hogy az üzenet az i -edik lépésben $u \neq t$ csúcsból v -be lép hosszú kapcsolaton keresztül, amely kielégíti a $d(u, v) > n^{\frac{1}{1-r}}$ feltételt.

Vezessük be a következő jelöléseket, hogy a bizonyítás leírása átláthatóbb legyen:

$$\begin{aligned}\beta &= \frac{\varepsilon}{1 + \varepsilon}, \\ \gamma &= \frac{1}{1 + \varepsilon}, \\ \lambda &= \frac{\min(\varepsilon, 1)}{8q}.\end{aligned}$$

Legyen $\mathcal{E} = \bigcup_{i \leq \lambda n^\beta} \mathcal{E}_i$. Az előzőekhez hasonlóan most is igaz, hogy

$$P(\mathcal{E}) \geq \sum_{i \leq \lambda n^\beta} (\mathcal{E}_i).$$

A fentiekben levezetett $P(d(u, v) > m)$ becslést felhasználva, $m = n^\gamma$ helyettesítéssel \mathcal{E}_i bekövetkezésének a valószínűsége is felülről becsülhető, így adódik a következő:

$$P(\mathcal{E}_i) \leq q\varepsilon^{-1}(n^\gamma)^{-\varepsilon}.$$

Ez alapján pedig $\sum_{i \leq \lambda n^\beta} \mathcal{E}_i$ is felülről becsülhető:

$$\sum_{i \leq \lambda n^\beta} (\mathcal{E}_i) \leq \lambda n^\beta q \varepsilon^{-1} n^{-\varepsilon \gamma}$$

Behelyettesítve a $\beta, \lambda, \varepsilon$ és γ értékeket ezt az értéket tovább egyszerűsödik a becslésünk:

$$\begin{aligned}\frac{\min(r-2, 1)}{8q} n^{-\frac{r-2}{1-r}} q (r-2)^{-1} n^{\frac{r-2}{1-r}} \\ = \frac{\min(r-2, 1)}{8q} q \frac{1}{r-2}.\end{aligned}$$

A tétel feltétele szeint $r > 2$, ami alapján $\varepsilon = r - 2 > 0$ értelemszerűen adódik. A további számoláshoz azonban szükséges, $\min((r-2), 1)$ értékét meghatároznunk, ehhez bontsuk két lehetőségre ezt az értéket:

1. eset $\varepsilon = r - 2 < 1$, azaz $\min((r-2), 1) = r - 2$

$$\frac{r-2}{8q} q \frac{1}{r-2} = \frac{1}{8}$$

2. eset $\varepsilon = r - 2 > 1$, azaz $\min((r-2), 1) = 1$

$$\frac{1}{8q} q \frac{1}{r-2} < \frac{1}{8}.$$

Tehát $P(\mathcal{E}) \leq 1/8$.

Jelölje \mathcal{F} esemény, az állítás első részének bizonyításához hasonlóan most is azt az esetet, hogy s és t között a távolság legalább $n/4$. Már megmutattuk, hogy ennek a valószínűsége legalább $1/2$. Továbbá már azt is beláttuk, hogy $P(\mathcal{F} \wedge \bar{\mathcal{E}}) \geq \frac{1}{4}$, ami a mostani \mathcal{E} definiálása alapján is igaz, hiszen az imént beláttuk, hogy $P(\mathcal{E}) \leq \frac{1}{8}$, tehát $P(\mathcal{E}) \leq \frac{1}{4}$ is jó felső becslés.

Legyen X egy valószínűségi változó, amely a már megszokott módon az algoritmus lépésszámát jelöli, hogy s -ből t -be hány lépés alatt jut el az üzenet. Jelölje \mathcal{G} pedig azt, hogy az üzenet pontosan λn^β lépésben éri el a t célszemélyt, ahol λ és β értéke a fentiekben meghatározott.

Tegyük fel, hogy \mathcal{F} teljesül, de \mathcal{E} nem teljesül, azaz s és u között a távolság minimum $n/4$, de az első λn^β lépésben legfeljebb n^γ távolságot tud megtenni egy lépésben. Nyilvánvaló, hogy ekkor \mathcal{G} sem teljesülhet, hiszen a lépésszám így legfeljebb $\lambda n^{\beta+\gamma} = \lambda n$. Erről a szorzatról viszont biztosan tudjuk, hogy értéke szigorúan kisebb, mint $n/4$. Tehát abból, hogy \mathcal{E} nem következik be egyértelműen következik az, hogy \mathcal{F} teljesül, azaz $P(\mathcal{F} \wedge \bar{\mathcal{E}}) = P(\bar{\mathcal{E}})$. Mivel az algoritmus teljes lépésszáma jelenti X várható értékét, ezért

$$P(X|\mathcal{F} \wedge \bar{\mathcal{E}}) = P(X|\bar{\mathcal{E}}) \geq \lambda n^\beta.$$

Ismételten használjuk ki a teljes várható érték tételét X várható értékének alsó becsléséhez:

$$EX \geq E(X|\mathcal{F} \wedge \bar{\mathcal{E}})P(\mathcal{F} \wedge \bar{\mathcal{E}}).$$

Az eddigi eredményeinket felhasználva pedig ez azt jelenti, hogy $EX \geq 1/4\lambda n^\beta$, ahová behelyettesítve az adott értékeket a következőt kapjuk:

$$EX \geq \frac{1}{4} \frac{\min((r-2), 1)}{8q} n^{\frac{n-2}{r-1}},$$

ahol $\alpha(p, q, r) = \frac{1}{4} \frac{\min((r-2), 1)}{8q}$ választással a tétel állítását kapjuk. Ezzel az állítást bebizonyítottuk.

4. fejezet

Összegzés

A szociális hálózatok a társadalom felépítésének egy, a matematika nyelvére lefordított és jelentősen idealizált ábrázolása a szakdolgozatomban bemutatott gráfmodellek mindegyike. Kleinberg az által javasolt modellben megmutatta, hogy hogyan és milyen paraméterek mellett valósul meg a Milgram-kísérlet. Bár nyilvánvalóan a valós kapcsolati hálózat egyáltalán nem a Kleinberg-modell szigorúan definiált szabályai mentén épül fel, a jövőbeni kutatásoknak jó alapot ad eredménye.

A Milgram-kísérletet azóta többen, többféleképpen megismételték. Például idén novemberben jelent meg egy cikk a PLOS ONE portálon, melyben a hat lépés elméletet egy twitteres adatbázison modellezték le egyetemünk dolgozói és hallgatói [4].

Irodalomjegyzék

- [1] Rónyai Lajos: *Véletlen és algoritmusok*, Typotex, 2011.
- [2] Jon Kleinberg: *The Small-World Phenomenon: An Algorithmic Perspective*, Proc. 32nd ACM Symposium on Theory of Computing, 2000, 163-170
- [3] Jon Kleinberg: Navigation in a Small-World, Nature **406**(2000), 846.
- [4] Szüle J, Kondor D, Dobos L, Csabai I, Vattay G: *Lost in the City: Revisiting Milgram's Experiment in the Age of Social Networks*, PLoS ONE **9**(11), (2014), e111973.
- [5] Duncan J. Watts, Steven H. Strogatz: *Collective dynamics of 'small-world' networks*, Nature **393** (1998), 440.
- [6] Erdős Pál, Rényi Albert: *On random graphs I.*, Publicationes Mathematicae **6** 1959, 290-297
- [7] Barabási Albert László, Albert Réka: *Emergence of scaling in random networks*, Science **286**(5439), (1999), 509-512